Simulating heavy rain damage in an insurance context

S. BUSCH

Institute of Water Resources Management, Leibniz University of Hanover, Appelstr. 91, D-30167 Hannover, Germany and E+S Rückversicherung AG, Karl-Wiechert-Allee 50, D-30625 Hannover, Germany

stefanie.busch@es-rueck.de

Abstract Rising cost of damages due to heavy precipitation brought up the need to consider this type of hazard in models used by the insurance industry to calculate monetary risk. Up to now risk maps quantify fluvial risk or the probability of being exposed to river flooding. However, the pluvial risk or potential damage due to heavy precipitation events has not been assessed in terms of its contribution to insurance losses. The assumption that damages contributed by heavy rain will occur additionally to river flooding will be examined. Attempts to create zonal maps as they exist in flood zones for fluvial risk did not succeed. Apart from topography that is clearly visible when looking at rain maps, no further distinguished areas were found. Conclusively, every building in Germany is potentially threatened by heavy rain events and as such, this paper concerns itself with the outline of a probabilistic model for assessing the component of pluvial flooding.

Keywords heavy precipitation; simulation; insurance; decision support

INTRODUCTION

In the insurance industry flood models are used to assess the risk of river flooding. These models are mostly constructed to best estimate the expansion of water on the flood plains. The off flood plain component is mostly only implicit or very crudely implemented. Recently, some primary insurance companies refrained from demanding a special coverage for elementary risks, because it is included the inclusion of heavy precipitation events in their residential building policies. Hence, the necessity of being able to assess losses due to heavy rain events is becoming more and more evident. This article divides inundation into fluvial and pluvial flooding; the first due to river flooding, and the latter due to local heavy precipitation events. Focus lies on the latter. The German KOSTRA Atlas (DWD, 1997) allows for quantifying the rain depth of a rain event for selected recurrence intervals and several duration intervals for each point in Germany. This statistical background can be used to resolve the equations in a way that a certain rain depth can be attributed a return period. However, these conclusions are only valid for one particular point in time. To make statements about the occurrence of a complete event and the spatial variability of the occurrences, an event set has to be constructed to distribute the incidences spatially and to consider all possible combinations of parameters. By means of simulation, events which have not yet been observed can be represented in the event set.

The goal is it to extract the necessary distributions of parameters from the analysed data to allow for simulating a synthetic event set of 10.000 event years. Research is still ongoing. In this study the fields of investigation are outlined and the concept of building the model is described. Definite results are expected in the near future.

DATA

The Emschergenossenschaft/Lippeverband (EGLV) provided a narrow mesh of 92 rain gauge recording series spanning up to 86 years in a 5-minute resolution. This data was used to extract damage prone events and find appropriate probability distributions for parameters such as duration, intensity and amount of rain.

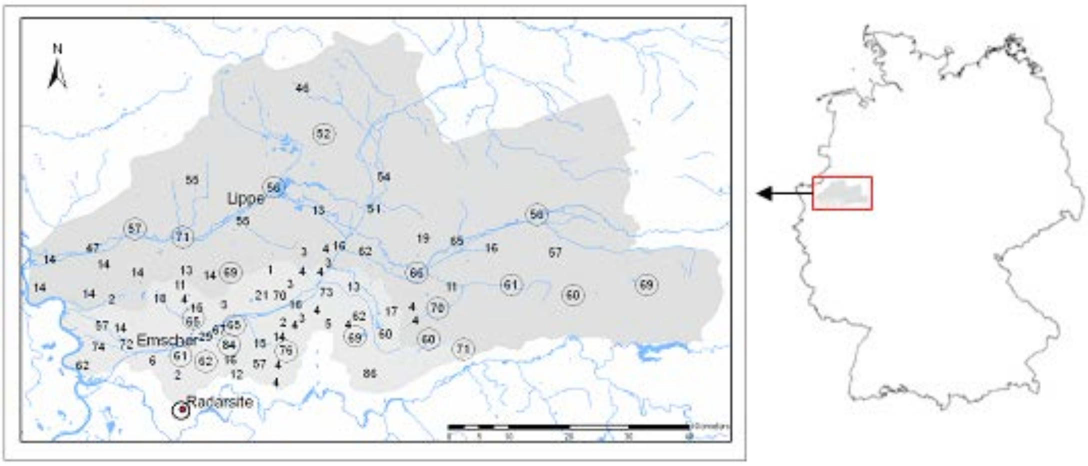


Fig. 1 The Emscher/Lippe study area. Numbers show the record lengths of stations.

In addition to the rain gauges, three different set of radar imagery are used. The first set was provided for 15 events in the study area which occurred between 1999 and 2005. It has a spatial resolution of 1 km^2 and a temporal resolution of 5 minutes. This high resolution set was made available for the Essen radar site. The second set comprises 11 years of radar data in a 4 km^2 resolution and 15 minute time steps. This set is a composite of the 16 operating radar sites in Germany and covers the whole country. The last set has already been processed by the German Meteorological Service (DWD), containing the pixel coordinates of the heavy rain pixels detected by the radar according to a definition of the DWD. This set was available for the years 2004 through 2007. Spatial and temporal resolutions are 1 km^2 and 5 minutes, respectively.

The above mentioned meteorological data were linked with the disaster relief operations associated with flooding by the emergency response units of the responsible fire departments. Furthermore, corresponding insurance claims data were examined to develop an algorithm that ties specific meteorological and surface conditions to damage prone situations. While fire department runs are only useful to get a qualitative notion of induced losses, the insurance claims help to shed some light on the monetary aspect.

The SRTM digital elevation with a 90 m resolution model was used to calculate height, slope or aspect of the affected areas. Census data served to determine the percentage of the affected buildings.

METHODS AND RESULTS

Rain gauge data

The rain gauge data provided by the EGLV exist in a five minute time step. Secluded events were extracted using the KOSTRA basis. An event is defined as continuous rainfall when it meets both criteria of at least one millimetre in five minutes and at least five millimetres per hour. The minimum value for the complete event has to be 30 millimetres. Duration intervals from 5 minutes to 24 hours were considered. The given return period, which had to be reached in any duration interval, was set to 0.4. This includes, on average, 5 events in 2 years per station.

For each of the resulting 22,349 record series, statistical distribution functions were fit to the variables *amount of rain* (mm), *duration* (min), *maximum return period* (year), *duration interval* (min) and *intensity* (l $^{-(\mathrm{ha}\ \mathrm{s})}$). The events spread over 3,817 separate days. Typical seasonal and diurnal distributions of the events were examined. Additionally, the interactions between these variables were investigated. Since most relationships are not linear but nevertheless show certain patterns, the Bernstein copula was used to show the interdependencies between the selected parameters.

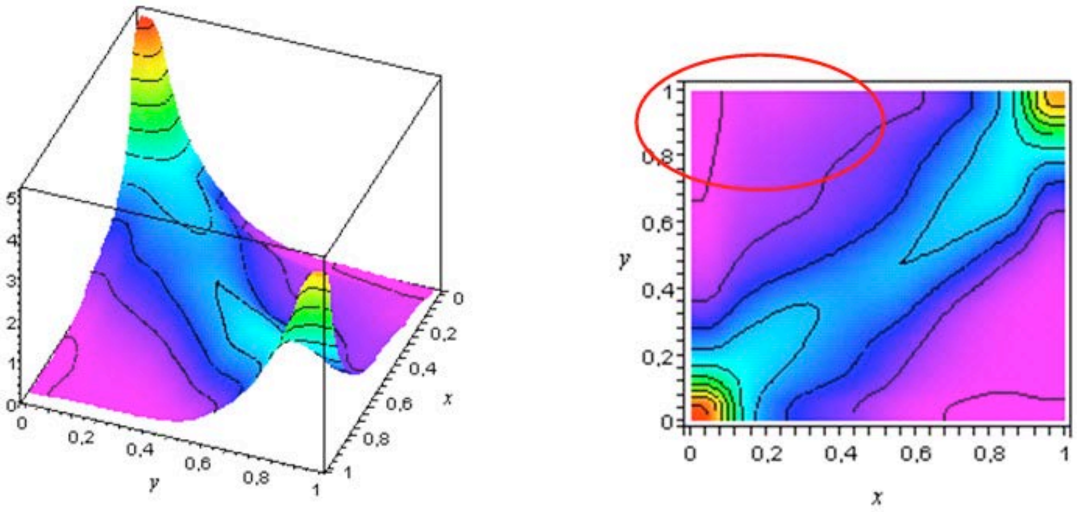


Fig. 2 left: density plot for duration (x) and amount of rain (y). right: copula contour plot

The density plot in Figure 3 shows that the bulk of the values are found in the diagonal (short duration – little rainfall, long duration – large rainfall). More importantly, it depicts the possibility of experiencing short rain events with a large amount of rain (high intensity, thus interesting for the simulation), as well as long rain events with a small amount of rain (less important).

Being non parametrical and non symmetrical, the Bernstein copula is very flexible (Pfeifer *et al.*, 2009). For the general concept of copulas, please refer to Nelsen (1999) or Salvadori *et al*. (2007). Studies with a hydrological focus can be found in Renard *et al*. (2007).

To get a notion of the extent of the events, a subset of 20 stations with continuous and parallel data for 47 years was selected and the concurrent recording of values representing return periods of 10 years or higher were examined.

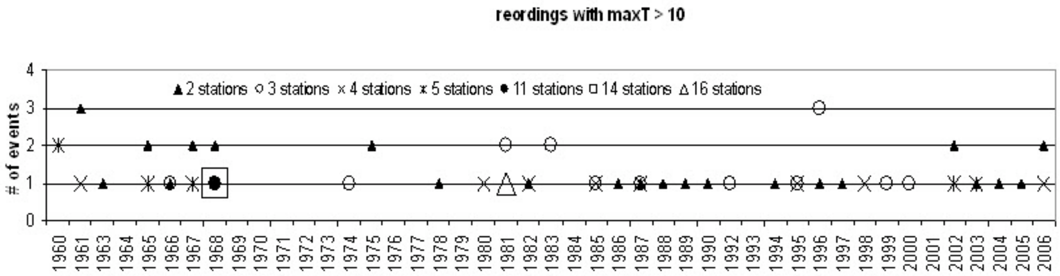


Fig. 3 Number of events per year that comprise of 2 stations and more (*maxT* > 10)

Events of a magnitude higher than 200 years in return period are mostly recorded at one station only; rarely, two stations registered these high values. Even for return periods (*maxT*) of 10 years or higher, the investigation shows that 60% of the cases are only recorded at one station. Thus, the extent of these extreme values seems to be relatively small. Figure 3 depicts the number of events per year that are recorded at two gauges or more. For the subset, the correlation of the parallel wet periods were determined according to distance and direction. With distances larger than 40 km the correlation falls under 0.3. Nevertheless, if the gauges lie in WSE ENE direction, the correlation increases slightly. For a similar study in Belgium, Vannitsem *et al.* (2007) observed a complete independence for the stations of 50 km in summer and 100 km in winter, for a temporal resolution of 1 hour.

Radar data

By means of the documentation of the emergency response units of the fire departments in the study area, originally 26 days with the highest frequency of disaster relief operations were selected. For these days the DWD provided high resolution radar data for the radar site Essen. The data were corrected for attenuation. For details please refer to Krämer (2007). Fifteen of these days showed local patterns and were analysed further.

Radar data are a good instrument to assess the spatial extent of an event; yet, when converting reflectivity to absolute values, they are still afflicted with large uncertainties. For this reason, no statements are made about absolute amount of rain, but the images are rather used to extract typical patterns that result in damages to buildings. Since the data are in a high temporal resolution, it is very likely that peaks are not diluted by averaging over a period of time. Each rain cell is approximated by an ellipse on different threshold parameters. It was experimented with various sills. A threshold of 37 dBZ seems to describe well the critical cell shape, when comparing the resulting footprint to the reported losses on the ground.(Busch *et al.*, 2009) However, when aggregating over the complete event, all of the analysed events also show a good concurrence of the accumulation of pixels larger than 46 dBZ and reported damage.

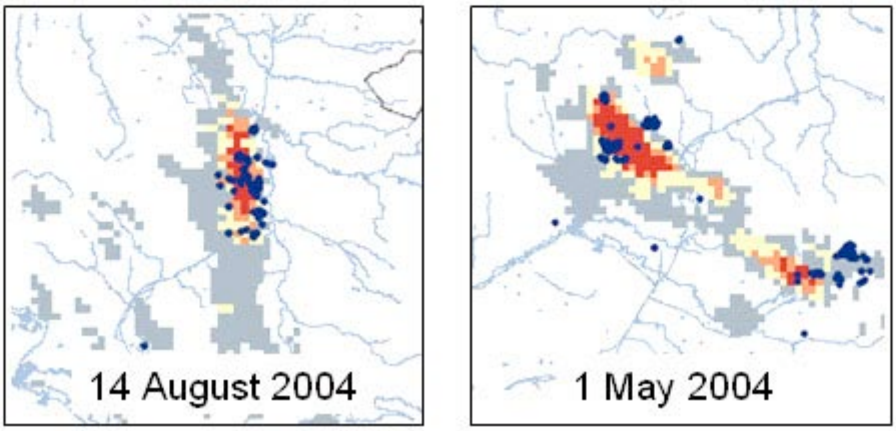


Fig. 4 footprints of damage areas for local cells in the study area.

Keeping in mind the target of simulating synthetic events, it is tried to determine one single ellipse for each detached pattern to approximate the course over time. The processed data of the DWD were treated in a way that all identified heavy rain cells, pixels that trigger a warning (representing 25 mm^{-h}), were visualised in each time step. Since these cells carry a cell-id they can be recognized in the following image. Ellipses with identical cell-ids are then merged to coherent features. Thus, a moving cell will overlap from image to image and build a new shape, which can again be approximated by an ellipse. The ratio between the single extent of each cell and the area the cell covered after its lifetime is examined as to whether it can be used as an intensity indicator. (Busch, 2010)

Furthermore, a peak over threshold (POT) analysis was conducted, counting each instance when the reflectivity was larger than 46 dBZ (cnt_{46}). Additionally each instance larger than 0 was counted to get a notion of the total duration (cnt_0). The resulting image was then overlain with the reported damage locations. The locations were buffered with a two kilometre radius to account for possible wind drift or slope, i.e. the losses may have occurred at the bottom of a hillside. Other than Schuster *et al.* (2006), the radar field was not shifted in one direction, but the search radius around the radar fields was extended.

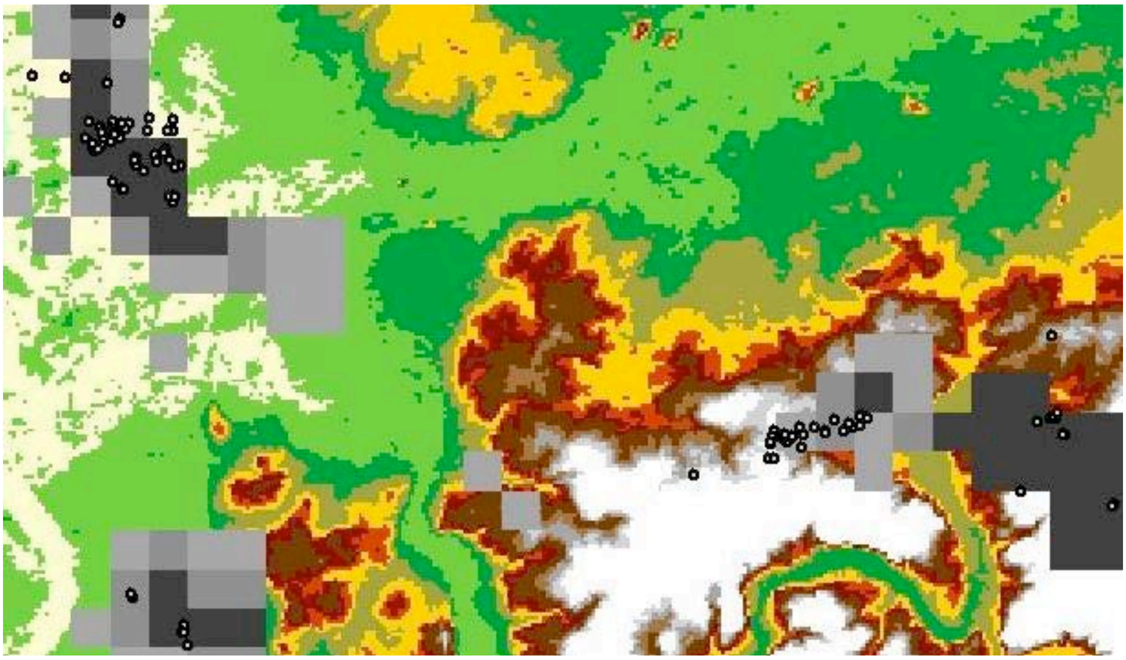


Fig. 5 Damage locations with underlying elevation model and radar cells

Figure 5 displays an example of reported losses (points), which correspond well with the largest ratio of $cnt_{46}{}^{-cnt0}$. For an interval width, one standard deviation was chosen, displayed from light grey to dark grey. As expected, the underlying digital elevation model shows well that the fit is better when topography is level (left hand cells in figure 5).

The VX-format, provided for 2004 to 2007 was examined to find the distribution of moving direction, ellipticity of the cells, statistical parameters of speed, area and duration. To illustrate the moving direction, the direction was classified into eight classes oriented at the wind rose. Figure 4 shows the total count of moving direction as the midpoints of each class, e.g. class NE stretches from NNE to ENE. The rose reflects well the typical atmospheric conditions in Germany. Winds from West of SW are the prevailing patterns. For this analysis 327,247 cells were considered. The remaining cells are starting point cells with no direction yet.

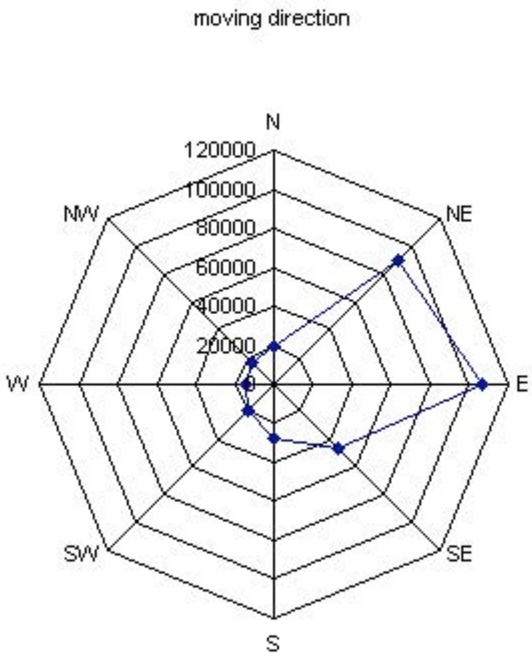
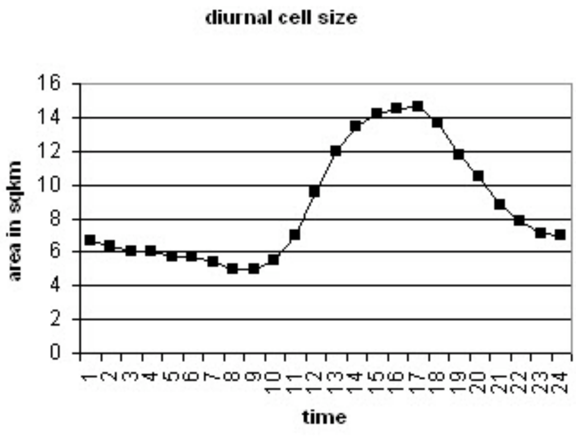


Fig. 7 Moving direction of the extracted rain cells (n = 337,247)

The radar composite data of Germany have a coarser resolution, but since they were available for 11 years (1999-2009) the images were analysed with the target of finding typical distribution patterns of rain cells. Reflectivity values are given in 6 classes, with a step size of 11, except for the first two classes with 0 and 19 dBZ. Hence, 46 dBZ are represented in class 5. For each category a POT analysis was conducted. The resulting count of values above 46 dBZ over 11 years does not allow for delineating a risk map. Except for the topography which shows through, features due to artefacts in the radar are more articulated than obvious regions with higher counts. The time series of available radar data is still very short, thus, one single exceptional year will have an enormous influence on the accumulation.

Ellipticity describes the ratio between the major and minor semi axis of an ellipse. It was found to be 0.6 with a standard deviation of 0.13 (n = 1,533,126). This value is in the same range of the values reported by von Rebora et al. (2006), who derived their values for cells of the C-ban-Radar of Mount Settepani in NW-Italy.



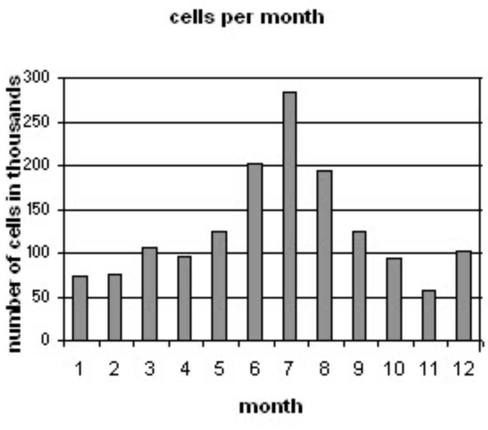


Fig. 8 left: average cell size throughout the day. right: number of cells per month

Finally, the diurnal and monthly patterns were verified. The assumption that a multitude of events happen in the summer months and are more severe in the late afternoon seems trivial (Fig 8). Nevertheless, for simulating an event set, each parameter has to be expressed as a function. During the simulation, the month July will be drawn more often than January. The occurrence throughout the day is more important when looking at movable risks such as cars, which might be exposed during daytime, but are mostly sheltered during the night.(Otto, 2009) Thus, the frequency of occurrence is not displayed, but rather the average spatial coverage nationwide.

Damage data

The emergency responses documented by the fire departments as well as the insurance claims data were geo-referenced in order to assign an xy-coordinate to each address. These locations were then attributed with their flood risk class using ZÜRS Geo, a web based GIS used by most insurance companies. It provides the four inundation zones: statistically flooded more than every 10 years, between 10 and 50 years, less frequent than every 50 years and finally, statistically flooded less than every 200 years. The intention was to verify the assumption that losses caused by local heavy rain events are hardly ever due to the flooding of a river. Thus, if it rains heavily but locally, the water body will be easily able to absorb the precipitation, and consequently, no damage will be caused by flooding of the watercourse. As expected, more than 99% of the data were attributed risk class 1, implying a statistical risk of being flooded (by a river) of less than every 200 years. Since heavy precipitation events were not part of general building insurances, very little data about damages due to heavy rain events were collected. For this study, only 5 primary insurance company provided data, amounting to 13,137 damage claims attributed to flooding.

In addition the insurance claims, data were analysed in regard to their monetary amount. The losses were averaged over each risk zone separately. Furthermore, the cause of loss could be distinguished between local pluvial flooding, backwater and fluvial flooding. In this study only losses to buildings are investigated. Damage to content is too sparsely reported to attempt any statistical analysis.

Losses did not differ much from one risk zone to the other. They range from 1% to 1.5% on average, whereas losses due to river flooding ranged from 1.3% to 4.2%. These values have to be looked at with care, since the data base is relatively small.

One primary insurance company provided data from 1995 until 2009 separated into pluvial and fluvial damage. Only 18% of the losses were attributed to fluvial damage. Thus, 82% of the losses were due to heavy rain events. Nevertheless, the average loss for fluvial flooding of 5,365 Euro was much higher than the 3,080 Euro for fluvial damage.

Since no information about building type, age, or existing cellar was available, the construction of a dependable damage function has to be very crude. This aspect will be elaborated on at a later point in time.

APPLICATION

Hazard module

The hazard module comprises all parameters necessary to create a synthetic event set.

This set should reflect the realistic characteristics well, but also contain events of an intensity that has not yet been observed. For the intensity, amount of rain and duration in a single point, the rain gauges with records up to 86 years build a very good base to fit statistical distributions.

The radar images are used to determine the extent more explicitly as well as the orientation of the cell and the direction of its path. To be able to simulate the large number of cells, a geometrical form was chosen. It was found, that an ellipse reflects well the shape of the cell. An ellipse can be described with three variables; two semi axes and its orientation in space. Minor and major axis were chosen in a way that the area of the cell remains unchanged. Via cell tracking, the corresponding cells of concurrent images can be aggregated to one cell over time. To reduce calculation time for the simulation, it is envisioned to simulate the feature over the complete time interval.

The analysis of the 11 years of the nationwide composite data did not aid in defining risk zones. Topography is visible; however, high frequency values due to artefacts are more present than the recognition of risk zones. Similar results were found in the URBAS study. The authors analysed the 16 individual radar sites and used corrective operations. Heat island effects are held responsible for a larger frequency over Hamburg and the Ruhr area and topography with its obstructive characteristics leading to precipitation on the windward hillside (URBAS, 2008).

Radar history does not seem to allow yet for any reliable statements concerning distribution of heavy rain cells. For this reason, the desired frequency map was not created. More research will be put into this issue. In the interim a uniform distribution, corrected for topographic effects, is used.

The second part of the module is the linking of characteristics to the cell structures. Each cell will obtain its characteristics according to its extent, time span and geographical location in Germany. Currently, the calculation time is not tolerable so that the setup has to be reworked, thus, individual results will be presented at a later stage.

Vulnerability module

In the vulnerability module the hazard components are linked to their impact to buildings. The first step is a Boolean operation. According to the ellipse parameters it will be defined whether the ellipse will result in damage on the ground or not. Secondly, the degree of damage is calculated. To get a notion of the possible loss dimension of the events, the insurance data served as a basis. The information is sparse, but until better damage data can be acquired, the 8,478 damage claims have to suffice. The fact, that more than 99% of the reported losses to buildings due to pluvial flooding were recorded in risk class 1, leads to the predication that pluvial flooding has to be accounted for additive to fluvial flooding. In the study area the emergency runs and insurance claims addresses were linked to the census data to identify the percentage of risks affected. Maximum values of 80% of the buildings were reached; this is insurable risks not insured risks. With improving accessibility to damage data, the vulnerability module will be improved constantly.

Exposure module

The exposure module contains the risks insured (in this case buildings covered for flooding) due to heavy rain events. These addresses are geo-referenced and then intersected with each simulated event. According to the characteristics of the synthetic event, losses will be produced. These losses are then sorted in descending order to create an events loss table. Each event can then be attributed a return period. These return periods are completely independent from the recurrence intervals of the meteorological event but depend on the underlying exposure. With increasing building density, chances of generating losses also rise.

CONCLUSION AND OUTLOOK

The risk of local flooding, defined as pluvial flooding, is a significant contributing factor to flood losses. Conventionally, flood plain risks zones are considered when assessing the risk of flooding. Inundations due to fluvial events are rarely addressed (Kron, 2010). The risk is frequently included in elementary coverages but not taken into account when calculating the rate for the policy. A model structure was outlined to be able to simulate the latter risk. The model consists of a hazard module, a vulnerability module, an exposure module and a financial module. The latter is designed for representing insurance structures and is not part of this study. The module approach allows for updating each module individually. With new data, the affected part of the model can be adjusted without having to redesign the other components. Evaluation of the model was still ongoing when this article passed for press. The target is to implement the heavy rain component in an existing flood model.

Acknowledgements The author thanks Emschergenossenschaft/Lippeverband, Essen and DWD (German Meteorological Service) for providing rain data from gauges and radar via Leibniz University Hannover. Thanks go also to E+S Rückversicherung AG, Hannover who gave access to insurance data, census data and software.

REFERENCES

Busch, S. & Willjes, S. (2009): Simulating heavy rain damage in urban areas on a long term scale. Proceedings Final conference of the COST Action C22 Urban Flood Management, 25. – 28. November 2009

Busch, S. (2010): Schäden durch Starkniederschläge. Tagungsband *Risikomanagement und Vorsorge für kritische Wasserinfrastrukturen.* Österreichischer Wasser- und Abfallwirtschaftsverband. 27 January 2010, Innsbruck

DWD (1997) Starkniederschlagshöhen für Deutschland – KOSTRA. Eigenverlag des Deutschen Wetterdienstes, Offenbach

Krämer, S. (2007) *Quantitative Radardatenaufbereitung für die Niederschlagsvorhersage und die Siedlungsentwässerung*. Ph.D. Dissertation submitted to the Faculty of Civil Engineering and Geodesy, Leibniz University of Hanover

Kron, W. (2010) Elementarschutz: Weit weg vom Gewässer – sicher vor Überflutung? *Versicherungswirtschaft* (3), 210 - 215

Nelsen, R. (1999): *An Introduction to Copulas*. Springer Verlag. New York

Otto, M. (2009): Modellierung von Hagelschäden in der PKW-Kakoversicherung in Deutschland. Diploma Thesis submitted to the Faculty of Forest-, Geo- and Hydro-Sciences, Technical University of Dresden

Pfeifer, D., Straßburger, D., Philipps, J. (2009) Modelling and simulation of dependence strutures in nonlife insurance with Bernstein copulas. Paper presented at the 39th International ASTIN Colloquium, Juni 1st to 4th 2009, Helsinki

Renard, B. & Lang, M. (2007) Use of a Gaussian copula for multivariate extreme value analysis: Some case studies in hydrology. *Advances in Water Resources* **30** (2007), 897 – 912

Salvadori, G., de Michele, C., Kottegoda, N. T. & Rosso, R. (2007) *Extremes in Nature – An Approach using Copulas,* Water Science an Technology Library (**56**), Springer

URBAS (2008) http://www.urbanesturzfluten.de/, visited 27 February 2008

Vannitsem, S. & Naveau, P. (2007) Spatial dependences among precipitation maxima over Belgium. *Nonlinear Processes in Geophysics.* **14**, 621 - 630

Willjes, S. (2009) *Copulas in der Naturgefahrenmodellierung*. Diploma Thesis, Faculty of Mathematics and Science, University of Oldenburg