10th International Conference on Hydroinformatics
HIC 2012, Hamburg, GERMANY

CORRELATIONS BETWEEN RAINFALL DATA AND INSURANCE DAMAGE DATA ON PLUVIAL FLOODING IN THE NETHERLANDS

SPEKKERS, M.H. (1), TEN VELDHUIS, J.A.E. (1), KOK, M. (1), CLEMENS, F.H.L.R. (1)
(1): Delft University of Technology, P.O. Box 5048, 2600GA, Delft, the Netherlands

The objective of this paper is to establish relationships between rainfall extremes and damage data from Dutch insurance industry. Rainfall data are based on a network of 33 automatic rain gauges held by the Royal Netherlands Meteorological Institute. Rainfall characteristics, such as peak rainfall intensity and rainfall volume, are correlated with damage statistics of claims in the vicinity of the rain gauges. The results show that rainfall-related damage mainly occurs during summer seasons. There is a weak relationship between property damage and rainfall intensities and between property damage and rainfall volumes for summer events. More data is needed to confirm these relationships. In a subsequent study this will be investigated by using weather radar data to obtain a higher spatial rainfall resolution and thus be able to include more insurance data.

INTRODUCTION

Intense rainfall may generate overland flows and pooling in urban areas, causing damage to buildings, infrastructure and inconvenience to people. This process is commonly known as pluvial flooding. Due to a lack of data on the consequences of flooding, assessment of damages related to pluvial flooding remains a challenge and so is the identification of explanatory factors for damage variations [1,2].

Existing methods to assess damage mainly focus on the estimation of tangible damage by applying stage-damage curves [3]. These functions are not applicable to floods characterised by small flood depths (<30 cm) and small geographical extension [4]. The assessment of damage related to small-scale flooding should take into account the relative short time scales at which urban drainage processes respond to rainfall and local characteristics, such as the functioning of urban drainage systems.

Insurance databases are a promising source of quantitative damage data related to pluvial flooding, although access to data is often difficult [5]. For this study, an insurance database containing a series of 20 years of information on water-related damages to private properties and content in the Netherlands has recently been made available for research by the Dutch Association of Insurers [6]. This offers opportunities to analyse actual data related of flood event damages. This study focuses on the relationships between rainfall and flood damage. This is done by correlating rain gauge data from the Royal Netherlands Meteorological Institute with insurance damage data. An attempt is made to establish relationships between rainfall peak intensity, volume and duration and damage statistics. A better understanding of relationships between rainfall and damage helps, for example, to improve extreme rainfall alerts at predicting flooding [7].

DATA DESCRIPTION

Insurance data

The insurance database covers water-related damages to private buildings and building content in the Netherlands. The database is systematically stored in a data warehouse environment by the Dutch Association of Insurers. Data related to property and content damage is available from 1986 until 2010 and from 1992 until 2010 respectively. The database consists of data from a number of large insurance companies in the Netherlands (about 20-30% of the Dutch market related to property and content policies). In the Netherlands, nearly everyone is insured for water-related damages through their property and content insurance. Commercial buildings are covered in a seperate database.

Water-related damages is a broad term for different kind of damages and can be divided into two main groups based on the cause of damage: 1) non precipitation-related damages (e.g. leakages of drinking water pipes) and 2) precipitation-related damages (e.g. rainfall leaking trough roofs, hail damage and flooding from sewer systems). Damage causes are poorly defined in the database; different insurance companies use different ways to categorize damage claims and the process of classifying damage claims highly depends on the employee at the insurance company that is responsible for processing the claims. This information therefore can only limitedly be used to filter records.

Damage due to flooding from sewer systems or regional watercourses is included in nearly all insurance policies after 2000 following an advice issued by the Dutch Association of Insurers. Damage due to flooding from sewer systems or regional watercourses should be directly and solely related to local extreme rainfall for a claim to be accepted. In addition, the rainfall event should have a minimum intensity to be considered as 'extreme'. The Dutch Association of Insurers defined 'extreme' rainfall when rainfall intensity is higher than 40mm in 24 hours, 53mm in 48 hours or 67mm in 72 hours. The reasoning behind this is to prevent reoccurring claims of damaged buildings that are built on very vulnerable locations. However, it is unclear how and to what extent fulfilment of this requirement is examined by the insurance companies.

The insurance database consists of four sub databases: 1) a damage claim database with records related to property, 2) a damage claim database with records related to building content, 3) a database with policy holder information related to property insurances and 4) a database with policy holder information related to content insurances. The variables from the damage databases that are used in this study are the damage value that has been paid out by the insurance company, the date on which the damage occurred and a unique policy identification key. The latter is used to lookup the address of the policy holder in the policy holder databases, which is given by means of a 4-position postal district code. The damage value is based on the replacement values of materials and objects. Small damages (<1000 euro) are often directly paid out by insurance companies, whereas for larger damage values (>1000 euro) a damage expert will assess the insured damages. This study uses a subset of the insurance database based on temporal and spatial availability of rain gauge data, which will be explained in more detail in section 'Methodology'.

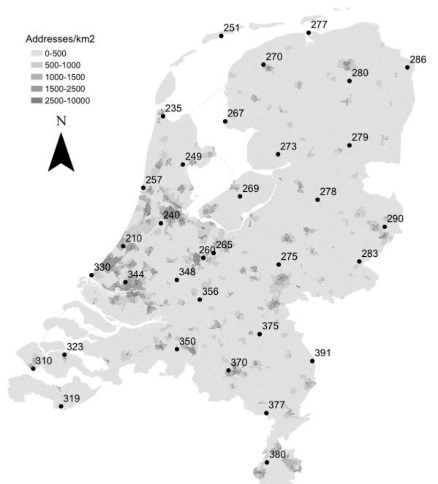

Figure 1. Location of 33 KNMI weather stations and their identification number as used by the KNMI. The urban density (=addresses/km2) is presented in greyscales.

Rain gauge data

Rainfall data that are used for this study are based on a rain gauge network, consisting of 33 automatic weather stations held by the Royal Netherlands Meteorological Institute (KNMI). The locations of the weather stations and their identification numbers are depicted in figure 1. The spatial density of the network is about 1 station every 1000 km^2. Most weather stations are located in rural areas or just outside urban areas. The low network density and poor coverage of urban areas impose some limitations. Rain gauge measurements are only representative within a limited radius from the rain gauge. Therefore only damage claims in the vicinity of weather stations can be used for analyses.

Rain gauge data is available since April 2003 with a 10-minute temporal resolution. Insurance data is available from 1986/1992 (property/content) until 2010 and therefore rainfall data from April 2003 until 2010 is used. The availability of rainfall data for the top 8 most populated areas is listed in table 1.

Table 1. Summary per station (April 2003 until 2010). Only the top 10 stations in areas with the highest number of households are listed (based on 10-km radius).

			Property database			Content database		
Station No.	Availability rainfall data	No. of households	No. of claims	No. of policies	Policies/ households	No. of claims	No. of policies	Policies/ households
344	100%	299374	3443	31534	0.11	6808	95615	0.32
240	100%	172155	2308	23573	0.14	2762	58682	0.34
260	100%	105715	1638	19346	0.18	1742	38772	0.37
370	81%	64071	2602	16423	0.26	1549	29329	0.46
350	98%	59372	1853	11849	0.20	1061	19861	0.33
380	100%	56421	2116	14838	0.26	1526	26675	0.47
210	92%	53184	2818	23458	0.44	1963	42451	0.80
290	73%	37528	1920	15451	0.41	1248	23522	0.63

METHODOLOGY

The processing of insurance and rainfall data compromises three steps. First, insurance data have undergone a number of pre-processing operations. Second, mutual independent rainfall events are determined from rain gauge data and for each rainfall event rainfall characteristics are derived. Third, rainfall characteristics and damage data are combined in to a single data set, which finally will be used for data analyses. All data-mining and statistical computing are carried out within R 2.13.2 software environment.

Pre-processing insurance data

The address of the policy holder is added to the damage records by looking up the address in the policy holder databases. The address is known at the level of 4-position postal districts. Typical surface areas of the postal districts are 1-5 km^2 for urban areas and 10-50 km^2 for rural areas. Duplicate records were removed, as well as records with missing or incorrect date, location or damage value. Damage values before 2002 were converted from guilder to euro (1 guilder = 0.45 euro). All values are in 2009 euros. Every value associated with a year before 2009 was adjusted for inflation using Consumer Price Index [8].

Damage records from April 2003 until 2010 were selected, restricted by the availability of the rain gauge data. Moreover, only damage claims in the vicinity of the rain gauges are used. Convective rainfall can be very local and intense and therefore rain gauge measurements are only representative within a certain radius from the rain gauge. The radius depends on many factors, such as the spatial distribution of the rainfall event and wind direction and velocity. The use of weather radar images could be useful to estimate these parameters, however this is outside the scope of this study. A subset of damage data has been made for records within a varying radius from a rain gauge: 2, 4, 6, 8 and 10 km. For each postal district the distance between the district's centroid and its nearest weather station was determined using a map shape file of all Dutch postal districts (version: March 2011). Figure 2 shows an example of a rain gauge location and its neighbouring postal

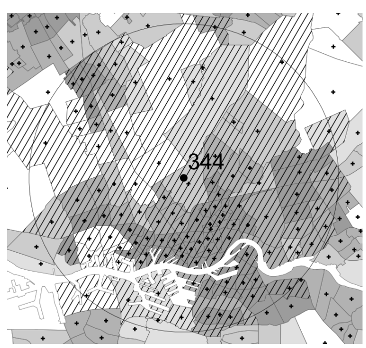

Figure 2. Selecting claims based on their distance to nearest weather station (black dot). The crosses represent the centroids of postal districts. The shaded areas are the postal districts which centroid is within the 10-km radius from the weather station (circle).

districts. In table 1 the number of claims, number of policies and the number of policies per household are listed for the subset related to the areas within 10-km radius. Subsequently, damage records are aggregated by calculating the number of claims, average height of a claim and the total damage of all claims by day and by station.

Data have been normalised based on the number of insurance policies per area and year to be able to compare damage statistics between areas and between years. The number of insurance policies in the 33 station areas for the years 2003-2009 are determined based on policy holder databases. The damage statistics are normalised by calculating the number of claims per 1000 policies and the total damage per 1000 policies (see table 2).

The database contains also records that are not related to rainfall; the number of claims on dry days is less than 20 claims per station area. Damage claims on days that have considerably more claims than average are more likely to be related to rainfall; for this analysis only records with at least 20 claims have been included.

Deriving rainfall characteristics

In order to derive rainfall characteristics from rain gauge data, independent rainfall events have been defined. Two rainfall events are considered independent when there is a certain dry period in between without or with minor rainfall. This definition is based on the knowledge that urban drainage systems have a lagged response to rainfall and need a certain time to restore to equilibrium state. If systems are in equilibrium state before a rainfall event, flood damage can be related completely to that rainfall event. A typical time for sewer systems to restore to equilibrium state is between 10 and 20 hours. The dry period is varied (6, 12, 24 and 48 hours) to study the sensitivity of the parameter on the results. Rainfall intensities smaller than 0.17 mm/10 min (=1 mm/hour) are treated as no rainfall. This prevents unrealistic long rainfall events to exist because of very small rainfall volumes between rainfall events. For each independent rainfall event, the start and end date/time, volume, peak intensity and duration were determined.

Linking rainfall characteristics to damage records

Given the day of a record and its nearest rain gauge (see table 2), the rainfall event that intersects with that day was looked up. A rainfall event intersects a day if the rainfall event starts before 24:00 and finishes after 00:10. The peak intensity, volume and duration of the rainfall event were added to the records in table 2. For three records rainfall data was missing, all other records could be linked to a rainfall event.

Table 2. Structure of the aggregated table. Only 5 observations are given.

		Damage statistics			Rainfall characteristics		
Station No.	Date	Average damage [euro]	Damage/ 1000 pol. [euro]	No. of claims /1000 pol.	Duration [hours]	Peak Intensity [mm/10min]	Volume [mm]
350	2005-09-10	1415	6622	4.7	5.3	16.7	54.2
210	2007-07-16	1790	5229	2.9	19.8	11.5	35.1
344	2009-05-26	1538	2392	1.6	18.3	7.9	28.5
380	2003-06-30	1500	2183	1.5	16.0	5.6	15.6
210	2003-06-30	1235	1461	1.2	7.2	1.4	8.4

RESULTS AND DISCUSSION

Damage statistics versus rainfall characteristics

In figure 3 the total property damage per 1000 insurance policies is plotted against the peak rainfall intensity (in mm/10-minute time window, left) and the rainfall volume (in mm, right). These results are based on a 12-hour dry period between rainfall events and insurance claim data within the 10-km radius of a weather station. The damages that occured in the period April-September ('summer') and the period October-March ('winter') are marked as circles and triangles respectively. Figures for damage related to building content are not shown here.

Most observations (21 out of 26) are related to summer events, with peak rainfall intensities up to 16 mm/10min and rainfall volumes up to 55 mm. Observations related to the period October-March (5 out of 26) are associated with low rainfall intensities (<5 mm/10min) and high rainfall volumes (>45 mm). There is a weak relationship between the total damage per 1000 insurance policies and the peak rainfall intensity and rainfall volumes of summer rainfall events. High damage values are related to high rainfall intensities and rainfall volumes. Low damage values are related to low rainfall volumes, but not necessarily to low rainfall intensities. It is unclear from the shape of data scatter whether data follows a linear or curvelinear relationship. This could be because of, for example, a lack of data in the vicinity of the rain gauge or because rainfall data is not representative for the entire area within the 10-km radius.

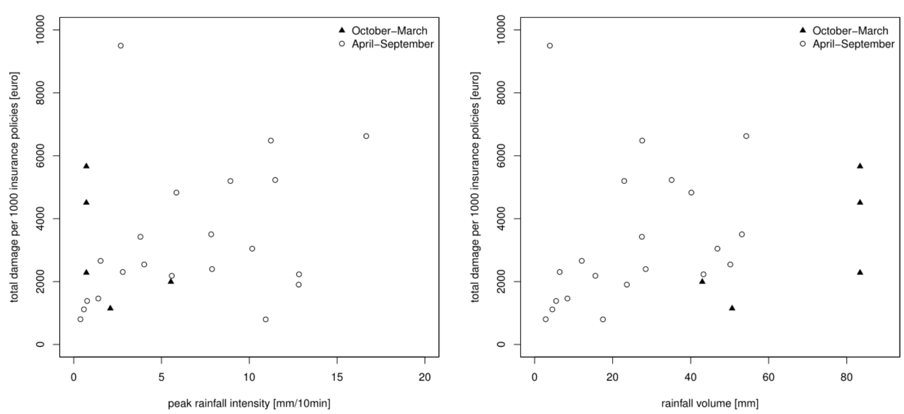

Figure 3. Total damage per 1000 insurance policies versus peak rainfall intensity (left) and rainfall volume (right). The results are based on a data within 10-km radius of a weather station and a 12-hour intermediate dry period. Circle and triangle markers respresent events between April-September and October-March respectively.

Sensitivity of intermediate dry period and radius

The damage-rainfall-relationships are affected by the selected dry period between rainfall events that and the radius within which insurance claims are selected. In figure 4 the total property damage per 1000 insurance policies is plotted against the rainfall volume. The squares represent the case for a 12-hour dry period and a 10-km radius. The crosses are the results for 8-km radius (left) and 6-hour dry period (right).

The total damage becomes slightly higher (15% on average) for most observations if the radius is reduced from 10 km to 8 km; some observations show a decrease in damage. The observation related to a total damage of 9500 euro/1000 insurance policies (point in upper left corner) decreases to 5300 euro/1000 insurance policies. This implies that in this case higher damage is related to claims in the area between the 8-km and 10-km circles. There are hardly any data within 2-km and 4-km radiuses from a rain gauge and therefore these are not plotted here.

Decreasing the intermediate dry period causes a slight decrease in rainfall volume (8-12%) for about half of the observations, whereas the other half is unchanged. Varying the dry period does not improve the rainfall-damage-relationships. Although not shown here, similar changes can be observed for rainfall durations. The peak intensities are hardly affected by the dry period.

It is assumed that the data plotted in figure 3 and 4 are partly or largerly related to precipitation, as the data are based on days with considerable more claims (>1-7 claims/1000 policies/day) than could be expected on days without rainfall (<0.5 claims/1000 policies/day). It has been checked that winter events are not related to snowfall, although the presence of hail could not be excluded. The database contains insufficient information to distinguish betweens claims related to roof leakages or dysfunctioning of urban drainage systems.

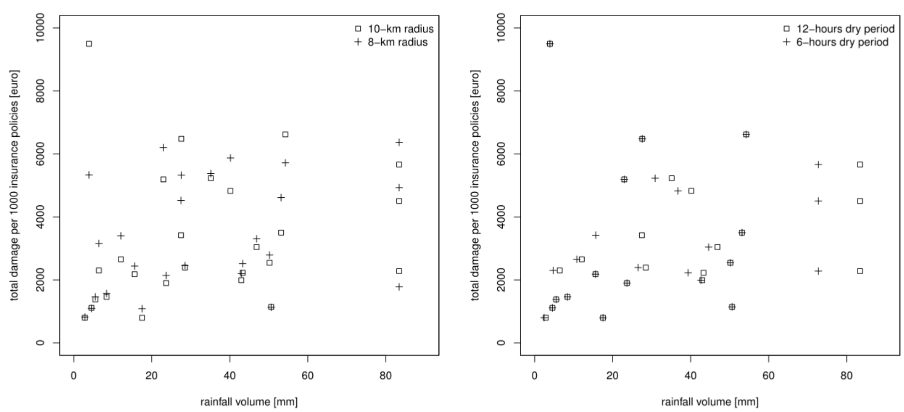

Figure 4. Effect of different choices of intermediate dry period and radius. The square markers are the results assuming a 12-hour dry period and 10-km radius. The crosses are the results for 8-km radius (left) and 6-hour dry period (right).

CONCLUSIONS AND RECOMMENDATIONS

Rainfall-related damage is mainly related to rainfall during summer seasons. Rainfall-related damage during winter seasons are associated with relatively high rainfall volumes and low rainfall intensities. There is a weak relationship between the rainfall-related property damage and rainfall intensities and volumes of rainfall events during summer seasons, but more data is needed to be able to establish relationships.

The weather stations that are used for this study are located in rural areas or just outside urban areas, which strongly limits the amount of insurance claim data that could be used for analyses. More detailed rainfall data is required to better establish rainfall-damage-relationships. In a follow-up study, a similar approach will be applied using a 10-year database of weather radar images. In that case, it would be interesting to study rainfall-damage-relationships for different seasons (e.g. for midsummer events).

The damage cause is often poorly defined or not provided by insurance companies. A better documentation of damage causes in insurance databases could therefore greatly improve future data analysis.

REFERENCES

[1] Thieken, A. H., “Methods for the Documentation and Estimation of Direct Flood Losses”, In: Zenz, G. and Hornich, R., editors, *Urban Flood Risk Management Conference*, Graz, Austria, (2011).

[2] Freni, G., La Loggia, G., and Notaro, V., “Uncertainty in urban flood damage assessment due to urban drainage modelling and depth-damage curve estimation”, *Water Science and Technology*, 61(12):2979–2993, (2010).

[3] De Moel, H. and Aerts, J. C. J. H. , “Effect of uncertainty in land use, damage models and inundation depth on flood damage estimates”, *Natural Hazards*, 58(1):407–425 (2011).

[4] Ten Veldhuis, J. A. E. and Clemens, F. H. L. R., “Flood risk modelling based on tangible and intangible urban flood damage quantification”, *Water Science and Technologyy*, 62(1):189–95, (2010).

[5] Busch, S., “Quantifying the risk of heavy rain: its contribution to damage in urban areas”, In: *Proceedings of the 11th International Conference on Urban Drainage*, Edinburgh, Scotland, UK, (2008).

[6] Ririassa, H. and Hoen, A., “Neerslag en schade: onderzoek naar het verband tussen neerslag en de schadelast voor brandverzekeraars met het oog op de klimaatverandering (in Dutch)”, Technical report, Verbond van Verzekeraars, (2010).

[7] Parker, D., Priest, S., and McCarthy, S., “Surface water flood warnings requirements and potential in England and Wales”, *Applied Geography*, 31(3):891–900, (2011).

[8] Statistics Netherlands, StatLine online database: http://statline.cbs.nl (viewed on December 2011).