

Collecting data for quantitative research on pluvial flooding

M.H. Spekkers^{1*}, J.A.E. ten Veldhuis¹ and F.H.L.R. Clemens¹

¹ *Delft University of Technology, Department of Water Management, Stevinweg 1, PO Box 5048, 2600GA, Delft, the Netherlands*

**Corresponding author, e-mail: M.H.Spekkers@tudelft.nl*

ABSTRACT

Urban pluvial flood management requires detailed spatial and temporal information on flood characteristics and damaging consequences. There is lack of quantitative field data on pluvial flooding resulting in large uncertainties in urban flood model calculations and ensuing decisions for investments in flood protection. In this paper four different data sources are discussed, based on literature and expert consultation, that are believed to be of value for the acquisition of quantitative data on pluvial flooding. Data assembled by insurance agencies on flood damage, call databases held by water authorities and emergency services and remote sensing images cover years of observational data that can be mined to obtain data on flood characteristics and occurrence. Flood monitoring using sensor technology can be effective to collect additional pluvial flood data, that is not captured by existing data sources.

KEYWORDS

Pluvial flooding; data collection, quantitative research

INTRODUCTION

Intense localised rainfall may generate overland flows and pooling in urban areas, causing damage to buildings, infrastructure and inconvenience to people. This happens when rainfall overloads the urban subsurface and surface drainage system, or is for some reason unable to enter the urban drainage system. This process is commonly known as pluvial flooding. Some severe cases show the catastrophic consequences of pluvial flooding. In the summer of 2007 the City of Hull (UK) suffered from severe pluvial flooding, causing damage to over 8600 houses and 1300 premises (Coulthard and Frostick, 2010). Similarly, in several parts of the Netherlands, intense rainfall in autumn of 1998 caused damage to 2470 houses, 1220 premises and 350 governmental agencies (Jak and Kok, 2000). However, not only severe events cause damage. The cumulative damage of smaller flood events over the lifetime of urban drainage systems can be considerable (Ten Veldhuis and Clemens, submitted). This is particularly true for lowland areas where pluvial floods have relatively high occurrence frequencies and small flood depths (<20 cm). Pluvial flood damage is likely to increase in the future because of predicted increase in rainfall extremes (IPCC, 2007) and urbanisation. This emphasizes the need to study strategies to cope with pluvial flooding.

Urban pluvial flood management requires detailed information on what causes pluvial flooding, what the consequences are, how frequently flooding occurs, what locations are vulnerable for flooding and how climate change and urbanisation effect flooding. In other words, how the urban drainage system works under extreme conditions, now and in the future. Literature addressing these questions focuses primarily on descriptive cases of severe pluvial flooding with large damages (e.g. Douglas et al., 2010; Coulthard and Frostick, 2010) or quantification by means of urban flood risk models (e.g. Zevenbergen et al., 2011). A

number of studies has been published questioning the reliability of urban flood risk models (e.g. Freni et al., 2010; Apel et al., 2008). They argue that, although a lot of advances are made in modelling flood physics and related damages for the urban context, field data are needed to calibrate models and reduce model uncertainties. Current large uncertainties are caused by incomplete knowledge of the mechanisms causing pluvial flooding and by lack of data on flood characteristics and relations between those and flood damage.

There is a clear need to better quantify flood risks and to extend and improve the collection of evidence. Field data to quantify pluvial flooding characteristics and consequences are scarce due to limited data collection efforts and low frequencies of occurrence. In this paper four different data sources are discussed which can be used for quantitative research on pluvial flooding. The overview is based on a literature review and expert consultation in the field of remote sensing and sensor technology. The paper focuses on data sources available in the Netherlands, where pluvial flooding has a high frequency of occurrence and therefore more observational data are available for research compared to hilly areas. The applicability for cases in other countries is discussed at the end of the paper.

OVERVIEW OF DATA SOURCES

Flood event data includes flood characteristics describing the probability of a certain flood event and related consequences. Typical flood variables are the frequency of occurrence, duration of the flood, water depth, flood extent, flow velocity and whether or not flood water is contaminated. Related consequences of pluvial flooding mainly consist of economic losses such as damage to buildings and their contents and to infrastructure and intangible damages due to traffic delays, road and public and commercial function closures, and evacuation of people (Hauger et al., 2006). Two different groups of flood event data are distinguished, based on definitions proposed in econometrics (e.g. Angrist and Pischke, 2009). The first group is data based on natural experiment: data collected ‘by nature’ which contain useful information on pluvial flooding. The second group consists of data related to real experiment: data collected by conducting a controlled experiment with the purpose of collecting required flood data. This section discusses four data sources, namely insurance databases, call databases, remote sensing observations and data from flood monitoring. The first three fit in the group ‘natural experimental data’, the fourth is a case of ‘real experimental data’. In this section an overview is given of the data sources, including relevant literature, use of data variables for research, limitations of the data and challenges to improve data collection. The different data sources are summarized in table 1 (see section ‘Discussion’).

Natural experimental data

Insurance databases. In most European countries, people can insure their property and content for the damaging consequences of pluvial flooding. The databases of insurance companies typically cover tens of years of information on flood damage and can therefore potentially be used for quantitative research. In the Netherlands, homeowners can insure both their property and their content; renters can only insure their content, while property damage is in that case the responsibility of the landlord. Property and content insurance is only obligatory if a mortgage is obtained from a bank. In practice most people are insured and therefore the damage records of all insurers cover a large part of all events where flood damage was claimed. The causes of pluvial flooding which are covered by property and content insurances are 1) direct rainfall on property, 2) flooding from sewer systems and 3) flooding from regional watercourses. The policies are practically the same for all insurance

companies. The third cause of flooding was included in most insurance policies after 2002 following an advice issued by the Dutch Association of Insurers in 2002 (Ministry of Transport, Public Works and Water Management, 2003). In addition, the damage should be directly and solely related to local extreme rainfall for a claim to be accepted. Flooding from rivers, sea or groundwater is not insurable in the Netherlands and therefore if pluvial flooding coincides with other flood types, the damage is not insured. Furthermore, 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 (Ministry of Transport, Public Works and Water Management, 2003). The reasoning behind this is to prevent reoccurring claims of damaged buildings that are built on very vulnerable locations. However, it is unclear on what data this condition is based and how fulfilment is examined. Nevertheless, it is likely that some flood damage is not recorded in insurance databases because they do not fulfil the policy conditions.

Damage data from insurance companies are difficult to collect. Based on a questionnaire, only 4 out of 48 insurance companies in Germany were willing to provide damage data for scientific research (Busch, 2008). In the Netherlands researchers experienced similar difficulties in the past (Kok, pers. comm.). The Association of British Insurers presently only aggregates national damage data and does not hold local damage data (Lawson and Carter, 2009). An explanation for this is that the databases contain private information of insured which insurers are not willing or able to share. Another difficulty is that insurance premiums are based on the statistics of historical records and insurers do not want to publish much information about the reasoning behind their premiums.

Nonetheless, Dutch insurers show an increasing interest in scientific research as climate change poses challenges to future insurance arrangements (Dutch Association of Insurers, 2010). They want to know if the predicted increase in rainfall extremes will cause more and/or higher damage claims. Recently, a database of flood damage to private buildings has been made available for research by the Dutch Association of Insurers. The database covers around 30% of the total sum of premiums in the Netherlands for the last 19 years. This gives opportunities to quantify spatial and temporal characteristics of flood damage and research relations with rainfall characteristics. The records include the claimed damage, location by means of a 4-digit zip code, date on which the damage occurred and policy holder data. The claimed damage is based on the replacement values of materials and objects. The damage records do not or only limitedly describe the flood cause. Inclusion of the causation could provide a more useful record. Damage data can be used to analyse explanatory variables for damage by correlating the damage data with precipitation data, building properties, digital elevation models and urban drainage characteristics. Insurance databases can also be a source of information on claim frequencies and can be used to map vulnerable locations. Similar databases as those for private flood damage to buildings also exists for damage to businesses; however, those are not available yet.

Call databases. Other sources of flood event data are call databases held by water authorities and emergency services. These databases cover several years of flood observations that can improve decision making on how to effectively solve observed problems.

In the Netherlands, municipalities are responsible for management of sewer systems to prevent pluvial flooding. More than half of the municipalities in the Netherlands have some sort of database in which they register calls from citizens experiencing pluvial flooding (Foundation Rioned, 2007). Citizens can call a call centre when there is a dysfunction of

urban drainage systems, for example if rainwater cannot be transported because of debris blocking the drainage infrastructure. The call is stored as a short text message describing the observation. The records include a date of the event and a street name and sometimes a house number. The messages contain information on causes and, to a lesser extent, consequences of the event. Occasionally, the database contains a small report of findings and solution after an on-site check. The text messages need to be categorized to make the database suitable for quantitative research, for example categories based on a number of defined causes or consequences. Municipal call databases have proven to be a suitable data source for quantitative flood risk analysis. Recent studies show that these text messages can be used to identify failure mechanisms of pluvial flooding (Ten Veldhuis, 2009; Caradot et al., 2010). The database can potentially be used to map vulnerable locations which are frequently flooded.

Although the text messages are subject to interpretation and classification errors, the database holds valuable information because of the large number of database records. The drawback of the data is, that probably not all flooded locations have an entry in the database. Other data sources can be used to validate the calls and to estimate what percentages of events is represented in the call database, for example by comparison with results from flood monitoring campaigns. At the present, databases are scarcely used by municipalities themselves to actually study pluvial flooding. Municipalities often do not have the time or resources for extensive data analysis. Another lack of the data source is that no flood depth, duration or extent are recorded. This makes it hard to judge the severity and magnitude of the event.

In addition, police and fire brigades keep track of flood events. In case of an emergency, for example to drain a flooded basement, a call is registered in the emergency database. Temporal and spatial resolutions are often limited to a date (and time) and locations by means of just a street name; they do not contain descriptive text messages of the observations like call databases and are limited to flooded buildings and do not describe street flooding. Lawson and Carter (2009) argue that by not recording the cause of the event, the usefulness of a potential valuable data source is reduced. Another drawback is that when extreme rainfall causes large number of calls in short time period, it is likely that not all events are recorded or that multiple calls are combined in one record (Busch, 2008). Police and fire brigade records can be used to cross-check with municipal call data.

Call and emergency record databases can provide more useful information in the future if flood causes and consequences are systematically recorded, for example, by improving and automating a classification system for calls. In addition, registration of flood characteristics (e.g. flood duration, flood depth) in the database can help to determine the severity of the flood event. This implies the need for additional measurements of flood characteristics (see section 'Monitoring').

Remote sensing. Remote sensing is the acquisition of information of an object or phenomenon, in this case the earth surface, from satellites or airplanes. Remote sensing can be used to gather flood data during or right after a flood event. There is a wide range of remote sensing techniques. The most promising technology in relation to flood data collection is synthetic aperture radar (SAR) (Mason et al., 2010). Other remote sensing techniques cannot be used either because the spatial resolution of the data is too coarse (e.g. satellite infrared) or the method requires daylight and/or good weather conditions (e.g. satellite photography) or the technique has low or irregular measurements frequencies (e.g. airborne photography)

(Bastiaansen, pers. comm.). SAR is a technique which is used to make high-resolution radar images of the earth surface and its properties. SAR is applied in a wide range of different applications from military to environmental sciences. There are many other potential applications and flood detection in urban areas is one of them. SAR measures time delay between emission and return of radar pulses, establishing the location and height of the surface. The radar pulses are processed into a single radar image. Flooded areas would generally appear black on radar images due to specular reflection and therefore the radar images are a potential source of flood extent data. Flood extents have been successfully determined from SAR observations for major fluvial floods in rural areas. Observed flood extents are, for example, used to calibrate flood models. In addition, it is possible to estimate flood depths if flood extent maps are combined with high-resolution ground elevation data.

Application of this technique is far more complicated for flooding in densely built-up areas; however, a recent study by Mason et al. (2010) shows promising results. Mason et al. (2010) studied the SAR observations (3 m² resolution) of the flood near Tewkesbury (UK) in 2007. Of the flood water that was visible on SAR image, 76% were correctly detected, with an associated false positive rate of 25%. They concluded that because of the side-looking nature of SAR, substantial areas of ground surface is not visible due to shadowing and unwanted reflection caused by buildings or taller vegetation. Another limiting factor is the frequency of flyovers by a satellite. Typical frequencies are once every 11 days to once every month. This questions the availability of measurements that coincide with the flood peak, in particular for pluvial floods with flood duration of only a few hours or a day. Another complicating factor is the micro climate in urban areas. Temperature, rain and wind affect the reflective properties of water surfaces and thus the way flood water appears on radar images. For rural areas where climate characteristics are relatively homogenous in space a distortion can be corrected easily, whereas urban areas with a more heterogeneous climate poses new challenges to correct climate effects (Hanssen, pers. comm.).

Over the last decade the resolution of radar images reached higher levels of detail. Currently, a small number of satellites is equipped with SAR with resolutions ≤ 3 m². Wide streets and gaps and large public spaces can clearly be seen on current radar resolutions (3 m²), whereas none of this was visible on previous radar observations (12.5 m²) (Mason et al., 2010). In the near future the resolution will reach even more detail (≤ 1 m²) (Hanssen, pers. comm.). With increasing radar detail it will be more likely that also smaller flow and pools will be detected. Nevertheless, radar shadowing and unwanted reflections continue to exist, and an increasing spatial resolution will not solve this.

Real experimental data

Monitoring. If a certain flood phenomenon is subject of study, but is not captured with natural experimental data, a monitoring campaign can be carried out to measure flood characteristics systematically, such as flood extent, depth and duration. For example, municipal call databases do not register flood depths and extents, which can be used for model development and calibration. In addition, measurements can be used to validate existing data and models. For example, measurements of flood locations can be used to check to what extent call databases cover flood events and to calibrate overland flow models.

Sensor technology can be effective to measure flood characteristics at predetermined locations. Nowadays, sensors can be very cheap and small and have relative low-energy consumption (Akyildiz et al., 2002), making it possible to deploy a dense network of ‘water’

sensors for years. The ‘water’ sensors detect the presence of water at a specific location or measures the local water height. Sensors can be installed virtually everywhere, for example, on streets or near vulnerable object to monitor the flood frequency, duration and extent. They can be modified in such a way that the sensor is most of its time in a standby modus and only measures when triggered (Hut, pers. comm.). A wireless interface can be installed in the sensor for data communication between sensors. The network as a whole can be connected to the internet for easy data collection through a web interface. Examples of such monitoring networks for flood detection are scarce, although they show promising results in collecting flood data (See et al., 2009; Chang and Guo, 2006). Monitoring is particularly sensible in areas where pluvial flooding occurs frequently throughout the year. This is a complicated factor because pluvial flooding is often very localised and unpredictable. A challenge is therefore to select suitable sensor locations and to keep some flexibility in the experimental set up. A less static way of monitoring can be the use of citizens as ‘dynamic sensors’. Nowadays most mobile phones are equipped with internet, GPS and photo camera, enabling people to participate in data collection (Hut, pers. comm.). Participants can send photos and meta data (e.g. their location and time) to a central database where ‘calls’ are systematically stored.

DISCUSSION

In this paper four data sources are discussed that are believed to be of value for future studies on pluvial flooding. The data sources mentioned are available for the Netherlands; it is expected that the results of this study are also applicable to cases in other countries, even if flooding frequencies are lower and longer data series must be collected for proper statistical analysis. Most European households and company buildings have similar content and property insurance arrangements and the insurers’ interest in thorough analysis of flood damage data in relation to explanatory variables is likely to increase in view of future developments, as it did in the Netherlands. Radar images are available world wide, while high-resolution SAR observations are limited because of the low time resolution. Therefore, the SAR observations are likely to be more useful for countries with large flood durations (>day), as this increases the likeliness of having an observation that coincides with the flood event. Instead of looking for observations that fit a specific flooded area, it is better to start studying the radar image as a whole and try to identify flooded areas on it, i.e. to look for cases where data is available. Precipitation data can be useful to identify days with intense rainfall and look for those radar images that are likely to show flooded areas. Call databases of water authorities and emergency services are likely to be more effective for countries with high frequency, low impact flooding, because call centres can be overwhelmed by calls in case of a high impact flood, making the database less reliable. To the knowledge of the authors, sensor technology has never been applied for the systematic monitoring of pluvial flood depth, extent and duration so far. Application of sensor technology combined with data-communication systems is believed to be a useful and relatively cheap manner to monitor flood events, in particular in countries with high frequency flooding. A first step could be a pilot study in an area which is known to be flooded frequently.

The data sources all capture different variables of pluvial flooding at different spatial and temporal scales, as summarized in table 1. None of them can solely cover all aspects of pluvial flooding. It depends on the research subject and approach and the nature of pluvial flood events studied which data source or combination of data sources are best to be used. Data can be combined, for example, for verification and calibration of flood and damage models. Results of flood monitoring can be used to check the extent to which municipal call

Table 1. Overview of data sources

Data source	Description	Type	Temporal resolution	Spatial resolution	Flood event data variables	Limitations	Challenges
Insurance databases	Damage claims by flood affected people and companies	natural	day	street-district	-material damage to building and content -claim frequency -identification of vulnerable locations	-Hard to collect -Privacy issues	-Inclusion of the flood cause in damage record
Call databases	Calls of flood observations recorded by water authorities or emergency services	natural	day	building-street	-flood frequency -flood locations -flood causes -consequences	-May not cover all events -No flood depth, nor extent measurements -Interpretation errors -Classification errors	-Improve and automate classification system based on text messages -Registration of flood characteristics in database
Remote sensing	Synthetic aperture radar observations for the detection of flood water.	natural	week-month	street	-flood extent -flood depth	-Low temporal resolution -Radar shadowing and unwanted reflections due to high buildings -Complex urban micro climate	-Higher radar resolution -More frequent overfly times
Monitoring	Measuring flood characteristics using sensors	real	minute-hour	building-street	-flood depth -extent -duration	-Selection strategy of sensor location -Vandalism	-Theoretical framework on selecting sensor locations

databases cover flood events. Flood extents from remote sensing observations can be used to select suitable areas for monitoring. In addition, municipal call databases and insurance databases can be compared for cross-checking. For example, to check correlation between call frequencies and claims frequencies.

The authors will carry out a case study for the City of Rotterdam (the Netherlands) where a combination of insurance, call and monitoring data will be used to aim for a more holistic and data-driven approach of pluvial flood management.

CONCLUSIONS

- There is lack of quantitative data on pluvial flooding resulting in large uncertainties in flood model calculations and ensuing decisions for investments in flood protection and mitigation.
- A more data-driven approach in pluvial flood management is required to improve understanding of flooding processes and to improve the quality of models.
- A number of data sources exists which potentially can be used for quantitative research on pluvial flooding: insurance databases, call databases, remote sensing observations and monitoring using sensor technology.
- Insurance databases cover tens of years of information on flood damage that can be used to analyse explanatory variables of damage characteristics and occurrences. Although damage data is hard to collect, insurance companies show an increasing willingness to contribute to scientific research.
- Call databases held by water authorities and emergency services cover years of flood observations and have proven to be useful for quantitative research.
- Synthetic aperture radar is a promising remote sensing technique for flood detection in urban areas, however, the most limiting factors at present are the low overfly frequency of the radar satellite, the complex urban climate distorting the radar images, radar shadowing and unwanted reflections.
- Monitoring using networks of sensors can be effective for the purpose of collecting required pluvial flood data that is not captured by existing data sources.

REFERENCES

- Akyildiz I., Su W., Sankarasubramaniam Y. and Cayirci, E. (2002). A survey on sensor networks. *IEEE Communications Magazine*, 40, 102-105.
- Angrist J.D. and Pischke J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, USA, ISBN 978-0-691-120355.
- Apel H., Merz B. and Thieken A. (2008). Quantification of uncertainties in flood risk assessments. *International Journal of River Basin Management*, 6, 149-162.
- Busch S. (2008). Quantifying the risk of heavy rain: its contribution to damage in urban areas. Proceedings of the 11th Int. Conference on Urban Drainage, Edinburgh, Scotland, UK.
- Caradot H., Granger D., Rostaing C., Cherqui F. and Chocat B. (2010). Risk assessment of overflowing sewage systems: methodological contributions of two case studies (Lyon and Mulhouse) (in French). NOVATECH 2010.
- Chang N.-B. and Guo D.-H. (2006). Urban flash flood monitoring, mapping, and forecasting via a tailored sensor network system. Proceedings of the 2006 IEEE International Conference on Networking, Sensing and Control, 757-761.
- Coulthard T. and Frostick L. (2010). The Hull floods of 2007: Implications for the governance and management of urban drainage systems. *J. of Flood Risk Man.*, 3, 223-231.
- Douglas I., Garvin S., Lawson N., Richards J., Tippet J. and White I. (2010). Urban pluvial flooding: a qualitative case study of cause, effect and nonstructural mitigation. *J. of Flood Risk Man.*, 3, 112-125.
- Dutch Association of Insurers (2010). Rainfall and damage. Report 2010/rap/1112/SJANK, Dutch Association of Insurers, The Hague, the Netherlands.
- Foundation Rioned (2007). Inquiry of flood problems in the built environment (in Dutch). Report Foundation Rioned, Ede, the Netherlands.
- Freni G., La Loggia G. and Notaro V. (2010). Uncertainty in urban flood damage assessment due to urban drainage modelling and depth-damage curve estimation. *Wat. Sci. Tech.*, 61, 2979-2993.
- Hauger M., Mouchel J.-M. and Mikkelsen P. (2006). Indicators of hazard, vulnerability and risk in urban drainage. *Wat. Sci. Tech.*, 54, 441-45.
- IPCC (2007). *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press.
- Jak M. and Kok M. (2000). A database of historical flood events in the Netherlands. *Flood Issues in Contemporary Water Management*. Ed: Watt, et al., eds. NATO Science Series 2, Environmental Security, 71, Dordrecht, the Netherlands: Kluwer Academic Publisher.
- Lawson N. and Carter J. (2009). Greater Manchester Local Climate Impacts Profile (GMLCIP) and assessing Manchester City Council's vulnerability to current and future weather and climate. Report, University of Manchester.
- Mason D., Speck R., Devereux B., Schumann G.-P., Neal J. and Bates P. (2010). Flood detection in urban areas using TerraSAR-X. *IEEE Transactions on Geoscience and Remote Sensing*, 48, 882-89.
- Ministry of Transport, Public Works and Water Management (2003). The insurability of damage due to rainfall and pluvial flooding (in Dutch). Report, Ministry of Transport, Public Works and Water Management.
- See C., Horoshenkov K., Tait S., Abd-Alhameed R., Hu Y., Elkhazmi E.A. and Gardiner J. (2009). A Zigbee based wireless sensor network for sewerage monitoring. Asia Pacific Microwave Conference, 731-734.
- Ten Veldhuis J.A.E. and Clemens F.H.L.R. (submitted). How the choice of flood damage metrics influences urban flood risk assessment. Submitted for a special issue of the Journal of Flood Risk Management.
- Ten Veldhuis J.A.E., Clemens F.H.L.R. and van Gelder P.H.A.J.M. (2009). Quantitative fault tree analysis for urban water infrastructure flooding. *Structure and Infrastructure Engineering*. doi: <http://dx.doi.org/10.1080/15732470902985876>.
- Zevenbergen C., Cashman A., Evelpidou N., Pasche E., Garvin S. and Ashley R. (2011). *Urban flood management*. CRC Press/Balkema, Leiden, the Netherlands, ISBN 978-0-415-55944-7.