Innovative approaches in operational flood management

Risk-based forecasting and application of social media in disaster response

Dirk Eilander, Daniel Bachmann, Peter Gijsbers and Arnejan van Loenen

Abstract

An integrative disaster management approach consists of four different phases: mitigation, preparation, response and recovery. Adequate information for decision support during all phases is a key to strengthen resilience against flooding. During the preparation phase forecasted information about the event can help to prepare and take adequate and effective emergency response measures. During the response phases reliable information about the situation on the ground is crucial for the allocation of affected areas and people. This paper presents two innovative approaches to generate such information. The first approach extends existing flood forecasting systems with elements of strategic flood risk analysis, such as probabilistic failure analysis, two dimensional flood spreading simulation and the analysis of flood impacts and consequences. With the help of this information emergency measures at the flood defence line (e.g. temporally dike strengthening) and within the protected area (e.g. evacuation of people) can be operationally planned, adapted and triggered. The second approach exploits observed information shared via social medias (e.g. Twitter) about the physical characteristics of floods, such as flood depth and the location. These observations are used in combination with a Digital Elevation Model (DEM) to derive flood extent in real-time during the response phase. An overview of the current situation gets available. A coordination of emergency measures is supported.

Zusammenfassung

Ein integratives Katastrophenmanagement besteht aus vier verschiedenen Phasen: Vorbeugung, Vorbereitung, Gefahrenabwehr und Wiederaufbau. In jeder dieser Phasen sind adäquate Informationen zur Entscheidungsunterstützung notwendig. In der Vorbereitungsphase können vorhergesagte Informationen über die Ausprägung und Folgen eines Ereignisses wesentlich die Entscheidungsfindung über geeignete und effektive Notfallmaßnahmen unterstützen. Während der Phase der Gefahrenabwehr sind Echtzeit-Informationen über die Lage der betroffenen Gebiete und die Situation der Personen entscheidend, um entsprechende Maßnahmen einleiten zu können. Dieser Artikel beschreibt zwei innovative Ansätze, um entsprechende Informationen zu generieren. Der erste Ansatz erweitert bestehende Hochwasservorhersagesysteme um

Analysen, welche bereits in einer strategischen Hochwasserrisikoanalyse eingesetzt werden, wie der probabilistischen Zuverlässigkeitsanalyse, die Modellierung der Überflutungsausbreitung und die Analyse der Konsequenzen. Mit Hilfe dieser Informationen kann die Planung und Ausführung von Notfallmaßnahmen an der Hochwasserschutzlinie (z. B. notfallmäßige Deichverstärkung) oder im betroffenen Gebiet (z. B. Evakuierung) unterstützt werden. Der zweite Ansatz kombiniert beobachtete Informationen über ein Überflutungsereignis (z. B. Wasserstände), die über soziale Medien (z. B. Twitter) veröffentlicht werden, mit einem digitalen Höhenmodell, um Echtzeit-Informationen über die Überflutungsausbreitung zu erstellen. Ein Überblick über die aktuelle Situation und damit eine bessere Koordination von Hilfsmaßnahmen wird möglich.

1 Introduction

An integrative disaster management approach consists of four different phases: mitigation, preparation, response and recovery (see Fig. 1). The objective of the mitigation phase is the reduction of the probability of a disaster and the mitigation of impacts. This phase includes primarily long term planning of mitigation measures (strategic planning). However, a residual risk remains. Historical figures prove the relevance of the residual risk: 5000 casualties and about 27 bn € damages are caused by hydrological disaster worldwide for the year 2014 (Munich RE, 2015). For Germany the flood of 2013 in the Elbe and Donau region also caused about 7 bn € damages at public infrastructure, private households, agriculture and industry (BMI, 2013).

The second phase (preparation) and third phase (response) try to reduce the residual risk. Preparedness plans, emergency training, risk communication and the set-up of forecasting- and warning systems are measures of the preparation phase, which can be implemented before a disaster. The response phase short before and during an event puts the measures of the preparedness phase into action to avoid casualties and damages.

The recovery phase after a disaster includes activities restoring the lives of affected people, the function of affected infrastructure and economy.

The boundaries between the different phases are fluid and can even overlap. For example the forecast of an event takes place before an event (preparation phase); warnings based on such a forecast can be used to trigger actions short before and during the event (response phase), such as dike strengthening with sand bags before or emergency evacuation during an event.

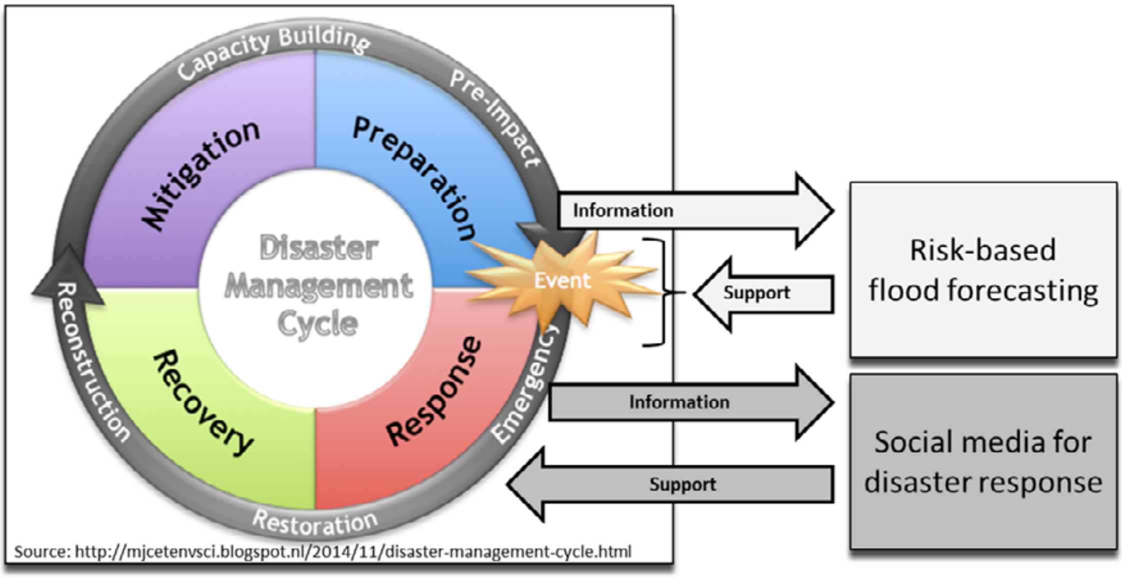

Figure 1: New approaches for disaster management

This paper presents two new approaches to improve information for decision support in disaster management (see Fig. 1). The first approach aims to enable probabilistic forecasts of the failure of flood defences as well as flood spreading, consequences and impacts short before an event.

The second approach uses new information from social media to monitor the situation on the ground during an event. Both approaches will support the decision maker with additional information for effective emergency response.

2 Risk-based flood forecasting

2.1 Concept

2.1.1 Current flood forecasting

In recent years model-based flood forecasting systems have been established all over the world for different spatial scales. In general, these model-based forecasting systems include a process chain starting with a meteorological model predicting the amount and spatial distribution of precipitation, wind and temperature over time, followed by a hydrological model which transfers the meteorological output data into hydrological data; mainly discharges in rivers (see Fig. 2).

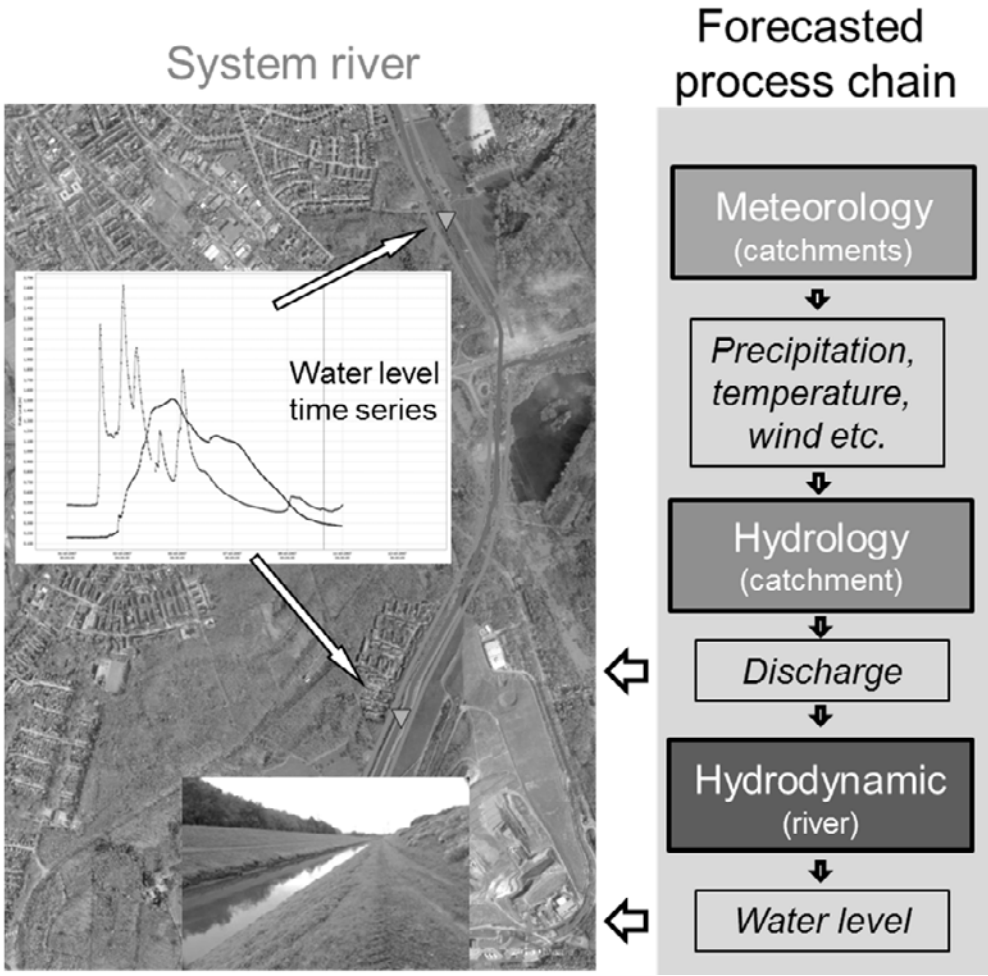

Figure 2: Current application of model–based flood forecasting: from meteorological forecast to a water level forecast

In some cases a next step in the process chain, the transfer of discharge into water level by a hydrodynamic model, is applied. However, the hydrological/hydrodynamic-based information alone barely supports answering crucial questions within an emergency response framework, such as:

- Where and when is a failure of my flood defence system possible?
- Where and when should I reinforce my flood defence system?
- What are the consequences and impacts of an overflow or a breach?
- Which areas should I evacuate first and how can I evacuate most effectively?

Currently, these questions are answered based on expert knowledge and/or – if available – pre-calculated information (e.g. about dike performance or flood spreading behaviour in the area).

The presented concept is based on extending the currently applied forecasted process chain (see Fig. 2) by modelling techniques already applied and combined for several years in strategic risk assessment. These methods are probabilistic failure analysis, two dimensional flood spreading simulation and the analysis of flood impacts and consequences. These extensions generate risk-based forecasted information to support decision makers within the emergency response by finding answers to the above mentioned questions.

2.1.2 Prediction of the status of the flood defence line (Workflow 1)

The first workflow extends the currently applied forecasting workflow (see Fig. 2), by an assessment of the performance of the riverine or coastal flood defence line in case of a high water event. This status is relevant to predict if a breach is likely to occur, resulting in an inundation of the protected area. The flood defence line is divided into sections to conduct the assessment. The concept of a fragility curve is used to summarize all geotechnical and hydraulic aspects of each section in one performance criteria to provide a fast accessible and comprehensible format (Bachmann et al., 2013).

The fragility curve $Frc(h_w)$ expresses the reliability of a structure as a function of a defined dominant stress variable (Hall et al., 2003). In this context, the water level at the structure is defined as the dominant stress variable. The curve shows the conditional probability of the occurrence of a failure event $P(failure|h_w)$ [-] on the vertical axis as a function of the water level h_w [m], represented on the horizontal axis (see Fig. 3).

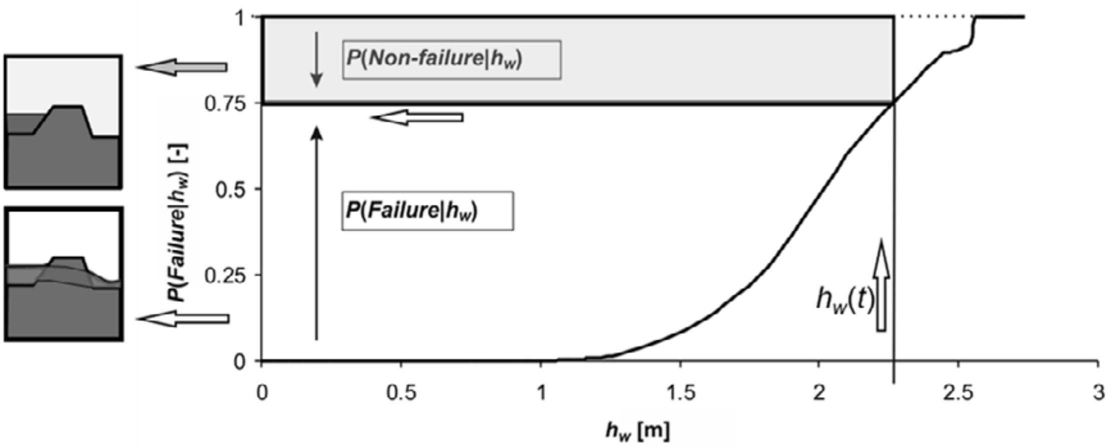

Figure 3: Concept of the fragility curve and the transformation from water level h_w to probability of failure $P(failure|h_w)$

Within workflow 1 the determination of the fragility curves is done in advance (see Fig. 4) assuming that the relevant input parameters are constant during a flood event. This ensures fast access to the information which, if suitable, could be updated during operational application using measurements (e.g. by sensors in the dike) or observations (e.g. by dike inspectors) during the event. This update may be necessary as emergency measures may have a direct influence on the dike performance (e.g. sand bags on the dike crest).

Using the fragility curve for each section, the water level forecast is transformed into a failure probability forecast over time per dike section (see Fig. 4). This transformation is performed for each water level h_w (see Fig. 3) over the time horizon of the water level forecast t.

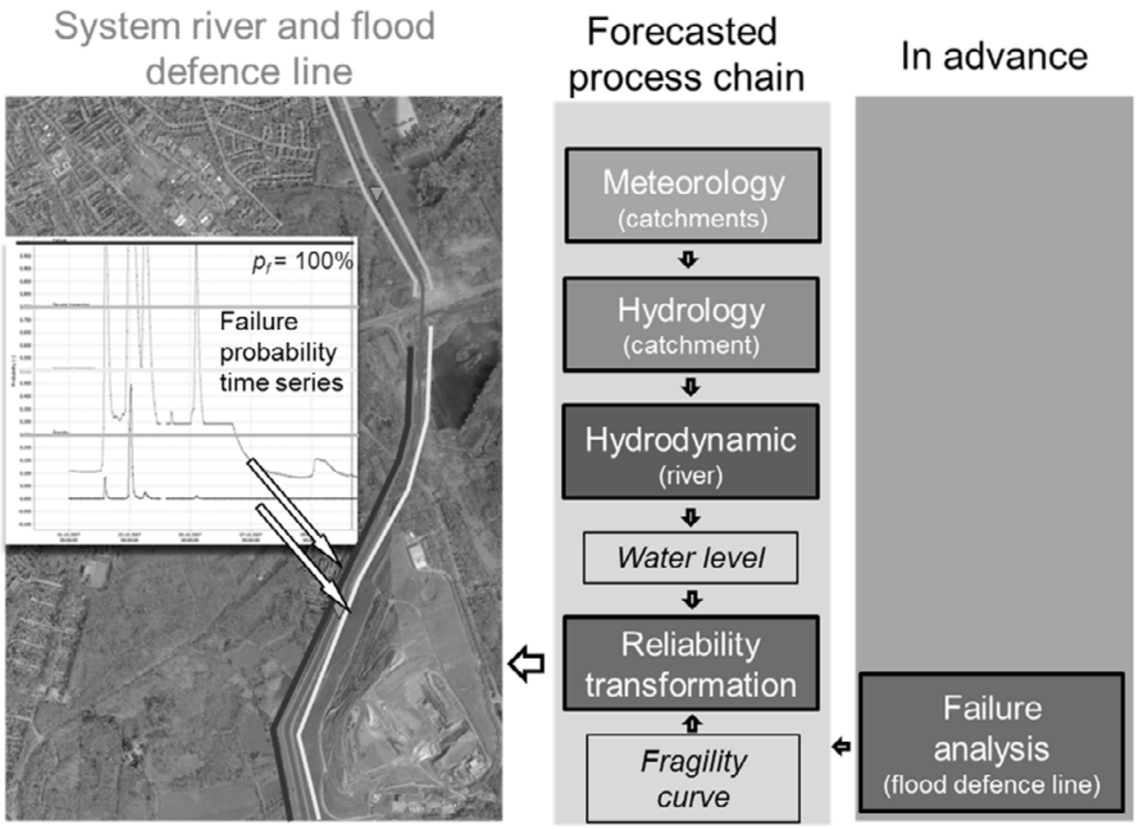

Figure 4: **Integration of probabilistic failure analysis into the process chain by workflow 1 to represent the performance of the flood defence line**

2.1.3 Prediction of impacts (Workflow 2)

The second workflow generates model-based forecasts about flood impacts. Flood impacts in the adjacent areas are expected by two possible events: a failure event including a breach development in the flood defence line or an overflow event over the flood defence line.

To assess the probable impact of flooding due to breach or an overflow (see Fig. 5) the process chain is further extended (see Fig. 5).

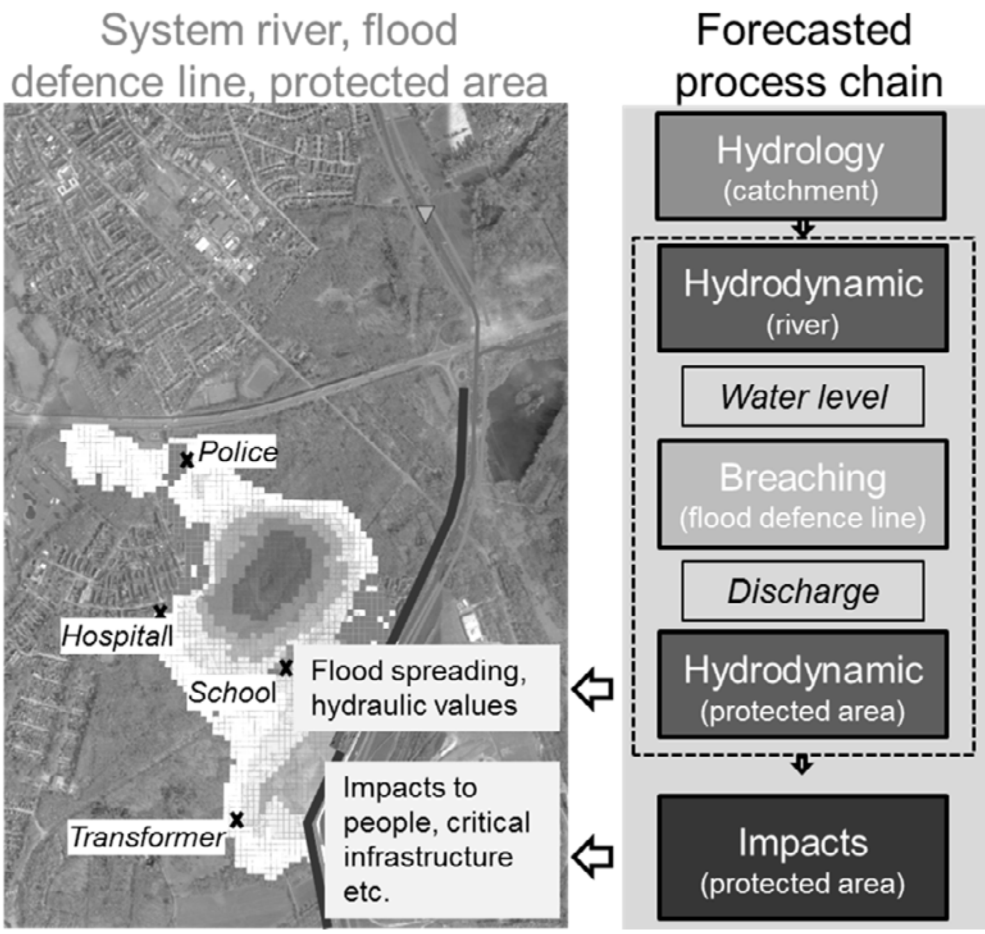

Figure 5: Integration of breaching, flood spreading and impact analysis into the process chain by workflow 2 to provide model-based forecasted information about the impacts of flooding

The objective of the hydrodynamic overland flow analysis is the transformation of the water level forecast into inundation characteristics such as water depths, flow velocities and associated travel times, taking into account the morphological characteristics of the river and protected area as well as the status of the flood defence line (failure or non-failure).

The objective of the impact analysis is the conversion of the hydraulic variables of a flood, like water depths, flow velocities and associated travel times, into consequences for people, assets and goods located in the affected areas. In an operational application a qualitative assessment of impacts via visualization is often adequate. The qualitative assessment of impacts is based on intercepting the flood spreading with spatial data about impact potential in the protected area. These spatial data provides information about densely populated areas, land use and critical infra-structure, such as fire or police departments, schools and hospitals, power supply units, cultural monuments or waste disposal areas.

2.2 Risk-based forecasted information for emergency response

2.2.1 General

The first prototype applications in the Netherlands (Rotterdam area) and Italy (Po river) have been conducted to test the extended forecasting workflows and to demonstrate the added value of the additional available information for emergency response organizations. The events are hypothetical, as the magnitude of a historical event has been adjusted.

Due to the flexibility and the wide-spread applications in operational forecasting domain, Delft-FEWS is chosen as a shell to manage the data and models required for the extended workflows. It provides a shell through in which an operational forecasting application can be developed specific to the requirements of an operational forecasting center (Werner et al. 2013).

First prototype applications in the Netherlands (Rotterdam area) and Italy (Po river) have been conducted to test the extended forecasting workflows and to demonstrate the added value of the additional available information for emergency response organisations.

Due to the flexibility and the wide-spread applications in operational forecasting Delft-FEWS is chosen as a shell, which manages the data and models required for the extended workflows. It provides a shell through in which an operational forecasting application can be developed specific to the requirements of an operational forecasting centre (Werner et al. 2013).

The following sections summarize the risk based forecasted information as a result of the extended workflows and the added value for the emergency response. More information about the prototypes is provided in Bachmann et al. (2016).

2.2.2 Prediction of the status of the flood defence line (Workflow 1)

Figure 6 shows an example for the prediction of the status of the flood defence line. In the upper part of Figure 6 the modelled water level time series for the 15 dike sections depending to the forecasted boundary conditions are shown. The lower part of Figure 6 presents the corresponding failure probability time series as a result of a transformation of the water levels using the fragility curves. The peaks of the water levels correspond to the peaks in the failure probabilities. By exceeding user-defined thresholds specific measures can be triggered. Dike inspectors can be warned with an extended lead time and deployed to the endangered sections. Emergency measures (e.g. sandbags) can be prepared.

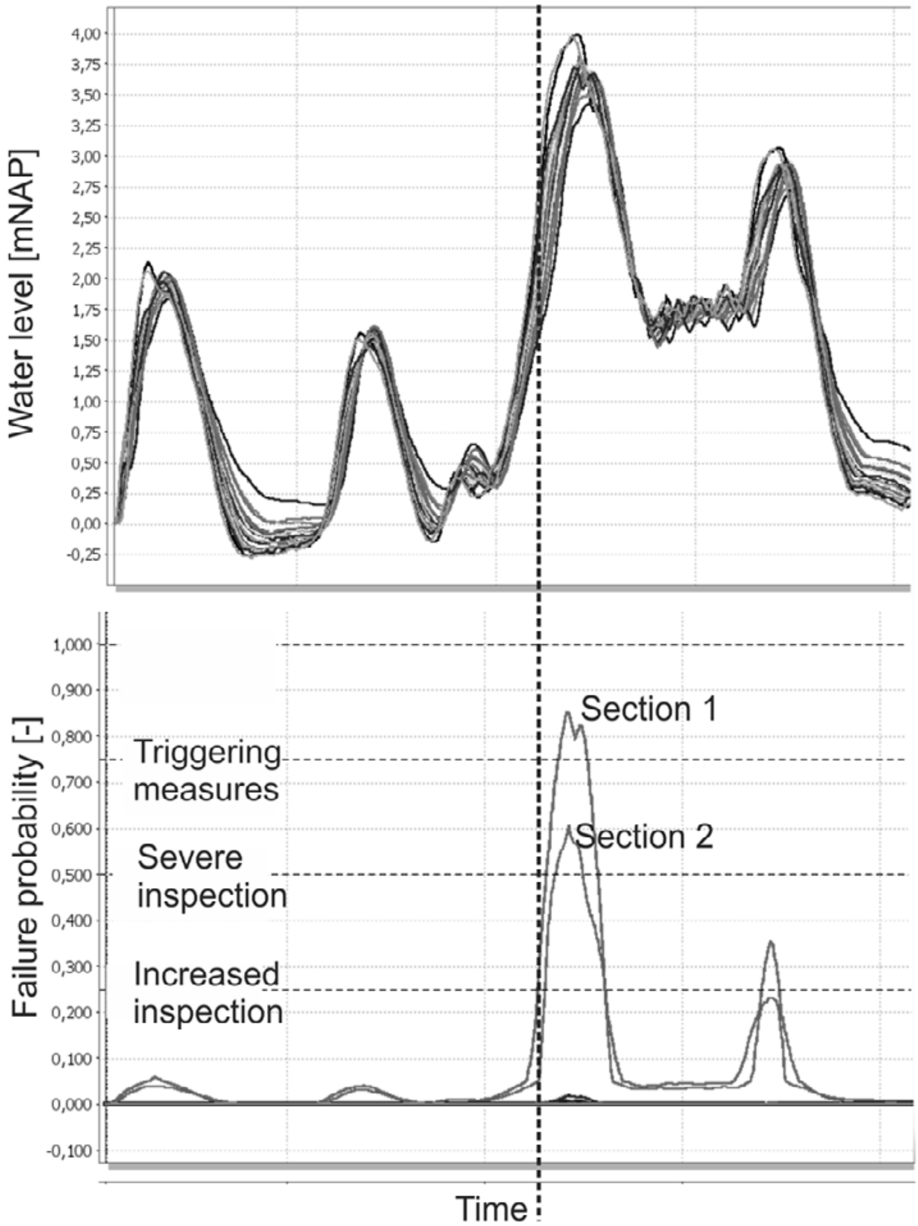

Figure 4: **Simulated water level time series (upper part), corresponding failure probability time series (lower part) and threshold level (Delft-FEWS display) for 15 dike sections**

Figure 6 refers to one set of predicted future conditions, e.g. predicted by a hydrological model (see Fig. 2). If a hydrological ensemble forecast is used (multiple set of predicted boundary conditions gets available), the failure probability forecast series shows also multiple possible future conditions of the flood defence line. Figure 7 shows an ensemble forecast of 9 ensemble members for one dike section. Thus, uncertainties within the forecasting process, like imperfect initial and boundary conditions of the models and imperfect models are taken into account.

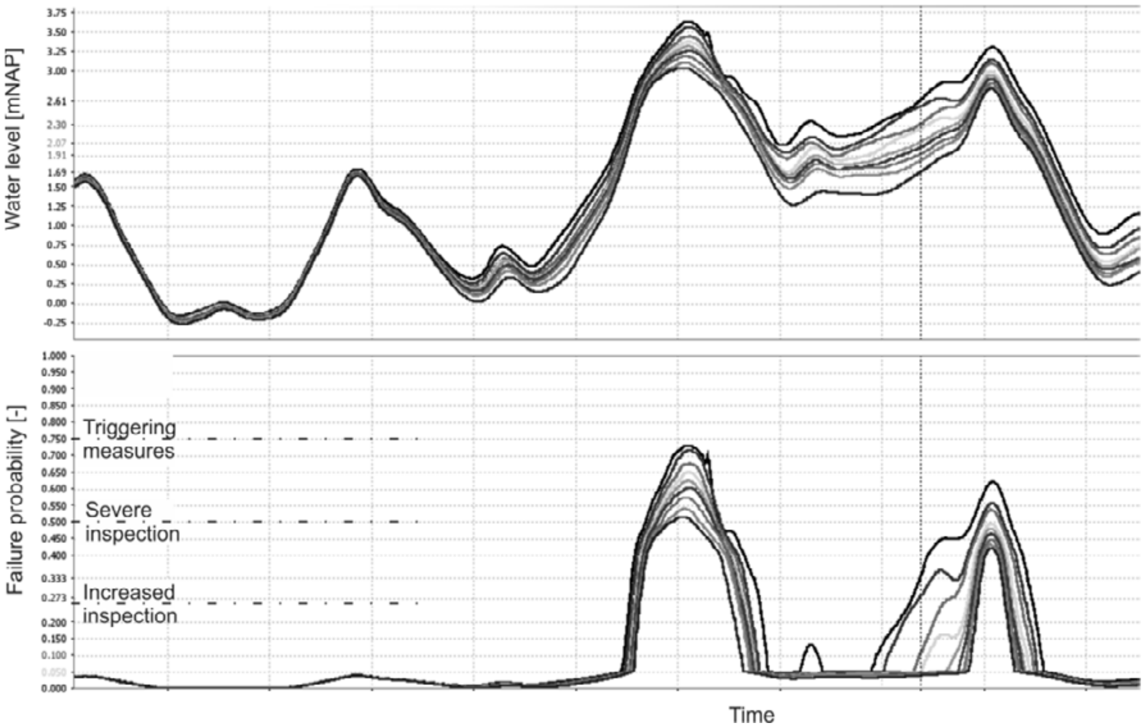

Figure 7: Simulated water level time series (upper part), corresponding failure probability time series (lower part) and threshold level (Delft-FEWS display) for 9 forecasted ensembles and one dike section

2.2.3 Prediction of impacts (Workflow 2)

The predicted time dependent flood spreading caused by a failure in two dike sections is shown in Figure 8. It shows two different points in time. The first one shows the flood spreading about two hours after failure of the dike (upper part), the second the flooded area about 9 hours after failure (lower part). Thus, information for the emergency response about the flood spreading, such as time and distribution of flooding, flow velocities and water depth are provided.

Figure 8 also demonstrates the added value for the emergency response by overlaying the flood spreading with information about the population density and critical infrastructure within a qualitative impact analysis. The dark grey shaded areas in Figure 8 indicate densely populated residential zones, the labels represent critical infrastructure, e.g. schools, hospitals, fire stations etc.

Predicted model-based information about the time when and how a specific area is flooded and about the impacts in this area become available during an emergency event. With the help of this information emergency measures within the protected area (e.g. horizontal or vertical evacuation of people and assets, temporally object protection) can be operationally planned, adapted and triggered.

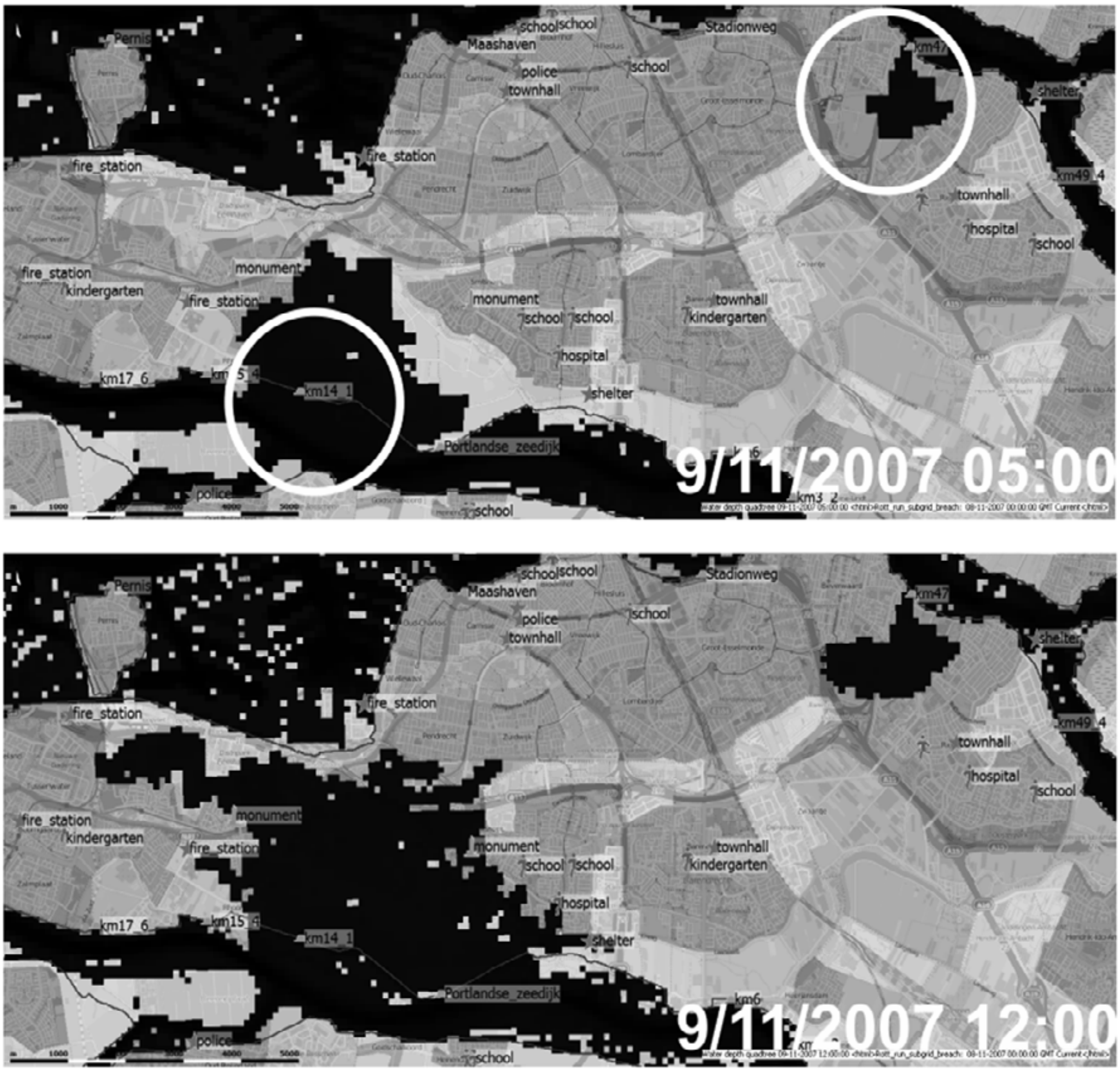

Figure 8: **Overview of flood spreading and qualitative impact analysis for dike failure in two dike sections (Delft-FEWS display)**

3 Social media for disaster response

People affected by natural hazards increasingly share their observations and needs through digital social media. It is recognized that hazards leave a footprint on social media (Guan and Chen, 2014). Twitter data has been used to gain information about the social impact (Lu and Brelsford, 2014), the temporal progression (Schnebele et al., 2014), the relative severity (Guan and Chen, 2014) of a hazard and extent of the hazard (Fohringer et al., 2015; Smith et al., 2015). All this is valuable information in the response phase (for water managers, disaster managers and insurance companies) using the perspective on-the-ground. But at present it is not analysed nor used in a structured manner for response.

In this paper we present results of a case study to use Twitter data for disaster response in the city of Jakarta, Indonesia. Jakarta is a large metropolis, which suffers yearly from recurring floods throughout the whole city and is therefore a very suitable case. The concept exploits observed information about the physical characteristics of hazards, such as flood depth and the location of flooding. These observations are used in combination with a Digital Elevation Model (DEM) to derive flood extent in real-time. As mentioned, the results could be used for disaster management and crisis relief organizations in the response phase (see Fig. 1). The same data also provides valuable information for model calibration during the mitigation phase. The generated maps are an addition to sources such as satellite images, areal images and post-flooding flood marks.

3.1 Method

Tweets were collected based on the mention of keywords, e.g. "*banjir*" the Bahasa Indonesia word for "flood". These tweets are then filtered for (1) any reference to a location based on an OSM based gazetteer and (2) a reference of water depth. The main geographical entity in Jakarta is the neighbourhood, also called "*keluharan*". This entity is often used by the inhabitants of Jakarta to refer to areas of interest, and is therefore found often in tweets on floods. In order to geolocate the observations in this dataset, we built a gazetteer with names and location of the 267 *kelurahans* in Jakarta and a large number of points of interest based on Open Street Maps. Any tweet that could be located in this way was given the respective *kelurahan* as geolocation attribute. This way, a database of flood observations per *kelurahan* was generated. The full process is shown in Figure 9. Note that we did not use georeferences of the (mobile) device that was used to send the tweet. The availability of this information is limited. Moreover, the location of a flood does not necessarily correspond to the location of the telephone. There are many known cases of people reporting floods from other areas (Jongman et al., 2015).

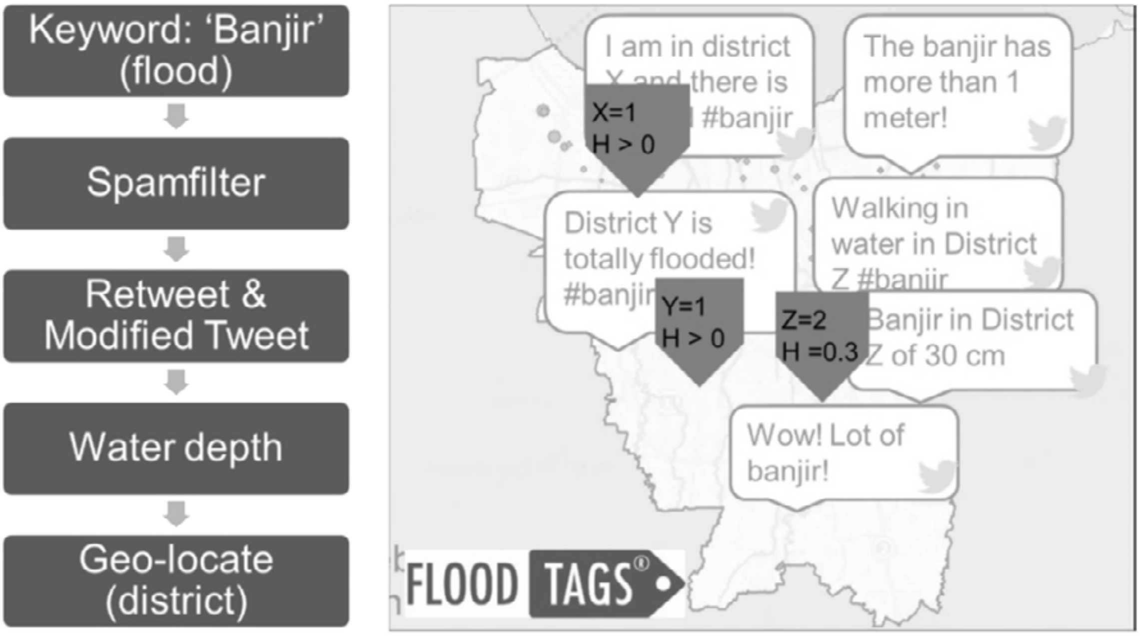

Figure 5: Flow diagram data mining flood observations from twitter; the figure right illustrates the workflow where Tweets for districts X,Y & Z are selected and water depth H is derived from the text

Observations by citizens are not made by validated instruments or reliable observers. Therefore, a single observation has to be considered as being unreliable. But we can assume that a large number of observations converge to the "truth". Here, this concept, also known as "wisdom of the crowd" (Surowiecki and Silverman, 2007) was implemented using the following assumptions. Firstly, a flood map is generated for each individual observation using a flood fill algorithm. Secondly, the reliability of each observation within a given time window is assessed. Finally the flood maps within the given time window are combined based on the reliability of each observation. These steps are discussed in the next paragraph. The complete workflow is shown in Figure 10.

The reliability of a single observation was based on the combined likelihood of a flood in a *kelurahan* and the likelihood of the observed maximum water depth in that *kelurahan* within a timeframe. The likelihood of a flood in a *kelurahan* was defined as the number of tweets in that *kelurahan* relative to the largest number of observations at any other *kelurahan* within the same timeframe. The median observed maximum water depth within a timeframe was taken as the most likely maximum water depth and given a value of one. It was assumed that the likelihood decreases linearly to zero at three standard deviations difference from the median. Based on these assumptions the median observed maximum water depth in the *kelurahan* with the largest number of observations is most reliable and has a likelihood of one. However, all observations are taken into account. These assumptions are only valid when there are sufficient flood observations within a time window.

For each observation a flood extend map was calculated and its likelihood value assigned. The flood extent was calculated by plotting the observed water depth on a Digital Elevation Map using a flood fill algorithm. The algorithm assumes a horizontal water level based on the observed water depth relative to the closest local depressions (i.e. lower lying areas in the terrain). The individual flood extent maps over a given period were combined. The result was a probabilistic flood map based over the analysed time window.

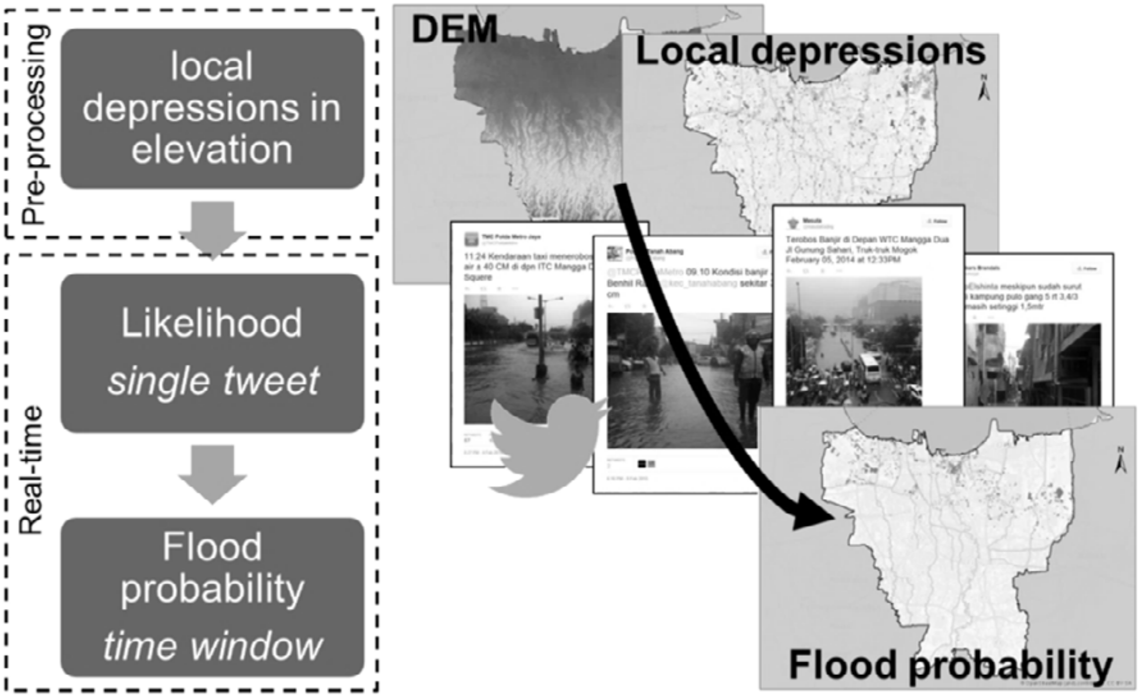

Figure 10: Workflow flood maps based on local depressions in elevation and flood observations from Twitter

3.2 Results case study Jakarta

On 9-11 February 2015, the city flooded due to heavy rainfall. During this event almost 728,000 Twitter messages related to the flood were recorded in the city, peaking at 893 tweets/minute. Data mining showed that 2.200 unique tweets sent during the flood contained information on the local water depth. Out of these tweets, 40% contained textual location information, resulting in a total of 888 unique geo-located flood observations. All flood observations for February 10th are shown as total number per *kelurahan* in Figure 11, top map; the resulting flood likelihood map is shown in the bottom map.

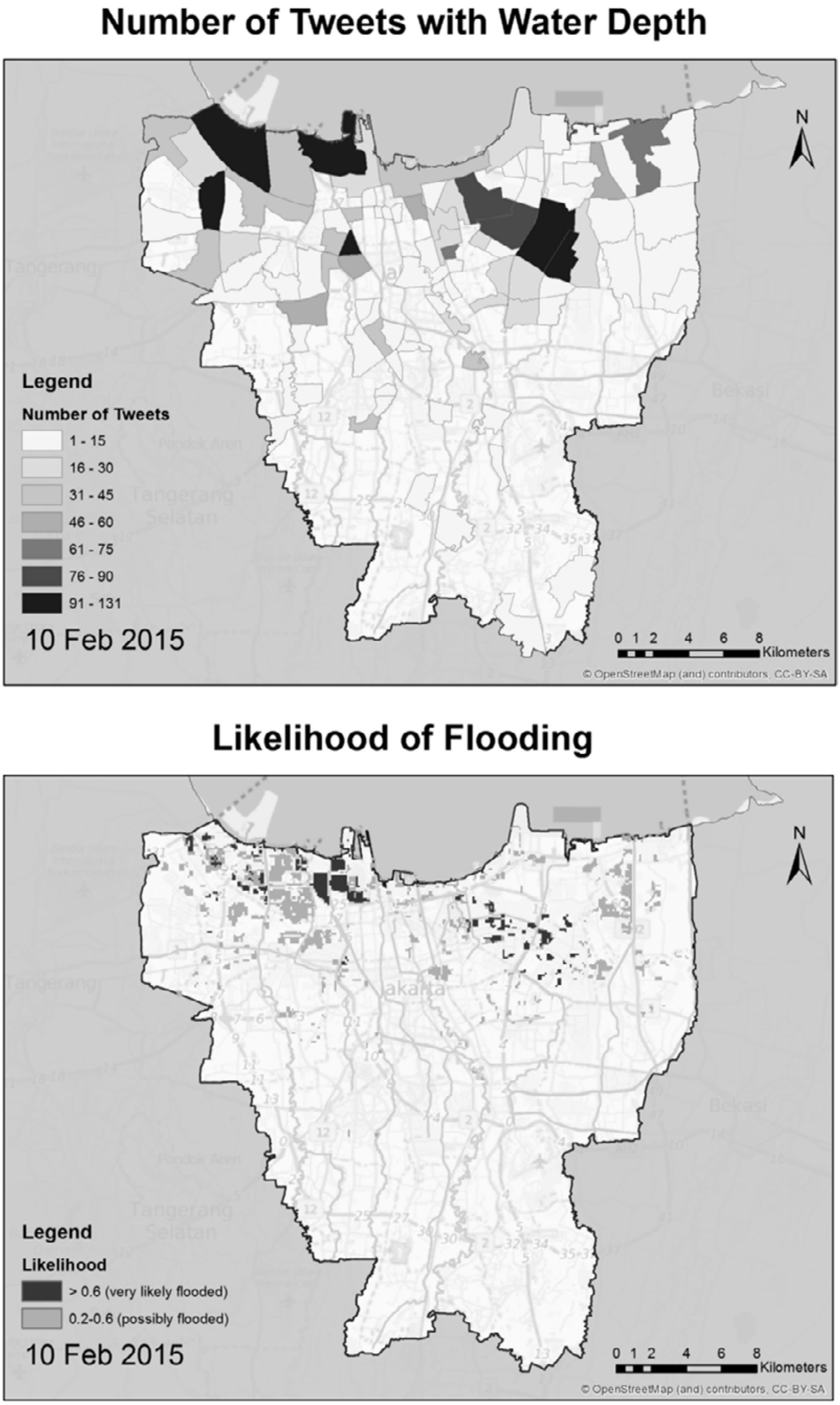

Figure 6: Total number of mined flood observations (top) and derived flood likelihood map (bottom)

The resulting flood probability maps were validated by cross-checking geo-located photos of the floods with the derived flood maps. Note that only true positives and false negatives can be detected with this method as the photo observations do not cover non-flooded areas. The photos were posted on Twitter, but neither the photo nor the accompanying message was used in the flood mapping algorithm. The confirmed flood photos were taken from Tweets that were filtered out at an earlier stage because a reference to a location was not automatically detected or the message did not contain flood depth observation. The photos were manually geo-located based on the message and the photo itself with an estimated accuracy of 500 m. In total 103 with confirmed flood photos were used for validation, providing a good spatial coverage of the flood area (see Fig. 12). The results of the validation are shown in Figure 12 and Table 1.

For 69% of the validation points (true positives), flooding was modelled within a range of 500 m (the accuracy of the validation points); and for two-third of these location the modelled water depth matched the photos. At 31% of the validation points (false negatives) no flood was modelled. Most of these 'missed' flood locations are at locations outside local depressions as derived from the DEM. In some cases a local depression was missed because of inaccuracy in the DEM. However, in most cases the flood was not captured as the developed flood fill algorithm does not take into account hydrodynamic aspects of flooding, nor does it capture fluvial floods.

Of the validation points, 93% was located in a *kelurahan* where flood observations were detected automatically (true positives). Assuming that the confirmed flood photos give a complete spatial coverage of the flood area, flood observations from Twitter filtered on the mention of a flood depth provide a good coverage of the flood extent in Jakarta. Of the photographical evidence, 7% was located in *kelurahans* where, in contradiction to this evidence, no floods were detected. The absence of automatically detected flood observations could be due to inaccurate (misinterpreting of the geo-location) or incomplete gazetteer (the place names that are used in the twitter messages, possibly abbreviations, are missing).

Table 1: Validation matrix

Validation category	Tweets [no.]	Percentage
Simulated water depth matches observation	46	45%
Simulated water depth is underestimation	25	24%
No flooding simulated, while data in *kelurahan*	25	24%
No flooding simulated and no data in *kelurahan*	7	7%
total	103	100%

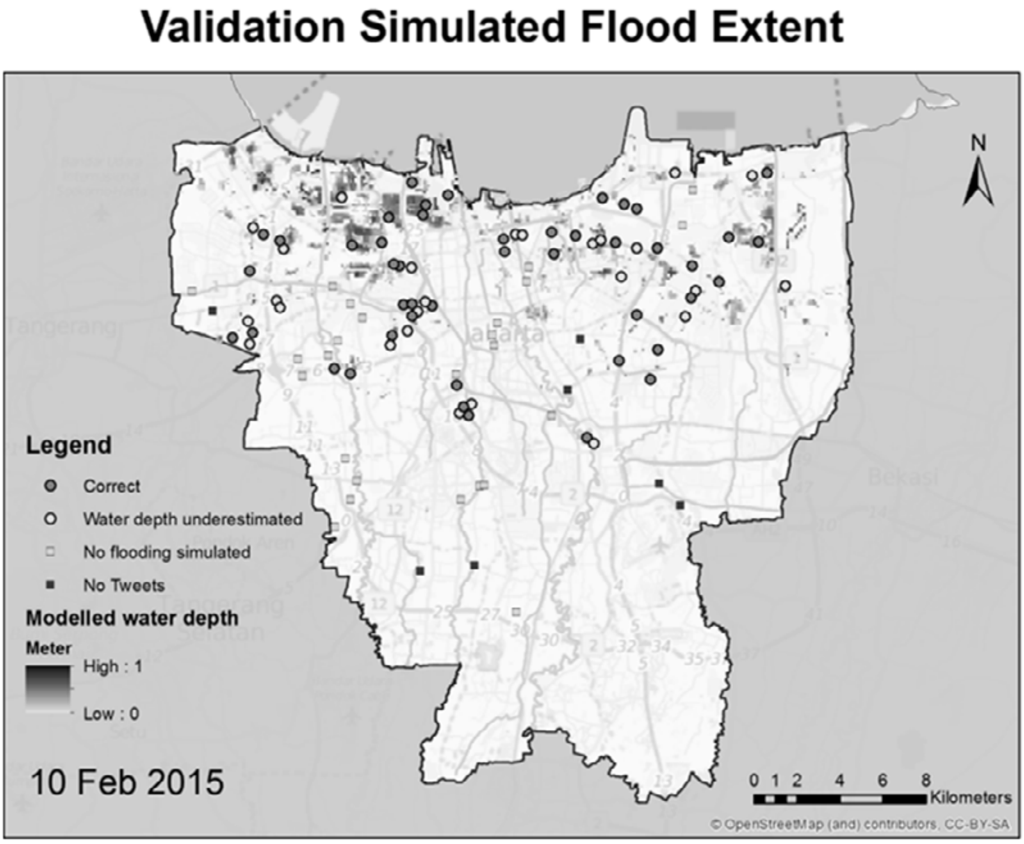

Figure 7: Validation of derived flood maps based on confirmed flood photos

The pilot area, Jakarta, proved to be a welcome test case as frequently struck by floods and the inhabitants tend to intensely use social media. However, it is not said that the developed method will be directly applicable in other flood prone areas as less people might be using social media, or posting location information, or sharing the flood depth that they are experiencing. At the same time, there is more information in the cloud than presently automatically harvested. Therefore, further research is required into using all the information present in social media.

4 Summary and Conclusion

An integrative disaster management approach consists of four different phases: mitigation, preparation, response and recovery. Adequate information for decision support during all phases is a key to strengthen resilience against flooding. The concepts of two innovative approaches to generate such forecasted and real-time information are presented and were successfully applied in two case studies.

The first approach consists of an extended flood forecasting system. Probabilistic failure analysis resulting in fragility curves is used to forecast the status of the flood defence line; two dimensional hydrodynamic simulations to forecast the extent of the hazard; and qualitative impact assessment to forecast impacts to people and critical infrastructures. With the help of this information emergency measures at the flood defence line (e.g. temporally dike strengthening) and within the protected area (e.g. evacuation of people) can be operationally planned, adapted and triggered.

The second presented approach shows that social media contain very useful information for flood disaster management. Exploiting physical characteristics of floods such as flood depth and location proved to be an added value for flood mapping. As data from social media becomes available in real-time and the applied method takes only seconds to compute, data processing is no limitation for applying this method in near-real time for disaster response. The availability of sufficient reliable data is however not guaranteed and may turn out to be a limitation, especially in places where social media are less frequently used.

Acknowledgement

The presented prototype applications of a risk-based forecasting (Rotterdam area and Po river) were part of the EU-funded FP7 RASOR project (Rapid Analysis and Spatialisation Of Risk). The dataset in the presented pilot study about the use of social media for disaster response was kindly provided by FloodTags.com.

5 References

Bachmann, D., Johann, G., Huber, N.P and Schüttrumpf, H. (2013): Fragility curves in operational dike reliability assessment. – In: Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 7:1, pp. 49-60.- ISSN 1749-9518.

Bachmann, D., Eilander, D., de Leeuw, A., de Bruijn, K., Diermanse, F., Weerts, A., Beckers, J. (2016): Prototypes of risk-based flood forecasting systems in the Netherlands and Italy. Submitted to FloodRisk Conference 2016, Lyon.

BMI (2013): Bericht zur Flutkatastrophe 2013 - Katastrophenhilfe, Entschädigung, Wiederaufbau (Bundesministerium des Innern).

Fohringer, J., Dransch, D., Kreibich, H., & Schröter, K. (2015). Social media as an information source for rapid flood inundation mapping. In: Natural Hazards and Earth System Sciences, 3, 4231-4264.

Guan, X., Chen, C. (2014) Using social media data to understand and assess disasters, In: Natural Hazards 74.2, 837-850.

Hall, J., Dawson, R., Sayers, P., Rosu, C., Chatterton, J. U. and Deakin, R. (2003): A methodology for national-scale flood risk assessment. In: Water & Maritime Engineering, Vol. 156, No. WM3, pp. 235-247. - ISSN 1472-4561.

Jongman, B., Wagemaker, J., Romero, B. R., & de Perez, E. C. (2015). Early flood detection for rapid humanitarian response: harnessing near real-time satellite and Twitter signals. In: ISPRS International Journal of Geo-Information, 4(4), 2246-2266.

Lu, X., Brelsford, C. (2014) Network Structure and Community Evolution on Twitter: Human Behavior Change in Response to the 2011 Japanese Earthquake and Tsunami, In: Scientific Reports 4, 6773

Munich RE (2015): GeoTopics - Naturkatastrophen 2014 Analysen, Bewertungen, Positionen.

Smith, L., Liang, Q., James, P., & Lin, W. (2015). Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. In: Journal of Flood Risk Management. doi:10.1111/jfr3.12154

Schnebele, E., Cervone, G., Waters, N. (2014) Road assessment after flood events using non-authoritative data, In: Natural Hazards and Earth System Science 14.4, 1007-1015

Surowiecki, J., & Silverman, M. P. (2007). The wisdom of crowds. In: American Journal of Physics, 75(2), 190-192.

Werner, M., Schellekens, J., Gijbers, P., van Dijk, M., van den Akker, O., Heynert, K. (2013): The Delft-FEWS flow forecasting system. In: Environmental Modelling & Software, Vol. 40, pp. 65-77. - ISSN 1364-8152.

Authors‘ addresses

Dirk Eilander
Deltares
Boussinesqweg 1
2629 HV Delft
Niederlande
Tel.:+31 88 3357672
Fax:+31 88 3358582
E-Mail:dirk.eilander@deltares.nl

Daniel Bachmann, Peter Gijsbers, Arnejan van Loenen
Deltares
Boussinesqweg 1
2629 HV Delft
Niederlande