Critical rainfall thresholds for urban pluvial flooding inferred from citizen observations

Tian, Xin; ten Veldhuis, Marie claire; Schleiss, Marc; Bouwens, Christian; van de Giesen, Nick

DOI
10.1016/j.scitotenv.2019.06.355

Publication date
2019

Document Version
Final published version

Published in
Science of the Total Environment

Science of the Total Environment 689 (2019) 258–268

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Critical rainfall thresholds for urban pluvial flooding inferred from citizen observations

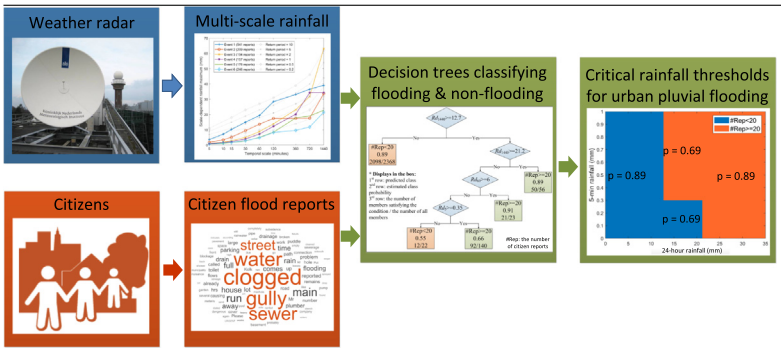
Xin Tian[a,*], Marie-claire ten Veldhuis[a], Marc Schleiss[b], Christian Bouwens[c], Nick van de Giesen[a]

[a]Department of Water Management, Delft University of Technology, Delft 2826 CD, The Netherlands
[b]Department of Geoscience and Remote Sensing, Delft University of Technology, Delft 2826 CD, The Netherlands
[c]Beheer Water en Riolen, Gemeente Rotterdam, Rotterdam 3072 AP, The Netherlands

HIGHLIGHTS

- Citizen observatories can provide valuable information for investigating urban pluvial flooding.
- A 10-year database composed of 70,000 citizen flood reports is used for analysis.
- Three decision tree learning models are built to predict flood occurrences.
- Dominating features are further identified based on a principal component analysis.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 20 May 2019
Received in revised form 22 June 2019
Accepted 22 June 2019
Available online 29 June 2019

Editor: Damia Barcelo

Keywords:
Citizen flood observations
Urban pluvial flooding
Critical rainfall thresholds
Decision tree learning
Weather radar

ABSTRACT

Urban pluvial flooding is one of the most costly natural hazards worldwide. Risks of flooding are expected to increase in the future due to global warming and urbanization. The complexity of the involved processes and the lack of long-term field observations means that many crucial aspects related to urban flood risks still remain poorly understood. In this paper, the possibility to gain new insight into urban pluvial flooding using citizen flood observations is explored. Using a ten-year dataset of radar rainfall maps and 70,000 citizen flood reports for the city of Rotterdam, we derive critical thresholds beyond which urban pluvial flooding is likely to occur. Three binary decision trees are trained for predicting flood occurrences based on peak rainfall intensities across different temporal scales. Results show that the decision trees correctly predict 37%–52% of all flood occurrences and 95%–97% of all non-flood occurrences, which is a fair performance given the uncertainties associated with citizen data. More importantly, all models agree on which rainfall features are the most important for predicting flooding, reaching optimal performance whenever short- and long-duration rainfall peak intensities are combined together to make a prediction. Additional feature selection using principal component analysis shows that further improvement is possible when critical rainfall thresholds are calculated using a linear combination of peak rainfall intensities across multiple temporal scales. The encouraging results suggest that citizen observatories, although prone to larger errors and uncertainties, constitute a valuable alternative source of information for gaining insight into urban pluvial flooding.

1. Introduction

Urban pluvial flooding occurs when stormwater produced by intense or long-lasting rainfall exceeds the capacity of urban

* Corresponding author.
E-mail address: x.tian@tudelft.nl (X. Tian).

https://doi.org/10.1016/j.scitotenv.2019.06.355
0048-9697/

drainage or sewerage systems (Kolsky, 1999; Butler and Davies, 2004; ten Veldhuis, 2010). It is one of the most common and costly natural hazards worldwide, posing a great threat to human safety and economic growth (Spekkers et al., 2015). The risks and impact of urban pluvial flooding are expected to increase with time as a result of rapid urbanization, aging infrastructure and climate change (Huong and Pathirana, 2013; Akter et al., 2018; Zhou et al., 2019), making the topic a very active field of research in urban water management (Rosenzweig et al., 2018).

For the purpose of flood risk management, one key issue remains the ability to predict the occurrence of flooding based on existing stormwater infrastructure and rainfall inputs (Douglas et al., 2010; Gaitan et al., 2015; Cherqui et al., 2015; Jiang et al., 2018). Accordingly, a variety of studies have focused on determining critical rainfall thresholds beyond which pluvial flooding is likely to occur (Martina et al., 2005). The thresholds are defined either in terms of rainfall intensities or total accumulations (Montesarchio et al., 2011). The inferred thresholds can be used to predict local flooding or design early warning systems (Martina et al., 2005). To determine the critical rainfall values, two main approaches have been proposed: (i) physically-based hydrological modeling (Norbiato et al., 2008; Yang et al., 2016) and (ii) statistical, data-driven analyses (Carpenter et al., 1999; Martina et al., 2005; Golian et al., 2010; Montesarchio et al., 2011). Both approaches require long-time records of spatially distributed high-quality hydrometeorological observations across various temporal and spatial scales. Initially, studies used rain gauges and simulated inundations to represent the rainfall and flooding data. However, rain gauges can be scattered very heterogeneously and sparsely across catchments, making it impossible to get a full picture of the spatial variability. This is a serious problem in urban areas where the hydrological response can be very sensitive to spatiotemporal rainfall variability (Schellart et al., 2012; Cristiano et al., 2017). Hydraulic inundation models also come with their own challenges, requiring a large number of field observations for calibration and validation, many of which are likely not to be available in practice (Eggimann et al., 2017).

In recent years, substantial improvement in terms of rainfall observations has been made thanks to the increased use of weather radar in hydrology. By merging radar and gauges, accurate and detailed characterization of the spatiotemporal variability of rainfall becomes possible. This has resulted in an improved understanding of flood responses (Thorndahl et al., 2017; Wright et al., 2014; Wang et al., 2015; Zhu et al., 2013). The gap in terms of high-resolution hydrological observations remains, however, a crucial limiting factor. To bridge this gap, various studies have been proposed to replace professional hydrological measurements by citizen observations (Smith et al., 2017; Overeem et al., 2016, 2013; Yang and Ling Ng, 2017; Mazzoleni et al., 2018; Weeser et al., 2018). Previous studies have shown that such an approach can potentially benefit hydraulic and hydrological modeling (Herman Assumpção et al., 2017; Starkey et al., 2017) and flood risk monitoring and management (Wang et al., 2018; Wehn et al., 2015; Liu et al., 2011), provided that the data are of sufficient quality. For a more detailed review of the potential of citizen observations applied to hydrology and water resources, see Buytaert et al., 2014, Muller et al., 2015, See, 2019 .

Although citizen observations can contribute to hydrological monitoring and modeling, they have not yet been explicitly employed to predict urban pluvial flooding, partly because long-term citizen datasets remain scarce. The present study aims to address this issue by exploring the potential of citizen flood observations for deriving critical rainfall thresholds and predicting the occurrence of urban pluvial flooding. The results are based on a 10-year database of citizen flood observations collected in Rotterdam, a pilot city in the Netherlands and very active in developing climate-adaptation and resilience strategies (Dunn et al., 2017). The dataset comprised about 70,000 citizen flood reports received via telephone, mobile app, email, and webpage. The reports are correlated to 5-min, 1×1 km spatial rainfall maps from the Dutch national C-band weather radar network (Overeem et al., 2009b). The data are used to train three different decision tree models for predicting the number of flood reports based on maximum rainfall intensities at different time scales, providing the basis for a simple early warning system for urban pluvial flooding.

The paper is organized as follows. Section 2 describes the materials, including the study area, radar-based rainfall observations, citizen flooding observations, and decision tree models. Section 3 presents the fitted decision trees and evaluates their performance. In Section 4, the advantages and limitations of the approach are discussed together with some ideas for future improvement. The conclusions are given in Section 5.

2. Materials and methods

2.1. Study area

The study focuses on the city of Rotterdam, which is located in a low-lying area in the delta of the Rhine and Meuse rivers in the Netherlands (see Fig. 1). The city counts more than 640,000 inhabitants spread over a heavily urbanized area of 257.6 km^2, with more than half of the total area being paved or semi-paved. Most areas of the city are drained by a combined sewer system (1800 km) whereas the remaining areas are drained by a separate sewer system (500 km) (Solomon, 2013). Historically, stormwater systems in Rotterdam have been designed to be able to cope with rainfall intensities of approximately 20 mm/h over short time windows of 10–15 min with an estimated in-sewer storage capacity of 10–12 mm (RIONED Foundation, 2017). However, the original infrastructure has aged considerably and some parts had to be renewed which means that the exact capacities are largely unknown. The city experiences a temperate oceanic and humid climate with prevailing westerly winds and regular precipitation throughout the year. The average annual precipitation in the area is about 850 mm (Royal Netherlands Meteorological Institute (KNMI), 2019) with 90% of the most intense rain events (at the 15–60 min time scales) occurring in summer, between June and September (Overeem et al., 2009b). Convection plays an important role in the development of heavy localized rain, with peak intensities predominantly occurring between 13:00 and 21:00 UTC (Overeem et al., 2009b).

2.2. Citizen observations of urban pluvial floods

In 2001, the municipality of Rotterdam established a citizen observatory database with the objective to improve the monitoring and management of urban sewerage and drainage systems (as well as other public services which are irrelevant and not included in this study). In the following years, a system was set up to collect reports by means of phone, email, mobile app, and webpage. Individual reports were stored in a database along with their date, location (geotagging) and a short description of the event. Each citizen report was attributed to one of seven different categories, five of which are related to flooding (see Fig. 2). In this study, we used a dataset of about 70,000 citizen reports collected between 2008 and 2017. The latter were subjected to basic quality control during which all incomplete reports (i.e., without a valid date, geotagging, category or description) and reports not directly pertaining to flooding were removed prior to analysis.

The fact that many reports were collected by human operators over the phone during working hours implies that weekends tend to be characterized by a lower number of reports compared to weekdays. Consequently, only the reports collected during the weekdays were considered for this study. The total number of reports after quality control and removal of weekends was approximately

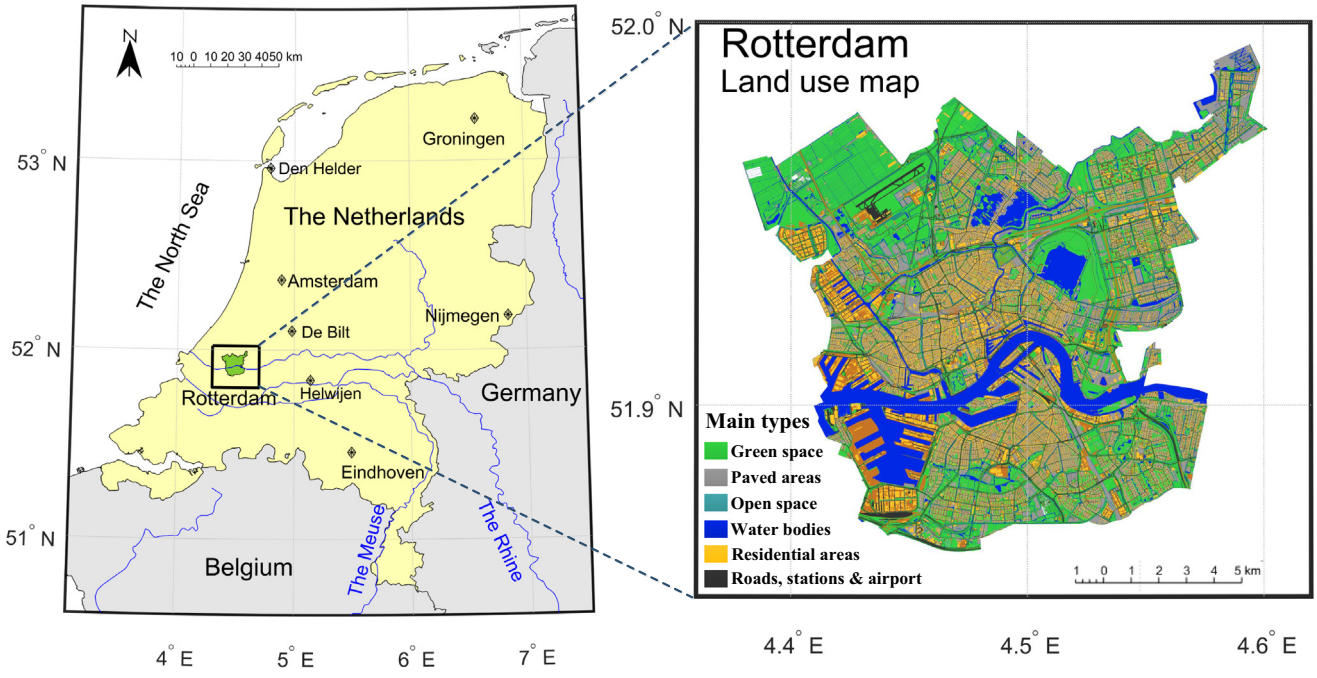


Fig. 1. The study area in Rotterdam, located in the delta of the Rhine and Meuse rivers. A large area of the city is highly urbanized and (semi-)paved.

38,300 over 2609 days. The large number of reports (i.e., ≈ 3830 per year) clearly highlights the extent to which Rotterdam suffers from water-related nuisances and how frequent these incidents are.

It is important to clarify that not all flood reports in the dataset are directly attributable to rainfall. Even on days without precipitation, citizens will call in to report flooding in their house or neighborhood associated with other causes not directly related to rain (e.g., high groundwater levels, bursts of water supply pipes). However, evidence shows that the average number of reports received during dry days tends to be much lower than on rainy days. To better distinguish between the two, a dry-weather baseline was established. Inspired by the work of Spekkers et al. (2013) , we derived the 97.5th percentile of the number of reports received on days with less than 0.1 mm of precipitation. This resulted in a baseline value of 20 reports/day on dry days, which is above the average number of reports per day in the database (≈ 14.7 reports/day). Using this baseline, 443 weekdays with urban pluvial flooding and 2166 days with non-pluvial flooding were identified.

2.3. Radar rainfall observations

Information about rainfall intensity on days with pluvial flooding was obtained from the Royal Netherlands Meteorological Institute (KNMI). Rainfall estimates were obtained from radar reflectivities

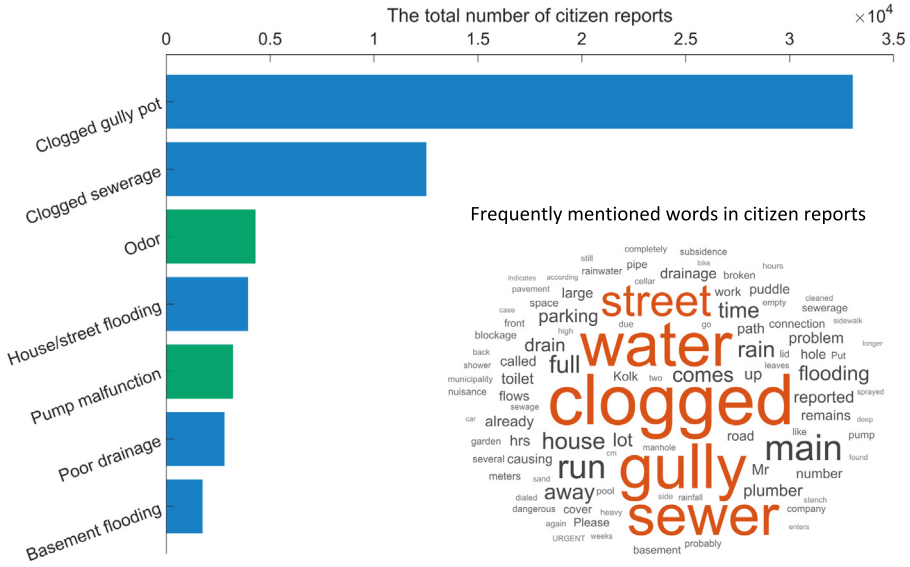


Fig. 2. Citizen reports about water nuisance, collected from 2008 to 2017 for the city of Rotterdam. Note that the categories “odor” and “pump malfunction” (in green) are not related to urban flooding and are not considered in this study. A word cloud of the remaining five categories (in blue) presents the most frequently mentioned words. Font size indicates the frequency of a given word.

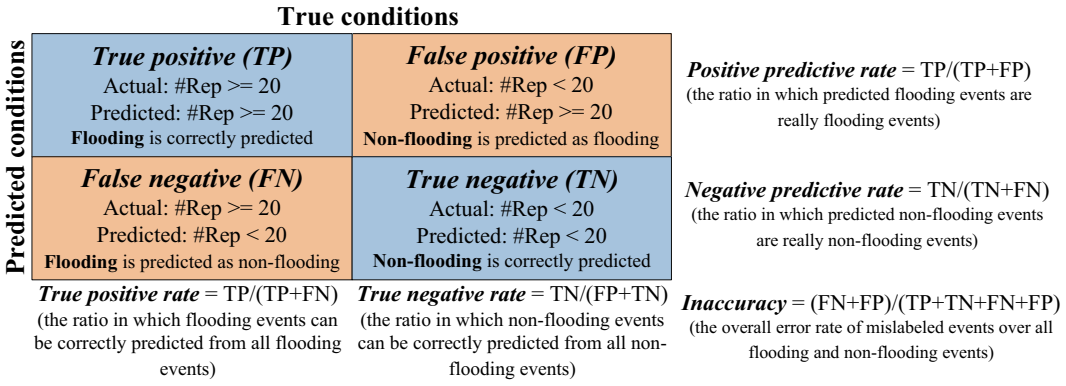


Fig. 3. Confusion matrix used to assess the performance of a decision tree. Only several indicators are included and others can be computed accordingly, e.g., false alarm/positive rate = 1 – positive predictive rate and miss rate = 1 – true positive rate.

measured by the C-band radars in De Bilt (52.103°N, 5.179°E) and Den Helder (52.955°N, 4.79°E) starting on January 1st of 2008. Note that the radar in De Bilt stopped contributing to the composite in the course of January 2017, at which point it was replaced by a new radar in Herwijnen (51.837°N, 5.138°E). All radar composites are adjusted and validated at hourly and daily time scales using the complete rain gauge network from KNMI (33 automatic stations and 326 manual gauges). Data are available for free in HDF5 format through KNMI's FTP server or in netCDF4 format via the Climate4Impact website. An extensive description and references can be found in Overeem et al. (2009a,b) .

For the purpose of this study, a ten-year time series of radar rainfall maps (from 2008 to 2017) over the study region at temporal and spatial resolutions of 5 min and 1 km, has been retrieved. This was used to compute areal-averaged rainfall intensities for each 5-minute time step of the entire municipal area (i.e., 257.6 km^2). To investigate the effect of temporal resolution, the data were aggregated from 5 min to eight coarser temporal scales (10, 15, 30, 60, 120, 360, 720 and 1440 min). Maximum rainfall accumulations Rd_x (in mm) were retrieved for all x-min aggregation time scales and used for further analysis:

$$Rd_x = \max_j \left\{ \underbrace{\sum_{k=1}^{x/5} R_5(k), \ldots, \sum_{k=1+j*x/5}^{(j+1)*x/5} R_5(k), \ldots, \sum_{k=(1441-x)/5}^{1440/5} R_5(k)}_{\text{Aggregation of 5-min rainfall to } x\text{-min rainfall}} \right\} \quad (1)$$

where $j = 0,1,\ldots,1440/x-1, x \in \{5, 10, 15, 30, 60, 120, 360, 720, 1440\}$, and R_5 is 5-min rainfall intensity.

2.4. Binary decision tree model

To represent the relationship between peak rainfall amounts (at different time scales) and the occurrence of pluvial flooding, a binary decision tree learning (DTL) model was selected (Breiman, 1984). The latter implements an easy to interpret tree-shaped hierarchical model to predict a binary outcome (yes/no) based on a set of features. In our case, "yes" means pluvial flooding (i.e., more than 20 reports per day) while "no" means non-pluvial flooding (i.e., fewer than 20 reports per day). The features used to make the decisions along the different branches of the tree are given by the peak rainfall amounts (expressed in mm) for each day, at nine different temporal aggregation scales.

The decision tree model is trained on the radar-rainfall and citizen report data and the model with the highest predictive performance is used to derive critical rainfall thresholds for urban pluvial flooding. The training part is done using the "fitctree" function in Matlab, starting at the root and moving to the leaves. At each node, all possible splits are examined and the one that minimizes the Gini impurity index of the two sub-populations is kept (Jost, 2006). The splitting stops when (i) a node only contains observations of a single class or (ii) any further split produces a child node with fewer than three observations. Performance is assessed by means of cross-validation, focusing on the following indicators:

- The confusion matrix: a 2×2 table in which each row represents a predicted outcome (i.e., more or less than 20 reports/day) while each column represents the actual observed outcome (Fawcett, 2006). This is used to count the number of true positives TP (i.e., correct flood warnings), true negatives TN, false positives FP and false negatives FN (i.e., missed flood events). From there, several performance indexes can be computed, as shown in Fig. 3. Note that a good flood warning system should have both low false negative rates and low false positive rates.
- The re-substitution error, which is the inaccuracy of the DTL model based on the training data (Schiavo and Hand, 2000). The inaccuracy is defined as the false alarms over all the alarms, i.e., (FN+FP)/(TP+TN+FN+FP). The re-substitution error rate measures how well the model performs on the training data. It tends to underestimate the true error rate that one would obtain by applying the tree to new unseen data, especially if the model is overfitted (i.e., tuned too tightly to the training dataset).
- The 10-fold cross-validation error, which is obtained by randomly partitioning the training dataset into ten equal-sized subsets and computing the mean value of the re-substitution error rates for all subsets. This provides a rough estimate of the error when applying the model to new data (Bengio Bengioy and Grandvalet Yvesgrandvalet, 2004). When the 10-fold cross-validation error is significantly larger than the re-substitution error, the tree is likely to be overfitted. Note that this approach is more robust than splitting the data into two independent training and testing datasets due to the low number of total flood days (443) in the dataset.

2.5. List of models

The number and variety of features used to train a decision tree model play an important role in determining its performance

(Guyon and Elisseeff, 2003). To investigate this aspect and get more insight into the most important rainfall features needed to predict urban pluvial flooding in Rotterdam, three different decision tree models were trained.

- Model 1: uses all nine available rainfall features at temporal aggregation scales from 5 min up to 24 h. During training, the algorithm automatically chooses the most relevant features and thresholds for classification.
- Model 2: is trained using only two rainfall features: Rd_5 and Rd_{1440} representative of the small-scale and large-scale peak intensities. Because of the lower number of features, this model is expected to perform slightly worse than the full model. However, since it only relies on two features for making its classification, it is operationally more straightforward to implement and also easier to interpret.
- Model 3: is trained using two alternative multiscale rainfall features determined by applying a principal component analysis (PCA) to the original 9 rainfall features (Jolliffe, 2002; Guyon and Elisseeff, 2003) and only keeping the two most meaningful components. The main purpose of this model is to assess whether linear combinations of peak rainfall intensities across different time scales are more useful predictors of urban pluvial flooding than Rd_5 and Rd_{1440} alone.

3. Results

3.1. Model 1: all nine rainfall features

The first trained decision tree model is shown in Fig. 4-(a). Although all nine rainfall features were available for training, only three of them (5 min peak, 60 min peak, and daily total) were needed to make the final classification, based on the algorithm described in Section 2.4. This was expected, as many of the rainfall features are highly redundant (i.e., correlated to each other) and only a few of them are needed to make the predictions. According to this tree, the first and most important rainfall threshold for distinguishing flood days is the total daily rainfall amount Rd_{1440}. Days with $Rd_{1440} < 12.7$ mm are labeled as non-flood days (i.e., fewer than 20 reports/days), resulting in a classification accuracy of 89% (2,98 correctly predicted days out of 2368). All days for which the daily rainfall accumulation was above 21.19 mm were labeled as flood days (likelihood of 89%), therefore providing a first straightforward way of predicting floods in cases where the only available data are daily rainfall accumulations. However, this is clearly not a good strategy as a unique daily rainfall threshold of 21.19 mm would result in a large number of flood events being missed. To address this issue, the tree considers two additional bifurcations depending on whether the hourly maxima Rd_{60} is larger than 6 mm and the 5-min rainfall amount Rd_5 larger than 0.35 mm (4.2 mm/h). These additional two thresholds increase the overall classification performance, although the last bifurcation based on the 5 min peak intensity does not seem to improve the results by much, resulting in rather large overall misclassification rates of 45% and 34% respectively. By contrast, the 6 mm threshold at the hourly scale clearly helps the classification by separating long-lasting low-intensity events with large accumulations from those with higher peak intensity and low to modest daily accumulations (e.g., localized thunderstorms).

Fig. 5-(a) shows the confusion matrix associated with the first decision tree. Among the 443 flood events in the database, only 163 were successfully predicted by the model, resulting in a high miss rate of 63%. On the other hand, the non-flood days were predicted very well (2110 out of 2166 or 97%). Note that the predictive rates for both cases are relatively good: among the 219 predicted flood events, 163 were actual flood days (positive predictive rate of 74%) and among the 2380 predicted non-flood days, 2110(88%) were correct. The inaccuracy, which is the overall error rate for both flood and non-flood days, is 13%. The 10-fold cross-validation error rate is 13.8% which indicates that the tree is probably not overfitted.

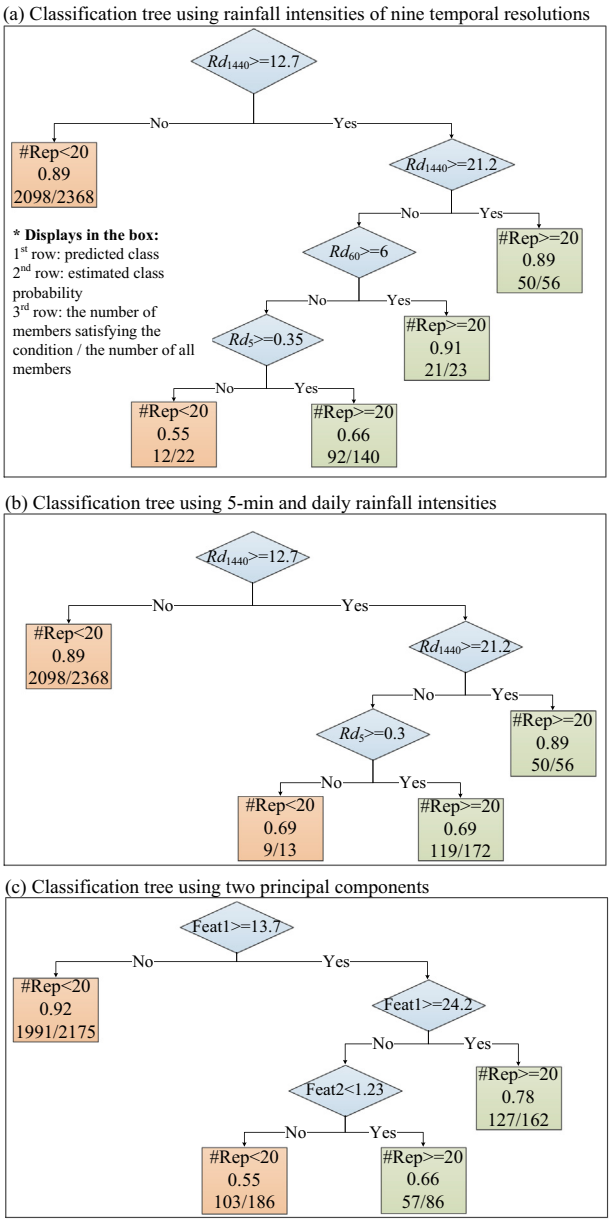
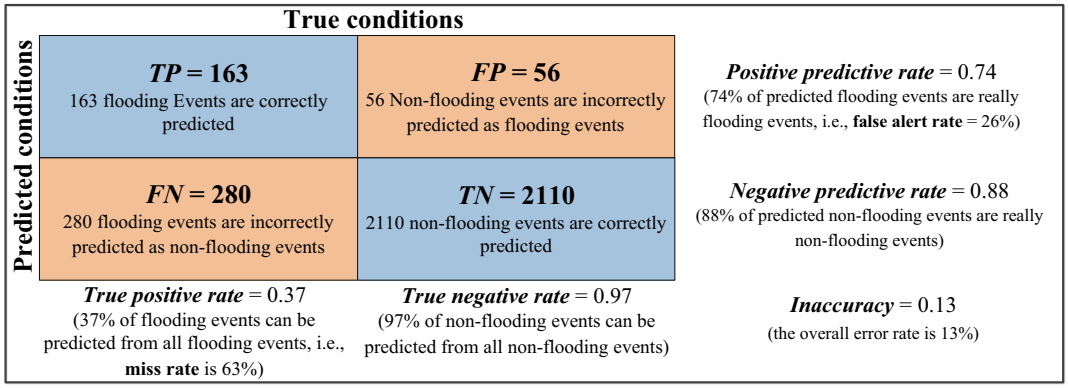


Fig. 4. Binary decision trees trained by using (a) rainfall intensities at nine selected temporal resolutions; (b) Maximum 5-minute (Rd_5) and 24-hour (Rd_{1440}) rainfall depths, and (c) two principal components by conducting a PCA. The decision tree predicts flooding and non-flooding by checking whether the number of reports exceeds the dry-days baseline of 20 reports per day. Note that the symbol '#Rep' stands for the number of reports.

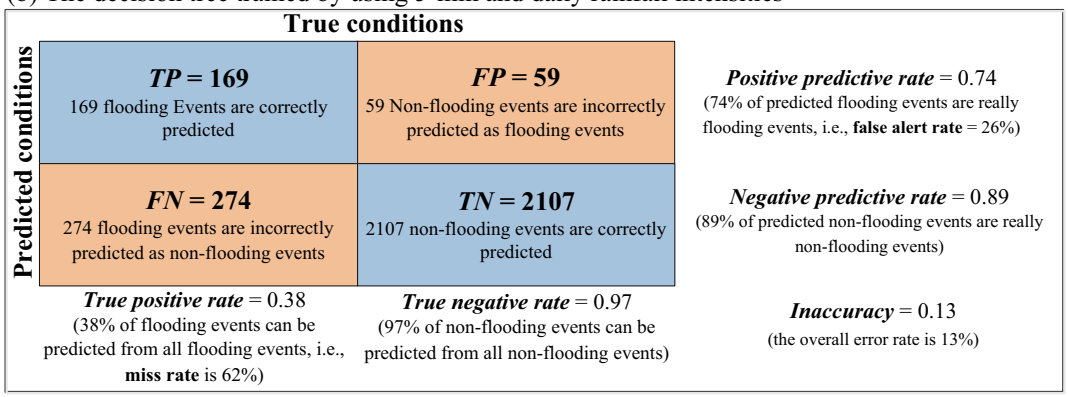
3.2. Model 2: 5-min peak intensity and daily accumulation

The second fitted decision tree model is shown in Fig. 4-(b). It was trained using only two of the nine available rainfall features: Rd_5 and Rd_{1440}. The first two bifurcations in the tree are identical to model 1 (Fig. 4-(a)). The main difference concerns the third bifurcation on $Rd_5 \geq 0.3$ which replaces the previous one based on hourly

(a) The decision tree trained by using rainfall intensities of nine temporal resolutions



(b) The decision tree trained by using 5-min and daily rainfall intensities



(c) The decision tree trained by using two principal components

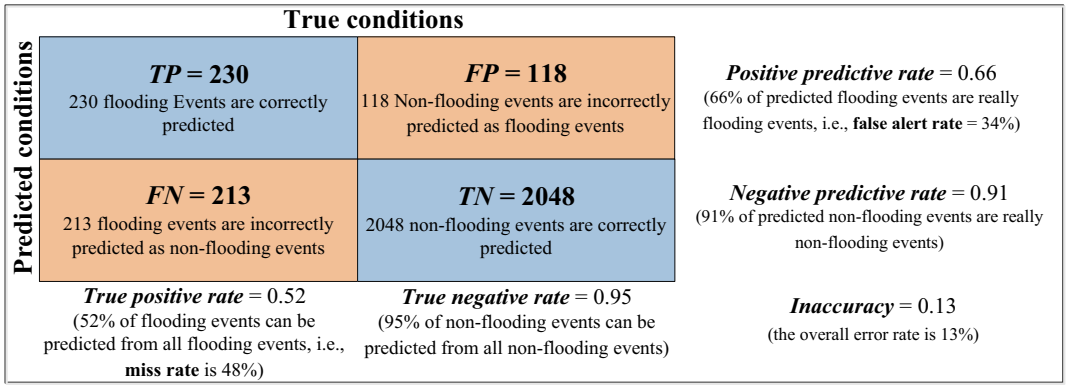


Fig. 5. The confusion matrix of the trained decision trees based on (a) rainfall intensities at nine selected temporal resolution; (b) Maximum 5-minute (Rd_5) and 24-hour (Rd_{1440}) rainfall depths, and (c) two principal components by conducting a PCA.

peak intensity. In model 2, this branch is replaced by two nodes with identical likelihoods of 0.69. The confusion matrix associated with this new tree is shown in Fig. 5-(b). Among the 443 flood events, 169 events (+6 with respect to model 1) were correctly predicted, resulting in a slightly lower miss rate of 62%, compared to model 1. The true negative rate remains almost the same, with 2107 correctly predicted days among 2166(−3 compared to model 1). The overall inaccuracy is 13% and the 10-fold cross-validation error of 14% suggests that the tree is probably not overfitted. Since the overall performance remains almost unchanged between model 1 and model 2, this suggests that there is little added value in including more than 2 temporal aggregation scales in the rainfall data when making the predictions. In other words, additional information about rainfall intensities at intermediate time scales does not significantly help the detection of flood events.

A two-dimensional diagram with the decision regions corresponding to model 2 is shown in Fig. 6-(a). The two colors are used to indicate the labels attributed to the observations as a function of Rd_5 and Rd_{1440}. Based on this diagram, a system operator or manager can rapidly estimate the likelihood of urban pluvial flooding based on the 5-min peak rainfall intensity and daily accumulations during the past 24 h. Alternatively, rainfall forecasts for the next 24 h can be used to help identify potentially dangerous situations and issue an early warning.

3.3. Model 3: features based on principal component analysis

The third and last model is trained using two new features obtained by principal component analysis (PCA). The new features are linear combinations of the nine daily peak rainfall intensities at

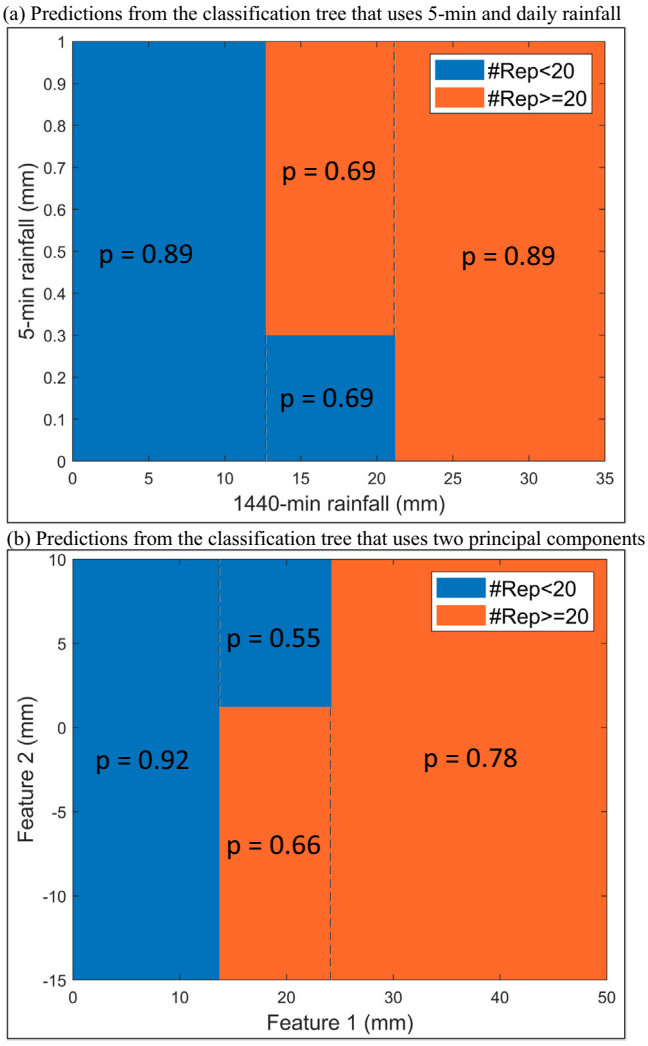


Fig. 6. Predictions made by two decision tree models based on their respective critical rainfall thresholds: (a) Rd_5 and Rd_{1440}, and (b) two principal components following from a principal component analysis. The letter p stands for probability of a correct model prediction.

5, 10, 15, 30, 60, 120, 360, 720, and 1440-min temporal resolutions with coefficients given in Table 1. The first feature corresponds to the first principal component, explaining 95% of the total variance while the second explains 3.1%. Note that the total variance in the case of PCA is the trace of the covariance matrix, i.e., the sum of eigenvalues, and the explanation rate is a ratio of a given eigenvalue to the trace. In other words, eigenvalues corresponding to the first two principal components take up 95% and 3.1% of the sum of all eigenvalues. Since they are uncorrelated, the two features taken together explain 98.1% of the total variance in the rainfall data. The interpretation of the new features is rather straightforward: feature 1 is a weighted sum of rainfall amounts at different time scales, with larger scales receiving more weight, therefore favoring total accumulation and persistence over peak intensity (e.g., large frontal systems). The second component puts more weight on scales between 60 and 120 min with total daily amounts only playing a minor role. It therefore specifically targets localized convective storms (e.g., thunderstorms) with high peak intensities but low to medium daily accumulations. The latter are less common than large frontal systems but obviously very relevant for local urban flooding, which has been reported in the literature (Daniels et al., 2014; Overeem et al., 2009a,b). Additional analyses (not shown) were also carried out to investigate the values of two features for every month, indicating that large feature 2 occurs exclusively in the summer (June–September) and is likely to be associated with summer thunderstorms.

The fitted decision tree with the two new rainfall features is shown in Fig. 4-(c). Its structure is identical to model 2, with feature 1 (weighted accumulation) taking over the role of Rd_{1440} and feature 2 the role of Rd_5. The main difference concerns the values of the critical thresholds used to make the decisions which are obtained by combining rainfall data at different scales and therefore harder to interpret physically. The decision regions of the tree as a function of the feature values are shown in Fig. 6-(b) and the confusion matrix is provided in Fig. 5-(c). Of the 443 actual flood days, 230 events were correctly predicted (52%). This is significantly larger than for the first two models. On the downside, the number of true negatives has decreased (2048 out of 2166) and the positive predictive rate has dropped to 66% (i.e., the model has a false alarm rate of 34%). Overall, the inaccuracy is 13% and 10-fold cross-validation error is 14%, which are comparable to the first two models. The comparison between model 2 and 3 suggests that a linear combination of rainfall intensities at different scales can improve the prediction. However, the gain in performance is only marginal and mostly concerns the prediction of non-flood events. When it comes to positive predictive rates, all models seem to suffer from large false alarm rates in the order of 26–34 %. Possible explanations for why the models fail to correctly predict the many flood events are discussed in the next section.

4. Discussion

4.1. Interpretation of critical rainfall thresholds for urban pluvial flood prediction

The trained decision trees show that a fair distinction between flood and non-flood days can be made using either a combination of 5-min peak rainfall intensity and daily accumulation (model 2) or a linear combination of rainfall intensities at nine different temporal scales (model 3). Model 2 is simpler and easier to interpret in terms of physical quantities but model 3 has a slightly better balance in terms of false positive and negative rates. In both cases, the combination of short-duration rainfall peaks at hourly or sub-hourly scales together with information about the total daily accumulations was necessary to achieve reasonable performance. The derived threshold values of 0.3 mm/5 min(equivalent to 3.6 mm/h) and 12.7 mm/day are consistent with the guidelines followed by the city of Rotterdam in the 1950–1980s (RIONED Foundation, 2017), when most of the currently existing stormwater and drainage systems were built. At that time, the system was designed for approximately 5 mm of rain over 15 min(i.e., 20 mm/h) and 10–12 mm of in-system storage capacity associated with longer time durations (e.g., 24 h). The short-duration value corresponded to a rainfall return period of approximately 1 year in the 1950–1980s ' standard (RIONED Foundation, 2017). The threshold of 12.7 mm daily accumulation in model 1 or 2, below which no flooding is expected to occur, is likely to be

Table 1
The coefficients for the two first principal components as a function of the maximum rainfall amounts at nine different scales.

Coefficients	Rd_5	Rd_{10}	Rd_{15}	Rd_{30}	Rd_{60}	Rd_{120}	Rd_{360}	Rd_{720}	Rd_{1440}	Explained
Feature 1	0.03	0.05	0.08	0.12	0.18	0.27	0.42	0.53	0.65	95%
Feature 2	0.09	0.17	0.22	0.33	0.42	0.45	0.30	−0.03	−0.58	3.1%

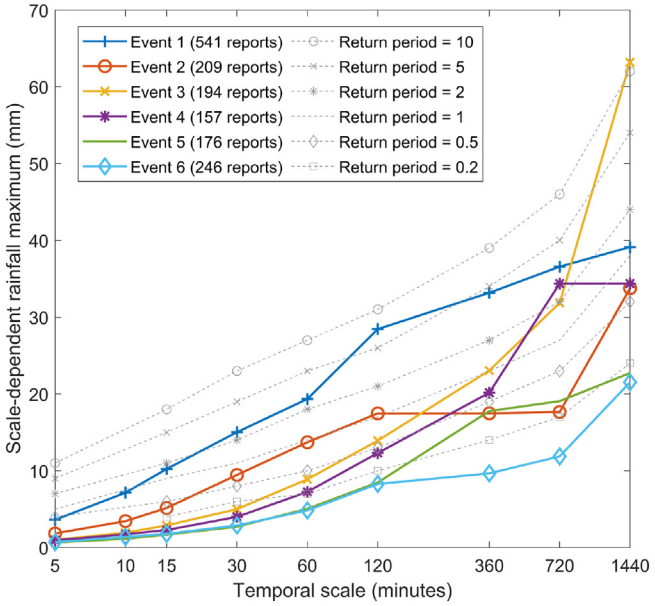


Fig. 7. Scale-dependent maximum rainfall intensities of top six events with the largest numbers of reports. Rainfall return periods are shown across the range of time-scales, for current climate (KNMI, 2018).

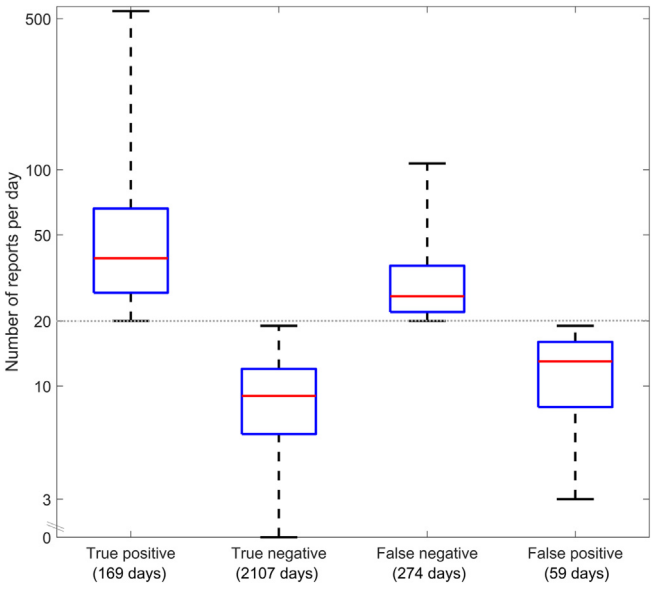


Fig. 8. Boxplots showing the number of reports for TP, TN, FN,and FP events, based on Model 2. Boxes indicate 75th and 25th percentiles and whiskers show minimum and maximum values.

associated with the 10–12 mm storage design guidelines. The short-duration threshold of 3.6 mm/h in 5 min is much lower than the design value, indicating that the effective transport capacity can be lower, locally, than the design value, triggering local flooding. Possible causes of this are blockages of sewer pipes or inlets, as discussed in ten Veldhuis (2010) .

Fig. 7 summarizes the characteristics of rain events responsible for the top 6 largest daily report numbers in the database (> 150 reports/day). In all cases, daily rainfall accumulations were above the 10–12 mm storage capacity. However, only four out of six events exceeded the critical threshold of 21.2 mm at the daily timescale, identified as the second split of the decision tree models 1 and 2. Event 1, associated with the largest number of reports, exhibits the highest peak intensities among all 6 events across all sub-daily time-scales, staying well above the 1-year return period value and reaching above 2-year return period at the half hour, hourly and 2-hour time-scales. Event 2 was also very intense at the hourly and 2-hour time-scales (rainfall accumulations ≥ T=1 year). Event 3 was characterized by high long-duration accumulations, exceeding T=10 years at the daily accumulation scale. Event 4 shows high long-duration accumulations too, in particular at the 6- and 12-hour time-scales. Events 5 and 6, despite having generated large numbers of reports, show low rainfall intensities and daily accumulations. Both rain events stay well below the critical thresholds and half-year return period across all time-scales. This suggests that flooding was associated with other types of failure mechanisms, such as blockages of sewer pipes, pumps or sewer inlets (ten Veldhuis, 2010).

4.2. False positives and negatives

To get a better understanding of the performance of the decision tree models, boxplots of the number of flood reports for all TPs, TNs, FPs, and FNs are shown in Fig. 8, for decision tree model 2. The boxplots clearly show that the majority of false negatives and false positives occur when the number of citizen reports is close to the baseline value of 20 reports/day used to distinguish pluvial flood reports from other non-rainfall related flood reports.

Clearly, the absence of information about the origin of the flooding in citizen data appears to be a limiting factor here. Indeed, in many cases, a mixture of rainfall-related flooding and other failure mechanisms are likely to play a role. Fig. 9 shows rainfall intensities associated with FP and FN events. Many of the FP events are associated with rainfall intensities in the range of $0.3 < Rd_5 < 1$ and $12.7 < Rd_{1440} < 21.2$ mm/h. Plausible explanations for this are (i) under-reporting when flooding happened in an area of lower population density, (ii) under-reporting due to public holidays, lost records or simply a lack of time or staff to attend the phones, and (iii) the possibility that flood reports may have been filed later, on the days following the rain event. In fact, this is hard to verify as citizen flood reports do not come with a precise time stamp except for the day when the report was filed and make no distinction between the time of the incident itself and the day on which it was reported. On the

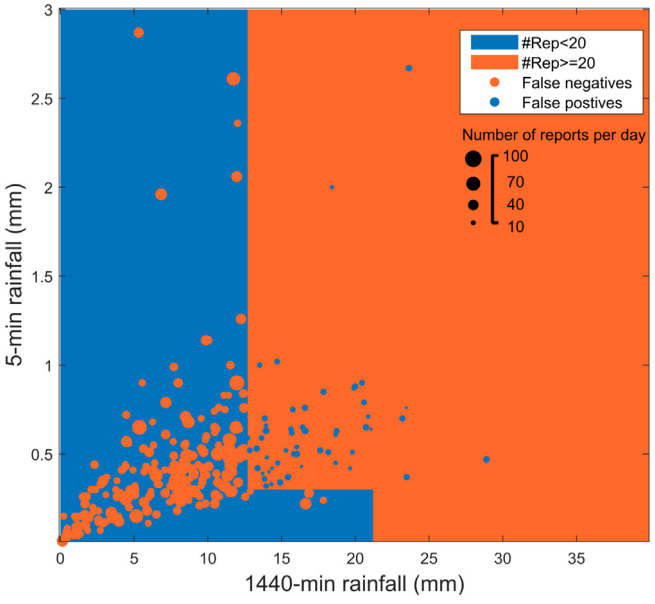


Fig. 9. Erroneously predicted events with the number of their corresponding reports, i.e., false negatives (274) and false positives (59), from Model 2, which is built based on Rd_5 and Rd_{1440} critical rainfall thresholds.

other hand, there were also many FNs in the model, which corresponded to days with no or little rain (e.g., $Rd_{1440} < 5$ mm). The most likely explanation for the flooding in these cases is a blockage in the stormwater system rather than excess rainfall (ten Veldhuis et al., 2011, 2010). Alternative explanations are (i) an underestimation of rainfall by radar, and (ii) large amounts of antecedent rainfall on the days prior to the flooding. Concerning the first point, it is important to remind that the radar rainfall estimates have been bias-corrected using gauge data at the hourly and daily time-scales which means that underestimation at daily time-scales rarely exceeds 10–20 % (Overeem et al., 2009a,b). Underestimation at shorter time-scales is, however, more likely to occur. The antecedent rainfall scenario is difficult to verify as citizen flood reports do not contain a precise time stamp, making it impossible to determine the exact amount of rain that fell on the days leading to the flood event. Another important fact to keep in mind is that reporting rates vary geographically depending on population density and sociodemographic factors such as age, education and cultural background. Also, one has to keep in mind that citizen observations are made by non-experts, which often evaluate situations differently than trained professionals.

4.3. Data quality from citizen observatory

Results of this study imply that citizen flood reports can be a valuable source of information for studying flood incidents and their link to rainfall inputs. However, the way that information is collected means that many practical limitations remain. The first is a general lack of details about what caused the floods (e.g., broken pipe, blocked sewer, malfunctioning pump, excess rain or a combination thereof), which would support a better distinction between pluvial floods and other non-rainfall related incidents. In this study, all days with more than 20 reports were assumed to be pluvial flood events. This baseline was established by taking the 97.5 percentile of the number of reports during non-rainy days. Results show that many days with flood report numbers close to the baseline were misclassified. Apart from that, cases of FPs and FNs (e.g., large rainfall amounts but small numbers of reports or vice-versa) may be a result of imperfect matching between rainfall and reports caused by the poor temporal resolution of citizen data (i.e., the only available information is the day on which the report was filed). For example, if a heavy rain event occurs late in the evening, the reports are likely to be logged on the next day only. If a rain event extended over two days or more, it is unclear how the reports will be distributed over time. To further investigate these timing issues, additional analyses (not shown) were carried out by expanding the time window for the rainfall inputs to 48 h prior to the reports. However, this new strategy did not improve the performance of the decision trees nor did it increase the correlation between rainfall and the number of reports.

Another potential limitation is that critical rainfall thresholds in this study were derived using the full datasets, without accounting for possible trends across years or seasons. The occurrence of blockages may have a seasonal signature, associated with blossom or leaf fall in spring and autumn, respectively. To examine the seasonality of the reports and predicted flooding/non-flooding events, a stacked bar chart is shown for all TPs, TNs, FPs and FNs of Model 2 with respect to their percentages in each month of a year (Fig. 10). It shows that the true positive rate appears to be correlated with the number of monthly reports, reaching a low in the relatively dry months of March, April and May and a high in July, August, and September. This makes sense as the flooding that happened during a relatively dry month is less likely to be a pluvial one. Similarly, the model also appears to be better at predicting non-flooding events in relatively dry months, as shown by the higher true negative rates and lower false negative rates between the months of March and May. On the other hand, the wetter months, e.g. January (with frontal rainfall) and August (with summer storms), contain significantly more pluvial flood events and also more TPs and fewer TNs. These seasonal variations in performance are not a shortcoming of the model itself but more likely the result of a general lack of information about the type of reported flood events. The possibility to distinguish between pluvial and non-pluvial flood reports would probably help significantly in achieving a lower false positive rate during the drier months. In the absence of this information, the model will naturally tend to fit better on months with more reports. The same comment applies to the distribution of model performance across weekdays, although differences there appear to be much smaller than over the months.

Lastly, it is worth pointing out that the results presented in this study are based on observations spanning a relatively long time period of ten years. Classification and tree training were performed on the entire dataset neglecting annual variations, which may not be

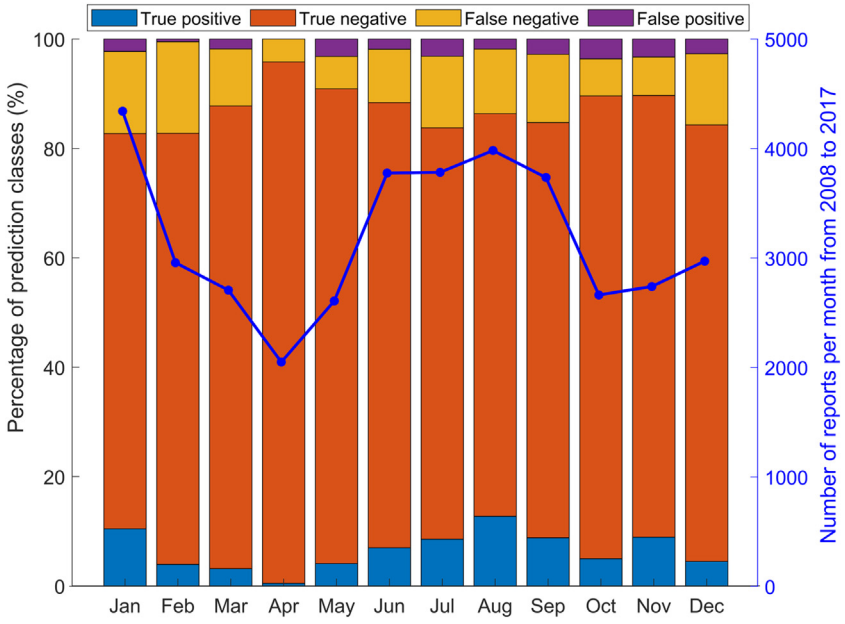


Fig. 10. Average true positive rate (TPR), true negative rate (TNR), false positive rate (FPR) and false negative rate (FNR) of model 2, with respect to each month of the year based on 10 years of data from 2008 to 2017. The blue line shows the average number of reports per month.

the case in reality and does not account for the growth in population nor possible shifts in the ways citizens report flood incidents. For example, a growing number of reports are now submitted through mobile apps in addition to the phone, resulting in approximately 1000 more reports per year since 2014. Reports submitted via the webpage and the app generally contain more details, including (very recently), the possibility to upload photos of the incident. This is an encouraging development that hopefully will help increase the overall quality and consistency of the reports.

5. Conclusion

In this study, we used a ten-year dataset of 70,000 citizen flood reports for the city of Rotterdam and radar rainfall maps at 1 km, 5-minute space-time resolution to derive critical thresholds beyond which urban pluvial flooding is likely to occur. Three decision tree models were trained to predict the occurrence of flooding based on peak rainfall intensities across a range of different temporal scales. Results show that the decision tree models correctly predict 37%–52% of all flood occurrences and 95%–97% of all non-flood occurrences. All models reach the best performance when short- and long-duration rainfall peak intensities are combined. Further improvement can be achieved when critical rainfall thresholds are calculated using a linear combination of peak rainfall intensities across multiple temporal scales, using principal component analysis.

Our results suggest that citizen observatories, although prone to larger errors and uncertainties, constitute a valuable alternative source of information for gaining insight into urban pluvial flooding. The study also highlighted a number of issues related to citizen flood observations which need to be improved. The first concerns the lack of information about the type and origin of the flooding, making it difficult to distinguish between events triggered by intense rain and those triggered by other failure mechanisms (in particular, sewer blockages). The second concerns timing between rainfall and flood reports; a more precise time stamp of flood occurrence (e.g., up to the minute) would enable better identification of rainfall conditions associated with the flood generation. Despite all these issues, results are highly encouraging, clearly highlighting the possibility to infer physically plausible critical rainfall thresholds responsible for urban pluvial flooding. And although the derived rainfall thresholds are specific to Rotterdam, the approach and methods are sufficiently general to be transferable to other urban catchments as well, where a citizen flood reporting system is in place.

Acknowledgments

This research was funded by 'JPI Urban Europe, ERA-NET Co-fund Smart Urban Futures' and 'NWO/VerDuS Smart Urban Regions of the Future (SURF)'. The third author acknowledges funding through 'Water JPI Europe', ERA-NET Co-fund WaterWorks2014 project MUFFIN (Multiscale Urban Flood Forecasting: From Local Tailored Systems to a Pan-European Service).

References

Akter, T., Quevauviller, P., Eisenreich, S.J., Vaes, G., 2018. Impacts of climate and land use changes on flood risk management for the Schijn River, Belgium. Environ. Sci. Policy 89 (August), 163–175. https://doi.org/10.1016/j.envsci.2018.07.002.

Bengio Bengioy, Y., Grandvalet Yvesgrandvalet, Y., 2004. No unbiased estimator of the variance of K-Fold cross-validation. J. Mach. Learn. Res. 5, 1089–1105.

Breiman, L., 1984. Classification and Regression Trees. Routledge, New York, USA.

Butler, D., Davies, J.W. (Eds.), 2004. Summary for policymakers. Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, pp. 1–30. Climate Change 2013 - The Physical Science Basis.

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T.C., Bastiaensen, J., De Bievre, B., Bhusal, J., Clark, J., Dewulf, A., Foggin, M., Hannah, D.M., Hergarten, C., Isaeva, A., Karpouzoglou, T., Pandeya, B., Paudel, D., Sharma, K., Steenhuis, T., Tilahun, S., Van Hecken, G., Zhumanova, M., 2014. Citizen science in hydrology and water resources: opportunities for knowledge generation, ecosystem service management, and sustainable development. Front. Earth Sci. 2 (October), 1–21. https://doi.org/10.3389/feart.2014.00026.

Carpenter, T.M., Sperfslage, J.A., Georgakakos, K.P., Sweeney, T., Fread, D.L., 1999. National threshold runoff estimation utilizing GIS in support of operational flash flood warning systems. J. Hydrol. 224 (1-2), 21–44. https://doi.org/10.1016/S0022-1694(99)00115-8.

Cherqui, F., Belmeziti, A., Granger, D., Sourdril, A., Le Gauffre, P., 2015. Assessing urban potential flooding risk and identifying effective risk-reduction measures. Sci. Total Environ. 514, 418–425. https://doi.org/10.1016/j.scitotenv.2015.02.027.

Cristiano, E., ten Veldhius, M.-C., van de Giesen, N., 2017. Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas - a review. Hydrol. Earth Syst. Sci. 21, 3859–3878. https://doi.org/10.5194/hess-21-3859-2017.

Daniels, E.E., Lenderink, G., Hutjes, R.W.A., Holtslag, A.A.M., 2014. Spatial precipitation patterns and trends in The Netherlands during 1951–2009. Int. J. Climatol. 34 (6), 1773–1784. https://doi.org/10.1002/joc.3800.

Douglas, I., Garvin, S., Lawson, N., Richards, J., Tippett, J., White, I., 2010. Urban pluvial flooding: a qualitative case study of cause, effect and nonstructural mitigation. J. Flood Risk Manag. 3 (2), 112–125. https://doi.org/10.1111/j.1753-318X.2010.01061.x.

Dunn, G., Brown, R.R., Bos, J.J., Bakker, K., 2017. The role of science-policy interface in sustainable urban water transitions: lessons from Rotterdam. Environ. Sci. Policy 73, 71–79. https://doi.org/10.1016/j.envsci.2017.04.013.

Eggimann, S., Mutzner, L., Wani, O., Schneider, M.Y., Spuhler, D., Moy de Vitry, M., Beutler, P., Maurer, M., 2017. The potential of knowing more: a review of data-driven urban water management. Environ. Sci. Technol. 51 (5), 2538–2553. https://doi.org/10.1021/acs.est.6b04267. arXiv:1011.1669v3.

Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27 (8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010.

Gaitan, S., ten Veldhuis, M.-C., van de Giesen, N., 2015. Spatial distribution of flood incidents along urban overland flow-paths. Water Resour. Manag. 29 (9), 3387–3399. https://doi.org/10.1007/s11269-015-1006-y.

Golian, S., Saghafian, B., Maknoon, R., 2010. Derivation of probabilistic thresholds of spatially distributed rainfall for flood forecasting. Water Resour. Manag. 24 (13), 3547–3559. https://doi.org/10.1007/s11269-010-9619-7.

Guyon, I., Elisseeff, A., 2003. An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182. arXiv:1111.6189v1.

Herman Assumpção, T., Popescu, I., Jonoski, A., Solomatine, D.P., 2017. Citizen observations contributing to flood modelling: opportunities and challenges. Hydrol. Earth Syst. Sci. Discuss. (July), 1–26. https://doi.org/10.5194/hess-2017-456.

Huong, H.T.L., Pathirana, A., 2013. Urbanization and climate change impacts on future urban flooding in Can Tho city, Vietnam. Hydrol. Earth Syst. Sci. 17 (1), 379–394. https://doi.org/10.5194/hess-17-379-2013.

Jiang, Y., Zevenbergen, C., Ma, Y., 2018. Urban pluvial flooding and stormwater management: a contemporary review of China's challenges and "sponge cities' strategy". Environ. Sci. Policy 80, 132–143. https://doi.org/10.1016/j.envsci.2017.11.016.

Jolliffe, I., 2002. Principal Component Analysis. Spinger, pp. 518.

Jost, L., 2006. Entropy and diversity. Oikos 113 (2), 363–375. https://doi.org/10.1111/j.2006.0030-1299.14714.x.

Kolsky, P., 1999. Storm Drainage: An Engineering Guide to the Low-Cost Evaluation of System Performance. Practical Action, pp. 152.

Liu, Y., Piyawongwisal, P., Handa, S., Yu, L., Xu, Y., Samuel, A., 2011. Going Beyond citizen data collection with Mapster: a Mobile+Cloud real-time citizen science experiment. 2011 IEEE Seventh International Conference on e-Science Workshops. IEEE, pp. 1–6.

Martina, M.L.V., Todini, E., Libralon, A., 2005. A Bayesian decision approach to rainfall thresholds based flood warning. Hydrol. Earth Syst. Sci. Discuss. 2 (6), 2663–2706. https://doi.org/10.5194/hessd-2-2663-2005.

Mazzoleni, M., Cortes Arevalo, V.J., Wehn, U., Alfonso, L., Norbiato, D., Monego, M., Ferri, M., Solomatine, D.P., 2018. Exploring the influence of citizen involvement on the assimilation of crowdsourced observations: a modelling study based on the 2013 flood event in the Bacchiglione catchment (Italy). Hydrol. Earth Syst. Sci. 22 (1), 391–416. https://doi.org/10.5194/hess-22-391-2018.

Montesarchio, V., Ridolfi, E., Russo, F., Napolitano, F., 2011. Rainfall threshold definition using an entropy decision approach and radar data. Nat. Hazards Earth Syst. Sci. 11 (7), 2061–2074. https://doi.org/10.5194/nhess-11-2061-2011.

Muller, C., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Overeem, A., Leigh, R., 2015. Crowdsourcing for climate and atmospheric sciences: current status and future potential. Int. J. Climatol. 35 (11), 3185–3203. https://doi.org/10.1002/joc.4210.

Norbiato, D., Borga, M., Degli Esposti, S., Gaume, E., Anquetin, S., 2008. Flash flood warning based on rainfall thresholds and soil moisture conditions: an assessment for gauged and ungauged basins. J. Hydrol. 362 (3-4), 274–290. https://doi.org/10.1016/j.jhydrol.2008.08.023.

Overeem, A., Buishand, T.A., Holleman, I., 2009. Extreme rainfall analysis and estimation of depth-duration-frequency curves using weather radar. Water Resour. Res. 45 (10), 1–15. https://doi.org/10.1029/2009WR007869.

Overeem, A., Holleman, I., Buishand, A., 2009. Derivation of a 10-year radar-based climatology of rainfall. J. Appl. Meteorol. Climatol. 48 (7), 1448–1463. https://doi.org/10.1175/2009JAMC1954.1.

Overeem, A., Leijnse, H., Uijlenhoet, R., 2016. Two and a half years of country-wide rainfall maps using radio links from commercial cellular telecommunication networks. Water Resour. Res. 52 (10), 8039–8065. https://doi.org/10.1002/2016WR019412. arXiv:2014WR016527.

Overeem, A., Robinson, J.C.R., Leijnse, H., Steeneveld, G.J., Horn, B.K.P., Uijlenhoet, R., 2013. Crowdsourcing urban air temperatures from smartphone battery temperatures. Geophys. Res. Lett. 40 (15), 4081–4085. https://doi.org/10.1002/grl.50786.

RIONED Foundation, 2017. Sewerage Guidelines: Design and construction (in Dutch). https://www.riool.net/detailontwerp-en-aanleg.

Rosenzweig, B.R., McPhillips, L., Chang, H., Cheng, C., Welty, C., Matsler, M., Iwaniec, D., Davidson, C.I., 2018. Pluvial flood risk and opportunities for resilience. Wiley Interdiscip. Rev. Water 5 (6), e1302. https://doi.org/10.1002/wat2.1302.

Royal Netherlands Meteorological Institute (KNMI), 2019. Daily weather data in the Netherlands. https://www.knmi.nl/nederland-nu/klimatologie/daggegevens.

Schellart, A., Shepherd, W., Saul, A., 2012. Influence of rainfall estimation error and spatial variability on sewer flow prediction at a small urban scale. Adv. Water Resour. 45, 65–75. https://doi.org/10.1016/j.advwatres.2011.10.012.

Schiavo, R.A., Hand, D.J., 2000. Ten more years of error rate research. Int. Stat. Rev. 68 (3), 295–310. https://doi.org/10.1111/j.1751-5823.2000.tb00332.x.

See, L., 2019. A review of citizen science and crowdsourcing in applications of pluvial flooding. Front. Earth Sci. 7, 1–7. https://doi.org/10.3389/feart.2019.00044.

Smith, L., Liang, Q., James, P., Lin, W., 2017. Assessing the utility of social media as a data source for flood risk management using a real-time modelling framework. J. Flood Risk Manag. 10 (3), 370–380. https://doi.org/10.1111/jfr3.12154.

Solomon, D., 2013. Rotterdam Open Database. http://rotterdamopendata.nl/dataset/het-rioolsysteem-van-rotterdam.

Spekkers, M., Kok, M., Clemens, F., ten Veldhuis, M.-c., 2013. A statistical analysis of insurance damage claims related to rainfall extremes. Hydrol. Earth Syst. Sci. 17, 913–922. https://doi.org/10.5194/hessd-9-11615-2012.

Spekkers, M.H., Clemens, F.H.L.R., ten Veldhuis, M.-C., 2015. On the occurrence of rainstorm damage based on home insurance and weather data. Nat. Hazards Earth Syst. Sci. 15 (2), 261–272. https://doi.org/10.5194/nhess-15-261-2015.

Starkey, E., Parkin, G., Birkinshaw, S., Large, A., Quinn, P., Gibson, C., 2017. Demonstrating the value of community-based ('citizen science') observations for catchment modelling and characterisation. J. Hydrol. 548, 801–817. https://doi.org/10.1016/j.jhydrol.2017.03.019.

ten Veldhuis, J.A.E., 2010. Quantitative Risk Analysis of Urban Flooding in Lowland Areas. Ph.D. thesis. Delft University of Technology.

ten Veldhuis, J.A.E., Clemens, F.H., van Gelder, P.H., 2011. Quantitative fault tree analysis for urban water infrastructure flooding. Struct. Infrastruct. Eng. 7 (11), 809–821. https://doi.org/10.1080/15732470902985876.

ten Veldhuis, J.A.E., Harder, R.C., Loog, M., 2010. Automatic classification of municipal call data to support quantitative risk analysis of urban drainage systems. Struct. Infrastruct. Eng. 9 (2), 1–10. https://doi.org/10.1080/15732479.2010.535543.

Thorndahl, S., Einfalt, T., Willems, P., Ellerbæk Nielsen, J., ten Veldhuis, M.-C., Arnbjerg-Nielsen, K., Rasmussen, M.R., Molnar, P., 2017. Weather radar rainfall data in urban hydrology. Hydrol. Earth Syst. Sci. 21 (3). https://doi.org/10.5194/hess-21-1359-2017.

Wang, L.P., Ochoa-Rodríguez, S., Van Assel, J., Pina, R.D., Pessemier, M., Kroll, S., Willems, P., Onof, C., 2015. Enhancement of radar rainfall estimates for urban hydrology through optical flow temporal interpolation and Bayesian gauge-based adjustment. J. Hydrol. 531, 408–426. https://doi.org/10.1016/j.jhydrol.2015.05.049.

Wang, R.-Q., Mao, H., Wang, Y., Rae, C., Shaw, W., 2018. Hyper-resolution monitoring of urban flooding with social media and crowdsourcing data. Comput. Geosci. 111, 139–147. https://doi.org/10.1016/j.cageo.2017.11.008.

Weeser, B., Stenfert Kroese, J., Jacobs, S., Njue, N., Kemboi, Z., Ran, A., Rufino, M., Breuer, L., 2018. Citizen science pioneers in Kenya – a crowdsourced approach for hydrological monitoring. Sci. Total Environ. 631-632, 1590–1599. https://doi.org/10.1016/j.scitotenv.2018.03.130.

Wehn, U., Rusca, M., Evers, J., Lanfranchi, V., 2015. Participation in flood risk management and the potential of citizen observatories: a governance analysis. Environ. Sci. Policy 48, 225–236. https://doi.org/10.1016/j.envsci.2014.12.017.

Wright, D.B., Smith, J.A., Villarini, G., Baeck, M.L., 2014. Long-term high-resolution radar rainfall fields for urban hydrology. J. Am. Water Resour. Assoc. 50 (3), 713–734. https://doi.org/10.1111/jawr.12139.

Yang, P., Ling Ng, T., 2017. Gauging through the crowd: a crowd-sourcing approach to urban rainfall measurement and stormwater modeling implications. 7 (2), 553–572. https://doi.org/10.1002/2017WR020682.

Yang, T.-H., Hwang, G.-D., Tsai, C.-C., Ho, J.-Y., 2016. Using rainfall thresholds and ensemble precipitation forecasts to issue and improve urban inundation alerts. Hydrol. Earth Syst. Sci. 20 (12), 4731–4745. https://doi.org/10.5194/hess-20-4731-2016.

Zhou, Q., Leng, G., Su, J., Ren, Y., 2019. Comparison of urbanization and climate change impacts on urban flood volumes: importance of urban planning and drainage adaptation. Sci. Total Environ. 658, 24–33. https://doi.org/10.1016/j.scitotenv.2018.12.184.

Zhu, D., Peng, D.Z., Cluckie, I.D., 2013. Statistical analysis of error propagation from radar rainfall to hydrological models. Hydrol. Earth Syst. Sci. 17 (4), 1445–1453. https://doi.org/10.5194/hess-17-1445-2013.