

Article

Study on the Classification of Urban Waterlogging Rainstorms and Rainfall Thresholds in Cities Lacking Actual Data

Bingyan Ma , Zening Wu, Huiliang Wang and Yuan Guo *

School of Water Conservancy Engineering, Zhengzhou University, Zhengzhou 450001, China; mabingyan@gs.zzu.edu.cn (B.M.); zeningwu@zzu.edu.cn (Z.W.); wanghuiliang@zzu.edu.cn (H.W.)

* Correspondence: g0628718@zzu.edu.cn

Received: 4 November 2020; Accepted: 24 November 2020; Published: 26 November 2020



Abstract: Extreme rainfall is the main influencing factor of urban waterlogging. Different types of rainfall often have different characteristics of waterlogging. In order to establish a more accurate urban flood control system, it is necessary to classify waterlogging rainstorms and divide their thresholds. This study proposes a method for applying web crawlers to identify waterlogging rainfall in cities lacking waterlogging observation data and classifying them using the rainfall intensity–duration curves. By selecting appropriate duration thresholds and return period, waterlogging rainstorms are divided into rainfall intensity waterlogging (IW), rainfall amount of waterlogging (AW), combined waterlogging (CW) and no waterlogging (NW). In the application of Zhengzhou City, China, the urban flood control standard and the rainfall time distribution characteristics are used as the basis for the selection of the return period and duration thresholds, and the storm water management model (SWMM) is constructed to simulate the 4 kinds of rainfall characteristics of waterlogging, which is similar to actual situations. It proves that the method is suitable for the classification and thresholds division of different waterlogging rainfall in cities. The results show that the best duration thresholds in Zhengzhou are 20 min (M_{20}) and 60 min (M_{60}), and the best return period standard is 2 a. The thresholds for the 4 types of waterlogging rainstorm are: $M_{20} \geq 26.47$ mm, $M_{60} \geq 43.80$ mm, CW; $M_{20} \geq 26.47$ mm, $M_{60} < 43.80$ mm, IW; $M_{20} < 26.47$ mm, $M_{60} \geq 43.80$ mm, AW; $M_{20} < 26.47$ mm and $M_{60} < 43.80$ mm, No waterlogging.

Keywords: urban waterlogging; rainfall thresholds; ID curves; SWMM; classification

1. Introduction

In recent years, with the rapid progress of global warming and urbanization, the heat island and rain island effects have concentrated, leading to an increase in extreme rainfall in urban areas [1–4]. In the rainy season, frequent extreme precipitation events cause urban waterlogging disasters, which seriously affect the lives and property safety of urban residents [5,6]. As the direct driving factor of urban waterlogging disasters, precipitation is also the most uncontrolled factor [7–9], and analyzing its impact on urban waterlogging is of great significance for exploring the characteristics of urban waterlogging and warning.

Many scholars have studied the characteristics of urban floods and waterlogging [10,11]. A general consensus is that excessive accumulated rainfall often causes urban waterlogging and that under the same accumulated rainfall conditions, the shorter the duration, the more likely it is to cause waterlogging disasters [12–14]. It can be seen that the amount and intensity of rainfall are two important factors affecting urban waterlogging, moreover, waterlogging dominated by two factors often has different characteristics. Studies on rainfall thresholds have also found that waterlogging events

are significantly correlated with short-term heavy rainfall and long-term continuous rainfall [15–17]. This indicates that urban waterlogging can be caused by the amount or intensity of rainfall exceeding the thresholds or that it may be caused by the mutual effect of two aspects. For cities lacking waterlogging observation data, it is impossible to directly collect waterlogging information with different characteristics. The application of crowdsourced data such as pipeline maintenance records, municipal center telephone calls, citizen observations and flood insurance provides new solutions to this problem. As a method to efficiently obtain Internet data resources with time and location attributes, web crawlers are used to obtain urban waterlogging information and have good applicability.

When the rainfall intensity exceeds the urban drainage capacity, the excess rainwater cannot be eliminated in time, which mostly occurs in areas where pipelines are blocked or with low design standards, and the surface water will quickly disappear as the rainfall weakens [18]. Long-term low-intensity rainfall weakens the drainage capacity of the pipe network by raising the water head downstream, and gradually causes a large number of nodes to overflow, this type of waterlogging exists for a long time and the water recedes slowly [19]. For extreme rainfall with peak height, large volume and long duration, the above two aspects will appear together [20]. The rational division of urban waterlogging with different characteristics is very important for targeted flood control measures. For short-term heavy rainfall waterlogging, it is only necessary to focus on urban “waterlogging points” or to carry out centralized drainage transformation. Long-term continuous rainfall and waterlogging need to appropriately improve the design standards of the downstream pipeline network to reduce urban waterlogging caused by the phenomenon of backwater jacking. When the two factors appear together, on the basis of the previous measures, it is also necessary to improve the drainage capacity of downstream pumping stations to minimize the loss of waterlogging caused by extreme rainfall.

The rainfall thresholds for the warning indicates the critical value of rainfall that causes a disaster, expressed in terms of accumulated rainfall at a specific time [21–24]. As shown in Figure 1, the lower part of the rainfall-duration (RD) curve is the safe zone, which means that rainfall will not cause waterlogging disasters, and the upper dangerous zone has the opposite meaning [25–27]. For rainfall with similar curvature characteristics, this curve can distinguish the occurrence of urban waterlogging well, but for rainfall with a relatively concentrated peak, the rainfall intensity in a short period of time is greater than the regional drainage capacity, even though it is generally in the safe zone, and it will also cause waterlogging and misjudgment. Therefore, this study has improved the method of dividing the thresholds of waterlogging, replacing the RD curve with the intensity–duration (ID) curves to highlight the impact of rainfall intensity on waterlogging.

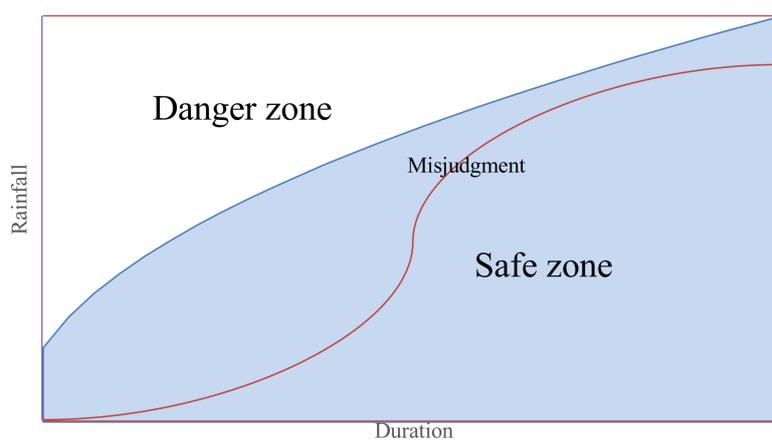


Figure 1. Schematic diagram of rainfall threshold curve.

Although there is a consensus that the characteristics of urban waterlogging dominated by rainfall intensity and amount are significantly different, the methods for effectively distinguishing different types of urban waterlogging rainstorms are not yet mature [28–30]. Compared with the RD curve, the ID curve

can highlight the impact of rainfall intensity on waterlogging. Therefore, it has obvious significance in terms of the improvement in the division of rainfall thresholds. This has been confirmed in the research fields of hydrogeological disasters such as mountain floods, landslides and debris flows. This study applies it to the study of urban waterlogging to comprehensively analyze the effects of rainfall intensity and accumulated rainfall in waterlogging [31–36]. This study is based on the urban waterlogging news obtained by web crawlers, combined with the rainfall ID curves to define and classify waterlogging rainstorms into 4 types: rainfall intensity waterlogging (IW), rainfall amount waterlogging (AW), combined waterlogging (CW) and no waterlogging (NW). Through the waterlogging process simulated by SWMM model, the characteristics of urban waterlogging under different rainfall conditions are further analyzed, and a method for calculating the thresholds value of waterlogging rainstorms is proposed. The application to the study area verifies the rationality and feasibility of the method.

2. Materials and Methods

2.1. Study Area

The study area is the urban area of Zhengzhou City, China, as shown in Figure 2. Zhengzhou is a representative large city in the North China Plain, with a resident population of more than 10 million and an urban area of 1010 km². This century, it has entered a stage of rapid urbanization. The built-up area of the central city area has rapidly expanded from 137.5 km² to 549.3 km², and the urbanization rate has exceeded 70%. The average annual rainfall is 542.15 mm; however, it is affected by the temperate monsoon climate and the hydrological effects of urbanization, more than 65% of the rainfall is concentrated in summer, and the center of heavy rain is mostly located in the central city, making the study area vulnerable to urban waterlogging every summer. From 2011 to 2018, there were 38 waterlogging events, an average of nearly 5 a year, of which 26 had a wide range of impacts. Therefore, it is necessary to study the characteristics of waterlogging rainstorms in this city, and provide a decision-making basis for urban waterlogging warning.

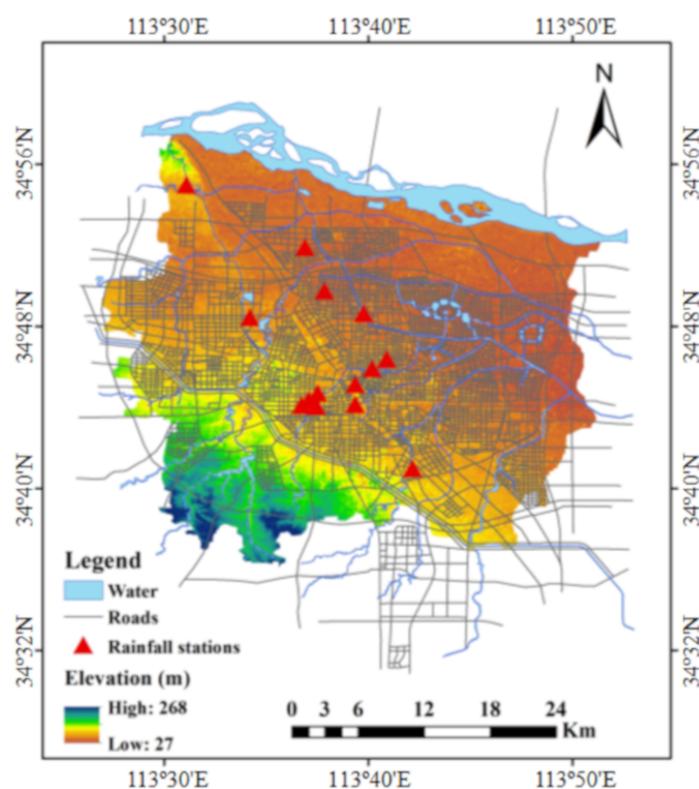


Figure 2. Location of the study area, rainfall stations and roads.

2.2. Data Description

This study needs rainfall and waterlogging data in the study area. Among them, the rainfall data is the 10 min rainfall sequence of the 14 rainfall stations in Zhengzhou and its surrounding areas from 2011 to 2018, and 117 rainfall processes are obtained after division. Table 1 shows the basic information of these rainfall stations.

Table 1. Rainfall station information in the study area.

Station Number	Latitude and Longitude	Time Interval (min)	Station Number	Latitude and Longitude (°)	Time Interval (min)
00	113.670, 34.766	10	07	113.623, 34.872	10
01	113.631, 34.830	10	08	113.673, 34.782	10
02	113.702, 34.761	10	09	113.662, 34.744	10
03	113.767, 34.764	10	10	113.773, 34.721	10
04	113.708, 34.688	10	11	113.670, 34.766	10
05	113.574, 34.817	10	12	113.631, 34.830	10
06	113.315, 34.809	10	13	113.631, 34.830	10

There is no special urban waterlogging measurement and monitoring facility, so it is impossible to obtain structured urban waterlogging sequence data. News media such as urban traffic broadcasting and urban residents have very active attention to the widespread urban waterlogging disasters. Related information will be fed back on the Internet with the occurrence of waterlogging disasters, and the amount of information will increase with the expansion of the scope of the disaster [37]. Therefore, crowdsourced data such as urban waterlogging news and information on the Internet are very valuable urban waterlogging data. The use of web crawlers to collect waterlogging information released by citizens and news media from the Internet can effectively compensate for the lack of data in this area.

Internet crowdsourced data is fragmented information, and it is difficult to describe the whole process of urban waterlogging caused by heavy rain. Constructing an urban rainstorm waterlogging model and calibrating it with waterlogging crowdsourced data can reproduce the urban waterlogging process reasonably well. This method is suitable for cities that lack waterlogging observation data. The basic data obtained from the urban construction management department are used for the construction of the SWMM model. Among them are tertiary highways and rivers used as the boundary of the subcatchments, a digital elevation model (DEM) with a resolution of 30 m, and land-use and pipe networks data which are used for the model runoff generation calculation, involving the setting of basic parameters and the subcatchments and the nodes hydraulic connection. The model is calibrated on the basis of reflecting the actual situation as much as possible to simulate the process of urban rainfall and waterlogging.

2.3. Methods Description

Three parts constitute this research. Part (1) describes the data preparation process, including rainfall data and waterlogging crowdsourced data obtained by web crawlers, as well as the underlying surface basic data used to build the simulation model. Part (2) introduces the analysis method of the characteristics of waterlogging rainstorms. By drawing the ID curves, the waterlogging rainstorms are classified. Part (3) introduces the simulation results of the SWMM corrected by the crawled waterlogging points information, which verifies the rationality and feasibility of the method.

2.3.1. Crowdsourced Data Acquisition of Urban Waterlogging Based on Web Crawlers

Search engines and scalable crawlers are two different ways of obtaining Internet crowdsourced data related to urban waterlogging. In terms of the customizable data acquisition and speed, the latter has more prominent advantages. Therefore, this study uses web crawlers to capture

waterlogging-related data from collection URLs (uniform resource locators) [38]. First, we put all these URLs in an ordered queue in a specific order, extract the URL and download the page, then we analyze the page content, extract the new URL and store it in the queue to be crawled. We repeat the above process until the URL queue is empty or meets specific crawl termination conditions, so as to traverse the web and achieve effective data collection [39].

The object of data acquisition is Sina Weibo V-certified urban waterlogging news. Compared with other types of crowdsourced data, such as municipal center telephone call volume, pipe network maintenance records, Internet flooding news, etc., these data have three advantages. (1) Weibo is a closed platform, V-certified news has high credibility. (2) The attributes of Weibo news data include geographic location and time information, making it easy to select data that meet the requirements. (3) Weibo news data are easier to clean up, with fewer redundant data and higher content value density. Input attributes are location, time and waterlogging keywords in the crawler program, and by simulating landing on the Weibo platform, the information that conforms to the attributes is automatically crawled into a fixed storage path, and the results are displayed in the form of rainfall information and waterlogging information. The method flow is shown in Figure 3.

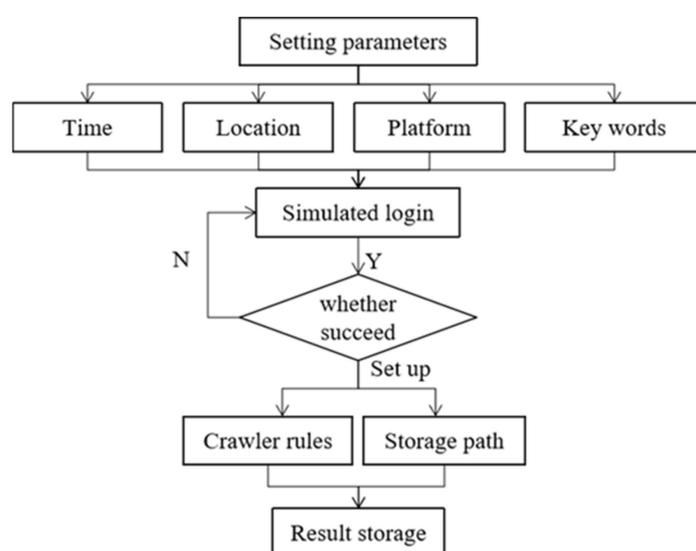


Figure 3. The process of web crawlers obtaining urban waterlogging information.

We change the time parameter in the code to collect multiple rainfall and waterlogging information. In this study, keywords such as rain, waterlogging, and flooding were used individually or in combination to expand the capacity of the target database. The reliability of internet information, collected by web crawlers to supplement the lack of flood-related data can greatly influence the accuracy of the results. From this point of view, data cleaning is considered an important step in collecting reliable data. The research achieves the purpose of data cleaning by eliminating redundancy and wrong information. Redundant information is mainly due to the same news reprinted by different news media, the wrong information mainly comes from news media's reports on flooding news in other regions or other times. For example, on 21 July 2012, Beijing, China suffered the worst rainfall event in the past 70 years, which caused widespread concern. The Weibo media in the research area also reported and reprinted this news in large numbers. These data are invalid for the research and need to be eliminated. In general, the more serious the urban waterlogging, the more information feedback from citizens and news. Therefore, the amount of waterlogging information can be used as an objective standard to reflect the severity of waterlogging. The geographic information involved in the text is the actual waterlogging point of the city.

2.3.2. Waterlogging Rainstorms Thresholds Based on Intensity–Duration (ID) Curves

The return period of rainfall is an important basis for the construction of urban flood control facilities, in a unit of year (a), and it represents the average interval time between the occurrence of rainfall greater than or equal to a certain intensity, and the value is equal to the reciprocal of the frequency of heavy rainfall. For example, the construction standard for road drainage in Zhengzhou is 2 a , which is determined based on the Zhengzhou rainstorm intensity formula, Equation (1).

$$i = \frac{40.1(1 + 0.794 \lg P)}{(t + 25.8)^{0.948}} \quad (1)$$

where i is the rainfall intensity of the design rainfall, mm/min; P is the return period, a ; t is the duration, min.

Urban waterlogging rainfall corresponds to 4 types: short-term heavy rainfall has excessive rainfall intensity, and poorly drained areas cannot drain the excess rainfall in a short time, which can be called rainfall intensity waterlogging; for rainfall with low intensity but long duration, rainwater gradually fills the downstream pipe network and overflow, which is called rainfall amount waterlogging; for rain with both strong intensity and long duration, waterlogging is called combined waterlogging; rainfall with a small amount and intensity always means no waterlogging.

Due to the independence between rainfall events, the characteristics of actual rainfall and designed rainfall are generally different. According to Equation (1), we calculate the return period of all rainfall events at 14 stations, and randomly select 12 events, the ID curves are shown in Figure 4, the dotted lines are the design rainfall of 0.5-, 1-, 2- and 5- a return periods. The ordinate is the maximum average rainfall intensity during the period, and the abscissa is its corresponding duration. By selecting appropriate long-short duration thresholds and return periods, waterlogging rainstorms can be divided into four categories. Taking the duration thresholds of 20 min (M_{20}) and 60 min (M_{60}) and the return period of 1 a as example, the four types of waterlogging rainfall are divided as follows. Table 2 shows the information and classification results of the 12 rainstorms in 4 return periods.

- (1) $M_{20} \geq 21.36$ mm and $M_{60} \geq 35.35$ mm. It shows that both of the intensity and amount of rainfall have reached the flood-causing conditions, which is CW.
- (2) $M_{20} \geq 21.36$ mm, $M_{60} < 35.35$ mm. It means rainfall is concentrated and rapidly attenuating, resulting IW.
- (3) $M_{20} < 21.36$ mm, $M_{60} \geq 35.35$ mm. It shows that the rainfall is uniform and lasts for a long time, the corresponding waterlogging is AW.
- (4) $M_{20} < 21.36$ mm and $M_{60} < 35.35$ mm. Which means the amount and intensity are both small and not enough to cause waterlogging disasters, corresponding to NW.

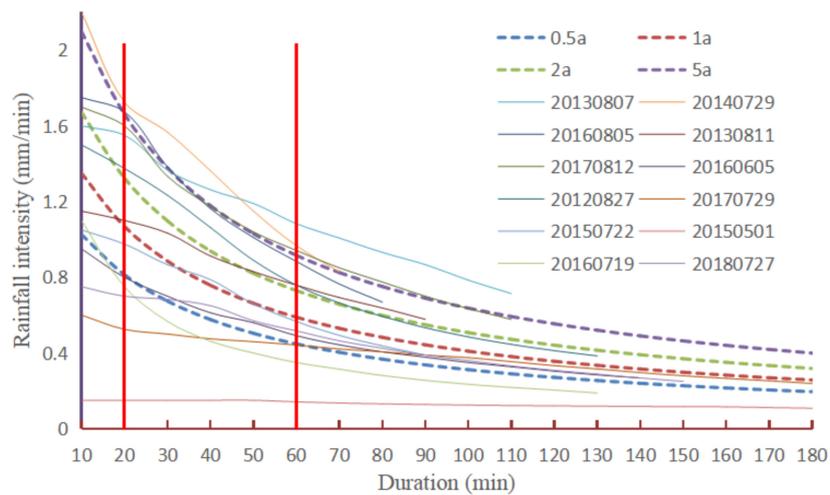


Figure 4. Intensity–duration (ID) curves of design rainfall and measured rainfall in the study area.

Table 2. Information and classification results of the 12 rainstorms.

Rainfall Events	Cumulative Rainfall (mm)	Duration (min)	Maximum Rainfall Intensity (mm/min)	Return Periods (a)			
				0.5	1	2	5
20140729	58.0	70	2.20	CW	CW	CW	CW
20160805	53.5	80	1.75	CW	CW	CW	IW
20170812	63.5	110	1.70	CW	CW	CW	AW
20130807	78.5	110	1.60	CW	CW	CW	AW
20120827	50.0	130	1.50	CW	CW	CW	NW
20130811	52.0	90	1.15	CW	CW	AW	NW
20150722	35.0	90	1.05	CW	NW	NW	NW
20160719	24.5	130	1.10	NW	NW	NW	NW
20160605	37.5	140	0.95	CW	NW	NW	NW
20180727	38.0	150	0.75	AW	NW	NW	NW
20170729	43.0	180	0.60	AW	NW	NW	NW
20150501	19.5	180	0.15	NW	NW	NW	NW

M_{20} and M_{60} indicate the maximum rainfall of 20 min and 60 min during the return period 1 a, which are 21.36 and 35.35 mm.

Different return periods and duration thresholds correspond to different waterlogging disaster discrimination standards. In practical applications, the duration thresholds can be determined according to the regional rainfall characteristics, and then the most accurate return period can be determined according to the waterlogging points information obtained by the crawlers and the flood simulation model, so as to obtain the appropriate waterlogging rainfall thresholds standard.

2.3.3. Analysis of Urban Waterlogging Process Based on SWMM

The urban waterlogging points data obtained by crawlers are often subject to the subjective influence of citizens. For example, rainfall of the same magnitude that occurs during peak hours and in the early morning will have completely different social responses, resulting in a lack of consistency in the data on waterlogging points obtained by crawlers. In contrast, the simulation results of the urban waterlogging model are not affected by data crowdsourcing, and are consistent and objective. Therefore, the use of waterlogging points data for model calibration to obtain reasonable urban waterlogging distribution characteristics is of great significance for judging the rationality of design rainfall in different return periods such as the waterlogging rainstorms thresholds standard.

SWMM, as a mature model used in urban waterlogging simulation research, is suitable for the needs of this research, which is an urban stormwater management model proposed by the US

Environmental Protection Agency (EPA). Since the model was developed in 1971, after more than 40 years of development and application, it has been widely recognized [40–44].

The model approximates the slope confluence as multiple sets of one-dimensional flow processes generated on the slope, and is calculated based on the motion wave equation. The basic principle comes from the simultaneous solution of water balance formula, Equation (2) and Manning formula, Equation (3) [45].

$$\frac{dV}{dt} = F \frac{dh}{dt} = Fr_s - Q \quad (2)$$

$$Q = W \frac{1.49(h - h_p)^{5/3} s^{1/2}}{n} \quad (3)$$

where, F is the subcatchment area, m^2 ; V is the storage capacity of the catchment area, m^3 ; h is the water storage depth of the catchment area, mm ; r_s is the surface runoff rate obtained from runoff analysis, m/s ; Q is the flow rate, m^3/s ; h_p is the water storage depth of the depression, mm ; W is the confluence width, m ; s is the slope of the subcatchment area; n is the roughness.

Substituting (3) into (2), using the Newton–Raphson iterative method to calculate the approximate solution of the finite difference scheme to obtain the water depth process, we obtain Equation (4). In the calculation of slope confluence, the water depth process is obtained by Equation (4), and the flow process can be obtained by introducing Equation (2).

$$\frac{h_2 - h_1}{\Delta t} = \bar{r}_s - K \left(\frac{h_1 + h_2}{2} - h_p \right)^{5/3} \quad (4)$$

where, h_1 and h_2 are the water depths at the beginning and end of the period Δt . K is the slope confluence coefficient, which is given by Equation (5).

$$K = \frac{1.49Ws^{1/2}}{Fn} \quad (5)$$

For the confluence of pipelines and rivers, due to its own linear characteristics of water flow, a one-dimensional water flow formula is used for calculation. Among the three currently popular methods, the constant flow method and the moving wave method are simple generalizations of the actual process. In contrast, the dynamic wave method is most suitable for the calculation of urban pipelines and river confluences. Its governing equations are the Saint-Venant equations composed of the continuous equations and the momentum equations, Equations (6) and (7).

$$\frac{\partial Q}{\partial t} + \frac{\partial A}{\partial t} = 0 \quad (6)$$

$$gA \frac{\partial H}{\partial x} + \frac{\partial(Q^2/A)}{\partial x} + \frac{\partial Q}{\partial t} + gAS_f = 0 \quad (7)$$

where, Q is the flow rate, m^3/s ; A is the cross-sectional area of the water, m^2 ; H is the water depth, m ; g is the acceleration of gravity, 9.8 m/s^2 ; S_f is the friction drop, which can be determined by Equation (8).

$$S_f = \frac{gn^2}{gAR^{4/3}} Q|v| \quad (8)$$

After simplified calculation, the flow rate is given by Equation (9).

$$Q_{t+\Delta t} = \frac{(Q_t + 2\bar{A}\Delta A + \bar{v}^2 \frac{A_2 - A_1}{L} \Delta t - g\bar{A} \frac{H_2 - H_1}{L} \Delta t)}{1 + (J\Delta t/R^{4/3}) \cdot |\bar{v}|} \quad (9)$$

where the subscripts 1 and 2 respectively represent the upstream and downstream nodes of the pipe section or river section; the upper horizontal line represents the average value of the Δt period; L is the pipe section or river section length, m. In addition, the nodes on the pipeline or river must also meet the continuity condition, Equation (10).

$$\frac{\partial H}{\partial t} = \frac{\sum Q_i}{\omega} \quad (10)$$

The finite difference format of the water level of the node can be expressed as Equation (11); where, H is the node water level (or head), m; Q_i is the flow of the node, m^3/s ; ω is the free water surface area at the node, m^2 .

$$H_{t+\Delta t} = H_t + \frac{\sum Q_i \Delta t}{\omega} \quad (11)$$

Combining Equations (9)–(11), the flow rate and node water level of each pipe section or river section can be obtained.

In summary, SWMM can perform nodal overflow calculations, and can also accurately simulate the generation and disappearance of stagnant points. The waterlogging simulation model of the study area is constructed based on SWMM and the waterlogging data obtained by reptiles can be used for model correction, which can carry out reasonable urban waterlogging simulation, judge the rationality of the thresholds division according to the simulation results, and form a complete set of urban waterlogging rainstorms thresholds division methods.

3. Results

3.1. Analysis of Web Crawler Results

Web crawlers crawled 117 rains from 2011 to 2018. After data cleaning, 70 rainstorms with waterlogging news were obtained, and the total number of news items was 2378, which represents the public's attention to rainfall and waterlogging events according to the principle that the higher the public attention, the more serious the waterlogging is. According to the amount of news, the rainstorms were divided into four equal parts, with greater news volume meaning that waterlogging caused more widespread concern, corresponding to a more dangerous disaster, as shown in Table 3.

Table 3. Waterlogging news and rainfall statistics analysis.

News Volume	Degree of Waterlogging	Rain Events	Average News Volume	Average Rainfall (mm)	Average Duration (h)	Average Rainfall Intensity (mm/h)
0–3	No waterlogging	18	2.11	10.94	1.55	7.08
4–10	Less waterlogging	18	6.61	13.50	2.61	5.17
11–39	Waterlogging	17	23.95	28.05	2.46	11.41
>40	Flood disaster	17	96.78	38.61	2.54	15.22

From this table, it can be considered that for a certain rainfall in the study area, if the amount of Sina Weibo waterlogging news reaches more than 10, it can be determined that the rainfall has caused urban waterlogging, and more than 40 can be identified as flood disaster. These two situations correspond to waterlogging rainstorms; therefore, from 2011 to 2018, there were 34 waterlogging rainstorms in the study area. This method is suitable for situations where the public pays more attention to waterlogging news, and there is greater uncertainty. It is necessary to conduct further analysis of the waterlogging rainstorms in order to obtain more reasonable identification indicators.

3.2. Urban Waterlogging Rainstorm Thresholds

Statistics provided the maximum rainfall in different periods of 34 waterlogging rainstorms events, as shown in Figure 5, M_{20} and M_{60} exceed 50% and 80% of the rainfall, so in this study area 20 min and 60 min were used as the duration thresholds.

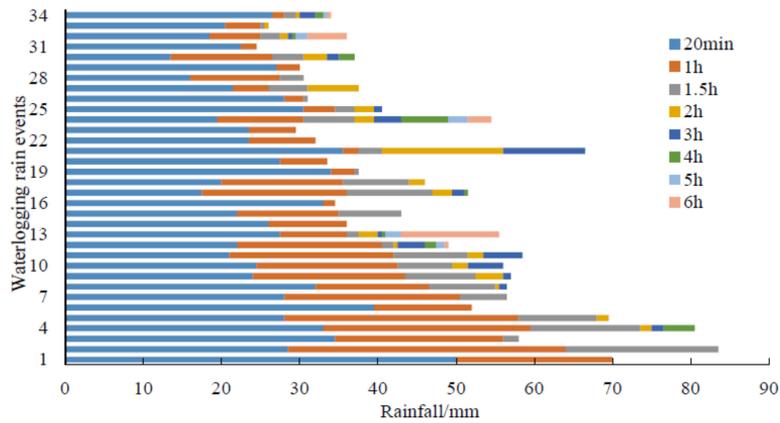


Figure 5. Time distribution of maximum rainfall.

The 26 waterlogging rains with an intermittent time of no more than 20% were regarded as continuous rainfall. The ID curves shown in Figure 6a–d, respectively indicate the return periods of 0.5-, 1-, 2- and 5 a, and the statistics of the 4 types of flooding condition corresponding to each return period are shown in Table 4. At 0.5 a, all waterlogging rainstorms were classified as CW, and this standard is too conservative when used in urban waterlogging warning. In contrast, at 5 a, 15 waterlogging rainstorms were judged as NW, which is too optimistic. The standard division results of 1 a and 2 a are more reasonable. Among them, at 2 a, the three types of waterlogged rainstorms are similar in quantity, and distinguish the characteristics more clearly, which is the most expressive result.

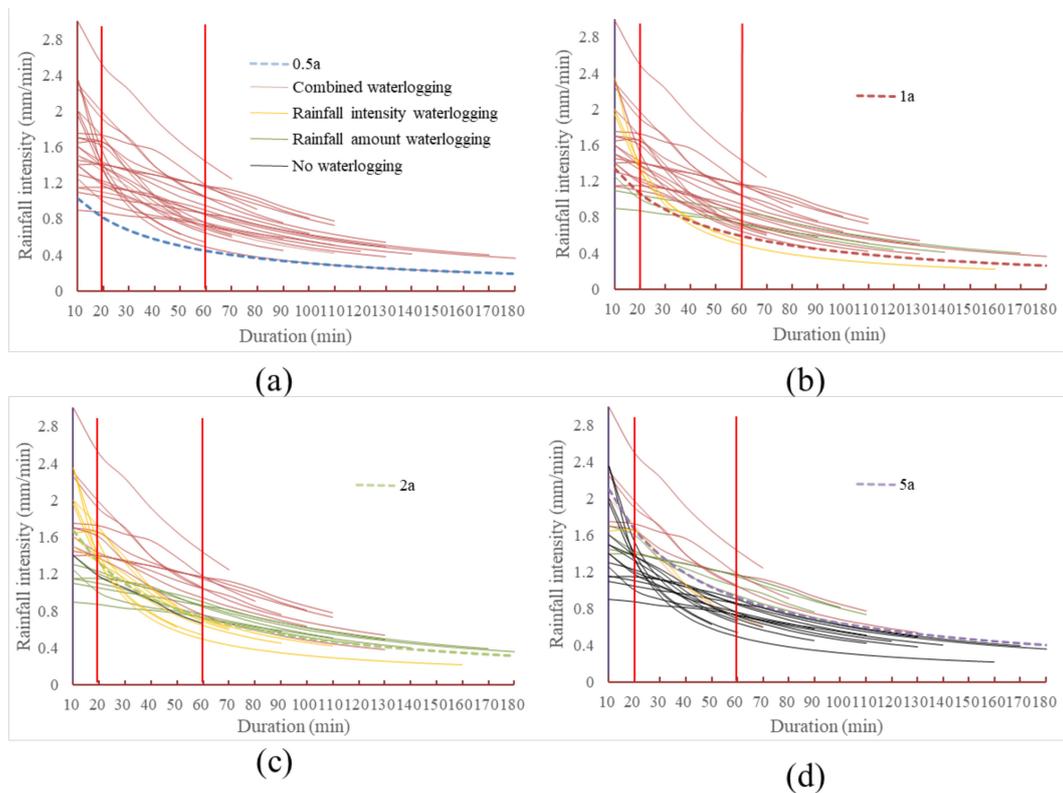


Figure 6. ID curves of waterlogging rainstorms in different return periods. (a) is the result of rainstorm type classification when the return period is 0.5 a; (b) is the result of rainstorm type classification when the return period is 1 a; (c) is the result of rainstorm type classification when the return period is 2 a; (d) is the result of rainstorm type classification when the return period is 5 a.

Table 4. Statistics of waterlogging rainstorms under different return period standards.

Return Periods (a)	Types of Waterlogging Rainstorms	Number	Return Periods (a)	Types of Waterlogging Rainstorms	Number
0.5	CW	26	2	CW	10
	IW	0		IW	8
	AW	0		AW	7
	NW	0		NW	1
1	CW	20	5	CW	5
	IW	3		IW	2
	AW	3		AW	4
	NW	0		NW	15

3.3. Analysis of Waterlogging Characteristics Based on SWMM

The waterlogging model in the study area was constructed based on SWMM and corrected with crawled results. Four waterlogging rainstorms were selected for simulation, representing the 4 possible classification results of rainstorm in different return periods, as shown in Table 5. We randomly select 6 nodes that are prone to waterlogging and plot their water level process. As shown in Figure 7, the ordinate represents the water depth of the node, the abscissa is the duration, the upper part of the abscissa is the ground surface, and the lower part is the underground. Under the condition of 2 a, different types of waterlogging rainstorm have different retreat speeds and waterlogging depths. Among them, the CW retreat process is slowest and the waterlogging is deepest, followed by AW, then IW, and finally NW, when there is only a small amount of waterlogging and it disappears quickly.

Table 5. Judgment results of different return periods of 4 flooding rainstorms.

Return Periods (a)	Rainstorms			
	20160805	20140729	20150803	20150829
0.5	CW	CW	CW	CW
1	CW	CW	CW	CW
2	CW	IW	AW	NW
5	CW	IW	NW	NW

It can be seen from Table 5 that when the return period is 2 a, different waterlogging characteristics of the 4 rainstorms can be distinguished most clearly. Therefore, the waterlogging warning return period of the study area is 2 a, and the waterlogging warning rainfall thresholds is shown in Table 6. In practical applications, it is possible to determine whether waterlogging occurs and the types of waterlogging according to the ID curves of the forecast rainfall or the actual rainfall, and provide corresponding warning or formulate a flood drainage plan.

The distribution of waterlogging points for the 4 rainstorms types of 2 a is shown in Figure 8. Points with different colors indicate the overflow of the nodes, and the larger the point, the greater the overflow. All of the 4 rainstorms occurred in the afternoon, and the waterlogging situation can be compared objectively.

The waterlogging points of CW are generated quickly, in large numbers, have large amount of water, and slowly disappear; IW can quickly generate waterlogging points in areas with poor drainage conditions, but the number of waterlogging points and the volumes of water are small and disappear quickly; the waterlogging points of AW appear in patches, but the occurrence is slightly slower and they exist for longer; and NW has few waterlogging points, disappearing quickly, the volume of water is small, and the impact does not meet the waterlogging warning standards.

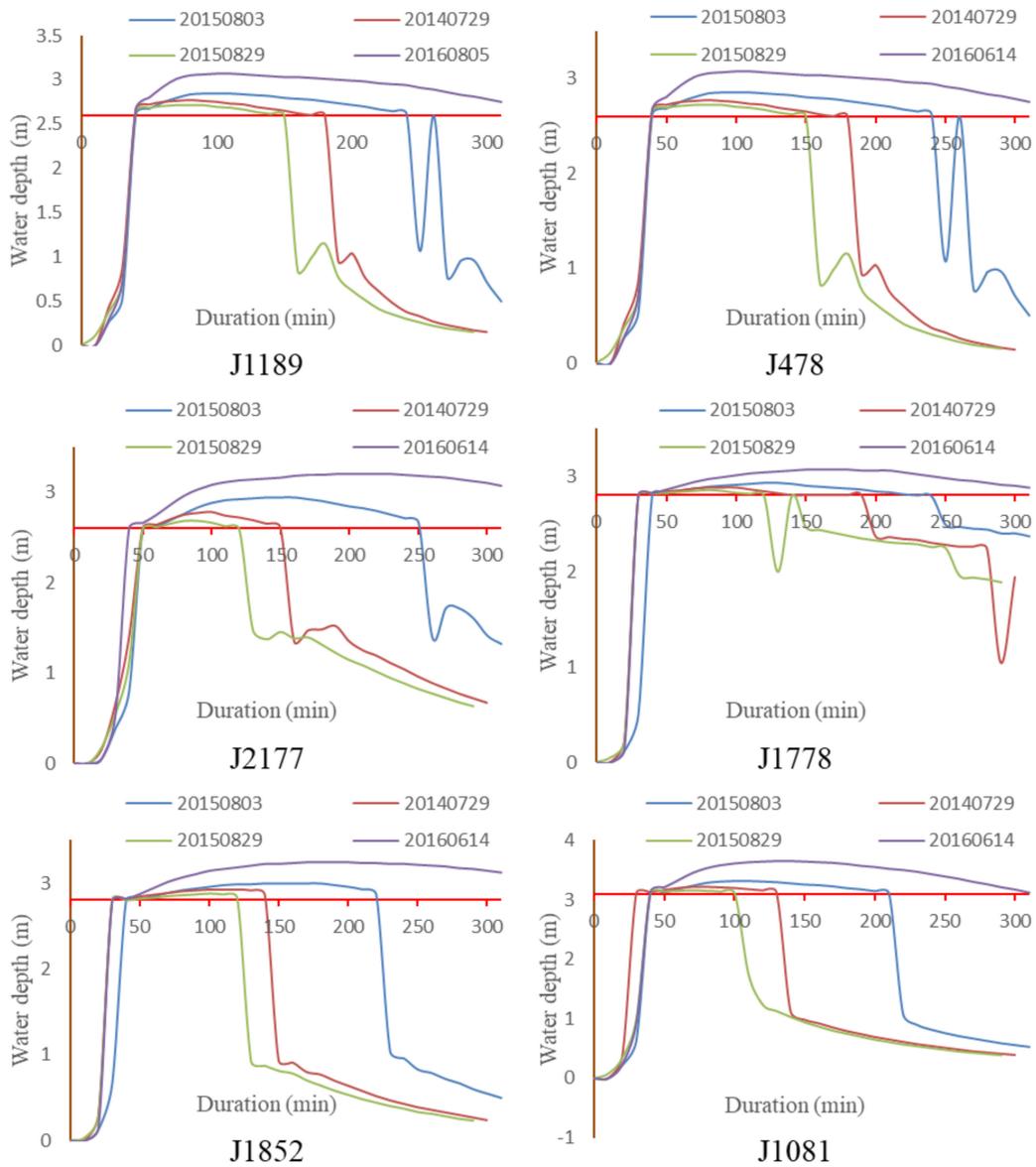


Figure 7. Nodes water level process of different waterlogging rainstorms.

Table 6. Rainfall thresholds for waterlogging warning with a return period of 2 a.

Types of Waterlogging Rainstorms	Maximum 20 min Rainfall (mm)	Maximum 60 min Rainfall (mm)
CW	≥26.47	≥43.80
IW	≥26.47	<43.80
AW	<26.47	≥43.80
NW	<26.47	<43.80

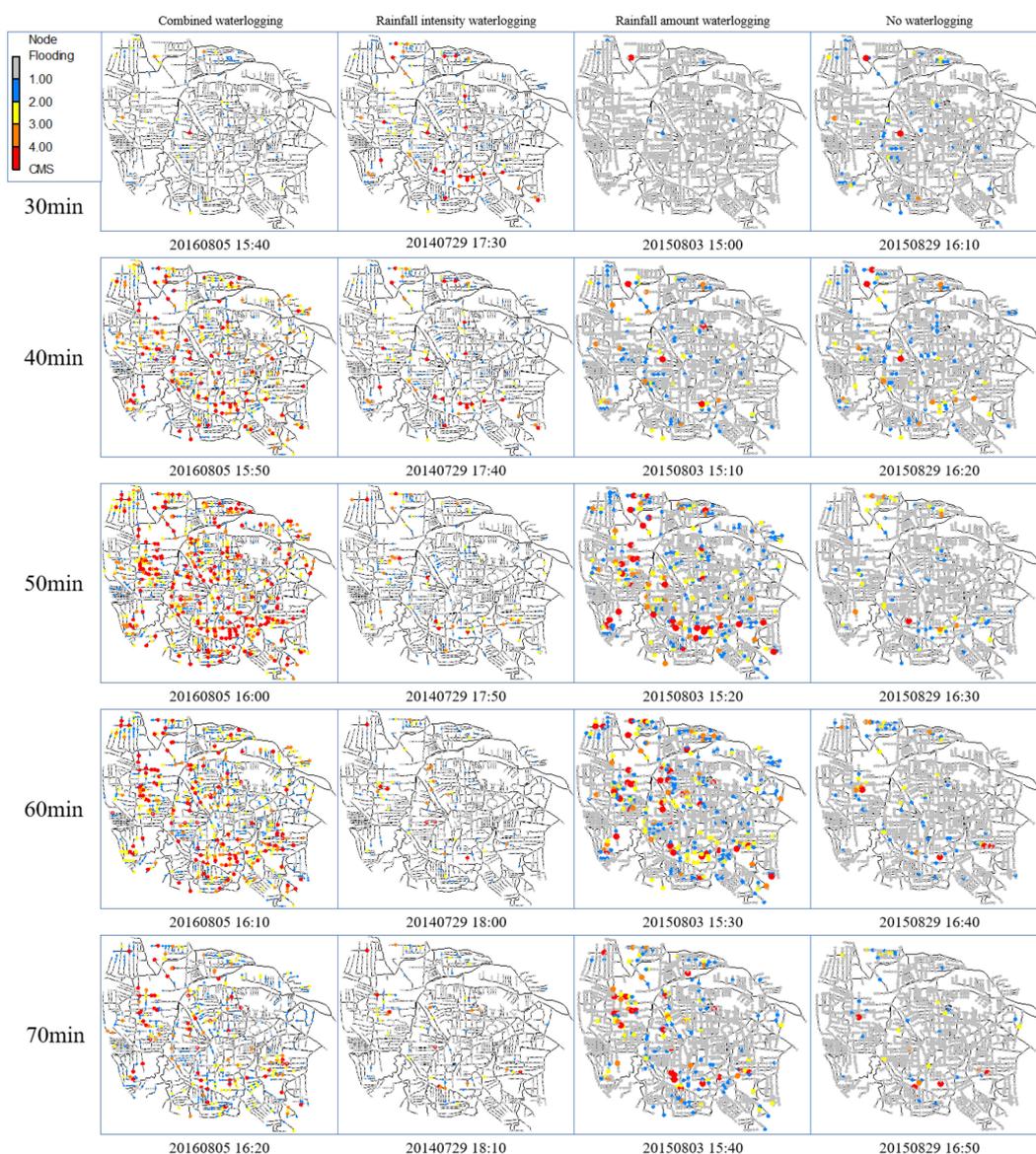


Figure 8. Temporal and spatial distribution of waterlogging points of different types of rainstorm.

4. Discussion

Determining the urban waterlogging rainfall thresholds based on the ID curves with different return periods firstly needs to determine the duration thresholds based on the rainfall characteristics of the study area, and secondly, select the appropriate design rainfall return period based on actual or simulated waterlogging information. The rainfall in this study area generally obeys the P-III distribution and has obvious single-peak characteristics, and M_{20} and M_{60} exceed 50% and 80% of the rainfall, so 20 min and 60 min are used as the duration thresholds.

The waterlogging points information obtained by web crawlers and SWMM simulation shows that although the thresholds of 0.5 a and 1 a can completely identify waterlogging rainstorms, these two standards are too conservative and cannot effectively distinguish the characteristic differences between different types of waterlogging rainstorm. When the return period is 5 a, the waterlogged rainstorms that many citizens are concerned about are judged as NW. Therefore, these three standards are not suitable for application in the study area. The thresholds under the design rainstorm standard of 2 a can accurately identify the waterlogged rainfall and also can reasonably reflect the difference between different types of waterlogging rainstorm, which is the most appropriate standard. In addition,

the duration thresholds and return period can be selected according to actual research or design requirements, in order to obtain the most suitable waterlogging classification standard.

Compared with the research of other scholars, the method is innovative. Chen and Liu [46] in a study of the rainfall thresholds in the urban area of Taiwan, China, regard the rainfall condition as a quantitative external load, and the drainage rate of the pipe network as the internal adaptability; when the load is greater than the adaptive capacity, it will cause urban waterlogging. Therefore, the waterlogging rainfall threshold is determined by calculating the drainage capacity of the rainwater pipe network under rainfall conditions. In fact, the drainage capacity of the pipe network is a complex parameter, and the drainage conditions of different rainfall characteristics are different, so it is difficult to calculate the drainage capacity effectively.

Tian et al. [37] used the municipal center telephone call volume in Rotterdam, the Netherlands as an indicator of urban waterlogging. When waterlogging affecting citizens' lives occurs, the call volume of municipal telephone calls will increase significantly. By analyzing this mutation point, the waterlogging rainfall thresholds can be obtained. This method provides a research plan for areas lacking actual waterlogging data. However, due to the lag of municipal telephone calls and the subjective influence of human behavior on this data, it is impossible to objectively evaluate the degree of waterlogging under different types of rainfall conditions. Under the premise that it is impossible to predict the influencing factors of citizen behavior, and is difficult to make effective application in urban waterlogging warning.

In contrast, this study judges the thresholds of waterlogging and its characteristics based on the intersection of actual rainfall and the ID curves, which more comprehensively reflect the mechanism of rainstorm waterlogging. In the case of forecast rainfall, it can be used as a warning method for urban waterlogging. Furthermore, this method is highly portable, and the use of crawlers to obtain waterlogging data is helpful for the study of areas lacking actual data. Different research areas can determine the duration thresholds and return period standards according to their rainfall characteristics in order to select the best city rainfall thresholds for waterlogging warning.

5. Conclusions

The purpose of this research is to provide a widely applicable method for the division of waterlogging rainfall thresholds in cities lacking actual waterlogging data. Sina Weibo news about urban waterlogging with a large amount of information and high value density are obtained through web crawlers, as a good supplement to urban waterlogging data; according to the quartile relationship between the news volume and the rainfall, waterlogging rainstorms are defined. By analyzing the ID curve relationship between actual and design rainstorms with different return periods, the rainstorms are divided into 4 types: IW, AW, CW and NW. SWMM is used to simulate the flooding process of the waterlogging rainstorms, which verified the rationality and feasibility of the method.

According to the results, in the study area, when the $M_{20} \geq 26.47$ mm and $M_{60} \geq 43.80$ mm, the rainstorm has the characteristics of peak height and large volume. Urban waterlogging will occur quickly and in large quantities, and waterlogging points will gradually increase over a long period of time and disappear slowly, which is called CW.

When the $M_{20} \geq 26.47$ mm and $M_{60} < 43.80$ mm, the peak rainfall is large and rapidly attenuates. Urban waterlogging occurs quickly in local areas with weak drainage capacity and quickly recedes. This type is IW.

If the $M_{20} < 26.47$ mm and $M_{60} \geq 43.80$ mm, this indicates that the rainfall is relatively uniform and there is no excessively high peak. The urban waterlogging is often caused by the water head passing upstream due to poor drainage downstream. The occurrence of urban waterlogging is slightly slower, and disappears slowly, without serious waterlogging disasters. The corresponding rainstorm is AW.

If $M_{20} < 26.47$ mm and $M_{60} < 43.80$ mm, the intensity and amount of rainfall will not reach the threshold values, and a small amount of stagnant water may be generated, but urban waterlogging does not occur, which is NW.

Author Contributions: Conceptualization, B.M. and Y.G.; methodology, B.M. and Z.W.; software, B.M. and H.W.; validation, Z.W., Y.G. and H.W.; data curation, Z.W.; writing—original draft preparation, B.M.; writing—review and editing, Y.G.; visualization, B.M.; supervision, Z.W.; project administration, Y.G.; funding acquisition, Z.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Key Program of National Natural Science Foundation of China (Grant No. 51739009), the National Natural Science Foundation for Young Scientists of China (Grant No. 51909240), the Science and Technology Innovation Talents Project of Henan Education Department of China (Grant No. 21HASTIT011) and the Young backbone Teachers Training Fund of Henan Education Department of China (Grant No. 2020GGJS005). The authors thank the anonymous reviewers for their valuable comments.

Conflicts of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. Miller, J.D.; Kim, H.; Kjeldsen, T.R.; Packman, J.; Grebby, S.; Dearden, R. Assessing the impact of urbanization on storm runoff in a pen-urban catchment using historical change in impervious cover. *J. Hydrol.* **2014**, *515*, 59–70. [[CrossRef](#)]
2. Apel, H.; Trepát, O.M.; Hung, N.N.; Chinh, D.T.; Merz, B.; Dung, N.V. Combined fluvial and pluvial urban flood hazard analysis: Concept development and application to Can Tho city, Mekong Delta, Vietnam. *Nat. Hazards Earth Syst. Sci.* **2016**, *16*, 941–961. [[CrossRef](#)]
3. Dan, M.; Huili, G.; Xiaojuan, L.L.; Siyao, Y. Spatiotemporal distribution of the rainstorm and the relationship between urban heat island and urban rain island in Beijing on July 21, 2012. *Remote. Sens. Land Resour.* **2017**, *29*, 178–185. [[CrossRef](#)]
4. Alexander, K.; Hettiarachchi, S.; Ou, Y.X.; Sharma, A. Can integrated green spaces and storage facilities absorb the increased risk of flooding due to climate change in developed urban environments? *J. Hydrol.* **2019**, *579*, 9. [[CrossRef](#)]
5. Su, M.; Zheng, Y.; Hao, Y.; Chen, Q.; Chen, S.; Chen, Z.; Xie, H. The influence of landscape pattern on the risk of urban water-logging and flood disaster. *Ecol. Indic.* **2018**, *92*, 133–140. [[CrossRef](#)]
6. Liu, J.; Shao, W.W.; Xiang, C.; Mei, C.; Li, Z. Uncertainties of urban flood modeling: Influence of parameters for different underlying surfaces. *Environ. Res.* **2020**, *182*, 108929. [[CrossRef](#)]
7. Merz, B.; Thielen, A.H. Separating natural and epistemic uncertainty in flood frequency analysis. *J. Hydrol.* **2005**, *309*, 114–132. [[CrossRef](#)]
8. Zhang, Q.W.; Yan, F.; Shen, J.; Ye, S.; Ren, B.; Zhang, X.K. A novel seed spread algorithm-based approach for the simulation of rainstorm water logging in urban area. *Desalin. Water Treat.* **2018**, *121*, 265–274. [[CrossRef](#)]
9. Freitag, B.M.; Nair, U.S.; Niyogi, D. Urban Modification of Convection and Rainfall in Complex Terrain. *Geophys. Res. Lett.* **2018**, *45*, 2507–2515. [[CrossRef](#)]
10. Zhang, X.; Hu, M.; Chen, G.; Xu, Y. Urban Rainwater Utilization and its Role in Mitigating Urban Waterlogging Problems—A Case Study in Nanjing, China. *Water Resour. Manag.* **2012**, *26*, 3757–3766. [[CrossRef](#)]
11. Saksena, S.; Dey, S.; Merwade, V.; Singhofen, P.J. A Computationally Efficient and Physically Based Approach for Urban Flood Modeling Using a Flexible Spatiotemporal Structure. *Water Resour. Res.* **2020**, *56*, e2019WR025769. [[CrossRef](#)]
12. Morrison, J.E.; Smith, J.A. Scaling Properties of Flood Peaks. *Extremes* **2001**, *4*, 5–22. [[CrossRef](#)]
13. Hurford, A.P.; Parker, D.J.; Priest, S.J.; Lumbroso, D.M. Validating the return period of rainfall thresholds used for Extreme Rainfall Alerts by linking rainfall intensities with observed surface water flood events. *J. Flood Risk Manag.* **2012**, *5*, 134–142. [[CrossRef](#)]
14. Seenu, P.Z.; Rathnam, E.V.; Jayakumar, K.V. Visualisation of urban flood inundation using SWMM and 4D GIS. *Spat. Inf. Res.* **2020**, *28*, 459–467. [[CrossRef](#)]
15. Egger, C.; Maurer, M. Importance of anthropogenic climate impact, sampling error and urban development in sewer system design. *Water Res.* **2015**, *73*, 78–97. [[CrossRef](#)]

16. Panziera, L.; Gabella, M.; Zanini, S.; Hering, A.; Germann, U.; Berne, A. A radar-based regional extreme rainfall analysis to derive the thresholds for a novel automatic alert system in Switzerland. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 2317–2332. [[CrossRef](#)]
17. Zhou, Z.; Smith, J.A.; Yang, L.; Baeck, M.L.; Chaney, M.; Ten Veldhuis, M.C.; Deng, H.; Liu, S. The complexities of urban flood response: Flood frequency analyses for the Charlotte metropolitan region. *Water Resour. Res.* **2017**, *53*, 7401–7425. [[CrossRef](#)]
18. Hou, J.; Du, Y. Spatial simulation of rainstorm waterlogging based on a water accumulation diffusion algorithm. *Geomat. Nat. Hazards Risk* **2020**, *11*, 71–87. [[CrossRef](#)]
19. Cheng, M.; Qin, H.; Fu, G.; He, K. Performance evaluation of time-sharing utilization of multi-function sponge space to reduce waterlogging in a highly urbanizing area. *J. Environ. Manag.* **2020**, 269. [[CrossRef](#)]
20. Chen, Z.; Yin, L.; Chen, X.; Wei, S.; Zhu, Z. Research on the characteristics of urban rainstorm pattern in the humid area of Southern China: A case study of Guangzhou City. *Int. J. Climatol.* **2015**, *35*, 4370–4386. [[CrossRef](#)]
21. Toth, E. Estimation of flood warning runoff thresholds in ungauged basins with asymmetric error functions. *Hydrol. Earth Syst. Sci.* **2016**, *20*, 2383–2394. [[CrossRef](#)]
22. Knighton, J.; Steinschneider, S.; Walter, M.T. A Vulnerability-Based, Bottom-up Assessment of Future Riverine Flood Risk Using a Modified Peaks-Over-Threshold Approach and a Physically Based Hydrologic Model. *Water Resour. Res.* **2017**, *53*, 10043–10064. [[CrossRef](#)]
23. Vorobevskii, I.; Al Janabi, F.; Schneebeck, F.; Bellera, J.; Krebs, P. Urban Floods: Linking the Overloading of a Storm Water Sewer System to Precipitation Parameters. *Hydrology* **2020**, *7*, 35. [[CrossRef](#)]
24. Arosio, M.; Martina, M.L.V.; Creaco, E.; Figueiredo, R. Indirect Impact Assessment of Pluvial Flooding in Urban Areas Using a Graph-Based Approach: The Mexico City Case Study. *Water* **2020**, *12*, 1753. [[CrossRef](#)]
25. Yang, T.-H.; Yang, S.-C.; Ho, J.-Y.; Lin, G.-F.; Hwang, G.-D.; Lee, C.-S. Flash flood warnings using the ensemble precipitation forecasting technique: A case study on forecasting floods in Taiwan caused by typhoons. *J. Hydrol.* **2015**, *520*, 367–378. [[CrossRef](#)]
26. Bouwens, C.; Veldhuis, M.C.T.; Schleiss, M.; Tian, X.; Schepers, J. Towards identification of critical rainfall thresholds for urban pluvial flooding prediction based on crowdsourced flood observations. *Hydrol. Earth Syst. Sci. Discuss.* **2018**, 1–24. [[CrossRef](#)]
27. Silva Cervantes, M.; Hernando, A.; Garcia-Abril, A.; Valbuena, R.; Velazquez Saornil, J.; Manzanera, J.A. Simulation of overflow thresholds in urban basins: Case study in Tuxtla Gutierrez, Mexico. *River Res. Appl.* **2020**, *36*, 1307–1320. [[CrossRef](#)]
28. Champion, B.B.; Venzke, J.-F. Rainfall variability, floods and adaptations of the urban poor to flooding in Kumasi, Ghana. *Nat. Hazards* **2013**, *65*, 1895–1911. [[CrossRef](#)]
29. Mei, C.; Liu, J.; Wang, H.; Li, Z.; Yang, Z.; Shao, W.; Ding, X.; Weng, B.; Yu, Y.; Yan, D. Urban flood inundation and damage assessment based on numerical simulations of design rainstorms with different characteristics. *Sci. China-Techol. Sci.* **2020**, *63*, 2292–2304. [[CrossRef](#)]
30. Cipolla, G.; Francipane, A.; Noto, L.V. Classification of Extreme Rainfall for a Mediterranean Region by Means of Atmospheric Circulation Patterns and Reanalysis Data. *Water Resour. Manag.* **2020**, *34*, 3219–3235. [[CrossRef](#)]
31. Rodriguez, R.; Navarro, X.; Casas, M.C.; Ribalaygua, J.; Russo, B.; Pouget, L.; Redano, A. Influence of climate change on IDF curves for the metropolitan area of Barcelona (Spain). *Int. J. Climatol.* **2014**, *34*, 643–654. [[CrossRef](#)]
32. Ashish, S.; Babel, M.S.; Weesakul, S.; Vojinovic, Z. Developing Intensity–Duration–Frequency (IDF) Curves under Climate Change Uncertainty: The Case of Bangkok, Thailand. *Water* **2017**, *9*, 145.
33. Sun, Y.; Wendi, D.; Kim, D.E.; Liong, S.-Y. Deriving intensity-duration-frequency (IDF) curves using downscaled in situ rainfall assimilated with remote sensing data. *Geosci. Lett.* **2019**, *6*, 17. [[CrossRef](#)]
34. Lutz, J.; Grinde, L.; Dyrrdal, A.V. Estimating Rainfall Design Values for the City of Oslo, Norway-Comparison of Methods and Quantification of Uncertainty. *Water* **2020**, *12*, 1735. [[CrossRef](#)]
35. Bezak, N.; Šraj, M.; Mikoš, M. Copula-based IDF curves and empirical rainfall thresholds for flash floods and rainfall-induced landslides. *J. Hydrol.* **2016**, *541*, 272–284. [[CrossRef](#)]
36. Zhou, W.; Tang, C. Rainfall thresholds for debris flow initiation in the Wenchuan earthquake-stricken area, southwestern China. *Landslides* **2014**, *11*, 877–887. [[CrossRef](#)]

37. Tian, X.; ten Veldhuis, M.-C.; Schleiss, M.; Bouwens, C.; van de Giesen, N. Critical rainfall thresholds for urban pluvial flooding inferred from citizen observations. *Sci. Total Environ.* **2019**, *689*, 258–268. [[CrossRef](#)]
38. Thelwall, M.; Vann, K.; Fairclough, R. Web issue analysis: An Integrated Water Resource Management case study. *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 1303–1314. [[CrossRef](#)]
39. Brin, S.; Page, L. The anatomy of a large-scale hypertextual web search engine. *Comput. Netw.* **2012**, *56*, 3825–3833, reprinted in *Comput. Netw. ISDN Syst.* **1998**, *30*, 107–117.. [[CrossRef](#)]
40. Huong, H.T.L.; Pathirana, A. Hydrology and Earth System Sciences Urbanization and climate change impacts on future urban flooding in Can Tho city, Vietnam. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 379–394. [[CrossRef](#)]
41. Palla, A.; Gnecco, I. Hydrologic modeling of Low Impact Development systems at the urban catchment scale. *J. Hydrol.* **2015**, *528*, 361–368. [[CrossRef](#)]
42. Behrouz, M.S.; Zhu, Z.; Matott, L.S.; Rabideau, A.J. A new tool for automatic calibration of the Storm Water Management Model (SWMM). *J. Hydrol.* **2020**, *581*, 124436. [[CrossRef](#)]
43. Dai, Y.; Chen, L.; Shen, Z. A cellular automata (CA)-based method to improve the SWMM performance with scarce drainage data and its spatial scale effect. *J. Hydrol.* **2020**, *581*, 124402. [[CrossRef](#)]
44. Jamali, B.; Bach, P.M.; Deletic, A. Rainwater harvesting for urban flood management-An integrated modelling framework. *Water Res.* **2020**, *171*, 115372. [[CrossRef](#)]
45. Gironas, J.; Roesner, L.A.; Rossman, L.A.; Davis, J. A new applications manual for the Storm Water Management Model (SWMM). *Environ. Modell. Softw.* **2010**, *25*, 813–814. [[CrossRef](#)]
46. Chen, C.-F.; Liu, C.-M. The definition of urban stormwater tolerance threshold and its conceptual estimation: An example from Taiwan. *Nat. Hazards* **2014**, *73*, 173–190. [[CrossRef](#)]

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).