Water Resources Management (2022) 36:3699–3714
https://doi.org/10.1007/s11269-022-03223-z

Application of Multinomial Logistic Regression to Model the Impact of Rainfall Genesis on the Performance of Storm Overflows: Case Study

Bartosz Szeląg[1] · **Roman Suligowski**[2] · **Grzegorz Majewski**[3] · **Przemysław Kowal**[4] · **Adrian Bralewski**[5] · **Karolina Bralewska**[5] · **Ewa Anioł**[3] · **Wioletta Rogula-Kozłowska**[5] · **Francesco De Paola**[6]

Received: 28 February 2022 / Accepted: 9 June 2022 / Published online: 28 June 2022

Abstract

In this study, a mathematical model was proposed to analyze the performance of storm overflows. The model included the influence of rainfall genesis on the duration of storm overflow, its volume, and the maximum instantaneous flow. The multinomial logistic regression model, which has not been used so far to model objects located in a stormwater system, was proposed to simulate the duration of storm overflow. The Iman–Conover method, using the theoretical cumulative distributions determined on the basis of 45 – year rainfall sequences, was adopted to simulate the rainfall characteristics describing the overflow performance (total and maximum 30-min rainfall depth and duration). The simulations showed a significant impact of rainfall genesis on the parameters of the storm overflow. The model and the results presented in this study can be used at the stage of dimensioning storm overflows and to create an early warning system against undesirable phenomena in the stormwater system within urban catchments.

Keywords Storm overflow · Multinomial logistic regression · Rainfall genesis · Iman–conover method

1 Introduction

Storm overflows are important elements located in separate and combined sewer systems. Stormwater discharges to receivers are a source of their pollution, because in the process of surface runoff from the catchment area, they are supplied with different pollutants, such as:

✉ Grzegorz Majewski
grzegorz_majewski@sggw.edu.pl

[1] Kielce University of Technology, 7 Aleja Tysiąclecia Państwa Polskiego St., 25-314 Kielce, Poland

[2] Jan Kochanowski University, 7 Uniwersytecka St., 25-406 Kielce, Poland

[3] Warsaw University of Life of Sciences, 166 Nowoursynowska St., 02-776 Warsaw, Poland

[4] Gdańsk University of Technology, 11/12 Narutowicza St., 80-233 Gdańsk, Poland

[5] The Main School of Fire Service, 52/54 Słowackiego St., 01-629 Warsaw, Poland

[6] University of Naples Federico II, via Claudio 21, 80125 Naples, Italy

suspension, heavy metals, pesticides, polycyclic aromatic hydrocarbons (PAHs). A asignificant volume of stormwater during high-intensity rainfall leads to an increase in the flow rate in receivers.

In order to reduce or eliminate the adverse impact of storm overflows on the receiver, the basis for design of storm overflows is usually the limit average rainfall intensity, which is necessary to determine the appropriate height of the overflow edge and its length (US EPA 1995; Zabel et al. 2001). Verification of the correct location of the overflow edge is carried out by analyzing its operation, as well as determining the annual number of storm overflows, their volume, and the maximum instantaneous flow, over a period of one or several years (DWA-A 118E 2006; Chen et al. 2013). To simulate the performance of storm overflows in catchments, hydrodynamic models are usually used (Aryal et al. 2016; Yao et al. 2021). In order to ensure the reliability of the results obtained by the model, detailed information on the catchment area, as well as data on rainfall depth and flows have to be collected for model calibration. The implementation of the above-mentioned models allows for the analysis of the impact of modernization of the rainwater sewage system on the conditions of its operation and selection of the optimal solutions (Jafari et al. 2020).

However, one of the few disadvantages with hydrodynamic models is that due to their local nature, the results obtained with the use of this model cannot be applied to other catchments. Therefore, statistical models are increasingly used to simulate stormwater systems and their components (including storm overflows) (Yao et al. 2021; Szeląg et al. 2020). Thorndahl and Willems (2008) and Grum and Aalderink (1999) developed models simulating the volume of overflow in small urban catchments. Using the First-Order Reliability Method (FORM), Thorndahl and Willems (2008) also developed a model for predicting storm overflow, while in the second case, the analyses also included stormwater quality. Among the statistical models simulating the storm overflow performance, there are few studies where the duration of overflow was analyzed. The research by Montserrat et al. (2015) showed a strong non-linear relationship between the depth of rainfall and the duration of the overflow. However, the obtained value of the coefficient of determination (R^2) ranged from 0.31–0.86, which confirmed the need for more detailed studies. Despite numerous scientific works on storm overflows and attempts to identify rainfall for their design using hydrodynamic models (Yang et al. 2020), a close relationship between the type of rainfall (expressed in its genesis) and the conditions for the operation of the storm overflow have not yet been established. Due to the high costs of the created measurement systems (high-resolution measurements of rainfall depth with implemented complex calculation algorithms for their simulation of the performance of stormwater networks and their components), it seems advisable to look for simplified solutions (Wu et al. 2021).

In this work, we present an innovative mathematical model that allows to analyze the impact of meteorological conditions (rainfall characteristics, taking into account their genesis) on the parameters of storm overflow performance (volume of storm overflows, their duration, and maximum instantaneous flow). To simulate the duration of storm overflow, a multinomial logistic regression model was used, which so far has not been used to model the stormwater system. A synthetic rainfall generator was used to simulate meteorological conditions, including the different genesis of rainfall (convective in the air mass, frontal or rainfall in the convergence zone), based on theoretical density distributions derived from 45-year rainfall series.

2 Material and Methods

2.1 Object of the Study

The analysis was based on a small urban catchment area (Fig. 1) located in the central part of Poland, in the south-western part of the city of Kielce, with a population of 200,000 residents. The average population density in the analyzed catchment is 21.4 people·ha^{-1}. The study area included housing estates, public utility buildings, and communication. The catchment is dominated by urban green zones covering 47.2% of its area while roads with parking lots and pavements and roofs of buildings occupy 37.3% and 14.3% respectively. The highest point of the catchment area was 271.2 m above sea level, while the lowest 260.0 m above sea level. The average decline in the area was 7.1%. The main sewer was 1569 m long and its diameter ranged from 600 to 1250 mm (at the outlet).

The annual rainfall was 537–757 mm. The number of days with rainfall ranged from 155 to 266, and with snow from 36 to 84. The average annual air temperature ranged from 8.1 °C to 9.6 °C. Stormwater outflow from the analyzed catchment to the Silnica River was transferred through the S1 collector. Before entering the river, outflow is brought to the stormwater treatment plant (STP) in the amount of $Q = 200$ dm^3·s^{-1}, where it is cleaned of suspensions. When the level of stormwater in the diversion chamber (DC) exceeded a height of 0.42 m, stormwater was discharged through the overflow (OV) to the discharge channel (S2), through which it was directed to the receiver.

An MES1 flowmeter was installed approximately 3 m from the S1 channel inlet to the separation chamber, which measured the flow values with a resolution of 1 min during intense rainfall events. Continuous monitoring of the volume of stormwater outflowing of the catchment area showed that in dry weather its values ranged 0.001–0.011 m^3·s^{-1}. As part of the research carried out within the city of Kielce, indications of the rainfall station (located 2.5 km from the catchment border) were also used, where continuous measurements of rainfall have been carried out since 2008, with a resolution of 1 min.

Fig. 1 Diagram of the urban catchment area under study

2.2 Rainfall Data

Data of rainfall were obtained from pluvial observations carried out at the meteorological station in Kielce, which was part of the Polish Institute of Meteorology and Water Management (IMGW-PIB) network, using a traditional float pluviograph in the warm half-year (from May to October) in 1961–2005. A detailed description of rainfall data and basic rainfall data are provided in Supplementary Information.

2.3 Analysis of the Influence of Rainfall Genesis on the Storm Overflow Performance

This work proposes an innovative methodology for analyzing the impact of rainfall genesis on the performance parameters of a storm overflow (duration, volume, maximum instantaneous flow) in individual rainfall events. Figure 2 shows the flow-chart of the methodology. The study also analyzed the impact of interaction of the duration of the storm overflow with the genesis of rainfall on other parameters of the overflow. In the adopted approach, the following rainfall types were distinguished – convective in the air mass, frontal (related to the cold front zone and warm front), and rainfall in the convergence zone. The model used an author's (original) solution, based on multinomial logistic regression, to simulate the duration of storm overflow. Due to the complex relationships between the duration of a storm overflow and the characteristics of rainfall, classification variables describing three main conditions, when no storm overflow occurred, the short and long-term of storm overflow was introduced. The adopted method made it possible to analyze the diversified impact of rainfall characteristics on the duration of the storm overflow, taking into account the variability of its numerical values.

The influence of rainfall genesis and the duration of storm overflow on the operational parameters of the storm overflow (volume and maximum instantaneous flow) were analyzed.

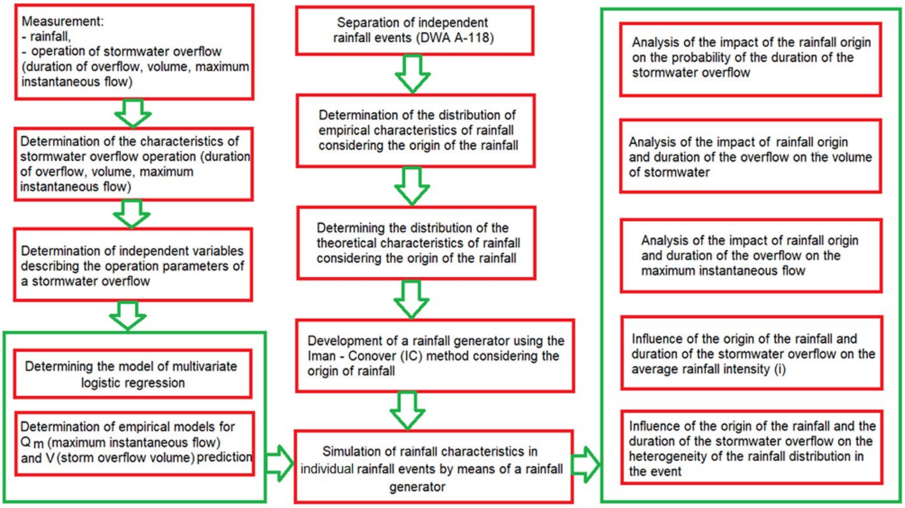

Fig. 2 Algorithm for determining the parameters of the overflow performance, including the genesis of rainfall and duration of storm overflow

In the presented approach, the variability of rainfall characteristics was identified along with the duration of storm overflow. The presented solution, on the one hand, is a simplification of the methodology of modeling the duration of a storm overflow (hydrodynamic models); on the other hand, it makes it possible to build an expert system that allows assess the performance of a storm overflow only on the basis of the established genesis of rainfall. The Supplementary Information explains the computational steps of the algorithm in Fig. 1.

2.3.1 Separation of Independent Rainfall Events

The DWA-A 118E (2006) guidelines were used to distinguish independent rainfall events in rainfall time series. In the adopted approach, the inter-event time separating two independent rainfall events was 4 h. Rainfall of at least 3.0 mm was assumed for the calculations, which was the basis for identification of the rainfall event. The adopted solution was used in the works of other authors dealing stormwater flooding modeling (Fu et al. 2011), or the performance of storm overflows (Szeląg et al. 2020).

2.4 Rainfall Simulation

2.4.1 Determination of Rainfall Characteristics in Independent Rainfall Events

For the separated rainfall events in the time series, the following characteristics were determined: total rainfall depth (P_t), duration of rainfall (t_r), maximum 20-min rainfall depth ($P_{t=20}$) and maximum 30-min rainfall depth ($P_{t=30}$). The relationship between the selected rainfall characteristics and storm overflow volume (V) and maximum instantaneous flow (Q_m) has been analyzed by many authors (Grum and Aalderink 1999; Thorndahl and Willems 2008). Moreover, in the case of the discharge duration, the dependency is more complex, as confirmed by previous works (Wei et al. 2019).

2.4.2 Determination of Theoretical Cumulative Distribution of Rainfall Characteristics

Based on the determined rainfall characteristics (P_t, $P_{t=20}$, $P_{t=30}$, and t_r), the empirical cumulative distribution functions for the separated rainfall events were determined. The cumulative distribution function was determined assuming the division of rainfall data taking into account the rainfall genesis, i.e., convective rainfall ($t_r \leq 150$ min), frontal I – related to the cold front ($t_r \in$ (150, 270 min), frontal II – related to the warm front ($t_r \in$ (270, 630 min), and rainfall in the convergence zone ($t_r > 630$ min). In the next step, the probability density functions described in the literature (Adams and Papa 2000) were adapted to these data. To obtain the best fit between the theoretical and empirical data, several distributions were tested: log-normal, gamma, Weibull, Frechet, Gumbel (Adams and Papa 2000). The χ^2 and Kolmogorov–Smirnov (KS) tests were used to assess the compliance of the empirical and theoretical distribution (Adams and Papa 2000).

2.4.3 Modeling of Synthetic Rainfall Series

The Monte-Carlo (MC) method was used to model the rainfall characteristics in individual rainfall events, constituting synthetic time series. The analyzed variables can be correlated, which was confirmed by many publications (e.g., Wu and Tsang 2004). Therefore, the Iman–Conover (IC) algorithm (Iman and Conover 1982) was used in computational experiments. In this method, the variability of the analyzed variables was expressed by theoretical marginal distributions; a measure of this relationship was the Spearman correlation coefficient. A detailed description of the subsequent steps of the algorithm development were provided in the Supplementary Information.

2.4.4 Empirical Models for Simulating the Performance Parameters of a Storm Overflow

To determine the storm overflow volume and maximum instantaneous flow, the measurement data showing the course of the storm overflow from 2015–2017 were used to select (for separate, independent rainfall events) empirical prediction models (Szeląg et al. 2018):

- storm overflow volume (V)

$$V = 283.12(\pm 7.83) \cdot P_{t=30} - 604.21(\pm 88.35) \quad R^2 = 0.90 \tag{1}$$

- maximum instantaneous flow (Q_m)

$$Q_m = \frac{268.12(\pm 14.93) \cdot \left(P_{t=30} - 2.01(\pm 0.13)\right)}{1 + 0.066(\pm 0.009) \cdot \left(P_{t=30} - 2.01(\pm 0.13)\right)} \quad R^2 = 0.89 \tag{2}$$

2.4.5 Multinomial Logistic Regression

Multinomial Logistic Regression (MLR) is a modification of the binomial logit model. In the classical model, the simulation results are divided into two groups, in which the probability of an event or its absence was identified (Hosmer et al. 2013). In the case of a multinomial model, the simulation results were divided into N classes (also called levels). Despite numerous implementations, it has not been used in modeling objects/elements of stormwater networks so far. In the MLR model, the output data for the operation of the storm overflow can be divided into several groups simultaneously, i.e., there was no storm overflow (2), the duration of storm overflow did not exceed the time limit (1) and the duration of storm overflow exceeded the limit (0). In the MLR model, the probability of overflow duration, for example, at three levels is described by the equations:

$$p(Y = 0) = \frac{exp\left(\beta_{10} + \sum_{i=1}^{k} \beta_{1i} \cdot x_i\right)}{1 + exp\left(\beta_{10} + \sum_{i=1}^{k} \beta_{1i} \cdot x_i\right) + exp\left(\beta_{20} + \sum_{i=1}^{k} \beta_{2i} \cdot x_i\right)} = \frac{exp\left(X_1\right)}{1 + exp\left(X_1\right) + exp\left(X_2\right)} \tag{3}$$

$$p(Y = 1) = \frac{exp\left(\beta_{20} + \sum_{i=1}^{k} \beta_{2i} \cdot x_i\right)}{1 + exp\left(\beta_{10} + \sum_{i=1}^{k} \beta_{1i} \cdot x_i\right) + exp\left(\beta_{20} + \sum_{i=1}^{k} \beta_{2i} \cdot x_i\right)} = \frac{exp\left(X_2\right)}{1 + exp\left(X_1\right) + exp\left(X_2\right)} \tag{4}$$

$$p(Y = 2) = \frac{1}{1 + exp\left(\beta_{10} + \sum_{i=1}^{k} \beta_{1i} \cdot x_i\right) + exp\left(\beta_{20} + \sum_{i=1}^{k} \beta_{2i} \cdot x_i\right)} = \frac{1}{1 + exp\left(X_1\right) + exp\left(X_2\right)} \tag{5}$$

where $Y = 2$ is the event when storm overflow not occurred, $Y = 1$ is the event describing storm overflow with a duration $t_{ov} \in (0, t_{ov,g})$, $Y = 0$ is the event describing storm overflow with a duration $t_{ov} > t_{ov,g}$. For the model determination, 3 cases were assumed (no overflow—inflow of polluted stormwater to the treatment plant; discharge of stormwater by overflow and inflow to the treatment plant, which corresponds to $t_{ov} = (0;75$ min$>$; stormwater outflow to the receiver having a negligible impact on the contamination of the receiver—$t_{ov} > 75$ min. The results of field studies of 10 storm overflows carried out by Ciupa (2009) in the catchments of the Silnica River confirmed that for the discharge time with an overflow longer than 75 min, the values of the suspended solids concentrations in the effluent from the catchment did not exceed 20 mg/L. Similar relationships have been found by Górska (2012) based on the analysis of 10 rainfall events with stormwater overflow.

To evaluate the predictive capacity of the designated MLR model, the results of measurements of storm overflow performance over the period of 2015–2017 were used. They included a total of 140 rainfall events, of which 64 were storm overflow. The duration of storm overflow in the analyzed rainfall events ranged from 6 to 528 min.

2.4.6 Analysis of the Impact of Rainfall Genesis on Overflow Operation Parameters and Rainfall Characteristics in a Single Event

As part of the conducted analyses, the parameters of storm overflow performance (storm overflow volume, maximum instantaneous flow, and duration of storm overflow) were determined as dependent on the genesis of rainfall (convective in the air mass, frontal in the cold front zone, frontal associated with the warm front, or rainfall in the convergence zone). For this purpose, the models presented in the sections: Sects. 2.4.4 and 3.2 were used to analyse the influence of the rainfall genesis and the duration of storm overflow on the variability of the overflow parameters (V, Q_m). Average values of the overflow performance parameters (storm overflow volume – $E(V)$, maximum instantaneous flow – $E(Q_m)$, probability of the duration of storm overflow – $E(p_{tv,g})_k$) were determined using the formulas presented in Supplementary Information.

2.4.7 SWMM Hydrodynamic Model and MLR Verification

To verify the simulation with the MLR model, a calibrated hydrodynamic model of the analyzed urban catchment with a storm overflow was used, developed with the SWMM 5.0 (Storm Water Management Model) program. The catchment area was 62 ha, while the sub-catchment area ranged from 0.12 ha to 2.10 ha. The number of nodes was 80, and the number of pipes was 72. The retention of the catchment areas was 2.5 mm, and that of the unsealed areas were 6.0 mm. The Manning roughness coefficient of the pervious areas was 0.025 $m^{-1/3} \cdot s$ and 0.25 $m^{-1/3} \cdot s$ for the impervious areas. The Manning channel roughness coefficient was 0.018 $m^{-1/3} \cdot s$.

The subject model of the catchment area has been used to simulate the runoff of stormwater, its quality and the dimensioning of the retention reservoirs. Model was calibrated with the deterministic and probabilistic Generalized Likelihood Uncertainty + Global Sensitivity Analysis (GLUE + GSA) method. On the basis of the discussed model (catchment – diversion chamber – storm overflow), probabilistic models were verified for prediction the storm overflow volume and maximum instantaneous flow (Szeląg et al. 2018).

2.4.8 Verification of Calculation Results with a Mathematical Model

In order to verify the calculations (average values) of the probability of the duration of storm overflow, storm overflow volume and maximum instantaneous flow, continuous simulations of the overflow operation were carried out based on a 45-year rainfall series, by calibrated hydrodynamic model of the catchment and storm overflow, developed with the SWMM program, was used.

3 Results

3.1 Duration of Storm Overflow

Based on the obtained measurements, it was found that the average value of the storm overflow duration in the analyzed rainfall events was 75 min (50% percentile), while the 5% and 95% percentile values were 11 min and 240 min, respectively (Szeląg et al. 2018). The values of the maximum instantaneous flows (Q_m) were 0.01–2.50 $m^3 \cdot s^{-1}$, while the average value was 0.87 $m^3 \cdot s^{-1}$. The storm overflow volume was 5–23,008 m^3, and the average was 1801 m^3. As part of the analyses, three levels were introduced to determine the MLR model using the measurement data, i.e., storm overflow not occurred ($t_{ov} = 0$), duration of storm overflow equal of $t_{ov} \in$(0, 75 min, and duration of storm overflow equal of $t_{ov} > 75$ min.

3.2 Multinomial Logistic Regression

The MLR model was developed on the basis of rainfall measurement data and performance of the storm overflow from 2015–2017. Equations (6) and (7) were obtained to determine the probability of: storm overflow not occurred, duration of storm overflow $t_{ov} \in$(0, 75 min, and duration of storm overflow $t_{ov} > 75$ min, described as:

$$X_1 = 0.467(\pm 0.012) \cdot P_t + 1.725(\pm 0.034) \cdot P_{t=30} - 0.0022(\pm 0.0004) \cdot t_r - 8.398(\pm 0.956) \tag{6}$$

$$X_2 = 0.257(\pm 0.018) \cdot P_t + 1.824(\pm 0.024) \cdot P_{t=30} - 0.0001(\pm 0.00002) \cdot t_r - 6.967(\pm 0.426) \tag{7}$$

The calculations showed that in the analyzed time intervals t_{ov}, the rainfall depth had a diversified impact on the probability of the duration of storm overflow. $P_{t=20}$ had no statistically significant effect on the duration of storm overflow (p-value was greater than 0.05). The obtained MLR models were characterized by high predictive abilities. In the case of the models described by Eqs. (3), (5), (6) and (7), the following results were obtained: for 75 episodes of storm overflow did not occur, the results of MLR calculations were consistent with the measurements in 71 events (SENS = 91.78%); for 40 events when the duration of storm overflow $t_{ov} \in$ (0, 75 min, the calculation results were consistent with the measurements for 37 rainfall events (SENS = 92.15%); and for 24 rainfall events when $t_{ov} > 75$ min, the MLR simulation results were consistent with the measurements for 23 rainfall events. Out of 140 rainfall events, the calculation results obtained using the MLR model were consistent with 85.2% of the measurement results. The validation included 10 rainfall events. Field studies and simulations showed that storm overflow did not occur in three of them. In four events, measurements confirmed that $t_{ov} \in$(0, 75 min and calculations by the MLR model showed that t_{ov} in this range fitted three episodes. Of the three episodes where $t_{ov} > 75$ min, two cases were consistent with the measured data. The results obtained in this manner confirm the possibility of using an MLR to model the duration of storm overflow. Examples of logit curves determined by using Eqs. (3)–(7) are shown in Fig. 3. The analysis of the curves shows that both P_t and $P_{t=30}$ have a significant impact on the probability of duration of storm overflow in the range of 0 min to $t_{ov} \in (0$, 75 min and $t_{ov} > 75$ min. The significant influence of the total rainfall depth on the CSO (Combined Sewer Overflow) duration was confirmed by the results of the analyses of Fortier and Mailhot (2015), who they developed a mathematical model for calculating the duration of a combined sewer on the basis of the measurement results of the CSO. However, the results of the calculations based on the database of 30 overflows showed that the obtained R^2 values ranged from 0.10–0.63. This may indicate that their model did not include all the significant independent variables influencing the duration of storm overflow.

With regard to the curves illustrating the probability of the duration of storm overflow $t_{ov} \in (0{,}75$ min, it was found that p values increased with increasing rainfall depth (P_t). Regarding the curves illustrating the probability of the duration of storm overflow $t_{ov} > 75$ min, it was shown that an increase in rainfall depth to an appropriate extent leads to an increase in the p-value. Once it is achieved, a further increase in the P_t value leads to a decrease in the probability of no occurrence of storm overflow and duration of storm overflow $t_{ov} > 75$ min. The above results illustrate the variability of the probability of an event eccurence in relation to all analyzed levels (storm overflow not occurred, duration of storm overflow up to 75 min, and above 75 min), which is a significant advantage of the MLR

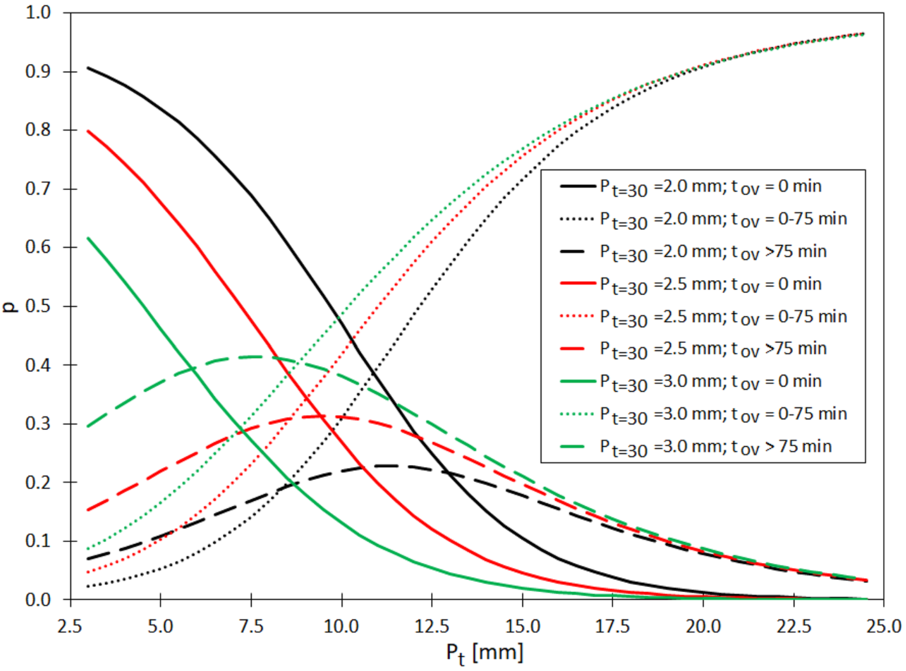

Fig. 3 Impact of total rainfall depth (P_t) and maximum 30-min rainfall in the event ($P_{t=30}$) on the probability the duration of storm overflow (p)

model. This allows for a detailed analysis of the overflow operating conditions in relation to the duration of storm overflow depending on the value of the rainfall characteristics.

3.3 Verification of the MLR Model Using SWMM

The results obtained with the MLR model were compared with the measurement data and simulation results obtained with a calibrated hydrodynamic model of the catchment area (SWMM). Cases where there was no storm overflow were compared, and the duration of storm overflow was equal t_{ov} ∈(0, 75 min and $t_{ov} > 75$ min (Table S1 in Supplementary Information). Based on these data, it can be concluded that the MLR model is characterized by high predictive abilities, which is confirmed by the high convergence of the measurement and simulation results (SWMM and MLR) of the storm overflow in 2015–2017. The difference of the simulation results between MLR and measurements did not exceed two overflows in consecutive years.

3.4 Identification of Empirical and Theoretical Distributions of Rainfall Characteristics

Based on the results of the polynomial logistic regression models, empirical distributions were determined, and theoretical distributions were adjusted to them. The results of the analyses and the obtained parameters of theoretical distributions in relation to rainfall depth (P_t) and its duration (t_r) were given in the previous study (Szeląg et al. 2020). Table S2 (Supplementary Information) presents the optimal theoretical distributions, their parameters, and the results of the Kolmogorov–Smirnov (KS) and χ^2 tests. The calculation results (p-value) for other statistical distributions (Frechet, Gumbel, Weibull, gamma) are given in Table S3 (Supplementary Information). The high adjustment of the theoretical distributions to the empirical data allowed for the calculations of the rainfall characteristics in events, which was confirmed by the results of the K-S and χ^2 statistical tests (as indicated by the calculated values of the test probability (p-value)).

3.5 Impact of Rainfall Genesis on the Probability of Duration of Storm Overflow

On the basis of theoretical distributions (Table S3 in SI) and the Iman–Conover method, sampling (10,000) of rainfall characteristics (P_t, $P_{t=30}$, t_r) was performed. The simulation results were replaced with Eqs. (3), (5), (6) and (7) and the probability p in the rainfall event was determined. The simulations included three cases: when there was no storm overflow ($t_{ov} = 0$), duration of storm overflow of $t_{ov} \in (0, 75 \text{ min})$, and duration od storm overflow of storm overflow of 75 min ($\mathrm{t_{ov}} > 75$ min). The results of the calculations are shown in Fig. 4. They were compared with the calculations in which the genesis of rainfall was omitted. Based on the calculations and determined curves (Fig. 4), it can be concluded that in the absence of a storm overflow, the highest possibility of exceeding the probability of no storm overflow for the 50% percentile was obtained for long-term rainfall associated with convergence zones (0.58). A clearly lower probability was found for frontal rainfall (frontal I – cold front: 0.28 and frontal II – warm front: 0.16), and a significantly low probability was obtained for short-term convective rainfall in the air mass (0.0001). The results of the simulation, in which the genesis of rainfall was omitted, gave the probability

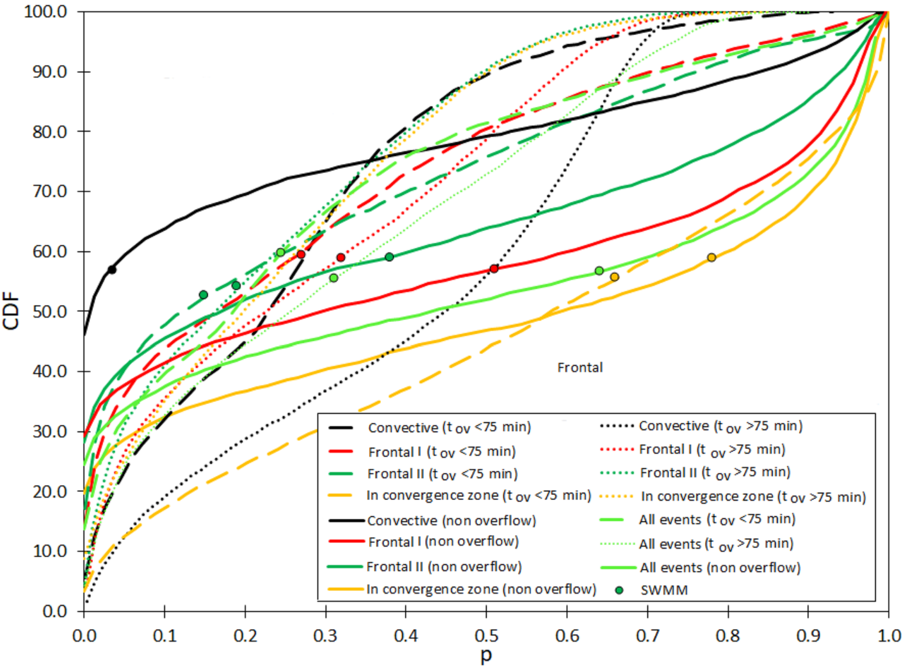

Fig. 4 Impact of rainfall genesis (convective in the air mass, frontal, and rainfall in the convergence zone) on the probability of storm overflow duration

of no storm overflow at the level of 43% for the 50% percentile. The calculation results obtained in two groups, t_{ov} ∈(0, 75 min and $t_{ov} > 75$ min, deserve special attention. In the first group, for the 50% percentile, the highest probability of the storm overflow occurrence was obtained for the rainfall in the convergence zone (0.58).

The omitting genesis resulted in a percentile value of 0.19. For storm overflow with durations longer than $t_{ov} > 75$ min, the highest percentile values were obtained for convective rainfall in the air mass (0.44) and rainfall related to the cold front (frontal I – 0.23). The probability of the duration of storm overflow (50% percentiles) obtained MLR model showed a good agreement with those of the calibrated hydrodynamic model of the catchment area (Fig. 4); this confirms that the model can be used in operation. It has been shown that duration of storm overflow is important for the performance of storm overflow construction (high and length of the overflow edge). In conjunction with the genesis of rainfall, it provided important information about function of the storm overflow for various rainfall conditions, which can be used to analyze the performance of a storm overflow and to design the height and length of its edge. Jean et al. (2018) made an attempt to select rainfall characteristics for the design of the overflow, by the use a calibrated model of a 410 ha urban catchment area in the province of Québec (Canada) to simulate the operation of the storm overflow. They used three approaches to model overflows: continuous simulations, separate independent rainfall events and IDF (Intensity–Duration–Frequency) curves. However, they did not establish explicit results, which could have resulted from the failure to take into account rainfall genesis in the analysis.

3.6 Impact of Genesis and Duration of Storm Overflow on the Volume and Maximum Instantaneous Flow

Following the computational algorithm (Fig. 2) and based on the obtained empirical models described by Eqs. (3)–(7), the maximum instantaneous flow and the storm overflow volume were determined. At the same time, rainfall genesis and duration of storm overflow were taken into account. The results of the calculations were presented in Fig. 5. Based on the analysis of the curves, it can be concluded that when modeling the parameters of the storm overflow, both the rainfall genesis and duration of storm overflow are important. The maximum instantaneous flow and storm overflow volume obtained for the calculation were $Q_m = 0.55\ \mathrm{m^3 \cdot s^{-1}}$ and $V = 628\ \mathrm{m^3}$, respectively, when all rainfall events were taken into account, excluding the genesis, for the 50% percentile. Considering the genesis of rainfall phenomena leads to a significant differentiation of the obtained percentile values. The highest values of the maximum instantaneous flow and storm overflow volume (50% percentile) were obtained for convective rainfall $Q_m = 1.03\ \mathrm{m^3 \cdot s^{-1}}$ and $V = 1422\ \mathrm{m^3}$, and the lowest for rainfall in the convergence zones $Q_m = 0.41\ \mathrm{m^3 \cdot s^{-1}}$ and $V = 442\ \mathrm{m^3}$. The calculation results for the maximum instantaneous flow and storm overflow volume (50% percentile) for convective rainfall in the air mass and in the cold front zone (frontal I) are greater than their counterparts when the rainfall genesis is omitted.

It is worth noting that the calculation results of the maximum instantaneous flow and storm overflow volume (50% percentile), omitting the rainfall genesis, differ by no more than 3% from the results of Q_m and V when the duration of storm overflow is longer than 75 min. The simulation results of the maximum instantaneous flow and storm overflow volume (50% percentile), obtained assuming a shorter duration of storm overflow (up to 75 min) for frontal rainfall (frontal I and frontal II) and rainfall in the convergence zone, differ by no more than 5% of the results of Q_m and V calculations, taking into account the genesis of rainfall. The calculated values of the maximum instantaneous flow and storm overflow volume obtained for convective rainfall $t_{ov} \in (0, 75$ min, when only the genesis of rainfall is included, are higher than the simulation results of Q_m and V by 20% and 37%, respectively.

Relating the obtained results to the aspects of designing the storm overflow, it can be stated that for convective and frontal rainfall, the obtained Q_m and V values are greater than those obtained without considering the genesis. An effort to determine the impact of the rainfall genesis on the volume of the storm overflow in the Beargrass Creek catchment area located in Louisville (USA) was demonstrated by Hyun et al. (2016). The conditions of rainfall formation in this region (humid subtropical climate) and the assumed types of rainfall (e.g. tropical convective rainfall and east cool strati form rainfall) are different from those defined in Central Europe (Szeląg et al. 2020). Their analyses were not aimed at developing relationships between the genesis of rainfall and certain parameters of the overflow performance, thus were limited to determining only simple relationships. Although the genesis of rainfall has a significant impact on the parameters of the stormwater system performance, including the facilities located on it, it has been neglected in most studies. In the context of the results obtained in the study, this led to an underestimation of calculation values of storm overflow volume, instantaneous flows in rainfall events in probabilistic models (Grum and Aalderink 1999; Szeląg et al. 2018). It also resulted in an underestimation of the amount of pollutants transported in the watercourse (mineral suspensions, heavy metals, etc.), which was modelled omitting the genesis by Grum and Aalderink (1999). High-intensity of rainfall events (convective, frontal I) may lead to local flooding and

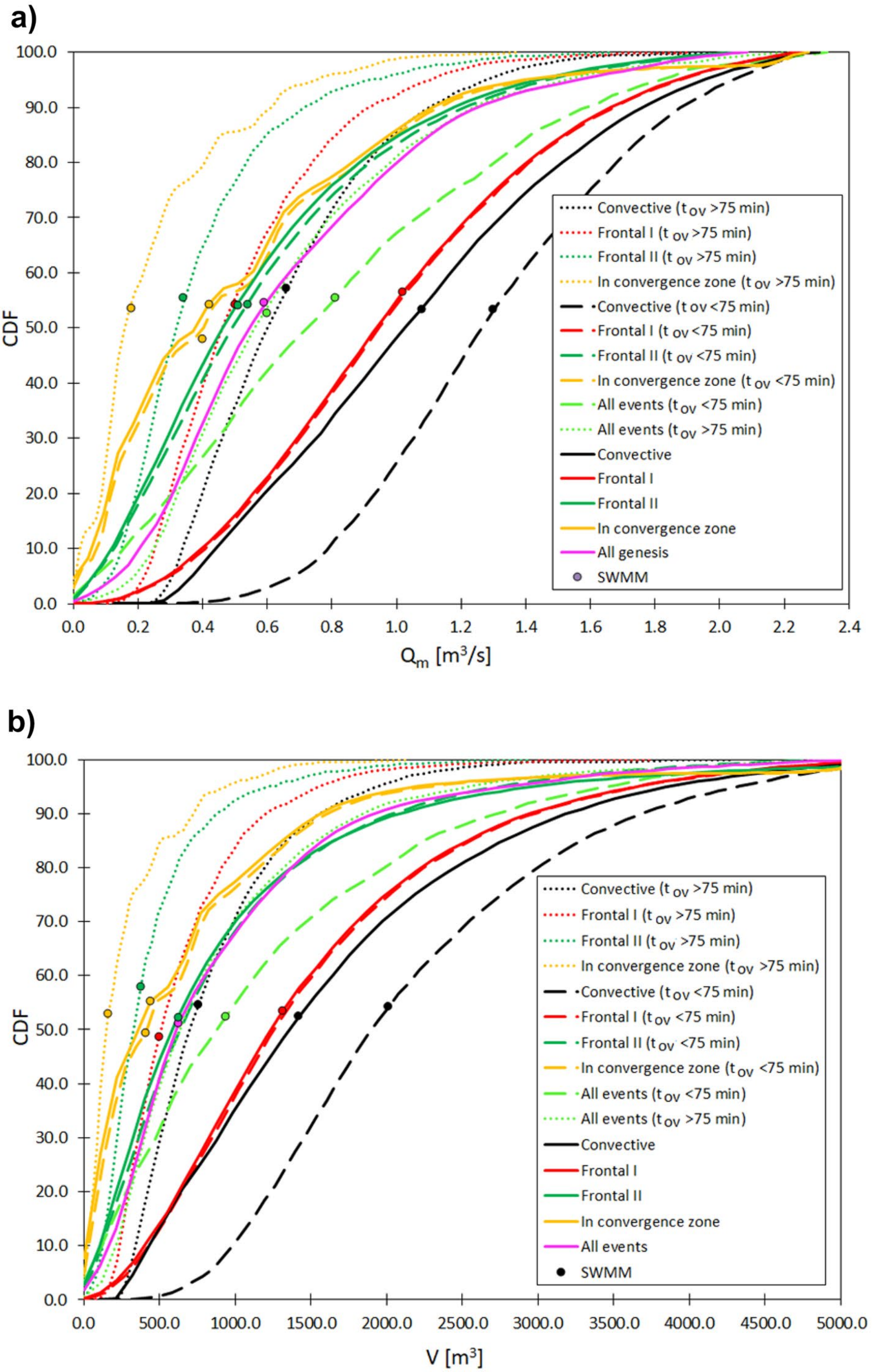

Fig. 5 Empirical distributions describing the probability of exceeding the maximum instantaneous flows – Q_m (**a**) and storm overflow volume – V (**b**) depending on the duration of storm overflow and genesis of the rainfall

bottom destabilization, and they may also have a negative impact on the living conditions of aquatic plants, invertebrates, and fish.

Taking into account the negative impact of storm overflows on the environment and the obtained results of calculations (Figs. 4 and 5), it seems advisable to modify the methodology of designing the height of the storm overflow edge and optimize its length, which should reduce the volume and maximum instantaneous flow directly to the receiver. However, this requires taking into account in the calculations other factors, such as the genesis of rainfall. To confirm this, further detailed analyses on urban catchments with physical and geographical characteristics other than those considered in the paper should be carried out.

4 Conclusions

The calculations and simulations presented in this paper showed that the MLR model can be used to model the duration of storm overflow. The conducted analyses showed that the duration of storm overflow is influenced by the total and maximum 30-min rainfall depth in a rainfall event, and the rainfall duration.

It was found that the genesis of rainfall has a significant impact on the parameters of the storm overflow (duration, volume, and maximum instantaneous flow). The highest storm overflow volume and maximum instantaneous flow were obtained for convective rainfall, while the lowest for low rainfall. Moreover, omitting the genesis of rainfall leads to the underestimation of the modeled maximum instantaneous flows and the storm overflow volume, compared with the calculation results obtained for convective rainfall in the air mass and frontal rainfall related to the cold front (frontal I). Higher values of the modeled storm overflow volume and maximum instantaneous flows were obtained for duration of storm overflow shorter than 75 min than those longer than 75 min.

It is advisable to take into account the genesis of rainfall when designing the edge of the storm overflow, which result in a reduction in the storm overflow volume, maximum instantaneous flow and have a significant effect on the receiver conditions. Considering the large diversity of parameters of storm overflow performance caused by rainfall of various genesis, further analyses are recommended to develop early warning systems against threats, such as overflows of reservoirs and flooding, in small urban catchments.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11269-022-03223-z.

Acknowledgements The authors would like to thank the Polish Institute of Meteorology and Water Management (IMGW-PIB) in Warsaw for access to precipitation data.

Authors' Contributions Conceptualization: B.S and G.M.; methodology: B.S., R.S., and F.D.P; formal analysis: B.S., R.S., writing—original draft preparation: B.S., P.K., G.M., R.S., A.B., K.B., E.A., W.R-K. and F.D.P.; visualization: B.S., P.K., R.S. and E.A.; supervision: B.S., P.K., G.M., and F.D.P. All authors have read and agreed to the published version of the manuscript.

Availability of Data and Material Materials and data used in the present paper are available under request to the corresponding author.

Declarations

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Consent to Participate Not applicable.

Consent to Publish All authors give their permission.

Conflicts of Interest The authors declare that they have no competing interest.

References

Adams BJ, Papa F (2000) Urban stormwater management planning with analytical probabilistic models. John Wiley & Sons Ltd

Aryal SK, Ashbolt S, McIntos BS, Petrone KP, Maheepala S, Chowdhur RK, Gardener T, Gardiner R (2016) Assessing and mitigating the hydrological impacts of urbanisation in semi-urban catchments using the storm water management model. Water Resour Manag 30:5437–5454. https://doi.org/10.1007/s11269-016-1499-z

Chen Z, Han S, Zhou FY, Wang K (2013) A CFD modeling approach for municipal sewer system design optimization to minimize emissions into receiving water body. Water Resour Manag 27:2053–2069. https://doi.org/10.1007/s11269-013-0272-9

Ciupa T (2009) The impact of land use runoff and fluvial transport in small river catchments: based on the Sufraganiec and Silnica Rivers (Kielce, Poland). Jan Kochanowski University, Kielce (in Polish)

DWA-A 118E (2006) Hydraulic dimensioning and verification of drain and sewer systems. German Association for Water, Wastewater and Waste, Hennef, Germany

Fortier C, Mailhot A (2015) Climate change impact on combined sewer overflows. J Water Resour Plan Manag 141:468. https://doi.org/10.1061/(ASCE)WR.1943-5452.0000468

Fu G, Butler D, Kh S, Sun S (2011) Imprecise probabilistic evaluation of sewer flooding in urban drainage systems using random set theory. Water Resour Res 47:8944. https://doi.org/10.1029/2009WR008944

Górska K (2012) Zmienność ładunków zanieczyszczeń w ściekach deszczowych na przykładzie wybranej zlewni. Kielce University of Technology, Kielce (in Polish)

Grum M, Aalderink H (1999) Uncertainty in return period analysis of combined sewer overflow effects using embedded Monte Carlo simulations. Water Sci Technol 39:233–240. https://doi.org/10.1016/S0273-1223(99)00063-3

Hosmer DW, Lemeshow S, Sturdivant RX (2013) Applied logistic regression (3rd ed.). John Wiley & Sons

Hyun JY, Rockaway TD, French MN (2016) Urban sewershed overflow analysis using super-resolution weather radar rainfall. Hydrol Earth Syst Sci Discuss. https://doi.org/10.5194/hess-2016-362

Iman RL, Conover WJA (1982) A distribution-free approach to inducing rank correlation among input variables. Commun Stat Simul Comput 11:311–334. https://doi.org/10.1080/03610918208812265

Jafari F, Mousavi J, Kim JH (2020) Investigation of rainfall forecast system characteristics in real-time optimal operation of urban drainage systems. Water Resour Manag 34:1773–1787. https://doi.org/10.1007/s11269-020-02528-1

Jean MÈ, Duchesne S, Pelletier G, Pleau M (2018) Selection of rainfall information as input data for the design of combined sewer overflow solutions. J Hydrol 565:559–569. https://doi.org/10.1016/j.jhydrol.2018.08.064

Montserrat A, Bosch L, Kiser MA, Poch M, Corominas L (2015) Using data from monitoring combined sewer overflows to assess, improve, and maintain combined sewer systems. Sci Total Environ 505:1053–1061. https://doi.org/10.1016/j.scitotenv.2014.10.087

Szeląg B, Chmielowski K, Dacewicz E (2018) Simulation of a storm overflow with probabilistic and hydrodynamic models. Urban Water J 15:662–670. https://doi.org/10.1080/1573062X.2018.1536759

Szeląg B, Suligowski R, Studziński J, De Paola F (2020) Application of logistic regression to simulate the influence of rainfall genesis on storm overflow operations: a probabilistic approach. Hydrol Earth System Sci 24:595–614. https://doi.org/10.5194/hess-24-595-2020

Thorndahl S, Willems P (2008) Probabilistic modelling of overflow, surcharge and flooding in urban drainage using the first-order reliability method and parameterization of local rain series. Water Res 42:455–466. https://doi.org/10.1016/j.watres.2007.07.038

US EPA (1995) Combined sewer overflows. Guidance for nine minimum controls. Office of Wastewater Management, Washington, D.C

Wei Z, Sun H, Xu H, Wu G, Xie W (2019) The effects of rainfall regimes and rainfall characteristics on peak discharge in a small debris flow-prone catchment. J Mt Sci 16:1646–1660. https://doi.org/10.1007/s11629-018-5260-3

Wu FC, Tsang YP (2004) Second-order Monte Carlo uncertainty/variability analysis using correlated model parameters: application to salmonid embryo survival risk assessment. Ecol Modell 177:393–414. https://doi.org/10.1016/j.ecolmodel.2004.02.016

Wu Z, Ma B, Wang H, Hu C, Lv H (2021) Identifcation of sensitive parameters of urban flood model based on artifcial neural network. Water Resour Manag 35:2115–2128. https://doi.org/10.1007/s11269-021-02825-3

Yang L, Li J, Kang A, Li S, Feng P (2020) The effect of nonstationarity in rainfall on urban flooding based on coupling SWMM and MIKE21. Water Resour Manag 34:1535–1551. https://doi.org/10.1007/s11269-020-02522-7

Yao S, Chen N, Du W, Wang C, Chen C (2021) A cellular automata based rainfall-runof model for urban inundation analysis under diferent land uses. Water Resour Manag 35:1991–2006. https://doi.org/10.1007/s11269-021-02826-2

Zabel T, Milne I, Mckay G (2001) Approaches adopted by the European Union and selected Member States for the control of urban pollution. Urban Water J 3:25–32. https://doi.org/10.1016/S1462-0758(01)00019-X