

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

Bartosz Szela

Kielce University of Technology: Politechnika Swietokrzyska

Roman Suligowski

Jan Kochanowski University in Kielce: Uniwersytet Jana Kochanowskiego w Kielcach

Grzegorz Majewski (✉ grzegorz_majewski@sggw.edu.pl)

Warsaw University of Life Sciences: Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

<https://orcid.org/0000-0002-0122-1409>

Adrian Bralewski

Szkoła Główna Służby Pożarniczej

Karolina Bralewska

Szkoła Główna Służby Pożarniczej

Ewa Anioł

Warsaw University of Life Sciences: Szkoła Główna Gospodarstwa Wiejskiego w Warszawie

Wioletta Rogula-Kozłowska

Szkoła Główna Służby Pożarniczej

Francesco De Paola

University of Naples Federico II: Universita degli Studi di Napoli Federico II

Research Article

Keywords: Storm overflow, Multinomial logistic regression, Rainfall genesis, Iman–Conover method

Posted Date: March 3rd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1398701/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Application of multinomial logistic regression to model the impact of rainfall genesis on the performance of storm overflows: case study

Bartosz Szela¹, Roman Suligowski², Grzegorz Majewski^{3*}, Adrian Bralewski⁴, Karolina Bralewska⁴, Ewa Anioł³, Wioletta Rogula-Kozłowska⁴, Francesco De Paola⁵

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

² Jan Kochanowski University, 5 Źeromskiego St., 25-406 Kielce, Poland

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

⁴ The Main School of Fire Service, 52/54 Słowackiego St., 01-629 Warsaw, Poland

⁵ University of Naples Federico II, via Claudio 21, Naples 80125, Italy

*Correspondence: grzegorz_majewski@sggw.edu.pl

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 to model objects located in a stormwater system, was used to simulate the duration of storm overflow. The Iman–Conover method, involving theoretical cumulative distributions determined based on 30-year rainfall sequences, was adopted to simulate the rainfall characteristics describing the overflow performance (total and maximum 30-min rainfall depth and rainfall duration). Simulations showed a significant impact of rainfall genesis on the parameters of the storm overflow. The model presented in this work and the results obtained can be used at the dimensioning stage of 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: suspension, heavy metals, pesticides, polycyclic aromatic hydrocarbons (PAHs). A significant volume of stormwater during high-intensity rainfall leads to an increase in the flow rate in receivers. Consequently, it may lead to disturbances in the stability of the watercourse bed, erosion of its bottom and banks, and flooding in its vicinity. To limit or eliminate the adverse impact of storm overflows on the receiver, it is necessary to select the optimal height and length of the overflow edge. In existing stormwater systems, this solution may be difficult to implement; therefore, it is often necessary to reconstruct it, for example, by building retention reservoirs, using devices to regulate the flow of stormwater, expanding systems for limiting the volume of storm overflow (e.g., rain gardens, green roofs, rainwater barrels, drainage ditches).

The basis for designing 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; Szeląg et al., 2020). Verification of the correctness of the overflow edge location is performed by analyzing its performance, as well as determining the annual number of storm overflow, their volume, and the maximum instantaneous flow, over a period of one or several years (DWA-A 118E, 2006; Kaźmierczak, 2013; Szeląg et al., 2016; Brzezińska, 2016). To simulate the performance of storm overflows in catchment areas, hydrodynamic models are usually used (DWA-A 117, 2006; Licznar, 2013; Strzec et al., 2018). Ensuring the reliability of the results obtained with the model requires the collection of detailed information about the catchment area, as well as data on the rainfall depth and flows. These data are the basis for model calibration. The use of the aforementioned models makes it possible to analyze the impact of modernization of the stormwater system on the conditions of its functioning and the optimal selection of solutions. However, the sole disadvantage associated 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. As such, statistical models are increasingly used to simulate stormwater systems and their components (including storm overflows) (Zoppou, 2001; Psarrou et al., 2018; Młyński et al., 2019; Szeląg et al., 2020). An existing literature review (Suárez and Puertas, 2005; Li et al., 2015; Me et al., 2017) shows that the use of statistical models in relation to storm overflows makes it possible to model the

pollutant load (chemical oxygen demand – COD, biological oxygen demand – BOD, ammonium nitrogen – NH₄-N, total suspended solids – TSS, etc.), the discharge volume, and the maximum instantaneous flow (based on rainfall data and catchment characteristics). Thorndahl and Willems (2008) and Grum and Aalderink (1999) developed models to simulate the volume of overflow in small urban catchments. Using the first-order reliability method (FORM), Thorndahl and Willems (2008) also developed a model to predict the storm overflow, while in the second case, the analyses also included the quality of stormwater. A number of statistical models to simulate outflow from the catchment area, the functioning of the retention reservoir, and the low-impact development system were also developed by Wang and Guo (2018). The modeling results obtained by these authors may be used when making decisions on the modernization of existing systems or when building models to assess the impact of local modifications in the catchment on the performance of stormwater system. Among the statistical models simulating the storm overflow performance, there are few studies in which the duration of overflow is analyzed. The results of the research by Montserrat et al. (2015) showed a strong non-linear relationship between rainfall depth and duration of overflow. However, the model they obtained does not consider all the variables that have a key impact on the performance of the storm overflow. The obtained value of the coefficient of determination (R^2) varied in the range of 0.31–0.86, which confirmed that it is necessary to conduct more detailed studies. Similar relationships were obtained by Fortier and Maihlot (2015), who analyzed the operation of 30 combined sewer overflow (CSO) in the Quebec catchment (Canada). Despite numerous scientific works on storm overflows and attempts to identify rainfall for their design with the use of hydrodynamic models (DWA-A 117, 2006; Licznar, 2013; Strzec et al., 2018), the strict relationship between the type of rainfall (expressed in its genesis) and the conditions of the activity of the storm overflow have not yet been established. From a practical point-of-view, 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. Taking into account local conditions (DWA-M 180E), it is not always advisable to create an extensive system for simulating and controlling the performance of the stormwater system from the beginning. In these systems, existing measurement or computational systems should be used to predict the performance of stormwater network elements based on weather forecasts obtained using regional numerical models (e.g., UM and WRF), which are well

known and work adequately in practice. The analysis by Szelağ et al. (2020) demonstrated that knowledge of the genetic type of rainfall can be useful for identifying the storm overflow.

Herein, we present an innovative mathematical model that enables the analysis of the influence of meteorological conditions (rainfall characteristics, taking into account their genesis) on the parameters of storm overflow performance (volume of storm overflow, its duration, and maximum instantaneous flow). The multinomial logistic regression model was used to simulate the duration of storm overflow, which so far has not been used to model the stormwater system. A synthetic rainfall generator was used to simulate meteorological conditions, including the diverse genesis of rainfall (convective, frontal, or rainfall in the convergence zone). It is based on theoretical density distributions determined from a 45-year rainfall series. The study also related the impact of rainfall genesis and duration of storm overflow on the volume storm overflow and the maximum instantaneous flow. This solution made it possible to easily identify the variability of the overflow parameters, without the need to perform long-term simulations using a hydrodynamic model.

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 southwestern part of the city of Kielce, with a population of 200,000. The average population density in the analyzed catchment is 21.4 people·ha⁻¹. There are housing estates, public utility buildings, and communication areas in the area studied. The catchment is dominated by urban green areas covering 47.2% of its area. Roads with parking lots and pavements cover 37.3% of the area and roofs of buildings cover 14.3%. The highest point of the catchment area is 271.2 m above sea level, while the lowest is at 260.0 m above sea level. The average decline in the area is 7.1%. The main sewer is 1,569 m long and its diameter varies from 600 mm to 1,250 mm (at the outlet).

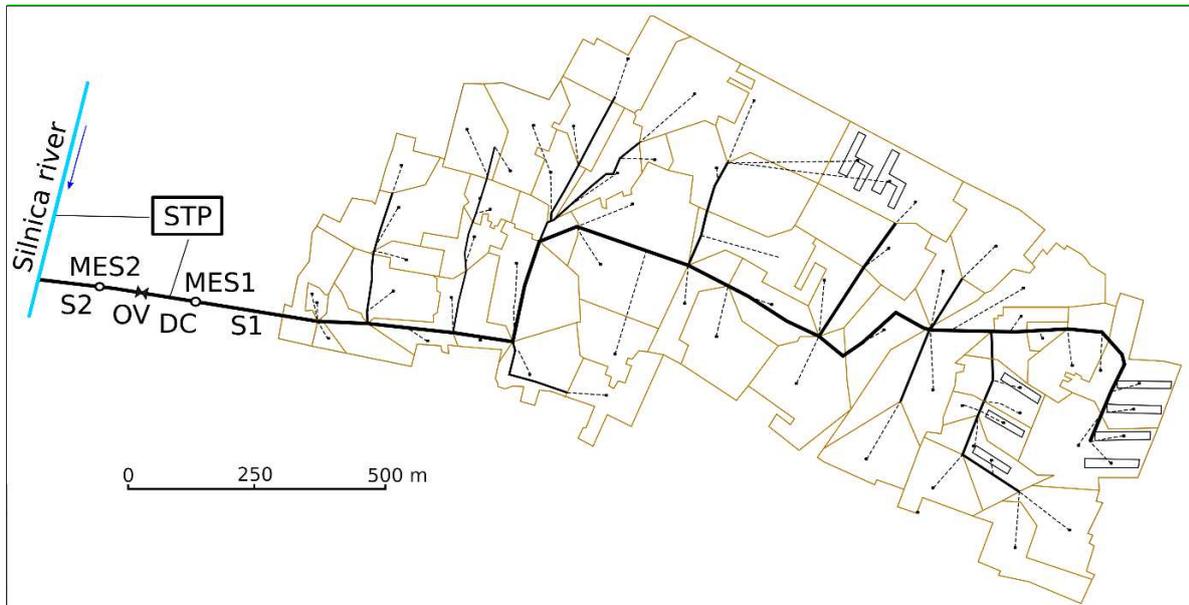


Fig. 1 Diagram of the urban catchment area under study.

The annual rainfall depth varies in the range of 537–757 mm. The number of days with rainfall range from 155 to 266, while those with snow range from 36 to 84. The average annual air temperature varies from 8.1 °C to 9.6 °C. Stormwater outflow from the analyzed catchment area to the Silnica River via the S1 collector. Before entering the river, it is brought to the stormwater treatment plant (STP) in the amount of $Q = 200 \text{ dm}^3 \cdot \text{s}^{-1}$, where it is cleaned of suspensions. When the level of stormwater in the diversion chamber (DC) exceeds the height of 0.42 m, stormwater is discharged through the overflow (OV) to the discharge channel (S2) through which the stormwater flows to the receiver. An MES1 flowmeter is installed at a distance of approximately 3 m from the S1 channel inlet to the separation chamber, which measures the flow values with a resolution of 1 min during intense rainfall events. Continuous monitoring of the amount of stormwater outflowing of the catchment area showed that in dry weather the amount of stormwater outflowing of the catchment area is $0.001\text{--}0.011 \text{ m}^3 \cdot \text{s}^{-1}$. 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 rainfall measurements have been carried out since 2008 with a resolution of 1 min for rainfall events.

2.2 Rainfall data

Rainfall data was obtained from pluvial observations conducted at the meteorological station in Kielce, which is part of the network of the Polish Institute of Meteorology and

Water Management-National Research Institute (IMGW-PIB). This information in the form of rainfall depth and its duration was recorded on paper tape strips using a traditional float pluviograph in the warm half-year (from May to October) from 1961 to 2005. Their analysis shows that during this period, there were 3,524 rainfall events with a rainfall depth of more than 0.3 mm, which means an average of 78.3 episodes per year. The analysis of the surface synoptic charts of Europe and calendar describing the types of atmospheric circulation together with air masses and atmospheric fronts in southern Poland (Niedźwiedź, 2019) indicates that convective air mass rainfall (lasting less than 150 minutes) was the most common (57 times on average) in Kielce. The convective air mass rainfall did not exceed a significant rainfall depth (max. 40.5 mm); however, it was characterized by high average intensity (up to $0.8 \text{ mm} \cdot \text{min}^{-1}$). Frontal rainfall occurred an average of 18 times between May and October. The dynamic rainfall processes taking place in the cold front zone resulted in a high average intensity of rainfall (max. $0.32 \text{ mm} \cdot \text{min}^{-1}$). In contrast, the processes occurring in the warm front zone caused rainfall with a low average intensity, but a high total sum (max. 73.9 mm). Rainfall in the convergence zone, related to deep low-pressure centers (with atmospheric fronts), was characterized by a long duration (more than 630 min), significantly lower frequency (on average two events in half a year) but generated significant total rainfall depth (up to 155.2 mm).

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 of storm overflow, volume, maximum instantaneous flow) in single rainfall events. Fig.2 shows the flow-chart of the methodology. The study also analyzed the impact of interaction between the duration of the storm overflow and the genesis of rainfall on other parameters of the overflow functioning. This aspect has not been considered so far, despite the fact that, in the context of the analyses performed by many researchers, it might be of significant importance at the phase of stormwater overflow dimensioning.

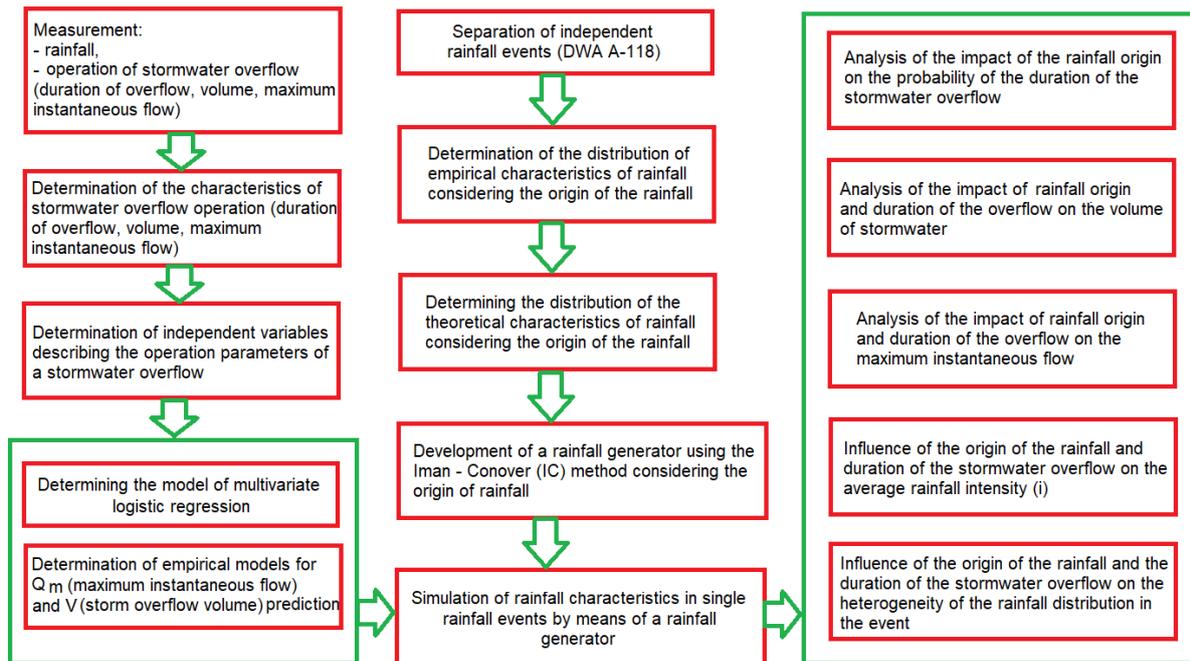


Fig. 2 Algorithm for determining the parameters of the overflow performance, including the genesis of rainfall and duration of storm overflow

The following rainfall types were distinguished in the adopted approach – convective in air mass, frontal (related to the cold front zone and warm front), and rainfall in the convergence zone. The Monte Carlo (MC) method with Iman–Conover (1982) modification was used to simulate rainfall characteristics, taking into account their genesis. This enabled the simultaneous modeling of many dependent variables, that is, correlated variables. The authorial (original) solution, based on multinomial logistic regression, was used in the model to simulate the duration of storm overflow. Due to the complex relationships between the duration of storm overflow and rainfall characteristics, it was reasonable to introduce classification variables describing three main conditions, when no storm overflow occurred, the short and long-term of storm overflow. The assumed method made it possible to analyze the varied impact of rainfall characteristics on the duration of storm overflow, taking into account the variability of its numerical values. The model of multinomial logistic regression adopted in this work is an innovative approach to modeling the operation of facilities/elements of stormwater systems. In the developed model, the influence of rainfall genesis on the duration of storm overflow in a single rainfall event was determined. Simultaneously, the influence of rainfall genesis and duration of storm overflow on the performance parameters of the storm overflow (volume and maximum instantaneous flow) was analyzed. In the proposed approach, the variability of rainfall characteristics was

identified along with the duration of storm overflow. On the one hand, the presented solution is a simplification of the methodology of modeling the duration of storm overflow (hydrodynamic models); on the other hand, it makes it possible to build an expert system that allows assessment of the performance of a storm overflow only based on the established genesis of rainfall. Additionally, this solution enables a typically quantitative analysis of the transfer performance based on qualitative data, i.e., genetic type of rainfall.

Separation of independent rainfall events

Guidelines DWA A-118 (2006) were used to distinguish independent rainfall events in rainfall time series. In the assumed 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 identifying the rainfall event. The adopted solution was used in the works of other authors who dealt with the modeling of stormwater flooding (Fu et al., 2011; Astaraie-Imani et al., 2013), the performance of storm overflows (Szeląg et al., 2020), or the operation of retention reservoirs (Balistrocchi et al., 2009).

2.4 Rainfall simulation

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-minute 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; Brzezińska et al., 2016); moreover, in the case of the duration of discharge, the dependency is more complex, as confirmed by previous works (Szeląg, 2016; Wei et al., 2019).

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), empirical cumulative distribution functions were determined for the separated rainfall events. The cumulative distribution function was determined assuming the division of rainfall data by taking into account the genesis of rainfall, that is, 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 (Ghosh, 2010; Härdle and Simar, 2012; Del Giudice et al., 2015) were adjusted to these data. To obtain the best fit between the theoretical

and empirical data, several distributions were tested: log-normal, gamma, Weibull, Frechet, Gumbel (Fu and Butler 2014, Balistocchi and Grossi 2020). The χ^2 and Kolmogorov–Smirnov (KS) tests were used to assess the compliance of the empirical and theoretical distribution (Adams and Papa, 2000; Guo and Dai , 2009; Fu et al., 2013).

Modeling of synthetic rainfall series

Monte-Carlo method (MC) were used to model the rainfall characteristics in separate rainfall events constituting synthetic time series. The analyzed variables, which are the basis for the simulation of the operation of storm overflow, can be correlated, as confirmed by many publications (e.g., Wu and Tsang, 2004); therefore, the Iman-Conover (IC) algorithm (Iman, Conover, 1982) was used in the computational experiments. In this method, the variability of the analyzed variables was expressed by theoretical marginal distributions; the measure of this dependence is the Spearman correlation coefficient. The obtained results of calculations using the IC method can be considered correct if the following conditions are met:

- a) Mean values $(\mu_1, \mu_2, \dots, \mu_i)_s$ and standard deviations $(\sigma_1, \sigma_2, \dots, \sigma_i)_s$ of the analyzed variables (x_i) in j – samples do not differ by more than 5%.
- b) Theoretical distributions of x_i variables obtained from calculations should be consistent with those obtained from measurements. To meet this condition, the Kolmogorov–Smirnov test (KS) is recommended (Iman and Conover, 1982).
- c) The value of the correlation coefficient (R) between the individual dependent variables (x_i) obtained for the MC simulation data does not differ by more than 5% from the R -value obtained for the empirical data.

If conditions (a), (b), and (c) are met, the simulation results of the rainfall characteristics in independent rainfall events constituting the time series can be considered correct. Otherwise, the size of the MC sample should be increased. To reduce the sample size and improve the effectiveness of the IC algorithm, a modification was developed, consisting of the use of the Latin Hypercube (LH) algorithm (Sriwastava et al., 2018) belonging to the layered sampling methods, aimed at improving the "uniformity" of the generated numbers from marginal distributions.

Empirical models for simulating the performance parameters of a storm overflow

To determine the storm overflow volume and maximum instantaneous flow, the measurement data presenting the performance of the storm overflow from 2015–2017 were used. Based on the above measurement results (for separate, independent rainfall events), empirical models for predicting were selected (Szeląg et al., 2018):

- storm overflow volume (V)

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

- maximum instantaneous flow (Q_m)

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

Statistical analyses presented in the works of Szelać et al. (2013, 2018) showed the limited possibility of simulating the duration of storm overflow, which resulted from the low adjustment of the calculation results to the measurements.

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 is identified. In the case of a multinomial model, the simulation results are divided into N classes (also called levels). This model, like the binomial model, is used in economic, social, medical, and natural sciences, to model phenomena occurring in the environment (Ingelmo et al., 2011; Salman and Salem, 2012; Khudair et al., 2019). Despite numerous implementations, it has not been used so far in modeling objects/elements of stormwater networks. In the MLR model, the output data for the operation of the storm overflow can be divided into several groups simultaneously, that is, no occurred 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 method (for example, for three levels, which can be identified with the overflow performance), the basis for determining the model is as follows:

$$\ln \frac{p(Y=0)}{p(Y=1)} = \ln \frac{p(Y=0)}{p(Y=2)} - \ln \frac{p(Y=1)}{p(Y=2)} \quad (3)$$

for which the following conditions are met:

$$\ln \frac{p(Y=0)}{p(Y=2)} = \beta_{10} + \sum_{i=1}^k \beta_{1i} \cdot x_i \quad (4)$$

$$\ln \frac{p(Y=1)}{p(Y=2)} = \beta_{20} + \sum_{i=1}^k \beta_{2i} \cdot x_i \quad (5)$$

$$\ln \frac{p(Y=0)}{p(Y=1)} = \beta_{30} + \sum_{i=1}^k \beta_{3i} \cdot x_i \quad (6)$$

Taking into account the equation (3) depending on (4), (5), (6), the following equations were obtained:

$$p(Y = 0) = \frac{\exp(\beta_{10} + \sum_{i=1}^k \beta_{1i} \cdot x_i)}{1 + \exp(\beta_{10} + \sum_{i=1}^k \beta_{1i} \cdot x_i) + \exp(\beta_{20} + \sum_{i=1}^k \beta_{2i} \cdot x_i)} = \frac{\exp(X_1)}{1 + \exp(X_1) + \exp(X_2)} \quad (7)$$

$$p(Y = 1) = \frac{\exp(\beta_{20} + \sum_{i=1}^k \beta_{2i} \cdot x_i)}{1 + \exp(\beta_{10} + \sum_{i=1}^k \beta_{1i} \cdot x_i) + \exp(\beta_{20} + \sum_{i=1}^k \beta_{2i} \cdot x_i)} = \frac{\exp(X_2)}{1 + \exp(X_1) + \exp(X_2)} \quad (8)$$

$$p(Y = 2) = \frac{1}{1 + \exp(\beta_{10} + \sum_{i=1}^k \beta_{1i} \cdot x_i) + \exp(\beta_{20} + \sum_{i=1}^k \beta_{2i} \cdot x_i)} = \frac{1}{1 + \exp(X_1) + \exp(X_2)} \quad (9)$$

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 $t_{ov} > t_{ov,g}$.

To assess the predictive abilities of the designated model of MLR, 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 varied from 6 to 528 min.

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; their values were determined to be dependent on the genesis of rainfall (convective, 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: "Empirical models for simulating the performance parameters of a storm overflow" and "Multinomial logistic regression" were used. Thus, in the conducted simulations, the influence of rainfall genesis on the probability of the duration of storm overflow was analyzed. Using the determined statistical models, the influence of the rainfall genesis and the duration of storm overflow on the variability of the overflow parameters (V , Q_m) were simultaneously analyzed. The mean values of the aforementioned overflow performance parameters were determined using the following formulas:

- storm overflow volume, $E(V)$

$$E(V) = \int_{P_{t=30(0)}}^{P_{t=30}} V f(P_{t=30})(g, p_{tv,g}(k)) dP_{t=30} \quad (10)$$

- maximum instantaneous flow - $E(Q_m)$

$$E(Q_m) = \int_{P_{t=30(0)}}^{P_{t=30}} Q_m f(P_{t=30})(g, p_{tv,g}(k)) dP_{t=30} \quad (11)$$

- probability of the duration of storm overflow - $E(p_{tv,g})_k$

$$E(p_{tv,g})_k = \int_{P_{1,0}}^{P_1} \int_{P_{2,0}}^{P_2} \dots \int_{P_{n,0}}^{P_n} p_{tv,g}(k) f(P_1, P_2, P_3, t_r \dots P_n, \theta_k) dP_1 dP_2 dP_3 dt_r dP_n \quad (12)$$

where g is a function describing rainfall characteristics for the k^{th} rainfall genesis, $p_{t_{ov},g(k)}$ is the probability function of the duration of storm overflow (storm overflow not occurred, duration of storm overflow in the range $t_{ov} \in (0, t_{ov,g(k)})$ and $t_{ov} > t_{ov,g(k)}$), θ_k a function that determines the relationships between the rainfall characteristics taking into account the genesis of rainfall, and n the rainfall characteristics influencing the parameters of the storm overflow depending on the genesis of rainfall.

Owing to the complex form of equations (10), (11), and (12), as well as the limited possibilities of their analytical solution, the MC method was used in this study, taking into account the IC modification. Additionally, to reduce the number of simulations performed, the LH algorithm was implemented, as discussed in the section "Modeling of synthetic rainfall series". The algorithm for solving the aforementioned stochastic integrals (10), (11), and (12) can be reduced to the following steps:

- N - sampling (5000) of rainfall characteristics in a single event of k^{th} genesis using the IC and LH,
- calculation of the probability of the duration of storm overflow using the obtained MLR models for rainfall characteristics of k^{th} genesis, determination of empirical cumulative distribution frequency (CDF), and appropriate distribution percentiles;
- calculations of the storm overflow volume (V) and the maximum instantaneous flow (Q_m) for the assumed k^{th} genesis of rainfall and assumed variability of the duration of storm overflow, calculations of cumulative distribution frequency (CDF), and determination of percentiles.

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 was 6.0 mm. The Manning roughness coefficient of the pervious areas was $0.025 \text{ m}^{-1/3} \cdot \text{s}$ and $0.25 \text{ m}^{-1/3} \cdot \text{s}$ for the impervious areas. The Manning channel roughness coefficient was equal to $0.018 \text{ m}^{-1/3} \cdot \text{s}$.

The subject model of the catchment area has been used many times to simulate the runoff of stormwater, its quality, and the dimensioning of the retention reservoirs. It was calibrated with the deterministic and probabilistic method using the Generalized Likelihood Uncertainty + Global Sensitivity Analysis (GLUE + GSA) method (Szeląg et al. 2016). The

model of the separation chamber (DC) and stormwater treatment plant discussed in the study was calibrated as described in detail in the works of Szeląg et al. (2018, 2020). Based on the model in question (catchment – diversion chamber – storm overflow), probabilistic models were verified to prediction the annual number of storm overflow, storm overflow volume, and maximum instantaneous flow (Szeląg et al., 2018, 2020).

Verification of calculation results with a mathematical model

To verify the calculation results (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 performed based on a 45-year rainfall series. For this, a calibrated hydrodynamic model of the catchment and storm overflow, developed with the SWMM program, was used.

3 RESULTS

Duration of storm overflow

Based on the obtained measurements, it was found that the average value of the duration of storm overflow 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., 2016; 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 5–23008 m^3 , and the average was 1801 m^3 . As part of the analyses, three levels were introduced to determine the model of a MLR using the measurement data, i.e., storm overflow not occurred ($t_{ov} = 0$), duration of storm overflow equal of $t_{ov} \in (0, 75 \text{ min})$, and duration of storm overflow equal of $t_{ov} > 75 \text{ min}$.

Multinomial logistic regression

Based on the measurement data of rainfall and the performance of the storm overflow from 2015–2017, an MLR model was established. Thus, equations (13, 14) were obtained to determine the probability of: storm overflow not occurred (Eq. 9), duration of storm overflow $t_{ov} \in (0, 75 \text{ min})$ (Eq. 8), and duration of storm overflow $t_{ov} > 75 \text{ min}$ (Eq. 7), described as:

$$X_1 = 0.467(\pm 0.012) \cdot P_t + 1.725(\pm 0.034) \cdot P_{t=30} - 0.0022(\pm 0.0004) \cdot t_r - 8.398(\pm 0.956) \quad (13)$$

$$X_2 = 0.257(\pm 0.018) \cdot P_t + 1.824(\pm 0.024) \cdot P_{t=30} - 0.0001(\pm 0.00002) \cdot t_r - 6.967(\pm 0.426) \quad (14)$$

The calculations showed that in the analyzed time intervals t_{ov} , the rainfall depth had a varied impact on the probability of the duration of storm overflow. $P_{t=20}$ did not have a statistically significant influence on the duration of storm overflow (p-value was greater than 0.05). The obtained models of MLR were characterized by high predictive abilities. In the case of the models described by equations (7), (9), (13), and (14), 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 $t_{ov} \in (0, 75 \text{ min})$, the calculation results were consistent with the measurements for 37 rainfall events (SENS = 92.15%); and for 24 rainfall events when $t_{ov} > 75 \text{ min}$, 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. To confirm the good fit of the results, the model was validated with 10 rainfall events. The performed field measurements showed that in three of them, storm overflow wasn't occurred; identical results were obtained with the calculations. In four subsequent events, the measurements confirmed that the duration of storm overflow $t_{ov} \in (0, 75 \text{ min})$, and the results of calculations using the logit model showed also showed the duration of storm overflow in this range for three episodes. Among the three episodes in which the duration of storm overflow lasted longer than 75 min ($t_{ov} > 75 \text{ min}$), two cases were found to be 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 equations (7), (9), (13), and (14) 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 $t_{ov} \in (0, 75 \text{ min})$ and $t_{ov} > 75 \text{ min}$. The significant influence of the total rainfall depth on the CSO (combined sewer overflow) duration was confirmed by the results of the analyses by Fortier and Maihlot (2015). Based on the results of measurements of 30 CSO, they developed a mathematical model for calculating the duration of a combined sewer overflow. However, the results of the calculations based on the database of 30 overflows showed that the obtained value was $R^2 = 0.10\text{--}0.63$. This may indicate that their model did not include all significant independent variables that influence the duration of storm overflow.

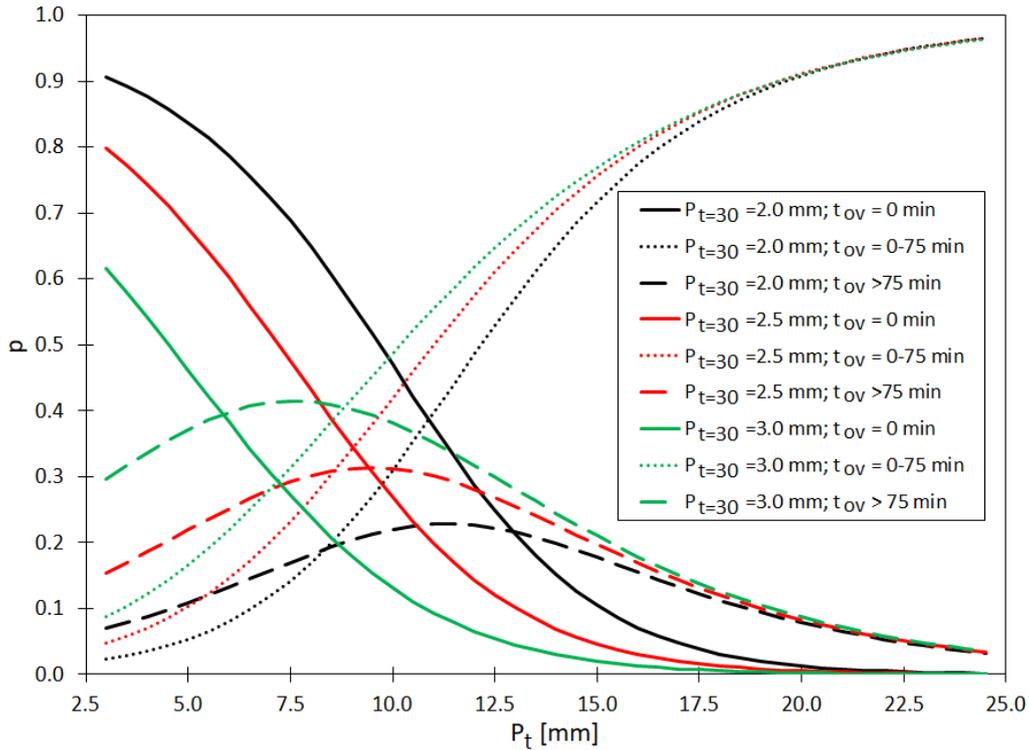


Fig. 3 Impact of total rainfall depth (P_t) and maximum 30-minute rainfall in the event ($P_{t=30}$) on the probability the duration of storm overflow (p).

Regarding the curves illustrating the probability of the duration of storm overflow $t_{ov} \in (0, 75 \text{ min})$, it was found that p values increase with increasing rainfall depth (P_t). Regarding the curves illustrating the probability of the duration of storm overflow $t_{ov} > 75 \text{ min}$, 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 $t_{ov} > 75 \text{ min}$. For example, while $P_{t=30} = 2.0 \text{ mm}$ and $t_{ov} \in (0, 75 \text{ min})$, an increase in rainfall depth from $P_t = 3.0 \text{ mm}$ to $P_t = 11.5 \text{ mm}$ reduces the probability of not occurred storm overflow from 0.91 to 0.35. At the same time, with an identical increase in the rainfall depth when $t_{ov} > 75 \text{ min}$, the probability of the duration of storm overflow increases from 0.08 to 0.21.

When $t_{ov} \in (0, 75 \text{ min})$, the probability of storm overflow increases from 0.04 to 0.45. A further increase in the rainfall depth from $P_t = 11.5 \text{ mm}$ to $P_t = 20.0 \text{ mm}$ reduces the probability of storm overflow not occurred, reduces the probability of the duration of storm overflow $t_{ov} > 75 \text{ min}$ to 0.005 and 0.087, and increases the probability of the duration of storm overflow $t_{ov} \in (0, 75 \text{ min})$ to the value of 0.911. The above results illustrate the variability of the probability of the occurrence of an event 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 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, which is one of the main topics discussed in this paper.

Verification of the MLR model using SWMM

The results obtained using the MLR model were compared with the measurement data and simulation results obtained with a calibrated hydrodynamic model of the catchment area (SWMM). The cases in which there was no storm overflow were compared, and the duration of storm overflow was equal $t_{ov} \in (0, 75 \text{ min})$ and $t_{ov} > 75 \text{ min}$ (Table 1). Based on the data presented in Table 1, it can be concluded that the MLR model is characterized by high predictive abilities. This is confirmed by the high convergence of the measurement and simulation results of the storm overflow in the years 2015–2017.

Table 1 Comparison of calculation results (SWMM, logit-MLR) with measurement data of storm overflow performance from 2015–2017

| Year | Number of events | Measurements | | | SWMM | | | logit (MLR) | | |
|------|------------------|--|----------------------------------|---------------------------|--|----------------------------------|---------------------------|--|----------------------------------|---------------------------|
| | | $t_{ov}=0 \text{ min}$ (no discharge) | $t_{ov} \in (0, 75 \text{ min})$ | $t_{ov} > 75 \text{ min}$ | $t_{ov}=0 \text{ min}$ (no discharge) | $t_{ov} \in (0, 75 \text{ min})$ | $t_{ov} > 75 \text{ min}$ | $t_{ov}=0 \text{ min}$ (no discharge) | $t_{ov} \in (0, 75 \text{ min})$ | $t_{ov} > 75 \text{ min}$ |
| 2015 | 58 | 34 | 15 | 9 | 32 | 14 | 10 | 35 | 15 | 9 |
| 2016 | 44 | 24 | 13 | 7 | 26 | 12 | 8 | 23 | 14 | 6 |
| 2017 | 38 | 18 | 12 | 8 | 19 | 14 | 6 | 17 | 13 | 7 |

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 according to them. The results of the analyses and the obtained parameters of theoretical distributions in relation to rainfall depth (P_r) and its duration (t_r) are given in a previous study (Szeląg et al., 2020). Table 2 lists of the optimal theoretical distributions, their parameters, and the results of the Kolmogorov–Smirnov (KS) and χ^2 tests. The calculation results (p - value) for the other statistical distributions (Frechet, Gumbel, Weibull, gamma) are given in Table 1 (Supplementary material).

Table 2 Results of fitting (p – value for test KS, χ^2) theoretical distributions to empirical data, where frontal I refers to the cold front and frontal II to the warm front

| Variables | Distributions | Parameters | Fitting (KS, χ^2) |
|-----------------------------|---------------|------------------------|---|
| P_t (all events) | Weibull | Szeląg et al. | Szeląg et al. (2020) |
| t_r (all events) | GEV | Szeląg et al. | Szeląg et al. (2020) |
| P_t (convective) | Weibull | Szeląg et al. | Szeląg et al. (2020) |
| t_r (convective) | Beta | Szeląg et al. | Szeląg et al. (2020) |
| P_t (frontal I) | Weibull | Szeląg et al. | Szeląg et al. (2020) |
| t_r (frontal I) | Weibull | Szeląg et al. | Szeląg et al. (2020) |
| P_t (frontal II) | Weibull | Szeląg et al. | Szeląg et al. (2020) |
| t_r (frontal II) | Beta | Szeląg et al. | Szeląg et al. (2020) |
| P_t (in convergence) | Log-normal | Szeląg et al. | Szeląg et al. (2020) |
| t_r (in convergence) | Weibull | Szeląg et al. | Szeląg et al. (2020) |
| $P_{t=30}$ (all events) | Log-normal | $\mu = 1.292 \sigma =$ | $p(\text{K-S}) = 0.173 \quad p(\chi^2) = 0.024$ |
| $P_{t=30}$ (convective) | Log-normal | $\mu = 1.037 \sigma =$ | $p(\text{K-S}) = 0.263 \quad p(\chi^2) = 0.178$ |
| $P_{t=30}$ (frontal I) | Log-normal | $\mu = 0.757 \sigma =$ | $p(\text{K-S}) = 0.513 \quad p(\chi^2) = 0.032$ |
| $P_{t=30}$ (frontal II) | Log-normal | $\mu = 2.202 \sigma =$ | $p(\text{K-S}) = 0.971 \quad p(\chi^2) = 0.713$ |
| $P_{t=30}$ (in convergence) | Log-normal | $\mu = 2.342 \sigma =$ | $p(\text{K-S}) = 0.487 \quad p(\chi^2) = 0.002$ |

The analysis of the data in Table 2 shows that the empirical distributions of $P_{t=30}$ obtained from the measurements are best described by the log-normal distribution given by the following equation:

$$f(x) = \frac{1}{x \cdot \sigma \cdot \sqrt{2 \cdot \pi}} \cdot \exp\left(-\frac{(\ln(x) - \mu)^2}{2 \cdot \sigma^2}\right) \quad (18)$$

The high adjustment of the theoretical distributions to the empirical data allowed for the calculations of the rainfall characteristics in events. This 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).

Impact of rainfall genesis on the probability of duration of storm overflow

Based on theoretical distributions (Table 2) and using the Iman - Conover method, sampling (10000) of rainfall characteristics (P_t , $P_{t=30}$, t_r) was performed. Simulation results were substituted to equations (7), (9), (13), (14) and the probability p in the rainfall event was determined. The simulations included three cases: when no storm overflow occurred ($t_{ov} = 0$), duration of storm overflow of $\in (0, 75 \text{ min})$, and duration of storm overflow of 75 min ($t_{ov} >$

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.

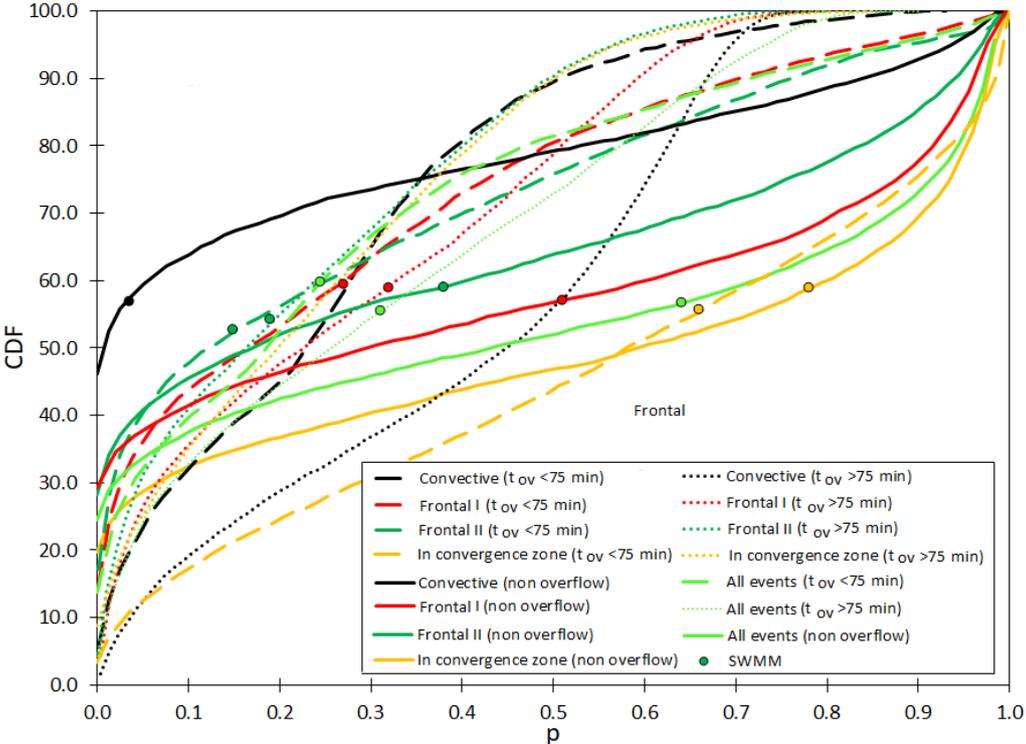


Fig. 4 Impact of rainfall genesis (convective, frontal, and rainfall in the convergence zone) on the probability of storm overflow-duration

Based on the calculations and determined curves (Fig. 4), it can be concluded that in the case when no storm overflow occurred, the highest possibility of exceeding the probability of not occurring of storm overflow for the 50% percentile was obtained for long-term rainfall related to convergence zones (0.58). A clearly lower probability was found for frontal rainfall (frontal I: 0.28 and frontal II: 0.16), and a significantly low probability was obtained for short-term convection rainfall (0.0001). The results of the simulation, in which the genesis of rainfall was omitted, give the probability of storm overflow not occurred as 43% for the 50% percentile. The obtained results are confirmed by the analysis results of Szeląg et al. (2020), who studied the impact of rainfall genesis on the probability of a storm overflow in the catchment area. The calculation results obtained in the two groups, $t_{ov} \in (0, 75 \text{ min})$ and $t_{ov} > 75 \text{ min}$, deserve particular attention. In the first of these cases, for the 50% percentile, the highest probability of the storm overflow occurrence was obtained for the rainfall in the convergence zone (0.58), convection zone (0.24), frontal type I (0.18), and frontal type II (0.13). For calculations where genesis was omitted, the percentile value was 0.19. For storm overflow with durations longer

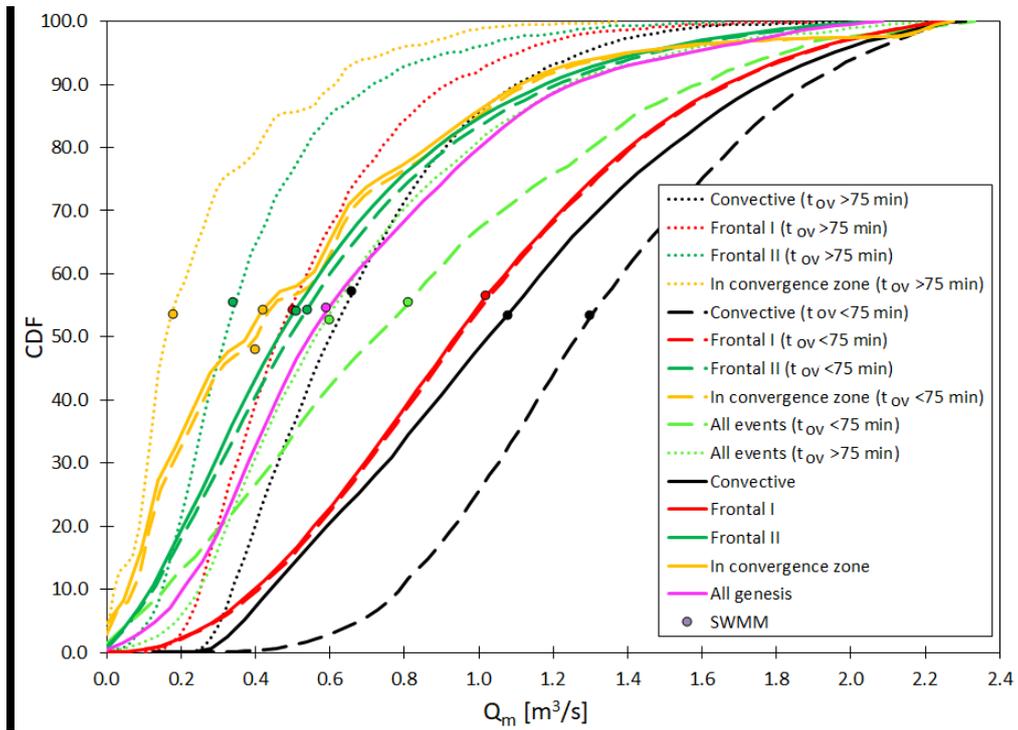
than $t_{ov} > 75$ min, the highest percentile values were obtained for convective rainfall (0.44), rainfall related to the cold front frontal I (0.23), rainfall in the convergence zone (0.19), and rainfall related to the displacement of the warm front frontal II (0.16). The calculation results for duration of storm overflow $t_{ov} > 75$ min, excluding the genesis of rainfall, are insignificantly higher than those obtained for rainfall that occurred in the cold front zone (0.26). The results of the calculations of the probability of the duration of storm overflow (50% percentiles) obtained using the developed mathematical model (Fig. 1) showed a good agreement with those of the calibrated hydrodynamic model of the catchment area. This was confirmed by the obtained curves (Fig. 4). The results obtained in this manner indicate the possibility of using the proposed mathematical model to analyze the impact of rainfall on the probability of storm overflow duration.

The calculations results confirm that considering 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 provides important information about function of the storm overflow for various rainfall conditions that can be used for analyzing the performance of a storm overflow and designing the height and length of its edge. Jean et al. (2016) attempted to select the rainfall characteristics for the design of the overflow. They used a calibrated model of a 410 ha urban catchment area in the province of Québec (Canada) to simulate the performance of the storm overflow. They used three approaches to calculate the performance of storm overflows: continuous simulations, separate independent rainfall events and IDF (intensity-duration-frequency) curves. However, they did not establish explicit results, which could be because rainfall genesis was not included in the analysis.

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 equations (7–12), the maximum instantaneous flow and storm overflow volume were determined. Rainfall genesis and duration of storm overflow were taken into account simultaneously. The calculations results are presented in Figure 5 a, b. 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.

a)



b)

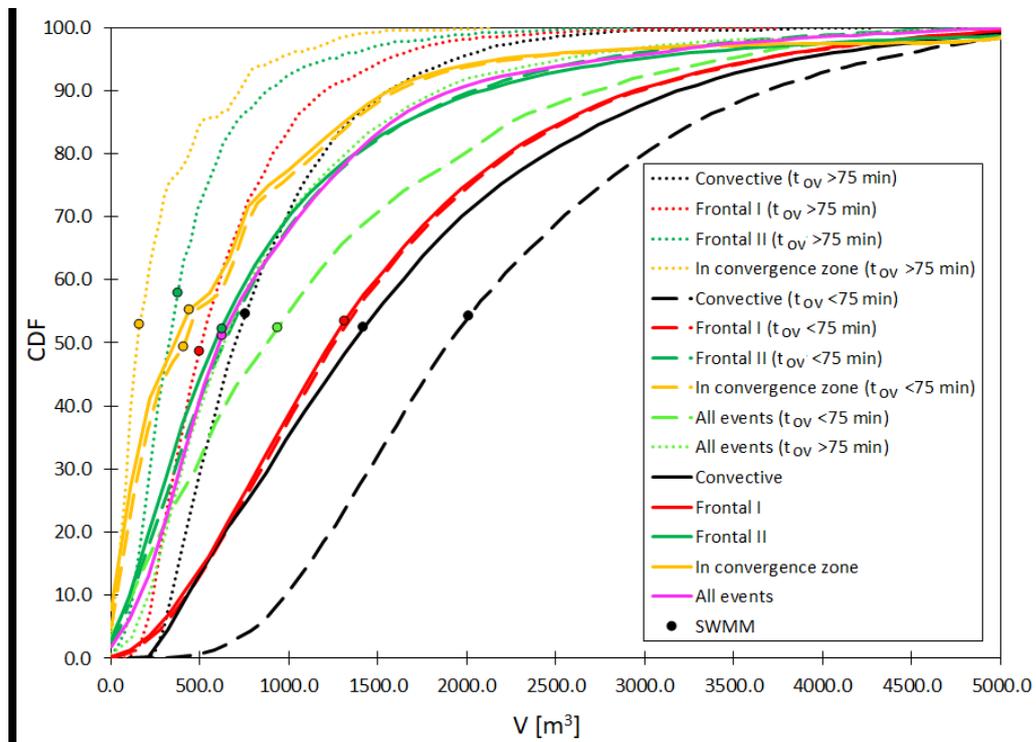


Fig. 5 Empirical distributions describing the probability of exceeding the maximum instantaneous flows – Q_m (a) and storm overflow volume – V (b) depending on the duration of storm overflow and genesis of the rainfall

The maximum instantaneous flow and storm overflow volume obtained for the calculation were $Q_m = 0.55 \text{ m}^3 \cdot \text{s}^{-1}$ and $V = 628 \text{ m}^3$, 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 $Q_m = 1.03 \text{ m}^3 \cdot \text{s}^{-1}$ and $V = 1422 \text{ m}^3$, and the lowest for rainfall in the convergence zones $Q_m = 0.41 \text{ m}^3 \cdot \text{s}^{-1}$ and $V = 442 \text{ m}^3$. The calculation results for the maximum instantaneous flow and storm overflow volume (50% percentile) for convective rainfall and in the cold front zone (frontal I) are greater than their counterparts when the rainfall genesis is omitted. The details of the conditions of the storm overflow operation were obtained by introducing the storm overflow duration into the analyses. 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 $t_{ov} \in (0, 75 \text{ min})$, when only the genesis of rainfall is included, are higher than the simulation results of Q_m and V by 20% and 37%, respectively. Based on the curves (Fig. 5), it can be concluded that the highest values of instantaneous flows and storm overflow (50% percentile) were obtained for convective rainfall and duration of storm overflow shorter than 75 min ($Q_m = 1.23 \text{ m}^3 \cdot \text{s}^{-1}$, $V = 1944 \text{ m}^3$). Frontal rainfall (frontal I) for the same duration of storm overflow has smaller values of the maximum flow ($Q_m = 1.03 \text{ m}^3 \cdot \text{s}^{-1}$) and storm overflow volume ($V = 1341 \text{ m}^3$).

An analysis of the obtained results (Fig. 5a, 5b) it shows that there a number of dependencies between the storm overflow volume (maximum instantaneous flow), taking into account the duration of storm overflow, and the genesis of rainfall and percentiles. For the percentiles below 5%, the values of Q_m and V obtained for convective rainfall and $t_{ov} > 75 \text{ min}$ were higher than those obtained for rainfall in the cold front zone - frontal I ($t_{ov} \in (0, 75 \text{ min})$). The lowest values of the maximum instantaneous flow and storm overflow volume were calculated for rainfall in the convergence zone and rainfall related to the warm front

(frontal II) when the duration of storm overflow was longer than 75 min. For percentiles with values above 32%, the values of Q_m and V obtained for the calculation variant, omitting the genesis of rainfall and the duration of storm overflow less than 75 min, are greater than those for storm overflow resulting from convective rainfall for the same duration of storm overflow. For percentiles above 45%, the Q_m and V values resulting from rainfall in the convergence zone ($t_{ov} \in (0, 75 \text{ min})$) were greater than those determined for rainfall in the warm front zone (frontal II) and for storm overflow shorter than 75 min. For percentiles above 57%, the values of Q_m and V resulting from rainfall in the convergence zone ($t_{ov} \in (0, 75 \text{ min})$) were greater than those established for the rainfall frontal I and a duration of storm overflow than 75 min. The values of Q_m and V (50% percentiles) obtained using the developed mathematical model converge with those obtained using the calibrated hydrodynamic model based on continuous simulations (45-year rainfall series). This is very important for the utility of the proposed mathematical model and its application as an alternative solution to the hydrodynamic model of the catchment area.

By relating the obtained results to the aspects of designing a storm overflow, it can be concluded 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 stratiform rainfall) are different from those defined in Central Europe (Szeląg et al. 2020). Hyun et al. (2016) concentrated on the analysis of storm overflow volume based on super-resolution meteorological radar rainfall data. Their analyses were not aimed at developing relationships between the genesis of rainfall and certain parameters of the overflow performance, therefore they were limited to determining only simple relationships. Though 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 (Aalderink and Grum 1998, 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 Aalderink and Grum (1998), Balistocchi et al. (2020). This leads to an imbalance in the ecosystem of the receiver water (Bashir et al., 2020). High-intensity of rainfall events

(convective, frontal I) may lead to local flooding and bottom destabilization, and they may also have a negative impact on the living conditions of aquatic plants, invertebrates, and fish. Further, they might result in changes in species composition and abundance of aquatic organisms. Both intense rainfall, accompanied by a high concentration of suspended solids in the receiver, and events with a longer duration of storm overflow and a lower concentration of suspended solids are dangerous for the ecological balance of waters and the health condition of fish living in them (Newcombe and Jensen, 1996).

Considering the negative impact of storm overflows on the environment and the obtained results of calculations (Fig. 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 among other factors, the genesis of rainfall in the calculations. To confirm this, further detailed analyses on urban catchments with physical and geographical characteristics other than those considered in the paper should be performed.

Impact of rainfall genesis and duration of storm overflow on rainfall characteristics

Empirical distributions describing the probability of exceeding the mean rainfall intensity (i) and the temporal rainfall distribution in the event ($P_{t=30}/P_t$) were determined. They determine the storm overflow of the appropriate volume and maximum instantaneous flow. The results of the obtained calculations are presented in Figs. 6 and 7. The average rainfall intensity conditioning the storm overflow caused by rainfall of an appropriate genesis (convective, frontal I, frontal II, or in the convergence zone) is lower than that obtained for duration of storm overflow lasting no longer than 75 min; however, it is greater for storm overflows with durations greater than 75 min (Fig. 6). The values of $P_{t=30}/P_t$ obtained for storm overflows caused by convective rainfall, frontal I, frontal II, and rainfall in convergence zones and with a duration of storm overflow longer than 75 min are also greater than those obtained in the simplified variant, that is, excluding the duration of storm overflow (Fig. 7).

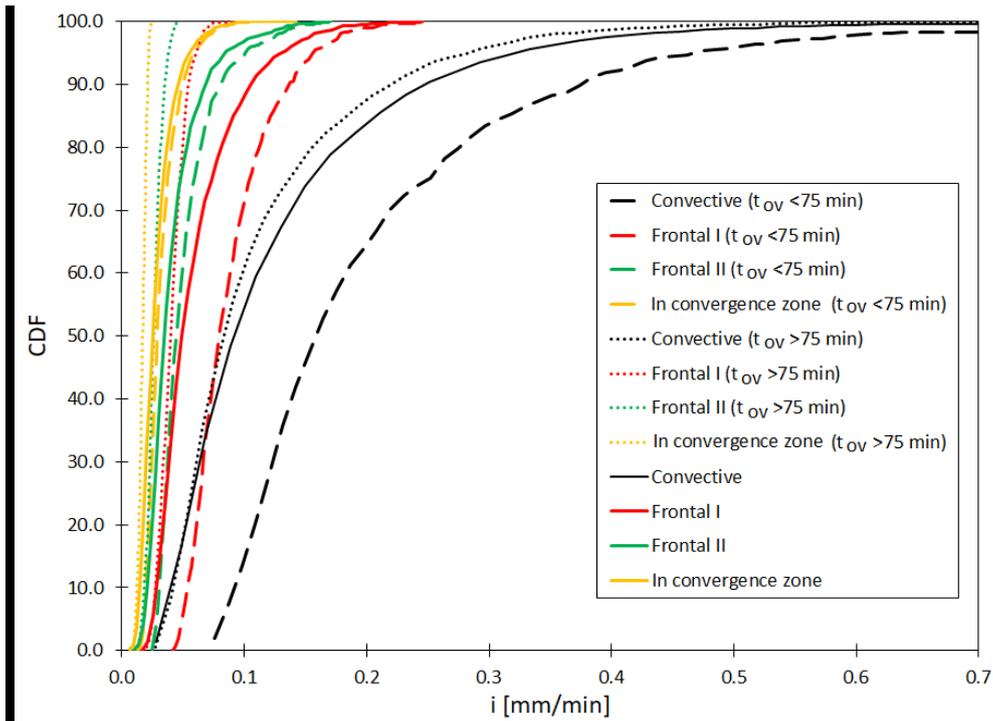


Fig. 6 Empirical distributions describing the probability of exceeding the mean rainfall intensity (i) depending on the duration of storm overflow and rainfall genesis

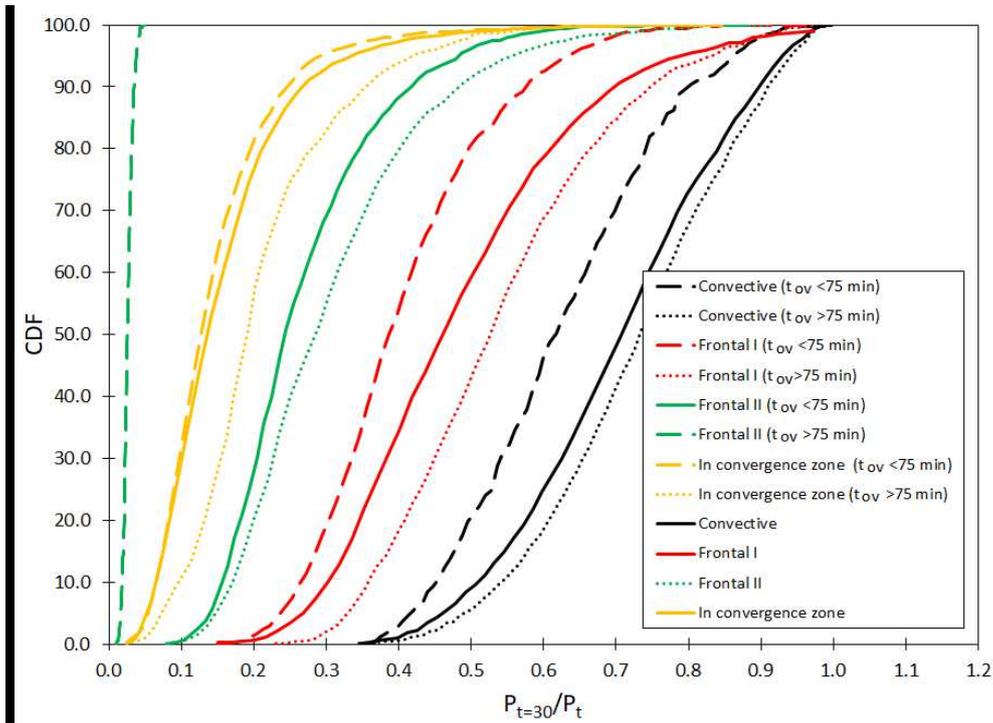


Fig. 7 Empirical distributions describing the probability of exceeding the value of $P_{t=30}/P_t$ (temporal distribution) in a rainfall event depending on the duration of storm overflow and genesis of rainfall

For example, the mean rainfall intensity (50% percentile) determining the storm overflow caused by convective rainfall, with a duration of storm overflow not longer than 75 min ($t_{ov} \in (0, 75 \text{ min})$) is $i = 0.16 \text{ mm} \cdot \text{min}^{-1}$, while the value of this intensity in storm overflow due to convective rainfall is lower ($i = 0.109 \text{ mm} \cdot \text{min}^{-1}$). In contrast, the value of $P_{t=30}/P_t$ obtained for convective rainfall and $t_{ov} \in (0, 75 \text{ min})$ is 0.73, while for the calculation variant without the duration of storm overflow, it is equal to 0.70. Based on the curves (Figs. 5 and 6), it can be concluded that the storm overflow ($t_{ov} \in (0, 75 \text{ min})$) caused by convective rain (50% percentile), conditioning the maximum instantaneous flow $Q_m = 1.26 \text{ m}^3 \cdot \text{s}^{-1}$ and the storm overflow volume $V = 1944 \text{ m}^3$, takes place at an average rainfall intensity of $i = 0.11 \text{ mm} \cdot \text{min}^{-1}$ and $P_{t=30}/P_t = 0.60$. Extending the duration of the storm overflow, that is, above 75 min, the occurrence of convective rainfall leads to a decrease in the average rainfall intensity to the value of $i = 0.09 \text{ mm} \cdot \text{min}^{-1}$. At the same time, for percentiles (storm overflow caused by convective rainfall, $t_{ov} > 75 \text{ min}$) exceeding 45%, the average rainfall intensity that determines the respective values of the maximum instantaneous flow (Q_m) and the storm overflow volume (V) are smaller than those obtained for rainfall frontal I, only if $t_{ov} \in (0, 75 \text{ min})$. When analyzing the overflow performance, the temporal distribution of rainfall should be taken into account. For convective rainfall conditioning of $t_{ov} < 75 \text{ min}$ (50% percentile), the values of $P_{t=30}/P_t$ are 90% higher than that for frontal I rainfall, leading to storm overflow with a duration longer than 75 min. The values of the percentile $p = 50\%$ of the average frontal II intensity and $t_{ov} < 75 \text{ min}$ ($i = 0.0448 \text{ mm} \cdot \text{min}^{-1}$) were 13% higher than those calculated for the rainfall frontal I ($i = 0.0396 \text{ mm} \cdot \text{min}^{-1}$). In the last case, the value of $P_{t=30}/P_t$ (50% percentile) for frontal II rainfall ($t_{ov} \in (0, 75 \text{ min})$) is equal to 0.24 and is lower than that obtained for frontal I rainfall ($P_{t=30}/P_t = 0.52$), only if $t_{ov} > 75 \text{ min}$. The mean rainfall intensity in the convergence zone and the storm overflow with a duration shorter than 75 min is $i = 0.0294 \text{ mm} \cdot \text{min}^{-1}$. This determines the maximum instantaneous flow $Q_m = 0.40 \text{ m}^3 \cdot \text{s}^{-1}$ and a storm overflow volume $V = 480 \text{ m}^3$. This intensity is higher than the value caused by the rainfall frontal II and $t_{ov} > 75 \text{ min}$ ($i = 0.026 \text{ mm} \cdot \text{min}^{-1}$). The results are confirmed by the analyses performed by Montserrat et al. (2015); they showed a strong non-linear relationship between the rainfall and the duration of storm overflow (the curve obtained was similar to the logit curve). However, these authors did not take into account the genesis of rainfall and its temporal distribution in the rainfall event, which implies that the relationships they determined are local and do not include all factors that influence storm overflow performance. The significant impact of the dynamics of rainfall on the duration of storm overflow and

performance parameters of storm overflow was demonstrated by Thorndahl and Willems (2008), who performed long-term simulations of the overflow operation and conducted statistical analyses.

Considering the obtained results and the small number of available research works, in which the rainfall characteristics related to their genesis are taken into account at the stage of stormwater systems modeling, the inclusion of the described aspects in the design of storm overflows seems justified. According to the obtained results, it seems to be possible to identify the operating conditions of storm overflows using the information about the genesis of rainfall. The use of models to forecast meteorological conditions from a long- and short-term perspective makes it possible to determine the basic characteristics of rainfall. Thus, an early warning system against threats in terms of the operation of stormwater networks can be developed without the need to expand the existing system and implement complex computing systems.

4 CONCLUSIONS

The calculations and simulations presented in this paper show that the MLR model can be used to model the duration of storm overflow. The conducted analyses show that the duration of storm overflow is influenced by the total and maximum 30-minute rainfall depth in a rainfall event, and the rainfall duration. Herein, a diversified impact of the above-mentioned rainfall characteristics was demonstrated for the duration of storm overflow longer and shorter than 75 min.

Based on the calculations, 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 convection rainfall, while the lowest for low rainfall. Moreover, the calculations show that 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 and frontal rainfall type I. When modeling the variability of the parameters of the storm overflow, the duration of storm overflow had a significant importance. Based on the calculations, it was found that 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.

Considering the calculation results of this work, it is advisable to take into account the genesis of rainfall when designing the edge of the storm overflow. This can result in a reduction in the storm overflow volume and maximum instantaneous flow. This can also 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. This requires further analysis in catchments with different physical, geographical, and rainfall characteristics.

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

Funding This work was supported by the Ministry of Science and Higher Education of Poland within statutory of funds.

Declarations

Consent to Publish All authors give their permission.

Competing Interests The authors declare no conflict of interest

References

- Adams, B. J., Papa, F., 2000. *Urban stormwater management planning with analytical probabilistic models*. John Wiley & Sons Ltd.
- Astaraie-Imani, M., Kapelan, Z., Butler, D., 2013. Improving the performance of an integrated urbanwastewater system under future climate change and urbanisation scenarios. *Journal of Water and Climate Change*, 4, 232–243. <https://doi.org/10.2166/wcc.2013.078>.
- Balistrocchi, M., Grossi, G., Bacchi, B., 2009. An analytical probabilistic model of the quality efficiency of a sewer tank. *Water Resources Research*, 45, Article 7822. <https://doi.org/10.1029/2009WR007822>.

- Balistrocchi, M., Grossi, G., 2020. Predicting the impact of climate change on urban drainage systems in northwestern Italy by a copula-based approach. *Journal of Hydrology: Regional Studies*, 28, Article 100670. <https://doi.org/10.1016/j.ejrh.2020.100670>.
- Bashir, I., Lone, F. A., Bhat, R. A., Mir, S. A., Dar, Z. A., Dar, S. A., 2020. Concerns and threats of contamination on aquatic ecosystems. In K. Hakeem, R. Bhat, & H. Qadri (Eds.), *Bioremediation and Biotechnology, Sustainable Approaches to Pollution Degradation*. Springer. https://doi.org/10.1007/978-3-030-35691-0_1.
- Brzezińska, A., 2017. Impact assessment of combined sewer overflows regulation on sewage volume discharged into surface waters – a case study. *Ochrona Środowiska*, 39(1), 23–31 (In Polish).
- Brzezińska, A., Zawilski, M., Sakson, G., 2016. Assessment of pollutant load emission from combined sewer overflows based on the online monitoring. *Environmental Monitoring and Assessment*, 188(9), Article 502. <https://doi.org/10.1007/s10661-016-5461-6>.
- Del Giudice, G., Padulano, R., Siciliano, D., 2015. Multivariate probability distribution for sewer system vulnerability assessment under data-limited conditions. *Water Science & Technology*, 73(4), 751–760. <https://doi.org/10.2166/wst.2015.546>.
- DWA-A 117, 2006. *Bemessung von Regenrückhalteräumen*. Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall Taschenbuch (In German).
- DWA-A 118E, 2006. *Hydraulic dimensioning and verification of drain and sewer systems*. German Association for Water, Wastewater and Waste.
- DWA-M 180E, 2005. *Framework for planning of real time control of sewer networks*. German Association for Water, Wastewater and Waste.
- Fortier, C., Mailhot, A., 2015. Climate change impact on combined sewer overflows. *Journal of Water Resources Planning and Management*, 141(5), Article 468. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000468](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000468).
- Fu, G., Butler, D., Khu, S.T., Sun, S., 2011. Imprecise probabilistic evaluation of sewer flooding in urban drainage systems using random set theory. *Water Resources Research*, 47(2), Article 8944. <https://doi.org/10.1029/2009WR008944>.
- Fu, G., Kapelan, Z., 2013. Flood analysis of urban drainage systems: probabilistic dependence structure of rainfall characteristics and fuzzy model parameters. *Journal of Hydroinformatics*, 15(3), 687–699. <https://doi.org/10.2166/hydro.2012.160>.

- Fu, G., Butler, D., 2014. Copula-based frequency analysis of overflow and flooding in urban drainage systems. *Journal of Hydrology*, 510, 49–58. <https://doi.org/10.1016/j.jhydrol.2013.12.006>.
- Ghosh, S., 2010. Modelling bivariate rainfall distribution and generating bivariate correlated rainfall data in neighbouring meteorological subdivisions using copula. *Hydrological Processes*, 24(24), Article 7785. <https://doi.org/10.1002/hyp.7785>.
- Gironas, J., Roesner, L.A., Rossman, L.A., & Davis, J. (2010). A new applications manual for the Storm Water Management Model (SWMM). *Environmental Modelling & Software*, 25(6), 813–814. <https://doi.org/10.1016/j.envsoft.2009.11.009>.
- Grum, M., Aalderink, H., 1999. Uncertainty in return period analysis of combined sewer overflow effects using embedded Monte Carlo simulations. *Water Science and Technology*, 39(4), 233–240. [https://doi.org/10.1016/S0273-1223\(99\)00063-3](https://doi.org/10.1016/S0273-1223(99)00063-3).
- Guo, Y., Dai, J., 2009. Expanded analytical probabilistic stormwater models for use in watershed and master drainage planning. *Canadian Journal of Civil Engineering*, 36(6), 933–943. <https://doi.org/10.1139/L09-037>.
- Härdle, W. K., Simar, L., 2012. *Applied multivariate statistical analysis* (4th ed.). Springer.
- Hosmer, D.W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied Logistic Regression* (3rd ed.). John Wiley & Sons.
- Hyun, J. Y., Rockaway, T. D., French, M. N. 2016., Urban sewershed overflow analysis using super-resolution weather radar rainfall. *Hydrology and Earth System Sciences Discussions*. <https://doi.org/10.5194/hess-2016-362>.
- Iman, R. L., Conover, W. J. A., 1982. A distribution-free approach to inducing rank correlation among input variables. *Communications in Statistics – Simulation and Computation*, 11(3), 311–334. <https://doi.org/10.1080/03610918208812265>.
- Ingelmo, F., Molina, M.J., de Paz, J.M., Visconti, F. 2011. Soil saturated hydraulic conductivity assessment from expert evaluation of field characteristics using an ordered logistic regression model. *Soil and Tillage Research*, 115–116, 27–38. <https://doi.org/10.1016/j.still.2011.06.004>.
- Jean, M.È., Duchesne, S., Pelletier, G., Pleau, M., 2018. Selection of rainfall information as input data for the design of combined sewer overflow solutions. *Journal of Hydrology*, 565, 559–569. <https://doi.org/10.1016/j.jhydrol.2018.08.064>.
- Kaźmierczak, B., 2013. Analysis of number of sewage discharges to the receiver. *Proceedings of ECOpole*, 7(2), 617–625. [https://doi.org/10.2429/proc.2013.7\(2\)081](https://doi.org/10.2429/proc.2013.7(2)081).

- Khudhair, B., Khalid, G., Jbbar, R., 2019. Condition prediction models of deteriorated trunk sewer using multinomial logistic regression and artificial neural network. *International Journal of Civil Engineering and Technology*, 10(1), 93–104.
- Li, D., Wan, J., Ma, Y., Wang, Y., Huang, M., Chen, Y. 2015. Stormwater runoff pollutant loading distributions and their correlation with rainfall and catchment characteristics in a rapidly industrialized city. *PLoS One*, 10(3), Article 118776. <https://doi.org/10.1371/journal.pone.0118776>.
- Licznar, P. 2013., Stormwater reservoir dimensioning based on synthetic rainfall time series. *Ochrona Środowiska*, 35(2), 27–32 (In Polish).
- Löwe, R., Vezzaro, L., Mikkelsen, P. S., Grum, M., Madsen, H., 2016. Probabilistic runoff volume forecasting in risk – based optimization for RTC of urban drainage systems. *Environmental Modelling & Software*, 80, 143–158. <https://doi.org/10.1016/j.envsoft.2016.02.027>.
- Löwe, R., Mair, M., Pedersen, A.N., Kleidorfer, M., Rauch, W., Arnbjerg – Nielsen, K., 2020. Impact of urban development on urban water quality management – Limits of predictability. *Computers, Environment and Urban Systems*, 84, 101546. <https://doi.org/10.1016/j.compenvurbsys.2020.101546>.
- Me, W., Hamilton, D. P., Abell, J. M., 2017. *Simulating discharge and contaminant loads from the Waipa Stream catchment under different irrigation scenarios using the SWAT model*. ERI Report No. 98, Faculty of Science and Engineering, University of Waikato, Hamilton, New Zealand.
- Montserrat, A., Bosch, L., Kiser, M. A, Poch, M., Corominas, L., 2015. Using data from monitoring combined sewer overflows to assess, improve, and maintain combined sewer systems. *Science of The Total Environment*, 505, 1053–1061. <https://doi.org/10.1016/j.scitotenv.2014.10.087>.
- Mustafa, A., Happenstall A., Omrani, H., Saadi, I., Cools, M., Teller, J., 2018. Modelling built-up expansion and densification with multinomial logistic regression, cellular automata and genetic algorithm. *Computers, Environment and Urban Systems*, 67, 147 – 156. <https://doi.org/10.1016/j.compenvurbsys.2017.09.009>.
- Newcombe, C. P., Jensen, J. O., 1996. Channel suspended solids and fisheries: a synthesis for quantitative assessment of risk and impact. *North American Journal of Fisheries Management*, 16(4), 693–727. [https://doi.org/10.1577/1548-8675\(1996\)016<0693:CSSAFA>2.3.CO;2](https://doi.org/10.1577/1548-8675(1996)016<0693:CSSAFA>2.3.CO;2).

- Niedźwiedź, T., 2019. *Catalogue of synoptic situations in the upper Vistula river basin (1873–2018)*. Retrieved from <http://klimat.wnoz.us.edu.pl>. Accessed September 10, 2020.
- Psarrou, E., Tsoukalas, I., Makropoulos, E., 2018. A Monte-Carlo-based method for the optimal placement and operation scheduling of sewer mining units in urban wastewater networks. *Water*, 10, Article 200. <https://doi.org/10.3390/w10020200>.
- Salman, B., Salem, O., 2012. Modeling failure of wastewater collection lines using various section-level regression models. *Journal of Infrastructure Systems*, 18(2), Article 75. [https://doi.org/10.1061/\(ASCE\)IS.1943-555X.0000075](https://doi.org/10.1061/(ASCE)IS.1943-555X.0000075).
- Sriwastava, A. K., Tait, S., Schellart, A., Kroll S., 2018. Quantifying uncertainty in simulation of sewer overflow volume. *Journal of Environmental Engineering*, 144(7), Article 1392. [https://doi.org/10.1061/\(ASCE\)EE.1943-7870.0001392](https://doi.org/10.1061/(ASCE)EE.1943-7870.0001392).
- Starzec, M., Dziopak, J., Słyś, D., Pochwat, K., Kordana, S., 2018. Dimensioning of required volumes of interconnected detention tanks taking into account the direction and speed of rain movement. *Water*, 10(12), Article 1826. <https://doi.org/10.3390/w10121826>.
- Suárez, J., Puertas, J., 2005. Determination of COD, BOD, and suspended solids loads during combined sewer overflow (CSO) events in some combined catchments in Spain. *Ecological Engineering*, 24(3), 199–217. <https://doi.org/10.1016/j.ecoleng.2004.11.005>.
- Szeląg, B., Kiczko, A., Dąbek, L., 2016. Sensitivity and uncertainty analysis of hydrodynamic model (SWMM) for storm water runoff forecasting in an urban basin – A case study. *Ochrona Środowiska*, 38, 15–22.
- Szeląg, B., Chmielowski, K., Dacewicz E., 2018. Simulation of a storm overflow with probabilistic and hydrodynamic models. *Urban Water Journal*, 15(7), 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. *Hydrology and Earth System Sciences*, 24(2), 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 Research*, 42(1–2), 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.

- Wang, J., Guo, J. C. Y., 2018. An analytical stochastic approach for evaluating the performance of combined sewer overflow tanks. *Water Resources Research*, 54(5), 3357–3375. <https://doi.org/10.1029/2017WR022286>.
- 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. *Journal of Mountain Science*, 16, 1646–1660. <https://doi.org/10.1007/s11629-018-5260-3>.
- Wu, F. C., Tsang, Y. P., 2004. Second-order Monte Carlo uncertainty/variability analysis using correlated model parameters: application to salmonid embryo survival risk assessment. *Ecological Modelling*, 177(3–4), 393–414. <https://doi.org/10.1016/j.ecolmodel.2004.02.016>.
- 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*, 3(1-2), 25–32. [https://doi.org/10.1016/S1462-0758\(01\)00019-X](https://doi.org/10.1016/S1462-0758(01)00019-X).
- Zoppou, C., 2001. Review of urban storm water models. *Environmental Modelling & Software*, 16(3), 195–231. [https://doi.org/10.1016/S1364-8152\(00\)00084-0](https://doi.org/10.1016/S1364-8152(00)00084-0).

Supplementary material

Table 1. Summary of the results (p - value) for selected statistical distributions (Gumbel, Weibull, Frechet, Gamma).

| Characteristic | p-test | | | | | | | |
|----------------------------------|--------|----------|---------|----------|---------|----------|--------|----------|
| | Gumbel | | Weibull | | Frechet | | Gamma | |
| | KS | χ^2 | KS | χ^2 | KS | χ^2 | KS | χ^2 |
| $P_{t=30}$ (all) | 0.0152 | 0.0122 | 0.0263 | 0.020 | 0.018 | 0.013 | 0.0207 | 0.0114 |
| $P_{t=30}$ (conv) | 0.0254 | 0.0214 | 0.0244 | 0.0211 | 0.017 | 0.012 | 0.0182 | 0.0102 |
| $P_{t=30}$ (frontal I) | 0.0253 | 0.0222 | 0.0255 | 0.0207 | 0.017 | 0.015 | 0.0192 | 0.0121 |
| $P_{t=30}$ (frontal II) | 0.0233 | 0.0231 | 0.0275 | 0.0200 | 0.016 | 0.012 | 0.0211 | 0.0123 |
| $P_{t=30}$ (in convergence zone) | 0.0273 | 0.0214 | 0.0285 | 0.0213 | 0.016 | 0.011 | 0.0256 | 0.0211 |