Development of Local IDF-formula Using Controlled Random Search Method for Global Optimization

Katarzyna WEINEROWSKA-BORDS

Faculty of Civil and Environmental Engineering, Gdańsk University of Technology, Gdańsk, Poland; e-mail: kwein@pg.gda.pl

A b s t r a c t

The aim of the study is to present the effective and relatively simple empirical approach to rainfall intensity-duration-frequency-formulas development, based on Controlled Random Search (CRS) for global optimization. The approach is mainly dedicated to the cases in which the commonly used IDF-relationships do not provide satisfactory fit between simulations and observations, and more complex formulas with higher number of parameters are advisable. Precipitation data from Gdańsk gauge station were analyzed as the example, with use of peak-over-threshold method and Chomicz scale for rainfall intensity. General forms of the IDF-function were chosen and the parameter calibration with use of CRS algorithm was developed. The compliance of the obtained IDF-formulas with precipitation data and the efficiency of the algorithm were analyzed. The study confirmed the proposed empirical approach may be an interesting alternative for probabilistic ones, especially when IDF-relationship has more complex form and precipitation data do not match "typical" hydrological distributions.

Key words: precipitation; data analysis; IDF formulas; optimization.

1. INTRODUCTION

Intensity-duration-frequency (IDF) analysis of extreme precipitation is an important part of hydrological study for many engineering problems. The re-

sults of such analysis, usually expressed by IDF-relations (IDF-formulas) and/or their graphical form – IDF-curves, enable to determine the so-called "design storms" and thus provide essential "input" information for many practical problems, such as: designing of drainage and irrigation systems, designing of sewer conduits, engineering device dimensioning (*e.g.*, dikes, culverts, *etc*.), flood protection, urban impact analysis, water quality management, and many others (*e.g.*, Arnell 1982, Koutsoyiannis *et al.* 1998, Akan and Houghtalen 2003, Elsebaie 2012, Walesh 1989, and others). Despite the fact that, thanks to the significant development of measurement and computing technology, much more advanced hydrological analyses of more sophisticated nature are available (*e.g.*, Molnar and Burlando 2005, Licznar *et al.* 2011a, b; Pui *et al.* 2012), the IDF-formulas, although "simple" and "basic", still pay an important role in hydrological applications (*e.g.*, Willems 2000, Vaes *et al.* 2001, Svensson *et al.* 2007, Venkata Ramana *et al.* 2008, Endreny and Imbeah 2009, El-Sayed 2011, Elsebaie 2012, Ariff *et al.* 2012, Hailegeorgis *et al.* 2013, *etc.*). They become particularly useful in analyzes concerning urban areas. The specificity of such basins expresses itself mainly in relatively small watershed areas (Schilling 1991) with diversified catchment characteristics and factors increasing rapidness and intensity of direct runoff (such as: low roughness, high share of impermeable surfaces, high slopes, existence of artificial conduits and open channel regulation, *etc*.; McCuen 2005 and others) and in local urban phenomena influencing rainfall characteristics (*e.g.*, urban heat island effect; Lutgens and Tarbuck 2004, Delleur 2003). These latter factors cause relatively large temporal and spatial variability of precipitation, which – in addition to other characteristic features of urban basins – entails a need of its more detailed analysis with higher spatial resolution (Schilling 1991, Mutzner 1991, and others).

The most common purpose of IDF-formulas development is to create the "source" for obtaining representative design rainfall hyetographs, which constitutes essential input information for many engineering problems. In many practical cases, such representative hyetographs take form of the "block rainfalls", very useful for the simplest hydrological runoff estimations, *e.g.*, based on rational method or its modifications (Chow 1964, Nagy 1991, Willems 2000, Akan and Houghtalen 2003, Ben-Zvi 2009, Elsebaie 2012). Furthermore, the IDF-formulas may serve also as a basis for the development of synthetic time-variant design storms (storm patterns), that are essential for more complicated hydrological rainfall-runoff models (of conceptual or hydrodynamic character) and for other studies (Licznar and Łomotowski 2007). Such approach is applied, *i.e.*, to determine synthetic hyetographs with use of Huff's, Yen and Chow's, Chicago, SCS, and other methods (Arnell 1982, Vaes *et al.* 2001, Akan and Houghtalen 2003, Grimaldi and Serinaldi 2006). The example of more sophisticated applica-

tions of IDF-relations is testing and calibrating of rainfall generators (Willems 2000). IDF-formulas may be also helpful in synoptic meteorological studies and hydraulic design.

2. THEORETICAL BACKGROUND

IDF-formulas express the relationship between mean precipitation intensity I [mm/min] (or q [dm^3/(s·ha)]) and its frequency of occurrence c [1/year] (or return period T [years], or exceedence probability p [%]) for different time intervals of rainfall duration t_d [min] (also called "aggregation-levels"). Due to the fact that in most of engineering problems the "extreme" values of rainfall intensity are essential (and thus computationally and designingly authoritative), the development of IDF-formulas requires the analysis of the largest extreme rainfall episodes. The basis of such analysis consists of the events of tempestuous and torrential rainfalls recorded during continuous observations in sufficiently long period. The length of historical rainfall data series must be vast enough to enable statistical recognition of temporal variability of precipitation, both annual and seasonal. The longer the return period T is considered (and thus – the smaller the frequency of rainfall occurrence c), the longer the period of rainfall observations is necessary. Thus, the needed length of the historical rainfall data series is strongly dependent on the aim of calculations (type of engineering problem), but usually not shorter than 10 years. The rainfalls observed over this time should constitute the set of data of independent and homogeneous character. In practice such conditions are very often impossible to fulfill.

2.1 Homogeneity and independence of the sample

Full homogeneity and independence of the sample is very difficult to verify and demonstrate, although there are mathematical tools developed to recognize these features of rainfall data sample, *e.g.*, autocorrelation analysis (Grace and Eagleson 1967, Wenzel and Voorhees 1978, Arnell 1982). Independence of the rainfall episodes is connected to their duration, the duration of rainless intervals and their succession in time. Although it is possible to assess the independence with the use of "impartial criterion" (*e.g.*, correlation factor related to different time-lags), it is a very laborious and time-consuming procedure. Moreover, the choice of this criterion is subjective and if not adequately determined it may lead to negligence of some important extreme rainfall episodes, subconsequently influencing the interpretation of return period and finally leading to mistaken IDF-formulas (*e.g.*, Arnell 1982, Willems 2000). Thus, many authors choose the independency criterion without any special mathematical analysis, by *a priori* establishment of the minimal rainless interval length enabling to consider two "neighboring" episodes as independent (*e.g.*, Willems 2000, Marsalek 1978, Johansen 1979).

The values of such criterion are usually equal to a few hours; however one can find analysis with 1 hour or even 0.5 hour criterion (Arnell 1982).

Homogeneity of the sample is connected to stable (over time) measurement conditions (*e.g.*, localization of gauge stations, measuring equipment, *etc.*), sufficiently long time of observations and the lack of breaks in observations. Schilling (1991) specified the "desirable" (ideal, optimal) characteristics of the set of rainfall data, *i.e.*, the length of the observation period equal to 20 years, temporal resolution of 1 min, spatial resolution of 1 km^2, and lack of breaks in observations. It is obvious that for many practical cases such conditions are impossible to fulfill. Moreover, the probability of totally uninterrupted work of the measuring instruments over several decades is very low. Thus, the above-mentioned author also described the "recommended" features of data set, depending on the purpose of the analysis. For example, for preliminary design of storm water conduits and storm water reservoirs, the 10-year observation period is sufficient and breaks in observations are acceptable. Temporal resolution of 10 min is usually considered as minimal for sewer system design (Vaes *et al.* 2001). For rivers the time-step in rainfall observations may be longer. However, the quality of the analysis is strongly influenced by the quality of data.

2.2 Storm episodes selection

For the purpose of IDF-formula determination, the largest episodes are selected from the whole "population" of recorded precipitation data on the basis of a predetermined criterion separating storms from the "typical" (negligible from an engineering point of view) rainfall episodes. Depending on the assumed method of analyses, two basic approaches can be considered: the "annual maxima series" approach and "partial-duration series" approach (Chow 1964, Arnell 1982, Vukmirović and Petrović 1991, Svensson *et al.* 2007, Reiss and Thomas 2007, Ben-Zvi 2009, and others). For the first one (also called "block maxima method") – a single largest extreme for each calendar year of the observation period is selected, while for the second one – the number of rainfall episodes taken into considerations may exceed the number of years in the period of observation, as it is allowed that several extreme episodes may occur in the same year. This last approach is also called "peak-over-threshold" (POT), as the criterion of selecting the storms from the whole "population" of rainfall events is the defined value of threshold in precipitation depth (corresponding with the considered interval duration t_d), over which the analyzed phase of the episode (of the considered duration equal t_d) is classified as a storm and thus taken into further consideration. The "peak-over-threshold*"* technique is more laborious and thus less often used. However, many authors consider this approach more suitable (*e.g.*, Kisiel *et al.* 1971, Todorovic 1978, Madsen *et al.* 1997a, b, Willems 2000,

Licznar and Łomotowski 2005), as it does not limit the number of storms to the number of years in observation period. This way, years which are "wetter" and with several intense storm episodes are not treated on the same level of importance as the years when no substantial storm occurred. Other advantages of the POT approach include higher accuracy of rainfall intensity prediction for lower return periods and possibility of predicting maximal intensities for return periods shorter than one year (Willems 2000). However, the POT technique requires special attention to rainfall events independency and adoption of the specific criterion of determining the value of threshold. As there are no strict and unified procedures concerning the latter of the mentioned aspects, this stage of data preparing is often connected to relatively high degree of subjectivity. The threshold can be set on the basis of a thorough analysis of the data (*e.g.*, Tanaka and Takara 2002, Svensson *et al.* 2007) – *e.g.*, from mean residual life plots (Coles 2001, Svensson *et al.* 2007), or assumed *a priori* on the basis of previous experiences and local specificity of the region. In some of the approaches the threshold value is determined on the basis of the arbitrarily assumed preferred average number of events per year (usually from 2 to 5) (*e.g.*, Tavares and Da Silva 1983, Trefry *et al.* 2005). Other authors suggest that the size of a "partial-duration series" sample should be at least 1.65 times greater than the number of years on record (*e.g.*, Cunnane 1973). More detailed list of different approaches to this question is presented by Ben-Zvi (2009).

In general, both approaches – "annual maxima series" and "partial-duration series" – lead to different IDF-formulas for the same gauge station and the same length of record period, thus it is important to take into account and state explicitly the method of estimation, and use recalculation formulas if necessary (Chow 1964).

2.3 IDF-formulas determination

After extracting the extreme rainfalls from the whole period of observations, the most intense rainfall phases of considered duration, t_d, are selected. Thus, from each selected rainfall episode, the extreme value of intensity (referred to the considered t_d) is derived by moving-average technique. For example, if the duration t_d equal to 30 min is considered, from each episode the worst (as to the rainfall intensity) phase lasting 30 min is selected, and the average intensity during this phase is recognized as the extreme value. This way, the set of extreme intensities is developed for each considered value of duration t_d. The statistical analysis is then carried out to obtain the formulas for the intensity-duration-frequency relationships.

In order to obtain the final result of IDF-formulas, two approaches may be applied – "empirical" (also called "physical") or "theoretical" (also called "probabilistic"). In the first one, the selected and structured storm data are

described by a chosen mathematical formula (set of formulas) on the basis of matching the general form of the formula and calibrating parameters, usually with use of optimization techniques. In the second one – theoretical distribution functions are applied and the parameters are estimated on the basis of statistical approach. The distribution functions most commonly used in IDF-formula development are Gumbel, Log-Pearson type III, and Generalized Extreme Value (GEV); however, other functions, such as Gamma, Log-normal, Pearson III, Fisher-Tippett, Generalized Pareto or exponential distributions have also been applied (*e.g.*, Sevruk and Geiger 1987, Katz *et al.* 2002, Koutsoyiannis 2004a, b, De Michele and Salvadori 2005, Venkata Ramana *et al.* 2008, Ben-Zvi 2009, Endreny and Imbeah 2009, Elsebaie 2012). Polish experiences indicate, either empirical approach or theoretical Pareto, Pearson III, Fisher-Tippett, and Log-normal distributions were usually adopted (*e.g.*, Wołoszyn 1991, Suligowski 2004 and others). More detailed mathematical background of IDF-formulas development was presented by Koutsoyiannis *et al.* (1998). Although the second (theoretical) way of IDF-formulas estimation may be considered as more justifiable from hydrological point of view, the first approach is also successfully applied and in some cases may even demonstrate better efficiency, especially if rainfall data do not "match" the theoretical distributions (*e.g.*, Langousis and Veneziano 2007). In such cases, the optimized empirical IDF-relation enables better fitting to the observations than "typical" hydrological approaches.

A generalized IDF relationship between rainfall maximal intensity I and duration of the rainfall interval t_d corresponding to specified return period T is usually presented in the literature in the form (Koutsoyiannis *et al.* 1998, Venkata Ramana *et al.* 2008, Endreny and Imbeah 2009):

$$I = \frac{A(T)}{B(t_d)} , \qquad (1)$$

where $A(T)$ and $B(t_d)$ are functions defining separable dependence of I on T and t_d. The function $B(t_d)$ in Eq. 1 is usually expressed as:

$$B(t_d) = (t_d + \theta)^{\eta} , \qquad (2)$$

where θ and η are parameters ($\theta \geq 0$, $0 < \eta < 1$).

Function $A(T)$ in Eq. 1 is usually presented in the literature in the form of two alternative relations of empirical type:

$$A(T) = A_1 T^k , \qquad (3a)$$

$$A(T) = A_1 \ln T + c , \qquad (3b)$$

where A_1, k, and c are empirical parameters. However, it can be also derived for typical distributions of maxima on the basis of hydrologic distribution functions (Koutsoyiannis *et al.* 1998), *e.g.*:

for Gumbel distribution:
$$A(T) = \lambda \left\{ \psi - \ln \left[-\ln \left(1 - \frac{1}{T} \right) \right] \right\} , \tag{4a}$$

for GEV distribution:
$$A(T) = \lambda \left\{ \psi + \frac{\left[-\ln \left(1 - \frac{1}{T} \right) \right]^{-\kappa} - 1}{\kappa} \right\} , \tag{4b}$$

for exponential distribution:
$$A(T) = \lambda \left(\psi + \ln T \right) , \tag{4c}$$

for Pareto distribution:
$$A(T) = \lambda \left(\psi + \frac{T^{\kappa} - 1}{\kappa} \right) , \tag{4d}$$

and others, where λ, ψ, and κ are the parameters of distribution functions.

Final formulas of IDF-relationships for each considered case are obtained after identification of all the parameters, the number of which is usually not greater than five. In practice, instead of the simultaneous consideration of maximal rainfall intensities for different values of rainfall durations – leading to combined formulas of a general type (Eq. 1) – a very common approach is to separate development of $I(T)$ formulas for selected specified rainfall durations. This makes the process of parameter identification much easier because of the simplicity of the $I(T)$ formula and the significant decrease of the number of parameters to identify. However, the obtained result has less universal character and can be applied in practice only for the specified rainfall durations. Therefore, a more general form of IDF-formula, expressing the dependence of I on both, T and t_d, is preferable.

Indeed, most of widely applied IDF-formulas can be expressed as combination of Eqs. 1 and 2 with one of the alternative relations for $A(T)$ presented above. Commonly known formulas of Lindley, Talbot, Sherman or Reinhold are in fact particular forms of general relationship (Eq. 1). However, some authors (*e.g.*, Arnell 1982) presented the studies in which the most satisfactory results were obtained for the general form of IDF relation different from Eq. 1. Similar conclusions are also provided by Polish experiences, *e.g.*, published by Lambor, Wołoszyn, Licznar, and Łomotowski (all in Polish), focused on development of local IDF-formulas for selected localizations in the country. Taking into account most of these works, more general form of IDF-relationship may be expressed as:

$$I = \frac{A(T)}{B(t_d)} + C(T) \ , \tag{5}$$

(with the range of $A(T)$ formulas wider than Eqs. 3-4). However, in some particular cases presented by the above-mentioned authors, the best results were obtained for even more complex forms of IDF-formulas, which are not a simple combination of functions defining separable dependence of I on T and t_d as in Eqs. 1 or 5. In such cases, the problem of IDF-formula estimation is more complicated.

The parameters appearing in a general form of IDF-relationship must be estimated in each particular case in order to obtain the final IDF-formula for the specified localization. Estimation of the parameters is often considered to be an easy and well recognized subject. It can be developed in different ways, ether with use of statistical estimators for known probability distribution functions or with use of any optimization technique, suitable for the case analyzed. The choice of the method of parameter estimation is certainly very much connected to the assumed form of IDF-relationship and the number of parameters. In most practical applications, commonly used methods of parameter estimation are: L-moments (LM), Maximum Likelihood (ML), Probability Weighted Moments (PWM) (*e.g.*, Koutsoyiannis *et al.* 1998, Endreny and Imbeah 2009, Ben-Zvi 2009, Hailegeorgis *et al.* 2013, and many others) or optimization techniques, *e.g.*, Powell method (Koutsoyiannis *et al.* 1998). Although in many cases (especially when the simple form of IDF-formula is chosen with a small number of parameters) the parameter identification is actually a very easy procedure from a "technical" point of view. However, there are also experiences showing that this "standard" approach can lead to significant errors in final solution (*e.g.*, Katz *et al.* 2002). In some cases the recommended optimization techniques may not lead to satisfactory results due to the relatively high number of parameters being estimated with use of local optimization techniques. For example, Koutsoyiannis *et al.* (1998) present the method of simultaneous estimation of the parameters of both – the distribution function $A(T)$ and duration function $B(t_d)$ – with use of Powell method. Taking into account that in such optimization problem, depending on the chosen form of IDF-relationship, even five or more parameters are searched, it seems that the Powell method (and other local minimum search methods) may lead to "not optimal" solution, as it is difficult to verify whether the objective function is unimodal. Some authors recommend performing the optimization using the solver tools in common spreadsheet packages (usually without pointing the optimization method used in such approach) or present other methods of solving the problem, *e.g.*, robust estimation proposed by Koutsoyiannis *et al.* (1998). However, appli-

cation of global optimization methods to the problem of IDF-formula development is not often considered in bibliography.

This paper proposes the effective and relatively simple empirical approach to IDF-formula development based on global optimization technique of Controlled Random Search (CRS). The approach may be an interesting alternative to other methods, as it leads to satisfactory solution obtained in relatively quick and simple way, especially in those cases in which IDF-formula has more complex form and the optimal values of higher number of parameters are searched simultaneously. In such cases the properties of the objective function (mainly its unimodal/multimodal character and the approximate position of the global optimum) are usually difficult to prior recognition, and thus the choice of global search method enables to achieve a satisfactory solution. CRS method proved to be effective in other applications in engineering (*e.g.*, Ali *et al.* 1997a, Dysarz and Napiórkowski 2002, Manzanares-Filho *et al.* 2005, Manzanares-Filho and Albuquerque 2008). It can be also successfully applied to hydrological problem of IDF-relationship determination.

3. CONTROLLED RANDOM SEARCH METHOD AS A GLOBAL OPTIMIZATION TECHNIQUE

The procedure of the controlled random search (CRS) for global optimization was first presented by Price (1977, 1978, 1983). The author proposed a preliminary version of the method (further referred as CRS1; Price 1977, 1978), which was then modified to CRS2 (Price 1983). The algorithms proposed by Price were later used, modified and improved by many researchers (*e.g.*, Ali *et al.* 1997a, b, Manzanares-Filho *et al.* 2005, Manzanares-Filho and Albuquerque 2008, Tsoulos and Lagaris 2006). In this study the scheme CRS2 in version adopted by Dysarz and Napiórkowski (2002) was applied to IDF-formula development.

The CRS technique combines the "random" algorithms of typical global optimization methods with "deterministic"-"controlled" – procedures of local optimization, increasing the efficiency of the process. The method represents the group of "direct" procedures, where no gradient analysis is involved. It is applicable for both unconstrained and constrained optimization.

Let *n* be a number of variables (parameters) that are searched, and thus *n* defines the size of the space of possible solutions (the search domain) *V*. The search domain *V* is constituted by specifying the limits to each variable. In optimization problem, the "optimal" set of parameter values is searched, meaning such for which the assumed criterion (objective function) takes the extreme (usually minimal) value.

In the first step of the CRS1 procedure, predetermined number N of trial points are randomly chosen over n-dimensional space V, with respect to additional constraints, providing there are any. In the original version of the procedure (CRS1), Price (1977, 1983) suggests (based on empirical experience) applying the relation $N = 25\,n$. Each of the N points represents one possible combination of the values of parameters that are searched. Let us denote this set of points as S. For each of the chosen trial points, the objective function F is evaluated, and the points are then arranged in the array $\mathbf{A}$, in which the position is dependent on the value of F (from 1 – "the best" to N – "the worst"). Let M denote the worst and L – the best point in $\mathbf{A}$. In each iteration, the set of $n + 1$ points R_i ($i = 1, 2, \ldots, n+1$) is randomly chosen from S, constituting the so-called "simplex" in n-dimensional space. The point R_{n+1} is arbitrarily considered as the pole of the simplex and the next trial point P is defined as the image point of the pole R_{n+1} with respect to centroid G of the remaining n points, according to the formula:

$$\boldsymbol{P} = 2\boldsymbol{G} - \boldsymbol{R}_{n+1} \;, \tag{6}$$

where $\boldsymbol{P}$, $\boldsymbol{G}$, and $\boldsymbol{R}_{n+1}$ represent the position vectors of corresponding points in n-dimensional space. Point P is called preliminary trial point. For the obtained point P the value of objective function F_P is calculated and compared with F_M – the value for the worst point M. If $F_P < F_M$, and if P satisfies the additional constraints (if there are any), the point M is replaced by P, the array $\mathbf{A}$ is rearranged and a new simplex is chosen. If P fails to improve the value of M, the secondary trial point $\boldsymbol{Q}$ is chosen, according to the formula:

$$\boldsymbol{Q} = \left(\boldsymbol{G} + \boldsymbol{R}_{n+1}\right)/2 \tag{7}$$

and the value of F_Q is calculated. If Q occurs better than M – the points are replaced; otherwise – the trial is discarded. The new simplex is then randomly chosen and so on. The "best" parameter values at each step of the CRS procedure are defined by the coordinates of point L, and are improved in successive iterations. The next iteration is performed as long as the stopping criterion is satisfied.

The modified procedure CRS2 is similar to CRS1 in general approach, however it differs in details. In CRS2 only primary trial points (Eq. 6) are chosen, no secondary trial points are generated. The second important difference is that the point R_1 of a simplex is not randomly searched, but it is always defined by the "best" point L from the array $\mathbf{A}$ of N points. Thus the simplex is constituted by choosing n points from $N-1$, because L is arbitrarily chosen. In this way the "local part" of the searching procedure is much more effective. In order to minimize the danger of premature convergence to local optimum (what may take place in the case of multimodal problems),

the sufficiently high number of trial points N should be randomly chosen and the stopping criterion should be carefully considered. For this version of the method, Price (1983) recommends determination of the optimal number of N according to the formula:

$$N = 10(n+1) \ . \tag{8}$$

Thus, if 4 parameters are searched, the initial set of 50 points is randomly chosen and each simplex is constituted of 5 points in 4-dimensional space. For 8 parameters, the size of the set of initial points increases to 90 and each simplex is built of 9 points.

CRS2 was successfully adapted to different applications, including complex multimodal problems (Ali *et al.* 1997a). It also has been modified many times by different authors (*e.g.*, Rinaudo *et al.* 1998, Manzanares-Filho *et al.* 2005, Tsoulos and Lagaris 2006, Manzanares-Filho and Albuquerque 2008). The modifications, leading to subsequent versions CRS3, CRS4, CRS5, and CRS6 (Ali *et al.* 1997a, b), usually concern the initial procedure of random choice of N points, the choice of $n+1$ points of the simplex, combining CRS with different local search algorithms and modifications of the procedure of controlled random search. Price (1977, 1983), Ali *et al.* (1997b), Tsoulos and Lagaris (2006), and other authors present several "mathematical" examples of test function minimization. One can also find several examples of CRS successful application to the practical problems, *e.g.*, inverse airfoil design (Manzanares-Filho *et al.* 2005, Manzanares-Filho and Albuquerque 2008), macromolecular modeling (finding least energy structure for a given molecular system, Ali *et al.* 1997a), chemical engineering (tank reactor optimization, Ali *et al.* 1997a), applied statistics (pig liver problem; Ali *et al.* 1997a), and optimal flood control in multi-reservoir river system (Dysarz and Napiórkowski 2003). The CRS method and its modifications were compared to other approaches of global optimization by several authors. The interesting study of such a comparison was presented by Ali (1994) and Ali *et al.* (1997a). The authors compared different versions of CRS algorithms with the methods of multilevel single linkage (MSL) (Rinnooy Kan and Timmer 1987), topographical multilevel single linkage (TMSL) (Ali and Storey 1994), simulated annealing (SA) (Dekkers and Aarts 1991) and aspiration-based simulated annealing (ABSA) (Ali and Storey 1997), all applied to selected practical problems. The efficiency of the algorithms was compared taking into account the number of function evaluations, the CPU time required to solve the problem and the value of obtained optimum in each analyzed case. Numerical tests showed that the algorithms of CRS-type were preferable to the others for problems with many optima, especially if the number of local optima was large, even if the optima were close to each other.

The CRS method can be also successfully applied to IDF-curves determination. The problem of IDF-relationship development, in comparison with the ones from the fields of material science or chemical engineering, is relatively easy. However, if the general form of IDF-relationship is more complex than the functions typically applied in such cases (*e.g.*, Eqs. 1-4), and if the number of parameters to be identified is higher (four, five or more) and finally – if the objective function is multimodal, the use of effective but relatively simple global optimization procedures is reasonable. As an example, the application of CRS2 algorithm to IDF-formula determination for the rainfall data collected in one of the gauge stations in Gdańsk (Poland) – in the Gdańsk University of Technology (GUT) station – is presented. The results are compared with the ones obtained with the use of successive search, Powell method (Powell 1964, Pierre 1969, Press *et al.* 2007), modified version of Powell method, Newton search (*e.g.*, Bonnans *et al.* 2006) and evolutionary algorithm (Goldberg 1989). In the two latter approaches, the embedded solver tools of commonly used spreadsheet packages (Excel) were applied. In the first case, Newton's search algorithm implemented in classical optimization tool proposed by Fylstra *et al.* (1998) was used. In the second one, The Evolutionary Solver (in Excel 2010), based on the combination of genetic algorithm and other search methods, was applied.

4. GDAŃSK STUDY CASE CHARACTERISTICS

The city of Gdańsk is located in the northern Poland, at the bank of Baltic Sea, in the region of Vistula river estuary (Fig. 1). Average yearly precipitation depth for Gdańsk (for the period 1951-2008) is 550 mm, thus it is lower than the corresponding value for the whole territory of Poland (*c.a.* 600 mm). The value of total annual precipitation depth for Gdańsk over the period of nearly 60 years of observations is highly variable; however, as a general trend, the well pronounced increase can be observed (Wołoszyn 2009). Precipitation in cold period (November-April) is on average 200 mm, thus it is nearly twice smaller than in warm season (May-October). Despite the above-mentioned proportion in precipitation depths, the number of days with measurable precipitation in winter is higher than in summer. Precipitation in cold period (especially in December, January, and February) has a more permanent character and usually takes the form of snow (in average 40-50 days per year). However, daily sums of precipitation in winter are low and snow cover lasts for a relatively short time. Thus, for the purposes of urban hydrology, precipitation in the form of rainfall, (occurring mainly in warm period, often as short but heavy storms) is much more important from an engineering point of view.

Fig. 1. Distribution of averaged total yearly precipitation in Poland.

Nowadays, on the territory of Gdańsk, there are three meteorological and climatological observation posts in charge of the national research institute (Institute of Meteorology and Water Management) and many private stations. Important part of the latter group was created by Gdańskie Melioracje (Gdańsk Meliorations Inc.), thanks to which the whole network of rainfall gauges, designed particularly for urban hydrology needs, was installed. Although the lengths of the observed precipitation records in these stations are still too short to support sufficient data for statistical analysis nowadays, the observations will be the important source of rainfall data in the closest future, both for temporal and spatial variability analysis.

The gauge station at the Gdańsk University of Technology (GUT), remaining in charge of a Faculty of Civil and Environmental Engineering, is located in the central district of the city. The GUT station was established in 1991, when Hellman's pluviometer and two pluviographs were installed. The observation post was additionally equipped with the automatic system in 2008, and since then both analog and automatic measurements of rainfall are carried out.

The observations in Gdańsk indicate that the mean monthly value of precipitation is the highest for July (70.0 mm) and August (63.3 mm) and the lowest for February (25.2 mm) and March (26.2 mm). The same tendency can be observed in the context of the mean daily maximum values – the

highest for July is about 24.1 mm, while the lowest for February is 6.5 mm (Wołoszyn 2009). However, the historical observations have shown that the region is not free from extreme phenomena. The most spectacular one, causing catastrophic flood in Gdańsk, took place in 2001, when during 24 hours (9 July 8.00 a.m. – 10 July 8.00 a.m.) the sum of 123.5 mm in GUT station was observed (*c.a.* 20% of average yearly precipitation in Poland). During the most intense phase of the rain, the amount of 65 mm (during 1 hour 40 min) was measured (Wołoszyn 2003).

The above-presented characteristics of precipitation in Gdańsk are additionally influenced by the factors resulting from local aspects. Observations indicate both: spatial and temporal variability of precipitation, caused by varied rate of urbanization and city development in different parts of the city. This confirms the strong need not only to determine the formulas characterizing "averaged" precipitation conditions in Gdańsk, but to develop the separate local IDF-formulas for several districts of the city as well.

5. PRECIPITATION DATA COLLECTION AND ANALYSIS

Rainfall data were collected in GUT climatological station during the period 1991-2010. Till 2008 daily and weekly pluviograph charts recorded by two Hellman's siphon pluviographs (float-type) were available. Since 2008 additionally the recordings of the automatic tipping-bucket rain gauge (with the time resolution of 10 min and the measurement accuracy of 0.2 mm) have been at disposal.

Data selection was carried out in several steps:
- identification and rough selection of rainfall episodes,
- analysis of chronological course of each selected episode,
- analysis of extreme rainfall phases during each episode,
- selection (for each assumed value of rainfall duration t_d) of the episodes relevant for IDF analysis, according to the pre-selected criterion.

For each day of the considered period of 20 years of observations, precipitation recorded at a resolution of 10-minutes was analyzed. In order to establish the criterion for storm selection, the Chomicz rainfall intensity scale, popular in Poland, was adopted. According to this scale, each rainfall episode (or each selected phase of the episode) can be classified in the context of its intensity to one of thirteen categories on the basis of the relation between rain/phase duration and the corresponding precipitation depth. Each category in this scale is given a grade number (from 0 to 12). The reference values of precipitation depth, distinguishing different types of rainfall, are defined as:

$$U_k = \alpha_k \sqrt{t} \ , \tag{9a}$$

where

$$\alpha_k = \sqrt{2^k} \quad , \tag{9b}$$

k is a grade of the Chomicz scale ($k = 0, 1, …, 12$), U_k is the upper limit value of precipitation depth [mm] for the rainfall of a kth grade and t [min] is duration time of a rainfall (or its selected phase). First nine lines describing the relation 9 for grades $k = 1…9$ are presented in Fig. 2. According to the Chomicz scale, rainfalls are classified as normal (grade 0), heavy (grade 1), storms (grades 2 and 3), heavy storms (grades 4 and 5), and torrential storms (grades 6-12). Knowing the duration of the analyzed episode/phase and total precipitation depth accumulated in this period, one can find the corresponding point in the plot and therefore determine the category of the rain intensity.

The approach applied in the Chomicz scale differs from the most popular method of rainfall classification, in which the category of rainfall intensity is determined on the basis of precipitation depth accumulated in predefined period (usually 1 hour or 24 hours) or average rainfall intensity. Differences between such approaches in different countries usually concern demarcation values of precipitation depth or rainfall intensities. For example, according to National Meteorological Institute of Spain, rainfall can be classified as light for intensities I not greater than 2 mm/hour, moderate – for $2 < I \leq 15$ mm/hour, heavy – for $15 < I \leq 30$ mm/hour, very heavy – for $30 < I \leq 60$ mm/hour and torrential for $I > 60$ mm/hour (Llasat 2001). In a similar Polish approach proposed by the Institute of Meteorology and Water Man-
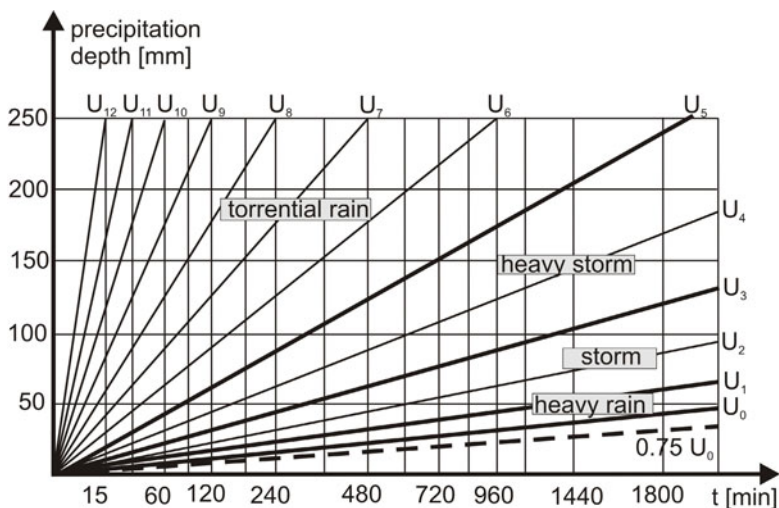

Fig. 2. Illustration of Chomicz scale for precipitation intensity.

agement, heavy rains refer to intensities $0.2 < I \leq 1.0$ mm/min $(12 < I \leq 60$ mm/hour), storms – $1.1 < I \leq 1.9$ mm/min and torrential rainfalls – $2.0 < I$ mm/min (with several grades in each of the mentioned categories). In all these approaches based on the predefined values of intensities, the ratio of precipitation depth to its duration, describing the limits of each rainfall category, is constant, while in the Chomicz scale not. The nonlinear character of formula 9 stresses not only the importance of intensity itself, but also duration of the period in which such intensity was observed.

Besides these simple approaches to rainfall classification, one can find more complex analyses. For example, Llasat (2001) presented the classification of rainfall events on the basis of their convective features described by the parameter $\beta^{*}_{L,\Delta T}$. This parameter, calculated on the basis of rainfall hyetograph (with the interval ΔT), expresses the ratio of the precipitation that exceeds the predefined intensity threshold L to the total rainfall in the episode. Based on the value of $\beta^{*}_{L,\Delta T}$ the episode is classified into one of four categories. The author discussed the relations between $\beta^{*}_{L,\Delta T}$, intensity and rainfall duration for the analyzed region. Maraun *et al.* (2008) presented the method for the classification of rainfall intensity based on 10 categories, each referred to 10% of the total rainfall amount (for the analyzed period of observations), from the 1st category containing the weakest precipitation events (in the number sufficient to make up the lowest 10% of the total rainfall amount) to the 10th category – the highest one. In this approach the values of thresholds depend on the number and intensity of the measured rainfall episodes, and not on the objective criterion, independent of the results of observations.

Different scales of rainfall intensities lead to different classifications of the same set of episodes therefore the choice of the objective criterion is difficult. In this study, Polish experiences in IDF-formula development were taken into account. The Chomicz scale, revised by recent Polish studies (Licznar and Łomotowski 2005), was finally chosen as the basis of the threshold selection. The experiences of Licznar and Łomotowski (2005) showed that "heavy rain" criterion referred to $k > 0$ of the Chomicz scale is too stringent, eliminating a large number of episodes of a lower precipitation depth which in practice should be taken into account. Consequently, the mentioned authors established the criterion:

$$P \geq 0.75 \cdot U_0 \ , \tag{10}$$

where P is the observed precipitation depth [mm], and U_0 is the reference value of Chomicz scale for $k = 0$. Thanks to such approach, each year in the analysis is represented by at least a few rainfall episodes. Following the sug-

gestions of the above-mentioned authors, the criterion 10 was also implemented in this study.

During the preliminary (rough) selection of the episodes (step 1), only the smallest, evidently below criterion episodes were neglected and 122 episodes observed during the period of 20 years were taken into account (Table 1). Average number of episodes per year was 6. The extreme frequency of storm episodes in a year was 13 and it was observed in 2008. The longest continuous rainfall episode was 1092 min long and was observed on 9/10 July 2001, when the already mentioned catastrophic flood in Gdańsk took place. However, the case of such a long rainfall was incidental, as most measured rainfall episodes were not longer than 100 min. Many of them were very short storms, lasting one hour or less, separated from other episodes by rainless period of at least several hours. In most cases, a single episode was observed per day; days with two or three rainfalls were incidental and the duration of rainless interval between the storms in such cases was usually a few hours.

Table 1

Number of extreme episodes from period 1991-2010 taken into considerations

Year	Number of extreme rainfall episodes								
	March	April	May	June	July	August	September	October	Total
1991	–	–	–	5	–	–	–	–	5
1992	–	–	–	1	3	–	–	–	4
1993	–	–	–	–	2	2	2	–	6
1994	–	–	1	–	–	1	2	–	4
1995	–	–	–	2	2	–	–	–	4
1996	–	–	1	1	3	–	–	–	5
1997	–	–	3	–	1	–	–	–	4
1998	–	–	2	–	1	–	–	1	4
1999	–	1	2	1	–	–	–	–	4
2000	–	–	1	2	–	1	2	–	6
2001	–	1	1	–	7	3	–	–	12
2002	–	–	3	1	–	2	–	–	6
2003	–	–	1	1	3	–	2	–	7
2004	–	–	–	1	2	–	1	–	4
2005	–	–	3	–	1	–	–	–	4
2006	–	–	–	3	–	–	1	–	4
2007	–	–	–	1	1	–	5	2	9
2008	1	–	2	3	4	2	–	1	13
2009	–	–	2	3	2	–	1	–	8
2010	–	–	4	1	1	2	1	–	9
Total	1	2	26	26	33	13	17	4	122

The selected episodes were analyzed in their chronological course in terms of the precipitation depth P [mm], intensity I [mm/min] and q [dm^3/(s·ha)] and cumulative precipitation depth ΣP [mm] in each 10-minute step of the episode duration. The example of such analysis (for the extreme episode recorded on 11 June 2010) is presented in Table 2. The analysis let observe that for very short rainfalls (20-30 min) maximal intensity usually fell on first impulse of rain. For short rainfalls (40-90 min), the maximal intensity is usually observed in the first half of the episode. Longer rainfalls usually have several peaks of intensity, at least one in the first and in the second half of the episode.

Table 2

Rainfall characteristics – the episode recorded on 11 June 2010

Clock time hours (p.m.)	Duration [min]	Precipitation depth [mm]	Cumulative precipitation depth [mm]	Intensity i [mm/min]	Intensity q [dm^3/(s·ha)]
4.20 – 4.30	10	0.3	0.3	0.03	5.00
4.30 – 4.40	10	8.6	8.9	0.86	143.36
4.40 – 4.50	10	1.8	10.7	0.18	30.01
4.50 – 5.00	10	0.5	11.2	0.05	8.34
5.00 – 5.10	10	0.3	11.5	0.03	5.00
5.10 – 5.20	10	0.5	12.0	0.05	8.34

Next stage of the study concluded of extreme rainfall phase analysis. In the study, the aggregation levels t_d equal to 10, 20, 30, 60, 90, 120, 150, 180, 210, and 240 min were considered. For each rainfall episode, the most intensive phases of pre-selected durations t_d (not exceeding the total duration of the episode) were found and analyzed as to their intensities. The assumed value of 240 min of the upper limit of t_d results from the specificity of the observed rainfall episodes, but it also has its practical justification. As it is known, the design rainfall duration, specified for the purposes of urban hydrology, is strongly correlated with the time of runoff concentration in the considered catchment. Most practical cases of urban drainage and rainwater system design concern relatively small basins with significant share of impervious areas, where time of concentration is short and runoff processes are very dynamic. Moreover, it has been proved that the vast majority of urban floods is caused by extremely intensive torrential rainfalls of short duration. Thus, the value of 240 min as the upper limit of rainfall duration in the presented considerations enables proper recognition of the rainfalls most important from the practical point of view and corresponding with local conditions in Gdańsk. The example of the analysis of the maximal intensity phases (episode of 11 June 2010) is presented in Table 3 and Fig. 3.

Table 3

Rainfall maximal intensities – the episode recorded on 11 June 2010

Clock time hours (p.m.)	Duration time t_d [min]	Maximal precipitation depth [mm]	Maximal rainfall intensity q [dm³/(s·ha)]
4.30 – 4.40	10	8.6	143.36
4.30 – 4.50	20	10.4	86.68
4.30 – 5.00	30	10.9	60.57
4.20 – 5.20	60	12.0	33.34
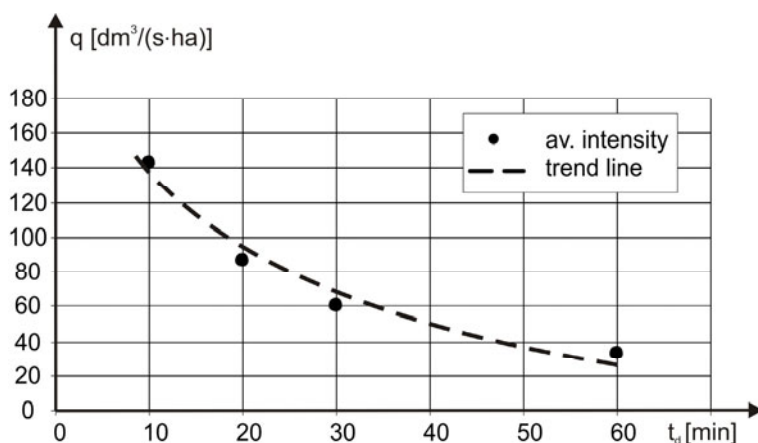
Fig. 3. Maximal intensity *versus* duration time –the episode recorded on 11 June 2010.

The intensities averaged for the analyzed rainfall phase durations were then segregated in tables with each duration t_d separately, and data in each table were arranged in the non-increasing sequence of string-values. For each table the condition 10 was rechecked and the data that did not match the criterion were removed. Such set of data was the starting point for IDF analysis.

6. PRELIMINARY IDF-FORMULA DEVELOPMENT

6.1 The $q(p)$ relationship

The observed rainfall extremes for each duration time t_d, arranged in the statistical order of non-increasing series q_i, $i = 1, 2, …., N$, where $q_1 \geq q_2 \geq … \geq q_N$, were assigned to the empirical quantiles corresponding to empirical probability of exceedence p_i according to the formula:

$$p_i = \frac{i}{N+s} \cdot 100 \ [\%] \ , \tag{11}$$

where N was the total number of extremes in the series and $0 \leq s \leq 1$. In this study $s = 1$, which corresponds with the so-called Weibull plotting position of a quantile plot. The formula obtained in this way is considered to be the most practical one, which fully satisfies Gumbel's conditions for plotting position (Yevjevich 1972). The example for t_d equal to 120 min is presented in Table 4. The comparison of the obtained relations for selected values of t_d is shown in Fig. 4.

Table 4

Sequence of distribution of extreme rainfall intensity
for duration time t_d equal to 120 min

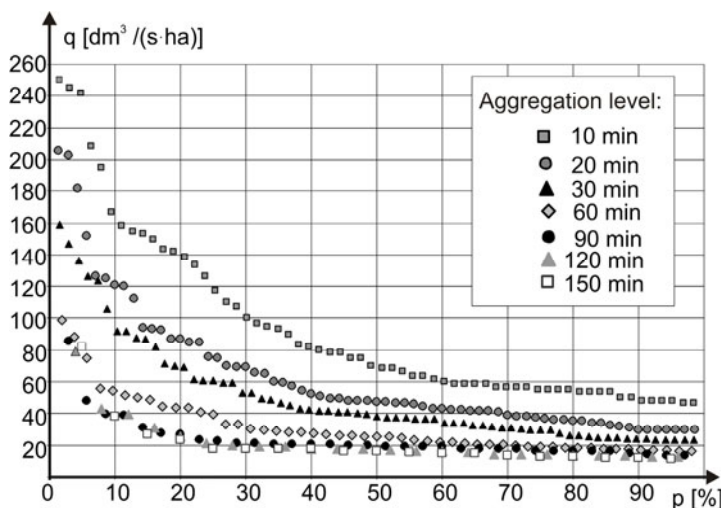| No. in series i	Episode number	Total precipitation depth [mm]	Average maximal rainfall intensity q [dm^3/s·ha]	Probability of exceedence p [%]
1	124	57.000	79.170	4.00
2	5	30.800	42.778	8.00
3	22	28.350	39.375	12.00
4	66	22.400	31.111	16.00
5	123	18.100	25.139	20.00
6	6	15.200	21.111	24.00
7	37	14.100	19.583	28.00
8	39	13.700	19.028	32.00
9	91	13.600	18.889	36.00
10	34	12.900	17.917	40.00
11	58	12.500	17.361	44.00
12	76	12.500	17.361	48.00
13	73	12.200	16.944	52.00
14	15	11.700	16.250	56.00
15	86	11.600	16.111	60.00
16	61	11.000	15.278	64.00
17	25	10.400	14.444	68.00
18	67	10.400	14.444	72.00
19	115	10.000	13.889	76.00
20	117	9.800	13.611	80.00
21	100	9.700	13.472	84.00
22	45	9.600	13.333	88.00
23	94	9.250	12.847	92.00
24	90	9.050	12.569	96.00

Fig. 4. Comparison of relationship $q(p)$ for different values of duration time.

Table 5

Results of preliminary approximation of $q(p)$ relationships

t_d [min]	$q = a \cdot \ln p + b$			$q = a \cdot p^{-b}$		
	a	b	R^2	a	b	R^2
10	57.931	303.43	0.9768	542.67	0.5194	0.9489
15	47.767	245.07	0.9706	435.57	0.5351	0.9675
20	43.401	220.92	0.9709	398.81	0.5464	0.9698
30	34.933	177.56	0.9778	333.35	0.5597	0.9609
60	19.870	103.66	0.9650	176.74	0.5076	0.9798
90	14.025	74.484	0.8100	105.34	0.4388	0.9605
120	15.945	80.253	0.8252	134.54	0.5303	0.9664
150	17.218	84.021	0.7366	144.88	0.5693	0.9293
180	20.092	95.855	0.7686	214.30	0.6675	0.9466
210	20.044	96.263	0.8159	247.06	0.6963	0.9577
240	22.490	105.150	0.8618	413.55	0.8480	0.9620

The relationships $q(p)$ obtained for each considered aggregation time t_d were analyzed in terms of the best fitting to rainfall data. Different forms of polynomial, logarithmic, and exponential functions were considered and least square method of approximation was applied. Preliminary analysis for each duration t_d showed that there was no type of function that would fit the rainfall data over the whole range of considered values of aggregation time.
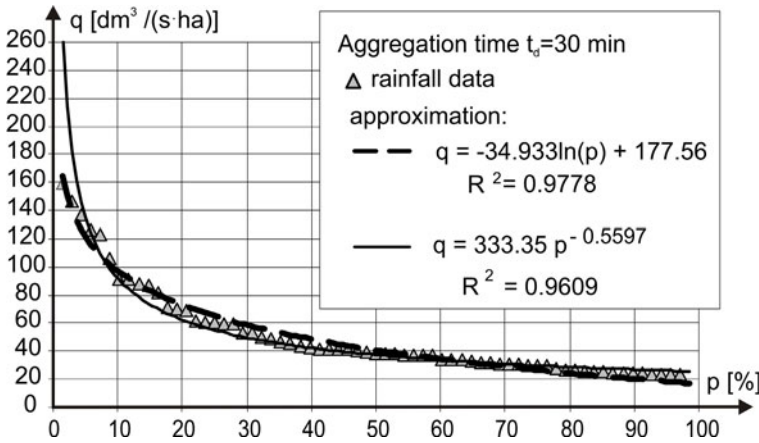
Fig. 5. Approximated $q(p)$ relationship for aggregation time equal to 30 min.

The best fit was obtained for logarithmic and exponential functions (equivalent to Eqs. 3a and b for q and p), however the first one proved to be the best for $t_d < 60$ min, while the second one – for $t_d \geq 60$ min (Table 5). In this last case, logarithmic function did not represent the shape of the relationship $q(p)$ in satisfactory way, while exponential function did not reproduce the $q(p)$ relation properly for the least values of p ($p < 2\%$), causing significant overestimation. The best (in the means of the least square error criterion) approximated functions were characterized by correlation coefficient $R^2 < 0.929, 0.979 >$. The example of approximated relationship $q(p)$ for t_d equal to 30 min is presented in Fig. 5.

6.2 The $q(t_d)$ relationship

Based on approximated relationships $q(p)$ for different values of duration time t_d, the $q(t_d)$ relations for selected values of probability p were analyzed. The obtained relationships were approximated with use of least square method. Exponential functions proved to provide the best matching to data points; however, the correlation between the data and approximated function was much poorer (with R^2 even less than 0.7 for some values of p). An example of approximation for $p = 20\%$ is presented in Fig. 6.

The results of the first stage of the analyses showed that the rainfall data required more complicated form of a function representing $q(t_d, p)$ relationship than the easiest forms based on Eq. 1, as the latter did not enable satisfactorily good fitting with observations in this case. Therefore, in order to consider more complex functions describing IDF-relations, more elaborate calibration methods will be desired.
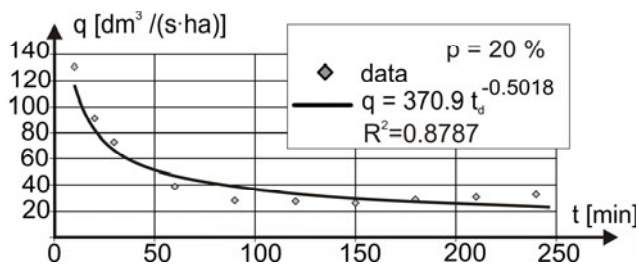
Fig. 6. Approximation of $q(t_d)$ relationship for $p = 20\%$.

6.3 Preliminary calibration of $q(t_d, p)$ relationship with the use of optimization

As mentioned before, the preliminary stage of the analysis showed the need for more complex functions describing IDF-relationship than basic relations expressed by Eqs. 1-3. In order to find the most suitable function describing the relation of q and both – aggregation time and exceedence probability, several forms of functions were taken into considerations:

a)
$$P_{\max} = \frac{a + b[-\ln(-\ln(1-p))]}{t_d^n} \ , \tag{12}$$

b)
$$P_{\max} = \left[a\left(\frac{1}{p}\right)^n + b \right] t_d^{-n} \ , \tag{13}$$

c)
$$P_{\max} = \left[a\left(\frac{1}{p}\right)^m + b \right] t_d^{-n} \ , \tag{14}$$

d)
$$P_{\max} = a\left(\frac{1}{p}\right)^m t_d^{-n} + b \ , \tag{15}$$

e)
$$P_{\max} = a\left(\frac{1}{p}\right)^m t_d^{-n} + b\,t_d \ , \tag{16}$$

f)
$$q = -a \ln p + b \ , \tag{17a}$$

where
$$a = a_1 t_d^2 + a_2 t_d + a_3 \ , \tag{17b}$$

$$b = b_1 t_d^2 + b_2 t_d + b_3 \ , \tag{17c}$$

and

g)
$$q = a\,p^{-b} \ , \tag{18}$$

where a and b are defined as in Eqs. 17b, c.

Functions a-e were selected on the basis of previous Polish experiences (mentioned in Section 2.3) and the obtained IDF-formulas for other localizations in the country. Formulas f and g were established on the basis of preliminary analysis. Parameters a and b of logarithmic and exponential relations $q(p)$ (Eqs. 17a and 18), estimated with use of least square method for each selected aggregation level t_d (Table 5), were then analyzed in terms of their dependence on t_d. As a result, it was concluded that the best fit was obtained when each parameter in Eqs. 17a and 18 (a and b) was a quadratic function of t_d. Therefore, formulas 17b, c were proposed. Finally, seven forms of IDF-relationship were taken into account.

In order to find the optimal values of parameters in formulas 12-18, the optimization was carried out. In the first stage, the successive search technique was used. Although time-consuming, it enabled recognition of the general features of the objective function selected as the optimization criterion in terms of its general shape and unimodal or polymodal character. This analysis was then supported by the calculations using Newton algorithm for local optimization. As the optimization criterion, the least square error objective function was assumed, which took the form:

$$F = \sum_{j=1}^{N_t}\left(P_{\mathrm{ob}} - P_{\mathrm{c}}\right)^2 \qquad (19)$$

or

$$F = \sum_{j=1}^{N_t}\left(q_{\mathrm{ob}} - q_{\mathrm{c}}\right)^2 , \qquad (20)$$

where N_t is the total size of a sample, index "ob" denotes "observations" and "c" – "calculations" of precipitation depth P or rainfall intensity q, respectively. The optimal set of parameters (for each formula a-g separately) was the one that minimized the value of F. To compare the obtained results and indicate the goodness-of-fit, several criteria were applied: the total square error (defined by the value of objective function F for the optimized set of parameters), the mean error ε defined as:

$$\varepsilon = \frac{\sqrt{F}}{N_t} \qquad (21)$$

and an "optimization cost" C equal to total number of objective function evaluations during the process of optimization. Additionally, the criteria of average relative errors:

$$E_1 = \frac{1}{N_t} \sum_{j=1}^{N_t} \left| \frac{P_{\mathrm{ob}} - P_{\mathrm{c}}}{P_{\mathrm{ob}}} \right| \, , \tag{22}$$

$$E_2 = \frac{1}{N_t} \sum_{j=1}^{N_t} \left| \frac{P_{\mathrm{ob}} - P_{\mathrm{c}}}{P_{\mathrm{c}}} \right| \tag{23}$$

were also analyzed. An interesting study of practical aspects concerning optimization, including objective functions and identifiability issues, was presented by Romanowicz *et al.* (2013).

After optimizing the IDF-formula parameters, the quality of the optimization (in the sense of the above-mentioned criteria) was checked (Table 6, "all data" tests). The study proved that even when the more complex form of IDF-relationship was applied, it was impossible to find one global formula matching to rainfall data in the whole range of considered values of aggregation time and probability. In all cases (a-g) the matching was not satisfactory, much worse than when $q(p)$ or $q(t_d)$ relations were searched separately. The analysis showed that the general form of relationship $q(t_d, p)$ describing short and very intensive rainfalls should be different from the corresponding formula for the rainfalls of longer duration. Better results than presented in Table 6 ("all data" tests) were obtained when two different formulas, Eq. 17a for $t_d < 60$ min and Eq. 18 for $t_d \geq 60$ min, were taken into account, even when the parameters a_i and b_i ($i = 1, 2, 3$) were estimated not *via* optimization but with use of simple quadratic approximation of $a(t_d)$ and $b(t_d)$ relations on the basis of the data presented in Table 5 only. For the values of a_i and b_i ($i = 1, 2, 3$) obtained in the latter case, the corresponding values of F defined by Eqs. 19 and 20, were 4609 and 24 661, respectively, and thus they were significantly lower than the values in Table 6 ("all data" tests). Therefore, a more complex mathematical description of IDF-relationship than one "global" equation should be applied in this case.

In the last step of the preliminary analysis, all the formulas 12-18 were applied to each of the ranges $t_d \leq 60$ min and $t_d \geq 60$ min separately. This time rainfall data for $t_d = 60$ min were taken into account twice (for both of the ranges considered), in order to obtain better coincidence in the surrounding of 60 min. The optimization was performed in the same manner as in the case of the whole range of t_d. The only difference was the lower number of rainfall data taken into account. The test confirmed that the best results were obtained when Eq. 17a for $t_d \leq 60$ min and Eq. 18 for $t_d \geq 60$ min were applied (Table 6). Thus, finally the Eqs. 17a and 18 were chosen as the best for the case considered.

Once two different formulas to obtain better matching for both, short and longer rainfall durations, are chosen, the main, dispositive stage of optimiza-

Table 6

Effectiveness of different IDF-formulas application

IDF formula (Eq. no.)	Optimal values of parameters	F Eq. 19 [mm²]	F Eq. 20 [dm³/(s·ha)]²	E_1 Eq. 22 [—]	E_2 Eq. 23 [—]	ε^* Eq. 21 [mm]	ε^{**} Eq. 21 [dm³/(s·ha)]
all data ($t_d \leq 240$ min)							
12	$a = 1.052$, $b = 0.662$, $n = -0.550$	23113	81 693	0.278	0.321	0.321	0.604
13	$a = -0.379$, $b = 4.130$, $n = -0.499$	20859	107 006	0.471	0.337	0.305	0.691
14	$a = 14.407$, $b = -9.476$, $n = -0.555$, $m = 0.078$	7451	42 005	0.126	0.110	0.182	0.433
15	$a = 8.382$, $b = -1.899$, $n = -0.443$, $m = 0.341$	10120	70 231	0.189	0.160	0.213	0.560
16	$a = 4.324$, $b = -0.246$, $n = -0.716$, $m = 0.309$	11328	56 018	0.171	0.526	0.225	0.500
17a-c	$a_1 = 0.003$, $a_2 = -0.783$, $a_3 = 60.127$ $b_1 = 0.015$, $b_2 = -4.045$, $b_3 = 312.846$	23215	83 144	0.237	0.429	0.322	0.610
18	$a_1 = 0.046$, $a_2 = -6.936$, $a_3 = 418.709$ $b_1 = 0.000$, $b_2 = 0.003$, $b_3 = 0.356$	14567	78 407	0.200	0.183	0.255	0.592
$t_d \leq 60$ min							
12	$a = 0.992$, $b = 0.623$, $n = -0.577$	3445	68 692	0.261	0.465	0.162	0.722
13	$a = -0.350$, $b = 4.000$, $n = -0.509$	1386	95 278	0.166	0.165	0.103	0.850
14	$a = 16.169$, $b = -11.547$, $n = -0.577$, $m = 0.063$	789	31 933	0.098	0.091	0.077	0.492
15	$a = 8.464$, $b = -2.122$, $n = -0.438$, $m = 0.331$	861	60 393	0.163	0.137	0.081	0.677
16	$a = 3.955$, $b = -0.566$, $n = -0.794$, $m = 0.231$	1179	40 241	0.107	0.146	0.095	0.553
17a-c	$a_1 = 0.015$, $a_2 = -1.747$, $a_3 = 72.597$ $b_1 = 0.088$, $b_2 = -9.999$, $b_3 = 387.700$	297	21 022	0.093	0.097	0.047	0.399
18	$a_1 = 0.083$, $a_2 = -10.347$, $a_3 = 465.369$ $b_1 = 0.000$, $b_2 = 0.004$, $b_3 = 0.364$	544	58 786	0.155	0.135	0.064	0.668
$t_d \geq 60$ min							
12	$a = 0.417$, $b = 0.278$, $n = -0.712$	7977	10810	0.244	0.343	0.554	0.646
13	$a = -0.894$, $b = 6.365$, $n = -0.406$	11571	13711	0.229	0.282	0.668	0.727
14	$a = 2.023$, $b = -0.112$, $n = -0.819$, $m = 0.425$	4304	5032	0.136	0.131	0.407	0.441
15	$a = 0.200$, $b = 6.027$, $n = -1.392$, $m = 0.797$	2077	4039	0.124	0.118	0.283	0.395
16	$a = 1.725$, $b = -0.085$, $n = -0.855$, $m = 0.377$	4164	4798	0.122	0.118	0.401	0.430
17a-c	$a_1 = 0.001$, $a_2 = -0.196$, $a_3 = 27.987$ $b_1 = 0.004$, $b_2 = -1.017$, $b_3 = 147.275$	4772	6147	0.191	0.239	0.429	0.487
18	$a_1 = 0.007$, $a_2 = -0.399$, $a_3 = 124.513$ $b_1 = 0.000$, $b_2 = 0.003$, $b_3 = 0.257$	1743	2923	0.108	0.108	0.259	0.335

[*)]referred to F (Eq. 19), [**)]referred to F (Eq. 20).

tion should be performed. Because of the relatively high number of parameters and thus very low efficiency of systematic search method, and due to the polymodal nature of the objective function (recognized in preliminary tests) and thus low efficiency of local search methods, more efficient optimization technique is desirable. Therefore, the CRS2 method for global optimization was applied.

6.4 CRS2 method application

The CRS2 method was applied to find the parameters of a global IDF-relationship described by formulas:

for $t_d \leq 60$ min:
$$q = -a \ln p + b \ , \tag{24}$$

for $t_d \geq 60$ min:
$$q = a\, p^{-b} \ , \tag{25}$$

where
$$a = a_1\, t_d^2 + a_2\, t_d + a_3 \ , \tag{26}$$

$$b = b_1\, t_d^2 + b_2\, t_d + b_3 \ . \tag{27}$$

The parameters a_i and b_i ($i = 1, 2, 3$) may take different values for Eqs. 24 and 25, and thus the optimal values of twelve parameters were searched. The limitations for the values of parameters defining the search domain were assumed on the basis of preliminary tests (see Section 6.3). According to Eqs. 26 and 27, the parameters a_i and b_i ($i = 1, 2, 3$) are related to individual components of the quadratic relations, wherein the values of t_d vary in a relatively wide range (from 10 to 240 min). Moreover, a and b describe ether logarithmic or exponential relation between q and p. These factors cause the parameters a_i and b_i of significantly different magnitudes and variations in different ranges. Based on the preliminary stage of analyses, the limitations of the possible parameter values, defining the search domain, were finally determined as:

for $t_d < 60$ min:
$$a_1 \in <\!-10, 10\!>, \quad a_2 \in <\!-20, 20\!>, \quad a_3 \in <\!-2000, 2000\!>,$$
$$b_1 \in <\!-10, 10\!>, \quad b_2 \in <\!-10, 10\!>, \quad b_3 \in <\!-2000, 2000\!>;$$

for $t_d \geq 60$ min:
$$a_1 \in <\!0, 1\!>, \qquad a_2 \in <\!-20, 0\!>, \quad a_3 \in <\!0, 2000\!>,$$
$$b_1 \in <\!-0.1, 0.1\!>, \quad b_2 \in <\!-1, 1\!>, \qquad b_3 \in <\!0, 2\!>.$$

The surrounding of the most probable location of the optimum was deliberately extended into a wider domain, in case the solution found in Section 6.3 (with the use of relatively not effective methods) was not the global optimum.

In order to decrease the number of the parameters to be estimated, the optimization was carried out in two steps: in the first, Eq. 24 for rainfall data $t_d \leq 60$ min, in the second – Eq. 25 for $t_d \geq 60$ min were optimized. For random choice of the trial points, sampling with no replacement according to the uniform probability distribution was applied. The relation $N = 10\,(n + 1)$ was adopted and the points were stored in an array **A** with the use of bubble sort algorithm. As the end-of-optimization criterion the following formula was applied:

$$\left| F_{\mathrm{av}} - F_L \right| < eps \; , \tag{28}$$

where *eps* is the predetermined value, F_L is the value of objective function for the best point in array **A**, and F_{av} denotes the average value of objective function calculated for each of the N points of **A**:

$$F_{\mathrm{av}} = \frac{\displaystyle\sum_{j=1}^{N} F_j}{N} \; . \tag{29}$$

The criterion defines the situation in which the solution is not corrected any more in iterative process, as all next simplexes are constituted of the points of very similar values of objective functions. However, two additional stop criteria were also included, in order to enable to break the calculations if the total number of iterations or/and total number of objective function runs ("cost") was too high. However, the experiments proved the criterion 29 was conclusive in the analyzed problem.

The method was applied with *eps* = 0.05 at first. The CRS2 procedure was applied ten times to recognize the effectiveness of the optimization in terms of finding the global optimum. As the values of optimized parameters for each attempt were very similar, for the final CRS2 run the search domain was narrowed to the closest surrounding of the expected optimum and the value of *eps* was decreased to 0.01. A few representative examples of the results of optimization runs and the final optimal parameter sets for $t_d \leq$ 60 min and $t_d \geq 60$ min are presented in Table 7. The values indicate relatively small differences between the parameters (and thus also corresponding error values) obtained in various optimization runs. In fact, the method found global optimum (or at least a point very close to it) in each run. This may confirm the credibility of the results and possibility to decrease the total number of runs leading to satisfactory solution. Taking into account the total number of starting points (here: 70 points in each run), very short duration of calculations (a few minutes for each run) and the obtained result – the method may be considered efficient and the solution reliable.

Table 7

Exemplary results of parameter optimization
with use of CRS2 method for $t_d \leq 60$ min and $t_d \geq 60$ min

No.	Optimal values of parameters [—]						F (Eq. 20) [dm³/ (s·ha)]²	ε (Eq. 21) [dm³/ (s·ha)]	Cost (number of obj. function runs) [—]
	a_1	a_2	a_3	b_1	b_2	b_3			
$t_d \leq 60$ min									
eps = 0.05									
1	0.011302	−1.5006	69.279	0.0757	−9.0489	374.919	20 651.710	0.395	14 086
2	0.011283	−1.4986	69.245	0.0756	−9.0410	374.783	20 651.704	0.395	14 896
3	0.011273	−1.4985	69.250	0.0756	−9.0386	374.773	20 651.717	0.396	14 187
4	0.011298	−1.5004	69.262	0.0757	−9.0480	374.867	20 651.724	0.396	15 106
eps = 0.01									
5	0.011306	−1.5005	69.263	0.0757	−9.0471	374.845	20 651.668	0.395	6 069
$t_d \geq 60$ min									
eps = 0.05									
1	0.017787	−2.7625	247.618	0.0000	0.0029	0.272	1 585.657	0.247	13 188
2	0.017740	−2.7551	247.367	0.0000	0.0029	0.272	1 585.678	0.247	15 064
3	0.017738	−2.7494	246.919	0.0000	0.0029	0.271	1 585.687	0.247	14 044
4	0.017788	−2.7614	247.600	0.0000	0.0029	0.271	1 585.664	0.247	12 819
eps = 0.01									
5	0.017767	−2.7563	247.284	0.0000	0.0029	0.271	1 585.654	0.247	9 828

For the optimal set of parameters, in the case of $t_d \leq 60$ the value of F (Eq. 19) was 281.8 mm² and the values of E_1 and E_2, defined by Eqs. 22 and 23, respectively, were 0.093 and 0.100. The value of ε referred to F (Eq. 19) was 0.046 mm. In the case of $t_d \geq 60$ min, the value of F (Eq. 19) was 927.1 mm², the values of E_1 and E_2 were equal to 0.110 and 0.115, respectively, and ε (referred to F; Eq. 19) was equal to 0.189 mm.

Finally, the application of global optimization technique CRS2 enabled determination of the IDF-relationship for GUT station. The obtained formulas are:

for $t_d < 60$ min:
$$q = -a \ln p + b \ , \tag{30a}$$

where
$$a = 0.011306 \, t_d^2 - 1.500482 \, t_d + 69.2635 \ , \tag{30b}$$

$$b = 0.075719 \, t_d^2 - 9.04707 \, t_d + 374.845 \ , \tag{30c}$$

for $t_d \geq 60$ min:
$$q = a \cdot p^{-b} \ , \tag{31a}$$

where $\qquad a = 0.017767\, t_d^2 - 2.756281\, t_d + 247.28426$,   (31b)

$$b = 0.002902\, t_d + 0.271315 \; . \qquad (31c)$$

The obtained intensity-frequency curves for different values of aggregation level are shown in Fig. 7. Intensity-duration curves for selected values of p are presented in Fig. 8.
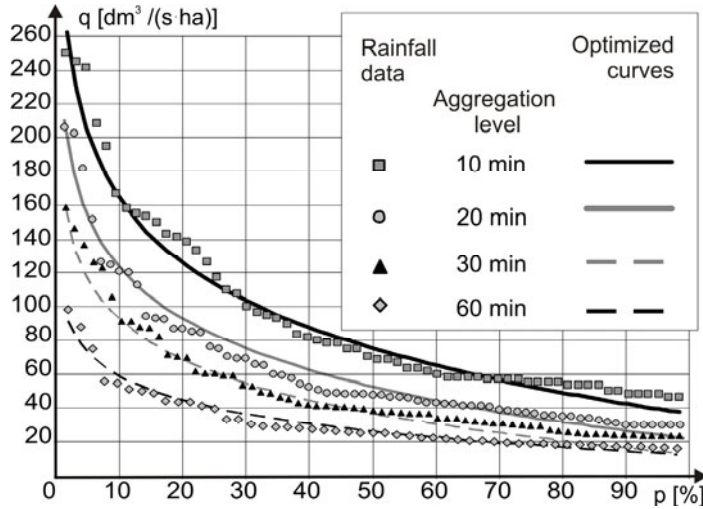
Fig. 7. Intensity-frequency curves for GUT station in Gdańsk (Poland).
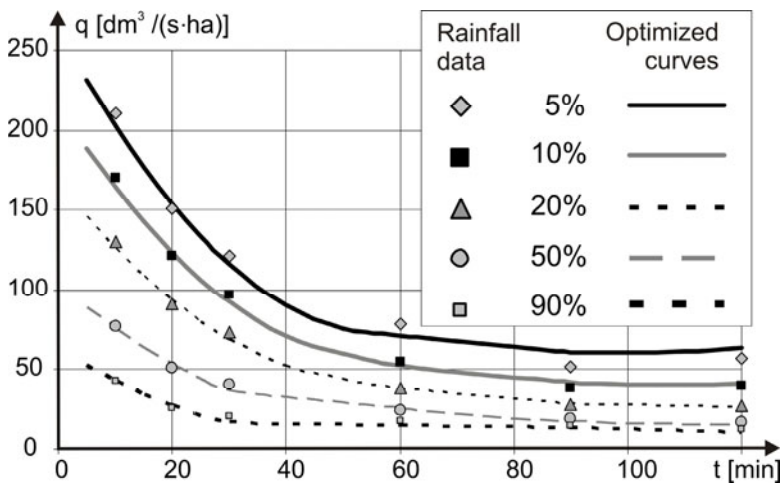
Fig. 8. Intensity-duration curves for GUT station in Gdańsk (Poland).

6.5 Comparison with other optimization methods

In order to verify the legitimacy of the CRS2 method application in the considered case, and for the purpose of its comparison with other algorithms, generally accepted for this type of problem, suitable numerical tests were carried out. In each of the tests the parameters a_1, a_2, a_3, b_1, b_2, and b_3 in Eqs. 26 and 27 were estimated (in the search domain defined as in the case presented in Section 6.4.) with use of different optimization method, but with the same accuracy requirements (defined in stopping criterion). Due to the different nature of the selected methods, the single iteration of the procedure was differently defined in each case. Thus, to enable objective comparison of the efficiency of the methods, the total cost of optimization, the values of the objective function F (Eq. 20), mean error ε (Eq. 21), and average relative errors E_1 (Eq. 22) and E_2 (Eq. 23) were compared for "optimal" sets of the estimated parameters. Each optimization method was applied at least ten times (increasing the number of runs was applied only if the results could affect the conclusions of the analysis). Selected representative examples of the results are presented in Table 8.

In the first stage, the parameters were estimated with the use of "successive search method". Because of a very low effectiveness of the algorithms, strongly depending on the size of the search domain and the length of the search step, the procedure was carried out in several stages, with successively narrowed search domain and reduced exploration step. However, even when the search step for each parameter was established only as 1/10 of the range of values defining boundaries of the search domain, in the case of six parameters to be estimated, the cost of 11^6 at each stage was needed

Table 8

Exemplary results of parameter optimization with use of systematic search (SS), Powell method (PM), modified Powell method (MPM), Newton optimization method (NOM), and evolutionary algorithm (EA) for $t_d \leq 60$ min and $t_d \geq 60$ min

Test no.	Optimal values of parameters [—]						F (Eq. 20) $[dm^3/(s{\cdot}ha)]^2$	ε (Eq. 21) $[dm^3/(s{\cdot}ha)]$	Cost (number of obj. function runs) [—]
	a_1	a_2	a_3	b_1	b_2	b_3			
	$t_d \leq 60$ min								
	Successive search method								
SS1	–0.018000	0.6200	37.200	–0.070	1.4600	224.000	101874.377	0.879	>7E+06
SS2	–0.016000	0.9536	22.400	–0.048	1.6960	179.200	138142.264	1.024	>8E+06
SS3	–0.040000	2.5920	11.200	–0.152	8.8000	128.000	172504.429	1.144	>91E+06
SS4	0.011000	–1.4800	69.200	0.075	–9.0000	375.000	20667.316	0.396	1.77E+06

to be continued

Table 8 (continuation)

					Powell method				
PM1	0.227188	–24.1799	414.772	0.8546	–91.7490	1627.687	6492794.640	7.010	385
PM2	–0.438900	15.7180	24.590	–1.3180	37.2449	388.400	19640656.910	12.210	379
PM3	–0.118034	12.3048	–141.716	0.5000	–18.9227	259.620	24533293.161	13.645	227
PM4	0.009999	–1.5000	69.382	0.0700	–9.0000	375.000	21606.116	0.405	62
					Modified Powell method				
MPM1	–0.147538	8.9843	–43.084	–0.3832	20.497	66.714	1002455.546	2.758	52917
MPM2	0.431536	–29.5952	375.202	1.8811	–135.7820	1826.212	9667618.690	8.566	54439
MPM3	0.502604	–36.2514	582.779	1.3530	–102.5665	1918.379	16688799.303	11.244	56279
MPM4	–0.021248	–6.1265	214.418	–0.4836	1.5628	631.197	9628908.819	8.548	25398
					Newton search method				
NOM1	0.014780	–1.7474	72.135	0.0891	–10.0000	385.922	20902.514	0.398	120
NOM2	0.014609	–1.7487	72.613	0.0884	–9.9999	387.699	20946.700	0.399	252
NOM3	–0.058240	3.4380	12.068	–0.1925	10.0000	154.274	138682.640	1.026	120
NOM4	0.011284	–1.4992	69.259	0.0756	–9.0400	374.831	20603.962	0.395	672
					Evolutionary algorithm				
EA1	–0.023090	1.0852	35.403	–0.0478	0.2859	251.027	54802.212	0.645	28780*
EA2	–0.014320	0.3238	47.772	–0.0214	–2.137	293.491	36249.521	0.524	28461*
EA3	–0.023595	1.0843	36.569	–0.0497	0.282	256.861	52920.362	0.634	31302*
EA4	–0.022700	0.9357	40.762	–0.0530	0.168	266.765	48456.261	0.606	16070*
					$t_d \geq 60$ min				
					Successive search method				
SS5	0.000800	–0.2400	172.800	0.000	0.0000	0.534	5412.163	0.456	>7E+06
SS6	0.000000	–0.1280	172.800	0.000	0.0000	0.533	5509.966	0.461	>7E+06
SS7	0.000000	–0.2000	170.000	0.000	0.0000	0.500	5713.677	0.469	>8E+06
SS8	0.017000	–2.7000	248.000	0.000	0.0028	0.280	1596.869	0.248	1.77E+06
					Powell method				
PM5	0.017000	–2.7500	251.766	0.0000	0.0029	0.270	1703.623	0.256	1541
PM6	–0.118000	10.500	591.000	0.0000	0.0000	1.000	761966.918	5.422	20013
PM7	0.019619	–1.4638	238.515	0.0007	0.0030	–1.979	91741.570	1.881	139
PM8	0.638100	–2.1253	515.044	0.0818	0.8810	1.831	143716.983	2.355	1329
					Modified Powell method				
MPM5	–0.232999	–9.3968	1420.001	0.0025	–0.1682	0.928	102248.066	1.986	105200
MPM6	0.012747	–1.0791	238.299	0.0007	0.0029	–1.964	87257.190	1.835	72739
MPM7	0.148486	–1.8437	667.270	0.0073	–0.0738	0.338	133992.829	2.274	66990
MPM8	0.019999	–3.3971	309.408	0.0000	0.0000	0.682	32287.420	1.116	66095
					Newton search method				
NOM5	0.007125	–0.3991	124.513	0.000	0.0028	0.257	2923.069	0.336	492
NOM6	0.999900	–6.5091	1588.983	0.000	0.0977	0.000	133526.940	2.270	48
NOM7	0.000000	–20.0000	0.000	0.000	1.0000	0.000	143717.015	2.355	24
NOM8	0.017756	–2.7516	246.986	0.000	0.0029	0.270	1586.663	0.247	288
					Evolutionary algorithm				
EA5	0.059317	–9.9158	556.111	0.000	0.0047	0.211	5280.886	0.452	39705*
EA6	0.038975	–5.7779	357.446	0.000	0.0047	0.174	3007.959	0.341	230259*
EA7	0.036106	–6.1234	400.184	0.000	0.0036	0.254	2833.594	0.330	28936*
EA8	0.029875	–4.8584	344.224	0.000	0.0036	0.227	2083.982	0.284	33705*

*)number of "subproblems" (solver evolutionary algorithm)

to find the "optimum", while the accuracy of such optimization was still very low. The effectiveness of the algorithm depended on the search step, the manner in which the search domain was successively narrowed at each subsequent stage of the calculations and the total number of stages. However, in all the tests conducted the process was very time-consuming and did not guarantee the success in finding the global optimum. In many cases the global optimum was omitted due to relatively large search step (Table 8, tests SS1, SS2, SS3, SS5, SS6, and SS7). The algorithm enabled obtaining the global optimum (with the objective function only slightly higher than in the case of CRS2 method) only in the case of the search domain initially narrowed to the closest neighborhood of the global optimum (Table 8, cases SS4 and SS8). However, even then the total cost of such search was much higher compared to CRS2 and other methods (more than a million objective function evaluations). In a typical situation, such precise localization of the global optimum is not known. Therefore, the only way to obtain a satisfactory result with the use of this method would be one-stage-search with a very small step. However, the cost of such calculation and the time required to obtain the solution would be unacceptable. Therefore, this algorithm may be useful only for rough estimation at the preliminary stage of the calculations and is not recommended in this case.

The second approach to the considered optimization problem was the "Powell method" implementation (Powell 1964, Fletcher 1965, Pierre 1969, Press *et al.* 2007). The typical, widely known concept of the conjugate directions was applied (Powell 1964), according to the procedures presented by Press *et al.* (2007). The initial bracketing of a minimum and the golden section search algorithm was applied as the procedures of one-dimension search for each of the conjugate directions (Press *et al.* 2007). The method required the specification of the initial point from which the searching procedure would start, thus the tests with different starting points were carried out. In each single iteration, the one-dimensional search along each of the conjugate directions was carried out, with the number of objective function evaluations depending on the function properties and required accuracy.

In the case of Powell method application, no satisfactory solution was obtained, as the impact of multimodal character of the objective function was clearly highlighted. In each case of the method run (from different starting points), premature convergence to the closest local optimum was observed. Thus, depending on the place in the search domain from which the optimization procedure was started, different optima were obtained (Table 8, exemplary tests PM1-PM8). The cost of the optimization was relatively low. Despite this, the quality of the solution was unacceptable. The values of the objective function in "optimal" points for most procedure runs were high. For example, for $t_d \leq 60$ min F was even higher than 24 000 000 (Table 8,

test PM3), while the corresponding value for the CRS2 algorithm was 20 651. The relatively good solution was obtained only when the starting point was close to the global optimum (tests PM4 and PM5). In typical situations, such a "lucky" guess is very difficult or impossible to provide. Therefore, the local search procedure (Powell method or any other) is not recommended in the analyzed problem.

In order to improve the quality of optimization using Powell method, the modification of the procedure was introduced. The algorithm of the local search (Powell method) was combined with random procedure of selecting the initial points in the manner similar to initial step of the CRS2 method. In this approach, the initial set of $N = 10 (n + 1)$ randomly chosen points was established and the Powell method was run from each of the points. Such a test was then repeated several times. Unfortunately, as there was no additional mechanism of random interference in the optimization process (apart from the initial stage), the success of such approach depended on the "chance" of the draw of the initial points (Table 8, exemplary tests MPM1-MPM8) and the algorithm still suffered from all the inconveniences described before. Additionally, the cost of such optimization was much higher compared with CRS2 method. In all tests the solution was of much worse quality than the one obtained with the use of the CRS2 method, which enabled the solution very close to the global optimum in each run.

The final step was to test the effectiveness of the optimization procedures accessible in the embedded solver tools of commonly used spreadsheet packages. The two procedures –Newton search and the evolutionary algorithm – were applied. In the case of Newton method, the procedure was very fast and enabled obtaining "optimal" solution (in the sense of local optimization) in a short time, but the quality (effectiveness) of such optimization (expressed in the value of the objective function calculated for the optimal point) depended on the initial point of the optimization process. Initial points for the tests were selected either deliberately (*e.g.*, the central point of the domain, points close to the global optimum, points previously selected for CRS2 or Powell methods) or randomly. The procedure proved more effective than the Powell method. For some of the starting points, for which the Powell procedure suffered from the premature convergence, the satisfactory solution was obtained. In some cases, the optimum found by the procedure was very close to the one obtained by the CRS2 method (Table 8, tests NOM1-NOM2, NOM5). The further improvement was possible (*e.g.*, Table 8, best results – tests NOM4 and NOM8) when the additional run of the algorithm was carried out started from the previously found "optimum". However, for many other trials, in which the starting points were less favorable, the false "optimum" was found (with the value of objective function close to 100 000 or more, *e.g.*, tests NOM3, NOM6, NOM7). The tests have

shown that this method of optimization may be effective, but the success in finding the global optimum depends on the starting point from which the procedure is initiated.

The comparison of the two approaches – CRS2 method (used once) and Newton search (applied many times) – shows several significant differences. In the latter approach, each single run is much quicker, however does not guarantee the success (in terms of the "guarantee" that non-deterministic method CRS2 can provide; see Section 6.4). Moreover, the restart of Newton search – in the approach applied in the study – suffers from "subjectivity" and some onerousness, because each starting point must be selected by the user, which may be done according to different criterions. What is more, in a typical situation (when the global optimum is not known), it is difficult to judge when to stop restarting the procedure of local optimization and how "good" the so-far "optimal" solution really is. The CRS2 method is more "objective", as it proposes random selection of a relatively high number of initial points (recognized by Price (1983) as sufficient). It also leads to satisfactory solution in each run. In order to compare the results of the methods in the same initial conditions, $10\,(n+1)$ runs of Newton search should be performed (for each run of CRS2 method). That would be much more time-consuming and burdensome, and thus the automation would be necessary. The "cost" of such optimization would be comparable but the time of calculations – longer.

To the authors mind, the CRS2 method is more reliable. The Newton search approach is worth considering, especially in relatively simple applications; however, it should be applied carefully, as there is always the risk of premature convergence.

The last of the analyzed approaches was the application of the evolutionary algorithm available in Excel 2010 package. Unfortunately, this method occurred to be ineffective in the case considered. The evolutionary procedure conducted on the population of $N=10\,(n+1)=70$ individuals, did not lead to global optimum in any of the tests (more than 10 trials for each t_d range), although in some of them it was relatively close, compared to successive search and Powell method (Table 8, exemplary results EA1-EA8). As the common spreadsheet packages enable only limited influence on the parameters of the procedures, the improvement of the effectiveness in this case is very difficult. On the other hand, the considered problem (the estimation of the IDF-formula parameters) is not complex enough to apply very sophisticated methods of optimization, especially when other algorithms (*e.g.*, the CRS2 procedure) occurred to be effective. The tests showed that the best approach to the case considered was the method which combined both random and deterministic search in order to overcome all the inconveniences encountered in the calculations.

7. SUMMARY AND CONCLUSIONS

IDF-relationships, in spite of their "simple" nature, play an important role in many engineering applications. The choice of IDF-formulas for design rainfall determination instead of more sophisticated analysis, is often a compromise between practical tendency to use simplified approaches (leading to solution in short, relatively easy and economically beneficial way), and the problem complexity and computational requirements. The presented study has developed the modified empirical approach to IDF-relationship development, based on the global optimization technique of controlled random search applied to the estimation of the parameters of the previously determined functions. The approach is relatively simple and efficient and it enables obtaining satisfactory solution in a relatively short time. In the problem of IDF-formulas determination – seen globally, as an engineering task – the most time-consuming is the first step – rainfall data collection and analysis. Currently available methods of parameter search may enable quick access to efficient solutions, even (or especially) in those cases where the global form of IDF-formula is more complex and typical theoretical distributions do not lead to satisfactory results, as the measured data do not "match" typical hydrological distributions.

The proposed procedure of IDF-relationship determination has proven effective in the case when commonly used simple formulas do not allow an acceptably good fit between simulated rainfall intensities and observations in the whole range of considered values of t_d and p. In this particular case, the "combined" relationship $q(t_d, p)$, in the form more complex than commonly used IDF-formulas, had to be taken into account and the higher number of parameters had to be identified. Therefore, global optimization method was applied. The proposed procedure CRS2 is obviously not the only one possible to use in this case. Numerous experiences in different fields of engineering enabled the development of many optimization procedures, from relatively simple to very sophisticated, dedicated to different kinds of problems. Therefore, the choice of the method should result from the complexity of the task (mainly the features of the objective function), as the chosen algorithm should enable obtaining satisfactory solution with reasonable labor input, corresponding to the problem complexity. The CRS2 algorithm applied in the study was effective in finding global minimum of the assumed function, although it was still simple, compared to further versions of the procedure (CRS3 and higher) and other global optimization techniques. Commonly used approaches to IDF-relationship determination (based on formulas 1-3) produced worse results in the considered case. Although the solution was achieved relatively easily and quickly, the goodness-of-fit of the obtained relationships to observations was much poorer, which was ana-

lyzed in the preliminary stage of IDF-formulas determination. Local optimization procedures (*e.g.*, the widely recommended Powell method) applied to parameter identification in case of more complex IDF-function, led to different, not always optimal solutions, depending on the assumed starting point, due to premature convergence to local optima. Systematic search procedure finally led to the correct result but the process was much more time-consuming. Evolutionary algorithm applied to the analyzed case was also ineffective. Such algorithms are also much more time-consuming and – because of their random nature – it is very difficult to judge whether the solution is really "optimal" in global meaning. The effectiveness of evolutionary algorithms may be improved by detailed recognition of the most suitable (to the analyzed case) procedures of selection, mutation and crossover. However, for the purpose of calculations considered in this paper (IDF-formulas determination), such effort seems redundant and application of such methods – unfounded. The CRS methods, which had been analyzed in detail and compared with selected global search procedures (Ali *et al.* 1997a), proved to be effective in many complex applications. Therefore, the choice of CRS2 in the study was a compromise between labor input and the expected quality of the results.

R e f e r e n c e s

Akan, A.O., and R.J. Houghtalen (2003), *Urban hydrology, Hydraulics, and Stormwater Quality: Engineering Applications and Computer Modeling*, John Wiley & Sons, Inc., Hoboken.

Ali, M.M. (1994), Some modified stochastic global optimization algorithms with applications, Ph.D. Thesis, Loughborough University of Technology, Loughborough, England, https://dspace.lboro.ac.uk/2134/13429.

Ali, M.M., and C. Storey (1994), Topographical multilevel single linkage, *J. Global Optim*. **5**, 4, 349-458, DOI: 10.1007/BF01096684.

Ali, M.M., and C. Storey (1997), Aspiration based simulated annealing algorithm, *J. Global Optim*. **11**, 2, 181-191, DOI: 10.1023/A:1008202703889.

Ali, M.M., C. Storey, and A. Törn (1997a), Application of stochastic global optimization algorithms to practical problems, *J. Optimiz. Theory Appl*. **95**, 3, 545-563, DOI: 10.1023/A:1022617804737.

Ali, M.M., A. Törn, and S. Viitanen (1997b), A numerical comparison of some modified controlled random search algorithms, *J. Global Optim*. **11**, 4, 377-385, DOI: 10.1023/A:1008236920512.

Ariff, N.M., A.A. Jemain, K. Ibrahim, and W.Z. Wan Zin (2012), IDF relationships using bivariate copula for storm events in Peninsular Malaysia, *J. Hydrol*. **470-471**, 158-171, DOI: 10.1016/j.jhydrol.2012.08.045.

Arnell, V. (1982), Rain fall data for the design of sewer pipe systems, Report Series A:8, Chalmers University of Technology, Department of Hydraulics, Göteborg, Sweden.

Ben-Zvi, A. (2009), Rainfall intensity-duration-frequency relationship derived from large partial duration series, *J. Hydrol*. **367**, 1-2, 104-114, DOI: 10.1016/j.jhydrol.2009.01.007.

Bonnans, J.F., J.C. Gilbert, C. Lemaréchal, and C.A. Sagastizábal (2006), *Numerical Optimization: Theoretical and Practical Aspects*, Springer-Verlag, Berlin Heidelberg.

Chow, V.T. (1964), *Handbook of Applied Hydrology: A Compendium of Water-resources Technology*, McGraw–Hill Book Co., New York.

Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics, Springer, London, 208 pp.

Cunnane, C. (1973), A particular comparison of annual maxima and partial duration series methods of flood frequency prediction, *J. Hydrol*. **18**, 3-4, 257-271, DOI: 10.1016/0022-1694(73)90051-6.

De Michele, C., and G. Salvadori (2005), Some hydrological applications of small sample estimators of Generalized Pareto and Extreme Value distributions, *J. Hydrol*. **301**, 1-4, 37-53, DOI: 10.1016/j.jhydrol.2004.06.015.

Dekkers, A., and E. Aarts (1991), Global optimization and simulated annealing, *Math. Program*. **50**, 1-3, 367-393, DOI: 10.1007/BF01594945.

Delleur, J.W. (2003), The evolution of urban hydrology: Past, present, and future, *J. Hydraul. Eng. ASCE* **129**, 8, 563-573, DOI: 10.1061/(ASCE)0733-9429(2003)129:8(563).

Dysarz, T., and J.J. Napiórkowski (2002), Flood control in Nysa Reservoir System by means of sequential optimisation and CRS method. **In:** *Proc. Workshop on Evolutionary Algorithms and Global Optimization, 23-25 September 2002, Cracow, Poland*, Warsaw Technical University Publ., Warszawa, 27-33.

Dysarz, T., and J.J. Napiórkowski (2003), Optimal flood control of Nysa Kłodzka Reservoir System, *Publs. Inst. Geophys. Pol. Acad. Sc*. **E-3**, 365, **(**J.J. Napiórkowski (ed.), *Modelling and Control of Floods*), 83-96.

El-Sayed, E.A.H. (2011), Generation of rainfall intensity duration frequency curves for ungauged sites, *Nile Basin Water Sci. Eng. J.* **4**, 1, 112-124.

Elsebaie, I.H. (2012), Developing rainfall intensity-duration-frequency relationship for two regions of Saudi Arabia, *J. King Saud Univ. – Eng. Sci.* **24**, 2, 131-140, DOI: 10.1016/j.jksues.2011.06.001.

Endreny, T.A., and N. Imbeah (2009), Generating robust rainfall intensity-duration-frequency estimates with short-record satellite data, *J. Hydrol*. **371**, 1-4, 182-191, DOI: 10.1016/j.jhydrol.2009.03.027.

Fletcher, R. (1965), Function minimization without evaluating derivatives – a review, *Comput. J*. **8**, 1, 33-41, DOI: 10.1093/comjnl/8.1.33.

Fylstra, D., L. Lasdon, J. Watson, and A. Waren (1998), Design and use of the Microsoft Excel Solver, *Interfaces* **28**, 5, 29-55, DOI: 10.1287/inte.28.5.29.

Goldberg, D.E. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison–Wesley Publ. Co., Reading, 412 pp.

Grace, R.A., and P.S. Eagleson (1967), A model for generating synthetic sequences of short-time-interval rainfall depths. **In**: *Proc. Int. Hydrology Symposium, 6-8 September 1967, Fort Collins, Colorado, USA*, Vol. 1, Colorado State University, Colorado, 268-276.

Grimaldi, S., and F. Serinaldi (2006), Design hyetograph analysis with 3-copula function, *Hydrol. Sci. J.* **51**, 2, 223-238, DOI: 10.1623/hysj.51.2.223.

Hailegeorgis, T.T., S.T. Thorolfsson, and K. Alfredsen (2013), Regional frequency analysis of extreme precipitation with consideration of uncertainties to update IDF curves for the city of Trondheim, *J. Hydrol*. **498**, 305-318, DOI: 10.1016/j.jhydrol.2013.06.019.

Johansen, L. (1979), Design rainfalls for sewer systems, Rep. 79-2, Dept. of Sanitary Eng., Technical University of Denmark, Copenhagen, Denmark.

Katz, R.W., M.P. Parlange, and P. Naveau (2002), Statistics of extremes in hydrology, *Adv. Water Resour*. **25**, 8-12, 1287-1304, DOI: 10.1016/S0309-1708(02)00056-8.

Kisiel, C.C., L. Duckstein, and M.M. Fogel (1971), Analysis of ephemeral flow in aridlands, *J. Hydraul. Div. ASCE* **97**, 10, 1699-1717.

Koutsoyiannis, D. (2004a), Statistics of extremes and estimation of extreme rainfall: I. Theoretical investigation, *Hydrol. Sci. J.* **49**, 4, 575-590, DOI: 10.1623/hysj.49.4.575.54430.

Koutsoyiannis, D. (2004b), Statistics of extremes and estimation of extreme rainfall: II. Empirical investigation of long rainfall records, *Hydrol. Sci. J.* **49**, 4, 591-610, DOI: 10.1623/hysj.49.4.591.54424.

Koutsoyiannis, D., D. Kozonis, and A. Manetas (1998), A mathematical framework for studying rainfall intensity-duration-frequency relationships, *J. Hydrol.* **206**, 1-2, 118-135, DOI: 10.1016/S0022-1694(98)00097-3.

Langousis, A., and D. Veneziano (2007), Intensity-duration-frequency curves from scaling representations of rainfall, *Water Resour. Res.* **43**, 2, W02422, DOI: 10.1029/2006WR005245.

Licznar, P., and J. Łomotowski (2005), Analysis of instantaneous intensities of design rainfalls in Wroclaw, *Ochr. Sr.* **27**, 2, 25-28 (in Polish).

Licznar, P., and J. Łomotowski (2007), *Rainfall Kinetic Energy Measurements with Impactometer Implementation*, Works and Studies, Vol. 69, Institute of Environmental Engineering, Polish Academy of Sciences, Zabrze, 70 pp.

Licznar, P., T.G. Schmitt, and D.E. Rupp (2011a), Distributions of microanonical cascade weights of rainfall at small timescales, *Acta Geophys.* **59**, 5, 1013-1043, DOI: 10.2478/s11600-011-0014-4.

Licznar, P., J. Łomotowski, and D.E. Rupp (2011b), Random cascade driver rainfall disaggregation for urban hydrology: An evaluation of six models and a new generator, *Atmos. Res.* **99**, 3-4, 563-578, DOI: 10.1016/j.atmosres.2010.12.014.

Llasat, M.-C. (2001), An objective classification of rainfall events on the basis of their convective features: application to rainfall intensity in the northeast of Spain, *Int. J. Climatol.* **21**, 11, 1385-1400, DOI: 10.1002/joc.692.

Lutgens, F.K., and E.J. Tarbuck (2004), *The Atmosphere. An Introduction to Meteorology*, 9th ed., Pearson Education Inc., New Jersey.

Madsen, H., P.F. Rasmussen, and D. Rosbjerg (1997a), Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 1. At-site modeling, *Water Resour. Res.* **33**, 4, 747-757, DOI: 10.1029/96WR03848.

Madsen, H., C.P. Pearson, and D. Rosbjerg (1997b), Comparison of annual maximum series and partial duration series methods for modeling extreme hydrologic events: 2. Regional modeling, *Water Resour. Res.* **33**, 4, 759-769, DOI: 10.1029/96WR03849.

Manzanares-Filho, N., and R.B.F. Albuquerque (2008), Accelerating controlled random search algorithms using a distribution strategy. **In**: *Proc. Int. Conf. Engineering Optimization (EngOpt 2008), 1-5 June 2008, Rio de Janeiro, Brazil*.

Manzanares-Filho, N., C.A.A. Moino, and A.B. Jorge (2005), An improved controlled random search algorithm for inverse airfoil cascade design. **In:** *Proc. 6th World Congress of Structural and Multidisciplinary Optimization (WCSMO-6), 30 May – 3 June 2005, Rio de Janeiro, Brazil*, paper No. 4451.

Maraun, D., T.J. Osborn, and N.P. Gillett (2008), United Kingdom daily precipitation intensity: improved early data, error estimates and an update from 2000 to 2006, *Int. J. Climatol.* **28**, 6, 833-842, DOI: 10.1002/joc.1672.

Marsalek, J. (1978), Research on the design storm concept, ASCE Urban Water Resources Research Program, Technical Memorandum No. 33, New York.

McCuen, R.H. (2005), *Hydrologic Analysis and Design*, 3rd ed., Pearson Prentice Hall, Englewood Cliffs, 859 pp.

Molnar, P., and P. Burlando (2005), Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model, *Atmos. Res.* **77**, 1-4, 137-151, DOI: 10.1016/j.atmosres.2004.10.024.

Mutzner, H. (1991), The significance of areal rainfall distribution for flows from a very small urban drainage catchment, *Atmos. Res*. **27**, 1-3, 99-107, DOI: 10.1016/0169-8095(91)90011-K.

Nagy, J.T. (1991), Evaluation of rainfall characteristics in Bratislava, *Atmos. Res*. **27**, 1-3, 209-217, DOI: 10.1016/0169-8095(91)90020-W.

Pierre, D.A. (1969), *Optimization Theory with Applications*, John Wiley &Sons Inc., New York.

Powell, M.J.D. (1964), An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Comput. J.* **7**, 2, 155-162, DOI: 10.1093/comjnl/7.2.155.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery (2007), *Numerical Recipes: the Art of Scientific Computing*, 3rd ed., Cambridge University Press, New York.

Price, W.L. (1977), A controlled random search procedure for global optimisation, *Comput. J*. **20**, 4, 367-370, DOI: 10.1093/comjnl/20.4.367.

Price, W.L. (1978), Controlled random search procedure for global optimization. **In:** L.C.W. Dixon, and G.P. Szegö (eds.), *Toward Global Optimization 2*, North-Holland Publ. Co., Amsterdam, 71-84.

Price, W.L. (1983), Global optimization by controlled random search, *J. Optimiz. Theor. Appl*. **40**, 3, 333-348, DOI: 10.1007/BF00933504.

Pui, A., A. Sharma, R. Mehrotra, B. Sivakumar, and E. Jeremiah (2012), A comparison of alternatives for daily and sub-daily rainfall disaggregation, *J. Hydrol*. **470-471**, 138-157, DOI: 10.1016/j.jhydrol.2012.08.041.

Reiss, R.-D., and M. Thomas (2007), *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*, 3rd ed., Birkhäuser Verlag, Basel.

Rinaudo, S., F. Moschella, O. Muscato, and M.A. Ahile (1998), Controlled random search parallel algorithm for global optimization with distributed processes on multivendor CPUs. **In:** L. Arkeryd, J. Bergh, Ph. Brenner, and R. Petterson (eds.), *Proc. 10th Conf. European Consortium for Mathematics in Industry "Progress in Industrial Mathematics at ECMI 98", 23-28 June 1998, Gothenburg, Sweden*, 271-278.

Rinnooy Kan, A.H.G., and G.T. Timmer (1987), Stochastic global optimization methods. Part II: Multi level methods, *Math. Program*. **39**, 1, 57-78, DOI: 10.1007/BF02592071.

Romanowicz, R.J., M. Osuch, and M. Grabowiecka (2013), On the choice of calibration periods and objective functions: A practical guide to model parameter identification, *Acta Geophys*. **61**, 6, 1477-1503, DOI: 10.2478/s11600-013-0157-6.

Schilling, W. (1991), Rainfall data for urban hydrology: what do we need?, *Atmos. Res.* **27**, 1-3, 5-21, DOI: 10.1016/0169-8095(91)90003-F.

Sevruk, B., and H. Geiger (1987), Frequency distributions preferred by hydrologists. **In**: B.C. Yen (ed.), *Proc. 4th Int. Conf. in Urban Storm Drainage "Topics in Urban Drainage Hydraulics and Hydrology", Session D, 31 August – 4 September 1987, Lausanne, Switzerland*, 51-52.

Suligowski, R. (2004), Temporal and spatial structure of the rainfall in Poland. Attempt at the regionalization, Publ. of Jan Kochanowski University, Institute of Geography, Kielce (in Polish).

Svensson, C., R.T. Clarke, and D.A. Jones (2007), An experimental comparison of methods for estimating rainfall intensity-duration-frequency relations from fragmentary records, *J. Hydrol.* **341**, 1-2, 79-89, DOI: 10.1016/j.jhydrol.2007.05.002.

Tanaka, S., and K. Takara (2002), A study of threshold selection in POT analysis of extreme floods. **In:** A. Snorasson, H.P. Finnsdóttir, and M. Moss (eds.), *Proc. Symp. "The Extremes of the Extremes: Extraordinary Floods", Reykjavik, Iceland*, IAHS Publications, No. 271, IAHS Press, Wallingford, 299-306.

Tavares, L.V., and J.E. Da Silva (1983), Partial duration series method revisited, *J. Hydrol.* **64**, 1-4, 1-14, DOI: 10.1016/0022-1694(83)90056-2.

Todorovic, P. (1978), Stochastic models of floods, *Water Resour. Res.* **14**, 2, 345-356, DOI: 10.1029/WR014i002p00345.

Trefry, C.M., D.W. Watkins, and D. Johnson (2005), Regional rainfall frequency analysis for the State of Michigan, *J. Hydrol. Eng. ASCE* **10**, 6, 437-449, DOI: 10.1061/(ASCE)1084-0699(2005)10:6(437).

Tsoulos, I.G., and I.E. Lagaris (2006), Genetically controlled random search: a global optimization method for continuous multidimensional functions, *Comp. Phys. Comm.* **174**, 2, 152-159, DOI: 10.1016/j.cpc.2005.09.007.

Vaes, G., P. Willems, and J. Berlamont (2001), Rainfall input requirements for hydrological calculations, *Urban Water* **3**, 1-2, 107-112, DOI: 10.1016/S1462-0758(01)00020-6.

Venkata Ramana, R., B. Chakravorty, N.R. Samal, N.G. Pandey, and P. Mani (2008), Development of intensity duration frequency curves using L-moment and GIS technique, *J. Appl. Hydrol.* **21**, 1-2, 88-100.

Vukmirović, V., and J. Petrović (1991), Statistical analysis of storms – a basis for urban runoff modeling. **In:** Č. Maksimović (ed.), *Proc. of Int. Conf. "New Technologies in Urban Drainage UDT'91", 17-21 June 1991, Dubrovnik, Yugoslavia*, Elsevier, London, 13-19.

Walesh, S.G. (1989), *Urban Surface Water Management*, John Wiley & Sons, Inc., New York.

Wenzel, H.G. Jr., and M.L. Voorhees (1978), Evaluation of design storm concept (presented at the AGU Fall Meeting, December 1978, San Francisco, California), *EOS Trans. AGU* **59**, 12, 1070.

Willems, P. (2000), Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types, *J. Hydrol*. **233**, 1-4, 189-205, DOI: 10.1016/S0022-1694(00)00233-X.

Wołoszyn, E. (1991), Polish rainfall-runoff investigations and modification of the rational method, *Atmos. Res*. **27**, 1-3, 219-229, DOI: 10.1016/0169-8095 (91)90021-N.

Wołoszyn, E. (2003), The catastrophic flood in Gdańsk on July 2001. **In:** R. Arsov, J. Marsalek, E. Watt, and E. Zeman (eds.), *Urban Water Management: Science Technology and Service Delivery*, NATO Science Series, Vol. 25, Springer, Dordrecht, 115-124, DOI: 10.1007/978-94-010-0057-4_12.

Wołoszyn, E. (2009), Analysis of rainfall data of Gdansk Meteorological Station. **In:** C. Popowska and M. Jovanovski (eds.), *Proc. 11th Int. Symp. on Water Management and Hydraulic Engineering, 1-5 September 2009, Univ. Ss. Cyril and Methodius, Faculty of Civil Engineering, Orhid, Macedonia,* 675-684.

Yevjevich, V. (1972), *Probability and Statistics in Hydrology*, Water Resources Publ., Fort Collins, USA.