Hydrol. Earth Syst. Sci. Discuss., 11, 5251–5285, 2014
www.hydrol-earth-syst-sci-discuss.net/11/5251/2014/
doi:10.5194/hessd-11-5251-2014

Hydrology and Earth System Sciences
Open Access
Discussions

Precipitation variability within an urban monitoring network in terms of microcanonical cascade generators

P. Licznar[1], C. De Michele[2], and W. Adamowski[3]

[1]Institute of Environment Protection Engineering, Wrocław University of Technology, Wrocław, Poland
[2]Department of Civil and Environmental Engineering, Politecnico di Milano, Italy
[3]Institute of Environmental Engineering, John Paul II Catholic University of Lublin, Stalowa Wola, Poland

Received: 3 May 2014 – Accepted: 12 May 2014 – Published: 20 May 2014

Correspondence to: P. Licznar (pawel.licznar@pwr.wroc.pl)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Understanding the variability of precipitation at small scales is fundamental in urban hydrology. Stochastic models of precipitation are required to feed hydrodynamic models with high resolution data, in order to obtain a probabilistic assessment of urban drainage networks. Microcanonical random cascades are considered here to represent precipitation time series collected in 25 gauges of a monitoring network in Warsaw, Poland. Breakdown coefficients (BDCs) are calculated separately for a hierarchy of subdaily timescales from 5 min (time resolution) to 1280 min, for all gauges. Strong deformations of BDC histograms in form of sharp peaks at small timescales are observed due to the truncation of precipitation depths recorded by gauges. Satisfactory smoothing of empirical BDC histograms is obtained statistically by a slight randomization of nonzero precipitation amounts. The scarce representation of BDCs at large timescales, due to the short period of observation, is solved by the introduction of an algorithm based on overlapping moving windows. BDC histograms are modeled by a 2N-B distribution, which combines two Normal (N) and one Beta (B) distribution. A clear evolution of the distribution from 2N-B at small timescales, to N-B at intermediate timescales, and finally to Beta distribution for large timescales is observed in all gauges. The performance of the microcanonical cascades is evaluated for the considered gauges. Synthetic time series are analyzed with respect to their intermittency and variability of intensity, and compared to observed series. BDC histograms, for each timescale, are compared among the 25 gauges in Warsaw, and with other gauges located in Poland and Germany. The cluster analysis is used to identify patterns of BDC histograms among analyzed set of gauges and timescales, as well as to detect outlier gauges.

HESSD

11, 5251–5285, 2014

1 Introduction

Urban hydrology demands the access to very precise information about the precipitation variability over small spatial and temporal scales. Widespread use of surface runoff models coupled to urban drainage networks increases the common request for inputs in the form of precipitation time series at high temporal resolution (∼ 1 min) and with a considerable record length (at least 20–30 years). Currently this quality of precipitation datasets is required especially from European perspective for the probabilistic assessment of the urban drainage network functioning (Schmitt, 2000; European standard EN 752), or the probabilistic assessment of retention volumes at hydraulic overloaded stormwater systems (Arbeitsblatt DWA-A 117).

From engineering perspective, the strategy of using local precipitation time series as basis of the probabilistic assessment of urban drainage systems has two important shortcomings. First of all, in case of local precipitation data shortage, this strategy fails completely. Whereas, in all other situations, when some local precipitation datasets are accessible, questions and doubts about the representativeness and reliability of data arise. Synthetic time series, generated from precipitation models, could be considered as probable precipitation scenarios to feed hydrodynamic urban drainage system models.

Thus there is a strong motivation for the development of local precipitation models at high temporal resolutions. Many of them are based on the idea of precipitation disaggregation in time. General overview of this class of models is given by Koutsoyiannis (2003). A central role is reserved to random cascade models, and especially to microcanonical cascade models (MCMs). The popularity of the latter ones could be explained by their appealing towards engineering applications, the assumptions like the mass conservation (i.e. rainfall depth conservation) across cascade levels, and straight rules for the extraction of cascade generators from local precipitation time series (Cârsteanu and Foufoula-Georgiou, 1996). Olsson (1998), Menabde and Sivapalan (2000), Ahrens (2003), Paulson and Baxter (2007) provide

HESSD
11, 5251–5285, 2014

contributions demonstrating the potentiality of MCMs in rainfall downscaling. Molnar and Burlando (2005) and Hingray and Ben Haha (2005) highlight the application of MCMs in urban hydrology. Hingray and Ben Haha (2005) applied a continuous hydrological simulation to produce from synthetic rainfall series a continuous discharge series used afterwards for the retention design. Recently, Licznar (2013) illustrated the possibility of substituting synthetic time series generated from MCMs to observed time series for the probabilistic design of stormwater retention facilities.

Two decades of random cascade applications to precipitation disaggregation brought progresses in the construction of generators. Quite soon, the assumption of independence and identical distribution of the cascade weight generators, at all timescales, was questioned and found suitable only for limited, rather narrow, range of analyzed scales (Olsson, 1998; Harris et al., 1998). As an alternative, Marshak et al. (1994), Menabde et al. (1997) and Harris et al. (1998) promoted the use of the so-called "bounded" random cascade, for which its weights distribution systematically evolves with the timescale decreasing the weights variance with the reduction of timescale. In addition, Rupp et al. (2009) suggested, that microcanonical cascade weights should not be timescale-dependent only, but also intensity-dependent. The common practice of assuming the Beta distribution for MCM generators was questioned by Licznar (2011a, b) especially for sub-hourly timescales. Alternatively MCM generators were assumed Normal-Beta (N-B) distributed with atom at 0.5, or 3N-B distributed, composed by three Normal and one Beta distribution.

Molnar and Burlando (2008) explored the variability of MCM generators on a large dataset of 10 min time resolution, including 62 stations across Switzerland. Molnar and Burlando (2008) investigated seasonal and spatial variability in breakdown distributions to give indications concerning the parameters' estimation of MCM in ungauged locations. To our knowledge, this is the only study considering the large-scale variability of MCM generators, and there is a lack of knowledge concerning the small-scale variability.

HESSD
11, 5251–5285, 2014

Having this in mind, we analyze here the spatial variability in the microcanonical cascade generators among gauges belonging to an urban precipitation network. We try to answer to the question "is it sufficient a single MCM for the whole city area?", or it is necessary to build a number of MCM generators varying in space. Here, we also explore the issue of calculating the MCM weights in case of only short time series, which we consider also very important for practical use perspective.

2 Data and methodology

2.1 Data

We use data belonging to a precipitation network of 25 gauges distributed throughout Warsaw, Poland (Fig. 1). The dataset is the same used by Rupp et al. (2012) and consists in a 1 min precipitation (both liquid and solid) time series recorded by electronic weighing-type gauges. All stations, TRwS 200E of MPS system Ltd. (Fig. 2), were installed and operated by the Municipal Water Supply and Sewerage Company (MWSSC) in Warsaw. Prior to the network installation, studies about the location of the stations have been done by MWSSC to identify the best configuration, representative of the precipitation variability within the urban area (Oke, 2006). All instruments were placed on grass, and their neighborhood met at least requirements of class 2 or 3, as recommended by WMO-No. 8. In the majority of gauges (i.e., R1, R3, R5, R7, R8, R10, R12, R17, R18 and R19) it was possible to install them on flat, horizontal surface, surrounded by an open area, meeting even requirements for class 1 instruments. In addition, gauge R15 was installed in perfect conditions on the ground at the Warsaw Fryderyk Chopin Airport.

Since the installation of the precipitation network in Warsaw was mainly motivated by the Real Time Control (RTC) of the drainage system, all gauges were connected to a single data acquisition system (Fig. 1). The accuracy of gauge measurements, as claimed by manufacturer is 0.1 %, and the data resolution is 0.001 mm for depth

HESSD
11, 5251–5285, 2014

and 1 minute for time. As it was already mentioned by Rupp et al. (2012), field tests, conducted prior to the operational use of the precipitation network, have shown good agreement between simulated and recorded totals, and have revealed a dampening/broadening of the input signal, evident over the range of a few minutes. The last phenomenon considered, known as "step response error", was studied in detail for different gauge types by Lanza et al. (2005). These found that the step error of TRwS gauge is quite small in comparison to other gauges, and equal to 3 min in laboratory conditions, which was consistent with our tests (Fig. 2). Having in mind the appearance of the step response, as well as timescales of previous microcanonical cascade studies concerning urban hydrology, we have aggregated the data to 5 min intervals.

Despite the limited timespan of available data, covering the period from the 38th week of year 2008 up to the 49th week of year 2010, we believe that the Warsaw precipitation network might support good probing ground for the variability study in the microcanonical cascade parameters over small-scale urban areas. In fact, the Warsaw precipitation monitoring network belongs to the biggest European urban gauge networks. Its size could be compared only with similar networks of 25 gauges in Vienna, or 24 gauges spread throughout Marseille and Barcelona (see Appendix B, Thames Tunnel Needs Report, 2010).

Finally, we compare the results of our study with those related to other Polish and German gauges. We limit our comparison to results previously published by Licznar et al. (2011a, b) for four gauges in Germany (gauges A, B, C and D – representing local climates of different parts of western Germany) and for one gauge in Wroclaw, Poland, and unpublished yet results by Górski (2013) for rain-gauge in Kielce, Poland (Fig. 3). Our choice is motivated by the similarity of the used methodology, and the investigated range of timescales, as well as by the indispensable accessibility to precise recordings of the breakdown coefficient histograms.

HESSD
11, 5251–5285, 2014

2.2 Microcanonical cascade models

We use microcanonical cascade models (MCMs) as in Licznar et al. (2011a, b). We limit our application to the disaggregation of precipitation totals from 1280 min (quasi daily) into 5 min times series, assuming the branching number b equal to 2, and constructing cascades assembled from only 8 levels ($n = 7, 6, \ldots, 1, 0$) corresponding to timescales $\lambda = 2^n$ from $\lambda = 128$ to $\lambda = 1$. Precipitation depth time series generated by such cascades are the products of the original precipitation total R_0 at timescale $\lambda = 128$ multiplied by the sequence of weights at the descending cascade levels:

$$R_{j,k} = R_0 \prod_{i=1}^{k} W_{f(i,j),i}, \tag{1}$$

where $j = 1, 2, \ldots, 2^{k-1}, 2^k$ marks the position in the time series at the kth cascade tier. The sequence of randomly generated weights $W_{f(i,j),i}$ is steered at the following ith cascade tier by the function $f(i,j)$, which rounds up $j/2^{k-i}$ to the closest integer. The weights in the microcanonical cascades are forced to sum to one, so their pairs are always equal to W and 1-W respectively, where W is a two-sided truncated random variable from 0 to 1. The microcanonical assumption conserves the mass (precipitation depth in our case) at each branch, and eliminates the risk of cascade degeneration. From engineering perspective, this means that the downscaling process could be seen as opposite to precipitation summation realized by Hellman gauges, recording daily totals only, and a pragmatic solution for the generation of synthetic precipitation time series at 5 min resolution.

Here, we do not however limit attention to disaggregation capabilities of microcanonical cascades. We rather focus on the small-scale variability of their generators W among gauges constituting the urban precipitation network. The obvious attractive of MCMs arises from the possibility of extracting the distribution of W from data on the base of breakdown coefficients studies (Cârsteanu and Foufoula-Georgiou, 1996). By

HESSD
11, 5251–5285, 2014

definition, BDCs are generally calculated using non-overlapping adjacent pairs of precipitation time series:

$$\mathrm{BDC}_{j,\tau} = \frac{R_{j,\tau}}{R_{j,\tau} + R_{j+1,\tau}} \qquad j = 1,3,5,\ldots,N_\tau - 1; \tag{2}$$

where $R_{j,\tau}$ is the precipitation amount for the time interval of length τ at position j in the time series, and N_τ is the length of timeseries at timescale τ. The calculation of BDCs with respect to Eq. (2) for Warsaw gauges is conducted only for nonzero pairs of R_j and R_{j+1}. Calculations are executed at aggregated intervals of length $2^{8-n}\tau_{\mathrm{org}}$, where τ_{org} is the original time step equal to 5 min and n is an integer, decreasing from 7 to 0, with increasing cascade timescales. Simultaneously, for all analyzed timescales, BDC couples equal to 0/1, or 1/0 (when only one between R_j and R_{j+1} is zero) are separated from resulting datasets and their occurrence probabilities, respectively p_0(LEFT) and p_0(RIGHT) are used to estimate intermittency probability p_0:

$$\Pr(\mathrm{BDC}_n(j) = 0 \text{ or } \mathrm{BDC}_n(j+1) = 0) = p_0(\mathrm{LEFT}) + p_0(\mathrm{RIGHT}) = p_0. \tag{3}$$

The probability p_0 is used within a MCM generator to take into account the intermittency, so characteristic of precipitation, forcing some portion of random weights W to be equal 0.

The preliminary results have revealed overrepresentation of BDC values equal to 1/2 or 1/3, 2/5, 1/4 (and 2/3, 3/5, 3/4 respectively) especially for small timescales, i.e. $\lambda = 1$ and $\lambda = 2$. Similar phenomenon was already reported by Rupp et al. (2009), and Licznar et al. (2011b), and explained as the result of instrument or recording precision of precipitation gauges. The magnitude of observed rounding errors for Warsaw gauges however smaller than in case of German gauges (Licznar et al., 2011b) since the precipitation depths were recorded with better resolution of 0.001 mm, still however resulted in irregularity of BDCs distribution, induced by sharp peaks at discrete BDC values, and making difficult the identification of some theoretical distribution. In order to

HESSD
11, 5251–5285, 2014

correct the rounding errors, a randomization procedure originally proposed by Licznar et al. (2011b) was applied. This type of procedure is fundamental in the analysis of data characterized by the presence of ties, see also De Michele et al. (2013). Thus, the original 5 min timeseries were slightly modified by adding to their precipitation depths, exceeding zero, some random corrections. Random correction values were sampled from the Uniform distribution in the range [−0.0005, 0.0005] mm.

Irregularities in BDC histograms were observed for timescales $\lambda > 8$. These are due to the decreasing sample size of BDC values, calculated on limited timespan of accessible data, slightly exceeding 2 years. This issue was rather irrelevant in former studies (Molnar and Burlando, 2005, 2008; Licznar et al., 2011a, b) realized on dataseries 10 or even 20 times longer. To solve this issue, we applied the overlapping moving window algorithm as an alternative to the classical non-overlapping moving window algorithm for the calculation of BDCs values. Figure 4 shows differences between the two algorithms on the first two timescales, however the real strength of the overlapping moving window algorithm is for the largest timescale $\lambda = 128$, where the number of windows for the calculation of BDC was almost 7 times greater than those with the non-overlapping moving window algorithm. We believed that the overlapping moving window algorithm has allowed the complete statistical exploration of BDC values even in presence of limited observational data, providing BDC datasets of almost equal size for all analyzed timescales.

To compare BDC histograms obtained for all analyzed timescales with theoretical functions, a probability distribution assembling 2 truncated (with truncation points at 0 and 1) Normal distributions (Robert, 1995), and 1 Beta symmetrical distribution was implemented. This distribution, indicated as 2N-B distribution, has the following density

HESSD
11, 5251–5285, 2014

function:

$$p(w) = p_1 \left\{ \frac{1}{\sigma_1 \sqrt{2\pi}} e^{\frac{-(w-0.5)^2}{2\sigma_1^2}} \right\} + (1-p_1) \left\{ p_2 \left\{ \frac{1}{B(a)} w^{a-1}(1-w)^{a-1} \right\} + (1-p_2) \left\{ \frac{1}{\sigma_2 \sqrt{2\pi}} e^{\frac{-(w-0.5)^2}{2\sigma_2^2}} \right\} \right\} \quad (4)$$

where p_1 and p_2 were weights characterizing the contribution of the individual distributions within the 2N-B distribution, σ_1 and σ_2 were the scale parameters of truncated Normal distributions, and $B(a)$ was the symmetrical Beta function, parameterized by a.

2.3 Cluster analysis

The shape of BDC histograms is compared among the stations of the monitoring network in Warsaw, and with other Polish and German gauges applying cluster analysis. In particular a *hierarchical clustering* is used. This is a data-mining tool, applied to segment data into relatively homogeneous subgroups, or clusters, where the similarity of the records within the cluster is maximized (Larose, 2005). Prior the application of the clustering technique, for each timescale and each site, the BDC histogram is sampled in 100 points, selected at equal distance one from the following one. These 100 values are considered components of a vector characterizing the empirical BDC distribution. The clustering of similar sites is operated using the Euclidean distance. It is computed as:

$$d_{\text{Euclidean}}(\boldsymbol{X}, \boldsymbol{Y}) = \sqrt{\Sigma_i (x_i - y_i)^2}, \quad (5)$$

where x_i and y_i with $i = 1, \ldots, 100$, represent respectively the i component of $\boldsymbol{X}$ and $\boldsymbol{Y}$ vectors.

HESSD
11, 5251–5285, 2014

Initially, each vector is considered to be a tiny cluster of its own. Then, in following steps, the two closest clusters are aggregated into a new combined cluster. By replication of this operation, the number of clusters is reduced by one at each step and eventually, sites are combined into a single huge cluster. During the agglomerative process, the distance between clusters is determined based on single-linkage criterion. In this case, the distance between two clusters A and B is defined as the minimum distance between any element in cluster A and any element in cluster B. With respect to this single-linkage is often termed the nearest-neighbor approach, and tends to form long, slender clusters, clearly indicating similarities among clustered elements. As a final result of agglomerative clustering a treelike cluster structure (named dendrogram) is created. Dendrograms show similarities of BDC distributions among the considered sites, separately for all analyzed timescales.

In addition, the cluster analysis is also applied to the intermittency parameter, comparing in this case, vectors of 8 components, each of these being the p_0 value for the 8 timescales λ = 1, 2, 4, 8, 16, 32, 64, 128.

3 Results and discussion

Results are presented relatively to gauge R7, for brevity. This station has been selected because of its localization in the strict city center and its installation in perfect meteorological conditions on the ground. Results for the other gauges are qualitatively similar to those shown for R7.

BDC histograms are calculated using the non-overlapping moving window algorithm, and plotted in Fig. 5 for gauge R7 and a sequence of analyzed breakdown times. It is clearly visible that despite the randomization procedure removes pronounced peaks of histograms at certain specific BDC values, like 0.5 or 1/3, 2/5, 1/4 and 2/3, 3/5, 3/4 respectively, the plots especially for timescales exceeding $\lambda = 8$ remain still irregular, reducing the possibility of identifying the proper theoretical distribution. Visible irregularities of BDC histograms increase with increasing timescales, which is an obvious

HESSD
11, 5251–5285, 2014

effect of decreasing datasets and thus decreasing populations of calculated BDC values not allowing to produce histograms of fine bins resolution. Similarly, Fig. 6 reports the distributions of BDC calculated through the overlapping moving window algorithm. The comparison between Figs. 5 and 6 shows how the change of algorithm from non-overlapping to overlapping moving window has brought to evident smoothing of BDC histograms especially for larger timescales, but occurring also at small timescales. Note that the smoothness of BDC histograms in Fig. 6 is comparable with the quality of BDC histograms showed by Licznar et al. (2011b) for German gauges, derived using non-overlapping moving window algorithm for much longer precipitation timeseries ranging from 27 to 46 years of continuous records. The introduction of overlapping moving window algorithm allows for the fitting of MCM parameters even when extremely short timeseries (say 2 years long) are available, as in the case of Warsaw gauges.

In Fig. 6, we report also the theoretical distribution for each timescale considered. The distribution changes from a Beta distribution (B) at $\lambda = 128$, to a joined double Normal-Beta distribution (2N-B) for the smallest value of λ, through joined Normal-Beta distributions (N-B). This is in agreement with previous studies by Licznar et al. (2011a, b). From the theoretical point of view a rapid increase in the number of BDCs, equal or close to 0.5, decreasing the timescale should be expected, as a symptom of enclosing a limit of precipitation homogeneity or variability recording. For bounded cascades (Marshak et al., 1994; Menabde et al., 1997; Harris et al., 1998) the variance of weights reduces with every descending cascade level. As a simple extension of this rule, the increasing frequency of weights at the central part of their distribution plots has to be observed. The increase in the number of BDCs equal or close to 0.5 with decreasing timescale is well illustrated by empirical histograms at well-known pioneering contributions to MCM applications for rainfall time series disaggregation, published by Olsson (1998), Menabde and Sivapalan (2000) and Güntner et al. (2001). Quite recently, this behavior was observed for empirical distributions of BDC and proved to be rainfall intensity dependent by Rupp et al. (2009).

HESSD
11, 5251–5285, 2014

For each analyzed timescale, we have estimated the 5 parameters of 2N-B probability distribution: p_1, p_2, a, σ_1 and σ_2. Table 1 gives the values for gauge R7 with their 95 % confidence limits. Figure 6 display a good visual fit between theoretical and empirical BDC distributions. As a proof of this, the 95 % confidence limits of the fitted parameters are quite narrow, not exceeding few percent of the estimated values, with the sole exception of parameter p_1 for $\lambda = 4$, where the differences range up to 27 %. Additionally, the scale parameters of normal distributions, σ_1 and σ_2, appear to be constant among analyzed timescales, not only for gauge R7, but also for the other Warsaw gauges.

The variability of p_1, p_2, a with λ is presented in Fig. 7 for gauge R7. A systematical decrease of p_1 up to 0 increasing the timescale is observed, denoting a decreasing importance of the first Normal within the 2N-B distribution. An opposite systematical increase of p_2 up to 1 increasing the timescale is observed, denoting an increasing importance of the second Normal within the 2N-B distribution. The evolution of the Beta parameter a shows a fast reduction with below 1 values noticed for the smallest scales, yielding the change of Beta distribution shape from convex to concave. At larger timescales, the reduction of a is hardly visible with the sole exception of $\lambda = 128$.

Figure 8 shows the variability of intermittency parameters p_0 with timescale λ. For all of them, the values of p_0(LEFT) match the values of p_0(RIGHT), which is in good agreement of previous studies of Molnar and Burnaldo (2005) and Licznar et al. (2011a, b). Systematical increase of p_0 with λ is observed with the sole exception of some small drop at $\lambda = 128$.

To check the correctness of the randomization procedure, the "moving windows" scheme and the 2N-B distribution fitting the empirical histograms of BDC, we build the microcanonical cascade and test their performance in disaggregating the precipitation for analyzed gauges. The MCM is used to generate 100 precipitation timeseries of 5 min resolution on the basis of the observed 1280 min precipitation totals. To evaluate the goodness of disaggregation, we calculate the survival probability function of 5 min synthetic precipitation amounts and compare it to the survival probability function

HESSD
11, 5251–5285, 2014

of 5 min observed precipitation amounts. Moreover we calculate and compare probabilities of zero precipitation at synthetic and observed time series for all analyzed timescales. An example of disaggregation for a 56.3 mm event is plotted in Fig. 9, for gauge R7.

Figure 10 shows the comparison between observed and simulated survival probability function of rainfall amount at 5 min, for gauge R7. In Fig. 10, we report the empirical survival probability function obtained from a synthetic series out of 100 and also the averaged function calculated using all the generated series. As expected the behavior of both the synthetic functions is very similar, with the sole exception of the extended and smoothed tail of the averaged function plot. Both the synthetic functions are plotted slightly above the observed function. Observed displacement reveals slight overprediction of 5 min precipitation depths, particularly at the range of small and moderate precipitation intensities from 0.3 mm/5 min to about 2.0 mm/5 min. In our opinion this overproduction does not disqualifies the use of generated time series for practical engineering tasks like e.g. probabilistic stormwater reservoir dimensioning, as proposed by Licznar (2013). In addition, the magnitude of dissimilarities between synthetic and observed functions does not exceed the one seen in other published works, see e.g., Molnar and Burlando (2005), Licznar et al. (2011a, b).

The probabilities of zero precipitation, $E(p_0)$, for observed and MCM genereted series are compared in. Figure 11, for gauge R7. The values of $E(p_0)$ for generated time series are calculated as average over all 100 MCM dissagregations. The differences among $E(p_0)$ values calculated on different synthetic series are negligible, and for this these are not reported in Fig. 11. Similar comment applies to the differences between observed and synthetic values.

Dendrograms summarizing the results of the cluster analysis for BDC histograms are produced for each timescale, and reported in Figs. 12 and 13 only for $\lambda = 1$ and $\lambda = 128$, respectively. Results for the first four timescales, i.e. $\lambda = 1, 2, 4, 8$, are unsurprising and easy to be interpreted. All Warsaw gauges are grouped in a single cluster with similar shapes of BDC histograms. For all Warsaw gauges their interconnection

HESSD
11, 5251–5285, 2014

on the dendrogram is placed at the level of binding distance equal to about 0.5. Only R25 seems to be characterized by slightly different pattern of BDC histogram. However, gauge R25 has a behavior, which is still much closer to other Warsaw gauges, rather than the behavior of the other cities considered. For example, at $\lambda = 1$, gauge R25 is merged into Warsaw gauges cluster at an Euclidean distance equal to 0.81, whereas the same occurs for Kielce (the closest considered Polish city) gauge at the Euclidean distance equal to 1.07. For other timescales, $\lambda = 2, 4, 8$, gauge R25 merges the cluster of Warsaw gauges at quite similar Euclidean distances: 0.89, 0.83 and 0.81 respectively. A plausible explanation of slightly outlying behavior of gauge R25 could be found in its location, on south-east suburban area, in close vicinity of forested area and Vistula river valley. This specific suburban area is also most frequently an origin place for the development of local convection processes (Prof. S. Malinowski, personal communication, 2013).

The dendrogram for $\lambda = 128$ is given in Fig. 13, being representative of timescales $\lambda = 16, 32, 64, 128$. From Fig. 13, it is possible to see the departure of gauge R15 from the cluster of other Warsaw gauges. The position of gauge R15 is isolated from other Warsaw gauges and its Euclidean distance from the closest one is big, and increases increasing the timescale; it is equal to 1.80, 3.19, 3.88, and 8.03 respectively for $\lambda = 16$, 32, 64 and 128. Simultaneously, the Euclidean distance from the cluster of Warsaw gauges to the nearest neighbor does not exceed 0.90, 1.00, 1.40 and 1.89 respectively for $\lambda = 16, 32, 64$, and 128.

This last observation puts in evidence that in general the variability of BDC shapes, among Warsaw gauges, increases increasing the timescale. It may partly be explained by the already mentioned evolution of histogram shapes, and the replacement of 2N-B distribution by less centered N-B and finally B distribution characterized by a higher variance of BDC.

In the specific case of gauge R15, its BDC histograms for the largest timescales are boldly concave (not shown for brevity) and their shapes are becoming similar to beta symmetrical distributions parameterized by very small values of a: 0.76, 0.64, 0.54, and

HESSD
11, 5251–5285, 2014

0.45 respectively for λ = 16, 32, 64, and 128. One possible explanation for the behavior of gauge R15 seems to arise from its specific location on the ground of the Warsaw airport. In the neighborhood of the instrument there are no high buildings and trees and the ground is covered only by short cut grass. The local atmospheric turbulence conditions, additionally influenced by taking off and landing aircrafts could have favored the different behavior of this station.

As a final step of our study, we used the cluster analysis to investigate the variability among the gauges in terms of the intermittency parameter p_0 considered as a vector having as the 8 components its values in correspondence of the considered timescales. Results are given in form of dendrogram in Fig. 14. With respect to p_0, all Warsaw gauges form one single chain-like cluster. Three gauges in the cluster, namely R14, R25 and R15, are characterized by the largest distances from the nearest neighbor with Euclidean distances equal to: 0.079, 0.064 and 0.0614 respectively. The distance of gauges R15 and R25 from the other stations in cluster is probably caused by already discussed conditions. A possible, but not certain, explanation for gauge R14 could be its location close to gauge R15, also in a weak-developed part of the city.

4 Conclusions

We have investigated the variability of precipitation in terms of microcanonical cascade generators among gauges of an urban precipitation network. We considered a network of 25 gauges deployed in Warsaw city (Poland) over an area of 517.2 km^2. An attempt was made to define the generators of a MCM applicable for producing 5 min time series, as requested by urban hydrologists, through the disaggregation of quasi-daily precipitation totals. We showed that smooth distributions of BDC are possible, for all analyzed timescales, even in case of limited length of time series, which in our case slightly exceeded 2 years only. This was made possible by the implementation of a randomization procedure and the overlapping moving window algorithm for the calculation of BDCs.

HESSD
11, 5251–5285, 2014

To adequately describe the shapes of BDC histograms, we have implemented a special joined probability distribution, 2N-B, assembled from 2 Normal distributions and 1 Beta symmetrical distribution. A systematical evolution of BDC histograms from joined double Normal-Beta distributions (2N-B), through joined Normal-Beta distributions (N-B) up to Beta distributions (B) was observed increasing the timescale. To verify the correctness of cascade generators, MCM was used to disaggregate 1280 min precipitation totals disaggregation into 5 min time series. The quality of the generated fine graining timeseries was checked comparing the statistical properties of these with the ones of observed series. In particular, we compared the survival probability functions of non-zero 5 min precipitation amounts and probabilities of zero precipitation for the considered timescales with a very good agreement.

As a core part of our study, we have conducted an intercomparison of BDC histograms among the 25 Warsaw gauges, and considering as a term of reference also other 6 gauges located in Poland and Germany. The intercomparison was made scale by scale by means of cluster analysis technique. Resulting dendrograms for small timescales (i.e. $\lambda = 1, 2, 4, 8$) revealed rather small variability of BDC histograms among all Warsaw gauges in comparison to the variability exhibited with respect to the other external gauges. Only gauge R25 seems to be characterized by a slightly different pattern. It might originate from the specific gauge location on the city boundary, in the vicinity of forested areas and Vistula river valley.

Dendrograms obtained for large timescales (i.e. $\lambda = 16, 32, 64, 128$) also delivered a general picture of similarity among Warsaw gauges, with the very clear exception of gauge R15. To our best knowledge a possible explanation of this was its installation on the ground of the Warsaw airport, strongly man-modified and with local turbulence conditions. In addition, R25, R15, and R14 were also identified as gauges presenting slightly different behavior in terms of the intermittency parameter p_0.

As final remarks, we can affirm that it is enough one single MCM for the precipitation disaggregation in Warsaw city. We suppose that this result could be valid even in larger urban areas, but verification is necessary. We dissuade from cascade generation fitted

HESSD
11, 5251–5285, 2014

on precipitation time series collected at instruments located out of the city center in unrepresentative sites, like in our case, the ground of the airport. We question the quite common practice of using gauges from airport gauges data adoption for urban hydrology. We recommend further research to assess the influence of the local conditions on BDC histograms to find clear explanation of anomalies. We also fully recognize the necessity of our observations verification for other cities and their precipitation monitoring networks, especially in case of cities with more complicated orography and hydrologic networks. Finally promoting the idea of the fast construction of local precipitation models based on MCMs with generators parameterized on short observational series, we see the urge for new research giving some insight into the influence of observational series length on BDCs distributions.

Acknowledgements. This project was financed by the Polish National Science Centre (NCN) funds allocated on basis of decision no 2011/03/B/ST10/06338. It was realized as a part of scientific project: “Spatio-temporal analysis and modeling of urban precipitation field”.

References

Ahrens, B.: Rainfall downscaling in an alpine watershed applying a multiresolution approach, J. Geophys. Res., 108, 8388, doi:10.1029/2001JD001485, 2003.

Arbeitsblatt DWA-A 117: Bemessung von Regenrückhalteräumen. Deutsche Vereinigung für Wasserwirtschaft, Abwasser und Abfall e. V., Hennef, 2006.

Appendix, B.: Report on Approaches to UWWTD Compliance in Relation to CSO’s in major cities across the EU, Thames Tunnel Needs Report, 2010.

Cârsteanu, A. and Foufoula-Georgiou, E.: Assessing dependence among weights in a multiplicative cascade model of temporal rainfall, J. Geophys. Res., 101, 26363–26370, doi:10.1029/96JD01657, 1996.

De Michele, C., Salvadori, G., Vezzoli, R., and Pecora, S.: Multivariate assessment of droughts: Frequency analysis and dynamic return period, Water Resour. Res., 49, 6985–6994, doi:10.1002/wrcr.20551, 2013.

EN 752, Drain and sewer systems outside buildings, 1997.

HESSD

11, 5251–5285, 2014

Górski, J.: Analysis of Precipitation Time Series for Needs of Urban Hydrology on Example of Kielce City. PhD thesis, Kielce University of Technology, 2013 (in Polish).
Harris, D., Seed, A., Menabde, M., and Austin, G.: Factors affecting multiscaling analysis of rainfall time series, Nonlin. Processes Geophys., 4, 137–156, doi:10.5194/npg-4-137-1997, 1997.
Hingray B. and Ben Haha, M.: Statistical performances of various deterministic and stochastic models for rainfall series disaggregation, Atmos. Res., 77, 152–175, 2005.
Koutsoyiannis, D.: Rainfall disaggregation methods: Theory and applications, in: Proc. Workshop on Statistical and Mathematical Methods for Hydrological Analysis, edited by: Piccolo, D. and Ubertini, L., Rome, 1-23, Universitá di Roma "La Sapienza", available at: http://itia.ntua.gr/en/docinfo/570, 2003.
Lanza, L., Leroy, M., Alexandropoulos, C., Stagi, L., and Wauben, W.: WMO laboratory intercomparison of rainfall intensity gauges, Final report, IOM Report No. 84, WMO/TD No. 1304, 2005.
Larose, D. T.: Discovering knowledge in data: an introduction to data mining, John Wiley & Sons, Inc., Hoboken, New Jersey, 2005.
Licznar, P.: Stormwater reservoir dimensioning based on synthetic rainfall time series, Ochrona Srodowiska, 35, 27–32, 2013.
Licznar, P., Łomotowski, J., and Rupp, D. E.: Random cascade driven rainfall disaggregation for urban hydrology: An evaluation of six models and a new generator, Atmos. Res., 99, 563–578, doi:10.1016/j.atmosres.2010.12.014, 2011a.
Licznar, P., Schmitt, T. G., and Rupp, D. E.: Distributions of microcanonical cascade weights of rainfall at small timescales, Acta Geophys., 59, 1013–1043, doi:10.2478/s11600-011-0014-4, 2011b.
Marshak, A., Davis, A., Cahalan, R., and Wiscombe, W.: Bounded cascade models as nonstationary multifractals, Phys. Rev. E, 49, 55–69, 1994.
Menabde, M. and Sivapalan, M.: Modeling of rainfall time series and extremes using bounded random cascades and Levy-stable distributions, Water Resour. Res., 36, 3293–3300, doi:10.1029/2000WR900197, 2000.
Menabde, M., Harris, D., Seed, A., Austin, G., and Stow, D.: Multiscaling properties of rainfall and bounded random cascades, Water Resour. Res., 33, 2823–2830, doi:10.1029/97WR02006, 1997.

HESSD
11, 5251–5285, 2014

Molnar, P. and Burlando, P.: Preservation of rainfall properties in stochastic disaggregation by a simple random cascade model, Atmos. Res., 77, 137–151, doi:10.1016/j.atmosres.2004.10.024, 2005.

Molnar, P. and Burlando, P.: Variability in the scale properties of high-resolution precipitation data in the Alpine climate of Switzerland, Water Resour. Res., 44, W10404, doi:10.1029/2007WR006142, 2008.

Oke, T.: Initial guidance to obtain representative meteorological observations at urban sites, Instruments and Observing Methods, Report no. 81, World Meteorological Organization, WMO/TD-No. 1250, 2006.

Olsson, J.: Evaluation of a scaling cascade model for temporal rain- fall disaggregation, Hydrol. Earth Syst. Sci., 2, 19–30, doi:10.5194/hess-2-19-1998, 1998.

Paulson, K. S. and Baxter P. D.: Downscaling of rain gauge time series by multiplicative beta cascade, J. Geophys. Res., 112, D09105, doi:10.1029/2006JD007333, 2007.

Robert, C. P.: Simulation of truncated normal variables, Stat. Comput., 5, 121–125, 1995.

Rupp, D. E., Keim, R. F., Ossiander, M., Brugnach, M., and Selker, J. S.: Time scale and intensity dependency in multiplicative cascades for temporal rainfall disaggregation, Water Resour. Res., 45, W07409, doi:10.1029/2008WR007321, 2009.

Rupp, D. E., Licznar, P., Adamowski, W., and Leśniewski, M.: Multiplicative cascade models for fine spatial downscaling of rainfall: parameterization with rain gauge data, Hydrol. Earth Syst. Sci., 16, 671–684, doi:10.5194/hess-16-671-2012, 2012.

Schmitt, T. G.: ATV-DVWK Kommentar, ATV-A 118 Hydraulische Berechnung von Entwässerungssystemen, DWA, Hennef, 2000.

WMO-No. 8: Guide to Meteorological Instruments and Methods of Observation, World Meteorological Organization – WMO, 2012.

HESSD
11, 5251–5285, 2014

Table 1. Values of p_1, p_2, a, σ_1 and σ_2 at analyzed timescales, for gauge R7. The fitted parameters values are reported in italic, whereas their 95 % confidence limits are given below.

Breakdown times	Timescale λ	p_1	p_2	a	σ_1	σ_2
5–10 min.	$\lambda = 1$	*0.1541*	*0.3479*	*1.3350*	*0.0559*	*0.1341*
		0.1474	0.3377	1.3097	0.0523	0.1300
		0.1608	0.3580	1.3604	0.0595	0.1383
10–20 min.	$\lambda = 2$	*0.0706*	*0.4036*	*1.0632*	*0.0559*	*0.1341*
		0.0644	0.3950	1.0474	0.0523	0.1300
		0.0768	0.4121	1.0789	0.0595	0.1383
20–40 min.	$\lambda = 4$	*0.0212*	*0.5036*	*0.9437*	*0.0559*	*0.1341*
		0.0155	0.4954	0.9325	0.0523	0.1300
		0.0270	0.5118	0.9548	0.0595	0.1383
40–80 min.	$\lambda = 8$	–	*0.6175*	*0.9484*	–	*0.1341*
		–	0.6091	0.9390	–	0.1300
		–	0.6259	0.9579	–	0.1383
80–160 min.	$\lambda = 16$	–	*0.7548*	*0.9170*	–	*0.1341*
		–	0.7494	0.9098	–	0.1300
		–	0.7601	0.9242	–	0.1383
160–320 min.	$\lambda = 32$	–	*0.8873*	*0.8929*	–	*0.1341*
		–	0.8827	0.8873	–	0.1300
		–	0.8919	0.8985	–	0.1383
320–640 min.	$\lambda = 64$	–	*0.9797*	*0.8799*	–	*0.1341*
		–	0.9758	0.8754	–	0.1300
		–	0.9835	0.8843	–	0.1383
640–1280 min.	$\lambda = 128$	–	*1.0000*	*0.7783*	–	*0.1341*
		–	0.9973	0.7754	–	0.1300
		–	1.0027	0.7813	–	0.1383

HESSD
11, 5251–5285, 2014

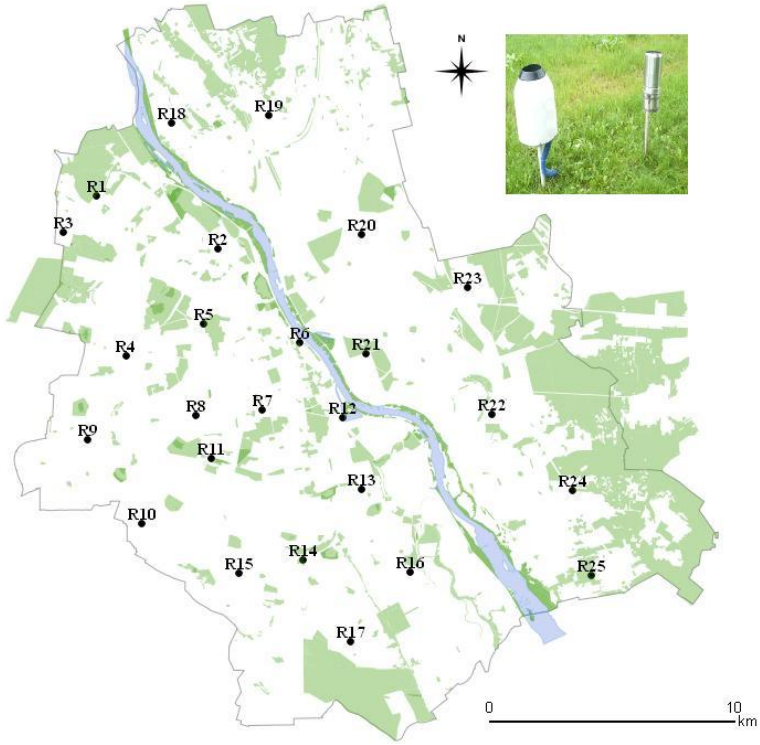

Figure 1. Map of 25 gauges composing the precipitation network in Warsaw. Administrative limits of Warsaw city and forested areas were marked in grey. The average density of network is 1 instrument over 20.7 km^2. MPS weighing-type TRwS 200E gauges were accompanied with standard Hellman gauges for the routine control of daily precipitation totals.

HESSD
11, 5251–5285, 2014

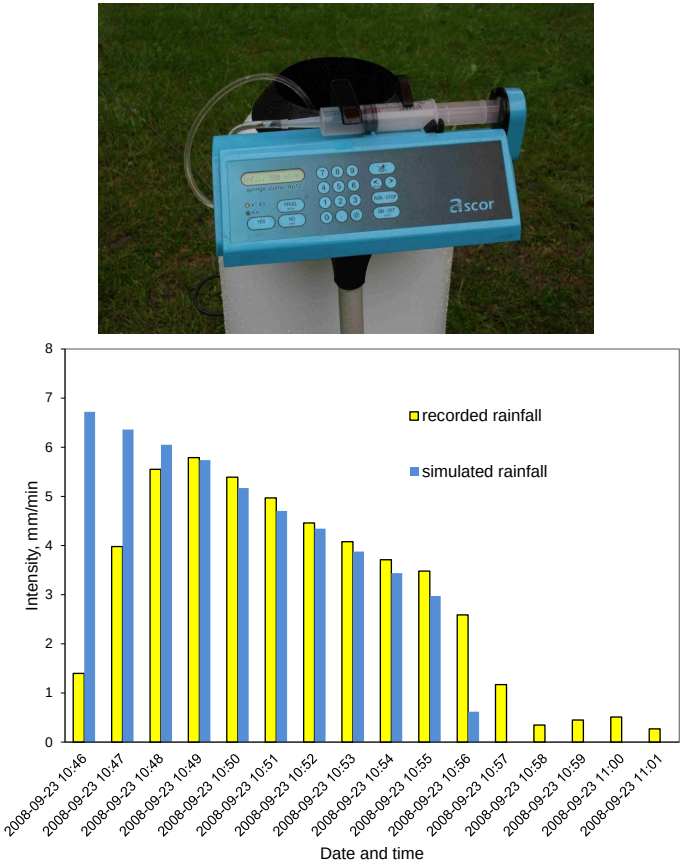

Figure 2. Weighing-type TRwS 200E gauge during some tests (upper panel). Rainfall is simulated by means of precise medical pump. Sample of test results reporting simulated and recorded rainfall depths (lower panel).

HESSD
11, 5251–5285, 2014

Figure 3. Location of Polish and German precipitation gauges used during the comparison of Warsaw results with other studies.

HESSD
11, 5251–5285, 2014

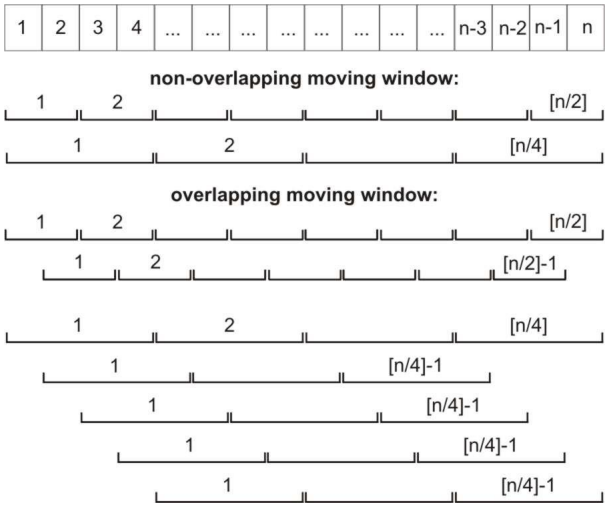

Figure 4. Differences between non-overlapping and overlapping moving window algorithms for the calculation of BDCs ([*n*] means the integer part of *n*).

HESSD
11, 5251–5285, 2014

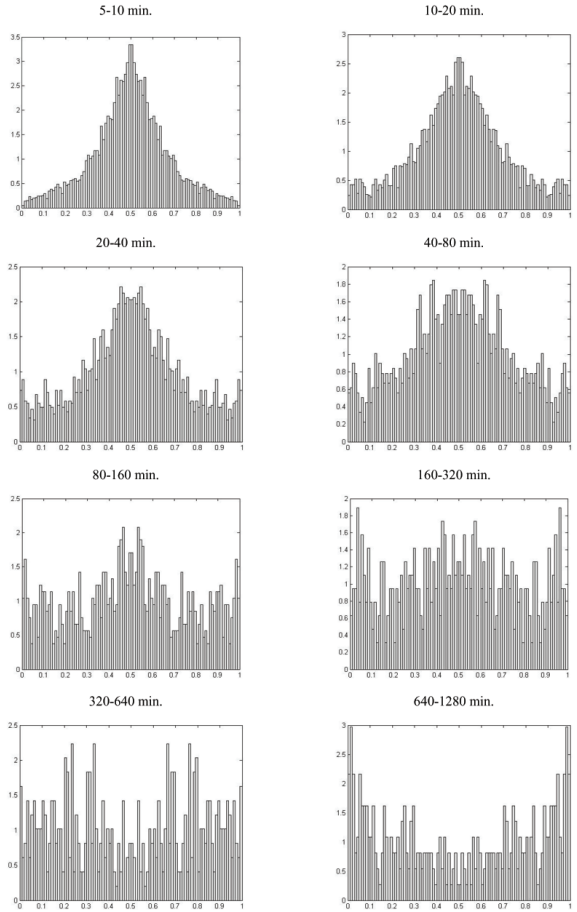

Figure 5. Histograms of BDC values for gauge R7 calculated according to the non-overlapping moving window algorithm and based on randomized precipitation timeseries. Horizontal axes show BDC ranges and vertical axes their respective frequency values.

HESSD
11, 5251–5285, 2014

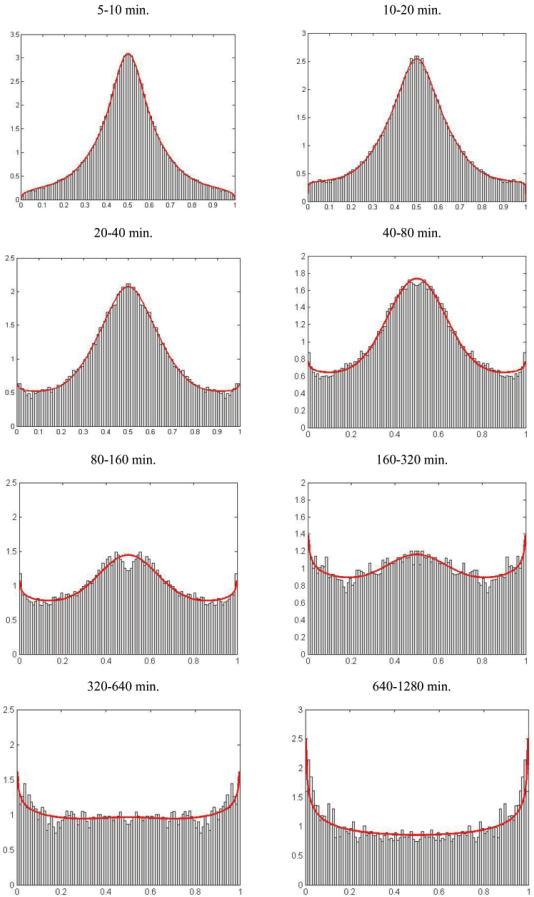

Figure 6. Histograms of BDC values calculated according to overlapping moving window algorithm and based on randomized gauge R7 precipitation times series. Horizontal axes show BDC ranges and vertical axes their respective frequency values. The solid curves represent the probability density functions (2N-B).

HESSD
11, 5251–5285, 2014

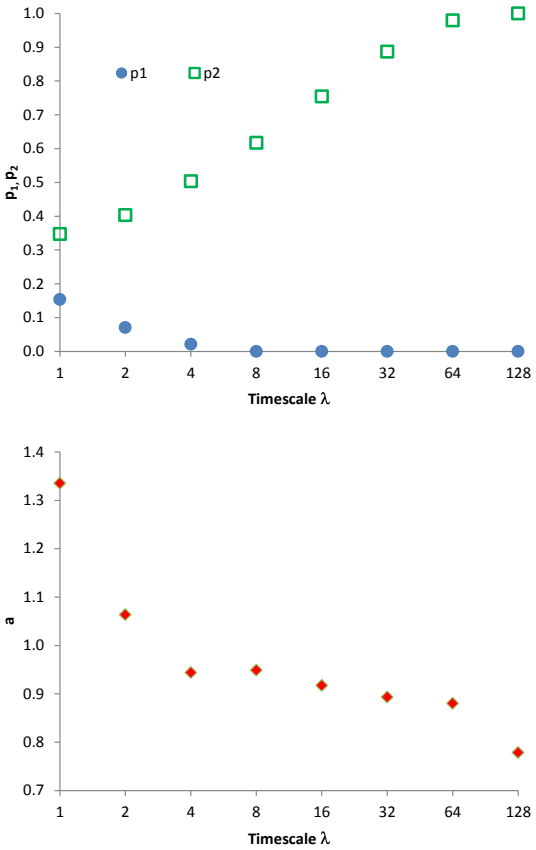

Figure 7. Variability of p_1, p_2 and a with λ, for gauge R7. Horizontal axes are plotted at binary logarithm scale $\log_2$.

HESSD
11, 5251–5285, 2014

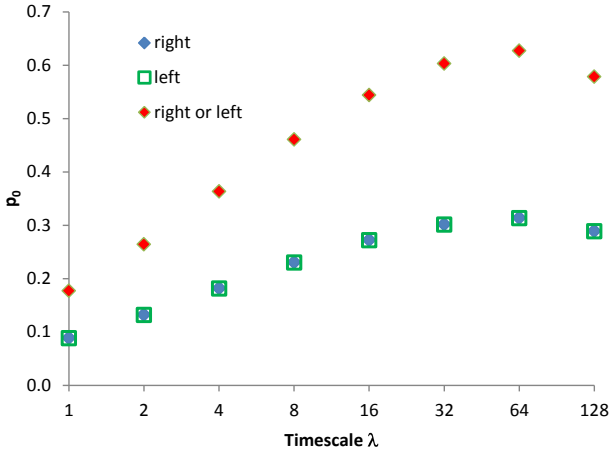

Figure 8. Variability of the intermittency parameter p_0 with λ, for gauge R7. Horizontal axis is plotted at binary logarithm scale $\log_2$.

HESSD
11, 5251–5285, 2014

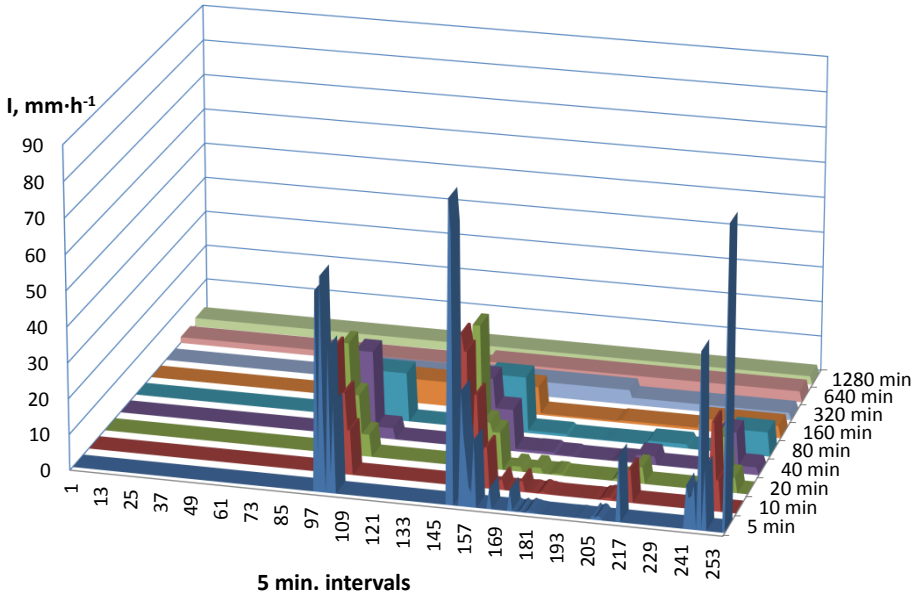

Figure 9. An example of precipitation disaggregation of a 56.3 mm event from 1280 to 5 min, for gauge R7.

HESSD
11, 5251–5285, 2014

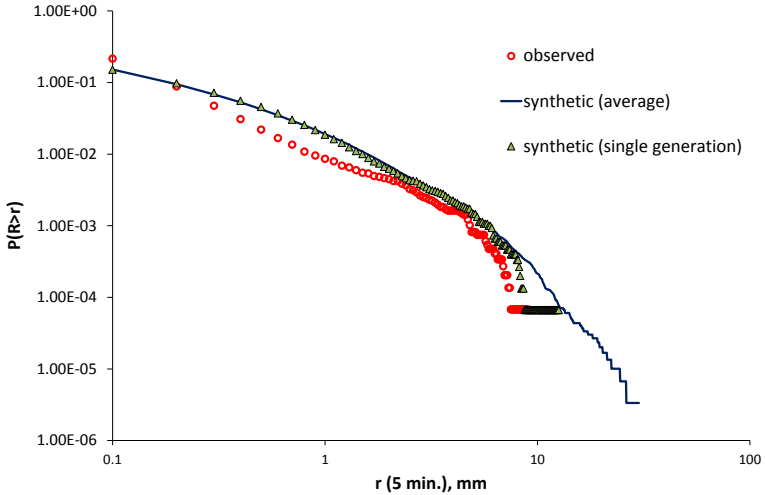

Figure 10. The survival probability function of 5 min precipitation amounts for the observed time series (circles) and the synthetic time series (triangles) generated by the disaggregation of 1280 precipitation amounts, for gauge R7. The line represents the average distribution calculated over the generation of 100 synthetic timeseries.

HESSD
11, 5251–5285, 2014

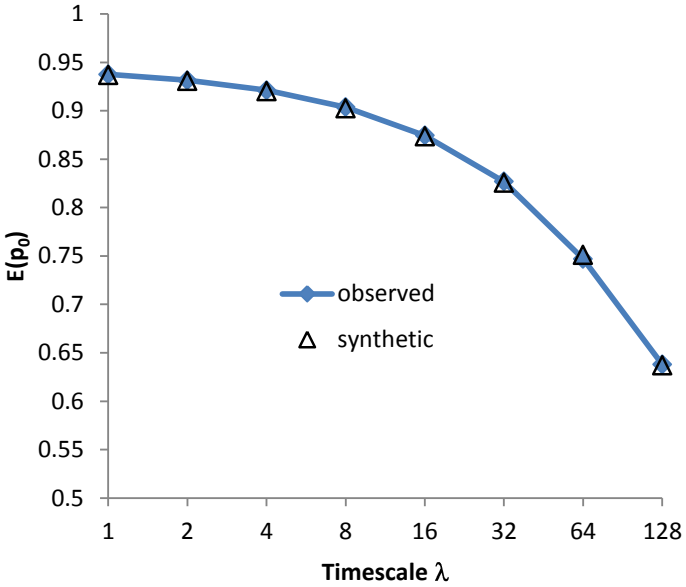

Figure 11. Comparison between observed and synthetic series in terms of intermittency $E(p_0)$ for the considered timescales, for gauge R7. The values for the generated data are calculated as average of 100 disaggregation runs. The variability between runs was negligible and so is not shown here.

HESSD
11, 5251–5285, 2014

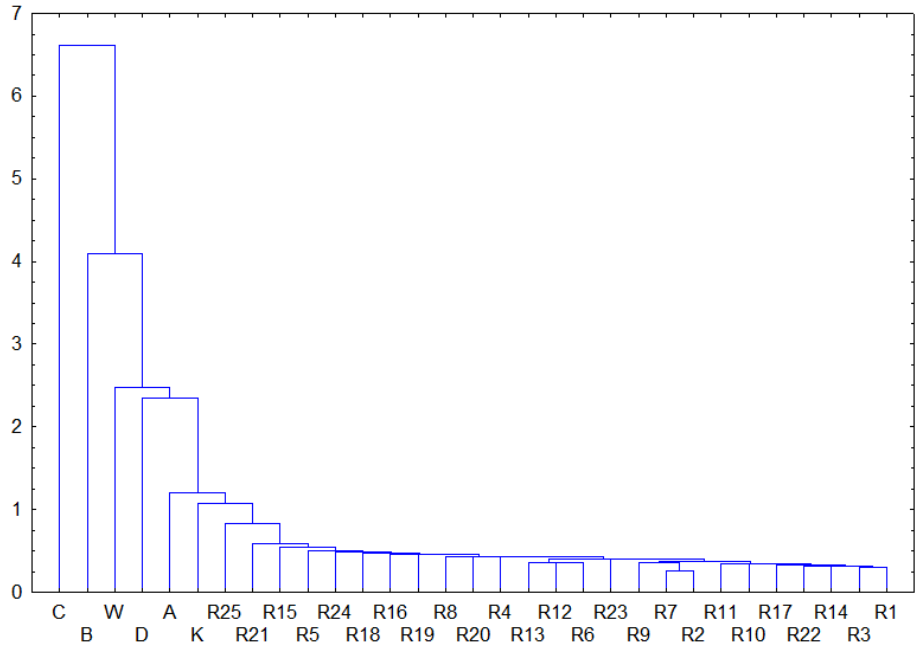

Figure 12. Dendrogram resulting from the cluster analysis of BDC histograms for $\lambda = 1$. The vertical scale shows binding distance, whereas names of gauges are given on horizontal scale (K stands for Kielce gauge, and W stands for Wroclaw).

HESSD
11, 5251–5285, 2014

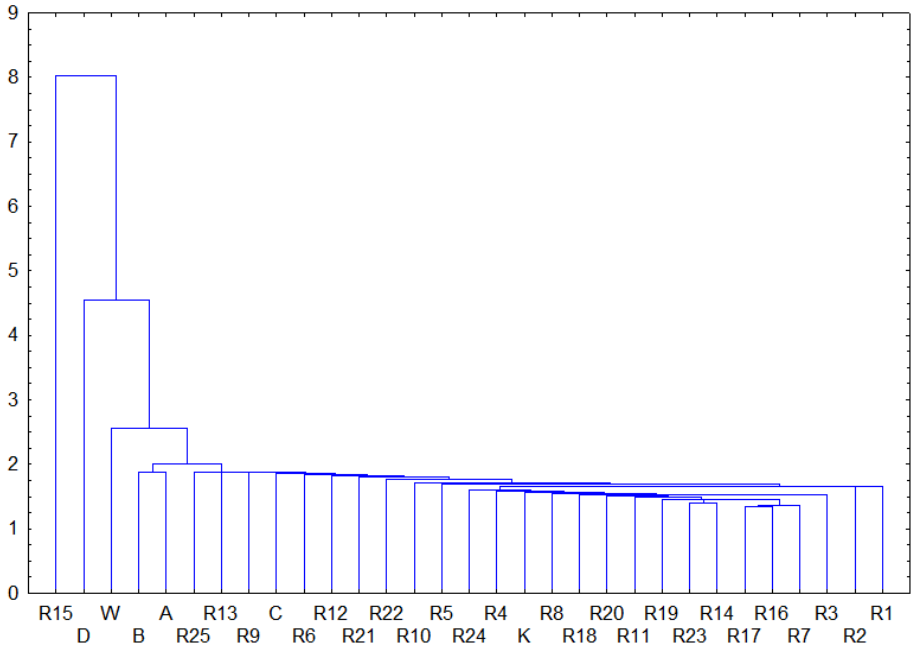

Figure 13. Dendrogram resulting from the cluster analysis of BDCs histograms for the timescale $\lambda = 128$. The vertical scale shows binding distance, whereas names of gauges are given on horizontal scale (K stands for Kielce gauge, and W stands for Wroclaw).

HESSD
11, 5251–5285, 2014

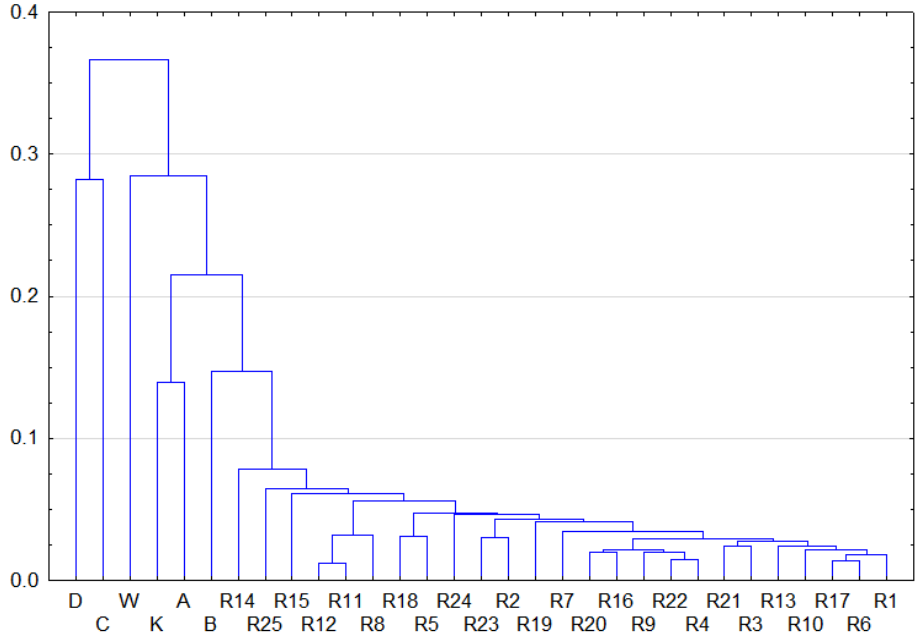

Figure 14. Dendrogram resulting from the cluster analysis of the intermittency parameter p_0. The vertical scale shows binding distance, whereas the name of gauges is given on horizontal scale (K stands for Kielce gauge, and W stands for Wroclaw).

HESSD
11, 5251–5285, 2014