A stochastic point rainfall model of single storm for urban drainage systems design based on 2-copula and dimensionless hyetograph

G. Brigandi*, G.T. Aronica¹

1 Department of Engineering, University of Messina, Italy

*Corresponding author: gbrigandi@unime.it

Abstract

In this paper a simple stochastic model to derive single synthetic sub-hourly rainfall events at a point is presented. Generated rainfall events are totally stochastic but with characteristics in terms of shape, duration and average intensity that have to satisfy the parameters derived by statistical analyses of the available historic records. The events are derived on the basis of a fixed inter-event time, and selected if their average intensity is bigger than a critical fixed one. In order to characterise the shape, dimensionless events are considered. Storm duration and average intensity of the observed events are modelled separately by fitting several probability distributions. Finally, their correlation is modelled using the Frank Copula. In order to test the methodology, two sites placed in Sicily, Italy, have been analysed. The good performance of the model has been tested by comparing the statistics of the simulated events with those of the measured data, and the obtained results confirm the effectiveness of the methodology.

1. Introduction

Simulation of sub-hourly rainfall time-series is an important area of hydrological research being, in fact, particularly useful, i.e., to design water control structures or urban drainage systems in urban areas, or also within the context of flood frequency estimation for urban catchments. Unfortunately, the temporal resolution of rainfall data usually available for practical application is often lower than the data required for the design procedures or mathematical models application. Moreover, when high resolution rain gauges are available, the registration period cannot be sufficiently long for obtaining practically usable statistical analyses. In these case the use of stochastic point process representations of rainfall which operate through the generation of random rainstorms can be a very practical alternative.

Goal of the present study is the implementation of a stochastic model that can be easily applied to derive single synthetic sub-hourly rainfall events at a point, without the need of a considerable size of rainfall data available. Some years of sub-hourly rainfall data in a site are, in fact, enough to calibrate the model.

2. Methods

Rainfall events are generated using a Monte Carlo technique with characteristics in terms of shape, duration and rainfall intensity derived by statistical analyses of the available historic records.

The first step of the analysis is the identification of the independent significant rainfall events in the collected time series. Independent rainfall events are derived on the basis of a fixed inter-event time, defined following the approach proposed by Koutsoyiannis & Foufoula-Georgiou (1993), and selected if their average intensity is bigger than a critical fixed one. The critical intensity, following the Rahman et al. (2002) approach, can be calculated as function of the 2-year return time

Intensity-Duration-Frequency (IDF) curve:
\[ I_{D,\text{thres}} = \alpha I_{D,T=2} \]

where \( \alpha \) is the reduction factor and the \( I_{D,T=2} \) is the IDF curve expressed with classical two-parameters power law. The use of appropriate values of the reduction factor \( \alpha \) allows the selection of the “significant” events in a way such that events with smaller intensity are not included. The value of this factor is in this study linked to the annual mean number of rainfall occurrences and can be evaluated by application of the Two Components Extreme Value (TCEV) distribution (Rossi et al. 1984) to the analysed sites deriving the annual mean number of rainfall occurrences. Once defined single events, the proposed rainfall model structure is based on the two following modules:

1. Temporal pattern sub-model (generation of within-storm temporal characteristics as time step intensity variations using simple statistical descriptors)
2. Intensity-Duration sub-model (statistical description and generation of storm characteristics using a multivariate model)

In order to characterise the shape dimensionless events are considered; in fact, for a particular location or within a meteorologically homogeneous regions, storms are expected to exhibit similarities in their internal structural despite their different durations and total storm depth (Koutsoyiannis and Foufoula-Georgiou, 1993). Therefore, in order to define the temporal patterns of rainfall for each selected event, we used the idea of mass curves followed by various authors (i.e. Huff, 1967; Koutsoyiannis & Foufoula-Georgiou, 1993). Following this kind of approach, the variability of precipitation within a rainy period is represented by dimensionless hyetographs. The normalised events obtained are the input for selecting an appropriate probability function for the hyetograph shape. Although any continuous density function could be appropriate to represents the shape for every analysed time-step between 0-1, here the choice has been orientated towards the "Beta distribution" because it is a very simple model that fits reasonably well the rainfall data.

For those concerning duration and average intensity (Intensity-Duration sub-model), Cordova and Rodriguez-Iturbe (1985) found that the correlation structure of rainfall intensity and storm duration has a significant effect on the probabilistic structure of storm surface runoff. Based on this study, many other authors analysed this correlation by applying several bivariate distribution and showing how duration and average intensity are variables of the same phenomenon and, consequently, they have to be correlated to each other. To assume an independence between duration and average intensity turned out to be a non-realistic hypothesis (De Michele and Salvadori, 2003). In this context, the application of copula theory to model the statistical dependence between average rainfall intensity and storm duration represents an effective choice (Salvadori et al., 2007). The advantage of the copula method is that no assumption is needed for the variables to be independent or have the same type of marginal distributions. Hence, following this approach, to generate pairs of duration and intensity, storm duration and average intensity of the observed selected events are firstly modelled separately by fitting several probability distribution and selecting the best one using several statistical criteria. Finally, they are linked to each other by using a joint probability distribution using copulas and in particular the Frank Copula (Yendra et al., 2016; De Michele and and Salvadori, 2003). De Michele and Salvadori (2003) found, in fact, how the Frank’s copula is the best candidate to model the dependence between average rainfall intensity and storm duration in comparison with other families of copulas, and also to reproduce the asymptotic statistical distribution of the storm depth.

3. Data

In order to apply the methodology, two sites placed in Sicily, Italy, were sub-hourly rainfall data were available, have been analyzed. Rainfall data at a 10 min time scale are, in fact, available for the selected sites from the database of the SIAS (Servizio Informativo Agrometeorologico Siciliano, www.sias.regione.sicilia.it). The selected sites are: Monreale raingauge station located in the North-West part of the Sicily, Italy (Fig.1), where rainfall data for a period from 2003 to 2009 are
available, and Palazzolo Acreide raingauge station located in the South-East part of the Sicily, Italy (Figure 1), where rainfall data for a period from 2002 to 2007 are available.

![Figure 1: Location of the raingauge stations used for the study.](image)

The analysed raingauge stations are characterised respectively by an average annual rainfall of 850 mm for Monreale and 600 mm for Palazzolo Acreide. Climate pattern is typical Mediterranean with heavy rainfall during the winter period and almost no rainfall during the summer period. As consequence, all the rainfall events occur in 50-55 rainy days and, in terms, of extreme events, an average value of about 15 significant rainfall events per year is observed.

4. Results

The first step of the analysis has been the selection of independent rainfall events in the collected time series. According to the approach proposed by Koutsoyiannis & Foufoula-Georgiou (2003) an inter-event equal to 2 h for Monreale raingauge station and equal to 4 h for Palazzolo Acreide raingauge station has been adopted.

The second step has been the selection of hydrologically “significant” events. According to Rahman et al. (2002) approach, the events have been selected if their average intensity was bigger than a critical fixed one, defined by applying eq. (1). The 2-year return time IDF curve has been estimated from the Report of Italian National Research Group for the Prevention of Hydro-Geological Disasters in which the Two Components Extreme Value (TCEV) distribution (Rossi et al. 1984) has been adopted for the analysis of extreme rainfall in Southern Italy.

The value of the reduction factor $\alpha$, set equal to 0.3 in both cases, has been evaluated by application of the TCEV to the analysed sites deriving the annual mean number of rainfall occurrences, equal to about 15 for Monreale station and equal to about 13 for Palazzolo Acreide station. With these assumptions, 105 independent rainfall events for Monreale and 83 independent rainfall events for Palazzolo Acreide were selected from the recorded dataset. Pertinent information and statistics of the rainfall events so derived are provided in Table 1.

Table 1: Information and statistics of the rainfall data.

<table>
<thead>
<tr>
<th>MONREALE</th>
<th>Intensity (mm/h)</th>
<th>Volume (mm)</th>
<th>Duration (min)</th>
<th>PALAZZOLO ACREIDE</th>
<th>Intensity (mm/h)</th>
<th>Volume (mm)</th>
<th>Duration (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of events</td>
<td>105</td>
<td></td>
<td></td>
<td>83</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Max</td>
<td>16</td>
<td>148.6</td>
<td>3600</td>
<td>Max</td>
<td>24.67</td>
<td>259.2</td>
<td>5430</td>
</tr>
<tr>
<td>Min</td>
<td>1.25</td>
<td>7.8</td>
<td>30</td>
<td>Min</td>
<td>0.63</td>
<td>7</td>
<td>30</td>
</tr>
<tr>
<td>Mean</td>
<td>4.37</td>
<td>31.45</td>
<td>611.14</td>
<td>Mean</td>
<td>6.38</td>
<td>41.40</td>
<td>775.42</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>3.12</td>
<td>23.70</td>
<td></td>
<td>Standard deviation</td>
<td>5.53</td>
<td>47.91</td>
<td>965.54</td>
</tr>
</tbody>
</table>
According to the non-parametric method, in order to calculate the Frank copula, the first step has been to obtain its generating function from bivariate observations; then the resulting copula has been calculated following the approach proposed by Genest and Rivest (1993). The Frank copula so obtained is characterised by a value of the characteristic parameter for the copula equal to -9.884 and -8.624 respectively for Monreale station and for Palazzolo Acreide station.

Use of copula requires, also, the determination of marginal distributions based on univariate data. The marginal distributions in this study used have been: Exponential, Gamma, Weibull and Lognormal; the parameters of these distributions have been estimated by applying the Maximum Likelihood method and the best fitted distribution has been selected using various criteria, as AIC and BIC criterion, the Relative Root Mean Square Error and the chi-squared test. The results, returned, for both stations, Lognormal probability distribution as best marginal distribution for average storm intensity, and Gamma distribution as best marginal distribution for storm duration. Finally, in order to define the temporal patterns of rainfall for each event, following the idea of mass curves, we derived the normalised events. The historical normalised mass curves have been sampled in 26 equal time-step (0; 0.04; 0.08; …; 0.96; 1) and, for each time-step considered, the parameter estimation of the Beta distribution has been carried out using the method of maximum likelihood.

In order to test the capability of the model to reproduce the rainfall events characteristics, 2x10⁴ single storm events have been generated. The synthetic events so generated have shown a very good reproducibility of the statistics of the historic observed events, both in terms of duration-intensity correlation, and also in the terms of shape (Figure 2).

![Figure 2: Comparison between average intensity and duration for generated and observed events (left); comparison between observed and generated hyetograph shape: mean and standard deviation (middle), 10th, 50th and 90th percentiles (right).](image)

This confirmed how Frank’s 2- Copula is well suited to describe the dependence structure between the available intensity-duration data, and how the assumptions of the selected marginal distributions are well respected.

In particular, the latter figure shows the performance of the generator of hyetographs of individual events. First of all, looking at the right graph of Figure 2 where the 10th, 50th and 90th percentiles of the hyetograph shape are plotted both for the generated and for the historical events, it is possible to verify how the generated curves have nonlinear shape in accordance with the observed ones. Moreover, all the curves plotted in Figure 2 show how generated and observed values at every step are very close to each other without any remarkable deviation, except, perhaps, for the standard deviation comparison where, in some point, a little discrepancy can be observed. The good results obtained provide a good indicator of how the temporal pattern sub-model allows to represent effectively the statistical characteristics of the observed event shape.
More significant for the aim of the present study are the results shown in Figure 3. In those graphs the comparison between mean of the rainfall depth for fixed duration, both for observed and for generated events, have been plotted. The historic points in the left graph refer to an average curve of the critical rainfall that has the same hydrological meaning of a Intensity Duration Frequency curve.

Figure 3: Comparison between observed and generated events: mean of the maximum depth for fixed duration.

The good agreement between historic and synthetic points is a significant indicator of the capability and the effectiveness of the proposed model to generate meaningfully extreme rainfall events as well. In addition, it’s worth to point out how there is a better agreement for the short durations than for the longer, according with one of the main goals of this study.

5. Conclusions

A stochastic rainfall model for simulating synthetic sub-hourly rainfall events at a given site, capable of replicating the main statistical characteristics of the recorded rainfall event, has been presented.

The main advantage of this kind of approach is its applicability if a sample of single historic events is available. This imply that also few years of records, not necessary continuous, are sufficient to implement the model. The model has been calibrated on two different sites in Sicily.

The application of the model has shown how the simulated events reproduce reasonably well the statistical behaviour of the observed events. The model is hence capable of representing duration, average storm intensity and temporal distribution characteristics of the point rainfall. The application of the proposed model to a given site allow to easily derive a long series of single synthetic rainfall event that can be useful for many purposes in hydrology of urban catchment.

References