Research Papers
Issue RP0256
April 2015

Regional Models and geo-Hydrological Impacts Division

Stochastic models for the disaggregation of precipitation time series on sub-daily scale: identification of parameters by global optimization

By **Veronica Villani**
Regional Models and geo-Hydrological Impacts Division, CMCC
veronica.villani@cmcc.it

Daniela di Serafino
Department of Mathematics and Physics, Second University of Naples (SUN)
daniela.diserafino@unina2.it

Guido Rianna
Regional Models and geo-Hydrological Impacts Division, CMCC
guido.rianna@cmcc.it

Paola Mercogliano
Regional Models and geo-Hydrological Impacts Division, CMCC
Meteo Systems and Instrumentation Laboratory, CIRA (Italian Aerospace Research Centre)
p.mercogliano@cira.it

SUMMARY Stochastic disaggregation model, based on coupling of the modified version of the Bartlett-Lewis Rectangular Pulse stochastic rainfall model and proportional adjusting procedure, is shown to disaggregate daily observed precipitation to hourly scale. Furthermore synthetic hourly time series are generated.This model requires the identification of a set of parameters that allow to reproduce, as well as possible, the statistical properties of the observed precipitation. The identification is formulated as a global optimization problem. A comparison between observed and modeled statistics of the precipitation time series is presented for the weather station of San Martino Valle Caudina (Southern Italy).

Keywords: Calibration, Optimization, Rainfall Disaggregation, Subdaily Precipitation, Temporal Downscaling, Synthetic Time Series Generation

The research leading to these results has received funding from the Italian Ministry of Education, University and Research and the Italian Ministry of Environment, Land and Sea under the GEMINA and Next Data projects and from EC, South East Europe Programme, OrientGate project. Observed data are property of Arpa Emilia Romagna.

INTRODUCTION

This work is part of the research activities performed at the REMHI division of CMCC with the aim of understanding and quantifying the potential impact of climate change on geo hydrological impacts regulated by sub-hourly dynamics (for instance, rainfall-induced shallow landslide and floods in small/impervious watersheds). Unfortunately, state-of-the-art climate models, including those with a high spatial-temporal resolution, do not show good predictability on this time scale, hence they need to be combined with other methods to provide useful information in this range. A common approach in literature is to take the model output at daily scale and downscale it using stochastic disaggregation rainfall models.

This work represents a first step toward the implementation of stochastic disaggregation rainfall models. In particular, this step concerns the optimization of a stochastic disaggregation model on a specific test case, in order to assess its capability of providing rainfall series on a sub-daily scale, starting from input data observed on a daily scale. Specifically, a stochastic precipitation model assuming that the precipitation events are distributed according to a Poisson process is considered [17] [18], together with a temporal disaggregation technique based on a proportional adjustment procedure [11]. The validation of the disaggregation model is performed by comparing the computed hourly precipitation with available observed hourly precipitation. This work will provide useful information for future application of the disaggregation model to simulated daily data.

In this context, a crucial issue is the identification of a set of parameters of the stochastic model that allow to reproduce, as well as possible, the statistical properties of the observed precipitation. The identification is formulated as a global optimization problem, i.e., finding the global minimum of a suitable function, which measures the “distance” between the values of appropriate statistics, expressed in terms of the model parameters, and the same statistics computed by using observed precipitation data. The minimization problem is solved using the Evolutionary Annealing-Simplex algorithm [5], which combines a direct search approach with a purely heuristic approach. A sensitivity analysis is also performed in order to understand how the model parameters computed by using the optimization algorithm are affected by variations of the input data and algorithmic parameters.

Hourly precipitation data coming from the meteorological station of San Martino Valle Caudina (AV) are used for validation. They cover about 12 years (2001-2012). These data are particularly significant as the meteorological station is close to a complex of slopes that have been affected by flow slides phenomena several times in recent years.

This work was done as part of a collaboration between the Department of Mathematics and Physics of the Second University of Naples (SUN), the Italian Aerospace Research Centre (CIRA) and the Euro-Mediterranean Centre on Climate Change (CMCC) .

THE STOCHASTIC DISAGGREGATION MODEL

The selected disaggregation model combines a modified version of the Bartlett-Lewis Rectangular Pulse stochastic rainfall model (Random Parameter Bartlett-Lewis Rectangular Pulse - RPBLRP) with a suitable rainfall disaggregation technique. The latter implements an empirical correction procedure called proportional adjusting procedure [8] [9] [14] [22], in order to modify the lower-level (e.g., hourly) time series, generated by the Bartlett-Lewis stochastic model,

so that it is consistent with a given higher-level (e.g., daily) time series.
The Bartlett-Lewis model, extensively tested with different climates and different time scales [10], has proven to properly reproduce the main features of the rainfall pattern from the hourly scale to the daily scale [17] [18] [20] [21]. For clarity and completeness, it is outlined in Appendix *The Bartlett-Lewis Rectangular Pulse Stochastic Rainfall Model*.
The adjusting procedure is described in Appendix *Proportional adjusting procedure*. We note that its application to precipitation events and cells can extend to more than a day. However, if applied over a long simulation period, the disaggregation model may require a very high computational effort. To avoid this, the simulation period should be divided into as many sub-periods as possible. To this aim, different sequences (clusters) of wet days, separated by at least one dry day, are considered as independent events. This empirical observation is consistent with the Bartlett-Lewis model, in which the arrivals of the rainfall events are modeled as a Poisson process. This allows the independent treatment of each cluster of wet days, which reduces the computation time. Then, the Bartlett-Lewis model is applied separately to each cluster of wet days. Furthermore, it is solved several times for each cluster and the generated sequence "closest" to the known value of the corresponding higher-level variable is chosen.
The Bartlett-Lewis model and the proportional adjusting procedure model are implemented in the Hyetos software system [12] [13], which is used in our experiments.

IDENTIFICATION OF THE PARAMETERS IN THE BARTLETT-LEWIS RECTANGULAR PULSE MODEL

The calibration of the BLRP model is a crucial step in the overall process. The model is calibrated using the Generalized Method of Moments (GMM), in which the model parameters are estimated by solving a suitable optimization problem. The objective function, i.e., the function to be minimized, is a weighted sum of the "distances" between the statistical moments computed using analytical expressions and the corresponding values obtained from observed data.
There are several approaches to calibration, but there are no general indications about the statistical moments to be considered, or the levels of aggregation in the objective function, or the weights to be used [23]. Several authors have empirically addressed some of these issues, coming to different conclusions [2] [1] [23]. The analytical formulas reported in Appendix *The Bartlett-Lewis Rectangular Pulse Stochastic Rainfall Model* allow to define the main statistical moments at different intervals of time aggregation. To take into account the possible variations in the rainfall pattern on the entire day, aggregations at 1, 6, 12 and 24 hours are used. In particular, for the calibration phase, the following statistics are considered at different time scales:

- mean (E_n),
- variance (Var_n),
- lag 1 covariance (Cov_n) (lag 1 stands for one hour delay),
- percentage of dry days (P_n),

where $n = 1, 6, 12, 24$ indicates the range of data aggregation used. The analytical expressions of such variables through the parameter of BLRP approach [21] are:

$$E_n = \frac{h_n \lambda \mu_c \mu_x \nu}{\alpha - 1} \tag{1}$$

$$Var_n = 2A_1 \left[h_n(\alpha-3)\nu^{2-\alpha} - \nu^{3-\alpha} + (\nu+h_n)^{3-\alpha}\right] - 2A_2 \left[h_n\phi(\alpha-3)\nu^{2-\alpha} - \nu^{3-\alpha} + (\nu+h_n\phi)^{3-\alpha}\right] \tag{2}$$

$$Cov_n = A_1[(\nu+2h_n)^{3-\alpha} - 2(\nu+h_n)^{3-\alpha} + \nu^{3-\alpha}] - A_2[(\nu+2h_n\phi)^{3-\alpha} - 2(\nu+h_n\phi)^{3-\alpha}\nu^{3-\alpha}] \tag{3}$$

$$P_n = \left\{-h_n\lambda - \frac{\lambda}{\phi}\frac{\nu}{\alpha-1} \cdot \left[1+\phi(\kappa+\phi) - \frac{1}{4}\phi(\kappa+\phi)(\kappa+4\phi) + \frac{1}{72}\phi(\kappa+\phi)(4\kappa^2+27\kappa\phi+72\phi^2)\right] + \frac{\lambda}{\phi+\kappa}\frac{1}{\alpha-1}\left[\nu+\frac{k\nu^\alpha}{\phi}(\nu+h_n)(\kappa+\phi)^{\alpha-1}\right] \cdot \left(1-\kappa-\phi+\frac{3}{2}\kappa\phi+\phi^2+\frac{1}{2}\kappa^2\right)\right\} \tag{4}$$

the constants A_1 e A_1 are defined by

$$A_1 = \frac{\lambda\mu_c\nu^\alpha}{(\alpha-1)(\alpha-2)(\alpha-3)}\left(2\mu_x^2 + \frac{\kappa\phi\mu_x^2}{\phi^2-1}\right) \tag{5}$$

$$A_2 = \frac{\lambda\mu_c\kappa\mu_x^2\nu^\alpha}{\phi^2(\phi^2-1)(\alpha-1)(\alpha-2)(\alpha-3)} \tag{6}$$

The model calibration is performed by minimizing the following function:

$$F = S_1 + S_6 + S_{12} + S_{24} \tag{7}$$

where

$$S_n = w_1\left(\frac{E_n}{E_{n^*}}-1\right)^2 + w_2\left(\frac{Var_n}{Var_{n^*}}-1\right)^2 + w_3\left(\frac{Cov_n}{Cov_{n^*}}-1\right)^2 + w_4\left(\frac{P_n}{P_{n^*}}-1\right)^2 \tag{8}$$

E_{n^*}, Var_{n^*}, Cov_{n^*} and P_{n^*} are the sample values of the statistics obtained from the observed data, and E_n, Var_n, Cov_n and P_n are the values obtained using (1) - (4). The scalars w_1, w_2, w_3 and w_4 are weights, the choice of which depends on the objective of the study, in particular on the importance of better reproducing certain statistics with respect to others. Therefore, the objective function minimizes the weighted relative square error of the selected statistics at different aggregation time scales.

The theoretical range of variability for each parameter of the RPBLRP model is the interval $[0, +\infty]$, except for α, which is assumed greater than 1 (see Appendix *The Bartlett-Lewis Rectangular Pulse Stochastic Rainfall Model*). In practice, the theoretical bounds on these parameters are restricted, identifying smaller intervals of variability for parameters with a clear physical meaning. In table 1 shows the bound on the parameters of the RPBLRP model, which define the feasible search space the four seasons.

The minimization problem is solved using the Evolutionary Annealing-Simplex (EAS) algorithm. EAS is a heuristic optimization method, which combines the Simulated Annealing algorithm with the Nelder-Mead algorithm [16]. It is based on a controlled random search technique, where a generalised Nelder-Mead method is coupled with an annealing strategy. The core of EAS is the evolution of a simplex through Nelder-Mead-type movements, according to a combination of deterministic and stochastic rules. An iteration cycle of the algorithm is reported in Appendix *Typical iteration cycle of EAS algorithm*.

Parameter	α	$\lambda(h^{-1})$	$\nu(h)$	κ	ϕ	$\mu_x(mm/h)$
Lower bound	1.01	0.001	0.1	0.001	0.001	0.001
Upper bound	40	0.1	48	2	1	12.5

Table 1
Lower and upper bounds on the RPBLRP model parameters for the autumn, winter and spring seasons.

CASE HISTORY OF SAN MARTINO VALLE CAUDINA (AV)

The weather station in San Martino Valle Caudina (Southern Italy) provides a series of precipitation data at hourly scale for about twelve years (2001-2012). The location is considered of particular interest for two main reasons: firstly, the series of available hourly rainfall data is among the most complete ones available for the Campania Region; secondly, it is installed very close to slopes that have been affected several times by flowslides (Cervinara, 1999) causing extensive damages and casualties. The hydraulic and mechanical characteristics of volcanic soils involved in these events make it relevant for triggering especially heavy precipitation on a daily/sub-daily time scale. For this reason, the development of statistical techniques to estimate the evolution of subdaily rainfall, even if the only data available are on a daily scale, can be of considerable interest, e.g., for the back-analysis of landslide events.
Figure 1 shows the evolution of the cumulative seasonal precipitation for the period 2001-2012. Winter cumulative values are consistently higher than in other seasons. The intermediate seasons (SON and MAM) return comparable cumulative values, although Autumn is characterized by more pronounced interannual variability. During the summer, precipitation values rarely exceed 200 mm. Figure 2 shows that maximum daily precipitation values are obtained during the summer and intermediate seasons and, therefore, they can be associated with convective phenomena. The maximum value is still observed during the wettest year (2010).
In Figure 3 the maximum values of hourly precipitation are shown; the higher values are recorded during the autumn season (SON). We note that, for the Mediterranean area, the autumn season is characterized by an increased frequency of extreme events, which can cause damage and casualties (Genoa, November 2011-2014; Sardinia, November 2013).

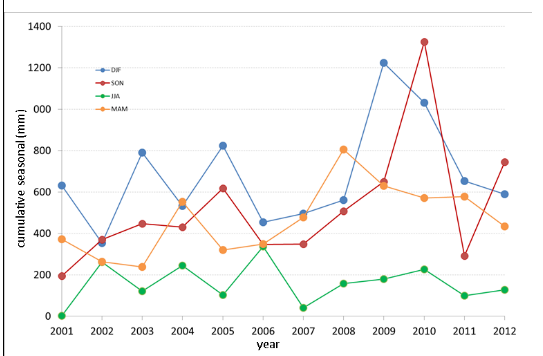

Figure 1:
Cumulative seasonal precipitation for weather station of San Martino Valle Caudina (AV).

CHOICE OF THE ALGORITHMIC PARAMETERS

A set of tests has been carried out to obtain a robust estimation of the parameters of the modified Bartlett-Lewis Rectangular Pulse model (RPBLRP). Since the EAS optimization method uses a heuristic approach and has several algorithmic parameters, the goal of the computa-

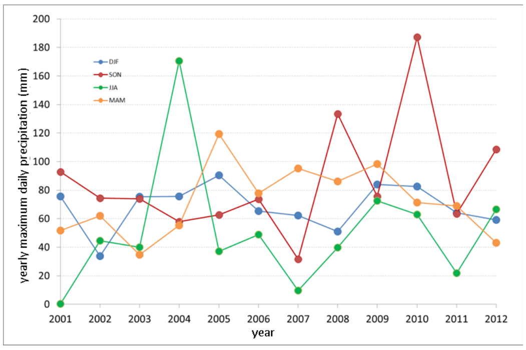

Figure 2:
Maximum daily precipitation, on seasonal scale, for weather station of San Martino Valle Caudina (AV).

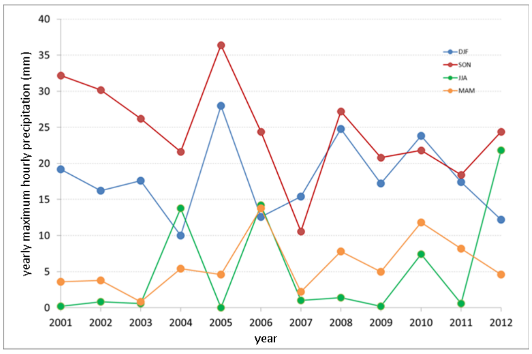

Figure 3:
Maximum hourly precipitation, on seasonal scale, for weather station of San Martino Valle Caudina (AV).

tional experiments is to define the values of the algorithmic parameters that allow to reproduce the statistics for the observed rainfall in a reasonably confident way. The experiments have been performed by using the function eas of the package HyetosR (Appendix *The Hyetos software system*), implementing an enhanced version of the EAS optimization method (Section *Identification of the parameters in the Bartlett-Lewis Rectangular Pulse model*).

A first reference test has been carried out setting the input parameters of the eas function to their default values and carrying out 20 executions of the algorithm.

Subsequently the input parameters, assumed as reference control variables, are:

- $ftol$: a positive number that specifies the fractional convergence tolerance to be achieved in the function value. Default is $ftol = 1.e - 07$.
- ratio: a positive number, typically between $0.80 - 0.99$, that specifies the fraction of temperature reduction, when a local minimum is found. Default is $ratio = 0.99$.
- $pmut$: a positive number, between $0.5 - 0.95$, that specifies the probability of accepting an offspring generated via mutation. Default is $pmut = 0.9$. Higher values are suggested for very hard problems, when it is essential to increase randomness.
- $beta$: a positive integer, greater than 1, that specifies the annealing schedule parameter. Default is $beta = 2$.
- $maxclimbs$: a positive integer, typically between $3 - 5$, that specifies the maximum number of uphill steps. Default is $maxclimbs = 5$.

As expected, by varying $ftol$ and setting all the remaining input parameters of eas to their default values, we found that decreasing the value of $ftol$ increases the execution time of the algorithm, while the variance of the values of the single parameters tends to decrease significantly (Figures 4-9). Using a very small $ftol$ value sometimes leads to exceeding the maximum number of function evaluations, without satisfying the tolerance. In order to evaluate the reliability and accuracy of the results, EAS has been run with different values of $ftol$ (ranging from $1.e - 1$ to $1.e - 12$) on 20 test sets concerning the winter season on the time span 2001-2012.

Seasons	α	$\lambda(d^{-1})$	$\nu(d)$	κ	ϕ	$\mu_x(mm/d)$
Autumn	2.0915	0.4164	0.0048	0.3181	0.0177	173.8233
Winter	2.5780	0.6030	0.0672	0.5703	0.1018	46.4391
Spring	2.7735	0.4326	0.0109	0.5882	0.0259	85.2918
Summer	40	0.0803	0.5817	0.1907	0.0473	284.8187

Table 2
Parameters of the RPBLRP model computed by minimizing the objective function.

The mean and variance of the resulting model parameters (i.e., of the solution obtained with EAS) have been computed. These statistics are reported in Figures 4-9.
According this preliminary study, it was decided to set $ftol = 1e-8$ for the winter season. This value of ftol has been chosen since it provides, for all the RPBLRP model parameters, average values slightly affected by variance and at same time sustainable computational time. With this choice of $ftol$ the relative distance of the parameters obtained from their average is generally about $1e-10$. For similar reasons, $ftol = 1e-8$ was chosen for the springer season too, while $ftol = 1e-10$ was selected for the summer and autumn seasons.
By varying $ratio$ and setting the remaining input parameters to their default values, we found that the more the temperature decreases, the faster the algorithm converges. On the other hand, the solution is more accurate when the temperature decreases more slowly. No significant changes have been observed in the solution by varying other algorithmic parameters.
In spite the parameters differ from their mean value with a variance of $1e-10$, among the 20 sets of RPBLRP model parameters resulting from the application of the EAS algorithm, those corresponding to the lower value of the objective function are chosen. Table 2 shows these parameters for the autumn, winter, spring and summer seasons.

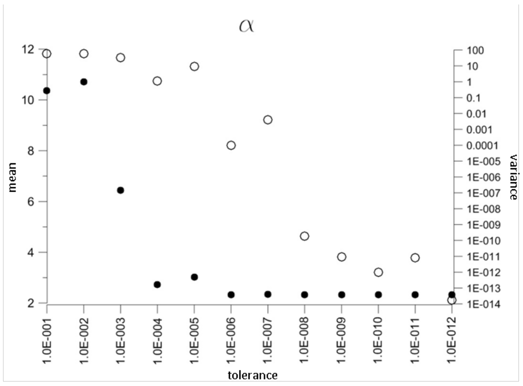

Figure 4:
Mean (●) and variance (○) related to the α parameter of the RPBLRP model to varying of ftol.

PHENOMENOLOGICAL EXPLANATION OF THE PARAMETERS OF THE RANDOM PARAMETER BARTLETT-LEWIS RECTANGULAR PULSE MODEL

The parameters of the RPBLRP model can be related to different aspects of the structure of precipitation events. Here, the effect of seasonal variability on the parameters of the RPBLRP model is examined. Table 3 shows the average duration of the interarrival time of events $(1/\lambda(h))$, the average duration of the cells $(E(1/\eta)(h))$, the average number of cells per event (μ_c), the average intensity of the cells $(\mu_x(mm/h))$, and the average duration of the event $(\mu_T(h))$ for the precipitation data of

Seasons	$1/\lambda(h)$	$E(1/\eta)(h)$	μ_c	$\mu_x(mm/h)$	$\mu_T(h)$
Autumn	57.6313	0.1065	19.0078	7.2426	6.0633
Winter	39.8018	1.0221	6.5996	1.9350	10.5898
Spring	55.4754	0.1481	23.7005	3.5538	5.7953
Summer	298.7913	0.3580	5.0318	11.8675	7.6477

Table 3

Statistical properties of precipitation series in question as a function of the calculated parameters of the RPBLRP model.

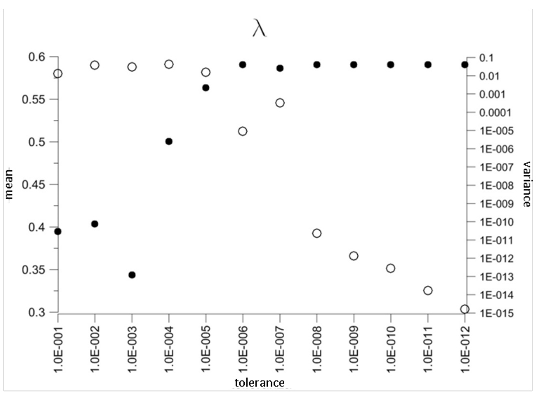

Figure 5:
Mean (•) and variance (○) related to the λ parameter of the RPBLRP model to varying of ftol.

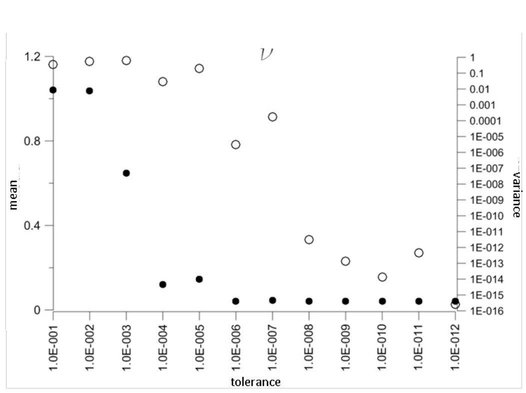

Figure 6:
Mean (•) and variance (○) related to the ν parameter of the RPBLRP model to varying of ftol.

the meteorological station of San Martino Valle Caudina. A comparison between the results on the observed data in Appendix *Identification of independent events of precipitation* and the results in Table 3 clearly shows that the RPBLRP model underestimates the duration of the inter-arrival time of the events, probably taking, on average, a "time of independence" of events (t_{b_min}) much smaller than the time obtained with the methods described in Appendix *Identification of independent events of precipitation*. Similar results have been found for weather stations in Belgium [24]. This underestimation is reflected in the low values found for the average time of the event and in the high values found for the average intensity of the event, for the four seasons. On the other hand, it must be taken into account that the identification of independent events is the result of a model too, and hence it can be obviously affected by uncertainties comparable with those found using the Bartlett-Lewis model.

APPLICATION TO THE PRECIPITATION DATA OF THE METEOROLOGICAL STATION OF SAN MARTINO VALLE CAUDINA (AV)

After fixing the parameters of the RPBLRP model, the hourly precipitation series from the daily precipitation historical series have been estimated. The performance of the disaggregation model has been tested first, i.e., the adequacy of the Bartlett-Lewis model and its parameters, and of the disaggregation technique,

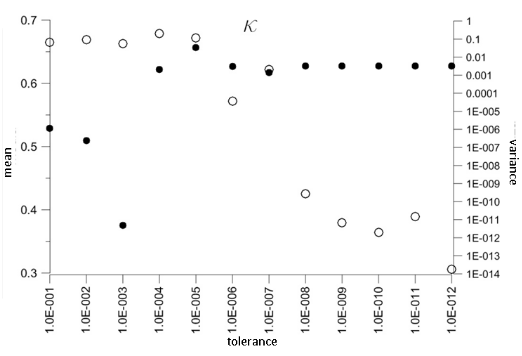

Figure 7:
Mean (•) and variance (○) related to the κ parameter of the RPBLRP model to varying of ftol.

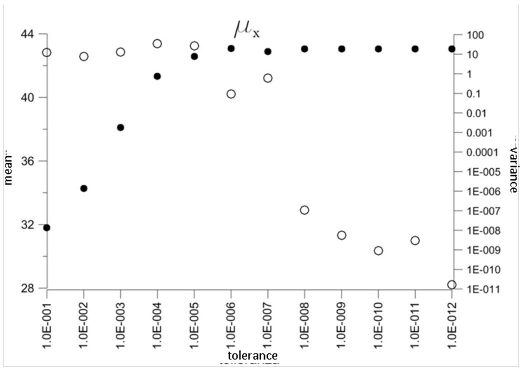

Figure 9:
Mean (•) and variance (○) related to the μ_x parameter of the RPBLRP model to varying of ftol.

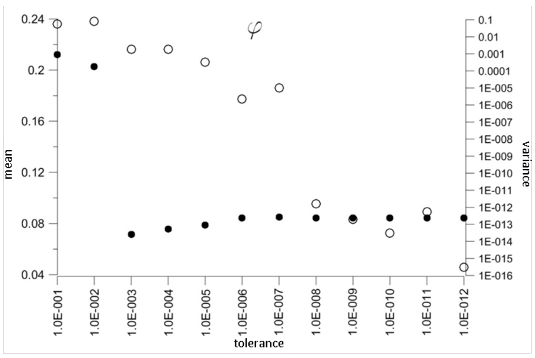

Figure 8:
Mean (•) and variance (○) related to the φ parameter of the RPBLRP model to varying of ftol.

to preserve the statistical properties of the precipitation time series. The test has been performed by using the function $DisagSimul.test$ of the package HyetosR (Appendix *The Hyetos software system*).

The argument $RepetOpt$ of the function $DisagSimul.test$ is a list of parameters, listed below, that specifying any changes to default values of the repetition schema and disaggregation illustrated in Appendix *Proportional adjusting procedure*.

- $DistAllowed$: a positive number that specifies the distance used to judge whether synthetic daily depths resemble the real ones. Default is $DistAllowed = 0.1$.
- $FacLevel1Rep$: a positive number that specifies the factor for Level 1 repetitions allowed. Default is $FacLevel1Rep = 20$.
- $MinLevel1Rep$: a positive integer that specifies the minimum number of Level 1 repetitions allowed. Default is $FacLevel1Rep = 50$.
- $TotalRepAllowed$: a positive integer that specifies the total repetitions allowed. Default is $TotalRepAllowed = 5000$.

The number of repetitions of Level 1 is determined by multiplying $FacLevel1Rep$ by the number of attempts made to establish an appropriate sequence of wet days (repetitions of the level are established by the user, and not determined by the program). In this way, the higher the number of required repetitions of Level 0, the greater the number of repetitions of Level 1, and the fewer the number of repetitions of Level 2. However, the number of repetitions of Level 1 cannot be smaller than $MinLevel1Rep$ and the total number of repetitions of Levels 1 and 2 cannot exceed $TotalRepAllowed$.

In our tests, the model is applied with the default parameters of the argument $RepetOpt$. In fact, no reasons arise to use a maximum allowed distance ($DistAllowed$) smaller than 0.1, because there is no gain in terms of reduction of data distortion. From the tests carried out, it was found that significantly larger values of $DistAllowed = 0.1$, even if leading to a smaller computation time, can significantly increase the variation and the asymmetry of the precipitation process [13].

Then, the $SequentialSimul$ function of the package HyetosR has been used for the generation of precipitation series on daily and hourly time scales, using the Bartlett-Lewis Rectangular Pulse precipitation model with the selected parameters, without performing any disaggregation. Tables 4-7 show the results of tests carried out applying $DisagSimul.test$ with its default parameter values, using the time series of daily and hourly precipitation, and the selected parameters of the RPBLRP model for autumn, winter, spring and summer. Theoretical statistics, computed by using the analytical expressions given in Section *Identification of the parameters in the Bartlett-Lewis Rectangular Pulse model* and the selected RPBLRP model parameters, as well as synthetical statistics, i.e., statistics concerning the synthetic precipitation series, obtained applying SequentialSimul with its default parameter values, are also reported. Note that $Autocorrelationi$ with $i = 0, ..., 10$ indicates $lagi$ autocorrelation, where $lagi$ stands for a delay of i hours.

Historical statistics are properly reproduced, although only four values are actually used for the estimation of the RPBLRP model parameters. The estimated parameters of the RPBLRP model, shown in Table 3, are physically reasonable, confirming the ability of the model to capture the general temporal structure of rainfall events.

In terms of variation, a generally satisfying preservation of the disaggregated series compared to the original ones is observed, but with a slight positive bias, due to the fact that the distance of the values of daily precipitation, obtained as the sum of the lower level variables of the disaggregated series, from the values of the original daily precipitation is below 0.01 [13].

In terms of asymmetry coefficients, the Bartlett-Lewis model fails to preserve the asymmetry of the temporal distribution of precipitation and, therefore, the disaggregation model cannot explicitly preserve the asymmetry of the distribution.

Figures 10(a)-10(d) show the trend of the autocorrelation coefficients of the hourly intensity of the precipitation for lags up to 10 hours. It worth noting that the disaggregation process simulates a decrease faster than that observed, provided by the synthetic series, or directly obtained from the RPBLRP model parameters. In particular, the last two series tend to coincide in most cases, showing how the generation of a synthetic series of 100 years effectively reproduces average characteristics of the precipitation through the stochastic approach. On average, the performances of the synthetic series and of the values obtained directly from the RPBLRP model parameters are better than those of the disaggregated series, showing that the time disaggregation is the step that causes the greatest error in the prediction. A RPBLRP model parameter critical for the function of $lag1$ autocorrelation is the shape parameter α. The inspection of the correlation structure reveals a dependence of the following type:

$$Corr_n \approx h_n(lag-1) + \nu^{3-\alpha} + (lagh_n + \nu)^{3-\alpha} + h_n(lag+1) + \nu^{3-\alpha} \quad (9)$$

which, for $\alpha \to 2$, presents a slow decay [15], while for $\alpha > 3$ shows a fast decay.

Statistics	Original	Disaggregated	Theoretical	Synthetical
Hours	26064	26064		
Mean(mm)	0.2408	0.2408	0.2516	0.2493
Standard Deviation(mm)	1.3142	1.4757	1.3173	1.3097
Variation	5.4583	6.1291	5.2352	5.2539
Asymmetry	10.0288	11.8781		10.8997
Autocorrelation 0	1	1	1	1
Autocorrelation 1	0.5032	0.4597	0.5079	0.5200
Autocorrelation 2	0.3463	0.2847	0.3416	0.3571
Autocorrelation 3	0.2721	0.2293	0.2799	0.2956
Autocorrelation 4	0.2280	0.1894	0.2412	0.2575
Autocorrelation 5	0.1795	0.1757	0.2132	0.2291
Autocorrelation 6	0.1619	0.1429	0.1916	0.2076
Autocorrelation 7	0.1537	0.1366	0.1741	0.1884
Autocorrelation 8	0.1381	0.1103	0.1596	0.1737
Autocorrelation 9	0.1132	0.1038	0.1474	0.1620
Autocorrelation 10	0.1196	0.0887	0.1370	0.1557
Proportion Dry	0.8921	0.9195	0.9035	0.9044

Table 4

Comparison, for the autumn season, between the values of the statistics of the disaggregated and synthetical historical series and the values of the theoretical statistics.

Statistics	Original	Disaggregated	Theoretical	Synthetical
Hours	25800	25800		
Mean(mm)	0.3130	0.3130	0.3280	0.3301
Standard Deviation(mm)	1.1600	1.2700	1.1386	1.1341
Variation	3.7100	4.0800	3.4710	3.4352
Asymmetry	7.0000	6.1200		5.3804
Autocorrelation 0	1	1	1	1
Autocorrelation 1	0.6600	0.7070	0.7244	0.7236
Autocorrelation 2	0.4900	0.4530	0.5076	0.5077
Autocorrelation 3	0.3910	0.3090	0.3988	0.4007
Autocorrelation 4	0.3230	0.2180	0.3322	0.3367
Autocorrelation 5	0.2600	0.1580	0.2865	0.2914
Autocorrelation 6	0.2250	0.1250	0.2527	0.2574
Autocorrelation 7	0.1930	0.1000	0.2264	0.2309
Autocorrelation 8	0.1630	0.0815	0.2051	0.2086
Autocorrelation 9	0.1380	0.0704	0.1876	0.1903
Autocorrelation 10	0.1320	0.0684	0.1727	0.1757
Proportion Dry	0.8140	0.8504	0.8219	0.8128

Table 5

Comparison, for the winter season, between the values of the statistics of the disaggregated and synthetical historical series and the values of the theoretical statistics.

Statistics	Original	Disaggregated	Theoretical	Synthetical
Hours	26376	26376		
Mean(mm)	0.2106	0.2106	0.2240	0.2242
Standard Deviation(mm)	0.9851	1.1325	0.9916	0.9923
Variation	4.6774	5.3773	4.4262	4.4249
Asymmetry	9.2444	11.3507	7.6248	
Autocorrelation 0	1	1	1	1
Autocorrelation 1	0.5403	0.5393	0.5465	0.5456
Autocorrelation 2	0.3704	0.3218	0.3851	0.3834
Autocorrelation 3	0.2993	0.2438	0.3203	0.3206
Autocorrelation 4	0.2408	0.2065	0.2759	0.2753
Autocorrelation 5	0.2064	0.1671	0.2419	0.2450
Autocorrelation 6	0.1731	0.1185	0.2146	0.2161
Autocorrelation 7	0.1574	0.1086	0.1921	0.1925
Autocorrelation 8	0.1502	0.0983	0.1731	0.1753
Autocorrelation 9	0.1477	0.0789	0.1570	0.1607
Autocorrelation 10	0.1415	0.0846	0.1432	0.1445
Proportion Dry	0.8821	0.9070	0.8933	0.8927

Table 6

Comparison, for the spring season, between the values of the statistics of the disaggregated and synthetical historical series and the values of the theoretical statistics.

Statistics	Original	Disaggregated	Theoretical	Synthetical
Hours	22824	22824		
Mean(mm)	0.0688	0.0688	0.0715	0.0699
Standard Deviation(mm)	1.0097	1.1780	0.9822	0.9698
Variation	14.6698	17.1150	13.7332	13.8810
Asymmetry	36.0875	49.9468		26.3492
Autocorrelation 0	1	1	1	1
Autocorrelation 1	0.3520	0.3311	0.3546	0.3413
Autocorrelation 2	0.1365	0.0831	0.1449	0.1392
Autocorrelation 3	0.0396	0.0300	0.1149	0.1093
Autocorrelation 4	0.0212	0.0248	0.0998	0.0974
Autocorrelation 5	0.0139	0.0351	0.0875	0.0851
Autocorrelation 6	0.0137	0.0336	0.0768	0.0767
Autocorrelation 7	0.0103	0.0308	0.0675	0.0644
Autocorrelation 8	0.0059	0.0769	0.0593	0.0568
Autocorrelation 9	0.0081	0.0271	0.0522	0.0491
Autocorrelation 10	0.0997	0.0231	0.0459	0.0473
Proportion Dry	0.9753	0.9816	0.9836	0.9838

Table 7

Comparison, for the summer season, between the values of the statistics of the disaggregated and synthetical historical series and the values of the theoretical statistics.

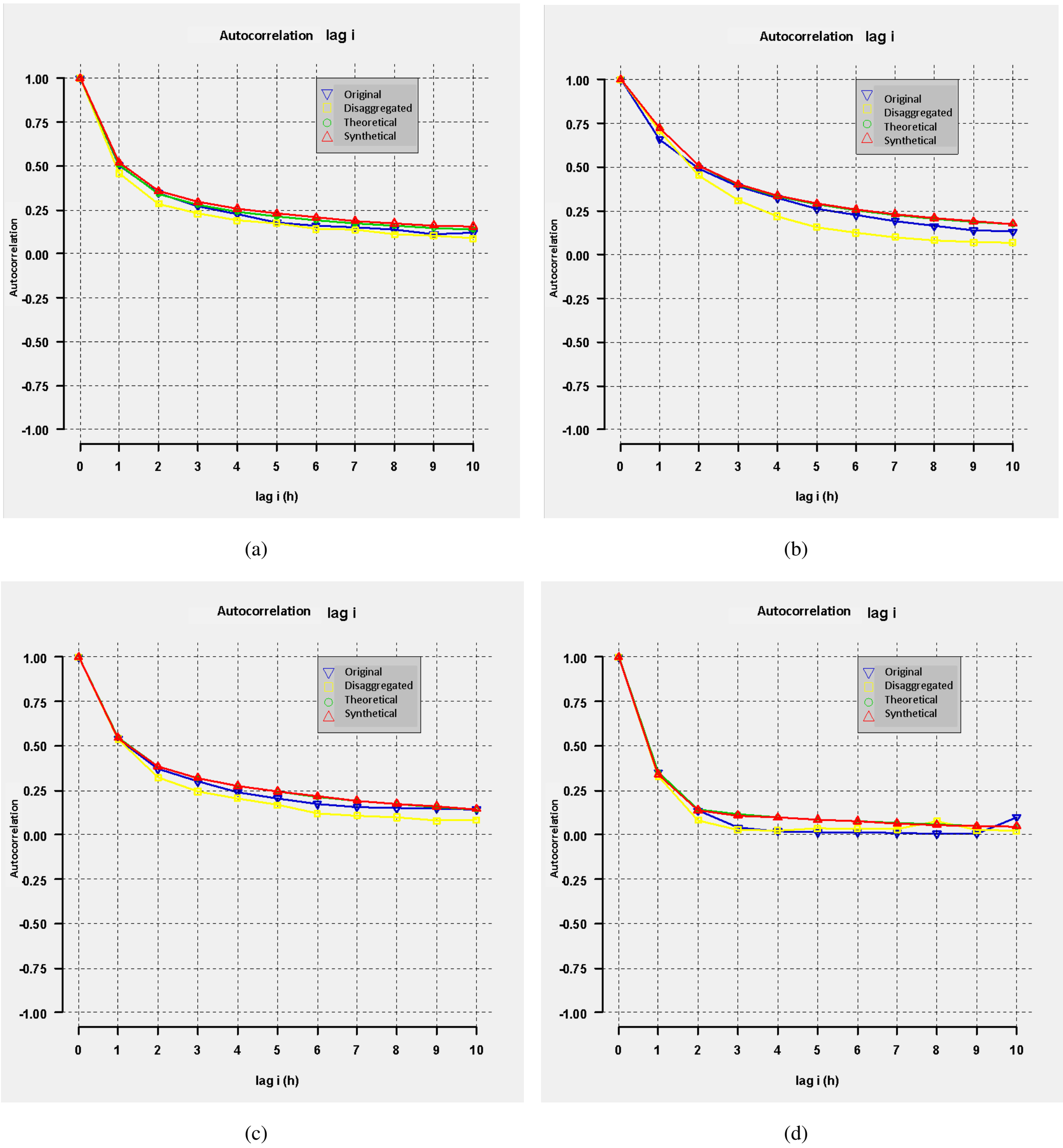

Figure 10:
Comparison, for the (a) autumn, (b) winter, (c) spring and (d) summer season, among the values of the autocorrelation, with a delay of i hours ($i = 0, \ldots, 10$), of the historical series, and the disaggregated, synthetical and theoretical ones.

CONCLUSIONS

The aim of this work was to validate the stochastic model of disaggregation of precipitation mentioned in Section *The stochastic disaggregation model* and described more in detail in the Appendix, using hourly observed data together with an analysis of the influence of the parameters of the Bartlett-Lewis model on the statistical properties of the simulated precipitation process, aimed at assessing the reliability of the model when there are no hourly observed data.

The identification of the parameters of the Bartlett-Lewis model was formulated as a global optimization problem and has been solved using the Evolutionary Annealing-Simplex algorithm. Given the stochastic nature of this algorithm an analysis has been carried out to understand the sensitivity of the optimal solution (i.e., the parameters of the Bartlett Lewis model) to some algorithmic parameters, and to choose the latter so that the computed Bartlett-Lewis model parameters can provide a reliable reproduction of the statistical properties of the observed precipitation.

The quality of the results was assessed by comparing the main statistics of observed, disaggregated, and synthetic precipitation data and those obtained by using analytical expressions with the computed values of the parameters of the Bartlett-Lewis model. This comparison showed that the statistics of the observed precipitation data are properly reproduced. On average, the predictive capabilities of the synthetic series and the values obtained directly by using the parameters of the Bartlett-Lewis model are better than those of the disaggregated series. Nevertheless the results are adequate to the target set.

APPENDIX

THE BARTLETT-LEWIS RECTANGULAR PULSE STOCHASTIC RAINFALL MODEL

In the Bartlett-Lewis Rectangular Pulse (BLRP model), arrivals of storms are modeled as a Poisson process with rate λ (by arrival we mean a random event that occurs in a temporal reference system). The probability $P(s)$ that s events occur in an interval Δt is therefore defined as

$$P(s) = (\lambda \Delta t)^s \frac{e^{-\lambda \Delta t}}{s!} \quad (10)$$

The interarrival time between successive storms, i.e., the time t_s between two consecutive arrivals, is exponentially distributed with rate λ:

$$P(T_s \leq t_s) = 1 - e^{-\lambda t_s} \quad (11)$$

A random number of cells is assigned to each storm. Each cell is a precipitation rectangular pulse, with random duration t_d and random intensity i. Cell arrivals are modeled as a Poisson process with parameter β. Thus, the probability $P(c)$ that c cells occur in a range Δt is

$$P(c) = (\beta \Delta t)^c \frac{e^{-\beta \Delta t}}{c!} \quad (12)$$

The interarrival time, t_c, between two subsequent cells is exponentially distributed with parameter β:

$$P(T_c \leq t_c) = 1 - e^{-\beta t_c} \quad (13)$$

The cells generation process finishes after a time t_g, exponentially distributed with rate γ:

$$P(T_g \leq t_g) = 1 - e^{-\gamma t_g} \quad (14)$$

Finally, for the duration t_d and the intensity i of the cells an exponential distribution is assumed too, with rates η and $\frac{1}{\mu_x}$, respectively:

$$P(T_d \leq t_d) = 1 - e^{-\eta t_d} \quad (15)$$

$$P(I \leq i) = 1 - e^{-\frac{i}{\mu_x}} \quad (16)$$

where μ_x is the average intensity of the cells of rainfall. Note that an alternative approach [18], assumes that the intensity i is a two-parameter gamma distribution with mean μ_x and standard deviation $\frac{1}{\sigma_x}$.

The model states that the average number of cells for each event is generated by a geometric distribution with mean

$$\mu_c = 1 + \frac{\kappa}{\phi} \quad (17)$$

where κ and ϕ are dimensionless parameters defined as follows:

$$\kappa = \frac{\beta}{\eta} \quad (18)$$

$$\phi = \frac{\gamma}{\eta} \quad (19)$$

The generation process is illustrated graphically in Figure 11.

The BLRP model does not return satisfactory results due to a significant overestimation of the occurrence of dry periods [21]; this does not make it suitable when an accurate description is required of the alternation of dry periods and wet periods. To overcome this problem, a modified version of the model has been developed, named Random Parameter Bartlett-Lewis Rectangular Pulse (RPBLRP).

The modified model improves the BLRP model by assuming that the parameter η, which the duration of the cells depends on, may vary randomly from event to event, according to a Gamma distribution with a shape parameter α and a scale parameter ν [21]. Therefore, the probability density function, the expected value and the variance of the distribution are described by the following expressions, respectively:

$$f(\eta) = \frac{\nu^\alpha}{\Gamma(\alpha)} e^{-\nu\eta} \eta^{\alpha-1} \quad (20)$$

where

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt \quad (21)$$

is the Gamma function,

$$E(\eta) = \frac{\alpha}{\nu} \tag{22}$$

$$Var(\eta) = \frac{\alpha}{\nu^2} \tag{23}$$

To ensure that the duration of events is greater than zero, $\alpha > 1$ is assumed [21]. While in the original version of the model all the parameters are assumed to be constant, in the modified version η is variable and the $\kappa = \frac{\beta}{\eta}$ and $\phi = \frac{\gamma}{\eta}$ are constant.
The modified model allows to have distinct events with a common structure, but on different time scales; different perturbations are composed of cells characterized by variable duration and variance for each event (respectively η^{-1} and η^{-2}). Similarly, the average interval of interarrival of the cells and the average duration of the perturbation (respectively β^{-1} and γ^{-2}) vary randomly, but keeping κ and ϕ constant throughout the process. Therefore, for larger values of η (i.e., for shorter durations of the cells), the interarrival time of the cells and the duration of the corresponding perturbation are smaller. This is physically plausible because longer perturbations are characterized by cells that tend to have longer "life" and longer interarrival times [21].
From the above expressions other physical characteristics can be derived, which contribute to determining the structure of the precipitation; in particular, it is possible to obtain the average number, duration and intensity of the cells forming the event: the average number of cells per event is provided by 17; the average duration of the cells is

$$E\left(\frac{1}{\eta}\right) = \frac{\nu}{\alpha - 1} \tag{24}$$

the average duration of the event can be approximated as follows:

$$\begin{aligned} \mu_T \cong \frac{\nu}{\phi(\alpha-1)} & \left[1 + \phi(\kappa+\phi) - \right. \\ & \frac{1}{4}\phi(\kappa+\phi)(\kappa+4\phi) + \\ & \left. \frac{1}{72}\phi(\kappa+\phi)\left(4\kappa^2 + 27\kappa\phi + 72\phi^2\right)\right] \end{aligned} \tag{25}$$

the average of the interarrival times of the cells is obtained from 18:

$$\beta = \kappa E(\eta) \tag{26}$$

and by 22 it becomes

$$\beta = \kappa \frac{\alpha}{\nu} \tag{27}$$

Therefore, in its simplest version the model uses five parameters, λ, β, γ, η and μ_x (or equivalently λ, κ, ϕ, η and μ_x) and in its improved version six parameters, κ, ϕ, α, ν, μ_x and σ_x. The latter version, called Random Parameter Bartlett-Lewis Gamma Model, provides a better reproduction of the extreme values of precipitation [18].

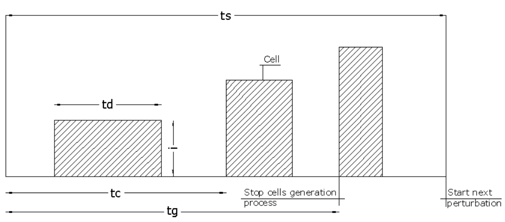

Figure 11:
Graphical representation of the Bartlett-Lewis Rectangular Pulse model.

PROPORTIONAL ADJUSTING PROCEDURE

The expression "adjusting a precipitation time series" refers to a modification of a lower-level (e.g., hourly) time series, generated for example by the Bartlett-Lewis stochastic model, so that it is consistent with a given higher-level (e.g., daily) time series, i.e. the sum of the lower-level values within a period must be equal to the corresponding observed upper-level value.

If a series of data Z_p $(p = 1, 2,)$ is known at an upper level time scale and a synthetic series $\tilde{X}_s$ is generated by a stochastic model at a lower-level time scale $(s = 1, 2,)$, the problem is to modify $\tilde{X}_s$ into another series X_s that is consistent with the higher-level series. As mentioned above, we use an adjusting procedure to preserve the additive property, i.e. to annihilate the difference between the sum of the lower-level variables within a period and the corresponding upper-level variable. The procedure should be accurate in the sense that it should preserve some statistics or even the complete distribution of the lower-level variables, at least under certain conditions. The proportional adjusting procedure is considered in this work [3] [11].

The proportional adjusting procedure modifies the original values $\tilde{X}_s$ to obtain the adjusted values X_s according to the formula

$$X_s = \tilde{X}_s \left(\frac{Z}{\sum_{j=1}^{k} \tilde{X}_j} \right) \quad (s = 1, ..., k) \tag{28}$$

where Z is the upper-level variable (for simplicity of notation $p = 1$ is assumed and the subscript p from Z is eliminated) and k is the number of lower-level variables within a higher-level period. The adjusting procedure proved to be the most appropriate for the disaggregation of the precipitation [13]. It is effective in preserving distributions, if the variables X_s are independent and follow a Gamma distribution with common scale parameter, and also provides a good approximation to the dependent variables with gamma distribution [3]. Moreover, it has the advantage of not generating negative values.

The simulation chain is reported in Figure 12, with reference to the disaggregation of daily rainfall depths of a period of L consecutive wet days (preceded and followed by at least one dry day). The scheme incorporates the four levels of repetition illustrated below.

- Level 0: the Bartlett-Lewis model is run several times until a sequence of L wet days equal to the observed one is generated.

- Level 1: the intensities of all cells and precipitation events are generated and the resulting daily depths are calculated; the latter are compared with the original ones through a "logarithmic distance" given by the following expression:

$$d = \left[\sum_{i=1}^{L} \ln \left(\frac{Z_i + c}{\tilde{Z}_i + c} \right)^2 \right]^{1/2} \tag{29}$$

  where Z_i and $\tilde{Z}_i$ are the original and generated daily depths of day i, respectively, and c is a constant $(c = 0.1mm)$. The logarithmic transformation is chosen to avoid the dominance of high values and the constant c is introduced to avoid too small values. If the "distance" d is greater than an accepted threshold d_a, then the intensities of the cells are regenerated (Level 1 repetitions) without modifying the time locations of the precipitation events and their cells. If, however, after a large number of Level 1 repetitions, the distance remains greater than the accepted threshold, the arrangement of precipitation events and cells is assumed not consistent with the original one. In this case we pass to the Level 2 repetitions.

- Level 2: the arrangement of precipitation events and cells not consistent with the original ones is deleted and a new one is generated.

- Level 3: if, for long sequences of wet days, a sequence of wet days for which the distance between the observed and simulated daily precipitation values and the daily precipitation values, obtained as the sum of the generated lower level variables (e.g. hourly), is lower than the accepted limit d_a, is not achievable, the sequence is divided into subsequences (randomly), to obtain a plausible sequence. The sequence with distance smaller than d_a is further processed to determine the lower level (hourly) rainfall depths through the application of the proportional adjusting procedure.

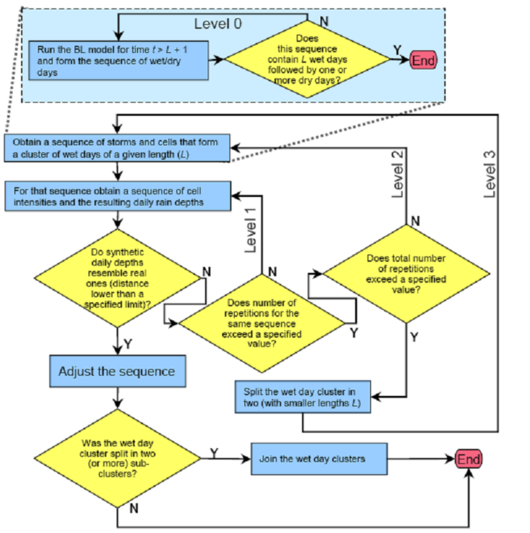

Figure 12:
Relationship between intensity and duration of the observed precipitation events series, identified by the method using the Pearson correlation coefficient [4]

THE HYETOS SOFTWARE SYSTEM

The disaggregation model described above is implemented in software Hyetos [12] [13]. It supports both the original version of the Bartlett-Lewis Rectangular Pulse model and the modified version. For practical reasons, the model implementation is specified for the daily higher level and the hourly lower level scales, although the methodology described above can also be used with other time scales. In this work, the tests are carried out using the package HyetosR, an updated version of the software Hyetos, developed in R programming environment (http://cran.r-project .org). The Hyetos software can operate in the following modes:

- *Disaggregation test mode* (without input). An initial sequence of precipitation events is generated using the Bartlett-Lewis model with pre-fixed parameters. The obtained data are aggregated to an hourly and daily scale, and the latter are used as source data for the disaggregation in hourly time scale. This mode can be useful to test the model and compare the statistics of the original and disaggregated data.

- *Full test mode* (with hourly input). The observed hourly data must be provided as an input to the model. Unlike the mode 1, the original sequence is read from the file rather than generated. This operation mode can be useful to verify the reliability of the Bartlett-Lewis model parameters obtained from the model calibration, by comparing the statistics of the observed data with those of the disaggregated data.

- *Operational mode* (with daily input). It is very similar to mode 2, but the input data are only daily. This is the ordinary operation mode.

- *Rainfall model test mode* (with hourly input). It is very similar to mode 2, but the obtained hourly data are not derived from the disaggregation of observed daily data: they are generated by the Bartlett-Lewis model with fixed parameters. The adjustment function must be disabled in the software to use this operation mode and the distance d defined by (7) must have a very large value (e.g. 1000). This mode is appropriate for testing how the Bartlett-Lewis model fits the observed data in terms of several statistics.

- *Simple rainfall generation mode* (without input and disaggregation). This mode is similar to mode 4, but without input data. This operation mode is used to generate rainfall series using the Bartlett-Lewis model with the fixed parameters without performing any disaggregation.

IDENTIFICATION OF INDEPENDENT EVENTS OF PRECIPITATION

To identify the ranges of variability of the parameters of the Bartlett-Lewis model, from the physical point of view, and thus to characterize duration and intensity of single rainfall events, the identification of minimum interval of no rain that ensures the independence of subsequent events represents a crucial issue. To this aim, in this work, three of the most well-known literature methods are selected. The first one is the "coefficient of variation" method [19]. In this method, the occurrence of precipitation events is characterized by a Poisson process of parameter λ (Appendix *Proportional adjusting procedure*). The interarrival time between successive independent precipitation events, indicated with t_s, is exponentially distributed with parameter λ. In fact, note that the derivative of the cumulative distribution function

$$F(t_s) = P(T_s \leq t_s) = 1 - e^{-\lambda t_s} \quad (30)$$

is the probability density function as follows.

$$f(t_s) = \frac{dF}{dt_s} = \lambda e^{-\lambda t_s} \quad (31)$$

This probability distribution function for the interarrival time is exponential distribution. Therefore the interarrival times of a Poisson process are independent and exponentially distributed with a mean of $\frac{1}{\lambda}$ and variance equal to $\frac{1}{\lambda^2}$. As known, in the case of an exponential distribution mean and standard deviation of the sample assume the same value, and therefore the coefficient of variation is equal to one.

$$\mu = \frac{1}{\lambda} \quad (32)$$

$$\sigma^2 = \frac{1}{\lambda^2} \quad (33)$$

$$CV = \frac{\sigma}{\mu} = 1 \quad (34)$$

This information can be used to identify independent precipitation events. the time between the end of an event and the beginning of subsequent independent one is well described by an exponential distribution, such as the interarrival time, because for a Poisson process exponentially distributed, the starting point of the time of arrival is arbitrary. However, not all periods without precipitation occurring in nature are exponentially distributed. The periods without precipitation, are analyzed progressively eliminating from the entire sample the period without precipitation respect to a specified time t. At each step is calculated the coefficient of variation of the remaining periods without precipitation. The sample containing values greater or equal to t, which has a coefficient of variation equal to one, determines the minimum time between independent precipitation events, t_{b_min}. The second method to detect statistically independent precipitation events is based on the Pearson correlation coefficient [7]. This coefficient is calculated between the series of precipitation under consideration and the same series of precipitation shifted by a specified time t. As minimum time between independent precipitation events, t_{b_min}, is chosen time t for which the Pearson correlation coefficient is less than

Seasons	$t_b(h)$	$t_{b\_min}(h)$	Average duration of the event(h)	Average intensity of the event(mm/h)	Height of average precipitation event(mm)	Average duration of interarrival time of events(h)
(1)						
Autumn	92.65	17	24.20	1.12	28.14	116.85
Winter	55.06	11	23.11	0.87	24.51	78.17
Spring	98.28	23	33.87	0.89	28.22	132.15
Summer	184.71	23	17.23	1.37	14.94	201.94
(2)						
Autumn	89.33	15	22.50	1.13	26.93	111.83
Winter	57.13	12	24.74	0.85	25.68	81.87
Spring	74.70	14	18.81	1.14	19.96	93.51
Summer	101.09	3	3.33	2.03	7.81	104.42
(3)						
Autumn	92.65	6	11.57	1.26	17.93	104.22
Winter	55.05	6	15.79	0.93	18.84	70.84
Spring	53.60	6	11.02	1.23	14.48	64.62
Summer	184.71	6	5.11	1.86	9.26	189.82

Table 8

Average characteristics of the events for the precipitation of San Martino Valle Caudina on season scale according to: **(1)** method of the coefficient of variation; **(2)** method using the Pearson correlation coefficient; **(3)** method of Huff.

10%.

The third method used is the method of Huff which assumes equal to 6 hours the minimum time span without precipitation identifying independent precipitation events [6]. This method has been used to completeness since, along with the two methods presented above, it is the most used method in the literature, but less accurate because it does not take into account neither of seasonal characteristics nor of climatology of the study area.

In Table 8 are reported the average interval of no rain, t_b; the minimum time for the identification of independent rain events, t_{b_min}; the average duration of the event; the average intensity of the event; the average height of the precipitation event; the average duration of the interarrival time of the events. It refers respectively to the method of the coefficient of variation, the method using the Pearson correlation coefficient and the method of Huff. Although in many works of literature is preferred a subdivision on a monthly scale, in this work a seasonal scale is adopted. The main reason for this choice lies in the possibility to have richer data sets despite the only 12 years available, while preserving the main physical dynamics of rainfall events.

The first two methods ensure that t_{b_min} is sufficiently long so that it becomes statistically insignificant for the autocorrelation of precipitation. However, select a very long t_{b_min} it may be useful to ensure the identification of statistically independent events, but, in turn, has a negative effect on the properties of the precipitation event, creating large distances intra-event and leading to a strong bias of the average duration and the average intensity of the event [19], as shown in the table below. Both approaches are able to define a consistent seasonal pattern. In winter, the average interval between events is comparable (55.06 h vs.57.13 h) with events of a duration of about 24 hours (23.11 h vs.24.74 h) and intensity less than a millimeter per hour (0.87 mm/h vs.0.85 mm/h). In the intermediate seasons, the time interval between independent events grows but, while according to the method of the coefficient of variation is for both seasons close to 4 days, the Pearson autocorrelation coefficient identifies, as a time interval between independent events for spring, a value slightly greater than 3 days. This difference also causes significant changes in the estimate of the average duration of the event that, for the spring, passes from almost 34 hours to about 19 hours with the respective intensity which increased from 0.89 to 1.14 mm/h. For autumn the two methods return similar estimations with comparable duration and intensity (24.20 h vs 22.50 h e 1.12 mm/h vs 1.13 mm/h). The difference observed in the average duration of the event in the spring, can be used to identify one of the limitations of the method of the coefficient of variation: in order to ensure, during the dryer season, larger time intervals between independent events, it is possible to overestimate the average duration of the event not physically based. In this perspective, the second method seems to be able to return estimates in line with the expected ones. This consideration is even more interesting for the summer season where the method of the coefficient of variation identified in 185 hours, the average duration of time in the absence of precipitation (t_b) compared with a minimum time between independent rainfall events (t_{b_min}) of 23 hours. Instead, the method using the Pearson correlation coefficient, currency in 101.09 hours t_b in front of a t_{b_min} of 3 hours. The dynamics of such events in the summer seasons, often linked to convective dynamics and then with short but intense phenomena, lead to consider, in this case, more reliable estimate returned by the second method. For the rest, in line with the observations carried out in Section *Phenomenological explanation of the parameters of the Random Parameter Bartlett-Lewis Rectangular Pulse model* , higher cumulative values occur during the autumn and winter that identify the average season, while in the dry season decreases both the number of events and the heap. In Figures 13-15 intensity-duration relationship are displayed . In these figures, two points clearly show a different behavior: the first, shown in blue, is relative to the Cervinara landslide and the second, shown in green, has similar characteristics to the first one but not inducing flowslide phenomena.

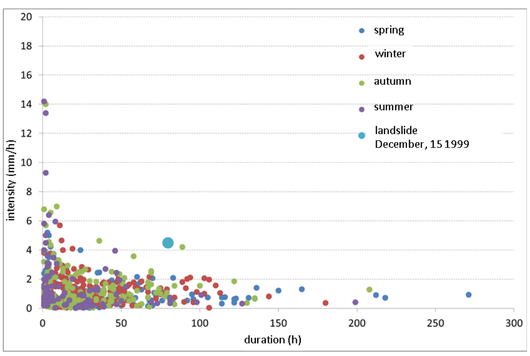

Figure 13:
Relation between intensity and duration of the observed precipitation events series, identified by the method of the coefficient of variation.

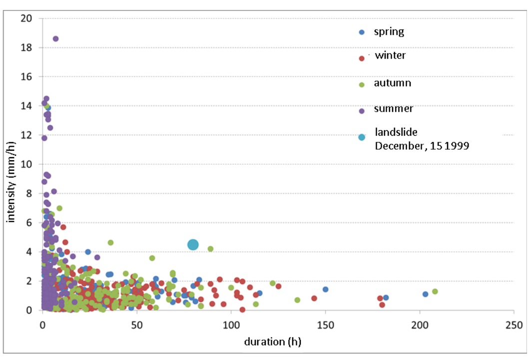

Figure 14:
Relationship between intensity and duration of the observed precipitation events series, identified by the method using the Pearson correlation coefficient.

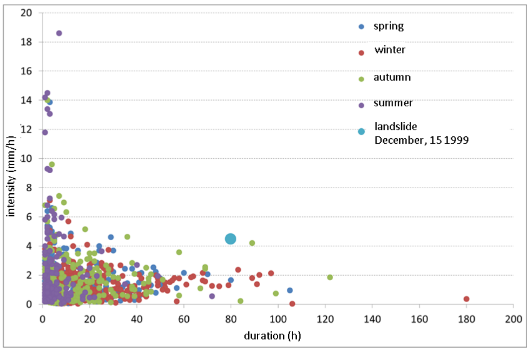

Figure 15:
Relationship between intensity and duration of the events of the observed precipitation series, identified by the method of Huff.

TYPICAL ITERATION CYCLE OF EAS ALGORITHM

A typical iteration cycle of EAS algorithm is given below.

- A random population of boundary points, P, is generated. It is represented by an $m \times n$ matrix, $m > n$, where m specifies is the size of the population and n is the size of the problem (the number of control variables).
- A simplex $S = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_{n+1}\}$ is defined by randomly selecting $n+1$ vertices from the current population P, where $\mathbf{x}_1$ corresponds to the best (smallest) value of the objective function f and $\mathbf{x}_{n+1}$ to the worst (greatest) one.
- The candidate point is the vertex that maximizes the modified function:

  $$g(\mathbf{x}_i) = f(\mathbf{x}_i) + uT \qquad (35)$$

  where u denotes a uniform random number from the interval $[0, 1]$ and T is the actual current temperature of the system.
- A vertex $\mathbf{w}$, the candidate point to be replaced, is selected from $\{\mathbf{x}_2, ..., \mathbf{x}_{n+1}\}$. A new point $\mathbf{r}$ is generated by reflecting the simplex away from $\mathbf{w}$ according to the rule

  $$\mathbf{r} = \mathbf{g} + (0.5 + u)(\mathbf{g} - \mathbf{w}) \qquad (36)$$

  where $\mathbf{g}$ is the centroid of the subset $S - \{\mathbf{w}\}$ and u is a uniform random number.
- If $f(\mathbf{r}) < f(\mathbf{w})$, the new point $\mathbf{r}$ replaces the vertex $\mathbf{w}$. Moreover, if $f(\mathbf{r}) < f(\mathbf{x}_1)$, a sequence of expansions steps is implemented according to the rule

  $$\mathbf{x}_{new} = \mathbf{g} + \varphi^{[s]}(\mathbf{r} - \mathbf{g}) \qquad (37)$$

  where $\varphi^{[s]} = \varphi^{[s-1]} + u$, with $\varphi^{[0]} = 1$. The expansion continues as long as the function value improves, thus accelerating the local searching procedure. If $f(\mathbf{r}) > f(\mathbf{x}_1)$, the simplex is contracted outside as follows:

  $$\mathbf{x}_{new} = \mathbf{g} + (0.25 + 0.5u)(\mathbf{r} - \mathbf{g}) \qquad (38)$$

  If either the expansion or the outside contraction succeeds, $\mathbf{x}_{new}$ replaces $\mathbf{r}$.
- If $g(\mathbf{r}) > g(\mathbf{w})$, the reflection point $\mathbf{r}$ is not accepted, the actual temperature is reduced by a factor λ and the simplex is contracted inside according to the equation

  $$\mathbf{x}_{new} = \mathbf{g} - (0.25 + 0.5u)(\mathbf{g} - \mathbf{w}) \qquad (39)$$

  If $f(\mathbf{x}_{new}) > f(\mathbf{x}_{n+1})$, i.e. the new point is worse than the current worst vertex, the simplex shrinks toward the best vertex $\mathbf{x}_1$, i.e., $\mathbf{x}_i^{'} = 0.5(x_1 + x_i)$.
- If $g(\mathbf{r}) < g(\mathbf{w})$, the reflection point $\mathbf{r}$ is accepted even if it worsens the value of the function. Then a given number of uphill movements are implemented according to (10). If some upward movement succeeds, the simplex escapes from the region of attraction of the current local

minimum and the new point replaces $\mathbf{r}$. Otherwise, a random point is generated on the boundaries of the population P. This new point replaces $\mathbf{r}$ if it improves the function value, otherwise $\mathbf{r}$ is replaced according to a mutation probability p_m. The new point is generated as follows:

$$\mathbf{x}_{new} = c + dy/\left\|y\right\| \tag{40}$$

where $\mathbf{c}$ is the centroid of P, d is the maximum Euclidean distance of the elements of P from the centroid and $\mathbf{y}$ is a random direction in the n-dimensional space.

- The algorithm stops if the relative distance between the current best and worst function values in P, f_{min} and f_{max} respectively, becomes smaller than a given tolerance ϵ. The initial temperature is set equal to $f_{max} - f_{min}$, while it is recomputed at the beginning of each cycle so that it never exceeds $\xi(f_{max} - f_{min})$, where $\xi \geq 1$ is a control parameter of the annealing schedule.

Bibliography

[1] A. Burton, C.G. Kilsby, H.J. Fowler, P.S.P. Cowpertwait, and P.E. O'Connell. RainSim: a spatial-temporal stochastic rainfall modelling system, Environ. *Model. Softw.*, 23:1356–1369, 2008.

[2] P. Cowpertwait, V. Isham, and C. Onof. Point process models of rainfall: developments for fine-scale structure. *P. Roy. Soc. A*, 463:2569–2587, 2007.

[3] Koutsoyiannis D. A stochastic disaggregation method for design storm and flood synthesis. *Journal of Hydrology*, 1994.

[4] Koutsoyiannis D. Rainfall Disaggregation Methods: Theory and Applications. *Workshop on Statistical and Mathematical Methods for Hydrological Analysis*, pages 1–23, 2003.

[5] A. Efstratiadis and D. Koutsoyiannis. An evolutionary annealing-simplex algorithm for global optimisation of water resource systems. *Proceedings of the Fifth International Conference on Hydroinformatics, Cardiff, UK, 1423-1428, International Water Association*, 2002.

[6] Huff F.A. Time distribution of rainfall in heavy storms. *Water Resour. Res.*, 3:1007–1019, 1967.

[7] R.A. Grace and P.S. Eagleson. A model for generating synthetic sequences of short-time-interval rainfall depths. *Proceedings of International Hydrology Symposium, Fort Collins, Colorado*, pages 268–276, 1967.

[8] J.C. Grygier and J.R. Stedinger. Condensed disaggregation procedures and conservation corrections for stochastic hydrology. *Water Resour. Res.*, 1988.

[9] J.C. Grygier and J.R. Stedinger. SPIGOT, A synthetic streamflow generation software package, Technical description. *School of Civil and Environmental Engineering, Cornell University, Ithaca, NY., Version 2.5*, 1990.

[10] S. Islam, D. Entekhabi, and R.L. Bras. Parameter-estimation and sensitivity analysis for the modified Bartlett-Lewis rectangular pulses model of rainfall. *J. Geophys. Res.*, 95:2093–2100, 1990.

[11] D. Koutsoyiannis and A. Manetas. Simple disaggregation by accurate adjusting procedures. *Water Resour. Res.*, 1996.

[12] D. Koutsoyiannis and C. Onof. A computer program for temporal rainfall disaggregation using adjusting procedures. *XXV General Assembly of European Geophysical Society, Nice, Geophys. Res. Abstracts*, 2, 2000.

[13] D. Koutsoyiannis and C. Onof. Rainfall disaggregation using adjusting procedures on a Poisson cluster model. *J. of Hydrol.*, 2001.

[14] W.L. Lane and D.K. Frevert. Applied Stochastic Techniques, User's Manual. *Bureau of Reclamation, Engineering and Research Center Bureau of Reclamation, Engineering and Research Center, Denver, Co., Personal Computer Version*, 1990.

[15] B.B. Mandelbrot and J.R. Wallis. Noah, Joseph, and operational hydrology. *Water Resour. Res.*, pages 909–918, 1968.

[16] J.A. Mead and R. Nelder. A simplex method for function minimization. *Computer Journal 7*, 1965.

[17] C. Onof and H.S. Wheater. Modeling of British rainfall using a Random parameter Bartlett-Lewis rectangular pulse model. *J. Hydrol.*, 149:67–95, 1993.

[18] C. Onof and H.S. Wheater. Improvements to the modeling of British rainfall using a modified random parameter Bartlett-Lewis rectangular pulses model. *J. Hydrol.*, 157:177–195, 1994.

[19] P.J. Restrepo-Posado and P.S. Eagleson. Identification of independent rainstorms. *J. Hydrol.*, 55:303–319, 1982.

[20] I. Rodriguez-Iturbe, D.R. Cox, and V. Isham. Some models for rainfall based on stochastic point processes. *Proc Royal Soc London A*, 1987.

[21] I. Rodriguez-Iturbe, D.R. Cox, and V. Isham. A point process model for rainfall: Further developments. *Proc Royal Soc London A*, 1988.

[22] J.R. Stedinger and R.M. Vogel. Disaggregation procedures for generating serially correlated flow vectors. *Water Resour. Res.*, 1984.

[23] W.J. Vanhaute, S. Vandenberghe, K. Scheerlinck, B. De Baets, and N.E.C. Verhoest. Calibration of the modified Bartlett Lewis model using global optimization techniques and alternative objective functions. *Hydrol. Earth Syst. Sci.*, 16:873–891, 2012.

[24] N. Verhoest, P.A. Troch, and F.P. De Troch. On the applicability of Bartlett-Lewis rectangular pulses models in the modeling of design storms at a point. *Journal of Hydrology*, pages 109–120, 1997.