RAINFALL DISAGGREGATION METHODS: THEORY AND APPLICATIONS

Demetris Koutsoyiannis

Department of Water Resources, School of Civil Engineering, National Technical University of Athens

Abstract

A large variety of disaggregation methods that have appeared in hydrological literature and used in hydrological applications are reviewed with emphasis in rainfall modelling. The general-purpose stochastic disaggregation models, which have been used at several applications including rainfall modelling but at time scales not finer than monthly, are summarised. The specialised models for rainfall disaggregation, in particular at fine time scales, are examined in more detail. A special disaggregation technique, which, instead of using simultaneously both coarser and finer time scales in one mathematical expression, couples two independent stochastic models, one at each time scale, is further analysed. Two examples of implementing this technique to fine scale rainfall disaggregation are given. In the first case the implementation results in a single variate rainfall disaggregation model (Hyetos) based on the Bartlett-Lewis process. In the second case it results in a multivariate rainfall disaggregation model (MuDRain). These two implementations are demonstrated with results from real world applications.

1. Introduction

Very often a hydrological stochastic process must be studied in different time scales. Therefore, the problem arises of how to generate consistent time series both in a coarser, or higher-level, time scale and a finer, or lower-level, time scale. A trivial solution of this problem is to model the process in the lower-level time scale only, and then aggregate to derive the process in the higher-level time scale. However, there are reasons to avoid this solution and model the process in different time scales separately, each time focusing on different important statistical properties of the process. For instance, at the annual scale a model focuses on the long-term persistence properties of the process; at the monthly scale another type of model should describe periodicity and short-term memory of the process; and at the hourly time scale a model should describe intermittency (e.g. wet versus dry state) and fine-scale structure of the process.

In other cases, the higher-level process may be the output of a specialised model (e.g., a meteorological or climatological model) or known from measurements (e.g., daily rainfall measurements); apparently in such cases the aggregation approach cannot work, but rather disaggregation is needed. Specifically when dealing with rainfall, there is a large number of daily raingauges (pluviometres), which have often been operational for a few decades. However, the number of raingauges providing hourly or sub-hourly resolution data (pluviographs, rainfall sensors) is smaller by about an order of magnitude. This situation reflects a general relative paucity of rainfall data for time-scales of one hour or less, both in number of gauges and length of the recorded series. The need for hourly data for hydrological applications, especially in flood studies, suggests the use of appropriate techniques to refine the available daily information and provide the user with possible realisations of hourly precipitation which aggregate up to the given daily data. Such techniques would provide a continuous simulation tool for use for design and management of hydrosystems.

This kind of problems is commonly tackled by disaggregation models. Several such models have been developed since the 1970s and utilised in numerous hydrological applications, including, among others, simulation of reservoir systems, either for design or operation purposes, storm and flood simulations, and even enhancement of hydrological data records. Some of the disaggregation models developed are general-purpose as they are not specific to a certain hydrological process or a certain application. General-purpose models, however, are not always applicable to the rainfall process, especially at fine time scales, due to certain peculiarities it exhibits. This triggered the development of specialised techniques for the rainfall process, whose study consisted a preferential field of disaggregation models.

Today, there is renewed interest in disaggregation methods as climate related issues (e.g. downscaling of climate change scenarios) have attracted the interest of many researchers. Usually, such scenarios are developed at a coarse time scale and there is a need to transform them into a finer scale. However, disaggregation is not identical to downscaling, as the latter aims at producing finer scale time series with the required statistics but that do not necessarily add up to any given coarse scale totals. Downscaling is in particular used for hydrological applications of general circulation models (GCM) output where the exact values of the large-scale GCM totals are not considered particularly reliable. In both problem types, synthetic fine-scale series should reproduce the important statistical features of the related hydrological processes at this time scale.

This paper reviews a large variety of disaggregation models that have appeared in hydrological literature and used in hydrological applications. The general-purpose stochastic disaggregation models, which have been used at several applications including rainfall modelling but at time scales not finer than monthly, are summarised in section 2. The specialised models for rainfall disaggregation, in particular at fine time scales, are examined in section 3. A special disaggregation technique, which, instead of using simultaneously both coarser and finer time scales in one mathematical expression, couples two independent stochastic models, one at each time scale, is studied in section 4. Two examples of implementing this technique to fine scale rainfall disaggregation are given in sections 5 and 6. In the first case the implementation results in a single variate rainfall disaggregation model (Hyetos) based on the Bartlett-Lewis process. In the second case it results in a multivariate rainfall disaggregation model (MuDRain). The conclusions and discussion, including future challenges of rainfall disaggregation models are given in section 7.

2. General-purpose stochastic disaggregation models

Disaggregation models were introduced in hydrology by the pioneering work of Valencia and Schaake (1972, 1973). Their multivariate model generates sequences of lower-level values at many locations as linear combinations of the related higher-level values and independently generated random components. Let $\mathbf{Z}_i := [\overset{1}{Z}_i, ..., \overset{n}{Z}_i]^T$ denote the vector of some hydrological quantities (e.g. rainfall, runoff, etc.), represented as discrete time stochastic processes, at n locations and at discrete higher-level time interval (e.g. year) i. Let also $\mathbf{X}_s := [\overset{1}{X}_s, ..., \overset{n}{X}_s]^T$ represent the same quantities but at the lower-level time interval (e.g. month) s. The superscript T denotes the transpose of a vector or matrix and time indexes i and s have common origin, i.e., at the time origin $i = 0$ and $s = 0$. Generally, in this paper we use upper case letters for random variables, and lower case letters for values, parameters, or constants. Furthermore, we use bold letters for arrays or vectors, and regular letters for their elements.

Higher- and lower-level time intervals will be referred to as periods and subperiods, respectively. Consistency of the processes at the two time scales requires that

$$\sum_{s = (i-1)k+1}^{i\,k} \mathbf{X}_s = \mathbf{Z}_i \tag{1}$$

where k is the number of subperiods within each period (e.g., for annual and monthly higher- and lower-level time scales, $k = 12$). It is assumed that the annual series $\mathbf{Z}_i$ are known (e.g. are given from measurement or are generated first). Then, according to Valencia and Schaake (1972, 1973), the series $\mathbf{X}_s$ can be generated from

$$\mathbf{X}_i^* = \mathbf{a}\,\mathbf{Z}_i + \mathbf{b}\,\mathbf{V}_i \tag{2}$$

where $\mathbf{X}_i^*$ is the vector containing all kn lower-level quantities of period i, i.e.

$$\mathbf{X}_i^* := [\mathbf{X}_{(i-1)k+1}^T, \ldots, \mathbf{X}_{i\,k}^T]^T \tag{3}$$

whereas $\mathbf{V}_i$ is a vector of kn independent identically distributed random variates and $\mathbf{a}$ and $\mathbf{b}$ are matrices of parameters with sizes $kn \times n$ and $kn \times kn$, respectively. These parameters depend on variance and covariance properties among higher- and lower-level variables, which are estimated from historical records. The parameter estimation procedure of this model ensures the resemblance of variance and covariance properties between historical and generated series. However the model makes no effort to preserve covariances of the lower-level variables belonging to consecutive periods. In fact, this model does not assume any connection between lower-level variables of different periods.

Different model structures and parameter estimation procedures intended to preserve the lagged covariance properties among lower-level variables belonging to consecutive periods have been suggested by Mejia and Rousselle (1976), Hoshi and Burges (1979) and Stedinger and Vogel (1984). However, as showed by Stedinger and Vogel (1984), the reproduction of the exact historical serial correlations of the lower-level variables is an impossible task within the framework of these models because of structural constraints imposed by that framework, thus giving rise to an inconsistency.

In cases where all variables are normally distributed, the basic linear generating scheme of the Valencia-Schaake model ensures the preservation of higher order moments and distribution functions as well; otherwise the model in its primary formulation fails to maintain such properties. Suggested modifications of the model, permitting the representation of non-normal distributions, can be classified into two categories. The methods of the first category are oriented towards the preservation of merely the skewness of the lower-level variables (Tao and Delleur, 1976; Todini, 1980). This is done by assuming that the auxiliary variables $\mathbf{V}_i$ in (2) have skewed distributions with appropriate coefficients of skewness determined in terms of the third order marginal and joint moments of historical data. However, as Todini (1980) notes, in practice it may be impossible to preserve the skewness of $\mathbf{X}_i$ because the required skewness coefficients of the auxiliary variables $\mathbf{V}_i$ may be arbitrarily high and impossible to be implemented in typical series with finite lengths. The methods of the second category utilise nonlinear transformations of the variables, so that the transformed variables have normal distributions (Valencia and Schaake, 1972; Hoshi and Burges, 1979; Stedinger and Vogel, 1984). However, as Todini (1980) notes, this means that the additive property, which is one of the main attributes of the original disaggregation scheme, is lost. To overcome this problem an empirical correction procedure, known as adjustment procedure, was suggested (Lane and Frevert, 1990; Stedinger and Vogel, 1984; Grygier and Stedinger, 1988, 1990). However, unless carefully studied (see section 4), such a procedure introduces bias in all statistics that are to be preserved.

Another drawback of the above summarised disaggregation models is their excessive number of parameters, because of the large number of cross correlations that they attempt to reproduce. For example in a problem of annual to monthly disaggregation (k = 12) involving n = 10 locations the number of parameters in model (2) is 15 600, even higher than the number of data values that are typically available to estimate them (e.g. for data records extending over 30 years the number of monthly data values are 3 600). The situation is even worse in later disaggregation models. To tackle this problem, several different procedures reducing the required number of parameters have been developed. The staged disaggregation models (SDM; Lane, 1979, 1982; Salas et al., 1980; Stedinger and Vogel, 1984; Grygier and Stedinger, 1988; Lane and Frevert, 1990) disaggregate higher-level variables at one or more sites to lower-level variables at those and other sites in two or more steps. The condensed disaggregation models (CDM; Lane, 1979, 1982; Pereira et al., 1984; Oliveira et al., 1988; Stedinger and Vogel, 1984; Stedinger et al., 1985; Grygier and Stedinger, 1988) reduce the number of required parameters by explicitly modelling fewer of the correlations among the lower-level variables. The dynamic disaggregation model (DDM) initially developed as a single site model (Koutsoyiannis, 1988; Koutsoyiannis and Xanthopoulos, 1990) and subsequently as a multivariate model (Koutsoyiannis, 1992) has been formulated as a generalised stepwise approach to disaggregation problems, allowing for a variety of configurations. DDM is in close connection with an associated particular sequential generating model (e.g. a seasonal autoregressive model like a PAR(1) model) and uses the same parameter set as in this sequential model. With the DDM approach lower-level variables are generated one at a time, given the total amount to be allocated across the remaining subperiods. Thus the disaggregation of a higher-level variable into its components (lower-level variables) is split in equivalent sequential steps, each corresponding to a specific subperiod. According to this method, in each step, two separate procedures are executed. First the parameters of the generation equations for this step are determined so as to preserve the specified marginal and conditional moments given the total amount remaining to be allocated and the already generated values at previous steps. Second, these parameters are used to generate the lower-level amount of the current step and to update the remaining amount to be allocated in next steps. DDM has the ability of representing non-Gaussian distributions and one of its configurations (Koutsoyiannis, 1992) is capable of preserving lagged covariance properties among lower-level variables related to consecutive periods, thus fixing the inconsistency problem of earlier disaggregation models mentioned before.

Among the more recent developments in disaggregation models we must mention the study by Kumar et al. (2000) who approximated the probability distribution of the vector of disaggregated flows nonparametrically using the k nearest neighbours of the monthly spatial flow pattern. Most of the problems involved with disaggregation models have been tackled by a different yet simpler framework by Koutsoyiannis and Manetas (1996) and Koutsoyiannis (2001); this will be discussed in more detail in section 4.

3. Rainfall disaggregation models

Most of the above mentioned disaggregation models have been used to disaggregate annual rainfall to monthly amounts. However, as first pointed out by Valencia and Schaake (1972), most modeling schemes of this kind are not suitable for the disaggregation of rainfall for time scales finer than monthly, due to the skewed distributions and the intermittent nature of the rainfall process at fine time scales. Other disaggregation models, different from those described above, have been proposed and used particularly for the disaggregation of rainfall. These, however, do not exhibit the generality of the Valencia-Schaake type linear schemes and in many cases are ad-hoc techniques rather than consistent generalised methods.

Grace and Eagleson (1966) proposed an urn model for the disaggregation of a storm's depth into shorter durations. Schaake et al. (1972) developed a Markov chain model for the disaggregation of monthly rainfall into daily. Woolhiser and Osborn (1985) presented a scheme for the disaggregation of an individual storm's depth into fractional depths, each corresponding to one tenth of the storm's duration; their scheme was based on a nondimensionalised Markov process, resulting from successive transformations of the real rainfall process. The issue of disaggregating a daily total into individual storm amounts in a day was studied by Hershenhorn and Woolhiser (1987). The problem of the simulation of the internal time distribution of a storm was studied also in the work of Marien and Vandewiele (1986), where the developed disaggregation scheme applies to properly defined fractional variables with gamma distributions. Similar techniques using nondimensionalised time and depth distributions of rainfall were also applied by Econopouly et al. (1990) and Garcia-Guzman and Aranda-Oliver (1993).

A more general mathematical method for disaggregating rainfall for time scales finer than monthly was proposed by Koutsoyiannis (1988) and Koutsoyiannis and Xanthopoulos (1990). Specifically, the DDM approach they developed (already mentioned above) can perform with an arbitrary and varying number of subperiods k (within one period), which facilitates the use of the model for the intermittent rainfall process. DDM disaggregates a total amount into k partial amounts in $k - 1$ steps, using a stepwise procedure. To disaggregate monthly to hourly rainfall, this procedure can be applied four times successively to (a) locate the starting points of storm events within a month; (b) disaggregate the monthly rainfall duration into event durations; (c) disaggregate the monthly rainfall into depths of individual events; and (d) to disaggregate the total depth of each individual event into shorter period (i.e., hourly) depths. This model could however not straightforwardly be applied to the case of the daily to hourly disaggregation.

The issue of the disaggregation of a specific storm event to finer time-scales later was addressed in a mathematically simpler yet accurate manner by Koutsoyiannis (1994). This was based on the assumption that the marginal distribution function of the incremental rainfall depth at a short time scale (e.g. hourly) is (approximately) gamma, and the mathematical proof that rescaling of incremental depths so as to match a given total amount does not alter the gamma distribution function (see also section 4 below). Although the aim of the latter technique was not continuous rainfall simulation but rather event based simulation (Zarris et al., 1998), Koutsoyiannis and Pachakis (1996; Appendix) managed to construct a continuous simulation tool based on this technique along with the assumption of an alternating renewal process for rainfall occurrences and a scaling hypothesis (Koutsoyiannis and Foufoula-Georgiou, 1993) for the intensity process.

This latter technique, however, did not consider explicitly daily rainfall amounts. This was done subsequently by Glasbey et al. (1995). Using a random parameter Bartlett-Lewis rectangular pulse model proposed by Rodriguez-Iturbe et al. (1987), the authors examine a method based upon simulating data until a good match of daily totals is obtained for the duration of a given event in the daily data. Rescaling is then required to reproduce daily totals exactly. This method leads to inflated hourly intensity variances. Another more ad hoc model consists in trying to reproduce the sequence of three daily rainfall totals with the simulation and then adjusting the data. This was more satisfactory, with a good reproduction of the main hourly statistics. However, the extreme values were not investigated and the decrease of the autocorrelations with the lag was somewhat faster than in the historical sequence. As the authors point out moreover, this method is not theoretically justified since three-day periods are not independent.

Among subsequent attempts for stochastic rainfall disaggregation at a fine scale we must mention the works by Cowpertwait et al. (1996; disaggregation of hourly rainfall into smaller time intervals by allocating pulses of a specified small depth each, at the different intervals); Bardossy (1997; disaggregation of daily precipitation into a number of wet subperiods conditioned on the total daily amount using the Polya distribution and a Markov-chain autocorrelation structure maintained by a Monte Carlo technique); and Connolly et al. (1998; disaggregation of daily rainfall into a number of events following a Poisson process). A more systematic model, based on the Bartlett-Lewis rectangular pulse process, was studied by *Koutsoyiannis and Onof* (2000, 2001). This will be discussed in more detail in section 5.

The above-mentioned models are generally based on classical probability and stochastic processes theory. In the last years there have been studied disaggregation techniques with different scientific bases. Thus, the development of multifractal simulation techniques has provided a potentially powerful tool for the exploration of problems such as disaggregation. An application of this approach to the disaggregation problem was proposed by Olsson (1996) and Olsson and Berndtsson (1997). The use of a self-similar microcanonical cascade enables the reproduction of the exact total daily rainfall, but it does not allow for the reproduction of the observed hourly autocorrelations (Tang, 1999). Bounded microcanonical cascades (Marshak et al., 1994) do however provide a tool, which could be used for disaggregation. Such approaches are promising, as illustrated by the successful reproduction of rainfall statistics with canonical bounded cascades (Menabde et al., 1997), but require more analysis, particularly in their ability to reproduce the dry period structure at different scales. A different approach has been followed by Burian et al., (2000, 2001) who used artificial neural networks to disaggregate hourly rainfall data into shorter time intervals. Yet another approach was followed by Sivakumar et al. (2001) who formed a simple chaotic model to disaggregate rainfall of five resolutions using techniques from the chaos literature. It must be mentioned however that the latter techniques that are not based on probability and stochastic processes may not be appropriate for large length simulations, as they do not perform well in extrapolation. For example they may result in poor description of extremes, whose study needs large length simulations.

All fine time scale rainfall disaggregation techniques summarised above have a common characteristic: they are single-site. The problem of multiple site rainfall disaggregation, as a means for simultaneous spatial and temporal disaggregation, is of significant practical interest but presents significant differences from that of single-site disaggregation, including increased mathematical complexity. The spatial correlation (cross-correlation among different sites) must be maintained in the multivariate problem, whereas it does not appear at all in univariate problems. A first attempt to incorporate more than one site into rainfall disaggregation was done by Socolofsky and Adams (2001) who disaggregated daily rainfall to hourly increments simulating the hourly rainfall trace for each storm from selected intensity patterns measured at a nearby station. More recently, Kottegoda et al. (2003) simulated daily rainfall through a two-station model at a key station and a satellite station whilst maintaining the first three moments and relevant correlations. Based on this two-station model they were able to perform disaggregation of daily rainfalls into hourly values through dimensionless accumulated hourly amounts generated by a beta distribution, also postulating that the occurrence process of hourly rainfall has a geometric distribution conditioned on the total daily rainfall. Application was made to the Tiber river basin in central Italy. A true multidimensional approach was developed by Koutsoyiannis et al. (2001, 2003) and applied initially to the Brue catchment in South-Western England and subsequently (Fytilas, 2002) to the Tiber river basin in central Italy. This will be presented in more detail in section 6.

4. A generalised stochastic framework for coupling stochastic models of different time scales

Most of the above-described models are purposely-designed to generate a process in the lower-level time scale given that in the higher-level. Specifically, they do not model the process of interest in the lower-level time scale itself, but rather they are hybrid schemes using simultaneously both time scales. Sometimes (owing to nonlinear transformations of variables, already mentioned in section 2) these models are not able to ensure consistency with the higher-level process. Then, adjusting procedures are necessary to restore consistency (Grygier and Stedinger, 1988, 1990; Lane and Frevert, 1990).

However, there is the possibility of not designing and implementing a special model for disaggregation as a hybrid scheme incorporating both time scales. On the contrary, there may be available a model of the lower-level time scale with no reference to the higher-level time scale. The problem is then how a time series generated by the lower-level model can be modified so as to be consistent with a given higher-level time series, without affecting the stochastic structure implied by the lower-level model. Practically, this is equivalent to the use of adjusting procedures mentioned before.

In a recent study, Koutsoyiannis and Manetas (1996) showed that this is possible without using any kind of disaggregation model but only using adjusting procedures on top of the separate lower-level model. Their adjusting procedures are accurate in the sense that they do not modify certain statistics of the lower-level process. In that study, a seasonal autoregressive (PAR(1)) model was used as the lower-level model.

Provided that a single-variate data series Z_i (i = 1, 2, …) is known at a higher-level time scale and a lower-level synthetic series $\tilde{X}_s$ (s = 1, 2, …) has been generated by some stochastic model, the problem is how to modify $\tilde{X}_s$ to obtain another series X_s, consistent with the higher-level one. In other words, the problem is to allocate the error in the additive property (1), i.e., the departure of the sum of lower-level variables within a period from the corresponding higher-level variable. A good adjusting procedure, in addition to reinstating the additive property, should preserve explicitly (at least under some specified conditions) certain statistics or even the complete distribution of lower-level variables.

A first example of such an adjusting procedure is the so-called proportional adjusting procedure which modifies the initially generated values $\tilde{X}_s$ to get the adjusted values X_s according to

$$X_s = \tilde{X}_s \left(Z / \sum_{j=1}^{k} \tilde{X}_j \right) \qquad (s = 1, \dots, k) \tag{4}$$

where Z is the higher-level variable (for notational simplicity we assumed i = 1 and also eliminated the subscript i from Z). As shown earlier by Koutsoyiannis (1994), the proportional adjusting procedure is exact for complete preservation of distributions if variables X_s are independent with two-parameter gamma distribution and common scale parameter.

Another example is the linear adjusting procedure, which is described by

$$X_s = \widetilde{X}_s + \lambda_s \left(Z - \sum_{j=1}^{k} \widetilde{X}_j \right) \qquad (s = 1, \dots, k) \tag{5}$$

where λ_s are unique coefficients depending on the covariances of X_s with Z. As shown by Koutsoyiannis and Manetas (1996), it is exact for the complete preservation of normally distributed, dependent or independent, variables. In addition, it is exact in preserving second order moments of (dependent or independent) variables with any distribution.

In addition, the methodology proposed by Koutsoyiannis and Manetas (1996) uses repetitive sampling in order to improve the approximations of statistics that are not explicitly preserved by the adjusting procedures (e.g. skewness). That is, the lower-level model is run not only once but several times and the generated sequence that results in closer agreement with the already known value of the higher-level variable Z is finally chosen.

This approach was more recently extended by Koutsoyiannis (2001), so as to form a generalised framework for coupling stochastic models of different time scales. This framework couples two independent stochastic models appropriate respectively for the coarser and finer time scales using a transformation that modifies the output of the latter to become consistent with the series produced by the former model (or the series known from measurements at the coarser scale). The finer-scale model is run independently of the coarser-scale one, without any reference to the known (generated first or measured) $\mathbf{Z}_i$, and produces lower-level series $\widetilde{\mathbf{X}}_s$. If we aggregate the latter at the coarser scale (by means of (1)) we will obtain some series $\widetilde{\mathbf{Z}}_i$, which will apparently differ from $\mathbf{Z}_i$. In a subsequent step, we modify $\widetilde{\mathbf{X}}_s$ thus producing $\mathbf{X}_s$ consistent with $\mathbf{Z}_i$ (in the sense that they obey (1)) without affecting the stochastic structure that characterises $\widetilde{\mathbf{X}}_s$. For this modification, a linear transformation $\mathbf{X}_s = \mathbf{f}(\widetilde{\mathbf{X}}_s, \widetilde{\mathbf{Z}}_i, \mathbf{Z}_i)$, termed the coupling transformation, is used. As shown by Koutsoyiannis (2001), this is given by an extension of (5), i.e.,

$$\mathbf{X}_i^* = \widetilde{\mathbf{X}}_i^* + \mathbf{h}\,(\mathbf{Z}_i^* - \widetilde{\mathbf{Z}}_i^*) \tag{6}$$

where $\mathbf{X}_i^*$ is defined in (3) and

$$\mathbf{Z}_i^* := [\mathbf{Z}_i^T, \mathbf{Z}_{i+1}^T, \mathbf{X}_{(i-1)k}^T]^T \tag{7}$$

$$\mathbf{h} = \mathrm{Cov}[\mathbf{X}_i^*, \mathbf{Z}_i^*]\ \{\mathrm{Cov}[\mathbf{Z}_i^*, \mathbf{Z}_i^*]\}^{-1} \tag{8}$$

whereas $\widetilde{\mathbf{X}}_i^*$ and $\widetilde{\mathbf{Z}}_i^*$ are defined in terms of $\widetilde{\mathbf{X}}_s$ and $\widetilde{\mathbf{Z}}_i$ in a manner identical to that of the definition of $\mathbf{X}_i^*$ and $\mathbf{Z}_i^*$.

It is clarified that the vector $\mathbf{X}_i^*$ contains the lower-level values of all k subperiods of period i for all examined locations and the vector $\mathbf{Z}_i^*$ contains (a) the higher-level values of the current period; (b) the higher-level values of the next period; and (c) the lower-level values of the last subperiod of the previous period. Items (b) and (c) of $\mathbf{Z}_i^*$ are included to ensure that the transformation will preserve not only the covariance properties among the lower-level values of each period, but the covariances with the previous and next periods as well (that is, to remedy the inconsistency of earlier disaggregation models mentioned in section 2). Note that at the stage of the generation at period i the lower-level values of period $i - 1$ are known (therefore, in $\mathbf{Z}_i^*$ we enter lower-level values of period $i - 1$) but the lower-level values of period i + 1 are not known (therefore, in $\mathbf{Z}_i^*$ we enter higher-level values of period i + 1, which are known).

It is also clarified that the covariance terms that are involved in the determination of the matrix **h** in (8) are not estimated directly from historical data but from mathematical equations that are associated to merely the lower-level model. In this way, the coupling transformation does not introduce any independent parameter, apart from those used in the lower-level model, thus keeping the number of independent parameters to a minimum and avoiding the problem of inflation of parameters that was discussed in section 2. Equations for estimating **h** for typical stochastic hydrological lower-level models (PAR(1), PAR(2), PARMA(1, 1) and nonlinear transformations thereof) are given by Koutsoyiannis (2001).

The quantity $\mathbf{h}\,(\mathbf{Z}_i^* - \widetilde{\mathbf{Z}}_i^*)$ in (6) represents the correction applied to $\widetilde{\mathbf{X}}$ to obtain $\mathbf{X}$. Whatever the value of this correction is, the coupling transformation will ensure preservation of first and second order properties of variables (means and variance-covariance matrix) and linear relationships among them (in our case the additive property (1)). However, it is desirable to have this correction as small as possible in order for the transformation not to affect seriously other properties of the simulated processes (e.g. the skewness). It is possible to make the correction small enough, if we keep repeating the generation process for the variables of each period (rather than performing a single generation only) until a measure of the correction becomes lower that an accepted limit. This measure can be defined as

$$\Delta = (1\,/\,m)\,\|\,\mathbf{Z}_i^{\prime *} - \widetilde{\mathbf{Z}}_i^{\prime *}\| \tag{9}$$

where $\mathbf{Z}_i^{\prime *}$ and $\widetilde{\mathbf{Z}}_i^{\prime *}$ are respectively $\mathbf{Z}_i^*$ and $\widetilde{\mathbf{Z}}_i^*$ standardised by standard deviation (i.e. $Z_i^{\prime *l} := Z_i^{*l} / \{\mathrm{Var}[Z_i^{*l}]\}^{1/2}$), m is the common size of $\mathbf{Z}_i^{\prime *}$ and $\widetilde{\mathbf{Z}}_i^{\prime *}$, and $\|.\|$ denotes the Euclidian norm.

5. Hyetos: A single variate fine time scale rainfall disaggregation model based on the Bartlett-Lewis process

The general framework presented in section 4 can be implemented directly in the single-site rainfall disaggregation at any time scale. A successful rainfall generation model capable of reproducing the particular fine scale structure of rainfall can be used as the lower-level model. This can then be combined with an appropriate procedure for adjusting the lower-level amounts so as to obtain the required higher-level totals.

Such an implementation was done by Koutsoyiannis and Onof (2000, 2001) and resulted in a computer program named Hyetos that can be easily applied at any location provided that a minimal amount of data is available that can support parameter estimation. The higher- and lower-level time scales in this implementation are respectively daily and hourly which are the most suitable for typical hydrological applications. As an appropriate rainfall model, the Bartlett-Lewis point process in any of its common configurations (constant or random parameter, exponentially or gamma distributed intensities) was chosen due to its wide applicability and experience in calibrating and applying it to several climates. Accumulated evidence on its ability to reproduce important features of the rainfall field from the hourly to the daily scale and above can be found in the literature (Rodriguez-Iturbe et al., 1987, 1988; Onof and Wheater, 1993, 1994). This type of model has the important feature of representing rainfall in continuous time. It is therefore particularly useful in a disaggregation framework where it may be used at a time-step different from that at which it is fitted. Another important feature justifying the choice of this model is the fact that the wet/dry structure can be generated independently of the intensity profile as well as the independence of successive storms; these reduce the number of computations required. It is recalled that the Bartlett-Lewis rectangular pulse model assumes that rainfall occurs in the form of storms of certain durations and each storm is a cluster of random cells, each having constant intensity during the time period it lasts. The general assumptions of the model are:

(1) Storm origins occur following a Poisson process.
(2) Origins of cells of each storm arrive following a Poisson process.
(3) Arrivals of each storm terminate after a time exponentially distributed.
(4) Each cell has a duration exponentially distributed.
(5) Each cell has a uniform intensity with a specified (exponential or gamma) distribution.

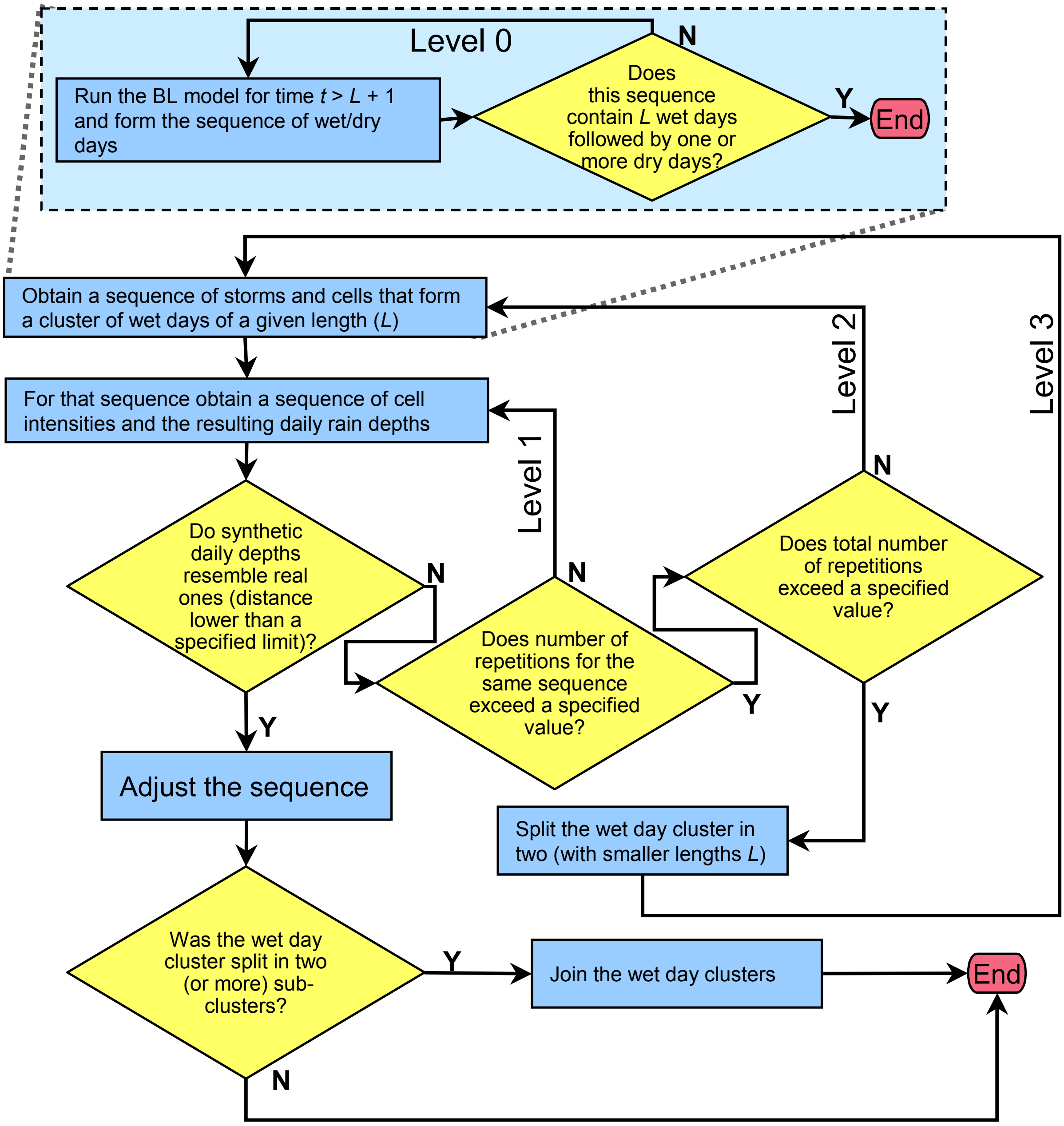

Figure 1 Flow diagram of the repetition scheme used in Hyetos (Koutsoyiannis and Onof, 2001).

Finally, as an appropriate adjusting procedure, the proportional adjusting procedure (equation (4)) was chosen due to its ability not to affect the dry periods (zero values map to themselves) and the fact that it does not generate negative values. In addition, the empirical observation that fine scale rainfall depths are approximately gamma distributed supports the selection of this particular adjusting procedure.

The adjusting procedure is combined with an appropriate repetition scheme. Details of the repetition and disaggregation scheme are shown in Figure 1, with reference to the disaggregation of daily rainfall depths of a cluster of L wet days (preceded and followed by at least one dry day). The scheme was assembled so as to optimise computer time and incorporates four levels of repetition. Initially (Level 0), the Bartlett-Lewis model runs several times until a sequence of exactly L wet days is generated. Then (Level 1), the intensities of all cells and storms are generated and the resulting daily depths are calculated. These are compared to the original ones by means of a cumulative logarithmic distance between original and generated daily depths of all days of the cluster. If the distance d is greater than an accepted limit d_a, then we re-generate the intensities of cells (Level 1 repetitions) without modifying the time locations of storms and their cells. If, however, after a large number of Level 1 repetitions, the distance remains higher than the accepted limit, this may mean that the arrangement of storms and cells is not consistent with the original (and unknown) one. In this case we discard this arrangement and generate a new one, thus entering Level 2 repetitions. Furthermore, in the case of a very long sequence of wet days it is practically impossible to get a sequence of wet days with a departure of the daily sum from the given daily rainfall smaller than the specified limit. In these cases the sequence is subdivided into sub-sequences (in a random manner), each treated independently from the others (Level 3 repetitions). The algorithm allows nested subdivisions. Eventually, the sequence with distance smaller than the accepted limit is chosen and further processed by determining the lower-level (hourly) rainfall depths through the application of the proportional adjusting procedure.

To validate the model, Koutsoyiannis and Onof (2001) performed two test case studies for two raingauges with extremely different climatic conditions: the Heathrow airport raingauge (London, UK) located in a wet region and the Walnut Gulch (Arizona, USA) Gauge 13 located in a semiarid region. Some of their results in the two test cases are reproduced here in Figures 2-6 in adapted graphical form. In each of Figures 2-5 results of four cases are plotted, namely (1) historical data; (2) simulated data using the Bartlett-Lewis rectangular pulse model with length equal to that of the historical record and parameters estimated from the historical data using the generalised method of moments; (3) data produced by disaggregating the historical data series 1; and (4) data produced by disaggregating the simulated data series 2. The inclusion of results from all four data series allows distinguishing the performance of the Bartlett-Lewis model, the disaggregation model, and the combination of the two, in preserving several characteristics of the historical data series. In cases (3) and (4) the Hyetos model was applied with maximum allowed distance d = 0.1 and maximum number of repetitions (total for all levels) 5000.

Figures 2, 3, 4 and 5 depict the good model performance in reproducing respectively the proportions of dry hours and dry days, the coefficients of variation (standard deviation divided by mean value) and skewness of the hourly rainfall intensities, the autocorrelation coefficients of the hourly rainfall intensities for lags up to 10, and the hourly maxima. In addition, Figure 6 concentrates the most important comparisons of Figures 2-4 for one of the studied cases (Walnut Gulch Gauge 13, month of July) also providing information for timescales greater than one hour. Specifically, it compares the coefficient of variation and skewness, probability of dry intervals and lag-1 autocorrelation coefficient of historical and synthetic data at timescales (aggregation levels) 1 to 24 hours. The agreement of historical and synthetic statistics is impressively good at all timescales.

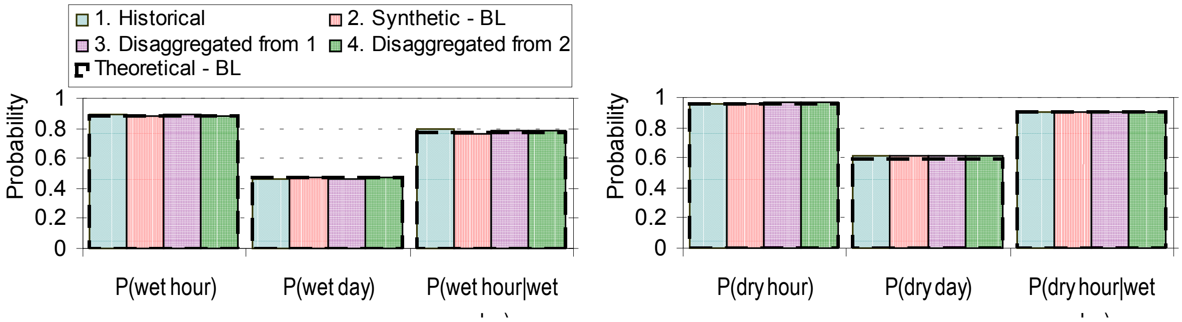

Figure 2 Comparison of dry/wet probabilities of the historical and synthetic data records for the case studies of Heathrow airport, month of January (left) and Walnut Gulch Gauge 13, month of July (right) (adapted from Koutsoyiannis and Onof, 2001).

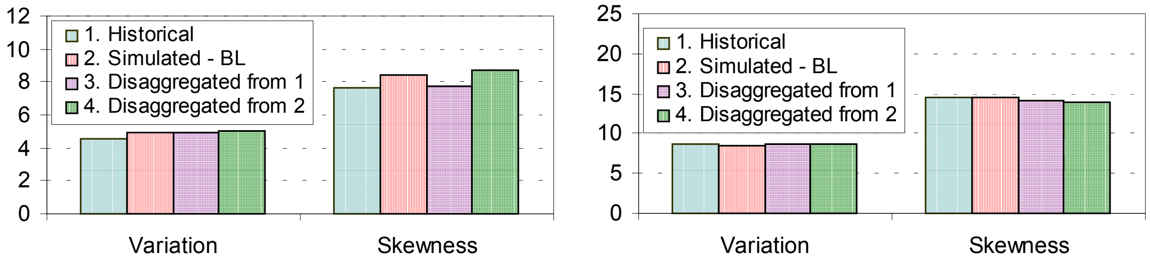

Figure 3 Comparison of coefficients of variation and skewness of the historical and synthetic data records for the case studies of Heathrow airport, month of January (left) and Walnut Gulch Gauge 13, month of July (right) (adapted from Koutsoyiannis and Onof, 2001).

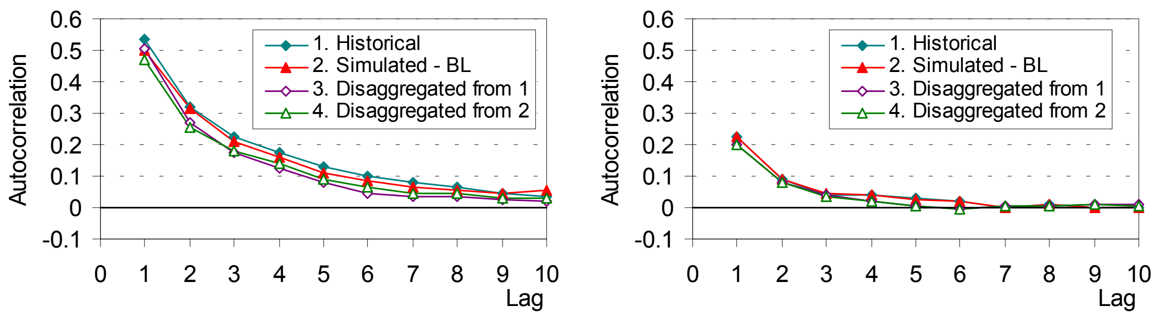

Figure 4 Comparison of autocorrelation functions of the historical and synthetic data records for the case studies of Heathrow airport, month of January (left) and Walnut Gulch Gauge 13, month of July (right) (adapted from Koutsoyiannis and Onof, 2001)..

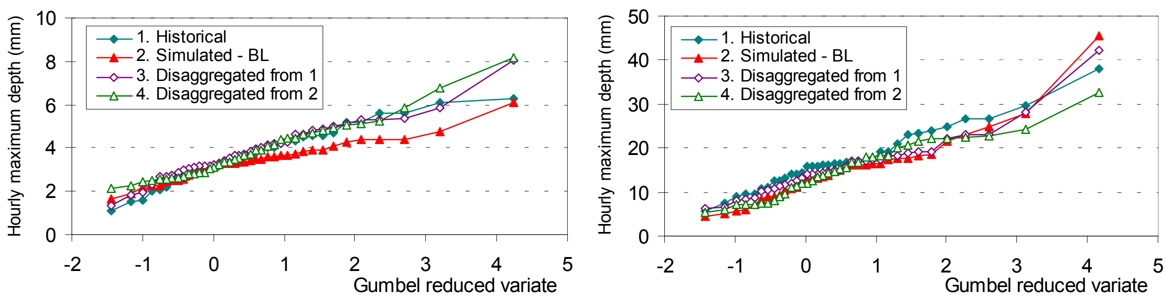

Figure 5 Comparison of empirical distributions of maximum hourly rainfall of historical and synthetic data records for the case studies of Heathrow airport, month of January (left) and Walnut Gulch Gauge 13, month of July (right) (adapted from Koutsoyiannis and Onof, 2001).

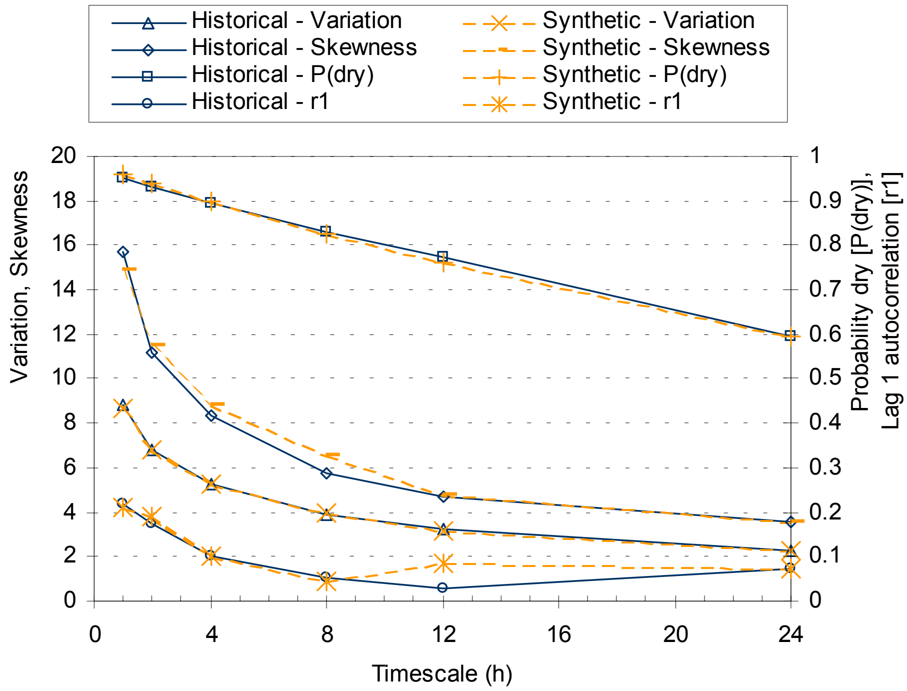

Figure 6 Comparison of coefficient of variation and skewness, probability of dry intervals and lag one autocorrelation coefficient of historical and synthetic data at timescales (aggregation levels) 1 to 24 hours for the case study of Walnut Gulch Gauge 13, month of July (Koutsoyiannis and Onof, 2001).

6. MuDRain: A model for multivariate disaggregation of rainfall at a fine time scale

As already mentioned in section 3, the problem of multiple site rainfall disaggregation presents significant differences from that of single-site disaggregation, including increased mathematical complexity. However, a multivariate approach to rainfall disaggregation is of significant practical interest even in problems that are traditionally regarded as univariate. Let us consider, for instance, the disaggregation of historical daily raingauge data into hourly rainfall. A univariate disaggregation model like Hyetos would generate a synthetic hourly series, fully consistent with the known daily series and, simultaneously, statistically consistent with the actual hourly rainfall series. Obviously, however, a synthetic series obtained by such a manner could not coincide with the actual one, but would be only a likely realisation. Now, let us assume that there exist hourly rainfall data at a neighbouring raingauge. If this is the case and, in addition, the cross-correlation among the two raingauges is significant (a case met very frequently in practice), then we could utilise the available hourly rainfall information at the neighbouring station to generate spatially and temporally consistent hourly rainfall series at the raingauge of interest. In other words, the spatial correlation is turned to advantage since, in combination with the available single-site hourly rainfall information, it enables more realistic generation of the synthesised hyetographs. Thus, for example, the location of a rainfall event within a day and the maximum intensity would not be arbitrary, as in the case of univariate disaggregation, but resemble their actual values.

With these thoughts in mind, a multivariate disaggregation model named MuDRain was built based on the general framework presented in section 4. The higher- and lower-level time scales are again daily and hourly, respectively. Here, as a lower-level rainfall model, a simplified multivariate AR(1) process was chosen. Such a choice may seem unjustified and oversimplified at first glance. A more sophisticated model like a multidimensional extension of a point process rainfall model seems to be a more consistent candidate as it would describe better features like intermittency, high skewness and clustering of rainfall. However, such a model cannot have the appropriate structure to utilise the given hourly information at one or more sites according to the hypothesis set above. In other words, it is not possible to force (or condition) such a multidimensional model to produce a given hourly rainfall depth (known for instance from measurement) at a certain point and a certain hour.

In contrast, this is possible with the multivariate AR(1) model. A careful application of an AR(1) model, even in single-variate setting reveals that this simple model can well describe intermittency and skewness during rain days (during dry days there is no need to run the model at all). This is demonstrated in the plots of Figure 7, where we can observe that, during the wet days, 73.6% of the hourly data values at a raingauge in Brue catchment, South-western England, are smaller than 0.1 mm. In fact all these data values have been recorded as zero values because typically raingauges do not record rainfall depths that are smaller than a threshold of 0.1-0.2 mm. The left panel of Figure 7 shows that if we regard all recorded zero values as small positive values (< 0.1 mm), then the complete data set can be well be modelled with a gamma distribution function. Based on this observation we can use a gamma-autoregressive (GAR) process or even an AR(1) process with skewed distribution of noise to generate hourly depths. Then, if we round-off to zero all generated values that are smaller than the given threshold (0.1 mm) we will obtain a number of zero values that are comparable to that of zero values in the historical record (right panel of Figure 7). As demonstrated further in Figure 8, if in this generation we use the historical autocorrelation coefficient of hourly depth (0.43), the resulting dry intervals during rainy days will have distribution function approaching that of the historical data.

This discussion justifies that a simple AR(1) model can perform well in reproducing intermittency and simultaneously skewness of the rainfall process in single-variate setting. In multivariate setting, its behaviour is expected to be even better provided that, at least in one location, the series of hourly rainfall is not generated by the AR(1) model but rather known either from measurement or generated already by a more consistent model like Hyetos.

As a coupling transformation, the full multivariate transformation (6) was chosen with **h** estimated from (8). In addition, the model uses repetitive sampling as described in section 4 to improve performance in preserving skewness and probability dry.

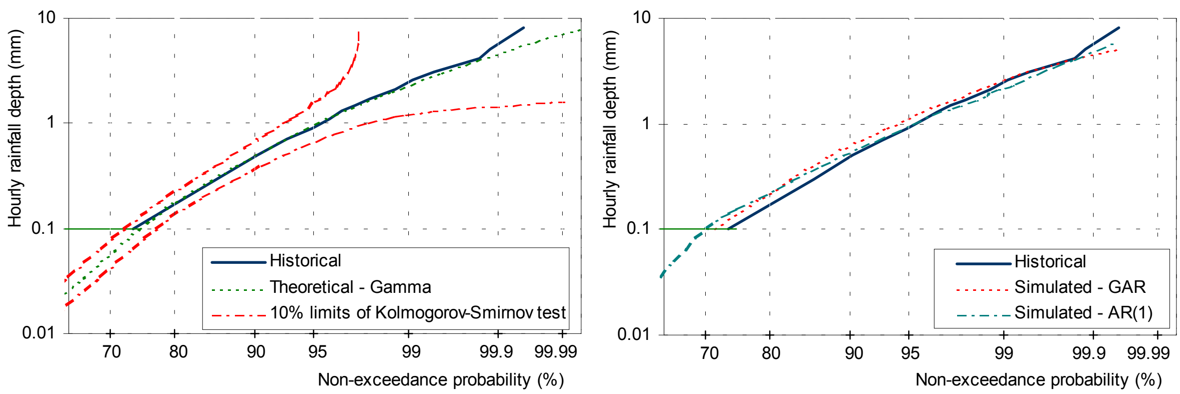

Figure 7 Plots on Weibull probability paper of probability distribution function of hourly rainfall depth during rainy days at gauge 24, Brue catchment, Southwestern England, for the month of January: (left) comparison of empirical (historical) and theoretical (Gamma) distribution functions; (right) comparison of historical and simulated distribution functions using the GAR model with rounding-off of resulting rainfall depths, and the AR(1) model without rounding-off of resulting rainfall depths (adapted from Koutsoyiannis et al., 2003).

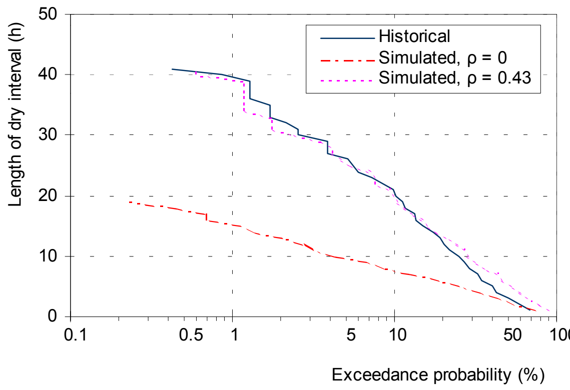

Figure 8 Plots on exponential probability paper of probability distribution function of the length of dry intervals (historical and simulated) during rainy days at gauge 24, Brue catchment, Southwestern England, for the month of January (adapted from Koutsoyiannis et al., 2003).

The program MuDRain automates most tasks of parameter estimation, performs the disaggregation and provides tabulated and graphical comparisons of historical and simulated statistics of hourly rainfall. In the parameter estimation phase, the program estimates all statistics to be preserved, apart from hourly cross-correlation coefficients whose estimation, is explained next. The program offers several options to fine-tune the model performance such as selection of the maximum allowed number of repetitions and the related maximum allowed distance, as well as a threshold for values that will be rounded-off to zero.

For the estimation of cross-correlations two options have been studied. The first option (Koutsoyiannis et al., 2003) is to fit a multidimensional point process model to the area under study, based on available daily and hourly historical data. The second option (Fytilas, 2002) is to estimate the cross-correlation coefficient r_{ij}^{h} between raingauges i and j at the hourly time scale from the corresponding cross-correlation coefficient r_{ij}^{d} at the daily time scale based on the regression equation

$$r_{ij}^{\mathrm{h}} = \left(r_{ij}^{\mathrm{d}}\right)^{m} \tag{10}$$

where the exponent m can be estimated using known pairs of cross-correlation at the hourly and daily time scale. Typical values of m found so far range between 2 and 3.

To validate the model, Koutsoyiannis et al. (2003) performed a case study in the Brue catchment in South-Western England and Fytilas (2002) performed another case study to the Tiber river basin and more specifically to the catchment of its primary tributary, the Aniene river located in Central Italy. Results of both studies indicate a good performance of the model. Some of the results of the second study (those for the month of December) are reproduced here in graphical form. The data set available was six years (January 1994 - December 1999) of hourly series from six raingauges (three of which were used as reference stations in the generation phase and the other three as test stations for comparisons between historical and simulated properties) and two daily series from other raingauges (Figure 9).

Figures 10 and 11 depict the good model performance in reproducing respectively the marginal statistics and the lag zero cross-correlations. Figure 12 (left) shows equally good performance in reproducing lag one autocorrelations, whereas Figure 13 shows that this extends also to higher lag autocorrelations, although these were not modelled explicitly (in fact they were modelled as Markovian). What forced the synthetic autocorrelations to agree with the historical ones was the reference hourly rainfall series at gauges 1-3 along with the significant cross-correlations among the different gauges. This indicates that in cross-correlated sites there is no need for a higher order AR or ARMA model, since in the multivariate framework studied (with at least one – in this case three – data series known), even the AR(1) model can reproduce historical autocorrelations (for lags > 1) adequately.

A graphical comparison of the entire distribution function of historical and simulated hourly rainfall depth during wet days is given in Figure 14 (left). The historical and simulated distribution functions in general, and the corresponding probabilities dry in particular, emerging as the values of the distribution functions for hourly rain depth 0.1 mm (also shown in Figure 12, right), are in good agreement to each other. The length of dry intervals is another variable that has not been explicitly modelled in this approach. Nevertheless, a comparison of historical and simulated probability distribution functions of this variable during wet days, depicted in Figure 14 (right) indicates an encouraging performance of the model.

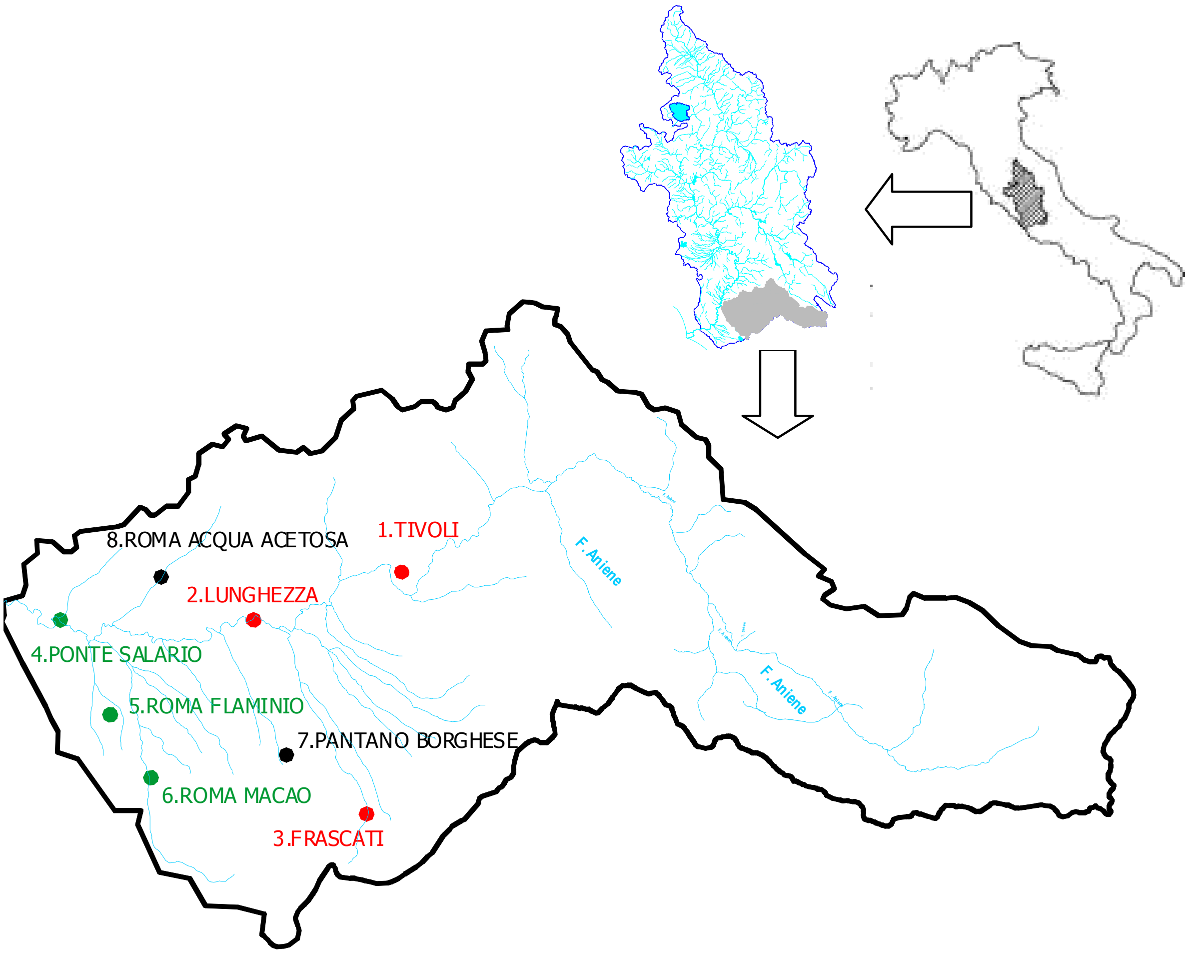

Figure 9 Schematic of the area and the raingauges used for the case study by Futilas (2002). Gauges 1-3 are reference stations (their hourly data was used in simulation), gauges 4-6 are test stations (their hourly data was used in validation) and gauges 7-8 had only daily data.

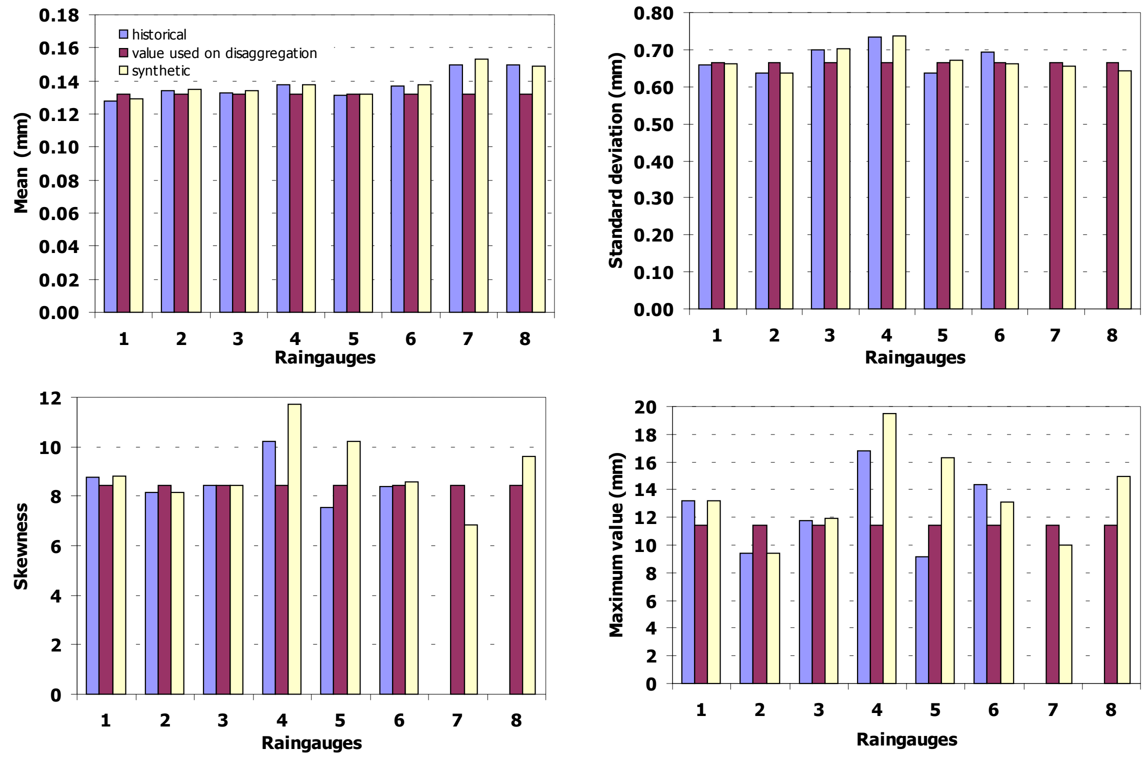

Figure 10 Comparison of marginal statistics of hourly rainfall for the application of MuDRain at the Tiber river basin (adapted from Fytilas, 2002). In each panel, the compared values are (a) the historical value, not used in the disaggregation model; (b) the value used in the disaggregation model, which is the mean of historical values of gauges 1-3; (c) the synthetic value.

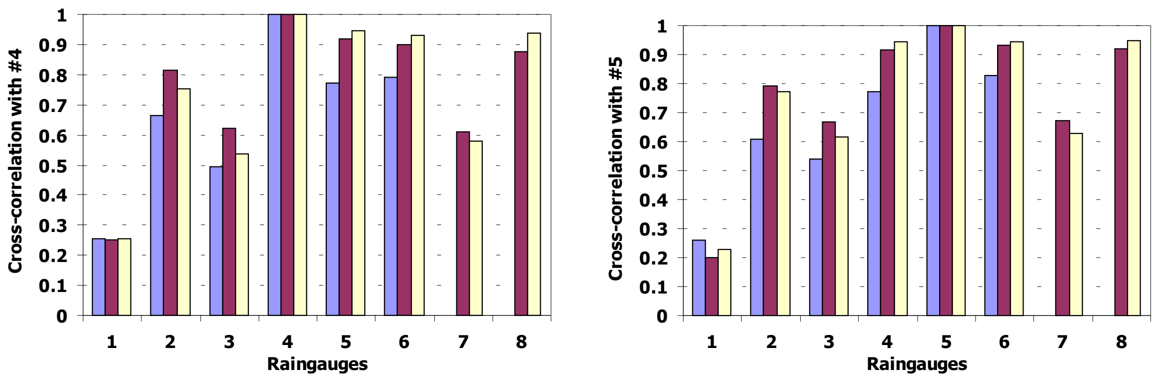

Figure 11 Comparison of cross-correlation coefficients of hourly rainfall for the application of MuDRain at the Tiber river basin (adapted from Fytilas, 2002). In each panel, the compared values are (a) the historical value, not used in the disaggregation model; (b) the value used in the disaggregation model; (c) the synthetic value.

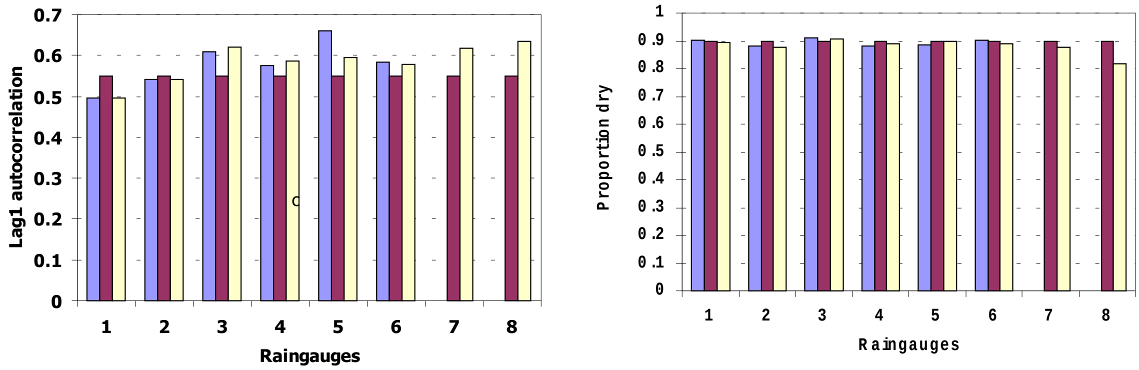

Figure 12 Comparison of lag one autocorrelation coefficients of hourly rainfall (left) and probabilities dry (right) for the application of MuDRain at the Tiber river basin (adapted from Fytilas, 2002). In each panel, the compared values are (a) the historical value, not used in the disaggregation model; (b) the value used in the disaggregation model, which is the mean of historical values of gauges 1-3; (c) the synthetic value.

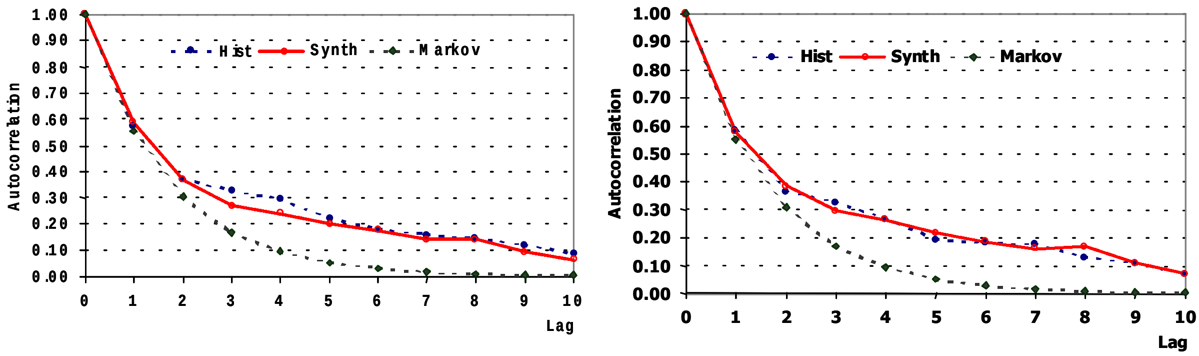

Figure 13 Comparison of autocorrelation functions of hourly rainfall for the application of MuDRain at the Tiber river basin (adapted from Fytilas, 2002), as determined from historical or simulated series, or theoretically implied from the AR(1) (Markov) model, for gauges 4 (left) and 6 (right).

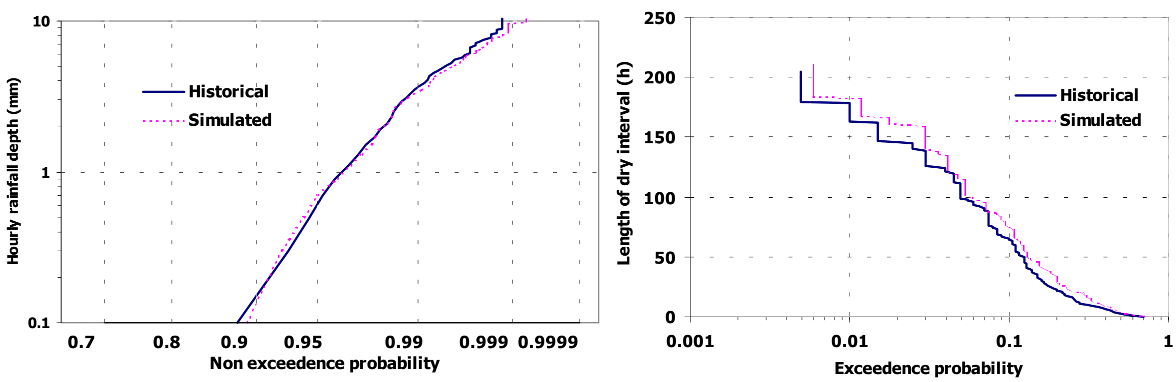

Figure 14 Comparison of historical and simulated probability distribution functions of: (left) hourly rainfall depth during wet days at gauge 4 on Weibull probability plot; (right) length of dry intervals during wet days at gauge 4 on exponential probability plot. The results are from the application of MuDRain at the Tiber river basin (adapted from Fytilas, 2002).

As an additional means of comparison, two hyetographs are given in Figure 15. Generally, the disaggregation model, by virtue of its multivariate character, the high cross-correlation coefficients and the conditioning on observed hourly series at three gauges, reproduced impressively well the actual hyetographs at all gauges (two of which are shown in Figure 15).

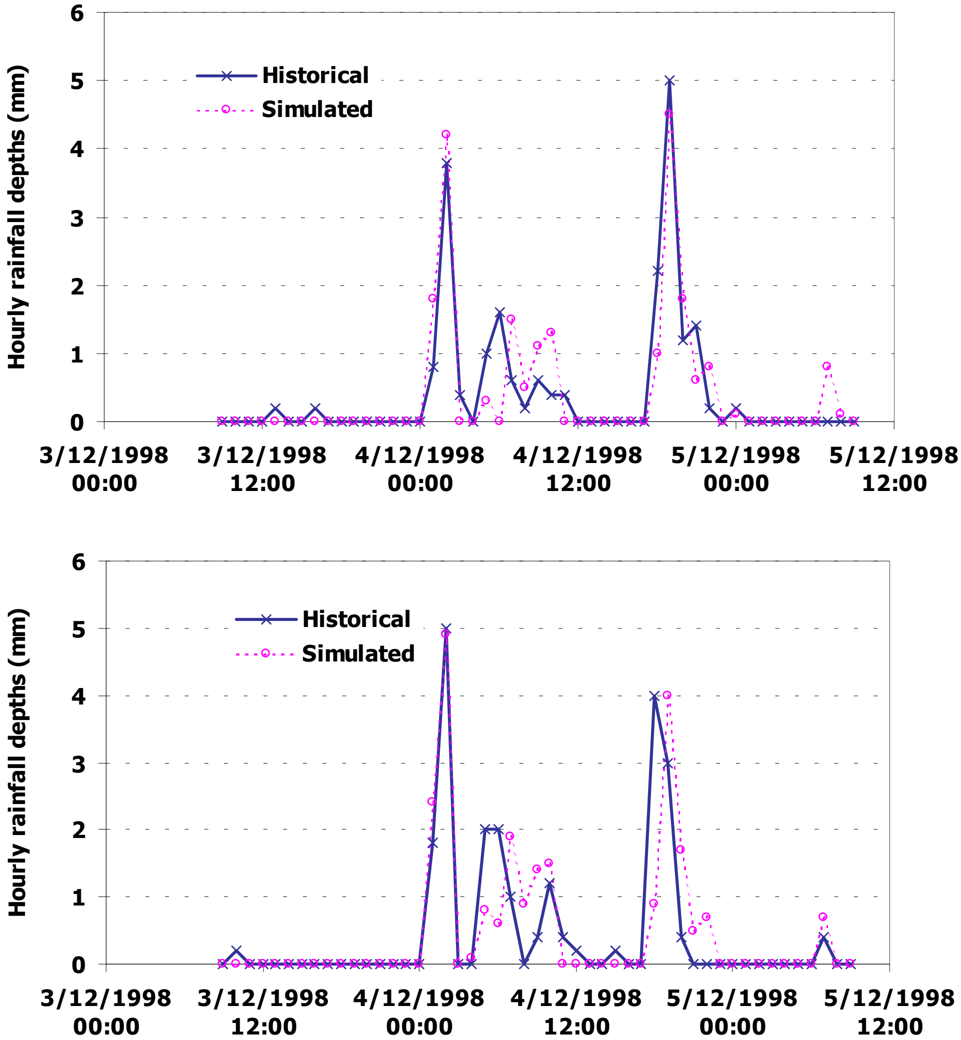

Figure 15 Comparison of historical and simulated hyetographs from the application of MuDRain at the Tiber river basin at a day with high rainfall for gauges 4 (upper panel) and 6 (lower panel) (adapted from Fytilas, 2002).

7. Conclusions and discussion

After more than thirty years of extensive research, a large variety of stochastic disaggregation models have been developed and applied in hydrological studies. Until recently, there was a significant divergence between general-purpose disaggregation methods and rainfall disaggregation methods at fine time scale. The general-purpose methods are generally multivariate while rainfall models were only applicable in single variate setting. Only recently there appeared significant developments in multivariate rainfall disaggregation along with a tendency of implementing general-purpose methodologies into rainfall disaggregation.

Multivariable rainfall disaggregation models have greater potential in hydrological applications including enhancement of historical data series and generation of simulated data series. Specifically, they can be applied to derive spatially consistent hourly rainfall series in raingauges where only daily data are available. In addition, in a simulation framework, they provide a way to take simulations of multivariate daily rainfall (incorporating spatial and temporal non-stationarity) and generate multivariate fields at fine temporal resolution.

The most recent results reproduced here are extremely encouraging. Specifically, the case studies presented, regarding the disaggregation of daily historical data into hourly series, showed that recent methodologies result in good preservation of important properties of the rainfall process such as marginal moments, temporal and spatial correlations, as well as proportions and lengths of dry intervals. In addition, multivariate methodologies provide a good reproduction of the actual hyetographs.

There is considerable flexibility in the recent methodologies, and hence potential for further refinement and remediation of weaknesses, which will certainly emerge in future applications in different climates.

Modern technologies in measurement of rainfall fields including weather radars and satellite images will undoubtedly improve our knowledge of the rainfall process and will provide more reliable and detailed information for modelling and parameter estimation of rainfall fields. One may anticipate that when such detailed information will be available the need for disaggregation models will cease. However, simulation studies will ever require investigation of the process at many scales. As a single model can hardly be appropriate for all scales simultaneously, it may be conjectured that there will be space for disaggregation methods even with the future enhanced data sets. Such future disaggregation methods should need to give emphasis to the spatial extent of rainfall fields in order to incorporate information from radar and satellite data.

Acknowledgments

The author expresses his sincere thanks to Francesco Napolitano for inviting this paper and to Paola Fytilas for providing some of the figures from her graduate thesis regarding the application of MuDRain to the Tiber river basin.

References

Bardossy, A. (1997). Disaggregation of daily precipitation. In: Casale, R., Borga, M., Baltas, E., Samuels, P. (Eds.) *River basin modelling, management and flood mitigation: concerted action.* Proc. Workshop 25 and 26 Sep. 1997, Monselice, Padua, Italy, pp.40–53. European Commission: Environment and Climate Programme.

Burian, S.J., Durrans, S.R., Tomic, S., Pimmel, R.L., Wai, C.N. (2000). Rainfall disaggregation using artificial neural networks. *J. Hydrol. Engng ASCE* 5, 299–307.

Burian, S.J., Durrans, S.R., Nix, S.J., Pitt R.E. (2001). Training artificial neural networks to perform rainfall disaggregation. *J. Hydrol. Eng.-ASCE*, 6(1), 43-51.

Connolly, R. D., Schirmer, J., Dunn, P. K. (1998). A daily rainfall disaggregation model. *Agricultural And Forest Meteorology*, 92(2), 105-117.

Cowpertwait, P.S.P., O'Connell, P.E., Metcalfe, A.V., Mawdsley, J.A. (1996). Stochastic point processes modelling of rainfall, II, Regionalisation and disaggregation. *J. Hydrol*., 175, 67-65.

Econopouly, T.W., Davis, D.R., Woolhiser, D.A. (1990). Parameter transferability for a daily rainfall disaggregation model. *J. Hydrol.* 118, 209–228.

Fytilas, P. (2002). Multivariate rainfall disaggregation at a fine time scale, Graduate thesis, 171 pp., Universita degli Studi di Roma "La Sapienza", Roma.

Garcia-Guzman, A., Aranda-Oliver, E. (1993). A stochastic model of dimensionless hyetograph. *Water Resour. Res.* 29, 2363–2370.

Glasbey, C.A., Cooper, G., McGehan, M.B. (1995). Disaggregation of daily rainfall by conditional simulation from a point-process model, *J. of Hydrol.*, 165, 1-9.

Grace, R.A., Eagleson, P.S. (1966). *The synthesis of Short-Time-Increment Rainfall Sequences*, Report no. 91. Hydrodynamics Laboratory, Massachusetts Institute of technology, Cambridge, Mass.

Grygier, J.C., Stedinger, J.R. (1988). Condensed disaggregation procedures and conservation corrections for stochastic hydrology. *Water Resour. Res.*, 24(10), 1574-1584.

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

Hershenhorn, J., Woolhiser, D.A. (1987). Disaggregation of daily rainfall. *J. of Hydrol.*, 95, 299-322.

Hoshi, K., Burges, S.J. (1979). Disaggregation of streamflow volumes. *Journal of the Hydraulics Division, Proceedings ASCE*, 105 (HY1), 27-41.

Kottegoda, N.T., Natale, L., Raiteri, E (2003). A parsimonious approach to stochastic multisite modelling and disaggregation of daily rainfall. *J. Hydrol.,* 274, 47–61.

Koutsoyiannis, D. (1988). *A disaggregation model of point rainfall*. PhD thesis, 310 pp., National Technical University of Athens, Athens.

Koutsoyiannis, D. (1992). A nonlinear disaggregation method with a reduced parameter set for simulation of hydrologic series. *Water Resources Research*, 28(12), 3175-3191.

Koutsoyiannis, D. (1994). A stochastic disaggregation method for design storm and flood synthesis, *Journal of Hydrology*, 156, 193-225.

Koutsoyiannis, D. (2001). Coupling stochastic models of different time scales, *Water Resources Research*, 37(2), 379-391.

Koutsoyiannis, D., Foufoula-Georgiou, E. (1993). A scaling model of storm hyetograph, *Water Resour. Res.*, 29(7), 2345-2361.

Koutsoyiannis, D., Manetas, A. (1996). Simple disaggregation by accurate adjusting procedures. *Water Resources Research*, 32(7) 2105-2117.

Koutsoyiannis, D., Onof, C. (2000). A computer program for temporal rainfall disaggregation using adjusting procedures, XXV General Assembly of European Geophysical Society, Nice, *Geophys. Res. Abstracts*, vol. 2. (http://www.itia.ntua.gr/e/docinfo/59/).

Koutsoyiannis, D., Onof, C. (2001). Rainfall disaggregation using adjusting procedures on a Poisson cluster model. *J. of Hydrol.*, 246, 109-122.

Koutsoyiannis, D., Onof, C., Wheater, H.S. (2001). Stochastic disaggregation of spatial-temporal rainfall with limited data, *26th General Assembly of the European Geophysical Society, Geophysical Research Abstracts, Vol. 3*, Nice, March 2001, European Geophysical Society.

Koutsoyiannis, D., Onof, C., Wheater, H.S. (2003). Multivariate rainfall disaggregation at a fine time scale, *Water Resources Research* (in press).

Koutsoyiannis, D., Pachakis, D. (1996). Deterministic chaos versus stochasticity in analysis and modeling of point rainfall series, *Journal of Geophysical Research-Atmospheres,* 101(D21), 26444-26451.

Koutsoyiannis, D., Xanthopoulos Th. (1990). A dynamic model for short-scale rainfall disaggregation. *Hydrol Sci. J.*, 35(3), 303-321.

Kumar, D. N., Lall, U., Petersen, M.R. (2000). Multisite disaggregation of monthly to daily streamflow. *Water Resour. Res.*, 36(7), 1823–1833.

Lane, W.L. (1979). Applied Stochastic Techniques, User’s Manual. Bureau of Reclamation, Engineering and Research Center, Denver, Co..

Lane, W.L. (1982). Corrected parameter estimates for disaggregation schemes, in *Statistical Analysis of Rainfall and Runoff*, edited by V. P. Singh, Water Resources Publications, Littleton, Colo..

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

Marien, J.L., Vandewiele, G.L. (1986). A point rainfall generator with internal storm structure. *Water Resour. Res.*, 22(4), 475-482.

Marshak, A., Davis, A., Cahalan, R. and Wiscombe, W. (1994). Bounded cascade models as nonstationary multifractals. *Phys. Rev. E*, 49(1), 55-69.

Mejia, J. M., Rousselle, J. (1976). Disaggregation models in Hydrology revisited. *Water Resour. Res.*, 12(2), 185-186.

Menabde, M., Harris, D., Seed, A., Austin, G. and Stow, D. (1997). Multiscaling properties of rainfall and bounded random cascades. *Water Resour. Res.*, 33(12), 2823-2830.

Oliveira, G.C., Kelman, J., Pereira, M.V.F., Stedinger, J.R. (1988). Representation of spatial cross-correlation in a seasonal streamflow model. *Water Resour. Res.*, 24(5), 781-785.

Olsson, J. (1996). Scaling and fractal properties of rainfall. PhD Thesis, University of Lund, Sweden.

Olsson, J. and Berndtsson, R. (1997). Temporal rainfall disaggregation based on scaling properties. *Third International Workshop on Rainfall in Urban Areas, ed. Fankhauser, R., Einfalt and Th., Arnbjerg-Nielsen, K., IHP-V, Technical Documents in Hydrology*, UNESCO.

Onof, C., Wheater, H.S. (1993). Modelling of British rainfall using a Random Parameter Bartlett-Lewis Rectangular Pulse Model, *J. Hydrol.*, 149, 67-95.

Pereira, M.V.F., Oliveira, G.C., Costa, C.C.G., Kelman, J. (1984). Stochastic streamflow models for hydroelectric systems. *Water Resour. Res*., 20(3), 379-390.

Rodriguez-Iturbe, I., Cox, D.R., Isham, V. (1987). Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond.*, A 410, 269-298.

Rodriguez-Iturbe, I., Cox, D.R., Isham, V. (1988). A point process model for rainfall: Further developments, *Proc. R. Soc. Lond.*, A 417, 283-298.

Salas, J.D., Delleur, J.W., Yevjevich, V., Lane, W.L. (1980). *Applied Modeling of Hydrologic Time Series*, Water Resources Publications, Littleton, Colo.

Schaake J.C., Ganslaw, M.J., Fothergill, J.W., Harbaugh, T.E. (1972). Multivariate rainfall generator for annual, seasonal, monthly and daily events, *Proceedings of the International Symposium on Mathematical Modeling Techniques in Water Resources Systems*, vol. 2, 437-460, Ottawa, Canada.

Sivakumar, B., Sorooshian, S., Gupta, V.J., Gao, X. (2001). A chaotic approach to rainfall disaggregation. *Water Resour. Res.*, 37, 61–72.

Socolofsky, S., Adams, E.E. (2001). Disaggregation of daily rainfall for continuous watershed modeling. *J. Hydrol. Engng ASCE*, 6, 300–309.

Stedinger, J.R., Vogel, R.M. (1984). Disaggregation procedures for generating serially correlated flow vectors, *Water Resour. Res.,* 20(1) 47-56.

Stedinger, J.R., Pei, D., Cohn, T.A. (1985). A condensed disaggregation model for incorporating parameter uncertainty into monthly reservoir simulations. *Water Resour. Res.*, 21(5) 665-675.

Tang, A.C.K. (1999). Multifractal rainfall disaggregation, Final Year Project, Dept. of Civil/Environment. Engin., Imperial College.

Tao, P.C., Delleur, J.W. (1976). Multistation, multiyear synthesis of hydrologic time series by disaggregation. *Water Resour. Res.*, 12(6), 1303-1312.

Todini, E. (1980). The preservation of skewness in linear disaggregation schemes, *J. Hydrol.*, 47, 199-214.
Valencia, D., Schaake, J.C. (1972). A disaggregation model for time series analysis and synthesis, Report no. 149, Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, Massachusetts Institute of Technology, Cambridge, Mass.
Valencia, D., Schaake, J.C. (1973). Disaggregation processes in Stochastic Hydrology. *Water Resour. Res.*, 9(3) 211-219.
Woolhiser, D.A., Osborn, H.B. (1985). A stochastic model of dimensionless thunderstorm rainfall. *Water Resour. Res.*, 21(4) 511-522.
Zarris, D., Koutsoyiannis, D., Karavokiros, G. (1998). A simple stochastic rainfall disaggregation scheme for urban drainage modelling, *Proc. Fourth Int. Conf. on Developments in Urban Drainage Modelling*, September 1998, International Association of Water Quality & Imperial College, London, pp. 85-92, 21-24.