UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA"

MULTIVARIATE RAINFALL DISAGGREGATION AT A FINE TIME SCALE

A dissertation submitted to the Department of Hydraulics
of the Faculty of Civil Engineering

ADVISER
Prof. Francesco Napolitano

COADVISER
Prof. Demetris Koutsoyiannis

Academic Year 2001-2002

Table of Contents

Introduction

The understanding of hydrological processes that occur in nature is one of the most important tasks for both hydrologists and civil engineers for the design of almost all hydrological applications and civil engineering works. Rainfall is the main input to all hydrological systems, and a wide range of hydrological analyses, for flood alleviation schemes, management of water catchments, water quality or ecological studies, require quantification of rainfall inputs at both daily and hourly time scales. This may be possible using empirical observations, but there is often a need to extend available data in terms of record length, temporal resolution and/or spatial coverage.

In Europe and many other countries in the world, there is a large number of daily raingages, which have often been operational for a few decades and offer a large amount of daily data but at the same time there is a lack of sub-daily information due to the absence of hourly raingages or to the fact that the existing ones have been operational only for a few years making the length of the recorded series insufficient for all hydrological purposes and statistical analyses.

Therefore a common problem in hydrological studies is the limited availability of data at appropriately fine temporal and/or spatial resolution.

Rainfall disaggregation emerged as an important tool for facing this problem.

Disaggregation techniques have the ability to increase the time or spatial resolution of certain processes, such as rainfall and runoff while simultaneously providing a multiple scale preservation of the stochastic structure of the hydrologic processes.

This definition of disaggregation distinguishes it from downscaling, which aims at producing hourly data with the required statistics but that do not necessarily add up to the observed hourly data.

The first developed disaggregation models were multivariate, i.e performed a simultaneous disaggregation at several sites. Typical examples of such models are those by *Valencia and Schaake* [1972, 1973], *Mejia and Roussrelle* [1976], *Tao and Delleur* [1976], *Hoshi and Burges* [1979], *Todini* [1980], *Stedinger and Vogel* [1984]. These models express the vector containing all unknown lower-level variables as a linear function of the higher-level variables of all sites and some innovation variates.

Thus they attempt to reproduce all covariance properties between lower-level variables as well as those between lower-level and higher-level variables among all sites and time steps. This results in a huge number of parameters because of the large number of cross correlations that they attempt to reproduce.

Different procedures reducing the required number of parameters have been developed. The staged disaggregation models [*Lane* 1979, 1982; *Salas et al*. 1980; *Stedinger and Vogel* 1984; *Grygier and Stedinger* 1988, 1990; *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 [*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 parameters by explicitly modeling fewer of the correlations among the lower-level variables. Stepwise disaggregation schemes [*Santos and Salas* 1992; *Salas,* 1993] perform the disaggregation always in two parts. For example an annual value is disaggregated into 12 monthly values by first disaggregating the annual value into the first monthly and the sum of the remaining 11 months. Then the latter sum is disaggregated into the monthly value and the sum of the remaining 10 values, and so on until all monthly values are obtained.

However, 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 have been proposed and used, particularly for the disaggregation of rainfall, but do not exhibit the generality of these linear schemes.

Another stepwise disaggregation approach was proposed and developed by *Koutsoyiannis and Xanthopoulos* (1990). This approach, called the "Dynamic Disaggregation Model" (DDM), at each step disaggregates a given amount in two parts, the variable of the next period and the amount to be disaggregated (in subsequent steps) across the remaining periods. In this respect, DDM is very similar to the *Santos and Salas* stepwise disaggregation scheme. However, there are some differences. At each step, DDM uses two modules from which one is a nonlinear generation module that disaggregates the given amount in two parts and the second is linear, closely related to a seasonal AR(1) also known as PAR(1), that determines the parameters required for the generation.

In this way the overall model uses exactly the same parameter set as the PAR(1) model, which involves the minimum number of parameters, significantly lower than direct disaggregation schemes.

A further approach proposed and developed by *Koutsoyiannis and Manetas* (1996) is the "Simple disaggregation by Accurate Adjusting Procedures". This method keeps some ideas of the Dynamic Disaggregation Model approach but is simpler. It initially retains the formalism, the parameter set, and the generation routine of the PAR(1) model and then uses an adjusting procedure to achieve the consistency of lower-level and higher-level variables by modifying the values of the PAR(1) model as to preserve the additive property. Another idea that is used in this method is repetition: Instead of running the generation routine of the PAR(1) model one for each period, we run it several times and choose that combination of generated values, which is in closer agreement with the known value of the higher-level variable.

The general idea of the adjusting procedures in disaggregation developed by *Koutsoyiannis and Manetas* was later on combined with a successful generation model based upon a Poisson cluster process (*Onof and Wheater 1993*). This is the approach proposed in "Rainfall Disaggregation using Adjusting Procedures on a Poisson Cluster model" by Koutsoyiannis and Onof (2001). The rainfall model chosen is the Bartlett-Lewis because of its proven ability to reproduce important features of the rainfall field from the hourly to the daily scale and above. Then using the adjusting procedures the time series of lower-level (hourly) generated by the rainfall model are modified so as to be consistent with the given higher-level time series (daily) and simultaneously preserve the stochastic structure implied by the rainfall model. This methodology is of important practical interest because it offers the possibility to extend short hourly time-series using longer-term daily data at the same point and because, at the same time it represents a theoretical basis for application when no hourly data are available.

More recently *Koutsoyiannis* (2001) studied the problem of multiple site rainfall disaggregation as a means for simultaneous spatial and temporal disaggregation at a fine time scale. He investigated the possibility of using available hourly information at one raingage to generate spatially and temporally consistent hourly rainfall information at several neighboring sites in the case, met very frequently, that the cross-correlation coefficients between the raingages are significant. The combination of spatial correlation and available single-site hourly rainfall information enables more realistic generation of the synthesized hyetographs, i.e. the synthetic series generated will resemble the actual one.

This particular case of general multivariate spatial-temporal disaggregation problem is the subject of this thesis, which is essentially based on the disaggregation methodology proposed and studied by *Koutsoyiannis.*

In Chapter I there is a schematic description of two of the models proposed and studied by *Koutsoyiannis et al.* in the last decade and more specifically The Dynamic Disaggregation Model, Simple disaggregation by accurate adjusting procedures.

These two models are introduced in the present work because the represent the theoretical basis that guided the development of two other models studied by *Koutsoyiannis et al.* regarding the problems of single-site and multiple-site spatial temporal disaggregation at fine scale.

The methodologies proposed for facing these problems are the subject of “Rainfall disaggregation using adjustment procedures on a Poisson cluster model” and “Multivariate Rainfall Disaggregation at a fine time scale” which are described in Chapters II and III respectively, along with their computer implementations, Hyetos and MuDRain.

A detailed case study, concerning the application of the disaggregation procedure to the basin of Tiber River and conclusions on the methodology proposed are included in chapter IV.

Some additional information on the implementation of the multivariate rainfall disaggregation, i.e. the help mode of the program MuDRain, is contained in the Appendix.

CHAPTER I

THE DISAGGREGATION METHODOLOGY DEVELOPED IN NTUA

Here are described two models regarding short-scale disaggregation, proposed and studied in the last decade by Koutsoyiannis. This section can be seen as a sort of path through time of the approaches chosen to achieve the multiple-site disaggregation at a fine scale.

A dynamic model for short-scale rainfall disaggregation

The single-site dynamic disaggregation model developed by *Koutsoyiannis and Manetas* (1990) is a generalized step-by-step approach to stochastic disaggregation problems. The model development was intended for application to short-scale rainfall disaggregation problems. Important features of the model are:

1. the modular structure (composed of two parts studied separately) allowing various configurations of the model
2. the flexible step-by-step approach allowing the use of side procedures, adjusting properly the generated values in each step without loss of the additive property;
3. the simple analytical equations allowing a varying number of low-level variables and varying scales.

A combination of the dynamic disaggregation model with a developed rainfall model gives a point rainfall generator, performing with monthly through hourly time scales. The rainfall model can incorporate a varying number of parameters (4 to 12), depending on the desired accuracy.

An advantaged of the combined model is that the same disaggregation procedure is used for four different purposes:

- the determination of the starting points of the rainfall events
- the generation of rain durations
- the generation of event rain depths
- the disaggregation of event rain depths into hourly depths.

The rainfall generator models the total rainfall regardless of intensity. The internal disaggregation part of the model (phase 2) may be applied independently to severe storms in order to simulate their time profiles. Therefore the model may be useful for simulation of severe flood-producing storms and estimation of design storms.

The Dynamic Disaggregation Model (DDM)

The essential elements of the model, described in detail by Koutsoyiannis (1988), are the following:

(a) The disaggregation of a high-level variables, Z, into its k components (low-level variables, X_i , i= 1,…, k), is performed in k-1 sequential steps.

(b) At the beginning of the ith step, the amount-still-to-go, Si, is known, and X_i is generated. The remaining quantity $S_{i+1}= S_i - X_i$ is transferred to the next step.

(c) In each step the distribution function of (X_i, S_i), conditional on previous generated information, is determined or approximated via *conditional moments*. It is assumed that the sequence of X_i has certain properties allowing the calculation of conditional moments, e.g. it is an autoregressive sequence (AR).

(d) The generation of X_i is performed by the so-called *bisection procedure*, which can take several forms depending on the particular marginal distribution of the low-level variables.

The realization of the model includes two parts, the *conditional moments determination* that is influenced by the type of stochastic structure, and the *bisection procedure* affected mainly by its marginal distribution type, which can be studied separately.

The configurations studied by the model concern mainly single-site problems, described by Markov sequences, with Gaussian or Gamma marginal distributions. Therefore it can be used in any single-site hydrological application fulfilling or approximating these conditions.

Model Equations

The two parts of the realization are the determination of conditional moments and the bisection procedure and are studied separately.

Conditional moments determination

Let the low-level variables X_i , i= 1,…, k, add up to the high-level variable, Z :

$$X_1 + X_2 + ... + X_k = Z$$

The low-level variables are considered as a sub-set of an infinite stochastic sequence, (…,X_{-1}, X_0, X_1,…, X_k, X_{k+1}, …), that characterizes the current stage of disaggregation. It is assumed that the disaggregation procedure has already been completed at the previous stages; thus all previous X_i's have known values (X_0=x_0, X_{-1}=x_{-1},…).

The initial parameters of the model, at the current stage are the first and second moments of the low-level variables and form the following groups:

a. mean values of X_i, μ_i ;

b. variances of X_i, σ_i^2;

c. covariances between X_i, X_j (i, j>0), σ_{ij} ; and

d. covariances between X_i, (i>0), with variables X_j (j<0) of previous stages.

The number of independent parameters of groups (a), (b), (c) is *k, k, k(k-1)/2,* respectively, and in total, $(k^2+3k)/2$. If k previous variables are considered as affecting the current stage, the number of parameters of group (d) is k^2. The total number of initial parameters is $3(k^2+k)/2$.

Consider now the ith disaggregation step of the current stage, concerning the generation of the low-level variable X_i based on:

$$X_i + S_{i+1} = S_i$$

where the amount still to go:

$$S_i = X_i + X_{i+1} + ... + X_k = Z - X_1 - ... - X_{i-1}$$

has a known value, given that previous steps have been completed. Consider as intermediate parameters of the ith step, the first and second moments of the remaining low-level variables X_i, $i \le j \le k$, conditional on the previous generated information $\Omega_{i-1} = (X_{i-1} = x_{i-1},...X_1 = x_1, X_0 = x_0, X_{-1} = x_{-1})$. These parameters form groups similar to the groups (a), (b), and (c) of the initial parameters.

Finally, consider, as final parameters of the ith step at the current stage of the model, the first and second moments of the variables X_i and S_i, conditional on the previously generated information. These final parameters are:

$$\mu_X = E[X_i|\Omega_{i-1}] \quad \sigma_X^2 = Var[X_i|\Omega_{i-1}]$$
$$\mu_s = E[S_i|\Omega_{i-1}] \quad \sigma_S^2 = Var[S_i|\Omega_{i-1}]$$
$$\sigma_{XS} = Cov[X_i, S_i|\Omega_{i-1}]$$

and they are fully determined by linear combinations of the intermediate parameters. These parameters are the link with the bisection procedure; at the ith step, their values (as well as the known value of S_i) are passed to the bisection procedure, which proceeds to the generation of the X_i value (as well as S_{i+1}).

Take now the case in which the sequence of low-level variables is "wide sense" Markov. The following relation is a consequence of the Markovian property:

$$Cov[X_i, X_j]Cov[X_j, X_l] = Cov[X_i, X_l]Var[X_j] \quad i < k \leq l$$

This property reduces the number of the initial parameters. Thus the initial parameters of group (c) can be determined in terms of the parameters of group (b) and the *(k-1)* lag-one correlation coefficients:

$$\rho_i = Corr[X_i, X_{i-1}] = \sigma_{i,i+1} / (\sigma_i \sigma_{i-1}) \quad i = 2,..., k.$$

Similarly, the independent parameters of group (d), are reduced to one, since the covariances with the low-level variables of previous stages can be determined in terms of groups (b) and (c) of the current and previous stages, and the lag-one correlation coefficient ρ_1. Therefore the total number of parameters in this case is $3k$. Any covariance between low-level variables is given by:

$$\sigma_{ij} = \rho_j ... \rho_{i+1} \sigma_j \sigma_i \quad i < j \leq k$$

The intermediate parameters of the ith step are easily derived, considering that the wide sense Markov sequence X_i satisfies the difference equation

$$X_i - a_i X_{i-1} = V_i$$

where V_i is a sequence of uncorrelated random variables and a_i is a sequence of constants. The resulting equations are:

$$E[X_j | X_{i-1} = x_{i-1}] = \mu_j + \rho_j ... \rho_i \sigma_j \tilde{x}_{i-1}$$
$$Var[X_j | X_{i-1} = x_{i-1}] = \sigma_j^2 (1 - \rho_j^2 ... \rho_i^2)$$
$$Cov[X_l, X_j | X_{i-1} = x_{i-1}] = \rho_l ... \rho_{j+1} (1 - \rho_j^2 ... \rho_i^2) \sigma_l \sigma_j$$
$$where \quad 1 \leq i < j < l \leq k; \quad and \quad \tilde{x}_{i-1} = (x_{i-1} - \mu_{i-1}) / \sigma_{i-1}$$

The final parameters of the ith step are:

$$E[X_i | X_{i-1} = x_{i-1}] = \mu_i + \rho_i \sigma_i \widetilde{x}_{i-1}$$
$$Var[X_i | X_{i-1} = x_{i-1}] = \sigma_i^2 (1 - \rho_i^2)$$
$$E[S_i | X_{i-1} = x_{i-1}] = \sum_{j=i}^{k} \mu_j + \widetilde{x}_{i-1} \sum_{j=i}^{k} \rho_j ... \rho_i \sigma_j$$
$$Var[S_i | X_{i-1} = x_{i-1}] = \sum_{j=i}^{k} \sigma_j^2 (1 - \rho_j^2 ... \rho_i^2) + 2 \sum_{j=i}^{k-1} \sum_{l=j+1}^{k} \rho_l ... \rho_{j+1} (1 - \rho_j^2 ... \rho_i^2) \sigma_l \sigma_j$$
$$Cov[S_i, X_i | X_{i-1} = x_{i-1}] = \sigma_i (1 - \rho_i^2) \left[\sigma_i + \sum_{j=i+1}^{k} \rho_j ... \rho_{i+1} \sigma_j \right]$$
$$where \quad 1 \le i \le k$$

Bisection Procedure

The bisection problem, i.e the generation of variables *X* and *Y*, such that

X+Y=S

where *S* has a known value, s, can be studied independently of the other part of the model. What is required here is to determine the conditional distribution of *X |S*. the variable X can then be generated by this conditional distribution, and Y is obtained by *X+Y=S*. Due to the difficulties of the determination of conditional distribution, a simpler-based approach is preferable. This is done by assuming a proper auxiliary random variable W, and an explicit form *R(S,W),* such as:

X =R(S, W)

with parameters being determined via marginal and joint moments of *(X, S)*.

The linear bisection scheme:

X =R(S, W)=aS+W

where *W* is a random variable independent of *S*, is ideal for joint normal variables. If *W* is assumed normal, then the bisection scheme preserves completely the distribution function of (*X, Y, S*). Moreover, when this scheme is combined with the other part of the model (conditional moments determination), the complete joint distribution function of low-level variables is preserved, if it is multidimensional normal. However, this bisection scheme is not proper for skewed distributions, since it cannot preserve non-zero skewness coefficients.

Another simple bisection scheme, the so-called proportional one, is defined by:

X =R(S, W)= W S

where *W* is a random variable generally dependent on *S*, referred to as proportional variable. The degree of correlation between *W* and *S,* and this simplifies the problem. The proportional scheme is ideal for gamma distributed variables, since it has been shown that when *X* and *Y* are independent gamma distributed, having equal scale parameters, and W is assumed independent of S and beta distributed, the complete distribution $F_{XYS}()$ is preserved.

This preservation expands to the whole sequence of low-level variables under the same assumptions. In the general case of dependent gamma marginal variables, with different scale parameters, the proportional scheme still gives satisfactory approximations of the gamma marginal distributions.

The parameters of the proportional scheme, i.e. the moments of W conditional on S, for the general gamma case are given by:

$$\mu_{W|S} = \theta - \eta\left(1 + \frac{\mu_S^2}{\sigma_S^2}\right) + \eta \frac{\mu_S^2}{\sigma_S^2} \frac{s}{\mu_S}$$

$$\sigma_{W|S}^2 = \frac{\sigma_X^2 + \theta^2 \sigma_s^2 - 2\theta\sigma_{XS}}{\sigma_S^2 + \mu_S^2} - \eta^2 \left(3 + \frac{\mu_S^2}{\sigma_S^2}\right)$$

where :

$$\theta = \frac{\mu_X}{\mu_S} \qquad \eta = \frac{\sigma_{XS} - \theta\sigma_S^2}{\mu_S^2 + \sigma_S^2}$$

The above equations have been obtained under the assumption of a linear dependence between S and W. when X and Y are independent with common scale parameter, S and W should be assumed independent and η=0

It must be emphasized that in the above analysis and the relevant equations, all variables are in their initial form (no differences from means). Hence, if the variables are positive, W should be bounded in [0,1]. The two parameter beta distribution is a proper representation for the distribution of $X|S$. finally if X and S have three parameter gamma distributions, they can be replaced in the above analysis with their respective differences from their lower bounds, and the same bisection procedure used.

The Rainfall Model

The rainfall model used in combination with the dynamic disaggregation model represents the rainfall process in discrete time, from an hourly to a monthly time scale, using as a base the intermediate scale of a rainfall event. The main parts of the rainfall model are summarized as follows:

Rainfall event-rainfall occurrence

A rainfall event is considered as an individual entity which can be identified in a historical rainfall record. Successive rainfall events were defined as statistically independent, with starting points forming a Poisson process. Values of the separation time (c), i.e. the minimum dry time interval for two successive rainfall pulses to be considered as indipendent events, were obtained by a developed criterion, based on the Kolmogorov-Smirnov test, and were found to lie in the range c=5-7h.

The complete description of the rainfall occurrence process requires that the joint distribution function $F_{VDB}\ (v,\ d,\ b)$ is known, where V is the rainfall inter-arrival time, D is the duration of the event and B is the time between events (dry interval). This was based on:

(a) the obvious relation:

$$D+B=V$$

(b) the consequence of the event definition (a property of the Poisson process) that the marginal distribution of V-c is exponential, that is:

$$f_V(v) = \omega e^{-\omega(v-c)} \quad v \geq c$$

(c) the assumption that the conditional distribution of d, given V, comprises two additive parts, an exponential part independent of V, and a triangular part dependent on V, that is:

$$f_{D|V}(d,v) = \begin{cases} \delta e^{-\delta d} + 2d e^{-\delta(v-c)} / (v-c)^2 & 0 \leq d \leq v-c \\ 0 & elsewhere \end{cases}$$

The marginal densities of the distributions of the event duration, D, and the time between storms, B, derived theoretically from these assumptions are:

$$f_D(d) = (\delta + 2\omega)e^{-(\delta+\omega)d} - 2\omega(\delta+\omega)d\varepsilon[(\delta+\omega)d] \quad d \geq 0$$

$$f_B(b) = \frac{\omega\delta}{\delta+\omega} e^{-\omega(b-c)} - 2\omega e^{-(\delta+\omega)(b-c)} + 2\omega[1 + (\delta+\omega)(b-c)]\varepsilon[(\delta+\omega)(b-c)] \quad b \geq c$$

$$where \quad \varepsilon(x) = \int_x^\infty \left(e^{-\xi}/\xi\right)d\xi$$

$f_D(d)$ is quite similar to the exponential distribution, but $f_B(b)$ deviates, mainly in its lower tail, from the exponential, Weibull and gamma distributions (see figures 2, 3).

The distribution of the event rain depth, H, conditional on D, has been assumed gamma, with mean and standard deviation linearly depending on the duration:

$$E[H|D] = (d+a)\mu_\phi - b$$
$$\{Var[H|D]\}^{1/2} = (d+a)\sigma_\phi$$

where $a, b, .\mu_\phi$ *and* σ_ϕ are constants.

Internal rainfall event structure

Given a specific rainfall event, with duration *D* (considered as an integer multiple of Δ=1h), and total depth *H*, the sequence of hourly depths X_i, in the interior of the event, is related with *H* by:

$$X_1 + X_2 + ... + X_{D|\Delta} = H$$

The following main assumptions concerning the structure of the hourly depths have been used:

(a) The sequence of hourly rain depths is non-stationary; the statistics of a specific X_i depend on duration of the event, as well as on its time position in the event. It is accepted that these two influences are separable, and can be described by:

$$X_i = k(D)g(\theta_i)Z_i$$

where θ_i is the non-dimensionalized time position (t_i / D); Z_i is a sequence of dependent, identically distributed random variables, referred to as homogenized hourly rain depths, with mean μ_Z and standard deviation σ_z; *k()* and *g()* are properly defined functions.

(b) The covariance structure of Z_i, in a specific event is assumed to be stationary Markovian:

$$Cov[Z_i, Z_{i+1}] = (\rho_1)^j \sigma_Z^2$$

The lag-one correlation coefficient, ρ_1, generally depends on duration. The covariance between variables of different events is zero.

Secondary assumptions concern the form of *g(θ)*, which has been assumed linear, i.e.:

$$g(\theta) = g_0 + g_1\theta$$

where g_0 and g_1 are constants, and the distribution of Z, which is J-shaped and has been considered as gamma or Weibull, depending on the fit historical data.

The distribution of X_i, marginal or conditional on *D*, may be approximated by the same type as one of *Z*. The conditional mean and standard deviation are:

$$E[X_i | D = d] = k(d) g(\theta_i) \mu_Z$$
$$\{Var[X_i | D = d]\}^{1/2} = k(d) g(\theta_i) \sigma_Z$$
$$k(d) = 1 + (a - b / \mu_\phi) / d$$

For estimating the correlation coefficient, ρ_1 :

$$\sum_{i=1}^{d/\Delta - 1} \xi_i(d) [\rho_1(d)]^i = \xi_0(d)$$
$$where \quad \xi_0(d) = \frac{(d+a)^2 \sigma_\phi^2}{2k^2(d) \sigma_Z^2} - (1/2) \sum_{i=1}^{d/\Delta} g^2(\theta_i)$$
$$\xi_i(d) = \sum_{i=1}^{d/\Delta - 1} g(\theta_i) g(\theta_{i+1}) \quad i = 1, 2, ..., d/\Delta - 1$$

Monthly rainfall

The three variables describing the monthly rainfall are the number of rainfall events, *N*, the monthly rainfall depth, *S*, and the monthly rain duration, *U*.

The marginal distribution of N is a modified Poisson, the modification caused by the lower bound (*c*) of the inter-arrival time; given the month duration, τ, the probability function is accurately approximated by:

$$P_n = \Pr(N = n | \tau) = (1 + \kappa) \frac{(\lambda - \kappa n)^n}{n!} e^{-(\lambda - \kappa n)} \quad n = 0, 1, ..., m$$
$$where:$$
$$\lambda = \omega \tau = \tau / (\mu_V - c)$$
$$\kappa = \omega c = c / (\mu_V - c)$$
$$m = [\tau / c] = [\lambda / \kappa]$$

As a satisfactory approximation for simulation purposes, the gamma distribution was used for both *S* and *U*; their moments are completely determined by the corresponding moments of *V, D* and *H*

Model parameters

All the rainfall model parameters can be expressed in terms of four main independent parameters, namely the separation time (*c*), and the mean values of rainfall inter-arrival time (μ_V), event duration (μ_D) and event depth (μ_H), and five secondary independent parameters, namely *a, b*, σ_ϕ, σ_Z and g_1.

Three more secondary parameters may be introduced concerning the mean and standard deviation of events with duration equal to 1 h, ($\mu_{H1,}$ σ_{H1}), because it was found that equations

$$E[H|D] = (d+a)\mu_\phi - b$$

$$\{Var[H|D]\}^{1/2} = (d+a)\sigma_\phi$$

may not apply to these events, and the probability that hourly depth equals zero (p_0), a possibility permitted by the event definition. This probability may be represented by the value of the continuous distribution function of X or Z (gamma or Weibull) at the point x=0.05 mm, since, in fact, values less than 0.05mm are interpreted as zero (see Fig. 5). However, if this representation is not satisfactory, then p_0 should be used as an independent parameters is 12, and thisnumber may be reduced to 4, by omitting secondary parameters. All parameters are season-dependent.

The Rainfall Disaggregation (combination of DDM with a rainfall model)

Consider now the problem of disaggregation of monthly rainfall into hourly depths. Because of the intermittent nature of the rainfall process, a two-phase disaggregation procedure has been adopted. The first phase, external disaggregation, is to generate the rainfall events, while the second, internal disaggregation, generates hourly depths within each event.

The disaggregation in both phases is a combination of the dynamic disaggregation model and rainfall model. The latter calculates the initial parameters for the former which performs the generation. The Markovian configuration of the model studied, with the proportional bisection scheme, is satisfactory for both phases; for cases in which low-level variables are independent, the correlation coefficients are set to zero. Particular additional procedures, causing proper side effects on the generated variables, have been designed and used along with disaggregation model. The step-by-step course of the model permits the use of such side procedures.

It is supposed that, at the start of use of model, monthly rainfall variables, i.e. the number of rainfall events, *N*, the monthly depth, *S*, and the monthly rain duration, U, have known values. Nevertheless, the implementation was designed to include a separate part, generating, if needed, one or more of these values, in order to be a complete rainfall generator, from monthly through hourly time scales.

Below is a coded description of each disaggregation phase.

External disaggregation

1(a) External disaggregation-section a

Input: Number of events, N=n

Output: Inter-arrival times V_i(=low-level variables) determining the starting points of events.

Basic relation:

$$\sum_{i=1}^{n}(V_i - c) = T^*$$

where $T^* = \tau - nc - A + B$, is the high-level variable, A is the time distance of the starting point of the first event of the current month, and B is the same for the next month.

Remarks: The distribution of (V_i-c) is exponential, a particular case of the gamma.Successive low-level variables are independent.

Side procedures: A and B are generated separately, using the same exponential distribution of V_i

1(b) External disaggregation-section b

Input: Number of events, N=n, monthly duration, U=u (=high-level variable), event times V_i=v_i

Output: Event durations D_i (=low-level variables)

Basic relation:

$$\sum_{i=1}^{n} D_i = U$$

Remarks: The disaggregation model uses only the exponential part of conditional distribution of equation

$$f_{D|V}(d,v) = \begin{cases} \delta e^{-\delta d} + 2de^{-\delta(v-c)}/(v-c)^2 & 0 \le d \le v-c \\ 0 & elsewhere \end{cases}$$

while the triangular part is left to a side procedure. Succesive low-level variables are independent.

Side procedures: If the value, d_i , generated by the disaggregation model is greater than v_i-c, it is rejected and new one is generated in the range (0, v_i-c), using the triangular distribution.

1(c) External disaggregation-section c

Input: Number of events, N=n, monthly rain depth, S=s (=high-level variable), event times, event durations, D_i=d_i

Output: Event durations H_i (=low-level variables)

Basic relation:

$$\sum_{i=1}^{n} H_i = S$$

Remarks: Distributions of H_i , conditional on D_i , is gamma, with moments depending on d_i . Succesive low-level variables are independent.

Side procedures: None.

Internal disaggregation

Input: Event rain depth, $H_i = h_i$ (=high-level variable), event duration, $D_i = d_i$

Output: Hourly rain depths X_{ij} (=low-level variables).

$$\sum_{j=1}^{d_i/\Delta} X_{ij} = H_i$$

Remarks: Distributions of X_{ij}, conditional on D_i , is gamma or Weibull, with moments depending on d_i and j.Both distribution types are treated with the proportional bisection scheme, and the adjustment of the Weibull distribution is left to a side procedure. The covariance structure of low-level variables is Markovian. The correlation coefficients depend on d_i.

Side procedures: An empirical procedure, based on the generation of uniform rando, numbers, handles the probability of zero depth, p_0 ; also it adjusts the short interval tail of $F_X(x_{ij})$, when it is Weibull. Moreover, the side procedure handles the number of successive zero rain depths, disallowing exceedance of the value *(c/Δ)*.

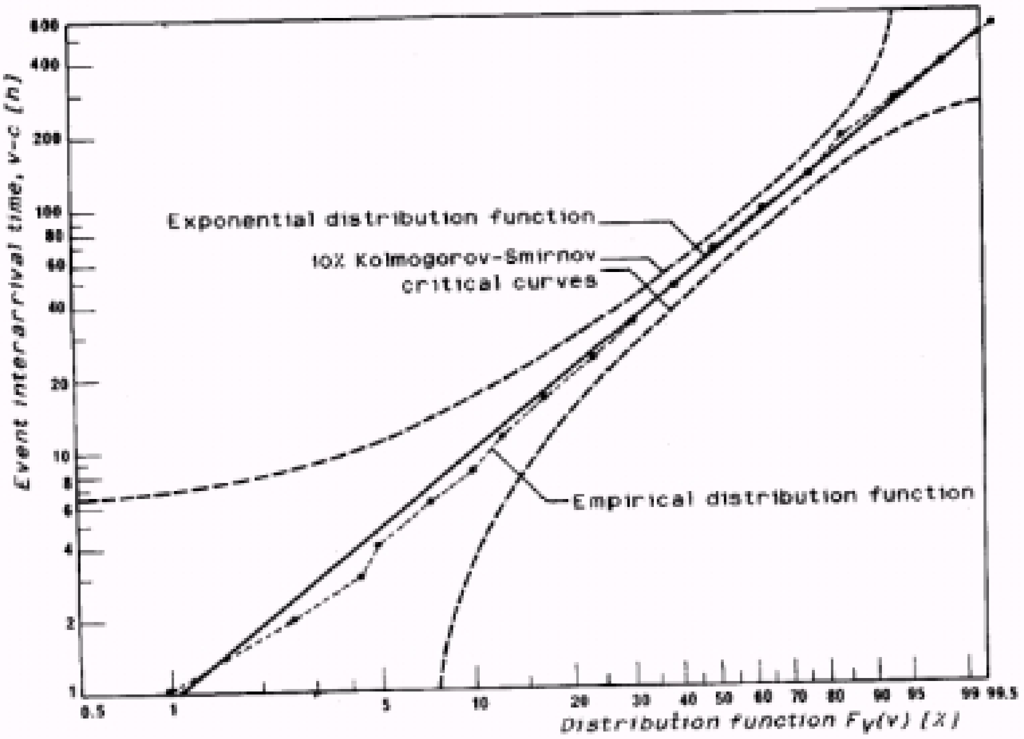


Fig. 1 Distribution function of rainfall event inter-arrival time, V.

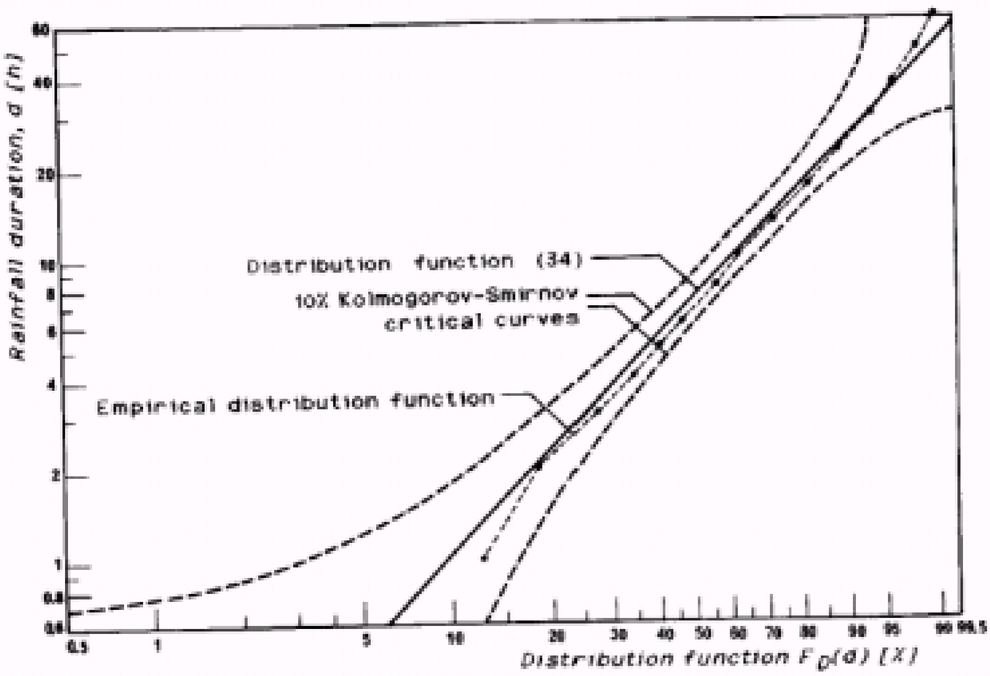


Fig. 2 Distribution function of rainfall event duration, D.

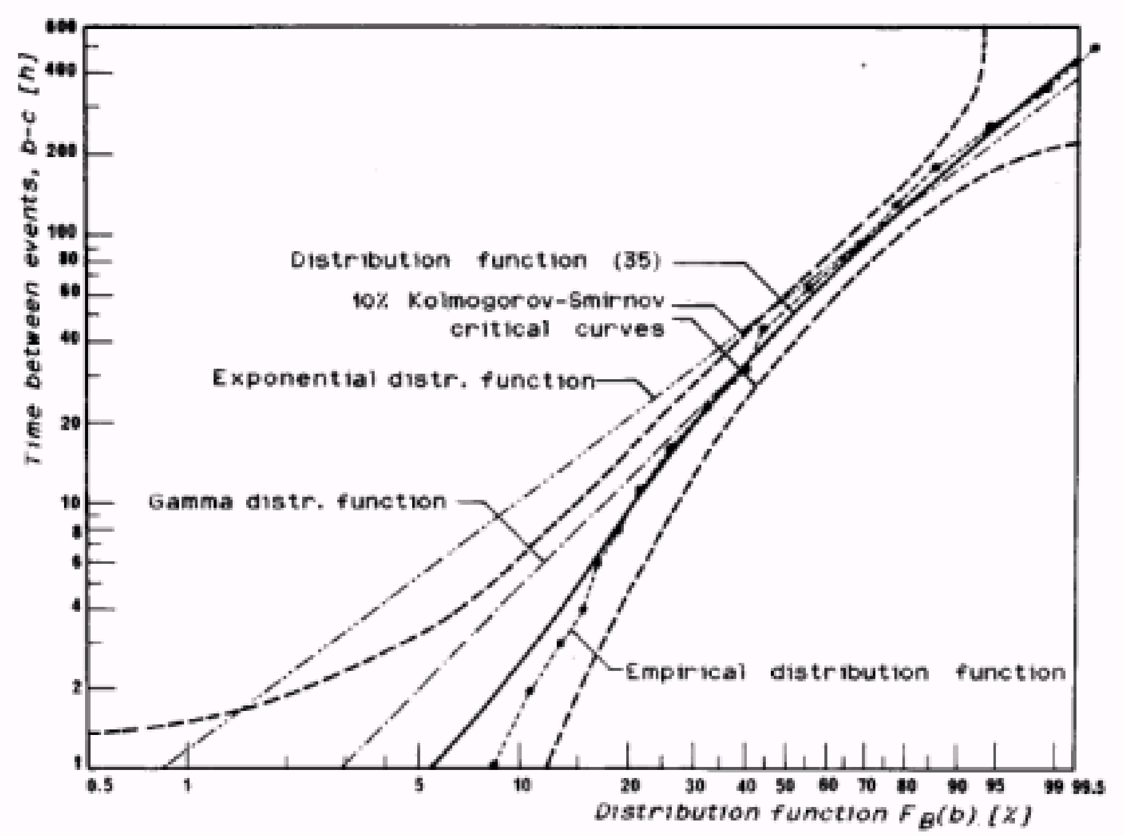


Fig. 3 Distribution function of time between events (dry interval), B.

Simple Disaggregation by Accurate Adjusting Procedures

This method keeps some ideas of the Dynamic Disaggregation Model (DDM) presented in the last section but it is simpler. The method is based on three simple ideas:

- First, it starts using directly a typical PAR(1) model (seasonal AR(1)) and keeps its formalism and parameter set, which is the most parsimonious among linear stochastic models, for generating a time series.
- Second, it uses accurate adjusting procedures to allocate the error in the additive property, i.e. to correct the generated lower-level times so that its terms add up to the corresponding higher-level variables.
- Third, it uses repetitive sampling in order to improve the approximations of statistics that are not explicitly preserved by the adjusting procedures.

There are three different methods of adjustment: the proportional, linear and power method. The proportional method is ideal as it uses a very simple proportional scheme to correct the time series but it is only fully accurate for lower-level variables with gamma distribution incorporating a similar scale parameter as the higher level variables. the linear method is able to cope with any distribution but has the disadvantage of returning negative values. The power method is a combination of the first two methods and is able to perform calculations with the logarithms of statistics, having the disadvantage however, of not being an exact procedure.

The model, owing to the wide range of probability distributions it can handle (from bell-shaped to J-shaped) and to its multivariate framework, is useful for a lot of hydrological applications such as disaggregation of annual rainfall or runoff into monthly or weekly amounts, and disaggregation of event rainfall depths into partial amounts of hourly or even less duration.

The main advantages of the model are the simplicity, parsimony of parameters, and mathematical and computational convenience (due to the simplicity of equations and the reduction in the size of matrices).

The Methodology

We consider a specific higher-level time step or period (e.g. 1 year), denoted with an index t=1, 2, …, and a subdivision of the period in k lower-level time steps or subperiods (e.g. 12 months), each denoted with an index s= 1, …,k. Let a specific hydrologic process with the symbols (rainfall, runoff, etc.) be defined at n locations specified by an index l= 1, …, n. We denote this process with the symbols X and Z for the lower- and higher-level time step, respectively. Generally, we use uppercase letters for random variables, and lowercase letters for values parameters, or constants. Furthermore, we use bold letters for arrays or vectors, and normal letters for their elements. In particular:

${}_tX_s^l$ lower-level variable at period t, sub-period s and location l

${}_tZ^l$ higher-level variable at period t and location l

${}_t\boldsymbol{X}_s$ vector of lower-level variables of sub-period s at all locations n

${}_t\boldsymbol{X}^l$ vector of lower-level variables of all sub-period at location l

${}_t\boldsymbol{Z}$ vector of higher-level variables at all locations n

The vectors ${}_t\boldsymbol{Z}$ ${}_t\boldsymbol{X}_s$ are related by the additive property:

$$\sum_{s=1}^{k} {}_t\boldsymbol{X}_s = {}_t\boldsymbol{Z} \tag{1}$$

The case studied concerns lower-level variables that are related by a contemporaneous seasonal AR(1) (or PAR(1)) model:

$$\boldsymbol{X}_s = \boldsymbol{a}_s\boldsymbol{X}_{s-1} + \boldsymbol{b}_s\boldsymbol{V}_s \tag{2}$$

Where $\boldsymbol{a}_s$ is an ($n\times n$) diagonal matrix; $\boldsymbol{b}_s$ is an ($n\times n$) matrix of coefficients and $\boldsymbol{V}_s = \left[V_s^1,\ldots,V_s^n\right]^T$ is a vector of independent random variates, not necessarily Gaussian.

The parameters that this specific model explicitly preserves are

1. the mean values of lower-lever variables, k vectors (size n) $\boldsymbol{\xi}_s$= E[$\boldsymbol{X}_s$]
2. the variances and the lag-zero cross-covariances of lower-level variables $\boldsymbol{\sigma}_{ss}$=Cov[$\boldsymbol{X}_s$, $\boldsymbol{X}_s$]= E[($\boldsymbol{X}_s$ -$\boldsymbol{\xi}_s$) ($\boldsymbol{X}_s$ -$\boldsymbol{\xi}_s$)^T]
3. the lag-one autocovariances of lower-level variables, k vectors: $Cov\left[X_s^l - X_{s-1}^l\right] = E\left[\left(X_s^l - \xi_s^l\right)\left(X_{s-1}^l - \xi_{s-1}^l\right)\right]$
4. the third moments of lower-level variables $\boldsymbol{\gamma}_s = \left[E\left[\left(X_s^l - \xi_s^l\right)^3\right]\right]$

The total number of second order parameters of this model configuration is *kn(n+3)/2*.

The parameters $\boldsymbol{a}_s$ and $\boldsymbol{b}_s$ are related with the variances, the lag-zero cross-covariances and the lag-one autocovariances of lower-level variables by:

$$a_s^l = Cov\left[X_s^l - X_{s-1}^l\right] / Var\left[X_{s-1}^l\right] = Cov\left[X_s^l - X_{s-1}^l\right] / \sigma_{s-1,s-1}^{ll} \quad (3)$$

$$\boldsymbol{b}_s (\boldsymbol{b}_s)^T = \boldsymbol{\sigma}_{ss} - \boldsymbol{a}_s \boldsymbol{\sigma}_{s-1,s-1} \boldsymbol{a}_s \quad (4)$$

The statistics of the variates V_s^l that are needed for complete determination of $\boldsymbol{X}_s = \boldsymbol{a}_s \boldsymbol{X}_{s-1} + \boldsymbol{b}_s \boldsymbol{V}_s$ are given by:

$$E[\boldsymbol{V}_s] = (\boldsymbol{b}_s)^{-1} \{ E[\boldsymbol{X}_s] - \boldsymbol{a}_s E[\boldsymbol{X}_{s-1}] \} \quad (5)$$

$$\mathrm{Var}\left[V_s^l\right] = 1 \quad (6)$$

$$\mu_3[\boldsymbol{V}_s] = (\boldsymbol{b}_s^{(3)})^{-1} (\gamma_s - \boldsymbol{a}_s^{(3)} \gamma_{s-1}) \quad (7)$$

where $\mu_3[\boldsymbol{V}_s]$ is the vector of third central moments of $\boldsymbol{V}_s$ and the superscript (3) in a matrix indicates that this matrix has to be cubed element by element.

A case alternative to $\boldsymbol{X}_s = \boldsymbol{a}_s \boldsymbol{X}_{s-1} + \boldsymbol{b}_s \boldsymbol{V}_s$ is the PAR(1) for some non-linear transformations of the lower-level variables:

$$\mathbf{X}_s^* = \boldsymbol{a}_s \mathbf{X}_{s-1}^* + \boldsymbol{b}_s \mathbf{V}_s \quad (8)$$

$$\mathbf{X}_s^* := \ln (\boldsymbol{X}_s - \boldsymbol{c}_s) \quad (9)$$

where $\boldsymbol{c}_s$ is a vector of parameters to be estimated in the manner that the $\mathbf{X}_s^*$ are Gaussian.

The models defined by (2) and (8)-(9) are proper for sequential generation of $\boldsymbol{X}_s$ but they do not take account of the known higher-level variables $\boldsymbol{Z}$ and apparently are not disaggregation models. What disaggregation models typically do is to incorporate $\boldsymbol{Z}$ into the model equations and perform a sort of conditional generation using the given values of higher-level variables. The method proposed is much simpler and consists of the following steps:

1. Use (2) or (8)-(9) to generate directly some $\widetilde{\mathbf{X}}_s$ within a period (s=1, …, k) without reference to the given higher-level variables $\boldsymbol{Z}$ of that period.
2. Calculate $\widetilde{\mathbf{Z}} = \sum_{s=1}^{k} \widetilde{\mathbf{X}}_s$ and the distance $\Delta Z = \left\| Z - \widetilde{Z} \right\|$ considering all sites.
3. Repeat steps 1 and 2 until the distance ΔZ is less than an accepted limit, and choose the final set of $\widetilde{\mathbf{X}}_s$ and $\widetilde{\mathbf{Z}}$, which has the minimum distance
4. Apply an adjusting procedure to correct the chosen $\widetilde{\mathbf{X}}_s$
5. Repeat steps 1-4 for all periods that have given $\boldsymbol{Z}$

An adjusting procedure may be viewed as a transformation or function:

$$X_s = f\left(\tilde{X}_s, \tilde{Z}, Z, statistics\right) \tag{10}$$

which given its arguments returns the correct lower-level variables that satisfy (1). The adjustment is done separately in each location.

Accurate Adjusting Procedures

Here are presented three different methods of adjustment: the proportional, linear and power method. First is the proportional procedure that is appropriate for lower-level variables with gamma distributions that satisfy certain constraints. Second is the linear procedure, which is more general, as it can apply to any distribution and it can preserve the first and second moments regardless of the type of distribution. Third is the power procedure, which is a modification of the linear one for positive lower-level variables and also incorporates, as a special case, the proportional procedure.

a. Proportional Adjusting Procedure

(Preservation of Gamma Marginal Distributions)

Proposition 1: Let $\tilde{\mathbf{X}}_s$ be independent variables with gamma distribution functions and parameters κ_s and λ (s=1, …, k).Let also Z be a variable of $\tilde{\mathbf{X}}_s$ with gamma distribution and parameters:

$$\kappa := \sum_{s=1}^{k} \kappa_s \tag{11}$$

and λ. Then the variables :

$$X_s := \frac{\tilde{X}_s}{\sum_{j=1}^{k} \tilde{X}_j} Z \quad s = 1,\ldots,k \tag{12}$$

are independent and have gamma distributions with parameters κ_s and λ. (*Koutsoyiannis 1994*). Note that the variables X_s have the same joint distribution function as their corresponding $\tilde{\mathbf{X}}_s$ and they add up to Z. Also, note that the attained result cannot be extended theoretically for gamma variables with different scale parameters or for normal variables even if they obey quite similar restrictions. This proposition gives rise to the adjusting procedure defined by:

$$X_s = \frac{Z}{\tilde{Z}} \tilde{X}_s \tag{13}$$

where $\widetilde{\mathbf{Z}}$ is the sum of all $\widetilde{\mathbf{X}}_s$ The contribution of proposition 1 is that it reveals a set of conditions that make the procedure exact in a strict mathematical sense. Apparently, this set of conditions introduces severe limitations :

1. the variables X_s should be two parameter gamma distributed
2. the variables X_s should have common scale parameter, i.e. $E[X_s]/Var[X_s]$ should be constant for all sub-periods and
3. all the X_s should be mutually independent

Koutsoyiannis (1994) after an empirical investigation of its performance proposed the following relaxed restrictions in order, for the proportional adjustment, to give satisfactory results, when applied without explicit dependence of lower-to-higher-level variables :

1. the variables X_s have a distribution approaching the two parameter gamma
2. the statistics $E[X_s]/Var[X_s]$ are close to each other for different sub-periods s and,
3. the variables X_s are correlated with $Corr[X_{s-1}, X_s]$ not too large (0.60-0.70)

All these relaxed restrictions are satisfied in the case of short-scale rainfall series, and thus the procedure was used for the disaggregation of total amounts of rainfall events into hourly depths with a stationary AR(1) model.

b. Linear Adjusting Procedure

(Preservation of Second-Order Statistics in the General Case)

Proposition 2 Let $\widetilde{X}_s$ (s=1,…, k) be any random variables with mean values $\boldsymbol{\xi}_s=E[\widetilde{X}_s]$ and variance-covariance matrix $\boldsymbol{\sigma}$ with elements $\sigma_{ij}=Cov[\widetilde{X}_s, \widetilde{X}_j] =E[(\widetilde{X}_s-\xi_s) (\widetilde{X}_{j-1}-\xi_{j-1})]$

Let also Z be a variable independent of $\widetilde{X}_s$ with mean :

$$\xi := E[Z] = \sum_{s=1}^{k} \xi_s \tag{14}$$

and variance

$$\sigma := E[Z] = \sum_{s=1}^{k} \sum_{j=1}^{k} \sigma_{sj} \tag{15}$$

Then the variables

$$X_s := \widetilde{X}_s + \lambda_s \left(Z - \sum_{j=1}^{k} \widetilde{X}_s \right) \quad s = 1,...,k \tag{16}$$

have identical mean values and variance-covariance matix with those of $\widetilde{X}_s$ if λ_s is defined properly:

$$\lambda_s = \sigma_s / \sigma \tag{17}$$

where

$$\sigma_s = \sum_{j=1}^{k} \sigma_{ij} \tag{18}$$

the proposition gives rise to the linear adjusting procedure defined by:

$$X_s := \widetilde{X}_s + \lambda_s \left(Z - \widetilde{Z}\right) \tag{19}$$

It is obvious from (17), (18) and (15) that λ_s always add up to unity, which assures the preservation of the additive property by the procedure

The linear procedure with adjusting coefficients is very general and can be applied regardless of the distribution function or the covariance structure of $\widetilde{X}_s$ It preserves both the means and the variance covariance matrix of lower-level variables. Notably, the procedure involves only linear transformations of the variables. This leads to the preservation of the complete multivariate distribution function of the lower-level variables in single-site problems, if they are Gaussian. In fact, if the higher-level variables Z and the initial low-level variables $\widetilde{X}_s$ have been generated with Gaussian distribution, then the adjusted variables will also have Gaussian distribution. The procedure must be applied for the real higher-and lower –level variables; if transformation are applied to the variables, then the inverse transformation must be applied before the procedure is used.

In this disaggregation scheme, the covariance structure is described by a PAR(1) model. The independent parameters in the PAR(1) model are lag-one covariances only ($\sigma_{s,s-1}$). Any other covariance σ_{sj} for j>s+1 can be computed by combining lag-1 covariances, i.e. using

$$\sigma_{sj} = \frac{\sigma_{s,s+1}\sigma_{s+1,s+2},\ldots,\sigma_{j-1,j}}{\sigma_{s+1,s+1},\ldots,\sigma_{j-1,j-1}} \tag{20}$$

c. power Adjusting Procedure

(Modification for Positive Lower Level Variables)

A weakness of the linear procedure is that it may result in negative values of the lower-level variables, while most hydrological values must be positive. The proportional procedure always results in positive variables, but it is strictly exact only in some special cases. The combination of these two procedures forms the power adjusting procedure which must:

1. result in positive values only
2. be identical to the proportional procedure if the related constraints are satisfied
3. be identical to the linear procedure in some area and preferably in the neighborhood of mean values, i.e. $(X_s, Z, \tilde{Z}) = (\xi_s, \xi, \xi)$

Both procedures, the linear and the proportional, can be written in a common form:

$$\frac{X_s}{\tilde{X}_s} = f_s\left(\frac{Z}{\tilde{X}_s}, \frac{\tilde{Z}}{\tilde{X}_s}\right) \tag{21}$$

for the linear procedure: $f_s(u, w) = u / w$ (22)

for the proportional procedure: $f_s(u, w) = 1 + \lambda_s (u - w)$ (23)

for the power procedure: $f_s(u, w) = u^{\mu_s} / w^{\nu_s}$ (24)

where μ_s and ν_s are parameters to be estimated. This equation always results in positive lower-level values and it is more general than (22). If the lower-level values are independent and

$$\xi_s / \sigma_{ss} = \xi / \sigma \tag{25}$$

then consistency with (22) demands $\mu_s = \nu_s$. To examine the consistency with (23), we linearize (24) in the neighborhood of means, taking the three terms of its Taylor series about the point $(u, w) = (1/\eta_s, 1/\eta_s)$ where $\eta_s = \xi_s/\xi$:

$$\begin{aligned} f_s(u, w) &\approx 1/\eta_s^{\mu_s - \nu_s} + \left(\mu_s / \eta_s^{\mu_s - \nu_s - 1}\right)\left(u - 1/\eta_s\right) - \left(\nu_s / \eta_s^{\mu_s - \nu_s - 1}\right)\left(w - 1/\eta_s\right) \approx \\ &\approx \left(1/\eta_s^{\mu_s - \nu_s}\right)\left(1 - \mu_s + \nu_s + \mu_s \eta u - \nu_s \eta_s w\right) \end{aligned} \tag{26}$$

By comparison with (24) we obtain:

$$\mu_s = \nu_s = \lambda_s / \eta_s \tag{27}$$

For independent lower-level variables satisfying (25), (27) becomes:

$$\mu_s = \nu_s = \frac{\lambda_s}{\eta_s} = \frac{\sigma_s}{\sigma}\frac{\xi}{\xi_s} = 1 \qquad (28)$$

Hence:

$$f_s(u,w) = (u/w)^{\lambda_s/\eta_s} = (u/w)^{\left(\lambda_s/\eta_s\right)} \qquad (30)$$

Finally the power adjusting procedure is:

$$X_s = \tilde{X}_s\left(Z/\tilde{Z}\right)^{\lambda_s/\eta_s} = \tilde{X}_s\left(Z/\tilde{Z}\right)^{(\sigma_s/\sigma)(\xi/\xi_s)} \qquad (31)$$

The adjusting procedure does not preserve the additive property at once. Thus its application must be iterative, until the calculated sum of the lower-level variables are equal to the given Z. Due to the iterative application and the approximations made for its development, the procedure is not exact in strict sense, except for special cases. However the power adjusting procedure may be a useful approximate generalization of the proportional procedure retaining the advantage of returning positive values.

In case X_s and Z have lower bounds c_s and c respectively, (31) can be modified to:

$$X_s - c_s = \left(\tilde{X}_s - c_s\right)\left(\frac{Z-c}{\tilde{Z}-c}\right)^{\lambda_s/\eta_s} \qquad (32)$$

with η_s now defined by $\eta_s = (\xi_s - c_s)/(\xi - c)$.

If we take the logarithms in the (32) we obtain:

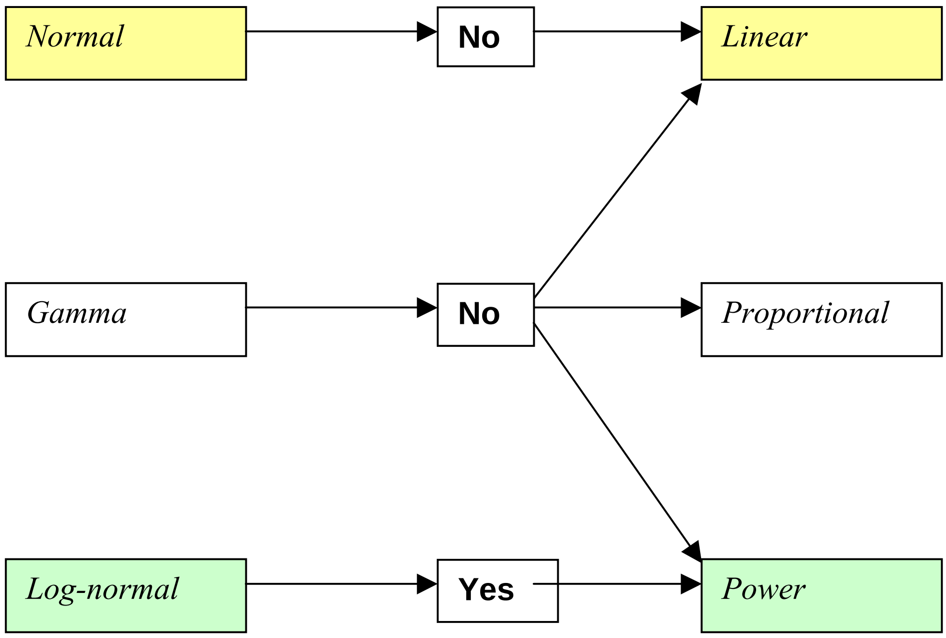
$$X_s^* = \tilde{X}_s^* + \lambda_s^*\left(Z^* - \tilde{Z}^*\right) \qquad (33)$$

$$\begin{aligned} &with: \\ &X_s^* = \ln(X_s - c_s) \\ &Z^* = \ln(Z - c) \\ &\lambda_s^* := \frac{\lambda_s}{\eta_s} = \frac{\sigma_s}{\sigma}\frac{\xi - c}{\xi_s - c_s} \end{aligned} \qquad (34)$$

Equation (33) is the linearized form of the power adjusting procedure.

Selection of the Adjusting Procedure

DISTRIBUTION OF LOWER-LEVEL VARIABLES	TRANSFORMATION OF LOWER-LEVEL VARIABLES	ADJUSTING PROCEDURE



- For normal variables the best choice is the linear procedure, which preserves the entire multivariate normal distribution. If negative values are meaningless, any generated negative value is either rejected (and the generation repeated) or set equal to zero. In the latter case the produced error is corrected by reapplying any of the adjusting procedures; if the linear procedure is chosen for this correction, it will require some iterations. The power adjusting procedure is also iterative and the proportional adjusting procedure performs the correction without iterations.
- For two-parameter or three-parameter gamma distributed variables, all three procedures can be used. More specifically if the constraints of the proportional adjusting procedure are satisfied, then this procedure is the best choice.

Otherwise, the linear adjustment is a good choice when the probability of generating negative values is small, which happens if the coefficients of variation of all variables are relatively low. The power procedure has no limitations and will work in any case.

- For log-normal variables one could follow the same method as in the case of gamma variables and use either the linear or the power procedure. However it is more reasonable to use the logarithmic transformation of the variables and perform the generation of the transformed variables applying the linearized form of the power adjustment (equation (33)).

Repetition scheme

The preservation of the skewness of the variables and of the cross-correlation coefficients among them is not assured by the adjusting procedures viewed, apart from very special cases. In fact, in the general case these procedures give some approximations of these statistics, which tend to be lower than the correct values and may not be adequate. However the approximation of these statistics can be improved by repetition. The idea behind the repetition is that of conditional sampling. The disaggregation may be viewed as the problem of generating lower-level variables $\boldsymbol{X_s}$ conditional on the given higher-level variables **Z** such that: $\sum_{s=1}^{k} X_s = Z$.

The method proposed by *Koutsoyiannis* is the repetitive generation of $\boldsymbol{X_s}$ from their unconditional joint distribution function until the error at the additive property is small, and then apply an adjusting procedure to allocate this error among the different sub-periods. For each period an unspecified number of generations through the PAR(1) model is performed, until the distance $\Delta Z = \left\| Z - \widetilde{Z} \right\|$ drops below a specified tolerance level. The lower the tolerance is the greater is the accuracy of the results and, also, the number of repetitions. To avoid the huge number of repetitions, it is advisable to set a maximum allowed number of repetitions.

The distance, in a dimensionless and independent of the number of locations form, is given by:

$$\Delta Z = \frac{1}{n} \sum_{l=1}^{n} \frac{\left| Z^l - \widetilde{Z}^l \right|}{\sqrt{Var\left[Z^l\right]}}$$

Repetition and minimization of the distance ΔZ helps preserving also the correlation coefficients.

CHAPTER II

RAINFALL DISAGGREGATION USING ADJUSTING PROCEDURES ON A POISSON CLUSTER MODEL

This disaggregation methodology concerns the combination of a rainfall simulation model based upon the Bartlett-Lewis process with techniques that aim to the adjustment of the finer scale (hourly) values generated by the rainfall model so as to be consistent with the coarser scale (daily) values, given. The adjusting procedures proposed are the proportional, linear and power procedure studied by *Koutsoyiannis* and *Manetas* (1996) and discussed in detail in the previous section.

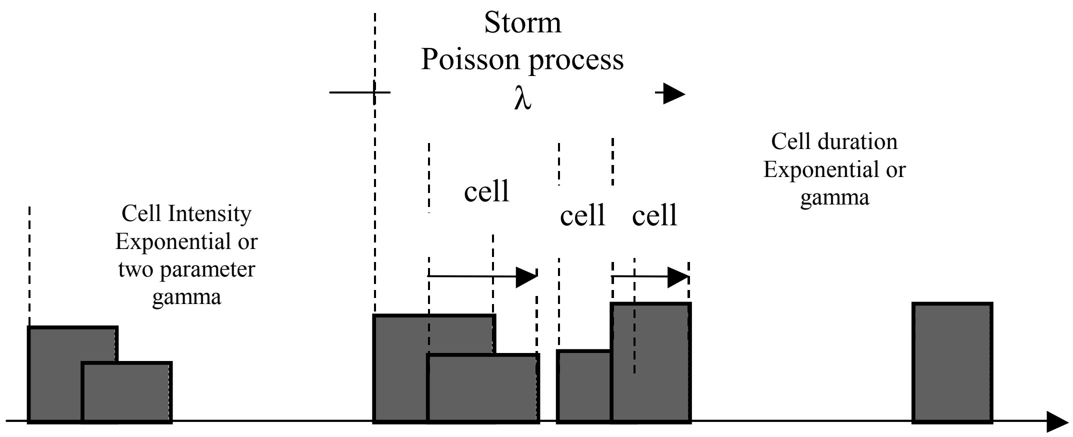
The Rainfall Model

The rainfall model used is the Bartlett-Lewis Rectangular Pulse Model (BL) and was chosen due to its wide applicability and experience in calibrating and applying it to several climates and to its ability on reproducing important features of the rainfall field from the hourly to the daily scale. The BL is a model that represents rainfall at a point in continuous time. Therefore it is particularly useful in a disaggregation framework where it may be used at a time-step different from that at which it is fitted. The model incorporates Poisson cluster processes, i.e. storms arrive according to a Poisson distribution and are constituted by clusters of cells or rectangular pulses with constant depth. The general assumptions of the model are:

1. Storm arrivals follow a Poisson process with rate λ;
2. cells arrivals follow a Poisson process with rate β;
3. Cell arrivals of each storm terminate after a time exponentially distributed with parameter γ;
4. Each cell has a duration exponentially distributed with parameter η; and
5. Each cell has a uniform intensity with a specified distribution.

In the original version of the model, all model parameters are assumed constant. In the modified version, the parameter η is randomly varied from storm to storm with a gamma distribution with shape parameter α and scale parameter ν. Subsequently, parameters β and γ also vary so that the ratios $\kappa := \beta / \eta$ and $\phi := \gamma / \eta$ are constant.

The distribution of the uniform intensity is typically assumed exponential with parameter 1 / μ_X. Alternatively, it can be chosen as two-parameter gamma with mean μ_X and standard deviation σ_X. Thus, in its most simplified version the model uses five parameters, namely λ, β, γ, η, and μ_X (or equivalently, λ, κ, ϕ, η, and μ_X) and its most enriched version seven parameters, namely λ, κ, ϕ, α, ν, μ_X and σ_X.



The adjusting procedures

The synthetic time series of the lower-level generated by the Bartlett-Lewis Rectangular Pulse model must be “modified” so as to be consistent with the higher-level time series given and simultaneously not affect the stochastic structure implied by the model.

Provided that a data series Z_p (p = 1, 2, …) is known at a daily time scale and an hourly synthetic series $\tilde{X}_s$ (s = 1, 2, …) has been generated by the Bartlett-Lewis model, disaggregation by adjusting procedures is a methodology to modify the lower-level series (thus getting a modified series X_s, s = 1, 2, …) so as to make it consistent with the higher-level one. To achieve this, it uses accurate adjusting procedures to allocate the error in the additive property, i.e., the departure of the sum of lower-level variables within a period from the corresponding higher-level variable. These procedures are accurate in the sense that they preserve explicitly (at least under some specified conditions) certain statistics or even the complete distribution of lower-level variables. In addition, the methodology uses repetitive sampling in order to improve the approximations of statistics that are not explicitly preserved by the adjusting procedures.

Three such adjusting procedures have been developed and studied (Koutsoyiannis, 1994; Koutsoyiannis and Manetas, 1996) and are summarily described below:

1. Proportional adjusting procedure

This procedure modifies the initially generated values $\widetilde{X}_s$ to get the adjusted values X_s according to

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

where Z is the higher-level variable and k is the number of lower-level variables within one higher-level period.

The proportional adjusting procedure gives exact, in a strict mathematical sense, results, only if the variables X_s are two parameter gamma distributed, have common scale parameter and are mutually independent. Nevertheless as shown by *Koutsoyiannis (1994)*, this procedure gives satisfactory results when applied to variables that have distributions approaching the two-parameter gamma, with scale parameters not necessarily common but close to each other for different sub-periods and with mutual correlation being not very high.

2. Linear adjusting procedure

The linear adjusting procedure modifies the initially generated values $\widetilde{X}_s$ to get the adjusted values X_s according to

$$X_s = \widetilde{X}_s + \lambda_s \left(Z - \sum_{j=1}^{k} \widetilde{X}_j \right) \qquad (s = 1, \ldots, k) \tag{2}$$

where λ_s are unique coefficients depending on the covariances of X_s with Z. This procedure is very general and can be applied regardless of the distribution function or the covariance structure of the generated values. It preserves both the means and the variance-covariance matrix of lower-level variables. In addition, in single-site problems with variables having a Gaussian distribution function, the procedure results in complete preservation of the multivariate distribution function of the lower-level variables. The main disadvantage of this technique is that it may result in negative values. In fact, if the term in parenthesis (equ. 2) is negative, all zero values become negative after the adjustment.

3. Power adjusting procedure

The power adjusting procedure modifies the initially generated values $\widetilde{X}_s$ to get the adjusted values X_s according to

$$X_s = \widetilde{X}_s \left(Z / \sum_{j=1}^{k} \widetilde{X}_j \right)^{\lambda_s / \eta_s} \qquad (s = 1,...,k) \tag{3}$$

where λ_s are appropriate coefficients depending on the covariances of X_s with Z and η_s are coefficients depending on the mean values of X_s and Z. This procedure is a combination of the proportional and linear methods and returns always positive values and it is also able to perform calculations with the logarithms of the statistics. This method does not preserve the additive property so the adjustment must be iterative. As well known, the iterative application requires many approximations and this makes the power adjusting procedure not fully exact in a strict mathematical sense, apart from very special cases.

Selection of the Adjusting Procedure

The choice of the appropriate adjusting procedure is conditional to the characteristics of the disaggregation problem:

The rainfall disaggregation problem at a short-scale is characterized by a large proportion of zero values, (an hourly series is formed by long sequences of zero values corresponding to dry intervals and relatively brief sequences of positive values that correspond to the rainy intervals).

In addition the stochastic structure of the rainfall model assumes that the rainfall process is stationary within a specific period (e.g. month) and that the rainfall depths in rainy intervals are approximately gamma distributed.

Under these conditions it can be easily observed that the linear adjusting procedure is inappropriate due to its tendency of returning negative values when the number of zero entries is high. Moreover it can be easily demonstrated that due to the stationarity of the rainfall process the power procedure becomes identical to the proportional one. In fact, the stationarity implies that the exponent λ_s/η_s in the power adjusting equation equals 1 making equations (1) and (3) identical (for the proof of this assumption see section 2 of this chapter). All these conditions in addition with the already mentioned gamma distribution of the rainfall depths makes the proportional adjusting procedure the most adequate for the examined rainfall disaggregation problem.

However, the proportional adjusting procedure, as well as any other adjusting method, does not explicitly preserve the skewness of the variables nor the cross-correlations among variables of different locations, instead it gives some approximations of these statistics. In order to improve these approximations a repetition scheme is required. This means that instead of running the generation routine of the rainfall model once for a certain period, it is run several times until the sequence of generated values reproduces the higher-order statistics that resemble the actual data available.

Coupling of the Bartlett-Lewis model with the adjusting procedure

The coupling of the Bartlett-Lewis rectangular pulse model with the adjusting procedures contains several problems.

First of all, the Bartlett-Lewis, models the rainfall process as continuous in time while the disaggregation operates on discrete time through two different time scales, the higher-level (e.g., daily) and lower-level (e.g., hourly) ones. The storms and cells generated by the Bartlett-Lewis model may lie on more than one higher- or lower-level time steps. Therefore, the application of the adjusting procedure on these storms and cells must extend to more than one day. In order to optimize the computational time of the adjusting and repetitions schemes especially in cases of long simulation periods, the simulation period is divided in sub-periods. This is possible due to the fact that different clusters of wet days separated by at least one dry day, can be assumed independent and therefore they can be handled separately. This assumption is in complete agreement with the stochastic structure of the BL model in which the storm arrivals occur accordingly to a Poisson process. Thus, the Bartlett-Lewis model runs separately for each cluster of wet days. Several runs are performed for each cluster, until the departures of the sequence of daily sums from the given sequence of daily rainfall becomes lower than an acceptable limit.

The whole disaggregation methodology is implemented in a computer program produced by *Koutsoyiannis and Onof* and described in detail below.

The implementation of the methodology : HYETOS Υετός

A computer program for temporal rainfall disaggregation using adjusting procedures.

The computer program Hyetos, produced by *Koutsoyiannis and Onof* (2000), is a software that performs temporal disaggregation of daily rainfall into hourly rainfall by adjusting procedures.

The method used is a process that can be represented by the following three steps:

Generate a time series according to an appropriate stochastic rainfall model.

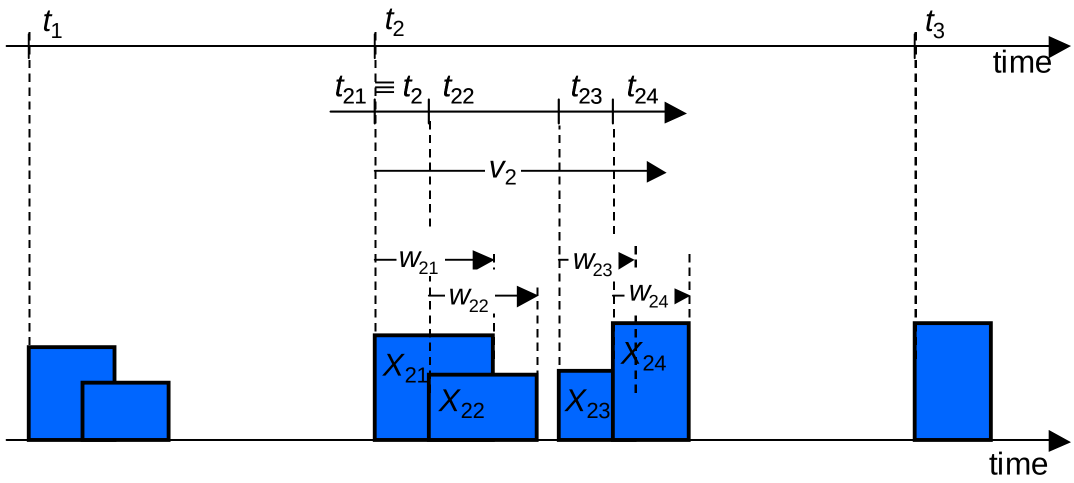
Use an accurate adjusting procedure to correct the generated lower-level time series so that its terms add up to the corresponding higher-level variables.

Repeat the process until a suitable time series can be obtained in order to improve the statistics that are not explicitly preserved in the adjusting procedure (e.g. skewness, cross-correlation coefficients.)

Hyetos follows exactly this process, it uses the Bartlett-Lewis rainfall model as a background stochastic model for rainfall generation. Then it uses repetition to derive a synthetic rainfall series, which resembles the given series at the daily scale, and, subsequently, an appropriate adjusting procedure, namely the proportional adjusting procedure to make the generated hourly series fully consistent with the given daily series.

In this program the user is required to enter in the parameters from the Bartlett-Lewis Rectangular Pulse Model. The actual historical rainfall time series can also be entered, so that the disaggregated and historical statistics can be compared. As an output, the program gives the fully calculated statistics of the hourly time series, as well as the simulated time series obtained. Statistics are calculated for wet and dry periods as well as the whole time period.

The stochastic rainfall model used in Hyetos : BARTLETT-LEWIS Rectangular Pulse Model



The general assumptions of the Bartlett-Lewis Rectangular Pulse model (Rodriguez-Iturbe Et Al., 1987, 1988; Onof and Wheater, 1993)are (see figure):

1. Storm origins t_i occur following a Poisson process (rate λ)
2. Cell origins t_{ij} arrive following a Poisson process (rate β)
3. Cell arrivals terminate after a time v_i exponentially distributed (parameter γ)
4. Each cell has a duration w_{ij} exponentially distributed (parameter η)
5. Each cell has a uniform intensity X_{ij} with a specified distribution

In the original version of the model, all model parameters are assumed constant. In the modified version, the parameter η is randomly varied from storm to storm with a gamma distribution with shape parameter α and scale parameter v. Subsequently, parameters β and γ also vary in a manner that the ratios $\kappa := \beta / \eta$ and $\varphi := \gamma / \eta$ be constant.

The distribution of the uniform intensity X_{ij} is typically assumed exponential with parameter $1 / \mu_X$. Alternatively, it can be assumed two-parameter gamma with mean μ_X and standard deviation σ_X.

Thus, in its most simplified version the model uses five parameters, namely λ, β, γ, η, and μ_X (or equivalently, λ, κ, φ, η, and μ_X) and its most enriched version seven parameters, namely λ, κ, φ, α, v, μ_X and σ_X.

Hyetos supports both the original and the modified model version with exponential or gamma intensities. The current software version does not support estimation of the Bartlett-Lewis model parameters, which should be estimated using other software.

The Adjusting Procedure used by Hyetos

After generating the synthetic rainfall series according to the BLRPM, Hyetos uses the proportional adjusting procedure to modify the generated series so as to be consistent with the given daily totals.

The proportional adjusting procedure is given by:

$$X_s = \widetilde{X}_s \left(Z / \sum_{j=1}^{k} \widetilde{X}_j \right) \qquad (s = 1,...,k)$$

Where $\widetilde{X}_s$ are the initially generated values ,X_s the adjusted values, Z is the higher-level variable and k is the number of lower-level variables within one higher-level period.

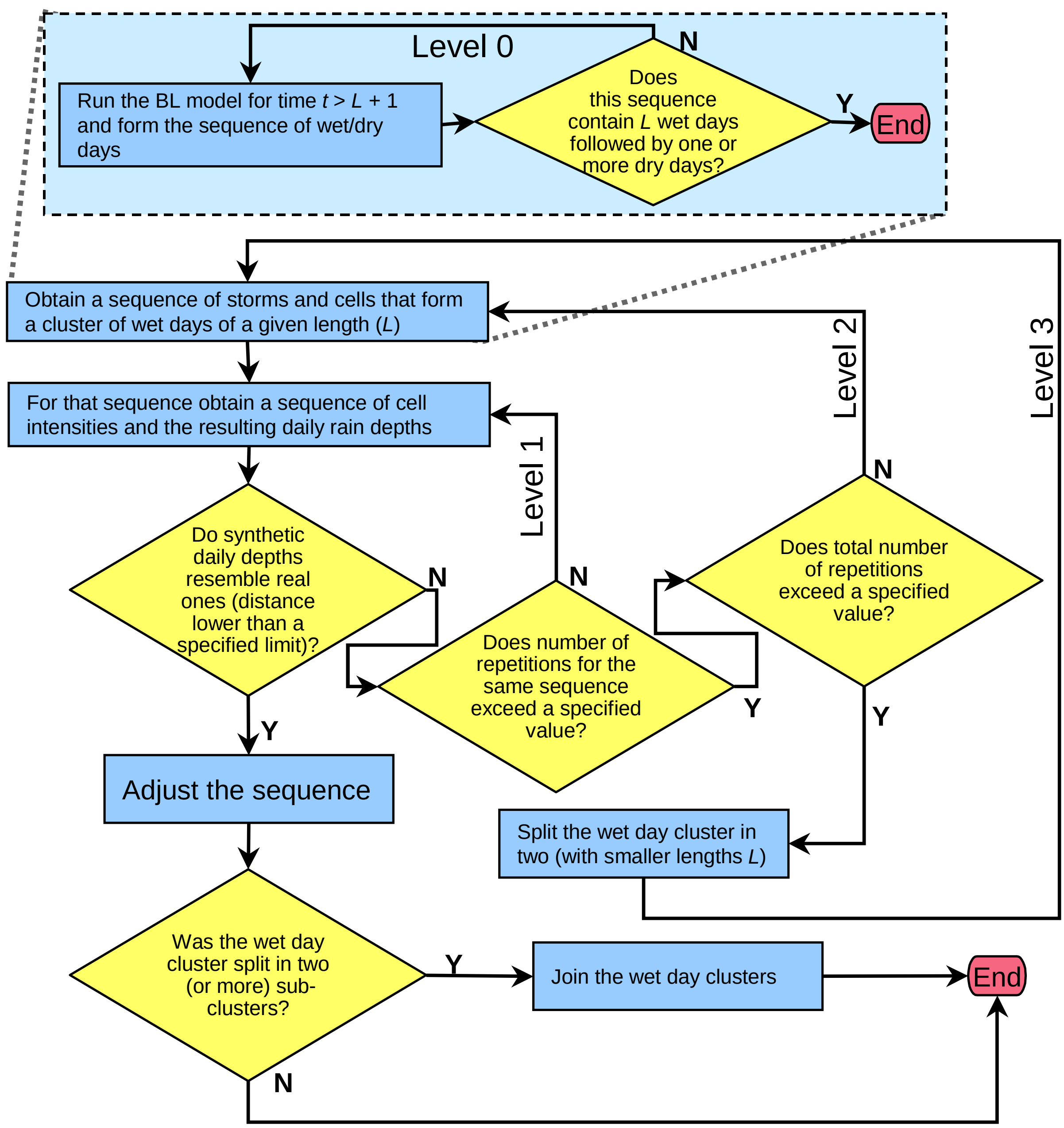
This procedure is exact given two conditions: the lower-level series must be gamma distributed with common scale factor as the higher-level variables. Therefore, Hyetos implicitly assumes that the rainfall is distributed according to a gamma distribution and under the condition of a stationary process within each month.

The proportional adjusting procedure was chosen due to its simplicity on application and also because it does not return negative values.

The Repetition Scheme

In order to handle the problems that arise in coupling the Bartlett-Lewis with the adjusting procedure Hyetos operates according to a scheme which incorporates four levels of repetitions. Repetition is used in the program to improve the statistics that are not explicitly reproduced by adjustment. In the process of repetition, a further refinement is made by considering a term that represents the difference between the generated higher level values and the actual higher level values, and is calculated for every repetition of the time series. Once this term reaches a minimum threshold the repetition stops and the adjusting procedure is applied. This results in preservation of the higher order statistics. The number of repetitions must be defined by the user, and if the limit of repetitions has been met, the model will try to use other methods to model the rainstorm. If an especially long storm is encountered, the model will randomly divide the storm in several portions, using a separate model to represent each portion. Hyetos then enters into a higher level of repetitions.

The repetition and disaggregation scheme, 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) is depicted in the following figure:



. 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 the logarithmic distance

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

where Z_i and $\tilde{Z}_i$ are the original and generated, respectively, daily depths of day i of the wet day sequence and c a small constant (= 0.1 mm). The logarithmic transformation is selected to avoid domination by the very high values and the constant c was inserted to avoid domination by the very low values. 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 (e.g., hourly) rainfall depths through the application of the proportional adjusting procedure.

Modes of operation

Hyetos can perform in each of the following modes depending on the user selections:

1. Disaggregation test mode (without input; **default mode**). An initial sequence of storms is generated using the Bartlett-Lewis model with the given parameters and then aggregated into hourly and daily scale. The daily sequence serves then as an "original" series, which is disaggregated, thus producing another synthetic hourly series. This mode is appropriate for testing the disaggregation model itself (e.g. by comparing original and disaggregated statistics

2. Full test mode (with hourly input).In this mode, an input file with the appropriate format containing **hourly historical data** must be available. The difference from the **Disaggregation test mode** is that the daily sequence is read from the file rather than generated. This mode is appropriate for testing (e.g. by comparing original and disaggregated statistics) the entire model performance including the appropriateness of the Bartlett-Lewis model and its parameters and the disaggregation model.

3. Operational mode (with daily input). This is similar to **Full test mode** the difference being that the input file contains no hourly data but only **daily**. This is the usual case for the model application. It cannot provide any means for testing.

4. Rainfall model test mode (with hourly input). This is similar to the **Full test mode** but with synthetic data not disaggregated but generated from the Bartlett-Lewis model with the given parameters. This mode is appropriate for testing whether the Bartlett-Lewis model fits the historical data (in terms of several statistics).

5. Simple rainfall generation mode (without input and without disaggregation). This is similar to the **Rainfall model test mode** but with no input provided (and consequently, no input file defined). This mode is appropriate for generation of rainfall series using the Bartlett-Lewis model with the given parameters without performing any disaggregation.

In all modes the Bartlett-Lewis model can be implemented either in its original or modified version with a number of parameters from 5 to 7.

The output of the program can be in both text and graphical form. Graphs showing comparisons between the historical and simulated statistics are drawn up, detailing the skewness and the proportion of wet periods, as well as the autocorrelation for lag n period.

CHAPTER III

MULTIVARIATE RAINFALL DISAGGREGATION AT A FINE SCALE

The disaggregation models proposed by *Koutsoyiannis et al*. and described in the previous sections, as well as most of the studies conducted by other authors on the problem of disaggregation, have the common characteristic of being univariate, i.e. they perform single-site temporal disaggregation. As discussed earlier, in all these approaches, the actual rainfall series is not known but its stochastic structure can be hypothesized according to a specific rainfall model . So, they are able to generate synthetic hourly series, fully consistent with the known daily series and simultaneously, statistically consistent with the actual hourly rainfall depths. In other words, the synthetic hourly series generated adds up to the given daily totals and agrees with the stochastic structure implied by the model. However, these synthetic series resemble the actual ones only in a stochastic sense, i.e. they represent a likely realization of the process and obviously do not coincide with the actual one.

In this section we describe a different approach to rainfall disaggregation. We investigate the possibility of generating spatially and temporally consistent hourly rainfall series at a raingage by using the available, at a neighbouring raingage, hourly information in the case that the cross-correlation between the two raingages is significant. This can be considered as a particular case of a general multivariate spatial-temporal rainfall disaggregation problem, i.e. the simultaneous rainfall disaggregation at several sites. Although there is substantial experience in multi-site disaggregation of rainfall from annual to monthly time scales and in single-site disaggregation of rainfall at finer scales, this multivariate fine-time-scale rainfall disaggregation problem has not been studied so far in the rainfall modeling literature. Multiple sites imply mathematical complexity and in contrast with the single-site problems another parameter is considered, the cross-correlation, that must be maintained. The problem involves the combination of several univariate and multivariate rainfall models operating at different time scales, in a framework that was studied by *Koutsoyiannis* [2001].

The Problem of spatial and temporal rainfall disaggregation at several sites

Before describing the methodology for facing the problem, we first need to standardize the problem itself. The description that follows is a more accurate and particularized formulation of the question arisen, that is whether we could utilize the available single-site fine scale information, in conjunction with the daily data, to generate spatially consistent rainfall series. We can consider two apparently different cases, the second being however the generalization of the first.

CASE 1

We assume that we are given:

1. an hourly point rainfall series at point 1, as a result of either:
    - measurement by an autographic device (pluviograph) or digital sensor,
    - simulation with a fine time scale point rainfall model such as a point process model (Bartlett-Lewis Rectangular Pulse model)
    - simulation with a temporal point rainfall disaggregation model applied to a series of known daily rainfall; (coupling of BLRPM with adjusting procedures)
2. several daily point rainfall series at neighbouring points 2, 3, 4, 5, … as a result of either:
    - measurement by conventional raingages (pluviometers with daily observations), or
    - simulation with a multivariate daily rainfall model.

We wish to produce series of hourly rainfall at points 2, 3, 4, 5, …, so that:

1. their daily totals equal the given daily values; and
2. their stochastic structure resembles that implied by the available historical data.

We emphasize that in this problem formulation we always have an hourly rainfall series at one location, which guides the generation of hourly rainfall series at other locations. If this hourly series is not available from measurements, it can be generated using appropriate univariate simulation models.

The essential statistics that we wish to preserve in the generated hourly series are:

1. the means, variances and coefficients of skewness;
2. the temporal correlation structure (autocorrelations);

3. the spatial correlation structure (lag zero cross-correlations); and
4. the proportions of dry intervals.

If the hourly data set at location 1 is available from measurement, then all these statistics apart from the cross-correlation coefficients can be estimated at the hourly time scale using this hourly record. To transfer these parameters to other locations, spatial stationarity of the process can be assumed.

The stationarity hypothesis may seem an oversimplification at first glance. However, it is not a problem in practice since possible spatial nonstationarities manifest themselves in the available daily series; thus the final hourly series, which are forced to respect the observed daily totals, will reflect these nonstationarities.

CASE 2

If hourly rainfall is available at several (more than one) locations, the same modeling strategy described above can be used without any difficulty with some generalizations of the computational algorithm. In fact, having more than one point with known hourly information would be advantageous for two reasons. First, it allows a more accurate estimation of the spatial correlation of hourly rainfall depths (see discussion below) or their transformations. Second, it might reduce the residual variance of the rainfall process at each site, thus allowing for generated hyetographs closer to the real ones. If more than one rainfall series are available at the hourly level, at least one cross-correlation coefficient of hourly rainfall can be estimated directly from these series. Then, by making plausible assumptions about the spatial dependence of the rainfall field an expression of the relationship between cross-correlation could be established.

Modeling Approach

The methodology proposed involves three categories of models :

1. models for the generation of multivariate fine-scale outputs
2. models associated with inputs
3. models associated with spatial-temporal parameters

models for the generation of multivariate fine-scale outputs

This category includes two models that are used in the generation phase in order to provide the required output, the hourly series.

The first model is a simplified multivariate model of hourly rainfall that is able to preserve the essential statistics of the multivariate rainfall process and, simultaneously, incorporate the available hourly information without any reference to the known daily totals at the other sites of interest. The essential statistics considered here are the means, variances and coefficients of skewness, the lag one autocorrelation coefficients and the lag zero cross correlation coefficients.

The second model used in the generation phase is the transformation model that is applied to the synthetic series generated by the simplified model modifying them so as to be consistent to the daily rainfall series, i.e. so that the daily totals are equal to the given ones.

Therefore the models are applied separately and more specifically, assuming that the sites in which we need to generate the hourly series are close to the site where hourly information is available and highly spatially correlated, then the given hourly series known can be used, with the simplified multivariate model, to:

- guide the generation of the hourly series at the sites with daily data, and act indirectly to preserve properties not modelled explicitly;
- properly locate the rainfall events in time
- produce initial hourly rainfall series at the daily sites, whose departures from the actual hourly depths at those sites are not large (even though daily totals are not considered at all at this stage).

The transformation model is then applied to adjust the synthetic series, this being possible because at this stage another source of information is additionally incorporated, that is the multi-site information. The modification of the generated series is essential for the preservation of the properties of the rainfall process which are not captured by the statistics considered in assuming a simplified rainfall model, i.e. nonstationarities of the rainfall field in space and time.

models associated with inputs

This second category contains models which may optionally be used to provide the required input, should no observed series be available. These may include

- a multivariate daily rainfall model for providing daily rainfall depths, such as the general linear model (GLM) [*Chandler and Wheater*, 1998a, b];
- a single-site model for providing hourly depths at one location such as the Bartlett-Lewis rectangular pulses model [*Rodriguez-Iturbe et al.*, 1987, 1988; *Onof and Wheater*, 1993, 1994];
- a single-site disaggregation model to disaggregate daily depths of one location into hourly depths [e.g. *Koutsoyiannis and Onof*, 2000, 2001].

Such models may be appropriate to operate the proposed disaggregation approach for future climate scenarios.

models associated with Spatial-temporal parameters

The third category includes models which play an auxiliary role in the disaggregation framework by providing some of the required parameters of the spatial-temporal rainfall process given the statistical properties of the available data.

In the case where the available hourly information concerns only one site some of the statistics at the hourly level (such as cross-correlation coefficients between raingages) required for the disaggregation cannot be estimated directly and must be inferred using a spatial-temporal rainfall model.

The estimation of the statistics of hourly level is made in the following manner:

1. The temporal and spatial correlations at the daily level are estimated using the daily data sets; in addition, if an hourly series is available, its marginal statistics are estimated.
2. The parameters of the spatial-temporal rainfall model are estimated by fitting the spatial-temporal rainfall model using the historical statistics estimated at point 1.
3. The remaining statistics that are required for disaggregation (e.g. spatial correlations at the hourly level) are inferred from the spatial-temporal rainfall model.

An example of a spatial-temporal model which may be used to provide the stochastic structure of the hourly rainfall is the GDSTM [Gaussian Displacement Spatial-Temporal rainfall Model, *Northrop*, 1996, 1998]. This model is a spatial analogue of the Bartlett-Lewis rectangular pulses model which is the point process model assumed in the single-site problems. According to the GDSTM, rainfall is realized as a sequence of storms that arrive following a Poisson process in space and time and each storm consists of a number of cells. Both storms and cells are characterized by their centers, durations and areal extents, and in addition cells have uniform rainfall intensity. The independent parameters necessary for the definition of the model are 11 and all the statistics to be estimated must be expressed in terms of these parameters. But, in practice the computational costs involved are relatively high and in addition the performance of the GDSTM was proven not to be so accurate. Therefore the use of another model would be more than justified.

If hourly rainfall is available at several (more than one) locations, the estimation of the spatial correlation of hourly rainfall depths can be more accurate. If more than one rainfall series are available at the hourly level, at least one cross-correlation coefficient of hourly rainfall can be estimated directly from these series. Then, by making plausible assumptions about the spatial dependence of the rainfall field an expression of the relationship between cross-correlation could be established.

An example of such an empirical expression is : $r_{ij}^{h} = \left(r_{ij}^{d}\right)^{m}$

Where:

r_{ij}^{h} is the cross-correlation coefficient between raingages i and j at the hourly time scale

r_{ij}^{d} is the cross-correlation coefficient between raingages i and j at the daily time scale

m is an exponent that can be estimated by regression using the known cross-correlation coefficients at the hourly and daily time scale or, in case no hourly data is available, its value can be assumed approximately in the range 2 to 3.

Generation Phase

As already mentioned in this phase two separate models are used, the simplified multivariate model which generates the synthetic hourly series and the transformation model which modifies these synthetic series in order to be fully consistent with the given daily series. The choice of a simplified multivariate rainfall model for producing the required outputs is justified by the following reasons:

1. A spatial-temporal rainfall model (such as GDSTM) would describe the rainfall process in continuous space and time but has not the appropriate structure to utilize the given hourly information.
2. assuming the simplified model means assuming a more parameter parsimonious approach which provides an efficient model that can be

indirectly corrected (by applying the transformation model which incorporates the multi-site daily information).

3. the implications given by assuming such a simplified model for the multivariate rainfall process are counterbalanced due to the additional source of information incorporated as model input (the hourly information).

The Simplified Multivariate model

This model preserves the statistics of the multivariate rainfall process and, simultaneously, incorporates the available hourly information at site 1, without any reference to the known daily totals at the other sites. The statistics considered here are the means, variances and coefficients of skewness, the lag-one autocorrelation coefficients and the lag-zero cross-correlation coefficients. All these represent statistical moments of the multivariate process. The proportion of dry intervals, although considered as one of the parameters to be preserved, is difficult to incorporate explicitly. However, it can be treated by an indirect manner.

For n locations,we may assume that the simplified multivariate rainfall model is an AR(1) process, expressed by

$$\mathbf{X}_s = \mathbf{a}\,\mathbf{X}_{s-1} + \mathbf{b}\,\mathbf{V}_s \tag{1}$$

where $\mathbf{X}_s := [X_s^1, X_s^2, \ldots, X_s^n]^T$ represents the hourly rainfall at time (hour) s at n locations, $\mathbf{a}$ and $\mathbf{b}$ are $(n \times n)$ matrices of parameters and $\mathbf{V}_s$ ($s = \ldots, 0, 1, 2, \ldots$) is an independent identically distributed (iid) sequence of size n vectors of innovation random variables (so that the innovations are both spatially and temporally independent). The time index s can take any integer value. $\mathbf{X}_s$ are not necessarily standardized to have zero mean and unit standard deviation, and obviously they are not normally distributed. On the contrary, their distributions are very skewed. The distributions of $\mathbf{V}_s$ are assumed three-parameter Gamma.

Alternatively, the model can be expressed in terms of some nonlinear transformations X_s^* of the hourly depths $\mathbf{X}_s$, in which case (1) is replaced by

$$X_s^* = aX_{s-1}^* + bV_s \tag{2}$$

Equations to estimate the model parameters **a** and **b** and the moments of $\mathbf{V}_s$ are given for instance by *Koutsoyiannis* [1999] for the most general case. For convenience, the parameter matrix **a** is assumed diagonal, which suffices to preserve the statistics The parameter matrix **b** is assumed lower triangular, which facilitates handling of the known hourly rainfall at site 1. The first row **b** will have only one nonzero item, call it b^1, so that from (1)

$$X_s^1 = a^1 X_{s-1}^1 + b^1 V_s^1 \tag{3}$$

which can be utilized to determine (rather than to generate) V_s^1, given the series of. X_s^1. This can be directly expanded to the case where several gages of hourly information are available provided that **b** is lower triangular.

The Transformation model

This is the model that modifies the series generated by the simplified multivariate model, so that the daily totals are equal to the given ones.

In the previous sections we have examined in detail transformation techniques that are able to modify a series generated by any stochastic process to satisfy some additive property (i.e. the sum of the values of a number of consecutive variables be equal to a given amount), without affecting the first and second order properties of the process, in the case of univariate problems. These techniques commonly known as adjusting procedures, are specialized, as already discussed, for the single-site disaggregation and have been studied previously by *Koutsoyiannis* [1994] and *Koutsoyiannis and Manetas* [1996] Nevertheless, they can be applied for multivariate problems, but in a repetition framework.

More recently, *Koutsoyiannis* [2001] has studied a true multivariate transformation of this type and also proposed a generalized framework for coupling stochastic models at different time scales.

This framework, specialized for the problem examined here, is depicted in the following *schematic representation* where $\mathbf{X}_s$ and $\mathbf{Z}_p$ represent the "actual" hourly- and daily-level processes, related by

$$\sum_{s=(p-1)k+1}^{pk} X_s = Z_p \tag{4}$$

where k is the number of fine-scale time steps within each coarse-scale time step (24 for the current application), and $\tilde{X}_s$ and $\tilde{Z}_p$ denote some auxiliary processes, represented by the simplified rainfall model in our case, which also satisfy a relationship identical to (4)

<u>*The problem is*</u>:

Given a time series $\mathbf{z}_p$ of the actual process $\mathbf{Z}_p$, generate a series $\mathbf{x_s}$ of the actual process $\mathbf{X_s}$. To this aim, we first generate another (auxiliary) time series $\tilde{x}_s$ using the simplified rainfall process $\tilde{X}_s$. The latter time series is generated independently of $\mathbf{z}_p$ and, therefore, $\tilde{x}_s$ do not add up to the corresponding $\mathbf{z}_p$, as required by the additive property (4), but to some other quantities, denoted as $\tilde{z}_p$. Thus, in a subsequent step, we modify the series $\tilde{x}_s$ thus producing the series $\mathbf{x}_s$ consistent with $\mathbf{z}_p$ (in the sense that $\mathbf{x}_s$ and $\mathbf{z}_p$ obey (4) without affecting the stochastic structure of $\tilde{x}_s$. For this modification we use a so-called coupling transformation, i.e., a linear transformation, $\mathbf{f}(\tilde{X}_s\ ,\ \tilde{Z}_p, \mathbf{Z}_p\ )$whose outcome is a process identical to $\mathbf{X}_s$ and consistent to $\mathbf{Z}_p$

Let $\mathbf{X}$ be the vector containing the hourly values of the 24 hours of any day p for all examined locations (i.e., the 24 vectors $\mathbf{X}_s$ for $s = (p - 1)k + 1$ to $s = p\,k$; for 5 locations, $\mathbf{X}$ contains $24 \times 5 = 120$ variables). Let also $\mathbf{Y}$ be a vector containing

(a) the daily values $\mathbf{Z}_p$ for all examined locations,

(b) the daily values $\mathbf{Z}_{p+1}$ of the next day for all locations, and

(c) the hourly values $\mathbf{X}_{(p-1)k}$ of the last hour of the previous day $p - 1$ for all locations.

This means that for 5 locations $\mathbf{Y}$ contains $3 \times 5 = 15$ variables in total. Items (b) and (c) of the vector $\mathbf{Y}$ were included to assure that the transformation will preserve not only the covariance properties among the hourly values of each day, but the covariances with the previous and next days as well. Note that at the stage of the generation at day p the hourly values of day $p - 1$ are known (therefore, in $\mathbf{Y}$ we enter hourly values of the previous day) but the hourly values of day $p + 1$ are not known (therefore, in $\mathbf{Y}$ we enter daily values of the next day, which are known). In an identical manner, we construct the vectors $\tilde{X}$ and $\tilde{Y}$ from variables $\tilde{X}_s$ and $\tilde{\mathbf{Z}}_p$.

Koutsoyiannis [2001] showed that the coupling transformation sought is given by

$$X = \tilde{X} + h\left(Y - \tilde{Y}\right) \tag{5}$$

where

$$\mathbf{h} = \mathrm{Cov}[\mathbf{X}, \mathbf{Y}]\ \{\mathrm{Cov}[\mathbf{Y}, \mathbf{Y}]\}^{-1} \tag{6}$$

The quantity in $h\left(Y - \tilde{Y}\right)$ (5) represents the correction applied to $\tilde{X}$ 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 (4) . 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

$$\Delta = \left\| h\left(Y - \tilde{Y}\right) \right\| / \left(m\sigma_x\right) \tag{7}$$

where m is the common size of $\mathbf{X}$ and $\tilde{X}$ **,** σ_X is standard deviation of hourly depth (common for all locations due to stationarity assumption) and **||.||** denotes the Euclidian norm..

Given the daily process $\mathbf{Z}_p$ and the matrix **h**, which determines completely the transformation, the steps followed to generate the hourly process $\mathbf{X}_s$ are the following:

1. Use the simplified rainfall model (1) or (2) to produce a series $\tilde{X}_s$ for all hours of the current day p and the next day $p + 1$, without reference to $\mathbf{Z}_p$.
2. At day p evaluate the vectors $\mathbf{Y}$ and $\tilde{Y}$ using the values of $\mathbf{Z}_p$ and $\tilde{X}_s$ of the current and next day, and $\mathbf{X}_s$ of the previous day.
3. Determine the quantity $h\left(Y - \tilde{Y}\right)$ and the measure of correction $\mathit{\Delta}$. If $\mathit{\Delta}$ is greater than an accepted limit $\mathit{\Delta}_{\mathrm{m}}$, repeat steps 1-3 (provided that the number of repetitions up to the current repetition has not exceeded a maximum allowed number r_{m}, which is set to avoid unending loops).

4. Apply the coupling transformation to derive **X** of the current period.
5. Repeat steps 1 and 4 for all periods.

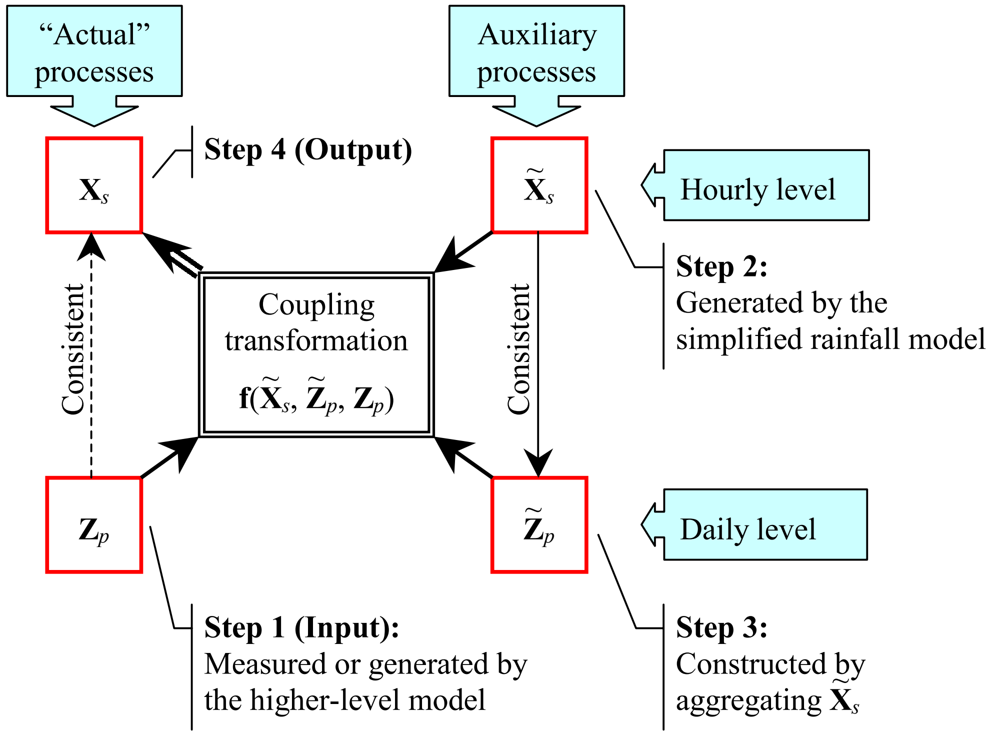


figure 1 : Schematic representation of actual and auxiliary processes, their links, and the steps followed to construct the actual hourly-level process from the actual daily-level process.

procedures for handling the specific difficulties of the methodology

A number of peculiarities of the rainfall process at fine time scale cause some specific difficulties which are listed below:

Negative values:

Linear stochastic models such as those used in the methodology proposed may generally generate negative values, which, of course, do not have any physical meaning. In practice, the probability of generating negative values depends on the coefficient of variation of the variables and it is negligible when these models operate at large time scales, where this coefficient is small. The probability becomes significant in the case of hourly rainfall where the coefficient of variation is very high. This problem can be resolved by setting to zero the negative values generated. This may have a beneficial effect in preserving the proportion of dry intervals but is also a potential source of bias to all statistical properties that are to be preserved. Specifically, it is anticipated to result in overprediction of cross-correlations as it is very probable that negative values are contemporary.

Dry intervals:

The proportion of dry intervals is an important characteristic of the rainfall process that must be preserved. This proportion cannot be preserved by the linear stochastic models in an explicit and theoretically consistent manner. However, after rounding the generated values a significant number of zero values emerge because of the high coefficient of skewness of the rainfall process. Additional zero values result from the truncation of negative values. The total percentage of zero values resulting this way can be comparable to (usually somewhat smaller than) the historical probability dry. We can match exactly the historical probability dry by using the following technique which was found effective: A proportion π_0 of the very small positive values, chosen at random among the generated values that are smaller than a threshold l_0 (e.g., 0.1-0.3 mm), are set to zero. The numbers π_0 and l_0 can be found by performing repetitions starting with different trial values until the proportion of dry intervals in the synthetic series matches that in the historical record.

Preservation of skewness:

Although the coupling transformation preserves the first and second order statistics of the processes, it does not ensure the preservation of third order statistics. Thus, it is anticipated that it will result in underprediction of skewness. However, the repetition technique described earlier can result in good approximation of skewness.

Homoschedasticity of innovations:

By definition, the innovations $\mathbf{V}_s$ in the simplified multivariate rainfall model are homoscedastic, in the sense that their variances are constant, independent of the values of rainfall depths $\mathbf{X}_s$. Therefore, if, for instance, we estimate (or generate) the value at location 2, given that at location 1, we assume that the conditional variance is constant and independent of the value at location 1. This, however, does not comply with reality: by examining simultaneous hyetographs at two locations we can observe that the variance is larger during the periods of high rainfall (peaks) and smaller in periods of low rainfall (heteroscedasticity). As a result of this inconsistency, synthesized hyetographs will tend to have unrealistically similar peaks. To mitigate this problem we can apply a nonlinear transformation to rainfall depths.

The first candidate nonlinear transformation is the logarithmic one,

$$X_s^* = \ln(X_s + \zeta) \tag{8}$$

with constants $\boldsymbol{\zeta} > \mathbf{0}$, where the logarithmic transformation should be read as an item to item one. The stationarity assumption allows considering all items of vector $\boldsymbol{\zeta}$ equal to a constant ζ. This transformation would be an appropriate selection if ζ was estimated so that the transformed series of known hourly depths have zero skewness, in which case the transformed variables could be assumed to be normally distributed. Then, preservation of first and second order properties of the untransformed variables is equivalent to preservation of first- and second-order statistics of the transformed variables [*Koutsoyiannis*, 2001]. However, evidence from the examined data sets shows that the skewness of the transformed variables increases with increasing ζ and it still remains positive even if very small ζ are chosen. This means that the lognormal assumption is not appropriate for hourly rainfall.

A second candidate is the power transformation

$$X_s^* = X_s^{(m)} \tag{9}$$

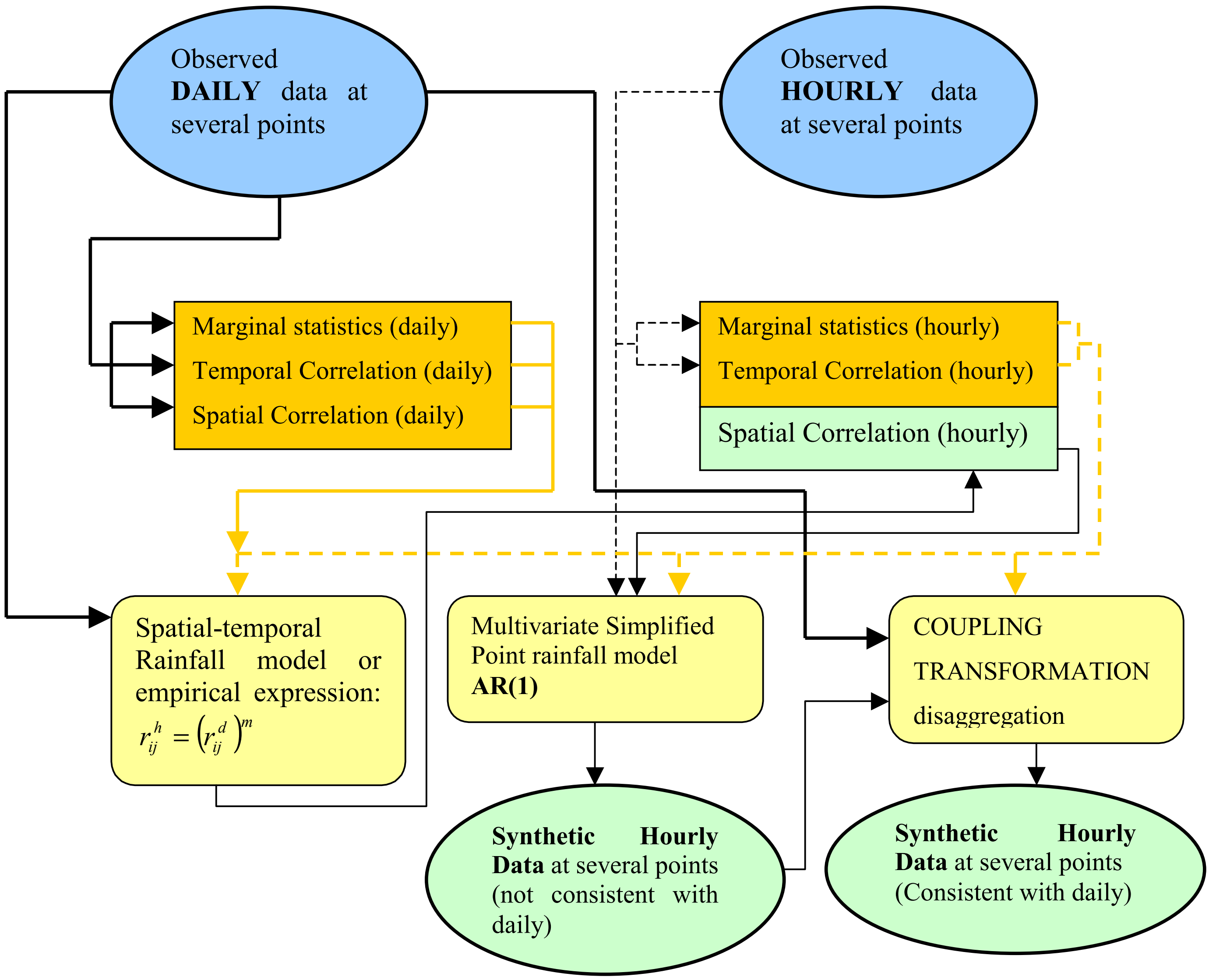
where the symbol (m) means that all items of the vector $\mathbf{X}_s$ are raised to the power m (item to item) where $0 < m < 1$. The stationarity assumption complies with the assumption that m is the same for all items. The preservation of the statistics of the untransformed variables does not necessarily lead to the preservation of the corresponding statistics of the transformed variables. However, the discrepancies are expected to be low if m is not too low (e.g., for $m \geq 0.5$).

The implementation of the Methodology : MuDRain

A computer program for multivariate rainfall disaggregation at a fine time scale

The methodology described in detail in the previous chapter was implemented in a computer program with the name MuDRain.

Therefore, Mudrain operates according to the following modeling scheme:



In order to perform the rainfall disaggregation from daily to hourly scale, a file with the appropriate format must be defined .The program operates using three different files:

1. a file named input.dat (text file) containing the cross-correlation coefficients between raingages at the hourly time scale, the number of gages with daily information, the number of gages with hourly information, the number of days to disaggregate and finally the name of two text files with the hourly and daily information available.
2. a file named daily.inp with the daily information available
3. a file named hourly.inp with the hourly information available.

The user must define only the file input.dat, MuDRain will automatically use the other two files with the information necessary for the disaggregation.

The program 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 (to be incorporated in the input file) whose estimation, requires other methods:

1. By using a spatial-temporal rainfall model
2. By using the empirical expression $r_{ij}^{h} = \left(r_{ij}^{d}\right)^{m}$

Where:

r_{ij}^{h} is the cross-correlation coefficient between raingages i and j at the hourly time scale

r_{ij}^{d} is the cross-correlation coefficient between raingages i and j at the daily time scale

m is an exponent that can be estimated by regression using the known cross-correlation coefficients at the hourly and daily time scale or, in case no hourly data is available, its value can be assumed approximately in the range 2 to 3.

If not specified, MuDRain operates with the simplified model in its linear form:

$\mathbf{X}_s = \mathbf{a}\, \mathbf{X}_{s-1} + \mathbf{b}\, \mathbf{V}_s$

Otherwise the user must define the use of one of the non linear transformations to be adopted:

for the logarithmic transformation $X_s^* = \ln(X_s + \zeta)$, by specifying the value of the constant ζ

for the power transformation $X_s^* = X_s^{(m)}$, by specifying the value of exponent m

Then the expression of the simplified model will be: $X_s^* = aX_{s-1}^* + bV_s$

The program offers other two categories of options, apart from the use or not of one of the transformations, and more specifically:

(a) the use or not of repetition in the generation phase,

whose adoption requires that the user must specify the maximum allowed distance Δ_{m} and the maximum allowed number of repetitions r_{m} and

(b) the use or not of the two-state representation of hourly rainfall, in which case the user must specify the probability φ_0, to stimulate dry state in each of the locations. Two additional parameters are used, which are related to the rounding off rule of generated hourly depths, i.e. the proportion π_0 and the threshold l_0.

In the current program configuration, the options and the additional parameters are specified by the user in a trial-and-error manner, i.e., starting with different trial values until the resulting statistics in the synthetic series match the actual ones. This can be seen as a fine-tuning of the model, which is manual.

For more details on the use of MuDRain see the help file of the program in the Appendix .

CASE STUDY

The methodology proposed by Koutsoyiannis in his computer program (MuDRain) has been applied to the Basin of Tiber River and more specifically to the catchment of its primary tributary the Aniene River located in Central Italy.

The catchment is equipped with several raingages of which only 8 were used in this case study and specifically those of Tivoli, Lunghezza, Frascati, Ponte Salario, Roma Flaminio, Roma Macao, Pantano Borghese and Roma Acqua Acetosa. The data set available was 6 years of hourly-recorded series coming from 6 of the 8 raingages and daily-recorded series from all raingages mentioned, covering the period from January 1994 to December 1999.

The disaggregation was performed using hourly data of three raingages only (raingage 1: Tivoli, raingage 2: Lunghezza, raingage 3: Frascati) and daily data from all raingages, as shown in figure 1.

figure 1

The historical hourly data of raingages at Ponte Salario (4), Roma Flaminio (5) and Roma Macao (6) was used for tests and comparisons with the simulated series obtained in this disaggregation framework to allow the effectiveness of the methodology to be evaluated.

Simulations were performed for each month separately to generate hourly synthetic series for gages 4,5,6,7,8 for all months of the year.

The first step was to determine the statistics of gages 1,2,3 at hourly and daily level among with the cross-correlation coefficients at hourly and daily time scale. We were able to estimate directly the cross-correlation coefficients between the raingages 1,2,3 at the hourly time scale and those between 1,2,…,8 at the daily time scale. The unknown cross-correlation coefficients at hourly level were estimated indirectly using the empirical relationship:

: $r_{ij}^{h} = \left(r_{ij}^{d}\right)^{m}$

Where:

r_{ij}^{h} is the cross-correlation coefficient between raingages *i* and *j* at the hourly time scale

r_{ij}^{d} is the cross-correlation coefficient between raingages *i* and *j* at the daily time scale

m is an exponent that was estimated by regression, using the known cross-correlation coefficients at the hourly and daily time scale, for each month.

In all simulations the single state approach was adopted, so the option of stimulating of dry condition and the related ϕ were deactivated.

The other options used as well as the results of the disaggregation framework for each month are illustrated as it follows.

Month of January

For January, the simplified multivariate model was used in terms of linear transformation $\mathbf{X}_s = \mathbf{a}\,\mathbf{X}_{s-1} + \mathbf{b}\,\mathbf{V}_s$

Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 6000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π were set to 0.3 mm and 0.1 respectively.

Applying the disaggregation modeling framework synthetic hourly rainfall series were produced for the eight raingages, those of gages 1, 2 and 3 being identical to the historical series. The statistics of the synthetic series are compared to the historical and to the values used in the disaggregation in *table 1* where it can be observed a good agreement. Graphical comparisons show that the probability distribution functions of historical and simulated hourly rainfall depth during wet days for gage 6 are also in good agreement with each other (see *figures 2 and 3*).

Lag-one cross correlation coefficients of the synthetic series are compared with those used in disaggregation and with those of the historical series in Table 2. It is shown that acceptable approximations of these statistics have been attained. The synthetic values tend to agree much better with the values used in disaggregation especially given the fact that the historical values have not been entered in the calculations and their preservation could not be assured. Nevertheless these discrepancies can be tolerated.

A further comparison is given in *figure 4* in terms of the autocorrelation function for higher lags, up to 10. It can be observed that even though in theory the synthetic autocorrelations should agree with those of the AR(1) model, they practically agree much better with the historical ones. In fact what forced the synthetic values to agree with the historical ones were the given hourly rainfall series at gages 1, 2, 3.

Hyetographs of the synthetic series given in figures 5-9 show that the disaggregation model predicted the actual hyetographs very well.

Table 1

Statistics of hourly rainfall depths at each gage for the month of JANUARY								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.93	0.92	0.93	0.92	0.92	0.93	-	-
value used on disaggregation	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
synthetic	0.93	0.92	0.93	*0.93*	*0.92*	*0.93*	*0.93*	*0.86*
Mean								
historical	0.09	0.09	0.09	0.08	0.08	0.08	-	-
value used on disaggregation	0.09	0.09	0.09	0.09	0.09	0.09	0.09	0.09
synthetic	0.09	0.09	0.09	*0.08*	*0.08*	*0.08*	*0.07*	*0.09*
Maximum value								
historical	7	8.6	10	9	9.4	8.8	-	-
value used on disaggregation	8.5	8.5	8.5	8.5	8.5	8.5	8.5	8.5
synthetic	7	8.6	10	*8.4*	*7.8*	*9.6*	*7.8*	*7.8*
Standard deviation								
historical	0.485	0.502	0.506	0.467	0.441	0.459	-	-
value used on disaggregation	0.497	0.497	0.497	0.497	0.497	0.497	0.497	0.497
synthetic	0.484	0.502	0.506	*0.458*	*0.413*	*0.447*	*0.415*	*0.407*
Skewness								
historical	7.66	9.57	9.37	8.56	9.37	8.83	-	-
value used on disaggregation	8.87	8.87	8.87	8.87	8.87	8.87	8.87	8.87
synthetic	7.67	9.58	9.38	*8.63*	*8.53*	*9.13*	*9.28*	*8.37*
Lag1 autocorrelation								
historical	0.58	0.54	0.50	0.53	0.54	0.58	-	-
value used on disaggregation	0.54	0.54	0.54	0.54	0.54	0.54	0.54	0.54
synthetic	0.58	0.55	0.50	*0.56*	*0.60*	*0.57*	*0.52*	*0.59*

Table 2

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of January

Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1								
historical	1.00	0.73	0.47	0.52	0.49	0.53	-	-
value used on	1.00	0.73	0.47	0.71	0.66	0.61	0.54	0.49
synthetic	1.00	0.74	0.47	0.73	0.75	0.67	0.63	0.59
2								
historical	0.73	1.00	0.52	0.72	0.60	0.71	-	-
value used on	0.73	1.00	0.52	0.84	0.72	0.77	0.71	0.68
synthetic	0.74	1.00	0.52	0.88	0.82	0.85	0.85	0.83
3								
historical	0.47	0.52	1.00	0.43	0.40	0.46	-	-
value used on	0.47	0.52	1.00	0.46	0.45	0.48	0.44	0.39
synthetic	0.47	0.52	1.00	0.51	0.56	0.58	0.50	0.50
4								
historical	0.52	0.72	0.43	1.00	0.82	0.87	-	-
value used on	0.71	0.84	0.46	1.00	0.90	0.90	0.65	0.78
synthetic	0.73	0.88	0.51	1.00	0.93	0.93	0.78	0.86
5								
historical	0.49	0.60	0.40	0.82	1.00	0.75	-	-
value used on	0.66	0.72	0.45	0.90	1.00	0.80	0.56	0.69
synthetic	0.75	0.82	0.56	0.93	1.00	0.89	0.74	0.82
6								
historical	0.53	0.71	0.46	0.87	0.75	1.00	-	-
value used on	0.61	0.77	0.48	0.90	0.80	1.00	0.63	0.75
synthetic	0.67	0.85	0.58	0.93	0.89	1.00	0.76	0.85
7								
historical	-	-	-	-	-	-	-	-
value used on	0.54	0.71	0.44	0.65	0.56	0.63	1.00	0.82
synthetic	0.63	0.85	0.50	0.78	0.74	0.76	1.00	0.89
8								
historical	-	-	-	-	-	-	-	-
value used on	0.49	0.68	0.39	0.78	0.69	0.75	0.82	1.00
synthetic	0.59	0.83	0.50	0.86	0.82	0.85	0.89	1.00

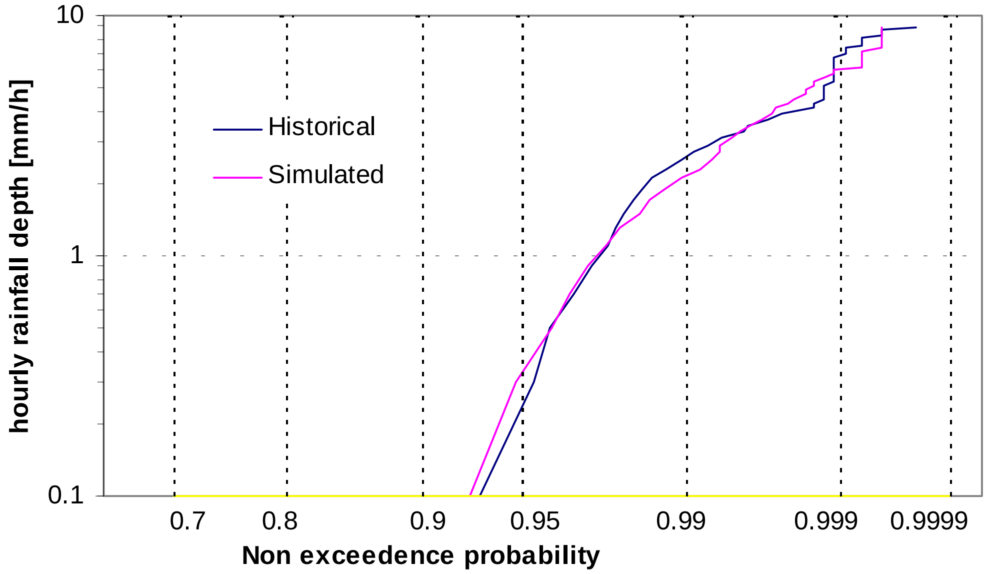


Figure 2: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 6 for the month of January

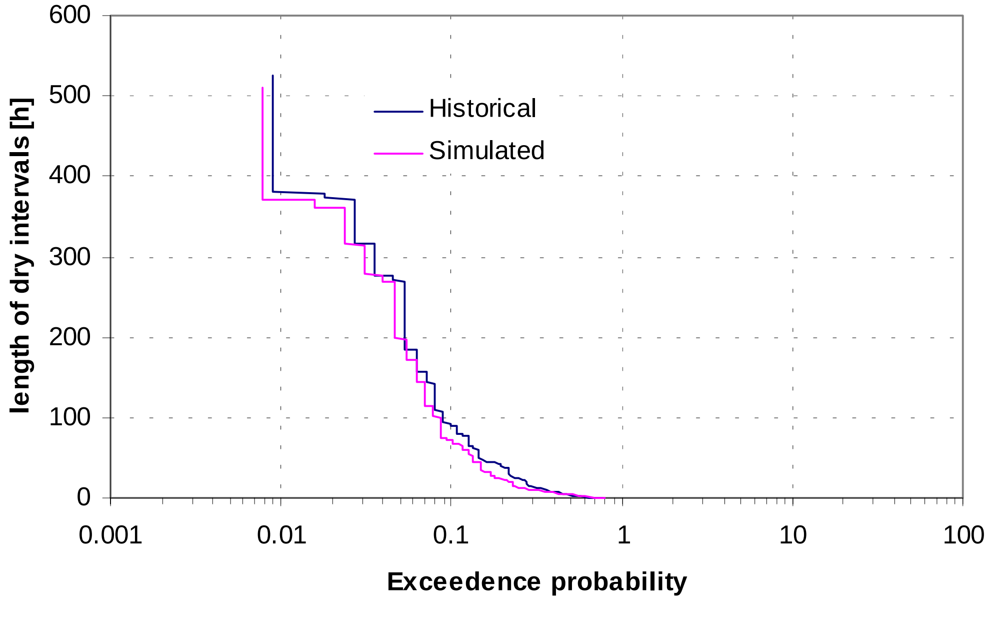


Figure 3: Comparison of historical and simulated probability distribution functions of of the length of dry intervals at gage 6 for the month of January

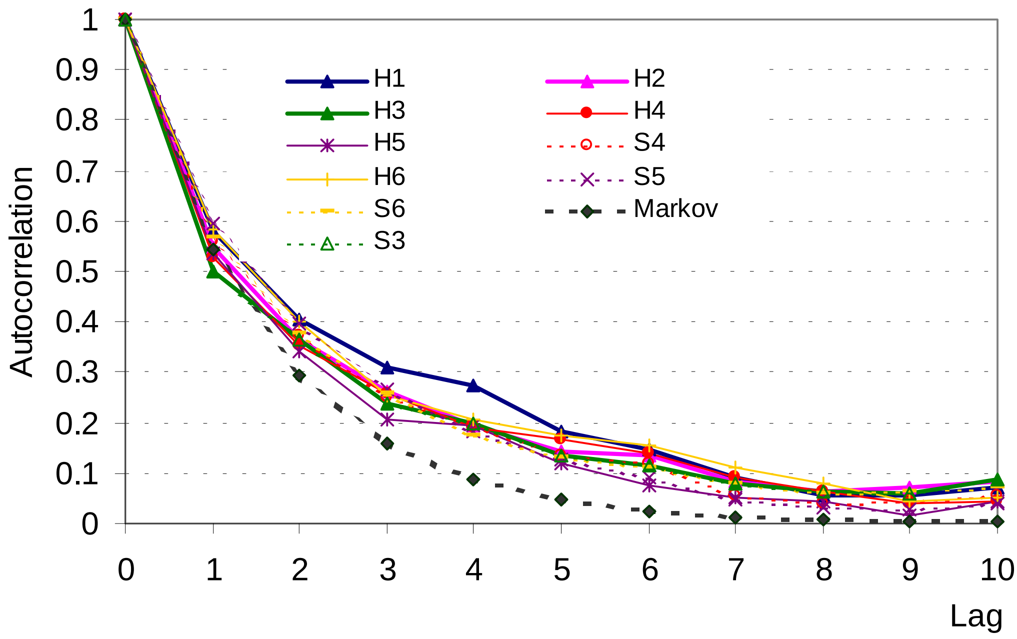


Figure 4: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of January.

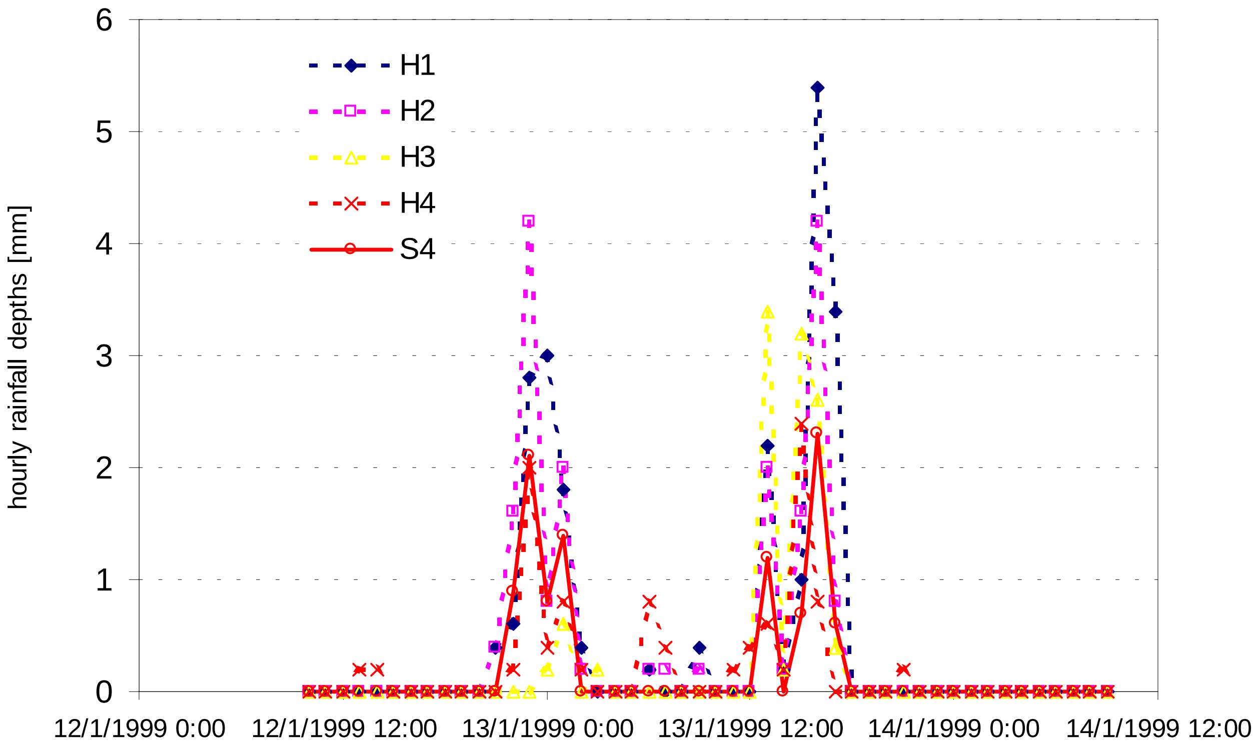


Figure 5: Comparison of historical and simulated hyetographs for raingage 4

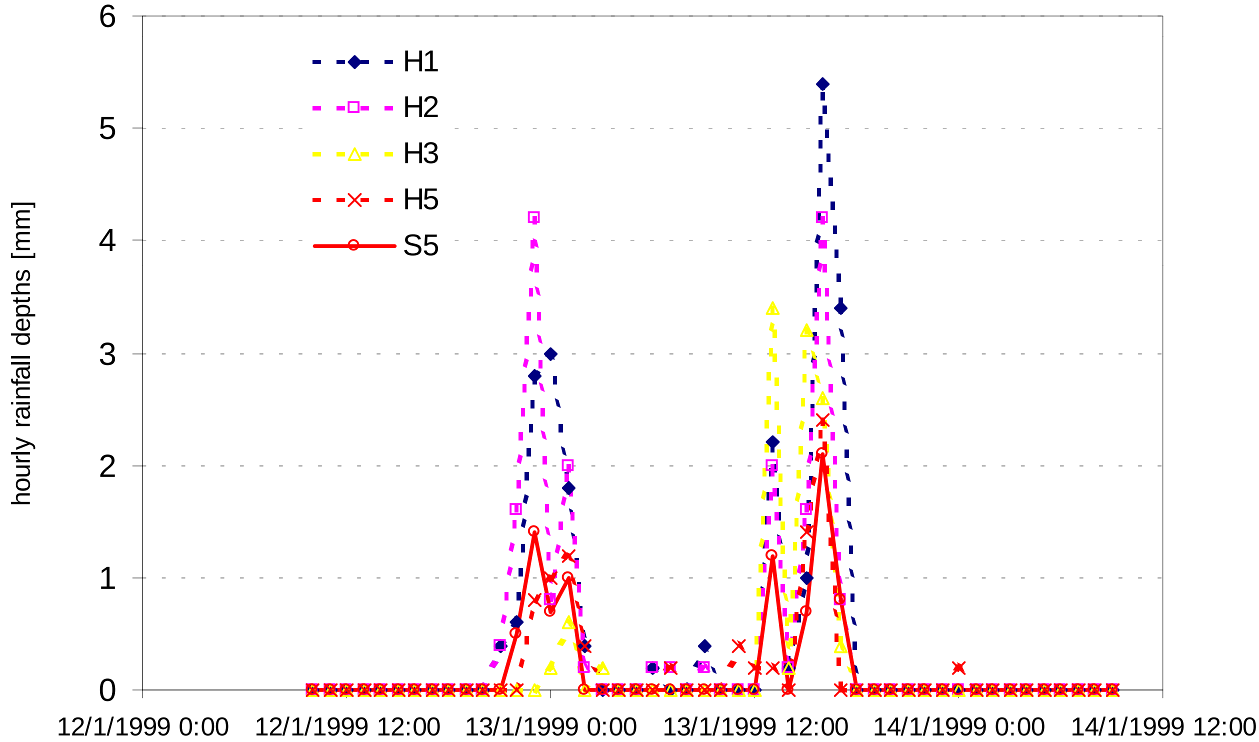


Figure 6: Comparison of historical and simulated hyetographs for raingage 5

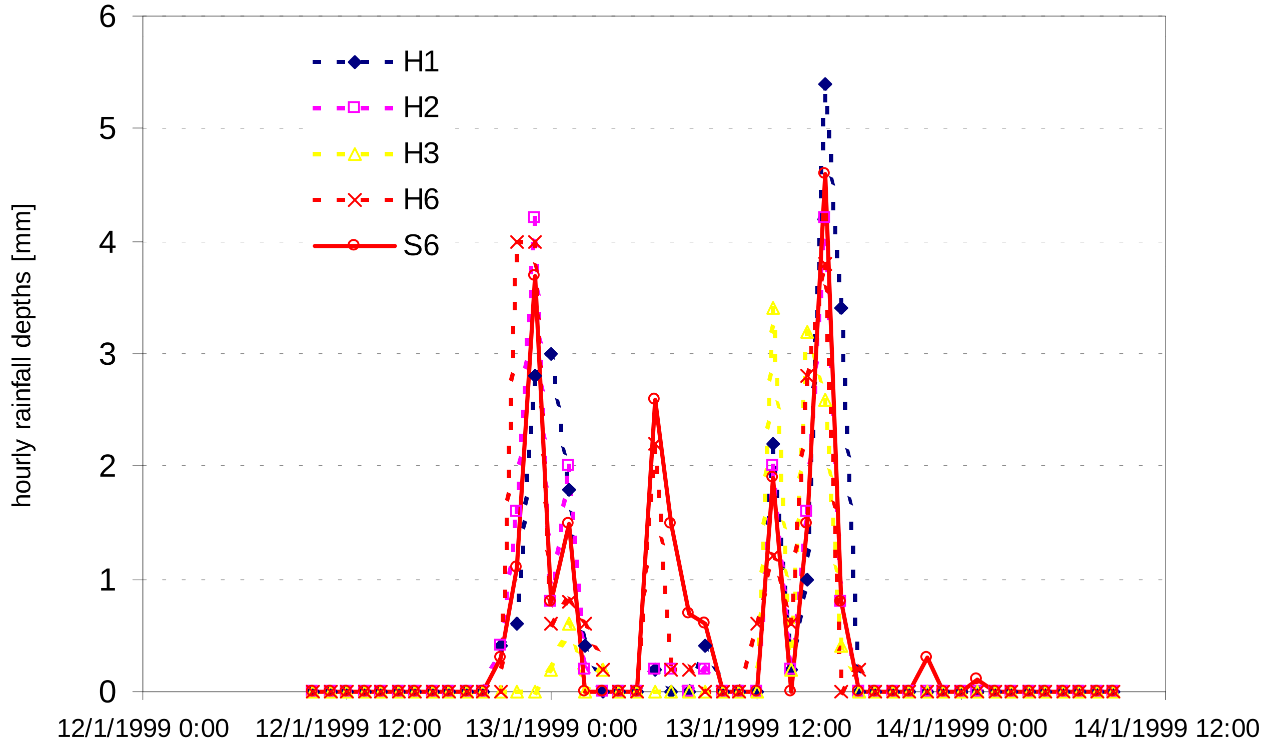


Figure 7: Comparison of historical and simulated hyetographs for raingage 6

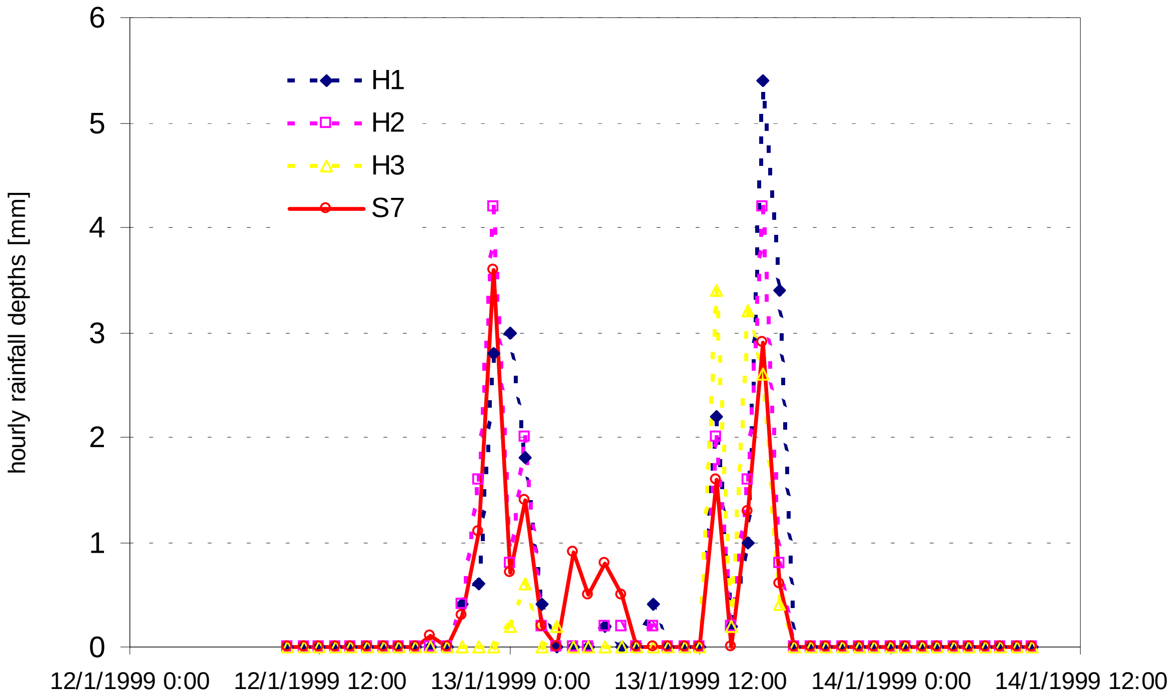


Figure 8: Simulated hyetograph for raingage 7

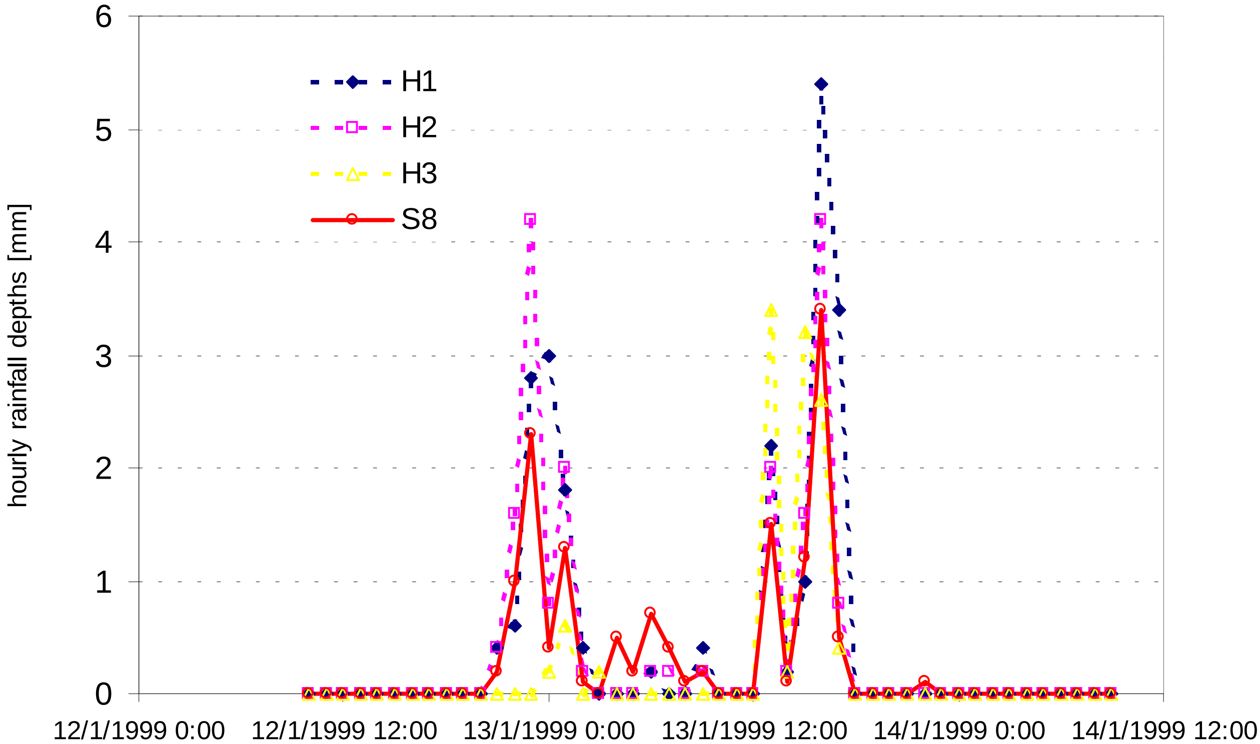


Figure 9: Simulated hyetograph for raingage 8

Month of February

For February, the simplified multivariate model was used in terms of linear transformation $\mathbf{X}_s = \mathbf{a}\,\mathbf{X}_{s-1} + \mathbf{b}\,\mathbf{V}_s$

Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 3000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π_0 were set to 0.1 mm and 0.1 respectively.

The statistics of the synthetic series are compared to the historical and to the values used in the disaggregation in *table 3* where it can be observed a good agreement. Graphical comparisons show that the probability distribution functions of historical and simulated hourly rainfall depth during wet days for gage 5 are also in good agreement with each other (*see figures 10 and 11*).

Lag-one cross correlation coefficients of the synthetic series are compared with those used in disaggregation and with those of the historical series in *Table 4*. It is shown that acceptable approximations of these statistics have been attained. The synthetic values tend to agree much better with the values used in disaggregation especially given the fact that the historical values have not been entered in the calculations and their preservation could not be assured. Nevertheless these discrepancies can be tolerated.

A further comparison is given in *figure 12* in terms of the autocorrelation function for higher lags, up to 10. It can be observed that even though in theory the synthetic autocorrelations should agree with those of the AR(1) model, they practically agree much better with the historical ones. In fact what forced the synthetic values to agree with the historical ones were the given hourly rainfall series at gages 1, 2, 3.

Hyetographs of the synthetic series given in *figure 13-17* show that the disaggregation model predicted the actual hyetographs well.

Table 4

Statistics of hourly rainfall depths at each gage for the month of FEBRUARY								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.91	0.90	0.93	0.91	0.92	0.92	-	-
value used on disaggregation	0.914	0.914	0.914	0.914	0.914	0.914	0.914	0.914
synthetic	0.91	0.90	0.93	*0.88*	*0.85*	*0.88*	*0.89*	*0.86*
Mean								
historical	0.11	0.10	0.08	0.10	0.10	0.10	-	-
value used on disaggregation	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
synthetic	0.11	0.10	0.08	*0.10*	*0.10*	*0.10*	*0.10*	*0.11*
Maximum value								
historical	11.2	20.8	14.8	14.6	13.8	13.2	-	-
value used on disaggregation	15.6	15.6	15.6	15.6	15.6	15.6	15.6	15.6
synthetic	11.2	20.7	14.8	*11*	*14*	*10.4*	*12.4*	*12.8*
Standard deviation								
historical	0.56	0.60	0.50	0.57	0.52	0.54	-	-
value used on disaggregation	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55
synthetic	0.56	0.60	0.50	*0.50*	*0.48*	*0.46*	*0.53*	*0.54*
Skewness								
historical	9.05	15.22	12.54	11.15	9.96	10.55	-	-
value used on disaggregation	12.273	12.309	12.309	12.309	12.309	12.309	12.309	12.309
synthetic	9.05	15.15	12.55	*9.18*	*12.17*	*8.71*	*10.38*	*9.77*
Lag1 autocorrelation								
historical	0.52	0.43	0.50	0.51	0.50	0.48	-	-
value used on disaggregation	0.484	0.484	0.484	0.484	0.484	0.484	0.484	0.484
synthetic	0.53	0.43	0.50	*0.51*	*0.58*	*0.53*	*0.61*	*0.56*

Table 5

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of February

Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1								
historical	1.00	0.68	0.31	0.58	0.33	0.53	-	-
value used on disaggregation	1.00	0.68	0.31	0.67	0.45	0.69	0.63	0.58
synthetic	1.00	0.68	0.31	0.66	0.47	0.72	0.71	0.61
2								
historical	0.68	1.00	0.33	0.69	0.37	0.68	-	-
value used on disaggregation	0.68	1.00	0.33	0.73	0.54	0.74	0.57	0.67
synthetic	0.68	1.00	0.34	0.75	0.54	0.82	0.60	0.70
3								
historical	0.31	0.33	1.00	0.27	0.27	0.30	-	-
value used on disaggregation	0.31	0.33	1.00	0.15	0.22	0.19	0.41	0.16
synthetic	0.31	0.34	1.00	0.25	0.40	0.31	0.45	0.26
4								
historical	0.58	0.69	0.27	1.00	0.59	0.83	-	-
value used on disaggregation	0.67	0.73	0.15	1.00	0.65	0.91	0.41	0.85
synthetic	0.66	0.75	0.25	1.00	0.66	0.97	0.54	0.88
5								
historical	0.33	0.37	0.27	0.59	1.00	0.52	-	-
value used on disaggregation	0.45	0.54	0.22	0.65	1.00	0.65	0.31	0.61
synthetic	0.47	0.54	0.40	0.66	1.00	0.70	0.43	0.64
6								
historical	0.53	0.68	0.30	0.83	0.52	1.00	-	-
value used on disaggregation	0.69	0.74	0.19	0.91	0.65	1.00	0.45	0.78
synthetic	0.72	0.82	0.31	0.97	0.70	1.00	0.60	0.86
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.63	0.57	0.41	0.41	0.31	0.45	1.00	0.42
synthetic	0.71	0.60	0.45	0.54	0.43	0.60	1.00	0.54
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.58	0.67	0.16	0.85	0.61	0.78	0.42	1.00
synthetic	0.61	0.70	0.26	0.88	0.64	0.86	0.54	1.00

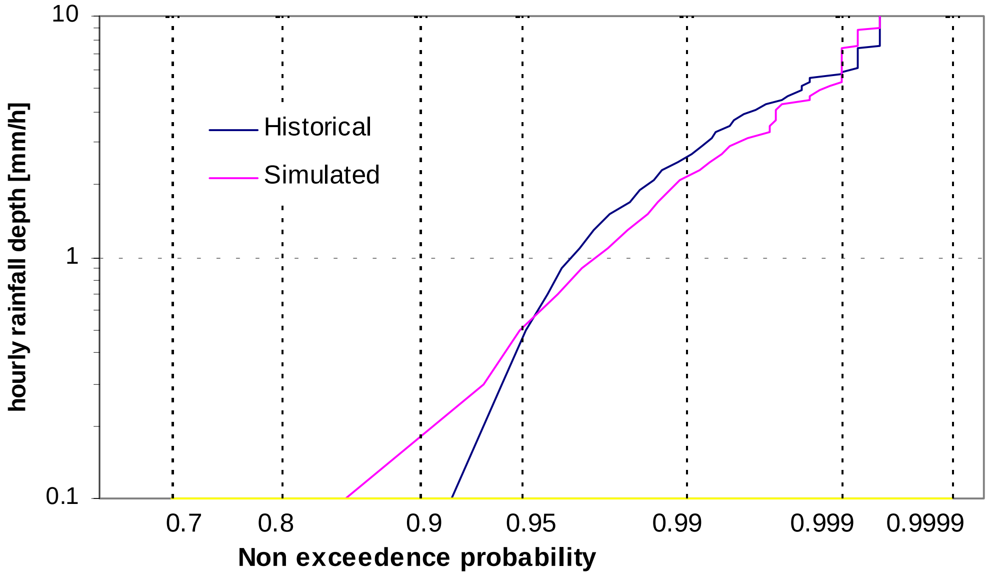


Figure 10: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 5 for the month of February

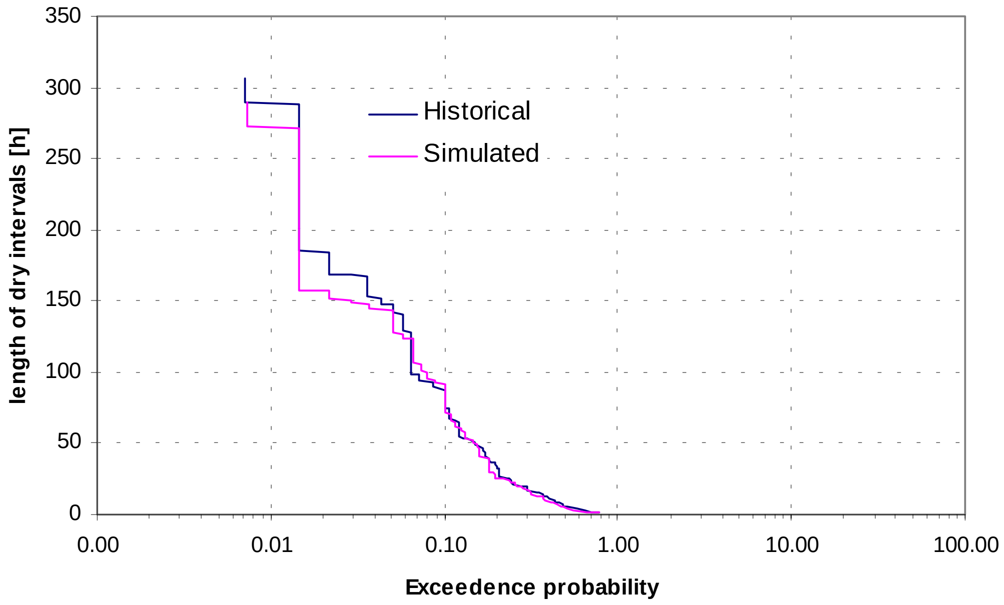


Figure 11: Comparison of historical and simulated probability distribution functions of of the length of dry intervals at gage 5 for the month of February

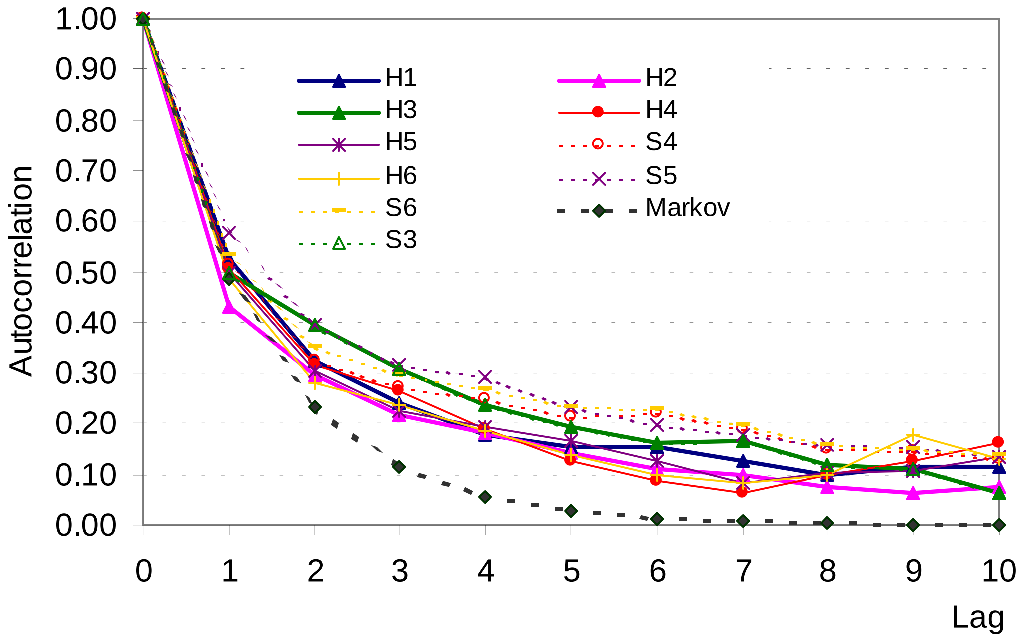


Figure 12: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of February.

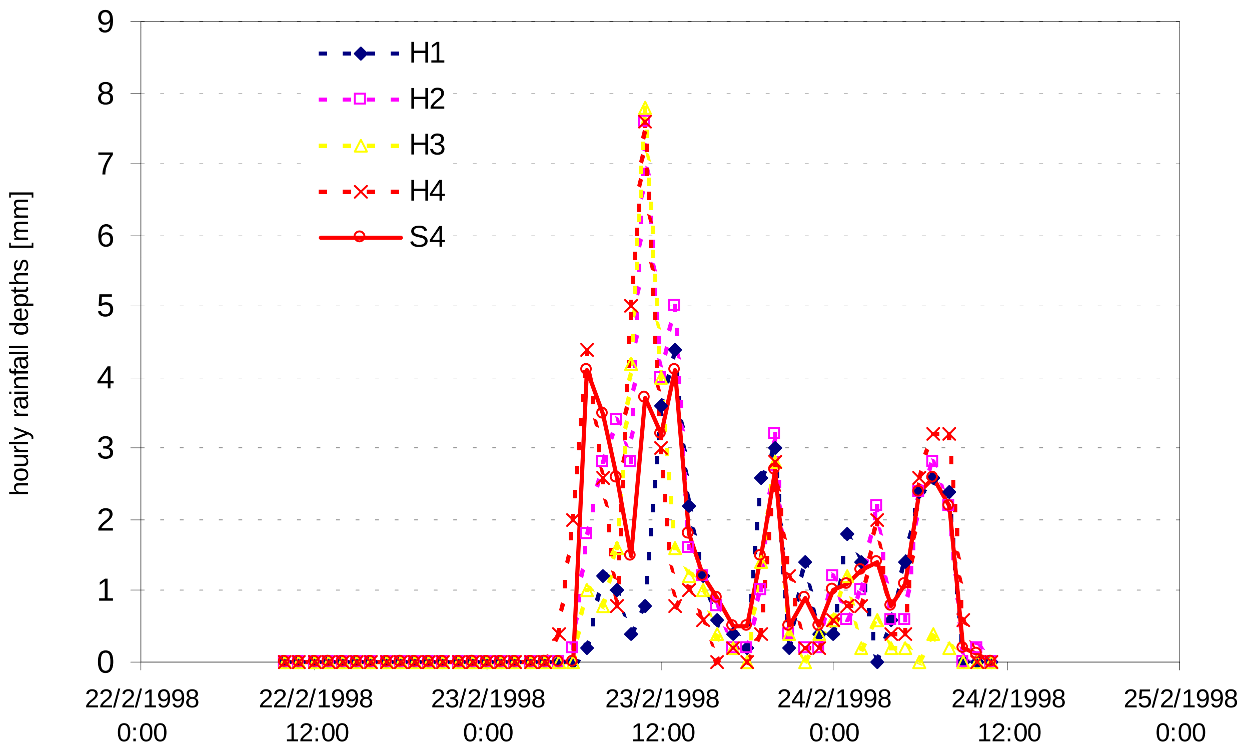


Figure 13: Comparison of historical and simulated hyetographs for raingage 4

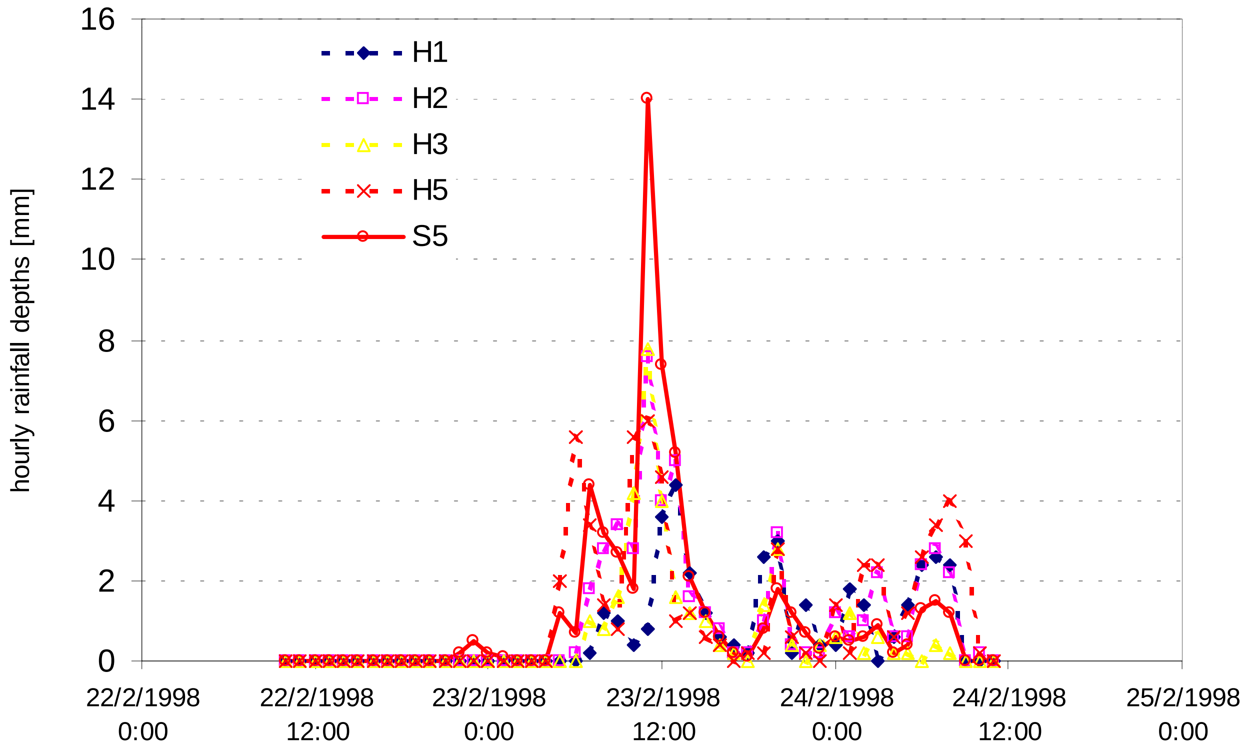


Figure 14: Comparison of historical and simulated hyetographs for raingage 5

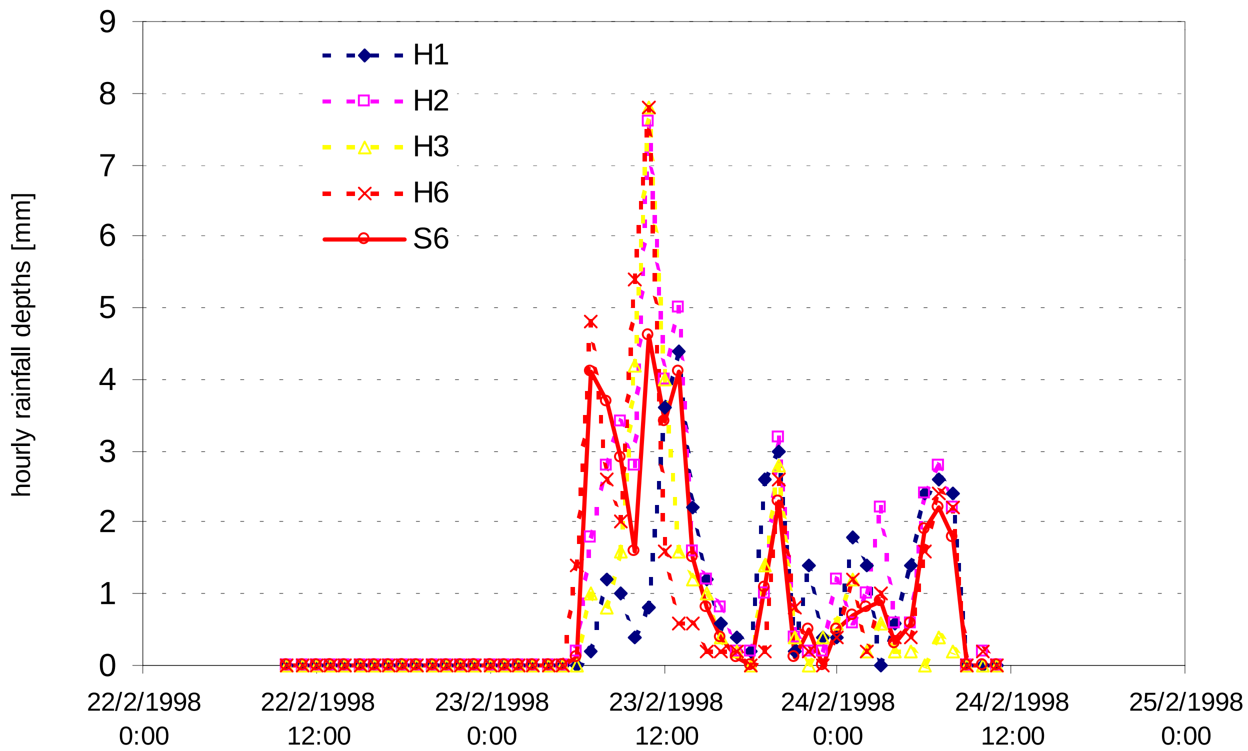


Figure 15: Comparison of historical and simulated hyetographs for raingage 6

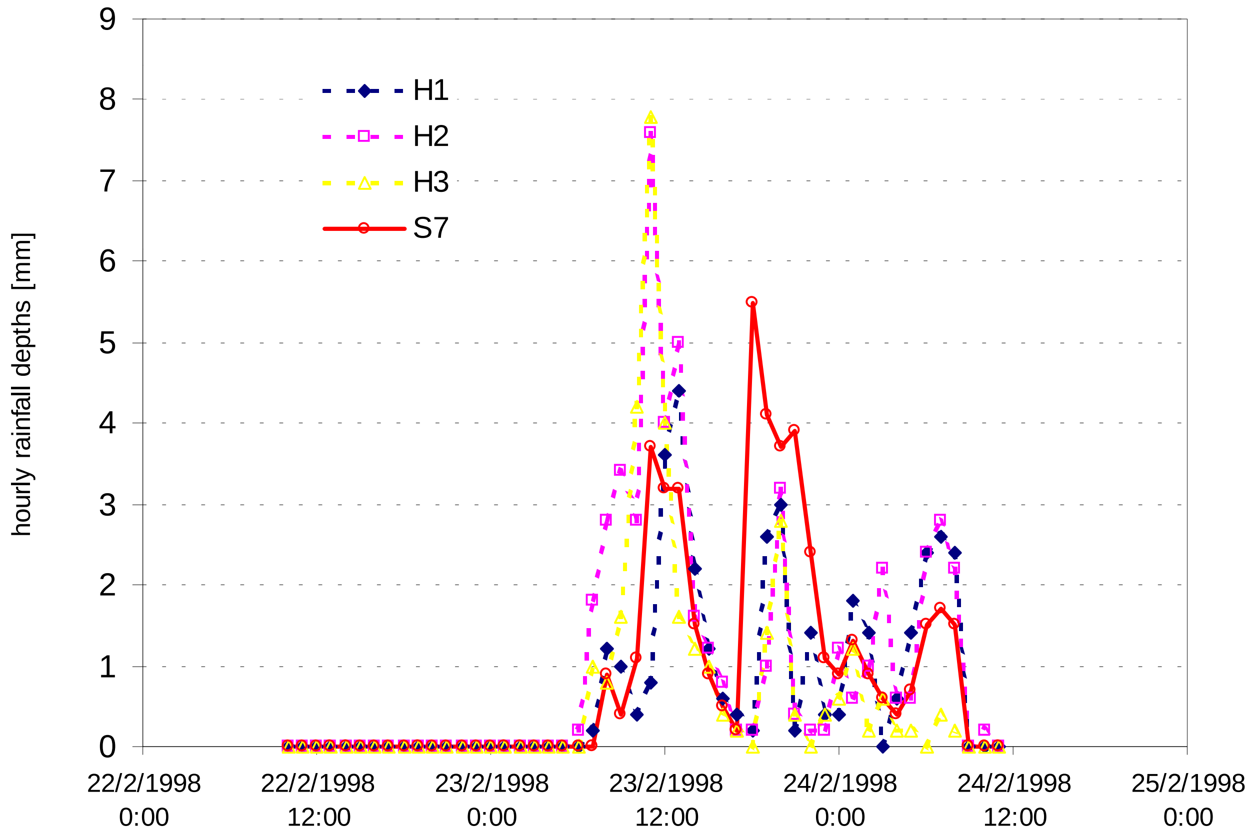


Figure 16: Simulated hyetographs for raingage 7

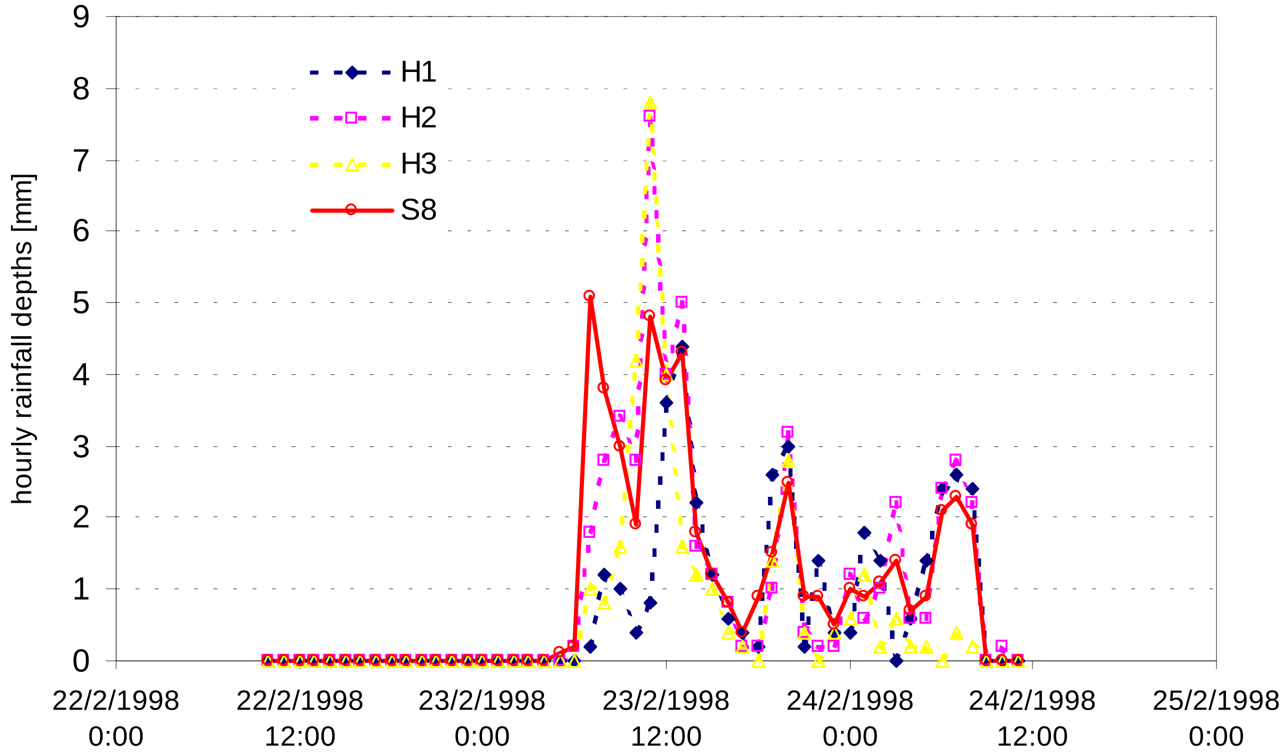


Figure 17: Simulated hyetographs for raingage 8

Month of March

For March, the simplified multivariate model was used in terms of linear transformation $\mathbf{X}_s = \mathbf{a}\,\mathbf{X}_{s-1} + \mathbf{b}\,\mathbf{V}_s$

Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 1000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π were set to 0.3 mm and 0.1 respectively.

Applying the disaggregation modeling framework synthetic hourly rainfall series were produced for the eight raingages, those of gages 1, 2 and 3 being identical to the historical series. The statistics of the synthetic series are compared to the historical and to the values used in the disaggregation in *table 5* where it can be observed a good agreement. Graphical comparisons show that the probability distribution functions of historical and simulated hourly rainfall depth during wet days for gage *4* are also in good agreement with each other (see *figures 18 and 19*).

Lag-one cross correlation coefficients of the synthetic series are compared with those used in disaggregation and with those of the historical series in *Table 6*. It is shown that acceptable approximations of these statistics have been attained. The synthetic values tend to agree much better with the values used in disaggregation especially given the fact that the historical values have not been entered in the calculations and their preservation could not be assured. Nevertheless these discrepancies can be tolerated.

A further comparison is given in *figure 20* in terms of the autocorrelation function for higher lags, up to 10. It can be observed that even though in theory the synthetic autocorrelations should agree with those of the AR(1) model, they practically agree much better with the historical ones. In fact what forced the synthetic values to agree with the historical ones were the given hourly rainfall series at gages 1, 2, 3.

Hyetographs of the synthetic series given in *figures 21-25* show that the disaggregation model predicted the actual hyetographs very well.

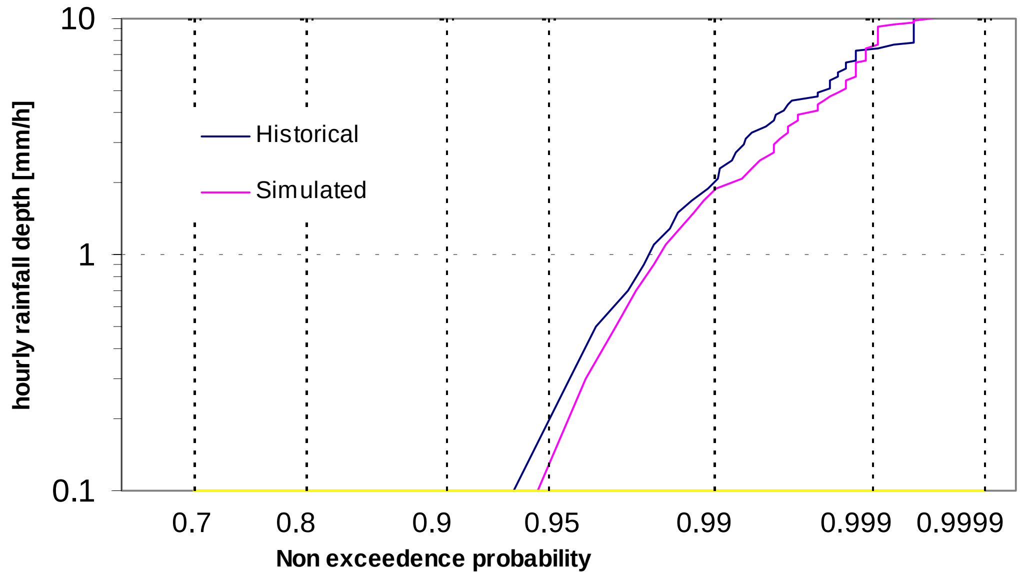
Table 5

Statistics of hourly rainfall depths at each gage for the month of MARCH								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.95	0.94	0.95	0.94	0.94	0.95	-	-
value used on disaggregation	0.946	0.946	0.946	0.946	0.946	0.946	0.946	0.946
synthetic	0.95	0.94	0.95	*0.93*	*0.92*	*0.94*	*0.94*	*0.90*
Mean								
historical	0.07	0.06	0.06	0.07	0.07	0.07	-	-
value used on disaggregation	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
synthetic	0.07	0.06	0.06	*0.07*	*0.07*	*0.07*	*0.07*	*0.08*
Maximum value								
historical	14.40	10.20	10.00	10.60	10.20	14.40	-	-
value used on disaggregation	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5
synthetic	14.40	10.20	10.00	*11.20*	*11.00*	*10.30*	*8.90*	*10.50*
Standard deviation								
historical	0.54	0.37	0.43	0.48	0.45	0.45	-	-
value used on disaggregation	0.448	0.446	0.446	0.446	0.446	0.446	0.446	0.446
synthetic	0.53	0.37	0.43	*0.50*	*0.43*	*0.43*	*0.44*	*0.47*
Skewness								
historical	13.57	11.83	13.08	10.91	11.20	13.27	-	-
value used on disaggregation	12.82	12.86	12.86	12.86	12.86	12.86	12.86	12.86
synthetic	13.62	11.82	13.10	*12.06*	*11.60*	*11.25*	*11.64*	*11.01*
Lag1 autocorrelation								
historical	0.52	0.41	0.49	0.48	0.53	0.47	-	-
value used on disaggregation	0.472	0.472	0.472	0.472	0.472	0.472	0.472	0.472
synthetic	0.52	0.41	0.49	*0.55*	*0.55*	*0.58*	*0.55*	*0.62*

Table 6

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of March

Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1								
historical	1.00	0.72	0.48	0.59	0.49	0.59	-	-
value used on disaggregation	1.00	0.72	0.48	0.73	0.74	0.76	0.74	0.73
synthetic	1.00	0.71	0.48	0.73	0.78	0.77	0.78	0.75
2								
historical	0.72	1.00	0.57	0.70	0.59	0.74	-	-
value used on disaggregation	0.72	1.00	0.57	0.75	0.85	0.82	0.65	0.77
synthetic	0.71	1.00	0.57	0.76	0.88	0.83	0.68	0.79
3								
historical	0.48	0.57	1.00	0.46	0.41	0.59	-	-
value used on disaggregation	0.48	0.57	1.00	0.52	0.48	0.59	0.52	0.55
synthetic	0.48	0.57	1.00	0.53	0.55	0.59	0.55	0.57
4								
historical	0.59	0.70	0.46	1.00	0.69	0.83	-	-
value used on disaggregation	0.73	0.75	0.52	1.00	0.89	0.93	0.51	0.96
synthetic	0.73	0.76	0.53	1.00	0.93	0.96	0.60	0.97
5								
historical	0.49	0.59	0.41	0.69	1.00	0.69	-	-
value used on disaggregation	0.74	0.85	0.48	0.89	1.00	0.87	0.50	0.88
synthetic	0.78	0.88	0.55	0.93	1.00	0.95	0.68	0.94
6								
historical	0.59	0.74	0.59	0.83	0.69	1.00	-	-
value used on disaggregation	0.76	0.82	0.59	0.93	0.87	1.00	0.57	0.94
synthetic	0.77	0.83	0.59	0.96	0.95	1.00	0.67	0.97
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.74	0.65	0.52	0.51	0.50	0.57	1.00	0.52
synthetic	0.78	0.68	0.55	0.60	0.68	0.67	1.00	0.64
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.73	0.77	0.55	0.96	0.88	0.94	0.52	1.00
synthetic	0.75	0.79	0.57	0.97	0.94	0.97	0.64	1.00



\\

Figure 18: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 4 for the month of March

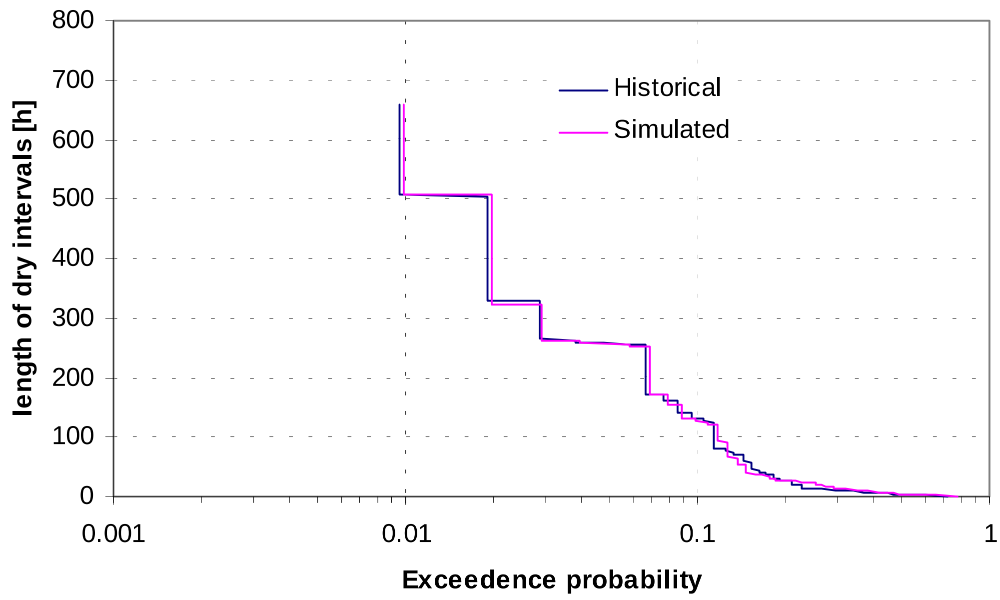


Figure 19: Comparison of historical and simulated probability distribution functions of of the length of dry intervals at gage 4 for the month of March

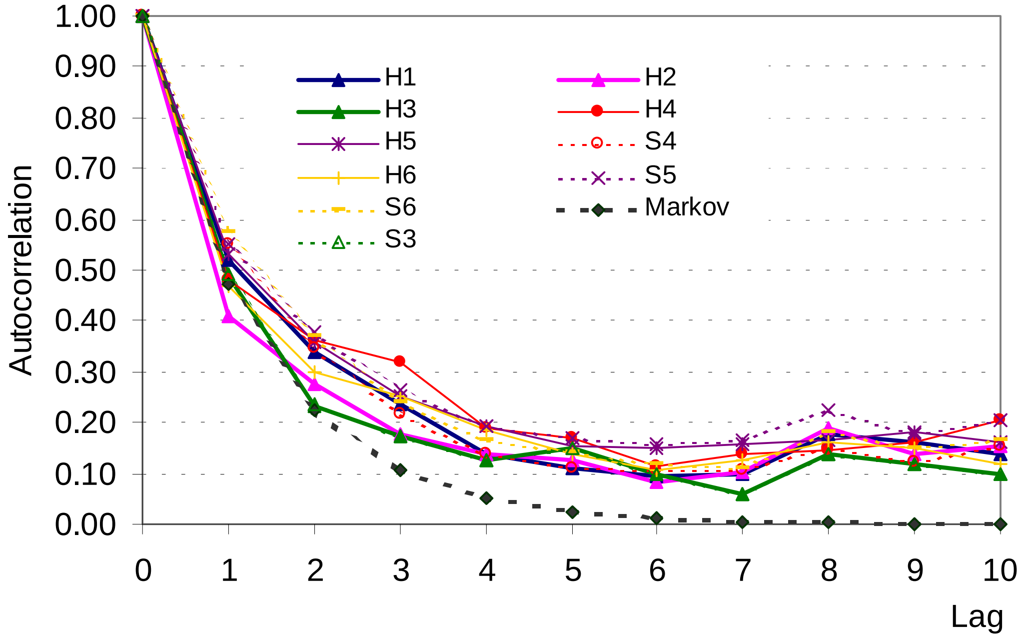
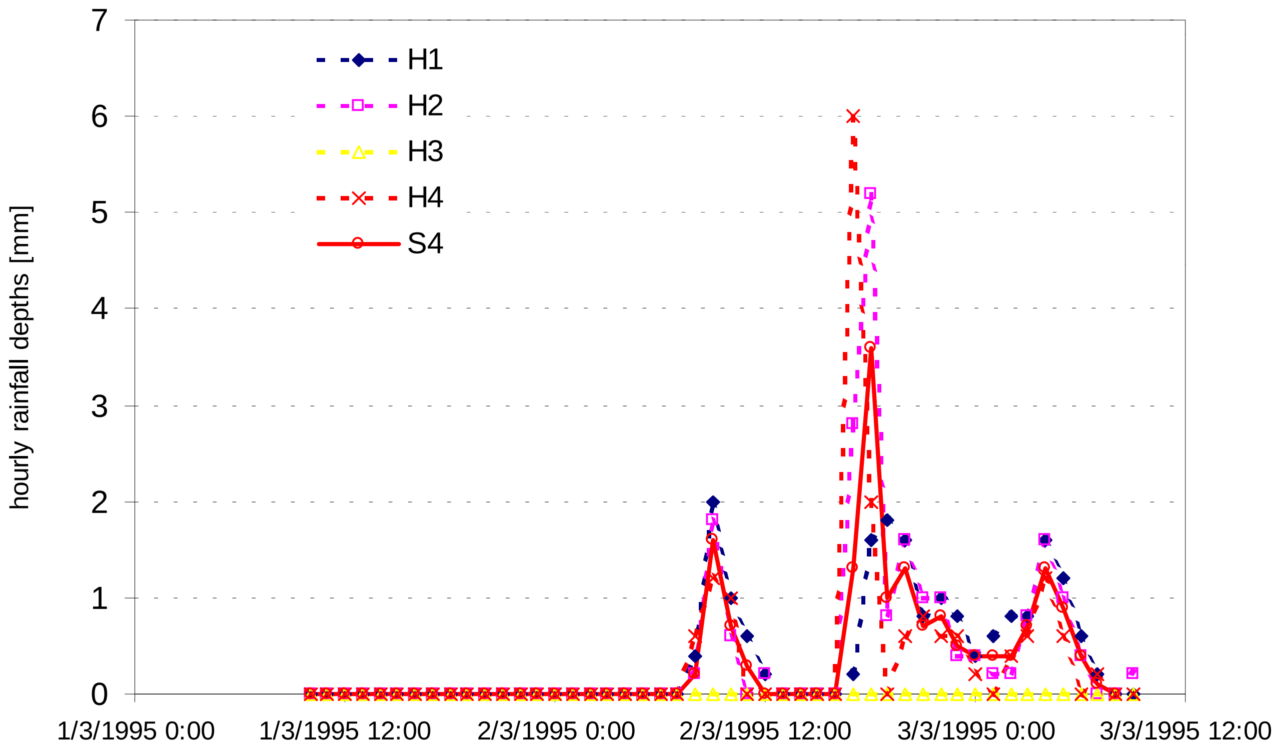


Figure 20: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of March.



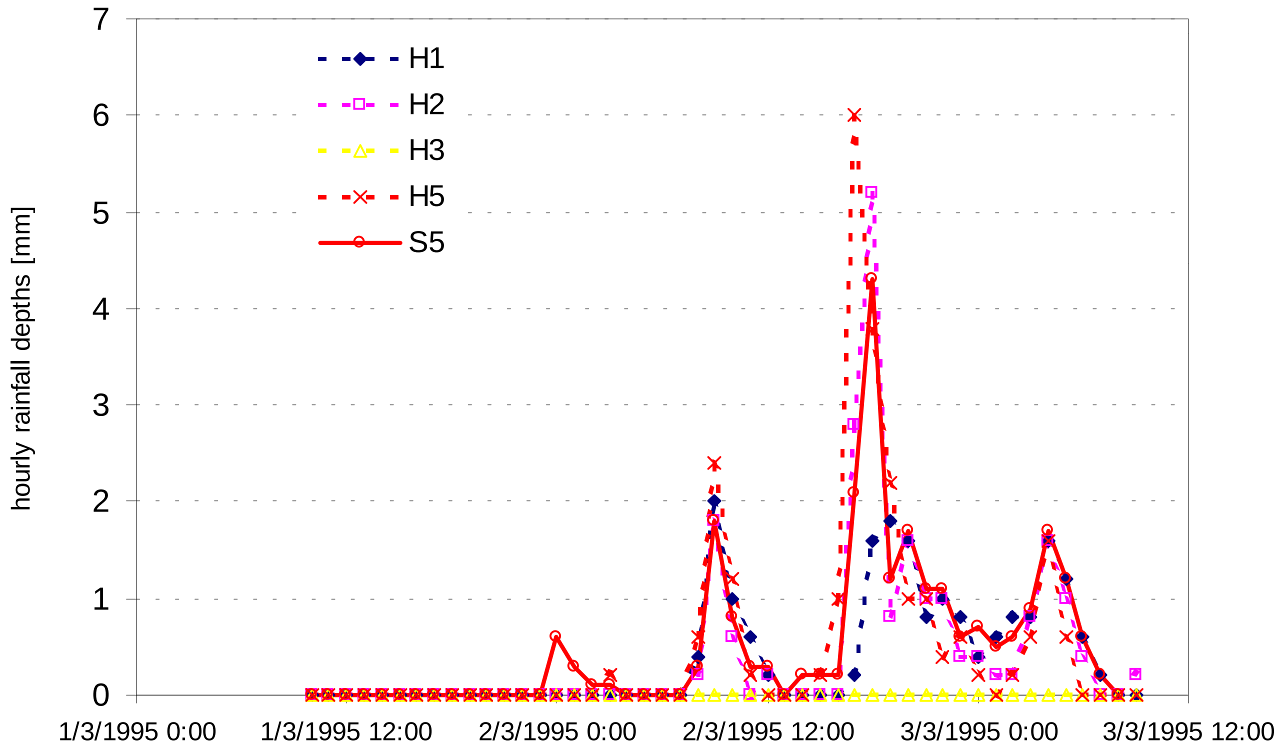


Figure 22: Comparison of historical and simulated hyetographs for raingage 5

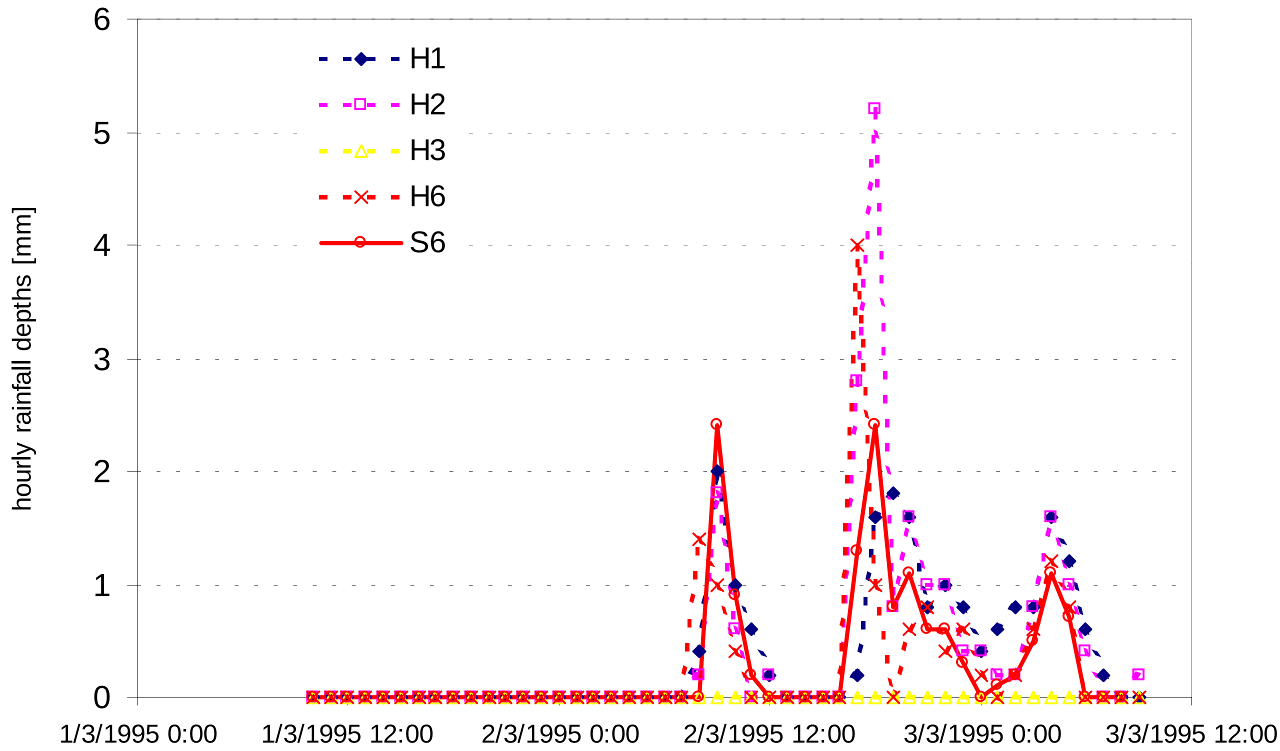


Figure 23: Comparison of historical and simulated hyetographs for raingage 6

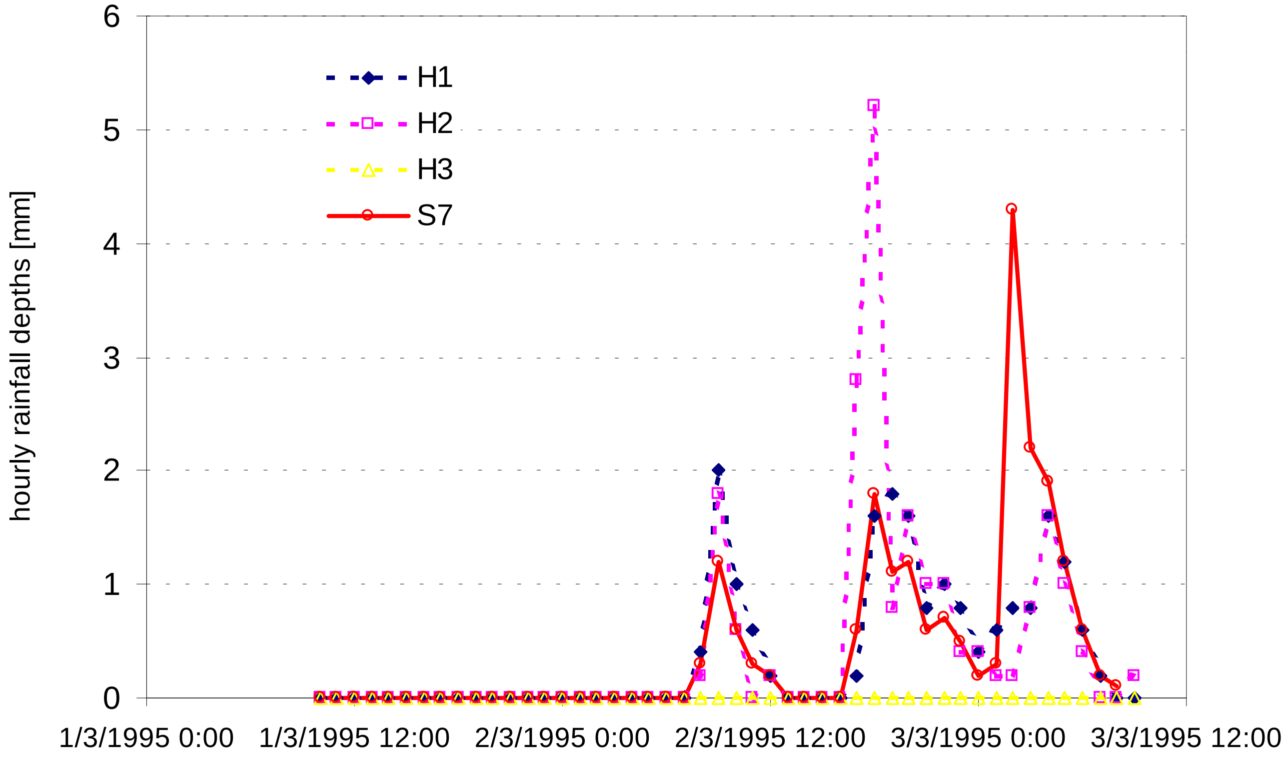


Figure 24: Simulated hyetographs for raingage 7

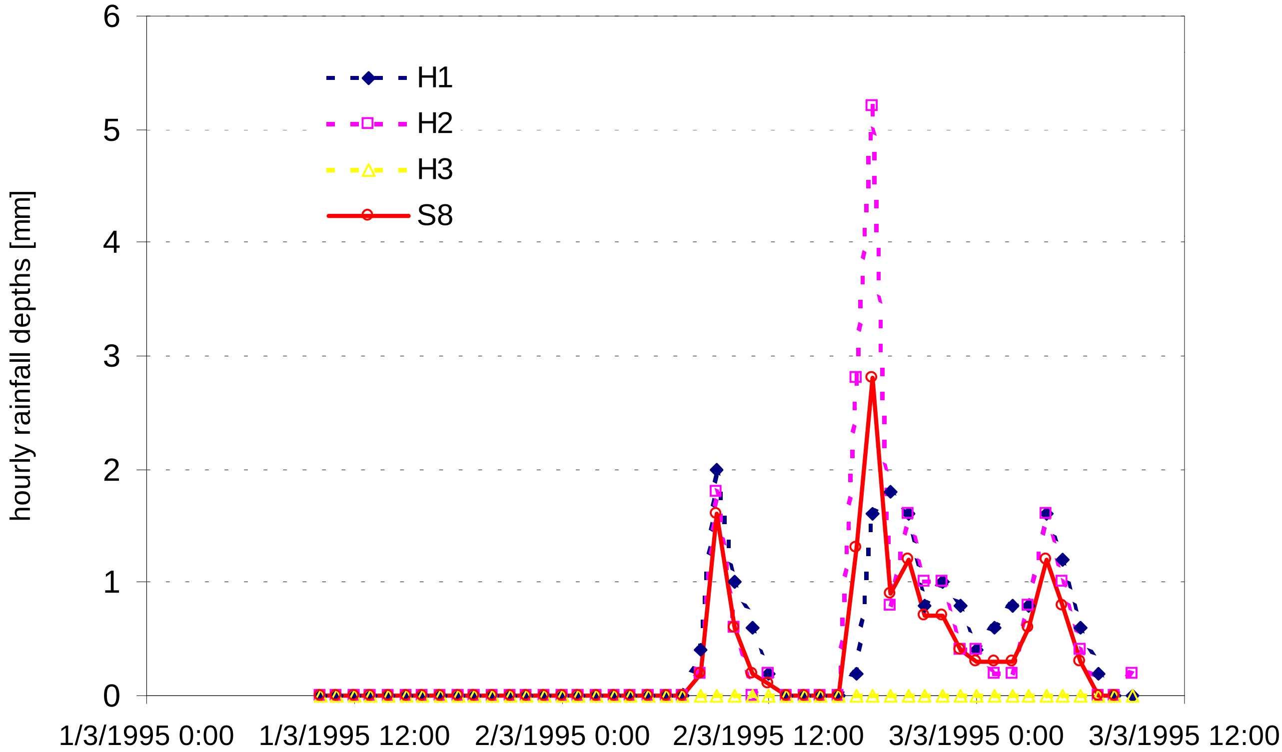


Figure 25: Simulated hyetographs for raingage 8

Month of April

For April, the simplified multivariate model was used in terms of linear transformation $\mathbf{X}_s = \mathbf{a}\,\mathbf{X}_{s-1} + \mathbf{b}\,\mathbf{V}_s$

Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 5000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π were set to 0.3 mm and 0.2 respectively.

Applying the disaggregation modeling framework synthetic hourly rainfall series were produced for the eight raingages, those of gages 1, 2 and 3 being identical to the historical series. The statistics of the synthetic series are compared to the historical and to the values used in the disaggregation in *table 7* where it can be observed a good agreement. Graphical comparisons show that the probability distribution functions of historical and simulated hourly rainfall depth during wet days for gage *4* are also in good agreement with each other (see *figures 26 and 27*).

Lag-one cross correlation coefficients of the synthetic series are compared with those used in disaggregation and with those of the historical series in *Table 8*. It is shown that acceptable approximations of these statistics have been attained. The synthetic values tend to agree much better with the values used in disaggregation especially given the fact that the historical values have not been entered in the calculations and their preservation could not be assured. Nevertheless these discrepancies can be tolerated.

A further comparison is given in *figure 28* in terms of the autocorrelation function for higher lags, up to 10. It can be observed that even though in theory the synthetic autocorrelations should agree with those of the AR(1) model, they practically agree much better with the historical ones. In fact what forced the synthetic values to agree with the historical ones were the given hourly rainfall series at gages 1, 2, 3.

Hyetographs of the synthetic series given in *figures 29-33* show that the disaggregation model predicted the actual hyetographs very well.

Table 7

Statistics of hourly rainfall depths at each gage for the month of APRIL								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.88	0.90	0.90	0.90	0.90	0.91	-	-
value used on disaggregation	0.894	0.894	0.894	0.894	0.894	0.894	0.894	0.894
synthetic	0.88	0.89	0.90	*0.91*	*0.90*	*0.91*	*0.89*	*0.85*
Mean								
historical	0.17	0.13	0.12	0.11	0.11	0.12	-	-
value used on disaggregation	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
synthetic	0.17	0.13	0.13	*0.12*	*0.11*	*0.12*	*0.11*	*0.12*
Maximum value								
historical	24.40	19.40	16.20	14.40	13.40	16.20	-	-
value used on disaggregation	20.00	20.0	20.0	20.0	20.0	20.0	20.0	20.0
synthetic	24.40	19.40	16.20	*15*	*13.3*	*16.2*	*11.2*	*13.7*
Standard deviation								
historical	0.85	0.68	0.65	0.63	0.63	0.73	-	-
value used on disaggregation	0.729	0.729	0.729	0.729	0.729	0.729	0.729	0.729
synthetic	0.86	0.69	0.65	*0.60*	*0.58*	*0.67*	*0.54*	*0.55*
Skewness								
historical	11.18	11.61	10.29	10.63	11.55	10.77	-	-
value used on disaggregation	11.027	11.027	11.027	11.027	11.027	11.027	11.027	11.027
synthetic	11.16	11.60	10.27	*10.21*	*10.64*	*10.83*	*8.84*	*9.28*
Lag1 autocorrelation								
historical	0.44	0.40	0.38	0.58	0.53	0.51	-	-
value used on disaggregation	0.407	0.407	0.407	0.407	0.407	0.407	0.407	0.407
synthetic	0.44	0.40	0.38	*0.42*	*0.44*	*0.43*	*0.46*	*0.51*

Table 8

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of April

Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1								
historical	1.00	0.41	0.43	0.31	0.25	0.28	-	-
value used on disaggregation	1.00	0.41	0.43	0.47	0.26	0.37	0.32	0.44
synthetic	1.00	0.46	0.48	0.50	0.38	0.41	0.44	0.52
2								
historical	0.41	1.00	0.48	0.68	0.47	0.62	-	-
value used on disaggregation	0.41	1.00	0.48	0.85	0.52	0.77	0.53	0.77
synthetic	0.46	1.00	0.48	0.91	0.62	0.82	0.53	0.84
3								
historical	0.43	0.48	1.00	0.41	0.30	0.39	-	-
value used on disaggregation	0.43	0.48	1.00	0.47	0.33	0.41	0.51	0.45
synthetic	0.48	0.48	1.00	0.48	0.40	0.46	0.59	0.51
4								
historical	0.31	0.68	0.41	1.00	0.68	0.83	-	-
value used on disaggregation	0.47	0.85	0.47	1.00	0.52	0.81	0.40	0.90
synthetic	0.50	0.91	0.48	1.00	0.61	0.83	0.49	0.92
5								
historical	0.25	0.47	0.30	0.68	1.00	0.53	-	-
value used on disaggregation	0.26	0.52	0.33	0.52	1.00	0.39	0.32	0.45
synthetic	0.38	0.62	0.40	0.61	1.00	0.52	0.41	0.55
6								
historical	0.28	0.62	0.39	0.83	0.53	1.00	-	-
value used on disaggregation	0.37	0.77	0.41	0.81	0.39	1.00	0.33	0.77
synthetic	0.41	0.82	0.46	0.83	0.52	1.00	0.46	0.82
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.32	0.53	0.51	0.40	0.32	0.33	1.00	0.47
synthetic	0.44	0.53	0.59	0.49	0.41	0.46	1.00	0.59
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.44	0.77	0.45	0.90	0.45	0.77	0.47	1.00
synthetic	0.52	0.84	0.51	0.92	0.55	0.82	0.59	1.00

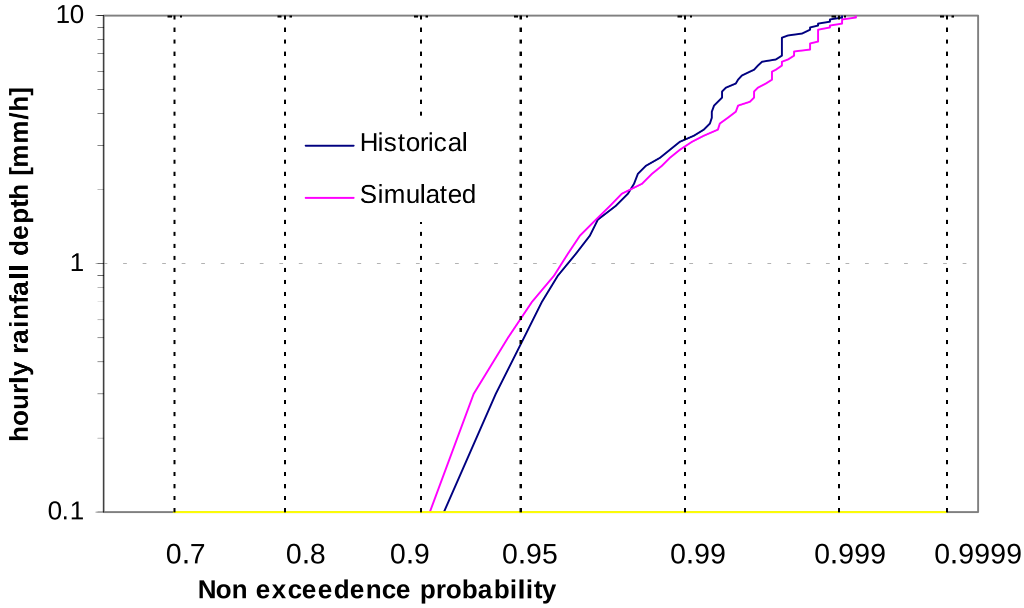


Figure 26: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 6 for the month of April

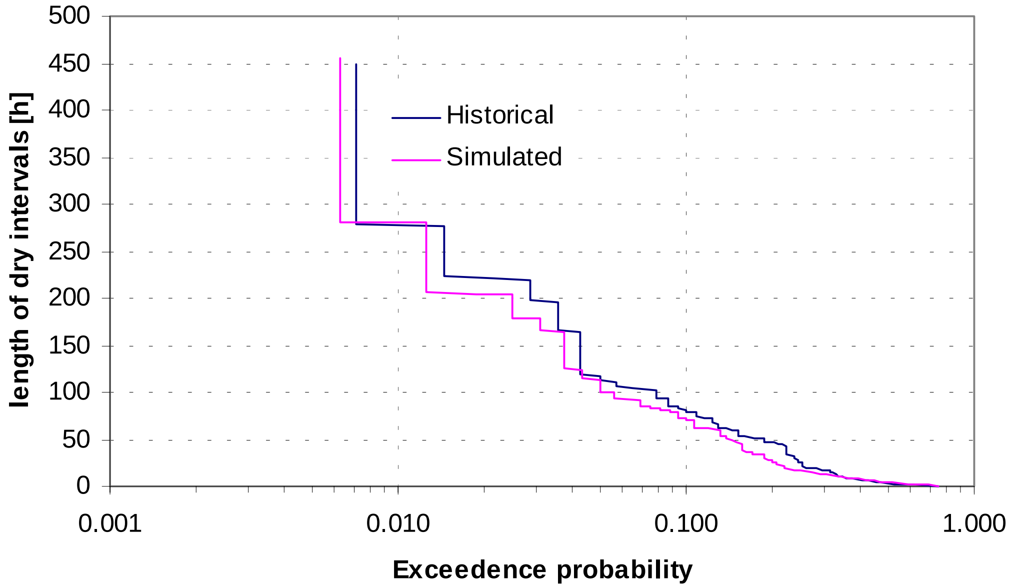


Figure 27: Comparison of historical and simulated probability distribution functions of of the length of dry intervals at gage 6 for the month of April

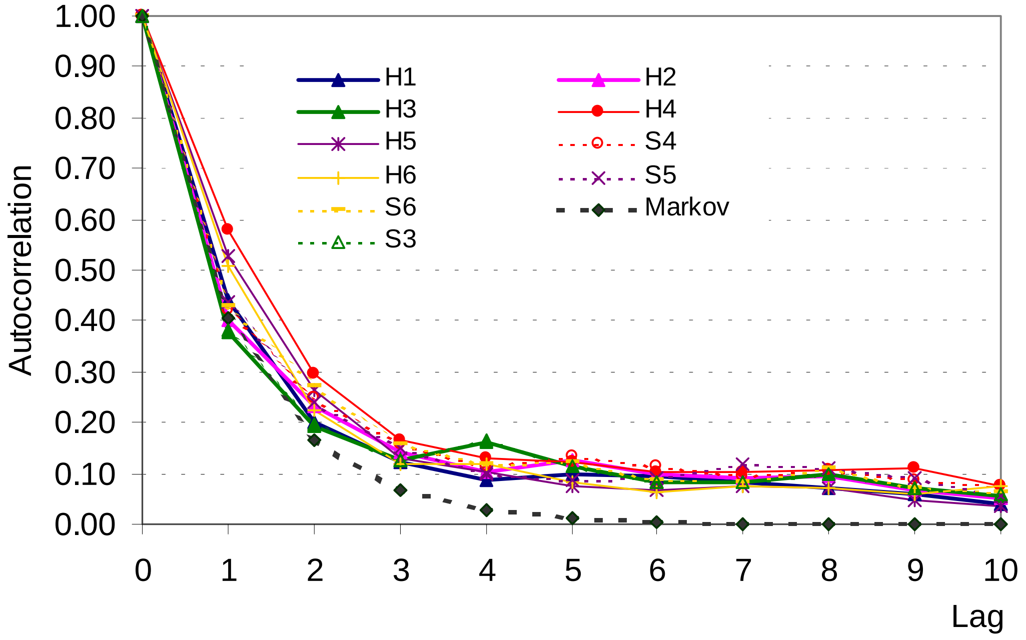


Figure 28: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of April.

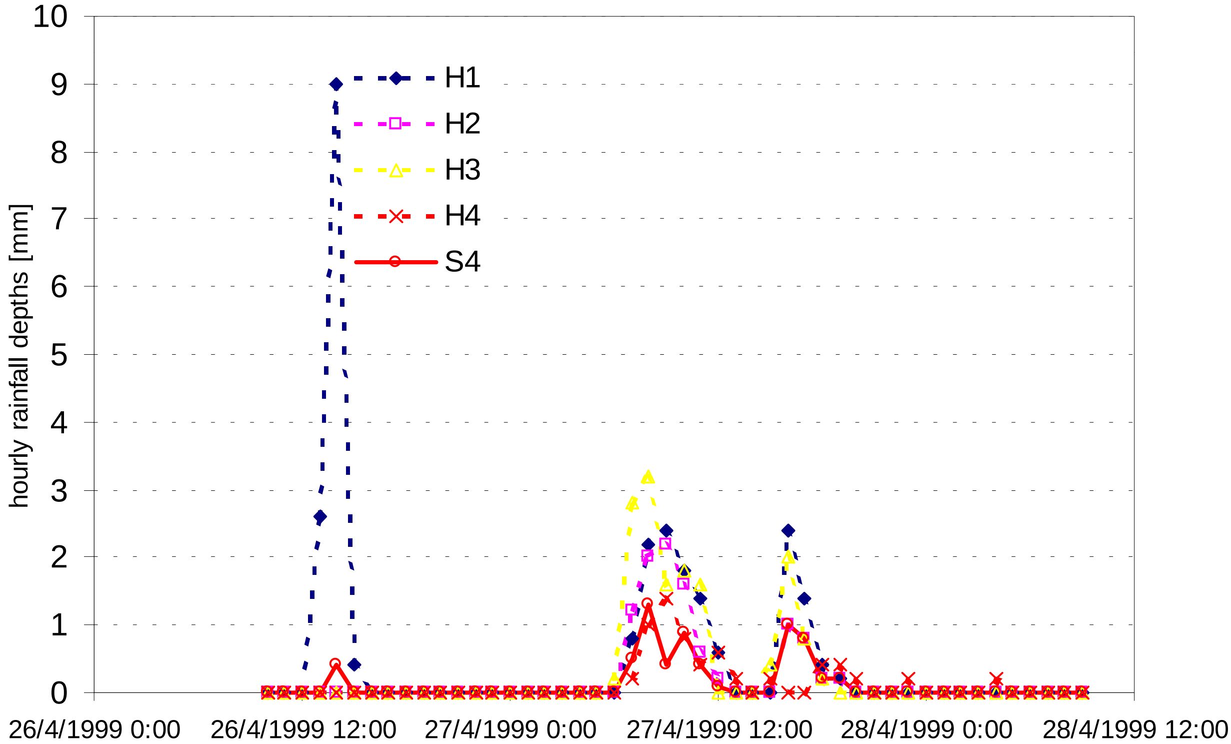


Figure 29: Comparison of historical and simulated hyetographs for raingage 4

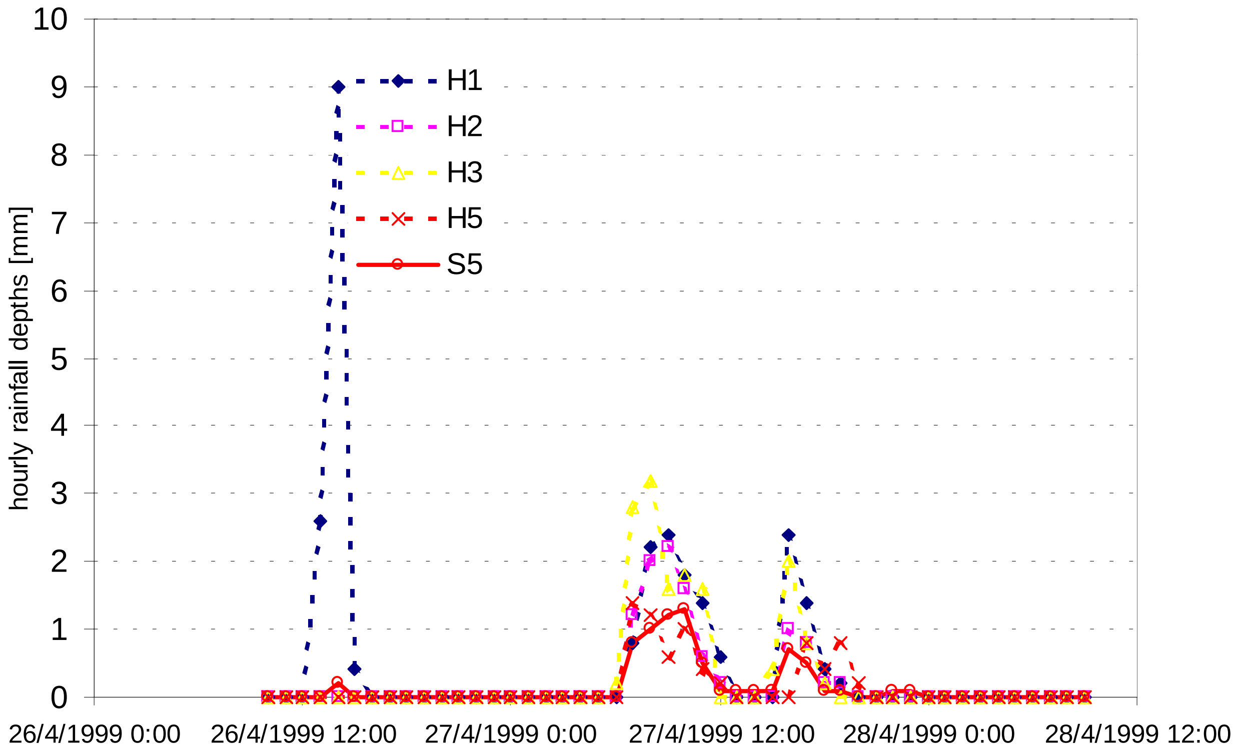


Figure 30: Comparison of historical and simulated hyetographs for raingage 5

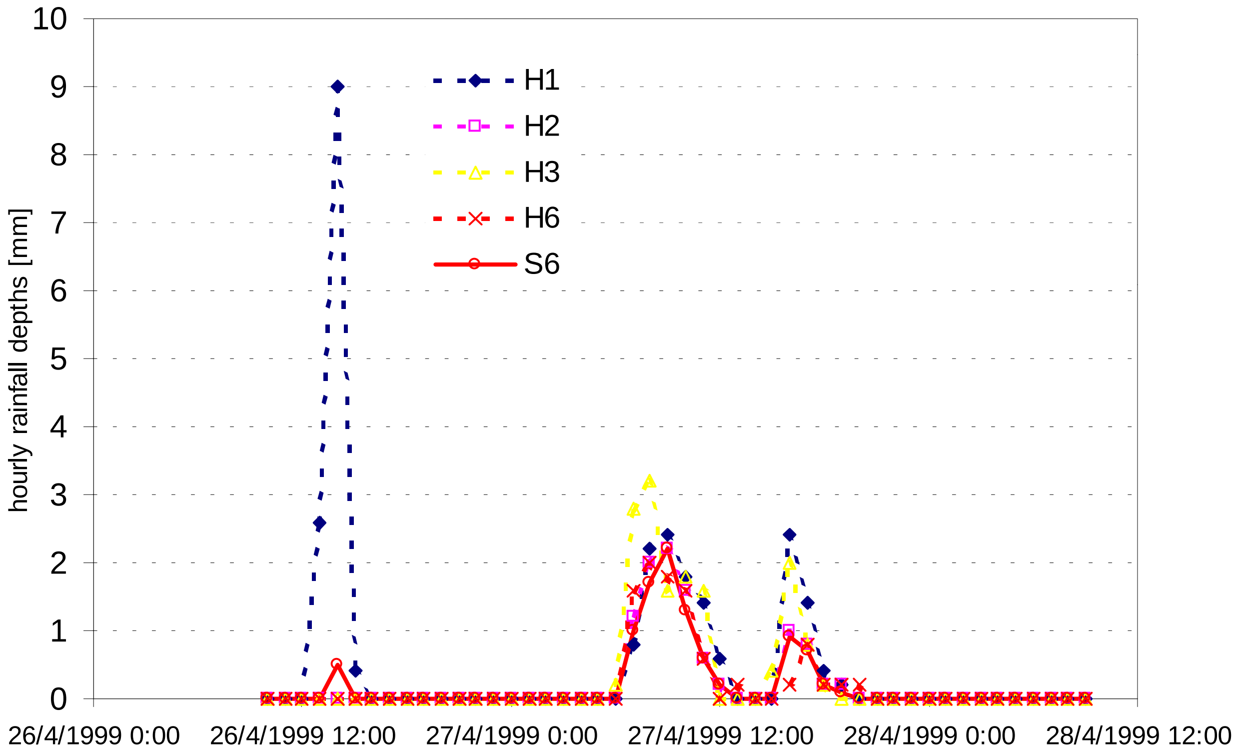


Figure 31: Comparison of historical and simulated hyetographs for raingage 6

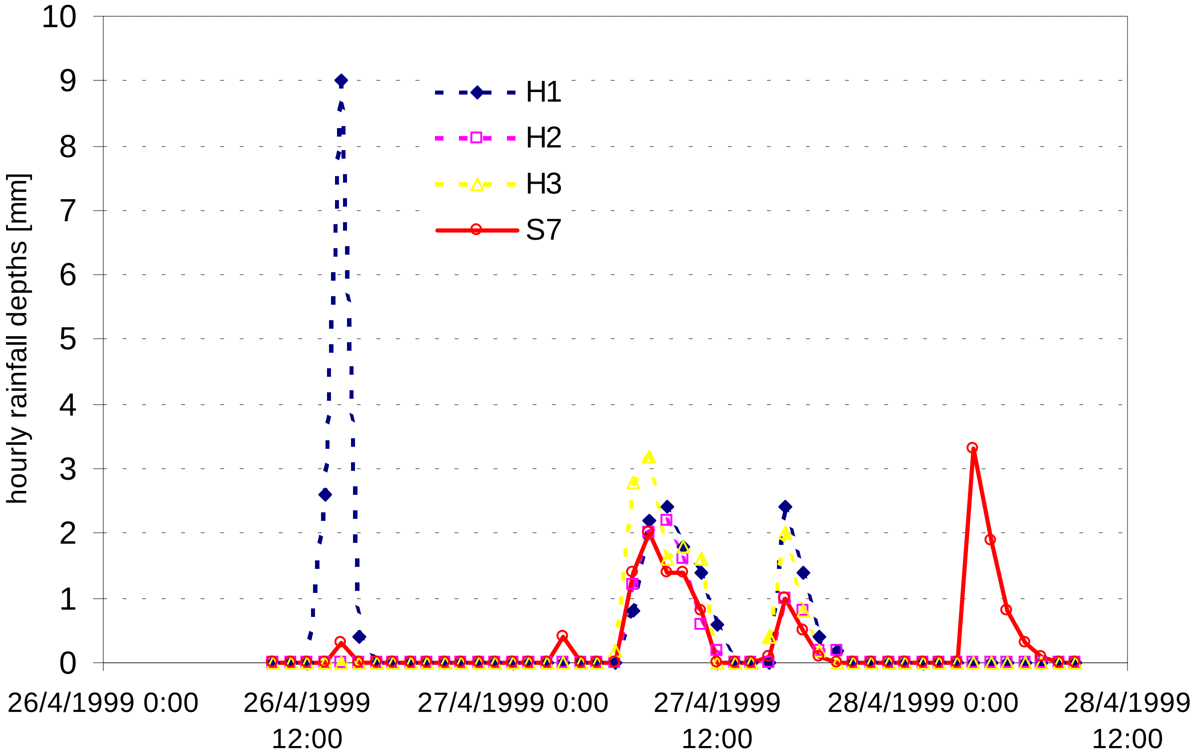


Figure 32: Simulated hyetographs for raingage 7

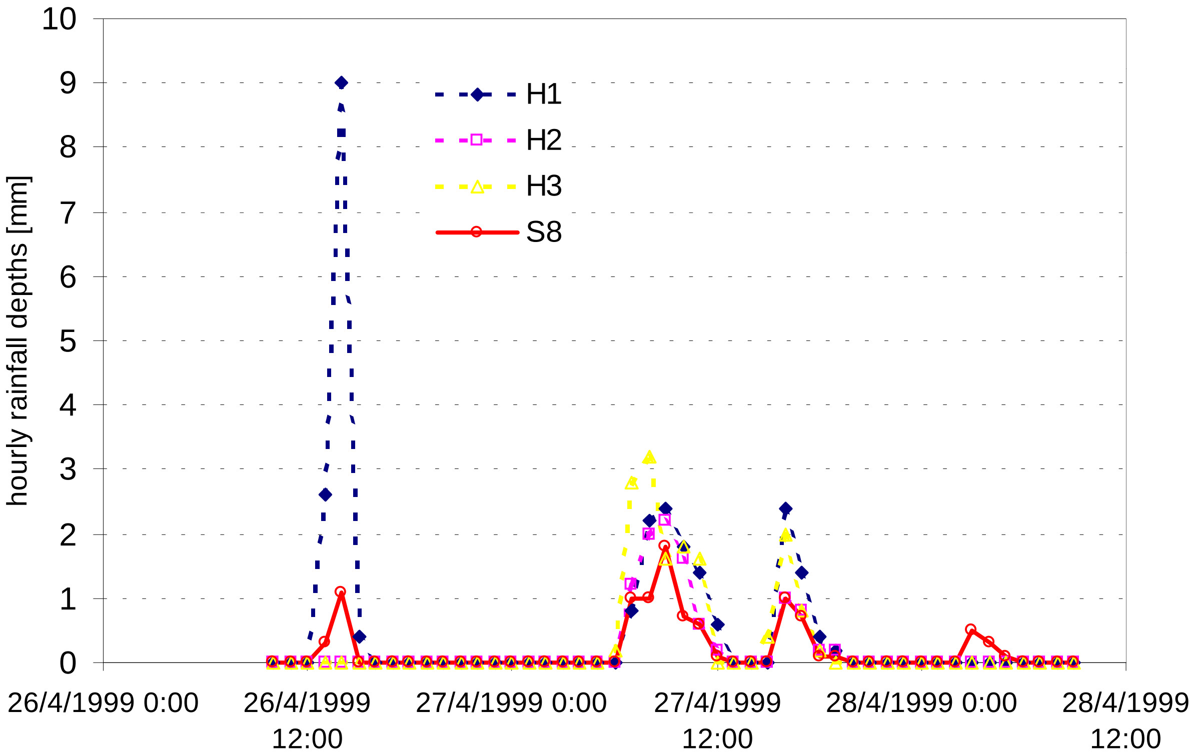


Figure 33: Simulated hyetographs for raingage 8

Month of May

For May, the simplified multivariate model was used in terms of linear transformation $\mathbf{X}_s = \mathbf{a}\,\mathbf{X}_{s-1} + \mathbf{b}\,\mathbf{V}_s$

Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 5000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π were set to 0.4 mm and 0.3 respectively.

Applying the disaggregation modeling framework synthetic hourly rainfall series were produced for the eight raingages, those of gages 1, 2 and 3 being identical to the historical series. The statistics of the synthetic series are compared to the historical and to the values used in the disaggregation in *table 7* where it can be observed a good agreement. Graphical comparisons show that the probability distribution functions of historical and simulated hourly rainfall depth during wet days for gage *4* are also in good agreement with each other (see *figures 34 and 35*).

Lag-one cross correlation coefficients of the synthetic series are compared with those used in disaggregation and with those of the historical series in *Table 8*. It is shown that acceptable approximations of these statistics have been attained. The synthetic values tend to agree much better with the values used in disaggregation especially given the fact that the historical values have not been entered in the calculations and their preservation could not be assured. Nevertheless these discrepancies can be tolerated.

A further comparison is given in *figure 36* in terms of the autocorrelation function for higher lags, up to 10. It can be observed that even though in theory the synthetic autocorrelations should agree with those of the AR(1) model, they practically agree much better with the historical ones. In fact what forced the synthetic values to agree with the historical ones were the given hourly rainfall series at gages 1, 2, 3.

Hyetographs of the synthetic series given in *figures 37-41* show that the disaggregation model predicted the actual hyetographs very well.

Table 9

Statistics of hourly rainfall depths at each gage for the month of MAY								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.94	0.94	0.95	0.94	0.94	0.95	-	-
value used on disaggregation	0.943	0.943	0.943	0.943	0.943	0.943	0.943	0.943
synthetic	0.94	0.94	0.95	0.95	0.95	0.96	0.95	0.88
Mean								
historical	0.10	0.08	0.08	0.09	0.07	0.07	-	-
value used on disaggregation	0.087	0.087	0.087	0.087	0.087	0.087	0.087	0.087
synthetic	0.10	0.08	0.08	0.09	0.07	0.08	0.09	0.10
Maximum value								
historical	33.8	19.8	15.8	31.6	19.8	20.8	-	-
value used on disaggregation	23.1	23.1	23.1	23.1	23.1	23.1	23.1	23.1
synthetic	33.8	19.8	15.8	23.1	15.1	20.6	14.7	17.3
Standard deviation								
historical	0.84	0.66	0.60	0.80	0.58	0.67	-	-
value used on disaggregation	0.703	0.703	0.703	0.703	0.703	0.703	0.703	0.703
synthetic	0.84	0.66	0.60	0.68	0.50	0.60	0.59	0.62
Skewness								
historical	20.67	16.04	14.57	21.94	17.64	18.44	-	-
value used on disaggregation	17.095	17.09	17.09	17.09	17.09	17.09	17.09	17.09
synthetic	20.63	15.93	14.54	17.09	13.47	17.32	12.12	13.01
Lag1 autocorrelation								
historical	0.25	0.38	0.38	0.29	0.30	0.32	-	-
value used on disaggregation	0.338	0.338	0.338	0.338	0.338	0.338	0.338	0.338
synthetic	0.25	0.38	0.39	0.39	0.38	0.37	0.42	0.45

Table 10

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of May

Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1								
historical	1.00	0.29	0.28	0.19	0.19	0.20	-	-
value used on disaggregation	1.00	0.29	0.28	0.29	0.27	0.21	0.48	0.23
synthetic	1.00	0.29	0.28	0.38	0.35	0.29	0.60	0.31
2								
historical	0.29	1.00	0.31	0.44	0.41	0.48	-	-
value used on disaggregation	0.29	1.00	0.31	0.49	0.47	0.46	0.41	0.38
synthetic	0.29	1.00	0.31	0.62	0.63	0.62	0.51	0.52
3								
historical	0.28	0.31	1.00	0.27	0.23	0.32	-	-
value used on disaggregation	0.28	0.31	1.00	0.36	0.32	0.36	0.46	0.38
synthetic	0.28	0.31	1.00	0.33	0.31	0.35	0.53	0.37
4								
historical	0.19	0.44	0.27	1.00	0.79	0.91	-	-
value used on disaggregation	0.29	0.49	0.36	1.00	0.65	0.87	0.52	0.85
synthetic	0.38	0.62	0.33	1.00	0.80	0.96	0.61	0.88
5								
historical	0.19	0.41	0.23	0.79	1.00	0.68		
value used on disaggregation	0.27	0.47	0.32	0.65	1.00	0.57	0.40	0.53
synthetic	0.35	0.63	0.31	0.80	1.00	0.78	0.54	0.70
6								
historical	0.20	0.48	0.32	0.91	0.68	1.00	-	-
value used on disaggregation	0.21	0.46	0.36	0.87	0.57	1.00	0.49	0.85
synthetic	0.29	0.62	0.35	0.96	0.78	1.00	0.58	0.87
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.48	0.41	0.46	0.52	0.40	0.49	1.00	0.60
synthetic	0.60	0.51	0.53	0.61	0.54	0.58	1.00	0.64
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.23	0.38	0.38	0.85	0.53	0.85	0.60	1.00
synthetic	0.31	0.52	0.37	0.88	0.70	0.87	0.64	1.00

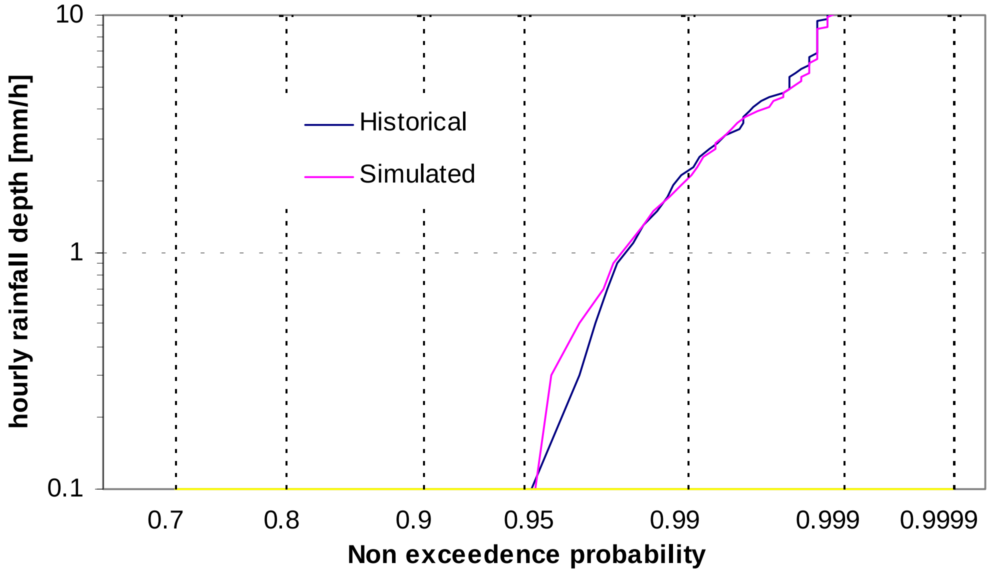


Figure 34: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 6 for the month of May

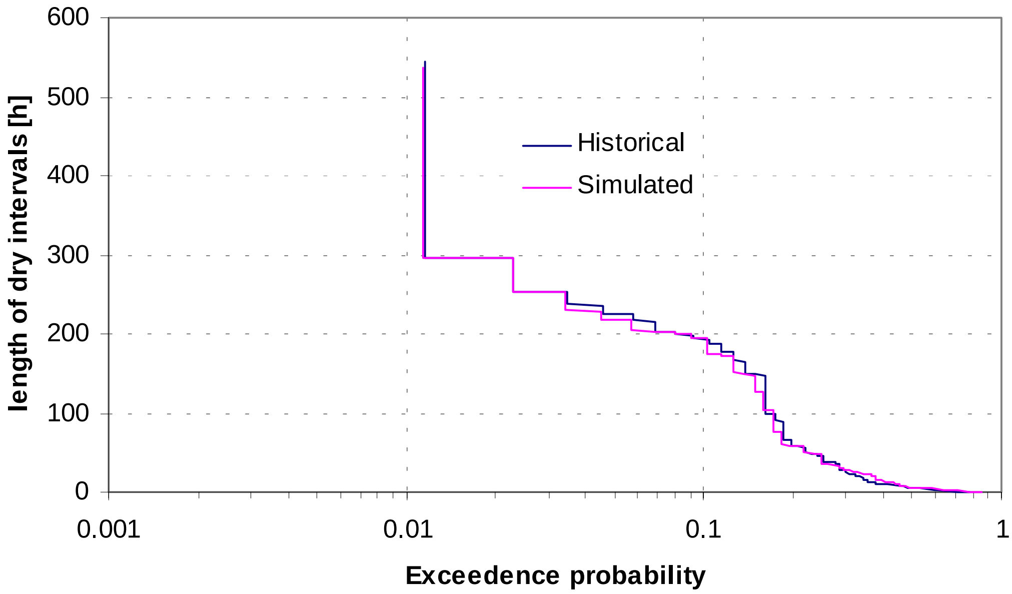


Figure 35 Comparison of historical and simulated probability distribution functions of of the length of dry intervals at gage 6 for the month of May

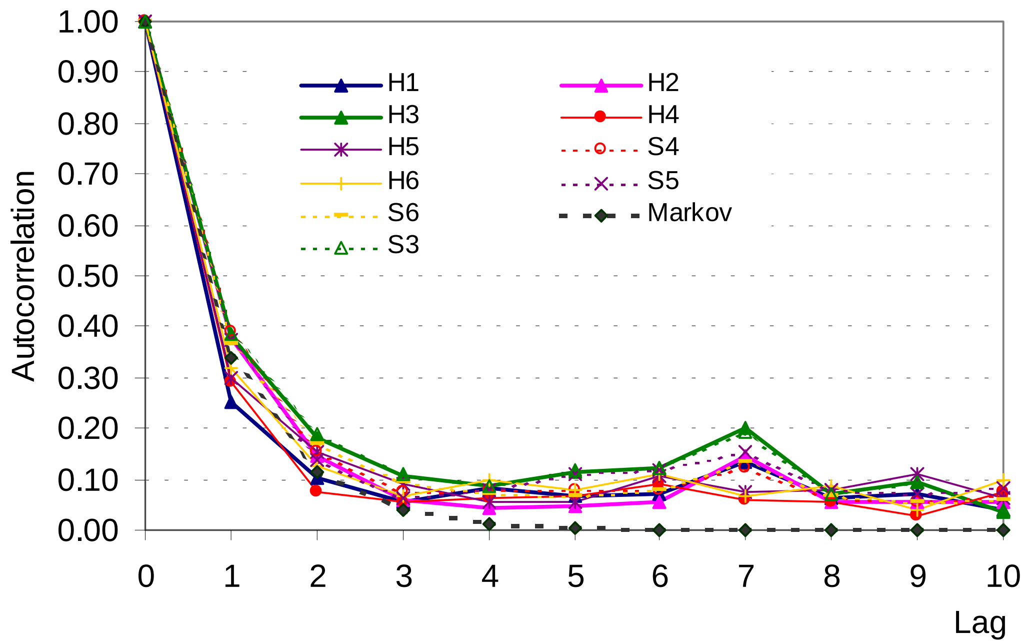


Figure 36: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of May.

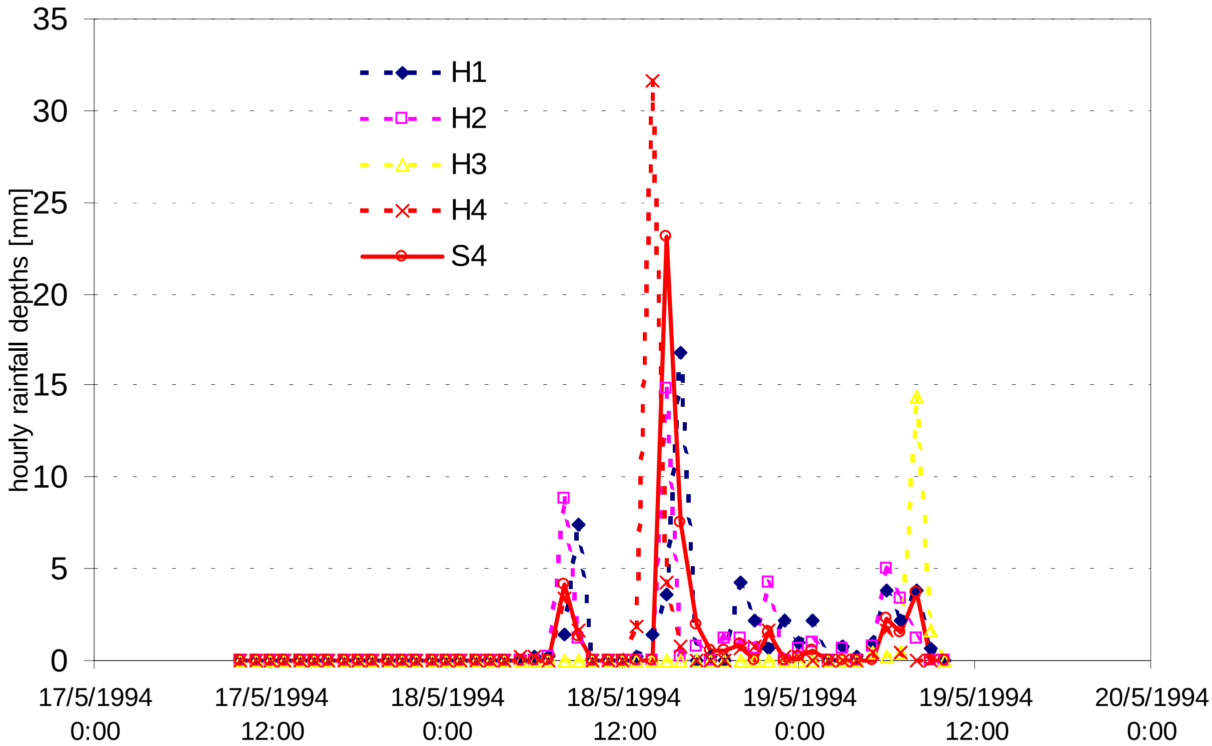


Figure 37: Comparison of historical and simulated hyetographs for raingage 4

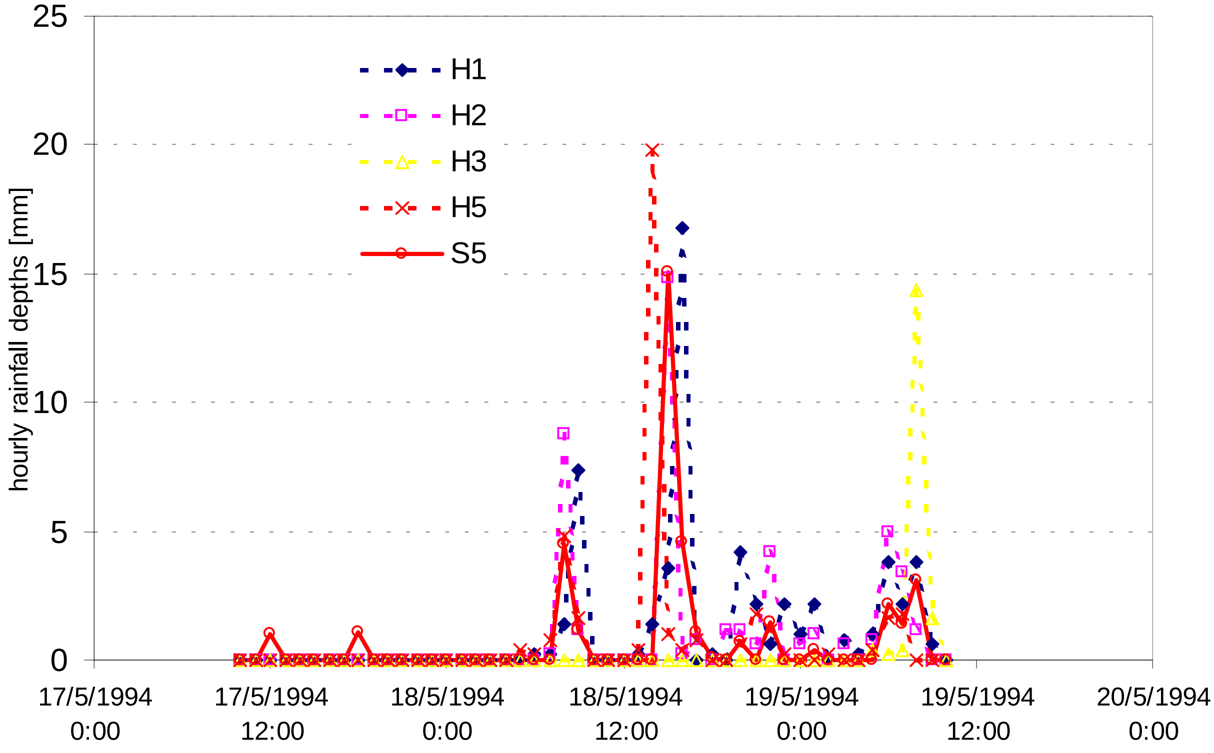


Figure 38: Comparison of historical and simulated hyetographs for raingage 5

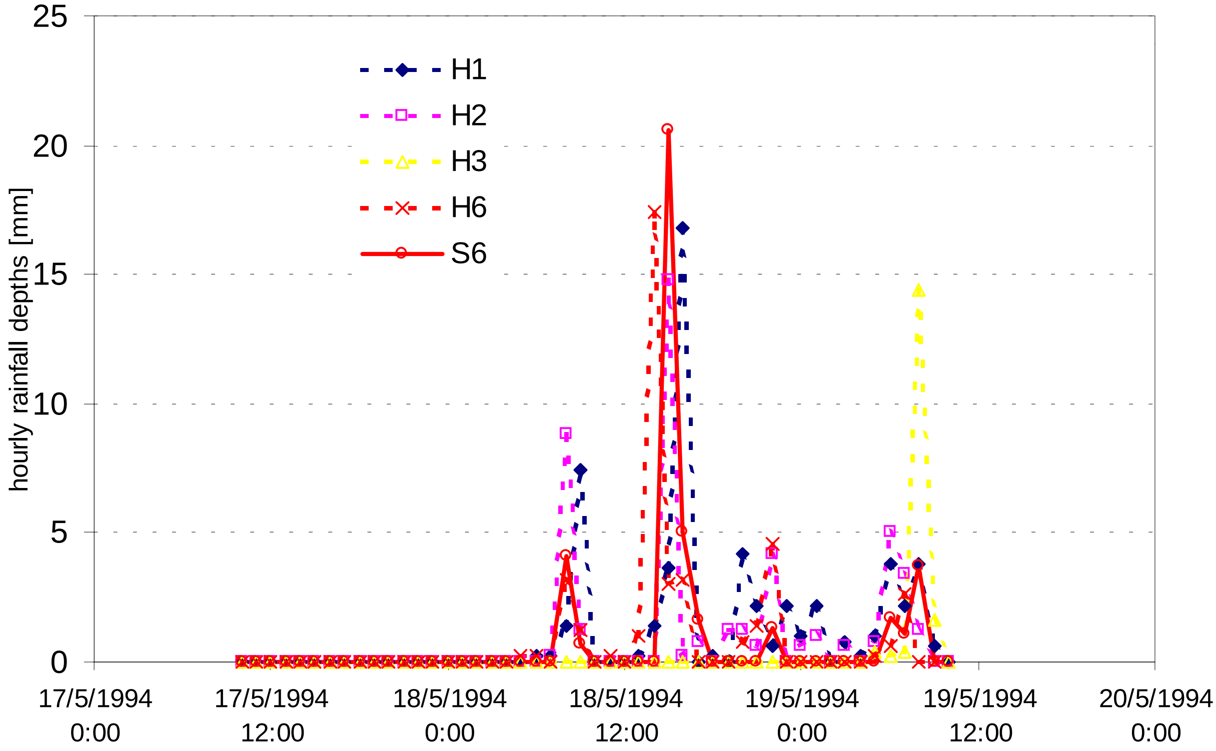


Figure 39: Comparison of historical and simulated hyetographs for raingage 6

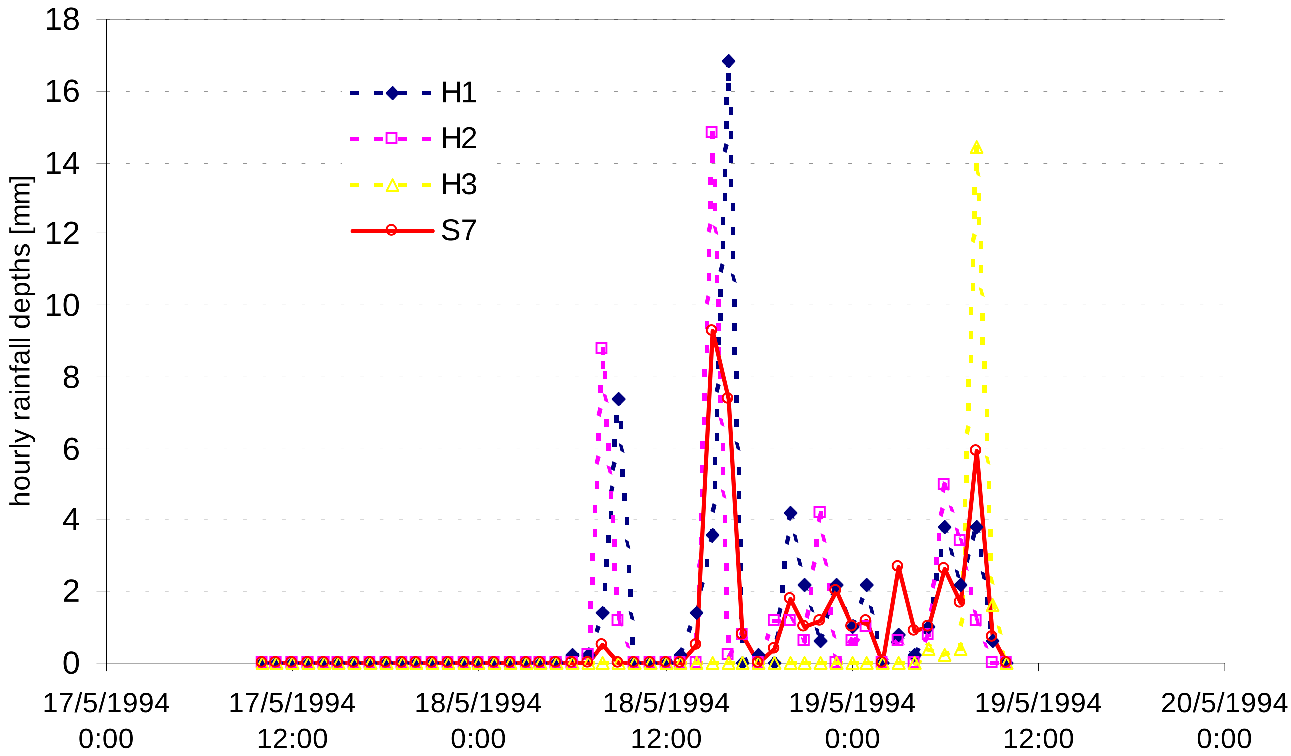


Figure 40: Simulated hyetographs for raingage 7

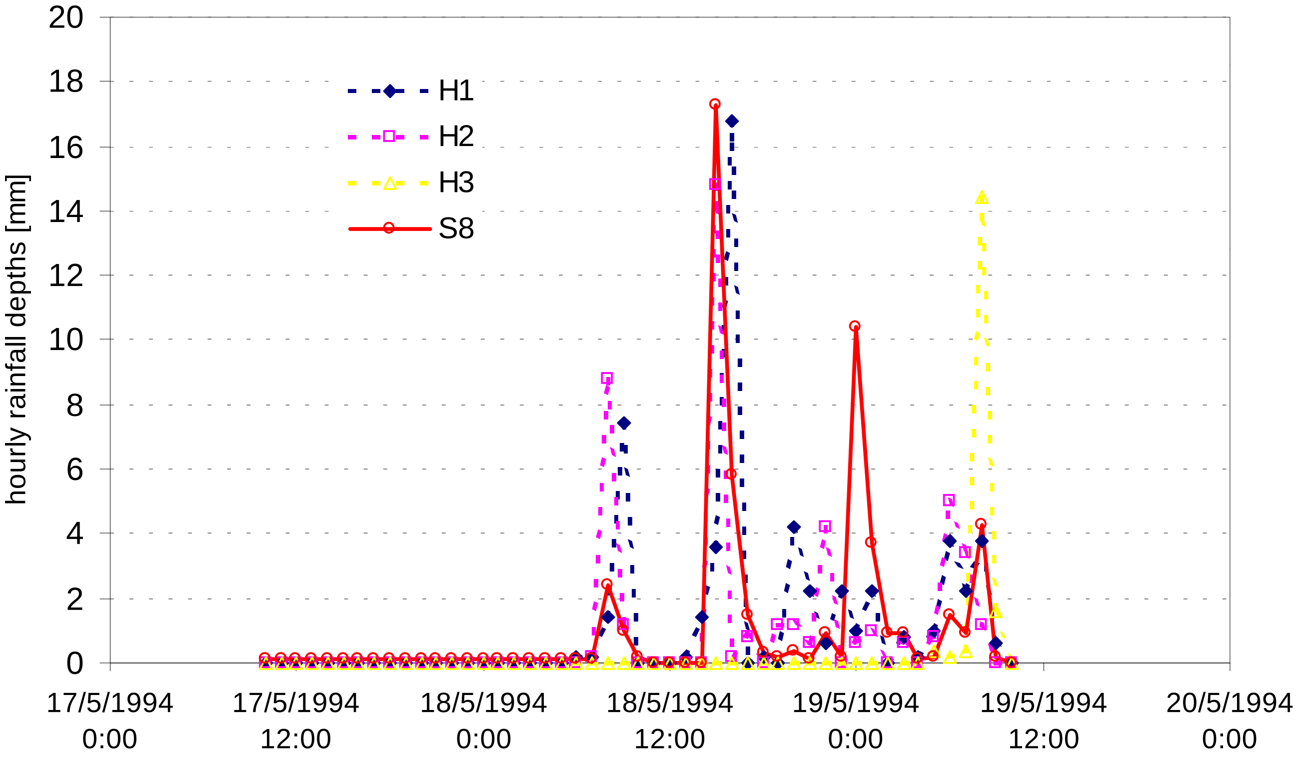


Figure 41: Simulated hyetographs for raingage 8

Month of June

For June, the simplified multivariate model was used in terms of power transformation $X_s^* = aX_{s-1}^* + bV_s$ where: $X_s^* = X_s^{(m)}$ and m=0.5

Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 3000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π_0 were set to 0.3 mm and 0.5 respectively.

Applying the disaggregation modeling framework synthetic hourly rainfall series were produced for the eight raingages, those of gages 1, 2 and 3 being identical to the historical series. The statistics of the synthetic series are compared to the historical and to the values used in the disaggregation in *table 11* where it can be observed a good agreement. Graphical comparisons show that the probability distribution functions of historical and simulated hourly rainfall depth during wet days for gage *5* are also in good agreement with each other (see *figures 42 and 43*).

Lag-one cross correlation coefficients of the synthetic series are compared with those used in disaggregation and with those of the historical series in *Table 12* It is shown that the synthetic values tend to agree much better with the values used in disaggregation especially given the fact that the historical values have not been entered in the calculations and their preservation could not be assured. Another fact worth mentioning is that the historical values of the cross correlations were extremely low and the essential hypothesis holding the entire framework, i.e. the significant spatial correlation between raingages, was not preserved. This fact could be encountered to the quality of the data , or to the particularities of the rainfall process during the month of June. A further comparison is given in *figure 44* in terms of the autocorrelation function for higher lags, up to 10. Hyetographs of the synthetic series given in *figures 45-49* show that the disaggregation model gives a good approximation of the actual hyetographs.

Statistics of hourly rainfall depths at each gage for the month of JUNE								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.97	0.95	0.97	0.97	0.98	0.97	-	-
value used on disaggregation	0.97	0.97	0.97	0.97	0.97	0.97	0.97	0.97
synthetic	0.97	0.95	0.97	0.97	0.98	0.97	0.97	0.94
Mean								
historical	0.05	0.05	0.05	0.05	0.04	0.04	-	-
value used on disaggregation	0.05	0.050	0.050	0.050	0.050	0.050	0.050	0.050
synthetic	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Maximum value								
historical	10.4	19	34.2	39.4	22	26.8	-	-
value used on disaggregation	21.2	21.2	21.2	21.2	21.2	21.2	21.2	21.2
synthetic	10.2	19	34.2	25	24.6	19.3	13.9	18.5
Standard deviation								
historical	0.47	0.52	0.74	0.74	0.59	0.56	-	-
value used on disaggregation	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58
synthetic	0.47	0.52	0.74	0.54	0.65	0.45	0.36	0.43
Skewness								
historical	15.348	23.903	30.672	39.735	23.450	32.329	-	-
value used on disaggregation	23.308	23.308	23.308	23.308	23.308	23.308	23.308	23.308
synthetic	15.28	23.86	30.59	29.00	25.99	25.26	22.76	24.99
Lag1 autocorrelation								
historical	0.24	0.43	0.26	0.07	0.15	0.16	-	-
value used on disaggregation	0.31	0.31	0.31	0.31	0.31	0.31	0.31	0.31
synthetic	0.24	0.43	0.26	0.31	0.28	0.31	0.38	0.33

Table 12

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of June

Gage	1	2	3	4	5	6	7	8
1								
historical	1.00	0.26	0.26	0.04	0.07	0.07	-	-
value used on disaggregation	1.00	0.26	0.26	0.20	0.37	0.27	0.29	0.30
synthetic	1.00	0.26	0.26	0.11	0.15	0.17	0.20	0.17
2								
historical	0.26	1.00	0.59	0.22	0.12	0.24	-	-
value used on disaggregation	0.26	1.00	0.59	0.54	0.34	0.46	0.18	0.52
synthetic	0.26	1.00	0.59	0.30	0.07	0.30	0.10	0.31
3								
historical	0.26	0.59	1.00	0.16	0.02	0.16	-	-
value used on disaggregation	0.26	0.59	1.00	0.34	0.13	0.33	0.19	0.37
synthetic	0.26	0.59	1.00	0.15	0.02	0.18	0.07	0.19
4								
historical	0.04	0.22	0.16	1.00	0.17	0.91	-	-
value used on disaggregation	0.20	0.54	0.34	1.00	0.23	0.89	0.53	0.92
synthetic	0.11	0.30	0.15	1.00	0.04	0.97	0.50	0.97
5								
historical	0.07	0.12	0.02	0.17	1.00	0.14	-	-
value used on disaggregation	0.37	0.34	0.13	0.23	1.00	0.15	0.21	0.30
synthetic	0.15	0.07	0.02	0.04	1.00	0.03	0.09	0.08
6								
historical	0.07	0.24	0.16	0.91	0.14	1.00	-	-
value used on disaggregation	0.27	0.46	0.33	0.89	0.15	1.00	0.53	0.89
synthetic	0.17	0.30	0.18	0.97	0.03	1.00	0.50	0.97
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.29	0.18	0.19	0.53	0.21	0.53	1.00	0.51
synthetic	0.20	0.10	0.07	0.50	0.09	0.50	1.00	0.51
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.30	0.52	0.37	0.92	0.30	0.89	0.51	1.00
synthetic	0.17	0.31	0.19	0.97	0.08	0.97	0.51	1.00

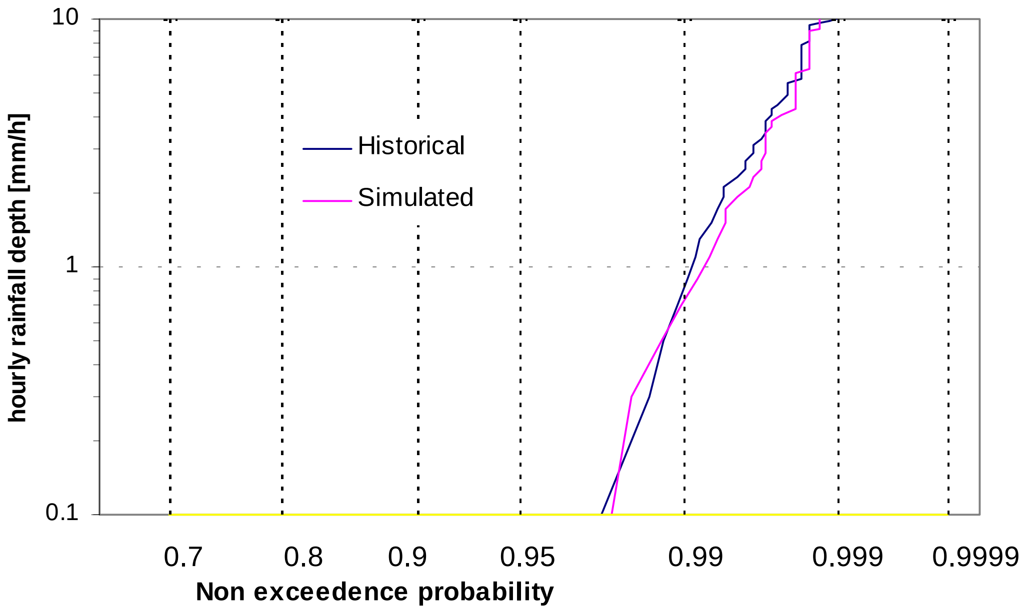


Figure 42: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 5 for the month of June

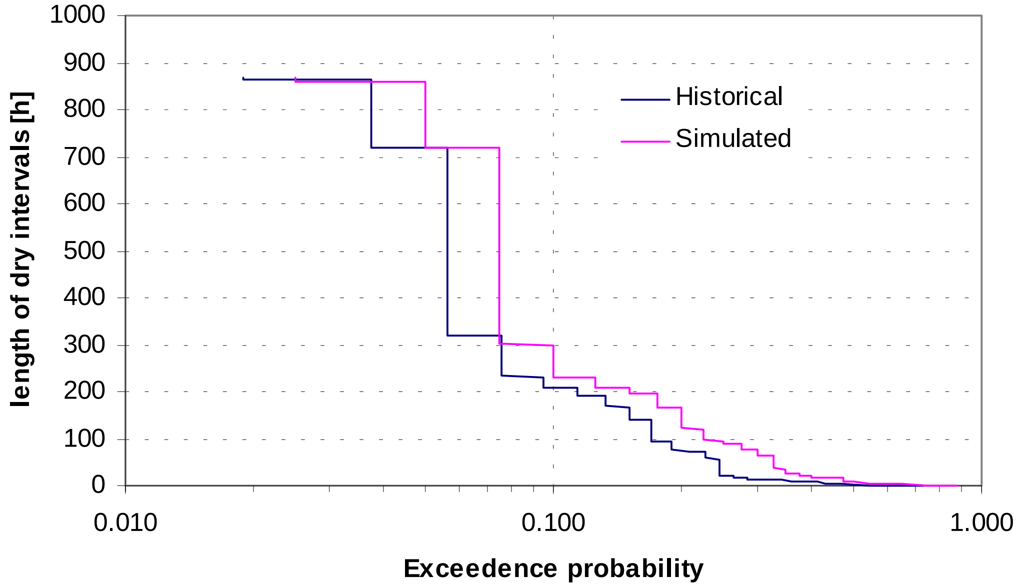


Figure 43: Comparison of historical and simulated probability distribution functions of of the length of dry intervals at gage 5 for the month of June

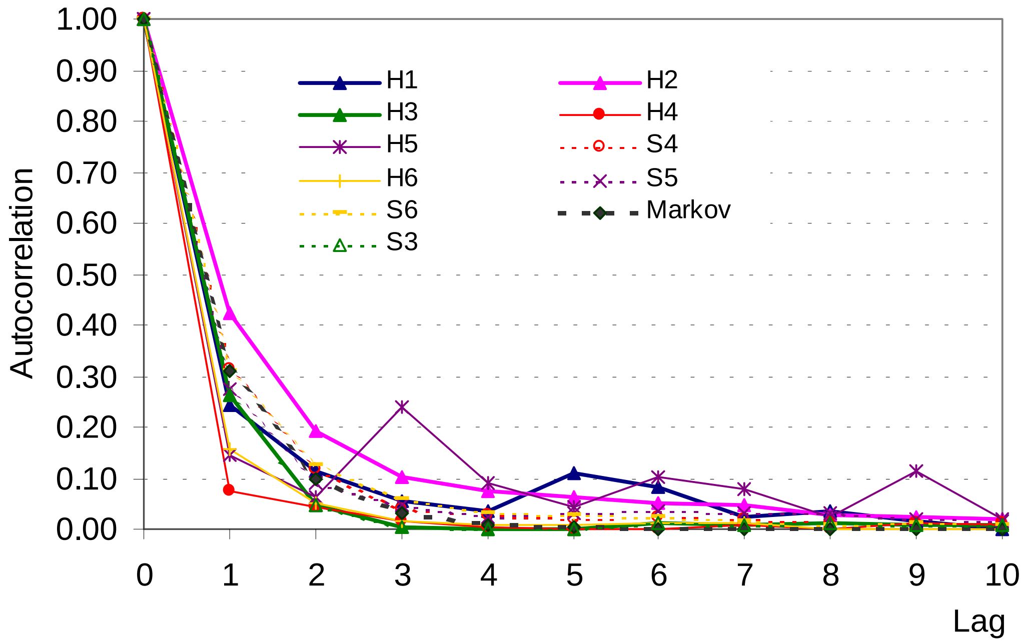


Figure 44: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of June.

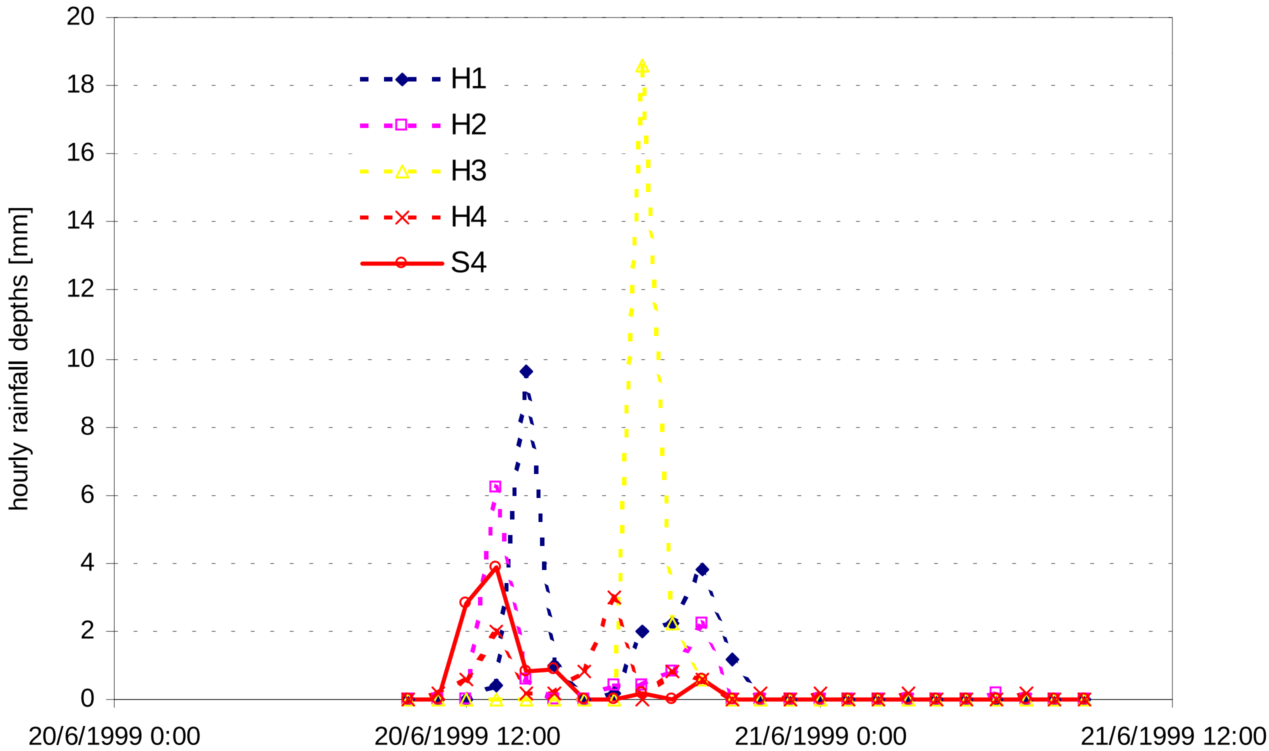


Figure 45: Comparison of historical and simulated hyetographs for raingage 4

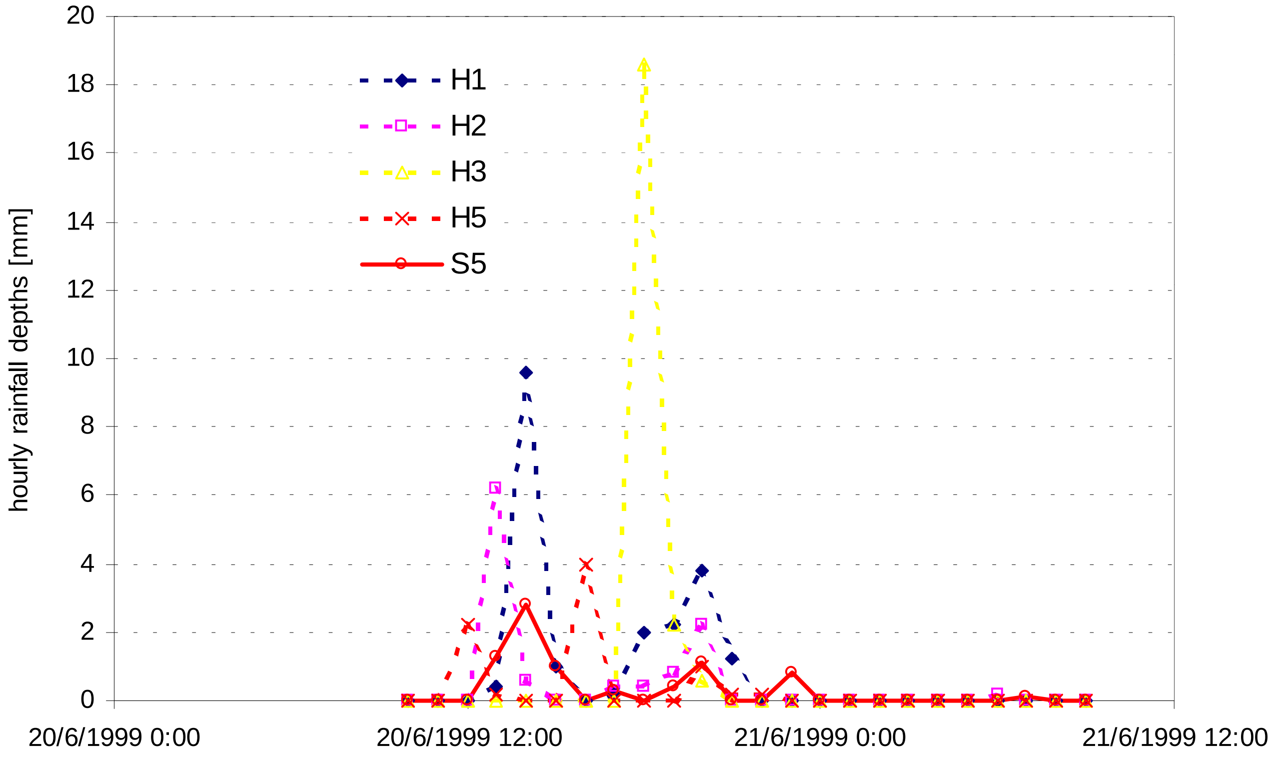


Figure 46: Comparison of historical and simulated hyetographs for raingage 5

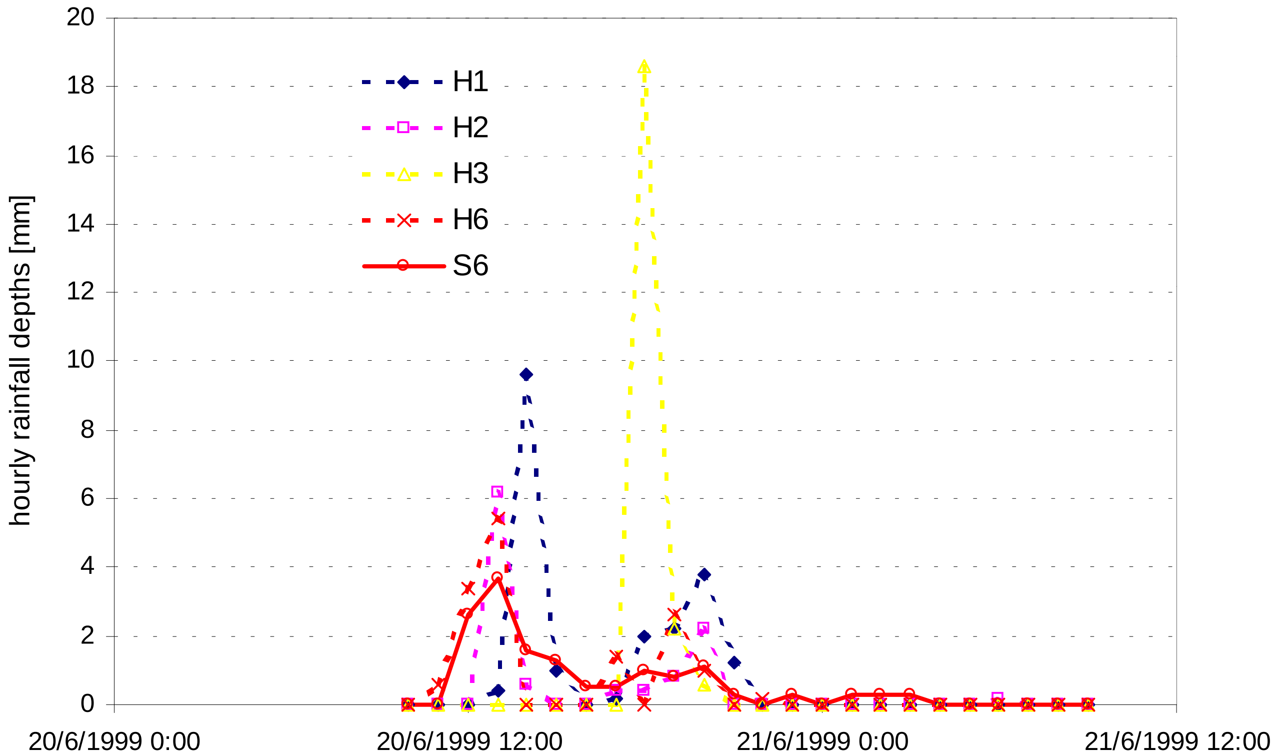


Figure 47: Comparison of historical and simulated hyetographs for raingage 6

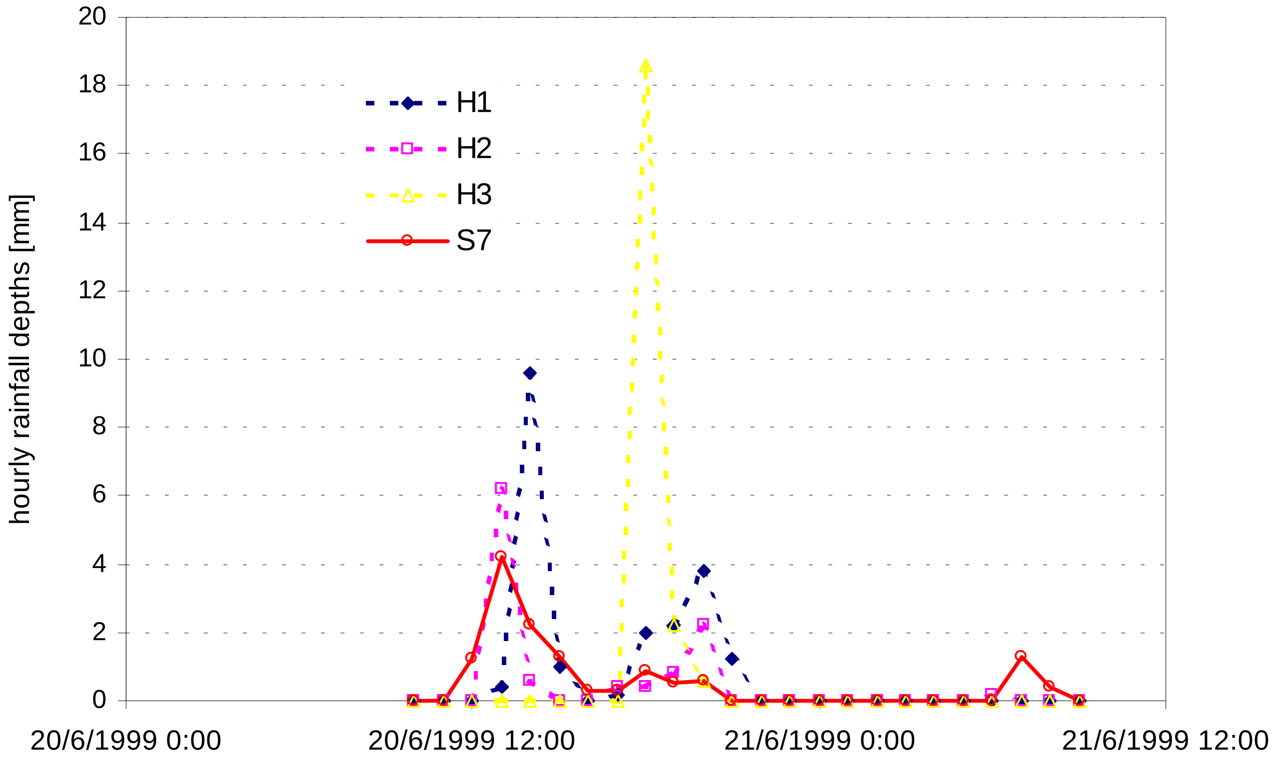


Figure 48: Simulated hyetographs for raingage 7

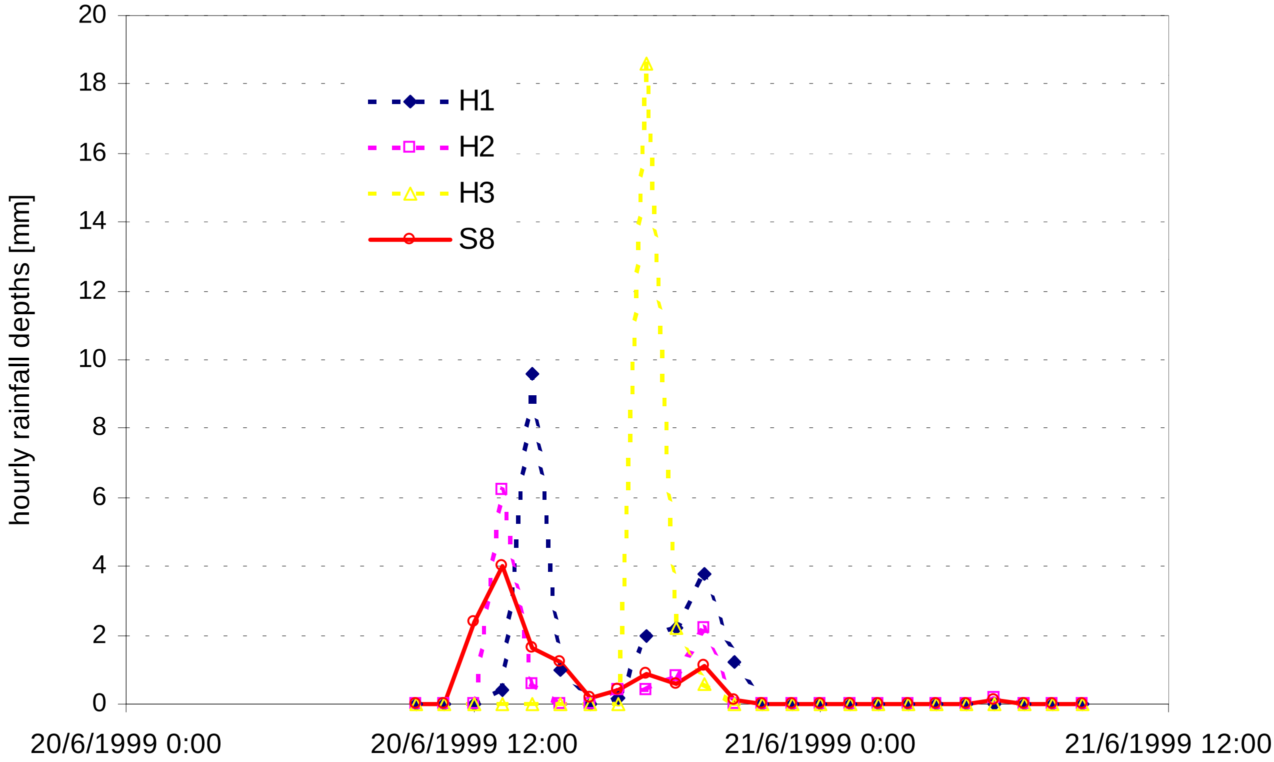


Figure 49: Simulated hyetographs for raingage 8

Month of July

For July, the simplified multivariate model was used in terms of power transformation $X_s^* = aX_{s-1}^* + bV_s$ where: $X_s^* = X_s^{(m)}$ and m=0.6

Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 3000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π_0 were set to 0.3 mm and 0.5 respectively.

Applying the disaggregation modeling framework synthetic hourly rainfall series were produced for the eight raingages, those of gages 1, 2 and 3 being identical to the historical series. The statistics of the synthetic series are compared to the historical and to the values used in the disaggregation in *table 13* where it can be observed a good agreement. Graphical comparisons show that the probability distribution functions of historical and simulated hourly rainfall depth during wet days for gage *4* are also in good agreement with each other (see *figures 50 and 51*).

Lag-one cross correlation coefficients of the synthetic series are compared with those used in disaggregation and with those of the historical series in *Table 14* It is shown that the synthetic values tend to agree much better with the values used in disaggregation especially given the fact that the historical values have not been entered in the calculations and their preservation could not be assured. Another fact worth mentioning is that the historical values of the cross correlations were extremely low and the essential hypothesis holding the entire framework, i.e. the significant spatial correlation between raingages, was not preserved. This fact could be encountered to the quality of the data , or to the particularities of the rainfall process during the month of July. The rainfall process in summer is characterized by intense storms of relatively small duration and extremely localized.

A further comparison is given in *figure 52* in terms of the autocorrelation function for higher lags, up to 10. Hyetographs of the synthetic series given in *figures 53-57* show that the disaggregation model gives a good approximation of the actual hyetographs.

Table 13

Statistics of hourly rainfall depths at each gage for the month of JULY								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.99	0.98	0.99	0.99	0.99	0.99	-	-
value used on disaggregation	0.984	0.984	0.984	0.984	0.984	0.984	0.984	0.984
synthetic	0.99	0.98	0.99	*0.98*	*0.99*	*0.99*	*0.98*	*0.97*
Mean								
historical	0.05	0.04	0.04	0.03	0.02	0.01	-	-
value used on disaggregation	0.041	0.041	0.041	0.041	0.041	0.041	0.041	0.041
synthetic	0.05	0.04	0.04	*0.03*	*0.02*	*0.01*	*0.04*	*0.03*
Maximum value								
historical	38.2	32	37.4	67.8	23	12.2	-	-
value used on disaggregation	35.9	35.9	35.9	35.9	35.9	35.9	35.9	35.9
synthetic	38.2	32	37.4	*68.4*	*23.4*	*15.6*	*46.5*	*67.7*
Standard deviation								
historical	0.87	0.66	0.88	1.13	0.46	0.30	-	-
value used on disaggregation	0.803	0.803	0.803	0.803	0.803	0.803	0.803	0.803
synthetic	0.87	0.66	0.88	*1.10*	*0.46*	*0.30*	*0.91*	*1.09*
Skewness								
historical	30.43	34.59	32.20	52.87	36.95	34.32	-	-
value used on disaggregation	32.176	32.176	32.176	32.176	32.176	32.176	32.176	32.176
synthetic	30.43	34.62	32.20	*56.09*	*38.69*	*39.24*	*39.00*	*56.44*
Lag1 autocorrelation								
historical	0.11	0.13	0.30	0.39	0.10	0.18	-	-
value used on disaggregation	0.181	0.181	0.181	0.181	0.181	0.181	0.181	0.181
synthetic	0.11	0.13	0.30	*0.33*	*0.16*	*0.06*	*0.36*	*0.34*

Table 14

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of July

Gage	1	2	3	4	5	6	7	8
1								
historical	1.00	0.24	0.17	0.02	0.04	0.04	-	-
value used on disaggregation	1.00	0.24	0.17	0.08	0.05	0.09	0.08	0.05
synthetic	1.00	0.25	0.17	0.05	0.02	0.07	0.07	0.03
2								
historical	0.24	1.00	0.33	0.70	0.58	0.20	-	-
value used on disaggregation	0.24	1.00	0.33	0.50	0.40	0.31	0.54	0.49
synthetic	0.25	1.00	0.34	0.72	0.60	0.61	0.66	0.72
3								
historical	0.17	0.34	1.00	0.49	0.31	0.44	-	-
value used on disaggregation	0.17	0.33	1.00	0.52	0.37	0.24	0.62	0.52
synthetic	0.17	0.34	1.00	0.46	0.31	0.2027	0.56	0.46
4								
historical	0.02	0.70	0.49	1.00	0.74	0.48	-	-
value used on disaggregation	0.08	0.50	0.52	1.00	0.62	0.57	0.79	0.99
synthetic	0.05	0.72	0.46	1.00	0.76	0.76	0.85	1.00
5								
historical	0.04	0.58	0.31	0.74	1.00	0.22	-	-
value used on disaggregation	0.05	0.40	0.37	0.62	1.00	0.33	0.52	0.63
synthetic	0.02	0.60	0.31	0.76	1.00	0.60	0.65	0.76
6								
historical	0.04	0.20	0.44	0.48	0.22	1.00	-	-
value used on disaggregation	0.09	0.31	0.24	0.57	0.33	1.00	0.39	0.57
synthetic	0.07	0.61	0.2027	0.76	0.60	1.00	0.60	0.75
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.08	0.54	0.62	0.79	0.52	0.39	1.00	0.80
synthetic	0.07	0.66	0.56	0.85	0.65	0.60	1.00	0.85
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.05	0.49	0.52	0.99	0.63	0.57	0.80	1.00
synthetic	0.03	0.72	0.46	1.00	0.76	0.75	0.85	1.00

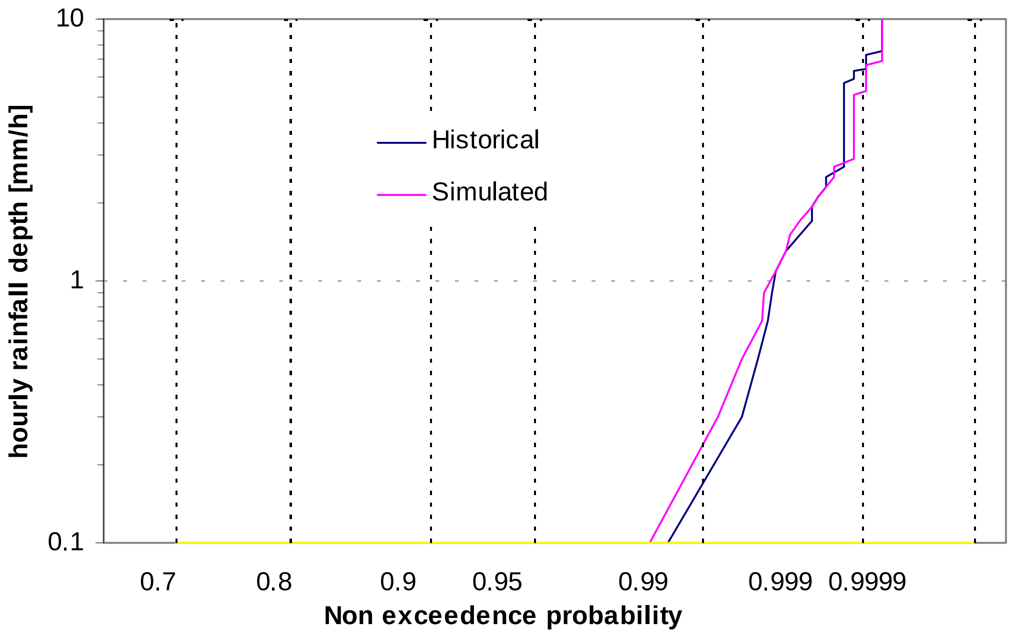


Figure 50: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 4 for the month of July

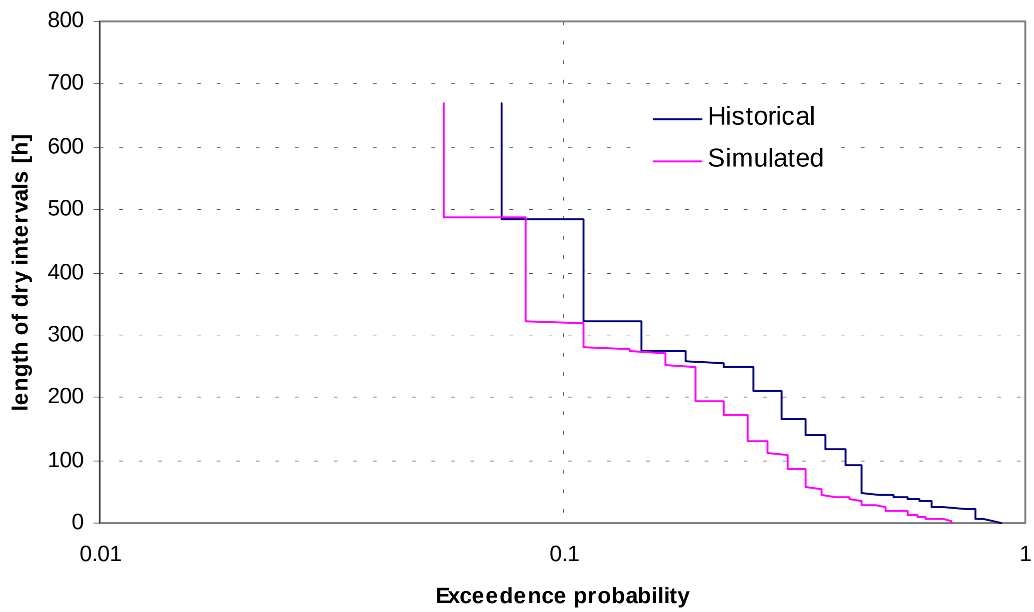


Figure 51: Comparison of historical and simulated probability distribution functions of of the length of dry intervals at gage 4 for the month of July

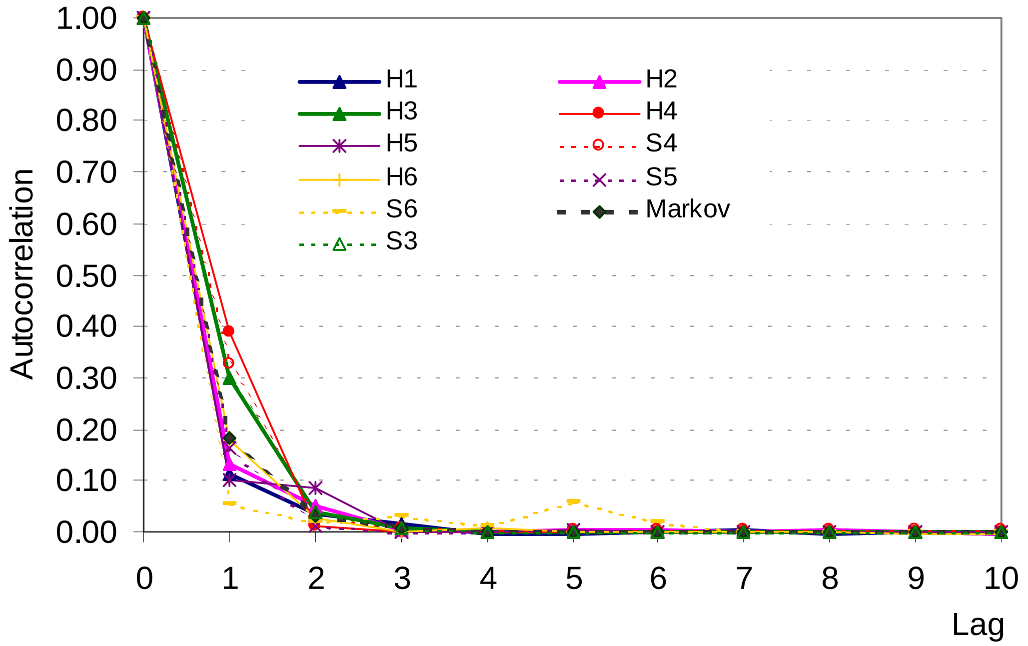


Figure 52: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of July.

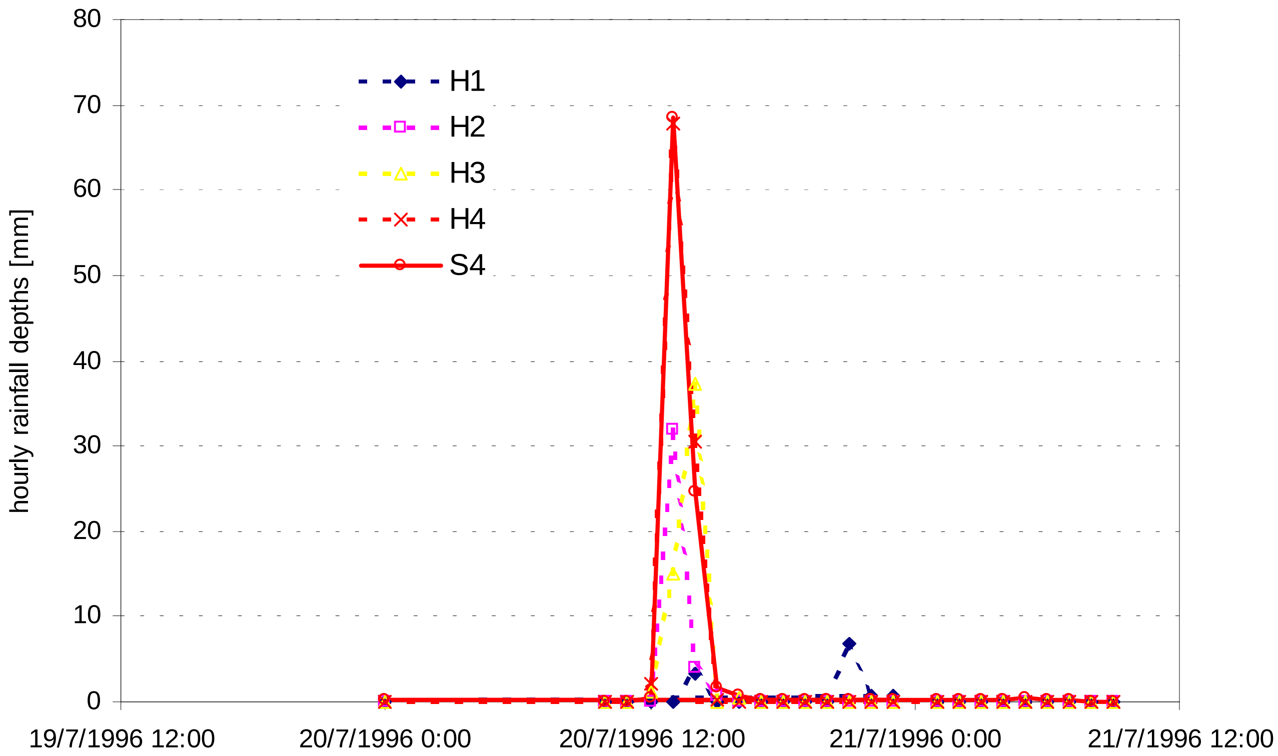


Figure 53: Comparison of historical and simulated hyetographs for raingage 4

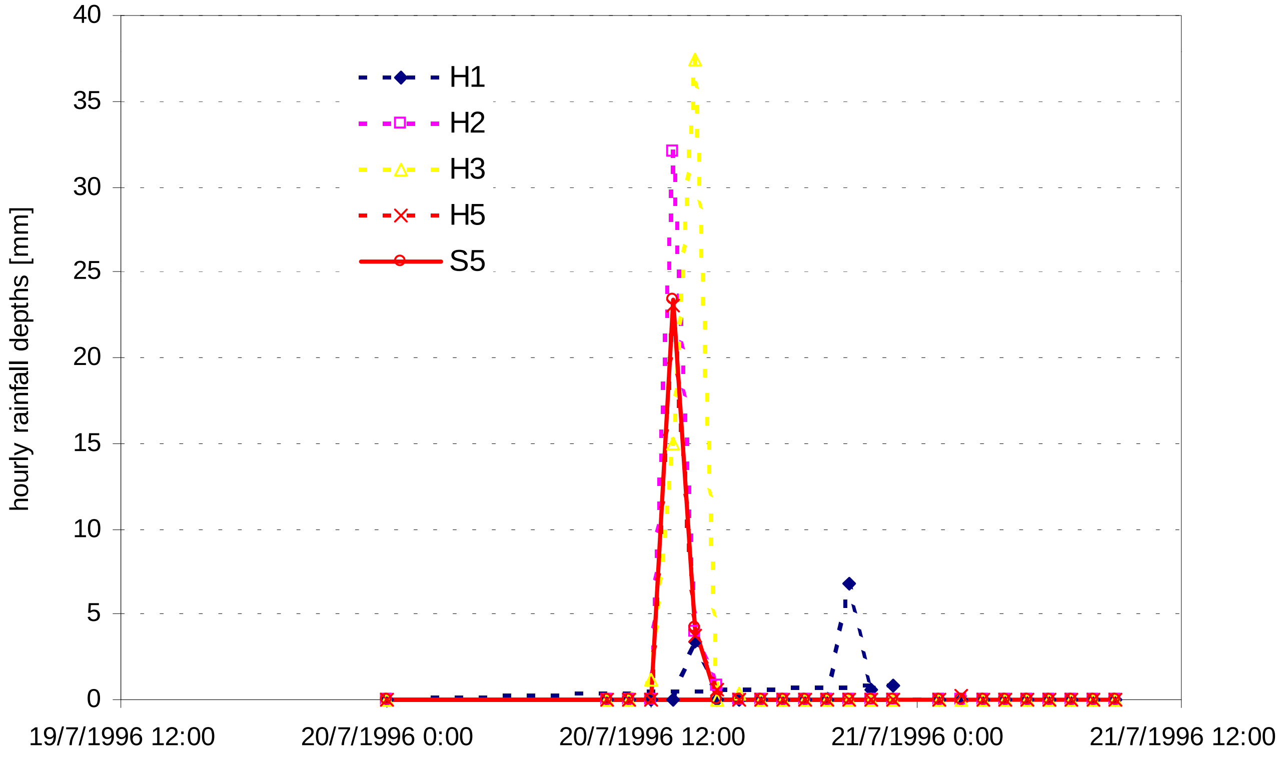


Figure 54: Comparison of historical and simulated hyetographs for raingage 5

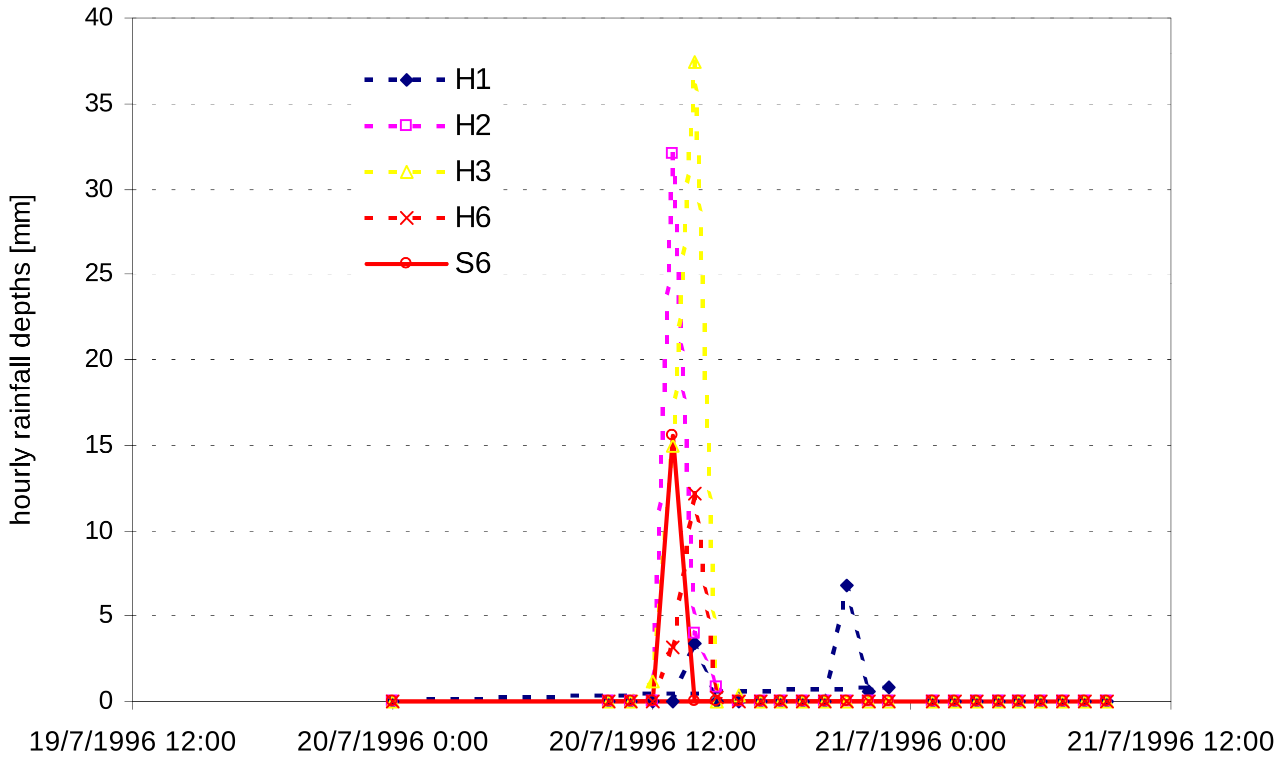


Figure 55: Comparison of historical and simulated hyetographs for raingage 6

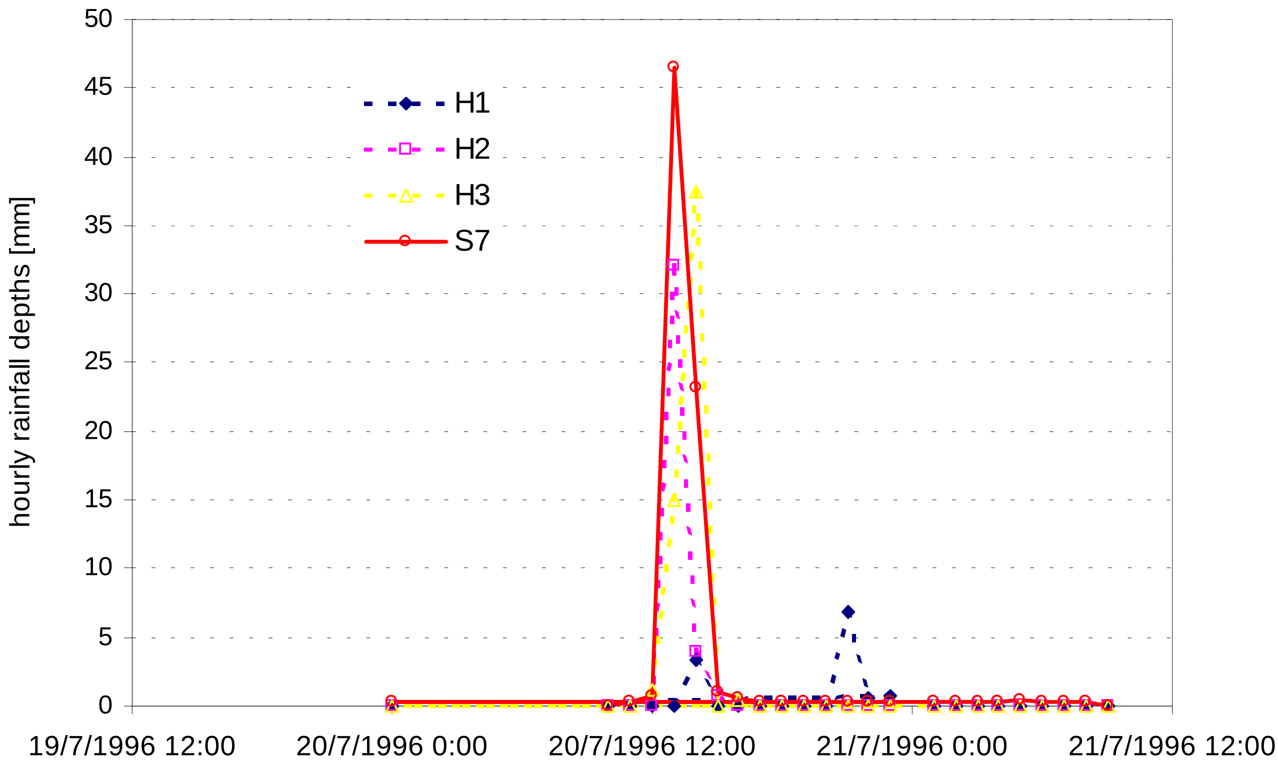


Figure 56: Simulated hyetographs for raingage 7

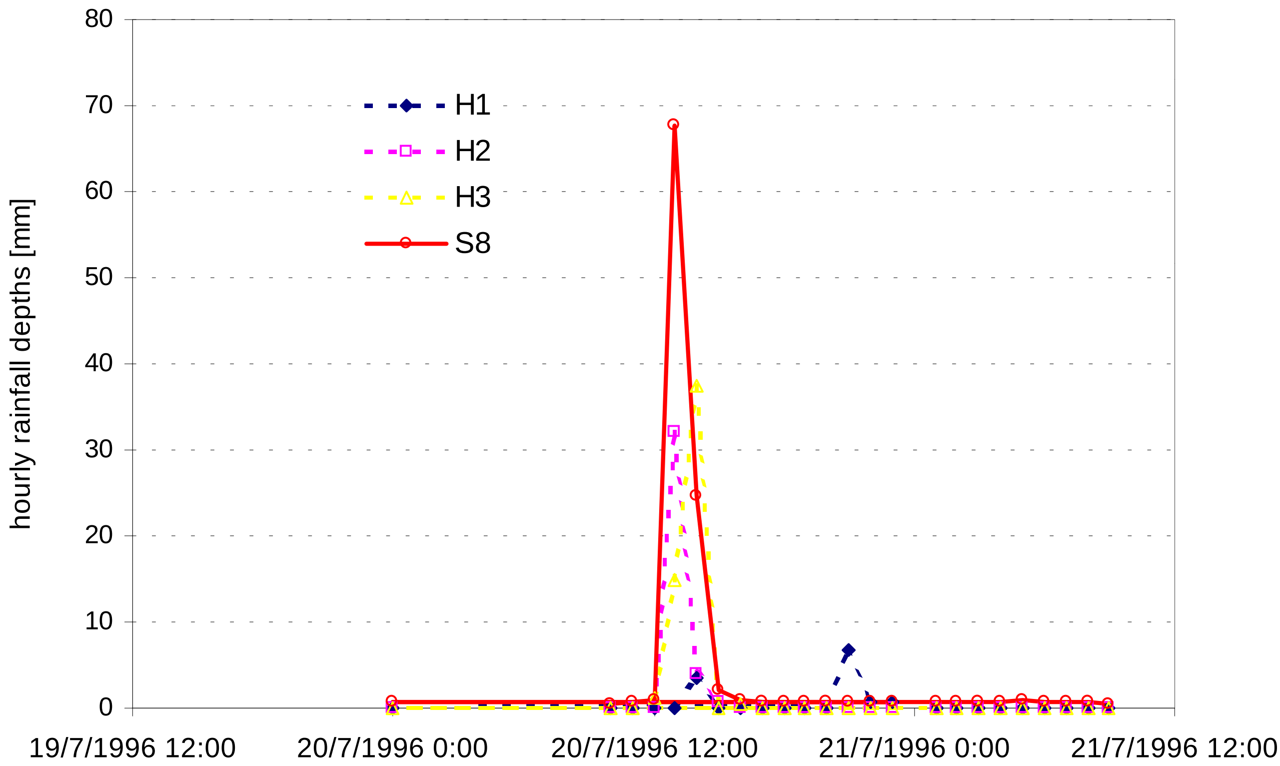


Figure 57: Simulated hyetographs for raingage 8

Month of August

For August, the simplified multivariate model was used in terms of power transformation $X_s^* = aX_{s-1}^* + bV_s$ where: $X_s^* = X_s^{(m)}$ and m=0.6

Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 3000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π_0 were set to 0.3 mm and 0.5 respectively.

Applying the disaggregation modeling framework synthetic hourly rainfall series were produced for the eight raingages, those of gages 1, 2 and 3 being identical to the historical series. The statistics of the synthetic series are compared to the historical and to the values used in the disaggregation in *table 15* where it can be observed a good agreement. Graphical comparisons show that the probability distribution functions of historical and simulated hourly rainfall depth during wet days for gage *5* are also in good agreement with each other (see *figures 58 and 59*).

Lag-one cross correlation coefficients of the synthetic series are compared with those used in disaggregation and with those of the historical series in *Table 16* It is shown that the synthetic values tend to agree much better with the values used in disaggregation especially given the fact that the historical values have not been entered in the calculations and their preservation could not be assured. Another fact worth mentioning is that the historical values of the cross correlations were extremely low and the essential hypothesis holding the entire framework, i.e. the significant spatial correlation between raingages, was not preserved. This fact could be encountered to the quality of the data , or to the particularities of the rainfall process during the month of August. The rainfall process in summer is characterized by intense storms of relatively small duration and extremely localized.

A further comparison is given in *figure 60* in terms of the autocorrelation function for higher lags, up to 10. Hyetographs of the synthetic series given in *figures 61-65* show that the maximum depths are predicted well but “with a temporal translation.

Table 15

Statistics of hourly rainfall depths at each gage for the month of AUGUST								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.98	0.97	0.98	0.98	0.98	0.98	-	-
value used on disaggregation	0.975	0.975	0.975	0.975	0.975	0.975	0.975	0.975
synthetic	0.97	0.97	0.98	*0.97*	*0.96*	*0.97*	*0.97*	*0.92*
Mean								
historical	0.06	0.08	0.08	0.05	0.05	0.06	-	-
value used on disaggregation	0.073	0.073	0.073	0.073	0.073	0.073	0.073	0.073
synthetic	0.06	0.08	0.08	*0.05*	*0.05*	*0.06*	*0.06*	*0.05*
Maximum value								
historical	23.2	36	34.8	24.8	24	19.8	-	-
value used on disaggregation	31.3	31.3	31.3	31.3	31.3	31.3	31.3	31.3
synthetic	23.2	36	34.8	*20.2*	*24.7*	*20.9*	*32.7*	*10.8*
Standard deviation								
historical	0.77	1.04	1.02	0.69	0.68	0.76	-	-
value used on disaggregation	0.945	0.945	0.945	0.945	0.945	0.945	0.945	0.945
synthetic	0.77	1.05	1.02	*0.56*	*0.63*	*0.67*	*0.78*	*0.34*
Skewness								
historical	18.18	22.65	19.56	24.32	20.58	19.12	-	-
value used on disaggregation	20.128	20.13	20.13	20.13	20.13	20.13	20.13	20.13
synthetic	18.15	22.59	19.48	*25.26*	*26.98*	*21.58*	*24.90*	*17.44*
Lag1 autocorrelation								
historical	0.29	0.32	0.20	0.19	0.22	0.26	-	-
value used on disaggregation	0.273	0.273	0.273	0.273	0.273	0.273	0.273	0.273
synthetic	0.29	0.32	0.20	*0.25*	*0.21*	*0.22*	*0.24*	*0.35*

Table 16

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of August

Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1								
historical	1.00	0.08	0.09	0.01	0.03	0.06	-	-
value used on disaggregation	1.00	0.08	0.09	0.11	0.07	0.45	0.36	0.25
synthetic	1.00	0.08	0.09	0.06	0.03	0.36	0.34	0.28
2								
historical	0.08	1.00	0.31	0.34	0.17	0.37	-	-
value used on disaggregation	0.08	1.00	0.31	0.17	0.18	0.35	0.13	0.38
synthetic	0.08	1.00	0.31	0.17	0.10	0.24	0.15	0.57
3								
historical	0.09	0.31	1.00	0.40	0.22	0.25	-	-
value used on disaggregation	0.09	0.31	1.00	0.19	0.14	0.18	0.17	0.24
synthetic	0.09	0.31	1.00	0.12	0.07	0.12	0.18	0.32
4								
historical	0.01	0.34	0.40	1.00	0.37	0.53	-	-
value used on disaggregation	0.11	0.17	0.19	1.00	0.39	0.27	0.09	0.45
synthetic	0.06	0.17	0.12	1.00	0.25	0.11	0.08	0.40
5								
historical	0.03	0.17	0.22	0.37	1.00	0.34	-	-
value used on disaggregation	0.07	0.18	0.14	0.39	1.00	0.15	0.04	0.35
synthetic	0.03	0.10	0.07	0.25	1.00	0.05	0.03	0.46
6								
historical	0.06	0.37	0.25	0.53	0.34	1.00	-	-
value used on disaggregation	0.45	0.35	0.18	0.27	0.15	1.00	0.31	0.39
synthetic	0.36	0.24	0.12	0.11	0.05	1.00	0.22	0.33
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.36	0.13	0.17	0.09	0.04	0.31	1.00	0.25
synthetic	0.34	0.15	0.18	0.08	0.03	0.22	1.00	0.30
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.25	0.38	0.24	0.45	0.35	0.39	0.25	1.00
synthetic	0.28	0.57	0.32	0.40	0.46	0.33	0.30	1.00

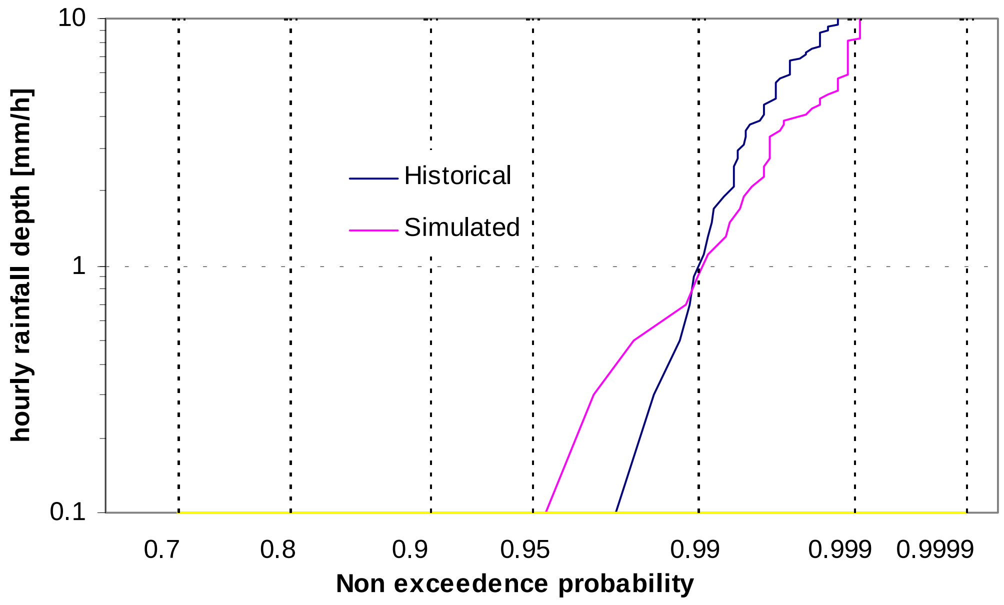


Figure 58: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 5 for the month of August

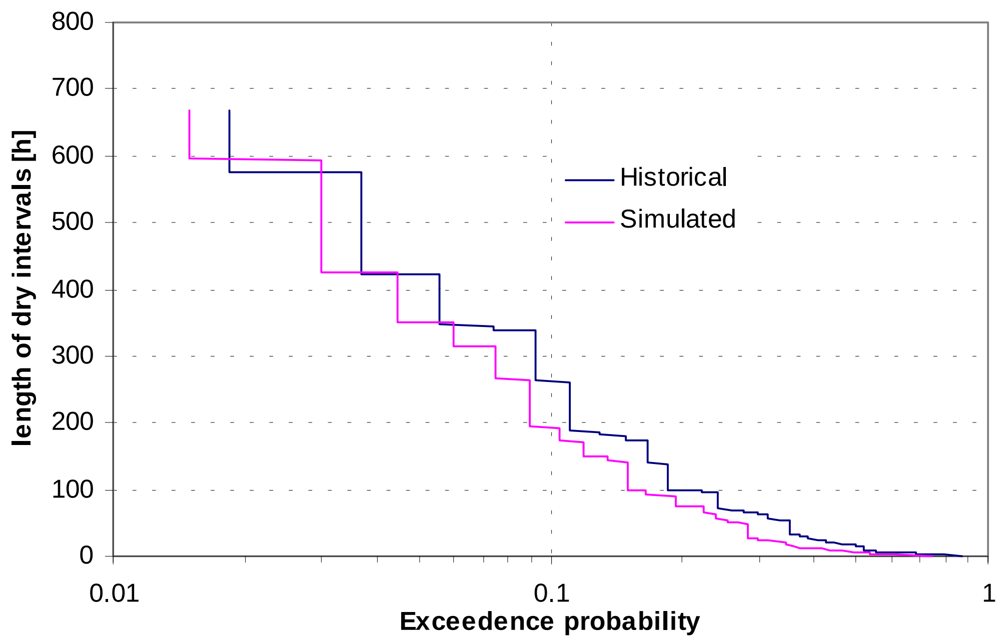


Figure 59: Comparison of historical and simulated probability distribution functions of of the length of dry intervals at gage 5 for the month of August

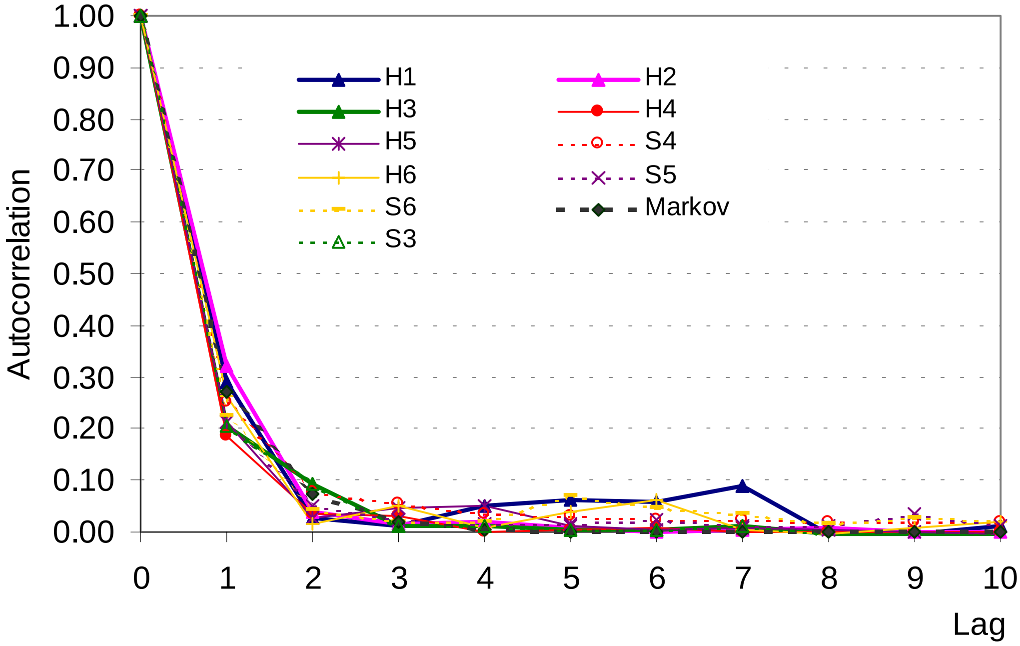


Figure 60: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of August.

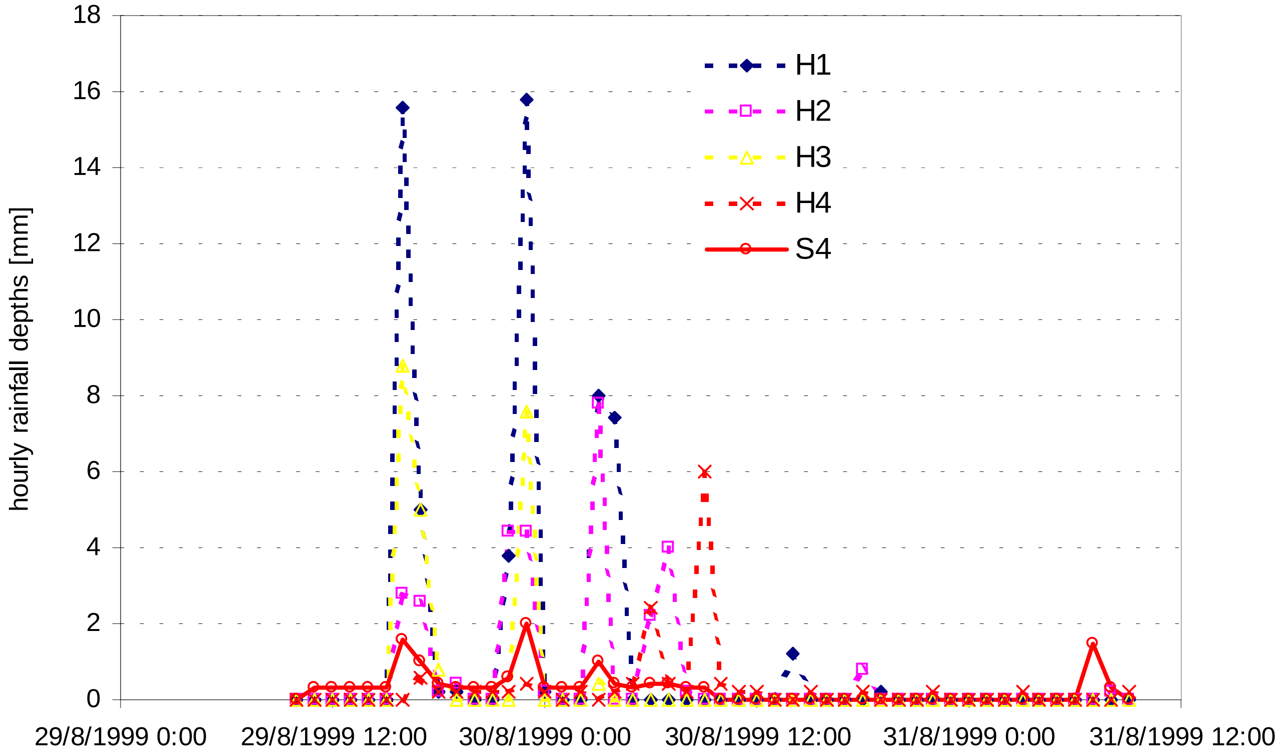


Figure 61: Comparison of historical and simulated hyetographs for raingage 4

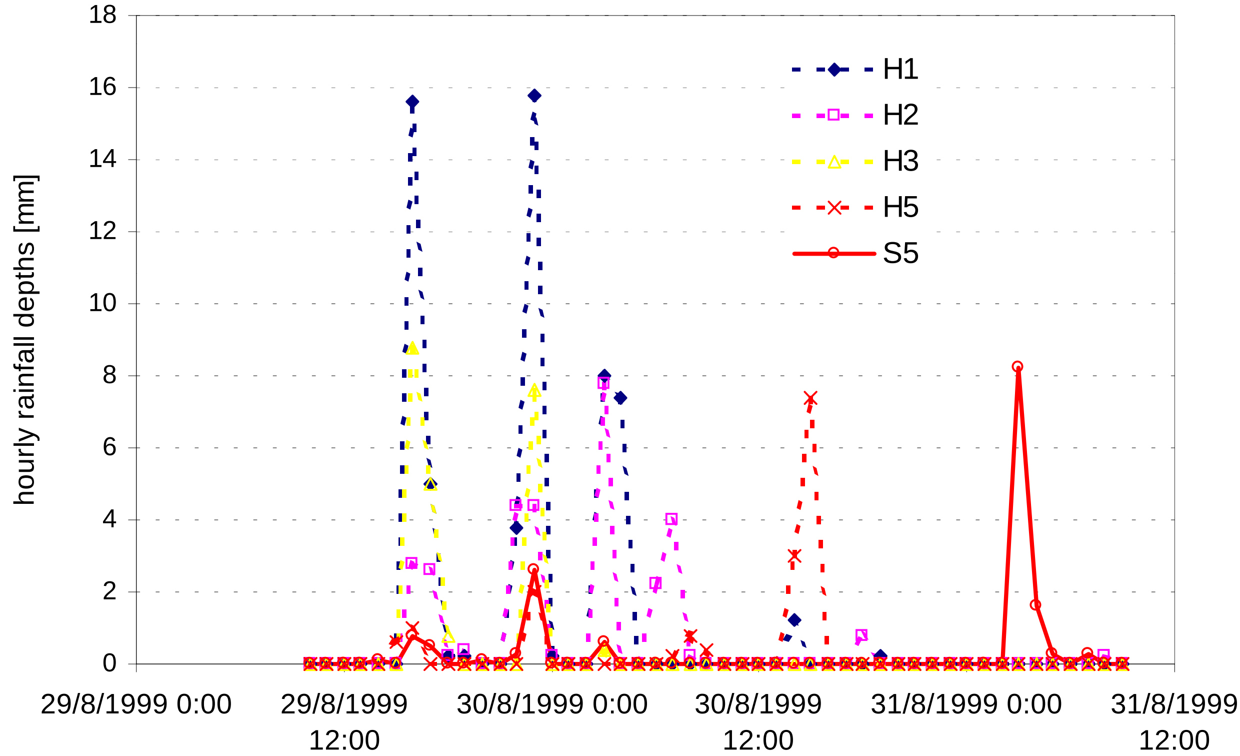


Figure 62: Comparison of historical and simulated hyetographs for raingage 5

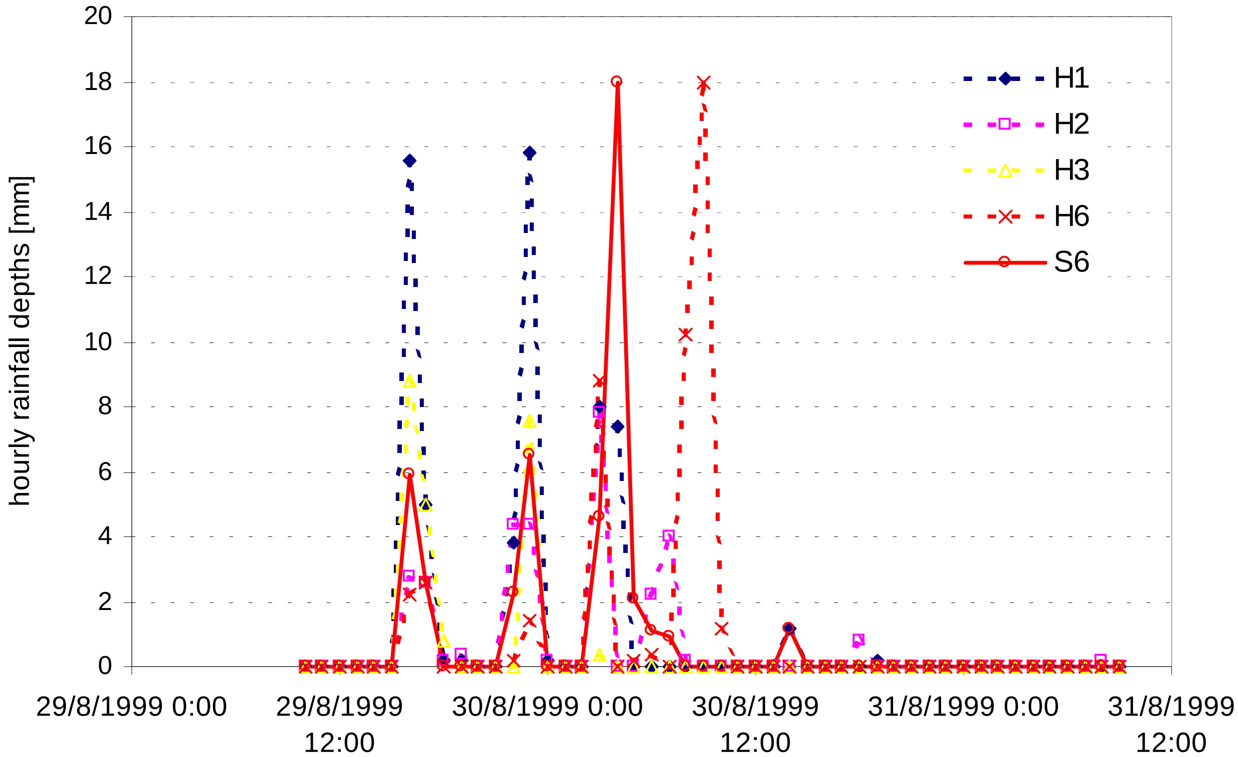


Figure 63: Comparison of historical and simulated hyetographs for raingage 6

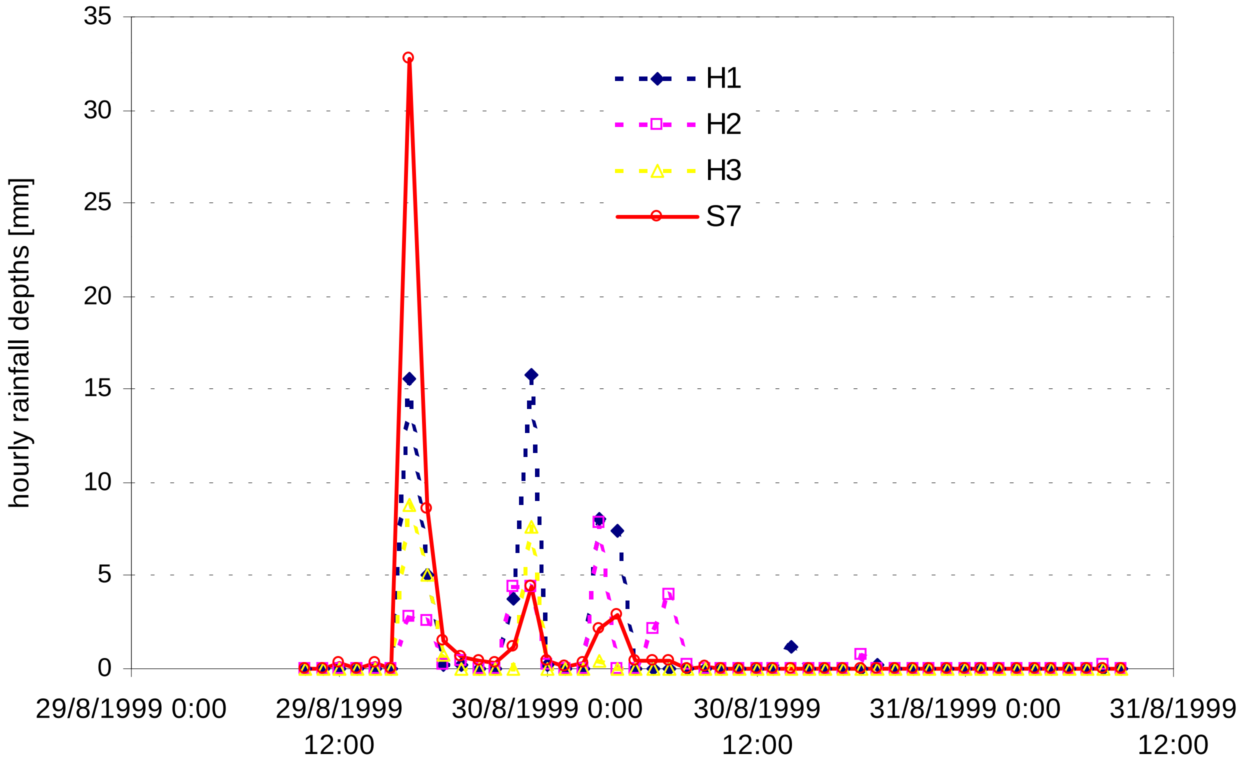


Figure 64: Simulated hyetographs for raingage 7

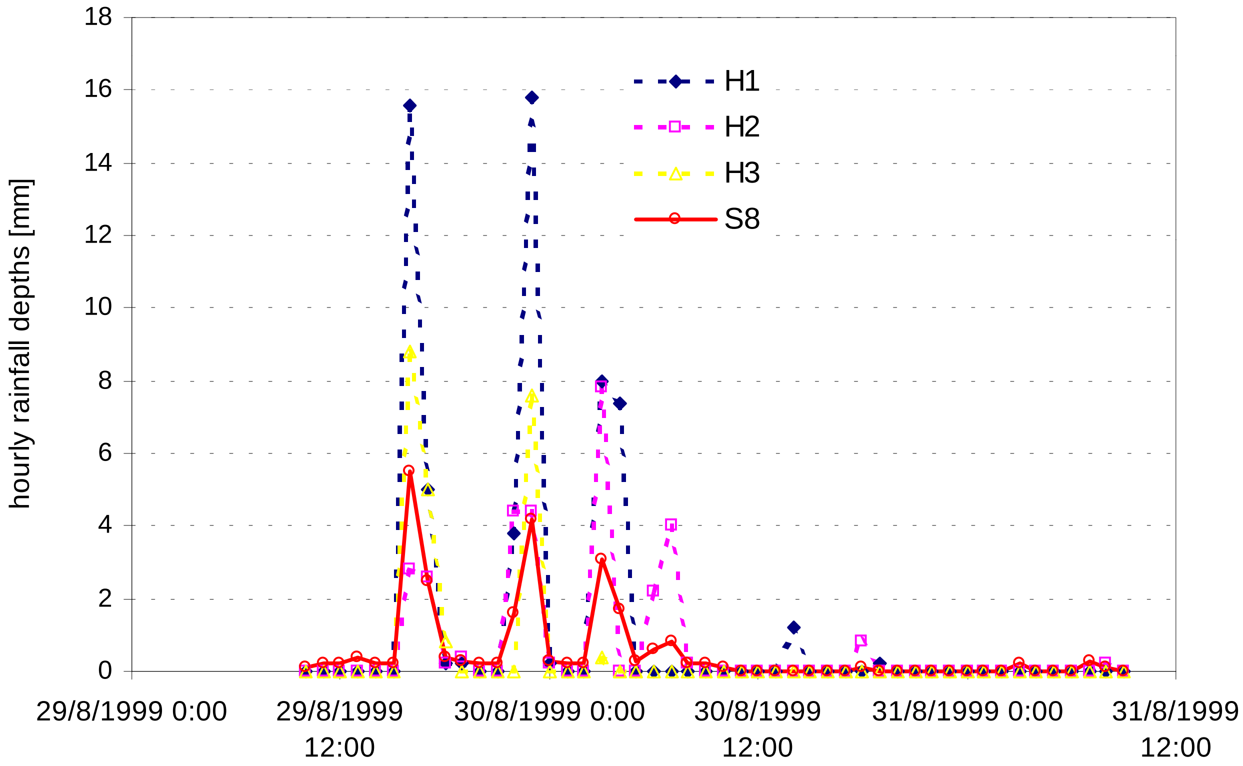


Figure 65: Simulated hyetographs for raingage 8

Month of September

For September, the simplified multivariate model was used in terms of power transformation $X_s^* = aX_{s-1}^* + bV_s$ where: $X_s^* = X_s^{(m)}$ and m=0.5

Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 3000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π_0 were set to 0.3 mm and 0.5 respectively.

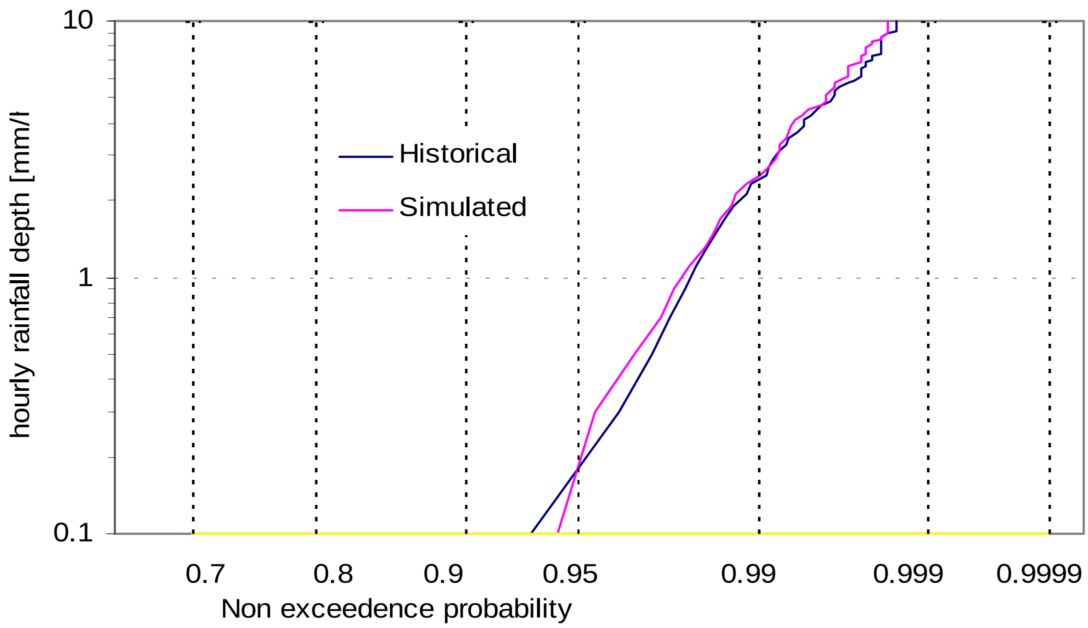


Figure 66: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 4 for the month of September

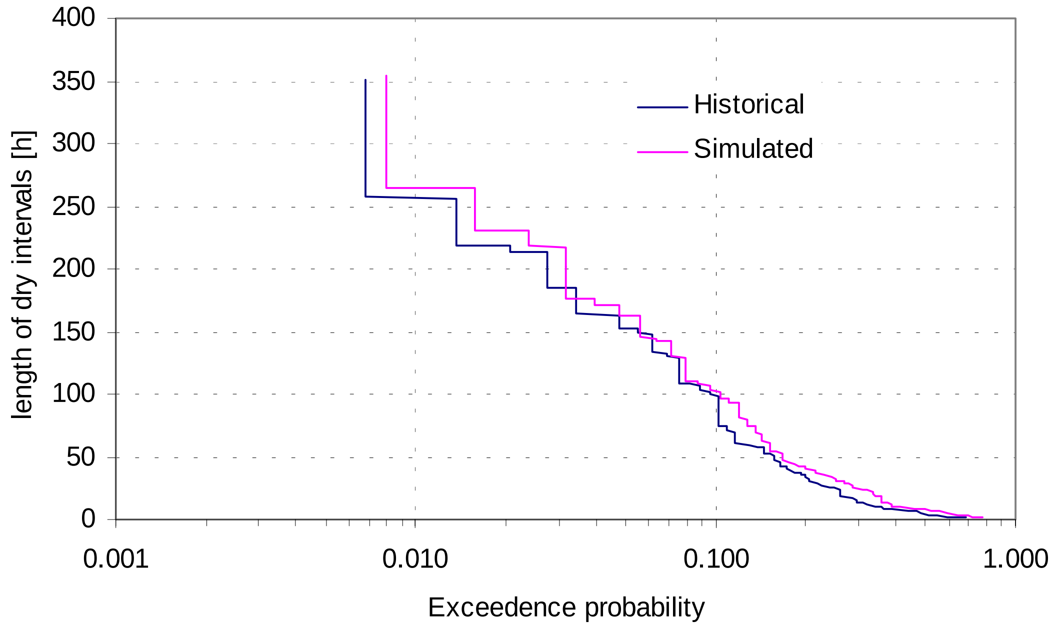


Figure 67: Comparison of historical and simulated probability distribution functions of the length of dry intervals at gage 4 for the month of September

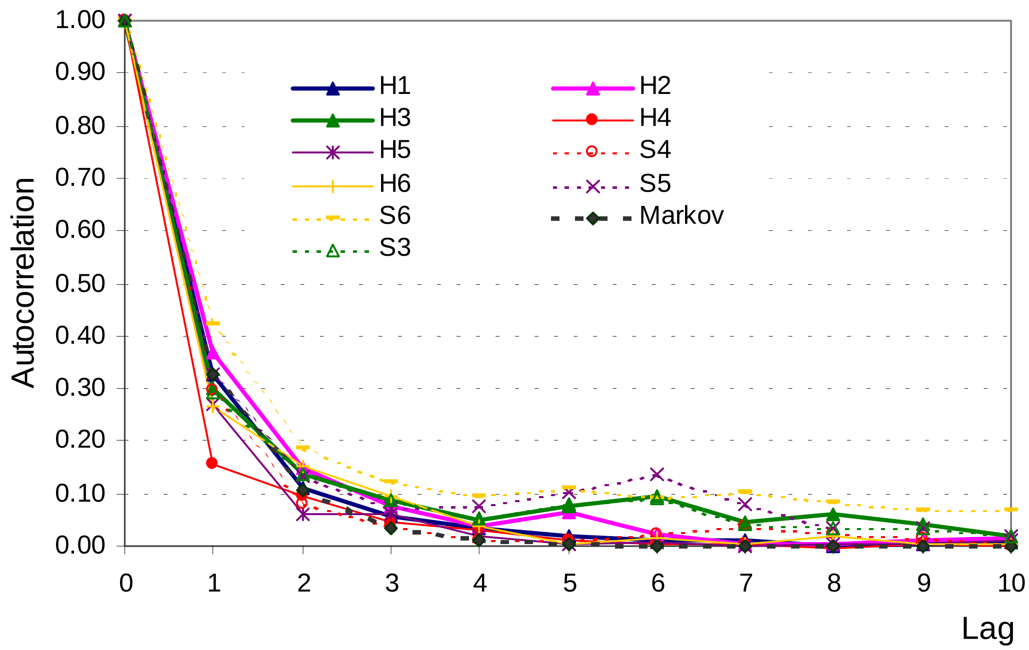


Figure 68: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of September.

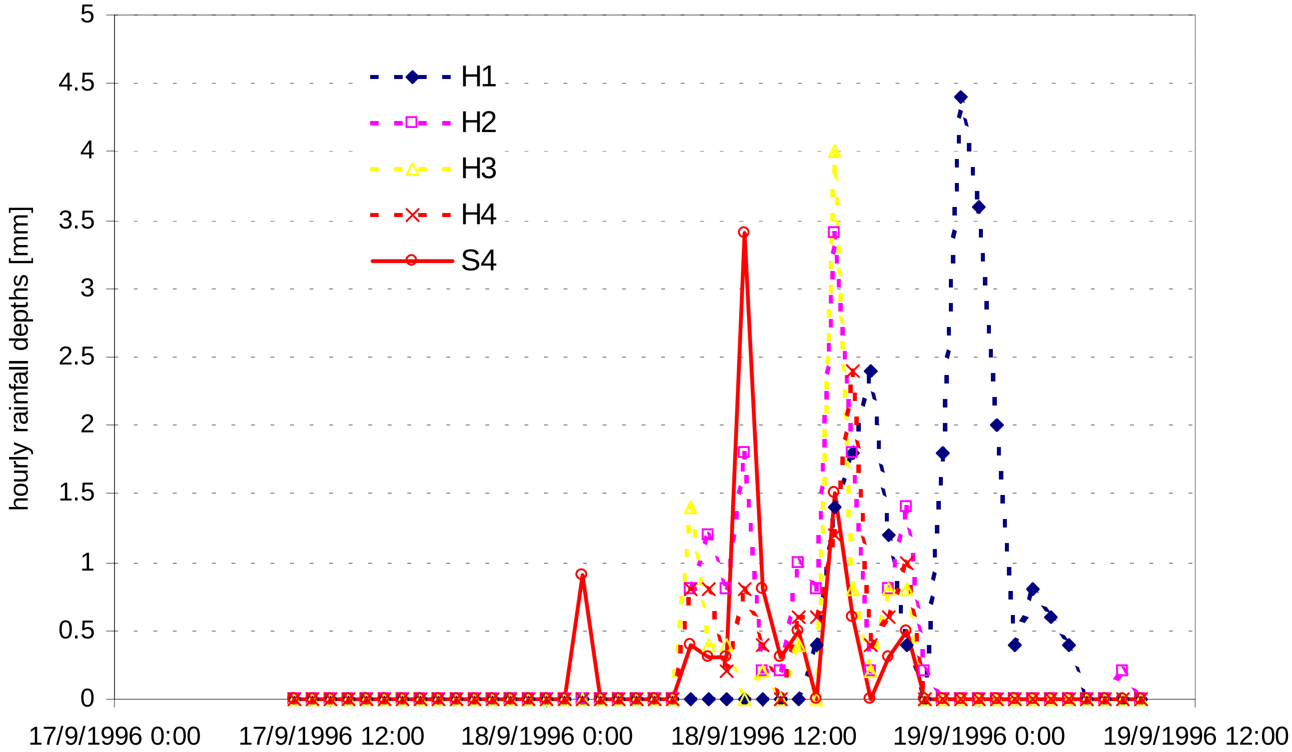


Figure 69: Comparison of historical and simulated hyetographs for raingage 4

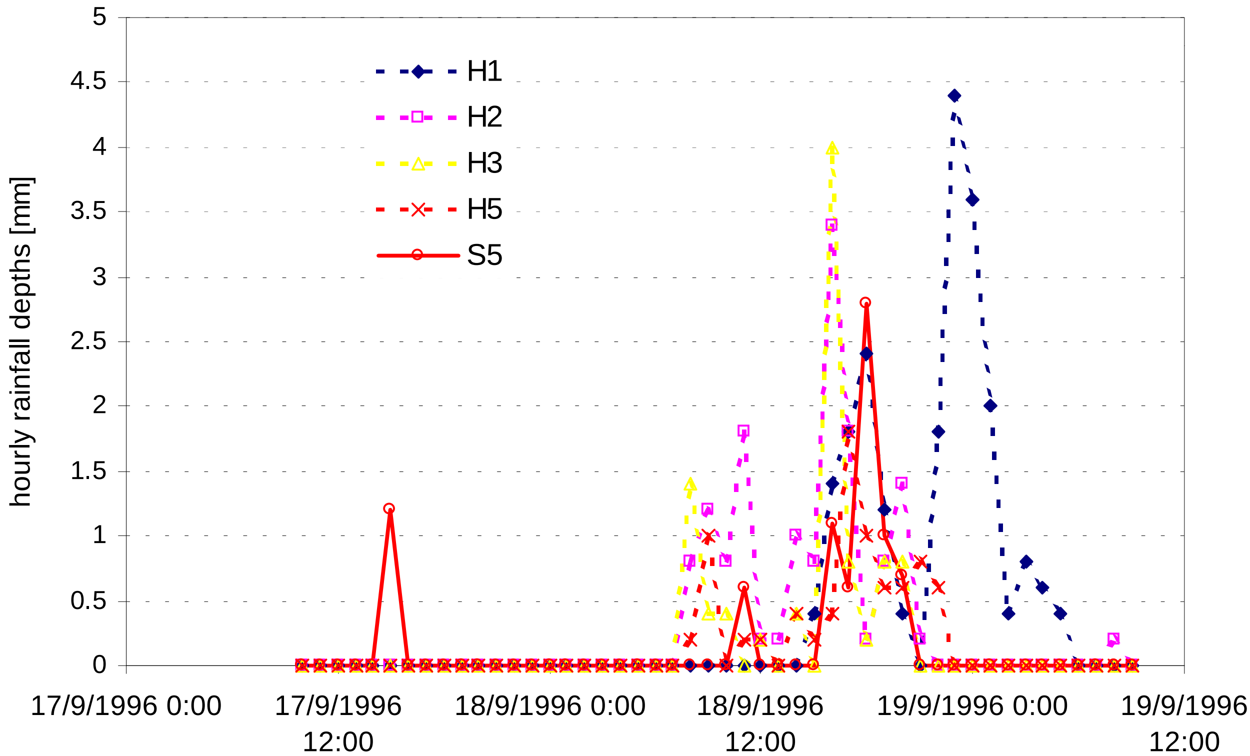


Figure 70: Comparison of historical and simulated hyetographs for raingage 5

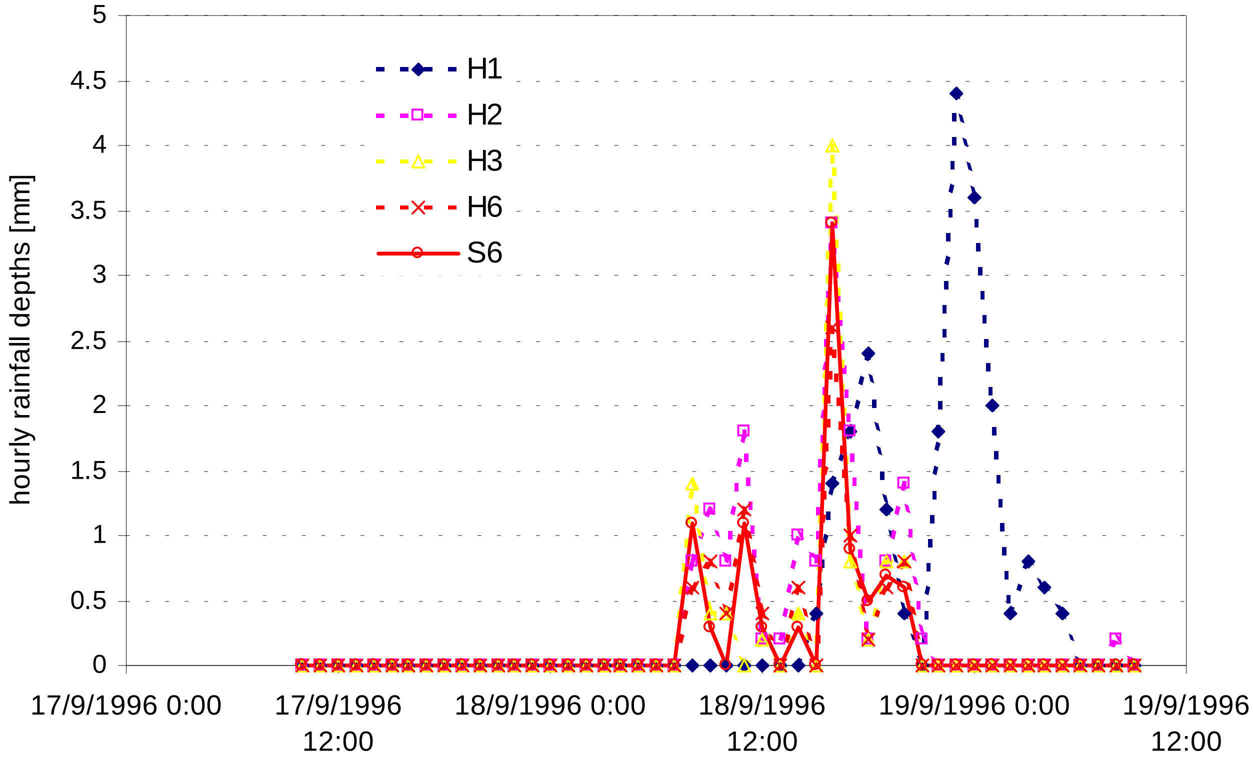


Figure 71: Comparison of historical and simulated hyetographs for raingage 6

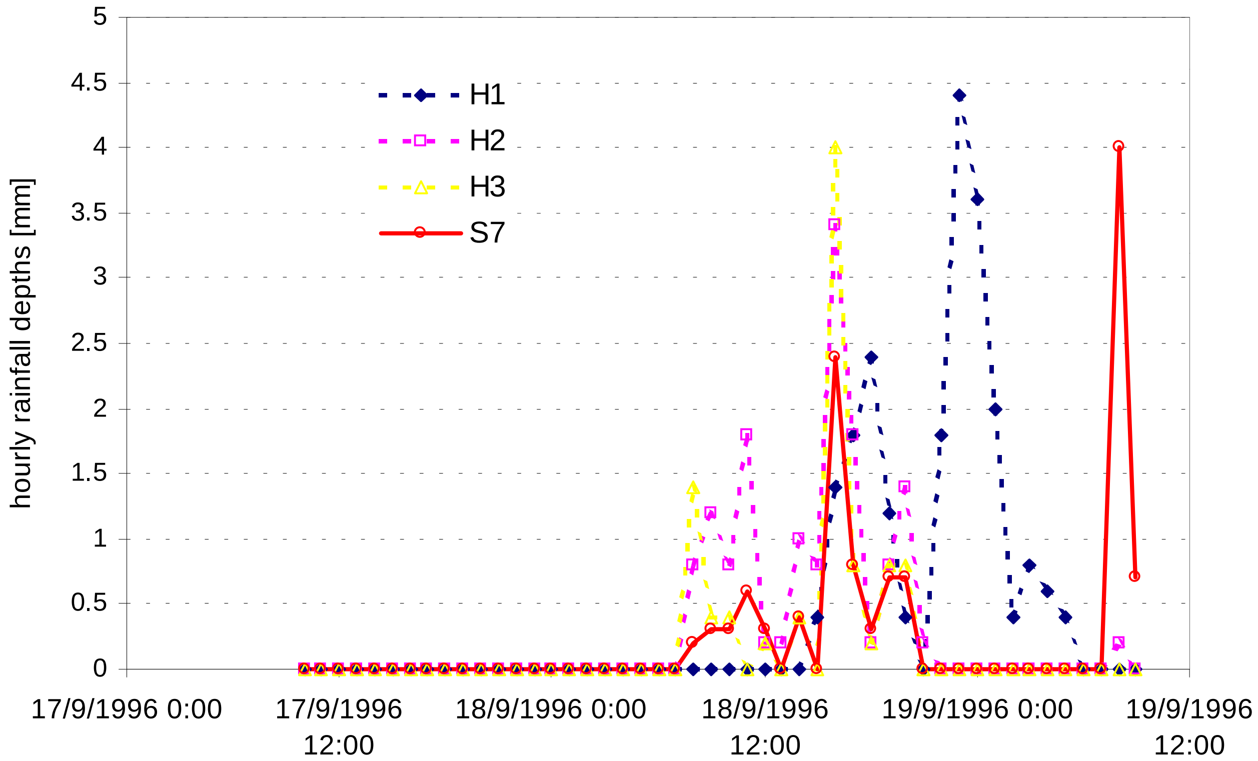


Figure 72:Simulated hyetographs for raingage 7

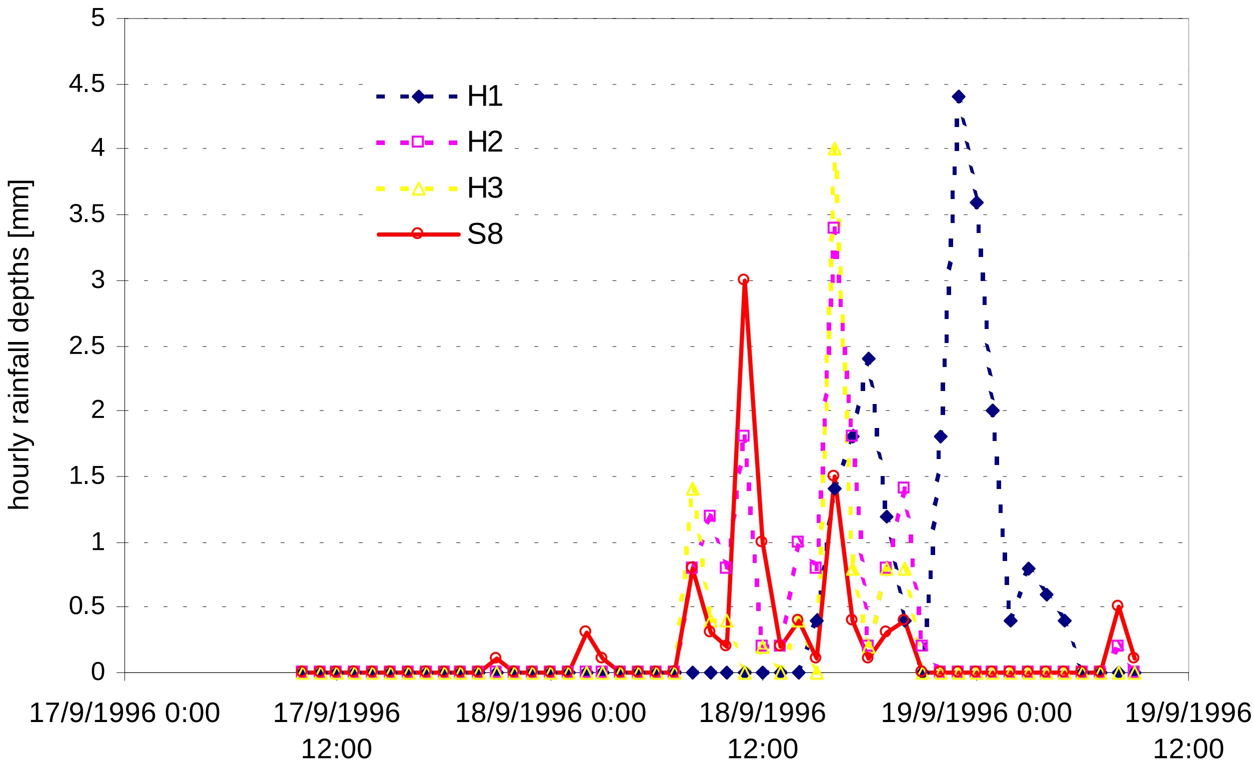


Figure 73:Simulated hyetographs for raingage 8

Table 17

Statistics of hourly rainfall depths at each gage for the month of SEPTEMBER								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.95	0.94	0.95	0.93	0.95	0.96		
value used on disaggregation	0.946	0.946	0.946	0.946	0.946	0.946	0.946	0.946
synthetic	0.95	0.94	0.95	*0.94*	*0.94*	*0.93*	*0.92*	*0.82*
Mean								
historical	0.12	0.11	0.11	0.10	0.13	0.09		
value used on disaggregation	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
synthetic	0.12	0.11	0.11	*0.10*	*0.13*	*0.09*	*0.12*	*0.12*
Maximum value								
historical	22.6	23.4	19.2	29.4	39.2	25	-	-
value used on disaggregation	21.7	21.7	21.7	21.7	21.7	21.7	21.7	21.7
synthetic	22.6	23.4	19.2	*22.5*	*27*	*13.9*	*16.8*	*22.3*
Standard deviation								
historical	0.92	0.88	0.85	0.93	1.29	0.80	-	-
value used on disaggregation	0.886	0.886	0.886	0.886	0.886	0.886	0.886	0.886
synthetic	0.92	0.87	0.85	*0.82*	*1.00*	*0.59*	*0.73*	*0.74*
Skewness								
historical	13.74	14.65	11.90	20.03	18.91	17.08	-	-
value used on disaggregation	13.429	13.429	13.429	13.429	13.429	13.429	13.429	13.429
synthetic	13.74	14.65	11.90	*16.41*	*15.22*	*12.89*	*11.64*	*15.28*
Lag1 autocorrelation								
historical	0.33	0.37	0.30	0.16	0.27	0.27	-	-
value used on disaggregation	0.333	0.333	0.333	0.333	0.333	0.333	0.333	0.333
synthetic	0.33	0.36	0.29	*0.30*	*0.33*	*0.42*	*0.39*	*0.38*

Table 18

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of September								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1								
historical	1.00	0.43	0.23	0.15	0.16	0.29	-	-
value used on disaggregation	1.00	0.43	0.23	0.22	0.33	0.42	0.30	0.20
synthetic	1.00	0.44	0.23	0.18	0.22	0.38	0.37	0.20
2								
historical	0.43	1.00	0.51	0.45	0.44	0.59	-	-
value used on disaggregation	0.43	1.00	0.51	0.51	0.49	0.60	0.59	0.51
synthetic	0.44	1.00	0.51	0.47	0.35	0.68	0.67	0.51
3								
historical	0.23	0.51	1.00	0.40	0.32	0.50	-	-
value used on disaggregation	0.23	0.51	1.00	0.45	0.35	0.46	0.64	0.49
synthetic	0.23	0.51	1.00	0.41	0.33	0.55	0.68	0.48
4								
historical	0.15	0.45	0.40	1.00	0.46	0.74	-	-
value used on disaggregation	0.22	0.51	0.45	1.00	0.51	0.69	0.61	0.79
synthetic	0.18	0.47	0.41	1.00	0.32	0.74	0.56	0.76
5								
historical	0.16	0.44	0.32	0.46	1.00	0.52	-	-
value used on disaggregation	0.33	0.49	0.35	0.51	1.00	0.58	0.54	0.55
synthetic	0.22	0.35	0.33	0.32	1.00	0.47	0.45	0.37
6								
historical	0.29	0.59	0.50	0.74	0.52	1.00	-	-
value used on disaggregation	0.42	0.60	0.46	0.69	0.58	1.00	0.64	0.59
synthetic	0.38	0.68	0.55	0.74	0.47	1.00	0.73	0.69
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.30	0.59	0.64	0.61	0.54	0.64	1.00	0.72
synthetic	0.37	0.67	0.68	0.56	0.45	0.73	1.00	0.64
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.20	0.51	0.49	0.79	0.55	0.59	0.72	1.00
synthetic	0.20	0.51	0.48	0.76	0.37	0.69	0.64	1.00

Month of October

For October, the simplified multivariate model was used in terms of linear transformation $X_s = aX_{s-1} + bV_s$ Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 6000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π_0 were set to 0.3 mm and 0.1 respectively.

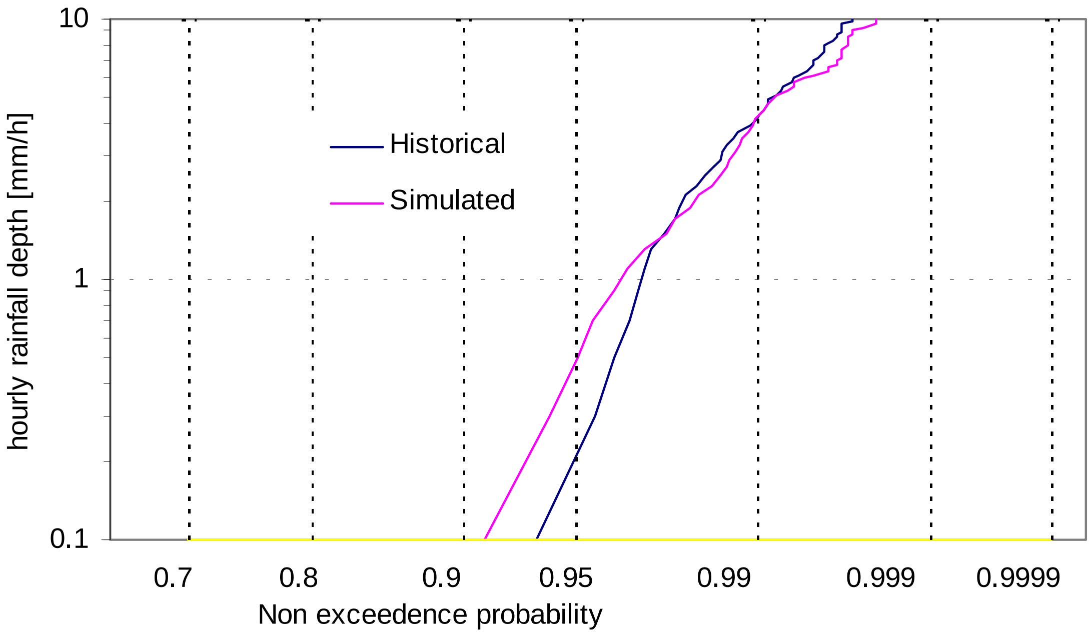


Figure 74: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 5 for the month of October

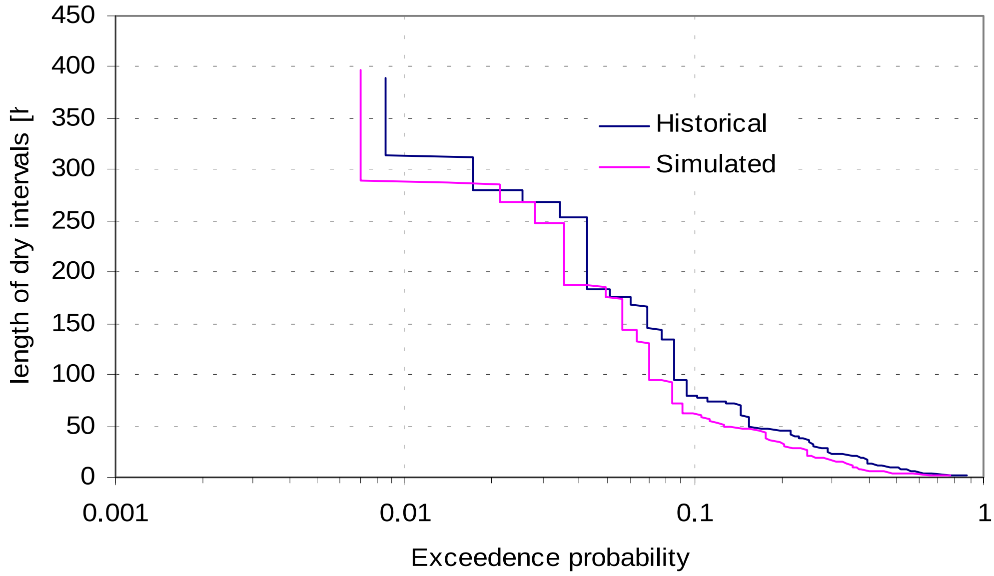


Figure 75: Comparison of historical and simulated probability distribution functions of the length of dry intervals at gage 5 for the month of October

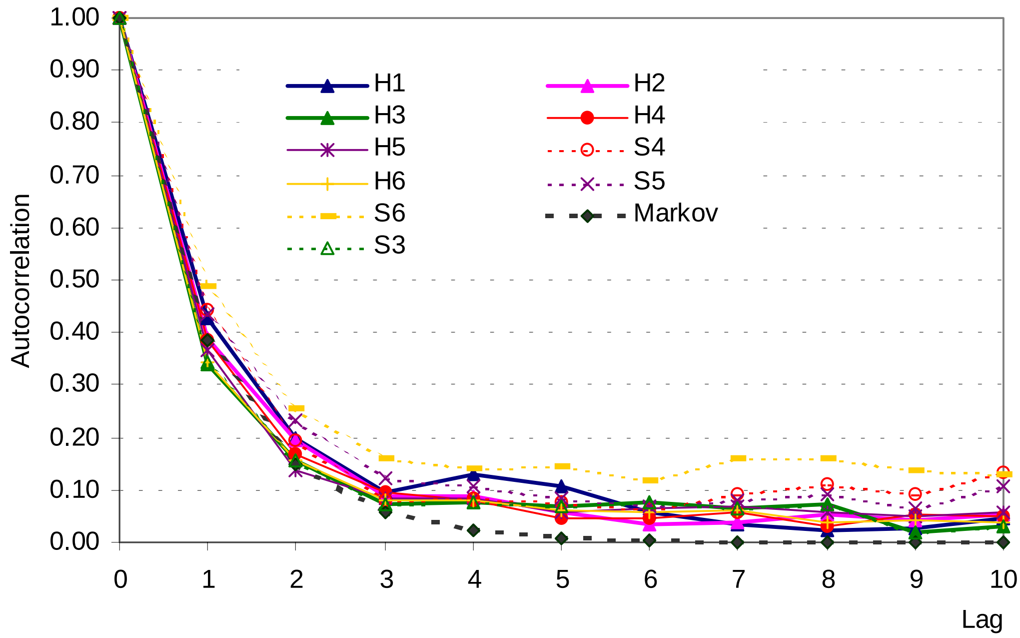


Figure 76: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of October.

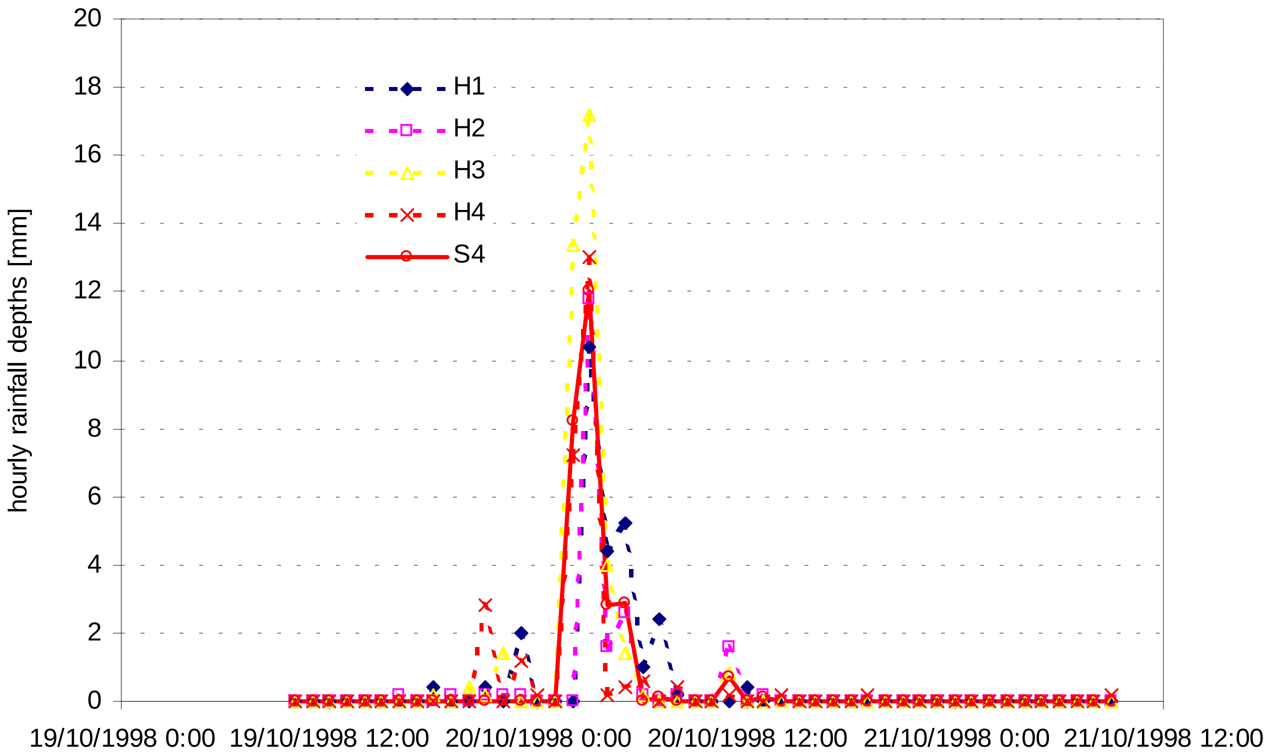


Figure 77: Comparison of historical and simulated hyetographs for raingage 4

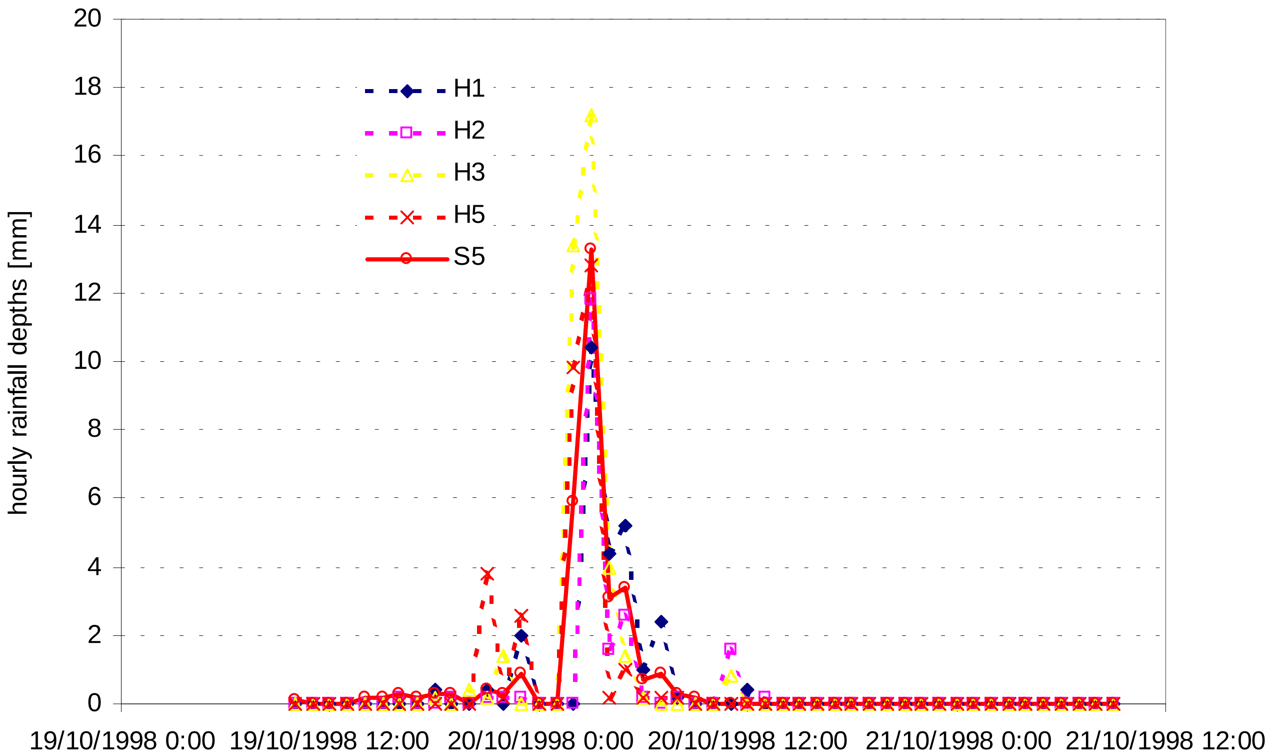


Figure 78: Comparison of historical and simulated hyetographs for raingage 5

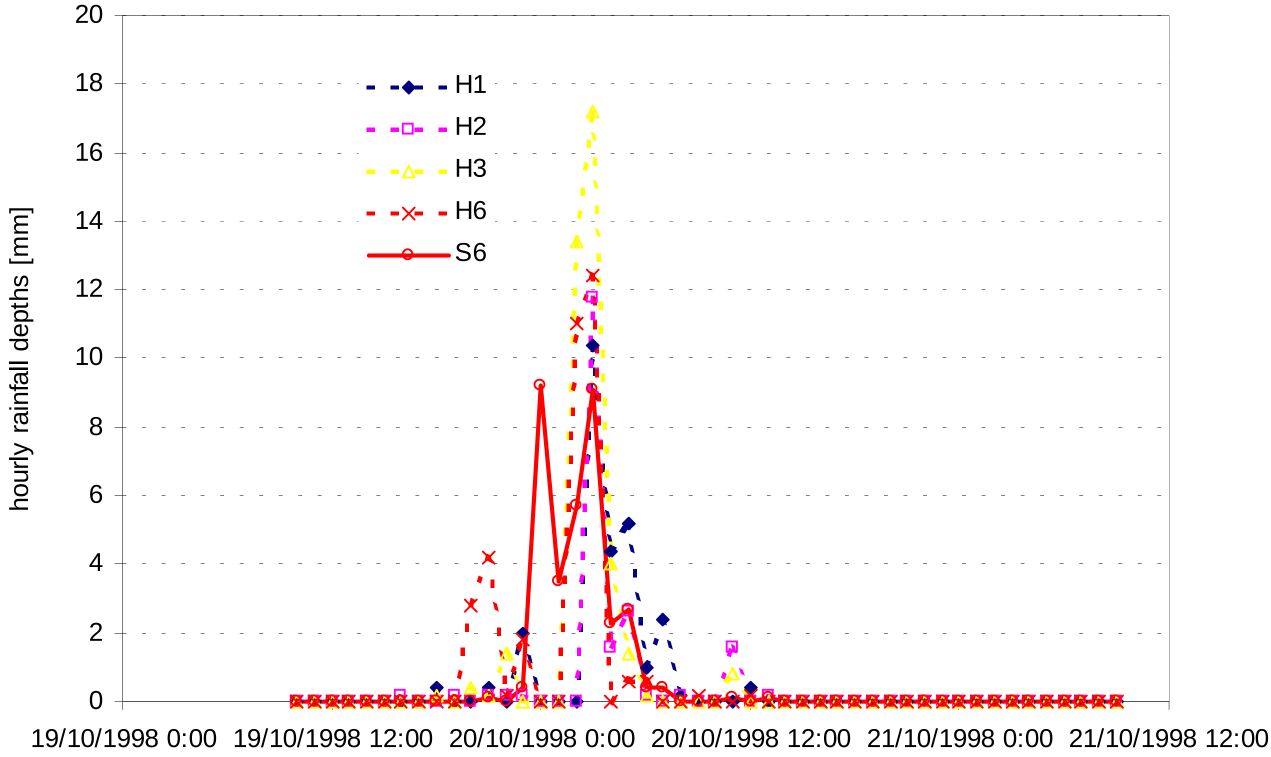


Figure 79: Comparison of historical and simulated hyetographs for raingage 6

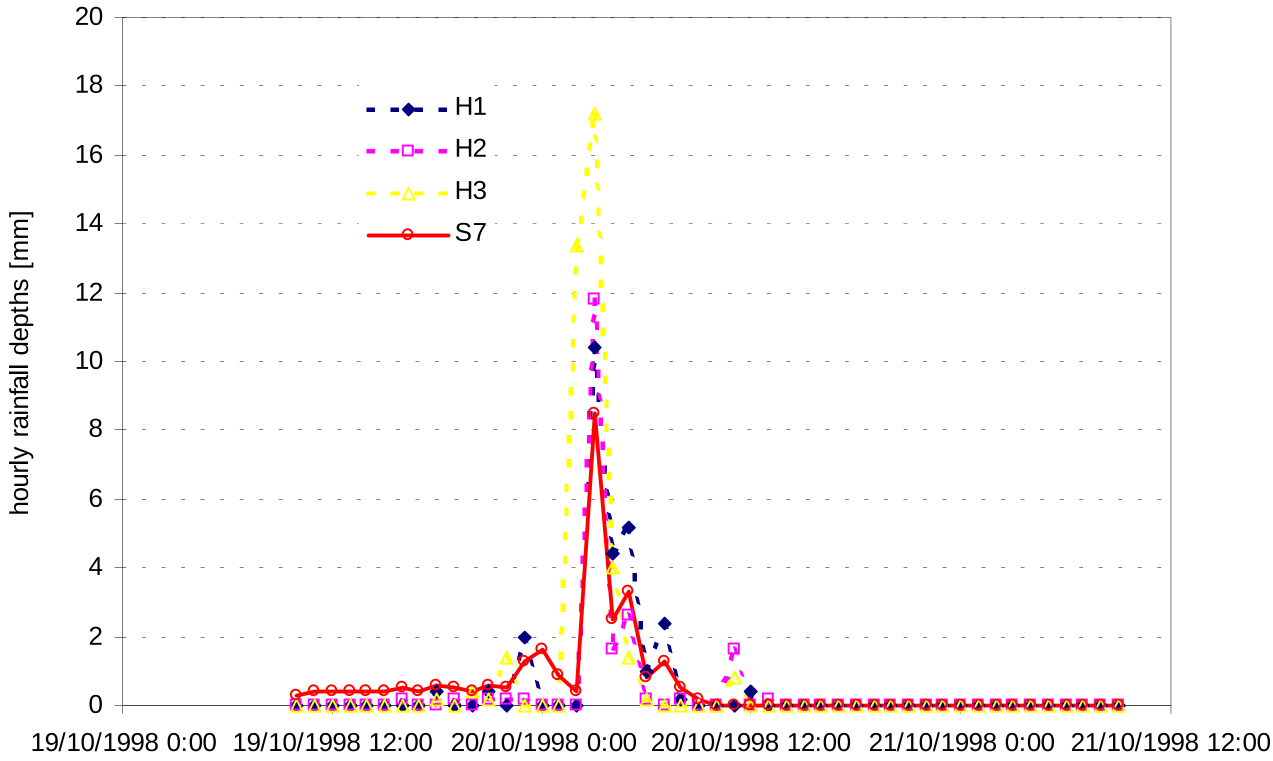


Figure 80: Simulated hyetographs for raingage 7

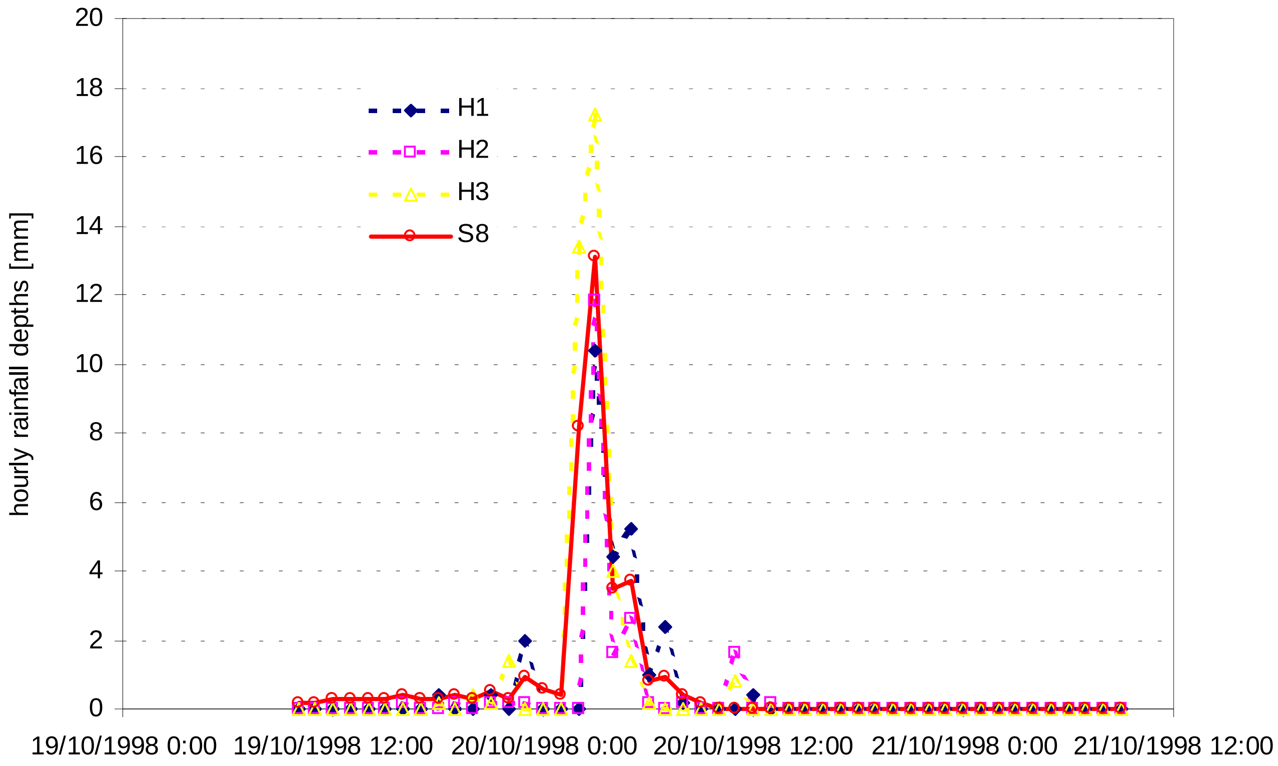


Figure 81: Simulated hyetographs for raingage 8

Table 19

Statistics of hourly rainfall depths at each gage for the month of OCTOBER								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.94	0.93	0.94	0.92	0.94	0.95	-	-
value used on disaggregation	0.938	0.938	0.938	0.938	0.938	0.938	0.938	0.938
synthetic	0.94	0.93	0.94	*0.92*	*0.91*	*0.92*	*0.92*	*0.89*
Mean								
historical	0.14	0.13	0.15	0.13	0.15	0.12	-	-
value used on disaggregation	0.14	0.14	0.14	0.14	0.14	0.14	0.14	0.14
synthetic	0.14	0.13	0.15	*0.13*	*0.15*	*0.12*	*0.15*	*0.15*
Maximum value								
historical	19.2	23.6	23.8	21.8	25.2	29.8	-	-
value used on disaggregation	22.2	22.2	22.2	22.2	22.2	22.2	22.2	22.2
synthetic	19.2	23.6	23.8	*17.1*	*26*	*15.8*	*26.2*	*15.8*
Standard deviation								
historical	0.98	0.91	1.14	0.95	1.09	1.01	-	-
value used on disaggregation	1.01	1.01	1.01	1.01	1.01	1.01	1.01	1.01
synthetic	0.98	0.91	1.15	*0.82*	*0.91*	*0.73*	*0.97*	*0.83*
Skewness								
historical	10.79	11.82	12.24	11.78	12.26	14.75	-	-
value used on disaggregation	11.617	11.617	11.617	11.617	11.617	11.617	11.617	11.617
synthetic	10.78	11.79	12.19	*10.42*	*12.26*	*10.87*	*12.63*	*10.15*
Lag1 autocorrelation								
historical	0.43	0.39	0.34	0.39	0.37	0.35	-	-
value used on disaggregation	0.386	0.386	0.386	0.386	0.386	0.386	0.386	0.386
synthetic	0.43	0.39	0.34	*0.44*	*0.43*	*0.49*	*0.42*	*0.48*

Table 20

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of October

Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1								
historical	1.00	0.43	0.23	0.27	0.28	0.22	-	-
value used on disaggregation	1.00	0.43	0.23	0.52	0.52	0.42	0.56	0.52
synthetic	1.00	0.44	0.23	0.52	0.53	0.42	0.48	0.54
2								
historical	0.43	1.00	0.34	0.72	0.52	0.72	-	-
value used on disaggregation	0.43	1.00	0.34	0.79	0.69	0.64	0.52	0.76
synthetic	0.44	1.00	0.34	0.82	0.73	0.69	0.45	0.81
3								
historical	0.23	0.34	1.00	0.36	0.32	0.30	-	-
value used on disaggregation	0.23	0.34	1.00	0.23	0.23	0.13	0.24	0.24
synthetic	0.23	0.34	1.00	0.31	0.32	0.19	0.22	0.33
4								
historical	0.27	0.72	0.36	1.00	0.72	0.84	-	-
value used on disaggregation	0.52	0.79	0.23	1.00	0.81	0.75	0.40	0.94
synthetic	0.52	0.82	0.31	1.00	0.83	0.81	0.45	0.97
5								
historical	0.28	0.52	0.32	0.72	1.00	0.60	-	-
value used on disaggregation	0.52	0.69	0.23	0.81	1.00	0.53	0.36	0.86
synthetic	0.53	0.73	0.32	0.83	1.00	0.64	0.41	0.91
6								
historical	0.22	0.72	0.30	0.84	0.60	1.00	-	-
value used on disaggregation	0.42	0.64	0.13	0.75	0.53	1.00	0.40	0.70
synthetic	0.42	0.69	0.19	0.81	0.64	1.00	0.42	0.80
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.56	0.52	0.24	0.40	0.36	0.40	1.00	0.40
synthetic	0.48	0.45	0.22	0.45	0.41	0.42	1.00	0.46
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.52	0.76	0.24	0.94	0.86	0.70	0.40	1.00
synthetic	0.54	0.81	0.33	0.97	0.91	0.80	0.46	1.00

Month of November

For November, the simplified multivariate model was used in terms of linear transformation $X_s = aX_{s-1} + bV_s$ Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 1000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π_0 were set to 0.2 mm and 0.1 respectively.

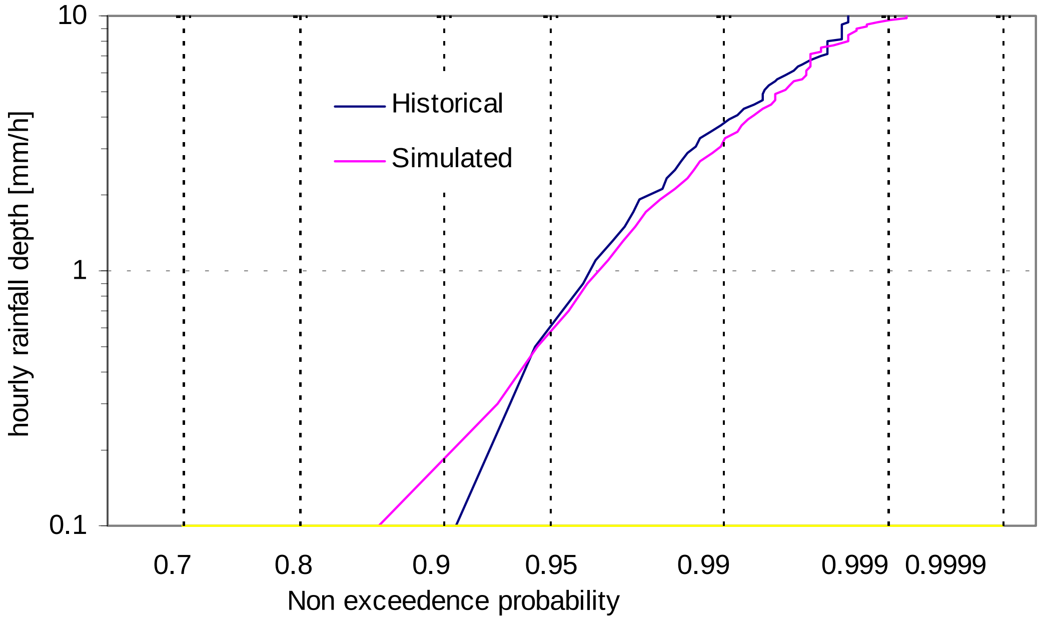


Figure 82: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 6 for the month of November

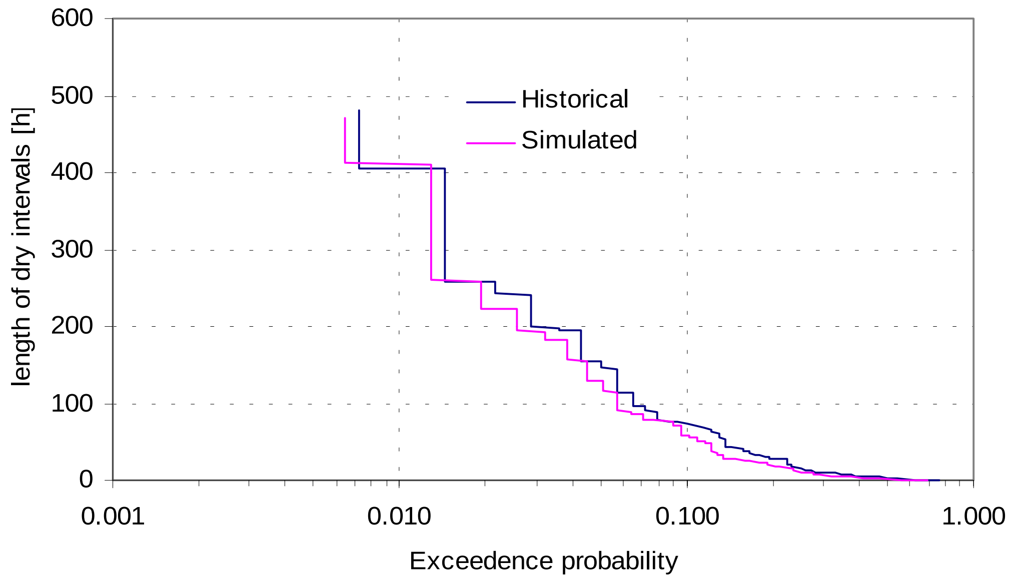


Figure 83: Comparison of historical and simulated probability distribution functions of the length of dry intervals at gage 6 for the month of November

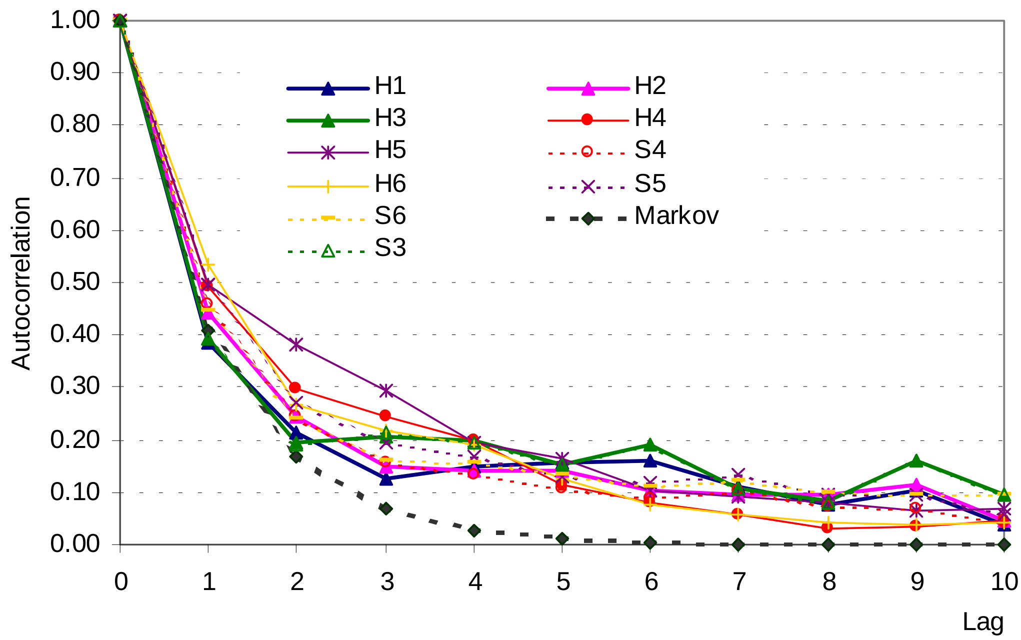


Figure 84: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of November.

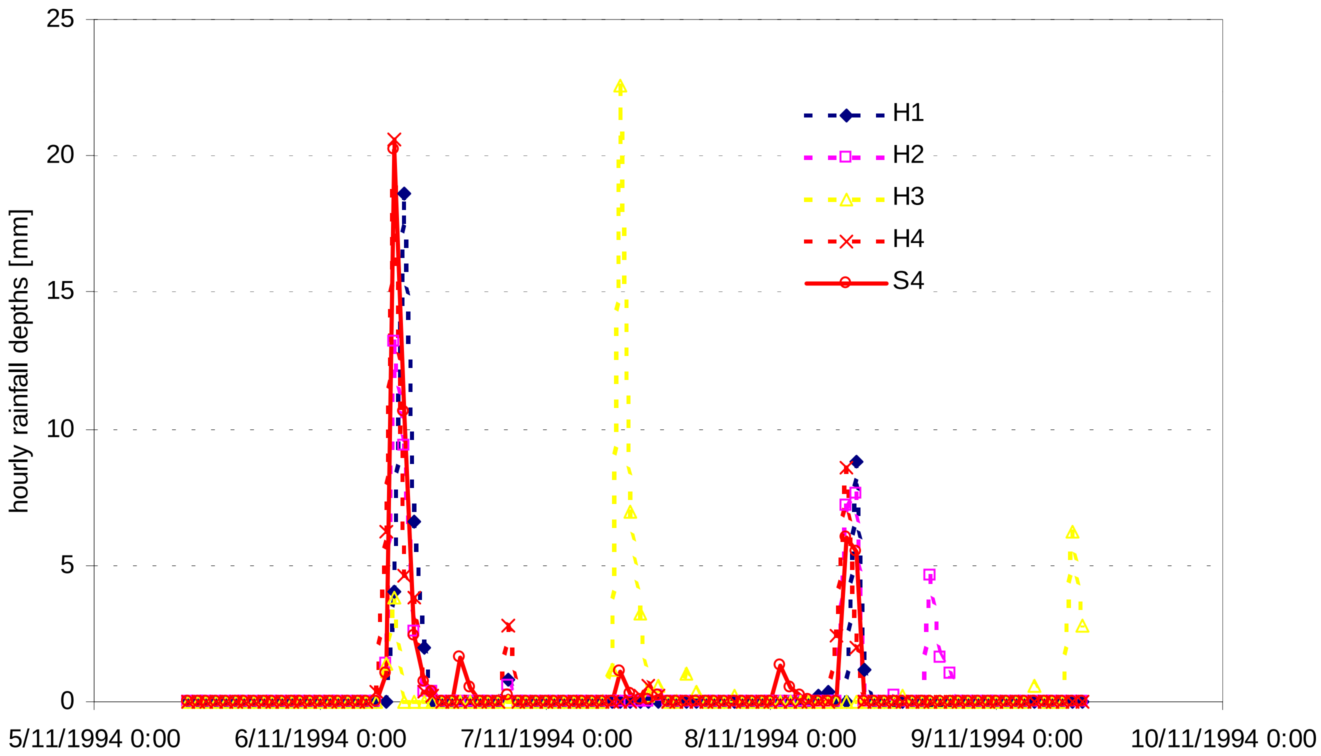


Figure 85: Comparison of historical and simulated hyetographs for raingage 4

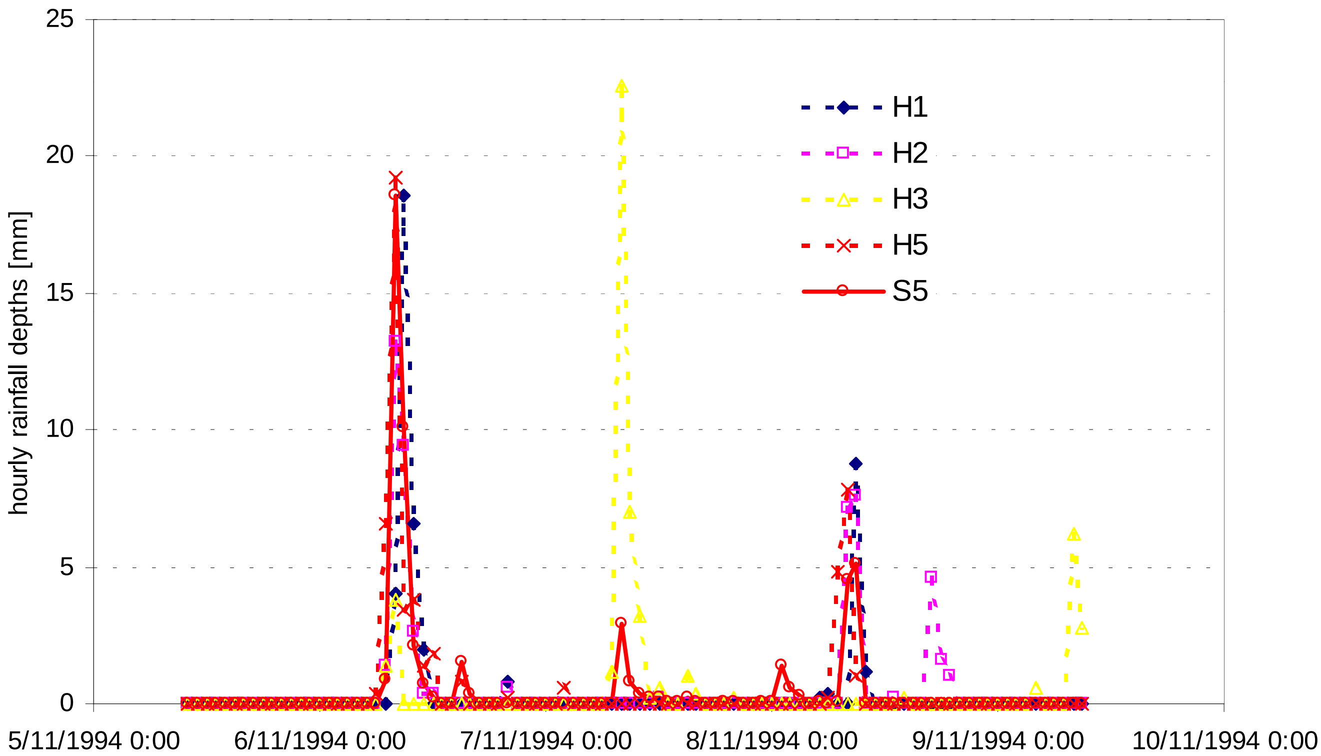


Figure 86: Comparison of historical and simulated hyetographs for raingage 5

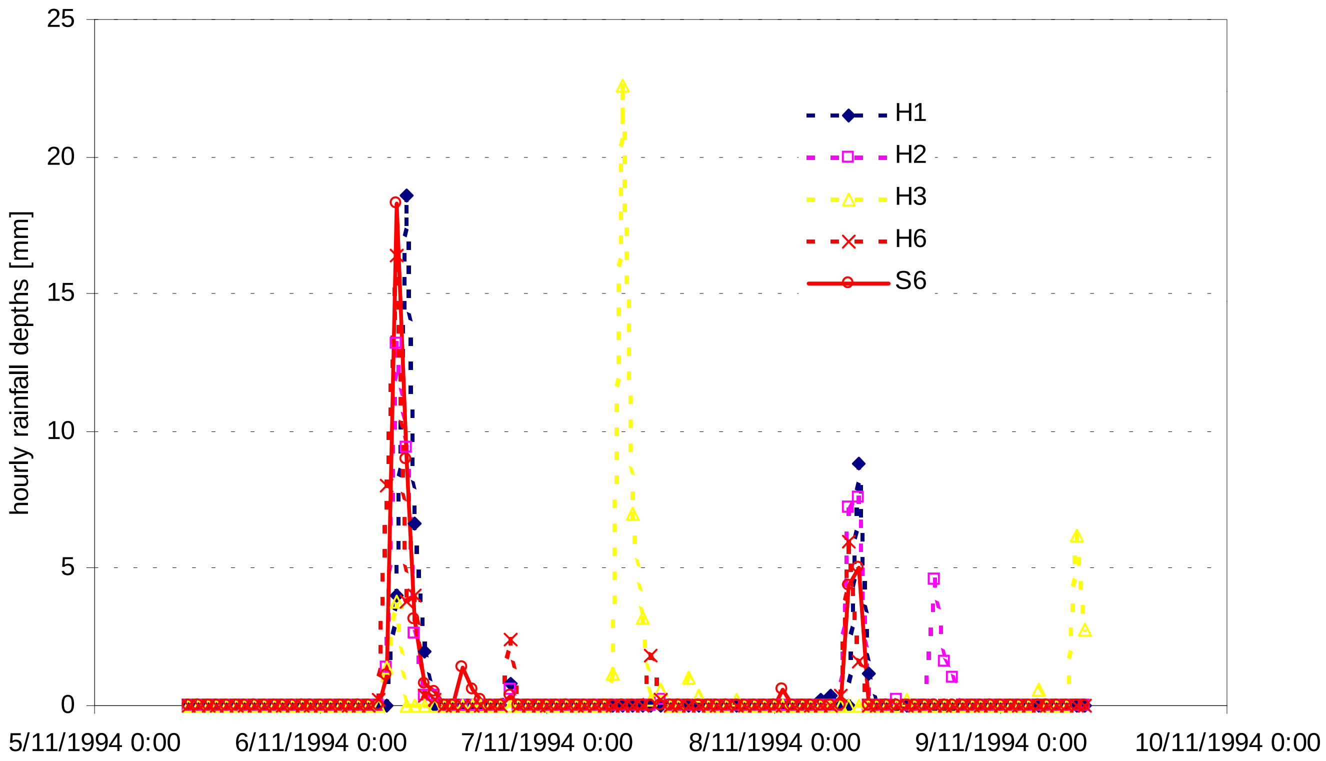


Figure 87: Comparison of historical and simulated hyetographs for raingage 6

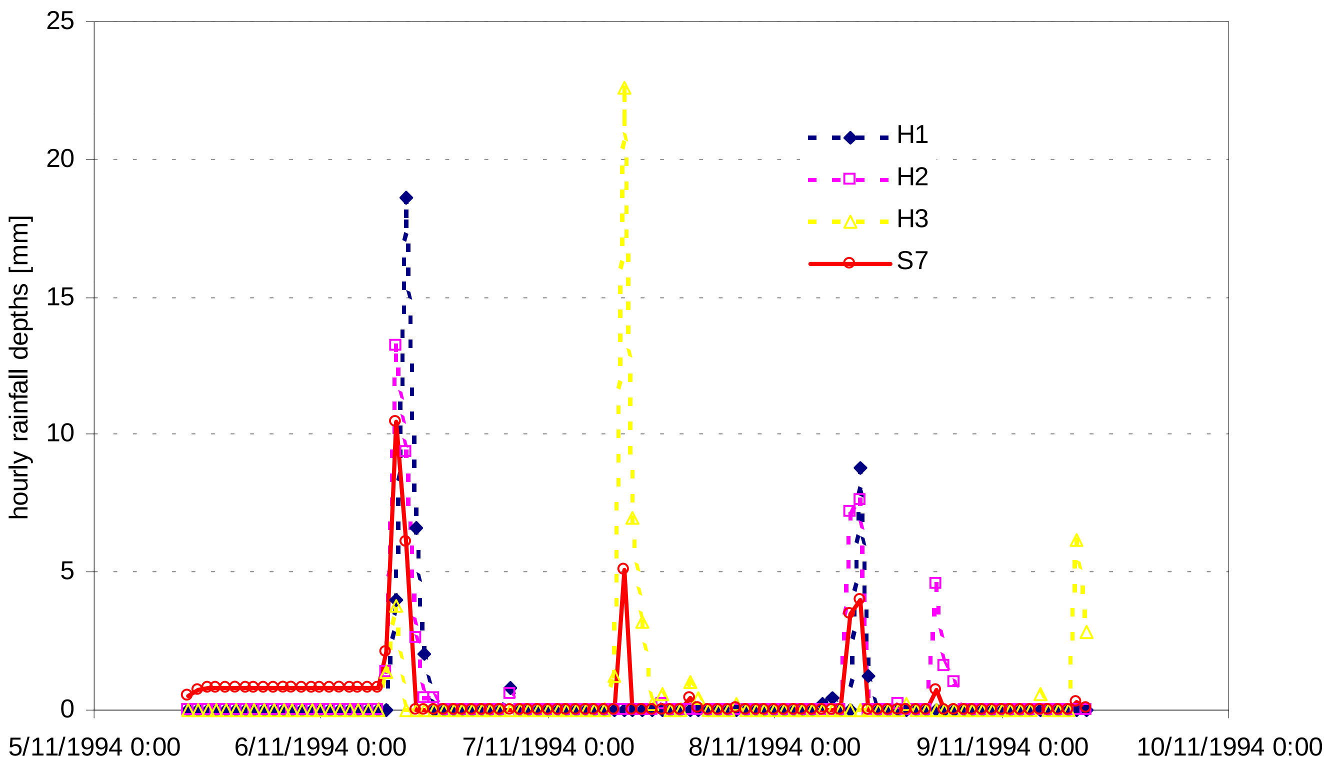


Figure 88: Simulated hyetographs for raingage 7

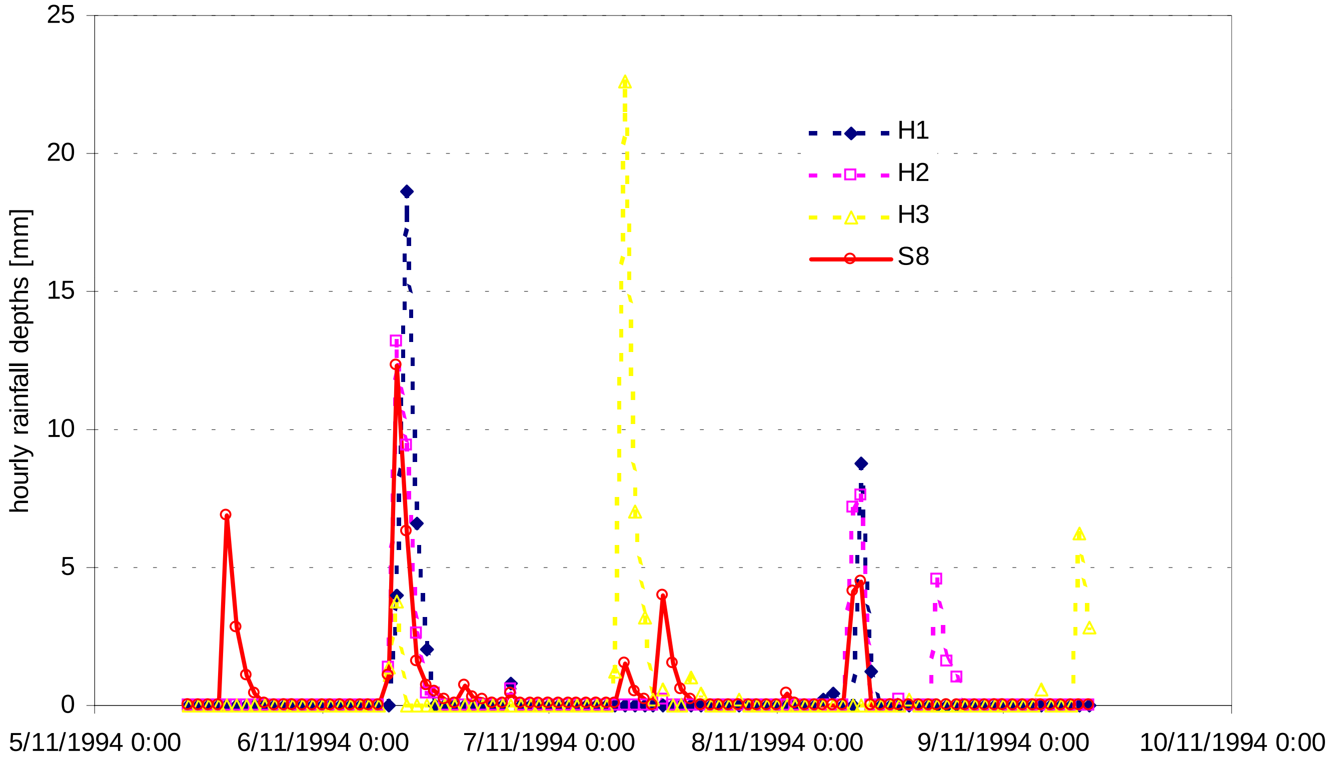


Figure 89: Simulated hyetographs for raingage 8

Table 21

Statistics of hourly rainfall depths at each gage for the month of NOVEMBER								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.90	0.89	0.91	0.88	0.89	0.91	-	-
value used on disaggregation	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
synthetic	0.90	0.89	0.90	*0.85*	*0.84*	*0.86*	*0.87*	*0.84*
Mean								
historical	0.15	0.14	0.14	0.15	0.15	0.14	-	-
value used on disaggregation	0.145	0.145	0.145	0.145	0.145	0.145	0.145	0.145
synthetic	0.16	0.14	0.15	*0.15*	*0.16*	*0.14*	*0.14*	*0.13*
Maximum value								
historical	21.6	21	24.6	20.6	19.2	16.4	-	-
value used on disaggregation	22.4	22.4	22.4	22.4	22.4	22.4	22.4	22.4
synthetic	21.6	21	24.6	*20.2*	*18.6*	*18.3*	*16.9*	*13.4*
Standard deviation								
historical	0.92	0.81	0.88	0.81	0.78	0.80	-	-
value used on disaggregation	0.873	0.873	0.873	0.873	0.873	0.873	0.873	0.873
synthetic	0.93	0.81	0.89	*0.81*	*0.77*	*0.76*	*0.75*	*0.67*
Skewness								
historical	12.06	11.71	13.90	12.04	9.81	10.18	-	-
value used on disaggregation	12.56	12.56	12.56	12.56	12.56	12.56	12.56	12.56
synthetic	12.05	11.71	13.88	*11.59*	*10.36*	*11.44*	*10.11*	*10.14*
Lag1 autocorrelation								
historical	0.39	0.44	0.39	0.49	0.50	0.53	-	-
value used on disaggregation	0.408	0.408	0.408	0.408	0.408	0.408	0.408	0.408
synthetic	0.38	0.45	0.39	*0.46*	*0.50*	*0.45*	*0.50*	*0.50*

Table 22

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of November

Gage	1	2	3	4	5	6	7	8
1								
historical	1.00	0.52	0.30	0.31	0.28	0.31	-	-
value used on disaggregation	1.00	0.52	0.30	0.37	0.36	0.32	0.43	0.29
synthetic	1.00	0.52	0.29	0.43	0.42	0.37	0.47	0.36
2								
historical	0.52	1.00	0.56	0.60	0.48	0.56	-	-
value used on disaggregation	0.52	1.00	0.56	0.82	0.70	0.75	0.77	0.71
synthetic	0.52	1.00	0.57	0.86	0.79	0.80	0.83	0.80
3								
historical	0.30	0.56	1.00	0.42	0.38	0.35	-	-
value used on disaggregation	0.30	0.56	1.00	0.50	0.49	0.42	0.75	0.44
synthetic	0.29	0.57	1.00	0.50	0.52	0.42	0.78	0.49
4								
historical	0.31	0.60	0.42	1.00	0.77	0.85	-	-
value used on disaggregation	0.37	0.82	0.50	1.00	0.91	0.86	0.68	0.70
synthetic	0.43	0.86	0.50	1.00	0.96	0.90	0.75	0.80
5								
historical	0.28	0.48	0.38	0.77	1.00	0.67	-	-
value used on disaggregation	0.36	0.70	0.49	0.91	1.00	0.74	0.66	0.53
synthetic	0.42	0.79	0.52	0.96	1.00	0.86	0.74	0.74
6								
historical	0.31	0.56	0.35	0.85	0.67	1.00	-	-
value used on disaggregation	0.32	0.75	0.42	0.86	0.74	1.00	0.53	0.62
synthetic	0.37	0.80	0.42	0.90	0.86	1.00	0.66	0.72
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.43	0.77	0.75	0.68	0.66	0.53	1.00	0.57
synthetic	0.47	0.83	0.78	0.75	0.74	0.66	1.00	0.70
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.29	0.71	0.44	0.70	0.53	0.62	0.57	1.00
synthetic	0.36	0.80	0.49	0.80	0.74	0.72	0.70	1.00

Month of December

For December, the simplified multivariate model was used in terms of linear transformation $X_s = aX_{s-1} + bV_s$ Repetitions was necessary, Δ_m was set 0.1 % and r_m was set 1000.For the preservation and control of the proportion of dry intervals the options zero threshold l_0 and probability of applying zero adjustment π_0 were set to 0.3 mm and 0.2 respectively.

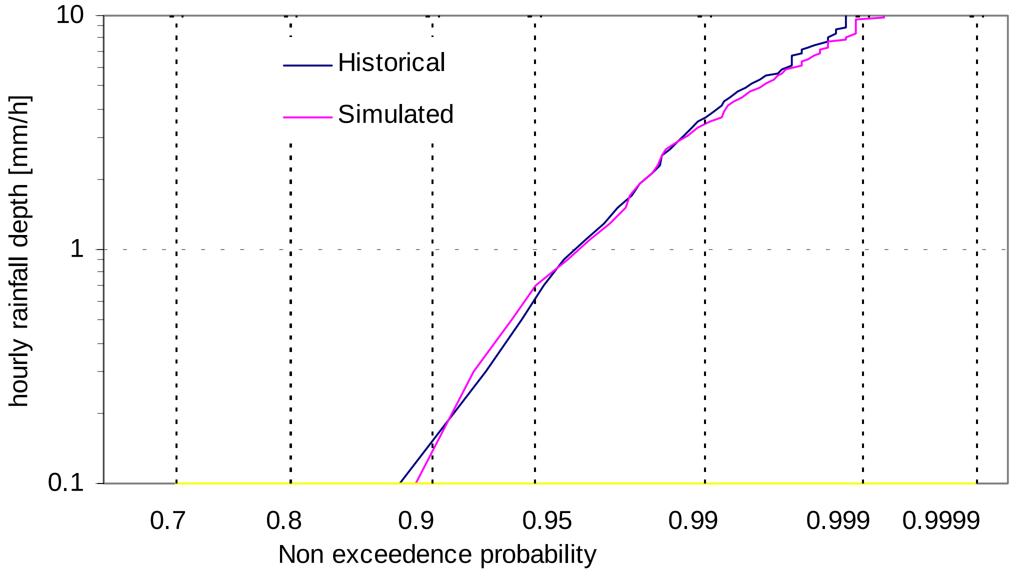


Figure90: Comparison of historical and simulated probability distribution functions of hourly rainfall depth during wet days at gage 4 for the month of December

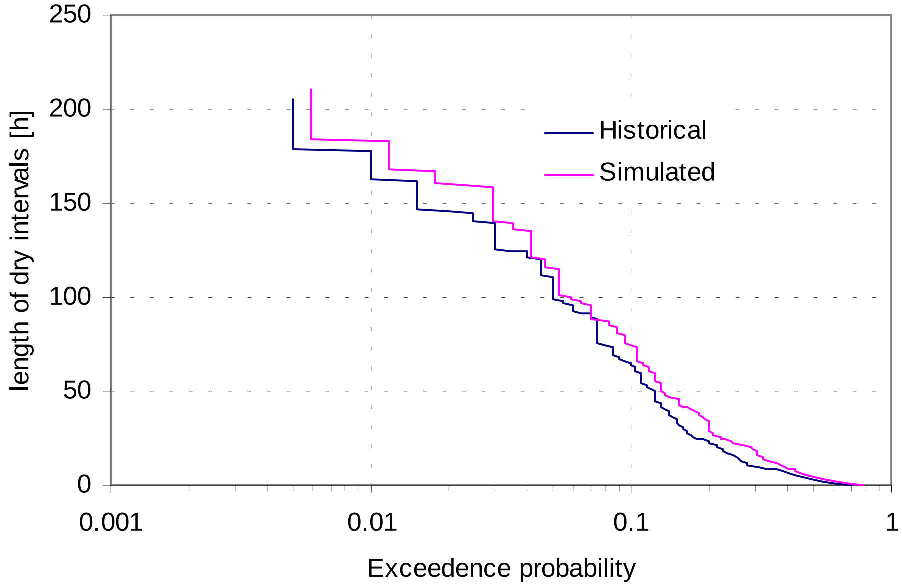


Figure 91: Comparison of historical and simulated probability distribution functions of the length of dry intervals at gage 4 for the month of December

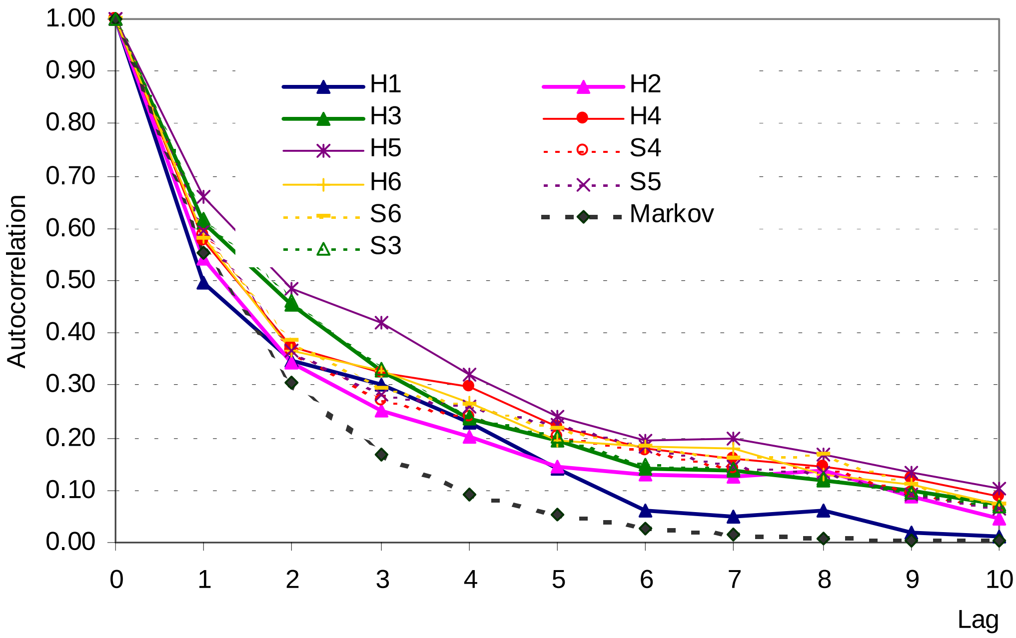


Figure 92: Comparison of autocorrelation functions of hourly rainfall as determined from historical (H1-H6) series or simulated (S3-S6) or predicted from the AR(1) for the month of December.

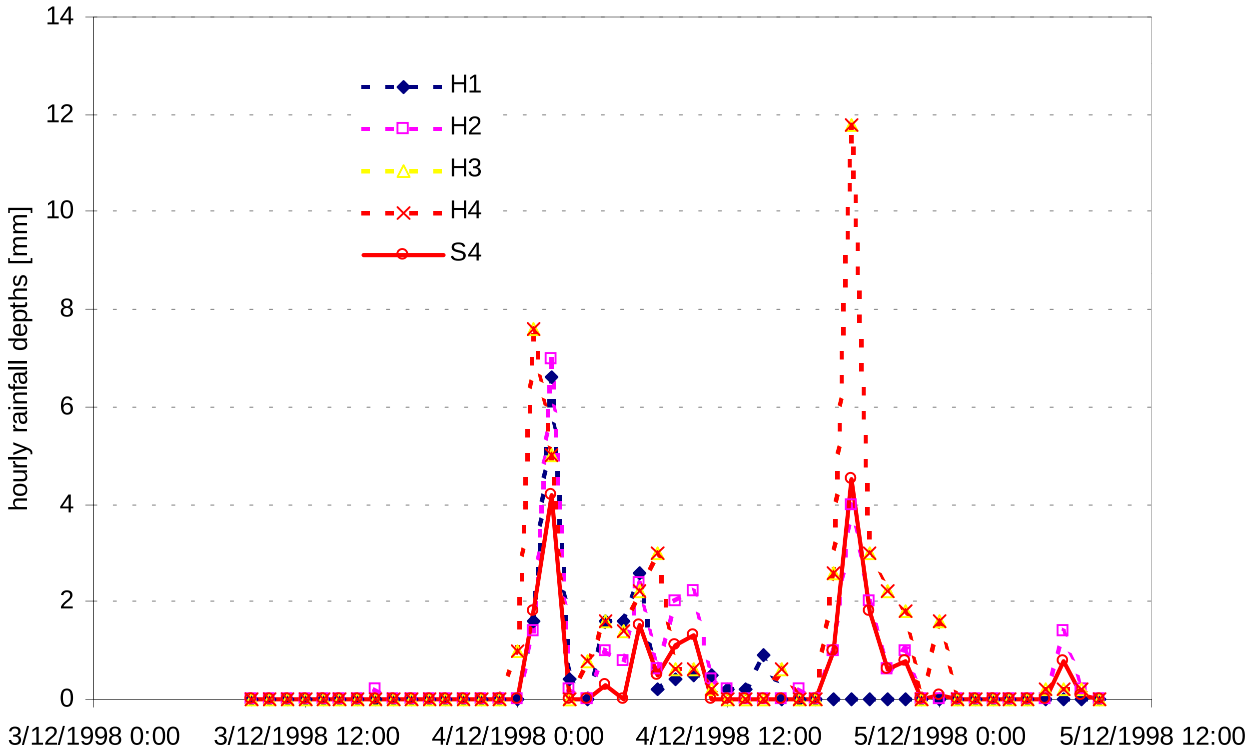


Figure 93: Comparison of historical and simulated hyetographs for raingage 4

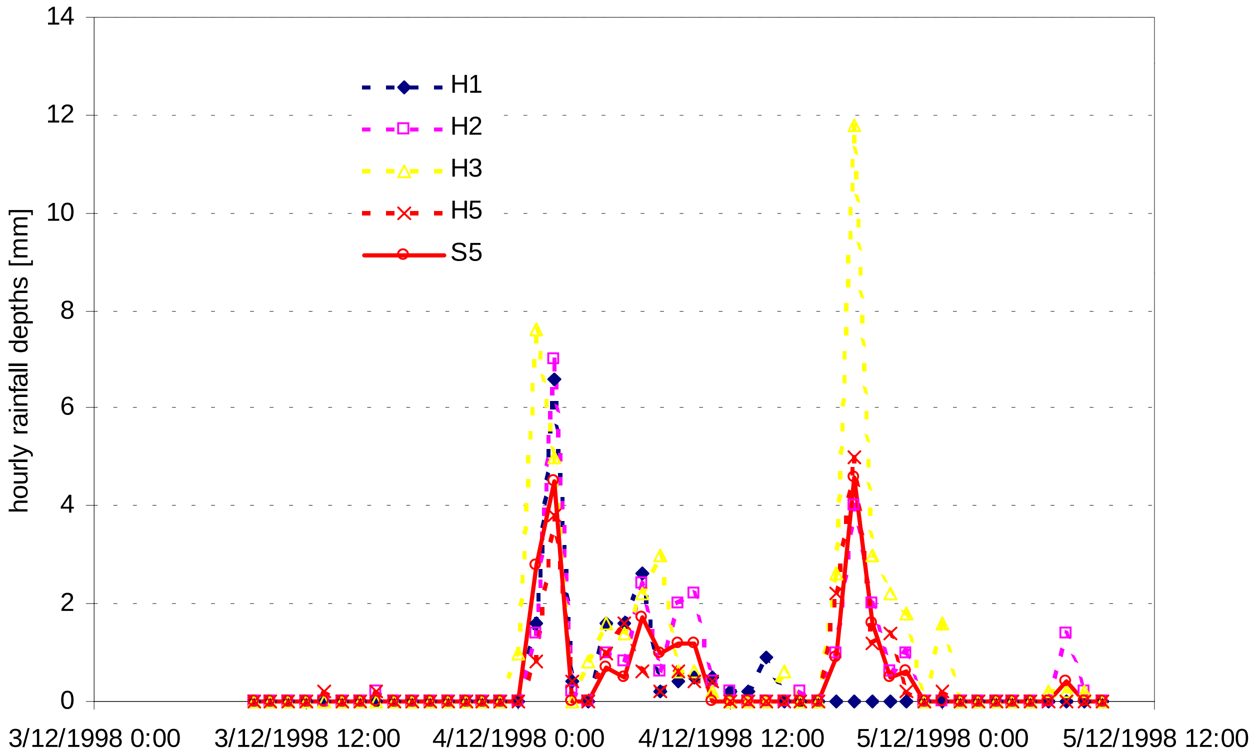


Figure 94 Comparison of historical and simulated hyetographs for raingage 5

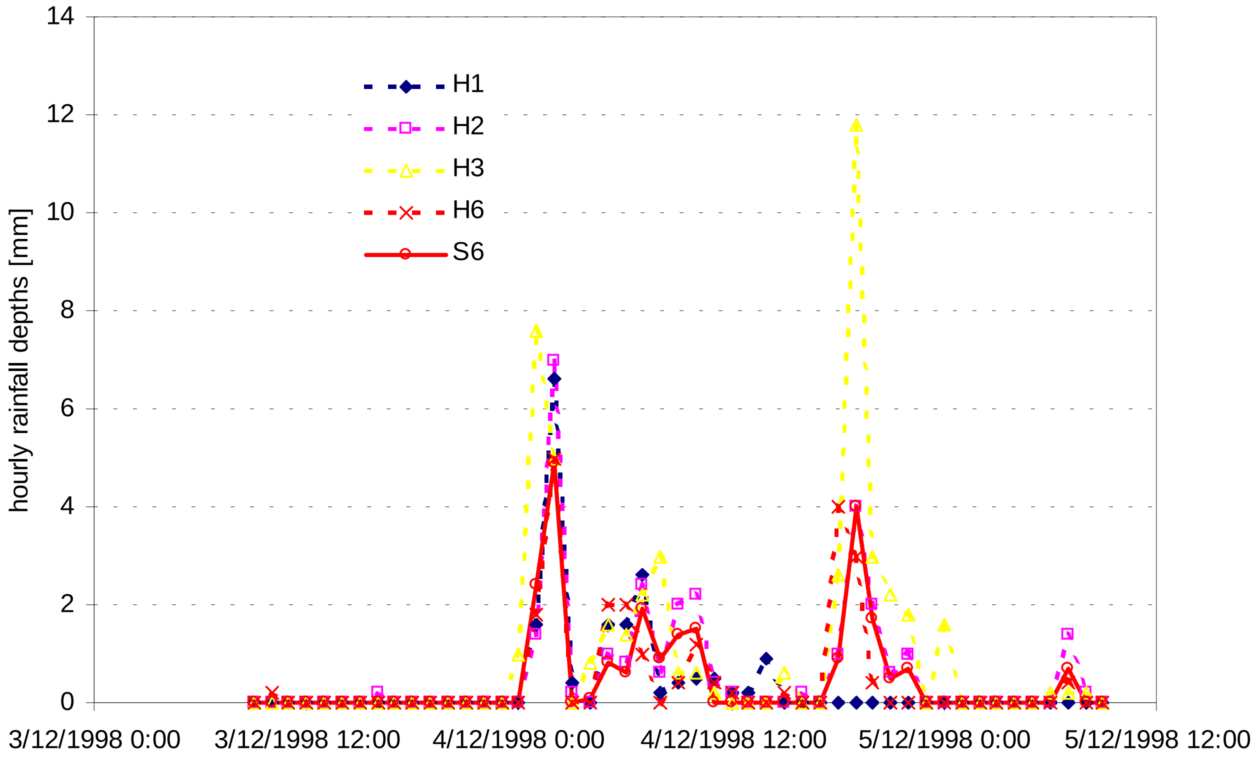


Figure 95: Comparison of historical and simulated hyetographs for raingage 6

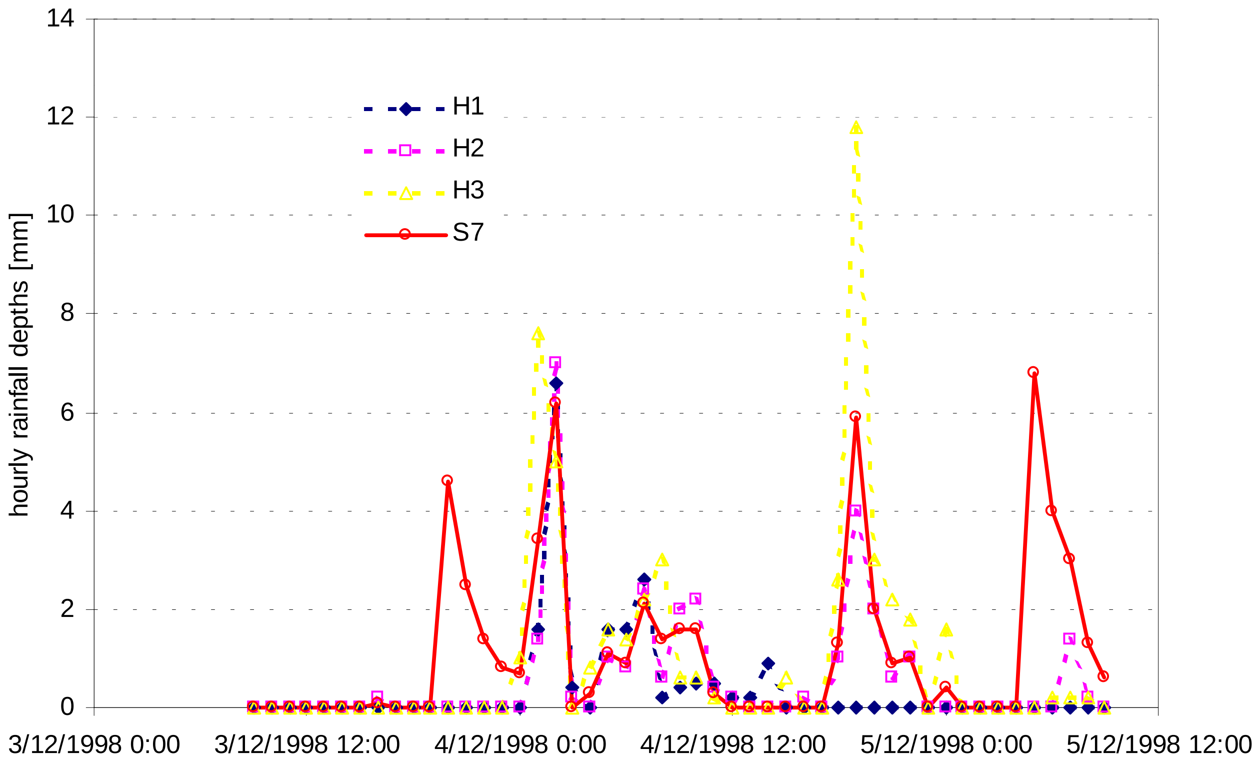


Figure 96: Simulated hyetographs for raingage 7

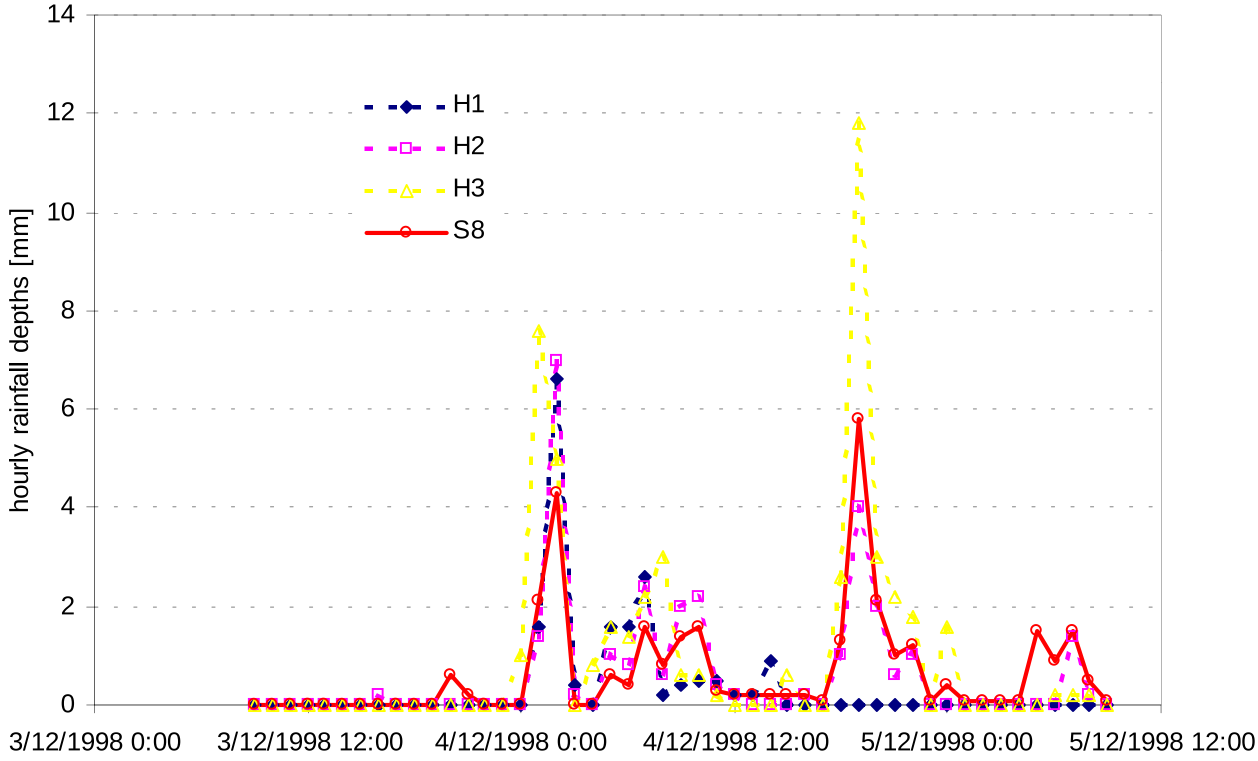


Figure 98: Simulated hyetographs for raingage 8

Table 23

Statistics of hourly rainfall depths at each gage for the month of DECEMBER								
Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Proportion dry								
historical	0.90	0.88	0.91	0.88	0.89	0.90	-	-
value used on disaggregation	0.898	0.898	0.898	0.898	0.898	0.898	0.898	0.898
synthetic	0.90	0.88	0.91	*0.89*	*0.90*	*0.89*	*0.88*	*0.82*
Mean								
historical	0.13	0.13	0.13	0.14	0.13	0.14	-	-
value used on disaggregation	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
synthetic	0.13	0.13	0.13	*0.14*	*0.13*	*0.14*	*0.15*	*0.15*
Maximum value								
historical	13.2	9.4	11.8	16.8	9.2	14.4	-	-
value used on disaggregation	11.5	11.5	11.5	11.5	11.5	11.5	11.5	11.5
synthetic	13.20	9.40	11.90	*19.50*	*16.30*	*13.10*	*10*	*15*
Standard deviation								
historical	0.66	0.64	0.70	0.74	0.64	0.69	-	-
value used on disaggregation	0.67	0.67	0.67	0.67	0.67	0.67	0.67	0.67
synthetic	0.66	0.64	0.70	*0.74*	*0.67*	*0.66*	*0.66*	*0.64*
Skewness								
historical	8.76	8.18	8.43	10.22	7.56	8.39	-	-
value used on disaggregation	8.45	8.45	8.45	8.45	8.45	8.45	8.45	8.45
synthetic	8.83	8.17	8.45	*11.71*	*10.20*	*8.58*	*6.82*	*9.61*
Lag1 autocorrelation								
historical	0.50	0.54	0.61	0.58	0.66	0.58	-	-
value used on disaggregation	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55
synthetic	0.50	0.54	0.62	*0.59*	*0.60*	*0.58*	*0.62*	*0.63*

Table 24

Lag-zero cross correlation coefficients for the eight gages at hourly level for the month of December

Gage	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
1								
historical	1.00	0.35	0.24	0.26	0.26	0.30	-	-
value used on disaggregation	1.00	0.35	0.24	0.25	0.20	0.23	0.27	0.20
synthetic	1.00	0.35	0.24	0.25	0.23	0.29	0.29	0.26
2								
historical	0.35	1.00	0.60	0.66	0.61	0.66	-	-
value used on disaggregation	0.35	1.00	0.60	0.81	0.79	0.85	0.77	0.78
synthetic	0.35	1.00	0.60	0.75	0.77	0.85	0.74	0.76
3								
historical	0.24	0.60	1.00	0.50	0.54	0.55	-	-
value used on disaggregation	0.24	0.60	1.00	0.62	0.67	0.64	0.66	0.60
synthetic	0.24	0.60	1.00	0.54	0.62	0.62	0.64	0.58
4								
historical	0.26	0.66	0.50	1.00	0.77	0.79	-	-
value used on disaggregation	0.25	0.81	0.62	1.00	0.92	0.90	0.61	0.88
synthetic	0.25	0.75	0.54	1.00	0.95	0.93	0.58	0.94
5								
historical	0.26	0.61	0.54	0.77	1.00	0.83	-	-
value used on disaggregation	0.20	0.79	0.67	0.92	1.00	0.93	0.67	0.92
synthetic	0.23	0.77	0.62	0.95	1.00	0.94	0.63	0.95
6								
historical	0.30	0.66	0.55	0.79	0.83	1.00	-	-
value used on disaggregation	0.23	0.85	0.64	0.90	0.93	1.00	0.72	0.92
synthetic	0.29	0.85	0.62	0.93	0.94	1.00	0.68	0.94
7								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.27	0.77	0.66	0.61	0.67	0.72	1.00	0.71
synthetic	0.29	0.74	0.64	0.58	0.63	0.68	1.00	0.67
8								
historical	-	-	-	-	-	-	-	-
value used on disaggregation	0.20	0.78	0.60	0.88	0.92	0.92	0.71	1.00
synthetic	0.26	0.76	0.58	0.94	0.95	0.94	0.67	1.00

Hyetos within MuDRain

Hyetos was used in this case study for generating a synthetic series to use within MuDRain. Hyetos performed in full test mode with a file containing the daily and hourly rainfall depths of raingage at Tivoli and the Bartlett-Lewis Rectangular Pulse Model Parameters estimated with the method of moments.

Estimation of the BLRPM parameters for Hyetos

Hyetos does not support the estimation of these parameters. The method used for fitting the BLRPM to the historical statistics of Tivoli is the method of moments. The set of parameters to be fitted is given by the set Θ, where Θ={λ, μ_x, κ, ϕ, α, ν}. Let k be the number of parameters to be fitted, p statistics are chosen from the historical data to fit the parameters, and these are denoted by the set *T*, where *T*={t_1, t_2, …, t_p). These can include the mean, variances, etc. of various time scales. The functions from which the various statistics can be calculated from the parameter values in the BLRPM are denoted by the set *S*={s_1(Θ), s_2(Θ), …, s_p(Θ)} and the equations for the modeled parameters are:

$$mean = \lambda\ \mu_x \mu_c \frac{\nu}{\alpha - 1} T \qquad \mu_c = 1 + \frac{\kappa}{\phi}$$

$$\mathrm{var}iance = \frac{2\nu^{2-\alpha}}{\alpha-2}\left(k_1 - \frac{k_2}{\phi}\right) - \frac{2\nu^{3-\alpha}}{(\alpha-2)(\alpha-3)}\left(k_1 - \frac{k_2}{\varphi^2}\right) + \frac{2}{(\alpha-2)(\alpha-3)}\left(k_1(T+\nu)^{3-\alpha} - \frac{k_2}{\phi^2}(\phi T + \nu)^{3-\alpha}\right)$$

$$k_1 = \left(2\lambda\mu_c\mu_x^2 + \frac{\lambda\mu_c\mu_x^2\phi}{\phi^2 - 1}\right)\left(\frac{\nu^\alpha}{\alpha - 1}\right)$$

$$k_2 = \left(\frac{\lambda\mu_c\mu_x^2\kappa}{\phi^2 - 1}\right)\left(\frac{\nu^\alpha}{\alpha - 1}\right)$$

$$\Pr(zero - rain) = \exp(-\lambda T - f_1 + f_2 + f_3)$$

$$f_1 = \frac{\lambda \nu}{\phi(\alpha-1)}\left[1+\phi\left(\kappa+\frac{\phi}{2}\right)-\frac{1}{4}\phi(\kappa+\phi)(\kappa+2\phi)+\frac{\phi(\kappa+\phi)\left(4\kappa^2+27\kappa\phi+36\phi^2\right)}{72}\right]$$

$$f_2 = \frac{\lambda \nu}{(\kappa+\phi)(\alpha-1)}\left(1-\kappa-\phi+\frac{3}{2}\phi\kappa+\phi^2+\frac{\kappa^2}{2}\right)$$

$$f_3 = \frac{\lambda\iota}{(\kappa+\phi)(\alpha-1)}\left(\frac{\nu}{\nu+T(\kappa+\phi)}\right)^{\alpha-1}\frac{\kappa}{\phi}\left(1-\kappa-\phi\frac{3}{2}\kappa\phi+\phi^2+\frac{\kappa^2}{2}\right)$$

$$Auto\operatorname{cov}ariance(lags)+\frac{k_1}{(\alpha-2)(\alpha-3)}\left\{T(s-1)+\nu\right)^{3-\alpha}+\left[T(s+1)+\nu\right]^{3-\alpha}-2(Ts+\nu)^{3-\alpha}\right\}+$$
$$+\frac{k_2}{\phi^2(\alpha-2)(\alpha-3)}\left\{2(\phi Ts+\nu)^{3-\alpha}-\left[\phi T(s-1)+\nu\right]^{3-\alpha}-\left[\phi T(s+1)+\nu\right]^{3-\alpha}\right\}$$

If $k=p$ the method of moments requires: $S=T \quad \forall p$

The equations were then formulated in an error-residual form, with an objective function such as:

$\min \sum_{i=1}^{p} w_i\left(s_i(\theta)-t_i\right)^2$ where w_i is the weight attributed to that particular statistic. The objective function is minimized in order to reduce the error between the calculated form of the statistics $s_i(\Theta)$ and the actual historical value t_i.

The BLRPM parameters for the month of march were found to be:

λ=0.1393352

μ_x=10.96869641

κ=1.2585539

ϕ= 0.105826

α= 95.13143094

ν = 8.411329581

With the synthetic series obtained by Hyetos, multivariate rainfall disaggregation from daily to hourly level was performed with MuDRain

Given:

1. an hourly point rainfall series at point 1, Tivoli, as a result of simulation with a fine time scale point rainfall model such as the BLRPM

2. several daily point rainfall series at neighbouring points 2, 3, 4, 5, namely Lunghezza, Frascati, Ponte Salario, Roma Flaminio, as a result of measurement by conventional raingages (pluviometers with daily observations).

We will produce series of hourly rainfall at points 2, 3, 4, 5, so that :

1. their daily totals equal the given daily values;
2. and their stochastic structure resembles that implied by the available historical data.

Estimation of the cross-correlation coefficients at the hourly level for MuDRain

We were able to estimate the cross-correlation coefficients between the all raingages 1 at the daily time scale, for the cross-correlation coefficients between the raingages at the hourly time scale we used the empirical expression:

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

Where:

r_{ij}^{h} is the cross-correlation coefficient between raingages *i* and *j* at the hourly time scale

r_{ij}^{d} is the cross-correlation coefficient between raingages *i* and *j* at the daily time scale

m was set equal to 3.

Graphical comparisons between disaggregated and theoretical values show that the methodology results in good preservation of the essential statistics of the rainfall process.

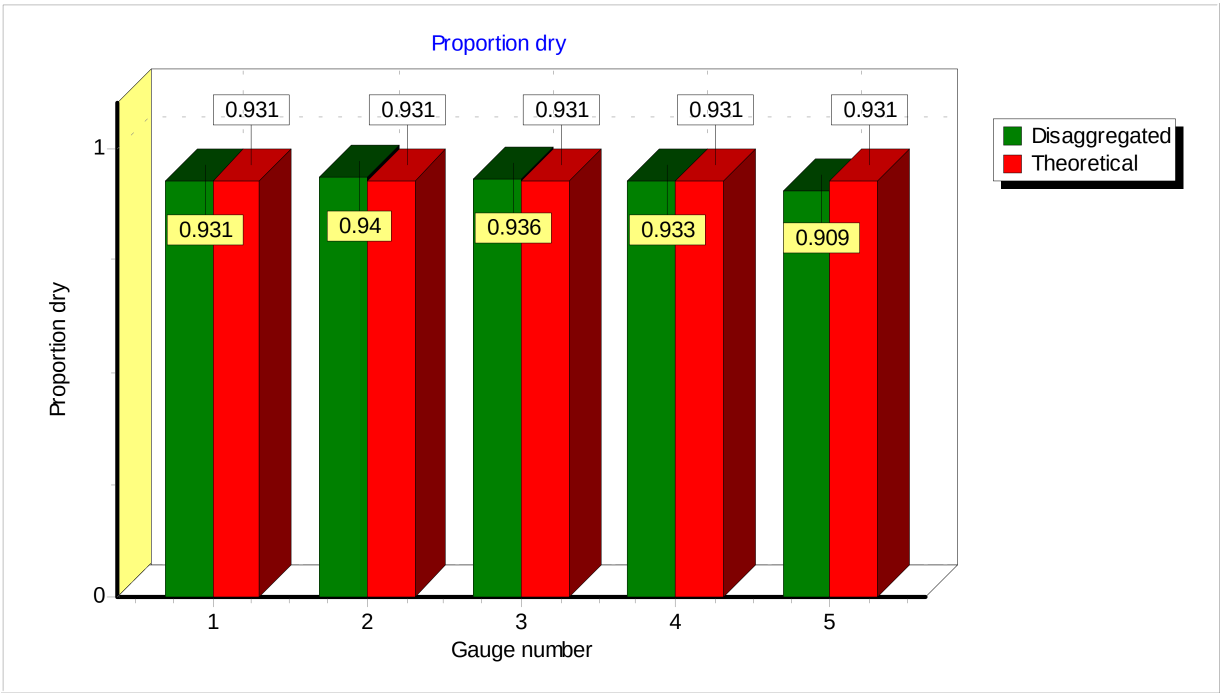


Figure 99: Proportion dry

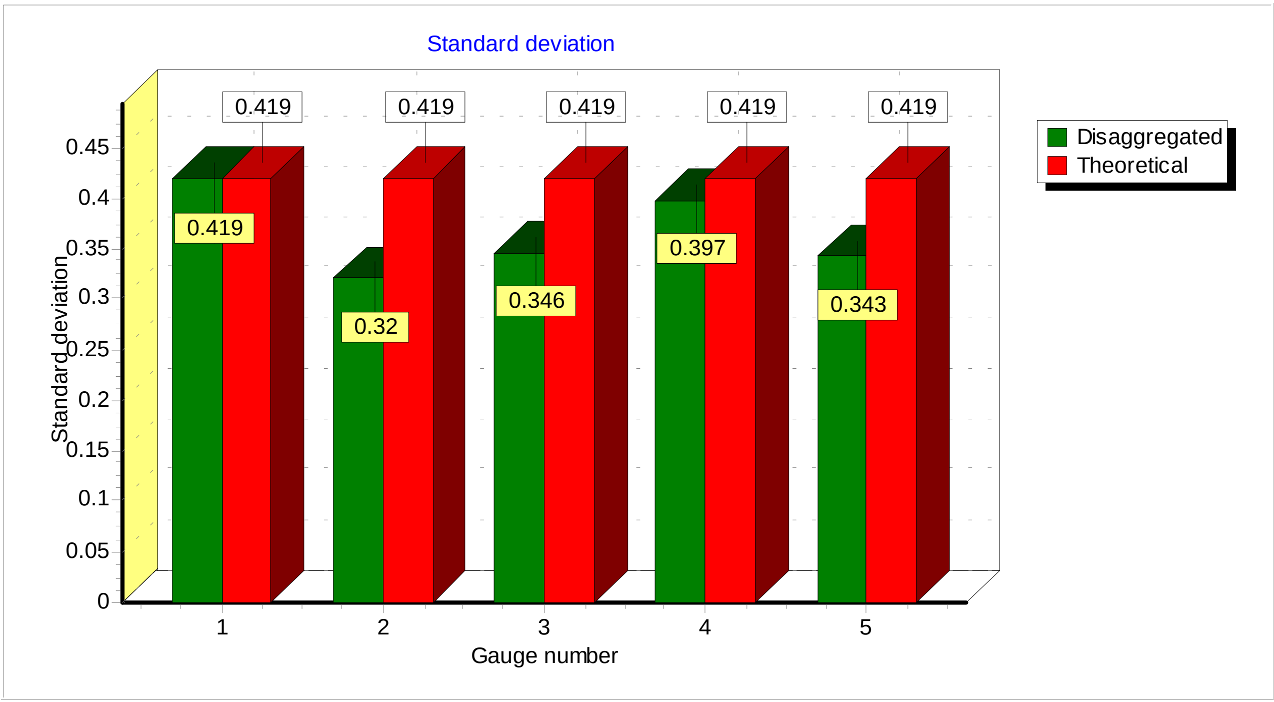


*Figure 100:*Standard deviation

*Figure 101:*Coefficients of skewness

Figure 102: Lag one autocorrelation coefficients

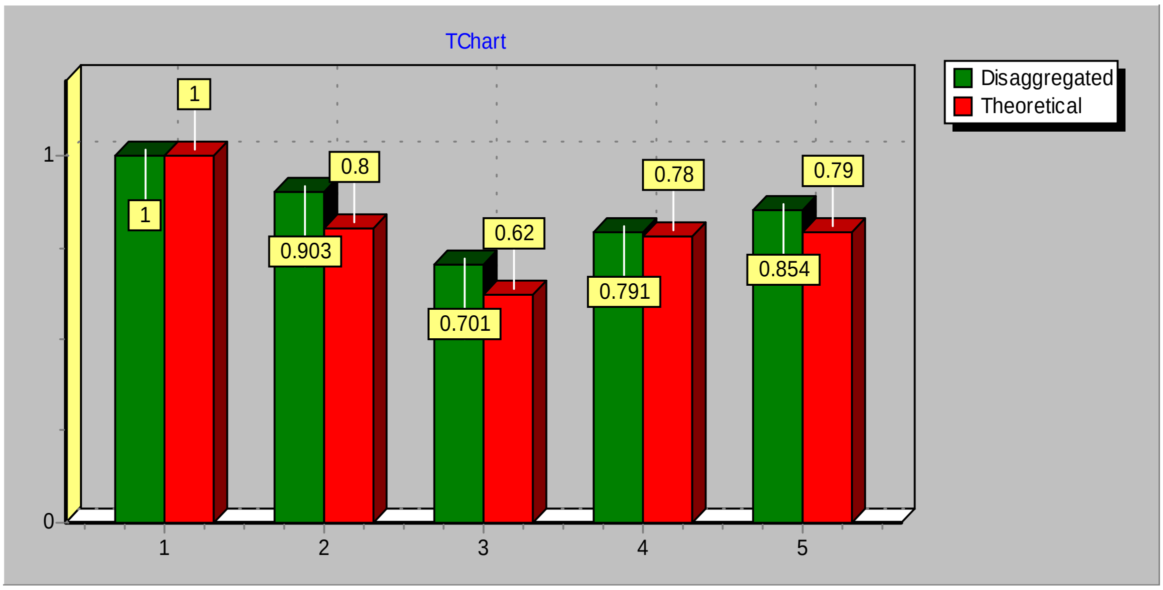


*Figure 103 :*Cross-correlation coefficients for raingage 1

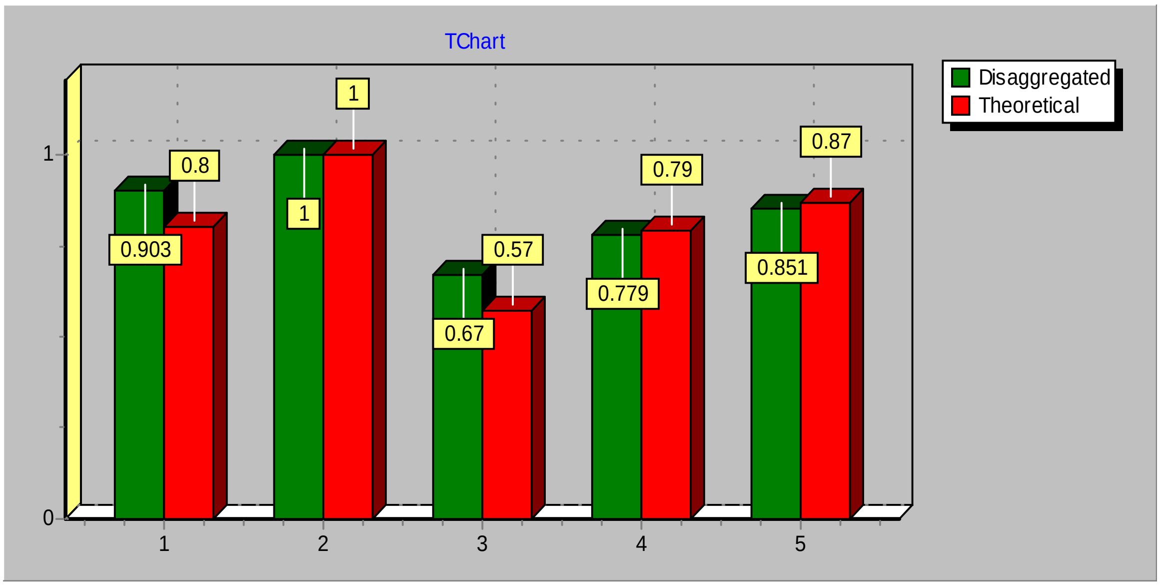


*Figure 104:*Cross-correlation coefficients for raingage 2

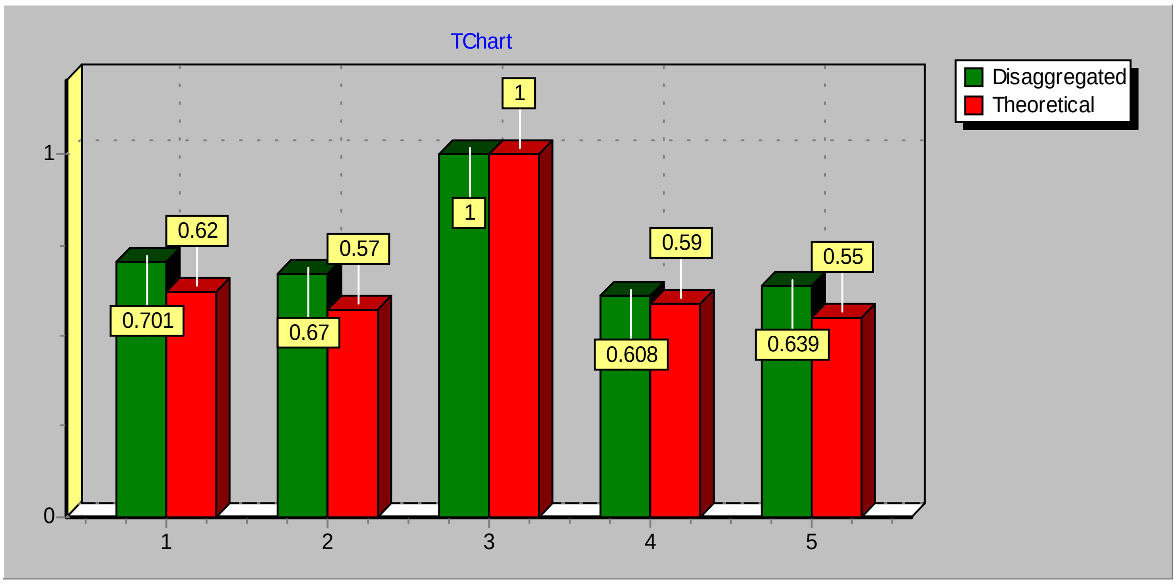


Figure 105: Cross-correlation coefficients for raingage 3

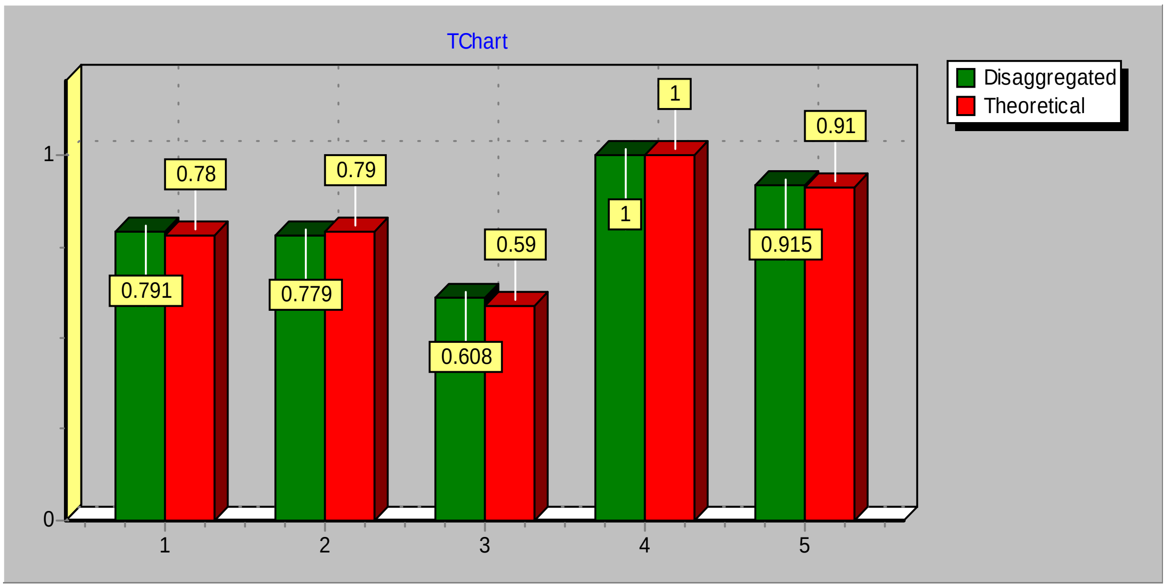


Figure 106: Cross-correlation coefficients for raingage 4

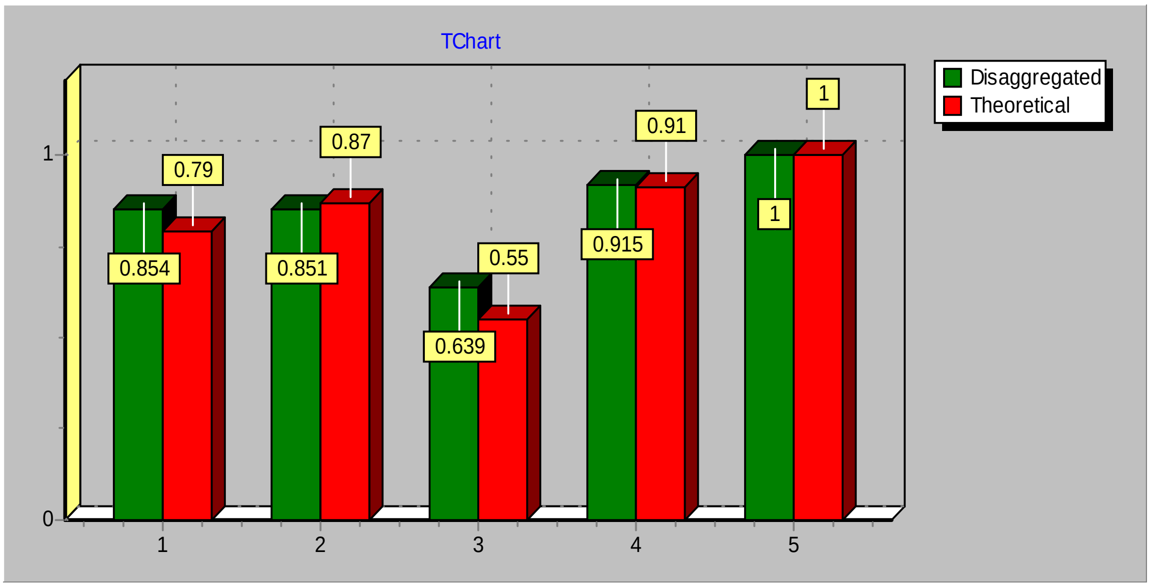


Figure 107: Cross-correlation coefficients for raingage 5

CONCLUSIONS

A common problem in hydrological studies is the limited availability of data at appropriately fine temporal and/or spatial resolution. In addition, in hydrological simulation studies a model may provide as output a synthetic series of a process (such as rainfall and runoff) at a coarse scale while another model may require as input a series of the same process at a finer scale. Disaggregation techniques therefore have considerable appeal due to their ability to increase the time or spatial resolution of hydrological processes while simultaneously providing a multiple scale preservation of the stochastic structure of the hydrological process.

The problem studied in this thesis is a particular case of a general multivariate spatial temporal rainfall disaggregation , the simultaneous rainfall disaggregation at several sites, in other words the problem is whether we are able to utilize the available single site information at the hourly scale to generate spatial and temporal consistent hourly rainfall series in other neighboring sites where only daily information is available. The methodology proposed to this aim, involves the combination of several univariate and multivariate rainfall models operating at different time scales, and was tested via a case study dealing with the disaggregation of daily historical data of eight raingages into hourly series. The data set available was six years of hourly-recorded series coming from six of the eight raingages and daily-recorded series from all raingages, covering the period from January 1994 to December 1999.

The disaggregation was performed using hourly data of three raingages only and the other three were used to allow the effectiveness of the methodology to be evaluated. Simulations were performed for each month separately to generate hourly synthetic series for gages 4,5,6,7,8 for all months of the year.

Graphical and tabulated comparisons showed that the methodology results in good preservation of important properties of the rainfall process such as marginal moments, temporal and spatial correlations and proportions and lengths of dry intervals, and in addition, in a good reproduction of the actual hyetographs.

Specifically the methodology seems to perform much better for the months characterized by a wet regime where the rainfall occurs frequently and it is more or less equally distributed over an area, i.e. rainfall events caused by stratiform

precipitation. Under these conditions the spatial cross-correlations between the raingages are high which is one of the essential hypothesis of the methodology.
For those months, characterized by a dry regime where also the cross correlations are extremely low the methodology gave good approximations of the essential statistics to be preserved. The actual hyetographs were predicted well considering the particularities of the rainfall process during the driest months. The simulated hyetographs have a realistic shape but they may depart from historical ones in their time distribution. Such departures are unavoidable as already mentioned, because of the relatively low cross-correlation coefficients.
Another variable that has not modeled explicitly in the approach followed is the length of dry intervals. Nevertheless the comparisons between historical and simulated probability distribution functions of this variable during wet days performed for each month, indicate an encouraging performance of the model.
There is considerable flexibility in the proposed scheme and hence potential for further refinement and remediation of these weaknesses.
The results attained by the methodology in combination with the empirical expression proposed and used instead of the GDSTM, for estimating the cross-correlation coefficients at hourly time scale between the gages, are extremely encouraging
Another advantage worth mentioning is the fact that the implementation of the methodology using MuDRain is extremely simple and immediate. The program automates most tasks of parameter estimation and provides tabulated and graphical comparisons of the essential statistics between theoretical and simulated values. There are also three categories of options for handling the specific difficulties and for optimizing the procedure.

Therefore the proposed methodology could be considered as useful, convenient and efficient in all hydrological applications and for all civil engineering designs where rainfall simulation is required.
Furthermore, the characterization of rainfall processes, both in space and in time, has become in the last few years of great importance for being able to solve all managing water related problems, like water catchments, water quality or ecological studies and flood alleviation schemes.
To this aim, the disaggregation model presented here has very encouraging results, and much more can be achieved in order to improve and refine statistical schemes used and spatial variability depiction.

What is MuDRain (Multivariate disaggregation of rainfall)?[#] [$] [K]

MuDRain is a methodology for spatial-temporal **disaggregation** of rainfall. It involves the combination of several univariate and multivariate rainfall **models** operating at different time scales in a disaggregation framework that can appropriately modify outputs of finer time scale models so as to become consistent with given coarser time scale series.

Potential hydrologic applications include **enhancement of historical data series** and **generation of simulated data series**. Specifically, the methodology can be applied to derive spatially consistent hourly rainfall series in raingages where only daily data are available. In addition, in a simulation framework, the methodology provides a way to take simulations of multivariate daily rainfall (incorporating spatial and temporal non-stationarity) and generate multivariate fields at fine temporal resolution.

[#] What_is

[$] What is MuDRain (Multivariate disaggregation of rainfall)?

[K] What is MuDRain (Multivariate disaggregation of rainfall)?

Why disaggregation? # $ K

A common problem in **hydrological studies** is the **limited availability** of data at appropriately fine temporal and/or spatial resolution. In addition, in hydrologic simulation studies a model may provide as output a synthetic series of a process (such as rainfall and runoff) at a coarse scale while another model may require as input a series of the same process at a finer scale. Disaggregation techniques therefore have considerable appeal due to their **ability to increase the time or space resolution** of hydrologic processes while simultaneously providing a multiple scale **preservation of the stochastic structure** of hydrologic processes.

Why_disaggregation

$ Why disaggregation?

K Disaggregation, usefulness

Why multivariate disaggregation? # $ K

The 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 raingage data into hourly rainfall. This is a common situation since detailed hydrological models often require inputs at the hourly time scale. However, historical hourly records are not as widely available as daily records. An appropriate univariate disaggregation model 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 disaggregation model could not coincide with the actual one, but would be a likely realization. Now, let us assume that there exist hourly rainfall data at a neighboring raingage. If this is the case and, in addition, the cross-correlation among the two raingages is significant (a case met very frequently in practice), then we could utilize the available hourly rainfall information at the neighboring station to generate **spatially and temporally consistent** hourly rainfall series at the raingage 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 synthesized 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.

Why_multivariate_disaggregation

$ Why multivariate disaggregation?

K Disaggregation;multivariate, usefulness

Problem formulation # $ K

CASE 1

We assume that we are given:

1. an hourly point rainfall series at point 1, as a result of either:
   - measurement by an autographic device (pluviograph) or digital sensor,
   - simulation with a fine time scale point rainfall model such as a point process model, using Hyetos (computer program for temporal rainfall disaggregation using adjusting procedures)
   - simulation with a temporal point rainfall disaggregation model applied to a series of known daily rainfall;(Hyetos)

2. several daily point rainfall series at neighboring points 2, 3, 4, 5, … as a result of either:
   - measurement by conventional raingages (pluviometers with daily observations), or
   - simulation with a multivariate daily rainfall model.

We wish to produce series of hourly rainfall at points 2, 3, 4, 5, …, so that:

1. their daily totals equal the given daily values; and
2. their stochastic structure resembles that implied by the available historical data.

We emphasize that in this problem formulation we always have an hourly rainfall series at one location, which guides the generation of hourly rainfall series at other locations. If this hourly series is not available from measurements, it can be generated using appropriate univariate simulation **models**

The essential statistics that we wish to **preserve** in the generated hourly series are:

1. the means, variances and coefficients of skewness;
2. the temporal correlation structure (autocorrelations);
3. the spatial correlation structure (lag zero cross-correlations); and
4. the proportions of dry intervals.

If the hourly data set at location 1 is available from measurement, then all these statistics apart from the cross-correlation coefficients can be estimated at the hourly time scale using this hourly record. To transfer these parameters to other locations, spatial stationarity of the process can be assumed. The stationarity hypothesis may seem an oversimplification at first glance. However, it is not a problem in practice since possible spatial

---

Problem_formulation

$ Problem formulation

K Problem formulation

nonstationarities manifest themselves in the available daily series; thus the final hourly series, which are forced to respect the observed daily totals, will reflect these nonstationarities.

CASE 2

If hourly rainfall is available at several (more than one) locations, the same modeling strategy described below can be used without any difficulty with some generalizations of the computational algorithm. In fact, having more than one point with known hourly information would be advantageous for two reasons. First, it allows a more accurate estimation of the spatial correlation of hourly rainfall depths (see discussion below) or their transformations. Second, it might reduce the residual variance of the rainfall process at each site, thus allowing for generated hyetographs closer to the real ones.

If more than one rainfall series are available at the hourly level, at least one cross-correlation coefficient of hourly rainfall can be estimated directly from these series. Then, by making plausible assumptions about the spatial dependence of the rainfall field an expression of the relationship between cross-correlation could be established (see **Estimation of crosscorrelation coefficients**.)

Modeling approach # $ K

Models involved

a. Models for the generation of multivariate fine-scale outputs. The first category includes two models that provide the required output (the hourly series).

The first model is a **the simplified multivariate rainfall model**of hourly rainfall that can preserve the statistics of the multivariate rainfall process and, simultaneously, incorporate the available hourly information at site 1, without any reference to the known daily totals at the other sites. **The statistics** considered here are the means, variances and coefficients of skewness, the lag-one autocorrelation coefficients and the lag-zero cross-correlation coefficients. All these represent statistical moments of the multivariate process. The **proportion of dry intervals**, although considered as one of the parameters to be preserved, is difficult to incorporate explicitly. However, it can be treated by an indirect manner .

The second model is a **transformation model** that modifies the series generated by the first model, so that the daily totals are equal to the given ones. This uses a (multivariate) transformation, which does not affect the stochastic properties of the series.

b. Models associated with inputs to a. above. The second category contains models which may optionally be used to provide the required input, should no observed series be available. These may include

- a multivariate daily rainfall model for providing daily rainfall depths, such as the general linear model (GLM) [*Chandler and Wheater*, 1998a, b **References**];
- a single-site model for providing hourly depths at one location such as the Bartlett-Lewis rectangular pulses model [*Rodriguez-Iturbe et al.*, 1987, 1988; *Onof and Wheater*, 1993, 1994 **References**];
- a single-site disaggregation model to disaggregate daily depths of one location into hourly depths [e.g. *Koutsoyiannis and Onof*, 2000, 2001 **References**].

Such models may be appropriate to operate the proposed disaggregation approach for future climate scenarios.

Modeling_approach
$ Modeling approach
K Modeling approach

Estimation of cross-correlation coefficients # $ K

We assume that we are given:

1. several hourly point rainfall series at points 1,2,3 as a result of measurement by an autographic device (pluviograph) or digital sensor,
2. several daily point rainfall series at neighboring points 4, 5, 6, 7,8 … as a result of measurement by conventional raingages (pluviometers with daily observations)

We are able to estimate the cross-correlation coefficients between the raingages 1,2,3 at the hourly time scale and those between 1,2,…,8 at the daily time scale.

We need to estimate the cross-correlation coefficients between all raingages at the hourly time scale.

For this purpose we use the empirical relationship: $r_{ij}^{h} = \left(r_{ij}^{d}\right)^{m}$

where:

r_{ij}^{h} is the cross-correlation coefficient between raingages *i* and *j* at the hourly time scale

r_{ij}^{d} is the cross-correlation coefficient between raingages *i* and *j* at the daily time scale

m is an exponent that can be estimated by regression using the known cross-correlation coefficients at the hourly and daily time scale or, in case no hourly data is available, its value can be assumed approximately in the range 2 to 3.(*Fytilas P*. 2002 **References**)

Estimation_of_crosscorrelation_coefficients

$ Estimation of cross-correlation coefficients

K Estimation of cross-correlation coefficients

The simplified multivariate rainfall model # $ K

For n locations, we may assume that **the simplified multivariate rainfall model** is an AR(1) process, expressed by

$$\mathbf{X}_s = \mathbf{a}\,\mathbf{X}_{s-1} + \mathbf{b}\,\mathbf{V}_s \quad (1)$$

where $\mathbf{X}_s := \left[X_s^1, X_s^2, \ldots, X_s^n\right]^T$ represents the hourly rainfall at time (hour) s at n locations, **a** and **b** are ($n \times n$) matrices of parameters and $\mathbf{V}_s$ (s = …, 0, 1, 2, …) is an independent identically distributed (iid) sequence of size n vectors of innovation random variables (so that the innovations are both spatially and temporally independent). The time index s can take any integer value. $\mathbf{X}_s$ are not necessarily standardized to have zero mean and unit standard deviation, and obviously they are not normally distributed. On the contrary, their distributions are very skewed. The distributions of $\mathbf{V}_s$ are assumed three-parameter Gamma.

Equations to estimate the model parameters **a** and **b** and the moments of $\mathbf{V}_s$ directly from the statistics to be preserved are given for instance by *Koutsoyiannis* [1999] for the most general case. In the special case examined here, for convenience, the parameter matrix **a** is assumed diagonal, which suffices to preserve the essential statistics, and is given by:

$$a = diag\left(Cov\left[X_s^l, X_{s-1}^l\right] / Var\left[X_{s-1}^l\right]\right), \; l = 1, \ldots, n \quad (2)$$

The parameter matrix **b** is determined from

$$b \cdot b^T = Cov\left[X_s, X_s\right] - a_s \cdot Cov\left[X_{s-1}, X_{s-1}\right] a_s \quad (3)$$

If **b** is assumed lower triangular, which facilitates handling of the known hourly rainfall at site 1, then it can be determined from $\mathbf{b}\,\mathbf{b}^T$ using Cholesky decomposition.

Another group of model parameters are the moments of the auxiliary variables $\mathbf{V}_s$. The first moments (means) are obtained by

$$E\left[V_s\right] = b^{-1}(I - a) \cdot E\left[X_s\right] \quad (4)$$

where **I** is the identity matrix. The variances are by definition 1, i.e., Var[$\mathbf{V}_s$] = $[1, \ldots, 1]^T$ and the third moments are obtained in terms of $\mu_3[\mathbf{X}_s]$, the third moments of $\mathbf{X}_s$, by

$$\mu_3\left[V_s\right] = \left(b^{(3)}\right)^{-1}\left(I - a^{(3)}\right) \cdot \mu_3\left[X_s\right] \quad (5)$$

where $\mathbf{a}^{(3)}$ and $\mathbf{b}^{(3)}$ denote the matrices whose elements are the cubes of **a** and **b**, respectively

At the generation phase, V_s^1, the first component of $\mathbf{V}_s$, is calculated from the series of X_s^1 rather than generated. Given that **b** is lower triangular, its first row will have only one nonzero item, call it b^1, so that from (1)

The_simplified_multivariate_rainfall_model

$ The simplified multivariate rainfall model

K The simplified multivariate rainfall model

$$X_s^1 = a^1 X_{s-1}^1 + b^1 V_s^1 \tag{6}$$

which can be utilized to determine V_s^1 This can be directly expanded to the case where several gages of hourly information are available provided that **b** is lower triangular.

Alternatively, the model can be expressed in terms of some nonlinear transformations X_s^* of the hourly depths $\mathbf{X}_s$ (see **<u>Specific difficulties</u>**), in which case (1) is replaced by

$$X_s^* = aX_{s-1}^* + bV_s \tag{7}$$

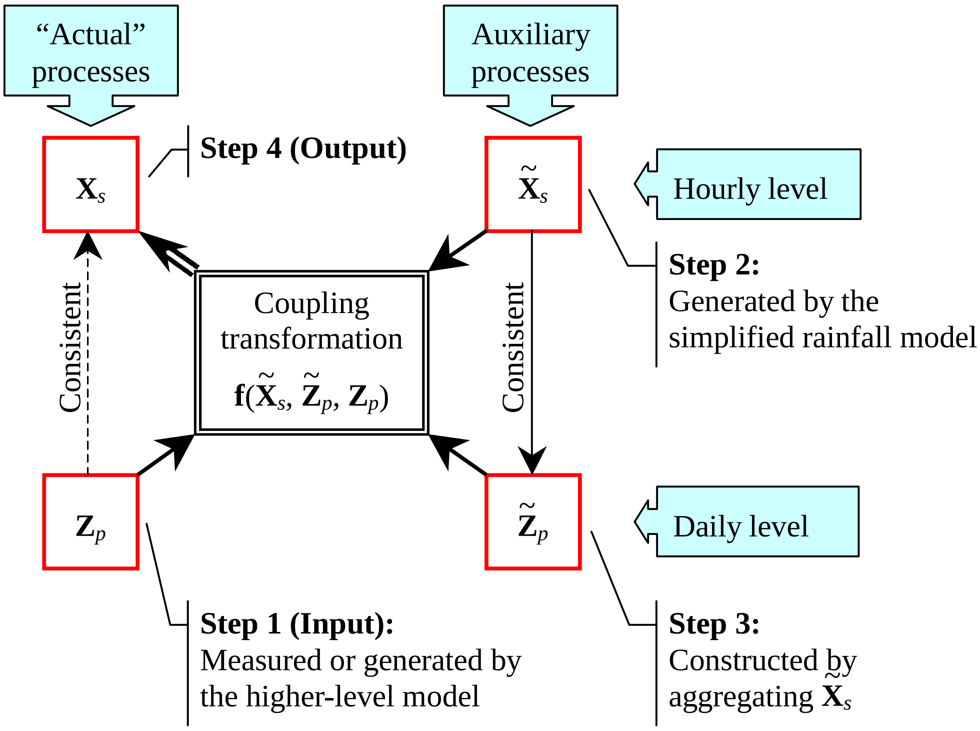
The transformation model # $ K

Transformations that can modify a series generated by any stochastic process to satisfy some additive property (i.e. the sum of the values of a number of consecutive variables be equal to a given amount), without affecting the first and second order properties of the process, have been studied previously by *Koutsoyiannis* [1994] and *Koutsoyiannis and Manetas* [1996]. These transformations, more commonly known as **adjusting procedures**, are appropriate for univariate problems, although they can be applied to multivariate problems as well, but in a repetition framework. More recently, *Koutsoyiannis* [2001](**References**) has studied a true multivariate transformation of this type and also proposed a generalized framework for coupling stochastic models at different time scales.

This framework, specialized for the problem examined here, is depicted in the following **schematic representation** where $\mathbf{X}_s$ and $\mathbf{Z}_p$ represent the "actual" hourly- and daily-level processes, related by

$$\sum_{s=(p-1)k}^{pk} X_s = Z_p \tag{8}$$

where *k* is the number of fine-scale time steps within each coarse-scale time step (24 for the current application), $\tilde{X}_s$ and $\tilde{Z}_p$ denote some auxiliary processes, represented by the simplified rainfall model in our case, which also satisfy a relationship identical to (8).



The_transformation_model
$ The transformation model
K Coupling transformation ;adjusting procedures

The problem is: Given a time series $\mathbf{z}_p$ of the actual process $\mathbf{Z}_p$, generate a series $\mathbf{x}_s$ of the actual process $\mathbf{X}_s$. To this aim, we first generate another (auxiliary) time series $\tilde{x}_s$ using the simplified rainfall process $\tilde{X}_s$. The latter time series is generated independently of $\mathbf{z}_p$ and, therefore, $\tilde{x}_s$ do not add up to the corresponding $\mathbf{z}_p$, as required by the additive property (8), but to some other quantities, denoted as $\tilde{z}_p$. Thus, in a subsequent step, we modify the series $\tilde{x}_s$ thus producing the series $\mathbf{x}_s$ consistent with $\mathbf{z}_p$ (in the sense that $\mathbf{x}_s$ and $\mathbf{z}_p$ obey $\sum_{s=(p-1)k}^{pk} X_s = Z_p$ (8)) without affecting the stochastic structure of $\tilde{x}_s$. For this modification we use a so-called coupling transformation, i.e., a linear transformation, **f(** $\tilde{X}_s$ **,** $\tilde{Z}_p$**,** $\mathbf{Z}_p$) whose outcome is a process identical to $\mathbf{X}_s$ and consistent to $\mathbf{Z}_p$.

Let $X_p^* = \left[X_{(p-1)k+1}^T, \ldots, X_{pk}^T\right]^T$ be the vector containing the hourly values of the 24 hours of any day *p* for all examined locations (i.e., the 24 vectors $\mathbf{X}_s$ for *s* = (*p* – 1)*k* + 1 to *s* = *pk*; for 5 locations, X_p^* contains 24 × 5 = 120 variables). Let also $Z_p^* = \left[Z_p^T, Z_{p+1}^T, X_{(p-1)k}^T\right]^T$ be a vector containing

- (a) the daily values $\mathbf{Z}_p$ for all examined locations,
- (b) the daily values $\mathbf{Z}_{p+1}$ of the next day for all locations, and
- (c) the hourly values $\mathbf{X}_{(p-1)k}$ of the last hour of the previous day *p* – 1 for all locations.

This means that for 5 locations Z_p^* contains 3 × 5 = 15 variables in total. Items (b) and (c) of the vector Z_p^* were included to assure that the transformation will preserve not only the covariance properties among the hourly values of each day, but the covariances with the previous and next days as well. Note that at the stage of the generation at day *p* the hourly values of day *p* – 1 are known (therefore, in Z_p^* we enter hourly values of the previous day) but the hourly values of day *p* + 1 are not known (therefore, in Z_p^* we enter daily values of the next day, which are known). In an identical manner, we construct the vectors $\tilde{X}_p^*$ and $\tilde{Z}_p^*$ from variables $\tilde{X}_s$ and $\tilde{Z}_p$.

Koutsoyiannis [2001] showed that the coupling transformation sought is given by

$$X_p^* = \tilde{X}_p^* + h\left(Z_p^* - \tilde{Z}_p^*\right) \tag{9}$$

where

$$h = Cov\left[X_p^*, Z_p^*\right]\left\{Cov\left[Z_p^*, Z_p^*\right]\right\}^{-1} \tag{10}$$

The quantity $h\left(Z_p^* - \tilde{Z}_p^*\right)$ in (9) represents the correction applied to $\tilde{X}$ to obtain **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 $\sum_{s=(p-1)k}^{pk} X_s = Z_p$). 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

$$\Delta = \left\| h\left(Z_p^* - \tilde{Z}_p^*\right) \right\| / \left(m\sigma_x\right) \qquad (11)$$

where *m* is the common size of X_p^* and $\tilde{X}_p^*$, σ_x is standard deviation of hourly depth (common for all locations due to stationarity assumption) and **||.||** denotes the Euclidian norm..

Given the daily process $\mathbf{Z}_p$ and the matrix **h**, which determines completely the transformation, the steps followed to generate the hourly process $\mathbf{X}_s$ are the following:

1. Use the simplified rainfall model (1) or (8) to produce a series $\tilde{X}_s$ for all hours of the current day *p* and the next day *p* + 1, without reference to $\mathbf{Z}_p$.

2. At day *p* evaluate the vectors Z_p^* and $\tilde{Z}_p^*$ using the values of $\mathbf{Z}_p$ and $\tilde{X}_s$ of the current and next day, and $\mathbf{X}_s$ of the previous day.

3. Determine the quantity $h\left(Z_p^* - \tilde{Z}_p^*\right)$ and the measure of correction *Δ*. If *Δ* is greater than an accepted limit Δ_m, repeat steps 1-3 (provided that the number of repetitions up to the current repetition has not exceeded a maximum allowed number r_m, which is set to avoid unending loops).

4.Apply the coupling transformation to derive X_p^* of the current period.

5.Repeat steps 1 and 4 for all periods.

Specific difficulties [#] [$] [K]

Here we describe how to handle the peculiarities of the rainfall process at a fine time scale in the multivariate modeling scheme.

Negative values. The negative values, unavoidably generated by any linear stochastic model when the coefficient of variation is high (possibly in a high proportion but with low values), are not a major problem in our case. They are simply truncated to zero, thus having a beneficial effect in preserving the proportion of dry intervals (as also shown in next paragraph). A negative effect is the fact that truncation may be a potential source of bias to statistical properties that are to be preserved. Specifically, it is anticipated to result in overprediction of cross-correlations, as it is very probable that negative values are contemporary.

Dry intervals. As already mentioned, the proportion of dry intervals cannot be preserved by linear stochastic models in an explicit manner. However, after rounding off the generated values, a significant number of zero values emerges, which is added to the significant number of zero values resulting from the truncation of negative values. The total percentage of zero values resulting this way can be comparable to (usually somewhat smaller than) the historical probability dry. It was demonstrated that we can match exactly the historical probability dry by slightly modifying the rounding-off rule. For the multivariate case, the following technique was found effective: A proportion π_0 of the very small positive values, chosen at random among the generated values that are smaller than a threshold l_0 (e.g., 0.1-0.3 mm), are set to zero.

An alternative technique, based on a two-state (wet-dry) representation of hourly rainfall within a rainy day, can be also used. According to this technique, at periods when the known hourly time series indicates dry condition (zero depth) the unknown hourly time series are stimulated, with a specified probability φ_0, to take zero depth as well.

Preservation of skewness. Although the coupling transformation preserves the first and second order statistics of the processes, it does not ensure the preservation of third order statistics. Thus, it is anticipated that it will result in underprediction of skewness. However, the repetition technique (see **transformation model**) can result in good approximation of skewness.

Homoscedasticity of innovations. By definition, the innovations $\mathbf{V}_s$ in the simplified multivariate rainfall model (see **the simplified multivariate rainfall model**) are homoscedastic, in the sense that their variances are constant, independent of the values of rainfall depths $\mathbf{X}_s$. Therefore, if, for instance, we estimate (or generate) the value at location 2, given that at location 1, we assume that the conditional variance is constant and independent of the value at location 1. This, however, does not comply with reality: by examining simultaneous hyetographs at two locations we can observe that the variance is larger during the periods of high rainfall (peaks) and smaller in periods of low rainfall (heteroscedasticity).

[#] Specific_difficulties
[$] Specific difficulties
[K] dry intervals; negative values; homoscedasticity; skewness

As a result of this inconsistency, synthesized hyetographs will tend to have unrealistically similar peaks. To mitigate this problem we can apply a <u>nonlinear transformation</u> to rainfall depths

<u>The first candidate</u> nonlinear transformation is the logarithmic one,

$$X_s^* = \ln(X_s + \zeta) \qquad (12)$$

with constants $\boldsymbol{\zeta} > \mathbf{0}$, where the logarithmic transformation should be read as an item to item one. The stationarity assumption allows considering all items of vector $\boldsymbol{\zeta}$ equal to a constant ζ. This transformation would be an appropriate selection if ζ was estimated so that the transformed series of known hourly depths have zero skewness, in which case the transformed variables could be assumed to be normally distributed. Then, preservation of first and second order properties of the untransformed variables is equivalent to preservation of first- and second-order statistics of the transformed variables [*Koutsoyiannis*, 2001] (**<u>References</u>**). However, evidence from the examined data sets shows that the skewness of the transformed variables increases with increasing ζ and it still remains positive even if very small ζ are chosen. This means that the lognormal assumption is not appropriate for hourly rainfall.

<u>A second candidate</u> is the power transformation

$$X_s^* = X_s^{(m)} \qquad (13)$$

where the symbol (*m*) means that all items of the vector $\mathbf{X}_s$ are raised to the power *m* (item to item) where $0 < m < 1$. The stationarity assumption complies with the assumption that *m* is the same for all items. The preservation of the statistics of the untransformed variables does not necessarily lead to the preservation of the corresponding statistics of the transformed variables. However, the discrepancies are expected to be low if *m* is not too low (e.g., for $m \geq 0.5$).

References[#$K]

Chandler, R. and H. Wheater, Climate change detection using Generalized Linear Models for rainfall, A case study from the West of Ireland, I, Preliminary analysis and modelling of rainfall occurrence, Technical report, no. 194, Department of Statistical Science, University College London, 1998a (http://www.ucl.ac.uk/Stats/research/abstracts.html)

Chandler, R. and H. Wheater, Climate change detection using Generalized Linear Models for rainfall, A case study from the West of Ireland, II, Modelling of rainfall amounts on wet days, Technical report, no. 195, Department of Statistical Science, University College London, 1998b (http://www.ucl.ac.uk/Stats/research/abstracts.html)

Fytilas P., Multivariate rainfall disaggregation at a fine time scale 2002,diploma thesis submitted at the *University of Rome "La Sapienza"*)

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

Koutsoyiannis, D., A stochastic disaggregation method for design storm and flood synthesis, *J. of Hydrol.*, 156, 193-225, 1994.

Koutsoyiannis, D., Optimal decomposition of covariance matrices for multivariate stochastic models in hydrology, *Water Resour. Res.*, 35(4), 1219-1229, 1999.

Koutsoyiannis, D., Coupling stochastic models of different time scales, *Water Resour. Res.*, 37(2), 379-392, 2001.

Koutsoyiannis, D., Multivariate rainfall disaggregation at a fine time scale (unpublished paper; in review for *Water Resour. Res.*); 2002.

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

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

Koutsoyiannis, D., 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. (Presentation also available on line at http://www.itia.ntua.gr/e/docinfo/59/)

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

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

Koutsoyiannis, D., and Th. Xanthopoulos, A dynamic model for short-scale rainfall disaggregation, *Hydrol. Sci. J.*, 35(3), 303-322, 1990.

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

Onof, C. 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.

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

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

[#] References

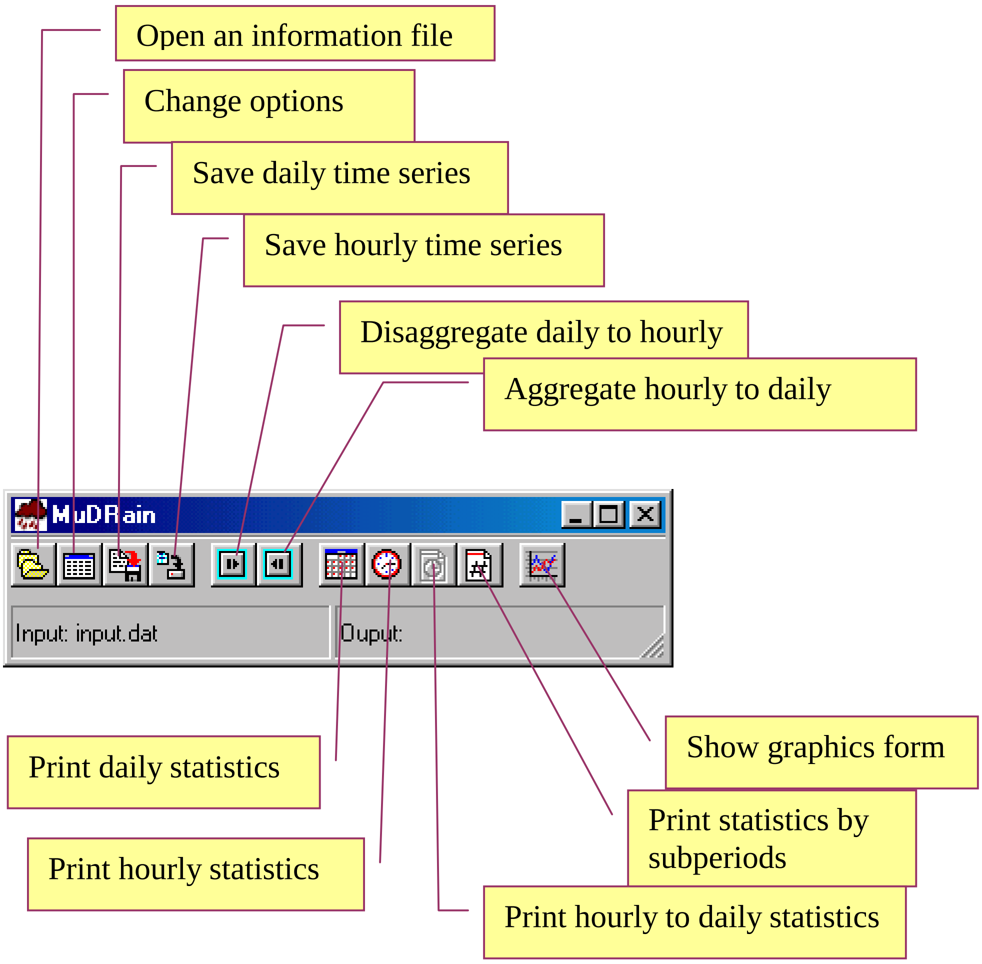
[$] References

[K] References

Main Form [#] [$] [K]

This is the main form of **MuDRain**. Use the on-screen hints of toolbar buttons displayed by placing and pausing the mouse pointer on them.
Here is a summary of toolbar buttons description:

The other forms of the software application **Options form**, the **Graphs form** and the Help About form appear by clicking the appropriate buttons of this form, whereas the **Visual output form** appears after opening an information file; see **Input files format**



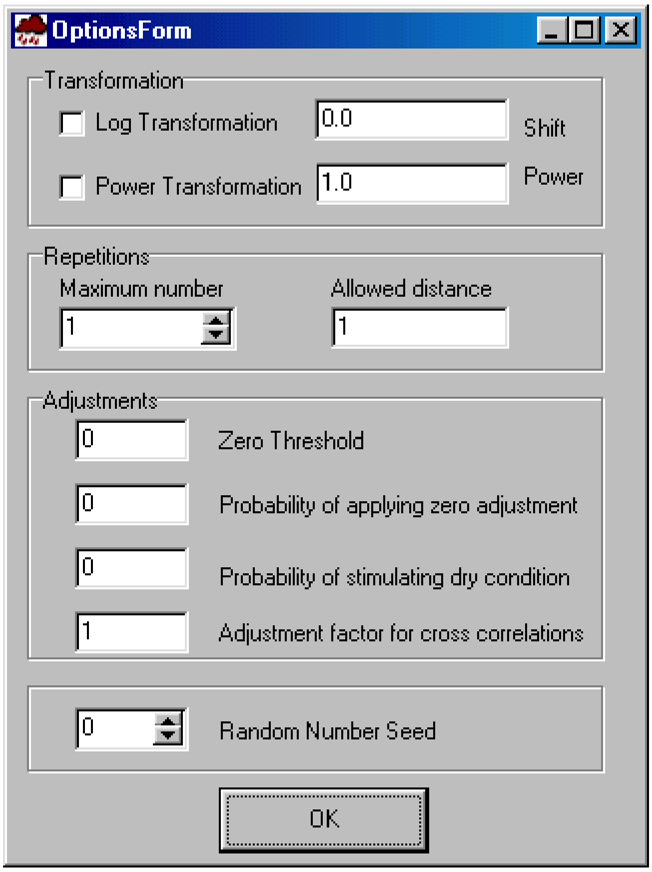
[#] Main_Form
[$] Main Form
[K] file, open, save; statistics, print

Options form[#] [$] [K]

Activate this form by pressing the appropriate button in the **Main form**.
The program offers three categories of options that must be specified by the user (for justification of these options see **specific difficulties**:
(a) the use or not of **repetition** in the generation phase,
(b) the use or not of one of the **transformations** and
(c) the use or not of the **two-state representation of hourly rainfall**.

In case of the adoption of each of these options, the user must specify some additional parameters for the generation, which are:
for (a), **the maximum allowed distance Δ_m** and **the maximum allowed number of repetitions r_m** (see **transformation model**);
for (b) the transformation constant ζ or m (as defined in equation (12) or (13), respectively see **specific difficulties**); and
for (c) the probability φ_0, to stimulate dry state in each of the locations. Two additional parameters are used, which are related to the rounding off rule of generated hourly depths, i.e. the proportion π_0 and the threshold l_0.

In the current program configuration, the options and the additional parameters must be specified by the user in a trial-and-error manner, i.e., starting with different trial values until the resulting statistics in the synthetic series match the actual ones. This can be seen as a fine-tuning of the model, which is manual. An automatic fine-tuning procedure, based on stochastic optimization, seems to be possible but has not been studied so far.



[#] Options_Form
[$] Options Form
[K] Options Form; options, repetitions; allowed distance; two state representation

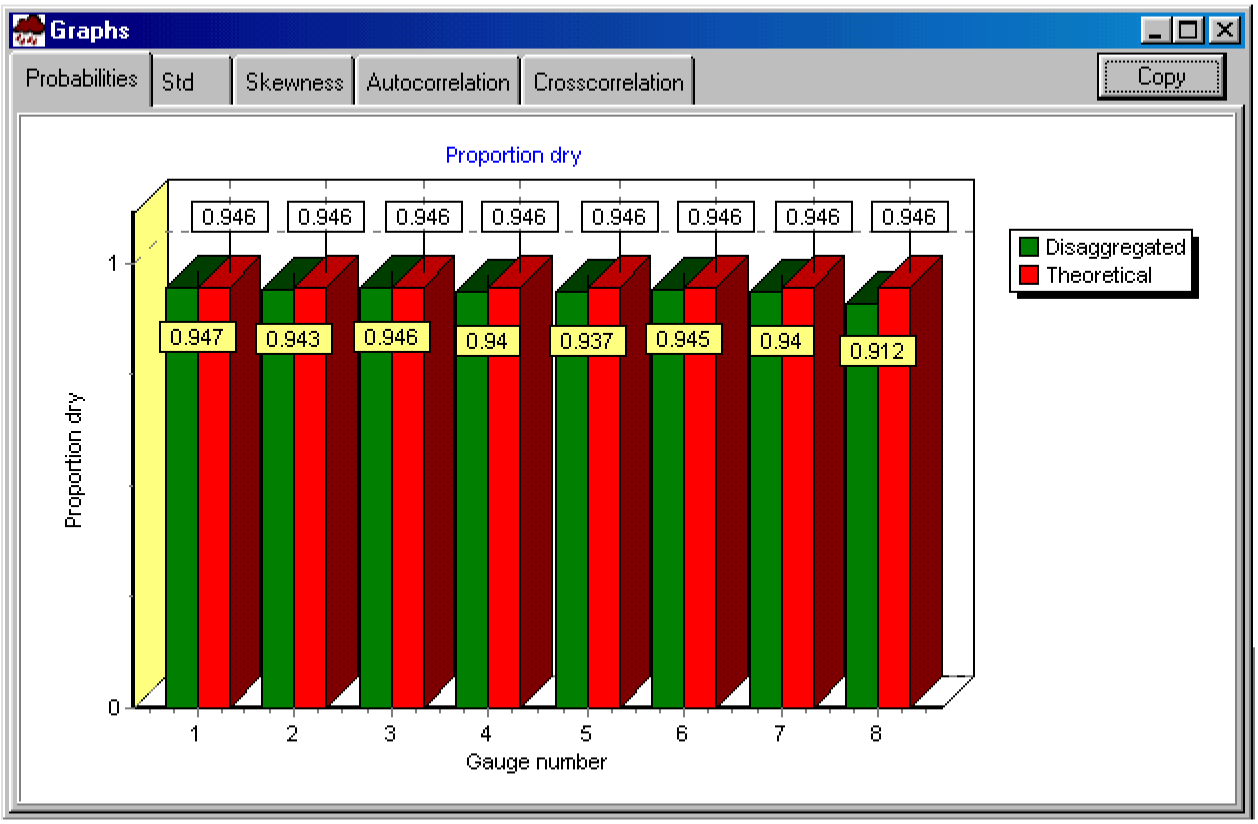
Graphs form#$K

Activate this form by pressing the appropriate button in the **Main form**.

Use this form, after performing the disaggregation, to visualize the graphical comparisons of historical and simulated statistics of hourly rainfall

To zoom in any of the graphs, drag on the region of interest downwards. To zoom out, drag on any region within the graph upwards. To move along the graph drag to the desired direction with the right mouse button pressed.

Using the Copy button, a graph is copied into the clipboard and can then be pasted to anywhere else (e.g. word processing programs etc.).



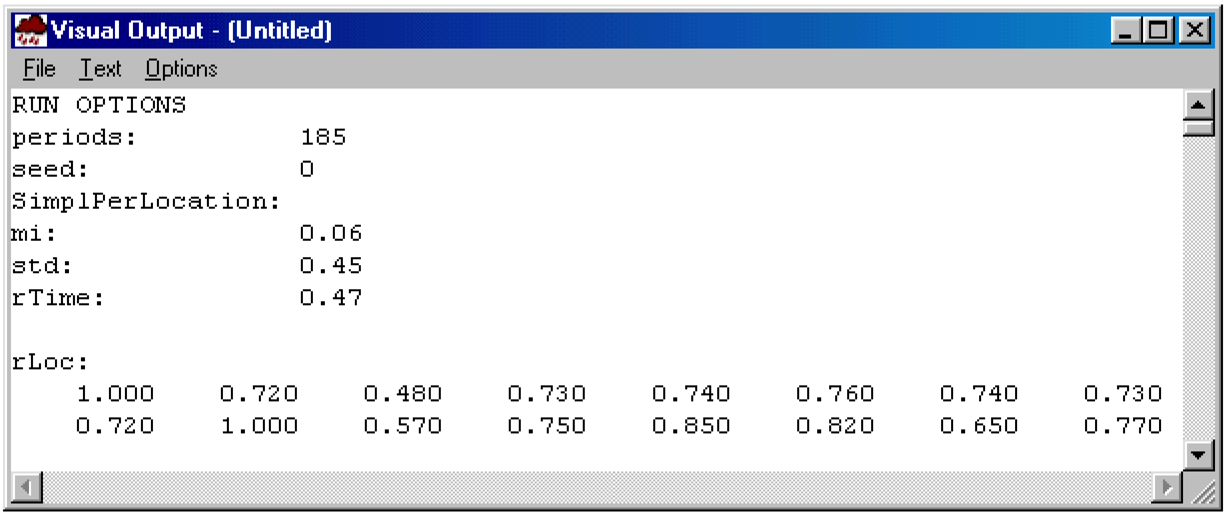
\# Graphs_form

$ Graphs form

K Graphs form; Form, graphs

Visual output form[#][$][K]

This form appears automatically when opening an information file (see **Input files format** The content of the form, **results of the disaggregation framework and printed hourly statistics** can be saved in a text file (use the file menu) or copied to the clipboard (press Ctrl-C).

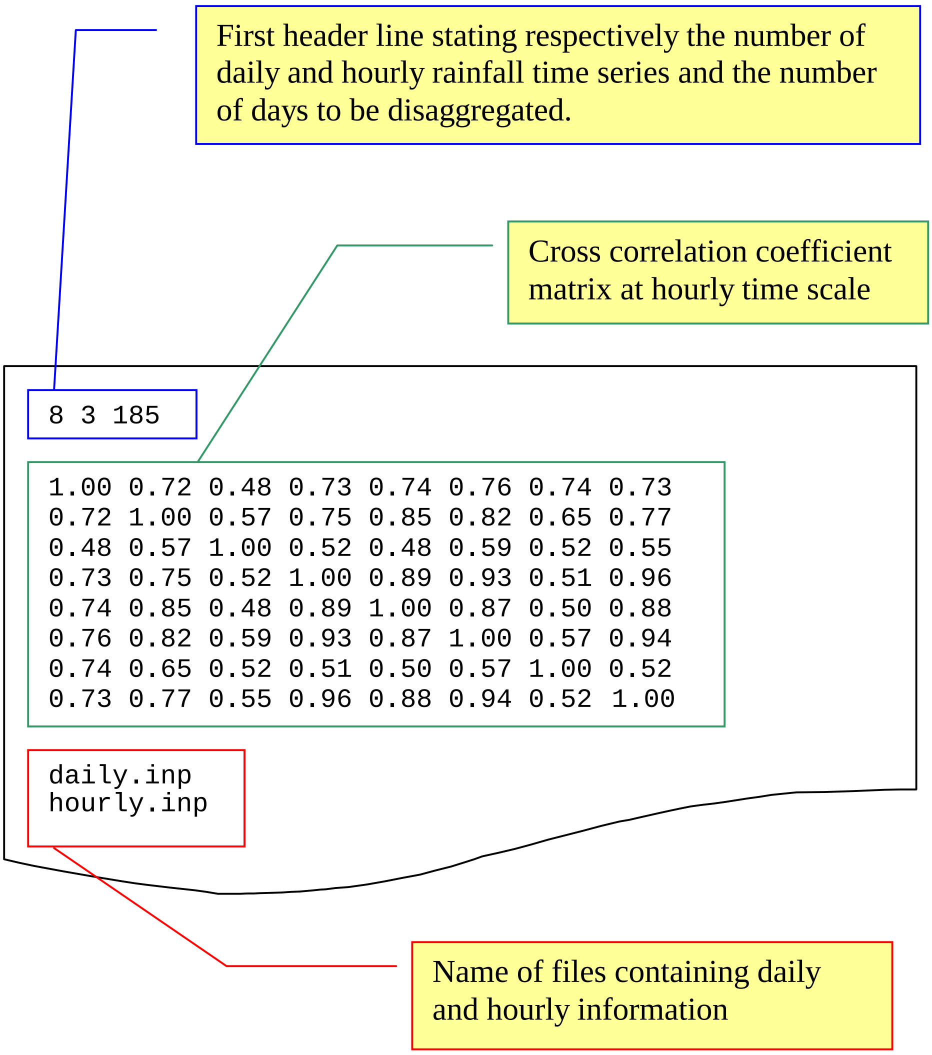


[#] Visual_output_form
[$] Visual output form
[K] Visual output form; Form, Visual output

Input files format[#][$][K]

File **input.dat :**

This is a text file that must be defined using the program **Main form** in order for the program to perform the disaggregation. The contents of the file are described below:



For the estimation of the unknown hourly crosscorrelation coefficients see related topic:

(**Estimation of crosscorrelation coefficients** .)

[#] Input_files_format
[$] Input files format
[K] input; rainfall depths; crosscorrelation coefficients matrix

File **daily.inp**: This is a text file containing the historical daily rainfall depths.In the current example we are considering 8 gages; historical hourly rainfall depths are available at gages 1,2,3 and historical daily rainfall depths are available for all gages.(The available hourly rainfall depths must be consistent with the daily rainfall depths of the same period).

Historical daily rainfall depths for raingages n, m…,

0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0.2	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0.2	0	0	0	0	0	0
0	0.2	0	0	0.2	0	0	0
0	0.2	0	0	0.4	0	0	0
0	0	0	0	0	0	0	0
0	0.2	0	0	0.2	0	0	0.2
0	0.2	0	0.2	0.2	0	0	0
0	0	0	0	0	0	0	0

Historical daily rainfall depths for raingages i , j, k, (in which historical hourly rainfall depths are also available from measurements)

File **hourly.inp**: : This is a text file containing the historical hourly rainfall depths available (3 in this example)

0	0	0
0	0	0
0	0	0
0	0	0
7.4	5.8	0
4.2	1.8	0
3.8	2.4	0
2	1.4	0
0.4	0.8	0
0.2	0	0
0	0	0

Output file format[#$K]

Press the appropriate button on the toolbar of the **Main form** to save the hourly (in case of disaggregation of daily to hourly) or daily time series .(in case of aggregation of hourly to daily).

The output file is a text file

0	0	0	0	0.1	0	0	0.1
0	0	0	0	0.1	0	0	0.1
0	0	0	0	0	0	0	0.1
0	0	0	1.4	0.8	0.8	3.3	1.1
0	0	0	0.7	0.4	0.4	1.5	0.6
0	0	0	0.4	0.2	0	0.7	0.3
0	0	0	0.3	0	0.1	0.3	0.2
0	0	0	0	0	0	0.1	0.1
0	0	0	0.2	0	0	0.6	0.1
7.4	5.8	0	7.1	6.9	5.6	4.9	5.1
4.2	1.8	0	2.5	2.4	2.3	2.7	2.3
3.8	2.4	0	2.6	2.7	2.4	2.5	2.5
2	1.4	0	1.5	1.6	1.3	1.3	1.4
0.4	0.8	0	0.6	0.7	0.5	0.3	0.6
0.2	0	0	0	0	0	0	0.1

[#] Output_file_format
[$] Output file format
[K] Output file format

REFERENCES

Chandler, R. and H. Wheater, Climate change detection using Generalized Linear Models for rainfall, A case study from the West of Ireland, I, Preliminary analysis and modelling of rainfall occurrence, Technical report, no. 194, Department of Statistical Science, University College London, 1998a (http://www.ucl.ac.uk/Stats/research/abstracts.html)

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

Grygier, J. C. 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.

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

Koutsoyiannis, D., A stochastic disaggregation method for design storm and flood synthesis, *J. of Hydrol.*, 156, 193-225, 1994.

Koutsoyiannis, D., Optimal decomposition of covariance matrices for multivariate stochastic models in hydrology, *Water Resour. Res.*, 35(4), 1219-1229, 1999.

Koutsoyiannis, D., Coupling stochastic models of different time scales, *Water Resour. Res.*, 37(2), 379-392, 2001.

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

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

Koutsoyiannis, D., 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. (Presentation also available on line at http://www.itia.ntua.gr/e/docinfo/59/)

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

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

Koutsoyiannis, D., and Th. Xanthopoulos, A dynamic model for short-scale rainfall disaggregation, *Hydrol. Sci. J.*, 35(3), 303-322, 1990.

Koutsoyiannis, D., Multivariate Rainfall Disaggregation at a fine time scale; 2002 (unpublished paper in review for *Water Resour. Res.*)

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

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

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

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

Northrop, P., Modelling and statistical analysis of spatial-temporal rainfall fields, PhD thesis, Department of Statistical Science, University College London, 1996.

Northrop, P. J., A clustered spatial-temporal model of rainfall, Proc. R. Soc. Lond. A454, 1875-1888, 1998.

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

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

Onof, C. 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.

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

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

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

Salas, J. D., Analysis and modeling of hydrologic time series, Chapter 19, *Handbook of hydrology*, edited by D. Maidment, McGraw-Hill, New York, 1993.

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

Santos, E. G., and J. D. Salas, Stepwise disaggregation scheme for synthetic hydrology, *Journal of Hydraulic Engineering*, 118(5), 765-784, 1992.

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

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

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

Todini, E., The preservation of skewness in linear disaggregation schemes, *J. Hydrol*., 47, 199-214, 1980.

Valencia, D. and J. C. Schaake, 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., 1972.

Valencia, D. and J. C. Schaake, Disaggregation processes in Stochastic Hydrology, *Water Resour. Res.*, 9(3) 211-219, 1973.

Wheater, H. S., V. S. Isham, C. Onof, R. E. Chandler, P. J. Northrop, P. Guiblin, S. M. Bate, D. Cox, and D. Koutsoyiannis, Generation of Spatially Consistent Rainfall Data, Volume I, Report to MAFF, Imperial College and University College London, 170 pp., London, 2000.

Acknowledgements

Vorrei ringraziare il mio professore Francesco Napolitano per la sua pazienza , la sua disponibilità e la fiducia riposta in me non solo durante il periodo della tesi ma anche negli ultimi tre anni di studi e per essere stato un esempio da seguire ma anche e soprattutto un amico.

Un ringraziamento speciale anche a Maria Chiara, Enrico, Alessandro, Evangelia e a tutti gli altri amici che mi sono stati vicini nei momenti più difficili (che non sono stati pochi ...) e che mi hanno sopportato con estrema pazienza.

E ancora un grazie alla mia famiglia che in questi anni mi e' stata vicina mi ha appoggiata ed ha creduto in me e nel mio sogno.

'Ενα πολύ μεγάλο ευχαριστώ στον καθηγητή Δημήτρη Κουτσογιάννη για τον χρόνο που μου αφιέρωσε, για την υποστήριξη, την εμπιστοσύνη, την υπομονή, για τα ενθαρρυντικά του λόγια ειδικά στις δύσκολες στιγμές και φυσικά για όλα όσα μου δίδαξε.

Και τέλος ευχαριστώ τον Αντρέα, το Βαγγέλη, τον Αντώνη, το Γιώργο, το Φοίβο και το Νίκο από τον τομέα που με έκαναν να νιώσω από την πρώτη μέρα στο ΕΜΠ σαν μέλος της ομάδας.