

Rainfall nowcasting by at site stochastic model P.R.A.I.S.E.

B. Sirangelo, P. Versace, and D. L. De Luca

Dipartimento di Difesa del Suolo, Università della Calabria – Rende, Italy

Received: 8 January 2007 – Published in Hydrol. Earth Syst. Sci. Discuss.: 29 January 2007

Revised: 18 April 2007 – Accepted: 28 April 2007 – Published: 15 May 2007

Abstract. The paper introduces a stochastic model to forecast rainfall heights at site: the P.R.A.I.S.E. model (Prediction of Rainfall Amount Inside Storm Events). PRAISE is based on the assumption that the rainfall height H_{i+1} accumulated on an interval Δt between the instants $i\Delta t$ and $(i+1)\Delta t$ is correlated with a variable $Z_i^{(v)}$, representing antecedent precipitation. The mathematical background is given by a joined probability density $f_{H_{i+1}, Z_i^{(v)}}(h_{i+1}, z_i^{(v)})$ in which the variables have a mixed nature, that is a finite probability in correspondence to the null value and infinitesimal probabilities in correspondence to the positive values. As study area, the Calabria region, in Southern Italy, was selected, to test performances of the PRAISE model.

1 Introduction

Rainfall is the main input for all hydrological models such as, for example, rainfall-runoff models and for forecasting landslides induced by precipitation. In the last few years catastrophic rainfall events have occurred in the Mediterranean area, leading to floods, flash floods and shallow landslides (debris and mud flows). Consequently there is the need for the implementation of forecasting systems able to predict meteorological conditions leading to disastrous occurrences. Nowadays, for achieving this goal, meteorological and stochastic models are used. The former (Chuang et al., 2000; Palmer et al., 2000; Untch et al., 2006) can be viewed as valid qualitative-quantitative rainfall forecasting tools at 24, 48 and 72 h (of course, at these forecasting horizons an absolute precision is not required, but rather an order of magnitude) when these phenomena occur on a considerable spatial scale. Nevertheless, they cannot yet be regarded as providers of quantitative rainfall forecasts in the

short term (6–12 h) to be used directly for forecasting systems, since the quantitative forecasting of precipitation, on the time and space scales of the hydrological phenomena, has not yet achieved the degree of precision necessary to avoid either the non-forecasting of exceptional small-scale situations or the issuing of unwarranted alarms. Consequently, in order to perform short term real-time rainfall forecasts for small basins (i.e. with size ranging 100–1000 km²), temporal stochastic models appear more competitive. Stochastic processes are widely used in hydrological variables forecasting (Waymire and Gupta, 1981; Georgakakos and Kavvas, 1987; Foufoula-Georgiou and Georgakakos, 1988) and, as regards at-site models, precipitation models can be categorized into two broad types: “discrete time-series models” and “point processes models”. Models of the former type, that include AutoRegressive Stochastic Models (Box and Jenkins, 1976; Salas et al., 1980; Brockwell and Davis, 1987; Burlando et al., 1993; Hipel and McLeod, 1994; Burlando et al., 1996; Toth et al., 2000), describe the rainfall process at discrete time steps, are not intermittent and therefore can be applied for describing the “within storm” rainfall. Models of the latter type (Lewis, 1964; Kavvas and Delleur, 1981; Smith and Karr, 1983; Rodriguez-Iturbe et al., 1984, 1987; Rodriguez-Iturbe, 1986; Cowpertwait et al., 1996; Sirangelo and Iiritano, 1997; Calenda and Napolitano, 1999; Montanari and Brath, 1999; Cowpertwait, 2004) are continuous time series models, are intermittent and therefore can simulate inter-storm periods also.

In the present work, a special kind of AutoRegressive model, named PRAISE (Prediction of Rainfall Amount Inside Storm Events), is described. It can be considered as a simple and useful tool for at site nowcasting precipitation, especially for applications in rainfall-runoff models, regarding small basins, and forecasting landslides, induced by rainfall, models. The paper is structured in three sections, excluding the introduction: in Sect. 2 the theoretical bases of the stochastic model are illustrated; while in Sect. 3 model

Correspondence to: D. L. De Luca
(davideluca@dds.unical.it)

calibration is shown; finally, Sect. 4 concerns the application of the model to the raingauge network of the Calabria region, in Southern Italy, with particular regard to the Cosenza raingauge.

2 The PRAISE model

In the PRAISE model, the rainfall heights h_n , cumulated over intervals $[(n-1)\Delta t, n\Delta t]$, are considered as a realisation of a weakly stationary stochastic process with discrete parameter $\{H_n; n \in I\}$, where I indicates the integer numbers. The H_n are non-negative random variables of mixed type, with finite probability on zero value and infinitesimal probability on positive values. In the following the instant $i\Delta t = t_0$ is assumed as current time, so that the observed rainfall heights are with subscripts less or equal to i and the probabilistic prediction will be referred to the rainfall heights with subscripts greater than i .

The main feature of the approach suggested in the PRAISE model is the identification of a random variable $Z_i^{(v)}$, a suitable function of the v random variables $H_i, H_{i-1}, \dots, H_{i-v+1}$, such that its stochastic dependence with the random variable H_{i+1} describes the whole correlative structure of the process $\{H_n; n \in I\}$. Two steps are required to identify $Z_i^{(v)}$:

- a) the individuation of the process “memory” extension v ;
- b) the optimal choice of the function depending on the random variables $H_i, H_{i-1}, \dots, H_{i-v+1}$ defining $Z_i^{(v)}$.

After this, the PRAISE model provides the identification of the joint probability density $f_{H_{i+1}, Z_i^{(v)}}(h_{i+1}, z_i^{(v)})$ and its utilisation for the real time forecasting of rainfall heights during a storm event.

2.1 Extension of the “memory”

In the PRAISE model, the linear stochastic dependence between the random variable H_{i+1} and the generic antecedent random variables $H_{i-j}, j=0, 1, 2, \dots$, is considered negligible when the correspondent sample coefficient of partial autocorrelation results less than a fixed value close to zero. The results given by this approach are similar to the more rigorous and less simple general method, in which the absence of a linear stochastic dependence should be tested verifying that the sample coefficient of partial autocorrelation, between such random variables, exhibits a value inside a confidence interval, which does not permit the rejection of the null value hypothesis for the correspondent theoretical quantity.

In order to determine the extension of the “memory” for the process $\{H_n; n \in I\}$, the absence of significant partial autocorrelation must be checked for increasing values of v . For each value of v must be verified the negligibility of all the sample coefficients of partial autocorrelation $\rho_{H_{i+1} H_{i-v+1-m} \cdot H_i \dots H_{i-v+1}}$ characterised by a couple of pri-

mary subscripts H_{i+1}, H_{i-v} ; H_{i+1}, H_{i-v-1} ; ... and by secondary subscripts $H_i, H_{i-1}, \dots, H_{i-v+1}$.

Practically, evaluation of $\rho_{H_{i+1} H_{i-v+1-m} \cdot H_i \dots H_{i-v+1}}$ must be performed using the sample coefficients of partial autocorrelation $r_{H_{i+1} H_{i-v+1-m} \cdot H_i \dots H_{i-v+1}}, m=1, 2, \dots$, obtained by an observed sample h_1, h_2, \dots, h_N of rainfall heights cumulated over time intervals of duration Δt , after estimation of autocorrelation ρ_k by sample autocorrelation coefficients r_k (Kendall and Stuart, 1969); for the hypothesis of weakly stationary process, autocorrelation depends only on lag k .

Finally, to estimate the extension of the “memory”, a simple way can be obtained introducing the sample maximum absolute scattering:

$$\chi_r(v) = \max_{1 \leq m < \infty} |r_{H_{i+1} H_{i-v+1-m} \cdot H_i \dots H_{i-v+1}}| \quad m=1, 2, \dots \quad (1)$$

The extension of the process “memory” can be assumed equal to the minimum value of v for which $\chi_r(v)$ results less than a fixed critical value $\chi_{r,cr}$.

2.2 Structure of the random variable $Z_i^{(v)}$

The criterion adopted in the PRAISE model to define a functional dependence between the random variable $Z_i^{(v)}$ and the random variables $H_i, H_{i-1}, \dots, H_{i-v+1}$ is the maximisation of the coefficient of linear correlation $\rho_{H_{i+1}, Z_i^{(v)}}$ between the same $Z_i^{(v)}$ and the random variable H_{i+1} . This choice allows, once identified the joint probability density $f_{H_{i+1}, Z_i^{(v)}}(h_{i+1}, z_i^{(v)})$, the best prediction of rainfall heights H_{i+1} during a storm event. If a functional relationship $Z_i^{(v)} = f_{01}(H_i, H_{i-1}, \dots, H_{i-v+1})$ satisfies the criterion here adopted, then all the functions:

$$f_{ab}(H_i, H_{i-1}, \dots, H_{i-v+1}) = a + b f_{01}(H_i, H_{i-1}, \dots, H_{i-v+1}) \quad (2)$$

with a and b generic constants, satisfy the same criterion. In other words, the criterion of maximisation of the coefficient of linear correlation between $Z_i^{(v)}$ and H_{i+1} identifies a “class of functions”, defined by Eq. (2), with all the members equivalent for the purposes of the suggested model.

As regards the analytical form of the relationship linking $Z_i^{(v)}$ to $H_i, H_{i-1}, \dots, H_{i-v+1}$, the more suitable choice is, clearly, the linear function:

$$Z_i^{(v)} = \beta + \sum_{j=0}^{v-1} \alpha'_j H_{i-j} \quad (3)$$

In this condition, in fact, the coefficients of the linear relationship that maximises $\rho_{H_{i+1}, Z_i^{(v)}}$ are simply the coefficients of the linear partial regression among H_{i+1} and $H_i, H_{i-1}, \dots, H_{i-v+1}$:

$$E(H_{i+1} | H'_i, H'_{i-1}, \dots, H'_{i-v+1}) = \mu_H + \sum_{j=0}^{v-1} \alpha'_j H'_{i-j} \quad (4)$$

where $E(\cdot)$ is the expected value operator, $\mu_H = E(H_n)$, constant value for every n because of the hypothesis of weakly stationary stochastic process, and $H'_i = H_i - \mu_H$, $H'_{i-1} = H_{i-1} - \mu_H$, $H'_{i-v+1} = H_{i-v+1} - \mu_H$.

The α'_j coefficients are evaluated as:

$$\alpha'_j = -\tilde{\lambda}_{1(j+2)}^{(v+1)} / \tilde{\lambda}_{11}^{(v+1)} \tag{5}$$

where $\tilde{\lambda}_{1(j+2)}^{(v+1)}$ e $\tilde{\lambda}_{11}^{(v+1)}$ are the cofactors of the Laurent matrix:

$$\left[\Lambda^{(v+1)} \right] = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_v \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{v-1} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{v-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho_v & \rho_{v-1} & \rho_{v-2} & \dots & 1 \end{bmatrix} \tag{6}$$

These coefficients can be readily evaluated utilising again an observed sample h_1, h_2, \dots, h_N of rainfall heights cumulated over time intervals of duration Δt . However, it should be highlighted that, as a consequence of the non-negative character of the involved random variables, the coefficients α'_j must satisfy the conditions $\alpha'_j \geq 0$ for $j=0, 1, \dots, v-1$.

Moreover, the coefficients α'_j identify, among all the possible linear combinations of $H_i, H_{i-1}, \dots, H_{i-v+1}$ variables, the one which maximizes the linear correlation coefficient with H_{i+1} .

Then, definition of $Z_i^{(v)}$ variable is very simple: let $Z_{*i}^{(v)}$ be a variable defined as:

$$Z_{*i}^{(v)} = \mu_H + \sum_{j=0}^{v-1} \alpha'_j H'_{i-j} \tag{7a}$$

it can be noted that it is equal to the conditional expected value $E(H_{i+1} | H'_i, H'_{i-1}, \dots, H'_{i-v+1})$; rewriting Eq. (7a) utilising $H_i, H_{i-1}, \dots, H_{i-v+1}$:

$$Z_{*i}^{(v)} = \sum_{j=0}^{v-1} \alpha'_j H_{i-j} + \left(1 - \sum_{j=0}^{v-1} \alpha'_j \right) \mu_H \tag{7b}$$

and introducing the standardised coefficients:

$$\alpha_j = \alpha'_j / \sum_{\kappa=0}^{v-1} \alpha'_\kappa \tag{8}$$

for which the conditions $0 < \alpha_j \leq 1$, for $j=0, 1, \dots, v-1$, and $\sum_{j=0}^{v-1} \alpha_j = 1$ are respected, $Z_{*i}^{(v)}$ can be expressed as:

$$Z_{*i}^{(v)} = \sum_{k=0}^{v-1} \alpha'_k \sum_{j=0}^{v-1} \alpha_j H_{i-j} + \left(1 - \sum_{j=0}^{v-1} \alpha'_j \right) \mu_H \tag{9}$$

Defining $Z_i^{(v)}$ as:

$$Z_i^{(v)} = \sum_{j=0}^{v-1} \alpha_j H_{i-j} \tag{10}$$

it is simple to verify that $Z_{*i}^{(v)}$ and $Z_i^{(v)}$ variables are connected by a linear transformation, identifying a class of function, defined by Eq. (2). Consequently the correlation coefficient $\rho_{H_{i+1}, Z_i^{(v)}}$ is equal to $\rho_{H_{i+1}, Z_{*i}^{(v)}}$.

Equation (10) shows that the random variable $Z_i^{(v)}$ can be regarded as a weighted average of the v antecedent rainfall heights with weights expressed by the coefficients α_j .

2.3 Joint probability density $f_{H_{i+1}, Z_i^{(v)}}(h_{i+1}, z_i^{(v)})$

Keeping in mind the character of random variable of mixed type for H_{i+1} and, therefore, for $Z_i^{(v)}$, the joint probability density $f_{H_{i+1}, Z_i^{(v)}}(h_{i+1}, z_i^{(v)})$ must be written in the form:

$$\begin{aligned} f_{H_{i+1}, Z_i^{(v)}}(h_{i+1}, z_i^{(v)}) &= p_{\cdot, \cdot} \delta(h_{i+1}) \delta(z_i^{(v)}) \\ &+ p_{+, \cdot} f_{H_{i+1}, 0}^{(+, \cdot)}(h_{i+1}) \delta(z_i^{(v)}) \\ &+ p_{\cdot, +} f_{0, Z_i^{(v)}}^{(\cdot, +)}(z_i^{(v)}) \delta(h_{i+1}) \\ &+ p_{+, +} f_{H_{i+1}, Z_i^{(v)}}^{(+, +)}(h_{i+1}, z_i^{(v)}) \end{aligned} \tag{11}$$

where $\delta(\cdot)$ is Dirac's delta function, $p_{\cdot, \cdot}$, $p_{+, \cdot}$, $p_{\cdot, +}$ and $p_{+, +}$ indicate, respectively, $\Pr[H_{i+1}=0 \cap Z_i^{(v)}=0]$, $\Pr[H_{i+1}>0 \cap Z_i^{(v)}=0]$, $\Pr[H_{i+1}=0 \cap Z_i^{(v)}>0]$ and $\Pr[H_{i+1}>0 \cap Z_i^{(v)}>0]$, with, obviously, $p_{\cdot, \cdot} + p_{+, \cdot} + p_{\cdot, +} + p_{+, +} = 1$, and:

$$\begin{aligned} f_{H_{i+1}, Z_i^{(v)}}^{(+, +)}(h_{i+1}, z_i^{(v)}) \cdot dh_{i+1} dz_i^{(v)} &= \\ \Pr [h_{i+1} \leq H_{i+1} < h_{i+1} + dh_{i+1} \cap \\ z_i^{(v)} \leq Z_i^{(v)} < z_i^{(v)} + dz_i^{(v)} | H_{i+1} > 0 \cap Z_i^{(v)} > 0] \end{aligned} \tag{12}$$

$$\begin{aligned} f_{H_{i+1}, 0}^{(+, \cdot)}(h_{i+1}) \cdot dh_{i+1} &= \\ \Pr [h_{i+1} \leq H_{i+1} < h_{i+1} + dh_{i+1} | H_{i+1} > 0 \cap Z_i^{(v)} = 0] \end{aligned} \tag{13}$$

$$\begin{aligned} f_{0, Z_i^{(v)}}^{(\cdot, +)}(z_i^{(v)}) \cdot dz_i^{(v)} &= \\ \Pr [z_i^{(v)} \leq Z_i^{(v)} < z_i^{(v)} + dz_i^{(v)} | H_{i+1} = 0 \cap Z_i^{(v)} > 0] \end{aligned} \tag{14}$$

A suitable mathematical form for $f_{H_{i+1}, Z_i^{(v)}}^{(+, +)}(h_{i+1}, z_i^{(v)})$ may be achieved considering the standardised bivariate probability density of Moran and Downton (Kotz et al., 2000):

$$\begin{aligned} f_{X,Y}(x, y) &= \theta \exp[-\theta(x+y)] I_0 \left[2\sqrt{\theta(\theta-1)xy} \right] \\ x > 0, y > 0; \theta &\geq 1 \end{aligned} \tag{15}$$

where $I_0(u)$ is the modified Bessel function I of zero order (Abramowitz and Stegun, 1970).

An important peculiarity of this distribution is the capacity to reproduce any positive correlative structure between X and Y variables; they will be independent for $\theta=1$ and connected in deterministic way for $\theta \rightarrow +\infty$.

Applying a double power transformation:

$$x = \alpha_{h,z} (h_{i+1})^{\beta_{h,z}} \quad h_{i+1} > 0; \alpha_{h,z} > 0, \beta_{h,z} > 0 \quad (16)$$

$$y = \gamma_{h,z} (z_i^{(v)})^{\delta_{h,z}} \quad z_i^{(v)} > 0; \gamma_{h,z} > 0, \delta_{h,z} > 0 \quad (17)$$

and rewriting θ as $\theta_{h,z}$, the bivariate probability density so obtained, that can be called Weibull-Bessel, is given by:

$$f_{H_{i+1}, Z_i^{(v)}}^{(+,+)}(h_{i+1}, z_i^{(v)}) = \theta_{h,z} \alpha_{h,z} \beta_{h,z} (h_{i+1})^{\beta_{h,z}-1} \gamma_{h,z} \delta_{h,z} (z_i^{(v)})^{\delta_{h,z}-1} \times \exp \left\{ -\theta_{h,z} \left[\alpha_{h,z} (h_{i+1})^{\beta_{h,z}} + \gamma_{h,z} (z_i^{(v)})^{\delta_{h,z}} \right] \right\} \times I_0 \left[2\sqrt{\theta_{h,z} (\theta_{h,z} - 1) \alpha_{h,z} (h_{i+1})^{\beta_{h,z}} \gamma_{h,z} (z_i^{(v)})^{\delta_{h,z}}} \right] \quad (18)$$

where $h_{i+1} > 0, z_i^{(v)} > 0; \alpha_{h,z} > 0, \beta_{h,z} > 0, \gamma_{h,z} > 0, \delta_{h,z} > 0, \theta_{h,z} \geq 1$. As can be simply verified, the marginal probability densities of the Eq. (18) are Weibull densities with parameters $\alpha_{h,z}, \beta_{h,z}$ and $\gamma_{h,z}, \delta_{h,z}$. Consequently, moments of the distribution are immediate, as regards means and variances:

$$\mu_{H_{i+1}}^{(+,+)} = E(H_{i+1} | H_{i+1} > 0 \cap Z_i^{(v)} > 0) = \frac{\Gamma(1 + 1/\beta_{h,z})}{\alpha_{h,z}^{1/\beta_{h,z}}} \quad (19)$$

$$(\sigma_{H_{i+1}}^2)^{(+,+)} = var(H_{i+1} | H_{i+1} > 0 \cap Z_i^{(v)} > 0) = \frac{1}{\alpha_{h,z}^{2/\beta_{h,z}}} \left[\Gamma(1 + 2/\beta_{h,z}) - \Gamma^2(1 + 1/\beta_{h,z}) \right] \quad (20)$$

$$\mu_{Z_i^{(v)}}^{(+,+)} = E(Z_i^{(v)} | H_{i+1} > 0 \cap Z_i^{(v)} > 0) = \frac{\Gamma(1 + 1/\delta_{h,z})}{\gamma_{h,z}^{1/\delta_{h,z}}} \quad (21)$$

$$(\sigma_{Z_i^{(v)}}^2)^{(+,+)} = var(Z_i^{(v)} | H_{i+1} > 0 \cap Z_i^{(v)} > 0) = \frac{1}{\gamma_{h,z}^{2/\delta_{h,z}}} \left[\Gamma(1 + 2/\delta_{h,z}) - \Gamma^2(1 + 1/\delta_{h,z}) \right] \quad (22)$$

Moreover, it can be proved that the coefficient of linear correlation between the random variables H_{i+1} and $Z_i^{(v)}$, distributed according to the Eq. (18), is:

$$\rho_{H_{i+1}, Z_i^{(v)}}^{(+,+)} = \frac{{}_2F_1(-1/\beta_{h,z}, -1/\delta_{h,z}; 1; 1 - 1/\theta_{h,z}) - 1}{\sqrt{\left[\frac{\Gamma(1 + 2/\beta_{h,z})}{\Gamma^2(1 + 1/\beta_{h,z})} - 1 \right] \cdot \left[\frac{\Gamma(1 + 2/\delta_{h,z})}{\Gamma^2(1 + 1/\delta_{h,z})} - 1 \right]}} \quad (23)$$

where $\Gamma(u)$ is the complete gamma function and ${}_2F_1(a, b; c; u)$ is the hypergeometric function (Abramowitz and Stegun, 1970). It must be pointed out how higher values of $\theta_{h,z}$ parameter give higher values of $\rho_{H_{i+1}, Z_i^{(v)}}^{(+,+)}$.

As regards the couple of probability densities $f_{H_{i+1}, 0}^{(+, \cdot)}(h_{i+1})$ and $f_{0, Z_i^{(v)}}^{(\cdot, +)}(z_i^{(v)})$, the PRAISE model, in the form here employed, assumes again Weibull densities with parameters, respectively, $\alpha_{h,\cdot}, \beta_{h,\cdot}$ and $\gamma_{\cdot,z}, \delta_{\cdot,z}$:

$$f_{H_{i+1}, 0}^{(+, \cdot)}(h_{i+1}) = \alpha_{h,\cdot} \beta_{h,\cdot} (h_{i+1})^{\beta_{h,\cdot}-1} \exp[-\alpha_{h,\cdot} (h_{i+1})^{\beta_{h,\cdot}}] \quad h_{i+1} > 0; \alpha_{h,\cdot} > 0, \beta_{h,\cdot} > 0 \quad (24)$$

$$f_{0, Z_i^{(v)}}^{(\cdot, +)}(z_i^{(v)}) = \gamma_{\cdot,z} \delta_{\cdot,z} (z_i^{(v)})^{\delta_{\cdot,z}-1} \exp[-\gamma_{\cdot,z} (z_i^{(v)})^{\delta_{\cdot,z}}] \quad z_i^{(v)} > 0; \gamma_{\cdot,z} > 0, \delta_{\cdot,z} > 0 \quad (25)$$

3 Parameter estimation

The parameter estimation for the bivariate probability density $f_{H_{i+1}, Z_i^{(v)}}(h_{i+1}, z_i^{(v)})$ can be made on the basis of a sample h_1, h_2, \dots, h_N of rainfall heights cumulated over time intervals of duration Δt . First of all, starting from the sample h_1, h_2, \dots, h_N and using the Eq. (10), a “sample” of $Z^{(v)}$ can be calculated, namely $z_1^{(v)}, z_2^{(v)}, \dots, z_N^{(v)}$. The probabilities $p_{\cdot,\cdot}, p_{+,\cdot}, p_{\cdot,+}$ and $p_{+,+}$, could be then estimated (according to the maximum likelihood method) by the frequencies $N^{(\cdot,\cdot)}/N, N^{(+,\cdot)}/N, N^{(\cdot,+)} / N$ and $N^{(+,+)} / N$ of the events $H_{i+1}=0 \cap Z_i^{(v)}=0, H_{i+1}>0 \cap Z_i^{(v)}=0, H_{i+1}=0 \cap Z_i^{(v)}>0$ and $H_{i+1}>0 \cap Z_i^{(v)}>0$ evaluated on the basis of the bivariate sample $h_1, z_1^{(v)}; h_2, z_2^{(v)}; \dots; h_N, z_N^{(v)}$. According to the method of moments, the parameters $\alpha_{h,\cdot}, \beta_{h,\cdot}$ and $\gamma_{\cdot,z}, \delta_{\cdot,z}$ could be estimated fitting means and standard deviations $m_{H_{i+1}}^{(+,\cdot)}, s_{H_{i+1}}^{(+,\cdot)}$ and $m_{Z_i^{(v)}}^{(\cdot,+)}, s_{Z_i^{(v)}}^{(\cdot,+)}$ of the bivariate sample $h_1, z_1^{(v)}; h_2, z_2^{(v)}; \dots; h_N, z_N^{(v)}$ restricted, respectively, by the conditions, $h_j > 0, z_j^{(v)} = 0$ and $h_j = 0, z_j^{(v)} > 0$:

$$\hat{\beta}_{h,\cdot} = \beta_{h,\cdot} : \quad \Gamma(1 + 2/\beta_{h,\cdot}) / \Gamma^2(1 + 1/\beta_{h,\cdot}) = 1 + (s_{H_{i+1}}^2)^{(+,\cdot)} / (m_{H_{i+1}}^{(+,\cdot)})^2 \quad (26)$$

$$\hat{\alpha}_{h,\cdot} = \left[\Gamma(1 + 1/\hat{\beta}_{h,\cdot}) / m_{H_{i+1}}^{(+,\cdot)} \right] \hat{\beta}_{h,\cdot} \quad (27)$$

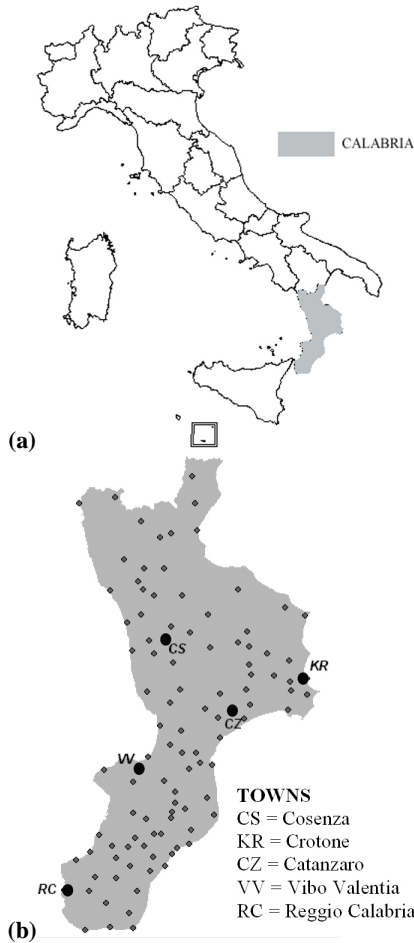


Fig. 1. (a) Location of the Calabria region, in the Southern Italy; (b) Telemetering Raingauge Network.

$$\hat{\delta}_{\cdot,z} = \delta_{\cdot,z} : \quad \Gamma(1 + 2/\delta_{\cdot,z}) / \Gamma^2(1 + 1/\delta_{\cdot,z}) = 1 + \left(s_{Z_i^{(v)}}^2 \right)^{(+,+)} / \left(m_{Z_i^{(v)}}^{(+,+)} \right)^2 \quad (28)$$

$$\hat{\gamma}_{\cdot,z} = \left[\Gamma(1 + 1/\hat{\delta}_{\cdot,z}) / m_{Z_i^{(v)}}^{(+,+)} \right]^{\hat{\delta}_{\cdot,z}} \quad (29)$$

Similarly, the parameters of the Weibull-Bessel density, $\alpha_{h,z}$, $\beta_{h,z}$, $\gamma_{h,z}$, $\delta_{h,z}$ and $\theta_{h,z}$ could be estimated fitting the marginal means and standard deviations $m_{H_{i+1}}^{(+,+)}$, $s_{H_{i+1}}^{(+,+)}$ and $m_{Z_i^{(v)}}^{(+,+)}$, $s_{Z_i^{(v)}}^{(+,+)}$, and the coefficient of linear correlation $r_{H_{i+1}, Z_i^{(v)}}^{(+,+)}$ of the bivariate sample $h_1, z_1^{(v)}; h_2, z_2^{(v)}; \dots; h_N, z_N^{(v)}$ restricted by the condition $h_j > 0, z_j^{(v)} > 0$:

$$\hat{\beta}_{h,z} = \beta_{h,z} : \quad \Gamma(1 + 2/\beta_{h,z}) / \Gamma^2(1 + 1/\beta_{h,z})$$

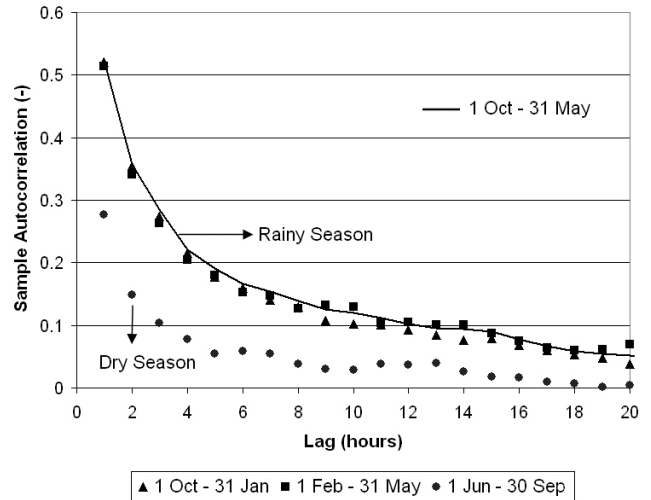


Fig. 2. Sample autocorrelograms referred to the Cosenza Rain-gauge.

$$= 1 + \left(s_{H_{i+1}}^2 \right)^{(+,+)} / \left(m_{H_{i+1}}^{(+,+)} \right)^2 \quad (30)$$

$$\hat{\alpha}_{h,z} = \left[\Gamma(1 + 1/\hat{\beta}_{h,z}) / m_{H_{i+1}}^{(+,+)} \right]^{\hat{\beta}_{h,z}} \quad (31)$$

$$\hat{\delta}_{h,z} = \delta_{h,z} : \quad \Gamma(1 + 2/\delta_{h,z}) / \Gamma^2(1 + 1/\delta_{h,z}) = 1 + \left(s_{Z_i^{(v)}}^2 \right)^{(+,+)} / \left(m_{Z_i^{(v)}}^{(+,+)} \right)^2 \quad (32)$$

$$\hat{\gamma}_{h,z} = \left[\Gamma(1 + 1/\hat{\delta}_{h,z}) / m_{Z_i^{(v)}}^{(+,+)} \right]^{\hat{\delta}_{h,z}} \quad (33)$$

$$\hat{\theta}_{h,z} = \theta_{h,z} : \quad {}_2F_1 \left(-\frac{1}{\hat{\beta}_{h,z}}, -\frac{1}{\hat{\delta}_{h,z}}; 1; 1 - \frac{1}{\theta_{h,z}} \right) = 1 + r_{H_{i+1}, Z_i^{(v)}}^{(+,+)} \frac{s_{H_{i+1}}^{(+,+)} s_{Z_i^{(v)}}^{(+,)}}{m_{H_{i+1}}^{(+,+)} m_{Z_i^{(v)}}^{(+,)}} \quad (34)$$

Equations (26), (28), (30), (32), (34) must be solved by numerical method (Press et al., 1988).

4 Application

Model calibration was performed using the hourly rain heights database of the tele-metering raingauge network of the “Dipartimento per la Protezione Civile – Centro Funzionale MeteIdrologico della Regione Calabria”. The raingauge network, located in the Calabria region, Southern Italy, is made up of 104 stations (Fig. 1). Approximately, 14 million hourly rainfalls form the database, of which about 7%

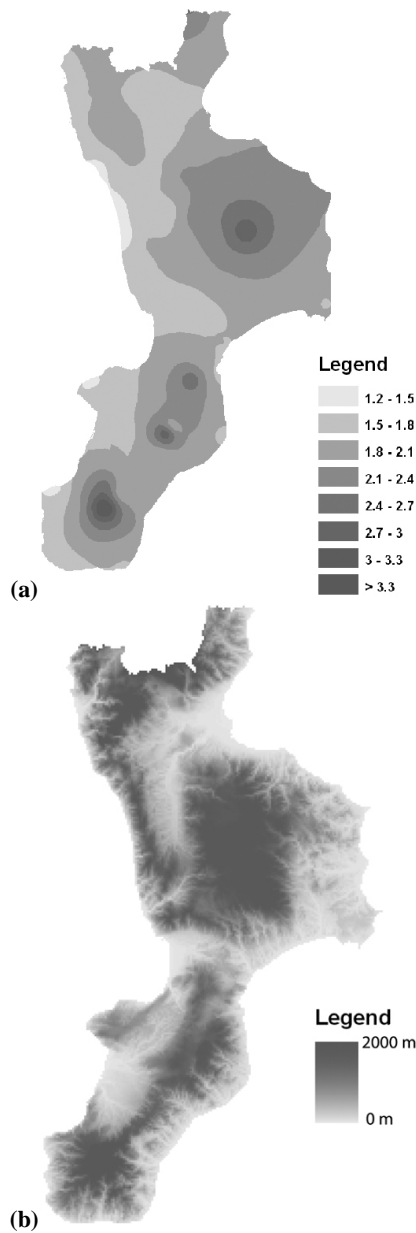


Fig. 3. (a) Mapping of $\theta_{h,z}$ parameter; (b) Digital Elevation Model of the Calabria region.

are rainy. In order to respect the hypothesis of stationary process, only the data measured during the “rainy season”, 1 October–31 May has been used (De Luca, 2005). In this period, correlation structure, mean and variance of the sample appear significantly homogeneous. In particular, considering hourly rain measurements of Cosenza raingauge, covering a period of about 15 years, and subdividing the year into three subperiods of four months (1 October–31 January; 1 February–31 May; 1 June–30 September), sample means and variances of the variable H are shown in Table 1, and

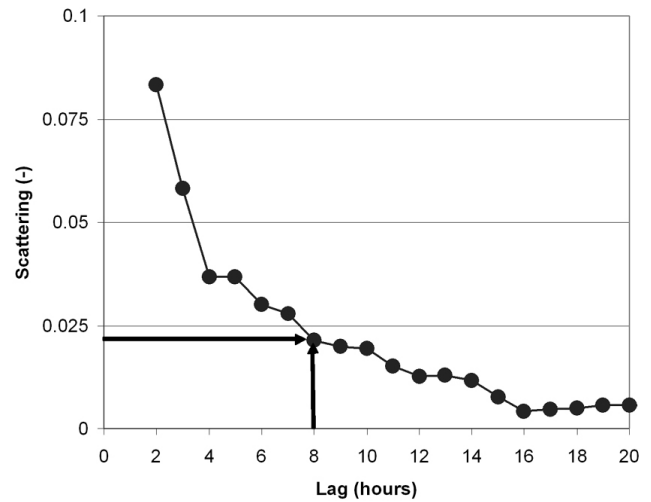


Fig. 4. Memory extension (rainy season) referred to the Cosenza Raingauge.

Table 1. Cosenza raingauge: sample means and variances for three subperiods of the year.

Season	SubPeriod	sample mean (mm)	sample variance (mm ²)
Rainy season	1 Oct–31 Jan	0.14	0.091
	1 Feb–31 May	0.12	0.084
Dry season	1 June–30 Sep	0.04	0.026

sample autocorrelograms are reported in Fig. 2. It appears plausible to consider as stationary the period comprising the first two subperiods. It must be highlighted that, because of the sample size, stationarity analysis based on a further subdivision into more subperiods appears unsuitable because of the greater uncertainty of moments evaluation.

The model parameters have been estimated for every raingauge. Subsequently, each parameter was mapped on the spatial regional domain by using a spline technique. The extension of the “memory”, determined by the technique described in the Sect. 2.1 fixing $\chi_{r,cr}=0.025$, has been found equal to $\hat{\nu}=8$ for all the tele-metering raingauges. An example of parameter mapping, referred to $\theta_{h,z}$ is represented in Fig. 3a, that shows greater values, and consequently higher values of $\rho_{H_{i+1}, Z_i^{(v)}}^{(+,+)}$ correlation, located in the part of region characterized by greater altitude, as depicted by the comparison with the Digital Elevation Model of the Calabria region (Fig. 3b).

In Fig. 4 memory extension, referred to the rainy season of the Cosenza raingauge, is depicted.

Table 2. Cosenza raingauge: estimated values of the α_j coefficients.

lag k	1	2	3	4	5	6	7	8
$\hat{\alpha}_j, \quad j=k-1$	0.717	0.092	0.056	0.040	0.031	0.025	0.021	0.018

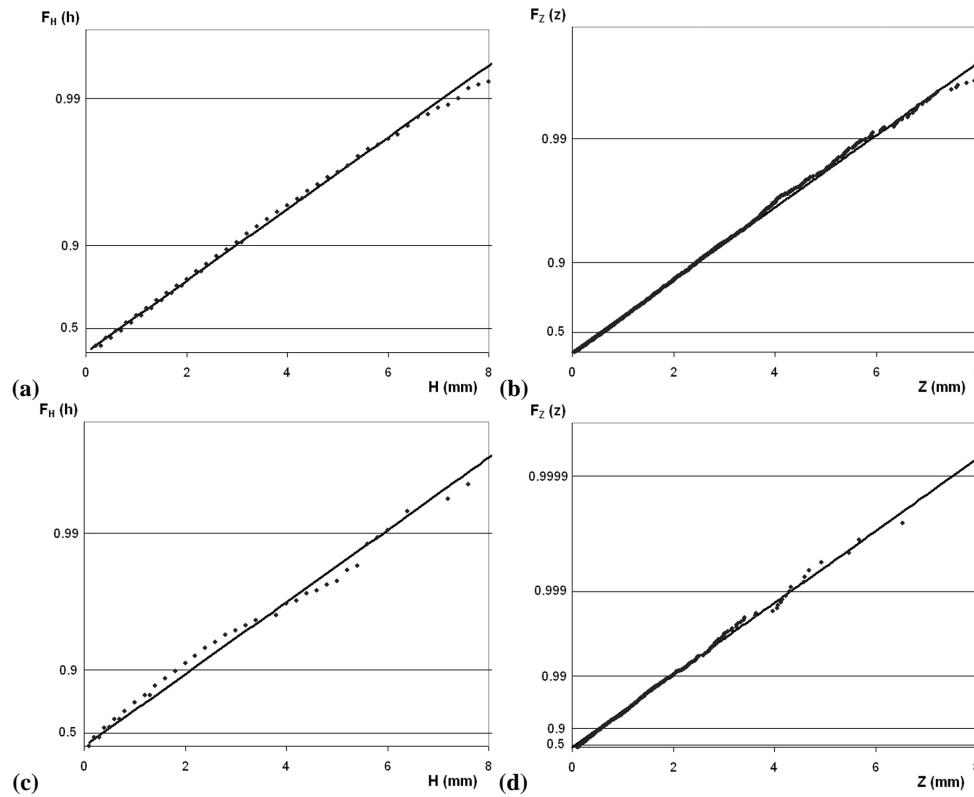


Fig. 5. Cosenza raingauge: plots on Weibull probabilistic papers of empirical and Weibull distributions for the variables (a) $H_{i+1} | H_{i+1} > 0 \cap Z_i^{(v)} > 0$; (b) $Z_i^{(v)} | H_{i+1} > 0 \cap Z_i^{(v)} > 0$; (c) $H_{i+1} | H_{i+1} > 0 \cap Z_i^{(v)} = 0$; (d) $Z_i^{(v)} | H_{i+1} = 0 \cap Z_i^{(v)} > 0$.

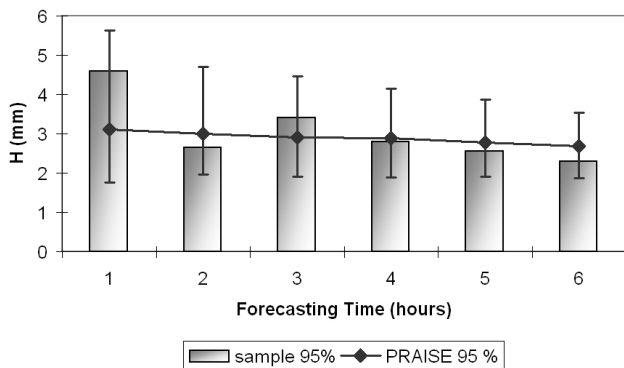


Fig. 6. Validation results for all tested events.

Table 3. Cosenza raingauge: estimated values of $p_{\cdot,\cdot}$, $p_{+,\cdot}$, $p_{\cdot,+}$ and $p_{+,+}$.

$\hat{p}_{\cdot,\cdot}$	$\hat{p}_{+,\cdot}$	$\hat{p}_{\cdot,+}$	$\hat{p}_{+,+}$
0.76	0.01	0.14	0.09

The estimated coefficients $\hat{\alpha}_j, \quad j=1, 2, \dots, 8$, given by Eqs. (5–8) once calculated the sample coefficients of the linear partial regression between H_{i+1} and $H_i, H_{i-1}, \dots, H_{i-v+1}$, are listed in Table 2. Finally, all the parameters of the bivariate probability density $f_{H_{i+1}, Z_i^{(v)}}(h_{i+1}, z_i^{(v)})$, estimated according to the methods described in Sect. 3, are reported in Tables 3 and 4. In

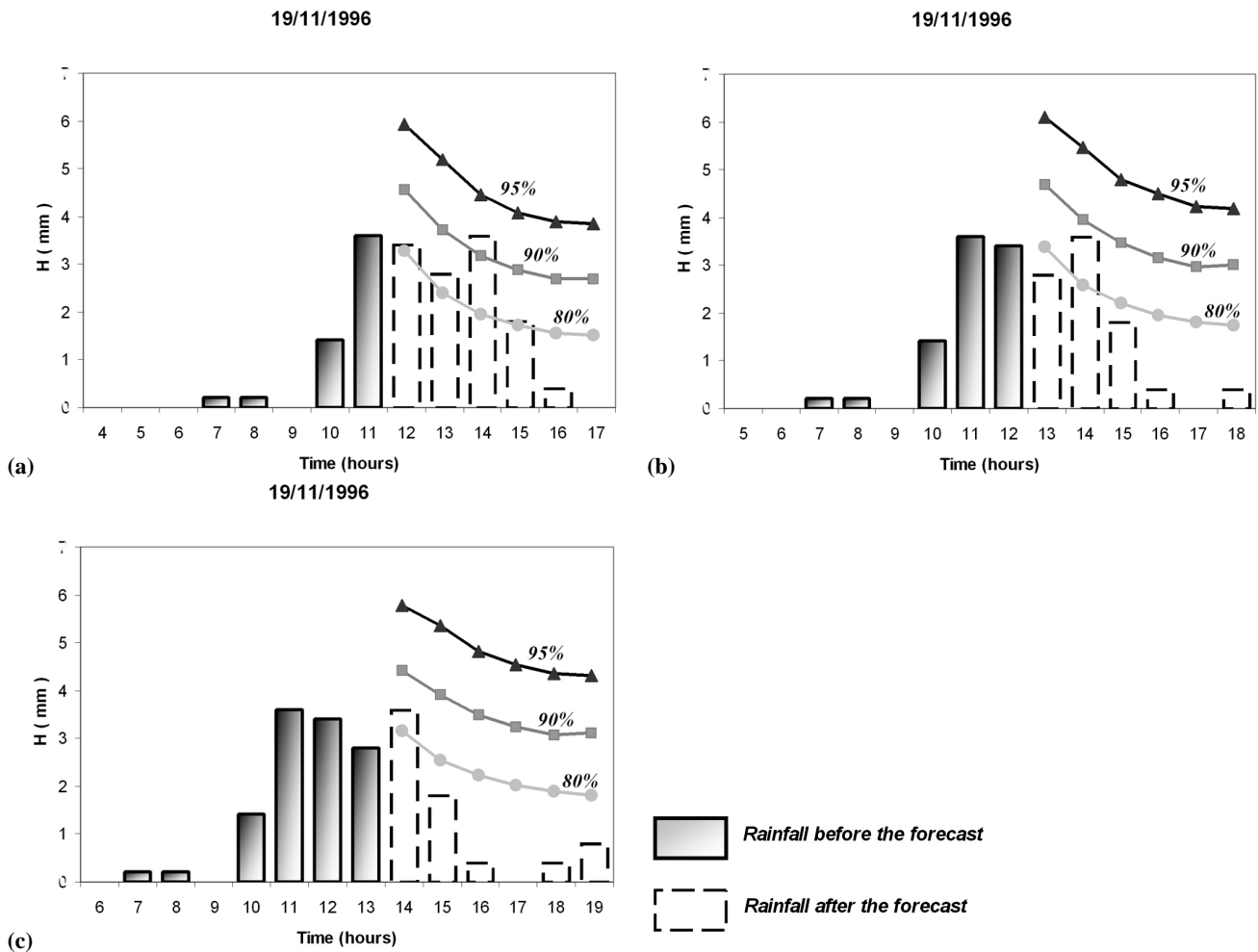


Fig. 7. Application of PRAISE Model relative to 19 November 1996, for the Cosenza raingauge, starting from (a) 11:00 LT; (b) 12:00 LT; (c) 13:00 LT.

Table 4. Cosenza raingauge: estimated parameters of the densities $f_{H_{i+1}, Z_i^{(v)}}^{(+,+)}(h_{i+1}, z_i^{(v)})$, $f_{H_{i+1}, 0}^{(+,\cdot)}(h_{i+1})$ and $f_{0, Z_i^{(v)}}^{(\cdot,+)}(z_i^{(v)})$.

Period	$1/\alpha_{h,z}$ (mm)	$\beta_{h,z}$	$1/\gamma_{h,z}$ (mm)	$\delta_{h,z}$	$\theta_{h,z}$	$1/\alpha_{h,\cdot}$ (mm)	$\beta_{h,\cdot}$	$1/\gamma_{\cdot,z}$ (mm)	$\delta_{\cdot,z}$
1 Oct–31 May	1.04	0.80	0.90	0.80	1.63	0.72	0.63	0.31	0.52
1 Oct–31 Jan	1.09	0.79	0.95	0.80	1.65	0.76	0.65	0.32	0.50
1 Feb–31 May	0.99	0.82	0.83	0.81	1.59	0.69	0.61	0.29	0.56

the latter table, temporary changes of the parameter over subperiods of the rainy season are also reported.

In Fig. 5 empirical distribution functions of the variables $H_{i+1} | H_{i+1} > 0 \cap Z_i^{(v)} > 0$, $Z_i^{(v)} | H_{i+1} > 0 \cap Z_i^{(v)} > 0$, $H_{i+1} | H_{i+1} > 0 \cap Z_i^{(v)} = 0$ and $Z_i^{(v)} | H_{i+1} = 0 \cap Z_i^{(v)} > 0$, for the Cosenza raingauge, are depicted. To these empirical distributions, Weibull distribution functions, with parameters $(\alpha_{h,z}$,

$\beta_{h,z}$), $(\gamma_{h,z}, \delta_{h,z})$, $(\alpha_{h,\cdot}, \beta_{h,\cdot})$ and $(\gamma_{\cdot,z}, \delta_{\cdot,z})$, are plotted. The plots indicate that the Weibull distributions are acceptable marginals for these variables.

4.1 Validation of the model

As regards rainfall forecasting, each simulation requires the knowledge of the rainfalls relative to the eight previous hours. Starting from these, simulations can be carried out for

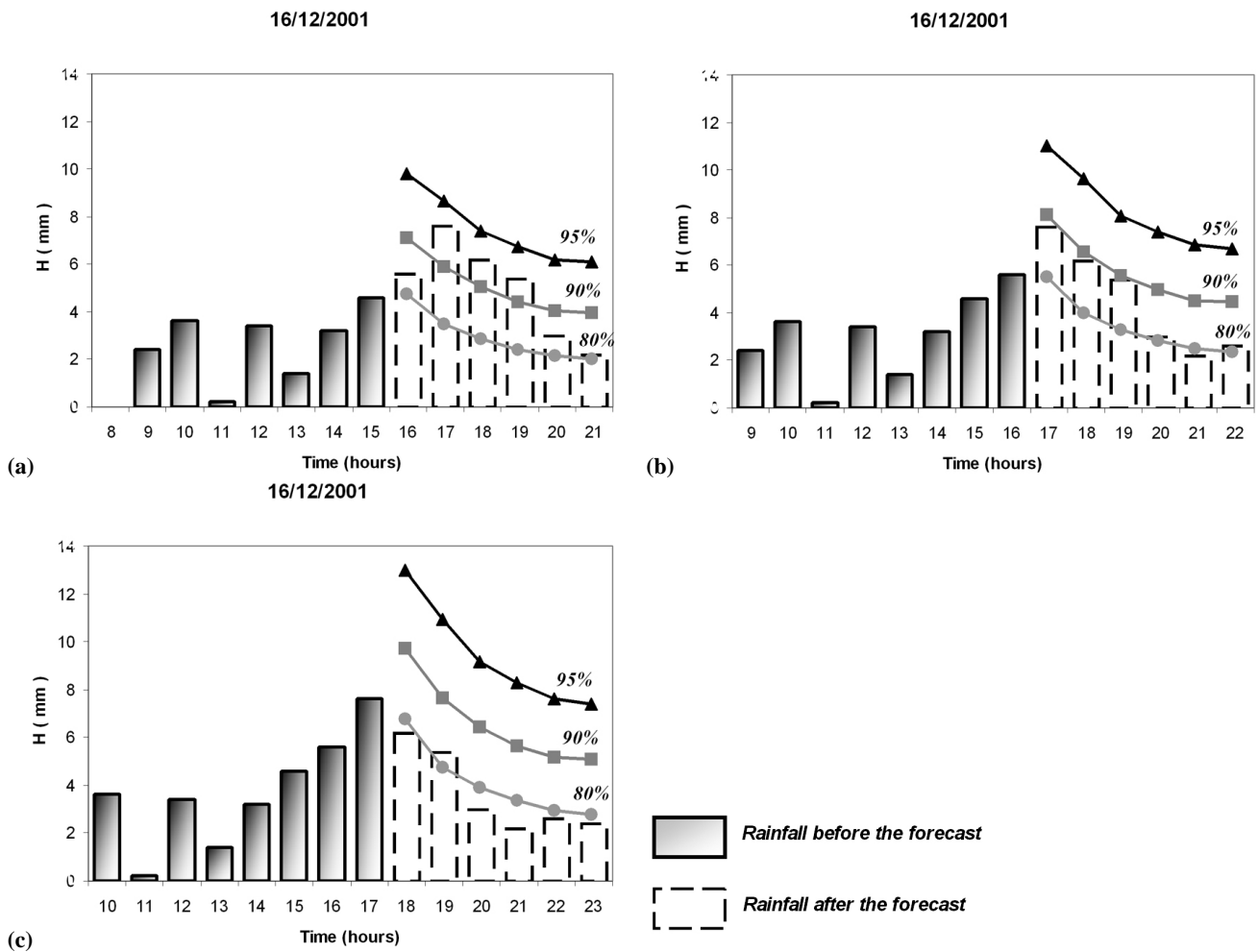


Fig. 8. Application of PRAISE Model relative to 16 December 2001, for the Catanzaro raingauge, starting from (a) 15:00 LT; (b) 16:00 LT; (c) 17:00 LT.

the successive hours. The temporal extension of the forecast should not exceed six hours. Beyond this limit the results become similar to the unconditional ones and, then, a model update by observed precipitation is necessary.

In the applications, using the Monte Carlo technique, the simulations are carried out repeating the process 10 000 times in order to have a large synthetic sample of rainfall heights for every forecasting hour. Rainfall heights are generated using the conditional probability distributions $f_{H_{i+1}|Z_i^{(v)}}(h_{i+1}|Z_i^{(v)} = z_i^{(v)})$, easily obtained from Eq. (11). The Monte Carlo technique is adopted because of the complexity of determining analytical probabilistic distributions for rainfall heights relative to forecasting the hours successive to the first one. For these distributions, convolution operations should be required.

PRAISE model has been validated considering 100 rainfall events out of about 10 000 ones. These events have been chosen in random way and weren't also used for model calibra-

tion. However, the number of these constitutes a small part of the whole ensemble of rainfall events and, consequently, their exclusion for model calibration doesn't modify parameter estimation.

In Fig. 6, the histogram, representing, for every forecasting hour, the percentile 95% of the events utilized for validation, is illustrated. Referring to the same percentile, in the figure, median value and 90% band comprising PRAISE simulated value are also reported.

In particular, as an example of model output, in this paper the applications relative to 19 November 1996 (Fig. 7) for the Cosenza raingauge, and 16 December 2001 (Fig. 8) for the Catanzaro raingauge are illustrated.

For every event, the real values of the precipitation have been compared with the percentiles 80%, 90% and 95% of the probabilistic distributions of every forecasting hour. Moreover, for every event the PRAISE model was applied shifting the starting point of forecasting for three successive hours.

The Figs. 7–8 show that rainfall heights of the real event fall between percentiles 80% and 90%. These results indicate the capability of the model to identify, for the forecast hours, statistical confidence limits containing the real rainfall heights.

5 Conclusions

This paper presents a new stochastic model named PRAISE to forecast rainfall heights at site. The mathematical background is characterized by a bivariate probability distribution, referred to the random variables H_{i+1} and $Z_i^{(v)}$, representing rainfall in a generic site and antecedent precipitation in the same site.

The peculiarity of PRAISE is the availability of the probabilistic distributions of rainfall heights for the forecasting hours, conditioned by the values of observed precipitation.

PRAISE was applied to all the telemetering raingauges of the Calabria region, in Southern Italy; the calibration model shows that the hourly rainfall series present a constant value of memory ν equal to 8 h, for every raingauge of the Calabria network. Moreover, analysing $\theta_{h,z}$ parameter mapping, it must be pointed out how higher values of $\rho_{H_{i+1}, Z_i^{(v)}}^{(+,+)}$ correlation are located in the part of region characterized by greater altitude.

The examples of validation, presented here, regarding the Cosenza and Catanzaro raingauges, indicate the capability of the model to identify, for the forecasting hours, statistical confidence limits containing the real rainfall heights. The PRAISE model therefore can be considered a very useful and simple tool for forecasting precipitation and consequently, using rainfall-runoff models or hydro-geotechnical models, floods or landslides, in planning and managing a warning system.

Edited by: A. Gelfan

References

- Abramowitz, M. and Stegun, I. A.: Handbook of mathematical functions, Dover, New York, NY, 1970.
- Box, G. E. P. and Jenkins, G. M.: Time series analysis: forecasting and control, Holden-Day, S.Francisco, 1976.
- Brockwell, P. J. and Davis, R. A.: Time Series. Theory and Methods, Springer Verlag, New York, NY, 1987.
- Burlando, P., Rosso, R., Cadavid, L. G., and Salas, J. D.: Forecasting of short-term rainfall using ARMA models, *J. Hydrol.*, 144, 193–211, 1993.
- Burlando, P., Montanari, A., and Ranzi, R.: Forecasting of storm rainfall by combined use of radar, rain gages and linear models, *Atmos. Res.*, 42, 199–216, 1996.
- Calenda, G. and Napolitano, F.: Parameter estimation of Neyman-Scott processes for temporal point rainfall simulation, *J. Hydrol.*, 225, 45–66, 1999.
- Chuang, Hui-Ya, and Sousounis, P. J.: A technique for generating idealized initial and boundary conditions for the PSU-NCAR Model MM5, *Mon. Wea. Rev.*, 128, 2875–2882, 2000.
- Cowpertwait, P. S. P., O'Connell, P. E., Metcalfe, A. V., and Mawdsley, J. A.: Stochastic point process modelling of rainfall. I. Single site fitting and validation, *J. Hydrol.*, 175, 17–46, 1996.
- Cowpertwait, P. S. P.: Mixed rectangular pulses models of rainfall, *Hydrol. Earth Syst. Sci.*, 8, 993–1000, 2004, <http://www.hydrol-earth-syst-sci.net/8/993/2004/>.
- De Luca, D. L.: Metodi di previsione dei campi di pioggia. Tesi di Dottorato di Ricerca, Università della Calabria, Italy, 2005.
- Foufoula-Georgiou, E. and Georgakakos, K. P.: Recent advances in space-time precipitation modelling and forecasting, *Recent Advances in the Modelling of Hydrologic Systems*, NATO ASI Ser., 1988.
- Georgakakos, K. P. and Kavvas, M. L.: Precipitation analysis, modeling and prediction in hydrology, *Rev. Geophys.*, 25(2), 163–178, 1987.
- Hipel, K. W. and McLeod, A. I.: Time series Modeling of Water Resources and Environmental Systems, Elsevier Science, 1994.
- Kavvas, M. L. and Delleur, J. W.: A stochastic cluster model of daily rainfall sequences, *Water Resour. Res.*, 17(4), 1151–1160, 1981.
- Kendall, M. G. and Stuart, A.: The advanced Theory of Statistics, Griffin, London, 1969.
- Kotz, S., Balakrishnan, N., and Johnson, N. L.: Continuous Multivariate Distributions – Models And Applications, Wiley, New York, NY, 2000.
- Lewis, P. A. W.: Stochastic Point Processes, Wiley, New York, NY, 1964.
- Montanari, A. and Brath, A.: Maximum likelihood estimation for the seasonal Neyman-Scott rectangular pulses model for rainfall, *Proc. of the EGS Plinius Conference*, 297–309, Maratea, Italy, 1999.
- Palmer, T. N., Brankovic, C., Buizza, R., Chessa, P., Ferranti, L., Hoskins, B. J., and Simmons, A. J.: A review of predictability and ECMWF forecast performance, with emphasis on Europe, ECMWF Research Department Technical Memorandum n. 326, ECMWF, Shinfield Park, Reading RG2-9AX, UK, 2000.
- Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.: Numerical Recipes in C. The art of scientific computing, Cambridge University Press, 1988.
- Rodriguez-Iturbe, I., Gupta, V. K., and Waymire, E.: Scale consideration in modelling of temporal rainfall, *Water Resour. Res.*, 20(11), 1611–1619, 1984.
- Rodriguez-Iturbe, I.: Scale of fluctuation of rainfall models, *Water Resour. Res.*, 22(9), 15S–37S, 1986.
- Rodriguez-Iturbe, I., Cox, D. R., and Isham, V.: Some models for rainfall based on stochastic point processes, *Proc. Royal Soc. London, A*, 410, 269–288, 1987.
- Salas, J. D., Delleur, J. W., Yevjevich, V., and Lane, W. L.: Applied modelling of hydrologic time series, Water Resources Publications, Littleton, CO, 1980.
- Sirangelo, B. and Iiritano, G.: Some aspects of the rainfall analysis through stochastic models, *Excerpta*, 11, 223–258, 1997.
- Smith, J. A. and Karr, A. F.: A point process model of summer season rainfall occurrences, *Water Resour. Res.*, 19(1), 95–103, 1983.
- Toth, E., Brath, A., and Montanari, A.: Comparison of short-term

- rainfall prediction models for real-time flood forecasting, *J. Hydrol.*, 239(1), 132–147, 2000.
- Untch, A., Miller, M., Hortal, M., Buizza, R., and Janssen, P.: Towards a global meso-scale model: the high-resolution system TL799L91 & TL399L62 EPS, Newsletter n. 108, ECMWF, Shinfield Park, Reading RG2-9A, UK, 2006.
- Waymire, E. and Gupta, V. K.: The mathematical structure of rainfall representations, 1. A review of the stochastic rainfall models, *Water Resour. Res.*, 17(5), 1261–1272, 1981.