A space-time generator for rainfall nowcasting: the PRAISEST model

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

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

Received: 22 January 2008 – Accepted: 27 January 2008 – Published: 18 March 2008

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

Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

The paper introduces a new stochastic technique for forecasting rainfall in space-time domain: the PRAISEST Model (Prediction of Rainfall Amount Inside Storm Events: Space and Time). The model is based on the assumption that the rainfall height $H$ accumulated on an interval $\Delta t$ between the instants $i\Delta t$ and $(i+1)\Delta t$ and on a spatial cell of size $\Delta x \Delta y$ is correlated either with a variable $Z$, representing antecedent precipitation at the same point, either with a variable $W$, representing simultaneous rainfall at neighbour cells. The mathematical background is given by a joined probability density $f_{H,W,Z}(h, w, z)$ in which the variables have a mixed nature, that is a finite probability for null value and infinitesimal probabilities for the positive values. As study area, the Calabria region, in Southern Italy, has been selected. The region has been discretised by $10 \times 10$ km cell grid, according to the raingauge network density in this area. Storm events belonging to 1990–2004 period were analyzed to test performances of the PRAISEST model.

1 Introduction

The risk mitigation in landslide or flood prone areas is one of the most important topics in environmental sciences. In the next future this relevance will increase owing to the development of mitigation and adaptation policies related to climatic change (Stern, 2006; IPCC, 2007). In this scenario the non structural measures, mainly based on early warning system, will play, increasingly, a relevant role. So all the related hydrological topics like rainfall-runoff modelling or rainfall-landslide relationships will be also developed, with the general scientific aim of realizing an accurate simulation of the real phenomena, but also in order to forecast landslide and flood events with a lag time large enough for activating civil protection measures.

Indeed, in all the cases where the phenomenon rapidly evolves, like flash floods or shallow landslides, the lag time between observed rainfall and flood or landslide results
too short and must be extended by rainfall fields forecasting. This is the main reason for the rise of the interest in this topic (Arnaud and Lavabre, 2002; D’Odorico and Fagherazzi, 2003; Sirangelo and Braca, 2004; Chiang et al., 2007; Reed et al., 2007; Bloschl et al., 2008).

In the technical literature rainfall forecasting models can be classified in time stochastic models, space-time stochastic models and meteorological ones.

In the first class, 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; Burlando et al., 1993; Bras and Rodriguez-Iturbe, 1984; Brockwell and Davis, 1987; Hipel and McLeod, 1994, Brath et al., 1999, 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 (Bartlett, 1963; Lewis, 1964; Kavvas and Delleur, 1981; Smith and Karr, 1983; Rodriguez-Iturbe et al., 1984; Rodriguez-Iturbe, 1986; Rodriguez-Iturbe et al., 1987b; Sirangelo and Versace, 1990; Cowpertwait, 1991; Onof and Wheater, 1994; Katz and Parlange, 1995, 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 interstorm periods also.

Space-temporal stochastic models can be classified in Multivariate models and Multidimensional ones. The former consider several rain gauges simultaneously and are intended to preserve the covariance structure of the historical rainfall data existing in the network points. On the temporal axis, forecasting can be made by autoregressive scheme (STARMA and CARMA models, Cliff et al., 1975; Pfeifer and Deutsch, 1980; Cressie, 1993; Burlando et al., 1996) or by point processes scheme (Multisite models, Cox and Isham, 1994, Wheater et al., 2000; Koutsoyiannis et al., 2003).

The latter models attempt to characterize the rainfall phenomenon at every point over the area of interest; space-time domain can be a set of “a priori” well known point (Bras and Rodriguez-Iturbe, 1976; Meiring et al., 1997) or defined in random way (Waymire et

Finally, meteorological models (Mesinger and Arakawa, 1976; Pielke, 1984; Buzzi et al., 1993; Chuang et al., 2000; Ebert and McBride, 2000; Palmer et al., 2000; Walser and Schar, 2004; Untch et al., 2006) solve in numerical way partial differential equations of atmosphere thermodynamics: they can be classified in GCM (Global Circulation Models) and LAM (Limited Area Models).

Meteorological models are useful qualitative-quantitative rainfall forecasting tools on 24–72 h interval and on large spatial scale. In such cases, indeed, absolute precision is not required for practical application, then the precision of forecasting model is quite enough. When both the forecasting lag time and spatial scale decrease the effectiveness and the precision of this kind of models also decrease (Koussis et al., 2003; Bartholmes and Todini, 2005; Sharma et al., 2007). Unfortunately this is the time space scale of the fast phenomena (flash floods and shallow landslide) that require rainfall forecast for civil protection measures. Then it is not plenty profitable to rely on meteorological models for quantitative rainfall forecasting, as the probability of both missed and false alarms may be too large.

Consequently, in order to perform short term real-time rainfall forecasts for small basins (i.e. with size ranging 100–1000 km$^2$), stochastic models appear to be competitive, as they take account of the hydrological characteristics of the investigated area.

Nevertheless, stochastic models input is only constituted by antecedent rainfalls, so they provide the same prevision, whether meteorological models forecast a wet period or a dry one. For these reasons, coupling stochastic and meteorological models appears a very interesting topic for rainfall forecasting in the small time space scale (Di Tria et al., 1999; Sirangelo et al., 2006).

This work introduce a new space-temporal models to forecast rainfall fields named PRAISEST (Prediction of Rainfall Amount Inside Storm Events: Space and Time). It is a multidimensional space-time model, that can be considered like the generalization of the at-site model PRAISE proposed by the authors (Sirangelo et al., 2007).
PRAISEST is based on the assumption that the evaluation of rainfall height $H$, accumulated over an interval $\Delta t$ and on a spatial cell of size $\Delta x \Delta y$, depends on antecedent precipitation at the same site, and on rainfalls of neighbour cells.

In the following sections the theoretical bases of the proposed model (Sect. 2), fitting techniques (Sect. 3) and rainfall generation algorithm (Sect. 4) are discussed. The application of the model to the case study of Calabria region, in Southern Italy, is reported in Sect. 5.

2 The PRAISEST Model

2.1 Identification of random variables

In the PRAISEST model, the rainfall heights $H_i$, accumulated over an interval $\Delta t$ between the instants $(i-1)\Delta t$ and $i\Delta t$, and on a spatial cell of size $\Delta x \Delta y$ are considered as a realisation of a weakly stationary stochastic fields with discrete parameter \{\(H_i; i \in I\)\}, where $I$ indicates the integer numbers. In the following the instant $i\Delta t=t_0$ is assumed as current time, so the observed rainfall heights have subscripts less or equal to $i$ while the forecasted ones have subscripts greater than $i$.

Besides, the model considers other two random variables:

1. $Z_i^{(v)}$, function of the $v$ variables $H_i$, $H_{i-1}$, ... $H_{i-v+1}$, representing rainfall heights accumulated over intervals $[(i-1)\Delta t ; i\Delta t]$, ... $[(i-v)\Delta t ; (i-v+1)\Delta t]$, on the same cell where the future rainfall value, $H_{i+1}$, must be estimated.

2. $W_{i+1}$, representing a weighed average of rainfall heights accumulated over temporal interval $[i\Delta t ; (i+1)\Delta t]$ on neighbour cells; in this case the Implicit scheme is adopted. An alternative is constituted by considering the variable $W_i$, weighed average of precipitations referred to the interval $[(i-1)\Delta t ; i\Delta t]$, and the adopted scheme is named Explicit (Fig. 1).
The random variable $Z_{i}^{(v)}$ is calculated as linear function of $v$ rainfall heights $H_i, H_{i-1}, \ldots, H_{i-v+1}$ referred to the reference cell:

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

with the conditions $0 < \alpha_j \leq 1$, for $j = 0, 1, \ldots, v-1$, and $\sum_{j=0}^{v-1} \alpha_j = 1$. The extension of the temporal “memory” $v$ of the rain field, for every cell, can be assumed equal to the minimum value of $v$ for which the sample maximum absolute scattering $\chi_r (v)$ (Sirangelo et al., 2007) results less than a fixed critical value $\chi_{r,cr}$.

The coefficients $\alpha_j$ can be estimated by maximization of the coefficient of linear correlation $\rho_{H_{i+1} Z_{i}^{(v)}}$. If $v$ is large the number of parameters may be too high, then a technique of linear filtering results convenient, as the coefficients $\alpha_j$ depend on a reduce number of parameters.

In this paper, the gamma-power function is used as filter (De Luca, 2005), depending on three parameters, that provides a good fitting of the estimated values of $\alpha_j$.

Then, the coefficients $\alpha_j$ can be calculated as:

$$\alpha_j = \frac{P [a, ((j + 1)b\Delta t)^c] - P [a, (jb\Delta t)^c]}{P [a, (vb\Delta t)^c]}$$

where $P (a, x) = \frac{1}{\Gamma (a)} \int_0^x e^{-t} t^{a-1} dt$ is the incomplete gamma function (Abramowitz and Stegun, 1970).

More precisely, for every set $(a, b, c)$, the sample variable $Z_{i}^{(v)}$, $i = \nu, \nu + 1, \ldots, N - 1$, can be evaluated and then the optimal values $\hat{a}, \hat{b}, \hat{c}$ can be obtained by maximizing the function:
\begin{equation}
\hat{H}(a, \hat{b}, \hat{c}) = \max_{a > 0, b > 0, c > 0} \left[ r_{HZ}(a, b, c) \right]
\end{equation}

using a numerical procedure (Press et al., 1988).

As previously reported, the random variables \(W_i \) and \(W_i\) represent a weighed average of the rainfall heights on the neighbour cells; in this paper there is the assumption that the spatial memory is limited to the four adjacent cells (Fig. 1), referred to the time forecast interval (Implicit scheme) or to the previous interval (Explicit scheme). This hypothesis allows to simplify the mathematical background and the results are similar to those which are obtained if a greater number of neighbour cells were considered.

The expressions of \(W_i\) are:

\begin{equation}
W_{i+1} = \sum_{j=1}^{4} \beta'_j H_{i+1}^{(j)} \quad \text{and} \quad W_i = \sum_{j=1}^{4} \beta''_j H_i^{(j)}
\end{equation}

where:

\begin{equation}
\beta'_j = \frac{\rho'_{j,0}}{\sum_{j=1}^{4} \rho'_{j,0}} \quad \beta''_j = \frac{\rho''_{j,0}}{\sum_{j=1}^{4} \rho''_{j,0}} \quad j = 1, 2, 3, 4
\end{equation}

and \(\rho'_{1,0}, \rho'_{2,0}, \rho'_{3,0}, \rho'_{4,0} \) and \(\rho''_{1,0}, \rho''_{2,0}, \rho''_{3,0}, \rho''_{4,0} \) indicate the linear correlation coefficients between the reference cell 0 and the neighbour cells 1, 2, 3 and 4 (Fig. 1), considering respectively “simultaneous rainfalls” and “1-h shift rainfalls”.

Obviously if the field can be considered locally isotropic, the coefficients \(\beta'_j \) and \(\beta''_j \) assume the same value equal to 1/4.

In the following, for notation simplicity, the subscripts of random variables \(H, W\) and \(Z\) will be removed where possible.

755
2.2 Structure of the joined probability density

To identify the joined probability density $f_{H,W,Z}(h,w,z)$ it is necessary to consider the mixed nature of random variables $H$, $W$ and $Z$. All the three variables are non-negative and characterized by a finite probability in correspondence of the null value and by infinitesimal probabilities in correspondence of the positive values.

Then, indicated with $p_{+,+,*}$, $p_{+,*,*}$, $p_{+,+,*}$, $p_{+,*,*}$, $p_{+,*,*}$ and $p_{*,*,*}$ the probabilities associated to the events:

\[
H>0 \cap W>0 \cap Z>0, H>0 \cap W>0 \cap Z=0, H=0 \cap W>0 \cap Z>0, H=0 \cap W>0 \cap Z=0, H=0 \cap W=0 \cap Z=0,
\]

the joined probability density $f_{H,W,Z}(h,w,z)$ assumes the form:

\[
f_{H,W,Z}(h,w,z) = p_{*,*,*} \delta(h) \delta(w) \delta(z) + p_{*,*,*} \cdot f_{H,0,0}^{(+,*,*)}(h) \cdot \delta(w) \delta(z) + p_{*,+,*} \cdot f_{0,W,0}^{(+,+,*)}(w) \cdot \delta(h) \delta(z) + p_{+,+,*} \cdot f_{H,0,Z}^{(+,+,*)}(h,w) \delta(z) + p_{+,+,+} \cdot f_{H,W,Z}^{(+,+,+)}(h,w,z) \delta(z)
\]

where the symbol $\delta(\cdot)$ indicates the Dirac’s delta function and:

\[
f_{H,W,Z}^{(+,+,+)}(h,w,z) dhdwdz =
\]

\[
Pr\left[h \leq H < h + dh \wedge w \leq W < w + dw \wedge z \leq Z < z + dz \mid H > 0 \cap W > 0 \cap Z > 0\right]
\]

\[
f_{H,W,0}^{(+,+,*)}(h,w) dhdw = Pr\left[h \leq H < h + dh \wedge w \leq W < w + dw \mid H > 0 \cap W > 0 \cap Z = 0\right]
\]

\[
f_{H,0,Z}^{(+,*,+)}(h,z) dhdz = Pr\left[h \leq H < h + dh \wedge z \leq Z < z + dz \mid H > 0 \cap W = 0 \cap Z > 0\right]
\]

\[
f_{0,W,Z}^{(*,+,+)}(w,z) dwdz = Pr\left[w \leq W < w + dw \wedge z \leq Z < z + dz \mid H = 0 \cap W > 0 \cap Z > 0\right]
\]
A space-time generator for rainfall nowcasting: the PRAISEST model
P. Versace et al.

2.2.1 Structure of probability density function $f_{H,W,Z}^{(+,+,+)}(h,w,z)$

Considering the event $H>0 \cap W>0 \cap Z>0$, the analytical expression of the density $f_{H,W,Z}^{(+,+,+)}(h,w,z)$ must be chosen among the trivariate probabilistic distributions defined for strictly positive variables. In this paper, the Al Saadi and Young’s trivariate exponential distribution (1982) is adopted:

$$f_{X_1,X_2,X_3}(x_1,x_2,x_3) = \theta^2 \exp \left[ -\theta (x_1 + x_2 + x_3) \right] \cdot \sum_{i=0}^{\infty} \frac{1}{(i!)^3} \left[ \theta^2 (\theta - 1) x_1 x_2 x_3 \right]^i$$

with $x_1>0, x_2>0, x_3>0, \theta \geq 1$ and linear correlation coefficients $\rho_{X_1X_2}=\rho_{X_1X_3}=\rho_{X_2X_3}=1-1/\theta$; this function is characterized by exponential marginal density functions with unitary scale parameter.

With the purpose of making more flexible the model, allowing either amodal marginal density functions either modal ones, a triple power transformation has been adopted for PRAISEST:

$$x_1 = \alpha_{H}^{(+,+,+)} h^{\beta_{H}^{(+,+,+)}} \quad h > 0, \alpha_{H}^{(+,+,+)} > 0, \beta_{H}^{(+,+,+)} > 0$$  

$$x_2 = \alpha_{W}^{(+,+,+)} w^{\beta_{W}^{(+,+,+)}} \quad w > 0, \alpha_{W}^{(+,+,+)} > 0, \beta_{W}^{(+,+,+)} > 0$$  

$$x_3 = \alpha_{Z}^{(+,+,+)} z^{\beta_{Z}^{(+,+,+)}} \quad z > 0, \alpha_{Z}^{(+,+,+)} > 0, \beta_{Z}^{(+,+,+)} > 0$$
from which the following density of probability distribution is obtained:

\[
f_{H,W,Z}(h,w,z) = a_H^{(+,+)} \beta_H^{(+,+)} h^{\theta_H^{(+,+)}} \cdot a_W^{(+,+)} \beta_W^{(+,+)} w^{\theta_W^{(+,+)}} \cdot a_Z^{(+,+)} \beta_Z^{(+,+)} z^{\theta_Z^{(+,+)}} - 1 \left( \theta^{(+,+)} \right)^2.
\]

\[
\exp \left[ -a_H^{(+,+)} \left( \theta^{(+,+) - 1} + a_W^{(+,+)} w^{\theta_W^{(+,+)}} + a_Z^{(+,+)} z^{\theta_Z^{(+,+)}} \right) \right].
\]

\[
\sum_{i=0}^{\infty} \frac{1}{(i!)^3} \left[ \left( \theta^{(+,+) - 1} + a_W^{(+,+)} w^{\theta_W^{(+,+)}} + a_Z^{(+,+)} z^{\theta_Z^{(+,+)}} \right) \right]^i.
\]

with \(a_H^{(+,+)} > 0, \beta_H^{(+,+)} > 0, a_W^{(+,+)} > 0, \beta_W^{(+,+)} > 0, a_Z^{(+,+)} > 0, \beta_Z^{(+,+)} > 0\) and \(\theta^{(+,+)} \geq 1\).

Means and variances for the variable \(H\) are the following:

\[
\mu_H^{(+,+)} = E \left( H \mid H > 0 \cap W > 0 \cap Z > 0 \right) = \frac{\Gamma \left( 1 + 1/\beta_H \right)}{\alpha_H^{1/\beta_H}} \tag{19}
\]

\[
\left( \sigma_H^{2(+,+)} \right) = \text{var} \left( H \mid H > 0 \cap W > 0 \cap Z > 0 \right) = \frac{1}{\alpha_H^{2/\beta_H}} \left[ \Gamma \left( 1 + 2/\beta_H \right) - \Gamma^2 \left( 1 + 1/\beta_H \right) \right] \tag{20}
\]

Similar expressions are for the variables \(W\) and \(Z\).

Linear correlation coefficient \(\rho_{H,W}^{(+,+)}\), ranging in the interval \((0; 1)\) can be written as:

\[
\rho_{H,W}^{(+,+)} = \frac{1}{\sigma_H^{1/\beta_H} \sigma_W^{1/\beta_W}} \sum_{i=0}^{\infty} \frac{1}{(i!)^2} \left( \frac{1}{\sigma_H^{1/\beta_H}} \cdot \Gamma(i+1+1/\beta_H) \cdot \Gamma(i+1+1/\beta_W) - \Gamma(1+1/\beta_H) \cdot \Gamma(1+1/\beta_W) \right) \cdot \sqrt{\Gamma(1+2/\beta_H) - \Gamma^2(1+1/\beta_H) \cdot \sqrt{\Gamma(1+2/\beta_W) - \Gamma^2(1+1/\beta_W) \right)} \tag{21}
\]

Equal ranges and analogous expressions hold for \(\rho_{H,Z}^{(+,+)}\) and \(\rho_{W,Z}^{(+,+)}\).
From Eq. (21) it seems clear that higher values of $\theta^{(+,+,+)}$ give higher correlation values among $H$, $W$ and $Z$ variables. Moreover if $\theta^{(+,+,+)} = 1$ the random variables are indipendent; consequently:

\[
 f_{H,W,Z}^{(+,+,+)} (h, w, z) = f_H^{(+,+,+)} (h) f_W^{(+,+,+)} (w) f_Z^{(+,+,+)} (z) \tag{22}
\]

where $f_H^{(+,+,+)} (h)$, $f_W^{(+,+,+)} (w)$, $f_Z^{(+,+,+)} (z)$ are the Weibull marginal densities, as easily verifiable.

2.2.2 Structure of the other probability density functions

When only one random variable assumes null value and the others are positive, Weibull-Bessel laws (Sirangelo and Versace, 2002) are chosen as probability density functions $f_{H,W,0}^{(+,+,*)} (h, w)$, $f_{H,0,Z}^{(+,*,+)} (h, z)$ and $f_{0,W,Z}^{(*,+,+)} (w, z)$ appearing in Eq. (6). For example, in the case $H>0 \cap W>0 \cap Z=0$ the expression is:

\[
 f_{H,W,0}^{(+,+,*)} (h, w) = \theta^{(+,+,*)} \alpha_H^{(+,+,*)} \beta_H^{(+,+,*)} (h)^{\beta_H^{(+,+,*)}-1} \alpha_W^{(+,+,*)} \beta_W^{(+,+,*)} (w)^{\beta_W^{(+,+,*)}-1} \cdot \
 \exp \left\{ -\theta^{(+,+,*)} \left[ \alpha_H^{(+,+,*)} (h)^{\beta_H^{(+,+,*)}} + \alpha_W^{(+,+,*)} (w)^{\beta_W^{(+,+,*)}} \right] \right\} . \tag{23}
\]

\[
 l_0 \left[ 2 \sqrt{\theta^{(+,+,*)} (\theta^{(+,+,*)} - 1)} \alpha_H^{(+,+,*)} (h)^{\beta_H^{(+,+,*)}} \alpha_W^{(+,+,*)} (w)^{\beta_W^{(+,+,*)}} \right]
\]
defined for $h>0, w>0$, in which $\alpha_H^{(+,+,*)}>0, \beta_H^{(+,+,*)}>0, \alpha_W^{(+,+,*)}>0, \beta_W^{(+,+,*)}>0, \theta^{(+,+,*)}\geq1$ and $l_0 (\cdot)$ is the modified Bessel function of zero order (Abramowitz and Stegun, 1970).

Similarly to the function in Eq.(18), $f_{H,W,0}^{(+,+,*)} (h, w)$ is derived by a double power transformation of a density function, that is, in this case, the Moran-Dowton bivariate exponential (Kotz et al., 2000). The obtained distribution is characterized by Weibull marginal densities, which can have either amodal behaviour either modal one.
The linear correlation coefficient $\rho_{HW}^{(+,+,+)}$ presents the value range $[0; 1)$ and higher values of $\theta^{(+,+,+)}$ give higher correlation values between $H$ and $W$ variables; it is easily verifiable that if $\theta^{(+,+,+)}=1$ the random variables are independent.

Analogous expressions are used as regards $f_{H,0,Z}^{(+,+,+)}(h,z)$ and $f_{0,W,Z}^{(+,+,+)}(w,z)$ probability density functions.

When only one variable is positive and the others assume null value, Weibull laws are adopted, as reported for $H>0 \cap W=0 \cap Z=0$ (similar formulas exist for $H=0 \cap W>0 \cap Z=0$ and $H=0 \cap W=0 \cap Z>0$):

$$f_{H,0,0}^{(+,+,+)}(h) = \alpha_{H}^{(+,+,+)} \beta_{H}^{(+,+,+)} h_{H}^{\beta_{H}^{(+,+,+)}-1} \exp \left[ -\alpha_{H}^{(+,+,+)} h_{H}^{\beta_{H}^{(+,+,+)}} \right]$$

(24)

with $h>0$; $\alpha_{H}^{(+,+,+)}>0$; $\beta_{H}^{(+,+,+)}>0$

3 Model calibration

The trivariate probability distribution function $f_{H,W,Z}(h,w,z)$ presents, either in the implicit scheme either in the explicit one, 42 parameters for every cell. This number is not too large, if we consider that in 10 years there are 87,600 hourly data related to every cell, i.e. the $d/p$ (data/parameters in a generic cell) ratio is approximatively equal to 2000, and remains high enough (about 150) also if positive rainfall data are only considered. Using raingauge data and cell domain the ratio does not change if the number of raingauges and cells are similar.

This $d/p$ ratio value allows consistent evaluations of PRAISEST parameters referred to the whole spatial domain.

The first calibration step, for every cell, is the evaluation of $(a, b, c)$ by numerical technique. Estimation of $\beta_{j}'$ or $\beta_{j}''$, (owing to implicit or explicit scheme) $j=1, 2, 3, 4$, follows by analyzing sample linear correlation coefficient $r_{1,0}', r_{2,0}', r_{3,0}', r_{4,0}'$ or $r_{1,0}'', r_{2,0}'', r_{3,0}'', r_{4,0}''$. 760
The probabilities $p_{+,.+}, p_{+,.+}, p_{+,.+}, p_{+,.+}, p_{+,.+}, p_{+,.+}$ and $p_{+,.+}$ can be estimated by the frequencies $\hat{F}_{+,.+}, \hat{F}_{+,.+}, \hat{F}_{+,.+}, \hat{F}_{+,.+}, \hat{F}_{+,.+}, \hat{F}_{+,.+}$ of the events $H > 0 \cap W > 0 \cap Z > 0, H > 0 \cap W > 0 \cap Z = 0, H = 0 \cap W > 0 \cap Z > 0, H > 0 \cap W > 0 \cap Z = 0, H = 0 \cap W > 0 \cap Z = 0$ and $H = 0 \cap W = 0 \cap Z > 0$.

The probability densities parameters can be estimated using the method of moments.

For the density $f_{H,W,Z}^{(+,+,+)}(h,w,z)$, the parameters $\alpha_{H}^{(+,+,+)}, \beta_{H}^{(+,+,+)}$, $\alpha_{W}^{(+,+,+)}, \beta_{W}^{(+,+,+)}, \alpha_{Z}^{(+,+,+)}, \beta_{Z}^{(+,+,+)}$ can be estimated fitting sample means and variances using classic expressions of Weibull distribution parameter estimation.

The estimation of the parameter $\theta^{(+,+,+)}$, is performed minimizing the following function:

$$R(\theta^{(+,+,+)} = \omega_{1} \left( \rho_{HW}^{(+,+,+)} - r_{HW}^{(+,+,+)} \right)^{2} + \omega_{2} \left( \rho_{HZ}^{(+,+,+)} - r_{HZ}^{(+,+,+)} \right)^{2} + \omega_{3} \left( \rho_{WZ}^{(+,+,+)} - r_{WZ}^{(+,+,+)} \right)^{2} \tag{25}$$

where $r_{HW}^{(+,+,+}), r_{HZ}^{(+,+,+)}, r_{WZ}^{(+,+,+)}$ are the sample linear correlation coefficients, and the sum of the weights $\omega_{1}, \omega_{2}$ and $\omega_{3}$ is unitary. The function Eq. (25) depends only on the parameter $\theta^{(+,+,+)}$, since the remaining parameters have been evaluated in a previous step.

With similar procedures, parameters of the remaining density functions of Eq. (6) can be evaluated.

### 4 Rain fields generation algorithms

Using the Explicit scheme of PRAISEST model, the values of both $Z$ and $W$ are known, so it is possible to generate the value of the variable $H$ on the whole domain, according to usual Monte Carlo techniques. In all the cells, rainfall heights are generated by means of the cumulate distribution function $F \left( H_{i+1} \leq h_{i+1} \mid W_{i} = w_{i}, Z_{i} = z_{i} \right)$, with ap-
appropriate expressions for the four possible cases: \( W_i \) null or positive and \( Z_i \) null or positive.

On the contrary, when Implicit scheme is used, the generation of the rainfall heights on the entire domain differs from the standard Monte Carlo approach. In fact, at the current time \( i \), the values of the random variable \( Z_i \) in every cell are known, but values of \( W_{i+1} \) and \( H_{i+1} \) on the entire domain must be generated. Such generations cannot be carried out independently cell by cell, because the variables are linked by congruence equations. This problem has been solved using the following “Chess-Board” algorithm (Fig. 2):

1. Starting from the “0” cells of the spatial domain, for everyone, knowing the value of \( Z_i \), generation is made using the random number \( R_U^{(0)} \), by the formula
\[
h_{i+1}^{(0)} = F^{-1}_{H | W} \left( R_U^{(0)} | W_{i+1}^{(0)} \geq 0, Z_i^{(0)} = Z_i^{(0)} \right),
\]
i.e. the variable \( H_{i+1} \) is generated supposing zero as lower bound for \( W_{i+1} \). This type of generation is justified because, in the model, rainfall heights in the “0” cells are independent among them.

2. As regards the “1” cells, knowing the value of \( Z_i \), \( W_{i+1}^{(1)} \) is set equal to the linear combination of the \( H_{i+1}^{(0)} \) in the neighbours “0” cells. Consequently, generation is made by the formula
\[
h_{i+1}^{(1)} = F^{-1}_{H | W, Z} \left( R_U^{(1)} | W_{i+1}^{(1)} = W_{i+1}^{(1)}, Z_i^{(0)} = Z_i^{(0)} \right)
\]
using the random number \( R_U^{(1)} \).

As in the Explicit scheme, the function
\[
F_{H | W, Z} \left( H_{i+1} \leq h_{i+1}^{(1)} | W_{i+1} = W_{i+1}, Z_i = Z_i \right)
\]
has different expressions in the four possible cases: \( W_{i+1} \) null or positive and \( Z_i \) null or positive.
5 Application

5.1 Parameters estimation

PRAISEST model has been applied using the hourly rain heights database of the tele-metering raingauge network of the “Centro Funzionale Meteorologico Idrografico Mareografico” of the Calabria region. The network, extended all over the Calabria and Basilicata regions has 92 stations for the period 1990–2001, and 126 stations from 2002 (Fig. 3). Approximately, 13 million of hourly rainfalls form the database, of which about 7% are rainy.

The region was discretized by 10 km × 10 km cell grid, according to the hourly raingauge network density in this area.

In order to respect the hypothesis of stationary process, only the data measured during the “rainy season”, 1st October–31st May have been used (De Luca, 2005). In this period, correlation structure, mean and variance of the sample appear significantly homogeneous (Sirangelo et al., 2007). So the d/p ratio is equal to about 2050 and about 140 considering only rainy intervals.

The historical series do not refer to a regular mesh, so the model parameters have been estimated for every raingauge and then mapped on the regular discretized domain by using a surface spline technique (Yu, 2001).

The extension of the “temporal memory”, i.e. the parameter ν, has been determined for every raingauge starting from sample partial autocorrelogram, and using the technique described at point 2.1. The value of $\chi_{r,cr}$ has been fixed equal to 0,025, and the estimate of ν has been $\hat{\nu}=8$, as depicted in Fig. 4, where the expected value of the scattering $\chi_r (\nu)$, considering all the telemeter-raingauges, is represented.

The coefficients $\alpha_j$ have been calculated, for every raingauge, applying the gamma-power function as filter (Eqs. 2–3).

For the evaluation of coefficients $\beta_j'$ and $\beta_j''$, defining $W_{i+1}$ and $W_i$ (Eq. 4), sample directional spatial correlograms have been analysed, using simultaneous rainfall for
implicit scheme and “1-h shift” rainfall for explicit one, and considering, on abscissa, distance between raingauges over the range (5 km; 15 km), representing spatial resolution of the region, discretized by 10 km × 10 km cell grid. Four circular sectors of 90 degrees, centred in the NE, SE, SW, NW directions have been considered. Sample correlation values appear to be independent on the directions, as depicted in Fig. 5 for the implicit scheme, so the rain fields can be considered as local isotropic, and then $\beta'_j = 1/4$ and $\beta''_j = 1/4$, $j = 1, 2, 3, 4$.

Figure 6 shows an example of parameter mapping in Calabria region, referred to $\theta^{(+,+)}$, for $H > 0 \cap W > 0 \cap Z > 0$ event, and for the Implicit scheme. Greater values are located in the Southern Calabria, so in this area the variables $H$, $W$ and $Z$ appear more strongly correlated.

5.2 Model validation

The hourly scattered data, referred to raingauges network, have been interpolated, by a surface spline technique, on regular mesh, of 10 × 10 Km square cells.

Each rain field simulation requires the knowledge of the rainfalls during the eight previous hours. The rain field simulations can be carried out for the successive hours, but the temporal extension of the forecast should not exceed six hours. Beyond this limit the uncertainty in rainfall evaluation increases, as the influence of recorded rainfall decreases.

In the application here described, 10 000 simulations of the process have been carried out, by Monte Carlo technique, in order to obtain a large synthetic sample. Rainfall heights are generated using the conditional probability distributions $f_{H|W,Z}(h|W=w, Z=z)$, obtained from Eq. (6). The Monte Carlo technique is adopted because of the complexity of determining analytical probabilistic distributions for forecasted rainfall during the hours successive to the first one. For these distributions, convolution operations are required.

For validation of PRAISEST model 100 rainfall events out of about 10 000 ones have
been considered. These events have been chosen in random way and have not been used for model calibration. However, the number of these constitutes a small part of the whole ensemble of rainfall events and, consequently, their exclusion for model calibration does not modify parameter estimation.

As examples the applications of the implicit scheme relative to 1st February 1998 (Figs. 7–9) and 24th November 1999 events in the Calabria region, are illustrated in Figs. 10–12.

For the first event the rain fields from 08:00 p.m. of 31st January to 03:00 a.m. of 1st February have been used as model memory, and the simulation period starts from 04:00 a.m. and finishes at 09:00 a.m. of 1st February. The second event is characterized by a rain field memory from 09:00 p.m. of 23rd November to 04:00 a.m. of 24th November, while the simulation period starts from 05:00 a.m. and finishes at 10:00 a.m. of 24th November.

On the abscissa, the cells are sorted from left to right, and from North to South. In the figures, besides the rain histograms effectively occurred, for every cell, percentiles 90% and 95% of the simulated fields are reported. Following the axis of the abscissas, the chief towns of Cosenza (CS), Crotone (KR), Catanzaro (CZ), Vibo Valentia (VV) and Reggio Calabria (RC) are met in this order.

One tail significance test, at 5% and 10% significance level, has been performed. The diagrams show that observed rainfall for all, but one, cells are inferior to percentiles 95% of forecasted values. In the most cases observed values are also inferior to percentiles 90% of forecasted ones. Then in all cases the results obtained by PRAISEST model seem in agreement with observed data.

6 Conclusions

The PRAISEST model, presented herein, is a multidimensional space-time model for forecasting rainfall fields. Mathematical background is characterized by a trivariate probability distribution, referred to the random variables $H$, $Z$ and $W$, representing...
rainfall forecast at the generic cell, antecedent precipitation in the same cell and rainfall in the adjacent cells. Two different schemes can be used, indicated as Implicit and Explicit. The former considers \( W \) at the forecast time, while the latter is referred to the value of \( W \) at the current time.

The results in simulation at regional scale show the capability of the model to forecast distribution functions of the rainfall in the next six hours that are in agreement with the observed values, at least at 10% significance level.

PRAISEST therefore can be easily coupled with other models like rainfall-runoff and rainfall-landslide ones for nowcasting of fast phenomena, characterized by short lag time, like flash floods and shallow landslides.

Moreover the model is highly flexible and can be usefully adopted with any cell grid, within the limit of the raingauge density. Then it is very suitable for the analysis of spatial rainfall data like the radar derived ones, which give a finer spatial description of the precipitation fields. Radar data are, in fact, compatible with the cell organization of the PRAISEST spatial domain.

The proposed model also seems appropriate for coupling with meteorogical models in order to realize a Bayesian approach to rainfall nowcasting.

References


A space-time generator for rainfall nowcasting: the PRAISEST model

P. Versace et al.
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.


Table 1. Example of estimated parameters set.

<table>
<thead>
<tr>
<th>event</th>
<th>$p$ (-)</th>
<th>$1/\alpha_H$ (mm)</th>
<th>$\beta_H$ (-)</th>
<th>$1/\alpha_W$ (mm)</th>
<th>$\beta_W$ (-)</th>
<th>$1/\alpha_Z$ (mm)</th>
<th>$\beta_Z$ (-)</th>
<th>$\theta$ (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(+, +, +)</td>
<td>0.062</td>
<td>1.27</td>
<td>0.76</td>
<td>1.25</td>
<td>0.81</td>
<td>1.10</td>
<td>0.75</td>
<td>1.94</td>
</tr>
<tr>
<td>(+, +, •)</td>
<td>0.008</td>
<td>0.81</td>
<td>0.66</td>
<td>0.83</td>
<td>0.73</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(+, •, +)</td>
<td>0.013</td>
<td>0.61</td>
<td>0.66</td>
<td></td>
<td>0.61</td>
<td>0.65</td>
<td>1.75</td>
<td></td>
</tr>
<tr>
<td>(•, +, +)</td>
<td>0.034</td>
<td></td>
<td>0.46</td>
<td>0.64</td>
<td>0.44</td>
<td>0.56</td>
<td>1.16</td>
<td></td>
</tr>
<tr>
<td>(+, •, •)</td>
<td>0.005</td>
<td>0.46</td>
<td>0.63</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(•, +, •)</td>
<td>0.028</td>
<td></td>
<td>0.45</td>
<td>0.59</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(•, •, +)</td>
<td>0.107</td>
<td></td>
<td>0.26</td>
<td>0.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 1. (a) Implicit Scheme; (b) Explicit scheme.
Fig. 2. “Chess-Board” algorithm.
Fig. 3. (a) Location of Basilicata and Calabria regions; (b) Rain gauge network; (c) Discretization of spatial domain.
Fig. 4. Evaluation of temporal memory extension.
Fig. 5. Sample directional spatial correlograms for the implicit scheme.
Fig. 6. Mapping of the parameter $\theta^{(\cdot,\cdot,\cdot)}$ for the Implicit scheme.
Fig. 7. 1st February 1998: 1st hour of simulation.
Fig. 8. 1st February 1998: 2nd hour of simulation.
Fig. 9. 1st February 1998: 3rd hour of simulation.
Fig. 10. 24th November 1999: 1st hour of simulation.
**Fig. 11.** 24th November 1999: 3rd hour of simulation.
Fig. 12. 24th November 1999: 6th hour of simulation.