Predicting rainfall fields from lightnings records: a hierarchical Bayesian approach

Edmondo Di Giuseppe[1], Giovanna Jona Lasinio[1],
Massimiliano Pasqui[3] and Stanislao Esposito[4]

Bayesian Young Statisticians Meeting (BAYSM), Milan June, 5-6, 2013
Paper no. 9

[1] University of Rome La Sapienza-Department of Statistical Sciences, Rome, Italy
`edmondo.digiuseppe@uniroma1.it`
`giovanna.jonalasinio@uniroma1.it`
[3] National Research Council-Institute of Biometeorology, Rome, Italy
`m.pasqui@ibimet.cnr.it`
[4] Consiglio per la ricerca e la sperimentazione in agricoltura- Research Unit for Climatology and Meteorology Applied to Agriculture, Rome, Italy
`stanislao.esposito@entecra.it`

Abstract

Mixed Models (linear and non linear) belongs to a class of models in which some of the effects are *fixed* and some are *random*, formalization of these models is easily achieved in a hierarchical Bayesian framework. Here we propose a space-time mixed model to link rain measures and lightnings counts in a given area of North-central Italy.
Keywords: lightnings; rainfall fields; space-time gaussian process; bayesian geostatistics.

1 Introduction

In this paper we aim at formulating a model for predicting the 15-minutes cumulated precipitation at unknown locations given lightnings counting. In particular, we assume that the cumulated precipitation at time t in cell p of a $10x10$ km regular grid is generated by a fixed component related to lightnings and a random $\mathbf{W}$ term structured in space and time. We refer to events during *convective storms*. We use lightnings records in the fixed component of the model. The study area is located in Central Italy, we analyze an event of 68 time units (May 9, 2006). The database is composed of lightnings records (instant-point fields) cumulated over a grid with $10 \times 10km$ cells and 179 rain gages. When two or

more rain gages belong to the same grid cell we take their median over the cell ending with 111 spatial measurements at each time point.
Data are affected by several problems, on one hand a very large number of zero values is recorded, on the other hand the rain gages precision (about 0.2mm) implies an almost discrete measurement of cumulated rain as shown in Table 1.

$[0, 0.2)$	$[0.2, 0.4)$	$[0.4, 0.6)$	$[0.6, 0.8)$	$[0.8, 1)$	≥ 1
9328	861	353	258	199	1173
$\lambda_0 = \log(0.1)$	$\lambda_1 = \log(0.3)$	$\lambda_2 = \log(0.5)$	$\lambda_3 = \log(0.7)$	$\lambda_4 = \log(0.9)$	

Table 1: Frequency distribution of observed values and discretization of the latent rainfall field.

2 The model

Let $X(t,p)$ be the latent rainfall field at cell p and time t. $L_{t,p}$ denotes the number of lightnings in cell p at time t. Given the partially discrete nature of the dataset, following [6] and [5], we discretize the latent process $X(t,p)$ below 1mm assuming that there exists five values $\lambda_i, i = 0, \ldots, 4$ described in Table 1, that occurs with positive probability whenever $X(t,p)$ belongs to one of the interval reported in the same table.
Let $Y(t,p)$ be the latent rainfall field on the log scale. $Y(t,p)$ is modeled as the sum of a fixed effect and a space-time random effect $\boldsymbol{W}$:

$$y(t,p) = \mu(t,p) + w(t,p) + \epsilon(t,p) \tag{1}$$

where $\mu(t,p)$ is as in Eq. 2, $w(t,p)$ is the (t,p) element of $\mathbf{W}$ a separable space-time random field such that $w(t,p) = T(t) + S(p)$ with $T(t) = \alpha T(t-1) + \eta(t)$, $\eta(t) \sim N(0, \sigma_\eta^2)$ and

$$\mathbf{S} \sim MN(\mathbf{0}, \sigma_s^2(\mathbf{I} - \rho_s \mathbf{C})^{-1})$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{C}$ an adjacency matrix describing the spatial neighborhood structure. $\epsilon(t,p) \sim N(0, \tau^2)$ are independent, identically distributed random variables.
The fixed component of the model relates precipitations and lightnings starting from the well known Tapia-Smith-Dixon relation [7]. More precisely, we assume that the number of lightnings in cell p depends on the number of lightnings occurring in neighbouring cells. We adopt a queen neighbouring structure [2] and $\omega_{i,p} = \dfrac{L_{i,p}}{L_p} + \frac{1}{8} * \dfrac{L_{i,N_p}}{L_p}$ are the corresponding spatial weights where $L_{i,N_p} = \sum_{P_s \in N_p} L_{i,P_s}$ and $L_p = \sum_{i=1}^{T} \left(L_{i,p} + L_{i,N_p}\right)$. Moreover, we assume that the number of lightnings at time t is a function of storms propagation speed

V with two different parameterizations depending on the stage of the event. In fact, the life of lightnings' pattern inside a rainfall convective event is composed of 3 stages: *Charging phase* (C), *Mature state* (M) and *Dissipating phase* (D). Consequently, the event duration interval can be partitioned into $[t_0, T_C)$, $[T_C, T_M)$ and $[T_M, T]$. Thus, the fixed effect can be described as follows:

$$\begin{aligned} \mu(t,p) &= \log\left(C * \sum_{i=1}^{T} L_{i,p} * \left(\exp\left\{-\frac{(a+bV)}{A_p^{1/2}}|t-T_i|^2\right\} I_{[T_C,T]}(t)\right.\right. \\ &+ \left.\left.\exp\left\{-\frac{(a+bV)}{A_p^{1/2}}|t-T_i|\right\} I_{[0,T_C]}(t)\right) + C * \sum_{i=1}^{T} \omega_{i,p}\right) \end{aligned} \tag{2}$$

where C is a mass-to-volume conversion factor which is linked to the rainfall lightning ratio [7], $I_{[\tilde{a},\tilde{b}]}(\cdot)$ is the indicator function of the time interval $[\tilde{a},\tilde{b}]$ and A_p is the area of a single cell. Here, t is a general time point and T_i is the observed time.

Priors specification complete the model. The model is implemented in JAGS [4]. First results are encouraging.

References

[1] S. Banerjee, B.P. Carlin, A.E. Gelfand **Hierarchical modeling and analysis for spatial data**. *Chapman and Hall/CRC*; 2004.

[2] N.A.C. Cressie. **Statistics for Spatial Data**. *Wiley-Interscience,revised edition edition*; January 1993.

[3] P.J. Diggle, J.A. Tawn, R.A. Moyeed. **Model-based geostatistics**. *Journal of the Royal Statistical Society. Series C*; 1998; 47(3); pp. 299-350.

[4] M. Plummer. **JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling**, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, March 20-22, Vienna, Austria. ISSN 1609-395X. 2003.

[5] G. Jona Lasinio, S.K. Sahu, K.V. Mardia. **Modeling rainfall data using a Bayesian Kriged-Kalman model**, *Bayesian Statistics and its application*, Anamaya Publisher, New Delhi, India; 2007.

[6] S. K. Sahu, G. Jona Lasinio, A. Orasi, K.V. Mardia. **A comparison of spatio-temporal Bayesian model for reconstruction of rainfall fields in cloud seeding experiment**. *Journal of Mathematics and Statistics*; 2005; 1(4); pp. 272-280.

[7] A. Tapia, J.A. Smith, M. Dixon. **Estimation of convective rainfall from lightning observations**. *J. Appl. Meteor.*; 1998; 37; pp. 1497-1509.