A NEW RANDOMISED POISSON CLUSTER MODEL FOR RAINFALL IN TIME

General Assembly of the European Geosciences Union, Vienna, Austria, 25-29 April 2005

Session HS38: *Stochastic rainfall models and their use in hydrological modeling*

C. Derzekos (1), D. Koutsoyiannis (1), C. Onof (2)

(1) National Technical University of Athens, Greece (2) Imperial College London, U.K.

(dk@itia.ntua.gr)

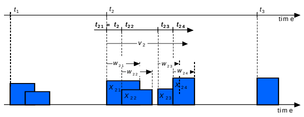
1. Introduction

The analysis and simulation of rainfall time series on fine time scales require the use of special types of stochastic models. This necessity is justified by the intermittent character of rainfall at these time scales. Among the successful model types are point process models. The purpose of the present study is to examine the behavior of a new randomized Poisson cluster model for rainfall in time. The new model is free to develop a negative or positive correlation between storm intensity and duration.

Two historical time series are used as case studies. The first one is from Denver airport (1949 – 1976), and the second is the National Technical University of Athens (NTUA) station data (1994 - 2003).

2. The Random Bartlett – Lewis Model (RBLM)

Two independent Poisson point processes are assumed: the first one determines the starting time, t_i, of storms, i, while the second one provides the (inner-storm) starting times, $t_{i,j}$, of cells, j, (rectangular pulses) generation.



Each group of rectangular pulses corresponds to a storm event, i.

According to this model, the intensity of each storm is independent of the storm duration.

3. The New Randomised Poisson Cluster Model

In the case of RBLM, the independence between storm intensity and duration, introduces a physical inconsistency. Additionally, this feature, seen from a mathematical point of view, is a restricting factor to the model efficiency.

The new, proposed, model is free to develop a negative or positive correlation between storm intensity and duration. All other aspects of RBLM are adopted.

This feature is assured by the introduction of the equation, $\beta_i = \kappa_1 \eta_i + \kappa_2 \eta_i^2$

($1/\beta_i$: mean distance between successive pulses, $1/\eta_i$: mean pulse duration within storm i, κ_1 and κ_2: parameters)

4. The complexity of the mathematical model

The efficiency of the new model is evaluated in terms of preserving historical rainfall statistics at different time scales – levels of aggregation. Four such levels are considered: 1, 6, 12 and 24 h.

Basic statistics include mean, variance, lag one auto-covariance and proportion dry (PDR) for each one of the selected aggregation levels.

The mathematical formulation of the new model introduces highly non-linear, relations with no closed form. The most difficult point is located in the PDR expression.

5. The non-closed PDR expression

$$E_\eta[\mu_\tau(\eta)] = \varphi \int_0^1 \int_0^1 \omega^{-1} t^{\varphi-1} \left(\frac{\nu}{\alpha-1}\right) \left[1 - (1-\omega t) e^{-\kappa_1 \omega (1-t)} \left(\frac{\nu}{\nu + \kappa_2 \omega (1-t)}\right)^{\alpha-1}\right] d\omega \, dt + \left(\frac{\nu}{\alpha-1}\right) \frac{1}{\varphi}$$

$$I_1(t) = E_\eta\left[\frac{e^{-(\kappa_1+\kappa_2\eta)(1-t)}}{\eta(\kappa_1+\kappa_2\eta+\varphi)}\right] = \int_0^{+\infty} \left\{\frac{\eta^{(\alpha-1)} \nu^\alpha e^{(-\nu\eta)}}{\Gamma(\alpha)} \left[\frac{e^{-(\kappa_1+\kappa_2\eta)(1-t)}}{\eta(\kappa_1+\kappa_2\eta+\varphi)}\right]\right\} d\eta$$

$$I_2(t) = E_\eta\left[\frac{e^{-(\kappa_1+\kappa_2\eta)(1-t)} e^{-(\kappa_1\eta+\kappa_2\eta^2+\varphi\eta)h}}{\eta(\kappa_1+\kappa_2\eta+\varphi)}\right] = \int_0^{+\infty} \left\{\frac{\eta^{(\alpha-1)} \nu^\alpha e^{(-\nu\eta)}}{\Gamma(\alpha)} \left[\frac{e^{-(\kappa_1+\kappa_2\eta)(1-t)} e^{-(\kappa_1\eta+\kappa_2\eta^2+\varphi\eta)h}}{\eta(\kappa_1+\kappa_2\eta+\varphi)}\right]\right\} d\eta$$

$$I_3(t) = E_\eta\left[\frac{e^{-(\kappa_1+\kappa_2\eta)(1-t)} e^{-(\kappa_1\eta+\kappa_2\eta^2+\varphi\eta)h}}{(\kappa_1+\kappa_2\eta+\varphi)}\right] = \int_0^{+\infty} \left\{\frac{\eta^{(\alpha-1)} \nu^\alpha e^{(-\nu\eta)}}{\Gamma(\alpha)} \left[\frac{e^{-(\kappa_1+\kappa_2\eta)(1-t)} e^{-(\kappa_1\eta+\kappa_2\eta^2+\varphi\eta)h}}{(\kappa_1+\kappa_2\eta+\varphi)}\right]\right\} d\eta$$

$$\mathrm{PDRY}(h) = \exp\left\{-\lambda\left(h + E_\eta[\mu_\tau(\eta)]\right) + \int_0^1 t^{\varphi-1}(1-t)\lambda[\varphi I_1(t) + \kappa_1 I_2(t) + \kappa_2 I_3(t)]dt\right\}$$

6. The use of a novel optimization approach

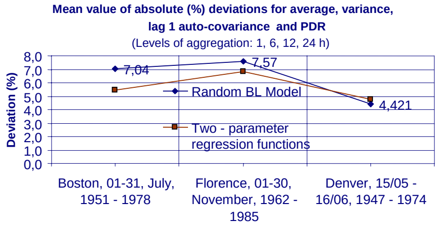
The complexity of the mathematical model, the non-closed relations and the presence of many local optima require the use of a direct search (global) optimization approach.

A new optimization algorithm is used, based on the simulated annealing method, proposed by Vanderbilt and Louie (1983). The (1st order) Markovian character of the method is altered to a "weight centered" generation mechanism for the production of new moves.

7. Decomposition of the optimization problem

The decomposition of the optimization model leads to convenient initial starting points.

Standard two-parameter regression functions are used for each group of statistical values. The optimal results provide a semi-empirical criterion (indication of the model efficiency), prior to the main optimization procedure.



Data taken from Isham *et al.*, 1990

This feature may lead to a prior selection of a homogeneous hydrologic period, for which the model can be applied successfully.

Additionally, the proper selection of the historical rainfall statistics needed to solve the non-linear system of n independent equations with n unknown parameters, can be based on this preliminary study.

8. Further simplifications

Further simplifications lead to the reduction of the unknown model parameters, from seven to six. The selection of the dependent parameter can be based either on the methodology of the LCR (Lee–Cristensen–Rudd) algorithm or on previous sensitivity studies for the RBLM (Isham *et al.*, 1990).

A NEW RANDOMISED POISSON CLUSTER MODEL FOR RAINFALL IN TIME

C. Derzekos (1), D. Koutsoyiannis (1) , C. Onof (2)

(1) National Technical University of Athens, Greece (2) Imperial College London, U.K.

Session HS38: *Stochastic rainfall models and their use in hydrological modeling*

9. The model parameters κ_1 and κ_2

Given the equation $\beta_i = \kappa_1 \eta_i + \kappa_2 \eta_i^2$, the positive character of the random variables β_i and η_i implies the following constraint:

$$\kappa_1 + \kappa_2 \eta_i > 0$$

Correspondingly, this constraint implies the following possibilities:

κ_1 & $\kappa_2 > 0$, Negative correlation between storm intensity and duration

$\kappa_1 > 0$ & $\kappa_2 < 0$, Positive correlation between storm intensity and duration

$\kappa_2 = 0$, Uncorrelated storm intensity and duration (the case of RBLM)

Mean "live" storm duration: $= (\varphi \eta_i)^{-1}$

Mean storm intensity: $= \mu_x [\varphi + (\kappa_1 + \kappa_2 \eta_i)]$

(φ, μ_x, κ_1, κ_2: constant parameters throughout the storm generation process)

10. Case Studies

i. Denver International Airport station data, 15/05–16/06, 1947–1974 *

ii. National Technical University of Athens station data, 15/12–15/01, 1994–2003

* Previous studies regarding RBLM and for the Denver station data have been conducted by

- Rodriguez-Iturbe et al. (1988)
- Velghe et al. (1994)

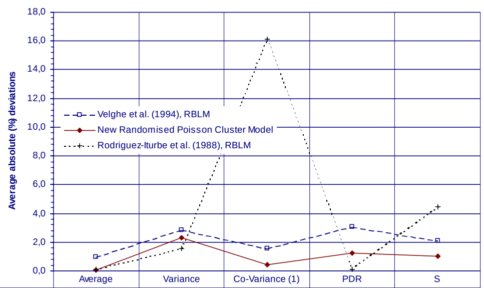
11. The Denver case

The results of the New Randomised Poisson Cluster Model are compared to those obtained for RBLM according to the previously mentioned studies.

(Our optimization results for RBLM are virtually identical to those obtained by Velghe *et al.*, 1994)

12. Results obtained for the Denver station data

Comparison between the RBLM and the New Model for the Denver case



	Average	Variance	Co-Variance (1)	PDR	S
Velghe et al. (1994), RBLM	0,9547	2,8443	1,5445	3,0397	2,0958
New Randomised Poisson Cluster Model	0,0400	2,2862	0,4281	1,2274	0,9954
Rodriguez-Iturbe et al. (1988), RBLM	0,0880	1,5372	16,1073	0,0962	4,4572

The y-axis values correspond to mean absolute (%) deviations from historical data, per solution – model and per statistical category (average, variance, lag one autocovariance, PDR), for four levels of aggregation (1, 6, 12 and 24 h).

"S" denotes the mean absolute (%) deviation of all statistical quantities (16 in total) and indicates the superiority of the new randomized model, in comparison to the RBLM.

13. The New Model's parameters for the Denver data

For the Denver Airport station data and the New Randomised Poisson Cluster Model the following optimal parameter values are determined:

$\lambda = 0{,}016$, $\nu = 0{,}303$, $\kappa_1 = 0{,}640$, $\mu_x = 3{,}921$, $\alpha = 3{,}089$, $\varphi = 0{,}060$, $\kappa_2 = -0{,}019$

The negative value of the parameter κ_2 implies a positive correlation between storm intensity and duration.

The simulation results confirm the validity of the resulting parameter values.

14. The Athens case

In the case of the Athens data, the new model also yields a better approximation of the historical statistics (in comparison to the RBLM).

However, in simulation mode, it did not provide any improvement due to unacceptable ratio of negative parameter values (The negative κ_2 value which was determined, yields an unacceptable ratio of β_i values).

As a result, RBLM, is preferable to the new model in the Athens case.

15. Conclusions

1. The New Randomised Poisson Cluster Model is free to develop a negative, zero or positive correlation between storm intensity and duration.
2. Zero values of the model parameter κ_2 yield storm intensity independent of the storm duration. In this case the new model reduces to RBLM.
3. In the Denver case, the optimal κ_2 value is negative, which implies a positive correlation between storm intensity and duration. This result was acceptable, given that the negative κ_2 did not yield negative β_i values (except in 0,3% of cases, which were rejected in simulation). The simulation results confirm the validity of this solution.
4. In this case, the solution of the new model is improved by about 50% (on the average), in comparison to the RBLM reported solutions.
5. In the Athens case, the optimization procedure also yields a negative κ_2 and a better approximation of the historical statistics (in comparison to the RBLM). However, this negative value of κ_2 was unacceptable, since it produced an unacceptable percentage of negative β_i values in simulation mode.

References

Isham S., D. Entekhabi and R. L. Bras, Parameter estimation and sensitivity analysis for the modified Bartlett – Lewis rectangular pulses model of rainfall, *Journal of Geophysical Research*, 95, 2093 - 2100, 1990.

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

Vanderbilt D. and S. G. Louie, A Monte Carlo simulated annealing approach to optimization over continuous variables, *Journal of Computational Physics*, 56, 259 - 271, 1983.

Velghe T., P. A. Troch, F. P. De Troch and J. Van de Velde, Evaluation of cluster-based rectangular pulses point process models for rainfall, *Water Resources Research*, 30, 2847 - 2857, 1994.