# Stochastic point process model for fine-scale rainfall time series

N. I. Ramesh and R. Thayakaran

School of Computing and Mathematical Sciences
University of Greenwich
Old Royal Naval College
Park Row, Greenwich, London SE10 9LS, UK
(e-mail: N.I.Ramesh@greenwich.ac.uk)

**Abstract.** A stochastic point process model, which is constructed from a class of doubly stochastic Poisson processes, is proposed to analyse point rainfall time series observed in fine sub-hourly time scales. Under the framework of this stochastic model, rain cells arrive according to a Poisson process whose arrival rate is governed by a finite-state Markov chain. Each cell of the point process has a random lifetime during which instantaneous random depths (pulses) of rainfall bursts occur as another Poisson process. The structure of this model enables us to study the variability of rainfall characteristics at small time intervals. The covariance structure of the pulse occurrence process is studied. Second-order properties of the time series of cumulative rainfall in discrete intervals are derived to model 5-minute rainfall data, over a period of 48 years, from Germany. The results show that the proposed model is capable of reproducing rainfall properties well at various sub-hourly resolutions.
**Keywords:** Doubly Stochastic Poisson process, Fine-scale rainfall, Point process, Stochastic models, Rainfall pulse.

## 1 Introduction

Stochastic point process models for rainfall have been studied extensively by many authors over the years. Much of the work has focused on models based on Poisson cluster processes (Rodriguez-Iturbe et al, 1987 [11], Cowpertwait 1994 [3], Onof 1994 [5], Chandler 1997 [1]) utilizing either the Neyman-Scott or Bartlett-Lewis processes. Rainfall models based on Markov process have also been considered by some authors (Smith and Karr 1983 [12], Ramesh 1998 [9], Onof et al 2002 [6], amongst others). However, the majority of the literature on this topic has concentrated on modelling rainfall data recorded at hourly or higher aggregation level. In some hydrological applications there is a need to reproduce rainfall time series at much smaller aggregation level. There has been some work lately on modelling fine-scale rainfall data using point process models. Cowpertwait et al 2007 [2] developed a Bartlett-Lewis pluse model to study fine-scale rainfall structure whereas Ramesh et al, 2011 [10] considered a class of doubly stochastic Poisson processes to study fine-scale rainfall intensity using rainfall bucket tip time series.

In this paper, following the approach suggested in Cowpertwait et al 2007 [2], we develop a simple point process model based on a doubly stochastic Poisson process to analyse rainfall time series collected at sub hourly fine-scale resolution. Expressions for the second-order properties of the accumulated rainfall in disjoint intervals are derived. The proposed model is fitted to 48 years of 5-minute rainfall time series from Germany. The results show that the model is capable of reproducing rainfall properties well at various sub-hourly resolutions.

## 2 Model framework

We shall start with a brief description of the doubly stochastic Poisson processes (DSPP), as the model we propose is derived from a special class of this process. A DSPP is a point process where the arrival rate of a Poisson process itself becomes a stochastic process. A special class of tractable DSPP emerges when the arrival rate of the point process is governed by a finite-state irreducible Markov chain. This process is also called a Markov-modulated Poisson process (MMPP), see for example, Ramesh 1995 [8] amongst others. The model we propose in this paper, to study fine-scale rainfall time series, is based on this class of DSPP.

Suppose that the rain cells arrive according to a DSPP on two states where the arrival rate is switching between the high intensity ($\phi_2$) and low intensity ($\phi_1$) states at random times controlled by the underlying Markov chain that has transition rates $\lambda$ (for $1 \to 2$) and $\mu$ (for $2 \to 1$). Each rain cell has a random lifetime of length $L$ and a cell originated at time $T_i$ terminates at time $T_i + L_i$. The cell lifetimes $L_i$ are taken to be independent and exponentially distributed with parameter $\eta$. During the lifetime of each cell, $[T_i, T_i + L_i)$, instantaneous random pulses of rainfall at times $T_{ij}$ occur according to another Poisson process at rate $\xi$. The process of pulse arrival terminates with the cell lifetime. Hence each cell of the DSPP generates a series of pulses during its lifetime and associated with each pulse is a random rainfall depth, $X_{ij}$, and therefore the process $\{T_{ij}, X_{ij}\}$ becomes a marked point process (Cox & Isham, 1980 [4]). In our derivation of model properties in Section 3, we treat the pulses in distinct cells as independent but allow those within a single cell to be dependent. We refer this model as the doubly stochastic pulse model (DSP).

## 3 Covariance structure of pulse arrival process

As the properties of the pulse arrival process are functions of those of the cell arrival process, we shall first see the properties of the cell arrival process. The second-order properties of the two state DSPP can be obtained as functions of the parameters $\{\lambda, \mu, \phi_1, \phi_2\}$ (Ramesh 1998 [9]). The mean arrival rate of this cell arrival process $M(t)$ is written as $E(M(t)) = m = \frac{\lambda\phi_2+\mu\phi_1}{\lambda+\mu}$ and the covariance density of $M(t)$, for $t > 0$, is given by

$$c^M(t) = Ae^{-(\lambda+\mu)t}, \qquad \text{where } A = \frac{\lambda\mu}{(\lambda+\mu)^2}(\phi_1 - \phi_2)^2. \tag{1}$$

This shows that the covariance of the cell arrival process decays exponentially with time. We shall now study the covariance structure of the pulse arrival process and focus our attention on deriving an expression for its covariance density which will then be used in the derivation of the statistical properties of the aggregated rainfall process in Section 4.

In this DSP model framework, the cell lifetimes $L_i$ are assumed to follow exponential distribution with parameter $\eta$ and therefore we have $E(L_i) = \frac{1}{\eta}$. Let $N(t)$ be the counting process of pulse occurrences from all cells. If a cell is active then it generates a series of instantaneous pulses at Poisson rate $\xi$ during its lifetime and therefore the mean number of pulses per cell is $\frac{\xi}{\eta}$. Hence the mean arrival rate of pulses is $E(N(t)) = \frac{m\xi}{\eta}$.

To derive an expression for the covariance density of this DLP process, we first studied the product density of the point process (Cox & Isham, 1980 [4]) at distinct time points. We considered two distinct pulses at time $t$ and $t+u$ ($u > 0$), which may come from the same cell or different cells, and obtained an expression for the product density which was then used to obtain the covariance density of this DLP process for $u \geq 0$ as

$$c(u) = \left(\frac{m\xi}{\eta}\right)\delta(u) + A_1 e^{-(\lambda+\mu)u} + [B_2 - B_1]\, e^{-\eta u} \tag{2}$$

where $A_1 = \left(\frac{\xi^2 A}{\eta^2-(\lambda+\mu)^2}\right)$, $B_1 = \left(\left(\frac{\xi m}{\eta}\right)^2 + \left(\frac{\xi^2 A}{\eta^2-(\lambda+\mu)^2}\right)\right)$ and $B_2 = \left(\frac{\xi^2 m}{\eta}\right)$. Here $A_1$ and $B_1$ correspond to the contribution from pulses generated by different cells whereas $B_2$ corresponds to the contribution from different pulses within the same cell, where the depths of these pulses may be dependent.

## 4 Properties of the aggregated rainfall

Although our DSP process evolves in continuous time, the rainfall data are usually available in aggregated form in equally spaced discrete time intervals. We, therefore, develop expressions for the second-order properties of the aggregated rainfall process which can be used for model fitting and assessment. Let $Y_i^{(h)}$ be the total rainfall in disjoint time intervals of fixed length $h$, for $i = 1, 2, \ldots$, then it can be expressed as

$$Y_i^{(h)} = \int_{(i-1)h}^{ih} X(t)dN(t),$$

where $X(t)$ is the depth of a pulse at time $t$. Let $E(X(t)) = \mu_x$ be the mean depth of the pulses. The mean of the aggregated rainfall can be written as

$$E\left[Y_i^{(h)}\right] = \int_{(i-1)h}^{ih} E(X(t))dN(t) = \left(\frac{m\xi}{\eta}\right)\mu_x h. \tag{3}$$

The variance and autocovariance function of the aggregated rainfall process can now be worked out using the covariance density of the pulse arrival process given in (2). In this derivation, we need to distinguish whether the pulses at time $t$ and $s$ belong to the same cell or come from different cells. This will allow us to accommodate some within-cell depth dependence. However, it is assumed that any two pulses within a cell, regardless of their location within the cell, have the same expected product moment of depths. In this set up, the variance function turns out to be

$$\begin{aligned}\mathrm{Var}\left[Y_i^{(h)}\right] &= E(X^2)\left(\frac{m\xi}{\eta}\right)h \\ &\quad + 2\,\mu_x^2 A_1\psi_1(\lambda+\mu) + 2\left[E\left[X_{ij}X_{ik}\right]B_2 - B_1\mu_x^2\right]\psi_1(\eta)\end{aligned} \tag{4}$$

where $\psi_1(\lambda+\mu) = \frac{\left[(\lambda+\mu)h-1+e^{-(\lambda+\mu)h}\right]}{(\lambda+\mu)^2}$ and $\psi_1(\eta) = \frac{\left[\eta h-1+e^{-\eta h}\right]}{\eta^2}$.

Similarly, the autocovariance function for the aggregated rainfall in two distinct intervals can be derived, by distinguishing the contributions from pulses within the same cell, and this is given below, for $k \geq 1$,

$$\begin{aligned}\mathrm{cov}\left[Y_i^{(h)}, Y_{i+k}^{(h)}\right] &= \int_{kh}^{(k+1)h}\int_0^h \mathrm{cov}\left[X(s)dN(s), X(t)dN(t)\right] \\ &= \mu_x^2 A_1\psi_2(\lambda+\mu) + \left[E\left[X_{ij}X_{ik}\right]B_2 - B_1\mu_x^2\right]\psi_2(\eta)\end{aligned} \tag{5}$$

where $\psi_2(\lambda+\mu) = e^{-(\lambda+\mu)(k-1)h}\frac{\left[1-e^{-(\lambda+\mu)h}\right]^2}{(\lambda+\mu)^2}$ and $\psi_2(\eta) = e^{-\eta(k-1)h}\frac{\left[1-e^{-\eta h}\right]^2}{\eta^2}$. When considering the special case where all pulse depths are independent $E\left(X_{ij}X_{ik}\right)$ can be replaced by $\mu_x^2$ in equations (4) and (5).

## 5 Model fitting and assessment

We use our DSP model to analyse 48 years (1960 - 2007) of 5-minute rainfall data from Dortmund (courtesy of Emschergenossenschaft/Lippeverband) in the Bochum region around the river Ems in Germany and assess how well the fitted model reproduces the properties of the rainfall over a range of sub hourly resolutions. In this work, we shall restrict ourselves to the special case where the pulse depths $X_{ij}^{'s}$ are independent random variables with an exponential distribution. Our model then has 7 parameters but we estimate the 6 parameters by the method of moment approach using the observed and theoretical values of the second-order properties. The parameter $\mu_x$ is estimated separately for each month using the sample mean by the following equation

$$\mu_x = \left(\frac{\eta}{m\xi}\right)\bar{x}$$

where $\bar{x}$ is the estimated average of hourly rainfall for each month.

The following dimensionless functions, coefficient of variation $\nu(h)$ and the autocorrelation at lag 1 $\rho(h)$ of the aggregated rainfall process, are used to estimate the remaining 6 parameters of the model.

$$\nu(h) = \frac{E\left[\left(Y_i^{(h)} - E\left[Y_i^{(h)}\right]\right)^2\right]^{1/2}}{E(Y_i^{(h)})} \quad , \quad \rho(h) = \mathrm{Corr}\left[Y_i^{(h)}, Y_{i+1}^{(h)}\right].$$

The above properties of the aggregated process at 4 different agrregation levels (at $h$=1/12, 1/3, 1/2 and 1 hour) are used in our estimation. The estimates of the functions from the empirical data, denoted by $\hat{\nu}(h)$ and $\hat{\rho}(h)$, are calculated for each month using 48 years of 5-minute rainfall series accumulated at appropriate scales. The estimated values of the model parameters $\{\hat{\lambda}$, $\hat{\mu}$, $\hat{\phi}_1$, $\hat{\phi}_2$, $\hat{\eta}$ and $\hat{\xi}\}$ for each month can be obtained by minimizing the weighted sum of squares of dimensionless functions as given below using standard routines. Here the weights are taken as the reciprocal of the variance of the empirical values of the functions calculated separately for the 48 years.

$$\sum_{h=\frac{1}{12},\frac{1}{3},\frac{1}{2},1}\left[\frac{1}{\mathrm{var}(\hat{\nu}(h))}\left(\hat{\nu}(h) - \nu(h)\right)^2 + \frac{1}{\mathrm{var}(\hat{\rho}(h))}\left(\hat{\rho}(h) - \rho(h)\right)^2\right].$$

The above objective function is minimized, using the simplex algorithm by Nelder & Mead, separately for each month to obtain estimates of the model parameters. Values of $\hat{\mu}$ are larger for summer months showing smaller mean sojourn times $(1/\mu)$ in higher rainfall intensity state. The estimates $\hat{\phi}_2$ and $\hat{\xi}$ are also higher, in general, for the summer months and show that the cell arrival rates vary from about 55 to 82 per hour whereas the pulse arrival rates range from 106 to 178 per hour throughout the year. The mean duration of cell lifetime $(1/\eta)$ falls between 1.3 to 2 minutes.

The empirical and fitted values of the mean, standard deviation, coefficient of variation and lag 1 autocorrelation of the aggregated rainfall are displayed in Figures 1 and 2. In almost all cases a near perfect fit, exact fit in some cases, was obtained for all properties. An exception is the lag1 autocorrelation at 1 hour aggregation level where there appears to be a slight underestimation. Nevertheless the differences in the correlations are less than 0.1 and the model does well at small time-scales. One point to note here is that h=10 minutes aggregation was not used in the fitting but the model has certainly reproduced all the properties well for this time-scale. This reveals that the model is capable of producing estimates of the quantities not used in the fitting which adds strength to this DSP modelling framework.

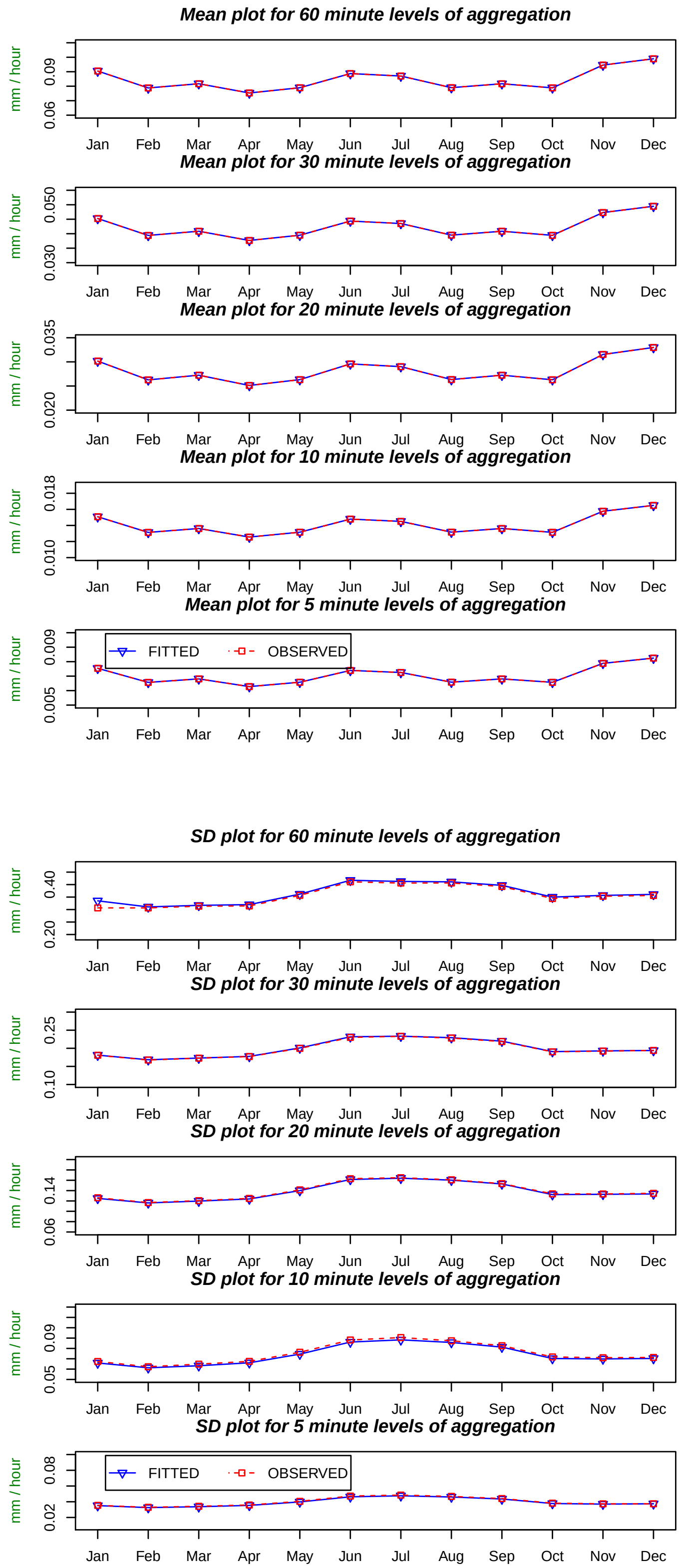


**Fig. 1.** Observed and fitted values of the mean and Standard deviation of the aggregated rainfall for DSP model at h=5, 10, 20, 30 and 60 minute aggregations.

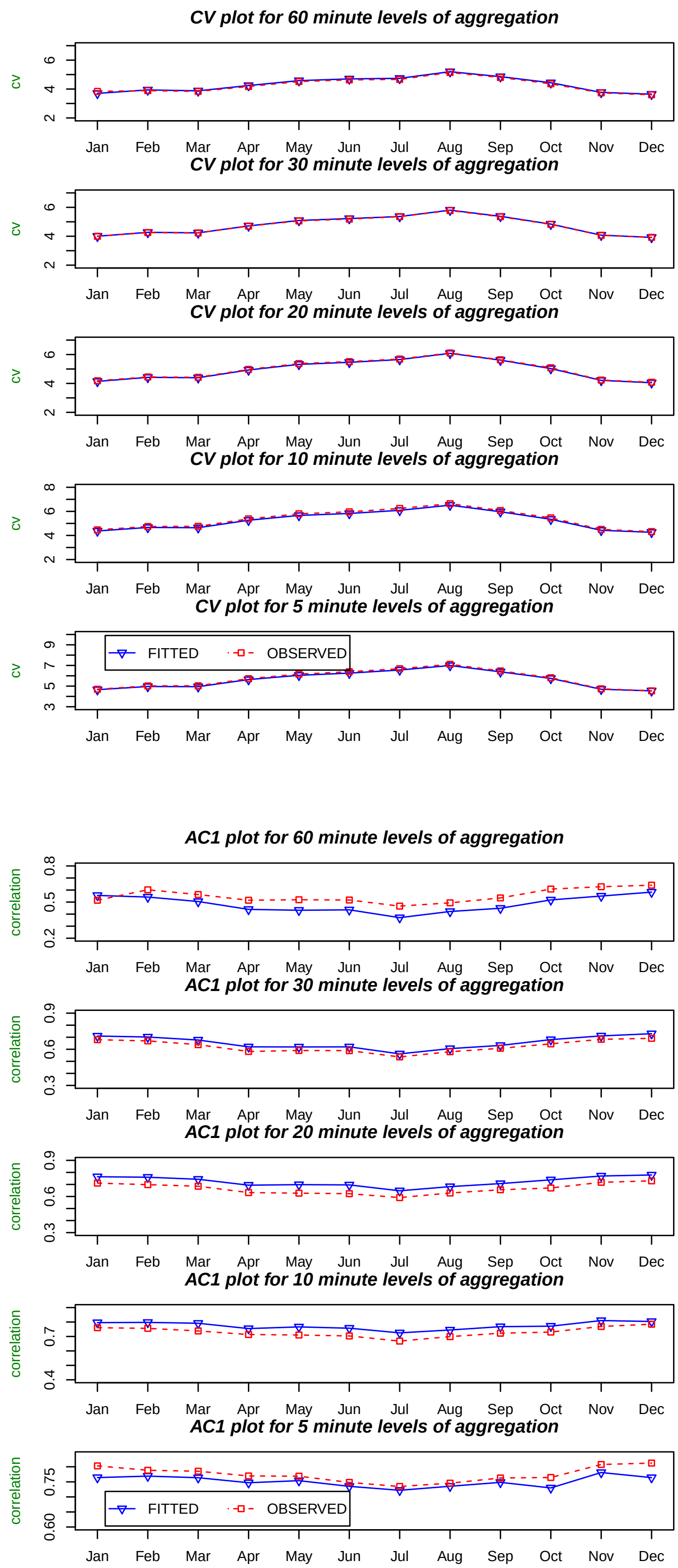


**Fig. 2.** Observed and fitted values of the coefficient of variation and autocorrelation of the aggregated rainfall for DSP model at h=5, 10, 20, 30 and 60 minute aggregations.

## 6 Conclusions and future work

A DSP model has been developed to study the properties of fine-scale rainfall time series. Second-order moment properties of the aggregated rainfall have been derived and used for model fitting and assessment. The empirical properties of the rainfall are in very good agreement with the fitted theoretical values over a range of sub hourly time scales, including those that are not used in fitting. This suggests that the model is capable of reproducing the fine scale structure of the rainfall process well and has potential application in many areas. Despite this, there is potential to develop the model further to accommodate third order moments and also to explore its capability to handle aggregations at higher levels. Further developments to explore other hydrological properties of interest are also envisaged.

## References


1. Chandler, R. E. "A Spectral Method for Estimating Parameters in Rainfall Model", *JSTOR: Bernoulli* 3(3), 301-322, (1997).
2. Cowpertwait, P.S.P., Isham, V. and Onof, C. "Point process models of rainfall: developments for fine-scale structure", *Proceedings of the Royal Society of London, Series A*, 463:2569-2587, (2007).
3. Cowpertwait, P.S.P. "A generalized point process model for rainfall", *Proc. R. Soc. Lond.*, A447, 23-37, (1994).
4. Cox, D.R & Isham, V. *Point processes*, London, UK: Chapmann and Hall (1980).
5. Onof, C. and Wheater, H.S. "Improvements to the modelling of British rainfall using a modified random parameter Bartlett-Lewis rectangular pulse model", *J Hydrol.*, Vol:157, Pages:177-195, (1994).
6. Onof, C., Yameundjeu, B., Paoli, J.P and Ramesh, N. I. "A Markov modulated Poisson process model for rainfall increments", *Water Science and Technology*, Vol 45, pp. 91-97, (2002).
7. Onof, C., Chandler, R. E., Kakou, A., Northrop, P., Wheater, H. S. and Isham, V. "Rainfall modelling using Poisson-cluster processes: a review of developments", *Stochastic Environmental Research and Risk Assessment* 14(6), 384-411, (2000).
8. Ramesh, N. I. "Statistical analysis on Markov-modulated Poisson processes", *Environmetrics* 6, 165-179, (1995).
9. Ramesh, N. I. "Temporal modelling of short-term rainfall using Cox processes", *Environmetrics* 9, 629-643, (1998).
10. Ramesh, N. I., Onof. C. and Xie, D. "Doubly stochastic Poisson process models for precipitation at fine time-scales", *Adv Water Resour* doi:10.1016/j.advwatres.2011.09.017, (2011).
11. Rodriguez-Iturbe, I., Cox, D.R and Isham,V. "Some models for rainfall based on stochastic point processes", *Proc. R. Soc. Lon. A* 410(1839), 269-288, (1987).
12. Smith, J.A. and Karr, A.F. "A point process model of summer season rainfall occurrences", *Water Resources Res.*; 19: 95-103, (1983).