CEE 599C
Homework 5
Brian Henn and Jiawei He
June 12, 2014

Rectangular Pulse Precipitation Modeling Assignment

Abstract

We used an observational hourly precipitation dataset to fit two different types of stochastic rainfall models' parameter: Poisson and a Bartlett-Lewis clustered rectangular pulse models. We were able to numerically solve for the Poisson model parameters but were not able to successfully fit parameters for the Bartlett-Lewis model; instead, a manual calibration approach was applied. The results from the Poisson simulations indicate that the models can match mean rainfall rates, though the match with rainfall rate variance and autocorrelation for these models is not quite exact due to unknown errors. The manually calibrated Bartlett-Lewis models was able to approximately match observations of mean rainfall rate, rainfall rate variance, storm inter-arrival period and rainfall rate autocorrelation. The main advantage of the Bartlett-Lewis approach would correctly simulating rainfall autocorrelation, though this benefit is offset by a less robust parameter fitting method.

Sea-Tac Precipitation Dataset

To carry out the analysis of the precipitation models, we downloaded hourly precipitation data from Seattle-Tacoma International Airport station from the NCDC web server. This source provided reasonable-looking data from the period 1965-present. We replaced all missing values with zero precipitation because it appeared that frequently no recordings were made during dry periods. The record indicated approximately 1000mm/yr of precipitation at this site which appeared reasonable.

We then extracted only the 49 Januaries from the period of record, and used these months to create a perpetual January just over four years long, or 36456 hours long. From this record, we could calculate various important statistics, such as mean precipitation rate, variance of precipitation rate, lag-1 hour autocovariance of the precipitation rate, mean dry period length, mean wet period length, mean total precipitation per wet period and mean dry/wet time fraction.

Fitting a Poisson Rectangular Pulse Model – Approach 1

Using equations 2.15 from Rodriguez-Iturbe et al. 1987, (below), we attempted to solve for the three parameters of a Poisson rectangular pulse model. These three parameters control pulse inter-arrival time, pulse duration, and pulse intensity. Because the equations are nonlinear, we used the Excel GRG (Generalized Reduced Gradient) nonlinear solver to find a solution to the three equations.

$$\left.\begin{aligned} E(Y_i^{(h)}) = h\psi\eta\rho,\ \operatorname{var}(Y_i^{(h)}) = 4\psi^2\eta\rho(h\eta - 1 + e^{-h\eta}),\\ \operatorname{cov}(Y_i^{(h)},\ Y_{i+k}^{(h)}) = 2\psi^2\rho(e^{h\eta} + e^{-h\eta} - 2)\,e^{-k\eta}\ (k \geqslant 1).\end{aligned}\right\} \tag{2.15}$$

In our first attempt, we solved for the parameters using the observed mean rainfall rate, the rainfall rate variance, and the lag-1 hour autocovariance of the observations. The solver was able to meet all three equations successfully, indicating that the calibrated model should successfully reproduce these three statistics. We refer to this solution as "Poission Solution 1" or "P1".

Fitting a Poisson Rectangular Pulse Model – Approach 2

Alternatively, Rodriguez-Iturbe et al. 1987 describe that the mean dry period between storms should be geometrically distributed and dependent only on the aggregate period and the mean storm inter-arrival parameter. Thus, we could use the observed mean dry period to estimate this parameter, and then use the equations for mean and variance to solve for the other two. In this case, we do not use the equation for lag-1 hour autocovariance to solve for the parameter set, so in theory this solution should match the dry period lengths but not the autocorrelation. Similar to the first approach, we solved these equations successfully in Excel, and this solution is referred to as "Poisson Solution 2" or "P2".

Fitting a Bartlett-Lewis Clustered Rectangular Pulse Model – Approach

Using equations (10) (11) (12) from Rodriguez-Iturbe et al. 1987, (below), we attempted to solve for the five parameters of a Bartlett-Lewis Clustered Rectangular Pulse Model. These five parameters control mean arrival rate of a storm, mean number of cells in a storm, mean intensity of a cell, mean displacement, and mean cell life span.

$$E[Y_i(h)] = h\rho\mu_c\mu_x \tag{10}$$

$$Var[Y_i(h)] = 2\rho\mu_c\{E[X^2] + \frac{\beta}{\gamma}\mu_x^2\}\frac{h}{\eta} - 2\rho\mu_c\{E[X^2] + \frac{\beta\gamma}{\gamma^2-\eta^2}\mu_x^2\}\bullet\frac{(1-e^{-\eta h})}{\eta^2} + 2\rho\mu_c\mu_x^2\frac{\beta}{\gamma^2-\eta^2}\frac{(1-e^{-\gamma h})\eta}{\gamma^2} \tag{11}$$

$$Cov[Y_i(h), Y_{i+k}(h)] = \rho\mu_c\{E[X^2] + \frac{\beta\gamma}{\gamma^2-\eta^2}\mu_x^2\}(1-e^{-\eta h})^2\bullet\frac{e^{-\eta(k-1)h}}{\eta^2} - \rho\mu_c^2\frac{\beta}{\gamma^2-\eta^2}\mu_x^2(1-e^{-\gamma h})^2 e^{-\gamma(k-1)h}\frac{\eta}{\gamma^2} \tag{12}$$

Normally, these systems are solved for the required parameters by unconstrained minimization of an objective function. These combinations range from the most basic (5 equations in the objective function to solve for the 5 parameters are used) to the more thorough (more than 5 equations in the objective function to solve for the 5 parameters are used). Only the former combinations were considered for this homework. Thus, the determination of the five parameters included the following equations:

1. hourly mean of rainfall depth ((10) cast in h = 1 hr)
2. variance of hourly rainfall depth ((11) cast in h = 1 hr)
3. lag -1 covariance of hourly rainfall depth ((12) cast in h = 1 hr)

4. lag-1 covariance of 6-hourly rainfall depth ((12) cast in h = 6 hr)
5. lag-1 covariance of 12-hourly rainfall depth ((12) cast in h = 12 hr)

The search algorithm used was the BFGS method, a gradient based method. But the result does not look good, because for different initial values, results could be different; some are not even close, and some are even unrealistic.

Simulating the Poisson Model

Now that the three parameters had been solved for, it was possible to simulate a precipitation series using the Poisson rectangular pulse model. This was done using the "rand()" pseudorandom number generator in MATLAB. A random number, drawn from an exponential distribution, was used to simulate the inter-arrival time between each pulse of rainfall. This process was repeated until the cumulative time reached the end of the simulation. Then, for each pulse time, two random numbers were drawn from two exponential distributions – one for pulse duration, and one for pulse intensity. The pulse was assumed to start at the pulse time and continue for the simulated duration with the simulated intensity.

This continuous series of discrete pulses was then aggregated to a one-hour observation timestep. This was done by considering at each hour how many pulses had fallen within the previous hour-long interval, and their intensity. Fractional pulses within an hour contributed were prorated to contribute correctly to the overall hourly accumulation. This accumulation time series was then used as the simulated rainfall record, upon which to calculate the relevant statistics for comparison with observations.

Simulating the Bartlett-Lewis Clustered Model

The Bartlett-Lewis clustered rectangular pulse model was simulated in a similar fashion to the Poisson model. Random numbers were drawn in MATLAB from an exponential distribution to simulate the inter-arrival time of storms (as opposed to pulses in the Poisson model). For each storm, an additional random number was then from an exponential distribution to determine the length of the storm.

Once the storm length was established, a third parameter was used to draw exponentially-distributed random numbers to determine the inter-arrival times of clustered pulses within the storm. Together, the storm length parameter and the pulse generation rate parameter dictate the average number of pulses per storm.

Finally, identically to the Poisson model, two final parameters were used to generate the pulse length and intensity from a random exponential distribution. Then the pulses were aggregated to hourly accumulations, again identically to the Poisson model method, and this simulation time series was used to calculate the relevant statistics.

Results: Poisson Solution 1

We show the results of the Poisson Solution 1 simulation in Figs. 1 and 2. Fig. 1 shows the entire 49-month simulation, while Fig. 2 shows only the first 15 weeks of the simulation. Qualitatively, it would appear that the simulation is approximately matching the peak rainfall intensities, as well as the average storm frequency and the like. This indicated that the model parameter solution and simulation was generally successful. However, more objective statistics for model success will be presented subsequently.

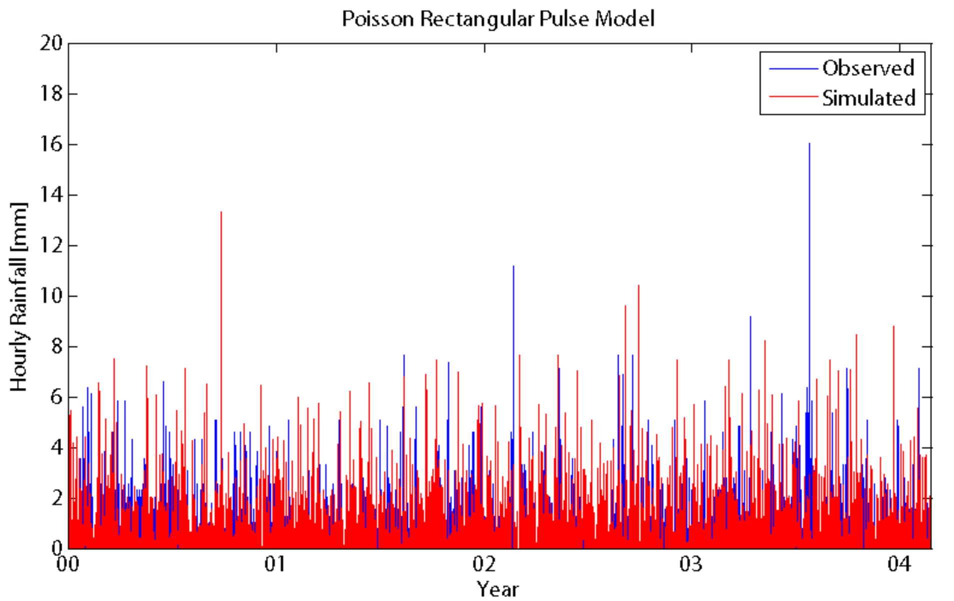

Fig. 1 – Poisson Solution 1, entire simulation.

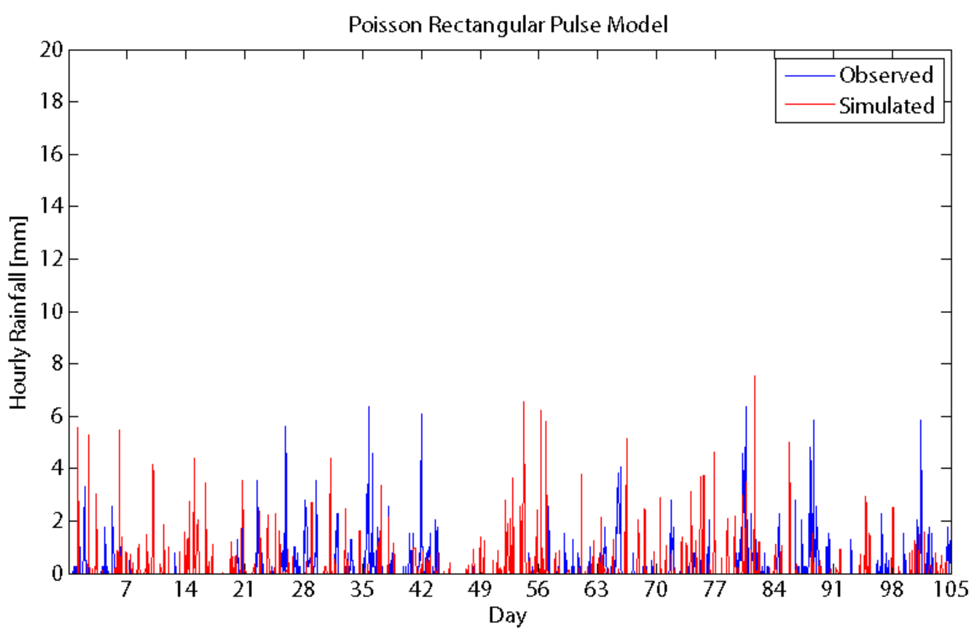

Fig. 2 – Poisson Solution 1, first fifteen weeks.

Results: Poisson Solution 2

The second Poisson solution, in which the parameters were estimated from the mean dry period and not the lag-1 hour autocovariance, is shown in Figs. 3 and 4. Again, a broadly successful simulation can be seen, though in comparison with Figs. 1 and 2, we find that this solution set produces perhaps greater peak intensities and less coherence between timesteps.

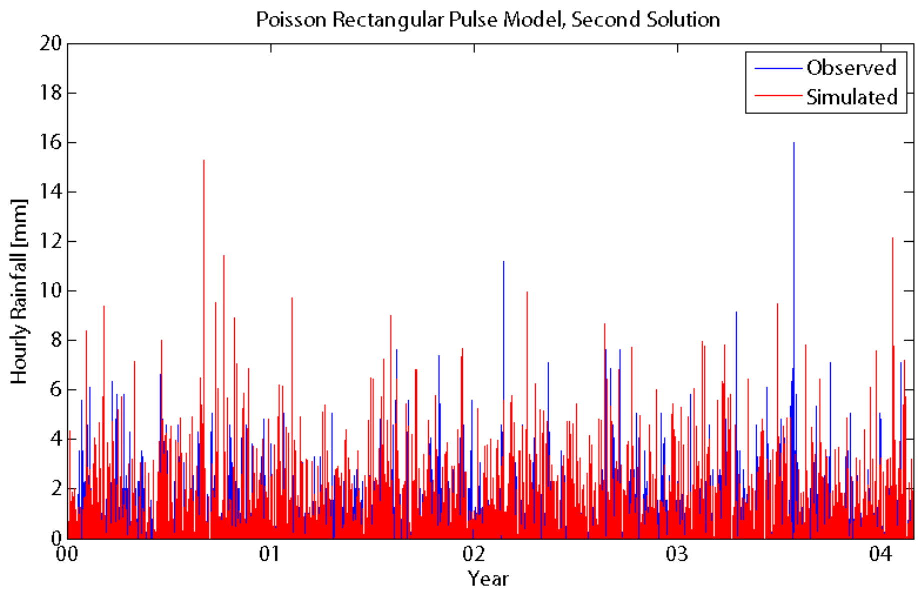

Fig. 3 – Poisson Solution 2, entire simulation.

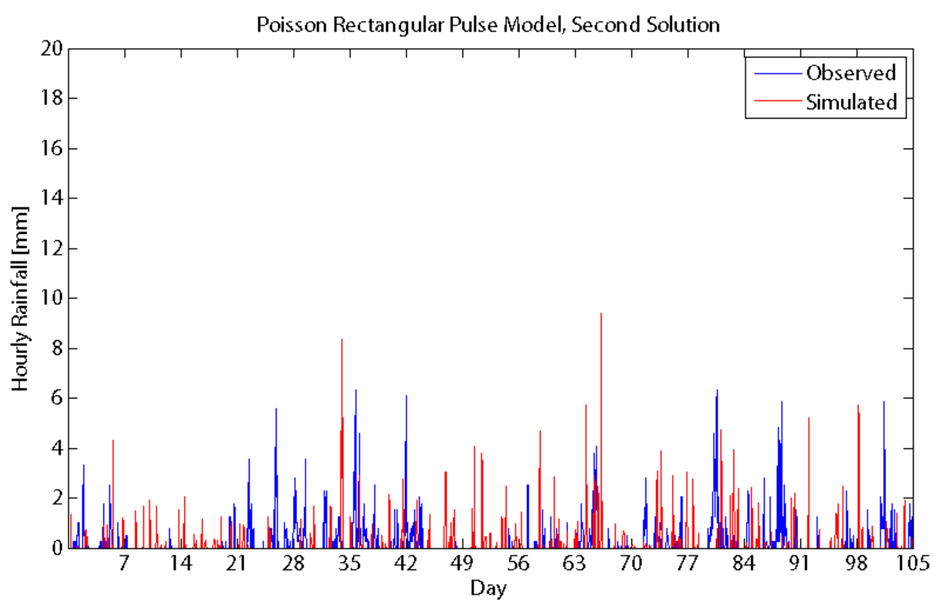

Fig. 4 – Poisson Solution 2, first fifteen weeks.

Results: Bartlett-Lewis Estimated Solution

Because we were unable to successfully solve the system of five equations to determine the five Bartlett-Lewis parameters, we present an alternative approach to model fitting. These simulations shown here were the result of a rough "calibration" by simulating various sets of model parameters, to see which gave the best fit to observed data. The initial values of the parameters were based on Table 12 in Rodriguez-Iturbe et al. 1987b, and from visual inspection of the Sea-Tac dataset. Then modifications were made to the parameters to improve the fit and various statistics. Because this was done qualitatively, exact agreement between simulated and observed statistics should not be expected.

The Bartlett-Lewis simulations are shown in Figs. 5 and 6. At a macro view (Fig. 5), the simulation appears similar to the Poisson simulations. However, at a smaller scale (Fig. 6), the clustering that is integral to this model is clearer. Periods of heavy rainfall tend to occur every few days, which more coherent dry periods in between, in comparison with Figs. 2 and 4.

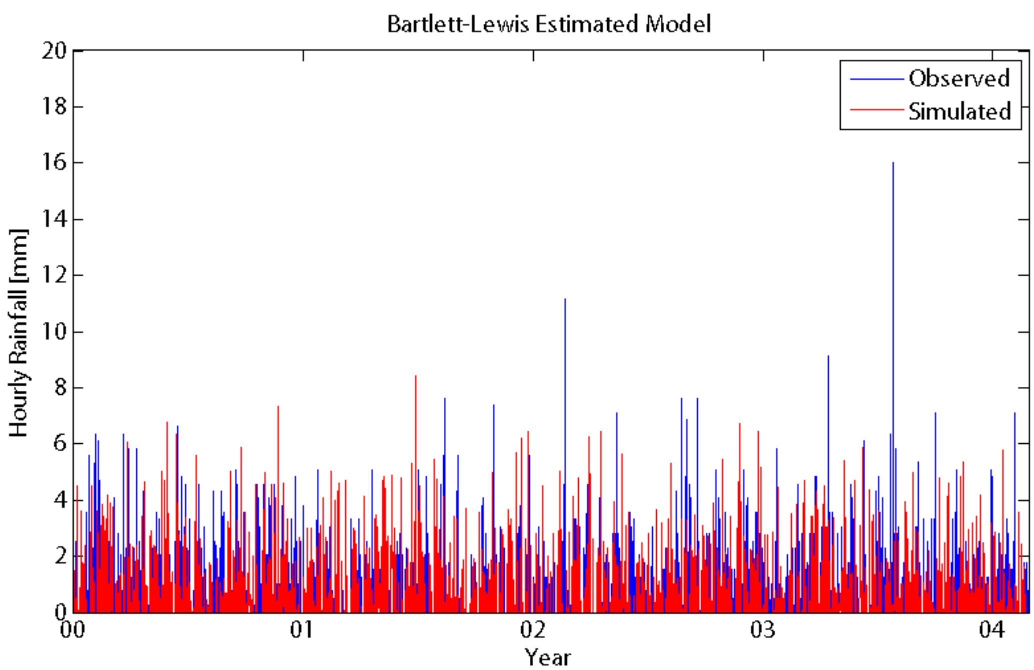

Fig. 5 – Bartlett-Lewis estimated model, entire simulation.

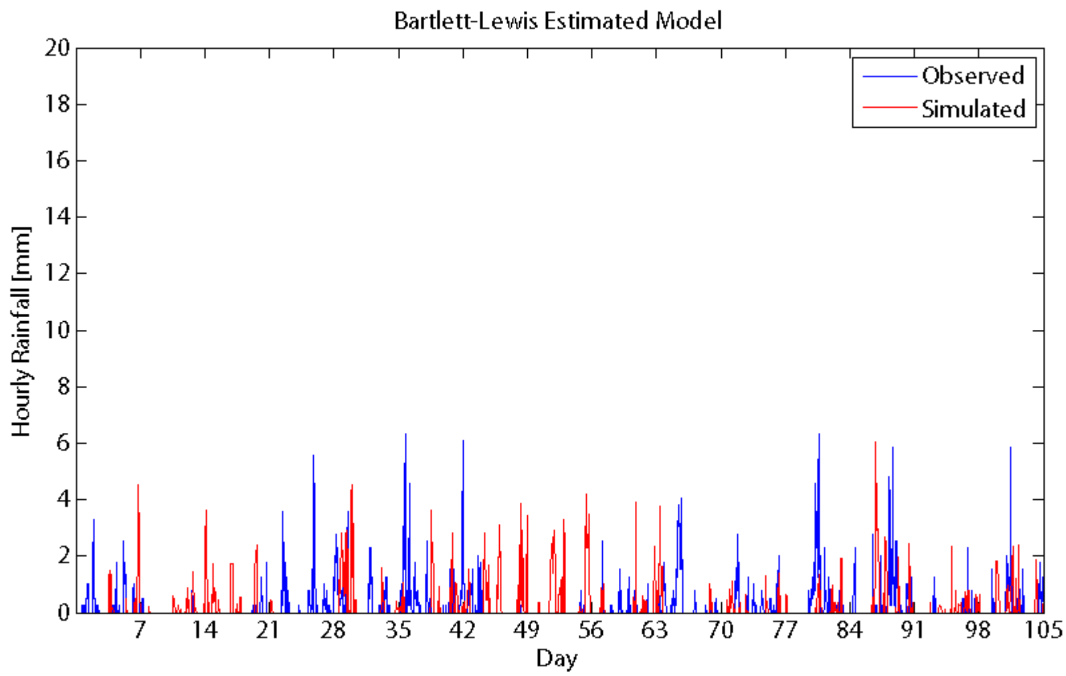

Fig. 6 – Bartlett-Lewis estimated model, first fifteen weeks.

Results: Model Comparisons to Observations and Each Other

Next, we performed a systematic comparison of the three model simulations against the observed precipitation record, and against each other. The first collection of findings from this comparison is shown in Fig. 7. The panels show comparative statistics for six major indicators of model performance:

- Mean rainfall at all times
- Rainfall rate variance at all times
- Mean dry period length
- Mean wet period length (defined as continuous hourly rainfall)
- Mean rainfall depth for wet periods
- Fraction of dry hours

In each panel, the statistics for the observations, the two Poisson solutions and the Bartlett-Lewis estimated solution simulation are shown.

First, it can be seen in the upper left panel that all four models relatively closely match the observed mean precipitation rate, although the Bartlett-Lewis simulation is less closely matched. However, the variance of the rainfall rate (bottom left panel) shows discrepancies between observations and two Poisson solutions. This is problematic because the fitting of the parameters should have matched the simulated variance exactly. It is not known why the simulated variances are higher in both Poisson simulations – presumably due to an error in our fitting and simulation routines, but we could not detect this.

Next, we can examine the mean dry periods (upper middle panel) to see how well each model simulates this behavior. The first Poisson solution underestimates the mean dry period substantially. However, this is not surprising because this statistic was not used in fitting the first Poisson solution. The second Poisson solution, which did use mean dry period in parameter fitting, does a much better job at this metric as would be expected. The Bartlett-Lewis “calibrated” simulation also manages to match this parameter fairly well.

In terms of the mean wet period (bottom middle panel), the two Poisson simulations both underestimate the observed period, while the Bartlett-Lewis solution overestimates the wet period length. This might be expected given that the Poisson rainfall pulses are random in time, while the Bartlett-Lewis pulses are explicitly clustered. In terms of mean storm depth (upper right panel), the Poisson solutions again underestimate observations, while the Bartlett-Lewis simulation overestimates observations, which is consistent with the findings in terms of the length of the wet period. Finally, each model is doing a fairly good job of simulating the dry fraction, though the Bartlett-Lewis simulation has a dry fraction which is somewhat too low.

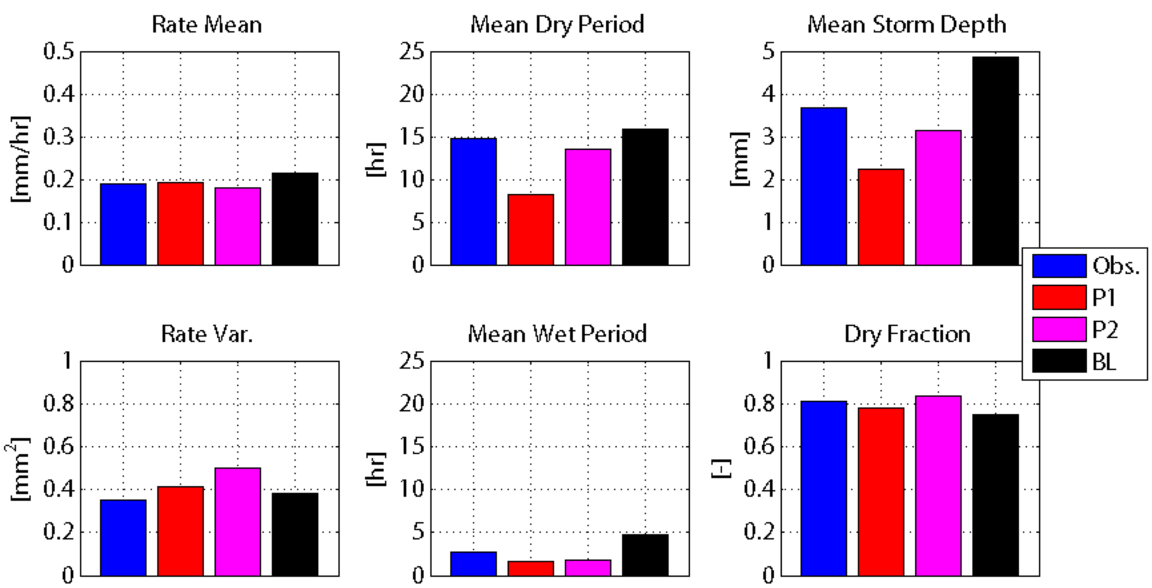

Fig. 7 – Summary of simulation statistics for the observations and the three model simulations. Upper left panel shows mean rainfall rate, bottom left panel shows rainfall rate variance, upper middle panel shows mean dry period length, bottom middle panel shows mean wet period length, upper right panel shows mean wet period total depth, and bottom right panel shows fraction of dry hours. “P1” indicates the first Poisson solution, “P2” indicates the second Poisson solution, and “BL” indicates the Bartlett-Lewis estimated solution.

Next, we present the results of the autocorrelation estimates for the observations and for each model (Fig. 8). Autocorrelation was calculated for each time series for lags of 1-12 hours. We note that the two Poisson solutions slightly underestimate lag-1 hour autocorrelation, while the Bartlett-Lewis approach somewhat overestimates lag-1 hour autocorrelation. The underestimation of lag-1 hour autocorrelation by the first Poisson solution is problematic, because this parameter set was fitted by using lag-1 hour autocovariance. Again, there seems to be a problem in the fitting and simulation of the Poisson process.

Beyond lags of 1-2 hours, the observed autocorrelation persists for some time, while the Poisson simulated autocorrelation rapidly falls to near zero. However, the Bartlett-Lewis simulation successfully matches the observed autocorrelation out to much longer time lags – as much as 6-8 hours. This would indicate that the clustered model structure is capable of simulating this aspect of rainfall much better than the independently distributed pulses in the Poisson models.

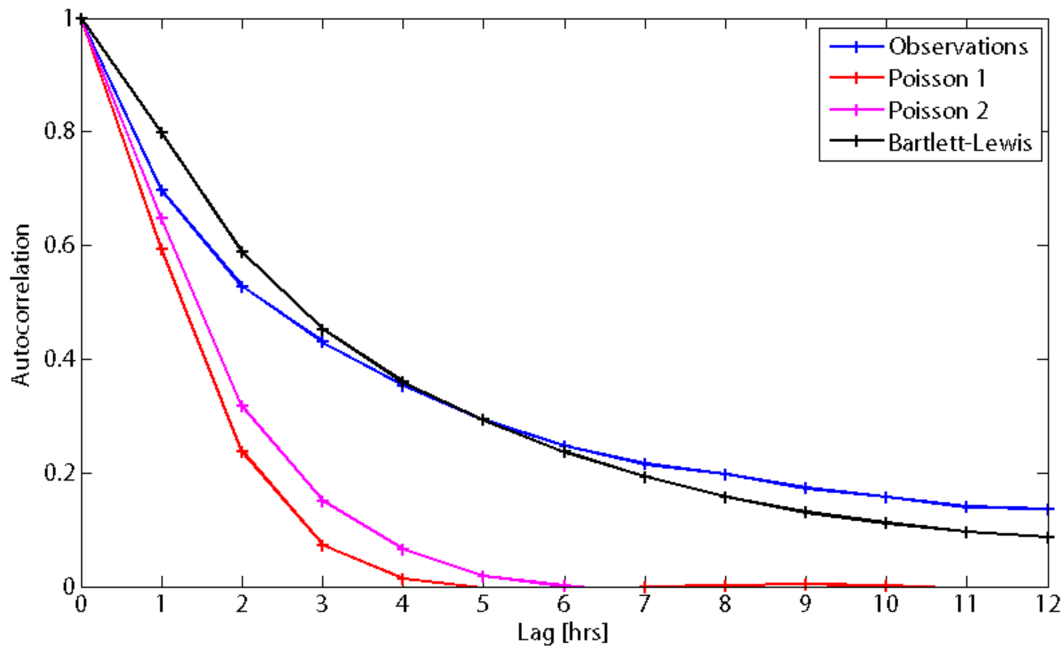

Fig. 8 – Autocorrelation for each precipitation time series.

Finally, we present the non-exceedence probability curves (CDFs) of each precipitation time series (Fig. 9), in order to show whether the model simulations are matching the distributions of the observed rainfall. To generate this plot, the Cunnane plotting position is used, and rainfall rates are log-transformed. It would appear that all three models actually are matching the non-exceedence curve of the observations fairly well, with the exception of extremely low rainfall rates which are below the minimum observable threshold. Thus, the different between the Poisson and Bartlett-Lewis models is not the distribution of precipitation intensities, but the difference in autocorrelation (Fig. 8), which is not represented by an non-exceedence probability curve (or CDF) such as Fig. 9.

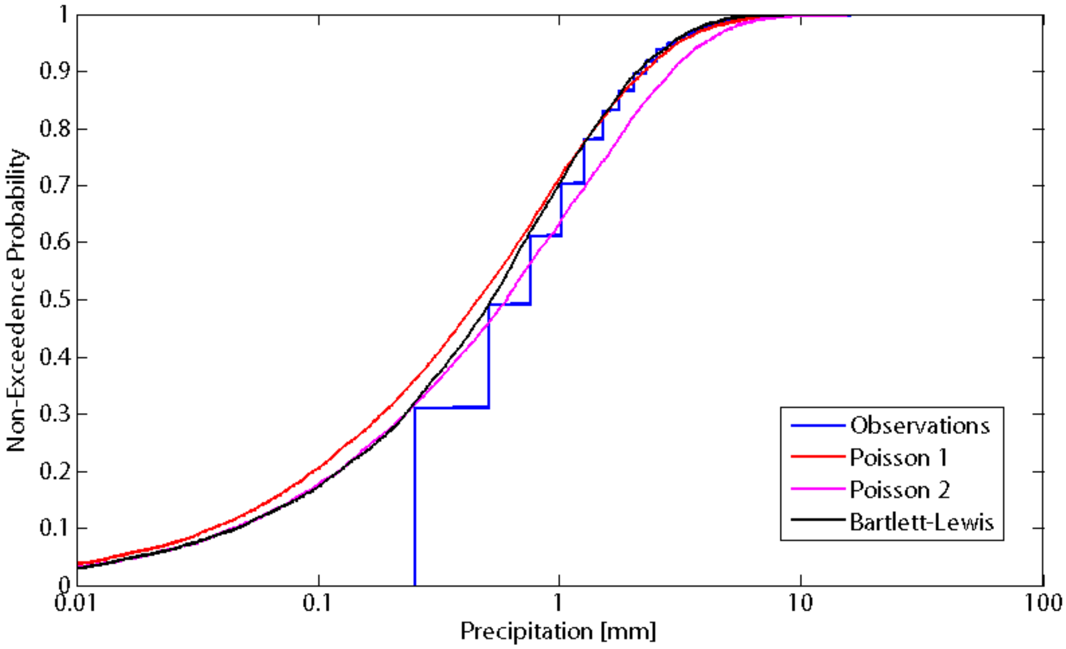

Fig. 9 – Precipitation non-exceedence probability curves each time series.

Sources

Rodriguez-Iturbe, I., D. R. Cox and V. Isham, 1987: Some models for rainfall based on stochastic point processes. *Proc. R. Soc. Lond.* A **410**, 269-288.

Rodriguez-Iturbe, I., B. Febres de Power and J. B. Valdés, 1987b: Rectangular pulses point process models for rainfall: Analysis of empirical data. *J. Geophys. Res.*, **92**, D8, 9645-9656.