CHAPTER 2

LITERATURE REVIEWS ON THE STOCHASTIC RAINFALL DISAGGREGATION

Stochastic Disaggregation of Rainfall is based on stochastic point processes. A simple stochastic point process of rainfall can be conceptualised by storms arriving in a Poisson process, and each storm is associated with a random and single rain cell of rectangular pulse with independent intensity and duration. The rain cells can overlap, and the total rainfall intensity at any instance can be given by the sum of the intensities of all active rain cells at that instance. One of the shortcomings of this type of conceptual model is that they are not capable of accounting for the temporal statistics of rainfall at different aggregation levels (Rodriguez-Iturbe et al. 1987).

Cluster based models such as Neyman-Scott and Bartlett-Lewis were introduced by Rodriguez-Iturbe et al. (1987 and 1988) as measures to overcome the shortcomings of the point processes. These cluster based rainfall models are actually based on the Poisson process with some adjustments. In the cluster based models, each storm produces a cluster of rain cells instead of one as in the simple Poisson point process, with each cell having random duration and intensity. From a statistical point of view both randomised and nonrandomised versions of the models are parametric analyses. Normally, these parameters are estimated by the method of moments. Velghe et al. (1994) have shown discrepancies in these parametric statistics at different aggregation levels.

Some researchers have suggested different approaches to improve the results of these models. Rodriguez-Iturbe et al. (1987) incorporated a high frequency jitter process to deal with more irregular rainfall traces. Onof and Wheater (1994) combined this jitter process with the 6-parameter randomised Bartlett-Lewis model and advised against seeking models with more than seven parameters for each month of the year. Gyasi-Agyei and Willgoose (1997) adopted a hybrid point rainfall model that incorporated two random processes, namely *{A(t)}* and *{Y(t)}*. The binary (wet and dry) process *{Y(t)}* was chosen as the non-randomised Bartlett- Lewis Rectangular Pulse simply because of its capacity to simulate continuous rainfall time series with fewer parameters and also due to its ability to aggregate and disaggregate into desired timescales. The intensity process *{A(t)}* was chosen as an autoregressive and autocorrelated jitter model. Cowpertwait (1991) developed the dry probability equations for use in the Neyman- Scott model with the same intention. Cowpertwait (1998) also made comparisons between these two models and showed their similarities up to certain order statistics.

There have been some initiatives to use these models in both space and time. Northrop (1998) developed a spatial-temporal model based on the Bartlett-Lewis model, while Cowpertwait (1995) derived spatial-temporal model properties based on the Neyman-Scott model. Cowpertwait (2006) proposed a fitting procedure to use the spatial-temporal Neyman-Scott model to disaggregate daily data into hourly data, to infill missing values, and to simulate data at sites where no historical data were available. To deal with the non-stationarity of the land surface, Margulis and Entekhabi (2001) proposed two temporal disaggregation models based on satellite-derived precipitation. Segond et al. (2006) introduced a spatially uniform temporal disaggregation pattern to a multi-site rain gauge network. This technique had some limitations in simulations of extreme rainfall and correlation structures. Elshamy et al. (2006) evaluated the Geng rainfall model (Geng et al. 1986) and advised of calibrating the Geng parameter estimation equations to reduce the overestimation of rainfall variability of the original model, but this technique was emphasised only on a regional basis using small datasets for temporal disaggregation of monthly rainfall to daily rainfall.

Gyasi-Agyei (1999) regionalised the Bartlett-Lewis based hybrid model (Gyasi-Agyei and Willgoose 1997, 1999) for daily rainfall disaggregation. Cowpertwait et al. (1996) used a fitting procedure for the Neyman-Scott based model to disaggregate daily rainfall. Some researchers did not take the seasonal variation of the rainfall data into account. Glasbey et al. (1995) proposed conditional simulation to calibrate the model parameters using statistics of hourly rainfall data. Their model was used to generate an archive of a large number of years of rainfall data. The best matches from the observed and the generated sequences, and the hourly data from the past events, were used to generate the disaggregated rainfall sequence. This approach did not necessarily predict the second order characteristics and dry probabilities due to the constraint put on the disaggregated sequence to follow the same trend in the mean rainfall.

Koutsoyiannis and Onof (2001) proposed a proportional adjusting procedure on a Bartlett-Lewis based model that preserved the individual observed daily totals while disaggregating rainfall into fine timescale. Koutsoyiannis et al. (2003) also discussed multivariate rainfall disaggregation schemes. Hansen and Ines (2005) advised in favour of constraining a stochastic generator to match a target monthly rainfall total while disaggregating monthly rainfall. Even though this approach does not directly manipulate rainfall frequency or intensity, it is yet to be validated for reproducing the relationships among variations of historical fine timescale rainfall totals, frequency and mean intensity.

Sivakumar (2000) and Sivakumar et al. (2001) proposed a chaotic model for temporal rainfall disaggregation based on correlation dimension method. However, this approach gave rise to much controversy on whether a deterministic chaotic process exists in the temporal evolution of rainfall (Koutsoyiannis and Pachakis, 1996; Schertzer et al. 2002). Gaume et al. (2006) established firmly that no low dimensional chaotic behaviour exists in the rainfall temporal disaggregation.

Cascade-based, fractal and multifractal approaches have been proposed for rainfall disaggregation into fine time scale (Ormsbee, 1989; Olsson, 1998; Olsson and Berndtsson, 1998; Günter et al., 2001; Molnar and Burlando, 2005). The regionalisation of parameters and preservation of second order statistics such as variance and autocorrelations need to be explored in these studies. Also the overestimation of rainfall extremes at longer durations are yet to be analysed in both canonical and micro-canonical cascade based models of disaggregation. Gaume et al. (2006) showed that neither a low dimensional chaotic model nor a multifractal multiplicative random cascade model can reproduce faithfully all the characteristics of rainfall time series. Gyasi-Agyei (2005) developed a regionalised disaggregation model based on the point process hybrid model of Gyasi-Agyei and Willgoose (1997). This model was the product of a binary chain (wet and dry sequence) model and an autocorrelated jitter as intensity process model. This model incorporated the use of repetition techniques and proportional adjusting procedure proposed by Koutsoyiannis and Onof (2001) to disaggregate daily rainfall data into hourly time scale. The model was evaluated with a 5-year time series of hourly rainfall observed at a small experimental site, and demonstrated the preservation of the multiple sub-daily time scale stochastic structure of rainfall after disaggregation. Gyasi-Agyei and Mahbub (2007) extended the original stochastic model (Gyasi-Agyei, 2005) by disaggregating daily rainfall data throughout Australia to any desired fine timescale down to 6-minute. This research will evaluate the regionalised disaggregation model (Gyasi-Agyei, 2005; Gyasi-Agyei and Mahbub, 2007) and explore its improvement to make it applicable for a broader Queensland region.

Another focus of this research would be a robust approach to the derivation of Intensity-Frequency-Duration (IFD) curves for a large region such as Queensland based on stochastic disaggregation of daily rainfall into fine timescale. Analytical IFD curves are determined from an assumed probability distribution function of either annual or seasonal maximum rainfall series (Muller et al., 2007; Leonard et al., 2007; Koutsoyiannis et al., 1998; Cowpertwait et al., 2002; Koutsoyiannis and Baloutsos, 2000). The Australian Rainfall and Runoff (ARR) have standardised the calculation of design IFD (Canterford et al., 1998) through a set of standard maps for the Australian continent. The analytical probability distribution functions are different in all approaches. Hence there is a need to derive the design IFD curves with a unifying approach. This research will evaluate the stochastic disaggregation of daily rainfall into fine timescale based on a point process model to derive the IFD curves at different fine timescale durations.

In summary, it can be seen that all of the proposed point process models for rainfall disaggregation are somewhat variants of either the Bartlett-Lewis rectangular pulse model or the Neyman-Scott model. Therefore, a detailed discussion on these two basic rainfall models is necessary. The following two sections describe in greater length both Neyman-Scott model and Bartlett-Lewis rectangular pulse model, respectively.

2.1 The Neyman-Scott Model

The schematic diagram of the Neyman-Scott Model is given in Figure 2.1 (Cowpertwait, 1991).

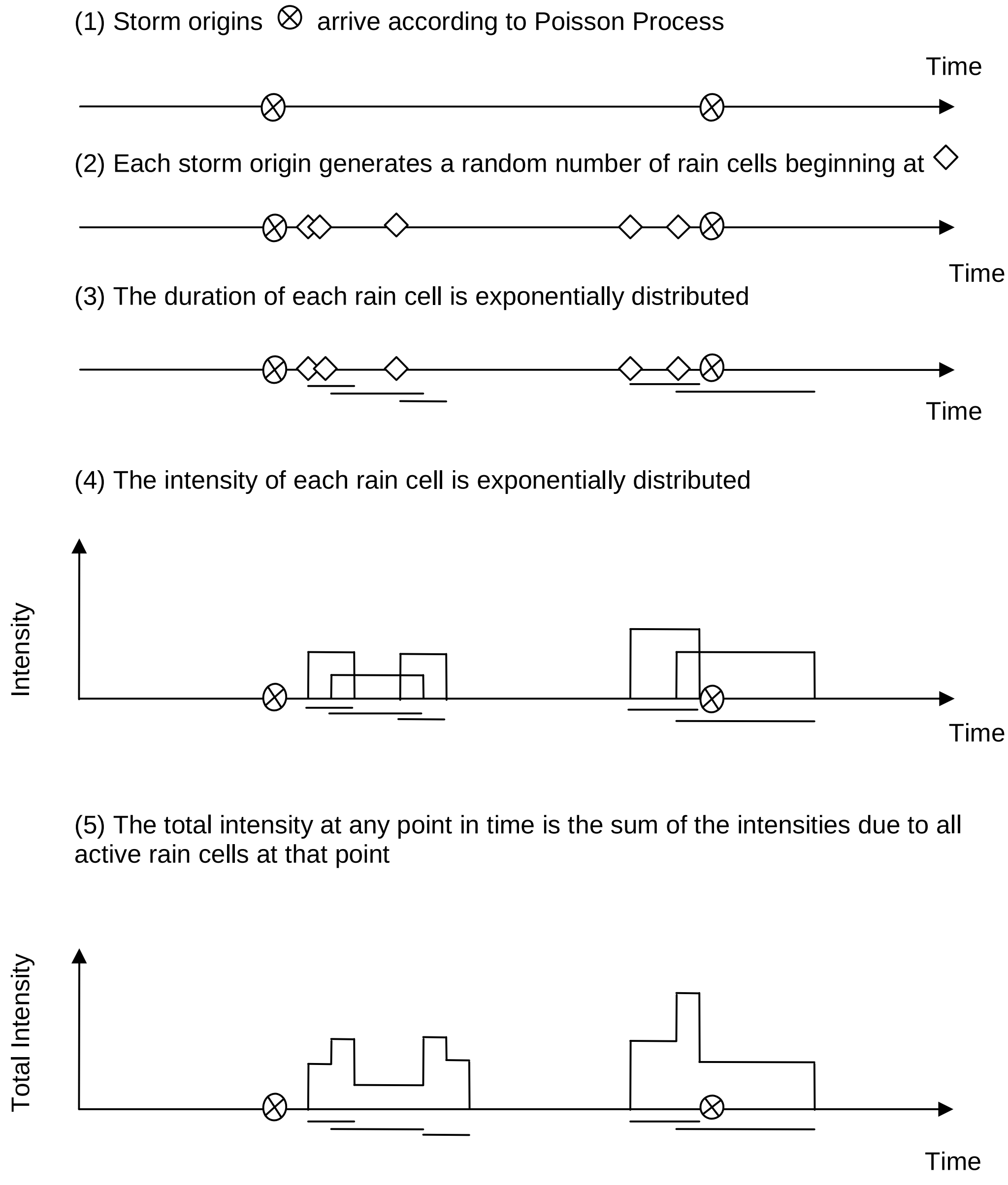

Figure 2.1 Schematic representation of the Neyman – Scott Model

The generating mechanism of any rainfall event is assumed as the storm origin. This storm origin may be passing fronts or some other criteria for convection storms, from which rain cells arise in our study area. The following are the main assumptions of the Neyman-Scott Model:

- The storm origins arrive according to a Poisson process with rate parameter λ.
- Each storm origin generates a random number of rain cells, for which the waiting time after the storm origin of each rain cell is exponentially distributed with parameter β.
- The duration of each rain cell is exponentially distributed with parameter η.
- The intensity of each rain cell is constant throughout its duration and is exponentially distributed with parameter ε.
- The total intensity at any instant in time is the sum of all the intensities due to all active cells at that time.

It is also assumed that the intensity, duration and waiting time after the storm origin of any rain cell are independent of each other and other rain cells.

2.2 The Bartlett-Lewis Rectangular Pulse Model

The schematic diagram of the Bartlett – Lewis Model is given in Figure 2.2.

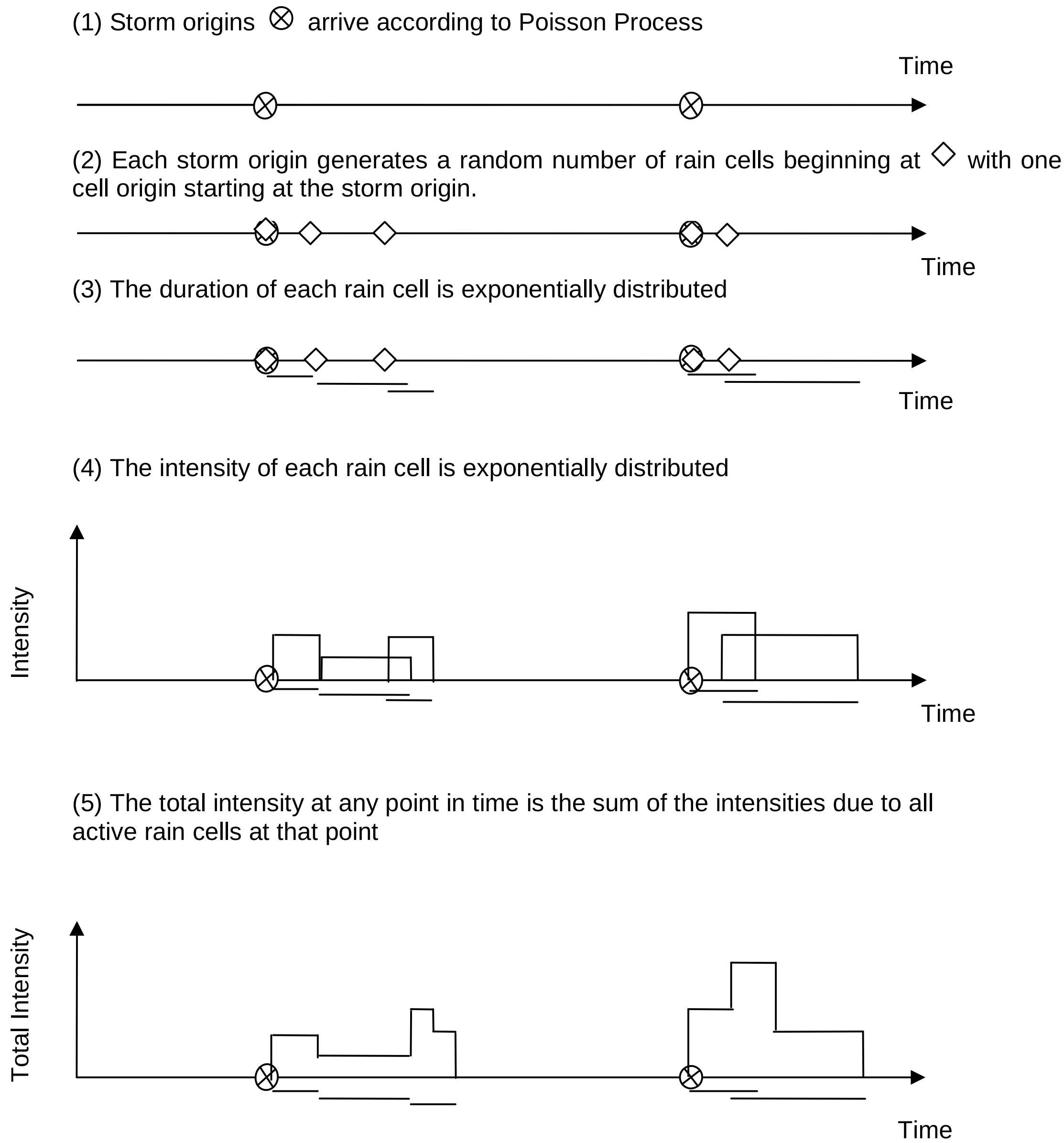

Figure 2.2 Schematic representation of the Bartlett-Lewis Model

The Bartlett-Lewis Model has two versions namely, randomised and non-randomised. Following are the principal assumptions of the non-randomised Bartlett Lewis Model:

- Storm origins occur with a Poisson process of rate λ.

- Each storm origin is associated with the arrival of cell origins which are governed by a Poisson process of rate β. It is assumed that one cell arrives at the storm origin.
- Cell arrivals of each storm terminate after a time exponentially distributed with parameter γ.
- Each cell has a duration exponentially distributed with parameter η.
- Each cell has uniform intensity with specified distribution, typically assumed as exponential with parameter μ_x.

In the randomised Bartlett Lewis Model (Rodriguez-Iturbe et al., 1988), the parameter η of the rectangular pulse duration varies randomly between storms. This model thus incorporates structurally different storms with cells having random durations. The parameter η is assumed to be Gamma distributed with index α and scale parameter v. In summary, the non-randomised version of the Bartlett-Lewis Model has five parameters namely, λ, β, γ, η and μ_x to calibrate, whereas the randomised version has six parameters namely, λ, β, γ, α, v and μ_x to calibrate. Gyasi-Agyei and Willgoose (1997) showed that both the randomised and the non-randomised Bartlett-Lewis Models gave satisfactory results for point rainfall estimation. As the randomised model has one parameter more than the non-randomised model it needs more calculation time for the processor, especially when the data are grouped on a monthly basis. For this reason, the non-randomised version of the Bartlett-Lewis Model was adopted for the generation of the wet and dry sequence of the rainfall process.

The next chapter describes the data and the rainfall statistics used in the model.