EPSRC
Engineering and Physical Sciences Research Council

Building Knowledge for a Changing Climate

UK Climate Impacts Programme

BETWIXT
Built EnvironmenT: Weather scenarios for investigation of Impacts and eXTremes

BETWIXT Technical Briefing Note 2

Version 2, January 2004

NEYMAN-SCOTT RECTANGULAR PULSES RAINFALL SIMULATION SYSTEM

A.BURTON, C.G.KILSBY*, A.MOAVEN-HASHEMI, P.E.O'CONNELL

Water Resource Systems Research Laboratory,
School of Civil Engineering and Geosciences
University of Newcastle upon Tyne
Newcastle NE1 7RU

*Corresponding author: c.g.kilsby@ncl.ac.uk

UNIVERSITY OF NEWCASTLE

WRSRL

1 Introduction

The Generalised Neyman-Scott Rectangular Pulses (GNSRP) model (Cowpertwait, 1994; 1995) is one of the most advanced approaches to stochastic rainfall simulation currently available. The model has been implemented in software at Newcastle University (UNEW) known as RainSim, which has been built in order to provide a state-of-the-art rainfall simulator for a variety of applications and to act as a sound basis for further research into rainfall modelling. This document describes the GNSRP model and gives a brief overview of its development and capabilities. A demonstration application is also reported to illustrate the utility of the software.

UNEW previously produced a modelling package called RainSim, coded in Visual Basic, which has the capability of simulating rainfall time series either at a single location or distributed across a region of up to about 200km in diameter using the GNSRP (Generalized Neyman Scott Rectangular Pulses) spatial-temporal stochastic model. The parameters of the model are fitted using a set of statistics which are determined from observed rainfall time series. This package has been used extensively in various projects (Cowpertwait *et al.*, 2002, Fowler *et al.*, 2000). Some of the functionality, for single sites, was made available in the WRc software package StormPac, used extensively throughout the UK water industry for storm sewer design.

There were a number of shortcomings with these packages, and RainSim has since undergone various improvements and modifications, notably under the EU FRAMEWORK project. This version, (RainSim V2) was implemented in C++, so enabling much faster simulation and analysis times. The model is now being further developed in the EPSRC/UKCIP Building Knowledge for Climate Change Programme, under the BETWIXT project.

2 The Generalized Neyman Scott Rectangular Pulses (GNSRP) Model

RainSim is an implementation of the spatial temporal Generalized Neyman Scott rectangular pulses (GNSRP) stochastic model for simulating rainfall (Cowpertwait, 1994; 1995). The spatial temporal and generalized aspects of this model are extensions of the Neyman Scott Rectangular Pulses (NSRP) model, described in the following section.

2.1 Definition of the Neyman-Scott Rectangular Pulses Model

In the NSRP model the positions of the rain cells are determined by a set of independent and identically distributed random variables representing the time intervals between the storm origin and the birth of the individual cells. The model structure is shown in Figure 1 and is based upon the following assumptions:

- storm origins arrive in a Poisson process with the arrival time represented by a parameter λ;
- each storm origin generates a (Poisson) random number *C* of raincells separated from the storm origin by time intervals that are each exponentially distributed with parameter β;
- the duration of each raincell is exponentially distributed with parameter η;
- the intensity of each raincell is exponentially distributed and represented with a parameter ξ;
- the rainfall intensity is equal to the sum of the intensities of all the active cells at that point;

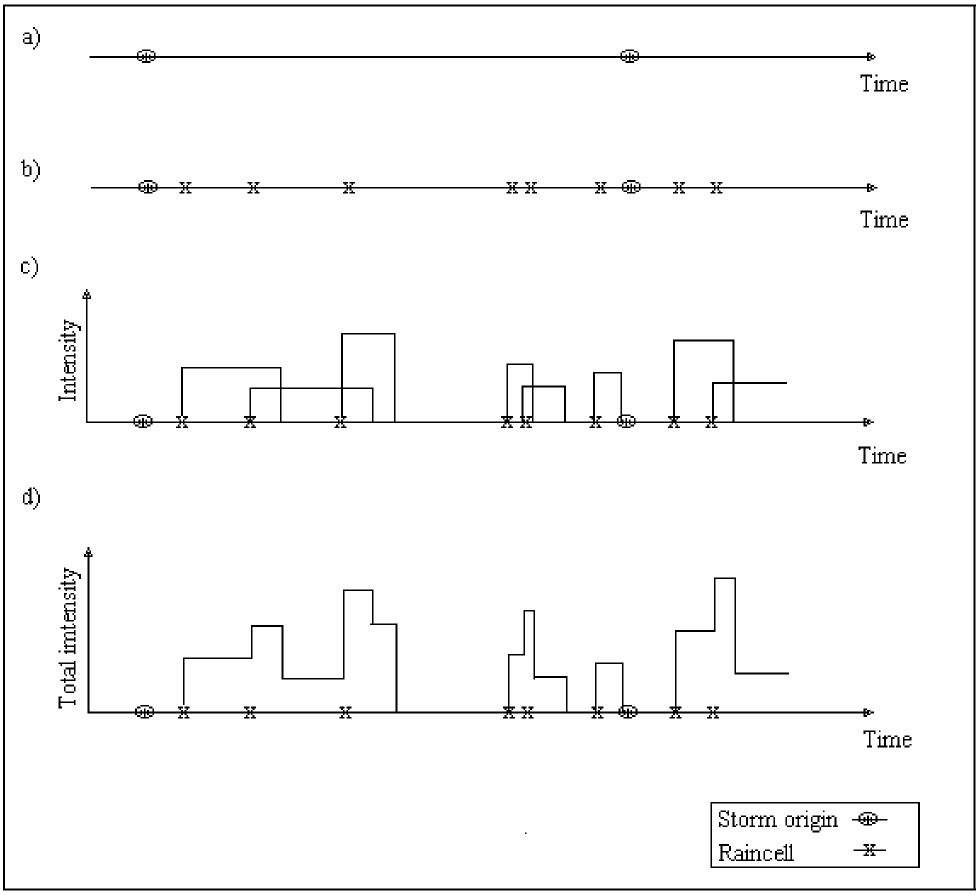


Figure 1. A schematic representation of the NSRP model

The parameters of the model can be summarised as follows:

1. λ^{-1} the average waiting time between subsequent storm origins (*hours*),
2. β^{-1} the average waiting time of the raincells after the storm origin (*hours*),
3. η^{-1} the average cell duration (*hours*),
4. ν^{-1} the average number of cells per storm,
5. ξ^{-1} the average cell intensity (*mm/hour*).

2.2 Fitting procedure

In general, the parameters of the NSRP model can be estimated by matching, as closely as possible, the analytical moments with their corresponding values estimated from historical rainfall time series. This is an appropriate method to ensure that these characteristics will be properly reproduced.

Therefore, the model parameters may be estimated by minimising the sum of squared differences between the estimated statistics $\hat{f}$ and the model statistics $\boldsymbol{f}$. This is carried out by minimizing the following function D:

$$D = \sum_{i=1}^{12} \sum_{h \in H} [1 - \frac{f_i(h)}{\hat{f}_i(h)}]^2$$

where H is a set of aggregation levels, $\boldsymbol{f_i} = f(\lambda, \beta, \eta, \nu, \xi)$ is a set of model statistics and $\hat{f}_i$ denotes the sample estimate of $\boldsymbol{f_i}$ for the ith calendar month and is taken from all of the available observation data for that specific month. There are five parameters to be estimated, so a natural procedure would be to equate five statistical properties, or 'moments', taken from the observed time series with their derived expressions for the model, and to solve the resulting set of simultaneous equations for the parameter estimates. The model would then fit five sample moments exactly, with the fit to other values not guaranteed. A more flexible fitting procedure is adopted here which assumes that it is more desirable to fit a larger set of sample moments approximately rather than a smaller set exactly. A standard numerical routine, such as the Simplex algorithm (used in RainSim V2), is used to find the parameter set that minimizes the *D* function subject to fixed upper and lower bounds applied to the parameters.

2.3 The Generalized and Spatial-Temporal versions of the NSRP model

The generalized model allows for the possibility of two (or more) types of raincell, each type with different properties: conceptually these may be considered as stratiform (*i.e.* large area and low intensity) and convective (smaller area and more intense). During the modelling process, each raincell has a probability (α_i) of being of type i, where the α_i are parameters of the model and $\sum \alpha_i = 1$. Furthermore, the different properties of the raincell types are modelled by making the parameters representing the distributions of intensity, duration and radius specific to the raincell type. That is, the NSRP model uses the parameter ξ to represent the intensity of all raincells, whereas the GNSRP model can use ξ_1 for convective cells and ξ_2 for stratiform cells. The generalized model is represented in Figure 2A.

In the spatial temporal version of the model, the cells are considered to be circular in space, with the radius being exponentially distributed with parameter γ and their locations in space described by a Poisson process with parameter ρ. During the cell's lifetime, there is a uniform rainfall intensity both in space and in time. The diagram in Figure 2B shows

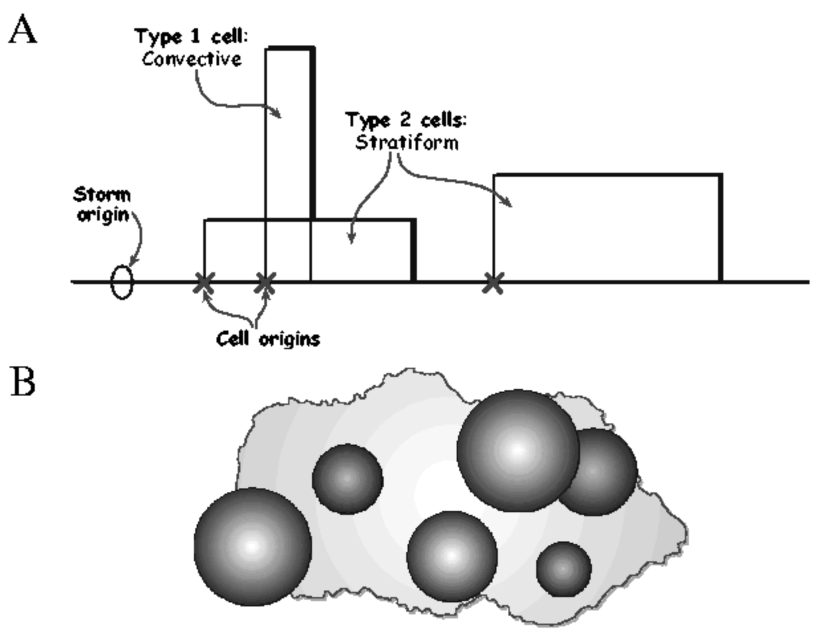


Figure 2 (A) GNSRP model at a point, (B) raincells over a catchment

simulated locations of raincells over a catchment during a simulated storm.

Both the spatial-temporal and generalized versions of the NSRP model are compatible and can be combined. The combined model has been demonstrated to successfully reproduce many of the quantitative and qualitative properties of rainfall time series observed in Europe. The major practical difficulty in implementing the model is in estimating a set of parameters. The same fitting procedure can be used as for the NSRP model but with the provision of analytical moments which take account of the new model forms and parameters. These relationships have been determined by Cowpertwait (1994, 1995, 1998). Consequently, a set of parameters can be estimated that best reproduces the chosen observed statistics.

3 Software design

RainSim V2 offers an integrated approach by combining the model fitting results in Cowpertwait (1994, 1995, 1998) and Cowpertwait *et al*. (1996a,b). One of the more recent developments included in the specification is that of fitting the single cell type version of the GNSRP model to the third order moment of the observed rainfall time series. Consequently, the specification defines a GNSRP model with either: multiple cell types fitted using up to second order statistics, or a single cell type fitted using up to third order statistics. In particular, the software can be considered to be divided into 3 main parts:

1. Time series analysis - To determine common statistics at various aggregation levels from time series of single or multiple raingauges, e.g. variance of 1 hour rainfall or probability of a 12 hour dry period.
2. Fitting - To choose a set of parameters for the model, so that the model will on average match a given set of rainfall statistics.
3. Simulation - To use the stochastic model to generate single or multiple time series of rainfall according to a given set of parameters.

3.1 Third order moment fitting

Fitting the GNSRP rainfall simulator using third order moment (or skewness) statistics is considered to be necessary because the evidence of the model fit to extreme precipitation using only second order statistics is not convincing. Third order moments represent the 'one-sidedness' of a distribution. Consequently, in a strictly positive distribution such as for rainfall they provide a measure of the intensity and probability of extreme high precipitation events. The inclusion of third order moments in RainSim V2 is therefore of high priority for any application with an emphasis on flood risk estimation.

The fitting procedure of the GNSRP model has been extended to include the third order central moment:

$$E\left[\left(Y_h - E\left(Y_h\right)\right)^3\right]$$

where Y_h is the quantity of rainfall observed in an h hour period and E refers to statistical expectation. That is, an expression has been derived so that the estimated value of this moment can be calculated from the model parameterisation. This expression is appropriate for versions of the GNSRP model which use only a single raincell type. When used in the fitting procedure, a greater emphasis can be placed upon matching the extreme tail of the rainfall distribution.

3.2 Testing the software

An extensive programme of testing of the software has been carried out. The modular structure of object-orientated software allows a first level of testing directly on a particular subroutine within the software with the assurance that subsequent revisions to other parts of the software will not affect the performance of the tested routine.

A general programme of testing was carried out to check that the statistical analysis of existing time series worked correctly. Having demonstrated that the software can be used to calculate the statistics of time series correctly, this ability was utilised to test the rest of the software using the following approach: a set of up to three statistics was chosen (S_0) (with the same values for each month of the year), a parameter set (P) was fitted to S_0 using the software and, from this, the software was used to calculate an expected set of statistics (S_E) for any simulation using P. The ability of the software to match S_0 and S_E using the fitting procedure is of great importance but simply demonstrates the utility of the minimisation algorithm. Figure 3 shows two examples of the match achieved by the minimisation routine.

Thorough testing of RainSim V2 has determined that all of the statistics used by the model were well matched and thus the accuracy of the methodology, its coding and the simulation engine were demonstrated. The exception to this is found to be in the fitting of the proportion of dry days in a month (PDRY). The discrepancy comes from an approximation made by Cowpertwait (1998) whereby a dry day is assumed to have zero rainfall. This is not actually the case in the model: rainfall is simulated in 0.1mm intervals and therefore a dry day is a day in which rainfall is simulated to be less than 0.05mm. This difference is sufficient to incorporate a bias in the simulated rainfall whereby a PDRY value in S_{sim} will exceed a PDRY value in S_E by on average 2 percent. Fortunately, this bias is small and can be corrected.

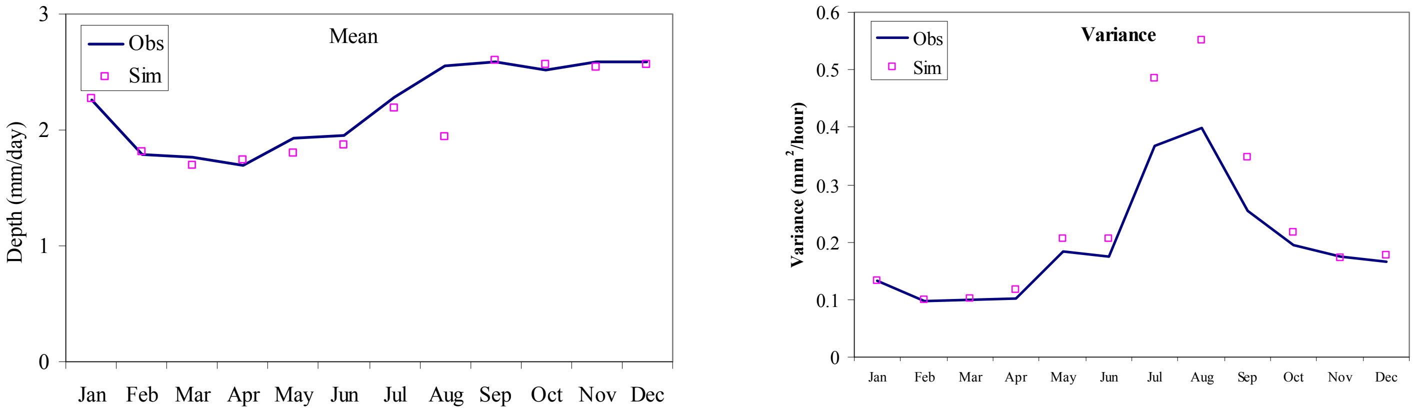


Figure 3 Fitted and simulated, daily mean and variance statistics compared with the observed data

3.3 Software implementation

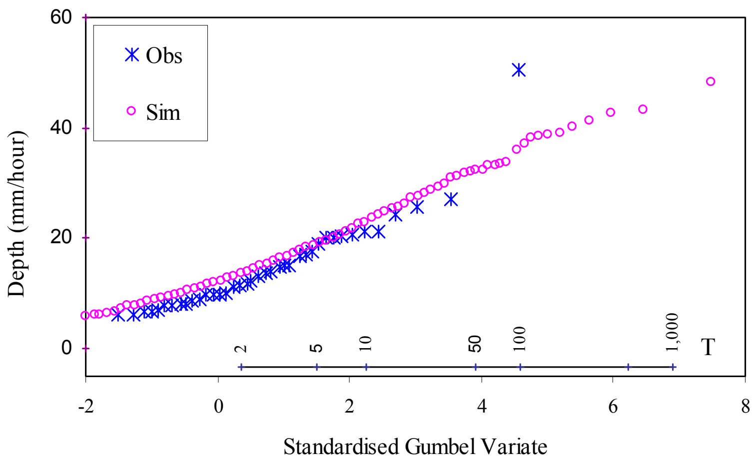
A most useful implementation for BETWIXT incorporates the ability to parameterize the GNSRP model at a single site using up to third order statistics and provide complementary utilities for time series analysis and for the rainfall simulation engine. Although a PC GUI will be highly desirable for users who do not wish to develop the rainfall model, a textual interface within a UNIX environment allows researchers the many advantages of that environment over the PC Windows visual interface. In particular, it is extremely easy to automatically carry out batch processing of RainSim tasks controlled by UNIX shell scripts, a facility which is extremely useful for research into rainfall modelling. For example, a script can be easily created to take 100 raingauges and to fit the model to each individually - quickly producing the information needed for a regionalization study of the GNSRP parameters.

4 Demonstration Application

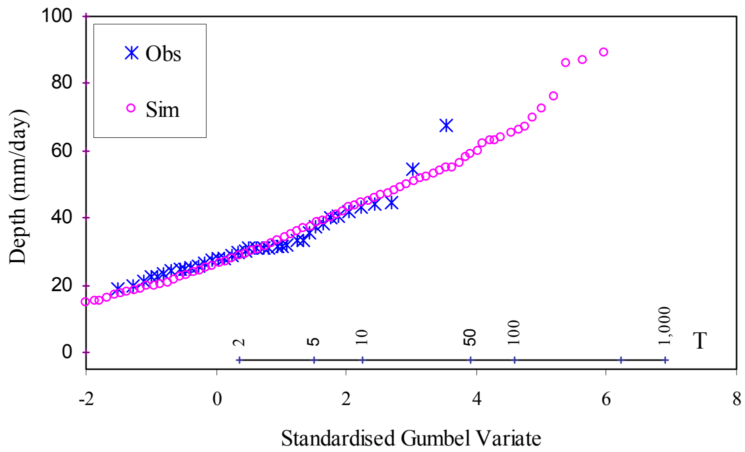
The model has been used to simulate present climate rainfall at Ringway (UK). The requirement was to parameterize the GNSRP model and generate a 1,000-year hourly-resolution rainfall time series that matched the statistical properties of the observed rainfall. This demonstrates the ability of RainSim to use the skewness statistic.

40 years of hourly rainfall data were available at Ringway. Statistics and weights were chosen for use in the fitting procedure so as to optimise the fit to the extreme values. Figure 3 shows a comparison between the selected observed statistics and the fitted and simulated statistics, clearly demonstrating excellent matches to the observed.

Figure 4 shows Gumbel plots comparing hourly and daily maxima with the observed record. This shows a good correspondence between the observed and the simulated rainfall for extremes.



(a) hourly



(b) daily

Figure 4 Extreme value plots comparing observed and simulated annual rainfall maxima at Ringway

5 Summary and planned developments

The RainSim package is able to: analyse the statistics of rainfall time series, fit the GNSRP model to a set of rainfall statistics (including the fitting of the skewness statistic) and to simulate 1000's of years of continuous rainfall using the GNSRP model.

Further capabilities are being developed as part of BETWIXT and have not been described here. These include the following features;

- Generation of rainfall series for future climate scenarios. This is achieved by modifying the rainfall statistics to be consistent with those of the UKCIP02 scenarios, and then refitting the model parameters;
- Simulation of rainfall at any site in the UK. This is being developed by generating 5km grids of rainfall statistics for the UK from the UKCIP02/UKMO climatology;
- Improved reproduction of extreme values, consistent with the Flood Estimation Handbook DDF (Depth-Duration-Frequency) model;
- Disaggregation of the hourly rainfall series down to 5-minute resolution using a separate stochastic model, to provide data suitable for use in urban drainage studies;
- Combination with the CRU daily weather generator.

6 References

Cowpertwait, P.S.P., 1994, A generalized point process model for rainfall, *Proc. R. Soc. Lond*., 447, 23-37

Cowpertwait, P.S.P., 1995, A generalized spatial-temporal model of rainfall based on a clustered point process, *Proc. R. Soc. Lond*., 450,163-175

Cowpertwait, P.S.P., P.E.O'Connell, Metcalfe, A.V. and Mawdsley, J.A. 1996a, Stochastic point process modelling of rainfall. I. Single site fitting and validation, *J. Hydrol.,*, 175, 17-46

Cowpertwait, P.S.P., 1998, A Poisson-cluster model of rainfall: high-order moments and extreme values, *Proc. R. Soc. Lond*., 454, 885-898

Cowpertwait, P.S.P., O'Connell, P.E., Metcalfe, A.V. and Mawdsley, J.A. 1996b, Stochastic point process modelling of rainfall. II. Regionalisation and disaggregation, *J. Hydrol.,* 175, 47-65

Cowpertwait, P.S.P., Kilsby, C.G., and O'Connell, P.E., 2002, A Spatial-Temporal Neyman-Scott Model of Rainfall: Empirical Analysis of Multisite Data, *Water Resour. Res,* 38, 8, 10.1029/2001WR000709

Fowler, H.J., Kilsby, C.G., and O'Connell, P.E., 2000 A stochastic rainfall model for the assessment of regional water resource systems under changed climatic conditions *Hydrol. Earth Sys. Sci*., 4, 261-280.