
PERFORMANCE OF MIXED EXPONENTIAL AND EXPONENTIAL DISTRIBUTION REPRESENTING RAIN CELL INTENSITY IN NEYMAN-SCOTT RECTANGULAR PULSE (NSRP) MODEL

Fadhilah Yusof^{1,*}, Zalina Mohd Daud²,
Nguyen V-T-V³, Zulkifli Yusop⁴

¹Department of Mathematics, Faculty of Science, Universiti Teknologi Malaysia, Skudai 81310 Johor.

²College Science and Technology, Universiti Teknologi Malaysia City Campus, Jalan Semarak, 54100 Kuala Lumpur.

³Department of Civil Engineering and Applied Mechanics, McGill University, Quebec, Canada

⁴Institute of Environmental and Water Resources Management, Universiti Teknologi Malaysia, Skudai 81310 Johor

*Corresponding Author : fadhilah@mail.fs.utm.my

Abstract: Sub-daily timescale data such as hourly data are needed for modeling urban systems. However such series are not readily available as compared to daily rainfall series. Stochastic rainfall models are useful in estimating input for design work. One of the models that applies the clustered point process theory is the Neyman-Scott Rectangular Pulses (NSRP) model. The model uses a flexible model fitting procedure which involves matching approximately a chosen set of historical statistics which exceeds in number of set of parameters to be estimated. An optimization technique called Shuffle Complex Evolution (SCE-UA) was used to estimate the parameters. The performance of NSRP model was evaluated using 10 years hourly data taken from a station in Wilayah Persekutuan. Two distributions, namely exponential (EXP) and mixed exponential (MEXP) were used to model the cell intensities in the model. The models were evaluated on a monthly basis regarding their ability to preserve the statistical properties as well as the physical properties of the rainfall time-series over timescales of 1 h, 6 h and 24 h. The performance of the models with the two different distributions was evaluated and compared. The model with the mixed exponential (MEXP) distribution perform better in preserving most of the statistical and physical properties of the observed data.

Keywords: *Neyman-Scott Rectangular Pulses (NSRP) Model; Shuffle Complex Evolution; Hourly Rainfall; Aggregation*

Abstrak: Data dalam skalar masa sub-harian seperti data setiap jam diperlukan untuk pemodelan sistem bandar. Walau bagaimanapun data sedemikian sukar didapati berbanding data hujan harian. Model hujan berstokastik adalah suatu kaedah penting dalam menentukan input rekabentuk. Salah satu model yang menggunakan teori proses titik berkelompok ialah model Neyman-Scott Rectangular Pulse (NSRP). Ia menggunakan prosedur kecocokan model yang fleksibel, melibatkan perbandingan suatu set statistik lampau yang dipilih, melebihi daripada

bilangan parameter model. Suatu teknik pengoptimum yang dinamai Shuffle Complex Evolution (SCE-UA) digunakan untuk menganggar parameter model. Keupayaan model ini diuji ke atas 10 tahun data setiap jam yang diambil dari sebuah stesen di Wilayah Persekutuan. Dua taburan iaitu taburan Eksponen (EXP) dan taburan Eksponen Bergabung (MEXP) digunakan untuk mewakili intensiti sel dalam model ini. Model-model dinilai secara bulanan dari segi keupayaan mereka mengekalkan sifat statistik dan fizikal siri-masa hujan dalam selang masa 1 jam, 6 jam dan 24 jam. Keupayaan model dengan dua taburan yang berasingan telah dibandingkan. Kajian menunjukkan bahawa model dengan taburan Eksponen Bergabung (MEXP) mempunyai keupayaan yang lebih baik dalam mengekalkan sifat cerapan.

Katakunci: *Model Neyman-Scott Rectangular Pulses (NSRP); Shuffle Complex Evolution (SCE-UA); Hujan Setiap Jam; Agregasi*

1.0 Introduction

Stochastic rainfall modeling involves using historical data to estimate model parameters which may then be used to simulate the desired length of rainfall series that mimic the properties of the historical data. It is widely acknowledged that more adequate information would be gathered from this generated series in assessing the response and reliability of a water resource system.

Rainfall data in daily aggregation is used in many hydrological applications. In relation to that many stochastic models that employ daily rainfall data have been developed (e.g. Todorovic et al, 1975; Han et al, 1976; Katz, 1977; Woolhiser et al, 1982). However, data at finer scales such as hourly data is even more crucial in rainfall-runoff modeling. Hence in this study, a cluster-based stochastic model namely, the Neyman-Scott Rectangular Pulse (NSRP) model is used to model the hourly rainfall series. This model uses a modest number of parameters to represent the rainfall processes and the underlying physical phenomena such as the arrival of storm or the rain cells (Rodriguez-Iturbe et al., 1987a, 1988). Previous studies assumed that rain cell intensity follows an exponential distribution due to its small number of parameter (e.g. Rodriguez-Iturbe et al., 1987a, 1988; Cowpertwait, 1996). However, the choice of distribution for the cell intensity in the NSRP model is arbitrary. Cowpertwait (1996) attempted a heavier-tailed distribution such as Weibull to improve the fit in the extremes. It was found that the mixed exponential distribution is the most accurate distribution for describing the distribution of the hourly rainfall amounts as compared to the other popular candidate distributions such as exponential, gamma and Weibull (Fadhilah et al., 2007). Hence, this distribution is used to represent the rain cell intensities. This study attempts to compare the original NSRP exponential (EXP) model with a modified NSRP mix exponential (MEXP) model. The NSRP (EXP) assumes that the rain cell intensities follow the exponential distribution while the NSRP (MEXP) assumes that the rain cell intensities follow the mixed exponential distribution. The performance of the models is assessed in terms of its ability in preserving the statistical and physical properties of the observed data.

2.0 Methodology

2.1 Model description

The Neyman-Scott is a clustered point process model. This model assumes that there exists a generating mechanism called the storm origin in any event which may be passing fronts or some other criteria for convection storms from which rain cells develop. The Neyman-Scott models are described by 3 independent elementary stochastic processes, namely

- A process that sets the origin of the events;
- A process that sets the number of rain cells generated by each event;
- A process that sets the origin of the cells.

Storm origins are governed by a Poisson process with parameter λ . At a point on the ground the storm is conceptualized as a random number of rain cells C . Natural candidates for the distribution of the number of cells C follow the Poisson distribution and the geometric distribution. The cell origins are independently separated from the storm origin by distances that are exponentially distributed with parameter b . No cell origins are assumed to be located at the storm origin. A rectangular pulse is associated independently with each cell origin with the duration and intensity (depth) being independent. The duration and intensity are assumed to be exponentially distributed with parameters h and $m_x = \frac{1}{x}$, respectively. A scheme for the NSRP model is shown in Figure 1.

2.2 The basic mathematical theory

The precipitation intensity at time t , $Y(t)$, is given by the sum of the intensities of the individual active cells at time t :

$$Y(t) = \int_{u=0}^{\infty} X_{t-u}(u) dN(t-u) \tag{1}$$

where $X_u(k)$ is the random depth of the pulse originating at time u measured at time k later and $\{N(t)\}$ counts occurrences in the Poisson process of pulse origins. Note that the intensity of $N(t)$ is λm_c , where m_c denotes the mean number of cells per storm $E[C]$. The derivative of the counting process is

$$dN(t - u) = \begin{cases} 1 & \text{if there is a cell at } t - u \\ 0 & \text{if otherwise} \end{cases} \quad (2)$$

and for the rectangular pulses, we have

$$X_{t-u} = \begin{cases} X & \text{with probability } R(x) \\ 0 & \text{with probability } 1 - R(x) \end{cases} \quad (3)$$

where $X_{t-u}(u)$ is the intensity of the rectangular pulse triggered at time u and $N(t)$ represents the counting stochastic process of the arrivals of the individual cells. $R(x)$ is the survival function of X .

Since rainfall data are usually available only as rainfall depths in discrete time intervals (e.g. historical records of hourly or daily totals), the aggregated properties are needed to estimate the model parameters. The aggregated process at time scale h (the total depth in a time interval h) is given by:

$$Y_i^{(h)} = \int_{(i-1)h}^{ih} Y(t) dt \quad (4)$$

Thus, if h is measured in hour, the series $\{Y_i^{(h)} : i = 1, 2, \dots\}$ is a rainfall time series at the h -hour level of aggregation, i.e. an h -hourly rainfall time series. The second-order properties of the aggregated process (Rodriguez-Iturbe et al., 1987a) are

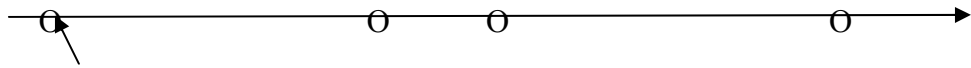
$$E\{Y_i^{(h)}\} = h\lambda E\{C\}E\{X\} / \eta \quad (5)$$

$$\begin{aligned} Var\{Y_i^{(h)}\} &= \lambda\eta^{-3}(\eta h - 1 + e^{-\eta h})[2\mu_c E\{X^2\} + E\{C^2 - C\}\mu_x^2\beta^2 / (\beta^2 - \eta^2)] \\ &- \lambda(\beta h - 1 + e^{-\beta h})E\{C^2 - C\}\mu_x^2\beta^{-1} / (\beta^2 - \eta^2) \end{aligned} \quad (6)$$

$$\begin{aligned} Cov\{Y_i^{(h)}Y_{i+k}^{(h)}\} &= \lambda\eta^{-3}(1 - e^{-\eta h})^2 e^{-\eta(k-1)h} \\ &\times [\mu_c E\{X^2\} + \frac{1}{2}E\{C^2 - C\}\mu_x^2\beta^2 / (\beta^2 - \eta^2)] \\ &- \lambda(1 - e^{-\beta h})^2 e^{-\beta(k-1)h} E\{C^2 - C\}\mu_x^2 / [2\beta(\beta^2 - \eta^2)] \end{aligned} \quad (7)$$

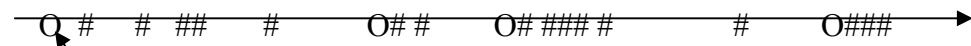
This study assumed the Poisson distribution to represent the distribution of C . Therefore, in the above equations the followings are to be substituted:

Storm origins arrive according to a Poisson Process



Time
Storm origin

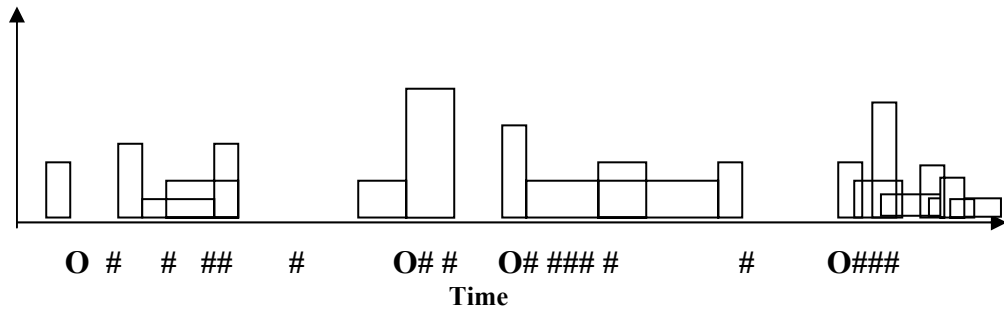
Each origin generates a random number of rain cells with cell origins at #



Time
Cell origin

A rectangular pulse is associated with each rain cell.

Intensity



The total intensity at any point in time is the sum of the intensities of all active rain cells at that point.

Total Intensity

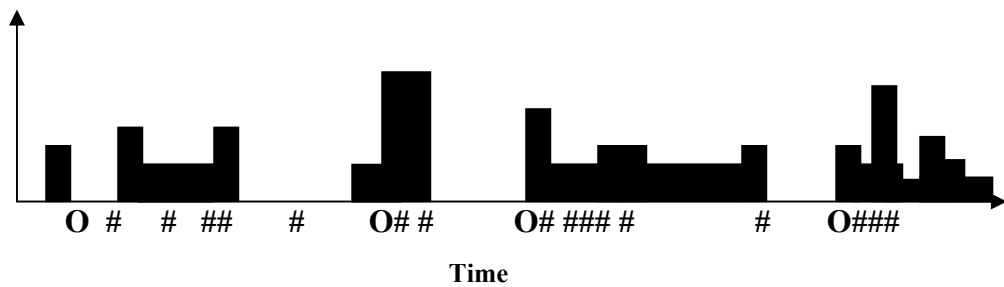


Figure 1: A scheme for Neyman-Scott rectangular pulses model

$$\mu_c \equiv E(C) = \nu; E(C^2 - C) = \nu^2 - 1; \mu_x \equiv E(X) = \xi^{-1}; E(X^2) = 2\xi^{-2}. \tag{8}$$

By assuming C-1 follows a Poisson distribution with mean (ν-1), the following expression for the probability that an arbitrary interval of length h is dry was derived by Cowpertwait (1991) :

$$P\{Y_i^h = 0\} = \exp\left(-\lambda h + \lambda \beta^{-1} (\nu - 1)^{-1} \{1 - \exp[1 - \nu + (\nu - 1)e^{-\beta h}]\} - \lambda \int_0^\infty [1 - p_h(t)] dt\right) \tag{9a}$$

where

$$p_h(t) = \left\{ e^{-\beta(t+h)} + 1 - (\eta e^{-\beta t} - \beta e^{-\eta t} / (\eta - \beta)) \right\} \times \exp\left\{ -(\nu - 1)\beta(e^{-\beta t} - e^{-\eta t}) / (\eta - \beta) - (\nu - 1)e^{-\beta t} + (\nu - 1)e^{-\beta(t+h)} \right\} \tag{9b}$$

With the above properties, the original model (EXP) has five parameters, namely λ, ν, β, η, and ξ.

2.2.1 NSRP(MEXP)

The mixed-exponential distribution is used as an alternative to the exponential distribution. The probability distribution function of a mixed exponential is given as:

$$f(x) = \frac{\alpha}{\xi} e^{\left(\frac{-x}{\xi}\right)} + \frac{(1 - \alpha)}{\theta} e^{\left(\frac{-x}{\theta}\right)} \tag{10}$$

$$x > 0; 0 \leq \alpha \leq 1; 0 < \xi < \theta$$

The mixed-exponential distribution is a weighted average of two one-parameter exponential distributions. The mixture distribution has three parameters, with α represents the mixing probability, ξ and θ represent the scale parameters and x represents the hourly rainfall.

The second-order properties of the NSRP(MEXP) are similar to equations (5) to (7) but the E(X) and E(X²) to be substituted are;

$$\begin{aligned} \mu_x &= E(X) = \alpha(\xi) + (1 - \alpha)(\theta) \\ E(X^2) &= 2\alpha(\xi^2) + 2(1 - \alpha)(\theta^2) \end{aligned} \tag{11}$$

and the $E(C)$ and $E(C^2 - C)$ to be substituted are

$$\begin{aligned} \mu_c &= E(C) = \nu \\ E(C^2 - C) &= \mu_c^2 - 1 = \nu^2 - 1 \end{aligned} \tag{12}$$

With the above properties, the NSRP (MEXP) has seven parameters, namely $\lambda, \nu, \beta, \eta, \alpha, \xi,$ and θ . The rain cell intensities are represented by two parameters: ξ to represent the heavy cell intensity and θ to represent the light cell intensity. The heavy cell intensity can be interpreted as ‘heavy’ short-duration convective cells and the light cell intensity can be interpreted as ‘light’ long duration stratiform cells. α then measures the proportion of cells being of those types.

2.3 Parameter estimation

In this study, parameter estimation procedure is achieved by minimizing the sum of squares, where the squared terms are the differences between the selected expressions of the model and their equivalent historical sampled values. The fitting procedure by Cowpertwait et al. (1996) is used which assumes that it is more desirable to fit a larger set of sample moments approximately rather than a smaller set exactly.

For the original NSRP (EXP) model, let $M_i \equiv M_i(\lambda, \nu, \beta, \eta, \xi)$ be a function of the NSRP model, and let M_i^s be its historical sampled value.

$$S = \sum_{i=1}^m w_i \left[1 - \frac{M_i}{M_i^s} \right]^2 \tag{13}$$

where w_i is a weight. The use of a ratio of model function is to ensure that large numerical values do not dominate the fitting procedure. Cowpertwait (1996) applied a weight of 100 to the term relating to the sample mean to ensure that this is matched almost exactly by the model. The most frequently used sampled moments are given in Table 1. They are 1 hour mean [$\hat{\mu}(1)$], 1 hour variance [$\hat{\gamma}(1)$], 1 hour autocorrelation of lag-1 [$\hat{\rho}(1,1)$], 6 hourly variance [$\hat{\gamma}(6)$], 6 hourly autocorrelation of lag-1 [$\hat{\rho}(6,1)$], 24 hourly variance [$\hat{\gamma}(24)$], 24 hourly autocorrelation of lag-1 [$\hat{\rho}(24,1)$] and probability of dry days [$\hat{\phi}(24)$] as suggested by Rodriguez-Iturbe et al. (1987), Entekhabi et al. (1989) and Cowpertwait (1991, 1996).

The Shuffled Complex Evolution (SCE) Method is used to minimize equation (13). It is a combination of genetic algorithm and Simplex Downhill search and the algorithm is well-structured especially for rainfall models. The parameters' upper and lower bounds must be known before the algorithm can be used. This technique is a global optimization technique which is not influenced by initial values as compared to local optimization method like Nelder-Mead Method. The later is said to be very sensitive to initial values (Duan, 1992). The same procedure is applied to the NSRP (MEXP) model.

2.4 Model Evaluation

The performance of the model was evaluated using hourly rainfall series for 10 years (1981-1990) recorded at KM 16 Gombak, Selangor (Station 321700). The rainfall station is being maintained by the Department of Irrigation and Drainage (DID). The sample moments were obtained by pooling all the available data for each calendar month. The sample moments evaluated consist of one hour mean, variances and autocorrelations (at one, six and twenty-four hourly), and probability of dry days. The parameters of the NSRP models were then estimated for each month in order to take seasonal variability into account, giving a total of 60 parameters for the NSRP (EXP) model and 84 parameters for the NSRP (MEXP) model (Table 2). Thirty sets of 10 years data were then simulated for each month and were compared with a 10-year historical data. The statistical properties evaluated include mean, variance, skewness and autocorrelations for 1-hour, 6-hour and 24-hour aggregation on monthly and annually basis. The physical properties include the annual maximum for daily and hourly series and the probability of dry days for monthly and annually basis.

Graphically, the comparison between the simulated properties and the observed properties was done using box plots. If the observed property falls in the rectangular box, then the proposed model is said to have an excellent ability in preserving the properties of the historical data. If the observed property falls on the whiskers and within the range defined by the simulated minimum and maximum, then the proposed model is said to have a fair ability in preserving the properties of the historical data. Otherwise, the model either underestimates or overestimates the observed statistical characteristics.

Quantitatively the NSRP (EXP) and NSRP (MEXP) models were compared using the root-mean-square errors (RMSE) calculated for each property. The RMSE formula is as follows:

$$R_M = \left\{ \frac{\sum_{i=1}^n (S - \hat{S}_m)^2}{n} \right\}^{1/2} \quad (14)$$

where R_M is the root-mean-square error, S is statistics of the observed, \hat{S}_m is the median of the simulated statistics and n is the number of simulated statistics.

3.0 Results and Discussion

Values of the sample moments for each of the 12 calendar months over 10 years are presented in Table 1. The parameter estimates for each month were estimated (Tables 2a and 2b).

Table 1: Sample moments of observed hourly rainfall data at station 3217001 from 1981-1990

Months	1-hr-mean $\hat{\mu}(1)$	1-hr-var $\hat{\gamma}(1)$	1-hr-auto $\hat{\rho}(1,1)$	6-hr-var $\hat{\gamma}(6)$	6-hr-auto $\hat{\rho}(6,1)$	24-hr-var $\hat{\gamma}(24)$	24-hr-auto $\hat{\rho}(24,1)$	Dry-Prob. $\hat{\phi}(24)$
Jan	0.099	1.034	0.0322	11.19	0.0038	47.68	0.0105	0.71
Feb	0.261	4.144	0.3336	42.46	0.0936	209.02	0.2485	0.61
Mar	0.279	3.842	0.3478	40.77	0.0039	149.42	0.0212	0.50
Apr	0.277	3.532	0.3637	41.34	0.0067	180.16	0.0850	0.48
May	0.380	4.312	0.4125	48.76	0.1616	265.50	0.1326	0.41
Jun	0.156	2.085	0.4211	23.58	0.1012	110.04	0.1354	0.67
Jul	0.240	3.045	0.5101	39.44	0.0718	185.87	0.0178	0.55
Aug	0.240	3.874	0.3173	36.82	0.0679	189.67	0.0033	0.61
Sep	0.394	4.996	0.3582	51.49	0.0652	214.54	0.0348	0.38
Oct	0.356	4.796	0.3143	44.19	0.0560	191.81	0.0320	0.40
Nov	0.416	4.130	0.4061	48.90	0.1192	241.48	0.0510	0.31
Dec	0.170	1.662	0.4104	18.80	0.1771	118.59	0.1132	0.58

The parameter estimates for λ , η , ν and β for both models were quite similar in most of the months (Figure 2). However, it is not possible to compare the rain cell intensities parameters for the two models because the distributions that represent them are different. Nevertheless, the proportions of heavy rain cell intensities are high in March, April and September (Table 2b). Interestingly, these months correspond to the inter-monsoon season where there is high occurrence of convective rainfall in Klang Valley. The NSRP (MEXP) has two parameters to represent the rain cell intensities. This is physically realistic in representation of rainfall because it allows there to be more than one cell type within the same storm. This also justifies the use of mixed exponential distribution to represent the rain cell intensity. The parameter estimates for the two cells type agree with the observational studies of convective rainfall in the Klang Valley (Noradilla et al., 2006). Figures 3 to 5 show the box plots that represent the simulated properties and the line graph for the observed properties. Both models matched excellently the 1 hourly mean, 24 hourly variance and 24 hourly autocorrelation. Similarly, the 6 hourly variance and 6 hourly autocorrelation show excellent matching in all months. However

for one hourly maximum, variance and skewness, all months were matched excellently by both models except for the months of March and April.

Table 2a: Parameter estimates for the NSRP (EXP) model.

Months	Lambda (λ)	Eta (η)	Nu (ν)	Epsilon (ξ)	Beta (β)
Jan	0.0202	2.1585	1.0196	0.0966	0.4998
Feb	0.0088	2.0096	4.4785	0.0752	0.0247
Mar	0.0342	4.9948	1.0009	0.0261	0.4989
Apr	0.0499	4.9993	1.0054	0.0363	0.3768
May	0.0202	1.7019	3.7622	0.1172	0.0648
Jun	0.0074	1.6466	3.3205	0.0962	0.0272
Jul	0.0303	1.2856	1.0525	0.1032	0.4999
Aug	0.0375	3.7762	1.0035	0.0415	0.4837
Sep	0.0416	2.3510	2.1015	0.0945	0.4908
Oct	0.0499	2.3154	1.313	0.0799	0.4486
Nov	0.0458	1.6929	2.0736	0.1315	0.1612
Dec	0.0108	1.7399	3.8218	0.1367	0.0922

Table 2b: Parameter estimates for the NSRP (MEXP) model.

Months	Lambda (λ)	Eta (η)	Nu (ν)	Epsilon (ξ)	Beta (β)	Alpha (α)	Teta (θ)
Jan	0.0224	2.1357	1.0132	8.228	0.4877	0.8988	18.550
Feb	0.0037	2.2083	10.5137	14.681	0.0100	0.9921	22.757
Mar	0.0468	4.9999	1.0016	12.061	0.3515	0.0003	29.635
Apr	0.0499	4.9997	1.0031	3.201	0.4999	0.2082	33.463
May	0.0199	1.6979	4.5927	4.939	0.0675	0.7239	12.548
Jun	0.0045	1.5409	6.3263	7.693	0.0117	0.9719	31.760
Jul	0.0352	1.2550	1.0215	7.613	0.4999	0.9517	23.573
Aug	0.0499	4.8328	1.0012	14.416	0.4961	0.6613	40.190
Sep	0.0499	2.1551	1.4740	10.623	0.1421	0.1689	11.673
Oct	0.0500	2.3352	1.2741	12.742	0.3058	0.5780	13.175
Nov	0.0462	1.6800	2.1239	6.564	0.1595	0.8202	10.706
Dec	0.0108	2.2477	5.0886	1.897	0.0961	0.3616	10.000

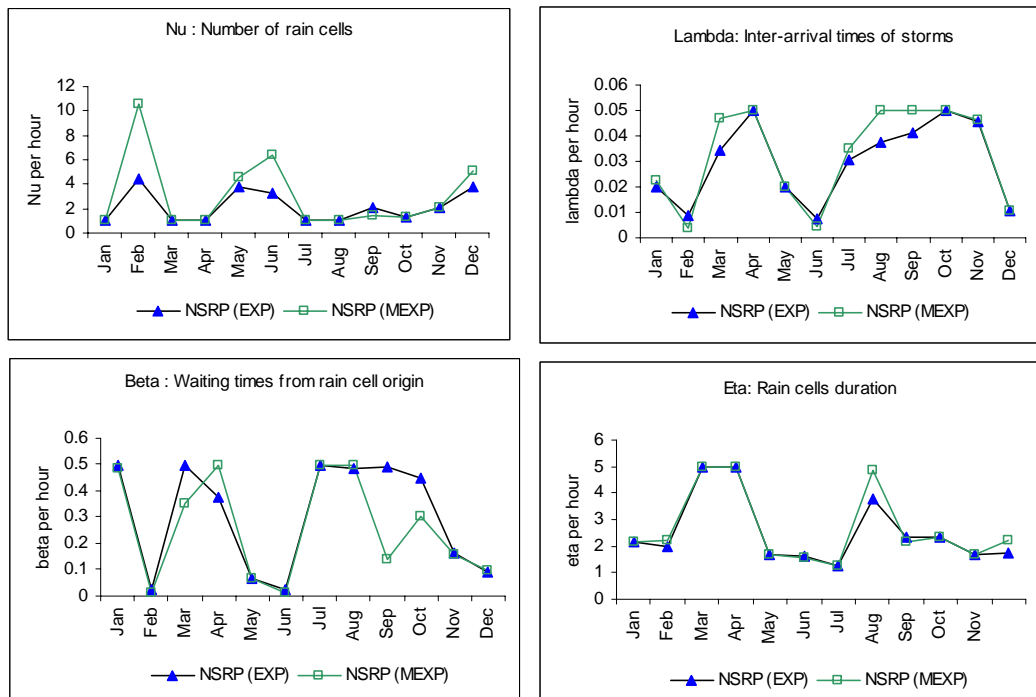


Figure 2: Comparison of parameter estimates between NSRP (EXP) and NSRP (MEXP).

Similarly in the one hourly autocorrelation, both models matched excellently in all months except for March, April and August. Nevertheless, the observed statistics were still within the sampling variability since the values fall in the range defined by the simulated minimum and maximum. The 24 hourly maximum and skewness were matched fairly in all months. The probability of dry days were also preserved within the sampling variability by both models. Therefore, it can be concluded that on a monthly basis, both models have the ability to preserve most of the properties of the historical hourly rainfall series.

The annual properties such as mean, variances and autocorrelations were also compared (Figures 6a and 6b). The 1 hourly and 24 hourly annual means were preserved excellently in both models. The variances and autocorrelations on the other hand were fairly matched by the models but still within the sampling variability. However the annual hourly and 24 hourly maximum seemed to be underestimated. The probabilities of dry days were comparable by both models. Hence, both models have a fair ability in preserving the properties of the observed on the annually basis.

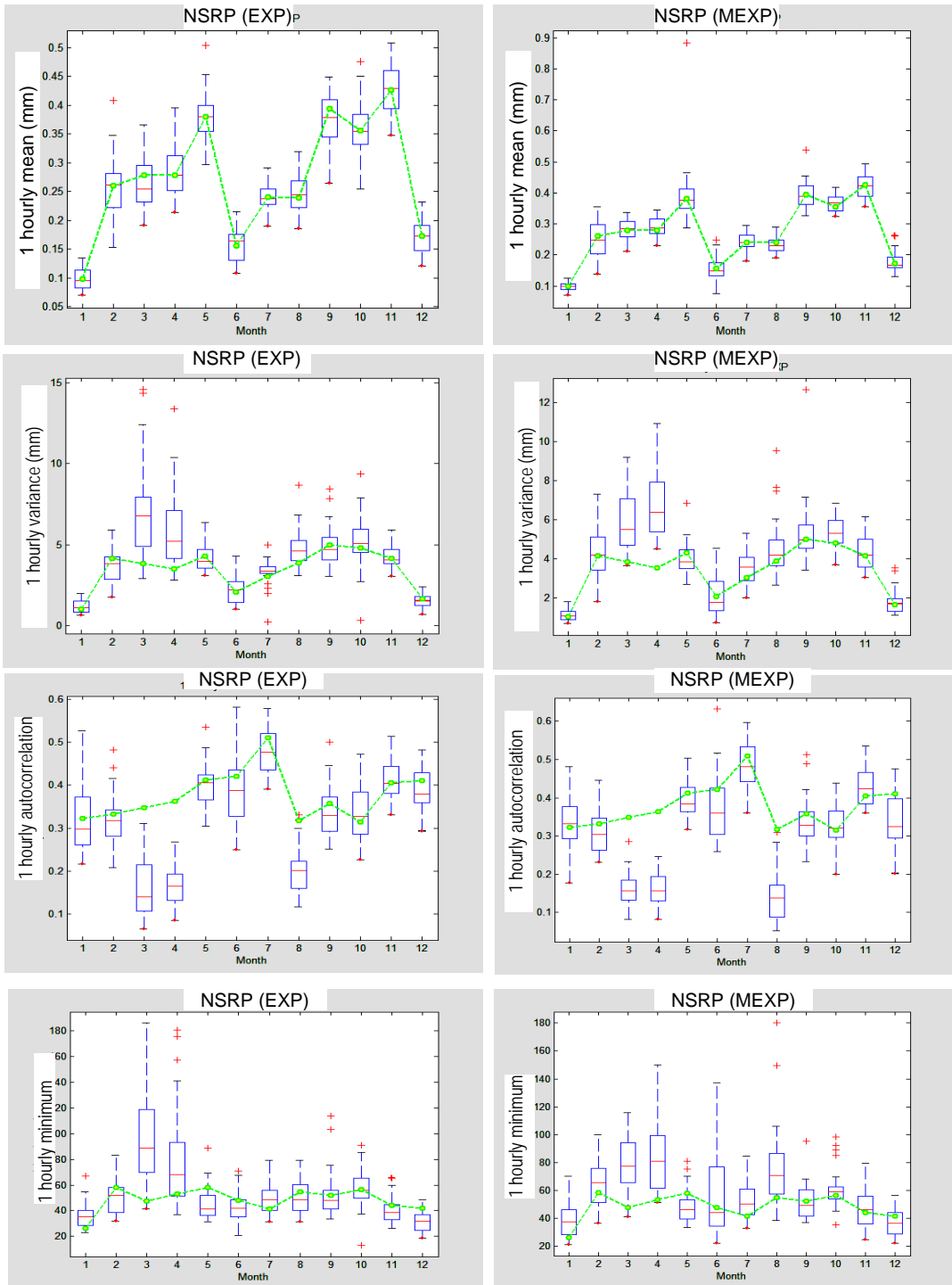


Figure 3 : Comparison between observed and simulated properties of hourly series on a monthly basis.

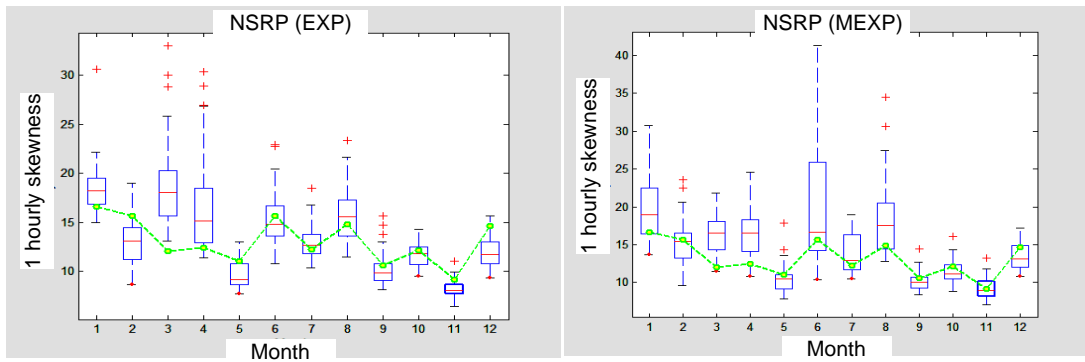


Figure 3: Continued

Table 3: Root mean square errors (RMSE) of NSRP (MEXP) and NSRP (EXP) models.

	1-h mean	1-h var	6-h var	24-h var	1-h auto
NSRP (MEXP)	0.0071	0.9972	4.2310	28.9791	0.1029
NSRP (EXP)	0.00899	1.0225	4.7177	24.6106	0.0916

	6-h auto	24-h auto	1-h skewness	24-h skewness	Annual hourly maximum
NSRP (MEXP)	0.0266	0.0450	2.1604	0.7037	14.2284
NSRP (EXP)	0.0259	0.0495	2.3414	0.7882	13.97

	Annual no. of dry days	Monthly 24-h maximum	Monthly 1-h maximum	Probability dry days	Annual daily maximum
NSRP (MEXP)	1.2663	26.7515	13.9978	0.0561	31.1734
NSRP (EXP)	1.4426	27.2017	14.6840	0.0425	32.1768

	Annual 1-h mean	Annual 1-h variance	Annual 1-h autocorrelation	Annual 24-h mean	Annual 24-h variance
NSRP (MEXP)	0.0346	0.6639	0.0903	0.7871	33.2113
NSRP (EXP)	0.0341	0.6214	0.0821	0.8173	35.7805

	Annual 24-h autocorrelation	Annual Probability of Dry days
NSRP (MEXP)	0.0971	0.0441
NSRP (EXP)	0.1028	0.0519

Note: Values in bold are the smaller RMSE

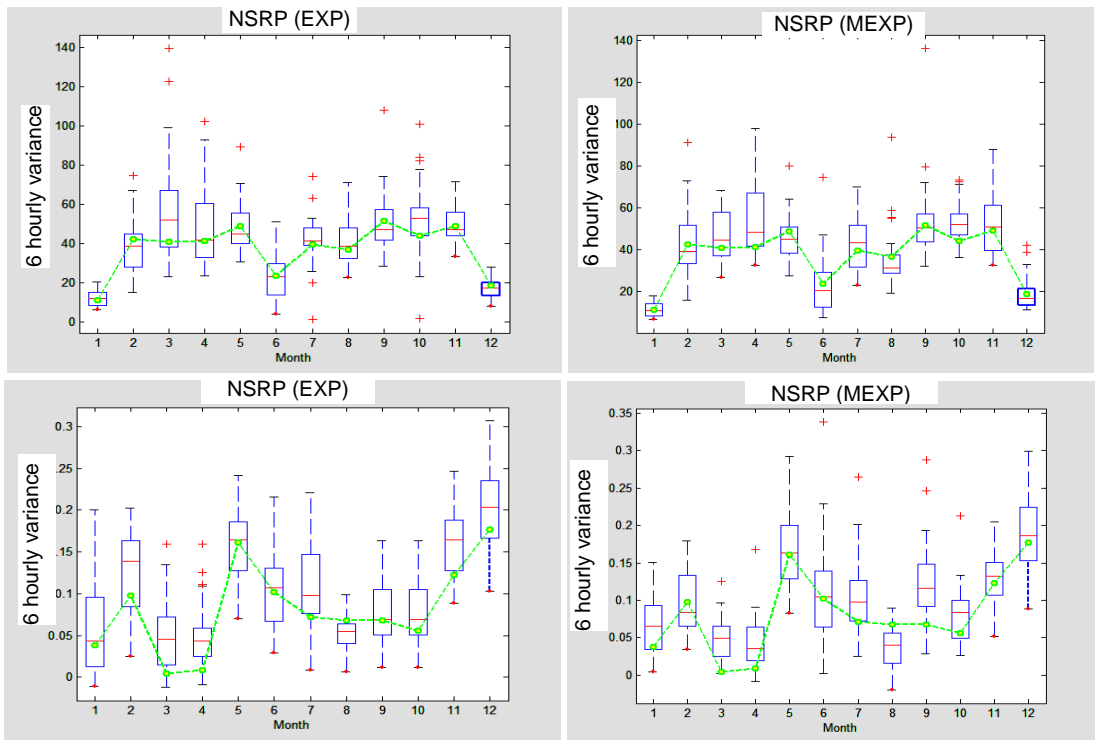


Figure 4: Comparison between observed and simulated properties of six hourly series on a monthly basis.

Table 3 presents the RMSE calculated for each property tested and verified further by the box-plots and graphs in Figures 3 to 5. The performances of both models in preserving the properties tested were equally well in most cases. Graphically it is quite difficult to see the difference. However, the NSRP (MEXP) gives smaller RMSE in most of the properties tested. Therefore the NSRP (MEXP) has better ability in preserving most of the properties in the observed rainfall series.

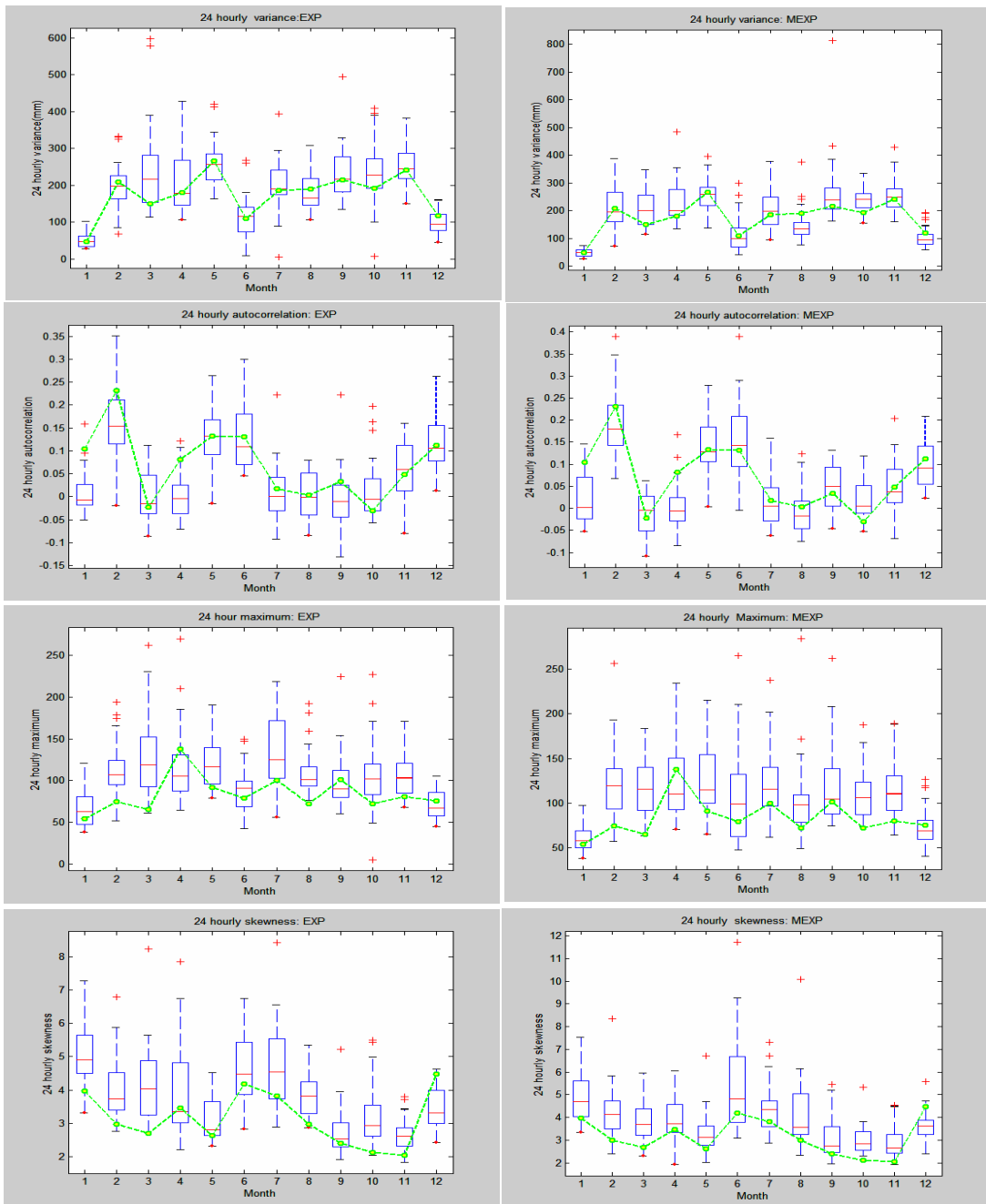


Figure 5: Comparison between observed and simulated properties of 24 hourly series on a monthly basis.

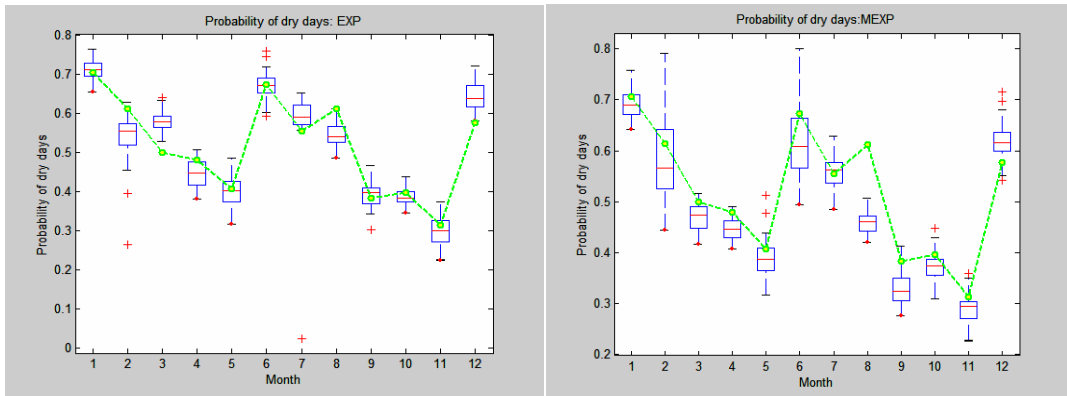


Figure 5: Continued

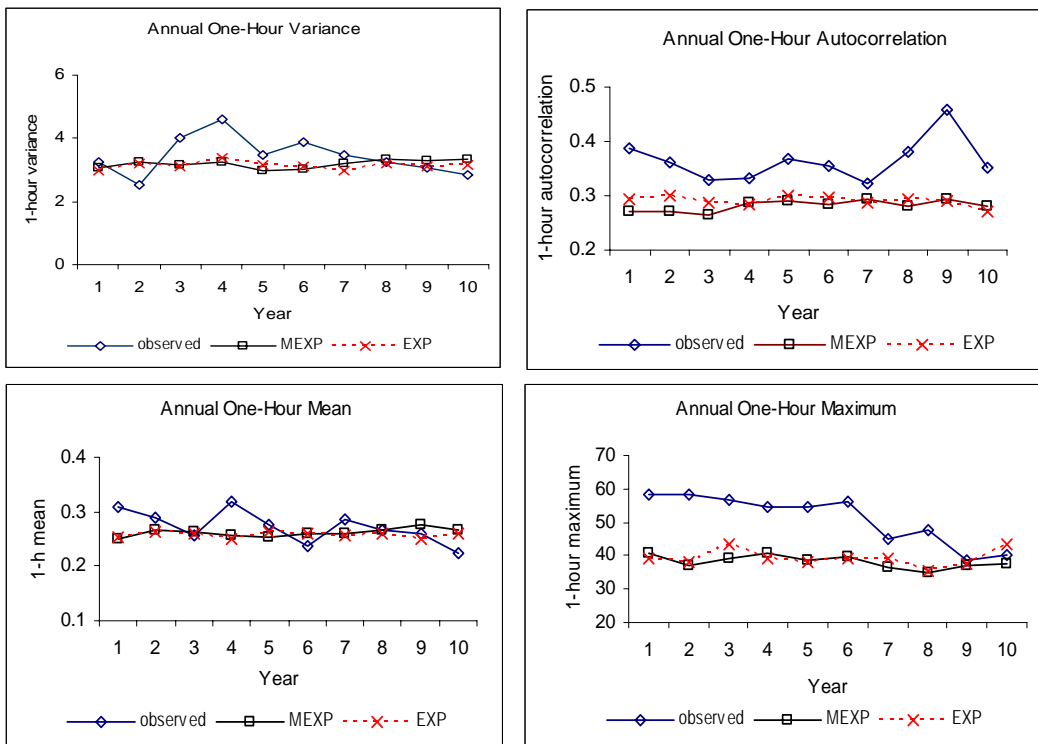


Figure 6a: Comparison between observed and simulated properties of hourly series on an annual basis.

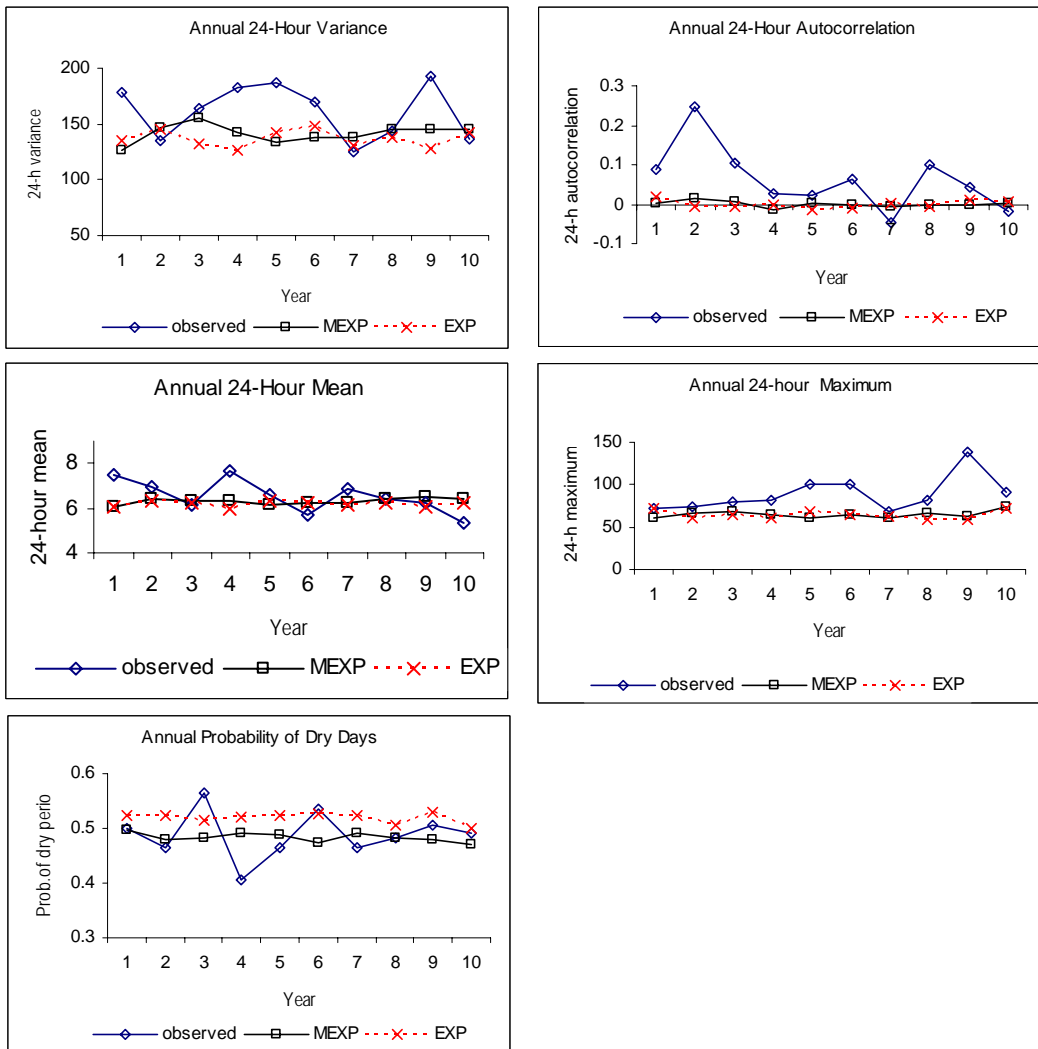


Figure 6b: Comparison between observed and simulated properties of 24 hourly series on an annual basis.

4.0 Conclusion

The performances of the NSRP (MEXP) model with the rain-cell intensity of mixed-exponential distribution and the NSRP(EXP) model with rain-cell intensity of exponential distribution were compared. Both models seemed to perform equally well in preserving the statistical and the physical properties. The differences in the performance of the two models were insignificant in some of the properties tested. However, based on the RMSE value for each property, the performance of the NSRP (MEXP) is better than the NSRP(EXP) in most of the properties tested. The NSRP

(MEXP) model demonstrates better ability in describing the seasonal trend of statistical characteristics at different time-scales. Hence, the NSRP (MEXP) model performed better in preserving the statistical and the physical properties at various time scales.

Acknowledgement

We thank the Department of Irrigation and Drainage for providing rainfall data. This study was funded by the Ministry of Science, Technology and Innovation (MOSTI). Support rendered by the Research Management Unit (RMC) UTM is highly appreciated.

References

- Cowpertime, P.S.P. (1991) Further developments of the Neyman-Scott clustered point process for modeling rainfall. *Water Resources Research*, 27:1431-1438.
- Cowpertime, P.S.P., O'Connell, P.E., Metcalfe, A.V. and Mawdsley, J.A. (1996) Stochastic point process modeling of rainfall: 1. Single-site fitting and validation. *Journal of Hydrology*, 175:17-46.
- Cowpertime, P.S.P., O'Connell, P.E., Metcalfe, A.V. and Mawdsley, J.A. (1996) Stochastic point process modeling of rainfall: 1. Regionalization and disaggregation. *Journal of Hydrology*, 175: 47-65.
- Duan, Q., Sorooshian, S. and Gupta, V. (1992) Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4):1015-1031.
- Entekhabi, D., Rodriguez-Iturbe, I., and Eagleson, P.S. (1989) Probabilistic representation of the temporal rainfall process by a modified Neyman-Scott rectangular pulses model: Parameter estimation and validation. *Water Resources Research*, 25:295-302.
- Fadhilah Y., Zalina MD., Nguyen, V-T-V, Suhaila S. and Zulkifli Y. (2007) Fitting the best fit distribution for the hourly rainfall series in the Wilayah Persekutuan. *Jurnal Teknologi*, (accepted)
- Haan, C.T., Allen, D.M. and Street, J.O. (1976) A Markov Chain Model of daily rainfall. *Water Resources Research*, 12(3): 443-449.
- Katz, R.W. (1977) Precipitation as a chain-dependent process. *Journal of Applied Meteorology*, 16:671-676.
- Nordila A., Zulkifli Y. and Zalina MD. (2006) Characterization of convective rain in Klang Valley. In Proceedings of National Conference on Water for Sustainable Development towards a Developed Nation by 2020, Port Dickson, Negeri Sembilan, Malaysia, 13-14 July 2006.
- Rodriguez-Iturbe, I., Cox, D.R. and Isham, V. (1987a) Some models for rainfall based on stochastic point processes. *Proc. Roy. Soc. London Ser. A*, 410(1839):269-288.
- Rodriguez-Iturbe, I., Febres De Power, B. and Valdes, J. (1987b). Rectangular pulses point process models for rainfall: analysis of empirical data. *Journal of Geophysical Research*, 92(D8): 9645-9656.
- Todorovic, P. and Woolhiser, D. (1975) A stochastic model of n-day precipitation. *Journal of Applied Meteorology*, 14(1): 17-24.
- Woolhiser, D.A. and Roldan, J. (1982) Stochastic Daily Precipitation Models 2. A comparison of Distributions of Amounts. *Water Resources Research*, 18(5):1461-1468.