RAINFALL PREDICTIVE MODELS FOR BUILDING SIMULATION II –RAINFALL ESTIMATION

Qinglin Meng[1], Yufeng Zhang[1] and Zhigang Wang[2]

[1] Building Environment Energy Laboratory (BEEL), State Key Laboratory of Subtropical Building Science, South China University of Technology, Guangzhou, 510640, P.R.China
[2] Dongguan Daye Construction Technology Consulting Co. Ltd, Dongguan, 523000, P.R.China

ABSTRACT

Rainfall simulation is key problem for the evaluation of the energy saving potential of passive cooling by natural rainfall evaporation in humid and hot areas. This paper used a set of monthly, daily and hourly rainfall data in Guangzhou Area to identify the statistic feathers and distribution of the monthly, daily and hourly rainfall. A one-order AR model was found to be suitable for monthly rainfall estimation by AIC criterion judgment. Using a Gamma distribution model and monthly rainfall, daily rainfall was estimated by stochastically assignment. Using a Beta distribution model and daily rainfall, hourly rainfall was estimated by stochastically assignment. A comparison of statistics features of simulated results to that of measured results validates the estimation model.

KEYWORDS

Rainfall predictive model, monthly rainfall estimation, daily rainfall estimation, hourly rainfall estimation

INTRODUCTION

The paper presents the second part of the study on the rainfall predictive model: rainfall estimation. The background and the methodology of the study are presented in another paper on the first part of the study and omitted in the present paper.

MONTHLY RAINFALL ESTIMATION

Statistic feature of monthly rainfall

Rainfall changes from month to month and Table 1 lists the statistics feature of monthly rainfall for each month based on the data from 1971 to 2003.

Table 1 Statistic feature of monthly rainfall (to be continued)

MONTH	SAMPLE SIZE	MEAN VALUE（0.1MM）	STANDARD DEVIATION
Jan	33	416.00	416.70
Feb	33	660.70	787.20
Mar	33	846.00	590.60
Apr	33	1940.90	957.70
May	33	2853.80	1486.20
Jun	33	2828.10	1146.90
Jul	33	2356.50	1219.00
Aug	33	2327.50	1096.20
Sep	33	1814.80	900.80
Oct	33	845.40	771.50
Nov	33	348.00	406.10
Dec	33	328.50	366.70

(continued)

MONTH	SKEWNESS	KURTOSIS	MIN VALUE（0.1MM）	MAX VALUE（0.1MM）
Jan	2.15	5.97	0.00	1995.00
Feb	2.01	3.85	0.00	3270.00
Mar	1.86	5.69	34.00	3116.00
Apr	0.49	0.39	339.00	4187.00
May	1.83	4.12	1081.00	7971.00
Jun	0.89	0.38	1146.00	6010.00
Jul	1.35	2.99	610.00	6539.00
Aug	0.48	-0.42	471.00	4914.00
Sep	0.64	-0.70	567.00	3678.00
Oct	0.89	0.36	0.00	3043.00
Nov	1.43	1.11	0.00	1447.00
Dec	1.69	3.19	0.00	1445.00

It can be seen that monthly rainfall is more unsteady compared with yearly rainfall. Mean value of monthly rainfall from April to September is comparably abundant, which accounts for 80% of annual rainfall. Skewness of monthly rainfall is positive for all months, and skewness of monthly rainfall in January, February, November and December is comparably smaller because of less rainfall in those months.

Autocorrelation of monthly rainfall series was analyzed and it was found that the absolute value of one-order autocorrelation coefficient was big for most of months, which indicates that the series are dependent with each other. None of one-order autocorrelation coefficients equals 0, and amplitude of autocorrelation coefficient sequence attenuates abruptly. Therefore, AR (P) model is chosen to estimate monthly rainfall (Salas et al. 1991). The coefficients of AR(P) model were estimated by distance evaluation method, which has been widely used in hydraulics.

Before the analysis of model, logarithm transformation was performed and the pattern of monthly rainfall sequence is transferred to the one close to normal distribution.

After the transformation, absolute values of skewness and Kurtosis decrease significantly, which is more close to normal distribution.

AIC criterion for model order determination

Model order determination is the key problem for the establishment of AP(P) model. AIC criterion, which is proposed by Akaike in 1974, is used to calculate the order of the model. The criterion can be expressed as:

$$\text{AIC} = n \ln \hat{\sigma}_{\varepsilon}^{2} + 2(p+1) = \min \quad (1)$$

where n is fixed, and $\hat{\sigma}_{\varepsilon}^{2}$ is correlated with P.

When the equation (1) reaches a minimum value, the sequence can be regarded as AR(P) model. For monthly rainfall sequence, the values of AIC for AR (1) to AR (7) were calculated by SAS and the results are listed in Table 2.

Table 2 AIC value for AR model in different order (to be continued)

MONTH	AR(1)	AR(2)	AR(3)
Jan	123.35	125.08	123.46
Feb	124.23	125.81	127.43
Mar	83.12	84.34	86.26
Apr	65.05	63.10	63.45
May	43.93	44.72	46.26
Jun	35.29	37.17	38.77
Jul	51.31	41.54	43.42
Aug	53.25	55.21	56.39
Sep	51.99	53.22	52.99
Oct	148.20	150.14	151.76
Nov	144.16	145.64	147.62
Dec	138.77	138.78	139.85

(continued)

MONTH	AR(4)	AR(5）	AR(6）	AR(7）
Jan	124.56	126.30	128.23	130.21
Feb	128.94	130.60	129.88	130.60
Mar	88.26	89.96	91.70	93.24
Apr	62.06	63.29	63.92	65.82
May	48.24	50.16	50.54	52.24
Jun	40.77	42.68	44.16	46.05
Jul	43.09	45.07	46.71	48.28
Aug	54.99	56.46	58.39	59.98
Sep	54.91	56.88	57.48	58.51
Oct	152.28	147.54	149.12	149.97
Nov	149.42	150.92	151.98	153.28
Dec	139.10	140.49	142.01	140.62

It can be seen that the value of AIC for AR(P) model increases with the model order. According to the AIC criterion, AR(1) model is suitable to describe the monthly rainfall for all 12 months.

The monthly rainfall estimation model can be expressed as:

$$Y_{i,j} = u_j + \rho_{1,j}(Y_{i-1,j} - u_j) + \varepsilon_j \quad (2)$$

where:

$Y_{i,j}$ —Logarithm transformation value of rainfall in the j-th month of the i-th year;

u_j — Mean of logarithm transformation value of rainfall in the j-th month for all years;

$\rho_{1,\ j}$ —One-order autocorrelation coefficient for the j-th month rainfall sequence after logarithm transformation;

ε_j — Independent stochastic item for the j-th month.

Coefficient estimation

1、Mean value

Mean value for each monthly sequence after logarithm transformation was calculated by the following equation:

$$\bar{u}_j = E(Y_j) = \frac{1}{n}\sum_{i=1}^{n} Y_{i,j} \quad (3)$$

2、One-order autocorrelation coefficient $\rho_{1,\ j}$

One-order autocorrelation coefficient were calculated by the following equations:

$$\hat{\gamma}_k = \frac{1}{n}\sum_{i=1}^{n-k}[(X(j)_{i+k} - u(j)(X(j)_t - u(j)] \quad (4)$$

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (5)$$

3、Independent stochastic item ε_j

For that monthly rainfall sequence is closed to normal distribution after logarithm transformation, the independent stochastic item ε_j can be substituted by the normal stochastic function with zero mean value and variance σ_j^2。

Therefore, AR(1) model can be written as:

$$Y_{i,j} = \bar{u}_j + \hat{\rho}_{1,j}(Y_{i-1,j} - \bar{u}_j) + \xi \cdot s_j \quad (6)$$

where:

ξ —Standard normal distribution stochastic function

s_j —Standard deviation of monthly rainfall sequence after logarithm transformation

Validation of monthly rainfall estimation model

Mean values of monthly rainfall in 33 years were used as the initial values for $Y_{i,j}$, and 10000 times of simulation were performed. The sequences which annual rainfall exceeds the maximum one in the 33 years were excluded and the simulated results are shown in Table 3. The measured results are shown as well for comparison.

Table 3 AIC value for monthly rainfall estimation model in different order (to be continued)

MONTH	MEAN VALUE (0.1MM)		STANDARD DEVIATION		SKEWNESS	
	Sim	Mea	Sim	Mea	Sim	Mea
Jan	36.0	41.6	41.8	41.7	1.77	2.15
Feb	53.6	66.1	65.6	78.7	1.88	2.01
Mar	81.9	84.6	62.0	59.1	1.34	1.86
Apr	174.6	194.1	88.9	95.8	0.7	0.49
May	279.6	285.4	128.8	148.6	1.08	1.83
Jun	274.4	282.8	101.5	114.7	0.67	0.89
Jul	231.9	235.7	115.4	121.9	1.04	1.35
Aug	217.5	232.8	97.9	109.6	0.65	0.48
Sep	165.3	181.5	72.8	90.1	0.63	0.64
Oct	50.2	84.5	67.0	77.2	1.84	0.89
Nov	21.3	34.8	29.3	40.6	1.99	1.43
Dec	22.3	32.9	29.9	36.7	2	1.69

(continued)

MONTH	KURTOSIS		MIN VALUE (0.1MM)		MAX VALUE (0.1MM)	
	Sim	Mea	Sim	Mea	Sim	Mea
Jan	2.66	5.97	0.1	0	199.3	199.5
Feb	3.24	3.85	0.1	0	326.1	327
Mar	1.48	5.69	3.0	3.4	309.9	311.6
Apr	-0.24	0.39	12.1	33.9	418.4	418.7
May	1.18	4.12	46.1	108.1	794.5	797.1
Jun	0.12	0.38	61.8	114.6	600.5	601
Jul	0.88	2.99	313	61.0	649.3	653.9
Aug	-0.22	-0.42	185	471	491.0	491.4
Sep	-0.23	-0.7	238	567	367.4	367.8
Oct	2.77	0.36	0	0	303.4	304.3
Nov	3.57	1.11	0	0	144.7	144.7
Dec	3.66	3.19	0	0	144.4	144.5

Note: 'Sim' means simulated result and 'Mea' means measured results.

It can be seen that there is no significant difference in main statistic feature between the simulated results and the measured results. The differences in Jan, Feb, Nov and Dec are comparably district because of less rainfall in those months. For that independent stochastic item in monthly rainfall estimation model is substituted by standard normal distribution stochastic function, distribution pattern of the simulated results is different from that of the measured results, especially for strewness and kurtosis. The range of monthly rainfall of the simulated results coincides well with that of measured results, such as maximum value. Based on the above analysis, the monthly rainfall estimation model is believed acceptable and suitable for the estimation of monthly rainfall in Guangzhou area. Further research will be performed to study the effect of monthly rainfall in the normal range on buildings energy consumption. If the effect does not significantly change with monthly rainfall, the mean value of 30 years observations can be used directly instead of monthly rainfall estimation model.

DAILY RAINFALL ESTIMATION

Daily rainfall distribution

Literatures (Allan and Haan 1975, Woolhiser and Roldan 1982, Coe and Stern 1982, Richardson et al. 1984, Chapman 1997) show that daily rainfall sequence in a month follows Weibull distribution, exponential distribution, Gamma distribution or skewed normal distribution respectively. Liao (2004) found that the daily rainfall in most part of China followed the two parameters Gamma distribution well.

Parametric estimation

Maximum likelihood method was used to estimate the parameters of the model and the results are shown in Figure 1 and 2. It can be seen that the value of shape parameter of Gamma distribution model is comparably steady with mean value of 0.548, while scale parameter varies greatly in a range from 9.2 to 35.8, which indicates the significant variation of daily rainfall.

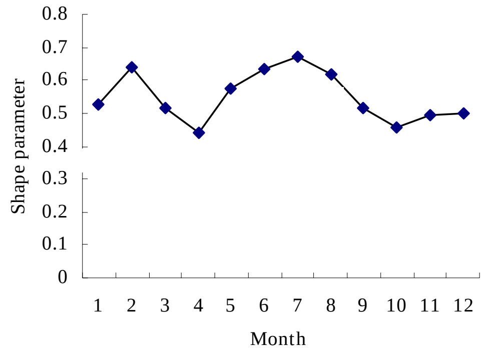

Figure 1 Monthly variation of shape parameter

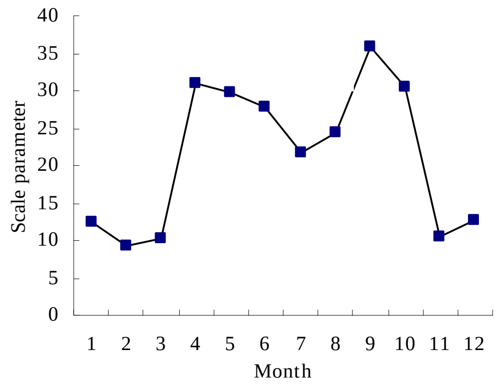

Figure 2 Monthly variation of scale parameter

Daily rainfall estimation

Daily rainfall has strong stochastic features and is less related to other weather elements. Therefore, the daily rainfall can be distributed stochastically based on the Gamma distribution model with given parameters. Supposing the rainfall of a month is M, and the number of rainy days in the month is n. Stochastic sequence $y_1, y_2, ... y_n$ which follows the Gamma distribution model of the month was derived by randomization and the daily rainfall estimation model can be expressed as:

$$x_i = M.\frac{y_i}{\sum_{i=1}^{n} y_i} \quad (i = 1,2,...n) \tag{7}$$

The simulated results were obtained by the equation (7) and compared with the measured results in Table 4.

Table 4 Comparison of simulated results and measured results (to be continued)

MONTH	MEAN NUMBER OF RAINY DAYS		MIN VALUE (MM)		MAX VALUE (MM)	
	Sim	Mea	Sim	Mea	Sim	Mea
Jan	6.2	4	0.1	2.4	58.3	12.9
Feb	8.5	8	0.1	0.2	61.0	19.3
Mar	12.6	15	0.1	0.1	63.6	19.8
Apr	14.3	14	0.1	0.9	134.0	48.0
May	14.7	17	0.1	0.3	157.0	74.4
Jun	17.2	15	0.1	0.1	129.5	87.4
Jul	16.3	12	0.1	1.4	86.5	108.2
Aug	16.5	12	0.1	1.0	239.0	75.8
Sep	10.3	13	0.1	1.0	156.4	49.8
Oct	5.6	6	0.1	0.9	110.7	43.4
Nov	5.8	6	0.1	0.3	116.3	19.2
Dec	5.4	3	0.1	4.0	46.1	20.2

(continued)

MONTH	MEAN VALUE (MM)		STANDARD DEVIATION	
	Sim	Mea	Sim	Mea
Jan	6.6	8.0	11.03	5.11
Feb	5.9	8.1	8.82	7.37
Mar	5.2	5.3	8.72	6.49
Apr	13.7	14.1	24.08	14.10
May	17.1	15.8	23.51	19.87
Jun	17.7	20.3	20.42	23.90
Jul	14.6	19.0	17.13	29.49
Aug	15.1	21.2	24.83	23.25
Sep	18.5	24.3	27.33	18.82
Oct	13.9	12.3	20.58	15.68
Nov	5.2	4.8	13.79	6.93
Dec	6.3	14.5	10.20	7.15

Note: ‘Sim’ means simulated result and ‘Mea’ means measured results.

The Table shows that the mean number of rainy days in simulation is close to the mean measured value of ten years and the daily rainfall obtained by simulation lies between the maximum and minimum rainfall obtained by measurement in ten years. The comparison indicates that it is feasible to identify the rainy days by distance discriminant analysis, and then estimate the monthly rainfall by AR(1) model, and finally distribute the monthly rainfall to each rainy day stochastically.

HOURLY RAINFALL ESTIMATION

Hourly rainfall distribution

In Guangzhou area, rainfall from April to September accounts for about 80% of annual rainfall. In those months, most rainfall results from strong convection of the airflow, which intensity is great and duration is short. In other months, rainfall is less in amount and most of the rainfall is frontal rain, which intensity is low and duration is long due to the advance and retreat of frontal cloud. Based on this feature of rainfall intensity in different season, a whole year is divided into rainy season and non-rainy season for hourly rainfall study in Guangzhou area.

The regularity of hourly rainfall distribution was analyzed by Quantile-Quantile chart of different distribution function for rainy season and non-rainy season respectively and it was found that the measured data sequence of hourly rainfall followed Gamma and Beta distribution well (see figure 3 and 4). Because the data fit the Beta distribution line better, Beta distribution function was used to estimate hourly rainfall percentage in the present study.

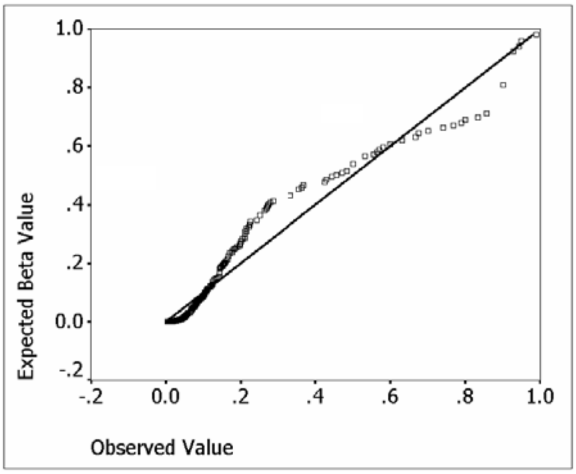

Figure 3 Quantile-Quantile chart of Beta distribution for rainy season

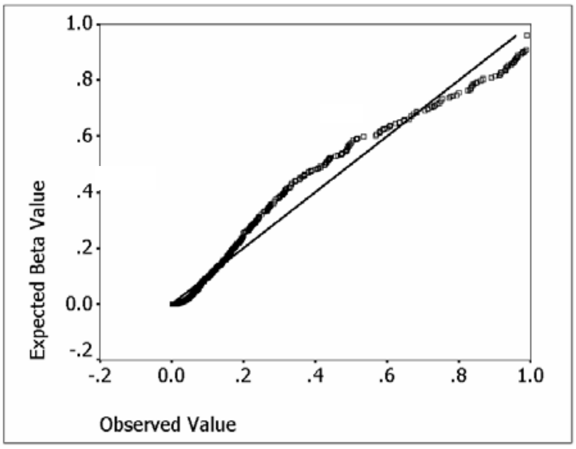

Figure 4 Quantile-Quantile chart of Beta distribution for non-rainy season

Maximum likelihood method was used to estimate the parameters in Beta distribution function and the results are shown in Table 5.

Table 5 Parametric estimation of Beta distribution function

PARAMETERS	RAINY SEASON	NON-RAINY SEASON
a	0.456	0.480
b	1.761	1.638

Hourly rainfall estimation

Similar with daily rainfall, hourly rainfall has strong stochastic features and is less related to other weather elements. Therefore, Parameters in Beta distribution model were generated stochastically.

Supposing the rainfall in a day is R, and the number of rainy hours in the day is n. Stochastic sequence $y_1, y_2, ... y_n$ which follows the Beta distribution model of the season was derived by randomization and then the hourly rainfall estimation model can be expressed as:

$$x_i = R.\frac{y_i}{\sum_{i=1}^{n} y_i} \quad (i = 1,2,...n) \tag{8}$$

Comparison of simulated results and measured results is shown in Table 6. From the viewpoint of statistic features, simulated results coincide with measured results well, which validates the estimation method.

Table 6 Coefficient estimation and error of January

		ANNUAL RAINFALL (0.1MM)	MAX VALUE OF HOURLY RAINFALL (0.1MM)	NUMBER OF RAINY HOURS (H)
Simulated results	No.1	18354	350	698
	No.2	18343	625	717
	No.3	18356	411	727
	No.4	18338	495	705
	No.5	18349	528	715
Measured results	2003	13830	433	590
	2004	16589	468	626

CONCLUSIONS

As an important component of weather database, hourly rainfall model is indispensable for evaporation cooling effect evaluation. In the present paper, the statistic feathers and distribution of the monthly, daily and hourly rainfall was identified and the following conclusions were drawn:

1. According to an autocorrelation analysis of monthly rainfall series, a one-order AR model is found to be suitable for monthly rainfall estimation by AIC criterion judgment.
2. Distribution of daily rainfall month by month shows a Gamma distribution model agrees well with daily rainfall distribution. Using a Gamma distribution model and monthly total rainfall, daily rainfall is estimated by stochastically assignment.
3. Distribution patterns of hourly rainfall percentage both in the rainy season and non-rainy season coincide well with the Beta distribution. Using a Beta distribution model and daily total rainfall, hourly rainfall is estimated by stochastically assignment.

ACKNOWLEDGMENT

The experiment was supported by National Natural Science Foundation of China (Project No. 50538040).

REFERENCES

Allan DM and Haan CT. 1975. “Stochastic simulation of daily rainfall” Research Report Water Resources Institute, University of Kentucky.

Chapman TG. 1997. “Stochastic models for daily rainfall in the Western Pacific” Mathematics and computers in simulation. 43: 351-358.

Coe R and Stern RD. 1982. “Fitting models to daily rainfall data” J. Appl. Meteorol. 21:1024–1031.

Liao Y, Zhang Q and Chen D. 2004. “Rainfall simulation for weather data generator in China” Journal of Geography. 59(5): 689~698. (in Chinese)

Richardson CW, Wright DA and WGEN. 1984. “A model for generating daily weather variables” ARS-8, US Department of Agriculture, Agricultural Research Service, Springfield, VA.

Salas JD et al. 1991. “Stream flow drought statistics by stochastic simulation” Proceeding of the United States-PRCRilateral Symposium on droughts and rid-region on Hydrology.

Woolhiser DA and Roldan J. 1982. “Stochastic daily precipitation models 2. A comparison of distribution of amounts” Water Resour. Res. 18: 1461-1469.