Random Simulation Study on Flood Process of Flood Control Engineering System in Nanning

Rong-yong Ma[1,a], Yi Du[2,b] and Jia-li Lin[3,c]

[1]College of Civil Engineering and Architecture，Guangxi University，China

[2]College of Civil Engineering and Architecture，Guangxi University，China

[3]Guangxi Key Laboratory of Disaster Prevention and Engineering Safety，Guangxi，China

[a]mry57@gxu.edu.cn, [b]duyi19930922@163.com, [c]563435554@qq.com

Keywords: Flood control engineering system; Flood random simulation; SAR(1) model; Characteristic parameter

Abstract: In order to better carry out the research on the flood control capacity of the flood control engineering system in Nanning, using the SAR(1) model to simulate various 50 thousands of flood process of Baise reservoir inflow and Baise-Nanning interval, and analyze the independence and applicability of the model. The independent autocorrelation coefficient and the autocorrelation analysis method are used to test the independence of the random term of various sections of the model. The results show that assumption that residual series of the model is independent are reasonable. By comparing the characteristic parameters of each section of the simulated and measured flood process, it is found that their coincidence degree is high and the simulated flood sequence well maintains the statistical characteristics of the measured flood, which fully shows that the SAR(1) model is suitable for random simulation of flood in Baise reservoir inflow and Baise-Nanning interval.

Introduction

The flood process line is an important data for hydrological analysis. It is generally necessary to have a sufficient amounts of flood process lines to carry out flood control calculations. However, due to the limited data of hydrological station observation, the data length of the general flood process line is insufficient, and it is difficult to carry out engineering analysis and calculation. Therefore, the establishment of random simulation model of hydrological process has become one of the important contents of hydrological process analysis. After a random simulation of the flood process is proposed as a new way of flood control safety design, a number of studies have been carried out at home and abroad, and a variety of flood random simulation models have been generated[1], such as autoregressive model, typical solution model, nonlinear model and so on. In recent years, emerging nonparametric models[2], wavelet analysis[3], random simulation based on Copula function[4-5] and other methods have been applied to the field of random flood simulation, which greatly enriched the theory and method of flood random simulation.

Thereinto, hydrological simulation model and method of the regressive sort is relative mature, can be applied in large and medium-sized watershed, has a strong applicability. Therefore, this study select the seasonal first order autoregressive model[6], which can reflect the seasonal characteristics of floods variation among regression models.

Model introduction

Model calculation method and thinking[7]

Assume that the flood process series $Q_{t,w}$ is expressed by the matrix:

$$\left\{Q_{t,w}\right\}_{n\times m}=\begin{bmatrix} Q_{1,1} & Q_{1,2} & Q_{1,3} & \mathbf{L} & Q_{1,m} \\ Q_{2,1} & Q_{2,2} & Q_{2,3} & \mathbf{L} & Q_{2,m} \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ Q_{t,1} & Q_{t,2} & Q_{t,3} & \dots & Q_{t,m} \\ \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} & \mathrm{M} \\ Q_{n,1} & Q_{n,2} & Q_{n,3} & \dots & Q_{n,m} \end{bmatrix} \tag{1}$$

Thereinto, $t=1,2,\mathbf{L},n$, n is the number of years; $w=1,2,\mathbf{L},m$, m is the number of sections.

The seasonal first order autoregressive model SAR(1) expressed in the original series $Q_{t,w}$ can be written as:

$$Q_{t,w}=j_{1,\mathrm{w}}Q_{t,w-1}+e_{t,w} \tag{2}$$

Thereinto, $j_{1,\mathrm{w}}$ is the autoregressive coefficient of the section w; $e_{t,w}$ is the random number of the normal distribution.

To eliminate the seasonal effects of mean m_w and variance s_{w}^{2}, the original series $Q_{t,w}$ should be standardized:

$$Y_{t,w}=\frac{Q_{t,w}-m_w}{S_w} \tag{3}$$

Thereinto, m_w, S_w are the mean and mean variance of the original section w respectively.

And then the standardized series $Y_{t,w}$ becomes the standard normal series by W-H inverse transformation :

$$Z_{t,w}=\frac{6}{C_{s,\mathrm{w}}}\left[\left(\frac{C_{s,w}}{2}Y_{t,w}+1\right)^{\frac{1}{3}}-1\right]+\frac{C_{s,w}}{6} \tag{4}$$

Thereinto, $C_{s,w}$ is the skew coefficient of the original series of intersection w.

After the original series is standardized and normalized, the SAR(1) model established for the standard normalization series is:

$$Z_{t,w}=r_{1,\mathrm{w}}Z_{t,w-1}+x_{t,w} \tag{5}$$

Thereinto, $r_{1,\mathrm{w}}$ is the autoregressive coefficient, which is estimated by the first-order autocorrelation coefficient of the corresponding section,that is

$$r_{1,\mathrm{w}}=\mathrm{r}_{1,\mathrm{w}} \tag{6}$$

$x_{t,w}$ is an independent random variable whose mean is 0 and variance is $s_{x,w}^{2}$.

$$s_{x,w}^{2}=\sqrt{1-r_{1,\mathrm{w}}^{2}} \tag{7}$$

Therefore:

$$Z_{t,w} = r_{1,\mathrm{w}} Z_{t,w-1} + \sqrt{1-r_{1,\mathrm{w}}^2}\, e_{t,w} \tag{8}$$

Finally, by the inverse transformation of the standardization and normalization, all $Z_{t,w}$ are conversed into a simulated flood process $Q_{t,w}$:

$$Y_{t,w} = \left(\frac{2}{C_{s,w}}\right)\left(1+\frac{Z_{t,w}C_{s,w}}{6}-\frac{C_{s,w}{}^2}{36}\right)^3 - \frac{2}{C_{s,w}} \tag{9}$$

$$Q_{t,w} = m_w + Y_{t,w} S_w \tag{10}$$

Formula (9) and (10) are the seasonal first-order autoregressive SAR(1) models finally obtained.

Analysis of model residual feature and applicability analysis

Comprehensive autocorrelation coefficient test

①According to the established random simulation model, the random term e_t of the sample series is obtained.

$$\varepsilon_t = \left(Z_{t,w} - r_{1,\mathrm{w}} Z_{t,w-1}\right) / \sqrt{1-r_{1,\mathrm{w}}^2} \tag{11}$$

②The autocorrelation coefficients $(k = 1, 2, \mathbf{L}, m)$ are calculated according to the series of random items e_t;

$$r_k = \frac{\sum_{t=1}^{n-k}\left(x_{t+k}-\overline{x}\right)\left(x_t-\overline{x}\right)}{\sum_{t=1}^{n}\left(x_t-\overline{x}\right)^2} \tag{12}$$

Thereinto, for value of m: when $n > 50$, $m < n/4$; when $n < 50$, $m = n/4$ or $m = n-10$.

③Construct statistic

$$Q = n\sum_{k=1}^{m} r_{k,e}^2 \tag{13}$$

If e_t is independent random series, Q gradually obeys c^2 distribution whose degree of freedom is m-p-q;

④Given a significant level of α, obtain $c_a{}^2$ by checking c^2 distribution table, if $Q \le c_a{}^2$ it can be judged that e_t is independent, and vice versa is not independent.

Autocorrelation analysis

First, reversely calculate random item e_t series by the model, calculate the autocorrelation coefficient $r_{k,e}$, and then calculate the upper and lower allowable limits whose tolerance is 95% by the formula. The upper and lower tolerance limits and autocorrelation coefficients are plotted on the same graph to determine whether the autocorrelation coefficient falls in the upper and lower limits. If yes, it indicates that e_t is independent ,and vice versa is not independent. Thereinto, the formula for calculating the allowable limit is:

$$r_k\left(a=5\%\right)=\frac{-1\pm1.96\sqrt{n-k-1}}{n-k} \qquad (14)$$

Applicability analysis mainly analyzes whether the results the model simulated maintain statistical characteristics of the measured series. If the statistical parameters of the simulation results have little difference with the corresponding statistical parameters of the measured series, the model results can be regarded as inferred totality.

Flood simulation

The SAR(1) model is used to simulate the flood of Baise reservoir and Baise-Nanning interval, and make residual independence test and suitability analysis of the model, due to space restrictions, the following takes simulation of inflow flood of Baise reservoir only as an example to introduce.

In the collected flood data of 46 years in Baise reservoir, every year select flood process whose duration is 7d ,take 3h for a section and divide it into 56 sections. There are exceptionally large flood year, the general medium flood year and small flood years in the selected flood process data of Baise reservoir. So the data is a good representative. MATLAB software programming is used to get 50,000 flood process of Baise reservoir.

Simulation of flood simulation model of Baise reservoir

The flood simulation of Baise reservoir is carried out by the way about sections. First obtain the random term from the model, and then calculate the autocorrelation coefficient whose delay time is $k(k=1,2,\mathbf{L},m)$. Flood simulation Baise reservoir, due to the number of sections greater than 50, takes m=13. The results are shown in Table 1 and Table 2.

Table 1 Randomness of baise reservoir flood simulation model

section	e_t	section	e_t	section	e_t	section	e_t
1	1.16160	15	-0.02827	29	-0.24867	43	-0.90042
2	0.71557	16	-1.01840	30	-0.84084	44	-0.09350
3	0.47132	17	1.97420	31	0.42810	45	-0.93018
4	-0.52042	18	1.01740	32	-0.01269	46	-0.79567
5	-0.09795	19	-0.57583	33	1.07210	47	1.14530
6	-0.15945	20	1.28810	34	0.87343	48	-0.28157
7	0.44658	21	0.17754	35	-1.99530	49	-0.04258
8	0.37822	22	-0.34276	36	0.17735	50	0.23605
9	-0.39596	23	-1.05830	37	-0.05637	51	-0.10833
10	-0.23490	24	-1.67300	38	-1.92990	52	0.40390
11	0.62386	25	-0.91167	39	0.28801	53	-0.44096
12	-0.38048	26	-0.87178	40	1.00560	54	1.34790
13	1.05300	27	-0.49123	41	-1.04060	55	1.03720
14	-0.76408	28	-0.39264	42	0.77340	56	-0.51616

Table 2 Autocorrelation coefficient of randomness

Lag time (k)	Autocorrelation coefficient	Lag time (k)	Autocorrelation coefficient
1	-0.02352	8	0.03188
2	0.02065	9	-0.18860
3	0.16418	10	0.03826
4	-0.17111	11	0.04949
5	0.01035	12	-0.05896
6	-0.08547	13	0.01651
7	-0.02864		

（1）Comprehensive autocorrelation coefficient test

By calculation ,obtain that statistic Q=6.143. Take the significance level α=0.05. Get that c_a^2= 21.0261 by checking c^2 distribution table. By comparison, $Q<c_a^2$, it indicates that the model is independent and can be used to describe the flood characteristics.

（2）Autocorrelation analysis

When using autocorrelation analysis to test, it is necessary to determine whether the autocorrelation coefficient falls between 95% upper and lower allowable limits. If so, it can be judged that is independent. Obtain the autocorrelation coefficient by calculation, 95% tolerance of the upper and lower tolerance limit, the results are shown in Table 3. As can be seen from Fig. 1, the autocorrelation coefficient falls within the allowable limit, so the model passes the independence test.

Table 3 The flood simulation model residual independence test of Baise reservoir

Lag time(k)	Autocorrelation coefficient and allowable limit		
	Upper allowable limit	Autocorrelation coefficient	Lower allowable limit
1	0.2437	-0.02352	-0.2801
2	0.2457	0.020648	-0.2828
3	0.2478	0.16418	-0.2855
4	0.2499	-0.17111	-0.2884
5	0.2521	0.010352	-0.2914
6	0.2544	-0.08547	-0.2944
7	0.2567	-0.02864	-0.2975
8	0.2591	0.03188	-0.3008
9	0.2616	-0.1886	-0.3041
10	0.2641	0.038257	-0.3076
11	0.2667	0.049494	-0.3111
12	0.2694	-0.05896	-0.3148
13	0.2721	0.016509	-0.3187

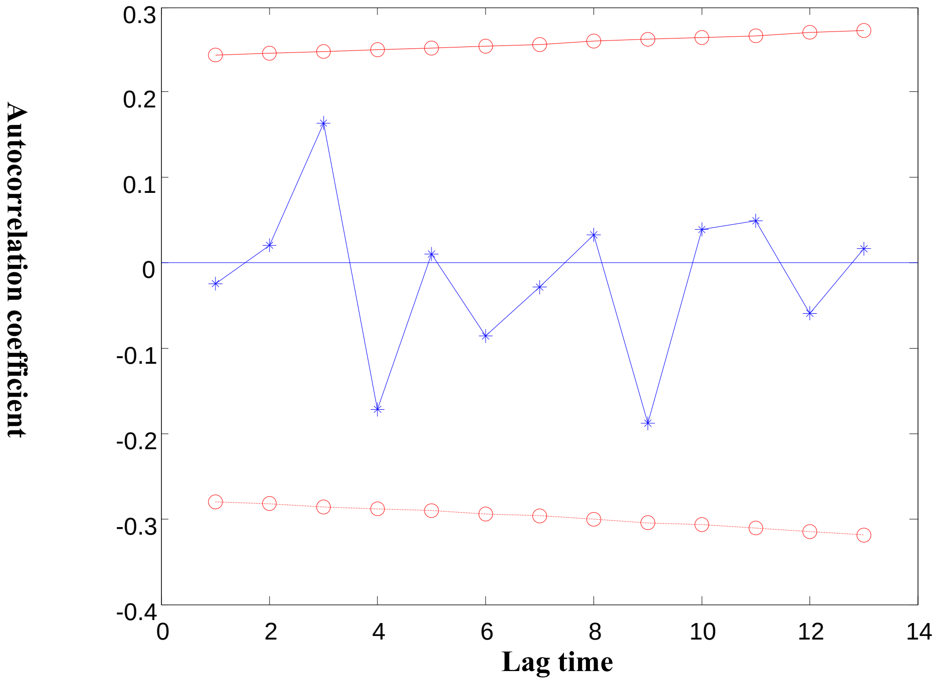

Fig.1 The flood simulation model of random item autocorrelation coefficient figure of Baise reservoir

Applicability Analysis of Flood Simulation Results of Baise Reservoir

In order to test whether the flood generated by the simulation has maintained the statistical characteristics of the measured floods or represent the measured samples well. Calculate the characteristic parameters of various sections of the flood process series and compare them with the measured flood series.

Table 4 comparison of simulated and measured parameters of Baise reservoir

section	mean			mean square error			coefficient of variation			coefficient of skewness		
	measured	simulation	relative error(%)	measured	simulation	relative error(%)	measured	simulation	relative error(%)	measured	simulation	relative error(%)
1	970.72	972.38	0.171	402.65	405.19	0.631	0.415	0.417	0.458	0.724	0.739	2.207
2	1005.30	1007.10	0.179	417.05	419.68	0.631	0.415	0.417	0.458	0.725	0.741	2.205
3	1022.70	1024.50	0.176	424.25	426.93	0.632	0.415	0.417	0.458	0.724	0.740	2.198
4	1040.00	1041.80	0.173	431.46	434.18	0.630	0.415	0.417	0.458	0.725	0.741	2.193
5	1057.40	1059.20	0.170	438.59	441.35	0.629	0.415	0.417	0.458	0.724	0.740	2.195
6	1074.70	1076.50	0.167	445.80	448.61	0.630	0.415	0.417	0.458	0.724	0.739	2.191
7	1092.00	1093.90	0.174	452.96	455.82	0.631	0.415	0.417	0.458	0.725	0.741	2.186
8	1109.30	1111.20	0.171	460.19	463.09	0.630	0.415	0.417	0.456	0.724	0.740	2.187
9	1123.60	1125.30	0.151	453.52	456.29	0.611	0.404	0.405	0.461	0.724	0.739	1.997
10	1137.90	1139.70	0.158	448.72	451.36	0.588	0.394	0.396	0.423	0.729	0.742	1.796
11	1152.10	1153.70	0.139	445.82	448.43	0.585	0.387	0.389	0.442	0.740	0.753	1.789
12	1166.50	1168.30	0.154	444.97	447.38	0.542	0.381	0.383	0.388	0.758	0.770	1.599
13	1180.40	1182.30	0.161	445.95	448.53	0.579	0.378	0.379	0.418	0.782	0.793	1.475
14	1194.70	1196.50	0.151	449.03	451.60	0.572	0.376	0.377	0.426	0.811	0.823	1.462

section	mean			mean square error			coefficient of variation			coefficient of skewness		
	measured	simulation	relative error(%)	measured	simulation	relative error(%)	measured	simulation	relative error(%)	measured	simulation	relative error(%)
15	1208.90	1210.60	0.141	453.89	456.34	0.540	0.375	0.377	0.405	0.846	0.857	1.305
16	1223.10	1224.60	0.123	460.73	462.81	0.451	0.377	0.378	0.329	0.883	0.889	0.655
17	1286.20	1287.80	0.124	485.14	486.84	0.350	0.377	0.378	0.225	0.718	0.723	0.667
18	1349.20	1351.00	0.133	522.93	524.78	0.354	0.388	0.388	0.217	0.664	0.669	0.789
19	1412.30	1414.10	0.127	571.59	573.71	0.371	0.405	0.406	0.240	0.723	0.730	1.015
20	1475.50	1478.10	0.176	628.39	631.52	0.498	0.426	0.427	0.322	0.853	0.864	1.319
21	1538.40	1541.30	0.189	691.55	695.57	0.581	0.450	0.451	0.392	1.011	1.027	1.572
22	1601.40	1604.50	0.194	759.54	764.11	0.602	0.474	0.476	0.411	1.172	1.194	1.868
23	1664.50	1667.80	0.198	830.87	836.10	0.629	0.499	0.501	0.433	1.321	1.351	2.233
24	1727.40	1731.20	0.220	905.04	911.15	0.675	0.524	0.526	0.456	1.453	1.491	2.567
25	1915.10	1918.50	0.178	979.49	986.32	0.697	0.511	0.514	0.516	1.492	1.525	2.259
26	2102.60	2107.40	0.228	1063.10	1070.30	0.677	0.506	0.508	0.453	1.498	1.518	1.322
27	2290.20	2295.40	0.227	1153.70	1161.90	0.711	0.504	0.506	0.480	1.489	1.506	1.169
28	2477.80	2483.70	0.238	1249.90	1259.20	0.744	0.504	0.507	0.500	1.471	1.490	1.264
29	2665.30	2672.20	0.259	1350.60	1361.80	0.829	0.507	0.510	0.570	1.452	1.470	1.288

section	mean			mean square error			coefficient of variation			coefficient of skewness		
	measured	simulation	relative error(%)	measured	simulation	relative error(%)	measured	simulation	relative error(%)	measured	simulation	relative error(%)
30	2852.80	2860.10	0.256	1454.70	1467.10	0.852	0.510	0.513	0.596	1.432	1.451	1.313
31	3040.40	3048.20	0.257	1561.60	1574.90	0.852	0.514	0.517	0.596	1.414	1.433	1.323
32	3227.80	3236.40	0.266	1670.70	1685.60	0.892	0.518	0.521	0.622	1.397	1.419	1.517
33	3073.50	3081.40	0.257	1544.80	1558.80	0.906	0.503	0.506	0.647	1.288	1.307	1.475
34	2919.20	2925.90	0.230	1422.00	1434.20	0.858	0.487	0.490	0.620	1.162	1.176	1.188
35	2764.50	2770.80	0.228	1303.50	1314.20	0.821	0.472	0.474	0.592	1.020	1.033	1.334
36	2610.20	2615.40	0.199	1190.30	1199.10	0.739	0.456	0.458	0.537	0.863	0.876	1.420
37	2455.70	2459.90	0.171	1084.40	1092.20	0.719	0.442	0.444	0.550	0.702	0.714	1.719
38	2301.40	2305.20	0.165	987.73	993.82	0.617	0.429	0.431	0.452	0.552	0.560	1.423
39	2146.80	2149.80	0.140	903.47	908.66	0.574	0.421	0.423	0.432	0.444	0.455	2.655
40	1992.20	1995.80	0.181	835.46	840.29	0.578	0.419	0.421	0.396	0.411	0.423	2.935
41	1949.80	1952.90	0.159	802.33	806.89	0.568	0.411	0.413	0.408	0.383	0.395	3.319
42	1907.20	1909.90	0.142	770.47	774.86	0.570	0.404	0.406	0.428	0.357	0.370	3.709
43	1864.50	1867.20	0.145	740.04	744.17	0.558	0.397	0.399	0.418	0.335	0.348	3.971
44	1822.00	1824.40	0.132	711.01	714.79	0.532	0.390	0.392	0.397	0.319	0.332	4.251

section	mean			mean square error			coefficient of variation			coefficient of skewness		
	measured	simulation	relative error(%)	measured	simulation	relative error(%)	measured	simulation	relative error(%)	measured	simulation	relative error(%)
45	1779.40	1781.90	0.140	684.01	687.32	0.484	0.384	0.386	0.346	0.311	0.323	3.810
46	1736.80	1739.10	0.132	658.96	662.04	0.467	0.379	0.381	0.337	0.315	0.327	3.755
47	1694.20	1696.20	0.118	636.17	639.06	0.454	0.376	0.377	0.333	0.333	0.344	3.222
48	1651.50	1653.50	0.121	615.77	618.64	0.466	0.373	0.374	0.346	0.367	0.380	3.519
49	1615.10	1617.20	0.130	596.22	598.98	0.463	0.369	0.370	0.336	0.351	0.366	4.190
50	1578.80	1580.60	0.114	578.26	581.06	0.484	0.366	0.368	0.366	0.340	0.354	4.180
51	1542.20	1544.20	0.130	562.02	564.70	0.477	0.364	0.366	0.348	0.333	0.346	3.740
52	1505.80	1507.80	0.133	547.61	550.18	0.469	0.364	0.365	0.338	0.333	0.347	4.235
53	1469.20	1471.20	0.136	535.41	537.91	0.467	0.364	0.366	0.329	0.337	0.352	4.247
54	1432.80	1434.90	0.147	525.45	528.19	0.521	0.367	0.368	0.379	0.347	0.363	4.626
55	1396.30	1398.40	0.150	517.72	520.68	0.572	0.371	0.372	0.421	0.360	0.376	4.464
56	1359.80	1361.70	0.140	512.50	515.50	0.585	0.377	0.379	0.446	0.375	0.391	4.191

It can be seen from Table 4 that the simulated flooding process is close to the statistical parameters of the measured flood process and the relative error is less than 5%. The simulation results are better and the flood generated by the model basically maintains the statistical characteristics of the measured flood.

According to the above steps, make simulation on the Baise-Nanning flood. It can be got from the calculation that the statistical amount Q=20.776. Take significance level α=0.05, then c_a^2=21.0261 and meets $Q \le c_a^2$, and the random term e_t is independent. The model passes independence test can be used for flood simulation by independence test. Calculate the autocorrelation coefficient and 95% upper and lower allowable limits of the random items of the sample sequence and all autocorrelation coefficients are between 95% upper and lower allowable limits, so that the random items can be judged to be independent. After counting measure flood process lines, simulating the statistical parameters of various sections of flood process lines, it can be found that the simulated flood sequence are close to the statistical parameters of the measured flood sequence and the relative error is less than 2%, which indicates that the simulated floods maintain the statistical characteristics of the measured floods, and the simulation results are excellent.

Conclusions

In this paper, the SAR(1) model is used to simulate the flood process of Baise reservoir inflow and Baise-Nanning interval by analyzing the measured flood data.

In the simulation calculation process, the independence and applicability of the model are tested by the autocorrelation coefficient method and the autocorrelation analysis method respectively.

In addition, by comparing characteristic parameters of flood process the simulation generated and various sections of measured flood serials, the relative error of the statistical data (mean, mean square error, coefficient of variation and skewness coefficient) of each section is less than 5%, and the coincidence degree is relatively high, which indicates sufficiently that it is reasonable to use SAR(1) model to simulate the inflow flood of Baise reservoir flood and Baise-Nanning interval.

Acknowledgements

This work was financially supported by the National Natural Science Foundation of China (51369005).

References

[1] DING Jing, DENG Yuren, HOU Yu, et al. Advances in Water Science, 1992, 3(1): 45-52. (In Chinese)

[2] YUAN Peng, WANG Wensheng, DING Jing. Journal of Sichuan University: Engineering Science Edition, 2000, 32(1): 82-86. (In Chinese)

[3] WANG Wensheng, YUAN Peng, DING Jing. Journal of Hydraulic Engineering, 2000, 31(11): 43-48. (In Chinese)

[4] XIAO Yi, GUO Shenglian, XIONG Lihua, et al. Journal of Sichuan University: Engineering Science Edition, 2007, 39(2): 55-60. (In Chinese)

[5] ZHANG Tao, ZHAO Chunwei, LUO Wensheng. Engineering Journal of Wuhan University, 2008, 41(4): 1-4. (In Chinese)

[6] WANG Wensheng, JIN Juliang, LI Yueqing, et al: *Hydrological water resources stochastic simulation technology* (Sichuan university press, Chengdu 2007). (In Chinese)

[7] ZHOU Yanlai, GUO Shenglian, LI Tianyuan, et al. Engineering Journal of Wuhan University, 2013, 46(2): 137-142. (In Chinese)