Journal of Hydroscience and Hydraulic Engineering
Vol. 25, No.2 November, 2007, 47-57

UNCERTAINTY OF CALIBRATED MODEL PARAMETER DUE TO TEMPORAL RESOLUTION OF HYDROLOGICAL DATA AND ITS REMOVAL

By

Takahiro YAMAMOTO

Nagaoka University of Technology, Nagaoka, Japan

and

Minjiao LU

Nagaoka University of Technology, Nagaoka, Japan

SYNOPSIS

The objective of this study is to evaluate the effects of temporal resolution of hydrological data on a rainfall runoff analysis. When the time interval of hydrological data, the temporal resolution, is longer than the time scale of the rainfall-runoff system, for example runoff concentration time, the calibrated model parameters may be different from their "true values" and include uncertainty caused by the poor temporal resolution. In order to investigate this uncertainty, rainfall data with temporal resolution about one minute are generated by means of a random cascade model, transformed into discharge data by using a storage-function model with known parameters. Then these data are re-sampled with different temporal resolutions and used to calibrate the parameter of the storage-function model. It is shown that the temporal resolution has significant effects on the calibrated parameter, and the calibrated parameter can be expressed by its true value and the temporal resolution. By using this relationship, the effects due to temporal resolution can be eliminated, and the true value parameter invariant with temporal resolution can be derived. Furthermore, the effects of the rainfall intensity on this relationship are also discussed.

INTRODUCTION

The uncertainties of the rainfall runoff analysis may result from the uncertainties of input data, model structure, model parameters and initial conditions. When the model parameters are calibrated from observed river flow, the uncertainty of the river flow also induces uncertainty through calibration of model parameters. As a real problem in hydrological practice, when the parameters of rainfall runoff model are calibrated from the hourly hydrological data in small or medium sized river basin with flood concentration time of several minutes to a few hours, the calibrated model parameters may include the uncertainty caused by the poor temporal resolution, and differ from their "true values". The selection of proper temporal resolution is practically of great importance. Takasao *et al.* (8) proposed a relationship between drainage area and suitable temporal resolution for the Yodo river basin, Japan. Furthermore the removal of such uncertainties is crucial for most hydrological research, for example, the synthesization of model parameters (e.g. Nagai *et al.* (3)).

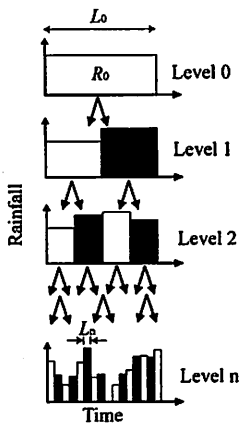

Fig.1 Schematic downscaling.

The objectives of this study are to evaluate the effects of temporal resolution of hydrological data on parameter estimation, and to develop a new methodology to eliminate them and obtain the true parameter.

METHODOLOGY

Random cascade model

Recently, random cascade models based on the fractal property of rainfall field have been widely studied (e.g. Schertzer and Lovejoy (7); Gupta and Waymire (1); Over and Gupta (6)). As for the temporal distribution of rainfall, Olsson (5) showed that the random cascade generator has uniform distribution for time scale ranging from one hour to several weeks. Based on Olsson's works (5), Lu and Yamamoto (2) proposed a new random cascade generator applicable to a wider range of time scale based on the analysis of AMeDAS 10 minute rainfall data at seven AMeDAS stations within Kagawa Prefecture, Japan. As shown in Fig. 1, the random cascade model divides rainfall R_0 in period L_0 (level 0) into b sub-periods (b=2 in this study) at each division step. After n times of divisions, the sub-period has length of

$$L_n = L_0/b^n \tag{1}$$

And rainfall $R_{n,i}$ in i-th sub-period can be expressed

$$R_{n,i} = R_0 \prod_{j=1}^{n} X_{j,i} \tag{2}$$

Then, the rainfall intensity $I_{n,i}$ becomes

$$I_{n,i} = R_{n,i}/L_n \tag{3}$$

where X is the so-called random generator, wchich determines the fraction of rainfall divided into sub-periods. Usually, this value is generated based on the statistical properties of the rainfall. In this study, a random cascade generator

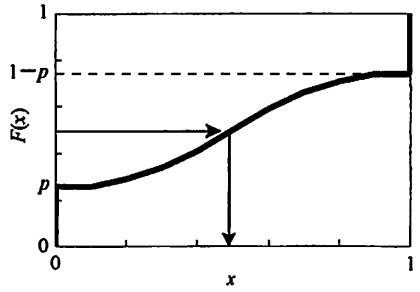

Fig.2 Comulative density function of random cascade generator.

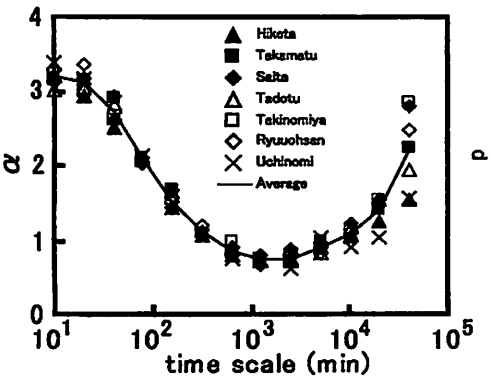

Fig.3 Relation of parameter α and time scale at all 7 AMeDAS stations in Kagawa Prefecture.

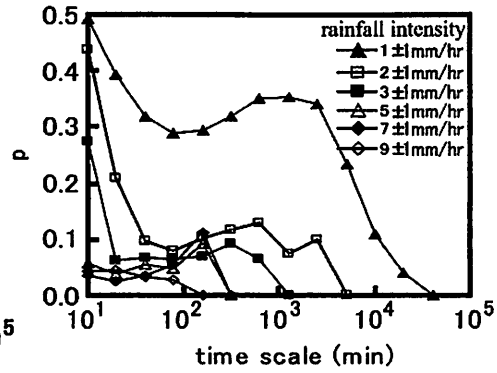

Fig.4 Relation of p and time scale for different rainfall intensity.

proposed by Lu and Yamamoto (2) is used. This generator is micro-canonical, and keeps the total rainfall amount unchanged after the division. In a division step, rainfall R is divided into R_1 and R_2 in two sub-periods.

$$R_1 = x_1 R \tag{4}$$

$$R_2 = x_2 R \tag{5}$$

where x_1 and x_2 are fractions of rainfall for first sub-period and second sub-period. The micro-canonical property requires

$$x_1 + x_2 = 1 \tag{6}$$

Actually, only x_1 should be generated by using a random cascade generator. The random cascade generator used in this study has probabilistic density function as follows :

$$f(x) = p\delta(x) + p\delta(1-x) + \frac{1-2p}{\beta(\alpha,\alpha)}[x(1-x)]^{\alpha-1} \tag{7}$$

where δ is delta function, β is beta function. p and α are parameters. They express the continuity and uniformity of the rainfall time series. The larger the parameter p, the more discontinuous the rainfall time series. The total rainfall

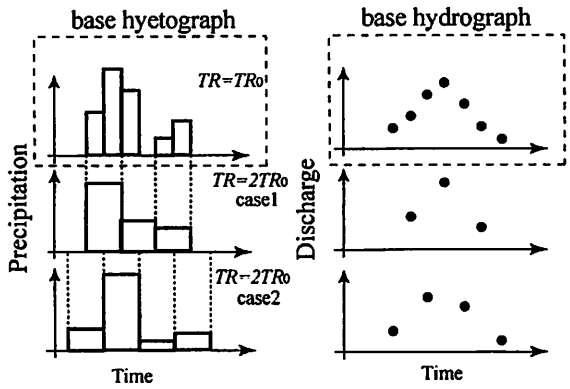

Fig.5 Extraction of hyetographs and hydrographs from base hyetographs and base hydrographs.

concentrates in part of the sub-periods. The larger the parameter α, the more concentrated around $x = 0.5$ the x value. The rainfall between two sub-periods becomes very close, hence the rainfall time series more uniform. As shown in Fig. 2, a random number with statistical density function of Eq.7 can be generated by calculating the inversion of its distribution function $F(x)$ as follows,

$$F(x) = \int_0^x f(x)dx \tag{8}$$

In this model, rainfall amount over initial period L_0 is divided into two rainfall $R_{1,1}$ and $R_{1,2}$ over $2^{-1}L_0$ hour sub-periods by using parameters at time scale L_0. This division procedure will be repeated until the rainfall time series has a temporal resolution high enough. Finally, a higher resolution rainfall time series can be generated. As shown in Fig. 3, Lu and Yamamoto (2) showed the relationship of the model parameter α and the time scale based on the analysis of AMeDAS 10 minute rainfall data at all seven AMeDAS stations within Kagawa Prefecture, Japan. In this study, the average value of all stations is used as a model parameter. Besides the time scale, the model parameter p also depends on the rainfall intensity as shown in Fig. 4, and p increases if rainfall intensity becomes high. Considering the intensity of design rainfall, the extremity, $p = 0$ is assumed in this study.

Base hyetographs and base hydrographs

By using this random cascade model, ten thousand hyetographs with temporal resolution 1.40625 minutes (level: 10) are generated from the daily rainfall amount 300 mm. Hereafter, these hyetographs are referred to as base hyetographs, and the temporal resolution 1.40625 minutes is referred to as base temporal resolution.

These base hyetographs are then inputted to a storage function method and are transformed to hydrographs that are referred to as base hyetographs. The storage function method used in this study is as follows.

$$r_e(t) = \begin{cases} f_1 r(t) & (R_{sa} > \sum r(t)) \\ r(t) & (R_{sa} < \sum r(t)) \end{cases} \tag{9}$$

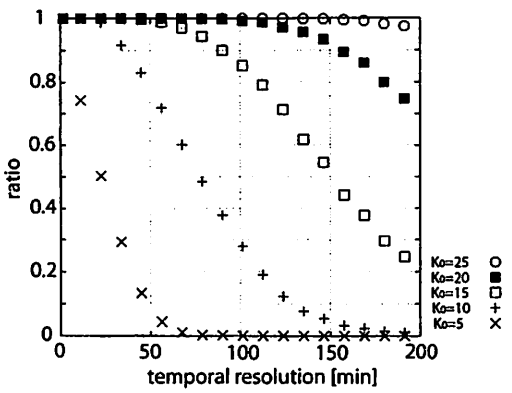

Fig.6 Relation of temporal resolution and ratio of reproduced events having Nash efficiencies larger than 0.9.

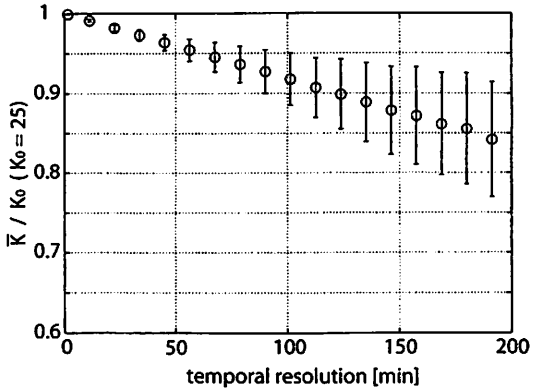

Fig.7 Relation of mean value $\overline{K}$ of the calibrated parameters and temporal resolution ($K_0 = 25$).

$$\frac{ds(t)}{dt} = r_e(t - TL) - q(t) \tag{10}$$

$$s(t) = Kq(t)^P \tag{11}$$

where t (h) is time, s (mm) is storage depth, r (mm h^{-1}) is observed rainfall depth, r_e (mm h^{-1}) is effective rainfall depth, f_1 is primary runoff ratio, q (mm h^{-1}) is river flow depth, TL (h) is time lag between rainfall and river flow, R_{sa} (mm) is saturation rainfall, and K ($mm^{1-P} h^P$) and P are parameters. In this study, the focus is concentrated at runoff routing. For runoff generation, the simple case, a saturated basin is considered so that $R_{sa} = 0$ mm. In this case, all rainfall becomes runoff, f_1 will not affect the calculation in this study. And the TL is set to 0 hour both in the transformation from hyetographs to hydrographs, and in the calibration of model parameters in following section. K and P are the only parameters that should be considered in this study. According to Nagai *et al.* (3), P equals to 0.6, and K is a parameter reflecting geomorphologic factors of the study basin such as drainage area. Therefore, only model parameter K is discussed in this study. K = 5, 10, 15, 20 and 25 $mm^{1-P} h^P$ are used in transformation from hyetographs to hydrographs. In the generation of the base hydrographs, a small computational time interval of one second is used. These are the true parameter values that should be calibrated in following section. They are denoted as K_0 hereafter.

Calibration of model parameters

As shown in Fig. 5, hydrological data with temporal resolution TR are extracted from base hyetographs and base hydrographs. Hereafter, the extracted hyetograph and hydrograph are used as observed hydrological data. From these data, the model parameter K is calibrated.

Adding up rainfall data with base temporal resolution TR_0 makes the hyetographs with large temporal resolution TR. When $TR = N \times TR_0$, there are N ways to start time period TR. In making 2.8125-minute ($2TR_0$) hyetographs, for exeample, two ways to add up the rainfall data exist as shown in Fig. 5. In this study, the start time is decided randomly. Once the start time is decided, the discharge data at the start time of period TR are sampled from base hydrograph to form a hydrograph having same time resolution and sampling time with rainfall data.

From the observed hyetograph and hydrograph, the storage can be calculated by

Table.1 Slope and correlation coefficients of regression lines.

K_0 ($mm^{1-P}h^P$)	5	10	15	20	25
a	-0.0031	-0.0019	-0.0014	-0.0010	-0.0008
Correlation	0.9552	0.9914	0.9977	0.9994	0.9996

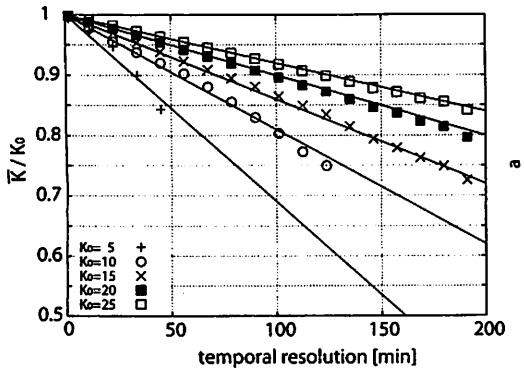

Fig.8 Relation of mean value $\overline{K}$ of the calibrated parameters and temporal resolution for different K_0.

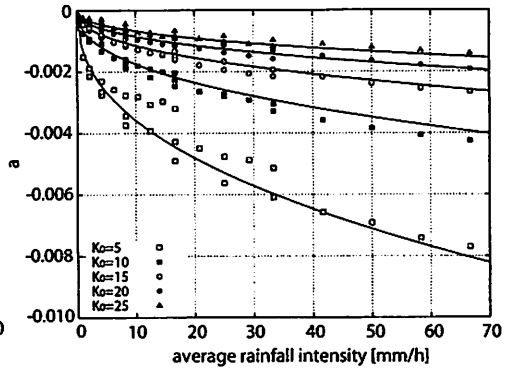

Fig.9 Relation of slope a of regression line and rainfall intensity

$$s(t+\Delta t) = s(t) + [r(t) - q(t)]\Delta t \tag{12}$$

In some cases, the storage may become negative. This may be caused by the rough temporal resolution. As temporal resolution increases, such cases increase. These cases are excluded from following parameter calibration.
From the storage and discharge, the model parameters can be estimated by minimizing

$$\sum[\ln s - (\ln K + P \ln q)]^2 \rightarrow \min. \tag{13}$$

Therefore, K can be calculated by using following equation:

$$K = \exp\left(\frac{\sum \ln s - P\sum \ln q}{n}\right) \tag{14}$$

In order to emphasize the fitness near the peak flow, only the discharge data larger than 50% of the peak flow are used in the parameter calibration. Using the estimated parameter K, the calculated hydrograph is reproduced, and its Nash efficiency (Nash and Sutcliffe (4)) is computed over the period twice the rainfall duration. For each case, a combination of temporal resolution and "true parameter" K_0, average and standard deviation of calibrated parameter K are computed from those values that make the Nash efficiency larger than 0.9. As shown in Fig. 6, the percentage of these values in all 10^4 events decreases as the temporal resolution decreases for each "true parameter" K_0. For the temporal resolutions at which these percentages are smaller than 10%, the parameter calibration is stopped, and no statistical data are calculated.

RESULTS AND DISCUSSIONS

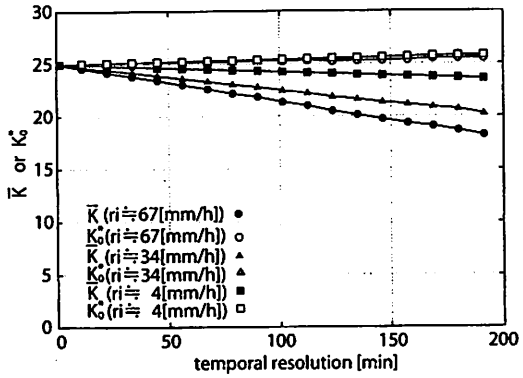

Fig.10 Mean value $\overline{K}$ of the calibrated parameters and their estimated "true" value K_0^*.

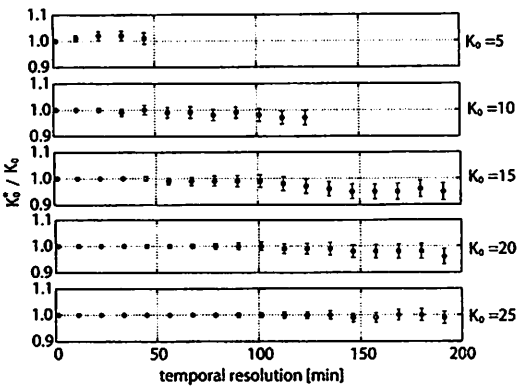

Fig.11 Relation of K_0^*/K_0 to temporal resolution for different K_0 when sample size n equals to ten.

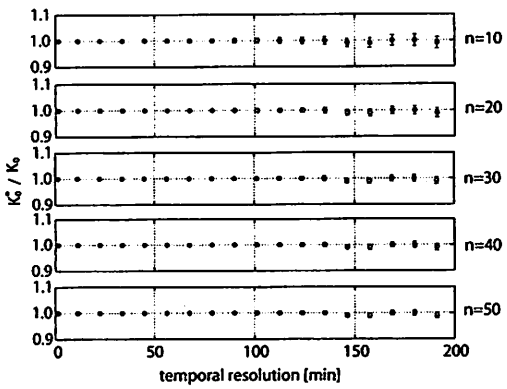

Fig.12 Relation of K_0^*/K_0 to temporal resolution for different sample size n when K_0 equals to 25.

Relationship between the temporal resolution and the calibrated model parameter

Fig. 7 shows the relationship between the mean value $\overline{K}$ of the calibrated model parameters K and the temporal resolution in case of K_0 = 25 $mm^{1-p}h^p$. However, the value $\overline{K}$ is normalized by its "true" value K_0 so that $\overline{K}/K_0$ becomes unity at the base temporal resolution (1.40625 minutes), and the error bar shows its normalized standard deviation. As temporal resolution decreases, the mean value $\overline{K}$ decreases and the standard deviation increases. This may be caused by the smoothing of the rainfall intensity at lower temporal resolutions. This indicates that the parameter calibrated from each flood event may include large uncertainty if the hydrological data does not have suitable temporal resolution for the study basin. Fig. 8 shows same relationships for different "true parameter" K_0. The straight lines on Fig. 8 are linear regression line derived by using the least square method, and they are expressed by

$$\frac{\overline{K}}{K_0} = a\,TR + 1 \qquad (15)$$

The relationship between $\overline{K}$ and temporal resolution is linear, and the slope a varies with K_0, the "true" parameter in Fig. 8. Table 1 shows the slope a and the correlation coefficient. For small K_0 which corresponds to small basin and quick response, the absolute value of the slope a is large. This means that smaller basins can be affected by the temporal

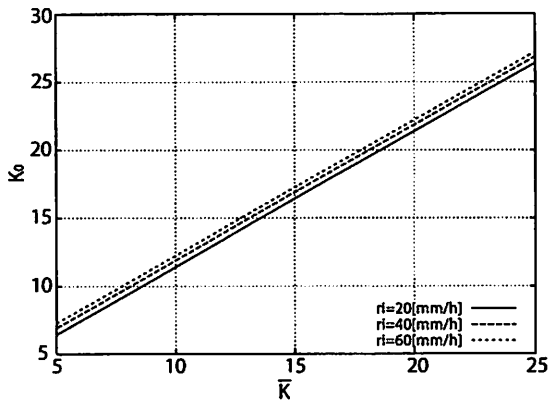

Fig.13 Relation of $\overline{K}$ and K_0 for fixed temporal resolution TR = 60 min.

resolution more significantly. In our previous study (Yamamoto and Lu (9)), the slope a is modelled as a function of the "true parameter" K_0.

Dependency of the slope a on rainfall intensity

In this study, the dependency of the slope a on rainfall intensity is also investigated by changing total rainfall amount R and initial rainfall period T in the generation of base hyetographs. The combination of T = 360, 720 and 1440 minutes, and R = 25, 50, 100, 150, 200, 250, 300, 350, 400, 450 and 500 mm are considered. For each R and T, base hyetographs are generated, and the analysis same with that using T = 1 day and R = 300mm are carried out. Fig. 9 shows the relationship between the slope a and average rainfall intensity. The slope a is related to K_0 and average rainfall intensity r_i by using following equation.

$$a = br_i^c K_0^d \tag{16}$$

As a result of multiple linear regression, b = 0.007, c = 0.427 and d = -1.031 are obtained. The coefficient of determination is 0.97. By substituting Eq.16 into Eq.15, the calibrated parameter $\overline{K}$ can be expressed as a simple function of both the "true" parameter and the resolution.

$$\frac{\overline{K}}{K_0} = br_i^c K_0^d TR + 1 \tag{17}$$

By using this equation, the "true" parameter K_0 can be estimated from the average value of the calibrated K, the temporal resolution and the average rainfall intensity. Fig. 10 shows both $\overline{K}$ and the estimation of K_0^* in case of K_0 = 25 $mm^{1-P}h^P$. Here K_0^* are estimated from $\overline{K}$, rainfall intensities, and TR. The K_0^* is very close to its "true" value K_0 at all temporal resolutions. Fig. 11 shows the relationship between the normalized K_0^*, the ratio of K_0^* to K_0, and the temporal resolution for all K_0 and r_i = 25 mmh^{-1}. From this figure, it is very clear that Eq.17 yields very good estimates K_0^* for all K_0 and temporal resolutions. Hopefully, this will allow improvements in analysis of the "true" model parameter, for example, the regionalization and/or synthesization of the model parameters.

Estimation of K_0 from calibrated parameter K

In the establishment of Eq.17, $\overline{K}$ is the average value of the parameters calibrated from at least one thousand flood events. However, the number of flood events used to calibrate model parameter is very limited in hydrological practice. The sample size is usually smaller than about one hundred. In order to apply Eq.17 to estimate the true parameter K_0, the effect of sample size, the number of flood events used to calculate $\overline{K}$, is analyzed. By randomly picking up n (n = 10, 20, 30, 40 and 50) flood events, mean value $\overline{K}$, and then $\overline{K}_0^*$ can be derived. Finally, the average value of $\overline{K}_0^*$ and its standard deviation are obtained by repeating above procedure. In Fig. 12, they are plotted against the temporal resolution when K_0 equals to 25 $mm^{1-P}h^P$. For all sample sizes, the $\overline{K}_0^*$ is very close to its true value and unbiased. As the sample size increases, the standard deviation of $\overline{K}_0^*$ decreases. However, even when n = 10, the error of $\overline{K}_0^*$ is less than 10 %. Fig. 13 shows the relationship between $K = 5.5A^{0.14}$ and K_0 for temporal resolution of 60 minutes that is the standard time interval of the hydrological data in Japan. In most cases, this relation can be used to convert the calibrated parameters to its "true" value.

CONCLUSIONS

In this study, the effects of temporal resolution of data on hydrological modelling are investigated. For a rainfall runoff system which is represented by a hydrological model and its parameters, high resolution rainfall series are generated by means of a random cascade model, and are transformed into discharge series with same resolution. By lowing the resolution of both rainfall and discharge, the model parameters are calibrated and compared with their "true values". It is shown that the temporal resolution of the hydrological data has significant effects on the calibrated model parameters, and then rainfall runoff analysis. For the model used in this study, the storage function method, the calibrated parameter depends on its "true value" that is temporal resolution invariant, the temporal resolution and the rainfall intensity. By applying this relationship inversely, the true value of the parameter can be obtained. As a result, the uncertainty caused by poor temporal resolution of hydrological data is removed. The removal of such uncertainties is crucial for most hydrological research and practice, for example, synthesization of model parameters. Hopefully, same methodology can be used to detect the temporal resolution dependencies of the parameters for other rainfall runoff models.

REFERENCES

1. Gupta, V. K. and Waymire, E.: A statistical analysis of mesoscale rainfall as a random cascade. Journal of Applied Meteorology, 32, pp.251-267, 1993.
2. Lu, M., and Yamamoto, T.: The applicability of random cascade model to estimation of design flood. Journal of Japan Society Hydrology and Water Resources, 18(2), pp.132-139, 2005.
3. Nagai, A., Kadoya, M., Sugiyama, H. and Suzuki, K.: Synthesizing storage function model for flood runoff analysis. Disaster Prevention Research Institute Annals, Kyoto University, 25, B-2, pp.207-220, 1982.
4. Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models. Part 1- A discussion of principles, Journal of Hydrology, 10, pp.282-290, 1970.
5. Olsson, J.: Evaluation of a scaling cascade model for temporal rainfall disaggregation. Hydrology and Earth System Sciences, 2(1), pp.19-30, 1998.
6. Over, T. M. and Gupta, V. K.: A space-time theory of mesoscale rainfall using random cascades. Journal of Geophysical Research, 101(D21), pp.26319-26331, 1996.
7. Schertzer, D. and Lovejoy, S.: Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes. Journal of Geophysical Research, 92, pp.9693-9714, 1987.

8. Takasao, T., Shiba, M., Tachikawa, Y., Fujita, A., Fujita and Nirupama: An analysis of a time scale on runoff phenomena. Disaster Prevention Research Institute Annals, Kyoto University, 38, B-2, pp.381-394, 1995.
9. Yamamoto, T. and Lu M.: The estimation of parameters of drain-storage function invariant with temporal resolution of hydrological data. Annual Journal of Hydraulic Engineering, JSCE, 50, pp.193-198, 2006.

APPENDIX – NOTATION

The following symbols are used in this paper:

a = slope of regression line;

c = parameter of relations of slope a and average rainfall intensity;

d = parameter of relations of slope a and average rainfall intensity;

$f(x)$ = probabilistic density function of x;

f_1 = primary runoff ratio (parameter of storage function model);

$F(x)$ = cumulative distribution function of x;

$I_{n,i}$ = rainfall intensity (level n, i is element of time series) ;

K = parameter of storage function model;

K_0 = true parameter in storage function model;

$\overline{K}$ = mean value of calibrated parameter in storage function model;

L_n = temporal resolution (level n);

P = parameter of storage function model;

p = parameter of cascade model;

$q(t)$ = river flow;

$r(t)$ = rainfall at t;

$r_e(t)$ = effective rainfall;

$R_{n,i}$ = rainfall amount (level n, i is element of time series);

R_{sa} = saturation rainfall (parameter of storage function model);

$s(t)$ = storage depth;

t = time;

TL = time lag between rainfall and river flow (parameter of storage function model);

TR = temporal resolution;

x_1 = fraction of rainfall for first sub-period;

x_2 = fraction of rainfall for second sub-period;

X = random cascade generator;

α = parameter of cascade model;

$\beta(\alpha,\alpha)$ = beta function; and

$\delta(x)$ = delta function.

(Received January 24, 2007 ; revised May 10, 2007)