Annual Journal of Hydraulic Engineering, JSCE, Vol.50, 2006, February

AN ASSESSMENT OF UNCERTAINTY IN PRECIPITATION ON RUNOFF SIMULATION

Khada Nanda DULAL[1], Kuniyoshi TAKEUCHI[2] and Hiroshi ISHIDAIRA[3]

[1]Student Member of JSCE, PhD student., Dept. of Civil and Environmental Engineering, University of Yamanashi (Takeda 4-3-11, Kofu, Yamanashi 4008511, Japan)

[2]Member of JSCE, PhD, Professor, Dept. of Civil and Environmental Engineering, University of Yamanashi (Takeda 4-3-11, Kofu, Yamanashi 4008511, Japan)

[3]Member of JSCE, PhD, Associate Professor, Dept. of Civil and Environmental Engineering, University of Yamanashi (Takeda 4-3-11, Kofu, Yamanashi 4008511, Japan)

The aim of this study is to assess the uncertainty in precipitation using a distributed hydrological model. Firstly, a true model is established and then error is applied to true precipitation. The uncertainty is analyzed by using sensitivity analysis approach for systematic error and Monte Carlo approach for random error. Next, the parameters of the model are calibrated with erroneous data. Finally, the impact of low precipitation is assessed by neglecting different levels of low precipitation. The result of the study for a Nepalese river basin shows that a systematic error exceeding +/-10% causes significant impact on simulated flows. The impact of normally distributed random error with standard deviation equals to 10% of observed precipitation is not substantial. The calibration of parameters can adjust the low error, but the higher errors should not be compensated by just fitting the curve. Error on low precipitation of amount less than 0.5mm measured with some error does not affect the flood discharge.

Key Words: *Uncertainty, Precipitation, Systematic error, Nepal, Error model, Monte Carlo*

1. INTRODUCTION

Due to the randomness in nature and the lack of knowledge, the result obtained from a hydrological model is always subjected to uncertainty. The uncertainty in output of the hydrological model is a function of uncertainty in data, parameter and model. Uncertainty in input data occurs due to the measurement errors and spatial and temporal sampling errors. There is uncertainty in parameter due to the lack of accuracy with which the parameters can be estimated or due to the limited understanding of the relationship. The model uncertainty arises due to the inability of the model to truly represent a natural process. An extensive review on sources of uncertainty in hydrological model and various methods for assessing the uncertainty can be found in Melching [1)].

In most of the hydrological modeling studies, parameter uncertainty is a major focus. Parameter calibration in hydrological modeling reduces parameter uncertainty as well as compensates data and model uncertainty to some extent. Besides calibration, additional information on parameters, good quality data of sufficient resolution, selection of appropriate model for a particular situation are also important to improve the model results.

Precipitation is the most important input of hydrological model as it is a major driver of the hydrological process. Two sources of uncertainty influence the precipitation measured by a rain gauge. Firstly, the uncertainty arises due to the systematic error (bias) or random error in measurement. The sources of systematic error are: human error, site error, instrumental error, evaporation error, wind error, wetting error, splashing error, drifting error. Random error occurs due to human error in observation, error in instrument and due to small variations in meteorological conditions. Secondly, spatial interpolation of point data to areal data adds uncertainty to the precipitation input.

The performance of any hydrological model is highly dependent on the precipitation data. For a model whose parameters have to be determined by calibration, the uncertainty in precipitation

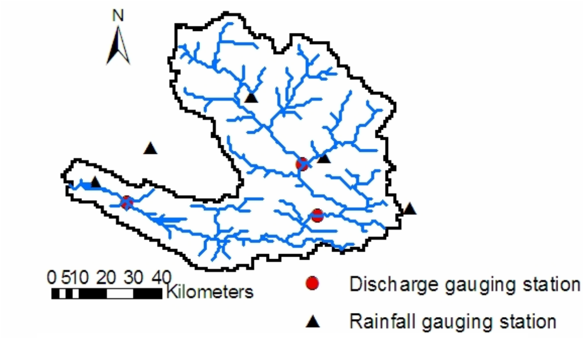

Fig.1 Map of the West Rapti River Basin

influences the model parameters as well. In hydrological modeling studies, a few papers have considered input data uncertainty, especially precipitation as one of the dominant sources of uncertainty, e.g.[2), 3), 4)]. Some papers have analyzed precipitation uncertainty as well as other sources of uncertainty[5), 6), 7)]. Review of various papers mentioned above makes us conclude that the research on both the systematic and random uncertainty in precipitation for a distributed hydrological model is still sparse. Therefore, the objectives of this study are to analyze the impact of systematic and random uncertainty in precipitation on the modeling results for a distributed hydrological model and to identify the extent of error beyond which it becomes more significant in modeling.

2. STUDY AREA

The study area for the research is the West Rapti River Basin (**Fig. 1**), which is located in the mid-western region of Nepal. The catchment area of the basin is 5450 square kilometres and the length of main stream channel is 208 km. The basin elevation ranges from 205 m to 3437 m above mean sea level. The river originates from the middle mountains of Nepal, then enters to the flat area and finally drains to India to join the Ganges River. The source of runoff is monsoon rainfall and groundwater. Daily data from 5 rainfall stations and 3 discharge stations from 1980-1993 is available for the study. The average annual rainfall during this period is 1580mm and mean annual discharge at Jalkundi (the most downstream station) is 113.7m^3/s. Landuse, topographic, soil and potential evaporation (PET) data are obtained from freely available global data set: specifically, topographic data from United States Geological survey (GTOPO30), land use data from International Geosphere-Biosphere Programme, soil data from Food and Agricultural Organization, PET data from United Nations Environment Programme, Global Resource Information Database.

3. HYDROLOGICAL MODEL

The hydrological model to be used for the study is BTOPMC (Version 1.0). The meaning of BTOPMC is “Blockwise use of TOPMODEL with Muskingum-Cunge routing”. This is a distributed hydrological model developed at the University of Yamanashi, Japan[8), 9)]. BTOPMC is an extension of TOPMODEL concepts[10)], which is developed in order to overcome the limitations of using the TOPMODEL for large river basins. For large river basins, spatial heterogeneity and timing of flow to outlet are the important factors. For representing spatial variability in BTOPMC, a basin is composed of grid cells, which can be divided into sub-basins, where each sub-basin is considered as a block or a unit. The runoff generation at each grid cell is based on TOPMODEL concepts. To consider timing of flow, flow from each grid cell is routed to the outlet using Muskingum-Cunge routing. BTOPMC can accommodate the spatial variability in forcing data, topographic, soil and vegetation data. The parameter variability in BTOPMC is considered in the following way: Transmissivity decay factor (m) and Manning roughness (n_0) for each sub-basin, Maximum root zone capacity (S_{rmax}) for each land use classes and Saturated transmissivity (T_0) for each soil texture.

4. METHODOLOGY

(1) Setting up of a true model

A set of precipitation and discharge is selected and the best set of parameters is identified by calibrating the BTOPMC model. The chosen set of precipitation is considered as true precipitation, the optimized parameters are considered as true parameters and the simulated discharge is considered as true discharge.

(2) Uncertainty analysis

a) Error model

Error model for systematic error: For studying the effect of systematic error, perturbation is applied to the observed precipitation using the following form of error model:

$$P_e = P_m + k.P_m = (1+k)P_m \quad (1)$$

where P_e = perturbed precipitation, P_m = observed precipitation, k = coefficient to indicate how much bias is added or subtracted from the observed precipitation. This approach is a sensitivity analysis approach used as a mean for uncertainty analysis. This form of equation is suitable for expressing systematic error because the systematic error is a fixed error and the equation expresses the systematic error as a fixed percentage of measured value for all measurements.

Error model for random error: In BTOPMC, spatial distribution of precipitation is obtained by using Thiessen polygon approach, where all the grids within the polygon have the same values of precipitation as the point precipitation gauge data. To study the uncertainty in discharge due to the random error in precipitation, Monte Carlo simulation approach is applied. To apply this approach, perturbation to precipitation is introduced for each gauging station data using the following form of error model:

$$P_e = P_m + \sigma.e\,; \quad \sigma = r.P_m \qquad (2)$$

where P_e = perturbed precipitation, P_m = observed precipitation, σ = assumed standard deviation of additive random error relative to the measured precipitation, r = coefficient, e = random error component assumed normally distributed with mean equals zero and standard deviation equals one. In the above formulation, the error is assumed to be independent in time and space. With this formulation if observed precipitation is zero, the error is also zero. Thus, the zero precipitation events become unaffected.

Error model for very low precipitation: Trace precipitation, which is beyond the resolution of rain gauge can not be measured and is usually neglected in modeling. To understand the effect of neglecting very low daily precipitation, performance of BTOPMC model is assessed by neglecting different levels of low precipitation.

b) Assessment of the impact of precipitation uncertainty on model results

For analyzing the impact of systematic error, different values of *k* are taken and BTOPMC model is run keeping other inputs and parameters at true values. For studying the influence of random error, Monte Carlo simulation is performed by taking different values of *r* keeping other inputs and parameters at true values. Then, uncertainty in discharge due to random error is analyzed from the outputs of *n* number of Monte Carlo simulations. Finally, the impact of neglecting very low precipitation is analyzed.

c) Assessment of the capability of parameter to absorb precipitation uncertainty

The purpose of calibration is to bring model results as close as observed values by tuning parameters. However, if the parameters of the model are determined by using erroneous precipitation data, the parameter will also be affected. To understand to what extent the parameter calibration can absorb precipitation uncertainty, the parameters of BTOPMC model are calibrated for different values of *k* and *r*, and the performance of the model with calibrated parameter is compared to the true model.

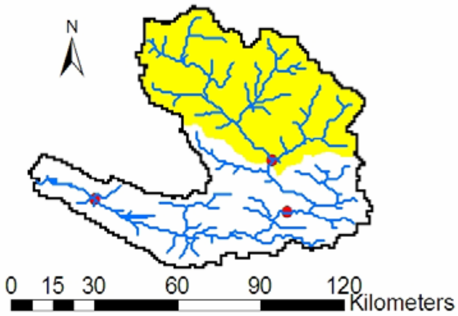

Fig.2 Sub-basin division of the West Rapti River Basin

5. RESULTS AND DISCUSSIONS

(1) True model

Digital Elevation Map (DEM) data, soil type data, land use data, precipitation data, potential evaporation data and flow data which are required for running BTOPMC, are formatted as per the requirements of BTOPMC model. The land use data is reclassified into 4 classes in order to reduce equifinality and increase efficiency in computation. The basin is divided into two sub basins (**Fig. 2**). Time series data from 1980 to 1987 is used for calibration and from 1988 to 1993 is used for validation. The calibrated parameters of the model are: m (sub-basin 1) = 0.06m, m (sub-basin 2) = 0.04m, n_0 (sub-basin 1) =0.02, n_0 (sub-basin 2) =0.01, S_{rmax} (Deep rooted) = 0.05m, S_{rmax} (Shallow rooted) = 0.04m, S_{rmax} (Shallow rooted & Irrigated) = 0.03m, S_{rmax} (Impervious) = 0.0001m, T_0 (Clay) = 0.5m^2/h, T_0 (Sand) = 7m^2/h, T_0 (Silt) = 3m^2/h. Nash-Sutcliffe coefficient of efficiency (NSE) at Jalkundi station is 61.55% for calibration and 66.06% for validation. The NSE in very high range could not be obtained because there is uncertainty in spatial distribution of precipitation (5 rainfall stations for 5450 km^2) and uncertainty in manual parameter estimation as no automatic optimization has been implemented in BTOPMC. The simulated and observed hydrograph for validation at Jalkundi station is shown in **Fig. 3**. The performance of the model is quite good for most parts of the hydrograph, except a few peaks, which are under-predicted by the model. For uncertainty analysis, the observed precipitation set from 1980-1987, the calibrated parameters and the simulated discharge of the same period are considered as true set of precipitation, parameters and discharge respectively.

(2) Assessment of precipitation uncertainty

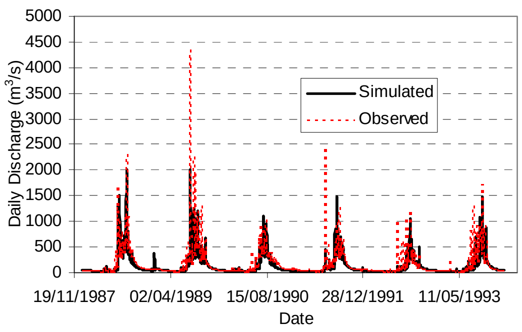

Fig.3 Validation at Jalkundi

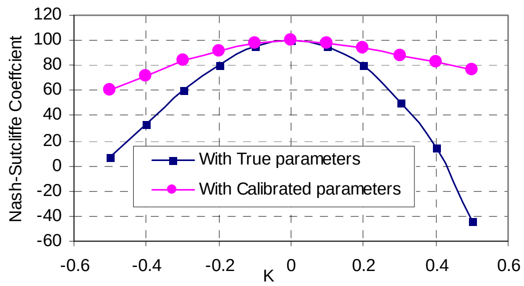

Fig.4 Impact of systematic error on overall model performance using $P_e = (1+k)P_m$

The impact of precipitation on model results is analyzed by using Nash-Sutcliffe coefficient of efficiency (NSE), Bias in runoff volume for overall time series and peak flow events, Normalized Root Mean Square Error (NRMSE) for low flow and high flow (normalized by average flow). Peak flow of magnitude greater than 500m^3/s is considered and the threshold value for separating low and high flow is set at 50m^3/s. The result for the most downstream station is presented here.

a) Impact of systematic error

The magnitudes of systematic error in precipitation from various factors are[11)]: wind error = 2%-10% for rain and 10%-50% for snow, wetting error = 2%-10%, evaporation error = 0%-4%, splashing error = 1%-2%. Including other errors, the maximum systematic error in precipitation is around 30% for rain and 70% for snow. In this study the maximum range of systematic error is kept at +/-50, i.e. k is varied from -0.5 to 0.5.

The result of the analysis (**Fig. 4**) shows that for k = -0.1 and 0.1, the NSE decreases by 5.4% and 4.9% of true model respectively. Bias in runoff volume (**Table 1**) for overall time series for k = -0.1 and 0.1 is -15.2% and 15.6% respectively, and for peak flow events, the bias is -19.1% and 19.6% respectively. Further increase of the error in either direction makes the performance of the model worse and worse.

Next, for the different values of k, the model parameters are calibrated. Comparison of model

Table 1 Bias in runoff volume (%) due to systematic error

k	For whole period	For peak flow events
-0.5	-71.5	-82.8
-0.4	-59.0	-70.4
-0.3	-44.5	-54.6
-0.2	-31.0	-38.3
-0.1	-15.2	-19.1
0.1	15.6	19.6
0.2	29.4	37.6
0.3	47.3	60.1
0.4	61.1	78.4
0.5	79.6	102

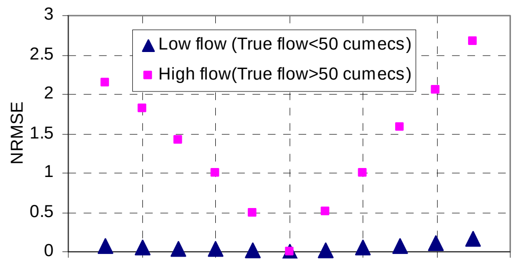

Fig.5 Impact of systematic error on low and high flows using $P_e = (1+k)P_m$ with true parameters

performance with true and calibrated parameters (**Fig. 4**) shows that for k = -0.5 to 0.5 in step of 0.1, the NSE is improved by 53.3%, 39%, 24.2%, 11.43%, 2.51%, 3.4%, 13.26%, 37.75%, 68.1%, 120.13% respectively. Though the performance of the calibrated model is better than the un-calibrated case due to curve fitting procedure, the model performance becomes worse for higher errors even for calibrated model. Parameter calibration can adjust some amount of error in data, but we can not expect that it can adjust any amount of data. The higher the error in data, the worse the performance of the model. Therefore, the NSE for recalibrated case degraded systematically.

The performance of the model for high flow and low flow is shown in **Fig. 5**. The result shows that increasing error in rainfall during high flow makes the model performance worse. This is obvious because the high flow period, which is the monsoon period, is the main rainfall period and any error on rainfall during this period propagates through the model thus affecting its performance. The impact of error on low flow is insignificant because during low flow period, there is no or very little rain.

It is seen from **Fig. 4** that the performance curve for positive k and negative k is asymmetric. Similarly in **Fig. 5**, the trend of the curves is asymmetric. Increasing k in the positive side degrades the model performance rapidly, while

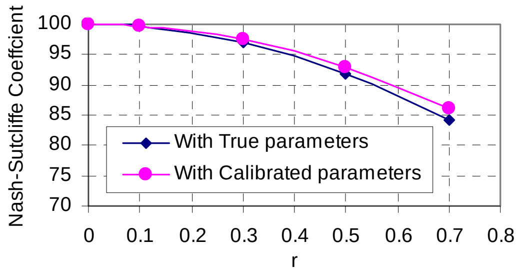

Fig.6 Impact of random error on overall model performance using $P_e = P_m + \sigma.e$; $\sigma = r.P_m$

Table 2 Bias in runoff volume (%) due to random error

r	0.1	0.3	0.5	0.7
For whole period	0.071	0.83	3.24	10.31
For peak flow events	0.01	0.3	2.22	10.17

increasing k in the negative side has lower effect than the positive side. This implies that negative k is safer side, while positive k is risky side. For linear model, the same bias in either direction has same impact. However, as BTOPMC is a non-linear model, the same bias in positive direction has higher impact.

b) Impact of random error

Though the exact amount of the random error can not be specified, its impact on modeling results can be quantified by uncertainty analysis. This study makes use of Monte Carlo framework for this purpose by taking r values equals to 0.1, 0.3, 0.5 and 0.7. For each case, 50 samples of precipitation data set are generated randomly according to equation (2). If the perturbed precipitation becomes negative due to the negative random number, then it is taken as zero. The occurrence of negative values is usually low. An example of one realization shows that negative values occurrence expressed as percentage of total number of rainy days are: for r = 0.1, no negative value, for r = 0.3, less than 0.3%, for r = 0.5, 2%-3%, for r = 0.7, 6%-8%. For each case, average performance indicator is computed from the performance of 50 simulations.

The result of the analysis (**Fig. 6**) shows that the decrease in the NSE from the true model for r = 0.1, 0.3, 0.5 and 0.7 is 0.4%, 3.1%, 8.2% and 15.9% respectively. As shown in **Table 2**, the bias in runoff volume for r = 0.1 and 0.3 is very low; 0.071% and 0.83% for overall time series and 0.01% and 0.3% for peak flow events of value greater than 500m^3/s. The result makes us clear that for r = 0.1, the impact is negligible and increase in r beyond that decreases the performance of model although the decreasing rate is small until r = 0.3.

Next, the parameters of model are calibrated for different values of r. The comparison of result with true parameters (**Fig. 6**) shows that the calibration improves the NSE by 0.1%, 0.7%, 1.1% and 2.2% for r = 0.1, 0.3, 0.5 and 0.7 respectively. As the impact of errors on un-calibrated model is low, the difference in performance of calibrated and un-calibrated model with the increase in r is not so big like the systematic error case. Similar to the systematic error case, the NSE is degraded systematically with the increase of errors even for recalibrated case.

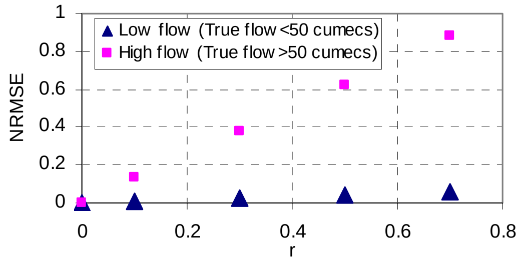

Fig.7 Impact of random error on low and high flows using $P_e = P_m + \sigma.e$; $\sigma = r.P_m$ with true parameters

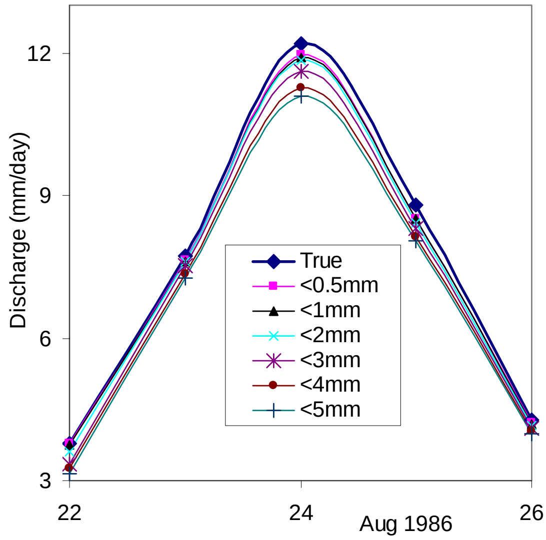

Fig.8 A sample plot of hydrograph neglecting low precipitation

NRMSE for both high flows and low flows due to increasing standard deviation of random error is shown in **Fig. 7**. As in the case of systematic error, the effect on low flow is negligible because it is a period with no or very little rain. As for high flow, the performance is decreasing with increasing r due to the propagation of error imposed on rainfall.

c) Very low precipitation

To study the impact of neglecting low precipitation, precipitation less than 0.5mm, 1mm, 2mm, 3mm, 4mm and 5mm are neglected respectively with other inputs and parameter same as true model. As an example, the output hydrograph for a small period with a peak is shown in **Fig. 8**. The result shows that the peak magnitude is decreased with neglecting higher values of precipitation. In particular for this peak, the peak magnitude is reduced by 1.1 mm/day for neglecting precipitation less than 5mm/day. When precipitation less than 0.5mm/day is neglected, the peak is reduced by 0.24 mm/day. Very little rain, beyond the precision limit of the instrument (0.1 mm in most cases), can not be measured well. This study concludes that even if there is some error in measuring such low precipitation (0.5mm/day in this case), this sort of error does not affect peak discharge as most of it is lost due to evaporation.

6. CONCLUSIONS

In general, the impact of input data error depends on the type of model, type of basin and the type of error model. This study assessed the precipitation uncertainty for a basin in Nepal using BTOPMC as an example of a distributed model. The conclusions of the study are summarized below:

I. For a systematic error of +10% and -10%, decrease in NSE from true model is 5.4% and 4.9% respectively; bias in runoff volume is -15.2% and 15.6% respectively; and bias in runoff volume for peak events is -19.1% and 19.6% respectively. It is intuitive that if the systematic error is very small, the effect is also small and if the error is very large, the effect is also very large. According to this study, a systematic error exceeding 10% of observed precipitation is significant in modeling. Therefore, systematic error should be identified and reduced. The ways of reducing systematic error are: implementation of quality control measures, application of correction methodology.

II. For a random error with standard deviation equals 10% of observed precipitation, decrease in NSE from true model is 0.4%; bias in runoff volume is 0.071%; and bias in runoff volume for peak events is 0.01%. The random error, which is unpredictable and non-constant, might be either positive or negative having long term expected value equals to zero. Slight increase or decrease of precipitation due to random effect acts as a compensating mechanism and hence the random errors have low impact on model results than the systematic error. However, if the random error is larger with standard deviation greater than 10% of observed precipitation, then its effect is detrimental to model results.

III. The calibration of model with systematic error of +10% and -10% in rainfall increases NSE by 2.51% and 3.4% respectively. In case of random error in precipitation, NSE is improved by 0.1% for r = 0.1. For larger error, though NSE is improved due to curve fitting, the model performance is deteriorating with increase of error. This implies that calibration can adjust the error of low magnitude, but errors of higher magnitude can not be just ignored. Therefore, input data uncertainty has to be given due consideration in hydrological modeling.

IV. The impact of neglecting precipitation less than 0.5mm/day does not affect peak discharge. This means the small error due to resolution limit is not very significant for flood discharge.

REFERENCES

1) Melching, C.S.: Reliability estimation, *Computer Models of Watershed Hydrology*, Singh, V.P. eds., Water Resources publications, pp.69-118, 1995.
2) Storm, B., Jensen, K.H. and Refsgaard, J.C.: Estimation of catchment rainfall uncertainty and its influence on runoff predictions, *Nordic Hydrology*, Vol.19, pp.77-88, 1988.
3) Andreassian, V., Perrin, C., Michel, C., Usart-sanchez, I. and Lavabre, J.: Impact of imperfect rainfall knowledge on the efficiency and the parameters of watershed models, *Journal of Hydrology*, Vol. 250, pp. 206-223, 2001.
4) Maskey S., Guinot V. and Price, R.K.: Treatment of precipitation uncertainty in rainfall-runoff modeling: a fuzzy set approach, *Advances in Water Resources*, Vol.27, pp.889 –898, 2004.
5) Nandakumar, K. and Mein, R.G.: Uncertainty in precipitation-runoff simulations and the implications for predicting the hydrologic effects for land-use change, *Journal of Hydrology*, Vol.192, pp.211-232, 1997.
6) Carpenter, T.M. and Georgakakos, K.P.: Impacts of parametric and radar rainfall uncertainty on the ensemble streamflow simulations of a distributed model, *Journal of Hydrology*, Vol.298, pp.202-221, 2004.
7) Butts, M.B., Payne, J.T., Kristensen, M. and Madsen, H.: An evaluation of the impact of model structure on hydrological modeling uncertainty for streamflow simulation, *Journal of Hydrology*, Vol.298, pp.242-266, 2004.
8) Takeuchi, K., Ao, T. and Ishidaira, H.: Introduction of block-wise use of TOPMODEL and Muskingum-Cunge method for the hydro-environmental simulation of a large ungauged basin, *Hydrological Sciences Journal*, Vol.44, pp.633-646, 1999.
9) Ao T.: Development of a distributed hydrological model for large river basins and its application to Southeast Asian rivers. Ph.D. Thesis, University of Yamanashi, Kofu, Japan, 2001.
10) Beven, K., Lamb, R., Quinn, P., Romanowicz, R. and Freer, J.: TOPMODEL, *Computer Models of Watershed Hydrology*, Singh, V.P. eds., Water Resources publications, pp. 627-668, 1995.
11) Sevruk, B.: Methods of correction for systematic error in point precipitation measurement, WMO Tech. Doc. 589, 91pp, Geneva, Switzerland, 1982.

(Received September 30, 2005)