

INTERCOMPARISON OF HYDROLOGICAL MODELING PERFORMANCE WITH MULTI-OBJECTIVE OPTIMIZATION ALGORITHM IN DIFFERENT CLIMATES

Qiaoling LI¹, Hiroshi ISHIDAIRA², Satish BASTOLA³ and Jun MAGOME⁴

¹Student Member of JSCE, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi (Takeda 4-3-11, Kofu 400-8511, Japan) E-mail: g07dea05@yamanashi.ac.jp

²Member of JSCE, Dr. of Eng., Associate Professor, Interdisciplinary Graduate School of Medicine and Engineering, University of Yamanashi (Takeda 4-3-11, Kofu 400-8511, Japan)

³Dr. of Eng., UMR HydroSciences Montpellier, Universite Montpellier II - Place Eugene Bataillon (Case courrier MSE, 34095 Montpellier Cedex 5, France)

⁴Dr. of Eng., International Centre for Water Hazard and Risk Management (ICHARM), under the auspices of UNESCO, Public Works Research Institute (PWRI) (1-6 Minamihara, Tsukuba, Ibaraki, 305-8516, Japan)

Performances of hydrological models vary from basin to basin in different climate conditions. The problem of further identifying how performances vary with climate conditions remains a research challenge for the discipline of hydrology. With the objective of studying this problem, several hydrological models such as HYMOD, NAM and TOPMODEL were applied to 11 basins located in Asian countries with different climate conditions. The multi-objective optimization method Non-dominated Sorting Genetic Algorithm (NSGAI) was used for model calibration and obtaining optimal parameter sets for ensemble modeling. A tendency between the model performances and the aridity index (the ratio of annual potential evaporation to precipitation) was found, which might help us improve understanding of the model performances in different climates and select appropriate models for catchments in different climates.

Key Words: hydrological modeling, intercomparison, climate, aridity index, NSGAI, uncertainty

1. INTRODUCTION

Watershed-scale hydrologic models are essential for flood and drought prediction, water resources planning and allocation, erosion and sedimentation studies, nonpoint source pollution and remediation, climate and land use change assessments, hydropower operations, *etc.*¹⁾. As improved understanding of hydrological process in catchments has become the primary focus of recent hydrological studies, application of hydrological models has been conducted in different climates of the world. Examples of comparative studies of catchments in different climates include those of R.Lidén *et al.*²⁾, Atkinson *et al.*³⁾, van Werkhoven *et al.*⁴⁾, and Samuel *et al.*⁵⁾. R.Lidén *et al.* applied a well established model, the HBV model, in four

climatologically different river basins in Europe, Africa, and South America. It was found that model performance increased and the demand of calibration period length decreased with increased catchment wetness and corresponding reduced climatic variability. Atkinson *et al.* studied four small catchments located in a relatively moderate climate in New Zealand and postulated a qualitative relationship between model complexity, timescale, and the climatic aridity index. In recent studies, van Werkhoven *et al.* applied the Sacramento Soil Moisture Accounting Model in 12 Model Parameter Estimation Experiment (MOPEX) catchments spanning a wide hydroclimatic gradient from arid to humid systems and found significant variation in parameter sensitivities that are correlated with the hydroclimatic characteristics of the catchments and

time periods analyzed. Samuel *et al.* conducted a comparative study in three diverse regions of Australia found that the biggest contributor to the differences between catchments is the distribution of soil depth and the soil's drainage characteristics and the second factor is climate.

These studies provided the basic ideas of difference in model performance and how the performance would differ under various hydroclimatic conditions. However, discussion of these researches was based on the application of a single model to several basins or application of several models to only one or limited number of basins. For more synthesized understanding, a multi-model application for multi-basins in different hydroclimatic conditions will be necessary. Furthermore, most of previous studies were mainly conducted in the regions of EU and US, focusing on qualitative analysis of model performance difference due to different parameter sensitivities and different levels of importance in model structure resulting from different hydroclimatic conditions. To improve our understanding of model performance difference in different climates, we need to address the following questions:

- (a) How do hydroclimatic conditions influence model performance? Is the influence significant?
- (b) Is there a significant relationship between model performance and hydroclimatic conditions such as aridity index⁶⁾?
- (c) Can we give a framework for selecting appropriate models for basins in different climates?

To address these questions, we applied several hydrological models such as HYMOD, NAM, and TOPMODEL respectively to 11 basins located in China, Nepal, Sri Lanka, and Philippines at a daily time step, analyzed the convergences of the optimization algorithm. We also analyzed ensemble modeling uncertainty and the tendency between the model performance and the ratio of annual potential evaporation to precipitation, referred to as the aridity index⁶⁾.

2. THE HYDROLOGICAL MODELS

Three hydrologic models are used in the present study: HYMOD⁷⁾, NAM⁸⁾, and TOPMODEL^{9),10)}. Brief descriptions of each model are presented in the following three sections. These models differ in their structure, simulated hydrologic processes, and number of calibration parameters, thereby allowing us to examine how different models affect the results in different climates. Models' parameters that need to be calibrated are listed in **Table 1**.

(1) The HYMOD model

This HYMOD model consists of a relatively simple two-parameter rainfall excess model connected with two series of linear reservoirs (three identical for the quick and a single for the slow response) in parallel as a routing component. The rainfall excess model is described in detail by Moore¹¹⁾. The model assumes that the soil moisture storage capacity, c , varies across the catchment and, therefore, that the proportion of the catchment with saturated soils varies over time step t . The spatial variability of soil moisture capacity is described by the following distribution function:

$$F(c)=1-(1-c(t)/C_{MAX})^{BEXP} \quad 0 \leq c(t) \leq C_{MAX} \quad (1)$$

(2) The NAM model

The NAM model (from the Danish: "Nedbor-Afstromnings-Model", which means precipitation-runoff-model) is a lumped and conceptual rainfall-runoff model originally developed at the Technical University of Denmark¹²⁾. The NAM model describes, in a simplified quantitative form, the behavior of the different land phase of the hydrological cycle, accounting for the water content in different mutually interrelated storages. These storages are the surface zone storage, the root-zone storage, the ground-water storage, and the snow storage. The river routing is done through linear reservoirs that represent the overland flow (two identical linear reservoirs in series), the interflow (a single reservoir), and the baseflow (a single reservoir).

(3) The TOPMODEL model

TOPMODEL is a quasi-physically based semi-distributed rainfall-runoff model in which distributed predictions of catchment response are made based on a simple theory of hydrological similarity of points in a catchment. In the original version of the model used here, the hydrological similarity comes from the use of the topographic index as $\ln \alpha / \tan \beta$, where α is the area draining through a point from upslope and $\tan \beta$ is the local slope angle. Total runoff in TOPMODEL is calculated as the sum of two major flow components: saturation excess overland flow from variable contributing areas and subsurface flow from the saturated zone of the soil. The river routing is conducted through a distribution function that is similar to isochrone method.

Among the above mentioned three hydrological models, HYMOD and TOPMODEL consider spatial variability of soil moisture by using the distribution function and topographic index respectively. All of the models adopt saturation excess runoff formation and relatively simple routing method.

Table 1 Parameters of the models used and their prior ranges

Parameter	Unit	Range	Description
HYMOD			
CMAX	[mm]	1-500	Max. storage capacity
BEXP	[-]	0.1-2	Degree of spatial variability of soil moisture capacity
ALPHA	[-]	0-0.99	Factor distributing the flow between slow/quick reservoirs
RS	[day]	0.0001-0.1	Residence time (slow response reservoir)
RQ	[day]	0.1-0.99	Residence time (quick response reservoir)
NAM			
UMAX	[mm]	1-50	Max. water content (size) of the surface storage
LMAX	[mm]	50-1000	Max. water content (size) of the root zone storage
CQOF	[0,1]	0-1	Fraction of excess rainfall
CKIF	[h]	0.01-2000	Time constant for drainage of interflow
CK12	[h]	3-100	Time constant for routing interflow and overland flow
TOF	[-]	0-0.99	Threshold value for overland flow
TIF	[-]	0-0.99	Threshold value for interflow
TG	[-]	0-0.99	Root zone threshold value for recharge
CKBF	[h]	0.01-5000	Time constant for baseflow
TOPMODEL			
SZM	[m]	0.0001-1	The parameter of the exponential Transmissivity function
lnT0	m ² /h	0.1-25	Effective lateral saturated transmissivity
Td	m/h	0-0.3	Unsaturated zone time delay per unit deficit
SRMAX	m	0.1-20	Max. root zone storage

3. DATA SETS

For an analysis of models performances in different climates preferably a large number of undisturbed data-intensive catchments of the same size and shape but located in different climate zones should be studied. However, it is difficult to obtain such data, especially in Asian regions. In this study, the eleven basins located in China, Nepal, Sri Lanka and Philippines were selected as the study area. The hydrological data of these basins was obtained under the collaboration with University of Yamanashi COE Virtual Academy (VA). The VA is an E-learning program for integrated river basin management. The data was provided by participants of YHyM/BTOPMC course in VA. The main selection criteria were accessible hydrological data of good quality, catchment areas in the same order and the studied basins representing a variety of climate. Basic characteristics of the basins are listed in **Table 2**. All of the models were forced with daily estimates of basin-average precipitation and potential evapotranspiration.

Table 2 List of basin location and data period used for the study

ID	River system	Country	Area (km ²)	Data period	Average Slope	Aridity Index
1	Angat	Philippines	781	1989-1993	0.267	0.60
2	Kalu	Sri Lanka	2719	1991-1994	0.048	0.61
3	Kankai	Nepal	1150	1996-2000	0.191	0.65
4	Bagmati	Nepal	2920	1996-2000	0.132	0.76
5	Hushan	China(Yangtze)	6374	1992-1996	0.070	0.79
6	Huangqiao	China(Zishui)	2660	1982-1985	0.076	0.85
7	WestRapti	Nepal	3380	1987-1991	0.215	0.87
8	Mumahe	China(Yangtze)	1224	1981-1985	0.138	0.89
9	Huangnizhuang	China(Huai)	805	1982-1986	0.104	1.01
10	Baohe	China(Yangtze)	3415	1980-1983	0.176	1.1
11	Fenghe	China(Yellow)	566	1992-1996	0.215	1.32

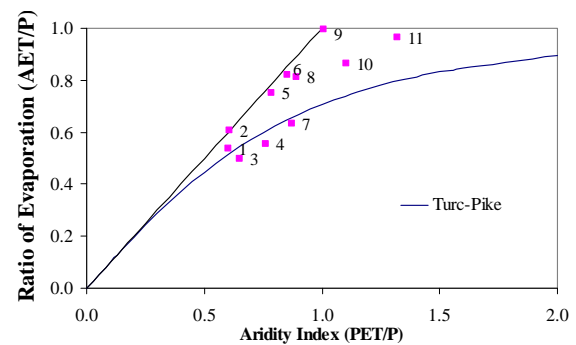


Fig.1 Scatter plot of evaporation ratio against aridity index. All basins have runoff ratios that are close to what is predicted by the Turc-Pike relationship.

Fig.1 plots evaporation ratio (calculated as the ratio of actual annual evapotranspiration (AET) to precipitation (P)) against the ratio of potential evaporation (PET) to precipitation commonly known as the aridity or dryness index (Both AET and PET data are obtained from UNEP GRID¹³). Thereby the larger the aridity index is, the drier the basin is. When the annual available energy expressed as potential evapotranspiration, is greater than the annual precipitation, the annual evaporation is limited by the annual supply of water. Conversely, when the available energy is less than the available precipitation, the annual evaporation is limited by the annual supply of energy^{14,15}).

4. METHODOLOGY

Schematic diagram of methodology is shown in **Fig.2**. Many hydrologic models must be calibrated to be useful for the solution of practical problems. It has not proved possible to clearly demonstrate that a particular function is better suited for calibration of a model than some other. A single objective function unavoidably leads to the loss of information¹⁶). Thereby in this study the widely used Nash Sutcliffe Efficiency (NSE, see Eq. (2)) and the

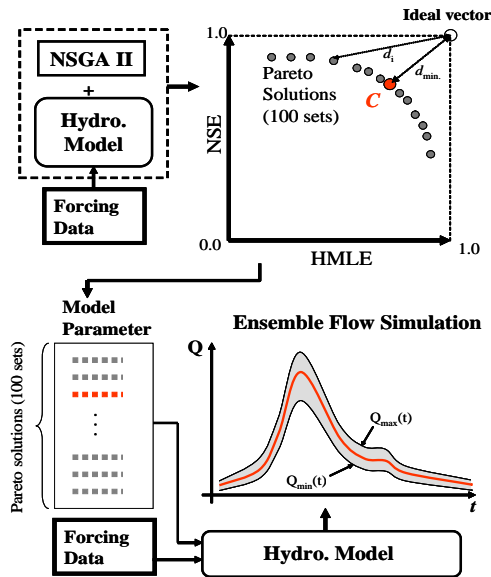


Fig.2 Schematic diagram of methodology used in this study.

Heteroscedastic Maximum Likelihood Estimator (HMLE) were used as multi-objective criteria.

$$NSE = 1 - \left(\frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^n (Q_{obs,i} - \overline{Q_{obs}})^2} \right) \quad (2)$$

where $\overline{Q_{obs}}$ is the average observed discharge and n is the number of time step during simulation. NSE for the transformed flow is referred to as HMLE¹⁷⁾. To consider the heteroscedastic variance in flow, the flow is transformed explicitly before evaluating the objective function by using $Q_{trans,i} = ((Q_{obs,i} + 1)^k - 1) / k$, where k is the transformation parameter and is selected to be 0.3.

A fast and elitist multi-objective genetic algorithm NSGAII was used for optimization, as it is capable of finding multiple Pareto solutions in a single optimization run and maintains a diverse set of elite population. Furthermore, it does not require any additional implicit or explicit parameters other than the standard genetic algorithm parameters such as population size and operator probabilities. Detailed description of the algorithm can be found in Dev *et al.*¹⁸⁾. To run the optimization algorithm, parameter ranges for each model were predetermined according to previous studies (see **Table 1**). In the present study, one hundred sets of Pareto solutions were obtained for each model in each basin.

Generational Distance¹⁹⁾(GD) is used to measure the convergence of the optimization algorithm. To derive this metric, firstly, the Euclidian distance between the “ideal vector” (it is assumed to be obtained when both NSE and HMLE are equal to 1.) and the solution vector is calculated for all the solution points lying in the Pareto optimal front. The average value is referred to as GD (see Eq. (3)). This measure indicates the overall progress of the

algorithm towards the true Pareto front. The smaller the value of generational distance is, the nearer the Pareto optimal front is to the ideal vector and vice versa. The smallest value is referred to as “compromised value” (see **Fig.2**). Hereafter, NSE and HMLE with a subscript c mean the corresponding compromised values.

$$GD = \left(\sqrt{\sum_{i=1}^n (d_i^2)} \right) / n \quad (3)$$

where d_i is the Euclidean distance, n is the number of solution points lying in the Pareto optimal front.

Prediction uncertainty arises from three sources: data uncertainty, model structural uncertainty and parameter uncertainty. Ensemble modeling can be used to quantify and reduce the prediction uncertainty. Thereby a hundred sets of optimal parameters were fed into each hydrological model in each basin and resulted in ensemble of simulated flow. To define the predictive uncertainty, the predicted outputs at each time step (days) are ranked to form a cumulative distribution of the output variable. Then 5% to 95% quantiles were chosen to define the prediction bounds. As three models were applied to 11 basins in different climates, to judge quantitatively the prediction uncertainty associated with each model, the Average Width of the Interval of Simulated Flow²⁰⁾ (AWISF) of the prediction bounds between 5% and 95% quantiles flow were derived using Eq. (4).

$$AWISF = \sum_{t=1}^n (Q_{max,t}(t) - Q_{min,t}(t)) / \overline{Q_{cal}} \quad (4)$$

where $Q_{max,t}(t)$ is the upper 95% quantile flow, $Q_{min,t}(t)$ is the lower 5% quantile flow, $\overline{Q_{cal}}$ is the average of estimated flow, and n is the number of time step during the simulation period.

5. RESULTS AND DISCUSSION

The parameters of the three hydrological models were calibrated respectively using 4 or 5 year daily hydrometeorological data of the 11 basins by utilizing the optimization algorithm NSGAII, the NSE and the HMLE being objective functions. Through this optimization, a hundred sets of optimal parameters were obtained for each model in each basin. Then hydrograph ensembles were obtained for each basin by forcing each model with the parameter sets. The shapes of Pareto front and distribution of solution points in Pareto front varied from model to model and basin to basin. Typically for a specified hydrological model they depend on the number of objective functions, interactions among objective functions, and optimization algorithm itself. Thereby generational distance was

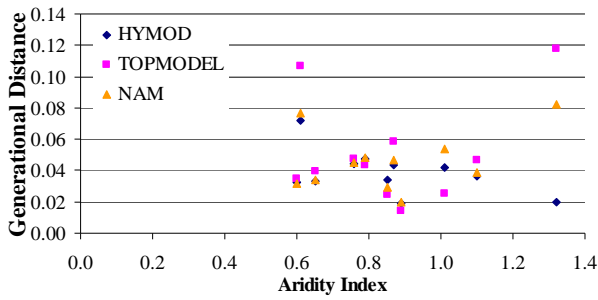


Fig.3 Scatter plot of the model performance measured in terms of generational distance against the aridity index.

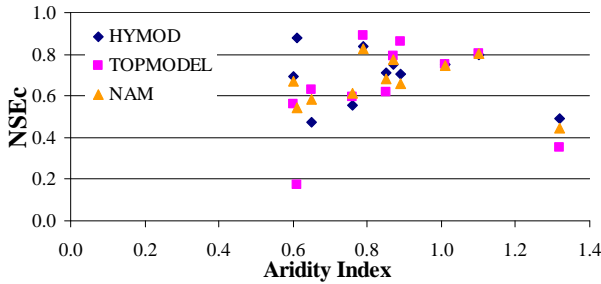


Fig.4 Scatter plot of model performance measured in terms of NSEc against basins' aridity index.

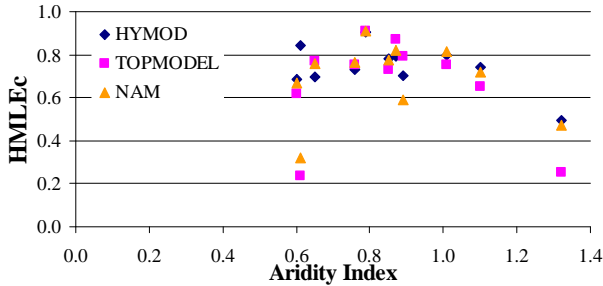


Fig.5 Scatter plot of model performance measured in terms of HMLEc against basins' aridity index.

used to judge the performance of NSGAI in calibrating HYMOD, TOPMODEL, and NAM models in the entire selected basins. **Fig.3** shows the variation in the GD at all basins for all the three models. The figure reveals that the generational distance varies among basins. Moreover, the values are larger in the wettest and driest basins, indicating that the NSGAI converges well and the objective criteria are more “conflicting” with each other. Furthermore, the GD is also found to be correlated with models performances measured in terms of NSE and HMLE (**Figs. 4 and 5**), having a high value in basin where NSE and HMLE are low.

Figs. 4 and 5 show that all of the three models performances are acceptable in humid basins where aridity index is less than 1 except the Kalu basin (ID=2) located in Sri Lanka. This is probably attributable to the hydroclimatic and geographic characteristics (e.g. the Kalu basin has the mildest slope among that of studied basins, see **Table 2**) making models more sensitive. The other probable reason for these discrepancies could be the wide ranges of parameters we set during optimization so that the NSGAI could not find suitable parameters

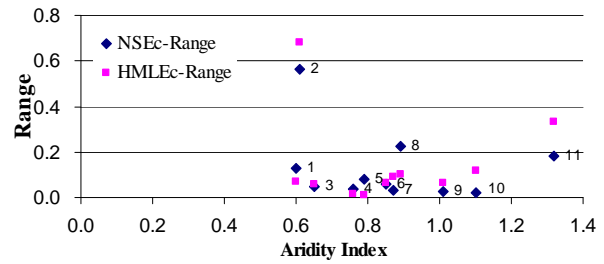


Fig.6 Scatter plot of model performance difference measured in terms of NSEc and HMLEc against basins' aridity index.

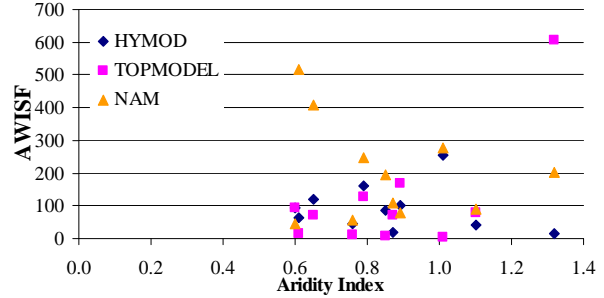


Fig.7 The Average Width of the Interval of (AWISF) in (%) calculated from the ensemble of simulated hydrograph corresponding to the hundred sets of parameters lying in the Pareto optimal front.

within the ranges. In terms of NSE, we can not find obvious relationship between models performances and aridity index, while in terms of HMLE, the tendency that models performances decreased with increased aridity index can be detected. The poor relationship between the aridity index and model performance can be attributed to the small sample of drier basins.

Fig.6 shows the performance difference of three models in each basin. Although we could not find a clear relationship between performance difference and aridity index, we could notice two dissimilar basins (ID=2&11) having relatively large differences indicating that the models are more sensitive in these areas and the model structure components are diversely important.

To quantify prediction uncertainty, scatter plots of AWISF for each model are shown in **Fig.7**. This figure shows that prediction uncertainty from NAM model is the largest while from TOPMODEL model is the smallest. This is probably due to the number of parameters needed to be calibrated. Over parameterized models bring larger uncertainties than parameter parsimonious ones. However, it is probably not the case for drier basins like Fenghe (ID=11) due to higher climatic variability.

6. CONCLUSIONS

With the objective of further identifying how model performances vary with climate conditions, three hydrological models were respectively applied to 11 basins located in Asian countries with

different climate conditions. The NSGAII and two widely adopted criteria viz. NSE and HMLE were used for calibration. Summarized results from this study are as follows:

(a) An automatic calibration procedure based on multi-objective optimization is implemented for HYMOD, NAM and TOPMODEL. The models performances are good in relatively humid basins except the basin with the mildest slope.

(b) The tendency that models performances decreased with increased aridity index was observed in terms of HMLE and this tendency might help us improve understanding of the model performances in different climates and select appropriate models for catchments in different climates. Furthermore, the poor relationship can be attributed to the selection of fewer drier basins. Future work is needed in the direction of including more basins located in a wider range of climatic conditions.

(c) The uncertainty in the model prediction was quantified from the ensemble of predicted model using an AWISF index. Results show that prediction uncertainties are larger for over parameterized models compared to parameter parsimonious ones.

Future work is needed to explore the relationship between model performance and hydroclimatic conditions and also geographic conditions. Understanding of differences in model performance and uncertainty, as shown in this study, might help us to identify the key hydrological processes/components for improvement of hydrological modeling. In addition, this would also be used to improve the multi-model ensemble and regionalization of hydrological parameters²⁰).

ACKNOWLEDGMENT: The authors express special thanks to VA, VA participants for providing data; MEXT and COE program at University of Yamanashi for financial support.

REFERENCES

- 1) Singh, V. P. and D. K. Frevert: *Watershed models*, CRC Press, Boca Raton, Fla, 2006.
- 2) R. Lidén and J. Harlin, Analysis of conceptual rainfall-runoff modeling performance in different climates, *J. of Hydrol.*, Vol.238, pp.231-247, 2000.
- 3) Atkinson, S. E., R. A. Woods, and M. Sivapalan, Climate and landscape controls on water balance model complexity over changing timescales, *Water Resour. Res.*, Vol.38, 1314, 2002.
- 4) Van Werkhoven K., T. Wagener, and P. Reed *et al.* Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resour. Res.*, Vol.44, 1429, 2008.
- 5) Samuel, J. M., M. Sivapalan, and I. Struthers, Diagnostic analysis of water balance variability: A comparative modeling study of catchments in Perth, Newcastle, and Darwin, Australia, *Water Resour. Res.*, Vol.44, 6403, 2008.
- 6) Vivek K. Arora, The use of the aridity index to assess climate change effect on annual runoff, *J. of Hydrol.*, Vol.265, pp.164-177, 2002.
- 7) Thorsten Wagener, Douglas P. Boyle, and Matthew J. Lees *et al.*, A framework for development and application of hydrological models, *Hydrol. Earth System Sci.*, Vol.5, pp.13-26, 2001.
- 8) Nielsen, S. A., Hansen, E., Numerical simulation of the rainfall runoff process on a daily basis, *Nordic Hydrol.*, Vol.4, pp.171-190, 1973.
- 9) K.J. Beven and M.J. Kirkby, A physically-based variable contributing area model of basin hydrology, *Hydrological Sciences Bulletin* Vol.24, pp.43-69, 1979.
- 10) Beven and J. Freer, A dynamic TOPMODEL, *Hydrological Processes* Vol.15, pp.1993-2011, 2001.
- 11) Moore, R.J., The probability-distributed principle and runoff production at point and basin scale, *Hydrological Sciences Journal* Vol.30, pp.273-297, 1985.
- 12) DHI, NAM technical reference and model calibration documentation, *DHI water and environment*, HØrsholm Denmark, 2000.
- 13) Ahn C.H. and Tateishi R., Development of Global Evapotranspiration and Water Balance Data Sets. *Journal of the Japan Soc. Photogrammetry and Remote Sensing*, Vol.33, pp.48-61, 1994.
- 14) Milly, P. C. D., and K. A. Dunne, Macroscale water fluxes: 2. Water and energy supply control of their interannual variability, *Water Resour. Res.*, Vol. 38, 1206, 2002.
- 15) Clark, M. P., A. G. Salter, D. E. Rupp, R. A. Woods, J. A. Vrugt, H. V. Gupta, T. Wagener, and L. E. Hay, Framework for Understanding Structural Errors (FUSE): A modular framework to diagnose differences between hydrological models, *Water Resour. Res.*, Vol. 44, W00B02, 2008.
- 16) Y. Tang, P. Reed, and T. Wagener, How effective and efficient are multiobjective evolutionary algorithms at hydrologic model calibration? *Hydrol. Earth Syst. Sci.*, Vol. 10, pp. 289-307, 2006
- 17) Sorooshian, S. and Dracup, J.A., Stochastic parameter estimation procedures for hydrologic rainfall-runoff models: correlated and heteroscedastic error cases. *Water Resour. Res.*, Vol.16, pp. 430-442, 1980.
- 18) Dev, K., Agrawal, S., Pratap, A., Meyarivan, T., A fast and elitist multi-objective Genetic Algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* Vol.6, pp. 182-197. 2002.
- 19) Veldhuizen, D., Multiobjective Evolutionary Algorithms: Classifications, Analyses, and New Innovations. PhD thesis, Air Force Institute of Technology, Wright-Patterson AFB, Ohio, 1999.
- 20) Bastola, S., H. Ishidaira, and K. Takeuchi, Regionalisation of hydrological model parameters under parameter uncertainty: A case study involving TOPMODEL and basins across the globe, *J. of Hydrol.*, Vol.357, pp.188-206, 2008.

(Received September 30, 2008)