EVALUATION OF PREDICTIVE UNCERTAINTY IN DISTRIBUTED RAINFALL-RUNOFF MODELS

Freddy SORIA[1], So KAZAMA[2] and Masaki SAWAMOTO[3]

[1]Member of JSCE, M. Eng., Graduate Student, Civil Engineering Dept., Tohoku University (Aobayama 6-6-06, Sendai, Japan)

[2] Member of JSCE, D. Eng., Associate Professor, Graduate School of Environmental Studies, Tohoku University (Aobayama 6-6-06, Sendai, Japan)

[3] Member of JSCE, D. Eng., Professor, Civil Engineering Dept., Tohoku University (Aobayama 6-6-06, Sendai, Japan)

This paper aims to raise the growing importance of predictive uncertainty evaluation in distributed modeling. The ideas are illustrated by applying a particular rainfall-runoff model in four catchments with different characteristics. Sensitivity analysis and parameter identifiability, as complimentary uncertainty measures allowed us to individually evaluate the suitability of model components. The Sobol implementation is affected by the sample size and the correlation degree among parameters, observed after the total sum of individual variance contributions exceeded a theoretical total variance of *1.0*, in magnitudes that varied from 5% in dry season to 60% in wet season, turning more difficult straightforward interpretations. Topographic index was used to judge the distributed performance of the model, adequately describing patterns of total discharge in homogeneous catchments with slopes ranging from 10% to 30%, and failing in less homogeneous catchments with low slope landscape.

Key Words: *Sensitivity indices, Sobol method, identifiability, topographic index.*

1. INTRODUCTION

Distributed models are an accessible tool in the analysis of hydrological processes and response prediction in watersheds. Model sophistication has greatly improved despite the small advances in data measurement techniques. Consequently our poor understanding of driving forces remains, and imposes a limit on our confidence in model outputs. Predictive uncertainty has increased together as well as the necessity of their evaluation.

Recognizing the limitations of current models and being aware of the necessity to give the final user a measure of the robustness in our predictions, the objective of this work is to set the importance of complimentary uncertainty measures in the evaluation of distributed models, having present the necessity to improve our knowledge on the model used and its components, increase the reliability in model results, and improve our knowledge on underlying hydrological processes. Uncertainty due to input data is not under the scope of this manuscript, therefore is not included in the analysis.

2. METHODOLOGY

(1) Study area and distributed model
a) Study area and data

Two systems with different topographic characteristics are studied: Natori basin (center of Miyagi Prefecture, northeastern Japan), and Caine basin (center of Bolivia). Natori basin ($1015km^2$) has mountainous landscape, with altitudes that vary from 0m to 1470m relative to mean sea level (msl), and land use characterized by urban areas, forests and crop fields (according to the Geographical Survey Institute, Japan). Caine River basin ($10214km^2$ at Molineros stream gauge station) is a semiarid region where evapotranspiration rate is higher than precipitation. Characteristics in Caine are heterogeneous, altitudes vary from 2500m to 2700m (relative to msl), and land use is agricultural with urban and industrial interferences. Two Caine sub basins are also considered: Huarmi ($1784km^2$), a valley with intermediate and fine granulometry (gravel, sand, lime and clay), and Taquiña River ($19km^2$), a mountainous steep catchment. Respect to

hydrometeorological data, hourly precipitation records in Natori basin are taken from 3 stations of the Automated Meteorological Data Acquisition System (AMeDAS), daily discharge records are taken from Yokata gauging station, and evapotranspiration is estimated from Normalized Difference Vegetation Indices derived from remote sensing data[1]. In Caine basin precipitation is taken from 27 stations of the National Meteorological and Hydrological Service (11 on daily basis and 16 with monthly resolution), and evapotranspiration is calculated using Penman Method with data from 9 stations[2]. Daily records at Molineros are used for analysis in Caine, meanwhile mean monthly discharge generated with a lumped model[2] is considered for Huarmi and Taquiña for all the evaluation period. In all cases, areal precipitation and areal evapotranspiration are estimated using the inverse distance method. Respect to topographic data, rectangular grids of elevation are considered at 250m resolution in Natori, and 250m to 1000m resolutions in Caine and sub basins. To simplify the analysis, the evaluation in Natori is carried for the period 1999-2000, and for the period 1974-1975 in Caine, Huarmi and Taquiña.

b) Distributed model

To illustrate the ideas, a distributed model structured with a subsurface and a groundwater reservoir[3] is used. There, overland flow routing is described by the kinematic wave equation, derived from the continuity equation as follows

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} - q = 0 \qquad (1)$$

$$Qdt + (r - E)dxBdt - \left(Q + \frac{\partial Q}{\partial t}dxdt\right) = \frac{\partial A}{\partial t}dxdt \quad (2)$$

where A is flow cross-sectional area, t time, Q discharge, x distance along the longitudinal axis of the watercourse, q lateral inflow or outflow per lineal distance along the watercourse, r rainfall rate, E evapotranspiration rate, B surface flow width, and h surface flow depth.

An infiltration model estimates abstractions due to infiltration. There, soil saturation (S_s) and initial infiltration (F_o) are assumed linearly varied as

$$S_s, F_o = S_{\min}, F_{\min} + \frac{S_{\max}, F_{\max} - S_{\min}, F_{\min}}{dayrate}dr \quad (3)$$

where $S_{max,min}$, $F_{max,min}$ are maximum and minimum conditions of S_s and F_o respectively, *dayrate* are the assumed days of rain to soil saturation, and *dr* are days of rain in the period. Later, the actual infiltration F that is subtracted to h, and recharges groundwater storage (when the soil has not reached saturation conditions S_{max}) is estimated as

$$F = F_o - \frac{F_o}{S_s}S_s \qquad (4)$$

Storage depth S is estimated with a storage function controlled by parameters K (dimensional) and m (dimensionless) as

$$S = Kq_s^{\ m} \qquad (5)$$

$$\frac{dS}{dt} = r_e - q_s \qquad (5a)$$

where q_s is runoff depth, and r_e is effective rainfall. In the model, above expression is adapted to account groundwater storage as

$$\frac{dS}{dt} = F + q_{in} - q_{out} \qquad (6)$$

where q_{in} and q_{out} are now referred to groundwater.

Overland flow is routed from cell to cell by flow direction vectors towards an explicitly defined river network. Once there, routing through main channels is described by a dynamic wave model that considers St. Venant equations of mass (Eq.1), and momentum conservation written as follows:

$$\frac{1}{g}\frac{\partial A}{\partial t} + \frac{1}{2g}\frac{\partial v^2}{\partial x} + \frac{\partial H}{\partial x} + \frac{n^2|v|v}{h^{4/3}} = 0 \qquad (7)$$

where g is gravity acceleration, H water level, v flow velocity, and n Manning's friction factor. Eq.7 is hence used to calculate water depth and velocity.

(2) Sensitivity analysis

Sensitivity analysis is commonly considered to identify sensitive factors. Methods are selected according to the problem addressed and the characteristics of the model used. One at a time (OAT) experiments[4] are screening tests, commonly used to evaluate feasible model parameter sets that may adequately describe a system. In OAT tests parameter values are varied one by one, within ranges and rates decided by modelers' experience, until is found a desired agreement between observations and calculations. When looking for more rigorous methods to quantitatively perform the analysis, variance based methods[4] are attractive, since can be used to estimate the variance contribution of model parameters to model output.

(3) Sensitivity indices

For uncorrelated factors, variance based techniques estimate the contributed variance from x model factors to model output $Y=f(x_i,x_j,..,x_k)$. To calculate sensitivity indices, the method establishes that the variance V of the model output can be decomposed into summands of decreasing dimensions as

$$V = \sum_i V_i + \sum_i \sum_{j>i} V_{ij} + .. + V_{ij..k} \qquad (8)$$

where

$$V_i = V\big(E\big(Y|x_i\big)\big) \qquad (9a)$$

$$V_{ij} = V\left(E\left(Y\middle|x_i, x_j\right)\right) - V_i - V_j \qquad (9b)$$

and $E(Y/x_i)$ is the expectation of Y conditional on x_i. The decomposition contains k terms of first order, $\frac{k(k-1)}{2}$ terms of second order and further until the last term. Each term can be computed by Monte Carlo integrations, and the results can be finally used to define sensitivity indices of different orders as

$$S_i = V_i / V \qquad (10)$$
$$S_{ij} = V_{ij} / V \qquad (11)$$
$$S_{Ti} = 1 - V_{-i} / V \qquad (12)$$

where S_i is the first order index denoting the main effect of parameter i, S_{ij} the second order index denoting the effect of the interaction of parameters i and j, V_{-i} denotes the influence of all the parameters except i, and S_{Ti} is the total index that represents the main effect of parameter i as well as its interactions. S_{Ti} is useful in the identification of input factors (i.e. model parameters) non-influential in the output variance. Monte Carlo-based Sobol' method was used to estimate the indices, and numerical schemes are detailed in the works of Saltelli[4] and Tang[5].

(4) Objective functions

The predictive potential of a model is usually evaluated using statistical measures. Current models are not capable to describe all aspects of the hydrograph with a single parameter set[6], therefore is common to consider many objective functions (OF). The measures used here are based on a normalized form of squared residuals (differences between observed O_i and computed data C_i) and ordinary error measures. The former emphasize performance during high flows since are sensitive to means and variances, meanwhile the latter evaluate concordance between observed and simulated time series. Nash-Sutcliffe dimensionless Coefficient of Efficiency CE (Eq.13) varies from minus infinity for poor model performance, to *1.0* for good model performance; zero would indicate that the observed mean O_m is as good predictor as the model, while negative values would mean the opposite[7]. The Root Mean Square Error RMSE (Eq.14) follows the same tendency, with the advantage that has the same units as the variable analyzed. For volumetric discharge evaluation, the dimensionless Relative Volumetric Error RVE (Eq.15) is more suitable.

$$CE = 1 - \frac{\sum(O_i - C_i)^2}{\sum(O_i - O_m)^2} \qquad (13)$$

$$RMSE = \sqrt{\frac{1}{N}\sum(C_i - O_i)^2} \qquad (14)$$

$$RVE = \frac{\sum(O_i - C_i)}{\sum O_i} \qquad (15)$$

(5) Parameter identifiability

Parameter identifiability is usually analyzed through "OF vs parameter value" plots, and can be used to verify the suitability of a model component to describe a process in a system[7]. This is sometimes difficult to evaluate when parameters highly sensitive seem not recognizable.

It has been observed that identifiability can be affected by parameter interdependences, their temporal and spatial variations, lack of information in the data used, or the inadequacy of the OF to represent their impact in model performance, therefore its evaluation has been suggested to be done during wet periods, where may occur stronger relationships between model response and hydroclimatic conditions [8].

(6) Prediction in ungauged locations

Continuous observations of hydrologic variables are fundamental for reliable model outputs. The problem arises when we are required to evaluate the model performance at ungauged sites. Prediction in Ungauged Basins PUB is an initiative of the International Association of Hydrological Sciences IAHS, and is an issue of current concern.

As an alternative to statistical methods requiring a high number of gauged locations, it is important to recognize the role of landscape properties in the spatial variability of hydrological processes. Those properties have been used by modelers, among which Beven's TOPMODEL (Topography based Hydrological Model) is one of the most cited in the hydrological literature. TOPMODEL's simplicity comes from the use of the topographic index $\kappa = ln(a/tan\beta)$, where a is the area draining to a given cell and $tan\beta$ is the slope angle. κ is a similarity index indicating that all cells with the same index value would have similar hydrological response. High values of κ will be caused by long slopes, upslope contour convergence or low slope angles that indicate saturation potential. Distributed patterns of those attributes can be compared against the outputs of a model to judge their distributed performance, helping us to improve our knowledge referred to the nature of process occurrence. The idea is simple and can help us to quantitatively analyze our results, although will not be applicable in dry regions where saturation is unlikely to occur.

3. RESULTS AND DISCUSSION

(1) Preliminary computational experiments

OAT experiments were used to estimate model sensitivity to variations in parameter values (i.e. model components). An initial warming-up computation determined initial conditions, and 9

model parameters (i.e. 3 model components) were varied on ±10% increments within assumed ranges. In Natori basin, it was assumed that the best model performance would be accomplished with a 250m grid resolution (the finest available), and 10s for calculation time step. Performance was evaluated based on CE, RVE and RMSE measures (Table 1), and parameters of the storage component (K, m) were the most sensitive. Infiltration component (i.e. parameters *Smax*, *Smin*, *fmax*, *fmin*, *dayrate*) was identified as insensitive.

Monte Carlo-based methods demand high computational effort in distributed models, then towards future steps was thought necessary to find the coarsest acceptable resolution for each basin. As a result, computational experiments continued in Natori for grid resolutions of 500m and 1000m. Later, additional empirical reasons (e.g. largest grid size to represent desired certain features) pointed to the 500m grid resolution as the most adequate.

Following the process and considerations above, model performance was evaluated on 3 systems arbitrarily classified: a large watershed (Caine, with channel runoff assumed as dominant), a medium size watershed (Huarmi, with predominant flat slopes, and groundwater processes assumed as dominant), and a small mountainous system (Taquiña, assumed dominated by overland flow processes). Model parameters spatially distributed within three ranges of slopes were considered in OAT tests with grid resolutions of 250m, 500m, 1000m, 2000m, 4000m and 8000m. Different performances were reached (Table 1), and the storage module was the most sensitive resembling the results in Natori despite the differences in geography. Finally, the coarsest acceptable grid resolutions selected were 8000m for Caine, 1000m for Huarmi, and 250m for Taquiña.

OAT tests are simple and the computational cost is low if the experiments are adequately planned.

Table 1. OAT experiments: Model performance evaluation (the coarsest resolutions empirically selected are marked).

		Grid size (m)					
		250	500	1000	2000	4000	8000
Caine	CE	---	---	0.65	0.84	0.84	0.91
	RVE	---	---	0.42	0.35	0.29	0.03
	RMSE (m^3/s)	---	---	32.66	22.53	22.36	16.43
Huarmi	CE	0.45	0.49	0.50	0.34	0.58	---
	RVE	-0.04	0.17	0.19	-0.22	0.05	---
	RMSE (m^3/s)	10.00	9.63	9.51	10.90	8.71	---
Taquiña	CE	0.85	---	---	---	---	---
	RVE	0.09	---	---	---	---	---
	RMSE (m^3/s)	0.10	---	---	---	---	---
Natori	CE	0.59	0.85	0.60	---	---	---
	RVE	0.10	0.08	0.55	---	---	---
	RMSE (m^3/s)	46.47	27.61	45.61	---	---	---

The main inconvenient might be that they may only evaluate a region within the universe of behavioral parameter sets; SA is then considered.

(2) Sensitivity indices

Sobol' method was used to estimate uncertainty contribution from input factors, in this work considered only as the model parameters, leaving input parameters out of the analysis. Sobol sampling generated 286 sample sets of 9 model parameters uniformly distributed (i.e. 3 model components). The hydrological model was run for the generated sample sets, and the outputs (total discharge) are used by Sobol' scheme to construct time series of sensitivity indices. Monte Carlo demanded high computational effort, and the degree of parameter correlations affected the implementation of the method here considered. The sum of variance contributions overestimated the total theoretical variance of *1.0* in magnitudes varying from time step to time step. In dry conditions overestimation was found to be between 5% to 25% in the worst case (June 1st), implying low predictive uncertainty. As expected, overestimation in wet conditions was higher (see Figure 2), with maximum values before hydrograph rise (33% of overestimation), and hydrograph recession (48% of overestimation), meanwhile low overestimation was observed during peak occurrence (22%, 0% and 10% respectively).

Hydrological processes vary in time, for instance time scale of the analysis might be defined by such aspect. Monthly time series (e.g. Fig.3) were easier to evaluate than daily ones (e.g. Fig.2), and can still be meaningful. Storage processes (i.e. slow time of occurrence) have permanent effect even during dry periods, and could be used on system characterization. Channel flow has faster response, therefore its effect in total variance vary according to soil moisture conditions, being important in catchments with well defined channel configuration. The latter was observed in all systems except on the fast-response mountainous catchment (not shown here), where storage and infiltration dominated system response.

Sensitivity indices were used to evaluate model components and to identify uncertainty sources, consequently improving our knowledge about underlying hydrological processes. The application of the methodology does not seem restricted by geographic conditions, and the quantitative properties of the method should be recognized. Finally, stronger relationships between model response and hydroclimatic conditions can be notorious during events occurrence[9], which as found here, can also increase predictive uncertainty and demand major attention.

(3) Parameter identifiability

Identifiability plots are complimentary measures that can be used to indicate model suitability to represent the response of a given system[9]. A parameter would be identifiable when different values drive different model responses, and describing recognizable trends[9].

The structure of a model will not be suitable to all kind of topographic configurations; uncertainty will also vary. Catchments with homogeneous topography and gentle slopes can present higher number of identifiable components (e.g. Fig.4) than catchments with steeper slopes and more heterogeneous characteristics (e.g. Fig.5). Storage component was the most sensitive and the most identifiable, hence can be used to describe variations in catchment state. Parameter m varied from 1.8 to 2.5 in Natori, implying more excess runoff stored than in the steep mountainous catchment where m varied from 2.5 to 5.0. In catchments where model performance was poor (Huarmi basin) identifiability plots did not carry much information but showing that particular model components are associated to high predictive uncertainty.
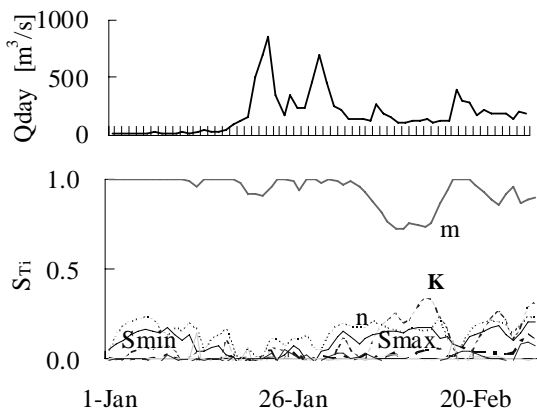
(4) Distributed performance of the model

Models conceptually based on the role played by landscape properties in catchment response can be useful to evaluate the distributed performance of a model. Topographic index is an interpretation of the reality that indicates potential to saturation (i.e. runoff occurrence) on locations with high values. It can be useful although may not hold for all cases. In this study, its distributed patterns were compared against those of total discharge during representative time steps, in order to qualitatively determine the distributed performance of a model. As a result, two catchments presented discharge patterns that seemed adequately described by the index (i.e. recognizable relationship): a middle-size homogeneous catchment (Fig.6, upper left), with gentle slope and discharges ranging from 100l/s to 900l/s, and by indices from 1 to 7 respectively, and a steep catchment with average slope of 30% (Fig.6, upper right), discharges in the range 5l/s to 75l/s, and described by indices from 0.5 to 6.5. On the other hand, discharge patterns from a middle-size heterogeneous catchment with predominant flat slopes (Fig.6, lower left), and from a large heterogeneous

Fig. 2 Daily time series comparison between observed hydrograph and total sensitivity index; Caine basin.

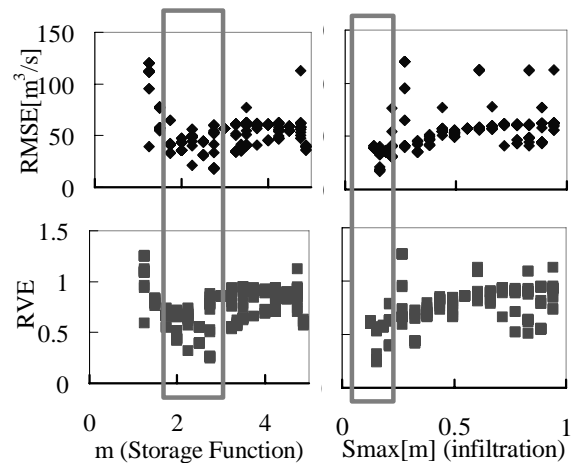**Fig. 4** Identifiability plots for September, 1999 in Natori basin.

Fig. 3 Monthly time series comparison between observed hydrograph and total sensitivity index; Caine basin.
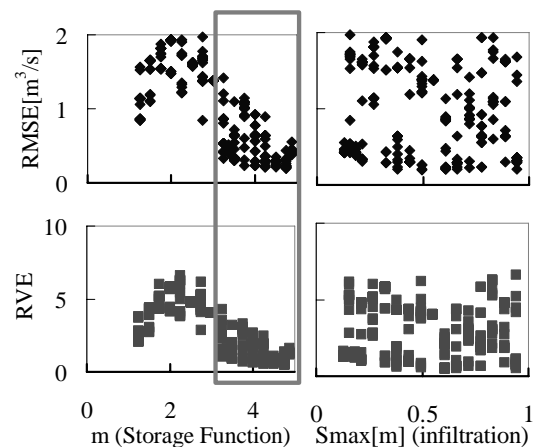

Fig. 5 Identifiability plots for Taquiña River basin.

catchment (Fig.6, lower right), were not adequately described since low slope angles caused high indices that blurred recognizable relationships. In the latter case, field data becomes fundamental for an adequate representation of discharge patterns.

4. CONCLUSIONS

The objective of the manuscript was to present the use of complimentary uncertainty measures that would help to improve the reliability on distributed model predictions, and determine the suitability of model structure to describe a given catchment. OAT experiments were first used to evaluate an initial behavioral range of model parameters, and later used to aid the evaluation of the most adequate grid scale to carry model runs. Finally, three uncertainty measures were considered: sensitivity indices, identifiability measures and landscape properties.

Sobol indices can be used on any mathematical model, although the high computational cost and the degree of correlation among parameters can complicate straightforward interpretations. In this study both aspects have influenced the temporal behavior of the indices, since the total theoretical variance was overestimated in magnitudes that varied from 5% in dry conditions, to 65% in wet conditions (when the complexity of processes interaction increase). On the other hand, the importance of the temporal evolution of sensitivity indices should be also recognized in terms of the inherent dynamic properties of a system.
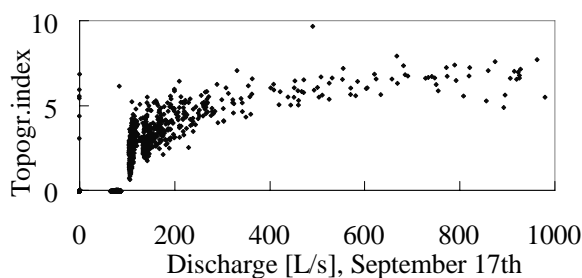

Fig. 6 Spatial model performance evaluated under the topographic index criteria. Natori basin.
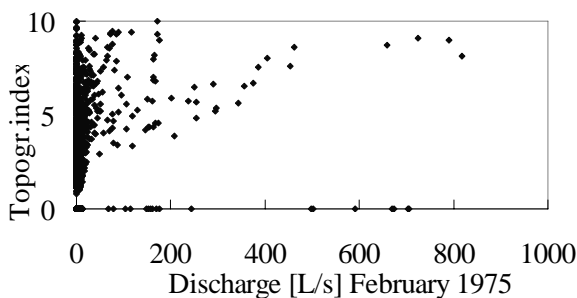

Fig. 7 Spatial model performance evaluated under the topographic index criteria. Huarmi basin.

Identifiability plots, although restricted to gauged locations, were used as complimentary measures that allowed us to evaluate the characteristics of the response surface, and the probable behavioral space. Later, distributed performance is also an uncertain aspect to evaluate, and was covered here by comparing distributed patterns of the topographic index and the total discharge calculated. The approach was found adequate on homogeneous catchments with indices in the range 0.1 to 6.5 and average slope from 10% to 30%, and inadequate in heterogeneous systems with preponderance of low slope angles, demonstrated to be highly uncertain, hence demanding more data for a reliable representation. Uncertainty due to input data will affect the results presented, and are out of the scope of this manuscript; in any case, to analyze both separately (model structure and input data) is though to be adequate. Finally, both sensitivity and identifiability measures can be highly influenced by the sample space size considered, and their application in distributed models still requires further research.

REFERENCES

1) Nourbaeva G., Kazama S., Sawamoto M.: Assessment of daily evapotranspiration using remote sensing data, *Environmental Informatics Archives*, Vol.1, pp.421-427, 2003.
2) Molina, J., Alurralde, J., Villarroel, E.: Water rights and water allocation in Andean basins, *Proc. of the Int. Symp. on Managing Water Supply for Growing Demand*, pp.51-58, 2006.
3) Kazama, S., Hyejin, K., Sawamoto, M.: Uncertainty of morphological data for rainfall-runoff simulation, *Proc. of the Int. Conf. on Sustainable Water Resources Management in the Changing Environment of the Monsoon Region*, Vol.1, pp.400-406, 2004.
4) Saltelli, A.: *Sensitivity analysis*, edited by Saltelli, A., Chan, K., and Scott, M., pp.3-12,167-197, John Wiley, Chichester, 2000.
5) Tang, Y., Reed, P., Wagener, T., van Werkhoven, K.: Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation, *Hydrol. Earth Syst. Sci. Discuss.*, Vol.3, pp.3333-3395, 2006.
6) Wagener, T., Wheater, H.S., Gupta, H.V.: *Rainfall-Runoff Modeling in Gauged and Ungauged Catchments*, Imperial College Press, pp. 83,91, 2004.
7) Wilcox, B., Rawls, W., Brakensiek, D., Wight, J.: Predicting runoff from rangeland catchments: A comparison of two models, *Water Resour. Res.*, Vol.26, pp.2401-2410, 1990.
8) Yapo, P., Gupta, H., Sorooshian, S.: Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data, *J. Hydrol.*, Vol.181, pp.23-48, 1996.
9) Wagener, T., and Wheater, H.S.: Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty, *J.Hydrol.*, Vol.320, Issue 1-2, pp.132-154, 2006.   **(Received September 30, 2007)**