SEQUENTIAL MONTE CARLO METHODS FOR REAL-TIME FORCASTING USING MULTIPLE HYDROLOGIC MODELS

Seong Jin Noh[1], Yasuto Tachikawa[1], Michiharu Shiiba[1], Sun Min Kim[1], Kazuaki Yorozu[1]

[1]Kyoto University, Kyoto, Japan

ABSTRACT

The applications of data assimilation techniques have been increasing to enhance forecasting considering various sources of uncertainties. As recently introduced data assimilation techniques, Sequential Monte Carlo (SMC) methods are applied for real-time forecasting of river discharge. Two hydrologic models, Storage Function (SF) model and Water and Energy transfer Processes (WEP) model, are applied for the middle-sized Japanese catchment by SMC methods. In case of SF model, two different approaches of SMC methods are implemented: state only updating and dual state-parameter updating. Various SMC methods such as ASIR, RPF with MCMC move step and SIR particle filter are also compared. In case of WEP model, accumulated soil moisture is used as hidden state variable and spatial distribution of soil moisture is updated by every time step. Despite limited information on physical characteristics of soil and aquifer, simulation results of WEP model using SMC show relatively enhanced forecast compared to those of SF model.

1 INTRODUCTION

Uncertainty in the predictions of science is an issue of great current relevance in relation to the impacts of climate change, protecting against natural and man-made disasters, pollutant transport and sustainable resource management (Beven, 2009). Recently, the applications of data assimilation techniques have been increasing to enhance forecasting considering various sources of uncertainties. Among data assimilation techniques, sequential Monte Carlo (SMC) methods are Bayesian learning process in which the propagation of all uncertainties is carried out by a suitable selection of randomly generated particles without any assumptions about nature of the distributions. SMC has the advantage of being applicable to non-linear, non-Gaussian state-space models. Over the last few years, the application of these powerful and versatile methods has been increasing, e. g., pattern recognition, weather forecasting, bioinformatics and hydrology (Moradkhani et al., 2005; Smith et al., 2009). In this study, SMC methods are applied for real-time forecasting of river discharge using multiple hydrologic models from conceptual one to process-based and spatially-distributed one. As hydrologic models, Storage Function (SF) model and Water and Energy transfer Processes (WEP) model (Jia et al., 2009) are implemented for the middle-sized Japanese catchment. In case of SF model, data assimilation is performed by two different approaches of SMC methods: state only updating and dual state-parameter updating. In dual state-parameter updating scheme, probabilistic distribution of parameters is treated as time-varying and structural uncertainty of hydrologic models is assessed using SMC, while state only updating scheme considers ensemble of different state variables having the same parameter set calibrated from previous events. In case of WEP model, state only updating scheme is performed in the preliminary simulation. Accumulated soil moisture is used as hidden state variable and spatial distribution of soil moisture is updated every time step.

2 METHODOLOGY

2.1 Sequential Monte Carlo

Consider a generic dynamic state-space model which can be described as follows:

$$x_t = f(x_{t-1}, \theta, u_t) + \omega_t \qquad \omega_t \sim N(0, W_t) \tag{1}$$

$$y_t = h(x_t, \theta) + v_t \quad v_t \sim N(0, V_t) \tag{2}$$

where $x_t \in \Re^{n_x}$ is the n_x dimensional vector denoting the system state at time t. The operator $f:\Re^{n_x} \to \Re^{n_x}$ and $h:\Re^{n_x} \to \Re^{n_y}$ express the system transition in response to the forcing data u_t, parameters θ. ω_t and v_t represent the model and the measurement error, respectively.

SMC is based on point mass ("particle") representations of probability densities with associated weights (Arulampalam et al., 2002).

$$p(x_t \mid Y_t) \approx \sum_{i=1}^{N} w_t^i \delta(x_t - x_t^i) \tag{3}$$

where x_t^i, w_t^i denote the i th particle and its weight, respectively, and $\delta()$ denotes the Direc delta function.

Since it is usually impossible to sample from the true posterior PDF, an alternative is to sample from a proposal distribution, also called importance density, denoted by $q(x_t \mid y_t)$. The recursive weight updating could be derived as follows:

$$w_t^i \propto w_{t-1}^i \frac{p(z_t \mid x_t^i) p(x_t^i \mid x_{t-1}^i)}{q(x_t^i \mid x_{t-1}^i, y_t)} \tag{4}$$

Several variants of SMC methods such as Sequential Importance Resampling (SIR), Auxiliary SIR (ASIR) and Regularized Particle Filter (RPF) have been developed to overcome the degeneracy phenomenon, selection of importance density and sample impoverishment.

2.2 Parameter inference

For the unknown parameters, the concept of "artificial evolution" could be applied for. That means, parameter θ is replaced at each time adding an independent, zero-mean normal increment to the parameter as follows:

$$\theta_t = \theta_{t-1} + \zeta_t \quad \zeta_t \sim N(0, s^2 Var_k^\theta) \tag{5}$$

where, ζ_t is random noise, Var_k^θ is the variance of parameter particles at time k before resampling and s is a small tuning parameter.

3 IMPLEMENTATION

3.1 Study area

Hydrologic models were applied to the Katsura river catchment (Figure 1) to forecast the river flow through SMC methods. This catchment is located in Kyoto, Japan and covers an area of 1,100 km^2 (887 km^2 at the Katsura station). There are 13 rainfall observation stations and 6 river flow observation stations within the catchment. There is the Hiyoshi dam whose outflow data was used as the upper boundary condition of river discharge in both SF model and WEP model.

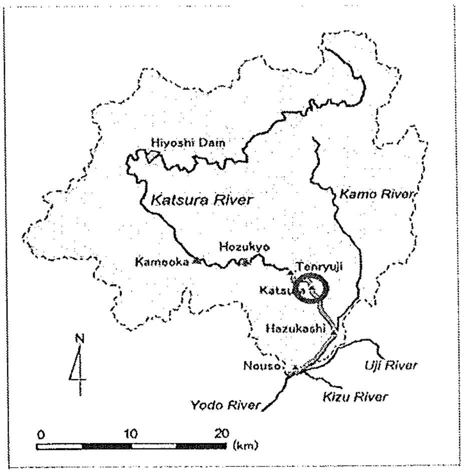


Figure 1. Katsura river catchment

3.2 Storage function model

The storage function (SF) model (Kimura, 1961) is one of the most commonly used conceptual models for flood runoff prediction in Asian countries due to its simple numerical procedure and proper regeneration of nonlinear characteristics of flood runoff. State-space form of SF model is as follows:

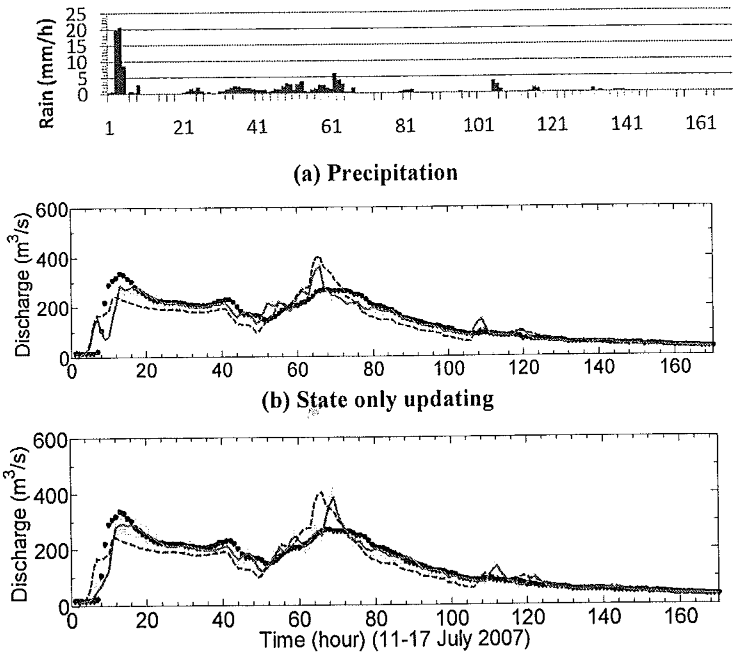
$$\frac{ds}{dt} = r_e(t - T_L) - \left(\frac{s(t)}{k}\right)^{1/p} - E(t) + \omega_t \tag{6}$$

$$q_{sim}(t) = \frac{A_{down}}{3.6}\left(\frac{s(t)}{k}\right)^{1/p} + q_{dam}(t - T_{dam}) \tag{7}$$

$$q_t = q_{sim}(t) + v_t \tag{8}$$

where, s is catchment storage, t is time (hour), r_e is effective rainfall (mm), E is evapotranspiration (mm) , A_{down} is the down area from the dam (km^2), q_{sim} is simulated river discharge (m^3/s), q_t and q_{dam} are observed discharge at Katsura station and Hiyoshi dam (m^3/s), T_L and T_{dam} are lag time parameter of catchment and dam (hour), respectively, and k, p are model parameter. ω_t and ν_t are the state and the measurement error, respectively.

SF model was simulated by SIR particle filter with 1,000 particles. Figure 2 shows 3-hour-lead forecasting results by two different approaches of SMC methods (state only updating and dual state-parameter updating) compared to observation and deterministic prediction. Forecast by dual state-parameter updating (Figure 2(c)) shows generally good conformity between observed and simulated discharge, while deterministic prediction and forecast by state only updating scheme (Figure 2(b)) do not to reproduce the river flow properly due to unsuitable parameter sets calibrated from previous events. Figure 3 illustrates traces of parameters during simulation. Initial distribution for parameters was given from prior information (calibrated from previous flood events).



(a) Precipitation

(b) State only updating

(c) Dual state-parameter updating

Figure 2. Hourly precipitation and river discharge from 11 to 17 July 2007. Black dots represent observed discharge. Blue line and area represent mean value and 95% confidential interval of 3-hr-lead forecast of the storage function model using SMC methods, respectively. Dashed line represents a deterministic modeling case.

3.3 WEP model

Water and Energy transfer Processes (WEP) model is a distributed hydrological model to be developed for simulating spatially variable water and energy processes in catchments with complex land covers (Jia et al., 2009). State variables include depression storage on land surfaces and canopies, soil moisture content, land surface temperature, groundwater tables and water stages in rivers, etc (Figure 4). Ensemble simulation of sixty particles was conducted on a multi-processing computer by Message Passing Interface (MPI) techniques. Accumulated soil moisture depth was selected as a hidden state variable and spatial distribution of soil moisture was updated every time step. Although information on physical characteristics of soil and aquifer was very limited in this catchment, simulation results of WEP model by SMC (Figure 5) show relatively enhanced forecasting compared to those of SF model.

4 CONCLUSION

Sequential Monte Carlo methods were applied for river flow forecasting using multiple hydrologic models. In case of SF model, data assimilation was performed by two different approaches of SMC methods: state only updating and dual state-parameter updating. Forecast by dual stat-parameter updating showed generally good conformity between observed and simulated discharge, while deterministic prediction and forecast by state only updating scheme did not to reproduce the river flow properly due to unsuitable parameter sets. In case of WEP model, accumulated soil moisture was used as hidden state variable and spatial distribution of soil moisture is updated by every time step. Despite limited information on physical characteristics of soil and aquifer in the applied catchment, simulation results of WEP model by SMC showed relatively enhanced forecasting compared to those of SF model.

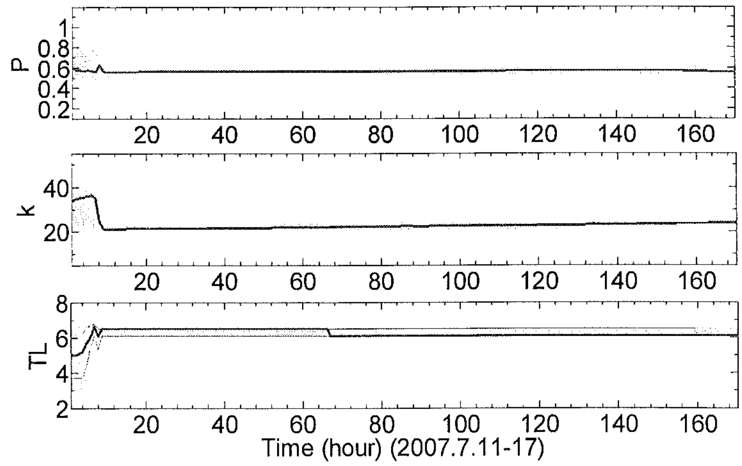


Figure 3. Traces of parameter k, P, TL of the storage function model using SMC methods from 11 to 17 July 2007. Black lines represent median value and gray area represents 95% confidential interval.

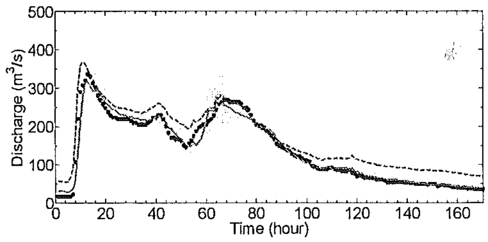


Figure 5. Hourly river discharge from 11 to 17 July 2007. Black dots represent observed values. Blue line and area represent mean value and 95% confidential interval of 3-hr-lead forecast of WEP model using SMC methods, respectively. Dashed line represents a deterministic case.

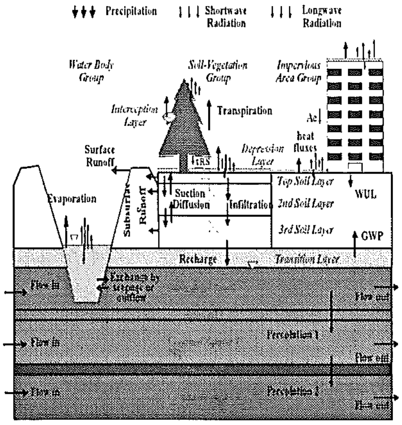


Figure 4. Vertical structure of WEP model

5 REFERENCES

Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T., 2002. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Trans. Signal Process., 50(2), 174-188.

Beven, K., 2009. Environmental modeling: an uncertain future?. Routledge: Oxon.

Kimura, T., 1961. The Flood Runoff Analysis Method by the Storage Function Model. The Public Works Research Institute, Ministry of Construction.

Jia, Y., Ding, X., Qin C., and Wang H., 2009. Distributed modeling of landsurface water and energy budgets in the inland Heihe river basin of China. Hydrol. Earth Syst. Sci., 13, 1849–1866.

Moradkhani, H., Hsu, K. L., Gupta, H., and Sorooshian, S., 2005. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using particle filter. Water Resour. Res., 41, W05012, doi:10.1029/2004WR003604.

Smith, P. J., Beven, K. J., and Tawn, J. A., 2008. Detection of structural inadequacy in process-based hydrological models: A particle-filtering approach. Water Resour. Res., 44, W01410, doi:10.1029/2006WR005205.