Simulation of Historical Rainfall over Indochina Peninsula

Hiroshi OKUMURA, Taichi TEBAKARI, Tianqi AO & Junichi YOSHITANI, **Public Works Research Institute of Japan**

Abstruct. *PWRI is developing a methodology to reconstruct historical rainfall over ungaged basins with only several ground observation data. In this study we present a new approach for PUB (Prediction in Ungaged Basins) applying the Kalman filter Hydrological Atmospheric Model (KHAM). This method is based on the state-space moedel of two parts. The one is the system model which can express the state of a hydrological phenomena and consider system noise. The other is the observation equation which can consider the observation noise. Putting hydrological observation data located at only several points into the KAHM, the whole spatial and temporal distritution of the hydrological state can be estimated . An optimal estimate can be expressed by the condional expectation of the state which is obtained from a computational solution of the stochastic method of least squares. KAHM will reconstruct past historical rainfall over the Chao Phraya river basin in 2000 with only 10 points observation rainfall data.*

State Space Model

The Kalman filter is based on a set of two equations. The one is the system model which can express the state of a hydrological phenomena, the other is observation equation. The system model can be written as;

$$x_{k+1} = F_k\, x_k + G_k\, v_k$$

where x_k is state vector at time k, F_k is state transition matirx, and G_k is driving matrix. In this study, rainfall is the state vector. The observation equation is given as;

$$y_k = H_k\, x_k + w_k$$

where y_k is observation vector at time k and H_k is observation matrix. The observation matrix specifies the position at which the observation is carried out.

Assumptions

The random variable v_k and w_k represent the system and observational noises. They are assumed to be independent each other, and with normal probability distributions:

$$v_k \sim N(0,\, Q), \quad w_k \sim N(0,\, R)$$

where Q is the system error covariance, and R is the observation error covariance. Optimal estimate $\hat{x}_k$ and estimate x_k^* are written as follows;

$$\hat{x}_k = E[x_k \,|\, Y_k], \quad x_k^* = E[x_k \,|\, Y_{k-1}]$$

These equations mean that the optimal estimate $\hat{x}_k$ and estimate x_k^* are the average of x_k knowing the measurements Y_k and Y_{k-1} respectively, where $Y_k = \{y_1, y_2, \cdots, y_k\}$. The covariances P_k and Γ_k are:

$$\begin{aligned} P_k &= \mathrm{cov}\{x_k \,|\, Y_k\} = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] \\ \Gamma_k &= \mathrm{cov}\{x_k \,|\, Y_{k-1}\} = E[(x_k - x_k^*)(x_k - x_k^*)] \end{aligned}$$

The whole process is assumed that the phenomena are steady state.

Formulation

The Bayes rule gives the following equation:

$$\mathcal{P}(x_k \,|\, Y_k) = \frac{\mathcal{P}(y_k \,|\, x_k)\, \mathcal{P}(x_k \,|\, Y_{k-1})}{\mathcal{P}(y_k \,|Y_{k-1})}$$

where $\mathcal{P}(x_k \,|Y_k)$ is the conditional probability. Using the asumption and the Bayes rule, optimal estimate, Kalman gain and estimate error covariance can be rewritten as follows;

Optimal Estimate

$$\hat{\boldsymbol{x}}_k = \boldsymbol{x}_k^* + \boldsymbol{K}_k\, (\boldsymbol{y}_k - \boldsymbol{H}_k\, \boldsymbol{x}_k^*)$$

Kalman Gain

$$\boldsymbol{K}_k = \boldsymbol{\Gamma}_k\, \boldsymbol{H}_k^T\, (\boldsymbol{R} + \boldsymbol{H}_k\, \boldsymbol{\Gamma}_k\, \boldsymbol{H}_k^T)^{-1}$$

Estimate Error Covariance

$$\boldsymbol{P}_k = (\boldsymbol{I} - \boldsymbol{K}_k\, \boldsymbol{H}_k)\boldsymbol{\Gamma}_k$$

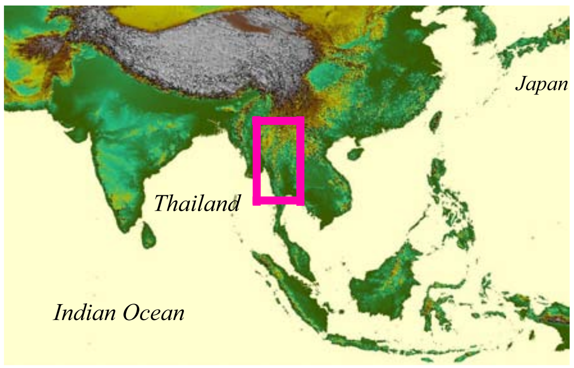
The algorithm of the Kalman filter Hydrlogic Atmospheric Model (KHAM) is written as follows:

do k, n :

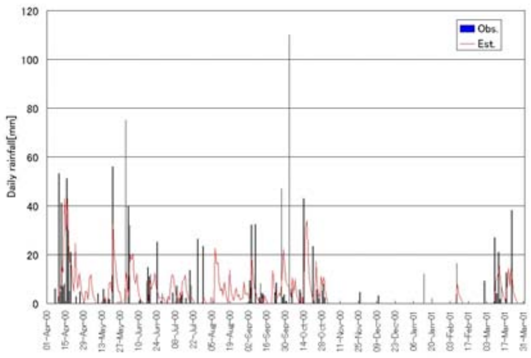
1. $\boldsymbol{\Gamma}_0 = \boldsymbol{v}_0, \quad \hat{\boldsymbol{x}}_{-1} = \hat{\boldsymbol{x}}_0$
2. $\boldsymbol{K}_k = \boldsymbol{\Gamma}_k\, \boldsymbol{H}_k^T (\boldsymbol{R} + \boldsymbol{H}_k\, \boldsymbol{\Gamma}_k\, \boldsymbol{H}_k^T)^{-1}$
3. $\boldsymbol{P}_k = (\boldsymbol{I} - \boldsymbol{K}_k\, \boldsymbol{H}_k)\boldsymbol{\Gamma}_k$
4. $\boldsymbol{\Gamma}_{k+1} = \boldsymbol{F}_k\, \boldsymbol{P}_k\, \boldsymbol{F}^T + \boldsymbol{G}_k\, \boldsymbol{Q}\, \boldsymbol{G}_k^T$
5. *if* $\|\mathrm{tr}\, \boldsymbol{P}_k - \mathrm{tr}\, \boldsymbol{P}_{k-1}\| < \epsilon$ then go to 6
   else go to 2
6. $\boldsymbol{x}_n^* = \boldsymbol{F}_n\, \hat{\boldsymbol{x}}_{n-1}$
7. $\hat{\boldsymbol{x}}_n = \boldsymbol{x}_n^* + \boldsymbol{K}_n(\boldsymbol{y}_n - \boldsymbol{H}_n\, \boldsymbol{x}_n^*)$

end do

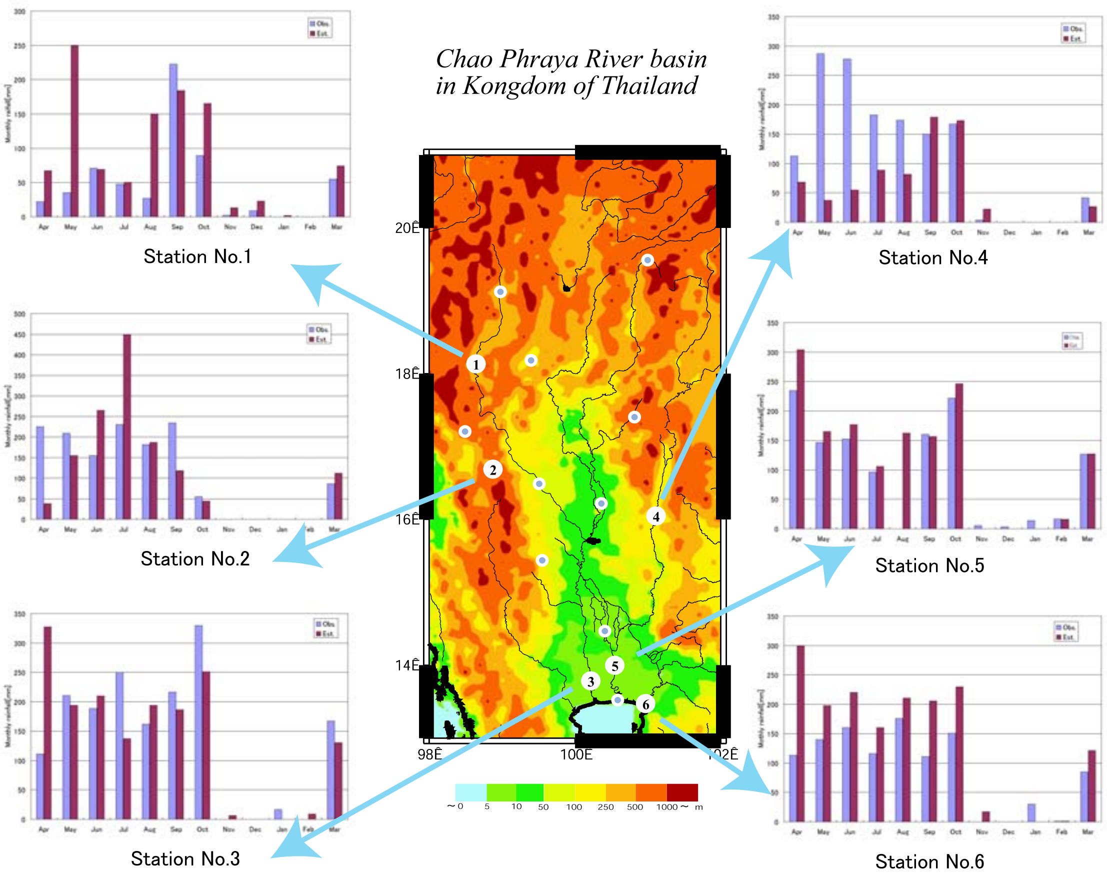
Algorithm of KHAM



Map of southern Asia and computational area



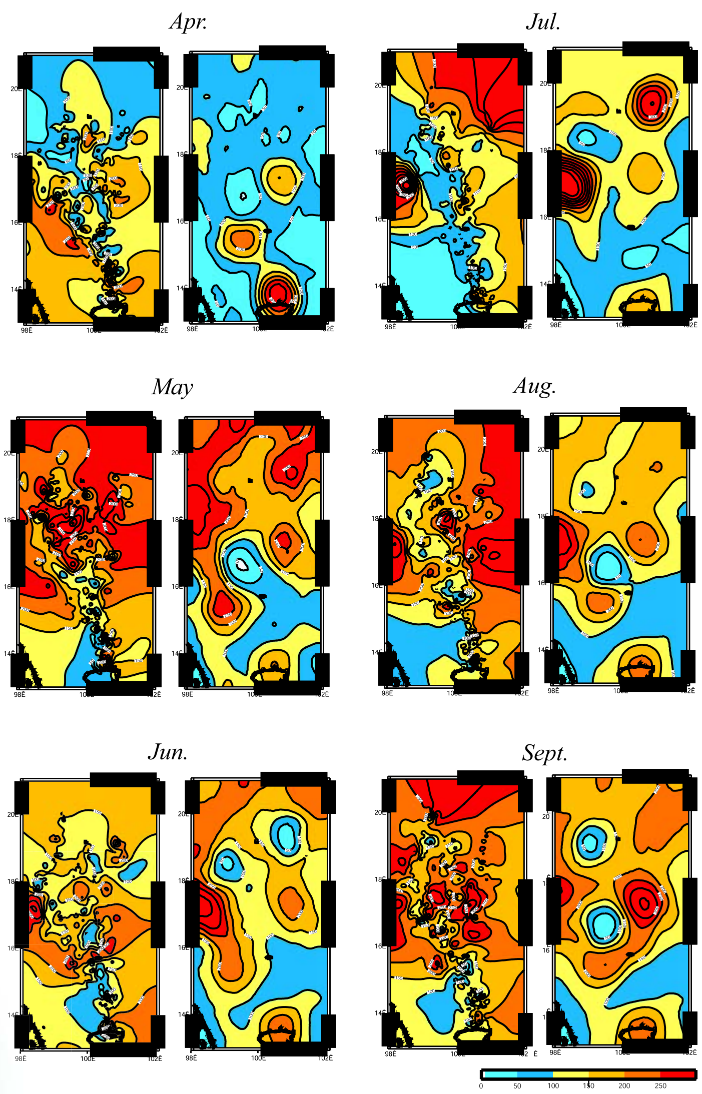
Time series of daily rainfall (observation vs estimation) at station No. 5



Comparison of observed and estimated monthly rainfall using by KHAM in 6 stations
(● denotes observation points. Number of observation points is 10 for input data.)

Concluding Remarks

KHAM can reconstruct historical rainfall of over 3,000 cells with only 10 points observation data of daily rainfall. In brind test for Chao Phraya River basin, monthly spatial and temporal distribution of past rainfall estimated by KHAM is good agreement with observation data. Using KHAM, the computation including the effect of the observation can be carried out. And it becomes possible to estimate rainfall not only in time but also in space considering the discrepancy between computation.



Comparison of observed and estimated monthly rainfall area. Each left figure is observed monthly rainfall. Each right figure is computed monthly rainfall obtained by KHAM.

This article is contributed to IUGG2003 at HW07.