

ENSEMBLE RAINFALL-RUNOFF PREDICTION WITH RADAR IMAGE EXTRAPOLATION AND ITS ERROR STRUCTURE

Sunmin KIM¹, Yasuto TACHIKAWA², Takahiro SAYAMA³ and Kaoru TAKARA⁴

¹Student Member, Graduate student, Dept. of Urban and Environmental Eng., Kyoto University
(Kyoto 606-8501, Japan) E-mail:sunmin@rdp.dpri.kyoto-u.ac.jp

²Member of JSCE, Dr. Eng., Associate Professor, DPRI, Kyoto University (Gokasho, Uji 611-0011, Japan)

³Member of JSCE, M. Eng., Assistant Professor, DPRI, Kyoto University (Gokasho, Uji 611-0011, Japan)

⁴Fellow of JSCE, Dr. Eng., Professor, DPRI, Kyoto University (Gokasho, Uji 611-0011, Japan)

New attempt of ensemble rainfall-runoff prediction is presented with radar rainfall prediction and spatial random error field simulation. A radar image extrapolation model gives deterministic rainfall predictions, and its prediction error structure is analyzed by comparing with the observed rainfall fields. With the analyzed error characteristics, spatial random error fields are simulated using covariance matrix decomposition method. The simulated random error fields successfully keep the analyzed error structure and improve the accuracy of the deterministic rainfall predictions; then the random error fields with the deterministic fields are given to a distributed hydrologic model to achieve an ensemble runoff prediction.

Key Words: radar rainfall prediction, error structure, random error field, ensemble simulation

1. INTRODUCTION

Nowcasting of precipitation is an essential prerequisite for real-time flood forecasting in operational hydrology. The term “nowcasting” is used to emphasize the specificity and shortness (0~3hrs) of rainfall forecast largely by radar image extrapolation, which has a history of nearly 50 years¹⁾. Even though the benefits of short-term precipitation forecasts are well known, this is acknowledged to be among the most challenging areas in hydrology and meteorology²⁾.

The early stage of the extrapolation methods takes the simple techniques in the extremes. It does not usually allow for the growth and decay of the rainfall intensities or nonlinear motion of the rainfall band. Main shortage of the techniques is that mainly because of its simplicity, the forecast accuracy decreases very rapidly within an hour³⁾. In a study of the forecasts accuracy improvement, elaborate nonlinear extrapolation schemes only give negligible improvement or even worse results than linear position extrapolation⁴⁾.

While the simple extrapolation techniques have been the mainstream of nowcasting until recently, vast research efforts by many hydrologists and meteorologists have been done over several decades

with introductions of many new schemes. These new schemes include mathematical and stochastic models integrated with a meteorological component^{5),6)} and hybrid models, which are combination of Numerical Weather Prediction (NWP) models and image extrapolation models^{7),8)}. In addition, there have been complex statistical approaches like with fractal generation algorithms⁹⁾ or artificial neural networks¹⁰⁾.

Even with the vast research efforts, it has been hard to find a notable improvement of the forecast accuracy from the chaotic nature. As an alternative to the accuracy improvement, numerical models have been run in an ‘ensemble mode’ for the last decade or so. Statistical outcome is obtained from the ensemble simulations mostly by adopting small perturbation in the initial values and initial condition¹¹⁾. The ensemble forecast of hydrographs is also a recent trend away from the conventional simple deterministic forecasts of hydrographs toward offering probabilistic forecasts, which include its prediction uncertainty.

As a step towards addressing the forecast accuracy improvement and ensemble hydrologic prediction, this study introduces a new attempt of ensemble rainfall-runoff simulations with a spatial random error field (SREF). Different to a

conventional ensemble simulation, which uses initial condition control to obtain statistical outcome, the SREF offers probable prediction error fields for the deterministic prediction results of a rainfall-forecasting model. The characteristics of SREF is based on an analyzed error structure of the spatial rainfall prediction, and it is simulated using the random field generation method¹²⁾.

The organization of this paper is as follow: section 2 illustrates the deterministic prediction of rainfall field using a radar extrapolation model. The prediction error fields are examined to extract the characteristics of prediction error structure, and the SREF is simulated based on the analyzed error structure. In section 3, probable extended prediction fields are formulated using the deterministic rainfall filed and the SREF. The extended prediction fields offer ensemble rainfall-runoff simulations with a distributed hydrological model. Discussions are on the accuracy of the rainfall prediction and its propagation to the discharge hydrograph through the distributed hydrological model in section 4.

2. SHORT-TERM RAINFALL FORECAST AND RANDOM ERROR FIELD

(1) Radar Image Extrapolation Model

The translation model by Shiiba *et al.*¹⁴⁾ simulates short-term radar rainfall predictions in this study. In the model, the horizontal rainfall intensity distribution, $z(x,y,t)$ with the spatial coordinate (x,y) at time t is modeled as:

$$\frac{\partial z}{\partial t} + u \frac{\partial z}{\partial x} + v \frac{\partial z}{\partial y} = w \quad (1)$$

$$\text{here, } u = \frac{dx}{dt}, \quad v = \frac{dy}{dt}, \quad w = \frac{dz}{dt}$$

where, u and v are advection velocity along x and y , and w is rainfall growth-decay rate. The vector u , v , and w are identified by the least square method using observed radar rainfall data and are updated on a real-time basis. Shiiba *et al.*¹⁴⁾ and Takasao *et al.*¹⁵⁾ give more details of the translation model and the function of the vector u , v , and w .

In this research, three consecutive observed rainfall fields, which have 3 km and 5 minutes of resolution, are used to determine u and v . When forecasting rainfall fields, the u and v are assumed spatially constant in each rainfall fields, but updated on a real-time basis. The growth decay rate w is always assumed to be zero.

Radar rainfall events used here are observed at the Miyama radar station located in the central part of Kinki District, Japan on 30th and 31st June 1993. Operated by the Japan Ministry of Land, Infrastructure, and Transportation, the radar data is converted to rainfall intensity covering a 240km×240km area with 3 km spatial resolution. A more

detailed description of the Miyama radar is given by Nakakita *et al.*⁶⁾.

(2) Prediction Error Structure Analysis

Tachikawa *et al.*¹⁶⁾ statistically analyzed the characteristics of absolute error and relative error defined as Equations 2 and 3.

$$E_{a,i} = R_{o,i} - R_{p,i} \quad (2)$$

$$E_{r,i} = (R_{o,i} - R_{p,i}) / R_{p,i} \quad (3)$$

The absolute prediction error $E_{a,i}$ on the grid i is calculated from the difference between predicted rainfalls $R_{p,i}$ and observed rainfalls $R_{o,i}$ on the grid, while the relative prediction error $E_{r,i}$ is the ratio of the absolute prediction error to its predicted rainfall. Tachikawa *et al.*¹⁶⁾ examined the time accumulated error values with variant spatial resolutions and found that the distribution of absolute and relative error are close to normal distribution and lognormal distribution respectively. Here we concentrate on the absolute prediction error $E_{a,i}$ and simulate the spatial random error fields in real-time basis.

Figure 1 shows spatially averaged rainfall intensities, on each time step, of the observed and predicted for 30, 60, and 120 min. In order to avoid an influence of non-rainfall area to the intensity calculation, the calculation only includes those grids in which rainfall was either predicted or observed within the study area. Note that there are overall delays of rainfall intensities as prediction time gets longer. Because the translation model used here only represents the movement of the rainfall bands without their growth or decay, the model assumes the same amount of current rainfall intensities lasts to the prediction target time. The slight differences of the intensities are only because of new included or excluded rainfall cells in the study area.

Two kinds of prediction performance measuring values are adopted: Correlation Coefficient (CC) of the observed and the predicted rainfall, and Critical Success Index (CSI). The CSI is given as:

$$CSI(\%) = \frac{X}{X + Y + Z} \times 100 \quad (4)$$

where, X is the number of correct forecasts rainfall cells (i.e. rainfall is observed and also predicted in the grid), Y is the number of misses (i.e. rainfall is observed, but not predicted), and Z is the number of false alarms (i.e. rainfall is predicted, but not observed). The CSI is especially appropriate as a summary measure of forecast skill for the case of extreme events, since the index value decreases with increasing both the number of misses and false alarms⁴⁾. A threshold rain-rate for the CSI in this research is 0.0 mm/hr within the radar range.

The CC values of the observation and prediction are under 0.5 even for the 30 min prediction and worse for the 60 min and 120 min predictions (see Fig. 2). On the other hand, the CSI values are

mostly over 50% and almost 100% for some times (see Fig. 3). The high CSI values are given by wide spread and slowly moving rainfall bands of the subject event. When a check of other events having variant rainfall area size is conducted, the CSI values are somewhat related to the size of rainfall area.

Spatial Correlation Coefficients (SCC) of the prediction error, which shows how much the error is spatially correlated to each other, for 60 min prediction is presented in Figure 4. The SCC is calculated by grouping each pairs of error values which apart one grid for 3 km, two grids for 6 km, etc. The SCC shows high values for short distance and decrease with long distance. The frequency distribution of the prediction error follows a normal distribution as the same pattern of Tachikawa *et al.*¹⁶⁾.

(3) Prediction Error Field Simulation

a) Outline of the Error Field Simulation

The main part of this study is to simulate possible error fields of the future prediction using the current prediction error structure, assuming temporal persistency of the error characteristics from the current time to the prediction target time. Figure 5 explains the procedure for possible error field simulation. The current characteristics of the prediction error are represented in the mean and standard deviation field of error (see Fig. 5(b) and (d)). The mean and standard deviation of the error on each grid formulate the mean and SD field, hereafter statistic fields, using certain duration of recent prediction error. The statistic fields convert Unit Random Error Field (UREF; Fig. 5(c)) to the target error fields (Fig. 5(a)), which are the aim of the error filed simulations. The UREF is a spatially correlated random field with the $N(0,1)$, and the spatial correlation for the UREF is determined by the current error fields. Generation of many UREF make it possible to get many target error fields. Finally, the deterministic prediction field obtained by the translation model is extended to many possible prediction fields by combining with the simulated prediction error fields.

b) Persistency of the Error Characteristics

To confirm the validity of the method, it was necessary to check how much the current statistic fields are close to the prediction error fields on the target time. Tests were conducted with three sets of statistic fields calculated using 10 min, 30 min, and 60 min of duration. After each statistic field was prepared from the current time and prediction target time, modified CSI values were calculated with each pair of the statistic field. The modified CSI was using the same form of Equation 4 only except a range concept. If the mean (or standard deviation)

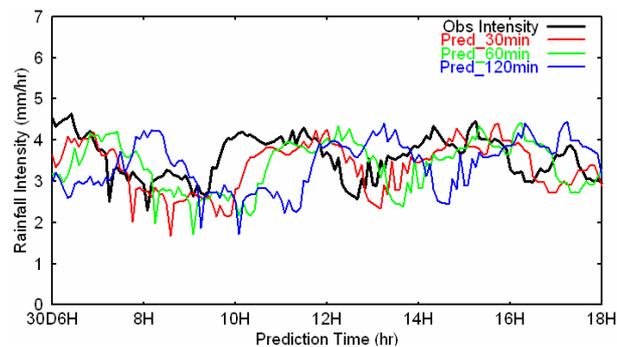


Fig.1 Average rainfall intensity of observation and prediction

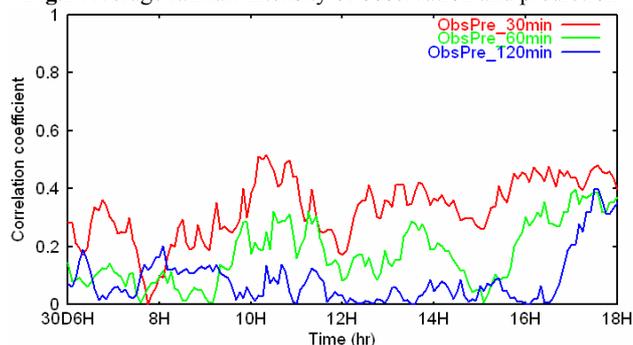


Fig.2 Correlation coefficients of observation and prediction

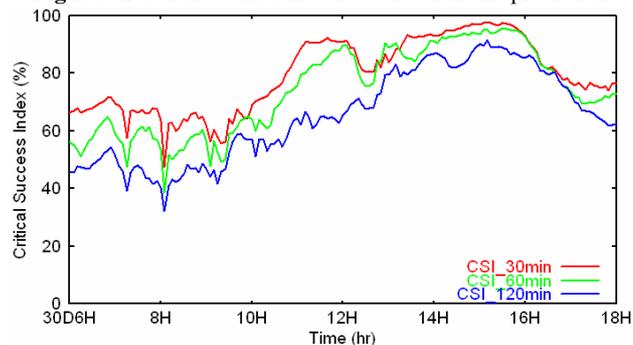


Fig.3 Critical success index (%) for observation and prediction

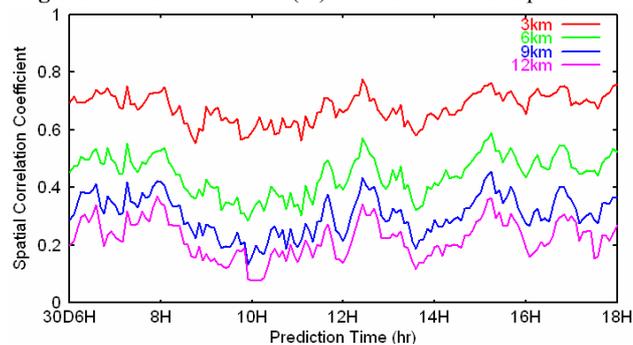


Fig.4 Spatial correlation coefficients of 60min prediction

value of future target time on a certain grid was within a certain range of the current mean (or SD) value on that grid, it was counted as correct forecast. If the values are out of range, it was counted as wrong forecast.

Figure 6 shows the modified CSI values with each pair of 10 min, 30 min, and 60 min statistic fields with the range of ± 3 mm. The CSI was over 60% for the mean field and over 80% for standard deviation field. With variant range, it was possible to check that the current statistic fields were good for simulating the prediction error on the target time

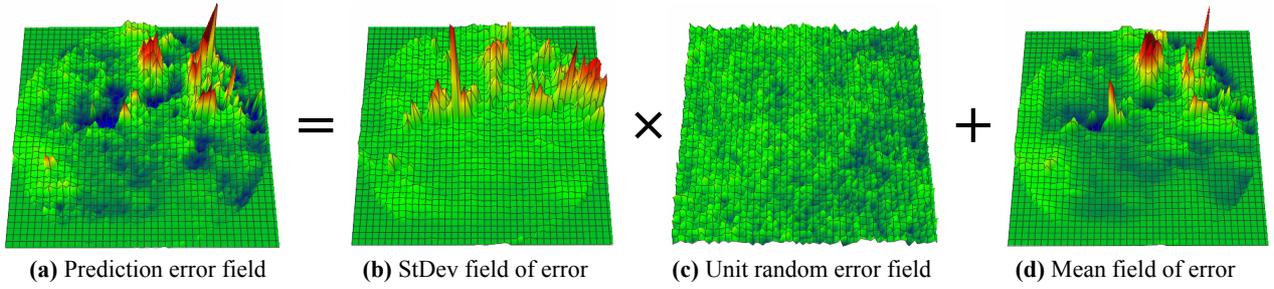


Fig.5 Simulation procedure of prediction error field using the statistic fields and UREF

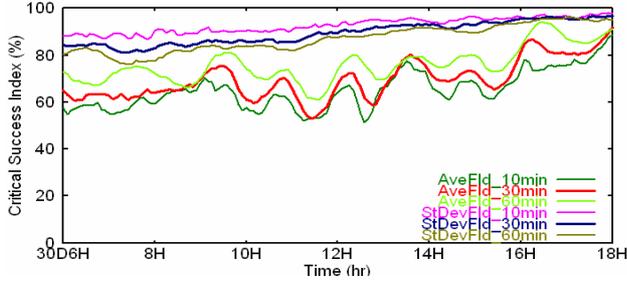


Fig.6 Modified CSI (%) for statistic fields of prediction error

The duration for calculating statistic fields does not give much effect to the CSI, thus the 30 min statistic fields were used.

The high CSI values may come from the characteristics of rainfall distribution in highly mountainous areas. Dynamics and microphysics of atmosphere are highly related to the topography for generation, growth, and decay of rainfall⁽⁶⁾. Because the translation model does not account this specific topography information, the prediction from the model contains inevitable errors, which relate to topographic information. The statistic fields made it possible to compromise the spatial and temporal pattern of these errors. Moreover, the information of the current prediction error can be updated in real-time basis with the statistic fields.

c) Prediction Error Field Simulation Algorithm

The UREF simulation uses the decomposition method of a matrix that includes spatial correlation characteristics of the prediction error. The matrix is decomposed approximately into its square root matrix with the matrix factorization technique^{(12),(13)}. Multiplying the square root matrix by a random vector $N(0,1)$ gives a non-conditional simulation of the UREF.

Davis⁽¹³⁾ proved a symmetric matrix \mathbf{B} that satisfies $\mathbf{K}=\mathbf{B}\mathbf{B}$ could be found when \mathbf{K} is symmetric and positive-definite. Considering the random vector \mathbf{Y} ; in this study, spatially correlated unit random error vector:

$$\mathbf{Y}=\mathbf{B}\mathbf{w}$$

where \mathbf{w} is uncorrelated random vector $N(0,1)$. The expected value of the matrix $\mathbf{Y}\mathbf{Y}^T$ ($n \times n$) is given by

$$\mathbf{E}[\mathbf{Y}\mathbf{Y}^T] = \mathbf{E}[\mathbf{B}\mathbf{w}\mathbf{w}^T\mathbf{B}^T] = \mathbf{B}\mathbf{E}[\mathbf{w}\mathbf{w}^T]\mathbf{B}^T$$

Because \mathbf{w} is a vector of independent random numbers, $\mathbf{E}[\mathbf{w}\mathbf{w}^T] = \mathbf{I}$, thus

$$\mathbf{E}[\mathbf{Y}\mathbf{Y}^T] = \mathbf{B}\mathbf{I}\mathbf{B}^T = \mathbf{K}$$

A covariance matrix \mathbf{K} is made out by using the spatial correlation coefficient, which is obtained from the absolute error $E_{a,i}$. The matrix \mathbf{K} is decomposed into the symmetric positive-definite matrix \mathbf{B} approximately by the *Chebyshev* polynomials^{(12),(16)}. A vector \mathbf{Y} (or UREF), which is a non-conditional simulation of spatially correlated random vector, can be generated by multiplying \mathbf{B} by an uncorrelated random vector \mathbf{w} . Figure 5(c) shows one example of UREF.

The statistic fields convert the UREF to the prediction error fields as Equation 5:

$$\begin{bmatrix} E_{s,1} \\ E_{s,2} \\ \vdots \\ E_{s,n} \end{bmatrix} = \begin{bmatrix} sd_1 & 0 & \cdots & 0 \\ 0 & sd_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & sd_n \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} + \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_n \end{bmatrix} \quad (5)$$

Here, the m_i and sd_i are the mean and standard deviation of the current prediction error on the grid i . The y_i is the unit random error of the vector \mathbf{Y} , and the $E_{s,i}$ is the simulated error for the prediction target time. Equation 5 is linear equation, thus the correlation structure of \mathbf{Y} , which is obtained from the $E_{a,i}$, is maintained in the $E_{s,i}$. The form of Equation 5 is identical with Figure 5. The total grid number of the Miyama radar image is 80×80 , thus the n of Equation 5 is 6400. This simple procedure allows each random error value y_i to have statistical characteristics of its grid. Fifty sets of prediction error fields for each time step were generated.

3. ENSEMBLE RAINFALL-RUNOFF SIMULATION

(1) Generation of Extended Prediction Fields

Deterministic prediction rainfall field from the translation model is extended to many possible prediction rainfall fields by combining with the simulated prediction error fields as:

$$R_{e,i} = R_{p,i} + E_{s,i} \quad (6)$$

where, $E_{s,i}$ is the simulated prediction error value on grid i , $R_{p,i}$ is the prediction from the translation model, and $R_{e,i}$ is the extended prediction. Because the simulated prediction error keeps the error statistics of the absolute prediction error ($E_{s,i} \approx E_{a,i}$),

the extended prediction can be close to the observed rainfall on the prediction target time as:

$$R_{e,i} \approx R_{o,i} = R_{p,i} + E_{a,i} \quad (7)$$

In other words, the properly simulated prediction error improves the forecast accuracy of the deterministic prediction.

Negative values could occur on the extended prediction fields, because the values on some points of the prediction error field could have larger minus value than the predicted rainfall value on that point. These minus rainfall set to zero, and the same amount of minus values compensate from the plus rainfall values for keeping the total rainfall amount as:

$$R'_{e,i} = \begin{cases} R_{e,i}(1+r) & (\text{if } R_{e,i} \geq 0.0) \\ = 0.0 & (\text{if } R_{e,i} < 0.0) \end{cases} \quad (8)$$

here, $r = \frac{\sum \text{Minus}R_{e,i}}{\sum \text{Plus}R_{e,i}}$

The value r in Equation 8 stands for the ratio of total minus rainfall amount on each extended prediction field to the total plus rainfall amount on it. The total amount of minus rainfall generally has 10% to 20% of the total plus rainfall amount, and the value r varies from -0.1 to -0.2.

The same error structure analysis described at 2.(2) is carried out for the extended prediction fields, and the result is compared with the original error structure of its corresponding time. Figure 7 shows the rainfall intensities of observation, deterministic prediction, and fifty sets of the extended predictions. The CC values in Figure 8 and CSI in Figure 9 show each value from the extended prediction fields. Average values (Fig. 10) and the standard deviation (Fig. 11) of prediction error, and SCC (Fig. 12) of the simulated prediction error show that the error characteristics were successfully kept through the whole procedure.

(2) Runoff Simulation Results and Discussion

The fifty sets of simulated extended prediction fields offer ensemble rainfall-runoff simulation with a distributed hydrological model. The model that was used here is the Yodo-river model using Ohymos system¹⁷⁾. The simulation is carried out for two different catchments located in the range of the Miyama radar; Ootori (156 km²) and Ieno (476 km²).

The fifty extended prediction fields for fifty-ensemble simulations were given at every 5 min during 6:00 and 18:00 on 30th Jun, 1993. Each extended prediction field was assigned to each runoff simulation independently. After 18:00 on 30th, the deterministic prediction was set to every fifty-ensemble simulations and continues the simulation for a while to observe the remaining effect of the extended prediction fields on the runoff.

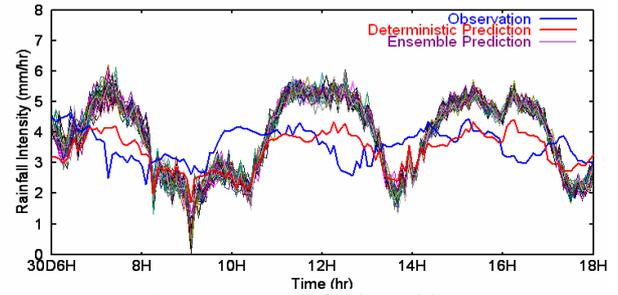


Fig.7 Average rainfall intensities

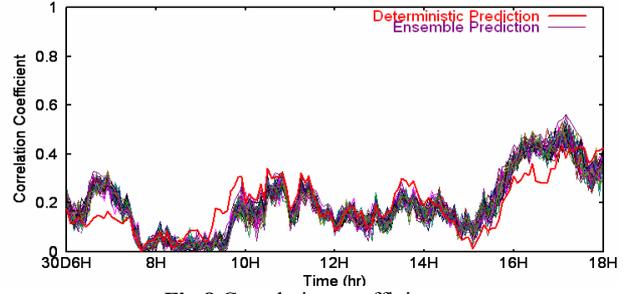


Fig.8 Correlation coefficients

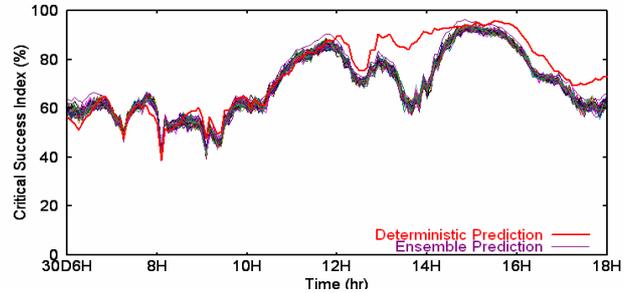


Fig.9 Critical success index (%)

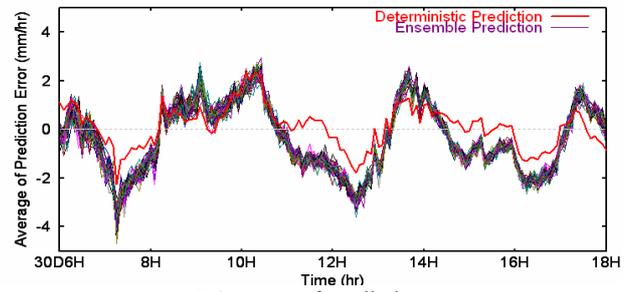


Fig.10 Average of prediction error

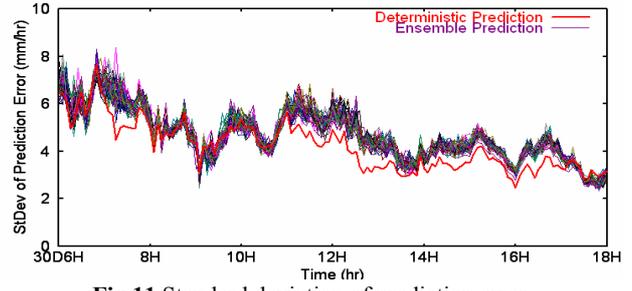


Fig.11 Standard deviation of prediction error

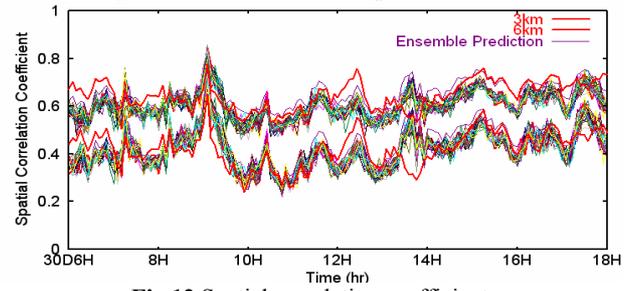


Fig.12 Spatial correlation coefficient

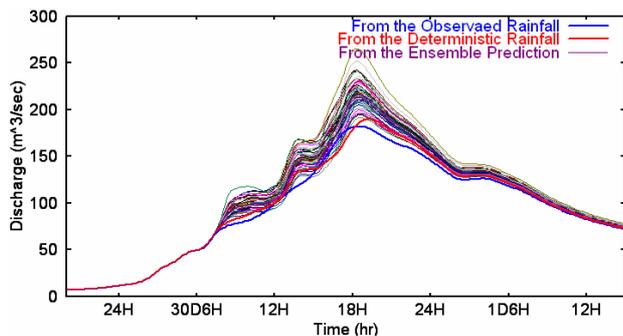


Fig.13 Rainfall-runoff Simulation Results for Ootori (156km²)

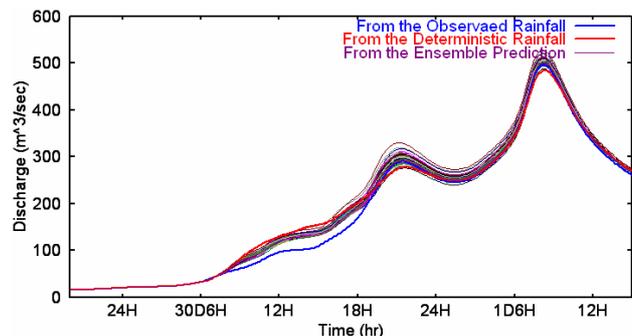


Fig.14 Rainfall-runoff Simulation Results for Ieno (476km²)

The discharges from the extended rainfall prediction were compared with the discharges from the original deterministic rainfall prediction to realize stochastic ensemble discharge simulations. In addition, we regarded the simulated discharge using the observed radar rainfall as the reference discharge to validate the improvement of the prediction accuracy by the extended prediction fields. Figures 13 and 14 show the ensemble runoff simulation results for two subject basins. The thick solid line stands for the discharge from the observed radar rainfall, and the other solid line is the discharge results from the original deterministic predicted rainfall for 60 min ahead at every time step. In the both simulation results, the discharges from the ensemble simulations (fifty dot lines) were not showing noticeable improvements. However, this results shows reliability of the prediction results with possible discharge bands.

It needs more study to validate carefully the reliability band of the ensemble simulation results using variant prediction lead-time. The event used in this research has rather low rainfall intensities compare to other events. Other rainfall event, which has higher rainfall intensities, for example typhoon event, need to be examined for checking the validity of the extended prediction field.

4. CONCLUSION

In this study, the translation model predicted radar rainfall field, and the prediction error structure was analyzed spatially and temporally. The random error fields were simulated using the error structure, and the extended rainfall prediction fields were generated. The extended prediction fields not only improved the accuracy of the original prediction by the translation model but also gave reliability with variant form of rainfall fields. The results of rainfall-runoff simulation with the distributed hydrologic model verified the hydrologic usefulness of the extended prediction.

REFERENCES

1) Fox, N. I. and Wilson, J. W.: Very short period quantitative precipitation forecasting, *Atmos. Science Letters*, Wiley Interscience, Vol.6, 7-11, 2005.

2) Collier, C. G. and Kzyzysztowicz, R.: Preface: Quantitative precipitation forecasting, *J. Hydrol.*, Vol.239, 1-2, 2000
 3) Bellon, A. and Austin G. L.: The accuracy of short-term radar rainfall forecasts, *J. Hydrol.*, Vol.70, 35-49, 1984
 4) Smith, K. T. and Austin, G. L.: Nowcasting precipitation-a proposal for a way forward, *J. Hydrol.*, Vol.239,34-45, 2000
 5) Georgakakos, K. P. and Bras, R. L.: A hydrologically useful station precipitation model: 1.Formulation, *Water Resour. Res.*, Vol. 20, No. 11, 1585-1596, 1984
 6) Nakakita E., Ikebuchi, S., Nakamura, T., Kanmuri, M., Okuda, M., Yamaji, A., and Takasao, T.: Short-term rainfall prediction method using a volume scanning radar and grid point value data from numerical weather prediction, *J. Geophys. Res.*, Vol.101, No. D21, 26181- 26197, 1996
 7) Golding, B. W.: Quantitative precipitation forecasting in the UK, *J. Hydrol.*, Vol.239, 286-305, 2000
 8) Ganguly, A. R. and Bras R. L.: Distributed quantitative precipitation forecasting using information from radar and numerical weather prediction models, *J. Hydrometeorol.*, Am. Meteor. Soc., Vol.4, 1168-1180, 2003
 9) Lovejoy, S. and Schertzer, D.: Scale invariance, symmetries fractals and stochastic simulations of the atmosphere, *Bull. Am. Meteor. Soc.*, Vol. 67, 21-32, 1986
 10) Grecu, M. and Krajewski, W. F.: A large-sample investigation of statistical procedures for radar-based short-term quantitative precipitation forecasting, *J. Hydrol.*, Vol.239, 69-84, 2000
 11) Du, J. and Mullen, S. L.: Shot-range ensemble forecasting of quantitative precipitation, *Monthly Weather Rev.*, Vol.125, 2427-2459, 1997
 12) Tachikawa, Y. and Shiiba, M.: Gaussaian and log normal random field generation based on the square root decomposition of a covariance matrix, *J. Hydraul. Coastal and Environm. Eng.*, JSCE, Vol.656(II-52), 39-46, 2000
 13) Davis, M. W.: Generating large stochastic simulation-The matrix polynomial approximation method, *Mathematical Geology*, Vol. 19, No. 2, 99-107, 1987
 14) Shiiba, M., Takasao, T. and Nakakita, E.: Investigation of short-term rainfall prediction method by a translation model, *Jpn. Conf. on Hydraul.*, 28th, 423-428., 1984
 15) Takasao, T., Shiiba, M. and Nakakita, H.: A real-time estimation of the accuracy of short-term rainfall prediction using radar, *Stochastic and Statistical Methods in Hydrol. and Environm. Eng.*, Vol.2, 339-351, 1994
 16) Tachikawa, Y., Komatsu, Y., Takara, K. and Shiiba, M.: Stochastic modeling of the error structure of real-time predicted rainfall and rainfall field generation, *Weather Radar Information and Distributed Hydrological Modelling* (ed. by Y. Tachikawa, B. E. Vieux, K. P. Georgakakos & E. Nakakita), IAHS Publ., No. 282, 66-7, 2003
 17) Sayama, T., Tachikawa, Y., Takara, K. and Ichikawa, Y.: Development of a distributed rainfall-runoff prediction system and assessment of the flood control ability of dams, *J. Hydraul. Coastal and Environm Eng.* JSCE, Vol. II-73, 2005 (in printing).

(Received September 30,2005)