Journal of Hydroscience and Hydraulic Engineering
Vol. 14, No. 2, November, 1996, 57-65

SNOWMELT RUNOFF FORECASTING BASED ON THE FUZZY REASONING METHOD WITH AN ERROR CORRECTION ALGORITHM

By

Shigemi HATTA

Tomakomai College of Technology, Tomakomai, Hokkaido, Japan

Mutsuhiro FUJITA

Dept. of Civil Eng., Hokkaido University, Sapporo, Hokkaido, Japan

and

Hiroshi SAGA

Dept. of Civil Eng., Hokkai-gakuen University, Sapporo, Japan

SYNOPSIS.

In this paper, the fuzzy reasoning method with prediction error correction is presented for snowmelt runoff prediction. The applicability of the model is assessed and compared by making a 3-hr lead time predictions of runoff in experimental basins. The results show that prediction using the proposed method has a higher degree of accuracy compared to snowmelt prediction using only the fuzzy reasoning method.

INTRODUCTION

In cold and snowy regions, snowmelt is the most important supply of water for agricultural and metropolitan areas, although under certain meteorological conditions, snowmelt runoff may cause flood disasters. Therefore, it is necessary to predict the timing and volume of snowmelt runoff not only for water use, but also for flood disaster protection. A wide range of snowmelt runoff models for short-term forecasting have been developed, based either on index methods (e.g. Rango and Katwijk(1)) or on an energy balance approach (e.g. Hatta et al.(2)). However, the accuracy of these methods are still not satisfactory, because the snowmelt rate, percolation of melt water and runoff process are extremely complex due to many vague factors involved.

Fuzzy sets were first introduced by Zadeh in 1965 (3), and have been applied to various fields, such as decision making and control. The characteristic of fuzzy theory which makes a distinction from classic mathematics is its operation on the membership function of variables with vagueness width instead of on the accurate values. This characteristic permits fuzzy theory to be a powerful tool for inaccurate data or an imprecise relationship between variables. The authors have proposed a new approach using the fuzzy reasoning method for the prediction of runoff from snowmelt (4). However, since the above method of snowmelt runoff prediction does not include algorithm to feed back prediction errors , an accumulative error in prediction develops especially for a long lead time prediction. Thus, in order to improve prediction accuracy, adjustments to the predicted hydrograph based on the level of prediction error are necessary. The purpose of this study is to develop the fuzzy reasoning method with a function of predicted error correction.

The basins used in this study were Kannonsawa river basin (drainage area 1.0km^2), Hyakumatsuzawa river basin (14.6km^2) and Misumai river basin (9.6km^2), which are all located in the Toyohira river drainage, near Sapporo (Fig. 1). In each basin, stream runoff was recorded at 1-hr interval in the 1992 and 1993 snowmelt period. This region has abundant snowfall, and the snowmelt accounts for a considerable percentage of the total water resources. The highest daily peak of snowmelt runoff usually occurs in early April. In this study, the correction algorithm for the prediction error was determined using data from the Kannonsawa river basin for the 1992 snowmelt season, and then in order to examine the correction algorithm, the proposed method was applied to 1993 flood data from Kannonsawa, Hyakumatsu and Misumai river basins.

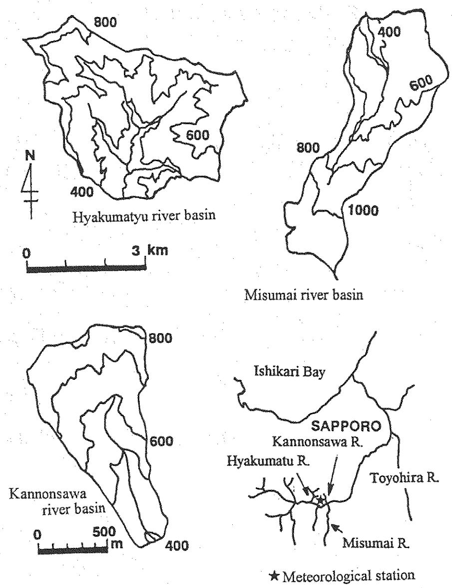

Fig.1 Outline of the study basins

FUZZY REASONING METHOD FOR SNOWMELT RUNOFF PREDICTION

Runoff Prediction Model

In this paper, we consider real-time runoff forecasting in which immediate runoff data for the current flood is constantly available. The general equation for the runoff system can be expressed as follows(Fujita and Zhu(5)):

$$O(t) = f\{O(t-1),\ldots,O(t-m), I(t-1),\ldots,I(t-n)\} \quad (1)$$

where t, O and I = time, outflow and inflow flux of the basin, respectively. Parameters m and n can be properly chosen by taking into account the basin characteristics. Since the observed data of hydrologic conditions unavoidably contain various types of errors, the function $f\{\ \}$ in Eq.(1) includes some uncertainty. Eq.(1) can be transformed into the following fuzzy conditional proposition after fuzzifying variables O and I using their membership functions M_O and M_I, respectively:

$$\begin{aligned} &if\ O(t-1)\ is\ M_{O(t-1)}\ and \ldots\ and\ O(t-m)\ is\ M_{O(t-m)}\ and\ I(t-1)\ is\ M_{I(t-1)} \\ &\qquad and \ldots\ and\ I(t-n)\ is\ M_{I(t-n)}\ then\ O(t)\ is\ M_{O(t)} \end{aligned} \quad (2)$$

This is a proposition with a compound premise. Employing the translating rule defined as the minimum of the fulfillment grades, the above proposition can be translated into the fuzzy relation P_t as follows:

$$P_t = M_{O(t-1)} \wedge \cdots \wedge M_{O(t-m)} \wedge M_{I(t-1)} \wedge \cdots \wedge M_{I(t-n)} \wedge M_{O(t)} \quad (3)$$

The fuzzy relation P_t varies with time. By current time t, it is possible to obtain a series of fuzzy relations $P_1, P_2, \ldots, P_t$, which may be combined to produce a general fuzzy relation Π_t through a conjunction operator, denoted by $\vee$;

$$\Pi_t = P_1 \vee P_2 \vee \cdots \vee P_t = \Pi_{t-1} \vee P_t \quad (4)$$

If flood data from the previous year is available, then Π_t obtained from the previous flood, denoted by Π_{et}, may be effectively and simply utilized by regarding it as an initial value:

$$\Pi_t = \Pi_{te} \vee P_1 \vee P_2 \vee \cdots \vee P_t = \Pi_{te} \vee \Pi_{t-1} \vee P_t \quad (5)$$

Fuzzy reasoning for the membership function $M_{O(t+1)}$ at the future time $(t+1)$ can be made based on Π_t as shown in Eq.(6). Here, the notation ' * ' is added to the prediction values.

$$M_{O*(t+1)} = \Pi_t \circ M_{O(t)} \circ \cdots \circ M_{O(t-m+1)} \circ M_{I(t)} \circ \cdots \circ M_{I(t-n+1)} \quad (6)$$

where, symbol ' o ' means max-min composition.

The center of gravity of the predicted membership function is adopted as a defuzzyfying procedure to obtain the value $O^*(t+1)$. Similarly, the membership function of a 2-hr lead time outflow can be inferred as follows:

$$M_{O*(t+2)} = \Pi_t \circ M_{O*(t+1)} \circ \cdots \circ M_{O(t-m+2)} \circ M_{I*(t+1)} \circ \cdots \circ M_{I(t-n+2)} \quad (7)$$

Snowmelt Runoff Prediction Model

The above method was applied to snowmelt runoff prediction. First, several types of runoff system equations were examined, as shown in Table-1 (Hatta et al. (4)).

The triangular membership function, as shown in Fig.2, was adopted to describe the membership function M_Q, $M_{\Delta Q}$, M_T and $M_{\Delta T}$. A training set obtained by the degree-hour model was used to create an initial fuzzy relation, Π_{te}. The applicability of each of the models using observed temperatures was assessed and compared by making 3-hr and 6-hr lead time forecasts of runoff during the snowmelt period in 1992.

The snow covered area decreases during snow melt season. The rapid decrease in the snow covered area is plays an important role in snowmelt runoff. Thus, in order to take into account of this effect, Eq.(8) was used instead of Eq.(4).

$$\Pi_t = P_{t-j} \vee \cdots \vee P_{t-2} \vee P_{t-1} \vee P_t \quad (8.a)$$

or

$$\Pi_t \wedge \neg P_j \rightarrow \Pi_t \quad (8.b)$$

where, symbol ' ¬ ' is a negation operator.

Eq.(8) means that the fuzzy relations $P_1, P_2, \ldots, P_j$ obtained in the remote part are not only useful but also have an adverse effect on the snowmelt runoff prediction. The following snowmelt runoff system equation was adopted for the Kannonsawa river basin with consideration to prediction error and computer time requirements:

$$\Delta Q = f\{Q(t-1), T(t-1), \Delta Q(t-1)\} \quad (9)$$

For forecasting, it is necessary to predict the temperature for any future period of time. In this study, temperatures were obtained from daily weather forecasts (Hatta et al. (6)).

The results of 1-hr, 3-hr and 6-hr lead-time predictions are shown in Fig.3. From this figure, it is clear that the prediction error gradually increases as the lead time increases from 1-hr to 6-hrs, and the phase difference between the predicted and measured discharges tends to correspond to the lead time. In the case of n-hr lead time prediction, computation of fuzzy reasoning is repeated n times, and the predicted membership function has a wide interval width. As a result, the center of gravity of the predicted membership function, which is an increment of discharge in this study, has a tendency to approach zero.

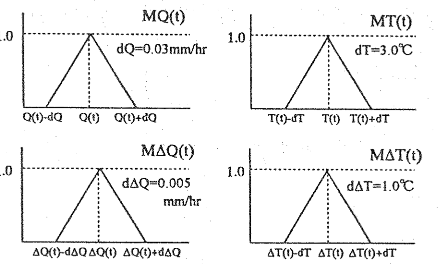

Fig.2 Triangular membership functions

Table-1. The various runoff system equations.
(Kannonsawa river basin, 1992)

		3-hr prediction	6-hr prediction
(1)	Q(t)=f {T(t-1)}	0.4754*	0.4718
(2)	Q(t)=f {ΔT(t-1)}	0.4986	0.4931
(3)	Q(t)=f {T(t-1),T(t-2)}	0.4653	0.4619
(4)	Q(t)=f {ΔT(t-1),ΔT(t-1)}	0.494	0.4886
(5)	Q(t)=f {T(t-1),ΔT(t-1)}	0.4596	0.4561
(6)	ΔQ(t)=f {Q(t-1),T(t-1)}	0.0446	0.065
(7)	ΔQ(t)=f {Q(t-1),ΔT(t-1)}	0.1274	0.2413
(8)	ΔQ(t)=f {ΔQ(t-1),T(t-1)}	0.0727	0.132
(9)	ΔQ(t)=f {ΔQ(t-1),ΔT(t-1)}	0.0719	0.1271
(10)	ΔQ(t)=f {Q(t-1),T(t-1),ΔT(t-1)}	0.0402	0.051
(11)	ΔQ(t)=f {ΔQ(t-1),T(t-1),ΔT(t-1)}	0.0413	0.0597

* lead time : 3-hr, relative prediction error : 0.4754

T(t): air temperature

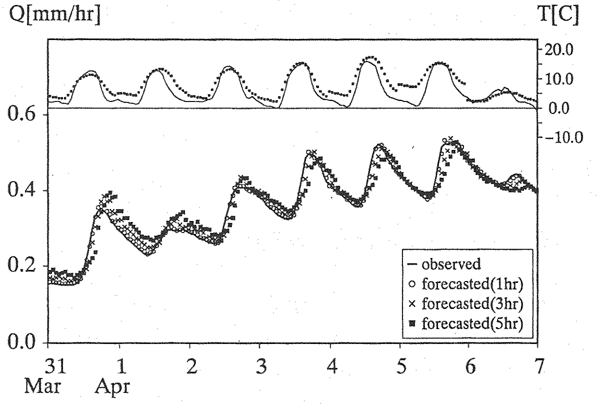

Fig.3 Prediction results by the fuzzy reasoning method.
(Kannonsawa river basin, 1992)
(··· predicted temperature based on daily weather forecast[6])

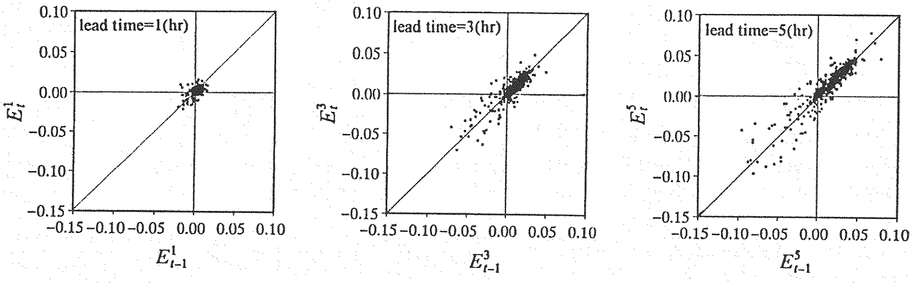

Fig.4 Relation between prediction error at present time and 1-hr before.

ERROR CORRECTION ALGORITHM

In order to make a long lead time forecast and to improve prediction accuracy, an updating procedure of the forecasted hydrograph is necessary. Our aim is to develop new methods in which the absolute deviation $|Q(t+i) - \mathbf{Q}^i(t)|$ between the actual discharge Q and forecasted values $\mathbf{Q}$ is as small as possible for each lead time i. $\mathbf{Q}^i(t)$ denotes the forecast made at origin t of discharge at some future time $t+i$. The deviation for a lead-time of i hours is:

$$E^i_t = \mathbf{Q}^i(t) - Q(t+i) \qquad (10)$$

In this paper, the authors propose two new methods to adjust the predicted hydrograph based on the level of prediction error:

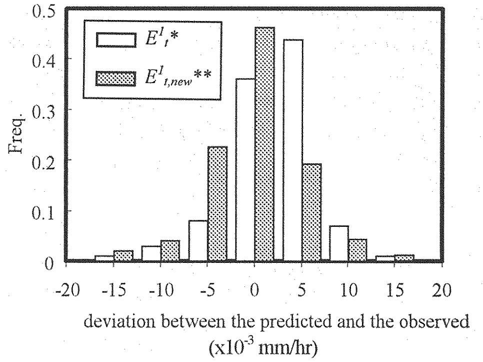

Fig.5 The frequency distribution of the 1-hr lead time prediction error deviation between the predicted and observed
(*$E^i_t = \mathbf{Q}^i(t) - Q(t+i)$, **$E^i_{t,new} = \mathbf{Q}^i(t)_{new} - Q(t+i)$)

Method A: Correction of predicted discharge based on 1-hr previous forecasting results

If it is possible to estimate E^i_t in Eq.(10), an adjusted prediction $\mathbf{Q}^i(t)_{new}$ can be written by

$$\mathbf{Q}^i(t)_{new} = \mathbf{Q}^i(t) - E^i_t \qquad (11)$$

Fig.4 shows the relationship between E^i_t and E^i_{t-1} for 1, 3 and 5-hour lead time predictions. Fig.4 leads to the following equation:

$$E^i_t = E^i_{t-1} \qquad (12)$$

For 1-hour lead time prediction, Eq.(11) can be rewritten as follows:

$$\mathbf{Q}^1(t)_{new} = \mathbf{Q}^1(t) - E^1_{t-1} = (\mathbf{Q}^1(t) - \mathbf{Q}^1(t-1)) + Q(t) \qquad (13)$$

The new deviation between the adjusted prediction and the actual discharge is:

$$E^i_{t,new} = \mathbf{Q}^i(t)_{new} - Q(t+i) \qquad (14)$$

Fig.5 indicates a frequency distribution for both $E^1_{t,new}$ and E^1_t. Fig.5 shows the effect of this method.

Eq(15) can be derived by generalizing Eq.(13).

$$\mathbf{Q}^i(t)_{new} = \mathbf{Q}^i(t) - E^i_{t-1} = (\mathbf{Q}^i(t) - \mathbf{Q}^i(t-1)) + Q(t+i-1) \qquad (15)$$

However, Eq.(15) includes an unknown term $Q(t+i-1)$, for which it is impossible to determine the value at time t. Therefore the following assumption is made:

$$Q(t+i-1) = \mathbf{Q}^{i-1}(t)_{new} \qquad (16)$$

Eq.(15) is rewritten as:

$$\mathbf{Q}^i(t)_{new} = (\mathbf{Q}^i(t) - \mathbf{Q}^i(t-1)) + \mathbf{Q}^{i-1}(t)_{new} \qquad (17)$$

$\mathbf{Q}^i(t)_{new}$ defined by Eq.(17) tends to fluctuate especially for longer lead time prediction. To solve this problem, a new $\mathbf{Q}^i(t)_{new}$ is redefined by Eq.(18).

$$\mathbf{Q}^i(t)_{new} = \frac{\mathbf{Q}^i(t-1)_{new} + \mathbf{Q}^i(t)_{new} + \mathbf{Q}^i(t+1)_{new}}{3} \qquad (18)$$

Fig.6 shows the adjusted prediction for a 3-hour lead time. Although the predicted peak discharge tends to be excessive, the predicted value is effectively corrected as a whole.

Method B: Correction of predicted discharge based on 24-hr previous forecasting results

Since snowmelt runoff is a periodic phenomenon which occurs in 24-hr cycles, it is possible to adjust the predicted runoff based on its periodicity. Basically, method B replaces E^i_{t-1} in Eq.(12) by E^i_{t-24}.

Fig.7 shows the relationship between time E^i_{t-1} and E^i_{t-24} for 1, 3 and 5-hour lead times. E^i_t is almost in proposition to E^i_{t-24}. Eq(19) can be assumed.

$$E^i_t = E^i_{t-24} \qquad (19)$$

Therefore, Eq.(15) is replaced by

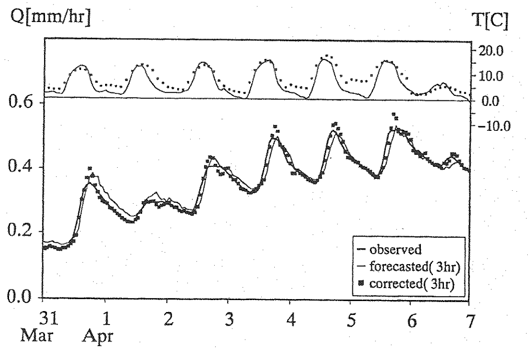

Fig.6 Corrected hydrograph for 3-hr lead time prediction based on prediction error at 1-hr before.

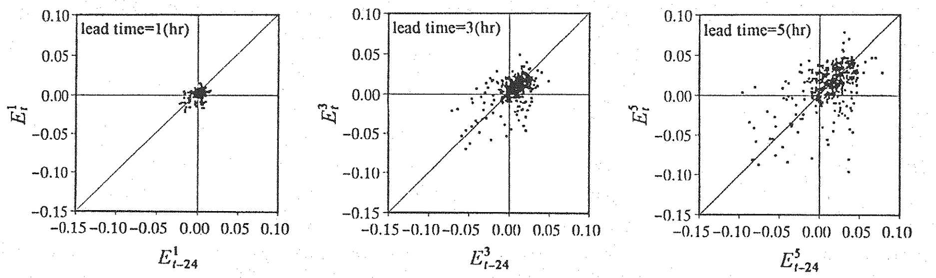

Fig.7 Relation between prediction error at present time and 24-hr before.

$$\mathbf{Q}^i(t)_{new} = (\mathbf{Q}^i(t) - \mathbf{Q}^i(t-24)) + Q(t-24+i) \quad (i<24) \tag{20}$$

Eq.(18) is also used to prevent fluctuations in the adjusted prediction.

Fig.8 shows the results for a 3-hour lead time predictions. Although the errors in flood peak and its timing are small, the corrected hydrograph is influenced by the hydrograph of the previous day.

Prediction accuracy was improved by using the above two methods. For practical applications, however, both methods have advantages and disadvantages. In Method A, the adjusted prediction agrees with the observed discharge on the rising limbs of the hydrograph, and the prediction accuracy of the peak discharge shows a tendency to decrease. On the other hand, Method B can forecast the peak discharge more accurately than Method A.

The advantages of both methods should be used. One simple way of doing this is to use a weight sum of both predictions.

$$\mathbf{Q}_c = k_1 \cdot \mathbf{Q}_1 + k_2 \cdot \mathbf{Q}_2 \tag{21}$$

where, $\mathbf{Q}_c$ = corrected values; $\mathbf{Q}_1$ = corrected values based on Method A; $\mathbf{Q}_2$ = corrected values based on Method B; and k_1, k_2 = weight coefficients. The weight coefficients in Eq.(17) are determined by the error of each last correction value. Thus, the weight coefficients are calculated as follows.

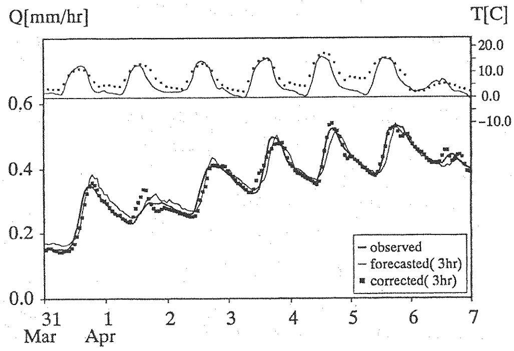

Fig.8 Corrected hydrograph for 3-hr lead time prediction based on prediction error at 24-hr before.

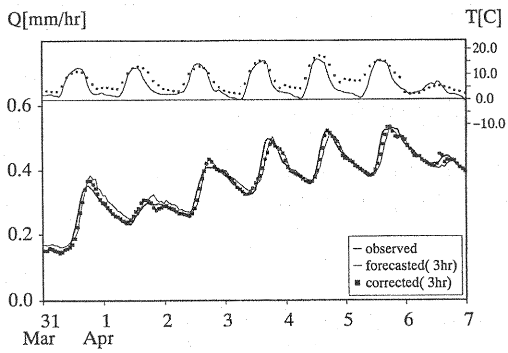

Fig.9 Corrected hydrograph for 3-hr lead time prediction using Eq.(21).

$$\varepsilon_1(t) = \sum \frac{1}{|\mathbf{Q}_1(t-1) - Q(t-1)|}; \qquad \varepsilon_2(t) = \sum \frac{1}{|\mathbf{Q}_2(t-1) - Q(t-1)|} \tag{22}$$

$$k_1(t) = \frac{\varepsilon_1(t)}{\varepsilon_1(t) + \varepsilon_2(t)} \tag{23}$$

$$k_2(t) = 1 - k_1(t) \tag{24}$$

Fig.9 shows the corrected results for a 3-hr lead time prediction using the above equations. The results show that this method can improve prediction accuracy.

APPLICATION RESULTS

In order to investigate the accuracy of the proposed correction algorithm, the proposed method Eq.(21) was applied to flood data from Kannonsawa, Hyakumatsu and Misumai river basins during the 1993 snowmelt season. Eq.(10) was used as the runoff system equation for each basin, and the runoff data for the 1992 snowmelt period was used to produce the initial fuzzy relation, Π_{te}. Figs.10 ,11 and 12 show a comparison of corrected prediction for a 3-hr lead time and the forecasted value based on only the fuzzy reasoning method and observed discharge. The results show good agreement between the corrected 3-hr lead time prediction value and the observed hydrograph.

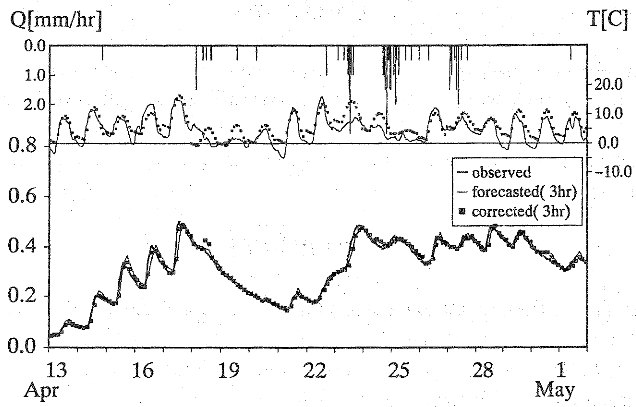

Fig.-10 Prediction results by the fuzzy reasoning method with error correction algorithm.
(Kannonsawa river basin, 1993)

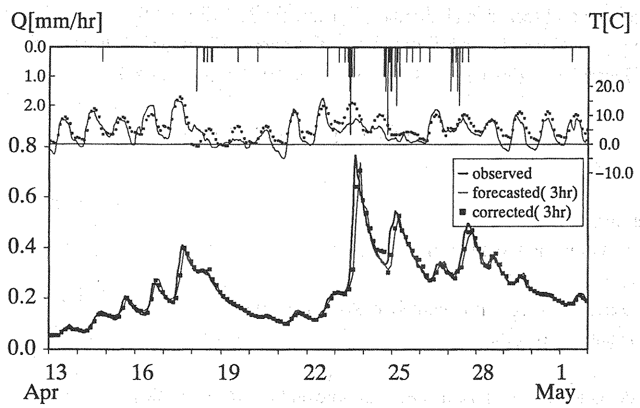

Fig.-11 Prediction results by the fuzzy reasoning method with error correction algorithm
(Misumai river basin, 1993)

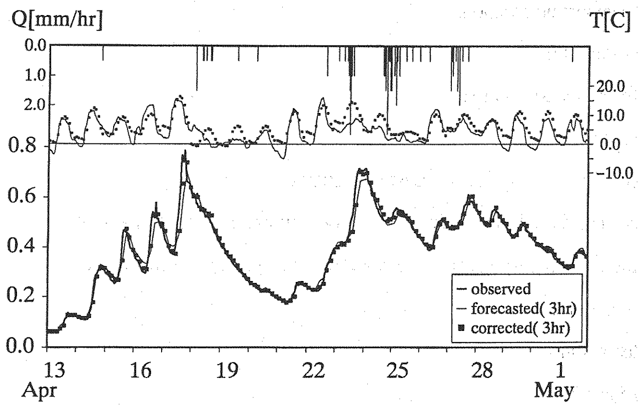

Fig.-12 Prediction results by the fuzzy reasoning method with error correction algorithm
(Hyakumatsu river basin, 1993)

CONCLUSIONS

The correction algorithm of predicted discharge based on the previous prediction error was reasonably successful when applied to data from the study basins. At present, the proposed method is effective for predictions within a 3-hour lead time.

In this study, small-scale study basins were used. From a practical standpoint, it is necessary to develop this method for a larger catchment such as a dam basin.

REFERENCES

1. Rango, A. and Katwijk,V. : Development and testing of a snowmelt-runoff forecasting technique, Water Resour. Bul., vol.26, pp.135-144, 1990.
2. Hatta, S., Nishimura, T., Saga, H. and Fujita, M. : Study on snowmelt runoff prediction using weekly weather forecast, Environment International, vol.21, No.5, pp.501-507, 1995.
3. Zadeh, L.A. : Fuzzy sets, Information and control, vol.8, pp.338-353, 1965.
4. Hatta, S., Fujita, M. and Saga, H. : Snowmelt runoff prediction using fuzzy reasoning method based on only air temperature, Journal of Japan Society of Hydrology & Water Resources, vol.7, No.4, pp.296-304, 1994 (in Japanese)
5. Fujita, M.and M. -L. Zhu : An application of fuzzy theory to runoff prediction, Proceedings of Sixth IAHR International Symposium on Stochastic Hydraulics, Taiwan, pp.727-734, 1992.
6. Hatta, S., Nishimura, T., Saga, H. and Fujita, M. : The snowmelt runoff analysis in the Kannonsawa river basin, Proceedings of JSCE Hokkaido Branch, vol.49, pp.515-518, 1993. (in Japanese)

APPENDIX - NOTATION

The following symbols are used in this paper:

$\wedge, \vee$	= minimum and maximum operations;
o	= composition operator, namely max-min operator;
$\neg$	= negation operator;
E	= deviation between the actual discharge and forecasted values;
i	= lead time;
I	= inflow flux;
k_1, k_2	= weight coefficients;
M	= membership function;
O	= outflow flux;
P_t	= fuzzy relation at time t;
Q	= runoff;
$\mathbf{Q}c$	= corrected value of the forecasted discharge;
$\mathbf{Q}_1$	= corrected predicted discharge based on 1-hr previous prediction results ;
$\mathbf{Q}_2$	= corrected predicted discharge based on 24-hr previous prediction results;
t	= present time;
T	= temperature;
ΔQ	= runoff increment;

ΔT = temperature increment;

Π_t = fuzzy relation by the current time t obtained; and

Π_{te} = fuzzy relation obtained from the previous flood events.

(Received January 8, 1996; revised September 13, 1996)