Uncertainty assessment for short-term flood forecasts in Central Vietnam

D. H. Nam, K. Udo & A. Mano
Disaster Control Research Center, Tohoku University, Japan

Abstract

Accurate flood forecasts with greater lead-times are very important in development of flood mitigation measures, especially in short response catchments. The flood forecasts based on numerical weather prediction (NWP) and runoff models have demonstrated its breakthrough to extend the forecast lead-time over traditional flood forecast methods, for instance, those are based on rainfall information from rain-gages. However, given the imperfectness either in the specification of initial states or in the formulation of NWP models, rainfall prediction for example, the driving factor for flood forecast, has been recognised as a major source of uncertainty in the generation of river flow. This paper presents the uncertainty assessment for a short-term flood forecast model that is coupled by the short-range global NWP model, 0.5 degree spatial resolution, with the distributed rainfall runoff model, for a large sized basin (Thu Bon River, 3,150km^2) located in Central Vietnam. To reduce uncertainty of runoff forecasts by means of increasing the rainfall prediction skill, first the model output statistic technique has been employed to downscale the large scale prediction forecasts directly derived from the NWP model output to the basin scale by using the artificial neural network with the back-propagation method. Skill scores of the downscaled precipitation are investigated with increasing lead-time and compared to those obtained using the large scale precipitation forecasts. Uncertainties of runoff prediction are assessed by quantifying the relative error of forecasts and estimates of confidence interval for the mean error. Results show that larger uncertainties along with the forecast lead-times are observed; however, the model is able to predict reliable river flows with lead-time of the order of 6-18 hours. This demonstrates great benefits in flood forecasting practices for many developing countries where ground weather observation is scarce and access to high resolution NWP models is limited.
Keywords: numerical weather prediction, flood forecast, uncertainty.

WIT Transactions on Ecology and the Environment, Vol 146, © 2011 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/RM110111

1 Introduction

Flood frequency and intensification are expected to alter considerably as a result of climate change. It is obviously that increase of extreme precipitation events, which usually cause severe flood disasters, has been observed in almost tropical regions. This trend is projected to be more severe in the future. Implementation of structure measures for flood risk reduction like the construction of reservoirs, dams, dykes and so on that has demonstrated its capability to reduce the impact of floods. This approach often requires huge investments and engineering technologies; however, not all floods have been entirely prevented given the change in flood intensification (Thielen *et al.* [1]). An operational flood forecast model is apparently one of the most essential tools for flood damage reduction. This aims to provide accurate and timely information of imminent floods to people at risk as well as flood control institutions for proper implementation of preparedness and mitigation plans (Micha *et al.* [2]). The benefit of forecast lead-time and experiences in response to the previous forecasts has significantly reduced loss of lives and damages to properties (Wind *et al.* [3]). With respect to the extension of the forecast lead-time, recently, the use of numerical weather prediction (NWP) in flood forecasting has been revealed as a promising alternative. However, flood forecast models are subject to uncertainty in general, especially those are based on rainfall prediction obtained directly from the NWP model output. The errors in rainfall, even predicted by high resolution models, have been observed as the most significant contribution to the total model uncertainty (Kardhana and Mano [4], Xuan *et al.* [5]).

Due to inherent errors in flood forecast models, quantification of the model uncertainties provides insights of imminent floods so that these uncertainties should be included throughout the decision making process for flood damage reduction. The objective of this paper is to assess uncertainties of the short-term flood forecast model that is coupled by the short-range global NWP model with the distributed rainfall runoff model. Large scale precipitation forecasts obtained directly from the NWP model output, hereinafter simply referred as direct model output (DMO), were downscaled to the basin scale using artificial neural network (ANN). Downscaled precipitation was then input to the super tank model for flood prediction. Skill scores of precipitation forecast with increasing forecast lead-time were explored for the downscaled precipitation as well as DMO. Uncertainties of runoff prediction were then assessed by quantifying relative errors of forecasts and estimates of confidence interval for the mean error. A river basin in Central Vietnam where floods are considered the most dangerous calamity to human lives and properties was selected as a case study.

2 Study area and methodology

2.1 Study area

It has been highlighted recently regarding the climate change impacts by the Intergovernmental Panel on Climate Change that tropical regions are likely to be

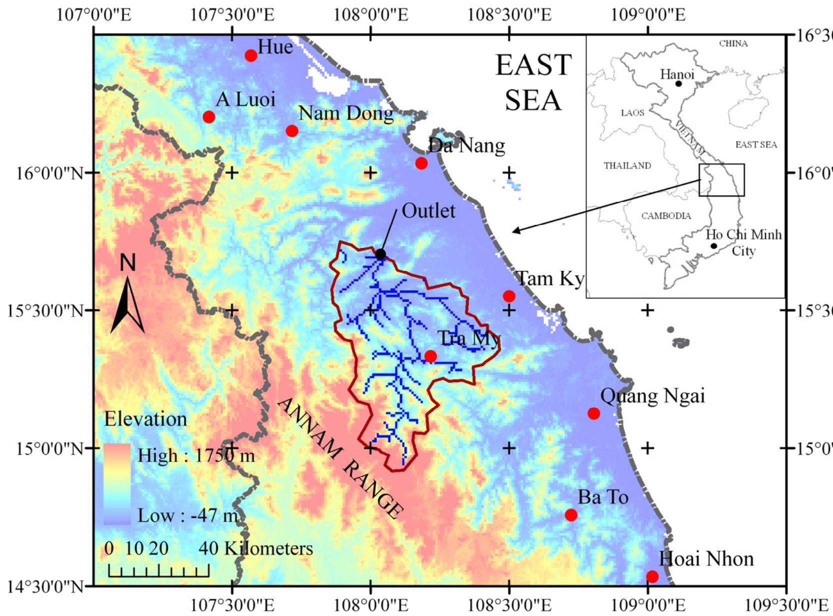

Figure 1: Locations of study area and weather observation points (●) and NWP model output on grid-point-value basis (+).

most influenced. This study selected the Central Vietnam as a case study. The study area located to the east of the Indochina Peninsula, is often hit by Pacific Ocean tropical storms associated with intense rainfall that usually causes large scale flooding in wet seasons, from September to December, almost every year. These storms are typically widespread orographic rainfall that generated on the windward of the Annam Range, known as the border between Vietnam and Laos (Figure 1). These orographic rainfalls were basically formulated through collaboration of the cold surges from northern continents and the tropical depressions from the Pacific Ocean (Yokoi and Matsumoto [6]). Given the topographical features, rivers in this region are generally very short and steep; therefore, catchment response is rapid, leaving a very short lead-time for people at risk to implement flood mitigation measures.

The catchment selected in this study is the upper reach of Thu Bon River, as seen in Figure 1, with the catchment area of 3,150 km^2, slightly larger than a grid cell size of the global NWP output of about 2,500km^2. The selection of a large size basin tends to reduce not only the effect of basin scale on runoff generation but also the effect of the coarse spatial resolution of the global NWP model.

2.2 Meteorological data

High resolution NWP models such as Mesoscale Model and Limited Area Model usually exhibit better forecast skills; however, its forecast domain and lead-time are often limited within a country boundary and short-range forecasts respectively. On the other hand, the coarser resolution NWP models tend to provide greater forecast lead-times and forecast domain, for instance, the global NWP models that are operational at European Centre for Medium-Range Weather Forecasts and Japan Meteorological Agency (JMA).

In the present study, given the unavailability of operational high resolution modes, atmospheric variables derived from the deterministic global NWP model output, issued by JMA, with spatial resolution of 0.5^{O} and 60 vertical layers were used. This NWP model provides forecasts for every 6-hour interval in the first 84-hour and every 12-hour interval for the next 132-hour. Forecasts are issued 4 times per day, at 00, 06, 12, and 18UTC.

In addition, this study addressed a convention that precipitation obtained from rain-gages was considered as reference rainfall (truth) for the comparison. Inverse distance weighting method was used to downscale precipitation and related atmospheric parameters either from a point representation (rain-gage) or grid-point-value representation (NWP) to the area average basis. Because the global NWP with 0.5^{O} spatial resolution has been effective since late 2007, this analysis was based on archived data for the wet seasons in 2008 and 2009.

2.3 Downscaling large scale precipitation

At global scale, even though NWP models have recently showed remarkable improvements in terms of spatial resolution, approximately 25-50km, this spatial resolution are still far away from requirement for hydrological simulations that usually require much finer resolutions, of the order of hundreds meter for small catchments to a couple of kilometers for large basins.

Model output statistic has been well known for a long history as a statistical downscaling tool for operational weather forecast. This approach is fundamentally based upon on the formulation of either linear or nonlinear relationships between large scale atmospheric variables and local or single-site scale variables. These relationships are then used to correct the outputs of the NWP models. In terms of learning skill, the non-linear regression model, such as ANN, which has demonstrated better skills than other linear regression models (Dawson and Wilby [7]). As a result, the present study utilized the most simple and widely used artificial neural network architecture, the feed-forward multilayer perceptron, hereinafter simply referred as ANN, for the learning process. Detail description of the network configuration, the selection of optimal predictors and learning methods were presented in [8].

2.4 Hydrological model

Downscaled precipitation is then input to the super tank model that was introduced by Kardhana *et al.* [9] for runoff prediction. The super tank model has been well-known as its nearly calibration-free parameters. A brief description of the model structure is presented here. The basin is divided into sub-basins that are represented by channel grids. The sub-basin represents a drainage area where precipitation throughfall reaches ground surface, partially infiltrates into ground, and the remaining turns into direct runoff, then lumped into channels.

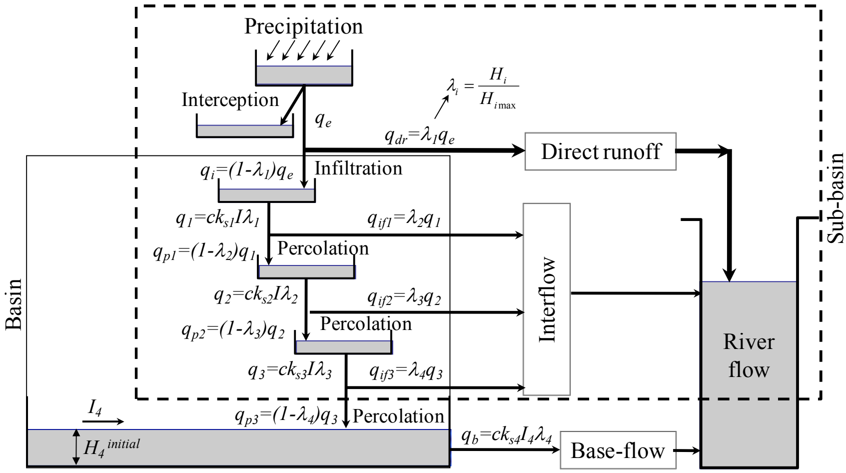

Figure 2: Structure of the super tank model and governing equations for determinations of infiltration rate, interflow, and base-flow.

Determination of infiltration rate and direct runoff are controlled by saturated hydraulic conductivity of the top soil layer. The lower efflux from the i^{th} tank in unit area is modelled by the modified Darcy Law as expressed by equations in Figure 2. Here k_s is saturated hydraulic conductivity. c is a modification coefficient on k_s that represents the assumed deviation between actual and estimation interflow from the Darcy Law due to the nature of soil structure. I is the slope of sub-basin. λ is saturation of a tank with water depth H and tank depth H_{max}, equal to soil layer thickness. Flood routing in streams and surface runoff uses one dimensional kinematics wave approximation scheme.

However, the super tank model merely includes the interflows from 3 tanks that represent the uppermost soil layers in the generation of river flow, as illustrated within the dashed line in Figure 2, while the contribution of ground water that controls the base-flow is omitted. The omission of the ground water is interpreted as the rapid saturation of the top soil layers because of heavy rainfall that throughfall mostly turns into direct runoff. In order to control the base flow, a lump tank that represents the contribution of groundwater of the whole basin was introduced for the complete process of flow generation in channels. Given the lack of hydro-geological information of the groundwater tank, calibration for lump parameters, the gradient (I_4) and initial storage ($H_4^{initial}$) of the groundwater tank in Figure 2, applied for all grid cells were required.

2.5 Model uncertainty

The overall uncertainty of a flood prediction model is usually defined as a result of errors in rainfall prediction, the incompleteness of model formulation and the improper assessment of model parameters (Maskey [10]). The propagation of uncertainty is understood as from the rainfall prediction through the runoff generation, then, to flood discharge in channels. In general, sources of uncertainty should be individually quantified and treated. In respect of rainfall forecast, however, substantial attempts have been made to enhance the accuracy of NWP models through the increase of model resolution as well as the better approximation of model initial condition, it has been recognised that the errors in rainfall prediction significantly influence the uncertainty of runoff prediction due to the intrinsic bias in the NWP models. Moreover, the longer forecast lead-time the larger error in rainfall prediction is expected.

As a result, in the context of this study errors in precipitation forecast and uncertainties in runoff prediction, considered as the total model uncertainty, are assessed against the forecast lead-times. First, the forecast accuracy of the downscaled precipitation is evaluated using skill score method and the similarity procedure for the DMO is also conducted for the comparison. Second, errors in peak discharge prediction and estimates of confidence interval for the mean error of runoff prediction are analysed. It should be noticed that the short-range global NWP model currently produces the forecasts up to 216-hour lead-time; however, the assessment of model uncertainty just focused on a short-term forecast so that forecast lead-time up to 48-hour was targeted.

2.5.1 Forecast skill score

Correlation coefficient that measures the degree of linear relationship between forecast and observation rainfall is usually used for forecast skill assessment. However, the correlation coefficient tends to represent the potential skill rather than the actual skill that is described as skill scores (Murphy [11]). The skill score measures the accuracy of a forecast relative to a reference forecast which are usually referred as the climatology forecast or the persistence forecast. In which, mean of squared differences between the forecast and the observation, or so-called mean square error (MSE), has been considered as a basic measure to evaluate the accuracy of the forecast. MSE is expressed in eqn. (1).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(f_i - o_i) \qquad (1)$$

where n is the number of forecasts, f_i and o_i denote the i^{th} forecast and observation of rainfall respectively.

This study utilised climatology as the reference forecast to compute the MSE skill scores. According to Murphy [11], climatological forecasts were defined based on observation from either a historical period or during the experimental period; it is so-called external climatological forecasts and internal climatological forecasts respectively. In addition, it also classifies the reference forecast of either a single constant forecast (single-valued) for all forecasting occasions or different forecasts (multiple-valued) for different occasions. Given the limitation of observation and prediction data availability, the most simplified form of MSE skill score (*SS*) based on the single-valued internal climatology was employed. It is expressed in eqn. (2).

$$SS = 1 - \frac{MSE(f,o)}{MSE(\mu,o)} \qquad (2)$$

where μ is the mean of climatology forecast.

2.5.2 Errors in runoff forecast and confidence interval for its mean

Perhaps the estimation of uncertainties of runoff forecast is most intended by decision-makers. So that they can include these uncertainties in the implementation of effective measures for flood risk reduction such as early flood warning, reservoir operation for flood control, and at last issuing evacuations if it is needed. In fact, the earlier decision is made the less flood impact is expected. As noted previously, however, the uncertainties are likely to be larger along with the forecast lead-time. It is necessary to evaluate the model accuracy against forecast lead-times for specific flood mitigation purposes. In the context of this study, forecasting for flood warning is targeted. It means timely and accurate predictions of imminent floods, especially peak discharges, are required. In addition, good estimations of total volume in early stages are useful for reservoir operation. Therefore, the relative errors of forecasted runoff and total runoff volume with various forecast lead-times were considered. It takes a form as expressed in eqn. (3).

$$\eta = \frac{\left| Q_{for} - Q_{obs} \right|}{Q_{obs}} \tag{3}$$

where Q_{obs} is observed river flow

Q_{for} is forecasted river flow.

A confidence interval with a certain confidence level in the estimate of a mean specifies a range of values within which the mean probably lies. In statistical analysis, the runoff, in general, can be characterised by the standard normal distribution. However, in the present study, the sample size was relatively small, of the order of 10 to 30, which depends on the durations of approximate 3 days for a single flood event and less than 10 days for continuous flood events. In this case, the T-distribution, as well known as normal distribution when the degrees of freedom get large, was applied to compute the confidence interval for the mean. Eqn. (4) shows the confidence interval for the mean of runoff forecast errors with (1-α) level of confidence for (N-1) degrees of freedom (Dingman [12]).

$$\Pr\left\{ m_X - \frac{t_{N-1,\alpha/2} s_X}{N^{1/2}} \le \mu_X \le m_X + \frac{t_{N-1,\alpha/2} s_X}{N^{1/2}} \right\} = 1 - \alpha \tag{4}$$

where $1-\alpha$ is the confidence level, and $0 \le \alpha \le 1$.

m_X is the mean of runoff forecast errors.

s_X is the standard deviation of runoff forecast errors.

N is sample size.

3 Results and discussion

3.1 Runoff simulation

Before utilising the super tank for runoff prediction, the calibration process for lump parameters, the gradient and initial storage of the groundwater tank that are applied for all grid cells, was conducted. The calibration process was simply based on trial and error approaches using historical rainfall and flow data of the wet seasons 2008 and 2009. The best values for the groundwater tank gradient and initial storage were found of 5.0E-2 and one fourth of the tank depth (approximate 10m) respectively. Results showed that the simulated hydrograph agreed very well with the observed hydrograph, as seen in Figure 3. The coefficient of model efficiency of 0.84 was attained for the overall model performance.

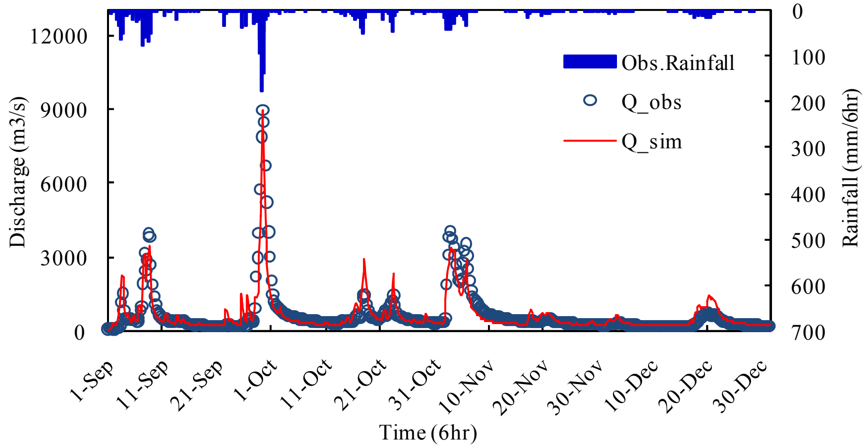

Figure 3: Time series of observed hydrograph (Q_obs) and simulated using rain-gages (Q_sim) for period Sep-Dec, 2009.

3.2 Downscaled precipitation

The large scale precipitation downscaling process was divided into learning and validation phases. In which the data set of twelve storm events that occurred in the wet seasons 2008 and 2009 were selected for the learning phase and the last event occurred in late 2009 was chosen for the validation phase. The optimal large scale predictors obtained from the NWP model output for the input layer of the ANN was finalised based on the predictor screening processes [8]. It includes vertical changes in atmosphere pressure at the layers 700hPa and 850hPa, and quantitative precipitation forecast for the surface layer. The output layer was the downscaled precipitation.

It is obviously that great attempts have been made to improve the forecast reliability of NWP models, however in terms of precipitation forecast; the NWP still tends to underestimate intense rainfall. This underestimation is clearly observed in Figure 4 for the precipitation forecast of 24-hour lead-time derived from DMO, mostly lower than those observed by rain-gages. Meanwhile, the downscaling results showed that the precipitation forecasts by ANN depicted better agreements with the observed rainfall during the learning phase. However, discrepancies were found mostly at the early period of the storm event. This is

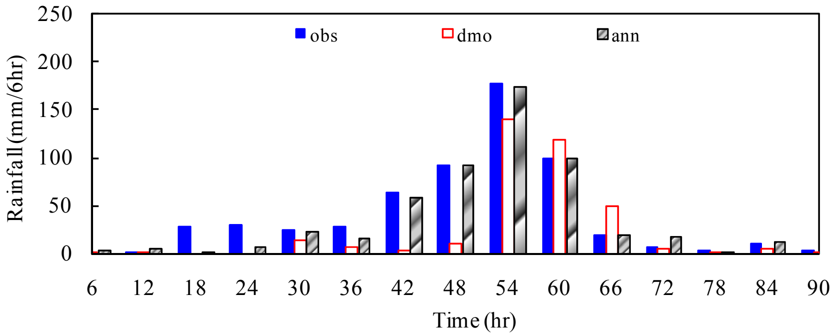

Figure 4: Hyetographs of areal-average rainfall obtained from observation (obs), direct model output (dmo) and downscaling model (ann) for the training event on Sep. 27-30th, 2009.

simply defined as a result of the errors either in the specification of the initial model state or in the model formulation. In which biases induced by the initial model state tend to be a major source due to small errors in initial conditions of the model are likely to increase the model total biases significantly (Ehrendorfer [13]). Because these initial conditions are mainly based on a very coarse spatiotemporal resolution of the existing weather observation networks, especially over large bodies of water such as the Pacific Ocean.

3.3 Flood forecast

This section demonstrates the experiment on flood forecast using the precipitation forecasts that either obtained from the DMO or resulted from the learning phase and the validation phase of the downscaling model by ANN. The precipitation forecasts were then input to the super tank model for runoff prediction. In present study, the storm event on Sep. 27th-30th, 2009, represented the learning phase. Meanwhile, the storm event on Nov. 1st-7th (the last storm in the wet season, 2009) was selected for the model validation. The predicted runoffs were compared to those obtained from actual observation. It is clearly observed in Figures 5a and 5b that the flood forecast with 24-hour lead-time based on DMO demonstrates significant underestimates to the observed hydrograph as a result of the underestimation of the DMO. In the meantime, the flood forecast using ANN driven precipitation forecast exhibits considerable improvement though underestimates of the peak discharges are observed for the validated event. However, the comparison of total volume showed a better estimation to the actual volume, approximate 25% lower for the validated event.

In this case, a good estimation of incoming volume might be useful information for operators to regulate the floods through reservoir systems.

3.4 Skill score of precipitation forecast

The reference forecast based on climatology in this study was defined as the mean precipitation of the wet seasons 2008 and 2009. Accordingly, the MSE skill scores of various forecast lead-times were computed for the downscaled precipitation which was selected from the learning phase and validation phase of the downscaling process. These skill scores were also compared to those obtained using DMO.

Results show that overall skill scores of precipitation forecast based on ANN outperformed those based on DMO, as illustrated in Figures 6a and 6b. The skill scores tend to approach to unity (the perfect score) in most cases during the learning phase (Figure 6a). Meanwhile, it was found a large variation with increasing lead-time for the validated event, two cases (6 and 18 hours lead-time) were observed showing lower forecast skills than those based on DMO (Figure 6b). However, it demonstrated relatively consistent forecast skills up to 18-hr lead-time.

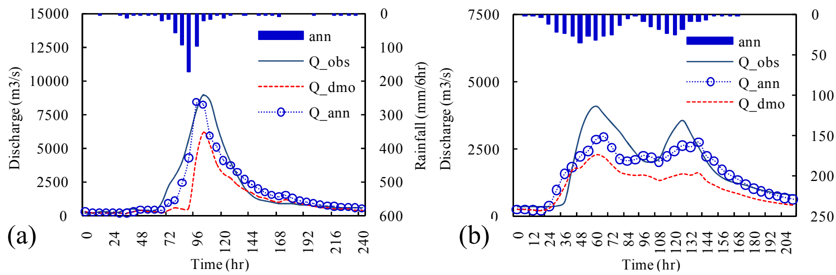

Figure 5: Time series of observed hydrograph (Q_obs) versus predicted hydrographs based on DMO (Q_dmo) and ANN (Q_ann) for (a) the learning phase and (b) validation phase.

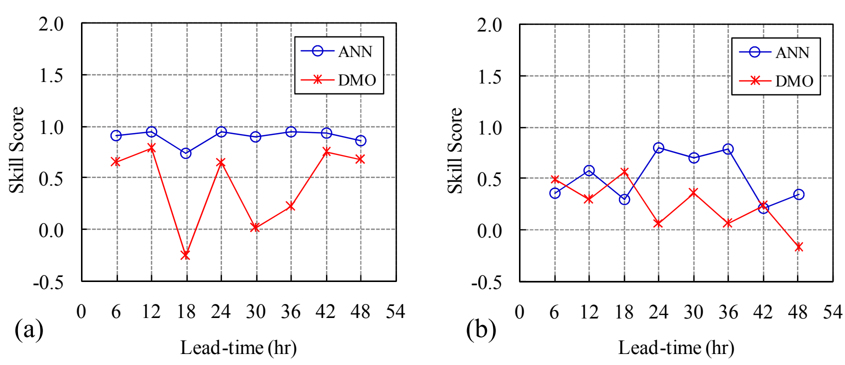

Figure 6: Skill scores of precipitation forecast by ANN and DMO for (a) the learning event on Sep. 27th-30th, 2009 and (b) the validation event on Nov. 1st-7th, 2009.

3.5 Uncertainty of runoff prediction

In this study, uncertainties of runoff prediction based on downscaled precipitation were assessed for the validated event. As stated previously, the accuracy of runoff prediction is very much dependent on the quality the precipitation forecasts. Therefore, errors in runoff prediction are considered to have the similar order of those from the precipitation forecast, the greater forecast lead-time the larger model uncertainty is expected.

Results showed that the relative errors of prediction of peak discharges and total volumes are found proportional to the increasing lead-time. An increasing trend of these errors was observed towards greater forecast lead-times (Figure 7a). Simultaneously, the similar tendency was found for the 95% confidence intervals of the mean of runoff prediction errors (Figure 7b). However, high forecast skills were found in both cases for the forecast lead-time of 6-18 hours.

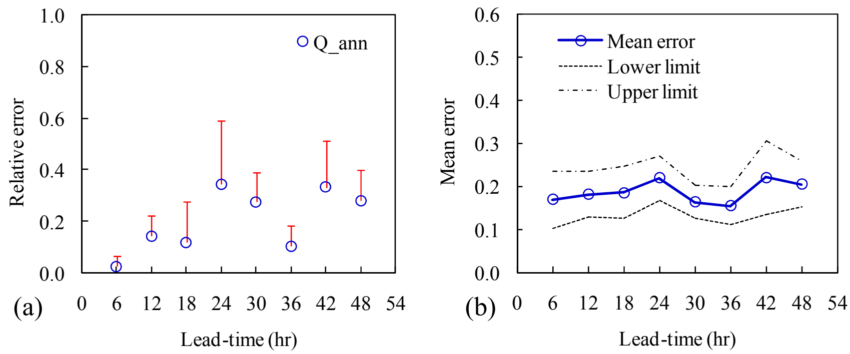

Figure 7: (a) Relative errors of peak discharge and total volume and (b) 95% confidence interval of the mean of runoff error versus forecast lead-time for the validation event on Nov. 1st-7th, 2009.

4 Conclusions and remarks

The paper has presented the quantification of model uncertainty for the short-term flood forecast based on the global NWP model and the distributed rainfall runoff model for a river basin in Central Vietnam. The findings are summarised as following:

(i) Errors in precipitation forecast were found as a major contribution to the total uncertainty of the runoff prediction model, especially those based on DMO.

(ii) The downscaled precipitation using ANN method has performed better forecast skills than the direct precipitation forecast by the NWP model. As a result, higher forecast skills have been observed for the runoff prediction based on the downscaled precipitation.

(iii) Though an increasing trend of uncertainty was observed as lead-time increased, the model can produce reliable forecast up to 18-hour lead- time.

However, it is essential to examine the forecast model with more validation events in the future in order to integrate the model results with other flood control measures for flood damage reduction.

Acknowledgements

This study is under the financial support of the Institute for International Advanced Research and Education of Tohoku University, Japan, and JST/JICA, SATREPS (Science and Technology Research Partnership for Sustainable Development).

Reference

[1] Thielen, J., et al.., The European flood alert system – Part 1: Concept and development. *Hydrol. Earth Syst. Sci.*, **13**, pp. 125-140, 2009.
[2] Micha, W., et al.., Flood forecasting and warning at the river basin and at the European scale. *Natural Hazard*, **36**, pp. 25-42, 2005.
[3] Wind, H., et al.., Analysis of flood damage from the 1993 and 1995 Meuse floods. *Water Resour. Res*, **35**, pp. 3459-3465, 1999.
[4] Kardhana, H. and Mano, A., Uncertainty evaluation in a flood forecasting model using JMA numerical weather prediction. *Journal of Disaster Research*, **Vol.4 No.4,** pp. 272-277, 2009.
[5] Xuan, Y., et al.., Uncertainty analysis of hydrological ensemble forecasts in a distributed model utilising short-range rainfall prediction. *Hydrol. Earth Syst. Sci.*, **13**, pp. 293-303, 2009.
[6] Yokoi, S. and Matsumoto, J., Collaborative effects of cold surge and tropical depression-type disturbance on heavy rainfall in Central Vietnam. *Monthly Weather Review*, **Sep**, pp. 3275-3287, 2008.
[7] Dawson, C.W. and Wilby, R.L., Hydrological modelling using artificial neural networks. *Progress in Physical Geography*, **25,1**, pp. 80-108, 2001.
[8] Nam, D.H., Udo, K. and Mano, A., Use of ANN in quantitative rainfall prediction for flood forecast in Central Vietnam. *Proceedings of the 9th International Conference on Hydroinformatic*, **Vol.3**, pp. 1811-1818, 2010.
[9] Kardhana, H., Tatesawa, H. and Mano, A., Flood forecast based on numerical prediction and distributed runoff model. *River Basin management IV - WIT Transactions on Ecology and the Environment*, **Vol.104**, pp. 201-211, 2007.
[10] Maskey, S., Modelling uncertainty in flood forecasting system. *PhD thesis, A.A Balkema Publisher, the Netherlands*, 2004.
[11] Murphy, A.H., Skill score based on the mean square error and their relationships to the correlation coefficient. *Monthly weather review*, **Vol.116**, pp. 2417-2424, 1988.
[12] Dingman, S.L., Physical hydrology – 2nd edition. *Prentice-Hall, Inc.*, pp. 573-574, 2002.
[13] Ehrendorfer, M., Predicting the uncertainty of numerical weather forecasts: A review. *Meteorol. Zeitschrift*, **N.F.6**, pp. 147-183, 1997.