WATER RESOURCES RESEARCH, VOL. 35 NO. 4, PAGES 1191–1197, APRIL 1999

River flood forecasting with a neural network model

Marina Campolo, Paolo Andreussi,[1] and Alfredo Soldati

Centro di Fluidodinamica e Idraulica and Dipartimento di Scienze e Tecnologie Chimiche
Università di Udine, Udine, Italy

Abstract. A neural network model was developed to analyze and forecast the behavior of the river Tagliamento, in Italy, during heavy rain periods. The model makes use of distributed rainfall information coming from several rain gauges in the mountain district and predicts the water level of the river at the section closing the mountain district. The water level at the closing section in the hours preceding the event was used to characterize the behavior of the river system subject to the rainfall perturbation. Model predictions are very accurate (i.e., mean square error is less than 4%) when the model is used with a 1-hour time horizon. Increasing the time horizon, thus making the model suitable for flood forecasting, decreases the accuracy of the model. A limiting time horizon is found corresponding to the minimum time lag between the water level at the closing section and the rainfall, which is characteristic of each flooding event and depends on the rainfall and on the state of saturation of the basin. Performance of the model remains satisfactory up to 5 hours. A model of this type using just rainfall and water level information does not appear to be capable of predicting beyond this time limit.

1. Introduction

Forecasting river flow after heavy rain is important for public safety, environmental issues, and water management. For these purposes, mathematical models have been developed based either on physical considerations [e.g., *Garrote and Bras*, 1995; *Refsgaard and Knudsen*, 1996; *Todini*, 1996; *Buchtele et al.*, 1996] or on a statistical analysis [see *Hsu et al.*, 1995; *Mukherjee and Mansour*, 1996; *Raman and Sunilkumar*, 1995]. Both approaches are difficult, and the type of forecast they provide is not completely satisfactory.

During heavy rain periods, some of the terms of the hydrological balance (evaporation, infiltration, and storage variation) can be neglected since they give no relevant contribution to the river flow rate in the short period. In contrast, accurate information on rainfall and on the state of the basin must be available. The rainfall gives a measure of the amount of water gathered by the basin and represents the perturbation experienced by the river system. The state of the basin, which is correlated, albeit indirectly, to the flow rate or, alternatively, to the water level, represents the capability of the river system to respond to rainfall perturbation. However, even in these simplified conditions, the usual approaches prove to be inefficient or too burdensome [*Woolhiser*, 1996]. In this paper we develop a fairly simple neural network model to predict the flow rate in a river during heavy rain periods.

In the hydrological context, as in many other fields, artificial neural networks (ANN) are increasingly used as black-box, simplified models [*Bishop*, 1994]. For hydrological applications, ANN models can take advantage of their capability to reproduce the unknown relationship existing between a set of input variables descriptive of the system, for example, rainfall, and a set of output variables, for example, river flow rate [*Chakraborty et al.*, 1992].

Previous work has demonstrated that ANNs are adequate to model the rainfall-runoff process [*Zhu et al.*, 1994; *Minns and Hall*, 1996; *Shamseldin*, 1997]. A comparison between ANN models and traditional models has been made by *Hsu et al.* [1995], who concluded that the ANN approach is more effective and more efficient whenever explicit knowledge of the hydrological subprocess is not required and when the object is to predict streamflow behavior from customary monitored time series of rainfall and flow rate.

Most of the previous work considered rainfall data averaged over the basin scale; this has the advantage of reducing the number of input variables but does not allow identification of the contribution of runoff routed from different subbasins. Recently, *Shah et al.* [1996] and *Smith and Eli* [1995] established the importance of considering the effect of spatially distributed rainfall. However, they used synthetic, stochastically generated rainfall patterns and runoff data.

In the present work we use an ANN to predict the occurrence of floods from available distributed rainfall and hydrometer data collected in the basin of the river Tagliamento, in Italy. The geography of the basin, which features various subbasins contributing to the main flow in different circumstances, is such that distributed rainfall information is required to obtain accurate streamflow prediction [*Shah et al.*, 1996].

2. Database

The data used in this work refer to the river Tagliamento, in Friuli, Italy, shown in Figure 1. The basin of the river (overall area is about 2480 km^2) includes various subbasins, and the monitoring system consists of a network of rain gauges and a series of hydrometers. A subset of rain gauges (located at Paularo, Ampezzo, Pesariis, Resia, and Moggio) and one hydrometer (located at Venzone) were used in this work.

A database of 20 different flooding events was selected from available historical records relative to the last 20 years. These events represent periods in which variation of the water level

[1]Now at Dipartimento di Chimica e Chimica Industriale, Università di Pisa, Pisa, Italy.

Paper number 1998WR900086.
0043-1397/99/1998WR900086$09.00

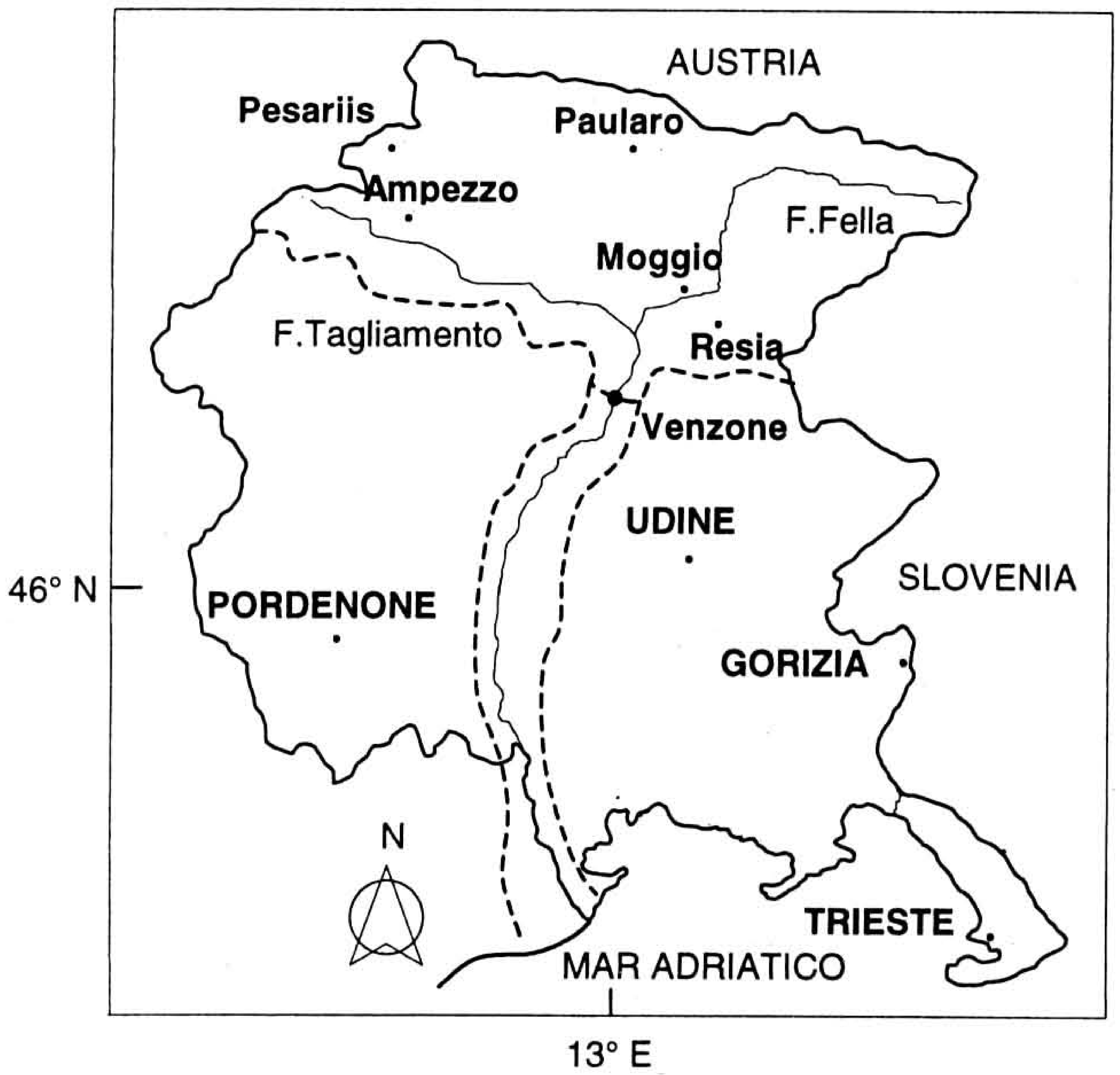

Figure 1. Map of the basin of the river Tagliamento showing rain gauge sites, hydrometer location, and extension of the basin under study.

of more than 100% of the starting value was recorded over a time period of less than a day at Venzone, shown in Figure 1, located at the end of the mountain district and selected as the closing section of the basin under study. This portion of the basin, with an area of 1950 km^2, is indicated in Figure 1 and covers approximately 80% of the entire catchment.

Since a flooding event may be generated by rainfall over the whole basin or by rainfall only occurring in certain regions of the basin, and the response of the basin depends on previous conditions in the catchment [*Shah et al.*, 1996], it seemed appropriate to use a spatially distributed description of the rainfall input without averaging over the basin. A preliminary analysis with basin-averaged and local data confirmed the need for such an approach. The rainfall distribution is very localized in time and space, with high values of the variance among local and basin-averaged values. The use of a basin-averaged datum would require the identification of a suitable averaging procedure, accounting for the different relevance of local contributing areas and for the time required for the contribution from each site to influence the flow rate at the closing section. On the one hand, this would simplify the model by reducing the input data, but, on the other hand, it would overburden data preprocessing. In the present case we chose to use the available local data without averaging. Data were recorded continuously; however, we decided to use data on an hourly basis, since this is an economic timescale enabling an appropriate description of the rainfall runoff process in the basin under study. For each event, rainfall and hydrometer data were collected for a time period starting before the heavy rain and ending when the rainfall contribution to the river flow rate had ceased. Although flooding events are concentrated in two different periods of the year, i.e., during spring and fall, a relationship between features of each event and seasonality was not clear from the available data. Note that we used the water level at the closing section of the basin rather than the discharge as the runoff variable since the stage-discharge relationship at the closing section was not available.

Table 1. Results of Correlation Analysis, Expressed in Hours

Rain Gauge Sites	Minimum Lag	Maximum Lag	Mean Lag
1, Paularo	6	18	13
2, Ampezzo	5	17	12
3, Pesariis	7	18	13
4, Resia	6	18	12
5, Moggio	7	17	13

In order to examine the contribution of the rainfall in each region to the flow rate at the closing section for each event, we calculated the cross correlation between rainfall and water level, and we determined the time lag between rainfall at each rain gauge and water level in the closing section, as reported in Table 1. We assumed a high correlation between rainfall and water level when the cross-correlation diagram for each flood event presented a peak, regardless of its value. Then the minimum, maximum, and averaged values of this lag were calculated from the frequency distribution of time lag over the events of the database. The lag is influenced first by the distance of the rain gauge from the hydrometer and then by the state of the basin, i.e., the lag is smaller in a saturated basin. However, it is hard, if not impossible, to distinguish properly how the local rainfall contributes to water level variation, since convolution of responses and nonlinearities should be considered. The correlation analysis indicates that minimum and maximum time lag are more or less the same; from 5 to 18 hours, for sites at different distances from the closing section (Table 1). This was somewhat expected, since when a flood is produced by widespread rainfall, the rainfall coming from the region nearer to the closing section is responsible for the steep rising limb, while the rainfall coming from outer regions contributes to the smearing of the falling limb. Correlation between rainfall at different sites and overall water level variation identifies only the averaged lag.

As previously observed by *Minns and Hall* [1996], rainfall information alone is not sufficient to compute flow rate, since the state of the basin plays an important role in determining flow rate behavior. For this reason we used the hydrometer recording at certain time intervals before the time of prediction as additional input information. In this way, information about the state of the basin is introduced.

Available records have been divided into two independent sets: calibration and validation. Grouping of flooding events was carefully done in order to ensure that statistical properties of each group, i.e., mean, variance, minimum, and maximum value of water level, were most similar. The calibration ensemble was further divided into two portions, the training set and the testing set, using the same criteria discussed above. The training set was used to minimize the error, and the testing set was used to avoid overfitting when implementing the neural model [see *Bishop*, 1994; *Haykin*, 1995]. From each data set, patterns, i.e., related couples of input-output data, were generated in order to describe the transformation of rainfall into runoff during the rising limb and the falling limb of the flood. Rainfall and hydrometer data were normalized in the range [0.1:0.9]. Normalization was obtained by linear scaling of the values of water level and rainfall from the range of variation of the available data, $[x_{min}:x_{max}]$, producing dimensionless variables, x_{norm}, given by

$$x_{\mathrm{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} (0.9 - 0.1). \tag{1}$$

3. Methodology

3.1. Neural Network Based Model

The model is based on a feed-forward neural net, with logistic activation function. The net is feed-forward in the sense that propagation of the signal is unidirectional, without feedback cycles between nodes. Input information coming into the model through special, nonprocessing input nodes is fed to a number of inner, hidden nodes. Hidden nodes perform a twofold function: First, they compute a signal z_j from all incoming information using the following relation:

$$z_j = \sum_i w_{i,j} x_i - \sigma_j \tag{2}$$

where $w_{i,j}$ is the weight associated with the connection from node i, the input node, and node j; x_i is the input variable value; and σ_j is a threshold value, which is different for each node. Second, they transform this signal using a nonlinear activation function f. In this work, f is the logistic or sigmoid function and has the following form:

$$f(z_j) = \frac{1}{1 + \exp[-z_j]}, \quad f(z_j) \in [0:1]. \tag{3}$$

The outgoing signal is then fed to the subsequent layer of nodes (this may be a hidden layer or an output layer, depending on the problem complexity) and transformed in the same way as described by (2) and (3).

The transformation of the input variables array, $\bar{I}$, into the corresponding output variables array, $\overline{O_c}$ can be expressed synthetically as

$$\overline{O_c} = F(\bar{I}, \bar{W}, \bar{\sigma}, N, f). \tag{4}$$

The adjustable parameters of the model are weights $\bar{W}$, thresholds $\bar{\sigma}$, and number of hidden nodes N for each hidden layer eventually present.

3.2. Calibration of the Model

Implementation and calibration of the neural model were realized using the software Stuttgart Neural Network Simulator (SNNS), developed by the University of Stuttgart and available as free software from the Internet. The calibration of the model, also called training, is performed by minimizing the function $E(\bar{W})$, representing the square of the error of the model with regard to real data, using the back propagation error algorithm to correct the weights [see *Bishop*, 1994; *Haykin*, 1995]. The error of the model is computed on all the calibration data at a time (batch training). In brief, variation of the weight $w_{i,j}$ from node i to node j is calculated as

$$\Delta w_{i,j} = -\eta \frac{\partial E}{\partial w_{i,j}}, \tag{5}$$

where η is the learning rate. Correction of the value of the weights corresponds to moving along the error surface $E(\bar{W})$ toward the local minimum following the direction of the steepest gradient.

In this work a decreasing learning rate (from 0.5 to 0.01) was used to accelerate convergence toward a global minimum. For each model structure the training procedure was repeated starting from independent initial conditions in order to select the best performing of the trained nets. The trend of $E(\bar{W})$ calculated using data from the training and testing sets was used to determine when learning was optimal. Training was stopped when no more sensible improvement in performance was found, i.e., the variation of $E(\bar{W})$ was negligible with proceeding of time, or when overtraining started, i.e., $E(\bar{W})$ calculated on the testing set started rising even if $E(\bar{W})$ calculated on the training was still decreasing [*Bishop*, 1994; *Haykin*, 1995].

4. Results and Discussion

4.1. Input Data and Structure of the Model

In this work we developed two models to predict the water level of the river at short times (1 hour in advance) and at longer times (up to 10 hours). Input data differ depending on the prediction to be made. For short-range prediction at time T we used (1) the water level at times $T-1$, $T-2$, $T-3$, and $T-4$; and (2) the rainfall at times $T-\Delta_i$, $T-\Delta_i+1$, $T-\Delta_i+2$, and $T-\Delta_i+3$, where Δ_i is the mean time lag of each rain gauge. The temporal trend of precipitation is used as input to better characterize the response of the basin, since cumulated rainfall and rate of variation of intensity are important as well as the intensity itself [*Mason et al.*, 1996]. The temporal trend of the water level is important since it determines the response of the basin to rainfall perturbation. For prediction at longer times, from time T to time $T+9$, we used (1) the water level from time $T-1$ to time $T-5$; and (2) the rainfall depths recorded by the rain gauges from time $T-15$ to time $T-1$. This time interval includes the mean time lag of each rain gauge and has been extended up to present time since prediction of the water level of future time intervals also requires rainfall information about the recent past.

On the basis of the structure identified for the input-output transformation, patterns, i.e., related couples of input-output data, have been generated from each data set. For the short-term prediction model the numbers of patterns generated were 1113, 333, and 591 from the training, testing, and validation data sets, respectively. For the long-term prediction model the numbers of patterns generated were 969, 268, and 503 from the training, testing, and validation data sets, respectively. These numbers are smaller than those of the short-term prediction model because the time string required in input was longer. Generated patterns describe the input-output transformation during the rising limb and the falling limb of the flood, with consecutive patterns representing a shift in time of the input and output temporal windows.

The selection of the final structure for the models, i.e., the number of hidden nodes, was a trial-and-error procedure. Starting from an oversized hidden layer of 15 nodes, neural network models were trained until a minimum for $E(\bar{W})$ for the particular configuration was obtained. A different error function $E(\bar{W})$ was used for short- and long-term prediction. For short-term prediction, $E(\bar{W})$ was calculated on a single output node, and minimizing $E(\bar{W})$ was equivalent to minimizing the error on the water level predicted 1 hour in advance. For long-term prediction, $E(\bar{W})$ was calculated over 10 output nodes, thus representing the error in prediction over the entire time horizon. Once the minimum for $E(\bar{W})$ was found, the number of hidden nodes was progressively decreased and the training repeated. The process continued until

Table 2. Error Standards Used to Evaluate Model Adequacy

	Calibration	Validation
RMSE	1.39×10^{-2}	1.34×10^{-2}
E_{max}	1.04×10^{-1}	1.02×10^{-1}
N_{RMSE}, percent	20.5	14.4
R^2, percent	98.4	98.5

Standards are root mean square error (RMSE), maximum error, number of patterns with error exceeding the RMSE, and model efficiency calculated during calibration and validation of model.

the point in which performance began to decrease was reached.

One layer of hidden nodes was sufficient to perform the input-output transformation for both short-range and long-range prediction. We found that the number of hidden nodes necessary to obtain satisfactory results was 3 and 10 for the 1-hour and longer time forecasts, respectively.

4.2. Forecasting at Short Times

The adequacy of the model was evaluated by estimating the maximum error E_{max}, the root mean square error RMSE, the percentage of patterns for which the error exceeded the RMSE, N_{RMSE}, and the model efficiency R^2, defined as

$$R^2 = \frac{F_0 - F}{F_0} \tag{6}$$

where F_0 and F are the variance of the water levels h_i about the mean $\bar{h}$, and the residual model variance, i.e., the model square error, respectively, given by

$$F_0 = \sum (h_i - \bar{h})^2$$

$$F = \sum (h_i - \hat{h}_i)^2$$

where $\hat{h}_i$ is the water level calculated by the model. The mean value of the water level computed over a segment of the time series can be considered a raw prediction for future values of the same time series. The model efficiency can be used to evaluate the capability of the model in predicting water level values different from the mean value, which is assumed to be the prediction, however available, in the worst case. The results obtained with the model are summarized in Table 2, in which parameters have been calculated for the 1-hour forward prediction and are shown with reference to the different data sets used for the calibration and the validation of the model. It is worth noting that errors are fairly small: RMSE is about 0.013 normalized units, corresponding to 6 cm. The maximum error also appears to be small, as it is bounded by a value of about 0.1 normalized units, which corresponds to 70 cm. The distribution of the error over the different patterns was also calculated. The value of N_{RMSE} indicates that almost 80% of patterns were reproduced with an error smaller than RMSE. In the bottom row the model efficiency R^2 is shown. A value of R^2 of 90% indicates very satisfactory model performance, a value in the range 80–90% indicates fairly good performance, and values below 80% indicate an unsatisfactory fit [*Shamseldin*, 1997]. According to this parameter, model predictions are extremely good. Results obtained from the network on data from the validation set are also plotted in Figure 2 in a dispersion diagram. Reduced scatter confirms that good results have been obtained.

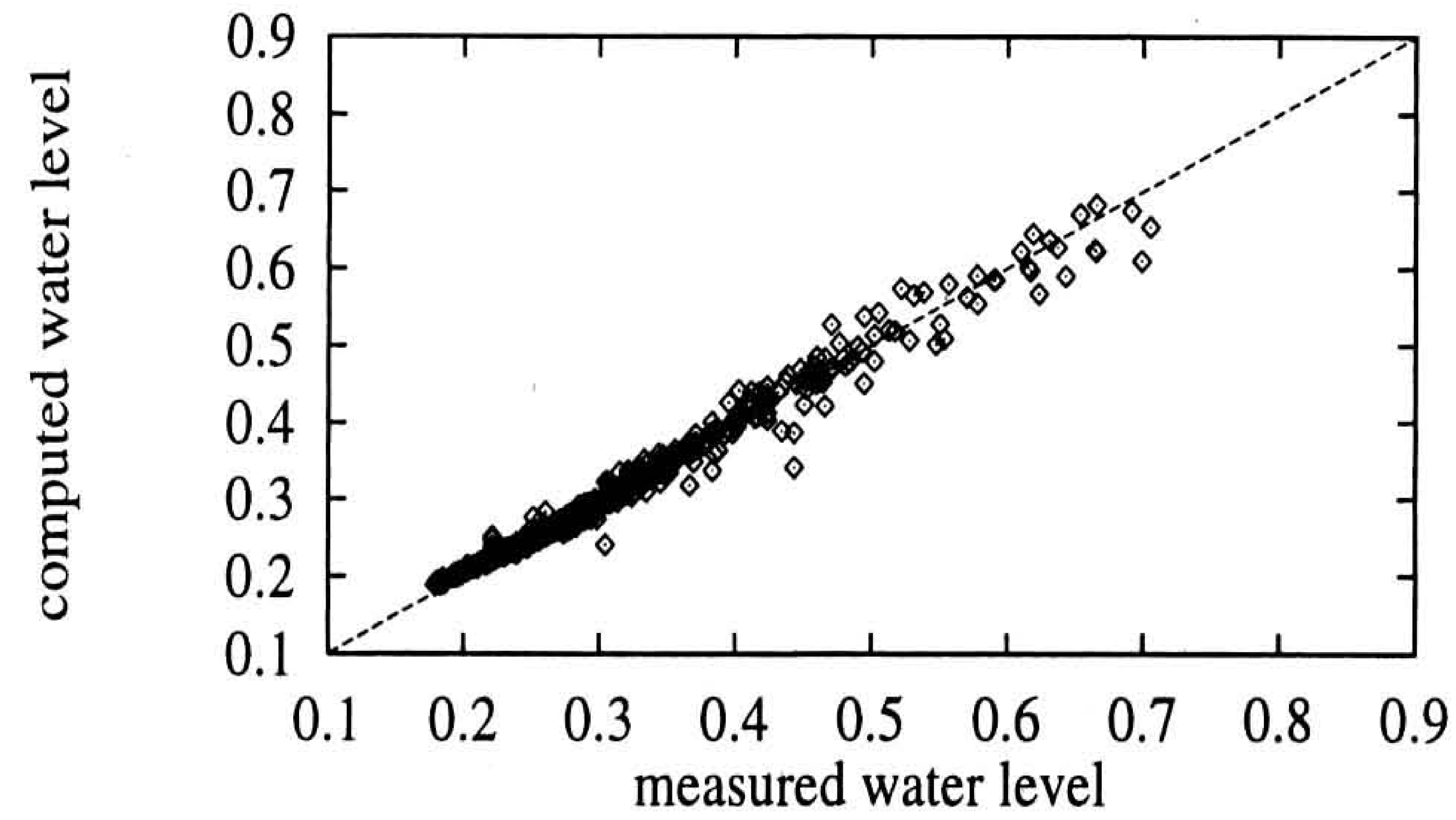

Figure 2. Comparison between measured and calculated water level for data used during model validation.

The behavior in time of the model is compared against experimental data for several events in Figure 3. Time is measured in hours and has a relative meaning only. The error between the values of computed and measured water levels is

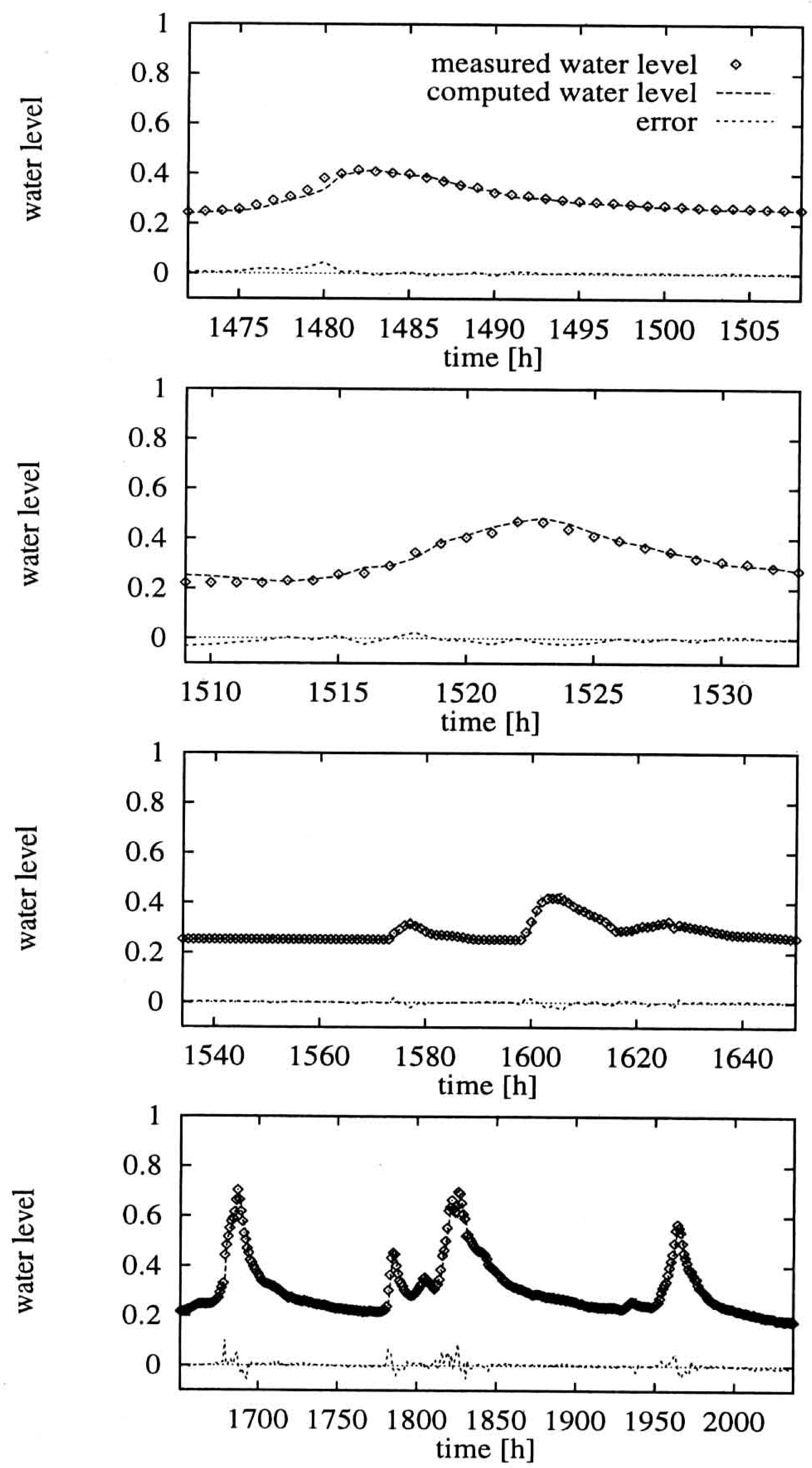

Figure 3. Time series of measured and computed water level for events from validation set. Water level is normalized.

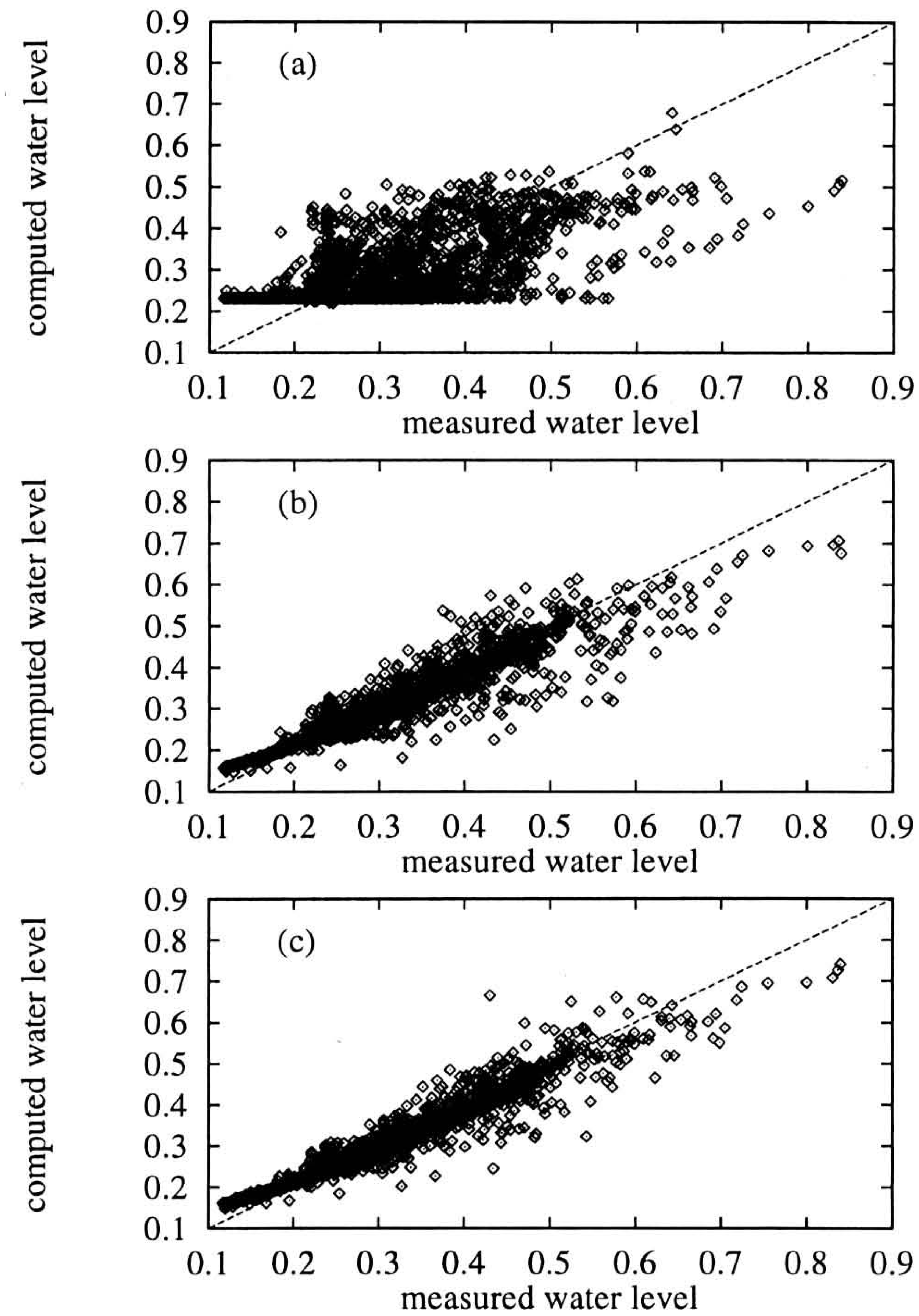

Figure 4. Importance of previous water level information. Shown are results obtained from calibration and validation sets using (a) no water level information, (b) water level 4 hours before, and (c) water level from 4 to 2 hours before.

also reported for each time step, and no preferential distribution of the error can be noted with respect to the value of the water level. An important characteristic of level behavior is the peak height. The error in the prediction of peak height, defined as the absolute difference between the observed and computed water level over the observed water level at the peak, can be used to evaluate the reliability of the model in forecasting floods. The peak error percent on predicted peaks is less than 10% on the validation data.

Results obtained with this model proved that information given as input is sufficient to capture the transformation. However, if our target is to predict the water level at time T, supplying the experimental value of the level at time $T - 1$ is too stringent a requirement for the model and makes it impractical to use for flood risk warning. As previously described, the model uses water level at times $T - 1$, $T - 2$, $T - 3$, and $T - 4$. In order to assess the importance of the level information, we examined the performance of the model in the cases in which (1) no water level data are supplied, (2) water level at $T - 4$ only is supplied, and (3) water level at $T - 4$, $T - 3$, and $T - 2$ is supplied. The results of this analysis are presented in Figure 4 relative to data from calibration and validation sets. Consider first Figure 4a, in which the model makes use only of rainfall information; results are extremely poor. In this condition it is impossible for the neural model to give good results because input-output mapping is not single-valued. Since the network is trained by exploiting data corresponding to the rising and falling limbs of the flood, during the falling limb the input information represented by zero rainfall should be mapped into a varying water level. This cannot be learned by the model, which associates to this input an output value most frequently found in the data set in the given conditions (about 0.22 normalized units for the water level). This is clearly visible in Figure 4a, where the flat distribution of points represents water level at time intervals in which the same value is calculated even if different values were measured. In Figures 4b and 4c the accuracy of the prediction increases, but results are not as satisfactory as those in Figure 2. This means that information about the capability of the basin to respond to a rainfall perturbation is more accurate when recent water level values are used.

4.3. Forecasting at Longer Times

In order to forecast the water level at longer times, from time T up to time $T + 9$, the parameters of the model introduced in section 4.1 were optimized again. The water level at the closing section measured at time T depends on the rainfall that occurred a few hours before in the mountain. The time required for the rain pulse to induce a change in the water level at the closing section was calculated with the correlation analysis, the results of which are presented in Table 1. The results are averaged over each single event history and over all events. The response time of the river depends on the state of saturation of the basin, which is a function of the rainfall history in the days or weeks preceding the flooding event. In a model of the type developed in this work, the only information available on the state of the basin before the heavy rain period is the water level at the closing section. This information is not sufficient to characterize the initial conditions of the basin and explains why water level forecasts for longer times become inaccurate.

In Figures 5a and 5b the trends of RMSE and R^2 are presented versus the prediction time horizon ΔT_p. As was expected, RMSE increases with ΔT_p, whereas R^2 decreases with ΔT_p. In particular, in Figure 5a the increase in RMSE is almost linear as the prediction time horizon increases, and a slight difference in the average rise rate of RMSE for calibration and validation is found. From comparison of the results obtained during calibration and validation, it can be observed that the rate of variation of R^2 (Figure 5b) increases slightly at 5 hours, i.e., the minimum time lag from the correlation analysis, for the validation set. Displacement between the two R^2 curves increases as the prediction time horizon becomes greater than 5 hours. This confirms previous considerations about the relation between the time horizon of prediction and information available for prediction (section 4.2). However, if we consider a prediction of water level 5 hours in advance to be satisfactory, we have a model efficiency of about 85%, which is more than adequate for this type of problem [see *Shamseldin*, 1997], and a root mean square error of about 0.04, which corresponds to 28 cm.

Results obtained in predicting the water level 1 hour in advance using the short-term or the long-term prediction models are different. This can be expected, since different evaluation criteria are used to decide when learning is optimal. For short-term prediction, minimizing $E(\bar{W})$ is equivalent to minimizing the error in the water level predicted 1 hour in advance. For long-term prediction, minimizing $E(\bar{W})$ is equiva-

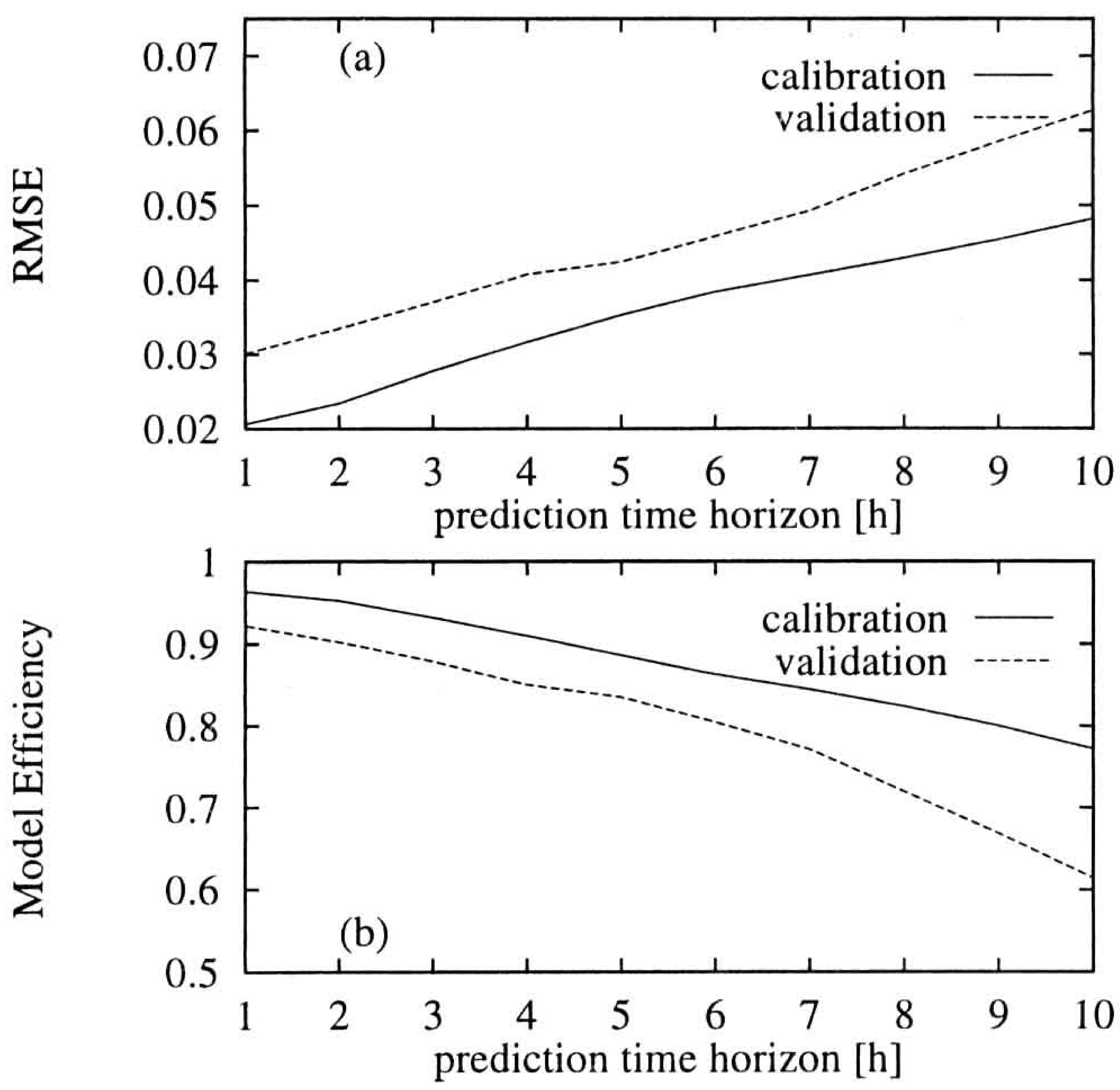

Figure 5. Root mean square error (RMSE) and efficiency of model (R^2) as function of prediction time horizon from calibration and validation sets.

lent to minimizing the error over the entire time horizon. For this reason, the final, optimized configuration of the short-term prediction model cannot be considered a subset of the final configuration of the long-term prediction model, and results are expected to be different.

In Figure 6 the behavior of the calculated water level is compared with the measured water level for different flooding events of the validation set. The largest error is found at the peaks; both height and location of the peaks are affected by an error which increases with ΔT_p. The error in the height of the peak can be analyzed by examining Figure 7. In Figure 7a the heights of the most representative peaks in the validation set are shown, and in Figures 7b, 7c, and 7d the error relative to the prediction of those peaks 1, 3, and 5 hours in advance, respectively, is also shown. The first observation is that almost all peaks are underestimated. The reason for such underestimation appears to be related to the underestimation of the rise rate of the level, i.e., the gradient of the rising limb, and of its rate of variation. In these cases, an inadequate representation of the state of saturation of the basin is given to the model. Many authors [*Todini*, 1996; *Garrote and Bras*, 1995] point out that the information about the changing state of saturation of a basin is vital in order to correctly predict the flow rate. Lack of this information is the main reason why traditional models experience difficulties in predicting the response of the basin until saturated conditions are achieved, since the major nonlinearities of the rainfall-runoff transformation can be ascribed to this component of the hydrological balance [*Todini*, 1996]. The model of the river system presented was used to identify the events for which the state of saturation of the basin is relevant to predict the water level. In the validation set the largest errors are related to peaks characterized by a rapid rising limb. We use level information to characterize the capability of the river to respond to rainfall. This information is sufficient to represent the changing state of the basin only if the time horizon is less than the response time of the basin for the event considered. In a very rapid flooding event the response time of the basin is very short, i.e., less than 5 hours for our river system. When trying to predict the behavior of rapidly rising limb with a ΔT_p larger than the response time of the basin for that event, water level information given in the input appears to be ineffective in describing the state of the basin.

The limit of the model when predicting rapidly rising limbs can be attributed to the limited set of data used as input. However, these data are sufficient to predict most events, at least when the prediction time horizon is short. Phreatic surface depth next to the closing section of the basin or rainfall gathered by the basin during a proper time period preceding the flood could be used as additional input information to describe, more accurately, the state of the basin, thus improving the predictive capability of the model.

5. Conclusions

This paper addresses the problem of forecasting the river flow rate on the basis of rainfall and water level data. The objective of the study was twofold: First, we wanted to develop a simple tool to analyze river behavior resulting from heavy rainfall; and second, we attempted to set up a model which would be able to forecast the water level of the river on the

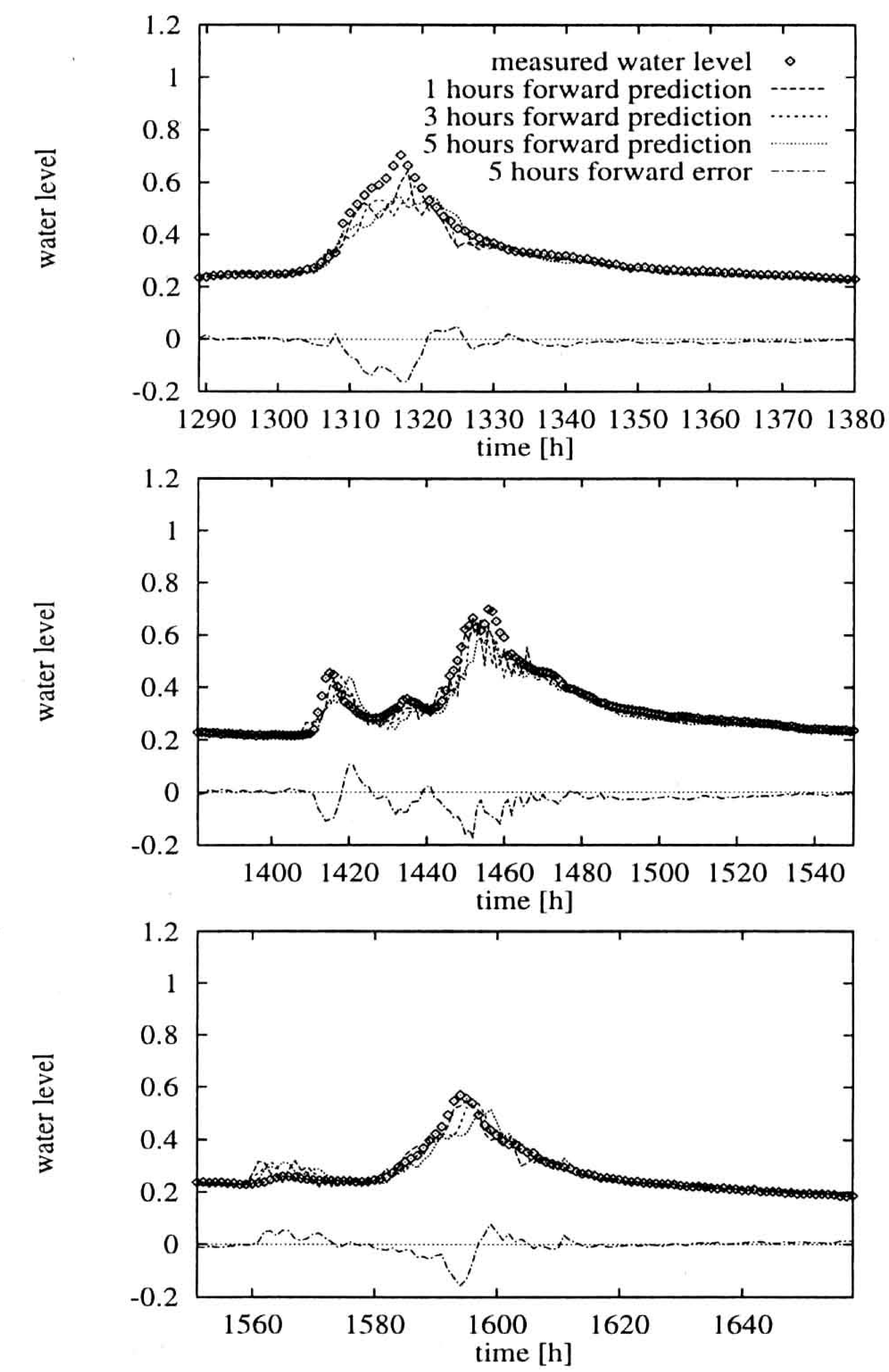

Figure 6. Time series trend of measured and computed water level: real data, model prediction 1, 3, and 5 hours in advance, and error on 5 hours forward prediction.

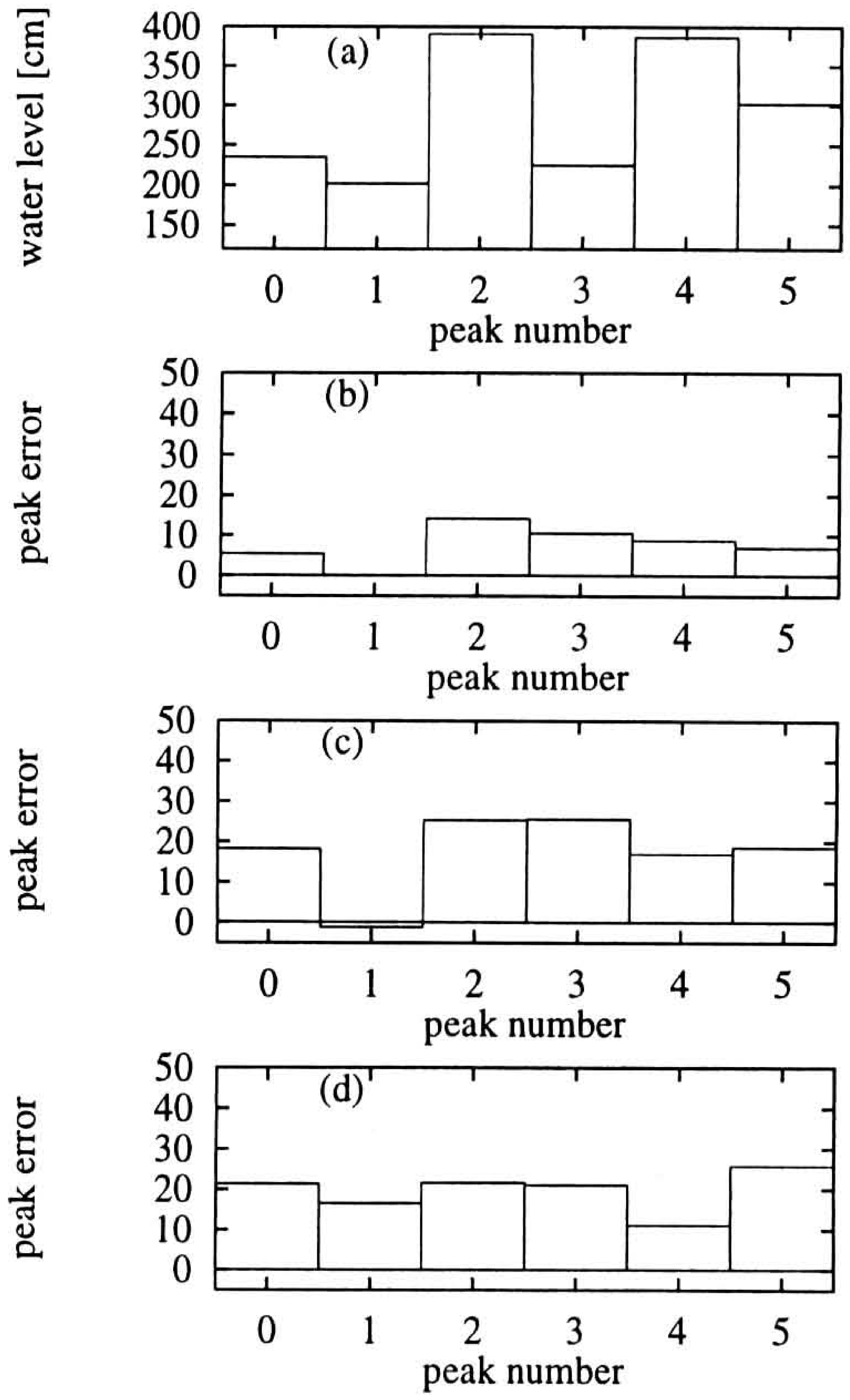

Figure 7. Peak error percent for different peak events from validation set. Shown are (a) value of water level at peak, and error on corresponding water level predicted (b) 1 hour in advance, (c) 3 hours in advance, and (d) 5 hours in advance.

basis of rainfall information and previous water level with a sufficient lead time in order to reduce the consequences of floods. The model is based on an artificial neural network and has the advantages of low cost and simplicity with respect to complex physically based models. The database on which the model was assessed comprised field data collected on the river Tagliamento, in northern Italy. To obtain a precise characterization of the floods, we used the rainfall information distributed over the basin, i.e., not averaged information. This introduces significant changes with respect to previous flood forecasting models [*Hsu et al.*, 1995; *Bertoni et al.*, 1992].

Predictions of the model for 1-hour forecasting are accurate, with a minimal error of the order of 4%. When trying to use the model to forecast water level with a longer time horizon, performance decreased rapidly. The main reason for this depends on the physical response time of the river basin to rainfall. In our model, rainfall is used as input and the river level is used to characterize the capability of the model to respond to rainfall. The water level at the closing section starts to respond to rainfall after a certain time lag characteristic of each rain gauge, i.e., distance, of the conditions of the basin and of the type of rainfall. The way in which the river responds to the rainfall depends first on the perturbation and then on the state of saturation of the basin. The information on the level at the closing section seems to be able to represent the state of the basin only after river flow rate started to be affected by the rainfall event under consideration.

Acknowledgments. We thank the Ufficio Idrografico e Mareografico di Udine for making available the data for this work and Nestor Zanon for collecting and preprocessing the time series.

References

Bertoni, J. C., C. E. Tucci, and R. T. Clarke, Rainfall-based real-time flood forecasting, *J. Hydrol.*, *131*, 313–339, 1992.
Bishop, C. M., Neural networks and their applications, *Rev. Sci. Instrum.*, *65*, 1803–1832, 1994.
Buchtele, J., V. Elias, M. Tesar, and A. Herrman, Runoff components simulated by rainfall-runoff models, *J. Hydrol. Sci.*, *41*, 49–60, 1996.
Chakraborty, K., K. Mehootra, C. K. Mohan, and S. Ranka, Forecasting the behavior of multivariate time series using neural networks, *Neural Networks*, *5*, 961–970, 1992.
Garrote, L., and R. L. Bras, A distributed model for real-time flood forecasting using digital elevation models, *J. Hydrol.*, *167*, 279–306, 1995.
Haykin, S., *Neural Networks: A Comprehensive Foundation*, Macmillan Coll., New York, 1995.
Hsu, K.-L., H. V. Gupta, and S. Sorooshian, Artificial neural network modeling of the rainfall-runoff process, *Water Resour. Res.*, *31*, 2517–2530, 1995.
Mason, J. C., A. Tem'me, and R. K. Price, A neural network model of rainfall-runoff using the radial basis function, *J. Hydrol. Res.*, *34*, 537–548, 1996.
Minns, A. W., and M. J. Hall, Artificial neural networks as rainfall-runoff models, *J. Hydrol. Sci.*, *41*, 399–417, 1996.
Mukherjee, D., and N. Mansour, Estimation of flood forecasting errors and flow-duration joint probabilities of exceedance, *J. Hydrol. Eng.*, *122*, 130–140, 1996.
Raman, H., and N. Sunilkumar, Multivariate modeling of water resources time series using artificial neural networks, *J. Hydrol. Sci.*, *40*, 145–163, 1995.
Refsgaard, J. C., and J. Knudsen, Operational validation and intercomparison of different types of hydrological models, *Water Resour. Res.*, *32*, 2189–2202, 1996.
Shah, S. M. S., P. E. O'Connell, and J. R. M. Hosking, Modelling the effects of spatial variability in rainfall on catchment response: Experiments with distributed and lumped models, *J. Hydrol.*, *175*, 89–111, 1996.
Shamseldin, A. Y., Application of a neural network technique to rainfall-runoff modeling, *J. Hydrol.*, *199*, 272–294, 1997.
Smith, J., and R. N. Eli, Neural network models of the rainfall-runoff process, *J. Water Resour. Plann. Manage.*, *121*, 499–508, 1995.
Todini, E., The ARNO rainfall-runoff model, *J. Hydrol.*, *175*, 339–382, 1996.
Woolhiser, D. A., Search for physically based runoff model—A hydrological El Dorado?, *J. Hydrol. Eng.*, *122*, 122–129, 1996.
Zhu, M. L., M. Fujita, and N. Hashimoto, Application of neural networks to runoff prediction, in *Stochastic and Statistical Methods in Hydrology and Environmental Engineering*, edited by K. W. Hipel et al., pp. 205–216, Kluwer Acad., Norwell, Mass., 1994.

P. Andreussi, Dipartimento di Chimica e Chimica Industriale, via Risorgimento 2, Università di Pisa, 56100 Pisa, Italy.
M. Campolo and A. Soldati (corresponding author), Centro di Fluidodinamica e Idraulica and Dipartimento di Scienze e Tecnologie Chimiche, Università di Udine, via del Cotonificio 108, 33100 Udine, Italy. (alfredo@euterpe.dstc.uniud.it)

(Received March 4, 1998; revised November 2, 1998; accepted November 5, 1998.)