Development of a Flood Forecasting Model For Kalu River and Kelani River Basins in Sri Lanka using Radial Basis Function Neural Networks

Seenipellage Chaminda Sugeeswara* **Supervisor: Prof. A. W. Jayawardena****
MEE09208

Abstract

This research is intended to investigate a kind of river stage/discharge forecasting model suitable for Sri Lankan flood prone rivers as an essential component of an effective early flood warning system. This work was carried out on a belief, that data driven river stage/discharge model would be more suitable for Sri Lankan rivers as there are considerable amount of past records of data while other physical information that are required to develop physically based models are rare. Artificial neural networks (ANN) have recently become popular in many fields including river stage/discharge forecasting in hydrology. Among various types of ANNs, radial basis function networks (RBFN) are popular over the more widely used multilayer perceptrone for its faster convergence, smaller extrapolation errors and higher reliability. In this study, generalized radial basis function network (GRBFN) with fully supervised algorithm (trained using daily data) was applied to two most flood prone river basins for forecasting one day ahead discharge at two respective river gauging stations. Upper Kalu River basin with smaller catchment area and smaller catchment response time did not show good forecasting results, but Kelani River catchment with comparatively large catchment area, with a response time of almost a day, showed good forecasting results. It was concluded that the GRBFN could be successfully applied to daily river stage/discharge forecasting for comparatively lager catchments with larger response time (close to a day). Data of smaller duration have to be selected for forecasting river stage/discharge for rivers with smaller catchments such as upper Kalu River basin. Additional input data such as discharge of upstream reservoirs could improve the forecasting accuracy of the model.

Key words: *Radial basis function networks, artificial neural networks, data clustering*

Introduction

It is obvious that timely flood warning can play a vital role in flood mitigation. In Sri Lanka, flood warning is still not based on sophisticated methods. The trend of the river stage is judged by continuous monitoring of river gauges. The current river stage, the trend noticed, the latest rainfall observations within the catchment, the rainfall forecasting made by the meteorological department, and the past experiences are used to issue flood warnings. Even though a reasonably accurate river flow model with reasonable lead time forecasting can effectively be used in flood warning systems, it is still a difficult task in Sri Lanka as efforts already taken to develop river flow models for flood prone rivers have failed due to lack of physical information. As a result of not using river flow models for issuing flood warnings, less accurate, less informative and delayed flood warnings are made most of the time. Therefore the main objective of this study is to develop a river discharge forecasting model at two river gauging stations situated in flood prone areas of the two most flood vulnerable river systems (Kalu River and Kelani River) in Sri Lanka as a supportive component of a reliable flood warning system of the two river systems. Out of many types of rainfall-runoff models, data driven models may be more suitable to Sri Lankan rivers due to two important reasons. The first is the lack of physical information and physical data of Sri Lankan catchments. The other is the availability of large amount of observed data such as rainfall and river flow. ANN models as a kind of data driven

* Irrigation Engineer, Department of Irrigation, Sri Lanka.
** Training and Research Advisor, International Centre for Water Hazard and Risk Management (ICHARM), Japan.

model, have recently become popular in many areas with complex input-output relationships. In the field of hydrology, it has been successfully applied by many researchers. Radial Basis Function Networks (RBFNs) are a class of feed forward neural networks that are used for classification problems, function approximation, noisy interpolation, and regularization. They have become increasingly popular in engineering applications due to their advantages over more widely used multilayer perceptron, namely faster convergence, smaller extrapolation errors, and higher reliability (Poggio & Girosi, 1990). Therefore in this research an attempt is made to establish a functional relationship between catchment rainfall and river discharge by using Radial Basis Function neural Networks.

Study Areas

Two river basins have been selected as study areas. They are Kalu and Kelani River basins of Sri Lanka. Kalu and Kelani river basins are the most flood vulnerable river basins in the country. Kalu River basin is the second largest river basin in Sri Lanka covering 2766 km^2 and most of the catchment is located in the area which receives highest annual rainfall in the country. The annual rainfall in the basin is averaged to 4000mm (Ampitiyawatta & Guo, 2009). Although the Kalu River has the second largest catchment in the country, it discharges the largest amount of water to the sea and it is in the order of 4000 million m^3 annually (Ampitiyawatta & Guo, 2009). The daily rainfall has sometimes recorded around 400mm in the area. The Kalu River originates from the central highlands in the wet zone at an altitude of 2250 m above Mean Sea Level (MSL), and runs through the western slopes and then through the western plains until falls out to the sea at Kalutara after traversing about 129 km. The basin has steep gradients in the upper basin area and mild gradients in the lower basin area.

Figure 1.1 Location map of Kalu River and Kalani River basins

Due to these hydrological and topographical characteristics of the river basin, its middle and lower flood plain starting from Ratnapura city suffers from frequent floods during the southwest monsoon season. The flood damages to the socio- economic profile are significantly high in areas such as Rathnapura city since the middle and lower flood plain of Kalu River is densely populated. In this study Rathnapura river gauging station is selected to model the river flow forecasting.

The Kelani River is a 145 km (90 miles) long river with a catchment area of 2,292km^2. It is ranked as the fourth longest river in Sri Lanka. The river is used for hydropower generation using three reservoirs situated in the upper catchment areas of the river. Depending on the operation of three reservoirs, the river flow varies around 25m^3/s in dry periods at Hanwella. The River flow variation during rainy seasons (monsoons) is from 800m^3/s to 1,500 m^3/s generally (Ministry of Irrigation and Water Resources, 2009). The river starts from Sri Pada Mountain situated in central hills. It flows through steep slopes from the start up to around 90km (close to Awissawella area) and then change to flow through a mild slope until it reaches the river mouth at the commercial capital Colombo. Because the river catchment is completely confined to the wet zone of the country, it gets high amount of annual rainfall close to around 3000mm on average. In the upper catchment area this figure sometimes goes up to around 4000mm per year. The area gets high amount of daily rainfall, especially in south-western monsoon season. These topographical and climatic situations cause frequent disastrous floods in its middle and lower reaches (from the middle reach city Yatiyantota up to Colombo). Floods are generally severe in lower reaches starting from Hanwella (situated around 27kms from the river mouth) to most downstream, the Colombo city. Kelani River floods receive highest attention as it causes floods in densely populated (highly urbanized) city Colombo. Therefore the economic damages caused by floods are generally very high. Hanwella river gauging station situated in the start of lower reaches of the river basin is selected to model the river flow. If the river flow at Hanwella could be predicted accurately, it can be effectively used in the river flood warning system in the lower reaches of Kelani River including Colombo city.

Data

Daily River discharge data for 14 years from 1982 to 1996 at Ratnapura river gauging station is used for training, validation and application of the model to Ratnapura gauging station. Daily rainfall data (from the year 1982 to 1996) of 5 rain gauging stations situated within the Kalu river upper catchment are used as input rainfall data of the model. They are Alupola, Palmadulla, Hapugastanna, Balangoda and Kuruwita rain gauging stations maintained by the Department of Meteorology, Sri Lanka.

Daily river discharge data of 16 years at Hanwella gauging station are used for training, validation and application of the model to Hanwella gauging station. In addition, river discharge data of the same period at two upstream river gauging stations are also used as inputs to the model. They are Glencourse (51kms upstream of the river mouth) and Kithulgala (90kms upstream of the river mouth). Daily rainfall data of 4 rainfall stations maintained by the Department of Meteorology is used for this study. They are Holombuwa, Uda Maliboda, Daraniyagala and Kithulgala.

Theory and Methodology

Radial Basis Function Neutral Networks are powerful techniques for interpolation in multidimensional space. This is based on Cover's theorem that is proved by him mathematically. Cover stated that a complex pattern-classification problem (or an interpolation problem) in multidimensional input space cast in high-dimensional space nonlinearly is more likely to be linearly separable than in a low dimensional space (Haykin, 1998). In radial basis function networks, input patterns are transferred to a higher dimensional space (which is hidden) than input space by using radial basis functions. In other words, RBFN have two layers of processing; in the first, input is mapped onto each radial basis function (RBF) in the 'hidden' layer which is the second layer. A RBF is a function which has built into a distance criterion with respect to a center. These centers can be select arbitrary or by clustering input patterns. The most popular and widely used type of radial basis functions are Gaussian functions. Radial basis functions have been applied in the area of neural networks where they may be used as a replacement for the sigmoid hidden layer transfer characteristic in Multi-Layer Perceptrons. In regression problems the output layer is then a linear combination of hidden layer values representing mean predicted output. The interpretation of this output layer value is the same as a regression model in statistics. The structure of the RBFN is shown in Figure 1.2.

In this study, two FORTRAN programs were written, one is for data clustering by using K-means clustering algorithm, the other for forecasting the output using the generalized radial basis function network (GRBFN) architecture. In the radial basis function network model, the input data are divided into 3 parts namely training data, validation data, and application data. Training input and relevant observed output data are used for learning (or training) the model. The solution surface obtained after training is used to predict the output for validation and application input data. The output given by the solution surface, were compared with the actual output (observed output) using performance indicators. The performance indicators used for output evaluation are Relative Root Mean Square Error (RRMSE), Coefficient of determination (CD), and R^2 value. The basic parameters of the network, the number of centers to the radial basis functions and the input data normalization range are selected by trial and error and the outputs obtained from the programs are evaluated using performance indicators. The parameters with the best performance indicators are selected for forecasting.

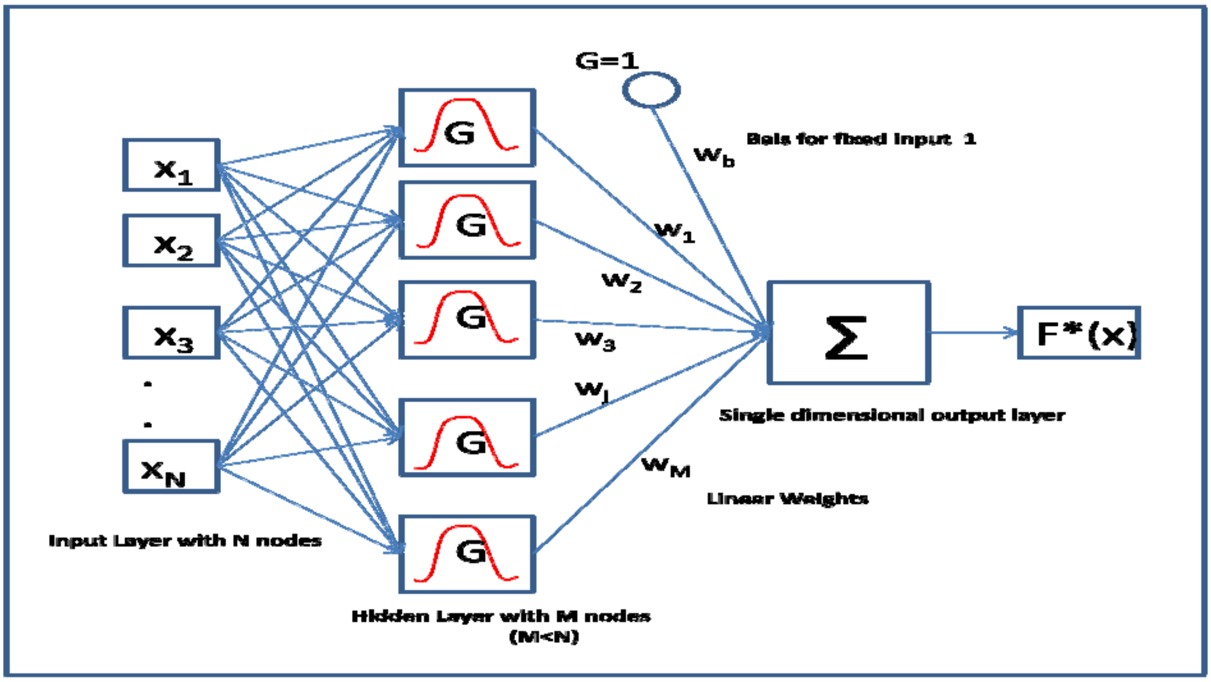

Figure 1.2 Structure of the GRBFN

Applications

First, for checking the validity, accuracy and applicability of the computer code developed, a generated series of data for known complex and nonlinear functions was applied to the model. Two Friedman's functions (Friedman's function 1 and 2) were used for this purpose. The results were evaluated by means of performance indicators explained above. By confirming the validity and the accuracy of the computer code developed, it was applied to forecast the river flow at Ratnapura and Hanwella river gauging stations. The results obtained from all three cases with the relevant values of performance indicators are presented in the Results and Discussions section.

Results and Discussions

The best results obtained from the application of the model to Friedman's function 1 are shown in Figure 1.3 and Table 1.1. The best results (measured by the performance indicators) were obtained by several trial runs of the model by changing parameters. The model gave the best results when the number of centers in the training input data was set to 25. Even though the numbers of centers for training input data have a significant effect on the results, very small effects (almost negligible) were noticed from the data normalization range. Similar results were obtained from the model application to the Friedman's function 2 also. These results confirm the validity and the accuracy of the computer code.

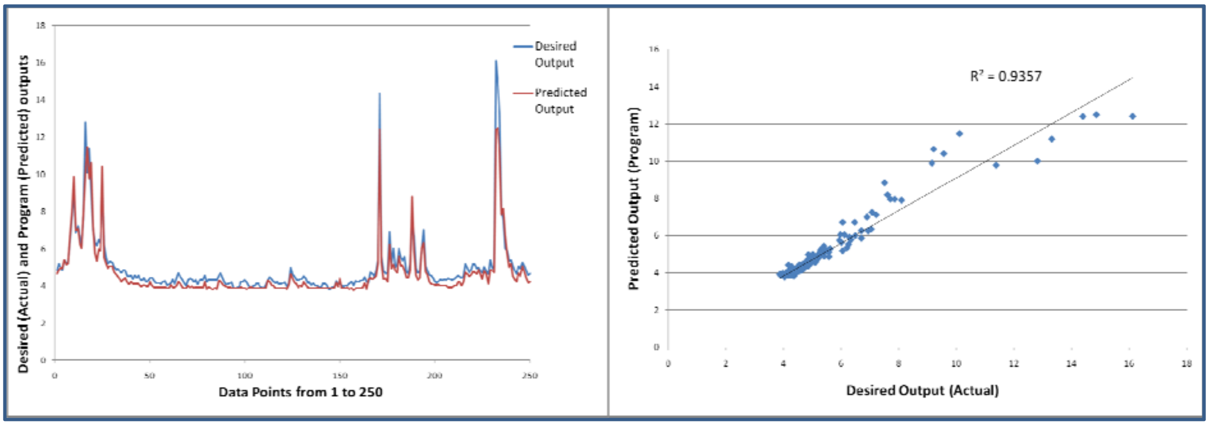

Figure 1.3 Time series plot and scatter plot of application results
for the Friedman's function 1.

Table 1.1 Values of Performance indicators for the best results of Friedman's function 1

Performance Indicator	Value
Reletive root mean square error (RRMSE)	0.16
Coefficient of Determination (CD)	1.13
Regression Coefficient (R^2 Value)	0.94

The best results of the model application to Ratnapura river gauging station in application are shown in Figure 1.4 and Table 1.2. Performance Indicators given in Table 1.2 show that the model is not good enough for one day ahead discharge prediction at Ratnapura river gauging station. The application results of the model to Hanwella river gauging station is shown in Figure 1.5 and Table 1.3. Values of Performance Indicators show that the model can reasonably well predict the daily discharge at Hanwella gauging station.

Table 1.2 Values of best Performance Indicators at model application to Ratnapura river gauging station

Number of centers selected for radial basis functions -25						
Data Normalization Range	Relative Root Mean Square Error (RRMSE)		Coefficient of Determination (CD)		R^2 Value	
	Training	Application	Training	Application	Training	Application
0 to 1	0.68	0.81	1.34	1.20	0.75	0.59

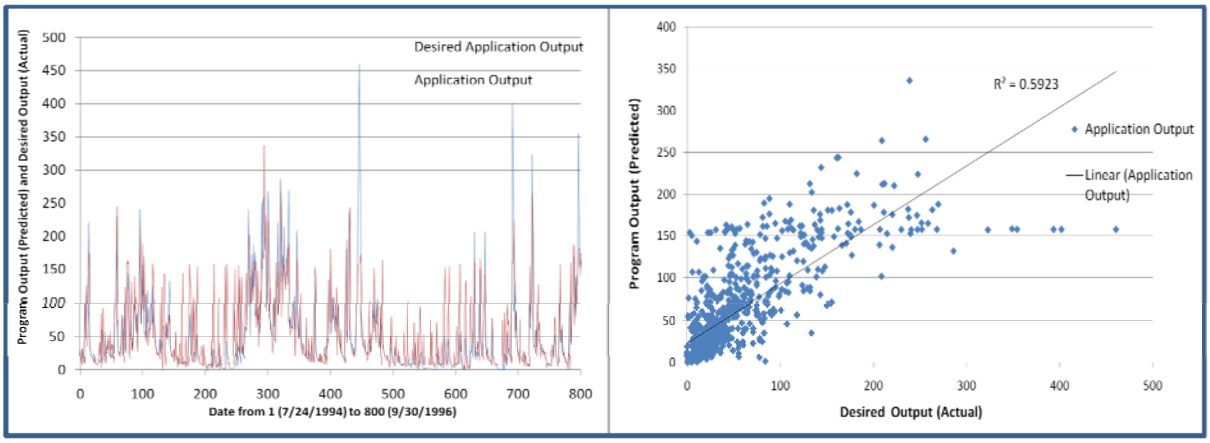

Figure 1.4 Time series and Scatter plots of model application results to Ratnapura river gauging station

Table 1.3 Values of Performance Indicators of the best results of the model application to Hanwella river gauging station

Number of Centers selected for radial basis functions - 25									
Data normalization range	Reletive Root Mean Square Error (RRMSE)			Coefficient of Determination (CD)			R^2 Value		
	Training	Validation	Application	Training	Validation	Application	Training	Validation	Application
0 to 1	0.37	0.59	0.57	1.14	1.31	1.37	0.87	0.76	0.76

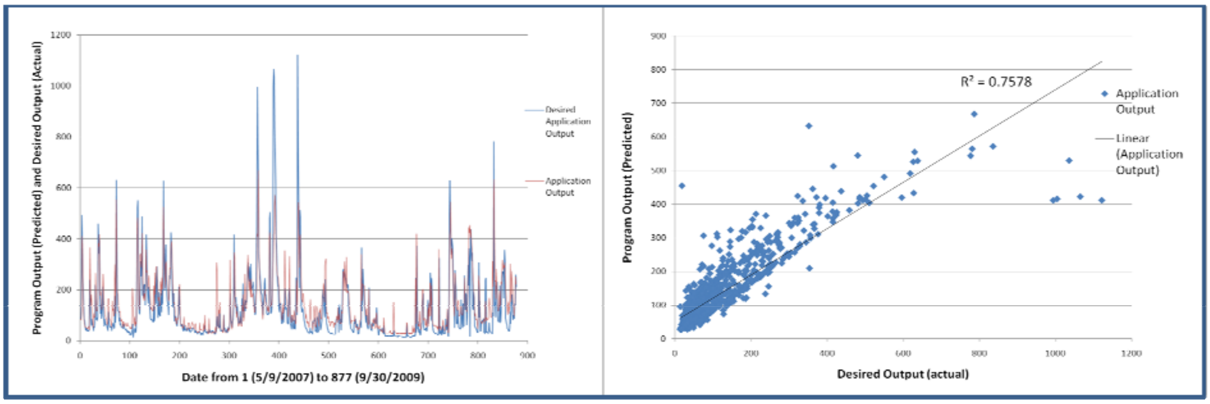

Figure 1.5 Time series and Scatter plots of model application results of Hanwella river gauging station at application

The catchment response times are estimated to be 9.5 hours at Ratnapura and 22 hours at Hanwella respectively. It is reasonable to assume that this is the main cause of the difference of model results of the two river gauging stations. In addition, daily river discharges are usually containing considerable amount of measuring and calculation errors. This fact may also contribute to the mismatch between actual observations and prediction results.

Conclusions

The data normalization range has no effect on the results of the model, and therefore can be kept constant. On the other hand, the model results were greatly dependent on the number of centers selected for the training input patterns at all 4 cases of model applications. Model results may be greatly improved if the model is used for water level prediction rather than daily river discharge prediction. If some shorter duration data could be used, model results may be improved reasonably at Ratnapura river gauging station.

Recommendations

Further studies are recommended to develop regular method for determining proper parameters (number of centers for the training input data) rather than trial and error method used in this study. It is also recommended to use the model for water level prediction rather than daily river discharge predictions.

Acknowledgement

I would like to express my solemn gratitude to my supervisor Prof. A. W. Jayawardena, Training and Research Advisor, ICHARM, for his overall supervision, invaluable suggestions and ardent encouragement in every aspect of this work.

References

Ampitiyawatta, A., & Guo, S. (2009). Precipitation trends in Kalu Ganga basin in Sri Lanka. *The Journal of Agricultural Science, Vol 4, No 1* , 10-18.

Haykin, S. (1998). *Neural Networks, A comprehensive foundation.* Ontario, Canada: Pearson Education.

Jayawardena, A.W, & Fernando, D. A.K. (1998). Use of Radial Basis Function Type Artificial Neural Networks for Runoff Simulation. *Computer-Aided Civil and Infrastructure Engineering, 13* , 91-99.

Ministry of Irrigation and Water Resources, S. L. (2009). *Proposed Kaluganga Multi-Purpose Development Project.* Irrigation Department, Water Resources and Project Planning Branch. Colombo, Sri Lanka: Ministry of Irrigation and Water Management.

Poggio, T., & Girosi, F. (1990). Regularization algoritms for learning that are equivalent to Multilayer Networks. *Science, New Series , Vol. 247*, 978-982.