IUTAM Symposium on the Dynamics of Extreme Events Influenced by Climate Change (2013)

Hydro-meteorological disasters: Causes, effects and mitigation measures with special reference to early warning with data driven approaches of forecasting

A.W. Jayawardena
Department of Civil Engineering, The University of Hong Kong, Hong Kong SAR, China, and, Research and Development Centre, Nippon Koei Co. Ltd (Engineering Consultants), Tsukuba, Japan

Abstract

In this paper, an attempt is made to highlight the causes, effects and mitigation measures of hydro-meteorological disasters with special reference to data driven approaches of forecasting. Recognizing the fact that the frequency of occurrence of water related disasters as well as the consequent damages including human casualties are on the increase in recent years, mitigation measures have become a high priority issue in all vulnerable countries. Structural measures taken by developed countries cannot be applied to developing countries because of the high capital investment. Non-structural measures such as early warning systems are more appealing to developing countries. One of the most important components of an early warning system is a mathematical model that links the input variables to the corresponding output variable. Several approaches of model formulation are discussed and some examples of the more recent fuzzy logic approach to flood forecasting is presented.

Key words: Hydro-meteorological disasters, physics-based models, conceptual models, data driven models, artificial neural networks, fuzzy logic approach

1. INTRODUCTION

Of the 3 main types of natural disasters in the world, geological, hydro-meteorological, and biological, hydro-meteorological disasters account for over 75% in terms of the damages including casualties, economic losses, infrastructure damage and disruption to normal life. They include floods, droughts, cyclones of all types, landslides, avalanches, heat waves, cold waves, and debris flow. Of the hydro-meteorological disasters, floods account for the majority of disasters followed by wind storms. Regionally, Asia suffers the most compared to other continents.

In recent years, flood disasters resulting from extreme rainfall have been on the increase in many regions of the world. In developed countries, the usual practice of mitigating flood disasters is by structural means which are unaffordable in most developing countries. The alternative then is to look for non-structural means that involve, among other things, early warning systems. They are cost effective and in some situations the only option.

The primary causes of all hydro-meteorological disasters are water and wind (風水). Precipitation, in many different forms at the upstream end leads to flooding when it is too high and droughts when it is too low. Wind systems caused by differential heating between the equator and the poles assisted by the Coriolis force lead to different forms of cyclones which have uncontrollable destructive power. Landslides and debris flow are triggered by rainfall whereas avalanches are triggered by excessive snowfall which is another form of precipitation. Heat waves are caused by stationary high pressure regions in the atmosphere which remain aloft for up to several weeks thereby trapping the heat instead of allowing it to lift. Cold waves occur when unusually cold and dense air near the surface in the high latitudes moves into the mid and lower latitudes. In addition to these primary causes, abnormal weather and climate patterns also cause natural disasters which many attribute to ‘climate change’.

1 Tel: +852 5413-5893
E-mail: hrecjaw@hkucc.hku.hk
The effects of all these disasters can be in many different forms. For example, frequent floods bring untold miseries, loss of lives, economic damages etc. in many parts of the world. Droughts on the other hand lead to crop failures, famines and diseases which may last for extended periods. Windstorms in the form of cyclones do not last for more than a few days but their consequences can be very destructive. The South-east Asian Region and some parts of USA are particularly vulnerable to windstorm disasters. Slides occur in mountainous areas due to slope failures caused by the weakening of soil cohesion when rainwater infiltrates into the sub-soil. Debris flows occur when movable solid particles in hilly areas are carried downslope by overland flow. In this paper the description is confined to flood disasters.

An early warning system is a set of procedures designed to protect human lives and minimize damages to be expected from a disaster which exceeds a certain critical level. It consists of a number of related and connected parts: forecasting, transformation of the forecast into a warning, transmission of the warning to local decision makers, conversion of the warning into remedial action. Forecasting of an impending event needs an understanding of the causes and effects in quantitative terms and formulation of a mathematical model that links the cause and the effect. The focus of this paper is how this can be achieved.

The basic technical components of an early warning system involves a measurable input data set that trigger a disaster, a measurable output data set that quantify the extent of the disaster and an appropriate mathematical model that transforms the input data set into a corresponding output data set. In the context of floods, the input that triggers a natural flood is the rainfall and the output is the runoff or discharge at a downstream point. There are many types of mathematical models that can be used to transform the input data into corresponding output data. They can be broadly classified into physics based, conceptual and data driven. In this paper, the emphasis is to highlight some of the recent developments in the latter type. In particular, the application of fuzzy logic systems to predict daily discharges in two rivers in two countries including the reliability and robustness of the approach are demonstrated.

2. PHYSICS-BASED APPROACH

In this approach, the starting point is the basic laws of physics: conservation of mass, momentum and energy. They can be described either by following a fixed mass of matter (water in this case) as it moves from a given point in space to another point leading to the Lagrangian approach, or by considering a control volume which would have inputs, outputs and changes in storages leading to the Eulerian approach. The latter is preferred as it leads to differential equations which will have only a single independent variable, or partial differential equations, which may have two or more independent variables. In the case of surface water flow, the governing equations are the St. Venant’s equations, or the shallow water equations, which for a one-dimensional physical domain take the following form:

\[
\frac{\partial q}{\partial x} + \frac{\partial h}{\partial t} = (i - f) = Q \quad \text{(Continuity Equation)} \tag{1}
\]

\[
\frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x} + g \frac{\partial h}{\partial x} + \frac{Qv}{h} = g(S_0 - S_f) \quad \text{(Momentum Equation)} \tag{2}
\]

where \(q\) is the discharge per unit width; \(h\) is the depth of flow; \(Q\) is the lateral inflow per unit length per unit width; \(v\) is the velocity of flow; \(S_0\) is the bed slope of the flow plane; \(S_f\) is the friction slope of the flow plane; \(i\) is the rainfall rate; \(f\) is the infiltration rate, and \(x, t\) are the distance along the flow plane, and time. These two equations in general have no exact mathematical solution. They are normally simplified by making certain assumptions. Two such simplifications lead to the diffusion wave equations and the kinematic wave equations which still have no analytical solutions.

Numerical solutions to Eq. 1 and 2, or their approximations can be obtained by using the finite difference method, the finite element method, or their combinations. Several examples of such numerical solutions of the shallow water equations for a one dimensional flow plane are available in the literature\(^{1,2,3,4}\). In order to seek numerical solutions to these or other partial differential equations, it is necessary to define a spatial domain which should be discretized.
Approximate solutions are then obtained in this domain by using an appropriate numerical method subject to applicable boundary and initial conditions.

Approaches based on physics lead to distributed type models which help to understand the processes that govern the phenomena. However, the downside is that the governing equations are usually over simplified, and very often the detailed data (input and output data and the applicable parameters) that would be needed for a realistic representation of the problem are unavailable. The parameters of such models which should have physical meanings and should be measurable are normally determined by a process of optimization thereby defeating the very purpose of adopting the physics-based approach. The advantages of the distributed approach are of a potential nature.

3. CONCEPTUAL APPROACH

In this approach, the principle is to assume some kind of concept, usually simple, to represent the real process behavior. Examples include the unit hydrograph theory⁵, Stanford Watershed Model⁶, Tank Model⁷, Xinanjiang Model⁸, HEC Model⁹ and VIC Model¹⁰ among several others. They can be lumped in which case there is no spatial variation of the inputs, outputs and parameters, or distributed in which case spatial variation of inputs, outputs and parameters can be accommodated, which in practice is rarely possible due to data restrictions. Parameters are estimated by some kind of optimization technique.

4. DATA DRIVEN APPROACHES

4.1 Historical review

There are many types of data driven models. The first rainfall-runoff model was probably the Rational Method¹¹ which relates the peak runoff to the rainfall intensity and the catchment area. Regression methods which take into consideration other influencing factors attempt to find a statistical relationship between the dependent runoff and the independent factors such as rainfall, antecedent rainfall and runoff and any other measurable variable. Regression methods can provide a time history of the runoff rather than the peak discharge only. The next significant development has been the introduction of the unit hydrograph theory⁵ which leads to a conceptual model that linearly relates the ‘rainfall excess’ to ‘direct runoff’ and which has stood the test of time due to its simplicity. Non-linear versions of the unit hydrograph¹² as well as many other versions that employ the systems theory approach have been suggested and used over the years¹³,¹⁴.

Considering the fact that the hydro-meteorological processes have some degree of randomness, models based on stochastic theory have led to many developments. Most hydro-meteorological and environmental data are measured at regular intervals of time and can therefore be represented as functions of time or Time Series. These, when observed over a long period of time, exhibit certain patterns which if identifiable can be used for forecasting purposes. Time series analysis involves the identification of such patterns or properties by a process of decomposition and subsequent extrapolation by synthesizing the decomposed components such that the statistical character of the generated series remains the same as that of the historical series. Stochastic methods coupled with Kalman Filtering can be used for real-time flood forecasting. The method has the advantage that the forecasting is by recursive equations which do not require heavy computer storage. Applications are too numerous to list and the reader may refer to a chapter on this topic in a recent book by the author¹⁵.

4.2 Recent advances

In recent years Artificial Neural Networks, or ANN’s have found applications in many areas of science and engineering. They emulate the brain which can be considered as a biological neural network. The processors operate on the data received via the connections. The transformation of an input to a corresponding output by a single neuron is relatively simple. The complexity arises as a result of the interactions of many neurons. They are more suited to theory weak data rich problems (Fig. 1)¹⁶.
There are several versions of artificial neural networks. They include multi-layer perceptrons (MLP’s), radial basis functions (RBF’s), recurrent neural networks, wavelet neural networks, and product unit neural networks. Among them, the multi-layer perceptron (MLP) is the widely used one followed by radial basis function (RBF) type. MLP’s have one input layer, one or more hidden layers and an output layer whereas the radial basis functions have only one hidden layer in addition to the input and output layers. The main problem in MLP function approximation is how to determine the number of nodes (neurons) in the hidden layer. Too few nodes may not model the process adequately and too many will require a long computational time as well as resulting in over-fitting. The objective of ANN’s is to model the signal, but over-fitting will fit into the noise as well producing a very good fit which lacks generalization properties when presented with unseen data. The optimal number of nodes in the hidden layer is determined by cross validation using different numbers of nodes which is a trial and error approach.

Radial basis functions constitute a global interpolation technique with good localization properties. Local radial basis function type models also have better approximating properties than local linear models. Radial basis functions can be Gaussian, multi-quadratic, inverse multi-quadratic and thin plate spline. Gaussian and inverse multi-quadratic are unbounded. Parameters of a radial basis function network include the number of basis points, the positions of basis points or centres, the receptive field widths, and the output layer weights which are linear and can therefore be estimated by the least-squares method. The performance of the RBF network depends not much on the non-linearity, but more on the choice of the centres which are sub-sets of the training data. After the selection of centres, RBF widths are estimated as some measure of the distance between cluster centres and the training data in that cluster. The distance (Euclidean) may be average or maximum.

Application of ANN techniques to hydrology started perhaps in the early 1980’s and within a short span of time, many researchers have produced a large number of publications (For a detailed list the reader is referred to a chapter on this topic in a recent book by the author).

There are several other data driven techniques that emerged in the recent years and which found applications in the flood forecasting problem. They include support vector machines (SVM’s), dynamical systems approach, genetic algorithms and genetic programming and fuzzy logic systems. In the remaining part of this paper, an attempt is made to highlight the background and application of fuzzy logic systems to flood forecasting.
4.3 Fuzzy logic approach

Traditional logic theory involves reasoning based on binary sets which have two valued logic, true or false, yes or no, zero or one. In real life, much of the information that we come across and process is not so crispy but involves some degree of fuzziness. The truth value may range between the completely true value and the completely false value, leading to a partial truth. The key idea in fuzzy systems is to allow a partial truth to prevail which can be numerically described by a specific function, referred to as a membership function that takes values between 0 and 1. For example, the discharge in a river may be perceived as high or low without a precise knowledge of the quantitative rate of flow. In other words, a quantitative description is translated into a qualitative linguistic description, and vice versa at a later stage. In this case the concept of ‘high’ or ‘low’ is subjective and context dependent. The mathematics of fuzzy set theory was introduced by Zadeh[17]. His idea was to replace the binary logic ‘yes/no’ by a five level classification of the form ‘definitely yes’, ‘probably yes’, ‘may be’, ‘probably no’, and ‘definitely no’. Fuzzy logic enables embedding uncertain or imprecise reasoning in everyday life to computers which operate in exact deterministic ways. Fuzzy logic models translate imprecise linguistic information sets into computer usable numerical language. However, they cannot learn well from the data. In general, since knowledge acquisition is difficult and the universe of discourse of each input variable needs to be divided into several intervals, fuzzy logic systems are restricted to fields where expert knowledge is available and the number of input variables is small.

The general structure of a fuzzy logic system is illustrated in Fig. 2. It consists of a knowledge base which includes a data base and a rule base, and 3 layers of information processing between the external input and output data. The main problems in building fuzzy systems include the selection of the relevant input and output variables, choice of the possible term sets for each linguistic variable, choice of the type of membership functions, fuzzification of the crisp input and output variables, derivation of the rule set, aggregation of the outcomes of the rules and de-fuzzification. It should be noted that the choice of membership functions is rather subjective but is not due to randomness. Membership functions can be triangular, trapezoidal, Gaussian, asymmetric Gaussian, generalized bell-shaped Gaussian, or Sigmoidal. The output of the aggregation process is a single fuzzy set for each output variable. The flow of information in a fuzzy system with 3 inputs and 3 rules is illustrated in Fig. 3.

![Fig. 2: Structure of a fuzzy logic system](image)
Fuzzy rules may be specified using the knowledge of experts directly and/or supplemented by available data, or may be not known explicitly but the variables are specified by experts, or, may have to be constructed purely from data. A fuzzy rule includes statement as 'IF-THEN' with two parts. The first part that starts with IF and ends before 'THEN' is referred to as the predicate (or, premise, or, antecedent), which combines in a harmonious manner the subsets of input variables. After the 'THEN' comes the consequent part, which includes the convenient fuzzy subset of the output based on the antecedent part. Sometimes, the input subsets within the antecedent part are combined, most often with the logical operator 'AND' whereas the rules are usually combined with the logical operator 'OR'. When the antecedent of a given rule has more than one part (such as for example, IF rainfall is high AND soil moisture is high THEN runoff is high), the fuzzy operator is applied to obtain one number that represents the result of the antecedent for that rule. This number is then applied to the output function. For the 'AND' operation, min (minimum) and prod (product) are supported by Matlab (Matlab Fuzzy Logic Toolbox). For the 'OR' operation, max (maximum) and the probor (probability) are supported. The activation of a rule is the deduction of the conclusion. The prod activation (multiplication) scales the membership functions, thus preserving the initial shape, rather than clipping them as the min activation does.

Interpretation of the IF-THEN rule consists of evaluating the antecedent after fuzzifying the input and applying fuzzy operators (when the antecedent consists of two or more parts), and then applying the result of the antecedent which should be a single number. The latter process is known as implication. In a multi-rule system, there could be several fuzzy outputs which should be combined to obtain a single output using very often the centroid method. It is also to be noted that the number of rules increases exponentially with increasing number of inputs leading to what is known as the 'curse of dimensionality'.

Fuzzy inference system (FIS) maps a given input to a corresponding output using fuzzy logic. It combines the components such as membership functions, fuzzy logic operators and rules. It can also be thought of as the rule evaluation of a fuzzy system. At the input stage, the input variables are mapped to appropriate membership functions. At the processing stage, the rules are invoked to generate outputs for each rule which are then combined in some manner to obtain an overall result for all the rules. At the output stage the combined result is converted to output values which become the end product. There are four well known inference mechanisms in fuzzy logic systems: Mamdani, Takagi-Sugeno-Kang (TSK), Tsukamoto, and Larsen. Of these, the widely used ones are the Mamdani type and the TSK type both of which are supported by Matlab Fuzzy Logic Toolbox.

4.4 Applications

In flood forecasting context, previous applications include, flood forecasting, rainfall-runoff modelling, hydrological time series modelling, amongst others. Many of the hydrological applications in the past use only one type of Fuzzy Inference System (FIS). This study aims to carry out a comparative analysis of two types of FIS’s using daily discharge data from two rivers having different climatological, geographical and land use characteristics.
4.4.1 Daily discharge prediction in Gin River, Sri Lanka

Daily discharge measurements at an upstream gauging station, a downstream gauging station and four rainfall measurements made in the catchment area for the periods 1997-1999 and 2004-2008 were used for model calibration while the corresponding data for the period 2000-2003 were used for validation with one- and two-day lead-times. The rainfall data of the four stations were averaged by the Thiessen polygon method. In this application, the two input variables were initially partitioned into three linguistic ranges, low, medium and high while the output variable was partitioned into five ranges, very low, low, medium, high and very high. Triangular and trapezoidal membership functions were used.

In general there could be $n^m$ rules where $n$ is the number of inputs and $m$ is the number of partitions. However, as the number of rules increases, the complexity of the formulation also increases leading to what is referred to as the ‘curse of dimensionality’. It is also possible that some rules may become superfluous. Therefore, the rule set was determined by a trial and error process from an initial value of 9 rules to a fine tuned value of 25 rules. The rule set consisting of 25 rules for this case, obtained by trial and error is as follows (RF is upstream rainfall; $Q_{u/s}$ is upstream discharge; $Q_{d/s}$ is downstream discharge):

1. IF $RF$ is very low AND $Q_{u/s}$ is very low THEN $Q_{d/s}$ is extremely low.
2. IF $RF$ is very low AND $Q_{u/s}$ is low THEN $Q_{d/s}$ is very low
3. IF $RF$ is very low AND $Q_{u/s}$ is medium THEN $Q_{d/s}$ is low
4. IF $RF$ is very low AND $Q_{u/s}$ is high THEN $Q_{d/s}$ is medium
5. IF $RF$ is very low AND $Q_{u/s}$ is very high THEN $Q_{d/s}$ is high
6. IF $RF$ is low AND $Q_{u/s}$ is very low THEN $Q_{d/s}$ is very low
7. IF $RF$ is low AND $Q_{u/s}$ is low THEN $Q_{d/s}$ is low
8. IF $RF$ is low AND $Q_{u/s}$ is medium THEN $Q_{d/s}$ is medium
9. IF $RF$ is low AND $Q_{u/s}$ is high THEN $Q_{d/s}$ is medium-high
10. IF $RF$ is low AND $Q_{u/s}$ is very high THEN $Q_{d/s}$ is very high
11. IF $RF$ is medium AND $Q_{u/s}$ is very low THEN $Q_{d/s}$ is low
12. IF $RF$ is medium AND $Q_{u/s}$ is low THEN $Q_{d/s}$ is low-medium
13. IF $RF$ is medium AND $Q_{u/s}$ is medium THEN $Q_{d/s}$ is medium
14. IF $RF$ is medium AND $Q_{u/s}$ is high THEN $Q_{d/s}$ is high
15. IF $RF$ is medium AND $Q_{u/s}$ is very high THEN $Q_{d/s}$ is very high
16. IF $RF$ is high AND $Q_{u/s}$ is very low THEN $Q_{d/s}$ is low-medium
17. IF $RF$ is high AND $Q_{u/s}$ is low THEN $Q_{d/s}$ is medium
18. IF $RF$ is high AND $Q_{u/s}$ is medium THEN $Q_{d/s}$ is medium-high
19. IF $RF$ is high AND $Q_{u/s}$ is high THEN $Q_{d/s}$ is very high
20. IF $RF$ is high AND $Q_{u/s}$ is very high THEN $Q_{d/s}$ is extremely high
21. IF $RF$ is very high AND $Q_{u/s}$ is very low THEN $Q_{d/s}$ is medium
22. IF $RF$ is very high AND $Q_{u/s}$ is low THEN $Q_{d/s}$ is medium-high
23. IF $RF$ is very high AND $Q_{u/s}$ is medium THEN $Q_{d/s}$ is high
24. IF $RF$ is very high AND $Q_{u/s}$ is high THEN $Q_{d/s}$ is very high
25. IF $RF$ is very high AND $Q_{u/s}$ is very high THEN $Q_{d/s}$ is extremely high

The performance indicators for the predictions using 9, 15 and 25 rules are shown in Table 1.

Table 1: Comparison of the performance indicator values for 9, 15 and 25 rules for the year 2008.
4.4.2 Daily discharge prediction in Fu River in China

A similar application using the Takagi-Sugeno-Kang (TSK) fuzzy inference system with hydrological data from the Fu River basin in east of Jiangxi province in China is presented next using daily discharge measurements made at upstream stations Liaojiawan across Fu River and Loujiachun across Linshui River, a tributary of Fu River, as inputs and the corresponding daily discharge measurements made at the downstream station Lijiadu across Fu River as outputs. The period of record for model calibration was the years 1960-1975 while that for validation was 1977-1979.

For a TSK fuzzy system with \( m \) inputs each of which has \( n \) partitions, the maximum number of fuzzy rules is \( n^m \). The number of inputs in this case is two thus requiring two cases with different partitions. With 3 partitions (low, medium, and high), there will be \( 3^2 = 9 \) rules. They take the form

\[
\text{IF } q_1 \text{ is } X \text{ AND } q_2 \text{ is } Y \Rightarrow Q = a_1q_1 + b_2q_2 + c_i
\]

where \( q_1 \) and \( q_2 \) respectively represent the upstream discharges at Liaojiawan and Loujiachun and \( Q \) represent the downstream discharge at Lijiadu; \( X \) and \( Y \) respectively indicate the fuzzy linguistic variables such as very low, low, medium, high, very high etc.; \( a, b, \) and \( c \) are the coefficients of the piece-wise linear functions fitted to the TSK consequents for each rule.

In this application, the partitioning was done by clustering the data. Cluster analysis is similar to principal component analysis and aims at reducing the dimensionality of the problem by grouping objects into subsets that have similar properties in the context of a particular problem. It is an iterative process that involves some trial and error. Popular methods include the \( k \)-means algorithm and the \( c \)-means algorithm. In this study, the centres of the antecedent fuzzy sets were determined by one dimensional clustering of the input-output space. It involves arranging \( N \) data points in a one dimensional space into \( k \) clusters where \( k \) is a user defined parameter. Two cases of clustering with 3 and 5 cluster centres corresponding to the peak values of the triangular and trapezoidal membership functions representing critical values for high, medium and low linguistic values for each input and very high, high, medium, low and very low respectively for the output were considered. The results of the two sets of predictions as measured by the 5 performance indicators (Table 2) show that the accuracy of prediction increases with increasing number of clusters. With the same number of cluster centres, the differences between different implication functions used to represent the fuzzy operators ‘AND’ and ‘OR’ is marginal.

### Table 2a: Performance indicators with 3 clustering centers

<table>
<thead>
<tr>
<th>Performance indicator</th>
<th>Minimum</th>
<th>Product</th>
<th>Optimum value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>Calibration</td>
<td>Verification</td>
<td>Calibration</td>
</tr>
<tr>
<td></td>
<td>84.176</td>
<td>84.495</td>
<td>84.882</td>
</tr>
<tr>
<td>RMSE</td>
<td>172.711</td>
<td>135.945</td>
<td>176.292</td>
</tr>
<tr>
<td>Performance indicator</td>
<td>Min and Max</td>
<td>Product and probability</td>
<td>Optimum value</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------</td>
<td>-------------------------</td>
<td>---------------</td>
</tr>
<tr>
<td></td>
<td>Calibration</td>
<td>Verification</td>
<td>Calibration</td>
</tr>
<tr>
<td>MAE</td>
<td>75.113</td>
<td>74.579</td>
<td>74.961</td>
</tr>
<tr>
<td>RMSE</td>
<td>148.517</td>
<td>133.085</td>
<td>148.028</td>
</tr>
<tr>
<td>RRMSE</td>
<td>0.375</td>
<td>0.413</td>
<td>0.374</td>
</tr>
<tr>
<td>EF</td>
<td>0.949</td>
<td>0.949</td>
<td>0.949</td>
</tr>
<tr>
<td>CD</td>
<td>1.054</td>
<td>1.032</td>
<td>1.053</td>
</tr>
<tr>
<td>R²</td>
<td>0.949</td>
<td>0.9509</td>
<td>0.9494</td>
</tr>
</tbody>
</table>

Table 2b: Performance indicators with 5 clustering centers

5. CONCLUDING REMARKS

Data driven approaches are not meant to replace other types of modelling but to supplement them. The fuzzy logic approach is particularly useful in situations where other methods are not feasible due to limitations in expertise available, and data resolution, quality, quantity and availability. In interpreting the results of this study it is important to bear in mind the fact that no model is perfect. All models depend on measured data for calibration. The reliability of measured data heavily contributes to the reliability of any model prediction. Discharge data are never measured on a long-term basis; they are estimated from stage measurements using rating curves. It would be far more practical if stage data are used for model development and application as they would be much easier to perceive than discharge data when it comes to issue early warnings of impending floods.

6. ACKNOWLEDGEMENTS

The last part of the study reported in this paper was carried out while the author was attached to the International Centre for Water Hazard and Risk Management (ICHARM) under the auspices of UNESCO and hosted by the Public Works Research Institute (PWRI) of Japan. The data used in the study as well as the numerical calculations were provided by graduate students J. D. Amarasekara and Zhu Bing who carried out their thesis studies under the supervision of the author.

7. REFERENCES