Stage–Discharge Relations for Low-Gradient Tidal Streams Using Data-Driven Models

Emad H. Habib[1] and Ehab A. Meselhe, M.ASCE[2]

Abstract: Development of stage–discharge relationships for coastal low-gradient streams is a challenging task. Such relationships are highly nonlinear, nonunique, and often exhibit multiple loops. Conventional parametric regression methods usually fail to model these relationships. Therefore, this study examines the utility of two data-driven computationally intensive modeling techniques namely, artificial neural networks and local nonparametric regression, to model such complex relationships. The results show an overall good performance of both modeling techniques. Both neural network and local regression models are able to predict and reproduce the stage–discharge multiple loops that are observed at the outlet of a 28.5 km^2 low-gradient subcatchment in southwestern Louisiana. However, the neural network model is characterized with higher prediction ability for most of the tested runoff events. In agreement with the physical characteristics of low-gradient streams, the results indicate the importance of including information about downstream and upstream water levels, in addition to water level at the prediction site.

DOI: 10.1061/(ASCE)0733-9429(2006)132:5(482)

CE Database subject headings: Neural networks; Regression models; Channel flow; Louisiana; Tidal effects; Streams; Water discharge; Coastal environment.

Introduction

Accurate information about rate of flow in rivers is important for a variety of hydrologic applications such as water resources planning and operation, water and sediment budget analyses, hydraulic and hydrologic modeling, and design of storage and conveyance structures. However, the logistics of collecting direct measurement of streamflow (discharge) on a continuous basis is costly and challenging, especially during large flood events. Therefore, it has been a common practice to convert records of water stages (which are easier and less expensive to measure) into discharges by using a pre-established stage–discharge relationship. In the literature, such relationships are often referred to as a rating curve. Unfortunately, the stage–discharge relationship is not always a simple unique relationship. This is due to the fact that in reality, discharge is not a function of stage alone. It also depends on water surface slope, channel geometry, bed roughness, and flow unsteadiness. In some situations, these factors may result in a nonunique relationship that is usually manifested as multiple loops in the observed stage–discharge measurements.

Rivers in coastal regions such as Louisiana are characterized by complex hydraulic and hydrologic characteristics. Streams such as the Vermilion River and its tributaries, for example, have a low gradient (typical slope is 1:10,000 for the main stem, and 1:1,000 for the tributaries) and often experience reverse flow during rainstorms and during periods of storm surge. For such streams, the concept of downstream and upstream is not well defined, especially when flow direction is often determined by a differential stage. The Vermilion River and its tributaries are strongly affected by the tidal signal and more importantly by the average water level in the Gulf of Mexico. For example, during strong southern wind events the Vermilion River becomes virtually a standing pool of water rather than a flowing river. Because of such strong backwater and tidal effects, the stage–discharge relationships for coastal streams are not unique and exhibit multiple looping curves. Fig. 1 shows a plot of pairs of discharge and stage (water level) measurements at the outlet of the Isaac–Verot low-gradient watershed (a subcatchment of the Vermilion River watershed; see Fig. 2). The plot shows multiple loops that have complex patterns. Similar complex relationships are also shown in Fread (1975) for different stations along the lower Mississippi River.

Two distinct approaches for modeling stage–discharge relations are available. The first approach is based on numerical solutions of the dynamic, unsteady, nonuniform flow equations (e.g., Fread 1973, 1975; Schmidt and Yen 2002). This approach will successfully mimic the observed stage–discharge relations if accurate information on boundary conditions and channel geometry are available. The second approach is data driven and is based on nonlinear regression techniques. This approach is the focus of the present study. Traditionally, simple least-squares regression techniques have been applied to establish a one-to-one relationship function between stage and discharge. A common form of this parametric function is the power relationship [discharge=a(stage)^b], where a and b=regression coefficients. For low-gradient streams that are subject to tidal and sea-level influence, such a relationship represents an oversimpli-

[1]Assistant Professor, Dept. of Civil Engineering, Univ. of Louisiana at Lafayette, P.O. Box 42291, Lafayette, LA 70504-2291 (corresponding author). E-mail: habib@louisiana.edu

[2]Associate Professor, Dept. of Civil Engineering, Univ. of Louisiana at Lafayette, P.O. Box 42291, Lafayette, LA 70504-2291. E-mail: meselhe@louisiana.edu

Note. Discussion open until October 1, 2006. Separate discussions must be submitted for individual papers. To extend the closing date by one month, a written request must be filed with the ASCE Managing Editor. The manuscript for this paper was submitted for review and possible publication on May 24, 2004; approved on July 8, 2005. This paper is part of the ***Journal of Hydraulic Engineering***, Vol. 132, No. 5, May 1, 2006. ©ASCE, ISSN 0733-9429/2006/5-482–492/$25.00.

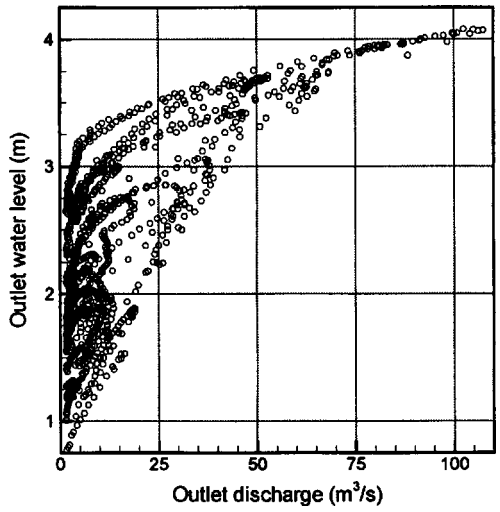

Fig. 1. Discharge and stage (water level) observations at outlet of low-gradient Isaac–Verot Coulee in southwestern Louisiana

fication of the complex physical processes and will result in significant errors in the estimated discharges. To account for these important factors, various input variables need to be included in development of the stage–discharge relationship. Such multivariate relations can be constructed using nonlinear regression techniques. However, this requires specification of a predefined parametric form of the relationship, which is not always possible. Preliminary analysis performed on the data shown in Fig. 1 has indicated that using such predefined relations fails to reproduce the nonuniqueness of these stage–discharge relationships.

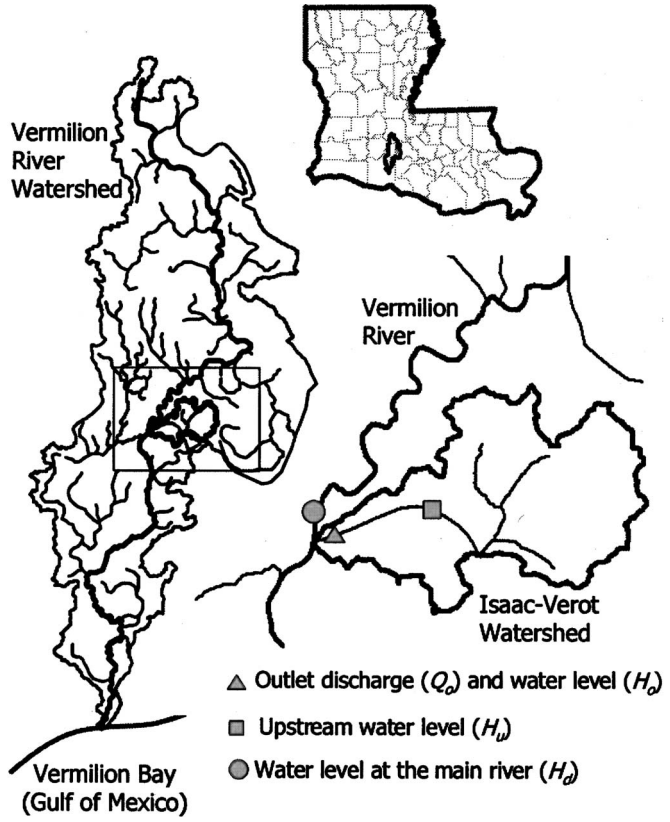

Fig. 2. Layout of Isaac–Verot watershed showing different discharge and stage gauges used in this study

An alternative approach to standard regression methods is based on data-driven, nonparametric, nonlinear models where no defined functional form of the rating curve is required. A popular example of such alternative models is the artificial neural network (ANN). The use of ANN models has received much attention in several water resources and hydrologic applications. Review of such applications is beyond the scope of this paper and the reader is referred to two comprehensive review studies (ASCE 2000a,b; Maier and Dandy 2000). Instead, only studies that are related to the problem of stage–discharge relationships are discussed herein. Despite the numerous studies on hydrologic applications of ANNs, only a few studies have addressed the problem of predicting hysteretic looping stage–discharge relations. Tawfik et al. (1997) used a neural network to model stage–discharge rating curves for two locations on the White Nile River and found ANN to be more accurate than three traditional techniques. However, their work was based on a simple fixed network structure and the results were validated using records that were included in training the network. The problem of setting stage–discharge relations using ANNs has also been addressed by Jain and Chalisgaonkar (2000), and Sudheer and Jain (2003). However, their analyses were based on idealized hypothetical looping rating curves rather than actual measurements. Supharatid (2003) developed a neural network model to forecast tidal-level variations at the mouth of a river in Thailand. The neural network model was then used to construct stage–discharge rating curves; however, no validation was given for the constructed curves. Bhattacharya and Solomatine (2000) also applied an ANN to model stage–discharge relations. However, discharge measurements from previous time steps were used as input to the network, which poses a limitation on the practical applicability of the developed ANN model.

Review of these previous studies indicated that the uniquely complex features observed in the multiloop stage–discharge relations of this study, as shown in Fig. 1, has yet to be addressed. The purpose of this paper is to examine the utility of data-driven computational techniques to model and predict such complex relations that are encountered in low-gradient coastal streams. In particular, the study will apply two modeling techniques: ANN and local nonparametric regression. To the best of the writers' knowledge, the latter technique has not been previously used to model stage–discharge relations. The study uses concurrent stage and discharge measurements for a low-gradient coastal watershed in southwestern Louisiana. In both modeling techniques, the model input is stage information at single or multiple sites in the watershed system, while the model output is the discharge at the stream outlet.

This paper is organized in the following manner. First, a description of the study site and the available stage and discharge measurements is given. Then, the two modeling methods are briefly explained. Afterward, the setup and the results of each model are presented. The paper discusses several related issues such as model generalization, data splitting, and selection of input variables. Assessment of the models' performance is carried out both graphically and statistically. The paper closes with conclusions and recommendations.

Study Site and Data

The outlet of Isaac–Verot Coulee was selected to demonstrate the complexity of stage–discharge relationships in low-gradient

coastal channels. The coulee, which is located in southwestern Louisiana, is a tributary of the Vermilion River (Fig. 2). The Vermilion River is approximately 108 km (67 mi.) long with a mouth on the Vermilion Bay of the Gulf of Mexico. The Isaac–Verot Coulee is the main channel that drains a 28.5 km^2 subcatchment. The coulee has a relatively low gradient of approximately 0.0008. According to Fread (1975), dynamic effects become significant if the channel bottom slope is less than 0.001. The outlet of the Isaac–Verot Coulee is approximately 60 km (37 mi.) north of the river's mouth. As shown in Fig. 2, there are two monitoring stations located on the Isaac–Verot Coulee: a continuous water level recorder at approximately 1,500 m upstream of its outlet to the river, and a continuous water level and discharge recorder near the outlet. The discharge measurements are taken by a bidirectional acoustic (side-looking) meter. Another continuous water level recorder is located on the main river upstream of the Isaac–Verot Coulee outlet.

Measurements from the coulee site are available for the period of October 2002 to November 2004. Preliminary examination of the discharge gauge measurements indicated that observations made with very low velocities might be subject to significant errors and therefore, any discharge observations less than 1.5 m^3/s are excluded from the analysis. This will also safeguard the stage–discharge samples from being dominated by such insignificant low flows. Throughout the observational period, the stage–discharge measurements at the various gauges have a recording frequency of either 15 or 30 min. After performing careful quality control processing of the data, and after excluding missing or suspect observations, the number of available concurrent stage and discharge measurements was 1,250. It should be noted that examining the frequency distribution (not shown) of these 1,250 records indicates that 75% of the records has discharge less than 20 m^3/s, 15% of the records has discharge between 20 and 50 m^3/s, and 10% of the data has discharge values greater than 50 m^3/s.

Methodology

Artificial Neural Networks

Artificial neural networks are mathematical models consisting of interconnected nodes that are capable of extracting complex nonlinear relationships and patterns in a set of input and output data. The network usually consists of several layers: one input layer, one or more hidden layers, and one output layer. Each layer is made up of several nodes that are connected to each other by links and a set of associated weights.

A common ANN model is the multilayer feedforward network (MLFF). Mathematical formulation of this and other ANN models are available in many references (e.g., Hagan et al. 1996) and only a brief description is given herein. The MLFF network has a unidirectional flow of information where the outputs from nodes in one layer are used as inputs to nodes in the next layer. The network is based on a supervised learning technique (also called training). The purpose of the training process is to find a set of optimal connection weights and thresholds that minimize a predetermined error function between the computed output and the target. Several algorithms are available for the process of network training. These algorithms vary in computational efficiency, rate of convergence, and storage and data size requirements. The mathematical and numerical details of these various algorithms are described in Hagen et al. (1996). A commonly used algorithm is back propagation (BP), which is a gradient descent algorithm that has been used in numerous ANN applications (see ASCE 2000a for a review). There are a number of variations on the basic BP algorithm that are based on optimization techniques such as conjugate gradient and Newton methods. In this study, the nonlinear least squares Levenberg–Marquardt optimization algorithm is adopted. This algorithm has been demonstrated to be more efficient than other BP algorithms (Hagan and Menhaj 1994). Furthermore, Supharatid (2003) illustrated the utility of the Levenberg–Marquardt algorithm for constructing stage–discharge relationships. This algorithm was found to have a faster convergence rate when compared to the standard gradient descent BP algorithm. The selection of a certain structure of the ANN model (i.e., number of hidden layers and nodes), and the choice of the input variables, are all critical factors for a successful application of ANN. These issues will be discussed and studied in further detail in the Model Identification and Development section.

Nonparametric Local Regression (Loess)

Local regression is an approach for fitting regression curves or surfaces to measurements of two or more variables by smoothing (Cleveland and Loader 1996). It is a nonparametric technique that, in contrast to parametric regression, does not require any assumptions about the specific functional model of the regression surface (Hardle 1990). A commonly used local regression method is a procedure known as the locally weighted least squares regression (Loess) (Cleveland 1979). Consider a sample of $i=1$ to n observations. In the sample, y_i represents a measurement of the response variable and x_i represents a corresponding vector of one or more independent variables. The underlying model for local regression is

$$y_i = f(x_i) + \varepsilon_i \tag{1}$$

where f=regression surface; and ε_i=random error. For each fitting point x, f takes the value of a parametric function fitted only to observations that are in a predefined neighborhood of x. The parametric function f is either a linear or a quadratic polynomial of degree p. The estimation of the parameters of the local polynomial function with $p=1$, for example, is done by minimizing a weighted least squares criterion

$$\sum_{i=1}^{n} W\left(\frac{x_i - x}{h}\right)\{y_i - [a_0 + a_1(x_i - x)]\}^2 \tag{2}$$

where W=weight function that gives a greater weight to the x_i points in the neighborhood that are closer to x, and h=bandwidth of the smoothing neighborhood. In the Loess fitting procedure, a tricube weight function is used

$$W(v) = \begin{cases} (1 - |v^3|)^3 & |v| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

In this application, $v=(x_i-x)/h$. Selection of this or other weight function forms is discussed in Cleveland and Loader (1996). It should be noted that the minimization process is performed locally at each individual point in the sample space.

The bandwidth, h, controls the smoothness of the fit. A simple choice is to make h constant for all values of x; however, this may lead to poor fits, especially in cases of sparse and nonuniform data density. An alternative choice is the nearest-neighbor bandwidth selection where h changes as a function of x based on the neighboring values of x_i. In this case, a prespecified smoothing parameter, α, which usually varies from zero to one, is multiplied

by n and rounded to an integer k. Then, the bandwidth $h(x)$ is taken as the distance from x to the kth closest x_i. In the Loess procedure, two parameters need to be specified: the smoothing parameter α, and the local polynomial degree p. In general, the criterion in choosing these two parameters is to produce a regression surface that is as smooth as possible but without missing important features and patterns that exist in the relationship between the response and independent variables. Details about selecting specific p and α values for this study are discussed in the "Model Identification and Development" section.

Model Generalization and Data Division

In any modeling approach, model testing using independent data is essential to evaluate its ability in providing accurate predictions beyond the data used in the training or fitting. This concept is usually referred to as the model's generalization ability. Split-sample is a common approach (McCuen 2005) for model testing where the available data are split into two subsets: a model identification set and a testing set. The first subset is used to determine the appropriate model structure, identify the input variables, and perform the model training (in the case of ANN) or fitting (in the case of Loess). The second subset is used to independently assess the model prediction accuracy. However, this approach suffers from some limitations (for a detailed discussion see Abrahart 2003; Bowden et al. 2002; McCuen 2005). Splitting the dataset into two parts will reduce the sample size used for model identification and optimization and, therefore, the accuracy of the estimated model parameters and coefficients. This may negatively impact the generalization ability of the developed model since data driven models are usually unable to extrapolate beyond the data ranges represented in the training sample. Such drawbacks are particularly relevant for the current study due to the relatively small sample size at high flow conditions.

In this study an approach known as the holdout method (Masters 1993; Maier and Dandy 2000) which avoids the limitations of the split-sample technique, is applied. In this method, small subsets of the data are identified and withheld in turn. Each withheld subset is reserved for independent testing. The remaining data are used for model training or fitting. The main advantage of this method is that, while maintaining independent testing samples, it maximizes the utilization of the available data and avoids the undesirable reduction in the sample size.

In the present study, the holdout approach is implemented on 12 significant runoff events that occurred over the observational period. Statistical description of these 12 events is summarized in Table 1. The observed peaks of these runoff events included one peak larger than 100 m^3/s, four peaks between 50 and 85 m^3/s, four peaks between 30 and 50 m^3/s, and the remaining events had peaks less than 20 m^3/s. Each of the 12 significant runoff events is withheld one at a time to independently test the model accuracy. For the case of Loess, the remaining data are used to fit the model and estimate its parameters. For the case of ANN, the remaining data are used to train the model and to compute and update the network weights. However, during training, and depending on the network complexity, running ANN for too many epochs (iterations) may cause the network to memorize the training sample without learning to generalize to new data patterns and relationships. Therefore, the remaining data are divided into two portions: a training set and a validation set. This approach is known as cross validation and is usually used to avoid the problem of overfitting in ANN models (Stone 1974; Bishop 1994). During the training process, the errors of both the training and the validation sets are computed and monitored. In the initial phases of training, errors of both the training and validations sets will continue to decrease with the increase of training iterations. However, when the network starts to overfit the training subset of data, the validation error will start to increase. At this stage, the training process is stopped and the current set of network weights and thresholds is retained and assumed to be the optimal parameters of the model.

Table 1. Statistical Description of Discharge Values Recorded during 12 Testing Runoff Events

Event	Date	Maximum (m^3/s)	Mean (m^3/s)	Standard deviation (m^3/s)
1	October 3–4, 2002	52.67	21.34	18.66
2	October 27–30, 2002	107.42	34.50	33.98
3	November 5–6, 2002	52.70	16.65	17.83
4	June 2–3, 2003	18.98	7.30	6.07
5	November 10, 2003	30.74	20.92	8.16
6	November 27–28, 2003	65.00	30.19	22.59
7	February 5–6, 2004	10.92	6.21	2.94
8	February 10–13, 2004	18.90	9.13	4.42
9	April 24–26, 2004	32.84	11.40	11.00
10	May 12–13, 2004	38.04	25.09	10.50
11	October 10–11, 2004	47.23	21.83	17.47
12	October 8–9, 2004	68.27	33.15	20.80

After finding the optimal set of the ANN's weights, and the optimal values of α and p for the Loess model, the withheld runoff event is used to assess the predictive ability of the models developed. This process is repeated for each of the 12 events and the performance of the developed ANN and Loess models is assessed both graphically and statistically.

Performance Measures

In any model development process, it is important to define the criteria by which the performance of the model and its prediction accuracy will be evaluated. Various statistical measures have been developed and used in the past. Following the findings of Legates and McCabe (1999), the current study will use various statistical measures to assess the model performance, namely, the root mean square error (RMSE), the coefficient of efficiency E, and the adjusted coefficient of efficiency E_1, defined as follows:

$$\text{RMSE}(\text{m}^3/\text{s}) = \sqrt{\frac{1}{n}\sum_i^n (Q_{\text{pi}} - Q_{\text{oi}})^2} \tag{4a}$$

$$\text{RMSE}(\%) = \frac{\sqrt{\frac{1}{n}\sum_i^n (Q_{\text{pi}} - Q_{\text{oi}})^2}}{\bar{Q}_o} \cdot 100 \tag{4b}$$

$$E(\%) = 1 - \frac{\sum_i^n (Q_{\text{pi}} - Q_{\text{oi}})^2}{\sum_i^n (Q_{\text{oi}} - \bar{Q}_o)^2} \tag{5a}$$

$$E_1(\%) = 1 - \frac{\sum_i^n |Q_{pi} - Q_{oi}|}{\sum_i^n |Q_{oi} - \bar{Q}_o|} \quad (5b)$$

where Q_{oi} and Q_{pi}=observed and predicted discharges at the outlet at record i, respectively; and n=size of the stage–discharge sample. The "overbar" in Eq. (5) denotes the mean of Q_o. The RMSE describes the average difference between model results and observations in units of the discharge variable [Eq. (4*a*)] and can be normalized to provide a relative measure with respect to the mean observed discharge [Eq. (4*b*)]. Physically, the coefficient of efficiency, E, measures the differences between the observations and predictions relative to the variability in the observed data itself. According to Eq. (5*a*), E may range from $-\infty$ to 1.0, where $E=1.0$ indicates a perfect model, $E=0.0$ indicates that the observed mean is as good a predictor as the model, and $E<0$ indicates that the model is worse than using the observed mean as a predictor. Legates and McCabe (1999) raised the issue of oversensitivity of E to extreme values (caused by squaring the difference terms), and therefore, introduced a modified coefficient of efficiency, E_1, which uses the absolute differences, rather than their squares. The combined use of RMSE, E, and E_1, will provide a sufficient assessment of each model's performance and will allow comparison of the accuracy of the two modeling approaches used in this study.

Model Identification and Development

ANN Structure and Selection of Input Variables

It should be noted that selecting a certain structure for an ANN, determining its parameters, and the choice of input variables, should be completed prior to the testing phase. In other words, results of the testing phase should not be used to guide the process of model identification and setup. In order to setup an ANN, it is necessary to determine the number of the hidden layers in the network and the number of nodes in each hidden layer. Determining an appropriate and optimal network design is an important but challenging task (ASCE 2000a; Maier and Dandy 2000). However, as demonstrated in several studies, an ANN with one hidden layer can be used to model real world functional relationships with unknown or poorly defined form and complexity. Therefore, a three-layer feed forward network (one input layer, one hidden layer, and one output layer) will be adopted in this study. The sufficiency of one hidden layer was confirmed for the present problem when additional hidden layers were added but did not result in significant improvements during the model development stage (results are not presented herein).

As discussed in ASCE (2000a), selection of the appropriate input variables for data-driven models, such as ANN and Loess, is not a trivial task. Discarding key input variables results in loss of information, while including unimportant variables leads to undesirable effects such as large network sizes, decrease in computational efficiency, and increase in the amount of data needed for model development. Several approaches have been developed to determine appropriate model elements (see Bowden et al. 2005 for a comprehensive review). In the present study, the input variables are selected using a combination of physical reasoning and a heuristic approach, which is discussed in Maier and Dandy (2000). In this approach, the number of input variables is increased gradually to assess their effect on the model performance.

An essential input variable in predicting the outlet discharge Q, is the water level at the outlet, H_o, which describes the local condition of the coulee outflow. However, there are other input variables in the watershed that should be considered as candidate variables in the ANN model. These variables include water levels at other gauge locations within the watershed (Fig. 2). The water level at the upstream stage station, H_u, is considered since it can indicate the slope of the water surface in the coulee. The water level at the Vermilion River, H_d, to which Isaac–Verot watershed drains, is also considered since it reflects the possible tidal and backwater effects from the Gulf of Mexico. Some studies have indicated that including information from previous time periods may improve the predictive ability of the models (e.g., Jain and Chalisgaonkar 2000). Therefore, water level records from up to four previous time steps are also considered as possible inputs.

For the present problem of predicting discharge–stage relations, the discharge at the outlet at time t_0 is predicted by a function with arguments that may include any combination of the different water level variables at present and past time steps

$$Q_p(t_0) = f[H_o(t_0), H_o(t_{-1}), \ldots, H_o(t_{-no}), H_u(t_0), H_u(t_{-1}), \ldots, H_u(t_{-nu}), H_d(t_0), H_d(t_{-1}), \ldots, H_d(t_{-nd})] + \varepsilon(t_0) \quad (6)$$

In this equation, $Q_p(t_0)$=predicted discharge at the outlet; $\varepsilon(t_0)$ represents the unknown model error term; and no, nu, and nd represent the unknown number of past inputs at the outlet, upstream, and downstream locations, respectively. The approach followed herein is to train the ANN using different combinations of the variables considered in Eq. (6) and to assess the prediction accuracy for the training and validation subsets. It should be noted that in this study, all of the considered combinations included the local water level $H_o(t_0)$. If a water level record from a past time step is considered in a certain combination, then all records from present up to that past step are included in the same combination. In this study, the maximum considered value for each of no, nu, and nd is four (i.e., input up to four past time steps is considered). Thus, this strategy makes the total number of possible combinations 180.

Another critical aspect in setting up an ANN is the choice of the number of nodes in the hidden layer. The more complex the function to be approximated, the larger the number of hidden nodes required. However, and as discussed in Maier and Dandy (2000), having too many free parameters in the model can result in overfitting. In this study, the number of nodes in the hidden layer is determined using the following approach. The model is trained for each of the 180 combinations by gradually increasing the number of hidden nodes starting from one node. The model

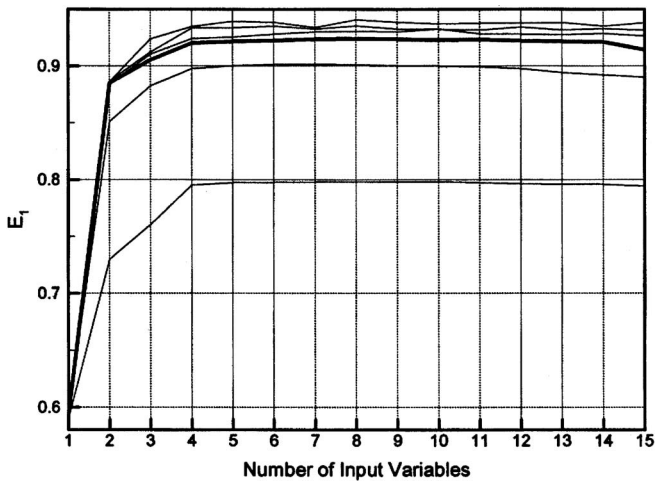

Fig. 3. Selection of input variables in neural network. Adjusted coefficient of efficiency (E_1) during validation subset of February 5–6, 2004 event is shown as function of number of input variables included in 180 combinations of present and past water level variables. Different lines (from bottom to top), represent cases with increasing number of hidden nodes (from 1 to 6). Bold line indicates case of three nodes.

accuracy during the training and validation stages is monitored to assess the significance of the added extra nodes. While this approach may become computationally demanding, it has been demonstrated (Hsu et al. 1995) that an acceptable solution to the model identification problem can be found.

An example of the results of this process is presented in Fig. 3. In this figure, the model prediction accuracy (represented by the coefficient E_1) over the validation subset of the February 5–6, 2004 runoff is plotted as a function of the number of input variables included in the 180 combinations. The results are also presented for different number of hidden nodes ranging from 1 to 6. It should be noted that different combinations exist for each number of input variables. For example, the case of two variables include the following three combinations $[H_o(t_0), H_u(t_0)]$, $[H_o(t_0), H_d(t_0)]$, and $[H_o(t_0), H_o(t_{-1})]$. However, the figure only shows the maximum E_1 value among all possible combinations at each total number of variables. In the figure, the minimum number of variables is 1 [which is $H_o(t_0)$], while the maximum number of variables is 15 (where each of no, nu, and nd is equal to 4). Examining Fig. 3 indicates that, in general, when more input variables are included, the model accuracy improves. The most significant improvement occurs when one additional variable is added. The model improvement continues with including a total of three and four variables. However, beyond four variables, no significant improvement, if any, is noticed. Similar observations were also evident for the other performance measures. Examining the results of the different combinations indicates that with two input variables, the best performance is reported when the water levels, $H_o(t_0)$ and $H_u(t_0)$, are selected as input variables. When adding an extra variable, it is found that the third most significant variable is $H_d(t_0)$. These results agree with the physical knowledge about this channel system where upstream and downstream conditions play a controlling factor in determining the magnitude of discharge at the outlet. The results also indicate that the fourth most significant variable is the upstream water level at a previous time step $[H_u(t_{-1})]$. It should be noted that these observations were consistent across the different tested number of hidden nodes with the exception of the case of two hidden nodes which showed that $H_u(t_{-1})$ and $H_o(t_{-1})$ give slightly better results when added as the third and fourth variables, respectively.

A similar argument can be used to determine the sufficient number of hidden nodes. As expected, the results indicate that a larger number of network nodes do not always lead to a significant reduction in the network errors. Fig. 3 shows a significant improvement in the model performance when the number of hidden nodes is increased from one to two. The improvement continues with the case of three nodes; however, the rate of improvement becomes much less when the number of nodes is increased beyond three. It should be noted that similar observations regarding the selection of input variables and the number of hidden nodes were also evident over the other runoff events.

Based on the above analyses, an ANN model that comprises one hidden layer with three nodes, and the input variables $[H_o(t_0), H_u(t_0), H_d(t_0), H_u(t_{-1})]$ is considered to be sufficient for the current problem. The values of the statistical measures obtained during the training and validation periods using this structure of ANN are summarized in Table 2. For each runoff event, the trained network is retained for later testing.

Table 2. Statistical Analysis of ANN Prediction Accuracy for 12 Testing Runoff Events

	Training sample				Validation sample				Testing sample			
	RMSE				RMSE				RMSE			
Event	(m^3/s)	(%)	E	E_1	(m^3/s)	(%)	E	E_1	(m^3/s)	(%)	E	E_1
1	1.54	10.4	0.99	0.93	1.65	11.0	0.99	0.93	4.12	19.3	0.95	0.81
2	1.66	13.7	0.99	0.90	1.79	14.6	0.99	0.90	3.20	9.3	0.99	0.93
3	1.71	11.0	0.99	0.93	1.75	11.9	0.99	0.92	2.31	13.9	0.98	0.90
4	1.89	11.7	0.99	0.91	1.72	11.2	0.99	0.92	1.26	17.3	0.96	0.82
5	1.70	11.2	0.99	0.92	1.80	11.8	0.99	0.93	1.14	5.5	0.98	0.87
6	1.70	11.6	0.99	0.92	1.77	12.3	0.99	0.92	2.26	7.5	0.99	0.91
7	1.75	11.4	0.99	0.92	1.80	11.0	0.99	0.92	0.43	6.9	0.98	0.86
8	1.94	11.9	0.99	0.92	1.64	10.9	0.99	0.93	0.75	8.2	0.97	0.84
9	1.71	11.2	0.99	0.93	1.89	11.9	0.99	0.92	0.75	6.5	0.99	0.94
10	1.78	11.9	0.99	0.92	1.72	11.4	0.99	0.92	1.67	6.6	0.97	0.86
11	1.59	10.5	0.99	0.93	1.43	9.6	1.00	0.93	4.88	22.3	0.92	0.79
12	1.69	11.5	0.99	0.92	1.75	12.3	0.99	0.91	4.17	12.6	0.96	0.80

Selection of Loess Parameters and Input Variables

Similar to the ANN, the number of input variables to the Loess regression model will be determined by examining different possible combinations of water level measurements at present and past time steps. However, the formulation of the Loess algorithm (Cleveland et al. 1992) allows for a maximum of four predictors (input variables) to be included. This limitation does not allow testing all the 180 combinations that were considered with the ANN model. Therefore, only possible combinations of four input variables or less will be considered. Similar to ANN, the local water level $H_o(t_0)$ is always included in each combination. To include a certain water level from a past time step, the present observation is also included in the same combination.

As discussed earlier, one needs to determine optimal values for the two parameters of the Loess model: the smoothing parameter, α, and the degree of the local fitting polynomial, p. A value of α that is too large may lead to oversmoothing (large bias); while a value that is too small may result in a noisy regression surface (large variance). To strike a balance between the bias and the variance, an optimal value of α needs to be determined. Such a value can be determined by optimizing a criterion that combines two competing measures: the closeness of model predictions to the observations, and the amount of smoothing in the model. While the first measure can be easily expressed by a measure such as RMSE, it is not quite clear how to quantify the degree of smoothness of the model. For this purpose, Cleveland et al. (1992) suggested a measure called the equivalent number of parameters, (NEP). Analogous to the number of parameters in a parametric fit, NEP provides a measure of the model smoothness, where its value decreases with the increase of α (mathematical details on the calculations of NEP are included in Cleveland et al. 1992).

An example of an optimization criterion that combines these two measures is the B information criterion (BIC) which has the following form:

$$\mathrm{BIC} = n \ln(\mathrm{RMSE}) + \mathrm{NEP} \ln(n) \tag{7}$$

where n=sample size of the fitting subset. In this function, RMSE will continuously decrease with the decrease of α, while NEP provides a penalty that increases with the decrease of α. A similar function to guide the selection among various fitting models of the rainfall-runoff process has been used (Hsu et al. 1995).

In search of an optimal value of the smoothing parameter, a series of possible values for α ranging from 0.01 to 0.9 with an increment of 0.01 is considered. Following the holdout approach, each of the 12 runoff events is withheld in turn, and the Loess procedure is applied on the remaining fitting subset over the full range of values of α. This process is implemented for each runoff event, for each of the input variable combinations, and for the two cases of $p=1$ and 2. For each fitting implementation, the model output $[Q_p(t_0)]$ is compared to the corresponding observation, and the statistics, RMSE, E, E_1, and BIC are computed for the fitting subset. An example of this analysis is shown in Fig. 4 for the runoff event on February 5–6, 2004 using the combination of $[H_o(t_0), H_d(t_0), H_u(t_0), H_u(t_{-1})]$ as input variables. The figure shows that the RMSE (E_1) of the fitting period continuously decreases (increases) with the decrease of α. As expected, a smaller value of α enables the model to closely follow and reproduce each point in the fitting sample. The RMSE and E_1 statistics also indicate better fitting with the case of $p=2$ than with $p=1$. However, BIC shows a different behavior where it initially decreases with the increase of α, but it starts to increase at larger values of α. The minimum value of BIC will determine the optimal value of α for each case of p. The results displayed in Fig. 4 indicate that an optimal value of α is less defined for the case of $p=2$ compared to the case of $p=1$. Despite not being well defined, the optimal value of α was significantly lower for $p=1$ than that obtained with $p=2$. This indicates that less smoothing is needed when a local fitting polynomial of degree 1 was used. It is also noticed that increasing the value of p from 1 to 2 has not resulted in any reduction of the minimum BIC value. This observation, which is consistent for all of the fitted runoff events, indicates that the penalty imposed on the model for using a higher degree of the local polynomial function outweighs the improvement in the fitting quality.

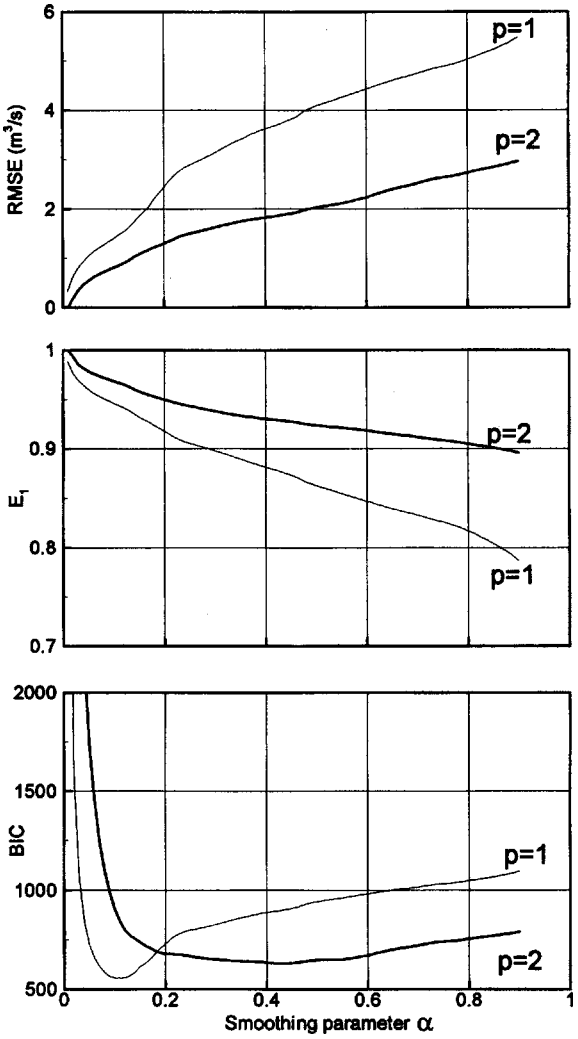

Fig. 4. Selection of optimal values of smoothing parameter, α, and degree of local polynomial, p, in Loess model. Performance measures during fitting subset of February 5–6, 2004 event are shown as function of α and p.

Examining the results obtained using the different combinations of input variables indicated that, similar to ANN, the best fitting quality over the fitting subsets is obtained when the four water levels $H_o(t_0)$, $H_d(t_0)$, $H_u(t_0)$, and $H_u(t_{-1})$ were used as input variables. These observations regarding the selection of α and the choice of the input variables were also confirmed for all other runoff events. The statistics obtained with each of the runoff

Table 3. Statistical Analysis of Loess Prediction Accuracy for 12 Testing Runoff Events

	Fitting sample						Testing sample			
		RMSE					RMSE			
Event	α	(m^3/s)	(%)	E	E_1	BIC	(m^3/s)	(%)	E	E_1
1	0.10	1.33	8.9	0.99	0.95	526	6.38	29.9	0.88	0.74
2	0.08	1.05	8.6	0.99	0.95	470	5.95	17.2	0.97	0.87
3	0.10	1.35	8.9	0.99	0.95	542	2.86	17.2	0.97	0.86
4	0.12	1.51	9.6	0.99	0.94	562	0.77	10.5	0.98	0.90
5	0.11	1.46	9.6	0.99	0.94	561	0.96	4.6	0.99	0.89
6	0.10	1.39	9.5	0.99	0.94	555	3.23	10.7	0.98	0.87
7	0.10	1.39	8.8	0.99	0.95	556	0.61	9.8	0.96	0.81
8	0.12	1.52	9.7	0.99	0.94	557	0.81	8.9	0.97	0.85
9	0.14	2.00	12.8	0.99	0.93	630	0.83	7.3	0.99	0.94
10	0.10	1.37	9.1	0.99	0.95	552	3.37	13.4	0.89	0.73
11	0.10	1.35	9.0	0.99	0.95	539	3.42	15.7	0.96	0.85
12	0.10	1.32	9.2	0.99	0.95	526	4.09	12.3	0.96	0.83

events, along with the corresponding optimal values of α, are summarized in Table 3. For each runoff event, the optimal value of α is retained for later testing of the model.

Independent Testing of ANN and Loess Models

Having selected the structure and the appropriate input variables, the ANN model is now ready to be tested using independent data. The testing procedure is performed on each of the 12 runoff events identified earlier. After withholding each event from the record, and after the training process is completed, the resultant ANN model is used to predict the discharge $Q(t_0)$ for the withheld event. This procedure is repeated for each event and the model prediction accuracy is evaluated both graphically and statistically. Comparisons of model predictions versus observations are presented in Figs. 5 and 6 for the 12 tested events. The scatter plot shown in Fig. 5 shows an overall good agreement between the observations and ANN predictions. A close examination of the plotted observed–predicted discharge pairs shows that the model slightly underestimates some of the observed extreme discharge values. This is possibly attributed to the paucity of high discharge records left in the remaining data subsets after withholding each of the testing events. It is also noticed that there is a higher degree of scatter at intermediate flow values. Examining the tested storms indicated that such values are mostly located on the rising limb of the hydrographs. The number of rising-limb records in the sample (in comparison to the falling limb records) is relatively limited mainly due to the steep rate of the rising limbs in most of these storms. Therefore, it is expected that the limited number of such records would eventually impact the training process and the resultant accuracy of the model. Finally, a comparison of the predicted and observed *Q-H* relationships (shown in Fig. 6) indicates that the model is generally able to reproduce such a complex relationship.

The overall good agreement between model predictions and the corresponding observations is also quantified by using the performance measures, RMSE, E, and E_1. The results of the statistical evaluation are summarized in Table 2 for each of the 12 storms. According to the table, the model RMSE is less than

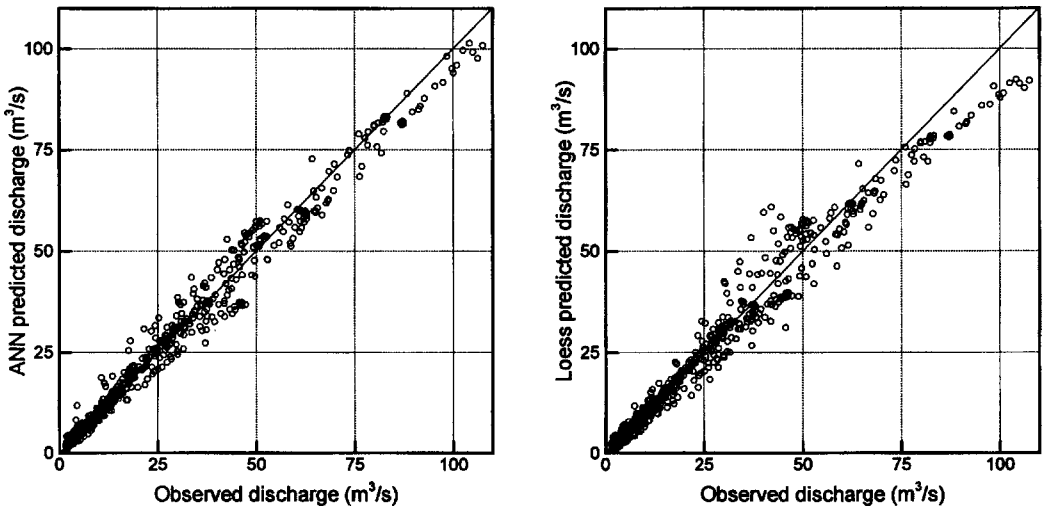

Fig. 5. Scatter plots of model predictions versus observations for 12 testing runoff events: (left panel) ANN results with one hidden layer and three hidden nodes; (right panel) Loess model using $p=1$ and optimal values of smoothing parameter α for each runoff event. Input variables are $H_o(t_0)$, $H_u(t_0)$, $H_d(t_0)$, and $H_u(t_{-1})$.

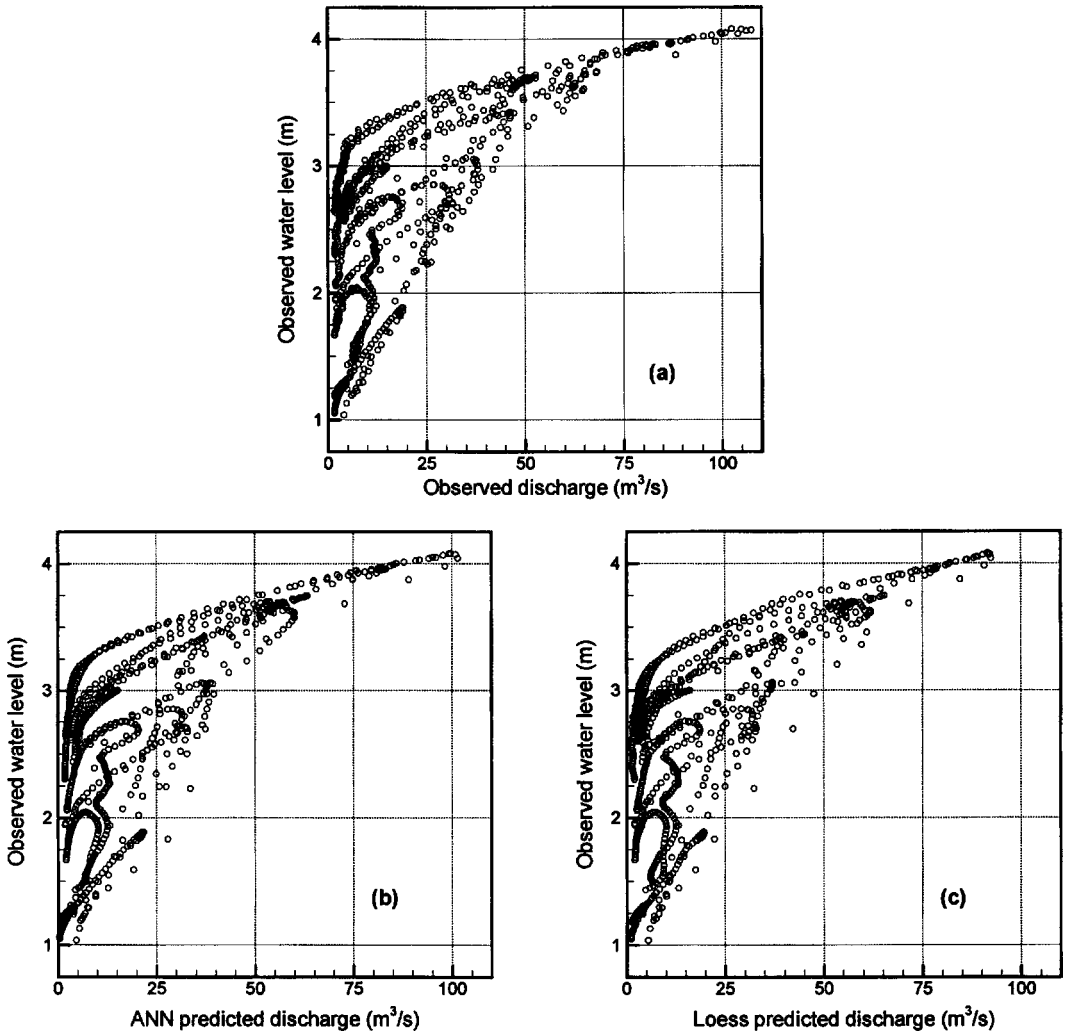

Fig. 6. Stage–discharge relations for 12 testing runoff events: (a) observed relation; (b) ANN model results; and (c) Loess model results. Model parameters and input variables are same as those described in Fig. 5.

2 m^3/s for most of the tested events and exceeds 4 m^3/s for two events only. When expressed as a percentage of the average event discharge, the RMSE is mostly less than 10% and exceeds 15% for three events only. Similarly, most of the events have a coefficient of efficiency E higher than 0.95. The adjusted coefficient, E_1, is mostly in the range of 0.85–0.95 with only few values between 0.80 and 0.85.

The same testing analysis is repeated for the Loess model. For each of the 12 runoff events, the Loess fitting model is applied using the optimal smoothing parameter α obtained during the model fitting stage (see Table 3), and with $p=1$. The Loess model results are compared to the observations as shown in Figs. 5 and 6. The scatter plot of observed versus predicted discharges, and the modeled stage–discharge relationship, indicate a reasonable performance of the Loess model. However, compared to the ANN model, it appears that the Loess model is less successful in providing accurate discharge predictions over the independent testing events. A significant degree of scatter is noticed at intermediate discharge values. Furthermore, the diagreement at extreme discharges is more pronounced than with the ANN model. It is likely that the limited number of rising-limb and high flows in the fitting subsets has significantly impacted the generalization ability of the local fitting model. The relatively reduced accuracy of the Loess model results is also reflected in the computed statistical measures as shown in Table 3 where, for most of the events, the RMSE (E) values are higher (lower) than those of the ANN model.

Discussion and Conclusions

Stage–discharge relationships for coastal low-gradient streams are characterized by complex patterns and multiple loops. Development of rating curves for such streams is challenging. Traditional methods, such as parametric regression, usually fail to model these relationships. The main objective of this paper is to identify appropriate data-driven methodologies capable of accurately relating discharge to observed water levels in coastal low-gradient streams. The study investigated the application of techniques such as artificial neural networks and local nonparametric regression as possible modeling tools. Both models were able to reproduce the complex nonlinear Q-H relationship and its multiple loops that

are observed at the outlet of a low-gradient subcatchment in southwestern Louisiana.

The results indicated that discharge in low-gradient streams cannot be predicted using the local water level alone, and that information about the surrounding water level conditions is needed. The downstream water level, which captures the backwater effects, is essential for accurate prediction by both models. The models' prediction accuracy was also increased significantly when water level measurements from an upstream location were included. It was also found that including information from one previous time step lead to significant improvements in predicting the stage–discharge relations. Findings of this study emphasize the concept that adding more complexity to the model structure (e.g., adding more hidden nodes in the ANN or using a higher degree of the local polynomial function in Loess) cannot compensate for including the physically necessary variables that control the hydraulic behavior of low-gradient systems.

Graphical and quantitative assessment of the results of both the ANN and Loess techniques indicate good agreement with the observed discharge measurements. ANN was characterized with better generalization ability for most of the tested events. Due to the relatively limited sample size and its skewed distribution, the Loess model suffered from a reduced accuracy in predicting extreme discharge values. Close examination of the results indicates that the two models have less accuracy in predicting discharges located on the rising limbs of the hydrographs (with Loess being less accurate than ANN). This can be attributed to the relatively few number of observations on the steeply rising parts of the hydrographs.

It was observed that the Loess approach suffers from some difficulties in selecting optimal values of its parameters (smoothing parameter and degree of local polynomial). This limitation indicates that local regression methods are more data demanding and require fairly good information about the local structures present in the data.

The writers believe that as more field measurements encompassing a wider range of flow conditions become available, the models' performance will improve. Future research will focus on the investigation of issues such as the effect of sample size and the use of other optimal methods of data splitting (e.g., Bowden et al. 2002).

As discussed above, the complexity of low-gradient systems indicates the need for continuous monitoring of their different hydraulic, hydrologic, and meteorological variables. It is only through high quality and high resolution observational setups that the underlying processes of such complex systems can be fully understood and modeled. To address such needs, the writers recently established an automated experimental network that provides continuous water level, discharge, precipitation, and other meteorological information in the coastal low-gradient Isaac–Verot subcatchment and its drainage channels. Such information is essential for further development of the data-driven models, as well as physically based approaches such as hydraulic and hydrologic numerical modeling.

Acknowledgments

This study was partially supported by the Army Research Office Grant No. DAAD19-00-1-0413 under the supervision of Dr. Russell Harmon, and by the Louisiana Transportation Research Center, Project No. 736-99-0918 under the supervision of Dr. Chester Wilmot. The writers would like to thank Ms. Tzvetlena Peeva and Ms. Sarada Kalikivaya for their help in data collection and developing some of the computer codes used in the study.

References

Abrahart, R. J. (2003). "Neural network rainfall-runoff forecasting based on continuous resampling." *J. Hydroinformatics*, 5, 51–61.

American Society of Civil Engineers (ASCE) Task Committee on Artificial Neural Networks in Hydrology. (2000a). "Artificial neural networks in hydrology. I: Preliminary concepts." *J. Hydrologic Eng.*, 5(2), 115–123.

American Society of Civil Engineers (ASCE) Task Committee on Artificial Neural Networks in Hydrology. (2000b). "Artificial neural networks in hydrology. II: Hydrologic applications." *J. Hydrologic Eng.*, 5(2), 124–137.

Bhattacharya, B., and Solomatine, D. P. (2000). "Application of artificial neural network in stage-discharge relationship." *Proc., 4th Int. Conf. on Hydroinformatics*, IAHR, Iowa City, Iowa.

Bishop, C. M. (1994). "Neural networks and their applications." *Rev. Sci. Instrum.*, 65, 1803–1832.

Bowden, G. J., Dandy, G. C., and Maier, H. R. (2005). "Input determination for neural network models in water resources applications. Part 1—Background and methodology." *J. Hydrol.*, 302, 75–92.

Bowden, G. J., Maier, H. R., and Dandy, G. C. (2002). "Optimal division of data for neural network models in water resources applications." *Water Resour. Res.*, 38(2), 2.1–2.11.

Cleveland, W. S. (1979). "Robust locally weighted regression and smoothing scatter plots." *J. Am. Stat. Assoc.*, 74, 829–836.

Cleveland, W. S., Grosse, E., and Shyu, W. M. (1992). "Local regression models." *Statistical models in S*, J. M. Chambers and T. Hastie, eds., Chapman and Hall, New York, 309–376.

Cleveland, W. S., and Loader, C. L. (1996). "Smoothing by local regression: Principles and methods." *Statistical theory and computational aspects of smoothing*, W. Haerdle and M. G. Schimek, eds., Springer, New York, 10–49.

Fread, D. L. (1973). "A dynamic model of stage-discharge relations affected by changing discharge." *NOAA Tech. Memo. NWS HYDRO-16*, National Weather Service, Silver Spring, Md.

Fread, D. L. (1975). "Computation of stage-discharge relationships affected by unsteady flow." *Water Resour. Bull.*, 11(2), 213–228.

Hagan, M., Demuth, H., and Beale, M. (1996). *Neural network design*, PWS, Boston.

Hagan, M. T., and Menhaj, M. B. (1994). "Training feedforward networks with the Marquardt algorithm." *IEEE Trans. Neural Netw.*, 5(6), 989–993.

Hardle, W. (1990). *Applied nonparametric regression*, Cambridge University Press, Cambridge, U.K.

Hsu, K.-L., Gupta, H. V., and Sorooshian S. (1995). "Artificial neural network modeling of the rainfall-runoff process." *Water Resour. Res.*, 31(10), 2517–2530.

Jain, S. K., and Chalisgaonkar, D. (2000). "Setting up stage-discharge relations using ANN." *J. Hydrologic Eng.*, 5(4), 428–433.

Legates, D. R., and McCabe, G. J. (1999). "Evaluating the use of goodness-of-fit measures in hydrologic and hydroclimatic model validation." *Water Resour. Res.*, 35(1), 233–241.

Maier, H. R., and Dandy, G. C. (2000). "Neural networks for the prediction and forecasting of water resources variables: A review of modeling issues and applications." *Environmental Modelling and Software*, 15(1), 101–124.

Masters, T. (1993). *Practical neural network recipes in C++*, Academic, New York.

McCuen, R. H. (2005). "Accuracy assessment of peak discharge models." *J. Hydrologic Eng.*, 10(1), 16–22.

Schmidt, A. R., and Yen, B. C. (2002). "Stage-discharge ratings revisited." *Hydraulic Measurements and Experimental Methods, Proc., EWRI and IAHR Joint Conf.*, (CD-Rom), T. L. Wahl, C. A. Pugh,

K. A. Oberg, and T. B. Vermeyen, eds., Estes Park, Colo.
Stone, M. (1974). "Cross-validatory choice and assessment of statistical predictions." *J. R. Stat. Soc. Ser. B. Methodol.*, 36, 111–147.
Sudheer, K. P., and Jain, S. K. (2003). "Radial basis function neural network for modeling rating curves." *J. Hydrologic Eng.*, 8(3), 161–164.
Supharatid, S. (2003). "Application of a neural network model in establishing a stage-discharge relationship for a tidal river." *Hydrolog. Process.*, 17(15), 3085–3099.
Tawfik, M., Ibrahim, A., and Fahmy, H. (1997). "Hysteresis sensitive neural network for modeling rating curves." *J. Comput. Civ. Eng.*, 11(3), 206–211.