ERDC/CHL CHETN-IV-58
March 2003

US Army Corps of Engineers®

Application of an Artificial Neural Network to Predict Tidal Currents in an Inlet

by Wenrui Huang and Catherine Murray

PURPOSE: The Coastal and Hydraulics Engineering Technical Note (CHETN) herein describes application of an Artificial Neural Network (ANN) that can be trained to predict currents at an inlet located within a larger regional system, based on water level measurements at a different and possibly distant location. Once developed, ANNs reduce the need for field gauging, and information may be hindcast for sites where data do not exist, or which have gaps in the historical record (Hsieh and Pratt 2001).

BACKGROUND: This CHETN describes an ANN developed to predict the horizontal current profile across Shinnecock Inlet, located on the south shore of Long Island, NY (Figure 1). The Regional Neural Network to Predict the Horizontal Velocity Profile across Coastal Inlets, or RNN_V, is a Matlab © program created to model the current at up to 9 points across an inlet within a larger regional system. For the Shinnecock Inlet application, the ANN was trained and verified based upon water level data from a remote monitoring station administered by the National Oceanic and Atmospheric Administration (NOAA) at Montauk (the input data), located approximately 65 km east of the inlet. The procedure applied current measurements, the target data, obtained from an Acoustic Doppler Current Profiler (ADCP) installed at Shinnecock Inlet (Figure 2) by the Coastal Inlets Research Program (CIRP). The network parameters were then saved and the model applied to predict the current profile across the inlet for a specific time period based solely on water level data for the same time period from the tide station at Montauk.

The RNN_V model has potential for many regional sediment management (RSM) applications. The correlation between measured currents at Shinnecock Inlet and the RNN_V model predictions is high. Linear correlation coefficient, r, values range from 0.80 to 0.98, with the majority falling between 0.90 and 0.96 for the nine monitoring stations across the inlet for all time periods for which the model was tested (the time intervals tested were March, May through December 2000; April 2000 was used to train the model). The root mean square errors (rmse) for the stations across the inlet ranged from 0.04 to 0.11 m/sec.

The RNN_V model may be applicable in the management of inlets located within a larger coastal region of interest. Within larger regions, many management actions such as dredging, sand bypassing, breach-contingency plans, and protection of beaches vulnerable to erosion must be integrated within a regional plan. The RNN_V model can support optimization of these Operation and Maintenance activities. The model was developed to predict the horizontal velocity profile across Shinnecock Inlet, but it may be applied to other inlet sites. The ANN was constructed in such a way that the network can be retrained with water level data from a monitoring site close to a particular inlet and velocity measurements across that inlet, and then applied to predict the current profile across the target inlet.

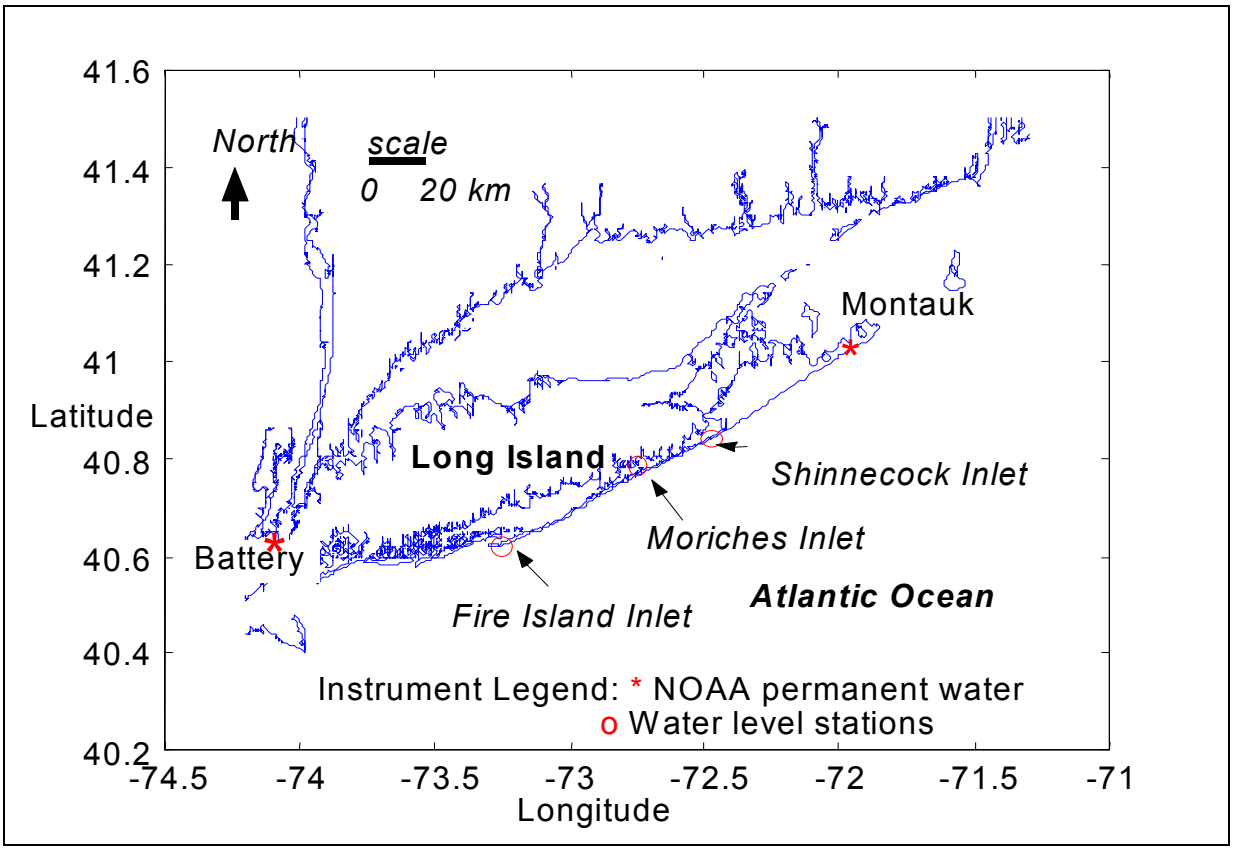

Figure 1. Location of Montauk gauge and Shinnecock Inlet

ARTIFICIAL NEURAL NETWORKS (ANNs): Artificial neural networks are designed to simulate the way simple biological nervous systems work (Hagan et al. 1995; Hsieh and Pratt 2001). ANNs are composed of large numbers of highly interconnected processing units that are analogous to neurons. The interconnected processing units are tied together with weighted connections that are analogous to synapses. ANNs have the capacity to learn, memorize, and create relationships among data. Several records of data are collected and input to the network, which evaluates them and essentially learns how the inputs of each record (water level data, in the present application) are related to the result (the measured current velocity). If there is an overall pattern to the data, or some consistent relationship between the inputs and results of each data set, the network creates an internal mapping of weights that accurately reproduce the expected output (velocity prediction across an inlet). Thus, learning in an ANN involves adjustment to the synaptic connections and the weights that exist between the network's neurons. These weights, and other network parameters created during network training, are then saved and recalled when water level data are entered into the model and current predictions are desired. Of particular advantage is that ANNs can discern correlations between nonlinear input and output signals, such as the ones between water level inputs and current response.

The RNN_V model is a feed-forward back-propagation network with two layers, a hidden layer with a sigmoid transfer function and output layer with a linear transfer function. A transfer function calculates a layer's output from its net input. The weights and biases of the network are iteratively

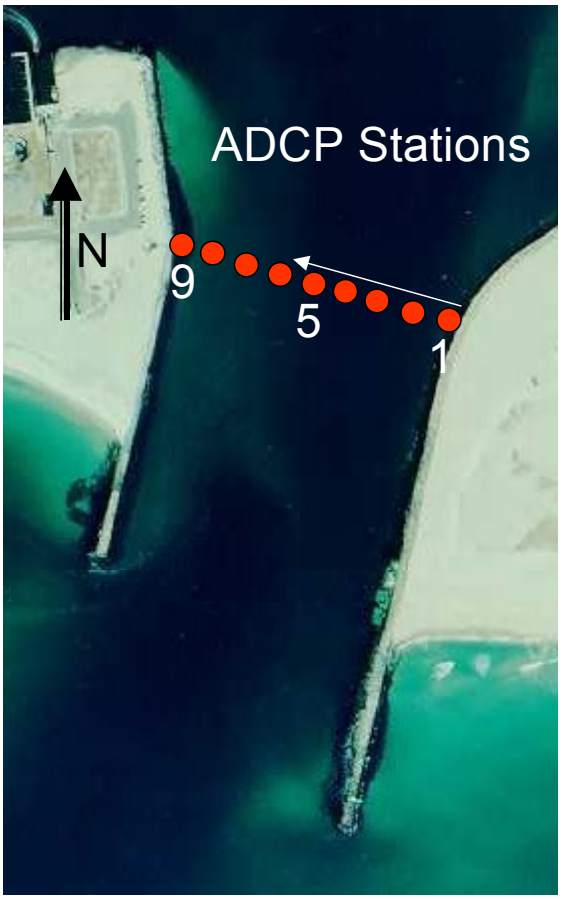

Figure 2. Location of current measurements across Shinnecock Inlet (photo dated April 22, 1997)

adjusted to minimize the network performance function during training. The mean square error between the network output (current predictions) and the target data (observed current data) are used as the performance function. A scaled conjugate gradient back-propagation training algorithm together with batch training are used in the training of the network. Back-propagation refers to the network training method and the manner in which the gradient of the error function is computed for nonlinear multilayer networks such as this one. The basic back-propagation algorithm adjusts weights in the steepest gradient descent direction (the negative of the gradient), which is the direction in which the performance function decreases most rapidly. Although the function decreases most rapidly in this direction, the fastest convergence is not necessarily produced in this direction. In the conjugate gradient algorithms, a search is performed along conjugate directions, which produces faster convergence than the steepest descent directions. Batch training, where the weights are updated after processing the entire training set, as opposed to incremental training, where the weights are updated each time an input is presented to the network, was also used in the training process.

PROCEDURE: The RNN_V model was trained with water level data from a NOAA tide gauge located in Fort Pond Bay at Montauk (Figure 1), and current measurements at Shinnecock Inlet (target data) for April 2000. The current data were collected by an ADCP that measured the current at 9 points across Shinnecock Inlet in April 2000 as a part of CIRP activities. Water level data

relative to mean lower low water (mllw) from the Montauk tide gauge starting from the 1940s have been processed and verified by NOAA and are available from *http://co-ops.nos.noaa.gov/coastline.shtml?region=ny*. Tables 1 and 2 show the format for these input data files.

Table 1
Format for Water Level Input File (Montauk Gauge, May 2001)

Date, month, and day	Time, hours	Time, % of month	Water Level, m, mllw
501	0	0	0.838
501	1	0.0417	0.754
501	2	0.0833	0.643
501	3	0.125	0.487
501	4	0.1667	0.337
501	5	0.2083	0.226
501	6	0.25	0.189
501	7	0.2917	0.245
501	8	0.3333	0.383
501	9	0.375	0.539
501	10	0.4167	0.648
501	11	0.4583	0.709
501	12	0.5	0.712
501	13	0.5417	0.635
501	14	0.5833	0.513
501	15	0.625	0.383
501	16	0.6667	0.266
…	…	…	…

Input files with water level data (Table 1) include the following information:

a. Column 1: Month and day (numeric).

b. Column 2: Time, in hour of the day.

c. Column 3: Time, in percentage of the month.

d. Column 4: Observed water level elevation relative to a specified datum.

Input and output files with current data (Table 2) include the following information:

a. Column 1-3: Same as specified for Table 1.

b. Columns 4-12: Current speed at station across inlet (see Figure 2 for location of stations).

The format shown in Table 2 will be used for the target and validation data files, and predicted currents.

The model training was verified with water level input and current data for the month of May 2000. Both the input and target data consisted of averaged hourly values. The correlation between the model predictions and the measured currents for a number of stations towards the center of the inlet from the verification time period are evident in Figure 3 (sta 4: r = 0.953, rmse = 0.077 m/sec; station 5: r = 0.951, rmse = 0.068 m/sec; station 6: r = 0.956, rmse = 0.057 m/sec). The points across the inlet at which the current has been recorded are numbered one through nine with one being the easternmost station and nine, the westernmost (Figure 4).

Table 2
Format for Files Containing: Target, Validation, and Predicted Currents (Measured Currents at Shinnecock Inlet, May 2001)

			Current Speed at Station (cm/sec)								
Date, month, and day	**Time hours**	**Time , % of Month**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
501	100	0.042	-54.8	-104.2	-87.5	-34.5	-31.2	-38.1	-43.4	-31.2	-31.3
501	200	0.083	-59.6	-107.3	-73.7	-64.1	-60.0	-62.3	-69.0	-54.2	-50.1
501	300	0.125	-46.1	-91.7	-68.0	-64.4	-50.3	-51.1	-52.7	-35.1	-33.8
501	400	0.167	-28.6	-62.4	-49.4	-49.1	-35.6	-35.2	-35.1	-21.7	-21.6
501	500	0.208	-4.9	-17.3	-16.3	-16.7	-7.0	-5.7	-6.9	-3.4	-6.0
501	600	0.250	18.1	33.4	27.7	21.5	28.2	30.8	25.9	19.1	13.5
501	700	0.292	35.6	76.9	70.9	58.8	61.7	63.6	56.1	44.2	37.4
501	800	0.333	47.0	101.6	98.7	86.1	82.3	79.9	72.3	61.3	54.8
501	900	0.375	48.5	96.6	97.0	86.7	77.4	71.8	65.4	58.0	51.7
501	1000	0.417	32.9	58.5	59.3	52.9	46.7	43.1	38.6	36.1	30.3
501	1100	0.458	2.1	0.9	-1.1	3.1	6.8	8.1	5.4	8.7	5.0
501	1200	0.500	-30.3	-53.2	-56.1	-37.1	-23.7	-22.3	-23.6	-14.1	-15.1
501	1300	0.542	-50.7	-87.8	-85.7	-62.1	-41.4	-45.7	-45.1	-31.5	-30.6
…	…	…	…	…	…	…	…	…	…	…	…

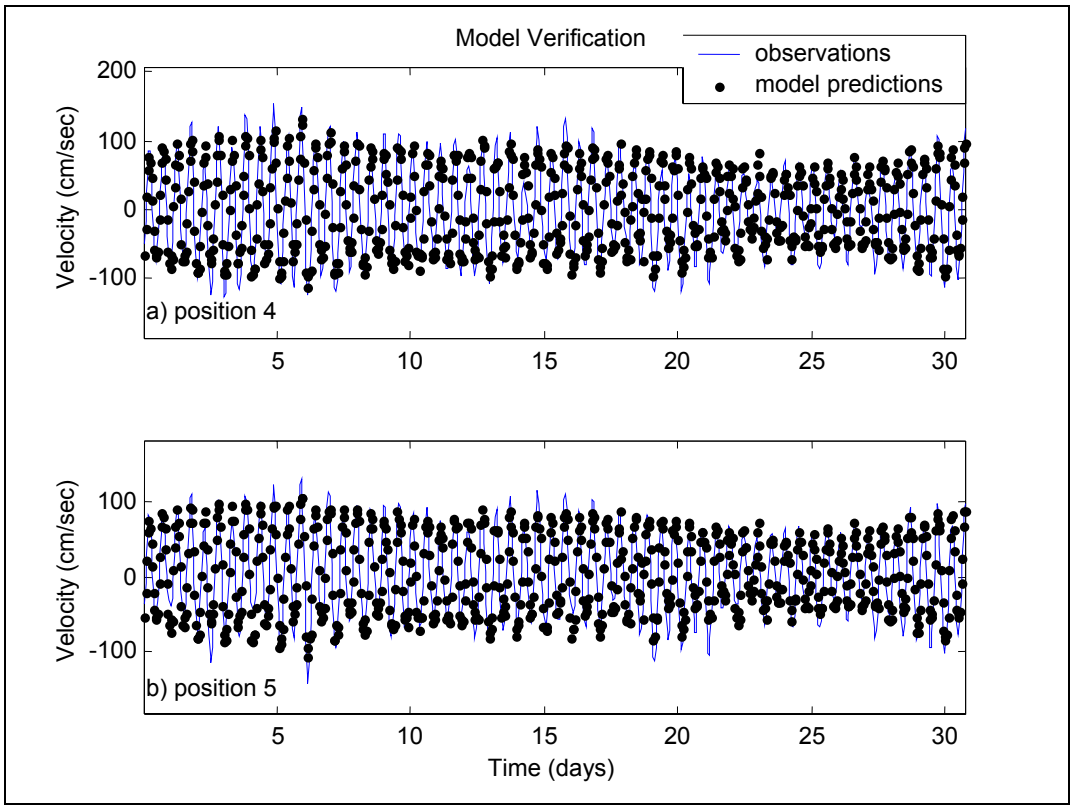

Figure 3. RNN_V model verifications using current data for ADCP monitoring positions 4 and 5 (central inlet) across Shinnecock Inlet during May 2000

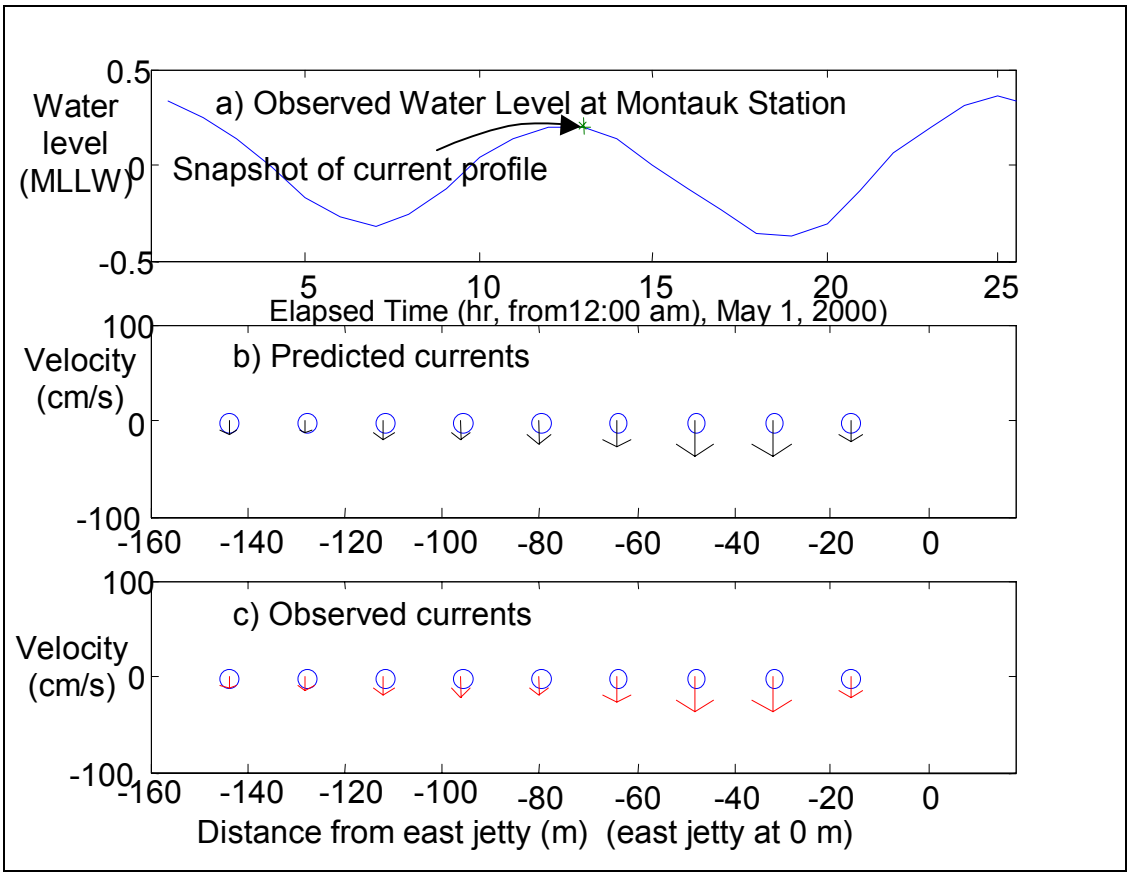

Figure 4. Snapshot of current profile across Shinnecock Inlet for May 2000 and corresponding water level at Montauk, NY. ADCP monitoring positions are 1-9 starting from right (east) at 16-m interval

After the model was satisfactorily verified using current data in May 2000, additional model verifications were conducted using current data in November 2000. Current data in November 2000 consists of gaps. Currents at position 7-9 during November 25-27 are extremely strong, which were different with current patterns in other positions during the same period. The model predicts the currents across Shinnecock Inlet reasonably well in comparison with observations (Figures 5 and 6). The highest correlation (r=0.92-0.95) between model predictions and observed current, however, tends to be towards the center of the inlet (positions 2-6). The correlation between the measured and predicted values tends to drop towards the shore on either side of the inlet, but especially on the western side of the inlet (positions 7 through 9). Correlation coefficients, r, on the western side of the inlet were more likely to fall in the upper 0.80s and lower 0.90s. The lowest correlation between the model predictions and the measured currents was observed at position 8 for November 2000 (r = 0.80). As is evident in Figure 6, there was a relatively strong current later in the month recorded at the western end of the inlet. The model does not predict these extreme values well, which contributes to the decreased correlation coefficients and increased rmse for these western positions. Exactly why correlation between the model predicted and the observed currents is lower closer to the shore is not clear. Possible explanations for the decrease in correlation include wave refraction and surface and/or bottom topography affecting the accuracy of the recordings made by the ADCP in the relatively shallow water.

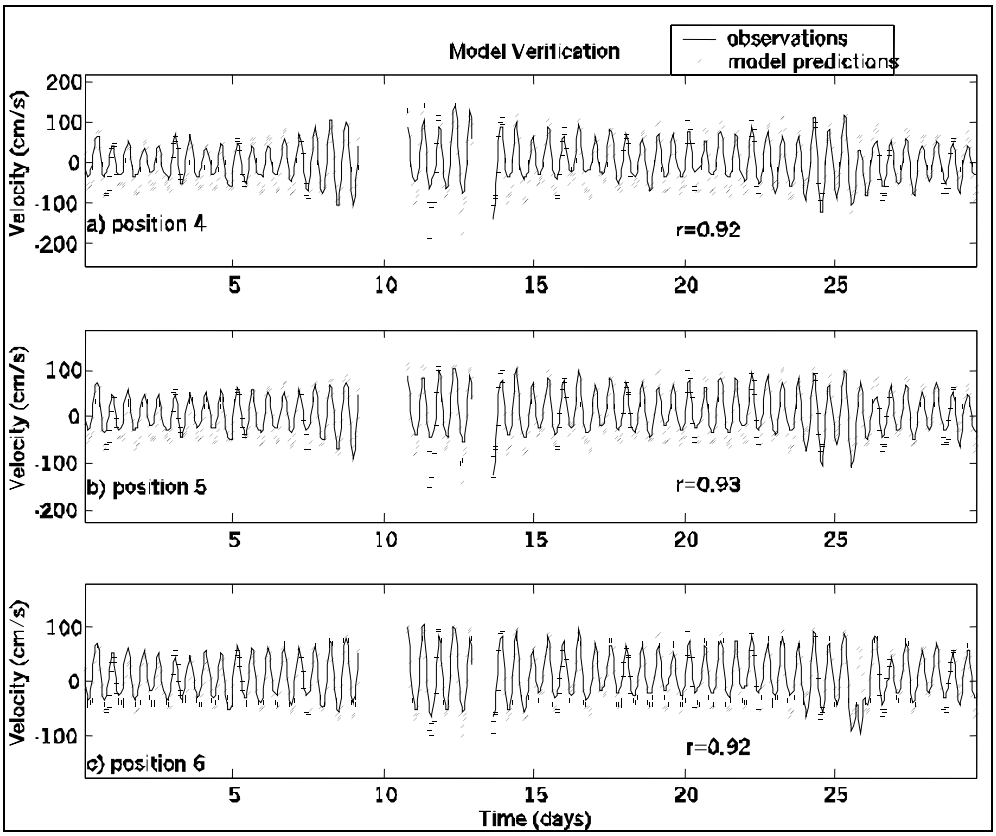

Figure 5. Additional model verifications using current data for ADCP monitoring positions 4-6 (central inlet) across Shinnecock Inlet for November 2000

After training and verification of the RNN_V model using water level data and the current data for a particular inlet/NOAA monitoring station pair, the network parameters are saved in a Matlab© file. This file is then called if the user requests current predictions to be made based only on water level data from the remote monitoring station. The network parameters saved in the Matlab© file are particular to a specific inlet/NOAA monitoring station pair. To obtain current predictions for an inlet, the model must be trained with current data from that inlet and water level data from that remote NOAA monitoring station. Once the training generated network parameter containing .mat file for a particular inlet/NOAA monitoring station pair has been generated, the model is capable of making predictions for any time period (using water level input data from the same time period). The RNN_V model then writes the predicted current at as many as 9 points across the inlet into an output file. The user shall specify the name of the output file.

TUTORIAL: The program RNN_V was written in Matlab© Technical Programming Language (Mathworks *http://www.mathworks.com/index.shtml*) and is available for beta testing from Ms. Julie Rosati, *Julie.D.Rosati@erdc.usace.army.mil*. The following section discusses use of the program. After loading RNN_V in Matlab©, the interface screen is displayed (Figure 7). From this window, select the *Training/Verification*, the *Predictions*, or the *Additional Verification* option from the list box (Figure 8).

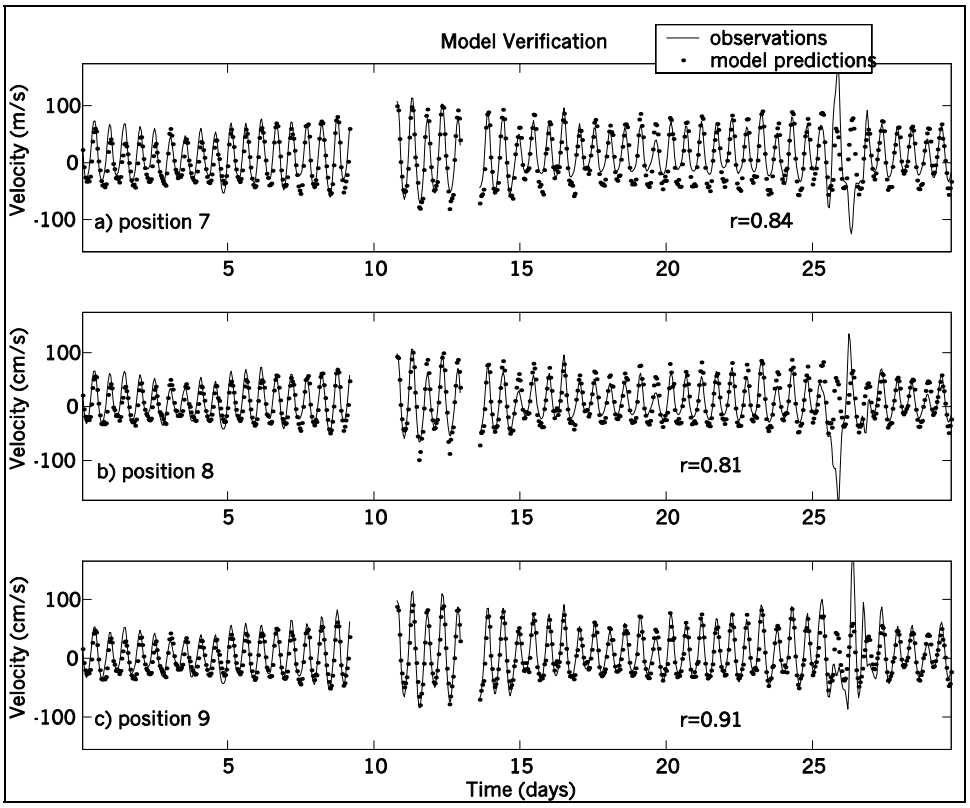

Figure 6. Additional model verifications using current data for ADCP monitoring positions 7-9 across Shinnecock Inlet for November 2000

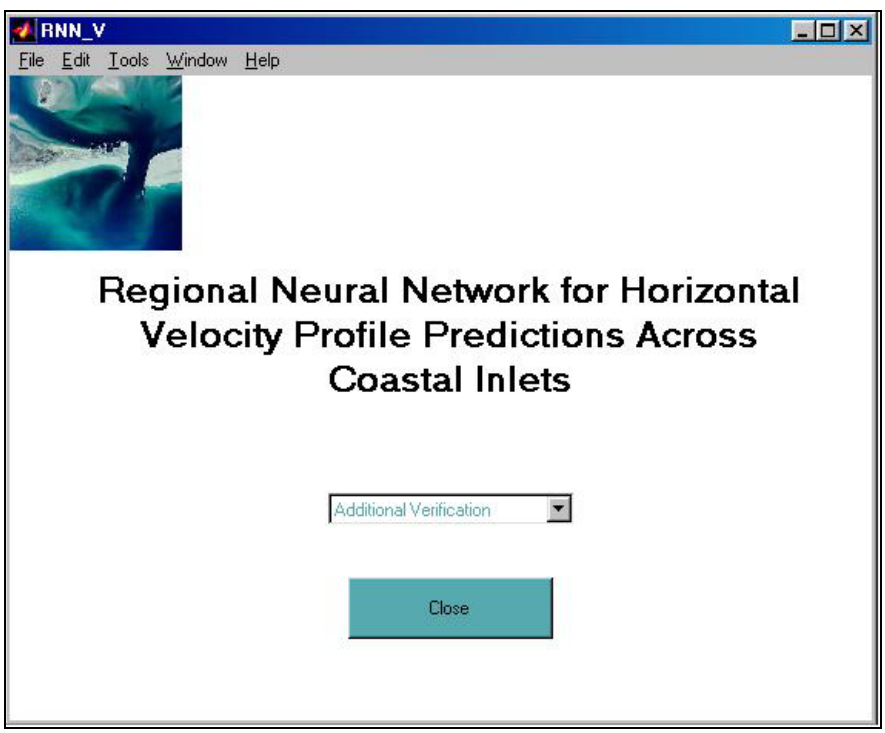

Figure 7. RNN_V interface screen

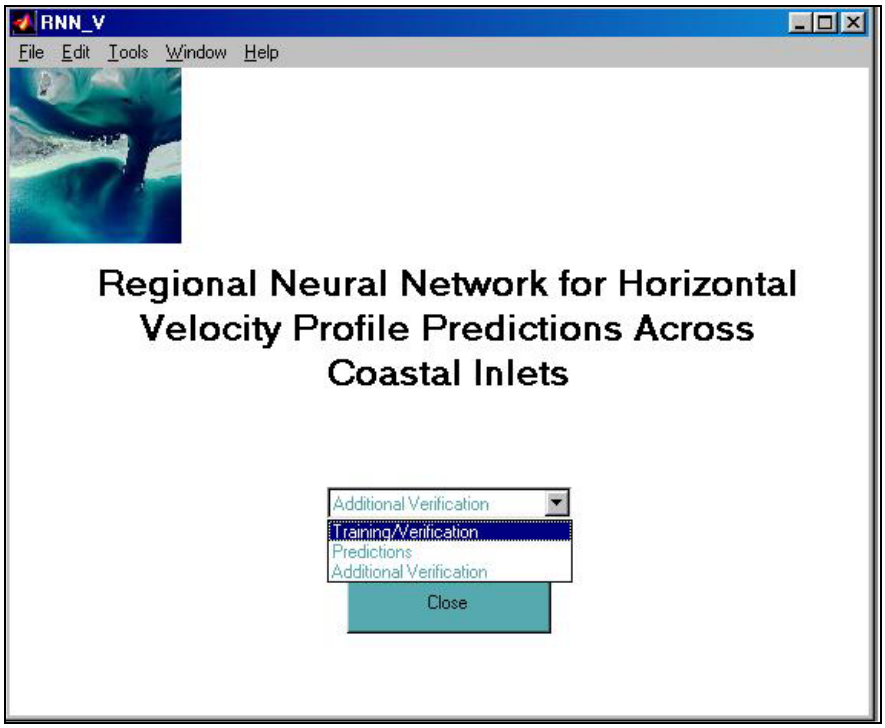

Figure 8. Selecting option for *Training/Verification*

When *Training/Verification* is chosen, the RNN_V *Training/Verification* window opens (Figure 9).

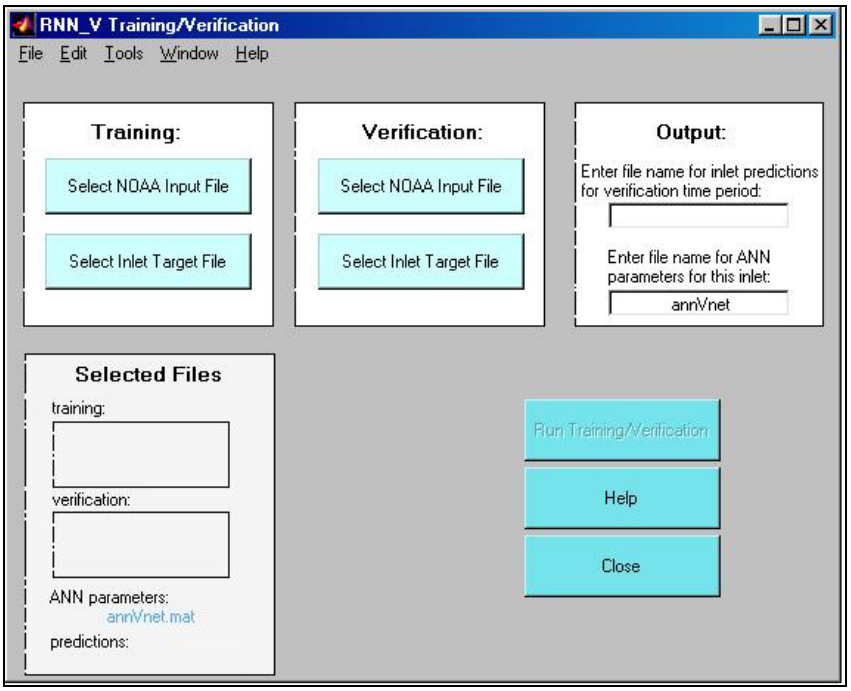

Figure 9. RNN_V *Training/Verification* window

The input (water level from remote NOAA monitoring station, see example shown in Table 1) and target (observed current across inlet, see example shown in Table 2) data sets are selected for both training and verification by pressing the input and target buttons in the training and verification boxes (Figure 10).

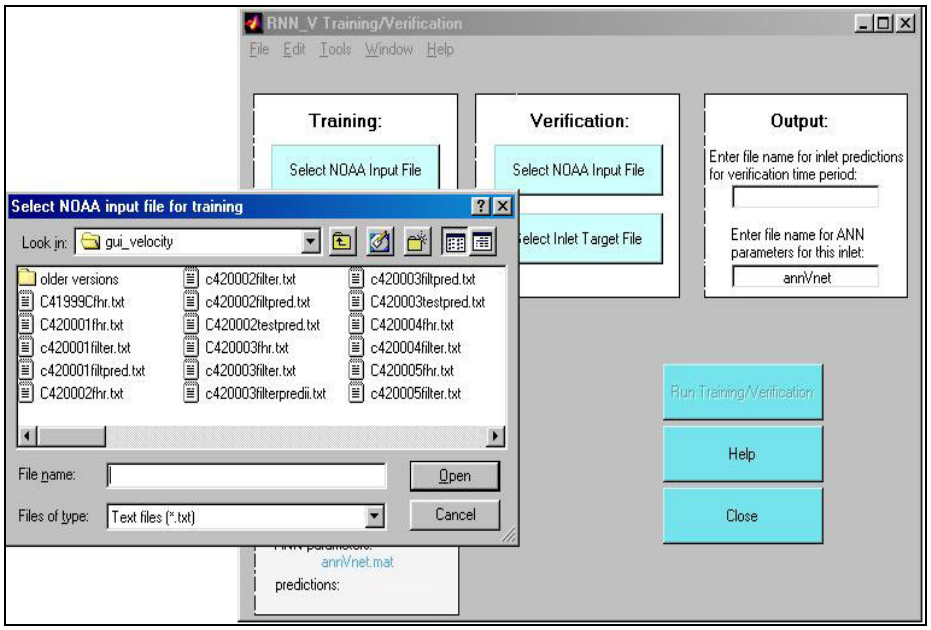

Figure 10. Selecting input and target data sets

As the input and target files are selected for training and verification, the file names are listed in the *Selected Files* box (Figure 11).

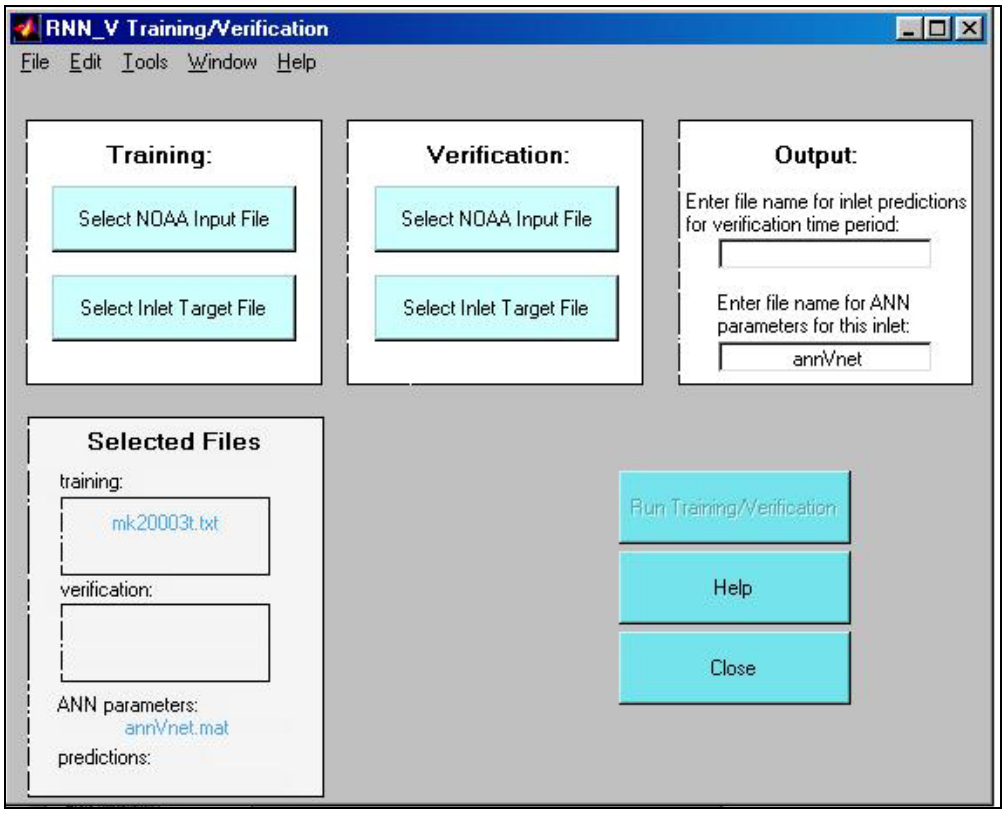

Figure 11. Screen showing input and target file names

After the target and input files have been selected, the output prediction file and the neural network parameters for the monitoring station and inlet in question must be named. The network parameters will be saved as a file named "annVnet" by default (Figure 12). The user can change the parameters file name if desired by entering a new file name in the *Output* box (Figure 13).

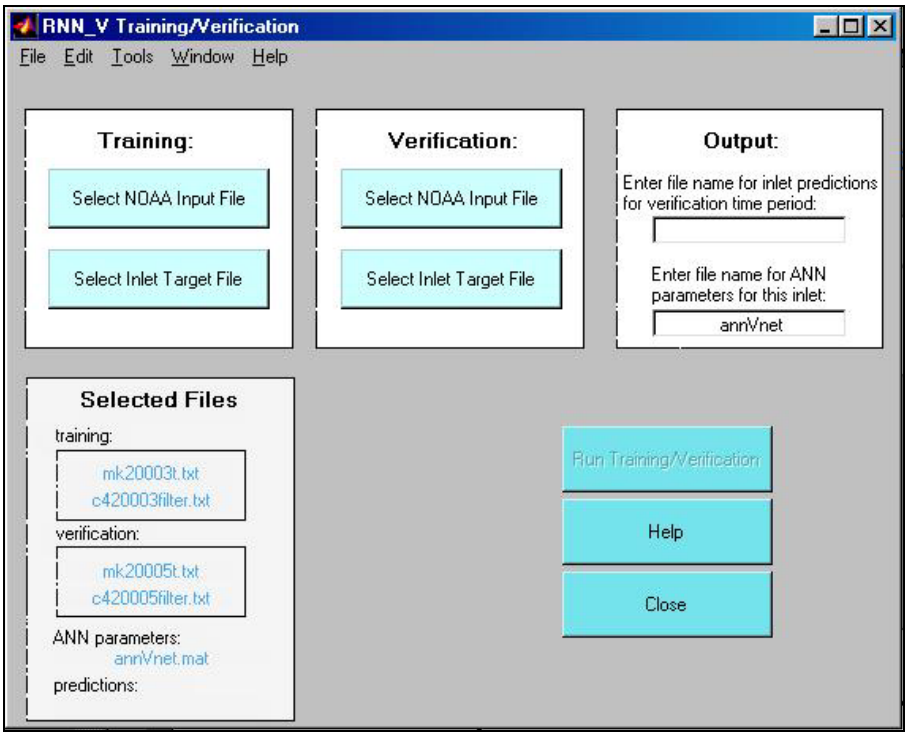

Figure 12. Default file for saving network parameters

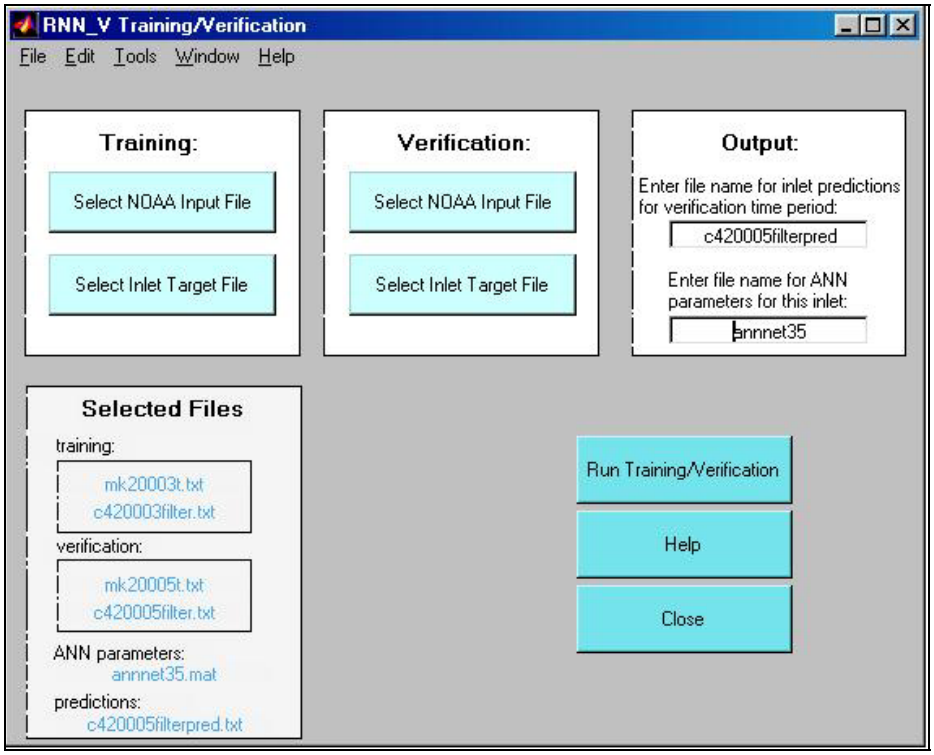

Figure 13. Naming output file names

Once all the training and verification input and target files have been chosen and the neural network parameters and output prediction files have been named, the **Run** button is enabled. By pressing the **Run Training/Verification** button, the *Training/Verification* program is run with the selected target and input files (Figure 14). The ANN parameters and the output predictions will be saved under the indicated file names.

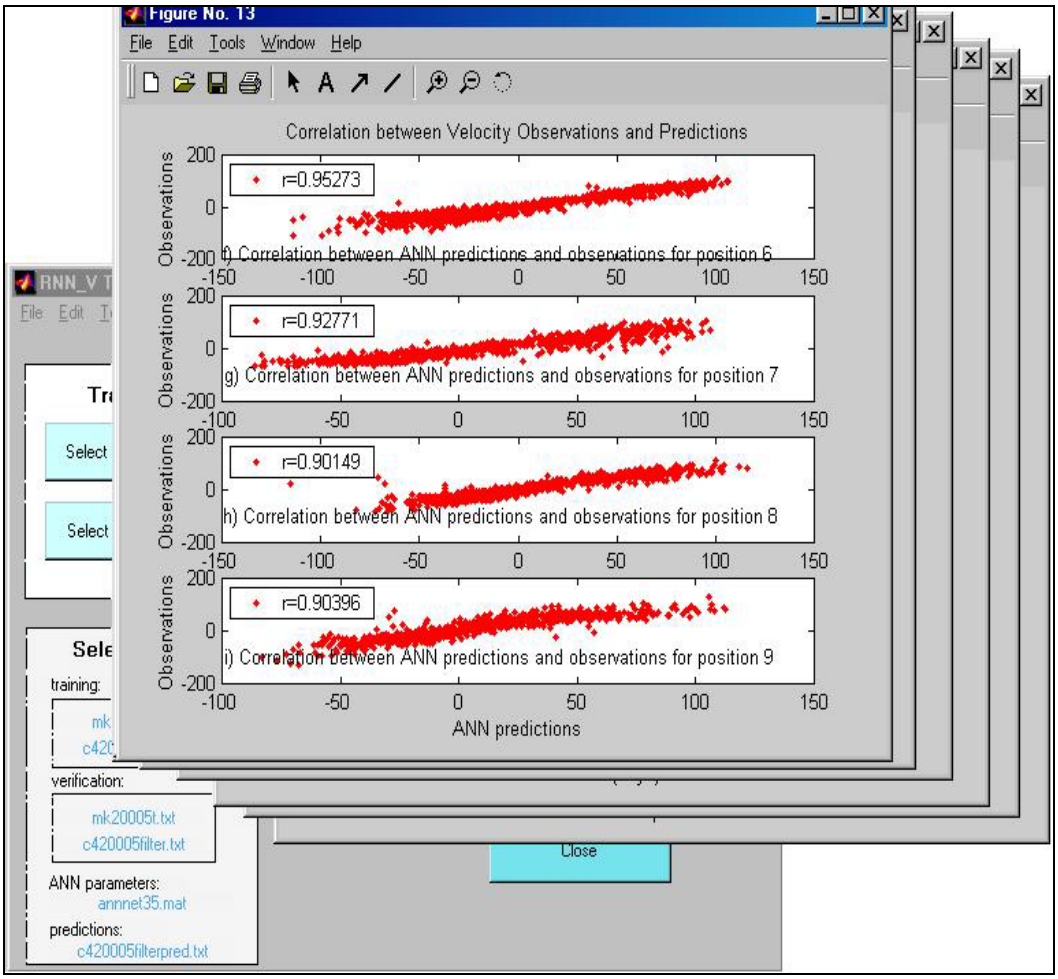

Figure 14. Comparison between predictions and observations

If any problems are encountered, there is a **Help** button to assist the user (Figure 15).

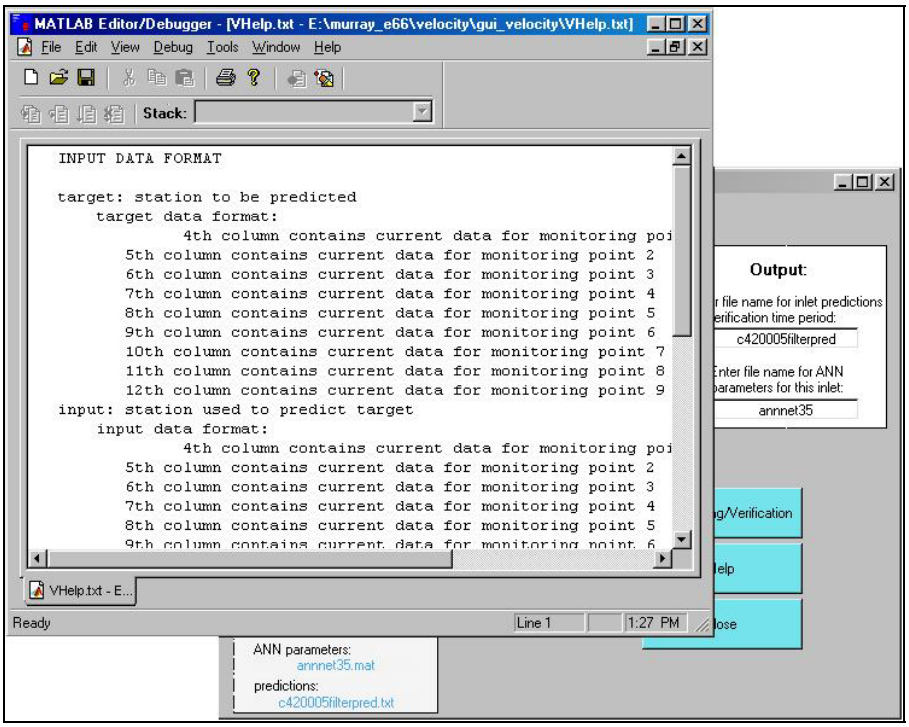

Figure 15. RNN_V help feature

After the relationship between the current across a particular inlet and the water level at a particular remote monitoring station has been determined and the network parameters saved, the user can create predictions for that inlet based upon water level input data from the monitoring station and the saved network parameters. To generate predictions, return to the RNN_V window and select *Predictions* from the listbox (Figure 16).

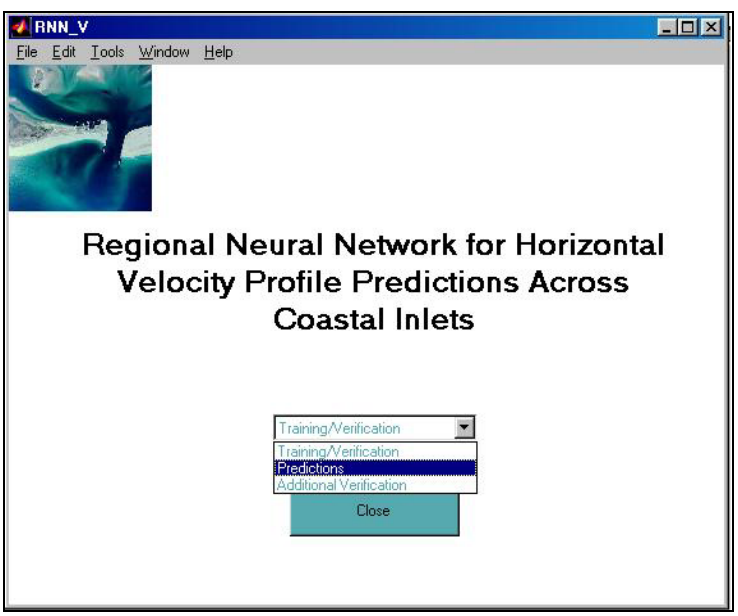

Figure 16. Selecting *Predictions* option

When *Predictions* is chosen, the *RNN_V Predictions* window opens (Figure 17).

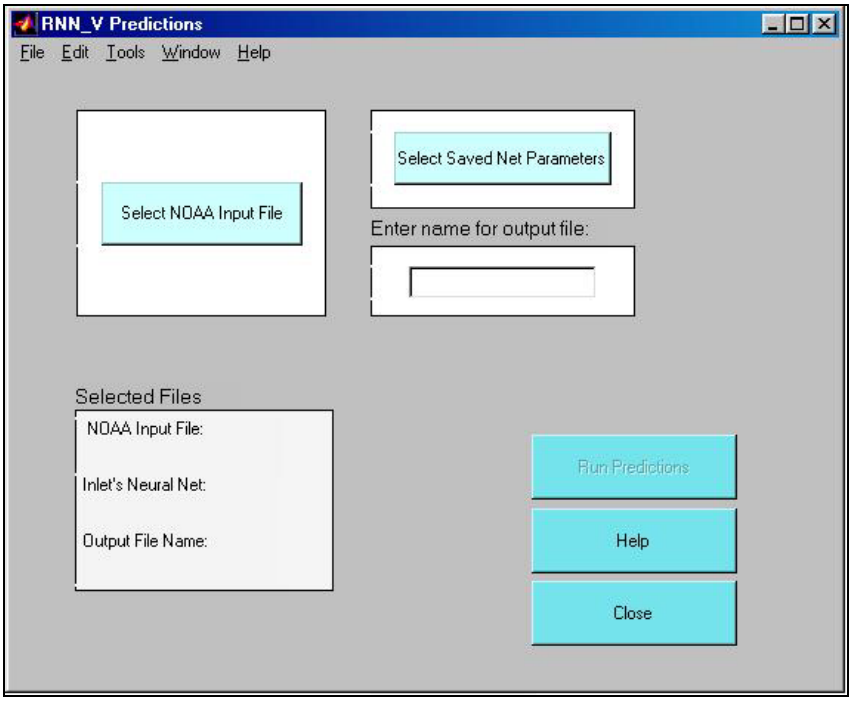

Figure 17. Predictions window

The user selects water level input data from a NOAA monitoring station for a particular time period and the saved network parameters (generated during *Training/Verification*) relating the inlet and monitoring stations in question using the labeled buttons to generate current predictions (see example format shown in Table 2) across the inlet for the same time period as the input data (Figure 18).

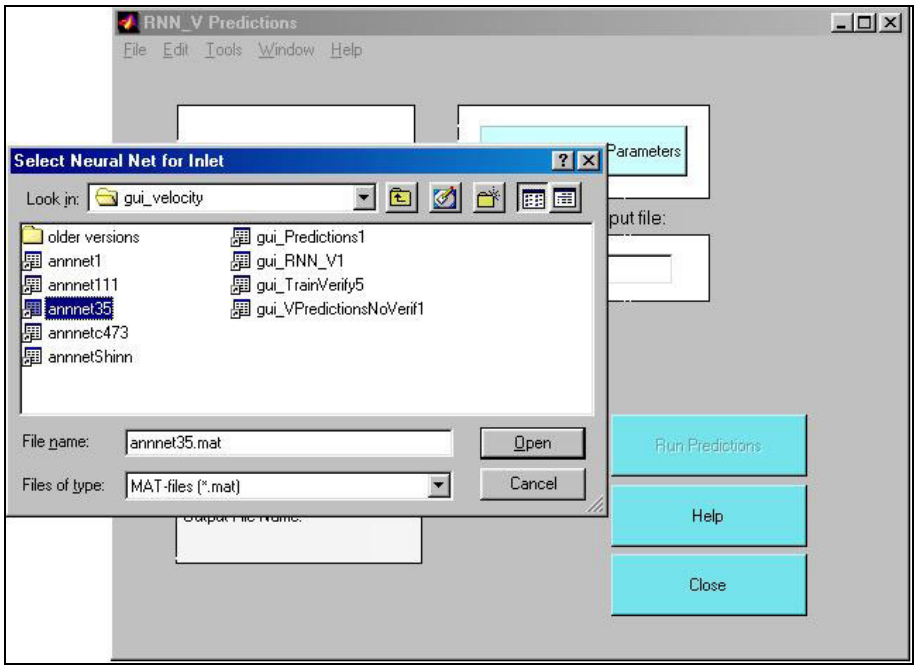

Figure 18. Selecting files that will be used to predict current velocities

Once the input data and the network parameters have been selected and the output predictions file named, the selected files are listed in the *Selected Files* box and the **Run Predictions** button is enabled (Figure 19).

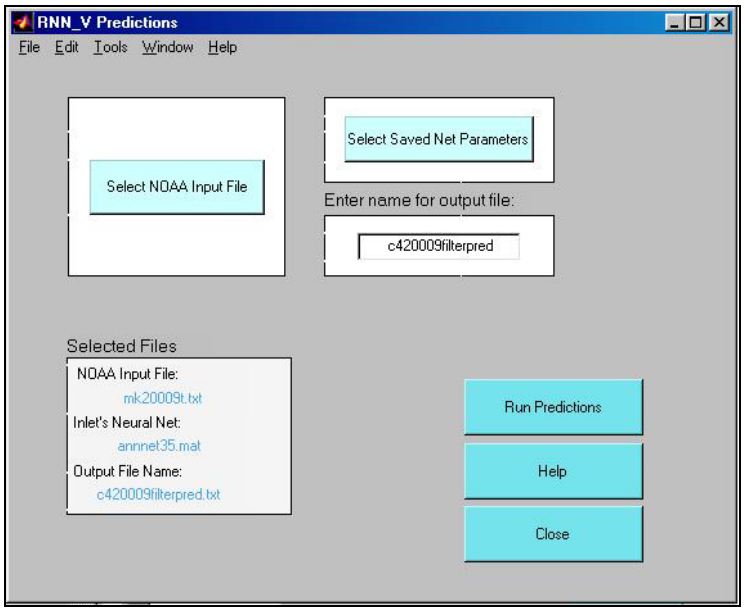

Figure 19. All files have been selected and RNN_V is ready to run predictions

Pressing the **Run Predictions** button generates current predictions across the inlet (Figure 20).

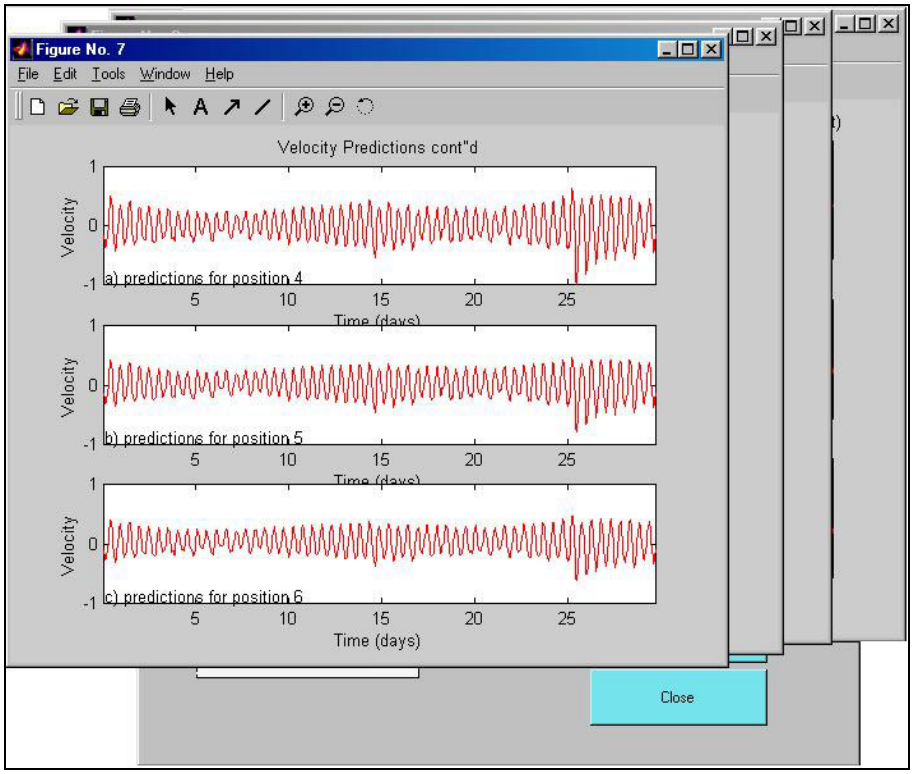

Figure 20. Predicted current speeds

CONCLUSIONS: One of the major advantages of artificial neural networks is that, in principle, they are capable of approximating any continuous function, so that the user does not need to have any hypotheses about the underlying structure, or even, to some extent, which variables matter (Haykin 1999). The resulting network developed in the learning process represents a pattern detected in the data. Thus, ANN methods can be applied to many processes in coastal engineering and oceanography (Hsieh and Pratt 2001).

As long as the training data set covers the maximum range of the forecasting boundary data, a short-term data set can be used to train an ANN model for long-term predictions. For a given point, a trained neural network can provide a faster simulation for forecasting long-term events than traditional hydrodynamic models because its calculation requires no computational iteration. In addition, a trained ANN model can provide information to fill gaps in measurements and can hindcast data for times when measurements are unavailable.

ADDITIONAL INFORMATION: This technical note was produced under the Coastal Sedimentation and Dredging Program work unit "Regional-Scale Modeling Sediment Transport and Morphology Change" by Dr. Wenrui Huang, Assistant Professor, Florida Agricultural and Mechanical University (FAMU), and Ms. Catherine Murray, Master's degree student, FAMU. Ms. Julie D. Rosati at the U.S. Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory (CHL) is the work unit Principal Investigator. Stimulating discussions with Dr. Nicholas C. Kraus at CHL are much appreciated. Mr. Bill Grosskopf, Offshore and Coastal

Technologies Inc. - East, provided assistance in creating Figure 2. Mr. Thomas Wilson provided the ADCP data. Dr. Liwah Lin, CHL, provided review comments and testing of the code. Questions can be addressed to Dr. Huang at *whuang@eng.fsu.edu* or to Ms. Rosati at *Julie.D.Rosati@erdc.usace.army.mil*. The Matlab© program RNN_V is available by contacting Ms. Rosati.

REFERENCES

Hagan, M. T., Demuth, H., and Beale, M. (1995). "Neural network design," PWS Publishing Co., Boston, MA.

Haykin, S. (1999). *Neural network: A compressive foundation*. Prentice Hall, NJ.

Hsieh, B. B., and Pratt, T. C. (2001). "Field data recovery in tidal system using artificial neural networks (ANNs)," ERDC/CHL CHETN-IV-38, U.S. Army Engineer Research and Development Center, Vicksburg, MS. *http://chl.wes.army.mil/library/publications/chetn*

NOTE: *The contents of this technical note are not to be used for advertising, publication, or promotional purposes. Citation of trade names does not constitute an official endorsement or approval of the use of such products*