Wave Height Forecasting Using Cascade Correlation Neural Network

Hamidreza Rashidy Kanan
Electrical Engineering Department, Amirkabir University of Technology, Hafez Avenue, Tehran, Iran, 15914
rashidykanan@aut.ac.ir

Karim Faez
Electrical Engineering Department, Amirkabir University of Technology, Hafez Avenue, Tehran, Iran, 15914
kfaez@aut.ac.ir

ABSTRACT

Forecasting of wave height is necessary in a large number of ocean coastal activities. Recently, neural networks are used for prediction and approximation of wave heights in sea and ocean due to their great convergence rate. In this paper a cascade correlation neural network is used for prediction of wave heights at given times due to the useful capability of this network for prediction and approximation. Results of different prediction for 500 data points in cascade correlation neural network are compared with those of the M.L.P. (Multi-layer Perceptron) neural network. These results show that cascade correlation network has larger convergence rate compared with M.L.P. network. Also various simulations show that the cascade correlation network has better performance with α=0.005 (Learning-rate), sigmoid activation function for hidden units and linear activation function for output units.

Keywords

Forecasting, Prediction, Wave Height, M.L.P. Network, Cascade Correlation Network, Quick-Propagation Training Algorithm.

1. INTRODUCTION

In recent years, artificial neural networks have been used in a number of coastal engineering applications such as forecasting of the Tide Level, Sea Level, Sea Currents and etc due to their ability to approximate the nonlinear mathematical behavior without a priori knowledge of interrelations among the elements within a system [Mak03, Puc01, Moh01].

Forecasting of Wave heights is currently done by numerically solving the differential equation representing wave energy balance. The procedure involved is extremely complex and calls for very large amounts of meteorological and oceanographic data [Deo99].

Due to the complexity of above methods, nowadays, artificial neural networks are used for forecasting of wave heights in sea and ocean, since this method has higher precision and convergence rate compared with the other methods [Deo99].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

WSCG POSTERS proceedings
WSCG'2004, February 2-6, 2004, Plzen, Czech Republic.
Copyright UNION Agency – Science Press

2. CASCADE CORRELATION NEURAL NETWORK

In addition to the probabilistic neural net, cascade correlation is another network that modifies its own architecture as training progresses [Hoe92]. It is based on the premise that the most significant difficulty with current learning algorithms (such as back propagation) for neural networks is their slow rate of convergence. This is due, at least in part, to the fact that all of the weights are being adjusted at each stage of training. A further complication is the rigidity of the network architecture throughout training [Hoe92, Fau94, Mik99, Pha94].

Cascade Correlation network addresses both of these issues by dynamically adding hidden units to the architecture, but only up to the minimum number necessary to achieve the specified error tolerance for the training set. Furthermore, a two-step weight-training process admits that only one layer of weights is being trained at any time. This allows the use of simpler training rules (the delta rule, perceptron etc.) than for multi layer training. In practice, a modification of back propagation algorithm known as Quick Propagation is usually used [Hoe92, Ada95].

2.1 Cascade Correlation Neural Network Architecture

A cascade correlation [Fau94] net consists of input units, hidden units, and output units. Input units are connected directly to output units with adjustable weighted connections. Connections from inputs to a hidden unit are trained when the hidden unit is added to the net and are then frozen. Connections from the hidden units to the output units are adjustable consequently.

Cascade correlation network starts with a minimal topology, consisting only of the required input and output units (and a bias input that is always equals to 1). This net is trained until no further improvement is obtained. The error for each output until is then computed (summed over all training patterns).

Next, one hidden unit is added to the net in a two-step process. During the first step, a candidate unit is connected to each of the input units, but is not connected to the output units. The weights on the connections from the input units to the candidate unit are adjusted to maximize the correlation between the candidate's output and the residual error at the output units. The residual error is the difference between the target and the computed output, multiplied by the derivative of the output unit's activation function, i.e., the quantity that would be propagated back from the output units in the back propagation algorithm. When this training is completed, the weights are frozen and the candidate unit becomes a hidden unit in the net.

The second step in which the new unit is added to the net now begins. The new hidden unit is then connected to the output units, and the weights on the connections being adjustable. Now all connections to the output units are trained. (Here the connections from the input units are trained again, and the new connections from the hidden unit are trained for the first time.)

A second hidden unit is then added using the same process. However, this unit receives an input signal from the both input units and the previous hidden unit. All weights on these connections are adjusted and then frozen. The connections to the output units are then established and trained. The process of adding a new unit, training its weights from the input units and the previously added hidden units, and then freezing the weights, followed by training all connections to the output units, is continued until the error reaches an acceptable level or the maximum number of epochs (or hidden units) is reached.

This process is shown in figures 1 through 5.

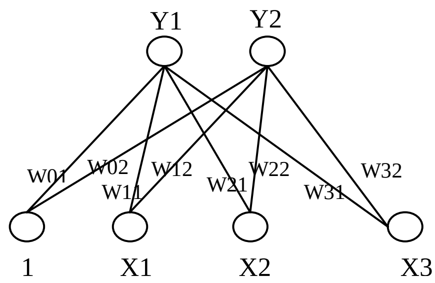

Fig 1. Stage 0, no hidden units

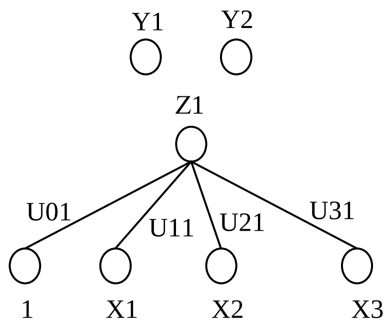

Fig 2. Stage 1, one candidate unit (z1)

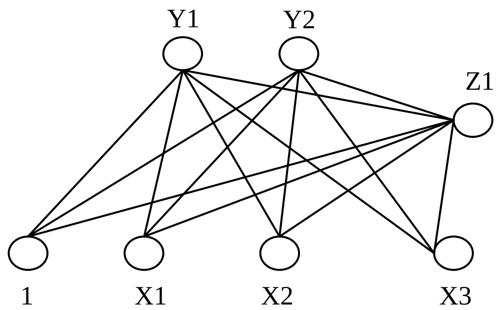

Fig 3. Stage 1, one hidden unit (z1)

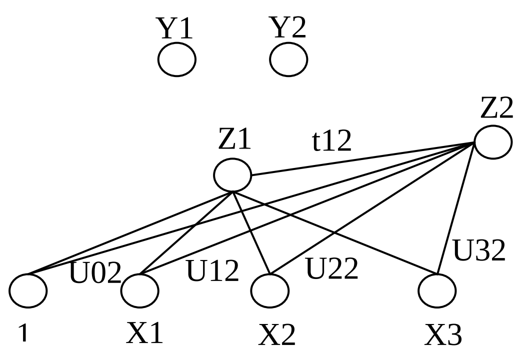

Fig 4. Stage 2, new candidate unit (z2)

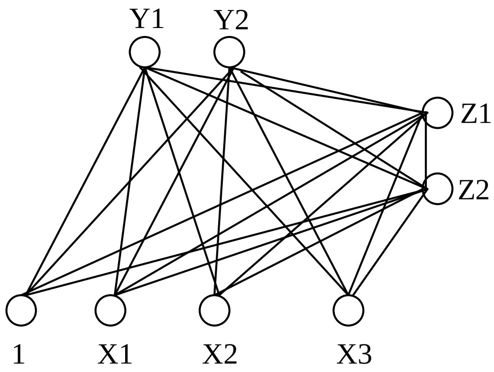

Fig 5. Stage 2, hidden unit (z2)

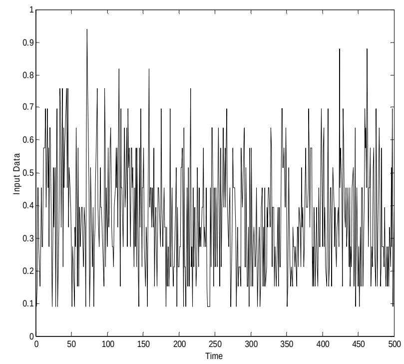

Fig 7. Target data for first case of training

3. SIMULATION

We have used 1000 data points of measured wave heights in a given position of sea, corresponding to 1000 sequential hours, to train the M.L.P. network (with one hidden layer, one input unit, five hidden unit and one output unit) and also to the cascade correlation network (with one input unit, one hidden unit and one output unit), in two different cases.

In the first case, the data of sequential odd hours (i.e. the data of first, third and … hour) are applied as input data and the data of sequential even hours (i.e. the data of second, forth and …hour) are applied as the corresponding target data to each network.

In the second case, the sequential data that each of them has five-hour delay with other (i.e. the data of first, sixth and … hour) are applied as input data and the data of one hour later are applied as the corresponding target data to each network.

For example, the input data and the corresponding target data of first case are shown in figures 6 and 7 respectively. Obviously the wave height is in meter and these data are normalized.

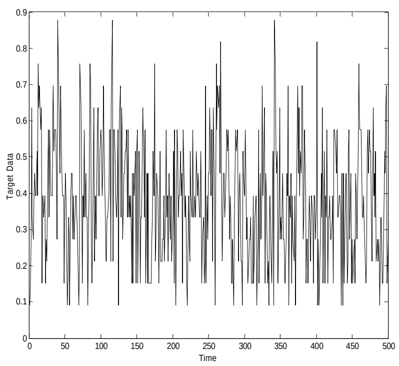

Fig 6. Input data for first case of training

The training phase results of the two mentioned cases are shown in Table 1 and 2. It is necessary to mention that for the best situation to compare the M.L.P. network with the cascade correlation network the number of hidden units of M.L.P. network achieved is five and for the optimized cascade correlation network, the best number of hidden units is one.

Network Type	**Network Configuration**	**Epochs**	**Time(s)**
M.L.P. (Back-Propagation)	1-5-1	25550	53655
Cascade-Correlation	1-1-1	30	24

Table 1. Comparison of convergence rate in M.L.P. and the cascade correlation network for first case of training

Network Type	**Network Configuration**	**Epochs**	**Time(s)**
M.L.P. (Back-Propagation)	1-5-1	25550	55114
Cascade-Correlation	1-1-1	51	33

Table 2. Comparison of convergence rate in M.L.P. and the cascade correlation network for second case of training

To obtain the best value of learning rate for cascade correlation network in the first case, the network is simulated with different values of learning rate and $\alpha=0.005$ is obtained as the best value and the corresponding curve is shown in figure 9.

Also, figure 8 demonstrate the output error versus the number of epochs for the first case of training in the cascade correlation network.

At last, in the test phase, the 500 remaining data of the whole data are used for the test in two different given cases and for the demonstration simplification, 70 data points of them in the first case are shown in figure 10.

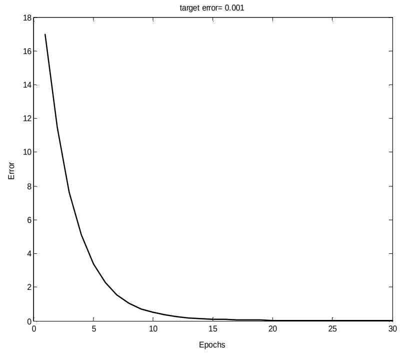

Fig 8. Output error versus the number of epochs for the first case of training in cascade correlation network

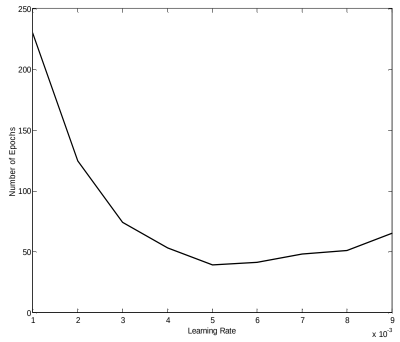

Fig 9. Number of epochs for convergence versus the learning rate in the first case of training

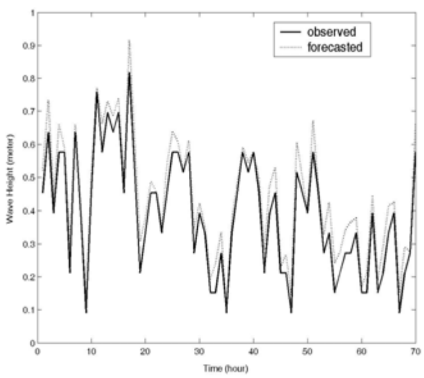

Fig 10. Forecasted and measured wave height for the first case of training

4. CONCLUSION

In this paper, M.L.P. and cascade correlation networks are used for wave height forecasting. Simulation results show that the cascade correlation network due to its characteristics has better performance in convergence rate and the training time compared with the M.L.P. network.

Considering the simulation results, the cascade correlation network has a better performance with sigmoid and linear activation function at the hidden and output units respectively. Also the network training with different learning rates demonstrates that α=0.005 is the best learning rate for this network.

5. REFERENCES

[Mak03] O.Makarynskyy, A.A.Pires-Silvad, Makarynska, C.Ventura-Soares, “Artificial Neural Networks in the Forecasting of Wave Parameters”, 7th International Workshop on Wave Hindcasting and Forecasting Preprints, canada, june 11-14, 2003.

[Puc01] S.Puca, B.Tirozzi, G.Arena, S.Corsini and R.Inghilesi, “A Neural Network Approach to the Problem of Recovering Lost Data in a Network of Marine Buoys”, Conference of ISOPE, Norway, pp. 615-620, June 17-22, 2001.

[Moh01] Mohamed A. Shahin, Mark B. Jaksa and Holger R. Maier, “Artificial Neural Network Applications in Geotechnical Engineering”, Australian Geomechanics, Vol. 36, No. 1, pp. 49-62, 2001.

[Deo99] M.C.Deo and C.Sridhar Naidu, “Real Time Wave Forecasting Using Neural Networks”, Ocean Engineering, 26(3), pp. 191-203,1999.

[Hoe92] Hoehfeld, M. and S. E. Fahlman "Learning with Limited Numerical Precision Using the Cascade-Correlation Learning Algorithm", IEEE Transactions on Neural Networks, Vol. 3, No. 4, pp. 602-611, 1992.

[Fau94] L.Fausset, *Fundamental of Neural Networks Architectures, Algorithms and Applications*, Prentice Hall, 1994.

[Mik99] Mikko Lehtokangas, “Fast Initialization for Cascade-Correlation Learning”, IEEE Transactions on Neural Networks, Vol. 10, No. 2, pp. 410-414, 1999.

[Pha94] D.S.Phatak and I.Koren, “Connectivity and Performance Tradeoffs in the Cascade Correlation Learning Architectures”, IEEE Transactions on Neural Networks, Vol. 5, No. 6, pp. 930-935, 1994.

[Ada95] A.Adams & S.Waugh, “Function Evaluation and the Cascade-Correlation Architecture”, ICNN), pp. 1126-1129. Perth, Western Australia, 1995.