Prediction of Oceanic Wave Heights using Artificial Neural Network

presented by

Akshat Shrivastava
01004018

under the guidance of

Prof. M.C. Deo

Introduction

- Objective : develop a NN to forecast oceanic wave heights on real time basis
- ANN – alternative to the current numerical techniques
- Records are of a chosen site, data of which is available for download

Site of Experiment

- 180 NM south of Southwest Pass, LA
- Station No. 42001
- Water depth 3246.0 mts
- Discus Buoy of 12m
- Data Available on NDBC website (http://www.ndbc.noaa.gov)

Data Available

- On an average 8700 records for each year are available – Hs (m) at every hour in 2001
- Hourly wave heights for year 2004 were to be predicted
- Monthly Statistical values were calculated.
- Skewness is high in months of August, September and October

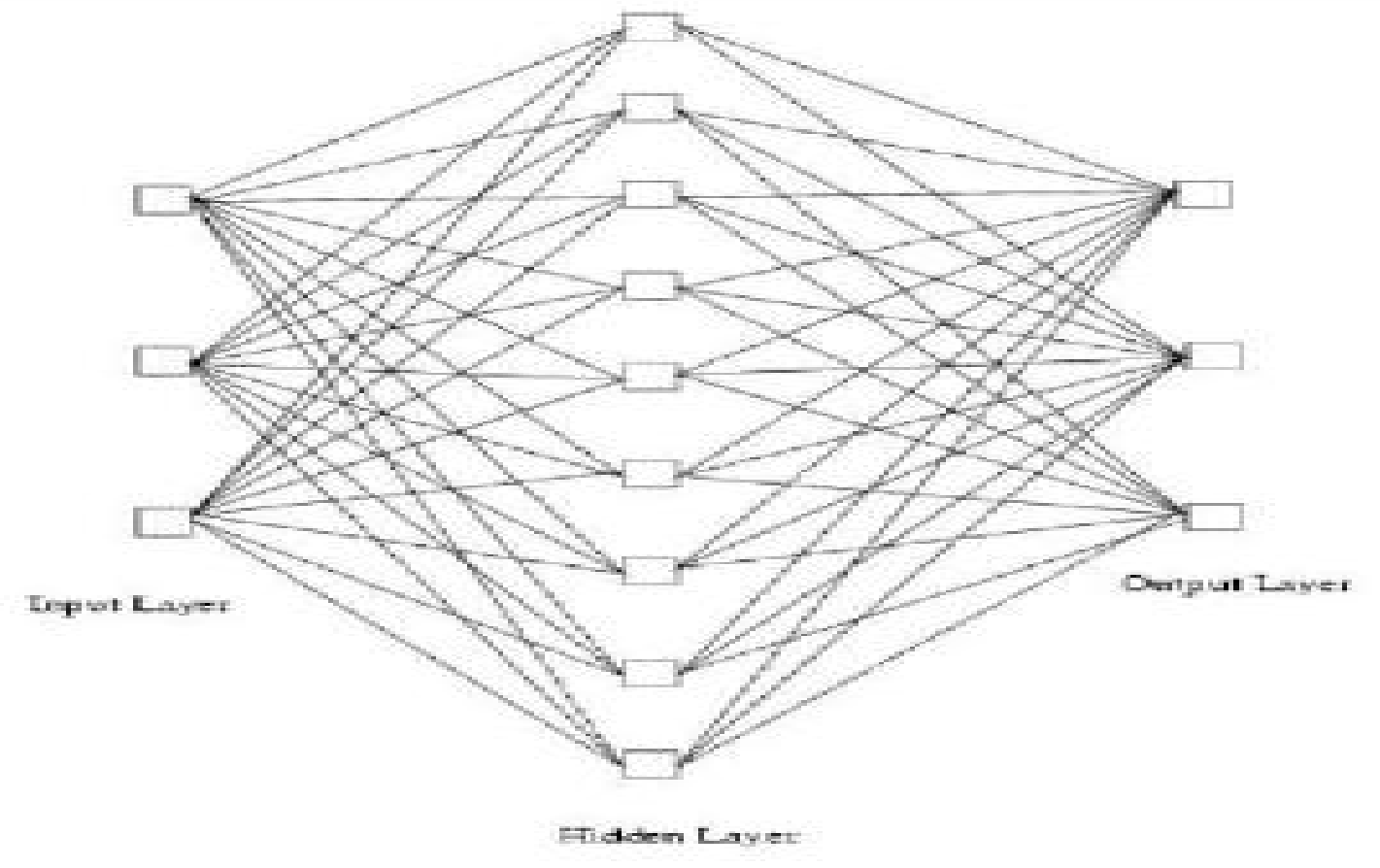
Year 2001

Month	Mean	Skewness	Std Dev.	Min	Max
Jan	1.4197	0.7207	0.7500	0.27	3.9
Feb	1.4519	0.9486	0.6586	0.53	3.84
Mar	1.2313	0.2385	0.5201	0.38	2.65
Apr	1.0418	1.0418	0.5407	0.23	3.34
May	1.0312	1.1729	0.3175	0.46	2.41
Jun	0.7689	1.1033	0.3869	0	2.64
Jul	0.4451	1.2958	0.1919	0.17	1.26
Aug	0.6597	2.5717	0.5035	0.23	3.64
Sept	1.284	3.0921	1.3518	0.2	8.77
Oct	1.0090	1.1895	0.5913	0.22	2.89
Nov	1.4274	1.0135	0.6266	0.51	3.92
Dec	1.6563	1.3457	0.8496	0.27	5.96

Methodology

- It was decided to try and test four types of networks
  - 3 Input nodes – 3 Output nodes network
  - 6 Input nodes – 6 Output nodes network
  - 12 Input nodes – 12 Output nodes network
  - 24 Input nodes – 24 Output nodes network
- A typical 3 input – 3 output network is shown

A typical Neural Network



Methodology

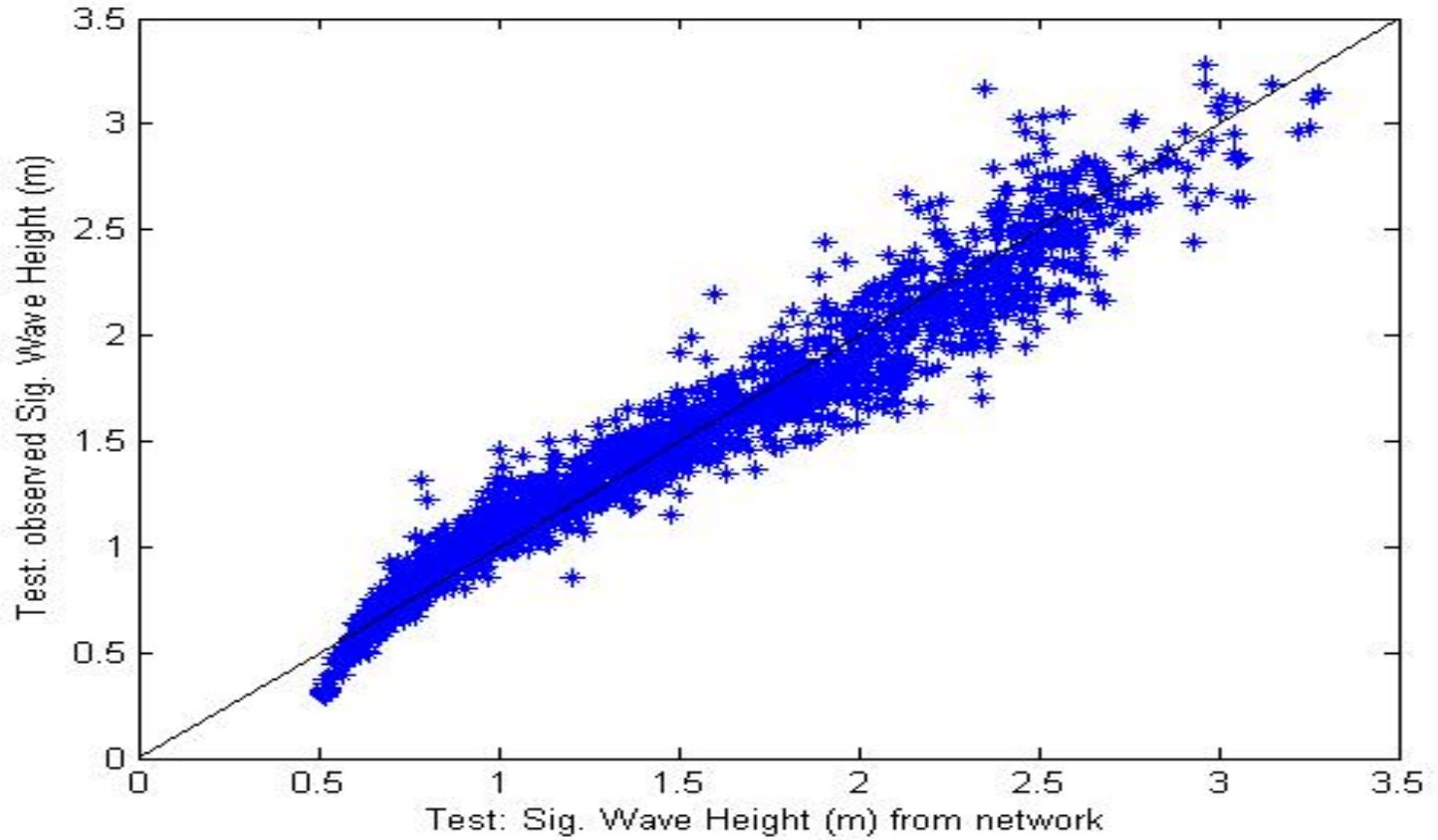
- NN Toolbox of MATLAB 7.0 is used
- Networks are trained using Backpropagation algorithm
- Different variations of backpropagation were tried for training e.g. Resilient Propagation
- No. of hidden nodes, momentum rate, learning rate were changed to get optimum network

Methodology

- Year 2001 data was divided into two sets - 9 months data for training and 3 months for testing
- All the four networks were trained and tested for same set of data
- 3-3 and 6-6 networks produced best results

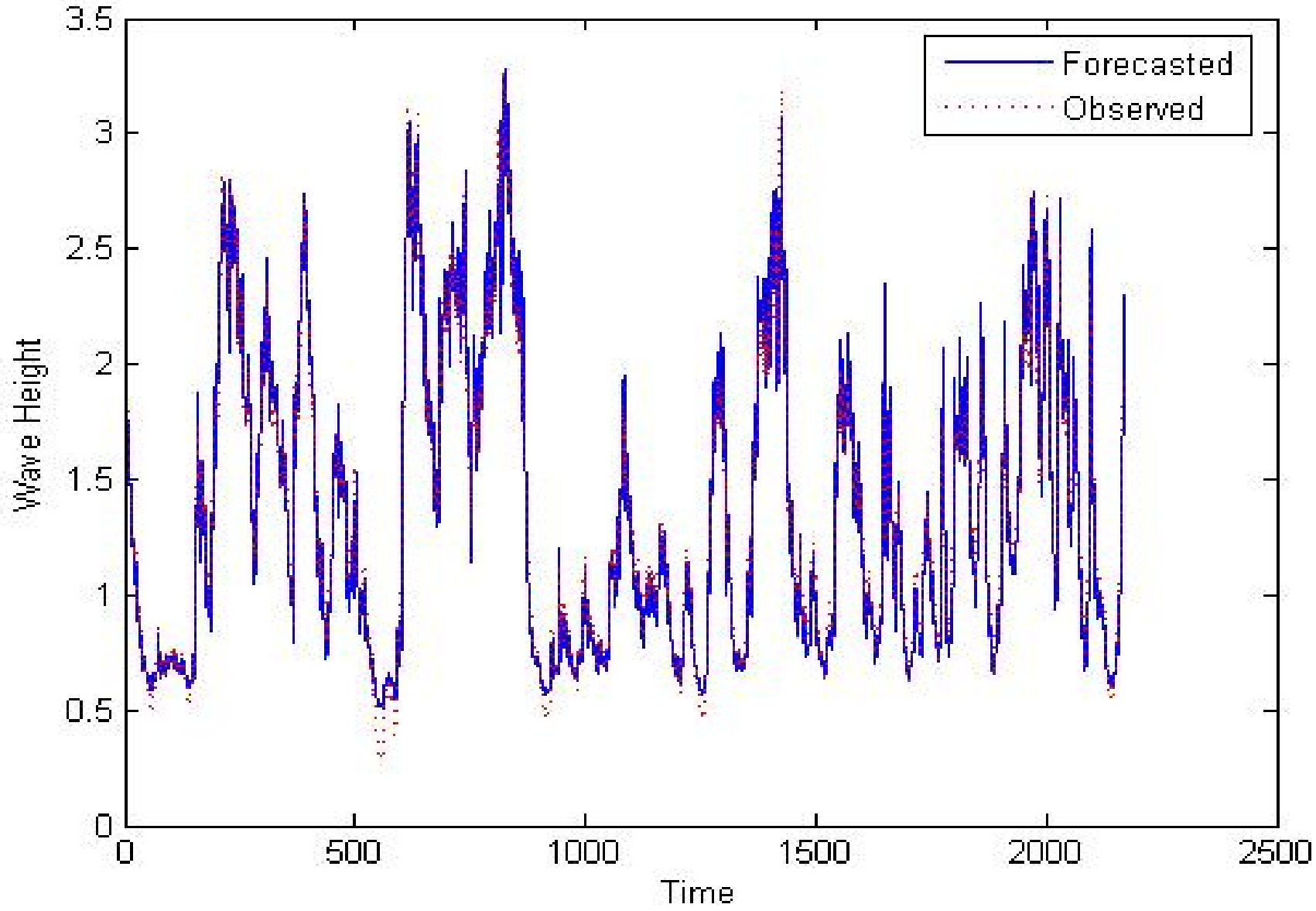
Parameter	3-3	6-6	12-12	24-24
R	0.9572	0.9446	0.8853	0.7432
MSE	0.0366	0.0400	0.0836	0.1787
MAE	0.1356	0.1362	0.2016	0.3065

Scatter Diagram (3-3 net 1hr lead time)

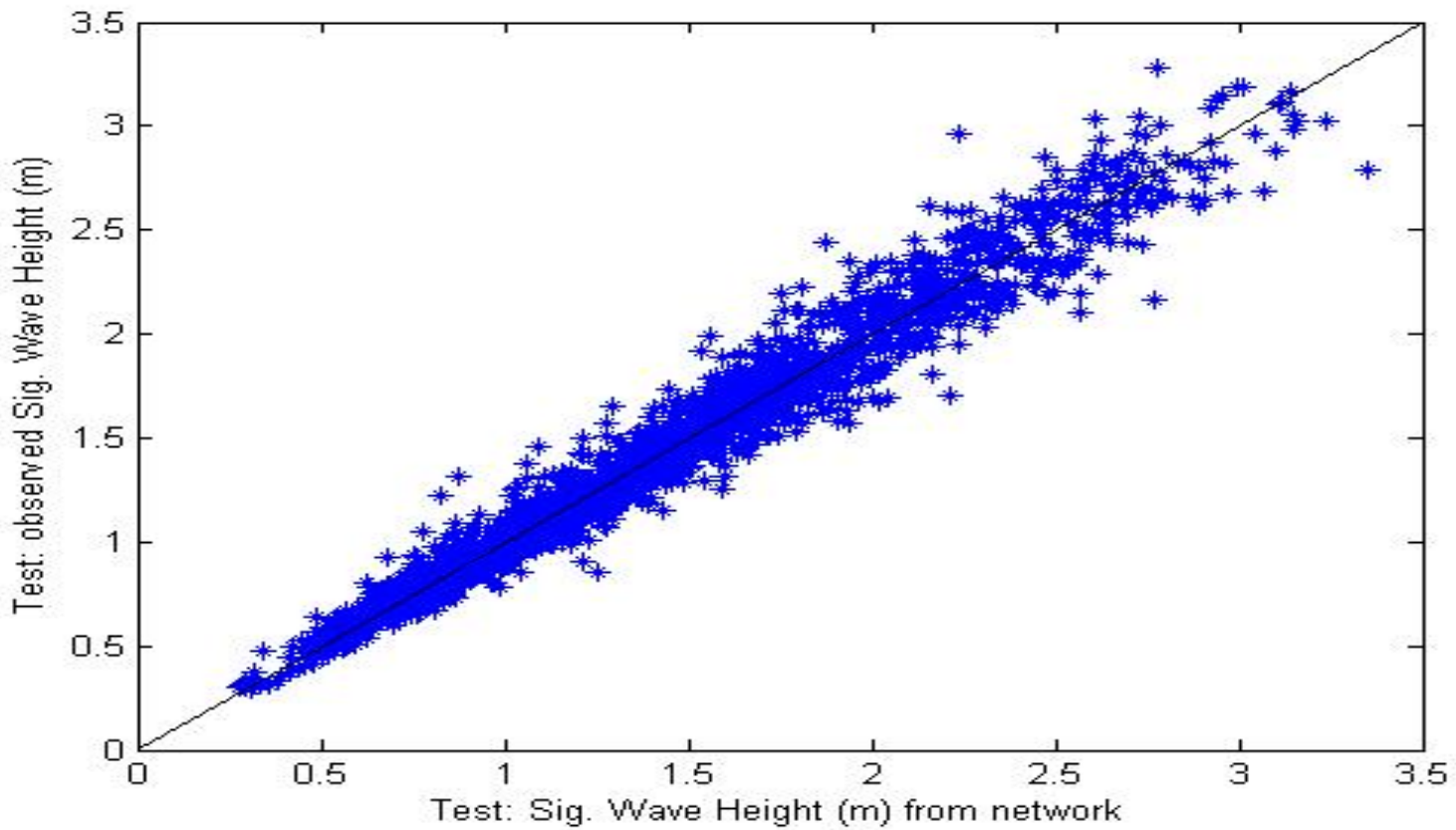


R = 0.9771, MSE = 0.0190 , MAE = 0.1006

Time History (3-3 net 1hr lead time)

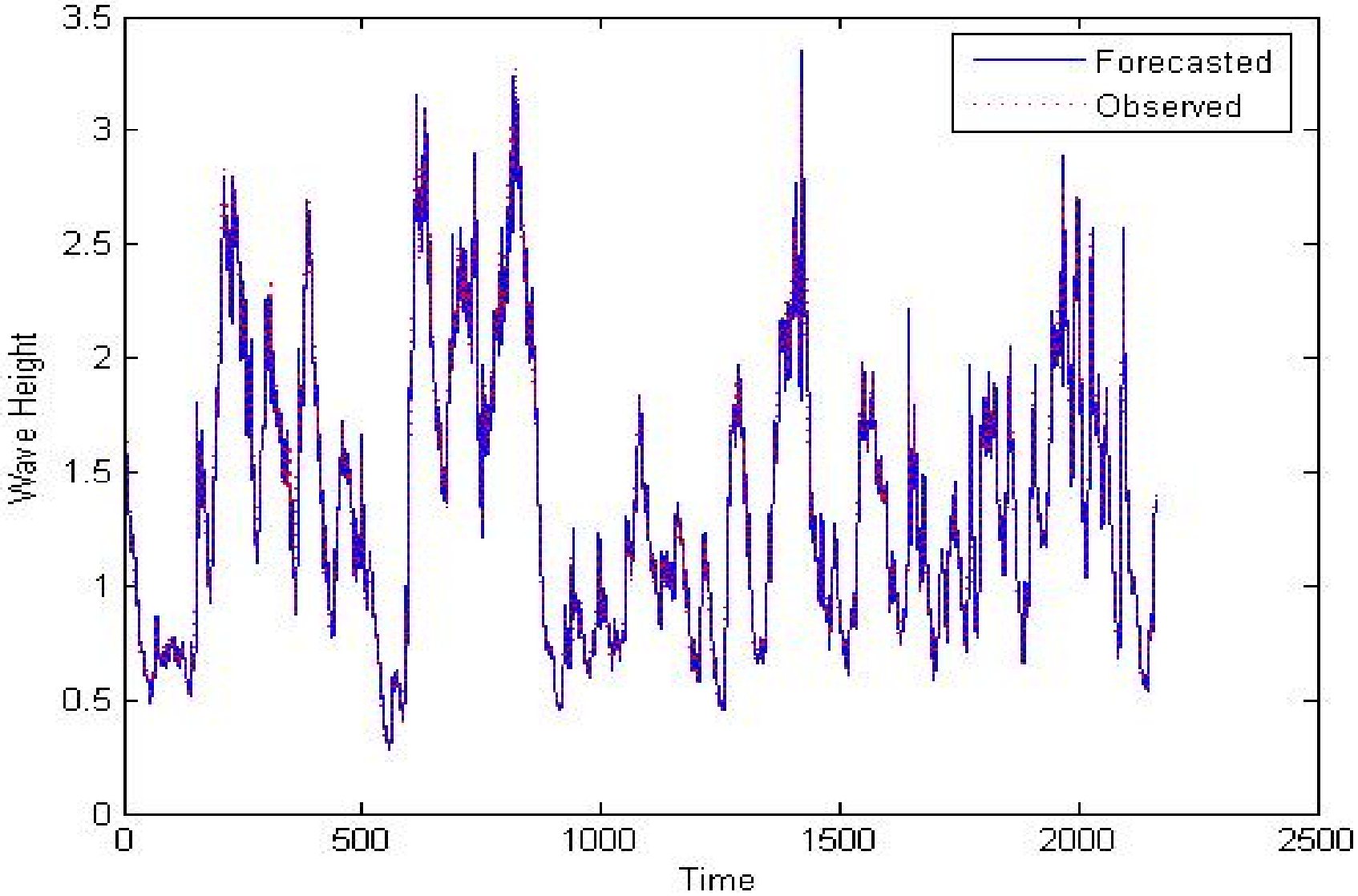


Scatter Diagram (6-6 net 1hr lead time)



R = 0.9828 , MSE = 0.0134 , MAE = 0.0817

Time History (6-6 net 1hr lead time)



Results comparison (R)

Lead Time (hr)	Network			
	3 - 3	**6 - 6**	**12 - 12**	**24 - 24**
1	0.9771	0.9828	0.9845	0.9800
2	0.9521	0.9720	0.9718	0.9669
3	0.9449	0.9548	0.9564	0.9521
4		0.9430	0.9379	0.9326
5		0.9245	0.9188	0.9122
6		0.9009	0.8969	0.8898

Results comparison (MSE)

Lead Time (hr)	Network			
	3 - 3	**6 - 6**	**12 - 12**	**24 - 24**
1	0.0190	0.0134	0.0118	0.0153
2	0.0457	0.0213	0.0214	0.0253
3	0.0452	0.0339	0.0328	0.0364
4		0.0425	0.0464	0.0507
5		0.0563	0.0601	0.0656
6		0.0727	0.0755	0.0814

Results comparison (MAE)

Lead Time (hr)	Network			
	3 - 3	**6 - 6**	**12 - 12**	**24 - 24**
1	0.1006	0.0817	0.0764	0.0872
2	0.1553	0.1007	0.1013	0.1118
3	0.1510	0.1291	0.1263	0.1339
4		0.1441	0.1511	0.1590
5		0.1669	0.1749	0.1850
6		0.1947	0.1989	0.2087

Methodology

- Previous training set was not an exhaustive one
- 3-3 and 6-6 networks were then chosen for further work
- Training Set – records of year 2003

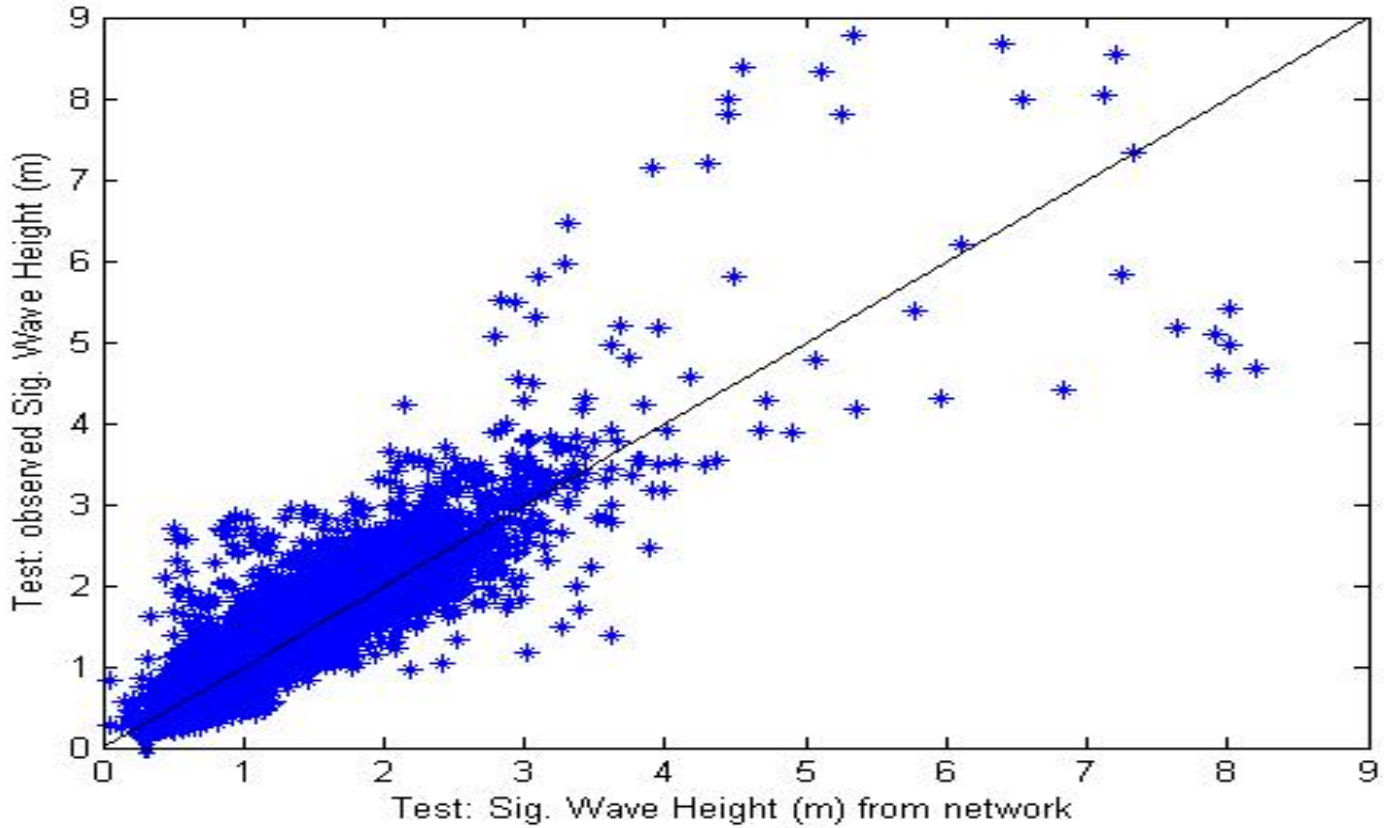
  Testing for year 2004
- Both networks were trained and tested for above data sets

Results Comparison

Parameter	3-3	6-6
R	0.9154	0.9428
MSE	0.0928	0.6306
MAE	0.1615	0.1359

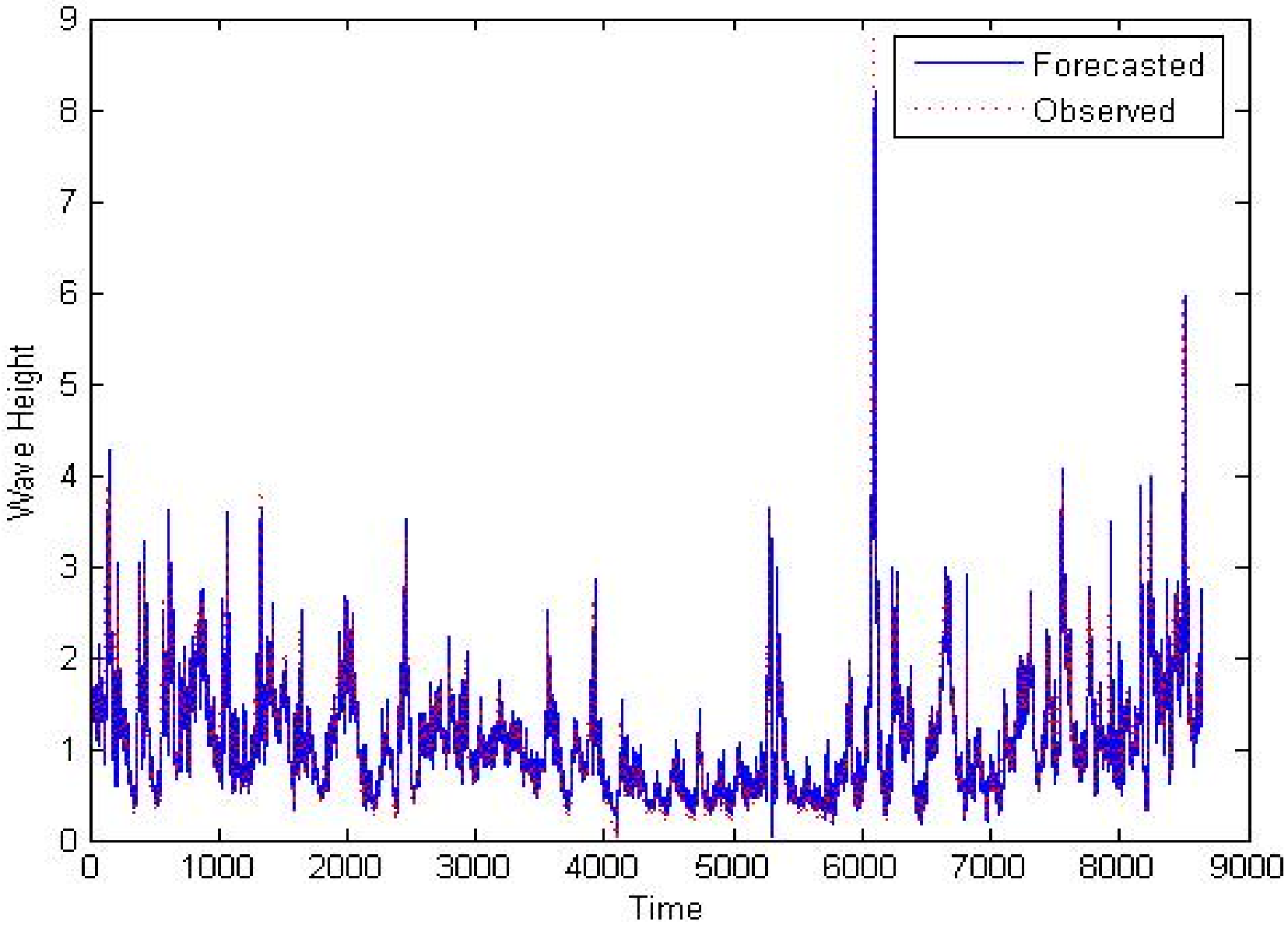
- 6-6 network produced better results
- It gives a good warning time
- Results of 6 hr lead time were also in acceptable limits

Scatter Diagram (6-6 net 6hr lead time)



R = 0.9072 , MSE = 0.0100 , MAE = 0.1928

Time History (6-6 net 6hr lead time)



Results comparison

Lead Time (hr)	Network	
	3 - 3	6 - 6
1	0.9286	0.9634
2	0.9303	0.9550
3	0.8874	0.9676
4		0.9473
5		0.9235
6		0.9072

'R'

Lead Time (hr)	Network	
	3 - 3	6 - 6
1	0.0787	0.0410
2	0.0786	0.0582
3	0.1213	0.0365
4		0.0583
5		0.0839
6		0.1004

'MSE'

Results comparison (MAE)

Lead Time (hr)	Network	
	3 - 3	6 - 6
1	0.1657	0.0874
2	0.1492	0.1046
3	0.1697	0.1192
4		0.1428
5		0.1688
6		0.1928

Methodology

- Out of four tried networks 6-6 network produced best results
- This network was then used for Selective testing
- Network was tested for 6 weeks of year 2004

Parameter	Value
R	0.9025
MSE	0.0971
MAE	0.1939

Results

- 6 input – 6 output nodes network was most efficient and accurate
- 12 hidden nodes in one hidden layer
- Weights and bias are tabulated
- Network gives a warning time of 6 hrs

Bias Values

Node	Bias
1	11.7090
2	-5.5844
3	-4.1450
4	4.5683
5	0.9203
6	4.1148
7	-0.7303
8	-1.0836
9	-1.0174
10	5.9476
11	0.1560
12	-7.0808

Node	Bias
1	-0.5314
2	0.8662
3	-0.8737
4	-0.4716
5	0.9991
6	-0.5760

Conclusion

- A neural network that uses input of 6 historical hourly Hs values to produce their hourly forecasts over the next 6 hours was found to be the most satisfactory.
- The deliverable of this project is the corresponding network configuration, complete with the matrix of trained weights and bias values, that can be used for online forecasting at the specified offshore site.
- The algorithm of RP yielded most adequate training.
- Forecasts were more accurate when the warning time was low.
- Future work may involve use of other network architecture like RBF, ANFIS in order to improve accuracy.

Thank You