PREDICTIONS OF SIGNIFICANT WAVE HEIGHT IN LAKE OKEECHOBEE, FLORIDA USING APPROACHES RELATED TO SIMPLIFIED STOCHASTIC PROCEDURE AND WAVE ENERGY SPECTRUM

A Thesis

Presented to

The Faculty of the Department of Civil and Environmental Engineering

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree

Masters of Science

in Civil Engineering

by

Ismat Tarin Khan

December 2012

PREDICTIONS OF SIGNIFICANT WAVE HEIGHT IN LAKE OKEECHOBEE, FLORIDA USING APPROACHES RELATED TO SIMPLIFIED STOCHASTIC PROCEDURE AND WAVE ENERGY SPECTRUM

Ismat Tarin Khan

Approved:

Chairman of the Committee
Keh-Han Wang, Professor
Civil and Environmental Engineering

Committee Members:

Jerry R. Rogers, Associate Professor,
Civil and Environmental Engineering

Kyle Strom, Assistant Professor
Civil and Environmental Engineering

Abdüsselam Altunkaynak,
Associate Professor, Civil Engineering
Istanbul Technical University, Turkey

Suresh K. Khator, Associate Dean,
Cullen College of Engineering

Abdeldjelil "DJ" Belarbi,
Professor and Chairman
Civil and Environmental Engineering

ACKNOWLEDGEMENTS

I wish to express my immeasurable and deepest gratitude to my graduate advisor, Dr. Keh-Han Wang, for his constant support and guidance throughout my graduate work. I am grateful to Dr. Jerry R. Rogers and Dr. Kyle Strom for serving on the committee and evaluating my thesis. I am also very grateful to Professor Abdüsselam Altunkaynak, who is a visiting scholar from Istanbul Technical University, Turkey for helping me in my thesis work and serving as committee member as well.

I am also indebted to my family for their constant encouragement and unwavering confidence in my ability which helped me achieve my academic dreams. Last but not the least, my utmost gratefulness to Almighty Allah, the most Merciful and Beneficent for successful completion of the work.

PREDICTIONS OF SIGNIFICANT WAVE HEIGHT IN LAKE OKEECHOBEE, FLORIDA USING APPROACHES RELATED TO SIMPLIFIED STOCHASTIC PROCEDURE AND WAVE ENERGY SPECTRUM

An Abstract

Of a

Thesis

Presented to

The Faculty of the Department of Civil and Environmental Engineering

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree

Masters of Science

in Civil Engineering

by

Ismat Tarin Khan

December 2012

ABSTRACT

Prediction of significant wave height is critically important to the physical and environmental impact study of coastal, estuarine or large lake environments. In this study, development of predictive models for the determination of time varying significant wave heights in Lake Okeechobee, Florida using the simplified stochastic procedure and wave energy spectrum method is presented.

The stochastic procedure related models are Regression Model 1 (RM1), Regression Model 2 (RM2) and Perceptron Least Square Method (PLSM). A new wave spectrum based model, Modified Pierson-Moskowitz (MPM) Spectrum is also developed. The predicted significant wave heights from each model are compared with the Artificial Neural Network (ANN) predictions obtained by Altunkaynak and Wang (2012). The comparisons between predicted significant wave heights from each model and observed data indicate that the proposed RM1, RM2, PLSM and MPM are effective models that are acceptable for predicting significant wave height in Lake Okeechobee.

TABLE OF CONTENTS

Acknowledgements……………………………………………………………………….iv

Abstract…………………………………………………………………………………...vi

Table of Contents………………………………………………………………………...vii

List of Figures…………………………………………………………………………….ix

List of Tables……………………………………………………………………………....xx

List of Symbols…………………………………………………………………………....xxii

Chapter 1 Introduction and Literature Review ……………………….........1

1.1 Introduction……………………………………………………………………….1

1.2 Regression Method (RM)..4

1.3 Wave Spectrum Methods..4

1.4 Modern Modeling Techniques……………………………………………………….7

1.4.1 Artificial Neural Network (ANN)………………………………………...7

1.4.2 Perceptron Neural Network (PNN)………………………………………10

1.4.3 Fuzzy Logic (FL)………………………………………………………...12

1.5 Contents of This Study……………………………………………………………..13

Chapter 2 Study Area and Data Collection…………………………………14

2.1 Study Area………………………………………………………………………….14

2.2 Data Collection……………………………………………………………………..17

2.3 Calculation of Significant Wave Height……………………………………………18

Chapter 3 Methodologies and Model Equations…………………………...19

3.1 Introduction…………………………………………………………………………19

3.2 Regression Method (RM)…………………………………………………………..20

3.3 Perceptron Least Square Method (PLSM)……………………………………36

3.4 Pierson-Moskowitz Spectrum (PM Spectrum)………………………………..43

3.5 Development of a new Wave Spectrum Model – Modified
Pierson-Moskowitz (MPM) Spectrum Model………………………………...50

3.6 Effects of time shifting of input variable on the performance of
model RM1 and RM2…………………………………………………………58

3.7 Summary………………………………………………………………………63

Chapter 4 Results and Comparison Discussions…………………………...64

4.1 Introduction…………………………………………………………………...64

4.2 Regression Method (RM1 and RM2)…………………………………………64

4.3 Perceptron Least Square Method (PLSM)……………………………………79

4.4 Modified Pierson Moskowitz (MPM) Spectrum Model………………………88

4.5 Comparison with ANN Model………………………………………………..98

4.5.1 Comparisons of RM1, RM2, PLSM, MPM and ANN at LZ40…………………99

4.5.2 Comparisons of RM1, RM2, PLSM, MPM and ANN at L006………………..102

4.5.3 Comparisons of RM1, RM2, PLSM, MPM and ANN at L005………………..105

4.5.4 Comparisons of RM1, RM2, PLSM, MPM and ANN at L001………………...108

Chapter 5 Conclusions and Future Studies………………………….........113

References………………………………………………………..................115

List of Figures

Figure 1.1 Simple architecture of ANN with a single input and output variable consisting of one hidden layer…………………………………...9

Figure 1.2 Perceptron configuration…………………………………………………10

Figure 1.3 Perceptron configuration of significant wave height and wind speed…...11

Figure 2.1 Satellite image of Lake Okeechobee (Internet 1)………………………..15

Figure 2.2 Lake Okeechobee and geographical locations of data collecting sites (L001, L005, L006 and LZ40) (Altunkaynak and Wang 2012)……16

Figure 3.1a Time variations of significant wave height and wind speed at LZ40 (year 2002 data)…………………………………………….....22

Figure 3.1b Correlation between significant wave height and wind speed at LZ40 (year 2002 data)…………………………………………………22

Figure 3.2a Time variations of significant wave height and wind speed at L006 (year 2002 data)...………………………………………...…..23

Figure 3.2b Correlation between significant wave height and wind speed at LZ40 (year 2002 data)………………………………………………...23

Figure 3.3a Time variations of significant wave height and wind speed at L005 (year 2002 data)…………………………………..……...…..24

Figure 3.3b Correlation between significant wave height and wind speed at LZ40 (year 2002 data)…………………………………….………..24

Figure 3.4a Time variations of significant wave height and wind speed at L001 (year 2002 data)…...……………………………………...…..25

Figure 3.4b Correlation between significant wave height and wind speed at LZ40 (year 2002 data)……………………………...……………….25

Figure 3.5 Predicted and observed time variations of significant wave heights at LZ40 using training data (RM1)……………………………………………27

Figure 3.6 Predicted and observed time variations of significant wave heights at L006 using training data (RM1)……………………….………………28

Figure 3.7 Predicted and observed time variations of significant wave heights at L005 using training data (RM1)………….…………………………28

Figure 3.8 Predicted and observed time variations of significant wave heights at L001 using training data (RM1) ………………………….………...29

Figure 3.9 Predicted and observed time variations of significant wave heights at LZ40 using training data (RM2)……………………………………………29

Figure 3.10 Predicted and observed time variations of significant wave heights at L006 using training data (RM2)…….……………………………………30

Figure 3.11 Predicted and observed time variations of significant wave heights at L005 using training data (RM2)……………………….………………30

Figure 3.12 Predicted and observed time variations of significant wave heights at L001 using training data (RM2)……………………………….……31

Figure 3.13 Verification of predicted and observed significant wave heights using training data (RM1, station LZ40)…...……………………………31

Figure 3.14 Verification of predicted and observed significant wave heights using training data (RM1, station L006)…………………………………32

Figure 3.15 Verification of predicted and observed significant wave heights using training data (RM1, station L005) ...………………………………32

Figure 3.16 Verification of predicted and observed significant wave heights using training data (RM1, station L001)……………………………………33

Figure 3.17 Verification of predicted and observed significant wave heights using training data (RM2, station LZ40) ...……………………………...33

Figure 3.18 Verification of predicted and observed significant wave heights using training data (RM2, station L006) ...………………………………34

Figure 3.19 Verification of predicted and observed significant wave heights using training data (RM2, station L005) ………………………………...34

Figure 3.20 Verification of predicted and observed significant wave heights using training data (RM2, station L001) ………………………...……35

Figure 3.21 Perceptron configuration of significant wave height and wind speed for Lake Okeechobee ..……………………………………..37

Figure 3.22 Predicted and observed time variations of significant wave heights at LZ40 using training data (Model PLSM) …………………………….38

Figure 3.23 Predicted and observed time variations of significant wave heights at L006 using training data (Model PLSM) ……….……………………39

Figure 3.24 Predicted and observed time variations of significant wave heights at L005 using training data (Model PLSM) ..……………………………39

Figure 3.25 Predicted and observed time variations of significant wave heights at L001 using training data (Model PLSM) ..……………………………40

Figure 3.26 Verification of predicted and observed significant wave heights using training data (Model PLSM, station LZ40) ………………………40

Figure 3.27 Verification of predicted and observed significant wave heights using training data (Model PLSM, station L006) ……………………….41

Figure 3.28 Verification of predicted and observed significant wave heights using training data (Model PLSM, station L005) ………….……………41

Figure 3.29 Verification of predicted and observed significant wave heights using training data (Model PLSM, station L001) ………………………42

Figure 3.30 Predicted and observed time variations of significant wave heights at LZ40 using PM Spectrum …………………………………………….46

Figure 3.31 Predicted and observed time variations of significant wave heights at L006 using PM Spectrum …………………………………………….46

Figure 3.32 Predicted and observed time variations of significant wave heights at L005 using PM Spectrum …………………………………………….47

Figure 3.33 Predicted and observed time variations of significant wave heights at L001 using PM Spectrum …………………………………………….47

Figure 3.34 Verification of predicted and observed significant wave heights using PM Spectrum at station LZ40 ……………………………………48

Figure 3.35 Verification of predicted and observed significant wave heights using PM Spectrum at station L006 ……………………………………48

Figure 3.36 Verification of predicted and observed significant wave heights using PM Spectrum at station L005 ……………………………………49

Figure 3.37 Verification of predicted and observed significant wave heights using PM Spectrum at station L001 ……………………………………49

Figure 3.38 Predicted and observed time variations of significant wave heights at LZ40 using Modified PM Spectrum ……………………………………54

Figure 3.39 Predicted and observed time variations of significant wave heights at L006 using Modified PM Spectrum ………………………..……………….54

Figure 3.40 Predicted and observed time variations of significant wave heights at L005 using Modified PM Spectrum …………………..………………….55

Figure 3.41 Predicted and observed time variations of significant wave heights at L001 using Modified PM Spectrum ……………………………………55

Figure 3.42 Verification of predicted and observed significant wave heights using Modified PM Spectrum at station LZ40 ………………………….56

Figure 3.43 Verification of predicted and observed significant wave heights using Modified PM Spectrum at station L006 ………………………….56

Figure 3.44 Verification of predicted and observed significant wave heights using Modified PM Spectrum at station L005 ………………………….57

Figure 3.45 Verification of predicted and observed significant wave heights using Modified PM Spectrum at station L001 …………………………..57

Figure 3.46 Time shifting effects on the performance of RM 1 (LZ40)………...……59

Figure 3.47 Time shifting effects on the performance of RM 1 (L006)…………..…..59

Figure 3.48 Time shifting effects on the performance of RM 1 (L005)……..……….60

Figure 3.49 Time shifting effects on the performance of RM 1 (L001)………...……60

Figure 3.50 Time shifting effects on the performance of RM 2 (LZ40)……….……..61

Figure 3.51 Time shifting effects on the performance of RM 2 (L006)……...………61

Figure 3.52 Time shifting effects on the performance of RM 2 (L005)…...……...….62

Figure 3.53 Time shifting effects on the performance of RM 2 (L001)…………...…62

Figure 4.1 Time variations of predicted and observed significant wave height using RM1 and RM2 (500 tested data applied to LZ40 – 2002) ….…….67

Figure 4.2 Time variations of predicted and observed significant wave height using RM1 and RM2 (500 tested data applied to L006 – 2002) ….……67

Figure 4.3 Time variations of predicted and observed significant wave height using RM1 and RM2 (500 tested data applied to L005 – 2002) ……….68

Figure 4.4 Time variations of predicted and observed significant wave height using RM1 and RM2 (500 tested data applied to L001 – 2002) ….…….68

Figure 4.5 Time variations of predicted and observed significant wave height using RM1 and RM2 (140 data applied LZ40 – 1996) ……….…………69

Figure 4.6 Time variations of predicted and observed significant wave height using RM1 and RM2 (140 data applied L006 – 1996) …….….………69

Figure 4.7 Time variations of predicted and observed significant wave height using RM1 and RM2 (140 data applied L005 – 1996) ……….…….…70

Figure 4.8 Verification of predicted and observed significant wave height using RM1 (500 testing data applied LZ40 – 2002) ……………...……..71

Figure 4.9 Verification of predicted and observed significant wave height using RM2 (500 testing data applied LZ40 – 2002) ……………………71

Figure 4.10 Verification of predicted and observed significant wave height using RM1 (500 testing data applied L006 – 2002) …...………………..72

Figure 4.11 Verification of predicted and observed significant wave height using RM2 (500 testing data applied L006 – 2002) ……...……………..72

Figure 4.12 Verification of predicted and observed significant wave height using RM1 (500 testing data applied L005 – 2002) ………...…………..73

Figure 4.13 Verification of predicted and observed significant wave height using RM2 (500 testing data applied L005 – 2002) ……………...……..73

Figure 4.14 Verification of predicted and observed significant wave height using RM1 (500 testing data applied L001 – 2002) ……………...……..74

Figure 4.15 Verification of predicted and observed significant wave height using RM2 (500 testing data applied L001 – 2002) …………………….74

Figure 4.16 Verification of predicted and observed significant wave height using RM1 (140 data applied LZ40 – 1996) …………………………….75

Figure 4.17 Verification of predicted and observed significant wave height using RM2 (140 data applied LZ40 – 1996) …………………...……..75

Figure 4.18 Verification of predicted and observed significant wave height using RM1 (140 data applied L006 – 1996) ………...…………………..76

Figure 4.19 Verification of predicted and observed significant wave height using RM2 (140 data applied L006 – 1996) ……………...……………..76

Figure 4.20 Verification of predicted and observed significant wave height using RM1 (140 data applied L005 – 1996) ………...…………………..77

Figure 4.21 Verification of predicted and observed significant wave height using RM2 (140 data applied L005 – 1996) …………………...………..77

Figure 4.22 Time variations of predicted and observed significant wave height using PLSM (500 tested data applied to LZ40 – 2002) …………………80

Figure 4.23 Time variations of predicted and observed significant wave height using PLSM (500 tested data applied to L006 – 2002) …………………80
Figure 4.24 Time variations of predicted and observed significant wave height using PLSM (500 tested data applied to L005 – 2002) …………………81
Figure 4.25 Time variations of predicted and observed significant wave height using PLSM (500 tested data applied to L001 – 2002) …………………81
Figure 4.26 Verification of predicted and observed significant wave height using PLSM (500 testing data applied LZ40 – 2002) …………………..82
Figure 4.27 Verification of predicted and observed significant wave height using PLSM (500 testing data applied L006 – 2002) …………………..82
Figure 4.28 Verification of predicted and observed significant wave height using PLSM (500 testing data applied L005 – 2002) …………………..83
Figure 4.29 Verification of predicted and observed significant wave height using PLSM (500 testing data applied L001 – 2002) …………………..83
Figure 4.30 Time variations of predicted and observed significant wave height using PLSM (140 data applied to LZ40 – 1996) ………………………..84
Figure 4.31 Time variations of predicted and observed significant wave height using PLSM (140 data applied to L006 – 1996) ………………………..84
Figure 4.32 Time variations of predicted and observed significant wave height using PLSM (140 data applied to L005 – 1996) ………………………..85
Figure 4.33 Verification of predicted and observed significant wave height using PLSM (140 data applied LZ40 – 1996) …………………………..85

Figure 4.34 Verification of predicted and observed significant wave height using PLSM (140 data applied L006 – 1996) …………………………..86

Figure 4.35 Verification of predicted and observed significant wave height using PLSM (140 data applied L005 – 1996) …………………………..86

Figure 4.36 Time variations of predicted and observed significant wave height using MPM (500 testing data applied to LZ40 – 2002) ………………..90

Figure 4.37 Time variations of predicted and observed significant wave height using MPM (500 testing data applied to L006 – 2002) ………………..90

Figure 4.38 Time variations of predicted and observed significant wave height using MPM (500 testing data applied to L005 – 2002) ………………..91

Figure 4.39 Time variations of predicted and observed significant wave height using MPM (500 testing data applied to L001 – 2002) ………………..91

Figure 4.40 Verification of predicted and observed significant wave height using MPM (500 testing data applied to LZ40 – 2002) ………………..92

Figure 4.41 Verification of predicted and observed significant wave height using MPM (500 testing data applied to L006 – 2002) ………………..92

Figure 4.42 Verification of predicted and observed significant wave height using MPM (500 testing data applied to L005 – 2002) ………………..93

Figure 4.43 Verification of predicted and observed significant wave height using MPM (500 testing data applied to L001 – 2002) …………..……93

Figure 4.44 Time variations of predicted and observed significant wave height using MPM (140 data applied to LZ40 – 1996) ……………………….94

Figure 4.45 Time variations of predicted and observed significant wave height using MPM (140 data applied to L006 – 1996) ………………………..94

Figure 4.46 Time variations of predicted and observed significant wave height using MPM (140 data applied to L005 – 1996) ………………………..95

Figure 4.47 Verification of predicted and observed significant wave height using MPM (140 data applied to LZ40 – 1996) ……………………….95

Figure 4.48 Verification of predicted and observed significant wave height using MPM (140 data applied to L006 – 1996) ………………………..96

Figure 4.49 Verification of predicted and observed significant wave height using MPM (140 data applied to L005 – 1996) ………………………..96

Figure 4.50 Comparison chart of time variations for RM1, RM2, PLSM, MPM and ANN (500 testing data of year 2002) ……………………...100

Figure 4.51 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (LZ40 – 140 data of year 1996) ……………………..101

Figure 4.52 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (L006 - 500 testing data of year 2002) ………………103

Figure 4.53 Comparison charts of time variations RM1, RM2, PLSM, MPM and ANN (L006 – 140 data of year 1996) ……………………..104

Figure 4.54 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (L005 - 500 testing data 2002) ……………………..106

Figure 4.55 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (L005 – 140 data 1996) …………………………….107

Figure 4.56 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (L001 - 500 testing data 2002) ……………………..109

List of Tables

Table 3.1 Calculated parameters of RM1 and RM2 for the station LZ40, L006, L005 and L001……………………….…………………….27

Table 3.2 Summary of AAE, RMSE and CE of Models RM1 and RM2 using training data for the stations LZ40, L006, L005 and L001 ……………………………………………………..…..36

Table 3.3 Network weights of Model PLSM for the station LZ40, L006, L005 and L001…………………………………………………..….38

Table 3.4 Values of AAE, RMSE and CE of model PLSM using training data for the station LZ40, L006, L005 and L001 ………………………..…...42

Table 3.5 Values $\boldsymbol{\alpha_1}$ and $\boldsymbol{\beta_1}$ for the stations LZ40, L006, L005 and L001………………………………………………….…..…52

Table 3.6 Values of AAE, RMSE and CE of MPM using training data for the station LZ40, L006, L005 and L001 ……………………………...58

Table 4.1 Summary of AAE, RMSE and CE of RM1 and RM2 using 500 testing data collected in year 2002 for the stations LZ40, L006, L005 and L001 ……………………………………………….....78

Table 4.2 Summary of AAE, RMSE and CE of RM1 and RM2 using 140 independent data testing data collected in year 1996 for the stations LZ40, L006 and L005 …………….……………………78

Table 4.3 Values of AAE, RMSE and CE of PLSM using 500 testing data of year 2002 for the stations LZ40, L006, L005 and L001 ……………..87

Table 4.4 Values of AAE, RMSE and CE of PLSM using 140 data (year 1996) for the stations LZ40, L006 and L005……………..……………..……..88

Table 4.5 Values of AAE, RMSE and CE of MPM using the 500 testing data of year 2002 for the stations LZ40, L006, L005 and L001 ……….……..97

Table 4.6 Values of AAE, RMSE and CE of MPM using 140 data of year 1996 for the stations LZ40, L006 and L005……...………………………………..97

Table 4.7 Summary of AAE, RMSE and CE for RM1, RM2, PLSM, PMP and ANN for 500 testing data of year 2002………………...……………….111

Table 4.8 Summary of AAE, RMSE and CE for RM1, RM2, PLSM, PMP and ANN for the year 1996 data…………...………………………………..112

List of Symbols

a and b	Model parameters
f	Wave frequency
g	Gravitational acceleration constant (9.81 m/s^2)
H_s	Significant wave height
H_{rms}	Root-mean-squared wave height
H_k, H_{k-1}	Significant wave height at a given time k and k-1
$S(f)$	Wave spectrum
W	Wind speed
W_k, W_{k-1}	Wind speed at a given time k and k-1
W_f	Friction velocity
X	Fetch
α	Phillips' constant (8.1x10^{-3})
υ	Kinetic viscosity of water at 20^o C (1.004 x 10^{-6} m^2/s)

Chapter 1

Introduction and Literature Review

1.1 Introduction

Prediction of the wave field, especially wind generated waves, is critically important to the physical and environmental impact study of coastal, estuarine, or large lake environments. Physical and ecological processes in the lakes and estuaries are known to depend on winds, waves, tides, fresh water inflows, human activities, etc. (Chen et al. 2005). Studies on Lake Okeechobee in Florida (Olila and Reddy, 1993; Reddy et al. 1995) have indicated that the internal loading of phosphorus from lake bottom sediments has become a major source of phosphorus to the water column.

Wind generated surface waves are the main source for re-suspension of bottom sediments and accordingly have a considerable impact on the water quality of the water bodies (Jin and Wang, 1998; Tehrani, 2001; Altunkaynak and Wang, 2012). To be able to reasonably estimate the re-suspended bottom sediment concentration in the water column, an accurate prediction on the wave properties, especially the time variation of a representative wave height, e.g., significant wave height is required (Altunkaynak and Wang, 2012). Despite the concern of the water quality, knowing the wave field and wave properties is also important to the design of coastal and offshore structures and other engineering works (McCormick, 2009).

Different from regular monochromatic waves, wind generated waves have the properties of randomness and irregularity and consist of waves with different wave heights and frequencies. Statistically defined wave heights or wave periods have been developed to characterize the random wave field. One of the most commonly adopted wave heights for the design and wave analysis purpose is the significant wave height. In terms of wave height, ordering the wave data ($\boldsymbol{N}$ waves) from the largest to the smallest with assigned number from 1 to $\boldsymbol{N}$, the average of the first highest one third ($\boldsymbol{N/3}$) waves is defined as the significant wave height ($\boldsymbol{H_s}$). This direct count approach can be expressed as

$$\boldsymbol{H_s = \frac{1}{N/3}\sum_{i=1}^{N/3} H_i} \, . \tag{1.1}$$

The other approach is to assume that the wave height probability density function satisfies the narrow banded Rayleigh distribution. Then, the significant wave height can also be determined using the following formulation

$$\boldsymbol{H_s = \sqrt{2}H_{rms} = \sqrt{2}\sqrt{\frac{1}{N}\sum_{i=1}^{N} H_i^2}} \, , \tag{1.2}$$

where $\boldsymbol{H_{rms}}$ is defined as the root-mean-squared wave height.

As the wind speed (or wind stress) is shown to have strong correlation with the significant wave height, the first empirical approach using wind speed as a major input parameter was presented by Sverdrup and Munk (1947). Later, comprehensive numerical modeling with inputs of wind speed was adopted to forecast the significant wave height. Numerical models became popular the past decades with the advancement of high speed computers and their ability to perform more crucial computations (Liu et al. 2002). These models were mostly based on the solutions of energy balance equations in finite difference form throughout a grid placed over the water area (Kazeminezhad et al. 2005). To predict reliable and accurate results using numerical models, a variety of meteorological data and high speed computers are required (Goda, 2003). Schwab et al. (1991) developed a wind wave model by applying several frequently used wind wave models to a case of active wave generation and growth in Lake Michigan and then compared the results with actual measurements (Liu et al. 2002). Other researchers like, Hasselmann (1962), Barnett (1968), Booji et al. (1999) and Chen et al. (2005) proposed numerical models using energy transfer equations to predict better wind generated waves. Jin and Wang (1998) developed a Boussinesq-type, three-equation wind wave model to simulate wind waves and determine the corresponding significant wave heights in Lake Okeechobee. A calibration and verification study of a spectrum based wind wave model SWAN (Simulation WAves Nearshore) was presented by Jin and Ji (2001). The issues related to wave growth and decay characteristics in SWAN was investigated by Rogers et al. (2003).

Even though numerical models are capable of delivering detailed temporal and spatial variation of wind induced wave elevation (Altunkaynak and Wang, 2012), these models may not be justified to use for preliminary or even for final design in some cases from the economical point of view (Goda, 2003). Simple and effective predictive models can be useful for forecasting significant wave height and practical applications. Those models include the use of stochastic based regression approaches, the development of fitted wave spectrum formulas, and data based modern modeling technology.

1.2 Regression Method (RM)

Regression Method (RM) is one of the simple methods that can be used in wave forecast studies. With noticed strong correlation between significant wave height and wind speed, a linear regression model for significant wave height and wind speed can be formulated as

$$\boldsymbol{H_K = a\,W_K + b}\,, \tag{1.3}$$

where $\boldsymbol{H_k}$ and $\boldsymbol{W_k}$ represent the significant wave height and wind speed at a given time $\boldsymbol{k}$ respectively. $\boldsymbol{a}$ and $\boldsymbol{b}$ are the model parameters to be calibrated.

There are some basic restrictive assumptions embedded in the regression approach (Şen et al. 2003; Uyumaz et al. 2006 and Altunkaynak, 2008). The following are some of the main assumptions (Şen et al. 2003):

i. Linearity: A straight line trend through the scatter of data points.

ii. Normality: The residuals have normal distributions.

iii. Means of conditional distribution: The expected mean of the difference between the measured and modeled values must be zero

iv. Homoscedasticity: The residuals have equal variance.

v. Lack of measurements errors: The previous and current measurements both are without errors.

Most of the time, the system dynamic of the regression method is restricted by a deterministic expression, as a least square method is used to minimize the sum of squared errors for parameter estimations (Şen et al. 2004; Altunkaynak and Wang, 2012). Even though the regression method has a few drawbacks, still it can provide acceptable results. From the economical point of view, estimating significant wave heights using regression method is suitable for the forecast and preliminary design purposes.

1.3 Wave Spectrum Methods

Development of wave spectrum formulas for characterizing the distribution of wave energy and the corresponding significant wave height of the random wave filed in in a water body is an empirical based approach. Those empirical expressions were formulated using correlations between dimensionless wave parameters and wind variables. The effectiveness associated evaluation time and cost of these methods is quite reasonable (Altunkaynak, 2008). Typical wave spectrum formulas applied in the ocean or enclosed large lakes are:

i. SMB (Sverdrup, Munk & Brestchneider) (Brestchneider,1970,1973)

ii. JONSWAP (Joint North Sea Wave Project) (Hasselmann et al. 1973)

iii. Donelan (Donelan, 1980; Donelan et al. 1985)

iv. Pierson-Moskowitz (PM) Spectrum (Pierson and Moskowitz, 1964)

Each of the above mentioned spectra was developed by fitting the collected wave data with a proposed frequency related wave energy formula. Both the significant wave height and wave period can be determined using SMB, JONSWAP and Donelan. Fetch length and other meteorological data such as wind speed and duration are required to estimate significant wave height and wave period. For the referred wave spectrum models, it is assumed that the direction of wind and wave is the same (Altunkaynak 2008). Accuracy of these methods was tested by some researchers at different occasions. According to Bishop (1983), Donelan model provides more accurate results than JONSWAP and SMB models. Another positive aspect of Donelan model is that it also provides the direction of estimated peak wave energy (Bishop, 1983).

Pierson-Moskowitz Spectrum is a widely used frequency spectrum for fully developed wind waves proposed by Pierson and Moskowitz (1964). PM spectrum is based on the power spectra for fully developed seas by Kitaigorodskii (1961). According to the assumptions of Kitaigorodskii (1961), the spectrum is a function of four variables only. Thus, the spectrum can be expressed as

$$\boldsymbol{S(f) = F\left(f, g, W_f, X\right)}, \tag{1.4}$$

where $\boldsymbol{f, g, W_f}$, and $\boldsymbol{X}$ are wave frequency, gravitational acceleration constant, friction velocity and fetch, respectively. Further modified by Schmitz (1962), Equation (1.4) became

$$\boldsymbol{S(f) = F\,(f, g, W, X)}, \tag{1.5}$$

where $\boldsymbol{W}$ is the wind speed. As the data collected by Moskowitz (1963) were for fully developed seas, fetch was vanished from the equation,

$$\boldsymbol{S(f) = F\,(f, g, W)}. \tag{1.6}$$

The spectrum was further modified by Pierson and Moskowitz (1963) by fitting the data with the Neumann Spectrum (Neumann, 1952) where -5 power laws for wave frequency in the equilibrium range proposed by Phillip (1958) was adopted (Liu et al. 2011; Pierson and Moskowitz, 1963). The final form of the P-M Spectrum is

$$\boldsymbol{S(f) = \frac{\alpha\, g^2}{16\, \pi^4\, f^5}\; e^{[\frac{-B}{f^4}]}}\,, \tag{1.7a}$$

where

$$\boldsymbol{B = \beta\, [\frac{g}{2\,\pi\, W}]^{\,4}}\;, \tag{1.7b}$$

$\boldsymbol{\alpha}$ (Phillips' constant) = **8.1x10^{-3}**, $\boldsymbol{\beta}$ = **0.74**, $\boldsymbol{g}$ (gravitational acceleration) = **9.81 m/s^2** and $\boldsymbol{W}$ is the wind speed measured in m/s. The integration of spectrum (1.7a) can be related to the significant wave height. Detail formulations are given in Chapter 3.

Considering the fetch limited statistics, the approach led to as named Joint North Sea Wave Project (JONSWAP) spectrum was established and reported by Hasselmann et al. (1973). The spectrum form was modified from PM spectral representation by adding the scale and shape related parameters as a five-parameter wave spectral density formula. The additional fetch related parameters modify the width and peak of the original PM spectrum. Different from the PM spectrum, the mathematical formulation of JONSWAP spectrum prevents it from obtaining an explicit expression of the integrated quantity of the spectrum for relating to the significant wave height. The other discussions on the JONSWAP spectrum can be found in Goda (1985), Hogben (1990), and McCormick (1999).

1.4 Modern Modeling Techniques

In recent years modern modeling techniques became very popular in the field of data analysis and provide simpler methods for mapping input variables to outputs (Uyumaz et al. 2006). Modern modeling methods are applicable to any field of studies where data are available for analysis. In study of ocean engineering, studies have been carried out by applying the data based modern modeling techniques to predict wave parameters. Some of the commonly used methods are:

i. Artificial Neural Network (ANN)
ii. Perceptron Neural Network (PNN)
iii. Fuzzy Logic (FL)

1.4.1 Artificial Neural Network (ANN)

Artificial Neural Network (ANN) provides non-deterministic and model free mapping between a given set of input and output values. Since the occurrence of wave is a random and unpredictable phenomenon, many researchers find this method more suitable to predict wind wave data (Deo and Kumar, 2000). Neural network modeling is primarily aimed to recognize the random pattern in a given set of input data and use the same to predict the desired property. To apply the neural network model, it is not required to have the knowledge regarding the underlying physical process (Deo et al. 2001).

The ANN method is used to forecast wave parameters by Deo et al. (2001), Agrawal and Deo (2002) and Tsai et al. (2002). Makarynskky (2004) extended the ANN approach to focus on the prediction of significant wave height. Modeling the significant wave heights in Lake Superior using ANN was presented by Etemad-Shahidi and Mahjoobi (2009). Altunkaynak and Wang (2012) also applied the commonly used three-layer ANN to predict the wind induced significant wave heights in Lake Okeechobee.

A simple three layered Artificial Neural Network is shown in Figure 1.1, where a single input node receives input data and pass them on to the hidden layer nodes. Inside the hidden layer, each node (neuron) includes the process of multiplying the input data with the weighting coefficients and summing them up. Each neural network has a threshold value. This threshold value is added to the summed results and then passed through a non-linearity transfer function to the output node (Deo and Kumar, 2000; Etemad-Shahidi and Mahjoobi, 2009).

In the neural network system, there is no limit on the number of input/ output or hidden layers. A particular network can have as many layers of input/output and hidden layers as required by the specific problem. Once the topology of the neural network is fixed, a training procedure is set up with the input and output data to determine the network weights and threshold. The aim of the training procedure is to reduce the error, ***E*** between the actual observed data and the network output (Etemad-Shahidi and Mahjoobi 2009, Deo et al. 2001).

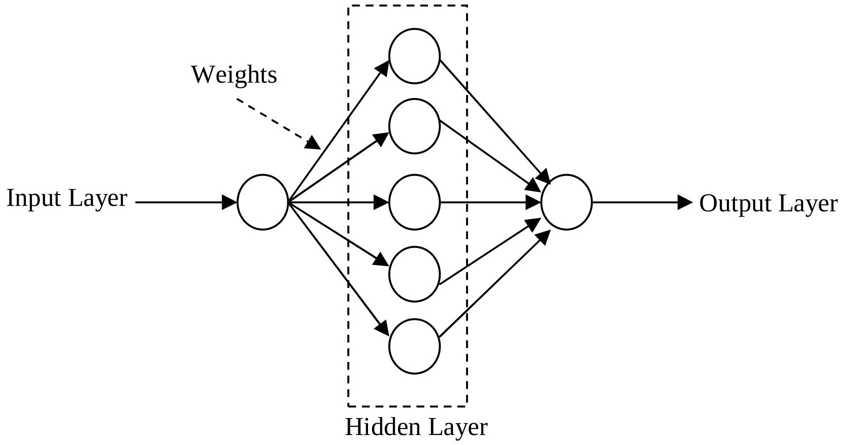

Figure 1.1 Simple architecture of ANN with a single input and output variable consisting of one hidden layer

The error, $\boldsymbol{E}$ can be defined as

$$\boldsymbol{E} = \sum(\boldsymbol{O_n} - \boldsymbol{O_t})^2, \tag{1.8}$$

where $\boldsymbol{O_n}$ and $\boldsymbol{O_t}$ are the network and target output respectively at the same node. All the output nodes are summed together for a given training pattern and then all over training patterns (Deo et al. 2001).

Different algorithms are available in the literature for the training procedure in the Neural Network, such as back propagation (Rumelhart et al. 1986; Altunkaynak and Wang, 2012), cascade correlation and conjugate gradient, etc. Back propagation method minimizes the error using steepest descent or the gradient descent approach. The network weights and threshold are adjusted by a small step towards negative gradient of the error function during iteration and repeated until a given number of iterations are done. In the cascade correlation, the weights between the layers of nodes are optimized using gradient ascent method where the correlation between output of the hidden layer and the residual error of the network is maximized. For the conjugate gradient approach to find the weights and threshold, the gradient descent is made along the conjugate or orthogonal direction from the previous step. Detailed mathematical discussion can be found in the literature by Fletcher and Reeves (1964), Rabelo (1990), Fahlman and Lebiere (1990), Yeh et al. (1993), and Deo et al. 2001.

1.4.2 Perceptron Neural Network (PNN)

Perceptron Neural Network (PNN) is the system similar to the Artificial Neural Network. Like Artificial Neural Network, Perceptron Network has the input and output layers. However, the PNN does not have any hidden layers within the network (Figure 1.2) (Holland 1975). Input and output layers are connected directly through the transition matrix weights. This structure forms a linear relationship between the input and output layers. Similar to ANN, the network weights can be calculated by the back propagation procedure.

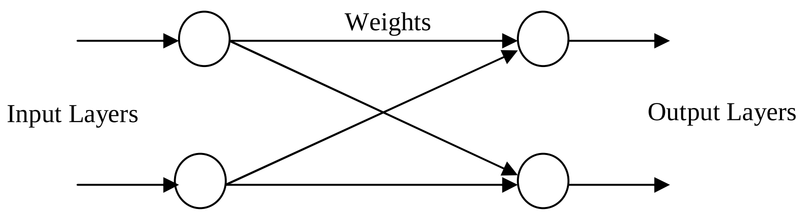

Figure 1.2 Perceptron configuration

Perceptron Neural Network can also be applied with other techniques and thus get a better result. Altunkaynak and Özger (2004) proposed a new concept where Perceptron and Kalman Filtering (KF) method is combined to predict significant wave height and named Perceptron Kalman Filtering (PKF). Şen et al. 2004 applied this combined method to predict sediment concentration. Kalman Filtering is an adaptive modeling technique of state variables (Altunkaynak and Özger, 2004) and also referred as modern least squares

(Şen, 1980; 1984) method. This method is capable of making predictions of past, present and even future states. Details on the Kalman Filtering techniques can be found in the literature by Kalman (1960) and Gelb (1974).

Prediction of significant wave height using PKF method is achieved by two steps. In the first step, a relationship between significant wave height and wind speed is derived using neural network system. Figure 1.3 shows the neural construction of significant wave height ($\boldsymbol{H}$) and wind speed ($\boldsymbol{W}$). In the network (Figure 1.3) $\boldsymbol{H}$ and $\boldsymbol{W}$ from previous time step ($\boldsymbol{k-1}$) are connected to the $\boldsymbol{H}$ and $\boldsymbol{W}$ of current time step ($\boldsymbol{k}$) by transition matrix weights $\boldsymbol{a_{11}}$, $\boldsymbol{a_{12}}$, $\boldsymbol{a_{21}}$ and $\boldsymbol{a_{22}}$. Same weights are used to connect $\boldsymbol{H}$ and $\boldsymbol{W}$ of current time step ($\boldsymbol{k}$) to the $\boldsymbol{H}$ and $\boldsymbol{W}$ of future time step ($\boldsymbol{k+1}$) for forecasting. Transition matrices are then determined by back propagation technique. Once the matrices are found, Kalman Filtering is used to predict significant wave height (Altunkaynak and Özger, 2004; Altunkaynak and Wang, 2012).

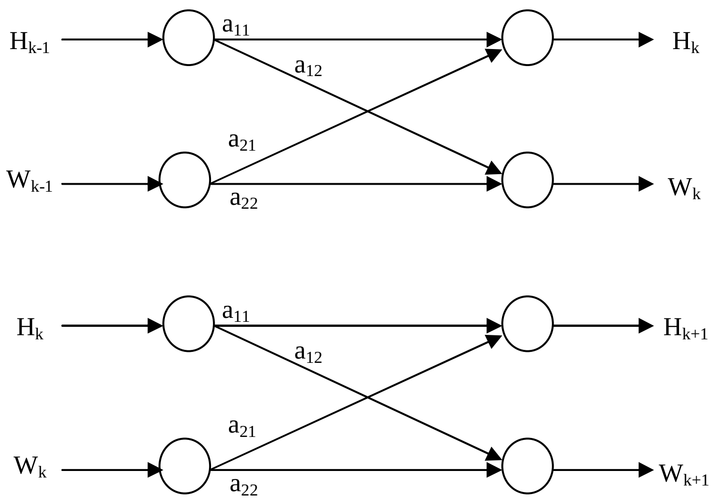

Figure 1.3 Perceptron configuration of significant wave height and wind speed

1.4.3 Fuzzy Logic (FL)

In the Fuzzy Logic (FL), a linguistics expressions is used which is the approach of ambiguity rather than numerical probabilistic or statistical approach. Fuzzy logic was originally presented by Zadeh in 1965. After that, many researchers adopted this method in various engineering problems (Mamdani, 1974; Pappis and Mamdani, 1977; Ross, 1995; Xiong et al. 2001; Şen and Altunkaynak, 2004; Uyumaz et al. 2005 ; Wang and Altunkaynak, 2012). Fuzzy logic approach applied particularly to ocean engineering filed for the prediction of wave parameters includes studies by Kazemiezhad et al. (2005) and Özger and Şen (2007).

According to Ross (1995), generally, three steps can be followed to develop a fuzzy logic model. The first step is “fuzzification” where selection of membership functions for the input and output variables and construction of fuzzy rules using the form of IF-THEN statement are conducted. The second step as referred as “inference” is to evaluate the membership degrees (antecedent) within the fuzzy rules of the input variables based on the max-min method. “Defuzzification” is the third or the final step to compute the outputs of the physical variables using the antecedent and consequent values obtained from each fuzzy set. As pointed out by Wang and Altunkaynak (2012), fuzzy logic is capable of relating the event total between input and output variables. With the limitation of the methodology, the fuzzy logic is generally unable to generate time varying outputs. Therefore, the fuzzy logic is not a recommended method to be used for prediction of time varying significant wave height.

1.5 Contents of This Study

Significant wave height prediction is one of the most important aspects in the study of ocean engineering. There are many different approaches available in the literature. Chapter 1 includes introduction and literature reviews which cover the details about the preferred methods for predicting wave parameters. Detailed information regarding the Lake Okeechobee and data collections are included in the study area and data collection of Chapter 2. The main focus of this study is the development of a model for the significant wave height using the concepts of Regression Method, Perceptron Neural Network (PNN) and Pierson-Moskowitz Spectrum. Even though linear regression has limitations, still it can be applied to predict significant wave heights and produces acceptable results.

A simplified approach of Perceptron Neural Network is Perceptron Least Square Method is also used to predict significant wave height. Then a new approach is developed by modifying the Pierson-Moskowitz Spectrum for Lake Okeechobee and the predicted significant wave heights are compared with those from other models used in this study. The above mentioned model approaches are described in Chapter 3. In Chapter 4, comparisons between the results obtained from the developed models and the data are presented. Comparisons are also made with the results of ANN model by Altunkaynak and Wang (2012). The error analyses to reflect the performance of each proposed model are also provided in Chapter 4. Conclusions and recommendation of future works are summarized in the Chapter 5.

Chapter 2

Study Area and Data Collection

2.1 Study Area

Lake Okeechobee is located in south-central Florida between 26º 41´ and 27 º 12´ North and 80º 36´ and 81º 05´ West (Jin and Wang, 1998; Altunkaynak and Wang, 2012). The Lake is at the center of South Florida's regional water management system. Lake Okeechobee is the seventh largest fresh water lake in the United States and the second largest freshwater lake contained entirely within the lower 48 states, behind only one of the Great Lakes- Lake Michigan. It is a relatively shallow lake with the area of 730 square miles (1900 km^2) and with an average depth of 9 feet (2.7 meters). Drainage basin of this massive lake covers more than 4600 square miles (11,913 km^2) (Tehrani, 2001). A satellite image of Lake Okeechobee is shown in Figure 2.1.

Lake Okeechobee and its watershed are integral components of South Florida's Kissimmee-Okeechobee-Everglades ecosystem. Kissimmee River is one of the main sources of the water coming to the Lake Okeechobee (Jin and Wang, 1998). The major outflows of the lake are Caloosahatchee River to the west, the St. Lucie Canal to the east and the Everglades Agricultural Area (Figure 2.2) (Tehrani, 2001).

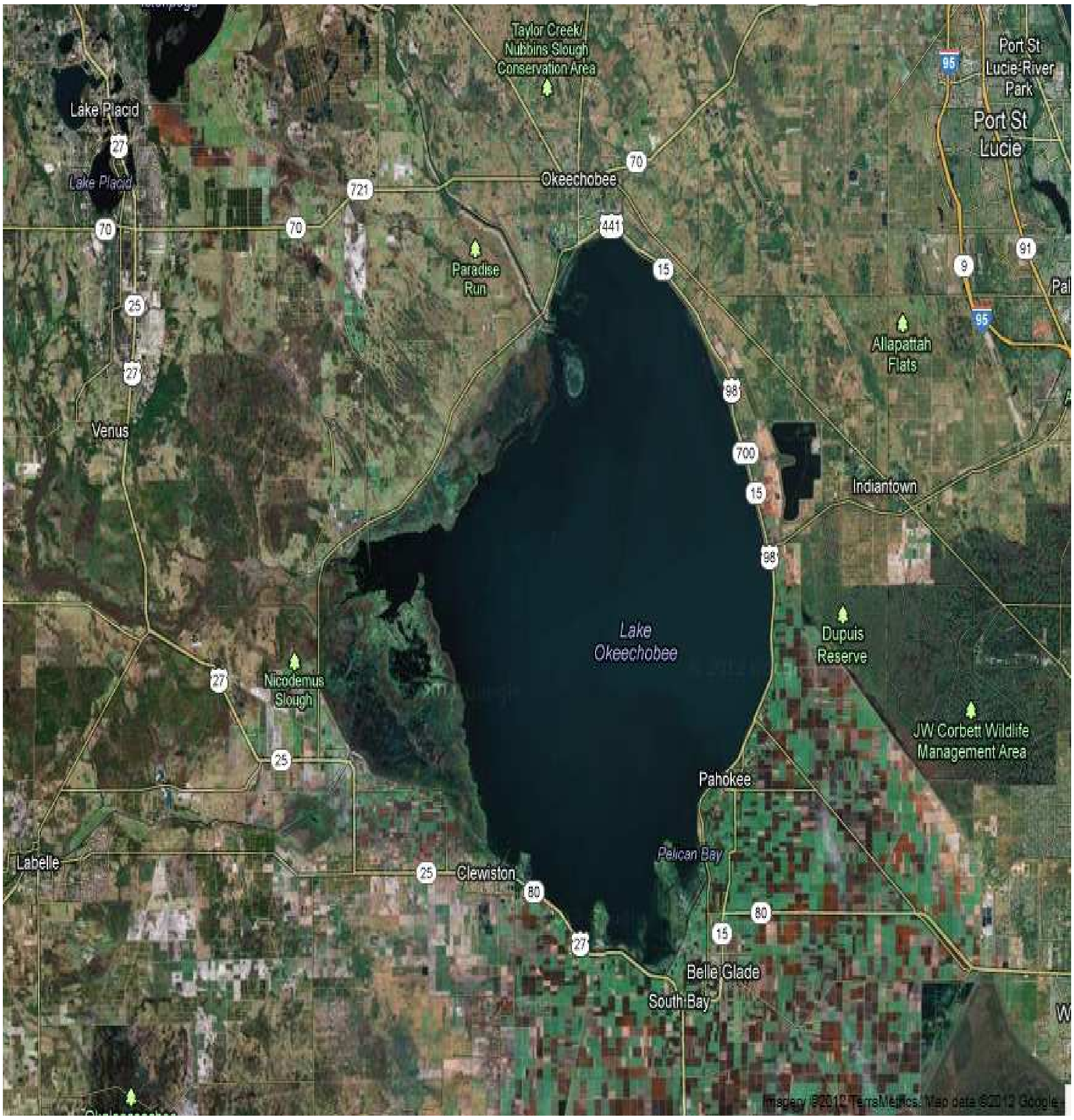

Figure 2.1 Satellite image of Lake Okeechobee (Internet 1)

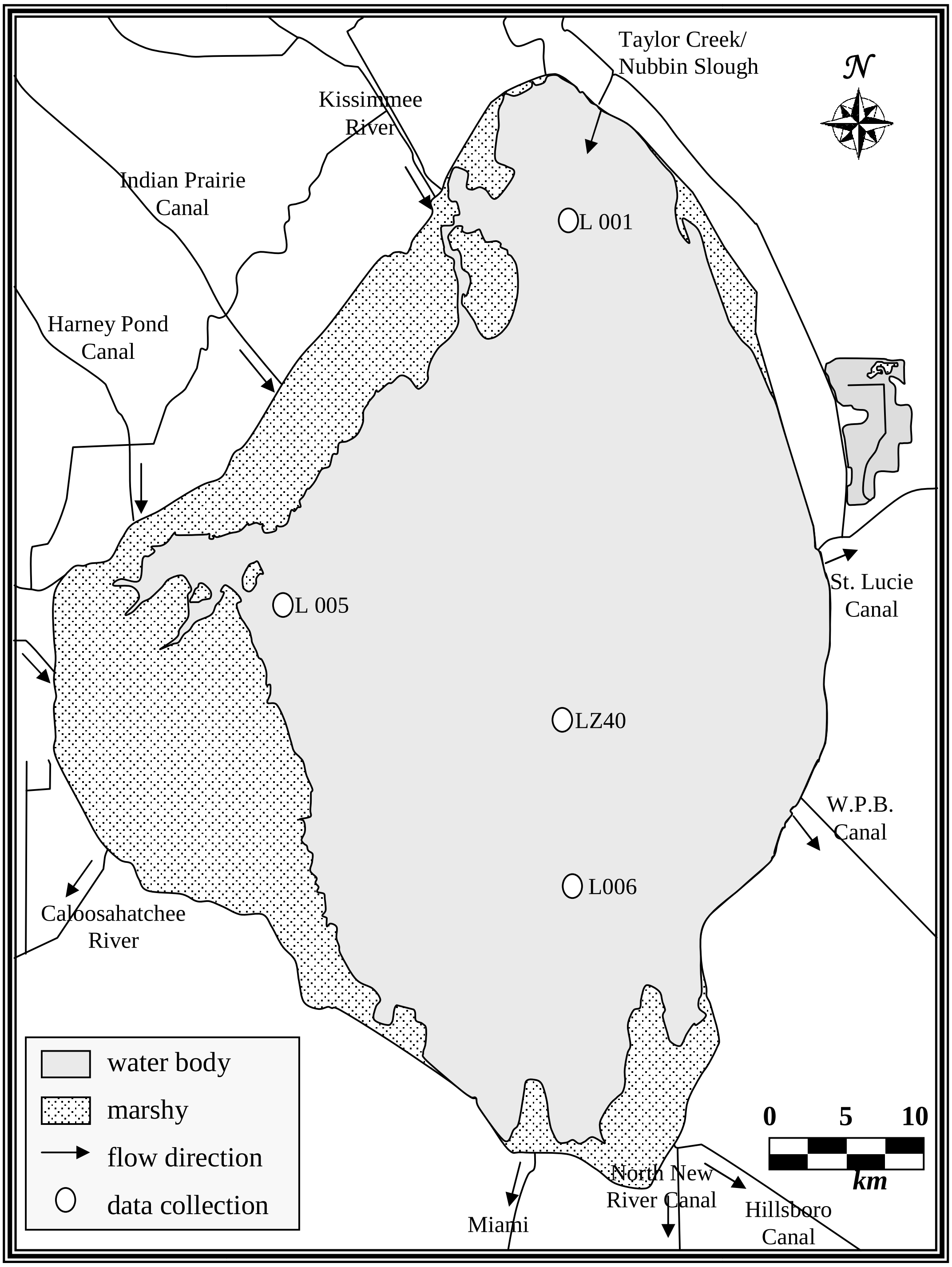

Figure 2.2 Lake Okeechobee and geographical locations of data collecting sites (L001, L005, L006 and LZ40) (Altunkaynak and Wang 2012)

2.2 Data Collection

Data collection stations for Lake Okeechobee are labeled as LZ40, L006, L005 and L001 in Figure 2.2. According to the measurements of Tehrani (2001) LZ40 (water depth 4.88m) and L001 (4m) are stationed comparatively in a deep region where L005 (3.2m) and L006 (3.66m) are near shallow region. Also the station L005 is near the marshy area, which has an effect on the wave parameters.

Two sets of independent wind-wave data from the year 1996 and 2002 are used for this study. The data from the year 1996 are collected by the South Florida Water Management District (SFWMD) from March 27th to April 2nd. During the data collection period, three strong events occurred which affected the wind speed. In those events, the measured wind speed is around 12-13m/s which is almost twice the regular wind speed. Wave elevation and wind speed are collected from three stations LZ40, L006 and L005 (Figure 2.2). Wind speeds were collected at every 15 minutes and wave elevations were measured at a recording rate of two seconds (Jin and Wang, 1998). The direct count method was used to generate the time-varying and hourly based significant wave height data used for this study.

The second set of data was collected by Wang (2002) in Lake Okeechobee from February 18 to March 7. For this data set, wave elevations were recorded at a 0.2 second sampling rate at four stations (LZ40, L006, L005 and L001) (Figure 2.2). The wind speed data was collected by the South Florida Water Management District (SFWMD) at a rate of every 15 minutes (Wang, 2002; Altunkaynak and Wang, 2012).

SI unit is adopted for all the models in this study. Wind speed data (2002) collected in the field are measured in mph (miles per hour) and converted to m/s (meter per second) for this study. Observed significant wave heights and times are measured in m (meter) and min (minute) respectively.

2.3 Calculation of Significant Wave Height

The time variation of significant wave height at all four stations (L001, L005, L006, and LZ40) for both the year 1996 and 2002 was computed by the direct count method. This method is based on the definition of significant wave height given by Sverdrup and Munk (1947). As shown in Equation (1.1), to calculate the significant wave height, individual wave heights are read from the recorded data and the average of the highest one-third heights is computed (Jin and Wang, 1998; Altunkaynak and Wang, 2012).

The number of data points for the significant wave heights are 1551, 1593, 1667 and 1425 for the stations LZ40, L006, L005 and L001 respectively (Year 2002). The data are divided into two parts, one for model testing (last 500 data points – testing data) and one for model training (rest of the data points – training data). The data points for the year 1996 are 140 for all three (L005, L006 and LZ40) stations. This set of data is used for further verifications of the models.

Chapter 3

Methodologies and Model Equations

3.1 Introduction

The methods adopted in this study to establish the predictive models for significant wave height in Lake Okeechobee are Regression Method (RM), Perceptron Least Square Method (PLSM) and a newly developed Modified Pierson-Moskowitz (MPM) Spectrum. Each section of this chapter, describes detailed mathematical derivations involved in developing the models. As mentioned in Chapter 2 (Section 2.3), data collected in 2002 (Wang, 2002) for the four stations (LZ40, L006, L005 and L001) are divided into two parts, training data and testing data. Training data are used to develop the model parameters and then they are tested by using the 500 testing data for each model. For further verification of the models, another independent set of data from 1996 (Jin and Wang 1998) are used.

Comparisons among different models are done based on Average Absolute Error (AAE), Root Mean Square Error (RMSE) and Coefficient of Efficiency (CE). They are defined as

$$\boldsymbol{AAE} = \frac{\mathbf{1}}{\boldsymbol{N}} \sum_{i=1}^{N} \left| \boldsymbol{H}_{pi} - \boldsymbol{H}_{oi} \right| , \tag{3.1}$$

$$\boldsymbol{RMSE} = \sqrt{\frac{\mathbf{1}}{\boldsymbol{N}} \sum_{i=1}^{N} \left(\boldsymbol{H}_{pi} - \boldsymbol{H}_{oi} \right)^2} , \tag{3.2}$$

$$\boldsymbol{CE} = \left[\mathbf{1} - \frac{\sum_{i=1}^{N} \left(\boldsymbol{H}_{pi} - \boldsymbol{H}_{oi} \right)^2}{\sum_{i=1}^{N} (\boldsymbol{H}_{ave} - \boldsymbol{H}_{oi})^2} \right] , \tag{3.3}$$

$\boldsymbol{H_{pi}}$ and $\boldsymbol{H_{oi}}$ are the predicted and observed significant wave height for data point $\boldsymbol{i}$ and $\boldsymbol{N}$ is the total number of the data points. $\boldsymbol{H_{ave}}$ is the averaged value of the significant wave height.

3.2 Regression Method (RM)

Generally in Regression Method, wind speed is used as independent variable and significant wave height as dependent variable. For this study two similar models are developed. One of them is the regular RM model which predicts significant wave height using wind speed at the current time and defined as RM1. Mathematical expression for model RM1 is given as

$$\boldsymbol{H_K = a_1\, W_K + b_1}\,, \tag{3.4}$$

where $\boldsymbol{H_k}$ and $\boldsymbol{W_k}$ represent the significant wave height and wind speed at current time respectively. $\boldsymbol{a_1}$ and $\boldsymbol{b_1}$ represent the model coefficients for model RM1.

The time variation plots showing the correlations between significant wave height and wind speed at stations LZ40, L006, L005, and L001 are shown in Figures 3.1a, 3.2a, 3.3a and 3.4a respectively using the year 2002 data. Scatter diagram of significant wave height and corresponding wind speed for LZ40, L006, L005 and L001 are shown in Figures 3.1b, 3.2b, 3.3b and 3.4b respectively. The value of $\boldsymbol{R^2}$ is given in the figures to represent the strength of the relationship between significant wave height and wind speed.

Strong correlation can be noticed for the station LZ40 (Figure 3.1a) and L006 (Figure 3.2a). The values of $\boldsymbol{R^2}$ for LZ40 and L006 are approximately 87% (Figure 3.1b) and 89% (Figure 3.2b) respectively, which represent the strong bond between significant wave height and wind speed. Station L005 (Figure 3.3a, marked with circle) shows weaker correlation comparing to LZ40 and L006. Figure 3.3b shows the value of $\boldsymbol{R^2}$ for L005 is 67%. It can be assumed because of the marshy area near L005, strong wind speed is not able to produce bigger wave height as it will do in the open area like LZ40 or L006. Figure 3.4a shows good correlation between wind speed and wave height over the time for L001, except at the end of the time series (marked with arrow) and the value of $\boldsymbol{R^2}$ is 60% (Figure 3.4b). This fluctuation is most likely due to some measurement error occurred during the data collection event.

In the second method, significant wave height from the previous time step is applied to predict the current significant wave height (defined as RM2). The equation for model RM2 can be expressed as

$$\boldsymbol{H_K = a_2\, H_{K-1} + b_2}\,, \tag{3.5}$$

where $\boldsymbol{H_k}$ and $\boldsymbol{H_{k-1}}$ represent the significant wave height in current and previous time step respectively. $\boldsymbol{a_2}$ and $\boldsymbol{b_2}$ represent the to-be-calibrated coefficients for model RM2.

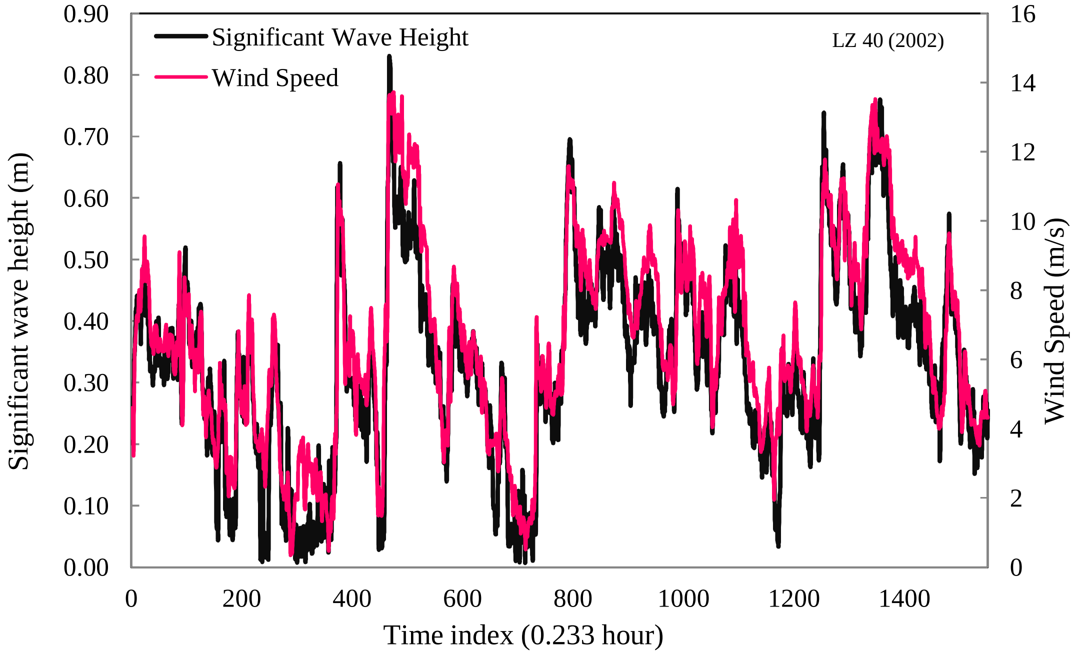

Figure 3.1a Time variations of significant wave height and wind speed at LZ40 (year 2002 data)

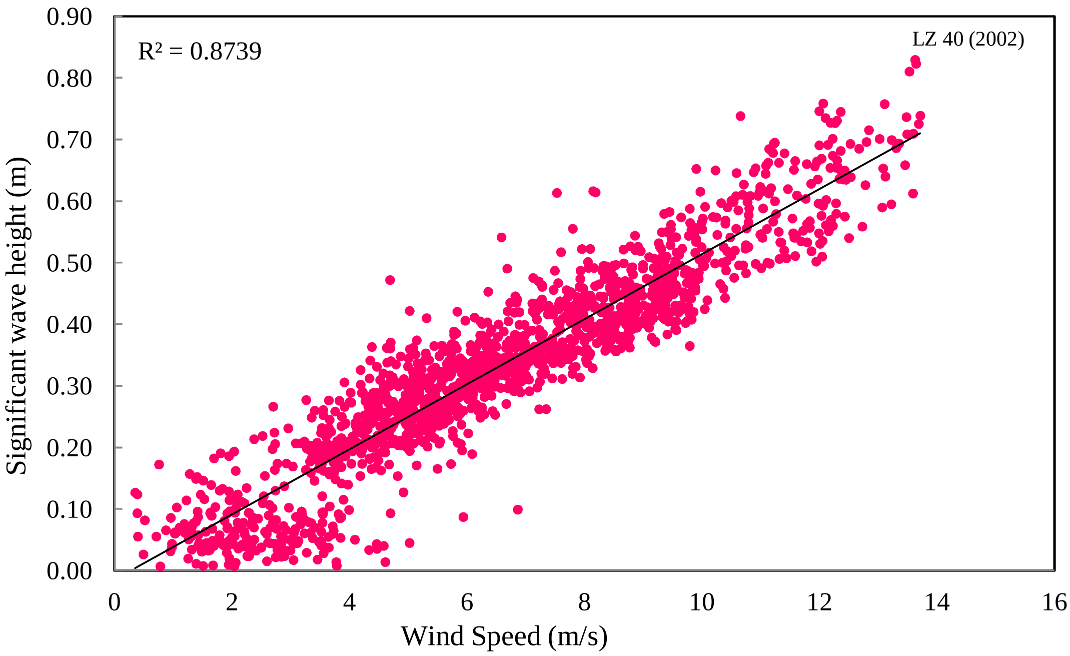

Figure 3.1b Correlation between significant wave height and wind speed at LZ40 (year 2002 data)

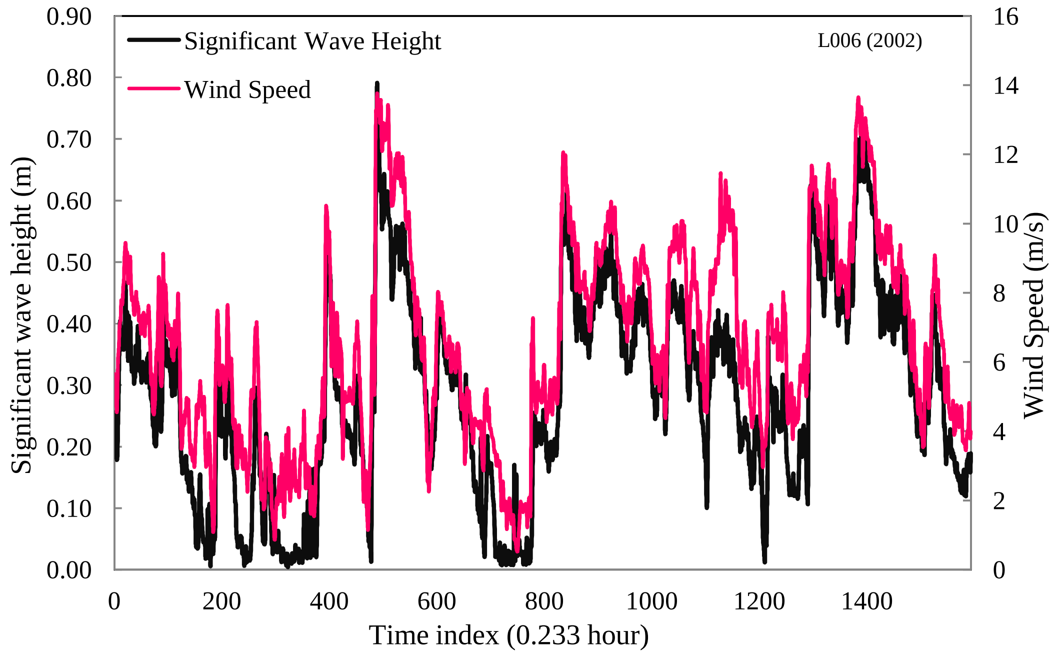

Figure 3.2a Time variations of significant wave height and wind speed at L006 (year 2002 data)

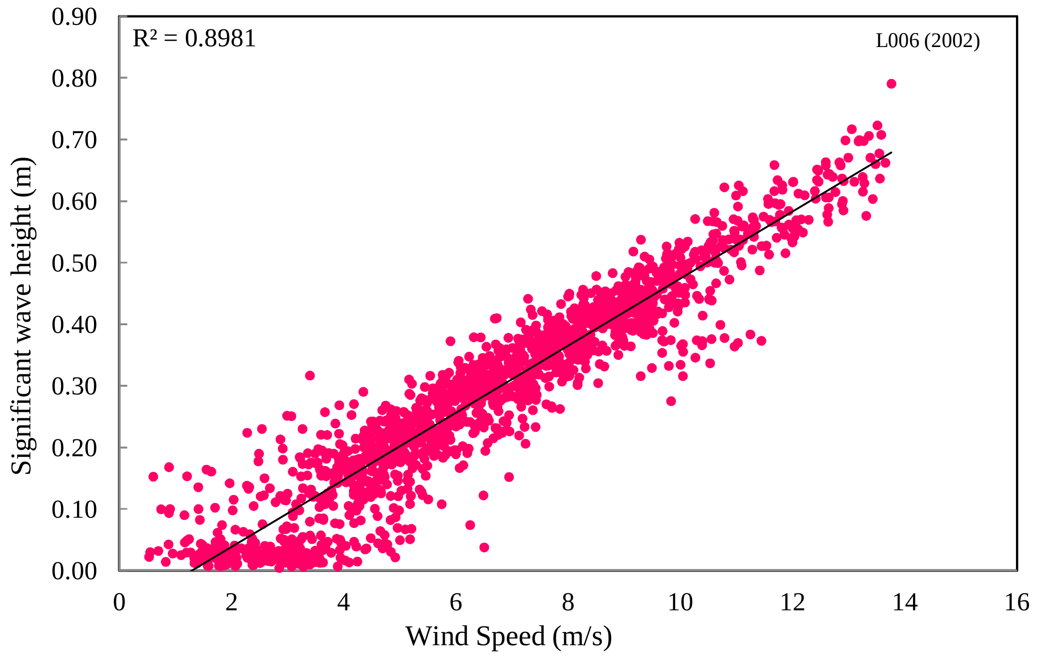

Figure 3.2b Correlation between significant wave height and wind speed at LZ40 (year 2002 data)

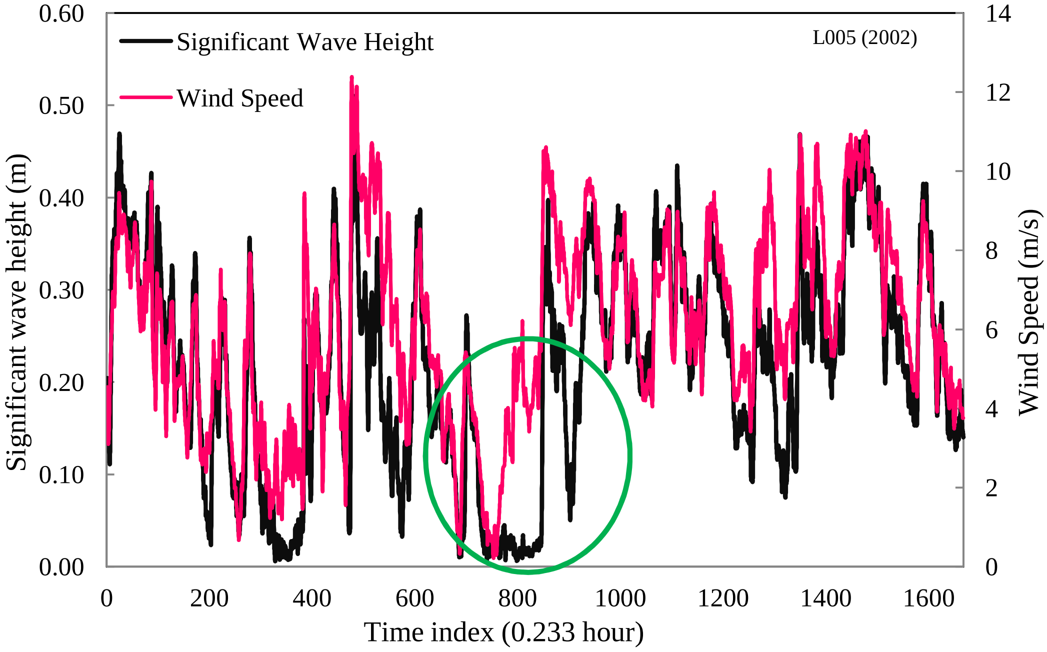

Figure 3.3a Time variations of significant wave height and wind speed at L005 (year 2002 data)

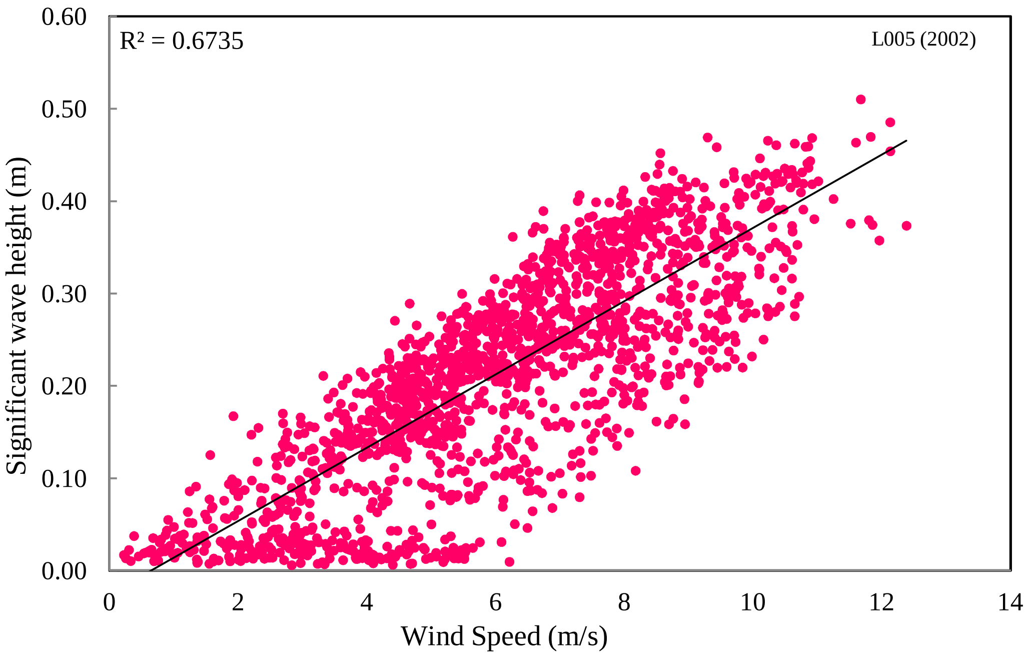

Figure 3.3b Correlation between significant wave height and wind speed at LZ40 (year 2002 data)

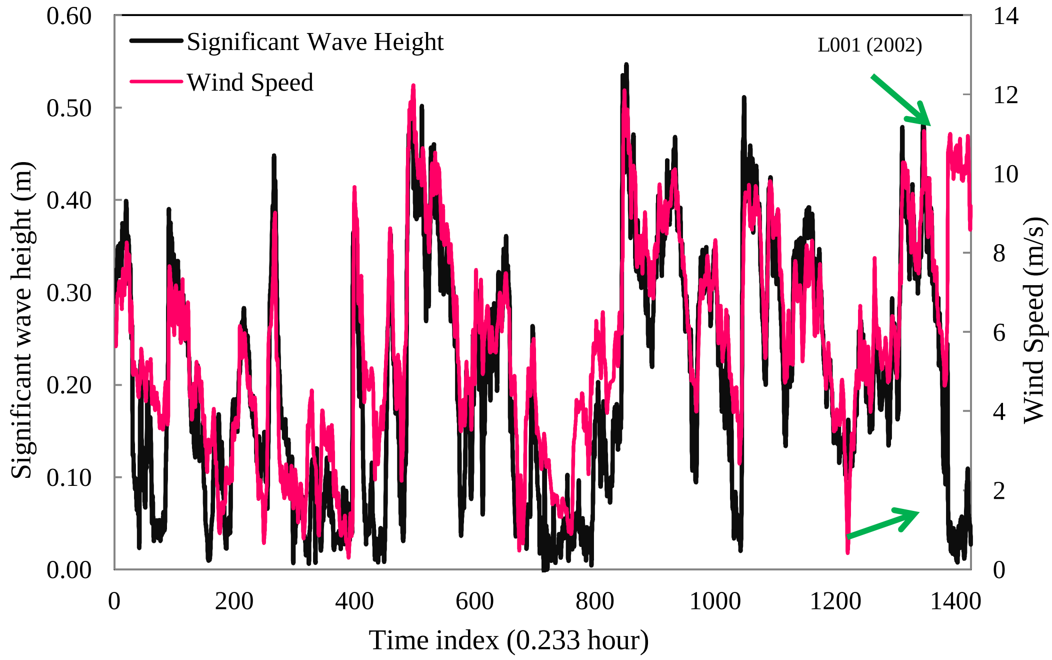

Figure 3.4a Time variations of significant wave height and wind speed at L001 (year 2002 data)

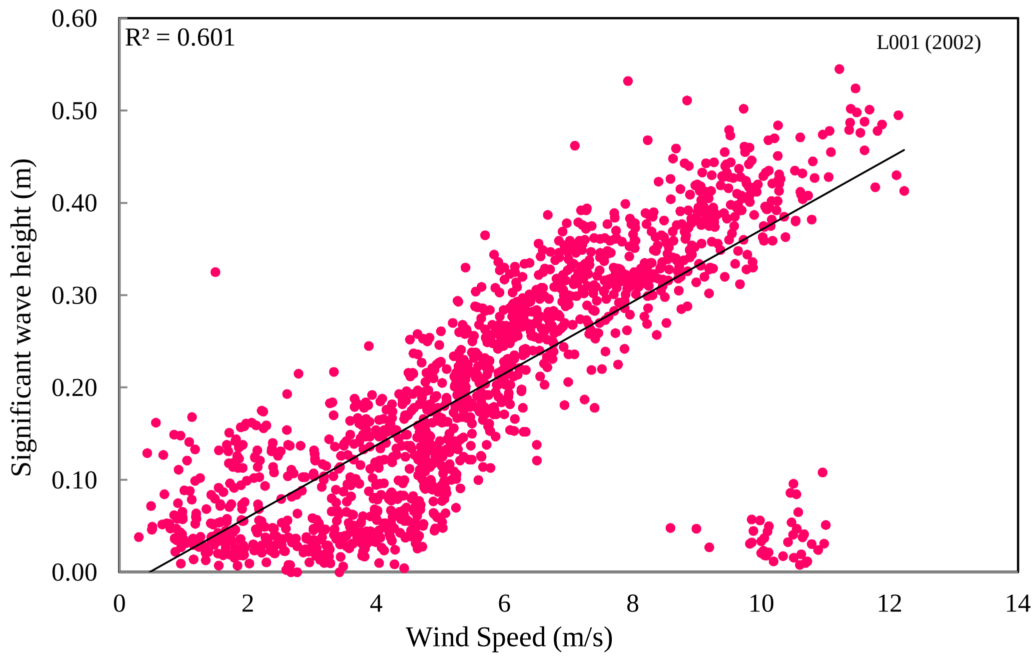

Figure 3.4b Correlation between significant wave height and wind speed at LZ40 (year 2002 data)

Calibrated model parameters of the models RM1 and RM2 for all the four stations are shown in Table 3.1. As a comparison, time variations of observed and predicted significant wave height from model RM1 using training data are shown in Figure 3.5 to Figure 3.8 for stations LZ40, L006, L005 and L00, respectively. Similarly, time variations of predicted and observed significant wave height at the four stations under the case of using model RM2 are shown in Figure 3.9 to Figure 3.12. For model RM1, the predicted time varying significant wave height generally fits well with the recorded data at LZ40, L006, and L001 stations. At station L005, the agreement is reasonable, however, with noticeable larger errors. Under very strong wind conditions, the model RM tends to slightly under-estimate the significant wave height. The predictions using model RM2 fit better than the RM1 results when compared with measured significant wave height for all modeled stations. This finding indicates that the current significant wave height depends heavily on that which occurred previously. To demonstrate further model performance, the perfect model line plots between RM1 results and observed significant wave height for four stations are presented in Figure 3.13 to Figure 3.16 and for the results from RM2 they are shown in Figure 3.17 to Figure 3.20. From the Figures, it is noticed that the predicted significant wave height using RM2 models generally fitted around the 45 degree perfect model line with good agreement with measured data (Figure 3.17 to 3.20), comparing to the results produced from RM1 model (Figure 3.13 to 3.16).

Table 3.1 Calculated parameters of RM1 and RM2 for the station LZ40, L006, L005 and L001

Station	Model / Parameters	RM1		RM2	
		a_1	b_1	a_2	b_2
LZ40		0.052	-0.010	0.962	0.012
L006		0.055	-0.067	0.981	0.005
L005		0.039	-0.022	0.980	0.004
L001		0.044	-0.045	0.968	0.006

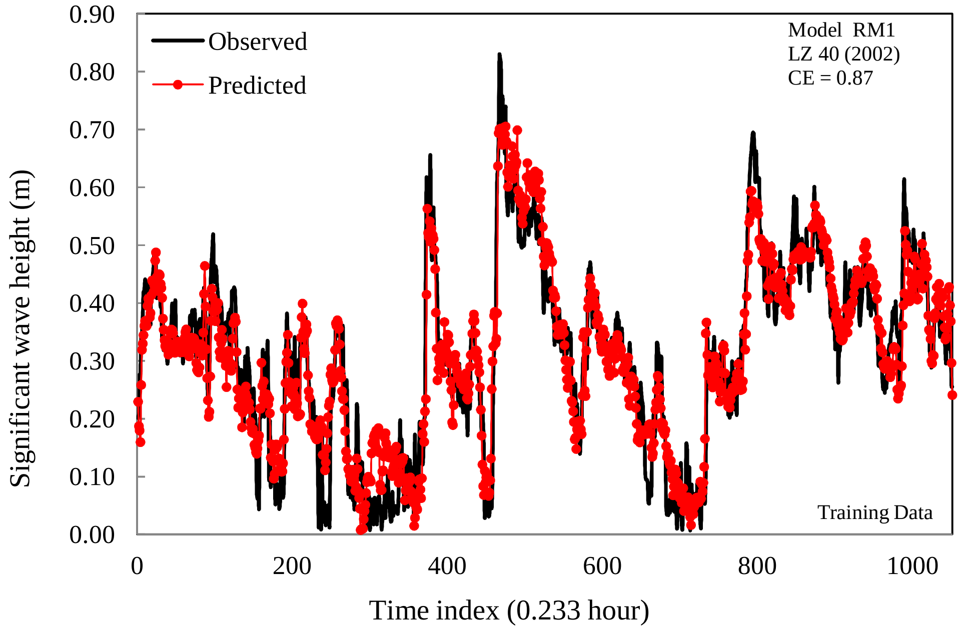

Figure 3.5 Predicted and observed time variations of significant wave heights at LZ40 using training data (RM1)

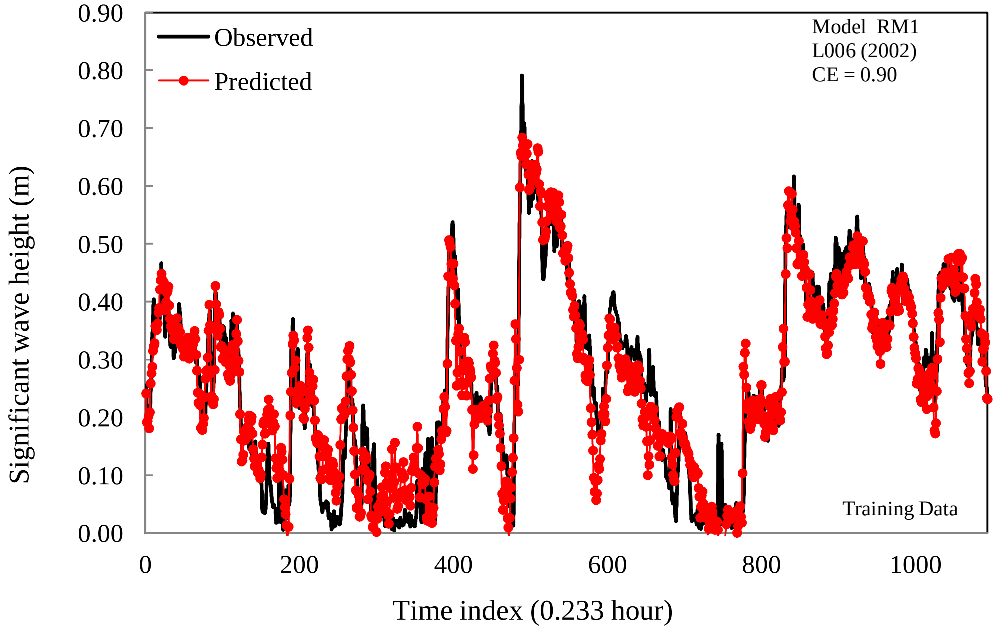

Figure 3.6 Predicted and observed time variations of significant wave heights at L006 using training data (RM1)

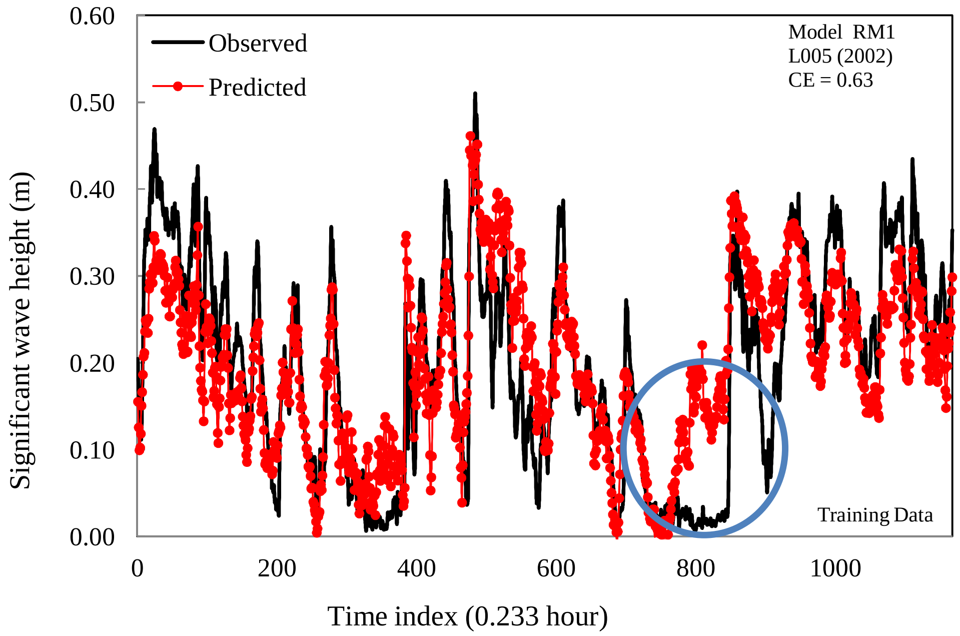

Figure 3.7 Predicted and observed time variations of significant wave heights at L005 using training data (RM1)

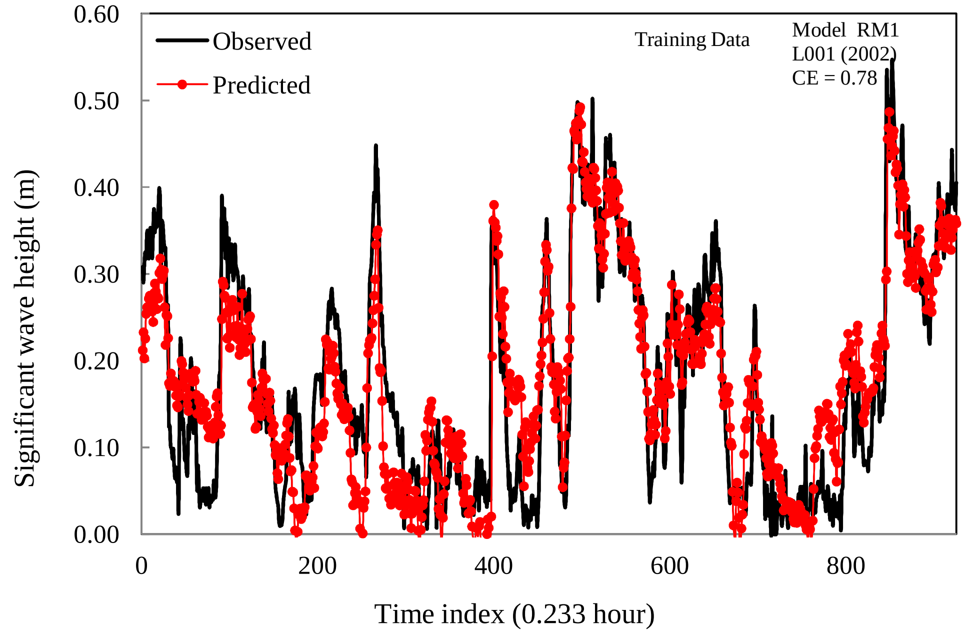

Figure 3.8 Predicted and observed time variations of significant wave heights at L001 using training data (RM1)

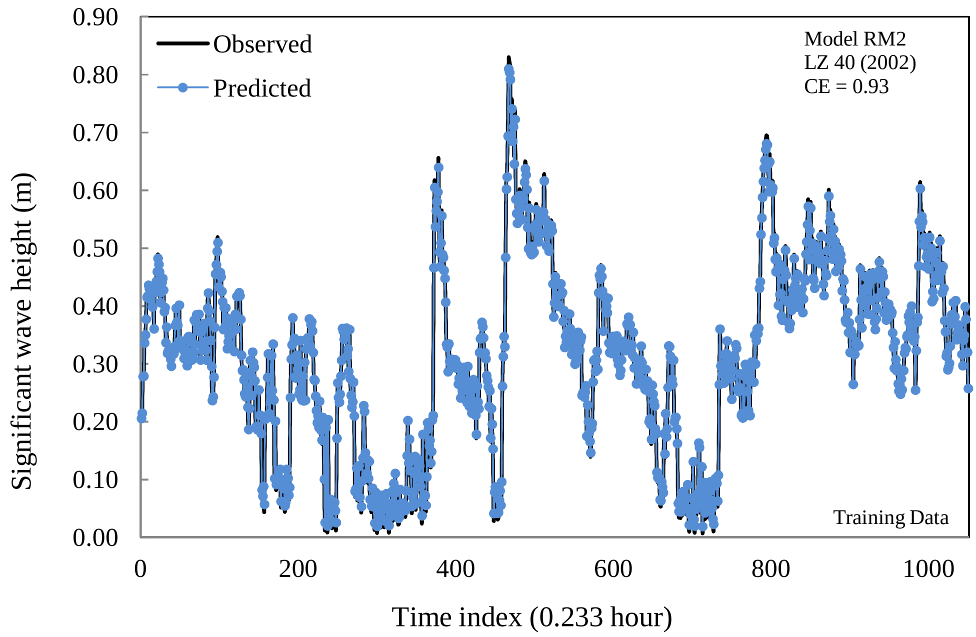

Figure 3.9 Predicted and observed time variations of significant wave heights at LZ40 using training data (RM2)

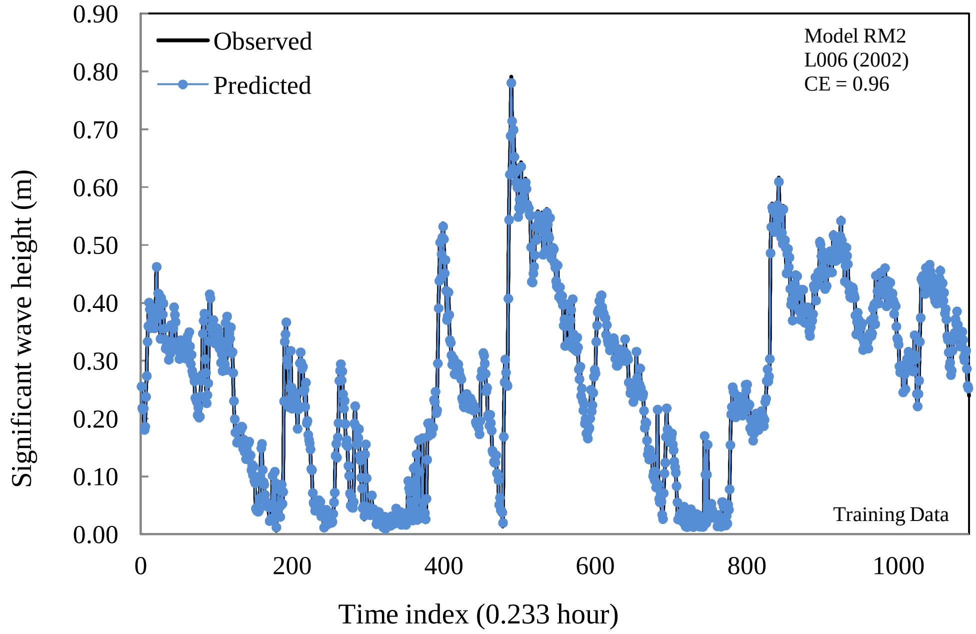

Figure 3.10 Predicted and observed time variations of significant wave heights at L006 using training data (RM2)

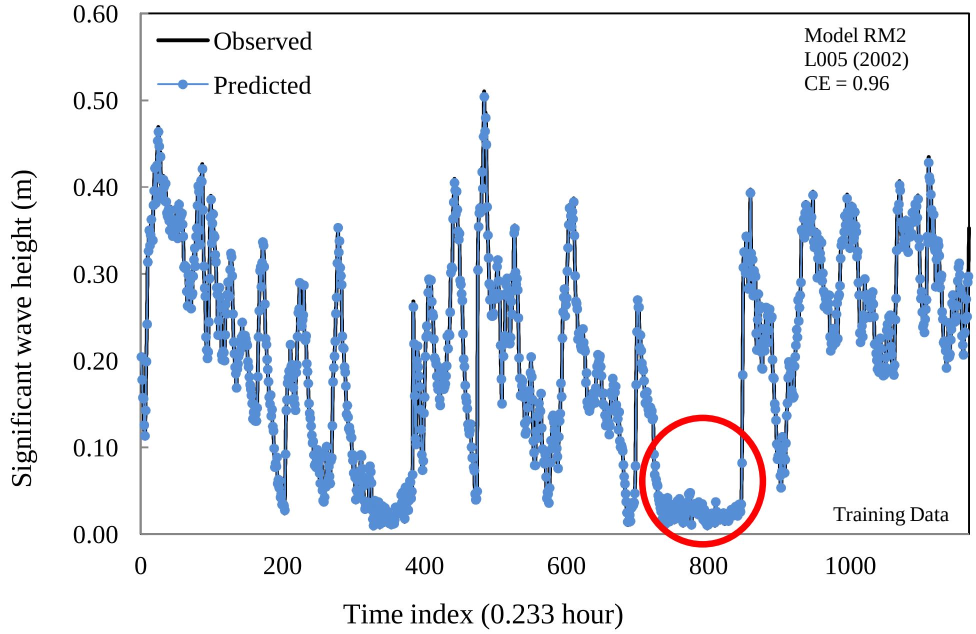

Figure 3.11 Predicted and observed time variations of significant wave heights at L005 using training data (RM2)

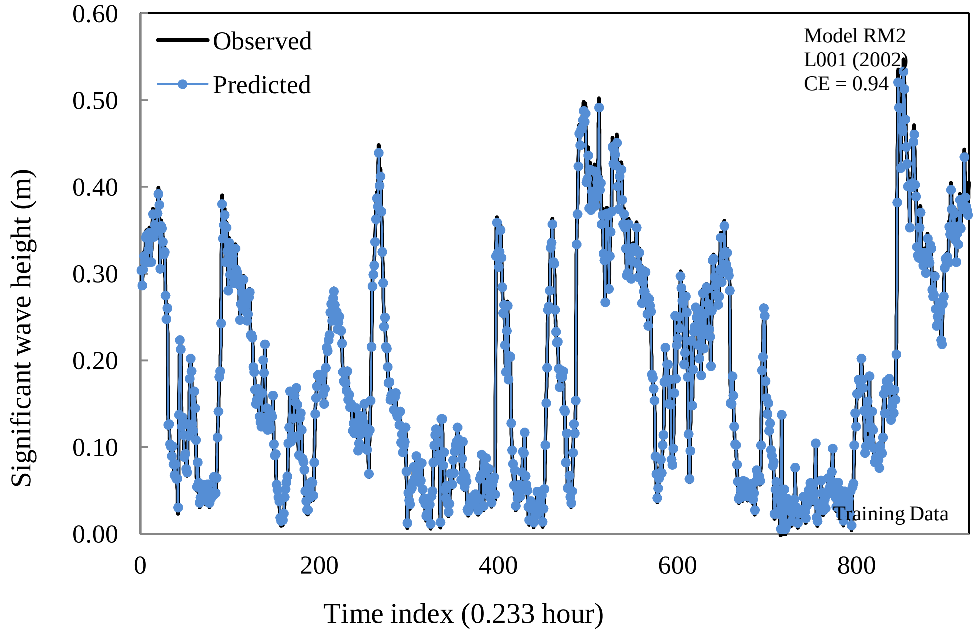

Figure 3.12 Predicted and observed time variations of significant wave heights at L001 using training data (RM2)

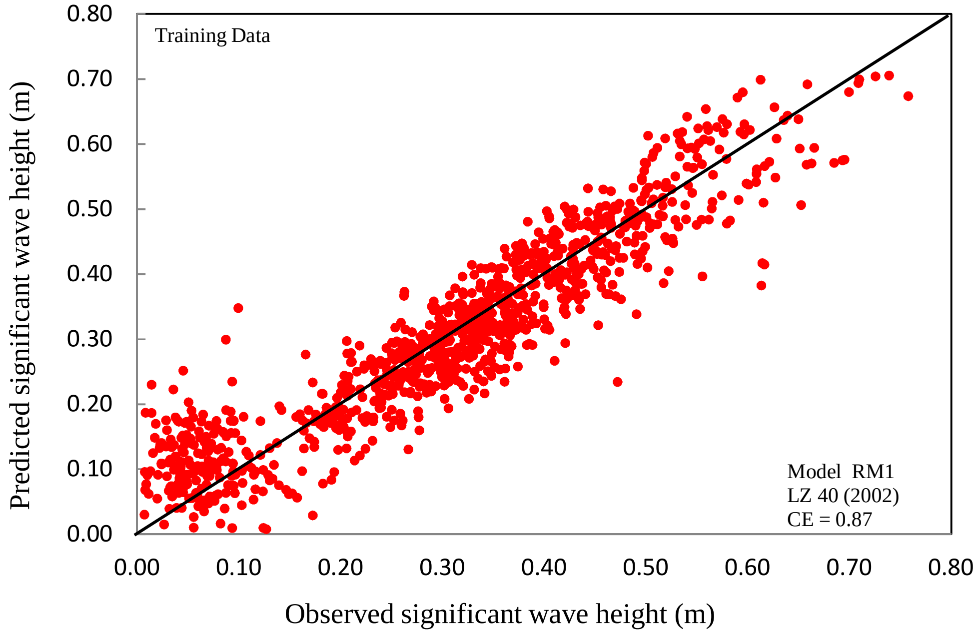

Figure 3.13 Verification of predicted and observed significant wave heights using training data (RM1, station LZ40)

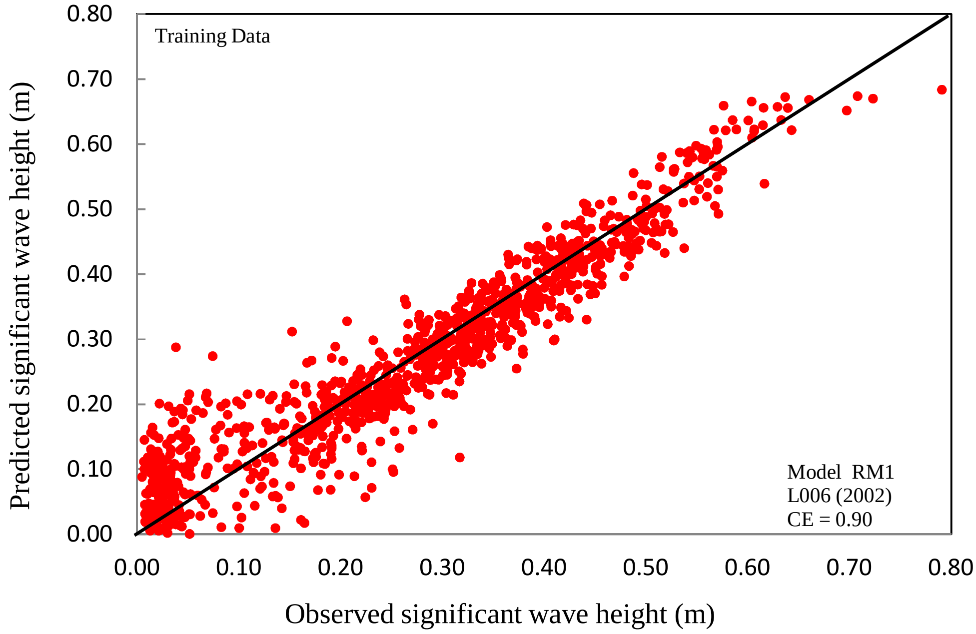

Figure 3.14 Verification of predicted and observed significant wave heights using training data (RM1, station L006)

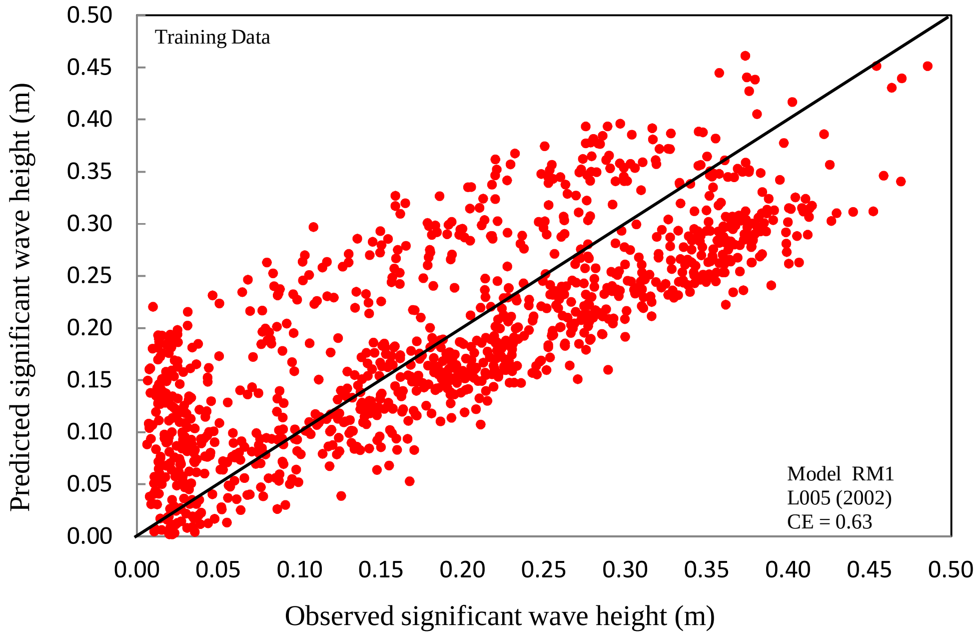

Figure 3.15 Verification of predicted and observed significant wave heights using training data (RM1, station L005)

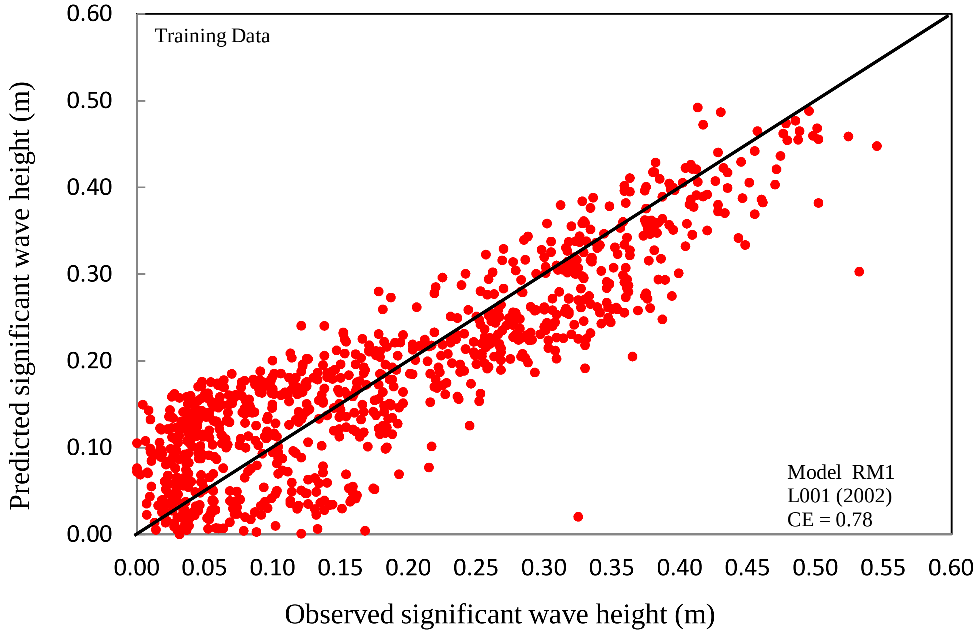

Figure 3.16 Verification of predicted and observed significant wave heights using training data (RM1, station L001)

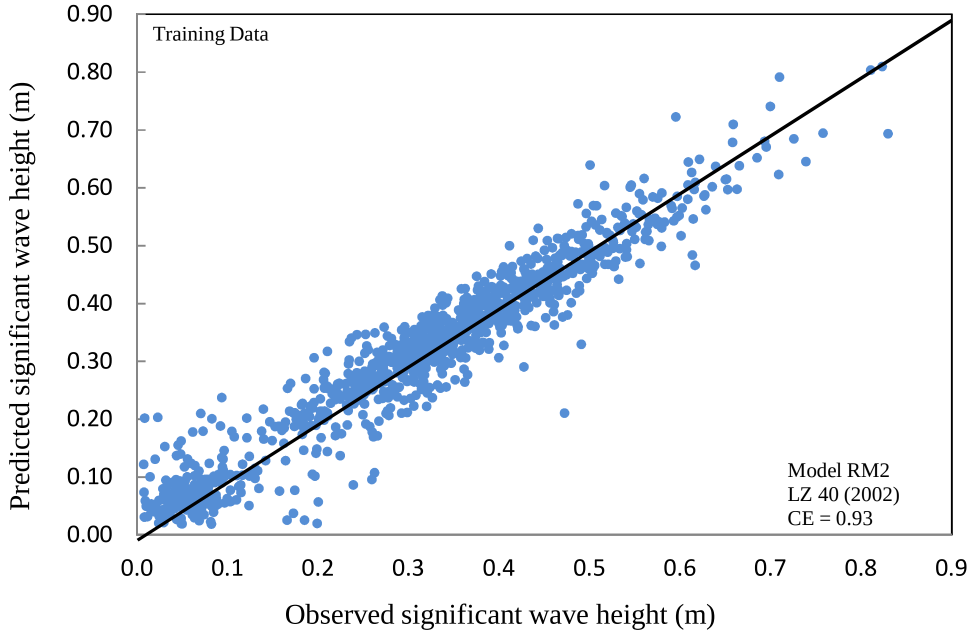

Figure 3.17 Verification of predicted and observed significant wave heights using training data (RM2, station LZ40)

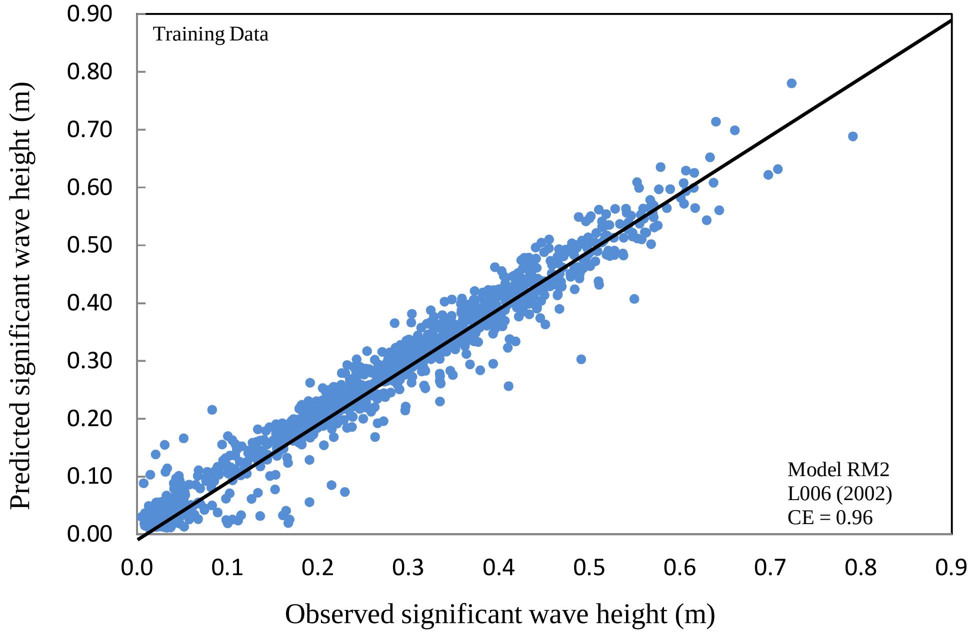

Figure 3.18 Verification of predicted and observed significant wave heights using training data (RM2, station L006)

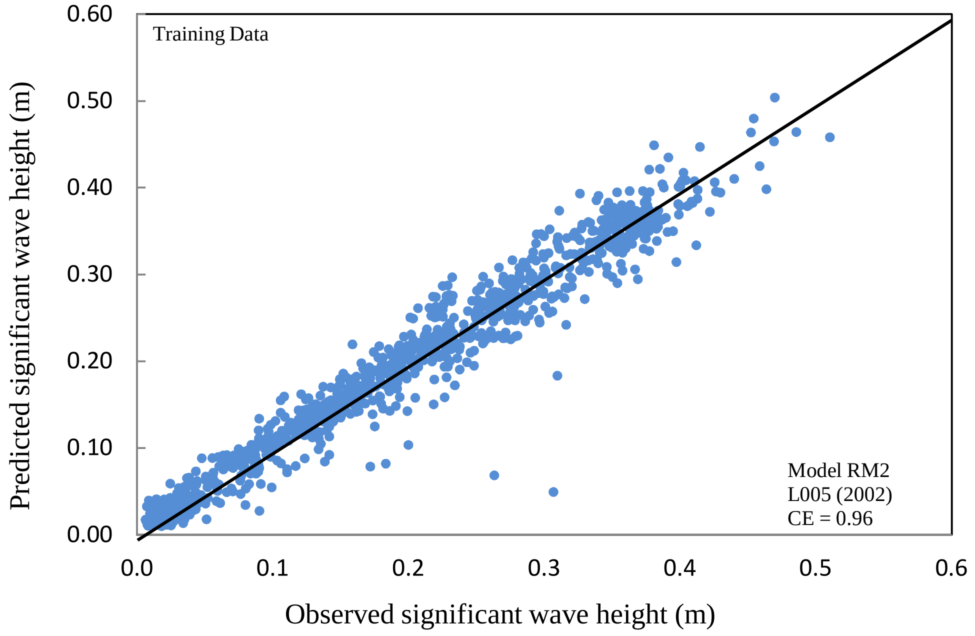

Figure 3.19 Verification of predicted and observed significant wave heights using training data (RM2, station L005)

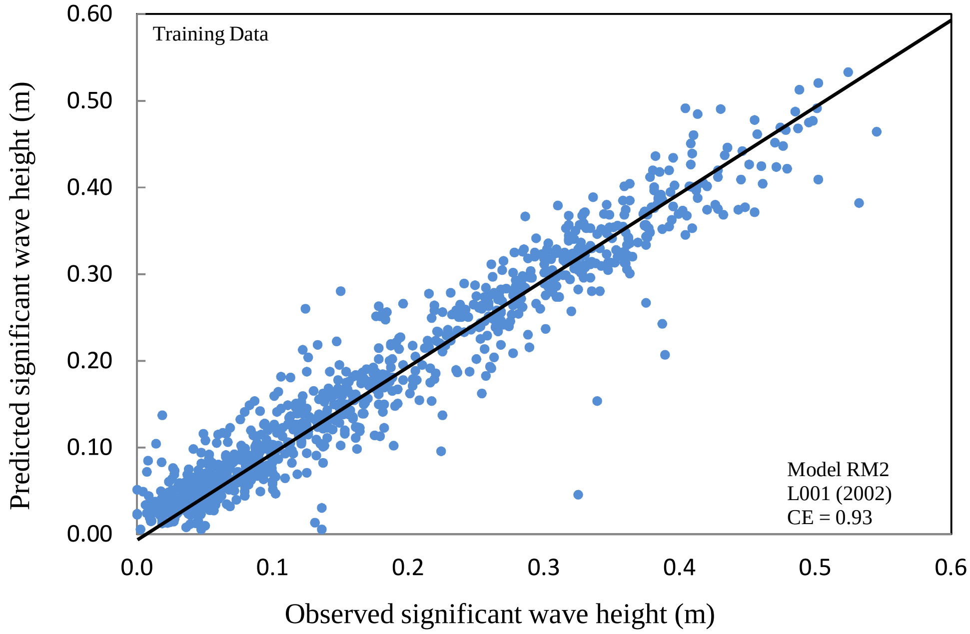

Figure 3.20 Verification of predicted and observed significant wave heights using training data (RM2, station L001)

Comparing results obtained from the training data, it can be seen that RM2 model worked better for all the four stations than model RM1. As mentioned in Section 2.2, station L005 is near the marshy area and that has an effect on wind speed. Wind speed is the independent variable in model RM1 and that effect of marshy area is not overcome by the model while predicting the significant wave height. As a result, RM1 model produced relatively poor results for station L005. As shown in Figure 3.7, it is visible (marked with circle in the figure) that if model RM1 is used to predict significant wave height, it does not produce accurate results. On the other hand, results from model RM2 for station L005 are more accurate (Figure 3.11). Further verifications of models RM1 and RM2 are done in Chapter 4.

To get a better understanding of the performance of models RM1 and RM2, a summary of AAE, RMSE and CE for all stations are provided in Table 3.2. Coefficient of efficiency (CE) produced by RM2 for all stations are over 90% where RM1 produced less than 80% of CE for L005 and L001. AAE and RMSE are also less for RM2 which indicates less error in predicting significant wave height comparing to RM1.

3.3 Perceptron Least Square Method (PLSM)

Perceptron Neural Network is also applied in this study to estimate the significant wave height occurred in Lake Okeechobee. To determine the network weights, least square method is applied. Since the network system is coupled with least square technique, it will be referred as Perceptron Least Square Method (PLSM). As only significant wave height will be calculated using PLSM, Figure 1.3 can be reconstructed as Figure 3.21 shown below.

Table 3.2 Summary of AAE, RMSE and CE of Models RM1 and RM2 using training data for the stations LZ40, L006, L005 and L001

	RM1			RM2		
	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE
LZ40	0.04	0.06	0.87	0.03	0.04	0.93
L006	0.04	0.05	0.90	0.02	0.03	0.96
L005	0.06	0.07	0.63	0.02	0.02	0.96
L001	0.05	0.06	0.78	0.02	0.03	0.93

Figure 3.21 demonstrates that significant wave height at current time step ($\boldsymbol{H_k}$) is determined by the significant wave ($\boldsymbol{H_{k-1}}$) and wind speed ($\boldsymbol{W_{k-1}}$) of previous time step. Network weights are denoted as $\boldsymbol{a_3}$ and $\boldsymbol{b_3}$ and least square methods are applied to calculate them. The mathematical form PLSM model is given as

$$\boldsymbol{H_K = a_3\, H_{K-1} + b_3\, W_{K-1}} \ . \tag{3.6}$$

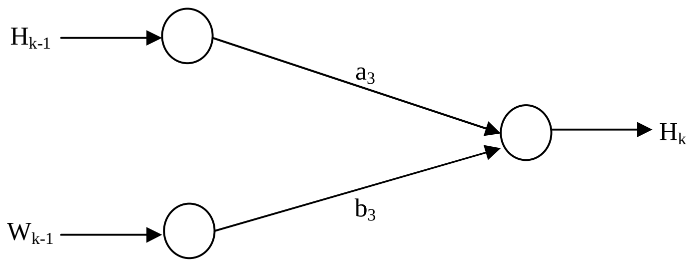

Figure 3.21 Perceptron configuration of significant wave height and wind speed for Lake Okeechobee

Network weight values for all four stations are showed in Table 3.3. Time variations of PLSM predicted and observed significant wave height using training data are showed in Figure 3.22 to Figure 3.25 for stations LZ40, L006, L005 and L00, respectively. It can be noticed, according to the comparisons with the training data, the PLSM model can produce accurate predictions on significant wave height. Good agreements between PLSM model results and observed significant wave heights are also presented in the perfect model line plots in Figure 3.26 to Figure 3.29 for stations LZ40, L006, L005, and L001 respectively. The results are shown to closely follow the 45 degree perfect model line.

Table 3.3 Network weights of Model PLSM for the station LZ40, L006, L005 and L001

Station	Model	PLSM	
	Parameters	a_3	b_3
LZ40		0.78	0.011
L006		0.86	0.007
L005		0.95	0.002
L001		0.90	0.004

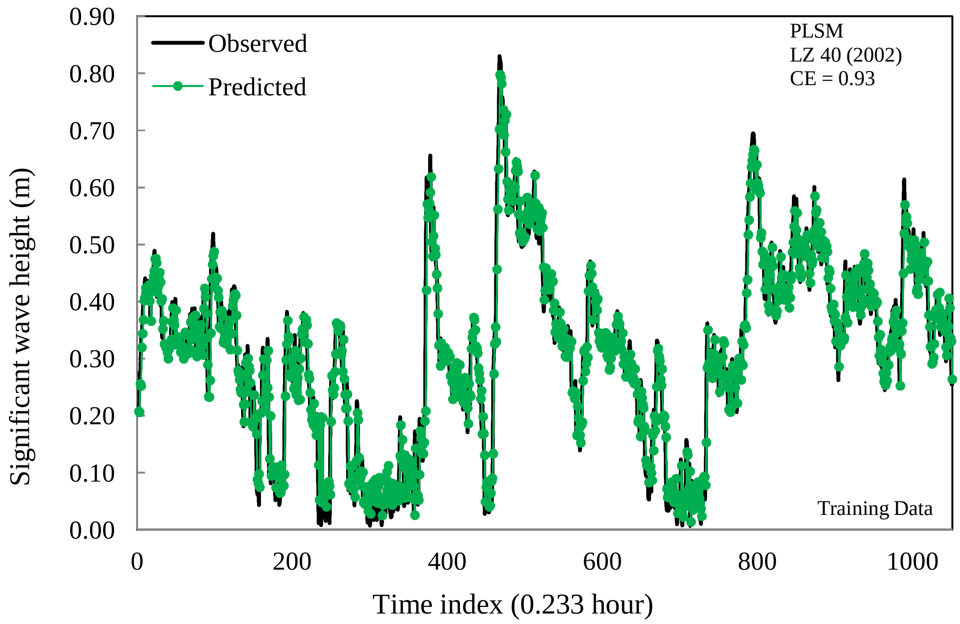

Figure 3.22 Predicted and observed time variations of significant wave heights at LZ40 using training data (Model PLSM)

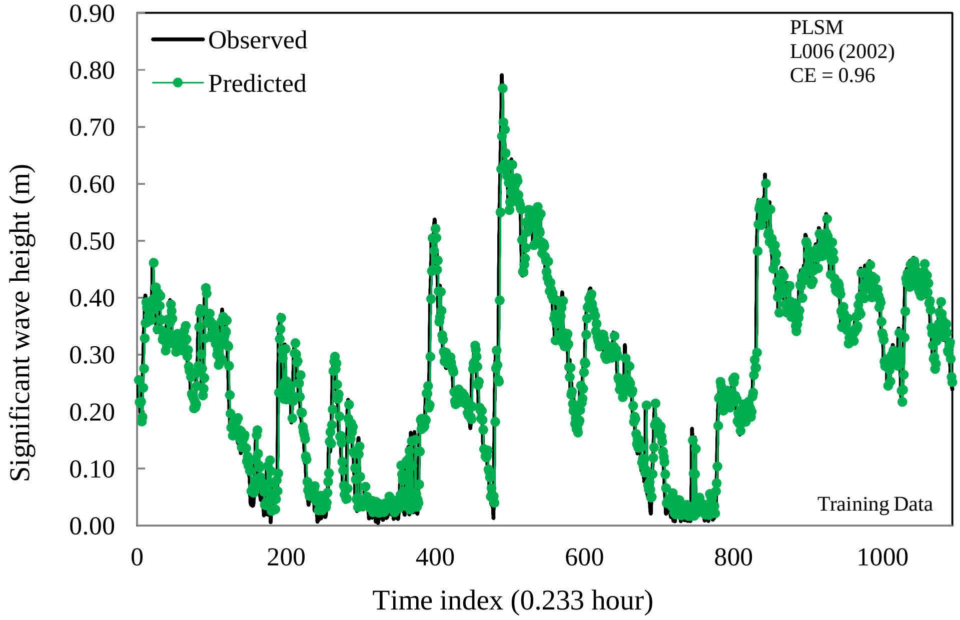

Figure 3.23 Predicted and observed time variations of significant wave heights at L006 using training data (Model PLSM)

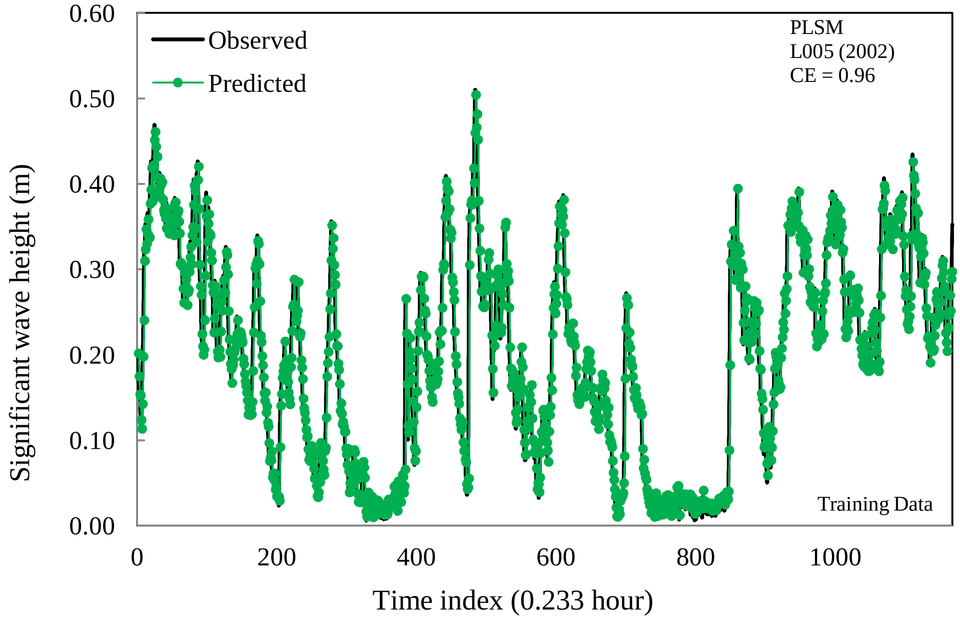

Figure 3.24 Predicted and observed time variations of significant wave heights at L005 using training data (Model PLSM)

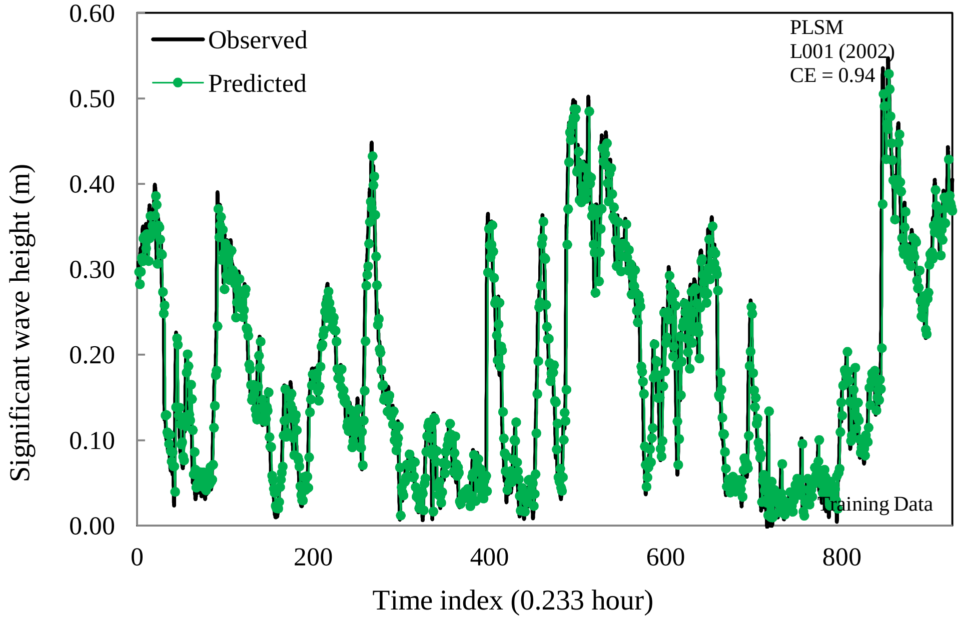

Figure 3.25 Predicted and observed time variations of significant wave heights at L001 using training data (Model PLSM)

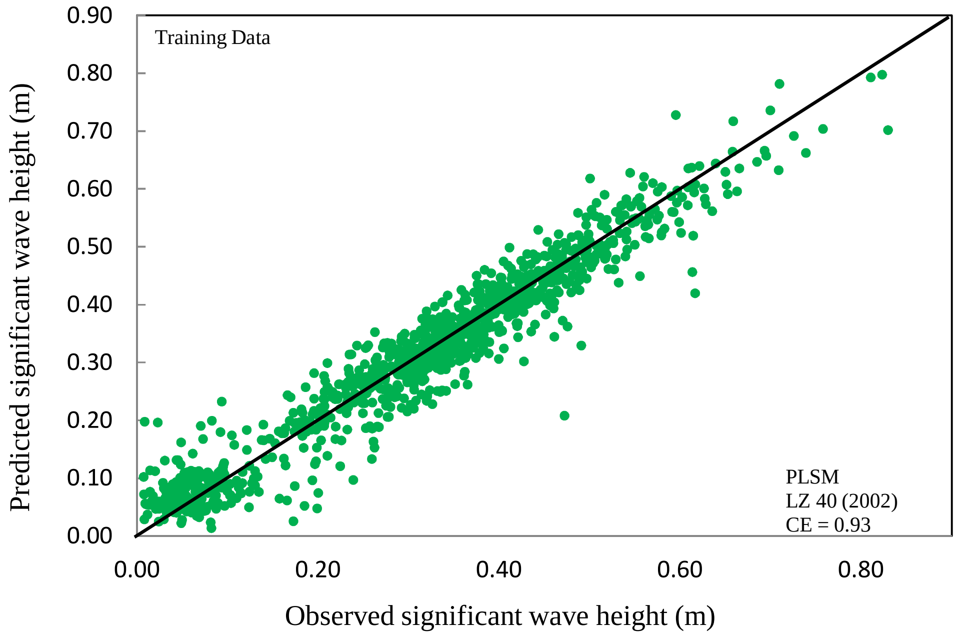

Figure 3.26 Verification of predicted and observed significant wave heights using training data (Model PLSM, station LZ40)

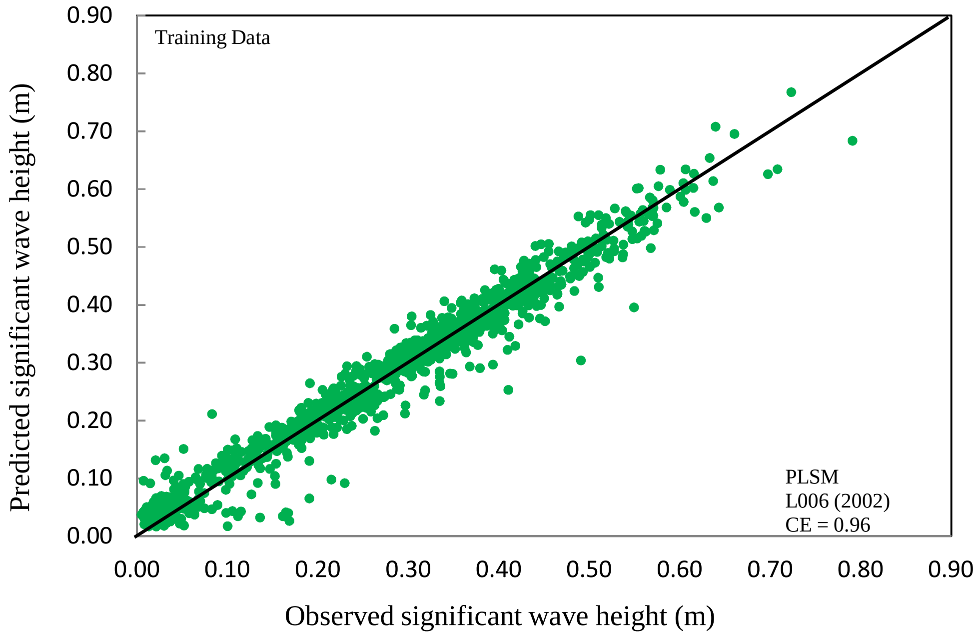

Figure 3.27 Verification of predicted and observed significant wave heights using training data (Model PLSM, station L006)

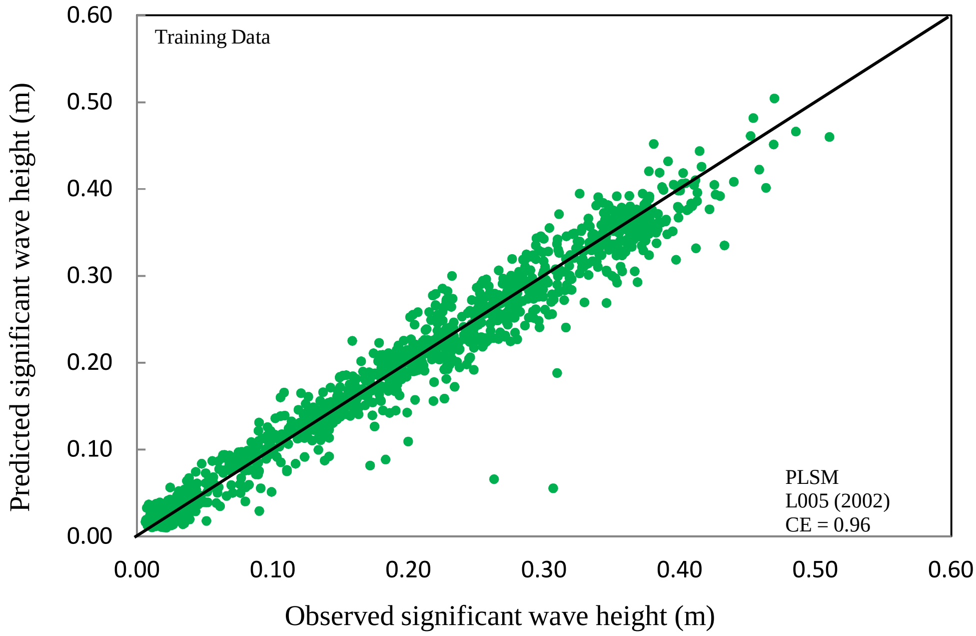

Figure 3.28 Verification of predicted and observed significant wave heights using training data (Model PLSM, station L005)

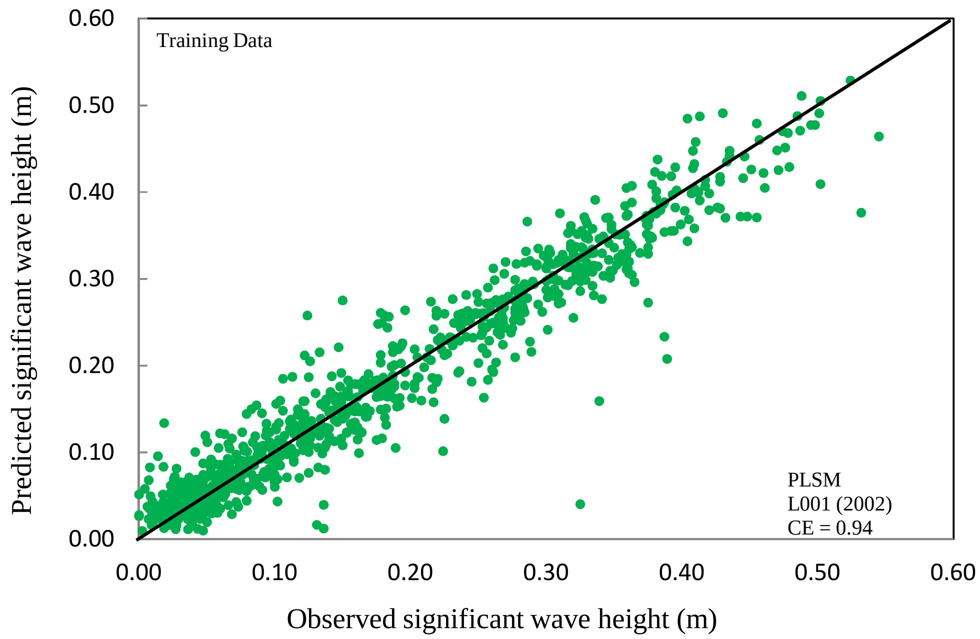

Figure 3.29 Verification of predicted and observed significant wave heights using training data (Model PLSM, station L001)

Table 3.4 Values of AAE, RMSE and CE of model PLSM using training data for the station LZ40, L006, L005 and L001

	Model PLSM		
	AAE (m)	RMSE (m)	CE
LZ40	0.03	0.04	0.93
L006	0.02	0.03	0.96
L005	0.02	0.02	0.96
L001	0.02	0.03	0.94

A summary of AAE, RMSE and CE of model PLSM for stations LZ40, L006, L005 and L001 are provided in Table 3.4. Model PLSM shows quite good predictions for all the sensor stations in Lake Okeechobee. Coefficient of efficiency is more than 90% for all stations using the training data. Error predicted by PLSM is also very less.

3.4 Pierson-Moskowitz Spectrum (PM Spectrum)

The original Pierson-Moskowitz (PM) Spectrum was developed and applied to fully developed sea. To the best of author's knowledge this approach was not commonly utilized to the study of shallow lakes. For the development of an appropriate predictive wave spectrum model, the PM spectrum is first applied to Lake Okeechobee to test its accuracy in predicting significant wave height. In order to do that, it is necessary to find out a relationship between the wave spectrum and the significant wave height. Detailed mathematical derivations are shown below.

From the Equation 1.7, PM wave spectrum can be restated as follow

$$\boldsymbol{S(f)} = \frac{\boldsymbol{A}}{\boldsymbol{f^5}}\, \boldsymbol{e}^{[\frac{-\boldsymbol{B}}{\boldsymbol{f^4}}]} \ , \tag{3.7}$$

where

$$\boldsymbol{A} = \frac{\boldsymbol{\alpha\, g^2}}{\boldsymbol{16\, \pi^4}} \ , \tag{3.8a}$$

$$\boldsymbol{B} = \boldsymbol{\beta}[\frac{\boldsymbol{g}}{\boldsymbol{2\, \pi\, W}}]^{\boldsymbol{4}} \ . \tag{3.8b}$$

In equation (3.8), $\boldsymbol{\alpha}$ (Phillips' constant) = **8.1x10^{-3}**, $\boldsymbol{\beta}$ = **0.74,** and $\boldsymbol{W}$ is the wind speed measured in m/s. The n-th moment of a wave spectrum $\boldsymbol{S(f)}$ can be defined as

$$\boldsymbol{m_n} = \int_0^\infty \boldsymbol{f^n\, S(f)\ df} \ , \tag{3.9}$$

where $\boldsymbol{f}$ is wave frequency and $\boldsymbol{m_n}$ is the n-th moment of spectrum $\boldsymbol{S(f)}$. When $\boldsymbol{n} = \boldsymbol{0}$, then equation 3.9 becomes

$$\boldsymbol{m_0} = \int_0^\infty \boldsymbol{S(f)\ df} \quad , \tag{3.10}$$

where $\boldsymbol{m_0}$, representing the total wave energy, is the zero-th moment of $\boldsymbol{S(f)}$ and can be related to the significant wave height. Replacing the value of $\boldsymbol{S(f)}$ from the equation 3.7 to equation 3.10, we have

$$\boldsymbol{m_0} = \int_0^\infty \frac{\boldsymbol{A}}{\boldsymbol{f^5}}\ \boldsymbol{e}^{\left[\frac{-\boldsymbol{B}}{\boldsymbol{f^4}}\right]}\ \boldsymbol{df} \ . \tag{3.11}$$

Now, through the following change of variable

$$\boldsymbol{f^{-4} = x}$$

$$\boldsymbol{-4\,f^{-5}df = x\ dx}$$

$$\boldsymbol{f^{-5}df = \frac{x}{-4}\ dx}$$

and replacing the value to equation 3.11 yields

$$\boldsymbol{m_0} = \int_0^\infty \frac{\boldsymbol{A}}{\boldsymbol{-4}}\ \boldsymbol{e^{-Bx}\ dx} \ . \tag{3.12}$$

After integrating the equation 3.12, it becomes

$$\boldsymbol{m_0} = \frac{\boldsymbol{A}}{\boldsymbol{4B}} \ . \tag{3.13}$$

Significant wave height, $\boldsymbol{H_s}$ (or $\boldsymbol{H_{1/3}}$) can be defined by the zero-th moment equation ($\boldsymbol{m_0}$) as follow

$$\boldsymbol{H_s} = \boldsymbol{4}\sqrt{\boldsymbol{m_0}} \ . \tag{3.14}$$

Replacing the value of $\boldsymbol{m_0}$ from equation (3.13) to equation (3.14), we have

$$\boldsymbol{H_s} = \boldsymbol{4}\sqrt{\frac{\boldsymbol{A}}{\boldsymbol{4B}}} \ . \tag{3.15}$$

Here,

$$\boldsymbol{A} = \frac{\alpha\, g^2}{16\, \pi^4} \ ,$$

$$\boldsymbol{B} = \boldsymbol{\beta}[\frac{g}{2\, \pi\, W}]^{\,4} \ ,$$

$$\boldsymbol{\alpha} = \mathbf{8.1} \times \mathbf{10^{-3}} \, .$$

By examining equation 3.15, it is interesting to point out that the significant wave height is proportional to the square of wind speed (or $\boldsymbol{W^2}$). With the inputs of wind speed, the derived PM model (equation 3.15) is applied to all four stations using data from the year 2002. The whole set of data is used in the model testing. Time variations of predicted and observed significant wave height for stations LZ40, L006,L005, and L001 are showed respectively in Figures 3.30, 3.31, 3.32, and 3.33. The comparison between predicted and observed significant wave height within the perfect fitted line plots are shown in Figure 3.34 to Figure 3.37.

It is visible from the figures that PM Spectrum cannot be applied to shallow lake like Lake Okeechobee. Significant wave heights are significantly over predicted for all stations by using the original PM spectrum. Then, the question is can the PM spectrum be modified or extended to develop a new wave spectrum model to reasonably predict the significant wave height in Lake Okeechobee.

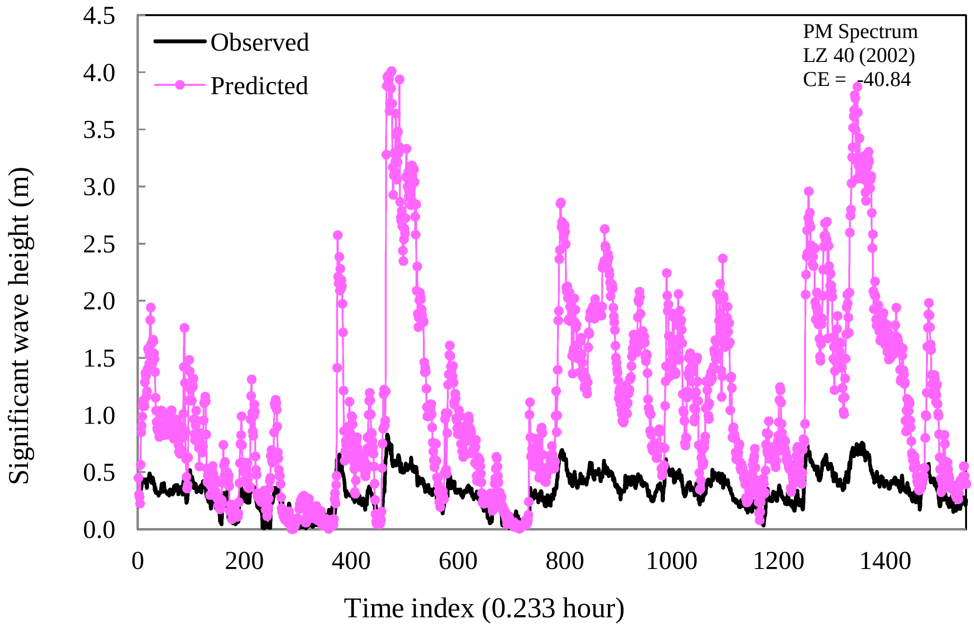

Figure 3.30 Predicted and observed time variations of significant wave heights at LZ40 using PM Spectrum

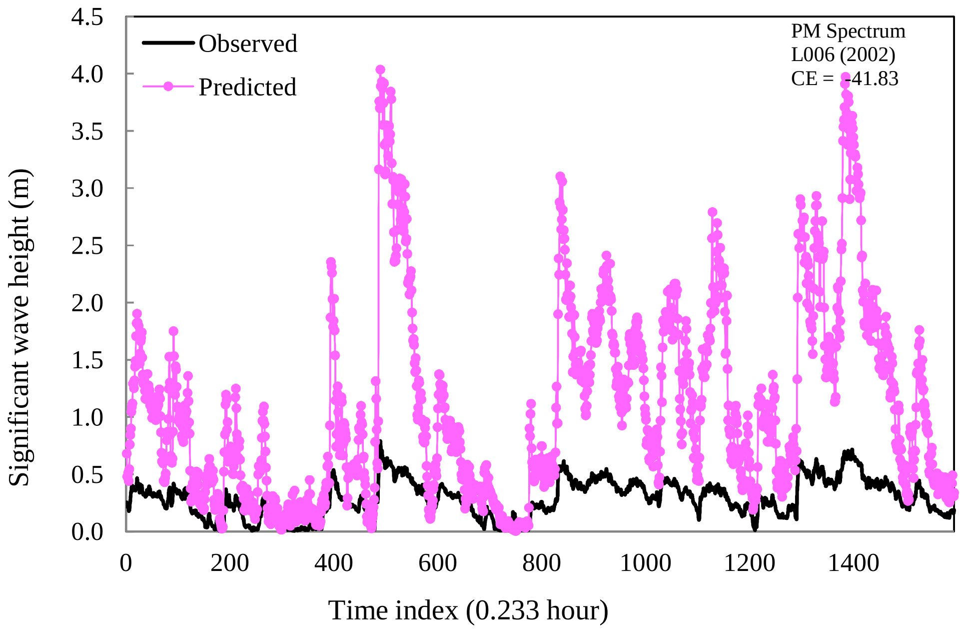

Figure 3.31 Predicted and observed time variations of significant wave heights at L006 using PM Spectrum

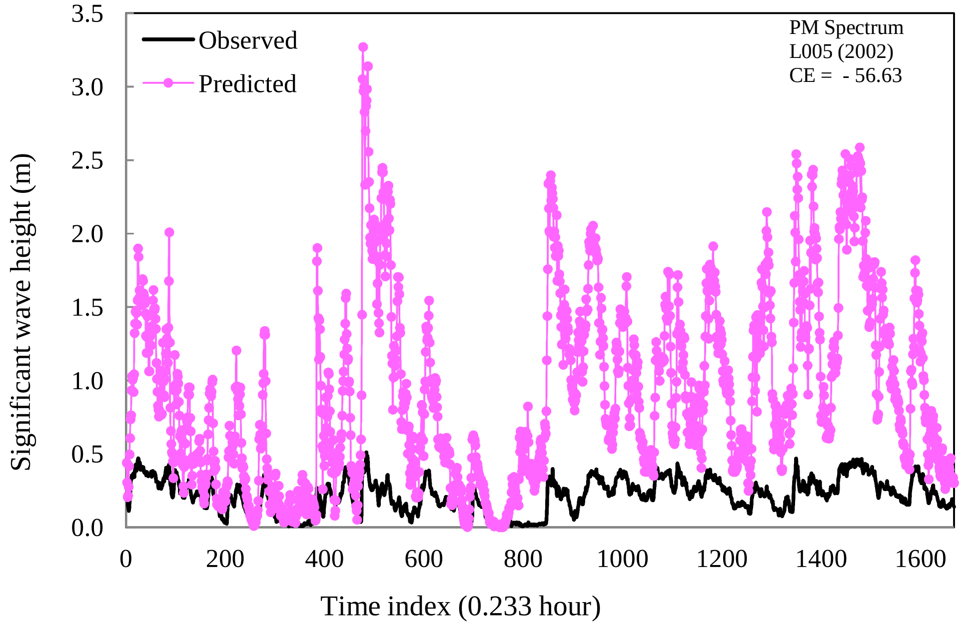

Figure 3.32 Predicted and observed time variations of significant wave heights at L005 using PM Spectrum

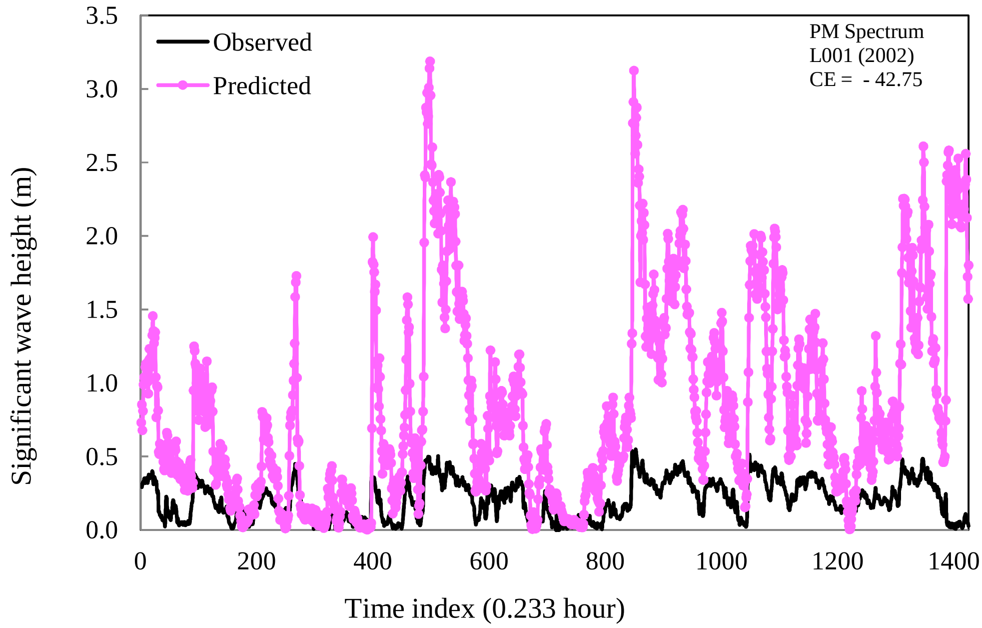

Figure 3.33 Predicted and observed time variations of significant wave heights at L001 using PM Spectrum

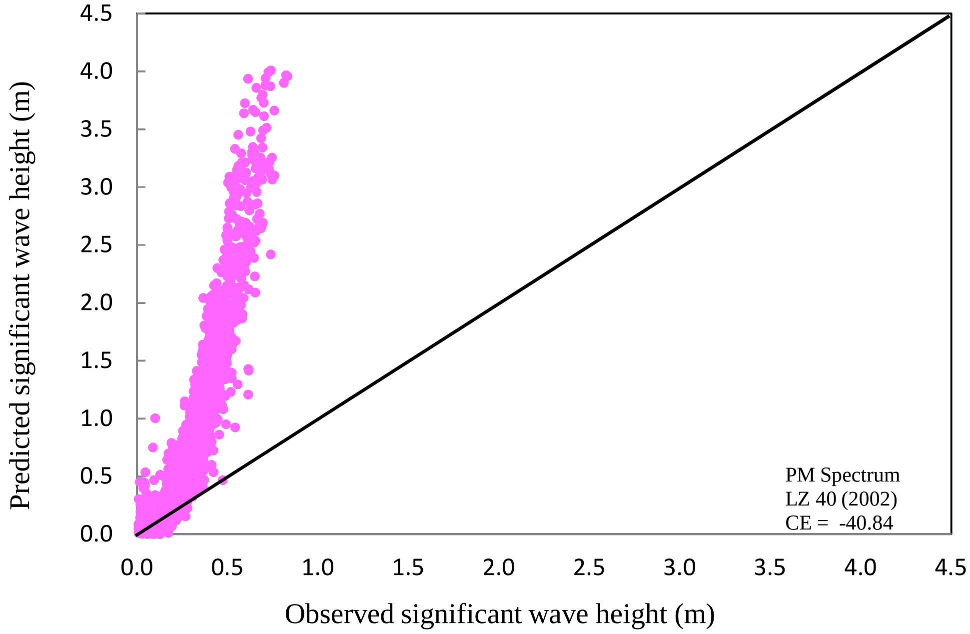

Figure 3.34 Verification of predicted and observed significant wave heights using PM Spectrum at station LZ40

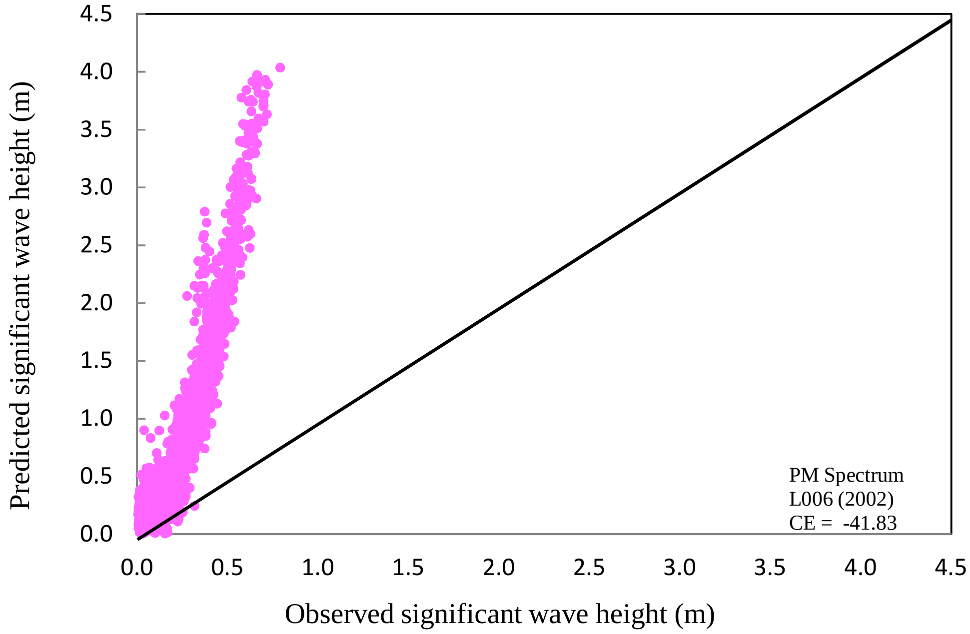

Figure 3.35 Verification of predicted and observed significant wave heights using PM Spectrum at station L006

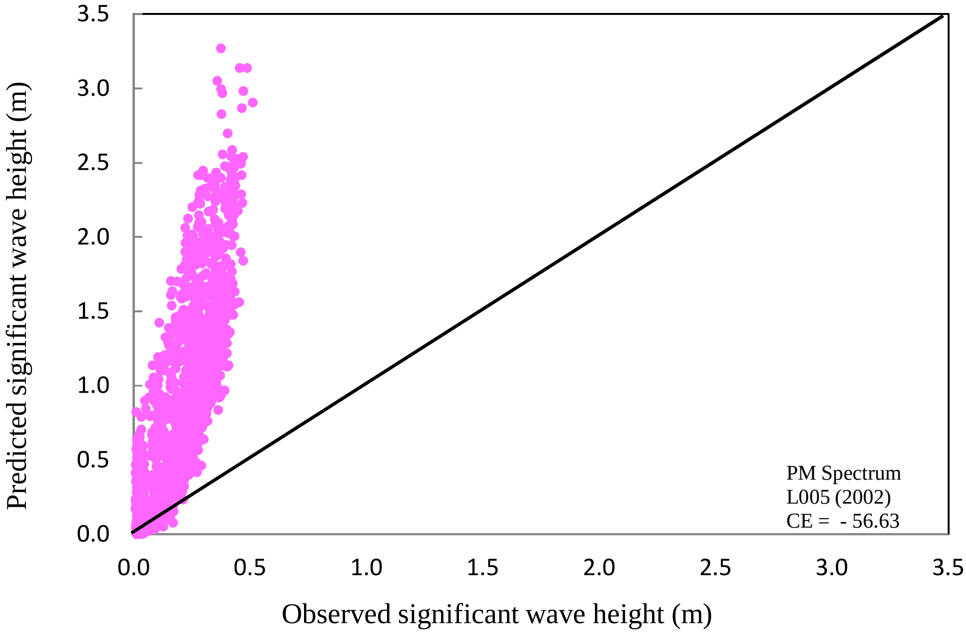

Figure 3.36 Verification of predicted and observed significant wave heights using PM Spectrum at station L005

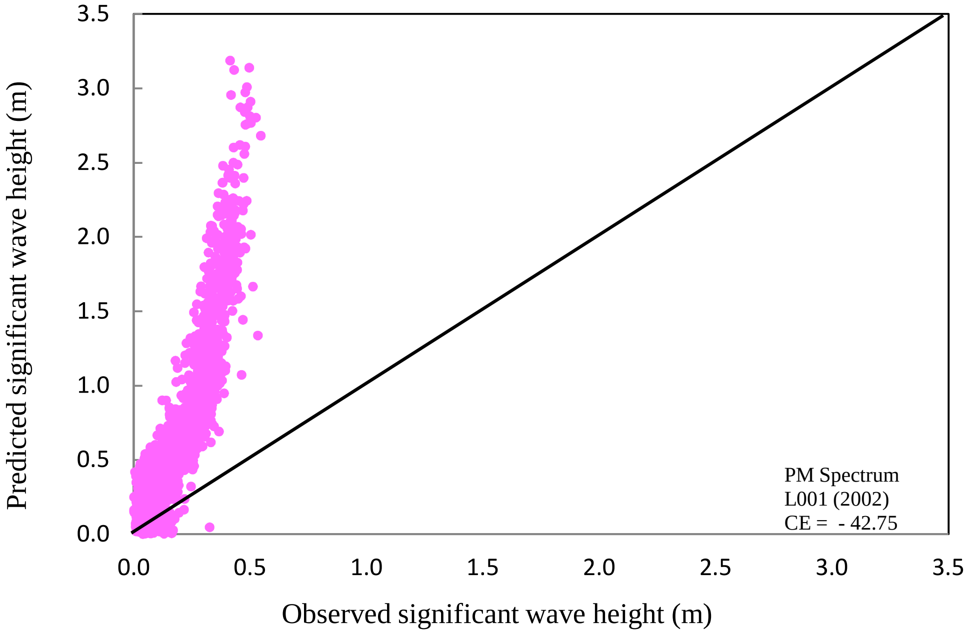

Figure 3.37 Verification of predicted and observed significant wave heights using PM Spectrum at station L001

3.5 Development of a new Wave Spectrum Model - Modified Pierson-Moskowitz (MPM) Spectrum Model

In the PM Spectrum, the Phillips constant, $\boldsymbol{\alpha}$ (= 8.1 x 10^{-3}) and -5 power laws for wave frequency $\boldsymbol{f}$ are adopted for the fully developed seas. In order to develop a new wave spectrum model for the prediction of significant wave height at stations LZ40, L006, L005 and L001in Lake Okeechobee, wave spectra with different powers of wave frequency have been tested. It is noticed that the -3.5 powers of wave frequency can be developed by modifying the coefficient $\boldsymbol{A}$ and $\boldsymbol{B}$ in PM spectrum. The explicit expression can also be obtained to calculate the significant wave height. The Modified PM (MPM) spectrum is proposed as

$$\boldsymbol{S(f)} = \frac{\boldsymbol{A_1}}{\boldsymbol{f^{3.5}}}\, \boldsymbol{e}^{[\frac{-\boldsymbol{B_1}}{\boldsymbol{f^{2.5}}}]} \quad , \tag{3.16}$$

where

$$\boldsymbol{A_1} = \frac{\boldsymbol{\alpha_1\, g} \sqrt{\boldsymbol{v}}}{(\boldsymbol{2\,\pi})^{2.5}} \quad , \tag{3.17a}$$

$$\boldsymbol{B_1} = \boldsymbol{\beta_1}\, [\frac{\boldsymbol{g}}{\boldsymbol{2\,\pi\, W}}]^{\,2.5} \; . \tag{3.17b}$$

Similar to equation (3.14), integration of equation (3.16) leads to

$$\boldsymbol{H_s} = \boldsymbol{4}\sqrt{\frac{\boldsymbol{A_1}}{\boldsymbol{2.5\, B_1}}} \; . \tag{3.18}$$

The values of $\boldsymbol{\alpha_1}$ and $\boldsymbol{\beta_1}$ are unknowns and need to be determined by the calibration procedure**.** $\boldsymbol{v}$ (kinetic viscosity of water at 20^{o} C) = 1.004 x 10^{-6} m^2/s, $\boldsymbol{g}$ is the gravitational acceleration, and $\boldsymbol{W}$ is again the wind speed measured in m/s. According to the modified PM spectrum, the significant wave height as shown in equation (3.18) is found to be proportional to the 1.25 power of wind speed ($\boldsymbol{W^{1.25}}$)**.** Different from PM spectrum model in equation (3.15), equation (3.18) is considered as the Modified PM

(MPM) model. To find out the similarity between these two equations, equation (3.18) can be written as follow

$$\boldsymbol{H_s = 4\sqrt{\frac{A_2}{4\,B_1}}}, \tag{3.19}$$

where

$$\boldsymbol{A_2 = \frac{4\,A_1}{2.5}}\ . \tag{3.20}$$

The mathematical expression in equation (3.19) is similar to equation (3.15). Then, $\boldsymbol{A_2}$ can be assumed to be equal to $\boldsymbol{A}$ (equation 3.8a) as

$$\boldsymbol{A_2 = A}\ , \tag{3.21}$$

and

$$\boldsymbol{\frac{4\,A_1}{2.5} = \frac{\alpha\,g^2}{16\,\pi^4}}\ . \tag{3.22}$$

Substituting equation (3.17a) into equation (3.22) gives

$$\boldsymbol{\frac{4}{2.5}\frac{\alpha_1\,g\,\sqrt{\upsilon}}{(2\,\pi)^{2.5}} = \frac{\alpha\,g^2}{16\,\pi^4}}\ . \tag{3.23}$$

We have

$$\boldsymbol{\alpha_1 = \frac{\alpha\,g^2(2\,\pi)^{2.5}\,(2.5)}{(4)\,(16\,\pi^4)\,g\,\sqrt{\upsilon}}}\ . \tag{3.24}$$

By replacing all the known values to the right side of the above equation, the value of $\boldsymbol{\alpha_1}$ is found to be

$$\boldsymbol{\alpha_1 = 3.147}\ .$$

From equation (3.18), we have

$$\boldsymbol{H_s = 4\sqrt{\frac{A_1}{2.5\,B_1}} = 4\sqrt{\frac{1}{2.5}\,\frac{\alpha_1\,g\,\sqrt{\upsilon}}{(2\,\pi)^{2.5}}\,\frac{(2\,\pi\,W)^{2.5}}{\beta_1\,g^{2.5}}} = 4\sqrt{\frac{\sqrt{\vartheta}}{2.5\,g^{1.5}}}\,\sqrt{\frac{\alpha_1}{\beta_1}}\,W^{1.25}}. \tag{3.25}$$

Finally, the MPM model equation for the computation of significant wave height can be expressed as

$$\boldsymbol{H_s} = \boldsymbol{D}^* \boldsymbol{W}^{\mathbf{1.25}} \ , \tag{3.26}$$

where

$$\boldsymbol{D}^* = \mathbf{4} \sqrt{\frac{\sqrt{\boldsymbol{v}}}{\mathbf{2.5}\, \boldsymbol{g}^{\mathbf{1.5}}}} \sqrt{\frac{\boldsymbol{\alpha_1}}{\boldsymbol{\beta_1}}} \ . \tag{3.27}$$

With known $\boldsymbol{\alpha_1}$ value, the least square method is applied with observed significant wave height data to equation (3.26), to obtain the calibrated parameter $\boldsymbol{\beta_1}$ for stations LZ40, L006, L005 and L001. The training data set described above is used to find the parameter $\boldsymbol{\beta_1}$. To further the comparisons, the testing data and the data of the year 1996 are used for model validation. The validation of the MPM model is detailed in Chapter 4. The individual values of $\boldsymbol{\beta_1}$ for all stations are summarized in Table 3.5.

Table 3.5 Values $\boldsymbol{\alpha_1}$ and $\boldsymbol{\beta_1}$ for the stations LZ40, L006, L005 and L001

Station	Model	Modified PM Spectrum (MPM)	
	Parameters	α_1	β_1
LZ40		3.147	0.71
L006			1.26
L005			1.81
L001			1.65

The Modified PM (MPM) Spectrum model for each station is applied to check the model performance in terms of the prediction of the significant wave height by comparing the model results with the training data. Time variations of predicted and observed significant wave heights at LZ40, L006, L005, and L001 are shown in Figures 3.38, 3.39, 3.40 and 3.41 respectively. The variation trend of the predicted significant wave height fits reasonable well with observed data at stations LZ40 (CE = 0.84) and L006 (CE = 0.81). For station L005, which is near the marshy area, the MPM model mostly underestimate the significant wave height. Except at the time index around 800 (circled in Figure 3.40), the model overestimate the wave height. The MPM model produces reasonable time varying trend for the significant wave height at station L001, however, the model also underestimates the values of the wave height. The direct comparisons between MPM predicted and observed significant wave height are shown in Figure 3.42 to Figure 3.45. Again, Figures 3.43 and 3.45 show the model's underestimation of the values of significant wave height for stations L006 and L001. Larger scatter of the data points when compared to the perfect 45 degree fitted line is shown in Figure 3.44 for station L005.

The statistical values of AAE, RMSE, and CE for the indication of the MPM model performance are summarized in Table 3.6. By examining the figures and the values in Table 3.6, Modified PM Spectrum shows acceptable results. Like regression model RM1, Modified PM Spectrum also produces relatively poor prediction of significant wave height at the marshy L005 area, especially the over-prediction shown at the time index around 500 (marked with a circle). More verification and validation study for the MPM model is given in Chapter 4.

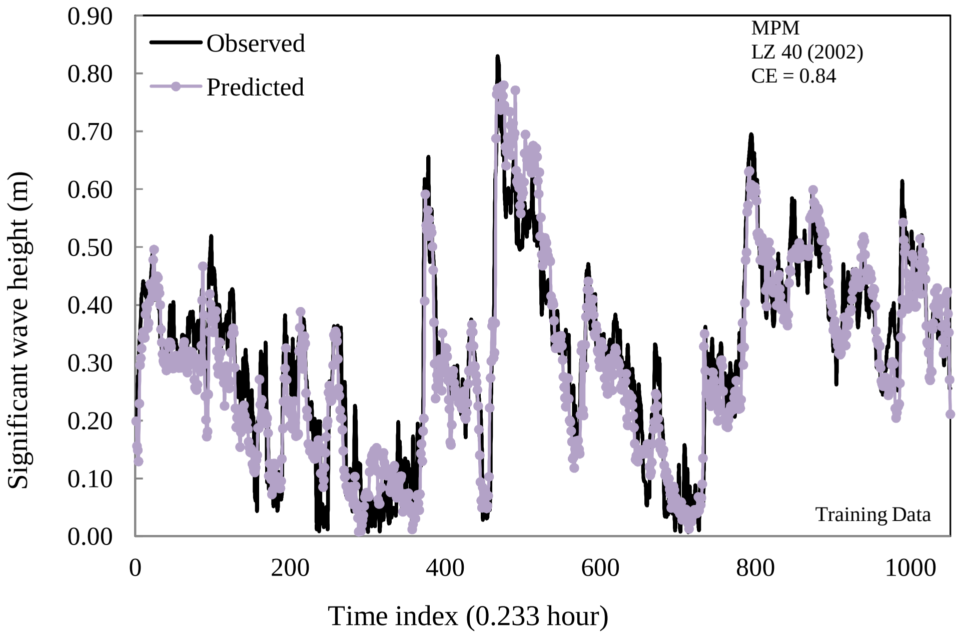

Figure 3.38 Predicted and observed time variations of significant wave heights at LZ40 using Modified PM Spectrum

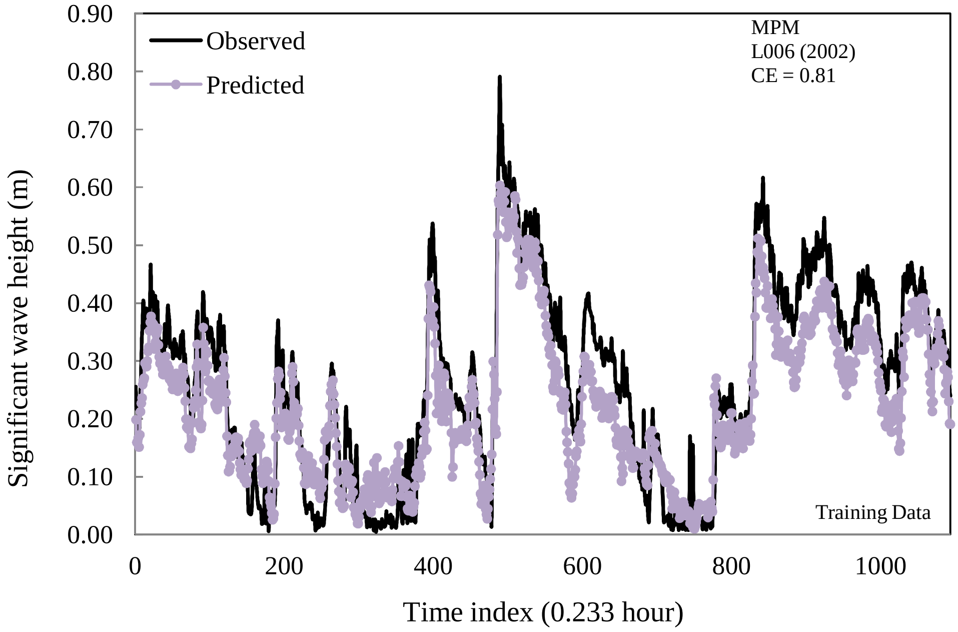

Figure 3.39 Predicted and observed time variations of significant wave heights at L006 using Modified PM Spectrum

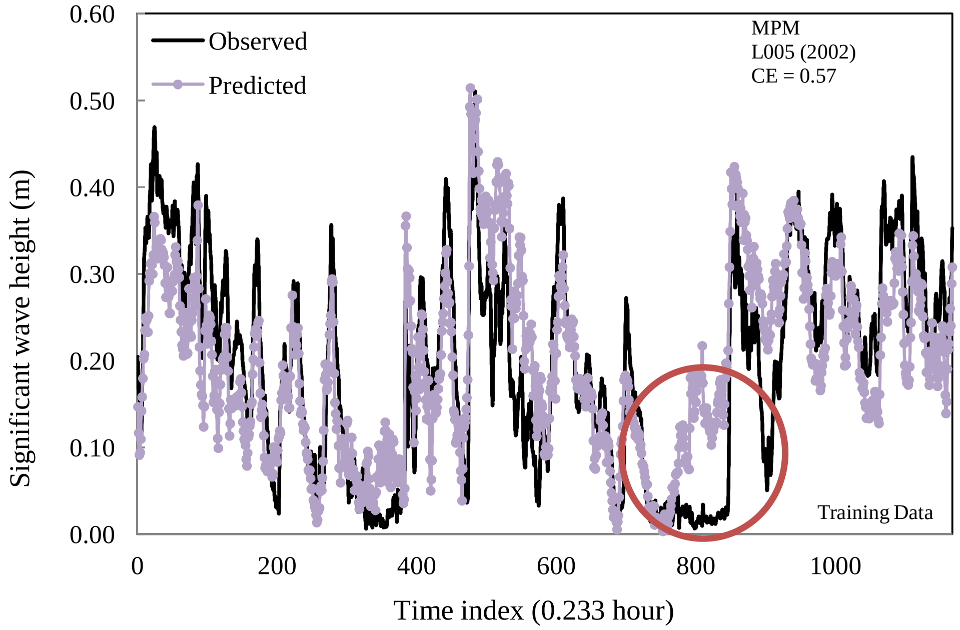

Figure 3.40 Predicted and observed time variations of significant wave heights at L005 using Modified PM Spectrum

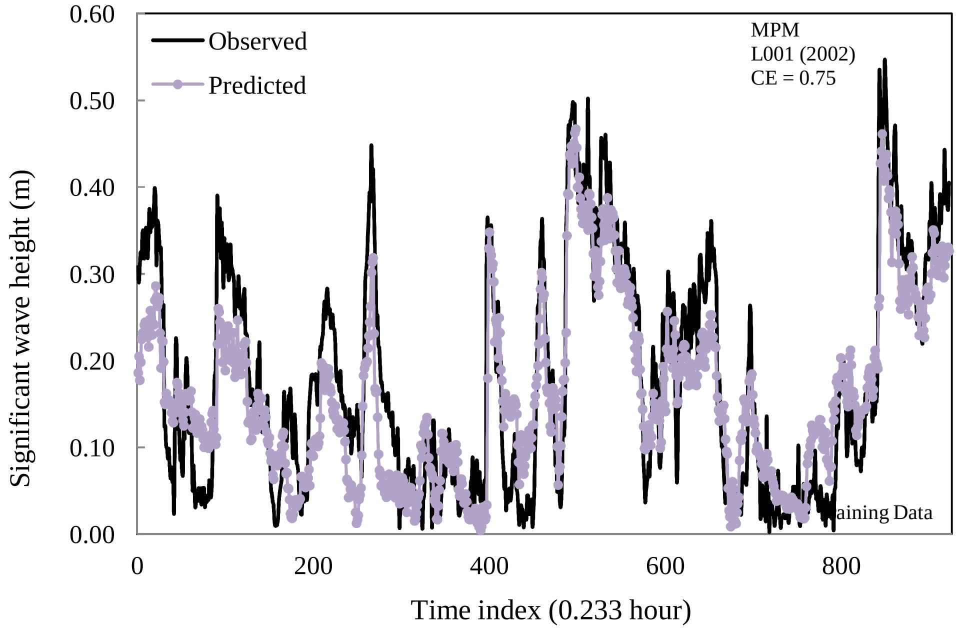

Figure 3.41 Predicted and observed time variations of significant wave heights at L001 using Modified PM Spectrum

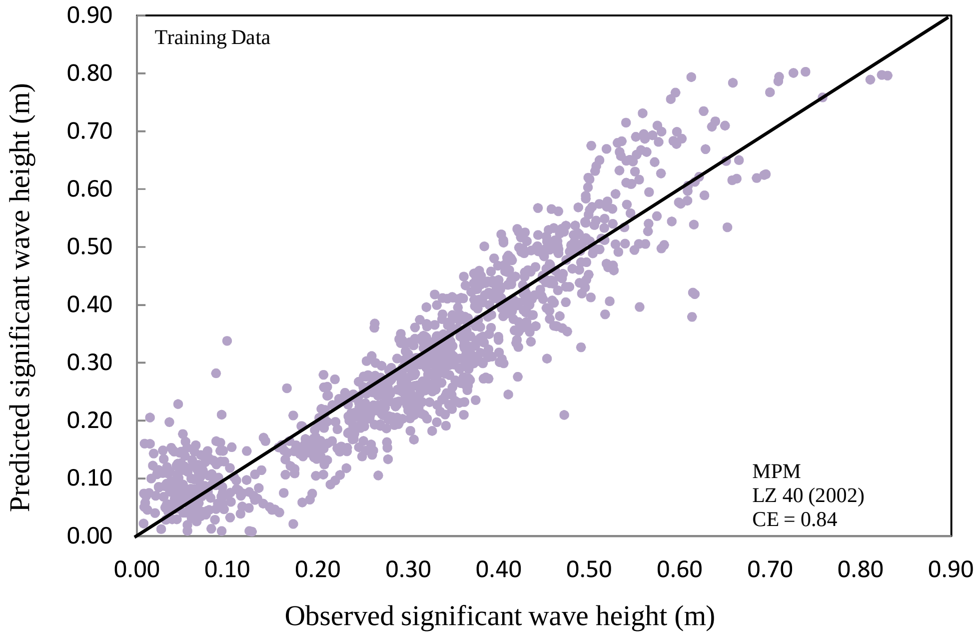

Figure 3.42 Verification of predicted and observed significant wave heights using Modified PM Spectrum at station LZ40

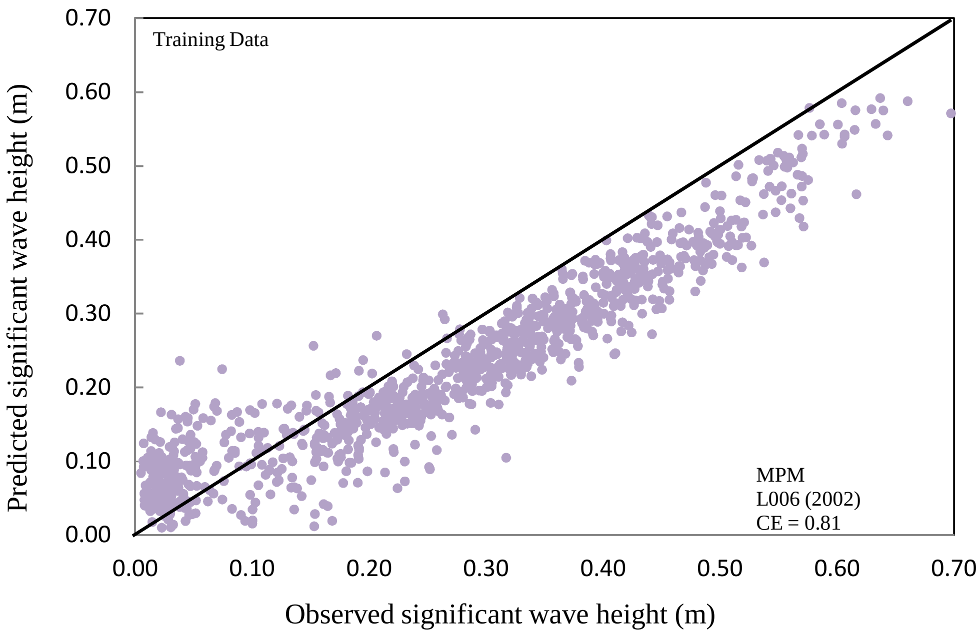

Figure 3.43 Verification of predicted and observed significant wave heights using Modified PM Spectrum at station L006

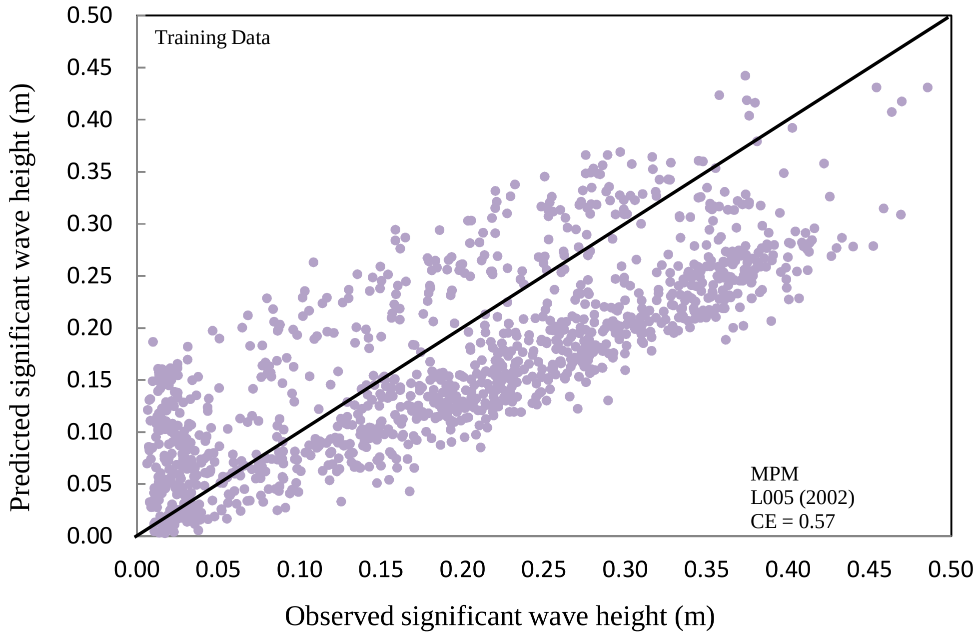

Figure 3.44 Verification of predicted and observed significant wave heights using Modified PM Spectrum at station L005

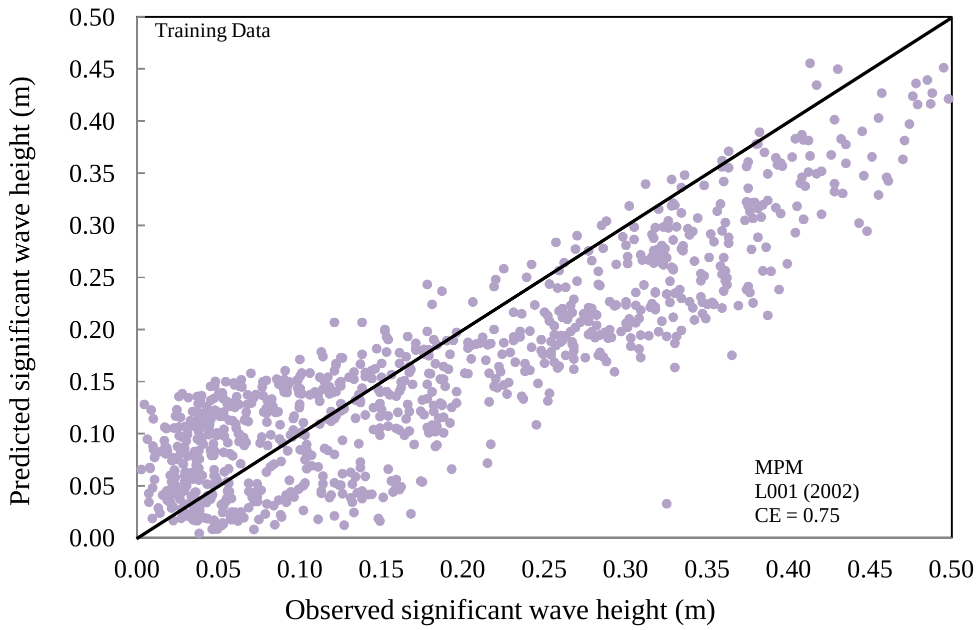

Figure 3.45 Verification of predicted and observed significant wave heights using Modified PM Spectrum at station L001

Table 3.6 Values of AAE, RMSE and CE of MPM using training data for the station LZ40, L006, L005 and L001

	MPM		
	AAE (m)	RMSE (m)	CE
LZ40	0.05	0.06	0.84
L006	0.06	0.07	0.81
L005	0.07	0.08	0.57
L001	0.05	0.07	0.75

3.6 Effects of time shifting of input variable on the performance of model RM1 and RM2

The time shifting effects on the Model RM1 and RM2 are tested and a chart with the values of coefficient of efficiency at different time steps are produced for all four stations. Training data from the year 2002 are used for the numerical experiments. The coefficient of efficiency is calculated for cases considering the time shift from current time step (e.g. wind speed as the input variable) to the 10^{th} time step prior to the current time. Figure 3.46 to Figure 3.49 represent the performance chart for RM1 at station LZ40, L006, L005 and L001 respectively. Similar charts are plotted for RM2 at all four stations (Figure 3.50 to Figure 3.53).

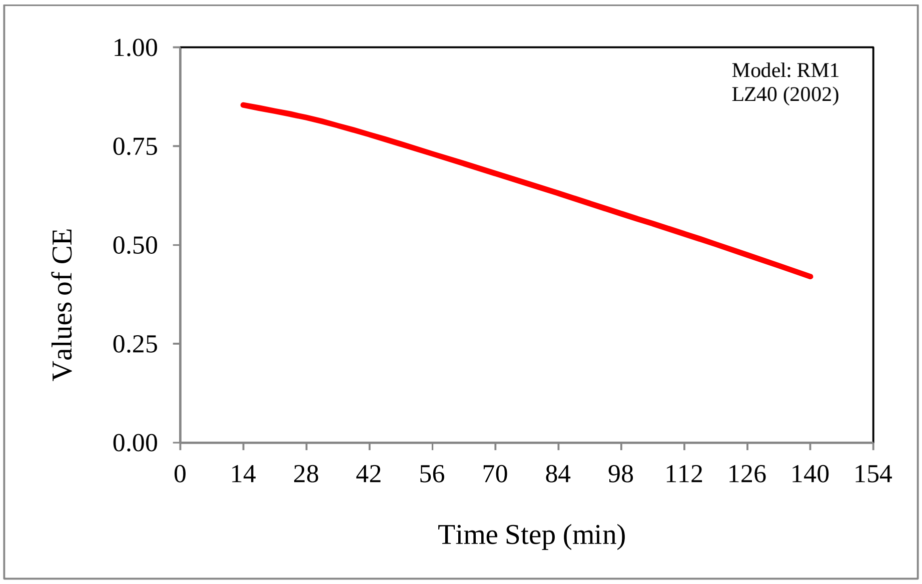

Figure 3.46 Time shifting effects on the performance of RM 1 (LZ40)

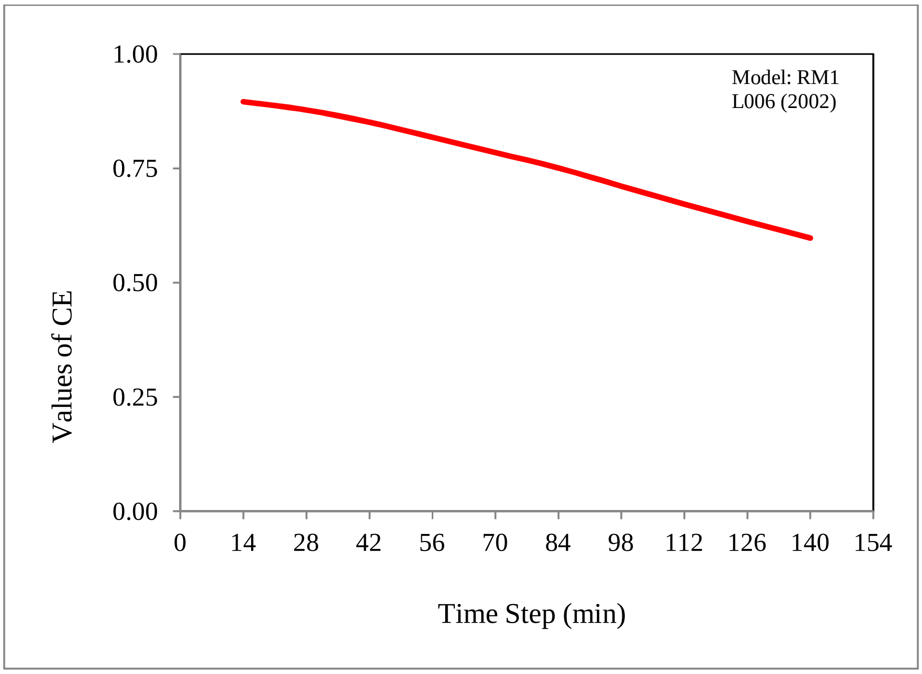

Figure 3.47 Time shifting effects on the performance of RM 1 (L006)

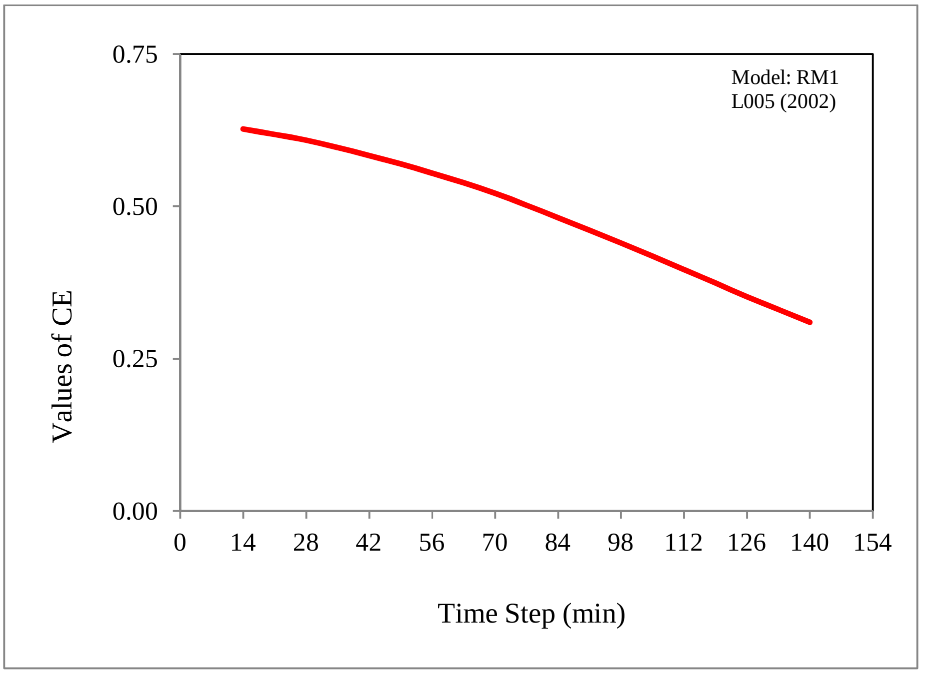

Figure 3.48 Time shifting effects on the performance of RM 1 (L005)

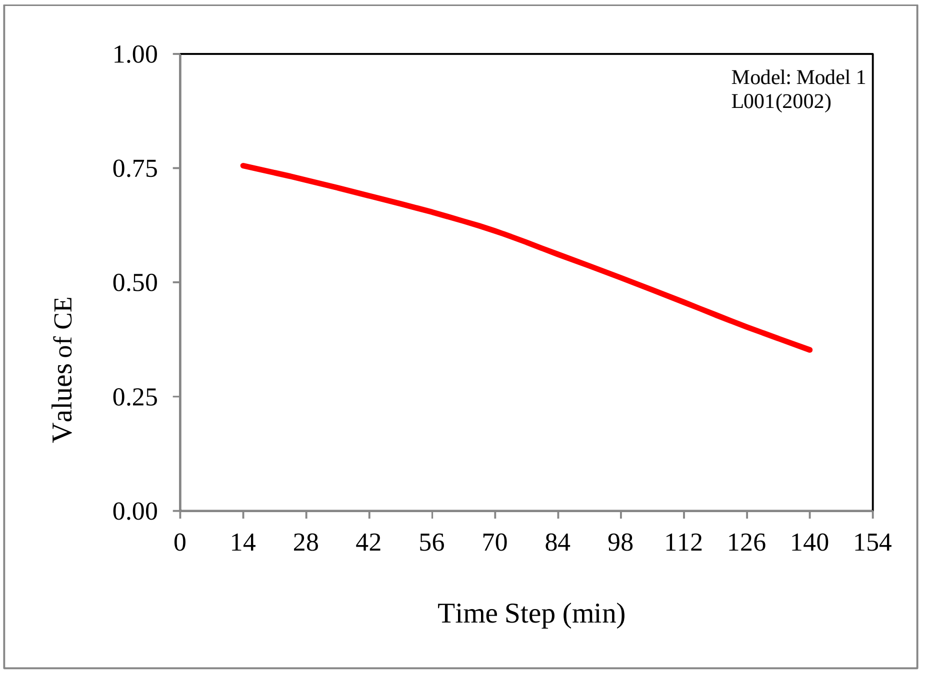

Figure 3.49 Time shifting effects on the performance of RM 1 (L001)

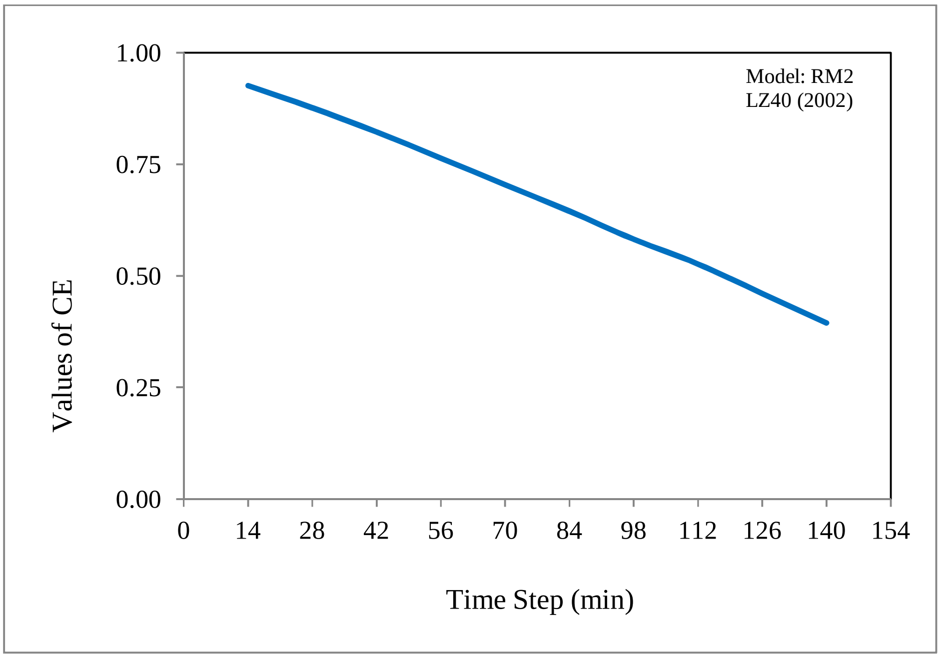

Figure 3.50 Time shifting effects on the performance of RM 2 (LZ40)

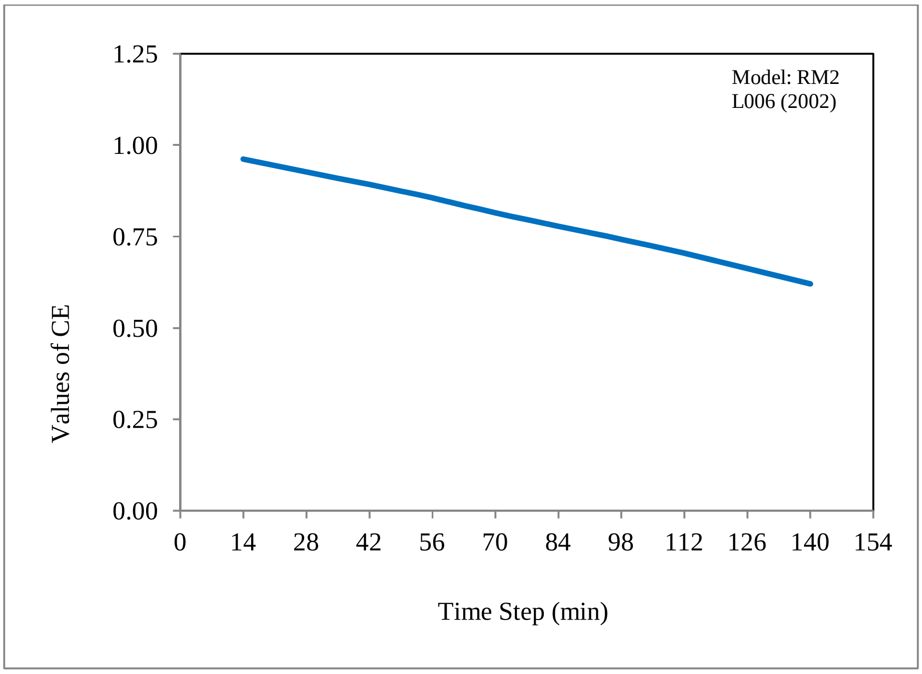

Figure 3.51 Time shifting effects on the performance of RM 2 (L006)

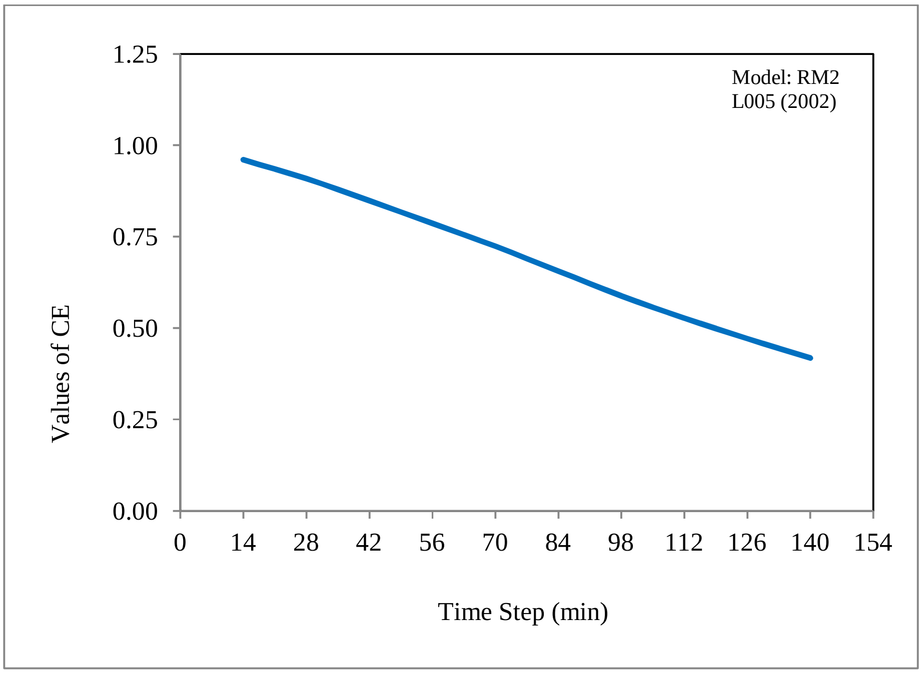

Figure 3.52 Time shifting effects on the performance of RM 2 (L005)

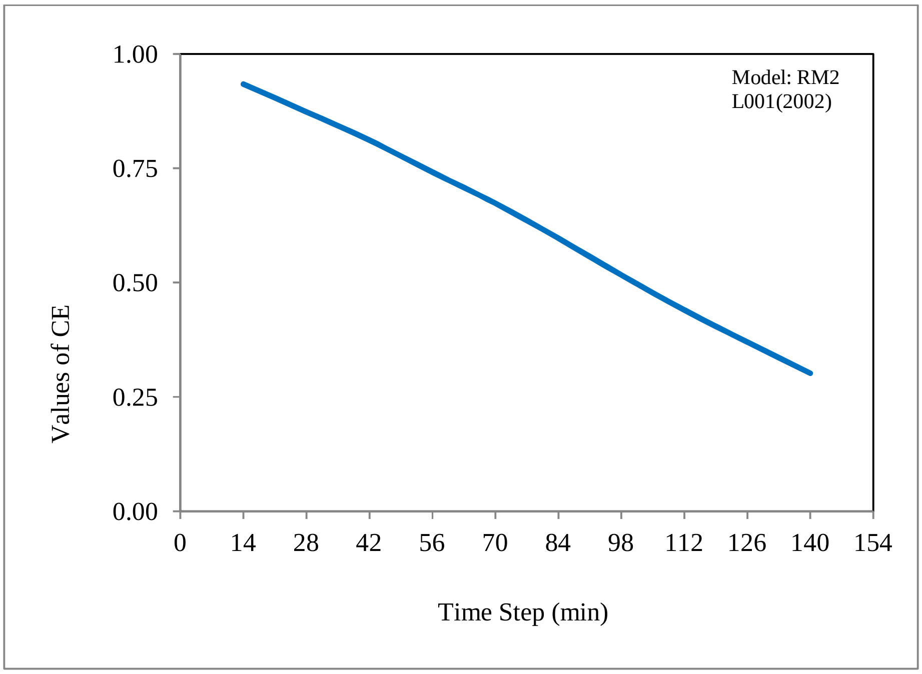

Figure 3.53 Time shifting effects on the performance of RM 2 (L001)

From the charts we can clearly perceive that, current data produces more accurate results. At the event of the unavailability of current data, the performance chart for any particular station can be used to determine whether the available data from prior time step are eligible to use to predict current or future significant wave height or not.

3.7 SUMMARY

Significant wave heights for Lake Okeechobee are predicted using four proposed models, which are RM1, RM2, PLSM and MPM models. Initial verification of the predictive models with obtained model parameters are done with the training data as discussed in previous sections. For further verification and validation of the performance and applicability of the developed models, they are tested using the remaining 500 testing data from the data collected in year 2002 and another set of 140 independent data from the year 1996. Study of the additional data and detailed comparisons of the predicated significant wave height from the four proposed models, as summarized below, are presented in Chapter 4,

RM1 model $\boldsymbol{H_K = a_1\, W_K + b_1}$,

RM2 model $\boldsymbol{H_K = a_2\, H_{K-1} + b_2}$,

PLSM model $\boldsymbol{H_K = a_3\, H_{K-1} + b_3\, W_{K-1}}$,

MPM Spectrum model $\boldsymbol{H_s = 4\sqrt{\frac{\sqrt{v}}{2.5\, g^{1.5}}}\sqrt{\frac{\alpha_1}{\beta_1}}\, W^{1.25}}$.

Chapter 4

Results and Comparison Discussions

4.1 Introduction

As mentioned previously, in this chapter results in the form of time variation of predicted significant wave height from all the models, including Regression Model 1 (RM1) and Model 2 (RM2), Perceptron Least Square Method (PLSM), and Modified Pierson-Moskowitz Spectrum (MPM) using other independent data in Lake Okeechobee are presented to further the validation of the proposed models in terms of prediction of significant wave height. The predicted and observed data of all models are also compared directly in 45 degree perfect-fit line plots. Summaries of AAE, RMSE and CE for each model are shown in Table format.

Altunkaynak and Wang (2012) applied Artificial Neural Network (ANN) model to Lake Okeechobee using also the 2002 and 1996 data described in this study. To get a better understanding about the performance of the models used in this study, comparisons are made with recorded data. Additionally a comparison is made with all the models (i.e. RM1, RM2, PLSM and MPM) and the ANN model developed by Altunkaynak and Wang (2012). Comparison in a single chart of time series of predicted and observed data for each station is shown in the last section of this chapter.

4.2 Regression Method (RM1 and RM2)

Time variations of predicted and observed significant wave height calculated by RM1 and RM2 using 500 testing data (year 2002) for stations LZ40, L006, L005 and L001 are showed in Figures 4.1, 4.2, 4.3 and 4.4 respectively. From Figures 4.1 to 4.4, it can be clearly seen that the prediction of significant wave height by RM2 fits well with observed data for all stations. Prediction of significant wave height by RM1 at LZ40 also provides good results (Figures 4.1). But at stations L006, L005 and L001, the RM1 results generally follow the trend of observed data, however, with some mismatched significant wave height values (Figures 4.2 and 4.4). Figures 4.5 to 4.7 represent the time variations of predicted and observed significant wave at LZ40, L006 and L005 respectively for the year 1996 data. Significant wave heights predicted by RM1 for the year 1996 data are shown to follow the variation trend with relatively larger oscillation. The RM1 results are generally acceptable while RM2 is able to produce a better variation trend and magnitude of predictions, however with a one hour shifting when compared to the recorded data (Figures 4.5 to Figure 4.7). From the results shown in Figures 4.1 to 4.7, it can be noticed that the significant wave height occurred at previous time level has a more dominant effect than the current wind speed for the prediction of current significant wave height.

Model validation is also carried out by plotting pairs of predicted and observed significant wave heights as a scatter diagram with an idealized slope equal to 45 degree (or the perfect fit line). The results are shown in Figures 4.8 to 4.15 for the 500 (year 2002) testing data and in Figures 4.16 to 4.21 for the year 1996 data. The scatter diagrams of RM2 using 500 testing data for all four stations show that the data points formed by the predicted and observed significant wave heights are narrowly distributed around the perfect fit line (Figures 4.9, 4.11, 4.13 and 4.15). This indicates that RM2 simulates better for Lake Okeechobee than RM1 (Figure 4.8, 4.10, 4.12 and 4.14). Figures 4.16 to 4.21 show the scatter diagram of RM1 and RM2 for the year 1996 data. In both cases, data are scattered with larger deviation away from the 45 degree line. The errors for the prediction of 1996 significant wave heights are greater than those for the 2002 data. Because of the time shifting in RM2 results, the RM2 model shows a wider scattering of the data comparisons in the range with larger significant wave height than the predictions from RM1 model.

Summary of AAE, RMSE and CE for the year 2002 testing data and the year 1996 data are shown in the Table 4.1 and 4.2 respectively. As it seen in the comparison figures, the statistical values shown in Table 4.1 also reveal that RM2 produces CE more than 90% for all four stations and comparatively smaller AAE and RMSE than RM1 for the year 2002 (500 testing data). RM1 produces poor CE of 0.06 for station L001. This may be caused by the measurement errors at the end of recorded data (See Figure 3.4 and 4.14). Coefficient of efficiency for LZ40, L006 and L005 using 1996 data are 0.82, 0.74 and 0.56 respectively for the model RM1 and 0.79, 0.75, and 0.72 for the model RM2 (Table 4.2). At station L005 near the marshy area, the performance of the wind speed based RM1 model as expected has a relatively low CE number. According to the CE values, the RM2 model can produce better predictions than those from the RM1 model at any location in Lake Okeechobee.

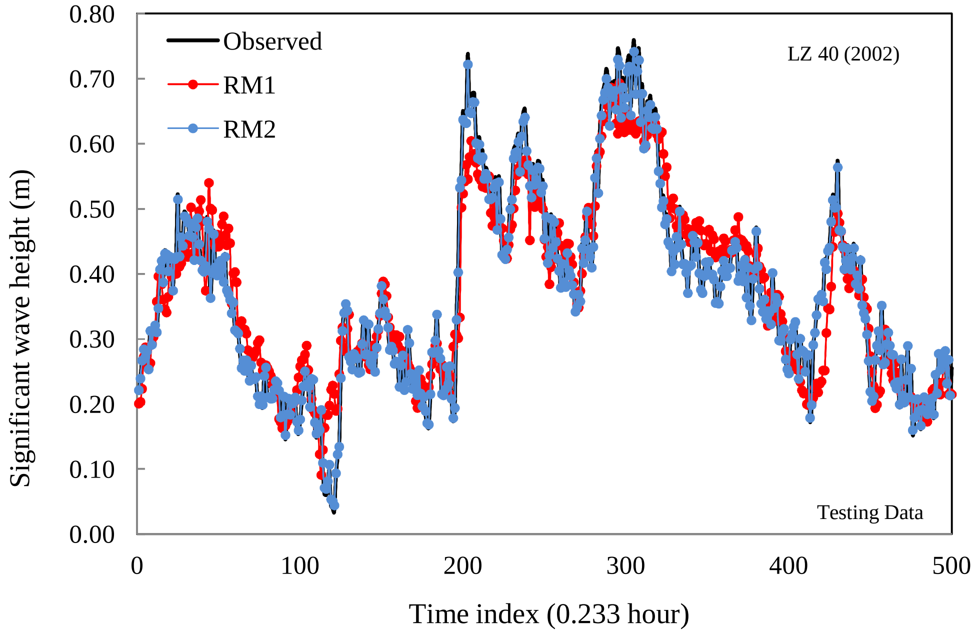

Figure 4.1 Time variations of predicted and observed significant wave height using RM1 and RM2 (500 tested data applied to LZ40 – 2002)

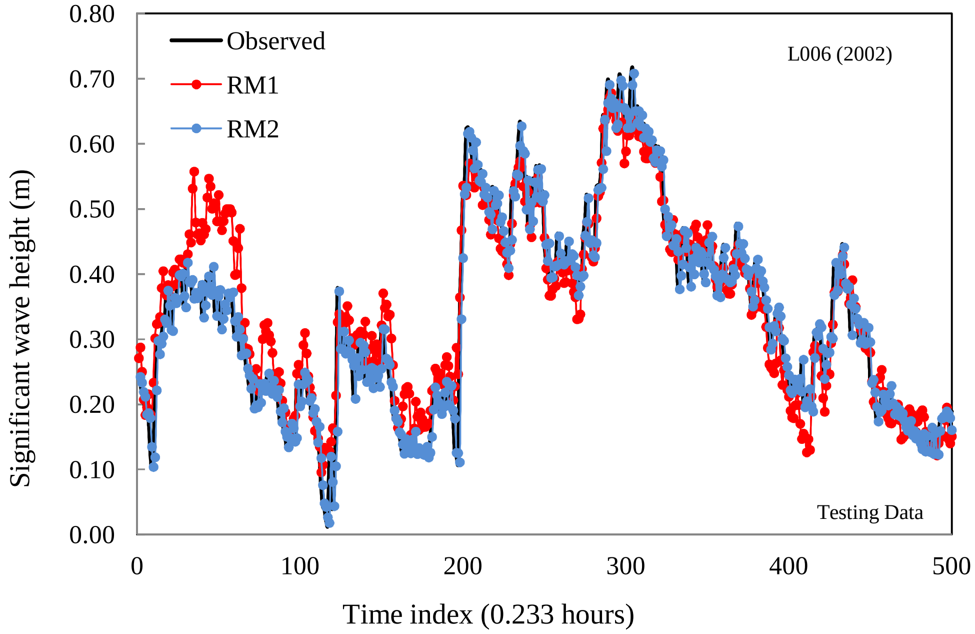

Figure 4.2 Time variations of predicted and observed significant wave height using RM1 and RM2 (500 tested data applied to L006 – 2002)

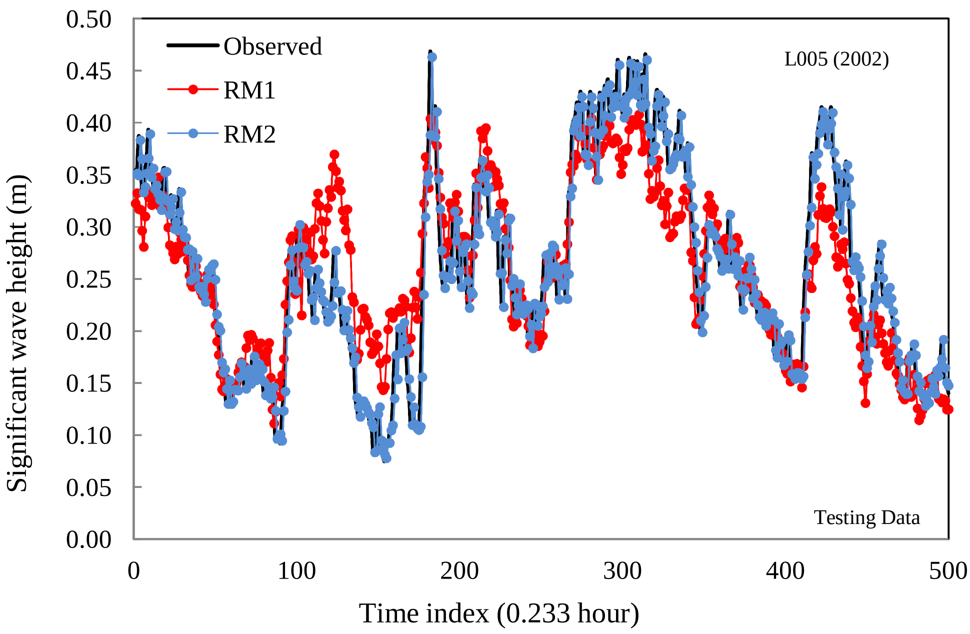

Figure 4.3 Time variations of predicted and observed significant wave height using RM1 and RM2 (500 tested data applied to L005 – 2002)

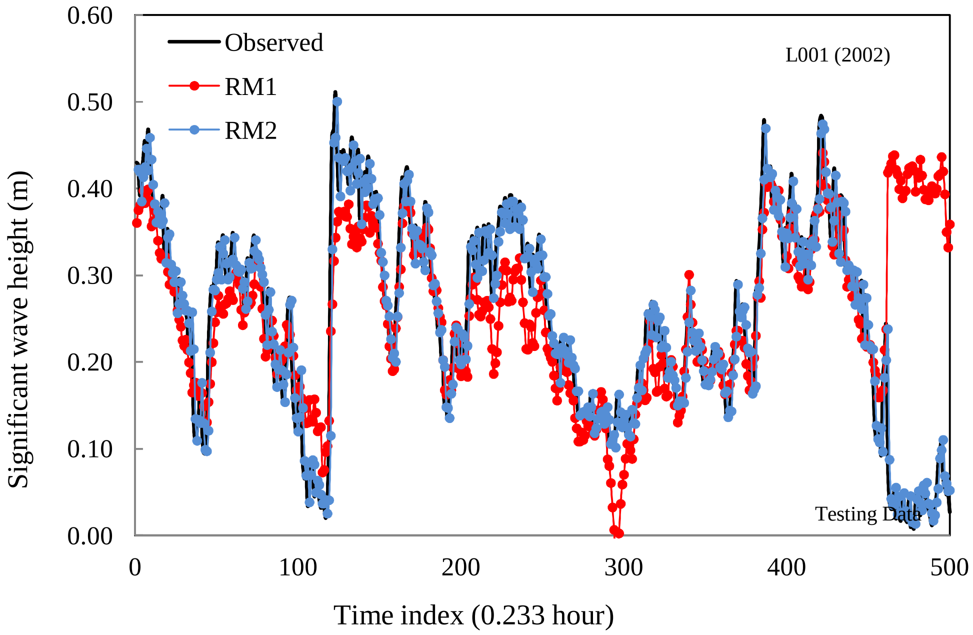

Figure 4.4 Time variations of predicted and observed significant wave height using RM1 and RM2 (500 tested data applied to L001 – 2002)

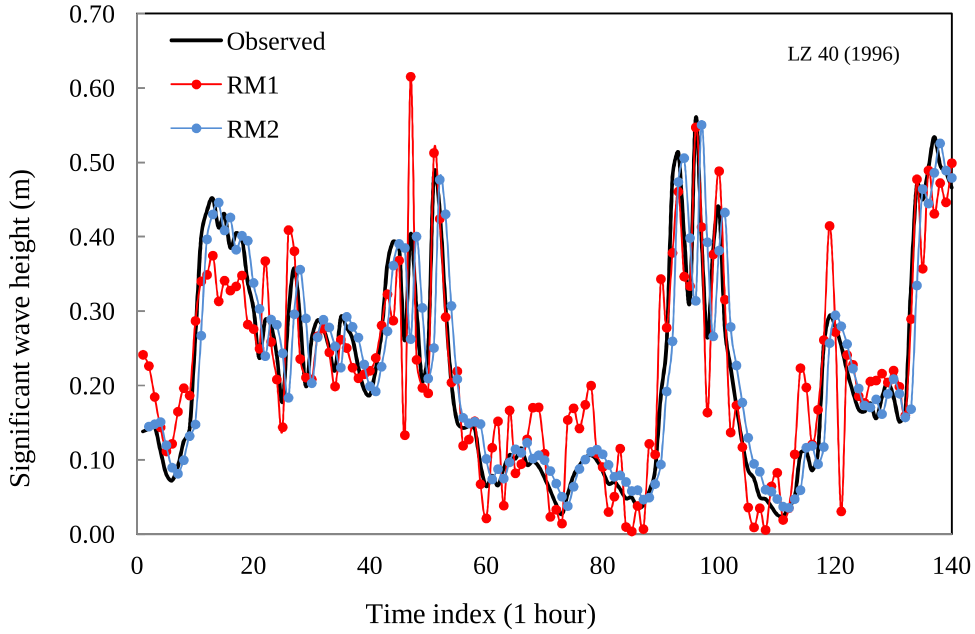

Figure 4.5 Time variations of predicted and observed significant wave height using RM1 and RM2 (140 data applied LZ40 – 1996)

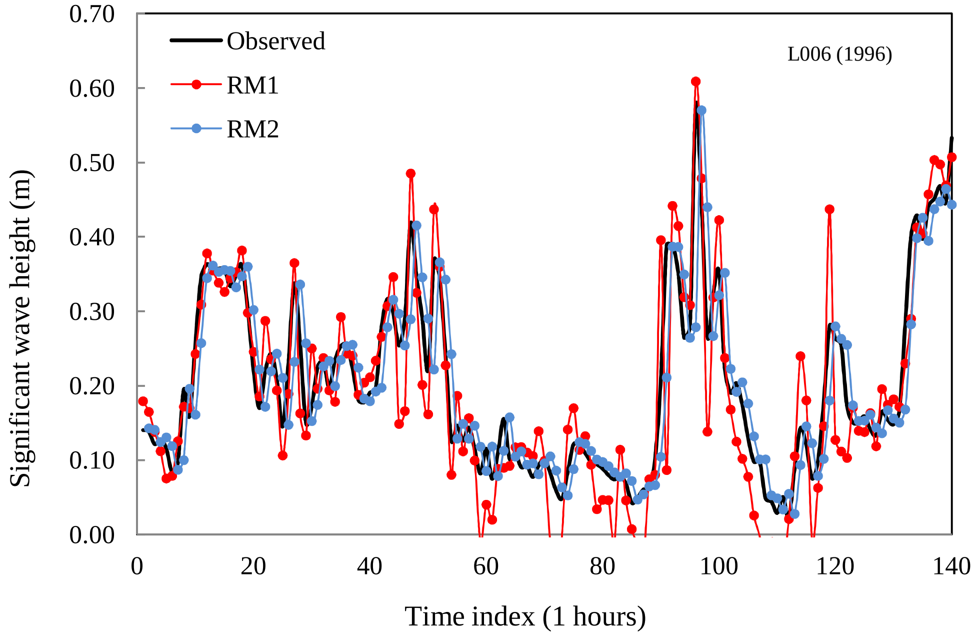

Figure 4.6 Time variations of predicted and observed significant wave height using RM1 and RM2 (140 data applied L006 – 1996)

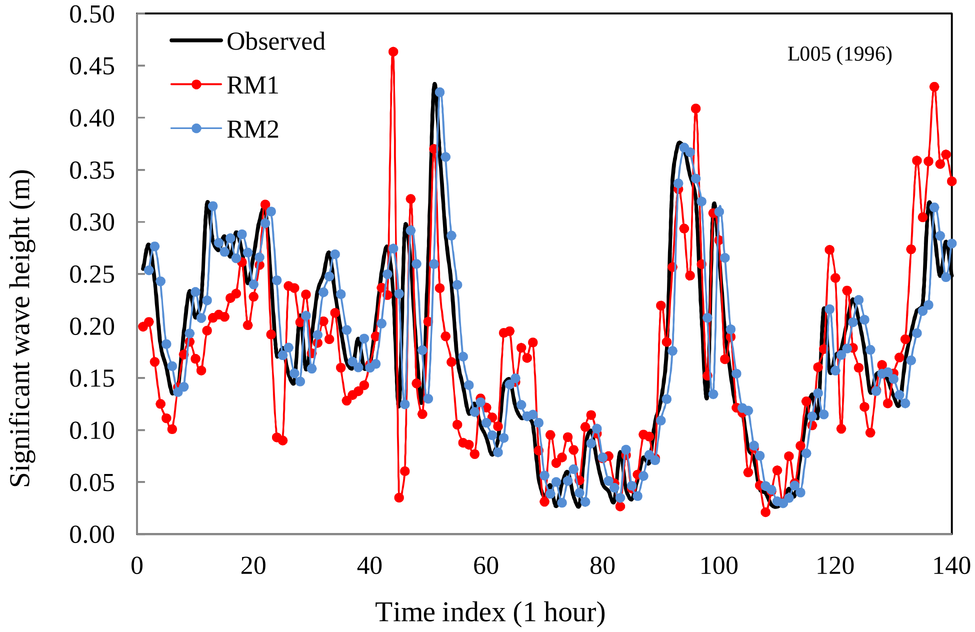

Figure 4.7 Time variations of predicted and observed significant wave height using RM1 and RM2 (140 data applied L005 – 1996)

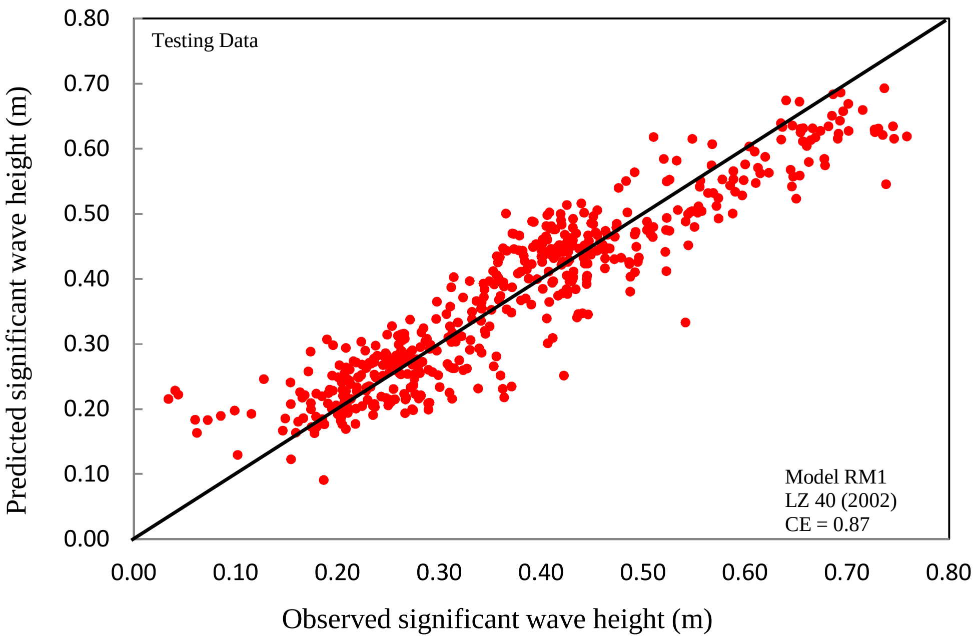

Figure 4.8 Verification of predicted and observed significant wave height using RM1 (500 testing data applied LZ40 – 2002)

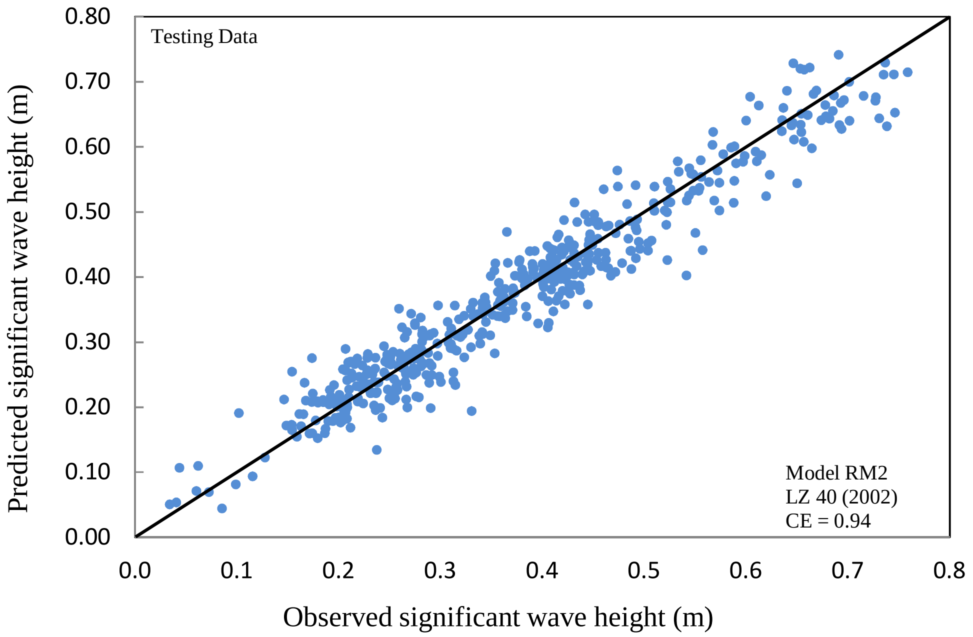

Figure 4.9 Verification of predicted and observed significant wave height using RM2 (500 testing data applied LZ40 – 2002)

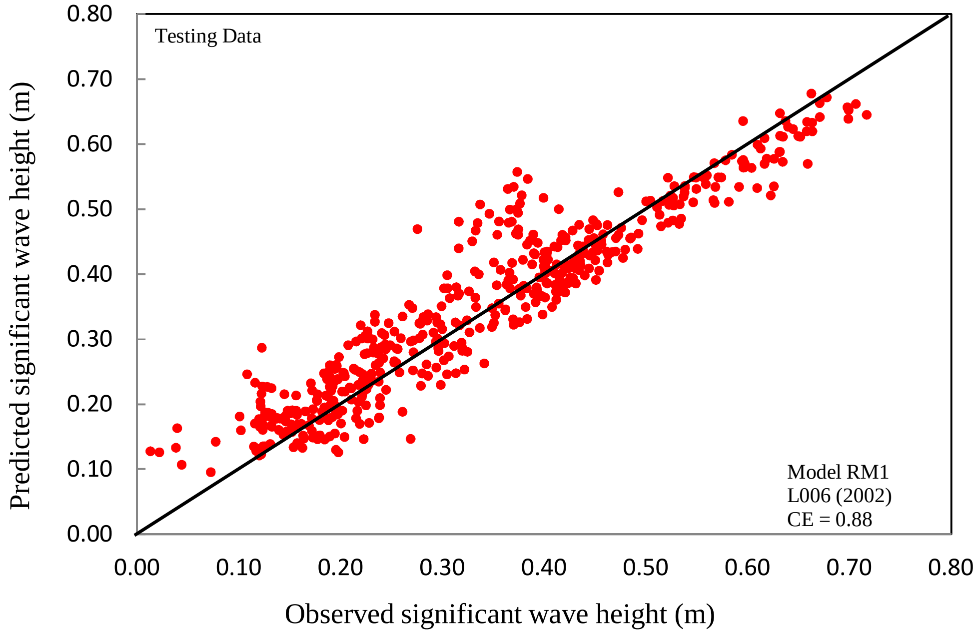

Figure 4.10 Verification of predicted and observed significant wave height using RM1 (500 testing data applied L006 – 2002)

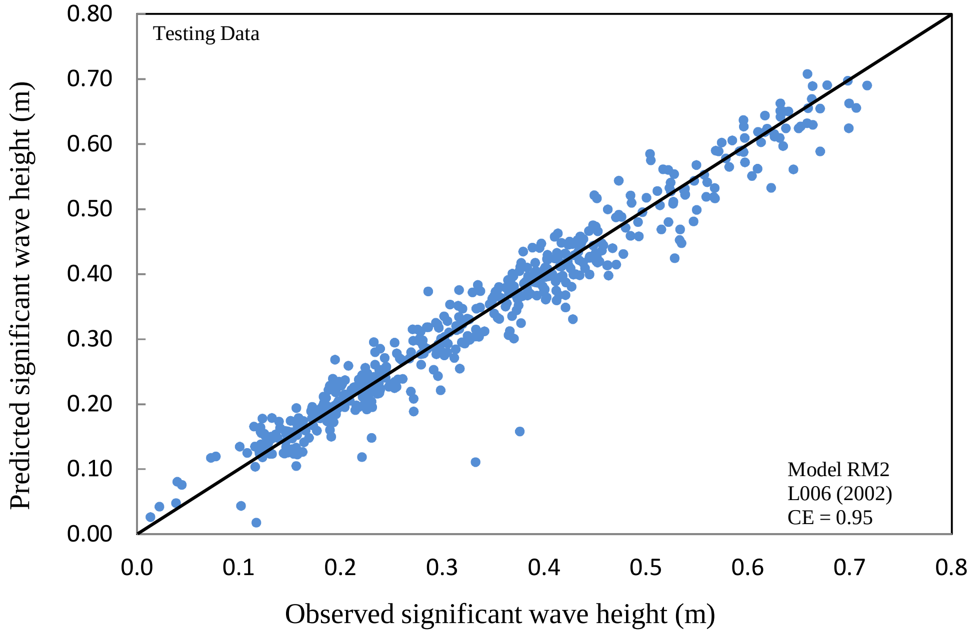

Figure 4.11 Verification of predicted and observed significant wave height using RM2 (500 testing data applied L006 – 2002)

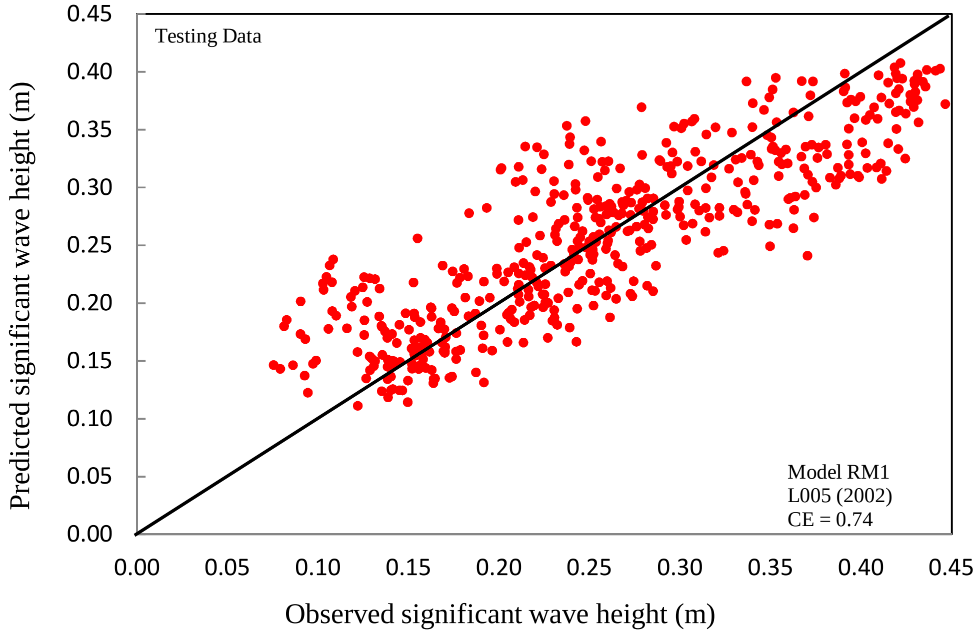

Figure 4.12 Verification of predicted and observed significant wave height using RM1 (500 testing data applied L005 – 2002)

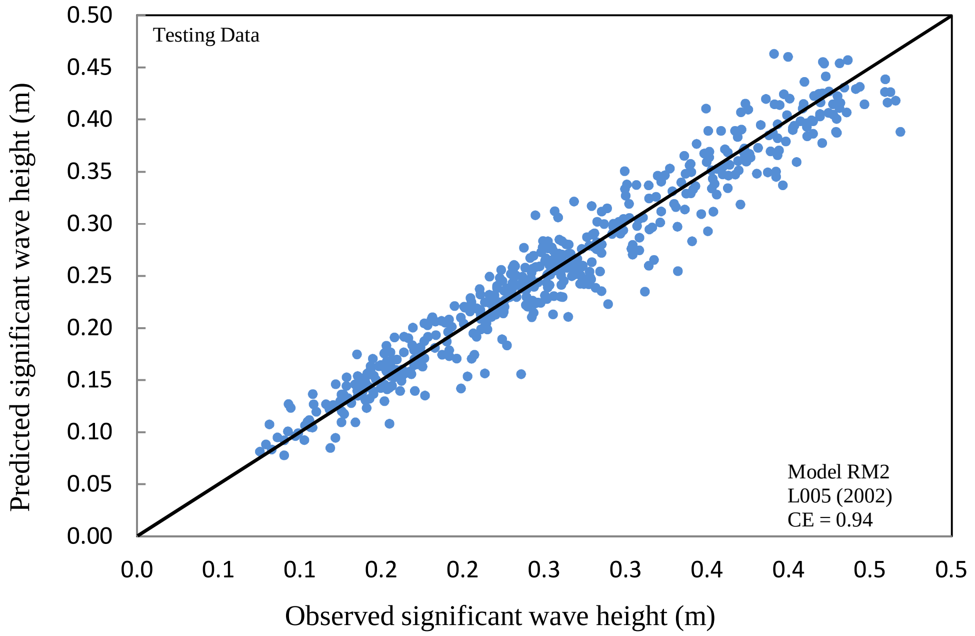

Figure 4.13 Verification of predicted and observed significant wave height using RM2 (500 testing data applied L005 – 2002)

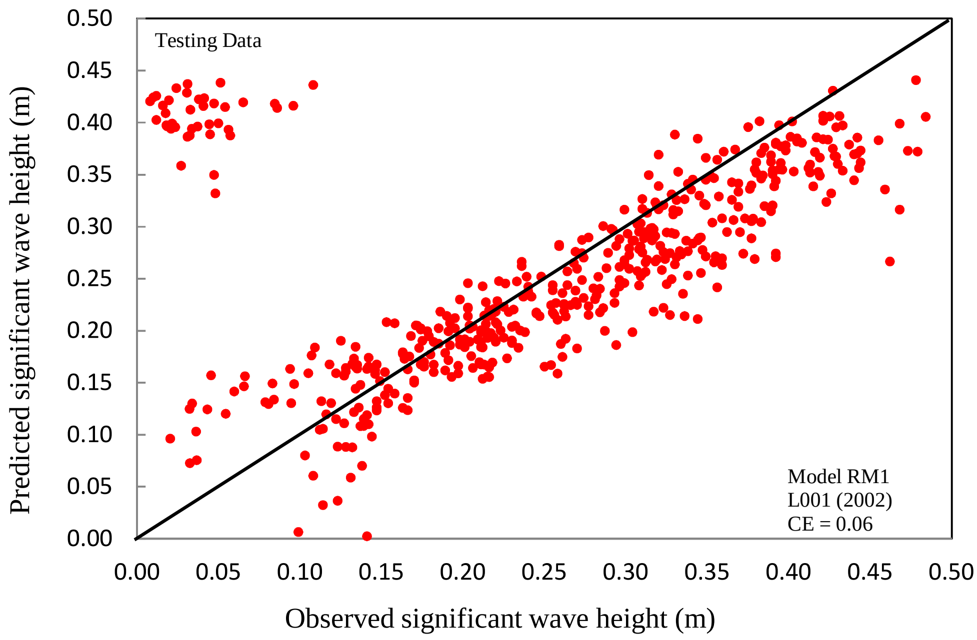

Figure 4.14 Verification of predicted and observed significant wave height using RM1 (500 testing data applied L001 – 2002)

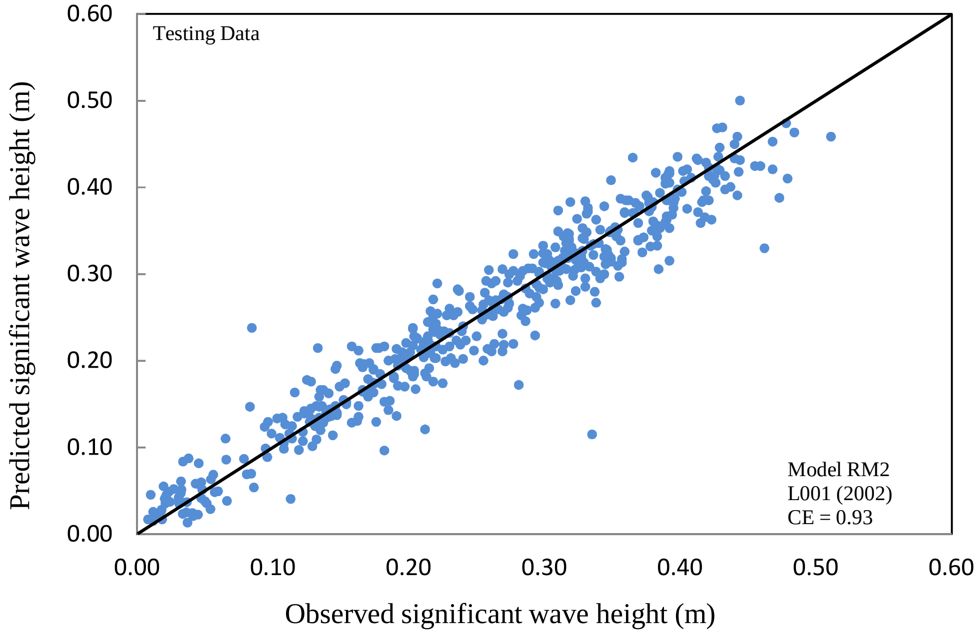

Figure 4.15 Verification of predicted and observed significant wave height using RM2 (500 testing data applied L001 – 2002)

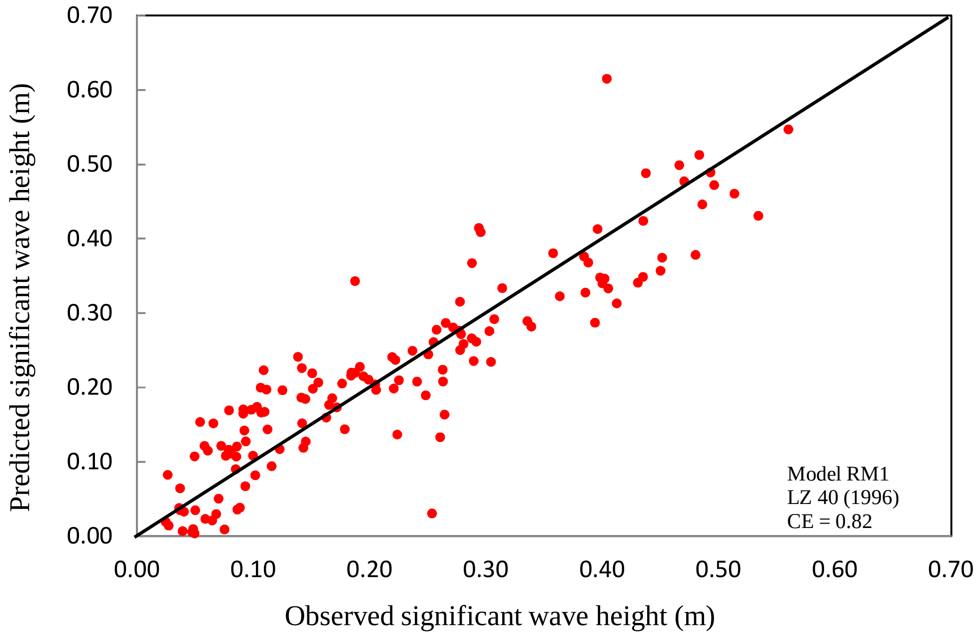

Figure 4.16 Verification of predicted and observed significant wave height using RM1 (140 data applied LZ40 – 1996)

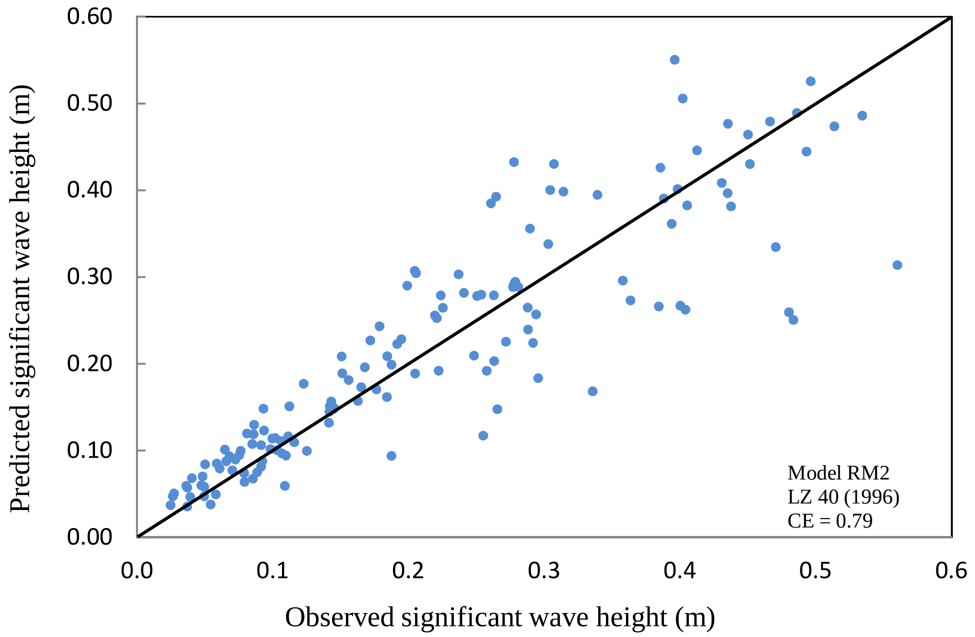

Figure 4.17 Verification of predicted and observed significant wave height using RM2 (140 data applied LZ40 – 1996)

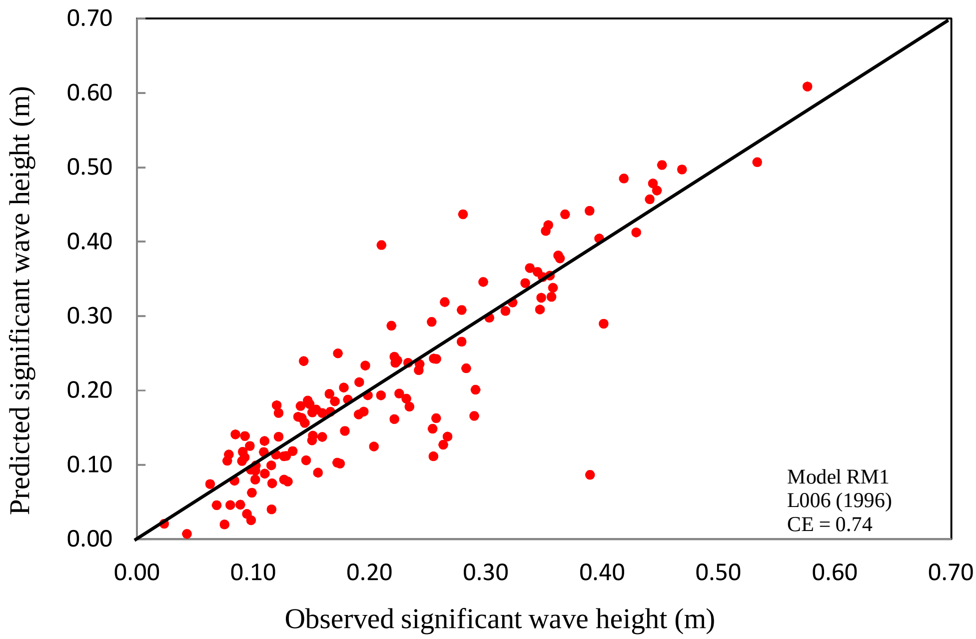

Figure 4.18 Verification of predicted and observed significant wave height using RM1 (140 data applied L006 – 1996)

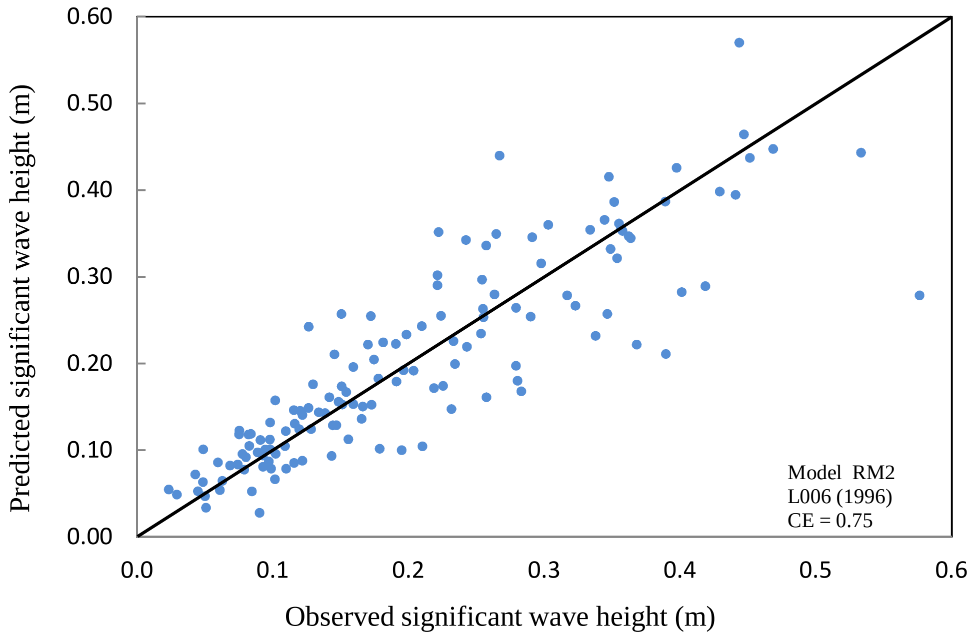

Figure 4.19 Verification of predicted and observed significant wave height using RM2 (140 data applied L006 – 1996)

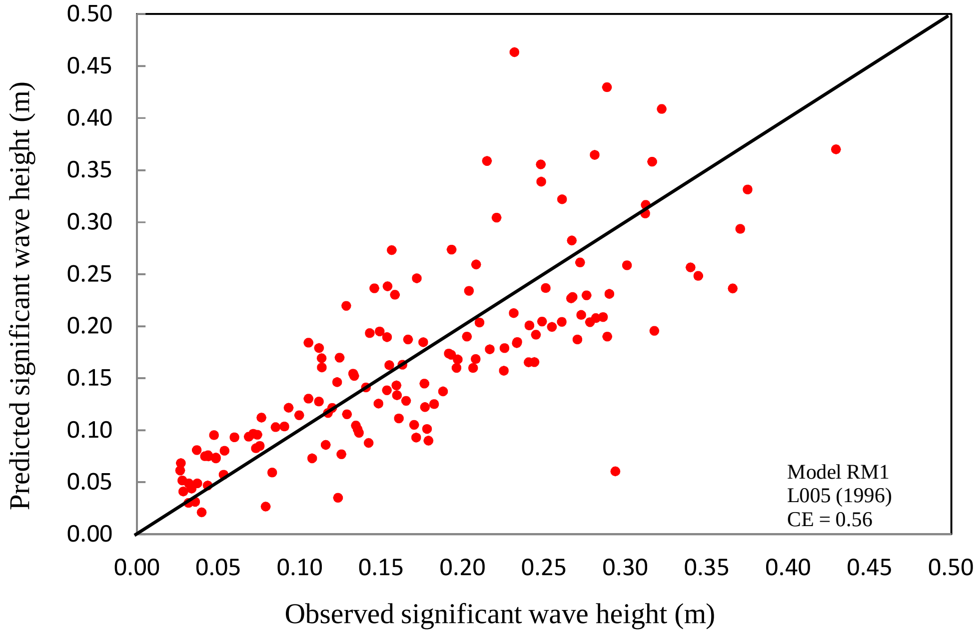

Figure 4.20 Verification of predicted and observed significant wave height using RM1 (140 data applied L005 – 1996)

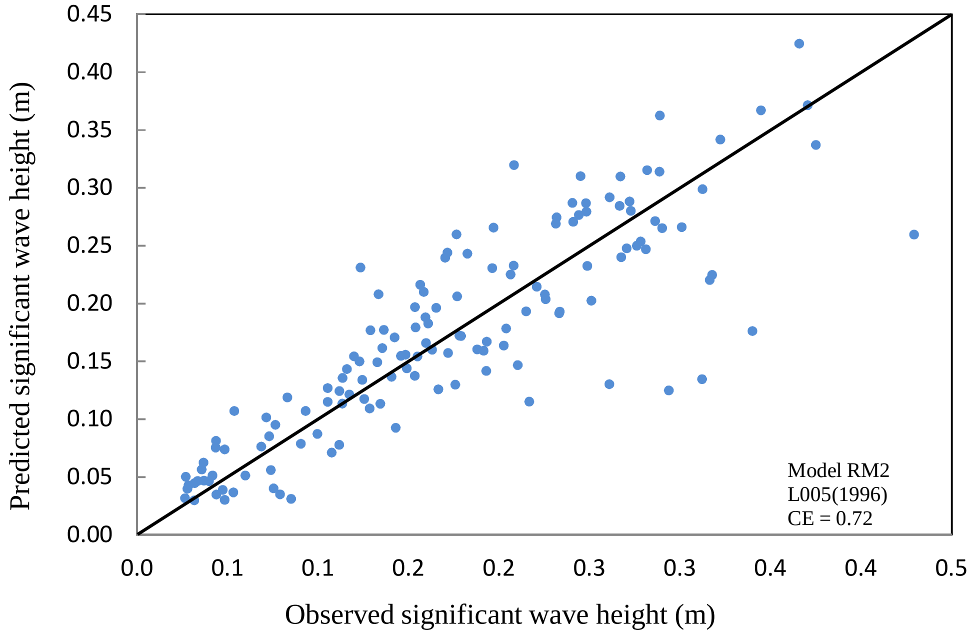

Figure 4.21 Verification of predicted and observed significant wave height using RM2 (140 data applied L005 – 1996)

Table 4.1 Summary of AAE, RMSE and CE of RM1 and RM2 using 500 testing data collected in year 2002 for the stations LZ40, L006, L005 and L001

500 testing data year 2002	RM1			RM2		
	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE
LZ40	0.04	0.06	0.87	0.03	0.04	0.94
L006	0.04	0.05	0.88	0.02	0.03	0.95
L005	0.04	0.05	0.74	0.02	0.02	0.94
L001	0.06	0.11	0.06	0.02	0.03	0.93

Table 4.2 Summary of AAE, RMSE and CE of RM1 and RM2 using 140 independent data testing data collected in year 1996 for the stations LZ40, L006 and L005

140 data year 1996	RM1			RM2		
	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE
LZ40	0.05	0.06	0.82	0.04	0.07	0.79
L006	0.04	0.06	0.74	0.04	0.06	0.75
L005	0.05	0.06	0.56	0.03	0.05	0.72

4.3 Perceptron Least Square Method (PLSM)

The validation of PLSM model are also performed using the same testing data as described in previous section, which include year 2002 and year 1996 data. Figures 4.22 to 4.25 present the time variations of predicted and observed significant wave heights for the year 2002 testing data at LZ40, L006, L005 and L001, respectively. There are 500 data points. Predicted significant wave heights at all four stations are well fitted to the observed ones. For the time series comparisons, the predictions of significant wave height for the year 1996 events are mostly accurate but with one hour shifting comparing to the recorded data for stations LZ40, L006 and L00 as shown respectively in Figures 4.30, 4.31 and 4.32.

With the reference of the perfect 45 degree fit line, the direct comparisons between predicted and observed significant wave heights for the case using the year 2002 testing data are presented in Figures 4.26 to 4.29. Similar plots for showing the performance of the PLSM model are also displayed in Figures 4.33 to 4.35 for the year 1996 data. Predicted data by PLSM are closely fitted to the 45 degree line for the validation data of 500 testing points at all stations (Figures 4.26 to 4.29). Figures 4.33 to 4.35 show more scatter data away from the 45 degree line because of the time shifting of predicted data for the year 1996. The PLSM model is proved to have a better applicability on the prediction of significant wave height in Lake Okeechobee. The results also suggest that with the inclusions of both inputs of current wind speed and previously occurred significant wave height the predictive model can improve the prediction of current significant wave height.

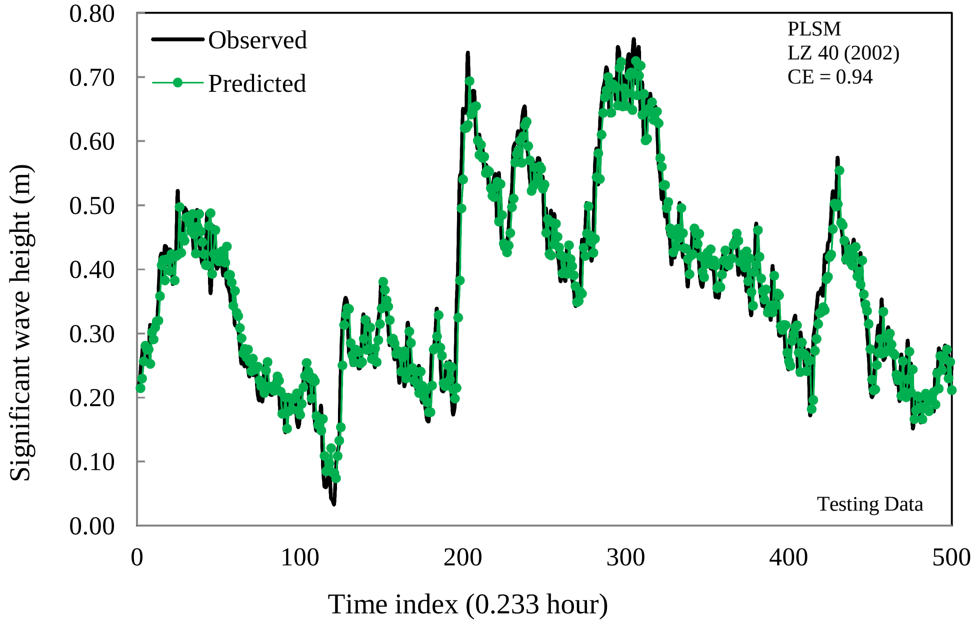

Figure 4.22 Time variations of predicted and observed significant wave height using PLSM (500 tested data applied to LZ40 – 2002)

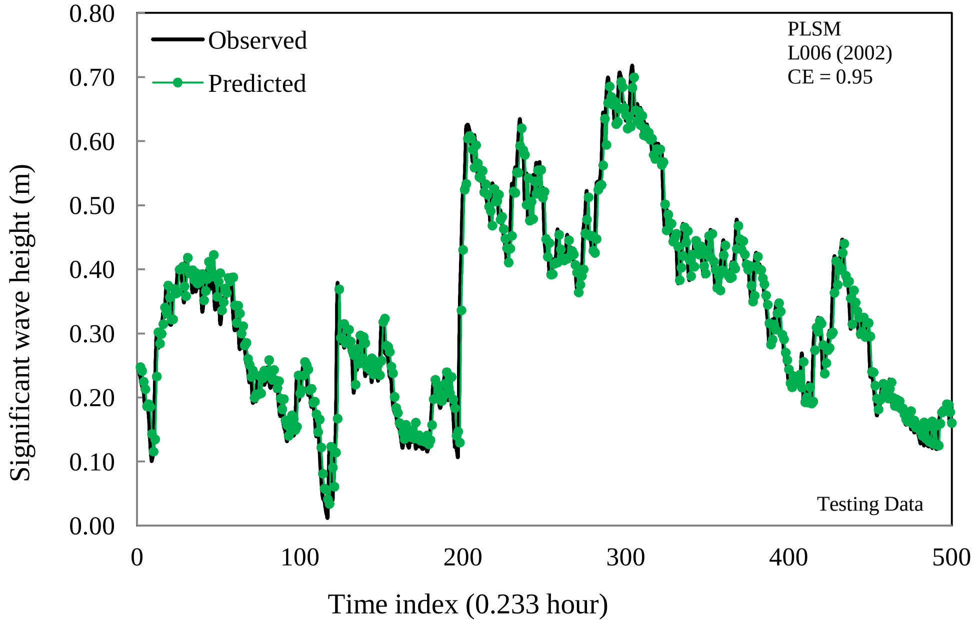

Figure 4.23 Time variations of predicted and observed significant wave height using PLSM (500 tested data applied to L006 – 2002)

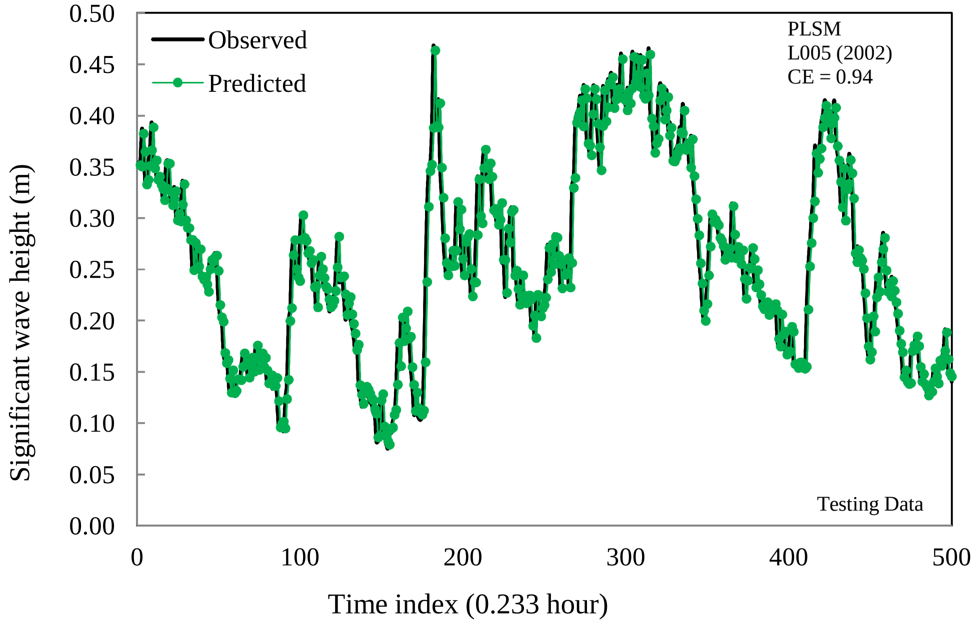

Figure 4.24 Time variations of predicted and observed significant wave height using PLSM (500 tested data applied to L005 – 2002)

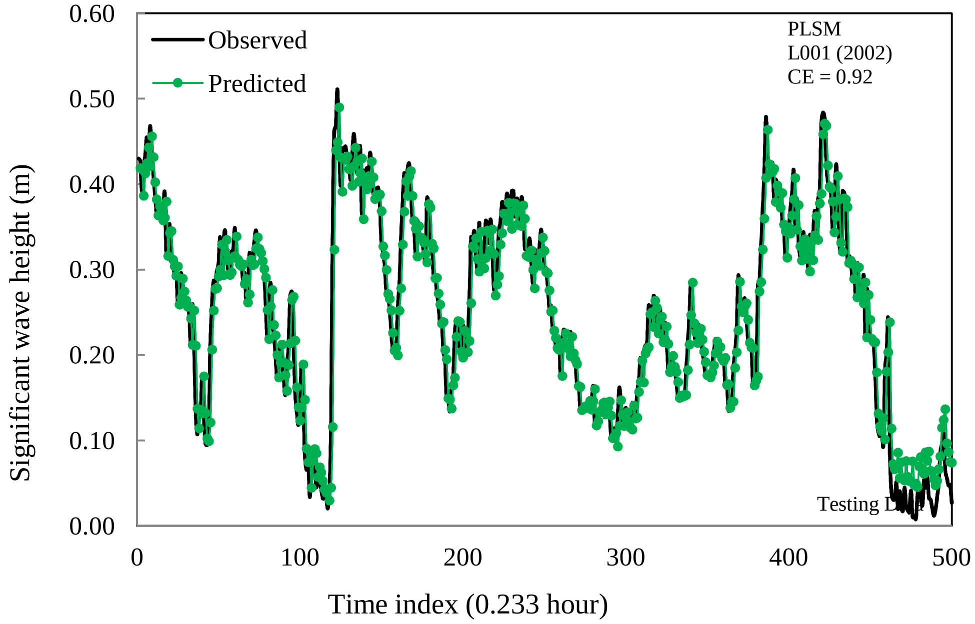

Figure 4.25 Time variations of predicted and observed significant wave height using PLSM (500 tested data applied to L001 – 2002)

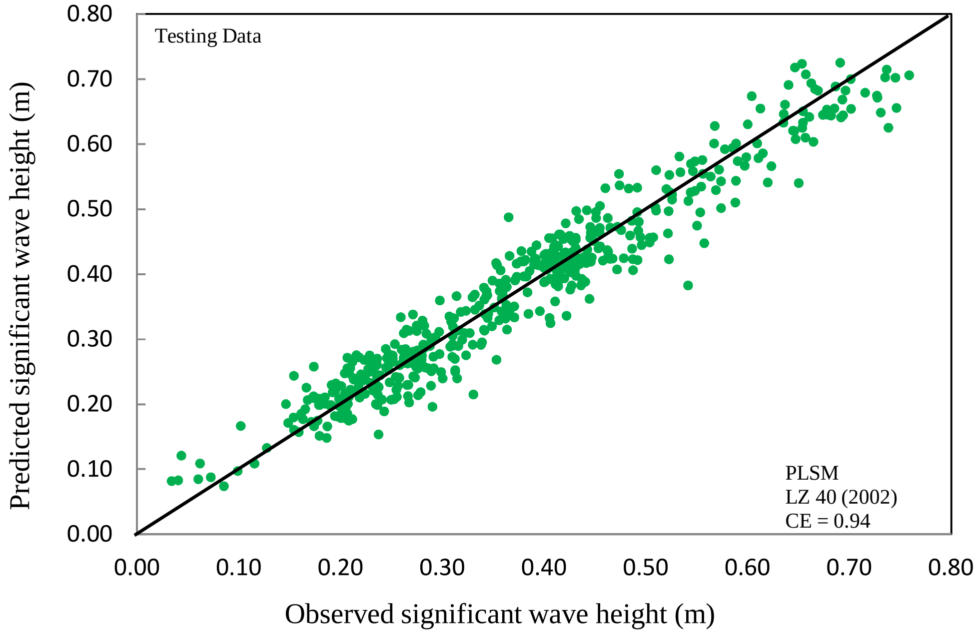

Figure 4.26 Verification of predicted and observed significant wave height using PLSM (500 testing data applied LZ40 – 2002)

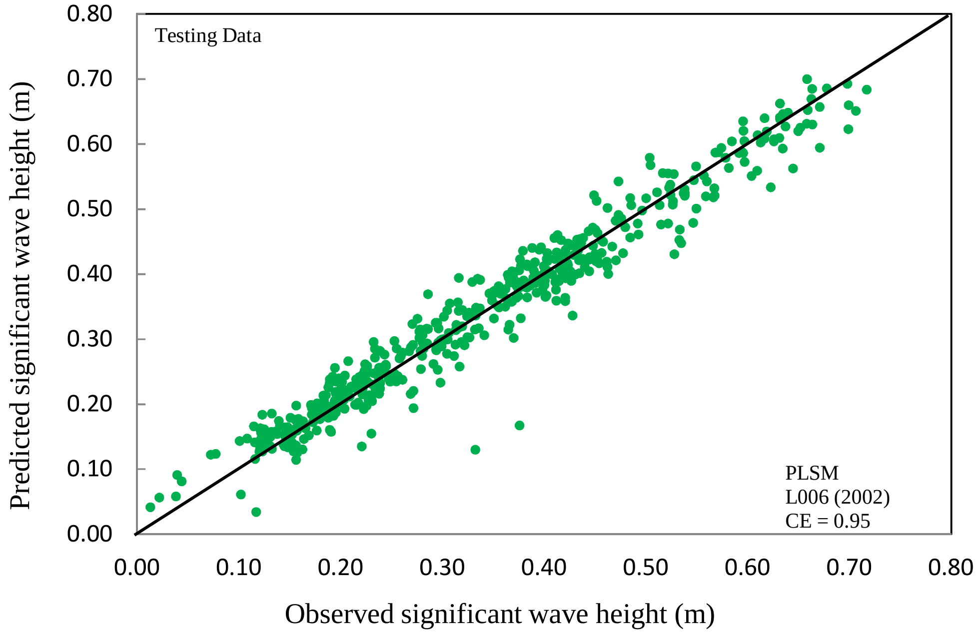

Figure 4.27 Verification of predicted and observed significant wave height using PLSM (500 testing data applied L006 – 2002)

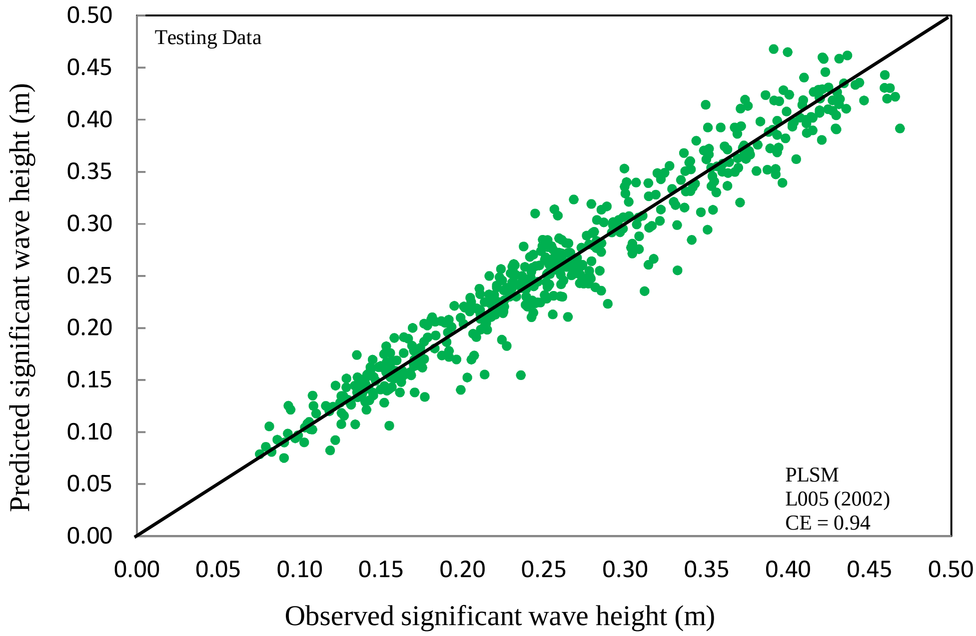

Figure 4.28 Verification of predicted and observed significant wave height using PLSM (500 testing data applied L005 – 2002)

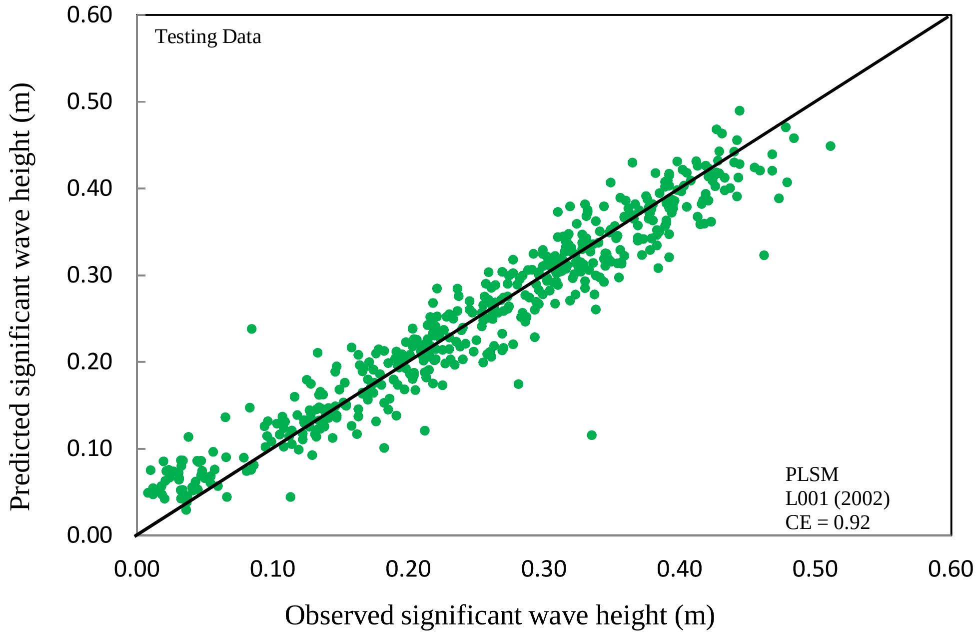

Figure 4.29 Verification of predicted and observed significant wave height using PLSM (500 testing data applied L001 – 2002)

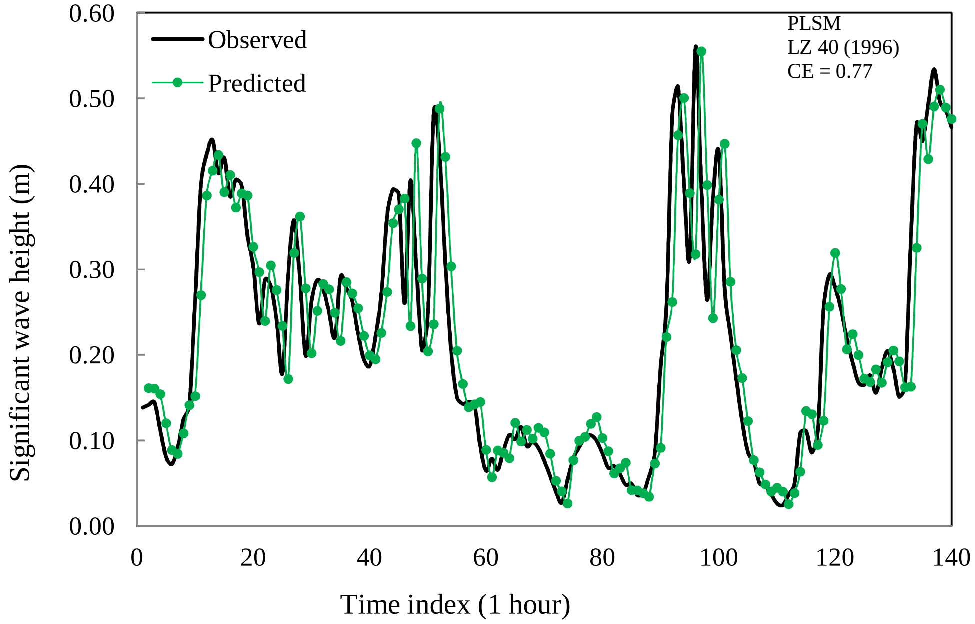

Figure 4.30 Time variations of predicted and observed significant wave height using PLSM (140 data applied to LZ40 – 1996)

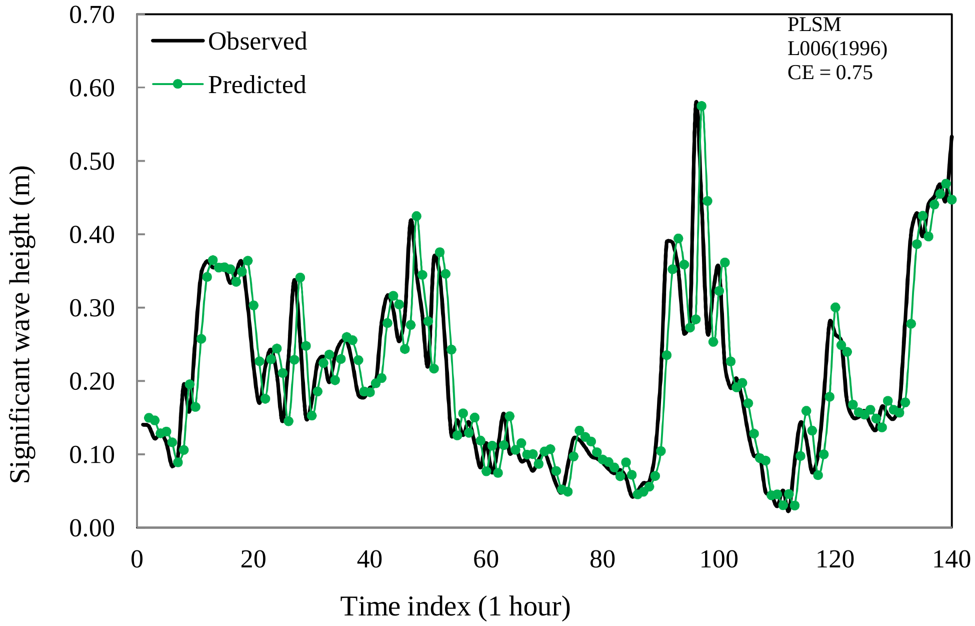

Figure 4.31 Time variations of predicted and observed significant wave height using PLSM (140 data applied to L006 – 1996)

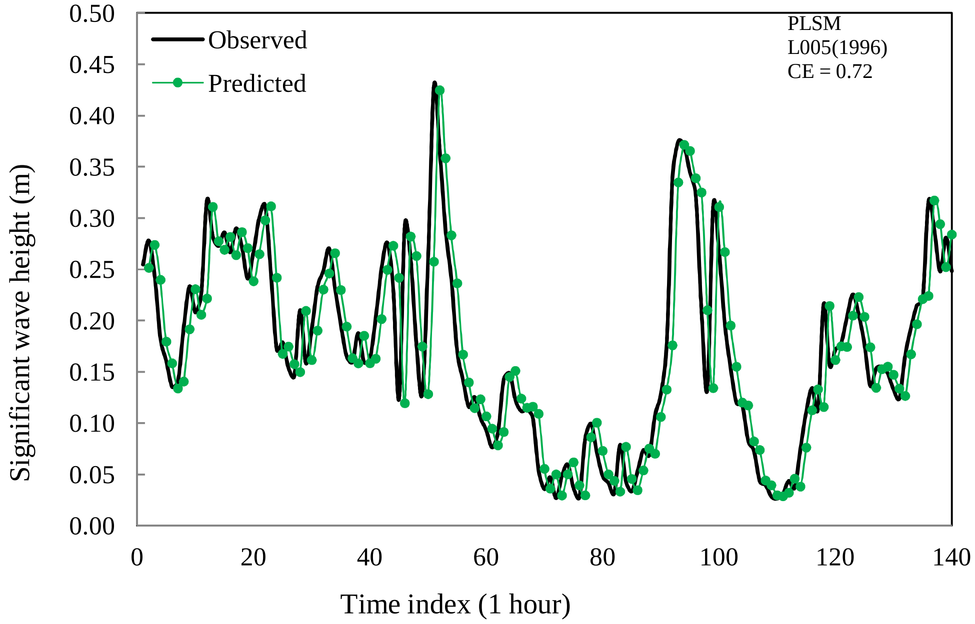

Figure 4.32 Time variations of predicted and observed significant wave height using PLSM (140 data applied to L005 – 1996)

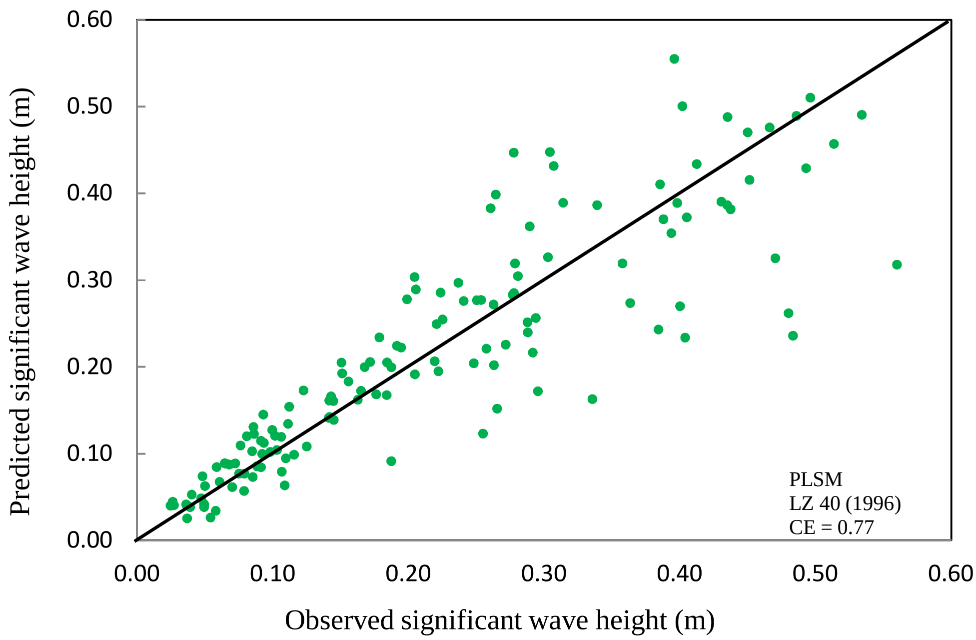

Figure 4.33 Verification of predicted and observed significant wave height using PLSM (140 data applied LZ40 – 1996)

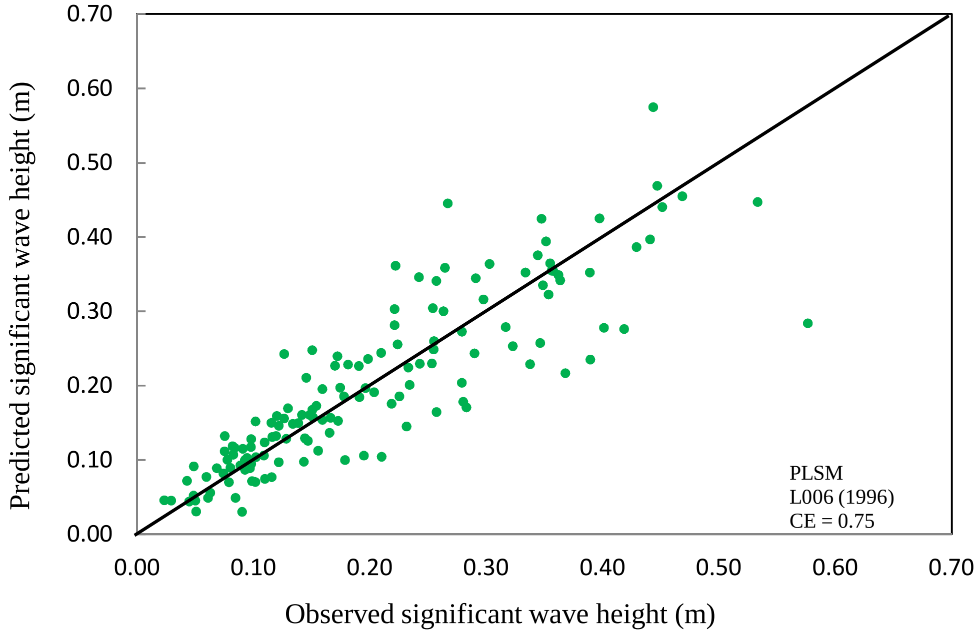

Figure 4.34 Verification of predicted and observed significant wave height using PLSM (140 data applied L006 – 1996)

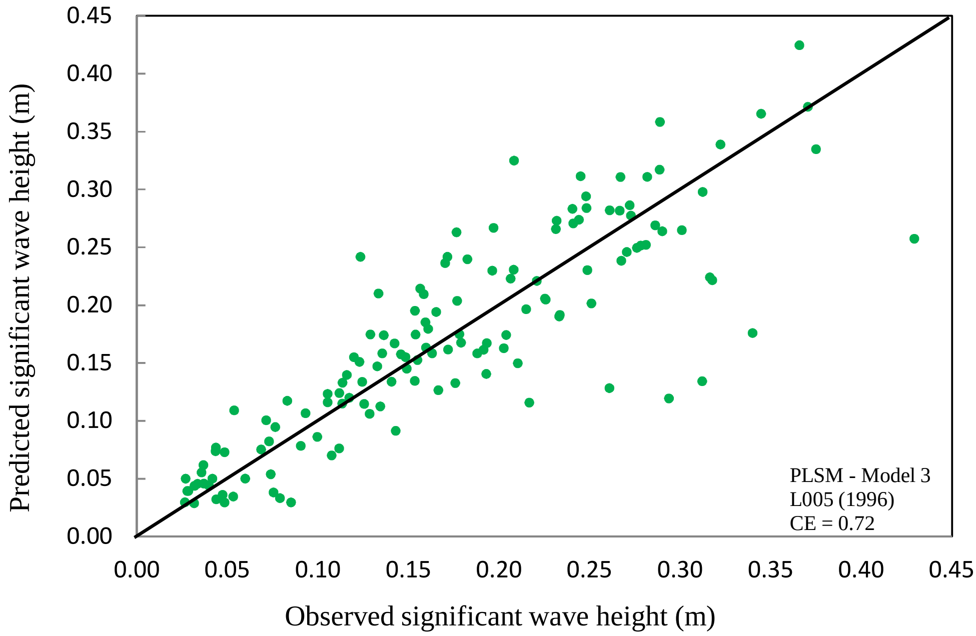

Figure 4.35 Verification of predicted and observed significant wave height using PLSM (140 data applied L005 – 1996)

Tables 4.3 and 4.4 show the model performance related statistical values of AAE, RMSE, and CE for testing the 500 data of the year 2002 and other independent data of the year 1996, respectively. Results from Tables 4.3 and 4.4 are also consistent with the comparison plots shown above. Errors produced by PLSM are less than those from RM1 model for both sets of data. Similar to RM2 model, PLSM produced CE values more than 90% for the 500 testing data and more than 70% for the year 1996 data.

Table 4.3 Values of AAE, RMSE and CE of PLSM using 500 testing data of year 2002 for the stations LZ40, L006, L005 and L001

	PLSM		
	AAE (m)	RMSE (m)	CE
LZ40	0.03	0.04	0.94
L006	0.02	0.03	0.95
L005	0.02	0.02	0.94
L001	0.02	0.03	0.92

Table 4.4 Values of AAE, RMSE and CE of PLSM using 140 data (year 1996) for the stations LZ40, L006 and L005

	PLSM		
	AAE (m)	RMSE (m)	CE
LZ40	0.04	0.07	0.77
L006	0.04	0.06	0.75
L005	0.03	0.05	0.72

4.4 Modified Pierson-Moskowitz (MPM) Spectrum Model

Modified PM Spectrum is proved in the model training (calibration) study (Section 3.5) to be able to predict reasonable results for all stations (Figures 3.38 - 4.41). It would be interesting to test the MPM model by applying other independent data. Following the similar procedures as described for the RM1, RM2, and PLSM models, the first validation test of the MPM model was performed using the 500 data points of the year 2002 (which are not used for training). The time variation plots of predicted and observed significant wave height using MPM for stations LZ40, L006, L005, and L001 are presented respectively in Figures 4.36, 4.37, 4.38, and 4.39. The predictions from MPM model are found to closely follow the variation trend of the observed data for the station LZ40 (Figure 4.36). The values of predicted significant wave height also fit reasonably well with measured data (Figures 4.36 and 4.40). For stations L006, L005 and L001, the MPM model still is able to predict reasonable variation trend as that appeared in the recorded data (Figures 4.37 to 4.39), however, the values are generally under estimated. These conclusions can also be seen in the comparison plots illustrated in Figures 4.41, 4.42 and 4.43 where most data points are situated below the perfect 45 degree fit line.

Prediction of significant wave height by the MPM model is solely dependent on the wind speed. During a time period near the end of the data collection at station L001 (marked with circle), it is noted that in Figure 4.39 the values of predicted significant wave heights are much greater than those from observations. As a result of recorded large values of wind speed (Figure 3.4), using the wind speed as the input data, the MPM model failed to overcome the inconsistent wind speed data to produce reasonable predictions. It is possible, at the time period marked in Figure 4.39, the data collection processes for either the wind speed or waves experienced large recording errors.

The second test was performed using the 1996 data. They are hourly data covering the time period of 140 hours. As shown in Figures 4.44 and 4.47, the MPM model again predicts reasonable results for station LZ40. The predicted significant wave heights generally follow the variation trend of observed data and distribute around the perfect fit line. Acceptable results (Figures 4.45 and 4.48) with less large oscillations in the time series plot for station L006 can also be noticed. For station L005, the results (Figures 4.46 and 4.49) are mostly underestimated when compared to the observations.

As summarized in Table 4.5, the MPM model produces coefficient of efficiency of 0.85, 0.8, 0.63 and 0.04 for LZ40, L006, L005 and L001 respectively for the 500 test data (year 2002). For the year 1996 data, values coefficient of efficiency are 0.8, 0.74, and 0.45 for LZ40, L006 and L005 respectively (Table 4.6).

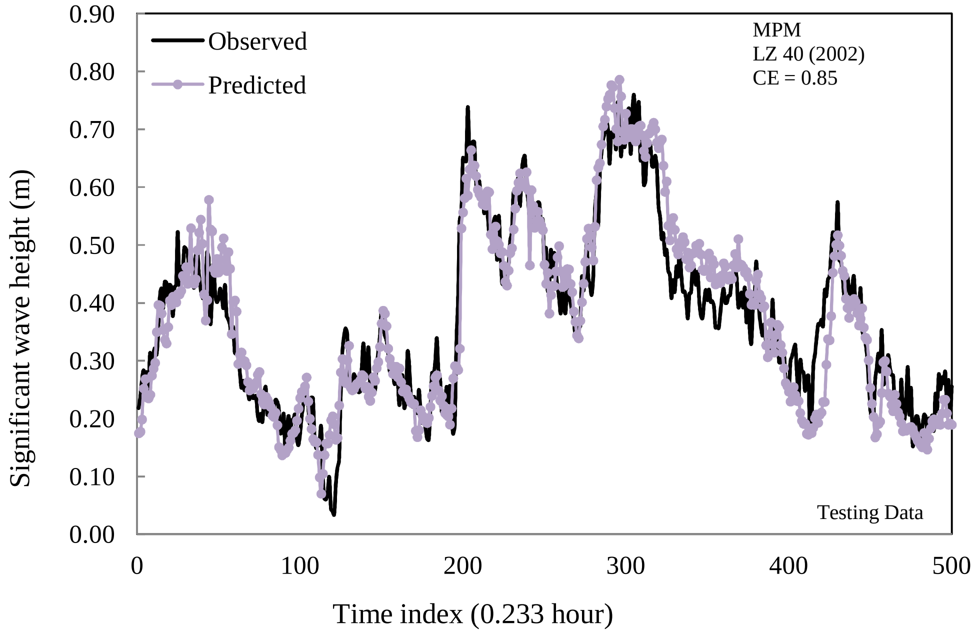

Figure 4.36 Time variations of predicted and observed significant wave height using MPM (500 testing data applied to LZ40 – 2002)

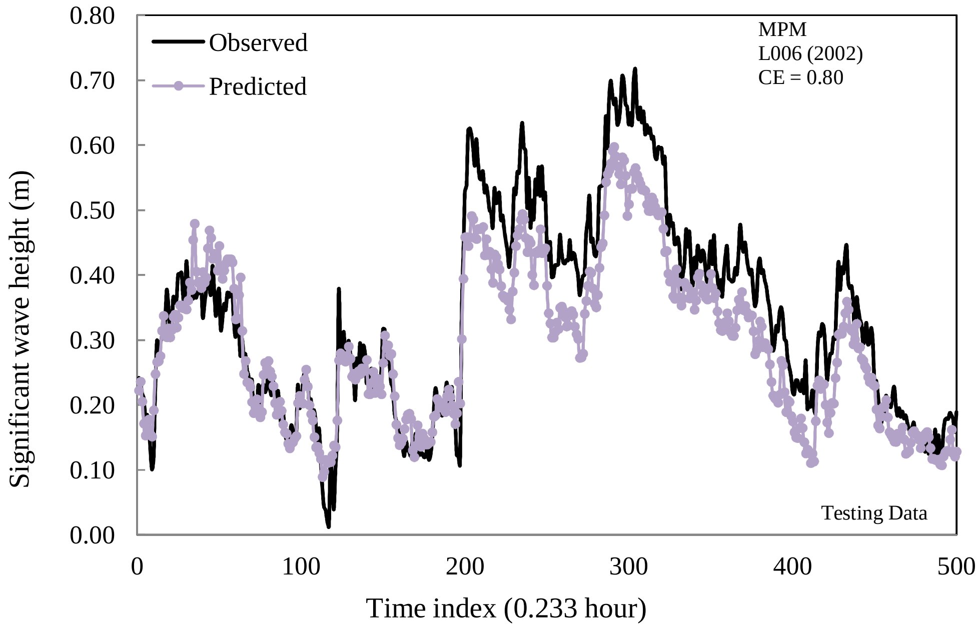

Figure 4.37 Time variations of predicted and observed significant wave height using MPM (500 testing data applied to L006 – 2002)

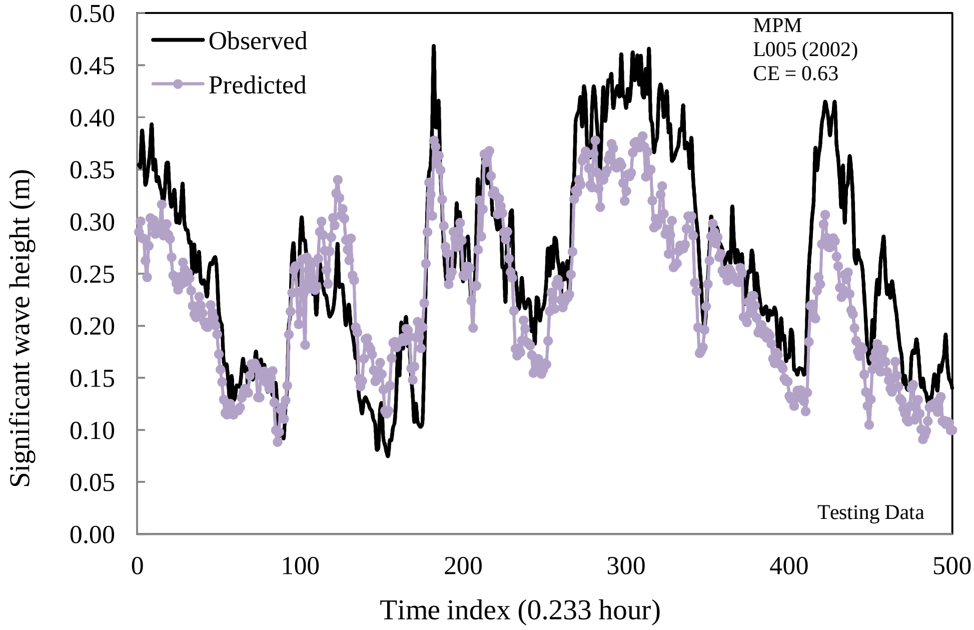

Figure 4.38 Time variations of predicted and observed significant wave height using MPM (500 testing data applied to L005 – 2002)

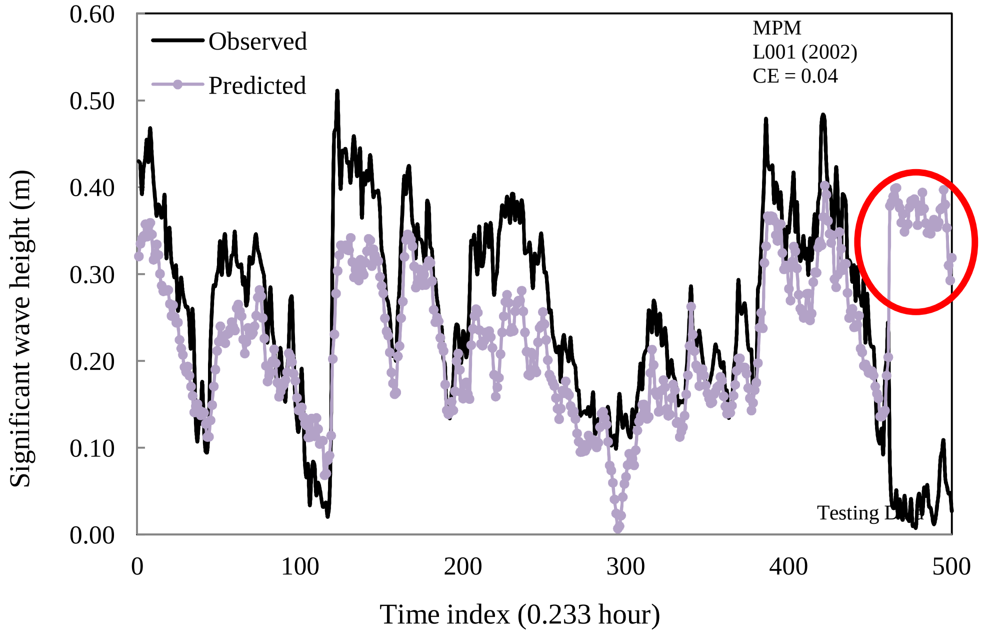

Figure 4.39 Time variations of predicted and observed significant wave height using MPM (500 testing data applied to L001 – 2002)

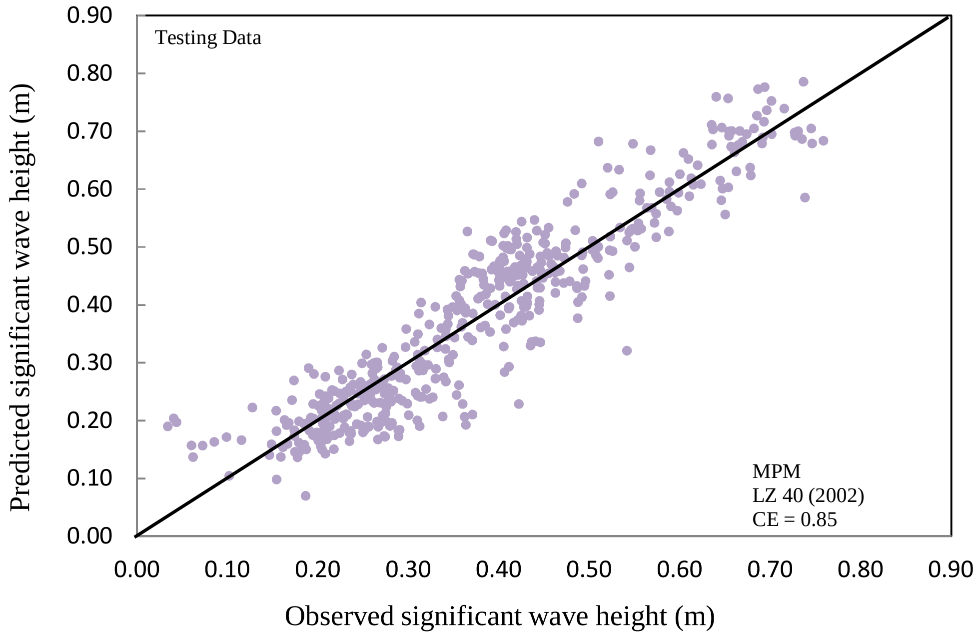

Figure 4.40 Verification of predicted and observed significant wave height using MPM (500 testing data applied to LZ40 – 2002)

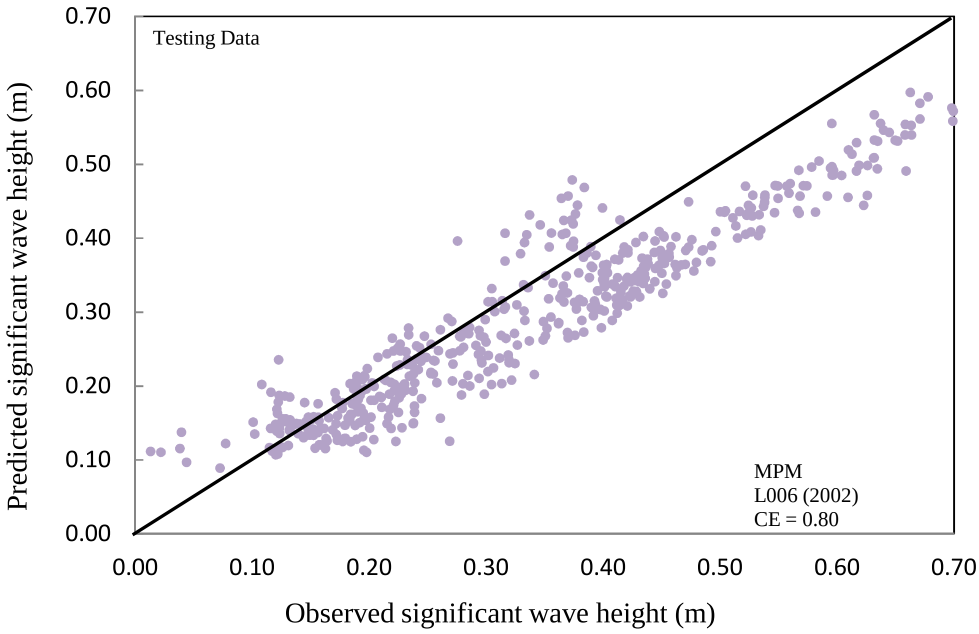

Figure 4.41 Verification of predicted and observed significant wave height using MPM (500 testing data applied to L006 – 2002)

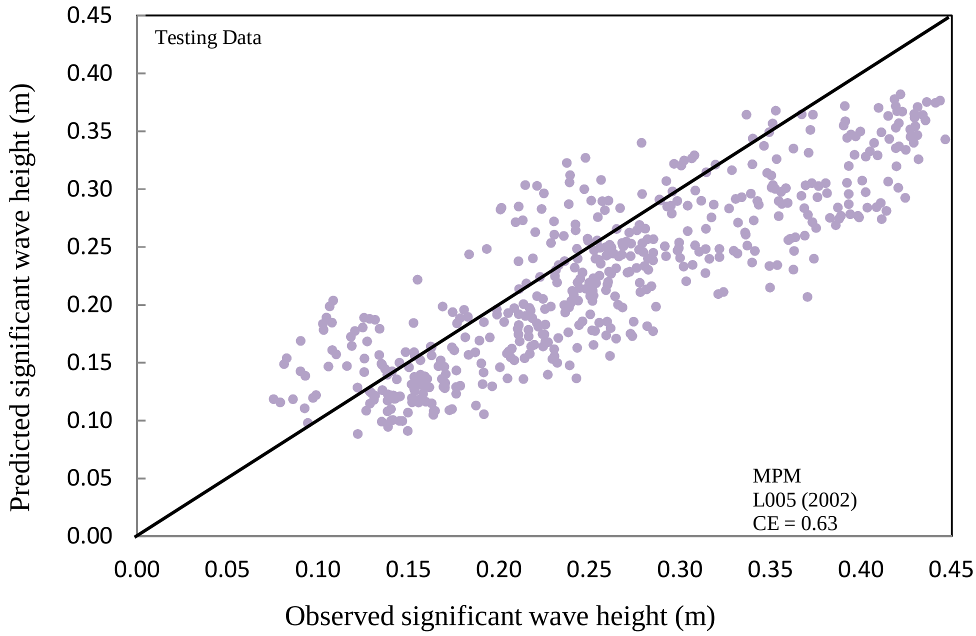

Figure 4.42 Verification of predicted and observed significant wave height using MPM (500 testing data applied to L005 – 2002)

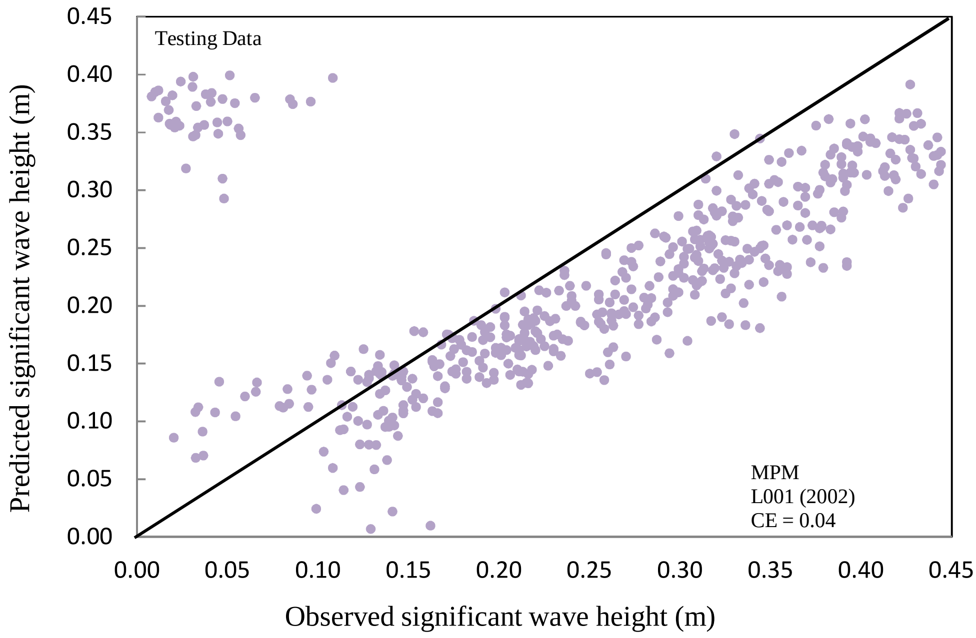

Figure 4.43 Verification of predicted and observed significant wave height using MPM (500 testing data applied to L001 – 2002)

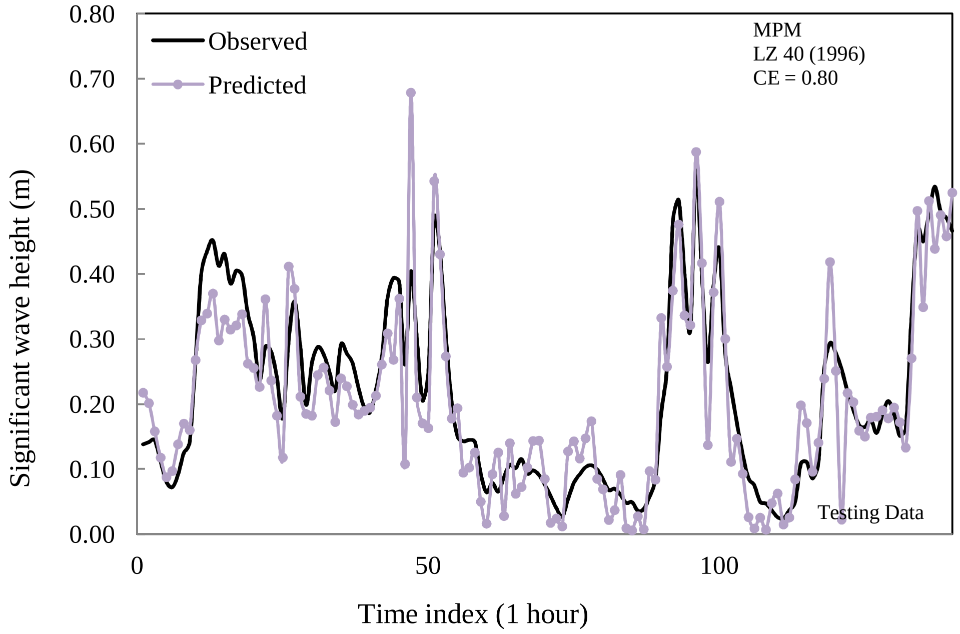

Figure 4.44 Time variations of predicted and observed significant wave height using MPM (140 data applied to LZ40 – 1996)

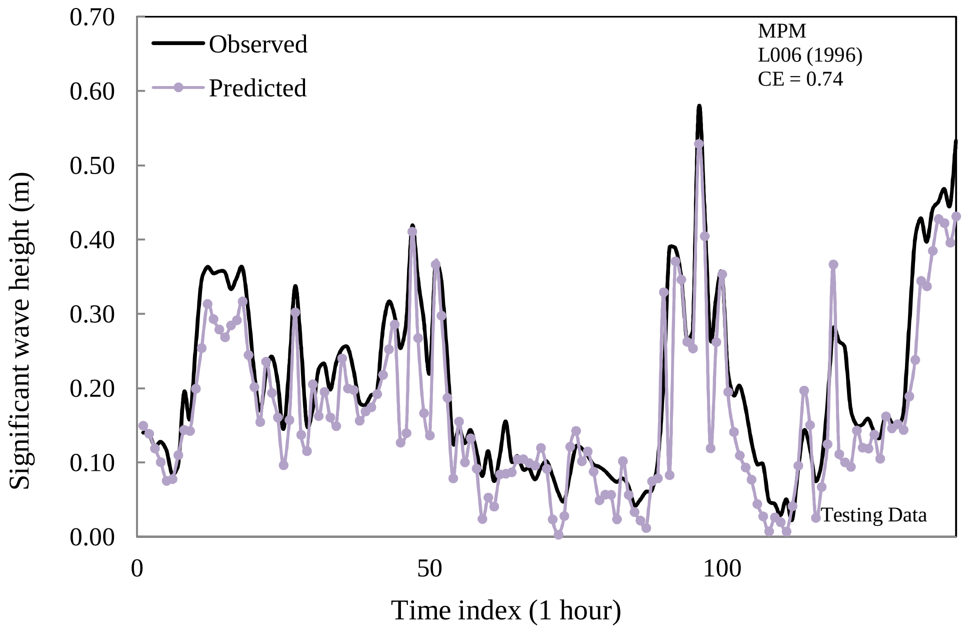

Figure 4.45 Time variations of predicted and observed significant wave height using MPM (140 data applied to L006 – 1996)

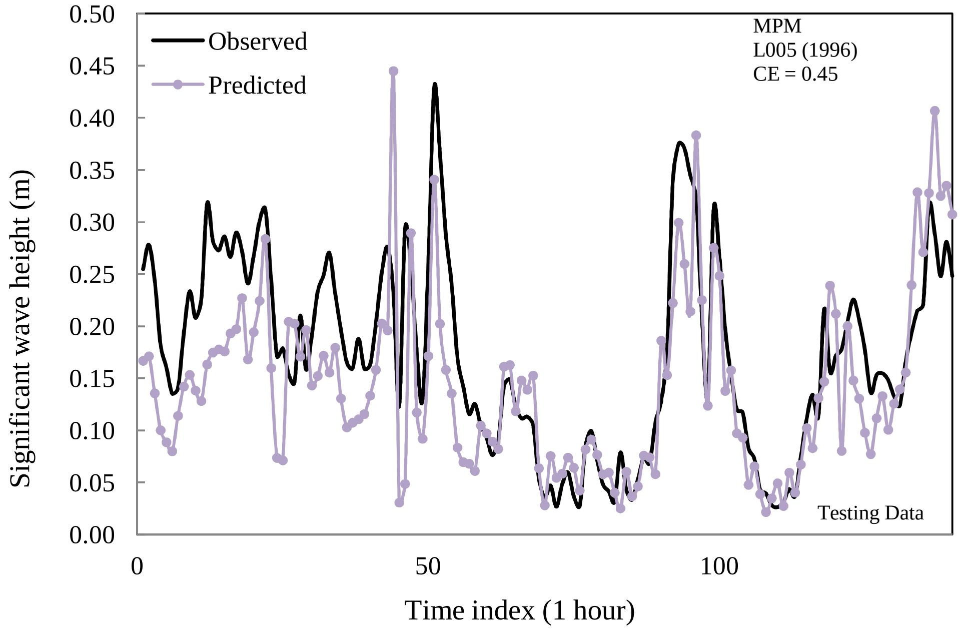

Figure 4.46 Time variations of predicted and observed significant wave height using MPM (140 data applied to L005 – 1996)

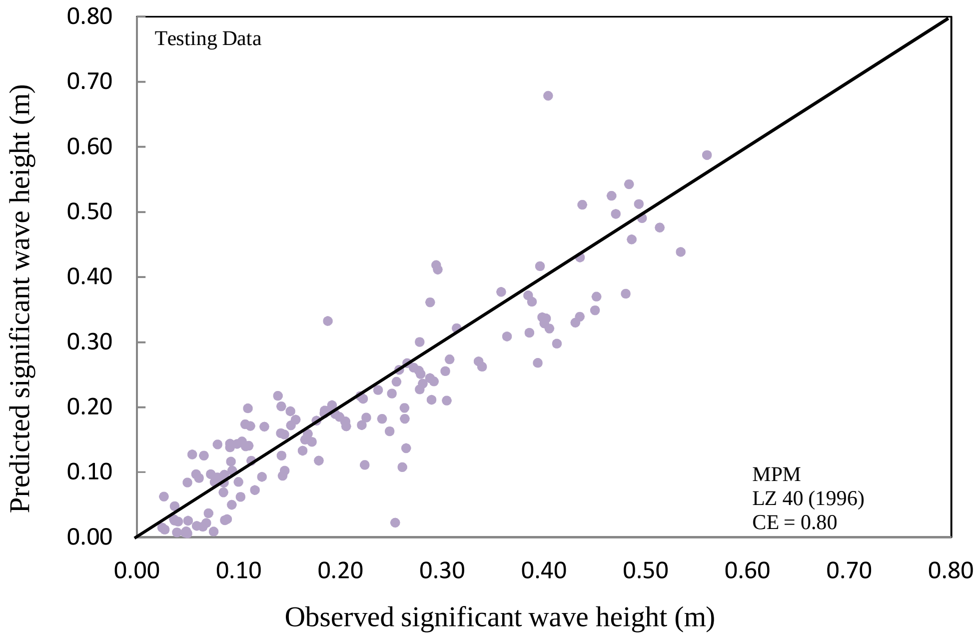

Figure 4.47 Verification of predicted and observed significant wave height using MPM (140 data applied to LZ40 – 1996)

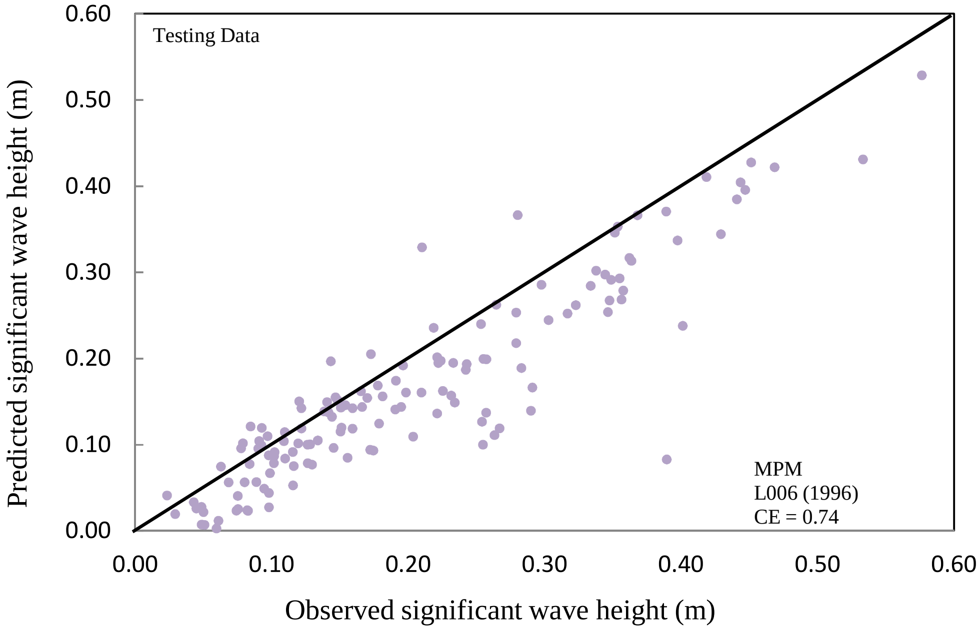

Figure 4.48 Verification of predicted and observed significant wave height using MPM (140 data applied to L006 – 1996)

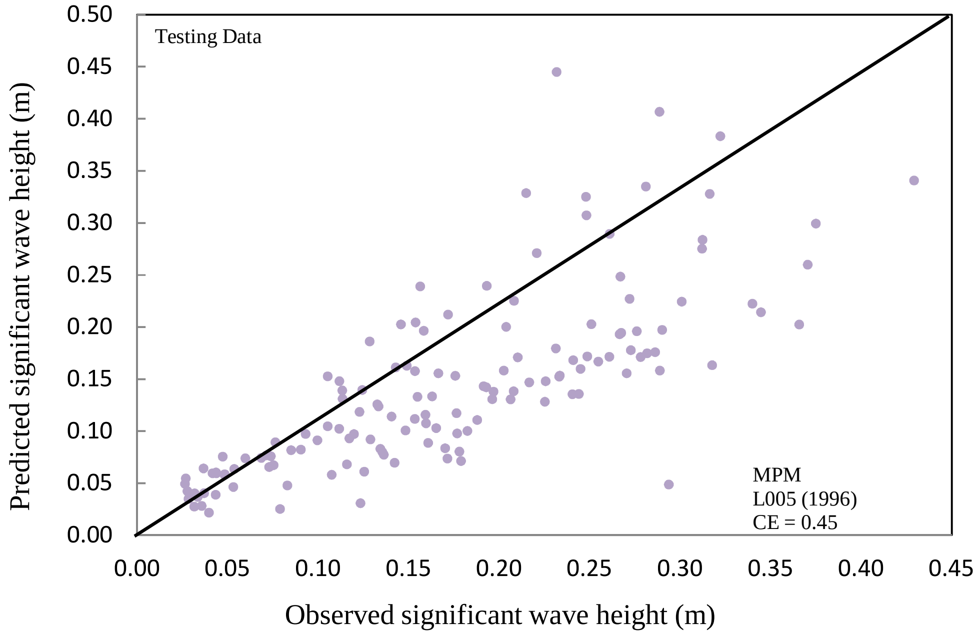

Figure 4.49 Verification of predicted and observed significant wave height using MPM (140 data applied to L005 – 1996)

Table 4.5 Values of AAE, RMSE and CE of MPM using the 500 testing data of year 2002 for the stations LZ40, L006, L005 and L001

	MPM		
	AAE (m)	RMSE (m)	CE
LZ40	0.05	0.06	0.85
L006	0.06	0.07	0.80
L005	0.05	0.06	0.63
L001	0.08	0.11	0.04

Table 4.6 Values of AAE, RMSE and CE of MPM using 140 data of year 1996 for the stations LZ40, L006 and L005

	MPM		
	AAE (m)	RMSE (m)	CE
LZ40	0.05	0.06	0.80
L006	0.05	0.06	0.74
L005	0.05	0.07	0.45

4.5 Comparisons with ANN Model

Four models for the prediction of significant wave height in Lake Okeechobee were developed and tested in this study. These are RM1 and RM2 models (Regression Method), Perceptron Least Square Method (PLSM) model and Modified Pierson Moskowitz Spectrum (MPM) model. All four models have their own cons and pros, but generally each model has been demonstrated to be able to provide reasonable predictions, especially for stations LZ40 and L006, near the center of the Lake Okeechobee. Among the four models, it is noticeable that RM2 and PLSM model are able to produce better estimations at four test stations. To get a better understanding of the performance of the proposed models, comparisons of time variations of predicted and observed significant wave heights for all four models at LZ40, L006, L005, and L001 stations are created for the 500 testing data of year 2002 (Figures 4.50, 4.52, 4.54 and 4.56) and 140 data of year 1996 (Figures 4.51, 4.53 and 4.55).

As the modern technology approaches, e.g. Artificial Neural Network (ANN), become popular in recent years for the prediction of time variation variables, it would be interesting to also compare the present model predictions with ANN results. In a study carried out by Altunkaynak and Wang (2012), an ANN method was applied to Lake Okeechobee for predicting significant wave height. The ANN results from Altunkaynak and Wang (2012) are also included in Figures 4.50, 4.52, 4.54 and 4.56 (the 500 testing data of year 2002) and in Figures 4.51, 4.53 and 4.55 (1996 data) for comparisons with the predictions from the four models presented in this study. A summary of AAE, RMSE and CE for both set of data is provided in Tables 4.7 and 4.8.

4.5.1 Comparisons of RM1, RM2, PLSM, MPM and ANN at LZ40

Figures 4.50 and 4.51 are shown respectively for the comparisons of the model results from RM1, RM2, PLSM, MPM and ANN for the year 2002 and year 1996 data. Significant wave heights predicted by RM2 and PLSM follow the actual trend of observed data. MPM and ANN follow almost same trend while predicting significant wave height. At two points (Figure 4.50 marked with arrow) ANN is under predicted while MPM is over predicted. The results obtained by RM1 are acceptable comparing to MPM and ANN.

For the year 1996 data (Figure 4.51), both RM2 and PLSM are also able to predict more accurate results. Rest of the models, RM1, MPM and ANN, although produceing reasonable results, are fluctuated more significantly at several occasions (marked with arrow) during the high wind event occurred that year.

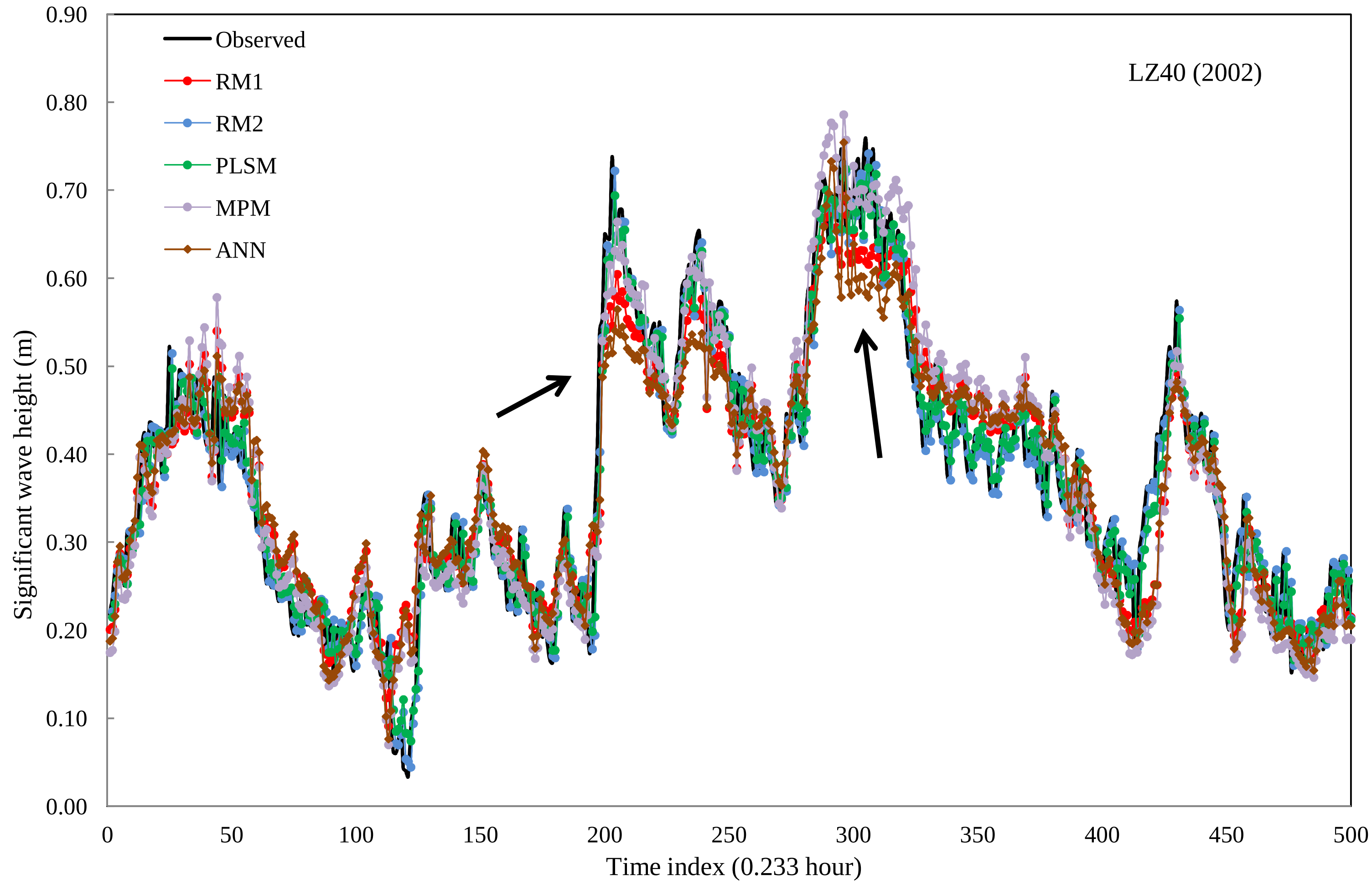

Figure 4.50 Comparison chart of time variations for RM1, RM2, PLSM, MPM and ANN (500 testing data of year 2002)

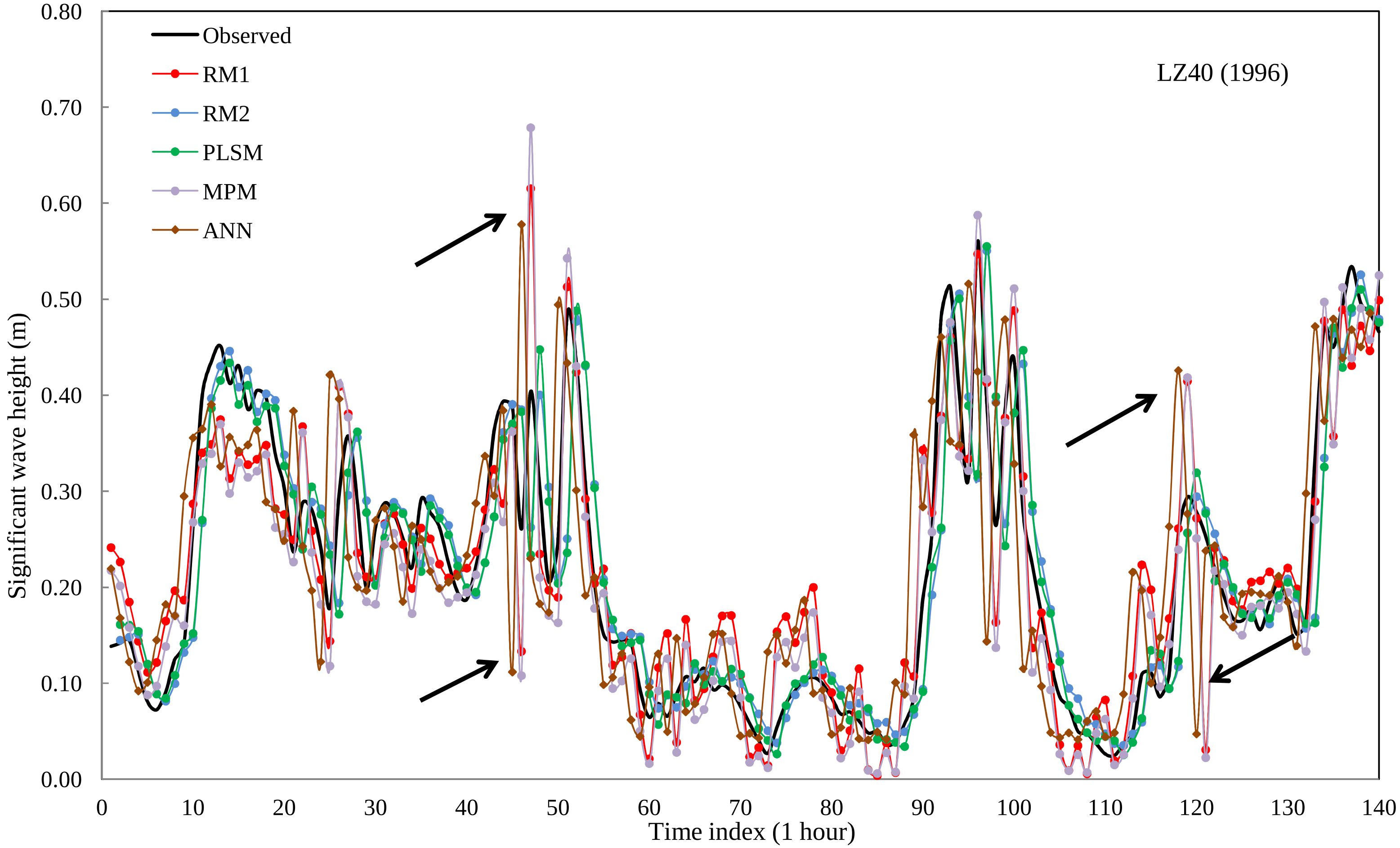

Figure 4.51 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (LZ40 – 140 data of year 1996)

4.5.2 Comparisons of RM1, RM2, PLSM, MPM and ANN at L006

Comparisons of RM1, RM2, PLSM, MPM and ANN predictions are shown in Figures 4.52 and 4.53 for the year 2002 and 1996 data, respectively. Figure 4.52 shows that the results produced by the models RM1 and ANN are almost same. At the beginning of the time series, RM1 and ANN are over predicted than other three models (RM2, PLSM and MPM) and after that both models follow the observed trend line closely. MPM results are slightly under predicted throughout the time series. Best results are found from RM2 and PLSM for both data sets.

Figure 4.53 presents the time series of predicted and observed significant wave heights for the year 1996 data. For this set of data, ANN has better predictions than those from RM1 and MPM. But, still RM1, MPM and ANN show some noticeable fluctuation at few points. RM2 and PLSM are the best predictive models for the year 1996 data.

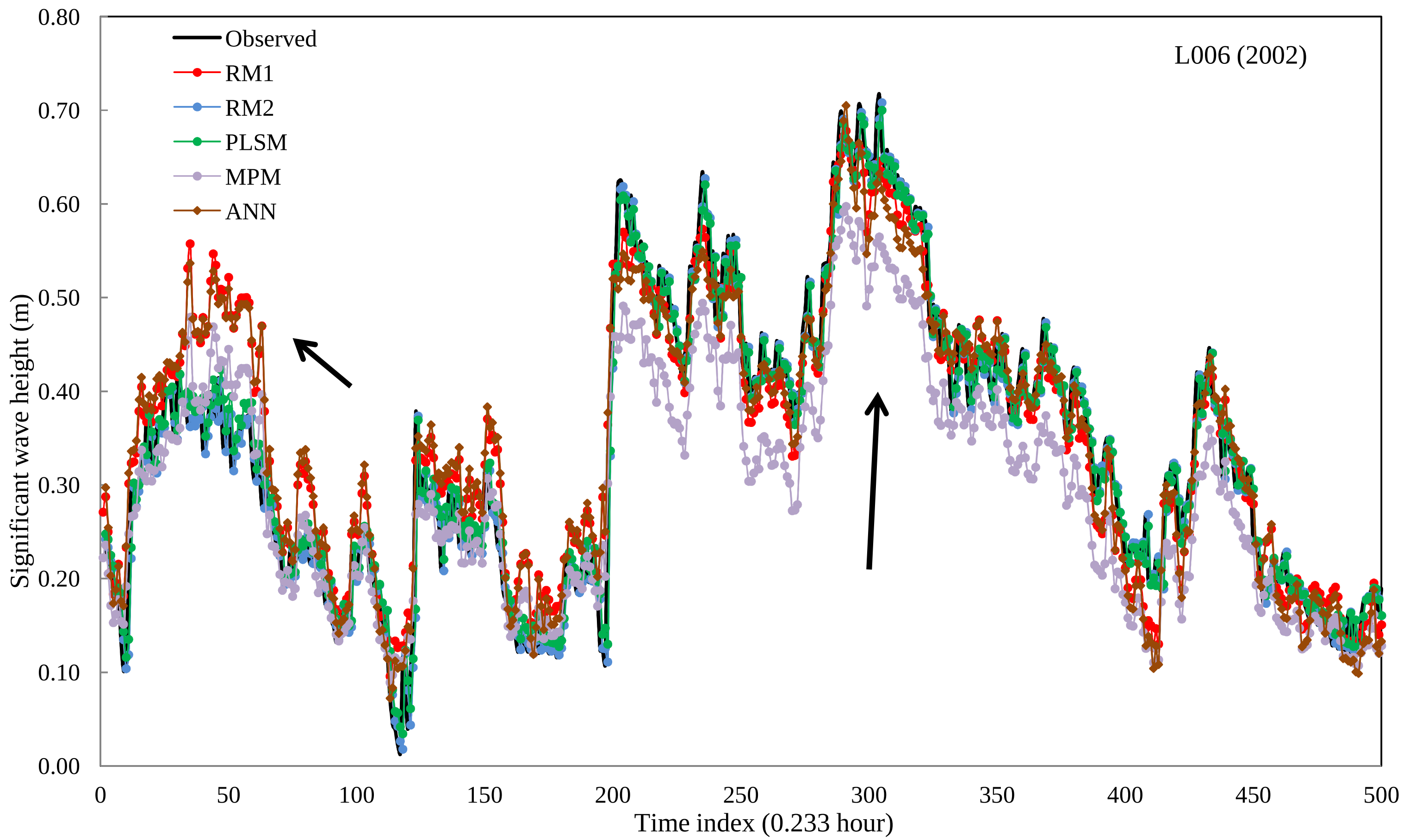

Figure 4.52 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (L006 - 500 testing data of year 2002)

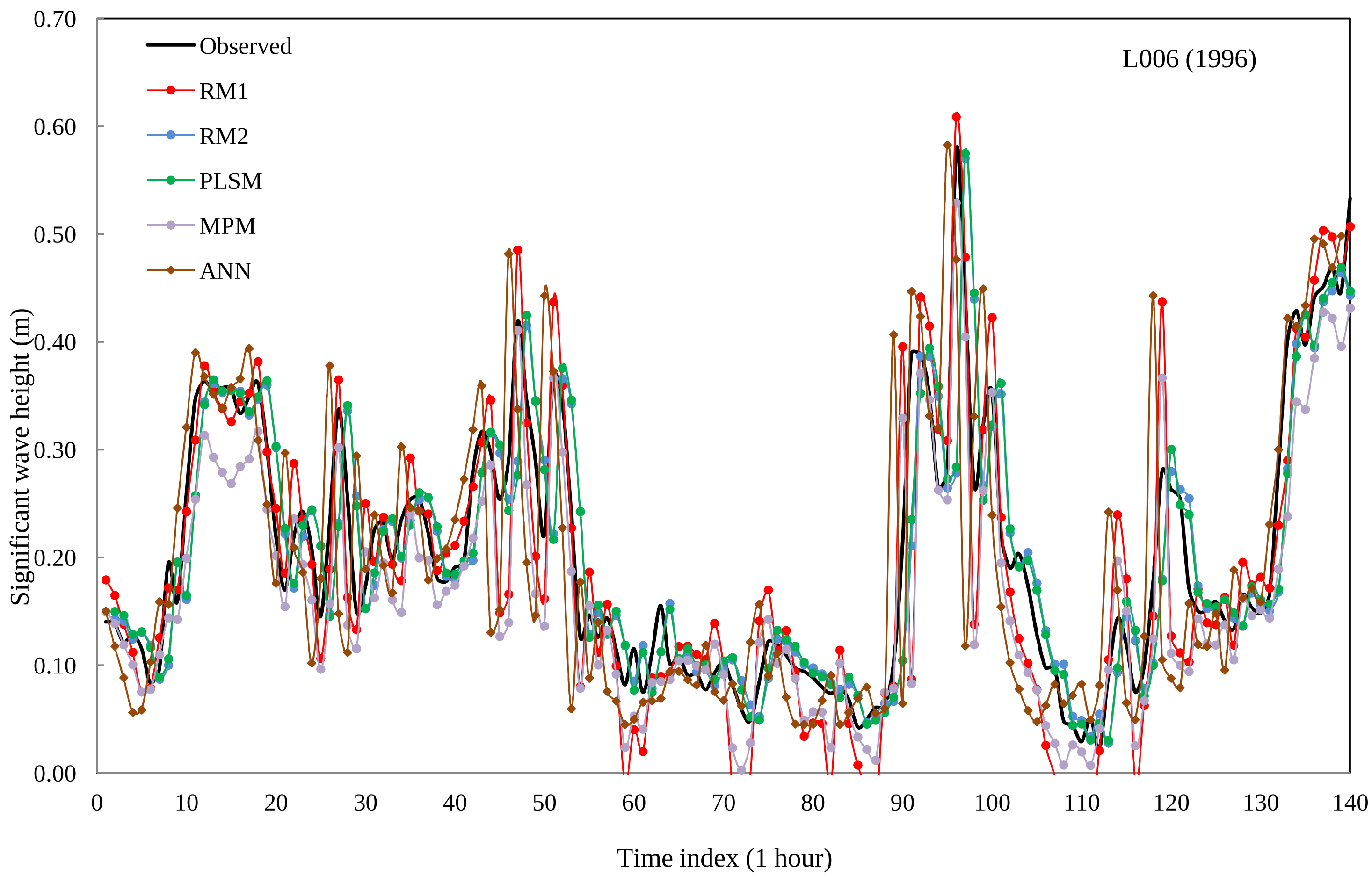

Figure 4.53 Comparison charts of time variations RM1, RM2, PLSM, MPM and ANN (L006 – 140 data of year 1996)

4.5.3 Comparisons of RM1, RM2, PLSM, MPM and ANN at L005

The predicted and observed time variations of significant wave heights for both 2002 and 1996 events are shown respectively in Figure 4.54 and 4.55. ANN results follow a medium trend for L005 (Figure 4.54) which leads to both over and under predictions at different time. Significant wave heights are under predicted by RM1 and MPM for most of the timeline. Again RM2 and PLSM produce most accurate result for the year 2002.

For the year 1996 (Figure 4.55), RM2 and PLSM are again following the original trend line but with slightly time shift. Other models show major fluctuations throughout the time series. The effect of marshy area and high wind event are clearly visible for the models RM1, MPM and ANN. The results from these three models are shown to have more pronounced fluctuations than the RM2 and PLSM. The main reason behind this is that RM1, MPM and ANN are all wind speed based models while predicting significant wave height. As seen earlier in the correlation plot of wind speed and significant wave height (Figure 3.3), correlations are comparatively weaker at L005.

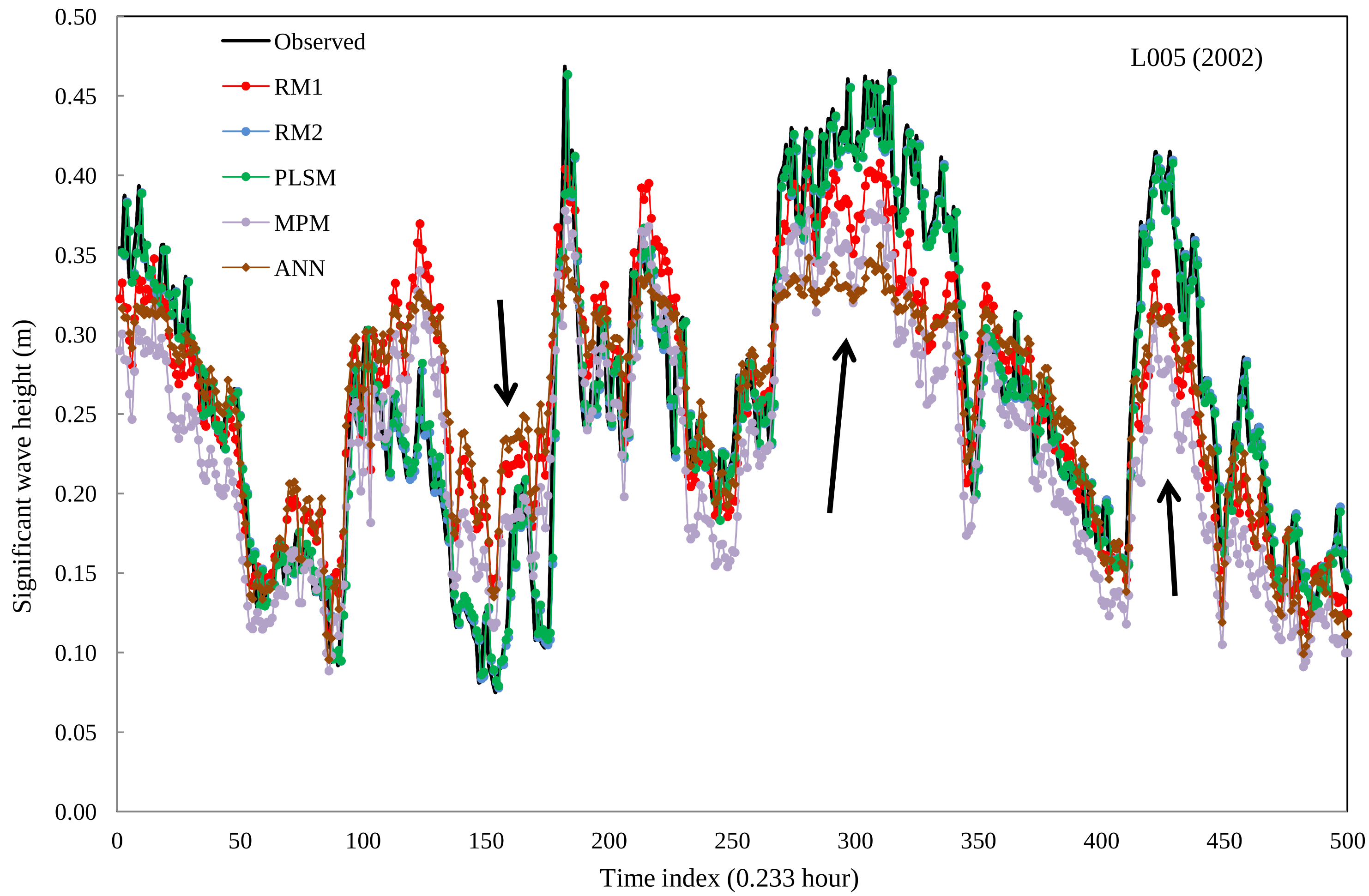

Figure 4.54 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (L005 - 500 testing data 2002)

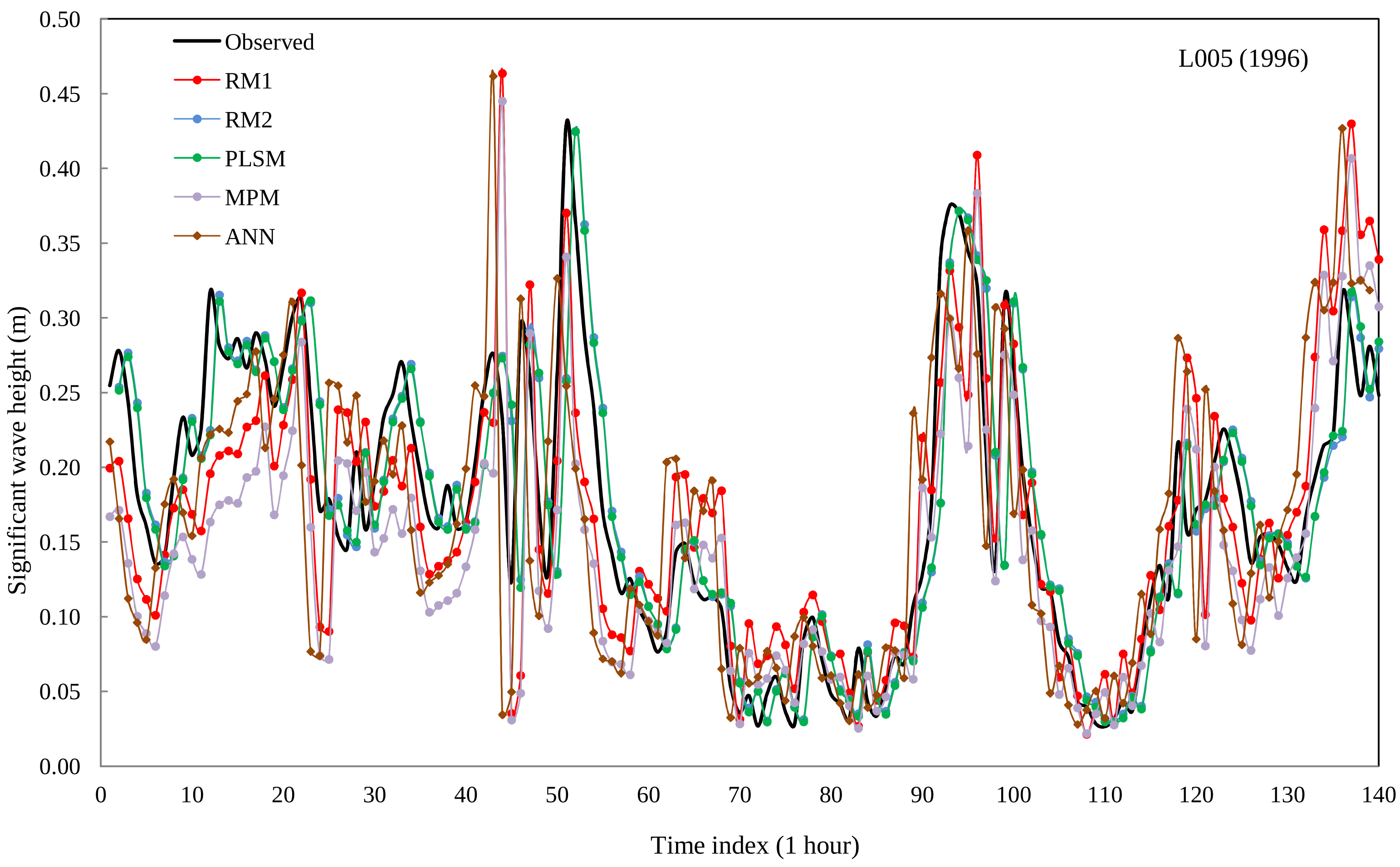

Figure 4.55 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (L005 – 140 data 1996)

4.5.4 Comparisons of RM1, RM2, PLSM, MPM and ANN at L001

At station L001, data from year 1996 is not available. Figure 4.56 presents the comparisons of time variations of significant wave heights generated by RM1, RM2, PLSM, MPM and ANN models. RM2 predicts the significant wave height more accurately. PLSM is also able to predict accurately but slightly over predicted at the end of the time series. Other three models, RM1, MPM and ANN show under predictions and a big jump of trend line at the end of the time series data (Figure 4.56 – marked with arrow). As these three models are dependent on wind speed, the predicted significant wave heights are also following the trend line of wind speed (as shown in Figure 3.4) rather than following the observed significant wave height.

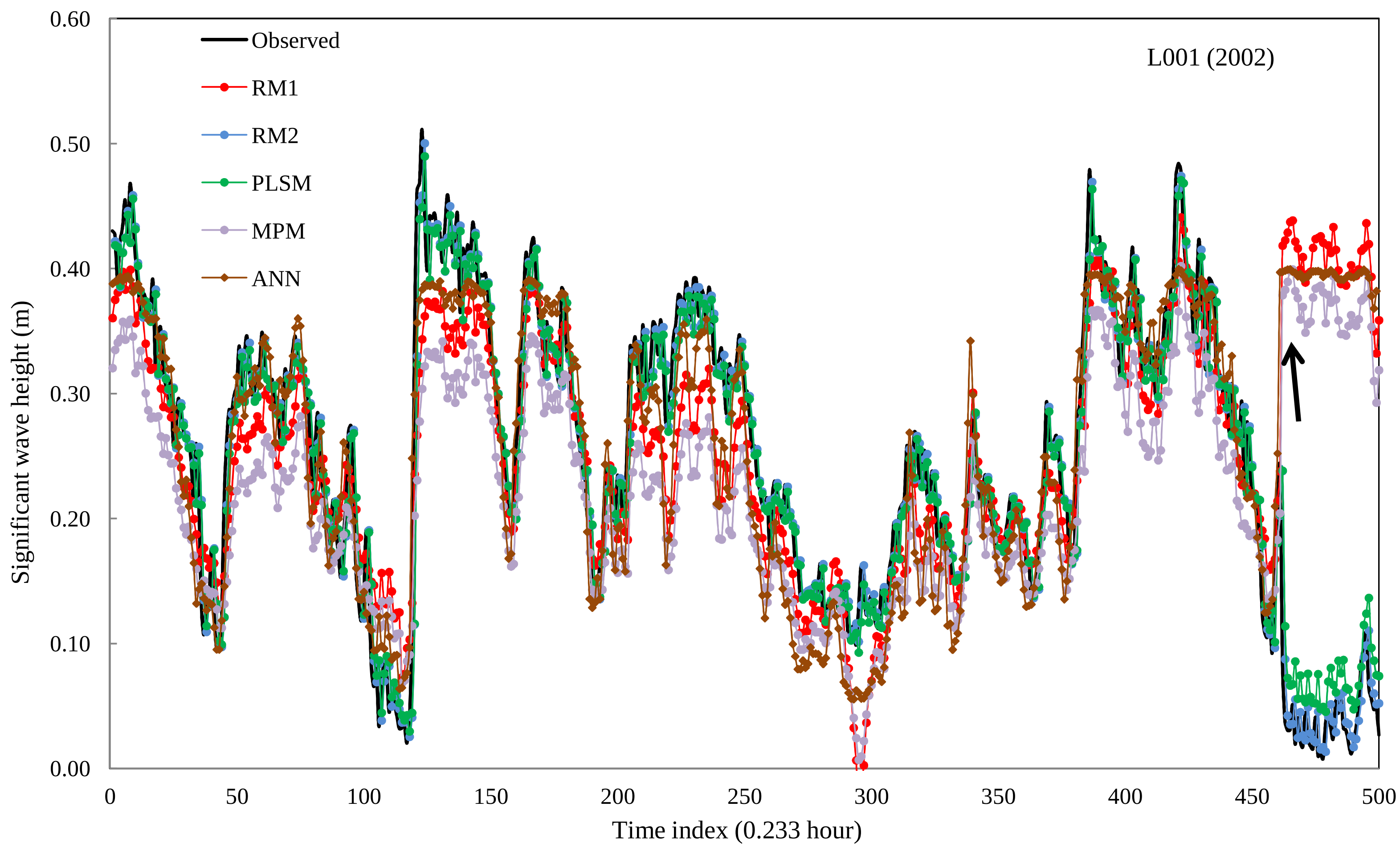

Figure 4.56 Comparison charts of time variations for RM1, RM2, PLSM, MPM and ANN (L001 - 500 testing data 2002)

Table 4.7 and 4.8 are the summaries of AAE, RMSE and CE received by the predictive models for the 500 testing data of year 2002 and the year 1996 data, respectively. Table 4.7 shows that the errors produced by RM1, MPM and ANN are within the same range, which are consistent with the time series comparison plots as discussed above. It can be concluded from both the comparison plots and tables that the models, which use wind speed as independent variable to predict significant wave height, require a condition of strong correlation between wind speed and significant wave height in order to predict accurately. Either the RM2 or PLSM is demonstrated to be a better predictive model with relatively small errors produced.

Table 4.7 Summary of AAE, RMSE and CE for RM1, RM2, PLSM, PMP and ANN for 500 testing data of year 2002

	RM1			RM2			PLSM			MPM			ANN		
	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE
LZ40	0.04	0.06	0.87	0.03	0.04	0.94	0.03	0.04	0.94	0.05	0.06	0.85	0.05	0.06	0.92
L006	0.04	0.05	0.88	0.02	0.03	0.95	0.02	0.03	0.95	0.06	0.07	0.80	0.04	0.05	0.87
L005	0.04	0.05	0.74	0.02	0.02	0.94	0.02	0.02	0.94	0.05	0.06	0.63	0.04	0.06	0.66
L001	0.06	0.11	0.06	0.02	0.03	0.93	0.02	0.03	0.92	0.08	0.11	0.04	0.06	0.11	0.14

Table 4.8 Summary of AAE, RMSE and CE for RM1, RM2, PLSM, PMP and ANN for the year 1996 data

	RM1			RM2			PLSM			MPM			ANN		
	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE	AAE (m)	RMSE (m)	CE
LZ40	0.05	0.06	0.82	0.04	0.07	0.79	0.04	0.07	0.77	0.05	0.06	0.80	0.06	0.05	0.87
L006	0.04	0.06	0.74	0.04	0.06	0.75	0.04	0.06	0.75	0.05	0.06	0.74	0.04	0.06	0.74
L005	0.05	0.06	0.56	0.03	0.05	0.72	0.03	0.05	0.72	0.05	0.07	0.45	0.05	0.06	0.57

Chapter 5

Conclusions and Future Studies

Lake Okeechobee is considered the "Liquid Heart" of South Florida (Tehrani 2001). This lake is an important part of Florida's water resources and eco system. The wind induced waves have a dominant effect on the suspended sediment transport, which indirectly affect the quality of the Lake. This is why it is important to study the wave parameters for Lake Okeechobee.

Four different models are established for this study. They are the regression based RM1 and RM2 models, the Perceptron Least Square Method (PLSM) model and Modified PM (MPM) spectrum model. The RM1, RM2, and PLSM approaches can be considered to follow the simplified stochastic principle while MPM model uses the concept of wave energy spectrum. The RM1 and MPM models include wind speed as the only input variable. On the other hand RM2 uses significant wave height of previous time step to predict the current wave height. The PLSM model however combines both the inputs of previous significant wave height and current wind speed for model prediction. Among all the models, RM1 is commonly used method and able to produce acceptable results. Perceptron Least Square Method is the simple form of Neural Network. The results from this model are best for all stations. This model is able to overcome the effect of marshy area at L005, and the strong wind event that occurred in 1996 for LZ40, L006 and L005. A fairly new approach using the concept of Pierson Moskowitz Spectrum (1964) is presented in study. PM Spectrum is modified (MPM model) to work with shallow lakes and then verified with two different sets of data. This model works fairly well for all the stations except L001 as the testing data collected in this station has some error. It should be noted that the MPM model is the first applicable spectrum model developed for Lake Okeechobee. RM1 and ANN also did not work well with L001. So it can be easily assumed that for these models, it is important to obtain a good set of observed input data.

As a new approach, Modified Pierson Moskowitz Spectrum has potential to work better for shallow lakes like Lake Okeechobee. More research on the wave spectrum models is recommended, especially develop models with water depth as an additional input parameter. In this study least square method is used to calibrate the model parameters. There are some other calibration techniques available in the literature. As an example, genetic algorithm is one of the modern optimization techniques available. For future study it is suggested to couple the Modified PM Spectrum with modern optimization techniques like genetic algorithm to predict the model parameters. The other suggested future study can focus on the comprehensive modeling approach by solving in the fluid domain a system of physically based wave energy transfer functions for the wave parameters, where the wave growth and decay characteristics, wave reflection, and wave-wave interaction can be included in the modeling process.

References

Agrawal, J.D., & Deo, M.C. (2002). “On-line wave prediction.” *Marine Structures,* 15, 57–74.

Altunkaynak, A. (2008). “Adaptive estimation of wave parameters by Geno-Kalman filtering.” *Ocean Engineering,* 35(11-12), 1245–1251.

Altunkaynak, A., & Özger, M. (2004). “Temporal significant wave height estimation from wind speed by perceptron Kalman filtering.” *Ocean Engineering*, 31(10), 1245–1255.

Altunkaynak, A & Wang, K. H. (2012). “Estimation of significant wave height in shallow lakes using the expert system techniques.” *Expert Systems with Applications*, 39, 2549–2559.

Barnett, T. P. (1968). “On the Generation, Dissipation, and Prediction of Ocean Wind Waves.” *Journal of Geophysical Research,* 73(2), 513-529.

Bishop, C. T. (1983). “Comparison of manual wave prediction models.” *Journal of Waterway, Port, Coastal and Ocean Engineering*, 109(1).

Booij, N., Ris, R.C., & Holthujsen, L.H. (1999). “A third generation wave model for coastal regions. Part I: Model description and validation.” *Journal of Geophysical Research*, 104(C4), 7649–7666.

Bretschneider, C.L. (1970). “Wave forecasting relations for wave generation.” Look Lab, Hawaii, 1(3).

Bretschneider, C.L. (1973). "Prediction of Waves and Currents." Look Lab, Hawaii, 3, 1-17.

Chen, Q., Zhao, H., Hu, K., & Douglass, S. L. (2005). “Prediction of wind waves in a shallow estuary.” *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 131(4), 137–148.

Deo, M. C., & Kumar, N. K. (2000). “Interpolation of wave heights.” *Ocean Engineering*, 27, 907–919.

Deo, M. C., Jha, A., Chaphekar, A. S., & Ravikant, K. (2001). “Neural networks for wave forecasting.” *Ocean Engineering*, 28(August 1999), 889–898.

Donelan, M.A. (1980). “Similarity theory applied to the forecasting of wave heights, periods and directions.” *Proceedings* of Canadian Coastal Conference, National Research Council of Canada, 47–61.

Donelan, M.A., Hamilton, J., & Hui, W.H. (1985). “Directional spectra of wind generated waves.” *Philosophical Transactions of Royal Society*, London, A 315, 509-562.

Etemad-Shahidi, A. & Mahjoobi, J. (2009). “Comparison between M5’ model tree and neural networks for prediction of significant wave height in Lake Superior.” *Ocean Engineering*, 36(15-16), 1175–1181.

Fahlman, S.E., & Lebiere, C. (1990). “The cascade correlation learning architecture. In: Touretsky.” *Advances in Neural Information Processing System*. Morgan Kaufmann, San Mateo, CA, 524–532

Fletcher, R., & Reeves, C.M. (1964). “Function minimization by conjugate gradients.” *Computer Journal* 7 (2), 149–153.

Gelb, A. (1974). “Applied Optimal Estimation.” MIT Press, Cambridge, MA.

Goda, Y. (1985). “Random Seas and Design of Marine Structures.” University of Tokyo Press.

Goda, Y. (2003). “Revisiting Wilson’s Formulas for Simplified Wind-Wave Prediction.” *Journal of Waterway, Port, Coastal, and Ocean Engineering*, ASCE 129(2), 93–95.

Hasselmann, K., (1962). “On the Nonlinear Energy Transfer in the Gravity-Wave Spectrum, 1. General Theory.” *Journal of Fluid Mechanics*, 12, 481-500.

Hasselmann, K., Barnett, T.P., Bouws, E., Carlson, H., Cartwright, D.E., Enke, K., Ewing, J.A., Gienapp, H., Hasselmann, D.E., Kruseman, P., Meerburg, A., Muller, P., Olbers, D.J., Richter, K., Sell, W., Walden, H. (1973). “Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP).” *Deutsche Hydrograph*. Zeit, Erganzung-self Reihe, A12, 95.

Hogben, N., (1990). “Long Term Wave Statistics.” *The Sea-Ocean Engineering Science*, 9, Part A, Chapter 8.

Holland, J. (1975). “Adoption in neural and artificial systems.” University of Michigan Press, Ann Arbor, Michigan.

Internet 1, https://maps.google.com, Searched for Lake Okeechobee, Source:

https://maps.google.com/maps?hl=en&q=lake+okeechobee+satellite+map&ie=UTF-8&hq=&hnear=0x88dbf8425d9cfad7:0x2377aaacc36bac15,Lake+Okeechobee&gl=us&t=h&ei=tOaAUJrPLOqU2QWVtIFQ&sqi=2&ved=0CBwQ8gEwAA

Jin, K.R & Ji, Z. G. (2001). “Calibration and verification of a spectral wind – wave model for Lake Okeechobee.” *Ocean Engineering*, 28, 571–584.

Jin, K.R. & Wang, K.H. (1998). “Wind generated waves in Lake Okeechobee.” *Journal of The American Water Resources Association*, 34(5), 1099–1108.

Kalman, R.E., (1960). “A new approach to linear filtering and prediction problems.” Transactions of the ASME Series D, *Journal of Basic Engineering*, 82, 35–45.

Kazeminezhad, M. H., Etemad-Shahidi, A., & Mousavi, S. J. (2005). “Application of fuzzy inference system in the prediction of wave parameters.” *Ocean Engineering*, 32(14-15), 1709–1725.

Kitaigorodskii, S. A., (1961). “Application of the theory of similarity to the analysis of wind generated wave motion as a stochastic process.” Bulletin of the Academy of Sciences, USSR, *Geophysics Series*, 73-80.

Liu, P. C., Schwab, D. J., & Jensen, R. E. (2002). “Has wind–wave modeling reached its limit?” *Ocean Engineering*, 29(May 2000), 81–98.

Liu, S., Jia, N., & Guan, C. (2011). “A proposed frequency spectrum for fully developed wind waves by revisiting P-M spectrum.” *Journal of Ocean*, University of China, 10(4), 331–335.

Mamdani, E. H. (1974). “Application of fuzzy algorithms for simple dynamic plant.” *Proceedings* of the IEEE, 121, 1585–1588.

Makarynskyy, O. (2004). “Improving wave predictions with artificial neural networks.” *Ocean Engineering*, 31(5-6), 709–724.

McCormick, M. E. (1999). “Application of the Generic Spectral Formula to Fetch-Limited Seas.” *Marine Technology Society Journal* (MTS), 33(3), 27–32.

McCormick, M. E. (2010). Ocean Engineering Mechanics - With Applications. Cambridge University Press. ISBN 978-0-521-85952-3.

Moskowitz, L. (1963). “Estimates of the power spectra for fully developed seas for wind speeds of 20 to 40 knots.” *Technical Report*, U. S. Naval Oceanographic Office, New York University, School of Engineering and Science, Research Division.

Neumann, G. (1952). “On Wind Generated Ocean Wave with Special Reference to the Problem of Wave Forecasting.” New York University, College of Engineering, Department of Meteorology, New York, 136.

Olila, O.G. & Reddy, K.R. (1993). “Phosphorous Sorption Characteristics of Sediments in Shallow Eutrophic Lakes of Florida.” *Archiv fur Hydrobiologie*, 129(1), 45-65.

Özger, M., & Şen, Z., (2007). “Prediction of wave parameters by using fuzzy logic approach.” *Ocean Engineering*, 34(3-4), 460–469.

Pappis, C.P., & Mamdani, E.H. (1977). “A fuzzy controller for a traffic junction.” IEEE Transactions on Systems, Man and Cybernetics 7 (10), 707–717.

Phillips, O. M. (1958). “The equilibrium range in the spectrum of wind-generated waves.” *Journal of Fluid Mechanics*, 4, 426- 434.

Pierson, W.J., & Moskowitz, L. (1963). “A proposed spectral form for fully developed wind seas based on the similarity theory of S. A. Kitaigorodskii.” *Technical Report*, U. S. Naval Oceanographic Office, New York University, School of Engineering and Science, Research Division., 1–26.

Pierson, W. J., & Moskowitz, L. (1964). “A proposed spectral form for fully developed wind seas based on the similarity theory of S. A. Kitaigorodskii.” *Journal of Geophysical Research*, 69 (24), 5181-5203.

Rabelo, L.C. (1990). “A Hybrid Artificial Neural Network and KBES Approach to Flexible Manufacturing System Scheduling.” UMI Dissertation Services, Ann Arbor, MI.

Reddy, K.R., Sheng, Y.P., & Jones, B. L. (1995). “Lake Okeechobee phosphorus dynamics study: Summary (volume I).” *Report*, South Florida Water Management District, West Palm Beach, FL.

Rogers, W. E., Hwang P.A., & Wang, D.W. (2003). "Investigation of Wave Growth and Decay in the SWAN Model: Three Regional-Scale Applications." *Journal of Physical Oceanography, 33*, 366–389.

Ross, J.T. (1995). "Fuzzy Logic with Engineering Applications." McGraw Hill, Inc., New York, 593.

Rumelhard, D.E., Hinton, G.E., & Williams, R.J. (1986). "Learning Internal Representation by Error Backpropagation." Parallel Distributed Processing: Explorations Microstructure of Cognition, vol. 1. MIT Press, Cambridge, MA.

Schmitz, H. P. (1962), "Über die Interpretation bodennaher vertikaler Geschwindigkeitsprofile in Ozean und Atmosphäre und die Windschobspannung auf Wasseroberflächen." *Deutsche Hydrograph*. Zeit, 15 (2), 45-72.

Schwab, D.J., Jensen, R.E., & Liu, P.C. (1991). "Comparative performance of spectral and parametric numerical wave prediction models in Lake Michigan." Adeli, H., Sierakowski, R. (Eds.), Mechanics of Computing in 1990's and Beyond. *American Society of Civil Engineers (ASCE).*

Sen, Z. (1980). ''Adaptive Fourier analysis of periodic stochastic hydrological sequences.'' *Journal of Hydrology,* 46, 239–249.

Sen, Z., (1984). ''Adaptive pumping test analysis.'' *Journal of Hydrology*, 74, 259– 270.

Şen, Z., and Altunkaynak, A., (2004). "Fuzzy awakening in rainfall-runoff modeling." *Nordic Hydrology*,35(1), 31–43.

Şen, Z., Altunkaynak, A., & Özger, M. (2003). "Autorun_Persistence_of Hydrologic Design." *Journal of Hydrologic Engineering*, 8(6), 329 – 338.

Sen, Z., Altunkaynak, A., & Özger, M. (2004). "Sediment Concentration and Its Prediction by Perceptron Kalman Filtering Procedure." *journal of Hydraulic Engineering*, 130(8), 816–826.

Sverdrup, H.V. & Munk, W. H. (1947). "Wind, Sea and Swell: Theory of relations for forecasting." U.S. Navy Hydrograph Office, Publication no. 601.

Tehrani, M. (2001). "Seasonal variation of flow velocity and suspended sediment concentration in Lake Okeechobee." *Thesis*, University of Houston.

Tsai, C.P., Lin, C., & Shen, J.N. (2002). "Neural network for wave forecasting among multi-stations." *Ocean Engineering* 29, 1683–1695.

Uyumaz, A., Altunkaynak, A., & Özger, M. (2006). "Fuzzy Logic Model for Equilibrium Scour Downstream of a Dam's Vertical Gate." *Journal of Hydraulic Engineering*, 132(10), 1069–1075.

Wang, K. H. (2002). "Wind-wave data collection for Lake Okeechobee." Final report, C-12744-016, South Florida Water Management District.

Xiong, L., Shamseldin, A.Y., & O'Connor, M.K. (2001). "A non- linear combination of the forecasts of rainfall-runoff models by the first order Takagi–Sugeno fuzzy system." *Journal of Hydrology*, 245, 196–217.

Yeh, Y.C., Kuo, Y.H., & Hsu, D.S. (1993). “Building KBES for diagnosing PC piles with artificial neural networks.” *ASCE Journal of Computing in Civil Engineering* 7 (1), 71–93.

Zadeh, L.A. (1965). “Fuzzy sets.” *Information and Control*, 12, 94–102.