School of Engineering & Design

NEURAL NETWORK BASED SIMULATION OF SEA-STATE SEQUENCES

A
thesis
submitted for the degree
of
Doctor of Philosophy

by

Hamid Bazargan-Harandi

July 2006

NEURAL NETWORK BASED SIMULATION OF SEA-STATE SEQUENCES

A thesis submitted for PhD degree

by

Hamid Bazargan

School of Engineering & Design
Brunel University

July 2006

Abstract

The present PhD study, in its first part, uses artificial neural networks (ANNs), an optimization technique called simulated annealing, and statistics to simulate the significant wave height (H_S) and mean zero-up-crossing period (T_z) of 3-hourly sea-states of a location in the North East Pacific using a proposed distribution called hepta-parameter spline distribution for the conditional distribution of H_S or T_z given some inputs. Two different seven- network sets of ANNs for the simulation and prediction of H_S and T_z were trained using 20-year observed H_S's and T_z's. The preceding H_S's and T_z's were the most important inputs given to the networks, but the starting day of the simulated period was also necessary. However, the code replaced the day with the corresponding time and the season. The networks were trained by a simulated annealing algorithm and the outputs of the two sets of networks were used for calculating the parameters of the probability density function (pdf) of the proposed hepta-parameter distribution. After the calculation of the seven parameters of the pdf from the network outputs, the H_S and T_z of the future sea-state is predicted by generating random numbers from the corresponding pdf.

In another part of the thesis, vertical piles have been studied with the goal of identifying the range of sea-states suitable for the safe pile driving operation. Pile configuration including the non-linear foundation and the gap between the pile and the pile sleeve shims were modeled using the finite elements analysis facilities within ABAQUS. Dynamic analyses of the system for a sea-state characterized by H_S and T_z and modeled as a combination of several wave components were performed. A table of safe and unsafe sea-states was generated by repeating the analysis for various sea-states. If the prediction for a particular sea-state is repeated N times of which n times prove to be safe, then it could be said that the predicted sea-state is safe with the probability of $100(n/N)\%$.

The last part of the thesis deals with the H_S return values. The return value is a widely used measure of wave extremes having an important role in determining the design wave used in the design of maritime structures. In this part, H_S return value was calculated demonstrating another application of the above simulation of future 3-hourly H_S's. The maxima method for calculating return values was applied in such a way that avoids the conventional need for unrealistic assumptions. The significant wave height return value has also been calculated using the convolution concept from a model presented by Anderson et al. (2001).

Keywords: Significant wave height; Mean zero-up-crossing period; Simulation; Prediction; Neural networks; Simulated annealing; Finite elements method, Pierson-Moskowitz spectrum; Safe pile-driving;

Table of Contents

List of Figures

Nomenclature

a	output of an artificial neuron
A	the FT-I, GEV, GPD & Weibull location parameter(see Appendix E)
a_j	the amplitude of the j^{th} wave component
ANN(s)	artificial neural network(s)
ASA	adaptive simulated annealing
b	1) the bias of an artificial neuron; 2) the Rayleigh distribution parameter; 3) a positive number
B	the FT-I, GEV, GPD & Weibull scale parameter (see Appendix E)
BP	back-propagation training algorithm
C	the Weibull, GEV, and GPD shape parameter (see AppendixE)
c	a high value H_s
c_0, c_1	constants in Eq. (1) of Chap. 4 and Eq. (22) of Chap. 6
c_2, c_5	constants in Eq. (1) of Chap. 4 and Eq. (22) of Chap. 6
CEM	Coastal Engineering Manual
C.F.	characteristic function
CDF	cumulative distribution function
chap.	chapter
d =day	a real number in the interval 0-365.25 with step 0.125, d=0 represents 0 AM 1 Jan.
D	pile outer diameter
DE	differential evolution optimisation technique
E	the energy of the system
EBSA	ensemble based simulated annealing
EDF	empirical distribution function
EV	extreme value
f	wave frequency (Hz)
f_{tr}	the neuron transfer function
$f(x)$	probability density function
F	wind fetch (km),
F_i	the plotting position for the i^{th} component of the ordered data
f_i	the frequency of the i^{th} wave component (Hz)

F^{-1} the inverse function of the CDF fitted to all the data

$f_{H_S(t_i)given8}(h_i) = f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};...,Hs(t_{i-8})=h_{i-8}}(h_i)$

the pdf of the distribution of H_s (t_i) given its 8 immediate preceding 3-hourly successive H_s's

$f_{Tz(t_i)|Tz(t_{i-1})=tz_{i-1},Tz(t_{i-2})=tz_{i-2},Hs(t_{i-1})=h_{i-1},Hs(t_i)=h_i,Hs(t_{i+1})=h_{i+1}}(tz_i)$

denotes the conditional pdf of T_z (t_i) given some H_s's and T_z's

FE(M) Finite Element (Method)

FT-1 Fisher-Tippet Type-I distribution

FT-II Fisher-Tippet Type-II distribution

FT-III Fisher-Tippet Type-III distribution

$F_X * F_Y$ the convolution of the distribution function of X and that of Y

F_{X+Y} the cumulative distribution function of X+Y

g gravity acceleration rate

GA genetic algorithm

GD gradient descent algorithm

GD networks The ANNs trained with gradient descent algorithm

GEV generalized extreme value distribution

GPD generalized Pareto distribution

GSA generalized simulated annealing

h time (hours)

H 1) neurons number of the hidden layer in a 2 layer network; 2) wave height

$H_1, H_2, ..., H_n$ a random sample of wave heights

h_i the value of significant wave height at time t_i; a particular value for H_s (t_i)

$\overline{H}$ the mean of the heights of all waves in a record

H_{rms} root mean square of sample wave heights

H_s significant wave height

$H_s(t_i)$ a random variable denoting the 3-hourly H_s measured at time t_i

$H_s(t_i) \mid H_s(t_{i-1}) = h_{i-1}, H_s(t_{i-2}) = h_{i-2}, ..., H_s(t_{i-8}) = h_{i-8}$

a random variable representing H_s (t_i) given its 8 preceding 3-hourly H_s's

H_{sNyr} N-year return value of H_s

i.i.d.	independent and identically distributed
IASA	integer augmented simulated annealing
Input-Vect	a general name for the columns of the input matrix of the implemented ANNs, corresponding to a desired T_z in the target vector
k	the Boltzmann constant
k_i	number of neuron of a layer
k_{in}	the number of the neuron of the input layer
k_{OUT}	the number of the neuron of the output layer
log	natural logarithm
L	1) pile length 2) distance from equivalent points on a wave
m	1) sampling period, time interval between two successive H_s 's (hours); 2) the order of Markov chain
M	number of observations per year
m_n	the n^{th} moment of energy density spectral function,
MAE	mean absolute error
MLE	maximum likelihood estimate
MLP	multi-layer perceptron
MSE	mean of squared errors
MSEREG	mean of squared errors with regularization
MWL	mean water level
n	1)number of data; 2)number of data above a threshold
N	1) the return period; 2) total number of zero up-crossings in the time history
p	1) input to an artificial neuron; 2) a given probability
pdf	probability distribution function
PM	Pierson-Moskowitz
PU	processing unit, neuron
POT	peaks over threshold
Q-Q	quantile-quantile
r	pile outer radius
R	1)number of inputs to the neural network, 2) correlation coefficient
RASA	real-coded augmented simulated annealing

res=residuals	denotes the residuals in Eq. (22) of Chap. 6
RMSE	root mean square error
rps	radiant per second
S	number of neuron in a layer
$\overline{S}$	the mean of the stress time history related to a pile element (MPa)
SA	simulated annealing algorithm
SA networks	the ANNs trained with simulated annealing algorithm
$S(f_j)$	energy density of the j^{th} wave component (m^2/ Hz)
S_e	most likely extreme stress occurred in the pile after h hours (MPa)
SF	safety factor
SI	scatter index
SOFM	Self-Organizing Feature Map
SPM	Shore Protection Manual
SSE	sum of squared errors
S_y	the pile yield stress (MPa)
t	pile thickness
T	1) a parameter in simulated annealing algorithm denoting the system *temperature*; 2) mean zero-crossing period for the time history of the stress of an element in the pile finite element model
t_0	the initial time or the starting day of the desired period for the prediction
t_i	the time of the i^{th} 3-hourly measurement of significant wave height
T_i	the current *temperature* in SA algorithm
T_{i+1}	the value of parameter *temperature* in the next iteration of SA
TF	transfer function
T_p	wave peak period
T_z	mean zero up-crossing period of sea-state
$T_{1/3}$	significant wave period
$T_z(t_i)$	a random variable denoting the 3-hourly T_z measured at time t_i
tz_i	a specific variate of $T_z(t_i)$
u	a threshold

v	the summing junction output in the artificial neuron
VFSA	very fast simulated annealing
VFSR	very fast simulated re-annealing
USNODC	United States National Oceanographic Data Center
W	wind velocity (km/h)
$w, w_1,\ldots,w_R$	artificial neuron synaptic weights
w_0	bias of the neuron
X	1) a continuous random variable having a range in $(0, \infty)$; 2) annual cycle in Eq.(29) of Chap. 6
x_i	1) output of i^{th} network in Chap. 4 ;2) observed H_s or T_z
$x, x_1,\ldots,x_R$	inputs of artificial neuron in some figures of Chap. 2
$x_{(1)}, x_{(2)},\ldots, x_{(n)}$	n values of the *res* exceeding the threshold *u*
$\overline{X}_u$	$\sum_{i=1}^{n} x_{(i)} - u$
$\overline{X}$	x_i's mean,
y_i	simulated H_s or T_z
Y	continuous random variable having a range in $(0, \infty)$; 2) *res* in Eq. (29) of Chap. 6
$\overline{Y}$	y_i's mean
Z	denotes (annual cycle | *res* > u + *res* | *res* > u) in Chap. 6
Z_{extr}	an extreme value of Z satisfying Eq. (19) of Chap. 6
β	a suitably small value used in SA algorithm
ΔC	the difference between the cost function of current and new configurations in generalised simulated annealing
ΔE	the change in the system energy e.g. E_2-E_1
ε	error
ε_j	phase angle of the j^{th} wave component
η_{rms}	root mean square of sea surface elevation
η (*x, y, t*)	sea surface height above the MWL at time *t* for the coordinates *x* and *y*
θ_j	the angle between the direction of wave propagation and the X–axis in counter clockwise direction
λ	mean rate that threshold is exceeded in the POT method

ρ	liquid density
σ	standard deviation of the stress time history of an element (MPa)
σ_{η}	the standard deviation of the sea surface elevation
σ_{η}^{2}	the variance of the sea surface elevation
σ_{all}	allowable stress in the pile (Mpa)
$\Phi_X(t)$	the marginal characteristic function of random variable X
$\Phi_{X,Y}(t, u)$	the joint characteristic function of random variables X and Y
$\Phi_Y(t)$	the marginal characteristic functions of random variable Y
ω_j	frequency of the j^{th} wave component (radiant per second).

Acknowledgment

The author is indebted extremely to the following kind people for their invaluable support:

- Dr. Peyman Adl, Former Academic Staff of Brunel University, UK
- Mr. A. Aminzadeh, student, UC, Berkeley, USA
- Mr. N. Aminzadeh, Shahid Bahonar University of Kerman, Iran
- The late Dr. Farzad Aryana, Shahid Bahonar University of Kerman, Iran
- Dr. H. Bahai, Senior Lecturer, Brunel University, UK
- Prof. N. Barltrop, Universities of Glasgow & Strathclyde, UK
- Prof. W. Balachandran , Brunel University, UK
- Mr. A. Bazargan, student, College of Engineering, Kerman, Iran
- Dr. D. Carter of Satellite Observing Systems, UK
- Mr. S. Ebrahimi-nejad, student, College of Engineering, Kerman, Iran
- Prof. Y. Goda, Professor Emeritus, Tokyo University, Japan
- Prof. G.G. Hamedani, Marquette University, USA
- Prof. E. Fakouri, Professor Emeritus, Indaina State University, USA
- Dr. C. Fakouri, Indiana State University, USA
- Prof. Isat, Brunel University, UK
- Prof. Khalil Hindi,Professor Emeritus, Brunel University
- Dr. O. Makaranskyy, Curtin University, Australia
- Prof. A. Munjiza, Queen Mary University of London
- Dr. Mustafa Ozbayrak, Brunel University, UK
- Dr. M. Rashidinejad, Shahid Bahonar University of Kerman, Iran
- Dr. S. Yasseri, Klegg Brown & Root (KBR), UK.

Appreciation is extended to Mrs. Linda Hedges and Mrs. Margaret Hodgson administrativestaff of our department at Brunel University. The sustained encouragement and financialassistance of my brothers and sisters especially Prof. A. Bazargan, academic staff of TehranUniversity and former-UN expert, is gratefully acknowledged. The support and patience of my family –wife, daughter and son- is highly appreciated. Finally I would like to thank those unknown experts who have produced the very helpful softwares of MATLAB and ABAQUS/AQUA.

Chapter 1
Introduction, Literature Survey, and Research Objectives

Nomenclature

ANN(s)	artificial neural network(s)
CEM	Coastal Engineering Manual
EV(A)	extreme value (analysis)
FE(M)	finite element (method)
H_s	significant wave height
POT	peaks over threshold (method) in extreme value analysis
SA	simulated annealing
SPM	Shore Protection Manual
T_z	wave mean zero-crossing period
T_p	wave peak period

1.1 Introduction

In coastal and open ocean human activities, there is an increasing demand for accurate forecasting of sea-state parameters. Some of these activities include driving offshore piles, planning offshore drilling operations, planning shuttle tanker take off times, and planning for transport and installation of offshore platforms. In such activities predictions of wave heights and periods are of particular importance. There are a number of methods, such as mathematical (numerical, statistical,) models which have been used for the forecasting. Mathematical models use wind wave relationships to forecast the environmental conditions e.g. a sea-state characterized by significant wave height (H_s) and mean zero up-crossing period (T_z) defined in Chapter 5. A variety of mathematical models have been presented by researchers for predicting wave heights and periods. However there are alternatives such as artificial neural network (ANN) modelling to the approximated mathematical models. ANN modelling is based on developing neural networks and belongs to that type of information-processing systems which allow an approximation to nonlinear behaviour that is characteristic in geophysical processes (Makarynskyy, 2002). ANNs have effectively been applied for predicting natural phenomena containing uncertain interrelationships among their physical parameters.

There are many fields that ANN has been applied to including aerospace, automotive, undersea mine detection, electronics, entertainment, speech processing (e.g. recognition and

generation), telecommunication and specifically in the recognition of speakers in communications, recovery of telecommunications from faulty software, linguistics (e.g. interpretation of multi-meaning Chinese words), texture analysis, three-dimensional object recognition, hand-written word recognition, facial recognition, medical (e.g. diagnosis of hepatitis), oil and gas industries, robotics, manufacturing, and ocean engineering. A part of this PhD work proposes an ANN-based method for simulating and forecasting sea-states characteristics i.e. H_S and T_z.

1.2 Literature Survey

1.2.1 Mathematical Modeling of Ocean Waves Characteristics

A number of studies have in the past years resulted in various mathematical methods for forecasting wave heights and periods mainly from the characteristics of the generating winds. Some of the more recent reported works are reviewed in this section. According to Liu et al (2002) Klaus Hassemann in 1963 laid out the framework for modern numerical modelling of wind waves. The first wave models, which were developed in the 1960s and 1970s, assumed that the wave components suddenly stopped growing. These so-called first generation wave models exhibit basic shortcomings including overestimating the wind input. Extensive wave experiments including Mitsuyasu(1969), Hasselmann et. al. (1973), Snyder et al. (1981), Hasselmann et al. (1986), led to the development of second generation wave models which attempted to simulate properly the so-called overshoot phenomenon and the dependence of the high-frequency region of the spectrum on the low frequencies. The SWAMP Group (1985) presented an intercomparison study of wind wave prediction models, discussing and documenting the shortcomings of the first and second generation models. Then the development of third generation models was suggested in which the wave spectrum was computed without any prior restriction on the spectral shape. As a result, the WAM group was established, whose main task was the development of such a third generation wave model. Komen et al. (1994) culminated an international effort of wind wave model development which lasted over a decade (Liu et. al., 2002). NESS (1997) gives an overview of the work carried out during the North European Storm Study (NESS). WMO (1990) & WMO (1998), two guides published by World Meteorological Organization, present a comprehensive review of various mathematical methods used for wind wave modeling. Khandekar's (1990) book also covers the prediction of wind waves.

Some of the well-known wind-wave models are: the JONSWAP model developed by Hasselmann et al. (1973), HYPA model presented by Gunther et al. (1979) which is an

extension of an earlier two-parameter model, UKMO model, the model by Golding (1983) and the so-called third generation WAM model. The description of the first, second and third generations of numerical models for forecasting waves are given in many references including WMO (1998). The details of the WAM model are covered by the WAMDI group (1988) and a comprehensive description of the model and its various applications are given by Komen et al. (1994). The CEM methods described in the Coastal Engineering Manual (2003) and the empirical SPM methods given in the Shore Protection Manual (1984) are also among favorable methods of wave predictions. These two kinds of numerical models predict wave parameters from wind properties using empirical formulae. Kaziminezhad et al. (2004) have investigated the performance of the CEM and SPM methods for predicting the H_s and the peak period (Tp) in a region in the Caspian Sea. Their results show that the models tend to overpredict H_s and underpredict Tp for the region. Bishop et al. (1992) have reviewed and compared the SPM methods, the JONSWAP, and the Donelan and Sverdrup-Munk-Bretschneider methods. Their results indicate that the 1984 version of the SPM, which uses an adjusted wind speed factor based on friction velocity, is the poorest predictor among the four methods and tends to overpredict wave height and period.

Lin et al. (2002) have tested two very different numerical wave models for the Chesapeake Bay, i.e. the Simulation of WAves Nearshore (SWAN) model and the Great Lakes Environmental Research Laboratory (GLERL) model. Both models behaved well in response to suddenly changing winds and, in general, both over-predicted H_s but the SWAN over-predicted more than the GLERL did. The SWAN model had a larger scatter index and a smaller correlation coefficient for wave height compared to those of the GLERL model. The formulae of correlation coefficient and scatter index have been given by Esq. (31) and (32) of Chapter 4. More over, both models slightly under-predicted Tp with a fairly large scatter index and low correlation coefficient but SWAN predicted mean wave direction better than GLERL. Another type of mathematical methods includes modelling time series of wave characteristics (e.g. H_s) into such models as auto-regressive models. Athanassoulis (1995), Cunha & Guedes-Soares (1999), Anderson et al. (2001) and Agrawal & Deo (2002) are typical works dealing with such modeling. For example Cunha & Guedes-Soares (1999) propose a transformation for the H_s time series and then fit an auto-regressive model to the transformed series. Anderson et al. (2001) applied a logarithmic transformation on H_s and then describes the behaviour

of log H_s in terms of a combination of sine and linear functions; but the residuals (errors) of this model are considerable.

Liu et al. (2002) compare four different numerical wind wave prediction models: the WAM, the GLERL/Donelab, the DWAVE and SHALWV models to illustrate that typical wave prediction models based on the concept of a wave energy spectrum may have reached a limit in the accuracy. All of these models assume that sea-states can be represented by a wave energy spectrum. This set of models includes a very complex model (the WAM) and a simple one (the GLERL). The paper reports that the four disparate models produce results that are more similar to each other than they are to the observed data. The paper concludes that there may be an underlying limitation to further improvement of the spectral models and therefore new approaches to wave may be required for further substantial improvements.

1.2.2 ANNs in Ocean Engineering

Numerical wave models, which have usually provided wave predictions are based on deterministic equations and in some cases, these do not entirely account for the complexity and uncertainty of the wave generation. Deo et al. (2001) state that despite considerable advances in forecasting environmental conditions using wind-wave relationships in the past five decades, the solutions obtained are neither exact nor uniformly applicable all times at all sites. This is because of the complexity and uncertainty of the wave generation phenomenon and also that wave prediction from the knowledge of generating wind is basically a random or uncertain process and hence difficult to model using deterministic equations. This makes it ideal to apply ANN modeling to, since ANNs are useful to model random inputs with the corresponding random outputs. ANNs have been effectively applied for predicting natural phenomena containing some ambiguous interrelationships among their physical parameters and have been successfully applied to many fields including coastal and ocean engineering. Some of the researches that have applied ANNs to coastal engineering are briefly described in the following paragraph.

Chan et al. (1995) use ANNs as an alternative to pile driving formulae. They have trained a neural network to predict the pile bearing capacity from dynamic testing data and report that the trained back-propagation ANNs produced better results than a pile driving formula approach. Mangal et al. (1996) present a ANN-aided scheme for automating the vibration monitoring method of detecting damage in offshore jacket platforms, and use a kind of neural network for damage diagnosis and investigate its advantages and limitations. The authors also

compare a back-propagation network and an ART network. Romero & Pamukcu (1996) have used ANNs to predict shear modulus of granular offshore sediments. Tsai and Lee (1999) present an application of ANNs with back-propagation training algorithm for accurate forecast of tidal-level variations. Unlike the conventional harmonic analysis, this model forecasts the time series of tidal levels directly based on a set of previous data. The results indicate that the hourly tidal levels over a long duration can be efficiently predicted using only a very short-term hourly tidal record. Roske (1997) and Ultsch & Röske (2002) are two works that have applied neural networks in the prediction of sea levels. The latter has used self-organizing map ANNs and compared the accuracy of the ANN results with those obtained from six conventional models including two hydrodynamic models, a statistical model, and three other models. The paper reports that the ANNs predict sea level better than the six models mentioned above. Rajasekaran et al. (2005) have used two neural procedures for the prediction of tides using very short-term observation. Results show that the hourly tides for even a month can be predicted efficiently. Thirumalaiah & Deo (2000) apply ANNs to the field of hydrology. The authors have dealt with the real-time forecasting of hourly flood runoff and daily river stage as well as the prediction of rainfall using ANNs; they report that in many situations the performance of ANNs methods was better than that of complicated statistical models. The results of Huang & Foo (2002) indicate that ANN modeling is capable of correlating the non-linear time series of salinity to the multiple forcing signals of wind, tides and freshwater input in the river being studied. According to the authors the model is an easy-to-use tool to obtain a quick assessment of salinity variation in response to the engineering modifications to the river system. In the paper there is also a recommended formula for the relationship between the number of neurons in the input layer and the hidden layer of a 3-layer ANN. This formula is given in Chapter 2. Lee and Jeng (2002) have developed an ANN model for forecasting the tidal level using a short-term tidal record. Unlike the conventional method of harmonic analysis, which requires a large amount of observed tidal data for estimating the appropriate harmonic parameters, the ANN model can easily decide the unknown parameters by learning the input–output interrelation of short-term tidal records. Toparli (2002) calculates residual stresses within steel bars after quenching using the finite element (FE) method and an ANN algorithm. Finite element method calculates the temperature distribution with time and thermal residual stress values in the samples after cooling. The calculated temperature and thermal residual stress values were used in training a multi-layer feed-forward backpropagation ANN. The results obtained from the ANNs have been compared with the FE results. Comparison showed good agreement. Makarynskyy (2005a) used a feed-forward three-layer ANN to retrieve and predict

sea level variations. This methodology demonstrated reliable results when compared with observed data in terms of such indices as the correlation coefficient and root mean square error.

1.2.2.1 ANNs in forecasting wave characteristics

As mentioned earlier an alternative approach in the complicated process of ocean wave generation methods is the one that uses neural networks. ANNs have recently been applied to predicting ocean wave periods and heights. Some recent chronological applications of the *ANN*s to predicting wave characteristics are reviewed here. Deo et al. (2001) use a 3-layered feed-forward ANN to obtain H_s and T_z (output) from wind speeds (input). They conclude that an appropriately trained network could provide satisfactory results for certain types of predictions and unlike deterministic models, wind fetch and duration do not seem to be necessary to be given as input to ANNs. Agrawal and Deo (2002) deal with on-line prediction of wave heights by ANNs as well as by first order auto-regressive moving average (ARMA) and auto-regressive integrated moving average (ARIMA) models. They reported that ANNs resulted in more accurate predictions of wave heights than the above two time series models when shorter prediction intervals were involved; but for long-range predictions both the stochastic and neural approaches showed similar performance. Makarynskyy et al. (2002) deal with forecasting H_s of subsequent 3-hourly intervals by training a set of ANNs with architecture $8 \times 17 \times 1$ given 24-hour H_s history as input. The closer the time, the better the reported forecast. Makarynskyy (2003) attempts to improve wave short-term forecasts using artificial neural networks. Hourly observations of significant wave heights and zero-up-crossing wave periods from two sites offshore the Atlantic and the Irish Sea coasts of Ireland are used to train and validate these networks. Two different approaches are involved. One of them corrects the predictions solely using the initial simulations of the wave parameters with leading times from 1 to 24 h. Another one allows merging the measurements and initial forecasts. The author report satisfactory results with the proposed procedures. Agrawal and Deo (2004) train some ANNs to estimate the values of mean zero-crossing wave period, peak-spectral period, maximum spectral energy density and maximum wave height from the given value of significant wave height and also to evaluate spectral width parameter from the spectral narrowness parameter. The trained network when tested revealed that it formed a useful tool in exploring the interdependency in between the parameters. Makarynskyy et al. (2005b) employ two different neural network approaches to forecast sea-state characteristics (H_s & T_z) for 3, 6, 12 and 24 hours in advance. In the first approach, 8 separate ANNs were implemented to

simulate every wave parameter over each prediction interval. In the second approach, only two networks provided simultaneous forecasts of the wave parameters for the four prediction intervals. The suitability of ANNs has been shown through verifying the short-term forecasts by the observed data. However the results obtained in simultaneous forecasts exhibited less accuracy than the predictions produced separately due to large numbers of input-processing-output nodes and thus more complicated interrelations among them. Makarynskyy (2005) presents a neural methodology allowing the estimation of wave conditions onshore and/or offshore on the basis of the wave information from distant sources. The simulated and the observed data has rerted to have overall satisfactory agreement and the author concludes that the methodology is reliable. Makarynskyy et al. (2005c) present a neural data interpolation methodology implemented to restore missing wave measurements. The methodology is based on the ability of ANNs to find and reproduce non-linear dependencies within complex geophysical systems. The restored wave records and actual measurements have demonstrated a fairly good overall agreement.

1.2.3 Simulated Annealing

The work reported in this thesis has used one of operations research techniques called simulated annealing (SA) to train the implemented neural networks. Its simplicity and good results over a wide range of problems have made it a highly popular optimisation tool, with hundreds of applications in a great variety of fields including some in engineering. Simulated annealing has been discussed widely in the literature including Dowsland (1995) and Salamon et al. (2002), which are two books on the facts and improvements of SA. Eglese (1990) also describes the algorithm and the physical analogy it is based on, and compares it with a few other optimisation algorithms such as repeated descent algorithm and reports that SA gives significantly better results than the algorithm.

Here are some papers dealing with SA applications in engineering. There are studies that have applied SA to mechanical problems, e.g. Kincaid & Padula (1990) and Nemec (1998) apply SA to minimize truss distortion and deformations respectively. Dhingra and Bennage (1995) use SA to solve a topological optimisation problem having discrete and continuous variables to optimize truss structures. The study presents numerical results using SA for both single and multi-objective topological optimisation of the structures. Hsancebi and Erbatur (2000) have tried to develop an SA-based efficient algorithm to achieve the simultaneous optimum design of truss type structures with respect to size, shape and topology design

variables when minimizing the weight of the structures. Cagan et al. (1998) have developed a method based on simulated annealing for optimizing components layout in a case that has multiple design goals and inter-component spatial and performance constraints. The proposed algorithm has successfully been used to optimize the placement of components of arbitrary geometry inside an arbitrarily shaped container. Yigit et al. (2004) and Heragu & Alifa (1992) deal with layout problem. In the latter paper, the results from applying three algorithms (a modified penalty algorithm, SA algorithm and a hybrid SA algorithm) on layout problems with facilities of unequal area and multi-row layout problems are presented. The hybrid simulated annealing algorithm appears to produce better quality solutions than the other two methods but requires more computation time than SA. Swarnkar and Tiwari (2004) use a hybrid SA/Tabu search method for solving a machine loading problem of flexible manufacturing systems (FMS). The problem has been formulated as 0–1 integer programming having the bi-criterion objectives of minimizing system unbalance and maximizing throughput in the presence of technological constraints such as available machining time and tool slots. The proposed methodology has been tested on ten standard problems and the results have been compared with those obtained from some of the existing heuristic algorithms. Allwright & Carpenter (1989) describe and analyze an implementation of SA to find a good solution to the travelling salesman problem (TSP). Sadeh and Nakakuki (1996), Sadeh et al. (1997), and Khator et al. (1997) are among the researchers who use SA for job-shop scheduling. Sadeh and Nakukaki (1996) present an SA search procedure to solve job-shop scheduling problems subject to tardiness and inventory costs simultaneously; the procedure is shown to significantly increase schedule quality. Three different variations of the procedure have been considered by them. One of these variations is shown to yield significant reductions in computation time, especially on the problems where search is more likely to get trapped in local minima.

Some researches have combined SA and other techniques for dealing with optimisation problems; e.g. Shalaby et al. (2003) combines a random search (SA) and a direct search (Simplex method) to solve the topology optimisation problem of a mechanism; they report a great saving in optimisation time compared to old methods. Nwana et al. (2005) apply hybrid SA/branch and bound algorithm to mixed 0–1 linear programming; their results demonstrate the effectiveness of using the hybrid approach.

1.2.4 Safe Pile Driving

A literature survey on the prediction of safe sea-states for driving piles resulted in no useful results. Contacting 3 consulting companies in the field of maritime structure designs led

to the paper published by Djahansouzi and Yasseri (1994). Djahansouzi and Yasseri (1994) analyse the dynamic response of the finite element model of a proposed vertical pile in its installation configuration with a view to identifying the possible safe sea-states for undertaking the pile operation and estimating the associated extreme value response parameters. It might be concluded that there are very few papers on the prediction of safe sea-state for offshore activities such as pile driving and therefore the part of this PhD thesis dealing with predicting safe sea-state for pile driving could also be considered a relatively new work.

1.2.5 Extreme Value Analysis and Extreme Wave Height Analysis

Although the extreme value theory is as old as nearly one century, but there has been a great interest over the past four decades on the application of the theory in many fields including ocean engineering. The work of Fisher & Tippet (1928) proposes three limiting distributions in a theorem for extreme values regardless of the parent distribution. However Jenkins (1955) derived a distribution called GEV and showed that the three distributions proposed by Fisher and Tippet could be expressed by it. Pickands (1975) showed that, if the parent distribution of the data falls within the domain of one of the extreme value distributions, the excess values of the data above a threshold level follow asymptotically a generalized Pareto distribution. Pickand's work followed by Smith (1984, 1989) and Davison & Smith (1990) could be regarded as the start point of peaks over threshold (POT) method in extreme value analysis which in turn extended ideas presented by Todorovic and Zelehansic (1970) and NERC (1975). Carter and Challenor (1983) propose to use the Weibull distribution (Weibull, 1951) besides the Fisher-Tippets distributions in extreme value analysis. Coles' (2001) book covers the fundamentals of extreme value theory and Castillo (1988) deals with this theory in engineering.

There are many references in the literature written on extreme wave analysis. Muir & El-Sharawi (1980) overviews excellently the methods presented by researchers for the calculation of extreme wave heights. Although the paper is rather old but the methods described there, are still being used. Barltrop (1998) and Goda (2000) also describe the methods of extreme wave analysis in some detail, and Anderson et al. (2001) deal with the analysis in full detail. Chapter 6 of this thesis is devoted to extreme wave analysis and the following paragraph lists some works that have applied extreme value methods.

Ferreira & Guedes-Soares (1998) apply POT method to estimate the extreme wave height of a region in Portugal and suggest a way for the application of the method to the calculation of extreme stresses of a structure. Dawson (2000) applies extreme value theory in

predicting the maximum crest amplitude in runs of ocean waves using data from extensive computer simulations of random linear waves. Lavenda & Cipollone (2000) use extreme value distributions to estimate the extreme values of such variables as the energy and magnitude of aftershocks in earthquake. Harris (2001) uses some methods for calculating the extreme of wind speed. The methods are the obsolete classical Gumbel method, Gumbel–Lieblein BLUE method, the Modified Gumbel method and the method of Independent Storms. The tests indicate that where sufficiently continuous data is available, the Method of Independent Storms should be used; otherwise if only annual maxima are available, Gumbel–Lieblein BLUE gives the best results for small data sets. Dunne & Ghanbari (2001) show that very accurate extreme exceedance probabilities associated with experimental measurements of highly non-linear beam vibrations could be obtained from relatively small amounts of measured data. Sanil and Deo (2004) estimate extreme value of H_S taking into account the directions of waves approaching a site unlike the conventional procedure; the paper reports that the new procedure has come up with a reduced return value. Li et al. (2005) use a POT method to model daily rainfall above a given threshold to calculate extreme rainfall in a dry region in Australia. Caires & Sterl (2005) estimate the wind speed and H_S extreme values using POT method and asses the effect of the space and time variability of H_S and wind speed on the prediction of their extremes.

1.3 Research Objectives

The thesis has followed three objectives:

- *Simulation and prediction of sea-state parameters by presenting a novel ANN-based method*

In Chapter 4 a novel approach to wave characteristics simulation has been presented; i.e. a combined application of neural networks, statistics and simulated annealing for real time on-line forecasting of the H_S and T_z of wind waves in a region in the North East Pacific. In this approach the parameters of a conditional distribution related to the H_S or T_z of a desired future sea-state, given its preceding 24-hour history of 3-hourly sea-states, are estimated from the outputs of a set of 3-layered ANNs trained by SA. Then the H_S or T_z is generated as a random variate from its conditional distribution. If this is done for a largish number of times, the most probable value or the mean of the generated values could be regarded as the forecast for the wave characteristic, though the mean value proved to be the better forecast. The H_S or the T_z

of a future sea-state could predicted for 3 hours, 6 hours, one week, 1 month, etc from the starting date; but the closer the time the better is the forecast.

- *Determining if the predicted sea-state is safe or otherwise for driving an offshore pile modeled by* FEM *and analysed by* ABAQUS/AQUA

In this part vertical piles have been studied with the goal of identifying the range of sea-states suitable for the safe pile driving operation. Pile configuration, including the non-linear foundation and the gap between the pile and the pile sleeve shims have been modeled using FE facilities within ABAQUS. Dynamic analyses of the system for various sea-states modeled as a combination of several regular waves components, have been performed. Repeating the above procedure has generated a table of safe and unsafe sea-states. If the prediction is repeated N times of which n times proved to be safe, then it could be said that the predicted sea-state is safe with a probability of $100 \times \frac{n}{N}$%.

- *Extreme wave analysis using the observed and simulated H_s's by conventional and two proposed methods*

Extreme wave analysis involves calculating the H_s return value (extreme wave height), which plays a significant role in determining the *design wave* when designing offshore structures. In this thesis two new methods are presented for the calculation of H_s return value. The first method uses the existing maxima method in an unconventional manner; i.e. the annual maxima of nearly 400-cetury simulated data are extracted; the empirical cumulative distribution of the maxima is developed and the return value is calculated from the empirical distribution without the need to use Fisher-Tippet theorem which have some requirements that do not hold for ocean waves characteristics. The other method calculates the return value from a mathematical model presented by some researchers for the behaviour of H_s in a region in the North East Pacific.

1.4 Outline of the Thesis

Chapter 1 deals with the literature survey after an introduction; it also describes the objectives of the present work. **Chapter 2** introduces ANNs. **Chapter 3** discusses simulated annealing in some details. **Chapter 4** presents a new method based on ANNs, statistics and SA for on-line prediction of sea-states characteristics. **Chapter 5** reports the results of a study into the behavior of vertical pile stickup. Pile configuration including the non-linear

foundation and the gap between the pile and its sleeve shims have been modeled using FEM. The final output of the chapter expresses the predicted future sea-state as being safe or unsafe for pile driving with a degree of certainty. **Chapter 6** is concerned with the calculation of the wave height extreme values. **Chapter 7** presents the conclusions and the recommendations for future work.

References (in alphabetical order)

ABAQUS 6.4 User's Manual, 2004.
Hobbit, Karlsson & Sorensen Inc.

Agrawal, J. D., and Deo, M. C., 2002.
On-line wave prediction.
Marine Structures, 15(1), pp 57-74.

Agrawal, J. D., and Deo, M. C., 2004.
Wave parameter estimation using neural networks
Mraine Structure, 17(7), pp 536-550.

Allwright, J. R. A., and Carpenter D. B., 1989.
A distributed implementation of simulated annealing for the travelling salesman problem.
Parallel Computing, 10 pp 335-338.

Anderson, C. W., Carter, D. J. T., & Cotton, P. D., 2001.
Wave Climate Variability and Impact on offshore Design Extreme.
Report prepared for Shell International and the Organization of Oil & Gas Producers.
http://info.ogp.org.uk/metocean/schedules.html.

Athanassoulis, G. A., and Stefanakos, C. N., 1995.
A nonstationary stochastic model for long-term time series of significant wave height.
Journal of Geophysical Research, 100(C8), pp 16149-62.

Barltrop, N. D. P., 1998.
Floating Structures: A Guide for design and analysis.
Oilfield Publication Ltd, ISBN: 1 870553 357.

Ben, C., and Gerwick, Jr., 2000.
Construction of Marine and offshore Structures.
2nd edition, CRC Press, New York, ISBN: 0-8493-7485-5.

Bishop, C. T., Donelan, M. A., and Kahma, K. K., 1992.
Shore protection manual's wave prediction reviewed.
J. Coastal Eng., 17, pp 25-45.

Cagan, J., Degentesh, D., and Yin, S., 1998.
A simulated annealing-based algorithm using hierarchical models for general three-dimensional component layout.
Computer-Aided Design, 30(10), pp 781-790.

Caires, S., and Sterl, A., 2005.
100-year return value estimates for ocean wind speed and significant wave height from the ERA-40 data.
J. Climate 18 (7), pp 1032-1048.

Carter, D. J. T., and Challenor, P. G., 1981.
Estimating return values of environmental Parameters.
Quart. J. Royal Meteor. Society 107 pp 259-266.

Carter, D. J. T., and Challenor, P. G., 1983.
Application of extreme value analysis to Weibull distribution.
Quart. R. Met. Soc. 109 pp 429-433.

Castillo, E., 1988.
Extreme Value Theory in Engineering.
Academic Press, S. Diego.

Chan, W. T., Chow, Y. K., and Liu, L. F., 1995.
Neural network: An alternative to pile driving formulas.
J. Computers and Geotechnics, 17 pp135-156.

Coastal Engineering Manual, 2003.
Chapter II-2, Meteorology and wave Climate.
Engineer manual 1110-2-1100 US Army corps of Engineers, Washington DC.

Coles, S., 2001.
An introduction to statistical modelling of extreme values.
Springer Verlag, London.

Cunha, C., and Guedes Soares, C., 1999.
On the choice of data transformation for modelling time series of significant wave height.
Ocean Engineering, 26(6), pp 489–506.

Dawson, T. H., 2000.
Maximum wave crests in heavy seas.
Jr. of Offshore Mechanics and Arctic Eng'g. ASME Trans., 122(3), pp 222-224.

Deo, M. C., Jha, A., Chaphekar, A. S., and Ravikant, K., 2001.
Neural networks for wave forecasting.
Ocean Engineering, 28(7), pp 889-898.

Davison,A.C. and Smith, R.L., 1990.
Models for exceedances over high thresholds.
J.R. Statis. Soc. B52, pp 393-442.

Dhingra, A. K., and Bennage, W. A., 1995.
Topological Optimization Truss Structures using Simulated Annealing.
Engineering Optimization, 24, pp 239-259.

Djahansouzi, B., and Yasseri, S. F., 1994.
Study of offshore pile stick-up.
Proc. 9th UK ABAQUS User Group Conference, Oxford, 1994.

Drexel, A., 1988.
A Simulated Annealing Approach to the Multi-constraint Zero-One Knapsack Problem.
Computing, 40, pp1-8.

Dunne, J. F., and Ghanbari, M., 2001.
Efficient extreme value prediction for nonlinear beam vibrations using measured random response histories.
Nonlinear Dynamics 24, pp 71-101.

Fisher, R. A., and Tippett, L. H. C., 1928.
Limiting Forms of the frequency by distribution of the Largest or Smallest Members of a Sample.
Proc. Cambridge, Phil Soc 24, 180-190.

Ferreira, J. A., Guedes-Soares, C., 1998.
An Application of the Peaks Over Threshold Method to Predict Extremes of Significant Wave Height.
Jr. of Offshore Mechanics and Arctic Engineering, 120, pp165-176.

Goda, Y., 2000.
Random Seas and Design of Maritime Structures.
2nd edition, World Scientific, Singapore.

Garibaldi, J. M., and Ifeachor, E. C., 1999.
Application of Simulate Annealing Fuzzy Model Tuning of Umbilical Cord Acid-Base. Interpretation, IEEE Trans Fuzzy Syst. 7, pp 72-83.

Gunther, H., Rosenthal, W., Weare, T. J., Worthingtown, B. A., Hasselmann, K., and Ewing, J.A., 1979.
A hybrid parametrical wave prediction model.
Jr. Geophysical Research 84(C9) pp 5727 -5738.

Golding, B., 1983.
A wave prediction system for real time sea-state forecasting.
Q. J. R. Met. Soc. 109(460) pp 393-416.

Harris, R. I., 2001.
The accuracy of design values predicted from extreme value analysis.
Jr. of Wind Engineering & Industrial Aerodynamics 89(2), pp153-164.

Hasselmann, K., Barnett, T. P., Bows, E., Carlson, H., Cartwright, E., Engke, K., Ewing, J. A., Gienapp, H., Hasselmann, D. E., Kruseman, P., Meerburg, A., Müller, P., Olbers, D. J., Richter, K., Sell, W., and Walden, H., 1973.
Measurements of Wind-Wave Growth and Swell Decay during the Joint North Sea Wave Project (JONSWAP). Deut. Hydrogr. A., Suppl A, 8, No. 12.

Heragu, S. S., and Alfa, A. S., 1992.
Experimental analysis of simulated annealing based algorithms for the layout problem.
European Journal of Operational Research, 1992, 57(2), pp 190-202.

Hsancebi, O., and Erbatur, F., 2000.
Layout optimization of trusses using simulated annealing.
2nd Inter. Conf. on Engineering Computational Technology, 2000, pp 175-190.

Huang, W., and Foo, S., 2002.
Neural network modeling of salinity variation in Apalachicola River.
Water research, 36(1), pp 356-362.

Jenkins, A. F., 1955.
The frequency distribution of annual maximum (minimum) values of meteorological Elements.
J R Met. Soc., 81, pp158-171.

Kazeminezhad, M. H., Etemad-shhidi, A., and Mousavi, S. J., 2004.
Appilcation of Simplified Methods in the Prediction of Wave Parameters in Neck.
Proceedings 6th Int. Conf. on Coats Ports and Marine Structures (ICOMPAS) 29 Nov-2 Dec 2004, Tehran, Iran.

Khator, S. K., Candido, M., and Barcia, R., 1997.
Applying Simulated Annealing to Real Job Shop Scheduling Problems.
2nd Annual Int. Conf. November 1997,
Proc. on Industrial Engineering Application and Practice, pp 1211-16.

Kinsman, B., 1965.
Wind Waves.
Prentice Hall, Englewood Cliffs, N. J.

Kincaid, R., and Padula, S., 1990.
Minimizing distortion and internal forces in truss structures by simulated annealing.
AIAA/ASME Structures, Structural Dynamics and Materials Conf. 1990, p 327.

Komen, G. J. L., Cavaleri, M., Donelan, K., Hasselmann, S., Hasselmann, P. A., & Janssen, E. M., 1994.
Dynamics and modelling of ocean waves.
Cambridge University Press, Cambridge, 532 pp.

Lavenda, B. H., and Cipollone, E., 2000.
Extreme value statistics and thermodynamics of earthquake: aftershock sequences.
Annali di geofisica, 43, pp 967-982.

Lee, T. L., and Jeng, D.-S., 2002.
Artificial neural networks for tide forecasting.
Ocean Engineering, 29(9), pp 1003-1022.

Li, Y., Cai, W., and Campbell, E. P., 2005.
Statistical modeling of extreme rainfall in southwest Western Australia.
Jr. Climate, 18(6), pp 852-863.

Lin, W., Sanford, L.P., and Suttles, S.E., 2002.
Wave measurement and modeling in Chesapeake Bay
Continental Shelf Research, 22 (18-19), pp 2673-2686.

Liu, P. C., Schwab, D. J., and Jenson, R. E., 2002.
Has wind-wave modeling reached its limit?
Ocean Engineering, 29 pp 81-86.

Makarynskyy, O., 2005.
Artificial neural networks for wave tracking, retrieval and prediction.
Pacific Oceanography, 3(1), 21-30.

Makarynskyy, O., Pires-Silva, A. A., Makarynska, D., and Ventura-Soares, C., 2002.
Artificial Neural Networks in the forecasting of wave Parameter.
International workshop on wave Hindcasting & Forecasting
B. Alberta Canada Oct 21-25 2002.

Makarynskyy, O., Makarynska, D., Kuhn, M., and Featherstone, W. E., 2005a.
Using artificial neural networks to estimate sea level in continental and island coastal environments.
Hydrodynamics IV: Theory and Applications, L. Cheng and K.Yeow (eds.),
Taylor & Francis Group, London, pp 451-457.

Makarynskyy, O., Pires-Silva, A. A., Makarynska, D., and Ventura-Soares, C., 2005b.
Artificial neural networks in wave predictions at the west coast of Portugal.
Computers & Geosciences, 31 (4), pp 415-424.

Makarynskyy, O., Makarynska, D., Rusu, E., and Gavrilov, A., 2005c.
Filling gaps in wave records with artificial neural networks.
Proc. Maritime Transportation and Exploitation of Ocean and Coastal Resources,
IMAM 2005, 26–30 September 2005, Lisbon, Portugal.

Mangal, L., Idichandy, V.G., and Ganapathy, C., 1994.
Neural Networks in damage detection of offshore structures.
Indian National Conference on Harbour and Ocean Engineering,
INCHOE-94, Vol.1, pp. B21-30, Awarded certificate of merit, 1994.

Mitsuyasu, H., 1969.
On the growth of the spectrum of wind-generated waves.
2. Rep. Res. Inst. Appl. Mech., Kyushu Univ. 17, 235-243.

Mangal, L., Idichandy, V.G., and Ganapathy, C., 1996.
ART based multiple Neural Networks for monitoring of offshore structures.
Applied Ocean Research 18(2-3), 137-143, 1996.

Monroy, L. I. G., and Cordora, A., 2000.
Optimization of Energy Supply Systems with Simulated Annealing: Continuos and Discrete Descriptions. Physica A, 284(2000), pp 433-477,
www.elsevier.com/locate/physa.

Muir, L. R., and El-Sharaawi, A. H., 1986.
On the Calculation of Extreme Wave Heights: A Review.
Ocean Engineering, 13(1) pp 93-118.

Nemec, L., 1998.
Solution of truss deformations by Simulated Annealing Method (in Czech).
Journal of Mechanical Engineering, 49(5), pp 301-306.

NESS, 1997.
Overview of the work carried out during the North European Storm Study: The hind-casting North Sea metocean parameters.
NUG/Matsu, UK.

NERC, 1975.
Flood Studies Report.
London: Natural Environmental Research Council.

Nwana, V., Darby-Dowman, K., and Mitra, G., 2005.
A co-operative parallel heuristic for mixed zero–one linear programming: Combining simulated annealing with branch and bound.
European Journal of Operational Research, 164(1) pp 12-23.

Pickands, J., 1975.
Statistical inference using order statistics.
Annals of Statistics, 3, 119-131.

Rajasekaran, S., Thiruvenkatasamy, K., and Lee, T., 2005.
Tidal level forecasting using functional and sequential learning neural networks.
Applied Mathematical Modelling, In Press, Corrected Proof, Available online.

Roske, F., 1997.
Sea level forecasts using neural networks.
German J. of Hydrography, 49(1), 71-99.

Smith, R.L., 1984.
Threshold methods for sample extremes.
In Tigo de Olivera, J., editor, *Statistical extremes and application*, pp 621-38.
Dordrecht: Reidel.

Smith, R.L.,1989.
Extreme value analysis of environmental time series: an application to trend detection in ground level ozone.
Statis. Sci., 4, pp 367-393.

Toparli, M., Sahin, S., Ozkaya, E., and Sasaki, S., 2003.
Residual thermal stress analysis in cylindrical steel bars using finite element method and artificial neural networks.
Computers & Structures, 80(23) pp 1763-1770.

Tsai, C. P., and Lee, T. L., 1999.
Back-propagation neural network in tidal-level forecasting.
Jr. Waterway, Port, Coastal and Ocean Eng'ng., ASCE, 125(4), 195-202.

Tsai, C. P., Hsu, J.R., and Pan, K. L., 2000.
Prediction of storm-built beach profile parameters using neural network.
Proc. 27th Int. Conf Coastal Engineering, ASCE V4 3048-3061.

Thirumalaiah, K., and Deo, M. C., 2000.
Hydrological Forecasting Using Neural Networks.
Jr of Hydrologic Engineering ASCE, 5(2), pp 180-189.

Todorovic,P. & Zelehansic, E. 1970.
A stochastic model for flood analysis.
Wat. Resour. Res., 6, pp1641-1648.

Sadeh, N., and Nakakuki, Y., 1996.
Focused Simulated Annealing Search: An Application to Job Shop Scheduling.
Technical Report, The Robotics Institute, Carnegie Mellon University,
Pittsburgh, PA; USA 1994.
Annals of Operations Research, 60, pp 77-103.

Sadeh, N. M., Nakakuki, Y., and Thangiah, S. R., 1997.
Learning to Recognize (Un) Promising Simulated Annealing Runs: Efficient Search Procedures for Job Shop Scheduling and Vehicle Routing.
Annals of Operations Research, 75, pp 189-208.

Romero, S., and Pamukcu, S., 1996.
Using Artificial Neural Networks to Predict Shear Modulus of Granular Offshore Sediments.
Proc 6th Int. Offshore and Polar Engineering Conf., v1 pp431-437.
Los Angeles, C.A. USA.

Sanil Kumar, V. and Deo, M. C., 2004.
Design wave estimation considering directional distribution of waves.
Ocean Engineering, 31(17-18), pp 2343-2352.

Salamon, P., Sibani, P., and Frost, R., 2002.
Facts, Conjectures and Improvements for Simulated Annealing.
Society for Industrial and Applied Mathematics ISBN: 0898715083.

Shalaby, M., Hegazi, H. A., Nassef, A. O., and Metwalli, S. M., 2003.
Topology optimization of a compliant gripper using hybrid simulated annealing and direct search.
Proc. ASME Design Engineering Technical Conf. Chicago Sept 2-6, 2003.
V 2 A, pp 641-648.

Shore Protection Manual, 2nd Vol, 1984.
Coastal engineering Research Center, Army Corps of Engineers, waterways.
Experiment Station, Washington DC.

Swarnkar, R., and Tiwari, M. K., 2004.
Modeling machine-loading problem of FMSs and its solution methodology using a hybrid tabu search and simulated annealing-based heuristic approach.
Robotics and Computer-Integrated Manufacturing, 20(3), pp199-209.

SWAMP Group, 1985.
Sea wave modeling project (SWAMP), an inter-comparison study of wind wave prediction models, Part 1 principal results and conclusions, ocean wave modeling. Plenum Press.

Ultsch, A., and Röske, F., 2002.
Self-organizing feature maps predicting sea levels.
Information Sciences, 144(1-4) pp 91-125.

van Gent, M. R. A., and van den Boogaard, H. F. P., 1998.
Neural Network Modelling of Forces on Vertical Structures.
Proc. 26th Int. Conf. on Coastal Engineering (ICCE '98), B. L. Edge, ed.
ASCE, V.2, pp 2096-2109.

WAMDI group, 1988.
The WAM –model, A third generation ocean wave prediction model.
Jr. of Oceanography, 18(12), pp 1775-1810.

Weibull, W., 1951.
A statistical distribution of wide applicability.
Jr. Applied Mechanics, vol. 18. pp 293-297

Wills, A. S., 2000.
A New Protocol for the Determination of Magnetic Structures Using Simulated Annealing and Representational Analysis (SARA).
Physica B 276-278, (2000) 680-681, www.elsevier.com/locate/physb.

WMO, 1998.
Guide to wave Analysis and forecasting.
2nd edition, World Meteorological Org. Publication no 702 Geneva.
www.wmo.int/web/aom/marprog/ Wordpdfs/WMO%20No%20702/WMO702.pdf.

WMO, 1990.
Proc. of the commission for marine meteorology.
Technical conference on ocean waves, report no. 24 wmo /td-no 359.

Yigit, V., Aydin, M. E., and Turkbey, O., 2004.
Evolutionary simulated annealing algorithms for uncapacitated facility location problems, In (Ed. Parmee, I.C.): *Adaptive Computing in Design and Manufacture,* VI Volume Package, 391 p.
ISBN: 1-85233-829-6, Springer-Verlag, NY.

Ziemianski, L., 2003.
Hybrid neural network/finite element modelling of wave propagation in infinite domains.
Computers & Structures, 81(8-11) pp 1099-1109.

Chapter 2
An Introduction to Artificial Neural Networks

Nomenclature

a	output of the neuron
ANN(s)	artificial neural network(s)
b	neuron bias
BP	back-propagation algorithm for learning
f_{tr}	the neuron transfer function
H	number of neurons of the hidden layer in a 2-layer network
k_{i}	number of the neurons of a network layer
k_{in}	the number of the neuron of the input layer
k_{OUT}	the number of the neuron of the output layer
MAE	mean absolute error
MLP	multi-layer perceptron
MSE	mean of squared errors
MSEREG	mean of squared errors with regularization
PU	processing unit; neuron
R	number of inputs to the network
S	number of neuron in a layer
SOFM	self-organizing feature neural networks
SSE	sum of squared errors
v	the summing junction output or the input to the transfer function of the neuron
w_0	bias of the neuron
$\mathrm{w},\mathrm{w}_1,\ldots,\mathrm{w}_R$	neuron synaptic weights
$x, x_1,\ldots, x_R$	inputs of artificial neurons

2.1 Introduction

An artificial neural network (ANN) is an extremely simplified biological neural network of human body. Complex interconnected biological neural networks have some 10^{11} neurons each connected to 10,000 others. A biological neuron, which has the complexity of a micro-

processor, is shown in Fig. 2.1. Dendrites are actually the inputs of the neuron and axons are its outputs.

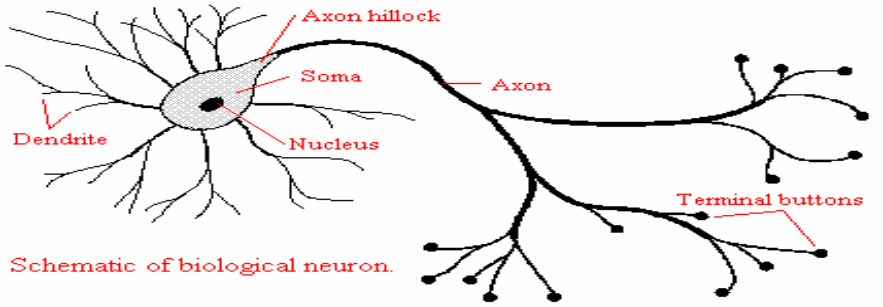


Fig. 2.1 A biological neuron.

Scientists have only just begun to understand how biological neural networks operate. It is generally understood that all biological neural functions, including memory, are stored in the neurons and the connections between them. Learning is viewed as the establishment of new connections between neurons or the modification of existing connections. The researchers and scientists have constructed a small set of simple artificial neurons to serve a useful function after being trained. Fig 2.2 shows an artificial neuron. Networks comprising these artificial neurons do not have a fraction of the power of the human brain but they can be trained to perform useful functions.

2.2 ANNs History

This part has mainly been adopted from Hagan (1996). Some of the background work for the field of artificial neural networks occurred in the late 19th and early 20th centuries. The modern view of the ANNs began in the 1940's. In 1943 McCulluch and Pitts showed networks of artificial neurons could compute any arithmetic or logical function (McCulluch & Pitts, 1943). In 1949 Donnald Hebb, following McCulluch and Pitts proposed a mechanism for learning in biological neurons (Hebb, 1949). In the late 50's and 60's first practical ANN i.e. the *perceptron* network and the associated learning rule by Frank Rosenblatt was invented (Rosenblatt, 1958). Bernard Widrow and Ted Hoff introduced a new learning algorithm and used it to train adaptive linear ANN (Widrow & Hoff, 1960). These basic perceptrons could solve only limited class of problems. Later some important work like that of Teuvo Kohonen (Kohonen, 1972) did continue during 1970's, however for a decade neural network research was suspended. This was mainly because of the belief that the research has reached a dead end and no powerful digital computers were available. During 1980's the researches in the field of

ANNs increased dramatically. Important new concepts were introduced, among which two concepts are most responsible for the rebirth of ANNs. The first was the use of statistical mechanics to explain the operation of a certain class of recurrent network described by John Hopfield (Hopfiled, 1982). The second was the back-propagation algorithm for training multi-layer perceptron (MLP) networks discovered by several researchers independently. The influential among them was that of Rumelhart & McClelland (1986). From mid 1980's until now thousands of papers have been published and ANN has been spreading with new theoretical and practical applications. The applications, the money invested, software, hardware related to ANN have been growing. ANNs have been applied to an increasing number of real-world coplicated problems of including those of ocean engineering. Chapter 1 of this thesis lists some papers introducing the applications of neural networks. ANNs have been covered by many authors including Haykin (1999), Hagan (1996), and Fausett (1994).

2.3 Mathematical Model of Neurons

The artificial neurons may have single input or multiple inputs as described below.

2.3.1 Single–Input Neuron

Figure 2.2, which has been extracted from Makarynskyy et al. (2002) with some modification, shows a single-input artificial neuron in its simplified mathematical model.

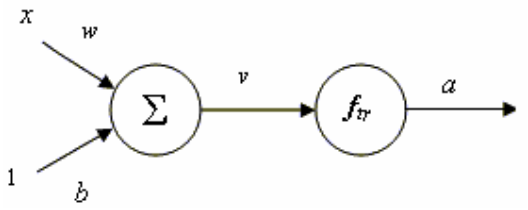


Fig. 2.2 An artificial neuron with a single scalar input and a bias.

where

x	input
w	synaptic weight
b	bias
a	output of the neuron
v	output of the summing junction i.e. Σ
f_{tr}	the neuron transfer function.

As Fig. 2.2 shows the neuron has two inputs. The first is *x,* a scalar input present in all neurons and the second input is 1, which is not present in a number of applications. The scalar input *x* is multiplied by the scalar weight *w* to form *wx.* The other input i.e. 1 is multiplied by a bias *b* and then passed into the summing junction. The output of the summing junction i.e. *v* goes into the transfer function f_{tr} and produces an output *a*. The input and output of the neuron satisfy the following equation:

$$a = f_{tr}(v) = f_{tr}(w \times x + b), \qquad (1)$$

As an example let w=4, x=2, b=-0.5, and $f_{tr}(v) = \dfrac{1}{1+e^{-v}}$; then

$$a = f_{tr}(4 \times 2 - 0.5) = f_{tr}(7.5) = \frac{1}{1+e^{-7.5}} = 0.9994.$$

2.3.2 Multiple-Input Neuron

Neurons usually have more than one input with each input having its own scalar weight. The inputs are fed into the summing junction. A typical multiple–input neuron, with a transfer (activation) function called symmetrical *Hard Limit*, is shown in Fig 2.3:

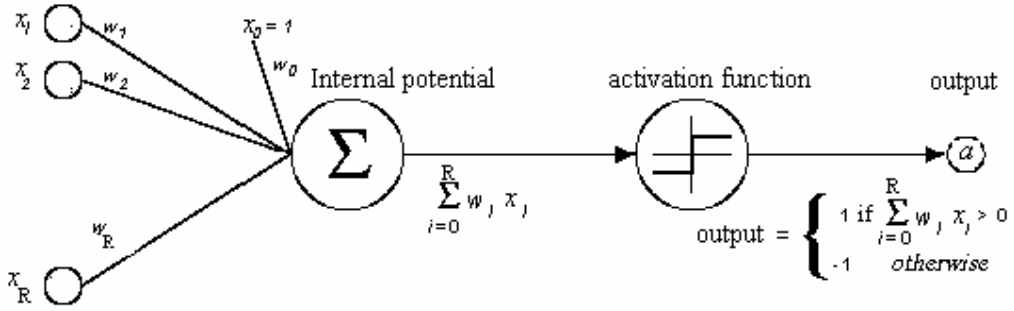


Fig 2.3 A multiple-input neuron.

Eq. (2) shows the output of this kind of neuron:

$$a = f_{tr}(v) = f_{tr}(w_1 x_1 + \cdots + w_R x_R + w_0), \qquad (2)$$

where

R	number of inputs
$x_1, \ldots, x_R$	inputs
$w_1, \ldots, w_R$	synaptic weights
$w_0 = b$	bias.

2.4 A Layer of Neurons

A single–layer network of S neurons is shown in Fig. 2.4:

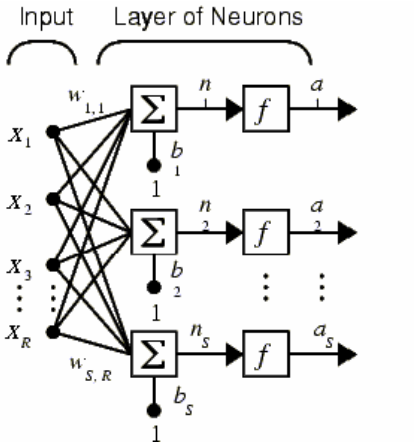


Fig 2.4 A layer of *S* neurons (after MATLAB\Help\toolbox\nnet).

The layer shown in Fig. 2.4 has R input elements and S neurons. Note that a layer is not constrained to have the number of its inputs equal to the number of its neurons (i.e., *R* is not necessarily equal to *S*). Note that each of the R inputs is connected to each the neurons and that the weight matrix has S rows and R columns.

The number of nodes in the hidden layer of the one-hidden-layer networks implemented in the first stages of this thesis was computed using the following empirical expression (Huang and Foo, 2002)

$$H = 2 \times R + 1, \tag{3}$$

where

H is the number of hidden layer nodes,

R is the number of input nodes.

2.5 Transfer (Activation) Function

The actual output of the neuron depends on the transfer or activation functions of the system. Hard-limiter (*hardlim*), symmetrical hard limit (*hardlims*), linear (*purelin*), log-sigmoid (*logsig*), step, and tan-hyperbolic are some transfer functions. Needless to say that each transfer function has its own formula. For example the log-sigmoid (logsig) function which is commonly used in the multi-layer networks has the following formula:

$$a = \frac{1}{1 + e^{-v}}, \tag{4}$$

where

v is the input to the transfer function, and

a is its output, in this case a number in the range 0 to1.

The formulae and the plots of *purelin, tansig,* and *logsig* transfer functions are shown in Fig. 2.5.

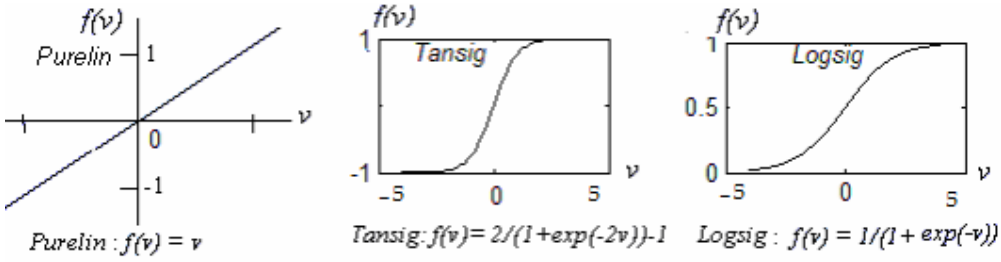


Fig. 2.5 Three different types of transfer functions.

2.6 Network

According to Haykin (1999) a neural network is a massively parallel distributed processor made up of simple processing units (PUs), which has a natural propensity for storing experiential knowledge and making it available for use. It resembles the brain in two respects:

1. Knowledge is acquired by the network from its environment through a learning process.
2. Inter-neuron connection strengths, known as synaptic weights, are used to store the acquired knowledge.

Artificial neural networks are arrangements of a number of richly interconnected nodes or processing neurons or PUs. These nodes are collected together in layers forming a network. Each layer has a weight matrix **W**, a bias vector **b**, and an output vector **a.** To distinguish between the weight matrices, output vectors, etc., for each of these layers in the figure, the number of the layer is appended as a superscript to the variable of interest. This notation has been used in the three-layer network shown in Fig. 2.6.

2.6.1 Classification of Network Layers

Network layers might be classified as input layer, hidden layer and output layer.
The activity of the input units represents the raw information that is fed into the network. The activity of each hidden unit is determined by the activities of the input layer and the weights on the connections between the input and the hidden layers. The behavior of the output layer depends on the activity of the hidden layers and the weights between the hidden and output layers.

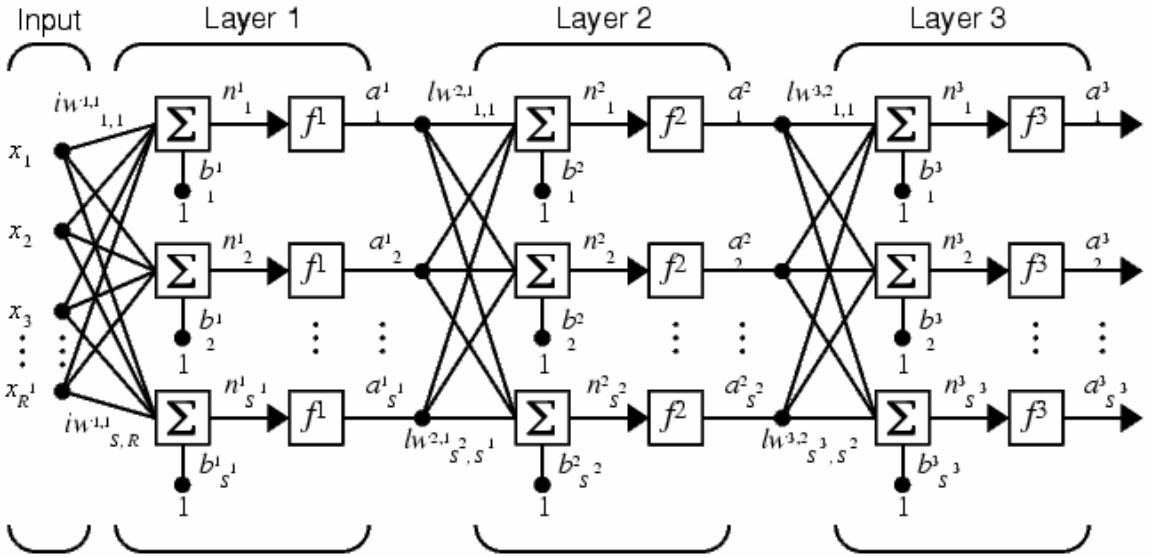


Fig 2.6 A multi-layer network (after MATLAB\help\toolbox\nnet).

2.7 Classification of Artificial Neural Networks

Artificial neural networks can roughly be divided into the following three categories, based on the arrangement of neurons and the connection patterns of the layers:
Feed-forward networks, Feedback networks, and Self-Organizing Maps (Kohnen models).

2.7.1 Feedforward Networks

Feed-forward networks transform sets of input signals into sets of output signals. The desired input-output transformation is usually determined by external supervised adjustment of the system parameters. A feed-forward network is a layered network in which each layer only receives inputs from their previous layers. In other words the input data only flows *forward* from layer to layer through the network. Figure 2.7 shows a typical feed-forward network. In this kind every neuron in a given layer receives inputs from layers below its own (that is from layers nearer the input layer) and sends output to the layers above its own i.e. to the layers nearer to the output layer. For such networks, given the input vector to the neurons in the input

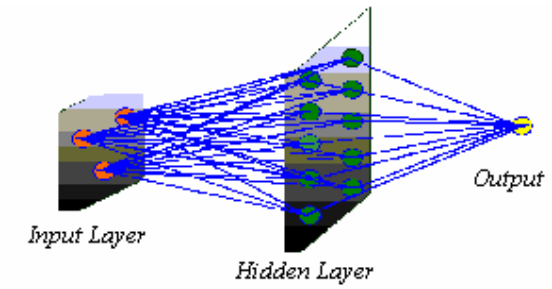


Fig. 2.7 A feed-forward network.

layer, the output vector is computed by a succession of forward passes, which in turn compute the intermediate output vectors of each layer using the previously computed signal values in the earlier layers. "In a feed-forward network, the input quantities are fed to input nodes, which in turn pass them on the hidden layer nodes after multiplying by a weight. A hidden layer node adds up the weighted input received from each input node, associates it with a bias, and then passes the result on through a nonlinear transfer function. The output nodes do the same operation as that of a hidden node. Before its application to any problem, the network is first trained, whereby the target output at each output node is compared with the network output and the difference (error) is minimized by adjusting the weights and biases though training algorithms" (Thirumalaiah & Deo, 2000).

A feed-forward network is a standard type of neural network and might be used for such applications as control systems and pattern identifiers. Perceptron, multilayer perceptrons (MLPs), and radial-basis function networks are three variations of feed-forward networks, though MLPs has been sometimes used as a synonym for general feed-forward networks.

2.7.1.1 Perceptron architecture

Perceptron is the earliest and simplest feed-forward networks having a single layer with a hard-limit transfer function. This network is often trained with the perceptron learning rule (learnrp). The perceptron network consists of a single layer of S neurons connected to *R* inputs through a set of weights $w_{i,j}$ as shown in Fig. 2.8. The network indices *i* and *j* indicate that $w_{i,j}$ is the strength of the connection from the j^{th} input to the i^{th} neuron.

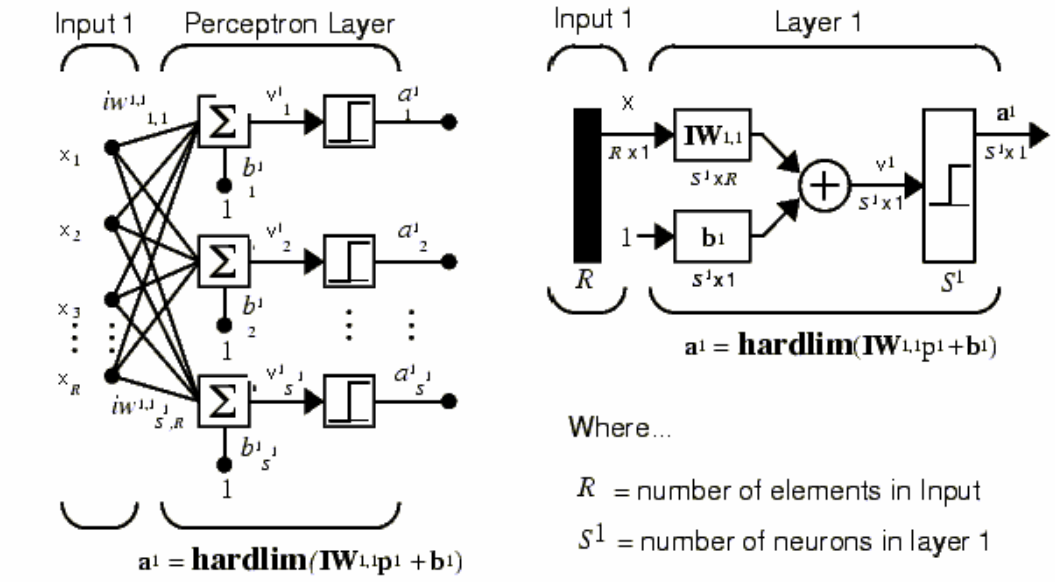


Fig. 2.8 Perceptron network shown in two forms (after MATLAB Help).

The perceptron learning rule is capable of training only a single layer. This restriction places limitations on the computations that a perceptron can perform.

2.7.1.2 Multilayer Perceptrons

Multilayer Perceptrons (MLPs) are feed-forward neural networks trained usually with the standard back-propagation algorithm. Figure 2.9 illustrates such a network. Some times

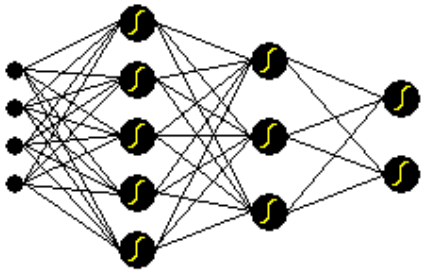


Fig. 2.9 Multilayer perceptrons.

the network architecture is shown as $k_{\mathrm{IN}} \times k_1 \times \ldots \times k_i \times \ldots \times k_{\mathrm{H}} \times k_{\mathrm{OUT}}$,

where

k_{IN} the number of inputs provided for the network,

k_i the number of the neurons of the i^{th} hidden layers, $i = 1, \ldots, H$

k_{OUT} the number of the neuron of the output layer.

2.7.1.3 Radial-basis function networks

The motivation behind radial basis function is the following viewpoint (Haykin,1999): "learning is equivalent to finding a surface in a multidimensional space that provides the best fit to the training data, with the criterion for 'best fit' being measured in some statistical sense. Correspondingly generalisation is equivalent to the use of this multidimensional space to interpolate the test data". A radial basis network is a neural network that can be designed directly by fitting special response elements where they will do the most good. Radial basis networks may require more neurons than standard feed-forward back-propagation networks, but often they can be designed in a fraction of the time it takes to train standard feed-forward networks. They work best when many training vectors are available. Generalized regression neural networks (GRNN) and probabilistic neural networks (PNN) are two variants of radial basis networks (MATLAB Help).

2.7.2 Feedback Networks

So far, input data has only flowed forward from layer to layer through the network; only corrections have flowed back. A feedback network is a network with connections from the output of a layer to its input. The feedback connection can be direct or pass through several layers. In feedback networks, the output information defines the initial activity state of a feedback system, and after state transitions the asymptote final state is identified as the outcome of the computation. The networks known as recurrent networks are feed-back networks. Some variations of recurrent neural networks are Elman, Hopfield networks, Regressive networks, and Jordan-Elman networks. Adaptive Resonance Memories is also a variation of feedback networks.

2.7.3 Self-Organising Maps

An ANN can also be *self-organizing* which means that it can create its own organisation or representation of the information it receives during learning time without any intervention from the outside world. In this third category of ANNs, neighboring cells in a neural network compare in their activities by means of mutual lateral interactions, and develop adaptively into specific detectors of different signal patterns. In this category, learning is called competitive, unsupervised, or self-organizing feature map (SOFM). SOFM, based on competitive learning, is a topology-preserving map (Kohonen, 1984), and can be adjusted to approach the probability distribution of the inputs (Zheng, 1996).

2.8 Learning Algorithms

2.8.1 Definition

The process by which the network weights and biases are adjusted to achieve some desired network behavior is called learning.

2.8.2 Classification of Learning Algorithms

There are many algorithms for training ANNs. Nobody knows exactly how many. New ones (at least variations of the existing ones) are invented every month or week. The main categorization of these methods is the distinction of *supervised* from *unsupervised* learning.

2.8.2.1 Supervised learning

In supervised learning, the learning rule is provided with a set of examples (the training set) of proper network behavior comprising inputs to the network, and their corresponding correct outputs (targets). As the inputs are applied to the network, its outputs are compared to the targets. The learning rule is then used to adjust the weights and biases in order to move the network outputs closer to the targets. In supervised learning changes in the weights and biases are due to the intervention of any external teacher. The teacher typically provides output targets. Supervised ANNs have a teacher i.e. the network is presented with a set of training patterns and corresponding *desired* output patterns. Supervised ANNs, depending on the network, may be self-organizing or not. The perceptron learning rule falls in this supervised learning category; in other words the perceptron is trained with supervised learning.

2.8.2.2 Unsupervised learning

In unsupervised learning, the weights and biases are modified in response to network inputs only. There are no target outputs available. Most of these kind of algorithms perform clustering operations i.e. they categorize the input patterns into a finite number of classes, for instance to identify groups of data. This is especially useful in such applications as vector quantization. Vector quantization is a technique that exploits the underlying structure of input vectors for the purpose of data compression. Specifically, an input space is divided into a number of distinct regions and for each region a reconstruction vector is defined (Haykin, 1999). Changes weights and biases in unsupervised learning are not due to the intervention of any external teacher. Common changes are a function of the current network input vectors, output vectors, and previous weights and biases. In unsupervised ANNs there is no teacher i.e. the network is not asked what to output for each of the input patterns; these ANNs imply that they are self-organizing.

2.8.3 The Implemented Learning Algorithms

Two learning algorithms were used to train the networks of the present study. At the beginning, back-propagation together with *early stopping* was tested for training three sets of networks: one set had seven networks of size $13\times30\times30\times1$. The other set had also seven $13\times45\times20\times1$ networks. The 7 networks of the third set were all of size $13\times50\times15\times1$; but back-propagation was not found rapid enough for the present study and a well known

optimisation algorithm i.e. simulated annealing was used for training the networks. The fundamentals of simulated annealing will be discussed in Chap. 3 but a short description of back-propagation, extracted from MATLAB Help, follows.

2.9 Back-propagation (BP) Algorithm, a brief discussion

Back-propagation (BP) is a well-known algorithm for training neural networks. It has lower memory requirement than most algorithms and usually reaches acceptable error level. It can be used on most types of networks although it is usually appropriate for training MLPs. "The name back-propagation derives from the fact that computations are passed forward from the input to the output layer following which calculated errors are passed forward from the input to the output layer" (Bose & Liang,1996). "Back-propagation was developed by generalizing the Widrow-Hoff learning rule (Delta rule) to multiple-layer networks and non-linear differentiable transfer functions. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function. Networks having biases, a sigmoid layer, and a linear output layer could approximate many functions. Standard back-propagation is a gradient descent algorithm, as is the Widrow-Hoff learning rule, in which the network weights are moved along the negative of the gradient of the performance function. There are a number of variations of the algorithm, based on other standard optimisation techniques, such as conjugate gradient method and Newton method.

The MATLAB command *newff* creates a feed-forward back-propagation network, it needs some information including

a) $R \times 2$ matrix of minimum and maximum values for its R input elements,

b) sizes of all layers,

c) type of the transfer functions of each layer,

d) backpropagation weight/bias learning function, and

e) performance function.

The command returns an N-layer feedforward ANNs. An illustration is given in Sec. 2.9.2

2.9.1 Performance Functions

The performance function is used to calculate network performance during training. This is useful for many algorithms, such as backpropagation, which operate by adjusting network weights and biases to improve performance. The following functions are conventional performance functions:

MAE	mean absolute error
MSE	mean of squared errors
MSEREG	mean of squared errors with regularization
SSE	sum of squared errors

The programmer could also introduce his/her own performance function.

2.9.2 A Typical Backpropagation Feedforward Network

Here is an illustration of how to create and use a back-propagation feed-forward network. The following MATLAB command creates a two-layer (excluding input layer) back-propagation feed-forward network called *net* with a one-element input ranging from -15 to 10, a hidden layer of eight neurons having *tansig* as transfer function(see Fig 2.5 of Chap. 2), and one output neuron with *purelin* transfer function(see Fig. 2.5 of Chap. 2).

net = newff([-15 10],[8 1],{'tansig','purelin'},'trainlm','learngd','mse').

Trainlm, the default of the BP training function, was used as training function, and 'learngd' was chosen as the BP weight/bias learning function, the performance function is MSE which is the default. Then the network is given input vector *x* and the target, and the error is calculated by subtracting the output y from target t, and MSE is calculated.

```
>>x = [-15 -5 0 5 10];
>> t = [0 0 1 1 2];
```

The network performance before training:

```
>>y = sim(net,x)
y =
-0.9194   -1.2470   0.0233   -0.3591    0.5348
>> e = t-y
e =
0.9194    1.2470    0.9767    1.3591    1.4652
>> Performance = mse(e)
Performance = 1.4697
```

The performance is intended to approach zero by training the network.

The network performance after training:

```
>>[net,tr] = train(net,x,t);
>> y = sim(net,x)
y =
0.0000   -0.0000    1.0000    1.0000    2.0000
```

>> **e** = **t**-**y**

e = 1.0e-015 * -0.1110 0.1665 0.2220 -0.4441 0

>> Performance = mse(**e**)

Performance = 5.7316×10^{-32}.

2.10 Overfitting

One of the problems that occur during network training is *overfitting*. In this case the error on the training set is driven to a very small value, but when new data is presented to the network the error is considerable. The network has memorized the training examples, but it has not learned to generalize to new situations. One method for preventing overfitting and improving generalisation is called early stopping. In this technique the available data is divided into three subsets. The first subset is the training set, which is used for computing the gradient and updating the network weights and biases. The second subset is the validation set. The error on the validation set is monitored during the training process. The validation error will normally decrease during the initial phase of training, as does the training set error. However, when the network begins to overfit the data, the error on the validation set will typically begin to rise. When the validation error increases for a specified number of iterations, the training process is stopped, and the weights and biases at the minimum of the validation error are returned. To use the early stopping function one could refer to MATLAB help, which introduce a sequence of commands .

References (in alphabetical order)

Bose, N. K., and Liang, P., 1996.
Neural network fundamentals with Graphs, Algorithms and Application.
McGraw-Hill, page 156, ISBN :0070066183.

Fausett, L. V., 1994.
Fundamentals of Neural Networks, Architecture, Algorithms & Application.
Printice Hall, Englewood Cliffs, NJ.

Kohonen, T., 1984.
Self-Organization and Associative Memory.
volume 8 of Springer Series in Information Sciences.
Springer, New York, pp 184-189.

Hagan, M. T., Demuth, H. B., and Bearle M., 1996.
Neural network design.
PWS Publishing Company, Boston, MA, ISBN: 0-534-94332-2.

Haykin, S., 1999.
Neural networks: a comprehensive foundation.
Prentice-Hall, Upper Saddle River, N.J. 842pp.

Hebb, D. O., 1949.
The Organization of Behavior.
Wiley, New York.

Huang, W., and Foo, S., 2002.
Neural network modeling of salinity variation in Apalachicola River.
Water research, 36(1), pp 356-362.

Hopfield, J. J., 1982.
Neural networks and physical system with emergent collective computational abilities.
Proc. Natl. Acad. Sci. U S A, April 1982, 79(8): 2554–2558.
http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=346238.

Makarynskyy, O., Pires-Silva, A. A., Makarynska, D., and Ventura-Soares, C., 2002.
Artificial Neural Networks in the forecasting of wave Parameter.
The 7th International workshop on wave Hindcasting & Forecasting.
Banff, Alberta Canada 21st-25th Oct 2002, pp 514-522.

McCulluch, W., and Pitts, W., 1943.
A logical calculus of the ideas immanent in nervous activity.
Bulletin of Mathematical Physics, 5, pp115-133.

Rosenblatt, F., 1958.
The perceptron: A probabilistic model for information storage and organization in the brain.
Psychological Review, 65, pp 386-408.

Rumelhart, D. E., and McClelland, 1986.
Parallel distributed Processing: Explorations in the Microstructure of cognition.
Vol 1 Cambridge MAMIT Press 323, pp. 533-536.

Widrow, B., and Hoff, M. E., 1960.
Adaptive switching circuits 1960 IRE WESCON Convention Record.
New York IRE, pp 96-104.

Zheng, Y., and Greenleaf, J. F., 1996.
The effect of concave & convex weight adjustments on self-organizing Maps.
IEEE Trans. Neural Networks, 7(1), pp 87-96.

Chapter 3
Simulated Annealing

Nomenclature

ANNs	artificial neural networks
ASA	adaptive simulated annealing
DE	differential evolution optimisation technique
E	the system energy
EBSA	ensemble based simulated annealing
k	the Boltzmann constant
GA	genetic algorithm
GD	gradient descent
GSA	generalized simulated annealing
IASA	integer augmented simulated annealing
MLE	maximum likelihood estimation method
OR	operations research
RASA	real-coded augmented simulated annealing
SA	simulated annealing
T	the system temperature
T_i	the current value of parameter *temperature* in SA algorithm
T_{i+1}	the temperature in the next iteration of SA
VFSA	very fast simulated annealing
VFSR	very fast simulated re-annealing
β	a suitably small value used in SA algorithm
ΔC	difference between the cost function of the current and the new configurations in simulated annealing algorithm
ΔE	the change in the system energy from E_2 to E_1 i.e. E_2-E_1

3.1 Introduction

Simulated annealing (SA) is one of the heuristic global optimisation techniques in operations research which has attracted significant attentions as being suitable for dealing with optimisation problems of large scale, especially the ones in which a desired global extremum is hidden among many local extrema of a multivariate function. The simplicity and good results

of this stochastic search algorithm over wide range of problems have made it a highly popular tool, with hundreds of applications in many fields. SA has been described widely in the literature; a literature survey on the algorithm was presented in Chap. 1. Salamon, et al. (2002) is a reference among many for an overview of SA.

Based on the principles of thermodynamics, SA is motivated by an analogy between physical annealing of solids and combinatorial optimisation problems. Because of choosing SA as the training algorithm of the artificial neural networks (ANNs) used in this study, a description of the SA algorithm is followed.

3.2 Physical and Simulated Annealing

3.2.1 Physical Annealing

Physical annealing refers to the process of finding low energy states of a solid by initially giving enough heat to the substance, lowering the temperature slowly, and then spending long time at temperatures close to the freezing point. If a solid is heated past the critical point and then be cooled, the structural properties of the solid will depend on the rate of cooling. If cooled slowly enough, large crystals will be formed. However, if it is quenched i.e. cooled quickly the crystals will contain imperfections.

3.2.2 Simulated Annealing

The motivation for SA stems from Metropolis et al. (1953). Metropolis's algorithm simulated the material as a system of particles. The algorithm simulated a kind of annealing process; i.e. the process of heating the substance to the melting point and cooling back by gradually lowering the temperature of the system until it converges to a steady frozen state. In other words, Metropolis and colleagues introduced a simple algorithm to simulate a collection of atoms at a given temperature. The algorithm works by iteratively proposing changes and either accepting or rejecting each change at each iteration; an atom is given a small random displacement and the resulting change in the energy of the system from say E_1 to E_2 is calculated as such $\Delta E = E_2 - E_1$. If $\Delta E < 0$, the change is accepted; but if $\Delta E > 0$, the change is accepted with the probability of $e^{-\frac{\Delta E}{kT}}$ where T is the temperature and k is a physical constant called the Boltzmann constant[1]. If a large number of iterations are carried out at each temperature, the system will reach thermal equilibrium at the temperature. The probability

[1] From kinetic theory, the average energy of a particle at the absolute temperature τ is proportional to $k\tau$.

distribution of the system follows the Boltzmann distribution where the probability function given in Eq. (1).

Kirkpatrick et al. (1983) suggested that the simulation idea of Metropolis could be used to search the feasible solutions of an optimisation problem, with the objective of converging to an optimal solution. They showed that the Metropolis algorithm could be applied to optimisation problems by mapping the elements of the physical cooling process onto the elements of an optimisation problem as shown in Table 3.1(Dowsland,1995). Kirkpatrick and colleagues used SA to search for a feasible solution space and convergence to an optimal solution.

Table 3.1 Analogy between different characteristics of the Metropolis simulation and an optimisation technique.

Metropolis Simulation	Combinatorial Optimisation
System states	Feasible solutions
Energy	Cost
Change of state	Neighbouring solutions
Temperature	Control parameter
Frozen state	Heuristic solution

3.3 Illustrative Description of Simulated Annealing

Before describing the algorithm, an illustration (Moins, 2002) is followed to illustrate how simulated annealing tries to find the global minimum of a multivariate function. Suppose the lowest valley in a geographical terrain is to be found. Simulated annealing approaches this problem similar to using a bouncing ball that can bounce over mountains from a valley to valley. Search starts at a high "*temperature*", where the temperature is a parameter in simulated annealing that mimics a bouncing ball. The ball can bounce over mountains from the effect of a fast moving particle in a hot object like a hot molten metal, thereby permitting the ball to make very high bounces and being able to bounce over any mountain to access any valley, given enough bounces. As the *temperature* is reduced, the ball cannot bounce very high, and might become trapped in relatively smaller valleys. Imagine that our mountain range is aptly described by a "*cost function*" which associates a cost with each state. The cost function used for the present work is based on a likelihood function of maximum likelihood estimation (MLE) method in statistics comprising of the product of a number of conditional density functions (pdf's) and therefore is a function of the parameters of the pdf's. The parameters are obtained directly or indirectly from the outputs of the ANNs. As it will be pointed out in Sec

3.7 the objective during training was to change the weights and biases of the networks continuously and randomly in order to maximize the cost function.

3.4 Standard Simulated Annealing Algorithm

The flowchart of the standard SA algorithm implemented in this study for training the seven networks whose outputs were used to estimate the 7 parameters of a distribution was adopted from Moins (2002) done at Linkping Institute of Technology in Sweden. Figure 3.1 shows the flowchart.

```
Initialization (Current_solution, temperature)

Calculation of the Current_Cost
LOOP
        New_State
        Calculation of the New_Cost
        IF ΔC  = (Current_Cost - New_Cost )≤ 0 THEN
                Current_State = New_State
         ELSE
                IF Exp (Current_Cost − New_Cost)/temprature >Random (0, 1)
        THEN
                        -- Accept
                        Current_State = New_State
         ELSE
                        -- Reject
        Decrease the temperature
EXIT when STOP_CRITERION
END LOOP
```

Fig.3.1 The standard SA algorithm for maximization problems (after Moins, 2002).

Simulated annealing process starts from a valid solution at a high value for one of its parameters called *temperature*; new states are randomly generated for the problem and the associated cost function is calculated. Bearing in mind that the cost function is calculated for each iteration, various researchers including Burke (1999) have shown that the cost function can be responsible for a large proportion of the algorithm execution time. A new state is randomly chosen and the difference in the cost function i.e. "Current_Cost – New_Cost" is calculated. If the difference is negative, the new state is accepted. This forces the system toward a state corresponding to a local or a possibly a global maximum. However, most large optimisation problems have many local maxima and the optimisation algorithm is likely to be trapped in one of them. To get out of a local maximum, a decrease of the cost function is

accepted with a certain probability based on a probability used in statistical thermodynamics which states that the probability of an increase in energy of magnitude $\Delta E = E_2 - E_1$ at temperature T is given by

$$\Pr(\Delta E) = e^{-\frac{\Delta E}{kT}}, \qquad (1)$$

where k is the Boltzmann constant. In SA algorithm it is usual to drop the constant as this was only introduced into the Eq.(1) to cope with different materials, and the probability of accepting a worse state in maximization problem, p, satisfies the following inequality

$$p = e^{\frac{Current\_Cost - New\_Cost}{Temprature}} > \text{rand}(0,1), \qquad (2)$$

where

rand (0,1) is a randomly generated number from uniform distribution in the interval 0-1.

That is if the probability is greater than rand (0, 1), the new state is accepted despite of being a worse solution. It is most probable that we explore all states of the system including the global optimum. Eq. (3) shows a new expression for p proposed by Tsallis & Stariolo (1996)

$$p = [1 - (1 - q_a)\frac{\Delta C}{\Gamma}]^{\frac{1}{1-q_a}}. \qquad (3)$$

where

q_a a real number parameter that is kept constant,

ΔC the difference between the cost function of the current and new configurations,

Γ a parameter analogous to kT in Eq. (1).

3.4.1 Illustration

Consider Fig 3.2. We would like to find the global minimum by simulated annealing. The current solution is at state X(i), the new state k_1 is accepted, but the state k_2 is only accepted with a certain probability. The probability of accepting a worse state is high at the beginning and decreases as the temperature decreases. For each temperature a number of new states must be tried before the temperature is reduced typically by 10%. The algorithm will find, under certain condition, the global optimum without getting stuck in a local one.

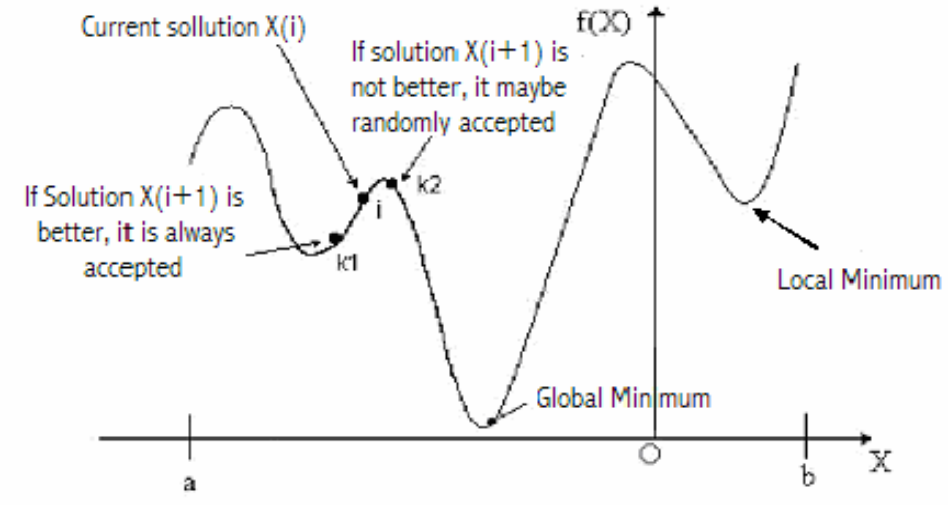


Fig.3.2 Selection of a new state in SA (after Moins, 2002).

3.4.2 The Cooling Process

The cooling schedule of a simulated annealing algorithm consists of four components: the starting temperature, the final temperature, the temperature decrement, and the number of iterations at each temperature.

3.4.2.1 Starting temperature

The starting temperature must be high enough to allow a move to almost all neighbour states. If this is not done then the ending solution will be the same as (or very close to) the starting solution. However, if the temperature starts at a high value then the search can move to any neighbour and thus transforms the search (at least in the early stages) into a random search. Effectively, the search will be random until the temperature is low. The problem is finding the correct starting temperature. At present, there is no known method for finding a suitable starting temperature for a whole range of problems. Therefore, we need to consider other ways. The method suggested by Rayward-Smith et al. (1996) is to start with a very high temperature and decrease it rapidly until about 60% of worst solutions is being accepted. This forms the real starting temperature and it can then be lowered more slowly.

Dowsland (1995) suggests a similar idea which is to heat the system rapidly until a certain proportion of worse solutions are accepted and then slow cooling can be started. Usually a value in the interval 1200-1400 was used as the starting temperature in the present work.

3.4.2.2 Final temperature0

It is common to let the temperature decrease until it reaches zero. However, this can make the algorithm run for a long time. In practise, it is not necessary to do so because as the temperature approaches zero the chances of accepting a worse move are almost the same as when the temperature is equal to zero. Therefore, the stopping criterion can either be a suitably low temperature or when the system is "frozen" at the current temperature i.e. no better or worse moves are being accepted.

3.4.2.3 Temperature decrement

Once the starting and stopping temperature are determined, it is needed to get from one temperature to the other i.e. to reduce gradually the temperature in order to arrive eventually at the stopping criterion. The way of decreasing the temperature is critical to the success of the algorithm. The SA theory states that enough iterations should be allowed at each temperature so that the system stabilises at the temperature. This can be performed by doing a large number of iterations at a few temperatures, a small number of iterations at many temperatures or a balance between the two.

3.4.2.4 Number of iterations at each temperature

Usually a certain number of iterations are carried out at each temperature and then the temperature is decreased. This is repeated until the system freezes into a steady state. There are different methods to reduce temperature iteratively. The method used in this work, first suggested by Lundy and Mees (1986), does only one iteration at each temperature and decreases the temperature very slowly according to the following formula:

$$T_{i+1} = \frac{T_i}{1+\beta T_i}, \qquad (4)$$

where

T_i the current *temperature,*

T_{i+1} the *temperature* of the next iteration,

β a suitably small value.

A suitable value for β might be found by trail and error; some of the values tried in this study were 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4} and 10^{-2}, but usually $\beta = 10^{-7}$ was used.

3.5 Storing the Best Solution so far

The SA algorithm sometimes accepts solutions worse than the current solutions. It is therefore possible, in any single run, the final solution to be worse than the best solution found during the run. It is advised to store details of the best solution found so far. In addition, Glover and Greenberg (1989) argue that with modification, there is less need for SA algorithm to rely on a strong stabilizing effect over time. This idea is supported by Connolley's (1988) modification of SA where having found a suitable fixed temperature, all the remaining iterations are carried out at that temperature (Eglese, 1990).

3.6 Some Variations of Simulated Annealing

In the last 15 years, the literature on the analysis of simulated annealing has grown significantly. Some various types of simulated annealing are listed below: Fast SA, adaptive SA (ASA) or very fast SA or VFSR, ensemble based SA (EBSA), Weighted-ensemble SA, nested simulated annealing, generalized simulated annealing (GSA), distributed simulated annealing and inverse planning simulated annealing, integer augmented simulated annealing (IASA). Real-coded augmented simulated annealing (RASA) is a combination of two stochastic optimisation techniques—simulated annealing and genetic algorithm. It should be added Adaptive Simulated Annealing (ASA), previously called very fast-simulated annealing (VFSA) only named so to distinguish it from the previous method of fast annealing. Gelfand and Mitter (1991) and Tsallis and Stariolo (1996) consider a continuous-space version of simulated annealing, and Moral and Miclo (1999), Cot and Catoni (1998) consider an even further generalisation of the Markov process in standard simulated annealing. However, the analysis of these generalisations of simulated annealing involves relatively sophisticated tools. Hrstka et al. (2003) compare IASA, RASA, simplified real-coded differential GA, and the differential evolution (DE) optimisation techniques.

References (in alphabetical order)

Burke, E. K., and Kendall, G., 1999.
Evaluation of Two-Dimensional Bin Packing Problem using the No Fit Polygon.
Proc. 26th Int. Conf. on Computers and Industrial Engineering, Melbourne, Australia, 15-17 Dec 1999, pp 286-291.

Cot, C., and Catoni, O., 1998.
Piecewise constant triangular cooling schedules for generalized simulated annealing algorithms.
Ann. Appl. Probab. 8(2), pp 375–396. MR 99f:65011.

Dowsland, K.A., 1995.
Simulated Annealing. In (Editor Reeves, C.R.)
Modern Heuristic Techniques for Combinatorial Problems.
McGraw-Hill.

Eglese, R.W., 1990.
Simulated annealing: A tool for operational research.
European Journal of Operational Research, 46, pp 271–281.

Gelfand, S. B., and Mitter, S. K., 1991.
Simulated annealing type algorithms for multivariate optimization.
Algorithmica 6 (3), pp419–436. MR 92d:65109.

Glover, F., and Green berg, H. J., 1989.
New approaches for heuristic search: a bilateral linkage with artificial intelligence.
European Journal of OR, 39, pp119-130.

Hajek, B., 1988.
Cooling schedules for optimal annealing.
Math. Oper. Res., 13(2), pp 311–329, MR 89f:90140.

Hrstka, O., Kucerova, A., Leps, M., and Zeman, J., 2003.
A competitive comparison of different types of evolutionary algorithms.
Computers & Structures, 81(18-19), pp1979-90.

Kirkpatrick, S., Gelatt, Jr C. D., and Vecchi, M. P., 1983.
Optimization by Simulated Annealing.
Science, No. 4598, 13 May.

Lundy, M. M., and Mees, A., 1986.
Convergence of an annealing algorithm.
Mathematical Programming, 34, pp 111-124.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M. N., and Teller, A. H., 1953.
Equation of State Calculation by Fast Computing Machines.
Jr. of Chem. Physics, 21, pp 1087 -1092.

Moins, S., 2002.
Implementation of a simulated annealing algorithm for MATLAB.
A report to Linköping Institute of Technology, Linköping, Sweden
Available in the Web <http://www.ep.liu.se/exjobb/isy/2002/3339/>.

Moral, P. D., and Miclo, L., 1999.
On the convergence and applications of generalized simulated annealing.
J. Control Optim. 37 (4), pp1222–1250 (electronic). MR 2000d:90125.

Rayward-Smith, V. J., Osman, I. H., Reeves, C. R., and Smith, G. D., 1996.
Modern Heuristic Search Methods.
John Wiley & Sons, Chichester, UK.

Salamon, P., Sibani, P., and Frost, R., 2002.
Facts, Conjectures and Improvements for Simulated Annealing
Society for Industrial and Applied Mathematic ISBN: 0898715083.

Tsallis, C., and Stariolo, D. A., 1996.
Generalized Simulated Annealing.
Physica A, 233, 395.

Chapter 4

Simulation and Prediction of Sea-states using Artificial Neural Networks trained with Simulated Annealing

Nomenclature

ANN(s) artificial neural network(s)

c_0, c_1, c_2, c_5 constants in Eq. (1)

d=day a number in the interval 0-365.25 with step 0.125; d=0 represents 0 AM 1 Jan.

CDF cumulative distribution function

$f_{H_S(t_i)given8}(h_i) = f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};\ldots,Hs(t_{i-8})=h_{i-8}}(h_i)$

the pdf of the distribution of H_s (t_i) given its 8 immediate preceding 3-hourly successive H_s 's

$f_{Tz(t_i)|Tz(t_{i-1})=tz_{i-1},Tz(t_{i-2})=tz_{i-2},Hs(t_{i-1})=h_{i-1},Hs(t_i)=h_i,Hs(t_{i+1})=h_{i+1}}(tz_i)$

denotes the conditional pdf of T_z (t_i) given some H_s's and T_z's

h_i the value of significant wave height at time t_i i.e. a particular value of H_s (t_i)

GD gradient descent algorithm

H_s significant wave height

H_s (t_i) a random variable denoting the 3-hourly H_s measured at time t_i

$H_s(t_k) \mid H_s(t_{k-1})=h_{i-1}, H_s(t_{k-2})=h_{k-2}, \ldots, H_s(t_{k-8})=h_{i-8}$

a random variable representing $H_s(t_k)$ given its 8 immediate successive previous 3-hourly H_s's

Input-Vect a general name for the columns of the input matrix of the implemented ANNs, corresponding to a desired T_z in the network target vector

MLE maximum likelihood estimation method

m the order of Hs Markov chain

n number (size) of data

pdf probability density function

R correlation coefficient; number of input node

RMSE root mean square error

SA simulated annealing algorithm

SI	scatter index
t_0	the initial time or the starting day of the desired period of prediction
t_i	the time of the i^{th} 3-hourly measurement of significant wave height
TF	transfer function
tz_i	the value of mean zero crossing period for time t_i
T_z	mean zero-crossing wave period
$T_z(t_i)$	a random variable denoting the T_z of the 3-hourly sea-state of time t_i
x_i	1) output of i^{th} network ; 2)observed H_s or T_z
y_i	simulated H_s or T_z
$\overline{X}$	x_i's mean
$\overline{Y}$	y_i's mean.

4.1 Introduction

Prediction of sea-states is a core requirement of naval meteorology and oceanography, because ocean waves on the sea surface affect virtually all aspects of modern naval operations. In other words the knowledge of heights and periods of oscillatory short waves generated by the action of wind is essential for almost any engineering activity in the ocean.

Some of the activities to which weather and sea-state forecasts are critical are ship routing, coordinating rescue services, fishing, marine engineering, and the management of offshore operations such as offshore drilling and driving operations. Sea-state forecast can also be used to determine if maritime structures have sufficient strength and fatigue resistance.

In the proposed approach of the present thesis for sea-states prediction, the characteristics of a future sea-state are generated as random variates from two different conditional distributions related to the characteristics. The parameters of the two different distributions are estimated from the outputs of two different sets of artificial neural networks (ANNs) trained by simulation annealing. The first 7–network set is trained using a 20–year collection of data observed the North East Pacific to estimate the seven parameters of a proposed distribution for approximating the conditional distribution of a desired future significant wave height (H_s), defined in Sec.5.3.2.2.1 i.e. the distribution of $H_s(t_i)$ given its 8 successive previous H_s's, the time, and four fuzzy season membership values of the desired H_s. The parameters of the probability density function (pdf) of the distribution are estimated from the trained networks outputs. The pdf could be used for H_s prediction and also for extreme

wave analysis (the calculation of the return value). In the past 25-30 years several statistical methods have dealt with H_s extreme value analysis, but these methods assume independence between the H_s's they used for the analysis. A **strength point** of the proposed procedure here for calculating a satisfactory return value is the fact that, contrary to conventional *maxima method*, the proposed method does not assume the H_s's it use are mutually independent. The mean zero-up-crossing period (T_z) of a future sea-state is generated from another distribution whose parameters are estimated from the seven outputs of another 7-network trained set.

4.2 Neural Network Modeling for the Simulation of H_s

4.2.1 Introduction

To develop a model for simulating H_s, those factors that affect ocean wave height are considered. Hourly or daily value of wave height depends on many known and unknown factors. As Fig 4.1 shows any algorithm for H_s simulation should consider two categories of factors: external and internal factors. The H_s's change as these factors change.

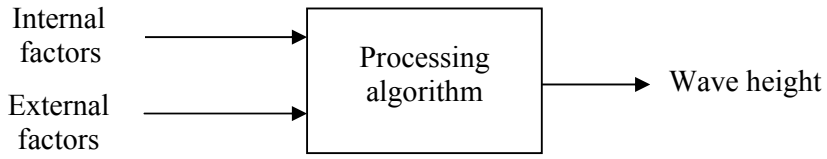


Fig. 4.1 A schematic general model for wave height simulation.

External factors

External factors are those occurrences which are completely uncontrollable. Examples of external factors affecting wave heights are the earth distance from the sun, the radiation angle, the season and the time. Factors such as the distance, angle, season and the time can be determined given the day on which the wave height is being measured. Hence the value of the day is chosen as representing the external factors. Day is usually considered an integer from 0 to 366, but some studies including the present one consider it a real number in the interval (0,365.25) with the step of 0.125. For example, taking 0 AM 1 January as the reference (day =0), day = 0.5 represents 12 noon of the first day i.e. 12 noon 1 Jan, day = 1 denotes 24 midnight of 1 Jan and day =1.625 represents 3 PM 2 Jan.

Internal factors

Internal factors are the factors due to the interactions within the climate. Examples of the internal factors are wind speed, and some unknown unwanted (disturbing) factors. Internal factors are not thoroughly and exactly known and it is even difficult to predict the known ones. On the other hand it is difficult to predict the exact value of H_s from the day, which represents the external factors, without any other information. As an illustration consider the following mathematical model presented by Anderson et al (2001) for significant wave heights, a model for the within-year variability of H_s with a long-term linear trend. The model is based on the time series of H_s's observed during 1978-99 by US National Oceanographic Data Center Buoy 46005 in the NE Pacific (near 46° N 131°W) :

$$\log Hs = c_0 + c_1 \sin(\frac{2\pi d}{365.25}) + c_2 \cos(\frac{2\pi d}{365.25}) + c_5 d + res, \qquad (1)$$

where

c_0, c_1, c_2, c_5	are constants,
d	represents the day in the 22-year period 1978-99 with $d = 0$ for 0 AM 1 Jan. 78, a number ranging from zero to 22*365.25 = 8035.5 with a step equal to 0.125
res = residual	is the random error having a normal distribution with approximately zero mean and variance = 0.1517.

Recalculating the coefficients of the above model using MATLAB gave the following values: $c_0 = 0.7964$, $c_1 = 0.1103$, $c_2 = 0.4408$, $c_5 = 0.00000219362$. Initially in an attempt to reduce the residuals term, additional sine terms were added to Eq.(1); this, however, did not reduce the variance of residuals significantly and therefore predicting the exact value of H_s is not possible from the model given the day. Hence internal factors have to be considered for the prediction of H_s and its return value, as well as external factors. Internal factors are not exactly known and difficult to quantify.

The following flowchart is proposed for predicting H_s using ANNs based on the fact that preceding H_s 's contain implicitly some useful information about internal factors.

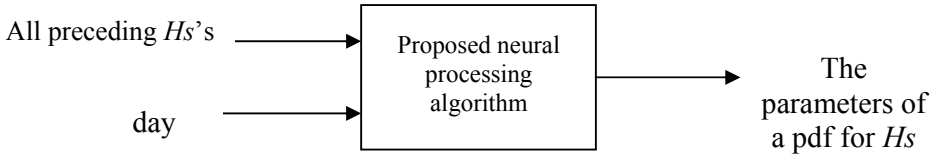


Fig. 4.2 The input and output of the proposed neural modeling for wave height simulation.

Figure 4.2 has been redrawn into Fig. 4.3 using the following notation:

1) $t_0, t_1, t_2, \ldots, t_{73050},$

represent the times at which the H_s's were measured where $t_n = t_{n-1}+3$, and also $t_{73051}, t_{73052}, \ldots$ represent the subsequent times. In general $t_n = t_0 + 3n$ where $t_0 = 0$ and $n = 0,1,2,3,\ldots,73050$.

2) h_i the H_s measured at time t_i, $i = 0,1,2,3,\ldots,73050$;

significant wave heights have not been reported for some i's the corresponding in the site of reference.

3) $H_s(t_0), H_s(t_1), H_s(t_2), \ldots, H_s(t_k),\ldots$

are random variables denoting H_s's occurring at times $t_0, t_1, t_2, \ldots, t_k, \ldots$

Note that h_k could be considered a random variate of $H_s(t_k)$ for $0 \le k \le 73050$.

4) $f_{Hs(t_0),Hs(t_1),\ldots,Hs(t_k)}(h_0,h_1,\ldots,h_k)$, $k = 0,1,2,\ldots,73050,73051,73052,\ldots$

the joint density function of $H_s(t_0), H_s(t_1), H_s(t_2), \ldots, H_s(t_k)$.

5) $f_{Hs(t_k)|Hs(t_{k-1})=h_{k-1},\, Hs(t_{k-2})=h_{k-2},\, Hs(t_{k-3})=h_{k-3},\ldots,\, Hs(t_0)=h_0}(h_k)$

the conditional pdf of $H_s(t_k)$ given its *entire* preceding H_s's.

6) $f_{Hs(t_k)|Hs(t_{k-1})=h_{k-1},\, Hs(t_{k-2})=h_{k-2},\ldots,\, Hs(t_{k-m})=h_{k-m}}(h_k)$

the conditional pdf of $H_s(t_k)$ given its m preceding 3-hourly successive H_s's.

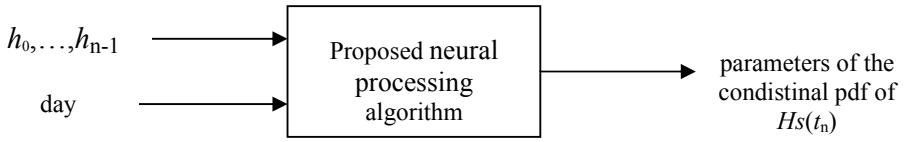


Fig. 4.3 The input and output of a proposed neural modeling for wave height simulation, revised.

4.2.2 The Data

The data implemented in this study were the 3-hourly H_s's measured by US NODC BOUY 46005 from 1/1/1978 to 1/12/2002. Therefore $(2002-1978+1)\times 8 \times 365.25 = 73050$ 3-hourly H_s's should have been observed, but there is no data for some years (e.g. for 1985) and also some data is missing in the period 1978 to 2002. In this period there are only 58507 H_s's whose 8 immediate preceding successive 3-hourly H_s's were available.

4.2.3 On the Number of Required Preceding H_s 's for Input

The proposed neural network algorithm is to find the parameters of the distribution of H_s (t_n) given [H_s (t_0)=h_0, H_s (t_1)=h_1,…, H_s (t_{n-1})=h_{n-1}, and the day on which H_s (t_n) occurs]. The sequence of *Hs* values constitutes a stochastic process. It is assumed that the process is (or could be approximated with) a finite m^{th} order Markov chain whose transient matrix is not necessarily time independent. Therefore according to Markovian property (Ash, 1990) we expect that the statistical dependence of $H_s(t_n)$ on its preceding significant wave heights i.e. on H_s (t_0), H_s (t_1)... H_s (t_{n-1}) should wear off as n approaches a very large number. Now the order of the chain of H_s has to be finding out. Let the order of our Markov chain be denoted by m, then for n =1, 2,…,73050,73051…

$$f_{Hs\,(t\mathrm{n})|Hs(t\mathrm{n}-1)=h\mathrm{n}-1,\,Hs\,(t\mathrm{n}-2)=h\mathrm{n}-2,\ldots,\,Hs(t_0)=h_0}(h_{\mathrm{n}}) \cong$$

$$f_{Hs\,(t\mathrm{n})|Hs(t\mathrm{n}-1)=h\mathrm{n}-1,\,Hs\,(t\mathrm{n}-2)\,=\,h\mathrm{n}-2,\ldots,\,Hs(t\mathrm{n}-m)=h\mathrm{n}-m}(h_{\mathrm{n}}) \qquad (2)$$

Determining m, the order of the significant wave height sequence, remains a problem to be tackled. Smith et al. (1997) have discussed how to find the order of a kind of Markov process. Makarynskyy et al. (2002) assume that a 24-hour history consisting eight measurements contains all necessary information to simulate the wave characteristics i.e. they have supposed that 8 preceding 3-hourly H_s's are enough to be given as input to the ANNs i.e. *order* = m = 8. This assumption has been accepted and implemented in the present approach and that is why Fig. 4.3 has been changed to Fig 4.4.

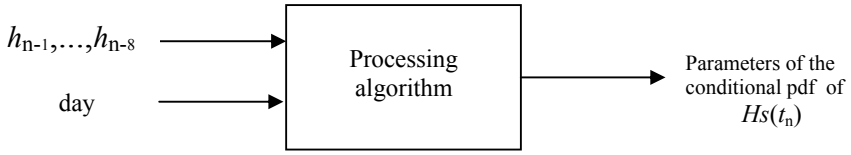


Fig. 4.4 The input and output of the neural algorithm for wave height simulation.

Experiments with some of the implemented ANNs showed that giving the *time* and *season* values of the desired 9^{th} H_s as input to networks instead of giving the *day* on which the H_s occurs increases the convergence of the neural networks and decreases their size. Therefore *the time* and *season* were included in the input of the networks (see Fig 4.5). The *time* of the desired (9^{th}) H_s is greater or equal to 0 and less or equal to 24 but as a preprocessing action was divided by 24 to increase the efficiency of training, hence the time values lie in the interval [0,1]. *Season* is usually considered as spring, summer, fall, and winter (1, 2, 3, 4), but this

study prefers to introduce it as a fuzzy value, such that each day having four membership fuzzy values belongs with some degree to all seasons.

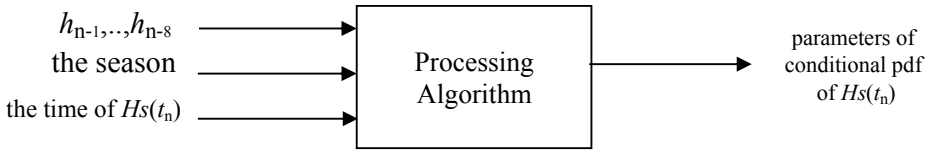


Fig. 4.5 The input and output of the algorithm for H_s simulation, revised.

The four different membership values assigned to each day of a year are: spring membership value, summer membership value, fall membership value, and winter membership value. Appendix B shows how the values are calculated from the day. For example the membership values (1, 0, 0, and 0) are assigned to midspring which is half past seven AM 5th May (d= 127.3125) and (0.0246, 0, 0, 0.9754) are assigned to six AM 7th Feb or 6 AM of the 38th day of Gregorian calendar(d= 38.25). Figure 4.6 shows a diagram related to the fuzzy values.

4.2.4 The Characteristics of the Distribution of H_S

As Fig. 4.5 shows, the objective is to use ANNs for estimating the parameters of the pdf of the distribution of any desired H_s given its 8 immediate preceding successive 3-hourly H_s's,

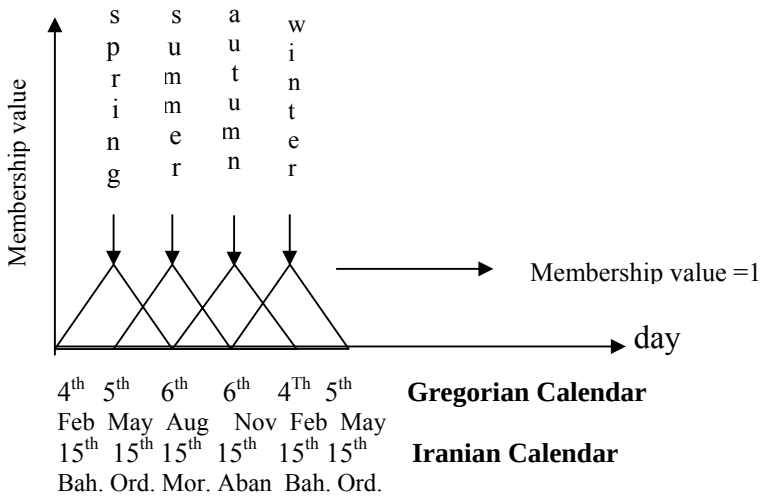


Fig. 4.6 A fuzzy diagram for the *season* membership values.

the time of the H_s and four fuzzy membership values related to the season. The conditional pdf is assumed to satisfy the following two properties:

1) $\lim_{x \to \pm \infty} f(x) = 0,$

2) smoothness i.e. having continuous derivatives.

The two properties imply that there are two points, say a and e ($a < e$) where the value of f(x) approaches zero for $x < a$ and $x > e$ and also there is a point c in the interval (a,e) for which $f'(c) = 0$, e.g. the point that maximises the function. The cubic interpolation according to the spline method will be used for approximating $f(x)$ on the four subintervals shown in Fig. 4.7 because of being a good method of approximating the values of a smooth function on some interval.

4.2.4.1 The subintervals of the distribution pdf

Suppose (a,e) is the interval which encompasses nearly all possible values of the H_s and $f(x)$ denoting the pdf of the conditional distribution of the H_s is defined on this interval, where $a > 0$ due to non-negativity of H_s. According to the Mean Value Theorem in mathematics there exists a point c in the interval (a,e) such that the derivative of the continuous function $f(x)$ for $x=c$ is equal to: $f'(c) = \frac{f(e)-f(a)}{e-a}$. Note that c is not necessarily the global maximum. Since $f(a) \approx 0$ and $f(e) \approx 0$ it is concluded that $f'(c)=0$.

(a,e) was divided into four subintervals i.e.

(a,b), (b,c), (c,d), and (d,e) such that

$b = \frac{a+c}{2}$, and $d = \frac{c+e}{2}$.

4.2.4.2 The parameters of the pdf

The pdf of the distribution of H_s (t_n) given [$H_s(t_{n-1})=h_{n-1,}\ldots$, $H_s(t_{n-8})=h_{n-8}$, the season fuzzy values & the time at which H_s (t_n) occurs] has the following parameters (Fig. 4.7):

a_1	the initial point = a,
a_2, m_1	are respectively the values of the pdf and its first derivative at point b,
a_4, m_2	are respectively the values of the pdf and its first derivative at point d,
a_3	the value of the pdf at point c,
Δ	the length of the interval (a , e), and
λ	the ratio of the length of subinterval (a , c) to that of (a , e).

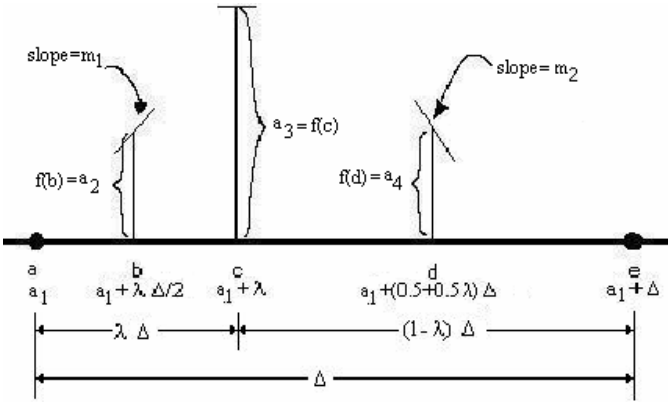


Fig. 4.7 The parameters of the proposed distribution.

4.2.4.3 The hepta-parameter spline distribution

The pdf of the above distribution on subintervals *(a,b), (b,c), (c,d),* and *(d,e)* shown in Fig. 4.7, were approximated by function *g(x)* defined below. *g(x)* is the pdf of a distribution that we called hepta-parameter spline.

$$g(x)=\begin{cases} 0 & x \le a \\ \text{Cubic-Inter.} \begin{bmatrix} a \\ 0 \end{bmatrix} \text{ with slope} = 0 \ \& \ \begin{bmatrix} b \\ a_2 \end{bmatrix} \text{ with slope } m_1 > 0 & a < x \le b \\ \text{Cubic-Inter.} \begin{bmatrix} b \\ a_2 \end{bmatrix} \text{ with slope } m_1 > 0 \ \& \ \begin{bmatrix} c \\ a_3 \end{bmatrix} \text{ with slope} = 0 & b < x \le c \\ \text{Cubic-Inter.} \begin{bmatrix} c \\ a_3 \end{bmatrix} \text{ with slope} = 0 \ \& \ \begin{bmatrix} d \\ a_4 \end{bmatrix} \text{ with slope } m_2 < 0 & c < x \le d \\ \text{Cubic-Inter.} \begin{bmatrix} d \\ a_4 \end{bmatrix} \text{ with slope } m_2 < 0 \ \& \ \begin{bmatrix} e \\ 0 \end{bmatrix} \text{ with slope} = 0 & d < x \le e \\ 0 & x \ge e \end{cases}$$

where

$a, b, c, d, e, a_2, a_3, a_4$ are shown in Fig. 4.7, and

Cubic-Inter. stands for the spline cubic interpolation for the above pairs of points.

.

The slopes m_1 and m_2, shown in Fig.4.7, are defined as follows

$m_1 = f'(b)$ = the slope of $f(x)$ at point $x = b$, and $m_2 = f'(d)$= the slope of $f(x)$ at point $x = d$.

The cubic spline interpolation between two points $(x_1, f(x_1))$ and $(x_2, f(x_2))$ is given by:

$$g(x) = f(x_2)\left(\frac{x-x_1}{x_2-x_1}\right)^3 + \left(f'(x_2) - \frac{3f(x_2)}{x_2-x_1}\right)\frac{(x-x_1)^2(x-x_2)}{(x_2-x_1)^2} + f(x_1)\left(\frac{x-x_2}{x_1-x_2}\right)^3 + \left(f'(x_1) - \frac{3f(x_1)}{x_1-x_2}\right)\frac{(x-x_2)^2(x-x_1)}{(x_1-x_2)^2} \quad (3)$$

where

$x_1 < x < x_2$

$f'(x_1)$ = the slope of $f(x)$ *at* $x = x_1$, and

$f'(x_2)$ = the slope of $f(x)$ *at* $x = x_2$.

The integral of $g(x)$ from x_1 to x_2 is given by Eq. (4)

$$\int_{x_1}^{x_2} g(x) = \frac{(x_2 - x_1)[f(x_1) + f(x_2) + (x_2 - x_1)(f'(x_1) - f'(x_2))]}{2} \tag{4}$$

Since the interpolated values must be positive and the total area under the probability density function is equal to 1, therefore some constraints have to be imposed on the distribution parameters. The parameters $a_1, a_2, a_3, a_4, \lambda, \Delta, m_1,$ *and* m_2 must satisfy the relationship derived from

$$\int_{-\infty}^{\infty} g(x)dx = 1. \tag{5}$$

Integrating g(x) given in Eq. (3) over different subintervals, we have:

$$\int_{-\infty}^{a} g(x)dx \cong 0, \tag{6}$$

$$\int_{a}^{b} g(x)dx = \frac{a_2\lambda\Delta}{4} - m_1\frac{(\lambda\Delta)^2}{48}, \tag{7}$$

and

$$\int_{b}^{c} g(x)dx = \frac{(a_2 + a_3)(\lambda\Delta)}{4} + m_1\frac{(\lambda\Delta)^2}{48}. \tag{8}$$

The sum of the right hand side of Eqs. (6)through (8) yields the area under $g(x)$ over $(-\infty, c)$ i.e.

$$\int_{-\infty}^{c} g(x)dx = \frac{\lambda\Delta}{4}(2a_2 + a_3). \tag{9}$$

Similarly, the area under the pdf over (c, ∞) is:

$$\int_{c}^{\infty} g(x)dx = \frac{(1-\lambda)\Delta}{4}(2a_4 + a_3). \tag{10}$$

Since

$$\int_{-\infty}^{\infty} g(x)dx = \int_{-\infty}^{c} g(x)dx + \int_{c}^{\infty} g(x)dx = 1,$$

then

$$\frac{\lambda\Delta a_2}{2} + \frac{\Delta a_3}{4} + \frac{(1-\lambda)\Delta a_4}{2} = 1. \tag{11}$$

Since a_2, a_3, a_4,λ,andΔ satisfy Eq. (11), therefore one of these five parameters is not independent from the other fours. Arbitrarily a_3 was chosen as the dependent variable on the others and it could be said that the pdf has 7 parameters and the might be therefore called hepta-parameter spline distribution.

Lemma: Since the interpolated values in the subinterval $(a,b) = (a_1, a_1 + \frac{\Delta\lambda}{2})$ must be positive, it is necessary and sufficient to have the following condition satisfied:

$$0 \le m_1 \le \frac{6a_2}{\Delta\lambda}. \tag{12}$$

Proof:

To show the necessary condition, notice that interpolated values of *g(x)* in the interval *(a,b)* must be positive, therefore the convexity of g(x) has to be upward at point *a*, or mathematically $g''(a) \ge 0$. Differentiating twice from Eq.(3) and substituting $x = a$ results in

$$g''(a) = \frac{2}{a-b}\left(\frac{3f(b)}{a-b} + f'(b)\right) + \frac{4f'(a)}{a-b} - \frac{6f(a)}{(a-b)^2} \tag{13}$$

Substituting $f(a) = 0, f'(a) = 0, f(b) = a_2, f'(b) = m_1$,and $a\text{-}b = -\frac{\lambda\Delta}{2}$ into Eq. (13) and noting that $g''(a) \ge 0$ yield

$$m_1 \le \frac{6a_2}{\Delta\lambda}. \tag{14}$$

On the other hand $m_1 \ge 0$, hence

$0 \le m_1 \le \frac{6a_2}{\Delta\lambda}$, and the inequality (12) is proved.

Similarly since interpolated values of *g(x)* in the interval (*d,e*) must be positive, therefore the convexity of *g (x)* is upward on point *e*, or mathematically $g''(e) > 0$ and in a similar way the following inequality could be derived:

$$\frac{-6a_4}{(1-\lambda)\Delta} \le m_2 \le 0. \tag{15}$$

Choosing right parameters for hepta-parameter spline distribution enables it to approximate many of the conventional distributions such as the chi square, positive normal, and the Weibull. For the purpose of comparison, a 2-parameter Weibull with scale $B = 2$, shape parameter $C = 5$ and also its approximating hepta–parameter spline distribution have been plotted in Fig.4.8.

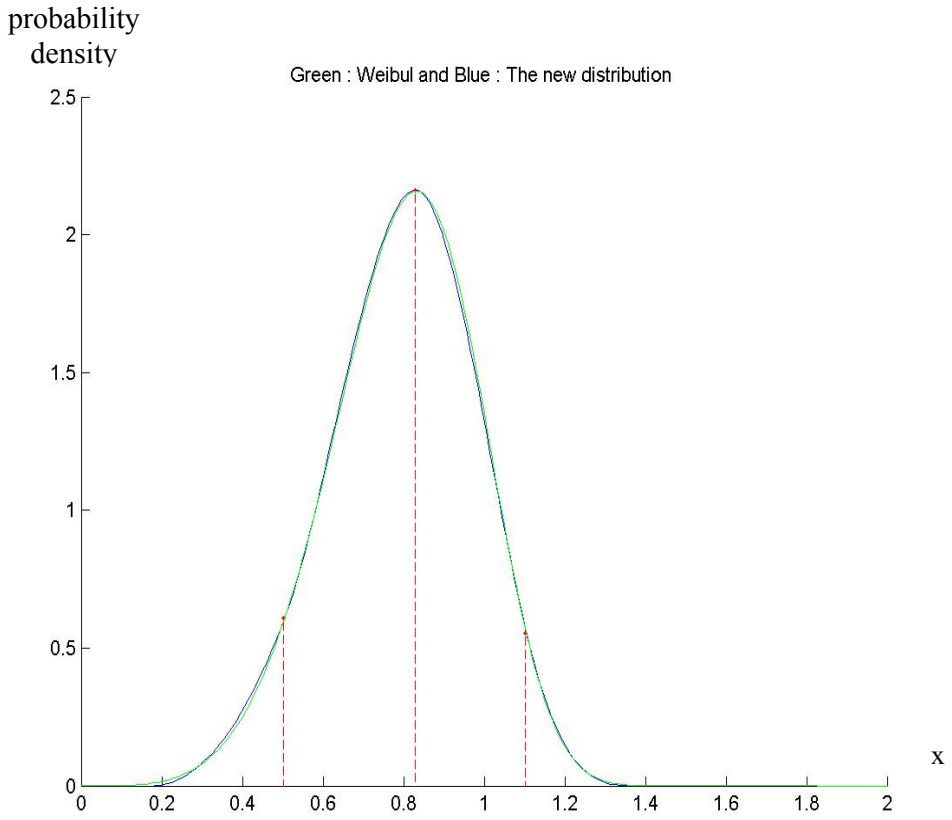


Fig. 4.8 A 2-parameter Weibull distribution (green) with scale $B = 2$ and shape $C = 5$ & its approximating hepta-parameter spline distribution (blue).

4.2.5 The ANNs Implemented for H_S Simulation

Now the problem is to determine the 7 parameters of the pdf which change as the time and the season and necessarily the eight preceding H_s's change. Seven ANNs were trained for estimating the 7 parameters. The criterion to see how well the ANN outputs estimate the parameter, a performance function was formed based on the likelihood function of the maximum likelihood estimation (MLE) method.

Two algorithms were used in training the ANNs utilized in estimating the parameters of the hepta-parameter distributions. In the first attempt the networks training algorithm was gradient descent (GD networks) and the other one consisted of the networks trained by simulated annealing (SA networks).

4.2.5.1 The specification of the networks trained with GD algorithm

An initial attempt was made to train some networks by gradient descent algorithm, but this algorithm was found too slow for our purposes and after a while simulated annealing was chosen as the training algorithm. The specifications of the seven three-layer networks trained using gradient descent for determining the 7 parameters were as follows:

The network type: feed-forward with early stopping
Early Stopping for prevent over-fitting was discussed in Sec. 2.10.

The transfer function (TF) of both hidden layers is *logsig* and that of the output layer is *purelin*(see Fig 2.5 of Chap. 2).

Number of the neurons of each layer

At the beginning the following three sets were chosen for training:
each set had 7 networks of 3 layers.

1) $\mathbf{13\times30\times30\times1}$ **2)** $\mathbf{13\times45\times20\times1}$ **3**) $\mathbf{13\times50\times15\times1.}$

After running the three sets for a while, it was found that the first set **(13× 30× 30× 1)** was being trained very slowly; and was discontinued before discontinuing the other two sets.

4.2.5.1.1 The Target Vector, Input Matrix, Output Vector, and Performance Function of GD networks

Input

A matrix having 13 rows was supplied as input for each of the seven networks. Each column of the matrix consists of 8 successive 3-hourly H_s measurements, 4 fuzzy membership values of the season of the desired H_s and also *time*, which is the time of the desired (9^{th}) H_s, divided by 24. The data set of significant wave heights are those mentioned earlier in Sec. 4.2.2. 58507 sets of 9 successive 3-hourly H_s's (the H_s + its 8 preceding ones) were extracted. The input matrix of size 13×58507 was created for the input was divided into the following 3 sets to use early stopping technique:

A data set with size13×47261 called **TRAIN-INPUT** was used for training, another data set with the name of **VALIDATION-INPUT** and size 13×2475 was considered for validation and the data set **TEST-INPUT** of size of 13×8771 for testing.

Target

The **target** vector was a vector having size 1×47261 consisting the H_s's whose 8 immediate preceding 3-hourly successive H_s 's were available in matrix **TRAIN-INPUT.**

Output

If we supply the matrix **TRAIN-INPUT** of size 13×47261 as input to each of the seven networks, each network would give a vector of size 1×47261 as output after training. These 7 output vectors were used in estimating the parameters of the proposed pdf's for the simulation of H_s according to some equations discussed later.

Performance function

The criterion of MLE method was used as the criterion of performance during the training stage to estimate the parameters of the density function, in other words for training the networks our goal was to maximize the joint density function of $H_s(t_0)$, $H_s(t_1)$,$H_s(t_2)$, …, $H_s(t_k)$, $k = 47261$ or to maximize

$$L = f_{Hs(t_0),Hs(t_1),...,Hs(t_i),...,Hs(t_k)}{}^{(h_0,h_1,...,h_i,...,h_k)}, \qquad (16)$$

where

$H_s(t_i)$ a random variable denoting the 3-hourly H_s measured at time t_i i =0 through 47261, and $k = 47261$,

h_i a particular value of $H_s(t_i)$.

Applying the chain rule for factorization on L and considering Eq. (2) yield

$$L = \prod_{i=47261}^{0} f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};...,Hs(t_{i-m})=h_{i-m},..,Hs(t_0)=h_0}{}^{(h_i)}, \qquad (17)$$

where

$$f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};...,Hs(t_{i-m})=h_{i-m},..,Hs(t_0)=h_0}{}^{(h_i)},$$

the pdf of the conditional distribution of $H_s(t_i)$ given its preceding 3-hourly H_s's,

m the order of the H_s Makov chain that is 8 in our study (see Sec. 4.2.3).

Now consider the following product:

$$L' = \prod_{i=47261}^{0} f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};...,Hs(t_{i-m})=h_{i-m},..,Hs(t_0)=h_0}{}^{'day = \text{ day of } Hs(t_1)^{(h_i)}}. \qquad (18)$$

Intuitively, the greater L', the better the likelihood function L. It could be shown that maximizing L' is equivalent to maximizing L and vice versa. The proof is similar to the one presented in Appendix A for the case in which the networks have been trained with SA.

Some other GD networks tried

As well as the above networks which did not give satisfactory results, the specifications of some more sets of networks used to estimate the parameters of the hepta-parameter spline distributions proposed in this thesis to approximate the distribution of a 3-hourly H_s were as follows. Each set consisted of seven feed-forward networks trained with GD algorithm and all 7 networks had the following size:

4) $13\times45\times20\times1$ input: 8 previous successive H_s's, 4 season values, and day,

5) $13\times30\times30\times1$ input: 8 previous successive H_s's, 4 season values, and day,

6) $13\times50\times15\times1$ input: 8 previous successive H_s's, 4 season values, and day,

7) $13\times45\times20\times1$ input: 8 previous successive ($\log H_s$)'s, 4 season values, & day.

4.2.5.2 ANNs trained with simulated annealing

When gradient descent training algorithm proved to be very slow for the present application, it was decided to use simulated annealing, described in Chapter 3, as the training algorithm.

4.2.5.2.1 The specification of the networks

Several different sets of 7 feed-forward ANNs including the following were trained by simulated annealing

$13\times13\times9\times1$ **$13\times45\times20\times1$** **$13\times28\times15\times1$** **$13\times65\times20\times1$** **$13\times28\times15\times1$**
$13\times20\times12\times1$ **...** **$13\times24\times13\times1$.**

During the training procedure, the simulations resulted from all of the above sets were compared and finally the training of ANNs with size $13\times24\times13\times1$ continued to be used in final estimation of the parameters of the hepta-parameter spline distribution used for simulating significant wave height. The transfer function of both hidden layers is *logsig* (Fig. 2.5 of Chap. 2) and that of the output layer is *purelin* (Fig .2.5 of Chap. 2). Fig. 4.9 shows the schematic form of the selected ANNs.

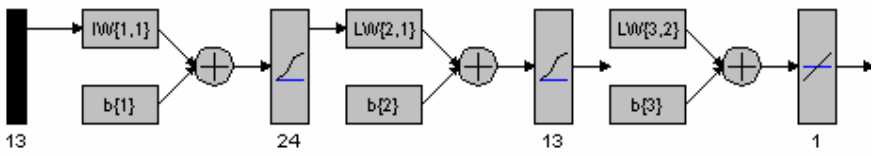


Fig. 4.9 The architecture of the ANNs used for H_s simulation.

4.2.5.2.2 The Target Vector, Input Matrix, and Output Vector of SA networks

Target

The first 49736 H_s's of the 58507 extracted H_s's mentioned in Sec. 4.2.2 was chosen as the **target,** a vector of size 1×49736 for training the networks. For each H_s in this vector its 8 immediate previous 3-hourly successive H_s's do exist.

Input

The input data set is a matrix of size 13×2475 consisting of the two data sets described in Sec. 4.2.5.1.1, i.e.**TRAIN-INPUT** with size 13×47261, and **VALIDATION-INPUT** of size 13×2475. The first eight elements of a column of the input matrix includes the eight preceding 3-hourly H_s's of the corresponding H_s in the target vector, extracted from the data described in Sec. 4.2.2. The time and the season of the corresponding H_s in the target vector constitute the other 5 rows of the column. Although the time values lie in the interval [0-24] but in order to increase the efficiency of training they were divided by 24. The season is usually considered as spring, summer, fall and winter (1, 2, 3, and 4); however this study expresses it as a fuzzy value, such that a particular day of the year belongs to all of the four seasons with 4 values each lying in [0 1]. Therefore 4 different membership values were assigned to each day of a year as described in Sec. 4.2.3. Each of its columns containing 13 pieces of information related to the corresponding H_s in the target vector.

The reason for choosing the 8 preceding H_s's and the other 5 variables as input for the H_s of a desired time t_i, stems from the fact that a very suitable way to predict the H_s of time t_i using the information included in all of its preceding 3-hourly H_s's is through finding the conditional distribution of $H_s(t_i)|H_s(t_{i-1})=h_{i-1},\ldots,H_s(t_0)=h_0$. On the other hand the statistical dependence of $H_s(t_i)$ and $H_s(t_{i-k})$ is expected to wear off as k approaches ∞(Ash,1990); therefore the distribution of $H_s(t_i)$ given $H_s(t_{i-1}) =h_{i-1},\ldots,H_s(t_{i-m})=h_{i-m}$ could be found instead, where m is an integer. As stated earlier in Sec. 4.2.3, m was chosen to be eight. According to

the above discussion the factors that affect the conditional distribution are the values of $h_{i-1},...,h_{i-8}$ and also t_i. Due to some experiments it was concluded that giving the year in which the desired H_s will occur as an input is not that useful and the day on which the H_s will occur, a real number in the interval [0,365.25], seemed a better choice for input. Experiments showed that using the *time* and the 4-value *season* of the desired H_s instead of *day* as input increase the convergence of the neural networks and decrease their size. Therefore the 5 values corresponding to the desired H_s were added to the input as well as its 8 previous H_s's. Hence for every H_s in the target vector there are 13 values in the input.

Output

The ***output*** of the 7 networks are 7 real numbers which are used to calculate the 7 parameters of the proposed pdf, described in Sec. 4.2.4.3, for the H_s of time t_i given its eight immediate successive previous 3-hourly H_s 's denoted by

$$f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};...,Hs(t_{i-8})=h_{i-8}}(h_i) \text{ or briefly } f_{H_S(t_i)given8}(h_i)$$

where

$H_s(t_i)$ a random variable denoting the significant wave height of time t_i,

$t_i = t_0 + 3i$ the time of the i^{th} 3-hourly measurement of H_s, where t_0 = the initial or starting time.

4.2.5.2.3 Performance function

For using SA as **the training algorithm** a performance (cost) function was proposed based on the MLE method. The proposed function, which is to be maximized, is the product of the conditional pdf's of all of possible *Hs*'s available in the data set described in Sec. 4.2.2 i.e.

$$P = \prod_{i=58507}^{8} f_{H_S(t_i)given8}(h_i), \qquad (19)$$

where

h_i is a particular observed value of $H_s(ti)$.

But those H_s's whose 8 previous H_s's exist were used:

$$P_1 = \prod_{i=49736}^{8} f_{H_S(t_i)given8}(h_i) \tag{20}$$

Equivalently

$$P_2 = \sum_{i=49736}^{8} \log f_{H_S(t_i)given8}(h_i), \tag{21}$$

was used as the performance function or the performance index, instead of P_1. The relationship between maximizing P, P_1, P_2 and maximizing the likelihood function of the MLE method is discussed in Appendix A.

As well as maximizing P_2, during the training procedure it was attempted to minimize the root mean square error (RMSE) between the forecasted H_s's for the period 78 -99 and their observed values according to Eq. (31) to get more accurate results. To calculate the RMSE, the mean of the corresponding conditional density function was taken as the forecast for the desired H_s.

The initial value of P_2 was less than -100,000 when training started, but increased gradually. During the training procedure the simulated data and the observed data were also compared every now and then to prevent problems such as over-fitting. The training was discontinued when the performance reached a steady point at P_2 = -9375. However the networks with P_2 = - 22174 was selected. With this performance the 25-year simulated H_s's showed a better fit compared with the corresponding observed 25-year data.

4.2.6 Parameter Estimation

Let x_i denotes the output of the i^{th} network where i = 1 through 7. Parameter a_1 is positive, therefore the outputs of none of the 7 networks can be used directly for estimating a_1 because the transfer function of the output layers is *purelin* and consequently the output of the networks is between - ∞ to ∞. The transformation $a_1 = x_1^2$ is used to arrive at a positive number.

All of the relationships used to calculate the parameters of the hapta parameter spline distribution from network outputs are as follows

Constraint	Relationship	
$a_1 > 0$	$a_1 = x_1^2$	(22)
$0 < a_2 < a_3$	$a_2 = a_3 \text{logsig} x_{10}$.	(23)

$$\Delta > 0 \qquad \Delta = x_3^2 . \tag{24}$$

$$0 < a_4 < a_3 \qquad a_4 = a_3 \text{logsig} x_4 . \tag{25}$$

$$0 \le m_1 \le \frac{6a_2}{\Delta\lambda} \qquad m_1 = \frac{6a_2}{\lambda \Delta} \text{logsig} x_5 . \tag{26}$$

$$\frac{-6a_4}{(1-\lambda)\Delta} \le m_2 \le 0 \qquad m_2 = -\frac{6a_4}{(1-\lambda)\Delta} \text{logsig} x_5 . \tag{27}$$

$$0 < \lambda < 1 \qquad \lambda = \text{logsig}\, x_7. \tag{28}$$

Cosidering Eqs.(11), (23) and (25), a_3 is calculated from the following relationship:

$$a_3 = \frac{2x_3^{-2}}{(\text{logsig} x_7)(\text{logsig} x_2) + 0.5 + (1 - \text{logsig} x_7)(\text{logsig} x_4)}, \tag{29}$$

where logsig x_j is given by

$$\text{logsig}\, x_j = \frac{1}{1+e^{-x_j}}, j = 1,...,7. \tag{30}$$

In practice, a MATLAB code calculates a_3 from Eq. (29) using x_2, x_3, x_4, x_7 i.e. the outputs of the networks no. 2, 3,4, and 7. The code then calculates a_2 and a_4 from Eqs. (23) and (25) respectively. Other parameters are calculated easily using the right relationship. The input of the networks in training phase is a matrix of 13×49736, then we are looking for 49736 pdf's. 49736 values will be produced for each of the 7 parameters of our 49736 pdf's i.e. 7 vectors of size 1×49736 will be obtained for a_1, a_2, Δ, a_4, m_1, m_2, and λ.

4.2.7 Simulation of Significant Wave Height

To simulate H_s the for a desired future period using the trained networks, the starting day of the period and the eight 3-hourly successive H_s's of its previous 24 hours are given to a MATLAB code. The code then calculates the 4 fuzzy membership values and the time. The output of the trained networks are obtained using the associated input vector of size 13×1 which we shall call **Input-Vect.** The 13 input values are 8 initial H_s's, the time value and the 4 fuzzy membership values for the season related to the H_s being predicted. From the 7 outputs of the 7 networks, the code will then calculate the estimates for the parameters of the conditional hepta-parameter spilne distribution of the desired H_s given its 8 previous successive H_s's. The distribution pdf is now completely known.

The random number generation method called the *inversion method* or *the inverse transformation method* (Naylor et al., 1968; MATLAB Help) has been used to simulate the H_s

of a desired time. Let $F(x)$ denotes the cumulative distribution function (CDF) of the above conditional pdf. According to the method, if we generate uniformly a distributed random number r from the interval [0 1] and set $r = F(x)$, it could easily be shown that $F^{-1}(r)$ will be a random variable that has the above conditional pdf as its density function. In general, if it is proved to be either impossible or extremely difficult to express x in terms of the inverse transformation $F^{-1}(r)$, a numerical approximation to the inverse of F and F^{-1} could be obtained or other random generation methods should be used.

To proceed the simulation for the subsequent steps, **Input-Vect**(i) for $i = 1,2,...,7$ is replaced with **Input-Vect**(i+1) and the previous simulated variate i.e. $F^{-1}(r)$ replaces Input-Vect(8). The code calculates the rest of the elements of **Input-Vect** from the time of the subsequent H_s being predicted. Using the new input vector **Input-Vect** of size 13×1 the above iteration is continued until the simulation of H_s for the desired period is completed. Figure 4.10 shows a typical simulation of H_s for 30 years.

4.2.7.1 Similarity of the simulated and the observed data

Appendix C shows that if the artificial neural networks are trained well, the set of simulated signifant wave heights and the set of observed data would be two random samples of the same joint distribution.

4.2.8 Forecasting H_s using the Simulated Data

Forecasting the H_s of a desired time is done using the previous history of the H_s. To forecast an H_s we could find its distribution given <u>all</u> of its previous observed H_s's and then calculate the mean or most probable value of the distribution as the forecast. As pointed out earlier, only 8 immediate 3-hourly successive H_s's preceding the H_s being predicted have been given in this work as input instead of its entire previous H_s's. According to the lemma stated in App. 4-C if the simulation for an H_s is repeated several hundred times, the empirical distribution function (EDF) of the simulated data will approach the actual conditional *CDF* of the H_s given its eight preceding H_s's. The most probable and mean values of the *EDF* are two point estimates for the H_s being forecasted (see Table 4.1).

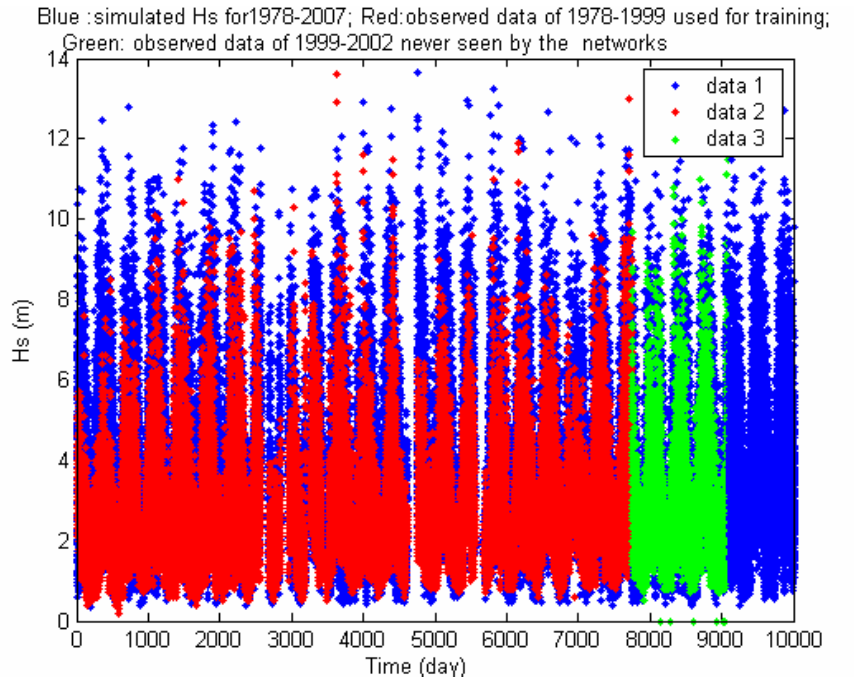


Fig. 4.10-1 A typical 30-year simulation of H_s for 1978-2007 and 25-year observed H_s's in the NE Pacific during 1978-2002,
Blue: Simulated H_s's for 1978-2007, Red: Observed H_s 's for 1978-98 used for training,
Green: Observed H_s's during 1999-2002, never seen by the ANNs (Test data).

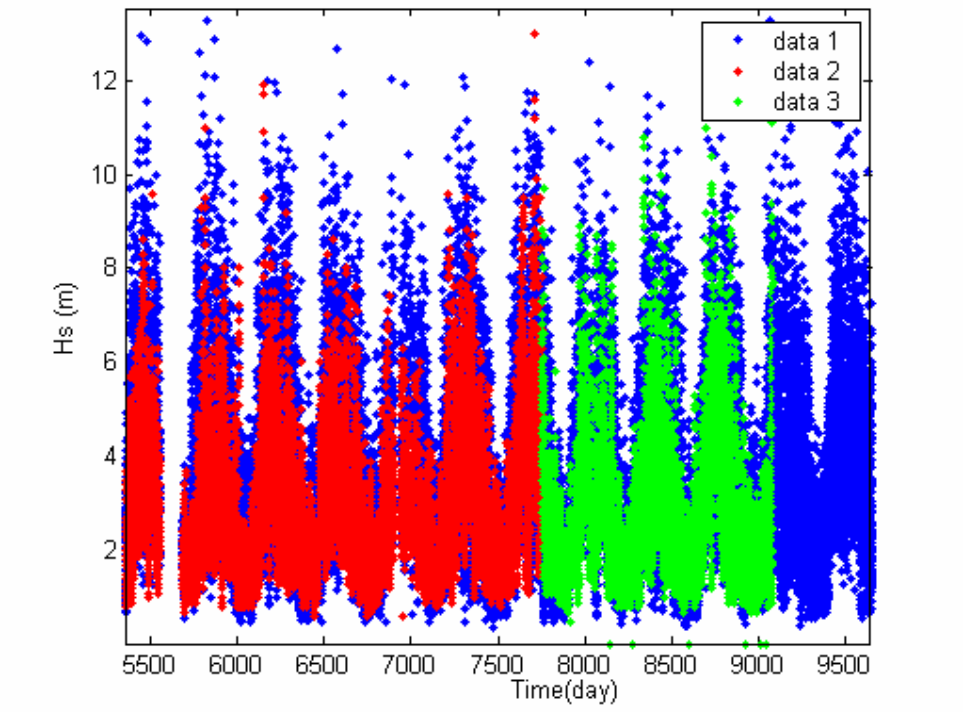


Fig. 4.10-2 A zoom form of Fig. 4.10-1.

4.2.9 The Networks Efficiency

The efficiency of the proposed ANNs was estimated in terms of RMSE, correlation coefficient (R), and scatter index (SI) calculated from Eqs. (31) through (33) respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - x_i)^2}{n}} \quad (31), \qquad R = \frac{\sum_{i=1}^{n}(x_i - \bar{X})(y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{X})^2 \sum_{i=1}^{n}(y_i - \bar{Y})^2}} \quad (32), \qquad SI = \frac{RMSE}{\bar{X}} \quad (33),$$

where

x_i , y_i observed & simulated data respectively,

n number of data,

$\bar{X}$ & $\bar{Y}$ x_i's & y_i's mean respectively.

As an illustration for long term forecasting, the simulation were carried 5378 times formore than 3 years from early 1999 through early 2002 using the trained ANNs. The reason for choosing this period is the fact that their corresponding observed H_s's, which had never been seen by the neural networks, had been reserved as test data. The inputs required by the MATLAB code used for the simulation were the eight 3-hourly H_s's observed prior to the starting day as well as the starting day of the period (day = t_0 = 84.25) and the duration of simulation period. Figure 4.11 & Table 4.1 compare the observed and simulated sets of data.

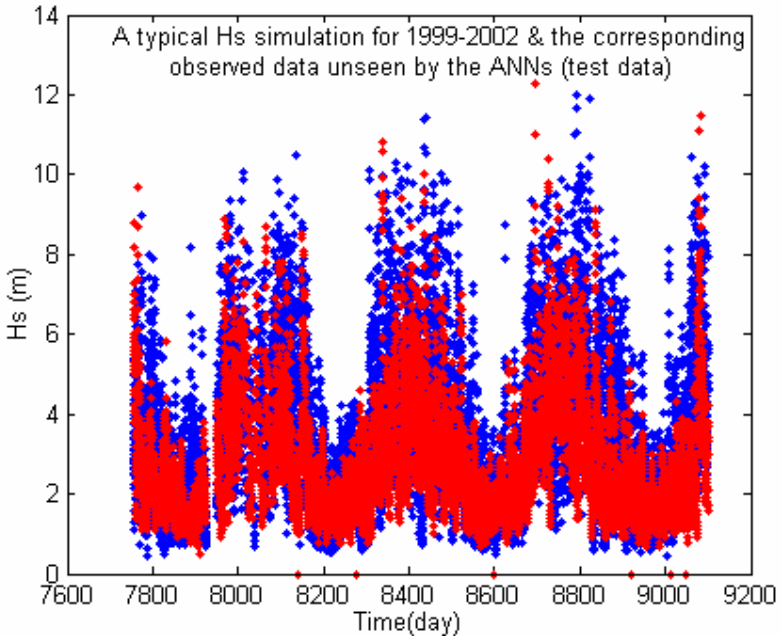


Fig. 4.11 Observed (red) and simulated (blue) H_s's for 1999-2002, a typical simulation.

Table 4.1 The comparison of the observed H_s 's versus predicted *Hs*'s for 1999-2002, simulated 5378 times.

Forecasting method for	**RMSE**(m)	**R**	**SI**
Most probable *Hs* among 5378 runs	1.31	0.72	0.48
***Hs* mean among** 5378 runs	1.45	0.59	0.53
A typical genrated *Hs*	2.10	0.40	0.76

As an illustration for short term forecasting**,** using 3 methods, the entire H_s's of the subsequent 3-hour interval for the period 1978-1999 were predicted given their corresponding 8 previous immediate successive 3-hourly observed H_s's. In other words every H_s in the target vector was predicted by 3 methods. One of the methods calculates the most probable value (mode) of the pdf; the other computes its mean and the third generates a variate from the pdf by the inversion method. All of the 3 methods use the conditional distribution of an H_s in the target vector given its 8 previous observed H_s's existing in the input matrix. The parameters of the pdf were readily estimated from the ANNs outputs. Table 4.2 shows the result of calculations of the indices of this simulation efficiency using the forecasted values and the target vector which includes the observed data. According to the indices of efficiency (RSME, SI and R) recorded in Table 4.2, the mean of the corresponding pdf of the H_s of a desired time sea-state is a good forecast for the H_s of the subsequent 3-hour interval.

Table 4.2 The indices of efficiency of H_s forecasts of subsequent 3-hour intervals for 1978-99.

Forecasting method for *Hs*	**RMSE**(m)	**R**	**SI**
pdf mode value	0.51	0.95	0.19
pdf mean value	0.43	0. 95	0.16
A genrated *Hs*	0.73	0.87	0.27

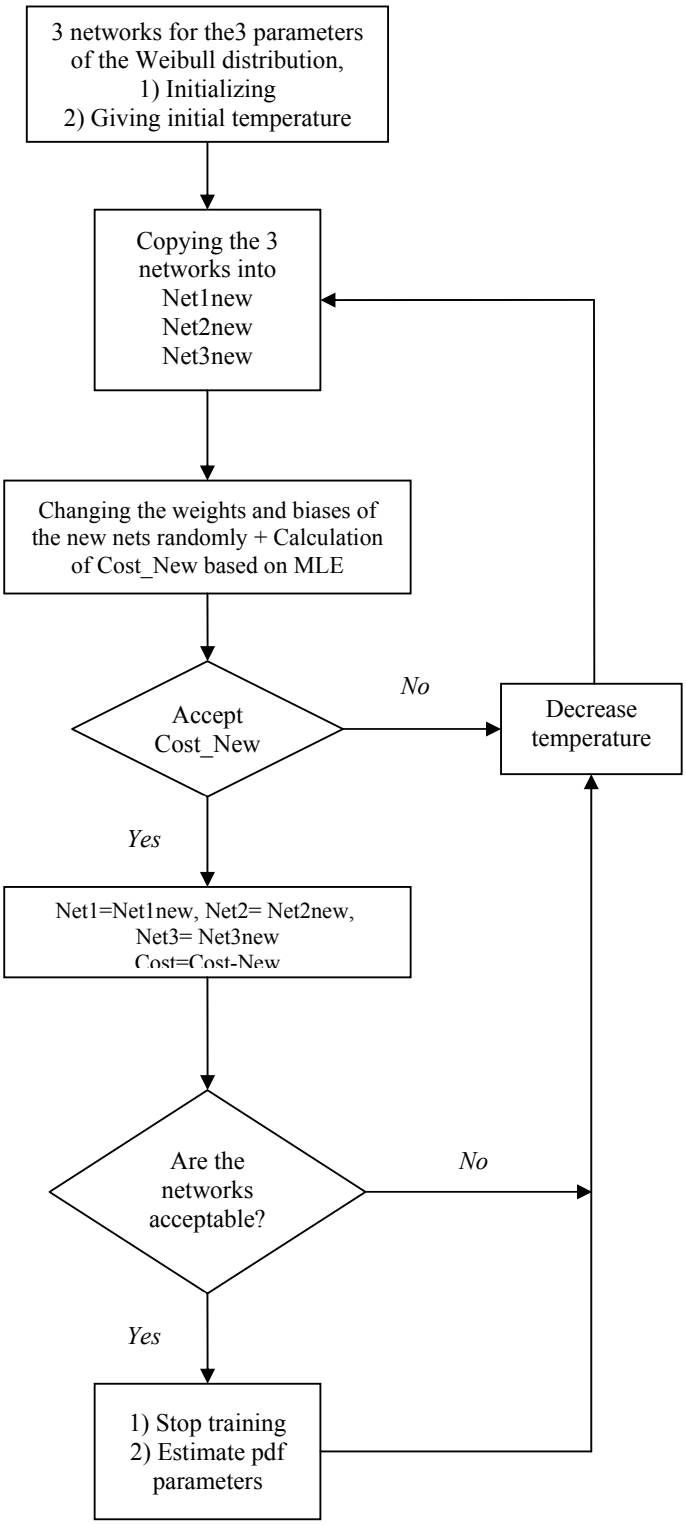


Fig 4.12 The SA training algorithm of the networks used to estimate the 3 parameters of a Weibull distribution for the simulation of H_s.

4.2.10 ANNs for the Weibull Parameters Estimation

As well as trying to estimate the parameters for the hepta-parameter spline distribution, two 3-network sets of the architectures $13\times24\times13\times1$ and $13\times24\times1$

were tried for estimating the parameters of a 3-p Weibull distribution. However, the outputs did not give satisfactory results for the maximum likelihood estimates of the parameters. The algorithm of the training phase for the 3-network set of ANNs of size $13\times24\times13\times1$ for the distribution is shown in Fig. 4.12.

4.3 The ANNs for the Simulation of Mean Zero Up-Crossing Period

To simulate the mean zero-crossing period(T_z), defined in Sec. 5.3.2.2.2, using the stochastic nature of T_z, the proposed hepta-parameter spline distribution was also chosen to approximate the conditional distribution of the T_z of the sea-state measured at time t_i, $T_z(t_i)$, given the following information:

1,2) its two immediate previous T_z's, i.e. $T_z(t_{i-1})=tz_{i-1}$, $T_z(t_{i-2})=tz_{i-2}$,

3) its immediate preceding H_s i.e. $H_s(t_{i-1})=h_{i-1}$,

4) the H_s of the sea-state at time t_i, $H_s(t_i)=h_i$,

5) next immediate H_s, $H_s(t_{i+1})=h_{i+1}$,

6) the time,

7, 8, 9, 10) four fuzzy membership values for the season in which the $T_z(t_i)$ occur.

The pdf of the distribution is denoted by

$$f_{Tz(t_i)|Tz(t_{i-1})=tz_{i-1},Tz(t_{i-2})=tz_{i-2},Hs(t_{i-1})=h_{i-1},Hs(t_i)=h_i,Hs(t_{i+1})=h_{i+1}}(tz_i)$$

4.3.1 The Data Employed

The H_s's and T_z's given as input data in this study were those measured by BUOY 46005 from 1978 to1999. Therefore $22\times8\times365.25 = 64284$ observed 3-hourly H_s's and their 64284 corresponding T_z's are expected to exist; however there is no data for some years e.g. for 1985. The details of the missing H_s's and T_z's are reported by Anderson et al. (2001). A successful attempt was done to interpolate those missing H_s's or T_z's whose immediate

previous and immediate next data were available. The interpolation was simply the mean of the previous and the next values. The others are still missing. On the whole, 55922 sea-states became available to be used for training the networks. The code written for training has the option to use all or a part of the 55922 sea-states for training the networks. A vector of size 1×25000 from the data was used as target.

4.3.2 The Specifications of the Networks used in the Simulation of T_z

The parameters of the hepta-parameter spline distribution used to approximate the conditional distribution of T_z were estimated from the outputs of seven trained networks. At the beginning, several different feedforward network architectures including
$10\times5\times10\times1$, $10\times5\times5\times1$, $10\times3\times10\times1$, $10\times1\times10\times1$, and $10\times1\times15\times1$
were tested and finally the size $10\times1\times15\times1$ was chosen (Fig 4.13). The transfer functions of the 2 hidden layers and the output layer were of '*logsig*', *'logsig' and 'purelin'* type respectively (see Fig. 2.5 of Chap. 2), and the training algorithm was simulated annealing.

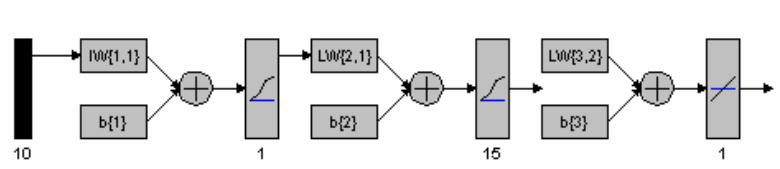


Fig. 4.13 The architecture of the ANNs used for T_z simulation.

4.3.3 Target Vector, Input Matrix, and Output Vector

The **target** could be a vector of maximum size 1×55922 observed T_z's from 1978 to 1999; however a vector of 1×25000 was used as the target to increase the speed of training. This set of data corresponds to 1 Jan 1978 to mid-1986. The target vector was scaled by the following expression to increase the training efficiency

$$T_{z\text{new}}=\left(\frac{Tz-\min(Tz)}{\max(Tz)-\min(Tz)}\times0.8\right)+0.1, \qquad (34)$$

where min(T_z) = 3.7 and max(T_z) = 17.5.

The **input** data for training the networks was a matrix of 10×25000.
Let $t_i = t_0+3i$ denote the time of the i^{th} 3-hourly observed T_z; where t_0 = the initial time. For any $T_z(t_i)$ in the target vector, i.e. the mean zero up-crossing period of the sea-state occurring at time t_i, the following data were included in the rows of the corresponding column of the **input** matrix.

Rows 1 through 5:

Tz(t_{i-1}), Tz(t_{i-2}), $H_s(t_{i-1})$, $H_s(t_i)$, and $H_s(t_{i+1})$ i.e. the next immediate H_s already predicted. To increase the efficiency, the T_z 's and H_s's were scaled using Eq.(34) and (35) respectively

$$H_{s\text{-new}} = \left(\frac{Hs - \min(Hs)}{\max(Hs) - \min(Hs)} \times 0.8 \right) + 0.1, \qquad (35)$$

where min (H_s) = 0.2 and max (H_s) = 13.6.

Rows 6 through 10: time and season.

The time of the desired T_z divided by 24 to increase the efficiency of training and 4 membership fuzzy values for the season of $T_z(t_i)$ i.e. spring value, summer value, fall value, and winter value were given. Experiments showed that if the time and the season values of the desired T_z are given as part of the input to the networks the convergence of the neural networks will increase and their sizes decrease.

The **output** of each of the 7 networks is a real number used to estimate, by Eqs. (22) through (29) the parameters of the of a distribution approximating that of the desired T_z given some data denoted by $f_{Tz(t_i)|Tz(t_{i-1})=tz_{i-1},Tz(t_{i-2})=tz_{i-2},Hs(t_{i-1})=h_{i-1},Hs(t_i)=h_i,Hs(t_{i+1})=h_{i+1}}(tz_i)$.

4.3.4 The Cost Function and theTraining Procedure

The objective during training the seven networks was to maximize the performance (cost) function P' given in Eq. (36). It could be shown that this function is based on the likelihood function of the MLE method in statistics.

$$P' = \sum_{i=25000}^{8} \log f_{Tz(t_i)|Tz(t_{i-1})=tz_{i-1},Tz(t_{i-2})=tz_{i-2},Hs(t_{i-1})=h_{i-1},Hs(t_i)=h_i,Hs(t_{i+1})=h_{i+1}}(tz_i) \qquad (36)$$

where

$H_s(t_i)$ a random variable denoting the 3-hourly H_s measured at time t_i,

h_i a specific value of $H_s(t_i)$,

$T_z(t_i)$ a random variable denoting the T_z of 3-hourly sea-state of time t_i,

tz_i a specific value of $T_z(t_i)$

However, simultaneously, the RMSE between the observed and predicted T_z's for a period starting from 1978 was calculated, with the goal to prevent overfitting in training by

minimizing the RMSE. To calculate the RSME, the mean of the corresponding conditional pdf was calculated and considered as the predicted value of the desired H_s.

The weights and biases of the 7 networks were changed randomly and the 7 parameters of the conditional pdf corresponding to a zero up-crossing period in the target vector i.e.

$f_{Tz(t_i)|Tz(t_{i-1})=tz_{i-1},Tz(t_{i-2})=tz_{i-2},Hs(t_{i-1})=h_{i-1},Hs(t_i)=h_i,Hs(t_{i+1})=h_{i+1}}(tz_i)$, where $i = 8$ to 25000,

were calculated from the outputs of the networks. The value of the function in Eq. (36) was also computed. If the function showed improvement the new weights and biases would be accepted; if not, the process of changing weights and biases would continue to improve the performance function and RSME simultaneously. The simulated T_z's were occasionally plotted versus the corresponding observed data and checked visually. Figure 4.14 shows the simulated T_z's in blue and the observed ones in red for 1978-1999 when the training process was terminated. At this time, the value of the performance function and the RSME had reached +70529 and 0.72731 seconds respectively. Needless to say that it is not claimed the networks of size $10\times1\times15\times1$ are the best architecture one could find for the above purpose. As stated before, SA was used as the **training algorithm** to maximize the performance function. Fig. 4.15 shows a zoom of the top part of Fig. 4.14.

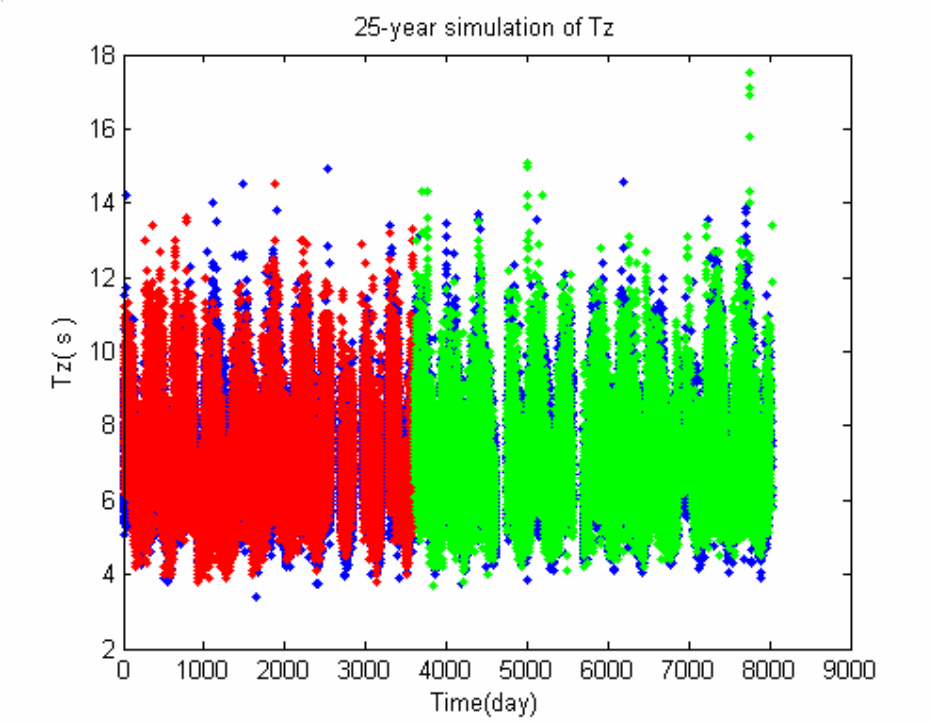


Fig 4.14 The simulated (blue) T_z's for 1978-99 in a region of the NE pacific using the trained networks of size $10\times1\times15\times1$ and the data used for training(red)and the test data(green).

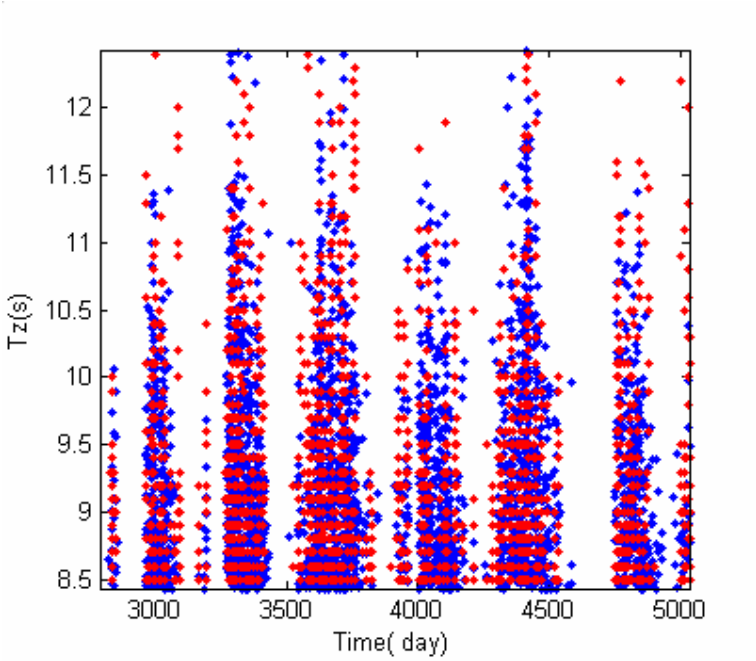


Fig. 4.15. A zoom of top part of Fig. 4.14, the observed *Tz* is in red and the simulated is in blue.

4.3.5 Simulation of T_Z using the Trained Networks

To simulate $T_z(t_i)$, the mean zero up-crossing wave periods for time t_i, the starting day of the period or the initial time t_0 and the following data are given as **input** to a MATLAB code: the T_z of 3-hour-ago sea-state or $T_z(t_{i-1})$, the T_z of 6-hour-ago sea-state or $T_z(t_{i-2})$, the H_S of 3-hour-ago sea-state or $H_s(t_{i-1})$, $H_s(t_i)$ and also the next predicted H_S or $H_s(t_{i+1})$.

$H_s(t_{i+1})$ will take place in future and is unknown, but it is forecast through generating a random variate from a conditional distribution whose probability density function is $f_{Hs(t_{i+1})|Hs(t_i)=h_i;\ldots,Hs(t_{i-7})=h_{i-7}}{}^{(h_i+1)}$ and its parameters are estimated by the help of the networks of size $13\times24\times13\times1$ already trained for the simulation of significant wave height (discussed in Sec.4.2.7). The four fuzzy membership values and the time are then calculated by the MATLAB code according to the procedure described in Appendix B.

The **output** of the trained networks are obtained using the associated input vector of size 10×1 called **Input-Vect**. The parameters of the conditional disribution of the desired T_z will then be estimated by the code from the 7 outputs of the 7 networks according to Eqs. (22) through (29). The conditional distribution of T_z is now completely known.

Inversion method described in Sec. 4.2.7 has been used to generate a variate from the distribution for the T_z of the sea-stae of the desired time. To proceed the simulation for subsequent steps say for time t_{i+2} the code replaces **Input-Vect**(1) with the already generated variate; **Input-Vect**(2) is replaced with the old **Input-Vect(**1); **Input-Vect**(3) is replaced with the old **Input-Vect**(4); **Input-Vect(**4) is replaced with the old **Input-Vect**(5). **Input-Vect**(5) has to be replaced with the H_s of the time of the $(i+2)^{th}$ 3-hourly sea-state i.e. $H_s(t_{i+2})$. This H_s is forecast through generating a random variate from the conditional probability density function $f_{Hs(t_{i+2})|Hs(t_{i+1})=h_{i+1};Hs(t_i)=h_i,\ldots,Hs(t_{i-6})=h_{i-6}}(h_{i+2})$ whose parameters are readily given by the trained networks of size $13\times24\times13\times1$ discussed in Sec. 4.2.7. **Input-Vect**(6) is the time of the T_z being predicted divided by 24. The MATLAB code calculates the fuzzy values of **Input-Vect** related to the season i.e. **Input-Vect**(7) through **Input-Vect**(10) according to Appendix B. Using the new 10×1 input vector **Input-Vect**, the above iteration is repeated until the simulation of T_z for the desired period is complete.

Figure 4.14 shows a typical long term T_z simulation for 22 years starting from 2 Jan 1978 to 1999. Figure 4.16, showing the cumulative distribution function of the above observed and simulated T_zs, conveys the degree of agreement between the observed and simulated T_zs.

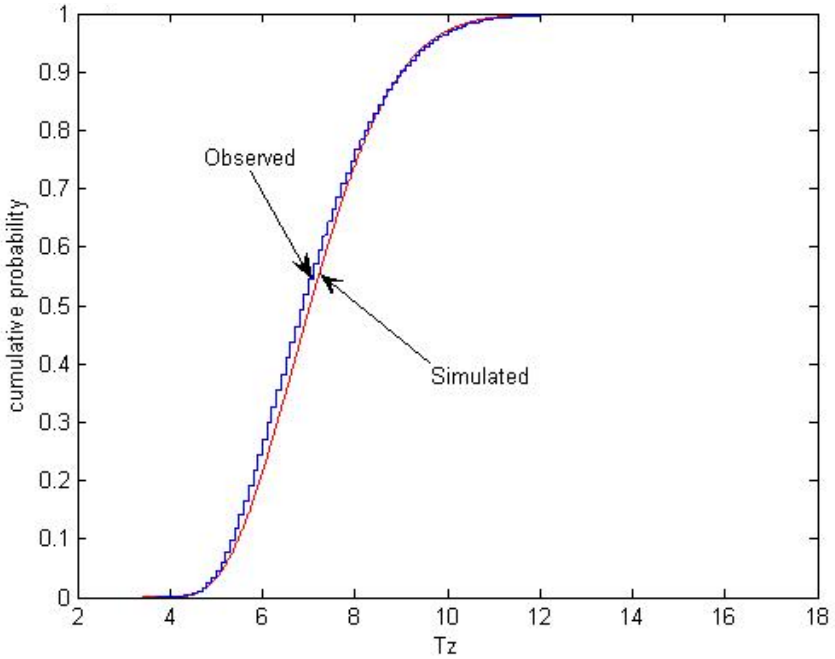


Fig. 4.16 The empirical cumulative distribution function of the observed and simulated T_zs of 1978-99.

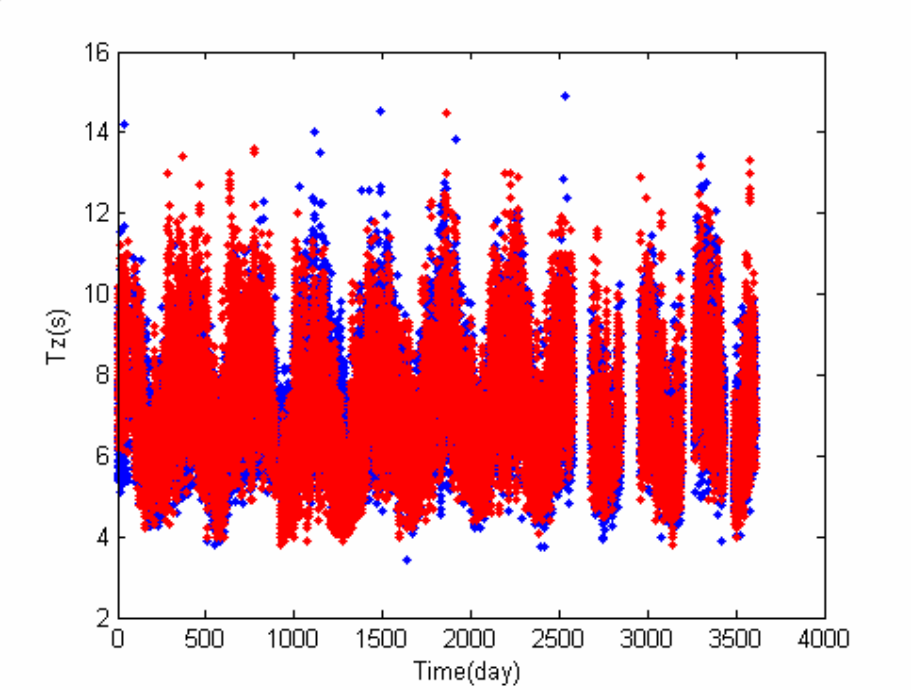


Fig. 4.17. Training data observed from 1978 to mid-86(red) and the corresponding simulated data (blue).

The training set of data and the test data together with their corresponding simulated data has been shown in Figs. (4.17) and (4.18). As the figures show the model has learned to imitate the behavior of the observed T_z's.

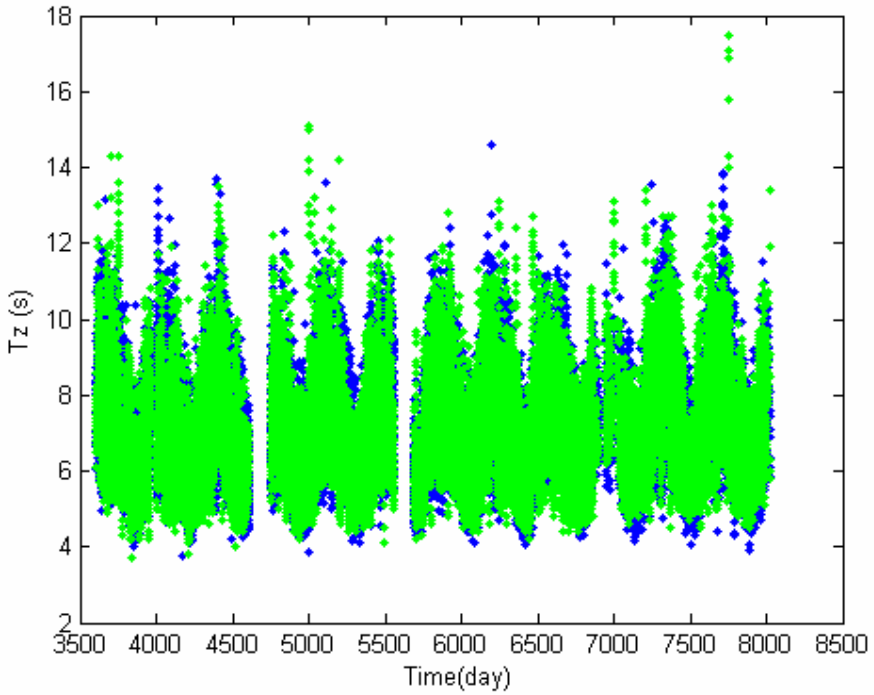


Fig. 4.18. Test data, observed during mid-1986-99 (green), and the corresponding simulated data (blue).

The scatter plot of the observed T_z s of 1978-99 and the corresponding predicted values obtained from the distributions means is shown in Fig.4.19. The Q-Q plot of Fig. 4.20 shows that the samples of the observed and predicted T_z's could be considered from the same distribution type.

4.3.6 The Networks Efficiency

RMSE, R, and SI, calculated according to Eqs. (31) through (33), were used as measures of the efficiency of the trained networks and also for comparing the different values obtained as forecasts for T_z. These indices for the **long term** simulation shown in Fig. 4.14 i.e. for 21 years beginning from 2 Jan 1978 is as follows: RSME = 1.46, R = 0.46, and SI = 0.21. The above RSME and the one appearing in Fig. 4.14 may be rather confusing and need clarification. RSME = 0.72731 shown in Fig 4.14 is the RSME of the simulated data and the observed ones for a selected period starting from 1 Jan 1978, but RSME = 1.47 refers to the root mean square error of the observed H_S's in the whole period 1978-99 and the simulated H_S 's for the period given some T_z's and H_S's.

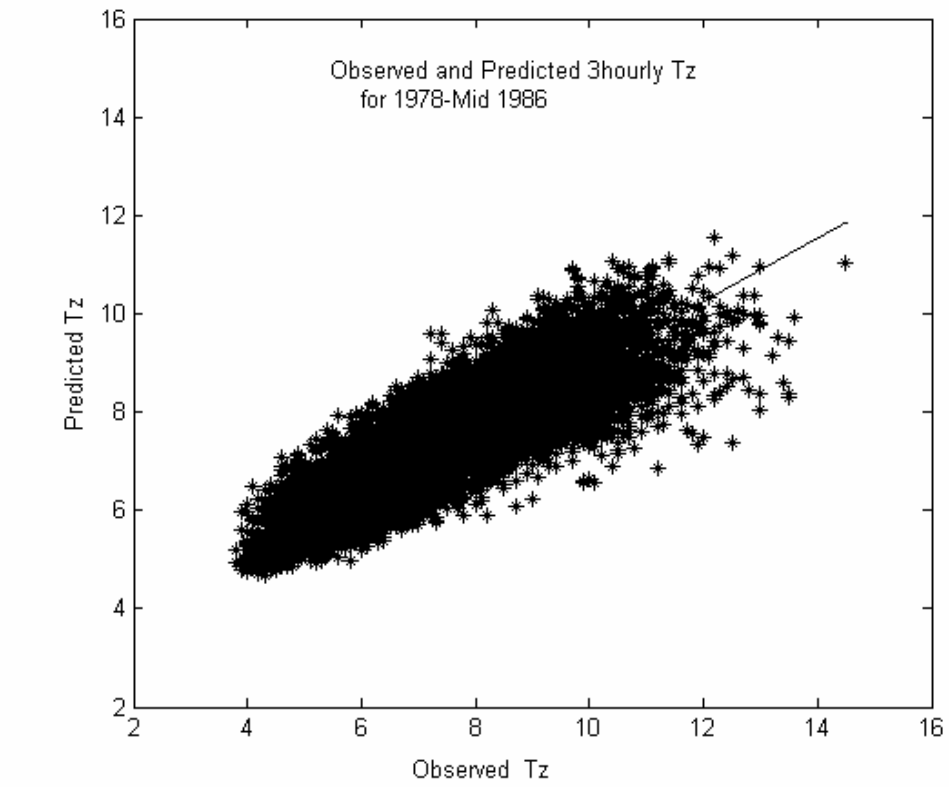


Fig.4.19 The scatter plot of the observed 3-hourly T_z s in the period 1978 through mid 1986 versus the predicted T_z s obtained from the mean of the distributions.

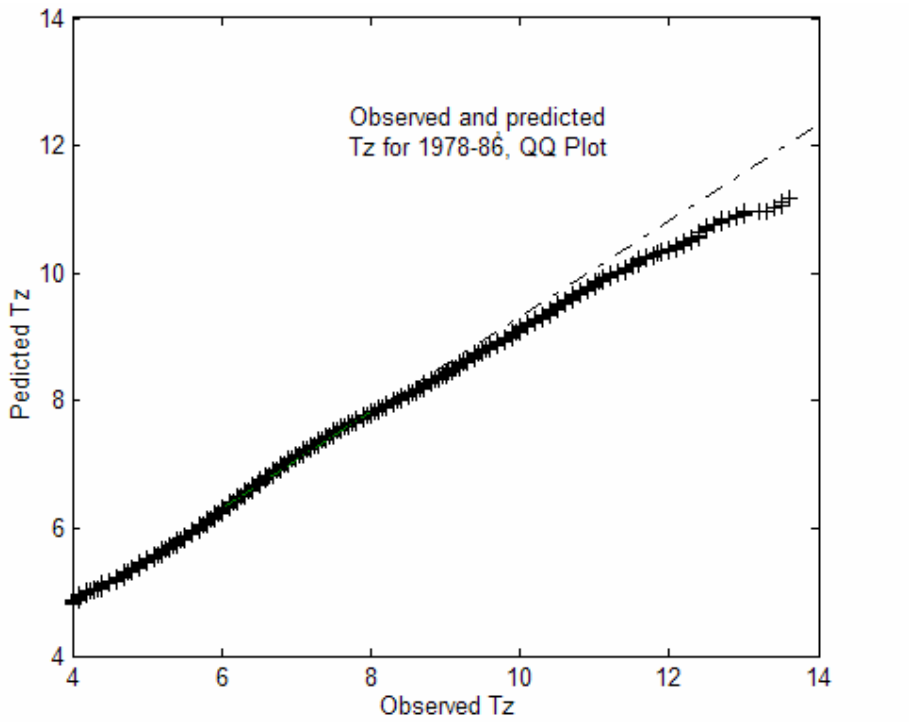


Fig. 4.20 The 3-hourly T_z s observed from 1978 to mid 1986 versus the predicted T_z s, QQ plot.

In this simulation the necessary T_z's and H_s's for the input of the trained networks to simulate the T_z's are simulated data except for simulating the first T_z i.e. the T_z of the sea-state occurring at 24 AM 2 Jan 1978 whose 4 of 5 necessary inputs are the observed ones. The fifth input i.e. the H_s of subsequent immediate 3-hourly interval is always a generated random variate. Having simulated the T_z's of the above period sea-states for a sufficient number of times, long term forecasting the T_z of a desired time in the period could be done by computing the most probable value or by calculating the mean of the simulated T_z's for the desired time, then the better forecast could be selected after calculatiing a performance index such as RSME.

As an illustration for **short term** forecasting**,** all of the T_z's of the sea-states of all exactly 3 hours ahead were forecasted for the period 1978-1999. For each T_z in the target vector the following 3 different values were obtained as 3 different forecasts for T_z:

1) the mode or most probable value of its corresponding hepta-parameter spline distribution whose parameters are readily estimated from the outputs of trained networks,

2) the mean of the distribution, and

3) a random variate generated from the distribution.

Table 4.3 compares the 3 forecasts by showing the indices of efficiency calculated using the predicted T_z's and the the ANNs target vector, which consists of the observed T_z's. According to the indices of efficiency recorded in the table, the mean of the corresponding conditional distribution of the T_z of the sea-state of a desired time is a better forecast for the T_z.

Table 4.3 The indices of efficiency of short term simulation of the T_z of the subsequent immediate 3-hourly sea-state by 3 methods for the period 1978- 99.

Forecasting method for T_z	RMSE(s)	R	SI
Most probable values of the disributions	0.0488	0.8609	0.1654
Mean of the distributions	0.0415	0.8776	0.1406
2 random generated T_z's	1.1145 1.1140	0.6859 0.6853	0.1578 0.1577

4.4 Empirical Joint Distribution of H_s and T_z

"In the design of offshore structures, a consistent method of choosing the height and period for the adopted design wave may be crucial for a reliable design. With this objective, the joint pdf of the height and period is considered "(Haver, 1987). As an illustration suppose the joint distribution of H_s & T_z of the sea-state at 6 AM 1 April 2004 given the H_s's and T_z's of 1 Jan 1978 is to be found. A two-dimensional histogram of relative frequency of a sample of $H_s(t_i)$ & $T_z(t_i)$ approaches their joint distribution if the sample size is largish. The H_s's & T_z's of the period 0 AM 2 Jan 1978 through 6 AM April 2004 could be simulated by using networks trained for the significant wave height and meanzero-up-crossing period. The simulated data is a sample of conditional joint distribution of

$$\begin{cases} \text{Hs (0 AM 2 Jan 1978)}, & \text{Hs (3AM 2 Jan 1978)}, \ldots, \text{Hs (6 AM 1 April 2004)}, \\ \text{Tz (0 AM 2 Jan 1978)}, & \text{Tz (3 AM 2 Jan 1978)}, \ldots, \text{Tz (6 AM 1 April 2004)}, \end{cases}$$

given the observed 3-horly H_s's & T_z's of 1 Jan 1978. Indeed the simulated sea-state characteristics (H_s and T_z) of 6 AM 1 April 2004 sea-state are random variates from the

conditional joint distribution of H_S and T_z related to 6 AM 1 April 2004 given the observed 3-hourly sea–states of 1 Jan 1978; since the time difference between these two dates i.e. the starting date and the desired time is rather long, therefore it would be quite logical to neglect the information conveyed by sea–states of the starting day and assume that the variates have been drawn from the joint distribution of sea-state characteristics of 6 AM 1 Jan 2004 instead of being variates from the conditional joint distribution. Repeating the above simulation for a sufficiently large number of times would yield a large sample of the characteristics of 6 AM 1 April 2004 sea-state and then their two-dimensional histogram of relative frequency, that represents approximately their joint distribution, can be developed.

4.5 Simultaneous Prediction of a Sea-State Characteristics

Two illustrations for the above procedure, which could simulate and predict the sea-states of 3 hours, 6 hours, 1 day, 1 week etc. from now, follow. The closer the time, the more accurate the predicted value.

4.5.1 1St Illustration

As the first illustration, suppose the sea-state of 9 AM 2 Jan (day = 1.5 of the Gregorian Calendar) for the region in NE Pacific is to be predicted. Table 4.4 shows the observed 3-hourly H_S 's of 24 hours before the desired time and also two necessary T_z's, all to be given as input to the MATALB code used in the prediction of the sea-state. To predict the sea-state, after inputting the data in Table 4.4 to the code, the sea-state was predicted 10 times by running the code which estimates the parameters of the hepta-parameter spline distributions from the outputs of the corresponding trained neural networks and generates the H_S and T_z as random variates from the corresponding distributions. The results are shown in Table 4.5. The mean of the ten times simulations for the H_S & T_z are 3.91 m and 7.57 s respectively and the mean of the observed H_S & T_z are 3.89 m & 7.67 s respectively for 9 AM 2 Jan. For practical purposes the number of runs should be increased to gain more accurate predictions.

4.5.2 2nd Illustration

As the second illustration, given the data in Table 4.4, the sea-state of 24 PM 30 Dec (365^{th} day of Gregorian calendar) for the region where Buoy 46005 operates has been simulated in a similar way. Table 4.6 shows the results. A suitable prediction for the sea-state

given the H_s's of the first day of the year could be the sea-state characterised by the mean of the H_S's and the T_z's in Table 4.6 respectively i.e. H_S = 5.02 m and T_z = 8.32 s.

Table 4.4 Inputs for sea-state prediction.

Time		H_s (m)	T_z (s)
9 PM	1st Jan	3.1	6.7
6 PM	1st Jan	3.2	6.8
3 PM	1st Jan	3.0	
12 Noon	1st Jan	2.6	
9AM	1st Jan	2.7	
6AM	1st Jan	2.0	
3AM	1st Jan	1.7	
00AM	1st Jan	1.8	

The mean of observed values of the H_s & T_z for 24 PM 30 Dec during 1978-1999 were 3.46 m & 7.52 s respectively. As it is evident from these two illustrations the closer the time of the sea-state being predicted to the time of the day whose sea-states are being given as input, the more accurate the prediction.

Table 4.5 Ten times generation of the sea-state for 9 AM 2nd Jan.

Run #	H_s (m)	T_z (s)
1	2.5	7.6
2	3.1	7.9
3	3.7	7.4
4	4.3	7.9
5	4.1	7.0
6	3.8	8.3
7	5.4	7.3
8	2.8	6.6
9	3.9	8.3
10	5.5	7.4

Table 4.6 Twenty times generation of the sea-state of 24 PM 30 Dec.

Run #	H_S (m)	T_z (s)	Run #	H_S (m)	T_z (s)
1	4.2	8.4	11	7.2	8.4
2	9.3	10.4	12	6.8	9
3	4.7	7.9	13	2	6.8
4	4.3	7.6	14	4.5	8.7
5	2.7	6.8	15	4.5	8
6	4.2	7.5	16	3.5	6.4
7	1.1	6.7	17	1.1	6
8	6.9	8	18	7.3	11.1
9	9.1	9.1	19	6.1	11
10	5.8	10	20	5.1	8.5

4.6 Description of the Attempt to Forecast H_S Directly from ANNs

It is worth stating that before creating and training the networks used to estimate the parameters of the conditional distributions of the H_S and the T_z of a future sea-state, as a first attempt some 55 different back-propagation feed-forward networks having 1 hidden layer were trained for predicting wave heights in the NE Pacific based on the model described in Eq. (1). Mean square error (MSE) between the forecast and the observed values was the error function used in training and its value was always around 0.15, which is equal to the variance of the residuals in Eq. (1). None of the 55 ANNs gave a good output as the forecast for significant wave height. The number of nodes (neurons) of the hidden layer of most of the ANNs was chosen according to Eq. (3) of Chapter 2. Although the network whose input consisted of the residuals of Eq. (1) resulted in good forecasts, but the attempt to train some suitable networks for the residuals gave no acceptable result. Therefore this attempt failed.

Network architecture

Some of the network architectures tested are :

1× 3 × 1	4× 9 × 1	3× 7 × 1	2× 5 × 1	4× 9 × 1	6× 13× 1
7× 15× 1	6× 30× 1	1× 5 × 1	7× 18× 1	8× 15× 1	9× 30× 1
5× 11× 1	4× 15× 1	3× 20× 1	2× 50× 1	2× 5 × 1	

Network type

Most of the networks used were feed-forward multilayer networks.

Transfer function

logsig for the hidden layer and *purelin* for the output layer (see Fig 2.5 of Chap. 5).

Training algorithm

The networks were trained by some variations of back-propagation learning rule such as Levenberg-Marquardt back-propagation (*trainlm*), Resilient back-propagation (*trainrp*) and gradient descent back-propagation (*traingd*). *Trainlm* algorithm updates the weight and bias values according to Levenberg-Marquardt optimisation. *Trainlm* algorithm is able to obtain better MSEs than any of the other algorithms tested. *Traingd* is an algorithm that updates the weight and bias values according to gradient. *Trainrp* is another algorithm, which is the fastest algorithm on pattern recognition problems.

Target

The target of the networks was a vector consisting of maximum 150,000 values of the left hand side of Eq. (1) i.e the natural logarithm of the observed 3-hourly H_s's for some 22 years starting from 1 Jan.1978. Some of these data, sometimes $\frac{1}{8}$ of the data, were allocated for testing the output of the trained network.

Input

Depending on the number of nodes chosen for the input layer, one, two or more inputs were given to the networks; some of the inputs given to the networks are the followings.

1)the time of the observed sea-sate, calculated from

$$\text{time} = x\text{-floor}(x), \tag{37}$$

with preprocessing according to Eq. (37):

$$\text{time} = \text{time} \times 0.8 + 0.1. \tag{36}$$

where *x is* a number between 0 and $22 \times 365.25 = 8035.5$ with step 1/24 representing the day of occurrence of the sea-state in the 22-year period.

2) year number.

3) xannual = $c_1 \sin(2\pi x/365.25) + c_2 \cos(2\pi x/365.25)$. (39)

4) $c_5 x$.

5) *day* in the form of a number between 0- 365.25 divided by 365.25 to lie in the interval 0-1 in order to increase the efficiency of training. This input was computed by MATLAB:

$$day = mod(x(1:132440,:),365.25)/365.25. \quad (40)$$

6) *season* in the form of 4 fuzzy membership values calculated as shown in Appendix A.

7) The H_s's of one year.

Figure 4.22 shows the H_s's obtained directly from a feed-forward back-propagation network of size $6 \times 13 \times 1$ (see Fig. 4.21) and the observed H_s's. The input matrix of the networks consisted some vectors including vector, $c_5 x = 0.000219362x$, day vector, time vector, *xannual* vector calculated from Eq. (33), and the significant wave heights of 1978. The networks were trained with *trainrp* and the transfer functions of the hidden and the output layers were *tansig* and *purelin* respectively. This ANN was also trained with *traingd* algorithm but much worse results were obtained.

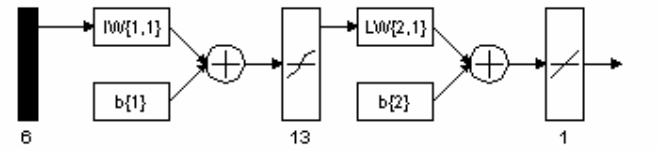


Fig 4.21 A typical architecture of the ANNs used for the direct simulation of H_s.

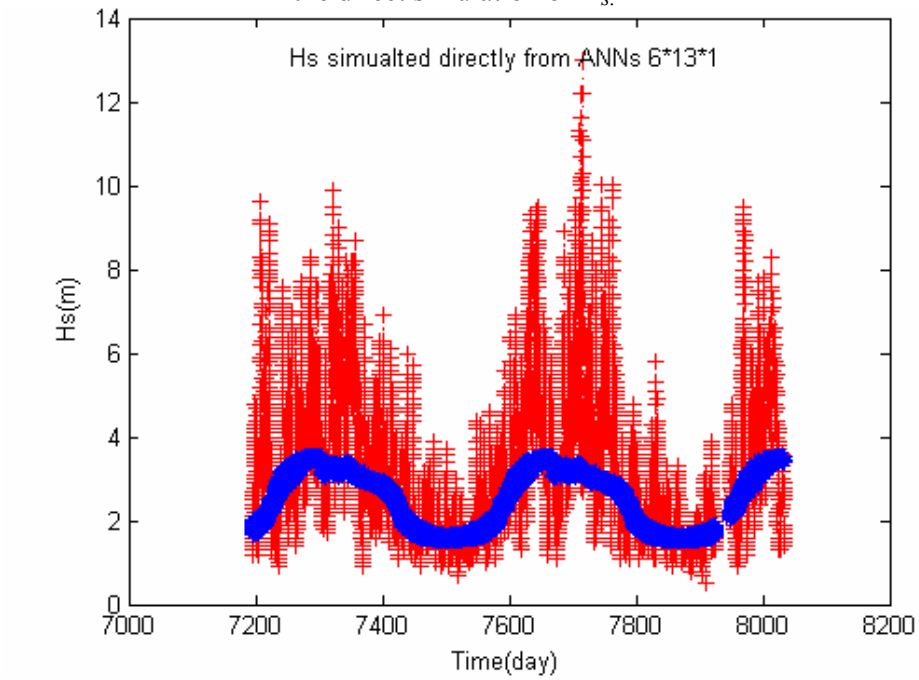


Fig 4.22 The H_s's predicted directly from a typical ANN (blue) and the observed H_s's (red).

As well as trying to predict the H_s for all days of a year, it was attempted to train networks for the H_s of a typical month (January) or a season (spring) but none of the attempts resulted in satisfactory forecasts.

4.7 Conclusions

Multi-layer feed-forward ANNs proved to be inefficient in the direct prediction of significant wave heights from neural network outputs given some input to the ANNs but was efficient in estimating the parameters of the hepta-parameter spline distributions proposed for sea-states characteristics and thereby in the simulation of sea-states sequences. Simulated annealing, due to its superior efficiency and speed over gradient descent, was chosen as the preferred algorithm for training the networks. In training the networks, a function based on the likelihood function of MLE method was used as the performance criterion, whilst conventionally other functions such as MSE are used for training ANNs.

The procedure proposed in this PhD thesis for simulating the characteristics of a sea-state is applicable to some other processes having instinct periodical variations versus time in order to simulate a desired random variable and calculate the most probable value and mean value of the process and also to do the extreme value analysis.

References (in alphabetical order)

Anderson, C. W., Carter D. J. T., and Cotton, P. D., 2001.
Wave Climate Variability and Impact on offshore Design Extreme.
Report prepared for Shell International and the Organization of Oil & Gas Producers.
http://info.ogp.org.uk/metocean/schedules.html.

Ash, R., 1990.
Information Theory.
Dover Publications, New York, ISBN: 0-486-66521-6 p189.

Haver, S., 1987.
On the joint distribution of heights and periods of sea waves.
Ocean Engineering, 14, 5, pp 359-376.

Huang, W., and Foo, S., 2002.
Neural network modeling of salinity variation in Apalachicola River.
Water Reasearch, 36(1), 356-362.

Makarynskyy, Oleg, Pires-Silva, A. A, Makarynska, D., and Ventura-Soares, C., 2002.
Artificial Neural Networks in the forecasting of wave Parameter.
International workshop on wave Hindcasting & Forecasting.
Banff Alberta Canada Oct 21-25 2002.

Smith, R.L., Tawn, J.A., and Coles, S.G., 1997.
Markov chain models for threshold exceedances.
Boimetrika, 84.2, pp 249-268.

Chapter 5

Prediction of safe sea-state for driving vertical piles using Finite Element Method and ANN/SA algorithm

Nomenclature

ANNs	artificial neural networks
a_j	the amplitude of the j^{th} wave component
F	wind fetch (km)
FEM	finite element method
f_j	the frequency of the j^{th} wave component (Hz)
g	gravity acceleration
h	time (hr)
H	wave height
H_{rms}	root mean square of a wave height sample
H_s	significant wave height (m)
MWL	mean water level
m_n	n^{th} moment of energy density function for n = 0,1,2,...
N	total number of zero-up-crossings in the time history
$\overline{S}$	the mean of the stress time history (MPa)
S_e	most likely extreme stress occurred in the pile after h hours (MPa)
S_y	the pile yield stress (MPa)
$S(f_j)$	wave energy density
SA	simulated annealing
SF	safety factor
T	1)mean zero-crossing period for the stress time history of an element, 2)regular wave period

T_z	mean zero-crossing period of a sea-state (s)
$T_{1/3}$	significant wave period
W	wind velocity (km/h)
ε_j	phase angle of the j^{th} wave component
η (*x,y,t*)	sea surface height above the MWL at time *t* for the coordinates *(x& y*)
η_{rms}	root mean square of sea surface elevation
θ_j	the angle between the direction of wave propagation and the X–axis (in counter clockwise direction)
ρ	fluid density
σ	standard deviation of the stress time history of an element (MPa)
σ_η	the standard deviation of the sea surface elevation
σ_η^2	the vaiance of the sea surface elevation
σ_{all}	allowable stress in the pile (Mpa)
ω_j	frequency of the j^{th} wave component in radiant per second (rps)

5.1 Introduction

Vertical piles are nowadays normally used for the installation of fixed jacket structures, offering the potential for cost saving. Vertical piling has in most instances been used for the piles for which the pile driving operation takes place well below the water surface. However, due to poor foundation conditions, the piles may extend well into the wave zone during the installation phase. Consequently, the dynamic response of these long unsupported vertical piles during installation becomes an important consideration. In this chapter vertical piles have been studied with the aim of identifying the range of sea-states suitable for safe pile driving operations. The results of a study into the behavior of vertical pile stickup just after stabbing but before driving have been reported. Pile configuration, including the non-linear foundation and the gap between the pile and the pile sleeve shims have been modeled using the finite elements (FE) facilities within ABAQUS. Figure 1 shows the configuration of such a pile. The lower end of the pile is embedded in soil and restrained from lateral movement by the soil and pile sleeves. The pile supports the weight of the pile-driving hammer at the top and is

subject to the wave and current loads as well as self-weight and buoyancy effects along its support length.

Dynamic analyses of the system for various sea-states characterized by significant wave heights and mean zero-up-crossing periods and modeled as a combination of several airy wave components, have been performed. Repeating the above procedure generated a table of safe and unsafe sea-states for pile driving operations. If, using the method described in Chapter 4, the prediction for a future sea-state is repeated N times from which n times proved to be safe then it could be said that the predicted sea-state is safe with a probability of $100(n/N)$%.

5.2 Modelling of the Pile

The pile being studied (Photo 5.1) is 72 m long with 8 cm wall thickness and an outer radius equal to 91.45 cm. The finite element model of the pile(see Appendix D) consisted of a number of two-node elements PIPE 21 beam (ABAQUS, 2004). The foundation was represented by non-linear spring elements. GAP elements were used to represent the gap between the pile and its sleeves.

The design and analysis of offshore platforms must be done taking into consideration many factors. An essential part of the design of offshore structures is the soil investigation, as it is the soil that ultimately resists the enormous forces and moments imposed by the presence of the platform in the hostile ocean environment. Generally, the seabed soils are clay, sand, silt or a mixture of these; however, calcareous soils are also common in some location.

Photo 5.1 A typical pile.

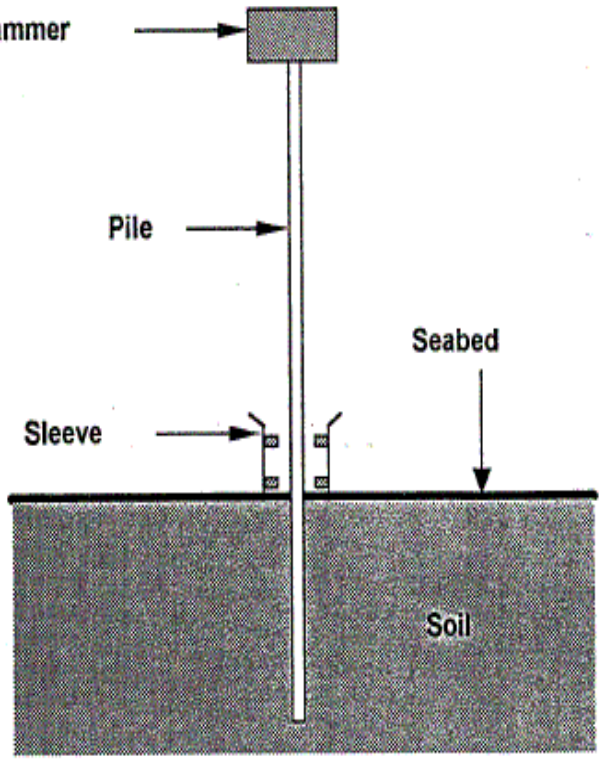


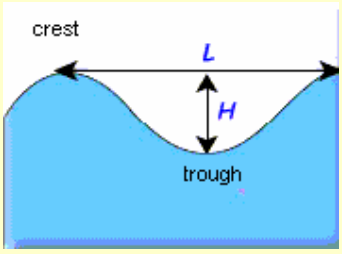
Fig. 5.1 Pile installation (after Djahansouzi &Yasseri, 1994).

5.3 Wave Types and Wave Theories

Waves can be classified as regular and random waves. Ocean waves are random and more complex in form than regular ones. Actually a large number of wavelets of different heights, periods and directions have been superimposed on one another leading to the confusing appearance of the sea water surface.

5.3.1 Regular Waves

The term regular refers to unidirectional train of waves with constant amplitude and frequency and hence constant length. If steepness = $\frac{amplitude}{length}$ is sufficiently small, the waves are said to be linear or Airy waves (Barltrop, 1998). Although regular (periodic) waves, which are identical waves following each other, are not found in real seas they can closely model some swell conditions. The characteristics of a regular wave with no current are described in Fig. 5.2.



Length (L): Distance from equivalent points on a wave (e.g. crest to crest)

Height (H): The vertical distance from the crest to the trough. Amplitude = $a = H/2$. Ocean wave heights are assumed to have a Rayleigh distribution.

Period(T) : The time that a complete wave takes to pass by a fixed point. Frequency = 1/Period

Surface elevation (η): The height of the water surface above mean water level (MWL). Sea surface elevation is assumed to be normally distributed with zero mean & variance σ_η^2.

Water depth : The depth of water from MWL

Clerity : The propagation of wave crests = L/T

Fig. 5.2 Some specifications of regular waves.

5.3.2 Irregular (Random) Waves

Waves in the real environment are sometimes referred to as random or irregular. They can effectively be modeled in terms of regular wave trains combining in random phase. An important advantage of such modeling is that the complete description of the wave motions including surface kinematics for the regular components can be linearly superposed to provide corresponding complete descriptions of the combined wave motion (Barltrop, 1998). Ocean waves, which are irregular, are supposed to follow a Rayleigh distribution to be discussed in Sec. 6.3. The statistical properties of random waves in a sea may be assumed to be approximately constant for short periods e.g. one to three hours (Barltrop, 1998).

5.3.2.1 Classification of ocean waves

Ocean waves are undulations of the water surface resulting from the transfer of energy. The disturbance is propagated by the interactions of disturbing forces (e.g. wind) and restoring ones(e.g. gravity). Water waves are classified by the disturbing and restoring forces involved. The energy in most ocean waves originates from the wind blowing across the surface of water, thus creating wind waves. Large tsunami or seismic sea waves are generated by earthquakes, volcanic eruptions or large marine landslides. On the other hand tides, largest of all ocean

waves, result from the combined gravitational force exerted on the oceans by the sun and the moon.

Wave conditions may also be classified as *sea* or *swell.* A *sea* is an area of water on which waves are being generated by the wind. A *swell* is a pattern of waves, which are no longer being acted on by the wind either because the wind is dropped or because they are propagated away from the generating area (Barltrop and Adams, 1991**)**.

5.3.2.2 Characteristics of a sea-state

The random nature of the surface of the oceans means that it can only be quantified statistically. The statistics do not change very much over an interval of about 3 hours and the sea-state is characterized by *significant wave height* and *mean zero-crossing period*, as defined below, and some times *peak period (Tp)*. A good discussion of different wave measures is given in IHAR (1986).

5.3.2.2.1 Significant wave height

The significant wave height denoted by H_s or sometimes by $H_{1/3}$, is a common representative height for a sea-state. H_s as measure of general sea-state is an important value for many reasons. The practice of reporting sea conditions on a significant wave height basis is wide–spread and most observed and recorded wave data are presented in the form of H_s and also most energy spectrum analysis are related to H_s. The effects that irregular seas have on many types of fixed and floating structures and on various shore processes such as littoral transport have been related to H_s with sufficient accuracy for many engineering applications. H_s is the average height of the waves that observers typically report. It is in fact biased towards the higher waves. Originally, it was defined as the mean value of the highest one third of a sample of the sea-state wave heights, which was thought to give about the same values as an experienced seaman's estimate by eye of "the average wave height". An example for such calculation of significant wave height is illustrated using the data in Table 5.1. To calculate the significant wave height, the highest one third of the 18 waves i.e. the first 6 highest waves with order numbers 1 through 6 are selected for the calculation; yielding the mean of 4.56 m as H_s. According to modern definitions, H_s is defined in terms of the variance of the sea surface elevation (Anderson et al. 2001) as such:

$$H_s = 4\sqrt{\sigma_\eta^2}\ , \qquad (1)$$

where σ_η^2 is the varinace of the sea surface elevation.

Table 5.2 The heights and periods of a wave record.

Wave No. -	Wave height H(m)	Wave Period T(s)	Order no.	Wave No.	Wave height H(m)	Wave period T(s)	Order No.
1	4.40	4.2	**4**	10	2.37	4.3	
2	2.05	8.0		11	1.03	6.1	
3	2.52	6.9		12	1.95	8.0	
4	4.50	11.9	**3**	13	5.0	7.6	**1**
5	3.20	7.3		14	3.22	7.0	
6	1.87	5.4		15	2.08	8.2	
7	1.90	4.4		16	3.89	8.0	
8	1.00	5.2		17	4.95	9.0	**2**
9	4.39	6.3	**5**	18	4.10	9.2	**6**

The H_s definition using Eq.(1) stems from the H_s estimate derived from Rayleigh distribution assumption for ocean wave height, which will be described in Sec. 6.3. The average of highest $(1/n)^{th}$ of a distribution having *f(x)* as its density function is obtained from $n\int_{x1/n}^{\infty} xf(x)dx$ where $x_{1/n}$ is such that the probability of exceeding it is equal to $1/n$. Recalling that H_s was originally defined as the mean value of one third highest waves and assuming Rayleigh distribution for wave height with an estimate of 8 σ_η^2 or H_{rms}^2 for its parameter(b^2) would yield

$$H_s = 3\int_{x1/3}^{\infty} xf_{Rayleigh}(x)dx \cong 1.416\,H_{rms} \cong 4\sqrt{\sigma_\eta^2}\,, \qquad (2)$$

where σ_η^2 is the sea surface elevation variance.

Eq.(1) is a useful for formula for calculating significant wave height when there is a measured time series of surface elevation (η). H_s is also estimated from the zeroth spectral moment *(m_0)* i.e. $H_s \cong 4\sqrt{m_0}$ as will be discussed in Sec. 5.4. The derivation of this formula is shown in many references including Sorenson (1997) and WMO (1998).

The statistics of the wave height within a sea-state are discussed by many authors including Cartwright & Longuet-Higgins (1956) and Ochi (1998). Some useful relationships

between various statistical measures of wave height and H_s are given in a table by Barltrop and Adams (1991), reproduced in Table 5.3.

In practice it is common to assume that the maximum observed height in a sea-state is taken 1.8 H_s or 1.86 H_s; this is based on the assumption that mean zero-up-crossing period is nearly 10 seconds i.e. $T_z \cong 10$ s for 3-hourly sea-states and therefore the number of waves in a 3-hourly sea-state would be $N = \frac{3\times 3600}{Tz} \cong 1000$ waves. Substituting N = 1000 in Eq. (6) would give the most probable (mode) extreme height of nearly 1.86 H_s. Therefore assuming that wave height (H) follows a Rayleigh distribution whose probability density function and

Table 5.3 Relationship between statistical measures of wave height and significant wave height (after Barltrop & Adams, 1991)

Description	formula	relationship with *Hs*	Eq. No.
Standard deviation of free surface [(a)]	$\sqrt{m_0}$	0.250 *Hs*	(3)
Mode height [(a)]		0.499 *Hs*	
Median height [(a)]		0.588 *Hs*	
Mean height [(a)]		0.626 *Hs*	
Root mean square [(a)]	$\sqrt[2]{\frac{\sum H^2}{n}}$	$0.705\ Hs \cong \sqrt{\frac{1}{2}}\ Hs$	(4)
Power weighted mean wave heights [(a)] (These are often useful for fatigue studies)	$\sqrt[3]{\frac{\sum H^3}{n}}$	0.776*Hs*	(5)
	$\sqrt[4.5]{\frac{\sum H^{4.5}}{n}}$	0.869*Hs*	
	$\sqrt[6]{\frac{\sum H^6}{n}}$	0.952*Hs*	
Most probable highest wave in 1000 waves(Eq. (6)) [(a)]		1.89 *Hs*	
Average highest wave in 1000 waves (Eq. (7)) [(a)]		1.93 *Hs*	
Maximum height in a sea-state [(b)]		1.8 *Hs*	

References of Table 5.3

(a) Rayleigh assumption,
(b) Wiegel(1949).

distribution function are given in Chapter 6 the prediction for the most probable (mode) extreme height is equal to 1.86 H_s and it could be shown the mean of the extreme height in

1000 waves is equal to 1.93 H_s (Barltrop and Adams, 1991). It is useful to be able to estimate the highest wave in a 3-hourly sea-state. If an estimate is required for the highest wave in N waves, the modal (most probable) value is given by Eq. (6) and the mean of the highest wave is obtained from Eq. (7) (Barltrop and Adams, 1991):

$$\text{Mode}(H_{Max\,N}) = Hs\sqrt{\frac{\ln N}{2}}\ , \tag{6}$$

$$\overline{H}_{Max\,N} = (\sqrt{2\ln N} + \frac{0.5772}{2\sqrt{\ln N}})\frac{2Hs}{4.005}\,. \tag{7}$$

5.3.2.2.2 Mean zero up-crossing period

The mean zero up-crossing period denoted by T_z is another characteristic of a sea-state. The customary practice in wave analysis is to utilize either the *zero-up-crossing* method or *zero-down-crossing* method as the standard techniques for defining wave. These two methods are described by Goda (2000) and Fig. 5.3 gives a feeling of the methods. The mean zero-crossing period is normally obtained from the mean time between up-crossings of water elevation, however the down-crossing method has also been used. Since before every up-crossing there must be a down-crossing apart from small errors caused by possible additional up or down crossing at the end of a record, both methods would presumably give the same result. In other words since in any reasonably long record the number of up-crossings must be the same as the number of down-crossing (apart from ±1 because of where the record starts and ends) so, mean zero-up-crossing period and mean zero-down-crossing period values should be equal.

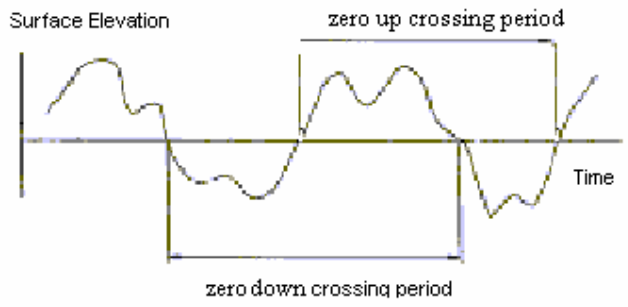


Fig. 5.3 down-crossing and up-crossing periods.

In practice the period is often routinely estimated from $T_z = \sqrt{m_0 / m_2}$ (when frequency is in Hz) or from $T_z = 2\pi \sqrt{m_0 / m_2}$ (when frequency is in rps), where m_0 and m_2 are the zeroth and second wave spectral moments. Description of the moments will be given in Sec. 5.4 and the proof of the formula is given by many references including Ochi (1998) and Goda (2000). The above formulae for estimating T_z from m_0 and m_2 is based on the assumption that the sea surface elevation follows a Gaussian process; i.e. can be represented by a normal distribution of zero mean and varaince σ_η^2. The T_z calculated from the formula and from a count will not be exactly the same due to some reasons including the fact that the distribution is not precisely Gaussian.

$T_{1/3}$, the H_s associated statistical parameter defined below, is sometimes introduced as the representing wave period. Recalling the definition of H_s as the arithmetic average of the highest one third of the wave heights in a wave record, the average period of the highest one-third of the waves in this record is called significant wave period and is denoted by $T_{1/3}$.

$T_{1/3}$ and the estimate of T_z from the formula will not be the same. Since T_z is the average of all the up-crossing periods and since the individual wave heights and periods are positively correlated, therefore $T_{1/3}$ is obviously greater than T_z. From the analysis of field wave data $T_{1/3}$ has been reported to be $(0.9\sim 1.4)\overline{T}$ where $\overline{T}$ is the mean of all of periods in a wave record, but T_z may become 0.70 of $\overline{T}$ (Goda, 2000) or $0.92\overline{T}$ (Ochi, 1998). For example based on the data of Table 5.1 the value of $T_{1/3}$ = 8.03 and since the table gives $\overline{T}$ =7.06 then $T_{1/3} = 1.1374\overline{T}$.

Relationship between H_s and T_z

As the wind blows over the sea, H_s increases and T_z also increases so that $\frac{Hs}{1.56Tz^2}$ is around $\frac{1}{10}$ or $\frac{1}{15}$; when the wind reduces the H_s reduces but the T_z does not. Nearby and faraway storms produce increases of H_s and T_z but $\frac{Hs}{1.56Tz^2}$ will be much smaller (Barltrop and Adams, 1991).

5.3.2.3 Superposition

If a time history record is made of the waves passing a point in the ocean, the resulting trace can in principle be viewed as a summation of a large number of regular sinusoidal wavelets each superimposed on the other (Fig. 5.4). The profile of a random wave composed of different components of different amplitudes a_j coming from divergent directions θ_j with various frequencies ω_j may be written as follows (Ochi, 1998):

$$\eta(x,y,t) = \sum_j a_j Cos[\frac{\omega_j^2}{g}(xCos\theta_j + ySin\theta_j) - \omega_j t + \varepsilon_j], \qquad (8)$$

where

$\eta(x,y,t)$	the height of sea water surface for the coordinates x and y at time *t* above the MWL,
a_j	the amplitude of the j^{th} wave component, $0 < a_j < \infty$
g	gravitational acceleration
ω_j	the frequency of the j^{th} wave component in radiant per second, $0 < \omega_j < \infty$
θ_j	the angle between the direction of wave propagation and the X – axis (in counter clockwise direction), $-\pi < \theta_j < \pi$
ε_j	phase angle of the j^{th} component, $-\pi < \varepsilon_j < \pi$.

5.3.3 Wave Theories- aim, types

The aim of a wave theory is to obtain expressions for surface elevation, wavelength, celerity, water particle velocity, energy, power etc. as a function of wave height, period and water depth. A historical overview of wave theories is mentioned by Goda (2000). Surface wave theories have been developed through works of Lagrange, Airy, Stokes, Kelvin, Rayleigh and Lamb (1945). Lemehaute (1976) and Lighthill (1978) made some contibutions. There are well-known wave theories including Stokes (nonlinear) theory, Airy (linear) theory, stream function theory and solitary wave theory to predict and describe wave behavior. Summaries of wave theories can be found e.g. in Barltrop & Adams (1991). Some points about three wave theories, with emphasis on Airy theory, due to being selected for using ABAQUS are followed:

- **Stokes (non linear) wave theory**: trochoidal waves
  - can be used for deep-, intermediate- and shallow-water waves
  - Stokes 5^{th} order frequently used for fixed structures, seldom used for the analysis of floating ones (Barltrop, 1998)

- mathematically complex
- takes into account the effects of wave height on velocity
- describes more accurately orbital velocity asymmetrie

o **Solitary wave theory**:
- an isolated crest moving in shallow water
- none oscillatory progressive waves

o **Airy (Linear) Wave theory**: sinusoidal waves
- first proposed by Laplace in 1776 and by Airy in 1845
- most accurate for low amplitude waves in deep water
- normally used for floating structure design (Barltrop, 1998)
- less accurate for predicting wave behavior in shallow water
- most commonly used wave theory because it is the least mathematically complex
- does not take into account the effects of wave height in determining wave velocity
- Some kinematic and dynamic properties of Airy (linear) waves are listed in Table 3A.1 of Barltrop(1998).
- a description of the theory is in Sec. 6.2.2 of the ABAQUS theory manual

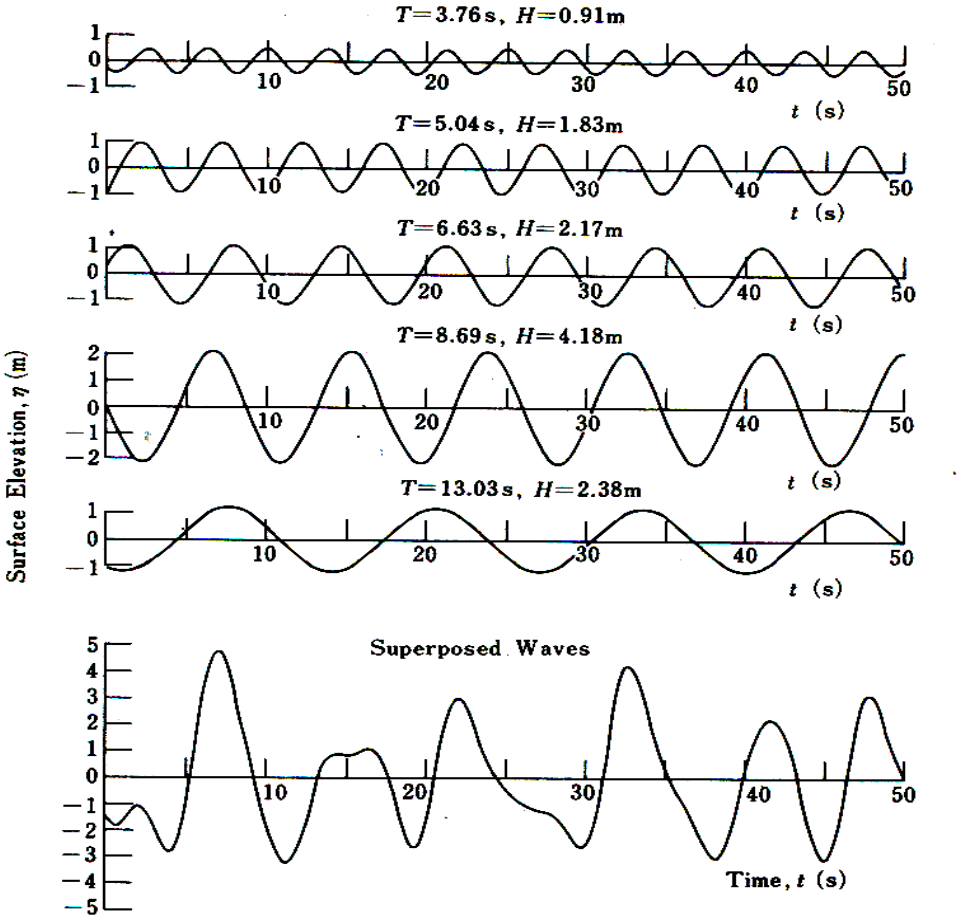


Fig 5.4 Random seas, as a summation of frequency components (after Goda, 2000)

Since design criteria depend on wave energy, it will be briefly discussed here. The wave energy is composed of two parts: potential and kinematic energy. The potential energy is due to the deviation of its profile from the mean water level. For Airy waves it could be shown (e.g. Sorenson, 1997) that its value per unit area, E_p, is equal to:

$$E_p = \frac{1}{16}\rho g H^2 \ (\mathrm{J/m^2}), \tag{9}$$

where

H	Airy wave height,
ρ	water density,
g	gravity acceleration.

The kinematic energy exists because of the oscillatory motion of the particles. For Airy waves the kinematic energy per unit area, E_k, is calculated from Eq. (10) (Sorenson,1997) :

$$E_k = \frac{1}{16}\rho g H^2 \ (\mathrm{J/m^2}). \tag{10}$$

Hence for Airy waves, the potential and kinematic energies are equal.

The total energy per unit area in airy wave theory, E, is given by

$$E = E_k + E_p = \frac{1}{8}\rho g H^2 \ (\mathrm{J/m^2}), \tag{11}$$

or

$$E = \frac{1}{2}\rho g a^2 \ (\mathrm{J/m^2}), \tag{12}$$

where $a = \frac{H}{2}$ is the Airy wave amplitude.

ABAQUS/AQUA has the option to use some wave theories including Stokes and Airy theories. Airy theory was selected for the present research because of being most accurate for low amplitude waves in deep water. Linear Airy wave theory is generally used when the ratio of wave height to water depth is less than 0.03, provided that the water is deep i.e. ratio of water depth to wavelength is greater than 20 (ABAQUS, 2004). This theory is normally used for floating structure design. The software requires sea-state to be decomposed into a number of regular waves when Airy theory is used. Decomposing a sea-state into Airy sine waves was done using a wind wave energy density spectrum.

5.4 Wave Energy Spectrums

Ocean waves in the real environment are random or irregular and the source of irregularity is usually the local winds; in other words the random fluctuation of the sea surface is generally attributed to energy transfer from wind to the sea. Sea waves could be assumed

consisting of an infinite number of sinusoidal wavelets with different frequencies and directions. The distribution of energy of these wavelets when plotted against frequency and direction is called the *wave spectrum*. More precisely the wave energy distribution with respect to frequency alone, irrespective of wave direction, is called the *frequency spectrum* whereas the energy distribution expressed as a function of both frequency and direction is called the *directional wave spectrum*. Figure 5.6 gives an example of an irregular wave profile constructed by adding five sinusoidal wave components of different heights and periods.

Since the total (potential and kinematics) energy of a regular wave of amplitude a_j is given by $E = \frac{1}{2}\rho \times g \times a_j^2$ according to Eq. (12), ignoring ρ & g and plotting the energy per unit frequency (energy density) for all of the components versus the corresponding frequency would give a basic spectrum. The diagram shown in Fig 5.5 is a spectral representation of the superposed waves illustrated in Fig. 5.4. This figure shows the way in which the component waves are distributed against the component frequencies.

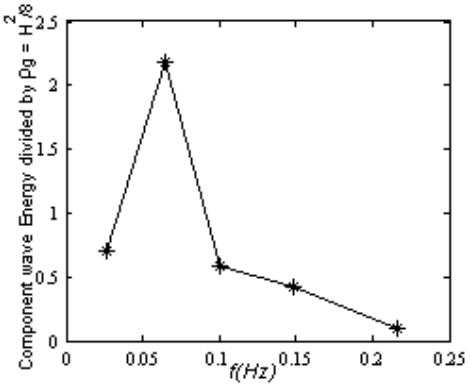


Fig. 5.5 Spectral representation of the irregular profile in Fig.5.6.

The wave spectral density function, which plays a significant role in evaluating the statistical properties of random seas, represents the potential and kinematic energies of random waves. Spectral density functions illustrate the magnitude of the time average of wave energy as a function of wave frequency. In practice there are a large number of frequencies present in the sea and it is more convenient to work with a spectrum that is a continuous function like the one shown in Fig 5.6. Measured spectra often exhibit complex shapes, which may include more than one peak. Theoretical spectra are inevitably simpler although bimodal models do exist (Ochi, 1998). The statistical properties of random waves in a sea may be assumed to be approximately constant for short periods of say one to three hours. During this time, the sea-state may be modeled by a frequency spectrum of water elevation. There are many frequency

spectra proposed for various ocean waves including the work or Bretschenider(1959), JONASWAP (Hasslemann et al.,1973) and the spectrum of Pierson & Moskowitz (1964). However it is believed that in most results the precise shape of the spectrum has only a small effect.

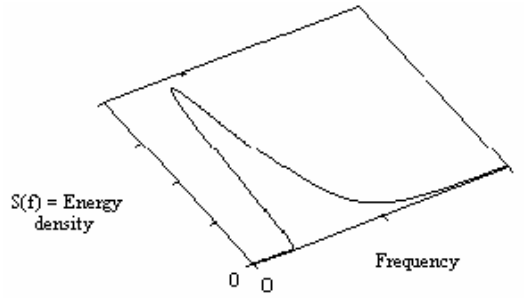


Fig. 5.6 A typical sea surface energy density- frequency spectrum.

It is worth pointing out that spectral models are often a combination of the theoretical model by Phillips (1958), which represents the upper tail, with an empirical lower frequency part. A summary of these spectra can be found in Goda (1990), and Ochi (1998) among others.

The Joint North Sea Wave Project (JONASWAP) spectral form is a well-known spectrum. The JONSWAP spectrum was developed from measurements made in the island of Sylt in the German Byte of the North Sea by Hasslemann et al. (1973). The description of the spectrum is found in many books on ocean engineering including Barltrop (1998). Goda (2000) has also presented a spectrum for wind waves developed over a long fetch such as over the Pacific. When waves are generated by very strong winds over a relatively short fetch such as in the case of North Sea, waves are better represented by the JONSWAP spectrum (Eq. 2.12 in Goda, 2000) with the peak enhancement factor = 1 - 7. For swell coming over a long distance, the peak enhancement factor = 10 is a good approximation. A peak enhancement factor of 3 for storm conditions is advised. In the present thesis Pierson–Moskowitz spectrum was used. This spectrum will be described in Sec 5.4.1.

Definition of spectral moment

The n^{th} spectral moment, m_n, is defined as :

$$m_n = \int_0^\infty f^n S(f)\,df\,, \qquad (13)$$

where

f wave component frequency, and

S(f) wave energy density.

The zeroth spectral moment (m_0), the variance of water surface elevation(σ_η^2), and the root square of surface elevation (η_{rms}) satisfy the following relationship (Goda, 2000, Sec. 2.4.1)

$$\sigma_\eta^2 = \eta_{\text{rms}}^2 = m_0 = \text{ the area under the spectrum.} \qquad (14)$$

Thus, the area under the spectrum is the variance of the sea water surface elevation. This is proportional to the energy per square metre of water surface. The energy is an average over a period of about an hour. The energy over short periods of e.g. a few minutes will be larger or smaller than this average.

The estimation of H_s and Tz from spectral moments

The sea-state characteristics are routinely estimated from the spectral moments as such:

$$T_z = \begin{cases} \sqrt{\dfrac{m_0}{m_2}} & \text{if frequency is in Hz,} \\ 2\pi\sqrt{m_0 / m_2} & \text{if frquency in rps} \end{cases}, \qquad (15)$$

and

$$H_s = 4.01 \times \sqrt{m_0}\ . \qquad (16)$$

The derivation of these formulae has been shown in many references including in Ochi (1998). WMO (1998) also deals with the formulae. It is worth pointing out that Eq. (16) in different references is slightly different. According to Barltrop and Adams (1991):

$$H_s = 4.005 \times \sqrt{m_0}\ , \qquad (17)$$

in Anderson et al. (2000), WMO (1998), and Sorenson (1997):

$$H_s = 4 \times \sqrt{m_0}\ , \qquad (18)$$

and Goda (2000) expresses the relationship as such:

$$H_s = 4.004 \times \sqrt{m_0}\ . \qquad (19)$$

To verify which of the above formula applies for the PM spectrum used in this study, the area under spectrum plotted in Fig. 5.9 was calculated: the area = m_0 = 0.0625. Since H_s =1 for the spectrum plotted in Fig 5.9, therefore $\frac{Hs}{\sqrt{m_0}}$ = 4, hence Eq. (18) is more accurate for our case.

The estimation of H_s from wind characteristics

For a fully developed storm, the significant height of wind-waves (H_s) can be roughly calculated by the following formula (Ben and Gerwick, 2000):

$$H_s = \sqrt{\frac{WF}{30}}, \qquad (20)$$

where

W is the wind velocity (km/h), and

F is the fetch (km).

5.4.1 Pierson-Moskowitz Energy Density Spectrum

Pierson-Moskowitz (PM) energy density spectrum is wildly used for characterising waves in the open sea (Barltrop, 1998). Pierson and Moskowitz (1964) put forward, on the basis of a similarity theory by Kitaigorodskii, some suggestions for deep-water wave spectra. The PM formulation was developed from the analysis of the measured data obtained by some wave recorders in North Atlantic. The analysis was carried out only on selected wave records considered to have been acquired in fully developed seas[1] (Ochi, 1998). PM density spectrum is dealt in many references including Ochi (1998). There are several expressions for PM energy density spectrum. The following formula of the spectrum (Barltop and Adams, 1991), which depends on the sea–state representing parameters i.e. H_s and T_z, has been used in this study to decompose a sea-state into several sinusoidal wave components. The decomposition procedure will be described in the next section.

$$S(f_j) = \left(\frac{A}{f_j^5}\right) \times e^{-\frac{B}{f_j^4}} \qquad 0 < f_j < \infty, \qquad (21)$$

[1] A sea state in which waves have reached its maximum energy.

where

$S(f_j)$ energy density of the wave component (m^2/Hz),

$A = \frac{Hs^2}{4\pi T_z^4}$, $B = \frac{1}{\pi T_z^4}$,

f_j the frequency of j^{th} wave component (Hz),

H_s significant wave height (m),

T_z mean zero-crossing period of a sea-state (s).

Zeroth, first, and second spectral moments are obtained from (Ochi, 1998):

$$m_0 = \frac{A}{4B}, \quad m_1 = \frac{0.306A}{B^{\frac{3}{4}}}, \quad m_2 = \frac{\sqrt{\pi}}{4}\frac{A}{\sqrt{B}}.$$

As an illustration, Fig. 5.7 shows the spectrum for a sea-state with H_s = 1 m and T_z = 3 s.

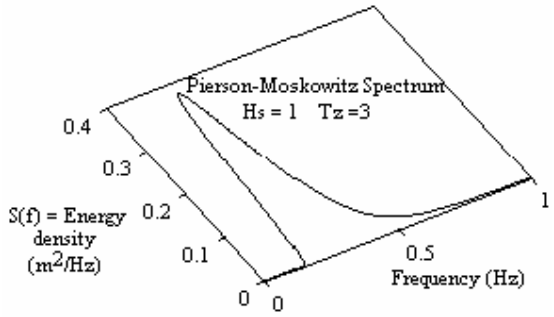


Fig. 5.7 A typical Pierson-Moskowitz energy density spectrum.

5.5 Decomposition of a Sea-State

Airy wave theory was used during the pile analysis with ABAQUS/AQUA because of being most accurate for low amplitude waves in deep water. The software requires the sea wave to be given as a series of sinusoidal wave components to determine the wave load when the Airy theory is used. It was assumed that the PM spectrum represents the sea-states of the location of interest in the North East Pacific. Therefore each sea-state was decomposed into several components using the PM wave spectral formulation described in Sec. 5.4.1. Each component is identified by three parameters that are amplitude, period and phase angle. Using a code written in MATLAB, each sea-state was decomposed into 30 sinusoidal components each having its own amplitude, frequency and phase angle. The phase angles were generated randomly between 0 and 360 degrees by MATLAB command *phase = rand (30,1)×360*. Then for each sea-state the range of frequency (bandwidth), i.e. zero to where the spectrum

approaches zero e.g. S (f) $\cong 0.0001$, was determined and divided into 30 subintervals. The frequency of each subinterval end point was selected as each component frequency. Then, given the frequency of the component, its amplitude was calculated using the following formula given by many references including Barltrop (1998) and Goda (2000):

$$a_j = \sqrt{2S(f_j) \times \Delta f_j}\ , \qquad (22)$$

where

a_j the amplitude of the j[th] component,

f_j the frequency of the j[th] component,

$S(f_j)$ energy density of j[th] component calculated from Eq. (21),

Δf_j $= \dfrac{f_{j+1} - f_{j-1}}{2}$.

Seventy sea-states were selected for the analysis and each of them was decomposed into 30 Airy waves. For example the thirty components of the sea-state with H_s = 1 m and T_z = 3 s are given in Table 5.4.

Table 5.4 Decomposition of a sea-state characterized by H_s = 1, T_z = 3.

Amplitude (m)	Period (s)	Phase angle (deg)
0.0000	27.43	342.0465
0.0000	13.72	83.2099
0.0000	9.14	218.4633
0.0135	6.86	174.9537
0.1006	5.49	320.8676
0.1603	4.57	274.3549
0.1619	.3.92	164.3284
0.1404	3.43	6.6613
0.1159	3.05	295.7066
0.0944	2.74	160.0932
0.0770	2.49	221.5557
0.0634	2.29	285.0973
0.0527	2.11	331.8527
0.0442	1.96	265.7546
0.0374	1.83	63.4558
0.0320	1.71	146.0542
0.0276	1.61	336.7691
0.0240	1.52	330.0856
0.0210	1.44	147.6973
0.0185	1.37	321.7138
0.0164	1.31	20.8409
0.0146	1.25	127.0325
0.0131	1.19	292.7399
0.0118	1.14	3.5501
0.0107	1.10	50.0007
0.0097	1.06	72.9955
0.0088	1.02	71.5398
0.0080	0.98	217.3653
0.0073	0.95	97.9877
0.0068	0.91	71.5731

5.6 Analysis with ABAQUS Software

The analysis was done for 70 sea-states to determine which states are safe for pile driving. For this purposes, applying the weight of the pile driving hammer at the top of the pile and the wave and current loads as well as self weight and buoyancy effects along its support length in the pile's finite element model, ABAQUS was asked to do dynamic analysis for a subset of elements thinking to include the critical one. The analyses were done in time domain covering 300 seconds with time increments of 0.05 s to ensure that the pile response has settled down. The combined (von Mises) stress has been calculated for nearly 300/0.05 = 6000 time increments for all of the elements in a subset of elements.

The pile is initially leaning against the pile sleeve. In the first step of a multi-step analysis the pile self-weight was applied and a non-linear static analysis was performed to determine the equilibrium position. In the next step the weight of hammer was applied. The third step was a non-linear time domain analysis in which the pile set up was exposed to the effect of wave loading.

Airy wave theory was used because of being most accurate for low amplitude waves in deep water. For each sea-state the sea surface was presented using 30 Airy waves. Then ABAQUS/AQUA was used to analyze the pile. This procedure was repeated for a total of 70 sea-states as shown in Table 5.5.

5.6.1 ABAQUA/AQUA

AQUA is a part of ABAQUS software which was used to analyze the pile response using the pile FE model. ABAQUS/AQUA contains some features that are specifically designed for the analysis of beam-like structures installed underwater and subject to loading by water currents and wave action. The ABAQUS/AQUA capabilities allow the application of buoyancy, drag, and inertia loads resulting from submersion in steady current, waves, and wind. AQUA loading can be applied in *STATIC steps and in *DYNAMIC steps.

5.6.2 Pile Response

To determine the pile response during pile driving operations note that, as Fig.5.1 shows, the pile supports the weight of the pile driving hammer at the top and is subject to the environmental loads due to wave and current as well as the self weight and buoyancy effects along its unsupported length.

The von Mises stress was used as a measure of pile strength (API, 1993) to compare it with allowable stress derived from the following well known formula

$$\sigma_{all} = \frac{S_y}{SF}, \tag{23}$$

where

σ_{all} allowable stress,

S_y the yield stress for the pile under study = 210 Mpa,

SF safety factor.

According to American Institute of Steel Construction (AISC) standards the general level of safety margin is 0.66 of the yield stress for the present case, which gives 1.67 as the safety factor. However in this tentative case a more conservative measure was taken; i.e. the allowable stress was selected as being 55% of the yield stress for the current research on pile driving operations which gives a safety factor of approximately 2.

The position of this maximum stress depends on the wave excitation, since the pile response is non-linear. The time history of von Mises stresses at a few elements of the pile, such as elements no. 8,9,14,15,16,17 collected in a subset called "*Max*", were obtained and examined to ensure that the critical element (the element with maximum response) is picked up. In nearly all of the analyses, element no.17 proved to be the critical. The analysis related to each sea-state was done in some 6000 separate increments which produced approximately 6000 von Mises stress values for both the top and bottom integration points of each of the

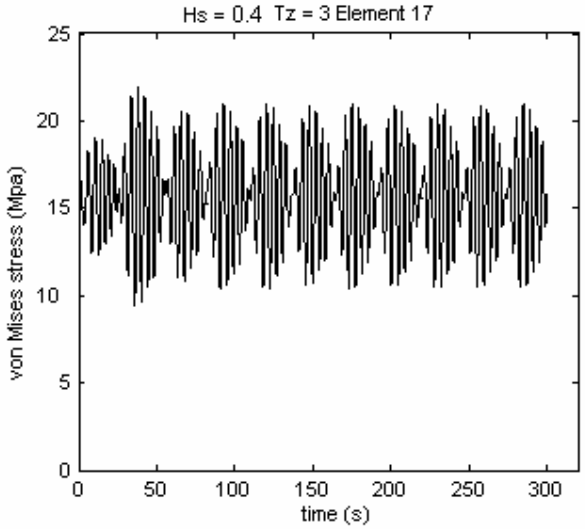


Fig. 5.8 Stress time history of an integration point of the critical element.

elements in subset "*Max*". Therefore for each element of the subset 2 series of 6000-von-Mises-stress values were produced by ABAQUS/AQUA. Figure 5.8 shows a typical stress time history at an integration point on the critical element resulted from applying the sea-sate characterized by H_s =1 m & T_z = 3 s.

As stated before more than 6000 von Mises stress values were obtained for each integration point. The mean and standard deviation of these data were computed and the most likely extreme amplitude of stress cycles after h =3 hours, using a long-term extreme prediction due to Barltrop and Adams (1991), was calculated from Eq. (24) for all elements in the previously defined subset "*Max*",

$$S_e = \bar{S} + \sigma\sqrt{2 \ln\frac{3600h}{T}}, \qquad (24)$$

where

S_e most likely extreme stress after h hours (MPa),

$\bar{S}$ the mean of the stress time history (MPa),

σ standard deviation of the stress time history of an element (MPa),

h time (hr); in the present study h =3,

T mean zero-crossing period for the stress time history related to an integration point of an element; in this study $T = 300/N$ where

N total number of zero up-crossings in the stress time history.

The analyses using Eq. (24) are shown in Table 5.5 for 75 sea-states. The most probable maximum von Mises stress predicted to occur in the pile after h = 3 hours is the value obtained from Eq. (24) for the critical element of the pile. In other words it is the maximum of all of the values obtained from Eq. (24) for all elements of the subset "*Max"*. For each sea-state the most probable maximum has been calculated and inserted in Table 5.5. Each result of the table shows the pile responses due to a particular sea-state as well as the self-weight and buoyancy effects calculated by ABAQUS/AQUA. In the present study a sea-state was considered safe for pile driving if the maximum predicted von Mises stress is less than 55 % of 210 MPa which is the yield stress of the pile material i.e. Steel M10. The sea-states with H_s's higher than 3.5 m were believed to be unsafe because any piling operation cannot proceed in a rough sea as indicated by H_s = 4 m or more. More details of the analysis for three sea-states using Eq.(24) are shown in Table 5.5.

Table 5.5 The results of the pile analysis.

Sea-State		Maximum von Mises predicted for 3 hours (MPa)	1= Safe 0= Uns.	Sea-state		Maximum von Mises predicted for 3 hours (MPa)	1= S 0= U	Sea-state		Maximum von Mises predicted for 3 hours (MPa)	1= S 0= U
Hs	*Tz*			*Hs*	*Tz*			*Hs*	**Tz**		
.4	2	19.62	1	1.2	4.3	79.50	1	2	8	59.41	1
.4	3	26.70	1	1.2	6	53.19	1	2	10	39.31	1
.4	4	29.96	1	1.2	8	48.52	1	2	12	42.24	1
.4	6	23.55	1	1.2	10	34.44	1	2.2	2	24.47	1
.4	7	26.82	1	1.2	12	29.71	1	2.2	4	69.64	1
.4	8	22.36	1	1.4	2	34.78	1	2.2	6	95.71	1
.4	10	21.12	1	1.4	4	52.66	1	2.2	8	51.46	1
.4	12	20.29	1	1.4	6	43.29	1	2.2	10	41.50	1
.6	2	28.33	1	1.4	8	35.62	1	2.2	12	42.21	1
.6	3.5	55.05	1	1.4	10	31.76	1	2.4	2	24.09	1
.6	4	43.94	1	1.4	12	34.00	1	2.4	3	77.03	1
.6	6	29.71	1	1.6	2	28.70	1	2.4	4	75.83	1
.6	8	26.03	1	1.6	4	120.86	0	2.4	6	78.36	1
.6	10	22.89	1	1.6	6	67.55	1	2.4	8	29.71	1
.6	12	22.60	1	1.6	8	52.70	1	2.4	10	54.03	1
.8	4	55.26	1	1.6	10	35.83	1	2.4	12	39.54	1
.8	10	30.12	1	1.6	12	30.68	1	2.8	3	66.87	1
1.0	2	20.48	1	1.8	6	56.20	1	2.8	7	112.25	1
1.0	4	45.57	1	1.8	8	41.36	1	2.8	8	79.36	1
1.0	6	39.25	1	1.8	10	51.77	1	3.2	4	227.90	0
1.0	8	36.31	1	1.8	12	32.44	1	3.2	8	69.1	1
1.0	10	31.65	1	2	2	26.06	1	3.2	12	57.5	1
1.0	12	26.74	1	2	4	77.63	1	3.5	4	129.97	0
1.2	2	150.58	0	2	6	50.96	1	3.5	8	116.93	0
1.2	4	59.68	1	2	6.5	64.49	1	3.5	12	53.8	1

Table 5.6 The details of analysis using Eq. (24) for 3 sea-states.

H_s	T_z	#elements in Set *"Max"*	Max Mises predicted for 3 hours	No of cycles available in the stress time series	Coeff. = $\sqrt{2\ln\frac{3600h}{T}}$	1=safe 0=unsafe	Critical element
0.4	3.0	6	26.7	75	3.98	1	17
1.8	10	6	51.8	69	3.95	1	17
3.2	4.0	2	227.9	134	4.12	0	17

It is worth mentioning that two more formulae, as well as Eq. (24), were also used to predict the most likely maximum von Mises stress occurring in the pile after 3 hours. Finally Eq. (24) based on Barltrop and Adams (1991) was chosen for the calculation of the most likely maximum. A comparison of the results of the three formulae is shown in Sec. 5.8.

5.6.2.1 Some other common formulas used instead of Eq. (24)

As well as using Eq. (24), two other formulae proposed by some experts were also used to calculate the most likely extreme stress, but the attempt to prove them or to find a source for their proof failed and therefore Eq. (24) was used for the purpose. The two formulae are given in Eqs. (25) and (26). Table 5.10 shows the results from the 3 different equations.

$$S_e = \overline{S} + \sigma \sqrt{2 \ln \frac{3600\ h}{T}}, \tag{24}$$

$$S_e = RMS_{\max} \sqrt{2 \ln \frac{3600\ h}{T}}, \tag{25}$$

$$S_e = \overline{S} + RMS_{\max} \sqrt{2 \ln \frac{3600\,h}{T}}, \tag{26}$$

where

S_e most likely extreme stress after h hours (MPa),

$\overline{S}$ the mean of the stress time history related to a pile element (MPa),

RMS_{max} root mean square of the von Mises of the stress time history produced by ABAQUS for the critical element during pile driving,

T 300/N,

N no. of cycles observed in the plot of von Mises stress versus time related to the critical element.

It is might be interesting to note that

$$\text{Coeff.} = \sqrt{2 \ln \frac{3600\ h}{T}} \tag{27}$$

has a value around 4 in all of the sea-states examined.

Table 5.10 Comparison of results from 3 Eqs. (24), (25) and (26).

Hs	*Tz*	Coeff.	Max von Mises stress predicted for 3 hrs (Se) From Eq. 25	1=Safe 0=Unsafe	Max von Mises stress predicted for 3 hrs (Se) From Eq. 24	1=S 0=U	Max von Mises stress predicted for 3 hrs (Se) From Eq. 26	1=S 0=U
0.4	3	3.98	63.254	1	26.698	1	78.923	1
0.8	4	3.96	71.955	1	55.264	1	87.021	1
1.2	2	4.08	181.539	0	150.580	0	215.709	0
1.2	4	4.01	73.571	1	59.681	1	88.062	1
1.2	4.3	3.96	92.602	1	79.497	1	109.927	1
1.6	2	3.99	63.799	1	28.703	1	79.458	1
1.6	4	4.05	143.963	0	120.857	0	170.911	0
1.8	10	3.95	70.548	1	51.773	1	85.8160	1
2.8	7	3.99	130.819	0	112.249	1	155.049	0
3.2	4	4.1	294.458	0	227.900	0	355.497	0

More details of Eq. (25)

This equation has been used in an analysis done on piles by a researcher but no source for its proof was found. The root mean square (RMS) of all of the 6000-value stress time histories related to set *Max* were calculated and the maximum RMS was determined and substituted in Eq. (25) to predict the most likely extreme von Mises stress occurring in the pile after h = 3 hours. If S_e is less than 55 % the yield stress of the pile, the sea-state will be considered safe. The sea-states having more than H_s = 3.5 m are believed to be unsafe. Table 5.11 shows the results for some sea-states.

Table 5.11 The details of analysis for 3 sea-states using Eq. (25)

H_s	T_z	no. of elements processed ×2= no. of integ. Points	Element & the integ. point on which RMS_{max} occurred	S_e= max von Mises predicted after 3 hr (MPa)	N= no. of cycles	Coeff..= $\sqrt{2\ln\frac{3600\,h}{T}}$	1=safe 0=unsafe
0.4	3	12	17 , 1	63.25	75	3.98	1
1.8	10	12	17 , 1	70.55	69	3.95	1
3.2	4	4	17 , 2	294	123	4.1	0

5.7 Application of Sea-State Simulation to Pile Driving Operations

Given the necessary inputs to a MATLAB code, the significant wave height and mean zero-up-crossing period of a desired future 3-hourly sea-state are generated as random variates from their corresponding hepta-parameter spline conditional distribution whose seven parameters are estimated from the outputs of the trained artificial neural networks described in Chapter 4. The probability density function of the distribution was also described earlier in Chapter 4. The information that belongs to the future needs not to be given as input; the code will generate them. To determine the level that a future sea-state is safe to drive a pile, the 3–hourly sea-state for the future is first predicted N times in this way and then it is determined from Table 5.5 which of the N times is safe or otherwise; if the sea-state proves to be safe **n** times and unsafe (**N-n**) times, then we could conclude that the sea-state is safe with 100(n / N) percent of certainty.

5.7.1 1st Illustration

As the first illustration, suppose it is desired to predict the sea-state of 6 PM 2 Jan for the region in NE Pacific, given 1 Jan sea-states. Table 4.4 of Chapter 4 shows the observed 3-hourly H_s 's for 1 Jan and two necessary T_z 's, all to be given as input to a MALAB code which generates the H_s and T_z of the sea-state from the corresponding conditional distributions.

To predict the sea-state for the time, after inputting the data of Table 4.4, the H_s and T_z of the sea-state were generated 20 times by a MATLAB code which calculates the parameters of the conditional distributions from the outputs of the trained networks and generates the H_s and T_z from their corresponding pdf. The results are shown in Table 5.6. The code has generated the H_s's and the T_z's (columns 2 and 3) in 20 runs. Fourth column, which indicates whether the generated sea-states are safe or otherwise, was extracted visually from the information in Table 5.5. According to Table 5.7 it is 5/20=25% probable that the sea-state of 6 PM 2 Jan be safe.

Table 5.7 Twenty times prediction for 6 PM 2 Jan sea-state.

Run #	$H_{s(m)}$	T_z (s)	1=Safe 0=Unsafe
1	4.3	6.7	0
2	4	8.2	0
3	5.6	7.6	0
4	5.3	9.1	0
5	7.3	10.8	0
6	7.4	9.8	0
7	9.8	10.	0
8	1.8	6.6	1
9	7.	9.7	0
10	2	6.3	1
11	5	7.9	0
12	1.8	7.5	1
13	5.9	6.8	0
14	3.4	7.9	0
15	3.2	7	1
16	2.2	8.1	0
17	5.9	8.9	0
18	5.9	8.9	0
19	6.6	11.1	0
20	1.8	6.8	1

5.7.2 2nd Illustration

As another illustration, given the data in Table 4.4 of Chapter 4, the sea-state of 9 AM 2 Jan was simulated in a similar way. The results are shown in Table 5.8. Therefore, based on the 10-time simulation shown in the table, the sea-state for the desired time is 30% safe. For practical purposes, the number of runs should be increased to get more accurate results.

Table 5.8 Ten times prediction of sea-state for 9 AM 2 Jan.

Run #	H_s (m)	T_z (s)	1=Safe 0=Unsafe
1	2.5	7.6	1
2	3.1	7.9	1
3	3.7	7.4	0
4	4.3	7.9	0
5	4.1	7	0
6	3.8	8.3	0
7	5.4	7.3	0
8	2.8	6.5	1
9	3.9	8.3	0
10	5.5	7.4	0

References (in alphabetical order)

ABAQUS 6.4 User's manual, 2004.
Hibbit, Karlsson & Sorensen Inc.

AISC
American Institute of Steel Construction (AISC) codes, and recommendations.
American Institute of Steel Construction Inc.
Chicago, Illinois, USA.

API RP 2A-LRFD, 1993.
Recommended Practice for Planning, Designing and Constructing Fixed Offshore Platforms–Load and Resistance Factor Design.
American Petroleum Institute.
1220 L Street, Northwest, Washington DC 20005 USA, Washington, D. C.

Ben, C. and Gerwick, Jr., 2000.
Construction of Marine and offshore Structures.
2nd edition, CRC Press ISBN: 0-8493-7485-5.

Barltrop, N. D. P., 1998.
Floating Structures: A Guide for design and analysis.
Oilfield Publication Ltd, ISBN: 1 870553 357.

Barltrop, N. D. P., and Adams, A. J., 1991.
Dynamics of Fixed Marine Structures.
Butterworth–Heinemann, Oxford.

Bretschenider, C. L., 1959.
Wave variability and wave spectra for wind generated gravity waves.
Technical memorandum 118, U.S. army beach erosion board, Washington D.C.

Cartwright, D. E., and Longuet-Higgins, M. S., 1956.
The Statistical Distribution of the Maxima of a Random Function.
Proc. Royal Soc. of London Series A Vol 237.

Djahansouzi, B., and Yasseri, S. F., 1994.
Study of offshore pile stick-up.
Proc. 9th UK ABAQUS User Group Conference, Oxford 1994.

Goda, Y., 1990.
"Random Waves and Spectra",
Chapter 4 of Vol 1 "*Handbook of Coastal and Ocean Engineering*".
Edited by Herbich, J., Gulf Publishing Company.

Goda, Y., 2000.
Random Seas and Design of Maritime Structures.
2nd edition, World Scientific, Singapore.

Hasselman, K., Barnet, T. P., Bouws, E., Carlson, H., Cartwright, D. E., Enke, K., Ewing, J. A., Gienapp, H., Hasselman, P., Kruseman, P., Meerburg, A., Muller, P., Olbers, D., Richter, K., Sell, W., and Walden, H., 1973.
Measurements of Wind Wave Growth and Swell Decay during the Joint North Sea Wave Project (JONSWAP).
Deutsches Hydrographisches Institut, Hamburg, Germany Series A. no. 12.

IAHR, 1986.
List of Sea State Parameters.
Supplement to PIANC Bulletin No. 52.
Int. Association for Hydraulic research.
Permanent International Association of Navigation Congress, Brussel.

Lamb, H., 1945.
Hydrodynamics.
Dover, 1945.

Lemehaute, B., 1976.
An Introduction to Hydrodynamics and Water Waves.
Springer-Verlag: New York.

Lighthill, J., 1978.
Waves in fluids.
Cambridge University

Ochi, M. K., 1998.
Ocean Waves-The Stochastic Approach.
Cambridge University Press, Cambridge.

Phillips, O. M., 1958.
The equilibrium range in the spectrum of wind-generated waves.
Journal of Fluid Mechanics, 4, pp 426-434.

Sorenson, R. M., 1997, Translated into Farsi by Bargi, K, 2001.
Basic Coastal Engineering.
Tehran University Press, Iran.

Pierson, W. J., and Moskowitz, L., 1964.
A proposed spectral form for fully developed wind seas based on similarity theory of S. A. Kitaigorodskii.
J. Geophys. Res., 69(24), pp 5181-5190.

Wiegel, R. L., 1949.
An analysis of data from wave records on the Pacific coast of the US.
Trans. American Geophys., Union, vol. 30, pp 700-704.

WMO, 1998.
Guide to wave Analysis and forecasting.
2nd edition, World Meteorological Org. Publication no. 702 Geneva.
www.wmo.int/web/aom/marprog/ Wordpdfs/WMO%20No%20702/WMO702.pdf.

Chapter 6

Extreme wave analysis using significant wave heights simulated by ANNs and using a model of H_s

Nomenclature

A	the FT-I, GEV, Weibull, and GPD location parameter (Appendix E)
ANNs	artificial neural networks
B	the FT-I, GEV, Weibull, and GPD scale parameter (Appendix E)
b	1)a positive number; 2) the parameter of the Rayleigh distribution
C	the GEV, Weibull, and GPD shape parameter (Appendix E)
c	a high value H_s
c_0, c_1	constants in Eq. (22)
c_2, c_5	constants in Eq. (22)
CDF	cumulative distribution function
C.F.	characteristic function
d=day	a number in the interval 0-365.25 with step 0.125; d=0 represents 0 AM 1 Jan.
EDF	empirical distribution function
EVA	extreme value analysis
F_i	the plotting position for the i^{th} member of the ordered data sample
F^{-1}	the inverse function of the CDF
$F_X * F_Y$	the convolution of the distribution functions of random variables X and Y
F_{X+Y}	the cumulative distribution function of X+Y
FT-1	Fisher-Tippet Type I distribution
FT-II	Fisher-Tippet Type II distribution
FT-III	Fisher-Tippet Type III distribution
GEV	generalised extreme value distribution
GPD	generalised Pareto distribution
$H_1, H_2, \ldots, H_n$	a random sample of wave heights
$\overline{H}$	the mean of the heights of all waves in a record.
H_{rms}	root mean square of a sample of wave heights
H_s	significant wave height
$H_{s\mathrm{Nyr}}$	N-year return value of H_s

i.i.d.	independent and identically distributed
M	number of the observed H_s's per year
m	sampling period, time interval between two successive H_s's (hr)
m_n	the n^{th} moment of the variance spectrum
MLEs	maximum likelihood estimate(s)
MSE	mean square error
N	the return period
n	number of the values above a threshold
POT	peaks over threshold
pdf	probability density function
p	a given probability
Q-Q	quantile-quantile
res	denotes the residuals in Eq. (22)
RMSE	root mean square error
u	a threshold
X	1)a continuous random variable having a range in $(0,\infty)$; 2) annual cycle in Eq.(29)
$x_{(1)},\ldots, x_{(n)}$	the values of *res* exceeding threshold u
$\overline{X}_u$	$\overline{X}_u = \sum_{i=1}^{n} x_{(i)} - u$
Y	1)a continuous random variable having a range in $(0,\infty)$; 2) *res* in Eq.(29)
Z	(annual cycle| *res*>u + *res* | *res* >u)
Z_{extr}	an extreme value of random variable Z satisfying Eq. (38)
ε	error
η	sea surface elevation
η_{rms}	root mean square of η
λ	the mean rate that a threshold is exceeded
σ_η^2	the variance of η
$\Phi_{X,Y}(t, u)$	the joint characteristic function of random variables X and Y
$\Phi_X(t)$	the marginal characteristic function of X
$\Phi_Y(t)$	the marginal characteristic function of Y.

6.1 Introduction

Maritime structures such as breakwaters for ports, sea walls for shore protection, and offshore platforms for oil exploitation are subject to waves, winds, tides, sea-quakes etc. in the natural environment. Waves, due to having the greatest influence on these structures, are the most important complex changeable phenomenon to be considered. For various random phenomena encountered in the field of naval and ocean engineering, it is often necessary to predict the largest (extreme) value of the environmental phenomena including ocean waves. The calculation of extreme values of various environmental parameters, which are important to the design of structures, is one of the most important problems in applied climatology.

6.2 Extreme Value Theory- History

Extreme value analysis (EVA), which deals with the methods of calculating the return values as will be defined in Sec. 6.6, goes back to 1928 when EVA was first formalized by Fisher and Tippet (1928). Later Rice (1944 &1945) discussed the statistical study of extreme values of random processes and Cartwright and Longuet-Higgins (1956) and Huston and Sponkinski (1956) dealt with the subject in detail. Extreme value analysis was greatly elaborated by Gumbel (1958). Coles' (2001) book, an outstanding recent publication on extreme value theory, introduces the following references on this subject: Leadbetter et al. (1983), Galambos (1987) and Resnick (1987) who all deal rigorously with extreme value models. Galambos et al.(1994) is the compendiums of conference proceedings that contain papers addressing specific theoretical and methodological aspects of extreme value theory. Castillo (1988) dealing with extreme values in engineering and Reiss and Thomas (2001) presenting the theory and applications of extreme value analysis in some fields including in hydrology are two more recent works that offer alternative viewpoint from that of Coles (2001).

This chapter has proposed a method for calculating significant wave height return value from a model presented by Anderson et al. (2002). It also deals with the conventional ways of extreme wave height analysis and also applies the *maxima method* in such a way that there is no need for unrealistic assumptions, as is the case with conventional maxima method in extreme wave height analysis. These assumptions are the requirements of using a theorem presented by Fisher and Tippet (1928).

6.3 Wave Height Distribution

Since the studies by Lord Rayleigh in 1880 and the paper by Longuet –Hinggin (1952), who showed that Rayleigh distribution is applicable to ocean wave height, this one-parameter distribution has universally been employed to describe the distribution of individual wave height (Goda, 2000). It is a special case of the Weibull distribution with the density function (pdf) and cumulative function (CDF) given in Eqs. (1) & (2) respectively where $0 \le x < \infty$.

$$f(x) = \frac{2x}{b^2} e^{\frac{-x^2}{b^2}}, \quad (1), \text{ and} \qquad F(x) = 1 - e^{-\frac{x^2}{b^2}}. \quad (2)$$

Figure 6.1 illustrates typically the pdf for x = wave height when $b = 5$.

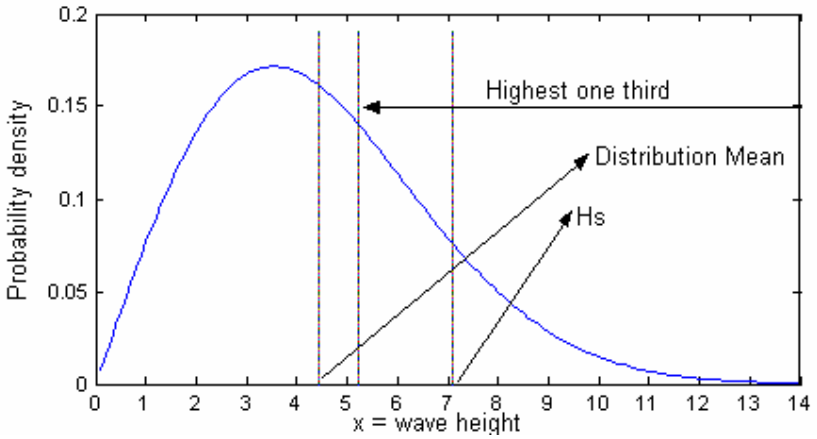


Fig 6.1 Rayleigh distribution with parameter $b = 5$.

Parameter estimate

The parameter of the Rayleigh distribution could be estimated from one of the following estimators:

$$\hat{b}^2 = \begin{cases} H_{\text{rms}}^2 \\ 8m_0 \quad (\text{Ochi}, 1997) \\ 8\sigma_\eta^2 \quad (\text{Guedes Soares}, 2003) \\ 8\eta_{\text{rms}}^2 \end{cases}, \quad (3)$$

where

$\hat{b}$ the distribution parameter estimate,

$H_{rms}^2 = \frac{\sum_{i=1}^{n} H_i^2}{n}$ root mean square of a sample of wave heights,

$H_1, H_2, ..., H_i, ..., H_n$ a random sample of wave heights,

m_0 zeroth moment of energy density spectrum,

η sea surface elevation,

σ_η^2 variance of the sea surface elevation,

η_{rms} root mean square of the sea surface elevation.

It could be easily verified that H_{rms} is the maximum likelihood estimate of b. The mean and variance of Rayleigh distribution is given in Eqs. (4) and (5) respectively. These formulas have been derived from the formulae for mean and variance of 2-parameter Weibull distribution (Montgomery & Runger, 1994) having the shape and scale parameters equal to 2 and b^2 respectively.

$$E(X) = \frac{b}{2}\sqrt{\pi}, \tag{4}$$

$$Var(X) = b^2 (1 - \frac{\pi}{4}). \tag{5}$$

Recall that the significant wave height (H_s) is a prominent representing parameter of a sea-state and its definitions and the related formulae were given in Chapter 5. The H_s of the sea-state from which the random sample is taken can be estimated from Eq. (6) (Sorenson, 1997, Barltrop & Adams, 1991):

$$H_s = 1.416\, H_{rms}, \tag{6}$$

or from (Goda, 2000)

$$H_s = 1.6\, \overline{H}, \tag{7}$$

$\overline{H}$ is the mean of the heights of all waves in a record.

As a common practice, sample size of at least n = 120 to150 is required for a reliable estimate of the root mean square value (Ochi, 1998). Assuming Rayleigh distribution for wave height and considering Eqs. (2), (3), and (6) would yield:

$$\Pr(H > H_s) = \exp(-1.416^2) \cong 0.135, \text{ or } H_s = 1.416\, b. \quad (8)$$

That is 13.5% of the wave heights recorded in a storm could be higher than a wave height equal to H_s. As proved in Chap. 5, the most probable highest wave in a record (of 1000 waves) is equal to

$$H_{\text{extreme most probable}} = 1.86\, H_s. \quad (9)$$

To illustrate some of the above equations suppose the parameter of the distribution plotted in Fig.6.1 has been estimated from a record of size 150 having H_{rms} = 5 m. Eq. (6) gives H_s =7.08 m. Eq. (8) or its following equivalent MATLAB command

weibinv(1-exp(-1.4162^2),1/b^2,2) or *weibinv((1-0.135),1/b^2,2)*

for b = 5 yields H_s = 7.075 m. The mean of the wave record has not been given, however calculating the mean of the distribution from Eq. (4) and substituting in Eq. (7) results in H_s = 7.090 m.

6.4 Methods of Fitting Data to a Distribution

Before dealing with H_s return value and its methods of calculation, some common ways of estimating the parameters of a distribution fitted to a sample of data is mentioned. In general it will be necessary to fit the data to a number of distributions and then to find the best-fitting one (Muir&El-Shaarawi,1986). There are many references on the fitting methods including Lettenmaier & Burges(1982), Muir&El-Shaarawi (1986), van Vledder (1993), Isaac-son and Mackenzie (1991), and Goda (2000).

In fitting a distribution to the data some methods are used for estimating the unknown parameters of the distribution. Maximum likelihood estimation (MLE) method, the method of moments, and least squares method are the most common. There are some others including graphical fitting method, mean slope, Pickands' estimator method. Carter & Challenor (1983) have illustrated using the first 3 methods for FT–I. Goda et al. (1993) show detailed application of the methods for the FT–I and 3-parameter Weibull. Goda et al. (1994) have discussed maximum likelihood method, the moment method, and the least squares with such goals as what fitting procedures should be used.

In fitting a probability density function having a set of parameter Θ by **maximum likelihood method**, which is a very popular method of parameter estimation, the joint

probability density function of random sample $x_1, x_2, ..., x_n$ denoted by $f_n(x_1, x_2, ...x_n \mid \Theta)$, known as likelihood function and factorized according to Eq. (10), is formed.

$$L = f_n(x_1, x_2, ...x_n \mid \Theta) = \prod_{i=1}^{n} f(x_i \mid \Theta). \quad (10)$$

The logarithm of the likelihood function, which is a function of the parameter set, Θ, is then maximized with respect to unknown Θ to estimate the unknown values of the parameters. Muir and El-Sharaawi (1986) introduce Johnson and Kotz (1970) for studying details of M.L.E method. Robustness of the estimates is discussed by Huber (1981). Maximum likelihood estimates are generally more robust (Huber, 1981) than other estimates i.e. are not so likely to produce erratic result when the observed distribution differs somewhat from the model distribution (Goda et al., 1994). The three parameters of the Weibull distribution can be estimated by MLE method only when its shape parameter is greater than 2 (Bury, 1975 & Goda et al., 1994). Carter & Challenor (1983) have discussed MLEs for FT-I distribution parameters in detail. The simultaneous likelihood equations for estimating the parameters of the FT-1, FT-II, FT-III, Weibull, Lognormal, and Poisson-Gumble distributions are given by Muir and El-Sharaawi (1986).

The **methods moments** equates the data sample moments to the corresponding moments derived from the desired distribution and solves for the unknown parameters. The number of equations is equal to the number of unknown parameters. This simple and old method has a small root mean square error (RMSE) compared to other methods except the maximum likelihood method. The method of moments is probably the most widely used parameter estimation method but its efficiency is less than that of the M.L.E method (Goda et al., 1994). The efficiency of an estimator is defined as the ratio of the minimum attainable variance to the variance of the parameter estimate. This method is theoretically inappropriate for fitting non-normal data (Muir & El-Sharaawi, 1986). The method is described for FT-I by Carter and Challenor (1983) in detail.

The **method of least squares**, widely used for 2-parameter distributions, requires a rearrangement of a given sample in the descending order i.e. $x_{(1)}<x_{(2)}<,...,<x_{(i)}<,...,<x_{(n)}$, where $x_{(1)}$ and $x_{(n)}$ are the minimum and the maximum of the sample respectively. The rearranged data are then given the probability of non-exeedance according to their order number, *i*, and the sample size *n*. The assignment of respective probabilities of non-exceedance to individual

values in the ordered data is made with the so-called *plotting position* formula. In fitting GPD, some researchers including Coles (2001) use the following formula for plotting position:

$$F_i = \frac{i}{n+1}, \tag{11}$$

where

F_i is the plotting position for the i[th] member of the ordered data sample i.e. $x_{(i)}$

i is the order of the member in the ordered sample,

n is the sample size.

However Eq. (11) is not appropriate for all distributions. The proposed plotting position formulae for five distributions, extracted from Goda (2000), are given by Eq. (12):

$$F_i = \frac{i-a}{n+b}, \tag{12}$$

where

F_i the plotting position of the i^{th} ordered statistic,

n number of data in the wave sample, and

a & b two constants given in Table 6.1.

Table 6.1 Constants for the plotting position formula given in Eq. (12)*

Distribution	a	b
FT-I	0.44	0.12
FT-II	$0.44 + \frac{0.52}{C}$	$0.12 - \frac{0.11}{C}$
Weibull	$0.20 + \frac{0.27}{\sqrt{C}}$	$0.20 + \frac{0.23}{\sqrt{C}}$
Normal	0.375	0.25
Lognormal	0.375	0.25

* C is the distribution shape parameter.

After calculating the plotting position, $\hat{x}_i$ is derived as a function of the parameters of the candidate distribution by equating $F_i = F(\hat{x}_i)$. Then the sum of the squared differences between x_i's and $\hat{x}_i$'s is calculated. The resulting sum, which is a function of the distribution parameters, is then minimized to estimate the candidate pdf parameters. The values of the parameters that minimize the sum are the estimates for the parameters obtained by least

squares method. Carter & Challenor (1983) give the details of applying the method to the FT-I distribution. The least squares method is widely used for fitting parameters of distributions such as FT-I. Goda (2000) have fixed shape parameters for FT-II and 3-parameter Weibull at one of the following values and then used least squares method for estimating other two parameters.

FT-I	2.5	3.33	5.	10
Weibull	0.75	1	1.4	2.

Carter & Challenor (1983) describe the details of applying least squares method to FT-I distribution.

Mean slope, a less known method of fitting data to a distribution, has been mentioned by Muir & El-Shaarawi(1986). Goda (2000) has described **graphical fitting** and Pickands (1975) has dealt with **Pickands' estimator** method.

6.5 Goodness of Fit Tests

Some distributions are fitted to the data. Is there any test or criterion to find the best one? A visual examination of the fitted distribution could be the first appropriate step. Wilk and Ganadersiken (1968) have developed some plotting methods including quantile-quantile (Q-Q) plots for comparing different distributions fitted to the data. In practice one uses the candidate distribution to calculate the wave heights corresponding to the observed wave height samples and then the computed and observed heights are plotted against each other. This should give a straight line passing the origin. Q-Q plots have the advantage that various distributions can be plotted with exactly the same scale and visually inspected and compared one against the other (Muir & El-Sharaawi 1986). Figures 6.3, 6.5, 6.6, 6.7, and 6.11 of this chapter shpw several Q-Q plots. A detailed practice of this kind of plot is illustrated by Muir & El-Sharaawi (1986).

There are some statistics including the correlation coefficient (R) that could be used as measure of goodness of fit. R is calculated from Eq. (31) of Chapter 4; the greater the R the better is the fit. For a broad distribution the absolute value of R tends to yield unfavorable

result. A remedy for above bias is to use the MIR (Minimum Ratio of residuals correlation coefficient) ratio to find the best fit. Goda 2000 has described this criterion.

There are several statistical tests for investigating goodness of fit including Pierson chi-square, Kolmogorov–Smirnov, Anderson-Darling, and Cramer von Mises tests. Lawless(1982) has given excellent explanation of some sophisticated tests for extreme wave data such as the types based on empirical distribution function (EDF) which is identical to plotting position mentioned earlier. Muir & El-Sharaawi (1986) give more details. Kolmogorov–Smirnov, Anderson-Darling, and Cramer von Mises perform best for large data sets (Harper, 1998). Kolmogorov–Smirnov test (kstest) is described in many resources including MATLAB help. van Vledder et al (1993) have used several goodness of fit tests including Kolmogorov–Smirnov, chi square, Anderson-Darling, correlation of coefficient and MIR ratio as well as visual inspection. Ferreira & Guedes-Soares (1999) present an application of Kolmogorov-Smirnov test. Mann and the Mann-Fetig test (Lawless, 1982) are specific to the Weibull and the FT-I. These tests and EDF tests require large number of data points to be sensitive enough. Some papers including Petrauskas & Aagaard (1971) and Goda (1988) have used Monte Carlo methods to determine goodness of fit (Harper, 1998). Harper (1998) states that non-parametric Boothrap method (Effron, 1979) can be thought as a variation of direct Monte Carlo.

6.6 Return Value- Definition and Calculation Methods

For various random phenomena encountered in the field of navel and ocean engineering, it is often necessary to predict the largest (extreme) value of the phenomenon. For instance predicting the largest motions and wave-induced forces is essential for the design of ships and maritime structures (Ochi, 1973). The first step in designing a maritime structure is the selection of design waves (Goda, 2000) which is based on the extreme wave. As mentioned before, H_s is a representative characteristic of ocean waves. Measuring devices such as buoys usually report the height of sea waves in terms of 3-hourly H_s's. The extreme values of significant wave heights are called H_s return values. The return value of H_s's has been widely used as an important factor in the design of maritime structures. "The extreme value is defined as the largest value of a random variable expected to occur in an hour, or the severest sea-state expected to be encountered in 50 years, for example" (Ochi 1998). The N-year return value, or return level, of a random variable is defined simply as " that value which will be exceeded once on average in N years" (Anderson, et al. 2001). In the simplest case, if the random variable is

the maximum value of wave height during one year denoted by H_{smax}, the probability that a value of H_{smax} is greater than the N-year return value denoted by $H_{s\text{Nyr}}$, is given by Eq. (13) (Anderson et al., 2001)

$$\Pr\left(H_{smax} > H_{s\,\text{Nyr}}\right) = \frac{1}{N}. \tag{13}$$

If M observations are available each year, $H_{s\text{Nyr}}$ is given by: (Anderson et.al.2001)

$$\Pr\left(H_s > \text{H}_{\text{Nyr}}\right) = \frac{1}{MN}, \tag{14}$$

If sampling is done every m hours per day then, based on Barltrop (1998)

$$\Pr\left(H_s > H_{s\text{Nyr}}\right) = \frac{m}{365.25 \times 24 \times N}, \tag{15}$$

where

m the sampling interval (hr),

N the return period (yr).

Depending on the method of selecting a sample (set of data) three methods have conventionally been used in the derivation of extreme values (Goda, 2000). Muir & El-Sharawi (1986) have conducted a comprehensive review of the methods of the derivation of H_s extreme values. Goda (2000) also deals with the methods. A brief description of them follows.

6.6.1 Method I: Total Sample Method (fitting all the data)

In total sample method the whole data of wave heights obtained during a number of years from an instrument such as a buoy are analysed in a form of CDF. In other words this method is the extrapolation of distribution functions fitted to the cumulative frequency distributions of the data (Barltrop, 1998) i.e. the extrapolation of some CDF fitted to all the data gathered in a single sample. Once a best fitting distribution such as the Fisher-Tippet Type-1 (FT-1), 2-parameter or 3-parameter Weibull is found, the return value is estimated by extrapolating the CDF to the level of probability which corresponds to a given period of years being considered in the design process. This method does not require the data be collected at regular intervals at one location as do methods II and III. Therefore it can be adapted to satellite data e.g. altimeter data, if it could be assumed that the data measured over sizeable

area around the location of interest are representative of the climate at that location. This method has also been called *initial distribution method* and *CDF method*. Once the best-fitting CDF has been chosen for the whole sample, $H_{s\mathrm{Nyr}}$ will be given by the following equation

$$H_{s\mathrm{Nyr}} = F^{-1}(1 - \frac{m}{365.25 \times 24 \times N}), \tag{16}$$

where

F^{-1} is the inverse function of the CDF fitted.

Total sample method has been covered by a number of researchers such as Anderson et al (2001), Guedes-Soares & Henroques (1996), van Vledder et al. (1993), Ochi (1992), Ochi and Whalen (1980), and Nolte (1973). Anderson et al. (2001) used FT-1 distribution to fit the data; Goda et al. (1993) and van Vledder et al. (1993) have tried fitting some distributions including FT-I and 3-parameter Weibull to a set of data and have compared them. Some other distributions used in this method are long-normal (Jasper, 1956), generalised gamma distribution (Ochi, 1992), and the beta distribution (Ferreira & Guedes-Soares, 1999). Some researches (e.g. Harver,1985) have used a combination of the longnormal and the Weibull distributions with this method.

6.6.1.1 Illustrations for Total Sample Method

In this section three illustrations are given for this method.

6.6.1.1.1 1st Illustration

Suppose the return value is to be calculated for the 3-hourly H_s's reported US NODC BUOY 46005 in the North East Pacific during 1996-1999. FT-I, having the CDF given in Appendix E, was selected as the candidate distribution and the maximum likelihood method as the fitting method. A MALAB code estimated the parameters of FT-I using MLE method. This resulted in location parameter A = 2.2238 and scale parameter B = 1.075. If the observed H_s's follow a FT-I distribution with these parameter then the $(H_s)_i$'s and $-ln(-lnF_i)$'s will satisfy the line $(H_s)_i = A + B \times [-\ln(-\ln F_i)]$,

where

$(H_s)_i$ is the i^{th} H_s in the ordered sample,

F_i is the H_s corresponding plotting position calculated from Eq. (11).

Therefore if the ordered H_s's are plotted against their corresponding probabilities on FT-1 scale i.e.-ln(-lnF_i)'s a line should fit the points; the better the fit, the more acceptable the FT-I fit. Figure 6.2 shows the fit.

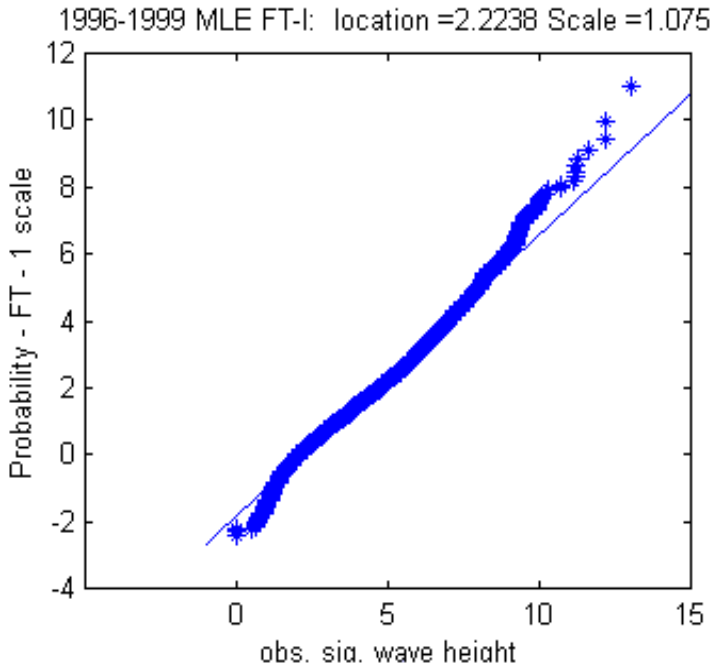


Fig. 6.2 Cumulative distribution for 3-hourly H_s 's measured by Buoy 46005 during 1996-99.

The result of fitting FT-I is A = 2.2238 and B = 1.075, therefore the *CDF* is given by

F(x) = $e^{-e^{-\frac{x-2.22388}{1.705}}}$, and according to Eq. (15) :

$$\Pr(H_s > H_{s100\text{yr}}) = \frac{3}{365.25\times 24\times 100} = 1 - e^{-e^{-\frac{H_{s100\,yr}-2.22388}{1.705}}} . \qquad (17)$$

Applying Eq. (16) on the fitted FT-I for the return period N =100 results in

$$H_{s\ 100\text{yr}} = 2.2238 - 1.705\times \ln[-\ln(1-\frac{3}{365.25\times 24\times 100})] = 15.753. \qquad (18)$$

The 100-year return value is therefore 15.76 m which is somehow greater than 15.74 m reported by Anderson et al. (2001). The difference is due to the difference of the replicated estimate of A and B (2.2238, 1.075), and the estimates calculated by the report for A & B (2.222, 1.074) which in turn could be due to the difference in number of our data extracted from the site of reference for analysis (i.e. 30718) and that of the report i.e. 30350.

6.6.1.1.2 2nd Illustration

As the second illustration the return value will be estimated by total sample method from the H_s's simulated for 1996-99 using the conditional pdf's whose seven parameters were estimated using the outputs of the trained neural networks described in Chapter 4. 3000 simulation runs were performed for the period 1996-1999**;** a 3-parameter Weibull fitted very well to each of the simulated data set. Figure 6.3 shows a quantile-quantile (Q-Q) plot for one of these 3000 simulations. 3000 return values were obtained from the 3000 runs using Eq. (16) with a maximum and a minimum equal to 17.35 m, 14.73 m respectively and a mean of 16.12 m. A return value computed by Anderson et al.(2000) for this period is 15.74 m.

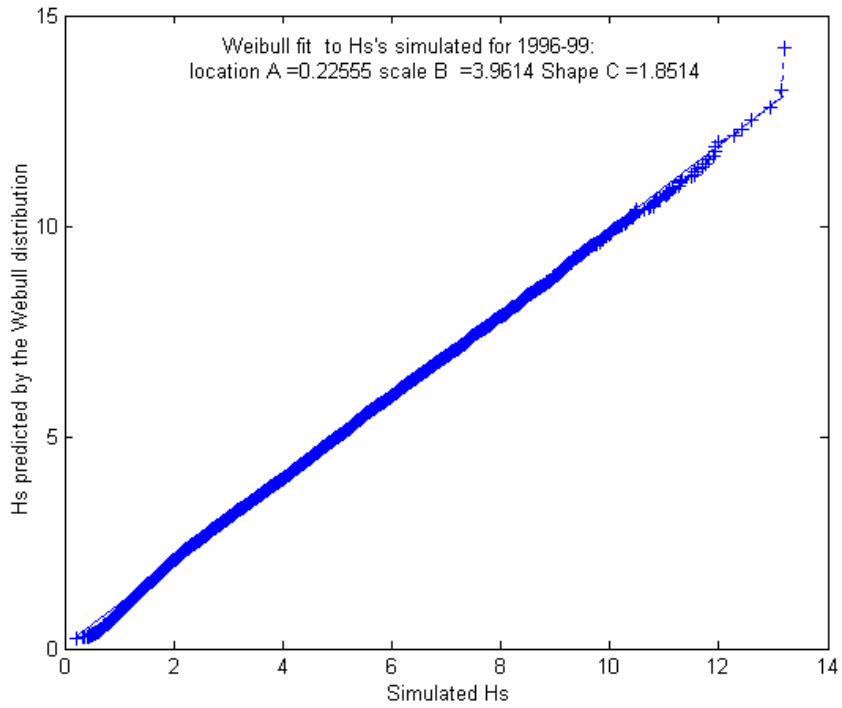


Fig. 6.3 A Weibull fitted to a typical simulation for the 3-houly H_s 's of 1996-99, Q-Q plot.

6.6.1.1.3 3rd Illustration *(using the empirical distribution of the simulated* $\mathbf{H_s}$ *'s)*

As the third illustration, a set of 3-houly H_s's for 1000 years beginning 1 Jan 78 was simulated from the hepta-parameter spline distributions whose parameters were estimated using the trained ANNs described in Chapter 4. This set was called Thousand-Year-Simulated-Vector. The empirical probability distribution function of the simulated data was

prepared (Fig.6.4). The empirical CDF or the empirical distribution function (EDF) of the H_s 's was calculated using MATLAB command :

[F, X] = ecdf (Thousand-Year-Simulated-Vector).

Applying Eq. (16) on the EDF yields H_{s100yr} = 15.29 m by the following MATLAB command

H_{s100yr} =*X (find (abs (F-3/(365.25*8*100)< ε)))*,

where

ε is the error i.e a small number such as 0.001 found by trial and error.

To compare this result with that of the conventional total sample method, a 3-parameter Weibull distribution with parameters: A = 0.14, B = 4.04, and C = 1.92 was found to fit well to all 3-hourly simulated H_s's for the 1000-year period. Figure 6.5 shows the quantile-quantile (Q-Q) plot for the Weibull distribution. This distribution resulted in H_{s100yr} = 14.77 m. A Rayleigh distribution with parameter b = 4.203 was also fitted to the 1000-year simulated data. The Rayleigh pdf and the above Weibull fit are shown in Fig. 6.4.

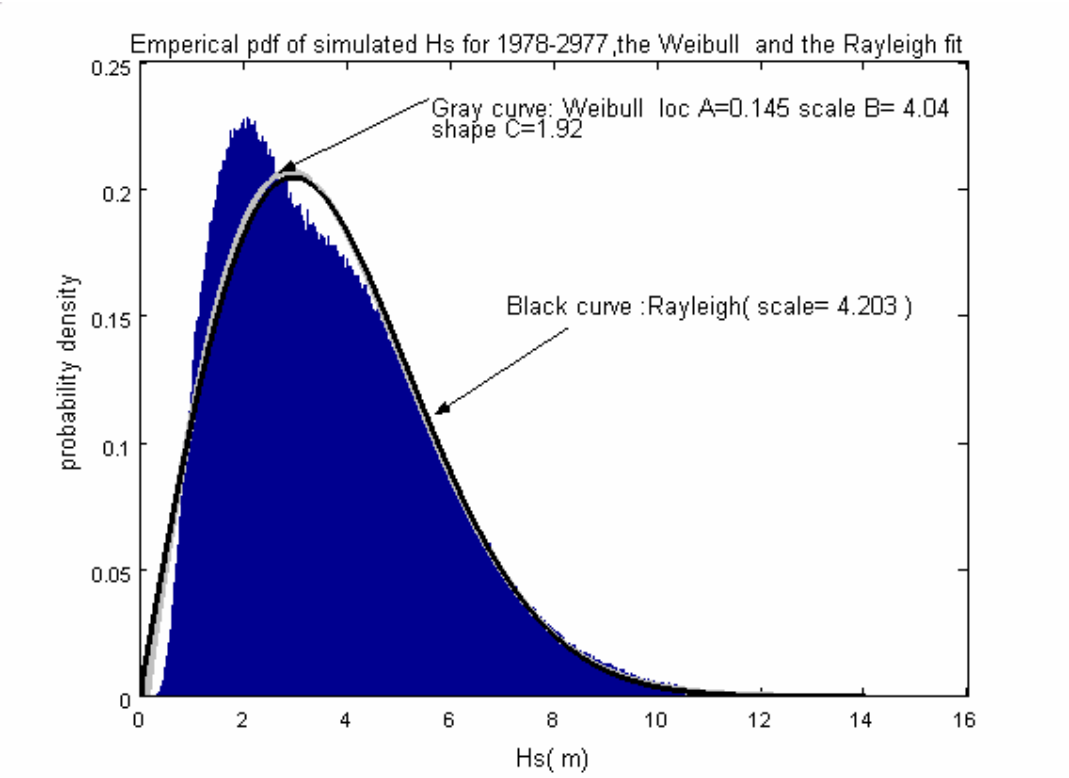


Fig. 6.4 The empirical pdf of 1000-year simulated H_s's and two fitted distributions.

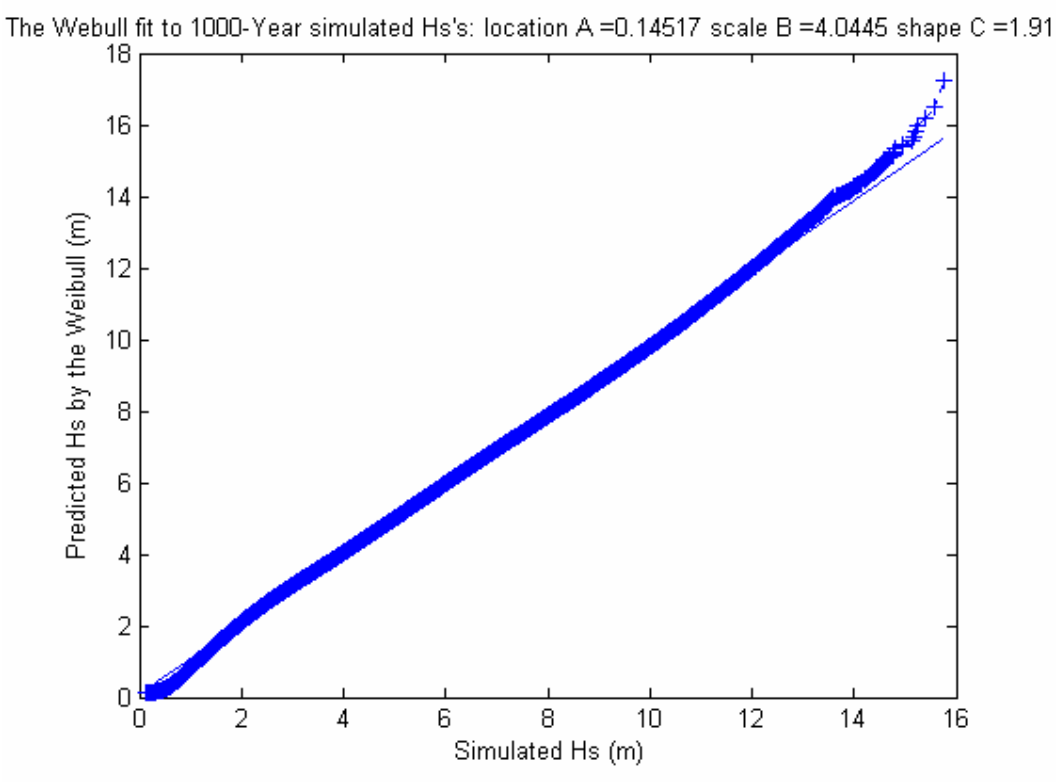


Fig. 6.5 A Weibull fit to 1000-year simulated H_s's beginning 1 Jan 78, Q-Q plot.

6.6.1.2 Lack of scientific justification of Total Sample Method

The successive 3-hourly data of the H_s time-series used in the total sample method are mutually correlated and do not conform to the hypothesis test of independent observations of a random sample; and also do not come from the same parent population. Therefore this method lacks scientific justification for dealing with non-iid H_s data in extreme wave analysis and is not recommended by a number of researchers. The problems with this method are discussed in more details by Goda (2000), and Ferreira & Guedes-Soares (1998).

6.6.2 Method II: Maxima Method

This method uses extremal distributions such as the FT-1 or the generalised extreme value (GEV) distributions fitted to the maxima of the H_s's, usually the annual maxima. N-year return value in this method is extrapolated from the CDF fitted to the annual maxima using Eq. (13) or equivalently from Eq. (19) (Anderson et al., 2001),

$$H_{s\,\mathrm{Nyr}} = F^{-1}(1-\frac{1}{N}), \qquad (19)$$

where

F^{-1} the inverse of the CDF fitted to the annual maxima

N the return period.

To perform the extreme analysis using the maxima method, the simulated data for 39996 years beginning from 1 Jan 1978 were used. In this section the conventional maxima method was done as well as a proposed method which uses the EDF of the maxima of nearly 400-century simulated H_s's.

Compounding monthly maxima method is a variation of the maxima method suggested by Carter and Challenor (1981) which analyses the monthly maxima separately and then deals with the annual maxima and calculating the return value. Compounding monthly maxima method has been illustrated in Sec. 6.6.2.2.

6.6.2.1 Conventional maxima method

Of central importance to the computation of extreme values is the work of Fisher and Tippet (1928). According to the Fisher-Tippet theorem, regardless of the parent distribution the maxima of data samples, as sample size increases, is asymptotically distributed as one of the distributions belonged to the so-called GEV distributions. The theorem requires that the data from which the maxima are drawn have some pre-conditions i.e. to be mutually independent and identically distributed. However Leadbetter et al. (1983) have shown that the theorem held even if there is some degree of independence. The family of GEV distributions includes the following three subclasses:

FT–I sometimes also called Gumbel, Double exponential, Extreme value1, and EV1,

FT–II sometimes also called Frechet, Extreme value 2, and EV2,

FT–III sometimes also called Extreme value 3, and EV3.

Jerkinson (1955) showed that the above three distributions might be expressed by GEV distribution whose pdf is given in Appendix E. EV3 is sometimes mistakenly called the Weibull distribution, although the Weibull distribution is in fact a different related distribution whose pdf and CDF are showm in Appendix E. Galambos (1978) showed that other asymptotic extreme distributions can exist if the earlier pre-conditions are not met but they will not necessarily be those of Fisher-Tippet (Harper, 1996). The GEV distributions are applicable only to regularly sampled data such as the annual or monthly maxima (Muir & El-Shaarawi, 1986) which are often not practically available in sufficient number (Harper, 1996). Muir &

El-Shaarawi (1986) state that "all of the well-known distributions such as the Rayleigh, the Wiebull, the normal and the long-normal distributions are in the domain of attraction of FT–I". They also mention some other distributions such as gamma and the log-pearson type III that have been used in the maxima method. This method has been covered in a number of references including Thom (1971), Carter & Challenor (1981), Muir & El-Shaarawi (1986), Goda et al. (1993), Goda (2000), and Anderson et al. (2001). Muir & El-Sharaawi (1986) in their excellent work deal the FT-I, FT-II, and FT-III, Weibull, Lognormal and Poissn-Gumbel distributions. The article also has applied some of the distributions to a set of data in order to compare them. Goda (2000) has discussed the FT-I, FT-II, FT-III, 3-parameter Weibull and lognormal distributions in detail. Goda et al. (2000) compare FT-I, four variations of FT-II and four variations of 3-parameter Weibull to find the best distribution for calculating return values in Japan Sea . Carter et al. (1986) present a catalogue of distributions (Barltrop, 1998).

6.6.2.1.1 Illustration for the conventional annual maxima method

Assuming that the nearly 400–century simulated data are i.i.d., a GEV distribution shown in Fig. 6.6 and Fig. 6.10 was fitted to the 39996 annual maxima of the simulated data using a MATLAB code of a downloaded wave package. The GEV parameters are A = 11.9865, B = 0.72564, and C = 0.090884.

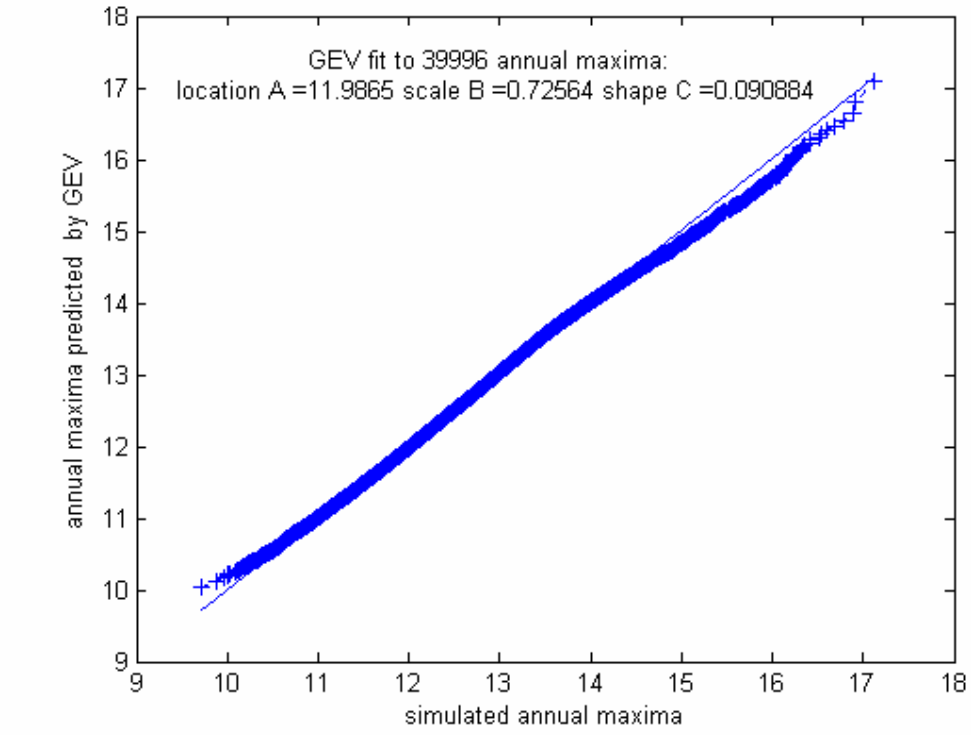


Fig. 6.6 A GEV fit to 39996 simulated annual maxima, Q-Q plot.

The code, having the option to use the method of moments and MLE method, was asked to use MLE method. The return value calculated from Eq.(19) by H_{s100yr} = *gevinv ((1-(1/100), 11.9880, 0.7187, 0.0904),* which uses the downloaded MATLAB code (gevinv.m), is H_{s100yr} = 14.67. Also a Weibull distribution with parametrs A = 9.7253, B = 2.9225, and C = 3.2692 was fitted to the 39996 simulated annual maxima (Fig. 6.7) using MLE method. Based on Eq. (19), the following MATLAB command

H_{s100yr} = *9.72+2.92*((-log(1/00)^(1/3.2692))* was used to calculate the return value using the inverse function of the 3-parameter Weibull distribution function. The result was H_{s100yr} = 14.39 m.

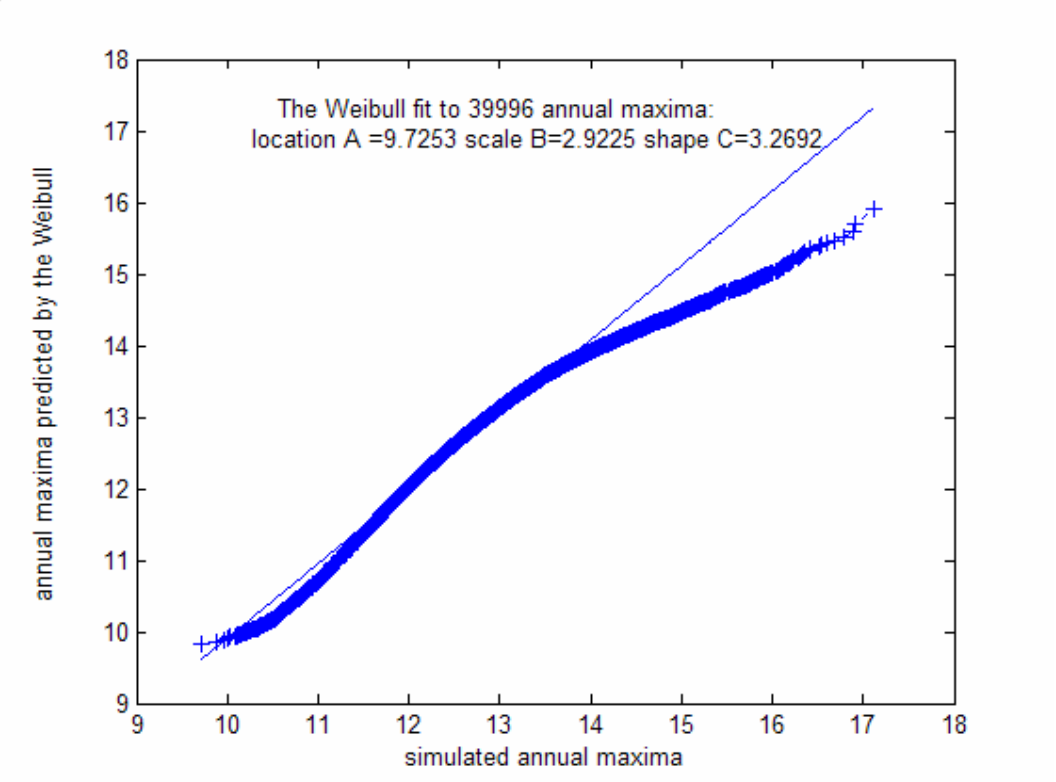


Fig. 6.7 A Weibull fit to 39996 simulated annual maxima, Q-Q plot.

6.6.2.2 Compounding monthly maxima method

Carter and Challenor in 1981 derived the cumulative distribution of annual maxima as the product of 12 monthly cumulative distribution functions (Goda, 2000) given by Eq. (20):

$$F_{\text{annual maxima}}(x) = \prod_{i=1}^{12} F_{\text{the maxima of month }(i)}(x), \qquad (20)$$

where

$F_{\text{annual maxima}}$ the CDF of the annual maxima,

$F_{\text{the maxima of month }(i)}$ the CDF of the i^{th} month maxima.

For simplicity one year was divided into 12 equal parts named M_1,..., M_{12}. The 12 corresponding EDF's were computed from the 12 sets M_i, i = 1,…,12 each consisting 39996 monthly maxima. To compute each of the 12 EDF's the following MATLAB commands were executed for ech i = 1 through 12:

[fMi , xMi] = ecdf(Mi);

H_s = 0:0.0003:15;

xMi(1) = xMi(1)-0.000000000001;

yMi = interp1(xMi,fMi, H_s);

yMi(H_s<min(xMi)) = 0;

yMi(H_s>max(xMi)) = 1

The EDF of the annual maxima, F, was calculated from multiplying the product of the 12 EDF's of the monthly maxima using Eq. (20) as such:

F = yM1.* yM2.* yM3.* yM4.* yM5.* yM6.* yM7.* yM8.* yM9.* yM10.*yM11.*yM12.

where

yMi, i =1,12 are sets of data whose derivations were described above.

Applying Eq. (19) on the grand EDF, F, resulted from the above multiplication yields

H_{s100yr} = 14.85 m by the following MATLAB command

H_{s100yr}= *X (find (abs (F- (1-(1/100) < ε))))*,

where

ε is the error, a small value e.g. 0.00001.

As typical pdf's of the monthly maxima, the empirical pdf 's of the 39996 simulated H_s monthly maxima for months 1 and 9 are shown in Fig. 6.8 and Fig. 6.9.

6.6.2.3 The proposed method: extrapolating from the annual maxima EDF

It is worth restating that according (Fisher & Tippet, 1928) there is a theoretical basis for the use of the extreme value dist- ributions to fit the maxima. The extreme value theory requires the data, whose maxim are used, to be identically distributed (i.d.), but annual wave height data are not i.d. so the theoretical basis of the method is questionable (Anderson et al.,

2001). The extreme analysis used in this study proposes an approach in which the empirical distribution of the 39996 simulated annual maxima is used in the maxima method, without the

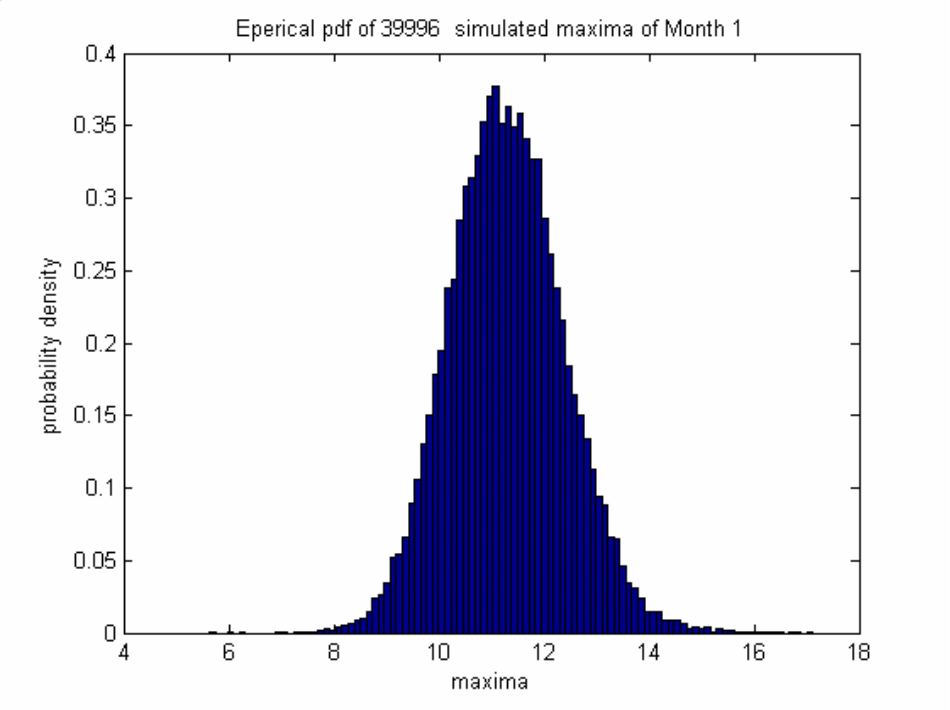


Fig. 6.8 Empirical pdf of 39996 simulated maxima for Month 1.

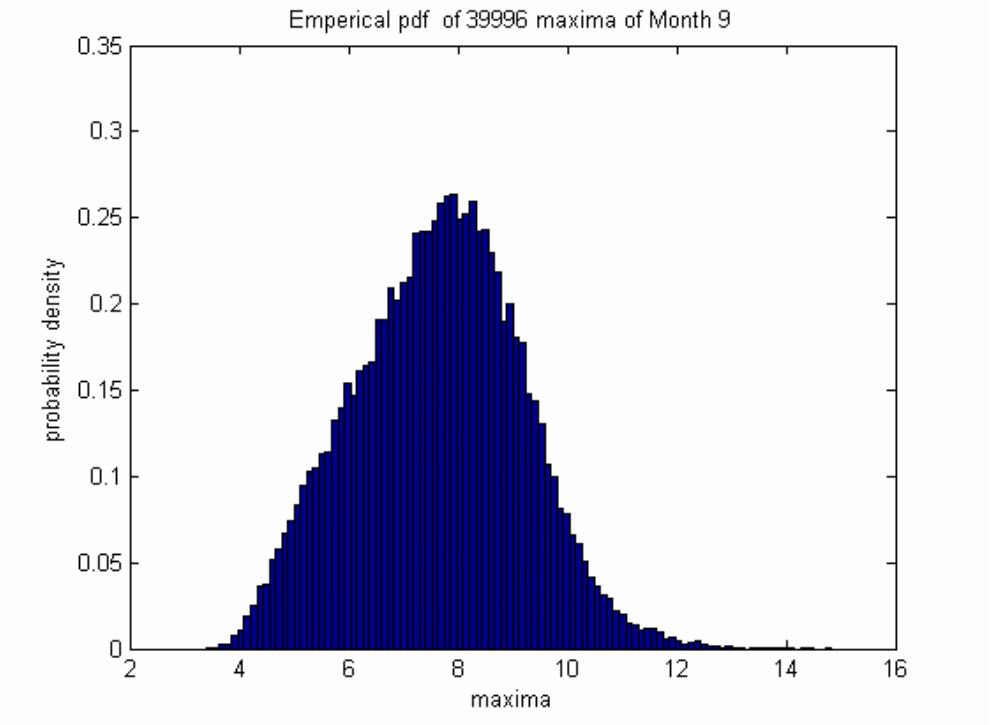


Fig. 6.9 Empirical pdf of 39996 simulated maxima for Month 9.

conventional need for the unrealistic assumption that they are drawn from identically distributed H_s's. To perform the extreme analysis using the maxima method, the simulated data for 39996 years beginning from 1 Jan 1978 were used. It could be shown that for well–trained networks, the EDF's of the simulated H_s's in this section will approach their corresponding actual CDF's as the number of simulation runs approaches a very large number. Since sufficient number of runs of simulation was available, the EDF's used here were taken as the representative of their actual CDF's. Figure 6.10 shows the empirical probability density function of 39996 simulated annual maxima and a generalised Pareto distribution (GEV) fit. Fig. 6.6 shows the Q-Q plot.

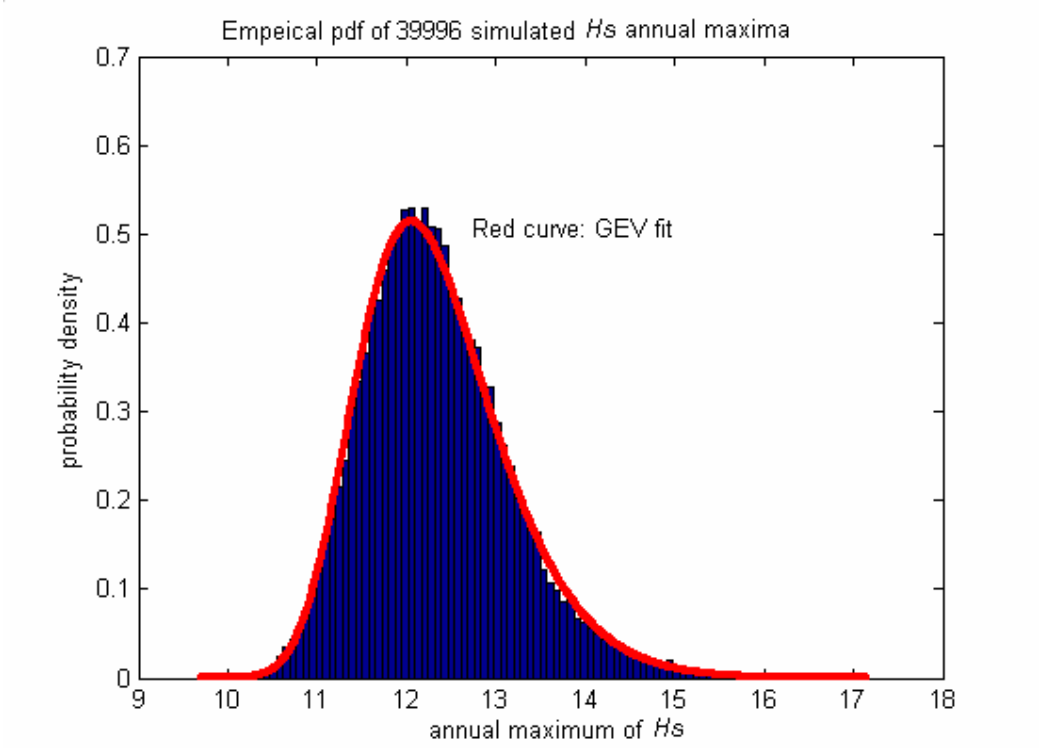


Fig. 6.10 The empirical pdf of 39996 annual maxima and a GEV fit.

Applying Eq.(19) to the EDF of 39996 annual maxima using the following MATLAB commands would yield a 100-year return value equal to 14.86 m.

[F,X] = ecdf(39996-Maxima-Vector); H_{s100yr} *= X (find(abs(F- (1-(1/100) <* ε *))))),*

where ε is a small value e.g. 0.00005.

In the above proposed method it was not assumed that the H_s's from which the maxima were extracted are independent and identically distributed, an assumption which is made in the conventional maxima method and is unrealistic in this case.

6.6.3 Method III: Peaks over Threshold (POT) Method

It is worth pointing out that the extreme value distributions are not suitable for the data sets containing the maxima above a certain level (Muir & El-Shaarawi, 1986). If the parent distribution of the data falls within the domain of one of the extreme value distributions, the excess values above a level follow asymptotically a generalised Pareto distribution (Pickands, 1975). The central idea of threshold analyses for high extremes is that attention should be focused on large values, i.e. only observations exceeding some high threshold = u are useful. If X denotes a continuous random variable, POT method is based on the distribution of X given X>u. For high enough u this distribution is taken to be generalised Pareto distribution (Anderson et al. 2001). Chapter 4 of Coles (2001) deals with the proof of this fact and its requirements. The pdf and the CDF of generalised Pareto distribution (GPD) are given in Appendix E where the location parameter A = u. Ferreira & Guedes-Soares (1998) describe this method as such : "The POT method consists of fitting the GPD to the peaks of *clustered excesses* over a threshold and of calculating the return value by taking into account the average number of *clusters*". It provides a more modern and soundly based solution to extrapolation problems. The POT method can address better the non-iid H_s data (Aderson et al., 2001) and enables the use of much more data than the annual maxima method. Goda (2000) and Coles (2001) cover the P.O.T method in detail. Goda et al. (1994), Ferreira & Guedes-Soares (1998) among others are researchers presenting the application of POT in the calculation of H_s return value. Anderson et al. (2001) proposes a new threshold method. POT method can be applied to calculating other extremes such as the extreme stresses of a marine structure. For this purpose time-series representative of the wave regime at the site of interest needs to be collected; the time-series of responses as a function of the time series of relevant sea-state and other parameters is constructed. After analyzing the resulting stress time series, the POT analysis is then applied if possible (Ferreira & Guedes-Soares, 1998). Goda (2000) 's attitude towards the distribution used with POT method is "There is no theoretical ground to recommend any distribution function a priori to the sample collected by peak-over-threshold method". The reason is that a sample collected by POT does not contain the extreme values in statistical sense. An extreme data is the maximum from a group of independent data and a sample of extreme data must be constructed from a number of sub-groups. Individual data of a sample preapred by POT represent merely the original data, which happen to be peaks of storm wave heights. Because the wave heights during a storm event are mutually correlated and not

independent, the peak height is not an extreme value. That is why 3-parameter Weibull, FT-I, FT-II, and log-normal distributions have been used with POT method besides generalised Pareto distribution by some researchers including Goda (2000).

6.6.3.1 Illustration for the conventional POT method

The simulated data of 39996 years were used for this illustration, several thresholds were chosen tentatively in order to use the conventional POT method and the peaks over the selected threshold were fitted to a GPD and inspected visually. Finally 14.3 was selected as threshold. The number of peaks over 14.3 in the 39996-year simulated H_s's was 989. The peaks were fitted to a GPD with parameters: A = 14.3, B= 0.65185 and C = 0.1684 using MATLAB. A Q-Q plot for the GPD fit is shown in Fig. 6.11. The return value for this method is calculated from the following formula extracted from Goda (2000)

$$H_{s\,N\,yr} = F^{-1}(1-\frac{1}{N\lambda}), \qquad (21)$$

where

F^{-1} the inverse function of the CDF of the peaks,

λ the mean number of peaks per year.

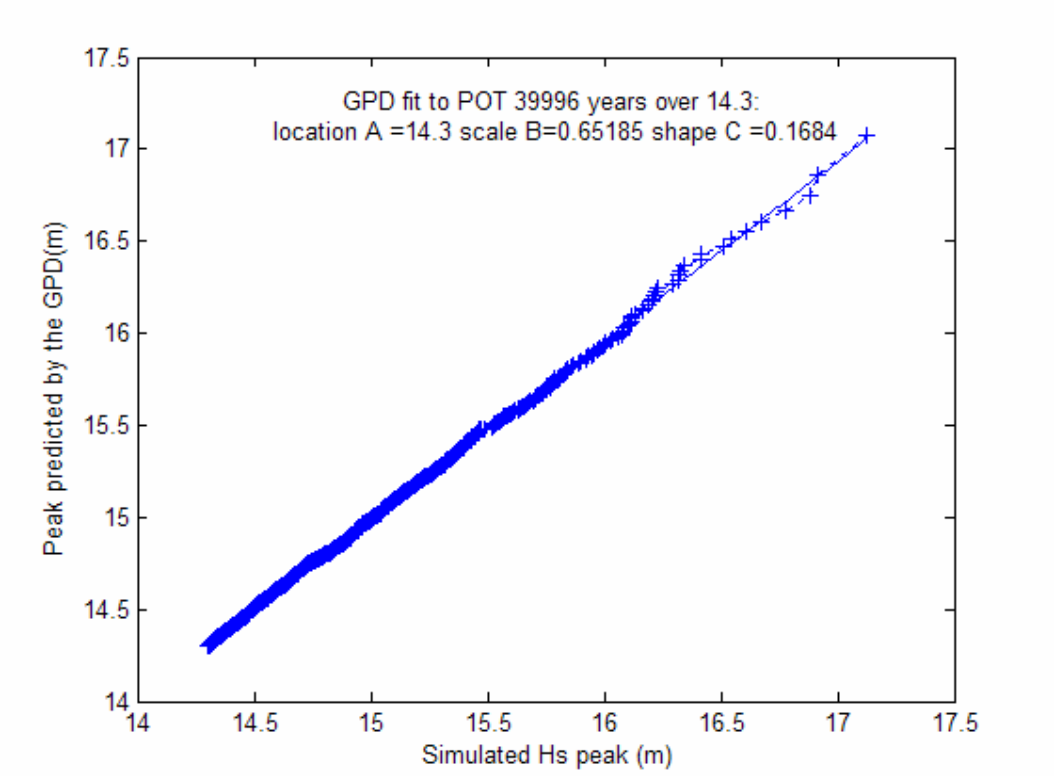


Fig. 6.11 A GPD fit to 989 peaks over 14.3 of the 39996 simulated H_s's, Q-Q plot.

In our case $\lambda = \frac{989}{39996}$ and the return period $N = 100$. The MTLAB command

H_{s100yr} = *gpdinv((1-1/(989/39966)*100), 14.3, 0.65185, 0.1684)* for applying Eq. (21) on the GPD results in $H_{s100\text{yr}} = 14.84$ m.

6.6.3.2 Illustration for using the peaks EDF

After finding the suitable threshold = 14.3 m by trial and error, the peaks extracted from the 39996-year simulated H_s's were stored in a vector named **peaks** 989×1. The empirical CDF (EDF) of these peaks was found using the MATLAB command

[CDFPOT, x] = ecdf (peaks).

The following MATLAB command which is based on Eq. (21) yields a 100-year return value equal to $H_{s\ 100\text{yr}} = 14.8$ m,

$H_{s\ 100\text{yr}}$ = X(find(abs(CDFPOT-1/((989/39996)*100))< ε))

where

ε is a small number found by trial and error e.g. 0.001.

6.7 Calculating the return value using a mathematical model of H_s

6.7.1 Discription of the Model

Anderson et al. (2001) have proposed the following model for significant wave heights observed during 1978-1999 at a US NODC Buoy in the North East Pacific (near 46°N, 131°W) showing the within-year variability of H_s with a long-term linear trend:

$$\log Hs = c_0 + c_1 \sin(\frac{2\pi d}{365.25}) + c_2 \cos(\frac{2\pi d}{365.25}) + c_5 d + res, \qquad (22)$$

where

log	natural logarithm
c_0, c_1, c_2, c_5	are constants,
d	is a real number from zero to 22*365.25 = 8035.5 representing the day in the period 1978-99 with $d = 0$ for 0 AM 1 Jan 78,
res = residual	is the error approximated with a normal distribution of zero mean and variance = 0.1517.

The coefficients of the above model were recalculated using MATLAB and the following values were obtained: $c_0 = 0.7964$, $c_1 = 0.1103$, $c_2 = 0.4408$, and $c_5 = 0.0000219362$. The sine portion of the Eq. (22) and the trend are sown in Fig. 6.12.

Attempting to reduce the residuals term, additional sine terms were added to Eq. (22). The terms added in two typical attempts are listed below

Attempt 1) $c_3\sin(4\frac{\pi d}{365.25}) + c_4\cos(4\frac{\pi d}{365.25})$,

Attempt 2) $c_3\sin(4\frac{\pi d}{365.25}) + c_4\cos(4\frac{\pi d}{365.25}) + c'_3\sin(\frac{\pi d}{365.25}) + c'_4\cos(\frac{\pi d}{365.25})$,

where

c_3, c_4,c'_3, and c'_4 are constants.

This, however, did not reduce the variance of residuals significantly and in one of the cases the residuals became worse; therefore the original model as appears in Eq. (22) was used for this study.

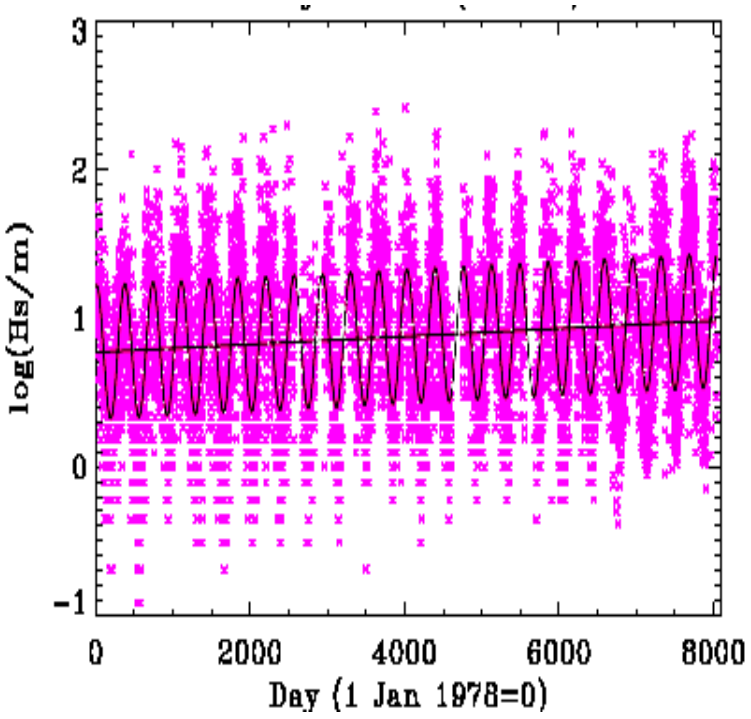


Fig. 6.12 Annual cycle and trend fitted to log(H_s's) of 1978-99 (after Anderson et al., 2001)**.**

It is worth pointing out that the distribution of $\log(H_s)$ cannot be found by applying the convolution concept directly to the annual cycle i.e. $c_1 \sin(\frac{2\pi d}{365.25}) + c_2 \cos(\frac{2\pi d}{365.25})$ and the residuals of Eq. (22), because the annual cycle is deterministic and furthermore applying a chi-square independence test on the values of the residuals and the values of the cycle revealed that these two sets of data, as expected, are not independent. Making both sides of Eq. (22) conditional on the event "*res* > *u*" where *c* is a high value of *res*, results in
($\log H_s$ | *res* > *u*) = c_0 + (annual cycle | *res* > *u*) + (*res* | *res* > *u*). As will be noticed later, there is a lesser degree of dependency between the distribution of *res* | *res* > *u*, which we call the distribution of the POT of *res,* and the distribution of (annual cycle | *res* >*u*) i.e. the distribution of those values of the sine portion of Eq. (22) that correspond to the POT of *res,* compared with the dependency of their parent distributions. The convolution was then conducted on the distribution function of the POT with that of the corresponding values of the sine portion using the concept of *sub-independence* described later. These two distributions will be shown to have a lesser degree of dependency on each other than that of their parent distributions.

Threshold Selection Method

For fitting POT of the residuals of Eq. (22) to a GPD, 1.331 was found to be the best threshold for the residuals, since at this value several measures of errors had their minimum. For finding this suitable threshold, the value of threshold was continually changed from zero to the maximum of the *res*, i.e. from 0 to 1.68, to arrive at a threshold that minimizes a measure of error such as root mean square error (RMSE) calculated from Eq. (30) of Chap. 4.

Table 6.2 The parameters(*A* = *u*,*B*, & *C*) of the *GPD* fitted to *res* > *u* & the corresponding *RMSE* for typical thresholds from the interval 0-1.68.

Threshold=*u*	0.90	0.98	1.01	1.07	1.10	1.15	1.19	1.25	1.29	**1.331**	1.37	1.40	1.46	1.50	1.54
no. of *res*>*u*	1462	860	716	510	427	299	221	145	109	**82**	62	49	24	17	8
A	0.90	0.98	1.01	1.07	1.10	1.15	1.19	1.25	1.29	**1.331**	1.37	1.40	1.46	1.50	1.54
B	.1572	.1719	.1796	.1544	.1422	.1469	.1619	.1510	.1578	**0.14563**	.1335	.0937	.1088	0.4562	.1072
C	-.0267	.1223	.2231	.1351	.0295	0.1188	.3289	.2647	.4424	**0.43045**	.4793	.1895	.7165	-.3797	1.1322
RMSE	0.0334	0.0138	0.0093	0.0098	0.0203	0.0139	0.0076	0.0075	0.0043	**0.0041**	0.0050	0.0095	0.0112	0.0489	0.0085

The RSME was calculated a measure of error between the observed peaks over threshold and their corresponding values predicted from the GPD fitted to the peaks over the threshold. For

these calculations the plotting position given in Eq. (11) was used. Table 6.2 shows the root mean square for typical thresholds and the values of the parameters of the GPD's fitted to the the residuals above a threshold. As the table shows RMSE has its minimum value for threshold $u = 1.331$.

Another threshold selection method has been extracted from Coles (2001) and is shortly described and illustrated here. In this method the mean $\overline{X}_u = \sum_{i=1}^{n} x_{(i)} - u$ is calculated for each $u <$ max (res), where $x_{(1)}, x_{(2)},..., x_{(n)}$ are the n values of the residuals that exceed the threshold = u. Then $\{u, \overline{X}_u : u < \max(res)\}$, which is termed the **mean residual life plot,** is plotted for $u = 0$ to the maximum of res. A confidence interval can be added to the plot based on the approximate normality of $\overline{X}_u$. Figure 6.13 shows the plot with the 95% confidence interval for the values of res in Eq. (22). The middle curve, called **mean residual life plot**, is the graph of mean excess versus threshold where *excess* equals a peak value minus threshold. The side curves constitute the 95 percent confidence interval of the excess. The graph appears to be approximately linear beyond 1.3. The procedure determines $u = 1.3$ as a good initial threshold.

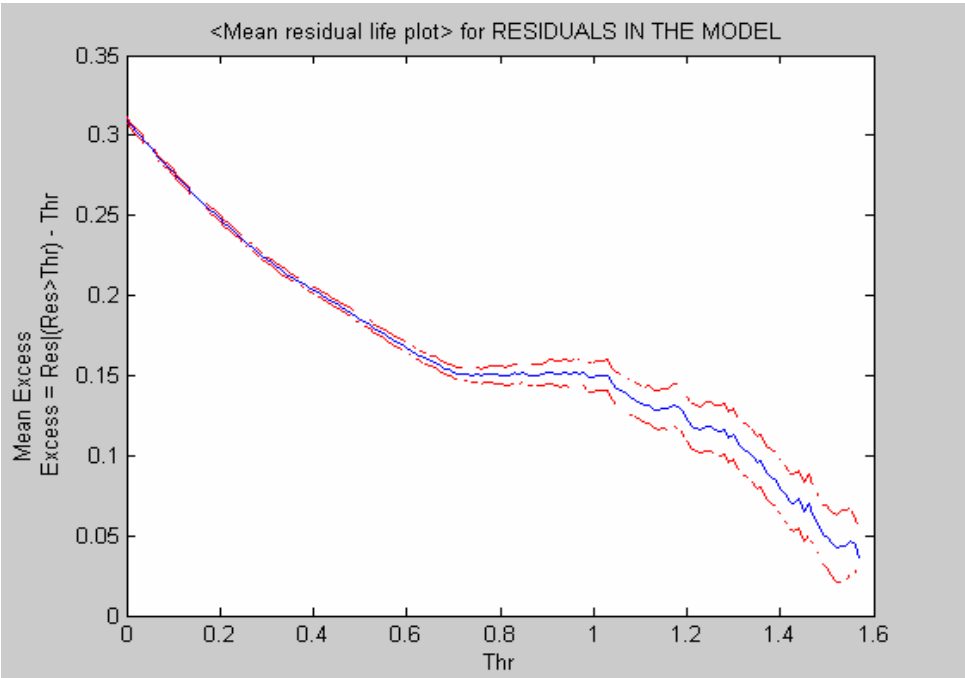


Fig. 6.13 "Mean residual life" plot for the residuals of Eq. (22).

6.7.2 The Proposed Approach for Calculating the Return Value

In the proposed approach for the calculation of the return value only those H_s's are considered whose corresponding residuals in the model are greater than a high threshold. This could result in a reasonably justified return value of significant wave height. In other words the return value is related to *res* greater than a high threshold due to the following facts:

The first thing is that the return values are among the high values of H_s, and also Pr ($H_s > c$) can be written as:

$$\Pr(Hs>c)=\Pr(Hs>c|res>u)\times\Pr(res>u)+\Pr(Hs>c|res\le u)\times\Pr(res\le u), \qquad (23)$$

where

res *res* in Eq. (22),

u a threshold,

c a high value of H_s.

We are interested to find the values of c satisfying $\Pr(Hs > c \mid res \le u) \cong 0$.

It could be proved that:

For $u = 1.331$, $\Pr(Hs > c \mid res \le u) \cong 0$ is satisfied with $c > 14$.

Proof:

The observed H_s's show that max $\log_e \{H_s \mid res \le 1.331\} = 2.5649$ or equivalently max $\{H_s \mid res \le 1.331\} = \exp(2.5649) = 12.99$ m, while the maximum of all observed H_s's during 1978-99 equals 13.6 m. It might be interesting that, based on the observed 151355 H_s's during 1978-99, the amount of the right hand side of $\Pr(Hs > c \mid res \le u) \cong 0$ for some values of c is as follows:

for $c = 12.0$ $\Pr(H_s > 12.0 \mid res \le 1.331) = \Pr(H_s > 12 \ \& \ res \le 1.331)/\Pr(res \le 1.331) = \dfrac{\frac{7}{151355}}{\frac{151328}{151355}} = \dfrac{7}{151328}$,

for $c = 12.5$ $\Pr(H_s > 12.5 \mid res \le 1.331) = \dfrac{3}{151328}$,

for $c = 13.0$ $\Pr(H_s > 13.0 \mid res \le 1.331) = \dfrac{1}{151328}$, and

for $c = 14.0$ $\Pr(H_s > 14.0 \mid res \le 1.331) = \dfrac{0}{151328} = 0$.

Therefore the fact that $\Pr(Hs > c \mid res \le u) \cong 0$ is satisfied for any $c > 14$ is strongly accepted. Note that the high number of observations guarantees that the statement could be even true for $c > 13$.

Furthermore it can even be argued that the statement is true for any c >12 or any c > 12.5 since only 7 observations out of 151328 observed H_s's whose $res < u$ are greater than 12 and only 3 of them are greater than 12.5. Thus $\Pr(Hs > c \mid res \le u) \cong 0$ is valid with a great certainty for any c >14 when u = 1.331. Calculating a return values implies identifying very high values of H_s such as c that satisfy

$$\Pr(Hs > c) = p, \tag{24}$$

where p is a given probability.

On the other hand, using the above reasoning, Pr ($H_s > c$) could be considered to be approximately zero for high values of c when the threshold $u \le 1.331$, then for $u \le 1.331$ and high values of c

$$\Pr(Hs > c \mid res \le u) \cong 0. \tag{25}$$

Hence from Eqs. (23) and (25) and the aforementioned lemma, for $u \le 1.331$ and c >14 we have:

$$\Pr(Hs > c) = \Pr(Hs > c \mid res > u) \times \Pr(res > u),. \tag{26}$$

According to Eq. (26) H_s high values such as the return values are related to *res* greater than a relatively high threshold u. This fact will be used in the calculation of the return value frm the model described in Eq. (22).

Now,it is demonstrated why the distribution functions of POT of the residuals were convovled with the distribution function of those values of the annual cycle, which correspond to the residuals greater than the threshold. Considering Eq. (22) one could write:

$$\Pr(\log Hs > \log(c) \mid res > u) = \Pr(c_0 + \text{annual cycle} + c_5 d + res > \log(c) \mid res > u), \tag{27}$$

or

$$\Pr(\log Hs > \log(c) \mid res > u) = \Pr(\text{annual cycle} + res > \log(c) - c_0 - c_5 d \mid res > u). \tag{28}$$

Let the " annual cycle" and "*res*" of Eq. (22) be denoted by Y and X respectively, it can be shown (see Appendix F) that

$$\Pr(X + Y > b \mid X > u) = \Pr[(X \mid X > u + Y \mid X > u) > b], \tag{29}$$

where b is a positive number.

Therefore from Eqs. (28) and (29):

$$\Pr(\log Hs > \log(c) \mid res > u) = \Pr(\text{annual cycle} \mid res > u + res \mid res > u > \log(c) - c_0 - c_5 d). \tag{30}$$

Once the probability distribution of H_s is known, given M observations each year for H_s, the following formula may be used for calculating the N-year return value of significant wave height, $H_{s\mathrm{Nyr}}$ (Anderson et al.,2001),

$$\Pr\left(H_s > H_{s\mathrm{Nyr}}\right) = \frac{1}{MN}. \tag{31}$$

From Eqs. (26) and (31):

$$\Pr(Hs > c \mid res > u) \times \Pr(res > u) = \frac{1}{MN}. \tag{32}$$

Pr ($res > u$) can also be estimated from

$$\Pr(res > u) = \frac{\lambda}{M}, \tag{33}$$

where

λ the mean number of peaks over threshold per year,

M number of observations per year.

Substituting $\Pr(res > u)$ from Eq.(33) in Eq. (32) results in

$$\Pr(Hs > c \mid res > u) = \frac{\frac{1}{MN}}{\frac{\lambda}{M}} = \frac{1}{N\lambda}. \tag{34}$$

From Eqs. (27), (30), and (34) one arrives at

$$\Pr[(\text{annual cycle} \mid res > u + res \mid res > u) > \log(c) - c_0 - c_5 d] = \frac{1}{\lambda N}. \tag{35}$$

To use Eq. (35) for calculating the return value, the distribution function of the sum of the two random variables annual cycle| $res > u$ and $res \mid res > u$ is needed. That is why the distribution function fitted to ($res \mid res > u$) and that of (annual cycle | $res > u$) were convolved to arrive at the distribution function of their sum.

6.7.2.1 The Distribution fitted to the peaks of the residuals

Among all values of the residuals of the model in Eq. (22), 82 values are greater than the high threshold u = 1.331. Assuming that these 82 values are independent of each other, according to Anderson et al. (2001) and theorem 4.1 in Coles (2001) a generalized Pareto distribution could be found to fit well to these peak values (i.e. to $res \mid res > u$). By using a downloaded MATLAB code, the three parameters of the generalized Pareto distribution fitted to these 82 values were estimated as such: A = 1.331, B = 0.14563, and C = 0.43045. The code

uses a fitting method called Pickands' estimator method as default for estimating the parameters. A Q-Q plot of the observed versus the predicted residuals is shown in Fig. 6.14. As the figure shows, the generalized Pareto distribution fits the values of the *res* over threshold $u = 1.331$ well.

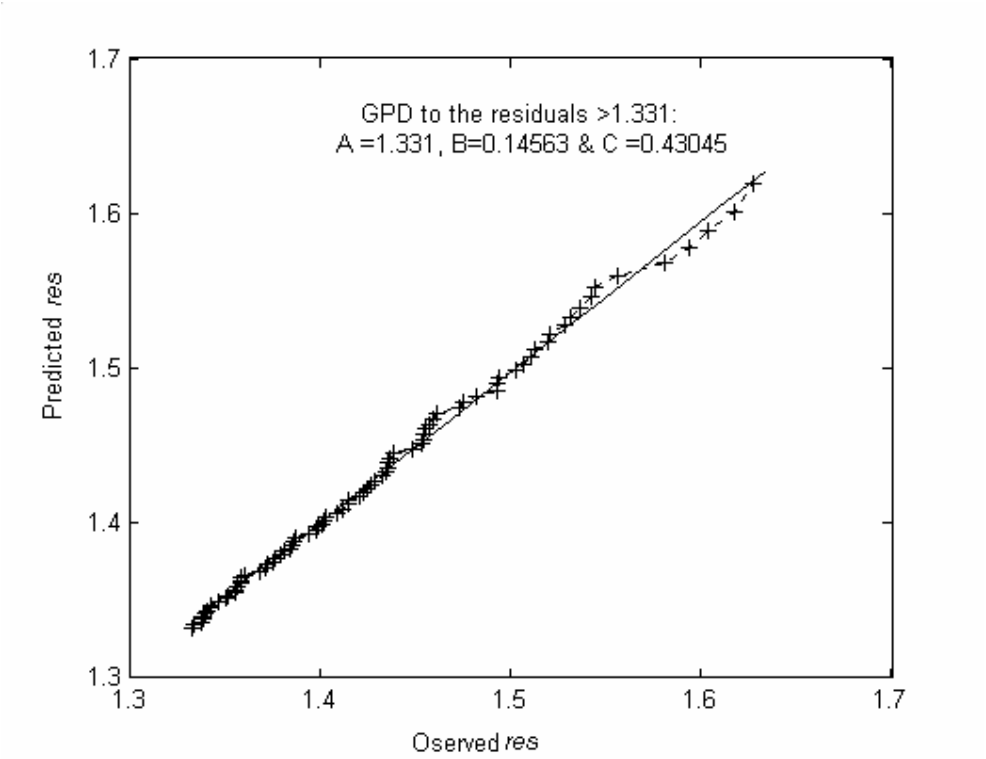


Fig. 6.14 The GPD fitted to the residuals over 1.331, Q-Q plot.

6.7.2.2 The Distribution of the Annual Cycle values corresponding to the peaks

The Weibull and the generalised Pareto distributions were both fitted to the 82 values of the annual cycle that correspond to *res* peaks over 1.331 (i.e. to annual cycle | *res* >1.331). As Fig. 6.15 shows the GPD fits well. The 3 parameters of the fitted GPD were $A = -0.38786$, $B = 0.27656$, and $C = 0.30129$. It is worth pointing out that visual inspection showed that the values of (annual cycle | *res* > *u*), fitted a GPD better as the threshold was changed continually from 0 to 1.331.

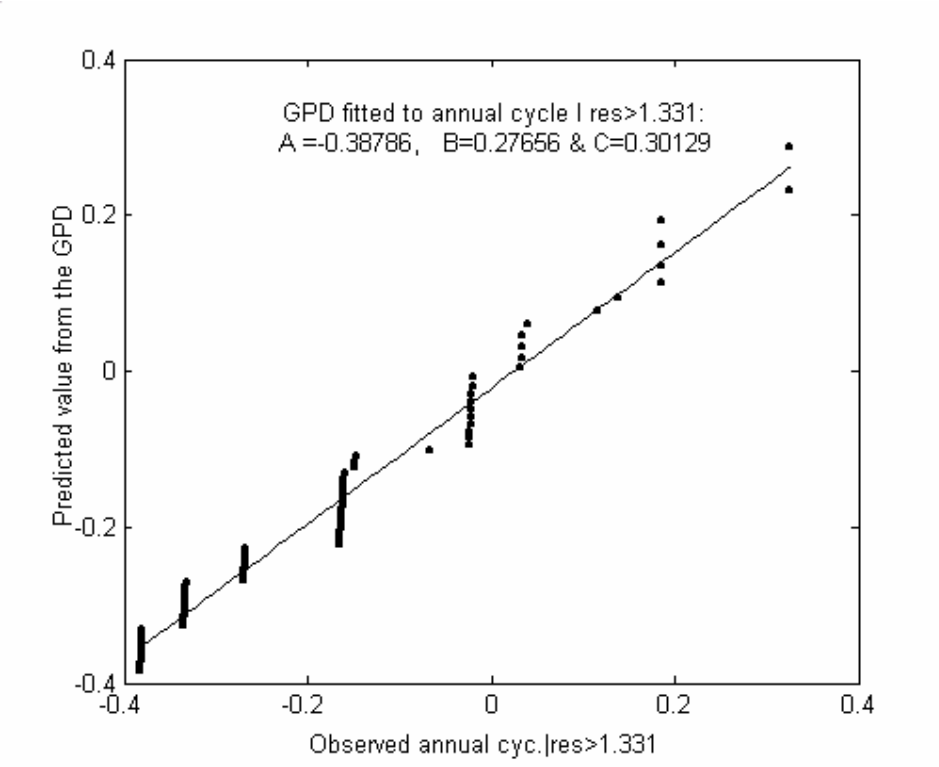


Fig. 6.15 A Q-Q plot for the GPD fitted to the annual cycle values for which the corresponding residuals are greater than 1.331.

6.7.2.3 Testing independence and sub-Independence

A chi square test for independence showed that (*res* | *res* > *u* = 1.331) fitted to a GPD distribution with *A* = 1.331, *B* = 0.14456, and *C* = 0.43045 and (annual cycle | *res* > *u* = 1.331) fitted to another GPD with *A* = 0.38768, *B* = 0.027656,and *C* = 0.30129 were not independent random variables. However they were approximately uncorrelated because their covariance is nearly zero (Cov. = 0.0049). On the other hand their joint characteristic function (C.F.) was calculated and found to be approximately equal to the product of the marginal C.F's. Hence these two random variables are stronger than two uncorrelated random variables and much weaker than two independent random variables therefore, the best that can be said is that they are *sub-independent.* Two random variables X and Y are said to be *sub-independent* if the distribution function of their sum is given by the convolution of their distribution functions (Hamedani & Walter, 1984) i.e.

$$F_{X+Y}(z) = (F_X * F_Y)(z), \qquad (36)$$

where

$F_X * F_Y$ denotes the convolution of the distribution function of X and that of Y,

F_{X+Y} is the distribution function of the sum of random variables X and Y.

Alternatively, in terms of characteristic functions, it is defined by

$$\Phi_{X,Y}(t, t) = \Phi_X(t)\, \Phi_Y(t), \tag{37}$$

where

$\Phi_{X,Y}(t, u)$ the joint characteristic function of X and Y

$\Phi_X(t)$ the marginal characteristic function of X

$\Phi_Y(t)$ the marginal characteristic function of Y.

Using the definition of sub-independent random variables, the two generalised Pareto distribution functions were convolved to obtain the distribution function of the sum of the two random variables i.e. (*res* | *res* > *u* = 1.331) and (annual cycle | *res* > *u* = 1.331) for calculating the return value of significant wave heights from Eq. (35).

6.7.2.4 Calculation of the return value

Now to calculate the return value of H_s, let Z = (annual cycle | *res*>*u* + *res* | *res*>*u*) and Z_{extr} be that value of Z which satisfies Eq. (38)

$$\Pr (Z > Z_{extr}) = \frac{1}{\lambda N}. \tag{38}$$

Comparing Eqs. (35) and (38) results in

$$Z_{extr} = \log (c) - c_0 - c_5 d. \tag{39}$$

Since, by definition, c is any high value of H_s including $H_{s\,Nyr}$, therefore

$$\log H_{sNyr} = c_0 + Z_{extr} + c_5 d. \tag{40}$$

It was observed that 82 values of *res* exceeded the threshold u = 1.331 during 22 years from 1978 to 1999 except some gaps in the data which are approximately 15 months or 1.25 years. Therefore λ is:

$$\lambda = \frac{82}{22 - 1.25} = 3.9518. \tag{41}$$

Substituting λ from Eq. (41) and N = 100 into Eq. (38) results in

$$\Pr(Z>Z_{extr}) = \frac{1}{100\lambda} = 0.0025. \quad (42)$$

The convolution of the two distribution functions and calculation of the extreme value from Eq. (42) were carried out using MATLAB, resulting in $Z_{extr} = 1.8589$.
Substituting $c_0 = 0.7964$, $c_5 = 0.0000219362$ and $Z_{extr} = 1.8589$ into Eq. (40) yields

$$H_{s100\,yr} = \exp(0.7964 + 1.8589 + 0.0000219362d), \quad (43)$$

$$H_{s100\,yr} = 14.2293 \times \exp(0.0000219362d). \quad (44)$$

Ignoring 0.0000219362 would yield H_{s100yr} =14.23 m. This return value compares well with the return values reported by Anderson et al. (2001) in which several samples from the full 1978-1999 data set of the NE Pacific have been selected for calculating the 100-year return value for the region. Depending on the method and data set used, the return value reported by them varies between 13.34 m to 17.3 m.
It should be pointed out that:

The above calculations for the return values from the model in Eq.(22) imply an assumption that if we shift the starting day i.e. $d = 0$ from 1 Jan 1978 to the first day of January of the year in which the structure is installed, a model similar to Eq. (22) is applicable again. And that although more observations were available, for comparison purposes we used the same data as Anderson et al.(2001) had used.

This method could be viewed as a modification of traditional methods since it explicitly takes the intrinsic variability of H_S into account. The method, in a general sense, could be best applied to calculating the return value for a non-stationary process for which a mathematical model could be developed. The model needs to be composed of a deterministic part representing the intrinsic variability of the process, and also a random residual part.

References (in alphabetical order)

Anderson, C. W., Carter D. J. T., & Cotton, P. D., 2001.
Wave Climate Variability and Impact on offshore Dsign Extreme.
Report prepared for Shell International and the Organization of Oil & Gas Producers
http://info.ogp.org.uk/metocean/schedules.html.

Barltrop, N. D. P., and Adams, A. J., 1991.
Dynamics of Fixed Marine Structures.
Butterworth–Heinemann, Oxford.

Bury, K., 1975.
Statistical Models in Applied science.
John Wiely & Sons.
Republished by Robert E. Krieger Pub. Co., Inc., Malabar, FL, 1986.

Carter, D. J. T. and Challenor, P. G., 1981.
Estimating return values of environmental Parameters.
Quart. J. Royal Meteor. Society 107 pp 259-266.

Carter, D. J. T., and Challenor, P. G., 1983.
Methods of fitting the Fisher –Tippett type 1 extreme distribution.
Ocean Engineering 10(3), 191-199.

Carter, D. J. T., Challenor, P.G., Ewing, J.A., Pitt, E.G., Srokosz, M.A. and Tucker M.J. 1986.
Estimating wave climate parameters for engineering applications.
Report OTH 86 228, Dept of Energy, London HMSO.

Cartwright, D.E., and Longuet-Higgins, M.S., 1956.
The statistical Distribution of the Maxima of a Random Function.
Proc. Royal Soc. of London Series A Vol 237.

Castillo, E., 1988.
Extreme value theory in Engineering.
Academic Press, San Diago.

Coles, S., 2001.
An Introduction to Statistical Modelling of Extreme Values.
Springer Verlag, London.

Dunne, J.F., & Ghanbari M., 2001.
Efficient extreme value prediction for nonlinear beam vibrations using measured random response histories.
Nonlinear Dynamics, 24, pp74-101.

Effron, B., 1979.
Bootstrap Method – Another Look at the Jackknife.
The Annals of Statistics, 7(1), pp 1-26.

Fisher, R. A., and Tippett, L. H. C., 1928.
Limiting Forms of the frequency by distribution of the Largest or Smallest Members of a Sample.
Proc. Cambridge, Phil. Soc. 24 180-190.

Ferreira, J.A., and Guedes Soares, C., 1998.
An Application of the Peaks Over Threshold Method to Predict Extremes of Significant Wave Height.
Jr. of Offshore Mechanics and Arctic Engineering 120, pp165-176.

Ferreira, J.A., and Guedes Soares, C., 1999.
Modelling the long term distribution of significant wave height with the Beta and Gamma models.
Ocean engineering, 26(8), pp 713 -725.

Ferreira, J.A., and Guedes Soares, C., 2000.
Modelling Distributions of Significant Wave Height.
Coastal Engineering, 40, pp.361-374.

Galambos, J., 1987.
The Asymptotic Theory of Extreme Order Statistics.
Kereiger, Florida, 2nd edition, 1st edition by John Wiley, New York.

Galambos, J., Leigh, S. and Simiu, E., 1994.
Extreme Value Theory and applications.
Kluwer, Amestredam.

Goda, Y., 1988.
On the methodology of selecting design wave height.
Proc. 21st Int. Conf. on Coastal Engineering ASCE pp 899-913.

Goda, Y., 2000.
Random Seas and Design of Maritime Structures.
2nd edition, World Scientific, Singapore.

Goda, Y., Hawkes, P., Mansard, E., Martin, M.J., Mathiesen, M., Poltier, E., Thompson, E., and van Vledder, G., 1994.
Recommended practice for extreme wave analysis.
Jr. of Hydraulic Research, 32(6), pp 803-814.

Goda, Y., Hawkes, P., Mansard, E., Martin, M.J., Mathiesen, M., Poltier, E., Thompson, E., and Van Vledder, G., 1993.
Intercomparison of extremal wave analysis methods using numerically simulated data
Proc. WAVES 1993 Conf. 26-28 July 1993.
New Orleans, USA pp 963-977.

Goda, Y., Konagaya, O., Takeshita, N., Hitomi, H., and Nagai,T. 2000.
Population Dist. of Extreme Wave Heights Estimated through Regional Analysis
Proc. 27th Int. Conf. Coastal Eng'g.

Gringorten, I., I., 1963.
A plotting rule for extreme probability paper.
J. Geophysic Resea. 68(3), pp 813-14.

Guedes Soares, C., 2003.
Advances in Coastal Modeling, Chapter 6
edited by C. Lakhm.
Elsevier Science B.V.
www.waterbouw.tudelft.nl/public/gelder/citatie77.pdf.

Guedes Soares, C., and Hennques, A. C., 1996.
On the Statistical Uncertainty in Long term Prediction of Significant Wave Height.
ASME Jr .of Offshore Mechanics and Arctic Eng'g ,18, pp 184-291.

Hamedani, G.G., and Walter, G.G. 1984.
A fixed point theorem and its application to the central limit theorem.
Arch. Math, 43, pp 258-264.

Harris, R. I., 2001.
The accuracy of design values predicted from extreme value analysis.
Jr. of wind Engineering & Industrial Aerodynamics, 89, pp153-164.

Harper, B. A., 1998.
Extreme Wave Height Data Analysis: Review and Recommendations.
Dept. of Environ. and Heritage, Conserv. Tech Rep No. 12, RE273, ISSN 1037-4701.

Huber, P. J., 1981.
Robust Statistics.
John Wiley and Sons, 308pp.

Huston, W. B., and Sponkinski, T. H., 1956.
Property and Frequency Characteristics of Some Flight buffet Loads.
NACA Tech. Note 3733.

Harver, S. 1985.
Wave climate off Northren Norway.
Appliied Ocean Research, 7, 85 , 85-92.

Jaspers, N. H., 1956.
Statistical Distribution Patterns of Ocean Waves and of Wave induced stresses and motions with engineering applications.
Trans. Society Nav. Archit. Mar. Eng., 64, pp 375-432.

Johnson, N.L., & Kotz, S., 1970.
Continuous Univariate Distributions.
Houghton_Miffin, Boston, Mass.

Jerkinson, A. F., 1955.
The Frequently Distribution of The Annual Maximum (or Minimum) Values of Meteorological Elements.
Quarterly Jr. of The Royal Met. Soc. Vol. 81.

Isaacson, M., & Mackensie, N.G., 1981.
Long term distributions of Ocean Waves – A Review.
Jr. Waterway, Port Coastal Ocean Div. Am. Soc. Civ. Eng'g., 107, pp 93-109.

Lawless, J. F., 1982.
Statistical Models and Methods for Lifetime data.
John Wiley, New York. Chapter 9.

Leadbetter, M. R., Lindgren, G. & Rootzen, H., 1983.
Extremes and Related Properties of Random Sequences and Series.
Springer, New York.

Lettenmaier, D. P., and Burges, S. J., 1982.
Gumbel's Extreme Value I Distribution: A New Look.
Jr. Hydraul. Div. Am. Soc. Civ. Engrs; 108(HY4), pp 502-514.

Longuet-Higgins, M.S., 1952.
On the Statistical Distributions of the Heights of Sea Waves.
Jr. Marine Research, 9(3), pp 245-266.

Montgomery, D. C., and Runger, G. C., 1994.
Applied Statistics and Probability for Engineers.
John Wiely.

Muir, L. R., and El-Sharaawi, A. H., 1986.
On the Calculation of Extreme Wave Heights: A Review.
Ocean Engineering, 13(1), pp 93-118.

Nolte K. G., 1973.
Methods for Detecting Extreme Sea States.
Proceedings 2nd international Conference on Port and Ocean Engineering Under Arctic Conditions pp, 705-742.

Ochi, M. K., 1973.
On Prediction of Extreme Values.
Jr. of Ship Research, 17(1), pp 29-37.

Ochi, M. K., 1992.
New Approach for Estimating the Severest Sea-state from statistical data.
Proc. 23rd Int. Conf. on Coastal Engineering, 1, 512-525, Venice.

Ochi, M. K., 1998.
Ocean Waves: The stochastic Approach.
Cambridge University Press, Cambridge.

Ochi, M.K., and Whalen J.E., 1980.
Prediction of the Severest Significant Wave Height.
Poc. 17th Coastal Engineering, 1, 587-599, Houston.

Petrauskas,C., and Aagaard, P.M., 1971.
Extrapolation of Historical Storm Data for Estimating Design Wave Heights.
Society of Petroleum Engineering Journal, 11, pp 23-37.

Pickands, J., 1975.
Statistical inference using order statistics.
Annals of Statistics, 3, 119-131.

Reiss, R-D., and Thomas, M., 2001.
Statistical Analysis of extreme values with applications to insurance, Finance, Hydrology and other fields.
2nd edition, Birkhauser Verlag AG, Boston, MA.

Resnick, S. I., 1987.
Extreme values Regular variation, and Point Processes.
Springer Verlag, New York.

Sorenson, R.M., 1997, Traslated into Farsi by Bargi, K., 2001.
Basic Coastal Engineering.
Tehran University Press, Iran.

Thom, H.C. S., 1970.
Asymptotic Extreme Value Distributions of Wave Heights in the Open Sea.
Jr. Marine Research ,29(1), pp 19-27.

Tryon, R.G., & Cruse, T.A., 2000.
Probabilistic mesomechanics for high cycle fatigue life prediction.
Jr. of Engineering Material and Technology, Transactions of ASME 122, pp 209-214.

Rice, O., 1944.
Mathematical Analysis of Random Noise.
Bell System Technical Jr., Vol 1944.

Rice, O., 1945.
Mathematical Analysis of Random Noise.
Bell System Technical Jr, Vol 1945.

van Vladder, G., Goda, Y., Hawkes, P., Mansard, E., Masrad Emartin, M. J., Mathiesen, M.,Peltier, E., and Thompson, E., 1993.
Case Studies of Extreme Wave Analysis- Comparative Analysis.
Proc. of 2nd Intern. Symposium on Ocean Wave Measu. & Analysis pp 978-992.

Tucker, M.J., and Pitt, E.G., 2001.
Waves in Ocean Engineering.
Elsevier Ocean Engineering Series.

Want, J., and Takefuji, Y., (Editor) 1993.
Neural Networks in Design and Manufacturing.
World Scientific.

Wilk, M.B., and Ganadersiken, R., 1968.
Probability Plotting Methods of the Analysis Data.
Biometrika, 55, pp 1-17.

Chapter 7
Conclusions and Recommendations for Future Work

Nomenclature

ANN(s)	artificial neural network(s)
FEM	finite element method
H_S	significant wave height
T_z	mean zero up-crossing period

7.1 Conclusions

7.1.1. The ANNs based Simulation

Based on the values obtained from the simulation and observation, it is concluded that:

1) The predicted 3-hourly significant wave heights and mean zero-up-crossing wave periods of the sea-state of a time in the near future compare well with their corresponding observed values.

2) In this study multi-layer feed-forward ANNs proved to be inefficient in the direct prediction of significant wave heights from neural network outputs given some input, however, these kind of ANNs proved their efficiency in estimating the parameters of the proposed distribution used to simulate the characteristics of the sea-state of a desired time and thereby in the simulation of sea-states sequences. Whilst conventionally other performance functions such as mean square error are used in ANNs training process, the performance criterion in training our networks which was a function based on the statistical likelihood function worked very well.

3) The proposed probability distribution called hepta-parameter spline approximating the conditional probability distributions used to simulate the significant wave height (H_S) and mean zero-up-crossing wave period (T_z) of a 3-hourly sea-state in the North East Pacific showed to be a good choice for the approximation.

4) The standard simulated annealing algorithm, which was applied successfully to the maximization of a multi-variable function utilized for training of the neural networks, showed

to be a rapid optimization technique compared to the gradient descent algorithm, that was used before simulated annealing to train the networks.

5) The mean of the conditional distribution or the mean of the simulated values obtained in several runs of simulation is a good forecast for the H_S and the T_z of a desired time sea-state, compared to the most probable value and a value obtained from one run of simulation.

6) The method introduced in this thesis for the simulation of sea-state characteristics is applicable to other time-dependent processes which have instinct periodical variations. The application could cover a variety of things such as simulating a related random variable, forecasting its most probable and mean values and also doing an extreme value analysis related to the process.

7.1.2 The Pile Analysis

The FEM and ANN modeling resulted an easy-to-use procedure in Chap. 5 to predict the safe sea-states for pile driving operation in a future period given the eight successive 3-hourly sea-state history preceding that period.

7.1.3 Extreme Wave Analysis

Some points are worth noting in the part of this thesis that deals with the extreme wave height analysis.

1) The simulated data used in the extreme wave analysis has been indirectly obtained from the output of some artificial neural networks whose inputs were all the observed 3-hourly significant wave heights. Therefore it could be concluded that, in a sense, the threshold method and the maxima method used in the present study have used all 3-hourly observed data contrary to the conventional peaks over threshold method and the maxima method. That is true also for the method proposed in Sec. 6.7 of Chap. 6, where a model presented by Anderson et al (2001) for significant wave heights in a region of North East Pacific was used to calculate the return value. Although the method is a kind of threshold method but the model constants were estimated using all of the observed 3-hourly H_s's. Therefore we conclude that, in a sense, the proposed method has used the entire observed significant wave heights contrary to peaks over threshold method.

2) A mathematical model of significant wave heights which could normally be used for other purposes, might be used reasonably for the calculation of H_s return value.

3) And finally, the return values calculated in this work [14.12 m as mean of 3000 return values obtained from applying total sample method to 3000 simulations, 15.29 m from the empirical distribution function of 1000-year simulation, 14.86 m from 39996 annual maxima empirical distribution, 14.8 m from the empirical distribution of the peaks of 39996-year simulated H_s's, and 14.23 m from the model presented by Anderson et al.(2001) for H_s] compare well with the return values calculated by Anderson et al. (2001). They have selected several samples from the full 1978-1999 observed data set of the North East Pacific in order to calculate the 100-year return value by some methods. Their reported return values vary from 13.34 m to 17.3 m. For example for applying total sample method they fitted all hourly observed H_s's during 1996-99 by the buoy to a Fisher-Tippet type I with location parameter A = 2.222 & scale parameter B = 1.074 and obtained H_{s100YR} = 15.74 m. A version of maxima method i.e compounding monthly maxima method used by them yielded H_{s100YR} = 17.3. Some of the 100-year return values obtained in this version for months are 14.92 m and 13.34 m.

7.2 Recommendations for Future Work

The following suggestions are put forward for future investigations.

7.2.1 The ANNs based Simulation

It should be pointed out that the selected networks for simulating H_S and T_z i.e. $13\times30\times30\times1$ and $10\times1\times15\times1$ are not necessarily the best network architecture that one could choose and train for the purposes. To get better trained networks and better results, while working with ANNs, needs enough time, patience, and enough number of computers to do more experiments with

other network structures,

other inputs,

other measures of terminating the training process,

other probability distributions, and

performing sensitivity analysis.

Other modern techniques such as genetic algorithm might be tested instead of ANNs.

7.2.2 The Pile Analysis

1) Performing the analysis with an improved version of the code used (Appendix D) might yield better results. For example it should be take into considerations that the number of

members with degraded material properties should be less than 2% of the total number of members in the model.

2) Using other wave energy spectrums to decompose the sea-states and comparing the results with those of obtained from Pierson-Moskowits'.

3) Repeating the dynamic analysis with the sea-states of Table 5.5 of Chap. 5 for a quite longer time than 300 s and calculation of the extreme von Mises stress occurring in the pile after 3 hours using Eq. (24) of Chap. 5 and comparing the results with those obtained in this study.

4)) If possible to repeat the dynamic analysis with the sea-states of Table 5.5 of Chap. 5 for 3 hours instead of 300 s with a very powerful computer and to obtain the extreme von Mises stress occurred in the pile after 3 hours directly from the analysis and comparing the results with those obtained in this study from Eq. (24) of Chap. 5 to check the accuracy of the equation.

3) Extracting the natural period of the system using nodal analysis, e.g. with *MODES and *FREQUENCY commands in ABAQUS, and running a few cases with a period around say 10 times the natural period. Note that the time of loading has must not be less than approximately 10 times the natural period. Since some wave periods are a high multiple of the natural period of the pile, as a result the structure may feel the wave load as a resonating load.

4) To run a few cases for each sea state but different phase angle, since each run (for a given sea state) gives one realisation of a random variable. Having more results for each sea-state helps to find a better estimate for the probable maximum Von Mises stress.

5) The pile in question is prone to buckling. It is recommended to extract the magnitude of of the axial forces, which can cause buckling, in all members under compression e.g. from FEM buckling analysis and compare it with a safe value recommended by an appropriate code such as API RP 2A-LRFD(1993) to see if the sea-state is safe or otherwise.

6) If possible to compare the maximum stresses obtained in this study for the pile with those obtained when designing the pile in order to give possible appropriate suggestions.

References (in alphabetical order)

Anderson, C. W., Carter, D. J. T., & Cotton, P. D., 2001.
Wave Climate Variability and Impact on offshore Design Extreme.
Report prepared for Shell International and the Organization of Oil & Gas Producers.
http://info.ogp.org.uk/metocean/schedules.html.

API RP 2A-LRFD, 1993.
Recommended Practice for Planning, Designing and Constructing Fixed Offshore Platforms–Load and Resistance Factor Design.
American Petroleum Institute.
1220 L Street, Northwest, Washington DC 20005 USA, Washington, D. C.

Appendix A :
The relationship between the cost function of ANN/SA algorithm for simulating Hs's, and the likelihood function of MLE method

To estimate the seven parameters of the hepta-parameter spline distribution by MLE method, the joint density function of the significant wave heights associated with 3-hourly H_s's in the period 1978-1999 i.e. the joint pdf of $H_s(t_0)$, $H_s(t_1)$, $H_s(t_2)$,…, $H_s(t_{62035})$ must be maximized; that is to maximize the likelihood function L given in Eq. (A-1)

$$L = f_{Hs(t_0),Hs(t_1),...,Hs(t_i),...,Hs(t_{62035})}{}^{(h_0,h_1,...,h_i,...,h_{62035})}, \qquad \text{(A-1)}$$

where:

$H_s(t_i)$ the random variable associated with 3-hourly significant wave height measured at time t_i where i = 0,1…62035,

h_i a particular observed or simulated value for $H_s(t_i)$.

The H_s's used as the target in training the ANNs had been observed for 21 years and 84 days i.e.7754.375 days from 1 Jan 1978. This data set was expected to consist of 7754.375×8 = 62035 H_s's. Applying the chain rule for factorization to Eq. (A-1) results in

$$L = \prod_{i=62305}^{8} f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};...,Hs(t_{i-m})=h_{i-m},..,Hs(t_0)=h_0}{}^{(h_i)} \times f_{Hs(t_0),Hs(t_1),...,Hs(t_7)}{}^{(h_0,h_1,...,h_7)} \qquad \text{(A-2)}$$

where

$f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};...,Hs(t_{i-m})=h_{i-m},..,Hs(t_0)=h_0}{}^{(h_i)}$

denotes the pdf of the conditional distribution of $H_s(t_i)$ given all of its previous H_s's (parents),

m an integer less than i, and

$f_{Hs(t_0),Hs(t_1),...,Hs(t_7)}{}^{(h_0,h_1,...,h_7)}$

is the joint distribution of the 8 initial H_s's.

Assume there exists a value for m such that all necessary information concerning the history of the H_s of time t_i, $H_s(t_i)$, is conveyed by the sequence $H_s(t_{i-1})$, H_s (t_{i-2}),…., $H_s(t_i\text{-}m)$. Therefore for all i = 62035 through m Eq. (A-3) hold:

$$f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};\ldots,Hs(t_{i-m})=h_{i-m},\ldots,Hs(t_0)=h_0}(h_i) \cong f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};\ldots,Hs(t_{i-m})=h_{i-m}}(h_i).$$

(A-3)

However finding the value of m, the order of H_s Markov chain, remains a problem to be tackled. It was assumed earlier that 8 immediate preceding measurements carry enough information to simulate $H_s(t_i)$; in other words $m = 8$. Therefore, from Eqs. (A-2) and (A-3), one could write

$$L \cong f_{Hs(t_0),Hs(t_1),\ldots,Hs(t_7)}(h_0,h_1,\ldots,h_7) \times \prod_{i=62035}^{8} f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};\ldots,Hs(t_{i-8})=h_{i-8}}(h_i),$$

or

$$L \cong f_{Hs(t_0),Hs(t_1),\ldots,Hs(t_7)}(h_0,h_1,\ldots,h_7) \times \prod_{i=62035}^{8} f_{H_S(t_i)given8}(h_i). \quad \text{(A-4)}$$

The joint density function $f_{Hs(t_0),Hs(t_1),\ldots,Hs(t_7)}(h_0,h_1,\ldots,h_7)$, which could be factorized by chain rule as the product of 7 conditional pdf's, is negligible in comparison with the remaining 62027 terms of Eq. (A-4). Hence we are able to conclude that when maximizing P as given in Eq. (A-5) the likelihood function L is being maximized.

$$P = \prod_{i=62035}^{8} f_{H_S(t_i)given8}(h_i) = \prod_{i=62035}^{8} f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1};\ldots,Hs(t_{i-8})=h_{i-8}}(h_i). \quad \text{(A-5)}$$

Since, among the 62035 data points, there were only 49736 points whose 8 preceding immediate 3-hourly H_s's were available in the site of reference ,P_1 as appears in Eq. (A-6) was used instead of P.

$$P_1 = \prod_{i=49736}^{8} f_{H_S(t_i)given\ 8}(h_i). \quad \text{(A-6)}$$

Finally the logarithms of the terms in P_1 i.e. the sum P_2 as shown in Eq. (A-7)

$$P_2 = \sum_{i=49736}^{8} \log f_{H_S(t_i)given\ 8}(h_i), \quad \text{(A-7)}$$

was used as the performance (cost) function to be maximized.

Appendix B : Calculation of the time & 4 season values from day values

The time at which the sea-state being predicted occurs is calculated from the following MATLAB command

time = *d*-floor(*d*),

where *d*=day is a number in the interval <0-365.25> with step 0.125 representing the day on which 3-hourly sea-states occur. *d*=0 represents 0 AM 1 Jan.

The four fuzzy membership values related to the season in which the sea-state occurs are calculated using the following four MATLAB instructions:

1) springvalue =Trianglefun(abs(*d*-midspring) /(365.25/4))

2) summervalue =Trianglefun(abs(*d*-midsummer) /(365.25/4))

3) fallvalue =Trianglefun(abs(*d*-midfall) /(365.25/4))

4) wintervalue =Trianglefun(abs (*d*-midwinter)/(365.25/4))

where

Trianglefun is the following MATLAB function:

function v = Trianglefun(*d*);

v = (1-abs (*d*)) (abs (*d)* <1)+ (1-abs (*d*-4)) (abs (*d*-4) <1)

and

since midwinter happens on the 36th day of the year (5th Feb.) ,therefore

midwinter =36; and also:

midspring =36+(365.25/4) = 127.3125 (5th May)

midsummer =36+(2)×(365.25/4) = 218.6250 (6th Aug.)

midfall =36+(3)×(365.25/4) = 309.9375 (6th Nov.)

Appendix C : A lemma and a collary

Consider $h_1,h_2,h_3,\ldots,h_8,\ldots,h_n$ a sequence of 3-hourly H_s's and suppose that the first 8 terms h_1, h_2, $h_3,\ldots,h_8$ are the initial observed H_s's which were given to the trained networks as part of input for the purpose of the simulation of H_s's of a period and $h_9,\ldots,h_n$ are the simulated significant wave heights for the times t_9, $t_{10},\ldots,t_n$.

Lemma

The sequence $h_9,\ldots,h_n$ of H_s's simulated by the trained networks is a random sample which has the conditional joint distribution of $H_s(t_9),\ldots, H_s(t_n)$ given $H_s(t_1) = h_1$, $H_s(t_2)=h_2, \ldots, H_s(t_8)=h_8$ as its density function..

Proof

Since a random sample, say $h'_9,h'_{10},\ldots,h'_n$, could be generated from the conditional joint distribution of $H_s(t_9)$, $H_s(t_{10}),\ldots, H_s(t_n)$ given $H_s(t_1)=h_1,H_s(t_2)=h_2,\ldots,H_s(t_8)=h_8$ by the following algorithm; and the code used for simulation in this thesis follows the steps of the algorithm, therefore the proof is complete; i.e. every sequence of H_s's obtained from the simulation of the well trained neural networks is considered a random sample from the conditional joint distribution of $H_s(t_9)$, $H_s(t_{10}),\ldots, H_s(t_n)$ given $H_s(t_1)=h_1$, $H_s(t_2)=h_2, \ldots, H_s(t_8)=h_8$.

The algorithm for generating a random sample of size n-8 for a desired time period from the joint distribution of the significant wave heights is as follows. The algorithm is based on the chain rule for factorization of a joint density function.

.Step 1 :

Generate random variate h'_9 from the conditional distribution of $H_s(t_9)$ given all its parents i.e. $H_s(t_1)=h_1$, $H_s(t_2)=h_2, \ldots, H_s(t_8)=h_8$.

.Step 2 :

Generate random variate h'_{10} from the conditional distribution of $H_s(t_{10})$ given $H_s(t_1)=h_1$, $H_s(t_2) =h_2,\ldots,H_s(t_8)=h_8,H_s(t_9)=h'_9$. In practice if a variate from the distribution of H_s (t_{10}) given its 8 preceding immediate successive H_s's i.e. $H_s(t_2)=h_2,\ldots, H_s(t_8)=h_8$, $H_s(t_9)=h'_9$ is generated, the variate is h'_{10} because substituting m=8 in Eq. (A-3) of Appendix A yields

$$f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1},Hs(t_{i-2})=h_{i-2},..,Hs(t_0)=h_0^{(h_i)}} \cong f_{Hs(t_i)|Hs(t_{i-1})=h_{i-1},Hs(t_{i-2})=h_{i-2}...,Hs(t_{i-8})=h_{i-8}^{(h_i)}}.$$

.

.
.

.Step j :

Generate random variate h'_j from the conditional distribution of $H_s(t_j)$ given all its 3-hourly previous H_s's. As stated in step 2, h'_j could be generated as a random variate from the conditional distribution of $H_s(t_j)$ given its eight 3-hourly preceding immediate H_s's.

.
.

.Step n-8 :

Generate another random variate say h'_n from the conditional probability distribution of $H_s(t_n)$ given $H_s(t_1)=h_1$, $H_s(t_2)=h_2$, …,$H_s(t_8)=h_8$,$H_s(t_9)=h'_9$, ….,$H_s(t_{n-1})=h'_{n-1}$. In practice this random variate could be generated from the conditional distribution of $H_s(t_n)$ given its 8 preceding immediate H_s's i.e. $Hs(t_{n-8})=h'_{n-8}$,…, $H_s(t_{n-2})=h'_{n-2}$, $H_s(t_{n-1})=h'_{n-1}$ according to Eq. (A-3) in Appendix A.

The above algorithm has been followed during the proposed simulation in this work, therefore the simulated sequence $h_9,h_{10},\ldots,h_n$ is a random sample from the conditional joint distribution of $H_s(t_9)$, $H_s(t_{10})$…, $H_s(t_n)$ given $H_s(t_1)=h_1$, $H_s(t_2)=h_2$, …, $H_s(t_8)=h_8$.

Collary

As stated earlier the value generated in step *j* is a random variate of the conditional distribution of $H_s(t_j)$ given $H_s(t_{j-1})=h_1$, $H_s(t_{j-2})=h_2$, ..., $H_s(t_{j-8})=h_8$. If this procedure is repeated for a very large number of times, the empirical distribution of this simulated H_s will approach the cumulative distribution function of the actual conditional distribution of H_s given its eight preceding immediate 3-hourly successive observed H_s's.

Appendix D: The code used for analysing the pile with ABAQUS/AQUA

```
**********************************************************************************************
**
**
** **
**
        ** **P I L E   S T I C K - U P   A N A L Y S I S
**
** **
**
             ** **O P T I O N     3
**
** **
**
: ** **
**
** **
**
              ** **May, 2004
**
** **
**
**
**********************************************************************
**
**       NOTES
**       *****
**       .1
 .2    **m & Newton  UNITS USED THROUGHOUT
**       .3
**
**
**************************************************************************
  **MODIFICATIONS
************  **
   **BY           DATE            COMMENTS
----------       -------       ----  **
                          **File modified to run with
 6.41version
   **SFY          4/05/04
**************************************************************************
**
**
**
**
**************************************************************************
**       JOB ATTRIBUTES
===============     **
**
 **Client contacts
---------------- **
**
*RESTART, WRITE, FREQ=1,OVERLAY
***HEADING
 **SABLE PILE STICK-UP ANALYSIS
**************************************************************************
**
```

```
**
**
          **NODES DEFINITIONS
**
**
**
**********************************************************************
*NODE
8,0.00000,0.00000
9,0.01709,1.51800
10,0.03941,3.50063
11,0.05021,4.45963
12,0.07318,6.50063
13,0.09007,8.00063
14,0.10696,9.50063
15,0.12384,11.00063
16,0.14073,12.50063
17,0.16278,14.45900
18,0.17450,15.50063
19,0.19139,17.00063
20,0.20828,18.50063
21,0.22516,20.00063
22,0.24205,21.50063
23,0.25894,23.00063
24,0.27582,24.50063
25,0.29271,26.00063
26,0.30960,27.50063
27,0.32648,29.00063
28,0.34337,30.50063
29,0.36026,32.00063
30,0.37715,33.50063
31,0.39403,35.00063
32,0.41092,36.50063
33,0.42781,38.00063
34,0.44469,39.50063
35,0.46158,41.00063
36,0.47847,42.50063
37,0.49535,44.00063
38,0.51224,45.50063
39,0.52913,47.00063
40,0.54601,48.50063
41,0.56290,50.00063
42,0.58576,52.03163
43,0.59667,53.00063
44,0.61356,54.50063
45,0.62489,55.50663
46,0.64733,57.50063
47,0.65781,58.43163
48,0.68106,60.49663
49,0.69862,62.05663
50,0.70515,62.63663
51,0.71544,63.55063
52,0.72573,64.46463
53,0.73602,65.37863
54,0.74631,66.29263
55,0.75660,67.20663
56,0.76689,68.12063
57,0.77718,69.03463
58,0.78747,69.94863
59,0.79776,70.86263
```

```
60,0.80860,71.82563
**
**
 *Nodes for gap elements
**
1.518        ,0.078   ,158
1.518  ,0.017089415   ,159
14.459  ,0.162777242   ,160
14.459        ,0.10177   ,161
**
**********************************************************************
**
**
          **ELEMENT DEFINITIONS
**
**
**
**********************************************************************
**
  **STRUCTURAL PILE ELEMENTS****
**
*ELEMENT, TYPE=B21 , ELSET=PILEALL
9     ,8      ,8
10    ,9      ,9
11    ,10     ,10
12    ,11     ,11
13    ,12     ,12
14    ,13     ,13
15    ,14     ,14
16    ,15     ,15
17    ,16     ,16
18    ,17     ,17
19    ,18     ,18
20    ,19     ,19
21    ,20     ,20
22    ,21     ,21
23    ,22     ,22
24    ,23     ,23
25    ,24     ,24
26    ,25     ,25
27    ,26     ,26
28    ,27     ,27
29    ,28     ,28
30    ,29     ,29
31    ,30     ,30
32    ,31     ,31
33    ,32     ,32
34    ,33     ,33
35    ,34     ,34
36    ,35     ,35
37    ,36     ,36
38    ,37     ,37
39    ,38     ,38
40    ,39     ,39
41    ,40     ,40
42    ,41     ,41
43    ,42     ,42
44    ,43     ,43
45    ,44     ,44
46    ,45     ,45
```

```
47    ,46     ,46
48    ,47     ,47
**
*ELEMENT, TYPE=B21 , ELSET=SOME_HAM
49    ,48     ,48
50    ,49     ,49
51    ,50     ,50
52    ,51     ,51
53    ,52     ,52
54    ,53     ,53
55    ,54     ,54
56    ,55     ,55
57    ,56     ,56
58    ,57     ,57
59    ,58     ,58
60    ,59     ,59
**
 **GAP ELEMENTS****
**
*ELEMENT, TYPE=GAPUNI, ELSET=GAP_X
158    ,9      ,79
*ELEMENT, TYPE=GAPUNI, ELSET=GAP_X1
160    ,17     ,81
*ELEMENT, TYPE=GAPUNI, ELSET=GAPL_-X
159    ,9      ,80
*ELEMENT, TYPE=GAPUNI, ELSET=GAPL_-X1
161    ,17     ,82
**
 **MASS ELEMENTS****
**
*ELEMENT, TYPE=MASS, ELSET=MASS
49     ,83
**
**
 **SRING ELEMENTS****
**
*ELEMENT, TYPE=SPRING1, ELSET=SP1
8    ,91
*ELEMENT, TYPE=SPRING1, ELSET=SP2
8    ,92
**
 **SPRING MEMBERS
**
*SPRING, ELSET=SP1
1
167.1E+06
*SPRING, ELSET=SP2
6
6.992E+09
**
**
**
**
*ELSET,ELSET=ALL,GENERATE
8,59,1
**
*ELSET,ELSET=PILE_WAVE,GENERATE
8,44,1
**
```

```
*ELSET,ELSET=LOW_HAM,GENERATE
45,49,1
**
*ELSET,ELSET=UP_HAM,GENERATE
50,59,1
**
*ELSET,ELSET=SPRS,GENERATE
91,92,1
**
**********************************************************************
**
**
    **GAP  DEFINITIONS
**
**
**
**********************************************************************
**
**
*GAP, ELSET=GAP_X
1,0,0  ,0.061
*GAP, ELSET=GAP_X1
1,0,0   ,0.0
**
**
*GAP, ELSET=GAPL_-X
1,0,0-   ,0.0
*GAP, ELSET=GAPL_-X1
1,0,0-  ,0.061
**
**********************************************************************
**
**
           **SECTIONS PROPERTIES
**
**
**
**********************************************************************
**
**
**
*BEAM SECTION, ELSET=PILEALL, SECTION=PIPE, MATERIAL=M10, POISSON=0.3
,0.08     ,0.9145
1.0-        ,0.0         ,0.0
**
**
 **no mass here, mass in added mass
**
*BEAM SECTION, ELSET=SOME_HAM, SECTION=PIPE, MATERIAL=DUM, POISSON=0.3
,0.08     ,0.9145
1.0-        ,0.0         ,0.0
**
**
**
**********************************************************************
**
**
    **MATERIAL DEFINITIONS
**
```

```
**************************************************************************
*MATERIAL, NAME=M10
**
*ELASTIC, TYPE=ISO
2.07    E+11,    0.3
*DENSITY
7850.0
**
*MATERIAL, NAME=DUM
**
*ELASTIC, TYPE=ISO
2.07    E+10,    0.3
*DENSITY
785.0
**
* **DAMPING,ALPHA=0.06987,BETA=0.002106
* **DAMPING,ALPHA=0.02404,BETA=0.003415
*DAMPING,ALPHA=0.0433,BETA=0.006819
**
**
*NSET,NSET=SLEEV
158,159,160,161
*NSET,NSET=END
8
*BOUNDARY
49,3
SLEEV,1,6
END,2,3
END,4,5
**
**
**
**************************************************************************
**
**
    **WAVE and CURRENT DEFINITIONS
**
**
**
**************************************************************************
*AQUA
0.0,75.4,9.81,1025
,0.0  , , ,0.24
,75.0 , , ,0.35
 ** Decomposed of Seat-state (Hs ≅ 3.2 Tz =4 ) into 30 components
*WAVE,TYPE=AIRY, WAVE PERIOD
1 ,298.39121 ,15.03181 ,0.00000
1 ,59.85724 ,7.51590 ,0.43719
1 ,141.80616 ,5.01060 ,0.77951
1 ,187.47269 ,3.75795 ,0.49638
1 ,258.52463 ,3.00636 ,0.30573
1 ,204.90823 ,2.50530 ,0.19897
1 ,165.89022 ,2.14740 ,0.13688
1 ,160.31054 ,1.87898 ,0.09856
1 ,31.58806 ,1.67020 ,0.07364
1 ,159.65396 ,1.50318 ,0.05668
1 ,131.86794 ,1.36653 ,0.04471
1 ,108.91217 ,1.25265 ,0.03599
1 ,306.66409 ,1.15629 ,0.02948
```

```
1 ,273.41258 ,1.07370 ,0.02450
1 ,341.91334 ,1.00212 ,0.02062
1 ,200.85786 ,0.93949 ,0.01755
1 ,5.12389 ,0.88422 ,0.01508
1 ,214.62375 ,0.83510 ,0.01308
1 ,293.83406 ,0.79115 ,0.01142
1 ,351.75325 ,0.75159 ,0.01005
1 ,79.88691 ,0.71580 ,0.00890
1 ,253.32612 ,0.68326 ,0.00792
1 ,187.94193 ,0.65356 ,0.00709
1 ,335.84294 ,0.62633 ,0.00637
1 ,256.80760 ,0.60127 ,0.00575
1 ,82.09401 ,0.57815 ,0.00522
1 ,161.87116 ,0.55673 ,0.00475
1 ,61.99189 ,0.53685 ,0.00433
1 ,348.77525 ,0.51834 ,0.00397
1 ,128.05780 ,0.50106 ,0.00365
**
**********************************************************************
**
     **                                                  **
    **MASS      DEFINITIONS
**
**
**
**********************************************************************
**
**MHU 1000 Total weight = 218Te   (weight controlled by g(
**
*MASS, ELSET=MASS
218.0     E+03,
**
**
**
*ELSET,ELSET=SHIM
17
*NSET,NSET=POINTS
17,49,152
*nset,nset=n1
60
**********************************************************************
**
**
    **S T A T I C   A N A L Y S I S   P R O C E D U R E
**
**
**
**********************************************************************
**
**************************************
  **S T E P 1

*AMPLITUDE,NAME=MAG
0.0,0.0,1.0,1.0,1.0,1.0
*STEP , NLGEOM, INC=1000
 INITIAL OFFSET
*STATIC
1.0 ,1.0 ,0.1E-30, 1.0
*DLOAD,AMPLITUDE=MAG,OP=NEW
```

```
**
  **APPLY SELF WIEGHT
**
 PILEALL , GRAV, 9.81, 0.0, -1.0, 0.0
  MASS   , GRAV, 8.53, 0.0, -1.0, 0.0
***
**
  **FLIUD DRAG
**
 PILE_WAVE ,   FDD,  1.0,  1.829,   0.65,  1.0,  MAG, MAG
 LOW_HAM   ,   FDD,  1.0,  2.555,   0.65,  1.0,  MAG, MAG
 UP_HAM    ,   FDD,  1.0,  1.835,   0.65,  1.0,  MAG, MAG
**
  **FLIUD INERTIA
**
 PILE_WAVE ,   FI ,  1.0,  1.829,    1.6,     0.894, MAG
 LOW_HAM   ,   FI ,  1.0,  2.555,    1.6,       0.6, MAG
 UP_HAM    ,   FI ,  1.0,  1.835,    1.6,       0.6, MAG
**
**
  **BOUYANCY
**
 PILE_WAVE , PB, 1.0, 1.829, 1025, 1.729, 75.4
 LOW_HAM   , PB, 1.0, 1.829, 1025, 1.709, 75.4
 UP_HAM    , PB, 1.0, 1.829, 1025, 1.709, 75.4
**
*NODE PRINT,FREQ=0
*EL PRINT,FREQ=0
*OUTPUT, FIELD,   FREQ=0
*OUTPUT, HISTORY, FREQ=0
**********************************************************
**
    **DEFAULT OUTPUT CONTROL
**
******************************************************
**
*MONITOR, NODE=60, DOF=1
**
*NODE PRINT,GLOBAL=YES,FREQ=0,TOTALS=YES
 RF,
**
**
*ELSET, ELSET=MAX
9,17
***
*EL PRINT, ELSET=MAX ,FREQ=1
MISES
** **
**EL FILE,FREQ=0
**SF,
**
*NODE FILE,FREQ=0
 U,
 **The following card saves results of a preselected  number of
**variables (ABAQUS's default) for all steps, since FREQ=1.
**OUTPUT,FREQ=0,FIELD,VARIABLE=PRESELECT
**
*OUTPUT,HISTORY,FREQ=1
```

```
*ELEMENT OUTPUT, ELSET=MAX
MISES
**
*END STEP
**
**********************************************************************
**
   **D Y N A M I C   A N A L Y S I S   P R O C E D U R E
**
**********************************************************************
**
************************************
   **S T E P 2
************************************
*STEP,NLGEOM,INC=150000
  WAVE LOADING
*DYNAMIC
.0.05,300
*END STEP
```

Appendix E : Four extreme value distributions

A = location parameter, B = scale parameter, C = shape parameter

	FT-1	GEV	GPD	Weibull
pdf	$B^{-1}(-\frac{x-A}{B}+e^{-\frac{x-A}{B}})$	$\begin{cases} 0 & x<A+\frac{B}{C}\ \&C<0 \\ (1-C\frac{x-A}{B})^{\frac{1}{C}-1}\times e^{-(1-C\frac{x-A}{B})^{\frac{1}{C}}} & for\begin{cases} x\geq A+\frac{B}{C}\&C<0 \\ x\leq A+\frac{B}{C}\&C>0 \end{cases} \\ 1 & x>A+\frac{B}{C}\ \&C>0 \end{cases}$	$\frac{1}{B}(1+C\frac{x-A}{B})^{-1-\frac{1}{C}}$ valid on $A<x<\infty$ when $C\geq 0$ or $A<x<A-\frac{B}{C}$ when $C<0$	$\frac{C}{B}(\frac{x-A}{B})^{C-1}\times e^{-(\frac{x-A}{B})C}$
CDF	$e^{-e^{-\frac{x-A}{B}}}$	$\begin{cases} 0 & x<A+\frac{B}{C}\ \&\ C<0 \\ e^{-(1-C\times\frac{x-A}{B})^{\frac{1}{C}}} & \\ 1 & x>A+\frac{B}{C}\ \&\ C>0 \end{cases}$	$1-(1+C\times\frac{x-A}{B})^{-\frac{1}{C}}$ valid on $A<x<\infty$ when $C\geq 0$ or $A<x<A-\frac{B}{C}$ when $C<0$	$1-e^{-(\frac{x-A}{B})^{C}}\quad x\geq A$
Mean	$A+\gamma B$ $\gamma=0.05772$		$E(x)=A+\frac{B}{1-C}$	$A+B\Gamma(1+\frac{1}{C})$
Var	$\frac{\pi^2B^2}{6}$			$B^2\Gamma(1+\frac{2}{C})-B^2[\Gamma(1+\frac{1}{C})]^2$

In Weibull distribution:

A = 0 results in 2-parameter Weibull distribution,

A = 0 & C = 2 gives the Rayleigh distribution,

A = 0 & C = 1 corresponds to an exponential distribution.

In GPD:

for A = 0 & C = 0 exponential distribution is obtained.

Appendix F: Proof of Eq. (29) of Chap. 6 i.e.

Pr (X+Y > b | X > u) = Pr[(X | X > u + Y | X > u) >b]

Let X and Y be two continuous random variables having ranges in $(0,\infty)$ and $f_{XY}(x,y)$, $f_X(x)$ *and* $f_Y(y)$ denote their joint and marginal density functions respectively and also b is a positive number, then

$$\Pr(X+Y<b|\ X>u) = \frac{\Pr(X+Y<b, X>u)}{\Pr(X>u)} = \begin{cases} 0 & \text{if } u \geq b \\ \dfrac{\int_{x=u}^{b}\int_{y=0}^{b-x} f_{XY}(x,y)\,dxdy}{\int_{x=u}^{\infty} f_X(x)dx} & \text{if } u < b \end{cases} \quad \text{(B-1)}$$

On the other hand two random variables X|X > u and Y|X > u have the following joint density function:

$$g(x,y) = \begin{cases} 0 & x \leq u \\ \dfrac{f_{XY}(x,y)}{\int_{x=u}^{\infty}\int_{y=0}^{\infty} f_{XY}(x,y)dxdy} & x > u \end{cases} . \quad \text{(B-2)}$$

Hence

$$\Pr[(X \mid X > u + Y \mid X > u) < b] = \begin{cases} 0 & u \geq b \\ \int_{x=u}^{b}\int_{y=0}^{b-x} g(x,y)dxdy & u < b \end{cases} . \quad \text{(B-3)}$$

Equations (B-1) and (B-3) show that Pr (X+Y < b | X > u) and Pr [(X | X > u + Y | X > u)< b] are equal.

Therefore 1- Pr (X+Y < b | X > u) = 1- Pr [(X | X > u + Y | X > u) < b] or

Pr (X+Y > b | X > u) = Pr [(X | X > u + Y | X > u)> b].

Appendix G: List of the papers by the author

Bazaragan, H., et al., 2005.
Prediction of Safe Sea-State for Pile Driving.
Proceedings of 24th International Conference on Offshore Mechanics and Arctic Engineering.(OMAE 2005), June 12-16, 2005, Halkidiki, Greece.

Bazaragan, H., et al., 2006.
Simulation of Mean Zero-upcrossing Period using Artificial Neural Networks trained by Simulated Annealing.
Proceedings of 25th International Conference on Offshore Mechanics and Arctic Engineering(OMAE 2006), June 4-9, 2006, Hamburg, Germany.

Bazaragan, H., et al., 2007.
Neural Network based Simultion of Significant Wave Height.
To be submitted to 26th International Conference on Offshore Mechanics and Arctic Engineering(OMAE 2007), June 10-15, 2007, San Diego, USA.

Bazaragan, H., et al., 2006.
Simulation of Mean Zero-upcrossing period using neural networks trained by simulated annealing.
Paper No: JMST-285 Accepted byJournal of Marine Science & Technology, Japan.

Bazaragan, H., et al., 2006.
Calculating the Return Value using a Mathematical Model of Significant Wave Height.
Paper No: JMST-281 Accepted by Journal of Marine Science & Technology, Japan.

Bazaragan, H., et al., 2006.
Simulation of Significant Wave Height by Neural Networks and its Application to Forecasting and Extreme Wave Height Analysis.
Submitted to a Journal in Norway.