ABSTRACT

Title of Document:	STATISTICAL CHARACTERIZATION AND PREDICTION FOR A STOCHASTIC SEA ENVIRONMENT
	Che-yu Chang, Doctor of Philosophy, 2012
Directed By:	Professor Bilal M. Ayyub Department of Civil and Environmental Engineering

Designing marine and maritime systems requires the probabilistic characterization of sea waves in the time-history and spectral domains. These probabilistic models include parameters that can be empirically estimated based on limited data in durations, locations and applicability to particular designs. Characterizing the statistical uncertainties associated with the parameters and the models is an essential step for risk-based design methods. A framework is provided for characterizing and predicting the stochastic sea-state conditions using sampling and statistical methods in order to associate confidence levels with resulting estimates. Sea-state parameters are analyzed using statistical confidence intervals which give a clear insight for the uncertainties involved in the system. Hypothesis testing and goodness-of-fit are performed to demonstrate the statistical features. Moreover, sample size is required for performing statistical analysis. Sample size indicates the number of representative and independent observations. Current practices do not make a distinction between the number of

discretization points for numerical computations and the number of sampling points, i.e. sample size needed for statistical analysis.   Sample size and interval between samples to obtain independent observations are studied and compared with existing methods.   Further, spatial relationship of the sea-state conditions describes the wave energy transferred through the wave movement.   Locations of interest with unknown sea-state conditions are estimated using spatial interpolations.   Spatial interpolation methods are proposed, discussed, and compared with the reported methods in the literature.   This study will enhance the knowledge of sea-state conditions in a quantitative manner.   The statistical feature of the proposed framework is essential for designing future marine and maritime systems using probabilistic modeling and risk analysis.

STATISTICAL CHARACTERIZATION AND PREDICTION FOR A STOCHASTIC
SEA ENVIRONMENT

by

Che-yu Chang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2012

Advisory Committee:
Professor Bilal M. Ayyub, Chair
Professor M. Sherif Aggour
Professor Amde M. Amde
Professor Charles W. Schwartz
Professor Michael S. Kearney, Department of Environmental Science and Technology,
Dean's Representative

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1. Background and Needs

After 100 years since the tragedy of the maiden voyage of the Titanic, many ships continue to sink and disasters continue to occur at sea. Despite the advances in ship designs, many ships have foundered due to bad weather, negligence, human errors, and the uncertainties of the sea environment, which have led to the loss of lives and economy impacts.

Designing marine and maritime systems requires the probabilistic characterization of sea waves in the time-history and spectral domains. These probabilistic models include parameters that can be empirically estimated based on limited data in durations, locations and applicability to particular designs. Characterizing the statistical uncertainties associated with the parameters and the models is the essential step for risk-based design methods. Cruz and Sarmento (2007) characterized sea-state by linear wave theory approach and using boundary element method. Goff (2009) analyzed the sea surface height noise for improving the altimetry processing algorithms. Auto-covariance analysis was applied to decompose the noise into uncorrelated and correlated components. It was found that the variance of the uncorrelated component is related to significant wave heights. Hamilton (2010) presented a method to characterizing spectral sea wave conditions by clustering the wave spectra.

Efforts are widely put into studies on the wave properties. However, the uncertainties involved in the system and characterization and prediction processes are not presented in the analysis procedure. A probabilistic framework is needed for

characterization and prediction of the sea environment based on statistical methods. Statistical confidence intervals should be studied for system parameters, such as the significant wave height and the modal period, as primary contributors to uncertainties associated with system characterization and prediction. In addition, statistical methods are based on the assumption that the data are independent and representative. Data dependency should be clarified in order to perform statistical analysis. Further, wave data variabilities present themselves in both temporal and spatial perspectives. A study demonstrates the connection of data properties in time and space domains is needed to enhance the knowledge of the applicability on estimation methods.

1.2. Related Research

Studies have been conducted to characterize the sea-state conditions which are necessary for marine and maritime systems design and vessels travel response estimations. Two key parameters to represent the sea-state characteristics are the modal period and the significant wave height. Statistical properties of the distributions and joint distribution of the wave modal period and the significant wave height have been investigated and compared with observations, such as the studies of Longuet-Higgins (1975, 1980), Hatori (1984), Mathisen and Bitner-Gregersen (1990), Sobey (1992), Ferreira and Guedes Soares (2000, 2002, 2003), Rodriguez et al. (1999, 2001, 2002), Goda et al. (2000, 2004), and Hou et al. (2006).

In addition, Forristall et al. (1996) showed that the maximum significant wave height in a storm increases as the length of the samples from which the maximum significant wave height is calculated decreases or the interval between samples decreases. It was indicated that there is an important bias when the maximum significant wave

2

height in a storm is estimated from short samples. Rodriguez and Guedes Soares (2001) investigated the dependency between wave heights and periods and compared the results with the theoretical joint distribution of the wave height and period. It was found that significant correlation between consecutive wave periods only presents in a swell dominated sea state, and the superposition of a swell wave and a wind-sea wave system enhances the correlation between successive wave heights. Wist, Myrhaug, and Rue (2004) studied the statistical properties of successive wave heights and wave periods. From their study, the distribution of the wave height given the previous wave height is independent of the wave height prior to the previous wave height. The distribution of successive wave periods can be estimated by a multivariate Gaussian distribution when the corresponding wave heights are larger than the root-mean-square of the wave heights.

Some predictions of the wave characteristics were made on the theory of wave grouping such as the work by Goda (1976), Kimura (1980), Longuett-Higgins (1984), Sobey and Read (1984), Ochi and Sahinoglou (1989(1) and (2)) and Rodriguez, Guedes Soares, and Ferrer (2000). Some other prediction techniques are summarized by Young and Sobey (1981) such as the Sverdrup-Munk-Bretschneider curve, the Bretschneider's hurricane wave curves, the Wilson's method for space and time varying winds, the Pierson-Moskowitz spectrum, and the Jonswap/Ross tropical cyclone spectrum. Each technique is applicable only on the conditions that the technique was developed.

Besides the studies focused on the wave modal period and the significant wave height, Hamilton, Hui, and Donelan (1979) proposed a nonspectral model to explain the statistical significance of the tail of the correlation function of wind waves and to obtain masking functions for the computation of smoothed wave spectra. Akaike (1981)

developed a computer program and underlying methods to condense observational data and predict the future behavior of locally stationary time series using least squares computations and the concept of Bayesian modeling. Donelan and Pierson (1983) showed that the sampling variability effects are present in spectral estimates computed from wave time histories. They demonstrated that the theory of stationary Gaussian processes provides accurate estimates of the sampling variability. Jensen and Vesecky (1993) indicated that it is inadequate to use the auto-correlation function for characterizing the ocean surface. Kazeminezhad et al. (2005) applied Adaptive-Network-Based Fuzzy Inference System on wave parameters prediction. Guedes Soares and Cherneva (2005) used the spectrogram based on the short-time Fourier transform to study the time frequency evolution of the ocean wind wave properties. Cruz and Sarmento (2007) characterized sea-state by linear wave theory approach and using boundary element method. Hamilton (2010) presented a method to characterize spectral sea wave conditions by clustering the wave spectra.

Properties in the sea environment have temporal and spatial variabilities. The temporal variabilities are represented based on time series analysis. The studies described above focused on the wave characteristics in time and frequency domains. On the spatial perspectives, geographic techniques such as inverse weight factor, semivariogram analysis, and Kriging estimation are often utilized for analyzing the influence of locations of interest in terms of distance. The properties at unobserved locations are estimated based on the properties and influence range of observed locations.

Matheron (1963) provided a procedure in mining reserve simulation studies which is known as the semivariogram method. Based on the semivariogram method, Sen

(1989) proposed an approach called cumulative semivariogram which presents the influence distance in a non-decrease model. The cumulative semivariogram is very similar to the semivariogram model except taking cumulative summations. In addition, a point cumulative semivariogram model was proposed by Sen (1992) to identify the spatial behavior around a reference site. A point cumulative semivariogram is a cumulative semivariogram with a reference site of interest. Sen and Sahin (2001) applied this approach on estimating the solar irradiation value of any point from sites where measurements of solar global irradiation already exist. Altunkaynak (2005) suggested a modified model considering the trigonometric point cumulative semivariogram for predicting significant wave height in a specific region. Altunkaynak and Ozger(2005) provided a standard regional dependence function for significant wave height assessment. This approach is based on the point cumulative semivariogram modified by dividing it by the maximum value and subtracting from unity. Therefore, the standard regional dependence function shows that locations in far distances have lower influence on the point of interest compared with the locations in close distances which have higher influence on the point of interest.

Regardless in time or spatial perspectives, the data, or the samples, used in the estimations are assumed to be representative and independent in order to apply the statistical methodologies. McCuen et al. (1988) addressed the needs to define the spacing between test points required to reach a desired level of testing accuracy on ultrasonic testing of bridge timber piles. Semivariogram analysis and Kriging estimation were taken place to determine the changes in accuracy according to the intervals between testing points. White and Ayyub (1990) also utilized the

semivariogram and Kiging technique for estimating the corrosion rate in steel plating to develop sampling strategy.   It was found that the benefit in taking more samples decreases when the number of sampling reaches certain level which is related to the area of the testing steel plate and the size of influence range.   Ayyub and McCuen (1990) demonstrated the number and location of sampling points for evaluating structural strength on columns and slabs using semivariogram analysis and Kriging estimation. Besides using the classic semivariogram model, Barry and Hoef (1996) proposed a flexible variogram model for spatial prediction using the Kriging concept.   They claimed that the classic variogram models such as linear, spherical, exponential, etc. might not represent the true variogram for the system and therefore lead to estimation errors.   The flexible variogram is in the form of cosine series.   By engaging the moving average concept, the modified flexible variogram has better fit at the origin.

The knowledge of sea-state conditions is essential for designing marine and maritime systems.   Current practices lack some important items required for providing accurate characterization and prediction in the sea environment.   Statistical methodologies are needed to estimate the uncertainties involved in the modeling and prediction procedures.   The independency of the analyzed data needs clarification before performing estimations.   The determination of the sample size of independent observations is required for applying statistical analyses.   Spatial dependency and estimation uncertainties need to be studied.

1.3. Research Purpose and Scope

This research provides a framework for characterizing and predicting the statistical uncertainties of parameters for a stochastic system such as the sea environment.

6

The statistical methodology characterizes the sea-state conditions in the time and spectral domains based on sea surface elevation data. The proposed framework also provides prediction of sea-state conditions and assesses confidence intervals for sea-state conditions at points of interest in the sea environment including points along the travel of cargo ships and naval combatants. This research also provides a method to determine the sample size of independent observations which is essential for performing statistical analyses. The sample size of independent observations determines the required interval between samples for data collection. Moreover, for a dynamic and nonlinear environment such as the ocean, spatial relationship of the sea-state condition is important for predicting the sea-state conditions at locations of interest based on properties at an observed point or points, such as the buoys. This research provides spatial interpolation and dependency study using geographic techniques and data correlation functions.

In this study, a sea environment is characterized in the time and spectral domains. The sea-state is usually characterized by two key parameters: the wave modal period and significant wave height. Therefore, these two parameters are selected for estimations. In the time domain, the wave modal period and significant wave height are estimated from the wave surface elevation time-history. In the spectral domain, the sea wave characteristics are presented in periodograms constructed from the time history using auto-covariance function. Chapter 2 describes the spectral analysis background and the assessment of the confidence intervals of the estimated parameters. The sample size which represents the number of independent observations is studied and discussed in this chapter as well. The interval between samples required for collecting efficient data to achieve desired estimation accuracy can be determined from the estimated sample size.

A comparison for the sample size of a given time-history estimated by the assessment approach provided in this research, the existing spectral analysis process, and the time-history auto-correlation function is provided.   Chapter 2 also provides descriptions of geographic techniques for analyzing wave data in the spatial perspective. Semivariogram analysis and Kriging estimation are introduced and discussed.

Chapter 3 presents the proposed framework and methodology.   This research provides statistical methods to assess confidence intervals for sea-state condition prediction at points of interest in the sea environment including points along the travel track of cargo ships and naval combatants.   The sea-state characteristics for a give buoy location can be obtained from its wave surface elevation time-history.   Two key parameters to describe the sea-state conditions are the wave modal period and the significant wave height described in Chapter 2.   To estimate the sea-state conditions for locations of interest, the wave characteristics of surrounding observed locations are utilized.   Prediction is performed based on the inverse distance weight factors according to the distance between the surrounding observed locations and the locations of interest such as the points of travel track of seagoing vessels.   Locations of interest are estimated as intermediate values among surrounding buoys using inverse distance weight factors applied on the buoys' periodograms.   A periodogram describes the sea-state characteristics in the frequency domain and is constructed from the time-histories. Statistical hypothesis testing is performed to define the confidence intervals of these two selected sea-state parameters for sea-state condition prediction.   Several approaches of defining the modal period are compared and discussed in this chapter using illustrative examples.   Chapter 3 also introduces a freeware called Simulated WAves of Nearshore

(SWAN) which simulates wave properties in a random field. Ris et al. (1997) and Booij et al. (1999, 2001) indicated that the numerical wave model SWAN can provide accuracy at desired level on wave simulations. The generated wave properties are used for verifying the sea-state characterization and interpolation procedures proposed in this chapter.

Chapter 4 provides numerical examples to illustrate the research methodology. Given wave surface elevation time-histories at specific buoy locations, the sea-state conditions described by the wave modal period and significant wave height are determined from these time-histories. Periodograms of these buoys are constructed from their time-histories, and the sea-state predictions for points of interest are performed using these surrounding buoys' periodograms along with the distances between the points of interest and these buoys. Predictions are presented in confidence intervals of the estimated parameters. Verification of the methodology is provided in this chapter as well using wave properties generated by SWAN as well as observations from the NOAA (National Oceanic and Atmospheric Administration) website. Chapter 5 provides the conclusions of this research.

1.4. Notations

C =   auto-covariance function

d =   distance between buoy and point of interest

$E[\ ]$ =   expected value of the term inside the brackets

Err =   error square per location

f =   wave spectrum

H_s = significant wave height

H_{st} = significant wave height from time history

H_{se} = estimated significant wave height of track point

H_{sL} = lower confidence limit of significant wave height

H_{sU} = upper confidence limit of significant wave height

LCL = lower confidence limit

UCL = upper confidence limit

M = truncation point; the number of auto-covariance coefficients considered, and also the number of discretized points for the periodogram

N = number of discretized points

n = sample size of independent observations

Pdg = periodogram

S^2 = variance

T_m = wave modal period

T_{mt} = wave modal period from time history

T_{me} = estimated wave modal period of track point

T_z = up-zerocrossing period

T_L = lower confidence limit of wave modal period

T_U = upper confidence limit of wave modal period

$var[\ ]$ = variance of the term inside the brackets

wb = weight factor

ω = angular frequency

μ = expected value

ρ $\quad = \quad$ auto-correlation function

v $\quad = \quad$ degrees of freedom

λ_k $\quad = \quad$ lag window

$f_p^{\,max}; f_p^{\,Dm}; f_p^{\,Mq}$ $\quad = \quad$ maximum spectral frequency estimated by different approaches

2. Time and Spatial Data Analyses

2.1. Random Process

A random function of a time parameter is called a random process which can be denoted as $x(t)$. $x(t)$ is a random, time-dependent quantity and represents one sample out of infinite possible samples. A collection of such samples is called an "ensemble," denoted as $\{x(t)\}$. Each random function $x^{(k)}(t)$ is a random process having probability density function $p_x^{(k)}(x)$, as shown in Figure 2-1. At a specific time, such as t_1, the density function for the ensemble can be expressed as $p_x(x,t_1)$. Figure 2-1 illustrates the random process and ensemble.
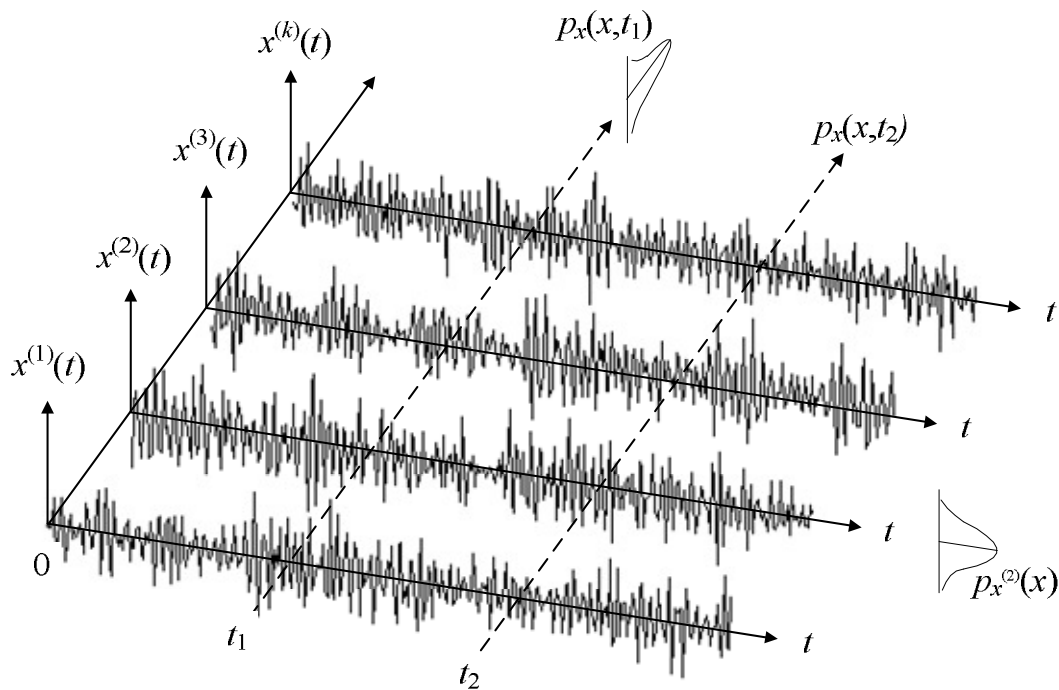
Figure 2-1. Representation of a random process $x(t)$ and the ensemble $\{x(t)\}$, i.e. each $x^{(k)}(t)$ is a sample of the ensemble.

2.2. Stationary and Ergodic

For a random process, if the statistical characteristics do not change with time, the process is called a stationary random process. In other words, if a process is stationary, the probability density functions at different times, say t_1, t_2 and so on, would all be the same. It has been found experimentally that the sea surface elevation is a stationary random process for short term observations, i.e. up to a few hours (Hughes, 1988), even though the random process $x(t)$ is a function of time.

Since a random process is a function of time, there are two ways to calculate the statistical characteristics. They can be calculated over all of the samples of the ensemble at a specific time, say t_1, which is referred to as ensemble averages, or they can be calculated over all time from -∞ to ∞ for a particular sample, say $x^{(1)}(t)$, which is referred to as temporal averages. In general, these ensemble and temporal averages would be different; however, for many random processes including ocean waves, the temporal averages computed from a single sample are equal to the ensemble averages. This type of processes is called ergodic process. An ergodic process means that a single sample $x(t)$ is typical enough to represent the entire process. This condition implies that an ergodic process must be stationary; while a stationary process might not be ergodic.

2.3. Auto-covariance Function

The auto-covariance function is the means to measure or represent the degree of association between values of the random variable $x(t)$ at times differing by a specific interval τ. For a stationary ergodic process, the expected value μ_x is a constant for all

times t, and the auto-covariance function is the expect value of the product of any two

values of $x(t)$, e.g. $x_1 = x(t_1)$ and $x_2 = x(t_2) = x(t_1 + \tau)$, expressed as follows:

$$C(\tau) = E[(x(t) - \mu)(x(t + \tau) - \mu)] = E[x_1 x_2] - \mu_x^2 \qquad (2\text{-}1)$$

The auto-covariance function at the origin, expressed in Equation 2-2, is the variance of

the process.   For a zero mean process, i.e. $\mu_x = 0$, the auto-covariance function at the

origin is the mean square.

$$
\begin{aligned}
C(0) &= var[x(t)], \quad \mu_x \neq 0 \\
C(0) &= E[x^2(t)] \quad, \quad \mu_x = 0
\end{aligned}
\qquad (2\text{-}2)
$$

In addition, a stationary process satisfies the condition expressed as follows:

$$E[(x(t_1) - \mu)(x(t_1 + \tau) - \mu)] = E[(x(t_2) - \mu)(x(t_2 + \tau) - \mu)] \qquad (2\text{-}3)$$

According to Equation 2-3, the auto-covariance function $C(\tau)$ is independent of the

starting point t and only depends on the interval τ.   The commutative property of $x_1 {*} x_2$

also leads to the relationship as follows:

$$C(\tau) = E[(x(t_1) - \mu)(x(t_1 + \tau) - \mu)] = E[(x(t_1) - \mu)(x(t_2) - \mu)]$$

$$= E[(x(t_2) - \mu)(x(t_1) - \mu)] = E[(x(t_2) - \mu)(x(t_2 - \tau) - \mu)] = C(-\tau) \quad (2\text{-}4)$$

Equation 2-4 indicates that the auto-covariance function is an even function of τ.

There is a function closely related to the auto-covariance function, which is call

the auto-correlation function.   When $C(0) > 0$, the correlation between two points

separated by τ is defined as

$$\rho(\tau) = C(\tau)/C(0) \qquad (2\text{-}5)$$

Equation 2-5 is called the auto-correlation function.   According to the definition in

Equation 2-5, the auto-correlation function has the value equals to one at the origin

expressed as follows:

$$\rho(\tau = 0) = \frac{C(\tau=0)}{C(0)} = 1 \qquad\qquad (2\text{-}6)$$

Two extreme cases of the auto-covariance functions are: (1) the values of a function $x(t)$ at different times are completely unrelated, and (2) each sample $x(t)$ is identical thus leads to perfect correlation. The former case which has the completely unrelated relationship between the values of a function at different times would have the auto-covariance value at the origin as the series variance while the auto-covariance values at all other times are zero. The latter case which has the perfect correlation would have the auto-covariance function as a constant which equals to the series variance value. Usually the relationship would be in between the two extreme conditions. That is, for small τ, the value $x(t+\tau)$ can be in a range of values that do not significantly differ from $x(t)$, and for large τ, the degree of association between $x(t)$ and $x(t+\tau)$ is very low.

In the frequency domain, if a process is made up of components of many different frequencies, the spectrum is quite wide due to the wide range of frequencies. The periodicity of the process is very little so that the auto-covariance function would have the shape such as Figure 2-2 that the values at times expect the origin are practically zeros and the value at the origin is the series variance. On the other hand, if the frequencies of a process are within a narrow range which is small compared with the magnitude of the center frequency of the range, the spectrum would have a single narrow peak at the center frequency of the range, denoted as ω_0. The periodicity would present regular peaks in the auto-covariance function such as Figure 2-3.

The auto-covariance function plays an important role in the subsequent sections. The auto-covariance function connects the random process from the time domain to the spectral domain. In addition, the dependency between data determined by the

15

auto-covariance function leads to the estimation of the sample size of independent

observations which provides a guideline of sampling.   The sampling guideline is

denoted as the interval between samples which determines the interval between each

independent sample.   Table 2-1 summarizes the temporal and ensemble averages such

as the mean and auto-covariance function for stationary and ergodic, stationary and

non-ergodic, and non-stationary random process conditions.
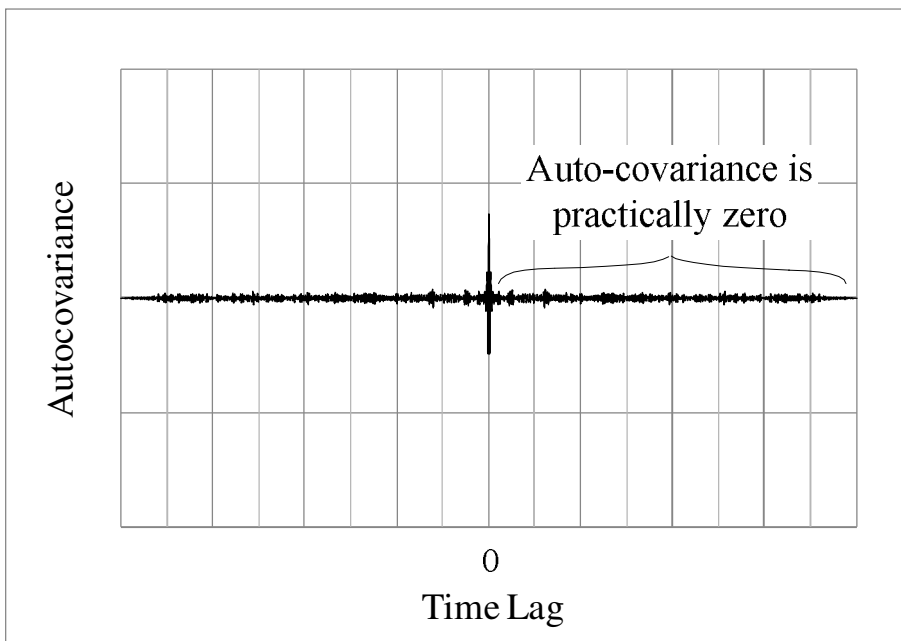
Figure 2-2. Illustration of the auto-covariance function for a wide-band process.
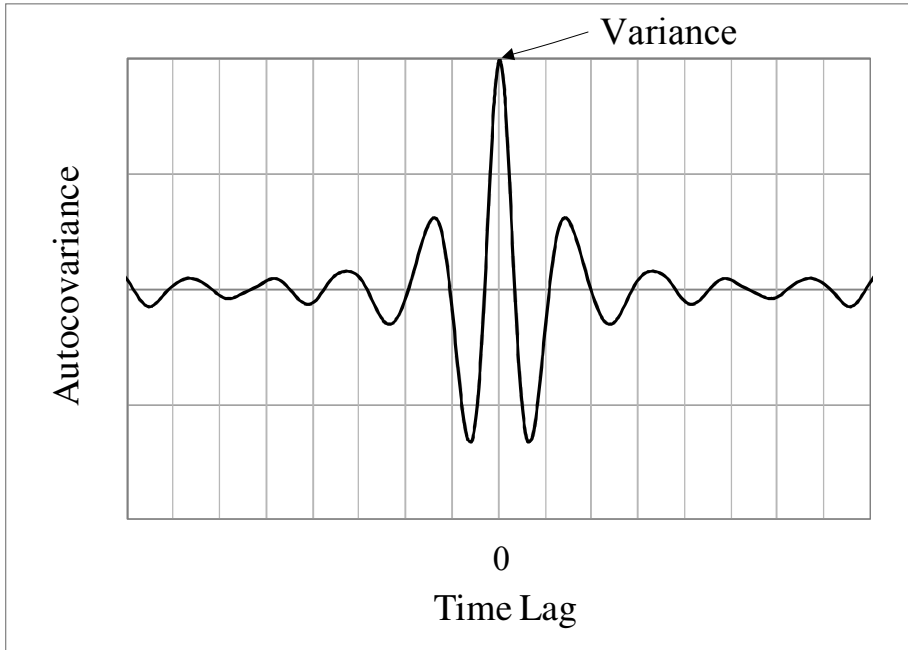
Figure 2-3. Illustration of the auto-covariance function for a narrow-band process.

Table 2-1. Summary of the mean and auto-covariance function for random process $\{x(t)\}$.

		Temporal averages	Ensemble averages
Random data classification	Stationary, ergodic	**Mean:** $$\mu_{x_k} = \lim_{T\to\infty}\frac{1}{T}\int_0^T x_k(t)dt = \mu_x$$ **Auto-covariance function:** $$C_{x_kx_k}(t_i,t_i+\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T [x_k(t)-\mu_{x_k}][x_k(t+\tau)-\mu_{x_k}]dt$$ $$= \lim_{T\to\infty}\frac{1}{T}\int_0^T [x_k(t)-\mu_x][x_k(t+\tau)-\mu_x]dt$$ $$= C_{xx}(\tau)$$	**Mean:** $$\mu_x(t_i) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N x_k(t_i) = \mu_x$$ **Auto-covariance function:** $$C_{xx}(t_i,t_i+\tau) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N [x_k(t_i)-\mu_x(t_i)][x_k(t_i+\tau)-\mu_x(t_i+\tau)]$$ $$= \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N [x_k(t_i)-\mu_x][x_k(t_i+\tau)-\mu_x] = C_{xx}(\tau)$$
	Stationary, non-ergodic	**Mean:** $$\mu_{x_k} = \lim_{T\to\infty}\frac{1}{T}\int_0^T x_k(t)dt$$ **Auto-covariance function:** $$C_{x_kx_k}(t_i,t_i+\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T [x_k(t)-\mu_{x_k}][x_k(t+\tau)-\mu_{x_k}]dt$$ $$= C_{x_kx_k}(\tau)$$	**Mean:** $$\mu_x(t_i) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N x_k(t_i) = \mu_x$$ **Auto-covariance function:** $$C_{xx}(t_i,t_i+\tau) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N [x_k(t_i)-\mu_x(t_i)][x_k(t_i+\tau)-\mu_x(t_i+\tau)]$$ $$= \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N [x_k(t_i)-\mu_x][x_k(t_i+\tau)-\mu_x] = C_{xx}(\tau)$$
	Non-stationary (implies non-ergodic)	**Mean:** $$\mu_{x_k} = \lim_{T\to\infty}\frac{1}{T}\int_0^T x_k(t)dt$$ **Auto-covariance function:** $$C_{x_kx_k}(t_i,t_i+\tau) = \lim_{T\to\infty}\frac{1}{T}\int_0^T [x_k(t)-\mu_{x_k}][x_k(t+\tau)-\mu_{x_k}]dt$$ $$= C_{x_kx_k}(\tau)$$	**Mean:** $$\mu_x(t_i) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N x_k(t_i)$$ **Auto-covariance function:** $$C_{xx}(t_i,t_i+\tau) = \lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^N [x_k(t_i)-\mu_x(t_i)][x_k(t_i+\tau)-\mu_x(t_i+\tau)]$$ $$= C_{xx}(t_i,\tau)$$

2.4. Time Series Analysis

2.4.1. Spectral Analysis

The spectral analysis is a modification of Fourier analysis making it suitable for stochastic rather than deterministic functions of time. It is assumed that the data are a time series made with N observations at equal time intervals. The number of observations, N, is assumed to be an even number, although this assumption is not a necessary condition. The N observations are denoted by $(x_1, x_2, ..., x_i, ..., x_N)$ or by $\{x_i\}$, where $x_i = x(t_i)$. By applying spectral analysis on the time series, the characteristics of a time series can be expressed in the frequency domain. It should be noted that the random process discussed here is assumed to be a stationary ergodic process unless specifically indicated.

As a basis of the spectral analysis, Section 2.4.2 introduces the simple sinusoidal model as well as the lowest and the highest frequencies for the spectral function to present a discrete random process. Further, the periodogram which is an application of the simple sinusoidal model is presented in Section 2.4.3. Periodogram analysis shows how the variance of a time series distributes over frequencies. There are some existing lag windows or spectral windows used to smooth the periodogram and eliminate the spurious peaks in the periodogram. The lag windows are applied on auto-covariance functions while the spectral windows are applied on the spectral functions. A few commonly used lag windows are introduced in the subsequent section. Smoothed, or modified, periodograms are also discussed.

In addition, the confidence intervals for the ratio of the periodogram and its original spectrum are described in Section 2.4.4.   To estimate the confidence intervals, the number of degrees of freedom, or the sample size, is needed.   Current practices to compute the degrees of freedom use the number of observations and the lag window applied on the auto-covariance function.   From statistical perspectives, the degrees of freedom, or the sample size, should be the same for the same time series of the same duration regardless how they are discretized.   For a time series having independent observations, the sample size is the number of observations; however, if the observations are correlated, it is necessary to obtain the number of independent observations before performing any statistical analysis or applying statistical methods.   Hence, the number of independent data points, or the sample size of independent observations, is studied and compared with the current practices in Section 2.4.5.

2.4.2. Simple Sinusoidal Model

The simple sinusoidal model describes the time series as a deterministic sinusoidal component at frequency ω with a random error term ε.   The following form can be used to present the simple sinusoidal model:

$$x_t = a + b * cos(\omega t) + c * sin(\omega t) + \varepsilon_t \tag{2-7}$$

in which ε_t is the white noise, and a, b and c are the parameters to be estimated from data using the least square concept.

For a discrete process measured at unit intervals without loss of generality, the spectral function argument can be restricted to the frequency range $(0, \pi)$, where the upper limit π is so-called the Nyquist frequency.   The Nyquist frequency presented as $\omega = \pi$ is the highest frequency that could be fitted into data.   Assuming measurements

20

taken at a unit time interval Δt, the minimum number of observations to complete one cycle is two to obtain meaningful information from a set of data, as illustrated in Figure 2-4. Therefore, the highest frequency (Nyquist frequency) can be found by equating the number of observations, 2, to the cycle length 2π divided by the frequency ω as $2 = 2\pi / \omega$ which leads to $\omega = \pi$. On the other hand, the lowest frequency, or the fundamental frequency, is the one that completes one cycle in the entire duration of the observations. The lowest frequency as shown in Figure 2-5 can be found by equating the number of observations, N, to the cycle length 2π divided by the frequency ω as $N = 2\pi / \omega$. Hence, the lowest frequency can be expressed as $\omega = 2\pi / N$.

Since the lowest frequency, or fundamental frequency, depends on the number of measurements, N, the lower the frequency we are interested in, the longer the time period over which we need to take observations for a unit time interval Δt. However, the higher the frequency we are interested in, the larger the number of measurements we should take over certain duration. In other words, the more frequently measurements should be taken.
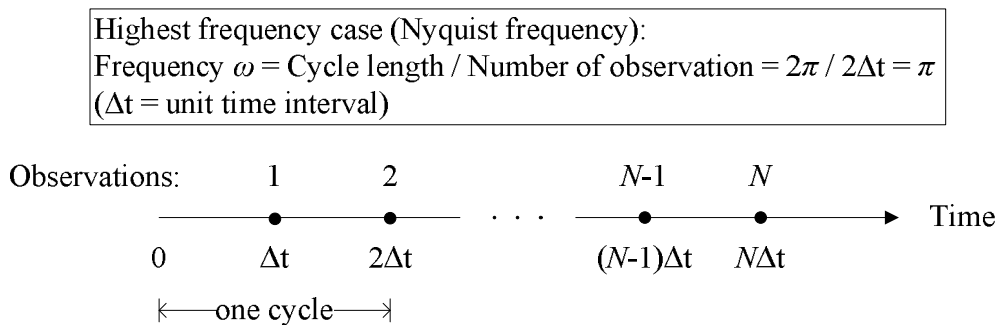
Figure 2-4. Illustration for the highest frequency, i.e. Nyquist frequency.

Lowest frequency case (fundamental frequency):
Frequency ω = Cycle length / Number of observation = $2\pi / N\Delta t$ = $2\pi / N$
(Δt = unit time interval)

Observations: 1   2   N-1   N

0   Δt $2\Delta t$ $(N-1)\Delta t$ $N\Delta t$

⊢——————————one cycle——————————⊣
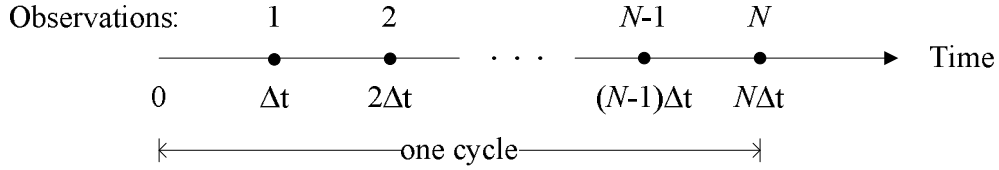
Figure 2-5. Illustration for the lowest frequency, i.e. fundamental frequency.

2.4.3. Periodogram Analysis

The characteristics of a time series can be presented in the frequency domain by applying spectral analysis on the time series as described by Chatfield (2004).   A periodogram shows how the variance of a time series is distributed over frequencies.   A discretized time series $\{x_i\}$ can be expressed using a finite Fourier series representation as follows:

$$x_t = a_0 + \sum_{p=1}^{\frac{N}{2}-1}\left(a_p \cos\left(\frac{2\pi p t}{N}\right) + b_p \sin\left(\frac{2\pi p t}{N}\right)\right) + a_{N/2}\cos(\pi t), t = 1, 2, \dots, N \quad (2\text{-}8)$$

in which N is the number of discretized points of the time series, and the coefficients are given by

$$a_0 = \frac{1}{N}\sum_{t=1}^{N} x_t$$

$$a_{N/2} = \frac{1}{N}\sum_{t=1}^{N}(-1)^t x_t$$

$$a_p = \frac{2}{N}\sum_{t=1}^{N} x_t \cos\left(\frac{2\pi p t}{N}\right) \;,\; p = 1, 2, \dots, (N/2 - 1)$$

$$b_p = \frac{2}{N}\sum_{t=1}^{N} x_t \sin\left(\frac{2\pi p t}{N}\right) \;,\; p = 1, 2, \dots, (N/2 - 1)$$

If the series $\{x_i\}$ is taken from a discrete pure random process where the observations are independent and normally distributed variables expressed as $N(\mu, \sigma^2)$, the coefficients a_p and b_p are independent and normally distributed, and each one has a zero mean and $2\sigma^2/N$ variance. The periodogram, denoted as Pdg, can be calculated from the time series data as follows:

$$Pdg(\omega_p) = \frac{\left(\sum x_t \cos\left(\frac{2\pi pt}{N}\right)\right)^2 + \left(\sum x_t \sin\left(\frac{2\pi pt}{N}\right)\right)^2}{N\pi} \quad , \quad p \leq N/2 \qquad (2\text{-}9)$$

Moreover, the periodogram is the Fourier transform of the auto-covariance function expressed as

$$Pdg(\omega_p) = \frac{1}{\pi}\left(C_0 + 2\sum_{k=1}^{N-1} C_k \cos(\omega_p k)\right) \qquad (2\text{-}10)$$

in which $\{C_k\}$ is the auto-covariance coefficient at time lag k defined as

$$C_k = \frac{1}{N}\sum_{t=1}^{N-k}(x_t - \bar{x})(x_{t+k} - \bar{x}) \qquad (2\text{-}11)$$

By the fact that the periodogram follows a chi-square distribution with two degrees of freedom and the variance of a two degrees of freedom chi-square distribution is four, the variance of the periodogram is a constant independent of the sample size. Hence, the periodogram requires modification to enhance estimation and prediction.

One approach to modify the periodogram is to apply a lag window on a truncated auto-covariance function as follows:

$$Pdg(\omega_p) = \frac{1}{\pi}\left(\lambda_0 C_0 + 2\sum_{k=1}^{M} \lambda_k C_k \cos(\omega_p k)\right) \qquad (2\text{-}12)$$

in which $\{C_k\}$ is the auto-covariance coefficient at time lag k defined in Equation 2-11, $\{\lambda_k\}$ is a set of weights called the lag window, and $M(<N)$ is the truncation point. The modified periodogram in the form of Equation 2-12 is a smoothed periodogram constructed by applying a lag window $\{\lambda_k\}$ on the raw periodogram expressed as

Equation 2-10.   There are some lag windows available such as the Parzen window, the

Tukey window and the Hamming window.   The Parzen and Tukey windows are the two

best-known lag windows.   The Parzen window is used in this study because of its

non-negative nature, and defined as follows:

$$\lambda_k = \begin{cases} 1 - 6(k/M)^2 + 6(k/M)^3 & 0 \le k \le M/2 \\ 2(1 - k/M)^3 & M/2 \le k \le M \\ 0 & M < k \end{cases} \qquad (2\text{-}13)$$

Because the precision of $\{C_k\}$ decreases as k increases since the coefficient is

based on fewer terms, it is reasonable to apply less weight to the values of $\{C_k\}$ as k

increases.   It should be noted that the auto-covariance values in the range of $M < k < N$

are no longer used.   The choice of the truncation point M can be subjective or based on

common practices.   A smaller value of M would result in smaller variance of the

periodogram, but some features of the spectrum might be smoothed out if the value of M

is too small.   However, if the value of M is too large, the periodogram would have too

many peaks which might be spurious.   A compromise value is chosen in this study as

(Chatfield, 2004):

$$M = 2\sqrt{N} \qquad (2\text{-}14)$$

2.4.4. Confidence Intervals

The periodogram as Equation 2-12 can be written as the following form

$$Pdg(\omega_p) = \frac{1}{\pi}\left(\sum_{k=-M}^{M} \lambda_k c_k \cos(\omega_p k)\right) \qquad (2\text{-}15)$$

Jenkins and Watts (1968) showed that the quantity

$$\frac{vPdg(\omega)}{f(\omega)} \qquad (2\text{-}16)$$

is approximately chi-square distributed with v degrees of freedom given by

$$\nu = \frac{2N}{\sum_{k=-M}^{M} \lambda_k^2} = \frac{2N}{\lambda_0^2 + 2\sum_{k=1}^{M} \lambda_k^2} \tag{2-17}$$

in which Pdg is the periodogram, f is its original spectrum, N is the number of discretized

points, M is the truncation point defined in Equation 2-14, and $\{\lambda_k\}$ is the lag window

such as the Parzen window defined in Equation 2-13.   The $100(1-\alpha)\%$ confidence

interval of the quantity defined in Equation 2-16 can be expressed as

$$\Pr\left(\chi_{\nu,1-\frac{\alpha}{2}}^2 < \frac{\nu Pdg(\omega)}{f(\omega)} < \chi_{\nu,\frac{\alpha}{2}}^2\right) = 1 - \alpha \tag{2-18}$$

The $100(1-\alpha)\%$ confidence limits, lower confidence limit (LCL) and upper confidence

limit (UCL), for $f(\omega)$ at different frequencies ω are then given by

$$LCL = \frac{\nu Pdg(\omega)}{\chi_{\nu,\frac{\alpha}{2}}^2} \tag{2-19}$$

$$UCL = \frac{\nu Pdg(\omega)}{\chi_{\nu,1-\frac{\alpha}{2}}^2} \tag{2-20}$$

The confidence interval defined by Equations 2-18, 2-19, and 2-20 represents the

point-estimation for $f(\omega)$ at frequency ω.

2.4.5. Sample Size of Independent Observations

For a time series that has independent observations, the number of discretized

points, N, can be treated as the sample size.   However, when the time series

observations are correlated, it is necessary to obtain the sample size which represents the

number of independent observations, denoted as n.   Equation 2-21 described in Wei

(2005) shows an estimator of the series variance.   Equation 2-22 demonstrates the

estimation from the definition of the variance of series variance for the sample size n

which represents the number of independent observations.

$$var[s^2] = \frac{2}{N}\left\{C_0^2 + 2\sum_{k=1}^{N-1}\left(1 - \frac{|k|}{N}\right)C_k^2\right\} \qquad (2\text{-}21)$$

$$var[s^2] = \frac{2S^4}{n} = \frac{2}{N}\left\{C_0^2 + 2\sum_{k=1}^{N-1}\left(1 - \frac{|k|}{N}\right)C_k^2\right\}$$

$$n = \frac{S^4}{\frac{1}{N}\left\{C_0^2 + 2\sum_{k=1}^{N-1}\left(1 - \frac{|k|}{N}\right)C_k^2\right\}} \qquad (2\text{-}22)$$

in which C is the auto-covariance function, N is the number of discretized points of the time series, and S^2 is the variance for the time series.

For the same time series of the same duration, the sample size or the degrees of freedom should be the same or vary within a limited range and should not be affected by how the time series is discretized. For instance, if a buoy records the sea surface elevation for an hour, the results presented in every minute have 60 data points; while results presented in every 30 seconds have 120 data points. These two sets of data are from the same buoy and cover the same period of time. In other words, these two sets of records represent the same event. The sample size which is the number of independent observations should be the same for these two sets of data in spite of the number of their discretization points.

Since the confidence intervals are evaluated based on the sample size which indicates the independent number of observations, the degrees of freedom v expressed by Equation 2-17 needs to be modified and replaced by n obtained from Equation 2-22. An illustrative example is shown below to demonstrate the necessity of modifying the degrees of freedom computation shown in Equation 2-17.

Table 2-2 shows the modal periods and significant wave heights used to simulate the time histories for three buoys. The simulated time histories have the same duration of 1500 seconds. Each of these three time histories is discretized by various time increments summarized in Table 2-3. The discretization time intervals vary from 0.25

26

sec to 1.5 sec, which lead to the number of discretized points between 6000 and 1000 points.   The degrees of freedom, or sample size, computed by Equation 2-17 is denoted as the "Lag window method" since it is based on the lag window and the time history discretization points N; while the sample size estimated by Equation 2-22 is denoted as the "Proposed method."   The sample sizes estimated by Equation 2-22, the proposed method, are different for three different buoys since Equation 2-22 is based on the time series variance, the number of discretization points N and the time series auto-covariance function.   Although the numbers of discretization points are the same for the three buoys, the time series variances and the auto-covariance functions are different. Therefore the sample sizes estimated by Equation 2-22, the proposed method, are different for the three buoys.   The sample size, or the degrees of freedom, computed by Equation 2-17, the lag window method, is the same for three buoys since it depends on the number of discretization points and the lag window which are the same for the three buoys.   Figure 2-6 shows that the sample size, or the degrees of freedom, obtained from Equation 2-17, the lag window method, increases when the number of discretization points increases.   On the other hand, the sample size obtained from Equation 2-22, the proposed method, varies within a limited rang and does not increase with an increasing number of discretization points.

Table 2-2. Locations, modal periods and significant wave heights of three buoys used to simulate the time histories.

Buoy	x-coordinate	y-coordinate	Modal period	Significant wave height
1	0 m	800 m	7 sec	2.0 m
2	350 m	500 m	6 sec	1.0 m
3	150 m	0 m	8 sec	1.5 m
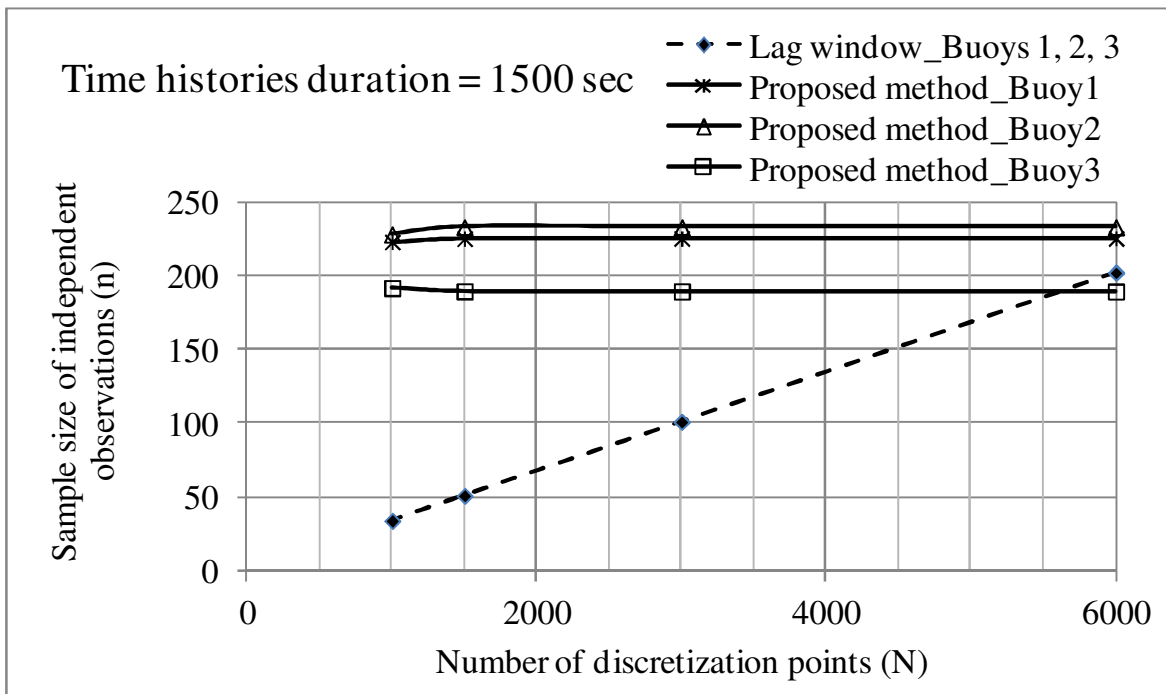
Figure 2-6. Sample size (or degrees of freedom) estimations using lag window and proposed methods.

Table 2-3. Sample size estimations using the lag window method, the proposed method and the auto-correlation function.

Time history duration = 1500 sec		Sample size of independent observations (n)						
Time increment	Number of discretization points (N)	Lag window method	Proposed method			Auto-correlation function		
		Buoys 1, 2 and 3	Buoy 1	Buoy 2	Buoy 3	Buoy 1	Buoy 2	Buoy 3
0.25 sec	6000	202	225	232	189	300	357	263
0.5 sec	3000	101	225	232	189	300	357	263
1.0 sec	1500	51	225	232	189	300	357	259
1.5 sec	1000	34	222	228	191	288	349	250

Table 2-3 also provides the sample size estimations using the auto-correlation function.   As demonstrated in Section 2.3, the auto-covariance function represents the degree of association between data defined in Equation 2-1.   The auto-correlation defined in Equation 2-5 is the auto-covariance function normalized by the series variance. When the auto-correlation function approaches zero, the data estimated are not correlated with each other.   That is, the observations become independent at which the auto-correlation function approaches zero.   Figure 2-7, Figure 2-8 and Figure 2-9 show the first 10 seconds time lags of the auto-correlation functions for the buoys 1, 2 and 3, respectively, defined in Table 2-2.   The case of time increment of 0.25 sec is shown, which is the case of 6000 discretization points.   The sample sizes summarized in Table 2-3 are evaluated by indicating the time lag of the second zero-crossing point of the auto-correlation function.   For example, if the second zero-crossing of the auto-correlation function is at 5 seconds, the sample size is estimated by dividing the total duration 1500 seconds by 5 seconds, which leads to the a sample size of 300.   Table 2-3 shows the comparison of sample size estimations using the lag window method defined in Equation 2-17, the proposed method defined in Equation 2-22 and the auto-correlation function.   The comparison includes cases of various discretization points of 1000, 1500, 3000 and 6000 which are determined by dividing the total duration 1500 sec by the time increments of 1.5, 1.0, 0.5 and 0.25 sec, respectively.   The results evaluated using the auto-correlation second zero-crossing locations do not depend on the number of the time series discretization points; however, there are differences between the estimations using the proposed method based on Equation 2-22 and the estimations using the auto-correlation function.

30
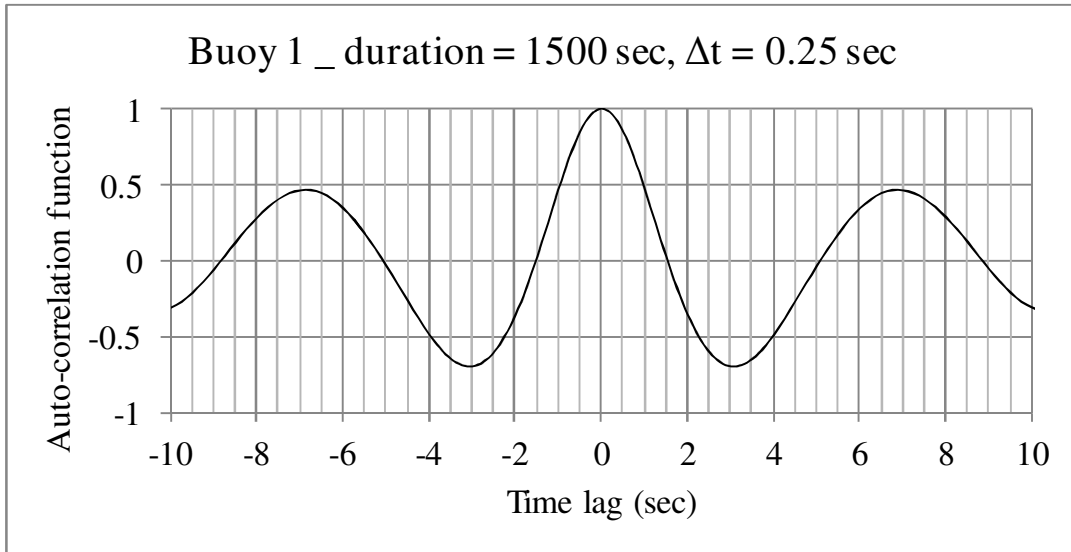
Figure 2-7. Auto-correlation function for buoy 1 at the first 10 seconds time lag.
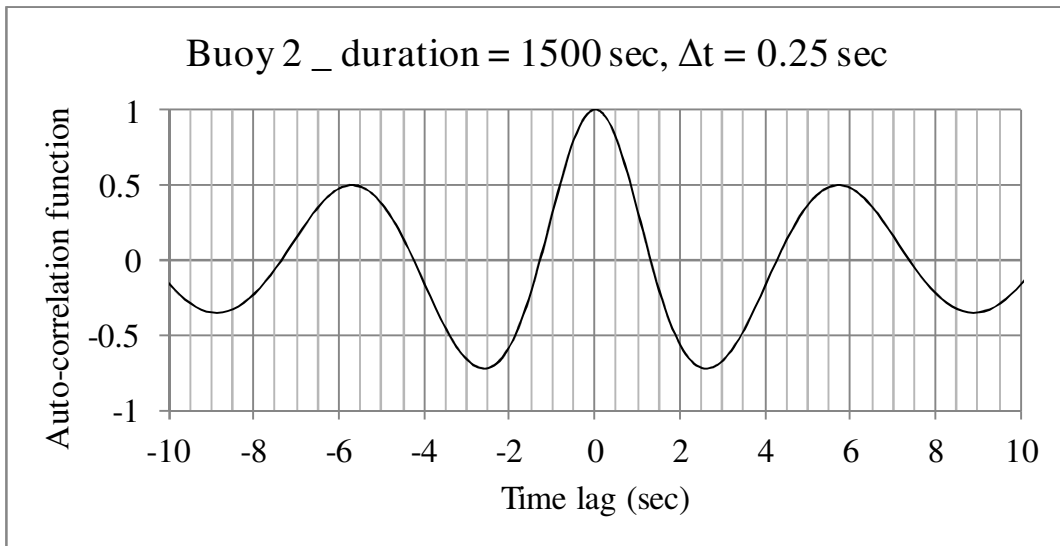
Figure 2-8. Auto-correlation function for buoy 2 at the first 10 seconds time lag.
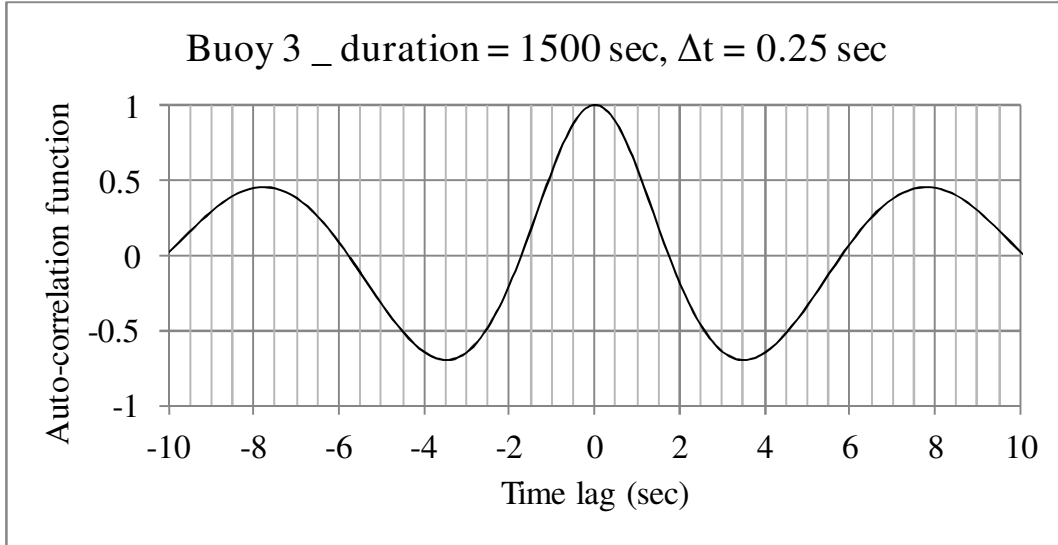
Figure 2-9. Auto-correlation function for buoy 3 at the first 10 seconds time lag.

Before getting into the discussion of the different sample size estimations results obtained by the proposed method and the auto-correlation function, it should be noted that there is the other method similar to Equation 2-22 to estimate the sample size of independent observations.   While Equation 2-22 estimates the sample size of independent observations from the variance of the time series variance, the properties of the variance of the auto-correlation function can also be utilized for the sample size estimation.

An auto-correlation function $\{\rho_k\}$ is defined in Equation 2-5.   For a random process with independent observations, the auto-correlation function value is one at the origin and zero for all non-zero values of lag k expressed as follows:

$$\rho_k = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

(2-23)

With a large number of observations, N, the variance of the time series auto-correlation function for a random process with independent observations is approximately distributed

with a zero mean and variance $1/N$, i.e. $N(0, 1/N)$.   If the observations are correlated, the following expression can be used to determine the sample size n of independent observations:

$$var[\rho] = \frac{1}{N}\left\{\rho_0{}^2 + 2\sum_{k=1}^{N-1}\left(1 - \frac{|k|}{N}\right)\rho_k{}^2\right\} \cong \frac{1}{n} \tag{2-24}$$

in which ρ_k is the auto-correlation function at lag k and ρ_0 is that at the origin.   Based on Equation 2-24, the sample size of independent observations, n, can be estimated as follows:

$$n = \frac{N}{\left\{\rho_0{}^2 + 2\sum_{k=1}^{N-1}\left(1 - \frac{|k|}{N}\right)\rho_k{}^2\right\}} \tag{2-25}$$

It should be noted that Equations 2-22 and 2-25 lead to the same estimation results for the sample size of independent observations n.   Therefore, the "Proposed method" can be referred to either Equation 2-22 or Equation 2-25.Background Problems

Estimations for the sample size of independent observations provided a guideline of the number of observations required to obtain independent and representative data. However, it would be easier to understand if the sampling recommendation is given in a way such as collecting one sample every second.   Equation 2-26 provides the required interval between samples to obtain independent data.   The interval between samples is determined by dividing the number of discretization points by the sample size of independent observations and multiplying by the discretization increment as the following form:

$$Interval\ between\ samples = \frac{N}{n}\Delta t \tag{2-26}$$

The sample size of independent observations can be estimated by either Equation 2-22 or

Equation 2-25.

Since the sample size of independent observations differs for different buoys as

shown in Table 2-3, it would be useful to present the interval between samples in terms

of the modal period T_m of the buoy itself.   Hence, Equation 2-26 can be re-written as

follows:

$$Interval\ between\ samples = \left(\frac{\frac{N}{n}\Delta t}{T_m} \right) * T_m \qquad (2\text{-}27)$$

Equation 2-27 determines the interval between samples based on the modal period T_m of

the series analyzed, which provides better demonstration on how to sample in order to

obtain independent observations.   The intervals between samples for the three buoys

defined in Table 2-2 are estimated and the results are summarized in Table 2-4.

According to the sample size shown in Table 2-3 estimated using the lag window method

defined in Equation 2-17, the proposed method defined in Equations 2-22 and 2-25 and

the auto-correlation second zero-crossing points, the interval between samples can be

computed by Equation 2-26.   For example, the estimated sample size of independent

observations n is 225 for buoy 1 with time series increment 0.25 sec and the number of

discretization points 6000, therefore, the interval between sample is calculated as ($N =$

6000) * ($\Delta t = 0.25$sec) / ($n = 225$) = 6.7sec.   Table 2-4 also provides the intervals

between samples as ratios of the modal period defined in Equation 2-27.   Considering

buoy 1 for example, the interval between samples calculated as 6.7 sec can be presented

by 6.7sec / ($T_m = 7.0$sec) * $T_m = 0.95T_m$.   The intervals between samples in the form of

Equation 2-27 are shown in parentheses. The results show that the intervals between samples are approximately 1.0 T_m.

As mentioned previously, the sample size of independent observations n estimated by the proposed method, Equations 2-22 and 2-25, and by the auto-correlation function second zero-crossing points show different results. For a random process, the auto-correlation function would stay within its standard deviation range during the beginning time lags. The auto-correlation functions for buoy 1 of various intervals between samples are shown in Figure 2-10 through Figure 2-17. The auto-correlation standard deviation is shown in dashed lines and is denoted as STD(corr). Based on Table 2-2, the modal period for buoy 1 is 7 sec. The auto-correlation function fluctuates exceeding its standard deviation for the case of the interval between samples as 6 sec shown in Figure 2-10. The cases of the intervals between samples as 6.5 sec, 7 sec and 7.5 sec show that the auto-correlation functions stay within the range of its standard deviation. As the interval between samples increases, such as the 8 sec, 8.5 sec and 9 sec cases, the auto-correlation functions fluctuate exceeding its standard deviation again.

According to Table 2-4, the interval between samples is approximately the modal period for obtaining independent samples. The auto-correlation functions shown from Figure 2-10 through Figure 2-17 verify the estimations based on the interval between samples defined in Equations 2-26 and 2-27. In other words, the proposed method for estimating the sample size of independent observations determined in Equations 2-22 and 2-25 is verified.

35

Table 2-4. Estimations for interval between samples using the lag window method, the proposed method, and the auto-correlation function.

Time history duration = 1500 sec		Interval between samples (sec)						
Time increment	Number of discretization points (N)	Lag window method	Proposed method			Auto-correlation function		
		Buoys 1, 2 and 3	Buoy 1	Buoy 2	Buoy 3	Buoy 1	Buoy 2	Buoy 3
0.25 sec	6000	7	6.7 (0.95 T_m)	6.4 (1.07 T_m)	7.9 (0.99 T_m)	5.0	4.2	5.7
0.5 sec	3000	15	6.7 (0.95 T_m)	6.4 (1.07 T_m)	7.9 (0.99 T_m)	5.0	4.2	5.7
1.0 sec	1500	29	6.7 (0.95 T_m)	6.4 (1.07 T_m)	7.9 (0.99 T_m)	5.0	4.3	5.8
1.5 sec	1000	44	6.8 (0.96 T_m)	6.5 (1.09 T_m)	7.8 (0.98 T_m)	5.2	4.3	6.0
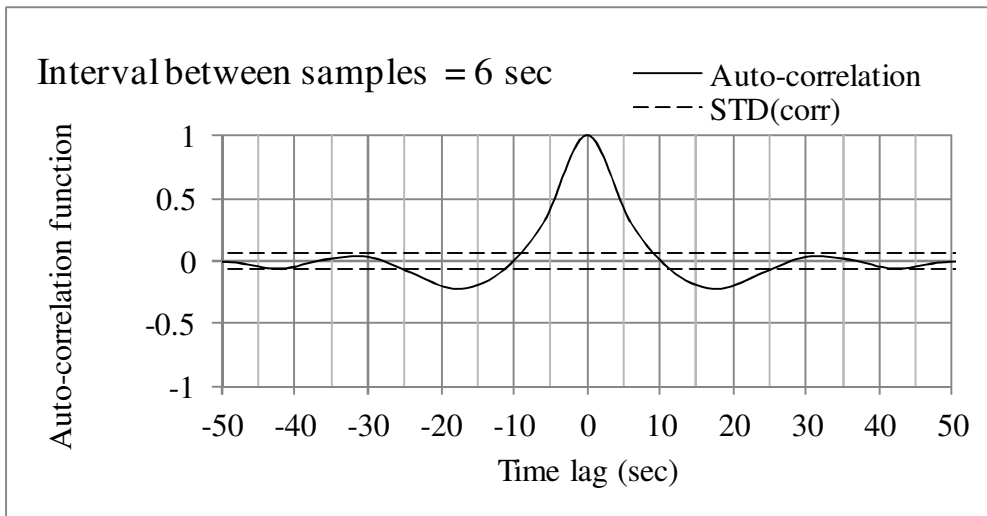
Figure 2-10. Auto-correlation function of buoy 1 for the interval between samples of 6 sec.
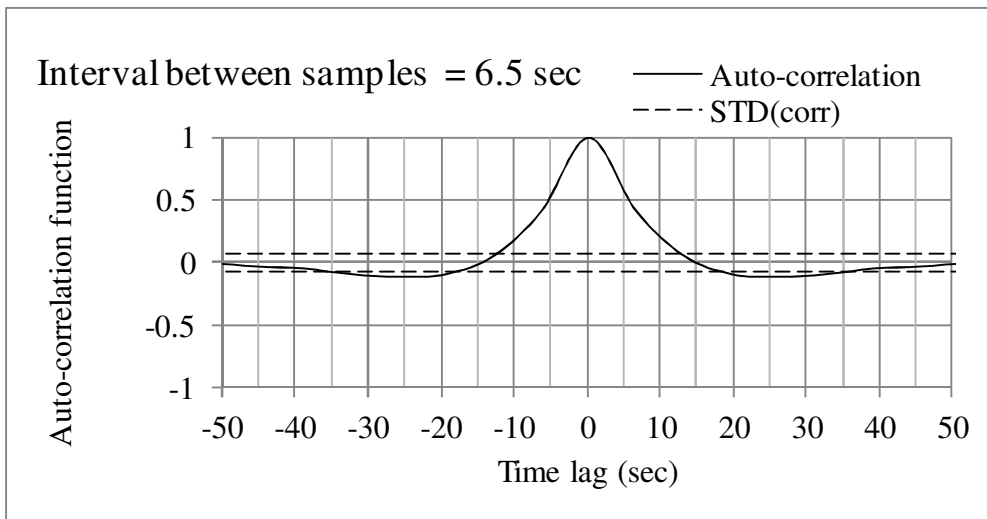
Figure 2-11. Auto-correlation function of buoy 1 for the interval between samples of 6.5 sec.
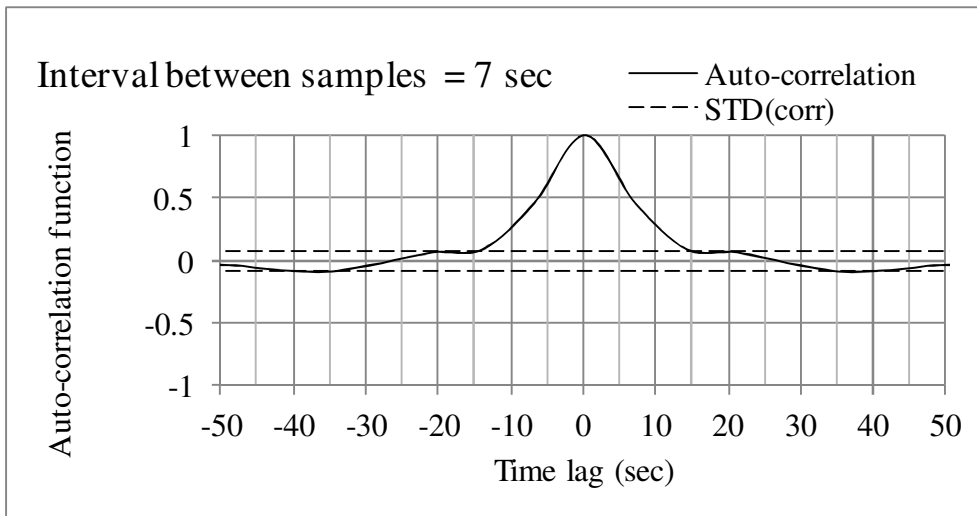
Figure 2-12. Auto-correlation function of buoy 1 for the interval between samples of 7 sec.
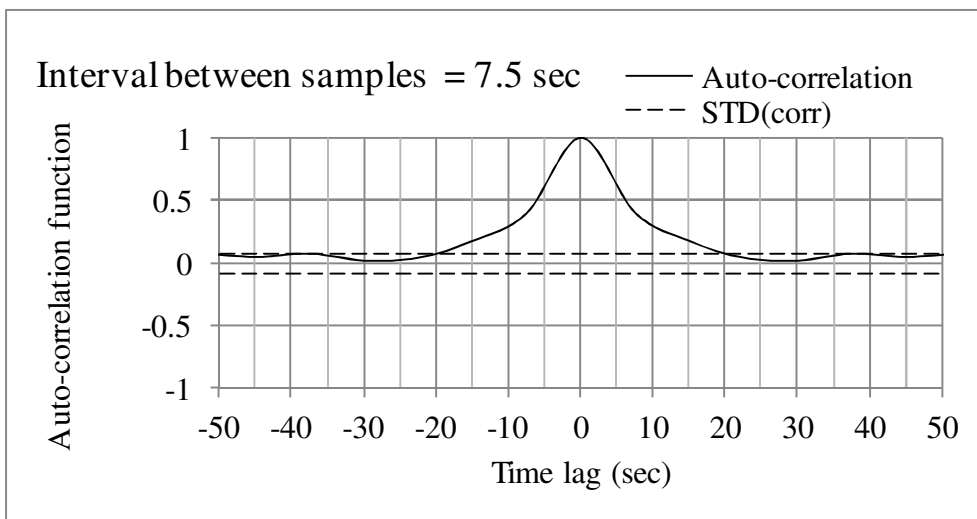
Figure 2-13. Auto-correlation function of buoy 1 for the interval between samples of 7.5 sec.
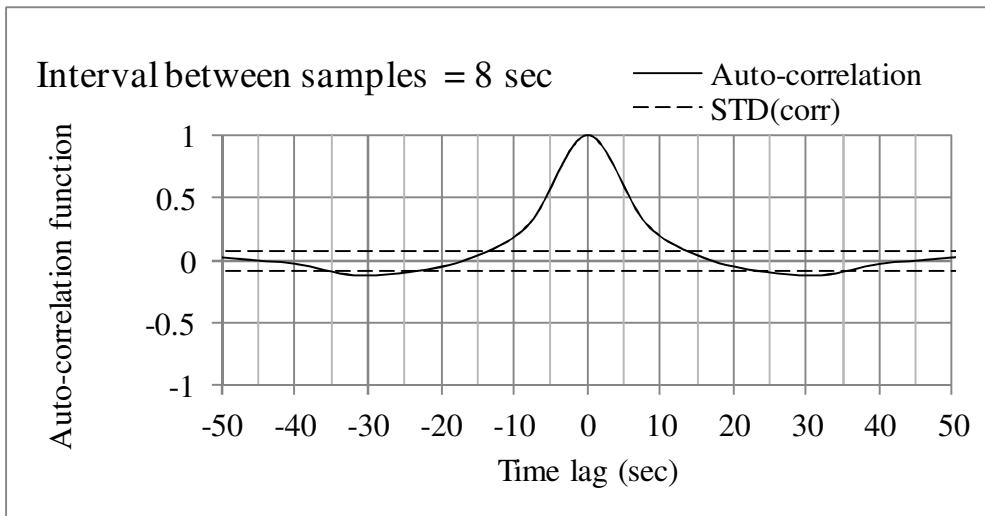
Figure 2-14. Auto-correlation function of buoy 1 for the interval between samples of 8 sec.
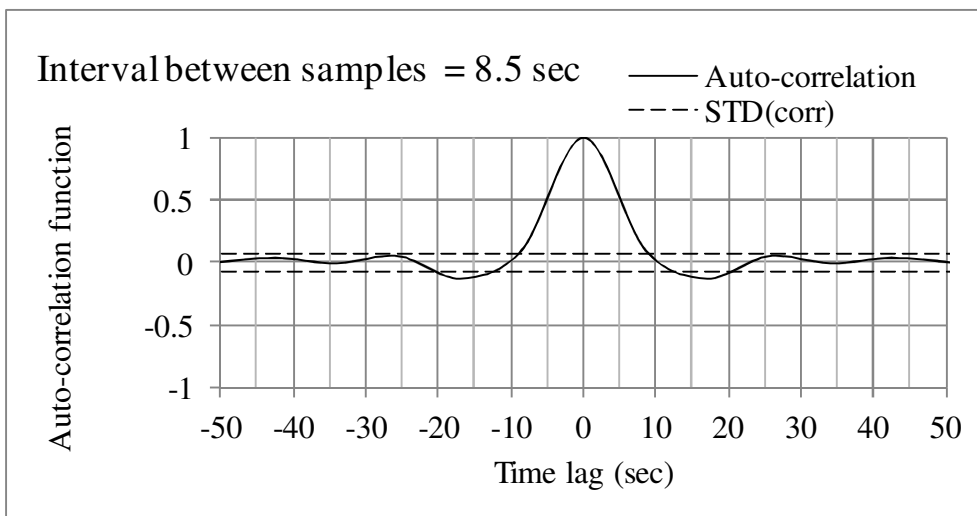
Figure 2-15. Auto-correlation function of buoy 1 for the interval between samples of 8.5 sec.
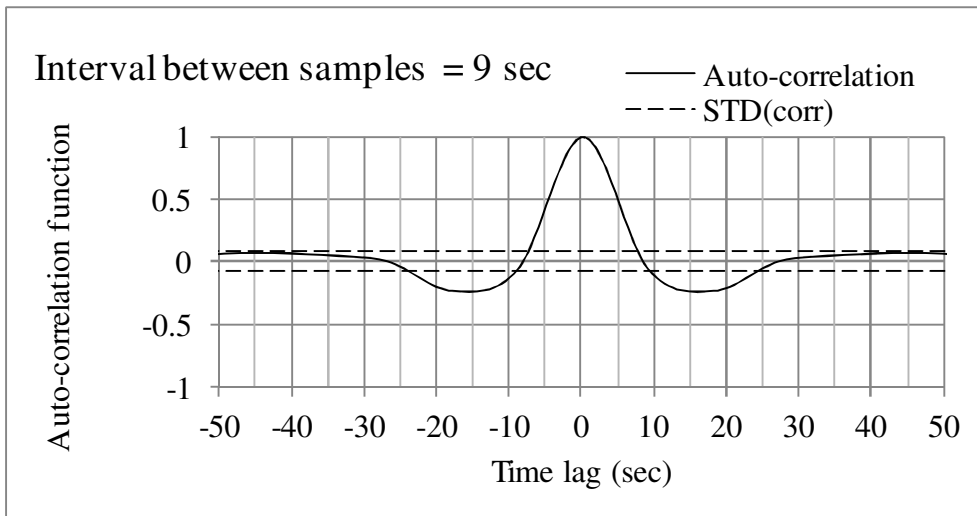
Figure 2-16. Auto-correlation function of buoy 1 for the interval between samples of 9 sec.
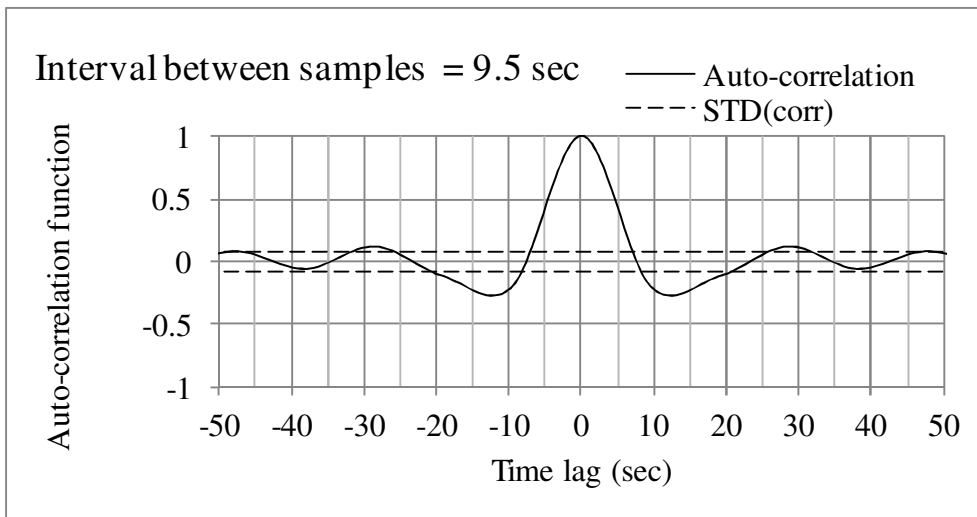
Figure 2-17. Auto-correlation function of buoy 1 for the interval between samples of 9.5 sec.

2.5. Spatial Data Analysis

2.5.1. Semivariogram

Recalling the auto-covariance function demonstrated in Section 2.3, the auto-covariance function is the means to measure or represent the degree of association between values of the random variable $x(t)$ at times differing by a specific interval τ. In the stationary random field, the degree of association between locations in the field can be demonstrated in the same way as the auto-covariance function defined in Equation 2-1 except the times denoted as t_i are now locations denoted as s_i. The distance interval h separating the two locations replaces the time interval τ. While the auto-covariance function is the term used in the time field, it is sometimes called the covariogram in the geographical field. The correlation between two points separated by a distance h is called the correlogram while it has the name as auto-correlation function in the time field.

The variance of the difference between two locations s_1 and s_2 is called a variogram and can be defined as follows:

$$2\gamma(s_1, s_2) = var[y(s_1) - y(s_2)] = E\left[(y(s_1) - y(s_2)) - (\mu(s_1) - \mu(s_2))\right]^2 \quad (2\text{-}28)$$

The function $\gamma(s_1, s_2)$ is called a semivariogram and is closely related to the covariogram (or auto-covariance function). For a stationary field, the variogram between two points separated by a distance h can be written as follows:

$$2\gamma(s, s + h) = \gamma(h) = var[y(s) - y(s + h)] = E[y(s) - y(s + h)]^2$$
$$= var[y(s)] + var[y(s + h)] - 2cov[y(s), y(s + h)] \quad (2\text{-}29)$$

Based on the stationary property, the variance does not depend on the location *s*, hence the variances at *s* and at *s+h* are the same. Equation 2-29 can be re-written as the following form:

$$\gamma(h) = C(0) + C(0) - 2C(h) = 2[C(0) - C(h)] \tag{2-30}$$

By dividing Equation 2-30 by two, the semivariogram is defined as follows:

$$\frac{1}{2}\gamma(h) = C(0) - C(h) \tag{2-31}$$

For a given random process such as wave surface elevation shown in Figure 2-18, the covariogram (or the auto-covariance function) is presented in Figure 2-19. The semivariogram values at different separating distances can be estimated using Equation 2-28 and is shown in Figure 2-20. According Equation 2-31, the semivariogram can also be estimated from the covariogram (or auto-covariance function). Figure 2-21 shows the comparison of the semivariograms estimated from the given random process using Equation 2-28 and from the covariogram (or auto-covariance function) of this process using Equation 2-31. The comparison shows that these two semivariograms agree well in the beginning. When the distance increases, the one estimated from the random process using Equation 2-28 fluctuates while the one estimated from the covariogram using Equation 2-31 approaches the variance. This is because that the covariogram values tend to approach zero when the distance increases since the degree of association between the estimated two points reduces. Hence, the semivariogram estimated from the covariogram would approach the variance of the random process with the increase in distance.
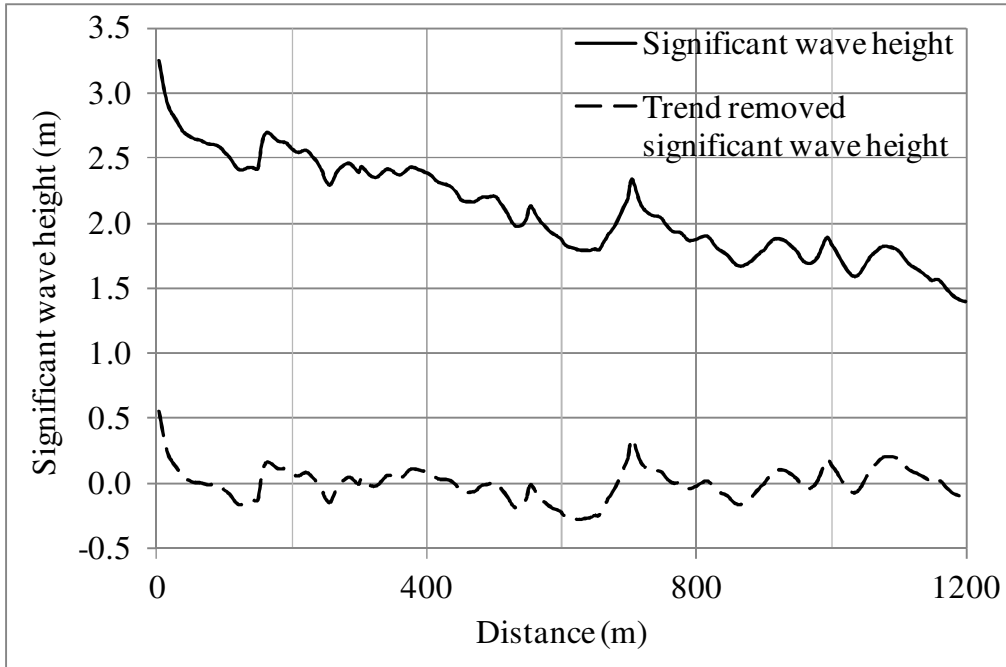
Figure 2-18. A random process of significant wave height.
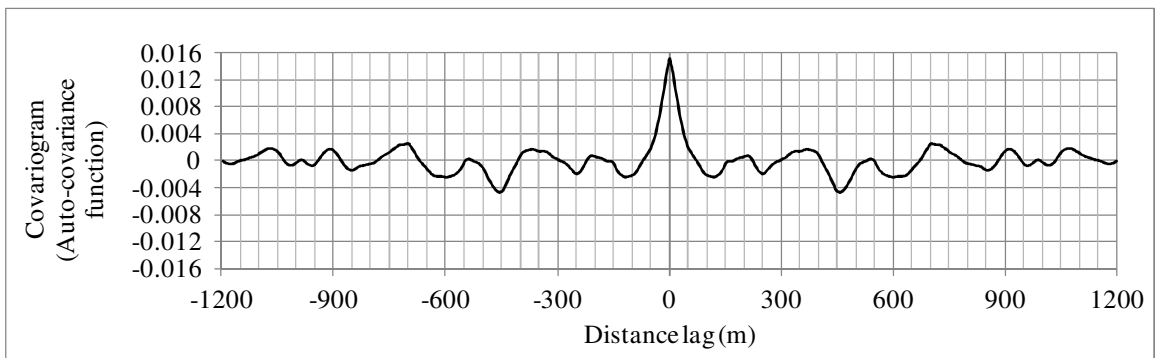
Figure 2-19. Covariogram (or auto-covariance function) for the random process shown in
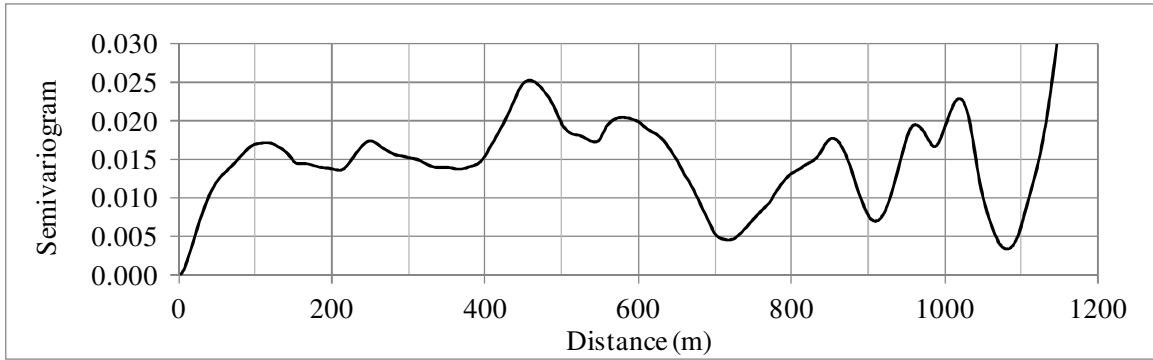
Figure 2-18.

Figure 2-20. Semivariogram for the random process shown in Figure 2-18.
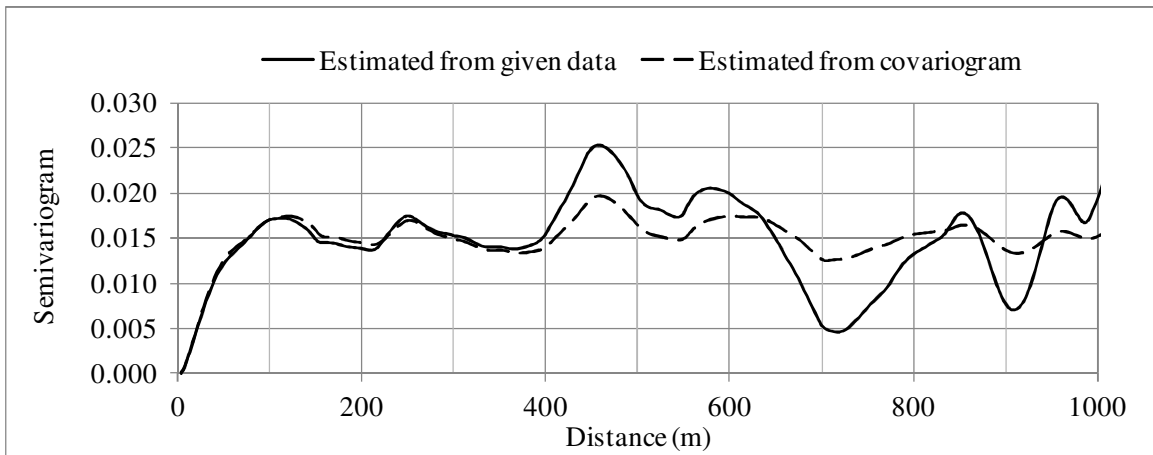
Figure 2-21. Semivariograms estimated from the given data using Equation 2-28 and estimated from the covariogram (or auto-covariance function) of the data using Equation 2-31.

Variograms are often described by parameters such as the sill, denoted as γ_r, and the radius of influence, denoted as r.   In a semivariogram model, the semivariogram value increases when the distance increases.   When reaching certain distance, the semivariogram value reaches its limit.   The variance of the difference between analyzed locations becomes negligible from this distance and beyond.   The semivariogram value

44

at this limit is called the sill (γ_r), and the distance of reaching this limit is called the radius

of influence (r).   Similar to the curve fitting for a set of data, there are some

semivariogram models, such as linear model, exponential model, spherical model, etc., to

represent the sample estimate.   The spherical model is often used and can be expressed

as follows (described by McCuen and Snyder, 1986):

$$\gamma(h) = \frac{\gamma_r}{2}\left[\frac{3h}{r} - \left(\frac{h}{r}\right)^3\right] \quad , h \leq r$$
$$\gamma(h) = \gamma_r \quad\quad\quad\quad\quad , h > r$$

(2-32)

in which γ_r is the sill, r is the radius of influence and h is the distance.   Figure 2-22 is an

illustrative spherical model with the sill (γ_r) of 0.1 and the radius of influence (r) of 650

m.   If the type of model, the sill and the radius of influence are known, a semivariogram

can be determined and spatial predictions can be performed.

According to Equation 2-1, when the distance h approaches infinity, the

covariogram (or auto-covariance function) tends to approach zero.   Given a random

process, the semivariogram can be estimated based on the covariogram (or

auto-covariance function) by using Equation 2-31.   Therefore, the semivariogram value

approaches the series variance when the distance h approaches infinity, i.e. $\gamma(h \to \infty) =$

$C(0)$.   In other words, the sill value γ_r is the variance of the process, i.e. $\gamma_r = C(0)$.   By

the fact that the sill value γ_r approaches the variance of the random process, the radius of

influence r can be determined.   Figure 2-23 shows an illustrative example to obtain the

radius of influence r from the series covariogram and the series variance.   The radius of

influence can be determined by the steps as follows: (1) indicate the series variance on

the semivariogram value axis; (2) draw a horizontal line from the semivariogram axis

which indicates the series variance value to meet the semivariogram curve; (3) draw a

vertical line from the point at which the semivariogram curve was met to the distance

axis; (4) locate the point on the distance axis which meets the vertical line coming down

from the semivariogram curve as the radius of influence.
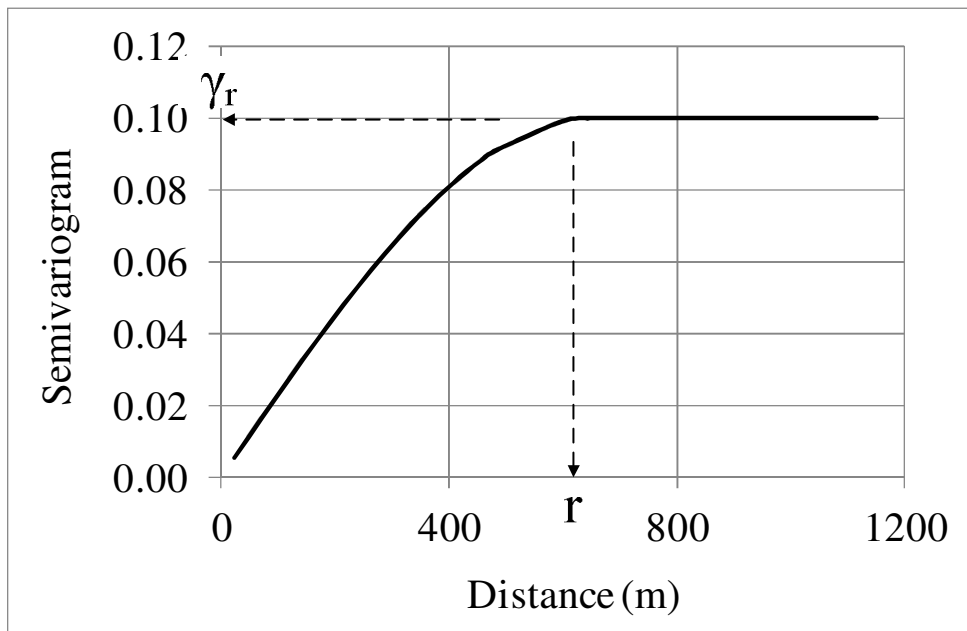
Figure 2-22. Illustration of the sill γ_r and the radius of influence r for a spherical

semivariogram model.
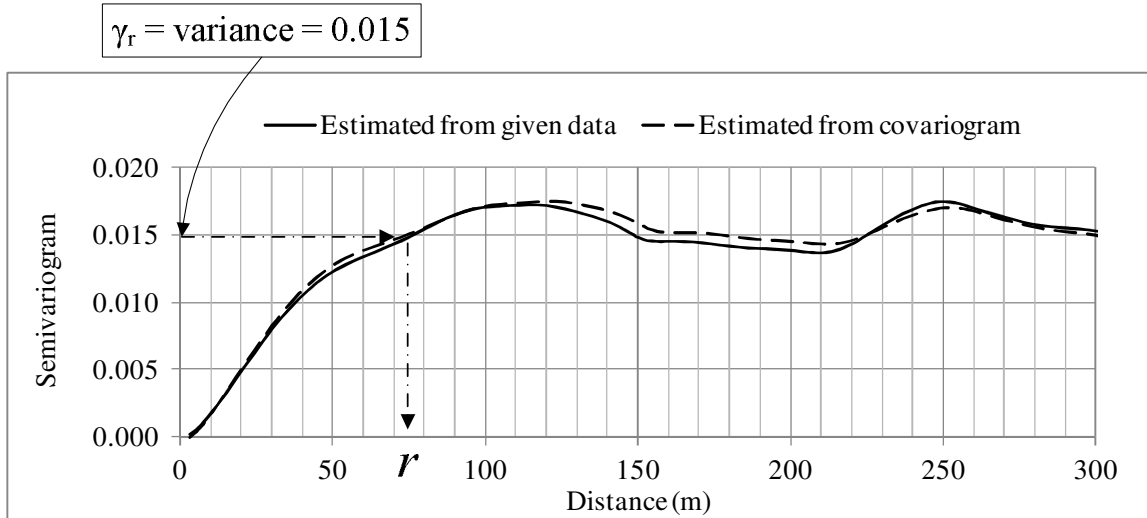
Figure 2-23. Determination for sill and radius of influence from the covariogram and the variance.

2.5.2. Kriging

Kriging is a technique used to interpolate the value of variables of interest at an unobserved location from observations of the interested variables at nearby locations. The variables of interest are functions of geographic locations.   Estimation by Kriging is to apply weights on the variable values at observed locations to estimate the variable value at the unobserved location.   The weight should be inversely proportional to the distance separating the estimated unknown location and the location that the weight applies on.   That is, the more weight should be applied on the locations which are nearby while the less weight should be applied on the locations which are far away.   The weights for the Kriging estimation can be denoted as wk_i.   The sum of the weights equals one expressed as:

$$\sum_{i=1}^{m} wk_i = 1 \tag{2-32}$$

47

in which m is the number of observations the weights applied on.   The estimation by

Kriging technique has the following form:

$$\hat{y} = \sum_{i=1}^{m} wk_i * y_i \tag{2-33}$$

in which $\hat{y}$ is the variable at the unobserved location s_A, y_i is the variable value at

location s_i and wk_i is the weight applied on the variable value at location s_i.   The weights

wk_i are unknown and are estimated as follows:

$$\begin{cases} \alpha + \sum_{i=1}^{m}\big(wk_i * \gamma(s_1, s_i)\big) = \gamma(s_1, s_A) \\ \alpha + \sum_{i=1}^{m}\big(wk_i * \gamma(s_2, s_i)\big) = \gamma(s_2, s_A) \\ \qquad\qquad \vdots \\ \alpha + \sum_{i=1}^{m}\big(wk_i * \gamma(s_m, s_i)\big) = \gamma(s_m, s_A) \\ \qquad \sum_{i=1}^{m} wk_i = 1 \end{cases} \tag{2-34}$$

in which α is a constant and γ is the semivariogram values demonstrated in Section 2.5.1.

Equation 2-34 has a set of $m+1$ equations to solve $m+1$ unknowns which are wk_1, wk_2, ...,

wk_m and α.

To illustrate how the Kriging estimation works, consider a random field shown in

Figure 2-24 which has the sill (γ_r) of 0.0303 and the radius of influence (r) of 700 m as

shown in Figure 2-25.   Assuming that the spherical semivariogram model represents the

data, the semivariogram values can be estimated by:

$$\begin{aligned} \gamma(h) &= \frac{0.0303}{2}\left[\frac{3h}{700} - \left(\frac{h}{700}\right)^3\right] \quad , h \le r \\ \gamma(h) &= 0.0303 \qquad\qquad\qquad\quad , h > r \end{aligned} \tag{2-35}$$

If the significant wave heights are known at buoys 1, 2, and 3 shown in Figure

2-24, the significant wave height at point 1 can be interpolated using Equations 2-33,

2-34 and 2-35.   The coordinates and significant wave height values are summarized in

Table 2-5.   Distances between buoys themselves and between buoys and the estimated

point are shown in Table 2-6.   Semivariogram values computed using Equation 2-35 and

the distance provided in Table 2-6 are summarized in Table 2-7 which has the
information needed to estimate the unknowns in the Kriging estimation process, i.e.
weights wk_i and the constant α in Equation 2-34.   In this example, there are four
unknowns, wk_1, wk_2, wk_3 and α.   Using the semivariogram values shown in Table 2-7,
the four unknowns in Equation 2-34 are solved as follows: $wk_1 = 0.3136$, $wk_2 = 0.5415$,
$wk_3 = 0.1450$ and $\alpha = 0.0044$.   The interpolated significant wave height at point 1 can be
computed using Equation 2-33 as:

$$\hat{y}_{point1} = 0.3136 * 3.13 + 0.5415 * 3.20 + 0.1450 * 2.42 = 3.06 \qquad (2\text{-}36)$$
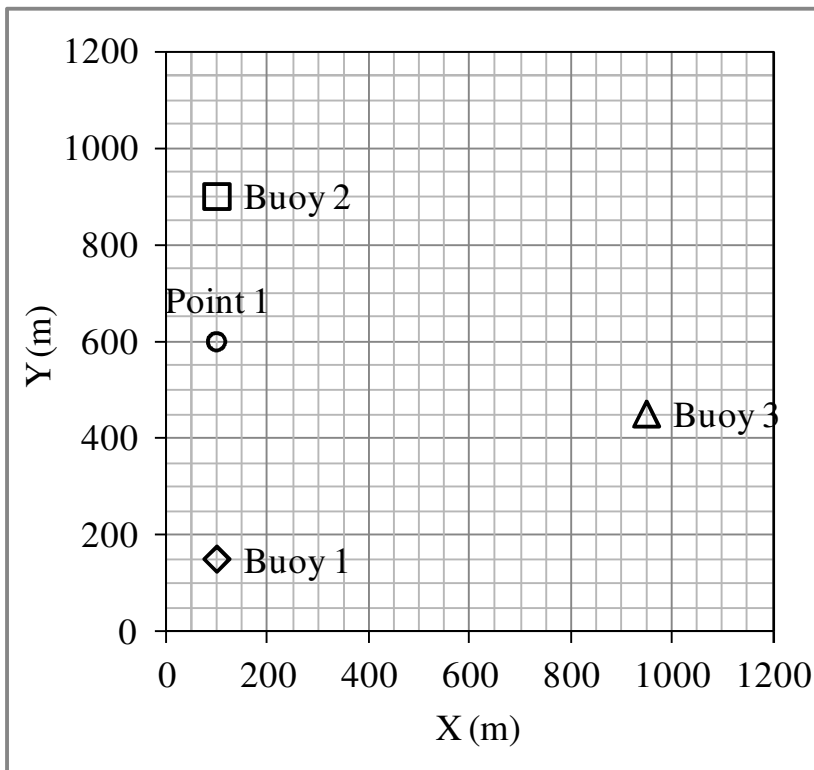
Figure 2-24. Studied random field and locations for estimation.

$\gamma_r = variance = 0.0303$
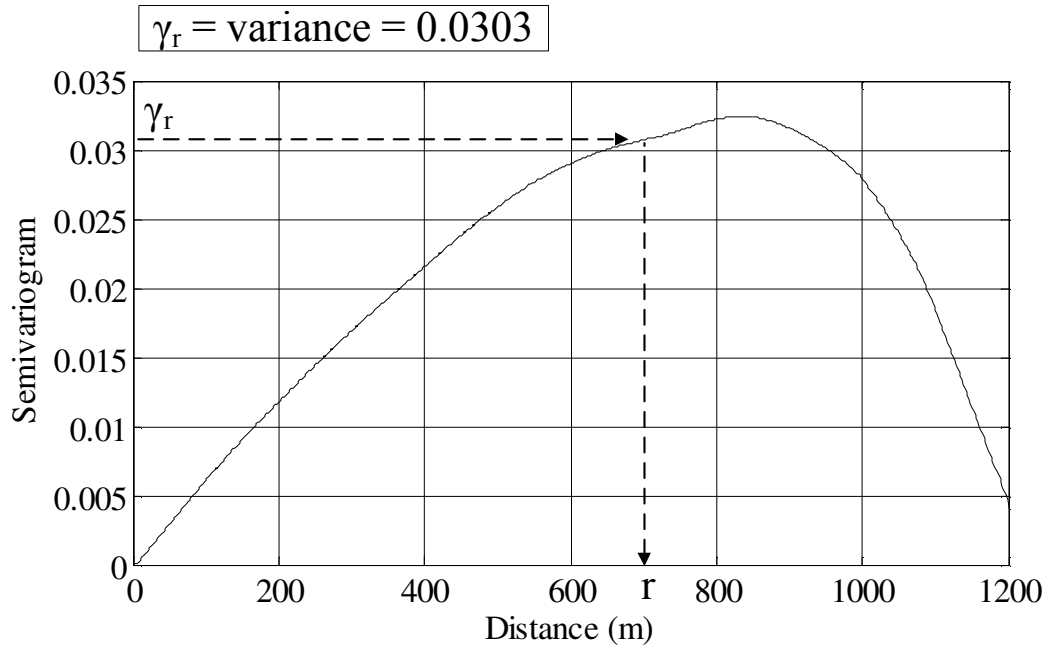
Figure 2-25. Semivariogram, sill and radius of influence of the studied random field.

Table 2-5. Coordinates and significant wave heights for locations used in estimation.

Location	X (m)	Y (m)	Hs (m)
Point 1	100	600	unknown
Buoy 1	100	150	3.13
Buoy 2	100	900	3.20
Buoy 3	950	450	2.42

Table 2-6. Distance between locations used in estimation.

Distance (m)	Buoy 1	Buoy 2	Buoy 3
Point 1	450	300	863
Buoy 1	0	750	901
Buoy 2		0	962
Buoy 3			0

50

Table 2-7. Semivariogram values for locations used in estimation.

Semivariogram	Buoy 1	Buoy 2	Buoy 3
Point 1	0.0252	0.0183	0.0303
Buoy 1	0	0.0303	0.0303
Buoy 2		0	0.0303
Buoy 3			0

2.5.3. Sample Size of Independent Observations

Considering a set of significant wave height data collected at multiple locations evenly separated along a line, the range for the interval between samples could be determined by Equations 2-26 and 2-27 based on the sample sizes of independent observations estimated using Equations 2-22 and 2-25 that demonstrated in Section 2.4.5.

Figure 2-18 shows a random process of significant wave heights collected at 400 locations evenly separated by 3 m along a line.   The number of discretization points N is 400 and the distance interval Δh is 3 m.   The sample size for independent observations is estimated as 25 based on Equations 2-22 and 2-25.   The interval between samples can be computed from the estimated sample size of independent observations using Equations 2-26 and 2-27 except the time increment Δt is now replaced by the distance increment Δh.   The interval between samples based on the sample size of independent observations estimated is $(N = 400) * (\Delta h = 3\text{m}) / (n = 25) = 48$ m.   The auto-correlation function is shown in Figure 2-26 with its standard deviation range shown in dashed line.   As mentioned in Section 2.4.5, the auto-correlation values exceeding its standard deviation are considered significantly different from zero.   The auto-correlation function shown in Figure 2-26 reaches into its standard deviation range at distance lag of approximately 48

m.   This observation agrees with the estimation of the interval between samples for

obtaining independent data using Equations 2-22, 2-25 and 2-26, which is 48 m as well.

Figure 2-27 through Figure 2-30 show the auto-correlation functions of several

different intervals between samples in the range of 24 m to 69 m.   The distance lag is

shown up to 120 m.   The auto-correlation function tends to stay within its standard

deviation range when the interval between samples approaches 48 m.   When the interval

between samples passes 48 m, the auto-correlation function fluctuates exceeding the

standard deviation range again.
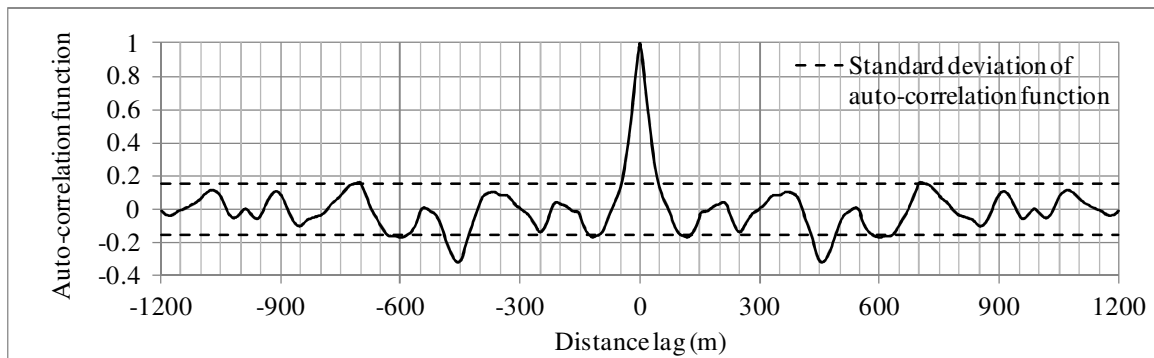
Figure 2-26. Auto-correlation function for the random process shown in Figure 2-18.
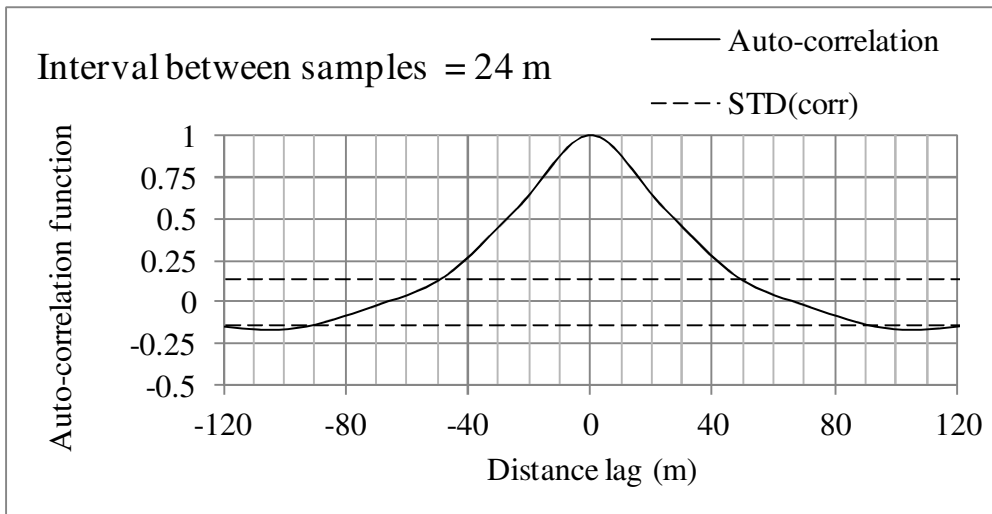
Figure 2-27. Auto-correlation function for the case of interval between samples = 24 m.
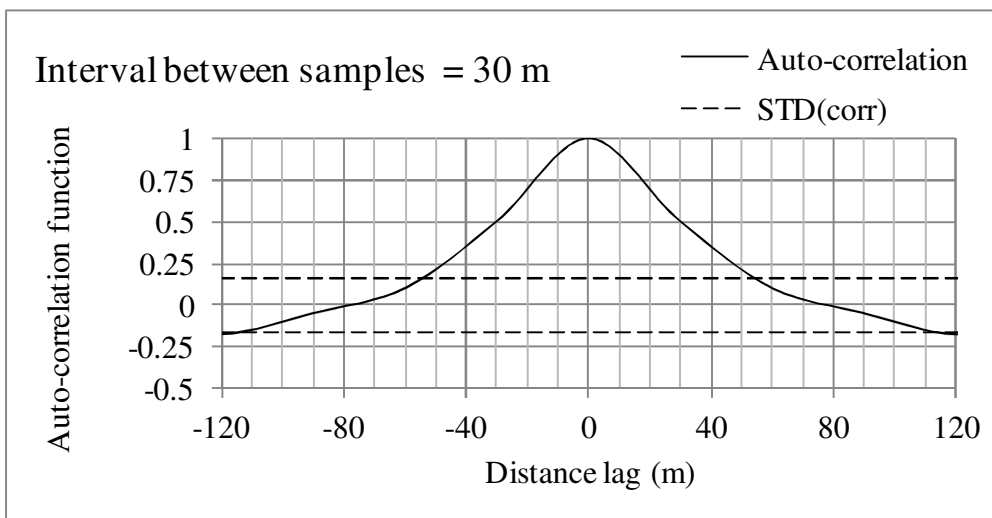
Figure 2-28. Auto-correlation function for the case of interval between samples = 30 m.
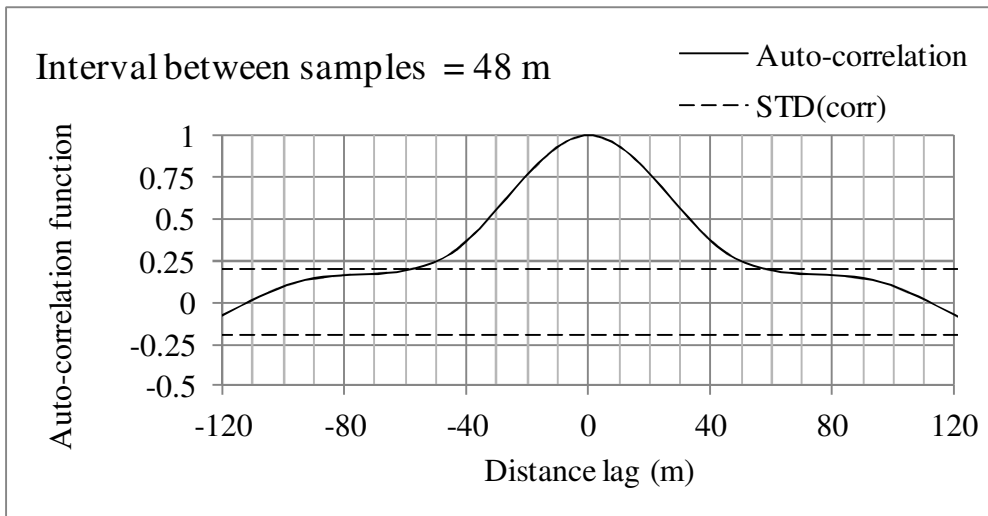
Figure 2-29. Auto-correlation function for the case of interval between samples = 48 m.
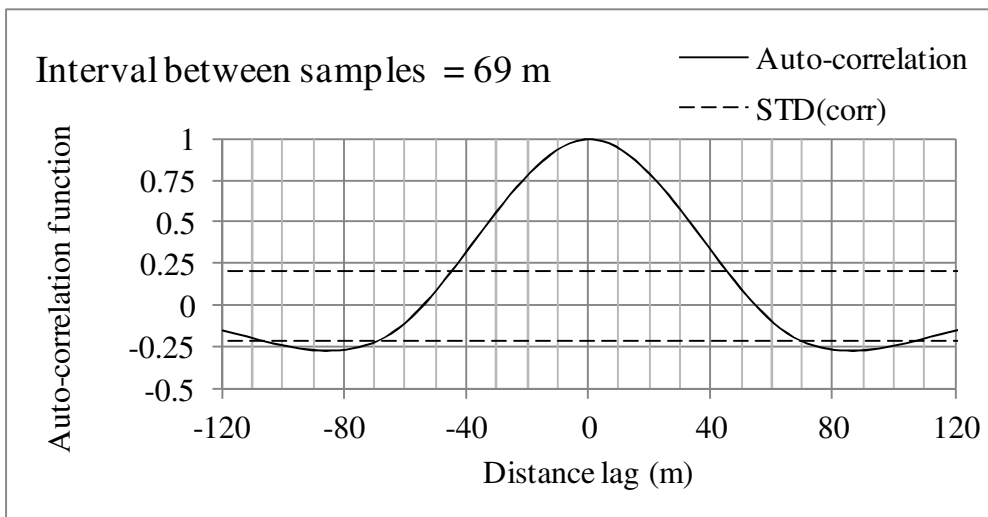
Figure 2-30. Auto-correlation function for the case of interval between samples = 69 m.

3. Methodology for Characterizing Sea Conditions

This chapter provides the statistical characterization procedure for predicting sea conditions in Section 3.1 and the numerical wave model SWAN for verifying the spatial data interpolation in Section 3.2. The numerical wave model SWAN is utilized to verify the spatial interpolation as well as the sea-state statistical characterization using the methods demonstrated in Section 3.1. Numerical examples will be provided in Chapter 4 for illustration and verification.

3.1. Statistical Characterization for Predicting Sea-state Condition

3.1.1. Overview

This section provides the statistical characterization procedure as a flowchart shown in Figure 3-1 for characterizing sea-state conditions using given buoy vertical elevation time histories and predicting sea-state conditions of unobserved points from data of observed buoys. The sea-state prediction starts with obtaining buoy vertical elevation time histories at points of interest followed by identifying key parameters of interest for analysis. In general, wave modal period and significant wave height are the two key parameters to characterize the sea condition. According to Chapter 2, the sample size of independent observations can be determined and the peroidograms can be constructed from the time histories. Upon obtaining the information needed, Sections 3.1.2 through 3.1.7 provide the procedure and demonstrations for the parameter estimations of the buoy locations. The unobserved locations within buoys range could

be interpolated by the procedure demonstrated in Section 3.1.8 and the parameters can be

estimated by the same procedure applied on the buoys.
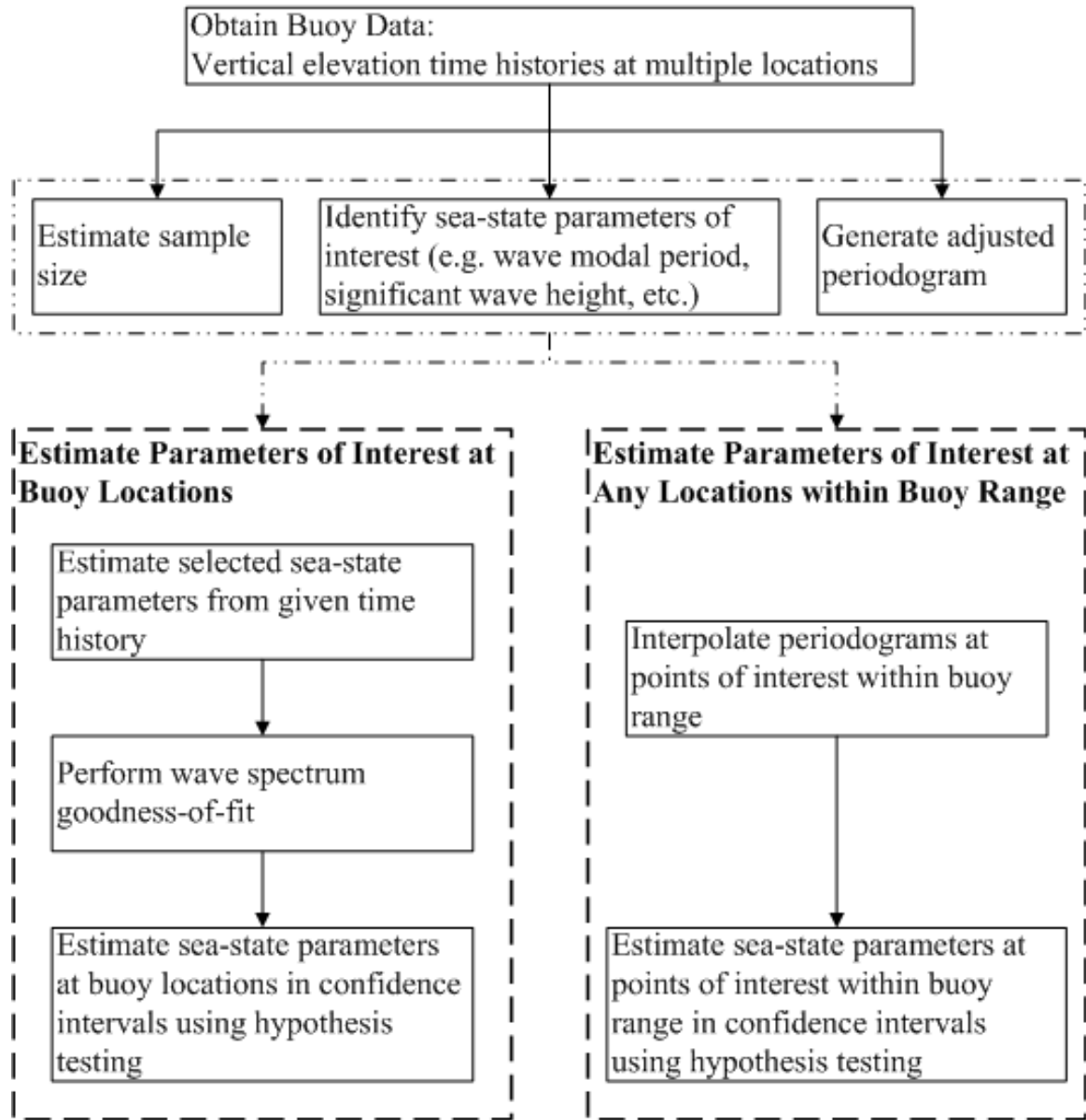
Figure 3-1. Statistical characterization and prediction for sea-state conditions of observed

and unobserved locations.

3.1.2. Sea-state Parameters

Sea-state prediction in this study utilizes the sea-state parameter estimates. Two-parameter wave model is chosen to represent the sea condition. The two parameters are the wave modal period and the significant wave height. The wave modal period T_m can be estimated from the time history using the zero-upcrossing period expressed as follows (described by Hughes 1988):

$$T_{mt} = \begin{cases} 1.41 T_z & \text{for Bretschneider spectrum} \\ 1.28 T_z & \text{for Jonswap spectrum} \end{cases} \tag{3-1}$$

in which T_{mt} is the modal period estimated from the time history and T_z is the zero-upcrossing period. The other key parameter, the significant wave height H_s, is related to the variance of the time history and is expressed as follows (described by Hughes 1988):

$$H_{st} = 4 \sqrt{\text{variance of time history}} \tag{3-2}$$

in which H_{st} is the significant wave height estimated from the time history.

The two parameters, modal period and significant wave height, can be estimated in the time domain from the time history and in the frequency domain from the wave spectrum. While using Equations 3-1 and 3-2 to estimate these two parameters in the time domain, the frequency at which the maximum spectrum magnitude locates represents the inverse value of the modal frequency, i.e. $2\pi / T_m$, and the area under the spectrum curve represents the variance of the time history which is the information needed to compute the significant wave height using Equation 3-2.

3.1.3. Adjusted Periodogram

In order to utilize the periodogram introduced in Chapter 2 for sea-state characterization, some adjustments on the periodogram are needed. Based on the parameters estimated from the buoy time history in Equations 3-1 and 3-2, the periodogram constructed from the time history needs to be shifted to match the peak at the estimated modal frequency $2\pi / T_{mt}$ in which T_{mt} is calculated by Equation 3-1. Then the periodogram needs to adjust the magnitude to match its unit the same as that of the wave spectrum and to have the area under the periodogram curve as the variance of the time history. The wave spectrum goodness-of-fit can be performed once the periodogram is adjusted. Since the wave spectra used to fit the periodogram are generated based on the estimated parameters using Equations 3-1 and 3-2, the peaks are at the modal frequency $2\pi / T_{mt}$. Besides, it wouldn't be possible to compare the periodograms and the wave spectra if they have different units. An illustrative figure of the periodogram shifting is shown in Figure 3-2.
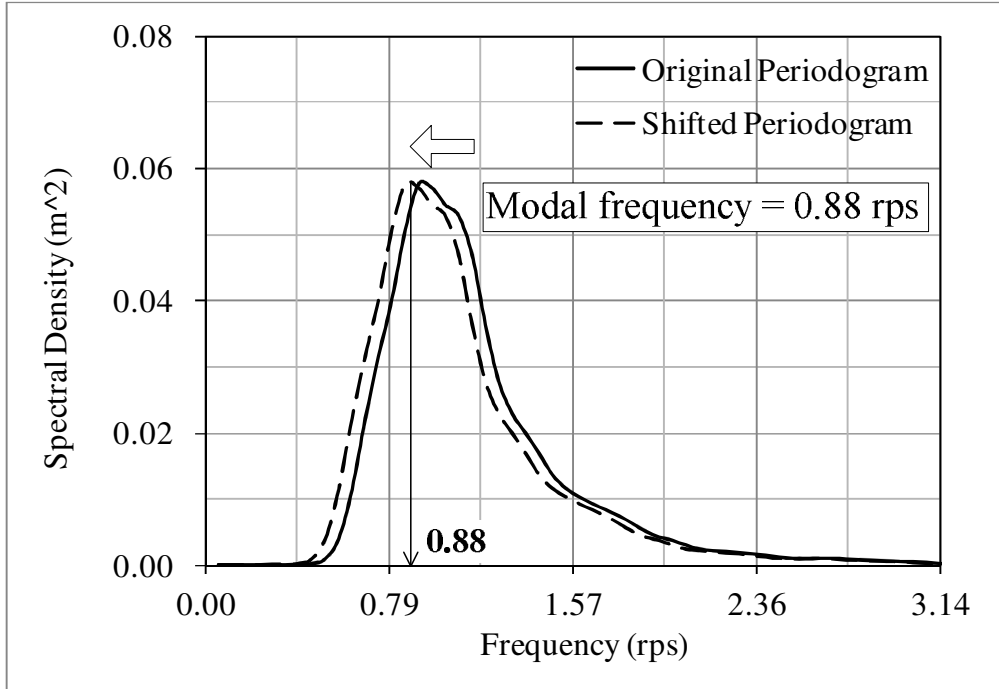
Figure 3-2.Periodogram shifts to match the peak at modal frequency $2\pi / T_{mt}$.

3.1.4. Wave Spectrum Goodness-of-fit

The wave spectrum goodness-of-fit is performed by fitting the periodogram constructed from the buoy time history using multiple wave spectra. Figure 3-3 shows an illustrative plot of wave spectrum goodness-of-fit. Two-parameter wave spectra such as the Bretschneider spectrum and the Jonswap spectrum are utilized to fit the periodogram. The wave spectra are generated using the two parameters, the wave modal period and the significant wave height, estimated from the given buoy time history based on Equations 3-1and 3-2. Figure 3-3 shows that Bretschneider spectrum is a better fit for the periodogram than the Jonswap spectrum in terms of the spectrum shape. In the quantitative manner, the energy estimated from the Bretschneider spectrum, which is the area under the spectrum curve, is closer to the energy estimated from the

periodogram compared with that estimated from the Jonswap spectrum.   Therefore, the

wave spectrum goodness-of-fit shows that the periodogram presented in Figure 3-3 is a

Bretschneider spectrum type.   This determination of spectrum type will be used to

generate spectra for estimating the confidence intervals of the parameters of interest in

the sea-state characterization procedure.
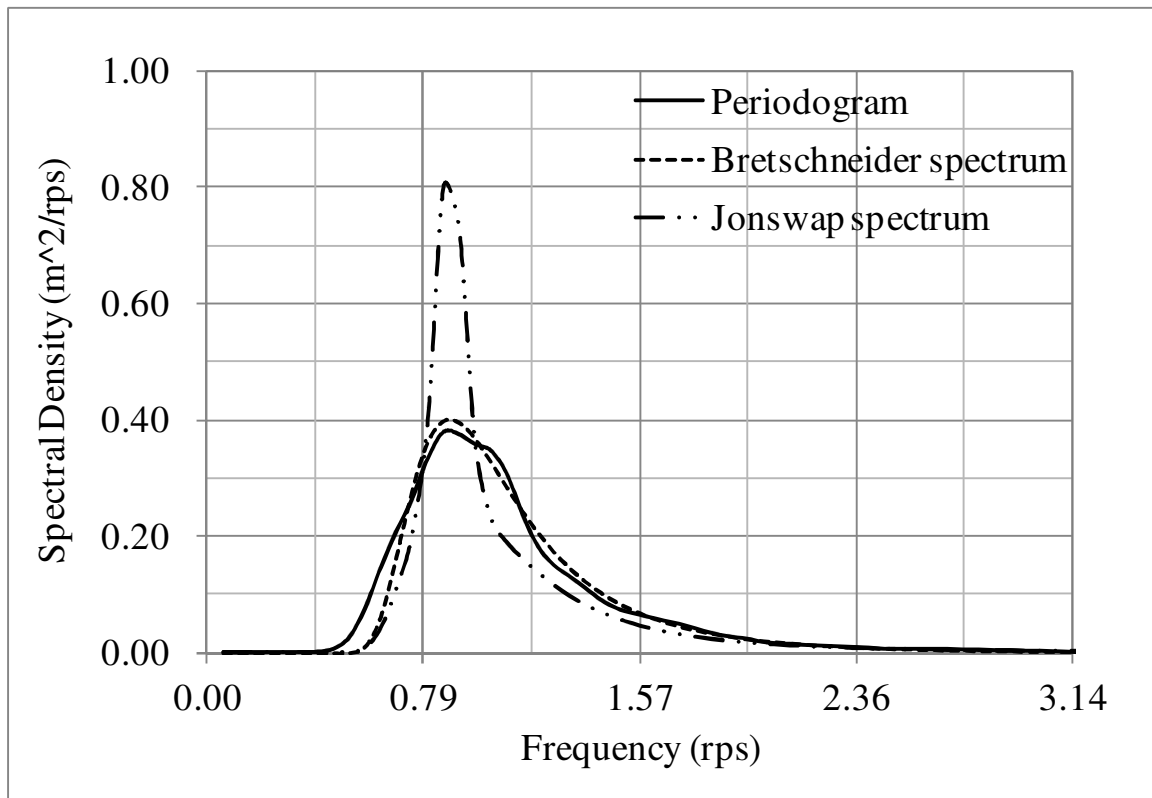
Figure 3-3. Wave spectrum goodness-of-fit using Bretschneider and Jonswap spectra.

3.1.5. Hypothesis Testing

Hypothesis testing is a decision making process for extrapolating information

from sample data to describe a population.   In this study, the periodograms constructed

from the buoy wave surface elevation time histories are used to illustrate the sea wave

characterization in the spectral domains.   Statistical hypothesis testing is performed to

describe the sea wave characteristics, to demonstrate the sea wave spectrum

goodness-of-fit test and to interpolate the buoy data to predict the sea-state characteristics

at locations of interest.   The hypotheses are defined as follows.   The null hypothesis,

denoted by H_0, represents the equality of two spectra $f_1(\omega)$ and $f_2(\omega)$; while the alternative

hypothesis, denoted by H_1, indicates that a significant difference between two spectra

exists.   The hypotheses can be expressed as:

$$H_0: f_1(\omega) = f_2(\omega) \tag{3-3}$$

$$H_1: f_1(\omega) \neq f_2(\omega) \tag{3-4}$$

As described previously, the quantity $vPdg(\omega) / f(\omega)$ follows a chi-square

distribution with v degrees of freedom.   Consider a statistic X given by the following

ratio:

$$X_i = \frac{v_1 Pdg_1(\omega_i)}{f_1(\omega_i)} / \frac{v_2 Pdg_2(\omega_i)}{f_2(\omega_i)} \quad , \quad i = 1, 2, \ldots, M \tag{3-5}$$

in which M is the number of auto-covariance coefficients considered.   It is assumed that

the two time history data sets have the same truncation points M and lag window $\{\lambda_k\}$.

The degrees of freedom v_1 and v_2 are calculated using Equation 2-22 which replaces

Equation 2-17 demonstrated in Chapter 2.

The random quantity of Equation 3-5 is distributed according to F-distribution

with v_1 and v_2 degrees of freedom, denoted as $F(v_1, v_2)$.   The mean and variance of

F-distribution with v_1 and v_2 degrees of freedom, denoted as $F_{v1, v2}$, are:

$$E[F_{v_1, v_2}] = \frac{v_2}{v_2 - 2} \quad , \quad v_2 > 2 \tag{3-6}$$

$$var[F_{v_1, v_2}] = \frac{2v_2^2(v_1+v_2-2)}{v_1(v_2-2)^2(v_2-4)} \quad , \quad v_2 > 4 \tag{3-7}$$

In the case when the null hypothesis $H_0: f_1(\omega) = f_2(\omega)$ is true, X_i does not depend on the underlying spectra and can be expressed in the following form:

$$X_i = \frac{Pdg_1(\omega_i)}{Pdg_2(\omega_i)} \quad , \quad i = 1, 2, \dots, M \quad , \quad \text{if } H_0 \text{ is true} \tag{3-8}$$

The following statistic is suggested to test the null hypothesis $H_0: f_1(\omega) = f_2(\omega)$ with the alternative hypothesis $H_1: f_1(\omega) \neq f_2(\omega)$ as:

$$Q = \sum_{i=1}^{M} X_i \tag{3-9}$$

The distribution of the quantity Q in Equation 3-9 is the M-fold convolution of F-distribution with v_1 and v_2 degrees of freedom. Since the quantity X_i in Equation 3-8 are independent and identically distributed, according to the central limit theorem, Q for a large sample size is normally distributed with the mean and variance as:

$$E[Q] = M\left(\frac{v_2}{v_2-2}\right) \quad , \quad v_2 > 2 \tag{3-10}$$

$$var[Q] = M\left(\frac{2v_2^2(v_1+v_2-2)}{v_1(v_2-2)^2(v_2-4)}\right) \quad , \quad v_2 > 4 \tag{3-11}$$

Note that M is the number of auto-covariance coefficients considered, or is called the truncation point.

3.1.6. Confidence Interval Estimation using Hypothesis Testing

Since the quantity Q expressed in Equation 3-9 is normally distributed with the mean and variance expressed in Equations 3-10 and 3-11, the $100(1-\alpha)\%$ confidence interval can be expressed as follows:

$$\text{Pr}\left(\begin{array}{l} E[Q] + \sqrt{var[Q]} * \Phi^{-1}\left(\frac{\alpha}{2}\right) < Q \\ < E[Q] + \sqrt{var[Q]} * \Phi^{-1}(1 - \alpha/2) \end{array}\right) = 1 - \alpha \qquad (3\text{-}12)$$

The 100(1-α)% confidence limits, lower confidence limit (*LCL*) and upper confidence

limit (*UCL*), for *Q* are:

$$LCL = E[Q] + \sqrt{var[Q]} * \Phi^{-1}\left(\frac{\alpha}{2}\right) \qquad (3\text{-}13)$$

$$UCL = E[Q] + \sqrt{var[Q]} * \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \qquad (3\text{-}14)$$

Figure 3-4 shows the estimation of the confidence interval for the quantity *Q*.

The quantities *Q* within the confidence limits indicate the range of the estimated

parameter of interest that satisfies the null hypothesis defined in Equation 3-3.   Hence,

the confidence interval of the estimated parameter of interest is determined.
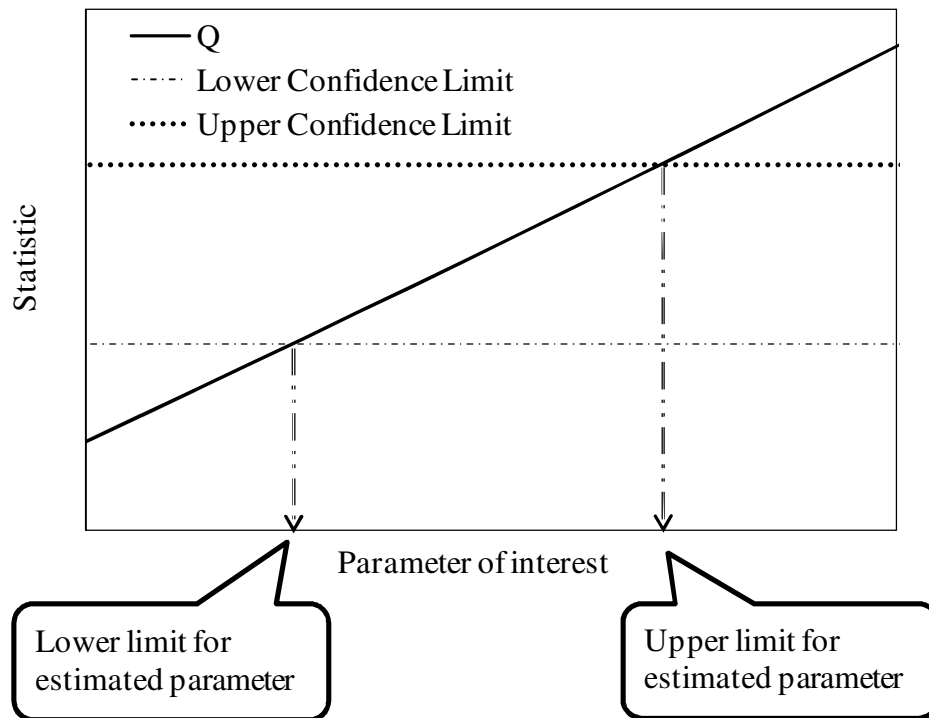
Figure 3-4. Illustration of confidence interval estimation for parameter of interest.

3.1.7. Probability Distribution of the Testing Statistic

The testing statistic Q demonstrated in Section 3.1.5 determines the confidence interval. This statistic is assumed to be normally distributed according to the central limit theorem. To verify that the assumption is correct, the histogram of one thousand quantities of Q is shown in Figure 3-5. Figure 3-5 shows a bell shape histogram and is normally distributed based on the goodness-of-fit test using chi-square critical value. Therefore, the assumption that the testing statistic Q defined in Equation 3-9 is normally distributed is verified and the lower and upper limits for the confidence interval can be determined by Equations 3-13 and 3-14.
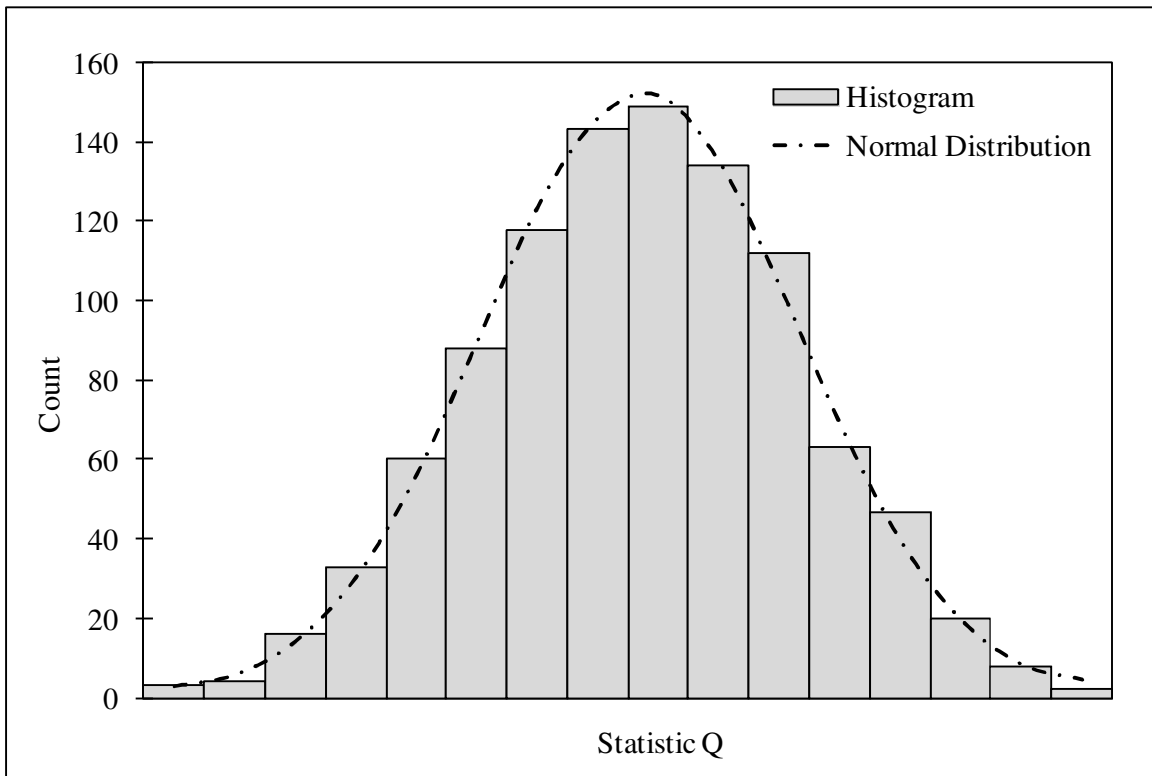
Figure 3-5. Histogram and distribution for quantity Q defined in Equation 3-9.

3.1.8. Spatial Analysis and Data Interpolation

Spatial analysis or spatial statistics is to use the geographic and/or geometric properties for analysis such as interpolation, regression, auto-correlation, etc. The spatial interpolation such as the inverse weight factor is to estimate variables at locations unobserved by applying weights on the properties at observed locations. This section introduces the inverse distance weight factor which is one of the spatial interpolation methods to estimate the sea-state parameters at points of interest within the buoy range from the obtained sea condition at buoy locations.

Inverse distance weight factors are defined inversely proportional to the distance between the points of interest, or the unobserved locations, and the surrounding buoys, or the observed locations. For point i, the weight factor, denoted as wb_{ij}, applied to the data or results of buoy j is defined as follows:

$$wb_{ij} = \frac{1/d_{ij}}{\sum_j 1/d_{ij}} \quad \text{for } d_{ij} \neq 0 \tag{3-15}$$

in which d_{ij} is the distance between point i and buoy j defined as follows:

$$d_{ij} = \sqrt{(ptx_i - Bx_j)^2 + (pty_i - By_j)^2} \tag{3-16}$$

where ptx_i and pty_i are the coordinates of point i, and Bx_j and By_j are the coordinates of buoy j. For each point i, the summation of the weight factors applied on observed buoy points j is one, i.e.:

$$\sum_j wb_{ij} = 1 \tag{3-17}$$

As the point of interest approaches one of the buoys, the weight factor applying on that buoy approaches one, and the weight factors applying on the other buoys approach zero.

The periodogram values $Pdg(\omega_i)$ defined previously in Section 2.4.3 represent the contribution to the wave variance and energy in the range of $\omega_i \pm \delta\omega / 2$, in which $\delta\omega$ is the frequency interval. The periodogram of unobserved point i, denoted as Pdg_i, is estimated as the summation of the periodogram Pdg_j, multiplied by the weight factor wb_{ij} of each surrounding buoy j accordingly expressed as follows:

$$Pdg_i = \sum_j (wb_{ij} * Pdg_j) \qquad (3\text{-}18)$$

This data interpolation process is a one-dimensional linear estimation. The inverse distance is taken to power of one. In Chapter 4, the data interpolation using distance square, and the interpolation in two-dimensional aspect will be discussed and summarized in the example using SWAN generated data.

Based on the estimated periodogram demonstrated in Equation 3-18, the sea-state parameters, the modal period T_m and the significant wave height H_s, of the point of interest i are estimated by the principle of lease squares as shown in Figure 3-6. The modal period and the significant wave height of the point of interest i are estimated in the range determined by the minimum and maximum modal period T_m and the significant wave height H_s of the buoys. As shown in Figure 3-6, T_{min} and T_{max} determine the modal period range for estimating the unobserved point of interest. T_{min} and T_{max} are the minimum and maximum modal period of the buoys, respectively. Similarly, the same definition applies on the significant wave height estimation range. By discretizing the estimation ranges, the matrix-like form such as Figure 3-6 is determined. Note that the discretization interval is subjective. Large intervals may not be able to provide accurate estimations. For each pair of the modal period T_q and the significant wave height H_{sp}, the summation of the squares of the errors between the adjusted periodgram of the point

estimated from the buoys, denoted as Pdg_i, and the adjusted periodograms constructed from this pair of sea-state parameters are calculated and expressed as follows:

$$error_{pq} = \sum (Pdg_i - Pdg_{pq})^2 \qquad (3\text{-}19)$$

The estimation according to Figure 3-6 and Equation 3-19 is performed using various wave spectrum types, such as the Bretschneider and the Jonswap spectra.   The fitted sea-state parameter set and the fitted sea spectrum type are defined at which the minimum summation of the squares of the errors defined in Equation 3-19 is produced.

	$T_1 = T_{min}$	T_2	..	$T_q = T_m$	..	$T_k = T_{max}$
$H_{s1} = H_{smin}$	$error_{11}$	$error_{12}$				$error_{1k}$
H_{s2}	$error_{21}$					
:					*minimum error*	
$H_{sp} = H_s$				$error_{pq}$		
:						
$H_{sn} = H_{smax}$	$error_{n1}$					$error_{nk}$

Figure 3-6. The best fit of sea-state parameters set defined by lease squares principle.

3.1.9. Comparison of Several Approaches for Modal Period Estimation

The modal period T_m, or the spectral peak frequency $2\pi / T_m$, is estimated in this study from the time history by using zero-upcrossing period expressed as Equation 3-1. Some other approaches to estimate the spectral peak frequency are available, such as the simple maximum method, the Delft method and the weighted mean method.   These

methods estimate the spectral peak frequency in the spectral domains.   To assess the

reasonableness of estimated modal period in this study, a comparison of estimated modal

periods using Equation 3-1, the simple maximum method, the Delft method and the

weighted mean method is presented in Table 3-1.   The original modal periods T_m used to

simulate the time histories for the three buoys are also shown in Table 3-1 for reference

and for computing the estimation relative errors.

Table 3-1. Several approaches to estimate modal periods of buoys 1, 2, and 3 with

relative errors to the original modal periods shown in parenthesis.

Estimation method for the modal period T_m					
Buoy	Original	Equation 3-1	Simple maximum	Delft method	Weighted mean
	T_m	$T_m = 1.41\ T_z$	$T_m = 2\pi/f_p^{max}$	$T_m = 2\pi/f_p^{D60}$	$T_m = 2\pi/f_p^{M4}$
1	7 sec	7.14 sec (2.04 %)	7.21 sec (3.06 %)	7.10 sec (1.42 %)	6.79 sec (-2.95 %)
2	6 sec	6.34 sec (5.74 %)	6.31 sec (5.21 %)	6.23 sec (3.91 %)	6.09 sec (1.57 %)
3	8 sec	7.95 sec (-0.57 %)	7.77 sec (-2.88 %)	7.66 sec (-4.23 %)	7.46 sec (-0.84 %)

The simple maximum method is a straightforward method which determines the

spectral peak frequency by simply selecting the frequency associated with the maximum

spectral ordinate, denoted as f_p^{max} (IAHR 1989; Young 1995).   The Delft method for

determining the spectral peak frequency is to find the centroid of the spectral band

between the lower and upper spectral densities which is the 80% or 60% of the maximum

spectral ordinate.   The lower and upper frequency thresholds are denoted as f_1 and f_2 and

the estimated spectral peak frequency is expressed as (IAHR, 1989; Young, 1995):

$$f_p^{Dm} = \frac{\int_{f_1}^{f_2} fF(f)df}{\int_{f_1}^{f_2} F(f)df} \qquad (3\text{-}20)$$

in which m is 80 or 60 meaning that 80% or 60% maximum spectral ordinate is used for

the estimation.   Young (1995) indicated that f_p^{D60} is a better estimation than f_p^{D80}, which

also observed in this study.   Therefore, the estimation results of f_p^{D60} are presented in

Table 3-1.   The weighted mean method estimates the spectral peak frequency by

applying a weighting exponent on the spectral densities and evaluating the spectral peak

frequency as follows (Sobey and Young, 1986; Young 1995):

$$f_p^{Mq} = \frac{\int fF^q(f)df}{\int F^q(f)df} \qquad (3\text{-}21)$$

This approach uses the entire spectrum instead of a portion between frequency thresholds

to estimate the spectral peak frequency.   Several choices of weighting exponent q have

been suggested.   The choice of $q = 4$ giving the spectral peak frequency as f_p^{M4} was

recommended by Young(1995) and is included in the comparison in Table 3-1.   Note

that the estimation comparison shown in Table 3-1 uses adjusted periodograms which

described in details in Section 2.4.3.   Comparisons show that these methods produce

estimations within 5.7% relative errors of the original modal periods used to simulate the

time histories.   The approach used in this study produced the best estimation for buoy 3

while the Delft method shows the best estimation on buoy 1 and the weighted mean

method shows the best results on buoy 2.   Overall, it is reasonable to use Equation 3-1

for the modal period estimation since the estimation relative errors are not significant, i.e.

within 5.7 %.

3.2. Verification for Spatial Data Interpolation using Numerical Wave Model SWAN (Simulating WAves Nearshore)

In order to verify the spatial interpolation and prediction procedure provided in Section 3.1, a numerical wave model SWAN (Simulating WAves Nearshore) is utilized. As shown in Figure 3-7, the results from the SWAN model play the role of providing verification for the data interpolation as well as the statistical characterization procedures. Similar to Figure 3-1, Figure 3-7 contains the same characterization and estimation procedures except the inputs at the beginning are the SWAN generated wave properties. These properties such as the modal period and the significant wave height are used to simulate the vertical elevation time histories for selected locations considered as the buoys to start the estimation process. The parameters estimation results at the end of Figure 3-7 are taken to compare with the SWAN generated wave properties which are the inputs from the beginning. The verification of the methodology demonstrated in Section 3.1 is taken place by the comparison of the parameters estimations and the wave properties generated by SWAN. Illustrative numerical examples will be provided in Chapter 4. An introduction of the SWAN model applications and commends used in this study is provided in this section.

The numerical wave model SWAN (Simulating WAves Nearshore) is software that developed at the Delft University of Technology and can be downloaded at www.swan.tudelft.nl. This model is for the simulation of waves in waters of deep, intermediate and finite depth. It accounts for the physics such as wave propagation, wave generation by wind, wave interactions, whitecapping, bottom friction, depth-induced breaking, dissipation due to vegetation, diffraction, transmission through

and reflection against obstacles and propagation from laboratory up to global scales. Outputs provided by the SWAN model include one-dimensional and two-dimensional spectra, significant wave height, wave period, average wave direction and directional spreading, diffraction parameter, dissipation, etc.

First step in SWAN commands is to define the simulation mode as stationary mode or non-stationary mode and also define the simulation would be one-dimensional or two-dimensional. Two-dimensional stationary mode is the default mode and is used in this study. Second step is to determine the coordinate in either Cartesian or spherical coordinates. Then, the computational grid and input grid should be determined as the next step. The computational grid (CGRID) defines the geographic computation range, how the computation range is meshed, the frequency range and the number of frequencies used in the calculation. In this study, the regular rectangular computational grid is used. The computational range is between 0 m to 1200 m in both x-coordinate and y-coordinate and the number of meshes is 50 in both directions. That is, there are 51 points from 0 m to 1200 m and the interval is 24 m.

The next step is to define the input grid which may provide the water level, current, bottom and friction at the grid points. In this study, the bottom grid which defines the bottom level is used, which has the command as INPGRID BOTTOM. The bottom grid has a origin of (0m, 0m). The mesh size in this study is 50 m and the number of meshes is 24 in both x-direction and y-direction, which make the bottom grid range from 0 m to 1200 m in both x and y directions. The bottom levels are defined in a text file that is read by the command READINP BOTTOM. The other command for the input field is the wind effect which is not considered in this study.

71
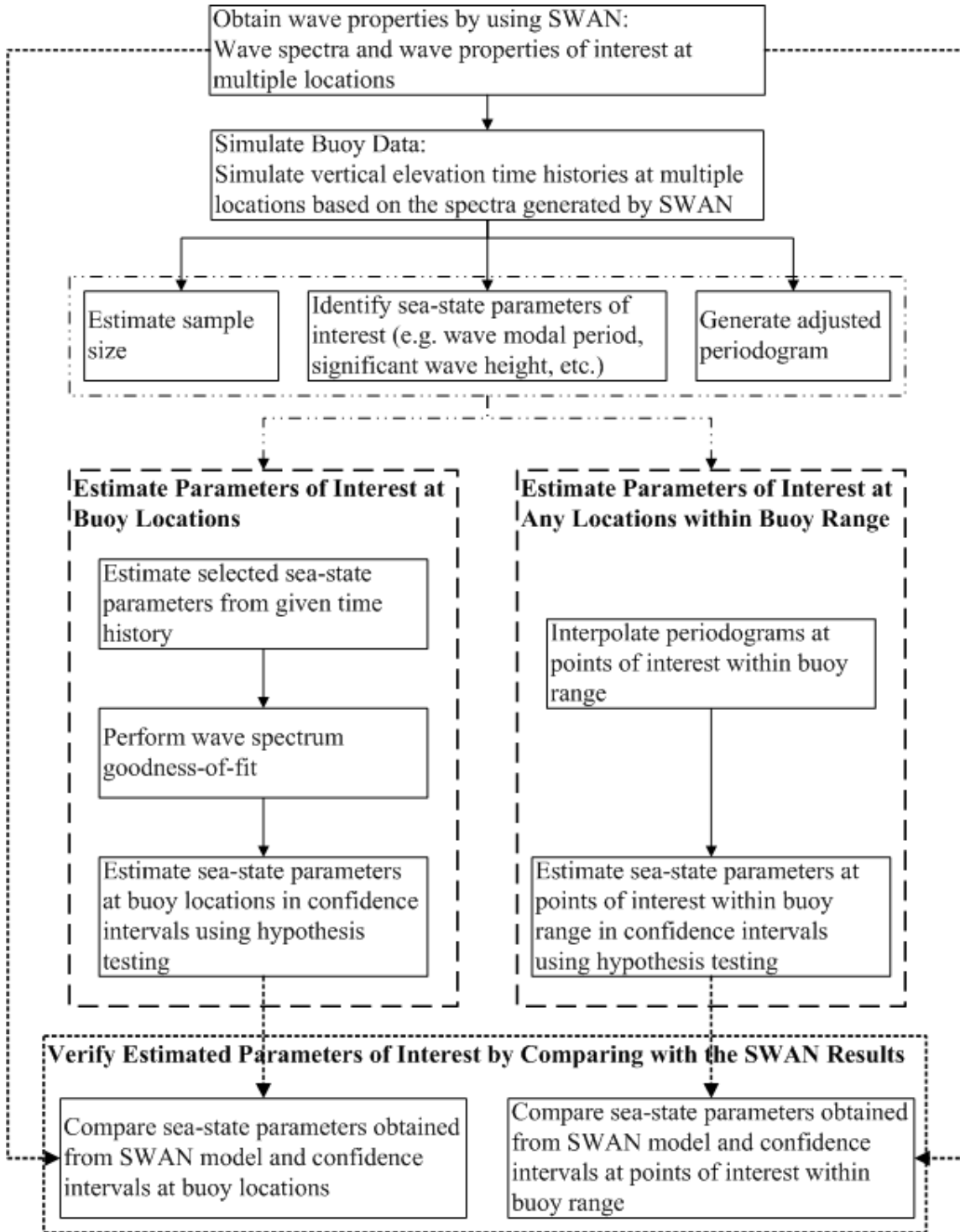
Figure 3-7. Statistical characterization for predicting sea-state condition based on SWAN generated wave spectra and verification using SWAN results.

As the next step, the initial condition can be given as a wave spectrum at the edge or the corner of the computational grid by giving the spectral densities in a text file or by determining the significant wave height, wave period and wave spectrum type. In this study, default wave spectrum type, the Jonswap wave spectrum, is used. The illustrative example in this study defined the initial condition at the edge shown in Figure 3-8 as a Jonswap spectrum with the significant wave height of 3.2 m and the wave modal period of 8.3 sec. Figure 3-9 shows the wave direction result in the computational range based on the input wave spectrum shown in Figure 3-8 and the bottom condition defined earlier in a text file and read by the command READINP BOTTOM.

Figure 3-10 represents one of the wave simulation output quantities, the significant wave height in meters, in a three-dimensional plot in the computational range. Figure 3-11 shows the locations of interested to obtain outputs which will be used for estimations illustrated in Figure 3-7. The locations, denoted as Loc in Figure 3-11, are defined in Cartesian coordinate in a text file and read by the command POINTS. The wave spectra and other wave properties of interest will be provided in the output files at the defined locations of interest. Figure 3-12 shows the outcomes of significant wave heights at locations defined in Figure 3-11 based on the computational grid, the input bottom levels and the initial condition defined in Figure 3-8.

The computed output quantities can be written in text files by requests, such as requesting a spectral output by the command SPEC and requesting a table output by the command TABLE. The spectral output file includes the locations of interest defined, the frequency discretization points, the variance densities, the wave direction and the directional spreading for each location. Figure 3-13 shows the variance densities

73

provided in the spectral output file for location 3 defined in Figure 3-11. The table

outputs shows the outputs quantities of interest of the points defined in Figure 3-11 in a

table. The output quantities could be the significant wave heights, the modal periods,

the wave direction, etc.

There is the other command called BLOCK provides the output quantities of

interest. The command BLOCK puts the outputs in a ".mat" file which can be read by

Matlab. Noye that the ".mat" file has the output quantities at computational grid points

instead of the interested locations defined earlier using the command POINTS. Figure

3-9 and Figure 3-10 are the plots of the quantities requested by the command BLOCK

which provides the outputs in the ".mat" file. Figure 3-14 is an example SWAN

command file which shows the commands described above. An example output file is

shown in Appendix A.
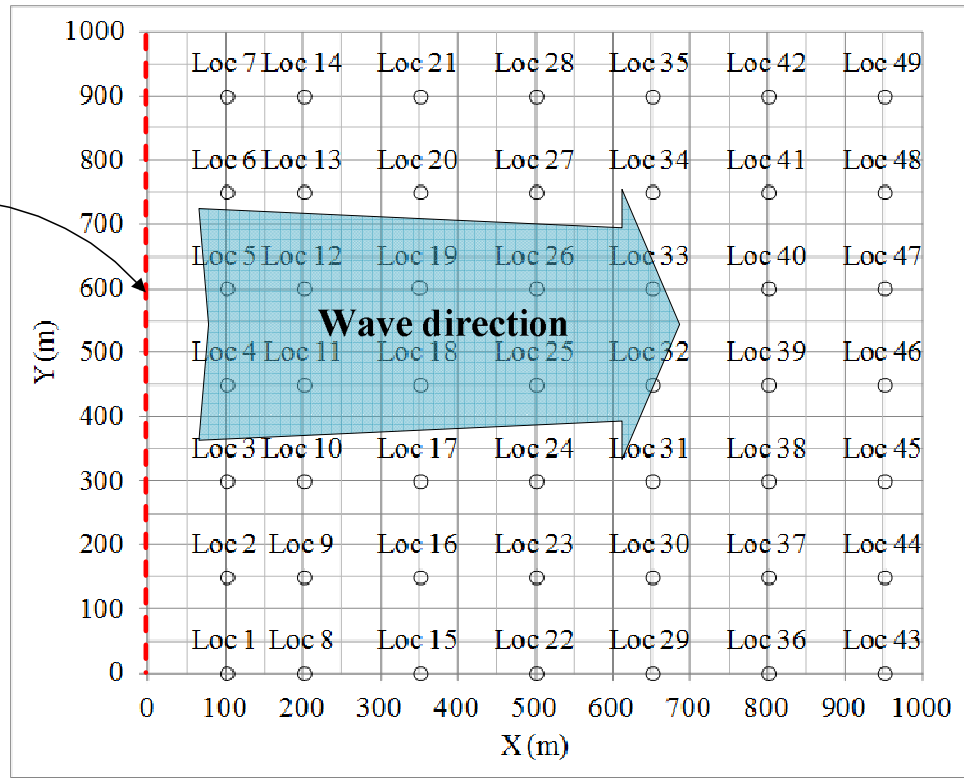
Figure 3-8. Initial conditions given in the wave simulating model SWAN in this study.
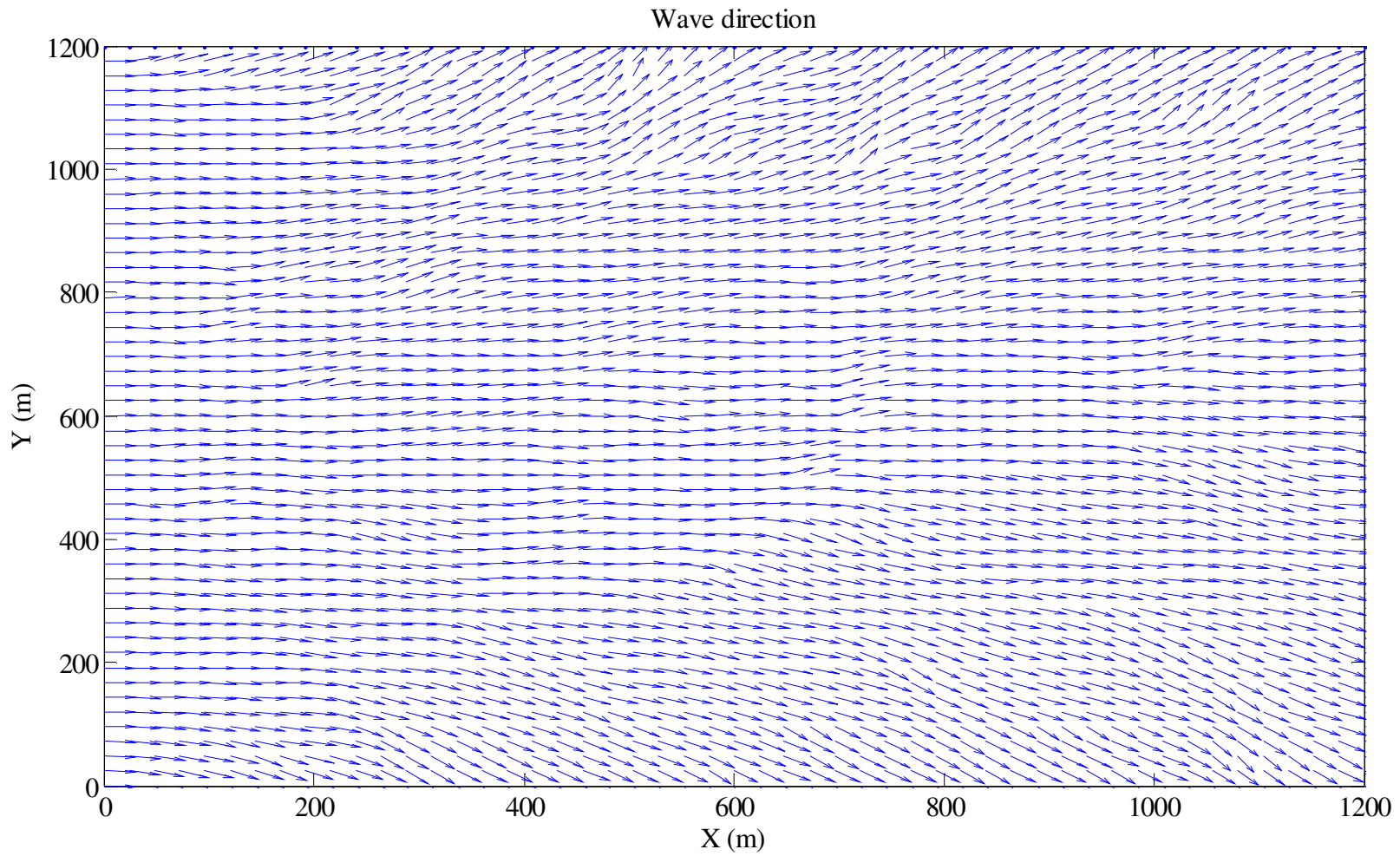
Figure 3-9. Wave direction of SWAN generated waves in the computational range.
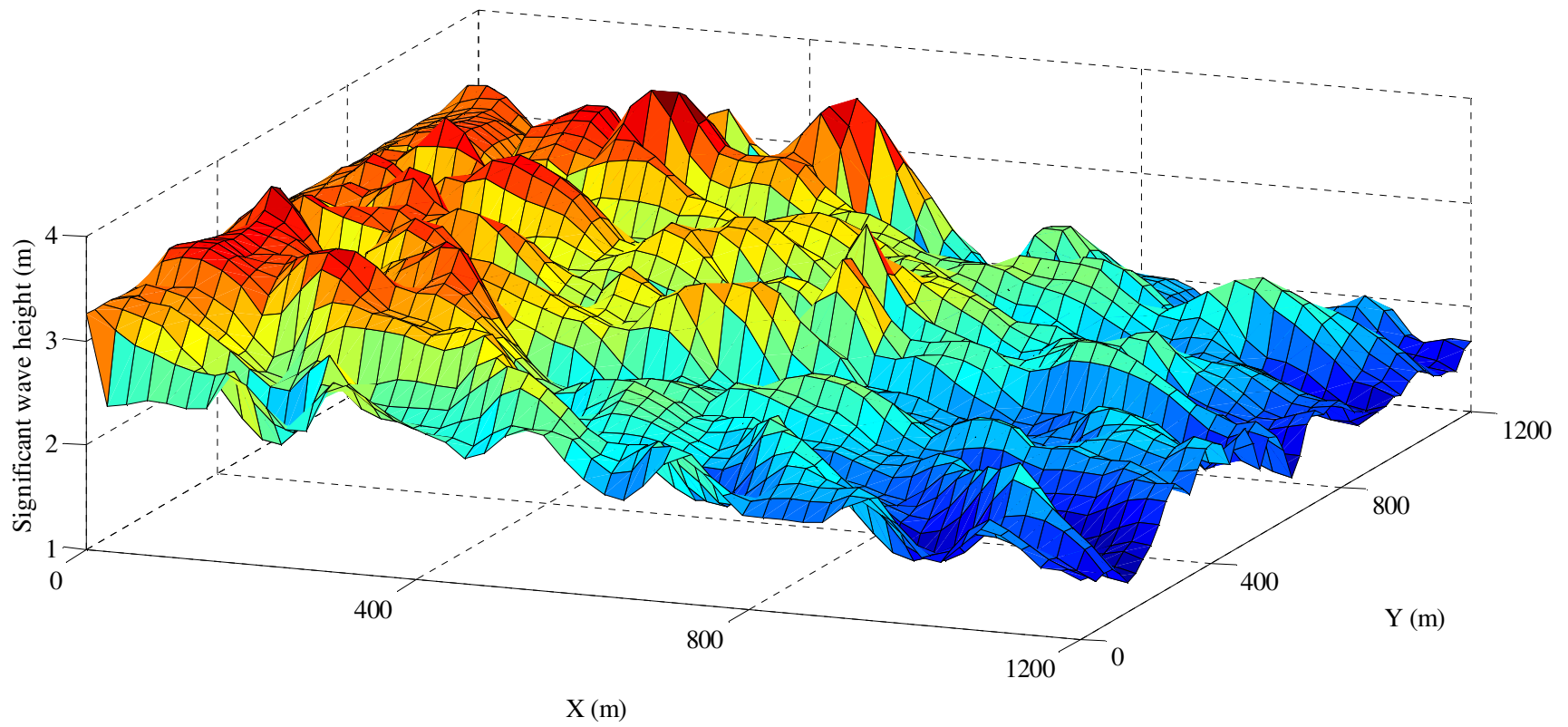
Figure 3-10. Significant wave heights computed by SWAN in the computational range.
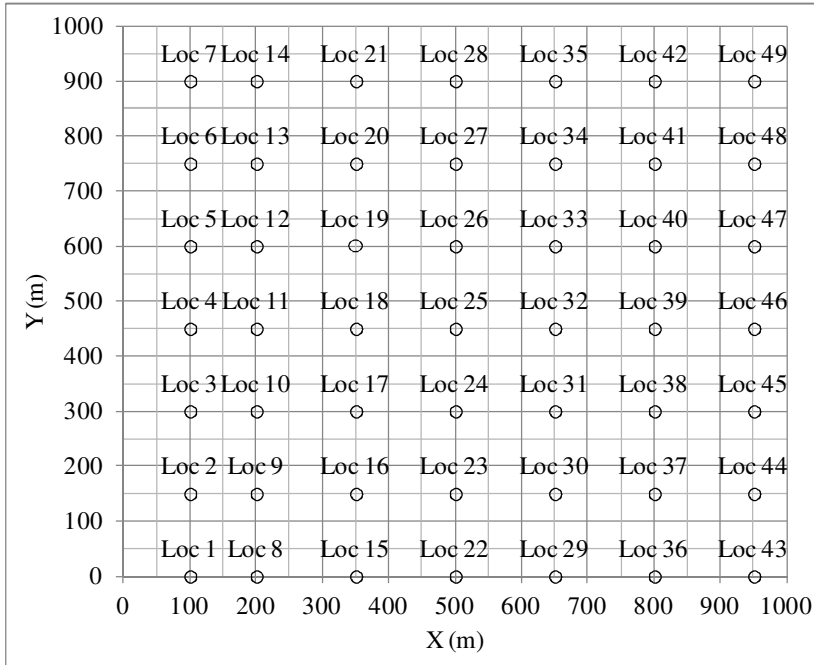
Figure 3-11. Locations of interest to obtain wave properties in this study.
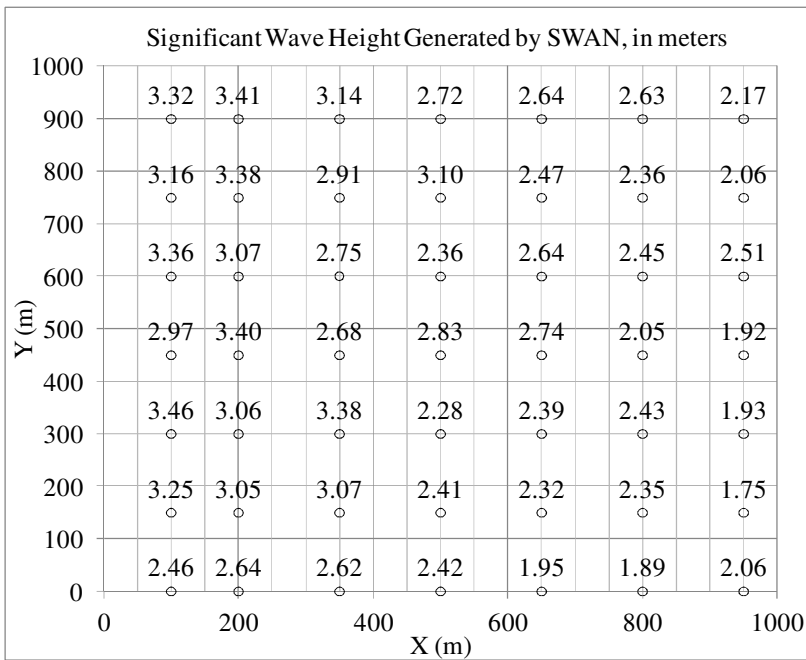
Figure 3-12. Significant wave heights generated using SWAN at locations of interest

defined in Figure 3-11.
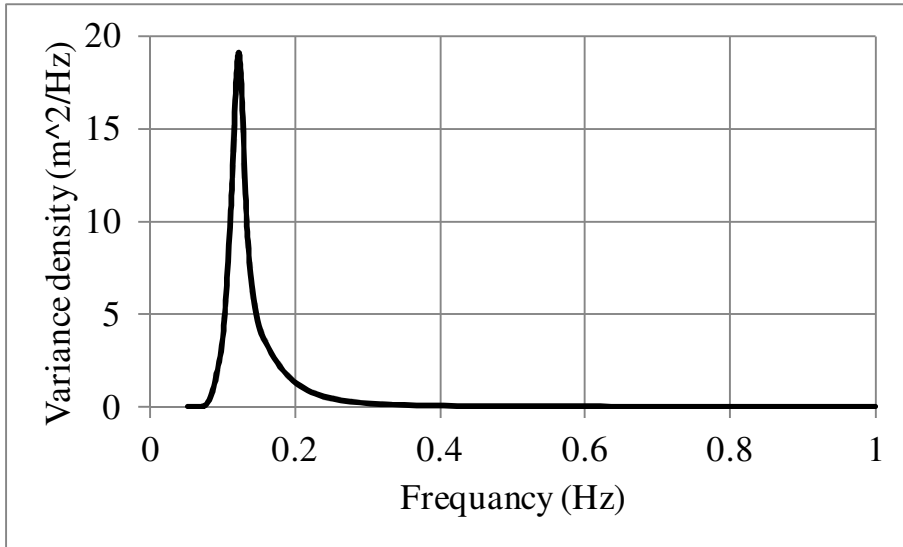
Figure 3-13. Generated variance densities at location 3.

```
$**************************HEADING***********************
PROJ 't5Trial1' 't5'
$********************MODEL INPUT***********************
CGRID 0 0. 0. 1200. 1200. 50 50 CIRCLE 36 0.0521 1. 31
INPGRID BOTTOM 0. 0. 0. 24 24 50. 50. EXC -99.
READINP BOTTOM 1. 'Trial1.bot' 3 0 FREE
WIND 0 0
BOUN SHAP JON 3.3 PEAK DSPR POWER
BOU SIDE W CCW CON PAR 3.2 8.3 0 2
GEN3
$*********************************************************
POINTS 'buoy' FILE 't5Trial1.loc'
TABLE  'buoy' HEAD   't5Trial1.tbl' HS RTP TPS FSPR
SPEC   'buoy' SPEC1D 't5Trial1.spc'
BLOCK  'COMPGRID' NOHEAD 't5Trial1.mat' LAYOUT 3 HS TPS FSPR
$
TEST 1,0
COMPUTE
STOP
$
```

Figure 3-14. Example of SWAN command file.

4. Case Studies

4.1. Sea-state Characterization using Simulated Buoy Data

A sea-state characterization example is presented in this section using simulated buoy vertical elevation time histories. Figure 3-1 shows the buoy data analysis procedure in a flowchart. The buoy vertical elevation is taken as the sea wave surface elevation simulated from Bretschneider wave spectrum in this study. Two sea-state parameters, sea wave modal period T_m and significant wave height H_s, are selected to describe the wave characteristics in the time domain. Upon estimating the two parameters using Equations 3-1 and 3-2 from the time history, the sea spectrum goodness-of-fit is performed on each buoy by applying statistical hypothesis testing on selected sea spectra types and on the periodogram which demonstrates the wave characteristics in the spectral domains. Bretschneider and Jonswap spectra are selected for the sea spectrum goodness-of-fit and are conducted using the sea-state parameters, the wave modal period T_{mt} and the significant wave height H_{st}, estimated from the time history by Equations 3-1 and 3-2.

The periodogram is constructed from the time history by using finite Fourier transform on the auto-covariance function of the time history defined in Equation 2-15. For the purpose of comparing the periodogram and sea spectra, the periodogram is adjusted to have the peak at the modal frequency $2\pi / T_{mt}$, the unit the same as the sea spectrum and the area under the periodogram the same as the variance of the time history, demonstrated in Section 3.1.3. The sea spectrum goodness-of-fit is performed by fitting the Bretschneider and Jonswap spectra to the adjusted periodogram. Once the fitted

spectrum type is defined, the confidence intervals of the selected sea-state parameters can be analyzed for sea-state characterization by applying statistical hypothesis testing on the buoy adjusted periodogram and the adjusted periodograms constructed from the fitted spectrum type for a range of sea-state parameters sets. Details are demonstrated in Section 3.1.6.

In addition, the buoys data are used for sea-state prediction of the points of interest such as the travel track points of seagoing vessels surrounded by the buoys. The vessels travel track points are arbitrarily chosen, and sea-state characteristics at these points are estimated as intermediate values among surrounding buoys by applying inverse distance weight factors on the adjusted periodograms of these buoys. Section 3.1.8 provides the detail procedure. The confidence intervals of the selected sea-state parameters for the travel track points are then analyzed using the same method used to analyze the buoy data.

4.1.1. Description of Simulated Buoy Data

The sea wave surface elevation time histories of three buoys are used in this example. Table 2-2 and Figure 4-1 summarize the locations, the modal periods and the significant wave heights of these three buoys. The time histories are generated using Bretschneider wave spectra with the modal periods and significant wave heights from Table 2-2. An example spectrum is shown in Figure 4-2. The duration of these buoy time histories is 1500 sec starting at 0.5 sec with a constant interval of 0.5 sec. The number of total data points of each buoy time history is 3000. Figure 4-3 shows an example of simulated time history.
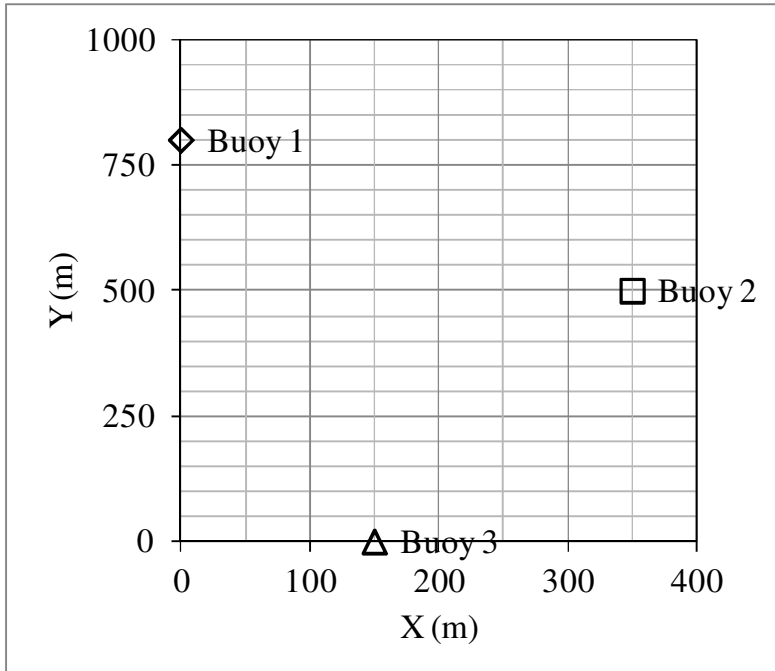
Figure 4-1. Locations of buoys 1, 2 and 3.
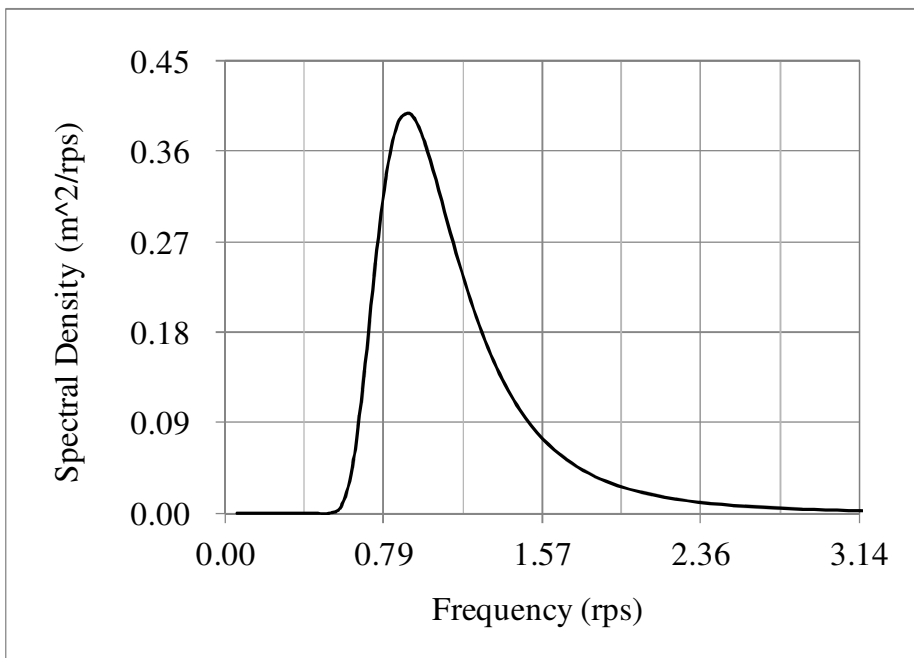
Figure 4-2. Wave spectrum.
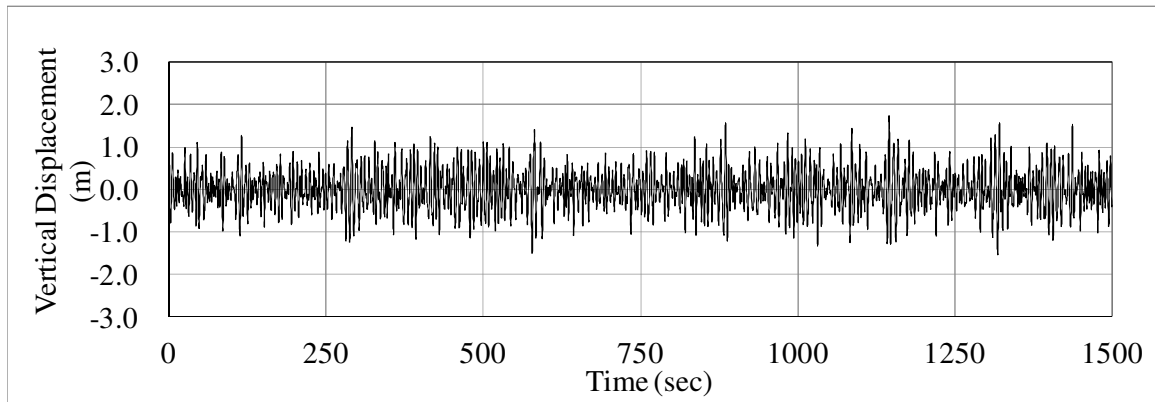
Figure 4-3. Simulated time history.

4.1.2. Parametric Analysis

The selected two sea-state parameters, the wave modal period T_m and the significant wave height H_s, are analyzed in this section. The confidence interval estimations, described in Section 3.1.6, are performed by fixing the value of one parameter when analyzing the other parameter within a range of values. Table 4-1 shows the estimation results of the confidence intervals at the 95% level for the three buoys on the modal period T_m for the significant wave height $H_s = H_{st}$ and on the significant wave height H_s for the modal period $T_m = T_{mt}$. Figure 4-4 and Figure 4-5 show the confidence interval estimations at the 95% level for the modal period and the significant wave height, respectively, for buoy 1; while Figure 4-6 and Figure 4-7 are for buoy 2 and Figure 4-8 and Figure 4-9 are for buoy 3.

Table 4-1. Two-sided confidence intervals at the 95% level for the modal period T_m and the significant wave height H_s for buoys 1, 2 and 3.

Two-sided confidence intervals at the 95% level	Buoy 1 T_{mt} =7.14 sec H_{st} =1.99 m	Buoy 2 T_{mt} =6.34 sec H_{st} =0.97 m	Buoy 3 T_{mt} =7.95 sec H_{st} =1.54 m
Lower modal period limit (T_L)	6.99 sec	6.24 sec	7.76 sec
Upper modal period limit (T_U)	7.83 sec	6.84 sec	8.76 sec
Lower significant wave height limit (H_{sL})	1.96 m	0.96 m	1.52 m
Upper significant wave height limit (H_{sU})	2.03 m	0.99 m	1.57 m
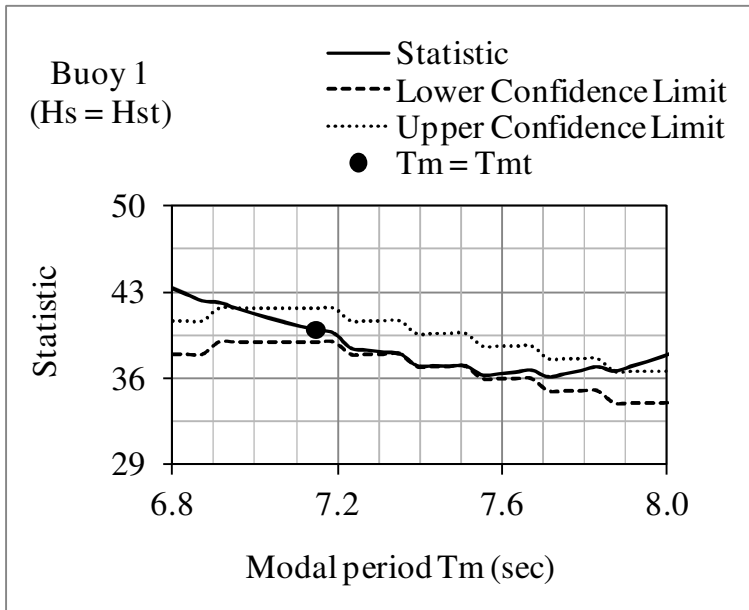
Figure 4-4. Two-sided confidence intervals at the 95% level of buoy 1 on the modal period T_m for the significant wave height $H_s = H_{st}$.
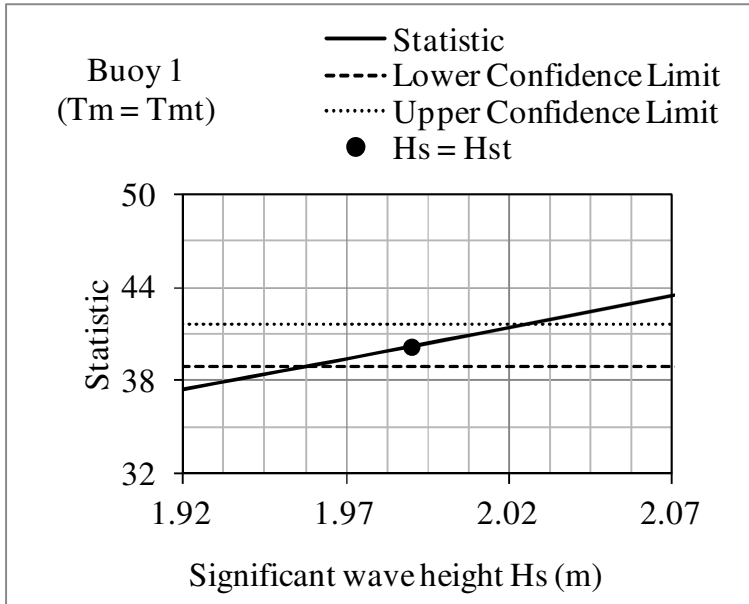
Figure 4-5. Two-sided confidence intervals at the 95% level of buoy 1 on the significant wave height H_s for the modal period $T_m = T_{mt}$.
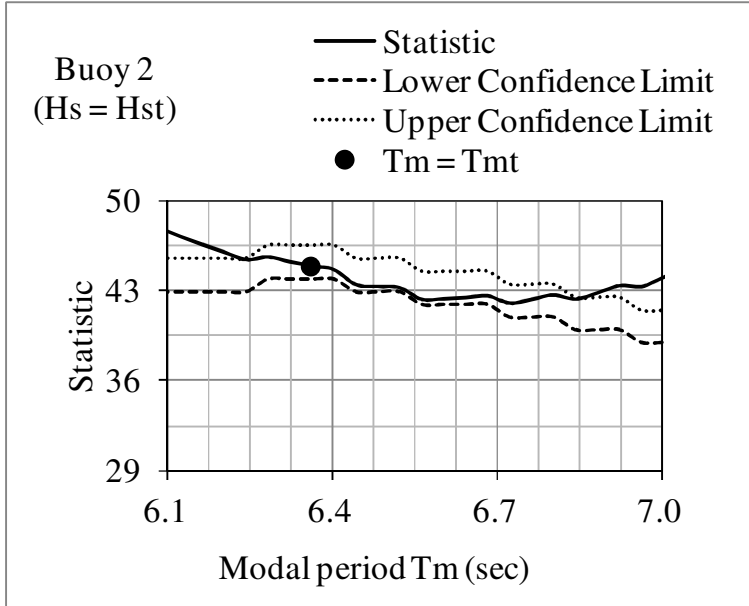
Figure 4-6. Two-sided confidence intervals at the 95% level of buoy 2 on the modal period T_m for the significant wave height $H_s = H_{st}$.

Figure 4-7. Two-sided confidence intervals at the 95% level of buoy 2 on the significant wave height H_s for the modal period $T_m = T_{mt}$.

Figure 4-8. Two-sided confidence intervals at the 95% level of buoy 3 on the modal period T_m for the significant wave height $H_s = H_{st}$.

Figure 4-9. Two-sided confidence intervals at the 95% level of buoy 3 on the significant wave height H_s for the modal period $T_m = T_{mt}$.

4.1.3. Data Interpolation

The points of the travel track of seagoing vessels can be treated as intermediate values among surrounding buoys, and the properties of the travel track points can be estimated from the properties of the buoys. Figure 4-10 and Table 4-2 show the locations of three track points of interest randomly selected for buoy data interpolation. According to Section 3.1.8, the inverse distance weight factors applied on buoys for each track point are computed using Equations 3-15 to 3-17and summarized in Table 4-3. These weight factors are applied on the adjusted periodograms of the buoys using Equation 3-18 to estimate the periodograms for the track points.
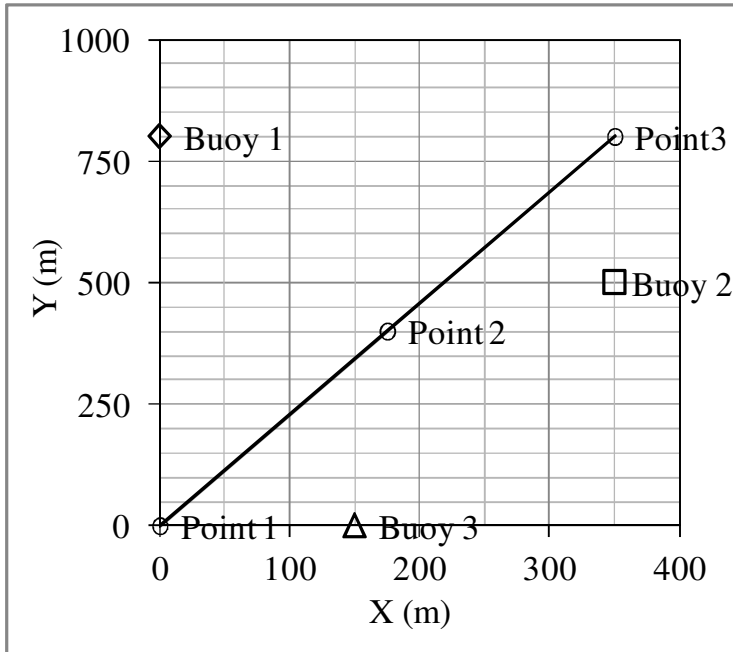
Figure 4-10. Locations of track points of interest.

Table 4-2. Locations of track points 1, 2, and 3.

Track Point	x-coordinate	y-coordinate
1	0 m	0 m
2	175 m	400 m
3	350 m	800 m

Table 4-3. Weight factors applying on the three buoys for track points 1, 2, and 3.

Track Point	Point 1	Point 2	Point 3
Weight Factor applied to Buoy 1	0.130820	0.234987	0.385936
Weight Factor applied to Buoy 2	0.171475	0.509022	0.450259
Weight Factor applied to Buoy 3	0.697705	0.255992	0.163806

The modal period, the significant wave height and the spectrum type for the track

points can be estimated according to Section 3.1.8.   Two sea-state parameters, the modal

period T_{mt} and the significant wave height H_{st}, estimated using Equations 3-1 and 3-2

from the time history for each buoy are taken as reference values to select a range of

sea-state parameters sets.   The ranges of ($T_{mt, \min}$, $T_{mt, \max}$) and ($H_{st, \min}$, $H_{st, \max}$) are used

to generate sea spectra, in which $T_{mt, \min}$ is the minimum T_{mt} among the buoys, $T_{mt, \max}$ is

the maximum T_{mt} among the buoys, and $H_{st, \min}$ and $H_{st, \max}$ are similar to the T_{mt} case.

The best set of sea-state parameters is determined by Equation 3-19 using the concept of

lease squares applied on the periodograms of the track points and the adjusted

periodograms generated from the sea spectra for various sets of parameters in the range

defined previously.   The estimated modal period, denoted as T_{me}, and the estimated

significant wave height, denoted as H_{se}, for the track points are shown in Table 4-4.

Figure 4-11 shows the estimated periodogram and the fitted adjusted periodograms

constructed from the Bretschneider and Jonswap sea spectra using the modal period T_{me}

and the significant wave height H_{se} for track point 1.   The results show that the

Bretschneider spectrum is a better fit compared with the Jonswap spectrum.   Figure 4-12

and Figure 4-13 are the results for track points 2 and 3, respectively.

   The confidence intervals analysis for the track points follows the same method

used to analyze the buoy data by applying statistical hypothesis testing on the

periodogram of the track point and the adjusted periodograms constructed from the fitted

spectrum type for a range of sea-state parameters sets.   Table 4-4 show the confidence

intervals at the 95% level for these three track points on the modal period T_m for the

significant wave height $H_s = H_{se}$ and on the significant wave height H_s for the modal

period $T_m = T_{me}$.   Figure 4-14 and Figure 4-15 show the confidence interval estimations

at the 95% level for the modal period and the significant wave height, respectively, for

track point 1; while Figure 4-16 and Figure 4-17 are for track point 2 and Figure 4-18 and

Figure 4-19 are for track point 3.

Table 4-4. Two-sided confidence intervals at the 95% level of the significant wave height

H_s and the modal period T_m of track points 1, 2, and 3.

Two-sided confidence intervals at the 95% level	Point 1 T_{me} =7.20 sec H_{se} =1.52 m	Point 2 T_{me} =6.83 sec H_{se} =1.40 m	Point 3 T_{me} =6.83 sec H_{se} =1.52 m
Lower modal period limit (T_L)	7.15 sec	6.80 sec	6.70 sec
Upper modal period limit (T_U)	7.58 sec	6.98 sec	6.99 sec
Lower significant wave height limit (H_{sL})	1.45 m	1.37 m	1.50 m
Upper significant wave height limit (H_{sU})	1.53 m	1.43 m	1.54 m
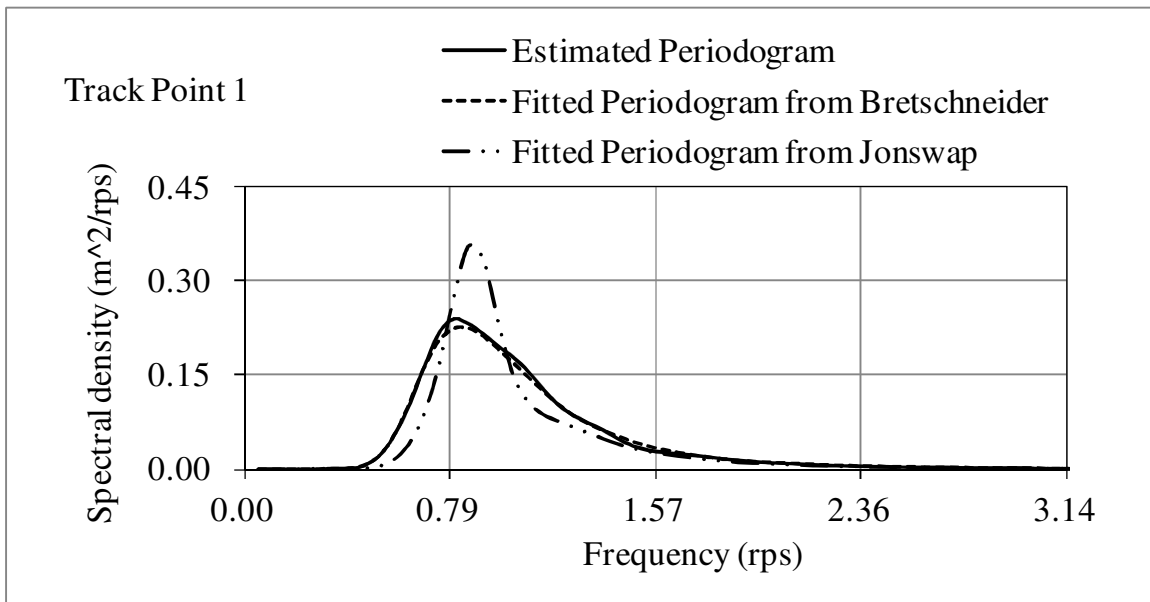
Figure 4-11. Estimated periodogram and fitted periodograms of different sea spectra for

track point 1.

90
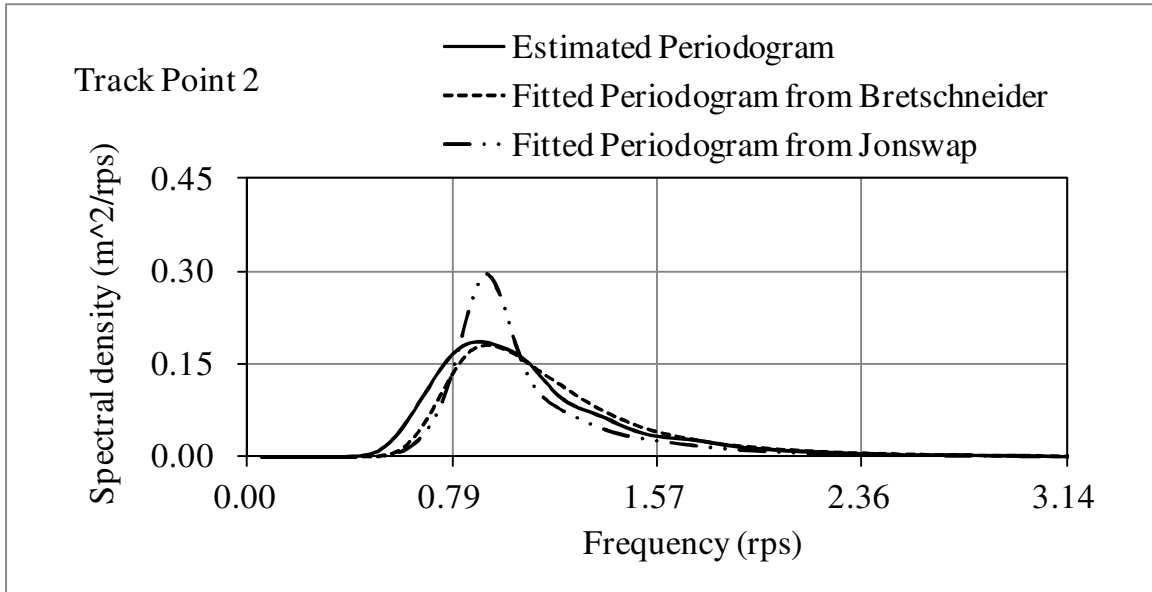
Figure 4-12. Estimated periodogram and fitted periodograms of different sea spectra for track point 2.
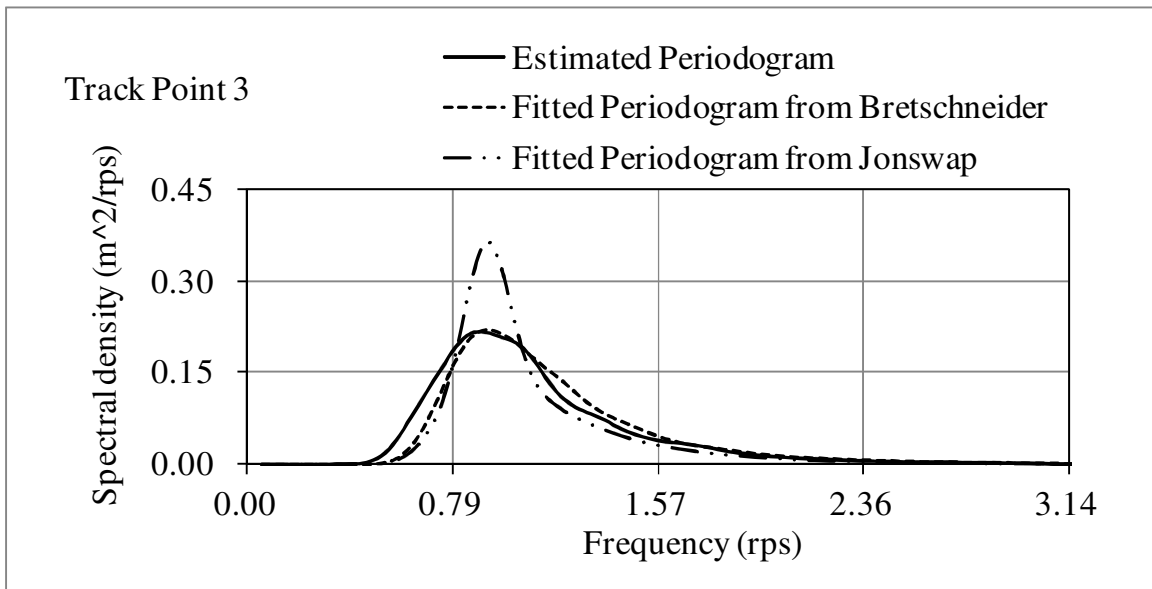
Figure 4-13. Estimated periodogram and fitted periodograms of different sea spectra for track point 3.
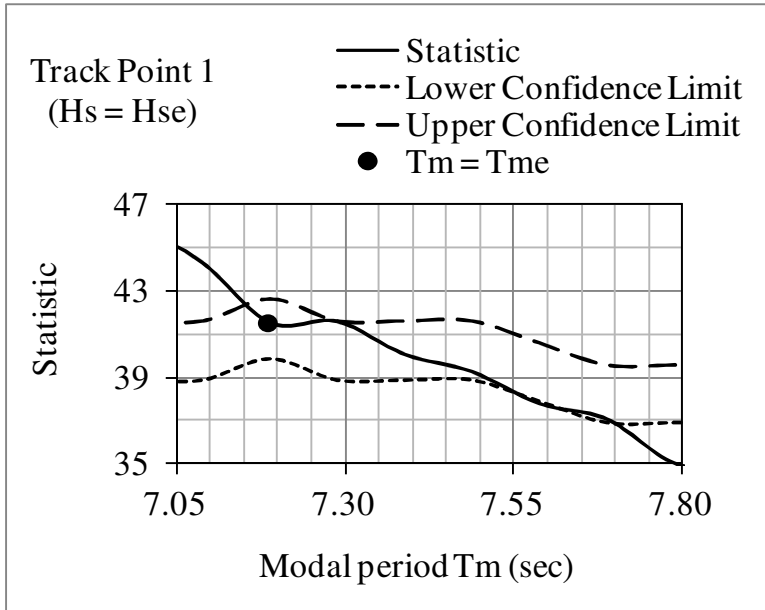
Figure 4-14. Two-sided confidence intervals at the 95% level of track point 1 on the modal period T_m for the significant wave height $H_s = H_{se}$.
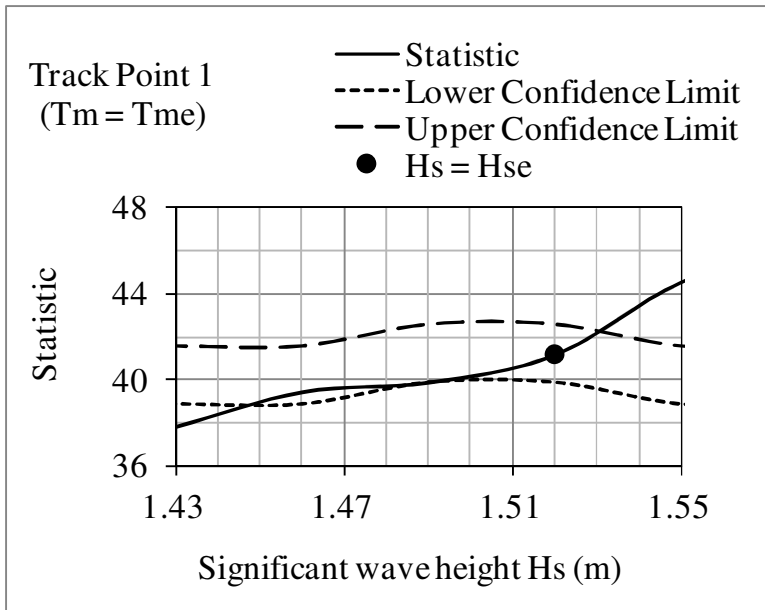
Figure 4-15. Two-sided confidence intervals at the 95% level of track point 1 on the significant wave height H_s for the modal period $T_m = T_{me}$.
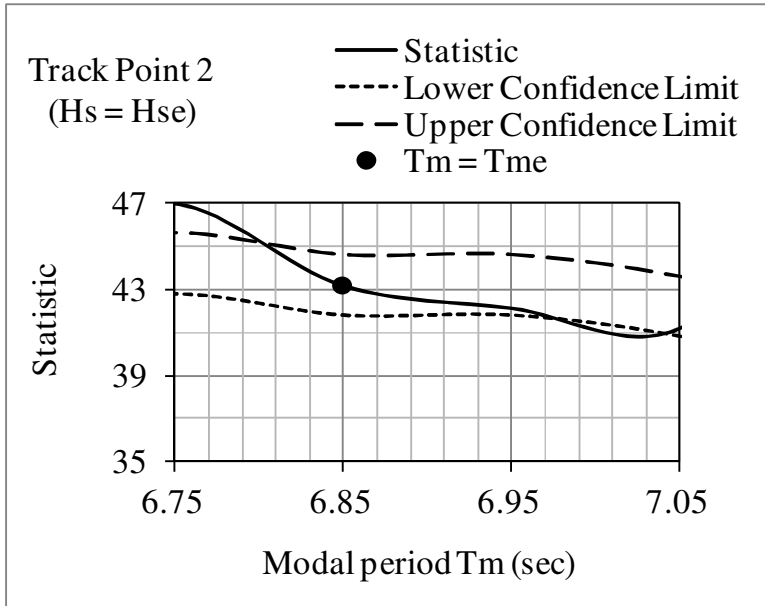
Figure 4-16. Two-sided confidence intervals at the 95% level of track point 2 on the modal period T_m for the significant wave height $H_s = H_{se}$.
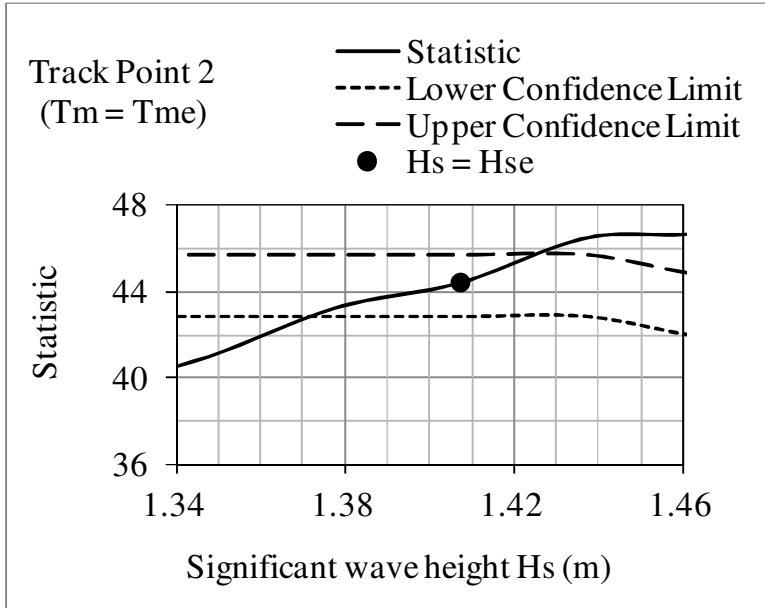
Figure 4-17. Two-sided confidence intervals at the 95% level of track point 2 on the significant wave height H_s for the modal period $T_m = T_{me}$.
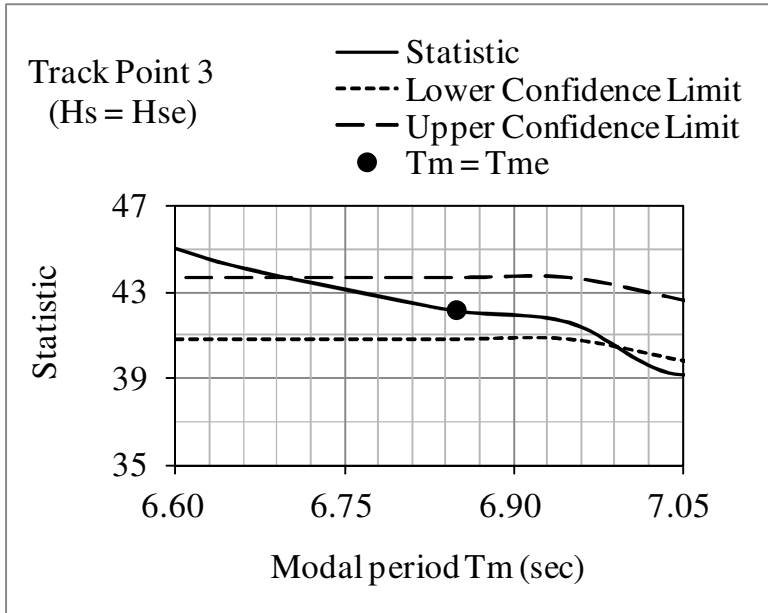
Figure 4-18. Two-sided confidence intervals at the 95% level of track point 3 on the modal period T_m for the significant wave height $H_s = H_{se}$.
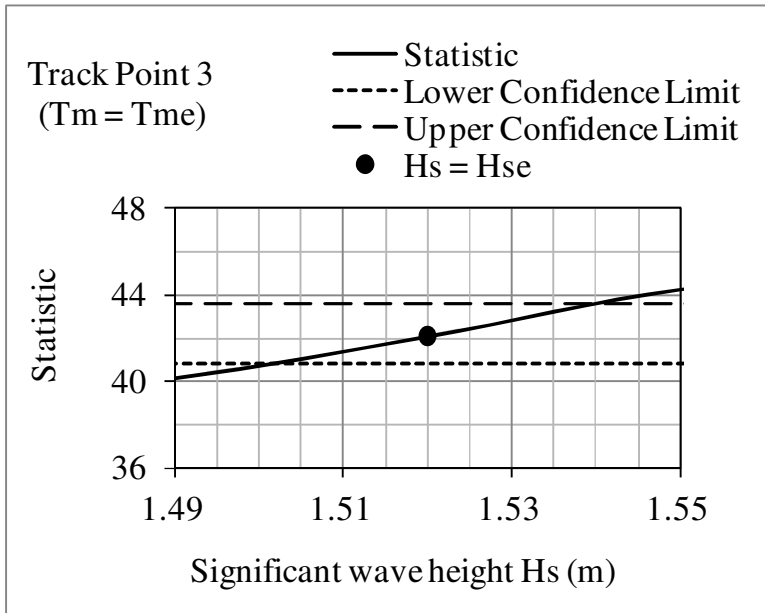
Figure 4-19. Two-sided confidence intervals at the 95% level of track point 3 on the significant wave height H_s for the modal period $T_m = T_{me}$.

This example illustrates the procedure of sea-state characterization and interpolation using spectral analysis concept described in Chapter 2 and the methodology introduced in Chapter 3. The prediction for unobserved locations is performed applying spatial interpolation on buoys, or observed locations, and the confidence intervals of the sea-state condition parameters are estimated for the buoys as well as the unobserved locations, or the points of interest. The statistical basis of the method enables the characterization to present the sampling variability and associated uncertainties by estimating the confidence intervals of the sea-state parameters. The estimation results show that the methodology proposed in Chapter 3 is able to provide accurate predictions by confidence intervals which cover the parameter values T_{me} and H_{se} estimated from the periodograms using least square concept demonstrated in Section 3.1.8.

4.2. Numerical Example and Verification using SWAN Generated Wave Data

This section provides sea-state characterization and interpolation as illustrated in Figure 3-7. The numerical wave model SWAN is utilized to generate the wave properties at locations defined in Figure 3-11. Six locations are selected from Figure 3-11 for this example to demonstrate and verify the methodology provided in Chapter 3. Three locations are arbitrary selected to be considered as the buoys and the other three locations within the buoys range are selected as the points on a travel track which need sea-state predictions. The sea-state parameters of these points of travel track will be interpolated based on the properties of the buoys. The wave properties generated by SWAN are considered as the true values. In other words, the methodology provided in

95

Chapter 3 is verified if the values generated by SWAN are within the confidence intervals of the parameters estimations.

4.2.1. Description of SWAN Data

Table 4-5 and Figure 4-20 summarize the coordinates for the six locations selected from Figure 3-11.   Figure 4-21 defines the locations considered as the three buoys and locations considered as the points of the travel track.

The initial and boundary conditions to generate the wave properties are shown in Figure 3-8.   Jonswap wave spectrum with the significant wave height 3.2 m and the modal period 8.3 sec is used as the initial condition.   Locations to obtain output quantities are defined in Figure 3-11, denoted as Loc 1, Loc 2, ..., Loc 49.   The output quantities such as the significant wave heights are shown in Figure 3-10 for the entire calculation range and in Figure 3-12 for the locations of interest.   The generated wave spectrum, for example at location 3, is shown in Figure 3-13.

The wave properties at the buoy locations defined in Figure 4-21 are provided in Table 4-6.   These properties are used to generate wave surface elevation time histories such as Figure 4-22 which is the time history for location 46, or buoy 3.   According to the procedure shown in Figure 3-7 and the details for each step in the procedure provided in Chapter 3, the significant wave height of the track points will be estimated using the buoy time histories generated using the given wave modal periods and the significant wave heights summarized in Table 4-6.   Sample size of independent observations and the adjust periodograms of the buoys are produced for the interpolation and estimation on the track points.

Table 4-5. Coordinates of locations selected from Figure 3-11 for estimation.

Location		x-coordinate	y-coordinate
Loc 2	(as Buoy 1)	100 m	150 m
Loc 7	(as Buoy 2)	100 m	900 m
Loc 47	(as Buoy 3)	950 m	450 m
Loc 5	(as Track Point 1)	100 m	600 m
Loc 19	(as Track Point 2)	350 m	600 m
Loc 33	(as Track Point 3)	650 m	600 m
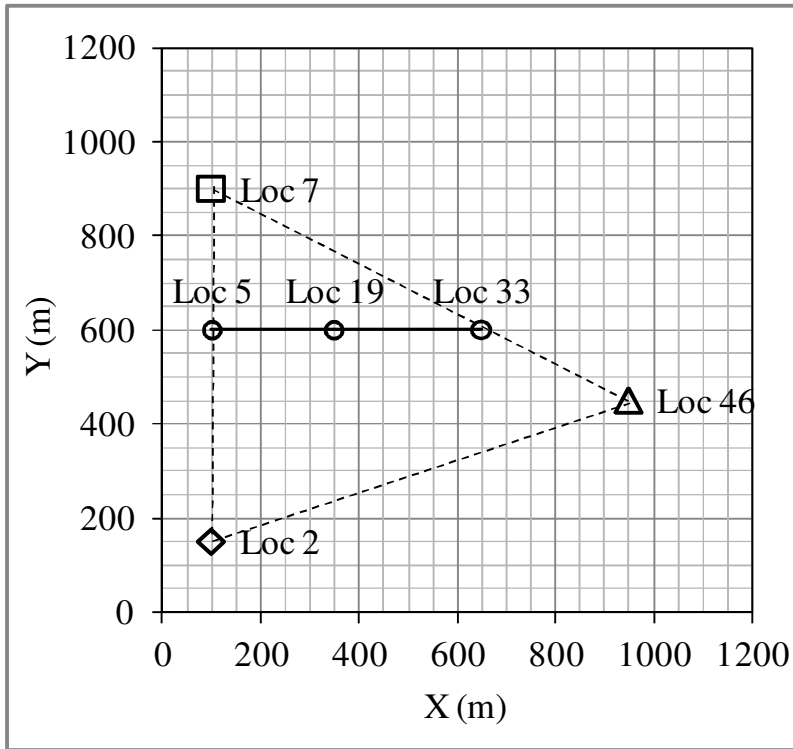
Figure 4-20. Locations of interest selected from Figure 3-11 for estimations.
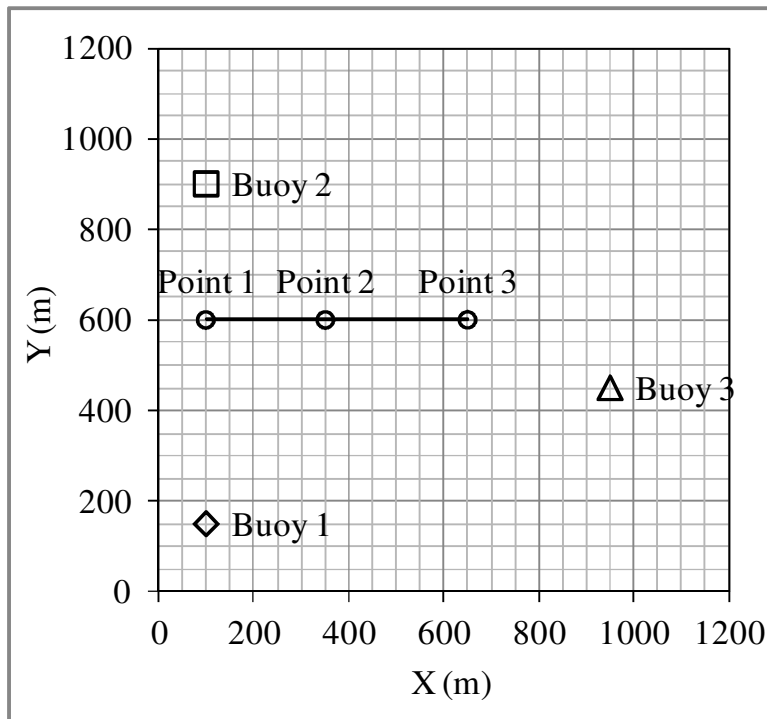
Figure 4-21. Determination of locations shown in Figure 4-20 selected as the buoys and

as the track points.

Table 4-6. Modal periods and significant wave heights of buoy locations defined in

Figure 4-21.

Location		Modal period	Significant wave height
Loc 2	(as Buoy 1)	8.15 sec	3.13 m
Loc 7	(as Buoy 2)	8.16 sec	3.20 m
Loc 46	(as Buoy 3)	8.14 sec	2.42 m
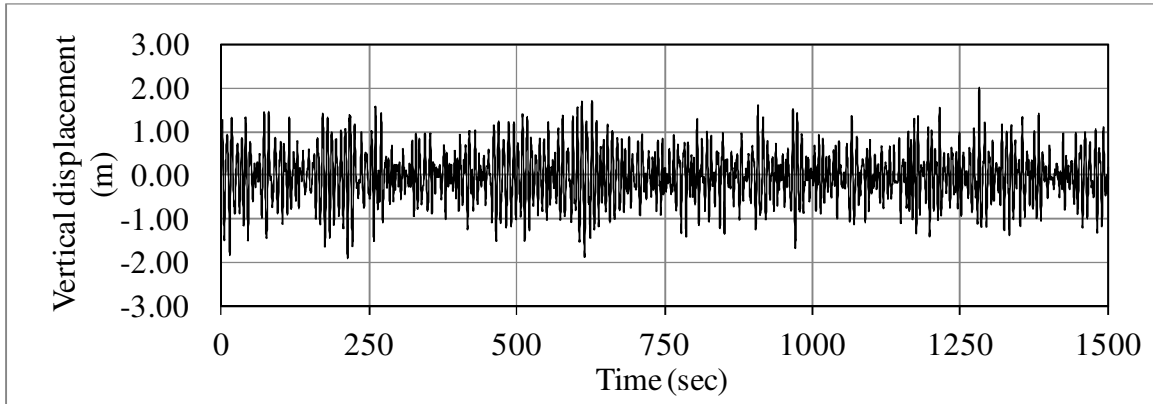
Figure 4-22. Generated time history for location 46 (or buoy 3) using the wave properties provided in Table 4-6.

4.2.2. Data Interpolation

Periodograms of the track points are interpolated from the adjust buoy periodograms using the inverse distance weight factors demonstrated in Section 3.1.8. By performing the wave spectrum goodness-of-fit, the spectrum type of the track points can be determined   Figure 4-23 show the estimated periodogram of track point 1 and the periodograms constructed from the Bretschneider and Jonswap spectra for the goodness-of-fit; while Figure 4-24 and Figure 4-25 are for the track points 2 and 3, respectively.   These three figures show that the Jonswap spectrum is a better fit for the three track points.   The results are reasonable since Jonswap spectrum is used as the initial condition for generating the wave properties in the estimation field.   The wave modal periods and the significant wave heights can be estimated using the least square concept and procedure demonstrated in Section 3.1.8.
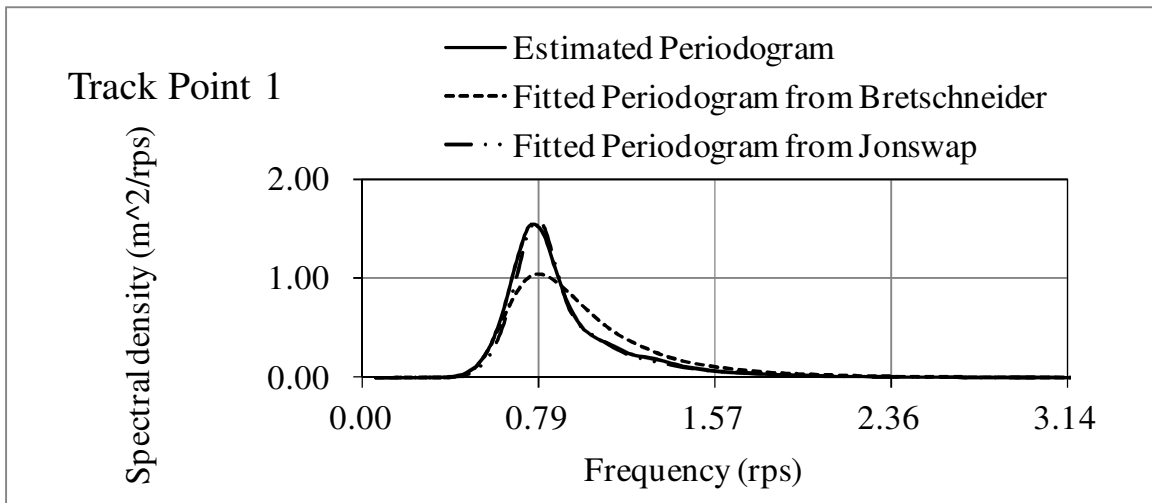
Figure 4-23. Estimated periodogram and fitted periodograms of different sea spectra for track point 1.
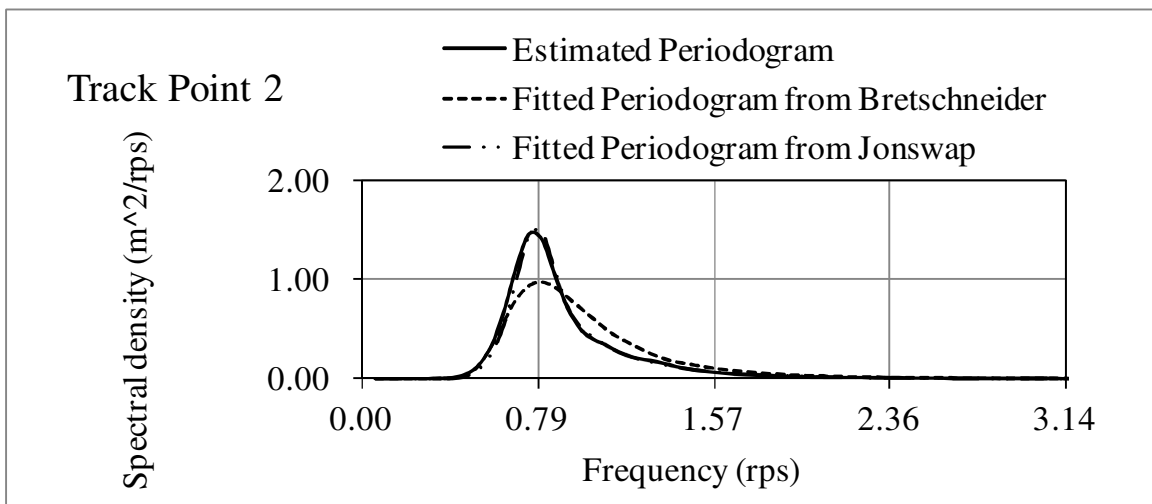
Figure 4-24. Estimated periodogram and fitted periodograms of different sea spectra for track point 2.
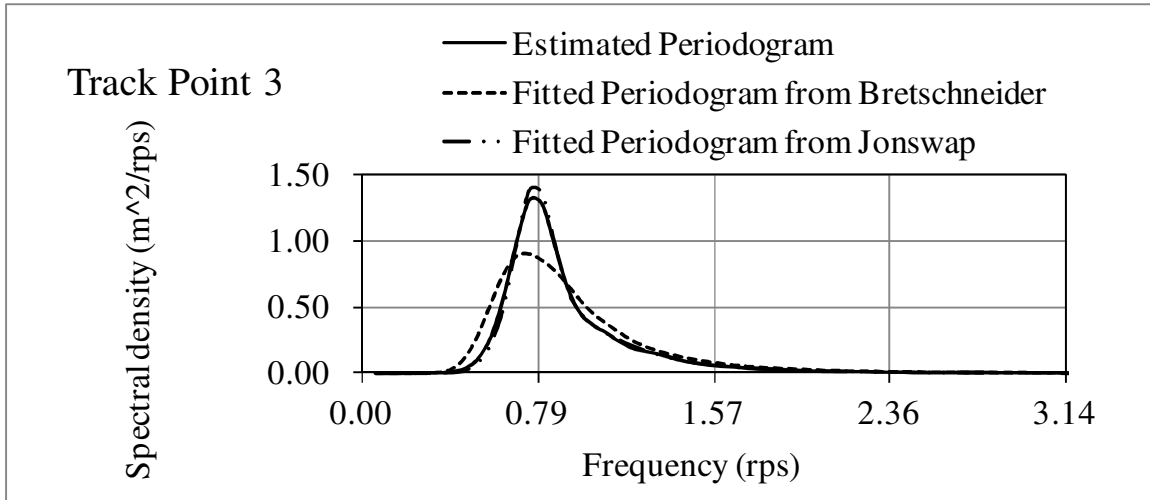
Figure 4-25. Estimated periodogram and fitted periodograms of different sea spectra for track point 3.

Table 4-7 provides the estimated wave modal periods T_{me} and the significant wave heights H_{se} along with the SWAN generated values for comparison. The results show that the absolute related errors of the modal period have a maximum value of 0.24%; while that of the significant wave height have a maximum value of 4.16%.

Table 4-8 summarizes the lower and upper limits of the confidence intervals for the modal period and significant wave height estimations. The lower and upper limits for the estimated modal period are denoted as T_L and T_U, respectively; while that for the significant wave height are denoted as H_{sL} and H_{sU}, respectively. The SWAN generated modal period and significant wave height are denoted as T_m and H_s. The results show that the values generated by SWAN are within the estimated confidence intervals for both parameters. In other words, the methodology demonstrated in Chapter 3 is verified for characterizing the sea-state conditions using the two key parameters, the modal period and the significant wave height. Figure 4-26 and Figure 4-27 show the two-sided

confidence intervals at the 95% level of track point 1 on the significant wave height H_s for the modal period $T_m = T_{me}$ and on the modal period T_m for the significant wave height $H_s = H_{se}$, respectively; while Figure 4-28 and Figure 4-29 present the results for track point 2 and Figure 4-30 and Figure 4-31 present the results for track point 3. The confidence intervals are able to capture the SWAN generated values of the modal periods and the significant wave heights.

Table 4-7. Comparison of the estimated and the SWAN generated modal periods and significant wave heights of track points 1, 2 and 3 with absolute relative errors presented in parenthesis.

| Wave properties	Point 1	Point 2	Point 3
Estimated modal period T_{me}	8.14 sec (0.24%)	8.15 sec (0.06%)	8.16 sec (0.05%)
SWAN generated modal period T_m	8.16 sec	8.15 sec	8.15 sec
Estimated significant wave height H_{se}	3.10 m (3.33%)	3.05 m (1.68%)	2.91 m (4.16%)
SWAN generated significant wave height H_s	3.21 m	3.14 m	2.79 m

Table 4-8. Two-sided confidence intervals at the 95% level of the significant wave height H_s and the modal period T_m of track points 1, 2 and 3.

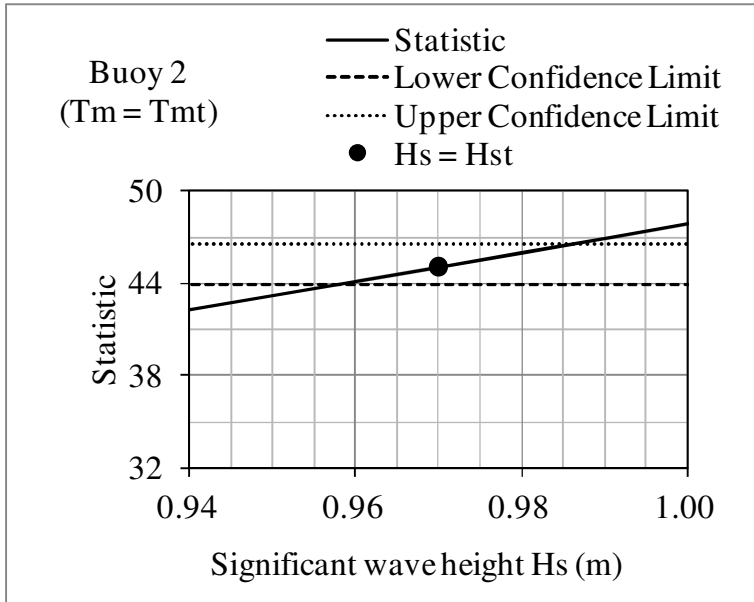| Two-sided confidence intervals at the 95% level	Point 1 T_m =8.16sec H_s =3.21 m	Point 2 T_m =8.15 sec H_s =3.14 m	Point 3 T_m =8.15 sec H_s =2.79 m
Lower modal period limit (T_L)	7.84 sec	7.85 sec	7.90 sec
Upper modal period limit (T_U)	8.20 sec	8.29 sec	8.38 sec
Lower significant wave height limit (H_{sL})	3.07 m	3.03 m	2.73 m
Upper significant wave height limit (H_{sU})	3.23 m	3.15 m	2.89 m
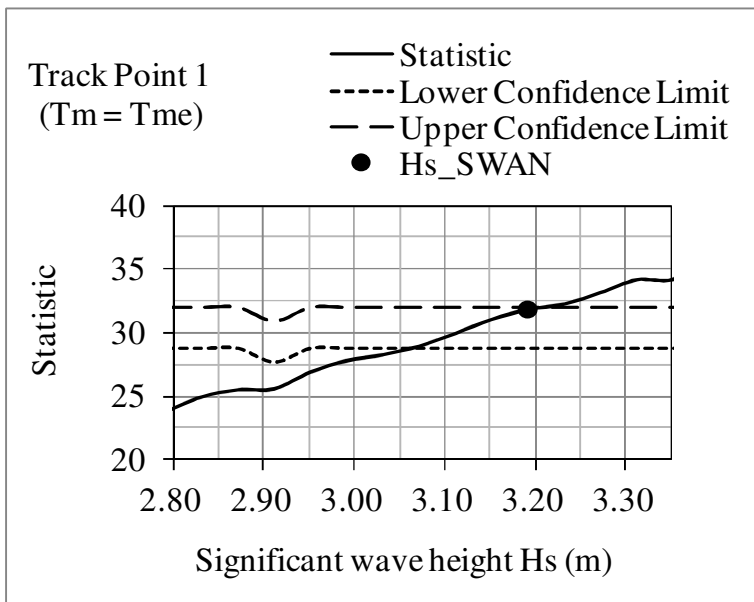
Figure 4-26. Two-sided confidence intervals at the 95% level of track point 1 on the significant wave height H_s for the modal period $T_m = T_{me}$.
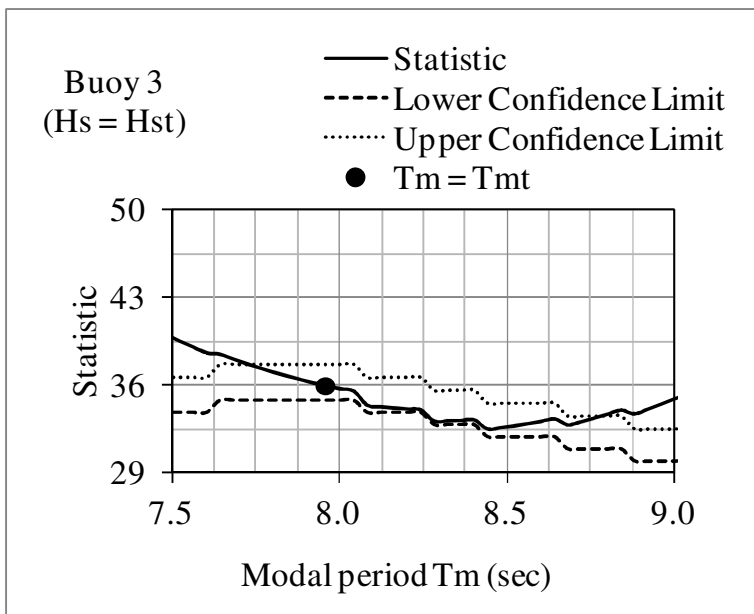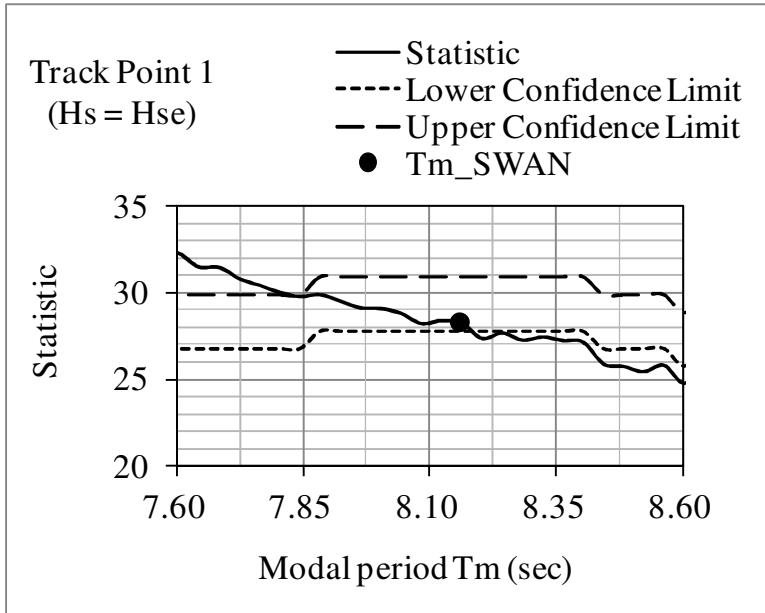
Figure 4-27. Two-sided confidence intervals at the 95% level of track point 1 on the modal period T_m for the significant wave height $H_s = H_{se}$.
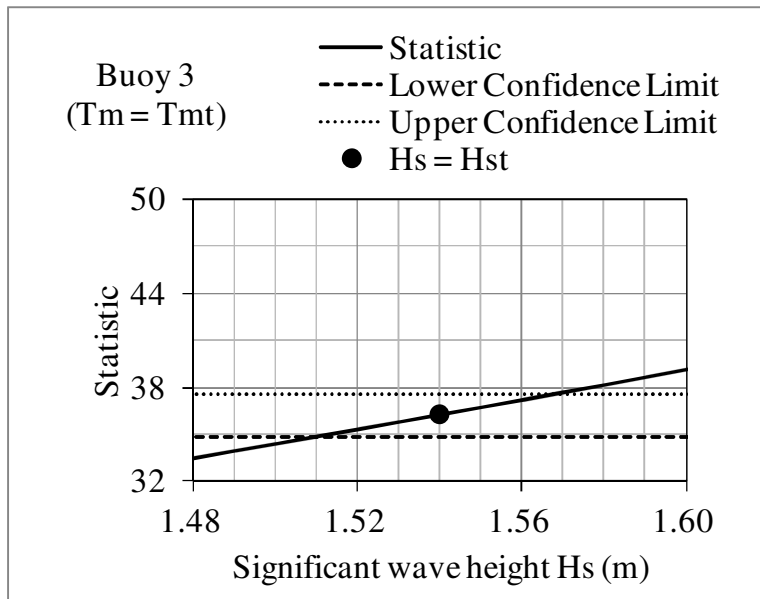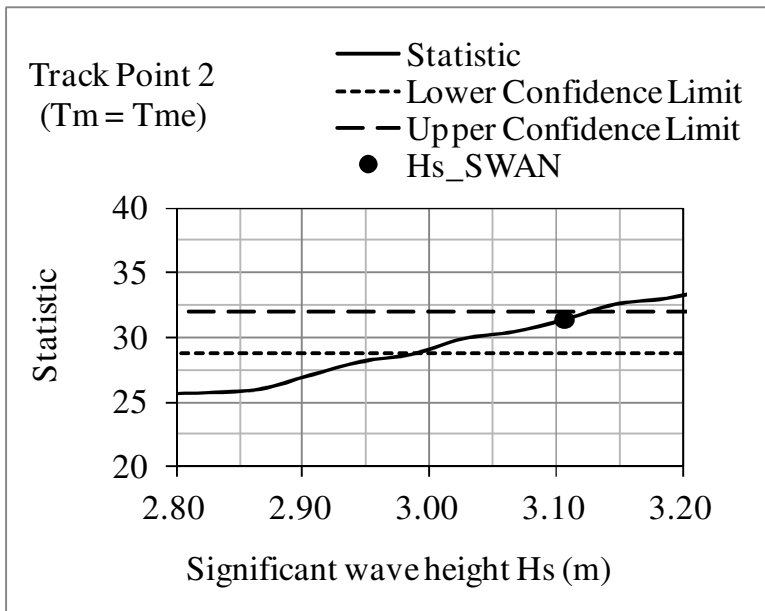

Figure 4-28. Two-sided confidence intervals at the 95% level of track point 2 on the significant wave height H_s for the modal period $T_m = T_{me}$.
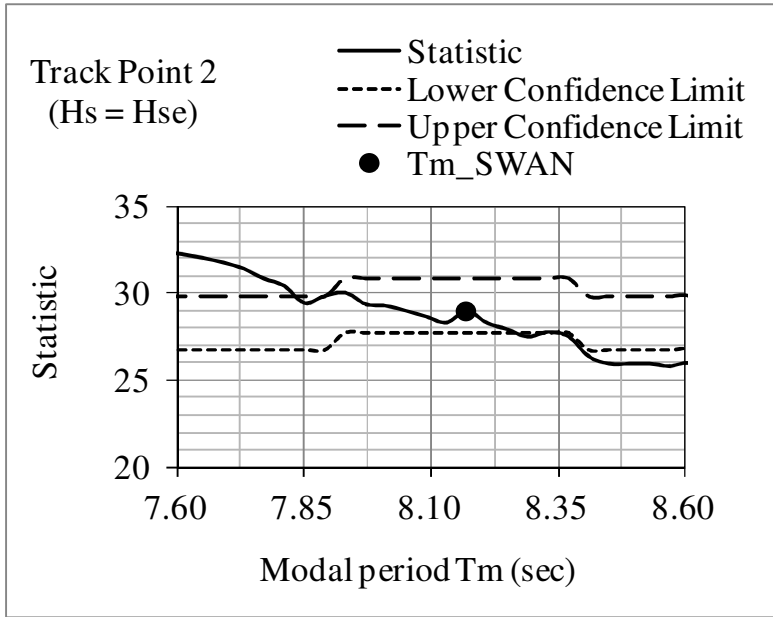
Figure 4-29. Two-sided confidence intervals at the 95% level of track point 2 on the modal period T_m for the significant wave height $H_s = H_{se}$.
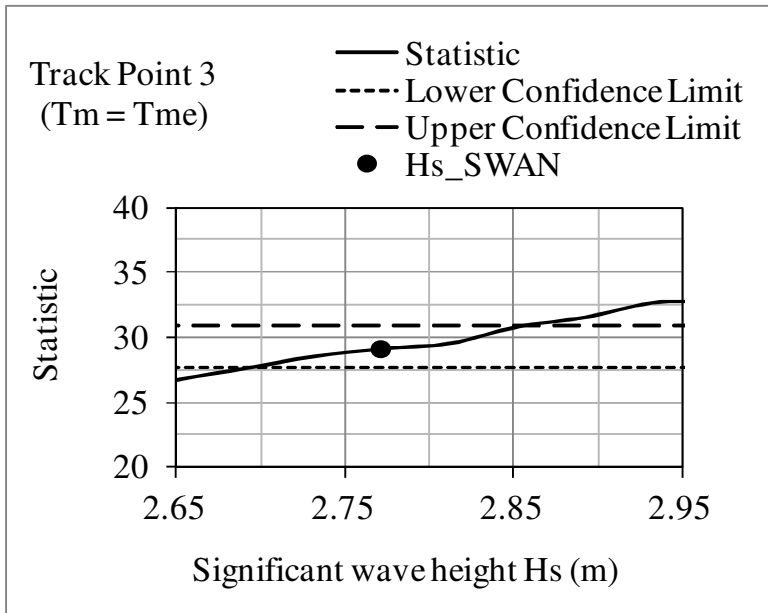
Figure 4-30. Two-sided confidence intervals at the 95% level of track point 3 on the significant wave height H_s for the modal period $T_m = T_{me}$.
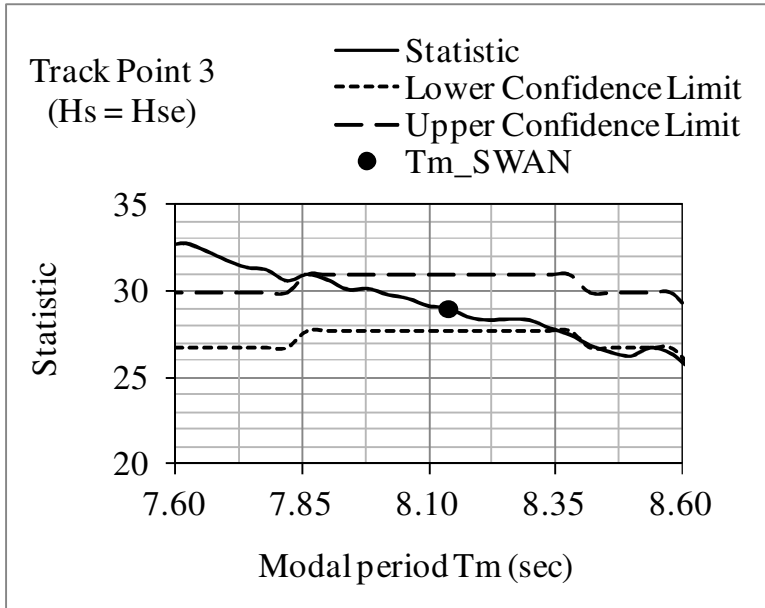
Figure 4-31. Two-sided confidence intervals at the 95% level of track point 3 on the modal period T_m for the significant wave height $H_s = H_{se}$.

4.2.3. Number of Reference Points

Estimations presented in previous sections are based on three buoys properties to interpolate other nearby locations of interest. This section provides comparisons of estimations using three, four and five buoys. In addition, inverse distance weight factor demonstrated in Section 3.1.8 uses the distance between buoys and locations of interest to the first order, i.e. d. This section discusses the results using distances to the first and second order, i.e. d^2.

Multiple locations from Figure 3-11 are selected as buoy locations and used to interpolate the sea-state parameters at locations within the buoy range. The data interpolation is performed by applying the inverse distance weight factors on the buoys spectra and interpolating the sea-state parameters based on the distance between

estimated location and the buoys.   Details can be found in Section 3.1.8.   Three, four

and five buoy locations are selected from Figure 3-11 for the data interpolation.   Figure

4-32 shows an example of selected buoy points and data interpolation range.   Locations

3, 28 and 44 marked as solid dots are the reference locations which are the buoys.   The

interpolation range is marked by the red dash lines which connect the three buoys, i.e.

locations 3, 28 and 44.   The sea-state parameter of locations 10, 11, 17, 18, 19, 24, 25,

26, 27, 31, 32, 33 and 38 will be estimated using the inverse distance weight factors

described in Chapter 3.

To verify the data interpolation results, comparisons are taken place.   Section

4.2.3 compares the significant wave heights interpolated using three, four and five

selected buoy locations and the significant wave heights generated by SWAN at these

interpolated locations.   The interpolation accuracy is determined by the relative error

square per location defined in Equation 4-1.
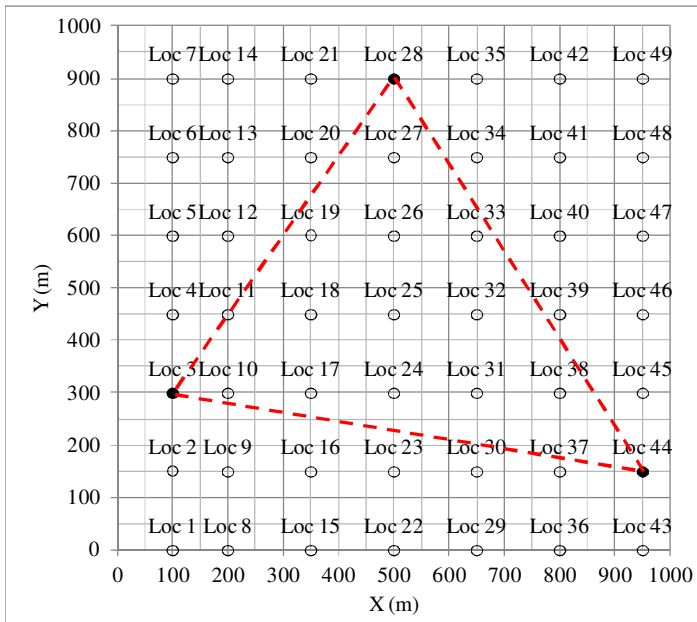
Figure 4-32. Three-point data interpolation range with buoys at locations 3, 28 and 44.

For location i, if the estimated significant wave height is denoted as Hs_i and the

SWAN simulated significant wave height is denoted as $Hs_{swan\ i}$, the error, or residual, of

the estimation can be represented as $Hs_i - Hs_{swan\ i}$, or in a unitless form as $(Hs_i - Hs_{swan\ i})/$

$Hs_{swan\ i}$, which is also called the relative error.   To estimate the interpolation error for the

entire area under consideration, the following quantity, denoted as Err, is suggested:

$$Err = \left[\sum_{i=1}^{k} \left(\frac{Hs_i - Hs_{swan\ i}}{Hs_{swan\ i}} \right)^2 \right] / k \qquad (4\text{-}1)$$

in which k is the number of interpolating data points.   Equation 4-1 shows the sum of

the relative error squares of all interpolated points and divided by the number of these

points.   Dividing the sum of the relative error squares by the number of estimation

points is to make the error estimation quantity, Err, as for one interpolation point.   The

reason of making the error term for one point is that since the interpolation range varies

due to the number of reference points, or buoys, and the selection of these reference

points, the number of estimation points will be different.   Therefore, by dividing the

sum of the relative error squares of the estimation points by the number of these points,

the error term Err is averaged and represents the relative error square per location.   It

then will be possible to compare interpolation results of different estimation ranges and

different number of estimation points by using the quantity Err.

For an area of interest such as which covered by the locations shown in Figure

3-11, the following analysis shows the one-dimensional data interpolation error square

per location (Err) for areas formed by three, four and five reference points, or buoys,

respectively.   Since the minimum number of points to form an area is three, the

estimation starts with three reference points, or say buoys.

Some examples of one-dimensional three-point estimations are shown in Figure 4-32 to Figure 4-39.   Figure 4-32 has the estimation buoy points at locations 3, 28 and 44.   The estimation points within the buoy points range are at locations 10, 11, 17, 18, 19, 24, 25, 26, 27, 31, 32, 33 and 38.   Total number of estimation locations is 13.   The data interpolation error square per location (Err) is 0.0094.   Figure 4-33 has the estimation buoy points at locations 1, 28 and 43.   The estimation points within the buoy points range are at locations 8, 9, 15, 16, 17, 18, 22, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 36, 37 and 38.   Total number of estimation locations is 20.   The data interpolation error square per location (Err) is 0.0175.   Figure 4-34 has the estimation buoy points at locations 2, 28 and 44.   The estimation points within the buoy points range are at locations 9, 10, 16, 17, 18, 19, 23, 24, 25, 26, 27, 30, 31, 32, 33, 37 and 38.   Total number of estimation locations is 17.   The data interpolation error square per location (Err) is 0.0081.   Figure 4-35 has the estimation buoy points at locations 2, 7 and 43.   The estimation points within the buoy points range are at locations 3, 4, 5, 6, 10, 11, 12, 13, 17, 18, 19, 20, 25, 26, 27, 32, 33 and 40.   Total number of estimation locations is 18.   The data interpolation error square per location (Err) is 0.0085.   Figure 4-36 has the estimation buoy points at locations 2, 7 and 46.   The estimation points within the buoy points range are at locations 3, 4, 5, 6, 10, 11, 12, 13, 17, 18, 19, 20, 24, 25, 26, 32, 33 and 39.   Total number of estimation locations is 18.   The data interpolation error square per location (Err) is 0.0101.   Figure 4-37 has the estimation buoy points at locations 7, 23 and 49.   The estimation points within the buoy points range are at locations 13, 14, 18, 19, 20, 21, 24, 25, 26, 27, 28, 32, 33, 34, 35, 41 and 42.   Total number of estimation locations is 17.   The data interpolation error square per location

(*Err*) is 0.0064.   Figure 4-38 has the estimation buoy points at locations 4, 44 and 49.

The estimation points within the buoy points range are at locations 11, 18, 25, 26, 31, 32,

33, 38, 39, 40, 41, 45, 46, 47 and 48.   Total number of estimation locations is 15.   The

data interpolation error square per location (*Err*) is 0.0131.   Figure 4-39 has the

estimation buoy points at locations 5, 44 and 49.   The estimation points within the buoy

points range are at locations 12, 19, 25, 26, 32, 33, 34, 38, 39, 40, 41, 45, 46, 47 and 48.

Total number of estimation locations is 15.   The data interpolation error square per

location (*Err*) is 0.0073.
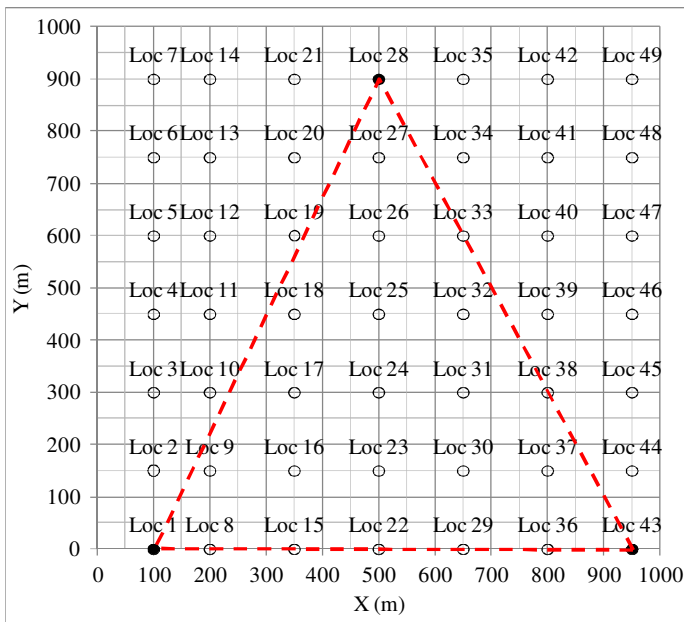
Figure 4-33. One-dimensional three-point data interpolation range with buoys at locations
1, 28 and 43.
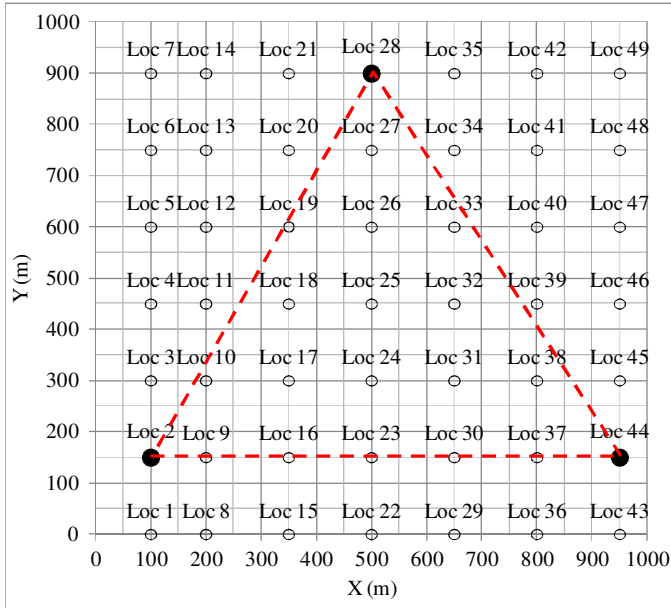
Figure 4-34. One-dimensional three-point data interpolation range with buoys at locations 2, 28 and 44.
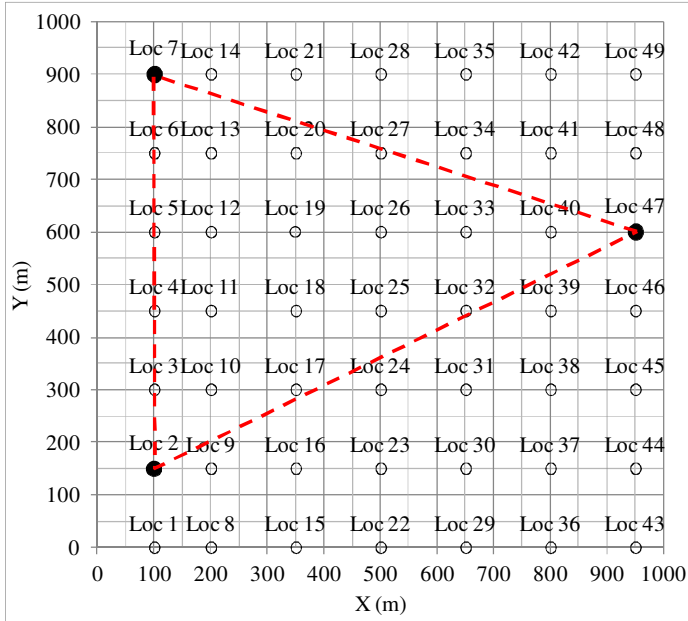
Figure 4-35. One-dimensional three-point data interpolation range with buoys at locations 2, 7 and 47.
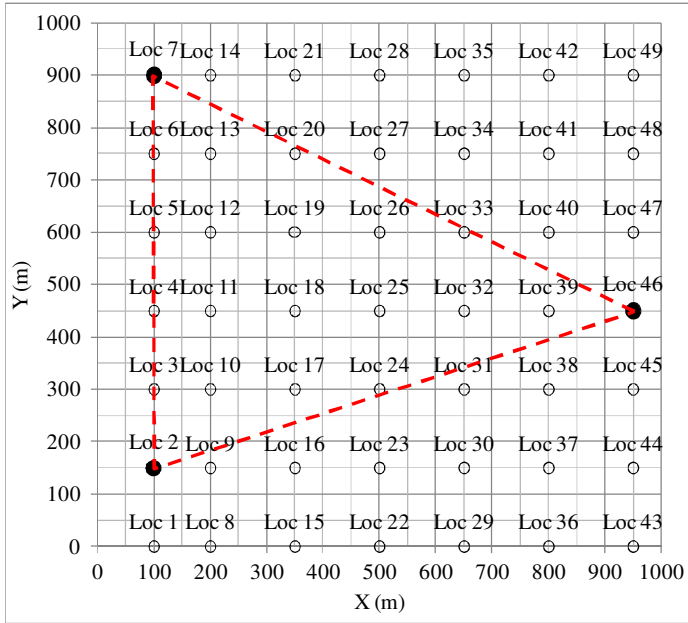
Figure 4-36. One-dimensional three-point data interpolation range with buoys at locations 2, 7 and 46.
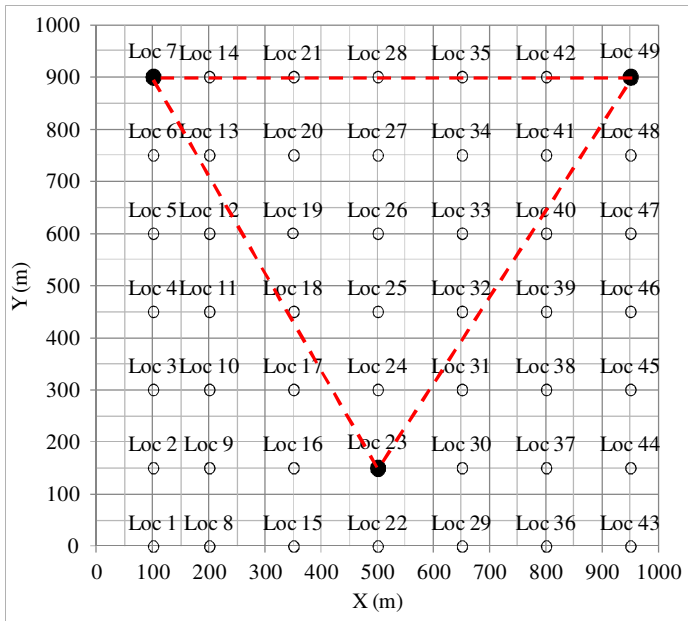
Figure 4-37. One-dimensional three-point data interpolation range with buoys at locations 7, 23 and 49.
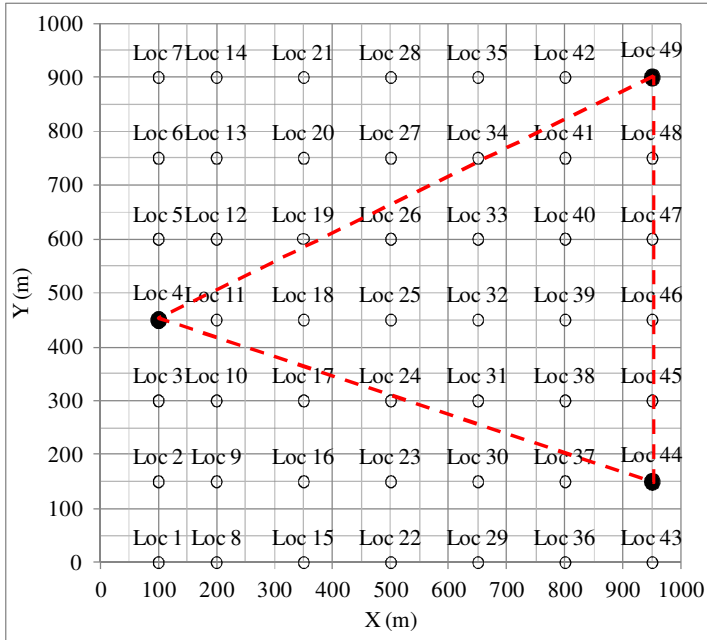
Figure 4-38. One-dimensional three-point data interpolation range with buoys at locations 4, 44 and 49.
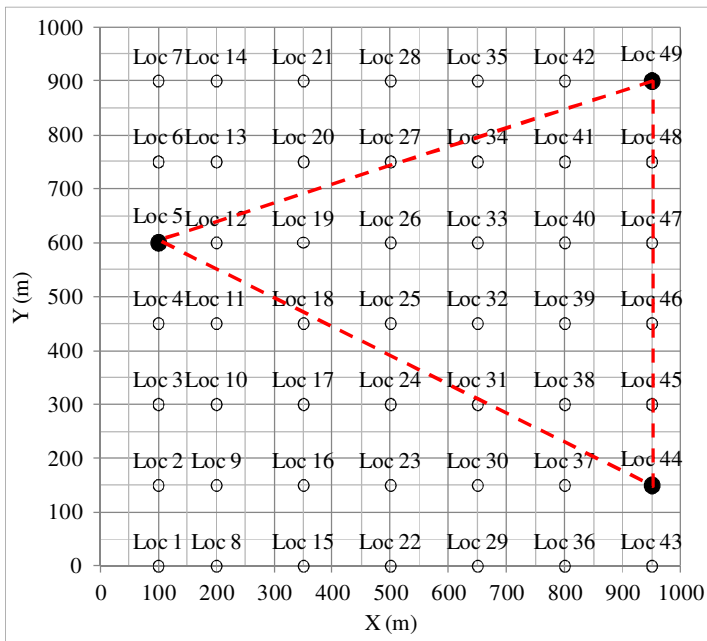
Figure 4-39. One-dimensional three-point data interpolation range with buoys at locations 5, 44 and 49.

Examples of two-dimensional three-point estimations are shown in Figure 4-40 to Figure 4-45.   Figure 4-40 has the estimation buoy points at locations 3, 25, 28 and 44. The estimation points within the buoy points range are at locations 10, 11, 17, 18, 19, 24, 26, 27, 31, 32, 33 and 38.   Total number of estimation locations is 12.   The data interpolation error square per location (Err) is 0.0112.   Figure 4-41 has the estimation buoy points at locations 1, 25, 28 and 43.   The estimation points within the buoy points range are at locations 8, 9, 15, 16, 17, 18, 22, 23, 24, 26, 27, 29, 30, 31, 32, 33, 36, 37 and 38.   Total number of estimation locations is 19.   The data interpolation error square per location (Err) is 0.0157.   Figure 4-42 has the estimation buoy points at locations 2, 25, 28 and 44.   The estimation points within the buoy points range are at locations 9, 10, 16, 17, 18, 19, 23, 24, 26, 27, 30, 31, 32, 33, 37 and 38.   Total number of estimation locations is 16.   The data interpolation error square per location (Err) is 0.0091.   Figure 4-43 has the estimation buoy points at locations 2, 26, 28 and 44.   The estimation points within the buoy points range are at locations 9, 10, 16, 17, 18, 19, 23, 24, 25, 27, 30, 31, 32, 33, 37 and 38.   Total number of estimation locations is 16.   The data interpolation error square per location (Err) is 0.0117.   Figure 4-44 has the estimation buoy points at locations 1, 7, 43 and 49.   The estimation points within the buoy points range are at locations 2 to 6, 8 to 42, and 44 to 48.   Total number of estimation locations is 45.   The data interpolation error square per location (Err) is 0.0188.   Figure 4-45 has the estimation buoy points at locations 2, 7, 44 and 49.   The estimation points within the buoy points range are at locations 3 to 6, 9 to 14, 16 to 21, 23 to 28, 30 to 35, 37 to 42, and 45 to 48.   Total number of estimation locations is 38.   The data interpolation error square per location (Err) is 0.0094.
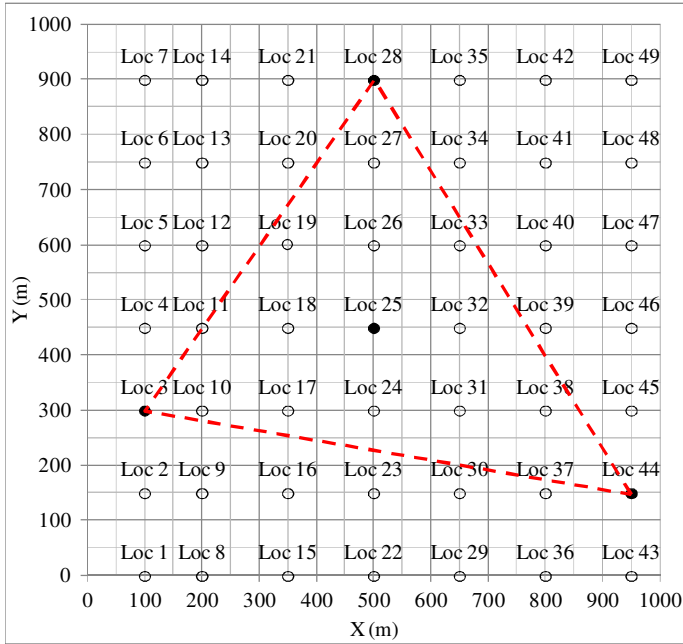
Figure 4-40. One-dimensional four-point data interpolation range with buoys at locations 3, 25, 28 and 44.
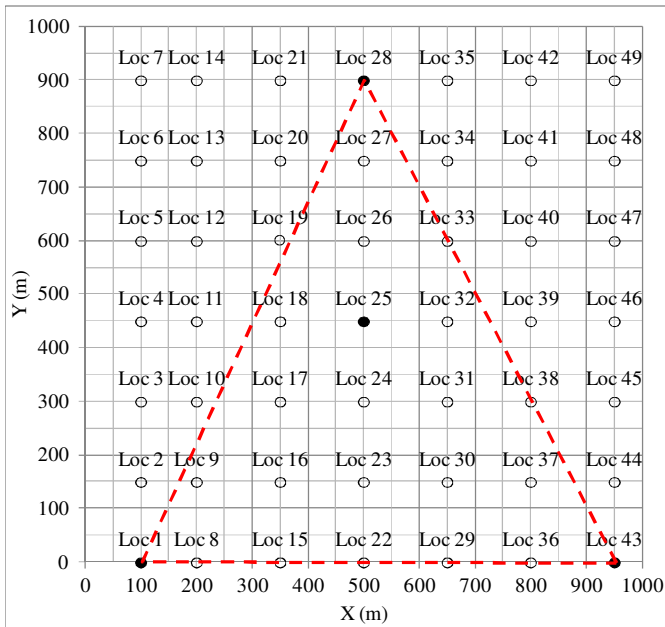
Figure 4-41. One-dimensional four-point data interpolation range with buoys at locations 1, 25, 28 and 43.
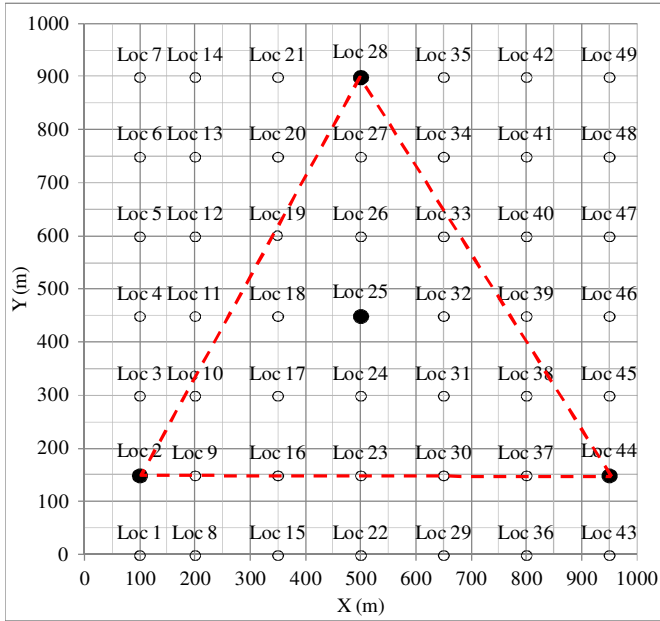
Figure 4-42. One-dimensional four-point data interpolation range with buoys at locations 2, 25, 28 and 44.
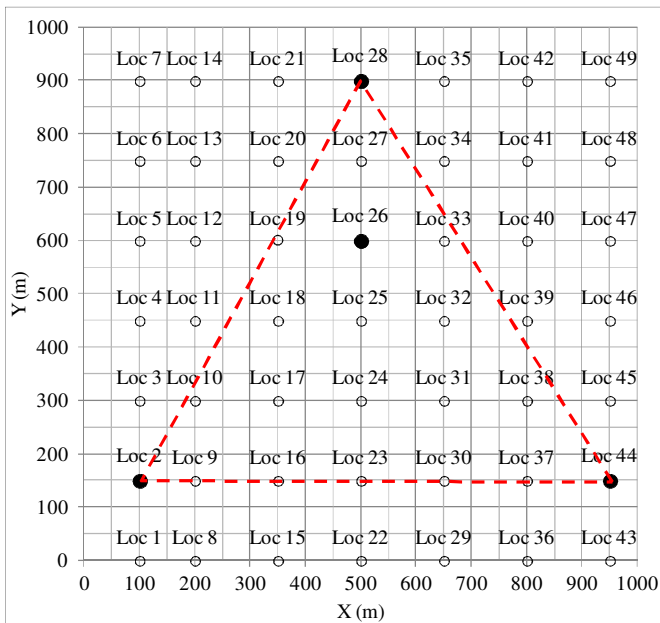
Figure 4-43. One-dimensional four-point data interpolation range with buoys at locations 2, 26, 28 and 44.
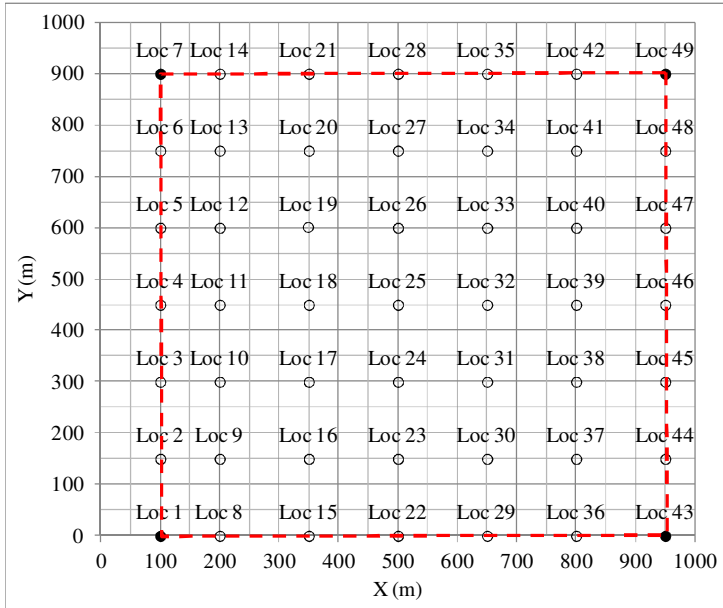
Figure 4-44. One-dimensional four-point data interpolation range with buoys at locations 1, 7, 43 and 49.
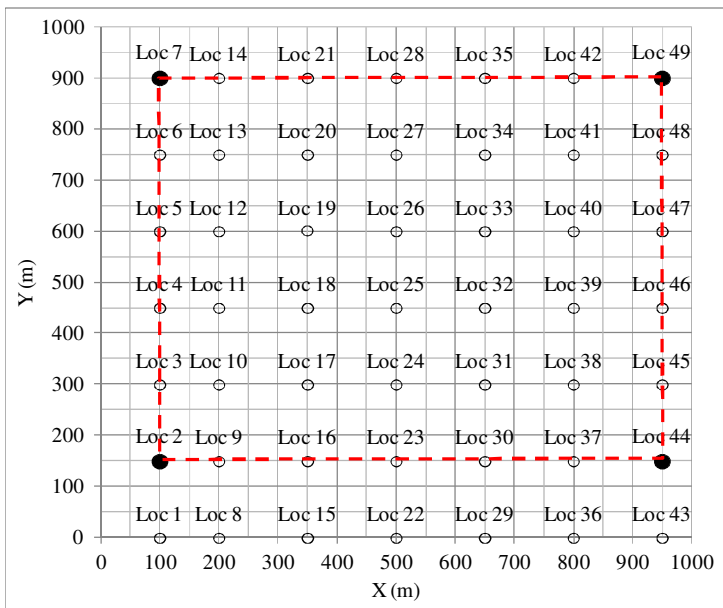
Figure 4-45. One-dimensional four-point data interpolation range with buoys at locations 2, 7, 44 and 49.

Figure 4-46 to Figure 4-48 show the one-dimensional five-point estimations. Figure 4-46 has the estimation buoy points at locations 1, 7, 25, 43 and 49.   The estimation points within the buoy points range are at locations 2 to 6, 8 to 24, 26 to 42 and 44 to 48.   Total number of estimation locations is 44.   The data interpolation error square per location (*Err*) is 0.0211.   Figure 4-47 has the estimation buoy points at locations 2, 7, 25, 44 and 49.   The estimation points within the buoy points range are at locations 3 to 6, 9 to 14, 16 to 21, 23, 24, 26 to 28, 30 to 35, 37 to 42 and 45 to 48. Total number of estimation locations is 37.   The data interpolation error square per location (*Err*) is 0.0125.   Figure 4-48 has the estimation buoy points at locations 2, 7, 26, 44 and 49.   The estimation points within the buoy points range are at locations 3 to 6, 9 to 14, 16 to 21, 23 to 25, 27, 28, 30 to 35, 37 to 42 and 45 to 48.   Total number of estimation locations is 37.   The data interpolation error square per location (*Err*) is 0.0130.
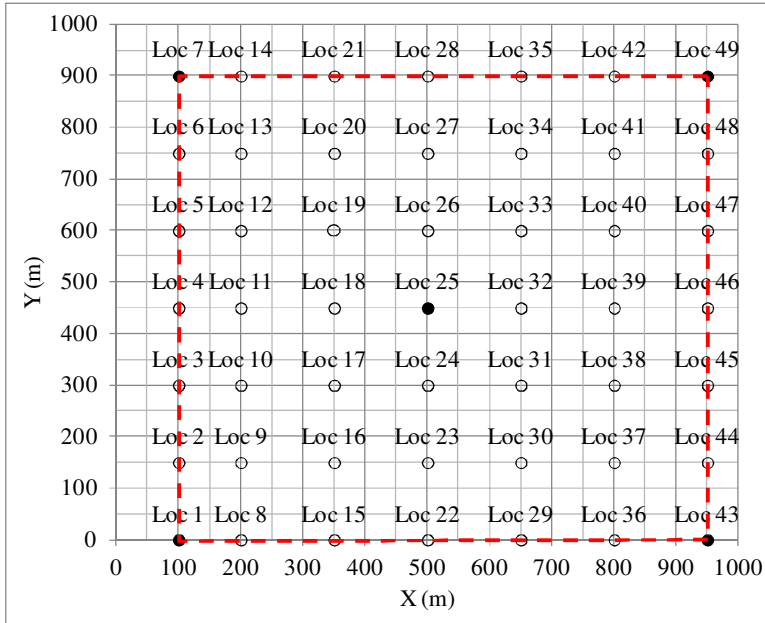
Figure 4-46. One-dimensional five-point data interpolation range with buoys at locations 1, 7, 25, 43 and 49.
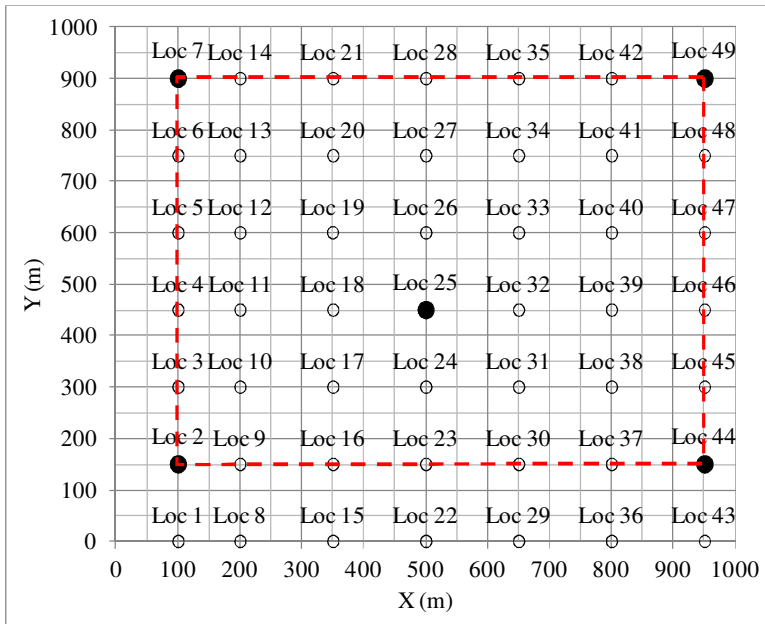
Figure 4-47. One-dimensional five-point data interpolation range with buoys at locations 2, 7, 25, 44 and 49.
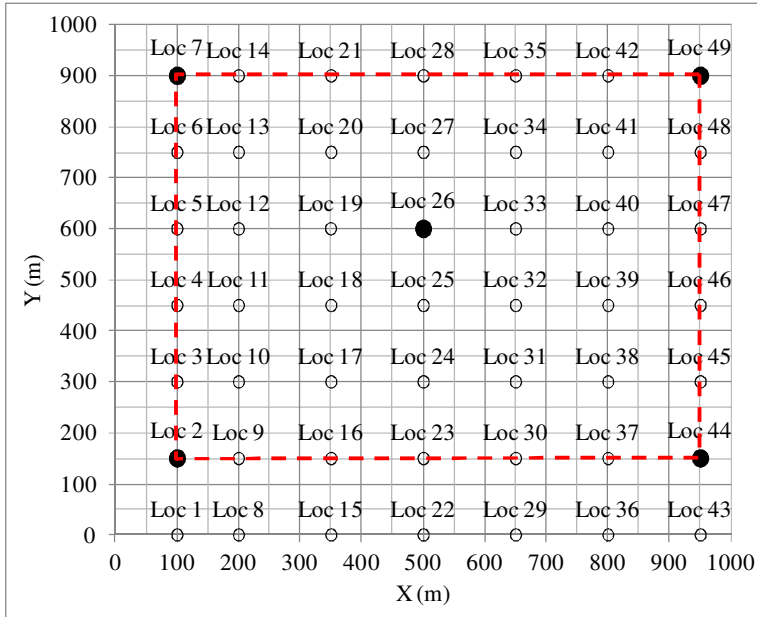
Figure 4-48. One-dimensional five-point data interpolation range with buoys at locations 2, 7, 26, 44 and 49.

Table 4-9 summarizes the examples of using three reference points which form triangular areas for data interpolation. The relative error square per location, *Err*, ranges between 0.0064 and 0.0175, and the number of estimation points are between 13 and 20. Table 4-9 shows that the one-dimensional three-point estimations produce interpolations within 1.8% relative error square per location.

Table 4-10 summarizes the examples of using four reference points for data interpolation. An area formed by four reference points can be a quadrilateral or a triangle with an additional reference point in the middle of the triangle. The relative error square per location, *Err*, of the examples shown in Table 4-10 have a range between 0.0091 and 0.0188, and the number of estimation points are between 12 and 45. Table 4-10 shows that the one-dimensional four-point estimations produce interpolations within 1.9% relative error square per location.

120

Table 4-11 summarizes the examples of using five reference points for data interpolation.   The area of the five-point interpolation is a quadrilateral with an additional reference point in the middle of the area.   The relative error square per location, *Err*, of the examples shown in Table 4-11 have a range between 0.0125 and 0.0211, and the number of estimation points are between 37 and 44.   From Table 4-9, Table 4-10 and Table 4-11, the three-point, four-point, and five-point interpolations show that the relative error square per location is about 2% regardless the number of reference points to be three, four, or five.

Table 4-9. Comparison of one-dimensional three-point data interpolation error square per location (*Err*) for various estimation ranges.

(in Figure 4-32)

Err = 0.0094

(in Figure 4-33)

Err = 0.0175

(in Figure 4-34)

Err = 0.0081

(in Figure 4-35)

Err = 0.0085

(in Figure 4-36)

Err = 0.0101

(in Figure 4-37)

Err = 0.0064

(in Figure 4-38)

Err = 0.0131

(in Figure 4-39)

Err = 0.0073
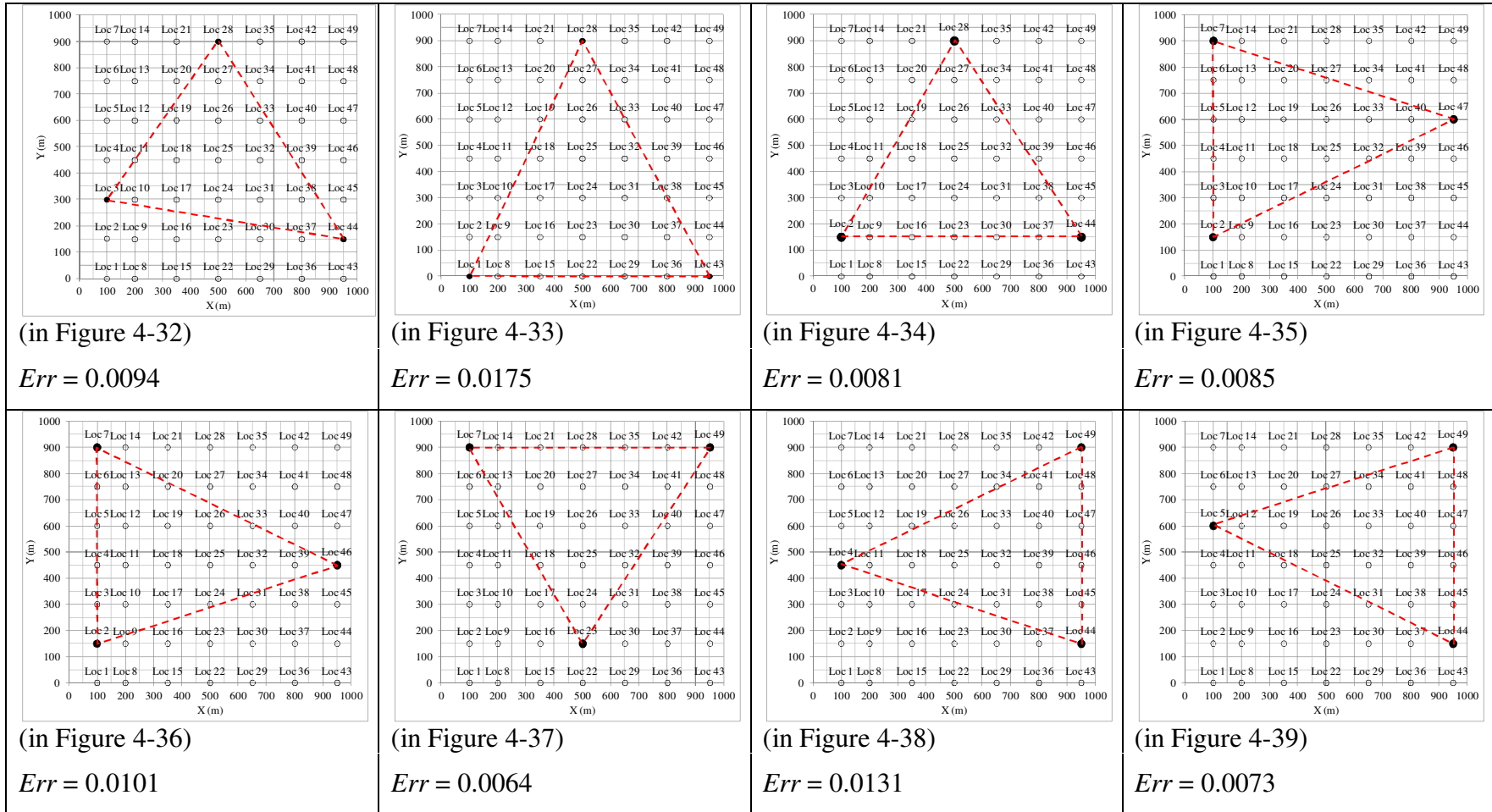
Table 4-10. Comparison of one-dimensional four-point data interpolation error square per location (*Err*) for various estimation ranges.

(in Figure 4-40)

Err = 0.0112

(in Figure 4-41)

Err = 0.0157

(in Figure 4-42)

Err = 0.0091

(in Figure 4-43)

Err = 0.0117

(in Figure 4-44)

Err = 0.0188

(in Figure 4-45)

Err = 0.0094

123
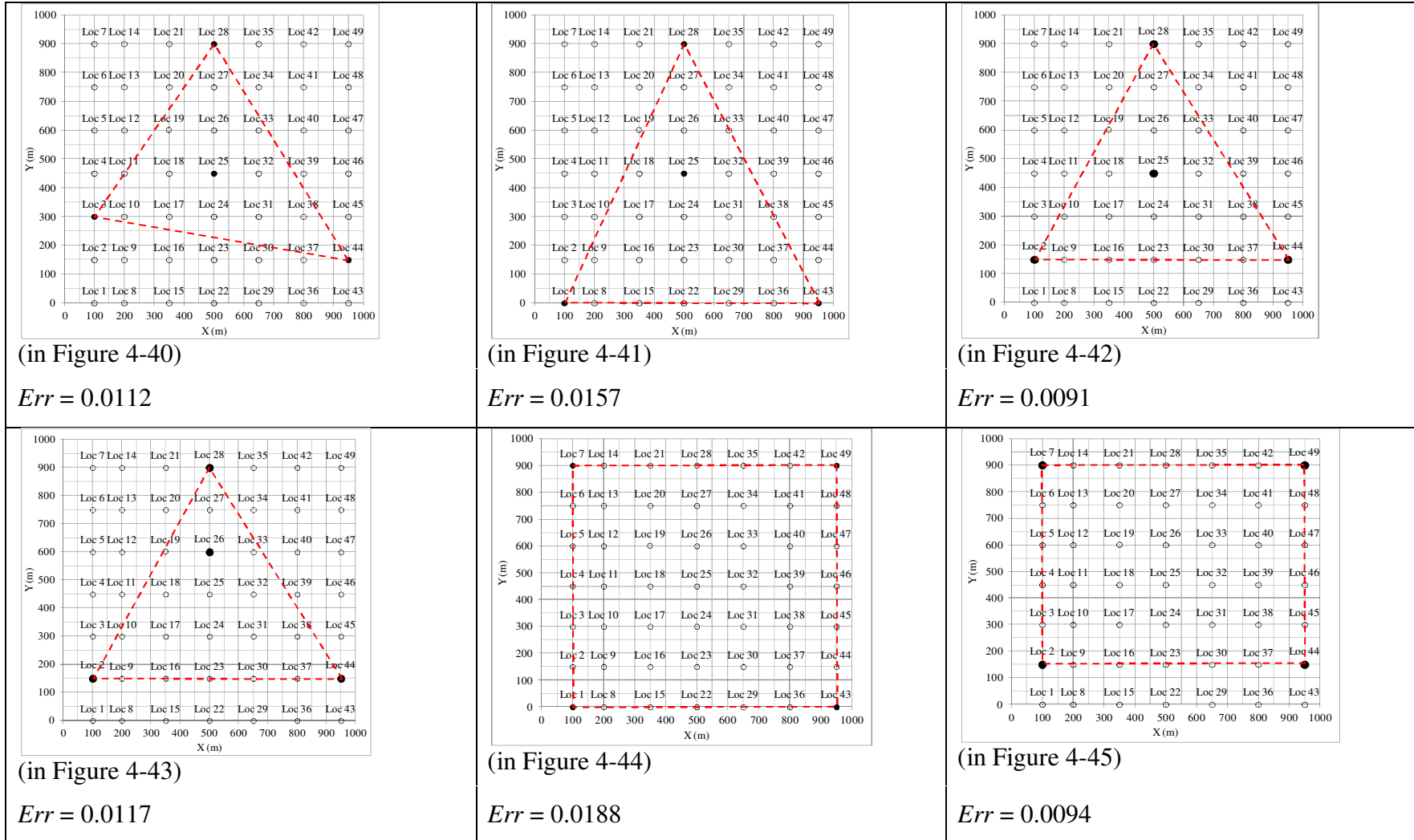
Table 4-11. Comparison of one-dimensional five-point data interpolation error square per location (*Err*) for various estimation ranges.

(in Figure 4-46)

Err = 0.0211

(in Figure 4-47)

Err = 0.0125

(in Figure 4-48)

Err = 0.0130
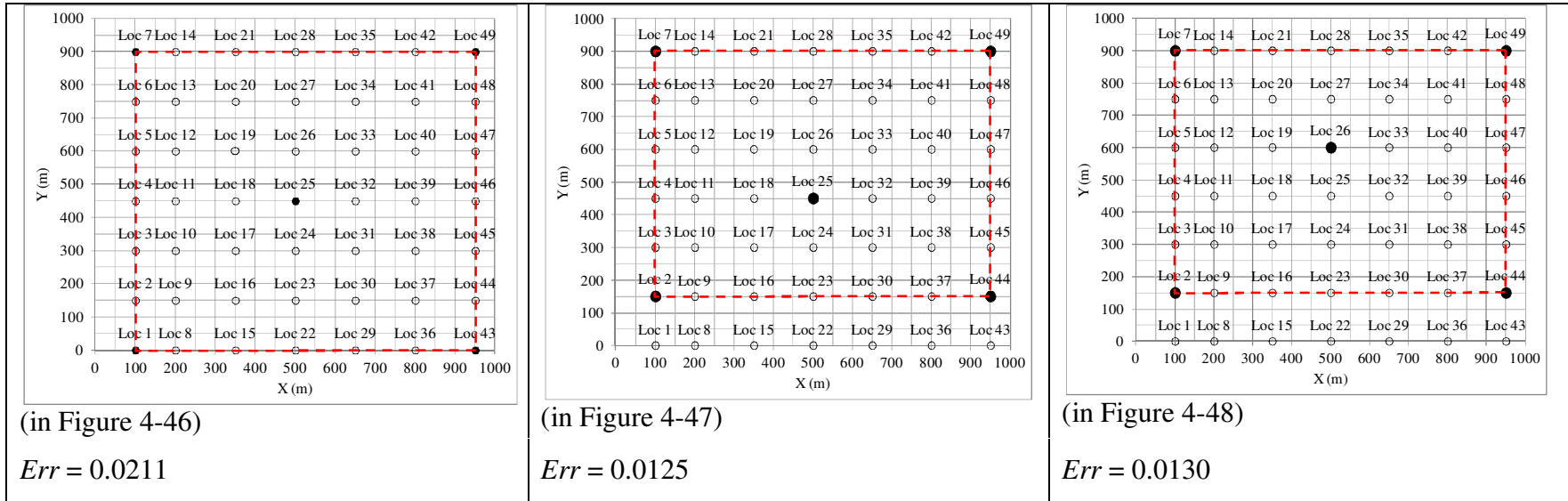
As the one-dimensional data interpolation shows that the relative error square per location (*Err*) is about 2%, the following summary is to show the two-dimensional data interpolation outcomes. To form a two-dimensional range, the minimum number of points is four, which make up a tetrahedron. By adding one additional point, five points can make a pyramid. The four points two-dimensional estimation shown in Figure 4-49 is using the same reference points as shown in Figure 4-40 which is the one-dimensional four-points interpolation having a triangular area and an additional points in the middle. The two-dimensional range is formed by considering the additional point in the middle of the area in a different plane. So the same four points form a tetrahedron now. For location 10 as an example, the two-dimensional estimation is performed by averaging the estimations interpolated using locations 3, 28 and 44 and using locations 3, 25 and 44. The estimations interpolated using locations 3, 28 and 44 and using locations 3, 25 and 44 are both one-dimensional three-point cases and provide interpolation results for location 10 at the plane of locations 3, 28 and 44 and at the plane of location 3, 25 and 44, respectively. By averaging the results of the two planes, the two-dimensional estimation for location 10 is the average of the one-dimensional estimations of the two planes cover location10.
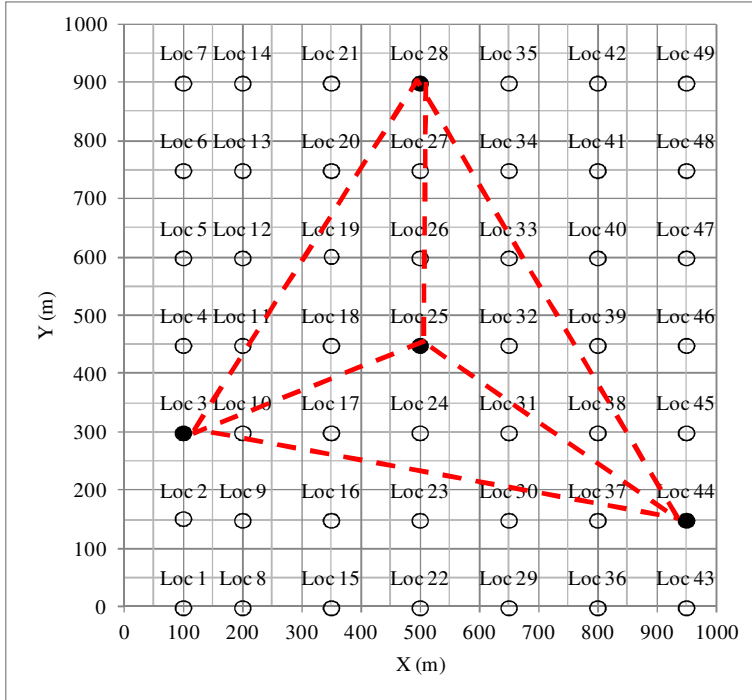
Figure 4-49. Two-dimensional four-point data interpolation range.

As shown above in Figure 4-40 and Figure 4-49, the two estimation cases have the same range of estimation except the four-point case has an additional buoy point in the middle at location 25. The three-points interpolation based on locations 3, 28 and 44 leads to the error square per location, *Err*, of 0.0094, and the four-points interpolation based on locations 3, 28, 44 and the middle point 25 leads to the error square per location, *Err*, of 0.0112. The results show that additional buoy points within the same range might not improve the estimation. If taking the same four points, 3, 25, 28 and 44 but estimating in two-dimensional as a tetrahedron, to the error square per location, *Err*, is 0.0110 which does not show significant difference from the one-dimensional estimations, either.

126

Moreover, if modify the inverse distance weight factor shown in Equation 3-15 by rising the power of distance to two as follows:

$$wb_{ij} = \frac{1/d_{ij}^2}{\sum_j 1/d_{ij}^2} \quad \text{for } d_{ij} \neq 0 \tag{4-2}$$

The data interpolation error square per location, *Err*, is 0.0140 for the one-dimensional three-point estimation based on locations 3, 28 and 44; while *Err*, is 0.0115 and 0.0127 for four-point, one-dimensional and two-dimensional estimations, respectively. The modified weight factor in Equation 4-2 does not show advantages compared with the factor defined in Equation 3-15.

4.3. Verification and Validation using Buoy Data

This section provides sea-state condition analysis on buoy locations selected from the National Oceanic and Atmospheric Administration (NOAA) website. Four buoy locations on the east coast of the United States are selected for analysis. Three of the four locations are considered as buoys with given sea-state information while the remaining location does not have sea-state information and needs estimations. Sea-state characterization is performed on this remaining location based on the methodology demonstrated in Chapter 3.

4.3.1. Description of Buoy Data

Figure 4-50 shows the four buoy locations selected for this example. Location A denoted as buoy 1 represents the buoy of Station ID 44009. Location B denoted as buoy 2 represents the buoy of Station ID 44025. Location C denoted as buoy 3 represents the buoy of Station ID 44011. Location D denoted as buoy 4 represents the buoy of Station

ID 44008.   Buoy 4 is considered the estimation point without given sea-state information.   Estimations for the estimation point are based on the sea-state conditions obtained from buoys 1, 2, and 3.   Table 4-12 summarizes the coordinates of these four locations.   The latitude and longitude coordinates are provided from the NOAA website. To indicate the locations of these four locations in Cartesian coordinate system, Table 4-12 shows the zones and the easting and northing coordinates in the Universal Transverse Mercator (UTM) coordinate system corresponding to the given latitudes and longitudes.   There are sixty zones in the Universal Transverse Mercator (UTM) coordinate system.   The width of each zone is about 1,000,000 m.   The easting coordinate of a specific zone indicates the distance from the west boundary of the zone. Since these four locations in this example are in two different zones, the easting coordinates of the Universal Transverse Mercator (UTM) system need to be adjusted to obtain the X and Y coordinates shown in Table 4-12.   For example, the coordinate easting 525,998 m in zone 18 N for buoy 1 means that buoy 1 is at the distance of 525,998 m from the west boundary of zone 18 N.   Similarly, buoy 3 is at the distance of 701,531 m from the west boundary of zone 19 N.   Since these two easting coordinates are not in the same zone, they couldn't be plotted in an X-Y plane using their easting coordinates.   By the fact that zone 19 is on the east side of and adjacent to zone 18, the easting coordinate 701,531 m of buoy 3 can be adjusted by adding 1,000,000 m to 701,531 m to obtain the easting distance from the west boundary of zone 18 as 1,701,531 m.   The adjusted easting coordinates are shown in Table 4-12 as the X coordinate which indicate the distance from the west boundary of zone 18 N.   The origin of the X axis, i.e.

128

X = 0 km, in Figure 4-50 indicates the west boundary of zone 18 N on the Universal

Transverse Mercator (UTM) coordinate system.

As shown in Figure 4-50 and Table 4-12, the buoy of Station ID 44009 is

considered as buoy 1 in this example, while buoys of Station ID 44025 and 44011 are

considered as buoy 2 and buoy 3, respectively.   The buoy of Station ID 44008 is buoy 4

and is treated as an unobserved location and needs estimation for the sea-state conditions.

This location is denoted as estimation point.   Table 4-13 provides the mean values of

modal period and significant wave height obtained from the NOAA website for these

four locations.   The modal period and significant wave height of buoy 4, the estimation

point, are provided for reference and verification of the estimation results.

The modal period and significant wave height data were collected hourly.   The

mean values were computed by adding up all the hourly collected data and dividing it by

the number of data.   In this example, the data collected in January are used.   Buoy 1

has data collected from 1986 to 2008.   Buoy 2 has data collected from 1991 to 2008.

Buoy 3 has data collected from 1984 to 2008.   Buoy 4 has data collected from 1982 to

2008.   Appendix B provides the obtained NOAA data used in this example.
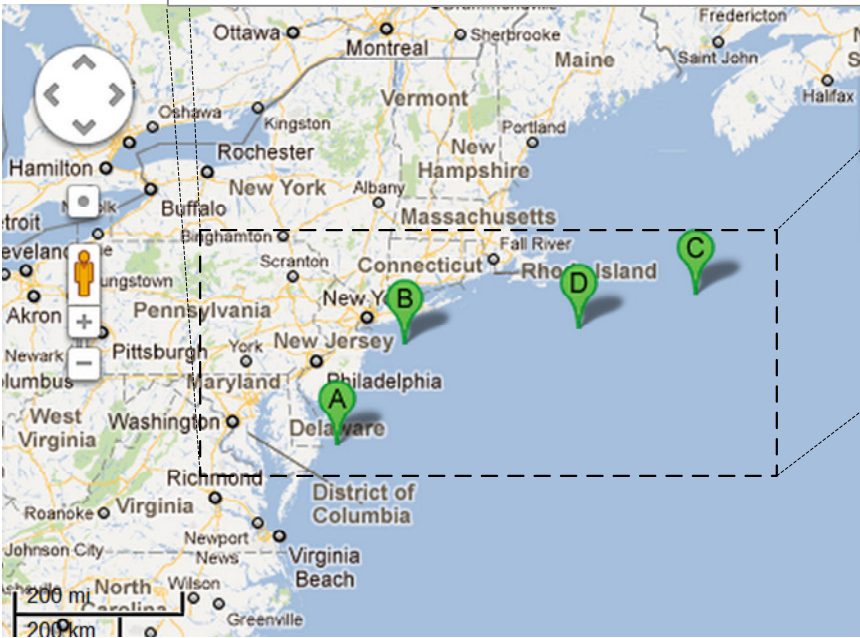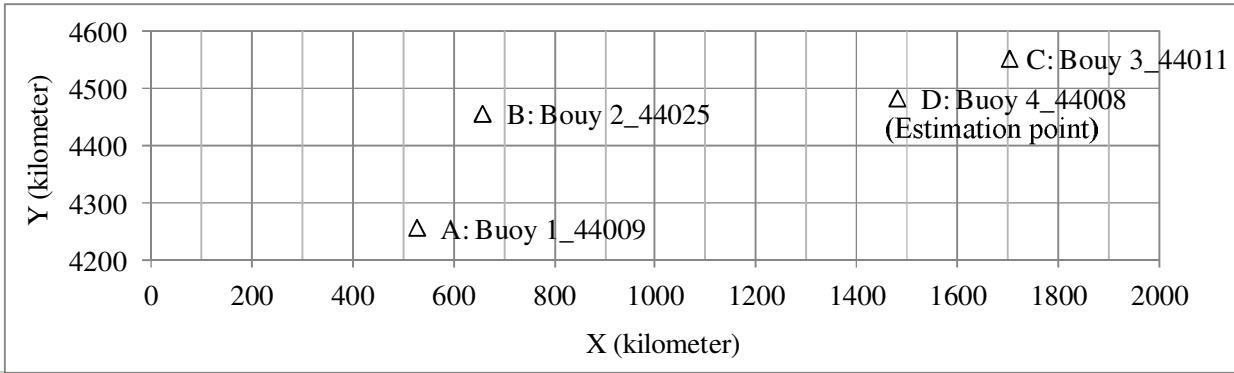
Figure 4-50. Locations of interest for estimations.

Table 4-12. Coordinates of locations shown in Figure 4-50.

Station ID	Latitude	Longitude	Zone	Easting (m)	Northing (m)	X (km)	Y (km)
44009: Buoy 1	38.464N	74.702W	18N	525,998	4,257,341	525.998	4257.341
44025: Buoy 2	40.250N	73.167W	18N	655,897	4,457,117	655.897	4457.117
44011: Buoy 3	41.105N	66.600W	19N	701,531	4,553,189	1701.531	4553.189
44008: Estimation point (Buoy 4)	40.502N	69.247W	19N	479,071	4,483,506	1479.071	4483.506

Table 4-13. Modal periods and significant wave heights of locations defined in Figure 4-50.

Location	Modal period T_m	Significant wave height H_s
44009: Buoy 1	7.2 sec	1.4 m
44025: Buoy 2	6.8 sec	1.5 m
44011: Buoy 3	8.5 sec	2.8 m
44008: Estimation point (Buoy 4)	8.0 sec	2.4 m

4.3.2. Data Interpolation

According to Section 3.1.8, the inverse distance weight factors applied on buoys for the estimating point are computed using Equations 3-15 to 3-17and summarized in Table 4-14.   Since the estimated point is closer to buoy 3 as shown in Figure 4-50, the weight factor applied to buoy 3 is larger than that applied to buoy 1 and buoy 2.   These weight factors are applied on the adjusted periodograms of the buoys using Equation 3-18 to estimate the periodograms for estimation point.   As demonstrated in Chapter 3, wave spectrum goodness-of-fit is performed to determine the spectrum type for estimation point.   Figure 4-51 shows the periodogram of estimation point and the periodograms constructed from Bretschneider and Jonswap spectra.   The modal period and significant wave height generating these spectra are based on the least square concept demonstrated in Section 3.1.4.   As shown in Figure 4-51, the periodogram constructed from the Bretschneider spectrum fits the estimation point periodogram better compared with the periodogram constructed from the Jonswap spectrum.   Therefore, the spectrum type for estimation point is determined as the Bretschneider spectrum.   The estimated modal period and significant wave height of Bretschneider spectrum are summarized in Table 4-15 and compared with the observation values provided from the NOAA website shown previously in Table 4-13.   Comparing with NOAA observations, the estimated modal period has an absolute relative error of 1.4 %; while the estimated significant wave height has an absolute relative error of 2.2 %.

Table 4-14. Weight factors applying on the three buoys for the estimation point (buoy 4).

Inverse distance weight factors	Estimation point (buoy 4)
Weight factor applied to buoy 1	0.1565
Weight factor applied to buoy 2	0.1861
Weight factor applied to buoy 3	0.6574
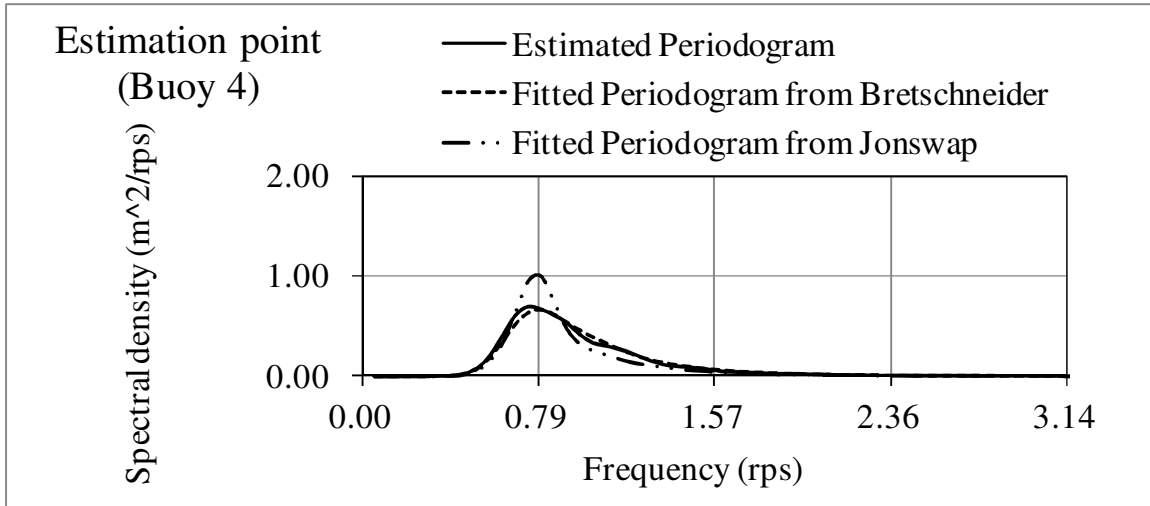
Figure 4-51. Estimated periodogram and fitted periodograms of different sea spectra for the estimation point (buoy 4).

Table 4-15. Estimated modal period and significant wave height of the estimation point (buoy 4) with absolute relative errors based on the NOAA observations.

Wave properties	NOAA observation	Estimation	Absolute relative error
Modal period	T_m = 8.0 sec	T_{me} = 8.11 sec	1.4 %
Significant wave height	H_s = 2.4 m	H_{se} = 2.45 m	2.2 %

According to the spectrum type determined by wave spectrum goodness-of-fit, hypothesis testing can be performed for parameter estimations to characterize the sea-state conditions for the estimated point. Hypothesis testing concept is demonstrated in Section 3.1.5. The parameter confidence intervals estimations follow the procedure described in Section 3.1.6. Table 4-16 summarizes the two-sided confidence intervals at the 95 % level of the modal period and significant wave height for the estimation point (buoy 4). The 95 % confidence interval of the modal period is between 7.96 sec and 8.18 sec. The estimated modal period based on the periodogram of the estimation point (buoy 4) is 8.11sec. The NOAA observation is 8.0 sec. Both 8.11 sec and 8.0 sec are within the confidence interval. As for the significant wave height, the 95 % confidence interval is between 2.38 m and 2.5 m. The estimated significant wave height based on the periodogram of the estimation point (buoy 4) is 2.45 m; while the NOAA observation is 2.4 m. Both 2.45 m and 2.4 m are within the confidence interval as well. Figure 4-52 shows the two-sided confidence interval at the 95% level of the estimation point (buoy 4) on the significant wave height H_s for the modal period $T_m = T_{me}$, where T_{me} is the modal period estimated from the periodogram. Figure 4-53 shows the two-sided confidence interval at the 95% level of the estimation point (buoy 4) on the modal period T_m for the significant wave height $H_s = H_{se}$, where H_{se} is the significant wave height estimated from the periodogram.

Table 4-16. Two-sided confidence intervals at the 95% level of the significant wave height H_s and the modal period T_m of the estimation point (buoy 4).

Two-sided confidence intervals at the 95% level	Estimation point (Buoy 4) Estimation: $T_{me} = 8.11$ sec, $H_{se} = 2.45$ m NOAA observation: $T_m = 8.0$ sec, $H_s = 2.4$ m
Lower modal period limit (T_L)	7.96 sec
Upper modal period limit (T_U)	8.18 sec
Lower significant wave height limit (H_{sL})	2.38 m
Upper significant wave height limit (H_{sU})	2.50 m
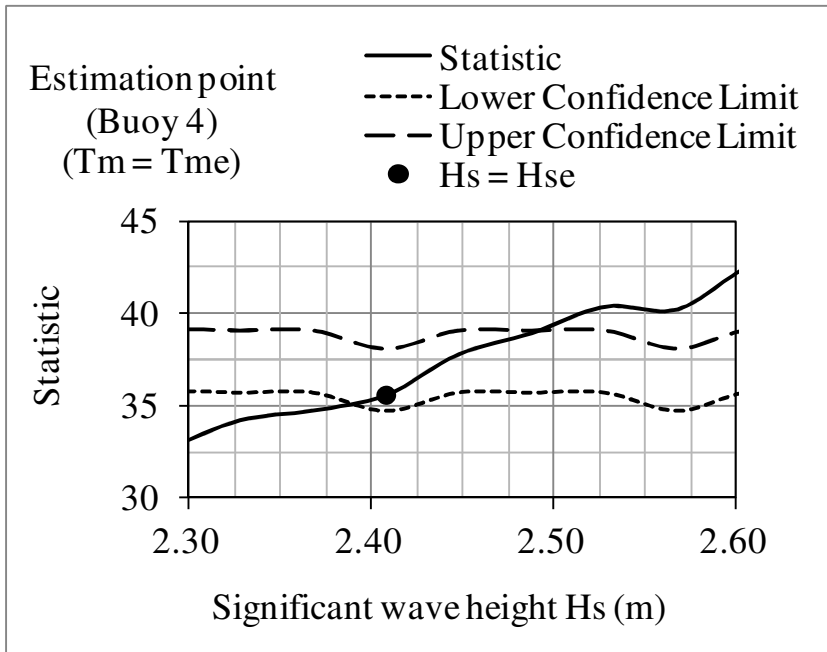
Figure 4-52. Two-sided confidence intervals at the 95% level of the estimation point (buoy 4) on the significant wave height H_s for the modal period $T_m = T_{me}$.
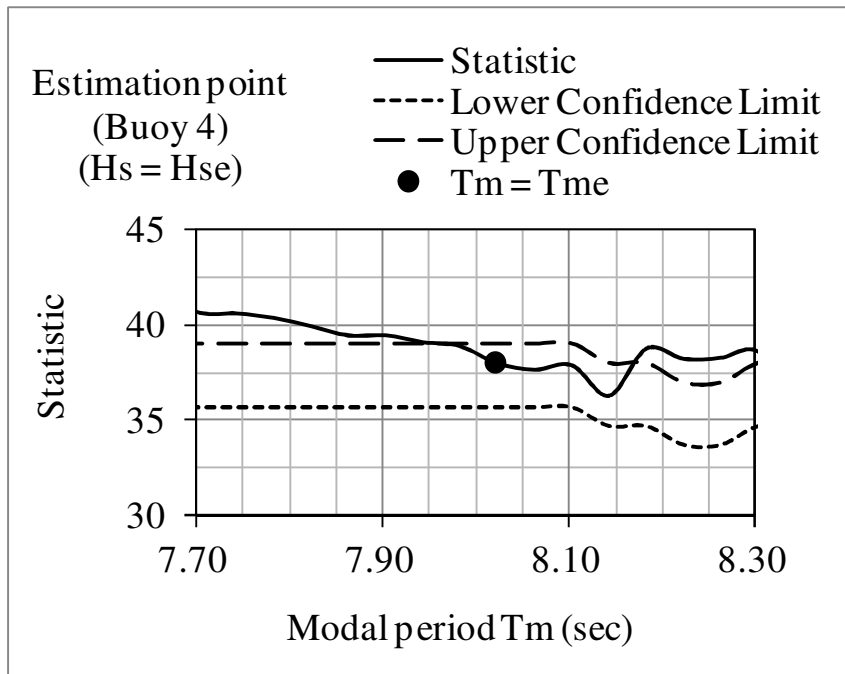
Figure 4-53. Two-sided confidence intervals at the 95% level of the estimation point (buoy 4) on the modal period T_m for the significant wave height $H_s = H_{se}$.

As mentioned in Chapter 1, Altunkaynak and Ozger(2005) provided a standard regional dependence function (SRDF) for assessing significant wave height. The standard regional dependence function (SRDF) is based on the point cumulative semivariogram (PCSV) modified by dividing it by the maximum point cumulative semivariogram value and subtracting from unity. A point cumulative semivariogram (PCSV) is a cumulative semivariogram (CSV) with a reference site of interest. The standard regional dependence function (SRDF) shows that locations in far distances have lower influence on the point of interest compared with the locations in close distances which have higher influence, which has the same idea as the inverse distance weight factors demonstrated in Chapter 3. Analysis of the locations defined in Figure 4-50 using the standard regional dependence function (SRDF) is presented in Table 4-17 to

136

compare with the estimations using the methodology proposed in this research.   The

estimated significant wave height for point 1 defined in Figure 4-50 is 2.34 m which has

an absolute relative error of 2.6 %, i.e. absolute value of (2.34 m - 2.4 m) / 2.4 m = 2.6 %.

Table 4-17. Estimation using standard regional dependence function (SRDF) for the locations defined in Figure 4-50.

Station ID	Significant wave height	Distance (km)	Standard distance (km)	Semivariogram (SV)	PCSV	Standard PCSV	SRDF
44008: Buoy 4 (Estimation point)	2.4 m	0	0				
44011: Buoy 3	2.8 m	233.118	0.24	0.080	0.080	0.08	0.92
44025: Buoy 2	1.5 m	823.597	0.84	0.405	0.485	0.49	0.51
44009:Buoy 1	1.4 m	979.54	1.00	0.500	0.985	1.00	0

5. Contributions, Limitations, and Future Work

5.1. Conclusions and Contributions

Risk-based methods are required for marine and maritime systems designs. Characterizing the statistical uncertainties associated with the system is essential for risk-based designs. The uncertainties for designing marine and maritime systems are embedded in the sea-state condition parameters and the modeling and prediction procedure. This study provides a statistical framework to characterize the sea-state conditions and associate uncertainties in confidence intervals on the estimated parameters.

Sea-state conditions are characterized by modal period and significant wave height which are two key parameters to represent the sea-state. At locations where the buoy elevation time histories are given, the estimation confidence intervals capture the modal period and significant wave height values, which verify and demonstrate the accuracy of the methodology. At locations of interest with no information provided, the sea-state conditions are interpolated from the nearby locations where the buoy data are given. Hypothesis testing and goodness-of-fit demonstrate the statistical features and uncertainties in the sea-state parameters, the wave model, and the characterization and prediction process. Verifications are taken place by utilizing a numerical wave simulation model called SWAN. Results show that the confidence intervals of the parameter estimations capture the values generated by SWAN model. That is, the proposed methodology is verified and demonstrated to provide accurate sea-state predictions. The other verification example uses sea-state properties based on historical

data obtained from the National Oceanic and Atmospheric Administration (NOAA) website. Spatial interpolation for the location of interest is presented in confidence intervals and verified by comparing with the NOAA observations. This example provides a verification and validation for the methodology.

Statistical and probabilistic methods are based on the assumption that the data are independent and representative. Current practices do not make a distinction between the number of discretization points for numerical computations and the number of sampling points, i.e. sample size needed for statistical analysis. Therefore, the correlation between data is discussed in this study. Approaches to estimate the sample size of independent observations are provided and examined. It is found that a series of independent samples has the interval between samples approximately the period of the series itself. In addition to characterizing the sea condition in time and frequency domains, the spatial interpolation techniques such as semivariogram analysis and Kriging estimation are discussed. In order to apply the semivariogram analysis and Kriging estimation, sufficient information on the field to establish the semivariogram model is required. The spatial interpolation procedure used in this research is compared with multiple existing methods reported in the literature. Comparisons show that the estimates reported herein have greater accuracy than the estimates by the existing methods. Moreover, the proposed estimators do not require as much information from the field as the existing methods.

This study provides methodologies for characterizing the sea-state conditions by estimating the sea wave parameters. The methodology applies on the locations with given wave properties and interpolates the locations of interest with unknown

information. The estimations take into account the uncertainties associated with the modeling and prediction processes and present the parameters characterizations in confidence intervals. Further, if the intervals between samples are too small, the data are most likely correlated. For intervals between samples too large, the information collected might not be sufficient. The interval between samples discussed in this study provides a guideline on how often, in time wise, to collect samples in order to obtain independent and representative data. Overall, the methodologies and discussions provided in this study can enhance the knowledge of the sea environment, provide statistical and probabilistic estimation framework, and improve the future risk-based marine and maritime designs.

5.2. Limitations and Future Work

This research is based on the assumption that the sea waves are stationary random processes. The sea-state condition characterization and prediction provided accurate estimations at the observed and unobserved locations. However, for extreme weather conditions such as storms, further examinations and modifications are required to ensure the achievement of accurate results at desired levels. In addition, the sea-state characterization methodology could be utilized for studies of wave-structure interactions. Further analysis is needed to ensure the applicability.

Appendix A.

Example output file from SWAN is shown as follows. The output quantities include the coordinates of the locations defined for obtaining output quantities, the spectral frequencies and variance densities for each location.

```
SWAN   1                        Swan standard spectral file, version
$  Data produced by SWAN version 40.81
$  Project: t5Trial1      ;  run number: t5
LOCATIONS                           locations in x-y-space
   49                           number of locations
    100.0000      0.0000
    100.0000    150.0000
    100.0000    300.0000
    100.0000    450.0000
    100.0000    600.0000
       .
       .
       .
    950.0000    600.0000
    950.0000    750.0000
    950.0000    900.0000
AFREQ                             absolute frequencies in Hz
   32                           number of frequencies
    0.0521
    0.0573
    0.0630
    0.0693
    0.0763
    0.0839
       .
       .
       .
    0.4241
    0.4665
    0.5132
    0.5645
    0.6209
    0.6830
    0.7513
    0.8264
    0.9091
    1.0000
```

```
QUANT
   3                      number of quantities in table
VaDens                       variance densities in m2/Hz
m2/Hz                    unit
  -0.9900E+02                exception value
CDIR                     average Cartesian direction in degr
degr                    unit
  -0.9990E+03                exception value
DSPRDEGR                      directional spreading
degr                    unit
  -0.9000E+01                exception value
LOCATION    1
 0.4528E-12  263.7  54.0
 0.2019E-09  265.3  54.0
 0.2319E-07  265.8  53.7
 0.9623E-06  266.6  53.3
 0.3515E-01  322.0  39.1
 0.2771E+00  322.4  37.1
 0.9987E+00  322.7  34.8
 0.2256E+01  323.2  31.8
 0.5719E+01  324.2  27.6
 0.9832E+01  325.1  25.1
 0.4393E+01  326.3  23.8
 0.2197E+01  327.9  22.8
 0.1484E+01  329.9  21.3
 0.1001E+01  331.6  20.0
 0.6622E+00  332.6  19.2
 0.4291E+00  333.2  18.8
 0.2728E+00  333.4  18.5
 0.1720E+00  333.4  18.4
 0.1078E+00  333.3  18.4
 0.6729E-01  333.0  18.5
 0.4193E-01  332.7  18.6
 0.2614E-01  332.4  18.7
 0.1632E-01  332.1  18.8
 0.1019E-01  331.8  18.9
 0.6369E-02  331.5  19.1
 0.3980E-02  331.1  19.3
 0.2487E-02  330.7  19.5
 0.1554E-02  330.2  19.7
 0.9717E-03  329.6  20.0
 0.6120E-03  328.9  20.3
 0.3875E-03  328.2  20.6
 0.2478E-03  327.2  20.9
```

Appendix B.

The mean values of the significant wave height and modal period data obtained from the NOAA website are provided as follows. The data are provided in the order of Buoy 1: Station ID 44009, Buoy 2: Station ID 44025, Buoy 3: Station ID 44011, and Buoy 4: Station ID 44008.

Buoy 1: Station ID 44009

```
STATION: 44009
1 - MONTHLY AND ANNUAL FREQUENCY AND CUMULATIVE PERCENT FREQUENCY (10THS)
ELEMENT: SIGNIFICANT WAVE HEIGHT (METERS) -- POR: (5/1986 - 12/2008) (177846 RECORDS,  88.2% HAVE ELEMENT)
```

	JAN		FEB		MAR		APR		MAY		JUN		JUL		AUG		SEP		OCT		NOV		DEC		ANN	
	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF
8.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	#
7.5	5	#	5	#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	10	#
7.0	6	999	14	999	-	-	-	-	-	-	-	-	-	-	-	-	1	#	1	#	1	#	-	-	23	999
6.5	14	999	8	998	-	-	-	-	1	#	-	-	-	-	-	-	7	999	1	999	5	999	2	#	38	999
6.0	15	998	12	998	-	-	-	-	6	999	-	-	-	-	-	-	10	999	5	999	8	999	3	999	59	999
5.5	18	997	18	997	9	#	1	#	4	999	-	-	-	-	3	#	7	999	2	999	8	999	8	999	78	999
5.0	24	996	16	995	29	999	7	999	8	999	-	-	-	-	13	999	29	998	9	999	28	998	11	999	174	999
4.5	41	994	40	994	51	997	43	999	10	999	-	-	2	#	17	999	55	997	68	999	65	996	33	998	425	998
4.0	67	991	61	991	116	994	105	996	36	998	3	#	4	999	32	998	91	993	128	995	105	992	75	996	823	995
3.5	193	986	144	986	266	986	179	989	134	996	2	999	1	999	84	996	138	987	206	987	116	984	201	991	1664	991
3.0	375	972	301	974	505	967	256	976	253	986	12	999	13	999	97	991	293	978	335	974	271	976	392	977	3103	981
2.5	815	945	728	950	965	933	593	958	439	969	124	999	98	999	178	986	551	960	577	953	763	957	846	950	6677	964
2.0	1696	886	1484	892	1793	866	1731	916	997	939	436	990	333	993	605	975	1317	925	1767	917	1616	903	1926	892	15701	926
1.5	3381	764	3110	772	3189	742	3125	793	2538	871	1618	961	1584	972	2226	940	3198	841	3431	806	3072	788	3438	759	33910	838
1.0	4582	520	3947	522	4689	522	5354	571	5698	698	6394	850	7196	876	7330	809	6223	637	5657	592	4745	570	4745	523	66560	648
0.5	2613	189	2529	205	2857	198	2675	190	4518	310	6011	413	7069	436	6442	378	3754	240	3789	239	3187	233	2797	196	48241	273
0.0	7	1	17	1	1	*	5	*	28	2	41	3	75	5	2	*	3	*	34	2	95	7	52	4	360	2
75PCTL	1.7		1.7		1.8		1.6		1.4		1.1		1.0		1.1		1.5		1.6		1.6		1.7		1.5	
50PCTL	1.2		1.2		1.2		1.1		1.0		0.8		0.8		0.8		1.0		1.1		1.1		1.2		1.0	
25PCTL	0.8		0.8		0.8		0.8		0.7		0.6		0.6		0.6		0.8		0.8		0.8		0.8		0.7	
MEAN	1.4		1.4		1.4		1.3		1.1		0.9		0.9		1.0		1.2		1.3		1.3		1.4		1.2	
S.D.	0.8		0.8		0.8		0.7		0.6		0.4		0.4		0.5		0.7		0.7		0.8		0.7		0.7	
TOTAL	13852		12434		14470		14074		14670		14641		16375		17029		15677		16010		14085		14529		177846	
MAX	7.6		7.7		5.7		5.4		6.3		4.1		4.5		5.5		6.8		6.9		6.9		6.4		7.7	
DATE	1992010414		2003021713		2004031100		2003040723		2008051218		2002060721		1996071313		1999083113		2006090205		2005102512		2006112301		2003120602		2003021713	
MIN	0.0		0.2		0.2		0.2		0.0		0.0		0.0		0.0		0.0		0.0		0.0		0.0		0.0	
DATE	1988011016		2007021307		1987030911		2007041116		1987053008		1994060914		1993071322		1990080816		1987091310		1993102518		1993112311		1997123018		1997123018	

(* < 0.05% , # = 100.0%)

STATION: 44009

1 - MONTHLY AND ANNUAL FREQUENCY AND CUMULATIVE PERCENT FREQUENCY (10THS)

ELEMENT: DOMINANT WAVE PERIOD (SECONDS) -- POR: (5/1986 - 12/2008) (177733 RECORDS, 88.1% HAVE ELEMENT)

	JAN		FEB		MAR		APR		MAY		JUN		JUL		AUG		SEP		OCT		NOV		DEC		ANN	
	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF
20	-	-	-	-	-	-	-	-	-	-	1	#	-	-	29	#	2	#	-	-	-	-	-	-	32	#
19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
17	30	#	12	#	19	#	13	#	5	#	41	999	48	#	52	998	198	999	28	#	46	#	22	#	514	999
16	-	-	1	999	23	999	1	999	30	999	38	997	1	997	26	995	-	-	3	998	-	-	-	-	123	997
15	2	998	1	999	38	997	3	999	48	998	60	995	4	997	9	994	31	987	16	998	1	997	12	998	225	996
14	147	998	110	999	369	994	176	999	160	994	129	990	82	997	316	993	710	985	412	997	169	997	350	998	3130	995
13	462	987	526	990	807	969	549	986	398	983	221	982	129	992	562	975	960	940	981	971	579	985	619	974	6793	977
12	88	954	95	948	127	913	90	947	60	956	19	967	41	984	32	942	95	879	108	910	117	943	46	931	918	939
11	1091	947	962	940	1449	904	1443	941	912	952	385	965	413	981	923	940	1432	873	1580	903	1317	935	900	928	12807	934
10	1398	869	1201	863	1597	804	1425	838	1356	890	788	939	607	956	969	886	1407	781	1700	805	1285	841	1118	866	14851	862
9	997	768	1149	766	1311	694	1604	737	1900	798	1655	885	1643	919	1538	829	1680	692	1704	698	1162	749	1229	788	17572	778
8	1409	696	1531	674	1875	603	2326	623	3417	668	3396	772	3519	819	3339	738	2598	584	2083	592	1573	667	1494	704	28560	679
7	1287	594	1257	551	1279	474	1848	458	1960	435	2485	540	3368	604	2853	542	1746	419	1140	462	974	554	1162	601	21359	519
6	2352	501	1895	449	1899	385	1864	327	1750	301	2555	370	3330	398	3109	375	1849	307	1943	390	2193	485	2692	520	27431	399
5	2904	331	2025	297	2140	254	1555	194	1383	182	1537	196	1860	195	2026	192	1793	189	2504	269	2767	328	2953	335	25447	244
4	1439	122	1343	134	1279	106	1023	84	1019	88	1049	91	1102	81	1072	73	1017	75	1505	113	1588	131	1601	131	15037	101
3	245	18	321	26	256	18	151	11	270	18	277	19	228	14	174	10	158	10	297	19	250	18	297	21	2924	17
2	-	-	5	*	1	*	3	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	*	10	*
75PCTL	9.1		9.1		10.0		10.0		9.1		8.3		8.3		9.1		10.0		10.0		10.0		9.1		9.1	
50PCTL	6.3		7.1		7.7		7.7		7.7		7.1		6.7		7.1		8.3		7.7		6.7		6.3		7.1	
25PCTL	5.0		5.0		5.3		5.9		5.9		5.9		5.9		5.9		5.9		5.3		5.0		5.0		5.6	
MEAN	7.2		7.3		7.9		7.8		7.7		7.3		7.1		7.5		8.3		7.9		7.3		7.2		7.5	
S.D.	2.6		2.6		2.8		2.5		2.4		2.2		2.0		2.4		2.9		2.8		2.7		2.7		2.6	
TOTAL	13851		12434		14469		14074		14668		14636		16375		17029		15676		16004		14021		14496		177733	
MAX	16.7		16.7		17.4		16.7		16.7		20.0		16.7		20.0		20.0		16.7		16.7		16.7		20.0	
DATE	2002012702		2004020305		2005031704		2000042509		1998052702		1992061802		1996071008		2003082917		1995091009		2002102513		2002110408		2002123023		2003082917	
MIN	2.5		2.1		2.4		2.4		2.5		2.5		2.5		2.6		2.5		2.5		2.6		2.4		2.1	
DATE	2004010211		2007021308		2008030308		2007041416		2007052221		2008060301		1993072312		2003082900		1997092000		2000101409		2008111310		2008122321		2007021308	

(* < 0.05% , # = 100.0%)

146

Buoy 2: Station ID 44025

```
STATION: 44025
1 - MONTHLY AND ANNUAL FREQUENCY AND CUMULATIVE PERCENT FREQUENCY (10THS)
ELEMENT: SIGNIFICANT WAVE HEIGHT (METERS) -- POR: (4/1991 - 12/2008) (142270 RECORDS,  78.4% HAVE ELEMENT)
```

	JAN F	CPF	FEB F	CPF	MAR F	CPF	APR F	CPF	MAY F	CPF	JUN F	CPF	JUL F	CPF	AUG F	CPF	SEP F	CPF	OCT F	CPF	NOV F	CPF	DEC F	CPF	ANN F	CPF
10.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	#
9.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	#	1	#
9.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
8.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	999	1	999
8.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	999	4	999
7.5	-	-	-	-	3	#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	999	4	999
7.0	1	#	-	-	6	999	-	-	-	-	-	-	-	-	-	-	1	#	-	-	-	-	6	999	14	999
6.5	8	999	-	-	10	999	-	-	-	-	-	-	-	-	-	-	2	999	-	-	1	#	2	999	23	999
6.0	5	999	6	#	9	998	-	-	-	-	-	-	-	-	1	#	-	-	1	#	-	-	12	999	34	999
5.5	10	999	15	999	18	998	5	#	1	#	-	-	1	#	1	999	-	-	7	999	5	999	22	998	85	999
5.0	35	998	24	998	23	996	8	999	2	999	-	-	2	999	-	-	11	999	29	999	11	999	24	996	169	999
4.5	56	994	32	996	87	994	10	999	5	999	-	-	2	999	1	999	22	999	71	997	59	998	69	993	414	998
4.0	110	989	72	993	116	987	34	998	7	999	-	-	4	999	4	999	53	997	104	991	129	993	110	987	743	995
3.5	204	978	139	986	220	978	113	995	49	999	2	#	4	999	31	999	89	993	197	983	230	981	227	978	1505	990
3.0	415	959	352	972	438	959	261	986	147	995	60	999	29	999	144	997	202	985	359	967	384	960	540	957	3331	979
2.5	889	919	938	939	941	923	733	964	404	984	164	995	110	997	296	985	492	969	609	938	706	925	1052	909	7334	956
2.0	1711	833	1725	849	1679	845	1242	902	959	953	485	982	463	988	546	961	1166	928	1232	889	1244	861	1591	815	14043	904
1.5	2535	669	2443	683	2586	706	2843	798	2512	878	1840	944	1946	953	1564	916	2709	832	2550	790	2497	747	2502	674	28527	805
1.0	2974	425	2774	448	3482	492	4102	560	5367	684	5789	798	6068	803	5078	787	4638	608	4435	586	3759	519	3167	451	51633	605
0.5	1412	139	1797	182	2341	204	2503	216	3434	270	4291	340	4338	336	4404	368	2693	224	2855	229	1921	176	1852	169	33841	242
0.0	34	3	100	10	131	11	75	6	62	5	-	-	33	3	58	5	23	2	1	*	8	1	39	3	564	4
75PCTL	2.0		1.9		1.9		1.6		1.4		1.2		1.1		1.2		1.5		1.6		1.8		2.0		1.6	
50PCTL	1.4		1.3		1.3		1.2		1.0		0.9		0.9		0.9		1.1		1.1		1.2		1.3		1.1	
25PCTL	1.0		0.9		0.8		0.8		0.7		0.7		0.7		0.7		0.8		0.8		0.9		0.9		0.8	
MEAN	1.5		1.5		1.5		1.3		1.1		1.0		1.0		1.0		1.2		1.3		1.4		1.5		1.3	
S.D.	0.8		0.8		0.9		0.7		0.6		0.4		0.4		0.5		0.6		0.8		0.8		0.9		0.7	
TOTAL	10399		10417		12090		11929		12949		12631		13000		12128		12101		12450		10954		11222		142270	
MAX	7.2		6.1		7.4		5.5		5.3		3.5		5.5		5.8		6.8		6.0		6.5		9.3		9.3	
DATE	1996010808		2003021716		1994030312		2007041602		2008051217		1994062800		1996071320		1991081916		1999091700		2005102513		1993112819		1992121116		1992121116	
MIN	0.0		0.0		0.0		0.0		0.0		0.3		0.2		0.0		0.0		0.0		0.2		0.0		0.0	
DATE	1992012717		2003022008		1993032706		1993040606		1998052321		2004062022		1994071809		1991081318		2003092917		2002100517		2000112512		1992122200		2003092917	

(* < 0.05% , # = 100.0%)

STATION: 44025

1 - MONTHLY AND ANNUAL FREQUENCY AND CUMULATIVE PERCENT FREQUENCY (10THS)

ELEMENT: DOMINANT WAVE PERIOD (SECONDS) -- POR: (4/1991 - 12/2008) (142186 RECORDS, 78.3% HAVE ELEMENT)

	JAN F	JAN CPF	FEB F	FEB CPF	MAR F	MAR CPF	APR F	APR CPF	MAY F	MAY CPF	JUN F	JUN CPF	JUL F	JUL CPF	AUG F	AUG CPF	SEP F	SEP CPF	OCT F	OCT CPF	NOV F	NOV CPF	DEC F	DEC CPF	ANN F	ANN CPF
25	1	#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	#
24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
20	-	-	-	-	-	-	-	-	-	-	1	#	-	-	-	-	3	#	-	-	-	-	-	-	4	999
19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
17	4	999	-	-	17	#	4	#	28	#	62	999	6	#	53	#	104	999	16	#	6	#	5	#	305	999
16	-	-	-	-	9	999	1	999	-	-	4	995	-	-	-	-	5	991	-	-	-	-	-	-	19	998
15	2	999	-	-	7	998	3	999	-	-	18	995	-	-	2	996	5	991	7	999	-	-	-	-	44	998
14	45	999	87	#	215	997	81	999	106	998	214	993	51	999	230	995	578	990	361	998	20	999	139	999	2127	997
13	220	995	275	992	418	979	210	993	153	990	178	976	104	996	293	976	739	943	597	969	298	998	300	987	3785	982
12	5	974	9	965	6	945	6	975	9	978	15	962	38	988	28	952	41	881	17	921	36	970	4	960	214	956
11	569	973	580	964	823	944	615	974	461	977	238	961	227	985	573	950	943	878	826	920	662	967	496	960	7013	954
10	900	919	1028	909	1179	876	919	923	934	941	456	942	339	967	549	903	955	800	1151	853	1035	907	845	916	10290	905
9	790	832	1145	810	1227	778	1554	846	1557	869	1092	906	938	941	901	857	1259	721	1321	761	979	812	1147	841	13910	833
8	1135	756	1328	700	1795	677	2695	715	3528	749	3149	820	3006	869	2528	783	1923	617	1697	655	1314	723	1355	738	25453	735
7	1074	647	922	572	1220	528	1445	489	1794	476	2181	570	3149	638	2145	574	1343	458	855	519	1001	603	1090	617	18219	556
6	2166	543	1745	484	1640	427	1670	368	1481	338	1965	398	2335	396	1972	397	1497	347	1658	450	1735	512	2145	520	22009	428
5	2068	335	1800	316	1886	291	1418	228	1417	223	1669	242	1437	216	1550	235	1470	223	2112	317	2171	353	2013	329	21011	273
4	1187	136	1185	143	1238	134	1002	109	1206	114	1133	110	1125	105	1055	107	1058	102	1525	147	1427	155	1363	150	14504	125
3	225	22	303	29	382	32	298	25	265	20	256	20	245	19	239	20	172	14	307	25	270	25	315	28	3277	23
2	-	-	-	-	1	*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	*
75PCTL	8.3		9.1		9.1		9.1		9.1		8.3		7.7		8.3		10.0		9.1		9.1		9.1		9.1	
50PCTL	6.3		6.7		7.1		7.7		7.7		7.1		6.7		7.1		7.7		7.1		6.3		6.3		7.1	
25PCTL	5.0		5.0		5.3		5.6		5.6		5.6		5.6		5.6		5.6		5.0		5.0		5.0		5.3	
MEAN	6.8		7.1		7.4		7.4		7.3		7.1		6.9		7.2		8.0		7.5		7.0		6.9		7.2	
S.D.	2.4		2.5		2.6		2.3		2.2		2.3		1.8		2.4		2.9		2.8		2.5		2.5		2.5	
TOTAL	10391		10407		12063		11921		12939		12631		13000		12118		12095		12450		10954		11217		142186	
MAX	25.0		14.3		17.4		16.7		16.7		20.0		16.7		16.7		20.0		16.7		16.7		16.7		25.0	
DATE	1994011111		2005020507		2008031900		2001041519		2007053103		1992061719		2007070319		2007082000		1996091417		2004100802		1998112521		1996122915		1994011111	
MIN	2.5		2.5		2.4		2.5		2.5		2.5		2.5		2.5		2.5		2.5		2.5		2.5		2.4	
DATE	2004011504		2001022721		2008031319		2007041201		2003051522		2007062422		1993071902		2003082910		1997092503		2006102800		2008111309		2008122322		2008031319	

(* < 0.05% , # = 100.0%)

148

Buoy 3: Station ID 44011

```
STATION: 44011
1 - MONTHLY AND ANNUAL FREQUENCY AND CUMULATIVE PERCENT FREQUENCY (10THS)
ELEMENT: SIGNIFICANT WAVE HEIGHT (METERS) -- POR: (5/1984 - 12/2008) (180611 RECORDS,  97.5% HAVE ELEMENT)
```

	JAN		FEB		MAR		APR		MAY		JUN		JUL		AUG		SEP		OCT		NOV		DEC		ANN	
	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF
14.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	#	-	-	1	#
13.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
13.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
12.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	999	-	-	1	999
12.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	#	1	999	-	-	2	999
11.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	999	-	-	-	-	3	999
11.0	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	2	#	4	999	2	999	-	-	8	999
10.5	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	999	-	-	-	-	5	999
10.0	-	-	1	#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	999	-	-	5	#	7	999
9.5	-	-	2	999	-	-	-	-	-	-	-	-	-	-	-	-	1	999	4	999	3	999	8	999	18	999
9.0	7	#	5	999	1	#	-	-	-	-	-	-	-	-	-	-	4	999	2	999	1	999	10	999	31	999
8.5	7	999	7	999	7	999	2	999	-	-	-	-	-	-	-	-	10	999	7	999	8	999	15	998	63	999
8.0	17	999	18	999	8	999	2	999	-	-	-	-	-	-	-	-	5	999	17	998	8	999	26	997	101	999
7.5	42	998	33	997	29	999	6	999	-	-	-	-	-	-	-	-	4	999	41	997	16	998	49	996	220	999
7.0	88	995	48	995	36	997	28	999	6	#	1	#	-	-	-	-	9	998	32	995	47	997	70	992	365	997
6.5	172	989	78	991	80	994	48	997	8	999	4	999	-	-	-	-	16	998	54	993	70	994	151	987	681	995
6.0	289	976	171	984	146	988	53	994	17	999	8	999	-	-	4	#	26	997	57	989	144	989	203	977	1118	992
5.5	416	956	322	971	217	977	90	991	33	998	21	999	3	#	7	999	36	995	107	986	223	980	299	962	1774	985
5.0	542	926	473	944	388	961	154	984	67	996	31	998	8	999	18	999	56	993	201	979	295	965	401	941	2634	976
4.5	670	888	626	906	518	933	255	974	124	992	37	996	13	999	25	998	96	989	296	967	459	945	556	913	3675	961
4.0	904	841	832	855	794	895	449	957	146	984	74	994	16	998	65	997	139	983	480	949	737	914	921	874	5557	941
3.5	1272	776	1095	788	1107	836	844	926	351	975	218	989	72	997	124	993	292	975	723	919	979	865	1311	809	8388	910
3.0	1676	686	1506	699	1575	755	1308	869	827	953	454	976	244	993	207	985	591	957	1104	875	1546	799	1634	717	12672	864
2.5	2045	568	1899	576	2062	639	1853	780	1288	902	826	948	573	972	1232	920	1811	807	2041	695	2144	603	2162	580	18406	793
2.0	2347	423	2103	422	2373	487	2852	655	2608	821	1933	898	1785	938	1495	936	2571	843	2959	695	2568	558	2494	452	28088	691
1.5	2503	256	2057	251	2382	313	3457	461	4276	659	4430	780	4227	827	3820	843	4277	682	4044	513	3049	385	2460	277	40982	536
1.0	1050	79	999	84	1659	138	2765	227	5151	392	6359	510	6693	562	6956	605	5154	415	3672	263	2002	180	1346	104	43806	309
0.5	62	5	34	3	209	16	578	39	1134	71	2019	123	2304	144	2752	172	1478	92	601	37	262	46	136	10	11569	66
0.0	2	*	2	*	3	*	2	*	1	*	-	-	-	-	3	*	2	*	1	*	420	28	-	-	436	2
75PCTL	3.6		3.5		3.2		2.6		2.0		1.7		1.6		1.5		1.9		2.5		3.0		3.4		2.5	
50PCTL	2.5		2.5		2.3		1.8		1.4		1.2		1.2		1.1		1.4		1.7		2.1		2.4		1.7	
25PCTL	1.7		1.7		1.6		1.3		1.0		0.9		0.9		0.8		1.0		1.2		1.4		1.7		1.1	
MEAN	2.8		2.8		2.5		2.1		1.6		1.4		1.3		1.3		1.6		2.0		2.3		2.7		2.0	
S.D.	1.4		1.3		1.3		1.1		0.8		0.7		0.6		0.6		0.9		1.2		1.3		1.4		1.2	
TOTAL	14111		12311		13594		14747		16037		16415		15997		16049		16001		16227		14883		14239		180611	
MAX	9.2		9.8		8.9		8.8		7.1		7.0		5.5		5.9		11.2		12.0		13.9		10.2		13.9	
DATE	2000012118		2004021915		1997030705		2007041618		2008051101		2006061516		1990073118		1998082918		2008092815		1991103016		2007110403		2004122707		2007110403	
MIN	0.0		0.0		0.0		0.0		0.0		0.3		0.3		0.0		0.0		0.0		0.0		0.0		0.0	
DATE	1998011515		2003020716		2002033000		2003041014		2003050104		1985062411		1989070215		2003082314		1998092906		1998101705		2003113023		1986123001		2003113023	

(* < 0.05% , # = 100.0%)

149

STATION: 44011
1 - MONTHLY AND ANNUAL FREQUENCY AND CUMULATIVE PERCENT FREQUENCY (10THS)
ELEMENT: DOMINANT WAVE PERIOD (SECONDS) -- POR: (5/1984 - 12/2008) (180173 RECORDS, 97.2% HAVE ELEMENT)

	JAN		FEB		MAR		APR		MAY		JUN		JUL		AUG		SEP		OCT		NOV		DEC		ANN	
	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF
20	-	-	-	-	-	-	-	-	11	#	11	#	-	-	24	#	7	#	3	#	-	-	-	-	56	#
19	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
17	8	#	-	-	4	#	8	#	10	999	69	999	11	#	100	999	136	999	63	999	10	#	31	#	450	999
16	-	-	-	-	-	-	-	-	4	999	30	995	-	-	3	992	8	991	3	996	4	999	-	-	52	997
15	-	-	-	-	7	999	2	999	26	998	46	993	25	999	19	992	34	991	4	996	6	999	-	-	169	997
14	74	999	63	#	137	999	97	999	33	997	116	990	85	998	218	991	604	988	336	996	61	999	174	998	1998	996
13	691	994	463	995	775	989	539	993	232	995	126	983	177	992	543	977	1061	951	661	975	443	994	580	986	6291	985
12	20	945	48	957	69	932	72	956	11	980	15	976	41	981	11	943	88	884	56	934	89	964	44	945	564	950
11	1434	944	1390	953	1507	927	1370	951	821	980	413	975	300	979	740	943	1386	879	1275	931	1473	958	1359	942	13468	947
10	2033	842	1718	840	2071	816	2110	858	1428	928	1009	950	589	960	871	897	1539	792	2018	852	1742	856	1615	846	18743	872
9	2183	698	1949	701	1990	664	2461	715	2445	839	2045	888	1605	923	1502	842	1811	696	2199	728	2056	735	1963	733	24209	768
8	3479	543	3179	543	3007	517	3602	548	5093	687	5051	764	5682	823	4304	749	3170	583	3419	592	3263	593	3652	595	46901	634
7	2045	297	1707	284	1890	296	1813	304	2464	369	3278	456	3779	468	3322	481	2352	385	2450	381	2293	367	2308	339	29701	373
6	1414	152	1140	146	1292	157	1483	181	2039	216	2605	256	2322	231	2573	274	2196	238	2207	230	1917	209	1556	176	22744	209
5	594	52	518	53	622	62	811	81	1026	88	1237	98	1000	86	1303	113	1220	100	1195	94	858	76	720	67	11104	82
4	127	9	124	11	195	16	331	26	368	25	336	22	353	24	447	32	358	24	314	21	223	17	221	17	3397	21
3	7	*	10	1	25	2	46	3	25	2	28	2	28	2	66	4	29	2	23	1	23	2	16	1	326	2
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	*
75PCTL	10.0		10.0		10.0		10.0		9.1		8.3		8.3		9.1		10.0		10.0		10.0		10.0		9.1	
50PCTL	8.3		8.3		8.3		8.3		7.7		7.7		7.7		7.7		8.3		8.3		8.3		8.3		8.3	
25PCTL	7.1		7.1		7.1		7.1		6.7		6.3		6.7		6.3		6.7		6.7		6.7		7.1		6.7	
MEAN	8.5		8.5		8.6		8.4		7.9		7.7		7.5		7.8		8.5		8.3		8.2		8.4		8.2	
S.D.	2.0		1.9		2.1		2.0		1.8		1.9		1.6		2.2		2.6		2.3		2.0		2.1		2.1	
TOTAL	14109		12309		13591		14745		16036		16415		15997		16046		15999		16226		14461		14239		180173	
MAX	16.7		14.3		16.7		16.7		20.0		20.0		16.7		20.0		20.0		20.0		17.4		16.7		20.0	
DATE	2001010411		2003022419		2002032012		2003041504		2002050619		1992061721		2000073106		1990081903		1996091419		1991103020		2007110405		1996122909		2002050619	
MIN	2.9		3.2		2.9		2.5		2.8		2.6		2.8		2.6		2.8		2.7		2.7		2.9		2.5	
DATE	1991010720		2008021721		1998031907		1989042720		1999051319		1991062402		1989070218		1990082801		1992090402		2008100606		1997112005		2007121017		1989042720	

(* < 0.05% , # = 100.0%)

150

Buoy 4: 44008

```
STATION: 44008
1 - MONTHLY AND ANNUAL FREQUENCY AND CUMULATIVE PERCENT FREQUENCY (10THS)
ELEMENT: SIGNIFICANT WAVE HEIGHT (METERS) -- POR: (8/1982 - 12/2008) (202771 RECORDS,  95.8% HAVE ELEMENT)
            JAN         FEB         MAR         APR         MAY         JUN         JUL         AUG         SEP         OCT         NOV         DEC         ANN
          F   CPF     F   CPF     F   CPF     F   CPF     F   CPF     F   CPF     F   CPF     F   CPF     F   CPF     F   CPF     F   CPF     F   CPF     F   CPF
 12.0     -    -      -    -      -    -      -    -      -    -      -    -      -    -      -    -      1    #      1    #      -    -      -    -      0    #
 11.5     -    -      -    -      -    -      -    -      -    -      -    -      -    -      1    #      1    #      -    -      -    -      -    -      2    #
 11.0     -    -      -    -      -    -      -    -      -    -      -    -      -    -      -    -      1   999     -    -      -    -      -    -      1   999
 10.5     -    -      -    -      -    -      -    -      -    -      -    -      -    -      -    -      -    -      -    -      2    #      1    #      3   999
 10.0     -    -      -    -      -    -      -    -      -    -      -    -      -    -      1   999     1   999     -    -      2   999      2   999      6   999
  9.5     4    #      -    -      -    -      -    -      -    -      -    -      -    -      -    -      2   999     5    #      -    -      4   999     15   999
  9.0     2   999     -    -      1    #      -    -      -    -      -    -      -    -      -    -      2   999     2   999     -    -      5   999     12   999
  8.5     4   999     -    -      2   999     -    -      -    -      -    -      -    -      1   999     3   999     3   999     1   999      7   999     21   999
  8.0    15   999     2    #      6   999     1    #      -    -      -    -      -    -      1   999     4   999     6   999     5   999      8   999     48   999
  7.5    18   998    17   999    18   999     6   999     -    -      -    -      -    -      -    -      4   999    12   999     6   999     13   998     94   999
  7.0    34   997    30   999    34   998    11   999     -    -      -    -      -    -      1   999     6   999    17   998    21   999     41   998    195   999
  6.5    64   995    45   996    35   996    23   999     -    -      -    -      2    #      1   999     5   999    36   997    39   998     47   995    297   998
  6.0   129   991    65   993    69   994    44   998     6    #      -    -      3   999     1   999     4   998    67   995    81   995     84   992    553   997
  5.5   154   983   120   989   126   990    81   995    14   999     -    -      2   999     7   999    11   998    92   992   118   991    167   987    892   994
  5.0   263   973   228   980   266   982   108   990    41   999     1    #      7   999     8   999    21   998   164   986   206   983    306   978   1619   989
  4.5   475   957   419   963   446   966   163   984    90   996     6   999     5   999    25   999    44   996   262   977   377   971    511   959   2823   981
  4.0   842   927   774   933   734   939   311   974   160   991    17   999     8   999    60   998   140   994   372   962   571   948    781   929   4770   968
  3.5  1233   874  1013   877  1092   894   661   955   248   982    77   999    25   999    66   995   245   986   627   941   940   914   1200   883   7427   944
  3.0  1676   797  1324   804  1525   828  1166   915   493   967   189   994    91   997   201   991   606   972  1075   905  1426   857   1724   812  11496   907
  2.5  2297   692  1959   708  2134   735  1755   845  1118   937   544   982   359   992   456   981  1025   938  1550   844  1883   771   2374   709  17454   851
  2.0  2927   548  2564   567  2860   606  2910   740  2007   871  1163   949  1128   973  1219   957  2254   880  2674   756  2730   657   2896   569  27332   765
  1.5  3383   364  2781   382  3527   432  4122   565  4308   752  3496   878  3403   912  3575   895  4186   753  4122   604  3668   492   3420   397  43991   630
  1.0  2171   152  2154   181  2935   217  4050   317  5975   496  7207   665  8418   727  8908   712  6402   518  5118   370  3608   270   2882   194  59828   413
  0.5   251    16   346    25   638    39  1213    73  2359   140  3687   225  4964   271  4962   255  2791   157  1404    80   827    52    392    23  23834   118
  0.0     1    *      1    *      -    -      1    *      -    -      -    -     26    1      3    *      -    -      -    -     26    2      -    -     58    *
 75PCTL     3.0         2.9         2.8         2.3         1.7         1.4         1.3         1.3         1.7         2.2         2.6         2.9         2.2
 50PCTL     2.1         2.0         1.9         1.6         1.3         1.0         1.0         1.0         1.2         1.5         1.8         2.0         1.4
 25PCTL     1.5         1.4         1.3         1.1         0.9         0.8         0.7         0.7         0.9         1.1         1.2         1.4         1.0
 MEAN       2.4         2.3         2.2         1.8         1.4         1.2         1.1         1.1         1.4         1.8         2.0         2.3         1.7
 S.D.       1.2         1.2         1.2         1.0         0.8         0.5         0.5         0.6         0.8         1.1         1.1         1.2         1.1
 TOTAL    15943       13842       16448       16626       16819       16387       18441       19497       17758       17608       16537       16865      202771
 MAX        9.7         8.0         8.8         7.8         6.2         4.9         6.7        11.4        11.5         9.6        10.3        10.7        11.5
 DATE 2000012607 1999022516 2005030909 1983040122 1986051217 1995060722 1996071402 1991081920 1999091705 1991103023 2007110323 1994122323 1999091705
 MIN        0.0         0.0         0.4         0.0         0.3         0.3         0.0         0.0         0.3         0.3         0.0         0.3         0.0
 DATE 2001010201 2000022002 2008031217 2000042601 2008052604 2008061502 2000072609 2005082721 1983091000 1994100813 1984112818 1998120523 2005082721
(* < 0.05% , # = 100.0%)
```

151

STATION: 44008
1 - MONTHLY AND ANNUAL FREQUENCY AND CUMULATIVE PERCENT FREQUENCY (10THS)
ELEMENT: DOMINANT WAVE PERIOD (SECONDS) -- POR: (8/1982 - 12/2008) (201301 RECORDS, 95.1% HAVE ELEMENT)

	JAN		FEB		MAR		APR		MAY		JUN		JUL		AUG		SEP		OCT		NOV		DEC		ANN	
	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF	F	CPF
25	-	-	-	-	-	-	-	-	-	-	-	-	1	#	-	-	-	-	-	-	-	-	-	-	1	#
24	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
22	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
21	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
20	-	-	-	-	-	-	-	-	-	-	16	#	-	-	27	#	4	#	-	-	-	-	-	-	47	999
19	-	-	1	#	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	999
18	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	999
17	-	-	3	999	2	#	3	#	11	#	79	999	36	999	76	999	189	999	41	#	2	#	32	#	474	999
16	-	-	-	-	15	999	1	999	20	999	57	994	4	998	22	995	18	989	4	998	5	999	-	-	146	997
15	2	#	1	999	17	999	3	999	53	998	73	991	45	998	13	994	53	988	16	997	2	999	-	-	278	997
14	67	999	29	999	102	998	66	999	39	995	146	986	104	995	326	993	703	985	451	997	33	999	168	998	2234	995
13	339	996	323	998	713	992	469	996	265	993	127	977	153	990	553	976	991	946	721	971	428	997	482	988	5564	984
12	40	974	73	974	155	948	96	967	46	977	29	970	27	981	52	948	160	890	111	930	116	970	58	958	963	957
11	1183	972	1243	969	1783	939	1413	962	934	974	481	968	392	980	934	945	1384	881	1501	924	1361	963	1087	954	13696	952
10	1710	898	1752	879	2061	831	2220	877	1454	919	931	938	587	959	1058	897	1585	803	2016	838	1601	877	1670	887	18645	884
9	2105	790	1917	753	2187	705	2865	743	2675	832	2079	882	1757	927	1972	843	2133	714	2243	724	1889	776	2112	783	25934	791
8	3838	658	3465	614	3575	572	4290	571	5026	673	4982	755	6081	831	4949	742	3698	593	3508	597	3430	657	4023	652	50865	662
7	2976	418	2377	364	2599	355	2272	313	2570	374	3274	451	4722	501	4095	488	2625	385	2716	397	2752	440	2980	403	35958	410
6	2345	231	1709	192	2056	197	1793	176	2118	222	2386	251	2870	245	3344	278	2489	237	2639	243	2611	267	2180	218	28540	231
5	1002	84	711	68	886	72	820	68	1156	96	1334	105	1316	89	1542	106	1361	97	1207	93	1184	102	993	83	13512	89
4	317	21	218	17	268	18	284	19	394	27	371	24	314	18	483	27	330	20	392	25	385	27	308	21	4064	22
3	18	1	19	1	25	2	30	2	58	3	22	1	16	1	41	2	33	2	40	2	44	3	33	2	379	2
2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0	*
75PCTL	9.1		9.1		10.0		10.0		9.1		8.3		8.3		9.1		10.0		10.0		9.1		9.1		9.1	
50PCTL	7.7		8.3		8.3		8.3		7.7		7.7		7.1		7.7		8.3		8.3		7.7		7.7		7.7	
25PCTL	6.7		6.7		6.7		7.1		6.7		6.3		6.7		6.3		6.7		6.7		6.3		6.7		6.7	
MEAN	8.0		8.2		8.4		8.3		7.9		7.7		7.5		7.8		8.5		8.3		7.9		8.1		8.0	
S.D.	1.9		1.9		2.1		1.9		1.9		2.0		1.7		2.2		2.6		2.3		2.0		2.0		2.1	
TOTAL	15942		13841		16444		16625		16819		16387		18425		19487		17756		17606		15843		16126		201301	
MAX	14.8		19.1		17.4		16.7		17.4		20.0		25.0		20.0		20.0		16.7		17.4		16.7		25.0	
DATE	2008012923		2004021112		2005031806		2000042710		2007052919		1992061721		1996070917		1990081900		1996091416		2002102507		2007110320		1996122917		1996070917	
MIN	2.9		2.6		2.6		2.6		2.5		2.9		2.9		2.6		2.9		2.8		2.6		2.6		2.5	
DATE	1985013123		1987021307		2005032321		1995041916		2007050611		2001062916		2005071218		2001082412		2003092315		1997100402		2003111123		1999121016		2007050611	

(* < 0.05% , # = 100.0%)

Bibliography

1. Akaike, H., 1981, "Statistical Information Processing System for Prediction and Control," Science Information Systems in Japan, 237-241.

2. Altunkaynak, A, 2005, "Significant Wave Height Prediction by Using A Spatial Model," Ocean Engineering, 32, 924-936.

3. Altunkaynak, A. and Ozger, M., 2005, "Spatial Significant Wave Height Variation Assessment and Its Estimation," Journal of Waterway, Port, Coastal, and Ocean Engineering, 131(6), 277-282.

4. Ayyub, B. M. and McCuen, R. H., 1990, "Optimum Sampling for Structural Strength Evaluation," Journal of Structural Engineering, 116(2), 518-535.

5. Barry, R. P. and Herf, J. M. V., 1996, "Blackbox Kriging: Spatial Prediction without Specifying Variogram Models," Journal of Agricultural, Biological, and Environmental Statistics, 1(3), 297-322.

6. Bendat, J. S. and Piersol, A. G., 1986, "Random Data: Analysis and Measurement Procedures," 2nd Edition, John Wiley & Sons, N. Y.

7. Booij, N., Ris, R. C., and Holthuijsen, L. H., 1999, "A Third-Generation Wave Model for Coastal Regions: 1. Model Description and Validation," Journal of Geophysical Research, 104(C4), 7649-7666.

8. Booij, N., Holthuijsen, L., and Battjes, J., 2001, "Ocean to Near-Shore Wave Modelling with SWAN," Coastal Dynamics, Proceedings of the 4th Conference on Coastal Dynamics, Lund, Sweden, 335-344.

9. Chatfield, C., 2004, "The Analysis of Time Series: An Introduction," 6th Edition, Chapman & Hall/CRC.

10. Cruz, J. M.B.P., and Sarmento, A. J.N.A., 2007, "Sea State Characterisation of the Test Site of An Offshore Wave Energy Plant," Ocean Engineering, 34, 763-775.

11. Donelan, M., and Pierson, W. J., 1983, "The Sampling Variability of Estimates of Spectra of Wind-Generated Gravity Waves," Journal of Geophysical Research, 88(C7), 4381-4392.

12. Ebuchi, N., Toba, Y, and Kawamura, H., 1992, "Statistical Study on the Local Equilibrium between Wind and Wind Waves by Using Data from Ocean Data Buoy Stations," Journal of Oceanography, 48, 77-92.

13. Ferreira, J. A., and Guedes Soares, C., 2000, "Modelling Distributions of Significant Wave Height," Coastal Engineering, 40(4), 361-374.

14. Ferreira, J. A., and Guedes Soares, C., 2002, "Modelling Bivariate Distributions of Significant Wave Height and Mean Wave Period," Applied Ocean Research, 24, 31-45.

15. Forristall, G. Z., Heideman, J. C., Leggett, I. M., Roskam, B. and Vanderschuren, L., 1996, "Effect of Sampling Variability on Hindcast and Measured Wave Heights," Journal of Waterway, Port, Coastal, and Ocean Engineering, 122(5), 216-225.

16. Goda, Y., 1976, "On Wave Groups," Proceedings of the International Conference on Behavior of Offshore Structures, Trondheim, Norway, 1-14.

17. Goda, Y., Konagaya, O., Takeshita, N.. Hitomi, H., and Nagai, T., 2000, "Population Distribution of Extreme Wave Heights Estimated Through Regional Analysis," Coastal Engineering, Proceedings of the 27th International Conference on Coastal Engineering (ICCE 2000), Sydney, Australia, 1078-1091.

18. Goda, Y., 2004, "Spread Parameter of Extreme Wave Height Distribution for Performance-Based Design of Maritime Structures," Journal of Waterway, Port, Coastal, and Ocean Engineering, 130(1), 29-38.

19. Goff, J. A., 2009, "Statistical Characterization of Geosat Altimetry Noise: Dependence on Environmental Parameters," Geochemistry, Geophysics, Geosystems, 10(8), 1-12.

20. Guedes Soares, C., 1990, "Effect of Spectral Shape Uncertainty in the Short Term Wave-Induced Ship Responses," Applied Ocean Research, 12(2), 54-69.

21. Guedes Soares, C. and Carvalho, A. N., 2003, "Probability Distributions of Wave Heights and Periods in Measured Combined Sea-States from the Portuguese Coast," Journal of Offshore Mechanics and Arctic Engineering, 125, 198-204

22. Guedes Soares, C. and Cherneva, Z., 2005, "Spectrogram Analysis of the Time-Frequency Characteristics of Ocean Wind Waves," Ocean Engineering,32(14-15), 1643-1663

23. Hamilton, J., Hui, W. H., and Donelan, M. A., 1979, "A Statistical Model for Groupiness in Wind Waves," Journal of Geophysical Research, 84(C8), 4875-4884.

24. Hamilton, LJ., 2010, "Characterising Spectral Sea Wave Conditions with Statistical Clustering of Actual Spectra," Applied Ocean Research, 32(3), 332-342.

25. Hatori, M., 1984, "Nonlinear Properties of Laboratory Wind Waves at Energy Containing Frequencies: Part 1. Probability Density Distribution of Surface Elevation" Journal of the Oceanographical Society of Japan, 40, 1-11.

26. Hou, Y., Guo, P., Song, G., Song, J., Yin, B., and Zhao, X., 2006, "Statistical Distribution of Nonlinear Random Wave Height," Science in China: Series D Earth Science, 49(4), 443-448.

27. Hughes, O. F., 1988, "Ship Structural Design- A Rational-Based, Computer-Aided Optimization Approach," The Society of Naval Architects and Marine Engineers.

28. IAHR, 1989, "List of Sea-State Parameters," Journal of Waterway, Port, Coastal, and Ocean Engineering, 115(6), 793-808.

29. Jenkins, G. M. and Watts, D. G., 1968, "Spectral Analysis and Its Applications," Holden-Day, San Francisco, C. A.

30. Jensen, G. A. and Vesecky, J. F., 1993, "Non-Fourier Spectral Characterizations of the Ocean Surface: New Theoretical Models of Ocean Surface Radar Scattering," OCEANS '93 Engineering in Harmony with Ocean Proceedings, 2, 19-24.

31. Kazeminezhad, M. H., Etemad-Shahidi, A. and Mousavi, S. J., 2005, "Application of Fuzzy Inference System in the Prediction of Wave Parameters," Ocean Engineering, 32, 1709-1725.

32. Kimura, A., 1980, "Statistical Properties of Random Wave Groups," Proceedings of the 17th International Conference on Coastal Engineering, 3, 2955-2973.

33. Longuet-Higgins, M. S., 1975, "On the Joint Distribution of the Periods and Amplitudes of Sea Waves," Journal of Geophysical Research, 80(18), 2688-2694.

34. Longuet-Higgins, M. S., 1980, "On the Distribution of the Heights of Sea Waves: Some Effects of Nonlinearity and Finite Band Width," Journal of Geophysical Research, 85(C3), 1519-1523.

35. Longuet-Higgins, M. S., 1984, "Statistical Properties of Wave Groups in a Random Sea State," Phil. Trans. R. Soc. Lond., Ser. A, 312(1521), 219-250.

36. Matheron, G., 1963, "Principles of Geostatistics," Economic Geology, 58, 1246-1266.

37. Mathisen, J. and Bitner-Gregersen, E., 1990, "Joint Distributions for Significant Wave Height and Wave Zero-up-crossing Period," Applied Ocean Research, 12(2), 93-103.

38. McCuen, R. H. and Snyder, W. M., 1986, "Hydrologic Modeling: Statistical Methods and Applications," Prentice-Hall, Englewood Cliffs, N. J.

39. McCuen, R. H., Aggour, M. S. and Ayyub, B. M., 1988, "Spacing for Accuracy in Ultrasonic Testing of Bridge Timber Piles," Journal of Structural Engineering, 114(12), 2652-2668.

40. Ochi, M. K. and Sahinoglou, I. I., 1989(1), "Stochastic Characteristics of Wave Groups in Random Seas. Part 1; Time Duration of and Number of Waves in A Wave Group," Applied Ocean Research, 11(1), 39-50.

41. Ochi, M. K. and Sahinoglou, I. I., 1989(2), "Stochastic Characteristics of Wave Groups in Random Seas; Part 2: Frequency of Occurrence of Wave Groups," Applied Ocean Research, 11(2), 89-99.

42. Ris, R. C., Holthuijsen, L. H., Booij, N., Andorka Gal, J. H., and de Jong, J. C. M., 1997, "The SWAN Wave Model Verified Along the Southern North Sea Coast," Ocean Wave Measurement and Analysis, 49-63.

43. Rodriguez, G., and Guedes Soares, C., 1999, "The Bivariate Distribution of Wave Heights and Periods in Mixed Sea States," Journal of Offshore Mechanics and Arctic Engineering, 121, 102-108.

44. Rodriguez, G. R., Guedes Soares, C. and Ferrer, L., 2000, "Wave Group Statistics of Numerically Simulated Mixed Sea States," Journal of Offshore Mechanics and Arctic Engineering, 122, 282-288.

45. Rodriguez, G. R., 2001, "Fitting the Long Term Bivariate Distribution of Wave Heights and Periods to A Theoretical Model," Ocean Wave Measurement and Analysis, Proceedings of the Fourth International Symposium Waves 2001, San Francisco, CA, 434-443.

46. Rodriguez, G. R. and Guedes Soares, C., 2001, "Correlation between Successive Wave Heights and Periods in Mixed Sea States," Ocean Engineering, 28, 1009-1030.

47. Rodriguez, G., Guedes Soares, C., Pacheco, M., and Perez-Martell, E., 2002, "Wave Height Distribution in Mixed Sea States," Journal of Offshore Mechanics and Arctic Engineering, 124, 34-40.

48. Sen, Z., 1989, "Cumulative Semivariogram Models of Regionalized Variables," Mathematical Geology, 21(8), 891-903.

49. Sen, Z., 1992, "Standard Cumulative Semivariograms of Stationary Stochastic Processes and Regional Correlation," Mathematical Geology, 24(4), 417-435.

50. Sen, Z. and Sahin, A. D., 2001, "Spatial Interpolation and Estimation of Solar Irradiation by Cumulative Semivariograms," Solar Energy, 71(1), 11-21.

51. Sobey, R. J. and Read, W. W., 1984, "Wave Groups in the Frequency and Time Domains," Nineteenth Coastal Engineering Conference, 695-707.

52. Sobey, R. J. and Young, I. R., 1986, "Hurricane Wind Waves-A Discrete Spectral Model," Journal of Waterway, Port, Coastal and Ocean Engineering, 112(3), 370-389.

53. Sobey, R. J., 1992, "The Distribution of Zero-Crossing Wave Heights and Periods in A Stationary Sea State," Ocean Engineering, 19(2), 101-118.

54. Wei, W. W. S., 2005, "Time Series Analysis: Univariate and Multivariate Methods," 2nd Edition, Addison Wesley.

55. White, G. J. and Ayyub, B. M., 1990, "Semivariogram and Kriging Analysis in Developing Sampling Strategies," First International Symposium on Uncertainty Modeling and Analysis Proceedings, 360-365.

56. Wist, H. T., Myrhaug, D. and Rue, H., 2004, "Statistical Properties of Successive Wave Heights and Successive Wave Periods," Applied Ocean Research, 26, 114-136.

57. Young, I. R. and Sobey, R. J., 1981, "Wave Prediction Techniques in the Nearshore Environment," Conferences on Environmental Engineering, 154-158.

58. Young, I. R., 1995, "The Determination of Confidence Limits Associated with Estimates of the Spectral Peak Frequency," Ocean Engineering, 22(7), 669-686.

59. SWAN (Simulating WAves Nearshore), www.swan.tudelft.nl.

60. NOAA (National Oceanic and Atmospheric Administration), http://www.ndbc.noaa.gov.