STATISTICAL FORECASTING MODEL OF SOLITON EVENTS IN THE SULU SEA

by

Lee Han Way

15114

Dissertation submitted in partial fulfilment of

The requirements for the

Bachelor of Engineering (Hons)

(Civil)

MAY 2014

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan

CERTIFICATION OF APPROVAL

Statistical Forecasting Model of Soliton Events in the Sulu Sea

by

Lee Han Way
15114

A project dissertation submitted to the
Civil Engineering Programme
Universiti Teknologi PETRONAS
In partial fulfilment of the requirement for the
BACHELOR OF ENGINEERING (Hons)
(CIVIL)

Approved by,

_____
Ir. Dr. M. Shahir Liew

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
May 2014

CERTIFICATION OF ORIGINALITY

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

_____
Lee Han Way

ABSTRACT

Sulu Sea is one of the areas heavily infested by soliton events, by which offshore and subsea exploration as well as operation can be heavily affected. In spite of numerous researches being done on this issue, its occurrence continues to impact the mentioned activities. As one of the potential oil and gas regions in the Southeast Asia, a statistical forecasting model for the soliton events is necessary to reduce downtime loses and onsite casualties. Tidal wave, responsible for the soliton events in this region, is highly seasonal, hence the exponential smoothing method is opted to generate the forecast model. Current speeds at 3 different water depths at Karupang have been measured and recorded at a constant interval of 10 minutes for approximately 109 days, from 1st September 2012 to 18th December 2012. The current speed forecast is first developed using the simple seasonal exponential smoothing method, as a parsimoniousy model is desired. Level and seasonality smoothing constant α and δ of 0.50 and 0.59 are chosen by numerical methodtrial and error. The seasonality of the model, 24 hours and 15 minutes, is obtained from the non-correlated power spectrum density plot. The forecasted values are compared with the actual observed current speed, and it is shown that the percentage of accuracy and the goodness of fit of the forecast model for 1 seasonal cycle of 3 different water depths are approximately 70%. Despite literaturepopular belief, no clear correlation can be observed between high speed current events and the occurrence of solitons in this study. Nonetheless, the accuracy and robustness of this model can potentially contribute to the oil and gas industryprediction of solitons.

Comment [S1]: grammar

Comment [S2]: parsimonious

Comment [S3]: numerical method!!!

Comment [S4]: remove

Comment [S5]: Despite LITERATURE!

Comment [S6]: Only for this stiudy

Comment [S7]: Prediction of solitons…don't be so heng…

ACKNOWLEDGEMENT

Over The duration of this project, I feel grateful for having AP Ir Dr Shahir Liew as my assigned university Final Year Project Supervisor. His guidance throughout the 8-month research project has assisted me in my progress.

Also, I would like to express my gratitude to PhD student, Mr Lim Eu Shawn for his continuous advice and guidance as my project advances. Besides, I would like to thank the outstanding researchers in the Universiti Teknologi PETRONAS Offshore Engineering Centre Unit (UTP OECU) for their support.

Last but not least, my appreciation to all my family and friends for their unconditional support and encouragement throughout the 8-month period.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background of Study

It was observed in the Malay Archipelago regions where narrow lines of propagating whitecaps appeared when ships sailed by (Wallace, 1869). Wallace also noted the occurrence of breaking waves on otherwise calm beaches. This phenomenon can be related to the internal gravity waves, by proving the existence of an 80 meters crest-to-trough internal wave through a mechanical bathythermographic experiment in the Andaman Sea (Perry and Schimke, 1965).

Throughout the second half of the seventies, Apel and his co-workers had carried out a series of researches on aerial observations of quasi-periodic, highly coherent variations of surface roughness. Their researches and experiments significantly indicated that the observations were caused by the underlying internal gravity waves (Apel *et al*., 1985). These internal gravity waves are usually highly nonlinear, behave in the form of "solitary waves" or "solitons" (Apel *et al*., 2006).

The impacts of solitons on offshore operation such as exploration drilling had been observed in many regions in the past few decades, especially in offshore West Africa and various parts of Southeast Asia (Jeans, 2013). According to Jeans, over compensation of dynamical positioning, large tilts and horizontal displacements beyond the water circle and excessive mooring line tensions are impacts of solitons on exploration drilling.

1.2. Problem Statements

1.2.1. Problem Identification

Research found out that the Sulu Sea is heavily infested by the soliton. Sulu Sea is one of the regions where the oil and gas exploration will be active. The presence of soliton however may cause damages to vessels, remotely operated vehicle (ROV), oil rigs and also subsea operations. Besides exploration, solitons affect operations such as vessel manoeuvres, installations and offloading. The sudden onset of a soliton current may bring severe impact on floating structures. The soliton event not only impairs assets owned by the operators, it may also lead to onsite casualties. Even though comprehensive researches had been done on this problem, solitons still bring impacts onto drilling operation in this region.

While the occurrence of soliton brings much impact to the offshore oil and gas activities, various efforts had been taken to predict or model this event. For example, the Massachusetts Institute of Technology General Circulation Model (MITgcm) is able to provide short term in-situ data. In addition to that, the Soliton Early Warning System from Fugro GEOS, a metocean service provider, is able to provide real-time prediction of soliton occurrence. However a minimum 10 hours soliton warning period is necessarily. The methods discussed above both predicts the solitons numerically. As the occurrence of soliton events is dependent on several physical oceanic conditions, a physical model to describe its nature is highly complicated. In contrary, a statistical approach is more appropriate to represent the highly stochastic soliton events.

1.2.2. Significance of Project

Sulu Sea is one of the most active region in terms of oil and gas exploration in the Southeast Asia. The most significant metocean phenomenon in the Sulu Sea is the occurrence of solitons. In order to reduce downtime loses and potential casualties caused by the soliton events, a statistical forecasting model on the soliton events is to be developed to support the offshore operation in the mentioned region.

1.3. Objectives

Based on the background research presented above, few objectives have been listed out to demarcate the direction of this research project:

a) To identify and characterize the statistical component of a soliton event.
b) To develop a statistical forecast model of soliton events in Sulu Sea.

1.4. Scope of Studies

The scope of studies will cover different types of soliton forecast mechanisms.

As a baseline, the exponential smoothing models will be studied to produce a parsimonious forecasting model. The types of for exponential smoothing models that will be covered are:

a) Simple Seasonal
b) Holt-Winters' Additive
c) Holt-Winters' Multiplicative

2. LITERATURE REVIEW

Soliton, or internal solitary wave, is one of the most common types of internal waves. Other internal waves include linear internal waves, linear baroclinic tide and baroclinic planetary Rossby wave.

A soliton moves along a pycnocline in the ocean. A pycnocline is the boundary between two stratified oceanic layers of water masses due to the change in density, often caused by the change in temperature and salinity. It is highly nonlinear, as governing variables such as current speed, current direction and wave frequency do not change at the same rate as the particle displacement, velocity as well as acceleration do, which is why a soliton is also known as a nonlinear internal wave. Solitons are non-sinusoidal because they do not repeat themselves indefinitely. In fact the occurrence of solitons is categorized as a stochastic process, or more commonly known as random process, because the governing variables discussed above also behave stochastically. Bathymetric as well as barotropic/baroclinic conditions also determine its occurrence. As long as all these conditions are present, a soliton propagates steadily whilst preserving its shape.

2.1. The Generation of Solitons

Soliton can be generated in several ways, some involving pycnocline shift whilst others due to conversion of tidal energy into pycnocline motion. The different ways of generation are discussed as follows.

2.1.1. Lee Wave Generation

The generation of soliton at a bank or an underwater sill has been explained using the lee wave mechanism (Jackson, da Silva and Jeans, 2012). The theory of supercritical flows has been used to describe this mechanism. The theory of supercritical flows uses internal Froude number to differentiate the hydraulic state of a stratified shear flow. The Froude number is defined as:

$$F = \frac{u_f}{c} \qquad (2.1.1\text{- }1)$$

where u_f = fluid velocity and

c = small-amplitude internal wave velocity.

When the Froude number is below 1, the flow is subcritical. Wave is able to move against the direction of flow as the internal wave velocity, c is greater

than fluid velocity (u_f). The flow on the other hand is supercritical when the Froude number is greater than 1. In this condition, wave and fluid move in unidirectional manner. When Froude number, F is equal to 1, flow condition is critical. The disturbance at the boundary remains stationary. It will gain energy from resonance. Critical flow can happen at the transition between supercritical and subcritical conditions when there is a sudden change in depth.

At a bank or sill, the sloping bathymetry displaces the pycnocline by imparting a vertical component to the stratified flow. Minimum depth change is also necessary to cause the transition between subcritical and supercritical flow. At this point, a stationary lee wave is generated in the pycnocline downstream of the sill. It acquires high phase velocity relative to the flowing water to remain stationary relative to the bank. The lee wave then flows over the sill and continues propagating after the flow slackens and becomes subcritical. In accordance with Korteweg-de Vries theory, the disturbance eventually turns into a series of solitons.

2.1.2. Internal Tide Evolution

The lee wave and internal tides are some of the normal occurrences due to tide-topography interaction (Jackson *et al*., 2012). An internal tide is made up of a linear internal wave at tidal period on the boundary between two stratified layers. Analogous to the lee wave mechanism, when there is a sudden drop in the bathymetric condition, the stratified layer of the water is displaced vertically downwards due to gravitational force. At baroclinic condition, due to hydroelasticity or flexible fluid-structure interaction (FSI), the pycnocline "bounces" and induces a baroclinic tide. Upon acquiring enough energy, the internal tide will steepen and break, changing into solitons.

Numerous studies had been done on the relationship between internal tide and soliton generations. It is found that in some locations, solitons propagate together with internal tides, maintaining a fixed phase. The soliton appears to be "trapped" within the trough of the tidal wave. It was observed that the solitons and the internal tide remained in phase even after three to four tidal cycles (Jackson *et al*., 2012).

2.1.3. Tidal Beam Generation

In 1992, generation of solitons from the interaction between internal tide linear internal waves was witnessed by New and Pingree in the central Bay of Biscay (Jackson *et al*., 2012). Continuous stratification yields a waveguide, allowing internal tidal energy to travel through the ocean vertically.

Airy wave theory is able to predict the occurrence of multiple vertical wave modes superposing on one another at tidal frequency, forming tidal beams (Gerkema and Zimmerman, 2008). Energy travels and particles accelerate along these beams. Energy propagates vertically along the beams and is reflected by the ocean surface and the seabed, comparable to how a laser beam is being reflected between two parallel mirrors. The tidal beams travel in a curve whose gradient varies with depth due to inconsistent stratification. Normally, when it is near to the seabed, stratification is the strongest, hence highest curvature is observed. On the other hand, curvature is least when stratification is less obvious, especially when it is near to the surface of the ocean. If the stratification is uniform, reflection of the beams along straight line can be observed.

The slope of the bathymetry may match the angle of the tidal beam at the continental shelf edges. At these "critical" slopes, a downward beam can be generated. It travels to the seabed and will be reflected upwards, affecting the pycnocline far away from the shelf. A study revealed that gigantic solitons at central Bay of Biscay occurred at where the thermocline intersected with internal tidal beam generated from the edge of a continental shelf approximately 150km away.

2.1.4. Resonant Generation

Research found that a stratified flow can be caused by not only vertical but also horizontal bathymetric contraction or minor unevenness in the seabed, leading to critical flow and the generation of solitons. Horizontal contractions in the ocean are less common than sills or banks, hence the relationship between the bathymetric change and soliton generation is not quite understood. However, weakly nonlinear Korteweg-de Vries equation had been used and "resonantly generated" upstream-propagating solitons had been predicted.

Based on the outcome of the studies, solitons can be generated in a plane channel with mildly varying lateral dimension when the flow is almost critical.

2.1.5. Plume Generation

Studies carried out by Nash and Moum (2005) and Matthews *et al*. (2011) had proven that plumes due to a throughflow current or river outflow can cause solitons (Jackson *et al.*, 2012). Existing coastal waters is interrupted by the fresh water discharged from the rivers in tidally modulated pulses. The pluming effect of the fresh water is determined by the initial momentum of the plume and the wind and current conditions at the moment. The plume flows at a higher velocity compared to that of the ocean water, hence causing a convergence at the leading edge. The plume, with a higher velocity, portrays as a supercritical flow. The internal wave resulted by the displacement on the other hand, has a lower velocity, and is hence trapped at the leading edge of the plume. As the plume loses speed, the internal wave, after separating from the plume's edge, propagates as a free wave.

2.2. Korteweg-de Vries Equation

The Korteweg-de Vries (or KdV) equation is a mathematical model of water waves. It is assumed that there is no dissipation of energy and the waves propagate indefinitely. In 1877, this equation was introduced by Joseph Valentin Boussinesq and was rediscovered by Diederik Johannes Korteweg and Gustav de Vries in 1895.

Theoretically the motion of a soliton can be described by using the 1-soliton KdV equation, which is defined as:

$$\phi(x,t) = \frac{1}{2} csech^2[\frac{\sqrt{c}}{2}(x - ct - a)]$$

where x = displacement**,**

t = time,

c = phase speed,

sech stands for hyperbolic secant and a is an arbitrary constants. This equation describes a soliton moving to the right whilst maintaining its fixed wave form.

2.3. Solitons in the Sulu Sea

In 1985, an exclusive study was done on the soliton events in the Sulu Sea. A research had been carried out on both in-situ experiment and theoretical studies. Satellite images of internal waves were captured using Synthetic Aperture Radar (SAR) and Defense Meteorological Satellite Program (DMSP) to provide clear pictures of the internal waves' nature throughout the research. Modified Korteweg-de Vries equation was also applied to simulate the behaviour of internal waves (Apel *et al*., 1985).

The internal waves in the Sulu Sea are formed at the sill between Doc Can Island and Pearl Bank, where, akin to the lee waves' formation mechanism, an internal hydraulic jump at the sill region was generated by a strong ebbing flow (southward) over the sill at Pearl Bank (Apel *et al.*, 1985; Liu *et al.*, 1985).

At this point, the internal tide weakens and a thermocline travels northward over the sill. As nonlinear and dispersive effects balance each other, soliton is formed on the depression hump. A train of solitons is then formed after travelling for nearly 200 km.

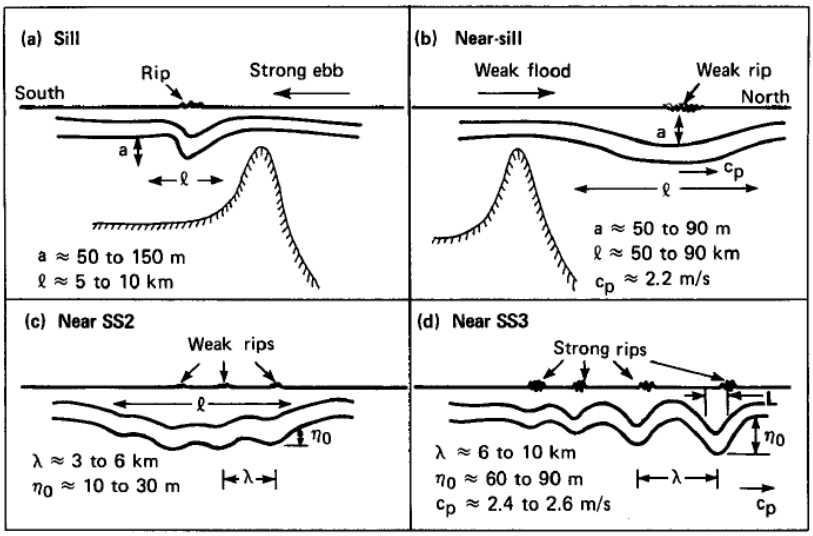


Figure 1: Schematic figure of soliton generation in the Sulu Sea: a) Ebbing tide flows southward over a submarine sill in Pearl Bank, induces a hydraulic jump, stationary lee wave is formed. b) Ebbing tidal flow slackens and eventually turns flooding, lee wave steepens nonlinearly and travels northward, moves over the sill. c) Nonlinear and dispersive effects take place and begin to balance, consistent with the KdV theory. d) After travelling for approximately 200 km, a train of solitons is fully developed (Apel *et al*., 1985).

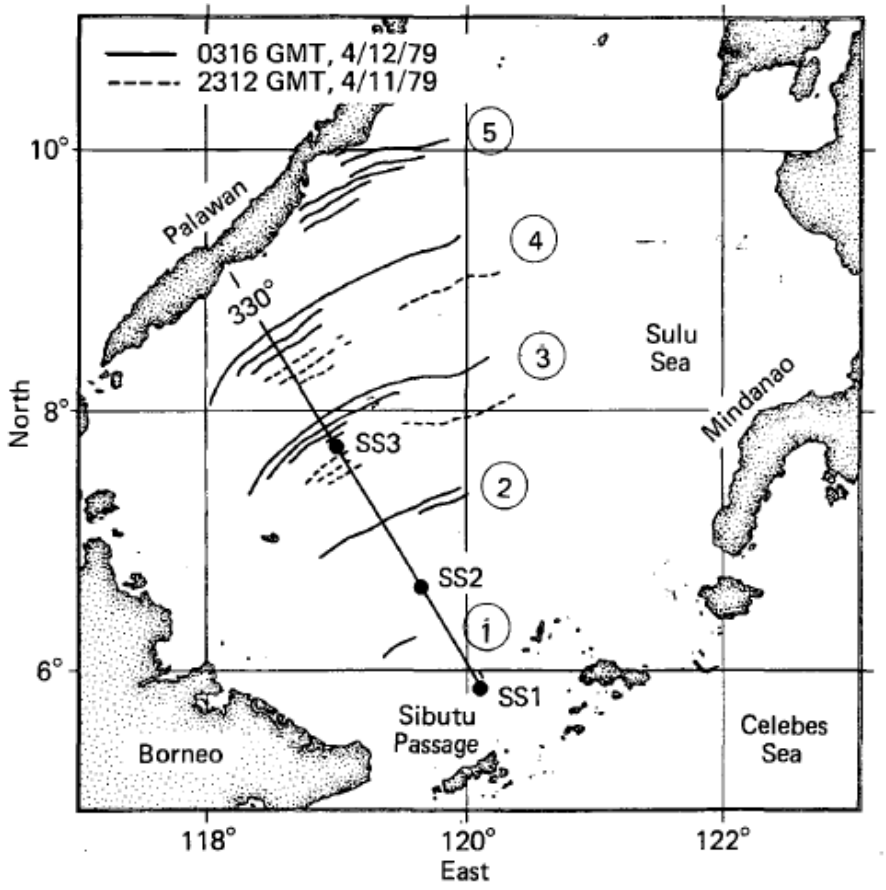


Figure 2: Line drawing rendition of DMSP image, illustrating the solitons observed in the Sulu Sea (Apel *et al*., 1985).

3. METHODOLOGY

The main objective of this project is to produce a statistically parsimonious and robust model to forecast the occurrence of soliton events in Sulu Sea. As far as statistical forecast is concerned, vast amount of time series data is often involved. The data obtained for this project is presented as current speed and direction with respect to time. The speeds and directions of currents at 3 different water depths, 26 meters, 100 meters and 200 meters below surface, are recorded at a 10-minute interval. These data specifications are based on the full-scale measured data collected from the site.

Several important methodologies are encompassed throughout this project. The details of these methodologies will be further discussed as follows.

3.1. Interpolation

Due to technical errors or limitations, approximately 7.24% of the data is not available. Suitable interpolation has to be made to acquire a complete time series data for forecasting purpose. SPSS (Statistical Package for the Social Sciences), a computer software used for statistical analysis developed by IBM, is used for interpolating the missing data. The types of interpolation functions SPSS is capable of generating include:

a) Series mean
b) Mean of nearby points
c) Median of nearby points
d) Linear interpolation
e) Linear trend at point

One interpolation method may not be applicable for every missing data in this series. For instance, a linear interpolation for one or two missing data points could be suitable. However, the time series data may be rendered inaccurate and unreliable if such method is applied on a series of continuous missing data points. Hence, choosing a suitable interpolation approach is essential as it would ensure the interpolated values fit in well with the existing data.

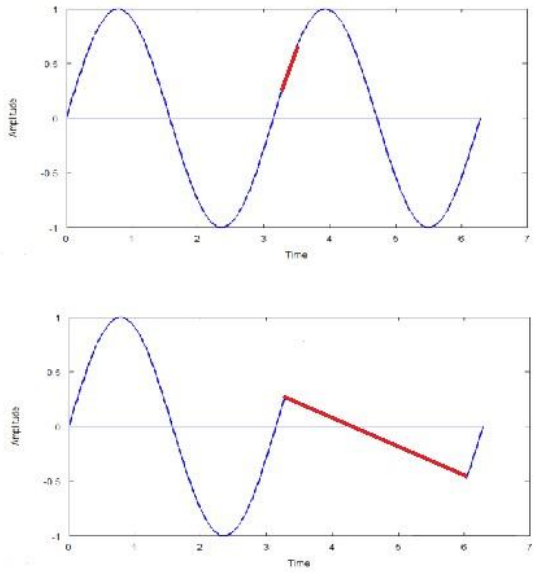


Figure 3: The same interpolation method may not work well for all missing data. The top figure shows how a linear interpolation fits well in a small gap missing data points. The bottom figure illustrates how the linear interpolation is not accurate for large gap missing data points.

3.2. Directional Rose Diagram

An alternative method to determine the dominating direction of the wind, wave or current is by using the directional rose diagram. A directional rose diagram shows the frequency of wind, wave or currents coming from a particular direction.

In this project, the directional rose diagram is divided into thirty six equal quadrants, each representing the incoming load from each respective direction.

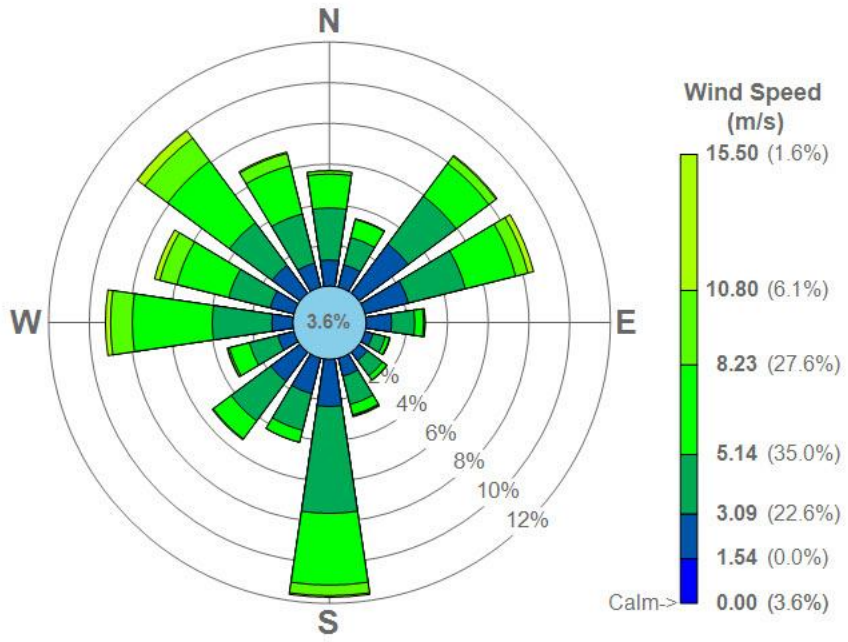


Figure 4: A typical 16-quadrant directional rose diagram.

3.3. Autocorrelation

Autocorrelation is a common method used to determine stochasticity of a time series. Such randomness can be shown by computing autocorrelations for data values at varying time lags. The autocorrelations for all time lag will be closed to zero if the data set is random.

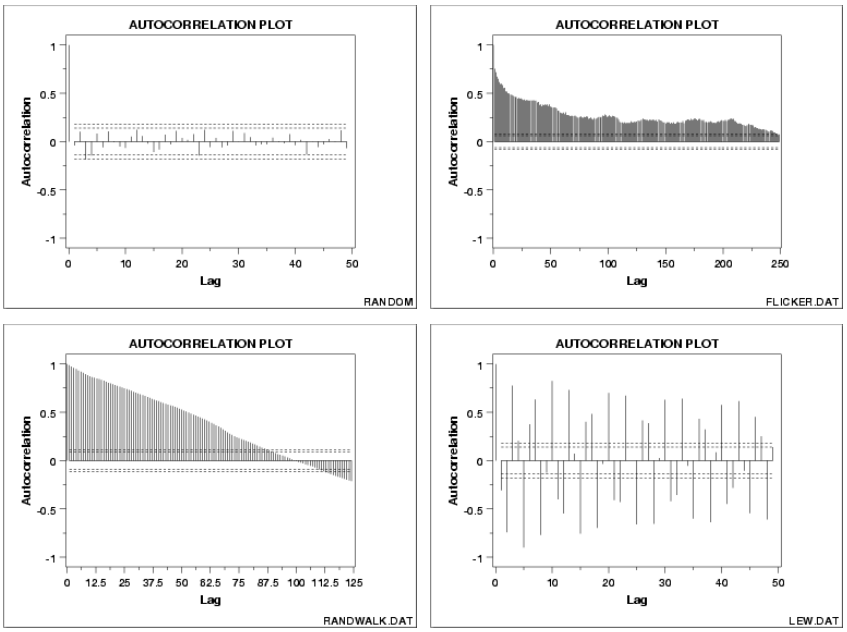


Figure 5: Types of autocorrelation plots. Clockwise from top left: random data, moderate autocorrelation, strong autocorrelation and autoregressive model and sinusoidal model. Reprinted from *Engineering Statistics Handbook: 1.3.3.1. Autocorrelation Plot*, n.d., Retrieved August 11, 2014, from www.itl.nist.gov/div898/handbook/eda/section3/autocopl.htm.

3.4. Spectrum Analysis

Researches show that the soliton event in the Sulu Sea is closely related to tidal current. Tide is highly cyclical, as it obeys several seasonal cycles. The seasonalities can include:

a) Daily diurnal or semidiurnal tide
b) Spring and neap tide
c) Annual monsoon seasons

It is difficult to determine all the seasonalities by studying solely on the speeds and directions of the tidal currents. Hence, a spectral analysis is adopted. Spectrum analysis is a plot of spectral density against frequency. It is suitable to determine the periodicities of a nondeterministic time series data.

3.5. Exponential Smoothing

Exponential smoothing is a method used in time series data. It is usually applied to make forecast. Normally, this smoothing method is adopted in economic market or financial data, it is however applicable for any discrete set of repetitive measurements. Raw data sequence and the output of the exponential smoothing

algorithm are represented by $\{x_t\}$ and $\{s_t\}$ respectively. The $\{s_t\}$ is taken as the best estimated x value. At $t = 0$, the series of measurements starts, the exponential smoothing of the simplest form is defined as:

$$s_0 = x_0$$

$$s_0 = s_t = \alpha x_{t-1} + (1-\alpha)s_{t-1}, t > 0$$

where α = smoothing factor, $0 < \alpha < 1$.

3.5.1. Simple Seasonal

The simple seasonal exponential smoothing formulae are made up of two parameters: level and seasonality.

$$Level, L_t = \alpha\,(Y_t - S_{t-1}) + (1-\alpha)L_{t-1}$$

$$Season, S_t = \delta(Y_t - L_t) + (1-\delta)S_{t-s}$$

$$Forecast, F_{t+k} = L_t + S_{t+k-s}$$

where s is the length of the seasonal cycle, for $0 \leqslant \alpha \leqslant 1$ and $0 \leqslant \delta \leqslant 1$.

3.5.2. Holt-Winters' Additive

For an exponential smoothing for data in which level, trend and seasonality parameters are taken into consideration, Holt-Winters' technique is now used. The trend equation is added to the Simple Seasonal method:

$$Level, L_t = \alpha(Y_t - S_{t-s}) + (1-\alpha)(L_{t-1} + T_{t-1})$$

$$Trend, T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1}$$

$$Season, S_t = \delta(Y_t - L_t) + (1-\delta)S_{t-s}$$

$$Forecast, F_{t+k} = L_t + kT_t + S_{t+k-s}$$

for $0 \leqslant \alpha \leqslant 1$, $0 \leqslant \gamma \leqslant 1$ and $0 \leqslant \delta \leqslant 1$.

3.5.3. Holt-Winters' Multiplicative

This is an alternative forecasting model generated by Holt-Winter, where the seasonal parameter is multiplied:

$$Level, L_t = \alpha \frac{Y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + T_{t-1})$$

$$Trend, T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}$$

$$Season, S_t = \delta \frac{Y_t}{L_t} + (1-\delta)S_{t-s}$$

$$Forecast, F_{t+k} = (L_t + kT_t)\, S_{t+k-s}$$

for $0 \leqslant \alpha \leqslant 1$, $0 \leqslant \gamma \leqslant 1$ and $0 \leqslant \delta \leqslant 1$.

Analogous to a linear Cartesian graph with a general equation y = mx + c, the level parameter would be the y-intercept, c. The trend on the other hand is represented by the gradient of the graph, m. Seasonality is defined as how often the output repeats itself.

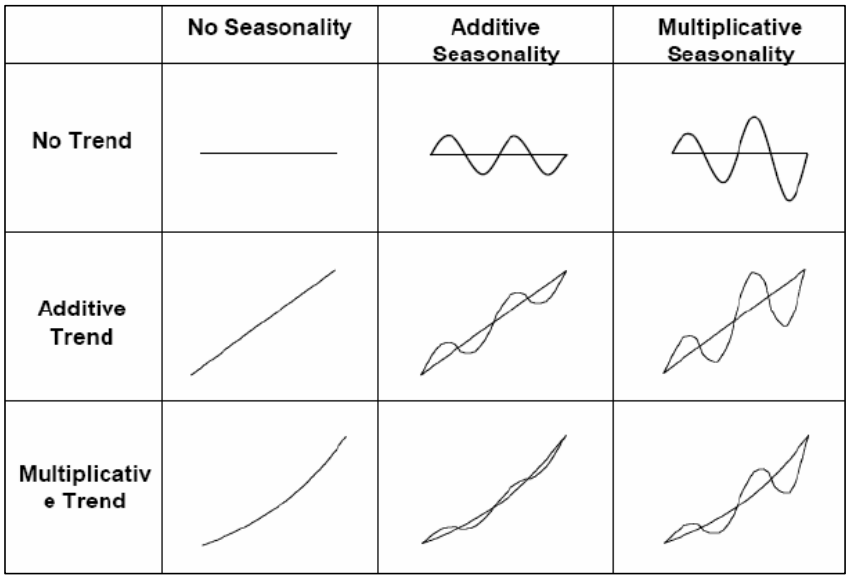


Figure 6: Relationship between trend and seasonality. Reprinted from *Holt-Winters' Exponential Smoothing with Seasonality*, n.d., Retrieved April 3, 2014, from http://s3.amazonaws.com/zanran_storage/www.cec.uchile.cl/ContentPages/107548415.pdf.

3.6. Initialisation

In order to achieve an accurate forecast model, a proper initialisation method is essential. In this forecast model, 2s observations are used for the initialisation, where s is the length of the seasonal cycle.

The level component L_s is taken as the average observation from t = 1 to t = s.

The trend component T_s is initialised by using the equation:

$$T_s = \frac{\sum_{i=s+1}^{2s} Y_i - \sum_{j=1}^{s} Y_j}{s^2}$$

The initialisation for seasonality components differ between additive and multiplicative. For additive, the seasonality is initialised by $S_i = Y_i - L_s$, where for multiplicative it is initialised by $S_i = Y_i / L_s$.

3.7. Pearson Product Moment Correlation

Correlation between two or more data sets describes how well they are related to one another. Pearson Product Moment Correlation (PPMC), sometimes known as Pearson Correlation, is one of the most widely used measure of correlation in statistics. The PPMC shows the linear relationship between these data sets.

In this studies, the Pearson correlation for sample, r is used, and it is defined as:

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{\left[n\sum x^2 - (\sum x)^2\right][n\sum y^2 - (\sum y)^2]}}$$

Result yield from the PPMC will range from -1.0 to 1.0. The r value approaches 0 when the data points around the line of best fit are of greater variations. The table below summarises the correlations between the data sets based on the r value.

Table 1: Relationship between r values and correlation between data sets.

r Value	Correlation
1.0 to 0.5 or -0.5 to -1.0	High
0.5 to 0.3 or -0.3 to -0.5	Medium
0.3 to 0.1 or -0.1 to 0.3	Low

3.8. Coefficient of Determination

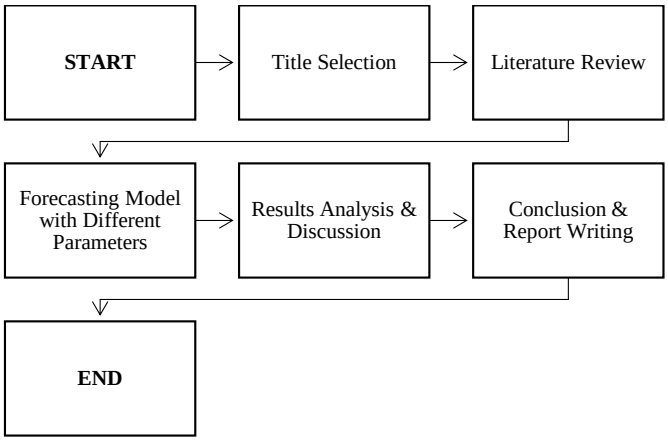
Without a proper testing approach, it is difficult to tell if a forecasting model is accurate and reliable. The coefficient of determination of a statistical model describes how closely forecasting model fits the observation. The result of this test summarises the discrepancy between the existing values and the forecasted values. In this project, the R^2 test will be used to test its goodness of fit.

3.9. Sensitivity Analysis

The robustness of this statistical forecasting model can be determined by conducting a sensitivity analysis. A sensitivity analysis serves to identify how the uncertainty in the input data impacts its result. Taking the simple seasonal exponential smoothing method as an example, a sensitivity analysis can encompass predicted seasonality (an independent variable) and a corresponding forecasted value (the dependent variable) based on different values for each independent variables. This can further be done for the level and trend variables for the Holt-Winters' approach.

3.10. Project Activities

The chart below illustrates the activities slated throughout the project:


START
Title Selection
Literature Review
Forecasting Model with Different Parameters
Results Analysis & Discussion
Conclusion & Report Writing
END

4. RESULTS & DISCUSSION

4.1. Interpolation

In this project, linear interpolation method by using SPSS (Statistical Package for the Social Sciences) software has been opted. A quick calculation on Microsoft Excel spreadsheet has shown that the difference between the mean and the standard deviation of the interpolated data and the raw uncorrected data are 0.176% and 1.079%. These values are within the allowable 5% confidence limit, ~~H~~hence~~,~~ the linear interpolation method is acceptable.

4.2. Current Rose Diagram

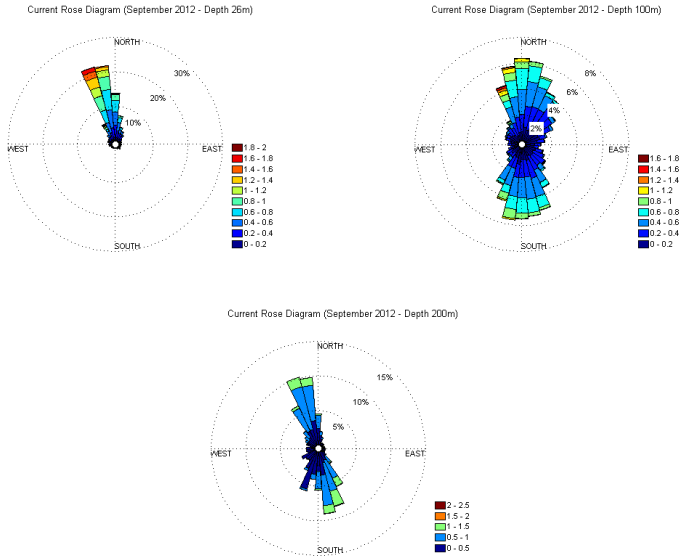


Figure 7: Current rose diagrams at Karupang in September 2012 at 26m, 100m and 200m depths

The figure above compares the directionalities, intensities as well as frequencies of the currents in Karupang in September 2012 at three different depths. The diagrams illustrate the likeliness of high speed current events (> 1.5m/s) in all depths. The current rose diagram suggests the occurrence of strong ~~ebbing~~ southward ~~(southwards)~~and weak ~~flooding (~~northward~~s)~~ flow at 26m water depth. As previously shown in Figure 1, southward flow represents the ebbing tide while the northward flow symbolizes the flooding tide.

In contrary to the ~~former diagram~~current rose diagram for 26m water depth, the ~~latter~~ ones for 100m and 200m water depth shows the strong northwards currents

occur as likely as the southwards ones, suggesting the possibilities of baroclinic flow condition.

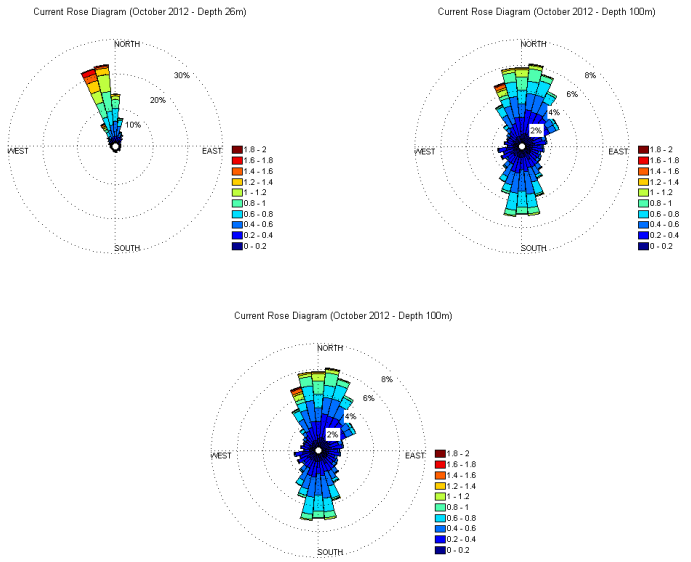


Figure 8: Current rose diagrams at Karupang in October 2012 at 26m and 100m depths

~~The figure above~~Figure 8 on the other hand compares the currents between 26m and 100m depths at Karupang in October 2012. It is shown that the trends of the rose diagrams are comparable to the ones in September. Also, it signifies the likelihood of baroclinic conditions between 26m and 100m water depths.

4.3. Autocorrelation Function

An autocorrelation function (ACF) determines the number of lags between two correlated observations. In this project, the autocorrelation graphs for each month at 26m, 100m and 200m water depths are plotted and compared.

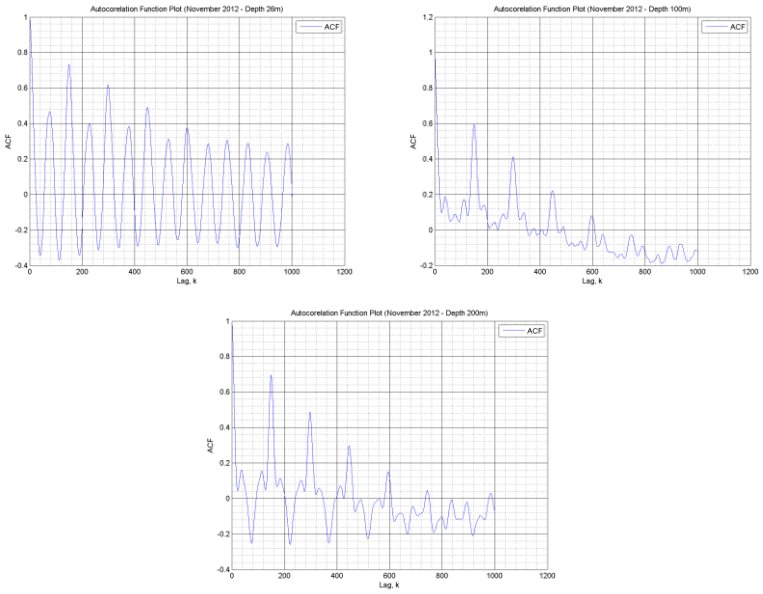


Figure 9: Autocorrelation function (ACF) plots for the current speed at Karupang in November 2012. Clockwise from top left: Depth 26m, Depth 100m and Depth 200m

From the plots above, the ACF plots of current speeds in November 2012 against the number of lags at 3 different depths: 26m, 100m and 200m can be observed. Each lag is equivalent to the amount of time (in seconds) between two consecutive observations. In this project, the time interval between each observation is 10 minutes. Hence, 1 lag is equivalent to 600 seconds.

From the ACF plots, it can be observed that the number of lags between two correlated data at each depth is approximately 150. The cyclical period is therefore 150 × 600 seconds = 90000 seconds or 25 hours. This phenomenon indicates the strong underlying process in a form of sinusoidal movement which repeats at a 24 hour period, which is believed to be strongly correlated to the daily tidal event. Other peaks represent the less dominant periodicities of the time series.

The plot for Depth 26m has shown a sinusoidal trend, whilst the ones for depth 100m and 200m can be considered as strong autocorrelation and autoregressive model.

Other dominating cyclical components in this time series can be obtained by carrying out spectrum analysis.

4.4. Spectrum Analysis

As previously discussed, the current trends that can lead to the occurrence of soliton are seasonal, and the seasonality components can be determined by carrying out a spectrum analysis, where dominant frequencies (or periods) can be observed as peaks in the spectral density diagram.

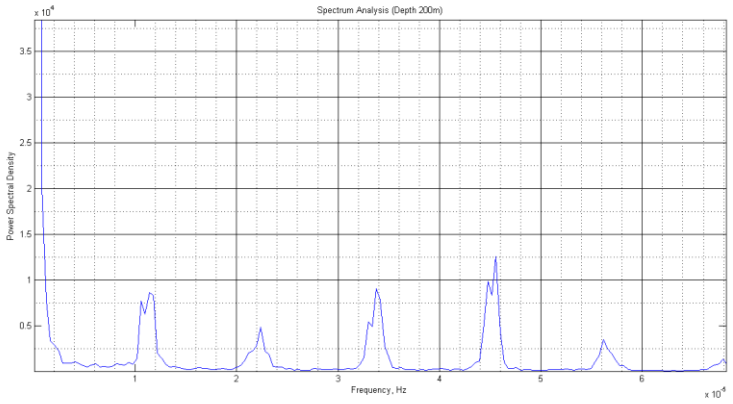


Figure 10: Power spectral density plot for water depth 200m.

The first power spectral density (PSD) plot is plotted based on the current speed data at 200m water depth. It is shown that the PSD plot captures more than one dominant frequencies. The results interpreted from this PSD plot are summarized in the table below:

Table 2: Dominant frequencies and periodicities from the PSD plots.

Water Depth (m)	Frequency ($x10^{-5}$ Hz)	Periodicity (hour)
26	1.15	24.15
	2.25	12.35
100	1.05	26.46
	2.25	12.35
200	3.40	8.17
	4.75	5.85

The periodicities obtained from the PSD plot are close to the diurnal (24 hours 50 minutes) and semidiurnal (12 hours 25 minutes) tidal cycle periods, suggesting the close relationship between the soliton events and the daily tidal cycle. ~~The table below summarises the frequencies and the periods of the first four peaks for both the plots:~~

Comment [S8]: Please relate this to tides! Shows that highspeed soliton events have a close relationship with tidal events..

~~4.5. Table 1: Frequencies and periods of the first four peaks in the spectrum analysis plots for October 2012 and the second new-moon-to-new-moon cycle~~

~~1.1.1.1.~~

~~4.6.~~4.5. Reported Events

Results from an oceanographic analysis report published by the Global Ocean Associates have been taken as reference to identify the relationship between the soliton event and oceanographic criteria.

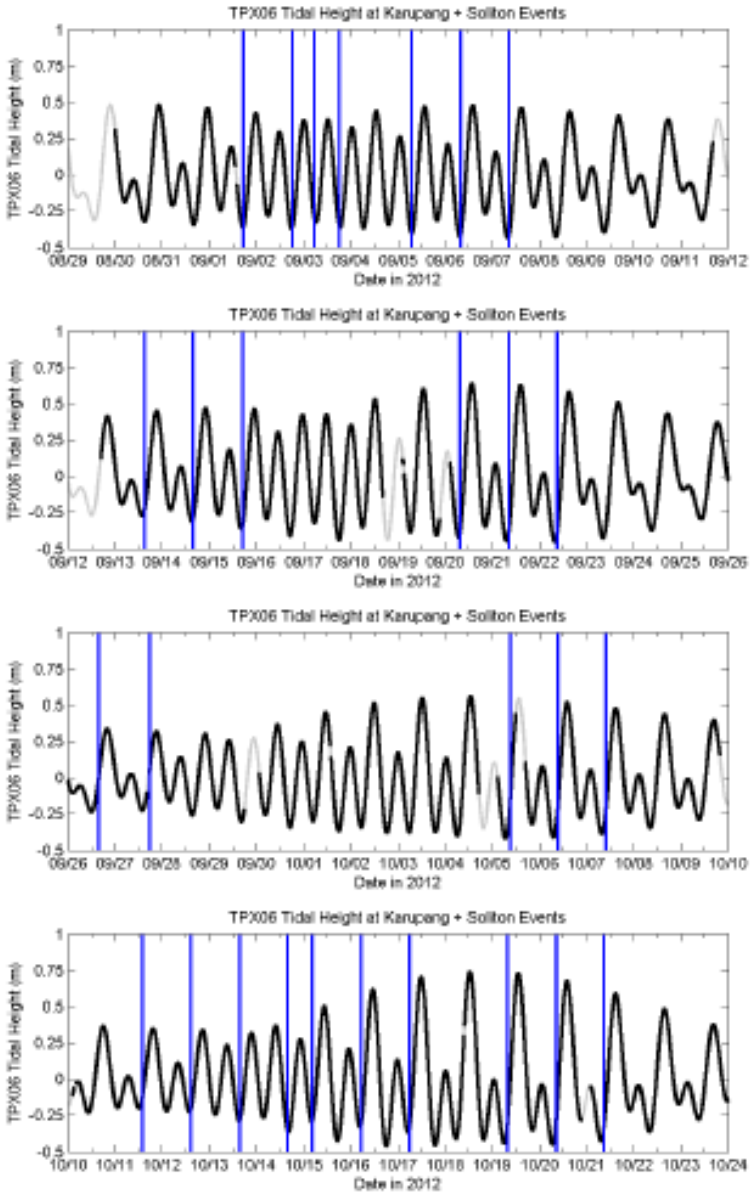


Figure 11: Soliton events with respect to the barotropic tidal height.

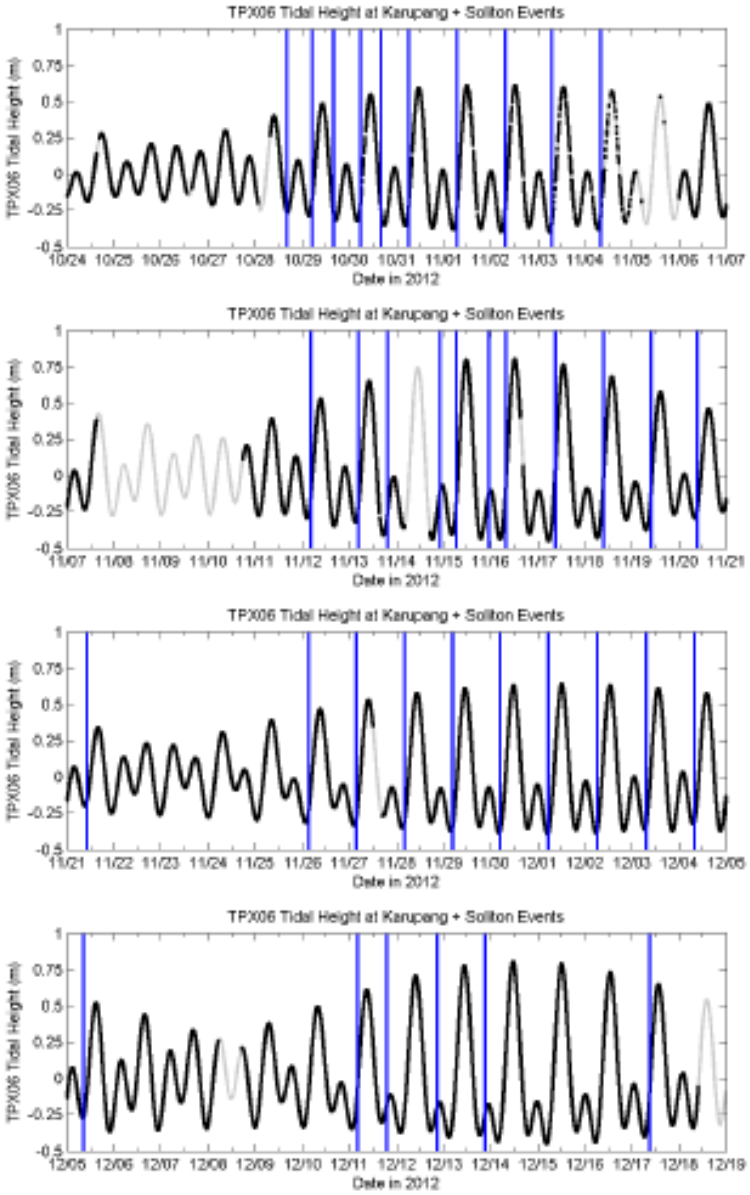


Figure 12: Soliton events with respect to the barotropic tidal height.

Figure 11 and 12, extracted from the oceanographic analysis report, illustrate the relationship between the tidal height and the occurrence of soliton. It can be concluded that solitons occur during minimum tidal height. However, soliton

events are not reported at every tidal cycle. There is also no clear indication that the soliton events are dependent on the magnitudes of the tidal heights.

It is believed that soliton events are closely related to high speed current events. In accordance with the lee wave generation method, soliton occurs during baroclinic condition, where resonated stationary internal wave slackens and becomes subcritical, during which soliton may occur. The current directionality time series is compared with the reported soliton events in Figure 11 and 12 to determine the number of events that coincide with baroclinic condition. Table 2 shows the number of events when the reported soliton coincides with baroclinic condition.

Comment [S9]: Why only baroclinic? Explain. Relate back to literature

Table 3: Number of reported soliton events in baroclinic and barotropic condition.

Reported Baroclinic Condition	Reported Soliton Events at Baroclinic Condition	Reported Soliton Events at Barotropic Condition
370	43	22

As opposed to the studies done on the lee wave generation methods, it is shown that only 66.2% of the soliton events coincide with baroclinic condition. This trend suggests that soliton events may take place independently of the baroclinic/barotropic condition.

Comment [S10]: As opposed to literature which indicates that baroclinic conditions should appear. So what does this mean???????

~~4.7.~~4.6. Statistical Forecast

In order to maintain the parsimony of the statistical forecast model, the simple seasonal exponential smoothing method has been adopted. The seasonal cycle length used in this model is based on the result from the power spectrum density plot discussed above, which is 24 hours and 15 minutes. Due to the 10-minute interval of the data, the seasonal cycle length has been rounded off to 24 hours and 20 minutes. The sensitivity analysis shows that using 0.50 and 0.49 as α and δ will yield the most optimal result.

Comment [S11]: After sensitivity analysis?

In the forecast model, the Level component is computed based on the observation in the previous seasonal cycle of the same phase, where:

$$L_t = \alpha(Y_{t-s} - S_{t-s}) + (1-\alpha)L_{t-1}$$

The level component is modified in such a way in order to provide a forecast for the current speed for at least one seasonal cycle ahead.

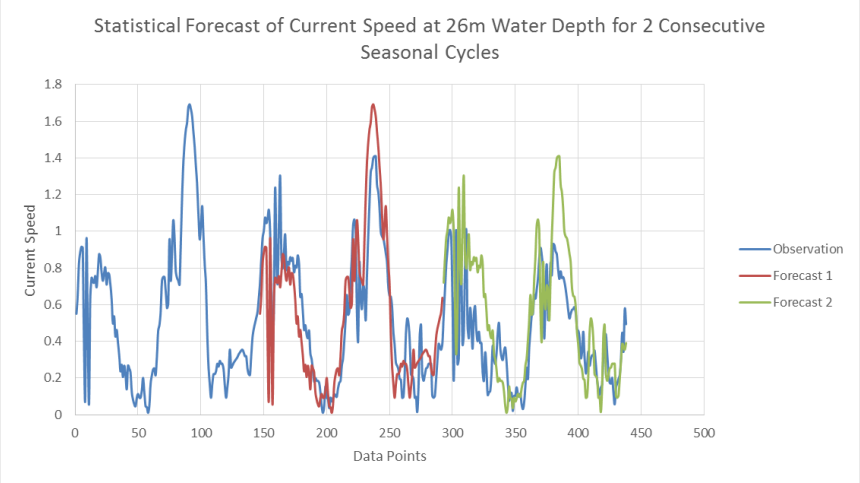


Figure 13: Statistical forecast of current speed at 26m water depth for 2 consecutive seasonal cycles.

The plot above illustrate the statistical forecast model using the simple seasonal method based on the current speed data at 26m water depth. The blue line represents the actual observed current speed, the red and green lines represent the forecasts based on the observation during the first seasonal cycle.

The forecasted values are compared to the values of the observation at respective time, and it is shown that 71.2% of the forecasted values of the first cycle are within a difference of 0.2 m/s, whereas the percentage drops to 61.6% for the second cycle. The R^2 coefficient of determination test has also been carried out for this model, and it is shown that the first forecast displays a R^2 of 69.6% whilst the second forecast shows 65.1%. The Pearson correlation, r, also decreases from 0.835 to 0.807 when comparing the values for the first and second cycle. The table below summarises the goodness of fit, Pearson correlation as well as percentage of accuracy for three different water depths.

Comment [S12]: U sure about this comment???????????

Table 4: Accuracy percentage and goodness of fit for different water depths.

Water Depth (m)	First Cycle Forecast			Second Cycle Forecast		
	R^2 (%)	Pearson Correlation, r	Percentage of Accuracy (%)	R^2 (%)	Pearson Correlation, r	Percentage of Accuracy (%)
26	69.631	0.835	70.548	54.565	0.739	61.644
100	45.881	0.677	71.233	28.712	0.536	68.493
200	67.349	0.821	66.438	65.112	0.807	71.233

~~4.8.~~4.7. Sensitivity Testing

In order to test the sensitivity of seasonal smoothing constant δ, a sensitivity test has been carried out and its impact onto the R^2, Pearson correlation, r and the percentage of accuracy has been observed. δ from 0 to 1 has been tested at a 0.1 interval.

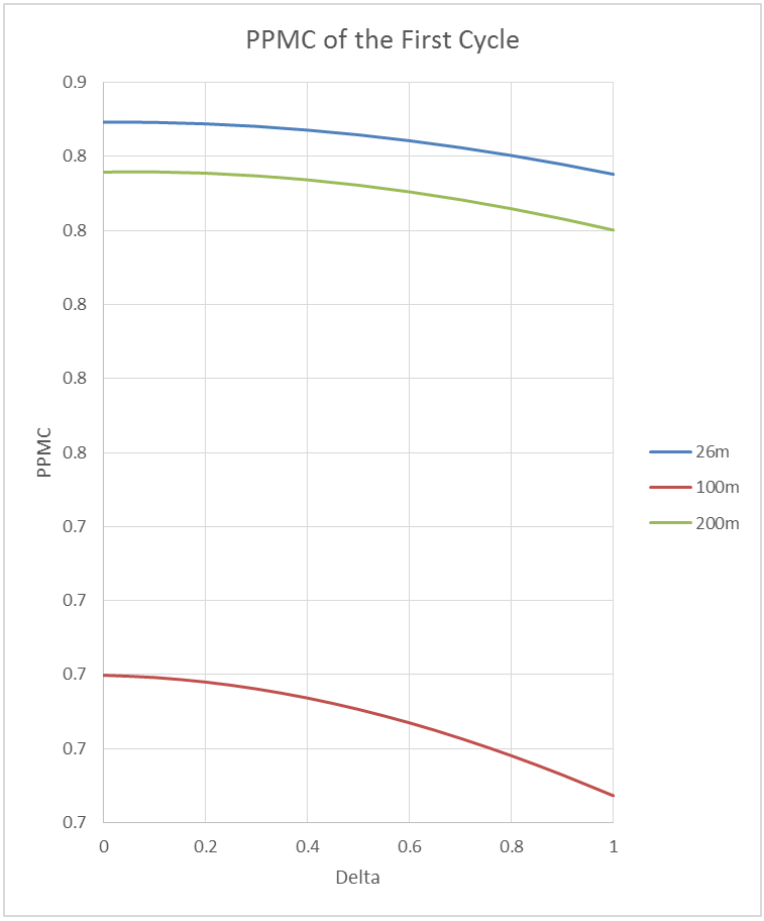


Figure 14: The Pearson product moment correlation values for the first forecast cycle for $\delta = 0$ to $\delta = 1$ at 0.1 interval.

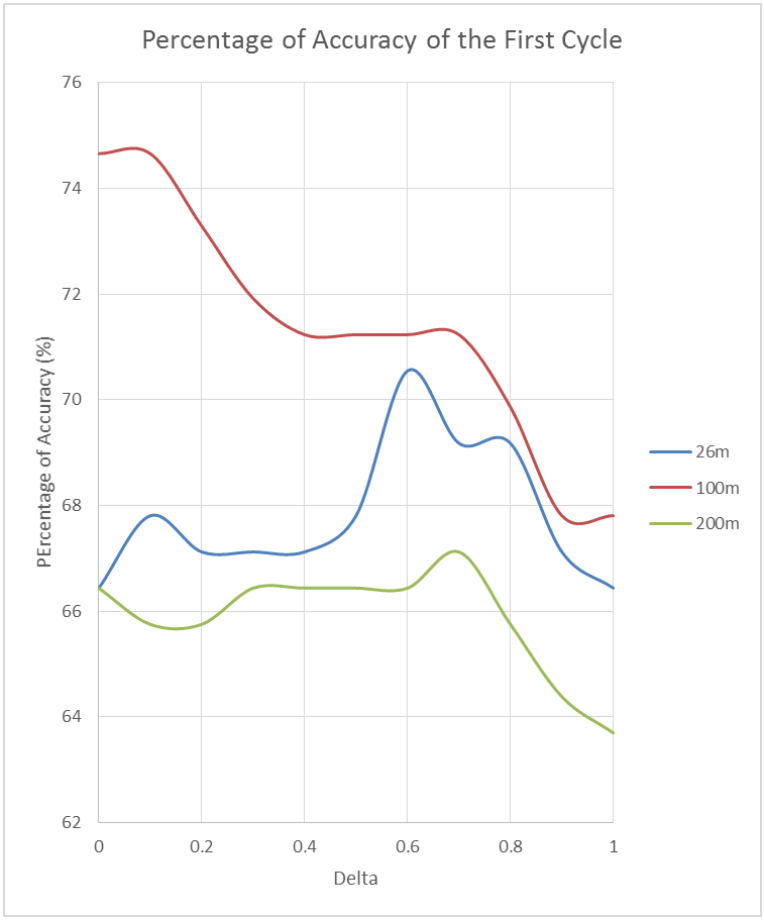


Figure 15: The percentage of accuracy for the first forecast cycle for $\delta = 0$ to $\delta = 1$ at 0.1 interval.

From the Pearson product moment correlation plot, it is observed that the correlation, r decreases as seasonality smoothing constant δ increases. From the percentage of accuracy plot however, optimal percentage of accuracy can be observed between $\delta = 0.4$ to $\delta = 0.6$. Hence, a more detailed sensitivity analysis is carried out at 0.01 interval from $\delta = 0.40$ to $\delta = 0.60$.

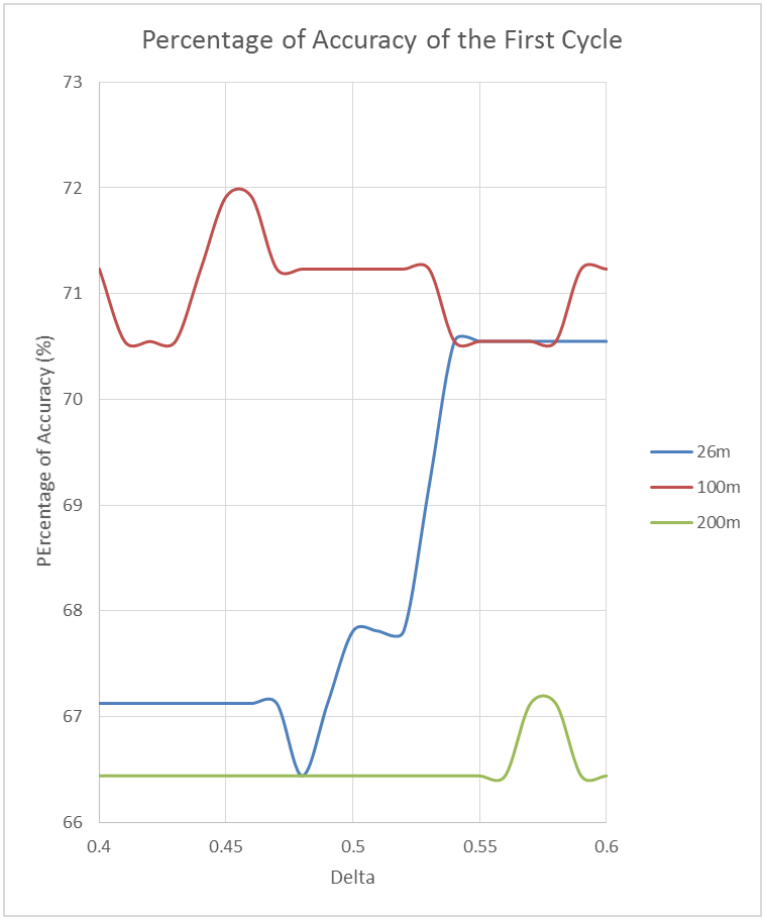


Figure 16: The percentage of accuracy for the first forecast cycle for $\delta = 0.40$ to $\delta = 0.60$ at 0.01 interval.

In the graph above, it can be observed that the percentage of accuracy shows optimal values seasonality smoothing constant $\delta = 0.59$.

5. CONCLUSION & RECOMMENDATION

The simple seasonal exponential smoothing method can produce a forecasting model of up to 1 seasonal cycle (24 hours and 20 minutes) ahead for the current speeds at Karupang site at different water depth with smoothing constants α and δ be 0.50 and 0.59 respectively. Results have shown that the percentage of accuracy and the goodness of fit are up to 70%.

Numerous literature has suggested that soliton events are dependent on high speed current event. In this study however, no clear correlation can be observed between the high speed current events and the occurrence of solitons. Nonetheless, the accuracy and robustness of this model can potentially contribute to the oil and gas industry. Downtime loses and potential casualties due to high speed current events can be reduced.

Comment [S13]: revise

Based on the findings in this studies, it is believed that the forecast model does not capture seasonalities other than the daily tidal cycle. It is recommended that a more comprehensive forecast should be constructed while considering seasonal components such as spring/neap tidal cycle or even monsoon cycle.

~~The exponential smoothing forecast can now be initiated.~~

Comment [S14]: WHERE IS YOUR RECOMMENDATIONS?

6. REFERENCES

Apel, J. R., Holbrook, J. R., Liu, A. K., & Tsai, J. J. (1985). *The Sulu Sea Internal Soliton Experiment*.

Apel, J. R., Ostrovsky, L. A., Stepanyants, Y. A., & Lynch, J. F. (2006). *Internal Solitons in the Ocean*. Woods Hole Oceanographic Institution.

Gerkema, T., & Zimmerman, J. T. F. (2008). *An Introduction to Internal Waves*. Retrieved February 16, 2014, from http://www.nioz.nl/public/fys/staff/leo_maas/course/book.pdf

Global Ocean Associates LLC., & Oceanalysis Ltd. (2013). *Oceanographic Analysis Report for the Sibutu Passage (2012).*

Holt-Winters' Exponential Smoothing with Seasonality. (n.d.). Retrieved April 3, 2014, from http://s3.amazonaws.com/zanran_storage/www.cec.uchile.cl/ContentPages/107548415.pdf

Hyndman, R. J. (2010, November 30). *Initializing the Holt-Winters method*. Retrieved April 11, 2014, from http://robjhyndman.com/hyndsight/hw-initialization/

Jackson, C. R., Da Silva, J. C., & Jeans, G. (2012). The Generation of Nonlinear Internal Waves. *Oceanography*, *25*(2), 108-123. Retrieved from http://dx.doi.org/10.5670/oceanog.2012.46

Jeans, G. (2013). *The Application of Internal Wave Physics to Offshore Engineering*.

Wallace, A. R. (1869). *The Malay Archipelago*. Dover Publications.

7. APPENDICES

7.1. Current Rose Diagrams

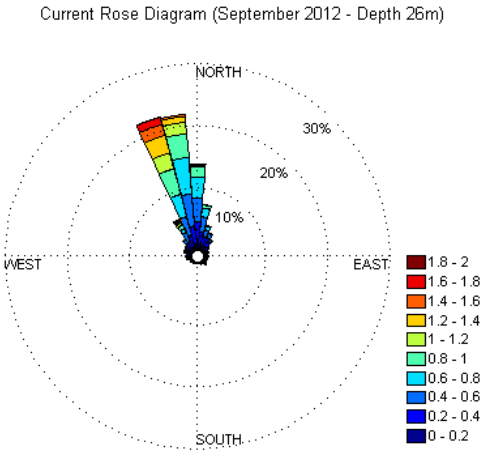


Figure 17: Current rose diagram for Sept 2012 at depth 26m.

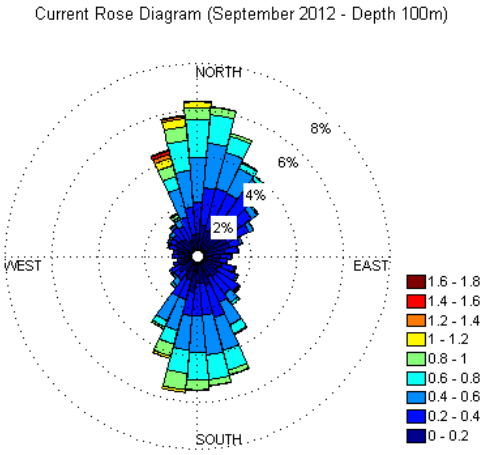


Figure 18: Current rose diagram for Sept 2012 at depth 100m.

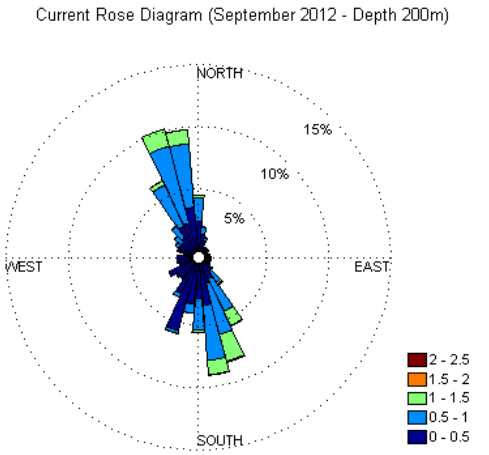


Figure 19: Current rose diagram for Sept 2012 at depth 200m.

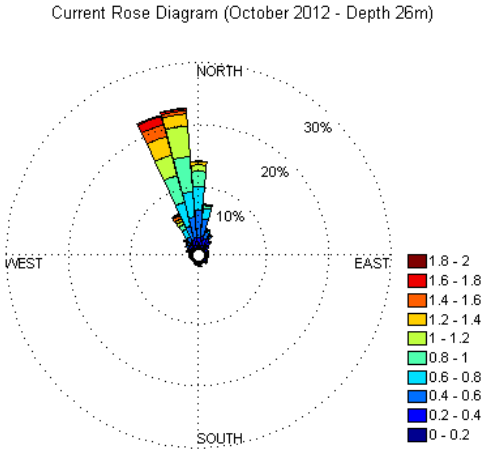


Figure 20: Current rose diagram for Oct 2012 at depth 26m.

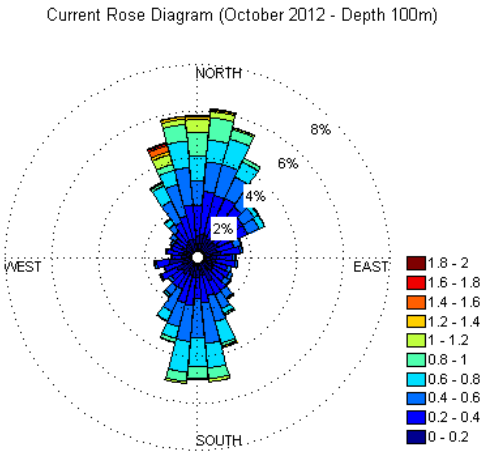


Figure 21: Current rose diagram for Oct 2012 at depth 100m.

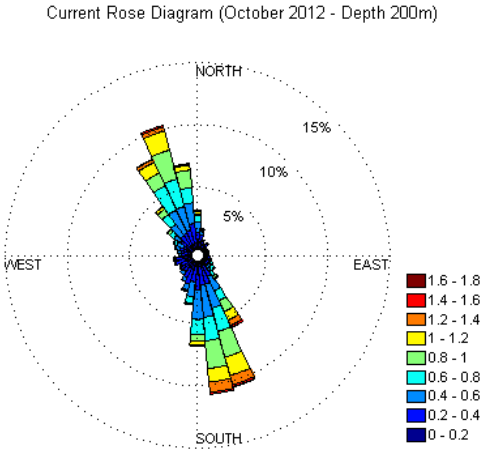


Figure 22: Current rose diagram for Oct 2012 at depth 200m.

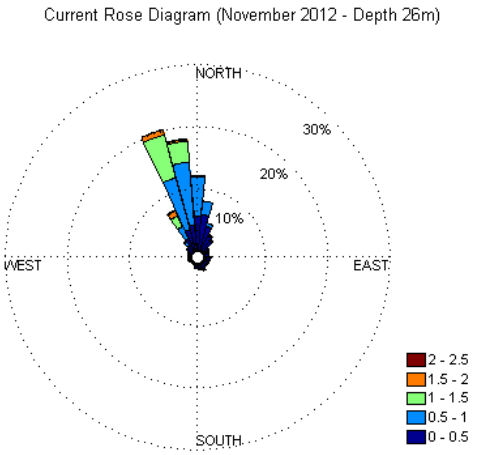


Figure 23: Current rose diagram for Nov 2012 at depth 26m.

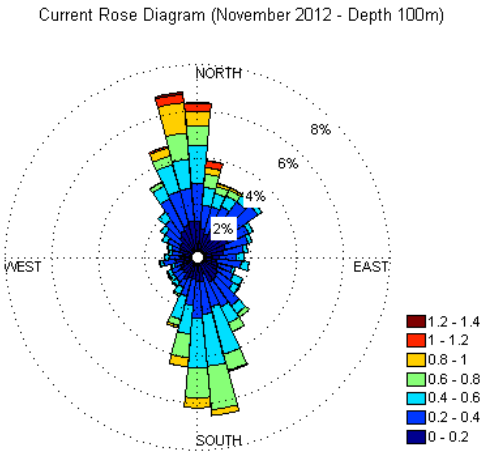


Figure 24: Current rose diagram for Nov 2012 at depth 26m.

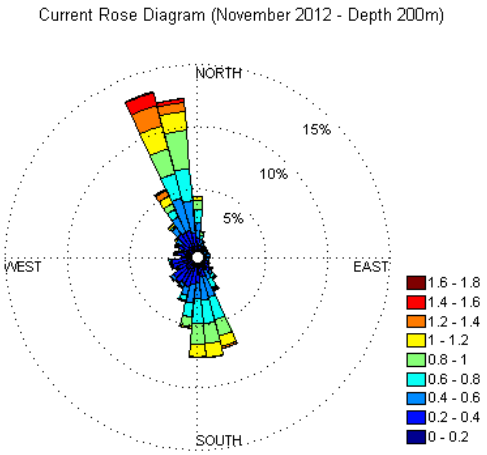


Figure 25: Current rose diagram for Nov 2012 at depth 200m.

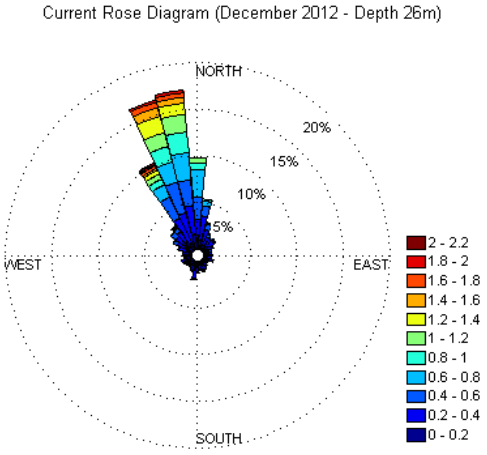


Figure 26: Current rose diagram for Dec 2012 at depth 26m.

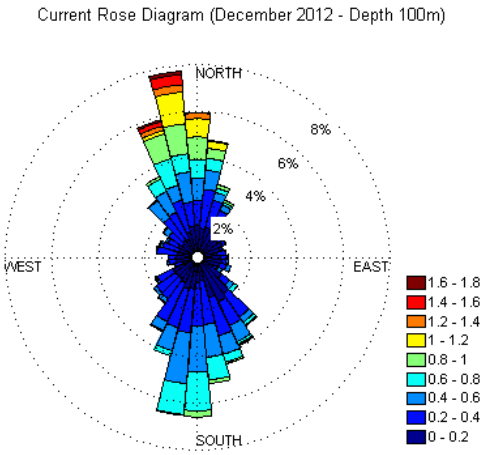


Figure 27: Current rose diagram for Dec 2012 at depth 100m.

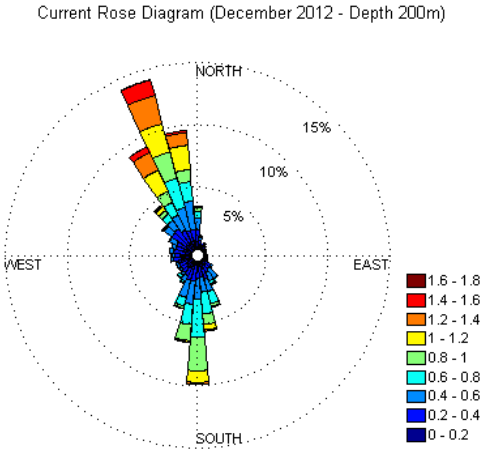


Figure 28: Current rose diagram for Dec 2012 at depth 200m.

7.2. Autocorrelation Plots

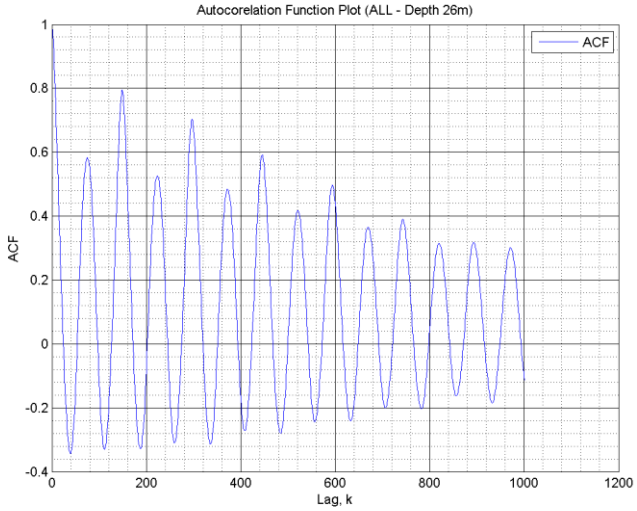


Figure 29: Autocorrelation plot for depth 26m.

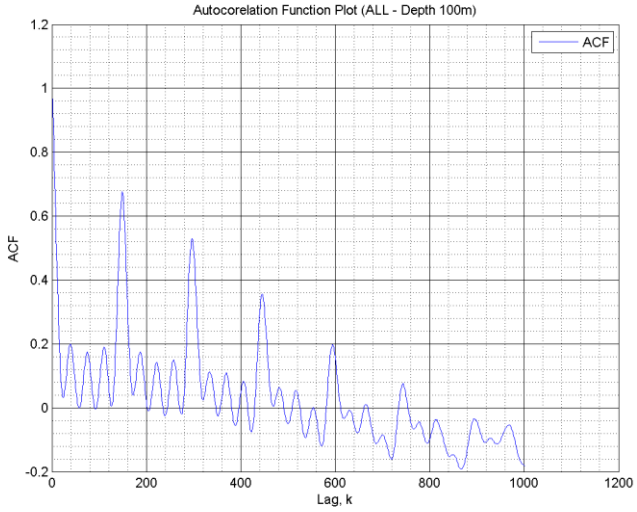


Figure 30: Autocorrelation plot for depth 100m.

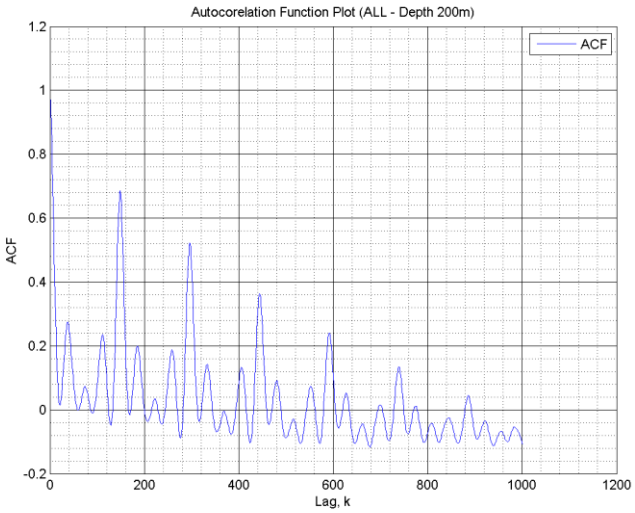


Figure 31: Autocorrelation plot for depth 200m.

7.3. Power Spectral Density (PSD) Plots

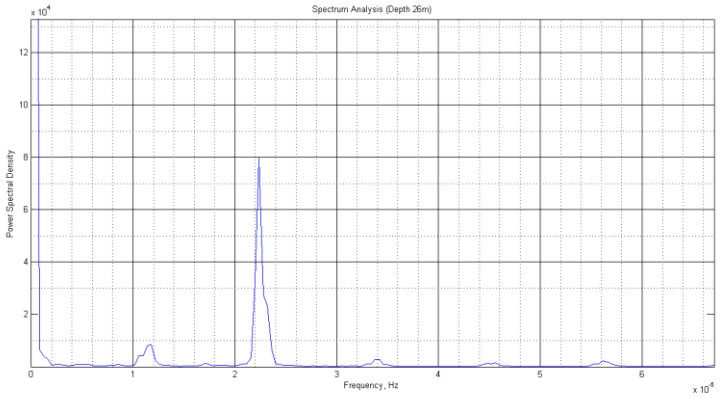


Figure 32: PSD plot of raw current speed data at 26m water depth.

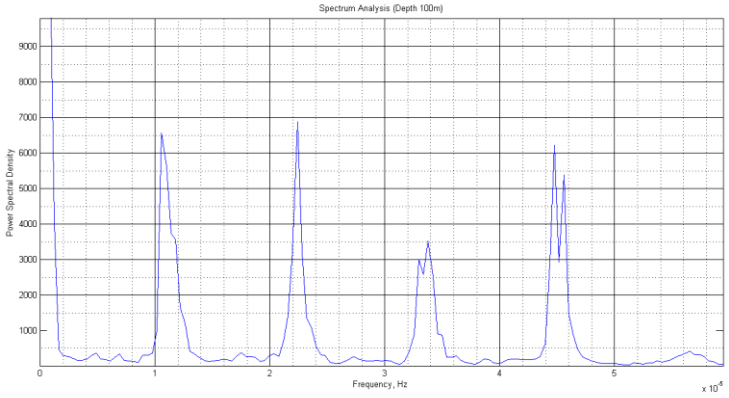


Figure 33: PSD plot of raw current speed data at 100m water depth.

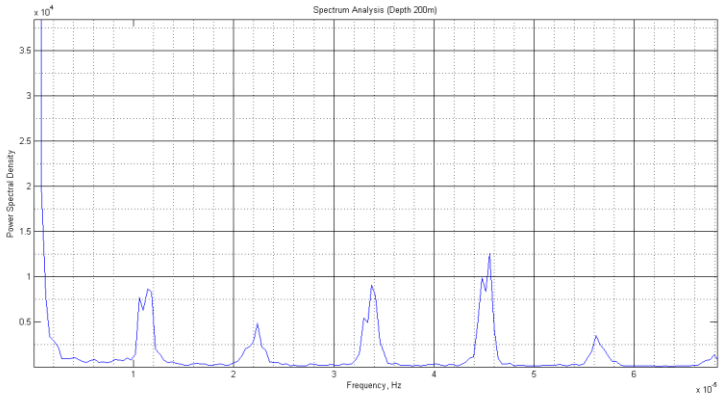


Figure 34: PSD plot of raw current speed data at 200m water depth.

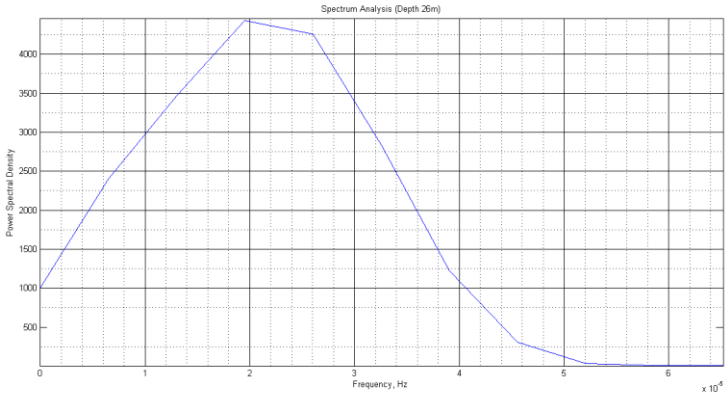


Figure 35: PSD plot of autocorrelated current speed data at 26m water depth.

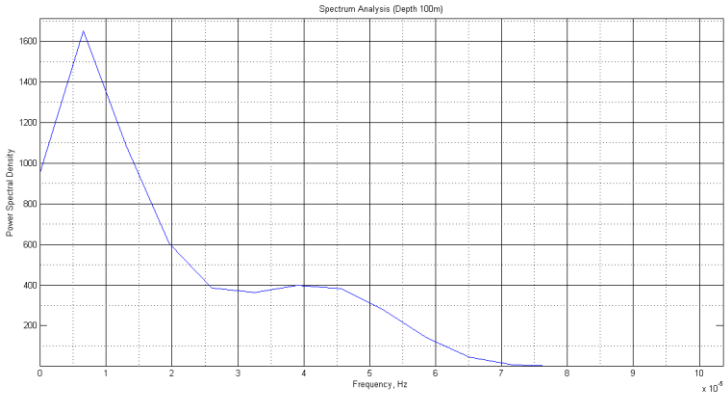


Figure 36: PSD plot of autocorrelated current speed data at 100m water depth.

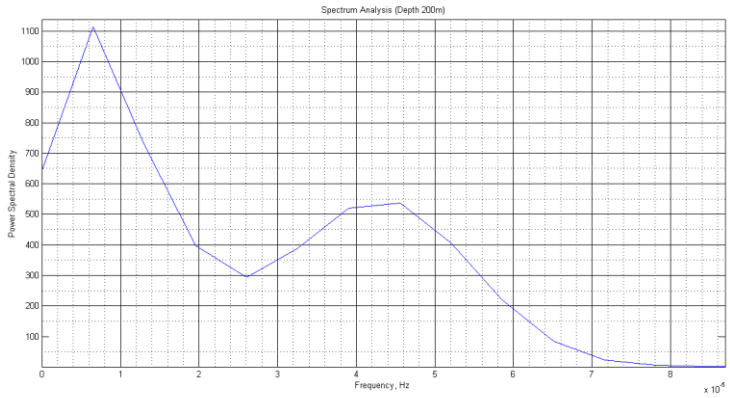


Figure 37: PSD plot of autocorrelated current speed data at 200m water depth.

7.4. Statistical Forecast

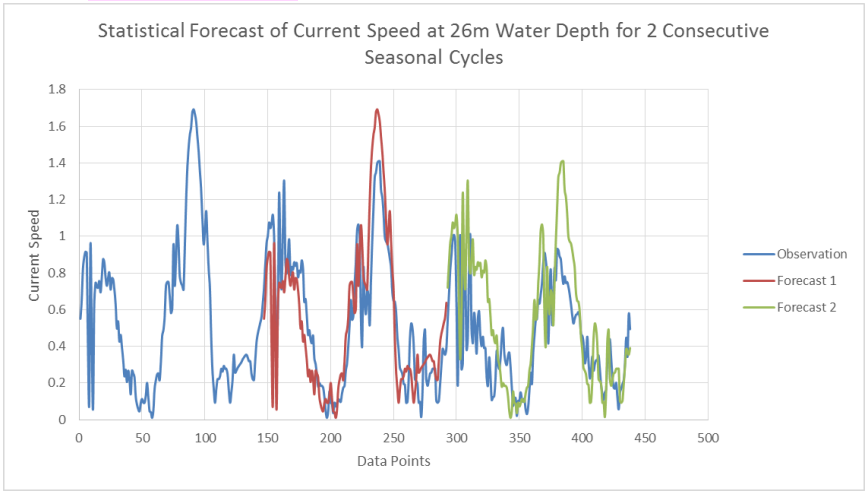


Figure 38: Statistical forecast of current speed at 26m water depth.

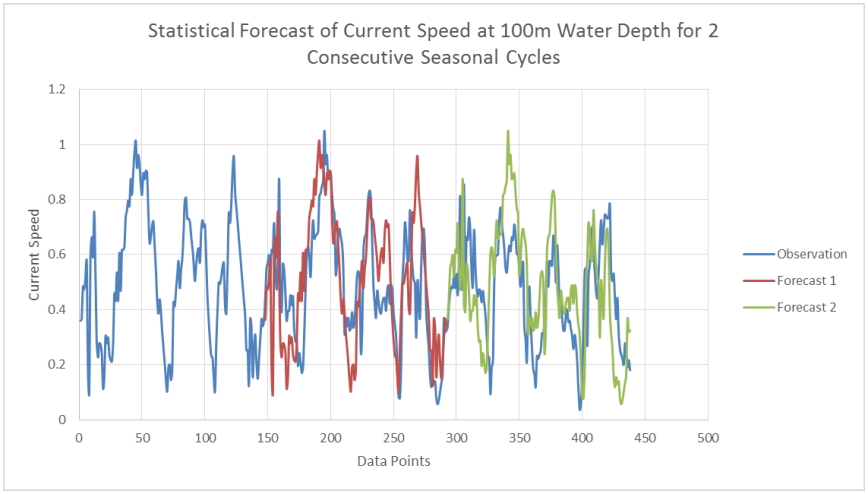


Figure 39: Statistical forecast of current speed at 100m water depth.

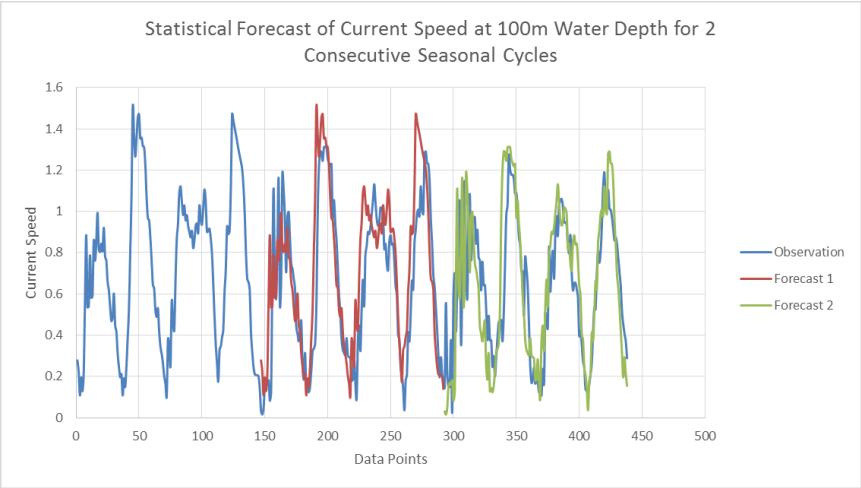


Figure 40: Statistical forecast of current speed at 200m water depth.

7.5. Sensitivity Analysis

Table 5: δ value from 0 to 1 at 0.1 interval for the first cycle forecast.

Delta	First Cycle									
	R^2			PPMC			Percentage of Accuracy			
	26	100	200	26	100	200	26	100	200	
0	70.444	47.584	68.197	0.839	0.690	0.826	66.438	74.658	66.438	
0.1	70.431	47.503	68.203	0.839	0.689	0.826	67.808	74.658	65.753	
0.2	70.366	47.334	68.146	0.839	0.688	0.826	67.123	73.288	65.753	
0.3	70.249	47.079	68.027	0.838	0.686	0.825	67.123	71.918	66.438	
0.4	70.083	46.740	67.848	0.837	0.684	0.824	67.123	71.233	66.438	
0.5	69.867	46.322	67.611	0.836	0.681	0.822	67.808	71.233	66.438	
0.6	69.603	45.828	67.317	0.834	0.677	0.820	70.548	71.233	66.438	
0.7	69.293	45.263	66.969	0.832	0.673	0.818	69.178	71.233	67.123	
0.8	68.937	44.633	66.570	0.830	0.668	0.816	69.178	69.863	65.753	
0.9	68.539	43.943	66.123	0.828	0.663	0.813	67.123	67.808	64.384	
1	68.099	43.199	65.630	0.825	0.657	0.810	66.438	67.808	63.699	

Table 6: δ value from 0 to 1 at 0.1 interval for the second cycle forecast.

Delta	Second Cycle									
	R^2			PPMC			Percentage of Accuracy			
	26	100	200	26	100	200	26	100	200	
0	57.153	29.504	68.968	0.756	0.543	0.830	62.329	69.863	72.603	
0.1	56.741	29.645	68.652	0.753	0.544	0.829	63.699	69.863	72.603	
0.2	56.302	29.658	68.157	0.750	0.545	0.826	63.699	69.178	73.973	
0.3	55.850	29.553	67.518	0.747	0.544	0.822	63.699	68.493	73.288	
0.4	55.397	29.345	66.765	0.744	0.542	0.817	62.329	68.493	71.918	
0.5	54.952	29.047	65.924	0.741	0.539	0.812	61.644	68.493	71.233	
0.6	54.523	28.672	65.018	0.738	0.535	0.806	61.644	68.493	71.233	
0.7	54.115	28.229	64.068	0.736	0.531	0.800	61.644	69.863	69.863	
0.8	53.731	27.731	63.089	0.733	0.527	0.794	60.959	68.493	67.808	
0.9	53.370	27.184	62.092	0.731	0.521	0.788	59.589	69.178	65.753	
1	53.034	26.594	61.087	0.728	0.516	0.782	58.904	67.123	65.068	

Table 7: δ value from 0.40 to 0.60 at 0.01 interval for the first cycle forecast.

Delta	First Cycle									
	R^2				PPMC				Percentage of Accuracy	
	26	100	200		26	100	200		26	100	200
0.40	70.083	46.740	67.848		0.837	0.684	0.824		67.123	71.233	66.438
0.41	70.063	46.702	67.827		0.837	0.683	0.824		67.123	70.548	66.438
0.42	70.043	46.663	67.805		0.837	0.683	0.823		67.123	70.548	66.438
0.43	70.023	46.623	67.783		0.837	0.683	0.823		67.123	70.548	66.438
0.44	70.002	46.582	67.760		0.837	0.683	0.823		67.123	71.233	66.438
0.45	69.981	46.541	67.737		0.837	0.682	0.823		67.123	71.918	66.438
0.46	69.959	46.499	67.713		0.836	0.682	0.823		67.123	71.918	66.438
0.47	69.937	46.456	67.688		0.836	0.682	0.823		67.123	71.233	66.438
0.48	69.914	46.412	67.663		0.836	0.681	0.823		66.438	71.233	66.438
0.49	69.891	46.367	67.637		0.836	0.681	0.822		67.123	71.233	66.438
0.50	69.867	46.322	67.611		0.836	0.681	0.822		67.808	71.233	66.438
0.51	69.843	46.276	67.584		0.836	0.680	0.822		67.808	71.233	66.438
0.52	69.818	46.229	67.556		0.836	0.680	0.822		67.808	71.233	66.438
0.53	69.793	46.182	67.528		0.835	0.680	0.822		69.178	71.233	66.438
0.54	69.767	46.133	67.500		0.835	0.679	0.822		70.548	70.548	66.438
0.55	69.741	46.084	67.471		0.835	0.679	0.821		70.548	70.548	66.438
0.56	69.714	46.034	67.441		0.835	0.678	0.821		70.548	70.548	66.438
0.57	69.687	45.984	67.411		0.835	0.678	0.821		70.548	70.548	67.123
0.58	69.659	45.933	67.380		0.835	0.678	0.821		70.548	70.548	67.123
0.59	69.631	45.881	67.349		0.834	0.677	0.821		70.548	71.233	66.438
0.60	69.603	45.828	67.317		0.834	0.677	0.820		70.548	71.233	66.438

Table 8: δ value from 0.40 to 0.60 at 0.01 interval for the second cycle forecast.

Delta	Second Cycle									
	R^2				PPMC				Percentage of Accuracy	
	26	100	200		26	100	200		26	100	200
0.40	55.397	29.345	66.765		0.744	0.542	0.817		62.329	68.493	71.918
0.41	55.352	29.319	66.684		0.744	0.541	0.817		62.329	68.493	71.233
0.42	55.307	29.292	66.603		0.744	0.541	0.816		62.329	68.493	71.233
0.43	55.262	29.265	66.520		0.743	0.541	0.816		62.329	68.493	71.233
0.44	55.217	29.236	66.437		0.743	0.541	0.815		61.644	68.493	70.548
0.45	55.173	29.207	66.353		0.743	0.540	0.815		61.644	68.493	70.548
0.46	55.128	29.176	66.269		0.742	0.540	0.814		61.644	69.178	70.548
0.47	55.084	29.145	66.184		0.742	0.540	0.814		61.644	69.178	70.548
0.48	55.040	29.113	66.098		0.742	0.540	0.813		61.644	69.178	70.548
0.49	54.996	29.081	66.011		0.742	0.539	0.812		61.644	69.178	71.233
0.50	54.952	29.047	65.924		0.741	0.539	0.812		61.644	68.493	71.233
0.51	54.909	29.013	65.836		0.741	0.539	0.811		61.644	69.178	71.233
0.52	54.865	28.978	65.747		0.741	0.538	0.811		61.644	69.178	71.233
0.53	54.822	28.942	65.658		0.740	0.538	0.810		61.644	69.178	71.233
0.54	54.778	28.906	65.568		0.740	0.538	0.810		61.644	69.178	71.233
0.55	54.736	28.868	65.478		0.740	0.537	0.809		61.644	69.178	70.548
0.56	54.693	28.830	65.387		0.740	0.537	0.809		61.644	69.178	70.548
0.57	54.650	28.792	65.295		0.739	0.537	0.808		61.644	68.493	70.548
0.58	54.608	28.752	65.204		0.739	0.536	0.807		61.644	68.493	70.548
0.59	54.565	28.712	65.111		0.739	0.536	0.807		61.644	68.493	71.233
0.60	54.523	28.672	65.018		0.738	0.535	0.806		61.644	68.493	71.233

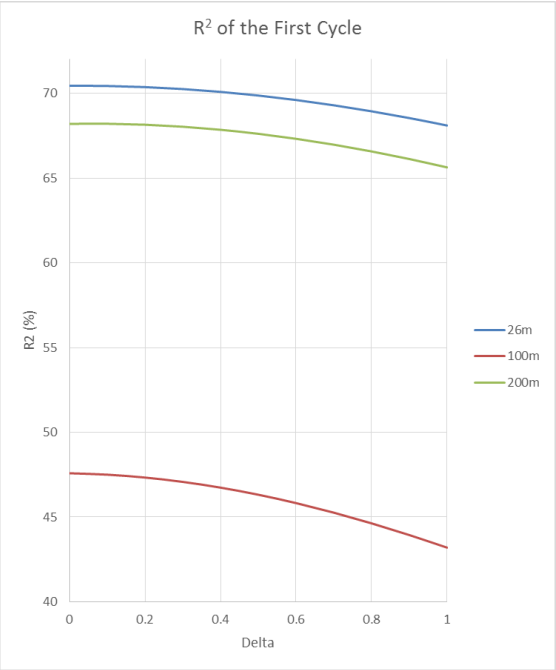


Figure 41: R^2 plots for δ from 0 to 1 at 3 different depths for the first cycle forecast.

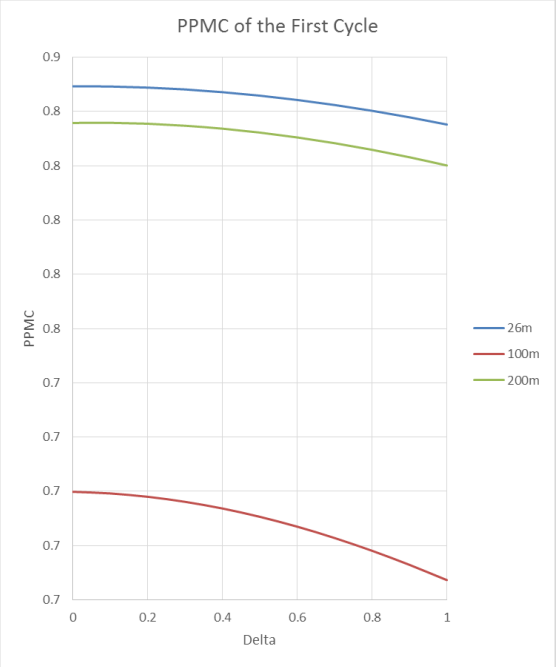


Figure 42: Pearson product moment correlation plots for δ from 0 to 1 at 3 different depths for the first cycle forecast.

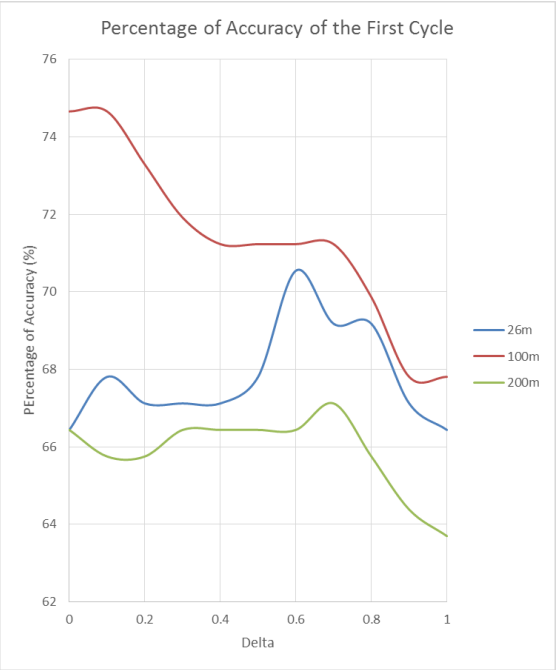


Figure 43: Percentage of accuracy plots for δ from 0 to 1 at 3 different depths for the first cycle forecast.

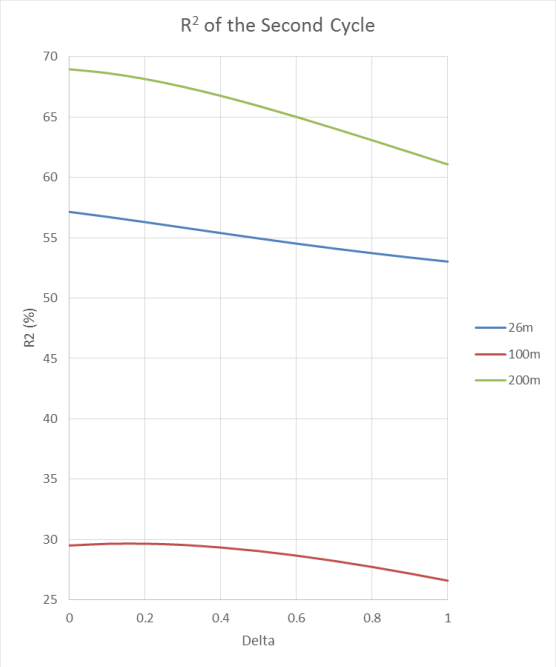


Figure 44: R^2 plots for δ from 0 to 1 at 3 different depths for the second cycle forecast.

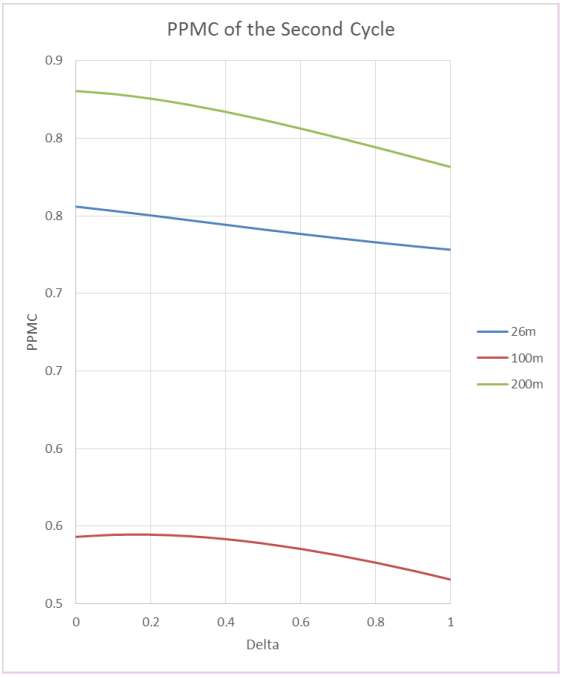


Figure 45: Pearson product moment correlation plots for δ from 0 to 1 at 3 different depths for the second cycle forecast.

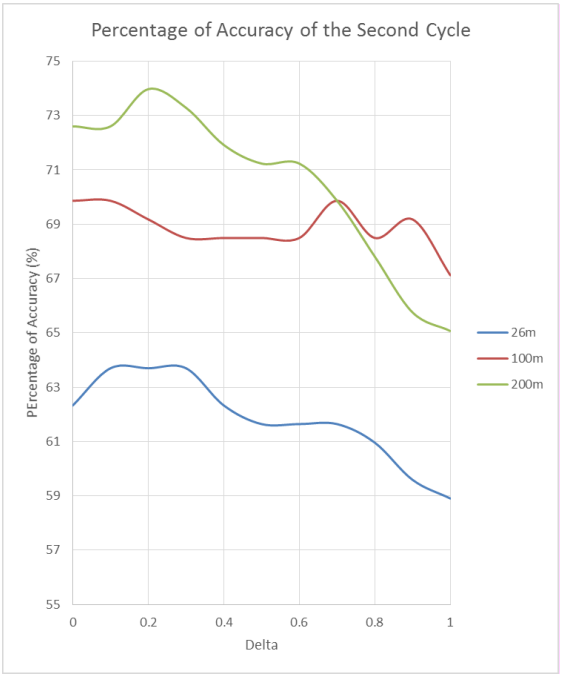


Figure 46: Percentage of accuracy plots for δ from 0 to 1 at 3 different depths for the second cycle forecast.

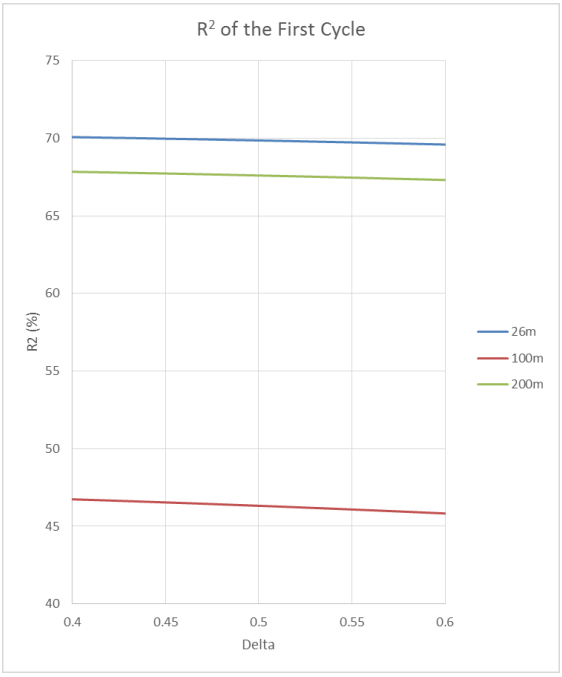


Figure 47: R^2 plots for δ from 0.40 to 0.60 at 3 different depths for the first cycle forecast.

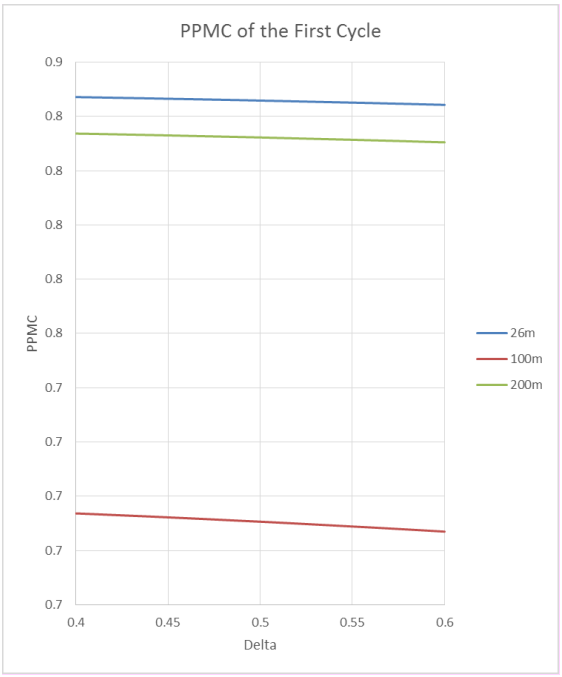


Figure 48: Pearson product moment correlation plots for δ from 0.40 to 0.60 at 3 different depths for the first cycle forecast.

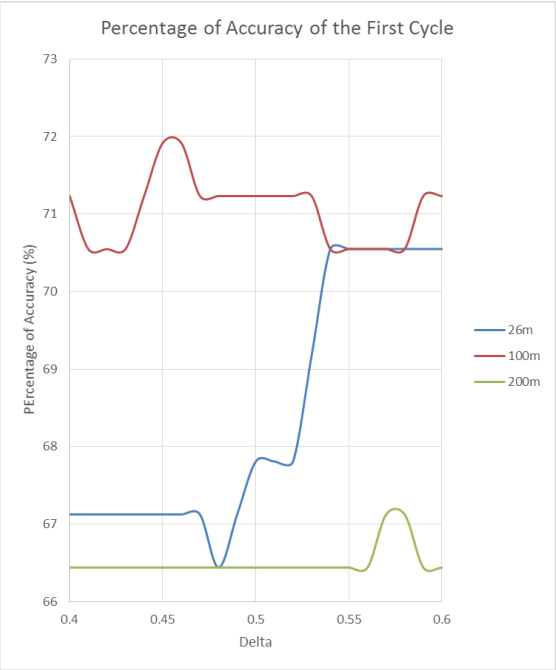


Figure 49: Percentage of accuracy plots for δ from 0.40 to 0.60 at 3 different depths for the first cycle forecast.

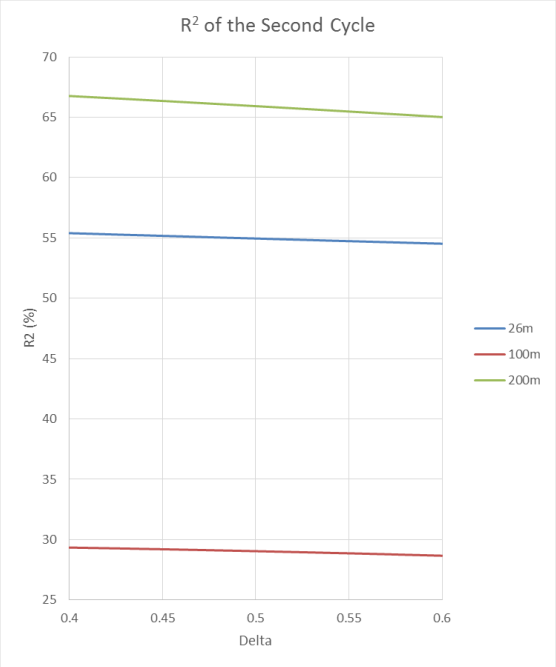


Figure 50: R^2 plots for δ from 0.40 to 0.55 at 3 different depths for the second cycle forecast.

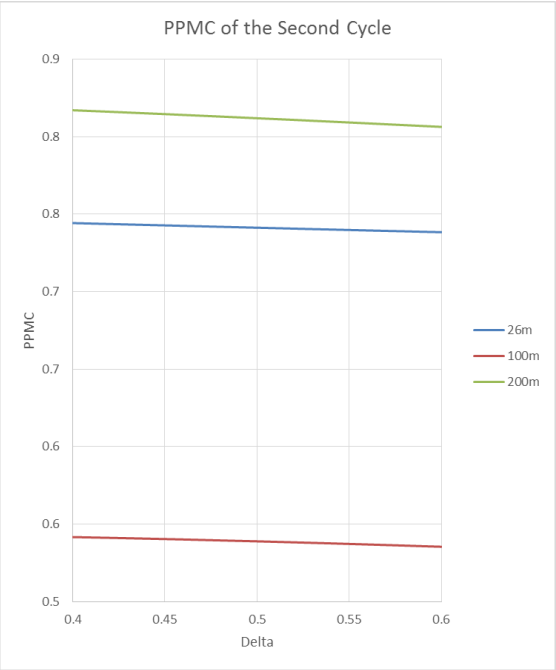


Figure 51: Pearson product moment correlation plots for δ from 0.40 to 0.60 at 3 different depths for the second cycle forecast.

Comment [S15]: LABEL AND DESCIRBE ALL YOUR APPENDIXES!!!! Use figures and tables

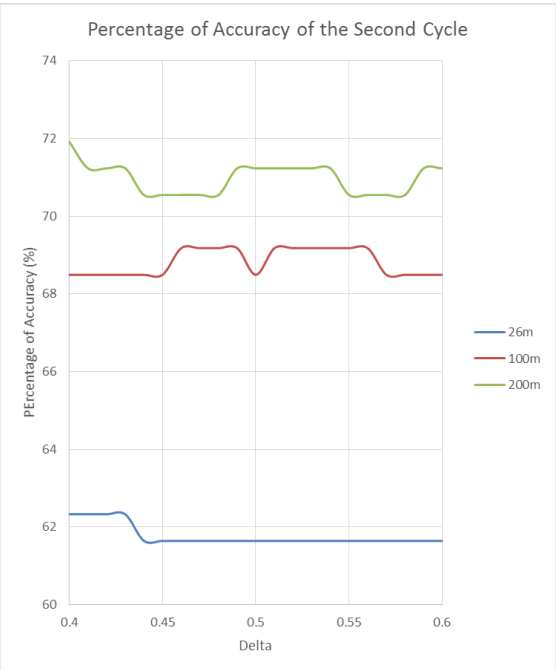


Figure 52: Percentage of accuracy plots for δ from 0.40 to 0.60 at 3 different depths for the second cycle forecast.