SHIP STATE AND COVARIANCE PROPAGATION USING TIME SERIES ANALYSIS AND FORECASTING

BY

ANIRUDDHA KATRE

Submitted in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical and Aerospace Engineering
in the Graduate College of the
Illinois Institute of Technology

Approved ______________________________
Advisor

Chicago, Illinois
May 2014

ACKNOWLEDGMENT

I would like to thank my advisor Dr. Boris Pervan for providing me the opportunity to work on such a unique project. His invaluable guidance, patience and enthusiasm during the ups and downs of this research was a great source of inspiration. I would also like to thank Professor Seebany Datta-Barua, and Professor David Williams for serving on the oral defense committee and for the time they set aside to review my thesis. Special thanks go to Mark Cooper for providing the sea trial data required for the analysis presented in this thesis. I would also like to acknowledge NAVAIR and Engility Corporation for funding this research.

I am extremely grateful to Dr. Steven Langel for his support and encouragement throughout this process. His willingness to spend time engaging in countless discussions as well as the ideas and suggestions he brought to the table have contributed significantly towards making this work possible. I would like to thank Dr. Samer Khanafseh and Dr. Fang Chan for their willingness to answer questions and improve my understanding of the background for this work, and my colleagues Dr. Mathieu Joerger, Stefan Stevanovic, Jing Jing, Naeem Roshan and Raisa Vitto for making this past year memorable. I also owe special thanks to my friends Jaya, Aditya, Suhasini, Manpreet and Faizan for making my years at IIT enjoyable.

Finally, my deepest thanks go to my family, especially my mother Aparna Katre and my sister Madhura for always supporting me and for being a source of energy, inspiration and encouragement. Words cannot describe my appreciation for all the sacrifices they have made to provide me with the opportunities that are now available to me. I dedicate this thesis to them.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The Navy's Unmanned Combat Aerial System (N-UCAS) program is currently developing technology for autonomous shipboard landing of unmanned aerial vehicles (UAVs). A high accuracy and high integrity relative navigation algorithm using carrier phase differential GPS measurements and high rate inertial sensor data has been implemented to land the UAV. Such an algorithm requires ship state information to be broadcast to the aircraft via a VHF data link. The data link is susceptible to failure for reasons such as interference from jamming. This thesis considers the problem of precise ship state propagation on board an aircraft during a data link outage occuring when it is too late for the aircraft to abort its landing approach.

Accurately quantifying the estimate error covariance is important for a high integrity and accuracy navigation algorithm. Therefore this thesis focuses on algorithms that can propagate the ship state as well as determine the propagation error covariance. Initially, a simple state propagation using kinematic equations for linear motion is tested. Seakeeping and Maneuvering theories used to model the dynamics of a sea vessel are also considered. However, analysis shows that due to constraints imposed by a data link outage and complexity in accurately modeling some parameters in the ship dynamic model, these two approaches are infeasible.

As an alternative to kinematic and dynamic modeling, Time Series Analysis and Forecasting methods for ship state propagation are investigated. This work introduces parametric modeling and forecasting of a time series using linear stochastic models. Maximum likelihood estimate (MLE) and outer product of gradients (OPG) algorithms are implemented for faithful parameterization of time series using ARIMA models. Expressions for forecasting and forecast error variance quantification are also developed. These algorithms are then tested using ship data provided by the N-UCAS program.

CHAPTER 1
INTRODUCTION

The Navy's Unmanned Combat Air System (N-UCAS) is currently being developed to support autonomous shipboard landing of unmanned aerial vehicles (UAV). Carrier phase differential GPS measurements and high rate inertial sensor data are used to provide high accuracy and high integrity relative navigation. At the core of the navigation architecture is a tightly coupled GPS/INS estimation algorithm that provides estimates of the relative position and velocity vectors. This algorithm was implemented in real-time software and successfully landed an unmanned combat air vehicle (X-47B) onboard the USS H.W. Bush [Sub13] on July 10, 2013. This demonstration proved the feasibility of shipboard landing of an autonomous UAV under nominal conditions. However, hardware such as sensors is susceptible to failures. Therefore, the possibility of these failures must be properly accounted for by the navigation system. One such failure that is particularly relevant is the possible outage of the data link onboard the aircraft carrier during a landing approach. The data link is responsible for broadcasting the processed ship GPS and inertial sensor data to the aircraft. Without the broadcast information, the aircraft is unable to estimate its position vector relative to the ship. To avert this problem, it is necessary for the aircraft to be capable of propagating the ship state.

The N-UCAS program has strict integrity risk and accuracy budgets associated with an aircraft landing approach. For an aircraft approaching the aircraft carrier at a glide slope of 3 degrees, the allowable vertical error is 1.8 meters (see [KL11]). Similarly the integrity risk, defined as the probability of the estimate error exceeding the allowable error budget, must be $I \leq 10^{-7}$. Considering all possible factors of the navigation algorithm that contribute to the error, it is necessary to minimize the state propagation error to ensure a minimal impact on the landing procedure. To evaluate if integrity and accuracy criteria are met the estimate error covariance for the state estimates needs to be quantified. Therefore any algorithm used to propagate the state must be capable of quantifying the estimate error covariance.

Prior to considering possible methods for state propagation, it is necessary to define the requirements for landing a UAV on the aircraft carrier. The problem is constrained in a realistic manner by making assumptions about the specifics of the landing approach. It is assumed that the maximum time requirement for state propagation in a communication outage situation is 15 seconds. A data link failure more than 15 seconds prior to touchdown will provide enough time for the aircraft to abort its landing procedure successfully. It is also assumed that the ship does not turn drastically in the event of a data link outage. However, the ship may exhibit motion characteristics resembling a *slight* turn. It is also expected for some ship states to display periodic trends.

The ship *state* as relevant to the aircraft for its landing approach consists of the position of the ship center of motion (COM) and ship orientation, i.e. roll, pitch and yaw angles. Other components such as the COM and angular velocities and accelerations may also be needed depending upon the method of propagation. Additionally, the aircraft must be able to estimate the ship position accurately using only the information known. The information about the ship known to the UAV at the time of data-link outage is the broadcasted past time history of ship position, velocity, acceleration, and the associated error variances on these quantities.

1.1 Kinematic Modeling

Kinematic equations may be used to model and propagate ship motion. Given knowledge of the ship COM acceleration, velocity and position at the time of the outage, the kinematic equations used to describe the motion of the ship COM are,

$$\begin{aligned} x_k &= x_{k-1} + v_k \cdot \Delta t \\ v_k &= v_{k-1} + a_k \cdot \Delta t \\ a_k &= const \end{aligned} \tag{1.1}$$

where, x is the position, v is the velocity, a is the acceleration, k is the time index and $\Delta t = 0.05s$.[1] Since there is no available method to update the acceleration term after the data link fails, the best position estimate using expressions in Eq. 1.1 is obtained by assuming a constant acceleration over the time of propagation. This method is a particularly simple but useful approach since it only uses the position, velocity and acceleration broadcast at the time of the outage to calculate the forecasts.

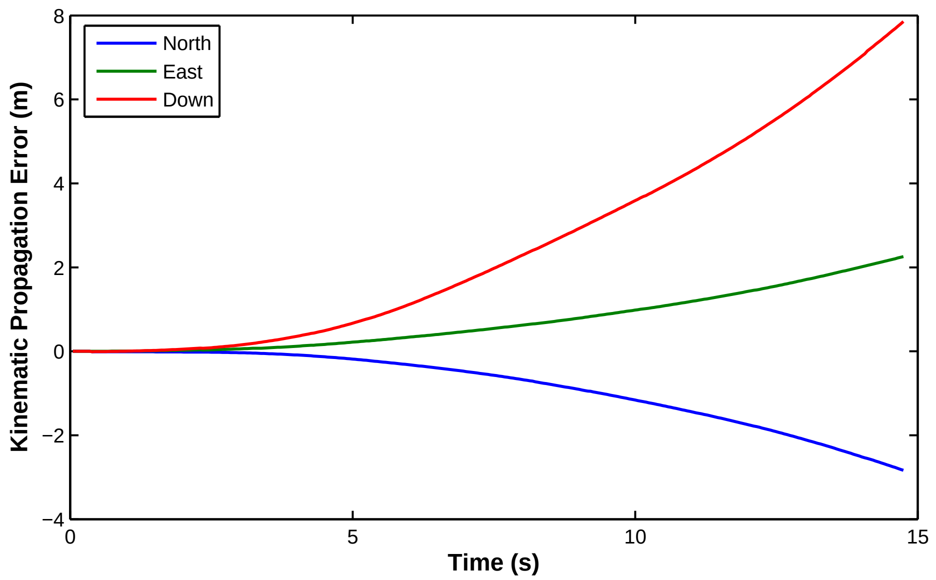

Figure 1.1. Ship state propagation using kinematics

However, the ship state estimate for position does not meet the accuracy and integrity requirements. Fig 1.1 shows an example of error in ship state propagation over a duration of 15 seconds. A random start point was chosen from the ship

[1] Δt is chosen to be 0.05 because the data is collected at a frequency of 20 Hz.

data provided by N-UCAS program and the position values from that point were propagated using expressions in Eq. 1.1. The error in the north, east and down components of the position estimate at the end of the propagation window exceeds the error budget for the complete navigation algorithm by a significant margin. This result shows that it is impossible to achieve the integrity requirements and therefore kinematic modeling is not a feasible approach to propagate the ship state.

Another simple approach for modeling the ship motion is to implement a curve fitting algorithm. However, this method experiences problems similar to the kinematic modeling approach. A curve fit for the data requires at least some prior knowledge of the trends in the data to obtain the best fit. Automating such a scenario will likely result in an inaccurate or inconsistent curve fit since a curve fitting approach cannot account for random disturbances such as gusty winds and changes in magnitude of sea waves and their effects on the ship. Therefore, it is very difficult to develop a robust method that achieves the integrity and accuracy requirements for ship state propagation by a significant margin using a curve fit.

1.2 Ship Dynamic Modeling

A more common method to modeling the ship state is by using a faithful dynamic model. A dynamic model is preferred over most other approaches because it takes into account the properties of the system such as mass and inertia, as well as the external forces affecting it through the system equations of motion. Well developed theories such as the Seakeeping and Maneuvering theories [Per05] are capable of modeling the dynamics of a ship and its response to external forces under different sea state conditions.

1.2.1 Maneuvering Theory. The Maneuvering Theory model for ship dynamics is typically used to design the control system of a ship. This theory deals with the motion of ships in absence of disturbance force, i.e. in calm water and is associated with maneuvers such as course changes, stopping, etc. In this theory, ship motion caused by the action of control devices such as control surfaces and propulsion units [Per05] is modeled. The equations of motion developed using this theory are formulated in a reference frame fixed to the ship (b-frame), and not in an inertial frame. Figures 1.2 and 1.3 show the specifics of the b-frame.

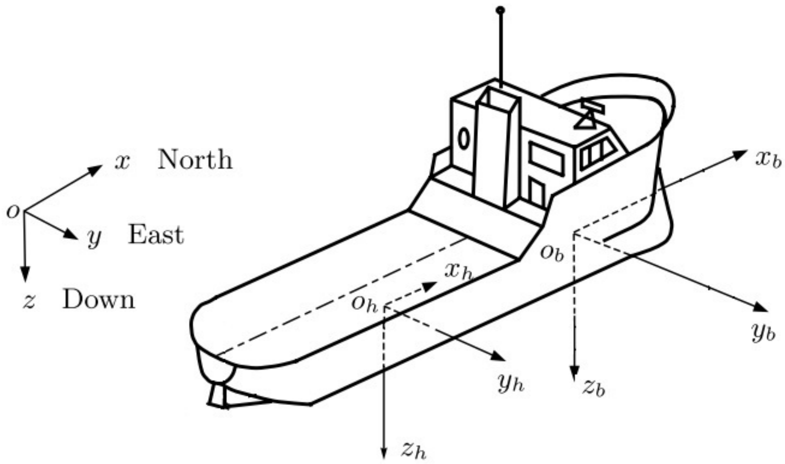

Figure 1.2. Hydrodynamic and Body Frames. Figure is adapted from [Per05]

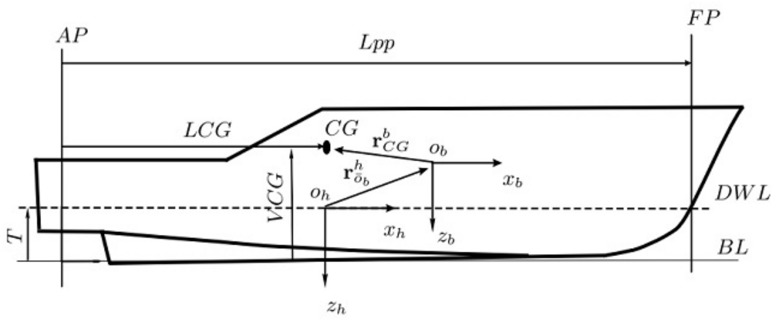

Figure 1.3. Ship geometry description and reference frames: hydrodynamic frame (origin o_h); body frame (origin o_b); CG - center of gravity; LCG - lateral center of gravity; VCG - vertical cener of gravity; AP - aft perpendicular; L_{pp} - length between perpendiculars; T - draught; DWL - design waterline and BL - baseline. Figure is adapted from [Per05]

The equations of motion obtained from Maneuvering theory as derived by [Fos94] are,

$$\begin{aligned}\left[\boldsymbol{M}_{RB}^{b}+\boldsymbol{M}_{A}^{b}\right]\dot{\boldsymbol{\nu}}+\boldsymbol{C}^{b}(\boldsymbol{\nu})\boldsymbol{\nu}+\boldsymbol{D}^{b}(\boldsymbol{\nu})\boldsymbol{\nu}+\boldsymbol{g}^{b}(\boldsymbol{\eta})&=\boldsymbol{\tau}^{b}\\ \dot{\boldsymbol{\eta}}&=\boldsymbol{J}_{b}^{n}(\boldsymbol{\Theta}_{nb})\boldsymbol{\nu}\end{aligned}\tag{1.2}$$

where, $\boldsymbol{M}_{RB}^{b}$ is the rigid body generalized mass and inertia matrix with respect to the origin of the b-frame; $\boldsymbol{M}_{A}^{b}$ is the generalized added mass matrix; $\boldsymbol{\nu}$ is the generalized (linear and angular) velocity vector given in the body frame; $\boldsymbol{C}^{b}(\boldsymbol{\nu})$ is the total (rigid body and added mass) Coriolis and centripetal acceleration matrix; $\boldsymbol{D}^{b}(\boldsymbol{\nu})$ is the damping matrix; $\boldsymbol{g}^{b}(\boldsymbol{\eta})$ is the restoring function; $\boldsymbol{J}_{b}^{n}(\boldsymbol{\Theta}_{nb})$ is the velocity transformation matrix from the inertial n-frame to the b-frame; $\boldsymbol{\tau}^{b}$ is the vector of forces and moments acting on the hull originating from the control devices, propulsion system and hydrodynamic effects.

The use of this theory and the derived equations of motion shown in Eq 1.2 is not a good method to estimate the ship state for computations done off-board the ship. [Per05] shows that many components of the equations of motion depend specifically on ship parameters such as the ship geometry and evaluating them requires computationally expensive and time consuming simulations. Furthermore, this theory is based on the assumption that the ship is conducting maneuvers such as turns and stops in calm water conditions. This not only complicates the aircraft landing approach, but it also does not satisfy the assumptions of constant heading and constant average velocity made to formulate the aircraft approach problem problem. Also, many terms contained in the equations of motion are not included in the data broadcast to the aircraft by the aircraft carrier. Therefore, it is not feasible to use the Maneuvering theory to accurately model and propagate the ship state.

1.2.2 Seakeeping Theory. In Seakeeping, the ship motion is studied under the assumption that the ship is moving at a constant heading and at a constant-average forward speed [Per05]. The constant heading and average-speed is defined as the equilibrium state. Since the assumptions that form the basis of the seakeeping theory match the assumptions made to formulate the problem in this thesis, the seakeeping theory is a potential choice for modeling the ship state.

To define the vector equations of motion, the hydrodynamic reference frame must first be described. As given in [PF07], the hydrodynamic frame (h-frame) is an inertial reference frame defined as forward-starboard-down. It has an origin at the coordinates x_h, y_h, z_h and the frame is fixed to the ship trajectory defined by its constant heading and constant average speed, i.e. it moves at the the ships average speed and follows the vessel path. Figures 1.2 and 1.3 show the location of the hydrodynamic frame on the ship. Therefore, the ship motion is then expressed as oscillations with respect to the h-frame.

The ship equations of motion are most commonly expressed in the hydrodynamic frame. [Per05] shows that the vector equation of motion for the ship in the *hydrodynamic frame* is,

$$\boldsymbol{M}^h_{RB}\ddot{\boldsymbol{\xi}} = \boldsymbol{\tau}^h_{hyd} \tag{1.3}$$

where, $\boldsymbol{\xi}$ are the seakeeping coordinates that describe the ship's oscillatory motion with respect to the hydrodynamic frame, $\boldsymbol{M}^h_{RB}$ is the rigid body mass matrix consisting of mass and inertia terms with respect to the h-frame, and $\boldsymbol{\tau}^h_{hyd}$ is the force vector consisting of hydrodynamic forces and moments expressed in the h-frame. The components of the mass matrix are,

$$\boldsymbol{M}_{RB}^{h} \triangleq \begin{bmatrix} m & 0 & 0 & 0 & mz_g^h & -my_g^h \\ 0 & m & 0 & -mz_g^h & 0 & mx_g^h \\ 0 & 0 & m & my_g^h & -mx_g^h & 0 \\ 0 & -mz_g^h & my_g^h & I_x^h & -I_{xy}^h & -I_{xz}^h \\ mz_g^h & 0 & -mx_g^h & -I_{yx}^h & I_y^h & -I_{yz}^h \\ -my_g^h & mx_g^h & 0 & -I_{zx}^h & -I_{zy}^h & I_z^h \end{bmatrix} \tag{1.4}$$

where, m is the mass of the ship, x, y and z are the coordinates of the ship C.G. expressed in the h-frame and I is the inertia of the ship. The information broadcast to the aircraft via the data link does not include the ship mass and inertia, thereby increasing the difficulty of implementing a dynamic modeling approach to state propagation.

The force vector affecting the ship motion can be decomposed as follows,

$$\boldsymbol{\tau}_{hyd} = \boldsymbol{\tau}_{1w} + \boldsymbol{\tau}_{2w} + \boldsymbol{\tau}_{r} + \boldsymbol{\tau}_{v} + \boldsymbol{\tau}_{hs} \tag{1.5}$$

where, the $\boldsymbol{\tau}_{1w}$ are first-order wave excitation forces, $\boldsymbol{\tau}_{2w}$ are second order wave excitation forces, $\boldsymbol{\tau}_{r}$ are the radiation forces, $\boldsymbol{\tau}_{v}$ are the viscous forces and $\boldsymbol{\tau}_{hs}$ are the hydrostatic forces (see [Per05]). Modeling these forces accurately is the most difficult part of constructing a faithful dynamic model of a ship. Apart from the viscous forces, all other forces can be studied within a linear framework. [Per05] shows that these forces are then best modeled using wave force response amplitude operators (FRAO). The response amplitude operators (RAO) quantify the response of a ship to sinusoidal inputs at varying frequencies and obtain force transfer functions. These force transfer functions are then combined with the sea spectrum and sailing conditions to obtain the forces acting on a ship.

To calculate the forces acting on a ship using FRAOs, the state of the sea needs to be quantified and a sea spectrum obtained for a particular sea state. [oW02] tabulates all the commonly used sea-spectra and the operating variables such as the wave frequency, wave height, windspeed that are needed to evaluate each spectrum. Sea spectra, such as the Pierson-Moskovitz spectrum, Bretschneider spectrum [Bre59], JONSWAP spectrum [ea73], Scott spectrum [Sco65], and many more are all valid choices for wave spectrum, however they are accurate for different sea states. The most commonly used definition of the sea state is the Douglas [Dou91] sea state which characterizes the roughness of the sea based on factors such as the significant wave height, i.e. the wave height with respect to the nominal sea level, and the frequency of the waves.

The wave spectra and sea state models are all based on empirical studies wherein the required information regarding sea states was measured to develop the models. Furthermore, different parts of the world experience drastically different sea state conditions over the period of a year. The data broadcast via the data-link does not include any information about the sea state thus making it impossible to choose the requisite models to accurately calculate the forces. Therefore, considering all these factors along with uncertainties in the empirical models, it is not easy to quantify the forces affecting a ship and the uncertainty in the force estimate.

These theories are typically used for designing the ship and its control systems. They are not ideal for implementation in real time estimation algorithms due to high requirements on computational power and time. Additionally, the ship equations of motion have significant unquantified errors due to uncertainties in the force modeling and the aircraft's lack of knowledge of the ship mass and inertia properties. Therefore, the dynamic models formed using various ship dynamics theories cannot be used for ship state propagation for the UAV landing problem.

1.3 System Identification

A completely different approach from both kinematic and dynamic modeling of ship motion, is using a system identification method. Identification of a system consists of generating a faithful mathematical and statistical model using the available data. A classical approach in system identification to model and predict a set of data is Time Series Analysis and Forecasting.

A time series is a sequence of data collected at successive uniform time intervals [BJR94]. A daily measure of the quantity of goods sold by a store, or data collected hourly of the yield of a chemical process are examples of a time series. Time series are used extensively in various fields such as signal processing [Ric63], control engineering [LW72], econometrics [Gud71], astronomy [Yul27], oceanography [GH68]. In order to use the sequence of data effectively in a variety of such applications, statistical properties and other characteristics of the time series must be extracted from the data and modeled faithfully. Time Series Analysis and Forecasting is used to determine such properties associated with the set of data, with the intention of predicting the future values of that time series. In this work, the ship data broadcast to the aircraft prior to data link outage can be treated as individual time series that have unique statistical properties thus making time series analysis and forecasting an applicable method for state propagation.

Time Series Analysis is comprised of two distinct methodologies; parametric modeling and non-parametric modeling. More common of the two due to various properties and resulting advantages is the idea of parametric modeling of a time series [Pol99]. If it was possible to exactly model a time series, such that any calculation relating to the time series is exact, the resulting model would be called a deterministic model [BJR94]. However, in most real applications random disturbances affect the system response and thus error free calculations are not possible. In such systems, it is possible to derive a model that can calculate the *probability* of a value of the time series lying between two specific limits. This method of modeling is called stochastic modeling [BJR94]. Assuming that the response to of a ship to disturbance forces can be modeled as a random process, stochastic models can be used to parameterize the ship state measurements. This work investigates Time Series Analysis and Forecasting as a method for robust modeling and propagation of ship state.

1.4 Contributions and Summary

This work investigates the use of different methods for propagation of ship state and covariance in an aircraft-based algorithm. The specific focus of this thesis is to investigate methods to land a UAV on an aircraft carrier during a data link outage conditions when it is too late to abort the landing approach. The methods are constrained by the available knowledge of ship motion which ends at the onset of communication outage as well as stringent integrity and accuracy requirements. Unfortunately, as discussed in this chapter, standard approaches such as kinematic and dynamic modeling are not implementable in this scenario due to the problem constraints. Therefore, alternate methods are explored in this work.

1.4.1 Contributions. The contributions of this work are as follows.

- The kinematic approach to ship motion propagation is investigated and shown to be infeasible for the problem of autonomous shipboard landing of unmanned aerial vehicles.
- The dynamic modeling approach to ship motion propagation is also shown to be infeasible for autonomous shipboard landing of unmanned aerial vehicles.
- Time Series Analysis is proposed as an alternative to kinematic and dynamic mod-

eling. For this method to be feasible, an optimal model identification method, a method for forecasting a time series, and a method to quantify the forecast error variance are required. In this work, the Maximum Likelihood Estimate (MLE) algorithm was implemented successfully for optimal model identification. An expression for the minimum mean squared error forecast was obtained and implemented for a set of sea trial data provided by the N-UCAS program. Finally, an expression to quantify the forecast error variance was obtained and tested using the sea trial data with encouraging results. The successful implementation of these three methods shows that Time Series Analysis is a feasible approach to autonomously propagate ship motion on an aircraft.

1.4.2 Summary of Chapters. The summary of chapters in this thesis is as follows.

- Chapter 1 forms the groundwork for the problem of ship state propagation in the high integrity and accuracy application of shipboard landing of unmanned aerial vehicles during communication outage. Several common methods for state propagation are investigated and their suitability for this problem evaluated. The methods studied include simple kinematic modeling, curve fitting and dynamic modeling. Work presented in this chapter shows that due to the constraints on information available as well as the fundamental nature of the problem, it is not possible to implement these methods for real time state propagation.
- Chapter 2 introduces the concept of time series analysis by identifying the categories of time series and expressions used to parameterize them. Specifically, the family of stochastic linear models such as the Autoregressive (AR) and Moving Average (MA) models are introduced. The constraints on the parameters of the models and their relevance to parameterizing a time series are also discussed.
- Chapter 3 discusses the Maximum Likelihood Estimation (MLE) method of iden-

tifying the parameters of the stochastic model accurately. Methods to quantify the covariance on the parameter estimate error are also discussed. Simulation results have been provided verifying that these algorithms for model identification satisfy the necessary criteria for implementation in high integrity applications.

- Chapter 4 includes a discussion of methods for forecasting a time series. Forms of the stochastic Autoregressive Integrated Moving Average (ARIMA) model are discussed and their properties highlighting different aspects of the forecasting problem are identified. Methods to quantify forecast error covariance are also introduced.
- Chapter 5 shows results from the implementation of time series modeling and forecasting methods on real ship data provided by N-UCAS program. The MLE algorithm is successfully implemented to parameterize the ship position data using the ARIMA model. The forecasting method is also successfully implemented and the forecast error variance associated with the forecasts is analyzed.

CHAPTER 2
TIME SERIES ANALYSIS

A general approach to parametric modeling of a time series involves optimal estimation of a set of parameters that uniquely defines the time series. It is advantageous to implement parametric modeling methods in the absence of a system dynamic model since a well defined mathematical and statistical model is important for a faithful state propagation under stringent integrity requirements. Furthermore, these methods are capable of quantifying the forecast error covariance, an important property that is essential for evaluating integrity risk and meeting integrity requirements. This chapter introduces the idea of time series classification and parametric modeling of these time series with the intentition of using them to model and forecast ship state data.

2.1 Time Series Classification

Prior to understanding how parametrization of a time series works, it is necessary to classify the different types of time series. A time series can be categorized as stationary or nonstationary and seasonal or nonseasonal.

2.1.1 Stationary and Nonstationary Time Series. A time series is always either stationary or nonstationary. A time series is considered to be stationary when the joint probability distribution of the time series is independent of a shift in time; i.e. the data contained within the time series has a constant mean, variance and autocorrelation. Intuitively, a time series is stationary if a region of the series "looks" the same as any other of its regions. Figure 2.1 shows an example of a stationary time series

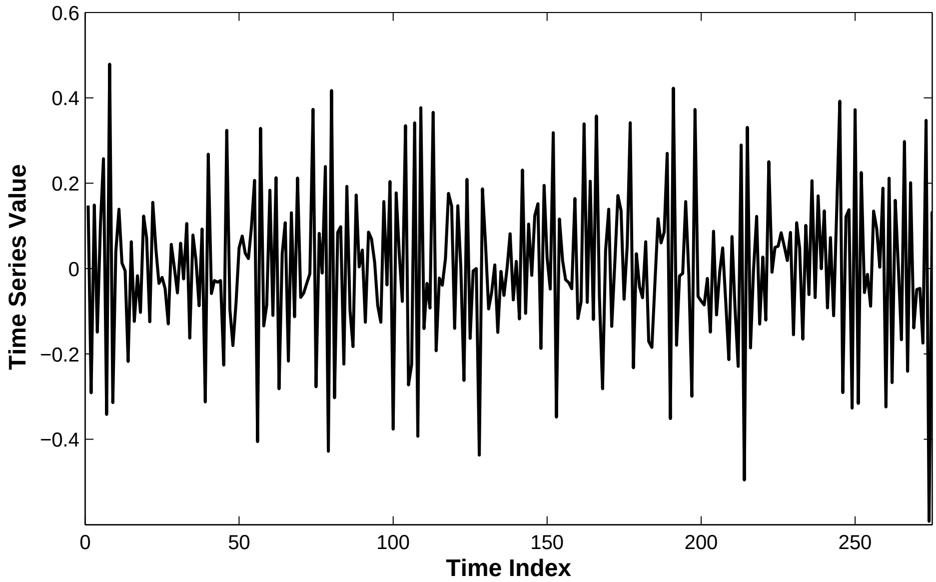

Figure 2.1. Stationary time series

Stationary time series occur frequently in many applications. However, nonstationary time series are more commonly observed in real physical processes. A time series is nonstationary if the mean or the variance of the time series changes over time. Another method to evaluate the stationarity of the time series is to check the autocovariance or autocorrelation (or both) at two different sections of the time series. If these quantities from the two different regions are different, then the time series is nonstationary.

The two of the most common types of nonstationary series seen in real industrial applications are ones nonstationary in level, i.e. nonstationary in mean, and ones that are nonstationary in both level and slope [BJR94]. Figures 2.2 and 2.3 show examples of these two types of nonstationarity. Note that a nonstationarity in level means that the mean of the time series will shift by some value, and then stay constant for a period of time. For example, it is easy to see in Figure 2.2 that the time series undergoes a shift in level after time index 400. Similarly, a time series

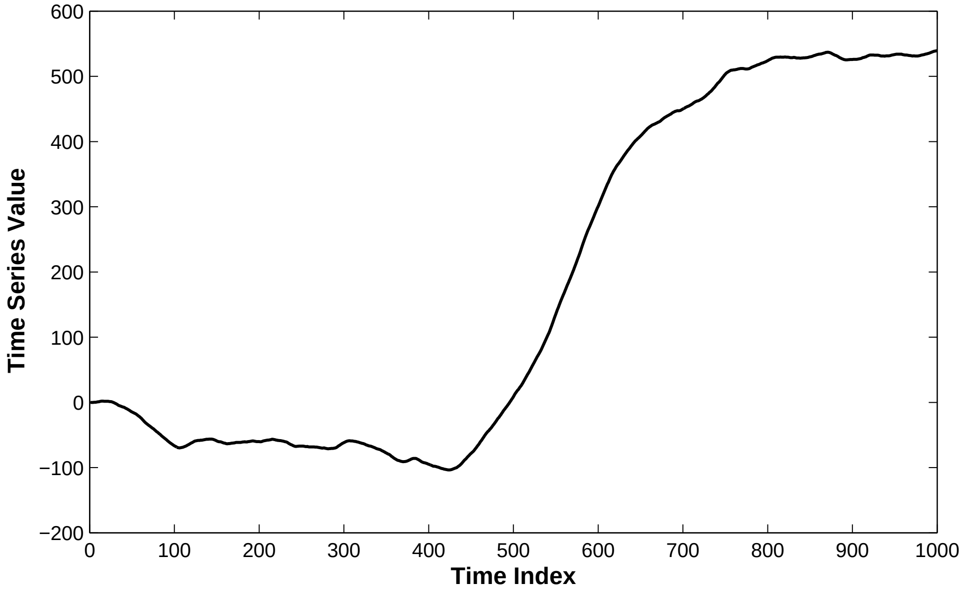

Figure 2.2. Time series nonstationary in level

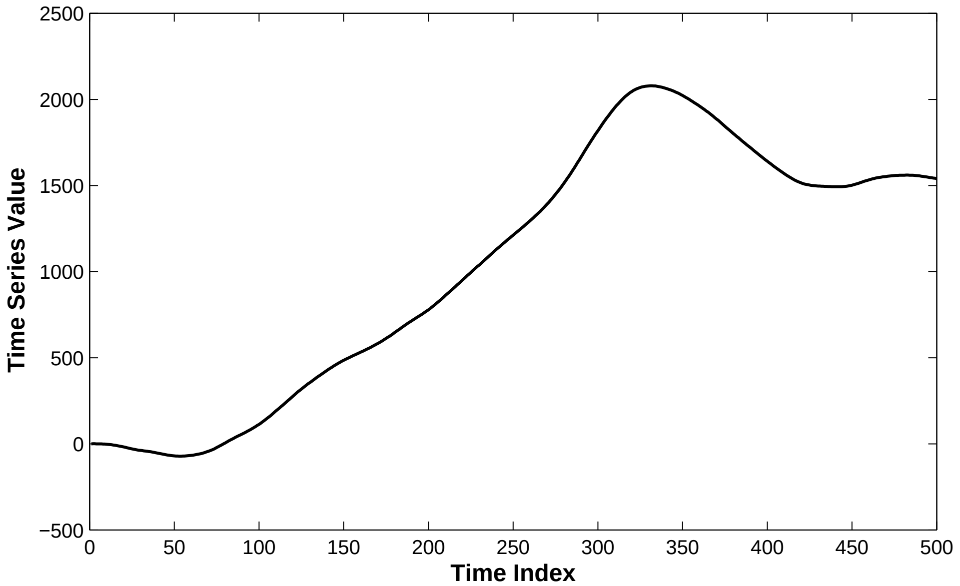

Figure 2.3. Time series nonstationary in both level and slope

with a drifting mean and sections of linear trends, such as the one in Figure 2.3, is nonstationary in both level and slope.

2.1.2 Seasonal and Nonseasonal Time Series. A time series, apart from being stationary or nonstationary, can also be classified as seasonal or nonseasonal. A time series is referred to as a seasonal time series if the data exhibits a periodic trend. The periodic trend does not need to be deterministic, i.e. caused by a deterministic function such as a sine wave. An example of a seasonal time series is shown in Figure 2.4 where the international airline passenger monthly total (in thousands) from January 1949 to December 1956 is modeled. The values in this time series are measured at monthly time intervals. It can clearly be seen that the time series has a periodicity of 12 months where a small peak followed by a larger peak appear in each period. If a time series does not follow a periodic trend, it is called a nonseasonal time series. Both seasonal and nonseasonal time series are commonly observed in a large variety of applications.

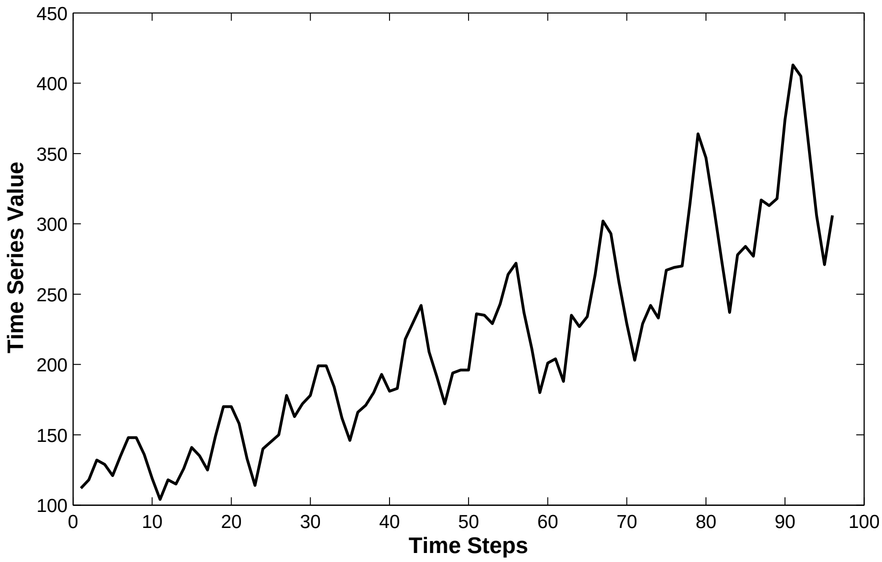

Figure 2.4. Seasonal time series: International airline passenger data in thousands from January 1949 to December 1956 [BJR94]

2.2 Parametric Models

Now that the possible representations of a time series have been categorized, the next step is to understand how the time series can be parameterized. Parametric modeling methods for stationary time series are based on the assumption that a time series is generated by a linear aggregation of random disturbances and thus can be modeled using a general linear stochastic model [BJR94]. The theory behind using linear models for stationary time series can be extended and applied to nonstationary time series as well.

The general linear stochastic model comes in two forms. The first representation of a stochastic process, as shown in [BJR94], models a value in a time series as the output $\tilde{z}_t$ from a linear filter with an input of white noise v_t. That is,

$$\begin{aligned}\tilde{z}_t &= v_t + \psi_1 v_{t-1} + \psi_2 v_{t-2} + \cdots \\ &= v_t + \sum_{j=1}^{\infty} \psi_j v_{t-j}\end{aligned} \tag{2.1}$$

Here, $\tilde{z}_t = z_t - \mu$ is the deviation of the process from some origin. For a stationary stochastic process, $\tilde{z}_t$ is the deviation from the mean of z_t.

The model in 2.1 suggests that, by recursively substituting a rearranged form that solves for v_t, the stochastic process can be represented as an infinite weighted sum of past values in the time series plus a random disturbance. This second representation of the general linear stochastic model is written as,

$$\begin{aligned}\tilde{z}_t &= \pi_1 \tilde{z}_{t-1} + \pi_2 \tilde{z}_{t-2} + \cdots + v_t \\ &= \sum_{j=1}^{\infty} \pi_j \tilde{z}_{t-j} + v_t\end{aligned} \tag{2.2}$$

[Wol38] established that the coefficients ψ_j, from equation 2.1, are absolutely summable for a stationary time series, i.e. $\sum_{j=0}^{\infty} |\psi_j| < \infty$ and thus a zero-mean stochastic stationary process can be represented by the models in equations 2.1 and 2.2.

The finite order representations of 2.1 and 2.2 are known as the Moving Average (MA) model and the Autoregressive (AR) model respectively. These models are based on the assumption that a time series has a finite order time dependency on past measurements or past random disturbances. The AR and MA models can be combined to form the Autoregressive Moving Average (ARMA) model. These models are capable of parameterizing nonseasonal and stationary time series. More complex scenarios such as modeling a seasonal time series or nonstationary time series are handled using more general forms of the ARMA model such as the Autoregressive Integrated Moving Average (ARIMA) model and the Seasonal ARIMA model. This section investigates the AR, MA, ARMA, ARIMA and Seasonal ARIMA models in more detail.

2.2.1 Autoregressive Model. The Autoregressive (AR) Model is one of the most basic models used to parameterize a stationary stochastic process. It models a stationary stochastic time series as a linear combination of a finite amount of past data contained in the sequence and a single random disturbance occuring at the same time as the measurement of interest. For a time series containing measurements z_t, collected at discrete time steps $t = k, k+1, k+2, ...$, the AR model is expressed as

$$\begin{aligned} z_t &= \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + v_t \\ &= \sum_{i=1}^{p} \phi_i z_{t-i} + v_t \end{aligned} \tag{2.3}$$

Here, ϕ_i's are weighting parameters determining the effect of past measurements on the current value of the time series and v_t is a random disturbance affecting the measurement. v_t is modeled as White Gaussian Noise (WGN), i.e. as a Gaussian or Normal random variable (RV) with a mean $\mu = 0$ and variance of σ^2. Standard shorthand notation for WGN is $v_t \sim N(0, \sigma^2)$ where N implies a Gaussian or Normal distribution. The amount of time dependency on past data of an AR model is known as the model order and is denoted by p. A stochastic process modeled using an AR model of order p is known as an AR process and is abbreviated as AR(p).

An alternative way to write the AR model is using the lag operator. For a lag operator defined as

$$B^j z_t = z_{t-j} \quad \text{where } j = 1, 2, ... \tag{2.4}$$

the AR model can be re-written in the form

$$(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) z_t = v_t$$

or

$$\Phi(B) z_t = v_t \tag{2.5}$$

In this form, the AR model can be thought of as the output z_t from a linear filter with a transfer function $\phi^{-1}(B)$ and a WGN input v_t.

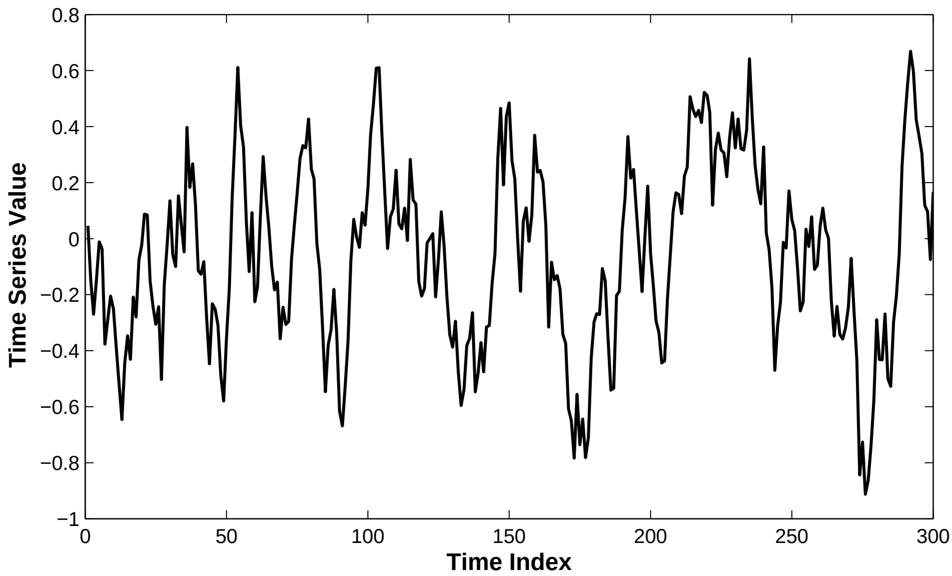

Figure 2.5. Sample AR(2) Process

The AR model is typically used when a time series has relatively low levels of noise and it is obvious that past data affects the current value more than any random disturbances. For example, figure 2.5 shows a time series generated using a second order AR model given by the equation $z_t = 0.95z_{t-1} - 0.15z_{t-2} + v_t$, with $v \sim N(0, 0.15^2)$. It can be seen easily that past values in the time series affect the future values more than any random disturbances.

2.2.2 Moving Average Model.

Similar to the AR Model is another basic model used to parameterize a stationary stochastic process called the Moving Average (MA) model. It models a measurement in a time series as a linear combination of a finite number of past random disturbances weighted by some parameters and a single random disturbance occuring at the same time as the measurement of interest. For a time series of measurements z_t, collected at discrete time steps $t = k, k+1, k+2, ...$, the MA model is expressed as

$$\begin{aligned} z_t &= v_t - \theta_1 v_{t-1} - \theta_2 v_{t-2} - \cdots - \theta_q v_{t-q} \\ &= v_t - \sum_{j=1}^{q} \theta_j v_{t-j} \end{aligned} \tag{2.6}$$

where, $v_t \sim N(0, \sigma^2)$ is the random disturbance at time t and θ_j's are the weighting parameters determining the effect of the current and past random disturbances on the current value of the time series. The amount of time dependency on past random disturbances of an MA model is known as the model order and is denoted by q. A stochastic process modeled using an MA model of order q is known as an MA process and is abbreviated as MA(q).

The MA model expressed in its transfer function form is

$$z_t = \Theta(B) v_t \tag{2.7}$$

where, $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ and B is the lag operator.

Based on the expression in 2.6, it is clear that an MA model is used to model a time series when the time series exhibits pronounced effects of random disturbances.

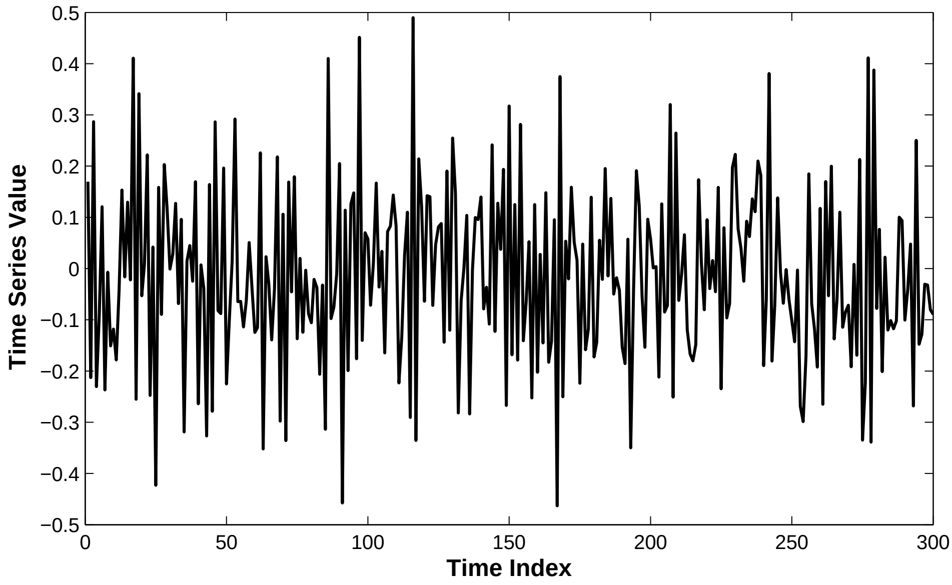

Figure 2.6. Sample MA(2) Process

Figure 2.6 shows a time series generated using a second order MA model given by the equation $z_t = -0.4v_{t-1} + 0.3v_{t-2} + v_t$ where $v \sim (0, 0.15^2)$. It is easily seen that the time series has a more random "look" to it than an AR process.

2.2.3 Autoregressive Moving Average Model.

A more comprehensive model capable of parameterizing a stationary time series is the Autoregressive Moving Average (ARMA) model. It is a combination of the AR and MA processes. In the difference equation form, the ARMA model is written as

$$z_t - \sum_{i=1}^{p} \phi_i z_{t-i} = v_t - \sum_{j=1}^{q} \theta_j v_{t-j} \tag{2.8}$$

where, the left hand side of equation 2.8 is the AR model and the right hand side is the MA model. A stochastic process modeled using an ARMA model with an AR part of order p and an MA part of order q is denoted by ARMA(p,q) and is called an ARMA process.

In its transfer function form, the ARMA model is expressed as

$$z_t = \frac{\Theta(B)}{\Phi(B)} v_t \tag{2.9}$$

In 2.9, the terms $\Phi(B)$ and $\Theta(B)$ are the AR and MA coefficient polynomials respectively.

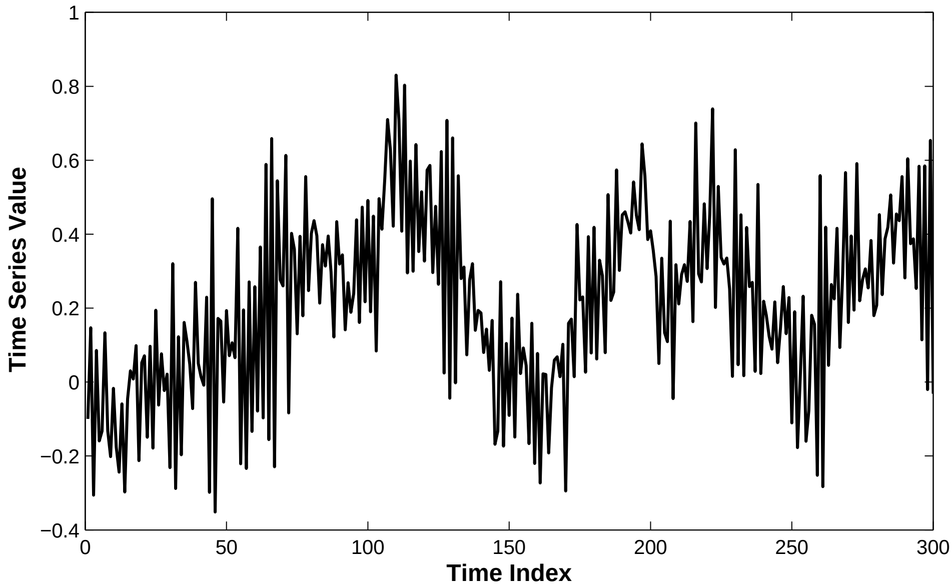

Figure 2.7. Sample ARMA(2,2) Process

The ARMA model is a more comprehensive approach to modeling a stationary stochastic time series. Since it is a combination of the AR and MA models, it inherits all the properties of both models. Consequently, it can account for the time dependency of a value in a time series on past measurements as well as on a sequence of random disturbances. For example, the time series in figure 2.7 is generated using an ARMA(2,2) model given by the equation $z_t = 0.4z_{t-1} + 0.55z_{t-2} - 0.7v_{t-1} + 0.3v_{t-2} + v_t$. It exhibits both the dependency on past values of an AR process and the dependency on past random disturbances of an MA process.

2.2.4 Autoregressive Integrated Moving Average Model. The previous parametric models, i.e. the AR, MA and ARMA models, are capable of handling only stationary stochastic processes. More common in real applications are nonstationary processes such as the ones depicted in Figures 2.2 and 2.3. If ARMA models are used to parameterize a nonstationary time series, the identified parameters will violate the stationarity and invertibility constraints.[2] This negates the purpose of using the parametric modeling approach of time series analysis and forecasting. Therefore, the Autoregressive Integrated Moving Average (ARIMA) model is developed to model nonstationary time series.

The ARIMA model accounts for the nonstationarity by adding an extra parameter to the model. It claims that a nonstationary time series can be modeled by a combination of the ARMA model and a finite number of roots placed at unity, i.e. on the unit circle. This idea is better expressed mathematically by the transfer function form of the ARIMA model.

$$\tilde{z}_t = \frac{\Theta(B)}{\varphi(B)} v_t \tag{2.10}$$

Here, $\tilde{z}_t = z_t - \mu$ with μ as the *level* or mean of the measurements z_t. $\varphi(B)$ is a nonstationary autoregressive operator such that d roots of $\varphi(B)$ are unity and the remaining roots p lie outside the unit circle like an AR model. Thus, the ARI part of the ARIMA model can also be expressed as

$$\varphi(B) = \Phi(B)(1 - B)^d$$

where, $\Phi(B)$ is the *stationary* AR operator of order p. By defining $\nabla^d = (1 - B)^d$ as the difference operator for $d \geq 1$, the ARIMA model can also be written as

$$\Phi(B)\nabla^d z_t = \Theta(B)v_t$$

Typically, an ARIMA model with degree of differencing $d = 1$ models a time series with type one nonstationary such as one depicted in Fig 2.2. Similarly, a time series with type two nonstationarity such as the one shown in Fig 2.3 is modeled by

[2]A comprehensive overview of model parameter constraints is provided in the next section.

setting $d = 2$.

2.2.5 Seasonal Autoregressive Integrated Moving Average Models. The Seasonal ARIMA model is used to augment the capabilities of the ARIMA model. An ARIMA model by itself is not capable of parameterizing periodic trends that may be encountered in the time series. For example, the time series shown in Fig 2.4 cannot be modeled using a simple ARIMA model. A modified version of the ARIMA model, commonly known as the Seasonal ARIMA model introduces extra parameters to handle such periodic trends. The seasonality of a time series, i.e. a periodic trend in the data can be modeled using a simplifying operator given by,

$$B^s z_t = z_{t-s} \tag{2.11}$$

where s is the period of the time series. Therefore, the Seasonal ARIMA model (see [BJR94]), is expressed as,

$$\Phi(B^s)\nabla_s^D z_t = \Theta(B^s)\alpha_t \tag{2.12}$$

where $\Phi(B^s)$ and $\Theta(B^s)$ are polynomials in B^s of degrees P and Q respectively which satisfy the stationarity and invertibility conditions, $\nabla_s^D = (1 - B^s)^D$ accounts for the nonstationary seasonality of the time series and α_t's are the error component of the model. These error components are generally correlated [BJR94]. Therefore, the error is modeled using a standard ARIMA model as,

$$\phi(B)\nabla^d \alpha_t = \theta(B)\upsilon_t \tag{2.13}$$

Multiplying these two models together, a general model used to parameterize a seasonal time series is obtained. This model is given by the expression,

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_s^D z_t = \theta_q(B)\Theta_Q(B^s)\upsilon_t \tag{2.14}$$

where $\phi(B)$ and $\theta(B)$ are polynomials in B of degrees p and q which satisfy the invertibility and stationarity conditions. A process with more periodic components can be obtained using similar methods [End95]. A more detailed discussion about the Seasonal ARIMA models can be found in [BJR94] and [End95].

2.3 Constraints on Model Parameters

For a time series to be successfully modeled using any of these parametric models, the parameters must satisfy certain conditions. The two most important constraints imposed on these parameters are the stationarity and invertibility constraints. The stationarity constraint applies only to parameters in the AR model. It ensures that an AR Process, i.e. a stationary stochastic process modeled using an AR model, is stable. Similarly, the invertibility constraint applies only to an MA process. When the AR and MA processes are combined to form the ARMA or ARIMA process, the stationarity and invertibility constraints still apply for the AR and MA parts of the new process respectively as per their original definitions. This section introduces these constraints and their importance to time series analysis. Another constraint, applicable to ARMA and ARIMA processes, called the common root test is also discussed in this section.

2.3.1 Stationarity Condition. To ensure that a time series can be successfully modeled using a linear stochastic model, it must be stationary [BJR94]. As discussed in a previous section, the condition that successfully ensures this stationarity is the convergence of the infinite sum series expressed in 2.1.

Consider the expression of a general AR process of order p given in equation 2.5 and rewrite it in the following form,

$$z_t = \Phi^{-1}(B)v_t$$

Comparing this result with the transfer function representation of the general linear stochastic model from 2.1

$$z_t = \Psi(B)v_t \tag{2.15}$$

where, $\Psi(B) = 1 + \psi_1 B + \psi_2 B^2 + \cdots$, it is evident that

$$\Psi(B) = \Phi^{-1}(B) \tag{2.16}$$

In section 2.2 it was stated that the infinite sum of the ψ-weights must converge to successfully model a stationary time series using a linear stochastic model. It follows, from equation 2.16, that $\Phi^{-1}(B)$ must also converge in order to model a stationary process using an AR model. [BJR94] shows that this condition, defined as the *stationarity condition*, is met if the roots of $\Phi(B) = 0$ lie outside the unit circle. In other words, the zeros of the transfer function $\Phi(B)$ must lie outside the unit circle to achieve a stationary AR process. This result implies that the linear system being modeled as an AR process will be stable and that an unstable system cannot be stationary.

The stationarity condition does not apply to an MA process. By direct comparison between equation 2.7 and equation 2.15, it is seen that $\Theta(B) = \Psi(B)$. By definition, $\Theta(B)$ is a polynomial of finite order q and therefore is equivalent to a convergent sum of ψ-weights. Hence, the stationarity constraint does not apply to MA parameters. Even when a nonstationary time series is being modeled, the MA part of the ARIMA process satisfies the stationarity constraint. This is easily seen by rewriting a nonstationary ARIMA process described by equation 2.10 as,

$$\Phi(B) w_t = \Theta(B) v_t \tag{2.17}$$

where, $w_t = (1 - B)^d \tilde{z}_t$ is the differenced time series. The resulting expression is obtained without modifying the AR and MA parts of the ARIMA process and resembles an ARMA process in that it is a stationary stochastic process with elements of both the AR and MA processes.

2.3.2 Invertibility. A time series can be modeled as a linear stochastic process only if it is stationary and invertible. It has been shown previously that the stationarity condition applies only to an AR process. This section investigates the invertibility condition.

Consider the expression of a general MA(q) process, i.e. a Moving Average process of order q given in equation 2.7 and rewrite it in the following form,

$$\Theta^{-1}(B)z_t = v_t$$

Comparing this result with the transfer function representation of the general linear stochastic model from 2.2,

$$\Pi(B)z_t = v_t \tag{2.18}$$

where, $\Pi(B) = 1 + \pi_1 B + \pi_2 B^2 + \cdots$, it is evident that

$$\Theta^{-1}(B) = \Pi(B) \tag{2.19}$$

For the infinite sum series $\Pi(B)$ to converge, [BJR94] shows that roots of $\Theta(B) = 0$ must lie outside of the unit circle. In other words, the zeros of the transfer function $\Theta(B)$ must lie outside the unit circle to achieve an invertible MA process. This result, known as the *invertibility condition*, is derived by the same process as the stationarity condition.

The invertibility condition does not apply to an AR process since the coefficient polynomial $\Phi(B)$ is by definition of finite order. Furthermore, using the results obtained in equation 2.17, it can be shown that the invertibility condition applies to an ARIMA process as well.

2.3.3 Common Roots. It is necessary to check for common elements between the AR and MA parts of the ARMA and ARIMA models to negate the possibility of the modeling procedure to result in degenerate models. Consider the transfer function form of the ARMA model rewritten as,

$$z_t = \frac{\Theta(B)}{\Phi(B)} v_t \tag{2.20}$$

If the identified polynomials $\Theta(B)$ and $\Phi(B)$ have roots in common, it will lead to a *pole-zero cancellation* and effectively model the time series as a reduced order ARMA process. In practical applications, a near cancellation between poles and zeros may also lead to a misidentification of a process. Such a misidentification can result in situations where combinations of parameter values yield similar v_t's and similar likelihood values. Potential difficulties arising from such a misidentification are discussed in more detail in [BJR94]. The same caveat applies for ARIMA models and its variants.

Now that the possible models that can be used to parameterize a time series have been defined and the constraints on the parameters of these models identified, the next step is to understand the process to identify the unknown variables of the model in a way that facilitates forecasting of time series.

CHAPTER 3
MODEL IDENTIFICATION

A complete model identification process includes identifying the optimal parameters that fit the model to the time series, quantifying any error in the variables being estimated and then assembling the results in a form that can then be used for estimation or forecasting. Various algorithms exist to estimate the parameters of a given model and quantify the estimation error such that the time series is modeled optimally.

The most common algorithms for identifying the model parameters are Least-Squares Estimate (LSE) and Maximum Likelihood Estimate (MLE). Out of the two, LSE is the most frequently implemented estimator due to ease of implementation. However, [FPW98] show that the LSE is an asymptotically biased and inconsistent estimator in a stochastic estimation problem involving ARMA models. Out of the different alternatives available, the MLE provides the most efficient solution to the parameter estimation of stochastic models [Kay93].

An added complication due to the use of an ARMA model is the lack of a closed form expression to quantify the parameter estimate error. In high integrity applications such as autonomous shipboard landing of aircraft, it is important to account for all contributing sources of error while staying within the error budget. Several algorithms that are capable of estimating the parameter estimate error covariance such as the Outer Product of Gradients (OPG), Hessian Estimate and the Sandwich Estimator exist.

For an ARMA model, the set of unknowns that need to be estimated are the weight coefficients θ_i's and ϕ_j's of AR and MA processes and the input WGN variance

σ^2. These unknowns form the parameter vector $\boldsymbol{\theta} = [\phi_1, \cdots, \phi_p, \theta_1, \cdots, \theta_q, \sigma^2]$. This chapter introduces the MLE and OPG algorithms as methods to optimally identify the elements of the parameter vector and quantify the parameter estimation error covariance.

3.1 Maximum Likelihood Estimation

The Maximum Likelihood Estimate (MLE) is defined as the set of parameter values $\boldsymbol{\theta}$ that maximize the likelihood function. The MLE works with both deterministic and stochastic processes. In the case of a deterministic process, the MLE maximizes a known function. However, to parameterize a stochastic process, the MLE requires an input probability density function (pdf), known as the *likelihood function*, defining the distribution of the input noise. For the purpose of this work it is assumed that the input noise is normally distributed with zero mean and an unknown variance σ^2, i.e. it is defined as WGN. Therefore, the normal distribution is chosen as the likelihood function for the MLE to maximize in order to determine the set of optimal parameters defining the time series.

An ARMA parameter estimation problem cannot be solved analytically using the MLE [FPW98]. Therefore, a numerical optimization algorithm needs to be implemented. Newton-Raphson and Expectation Maximization are the two most commonly used methods to implement the MLE [Kay93]. The Expectation Maximization (EM) algorithm guarantees convergence to at least a local maximum, but not to a global maximum. It also needs to be customized for the problem at hand. The Newton-Raphson method, however, is more commonly used and easily implemented.

This section introduces the MLE algorithm and its implementation using the Newton-Raphson iteration as a method to optimally determine the parameter vector defining a stochastic time series. It also investigates various properties of the MLE.

3.1.1 Maximum Likelihood Estimate Algorithm.

Consider the inverse ARMA model,

$$v_t = z_t - \sum_{i=1}^{p} \phi_i z_{t-i} + \sum_{j=1}^{q} \theta_j v_{t-j} \tag{3.1}$$

where v_t's are random disturbances characterized as WGN with variance σ^2, z_t's are the measurements in a time series, and ϕ's and θ's are the parameters defining the ARMA process. The process can be expressed in an inverse form as shown in Eq 3.1 due to the imposed stationarity and invertibility conditions. For a time history of the disturbances $\boldsymbol{V} = [v_n, v_{n-1}, v_{n-2}, \cdots, v_1]$ where n is the time index, [BJR94] shows that the exact likelihood function is,

$$p(\boldsymbol{V}|\boldsymbol{\theta}) = \prod_{t=1}^{n} (2\pi\sigma_t^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2} \sum_{t=1}^{n} \frac{(v_t - \hat{v}_{t|t-1})^2}{\sigma_t^2} \right] \tag{3.2}$$

Here, $\hat{v}_{t|t-1}$ is the predicted value of the disturbance at time t given the value at time $t-1$. The specifics of forecasting methods used to obtain the predicted value are discussed in chapter 4 of this work.

Taking the natural logarithm of Eq. 3.2 gives us

$$l(\boldsymbol{V}|\boldsymbol{\theta}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n} \log(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{n} \frac{(v_t - \hat{v}_{t|t-1})^2}{\sigma_t^2} \tag{3.3}$$

where the function $l(\boldsymbol{V}|\boldsymbol{\theta})$ is known as the log-likelihood function and $\boldsymbol{\theta}$ is the parameter vector consisting of the AR, MA and variance parameters ϕ's, θ's and σ respectively.

The advantage of using the MLE with the ARMA model is that the likelihood function has a unique maximum. Fig 3.1 plots the log-likelihood function for an ARMA(1,1) process. For a likelihood function maximized using a Newton-Raphson method, if the algorithm converges, it will always obtain the global maximum.

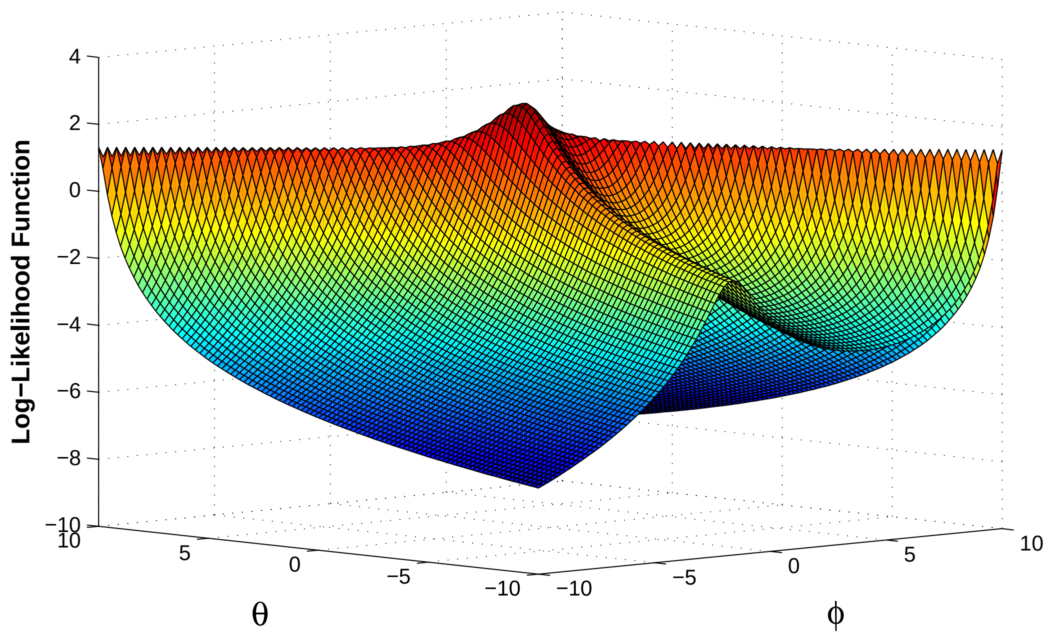

Figure 3.1. Log-Likelihood function for an ARMA(1,1) process

The Newton-Raphson iteration for the MLE is given by the expression

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - \left[\frac{\partial^2 l(\boldsymbol{V}|\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^T}\right]^{-1} \left.\frac{\partial l(\boldsymbol{V}|\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\boldsymbol{\theta}_k} \tag{3.4}$$

As with all implementations of the Newton-Raphson method, it is important to provide a good initial estimate of the parameters to start the process. This can be confirmed by observation of Fig. 3.1. It is evident that if the intial guess is far enough away from the maximum, the Newton-Raphson method will cause the algorithm to follow either of the two ridges that diverge to infinity. A Least Squares Estimate (LSE) may be used to obtain an initial guess needed for the Newton-Raphson method [FPW98].

Fig 3.3 shows results of a simulation where a single time series of length $N = 5000$ was parameterized multiple times using the Newton-Raphson method given a grid of initial guesses. The initial guesses are the set of parameters that satisfied the stationarity, invertibility and common root constraints. For this simulation, Fig. 3.2

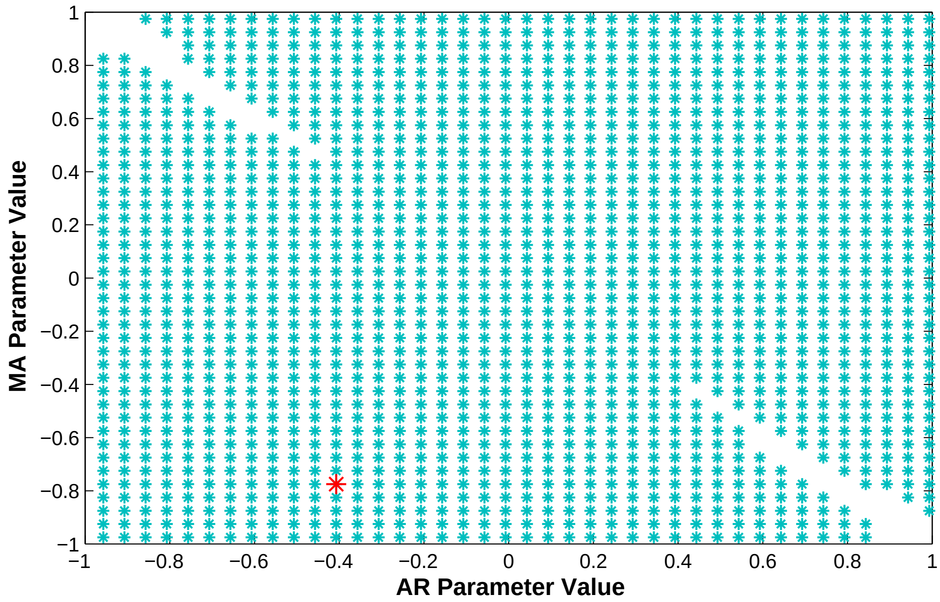

Figure 3.2. Grid of initial guesses that satisfy the constraints

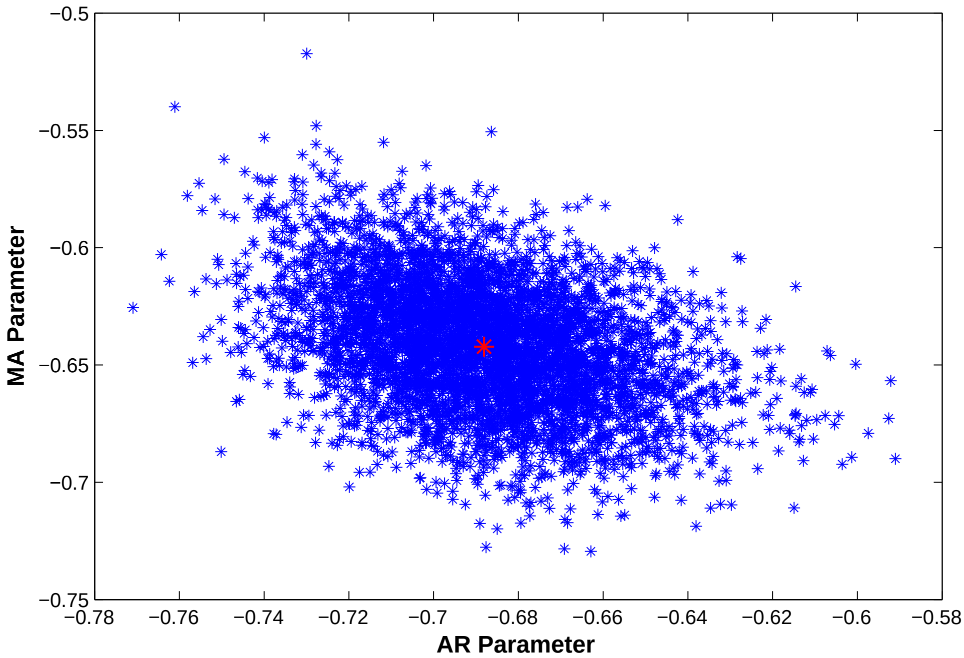

Figure 3.3. Convergence of MLE for valid initial guess

shows the initial guesses for the MLE algorithm that satisfy the constraints. The time series was generated using an ARMA(1,1) model parameterized as $\phi_1 = -0.25$ and $\theta_1 = -0.58$. The values of the true parameters are shown in red in both figures 3.2 and 3.3. In figure [**?**], convergent values for the initial guesses are shown in blue. It is clearly seen that the Newton-Raphson iteration converges in the neighborhood of the true parameters for all valid initial guesses. Therefore, regardless of the method used to obtain an initial guess, as long as the initial guess satisfies the parameter constraints, the MLE will converge successfully. Note that for a simple ARMA(1,1) process, a brute force method may be implemented to obtain parameter values by searching for points that satisfy the stationarity and invertibility constraints to initialize the Newton-Raphson iteration with satisfactory performance. However, for higher order systems, due to additional computational complexity it is necessary to implement more robust methods such as LSE to obtain the initial guesses.

3.1.2 Properties of Maximum Likelihood Estimate. The MLE also has a set of properties that make it a desirable option compared to other estimators such as LSE and BLUE. The most important properties are that the MLE is a consistent, unbiased, asymptotically efficient, asymptotically normal and invariant estimator [Kay93]. However, these properties hold true only if certain regularity conditions are satisified.

The regularity conditions, as stated in [Gre03], that need to be satisfied for the MLE to have the desirable properties are,

1. The first three derivatives of $l(\boldsymbol{V}|\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ are continuous and finite for almost all v_t's and for all $\boldsymbol{\theta}$. This condition ensures that a Taylor series approximation exists and the variance of the derivatives of $l(\boldsymbol{V}|\boldsymbol{\theta})$ is finite.
2. The conditions necessary to obtain the expected values of first and second derivatives of the log-likelihood function are met.

3. $\partial^3 l(\boldsymbol{V}|\boldsymbol{\theta})/(\partial\theta_j\,\partial\theta_k\,\partial\theta_l)$ is less than a function that has a finite expectation for all values of $\boldsymbol{\theta}$. This condition allows the truncation of the Taylor series.

There are also three properties of the probability density functions that need to be introduced prior to discussing the properties of MLE [Gre03].

1. The log-likelihood function $l(v_t|\boldsymbol{\theta})$, its first derivative $\boldsymbol{g}_t = \partial\, l(v_t|\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ and second derivative $\boldsymbol{H}_t = \partial^2\, l(v_t|\boldsymbol{\theta})/\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T$ for $t = 1, \cdots, n$ are all random samples of random variables.
2. $E_0\left[\boldsymbol{g}_t(\boldsymbol{\theta}_0)\right] = \mathbf{0}$.
3. $Var\left[\boldsymbol{g}_t(\boldsymbol{\theta}_0)\right] = -E\left[\boldsymbol{H}_t(\boldsymbol{\theta}_0)\right]$. This equality is also known as the *information matrix equality.*

These properties of the probability density function also need to be satisfied for the MLE properties to hold true.

3.1.2.1 Consistency. Now that the regularity conditions have been defined, the properties of the maximum likelihood estimator can be assessed. The first important attribute of the MLE is that it is an asymptotic *consistent estimator*. An estimator is termed as consistent if $\text{plim}\,\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_0$ where plim indicates a probability limit. This equality says that as the length of available data increases, the parameter estimate $\hat{\boldsymbol{\theta}}$ converges in probability to their true values $\boldsymbol{\theta}$. A rigorous proof of this property is provided in [Gre03].

3.1.2.2 Unbiased. Another important property that differentiates the MLE from other estimators like LSE is that it is asymptotically unbiased. An estimator is unbiased when the expected value of an estimator is equal to the true value of the parameters. For applications involving ARMA models, the LSE may result in biased parameter estimates [BJR94]. Therefore, this property makes the MLE a useful subsitute as an estimator. A simple way of showing this property is to observe the parameter estimation error for a simple process. If an estimator is unbiased, then the parameter estimate error must have an expected value of zero, i.e. it must be zero mean.

To verify this property, a total of 5000 ARMA(1,1) time series of length 1000 were generated using the same set of parameters and were then reparameterized using MLE. The values provided to initialize the MLE algorithm satisfied the stationarity, invertibility and common root conditions. The parameter estimate error was then evaluated. Figure 3.4 plots the parameter estimate error for both AR and MA parameters in blue and their mean in red. The estimate error for the AR parameter

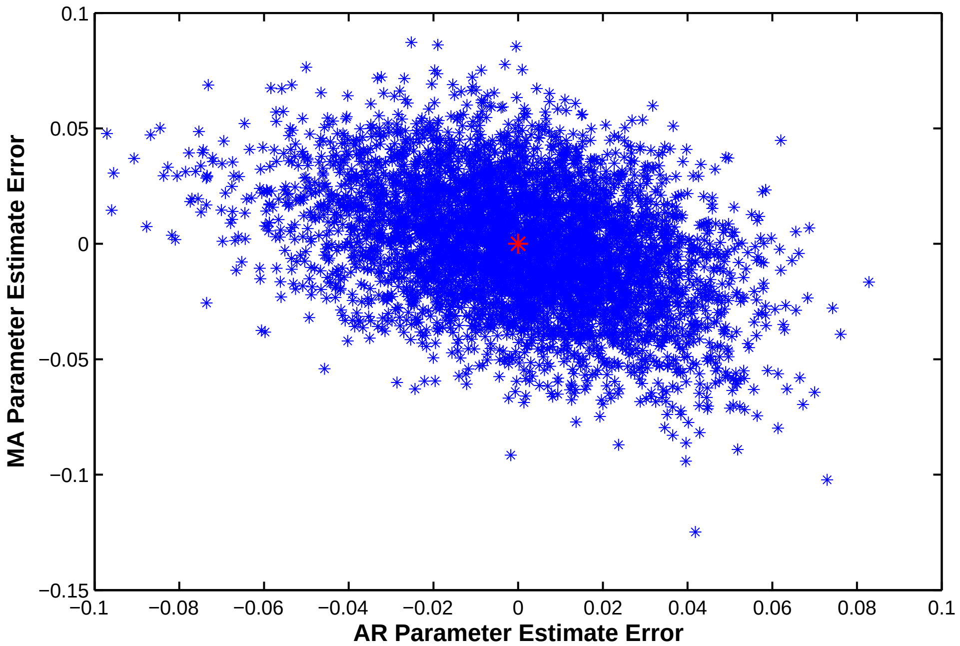

Figure 3.4. Unbiased Maximum Likelihood Estimate Simulation

has a mean of $E[\phi - \hat{\phi}] = 4.17 \cdot 10^{-4} \approx 0$ and the MA parameter estimate error has a mean $E[\theta - \hat{\theta}] = -6.39 \cdot 10^{-4} \approx 0$. This result, which can be extended to higher order ARMA processes, verifies the property that a maximum likelihood estimate is unbiased. A detailed proof of this property is provided in [Kay93].

3.1.2.3 Efficiency. For a sufficiently long time history of data, the parameter estimate error obtained from MLE achieves the Cramer-Rao Lower Bound (CRLB) [Kay93]. The CRLB bounds the minimum value of the variance an estimator can achieve. An estimator, such as the MLE, which achieves the CRLB is known as an efficient estimator. The CRLB increases as the number of parameters being estimated increases. Assuming that the log-likelihood function satisfies the regularity conditions, the covariance of the estimates of the true parameter vector $\boldsymbol{\theta}_0$ will always be at least as large as,

$$[\boldsymbol{I}(\boldsymbol{\theta}_0)]^{-1} = \left(E_0 \left[\left(\frac{\partial l(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}_0} \right) \left(\frac{\partial l(\boldsymbol{\theta}_0)}{\partial \boldsymbol{\theta}_0} \right)^T \right] \right)^{-1} \tag{3.5}$$

Since the MLE is unbiased and achieves the CRLB, it is known as a minimum variance unbiased (MVU) estimator.

3.1.2.4 Asymptotic Normality. The parameter estimate errors obtained from the MLE are asymptotically normal [Gre03]. As the length of the time series increases, the distribution of the parameter estimates will approach a Gaussian distribution with the following properties,

$$\hat{\boldsymbol{\theta}} \overset{a}{\sim} N \left[\boldsymbol{\theta}_0, \{\boldsymbol{I}(\boldsymbol{\theta}_0)\}^{-1} \right] \tag{3.6}$$

where $\boldsymbol{I}(\boldsymbol{\theta}_0)$ is called the information matrix and is defined as,

$$\boldsymbol{I}(\boldsymbol{\theta}_0) = -E \left[\frac{\partial^2 l}{\partial \boldsymbol{\theta}_0 \, \partial \boldsymbol{\theta}_0^T} \right] \tag{3.7}$$

A verification study for this property was conducted by generating a set of 1000 time series using a common set of parameters for an ARMA(1,1) process. These time series were then reparameterized assuming that the true parameters are unknown. The parameter estimation error, i.e. the deviation from truth of the estimated parameter values were then compared with an ideal Gaussian distribution. Figure 3.5 shows the histogram of the AR (i.e. ϕ_1) and MA (i.e. θ_1) parameter estimate error.

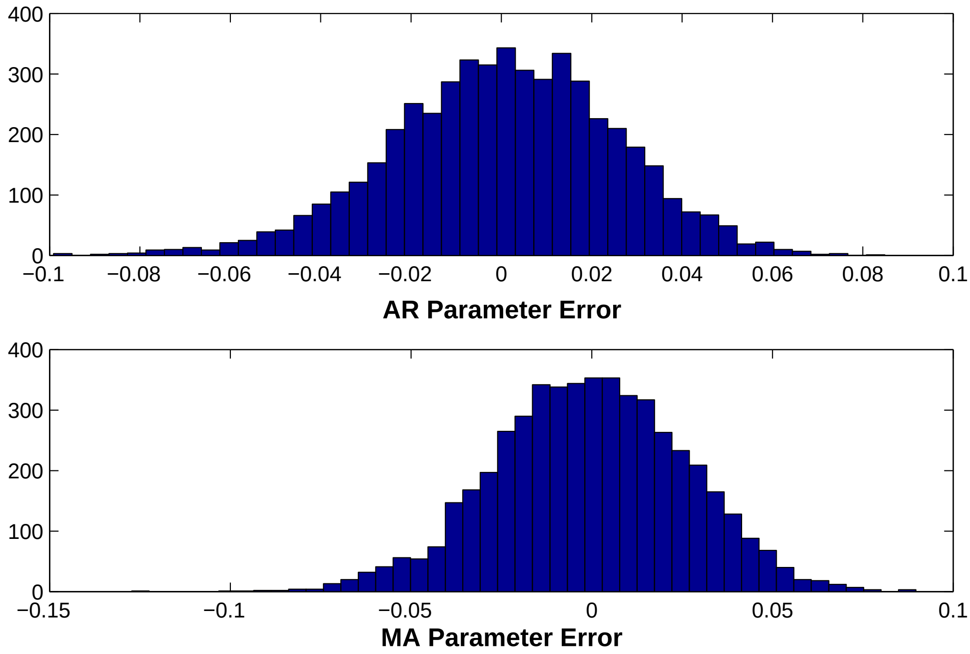

Figure 3.5. AR and MA parameter estimate error histogram

The Normal distribution of the estimate error can be seen easily from these plots. However, this result can be analyzed more efficiently via a quantile plot as shown in Fig 3.6. The quantile plot is a semi-log scale plot that compares the the probability of the values contained in a data set with a Normal distribution generated using its mean and variance. In figure 3.6, the ideal Normal distribution is depicted by the red line and the data, which is in blue, is overlaid on the plot. The results from this simulation show that the parameter estimate error is almost perfectly Gaussian for both AR and MA estimates thus verifying the MLE property of asymptotic normality. There is slight non-Gaussian behavior towards the extremities of the data set, however that is beyond 3-sigma or 99.7% probability limits.

3.1.2.5 Invariance. The last important property of the MLE is that it is invariant to one-to-one transformations of $\boldsymbol{\theta}$ [Gre03]. This property is not a statistical result, but one that increases computational efficiency of determining the MLE. It can be

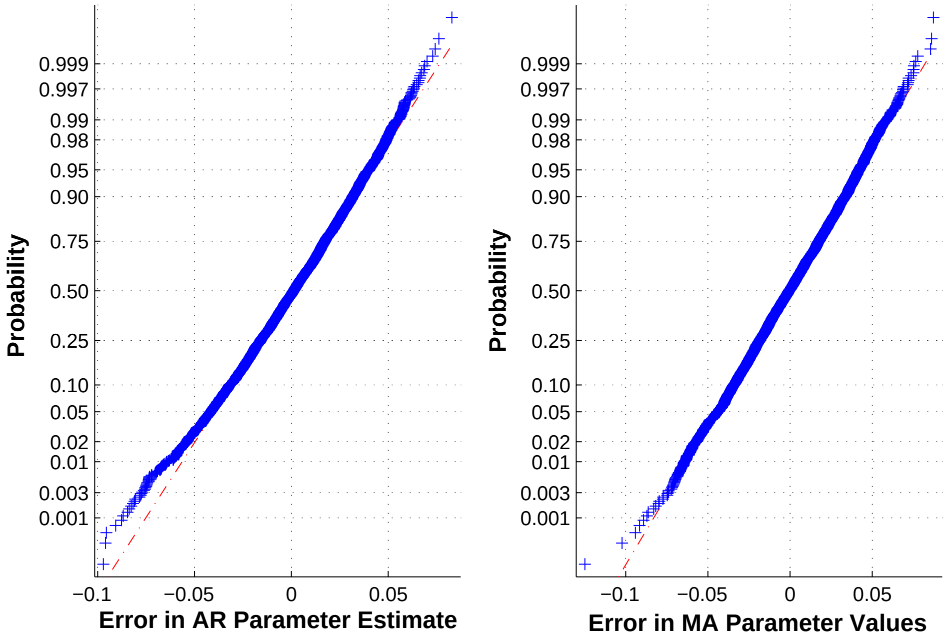

Figure 3.6. Verifying the Gaussian distribution of the parameter estimation error

used to great effect in simplifying the mathematical formulation of an expression to be used for parameter estimation. For example, if a parameter appears in a likelihood function in the form $1/\theta$, the model can be reparameterized in terms of $\eta = 1/\theta$. This may help with simplifying the expressions of the first and second derivatives of the likelihood function thereby increasing computational efficiency. Examples of this property and its impact on the computing of MLE are provided in [Gre03] and [Kay93].

3.2 Estimation Error: Outer Product of Gradients

Once the model parameters have been identified, it is important to quantify any possible errors in the model identification process. Assuming that the model order has been identified correctly, the only other source of error is the parameter estimation error. The asymptotic covariance matrix of the MLE is a matrix of parameters that are being estimated and is represented by the information matrix expression given in equation 3.7, i.e. it is the expected value of the second derivatives matrix. The second derivatives of the log-likelihood function are almost always complicated functions whose exact expected values will be unknown [Gre03]. Therefore, the estimate error covariance needs to be estimated.

There are three different methods available to estimate the covariance matrix. These methods (see [NM94]) involve evaluating second derivatives of the log-likelihood function with respect to the estimated parameter vector. These methods are computationally expensive and therefore not the ideal approaches for real-time applications. However, a method developed in [BHHH74] formulates the expression for the covariance matrix as the outer product of the gradient vector. This estimator, commonly known as BHHH estimator or Outer Product of Gradients (OPG) is given by,

$$\left[\hat{\boldsymbol{I}}(\hat{\boldsymbol{\theta}})\right]^{-1} = \left[\sum_{t=1}^{n}\left(\left\{\frac{\partial l(v_t|\hat{\boldsymbol{\theta}})}{\partial\hat{\boldsymbol{\theta}}}\right\}\left\{\frac{\partial l(v_t|\hat{\boldsymbol{\theta}})}{\partial\hat{\boldsymbol{\theta}}}\right\}^T\right)\right]^{-1} \tag{3.8}$$

The OPG is a consistent estimator of the asymptotic variance provided some regularity conditions (see [NM94]) are satisfied. For the ARMA model parameter identification, the OPG also provides an unbiased estimate of the parameter error variance. This property of the OPG estimator can be easily verified using a Monte-Carlo simulation. A total of 5000 generated time series were parametrized using an ARMA model and the parameter estimate error variance was quantified using the OPG. The variance of the error obtained from the Monte-Carlo simulation was assumed to be the true value and was compared with the estimates provided by the OPG. Figure 3.7 plots the variance estimated by the OPG for each time series against the variance obtained from the Monte-Carlo simulation. It is clearly seen that the Monte-Carlo variance is the mean of the estimates obtained from the OPG estimator.

Obtaining the parameter estimate error covariance matrix is an important part of quantifying the contribution of the model identification process to the total error

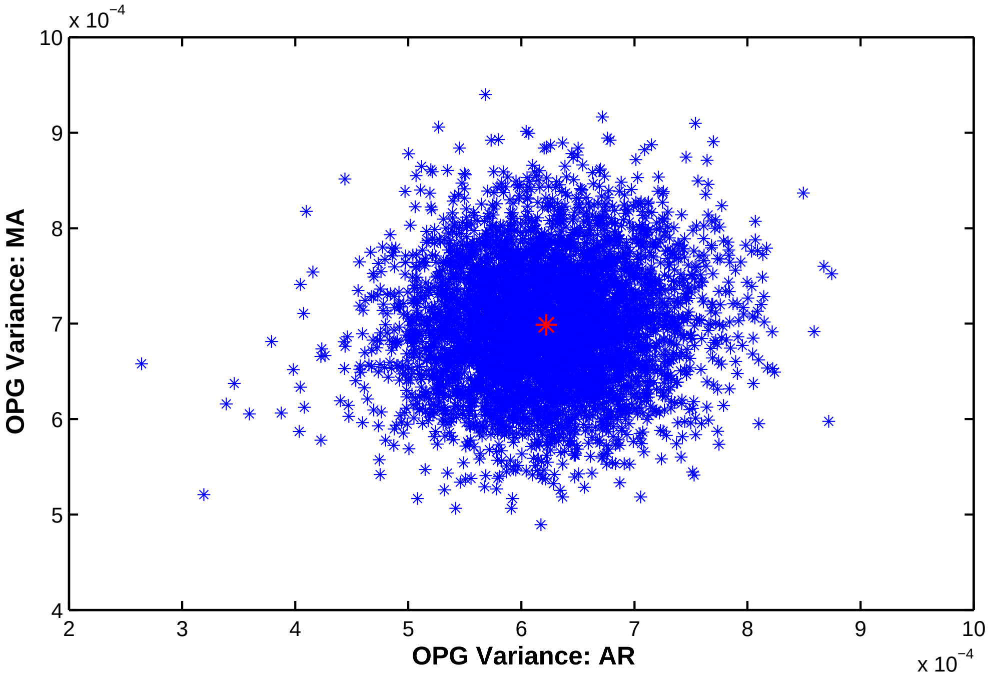

Figure 3.7. Outer Product of Gradients variance estimate

in ship state propagation. However, these OPG estimates also provide other benefits. [BJR94] shows that errors in model identification are reflected in the OPG estimates of the estimate error covariance. For example, in case of a pole-zero cancellation due to existence of common roots, the OPG estimate of parameter error covariance for those roots will tend to infinity. A detailed discussion of this and other applications of the OPG estimate is provided in [BJR94] and [NM94].

CHAPTER 4
FORECASTING

Once the time series has been parameterized using the MLE and the estimate error covariance is obtained using the OPG, this information can be used to propagate the ship state. Propagating the ship state for shipboard landing of aircraft requires forecasting the time series for position and orientation of the ship and quantifying the forecast error covariance. An ARIMA process modeled as,

$$\varphi(B)z_t = \theta(B)v_t \tag{4.1}$$

where $\varphi(B) = \phi(B)\nabla^d$, has three explicit forms that can be used for forecasting the time series. These forms provide different insights into the forecasting process. [BJR94] derives the minimum mean squared error (MMSE) forecast of an ARIMA process and expressions to quantify the forecast error at a certain prediction time. This chapter discusses the different forms of the ARIMA model that facilitate forecasting in order to obtain an MMSE forecast and quantify the error covariance on these forecasts. Since this work focuses on propagating state under communication outage conditions, updating the forecast functions upon arrival of new measurements is not considered.

4.1 MMSE Forecast

Consider a forecast $\hat{z}_{t+l}$ of the future measurement z_{t+l} made at time t. This forecast can be written using the general linear stochastic model where a value in a time series is a linear function of current and previous random disturbances. The expression for the forecast at lead time l originating at time t is

$$\hat{z}_{t+l} = \psi_l^* v_t + \psi_{l+1}^* v_{t-1} + \cdots$$

where the weights $\psi_l^*, \psi_{l+1}^*, \ldots$ are to be determined. [BJR94] and [Wol38] both show that the mean squared error of the forecast at time l is,

$$E\left[z_{t+l} - \hat{z}_{t+l}\right] = \left(1 + \psi_1^2 + \cdots + \psi_{l-1}^2\right)\sigma^2 + \sum_{j=0}^{\infty}\left(\psi_{l+j} - \psi_{l+j}^*\right)^2\sigma^2 \tag{4.2}$$

where the error is minimized by setting $\psi_{l+j}^* = \psi_{l+j}$. Therefore,

$$\begin{aligned} z_{t+l} &= \left(v_{t+l} + \psi_1 v_{t+l-1} + \cdots + \psi_{l-1} v_{t+1}\right) + \left(\psi_l v_t + \psi_{l+1} v_{t-1} + \cdots\right) \\ &= e_{t+l} + \hat{z}_{t+l} \end{aligned} \tag{4.3}$$

where e_{t+l} is the error of the forecast $\hat{z}_{t+l}$ at the time l made from time t.

4.1.1 Facts about MMSE Forecast. [BJR94] lists a few important facts which emerge from the derivation. However, prior to discussing these facts, certain notation and assumptions must be established. $E\left[z_{t+l}|z_t, z_{t-1}, \ldots\right]$, the conditional expectation of z_{t+l} given that all measurements z up to time t are known, is denoted by $E_t\left[z_{t+l}\right]$ and the random disturbances v_t are assumed to be a sequence of independent random variables.

4.1.1.1 Fact One. Based on the assumptions and equations, $E\left[v_{t+j}|z_t, z_{t-1}, \ldots\right] = 0$ for $j > 0$. Therefore,

$$\hat{z}_{t+l} = \psi_l v_t + \psi_{l+1} v_{t-1} + \cdots = E_t\left[z_{t+l}\right] \tag{4.4}$$

Therefore, the MMSE forecast for lead time l made at time t is the conditional expectation of the measurement z_{t+l} at time t. Here, $\hat{z}_{t+l}$ is known as the *forecast function* for time t with lead time l. Note that for the MMSE forecast to coincide with the MMSE linear forecast, the random disturbances v_t must satisfy the condition $E\left[v_{t+j}\right] = 0$ for $j > 0$.

4.1.1.2 Fact Two. Given a forecast error,

$$e_{t+l} = v_{t+l} + \psi_1 v_{t+l-1} + \cdots + \psi_{l-1} v_{t+1} \tag{4.5}$$

it is evident that

$$E_t\left[e_{t+l}\right] = 0 \tag{4.6}$$

Therefore, the forecast for lead time l at time t is unbiased.

4.1.1.3 Fact Three. [BJR94] shows that not only is $\hat{z}_{t+l}$ the MMSE forecast of z_{t+l}, but that a linear function of the forecasts is also the MMSE forecast of the corresponding linear function of measurements. That is, $\sum_{l=1}^{L} w_l \hat{z}_{t+l}$ is the MMSE forecast of $\sum_{l=1}^{L} w_l z_{t+l}$.

4.1.1.4 Fact Four. The random disturbances v_t's which until now were defined only as WGN input, are also the forecast errors for a prediction of one step. Consider the expression

$$e_{t+1} = z_{t+1} - \hat{z}_{t+1} \tag{4.7}$$

Assuming that the one step forecast $\hat{z}_{t+1}$ is exact, it is evident from direct substitution for expressions of z_{t+1} and $\hat{z}_{t+1}$, that

$$e_{t+1} = v_{t+1} \tag{4.8}$$

It follows, that for an MMSE forecast, the one step prediction errors must be uncorrelated since v_t's are defined as WGN[BJR94]. This is a sensible result, because if the errors were correlated, the forecast errors could be predicted from available forecast errors and a better forecast would be $\hat{z}_{t+1} + \hat{v}_{t+1}$.

4.1.1.5 Fact Five. Forecast errors for lead times much longer than one time step as well as forecast errors for the same lead time from different origins will be correlated. [BJR94] derives expressions for these correlations. This correlation may result in the forecast function evaluating either completely above or below the true future values of the time series. That is, the forecast values may all be above or below the true values that may be measured in the future.

4.2 Time Series Forecasting

The ARIMA model given in Eq 4.1 has three explicit forms that can be used for forecasting the ARIMA process. These forms are the difference equation form, the ψ-weights form and the π-weights form. The difference equation form is the most convenient form of the ARIMA model to obtain the forecasts. However, the other two forms provide different insights into the forecasting process as well and therefore are worth exploring. This section looks into evaluating the MMSE forecast of an ARIMA process using the different forms of the ARIMA model.

4.2.1 Forecasts using ARIMA Difference Equation form. A measurement z_{t+l} at lead time l for an ARIMA process can be expressed as

$$z_{t+l} = \varphi_1 \, z_{t+l-1} + \cdots + \varphi_{p+d} \, z_{t+l-p-d} - \theta_1 v_{t+l-1} - \cdots - \theta_q v_{t+l-q} + v_{t+l} \tag{4.9}$$

where, p, q and d are the orders of the AR, MA and differencing respectively. The expression for the predicted value of z_{t+l}, using the MMSE forecast derived by [BJR94], is

$$\begin{aligned} \hat{z}_{t+l} = \varphi_1 \left[z_{t+l-1}\right] \quad &+ \cdots + \varphi_{p+d} \left[z_{t+l-p-d}\right] - \theta_1 \left[v_{t+l-1}\right] - \cdots \\ &- \theta_q \left[v_{t+l-q}\right] + \left[v_{t+l}\right] \end{aligned} \tag{4.10}$$

where $[v_{t+l}] = E_t\,[v_{t+l}]$ and $[z_{t+l}] = E_t\,[z_{t+l}]$.

Assuming that each term can be evaluated, it is easy to see that forecasting using the difference equation is a more computationally efficient method compared to forecasts obtained using ψ-weights form and π-weights form. The specifics of how to evaluate the various term in the difference equation is discussed in section 4.2.4.

4.2.2 Forecasts using ARIMA ψ-Weights form. Although the difference equation form is the most efficient form for forecasting, an ARIMA process modeled using the ψ-weights form can not only be used for time series prediction, but also to enable quantification of the forecast error covariance. A measurement at a future time of an ARIMA process can be expressed in the form of an infinite weighted sum of current and previous random disturbances v_k's, similar to the general linear stochastic model shown in equation 2.1. The expression for the measurement z_{t+l} at lead time l is,

$$z_{t+l} = \sum_{k=0}^{\infty} \psi_k \, v_{t+l-k} \tag{4.11}$$

where $\psi_0 = 1$ and the remaining ψ-weights are obtained by equating coefficients of

$$\varphi(B)\,(1 + \psi_1 B + \psi_2 B^2 + \cdots) = \theta(B)$$

Then, the forecasting verion of this expression that calculates the predicted value of the time series at the lead time is,

$$\hat{z}_{t+l} = [v_{t+l}] + \psi_1\,[v_{t+l-1}] + \cdots + \psi_{l-1}\,[v_{t+1}] + \psi_l\,[v_t] + \psi_{l+1}\,[v_{t-1}] + \cdots \tag{4.12}$$

which is the same form as obtained from the derivation of the MMSE forecast.

Recall that the forecast error e_{t+l} is,

$$\begin{aligned} e_{t+l} &= \hat{z}_{t+l} - z_{t+l} \\ &= v_{t+l} + \psi_1 v_{t+l-1} + \cdots + \psi_{l-1} v_{t+1} \end{aligned} \tag{4.13}$$

Therefore, the variance of the forecast error is,

$$\begin{aligned} V(l) &= E_t\left[e_{t+l}^2\right] \\ &= \left(1 + \psi_1^2 + \psi_2^2 + \cdots + \psi_{l-1}^2\right)\sigma^2 \end{aligned} \tag{4.14}$$

where, σ^2 is the variance of the input WGN.

4.2.3 Forecasts using ARIMA π-Weights form.

The MMSE forecast at a lead time l originating at time t using the π-weights form, similar to the general model shown in equation 2.2, is

$$\hat{z}_{t+l} = \sum_{k=1}^{\infty} \pi_k\,[z_{t+l-k}] + [v_{t+l}] \tag{4.15}$$

Although the π weights form an infinite sum, the invertibility constraint imposed on the ARIMA model ensures that these weights form a convergent series. That is, the forecast depends finitely on past measurements of the time series. In practice, the π weights decay quickly and therefore only a moderate length time series is required to evaluate the forecasts [BJR94]. Furthermore, this method simplifies the implementation of the forecasting process because it depends for the most part on past values of the time series, where the expected value is the measurement itself, i.e $[z_t] = z_t$.

However, depending on the number of π weights needed to model the process, this method may not be as computationally efficient as the difference equation form, nor is it as intuitive.

4.2.4 Evaluating Terms in Forecast Equation. In order to implement these forecasting methods, it is necessary to calculate the the conditional expectations. Given that j is a nonnegative integer, the conditional expectations are evaluated as

$$\begin{aligned}
&[z_{t-j}] = z_{t-j} && j = 0, 1, 2, \ldots \\
&[z_{t+j}] = \hat{z}_{t+j} && j = 1, 2, \ldots \\
&[v_{t-j}] = v_{t-j} = z_{t-j} - \hat{z}_{t-j-1}(1) && j = 0, 1, 2, \ldots \\
&[v_{t+j}] = 0 && j = 1, 2, \ldots
\end{aligned} \tag{4.16}$$

where the square brackets imply the conditional expectation of that variable given past measurements at time t and $\hat{z}_{t-j-1}(1)$ is the one-step-ahead forecast, i.e. the forecasted time series value at time $t - j$ given the value at time $t - j - 1$.

Therefore, given the expressions in Eq. 4.16, the forecast $\hat{z}_{t+l}$ is obtained according to the following rules [BJR94]:

1. The z_{t-j} for $j = 0, 1, 2, \ldots$, which have already happened at time t are left unchanged.

2. The z_{t+j} for $j = 1, 2, \ldots$, which have not yet happened, are replaced by their forecasts $\hat{z}_{t+j}$.
3. The v_{t-j} for $j = 0, 1, 2, \ldots$, which have happened, are calculated from $z_{t-j} - \hat{z}_{t-j-1}(1)$.
4. The v_{t+j} for $j = 1, 2, \ldots$, which have not yet happened, are set to zero.

For a calculation using the difference equation form, the forecasts are calculated by

$$\hat{z}_{t+l} = \sum_{j=1}^{p+d} \varphi_j \hat{z}_{t+l-j} - \sum_{j=1}^{q} \theta_j v_{t+l-j} \tag{4.17}$$

where the moving average terms v_{t+l-j}'s are not present for lead times $l > q$. The evaluation of these disturbance terms is initialized by setting the unknown v's to their unconditional expected values of zero [BJR94]. Assuming that data contained in a time series is available starting at time index 1, the v_s's are computed recursively from the difference equation form of the ARIMA model as,

$$\begin{aligned} v_s &= z_s - \hat{z}_{s-1}(1) \\ &= z_s - \left(\sum_{j=1}^{p+d} \varphi_j z_{s-j} - \sum_{j=1}^{q} \theta_j v_{s-j} \right) \qquad s = p+d+1, \ldots, t \end{aligned} \tag{4.18}$$

where for the start time $s = p + d + 1$, the v_{s-j} terms are set to zero.

4.2.5 Example ARIMA(1,1,1).

Consider a stable ARIMA(1,1,1) process written in its transfer function form as,

$$(1 - 0.5B)(1 - B)z_t = (1 - 0.1B)v_t$$

By expanding the above equation, it can be expressed in its difference equation form,

$$z_t = 1.5z_{t-1} - 0.5z_{t-2} - 0.1v_{t-1} + v_t$$

Then, the forecast equation originating at time t for a lead time $l = 1$ is given by,

$$\hat{z}_{t+1} = 1.5z_t - 0.5z_{t-1} - 0.1v_t \tag{4.19}$$

where the terms $v_t = z_t - \hat{z}_{t-1}(1)$ as shown in Eq 4.18 and $v_{t+1} = 0$ because $[v_{t+1}] = 0$.

Similarly, forecast equations for higher lead times until the forecast process stabilizes (i.e. the v terms run out) are given by,

$$\begin{aligned}
\hat{z}_{t+2} &= 1.5\hat{z}_{t+1} - 0.5z_t \\
\hat{z}_{t+3} &= 1.5\hat{z}_{t+2} - 0.5\hat{z}_{t+1} \\
&\vdots \\
\hat{z}_{t+l} &= 1.5\hat{z}_{t+l-1} - 0.5\hat{z}_{t+l-2}
\end{aligned} \tag{4.20}$$

Notice that the direct effect of random disturbances lasts only until measured values are present. In this case, at forecast time $t + 2$, the forecast equation no longer has a v_t term present since the lead time is greater than the MA order q. The forecast from that time depends only on prior measurements and past forecasts. After times $t+p+d$, i.e. at times greater than $t+2$, the forecasts are a function of past forecasts only since there are no more measured values remaining in the time series.

In order to evaluate the accuracy of these forecasts, the forecast error covariance must be quantified. Equation 4.14 shows that the probability limits of the forecast depend on the ψ-weights. Therefore the ψ weights for the corresponding forecast times need to be calculated. These terms are calculated equating the coefficents of the $\varphi(B)$, $\psi(B)$, and $\theta(B)$ polynomials.

$$\varphi(B)\psi(B) = \theta(B) \tag{4.21}$$

Equating the coefficients of powers of B in

$$\begin{aligned}
(1 - \varphi_1 B - \cdots - \varphi_{p+d}B^{p+d})(1 + \psi_1 B + \psi_2 B^2 + \cdots) \\
= (1 - \theta_1 B - \cdots - \theta_q B^q)
\end{aligned} \tag{4.22}$$

provides a structure for evaluating the ψ's. The expressions for the ψ-weights after equating the coefficients are,

$$
\begin{aligned}
\psi_1 &= \varphi_1 - \theta_1 \\
\psi_2 &= \varphi_1\psi_1 + \varphi_2 - \theta_2 \\
&\ \vdots \\
\psi_j &= \varphi_1\psi_{j=1} + \cdots + \varphi_{p+d}\psi_{j-p-d} - \theta_j
\end{aligned}
\tag{4.23}
$$

where $\psi_0 = 1$, $\psi_j = 0$ for $j < 0$ and $\theta_j = 0$ for $j > q$. For a time index j such that $j > \max(p + d - 1\,,\, q)$, the ψ-weights satisfy the difference equation

$$\psi_j = \varphi_1\psi_{j-1} + \varphi_2\psi_{j-2} + \cdots + \varphi_{p+d}\psi_{j-p-d} \tag{4.24}$$

For this example, the ψ's are obtained by comparing coefficients of

$$(1 - 1.5B + 0.5B^2)(1 + \psi_1 B + \psi_2 B^2 + \cdots) = (1 - 0.1B)$$

Therefore, the ψ-weights evaluate to

$$
\begin{aligned}
\psi_0 &= 1 \\
\psi_1 &= 1.4 \\
\psi_2 &= 1.5\psi_1 - 0.5 \\
\psi_j &= 1.5\psi_{j-1} - 0.5\psi_{j-2} \qquad j = 3, 4, 5, \ldots
\end{aligned}
\tag{4.25}
$$

For each lead time l, the weights are calculated for $j \leq l - 1$ as seen from equation 4.14 which provides the expression for the forecast variance.

CHAPTER 5
RESULTS AND DISCUSSION

Time series analysis and forecasting methods discussed in the previous chapters were tested on sea trial data for an aircraft carrier provided by the N-UCAS program. The ship trajectory for the sea trial as expressed in a local North-East-Down (NED) body frame is shown in figure 5.1. The start of the ship trajectory is reflected in

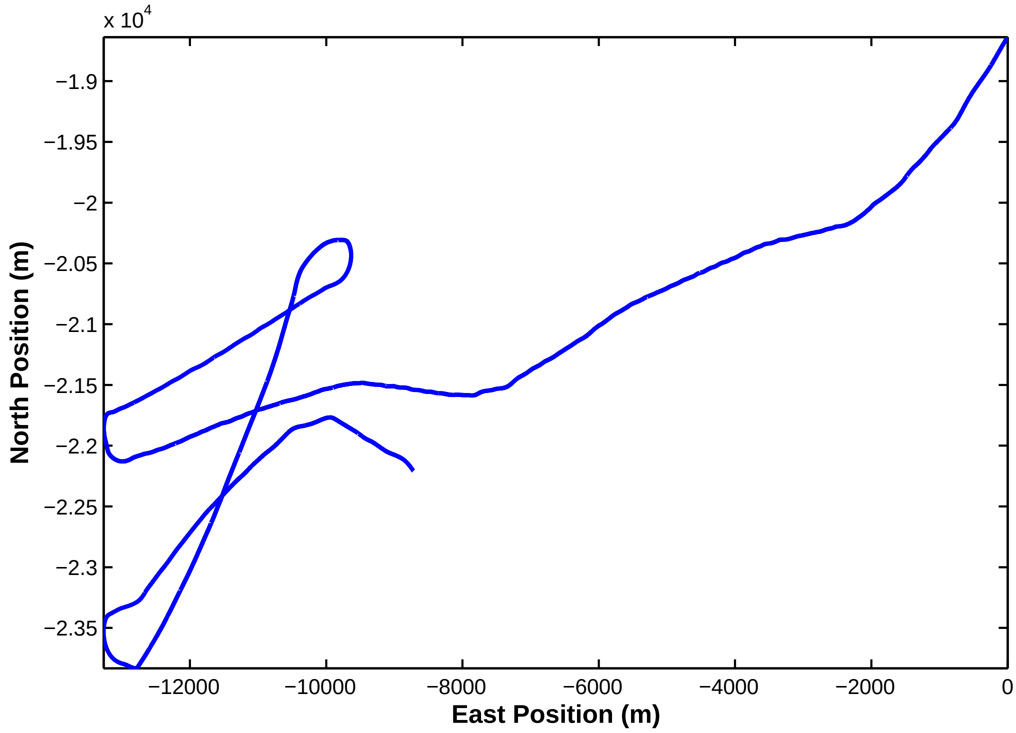

Figure 5.1. Ship trajectory expressed in the NED reference frame

the top right corner of the figure. In the second half of the sea trial, it can be seen that the ship executes maneuvers involving turns at least 4 times. The sections of data chosen for analysis must resemble the ship motion during an aircraft landing approach. As per the assumptions stated in the introduction, an aircraft will land on a ship only when it is not performing any maneuvers. Therefore, any sections of the ship trajectory where the ship is not performing turning maneuvers can be used for analysis. Fig 5.2 identifies, in the color red, sections of the ship trajectory from available sea trial data that are representative of ship motion for an actual UAV landing.

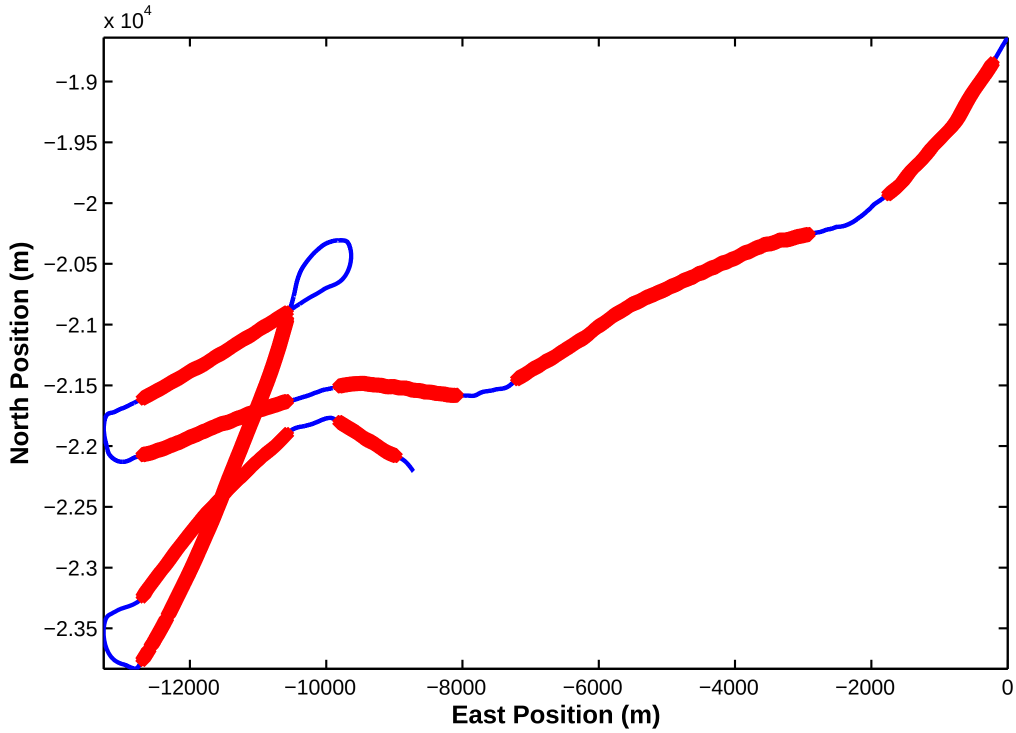

Figure 5.2. Sections of ship trajectory that satisfy the landing approach conditions

The analysis presented in this chapter chooses a point from the highlighted region in the figure as a point of data link outage to start ship motion forecasting. The ship state being propagated consists of ship position in a local North-East-Down (NED) reference frame. Since the available data does not include the ship orientation, i.e. the Euler angles, the analysis presented in this chapter does not include modeling and forecasting of the ship orientation. Similarly, velocity and acceleration of the ship are of secondary importance to position forecasting and therefore are not included in the analysis.

[BJR94] suggests that a time series consisting of at least 50 data points is typically necessary for a faithful model identification meaning that the ARIMA parameter estimate errors will be negligible. Therefore, sections of data with length 200 are used for analysis and the results are presented. The analysis presented is done for individual time series only. The time series for north, east and down components of the ship position are shown in figures 5.3, 5.4, 5.5.

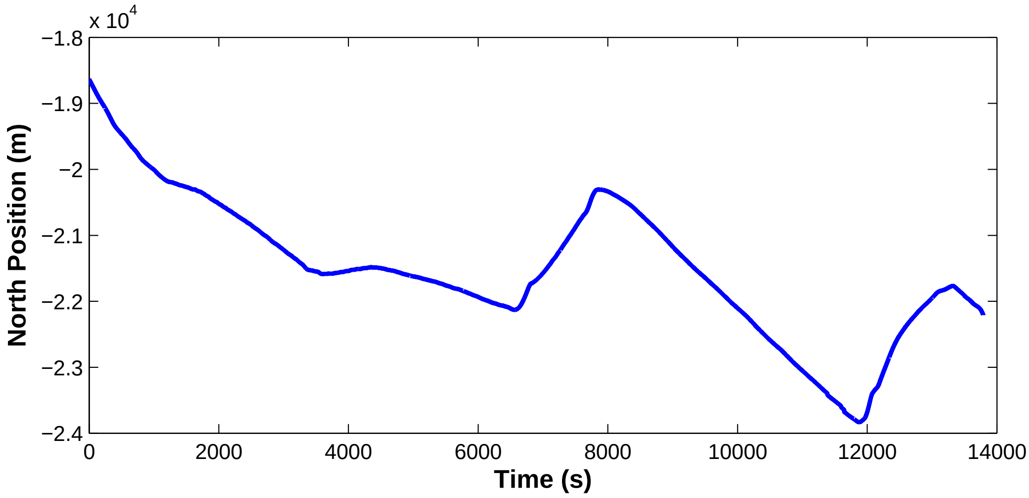

Figure 5.3. North component of ship trajectory

The analysis presented in this chapter focuses on model identification, forecasting the data for a lead time of 15 seconds and quantifying the forecast error.

The model identification process is sensitive to the chosen order for the AR, MA and degree of differencing. As discussed in the previous chapters, an initial estimate of the model order necessary to best describe the time series can be obtained by observing trends in the data. Figures 5.3 and 5.4 show the time series corresponding to the north and east components of the ship trajectory for the sea trial. It can be easily seen that the data is affected more by past measurements than by past random disturbances. Furthermore, the time series are nonstationary in both level and slope

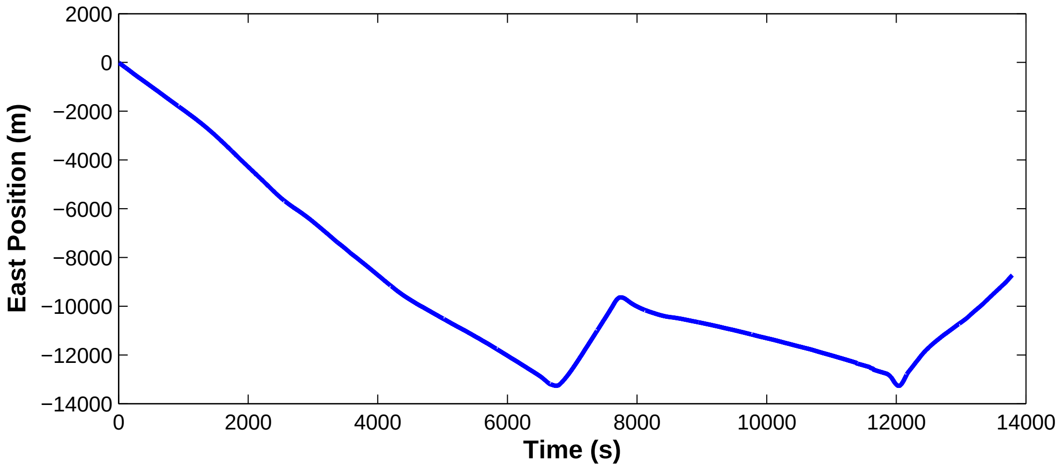

Figure 5.4. East component of ship trajectory

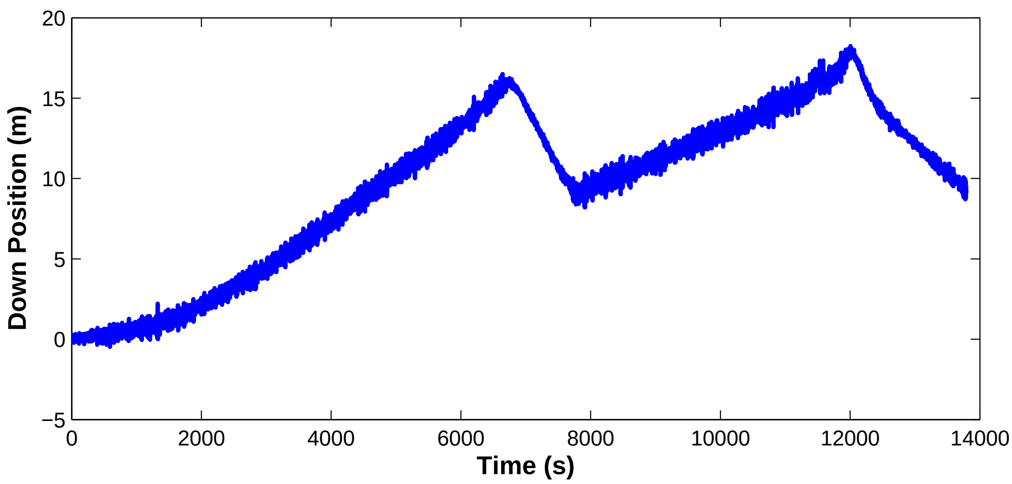

Figure 5.5. Down component of ship trajectory

which can be characterized by a degree of differencing $d = 2$. Therefore, the model is expected to be of the form ARIMA(p,2,q). Since the data does not seem to display any effects due to random disturbances, the order of the MA process is most likely small or equal to zero.

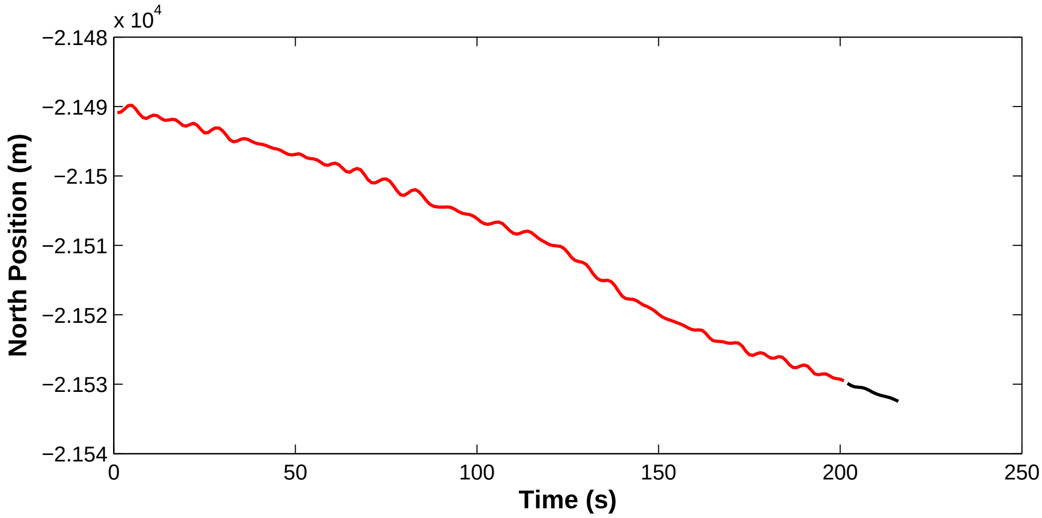

Figure 5.6. Forecasted values for north component of ship position

Figure 5.6 shows the mean square error (MSE) forecast for the north component of ship position based on a model identified using a 200 second section of data. The red time series is the data used for model identification and the black represents forecasted values. The accuracy of this prediction is better evaluated by observing the error between forecast values and the values measured at times corresponding to the predictions. The error displayed in figure 5.7 is depicted in black and the red dotted line represents the 1-sigma envelope evaluated from the forecasted mean square error variance.

In this case, the error does not exceed 0.5 meters for the forecasting time interval and the 1-sigma envelope completely bounds the error. This result is a significant improvement over the capabilities of the kinematic model discussed in the introduction. The identified ARIMA model is of the order ARIMA(4,2,1) and the

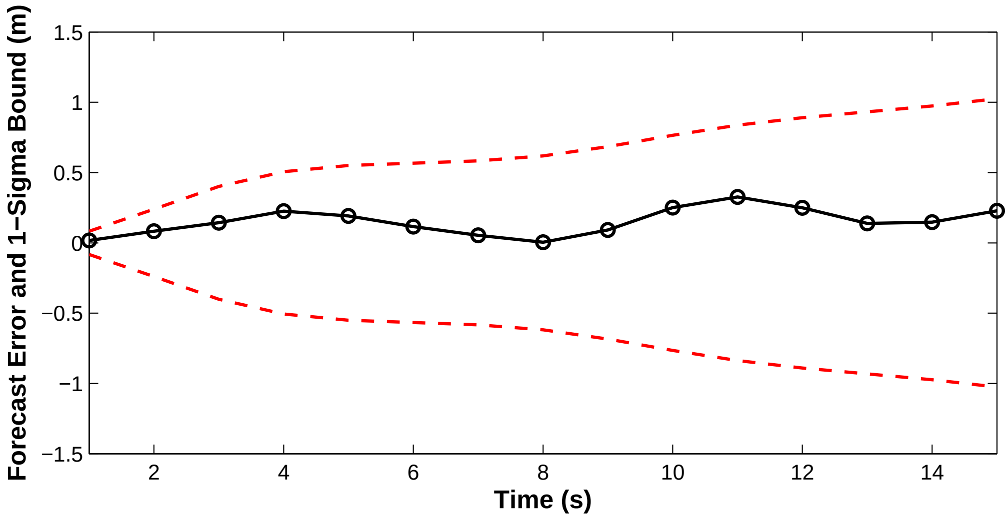

Figure 5.7. Forecast position error with 1-Sigma error bound for well identified model

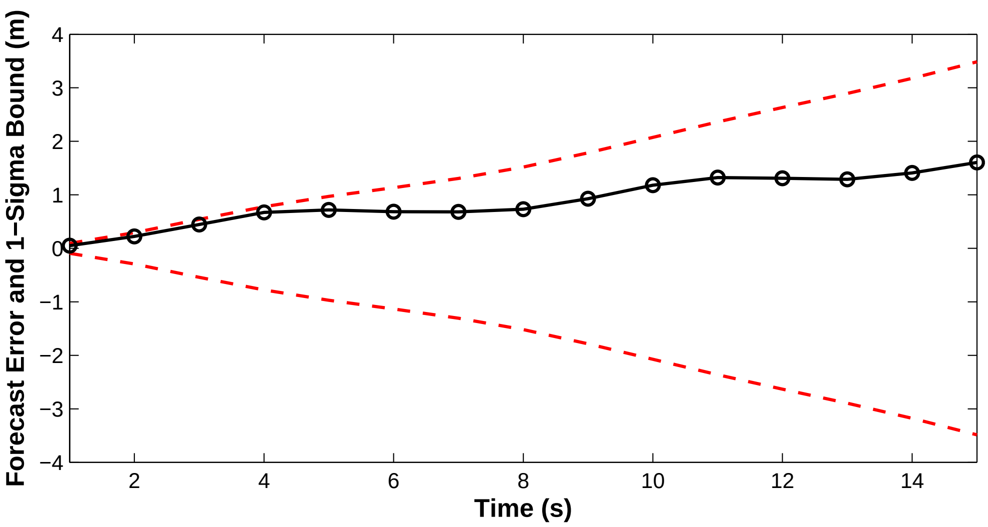

Figure 5.8. Forecast position error with 1-Sigma error bound for misidentified model

parameter values and parameter estimation errors computed using OPG are shown in the table 5.1. Figure 5.8 shows the results for the same set of data modeled using an ARIMA(4,2,0) process results. For this scenario, the forecast error at the end of 15 seconds lead time is approaching 2 meters, a significantly higher number than the 0.5 meter error obtained from the ARIMA(4,2,1) model. However, note that this case is one of the few exceptions in a trend seen in modeling the ship position data from this particular sea trial since an MA term was needed to accurately parameterize the time series.

Table 5.1. ARIMA(4,2,1) Model Parameter Values and OPG Error Estimates

Parameter	Value	OPG Error Estimate
AR(1)	1.47133	0.265
AR(2)	-1.5	0.351
AR(3)	0.5657	0.351
AR(4)	-0.2157	0.273
MA(1)	-0.8557	0.206
Input WGN Variance	0.00708	0.00080

A further detailed sensitivity analysis highlights the importance of a faithful model order identification prior to estimating the parameters. Figure 5.9 shows the results of the sensitivity analysis for different degrees of differencing d. The first subplot in the figure shows the actual predictions for a 15 second lead time for increasing values of d. The set of data in black is the last 50 seconds of a 200 second section of data used to identify the model parameters. In the second subplot, the error associated with each forecast is displayed. It is clear that as the degree of differencing increases, the forecast quality drastically decreases. For a degree of differencing $d = 2$, the forecast error is the smallest. This verifies the expected result that since the modeled time series is nonstationary in level and slope, the best result will be obtained by setting $d = 2$.

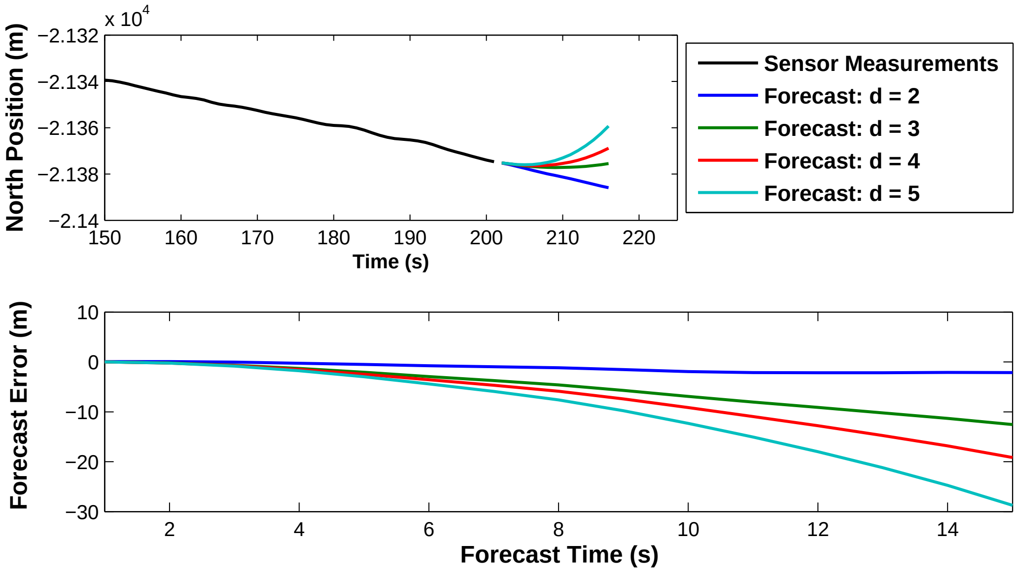

Figure 5.9. Sensitivity analysis for degree of differencing

As with the degree of differencing, the forecast process is sensitive to the order of AR parameters as well. Figure 5.10 shows the forecast (in the first subplot) and forecast error (in the second subplot) as the order of AR parameters is changed. As with the degree of differencing, it can be seen that for this specific set of ship data, the forecast quality degrades as the number of AR parameters increases. The best forecast is obtained by setting the order of the AR model to $p = 2$. Similarly, figure 5.11 shows the same AR parameter sensitivity analysis for a different section of the ship data. As opposed to the previous result, it can be seen that as the number of AR parameters increases, the forecast quality increases. Therefore, a high order AR model choice may not always result in the optimal forecast. This emphasizes the importance of accurate model order identification in order to obtain the best forecast. The model order for the results presented in this thesis was obtained manually. However, for a real time implementation, an autonomous model order identification method needs to be developed.

For high integrity applications such as shipboard landing of UAVs, it is nec-

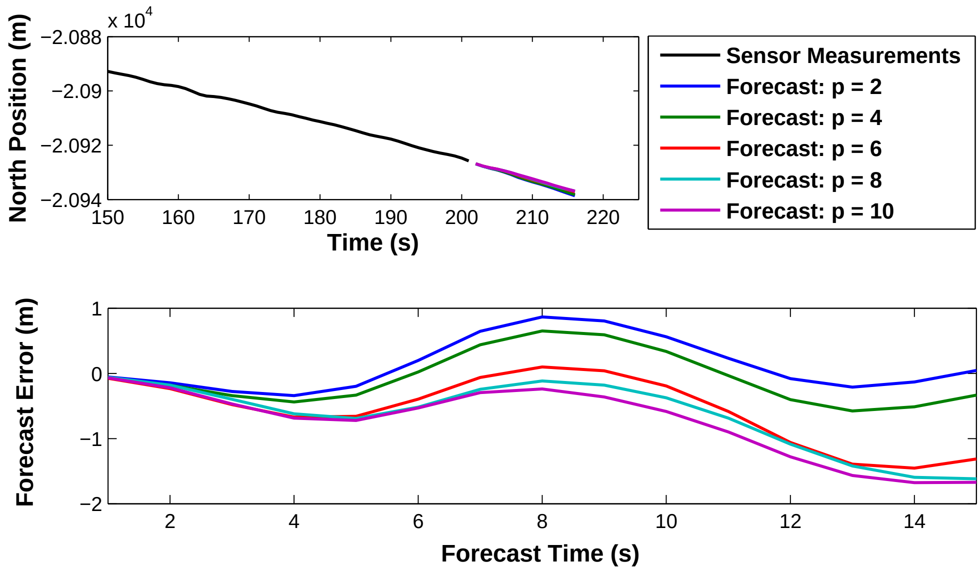

Figure 5.10. First sensitivity analysis for AR model order

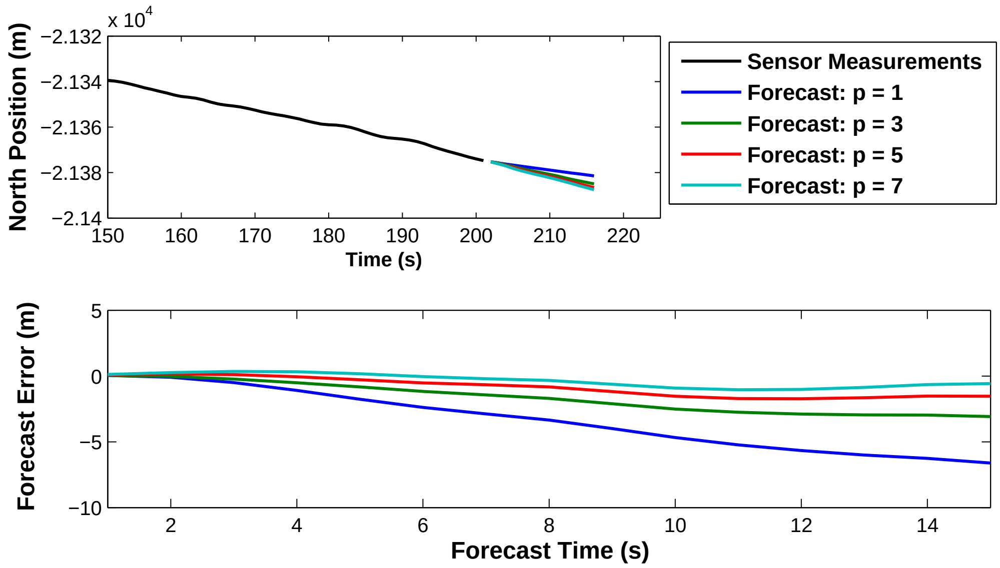

Figure 5.11. Second sensitivity analysis for AR model order

essary for the forecasting algorithm to be capable of quantifying the forecast error covariance faithfully. As discussed in the previous chapter, the ψ-weights calculated recursively at each lead time for the identified ARIMA model can be used to quantify the forecast error variance. The forecast error variance calculated for successive lead times using equation 4.14 is especially interesting, because it is not only capable of accounting for forecast errors but is also capable of accounting for any errors in the forecasting that may result from a misidentified model order.

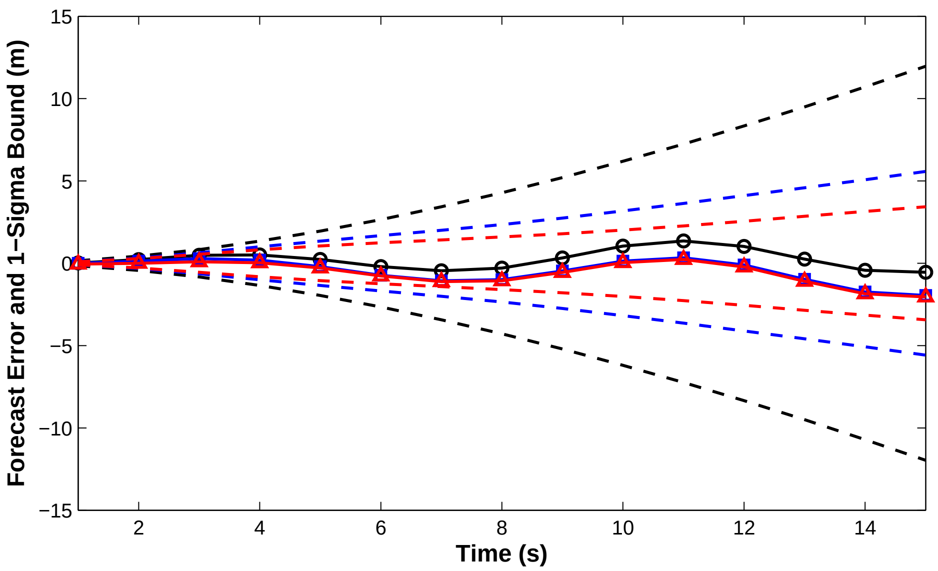

Figure 5.12. Forecast error variance for misidentified models [Black: ARIMA(1,2,0), Blue: ARIMA(3,2,0), Red: ARIMA(5,2,0)]

Figures 5.12 and 5.13 shows the forecast errors and 1-sigma forecast error variance bounds. In the two plots the solid lines represent forecast error and the dashed lines are the corresponding 1-sigma error envelopes. Model order identification with increasing number of parameters was carried out for the same section of the sea trial data. As the magnitude of error increases, the 1-sigma error envelope corresponding to each set of errors increases as well. This result shows that the equation for the mean squared error variance is capable of quantifying the variance even when the

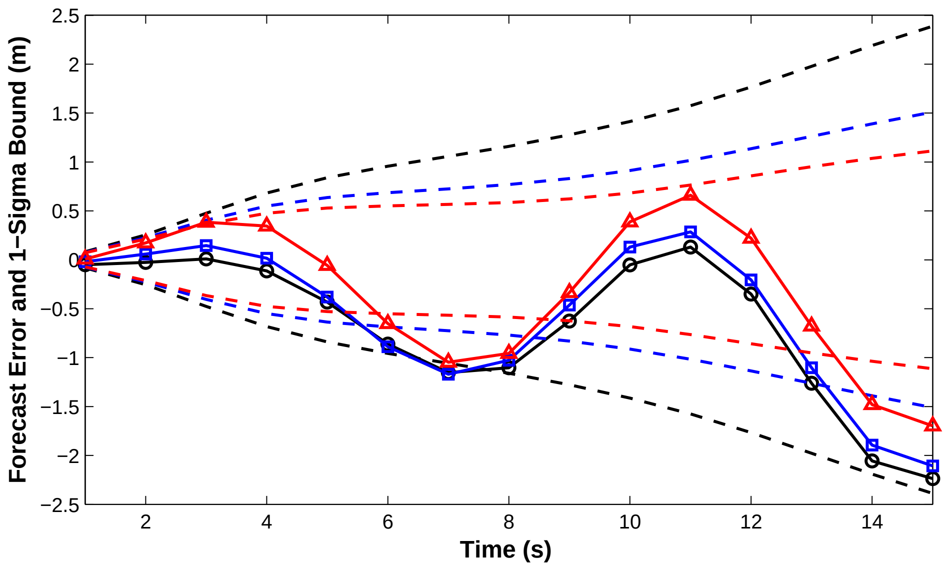

Figure 5.13. Forecast error variance for misidentified models [Black: ARIMA(7,2,0), Blue: ARIMA(9,2,0), Red: ARIMA(11,2,0)]

model order used to parameterize the data is erroneous. It makes intuitive sense that the calculated mean squared error variance is accurately modeling the forecast error, because the variance at each lead time depends on the the ψ-weights being calculated recursively using the ϕ and θ parameters of an identified ARIMA model. If the order of the ARIMA model order is misidentified, the change in the number of ϕ and θ parameters will be reflected in the ψ-weight calculations.

Another method to confirm the ability of equation 4.14 to quantify the forecast error variance is by comparing it to a sample variance of actual forecast errors. As seen from the prior sensitivity analysis, it is important to choose the optimal model order to obtain best forecasts since the model order can change significantly depending on the section of the ship sea trial data chosen for analysis. Therefore, to perform such an analysis, a method to autonomously determine the model order is required. Methods presented in the literature (see [BJR94], [Gre03] and [End95]) require a human-in-the-loop approach where the identified model order needs to be verified manually. Due to these constraints an exhaustive analysis could not be completed.

However, figure 5.14 shows results for 15 second lead time forecasts of 30 sections of the north component of ship position. The plot displays forecast errors normalized using the predicted 1-sigma envelopes for each forecast. The dashed black line shows the normalized 1-sigma bound and the solid colored lines are the normalized errors for each forecast. At lead time $t = 1$, 9 out of 30 total data points are outside

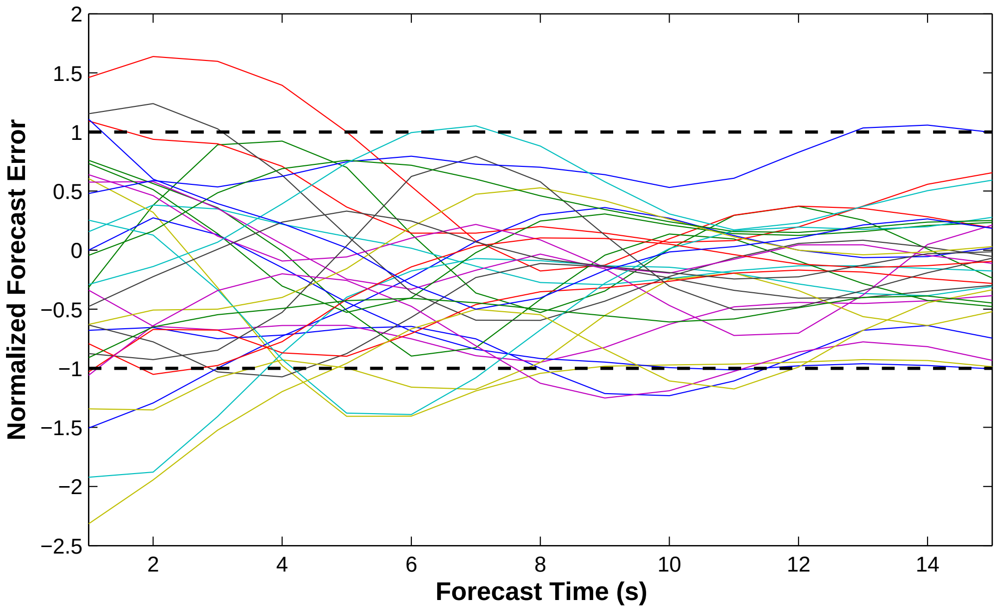

Figure 5.14. Normalized errors for 30 forecast simulations with 1-Sigma limits

the 1-sigma limits. This corresponds to 70% of the values being inside the limits, i.e. approximately equal to a standard 1-sigma limit. This preliminary result is encouraging since it shows that the 1-sigma estimate for these time series seems to be accurate. At lead time $t = 15$ only 1 out of 30 total data points remains outside the 1-sigma limits indicating that the 1-sigma limit conservatively bounds the error at longer lead times.

These results obtained from the model order identification, parameter estimation and forecasting analysis on the ship sea trial data shows that parameterizing the ship motion by ARIMA models and then forecasting the motion using the identified model is a valid approach for shipboard landing of UAVs under communication outage conditions. The results indicate that the ship position data can be accurately modeled as an ARIMA process and can also be forecasted easily as well. This method of time series analysis is also capable of faithfully quantifying the forecast error variance, although some more work needs to be done to produce a comprehensive result.

CHAPTER 6
CONCLUSION AND FUTURE WORK

This thesis focuses on developing methods to sucessfully land an Unmanned Aerial Vehicle on an aircraft carrier under communication outage conditions. Since the aircraft carrier broadcasts parameters describing the vessel motion to the aircraft for its navigation algorithm, if the data link that transmits this information fails, the aircraft needs to be able to autonomously predict ship motion until it lands safely. The two most important constraints are the error limit and the integrity risk limit for the entire navigation algorithm. The navigation algorithm has a vertical error limit of 1.8 meters for a UAV landing approach and an integrity risk requirement of $I < 10^{-7}$.

The initial focus of this thesis is on implementing standard methods such as kinematic modeling and dynamic modeling to propagate the ship state for a lead time of 15 seconds. However, these methods are shown to be infeasible. For the kinematic modeling approach, the most robust solution is obtained by assuming constant acceleration. The method utilizes the last set of position, velocity and acceleration data transmitted by the ship to the aircraft. Unfortunately, this method can result in large errors in the propagated state as shown in figure 1.1. Therefore kinematic modeling is not an ideal approach for this problem. The dynamic modeling approach was considered as an alternative method to propagate the ship state. This method uses parameters such as the ship mass, inertia and a force model to accurately represent the motion of a ship. This approach is limited in its implementation since the ship does not broadcast these parameters via the data link. More importantly, the forces affecting the ship cannot be modeled accurately due to lack of accurate models for describing the sea states and the aircraft cannot measure the sea state, therefore making it difficult to choose the correct sea state model. For these reasons the dynamic modeling approach for ship state propagation is also an infeasible method for ship state propagation.

As an alternative to the prior two methods, this work proposes a completely empirical method to predicting the ship state. The past data transmitted to the UAV by the aircraft carrier can be parameterized. The resulting model can then be used to predict the ship state. This method of parameterizing a set of data with the idea of forecasting it is known as Time Series Analysis and Forecasting. This thesis investigates the use of time series modeling for high integrity and high accuracy applications. Specifically, it focuses on the ability of the modeling methods to faithfully parameterize the time series, quantify the parameter estimation error, forecast the time series and quantify the forecast error.

The maximum likelihood estimation algorithm (MLE) is presented as the best method to optimally parameterize the time series using an Autoregressive Integrated Moving Average (ARIMA) model. The outer product of gradients (OPG) is implemented to quantify the parameter estimate error covariance. Theory for the ARIMA model is also presented by identifying the different forms of the model that facilitate modeling of a time series and forecasting it faithfully. The theory of MLE, OPG and ARIMA developed in the first few chapters has also been implemented on sea trial data provided by the N-UCAS program. The results from this analysis indicate that parametric modeling of time series using ARIMA models is a valid approach for forecasting the time series. Furthermore, preliminary analysis of the method to quantify forecast error variance shows that the forecast error seems to be quantified well.

6.1 Future Work

For a faithful model identification and forecasting process in a real time algorithm implementation, optimal identification of the model order, i.e orders of the degree of differencing d, number of AR paramters p and number of MA parameters q, is important. The algorithm must also be capable of determining these orders autonomously. One method to determine the order is discussed in [BJR94]. This approach involves interpreting trends in the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the time series. This method can potentially be automated since the criterion by which the model order is determined is evaluated using a mathematical expression that depends on an expression based on the ACF and PACF values.

Once the algorithm to autonomously determine the optimal model order for a set of data is developed, the forecast error variance expression needs to be comprehensively verified. As discussed earlier, this may be done by direct comparison with a sample variance obtained from the entire ship data set. Based on the theory discussed in pervious chapters, the forecast error is modeled faithfully if at any individual lead time the errors in all forecast traces is normally distributed. The intial results with 30 sample forecasts are encouraging.

The down component of the ship motion shown in figure 5.5 displays multiple periodic trends as well as nonstationarity in level and slope. For example, a look at a 400 second section of the data is shown in figure 6.1. A seasonal ARIMA model can be used to model data with a single periodic trend. However, data with multiple periods such as this is modeled using multiplicative seasonal ARIMA models. A more rigorous understanding of these models and how they are used to model multiple dominant periods in a data set needs to be developed prior to implementation in ship state modeling and forecasting. A brief discussion on multiplicative seasonal ARIMA models is provided in [BJR94].

The north (see figure 5.3) and east (see figure 5.4) components of the ship

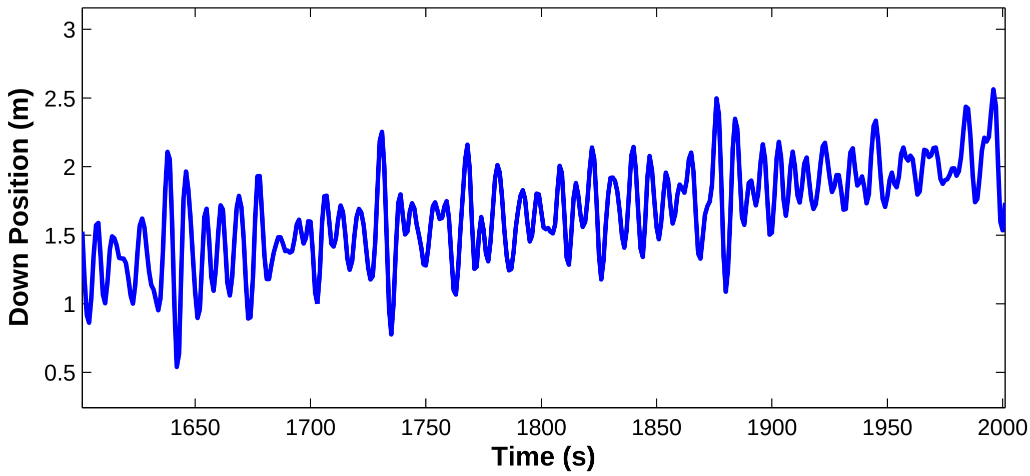

Figure 6.1. Periodic trends in down component of ship position

trajectory can be directly modeled using an ARIMA model of a suitable order. Based on the descriptions of time series behavior discussed in previous chapters, it can be observed that both time series are nonstationary in level and slope. Therefore, they may be modeled by using an ARIMA model with $d = 2$ plus a certain set of AR and MA parameters. It is interesting to note that these ARIMA(p,2,q) processes together represent not only the landing approach sections of the sea trial data but also the sections of data representing ship maneuvers such as turns. Therefore, these models have the potential of successfully parameterizing all sections of the ship trajectory. This has many potential benefits, with the primary benefit being that the aircraft would be able to predict ship motion regardless of the complexity of the ship motion, i.e. it would be capable of predicting ship maneuvers such as turns as well.

BIBLIOGRAPHY

[BHHH74] E. Berndt, B. Hall, R. Hall, and J. Hausman. Estimation and inference in nonlinear structural models. *Annals of Economic and Social Measurement*, 3:103–116, 1974.

[BJR94] G. Box, G. Jenkins, and G. Reinsel. *Time Series Analysis: Forecasting and Control.* Prentice Hall, Englewood Cliffs, NJ, 1994.

[Bre59] C. Bretschneider. Wave variability and wave spectra for wind-generated gravity waves. *Tech. Memo. No. 113. Beach Erosion Board, US Army Corps of Engineers*, 1959.

[Dou91] B. Douglas. Global sea level rise. *Journal of Geophysical Research*, 96:6981–6992, 1991.

[ea73] K. Hasselmann et al. *Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP).* Deutches Hydrographisches Institut, Hamburg, Germany, 1973.

[End95] W. Enders. *Applied Econometric Time Series.* John Wiley and Sons, Hoboken, NJ, 1995.

[Fos94] T. Fossen. *Guidance and Control of Ocean Marine Vehicles.* John Wiley and Sons Ltd., New York, 1994.

[FPW98] G. Franklin, J. Powell, and M. Workman. *Digital Control of Dynamic Systems.* Ellis-Kagle Pressl, Half Moon Bay, CA, 1998.

[GH68] G. Groves and E. Hannan. Time-series regression of sea level on weather. *Review of Geophysics*, 6:129–174, 1968.

[Gre03] W. Greene. *Econometric Analysis.* Prentice Hall PTR, Upper Saddle River, NJ, 2003.

[Gud71] G. Gudmundsson. Time-series analysis of imports, exports and other economic variables. *Journal of Royal Statistical Society, Series A*, 134:383–412, 1971.

[Kay93] S. Kay. *Fundamentals of Statistical Signal Processing Estimation Theory.* Prentice Hall PTR, Upper Saddle River, NJ, 1993.

[KL11] S. Khanafseh and S. Langel. Implementation and experimental validation of cycle ambiguity resolution with position domain integrity risk constraints. *NAVIGATION: Journal of the Institute of Navigation*, 58:45–58, 2011.

[LW72] T. Lee and S. Wu. An application of stochastic and dynamic models for the control of a papermaking process. *Technometrics*, 14:481–496, 1972.

[NM94] W. Newey and D. McFadden. Chapter 35: Large sample estimation and hypothesis testing. In *Handbook of Econometrics*, pages 2153–2162, 1994.

[oW02] The Specialist Committee on Waves. Final report and recommendations to the 23rd ITTC. *23rd International Towing Tank Conference*, 2, 2002.

[Per05] T. Perez. *Ship Motion Control: Course Keeping and Roll Stabilisation using Rudder and Fins*. Springer, London, 2005.

[PF07] T. Perez and T. Fossen. Kinematic models for seakeeping and manoeuvring of marine vessels. *Modeling, Identification and Control*, 28:1–12, 2007.

[Pol99] D. Pollock. *A Handbook of Time-Series Analysis, Signal Processing And Dynamics*. Academic Press, Oval Road, London, 1999.

[Ric63] S. Rice. *Noise in FM Receivers, Chapter 25 in Time Series Analysis*. John Wiley and Sons, New York, 1963.

[Sco65] J. Scott. A sea spectrum for model tests and long-term ship prediction. *Journal of ship research*, 9:145–152, 1965.

[Sub13] N Subbaraman. X-47b navy drone completes first ever unmanned carrier landing. *NBC News*, 2013. http://www.nbcnews.com/tech/innovation/x-47b-navy-drone-completes-first-ever-unmanned-carrier-landing-f6C10591335.

[Wol38] H. Wold. *A Study in the Analysis of Stationary Time Series*. Almqvist and Wiksell, Uppsala, Sweden, 1938.

[Yul27] G. Yule. On a method of investigating periodicities in disturbed series with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society, Series A*, 226:267–298, 1927.