

Events Identification Using Box-Jenkins
Methodology with Application to Accelerated
Durability Tests of Ground Vehicles.

by

Md. Mostofa Ali Sarkar

A practicum submitted to the Faculty of Graduate Studies of
the University of Manitoba
in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Statistics
University of Manitoba
Winnipeg, MB, Canada

Copyright © 2012 by Md. Mostofa Ali Sarkar

Abstract

Durability tests are important to ensure the safety and reliability of a ground vehicle and involve frequently driving a vehicle through a series of events that simulate different road conditions or obstacles encountered during actual driving. Since durability tests are costly in-terms of time and money, accelerated durability lab tests can be used to spot failures before actual road tests. Signals of different events of the actual durability road tests generate three continuous time series data, that can be used to conduct accelerated durability lab tests. The actual analysis of these time series is very challenging because they are (i) of high frequency (ii) very noisy and (iii) inconsistent.

The purpose of this study was to identify the patterns of signals from the noisy and inconsistent time series data collected from the field tests. The Box-Jenkins methodology was used to identify models corresponding to different events. Due to complex structures of the real data, ARMA modelling was considered after testing stationarity of the given time series. While the time series data in vertical direction was used to identify the first three events, the time series in vertical, longitudinal and lateral directions were used to identify other four events.

Acknowledgments

It is a great pleasure to take this opportunity to convey my sincere thanks to those people who made this dissertation possible. First, I thank my supervisors Dr. Liqun Wang, Professor, Department of Statistics and Dr. Christine Q. Wu, Professor, Department of Mechanical and Manufacturing Engineering, University of Manitoba for their continuous support in the Master program. Their enthusiasm, inspiration, and great efforts to explain things adequately helped to make the research easier for me. They taught me how to analyze data and how to express ideas in non-technical words. Without their encouragement and constant guidance, I could not have finished this dissertation. I am also thankful to my committee members, Dr. Saumen Mandal and Dr. Julieta Frank for their valuable time and support.

I am greatly indebted to MITACS, MCI, Department of Statistics and Department of Mechanical and Manufacturing Engineering of University of Manitoba for their contribution to this long journey.

Last, but not least, I thank my family: my parents, for giving me life in the first place, for educating me, for unconditional support and encouragement to pursue my interests. I am indebted to my wife and wish to thank her especially for her continuous support, encouragement and enthusiasm to complete my research works. And finally, I wish to thank my son for refreshing me by his innocent smile.

Dedications

To my parents, my wife and my son Ehan.

Contents

Contents	iii
List of Tables	vi
List of Figures	viii
1 Introduction to the real data problem	1
1.1 Introduction	1
1.2 Literature review	5
1.2.1 Accelerated durability test	5
1.2.2 GlyphWorks software and events identification	8
1.3 Events description and time series information	11
1.4 Objectives	18
1.5 Organization of the study	18

2	Time series models and Box-Jenkins methodology	19
2.1	Introduction	19
2.2	Some basic concepts	20
2.2.1	ARIMA models	20
2.2.2	Stationarity	21
2.2.3	The augmented Dickey-Fuller unit root test	22
2.2.4	The autocovariance and autocorrelation functions	23
2.2.5	Parsimony	25
2.2.6	Model selection criteria	25
2.2.7	Model checking tools	27
2.3	Stationary time series models	28
2.3.1	Autoregressive (AR) models	28
2.3.2	Moving-average (MA) models	33
2.3.3	Autoregressive moving average (ARMA) models	36
2.4	Box-Jenkins methodology	39
2.4.1	Identification	39
2.4.2	Estimation	41
2.4.3	Diagnostic checking	41
2.4.4	Model selection	42
2.4.5	Summary of the analysis procedure	42

3	Modelling the events time series	44
3.1	Introduction	44
3.2	Data preparation	45
3.3	Model identification of event-1: 4” Deep Chuck Hole	47
3.3.1	Step-1	47
3.3.2	Step-2	49
3.3.3	Step-3	49
3.3.4	Step-4 and Step-5	51
3.3.5	Step-6	59
3.4	Fitted models and event identification procedure	60
4	Results	65
5	Conclusion	74
	Bibliography	76

List of Tables

1.1	Travelling speed for each of the seven events of the test track [33] .	13
2.1	Summary of ACF and PACF patterns for stationary models [21] . .	38
3.1	Summary of the augmented Dickey-Fuller (DF) unit root test for stationarity	49
3.2	Estimation results of an AR(7) model	52
3.3	Estimation results of a reduced AR(7) model	53
3.4	Estimation results of an AR(2) model	56
3.5	Estimation results of an ARMA(2,2) model	58
3.6	Summery results of the reduced AR(7) and AR(2)	59
3.7	Fitted models for all seven events from counter clockwise laps . . .	61
3.8	Fitted models for all seven events from clockwise laps	63
4.1	Identified start and end time for the three events of counter-clockwise laps using the acceleration data in the Z direction	66

4.2	Identified start and end time for the four events of counter-clockwise laps using the acceleration data in the Z, Y and X directions	66
4.3	Identified start and end time for the three events of clockwise laps using the acceleration data in the Z direction	67
4.4	Identified start and end time for the four events of clockwise laps using the acceleration data in the Z, Y and X directions	67
4.5	Identified duration for the three events of counter-clockwise laps using the acceleration data in the Z direction	68
4.6	Identified duration for the four events of counter-clockwise laps using the acceleration data in the Z, Y and X directions	68
4.7	Identified duration for the three events of clockwise laps using the acceleration data in the Z direction	69
4.8	Identified duration for the four events of clockwise laps using the acceleration data in the Z, Y and X directions	69

List of Figures

1.1	Typical testing facility for automotive vehicles [5]	2
1.2	WESTEST's MAST table [4]	4
1.3	A typical field test track for ground vehicle [5]	12
1.4	Event Geometry of (a) 4" Single Chuck Hole, (b) (3/4)" Chatter Bumps, (c) 1" Deep Random Chuck Holes [33]	14
1.5	Event Geometry of (d) Railroad Crossing (e) Staggered Bumps (f) Frame Twist (g) High Crown Intersection [33]	15
1.6	Time series of the acceleration data measured from the field test [33]	17
2.1	An example of autocorrelation function (ACF)	24
2.2	Theoretical ACF and PACF patterns for AR(1) model	30
2.3	Theoretical ACF and PACF patterns for AR(2) model	31
2.4	Theoretical ACF and PACF patterns for AR(2) model (continued)	32
2.5	Theoretical ACF and PACF patterns for MA(1) models	34
2.6	Theoretical ACF and PACF patterns for MA(2) models	35

2.7	Theoretical ACF and PACF patterns for MA(2) models (continued)	35
2.8	Theoretical ACF and PACF patterns for ARMA models	37
2.9	Theoretical ACF and PACF patterns for ARMA models(continued)	37
3.1	Time series of the acceleration data measured from the field test for the 1 st lap [33]	46
3.2	Vertical accelerations of event-1: 4 th Deep Chuck Hole from counter clockwise first lap	48
3.3	Correlogram of the vertical acceleration time series data of event-1 .	50
3.4	Residuals plot of the reduced AR(7) model	55
3.5	Residuals correlogram of the reduced AR(7) model	55
3.6	Residuals plot of AR(2) model	57
3.7	Residuals correlogram of model AR(2)	57
3.8	Event identification procedure for counter-clockwise time series . . .	62
3.9	Event identification procedure for clockwise time series	64
4.1	Identified three events from the 1st counter-clockwise lap in Z direction	70
4.2	Identified four events from the 1st counter-clockwise lap in Z direction	71
4.3	Identified three events from the 1st clockwise lap in Z direction . . .	72
4.4	Identified four events from the 1st clockwise lap in Z direction . . .	73

Chapter 1

Introduction to the real data problem

1.1 Introduction

It is very important for a ground vehicle to ensure the safety and reliability. In addition to reliability, comfort and performance, growing environmental and economic concerns are elevating demand for durable vehicles. The durability test of a ground vehicle is one way to ensure the demand for durable vehicles. It enables a company to determine the stresses under which their products can continue to perform - in other words, to make sure components are not going to fail under anticipated operating conditions.

Testing the long-term durability could mean driving the vehicle over a test track 24 hours a day, seven days a week, for more than a month. Durability test of a ground vehicle can be particularly challenging because of the way the vehicle is used and the roads on which the vehicle is driven and the product design can change significantly during both the design phase and the product lifecycle. Testing the

vehicle is the most common way to verify durability, but the testing is also expensive and time-consuming. Hence, field tests are designed to represent the average road condition of the real world.

The durability field test tracks are precious to determine the safety and durability of a vehicle components over its lifetime. For automotive vehicles, the map of a typical field test track of Pennsylvania Transportation Institute [5] is shown in Figure 1.1 (on page 2).

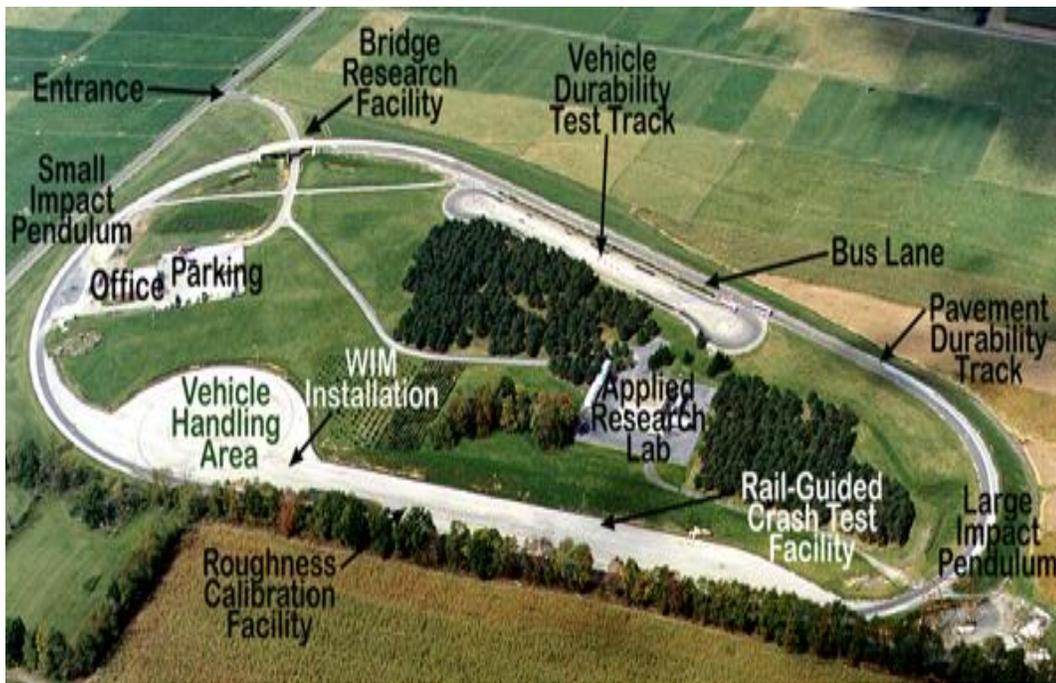


Figure 1.1: Typical testing facility for automotive vehicles [5]

To evaluate the vehicle there are different departments in the testing facility and the durability test track is used for the vehicle's durability assessment. The durability test track shown in Figure 1.1 would contain various events that would simulate typical real world driving conditions.

While doing the field test, occurrence of any failure will increase the time required to perform the test. The test is required to be repeated from the beginning once the problem is found and resolved. Since, the testing process is costly and time consuming, a sub-scaled accelerated durability test is essential to ensure that the testing is completed in a timely manner [15].

The accelerated durability tests are required to be used before the field tests to find the possible failure components and to evaluate the vehicle's performance within a much shorter time. At same time, the results of accelerated test should reveal the same potential damage content as the product would encountered over its service life. Other important information such as the determination of the warranty timeframe and minimization of product recalls and complaints after market can be obtained once the accelerated durability test is conducted [3].

The accelerated durability test is carried out by the use of Multi-axis Simulation Table (MAST) [33]. MAST system is often used to test the performance of various components in a realistic operating environment. Other important applications of MAST include the measurement of noise and vibration test. MAST is a mechanical system that works in a controlled laboratory environment by repeatedly replicate and analyze in service vibrations and motions of a testing component [2]. The MAST has a six degree of freedom control system [1]; it provides the mounting surfaces for the testing component analysis. This MAST operated through the uses of the hydraulic actuator and advanced flex test controller. In addition, different forms of data such as sinusoidal, random Power Spectrum Density (PSD), swept sine wave can be taken as the input of the MAST. Many advantages of the lab test comparable to the field test includes: individual test component can be tested

instead of the whole system, the testing time and cost commitment are significantly low, the modification of the road profiles for the field test can be conducted [1]. The prototype of the MAST that is used for the sub-scaled accelerated durability test is shown in Figure 1.2 (on page 4).

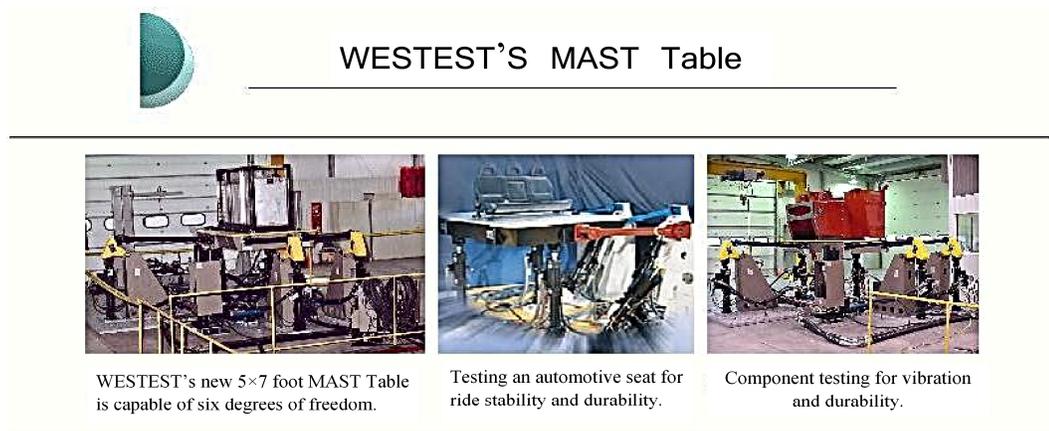


Figure 1.2: WESTEST's MAST table [4]

The MAST system that implements the accelerated durability test can be directly driven using the acceleration time series format data [16]. As a result, the driven profile of acceleration data with shorter duration needs to be created to implement the actual test on MAST, which is a crucial part of developing accelerated durability tests [33].

To ensure that the accelerated durability test is comparable to the field test,

initial data must be recorded from various road surfaces on the proving ground. It is remarkable that the field test data are: (i) of high frequency¹ (ii) very noisy² and (iii) inconsistent³. Accuracy of the accelerated durability test is mostly affected by the identified events' information.

1.2 Literature review

Accelerated test is mainly accomplished to assess the characteristics of a ground vehicle component under regular use conditions by testing the component at higher stress levels to accelerate the occurrence of failures. The challenge is to create the driven profile for MAST [33]. To create such driven profile for MAST, GlyphWorks software has been developed and used for many years. To create the mission profile and to develop the accelerated loading profile are the two important tasks to generate the accelerated loading profile using the commercial GlyphWorks software. The generated driven profile can be used repetitively for the durability assessment of the test component.

1.2.1 Accelerated durability test

Accelerated durability test is to test ground vehicles or their components included dynamic loads is often accomplished in the laboratory using road roughness simulation facilities [10]. Such tests can be carried out by a carefully designed field test,

¹The accelerations were measured at 200 Hz (200 times per second) and collected for 695 seconds.

²Along with accelerations, the time series also contain vehicle vibration due to the road surface and vehicle speed.

³The acceleration patterns for any of the events obtained from different laps of any driving directions are not exactly similar.

where both the roughness of the proving ground and the speed of the vehicle vary and control the degree of the test acceleration. To predict the service life of the vehicle, a representative loading profile is essential whose determination is a challenging task. Generally, two different types of loads are used to simulate the service loading environment, which are the sinusoidal loading and the random loading. The input random loading profiles are generally the accelerations, displacements, forces or strains [33]. Gopalakrishnan and Agrawal [20] provide an example of input load measurement. The measurement of tri-axial wheel forces, including the vertical, longitudinal and lateral is made using load-cell that is attached to the wheel hub. When the measurement approach is not realistic, Multi-body dynamic simulation technique can be used as an alternative to establish the dynamic input load [20].

In the automotive industry, two different approaches are used for assessing the accelerated durability once the loading profile is generated. The two approaches are: the numerical technique and the experimental testing [33]. Both the approaches have the same objectives of design cycle times and reduced development cost. Over the decade, many methods have been developed in literature regarding each approach, especially in the numerical approach.

In literature, the availability of the experimental durability assessment methods is very limited. Dressler, Spekert and Bitsch [22] has described three essential ways to systematize and speed up the experimental accelerated durability assessment. The first method is to use special test tracks which are designed to permit the ground vehicle to travel under the same load as the public roads in a much shorter time. However, it is a difficult task to derive a relation between the loads from the public roads and the loads from the test tracks. The second method is to bring

the test into the laboratory using servo-hydraulic test rigs [22]. The benefits of this approach are, the tests can be better reproducible and much better observable and can be performed within reduced time, as the individual testing component or subsystem is tested under the controlled environment other than the full vehicle prototype. However, the difficulty exists as the loads to be used for the rig testing have to be derived from measurements performed either on the tests tracks, public roads or a mix of both. The third method is to simulate the transfer of system loads (like the road profile) into component loads using multi-body models and termed as full numerical simulation. A major challenge here is to confirm that the subsystem and component loads in rig test and numerical simulation are exactly same and correspond to the loads created by the full vehicle driving on the test track [22].

A technique for developing a high accuracy laboratory durability test is demonstrated in [34]. The test was performed for a light-duty pickup truck on a six-degree-of-freedom road test simulator (RTS). Various transducers were instrumented on the vehicle under studied. The road data from the proving ground was used to drive file development on a RTS [34], for the subsequent accelerated durability testing. Vehicle response control channels configurations were used to compare the correlation of the transducer responses between those achieved on the RTS and the road data. The configuration that produced the best correlation with proving ground responses was adopted to develop the driven files used for subsequent accelerated testing [34].

Wannenburg and Heyns [31] have demonstrated a framework for mapping numerical methods and methodologies for the durability assessment of the ground vehicles. They have described the three steps for the numerical durability assess-

ment process, which are the determination of the appropriate input loading, stress analysis and fatigue analysis. The input loading is obtained either from different measurements (force input measurement, acceleration measurements, strain measurements) or dynamic simulations. The stress analysis can be preceded based on the input loading. Lastly, the fatigue analyses can be carried out based on the stress analysis or the strain analysis to predict the fatigue life of critical areas of the structure [31].

Each of the three steps for the numerical durability assessment process involves different methods. The methods of stress analysis are developed either in the time domain or in the frequency domain. Stress analysis methods available in literature are quasi-static finite element analysis, co-variance method, eigenvalue finite element analysis, transient dynamic direct integration method, etc. [31]. In these stress analysis methods different algorithm are employed for the determination of the stress level, and the choice of the algorithm depends on the type of input loading. For the fatigue analysis different methods are developed and the popular methods are Cycle Counting [18], Stress-life approach, strain-life approach [12] and fracture mechanics approach [31]. Wannenburg and Heyns state that in the automotive industry the strain-life approach is preferred due to its advanced theoretical background. But in the case of heavy vehicles the stress-life approach is preferred [31].

1.2.2 GlyphWorks software and events identification

The GlyphWorks is a software developed by HBM nCode, which is the leading data processing system for engineering test data analysis. It is designed to handle huge amounts of data with multi-file, multi-channel, multi-format environment

[33]. It uses for power processing and engineering test data analysis, with specific application to durability assessment and fatigue analysis [3]. Using this commercial GlyphWorks software, the accelerated durability test profiles can be achieved by using a so called Mission Profiling and Test Synthesis processes [33]. Generating the mission profiles using the Glyphworks software is a challenging job. The mission profiles can be efficiently created by using the acceleration data of the corresponding events (real life road obstacles) experienced by the vehicle. Hence, events identification is an important step to effectively conduct the accelerated durability test. Initial developments of events identification include the visual identification method [15] and the curve fitting method [35]. But these methods were not completely successful to identify the events. Based on wavelet analysis, clustering, and Fourier analysis, another event identification approach proposed in [33] to generate the mission profiles for the accelerated durability test. Accuracy of the mission profiles depends on the exactly identified events of the field test track. In this study Box-Jenkins time series methodology has applied to identify the events signals from the acceleration time series data.

Box-Jenkins time series methodology

The Box-Jenkins approach was described and employed to determine models for actual time series in a highly influential book by statisticians George Box and Gwilym Jenkins in 1970 [9]. The statistical nature and wide applicability of the Box-Jenkins models are well known. Box-Jenkins approach covers a large class of models and it is the standard and systematic approach to model identification, in which the validity of the model can be verified and the forecast accuracy can be measured

[21].

There are many published examples of Box-Jenkins approach in the areas of modelling and forecasting. A successful use of this approach to forecasting and control problems based on the modelling of time series is described in [29]. Another application of this approach is to estimate and to simulate earthquake ground motion and structural response [27], [6]. It is found that the time-varying autoregressive moving average (ARMA) model is an efficient method for estimating the observed ground motion as well as for simulating earthquake ground motion. Aghda, Gandomi and Beitollahi [7] have applied Box-Jenkins model to develop the artificial accelerograms for assessing the dynamic response of structures. They argued the simplicity and applicability of the AR model in signal processing. Box-Jenkins ARMA modelling can also be applied to predict the wind power output for producing electricity [24].

Box-Jenkins modelling involves identifying an appropriate ARMA model, fitting it to the data, and then using the fitted model for forecasting. One of the attractive features of the Box-Jenkins approach is that ARMA processes are a very rich class of possible models and it is usually possible to find a process which provides an adequate description of the data [17]. The Box-Jenkins modelling involved an iterative three-stage process of model identification, parameter estimation and model checking. Recent explanations of the process (e.g., Makridakis, Wheelwright and Hyndman [25]) often add a preliminary stage of data preparation and a final stage of model application (or forecasting).

Data preparation involves transformations of the data (such as square roots or logarithms) or differencing the data that can help stabilize the variance in a series

where the variation changes with the level.

Model identification in the Box-Jenkins framework is the most important and also the most difficult task. This stage consists of using various graphs based on the given time series to identify potential ARMA models which might provide a good fit to the data [17].

Parameter estimation means finding the values of the model coefficients which provide the best fit to the data. There are sophisticated computational algorithms designed to do this [25]. The common approaches to fitting Box-Jenkins models are non-linear least square estimation and maximum likelihood estimation [17].

Model checking involves testing the assumptions of the model to identify any areas where the estimated model is inadequate. This diagnostic checking involves a residual analysis to ensure that the residuals from the estimated model mimic a white noise process [19]. If the model is found to be inadequate, it is necessary to go back to identification step and try to identify a better model.

1.3 Events description and time series information

The field test time series used in this study to identify the events experienced by a ground vehicle are actually a 695 second time series of accelerations recorded by MCI from Altoona durability test track [16]. A typical proving ground field test track is shown in Figure 1.3 (on page 12). This test track contained seven different events that are separated into two sections: the three-event section (lower section as shown in Figure 1.3) and the four-event section (upper section as shown

in Figure 1.3). The seven events are : 4" Single Chuck Hole, (3/4)" Chatter Bumps, 1" Deep Random Chuck Holes, Railroad Crossing, Staggered Bumps, Frame Twist and High Crown Intersection.

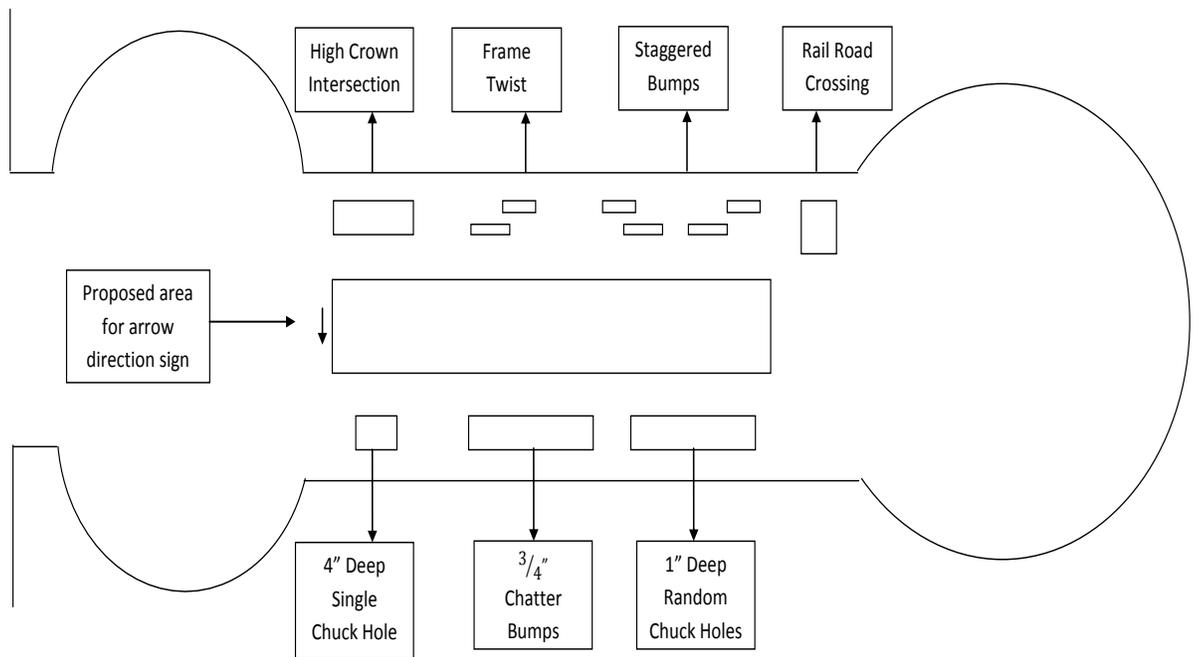


Figure 1.3: A typical field test track for ground vehicle [5]

As shown in Figure 1.3, there are two transitions which can be observed between the three-event section and the four-event section. Including both of the transitions and events sections the actual field test duration is 650 hours in total [33]. For the event sections the field test duration is 330 hours. Corresponding to each event the measured travelling speed of the vehicle on the test track is shown in Table 1.1 (on page 13).

Table 1.1: Travelling speed for each of the seven events of the test track [33]

Events Order	Events	Vehicle speed
Event-1	4" Deep Chuck Hole	5 mph
Event-2	(3/4)" Chatter Bumps	20 mph
Event-3	1" Random Chuck Holes	20 mph
Event-4	Rail Road Crossing	8 mph
Event-5	Staggered Bumps	10 mph
Event-6	Frame Twist	10 mph
Event-7	High Crown Intersection	10 mph

The estimated durations of each seven events was calculated in [33] based on the measured driving speed given in Table 1.1, and on the information about the events geometries. The events geometries are shown in Figure 1.4(on page 14) and in Figure 1.5(on page 15).

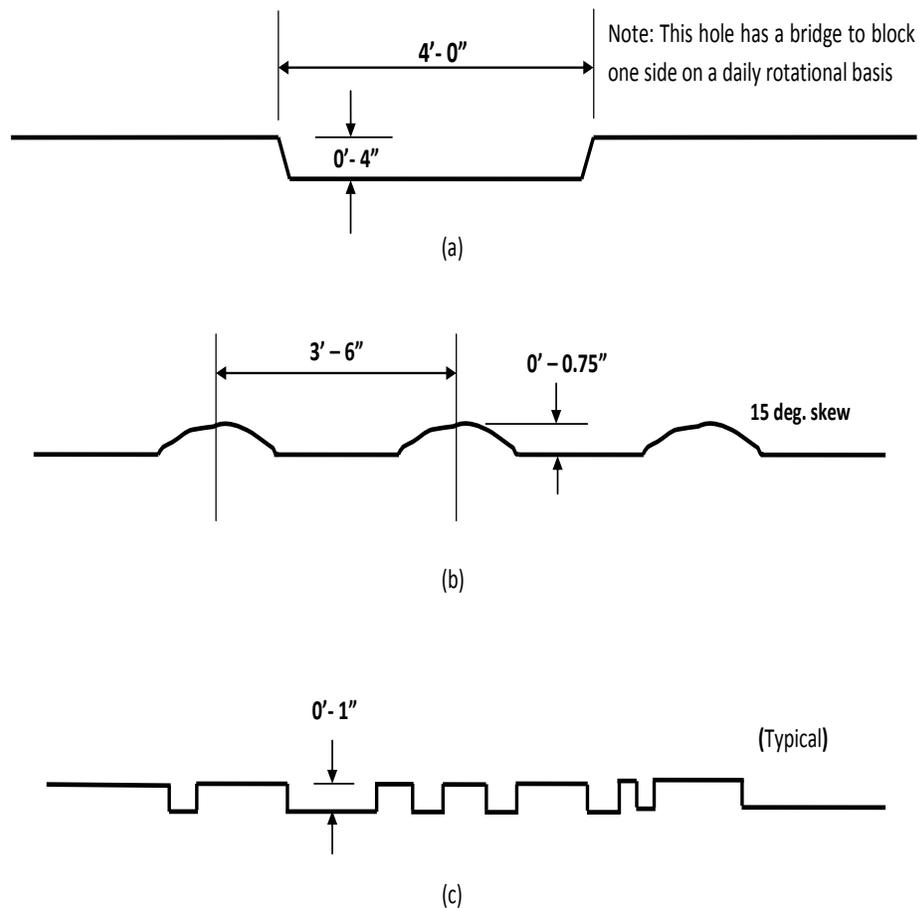


Figure 1.4: Event Geometry of (a) 4" Single Chuck Hole, (b) (3/4)" Chatter Bumps, (c) 1" Deep Random Chuck Holes [33]

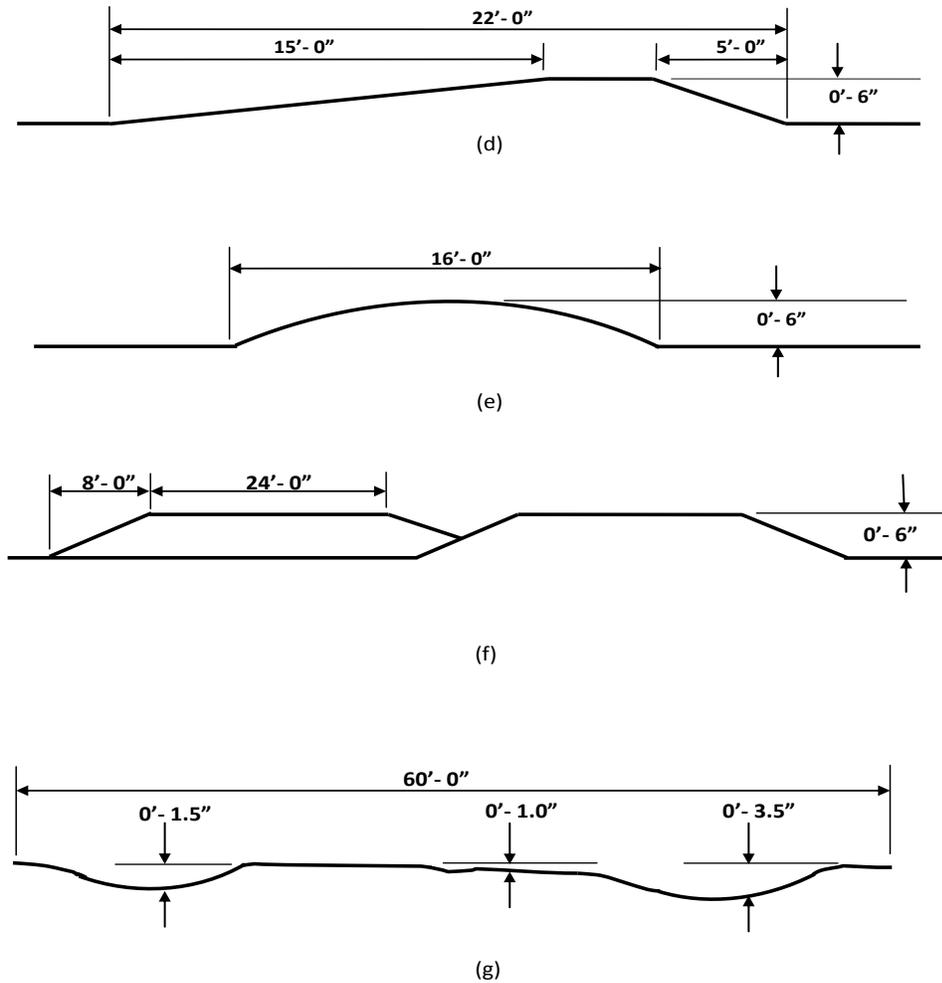


Figure 1.5: Event Geometry of (d) Railroad Crossing (e) Staggered Bumps (f) Frame Twist (g) High Crown Intersection [33]

The set of recorded field test time series data used for the events identification are shown in Figure 1.6 (on page 17). The horizontal axis is the time in the unit of second and the vertical axis is the acceleration in the unit of g . As mentioned earlier of this section, these time series are collected accelerations of the vehicle which includes six laps: three counter-clockwise laps followed by three clockwise

laps. Each of the six laps contains seven transient events [33]. In total, there were forty-two events among which 28 events from the first two laps of each driving directions were separated by visual inspection of the data and by the use of GPS data [15]. Last two laps, one from the counter clockwise and the other from the clockwise driving directions were kept to test the identification procedures. As shown in Figure 1.6, the X, Y and Z direction used to represent various components of the acceleration measured at a point near to the left front axle of the vehicle. The X direction represents longitudinal accelerations, the Y direction represents lateral accelerations and the Z direction represents the vertical accelerations [16]. The acceleration time series in all three directions⁴ will be used in this study for the events identification.

⁴The three directions include X, Y and Z where, X represents the longitudinal acceleration time series, Y represents the lateral acceleration time series and Z represents the vertical acceleration time series.

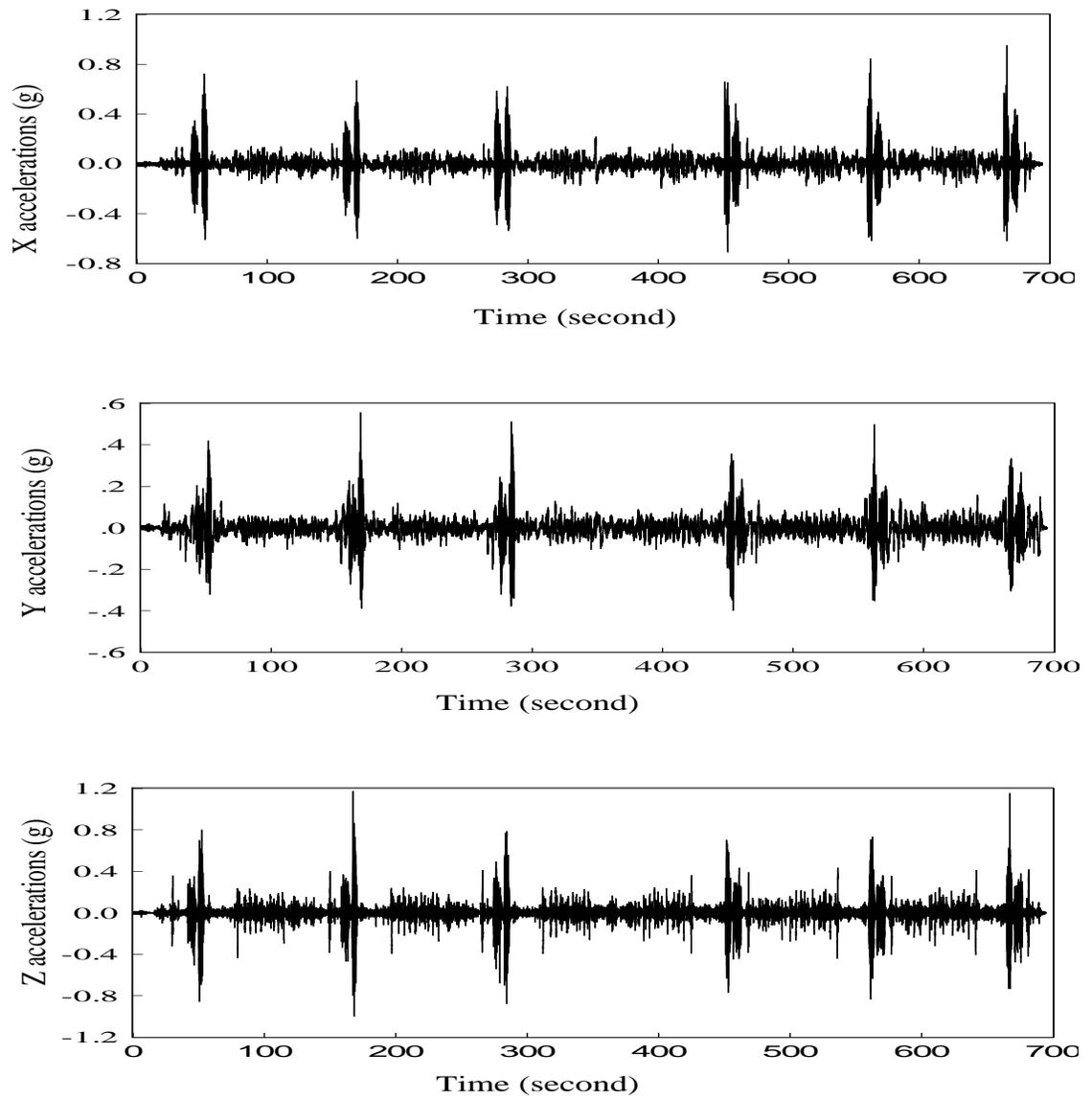


Figure 1.6: Time series of the acceleration data measured from the field test [33]

1.4 Objectives

The objective of this study was to identify the seven events (Table 1.1) of the actual road test data from statistical platform. The Box-Jenkins' methodology was used for modelling the given time series of each event. Thus, the purpose of this project was to identify the patterns of signals from these noisy and inconsistent time series.

ARMA modelling was considered for the data analysis as the real data has a noisy and complex structures. Based on the fitted models of the road test events, model identification procedures were proposed to identify these events.

1.5 Organization of the study

ARMA modelling, the Box-Jenkins methodology and the relevant tools used for data analyses have described in chapter-2. Modelling of the acceleration time series have discussed in chapter-3. Chapter-4 displayed the results obtained through the data analyses and chapter-5 described the concluding remarks, relevant limitations and future studies on this study.

Chapter 2

Time series models and Box-Jenkins methodology

2.1 Introduction

A time series is a collection of observations of a particular variable in an ordered sequence. The ordering is usually through time intervals of equal length but may also be taken through other dimensions, like space [30]. Time series occur in a variety of fields like in business and economics, in engineering, in natural and social sciences.

Time series analysis specially involves choice of an appropriate family of models, estimation of unknown parameters and diagnostic check for goodness of fit. One of the objectives of time series analysis is to use the fitted model to enhance our understanding and description of the mechanism generating the data [30], [13]. The model may be used simply to provide a compact description of the data.

In this study time series data for the test track acceleration measurements are analyzed separately. As an analysis of single time series is called a univariate time

series and this is the topic of this chapter. In this chapter the univariate ARMA models will be discussed and the focus will be given on how to identify these models followed by their estimation and diagnostic checking for adequacy.

2.2 Some basic concepts

This section will discuss some concepts that are necessary for proper understanding of time series models, model selection and model checking techniques.

2.2.1 ARIMA models

Box and Jenkins (1976) first introduced the Autoregressive Integrated Moving-Average (ARIMA) models [11]. In theory, ARIMA is a general class of models for describing and forecasting a time series. Lags of differenced series appearing in the forecasting equation are called Autoregressive, lags of the forecast errors are called Moving Average, and the time series which needs to be differenced to be made stationary is said to be an Integrated version of a stationary series [26], [23]. A general ARIMA model is described below,

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d(Z_t - \mu) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)u_t$$

or

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - B)^d(Z_t - \mu) = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)u_t$$

or

$$\phi(B)(1 - B)^d(Z_t - \mu) = \theta(B)u_t$$

where,

Z_t , is the observed series,

μ , is the mean of the series,

$\phi(B)$, is the autoregressive polynomial of order p ,

$\theta(B)$, is the moving-average polynomial of order q ,

d , is the number of non-seasonal differences,

B , is the backshift operator ($B^k Z_t = Z_{t-k}$),

u_t , is a white noise process $WN(0, \sigma_u^2)$.

Practically, many time series are so-called stationary and can be described by an ARMA model, that is $d=0$. In this case the model has the form,

$$Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} = \mu + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q}$$

or

$$Z_t - \phi_1 Z_{t-1} - \dots - \phi_p Z_{t-p} = \mu + u_t + \theta_1 u_{t-1} + \dots + \theta_q u_{t-q}$$

2.2.2 Stationarity

A key concept underlying time series processes is stationarity. Stationarity is important because if the series is non-stationary then most typical results of the standard regression analysis are not valid. Regression with non-stationary time series may have no meaning and are therefore called spurious.

Generally speaking, a covariance stationary time series has the following three characteristics [19]:

(a) exhibits mean reversion in that it fluctuates around a constant long-run mean;

- (b) has a finite variance that is time invariant; and
- (c) the correlation between any two observations depends only on the time lag and diminishes as the lag length increases.

Mathematically, a time series Z_t is said to be stationary (or covariance stationary) if:

- (a) the mean $E(Z_t) = \mu$ is constant for all t ;
- (b) the variance $Var(Z_t) = E[(Z_t - \mu)^2] = \sigma^2$ is constant for all t ;
- (c) the covariance $Cov(Z_t, Z_{t-k}) = E[(Z_t - \mu)(Z_{t-k} - \mu)] = \gamma_k$ depends on k only.

In the literature, a covariance stationary process is also referred to as a weakly stationary, second-order stationary, or wide-sense-stationary process. A stationary ARIMA model is called an autoregressive moving average (ARMA) model [11].

2.2.3 The augmented Dickey-Fuller unit root test

The augmented Dickey-Fuller (DF) unit root test is the regular t -test on the coefficient of the lagged dependent variable and does not have a conventional t -distribution. So, special critical values (which were originally calculated by Dickey and Fuller) are to be used in decision rule. This test actually performed to test the null hypothesis that the time series contains a unit root against the alternative hypothesis which states that the series is stationary. If the augmented DF test statistic value is smaller than the critical value then the null hypothesis of a unit root will be rejected and it can be concluded that the time series is stationary [11].

2.2.4 The autocovariance and autocorrelation functions

For a stationary process Z_t , we have the mean $E(Z_t) = \mu$ and variance $Var(Z_t) = E[(Z_t - \mu)^2] = \sigma^2$, which are constant, and the covariances $Cov(Z_t, Z_{t-k})$, which are functions only of the time difference k . Hence, we can write the covariance between Z_t and Z_{t+k} as,

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E[(Z_t - \mu)(Z_{t+k} - \mu)],$$

where γ_k is called the autocovariance function of lag k .

Autocorrelations are statistical measures of how a time series is related to itself over time. The general concept of measuring the correlation between two sets of data is the basis for measuring the autocorrelation of a time series. While calculating the autocorrelation for a given time series at some lag k , we actually calculating the correlation between two distinct data sets, namely, the original series (Z_t) and the same series (Z_{t+k}) moved forward in time a specified number of periods (lag). Hence, the correlation between Z_t and Z_{t+k} is defined as,

$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0}$$

where $Var(Z_t) = Var(Z_{t+k}) = \gamma_0$ and ρ_k is called the autocorrelation function (ACF).

The autocorrelation function of a stationary process is an important tool for assessing its properties. The graph obtained by plotting the autocorrelation function ρ_k against the (non negative)lag k is known as **correlogram**. Figure 2.1 (on page24)

shows an example of autocorrelation function for AR(1) model with a negative autoregressive parameter. Horizontal axis shows the lag values and vertical axis shows the corresponding correlation against the lag values.

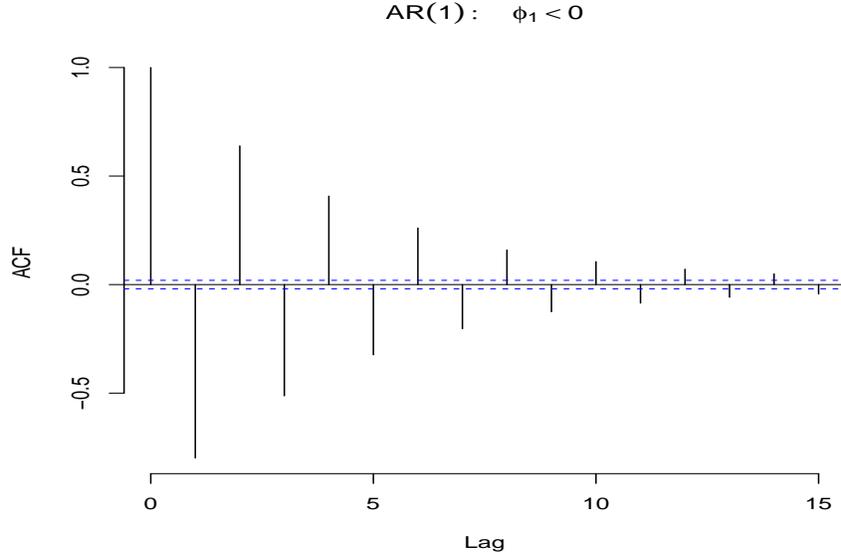


Figure 2.1: An example of autocorrelation function (ACF)

The partial correlations are another set of statistical measures, similar to autocorrelations, that are used to evaluate relationships among the series values. While the autocorrelation between Z_t and Z_{t+k} accounts for their *total* correlation, the partial correlation accounts for their *net* correlation after their mutual linear dependency on the intervening variables Z_{t+1}, Z_{t+2}, \dots , and Z_{t+k-1} has been removed [32]. Thus the partial autocorrelation function (PACF) between Z_t and Z_{t+k} is defined as,

$$\varphi_k = \text{Corr}(Z_t, Z_{t+k} | Z_{t+1}, \dots, Z_{t+k-1}) = \frac{\text{Cov}(Z_t, Z_{t+k} | Z_{t+1}, \dots, Z_{t+k-1})}{\sqrt{\text{Var}(Z_t)} \sqrt{\text{Var}(Z_{t+k})}},$$

where the term *Corr* stands for correlation.

Partial correlations are complementary to autocorrelations with respect to patterns they produce for autoregressive and moving-average relationships. Time series generated by ARMA models have easily identifiable autocorrelation and partial autocorrelation patterns. Theoretical ACF and PACF patterns of the ARMA models can be used as standard for comparing with the sample ACF and the sample PACF to identify the required number of moving average (MA) and autoregressive (AR) parameters in the initial models.

2.2.5 Parsimony

A fundamental idea in the Box-Jenkins procedure is the principle of parsimony [19]. The principle refers to representing the time series with as few parameters as possible, this means simpler representations of a time series process are more desirable than more complex one if both are adequate. Box and Jenkins argued that parsimonious models produce better forecasts than overparameterized models. A parsimonious model fits the data well without incorporating any redundant coefficients.

2.2.6 Model selection criteria

In time series analysis there may be several adequate models that can be used for representing a given data set. Sometimes it is difficult to choose a unique model. Thus various criteria have been introduced for model selection and can be viewed as measures of goodness-of-fit [19]. These criteria are different from the model identification methods. Theoretically, it is assumed that the residuals from all

adequate models should be white noise for a given series. For multiple adequate models, the selection criterion is normally based on summary statistics from the residuals computed from a fitted model or on forecast errors calculated from out-of-sample forecast. Some of the model selection criteria are described below.

Akaike's information criterion (AIC)

A general criterion for model selection is to minimize a quantity called Akaike's Information Criterion,

$$-2\log(\textit{likelihood}) + 2p$$

where, *likelihood* stands for maximized likelihood and p is the number of estimated parameters.

The AIC can be used to compare ARMA models. It is known that AIC criterion tends to overestimate the order of the model. Then Akaike has also developed a Bayesian modification of AIC, denoted by BIC, which penalizes models with large numbers of parameters in a more severe way than the AIC. If N denotes the number of observations to which the model is fitted, then BIC replaces the term $2p$ in the AIC criterion by $(p + p \log(N))$ [14].

Schwartz's SBC criterion

Similar to Akaike's BIC, Schwartz (1978) [28] suggested another Bayesian criterion of model selection, which has been called Schwartz's Bayesian Criterion (SBC). Ideally, the AIC and SBC will be as small as possible (both can be negative). The SBC will always select a parsimonious model than will the AIC. Again, the SBC is

asymptotically consistent while the AIC is biased toward select an overparameterized model. However, in small samples, the AIC can work better than the SBC. We can be quite confident in our results if both the AIC and the SBC select the same model.

Adjusted- R^2

The R-squared (R^2) statistic measures the success of the regression in predicting the values of the dependent variable within the sample. In standard settings, R^2 may be interpreted as the fraction of the variation of the dependent variable explained by the independent variables. The statistic will equal to 1 if the regression fits perfectly, and zero if it fits no better than the simple mean of the dependent variable. One problem with using R^2 as a measure of goodness of fit is that the R^2 will never decrease as we add more regressors.

The adjusted- R^2 , penalizes the R^2 for the addition of regressors which do not contribute to the explanatory power of the model. The adjusted- R^2 is never larger than the R^2 , can decrease as we add regressors, and for poorly fitted models, may be negative.

2.2.7 Model checking tools

When a model has been fitted to a time series, it is advisable to check that the model really does provide an adequate description of the data. After estimation of parameters we have to assess the model adequacy by detecting whether the model assumptions are violated. The basic assumption is that the residuals are white noise,

that is, the residuals are uncorrelated random shocks with zero mean and constant variance. Hence, model diagnostic checking is accomplished through careful analysis of residual series. Two obvious steps are to plot the residuals in a time plot, and to calculate the correlogram of the residuals. If the residual analysis indicates that the fitted model is inadequate in some way then alternative models may need to be tried.

2.3 Stationary time series models

It is necessary to have basic ideas about the stationary time series models for appropriate use of the time series analysis and Box-Jenkins methodology. Stationary time series models are briefly discussed here along with their patterns of autocorrelation function and partial autocorrelation function.

2.3.1 Autoregressive (AR) models

A simple time series model is the autoregressive of order one, or AR(1) model. This model is written in the following form:

$$Z_t = \mu + \phi_1 Z_{t-1} + u_t$$

where μ is the constant term and $|\phi_1| < 1$ and u_t is a white noise error term at period t , and Z_t is the stationary series.

The implication behind the AR(1) model is that the time series behaviour of Z_t is largely determined by its own value in the preceding period (Z_{t-1}). That is, what

happens this period (t) is largely dependent on what happened in the last period ($t - 1$), plus some current random error. Or, alternatively what will happen in next period ($t + 1$) will be determined by the value of the series in the current period (t) plus some random error from next period.

It is possible of course, that Z_t could be directly related to more than just one past value. For example, the AR(2) model will be an autoregressive model of order two, and which has the form:

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + u_t$$

This model indicates that Z_t is related to a combination of the two immediately preceding values Z_{t-1} and Z_{t-2} , plus some current random error u_t . Extending this idea further, we may write a general autoregressive model of order p as follows:

$$Z_t = \mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \phi_3 Z_{t-3} + \dots + \phi_p Z_{t-p} + u_t,$$

where μ is the constant term and $\phi_1, \phi_2, \dots, \phi_p$ are the AR parameters.

The subscripts on the ϕ 's are called the orders of the AR parameters. The highest order p is referred to as the order of the model.

ACF and PACF of AR models:

The autocorrelation function of an AR(p) process tails off as a mixture of exponential decays and/or damped sine waves depending on the nature of the parameters. The partial autocorrelation function of the general AR(p) process cuts off after lag p . This important property enables us to identify whether a given time series is generated by an autoregressive process

Figure 2.2 (on page 30) shows the theoretical ACF and partial-acf (PACF) patterns of an AR(1) model. As shown in Figure 2.2, top-left plot shows the ACF and the top-right plot shows the partial-acf (PACF) of the AR(1) model with positive autoregressive parameter ($\phi_1 > 0$). The ACF decreases as the lag number increases and the partial-acf has a positive value (single spike) at lag one. The lower two plots in Figure 2.2 show the ACF and the PACF of the AR(1) model with negative autoregressive parameter ($\phi_1 < 0$). The bottom-left plot shows that the ACF has an alternating decreasing pattern as lag number increases and bottom-right plot shows that the PACF has a negative value at lag one.

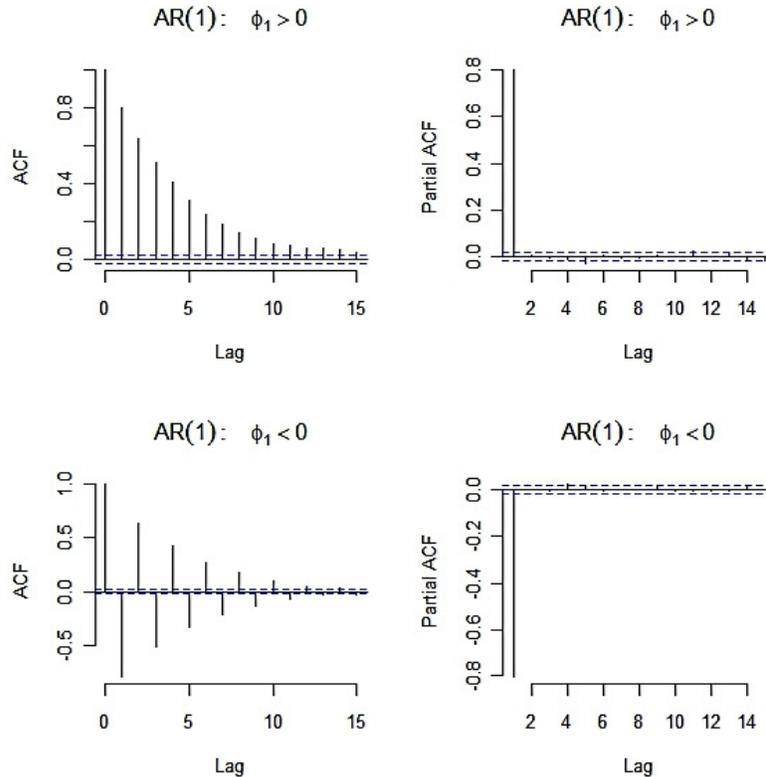


Figure 2.2: Theoretical ACF and PACF patterns for AR(1) model

Figure 2.3 (on page 31) shows the theoretical ACF and PACF patterns of an AR(2) model. Top two plots show the ACF and PACF pattern of an AR(2) model with two positive autoregressive parameters ($\phi_1 > 0, \phi_2 > 0$). The ACF has a decreasing pattern and the PACF has two spikes on the positive side of the axis. Similarly, bottom two plots show the ACF and PACF pattern of an AR(2) model with one negative ($\phi_1 < 0$) and one positive ($\phi_2 > 0$) autoregressive parameters.

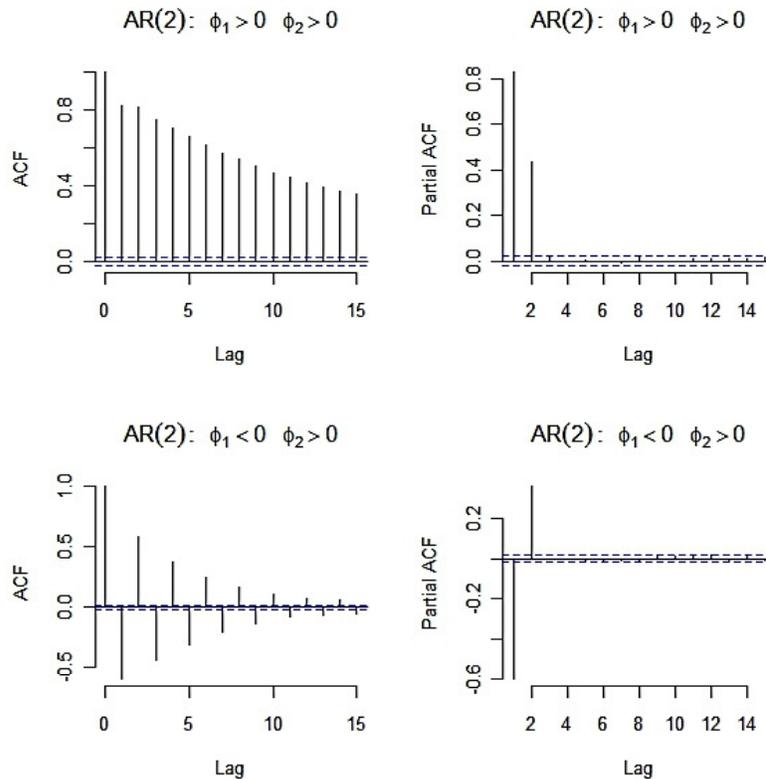


Figure 2.3: Theoretical ACF and PACF patterns for AR(2) model

The ACF has an alternating decreasing pattern and the PACF has two alternating spikes depending on the nature of the autoregressive parameters. Figure 2.4 (on page 32) also shows the theoretical ACF and PACF patterns of an AR(2)

model. Top two plots showing the ACF and PACF patterns for one positive and one negative autoregressive parameters and the bottom two plots showing the ACF and PACF patterns for two negative autoregressive parameters.

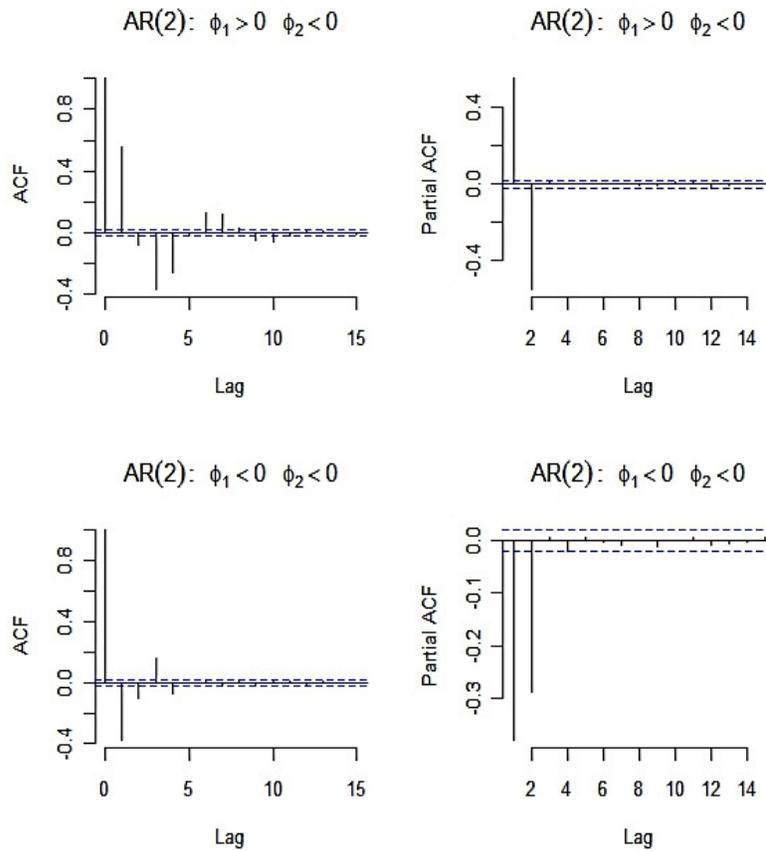


Figure 2.4: Theoretical ACF and PACF patterns for AR(2) model (continued)

2.3.2 Moving-average (MA) models

Moving average models closely resemble AR models in appearance but the concept behind the use of MA parameters is quite different [21]. Moving average parameters relate what happens in period t only to the random errors that occurred in past periods, i.e., to u_{t-1}, u_{t-2}, \dots (as opposed to being related to the actual series values Z_{t-1}, Z_{t-2}, \dots). Because any MA(q) process is by definition, an average of q stationary white noise processes, it follows that every moving average process is stationary, as long as q is finite.

A simple moving average model is that of order one, or MA(1) model, which has the following form:

$$Z_t = \mu + \theta u_{t-1} + u_t,$$

where μ is the constant term and θ is the MA parameter of order one. The above model simply says that any given value Z_t in the series is directly proportional only to the random error u_{t-1} from the previous period plus some current random error u_t .

The general form of the MA model is an MA(q) model of the following form:

$$Z_t = \mu + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q}$$

where $\theta_1, \theta_2, \dots, \theta_q$ are the MA parameters. The highest order q is referred to as the order of the MA model.

ACF and PACF of MA models:

The autocorrelation function of an MA(q) process cuts off after lag q . This important property enables us to identify whether a given time series is generated by a

moving average process [32].

The partial autocorrelation function of the general $MA(q)$ process tails off as a mixture of exponential decays and/or damped sine waves depending on the nature of the parameters.

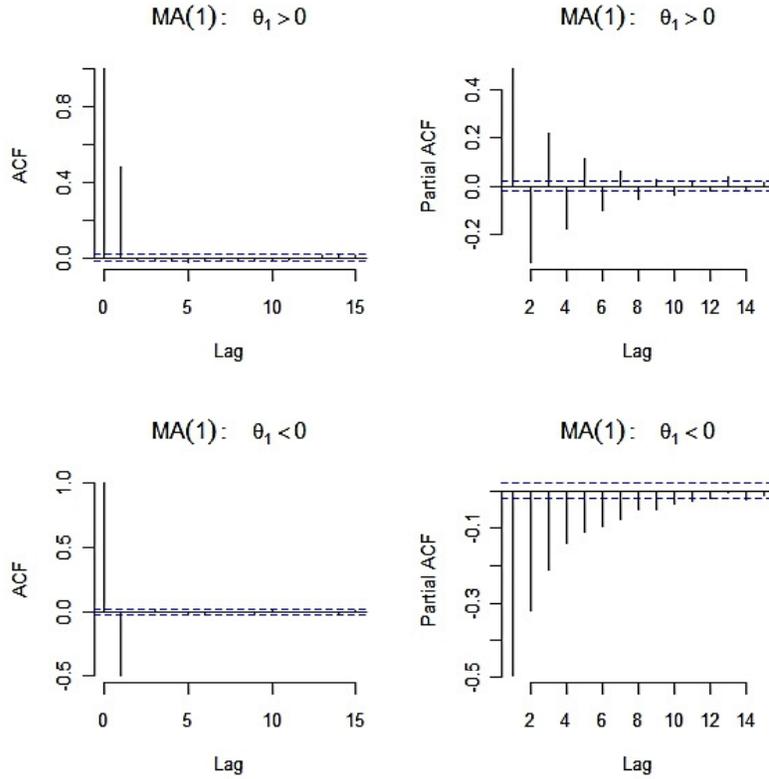


Figure 2.5: Theoretical ACF and PACF patterns for MA(1) models

Figure 2.5 (on page 34) shows the theoretical ACF and PACF patterns of a MA(1) model where, the ACF and the partial-acf (PACF) plots showing different patterns depending on the nature of the parameter θ . Figure 2.6 (on page 35) and Figure 2.7 (on page 35) show the theoretical ACF and PACF patterns for a MA(2) model.

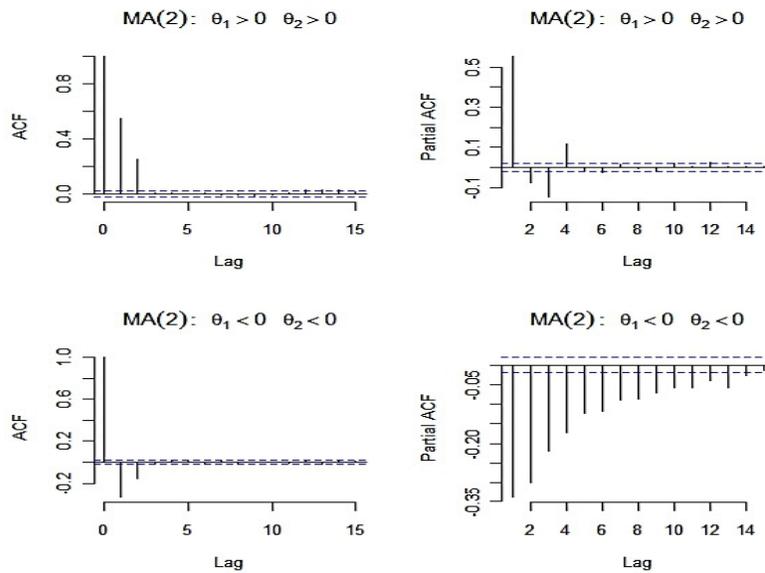


Figure 2.6: Theoretical ACF and PACF patterns for MA(2) models

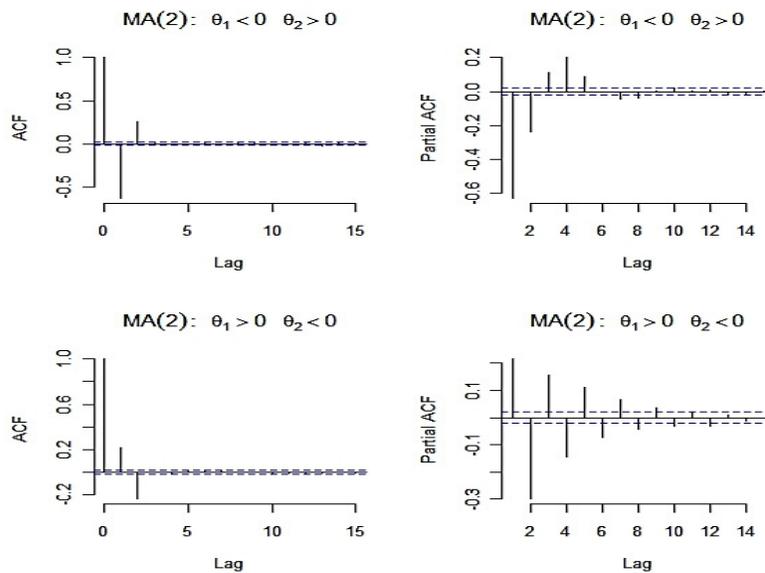


Figure 2.7: Theoretical ACF and PACF patterns for MA(2) models (continued)

2.3.3 Autoregressive moving average (ARMA) models

A useful class of models for time series is formed by combining MA and AR processes. A mixed autoregressive moving-average model containing p AR terms and q MA terms is said to be an ARMA model of order (p, q) and is written in the following form:

$$Z_t - (\mu + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p}) = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q}$$

The order of an ARMA model is expressed in terms of both p and q . The importance of ARMA process lies in the fact that a stationary time series may often be described by an ARMA model involving fewer parameters than a pure MA or AR process by itself.

Figure 2.8 (on page 37) and Figure 2.9 (on page 37) show the theoretical ACF and PACF patterns of an ARMA(1,1) model. The ACF and the partial-acf(PACF) plots show different patterns depending on the nature of the model parameters.

Table 2.1 (on page 38) summarizes some possible combinations of ACF and PACF forms [19],[30]. These combinations will be used to detect the order of ARMA models as well as for model identification.

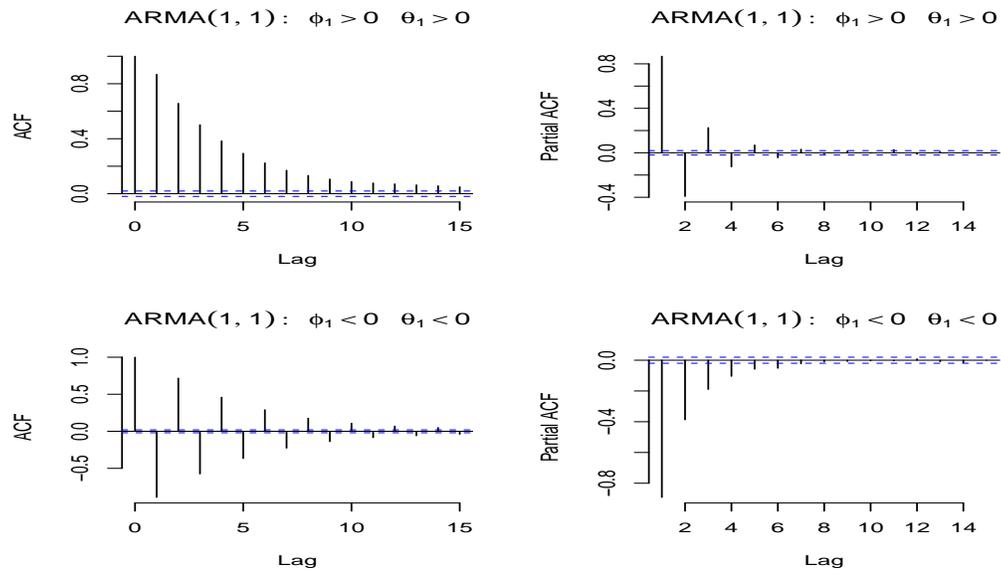


Figure 2.8: Theoretical ACF and PACF patterns for ARMA models

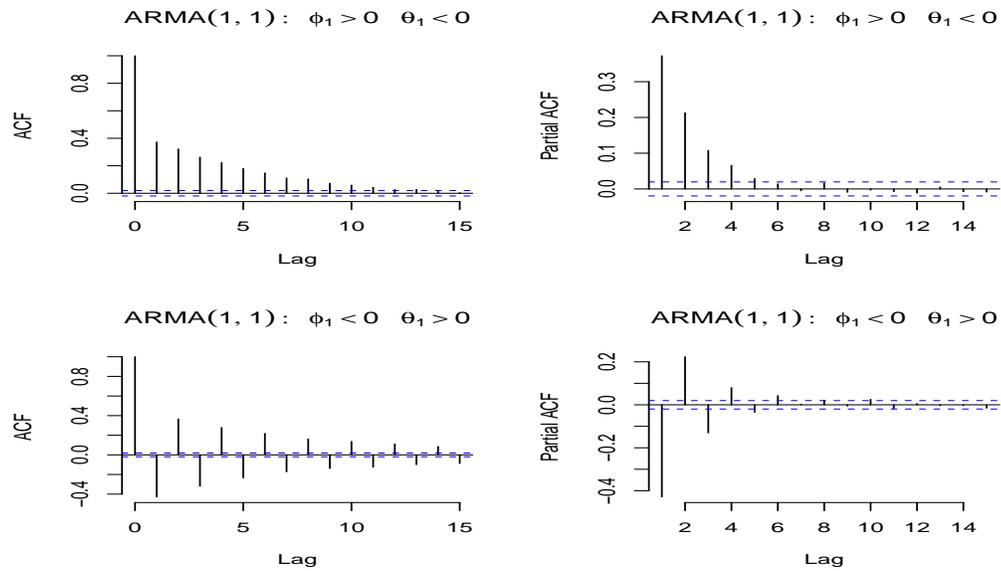


Figure 2.9: Theoretical ACF and PACF patterns for ARMA models(continued)

Table 2.1: Summary of ACF and PACF patterns for stationary models [21]

Process	ACF(ρ)	PACF (ψ)
Pure white noise	All autocorrelations are zero.	All partial autocorrelations are zero.
AR(1)	Tails off as exponentially decay on positive side if $\phi_1 > 0$. An alternating sign on negative side or oscillating decay if $\phi_1 < 0$.	Single spike at lag 1, then cuts off to zero: spike positive if $\phi_1 > 0$ and negative if $\phi_1 < 0$.
MA(1)	Single spike at lag 1, then cuts off to zero: spike positive if $\theta_1 < 0$ and negative if $\theta_1 > 0$.	Tails off as exponentially decay on negative side if $\theta_1 > 0$. An alternating sign on positive side or oscillating decay if $\theta_1 < 0$.
AR(p)	Exponential decays towards zero, may contain damped oscillations depends on the sign and sizes of $\phi_1, \phi_2, \dots, \phi_p$.	Spikes through lag p , then cuts off to zero.
MA(q)	Spikes through lag q , then cuts off to zero.	Exponential decays towards zero, may contain damped oscillations depends on the sign and sizes of $\theta_1, \theta_2, \dots, \theta_q$.
ARMA(1,1)	Exponential decay after lag 1 if $\phi_1 > 0$. Oscillating decay after lag 1 if $\phi_1 < 0$.	Oscillating decay after lag 1. Exponential decay after lag 1.
ARMA(p,q)	Decay (either direct or oscillatory) after lag q .	Decay (either direct or oscillatory) after lag p .

2.4 Box-Jenkins methodology

Box and Jenkins introduced time series models for autoregressive (AR), moving averages (MA), or autoregressive moving average (ARMA) processes [30]. A fundamental idea in the Box-Jenkins approach is the principle of parsimony. Box and Jenkins argued that parsimonious models produce better forecasts than over-parameterized models. In general Box and Jenkins modelling procedure involves an iterative three-stage method aimed at selecting an appropriate (parsimonious) ARMA model for the purpose of estimating and forecasting a univariate time series. The three stages are : (a) identification; (b) estimation and (c) diagnostic checking.

2.4.1 Identification

The essence of identification problem is that any model may be given more than one (and in most cases many) different representations, which are essentially equivalent [11]. Defining the best representation is fairly easy and here we use the principle of parsimony, this means that we pick the model with the minimum number of parameters to be estimated.

In identification stage we visually examine the time plot of the series autocorrelation function, and partial autocorrelation function. Plotting each observation of the sequence against time t provides useful information concerning outliers, missing values, and structural breaks in the data. A comparison of the computed sample ACF and sample PACF patterns to those of various known theoretical ARMA models may suggest several plausible models [21]. For non-stationary data common stationarity-inducing transformations are applied to make them stationary.

After achieving stationarity, the next step is to identify the orders (p and q) of the model. For a pure MA(q) process, the ACF will tend to show estimated autocorrelations which are significantly different from zero up to lag q and then it will die down immediately after the q^{th} lag. The PACF of a MA(q) process will tend to die down quickly either by an exponential decay or by a damped sine wave [30]. Contrary to the moving average (MA) process, the pure AR(p) process will have an ACF which will tend to die quickly either by exponential decay or by a damped sine wave, while the PACF will tend to show spikes (significant autocorrelations) for lags up to p and then it will die down immediately.

If neither the ACF nor the PACF show a definite cut off, then a mixed process is suggested [11]. In this case it is difficult to identify the AR and MA orders, but not impossible. The idea is that we should think of the ACF and PACF of pure AR and MA processes as being superimposed onto one another. For example, if both ACF and PACF show signs of slow exponential decay, then an ARMA(1,1) process may be identified. Similarly, if the ACF shows three significant spikes at first three lags and then an exponential decay, and the PACF spikes at the first lag and then shows an exponential decay, then an ARMA(3,1) process should be considered. Some possible combinations of ACF and PACF forms are shown in Table 2.1 (on page 38) which will be used to detect the order of ARMA processes [19],[30]. In general, it is difficult to identify mixed processes, so sometimes more than one ARMA(p,q) model might be estimated. This is why the estimation and the diagnostic checking stages are important and necessary.

2.4.2 Estimation

After identification of the tentative models, the next step is to estimate the parameters in the models [32]. In this stage each of the tentative models is estimated and the various coefficients are examined by a t -test. The common methods of estimations used in time series analysis are non-linear least squares estimation and maximum likelihood estimation [19], [32]. At this stage we have to be aware of the common factor problem. The Box-Jenkins approach necessitates that the series is stationary and the model is invertible.

2.4.3 Diagnostic checking

In the diagnostic stage we examine the goodness of fit of the model. The main objective is to check the adequacy of the model(s) selected in the estimation stage by checking whether the model assumptions are satisfied. The basic assumption is that the residuals u_t are white noise. If the fitted model is a good model for the data, the residuals should satisfy this assumption. Hence, model diagnostic checking is accomplished through a careful analysis of the residual series u_t [32]. To check whether the residuals are white noise, the sample ACF and sample PACF of the residuals are computed to see whether they do not form any pattern and are all statistically insignificant [32]. If the residual analysis indicates that the fitted model is inadequate in some way then we need to fit a more appropriate model. That is, we go back to the model identification step and try to develop a better model.

2.4.4 Model selection

After performing the three stages, there may be several adequate estimated models to represent the given time series. Sometimes the best choice is easy, other cases the correct choice can be very difficult. The estimated adequate models are compared using the Akaike Information Criterion (AIC) and Schwartz Bayesian Criterion (SBC). A model with smallest AIC and SBC values will be chosen as a parsimonious model. Of the two criteria, the SBC is preferable. Sometimes the adjusted- R^2 (Adj- R^2) is also compared to select the model.

2.4.5 Summary of the analysis procedure

Box-Jenkins methodology is an iterative procedure [32]. A typical procedure may be the following [30]:

- Step-1: Calculate the sample ACF and sample PACF of the given time series data and check the stationarity. For stationary series go to step-3 otherwise go to step-2.
- Step-2: For a non-stationary time series take the logarithm and the first differences. Calculate the sample ACF and sample PACF for the first logarithmic differenced series.
- Step-3: Examine the graphs(correlogram) of the sample ACF and sample PACF and determine which models would be good starting points.
- Step-4: Estimate all plausible models.

- Step-5: For each of the estimated models check the following to detect the adequate models.

- (a) estimated parameters and their orders;

- (b) sample ACF and sample PACF of the residuals.

If the estimated model found inadequate or if changes are needed, go back to step-4.

- Step-6: Compare all estimated adequate models by checking their AIC and SBC together with the $\text{Adj-}R^2$ to detect which model is the parsimonious one.

Chapter 3

Modelling the events time series

3.1 Introduction

A mathematical or statistical model, of course, is only an idealization of some process or activity in real life. The model tells us how we can expect the activity or process to most likely behave. But if we are examine any sample occurrence of this activity in real life, it would seldom turnout to be precisely what the model said it would be - probably close, but not the same. For example, flipping a fair coin 100 times will seldom produce exactly 50 tails and 50 heads, although that is expected to happen.

In a similar manner, a real life time series is only one sample of the process or activity that is represented by a given Box-Jenkins model. It follows that the autocorrelations computed from such a time series will never precisely match the theoretical autocorrelations associated with the correct model for the series. They will, however, be close to, or approximate, the theoretical autocorrelations. Autocorrelations computed from real life series called sample autocorrelations to dis-

tinguish them from the theoretical autocorrelations associated with a given model [21].

In this chapter Box-Jenkins modelling approach has described to identify the patterns of the acceleration time series corresponding to the field test events. The modelling has been described only for the first event using the acceleration time series obtained from the first counter clockwise lap. Following the similar modelling approach the rest events were identified and summarized in tables. The data analysis performed in this study by using computer software package Eviews-7.

3.2 Data preparation

As described in earlier, the acceleration time series used in this study to identify patterns for different events of interest. As shown in Figure 1.6(on page 17), X, Y and Z stands for representing the longitudinal accelerations time series, lateral accelerations time series and vertical accelerations time series respectively. Figure 3.1(on page 46) shows the acceleration time series for the 1st lap, where the horizontal axis is the time in the unit of *second*, and the vertical axis is the acceleration in the unit of *g*. Using visual inspection of the data along with the GPS information¹, the acceleration time series were initially separated for model identification. A similar visual identification method is described in [15].

¹GPS information used to detect the geographical position of the vehicle while travelling.

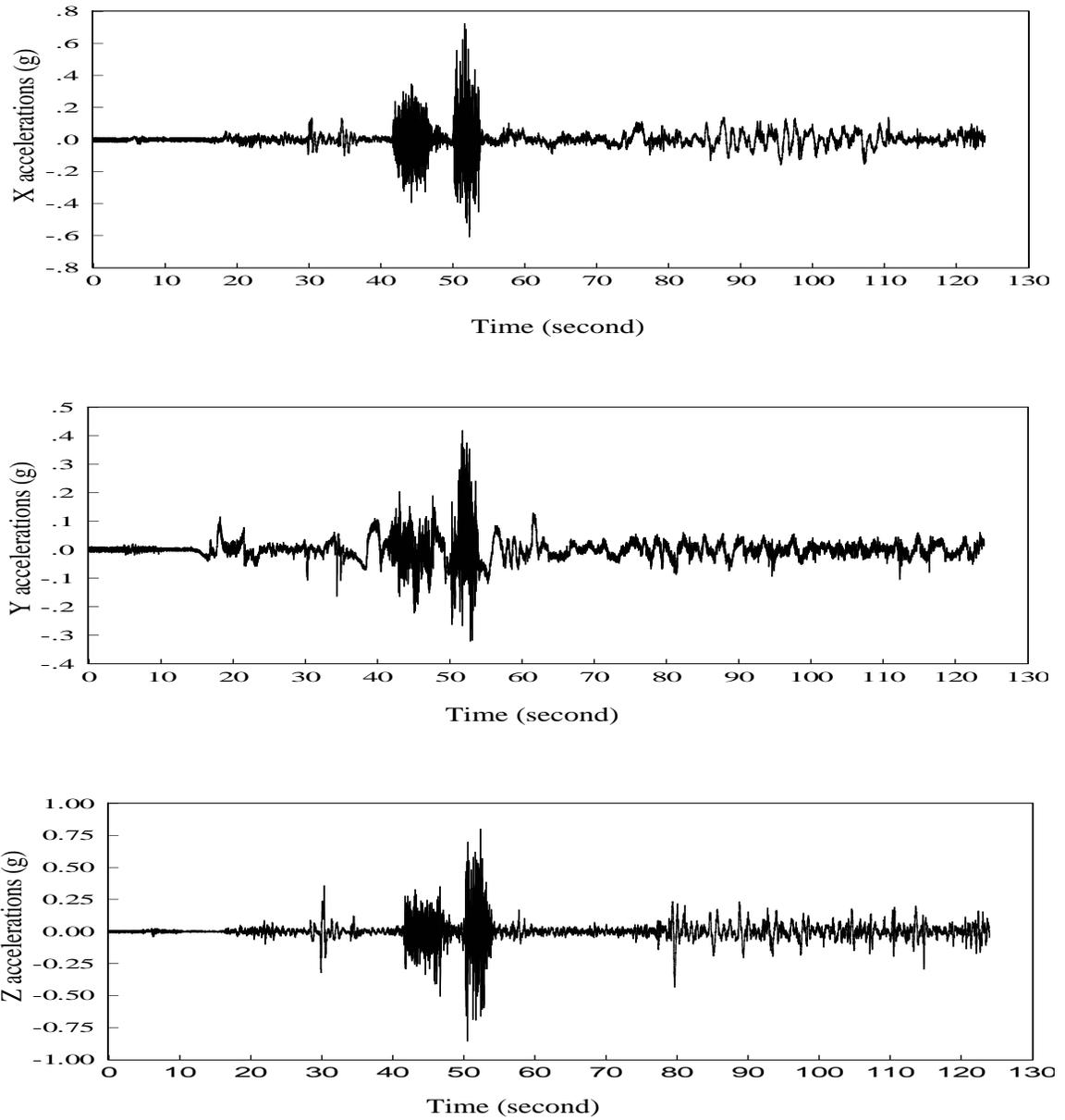


Figure 3.1: Time series of the acceleration data measured from the field test for the 1st lap [33]

As shown in Figure 1.3 (on page 12) and in Figure 3.1 (on page 46), the three

events are easily distinguishable (in Figure 3.1: in between 28th second to 56th second on Z-directional time series), where as the four events are not easily distinguishable (in Figure 3.1: in between 75th second to 120th second on Z-directional time series). Hence, the vertical accelerations time series (Z) was used to identify the patterns for the first three events and the complete time series set (X, Y and Z) was used to identify the rest four events.

3.3 Model identification of event-1: 4” Deep Chuck Hole

The steps of the Box-Jenkins approach on event-1: 4” Deep Chuck Hole are described in this section. To identify the model of event-1 from counter-clockwise first lap the vertical acceleration time series of the event was used. The steps are as described below.

3.3.1 Step-1

The time series plot of the vertical acceleration data for event-1 is shown in Figure 3.2 (on page 48), where, the horizontal axis is the time in the unit of *second*, and the vertical axis is the acceleration in the unit of *g*. Figure 3.2 shows that there is no indication of trend (specific upward or downward direction) in the data and we can assume that the given time series is approximately a stationary series. However, the augmented Dickey-Fuller (DF) unit root test [11] was carried out to test the stationarity of the time series data.

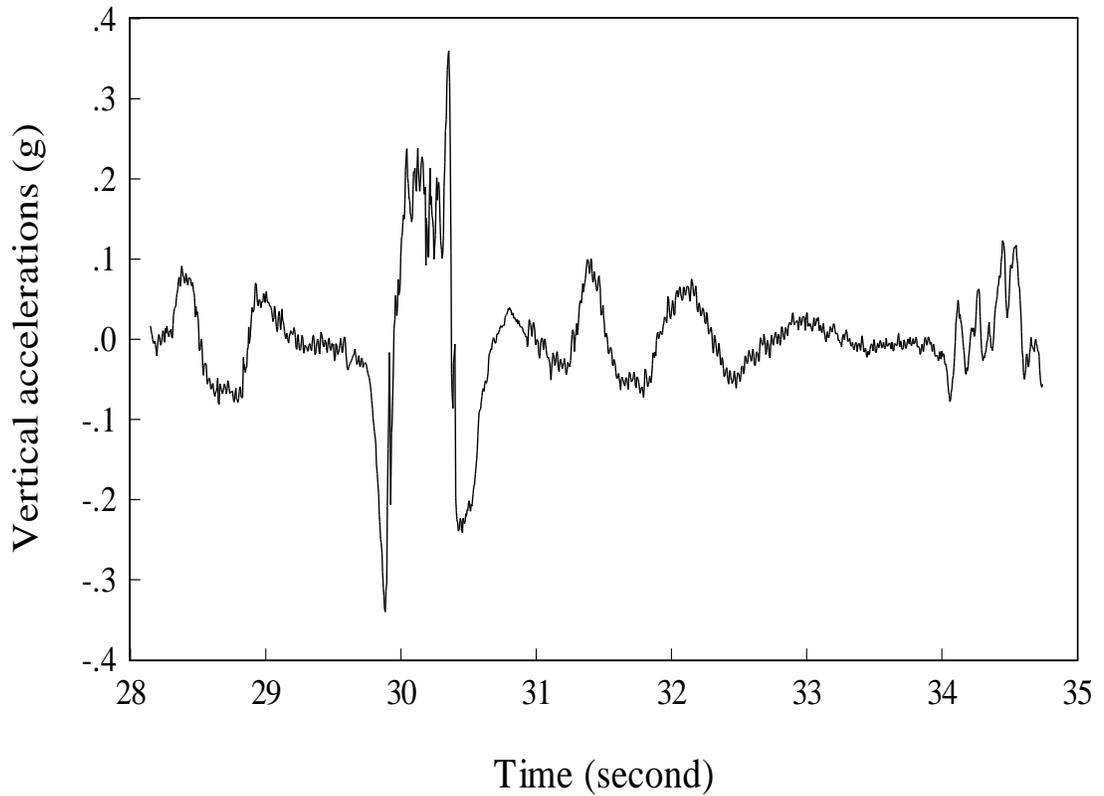


Figure 3.2: Vertical accelerations of event-1: 4" Deep Chuck Hole from counter clockwise first lap

Table 3.1 (on page 49) shows the summarized augmented Dickey-Fuller (DF) unit root test results. The test statistic values are smaller than the test critical values for different level of significance. From this test results it can be concluded that the given time series is stationary. Hence, ARMA modelling was considered for the given time series and Box-Jenkins approach was applied to identify model of the corresponding event.

Table 3.1: Summary of the augmented Dickey-Fuller (DF) unit root test for stationarity

Null Hypothesis: the series has a unit root Lag Length:1 (Automatic: based on SIC, maxlag=24)	Exogenous: None	Exogenous: Constant	Exogenous: Constant, Linear Trend
Augmented Dickey-Fuller test statistic	-4.914548	-4.912234	-4.908741
*MacKinnon (1996) one-sided p-values	0.0000	0.0000	0.0003
Test critical values:			
1% level	-2.566714	-3.45089	-3.965007
5% level	-1.941063	-2.863520	-3.413216
10% level	-1.616538	-2.567874	-3.128628
Durbin-Watson statistic	1.986960	1.986957	1.986957

3.3.2 Step-2

As the given time series found stationary, this step can be skipped.

3.3.3 Step-3

The sample ACF and sample PACF of the time series for event-1 were calculated. Figure 3.3 (on page 50) shows the correlogram and three statistics : (i) the AC (autocorrelation coefficient), (ii) the PAC (partial autocorrelation coefficient) and (iii) a Box-Pierce Q -statistic with its probability. The two dot lines in the graphs

of autocorrelation and partial correlation represents the approximate confidence bounds. That is, any value which is beyond these lines will be regarded as non-zero.

Correlogram of of Event-1: 4" Deep Chuck Hole

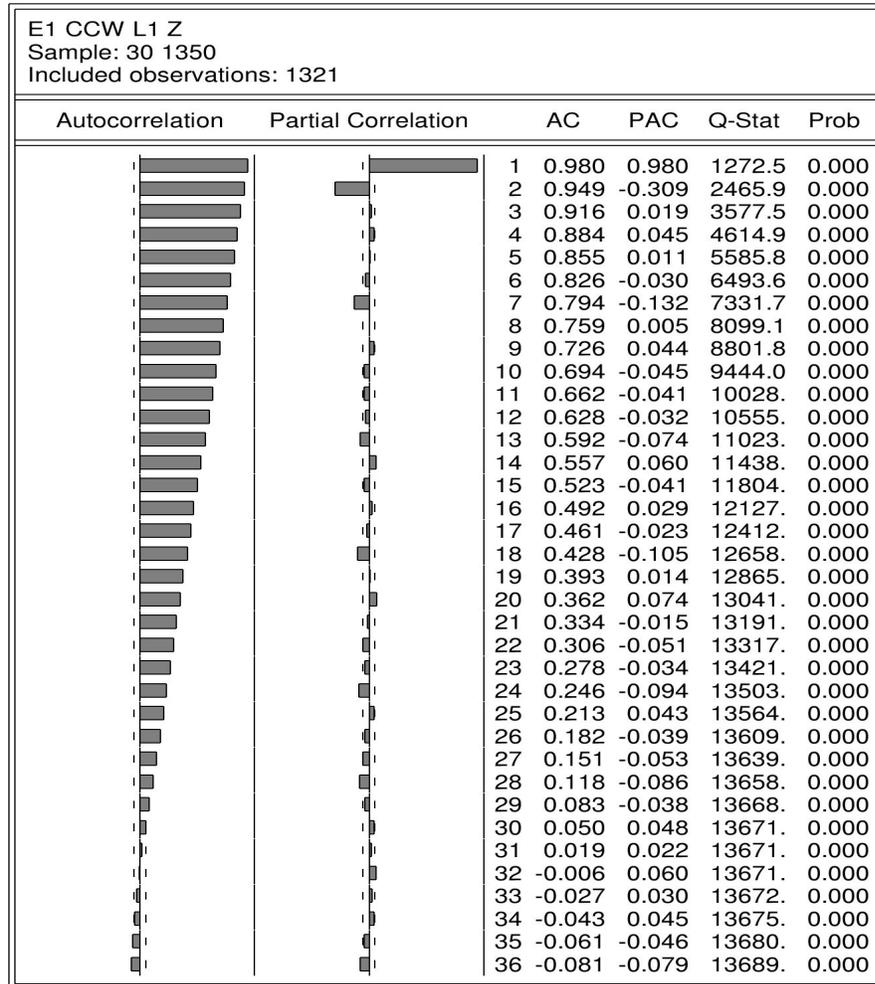


Figure 3.3: Correlogram of the vertical acceleration time series data of event-1

The correlogram in Figure 3.3 shows that the sample ACF has a pattern of moderately slow decay and the sample PACF has a few significant spikes at lags one, two and seven. The sample ACF and sample PACF patterns guided us initially

to choose possible candidate models of the time series data of event-1.

The decreasing pattern of the sample ACF (as shown in Figure 3.3) indicates that the series may follow an autoregressive model. The significant spikes of the sample PACF at lag-2 and lag-7 guide to choose the order of the autoregressive model. Hence, the possible models for event-1 are AR(7) and AR(2). Based on the slowly dying pattern of the sample ACF after lag-2, an ARMA(2,2) model was also chosen as a possible model for event-1.

3.3.4 Step-4 and Step-5

For each possible model of the event: 4" Deep Single Chuck Hole, model parameters were estimated and diagnostic checking were performed to verify their adequacy as a candidate model. The estimation and checking procedure are described below.

AR(7) model

The estimation steps began with estimating the AR(7) model. The estimation results in Table 3.2 (on page 52), shows that the t -statistic scores of the coefficients for autoregressive orders 3, 4 and 5 are -0.742207, 0.413892 and 0.158448 respectively, are less 2, implies that the coefficients for autoregressive orders 3, 4 and 5 are insignificant to be included in the model. That is, this AR(7) model is inadequate and need to be modified by reducing the three coefficients.

Table 3.2: Estimation results of an AR(7) model

E1 CCW L1 Z				
Method:	Least Squares			
Sample:	30 1350			
Included observations:	1321			
Convergence achieved	after 3 iterations			
<u>Variable</u>	<u>Coefficient</u>	<u>Std. Error</u>	<u>t-Statistic</u>	<u>Prob.</u>
C	-0.000586	0.013819	-0.042403	0.9662
AR(1)	1.287067	0.027353	47.05408	0.0000
AR(2)	-0.313600	0.044666	-7.021050	0.0000
AR(3)	-0.033766	0.045494	-0.742207	0.4581
AR(4)	0.018833	0.045503	0.413892	0.6790
AR(5)	0.007209	0.045496	0.158448	0.8741
AR(6)	0.139730	0.044671	3.127946	0.0018
AR(7)	-0.132995	0.027367	-4.859629	0.0000
R-squared	0.965969	Mean dependent var		0.000489
Adjusted R-squared	0.965788	S.D. dependent var		0.074719
S.E. of regression	0.013820	Akaike info criterion		-5.719300
Sum squared resid	0.250789	Schwarz criterion		-5.687892
Log likelihood	3785.597	Hannan-Quinn criter.		-5.707525
F-statistic	5324.230	Durbin-Watson stat		1.997689
Prob(F-statistic)	0.000000			
Inverted AR Roots	.95	.78	.43-.62i	.43+.62i
	-.31-.61i	-.31+.61i		-.68

Since the AR(7) model found inadequate, a reduced AR(7) model that excluded the three insignificant autoregressive orders was considered for modelling. The model parameters were estimated in the similar manner and the estimation results are shown in Table 3.3 (on page 53). As shown in Table 3.3, the t -statistic scores

indicate that all the coefficients are adequate to be in the model.

Table 3.3: Estimation results of a reduced AR(7) model

<u>Variable</u>	<u>Coefficient</u>	<u>Std. Error</u>	<u>t-Statistic</u>	<u>Prob.</u>
E1 CCW L1 Z				
Method:	Least Squares			
Sample:	30 1350			
Included observations:	1321			
Convergence achieved	after 3 iterations			
C	-0.000583	0.013802	-0.042213	0.9663
AR(1)	1.290262	0.026145	49.34986	0.0000
AR(2)	-0.332351	0.029089	-11.42533	0.0000
AR(6)	0.150001	0.029093	5.155909	0.0000
AR(7)	-0.135441	0.026162	-5.177048	0.0000
R-squared	0.965953	Mean dependent var		0.000489
Adjusted R-squared	0.965849	S.D. dependent var		0.074719
S.E. of regression	0.013808	Akaike info criterion		-5.723363
Sum squared resid	0.250909	Schwarz criterion		-5.703733
Log likelihood	3785.281	Hannan-Quinn criter.		-5.716003
F-statistic	9334.063	Durbin-Watson stat		2.004505
Prob(F-statistic)	0.000000			
Inverted AR Roots	.95 -.31-.62i	.78 -.31+.62i	.43-.61i -.68	.43+.61i

The diagnostic check of a time series model is based on the analysis of residuals. As shown in Figure 3.4 (on page 55), the residual time plot of the reduced-AR(7) model shows no specific pattern. The residual plot also shows that, it has a mean close to zero and a constant variance, which approximately satisfy the model assumptions. To check the model adequacy, the ACF and PACF of the residuals were also calculated and verified. For the reduced-AR(7) model, Figure 3.5 (on page 55) shows the residual correlogram and three statistics:(i) the AC (autocorrelation coefficient), (ii) the PAC (partial autocorrelation coefficient) and (iii) a Ljung-Box Q -statistic with its probability. The two dot lines in the graphs of autocorrelation and partial correlation represents the approximate two standard error bounds computed as $\pm 2/\sqrt{T}$, where, T is the number of observations [8]. The residual ACF and PACF strengthen the model assumptions, that the residuals mimic a white noise process. Hence, the fitted reduced-AR(7) model can be considered as an adequate model which is given below.

$$Z_t = -0.000583 + 1.29Z_{t-1} - 0.33Z_{t-2} + 0.15Z_{t-6} - 0.135Z_{t-7}$$

AR(2) model

The next possible model AR(2) was estimated in a similar manner. The estimation results in Table 3.4 (on page 56), shows the t -statistic scores of the the coefficients for autoregressive orders 1 and 2 are 49.13830 and -11.89967 respectively. The t -statistic scores for both the coefficients are greater than 2, which implies that the coefficients are adequate to be in the model.

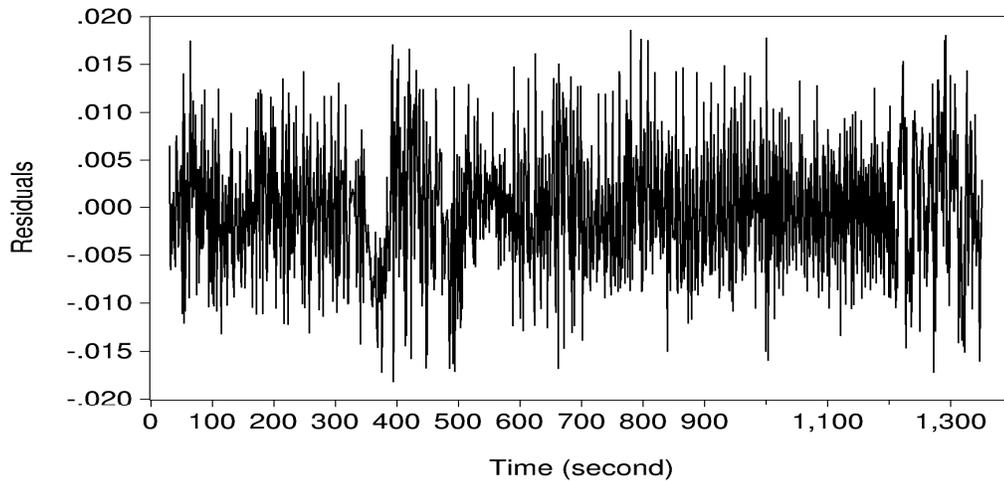


Figure 3.4: Residuals plot of the reduced AR(7) model

Correlogram of Residuals of AR(7) model

E1 CCW L1 Z						
Sample: 30 1350						
Included observations: 1321						
Q-statistic probabilities adjusted for 4 ARMA term(s)						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.003	-0.003	0.0089	
		2	0.021	0.021	0.5834	
		3	-0.019	-0.019	1.0748	
		4	-0.011	-0.012	1.2419	
		5	0.000	0.001	1.2420	0.265
		6	-0.009	-0.009	1.3442	0.511
		7	0.017	0.016	1.7263	0.631
		8	-0.052	-0.052	5.3123	0.257
		9	0.022	0.021	5.9584	0.310
		10	0.032	0.035	7.3478	0.290
		11	0.004	0.001	7.3654	0.392
		12	0.093	0.091	18.836	0.016
		13	-0.037	-0.035	20.640	0.014
		14	0.028	0.025	21.709	0.017
		15	-0.046	-0.040	24.592	0.010
		16	-0.026	-0.030	25.476	0.013
		17	0.080	0.085	34.073	0.001
		18	0.042	0.046	36.451	0.001
		19	-0.070	-0.081	42.957	0.000
		20	-0.047	-0.038	45.957	0.000

Figure 3.5: Residuals correlogram of the reduced AR(7) model

To check the adequacy of the fitted AR(2) model, the residuals time plot was

Table 3.4: Estimation results of an AR(2) model

<u>Variable</u>	<u>Coefficient</u>	<u>Std. Error</u>	<u>t-Statistic</u>	<u>Prob.</u>
E1 CCW L1 Z				
Method:	Least Squares			
Sample:	30 1350			
Included observations:	1321			
Convergence achieved	after 3 iterations			
C	-0.000632	0.015128	-0.041777	0.9667
AR(1)	1.286189	0.026175	49.13830	0.0000
AR(2)	-0.311550	0.026181	-11.89967	0.0000
R-squared	0.965235	Mean dependent var		0.000489
Adjusted R-squared	0.965183	S.D. dependent var		0.074719
S.E. of regression	0.013942	Akaike info criterion		-5.705538
Sum squared resid	0.256196	Schwarz criterion		-5.693760
Log likelihood	3771.508	Hannan-Quinn criter.		-5.701122
F-statistic	18297.06	Durbin-Watson stat		1.986957
Prob(F-statistic)	0.000000			
Inverted AR Roots	.96	.32		

checked (as shown in Figure 3.6 on page 57) and the ACF and PACF of the residuals were calculated and verified. The residual plot shows no patterns and has a mean close to zero and a constant variance, which approximately satisfy the model assumptions. Again, for the AR(2) model, Figure 3.7 (on page 57) shows the residual correlogram and three statistics: (i) the AC (autocorrelation coefficient), (ii) the PAC (partial autocorrelation coefficient) and (iii) a Ljung-Box Q -statistic with its probability. The residual ACF and PACF (in Figure 3.7) indicates that the residuals mimic a white noise process. Hence, the fitted AR(2) model can be considered

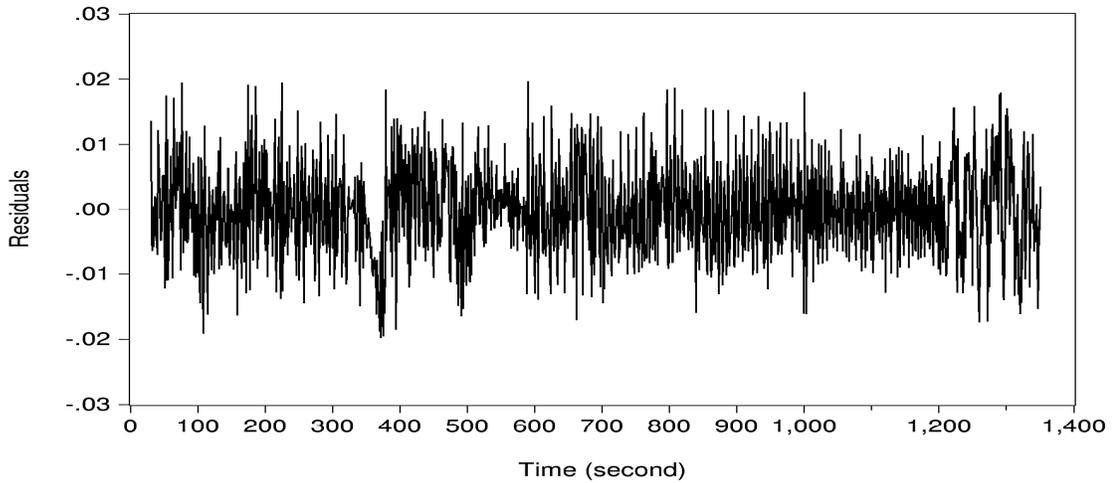


Figure 3.6: Residuals plot of AR(2) model

Correlogram of Residuals of AR(2) model

E1 CCW L1 Z						
Sample: 30 1350						
Included observations: 1321						
Q-statistic probabilities adjusted for 2 ARMA term(s)						
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.006	0.006	0.0498	
		2	-0.003	-0.003	0.0614	
		3	-0.042	-0.042	2.4061	0.121
		4	-0.029	-0.029	3.5557	0.169
		5	-0.016	-0.016	3.8850	0.274
		6	0.131	0.129	26.621	0.000
		7	0.047	0.044	29.608	0.000
		8	-0.044	-0.047	32.152	0.000
		9	0.005	0.015	32.192	0.000
		10	0.021	0.032	32.766	0.000
		11	0.007	0.010	32.831	0.000
		12	0.102	0.086	46.616	0.000
		13	-0.042	-0.055	48.934	0.000
		14	0.012	0.025	49.141	0.000
		15	-0.048	-0.039	52.216	0.000
		16	-0.023	-0.031	52.896	0.000
		17	0.091	0.094	63.926	0.000
		18	0.048	0.021	67.073	0.000
		19	-0.078	-0.084	75.308	0.000
		20	-0.049	-0.038	78.474	0.000

Figure 3.7: Residuals correlogram of model AR(2)

as another adequate model which is given below.

$$Z_t = -0.000632 + 1.286Z_{t-1} - 0.311Z_{t-2}$$

Table 3.5: Estimation results of an ARMA(2,2) model

E1 CCW L1 Z				
Method:	Least Squares			
Sample:	30 1350			
Included observations:	1321			
Convergence achieved	after 11	iterations		
MA Backcast:	28 29			
<u>Variable</u>	<u>Coefficient</u>	<u>Std. Error</u>	<u>t-Statistic</u>	<u>Prob.</u>
C	-0.000664	0.015612	-0.042505	0.9661
AR(1)	0.962580	0.276866	3.476707	0.0005
AR(2)	0.002304	0.269023	0.008564	0.9932
MA(1)	0.328577	0.275528	1.192537	0.2333
MA(2)	0.100811	0.084786	1.189012	0.2346
R-squared	0.965296	Mean dependent var		0.000489
Adjusted R-squared	0.965190	S.D. dependent var		0.074719
S.E. of regression	0.013941	Akaike info criterion		-5.704253
Sum squared resid	0.255750	Schwarz criterion		-5.684623
Log likelihood	3772.659	Hannan-Quinn criter.		-5.696894
F-statistic	9151.159	Durbin-Watson stat		1.999181
Prob(F-statistic)	0.000000			
Inverted AR Roots	.96	-.00		
Inverted MA Roots	-.16-.27i	-.16+.27i		

ARMA(2,2) model

Lastly, the parameters of the third possible model ARMA(2,2) was estimated. The estimation results in Table 3.5 (on page 58), shows that the t -statistic scores of the coefficient for autoregressive orders 2 and the coefficients for moving-average orders 1 and 2 are 0.008564, 1.192537 and 1.189012 respectively. The t -statistic scores for the three coefficients are less 2, which is an indication that the coefficients are insignificant to be included in the model. That is, the ARMA(2,2) model is inadequate to be a candidate model for the event-1.

3.3.5 Step-6

After estimation and diagnostic checking, there was two adequate candidate models corresponding to the vertical acceleration time series of the event-1. The final model was selected based on the model selection tools mentioned in chapter-2. Model selection criteria was compared for both two adequate models reduced-AR(7) and AR(2). The summarized Table 3.6 (on page 59) shows the estimated values of the coefficients along with their t -statistic scores, AIC, SBC and Adjusted- R^2 of the two models.

Table 3.6: Summery results of the reduced AR(7) and AR(2)

	AR(1,2,6,7)*	AR(1,2)
ϕ_1	1.290 (49.35)**	1.286 (49.14)
ϕ_2	-0.332(-11.43)	-0.312 (-11.899)
ϕ_6	0.150 (5.155)	
ϕ_7	-0.135 (-5.177)	
AIC	-5.723	-5.705
SBC	-5.704	-5.693
Adjusted- R^2	0.9658	0.9651

*reduced AR(7) model

** t -statistic

As shown in Table 3.6, the reduced-AR(7) model has smaller AIC and SBC

compared to those of AR(2) model. From this comparison it can be concluded that, to identify the patterns of the vertical acceleration time series of event-1 the reduced-AR(7) model is more convincing model than AR(2) model.

3.4 Fitted models and event identification procedure

The model fitting steps just described in earlier subsection were followed to fit models for the rest of the events using acceleration time series of different directions from different driving directions of the vehicle.

Vertical acceleration time series of the first two counter clockwise laps have used to fit models for each of the seven events. From the analyses it is found that event-4, event-5 and event-6 have the same fitted models. Then the longitudinal acceleration time series of these three events were fitted and found that event-4 and event-5 have the same fitted models. Finally, using the lateral acceleration time series of event-4 and event-5 two different fitted models were found for them. Table 3.7(on page 61) shows the fitted models corresponding to the events and acceleration time series.

Based on the fitted models in Table 3.7, an event identification procedure was proposed to identify models for the seven events of the acceleration time series collected from the counter clockwise driving direction. The proposed identification procedure was developed using the first two laps acceleration time series of counter clockwise driving direction and its accuracy was tested on acceleration time series collected from the third lap of the same driving direction. The proposed event identification procedure is shown in Figure 3.8 (on page 62).

Table 3.7: Fitted models for all seven events from counter clockwise laps

Event	Z-directional Time Series	Y-directional Time Series	X-directional Time Series
Event-1	AR(7)		
Event-2	AR(1)		
Event-3	AR(5)		
Event-4	AR(14)	AR(15)	AR(10)
Event-5	AR(14)	AR(15)	AR(6)
Event-6	AR(14)	AR(14)	
Event-7	AR(12)		

As shown in Figure 3.8, the proposed identification procedure for counter clockwise time series has three stages, the first stage will begin with the modelling an AR(14) model using Z-directional time series. The second stage need to be proceeded depending on the significance of the coefficient corresponding to order 14 of the fitted AR(14) model. Otherwise, start fitting an AR(12) model and event-7 will be reported as identified based on significantly fitted AR(12) model. The first stage will be continued until a significantly fitted model for event-2 is identified.

The second stage will begin with the modelling of an AR(15) model using Y-directional time series. Again, the third stage need to be proceeded depending on the significance of the coefficient corresponding to order 15 of the fitted AR(15)

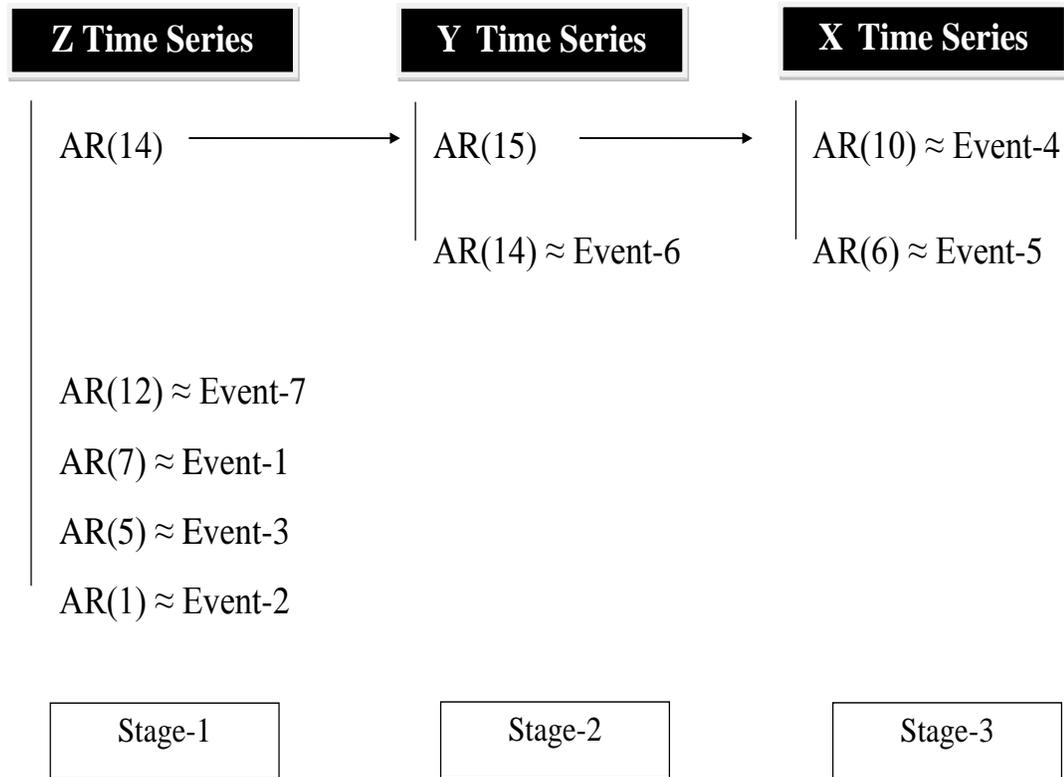


Figure 3.8: Event identification procedure for counter-clockwise time series

model. Otherwise, start modelling an AR(14) model and event-6 will be reported as identified based on significantly fitted AR(14) model.

Similarly, the third stage involves start modelling of an AR(10) model using X-directional time series. For significantly fitted AR(10) model, event-4 will be reported as identified by the fitted model. Otherwise, an AR(6) model need to be significantly fitted to identify event-5.

Based on the fitted models of the clockwise directional time series in Table 3.8(on

page 63), another event identification procedure was also proposed. As shown in Figure 3.9, this procedure also has three stages to identify models for events of the clockwise directional time series. Again, the accuracy of the identification procedure was tested on acceleration time series collected from the third lap of clockwise driving direction.

Table 3.8: Fitted models for all seven events from clockwise laps

Event	Z-directional Time Series	Y-directional Time Series	X-directional Time Series
Event-1	AR(7)		
Event-2	AR(1)		
Event-3	AR(5)		
Event-4	AR(14)	AR(15)	
Event-5	AR(14)	AR(11)	AR(14)
Event-6	AR(14)	AR(11)	AR(11)
Event-7	AR(12)		

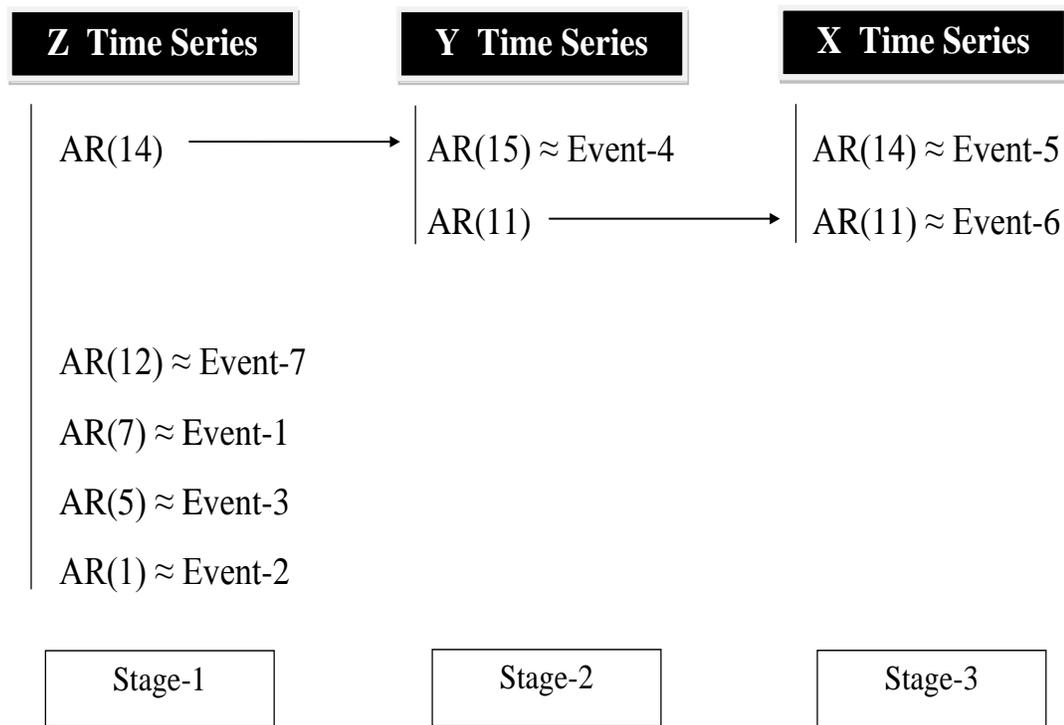


Figure 3.9: Event identification procedure for clockwise time series

Chapter 4

Results

Results of the event identification has presented in this chapter. The start and end time points of each event was found while fitting models for corresponding time series. Time series of the first two laps from each driving direction was used for modelling and the time series of the third laps was used to test the accuracy of the event identification procedure.

Initial start and end times of the seven events are the start and the end points of the segmented time series of the corresponding events. While doing the modelling the actual length of the time series of an event was used to identify the start and the end time points for that event. The seven events used in this study are already shown in Table 1.1. It is to be noted that the Z-directional (vertical acceleration) time series was used to identify the first three events and all the three directional (Z, Y and X) time series were used to identify the rest four events. The identified start and end points of all seven events of counter-clockwise and clockwise laps are shown from Table 4.1 to Table 4.4.

Table 4.1: Identified start and end time for the three events of counter-clockwise laps using the acceleration data in the Z direction

Events	Transition	Lap-1 (s)	Lap-2 (s)	Lap-3 (s)
Event-1	Start	28.20	148.00	264.20
	End	34.80	154.00	270.49
Event-2	Start	41.50	158.20	274.20
	End	47.50	164.20	280.40
Event-3	Start	50.12	166.97	282.40
	End	54.07	170.69	286.92

As mentioned earlier, the first three events were easy to distinguish and the rest four events were hard to distinguish. Table 4.1 shows the identified start and end time points of the three events of counter-clockwise laps based on vertical acceleration time series. Table 4.2 shows the successfully identified start and end points of the four events of counter-clockwise laps based on the three acceleration time series.

Table 4.2: Identified start and end time for the four events of counter-clockwise laps using the acceleration data in the Z, Y and X directions

Events	Transition	Lap-1(s)			Lap-2(s)			Lap-3(s)		
		Z	Y	X	Z	Y	X	Z	Y	X
Event-4	Start	79.22	79.23	78.99	196.4	196.5	196.00	311.6	311.7	311.0
	End	83.97	84.01	83.89	201.2	201.6	202.00	316.4	316.5	316.0
Event-5	Start	84.50	84.60	85.10	202.3	205.7	202.80	317.1	317.5	317.9
	End	98.80	99.50	99.20	216.2	216.9	216.70	331.1	332.4	331.6
Event-6	Start	100.80	102.60	101.50	219.1	220.7	219.50	334.5	333.7	334.9
	End	110.90	111.60	111.20	228.7	229.1	228.30	343.5	344.8	343.5
Event-7	Start	113.10	112.60	112.90	230.8	231.2	231.60	345.9	345.5	346.8
	End	118.00	118.40	118.35	234.4	235.4	234.70	349.6	350.5	350.6

Table 4.3 shows the identified start and end time points of the three events of clockwise laps based on vertical acceleration time series and Table 4.4 shows the identified start and end time points of the four events of clockwise laps based on the three acceleration time series.

Table 4.3: Identified start and end time for the three events of clockwise laps using the acceleration data in the Z direction

Events	Transition	Lap-1 (s)	Lap-2 (s)	Lap-3 (s)
Event-1	Start	467.70	575.80	679.90
	End	472.50	581.80	685.50
Event-2	Start	456.70	566.20	671.20
	End	462.70	572.30	676.90
Event-3	Start	450.60	560.00	664.80
	End	455.20	564.30	669.10

Table 4.4: Identified start and end time for the four events of clockwise laps using the acceleration data in the Z, Y and X directions

Events	Transition	Lap-1(s)			Lap-2(s)			Lap-3(s)		
		Z	Y	X	Z	Y	X	Z	Y	X
Event-4	Start	423.2	422.7	423.4	534.55	534.83	533.99	639.8	639.00	638.7
	End	426.6	426.9	426.5	537.24	537.97	537.57	643.1	643.30	642.6
Event-5	Start	406.5	405.8	408.1	517.60	517.40	519.70	623.3	622.80	624.3
	End	420.3	419.6	420.5	531.40	531.80	531.80	636.9	636.60	637.3
Event-6	Start	392.9	394.0	396.7	504.20	504.70	508.10	610.1	613.00	612.7
	End	402.6	402.8	404.3	514.00	513.90	514.60	620.0	621.60	621.4
Event-7	Start	388.5	388.6	388.5	499.80	500.00	499.40	606.1	606.60	606.3
	End	391.6	392.2	391.9	502.30	503.30	503.20	609.5	609.20	609.7

Identified event's durations for counter clockwise and clockwise laps are shown in Table 4.5 to Table 4.8. These tables show that the durations of the seven identified events for three different laps are approximately close, which has strengthened the accuracy of the proposed identification procedure.

Table 4.5: Identified duration for the three events of counter-clockwise laps using the acceleration data in the Z direction

Z Direction			
	Lap-1 (s)	Lap-2 (s)	Lap-3 (s)
Event-1	6	6	6.29
Event-2	6	6	6.2
Event-3	3.95	3.72	4.52

Table 4.6: Identified duration for the four events of counter-clockwise laps using the acceleration data in the Z, Y and X directions

Events	Lap-1(s)			Lap-2(s)			Lap-3(s)		
	Z	Y	X	Z	Y	X	Z	Y	X
Event-4	4.75	4.78	4.9	4.8	5.1	6	4.8	4.8	5
Event-5	14.3	14.90	14.1	13.9	11.2	13.9	14	14.9	13.7
Event-6	10.1	9.00	9.7	9.6	8.4	8.8	9	11.1	8.6
Event-7	4.9	5.80	5.45	3.6	4.2	3.1	3.7	5.1	3.8

Table 4.7: Identified duration for the three events of clockwise laps using the acceleration data in the Z direction

Z Direction			
	Lap-1 (s)	Lap-2 (s)	Lap-3 (s)
Event-1	4.8	6	5.6
Event-2	6	6.1	5.7
Event-3	4.6	4.3	4.3

Table 4.8: Identified duration for the four events of clockwise laps using the acceleration data in the Z, Y and X directions

Events	Lap-1(s)			Lap-2(s)			Lap-3(s)		
	Z	Y	X	Z	Y	X	Z	Y	X
Event-4	3.4	4.2	3.1	2.69	3.14	3.58	3.3	4.3	3.9
Event-5	13.8	13.8	12.4	13.8	14.4	12.1	13.6	13.8	13
Event-6	9.7	8.8	7.6	9.8	9.2	6.5	9.9	8.6	8.7
Event-7	3.1	3.6	3.4	2.5	3.3	3.8	3.4	2.6	3.4

Identified start and end points of the seven events obtained from the first laps of each driving directions were plotted to describe the results visually. Figure 4.1 and in Figure 4.2 show the identified three events and four events respectively from counter clockwise first lap. Similarly, Figure 4.3 and Figure 4.4 show the identified three events and four events respectively from clockwise first lap. In these figures, the horizontal axis is the time in the unit of second and the vertical axis is the accelerations in the unit of g .

As shown in Figure 4.1, the identified three events from the 1st counter-clockwise lap in Z direction are: 4" Deep Chuck Hole, (3/4)" Chatter Bumps and 1" Random Chuck Holes respectively. The first event 4" Deep Chuck Hole started at 28.20

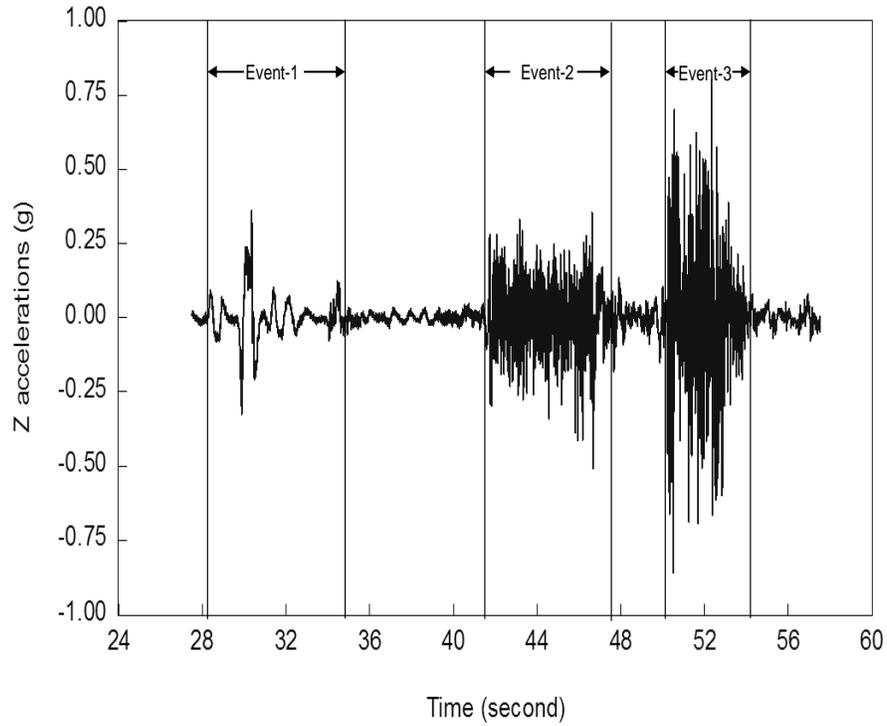


Figure 4.1: Identified three events from the 1st counter-clockwise lap in Z direction second and ended at 34.80 second, the second event (3/4)" Chatter Bumps started at 41.50 second and ended at 47.50 second and the third event 1" Random Chuck Holes started at 50.12 second and ended at 54.07 second.

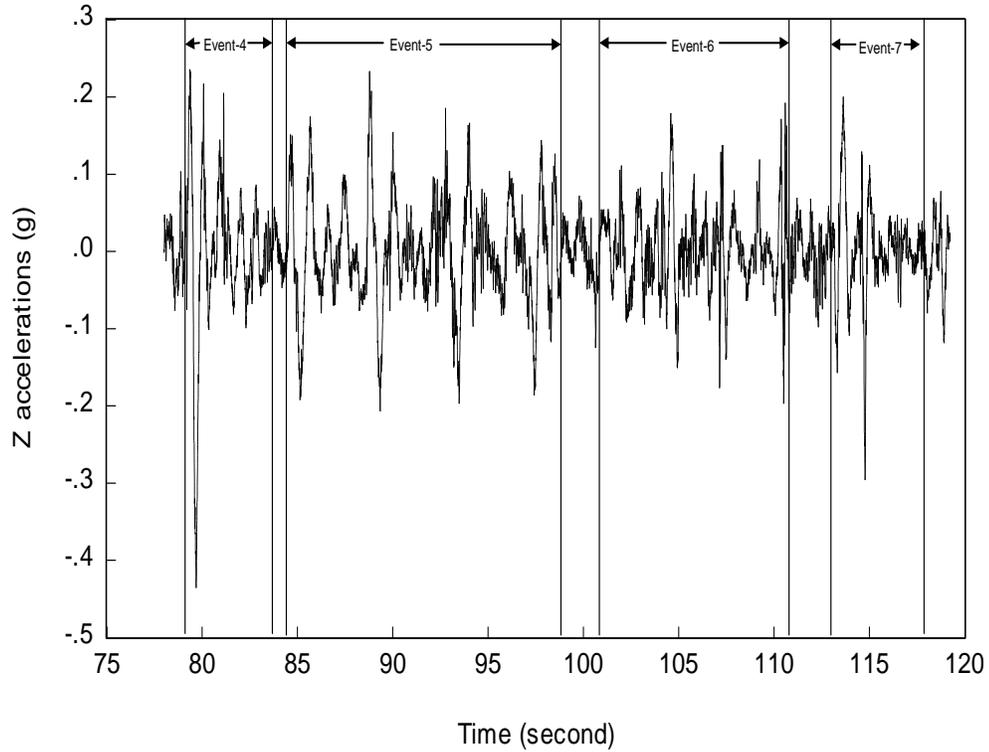


Figure 4.2: Identified four events from the 1st counter-clockwise lap in Z direction

Figure 4.2 shows the identified four events from the 1st counter-clockwise lap in Z direction. The fourth event Rail Road Crossing started at 79.22 second and ended at 83.97 second, the fifth event Staggered Bumps started at 84.50 second and ended at 98.80 second, the sixth event Frame twist started at 100.80 second and ended at 110.90 second and the seventh event High Crown Intersection started at 113.10 second and ended at 118.00 second.

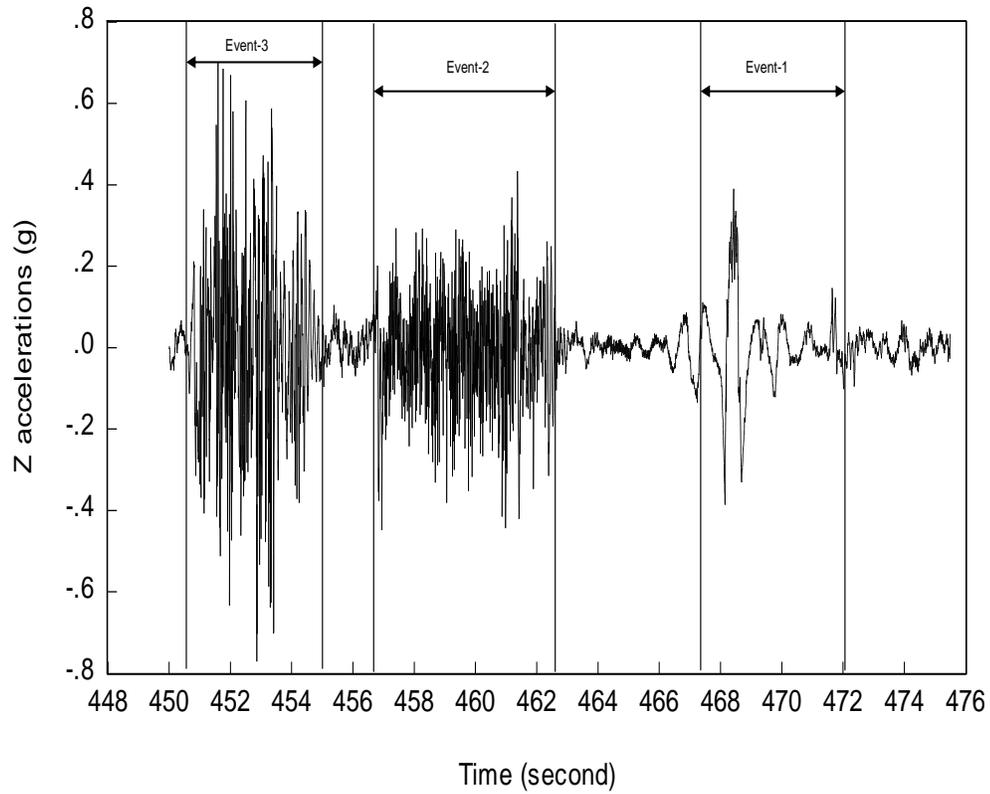


Figure 4.3: Identified three events from the 1st clockwise lap in Z direction

Identified three events and four events from the 1st clockwise lap in Z direction are shown in Figure 4.3 and in Figure 4.4 respectively. Contrary to counter-clockwise lap, the vehicle was encountered the event-7 at the beginning of clockwise lap and encountered the event-1 at the end of the clockwise lap.

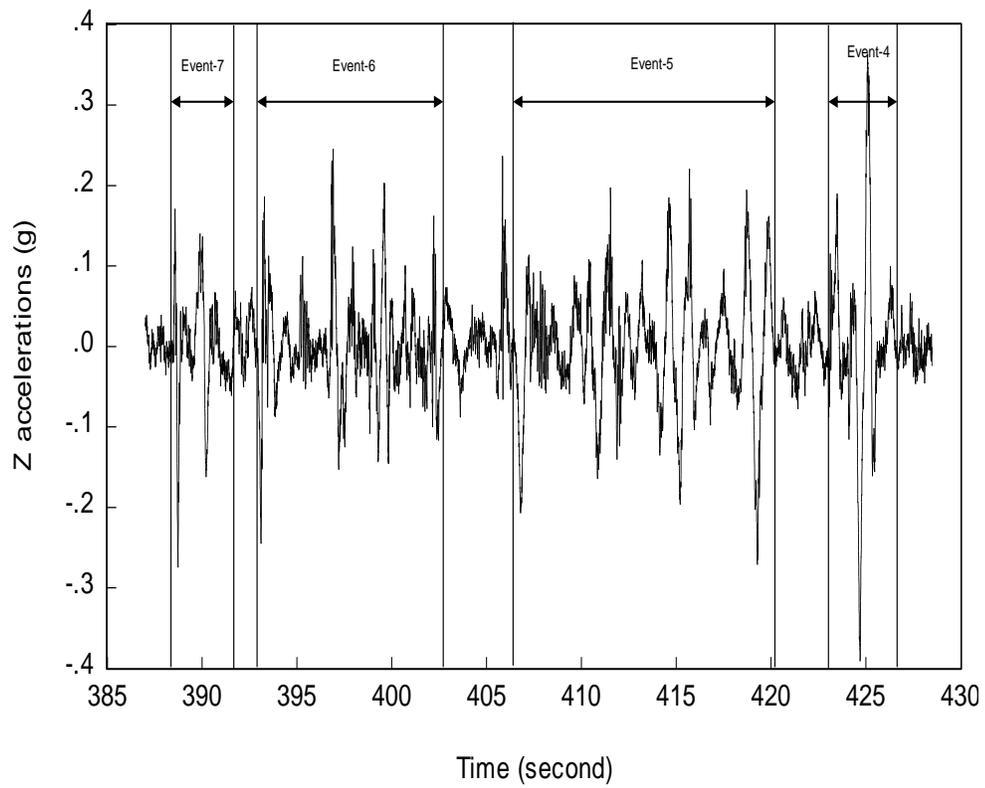


Figure 4.4: Identified four events from the 1st clockwise lap in Z direction

Chapter 5

Conclusion

Durability tests are important for measuring the characteristics and longevity of a ground vehicle. In laboratory, the MAST is used for carrying out the sub-scaled accelerated durability test. The accuracy of the accelerated durability tests heavily depends on the accurately generated mission profiles using the acceleration time series of the events experienced by the vehicle during the field test. Therefore, accurately separated acceleration data can efficiently generate accurate mission profiles for the accelerated test analysis.

Box-Jenkins methodology based event identification procedures have been proposed to identify the events for both the counter-clockwise and clockwise laps. Both the procedures identified the events efficiently and accurately.

It is already mentioned that the field test time series are noisy and of high frequency. Wavelet based dnoised data was used to fit models for event's time series using Box-Jenkins approach but results were not satisfactory. Hence, the original data (not dnoised) was used to identify events. Therefore, a chance of misclassification of the events is always exist, though the proposed identification procedure has

been tested on each of the driving directional time series. However, accuracy of the proposed procedures depend on the quality of the collected acceleration time series from the field tests, that is, the procedure will give the best result for completely noise free field test acceleration time series. In reality, complete noise free accelerations may not be obtainable but the noise can be reduced by controlling the driving speed of the vehicle.

As a limitation of this study, it can be said that the fitted models are sensitive to the data, that is the fitted models may changes with different acceleration time series recorded from different types of vehicle. The probability of misclassification of events using the proposed procedures may be obtainable but time consuming and costly as new set of data are needed from the test track.

Use of multivariate time series analysis and suitable noise filtering method, a simpler event identification procedure can be developed. Then the accuracy of that procedure can also be verified by using the upcoming acceleration data obtained from the field tests.

Bibliography

- [1] *Applying technology for industry*, Slides from MITACS. (Cited on pages 3 and 4.)
- [2] *An introduction to multi-axis simulation tables (mast)*, Slides from MITACS. (Cited on page 3.)
- [3] *Glyphworks worked examples 5.1*, Electronic, 2007. (Cited on pages 3 and 9.)
- [4] *Westest's mast table*, <http://www.westest.ca/mast.htm>, Feb, 2008, Online; accessed 12-January-2012. (Cited on pages viii and 4.)
- [5] *Bus research and testing facility (test track)*, http://www.vss.psu.edu/BTRC/btrc_test_track.htm, Nov, 2007, Online; accessed 12-January-2012. (Cited on pages viii, 2 and 12.)
- [6] Mobarakeh A.A., Rofooei F.R., and Ahmadi G., *Simulation of earthquake records using time-varying arma(2,1) model*, Probabilistic Engineering Mechanics **17** (2002), no. n/a, 15–34. (Cited on page 10.)
- [7] S.M.F. Aghda, A.J. Gandomi, and A. Beitollahi, *Arma modelling of artificial accelerograms for western iran*, Iranian Journal of Science and Technology **29** (2005), no. A1, 107–116. (Cited on page 10.)

- [8] I. Gusti Ngurah Agung, *Time series data analysis using eviews*, John Wiley and Sons, 2011. (Cited on page 54.)
- [9] McLeod A.I., Hipel K.W., and Lennox W.C., *Advances in box-jenkins modeling 2.application*, Water Resources Research **13** (1977), no. 3, 577–586. (Cited on page 9.)
- [10] S.C. Ashmore, A.G. Piersol, and J.J. Whitte, *Accelerated service life testing of automotive vehicles on a test course*, Vehicle System Dynamics **21** (1992), no. 2, 89–108. (Cited on page 5.)
- [11] Dimitrios Ateriou and Stephen G. Hall, *Applied econometrics a modern approach using eviews and microfit*, Palgrave McMillan, 2007. (Cited on pages 20, 22, 39, 40 and 47.)
- [12] J.A. Bannantine, J.J. Comer, and J.L. Handrock, *Fundamentals of metal fatigue analysis*, Prentice Hall, 1990. (Cited on page 8.)
- [13] Peter J. Brockwell and Richard A. Davis, *Introduction to time series and forecasting*, 2nd ed. ed., Springer, 2003. (Cited on page 19.)
- [14] Chris Chatfield, *The analysis of time series an introduction*, Chapman and Hall/CRC., 1999. (Cited on page 26.)
- [15] Stephen Cull, Ke Xu, and C. Wu, *Generating accelerated loading profiles from measured time series data*, Internal Report, 2009. (Cited on pages 3, 9, 16 and 45.)

- [16] Stephen Cull and Caisia Yang, *Generation and verification of accelerated durability tests*, Internal Report, 2010. (Cited on pages 4, 11 and 16.)
- [17] I. Dobre and A.A. Alexandru, *Modelling unemployment rate using box-jenkins procedure*, Journal of Applied Quantitative Methods **3** (2008), no. 2, 156–166. (Cited on pages 10 and 11.)
- [18] K. Dressler, V.B. Kottgen, and H. Kotzle, *Synthesis of realistic loading specifications*, European journal of mechanical and environmental engineering **41** (1996), no. 3, 153–168. (Cited on page 8.)
- [19] Walter Enders, *Applied econometric time series*, Wiley, 2010. (Cited on pages 11, 21, 25, 36, 40 and 41.)
- [20] R. Gopalakrishnan and H.N. Agrawal, *Durability analysis of full automotive body structures*, International Congress and Exposition **n/a** (1993), no. n/a, n/a. (Cited on page 6.)
- [21] John C. Hoff, *A practical guide to box-jenkins forecasting*, Lifetime Learning Publications, 1983. (Cited on pages vi, 10, 33, 38, 39 and 45.)
- [22] Dressler K., Speckert M., and Bitsch G., *Virtual test rigs for automotive engineering*, Vehicle System Dynamics **47** (April 2009), no. 4, 387–401. (Cited on pages 6 and 7.)
- [23] Scott Klippenstein, *Multivariate time series modeling and forecasting winnipeg’s electrical load-temperature relationship*, Master’s thesis, University of Manitoba, Department of Statistics, 2005. (Cited on page 20.)

- [24] Milligan M., Schwartz M., and Wan Y-H., *Statistical wind power forecasting for u.s. wind farms*, National Renewable Energy Laboratory, U.S. **n/a** (2003), no. n/a, NREL/cp-500-33956. (Cited on page [10](#).)
- [25] S. Makridakis, S.C. Wheelwright, and R.J. Hyndman, *Forecasting: Methods and applications*, 3rd edition ed., John Wiley and Sons, 1998. (Cited on pages [10](#) and [11](#).)
- [26] A. Pankratz, *Forecasting with univariate box-jenkins models*, John Wiley and Sons, 1983. (Cited on page [20](#).)
- [27] Olafsson S. and Sigbjornsson R., *Application of arma models to estimate earthquake ground motion and structural response*, Earthquake Engineering and Structural Dynamics **24** (1995), no. 7, 951–966. (Cited on page [10](#).)
- [28] G. Schwarz, *Estimating the dimension of a model*, The Annals of Statistics **6** (1978), no. 2, 461–464. (Cited on page [26](#).)
- [29] Fearn T. and Maris P.I., *An application of box-jenkins methodology to the control of gluten addition in a flour mill*, Center for Quality and Productivity Improvement **n/a** (1990), no. n/a, Report No.50. (Cited on page [10](#).)
- [30] T.M.J.A.Cooray, *Applied time series analysis and forecasting*, Alpha Science International Ltd., 2008. (Cited on pages [19](#), [36](#), [39](#), [40](#) and [42](#).)
- [31] J. Wannenburg and P.S. Heyns, *An overview of numerical methodologies for durability assessment of vehicle and transport structures*, Int. J. Vehicle System Modelling and Testing **5** (2010), no. 1, 72–101. (Cited on pages [7](#) and [8](#).)

- [32] William W.S. Wei, *Time series analysis univariate and mutivariate methods*, Addison-Wesley Publishing Company, Inc., 1994. (Cited on pages 24, 34, 41 and 42.)
- [33] Ke Xu, *Develpoment of vibration loading profiles for accelerated durability tests of ground vehicle*, Master's thesis, University of Manitoba, Department of Mechanical and Manufacturing Engineering, 2010. (Cited on pages vi, viii, ix, 3, 4, 5, 6, 9, 13, 14, 15, 16, 17 and 46.)
- [34] Peijun Xu, G. Peticca, and D. Wong, *A technique for developing a high accuracy durability test for a light truck on a six degree-of-freedom road test simulator*, International Journal of Vehicle Design **47** (2008), no. 1-4, 290–304. (Cited on page 7.)
- [35] P. Zhang and C. Wu, *Events identification of curve fitting method*, Internal Report, 2010. (Cited on page 9.)