FACULTY OF ENGINEERING AND TECHNOLOGY
DEPARTMENT OF CIVIL ENGINEERING

STOCHASTIC MODELLING AND FORECASTING OF MONTLHY GROUNDWATER LEVELS IN RAMOTSWA WELLFIELD

By

Jackson Matsaunyane

200301386

Supervision Team: Prof. B. P. Parida
Dr P.K. Kenabatho
Mr N.M. Sebusang

July 2015

ABSTRACT

Ramotswa wellfield located in South Eastern Botswana is undergoing rapid urbanization with a significant population growth rate between 2-3% per annum. This along with unreliable rainfall pattern, has escalated the demand on fresh water supplies, hence pressure on the well field and supplies from the Gaborone Dam. Use of historical groundwater level records from observation boreholes provides valuable source of information for understanding the hydrological dynamics due to different stresses within the aquifer system and helps in anticipating future challenges that are likely to occur due to these stresses. In this study, an attempt has been made to model the fluctuation of monthly groundwater table data in Ramotswa wellfield both in space and time using geostatistical and stochastic models. By conducting time series modelling, monthly groundwater level data collected from 2002 to 2012 at 13 different wells were subjected to intervention analysis. This was done to detect any changes in the data due to natural and manmade causes. Cumulative Summation (CUSUM) results revealed that groundwater level data had undergone intervention at all the boreholes. Change which was confirmed by T-Statistic test at 5% significance level was identified at end of year 2007 and beginning of year 2008.Trend analysis was conducted using Mann Kendall test for data after time of intervention. The results revealed that trend was not statistically significant in most boreholes. For stochastic modelling, data at each borehole were subjected to two approaches namely Autoregressive Integrated Moving Average (ARIMA) and Thomas Fiering models to make three months forecasts. The most suitable model was chosen. It was found that in almost all cases ARIMA models gave the least error estimates in forecasting and hence was recommended for forecasting.

For spatial interpolation of groundwater levels at unknown locations within the well field, geostatistical modelling approach was used for two scenarios; winter (July 2005) and wet season (i.e. after February 2006 floods). Based on the results of 3 semi-variogram models to the observed data, exponential model was found to provide the best fit for July 2005 scenario while for February 2006 it was the spherical model. The choice of the two models was supported by reasonable values of r^2; which were 0.7 for July 2005 and 0.9 for February 2006. The nugget- to -sill ratios of 0.1 ($<$ 0.25) for July 2005 scenario groundwater levels have strong spatial dependence while the ratio was 0.29 for February 2006 scenario indicating moderate spatial dependence. After interpolation by ordinary kriging, the results of the groundwater level at unknown were cross validated with known values at 5 boreholes within the radius of influence and the percentage errors were found to be low. The results indicated that groundwater levels are affected by topography and presence of water bodies and demonstrated the usefulness of stochastic modelling for temporal and geostatistics in spatial modelling of groundwater table in the study area, hence in water resources planning and management.

Keywords: Stochastic models, ARIMA, Thomas-Fiering model, Geostatistics, Ordinary Kriging, groundwater level

ACKNOWLEDGEMENTS

First of all I would like to thank God for his endless blessings throughout my life.

I would like to express my sincere thanks and appreciation to Barclays Bank through the Barclays F.G Mogae scholarship for granting me this opportunity to study for my Masters degree. Without your financial support, it would have been very challenging to finance my education.

I would like to express my sincere gratitude to the supervision team especially Prof B. Parida for the support, constructive comments, guidance, encouragement and supervision of my work throughout this study. My thanks and appreciation goes out to all staff members in the Department of Civil Engineering who made a significant contribution towards my postgraduate study.

I wish to extend to extend my thanks and appreciation to Department of Water Affairs especially Mr N.Ramotsoko for their willingness in sharing groundwater level data used in this study. I would also like to thank the Department of Meteorological Services for providing me with monthly rainfall data.

Lastly I would like to thank my family friends, and colleagues for their support and words of encouragement throughout my studies.

TABLE OF CONTENTS

Page Numbers

Chapter 3- Methodology

Chapter 4-Trend and Intervention Analysis

Chapter 5- Stochastic and Geostatistical Modelling

Chapter 6- Summary, Conclusion and Recommendations

LIST OF TABLES

Page Numbers

LIST OF FIGURES

Page Numbers

ABBREVIATIONS

ACF	Autocorrelation Function
ASTSA	Applied Statistical Time Series Analysis
AIC	Akaike Information Criterion
PACF	Partial Autocorrelation Function
ARIMA	Autoregressive Integrated Moving Average
TF	Thomas Fiering
BNWMP	Botswana National Water Master Plan
SADC	Southern African Development Community
UNDP	United Nations Development Program
GWL	Groundwater level
UNEP	United Nation Environment Program
NSC	North South Carrier
WUC	Water Utilities Corporation
DWA	Department of Water Affairs
RMSE	Root Mean Square error
ME	Mean Error
MDG	Millennium Development Goals
ANN	Artificial Neural Network approach
CUSUM	Cumulative Summation
P	Order of autoregressive Component
Q	Order of Moving Average Component
d	Non-Seasonal Differencing
D	Seasonal Differencing
ILWIS	Integrated Land and Water Information System
OBS	Observed groundwater levels
SPSS	Statistical Package for the Social Sciences

CHAPTER 1-INTRODUCTION

1.1 BACKGROUND AND RATIONALE

Water is one of the most important natural resources which supports all forms of life in our planet. Reliable access to adequate clean fresh water is now regarded as a universal human right (United Nations Committee on Economic, Social and Cultural Rights, 2003). The Millennium Development Goals (MDG's) recognizes the need for reliable supply of clean water and have set a target of reducing the number of people not having access to clean water and sanitary facilities to half by year 2015 (UNDP, 2006).

It has been estimated that by 2025, 5 million of the world's population will be living in water stressed countries and most of these countries are in Africa and South East Asia (Arnell *et al*., 1999). This water scarcity would be a direct effect of a combination of many factors such as increase in population, the ever increasing demand for water in this industrialization era as well as changes in climate. Climate change has over the years been linked to a number of various changes in different components of the hydrological system. This change is expected to result in an increase in temperature and an intensification of the hydrologic cycle. This leads to frequent occurrence of extreme events such as tropical typhoons, droughts, floods which cause catastrophic damage to human beings and other living organisms. According to simulations results of global hydro climatic variables conducted using mid-range emission scenario (IS92A), mean air temperature, precipitation, evaporation and runoff are likely to increase by 2.3℃, 5.2%, 5.2% and 7.3% respectively by 2050 (Wetherald *et al*.,2002). A lot of precipitation and evapotranspiration from the intensified cycle however would not be equally distributed in space. According to Arnell (1999), it is expected that increase in runoff will be experienced in high latitude, equatorial and tropical areas and will decrease in mid latitude and subtropical regions.

Groundwater is a very important source of water and forms part of the hydrologic cycle. Groundwater resources represent 15% of Africa's renewable water resources, and because of its hidden nature it has been undervalued and unutilized (UN Economic Commission for African Climate Change Policy Centre, 2011). Even though it has been second to surface water for a very long time in terms of use by general public and water sector managers (Villholth *et al.*, 2007), it plays an important role in maintaining important surface water systems and the ecosystem.

Groundwater level is controlled by a balance between storage, recharge, and discharge (Taylor *et al*., 2001). This balance is affected by physical factors such as permeability, porosity and sizes of rocks and sediments. The balance is also affected by climatic conditions, groundwater abstraction rates as well as land use/land cover changes. Groundwater monitoring needs to be conducted over long periods of time to enable proper development, management and protection of water resources. Water use data such as pumping rates, volumes of water pumped can enhance the interpretation of observed trends in water level and changes in storage capacities associated with withdrawals over time.

Failure of groundwater supply system is almost always due to failure of infrastructure or unsustainable and uncontrollable pumping rates over a short period of time (Department of Water Affairs, South Africa, 2010). The adjustment of groundwater storage in aquifers due to the factors already mentioned above can be easily described from measurements of water table. There is a massive gap in the level of available information and documentation of the state of groundwater in both developed and undeveloped countries which hinders groundwater monitoring and planning (Villholth *et al*., 2007).

Developed countries like the United States of America are experiencing challenges of groundwater monitoring due to data unavailability in certain locations. It is for this reason that the US Geological Survey has called for setting up of a nationwide program for a systematic and comprehensive record of water levels at observation wells (Taylor *et al*., 2000). The Minnesota Groundwater Monitoring Network found the importance of groundwater level monitoring by recommending the expansion of their monitoring network of wells from 750 to 7000 (Minnesota Groundwater level monitoring, 2011). This expansion was necessary because there were large unmonitored areas. The expanded network would become a long term investment to fully understand and manage groundwater in Minnesota.

In the SADC region, there have been some attempts to conduct regional groundwater monitoring for sustainable development through Groundwater Management Program (GMP) after recognizing the increase in demand for groundwater. This program was mainly focused on assessment, exploitation and protection with emphasis on groundwater resource management and coping strategies to deal with drought. It was not addressing the inadequacy and inhomogeneity of regional groundwater data. According to their study, a survey conducted in 2002 on the regional groundwater analysis situation for Southern Africa found that very few countries in the region conduct groundwater monitoring despite groundwater being such an important resource for rural communities. One of the recommendations made was the setting up of a comprehensive regional groundwater database in which there would be free dissemination of information to member states.

The importance of groundwater monitoring was highlighted in the South African Groundwater strategy for 2010 which stated that lack of monitoring data makes it challenging to make an accurate estimate of availability of groundwater as well as its abstraction rates. South Africa has a national network of monitoring boreholes and a National Groundwater Database. The network is still deemed inadequate and unreliable for water resources assessment (Department of Water Affairs, 2010). In an attempt to monitor groundwater, Water Resource Council of South Africa had a project of developing Groundwater Management Framework which included many aspects of groundwater management at municipal level (Riemann *et al*., 2012). The aim was to improve management of groundwater by equipping water authorities with necessary tools and capacity to conduct groundwater data collection and monitoring.

There is a clear indication of the need to conduct a comprehensive groundwater monitoring for our groundwater resources. Part of groundwater monitoring plan is monitoring of groundwater table data. Research on groundwater level fluctuation with time as a way of understanding and quantifying this depleting resource would go a long way in providing effective and sustainable solutions in the ever changing climatic condition. Due to the depletion of groundwater resources in many areas especially the semi-arid part of Africa, hydrological modelling would serve as an important tool for planning and management of groundwater especially during dry seasons. Groundwater plans and policies can be developed on the reliable basis of a countrywide assessment of trends and future estimates of groundwater.

1.2 PROBLEM STATEMENT

Botswana, like other Southern African countries located in the semi-arid region, is water stressed, with its fresh water resource ranging between 1000 to 1700 m^3 per person per year (UNEP, 1999). Due its accelerated economic and population growth in recent years, the problem of water scarcity is expected to continue in the near future with the ever changing climate having an adverse effect on water resources. With surface water development being limited by many factors such as lack of dam sites, low rainfalls and high evaporation rates, groundwater is a very critical resource required to meet the daily needs including domestic, agricultural and industrial use. It is estimated that Botswana's groundwater potential stands at 1.7×10^9 m^3/year (Kgathi, 1999) and it is capable of supplying 80% of Botswana's population (Department of Water Affairs, Botswana, 2006). Groundwater which is the main source of water accounts for 66% of portable water supply (Majelantle, 2009) mainly to rural areas while the rest is being accounted for by surface water for major towns and cities. There is a serious concern regarding the rate of abstraction as compared to recharge capacities of the aquifer. The annual rate of abstraction is estimated to be 76×10^3 m^3 and it is expected to increase (Du Plessis et al., 2003). The implication of this is that the resource is likely to last for a few decades, hence careful monitoring as well as knowledge is required for its sustainable use.

Botswana recently has been experiencing frequent droughts which have been affecting different sectors and continue to frustrate Government's efforts of diversifying the economy. The most hardly hit areas are the Western and the Southern parts. In the Southern region, the main sources of water for urban areas is surface water. The capital city Gaborone located in the South East District has been one of the most affected areas. It has over the years been experiencing rapid urbanization with an influx of people from remote areas of Botswana and other neighbouring countries. It is the central focus of the country's economy and hosts important governmental structures like government enclave, ministerial headquarters, government departments, embassies, and many other commercial, industrial and business centres.

With the city's main source of water, the Gaborone Dam now completely dry as shown on Figure 1.1, it is unable to meet the ever increasing water demand from these different sectors of the economy.

Fig 1.1: The drying Gaborone Dam, the main supply of potable water to Gaborone and surrounding areas.

The dam is also a source of portable water to other major villages in its surrounding such as Mogoditshane, Tlokweng, Ramotswa, Gabane and Mochudi which are also experiencing unprecedented population growth due to their proximity to the capital city. Other nearby dams (e.g. Nywane and Bokaa Dams) have also been drying up due to low rainfalls and increases in demand for water in the rapidly growing towns. This has triggered rationing of water by Water Utilities Corporation which has resulted in less water available mainly for domestic and industrial use. The drying up of dams due to low rainfall and increase in demand has demonstrated that surface water resources are becoming more and more unreliable. Even though plans are in place to augment supply through the North South Carrier (NSC) 2 pipeline from Dikgatlhong Dam, the project is still at implementation stage and because of that, water scarcity will continue to persist in the Southern Region. There is therefore a need to intensify efforts of considering groundwater as a supplementary source to minimize the current overreliance on surface water sources especially in urban areas.

Ramotswa Wellfield located in South–Eastern part of Botswana is one of the most productive wellfields in the country. It has in the past been used as emergency supply during drought periods as far back as the 1970's (Geotechnical Consulting Engineers, 2000). As part of the Gaborone Emergency Water Supply Scheme, boreholes were connected to the Lobatse-Gaborone pipeline. Following the completion of Gaborone dam in 1984, supply of water from Ramotswa to Gaborone was discontinued. Some of the boreholes were decommissioned while the remaining ones were used solely for the supply of Ramotswa village (Geotechnical Consulting Engineers, 2000). The wellfield was completely abandoned in 1996 due to high concentrations of nitrates from seepage of pit latrines

However to avert the recurring drought situation, efforts are being made by Water Utilities Corporation (WUC) to utilize groundwater from the wellfield as a mitigation strategy. There is already some abstraction being done in some of the production boreholes whereby groundwater is being blended with surface water from NSC in an effort to address the current water crisis. In addition to that, WUC plans to construct a reverse osmosis treatment plant to treat groundwater from the wellfield. According to the Corporation, plans are at an advanced stage to construct the plant and the project is expected to cost approximately P50 million.

With the current depletion of surface water resources in the area and the plans to fully utilize groundwater resources from Ramotswa Wellfield, a lot of attention has been focused on the nitrate concentrations in the water. However, it is also important to monitor the quantity of groundwater in the wellfield especially considering the ever increasing demand of water and the amount of investment that is being proposed for the treatment plant. According to the Minnesota Groundwater level monitoring report (2011), groundwater level measurement is currently the only reliable to tool to measure changes in quantity of groundwater.

Modelling and forecasting of groundwater levels can provide vital information on how the aquifer storage is changing with time. This information can be useful in planning and management of groundwater at the Ramotswa wellfield. No attempt has been done to model the fluctuation of groundwater level at this wellfield even though it is of paramount importance as a source of water to Ramotswa village and the surrounding areas. An attempt was therefore made through this study to model the spatial and temporal fluctuation of groundwater levels in the wellfield.

1.3 STUDY AREA

1.3.1 Location

Ramotswa village is located in the South Eastern part of Botswana, about 35 km south of nation's capital Gaborone. Because of its proximity to the capital city it has experienced significant population increase of about 21% between 1991 and 2011 due to migration of people from the city and surrounding areas (Kholoma, 2011). Ramotswa wellfield is located approximately 25 km upstream of Gaborone dam. It covers an area of 29 km^2 of which part of it is within the village (Geotechnical Consulting Engineers, 2000).

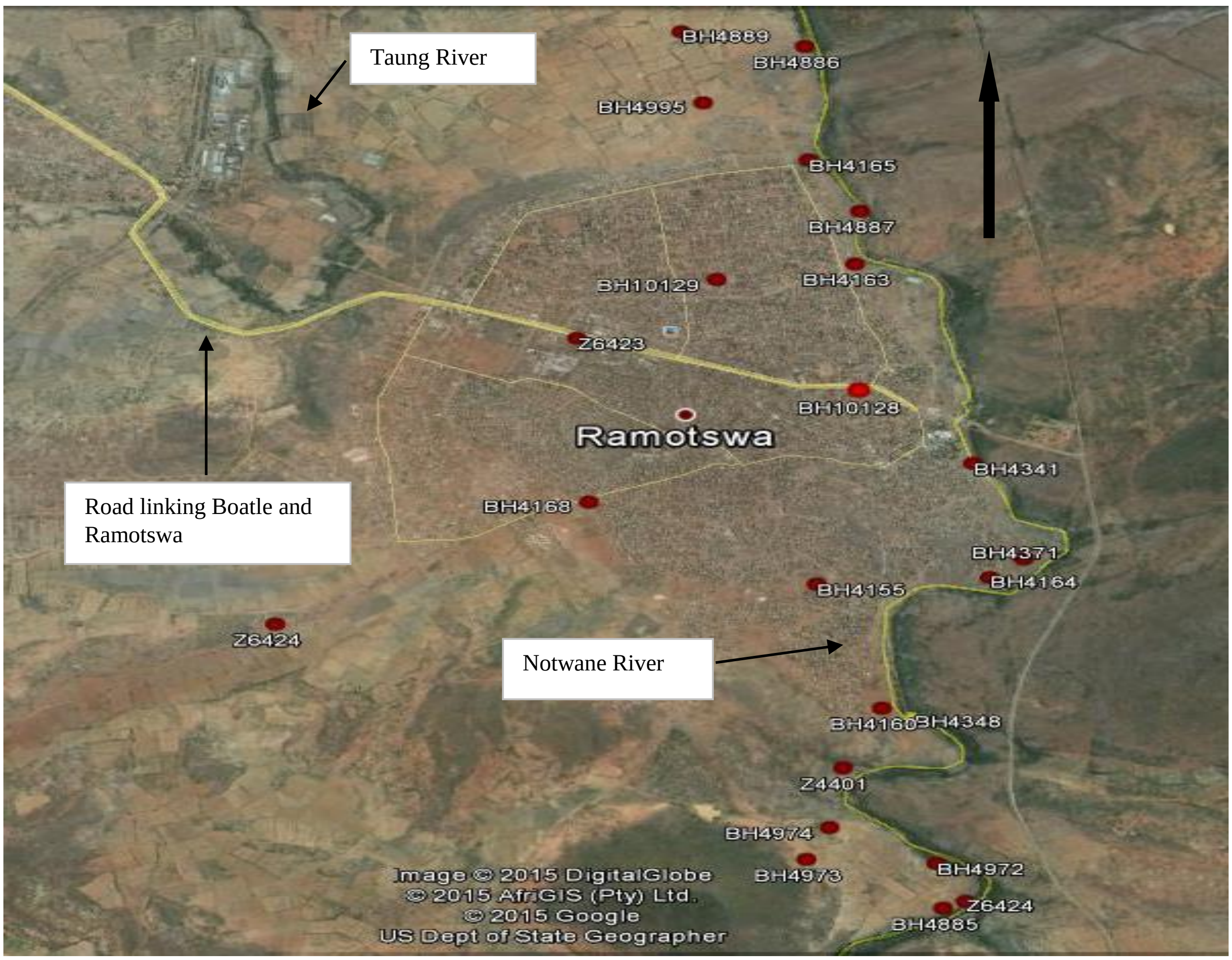


Fig 1.2 Location of the wellfield (Google Earth, Google 2015)

1.3.2 Climate

The climatic conditions of the study area are typical of a semi-arid climate, an area closer to the desert with no maritime influence because of the country's geographical location. The area experiences a wet season between October and March, and a dry season between April and August. Kholoma (2011) reported mean annual maximum temperature to be 28.9℃ and mean annual minimum temperature to be 12.9℃. The mean annual rainfall ranges between 475 mm and 525 mm. The rainfall events in this type of climate are highly intensive and have a short duration. High levels of evaporation occurs mostly in summer with the daily peak between 1300 hrs and 1600hrs.This is supported by evidence of high average daily pan evaporation of approximately 5.5 mm per day as reported by Department of Water Affairs (DWA) (2003).

1.3.3 Topography

The terrain is generally a hilly terrain, with landforms such as Sepitwane, and Rankepa hills. The altitude ranges from 1012 m to 1189 m above sea level (Kholoma, 2011).The area is drained by two ephemeral rivers namely; Notwane, and Taung Rivers. Notwane which is the main river while the other two are tributaries located in the western part of the area. The wellfield is located along this main stream which slopes gently in the northern direction with an overall hydraulic gradient of 1:300 (Staudt, 2003). The vegetation cover is predominantly shrubs and tree savanna but has been affected by human interventions such as cutting down of trees and overgrazing. This has resulted in erosion due to high intensity storms during rainy seasons.

1.3.4 Land Use and Development

Ramotswa which hosts the administrative headquarters of the South East district has undergone rapid development over the years. This includes provision of educational, social, industrial, medical and utility services making it highly attractive for migrants from areas within its vicinity. Industries such as steel factories, meat processing, and flour production can be found in the village. Agriculture still remains an important activity for sustaining the livelihoods of the residents. Sorghum and maize are the main crops while livestock keeping is mainly cattle, sheep and goats. Higher stocking in some areas has resulted in overgrazing and eventually land degradation due to soil erosion.

1.3.5 Hydrogeology

As reported by Staudt (2003), the area consist of three lithological supergroups namely; Otse Waterberg, Transvaal and Ventersdorp. Otse Waterberg covers the south western part, while Transvaal supergroups which are intensively faulted with deep normal faults covers the middle and southern part of the study area. Lobatse Ventersdorp which are the oldest in the area covers the middle and the northern part.

There are two aquifers in the study area namely; Ramotswa Dolomite and Lephala formation and these aquifers are considered to be hydraulically connected (Geotechnical Consulting Engineers, 2000). The Dolomite aquifer has two different karstic zones; the upper and the deeper zone. The upper zone which has a variable thickness of 20 to 50 m recharges from river and infiltration. The deeper zone has a thickness between 25 to 50 m. Dolomite have high transmissivity and storativity due to local karstification. For the study area within the Dolomite aquifer, high yielding wells are located along the major linear karst while low yield ones are located in the rock. It is this fracturing and intersection with minor side valleys in the East-West direction that produces favourable permeability conditions. Lephala formation is similar to Dolomite one but without karstification. It is found in the Southern and North Eastern part of Ramotswa. It consist of two fissured zones which are separated by a less fissured zone. The thickness of the upper zone is 30-40 m while that of lower zone is 30 m. Yield of boreholes in the Lephala formation depends on their proximity to the river, intersection of fissured zones and the extent of secondary infills (Staudt, 2003).

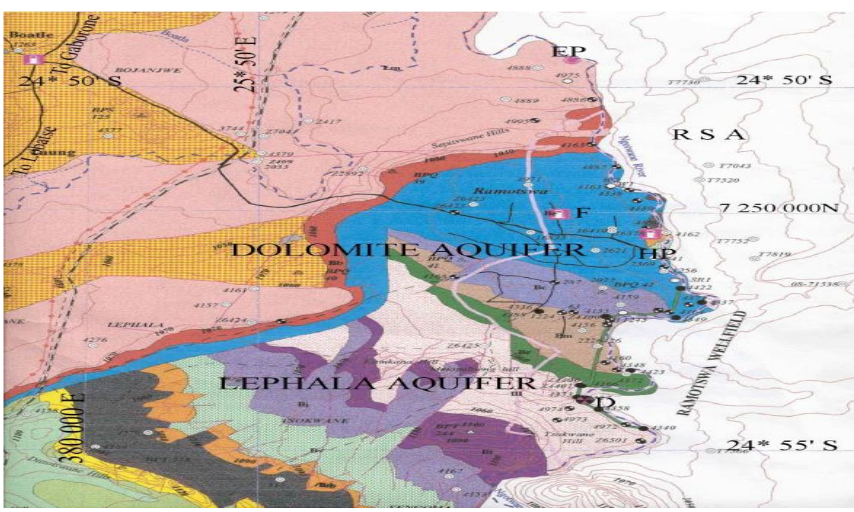


Fig 1.3: Map showing the two aquifers at the study area (Rangani et al., 2002)

1.4 OVERALL OBJECTIVE:

The main objective of this study is to use stochastic and geostatistical methods to model the temporal and spatial variation of monthly groundwater levels in Ramotswa wellfield.

1.4.1 Specific Objectives:

i. To investigate the homogeneity in the groundwater level data.

ii. To develop and forecast monthly groundwater levels using appropriate stochastic models

iii. To compare the forecasts from the chosen stochastic models and recommend the appropriate one for use in forecasting groundwater levels in the wellfield.

iv. To develop appropriate geostatistical techniques for spatial interpolation of groundwater levels in the wellfield during wet and dry season which can be used for determination of groundwater levels at unmonitored areas.

1.5 JUSTIFICATION OF STUDY:

The South Eastern part of Botswana has been experiencing a high rate of urbanization and industrialization. This is indicated by a rapid increase in population in the Gaborone and the surrounding areas. Due to the aridity of the region caused by high rates of evaporation and low rainfall, pressure has been exerted on the limited surface water sources to an extent of drying up. Because of the criticality of water for socio-economic development, inadequate supply of water hampers the chances of achievement of the country's vision 2016 pillars and some of the Millennium Development Goals (MDG's). In this regard, groundwater in the Ramotswa wellfield is a valuable resource that needs to be carefully monitored and utilized for sustainable supplementary supply of water in the district. However, for the resource to be sustainable, knowledge on the status of groundwater resource is of paramount importance to water resources managers to ensure adequate supply for present and future generations. It is for this reason that this study is undertaken in order to have a better understanding of what is likely to happen in the near future with regards to water storage in the wellfield aquifer so that appropriate mitigation strategies could be implemented.

1.6 LIMITATIONS

In this study, an attempt is being made to make use of available time series data to model and forecast groundwater levels at various observation boreholes. This could also be done through modelling of the physical processes involved. However physical modelling of the entire processes will involve measurement of hydrogeological variables for the entire aquifer which would be expensive and time-consuming.

Another limitation of the study is in terms of water quality. High concentrations of nitrates in the wellfield has been a subject of many discussions after the closure of the wellfield in 1996. Because of lack of adequate nitrate monitoring data and other water quality variables that need to be monitored for water to be considered safe for drinking, the study will only focus on the quantity aspect. The areas mentioned above which are not covered would provide an opportunity for further research in the wellfield.

1.7 LAYOUY OF THE DISSERTATION

The following is an outline of the contents of the dissertation;

Chapter 1 deals with the general introduction which consists of background and rationale, problem statement, study objectives and justification, study limitations as well as a brief description of the study area. **Chapter 2** comprises of the review of relevant literature on time series forecasting and geostatistical analysis with an emphasis on their hydrological application. These studies are being discussed in relation to the main objective of this dissertation. **Chapter 3** describes the methods recommended for addressing the main objective and how they will be applied. The 3 main methods described are ARIMA, Thomas-Fiering and geostatistical methods. **Chapter 4** describes the time series hydrographs of groundwater levels at various observation boreholes. Results and discussions of trend and intervention analysis are also presented in this chapter. **Chapter 5** covers presentation of results and discussions for both stochastic and geostatistical analysis. **Chapter 6** summarizes, concludes and makes recommendations for further research

CHAPTER 2-LITERATURE REVIEW

2.1 INTRODUCTION

As persistent droughts becomes more frequent and the demand for water continues to increase especially in semi-arid regions, the need for more efficient models and forecasting techniques of groundwater availability has become increasingly important. This literature review covers some of the most important and relevant work conducted by various researchers in an attempt to address the main objective of the study. These include trend and intervention analysis to ascertain the homogeneity of the observed series, and forecasting methods used in hydrology. The last part of the review focuses on stochastic models (ARIMA and Thomas-Fiering model) and geostatistical modelling.

2.2 INTERVENTION ANALYSIS

Before use of a data driven model like time series models for hydrological forecasting, it is of great interest and importance to understand the characteristics of the data. This can be done by subjecting the data to intervention analysis. Intervention here refers to a statistical approach used to analyse the effect of natural or manmade change in a time series (Hipel *et al*., 1978). Through this method, the actual change in the data can be statistically determined. After a point of intervention has been determined, it can be decided whether to choose the whole data for modelling and forecasting or part of the data. Cumulative Summation (CUSUM) technique by Woodwars and Gold Smith (1964) provides a graphical procedure for detecting intervention in a time series data. However, Hipel *et al*., (1978) suggested CUSUM as only an identification tool for estimating the time of intervention, therefore not statistically describing intervention effects. Kampata *et al*., (2008) used CUSUM technique to investigate the possibility of intervention in rainfall data from upper Zambezi basin. The step change analysis method was used to confirm any change in the homogeneity of data since CUSUM only identifies change graphically. Kenabatho *et al*., (2012) used CUSUM to investigate the suspected point of intervention using Botswana's climatic data. T-statistic method was used for confirmation of intervention using two samples (i.e. before and after the suspected point of intervention). Parida *et al*., (2008) analysed rainfall time series using intervention method to check whether the rainfall series came from the same population. CUSUM technique was once again used to detect a suspected point of intervention which was confirmed by T-statistic method. Based on literature review on intervention analysis, it was therefore important to test whether the groundwater level time series in the Ramotswa wellfield have been subjected to any intervention before using the series for forecasting.

2.3 TREND ANALYSIS

Trend is a tendency for successive values to be increasing or decreasing over time (Haan, 2002). It is one of the statistical analysis techniques that need to be carried out for almost all water resources studies involving use of hydrological time series data. Results of trend analysis are important in management of water resources and helpful in forecasting patterns of future hydrological events such as floods and droughts. Trend analysis approaches can be parametric or non-parametric. According to Machiwal *et al*., (2012), parametric approach is usually used by researchers in the time domain such as economics who make certain assumptions about the nature of the data. Parametric approach is a very powerful technique but requires data to be independent and normally distributed (Fathian *et al*., 2014). On the other hand, non-parametric approach is more widely applicable in hydrology than parametric approach since hydrological data often have missing segments and not normally distributed.

Moalafhi *et al*., (2012) used Mann Kendall test to analyse rainfall and temperature data from four meteorological stations in Botswana. They found out that rainfall was decreasing while temperature was showing an increasing trend. In both variables, the trend was not statistically significant. In order to estimate trend in hydroclimatical variables which affected flow into the Urmia Lake in Iran, Mann Kendall statistical test was one of the non-parametric tests used by Fathian *et al*., (2014). Their findings revealed that there was a significant increase in temperature throughout the basin and 75 % of the monitored rainfall stations were not showing any trend. Kampata *et al*., (2008) used Mann Kendall test to conduct trend analysis of rainfall in the headstreams of the Zambezi river basin. They concluded that even though the results were showing downward trends in rainfall at the five monitoring stations, the trend was not significant. In order to address some of the objectives stated above, a non-parametric approach was used to conduct trend analysis on the groundwater levels for Ramotswa wellfield. As a widely used method for checking the existence of linear trends by various studies in literature, Mann Kendall test was recommended for the study.

2.4 FORECASTING METHODS

Use of accurate models in hydrology is of fundamental importance for long term prediction of future events in order to make informed decisions regarding sustainable utilization of water resources. Some of the common applications of modelling and forecasting in operational hydrology include optimal reservoir operation, drought management, environmental protection, operation of water utilities, sustainable water resources development. It is for this reason that the main objective of this study is aimed at forecasting monthly groundwater levels in the Ramotswa wellfield using stochastic models.

Hydrological modelling and forecasting has been a subject of interest and has received tremendous attention from various researchers. Nayak *et al*., (2006) stated that hydrological models can be categorized into empirical time series models and physically descriptive modelling. Empirical modelling is done based on the collected data (i.e. data driven) while physical modelling uses the laws of physics to describe the physical processes involved. Empirical time series models have been used for hydrological applications in studies such as the one conducted by Knotters et al., (1997) in which the relationship between water table depth and excess precipitation was established. The advantage of empirical time series model is that only data on the input and output variable are required for calibration while its downfall being the fact that it cannot be used for forecasting when the dynamic behaviour of the hydrologic system changes with time (Bierkens, 1998). On the other hand physical models require enormous data, that is generally difficult or expensive to source making it challenging to apply in developing countries where the resources are limited (Nayak *et al*., 2006).

Linear regression analysis is a statistical tool to analyse data, estimate model parameters and make forecasts. The statistical association between two variables is determined by drawing a line of best fit using least square method technique. Regression can be expanded into one or more independent variables in which it is referred to as multiple linear regression (Guerard, 2013). Multiple linear regression models can be used to simulate the association between the explanatory and the response variables by fitting a linear equation to observed data (Markridakis *et al*., 2008). These models have been used in hydrological applications e.g. Shao *et al*., (2002), Hodgson, (1978) because they are easy to use, very flexible in inclusion of any number of independent variables, and easy interpretation of relationship between parameters (Sahoo *et al*., 2014). However the disadvantage in these models is that they assume linearity of variables but the relationship between hydrological variables may be non- linear.

ARIMA models have been applied by many researchers in modelling and forecasting of hydrologic data. Most of the hydrological time series are non-stationary due to trend and seasonal effects and this implies that the mean and variance of the series changes with time. ARIMA is often preferred for modelling and forecasting using such kind of data due to their flexibility and inclusion of both autoregressive and moving average terms. They have been found to be highly capable of describing the time related changes and forecasting of hydrological variables (Kuruc *et al*., 2004). However just like the linear regression model these models assume that the data has a linear relationship and cannot be used to address the nonlinearity which is often encountered in hydrology. Another disadvantage is that the model is complex, requires an experienced modeller to produce satisfactory results and also requires a lot of data to build a reasonable model (Bails *et al*., 1993).

Some studies e.g. Nayak *et al*.,(2006) have recommended the use of models which account for non-linearity which is a prime characteristic of issues relating to atmospheric and hydrological science since the models discussed above are only limited to linear assumption. This has led to the emergence of Artificial Neural Network approach (ANN) as an important tool for forecasting in many areas of science and engineering. Neural Networks have remarkable ability to derive meaning from imprecise data and can be used to extract patterns that are too complex to be noticed by humans or other computer techniques (Moalafhi *et al*., 2014). However even with their ability to overcome non-linearity in the data, they have two main disadvantages; their computational time and overfitting. ANN approach usually require a trial and error approach for parameter estimation and one is never sure whether an optimal model has been obtained (Haijie, 2010). Over memorizing (overfitting) can happen when the model is over trained with training data (when there are too many hidden neurons) leading to loss of generalizing to forecast future data.

Even though these models vary from simple to more sophisticated, it can be seen from the above discussion that none of the model is perfect for forecasting .Therefore given these alternatives, a choice needs to be made so that appropriate forecasting model can be selected (Makridakis *et al*.,1982). The selection of an appropriate model for a particular problem depends on many factors such as the number of series to be modelled, required accuracy, modelling costs, ease of use, ease of interpretation of the results (Adhikary *et al*., 2012). For instance, ARIMA models are linear and parametric, i.e. data is required to be stationary, sampled at equal time intervals and have a significant degree of serial dependence (Toth *et al*., 2000). It has been found through the Makridakis forecasting competition in 1982 that complex or highly sophisticated models did not in general outperform the simple ones therefore the complexity of the model does not guarantee the most accurate and reliable forecast (Weatherford *et al*.,2003).

In an attempt to fulfil the main objectives of this study, the methods discussed above have been considered. Time series modelling have been considered better option for areas where nothing but hydrological time series data is available (Adhikary *et al*., 2012). It is well appreciated that hydrogeological parameters and domain boundary conditions are often not available for physical modelling as in the case of Ramotswa wellfield. Most of these parameters are very difficult to obtain because of several natural and anthropogenic factors (Kim *et al*, 2005). Since time series models are popular tools for medium range forecasting and generating synthetic data, they were recommended to be applied for the study area. The two time series models that were applied for the study area are ARIMA and Thomas Fiering Models. Although this study mainly covers time series models for forecasting groundwater levels, geostatistical techniques were utilized to model the spatial distribution of groundwater level in the wellfield. The discussions in the next sections focus on studies in literature which utilized ARIMA, T-F as well as geostatistical modelling for solving various hydrological problems with more emphasis on groundwater.

2.4.1 ARIMA models

Extensive research work has been conducted in the area of time series analysis of hydrological variables. This involves use of historical observations taken at uniform intervals to develop stochastic models with an aim of understanding the past hydrological series as well as attempting to forecast the likelihood of occurrence of future events. In order to achieve this, a statistical description of a population is conducted based on a limited number of samples (Salas *et al*., 1993). ARIMA models developed by Box and Jenkins (1976) have been extensively used to simulate and forecast hydrological variables and processes such as stream flow, precipitation, water quality data etc.

In one of the studies conducted in Bangladesh, Adhikary *et al*., (2012) used ARIMA to model the fluctuation of groundwater table in an unconfined aquifer. They used weekly water level monitoring data from 5 observation wells from 1999 to 2006. In their approach, they used linear regression and periodogram to examine the trend and periodic behaviour of the groundwater series respectively. The appropriate model was selected from the other candidate models after verifying its performance using error indices. The outcome of the results showed a decline in groundwater level trend and cyclic annual periodicity. The study was successfully completed after generating reasonable forecast from 2005 to 2006.

The performance of stochastic models was tested in a study conducted in Kashan aquifer in Iran (Mirzavand *et al*., 2014). They used time series analysis to select the optimum model for water table prediction using observation data from 36 wells. In order to achieve this, they used water level data from 1990 to 2004 for the modelling part, and then used data from 2005 to 2010 for model prediction. From a list of five candidate models (AR,MA,ARMA,ARIMA and SARIMA) with 11 different structures, it was decided based on model evaluation using Akaike Information Criterion (AIC) and correlation coefficients that AR(2) was the most accurate model for forecasting of water table in Kashan aquifer.

Other studies have applied stochastic modelling techniques for time series analysis and prediction of water quality data for proper river basin management (Durdu *et al*., 2009). In that study, ARIMA modelling approach was used for prediction of boron concentration in Buyuk Menderes River, western Turkey using boron concentration data from 1996 to 2004. The study was meant to address issues of river water pollution due to high concentrations of boron which limited its use for irrigation purposes by developing a time series model for forecasting of boron pollution level. In order to check whether the data can be subjected to standard time series analysis, the data was subjected to trend and intervention analysis which were both found to be insignificant. After comparison of the observed and the predicted boron concentrations between 2002 and 2004 using Z test, it was observed that there was no significant difference between the mean and variance of the observed and predicted concentrations. This confirmed indeed confirmed that ARIMA models can be safely used for forecasting of boron concentrations.

Another study of ARIMA modelling of water quality parameters was applied using data from two inlets and an outlet of Latian Dam, in Tehran using 24 years of monthly data from 1981 to 2005 (Asadollahfardi *et al*., 2012). In addition to modelling of water quality parameters time series, they used the developed models to predict the variation of future water quality at the sampled locations. The outcome of their study showed that some water quality parameters showed seasonal behaviour while others were non-seasonal. The trend of TDS, Mg^{2+}, Na^{+} and SO_4^{2-} revealed a maximum amount in April and a minimum in September. The models showed consistency between observed and synthetic monthly predictions hence could be useful for water quality management in the inlet and outlet of the dam.

ARIMA models have been widely used for modelling and forecasting of stream flow by various researchers. Saleh *et al*., (2009) applied these models to study inflow into the Haditha Dam using monthly inflow data from 1999 to 2008. After testing the model's goodness of fit using Port Manteau Lack of fit and Residual Auto Correlation Function (RACF) test, a 3 year monthly flow forecast was determined. The outcome of the study revealed that the inflow into the dam was successfully fitted by ARIMA $(0,1,2) \times (0,1,1)_{12}$ which gave rise to the least errors compared to the other candidate model.

The adequacy of ARIMA family models was once again demonstrated in the study conducted by Mohan *et al*., (1994) using monthly flows into the Bhadra Reservoir system in India. The authors used 25 years of data for model development and 27 years for forecasting of inflow in the monsoon climatic conditions. They found that the inflows into the reservoir can be adequately modelled by ARIMA $(2, 0, 0) \times (0, 1, 1)_{12}$, and this was the model which was subsequently adopted for forecasting of 27 years of monthly inflows. They advocated for long historical data for building of the ARIMA model to avoid frequent modification of the model parameters. In order to demonstrate the flexibility of their developed model, the model was modified and applied in a different location where flows were unknown to evaluate real-time reservoir operating policies and found it to be effective. The application of the ARIMA model not only in optimal operational policies but also in optimal cropping patterns was also highlighted.

Mujumdar *et al*., (1990) used time series modelling to investigate the best suited models for representation of for three South Indian rivers; Cauvery, Malaprabha and Hemavathy. Before the series were used for model development, standardization technique was used for removal of periodicities inherent in the process. Most of the authors in this literature review used Akaike Information Criterion (AIC) as a decision rule for model selection. However, Mujumdar *et al*., (1990) highlighted downfalls of using AIC which include not only the fact that it does not minimize the average value of any criterion function but it is also not consistent. Instead, they recommended the use of maximum likelihood criterion for simulation purposes and mean square error criterion (MSE) also known as the prediction approach for forecasting.

2.4.2 Thomas Fiering model

Thomas Fiering model has been widely used for synthetic generation of hydrological data. This mathematical model has been used by a number of researchers mostly for sequential generation of stream flows in which flow at any time is treated as a linear function of flow in the preceding time step. The model which can be used for weekly, monthly, seasonal or annual observations does not require data to be normally distributed.

Thomas Fiering model was used by Boughto *et al*., (1968) for synthetic generation of stream flow data in New Zealand for gauging stations with more than 20 years of data. Although at that time their paper reported work in progress than a completed study, they did highlight possible errors that could occur when short records are used to estimate statistical parameters such as mean, standard deviation, correlation coefficient which form the basis of the T-F model. In an attempt to model the flow of Khassa Chi river in Kirkuk, Iraq between 1941 and 2001, Cheleng *et al*., (2011) used T-F model to simulate streamflow Synthetic data generated were very much close to the original data. In the synthetic generation of monthly and annual rainfall data at the Goztepe meteorological station, Thomas-Fiering model was one of the models reviewed by Unal *et al*., (2004). They found out that the wavelength approach is capable of preserving the statistical characteristics of the observed series just like the classical approaches in hydrological data generation schemes.

2.4.3 Comparison of ARIMA and Thomas- Fiering models

Most of the literature reviewed above used only one individual stochastic model (either the ARIMA or Thomas-Fiering Model) to forecast the behaviour of a hydrological variable and were not compared with other models to investigate their performance. Consideration of other models would be helpful in advising on the best model to use in order to describe the hydrological behaviour of a series and make informed decisions from reliable forecasts. However some researchers have attempted to compare different forecasting approaches with an aim of evaluating their forecasting capabilities. In a study conducted by Kurunc, *et al*., (2004), a best fit model was selected between ARIMA and T-F approaches using water quality and stream flow data for Yesilirmak River at Durucasu monitoring station, Turkey from 1984 to 1996. After conducting their analysis, it was concluded that Thomas Fiering model performed slightly better than ARIMA model and hence recommended for forecasting.

Similarly another study was conducted by Ahmad *et al*., (2001) for Gange River water quality forecasting. However unlike Kurunc *et al*., (2004) where only two approaches were considered, the authors added a deseasonalised ARIMA model to the ARIMA and T-F approach in their model performance evaluation. For their choice of the most suitable ARIMA model, Akaike Information Criterion was used. The overall conclusion was that in terms of overall error estimates, the deseasonalised approach with the Fourier series technique was recommended as the most suitable model for forecasting of water quality parameters.

2.5 GEOSTATISTICAL ANALYSIS

Geostatistical modelling technique has been widely used in hydrology to investigate the spatial variation of observed hydrological variables. This technique involves interpolation of hydrological variables to produce a prediction surface based on measurements at known locations. Sahoo *et al*., (2014) used geostatistical techniques for investigation of the spatial variation of groundwater depths in Eastern Odisha, India using data from 24 observation wells. The investigation was performed using the pre and post-moon soon data from 1997 to 2011. From the five semi-variogram models that they considered in their model development, they recommended the exponential model to be the best for spatial interpolation using ordinary kriging interpolation for their study area. It was concluded based on the goodness of fit criteria used which are; Root Mean Square error (RMSE), coefficient of determination (R^2) and Mean Error (ME) for model evaluation.

In a study conducted to investigate the spatial and temporal behaviour of groundwater level in the coastal aquifers of Tamilnadu, Mini *et al*., (2014) used geostatistical modelling. They used data for the pre and post monsoon period of 1999 and 2008. The circular model was found to be the best model for fitting data in unconfined aquifer for all cases except for post monsoon 1999. The spherical model fitted well during 1999 while Gaussian fitted best in year 1998 for semi-confined aquifer. However, unlike Sahoo *et al*., (2004), they were not only able to determine the best fit model, but also the nugget/sill ratio to determine the extent of spatial dependence. The overall result revealed that groundwater level has very strong spatial dependence, and groundwater level was below the mean sea level which resulted in the existence of reverse hydraulic gradient. The study hence concluded that the hydraulic gradient was the main reason behind the intrusion of sea water into the aquifers.

Ahmadi *et al*., (2006) made an attempt to use geostatistics to model the spatial and temporal structure of groundwater level fluctuation using piezometric well data for Darab plain, Iran from 1993 to 2004. After performing cross correlation to determine the most suitable model to construct the semi-variogram, they found that the spherical model provided the best fit and hence it was used for performing ordinary kriging. The outcome of the study was that groundwater level showed strong spatial correlation due to low nugget effects and the model underestimated the groundwater level drop by 3 %. This is an acceptable error therefore supports the unbiasedness hypothesis of kriging. The study was therefore able to prove the capability of geostatistics as a reliable tool for revealing the structure of the groundwater fluctuation in space and time. Taany *et al*., (2007) used geostatistical modelling to analyse the spatial and temporal variation of groundwater level fluctuations in the Amman –Zarqabasin, Jordan using monthly observation well data from 2001 to 2005. They found that the gaussian model provided the best fit for constructing the semi-variogram and this was confirmed by the cross-validation method. After performing interpolation by kriging method, it was found that the interpolation error was less than 5% and the strong spatial dependence was confirmed by the low nugget to sill ratio (<0.25).

In a study conducted by Manchiwal *et al*., (2012), an integrated geostatistics and GIS techniques were used to model the spatial and temporal variations of groundwater levels using monthly ground level data from May 2006 to June 2009 in the semi-arid hard rock basin of western India. They followed a standard procedure of fitting four geostatistical models to the experimental semi-variogram. Based on two goodness of fit criteria (RMSE and correlation coefficient), the exponential model was selected as the best among the four models. In order to identify critical areas suffering from a decline in groundwater levels, spatial interpolation by ordinary kriging was performed in a GIS environment to estimate groundwater elevations at unknown points. A nugget-to-sill ratio of less than 0.25 suggested that groundwater levels have strong spatial dependence, hence geostatistics and GIS were recommended as reliable tools for sustainable management of groundwater resources.

2.6 Summary

In general, stochastic models and geostatistical techniques are very useful in modelling both the spatial and temporal variation of hydrological variables. It is therefore reasonable to apply them for groundwater forecasting and sustainable management of water resources in the Ramotswa wellfield.

CHAPTER 3-METHODOLOGY

3.1 INTRODUCTION

In the previous chapter (i.e. chapter 2), the methods recommended for fulfilling the objectives outlined in chapter 1 are; ARIMA, Thomas-Fiering and Geostatistical methods. This chapter describes those methods and how they were applied starting with two data analysis tools namely intervention and trend analysis. All the proposed models require a single variable, i.e. monthly ground water level series from observation boreholes. This data is available from the Department of Water Affairs sampled from 2002 to 2012.

3.2 INTERVENTION ANALYSIS

As it has already been described above, intervention analysis is a flexible tool for rigorously ascertaining the effects of change upon the mean of a series. As specified under section (2.2) of the literature review, groundwater table data from the observation boreholes in the Ramotswa wellfield were investigated for the possibility of intervention either due to man-made or natural factors. The recommended method was Cumulative Summation (CUSUM) technique which was used to determine the suspected point of intervention. The results were then verified by student t-test. Both methods are described below;

3.2.1 CUSUM Test

This method was developed by Litchfield and Wilcoxon (1949) and it is used to determine whether the means of the two parts before and after the unknown point of change are different (Shanmugasundram, 2012). From a time series data of n observations ($x_1 x_2, x_3 \ldots x_n$), representing monthly groundwater levels from a borehole over n months, the CUSUM value y_i at any time i is given by;

$$y_i = (x_i + x_{i-1} + x_{i-2} + + + \cdots + x_n) - n_i \times \overline{x} \qquad (1)$$

Where; y_i = CUSUM value at time i

n_i= Time scale position of the datum x_i

n = Sample Size

CUSUM values are plotted against time and then graphically investigated for any suspected point of intervention. When there is no intervention in the series, there will be oscillation of the CUSUM value about the horizontal axis. Any drastic rise or decline of the plot would indicate possible intervention from the time of observation of such a change (Parida *et al*., 2008). At the suspected point of intervention, the series is split into two parts for further analysis using the split sample test.

3.2.2 Student's t-test

The t- statistic is given by;

$$T = \frac{|\overline{x_2} - \overline{x_1}|}{\sqrt{\frac{{S_1}^2}{n_1} + \frac{{S_2}^2}{n_2}}} \tag{2}$$

Where $\overline{x_1}$ and $\overline{x_2}$ are means and ${S_1}^2$ and ${S_2}^2$ are variances of split samples 1 and 2 respectively. n_1 and n_2 are sample sizes such that $n_1+n_2=N$. The T-Statistic is tested against the critical values of 1.96 and -1.96 at 5 % significance level. If the value of T lies within that range, then it shows that the intervention is insignificance hence the series belong to the same population. Otherwise there is intervention, i.e. data is non-homogenous.

3.3 Trend Analysis

Groundwater level time series were subjected to a non-parametric Mann-Kendall test. This test was originally used by Mann (1945) and the test statistic distribution was derived by Kendall (1975) (Burn *et al*., 2001). As it has already been explained in section (2.3), non- parametric test does not make any assumption about the underlying distribution of the data (Hipel, 1994). It is therefore advantageous to use it on hydrological data since one does not need to know the type of the distribution.

Mann-Kendall S is given by;

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sgn(x_j - x_k)$$

(3)

Whereby;

$$x_j - x_k = \begin{cases} 1 \ if \ x_j - x_k > \ 0 \\ 0 \ if \ x_j - x_k = \ 0 \\ -1 \ if \ x_j - x_k < \ 0 \end{cases}$$

The variance of S is given by;

$$\sigma_s{}^2 = \frac{n(n-1)(2n+5) - \sum_{j=1}^{q} t_j\,(t_j-1)(2t_j+5)}{18} \tag{4}$$

Where; n is the sample size, q is the number of tied groups in the data set, t_j is the number of data points in the jth tied group.

The test statistic Z_s is computed using the value of s and $\sigma_s{}^2$ as follows;

$$Z_s = \begin{cases} \frac{S-1}{\sigma} \ if \ s > 0 \\ 0 \quad if \ s = 0 \\ \frac{S+1}{\sigma} \ if \ s < 0 \end{cases} \tag{5}$$

When the value of Z is positive, it indicates an increasing trend, while a negative value indicates a decreasing trend. A null hypothesis H_o that there is no trend in the data is either rejected or accepted depending on whether *Zs* is less than or more than the critical value from the normal distribution table at 5% significance level (Kampata *et al*., 2008).

3.4 STOCHASTIC MODELLING

A time series is a set of observations of a phenomena done in a chronological order (Hipel, 1994). The intrinsic feature of time series is that the observations are dependent and it is the dependency of these of observations that is of interest and practical significance (Box and Jenkins, 1976). Therefore a lot of emphasis is placed on the order of the observations for reliable forecasting. If a phenomena that is time dependent can be predicted precisely using laws of physics, then deterministic models can be applied. However, most natural phenomena in hydrology like streamflow, precipitation cannot be predicted with absolute certainty. Randomness/uncertainty in observations of a natural variable can be accounted for by use of stochastic models. Because these observations evolve in time according to a probabilistic structure, a mathematical description of this structure is referred to as a stochastic process (Hipel, 1994).

3.4.1 AR and ARMA models

In this study, the intention was to use ARIMA models to forecast monthly groundwater levels in the Ramotswa wellfield. The model used in this case was a univariate model, i.e. only one variable is used to forecast future values. However, it is important to describe a brief historical background of evolution of these models. Early attempts in time series studies were generally characterized by a deterministic world until the works of Yule (1927) who postulated that every time series can be regarded as a realization of a stochastic process (Gooijer *et al*., 2006). Based on this idea, Autoregressive Model (AR) was introduced given by;

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + a_t \quad (6)$$

In this relationship, an observed value z_t is an output which is dependent on the previous observations namely z_{t-1} and z_{t-2} up to z_{t-p} and the random component a_t. This random component accounts for the error in the model and it's assumed to be normally distributed. Autoregressive model parameters are ϕ_1, ϕ_2 up to ϕ_p. "P" indicates the order of the model (persistency, i.e. the extent of serial dependency). It indicates up to what lag do the previous observations have an influence on the current observations and hence enable determination of values in the next time step.

A number of time series models were subsequently developed. In supplementing AR models, Slusky (1937) came up with a concept of moving average model (MA) given by;

$$z_t = a_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_p \varepsilon_{t-q} \quad (7)$$

θ Represents the moving average parameter in the MA model.

Wold (1938) combined the two theories to form Autoregressive Moving Average (ARMA) models and stated that they could be applied for all stationary time series as long as the order of the AR and that of the MA can be specified appropriately (Makridakis *et al*.,1995). The ARMA models are given by;

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \cdots + \phi_p z_{t-p} + a_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_p \varepsilon_{t-q} \quad (8)$$

This means that the series can be modeled by the past values and the error term. The ARMA model proposed by Wold became a challenge to model real life time series in that it required: i) transformation of original series to make it stationary, ii) specifying the appropriate order of p and q, iii) parameter estimation by nonlinear optimization procedures (Makridakis *et al*.,1995).

3.4.2 Autoregressive Integrated Moving Average (ARIMA) Models

The advent of computers in the 1960's made it possible for computations and optimization of parameters. Box and Jenkins (1970 and 1976) integrated work done by previous researchers and popularized ARIMA models. In most instances in hydrology, observation data may be subjected to trend and periodicity due to seasonal behaviour making it non-stationary in both mean and variance hence cannot be fitted using ARMA models. The success of Box and Jenkins approach was based on its ability to overcome these problems by transforming a time series into stationary series leading to a model requiring few parameters to be estimated in the final choice of the model (Makridakis et al., 1995). In generalizing the model, they obtained a multiplicative Autoregressive Integrated Moving Average (ARIMA(p,d,q)×$(P,D,Q)_w$) model.

It consist of a seasonal ARMA (P,Q) fitted to a Dth-seasonal difference of data coupled with a non-seasonal ARMA(p,d) fitted to the d^{th} difference of the residuals of the former model (Salas *et al*.,1980). The following operations are described below prior to the formulation of the generalized multiplicative ARIMA model;

i) Seasonal AR operator; $\Phi(B^w)Z_t = \left(1 - \Phi_1 B^w - \Phi_1 B^{2w} - \cdots \ldots . . \Phi_p B^{pw}\right)Z_t$ (9)

ii) Seasonal MA operator; $\Theta(B^w)Z_t = \left(1 - \Theta_1 B^w - \Theta_1 B^{2w} - \cdots \ldots . . \Theta_Q B^{Qw}\right)\alpha_t$ (10)

iii) Non Seasonal AR operator; $\phi(B)Z_t = \left(1 - \phi_1 B - \phi_2 B^2 - \cdots \ldots . . \phi_p B^p\right)Z_t$ (11)

iv) Non Seasonal MA operator; $\theta(B)Z_t = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots \ldots\ldots \theta_q B^q\right)\varepsilon_t$ (12)

When the ARMA (P, Q) model is fitted in the Dth seasonal difference of period w, the results is a seasonal ARIMA (P, D, Q) $_w$ model given by;

$$\left(1 - \Phi_1 B^w - \Phi_1 B^{2w} - \cdots \ldots\ldots \Phi_p B^{pw}\right) Z_t (1 - B^w)^D Z_t = \left(1 - \Theta_1 B^w - \Theta_1 B^{2w} - \cdots \ldots\ldots \Theta_Q B^{Qw}\right)\alpha_t \tag{13}$$

When the ARMA (p, q) model is fitted in the dth difference of the α_t series, the result is a non-seasonal ARIMA (p, d, q) model given by;

$$\left(1 - \phi_1 B - \phi_2 B^2 - \cdots \ldots\ldots \phi_p B^p\right)(1 - B)^d \alpha_t = \left(1 - \theta_1 B^w - \theta_2 B^{2w} - \cdots \ldots\ldots \theta_q B^{qw}\right)\varepsilon_t \tag{14}$$

By solving equation (14) for α_t and replacing it in equation (13) , the result is a multiplicative ARIMA(p,d,q)×(P,D,Q)$_w$ model which can be written in a condensed form;

$$\Phi(B^w)\phi(B)(1 - B^w)^D (1 - B)^d Z_t = \Theta_1 B^w \theta(B)\varepsilon_t \tag{15}$$

3.5 BOX JENKINS APPROACH FOR ARIMA MODELLING

This section outlines a systematic approach to hydrological time series modelling and how it was applied for forecasting of groundwater levels in Ramotswa wellfield. The first step was identification of the composition of the model. This step decides whether the model will be univariate or multivariate or a combination of univariate and disaggregation (Salas *et al*., 1980). This identification depends of a number of factors which include characteristics of the whole water resource system and that of the hydrologic time series data. In the case of the Ramotswa wellfield, monthly groundwater level was available at thirteen boreholes, hence a univariate model was used.

Once the form of the model has been identified, the type of the model has to be decided from different alternatives such as AR, ARMA, ARIMA, and disaggregation models. Once again the characteristics of the times series data plays an important role in deciding the type of model to use. Time series data is said to be stationary when its statistical parameters such as mean and variance remain constant with time (i.e. there is statistical equilibrium), otherwise data is non-stationary. For instance, ARMA models are used to model stationary data such as annual flows (Salas *et al*, 1980). However as already mentioned in the proceeding chapter, most hydrological processes such as stream flow and rainfall which are observed daily, weekly or monthly are non-stationary due to the presence of trends and seasonal effects. ARIMA models do account for trend and seasonality in the data through differencing in order to achieve stationarity. The level of differencing which usually varies from 0 to 2 is highly dependent on the level of stationarity of the data (Muhammad, 2012). Since groundwater level data was used in this study, it was more likely to be non-stationary hence ARIMA model was used.

Having recommended a univariate ARIMA model, a three step iterative procedure developed by Box and Jenkins (1976) is described. It consist of three main iterative steps which are;

i) Model Identification
ii) Parameter Estimation
iii) Diagnostic checking(Testing Goodness of fit)

3.5.1 Model identification

The important tools for model identification are the visual inspection of the original series and the behaviour of both the ACF and PACF. This is a stage in which a tentative model is chosen from a family of multiplicative ARIMA(p,d,q)×$(P,D,Q)_w$ models. It involves specifying the orders of the non-seasonal component (p, d, q) and those of the seasonal component (P, D, Q). w represents seasonality which could be days, weeks, or months. p, d and q are autoregressive, required differencing, moving average respectively for the non-seasonal component, while P,D,Q represent the autoregressive, differencing and moving average respectively for seasonal component. Model identification is done by plotting graphs of ACF (correlogram) and PACF from the series. The ACF indicates the extent of persistence (serial dependence) within the series for lags 1, 2, 3... up to lag K. It is given by;

$$r_k = \frac{\sum_{t=1}^{n-k}((x_t-\overline{x})(x_{t+k}-\overline{x}))}{\sum_{t=1}^{n}(x_t-\overline{x})^2} \qquad (16)$$

where; r_k is the autocorrelation function and *k* is the lag. The partial autocorrelation function (PACF) also indicate the extent of serial correlation at a given lag after accounting for the correlation from intervening lags. In addition to the ACF's and PACFs, the 95% interval limits are plotted for each lag. They indicate the significance of the correlation. The first serial correlation coefficient r_1 is currently the most useful measure of time dependence of a series (Salas *et al*., 1980). The 95% confidence limits for corellogram are given by;

$$r_k = \frac{-1 \pm 1.96\sqrt{N-k-1}}{N-k} \tag{17}$$

The order of AR (p) is determined by the number of positive successive spikes at lags greater than zero which are significant, while the order of the MA(q) is determined by the number of successive negative spikes at lags greater than zero which are significant. When the spikes are within the 95 % confidence limits, the correlation is considered insignificant indicating the absence of serial dependence between the observations.

3.5.2 Parameter Estimation

Following the identification of a tentative model, parameters such as ϕ, Φ , θ and Θ are estimated and their number depends on the order of the model and the characteristic of the series being studied. These can be estimated at three levels of increasing accuracy (Salas *et al*., 1980);

i) Preliminary estimates using Yule -Walker equation
ii) Maximum likelihood method-Steepest descent algorithm may be used to obtain the maximum likelihood of parameters and residuals
iii) Non-linear estimation-By using a Taylor series expansion

The parameters estimated should not be highly correlated. This can be checked through a correlation matrix among the AR and MA parameters. A model that has highly correlated parameters is not robust to changes in the data hence not suitable for forecasting (Reagan, 1984). In addition to that, the parameters must be significantly different from 0. Parameters which are not significantly different from zero must be removed from the model. When all these conditions are met, the analysis should proceed to the next stage, otherwise the parameters have to be re-estimated.

3.5.3 Goodness of fit/Diagnostic Stage

In the diagnostic stage, the modeller has to ensure that the modelling assumptions of normality and independence of residuals are verified, i.e. the model is parsimonious. A parsimonious model is the one which contains a minimum number of parameters to achieve white noise residuals. For instance, for an AR (P) model, the residual can be written as;

$$\varepsilon_t = z_t - \phi_1 z_{t-1} - \phi_2 z_{t-2} - \cdots - \phi_p z_{t-q} \tag{18}$$

This yields residuals $\varepsilon_1, \varepsilon_2 \ldots \varepsilon_n$ which would be tested for independence and normality. This can be achieved by;

i) Corellogram; Plotting of ACF and probability limits of residuals and verifying that they are white noise. At any lag, the white noise residuals are not supposed to show any significant correlation, i.e. they have to lie within the 95 % confidence limits

ii) Ponte Manteau Lack of Fit test; The Q statistic is given by;

$$\mathrm{Q} = N \sum_{k=1}^{L} r_k{}^2\,(\varepsilon) \tag{19}$$

where r_k is the autocorrelation of residuals and L may be of order between 10 and 30% of sample size N. Q is approximately χ^2 (L-P). If Q < χ^2 (L-P), then the residual is independent and the model is adequate, else a different model needs to be tested (Salas *et al*., 1980).

Normality test can be done by plotting the empirical distribution of residual on a normal distribution paper and verify whether the plotted points lie in a straight line. The modeller also has to ensure that that between competing models, the right order of the model is selected. Akaike (1974) came up with such a criteria which considers the principle of parsimony to ensure that the fitted model is more adequate compared to others. For instance, for ARMA (p, q) model,

$$\mathrm{AIC} = N ln(\sigma_\varepsilon{}^2) + 2(p+q) \tag{20}$$

where *N* is the sample size and $\sigma_\varepsilon{}^2$ is the residual variance. The model that gives a minimum AIC is the one that is selected. Once the candidate model passes the goodness of fit stage, it can now be used for forecasting. Otherwise the procedure has to be repeated again until an adequate model is formulated.

The summary of the iterative systematic approach for ARIMA modelling by Box and Jenkins (1970) is shown on Figure 3.1. This approach was applied for modelling and forecasting of groundwater level time series in the Ramotswa wellfield.

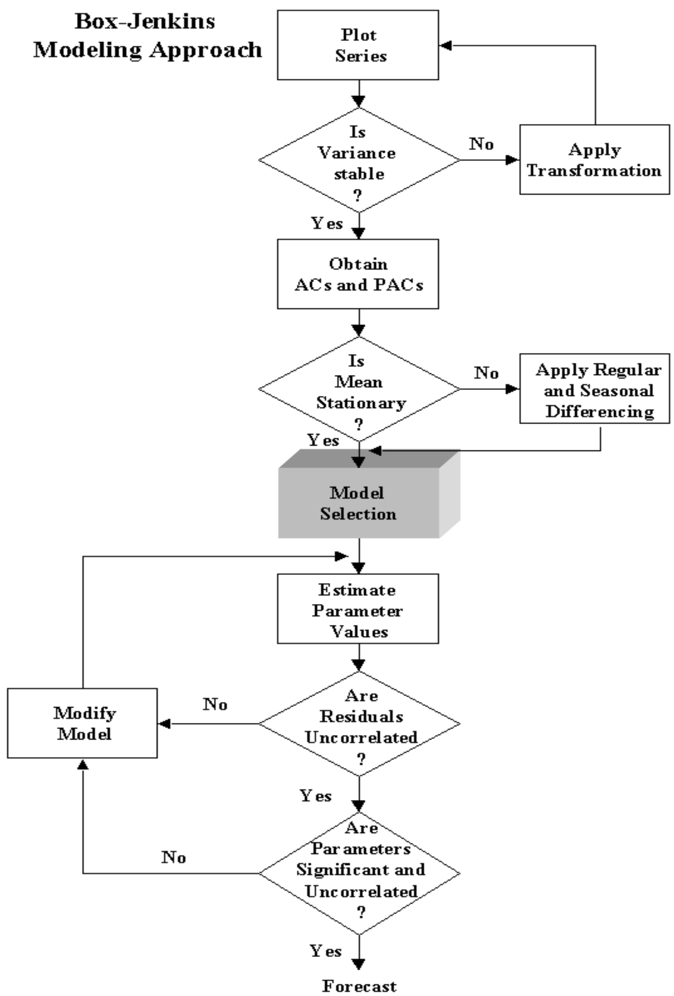


Fig 3.1 Box Jenkins approach

(http://home.ubalt.edu/ntsbarsh/business-stat/stat-data/forecast.htm)

3.6 THOMAS-FIERING (T-F) MODEL

Non stationary time series data brings a lot of complications in development of mathematical models for hydrological applications. Thomas and Fiering (1962) came up with a mathematical model which allows for non-stationary in observation data for synthetic generation of monthly flow sequences. Since then, a lot of researchers have applied this model for forecasting of hydrological variables especially stream flow and water quality data. However, in this study, an attempt was made to model and forecast monthly groundwater levels data using T-F model. The model is of Markovian nature with periodic parameters such as monthly mean, standard deviation, and the lag zero cross correlation between successive months (Subagadis, 2009). In its simple form, the model consists of twelve regression equations, one for each month of the year. Thomas Fiering model is given by;

$$x_{i,j} = \overline{x}_j + b_j\left(x_{i,j-1} - \overline{x}_j\right) + z_i s_i \sqrt{(1 - r_j{}^2} \quad (21)$$

For each month, (j= 1, 2, 3……………up to j = 12)

$\overline{x}_j$ is the mean given by; $\overline{x}_j = \frac{\sum_{i=1}^{N} x_{ji}}{N}$ (i = j, 12+ j, 24+ j…………….) (22)

S is the standard deviation given by; $S_j = \sqrt{\sum_{i=1}^{N} \frac{(x_{ji} - \overline{x_j})^2}{(N-1)}}$ (23)

r_j is the correlation coefficient with the proceeding observation given by;

$$r_j = \frac{\sum_{i=1}^{N}(x_{ji} - \overline{x_j})((x_{j+i,i} - \overline{x_{j+1}})}{\sqrt{\sum_{i=1}^{N}(x_{j,i} - \overline{x_j})^2 \sum_{i=1}^{N}(x_{j+1,i} - \overline{x_{j+1}})^2}} \quad (24)$$

b_j is the slope of the regression equation relating months observation to observation in the preceding month given by;

$$b_j = \frac{r_j S_{j+1}}{S_j} \quad (25)$$

z_i is the random normal deviate of mean and unit standard deviation.

From the Thomas –Fiering model (equation 21) the first term represents the mean, the second term represents the regressed component on the previous observation, the third term is the random component. The model accounts for the persistence up to lag 1, therefore may be regarded as a non-stationary first order Autoregressive model (Subagadis, 2009).

3.7 GEOSTATISTICAL MODELLING

Geostatistical modelling is an interpolation technique that is based on statistical relationships among measured points in space. It has many applications in hydrology which includes delineation of critical areas where measures like controlled pumping, artificial recharge, pollution control, have to be implemented. In this study, geostatistical method was used to determine the spatial distribution of groundwater table during the dry and wet season. This was achieved by obtaining a prediction surface of groundwater level based on the available data from 21 observation wells and determination of accuracy of those predictions.

A very important tool in geostatistical analysis is the semi-variogram. It expresses the spatial dependence between neighbouring observations separated by a distance h (Ahmadi *et al*., 2007). A semi variogram is defined as half of the variance of the difference between attribute values of all points separated by h and can be determined using the following relationship;

$$\gamma(\mathrm{h}) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \qquad (26)$$

Z(x) is the magnitude of the variable, *N (h)* is the total number of pairs of attributes that are separated by a distance h. The important semi -variogram parameters are; sill, nugget and the range as shown by the graphical representation below;

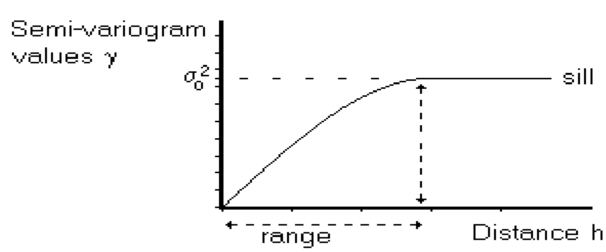


Fig 3.2: Semi Variogram with its parameters

Source :(http://spatialanalyst.net/ILWIS/htm/ilwismen/graph_window_add_semivariogram_model.htm)

Different semi variogram models are;

i) Exponential Model

$$\gamma(h) = C_0 + C_1\left\{1 - exp\left(-\frac{3h}{a}\right)\right\} \tag{27}$$

ii) Gaussian Model

$$\gamma(h) = C_0 + C_1\left\{1 - exp\left(-3\left(\frac{h}{a}\right)^2\right)\right\} \tag{28}$$

iii) Spherical Model

$$\gamma(h) = \begin{cases} C_0 + C_1\left\{\frac{3}{2}\frac{h}{a} - \frac{1}{2}\left(\frac{h}{a}\right)^3\right\} & 0 \leq h < a \\ C_0 + C_1 & a \leq h \end{cases} \tag{29}$$

where; C_0 is the nugget effect which shows the non-spatial variance, C_1 is the spatial variance and a is the range which corresponds to the separation distance at which the spatial autocorrelation diminishes. The following are four important steps in geostatistical modelling;

i) Determination of experimental semi-variogram using the available data
ii) Fitting of different semi-variogram models to the experimental one
iii) Checking the adequacy of the model based on statistical measures, i.e. cross validation. In cross validation, interpolated and actual values are compared the model that yields the most accurate prediction is retained. Cross validation method used is;

$$ME = \frac{1}{N}\sum_{i=1}^{N}(z^{*}(x_i) - z(x_i)) \tag{31}$$

where $z^*(x_i)$ and $z(x)$ are estimated value and observed value respectively.

Interpolation comes after the most accurate model has been selected as described above. In this dissertation, ordinary kriging method was be used. A brief description of this interpolation technique and how it is used to estimate variables at unknown locations based on the measure variables at known locations is presented below;

Ordinary Kriging model*:* $z(s) = \mu + \varepsilon(s)$ (32)

μ is the mean for the data (no trend) and $\varepsilon(s)$ is random errors with spatial dependence. The predictor at a given location S_0 is given by;

$$z^*(s_0) = \sum_{i=1}^{N} \lambda_i \, z(s_i) \quad (33)$$

S_0 is the prediction location, while is λ_i weight for measured value at location i. λ_i depends on semivariogram, distance to prediction location, and the spatial relationship around the prediction location. To have an unbiased estimate, the difference between the prediction and actual value should be as small as possible hence minimize the statistical expectation given by;

$$(z(s_0) - \sum_{i=1}^{N} \lambda_i \, z(s_i))^2 \quad (35)$$

CHAPTER 4-TREND AND INTERVENTION ANALYSIS

4.1 INTRODUCTION

As stated in section 1.4 the main objective of this study is to model the spatial and temporal variation of groundwater level in the wellfield using monthly groundwater level data from observation boreholes. This data was collected during a monitoring period between 2002 and 2012. Before the modelling part, this chapter covers the following aspects;

- Data availability
- Time series hydrographs of during the monitoring period
- Intervention analysis
- Trend analysis

4.2 SOURCE AND AVAILABILITY OF DATA

Modelling of the spatial and temporal fluctuation of groundwater level at Ramotswa wellfield relied on the availability of reliable groundwater level data since all the models used are data driven models. This data was obtained from Department of Water Affairs (DWA), Groundwater Division from 21 observation boreholes. The data was taken on monthly basis during a 10 year monitoring period (i.e. from 2002 to 2012). The length of the monitoring period differs slightly between individual boreholes. For stochastic modelling and forecasting, 13 of the 21 observation boreholes shown on table 4.1 were used based on their continuous records of data during the monitoring period. For geostatistical modelling, all the 21 boreholes were used for analysing spatial groundwater level scenario during wet and dry seasons. Detailed spatial and temporal modelling has been presented under chapter 5. There were some missing records due to blockage of some boreholes and inaccessibility due to flooding during rainy seasons. The missing records encountered were filled by linear interpolation on SPSS. After subjecting data to intervention analysis, the time of intervention was determined for each borehole data series. The data after the time of intervention was used for trend analysis and forecasting. The table 4.1 below shows the locations of observation boreholes used for both spatial and temporal analysis.

Table 4.1 Observation boreholes and their locations

Borehole	Latitude	Longitude	Period	Length of Records (Month)	Remarks
BH4155	25.87	-24.89			Too many gaps
BH4160	25.88	-24.9	2005-2012		Too many gaps
BH4163	25.88	-24.86	2002-2012	130	Reliable data
BH4164	25.88	-24.89	2002 -2012	130	Reliable data
BH4165	25.87	-24.85	2003-2012	102	Reliable data
BH4168	25.86	-24.88			Too many gaps
BH4341	25.88	-24.88	2005-2012	93	Reliable data
BH4348	25.88	-24.9	2002 -2012	131	Reliable data
BH4371	25.89	-24.88	2002 -2012	130	Reliable data
BH4885	25.88	-24.92	2002-2012		Too many gaps
BH4886	25.87	-24.84	2002 -2012	129	Reliable data
BH4887	25.88	-24.85	2002 -2012	127	Reliable data
BH4972	25.88	-24.91	2003-2012		Too many gaps
BH4973	25.87	-24.91	2002-2012	129	Reliable data
BH4974	25.87	-24.91	2004-2012		Too many gaps
BH4995	25.87	-24.84	2002-2012	129	Reliable data
Z 4401	25.87	-24.9	2005-2012	93	Reliable data
Z 6423	25.86	-24.86	2002-2012	126	Reliable data
BHZ 6424	25.84	-24.89	2002 -2012	132	Reliable data
BH 6501	25.88	-24.92	2002-2012		Too many gaps
BH10128	25.88	-24.87	2005-2012		Too many gaps

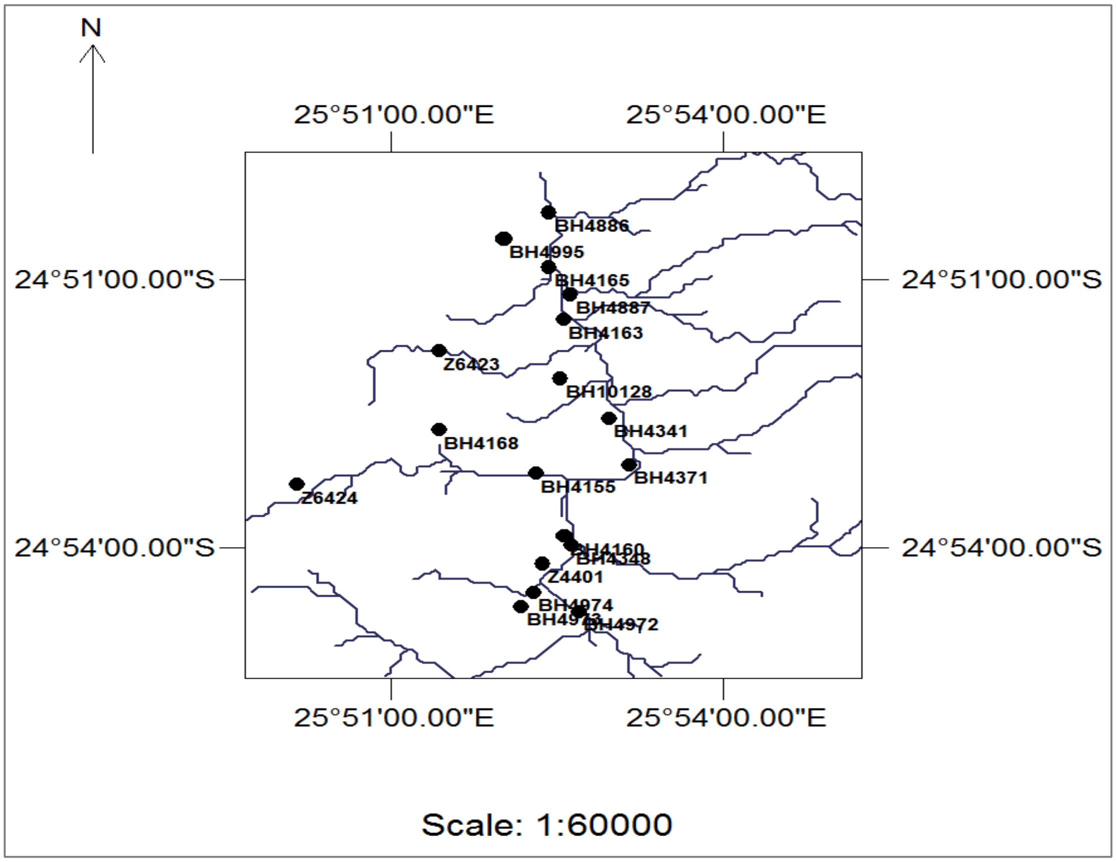


Fig 4.1-Map showing location of observation boreholes in Ramotswa wellfield

4.3 GROUNDWATER LEVEL HYDROGRAPHS FOR VARIOUS BOREHOLES

Ramotswa wellfield was abandoned in 1996 due to groundwater pollution mainly from pit latrines. Work has been done recently to refurbish the wellfield and utilize some of the selected production borehole water as a drought mitigation strategy. During this monitoring period from 2002 to 2012, there was no abstraction of water from production boreholes. This implies that groundwater level series used in this study are static levels and were not influenced by pumping. Therefore it can be safely assumed that water table response was mainly due to natural hydro-meteorological factors and changes in climatic conditions. The groundwater level and monthly rainfall series during the 10 year monitoring period for 13 boreholes are shown by the 3 graphs on Figures 4,2 a) b) and c).

4.3.1 Discussions

The graphs show the groundwater level hydrographs with the superimposed monthly rainfall hydrographs. The boreholes in Figures 4.2a and 4.2b are generally shallower with groundwater levels ranging between 2 and 7 m. Boreholes in Figure 4.2c are deeper with groundwater levels between 7 and 25 m. The following observations can be made from these hydrographs;

i) A general decline in groundwater levels between years 2002 and 2006 and the rate of decline decrease significantly between 2006 and 2012. In fact the hydrographs appears to stabilize after year 2006.

ii) There were 4 significant rainfall events in January 2006 (209 mm), December 2007 (129mm), January 2009(219 mm) and January 2010 (304.5 mm).

iii) In all the cases, the periods of high rainfall events coincided with a significant rise in groundwater levels.

Since these water levels were taken during periods of no pumping due to shutting down of production boreholes, the following points can be used to justify the above observations;

i) The response in groundwater levels is similar in all the boreholes i.e. periodic discharges and recharges charges are similar but different in magnitude. The steady decline in groundwater level between 2002 and 2005 can be directly linked to a period of lower average monthly rainfall (i.e. less than 100 mm). The recovery of groundwater levels between 2006 and 2012 can be linked to an increase in monthly rainfall especially due to the four periods of significant rainfalls mentioned above. Moreover, it can be generally concluded that the similar behaviour in terms of response to rainfall implies that the hydrogeological conditions underneath are similar. In terms of depth of groundwater table, boreholes which are shallower are closer to the Notwane river as shown in drainage map on Figure 4.1 while those further away are deeper (e.g. BHZ6424). This shows that groundwater flow is highly influenced by topography and it is flowing in the north and north eastern direction towards the Notwane River. The topographical effect was also confirmed by geostatistical results in chapter 5.

ii) The significant and immediate response due to heavy rainfall events which caused flooding in 2006 and 2010 can be described in terms of the hydrogeology of the area. As it has been already stated in section 1.3.5, the wellfield sits on high yielding aquifer system of Ramotswa Dolomite and Lephala formations shales. However most of the boreholes are located in the dolomite aquifer (Wellfield Consulting Services, 2011). Ramotswa Dolomite is a karst aquifer consisting of a network of interconnected fractures and water collecting channels. One of the most distinctive features of karst aquifers is their capacity to be filled and emptied fast due to large fracture channels, good interconnection and high permeability of surface zones (Milanovic *et al*,2004). During the rainy season, there are vertical changes in groundwater level hydrographs while dry season is a recessive period in which water table is constantly lowering. Ramotswa observation boreholes groundwater hydrographs hence demonstrate a classical representation of a karst aquifer indicated by the vertical response of groundwater level during wet season and water level during dry season. Figures 4.3a) and 4.3 b) taken at the South Eastern part of the study area during a site visit give an indication of the hydrogeological conditions and the sloping terrain which supports high recharge rates during rainy season.

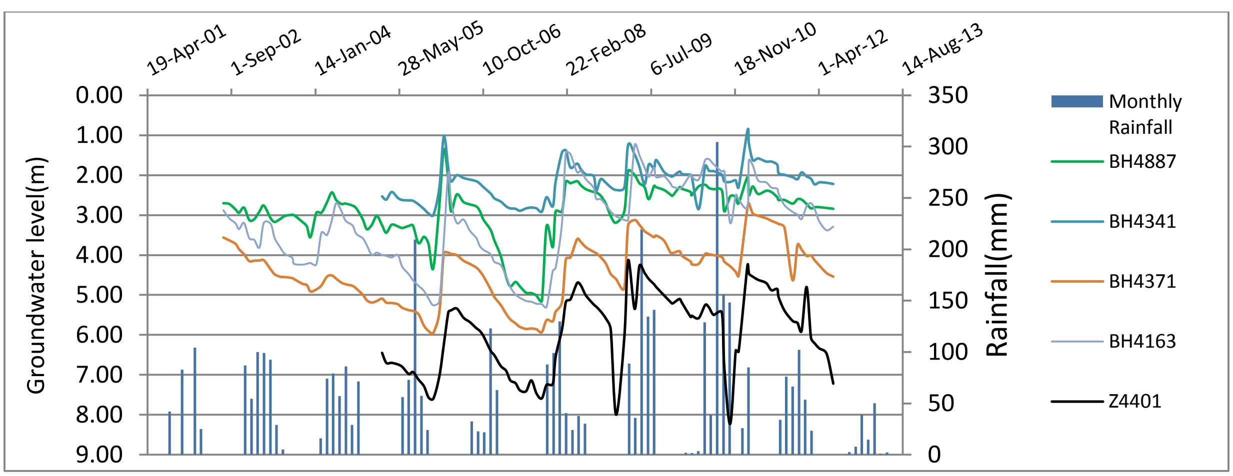


Fig 4.2 a) Groundwater level and rainfall hydrographs

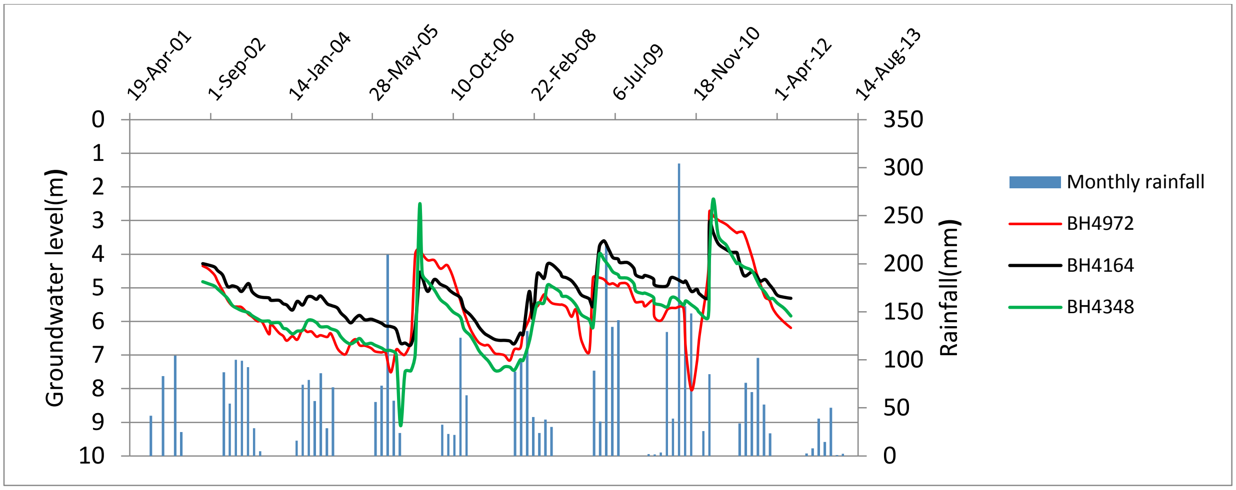


Fig 4.2 b) Groundwater level and rainfall hydrographs

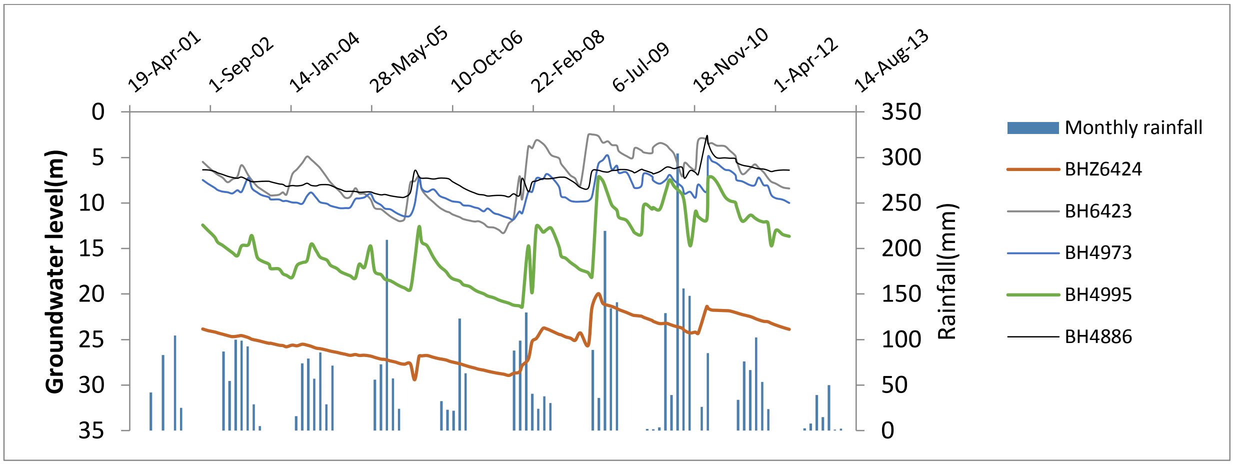


Fig 4.2 c) Groundwater level and rainfall hydrographs

Fig 4.3 a) Karst landscape is common throughout the study area

Fig 4.3 b) Terrain slopes in the west-east direction towards the main stream

4.4 GROUNDWATER LEVEL INTERVENTION ANALYSIS

The importance of understanding the characteristics of data and analysing the effect of natural and man-made changes in hydrological time series was described in section 2.2. In this section, intervention analysis on groundwater level series at various boreholes shown in Table 4.2 for the study period is presented. The first presentation is the results from the graphical approach in which CUSUM values are plotted against time. Fig 4.4 gives an indication of the suspected time of intervention for BH4341. The rest of the plots are shown appendix A for other boreholes. As already described under section 3.2.1, the time series is free from intervention when the CUSUM values oscillates about the axis (Parida *et al.*, 2008). The groundwater level data was further analysed using student t-test (split sample test) described under section 3.2.2. The reason behind this further analysis was to ascertain whether there was intervention in data or not. For this test, the samples were split into two groups at the suspected time of intervention and then it was determined whether the two samples are statistically different at 5 % significance level. The results of the student's t-test are summarized on Table 4.2.

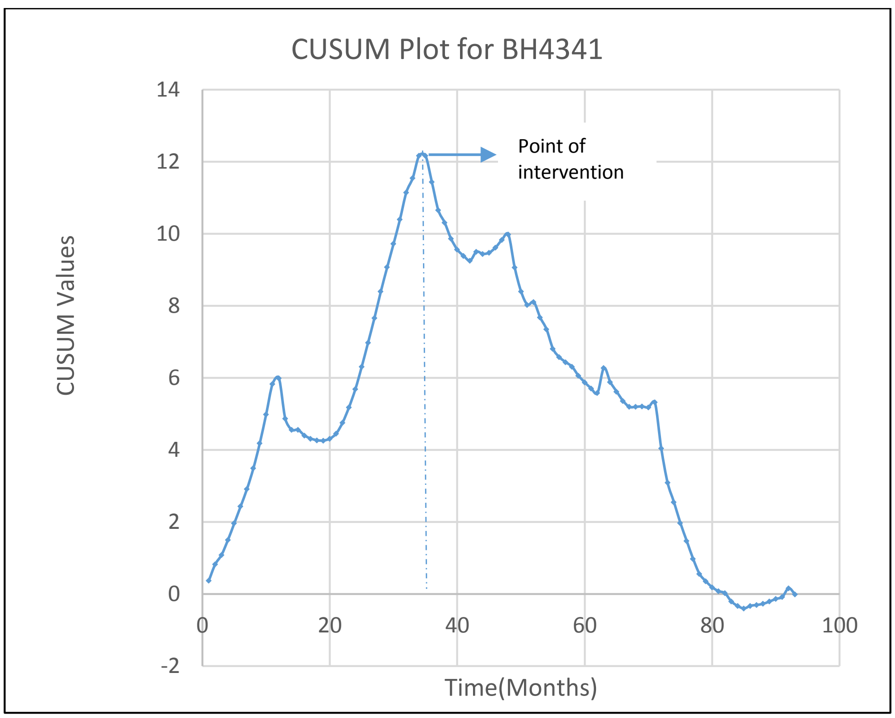


Fig 4.4 CUSUM Plot using observed GWL for BH 4341(2005-2012)

Table 4.2 Results of Student t-test

Boreholes	Student t-Test	Critical Value	Result	Time of Intervention
BH4341	7.09	1.96	Significant Intervention	Nov-07
BH6423	9.13	1.96	Significant Intervention	Nov-07
BH4887	-5.64	1.96	Significant Intervention	Sep-07
BH4973	5.06	1.96	Significant Intervention	Sep-07
BH4165	8.89	1.96	Significant Intervention	Sep-07
BH4164	6.75	1.96	Significant Intervention	Nov-07
BH4163	10.24	1.96	Significant Intervention	Oct-07
BHZ6424	12.72	1.96	Significant Intervention	Jan-08
BH4886	10.72	1.96	Significant Intervention	Dec-08
Z4401	35.4	1.96	Significant Intervention	Dec-07
BH4371	5.67	1.96	Significant Intervention	Oct-07
BH4348	6.08	1.96	Significant Intervention	Dec-07
BH4995	8.19	1.96	Significant Intervention	Nov-07

4.4.1 Discussion

Figure 4.4 shows a steady rise in the CUSUM values between years 2002 and end of year 2007. This rise indicates a period in which the groundwater table in almost all the 13 boreholes was generally declining. The steady decline in the CUSUM plot was evident between year 2007 and 2012. This indicates a period in which there was generally a reduction in the rate of decline of groundwater levels hence it can be interpreted as a recovery period. The time of intervention for most boreholes according to the CUSUM plot was detected at the end of 2007. This needed to be confirmed by the results of the split sample test at 5 % significance level. The results of the split sample test presented in table 4.2 shows that there was significant intervention in groundwater level data for all the boreholes. The t- statistics computed are outside the range of -1.96 and 1.96 at 5 % significance level. The time of intervention of groundwater level series for most boreholes is at the end of year 2007 except for BHZ6424 and BH4886 which are January and December 2008 respectively. This implies that in all the 13 boreholes analysed in the wellfield, their groundwater level time series data do not come from the same population (i.e. the data is not homogenous). It is for this reason that groundwater data after intervention was used for modelling as shown in the next section.

4.5 GROUNDWATER LEVEL TREND ANALYSIS

Mann Kendall test which was presented in section 3.3 was the method used for trend analysis at the selected boreholes. Trend analysis of groundwater levels provides valuable information in describing hydrological dynamics within the aquifer system due to interaction of different surface and subsurface processes. It also helps in identifying future challenges that are likely to arise due to different stresses in the groundwater system of Ramotswa wellfield. Trend analysis was conducted using data after the time of intervention, therefore it is of relevance to the current climatic conditions of the study area. The Mann Kendall test results for 13 boreholes have been summarized and presented in Table 4.3.

Table 4.3 Trend Results at various observation boreholes

Borehole	Mann Kendall S	Variance of S	Test Statistic Zs	Critical Value	Result
BH4341	277	21076	1.91	1.96	Trend is not significant
BH6423	264	20018	1.86	1.96	Trend is not significant
BH4887	370	24557	2.35	1.96	Trend is significant
BH4973	119	23377	0.77	1.96	Trend is not significant
BH4165	-638	16984	-4.89	1.96	Trend is significant
BH4164	348	24571	2.21	1.96	Trend is significant
BH4163	604	24577	3.85	1.96	Trend is significant
BHZ6424	-42	27101	-0.25	1.96	Trend is not significant
BH4886	-58	9754	-0.6	1.96	Trend is not significant
Z4401	509	22220	3.41	1.96	Trend is significant
BH4371	242	27101	1.46	1.96	Trend is not significant
BH4348	286	23374	1.86	1.96	Trend is not significant
BH4995	-130	23378	-0.84	1.96	Trend is not significant

4.5.1 Discussion

The null hypothesis (Ho) that there is no trend in the data is either accepted or rejected depending on the computed test statistic Z_s (Kampata *et al*., 2008). From the table above, positive Z values are an indication that groundwater levels are declining with respect to the ground surface while the negative values indicate an increase in groundwater levels. On the basis of this description, it can be seen that 9 of the boreholes are showing a decreasing trend in groundwater levels as indicated by the positive values of Z_s. Of those nine boreholes showing a decrease in groundwater levels, trend is statistically significant at 4 boreholes which are BH4887, BH4164, BH4163, and Z4401. This is indicated by the Z_s values which are outside the critical Z-table limit of ±1.96 at 5 % significance level. An increase in groundwater level is evident at 4 boreholes as indicated by the negative Z_s values (i.e. BH4165, BH4886, BHZ6424 and BH 4995). Among these boreholes, only one borehole series is showing a statistically significant increasing trend (i.e. BH4165).

4.6 Summary

The CUSUM findings have revealed that there was significant change in groundwater level series at all the boreholes considered in this study. In most of the boreholes, change in data was detected at the end of year 2007. These changes were confirmed by split sample analysis in a manner described in section 3.2.2 in which the t-statistic was tested against a specified critical value of 1.96 at 5 % significance level. This change in groundwater level data is likely to be due to change in hydroclimatic conditions especially the significant recharges due to rainfall events that caused flooding in years 2006, 2007, 2009 and 2010 as highlighted in section 4.3.1. In terms of trend analysis, trend is statistically significant at 5 boreholes (BH4887, BH4165, BH4164, BH4163 and Z4401) while it is insignificant at the remaining 8 boreholes (61 % of monitoring boreholes). Therefore it could be concluded that the period after year 2006 was a recovery period of groundwater levels and this can justified by the fact that most boreholes were not showing significant trends compared to a period of declining levels before 2006.

CHAPTER 5- STOCHASTIC AND GEOSTATISTICAL MODELLING

5.1 INTRODUCTION

In this chapter, modelling results from two stochastic models for groundwater level forecasting at 13 observation boreholes in Ramotswa wellfield are presented. Data taken after the time of intervention was used for model development. As already stated, data after the time of intervention is considered relevant to the current climatic conditions of the study area and hence it is very useful in water resources planning and management. The results from the best fitted model were then used for three months forecasting. This investigation enables recommendation of the best model between these two linear stochastic models for forecasting of groundwater level. The first part of the chapter focuses on describing fitting of ARIMA model followed by Thomas-Fiering model. The last part (i.e. Section 5.5) presents spatial modelling of monthly groundwater levels for dry season scenario (July 2005) and wet season (February 2006).

5.2 ARIMA MODELLING

The detailed Box-Jenkins approach for ARIMA modelling has been presented under section 3.5 and summarized under Figure 3.1. This iterative approach was used for modelling groundwater level fluctuation. One of the most important requirements for using this approach is that data is normally distributed. In boreholes where data was not normally distributed, data had to be transformed to normal using logarithmic transformation method.

ASTSA software was utilized for necessary computations such as Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), parameter estimation and AIC. The ACF and PACF plots are necessary to determine the possible persistence structure of the data and hence identifying the order of the model (Ahmed *et.al,* 2001). In order for ASTSA to give reliable results, correct inputs of possible model orders, differencing and the presence of seasonal components are very important. Therefore the ACF and PACF had to be carefully analysed. After the inputs of the possible model orders and differencing (p, d, q, P, D, and Q) as defined in chapter 3, ARIMA search was conducted to identify all possible models for a particular borehole series.

As stated by Hipel *et al*., (1994), the basis of model building is to keep the model as simple as possible but the model should provide a good fit to the modelled data. Different selection criteria such as Akaike Information Criterion (AIC) and Bayes Information Criterion (BIC) are available. Both criterions use the principle of parsimony to ensure that the fitted model is the most adequate from a range of other possible models. For this study, AIC was used for selecting the best model based on the model giving minimum AIC.

The following is a presentation of a summarized procedure and results obtained with an aid of ASTSA software for BH4341. The same procedure was used for the other 12 boreholes in the wellfield. The ACF s and PACFs of data were obtained and are graphically presented below (Fig 5.1 and Fig 5.2).

This was the first step in the model identification process in which all possible candidate models were identified. It was on the basis of these plots that the possible orders of the model were determined as inputs for conducting an ARIMA search shown on Fig 5.3 below. For BH 4341, it can be seen from Fig 5.1 and 5.2 that the series is stationary. The significant spike at lag 1 for both plots is an indication of significant autocorrelation at lag 1. There were no other significant correlations from lag 12 and beyond, an indication that there was no seasonal component. The absence of a significant moving average is also clear for all the lags except at lag 9 and 10. This analysis of the ACF and PACF plots informed what values to use as inputs in ASTSA.

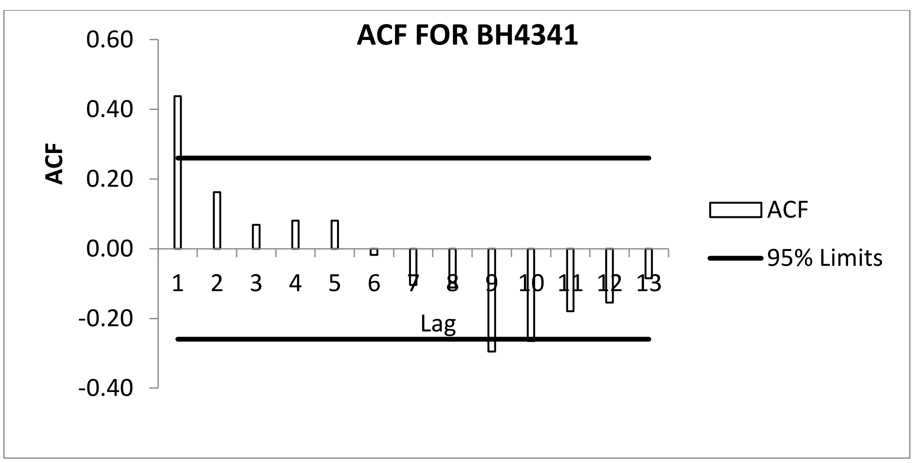


Fig 5.1 ACF Plot for BH4341

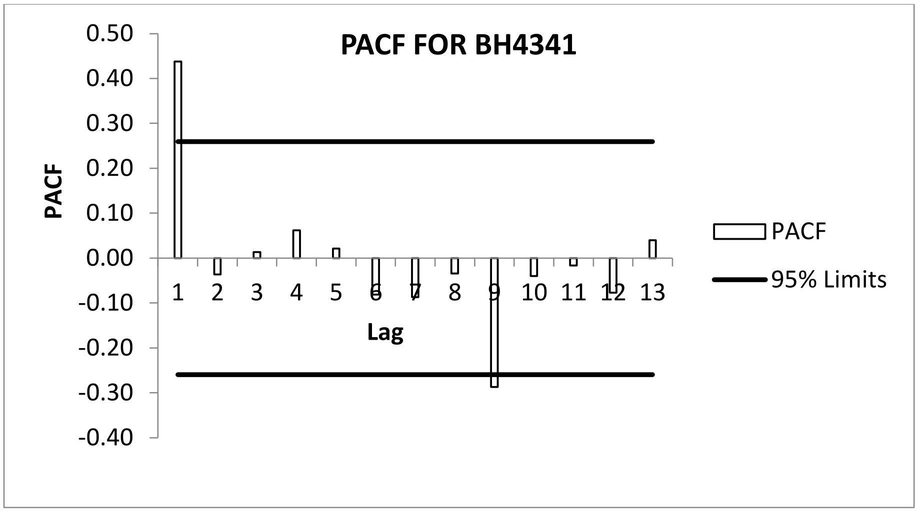


Fig 5.2 PACF Plot for BH4341

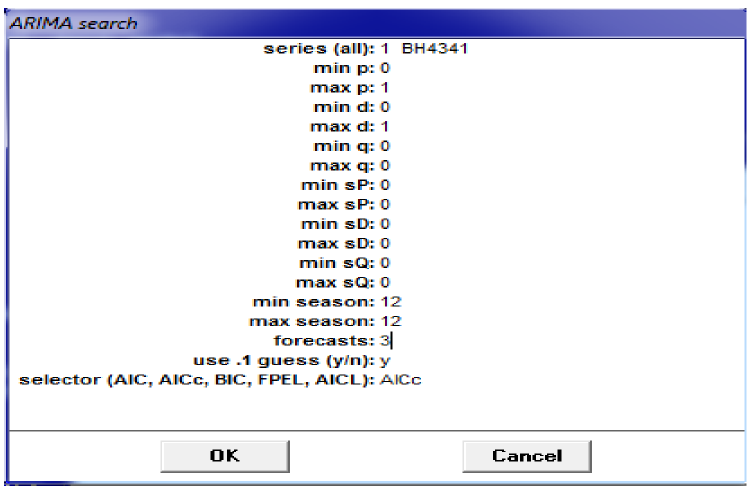


Fig 5.3 ARIMA search window on ASTSA

From Fig 5.3, the maximum possible Autoregressive (AR) on the basis of ACF and PACFs was 1. Non-seasonal differencing (d) ranged between 0 and 1. There was no Moving Average (MA) for the non-seasonal part. All the seasonal components were not significant because all the correlations were lying within the 95% confidence limits for lag 12 and beyond as shown on Fig 5.1 and Fig 5.2.

Table 5.1 Comparisons of candidate models for BH4341

ARIMA search on BH4341							
p	d	q	P	D	Q	s	AICc
0	0	0	0	0	0	12	.2195
0	1	0	0	0	0	12	-2.1719
1	0	0	0	0	0	12	-2.1474
1	1	0	0	0	0	12	-2.1912 best model

From Table 5.1 that four possible models were under consideration. The best model for BH 4341 was ARIMA (1,1,0) with a minimum AICC of -2.19 as highlighted. For this case, estimation of model parameters was done as shown in table 5.2 below. The model parameters were calculated using ASTSA software package.

Table 5.2 Parameter estimation results for BH4341

Model	Autoregressive Parameter	Moving Average Parameter	Variance	Standard Error
ARIMA(1,1,0)	0.271	No Moving average	0.39	0.13

5.2.1 ARIMA Forecasting

For all the groundwater level time series at each borehole, the last 3 monthly observed records were used for model verification. This was done by computing the forecasting error using the developed model and the actual observed data. The forecasting errors for the 3 months forecast for BH4341 are presented on Table 5.3 below.

Table 5.3 ARIMA forecast errors for BH4341.The rest are displayed under Appendix E

Borehole	Model	Forecast Time	Forecast Error (%)
BH4341	(1,1,0)	October 2012	3.6
		November 2012	12.9
		December 2012	4.02

Table 5.4 Summary of ARIMA results for all the boreholes

Borehole	Model	Parameters	Observed GWL(m)	Forecast (m)	95% Limits Lower	95% Limits Upper	AIC
BH4341	(1,1,0)	AR(1) = 0.27	Oct 2012 = 2.21	Oct 2012 = 2.13	1.513	3.274	-2.19
			Nov 2012 =2.40	Nov 2012 =2.09	1.38	3.59	
			Dec 2012 =1.99	Dec 2012 = 2.07	1.26	3.93	
BH6423	(1,0,0)	AR(1) = 0.97	Sept 2012 =8.39	Sept 2012 =8.25	6.09	10.39	1.24
			Oct 2012 =8.36	Oct 2012 =8.22	5.18	11.26	
			Nov 2012 = 8.42	Nov 2012 =8.20	4.48	11.91	
BH4887	(1,1,0)	AR(1) = 0.24	Sept 2012 =2.84	Sept 2012 = 2.82	2.26	3.51	-3.33
			Oct 2012 =2.86	Oct 2012 =2.82	2.14	3.71	
			Nov 2012 =3.03	Nov 2012 =2.82	2.03	3.91	
BH4973	(2,0,0)	AR(1) =1.02	Aug 2012=9.64	Aug 2012=10.07	8.10	12.58	-3.26
		AR(2) = -0.02	Sept 2012 =10	Sept 2012 = 10.18	7.45	14.01	
			Oct 2012 =10.01	Oct 2012 =10.35	6.98	15.34	
BH4165	(1,0,0)	AR(1) = 0.98	Dec 2011 = 3.29	Dec 2011 = 3.2	2.34	4.38	-2.57
			Jan 2012=3.45	Jan 2012=3.13	2.01	4.90	
			Feb 2012=3.25	Feb 2012=3.09	1.79	5.30	
BH4164	(1,0,0)	AR(1)=0.99	Oct 2012 =5.46	Oct 2012 =4.77	4.32	6.64	-3.37
			Nov 2012 =5.63	Nov 2012 =4.75	3.95	7.25	
			Dec 2012 =5.54	Dec 2012 =4.72	3.69	7.75	
BH4163	(1,0,0)	AR(1)= 0.96	Oct 2012 =3.38	Oct 2012 =3.33	2.13	5.14	-1.95
			Nov 2012 =3.49	Nov 2012 =3.18	1.73	5.82	
			Dec 2012 =2.96	Dec 2012 =3.04	1.47	6.29	
BHZ6424	(1,0,0)	AR(1)= 0.99	Sept 2012 =24.29	Sept 2012 =24.26	22.67	25.97	-5.66
			Oct 2012=24.24	Oct 2012=24.2	22.02	26.68	
			Nov 2012 =24.62	Nov 2012 =24.22	21.52	27.24	
BH4886	(1,0,0)	AR(1)=0.99	Aug 2012=6.4	Aug 2012=6.3	4.49	8.83	-2.43
			Sept 2012 =6.45	Sept 2012 =6.23	3.88	10.05	
			Oct 2012 =6.7	Oct 2012 =6.17	3.46	11.08	
Z4401	(1,0,0)	AR(1)=0.995	Sept 2012 =6.65	Sept 2012 =6.74	4.49	8.83	-2.43
			Oct 2012=6.8	Oct 2012=6.72	3.88	10.05	
			Nov 2012 =6.92	Nov 2012 =6.73	3.46	11.08	

BH4371	(1,0,0)	AR(1)=0.994	Oct 2012 =4.72 Nov 2012 =4.87 Dec 2012=4.81	Oct 2012 =4.62 Nov 2012 =4.58 Dec 2012=4.54	3.75 3.41 3.16	5.69 6.15 6.50	-3.42
BH4348	(1,1,0)	AR(1)= 0.997	Sept 2012 =6.05 Oct 2012 =6.17 Nov 2012 =6.33	Sept 2012 =6.02 Oct 2012 =5.99 Nov 2012 =5.95	4.70 4.23 3.89	7.69 8.46 9.09	-3.1
BH4995	(1,0,0)	AR(1)=1.00	Oct 2012 =14.68 Nov 2012 =15.06 Dec 2012=15.22	Oct 2012 =15.06 Nov 2012 =15.17 Dec 2012=15.28	12.10 10.97 10.12	18.01 19.37 20.44	1.9

5.2.2 Discussions

The selected best fitted ARIMA models for forecasting groundwater table at each borehole are summarized on Table 5.4. It can be observed in all the boreholes, the structure of the models shows only the non-seasonal component which implies that groundwater table series do not show any effect of seasonality. This means that all the boreholes groundwater level series are fitted by the simpler form of ARIMA model in the form of (p,d,q) as opposed to the multiplicative ARIMA model of the form $(p,d,q)\times(P,D,Q)_w$. This can be confirmed by observing the ACF and PACF's at lag 12. The autocorrelations at lag 12 lie within the 95% confidence limits, which signifies an insignificant autocorrelation. The second observation from the result is that all the boreholes are fitted by first order autoregressive model i.e. AR(1) except for BH4973. AR(1) in this case implies that the depth of groundwater for a particular month has a strong dependence on the previous month's observations. This is reflected by a significantly strong autocorrelation for most boreholes. However, BH4341 and BH4348 are slightly different from the rest of the AR(1) models. First order differencing had to be performed on the series before modelling. The structure of the models indicate that groundwater level at a particular time *t* depends on the previous 2 monthly observations. The time series at BH4973 was the only one fitted with an AR(2) model. This implies that groundwater level series have a significant autocorrelation up to lag 2 as opposed to lag 1 like in the rest of the boreholes. Lastly, the moving average component in almost all boreholes is insignificant. This can be observed from ACF and PACF under appendix B.

Percentage errors estimates of forecasts have been presented on table 5.3 for BH4341 while the rest are on appendix E. These are comparisons of observed data and model's three months forecasts at the individual boreholes. For ARIMA model, the results reveal that all forecasting error estimates are mostly less than 15% except for BH4164. This borehole recorded the largest percentage error of 15.6 %. The lowest percent error estimate was found to be 0.12% for BHZ6424. Therefore there was generally a small difference between forecasted and observed values in all the 13 boreholes under consideration.

5.3 THOMAS FIERING MODEL

Thomas Fiering model is of a Markovian nature consisting of periodic parameters such as mean, standard deviation and correlation between successive monthly observations (Subagadis, 2009). It consist of 12 regression equations, one for each month. For the purpose of modelling and forecasting of groundwater levels in the study area, a set of Thomas Fiering models were developed for 12 months at each borehole. The results are presented on table 5.6 for BH4341. A 3 months forecast was also conducted as shown on Table 5.5. The results for other boreholes obtained in the same manner have been presented in appendix D.

Table 5.5 Statistical parameters for BH4341 GWL time series

Month	Standard deviation	Mean	Z	Correlation
January	0.63	1.65	0.437	-0.2
February	0.37	1.45	-2.12	0.35
March	0.24	1.74	1.085	0.91
April	0.58	1.99	-0.277	0.79
May	0.27	1.86	-2.17	-0.09
June	0.18	1.80	0.018	0.165
July	0.15	1.86	-0.722	0.98
August	0.36	1.92	0.21	0.87
September	0.19	1.98	-0.556	0.73
October	0.12	2.08	0.465	0.96
November	0.14	2.11	-1.812	0.89
December	0.21	2.16	1.526	0.85

Table 5.6 Thomas Fiering Models for BH4341

Thomas-Fiering Model - BH 4341
$X_{jan} = 1.65 - 0.6(X_{dec} - 2.16) + 0.98Z_1$
$X_{Feb} = 1.45 + 0.21(X_{jan} - 1.65) + 0.34Z_{2}$
$X_{mar} = 1.74 + 0.59(X_{feb} - 1.45) + 0.234_{3}$
$X_{Apr} = 1.99 + 1.91(X_{mar} - 1.74) + 0.35Z_{4}$
$X_{may} = 1.86 - 0.04(X_{Apr} - 1.99) + 0.27Z_{5}$
$X_{jun} = 1.8 + 0.11(X_{may} - 1.86) + 0.19Z_{6}$
$X_{jul} = 1.86 + 0.82(X_{jun} - 1.8) + 0.03Z_{7}$
$X_{Aug} = 1.92 + 2.09(X_{jul} - 1.86) + 0.18Z_{8}$
$X_{sept} = 1.98 + 0.39(X_{Aug} - 1.92) + 0.13Z_{9}$
$X_{oct} = 2.08 - 0.61(X_{sept} - 1.98) + 0.03Z_{10}$
$X_{nov} = 2.11 + 1.04(X_{oct} - 2.08) + 0.06Z_{11}$
$X_{dec} = 2.16 + 1.28(X_{nov} - 2.11) + 0.11Z_{12}$

Table 5.7 Forecasting results for BH4341 using T-F model

Borehole	**Months**	**T-F Results**	**Observed Values**	**Forecast Error (%)**
BH4341	October 2012	1.98	2.21	10.4
	November 2012	1.96	2.4	18.3
	December 2012	1.9	1.99	4.5

5.3.1 Discussions

The structure of T-F model has been described in details under section 3.6. In its structure, the month to month correlation plays an important role in accuracy of fit to the observed data and hence forecasting. Poor values of month to month correlation and large standard deviations are likely to provide inadequate fit making forecasts to be highly dependent on Gaussian random number (Ahmed *et al.*, 2001). The performance of the model in terms of forecasting was done based on percentage error estimates by comparing forecasted values with observed groundwater table values. The forecasting results on Table 5.7 can be described as reasonable .The same can be said for most boreholes except for BH6423 where the forecasting errors are generally higher than the rest (Appendix E). These reasonable forecasts can be attributed to generally high month to month correlations and low standard deviations in most of the boreholes hence the forecasts are not just dependent on Gaussian random number.

5.4 Summary

Two stochastic models were investigated for modelling and forecasting of monthly groundwater levels in Ramotswa wellfield at 13 observation boreholes. For ARIMA modelling, it was found out that groundwater levels at all boreholes were fitted by simpler form of ARIMA model as opposed to multiplicative ARIMA form. There was no evidence of seasonality since all ACFs and PACFs for lag 12 and beyond were insignificant for most of the boreholes. All the models chosen for forecasting were AR(1) models except for BH4973 which was fitted by AR(2) model. For T-F model, the data generally has high month-to-month correlation and low standard deviation. This high correlation indicated that T-F model provided adequate fit. In terms of forecasting capability, both models managed to reasonably forecast groundwater level at each borehole. This is indicated by low forecast errors after verification with observed values for 3 months forecasting. For ARIMA model, the percentage errors range was 0.12 to 15.6 %. For T-F model, the percentage errors were higher and had a wider range (i.e. 2.37 to 53.38 %). Based on the comparisons between these stochastic modelling in terms of their forecasting capabilities presented in appendix E, it can be concluded that both models have shown their capability to produce reliable forecasts. However, ARIMA models which produced better forecasts are therefore recommended for forecasting monthly groundwater levels in the study area.

Since these two models can produce reasonable forecasts of groundwater levels in the wellfield, they are very useful tools to water resources managers and can contribute towards sustainable groundwater resource management in a number of ways;

i) The recent operation of the production boreholes after a long time could change the groundwater flow dynamics especially if there is uncontrolled abstraction and development of a huge cone of depression. This may result in reversals in flow directions and further contamination in an area that is located on a dolomite aquifer which is highly karstified and vulnerable to pollution. These problems can be detected by using reliable ARIMA models to forecast changes in groundwater levels which may result in changes in water quality.

ii) Drought Monitoring –Water resources managers can utilize forecasting results in development of drought warning systems when groundwater levels reach critical levels. This would prompt them to make informed decisions such as putting control measures to regulate pumping rates and where possible shutting down production in selected areas to enable recovery of depleted boreholes which would contribute towards extension of the aquifer's lifetime. Planning and control of future groundwater developments relies on accurate forecasting of groundwater levels hence the study would assist water regulating authority in that regard.

iii) Forecasted groundwater levels may be used together with water quality data for decision making including determining the rate and extent to which abstraction could be sustainably done on a borehole without compromising water quality since changes in groundwater levels affect groundwater quality.

5.5 GEOSTATISTICAL MODELLING

In sections 5.2 and 5.3 of this chapter, forecasting results were presented based on fitting two stochastic models to the groundwater level time series. This section presents results of using geostatistical techniques to model the spatial distribution of groundwater level in Ramotswa wellfield. These techniques are pivotal tools for sustainable management of groundwater resources in order to meet the continuously increasing demand for fresh water. Geostatistical techniques were utilized to determine the spatial distribution of groundwater table during the dry and wet season scenarios based on groundwater table observations at 21 boreholes. Groundwater levels for July 2005 were taken as representation of dry season scenario. This corresponds to a period in which there was less recharge from rainfall and therefore groundwater levels were generally low. February 2006 was chosen to represent the response of groundwater during a wet season. This was a time when Ramotswa village experienced extremely heavy rainfall and flooding. As a result, an attempt was made to obtain a groundwater level prediction surface under those two scenarios and hence interpolate groundwater level at unknown locations.

5.5.1 Summary of ordinary kriging

Before the data set for July 2006 and February 2007 were used for geostatistical analysis, normality checks were conducted on SPSS to check distribution of the data. This is a pre-requisite for applying geostatistical analysis on spatial data (Machiwal *et al*., 2012). After conducting logarithmic transformation and ensuring that data is approximately normally distributed, spatial interpolation process began.

In generating the spatial groundwater level maps, ordinary kriging method available on Integrated Land and Water Information System software (ILWIS) 3.3 was used. The point map was first created by importing groundwater table point data at each borehole into the GIS environment for each scenario. In order to understand the spatial autocorrelation and variance among observed groundwater level values at each borehole, experimental semi-variograms as well as their parameters were determined (i.e. nugget, range and sill).

Three geostatistical models namely spherical, exponential and Gaussian models were fitted to the experimental semi variograms. The best fitted model was then used for spatial interpolation. The semi variogram model which fitted data the best was tested by calculating goodness of fit (R^2). The model with the highest value of R^2 was chosen for spatial interpolation. The accuracy of interpolation was confirmed by cross validation for a sample of boreholes. This was done by removing one borehole at a time then interpolating with the remaining ones and thereafter checking the error of interpolation. The flow chat on Fig 5.4 illustrates summarizes the steps taken in spatial modelling.

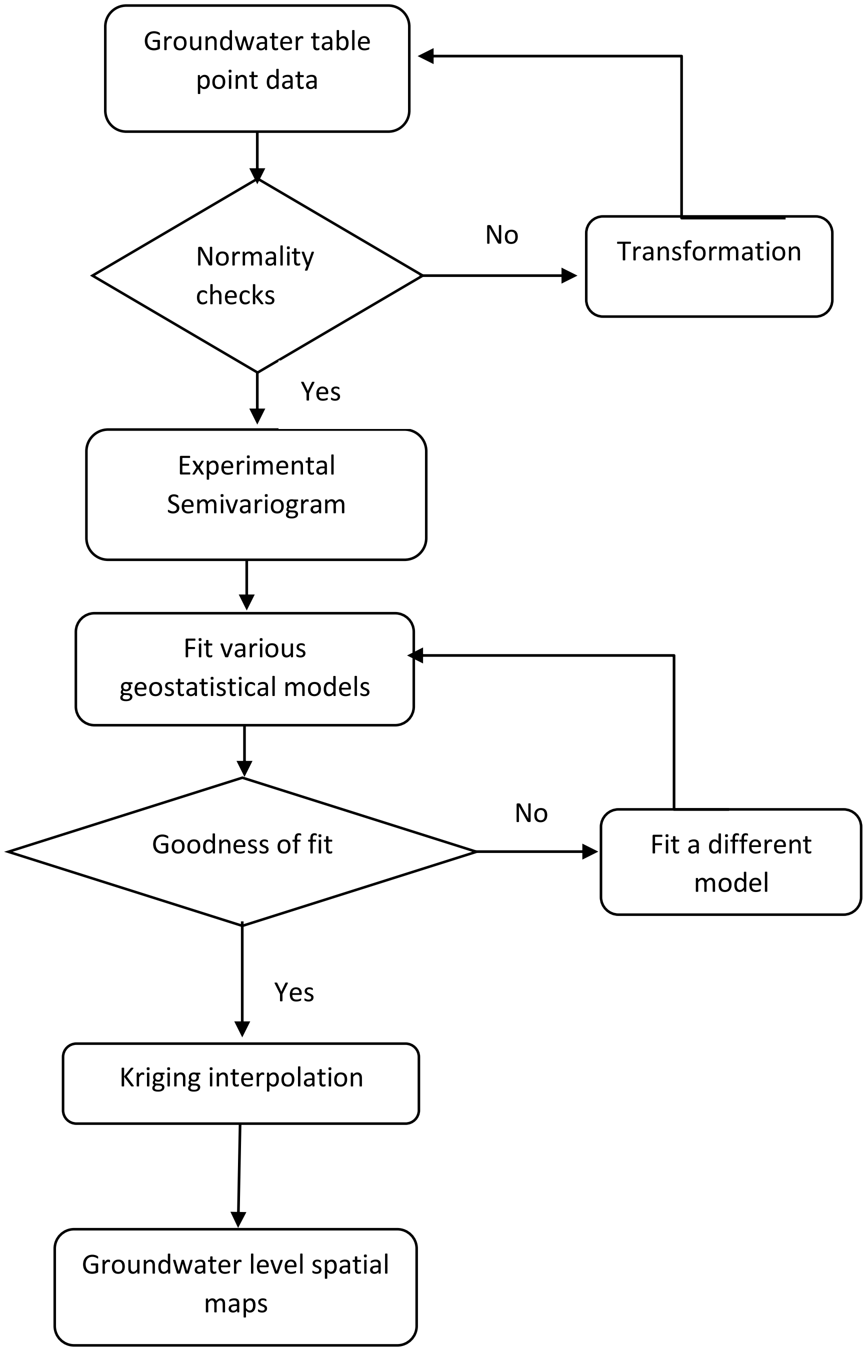
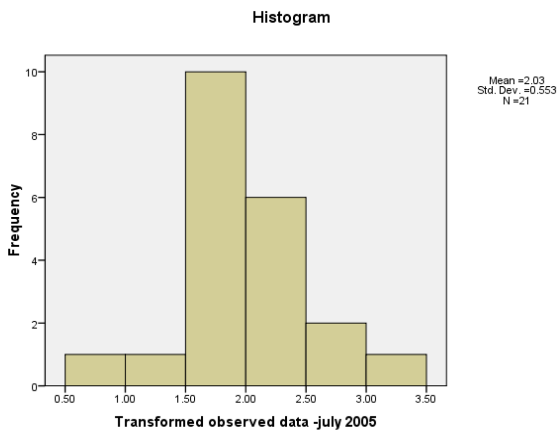


Fig 5.4 Summary of methodology for spatial modelling of groundwater levels

5.5.2 Normality Tests

In verifying that the data is indeed normally distributed, the histogram and Q-Q plots for the log transformed data for July 2005 and February 2006 are show on Figures 5.5 and 5.6. From visual inspection, it was concluded that data for both scenarios the transformed data was approximately normally distributed and therefore can be used for spatial interpolation.



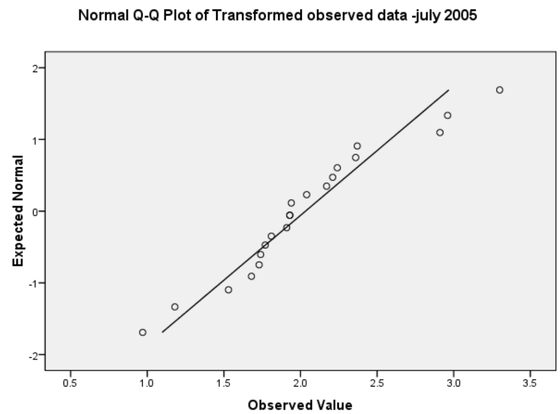
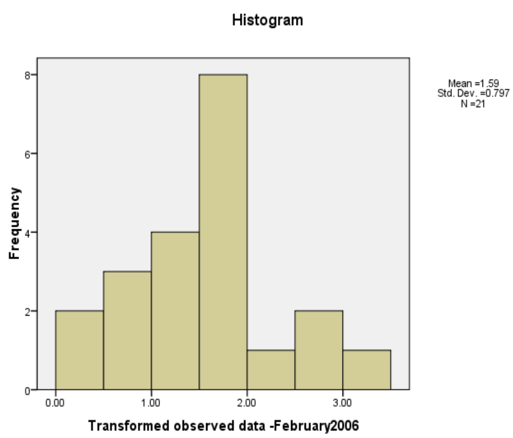


Fig 5.5a) Histogram –July 2005

Fig 5.5b) Q-Q- plots for July 2006



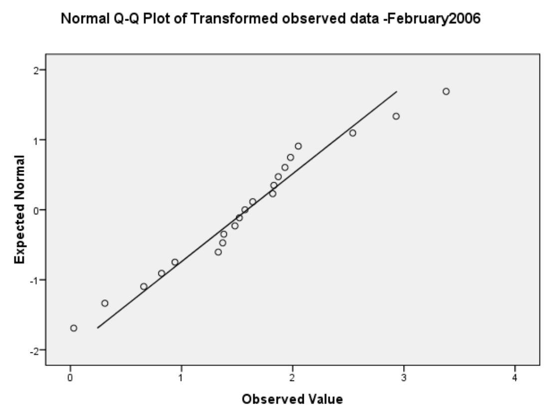


Fig 5.6 a) Histogram –February 2006

Fig 5.6b) Q-Q- plots for February 2006

5.5.3 Spatial interpolation-July 2005

The experimental semi-variogram for the dry season scenario (July 2005) has been shown on Fig 5.7. It shows the extent of spatial dependence of groundwater levels at the observation boreholes during the observation month. The variance of groundwater levels increases up to a distance between pairs of points of 4km then remains constant. This implies that there is spatial autocorrelations between groundwater levels at boreholes within a radius of 4Km (i.e. the radius of influence). Beyond this range there is no spatial correlation. Table 5.8 summarizes the parameters of 3 geostatistical models used in this study, i.e. spherical, exponential and Gaussian models. All the models show that the range is 4 km, which means that groundwater levels at unknown sites can be estimated by interpolating known groundwater levels within that radius of influence. The strength of spatial dependence of a variable can be determined by the nugget-to-sill ratio. A ratio that is less than 0.25 indicates strong spatial dependence, moderate if it is between 0.25 and 0.75 and weak spatial dependence if it beyond 0.75 (Machiwal *et al*., 2012). The nugget-to-sill ratio was found to be 0.1 which is less than 0.25, an indication that groundwater levels during this observation period (July 2005) has strong spatial dependence. From fitting of various models to the experimental semi-variogram, the best model was chosen on the basis of the model giving the highest R^2 value. The exponential model was chosen for spatial interpolation due to a higher R^2 value of 0.7 which implies that there is a 70% correlation between calculated semi-variogram values and experimental semi-variogram values. The goodness of fit calculations are shown on appendix F. Spherical model provided another alternative since it also gave the same goodness of fit as the exponential model for the dry season scenario.

Table 5.8 Parameters of three geostatistical models for Groundwater levels for July 2005

Month	**Model**	**Sill(m^2)**	**Nugget(m^2)**	**Range(m)**	**R^2**
July 2005	Spherical	0.05	0.007	4000	0.7
	Exponential	**0.06**	**0.006**	**4000**	**0.7**
	Gaussian	0.08	0.007	4000	0.6

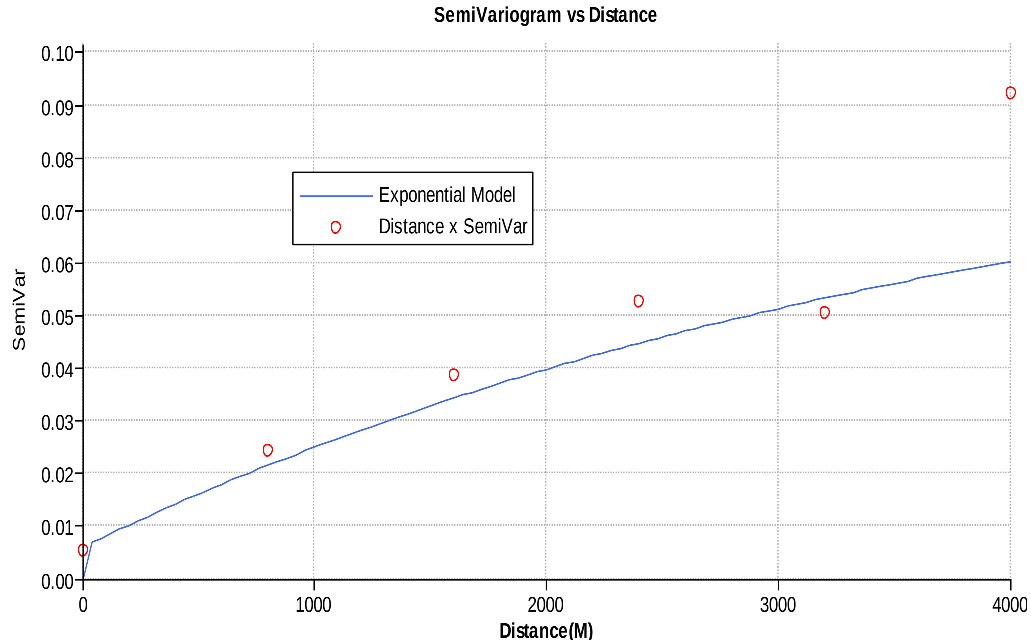


Fig 5.7 Experimental Semi-Variogram and the fitted Exponential model

Groundwater level interpolation maps were derived using observed data to estimate groundwater levels at unknown locations. In order to verify that the interpolated groundwater levels were reasonably estimated, Table 5.9 presents a comparison between the observed and estimated groundwater levels. The results for this scenario shows that even though interpolation using ordinary kriging slightly overestimated the depth of groundwater, the error of estimate ranged between 1and 18%. Therefore it could be reasonably stated that the spatial map represented the state of groundwater storage over the wellfield during the monitoring period of July 2005. It can be seen from the Fig 5.8 that in terms of spatial distribution, the observation boreholes are located along the Notwane River. These boreholes are shallow even in the dry season with the groundwater levels ranging between 3 and 5 m. For instance, BH 4341, BH 4163 and BH4371 recorded 2.63m, 4.63m and 5.38 m. However deeper boreholes are located further away from the river (e.g. BHZ6424 recorded 27.13 m). There is therefore evidence of groundwater level drop across the wellfield in the North-Eastern direction. This shows that groundwater flow is driven by the difference in elevation. The spatial interpolation maps also concurs with the findings by Staudt, 2003 in Ramotswa wellfield which revealed through the contours maps that groundwater flows in the North and North Eastern direction towards the river.

Table 5.9 Comparison between observed and interpolated groundwater levels (July 2005)

BOREHOLES	Observed GWL(m)	Estimated(m)	% Error
BH6423	10.67	12.2	14.3
BH4972	6.92	7.00	1.16
BH4163	4.63	4.91	6.05
BH4165	5.65	6.69	18.4
BH4371	5.38	4.93	8.36

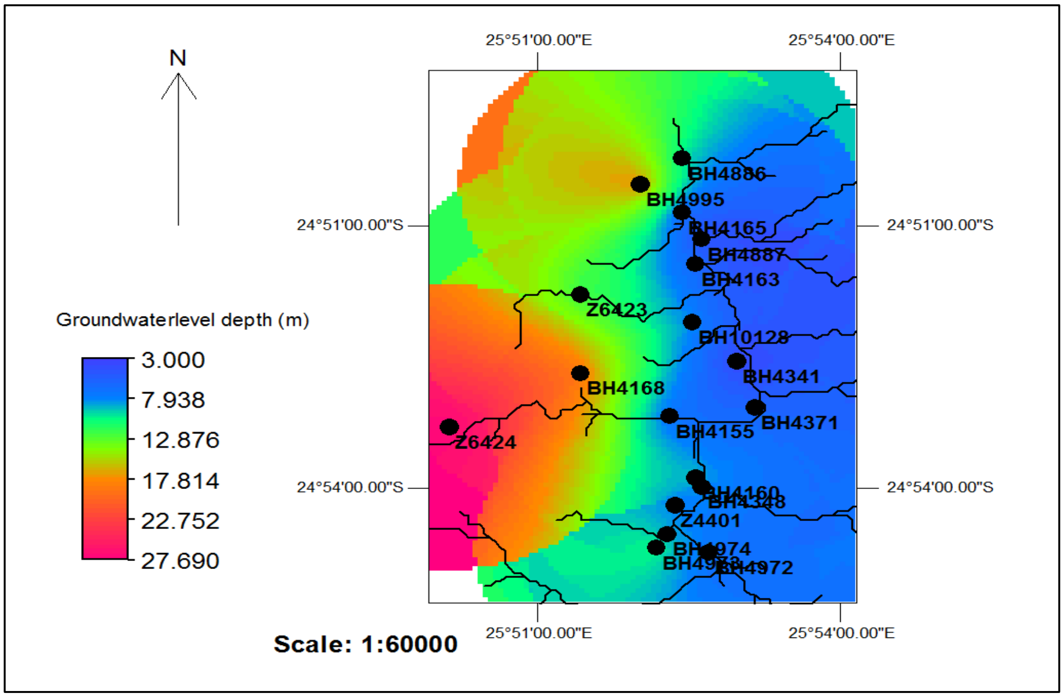


Fig 5.8 Spatial interpolation map for July 2005

5.5.4 Spatial interpolation-February 2006

Similar to July 2005 scenario, the results of groundwater level interpolation have been presented below. The interpolation map in Fig 5.10 shows the spatial distribution of groundwater table after flood events in February 2006. In terms of spatial correlation, the experimental semi-variogram shows that the radius of influence is 3 km. The geostatistical model parameters for spherical, Gaussian and exponential models have been presented on Table 5.10. In this case, the nugget-sill ratio of 0.29 is greater than 0.25 which indicates that the observed data has moderate spatial dependence. In terms of goodness-of-fit of geostactical models to the experimental semi-variogram, the Gaussian model gave the lowest R^2 value as low as 0.4, hence it was not used for interpolation. The exponential model gave an R^2 of 0.7 while spherical model gave 0.9 which was eventually chosen and used for groundwater level interpolation.

Table 5.10 Parameters of three geostatistical models for Groundwater levels for February 2006

Month	Model	Sill(m^2)	Nugget(m^2)	Range(m)	R^2
February 2006	**Spherical**	**0.59**	**0.17**	**3000**	**0.9**
	Exponential	0.55	0.15	3000	0.7
	Gaussian	0.6	0.17	3000	0.4

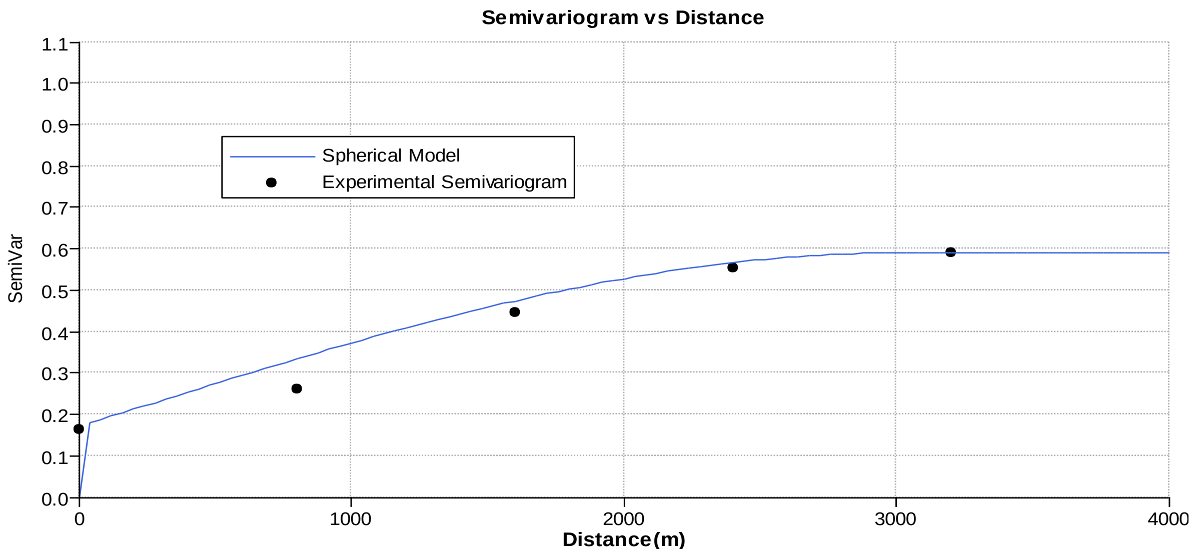


Fig 5.9 Experimental Semi-Variogram and the fitted model (Spherical model)

Table 5.11 shows the comparison between the interpolated groundwater level data and the observed data for month of February 2006.The results show that ordinary kriging in this case generally underestimated the depth of groundwater except for BH4155. The percentage error of estimation range was from 5- 31%. This is once again an indication that ordinary kriging method performed reasonably well in terms of estimating groundwater table at unknown locations. Similar to the July 2005 scenario, the existence of hydraulic gradient is evident in the North and North Eastern direction and this clear shows the influence of topography on groundwater flow.

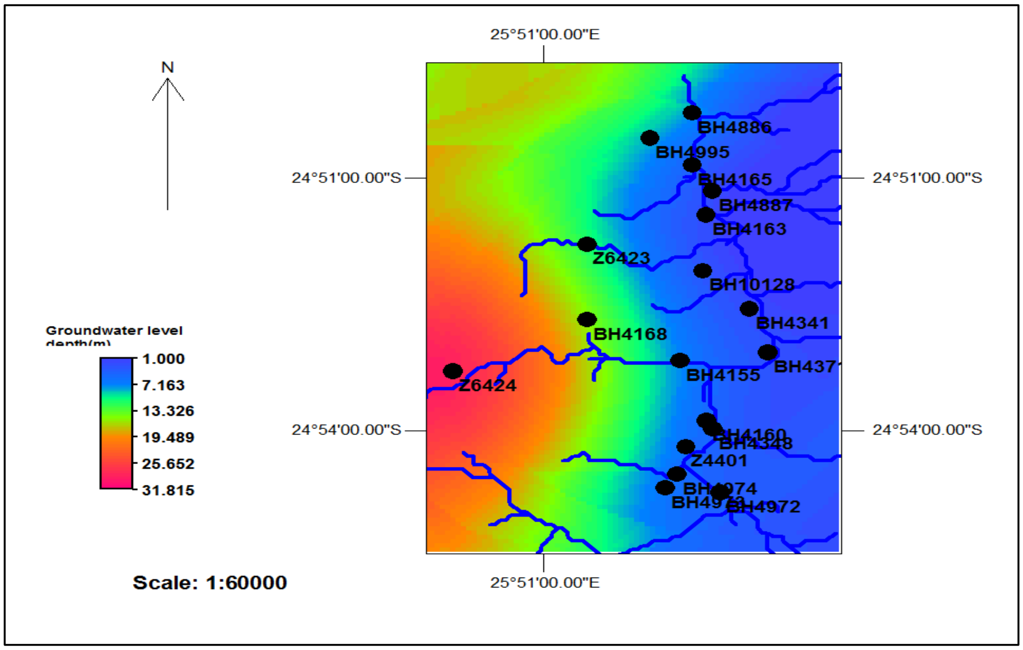


Fig 5.10 Spatial interpolation map for February 2006

Table 5.11 Comparison between observed groundwater levels and interpolated ones (February 2006)

Borehole	Observed GWL(m)	Interpolated(m)	% Error
Z4401	6.23	5.32	14.6
BH4371	3.95	3.56	9.87
BH4155	4.38	4.59	4.79
BH4164	4.55	3.79	16.7
BHZ6501	6.89	4.78	30.6

5.5.5 Response of groundwater levels due to flooding

In the previous sections, the groundwater level conditions in the Ramotswa wellfield has been presented under two scenarios namely;

- Dry season represented by season represented by observation data for July 2005
- Wet season represented by observation data for February 2006 after heavy floods

As already shown in section 4.3 by the groundwater level and rainfall hydrographs, flooding in 2006 caused a significant rise in groundwater table, in fact all the flooding years thereafter coincided with and immediate rise in water levels. For the scenarios mentioned above, the response can be visualized through the interpolation maps on Figs 5.8 and 5.10 that the area with high groundwater levels increased between July 2006 and February 2007. This was supported by numerical data in table 5.12 below in which the percentage change in groundwater level were computed for each borehole. Heavy rainfalls at the beginning of year 2006 led to 6 boreholes recharging by more than half of their previous levels in just 6 months. Percentage increase in groundwater level was significantly high in boreholes downstream which are closer to the river than those further away upstream. This is not surprising since groundwater flow follows the terrain which is sloping towards the river and majority of the boreholes along the river are at a discharge point. Boreholes further upstream which are located at high elevations received less or no recharge at all (e.g. BHZ6424). This is due to the fact that water is released very fast at high elevations due to gravity. These observations are supported by hydrogeological characteristics of the terrain. Typical dolomite characteristics support fast release of ground water from high elevations and cause less recharge in those boreholes upstream. The fracturing and intersection with minor valleys in the West-East direction as reported by Staudt, 2003 is responsible for high recharge rates in the boreholes which are in the proximity to the Notwane River (e.g. BH4341). This also explains the location of the production boreholes which are also close to the river in the south eastern part of the wellfield. This makes groundwater on the downstream side highly vulnerable to pollution from pit latrines as has been the case in the past.

Table 5.12 Change in groundwater levels between July 2005 and February 2006

Borehole	GWL-July 2005	GWL-February 2006	Percentage recharge (%)
BH4155	5.67	4.38	23
BH4160	9.41	6.18	34
BH4163	4.63	1.94	58
BH4164	6.13	4.55	26
BH4165	5.65	2.28	60
BH4168	19.25	18.73	3
BH4341	2.63	1.03	61
BH4348	6.86	2.55	63
BH4371	5.38	3.95	27
BH4885	5.89	5.15	13
BH4886	9.11	6.46	29
BH4887	3.27	1.36	58
BH4972	6.92	3.99	42
BH4973	10.62	7.24	32
BH4974	7.69	3.79	51
BH4995	18.37	12.62	31
BH10128	8.73	4.81	45
Z6424	27.13	29.39	-8
Z6501	6.72	6.87	-2
Z4401	6.98	6.24	11
Z6423	10.67	7.75	27

5.5.6 Summary

Geostatistical techniques were used to determine the spatial distribution of groundwater levels in Ramotswa wellfield for the dry (July 2005) and wet season (February 2006) using data from 21 boreholes. Explanatory analysis shows that data for both scenarios is approximately normally distributed after transformation. In order to interpolate groundwater levels at unknown locations through ordinary kriging, the spatial correlation structure had to be investigated through semi-variogram analysis. A strong spatial dependence for July 2005 with a nugget-to-sill ratio of 0.1 was revealed while February 2006 showed moderate spatial dependence (Nugget-to-sill ratio of 0.28). R^2 values of 0.7 and 0.9 for July 2005 and February 2006 respectively for chosen geostatistical models were reasonable enough to conclude that the models were correctly chosen for fitting to experimental semi-variograms. Interpolation results show that groundwater flow is in the north and north eastern direction, and therefore groundwater fluctuation is highly influenced by topography since land is sloping in that direction towards the Notwane river. The error estimates of interpolation shown in Tables 5.8 and 5.10 were in the range of 1 to 18 % for July 2005 and 5 to 31 % for February 2006 respectively. These minimal errors are an indication that ordinary kriging performed well in terms of interpolation of groundwater levels for the study area. This implies that ordinary kriging can be recommended to interpolate groundwater levels at unknown areas .Therefore this study can be used to assist in deciding on possible areas for the expansion of the monitoring network since a denser network would improve the accuracy of groundwater level spatial interpolation. Looking at the spatial interpolation maps on figs 5.9 and 5.11, it can be seen that most of the boreholes are located in the N-S direction closer to the river. Hence to improve groundwater monitoring, the borehole network should be expanded in the north and south western areas. In addition to that, critical areas where flooding may cause damage to property especially looking at areas where groundwater level is shallow and areas which have been critically affected by groundwater pollution in the village can be identified by making use of these maps.

CHAPTER 6-SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

6.1 SUMMARY

There is no doubt that adequate supply of water is of paramount importance for the livelihood of all living species on earth. Water resources availability will continue to be a serious challenge especially in semi-arid countries like Botswana and the rest of Southern Africa. Persistent droughts and the ever-increasing demand for water have resulted in depletion of surface water resources especially in South East region where Ramotswa (District administrative capital) and Gaborone (the nation's capital) are located. The drying of Gaborone dam which is a major source of water in the South East District escalated water demand. This led to serious consideration of alternative sources of water especially groundwater resources to supplement supply from North South Carrier (NSC) 1. Ramotswa wellfield which is one of Botswana's most productive wellfield was abundant in 1996 due to contamination of some boreholes from pit latrines. However, despite water being of questionable quality, the importance of groundwater in the study areas is highlighted by the refurbishments currently on-going at the wellfield and the blending of borehole water with surface water to mitigate the present drought. The construction of a P50 million reverse osmosis plant for treating this water is also on the pipe-line. Information on groundwater storage fluctuation with time due to external natural and man-made forces and how they affect the groundwater system needs to be understood, documented and presented to water authorities (DWA, WUC). Groundwater resource monitoring in both space and time by analyzing monthly groundwater level data provides a significant contribution towards managing Botswana's scares water resources. This information is critical for decision making with regards to the sustainable use of groundwater in the wellfield. The main aim of this dissertation is to contribute towards this by presenting the results of stochastic and geostatistical modelling of groundwater level fluctuations at observation boreholes.

The study first started by conducting intervention and trend analysis of groundwater level data collected between 2002 and 2012. Intervention analysis was done to ascertain homogeneity of data. Inconsistencies or shifts in the data maybe be due to several factors including changes in local climate or anthropogenic factors such as observational errors, changes in GWL recording methods and equipment used (Kampata *et al.*, 2008). Groundwater level data after the determined time of intervention was used for trend analysis and development of stochastic models. This data was considered to be recent and of relevance to the current climatic conditions of the study area. Trend analysis was conducted for each borehole data series to not only investigate groundwater level changes with time but to also verify the statistical significance of these changes. In the second part of the study, groundwater level fluctuations were modelled both in space and time using geostatistical and stochastic methods respectively. For time series modelling, two stochastic approaches were used; ARIMA and Thomas -Fiering models. These models account for the effects of seasonality differently (Ahmed *et al*., 2001). The developed models for each method were then used for 3 months forecasting for each observation borehole. The error estimate in groundwater level forecasting were quantified in order to recommend the model which produced better forecasts. The results of these forecasts are displayed in appendix E. For geostatistical modelling, groundwater level maps were produced to visualize the response of groundwater levels under 2 different scenarios; July 2005 (dry season scenario) and February 2006 (wet season after flooding). The ability of ordinary kriging to estimate groundwater levels at unknown locations using known locations was verified by quantifying the error of interpolation.

6.2 CONCLUSIONS

The following conclusions can be drawn from this study;

i) Since there was no pumping during the monitoring period from 2002 to 2012 because of the shutting down of the wellfield in 1996, changes in groundwater levels were a direct consequence of recharge from rainfall. From Figures 4.2a), 4.2b) and 4.2c) there was a steady decline in groundwater levels before year 2006 while afterwards the rate of decline of groundwater levels reduced. This could reasonably be linked to flooding events that occurred between 2006 and 2012. The quick response of groundwater levels to rainfall is due to hydro-geological characteristics of the study area. Most of the boreholes are located in the dolomite aquifer (Geotechnical Consulting Engineers, 2000) which is highly fractured and permeable by its nature. It is those fractures which formed water collecting channels which led to a general recovery of the wellfield due to significant recharge from flooding after year 2006.

ii) Results of intervention analysis reveal that in all the boreholes, groundwater level data collected during the 10 year monitoring period had undergone a significant change. This implies that the data set does not belong to the same population hence the entire data could not be used for stochastic modelling. The time of intervention was detected between end of 2007 and beginning of year 2008. This change in groundwater level data is likely to be due to changes in hydro climatic conditions especially the significant recharges due to rainfall events that caused flooding in years 2006, 2007, 2009 and 2010 as highlighted in section 4.3.1. In terms of trend analysis, trend was statistically insignificant in most of the monitoring boreholes (62%), while few boreholes were showing significant trend (38%).

(iii) Both stochastic models were able to successfully model the fluctuation of groundwater levels using collected monthly data. This conclusion was made based on their ability to produce 3 months forecast with reasonable error of estimates. The forecasting errors were ranging from 0.12 - 15.6 % for ARIMA models while the range for Thomas Fiering model was ranging from 2.37 - 53%. Therefore ARIMA model was recommended for forecasting of groundwater levels in the wellfield.

iv) Groundwater levels measured at observation boreholes for July 2005 and February 2006 scenarios have strong and moderate spatial dependence respectively as indicated by the nugget-to-sill ratios. The best geostatistical models were exponential model and spherical model based on their goodness of fit to experimental semi-variograms. Spatial interpolation by ordinary kriging gave error estimates ranging from 1%- 18 % for July 2005 and 5 - 31% for February 2006 therefore can be recommended for estimation of groundwater levels at unknown areas in the wellfield.

v) Since groundwater levels used to generate interpolation maps were not affected by the effects of pumping, therefore there is evidence that groundwater levels are highly influenced by topography and the presence of natural water bodies. The boreholes in the proximity of Notwane River are located at a lower elevation (i.e. discharge point downstream) and hence groundwater level is shallower. Groundwater levels at upstream boreholes are deeper. It is the existence of this hydraulic gradient created by elevation difference that drives groundwater in the West-East direction towards the river and the groundwater level maps clearly indicate that.

6.3 RECOMMENDATIONS

The following are recommendations that can be considered for future research studies;

i) In Ramotswa wellfield, groundwater abstraction through pumping has not been done since the closure of the wellfield in 1996 until recently when some of the production boreholes became operational for mitigating the current drought. This study was focused on groundwater level fluctuation during a 10 year monitoring period when there was no abstraction. However, there is no doubt that the effect of pumping would have a significant impact on groundwater storage now that the wellfield is operating. A study of this nature is equally of relevant importance in terms of controlling groundwater exploitation which would contribute towards water resources management.

ii) This study was only focused on modelling and forecasting of groundwater levels. There was no consideration of groundwater quality due to insufficient sampling done for chemical analysis. However, the water quality aspect especially pollution of groundwater due to nitrates from pit latrines has been a subject of many discussions. As already mentioned in chapter 1, the closure of the wellfield was due to high concentrations of nitrates in borehole water. Therefore time series and spatial modelling of groundwater quality parameters is of relevant importance in the study area.

iii) Stochastic and geostatistical models used in the study are mainly data driven models and were preferred due to availability of spatial and time series data. Their advantage as compared to physical models is the fact that it is cheaper and less time consuming and they do not require detail hydrogeological analysis of the aquifer. However, groundwater physical modelling of the wellfield can be considered for future studies to complement time series models. These models can to a large extent account for the hydrogeological characteristics as well as their heterogeneity in space to simulate groundwater flow and solute transport through the unsaturated zone. Fluctuation of groundwater levels can therefore be described in details through physical processes in the unsaturated zone. This would be a very significant research since Ramotswa dolomite is highly vulnerable to pollution especially from pit latrines in the village.

REFERENCES

Adhikary S.K, Rahman M.M, Gupta A.D (2012), A Stochastic Modelling Technique for predicting groundwater table Fluctuation with Time Series Analysis. International Journal of Applied Sciences and Engineering Research Vol. 1

Ahmadi SH, Sedghamiz A (2007), Geostatistical Analysis of Spatial and Temporal Variations of Groundwater level. Environmental Monitoring and Assessment 129:277-294

Ahmed S, Khan Q.H, Parida B.P (2001), Performance of Stochastic approaches for Forecasting River Water Quality, Water Research Vol 35, pp 4261-4266

Arnell L.W (1999), Climate Change and Global Water Resources, Global Environmental Change, Vol 9, pp31- 49

Asadollahfardi G,Rahbar M,Atemiaghda M (2012), Application of Time Series models to predict water quality upstream and downstream of Latian Dam ,Universal Journal of Environmental Resources Technology 2(1) 26-35

Bierkens F.P, (1998), Modelling Water table fluctuations by means of a stochastic differential equation, Water Resources Research Vol 34, Pages 2485-2499

Box G.E.P and Jenkins (1976), Time Series Analysis: Forecasting and Control, San Francisco

Burn, D.H and Hag Elnur, M.A. (2001), Climate change impact using hydrologic variables in CCAF Workshop Meeting.

Clarke R.T, (1984), Mathematical Models in Hydrology. FAO of United Nations. Rome

De Gooijer G.J, Hynman R.J (2006), 25 years of time series forecasting. International Journal of Forecasting Vol 22,443-473

Department of Water Affairs (2006), Botswana National Water Master Plan Review, Final Report, Ministry of Minerals, Energy and Water Affairs

Department of Water Affairs, South Africa (2010), Groundwater Strategy 2010

Du Plessis A and Rowntree A (2003), Water Resources in Botswana with particular Reference to Savanna Regions. South African Geographical Journal

Durdu O.F (2008), Stochastic Approaches for Time Series Forecasting of boron, A case Study of Western Turkey. Environmental Monitoring Assessment 169:687-701

Fathian F, Morid S, Kahya E (2014), Identification of trends in hydrological and climatic variables in Urmia Lake basin, Iran Theoretical and Applied Climatology

Geotechnical Consulting Engineers, (2000), Groundwater Monitoring Project. Final Report, Vol 19, Ramotswa Wellfield

Haan C.T (2002) ,Statistical Methods in Hydrology.2nd Edition Iowa State University Press, Iowa, USA

Hipel K.W, Mcleod A.I, Lettenmaier D.P (1978), Assessment of Environmental Impacts Part one: Intervention Analysis. Environmental Management Vol 2, pp529-535

Hipel K.W & McLeod A.I (1994), Time Series Modelling of Water Resources and Environmental Systems, Developments in Water Science, Elsevier Science, Amsterdam, Netherlands

Integrated Land and Water Information System software Manual (ILWIS) (2001)

Kampata J.M, Parida B.P, Moalafhi D.B (2008), Trend Analysis of Rainfall in the Headstreams of the Zambezi River Basin in Zambia. Physics and Chemistry of the Earth Vol 33, pp 621-625

Kenabatho P.K, Parida B.P, Moalafhi D.B (2012), The value of large scale climate variables in climate change assessment: The case of Botswana's rainfall. Physics and Chemistry of Earth Sciences, pp 64-71

Kgathi D.L (1999), Water Demand Population and Sustainability in Botswana: Implications for development policy

Kholoma E (2011), A survey of water losses, the case study of Ramotswa village in Botswana

Knotters M and Bierkens F.P, (2000), Physical basis of time series models for water depths, Water Resources Research Vol 36, Pages 181-188

Kurunc A, Yurekli K, and Cevick O (2004), Performance of two Stochastic approaches for forecasting water quality and streamflow data from Yesilimak river, Turkey. Environmental Modelling and Software. Vol 20 pp 1195-2000

Machiwal D, Jha M.K. (2012), Hydrologic Time Series Analysis: Theory and Practice Springer/Capital Publishing Company, Germany/New Delhi, India (2012)

Machiwal D, Jha M.K, Mishra A, Sharma A, Sisodia S (2012), Modelling short term Spatial and Temporal Variability of Groundwater level using Geostatistics and GIS Natural Resource Research Vol 21

Majelantle A (2009), Botswana Water Statistic, Central Statistics Office, Gaborone

Makridakis S and Hibon M (1995). ARMA models and Box–Jenkins Methodology

Makridakis S, Andersen A, Carbone R, Fildes R, Hibon M, Lewandoski R, Newton J, Winkler R,(1982).The Accuracy of Extrapolation(Time Series) Methods: Results of a Forecasting Competition, Journal of Forecasting,Vol 1,pp111-153

Milanovic P (2004), Water Resources Engineering in Karst, Taylor and Francis Group

Mini P.K, Singh D.K, Sarangi A (2014) Spatial and Temporal Analysis of Groundwater level in Coastal Aquifer using Geostatistics. International Journal of Environmental Research and Development, Vol 4, pp 329-336

Minnesota Groundwater level Monitoring Network, (2011). Guidance Document for Network Development, Department of Natural Resources, Minnesota

Mirzavand M, Sadatinejad S J, Ghasemieh H,Imani R, Motagh S M.(2014) Prediction of Groundwater level in Arid Environment using a Non deterministic model. Journal of Water Resources and Protection, 6,669-676

Moalafhi D.B, Parida B.P, Kenabatho P.K (2014) A hybrid Stochastic-ANN approach for flow partitioning in the Okavango Delta of Botswana. Global Nest Journal, Vol 16, pp68-80

Moalafhi D.B, Tsheko R, Atlhopheng J.R, Odirile P.T, Masike S (2012). Implications of Climate Change on Water Resources of Botswana. Advanced Journal of Physical Sciences Vol 1(2) pp4-13

Modarres R. Eslamian S.S. (2006) Iranian Journal of Science & Technology, Transaction B, Engineering, Vol. 30, No. B4

Mohan S.Vedula S (1994) Multiplicative Seasonal Arima Model for long-term forecasting of Inflows, Water Resources Management 9,115-126

Nayak P.C, Rao Y.R & Sudheer K.P ,2006.Groundwater forecasting in a Shallow Aquifer using Artificial Neural Network Approach, Water Resources Management, pp 77-90

Parida B.P , Moalafhi D.B (2008). Regional Frequency Analysis for Botswana using L-moments and radial basis function and network. Physics and Chemistry of the Earth Vol 33, pp 614 – 620

Ranganai T, Atekwana E, King J, Tladi B (2002), Case Histories of Environmental and Engineering Geophysics in Botswana, Southern Africa. Symposium on the Application of Geophysics to Engineering and Environmental Problems, 10-14 February 2002, Las Vegas

Reagan K.M (1984). An Evaluation of ARIMA (Box-Jenkins) Models for Forecasting Wastewater Treatment Process Variables

Sahoo S and Jha M.K (2012) Analysis of Spatial Variation of Groundwater Depths Using Geostatistical Modelling.Vol 9, International Journal of Applied Engineering Research pp 317-322

Salas J.D, Delleur J.W, Yevjevich V & Lane W.L. (1980), Applied Modelling of Hydrologic Time Series. Water Resources Publication, Littleton, Colorado.

Shanmugasundram S (2012). Statistical Analysis to Detect Climate Change and its Implications on Water Resources

Staudt M (2003). Environmental Hydrogeology of Ramotswa, South East District, Republic of Botswana

Subagadis Y.S (2009) Stochastic Simulation of Streamflow and Hydrologic Drought Analysis

Taany R A, Tahboub A B Saffarini G A (2009), Geostatistical analysis of spatiotemporal variability of groundwater level fluctuations in Amman–Zarqa basin, Jordan: a case study. Environmetal. Geololgy, 57, pp. 525–535.

Tabari H, Nikbakht J, Some'e B.S (2011) Investigation of Groundwater level Fluctuations in the North of Iran. Environmental Earth Science

Taylor C.J & Alley W.M, (2001), Groundwater level monitoring and long term importance of water level data, US geological Survey Circular 1217.

UNECA, (2011), Working Paper 4, Climate Change and Water Resources in Africa, Analysis of knowledge, gaps and needs.

United Nations Development Programme (2006), Human Development Report: Beyond Scarcity, Power, poverty and the global water Crisis

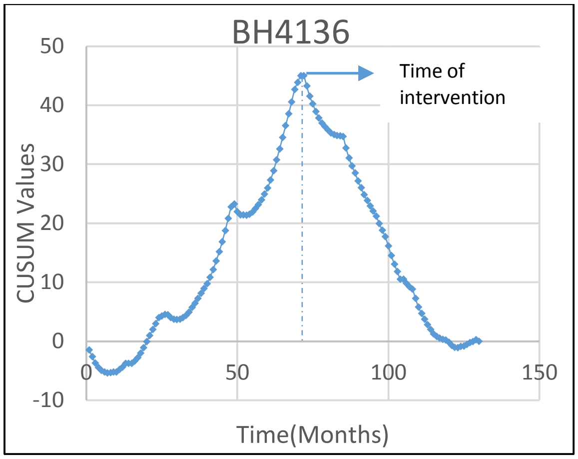
Villholth K.G, Giordano M, (2007). Groundwater use in a global perspective-Can it be Managed, The Agriculture Revolution, Opportunities and Threats to developments

Milanovic P (2004), Water Resources Engineering in Karst, Taylor and Francis Group

Wellfield Consulting Services (2011), Component 2 Regional Groundwater Drought Management Support, Contract 003D: Regional Groundwater Monitoring Network, Groundwater Monitoring Pilot Study: Final Report

Wetherald R.T, Manabe S, (2002), Simulation of Hydrologic changes associated with global warming, Journal of Geophysical Research 107 (D19), 4379

APPENDIX A – CUSUM Plots



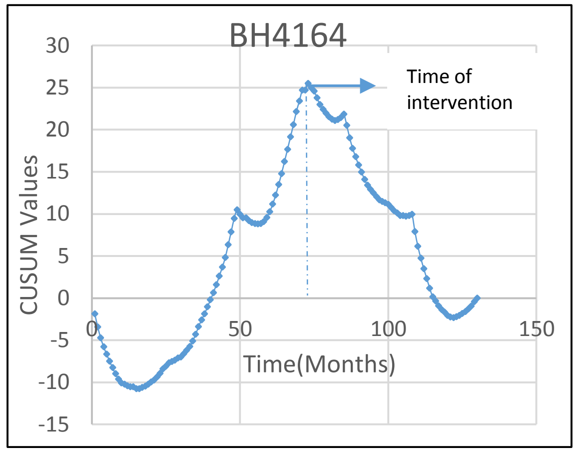
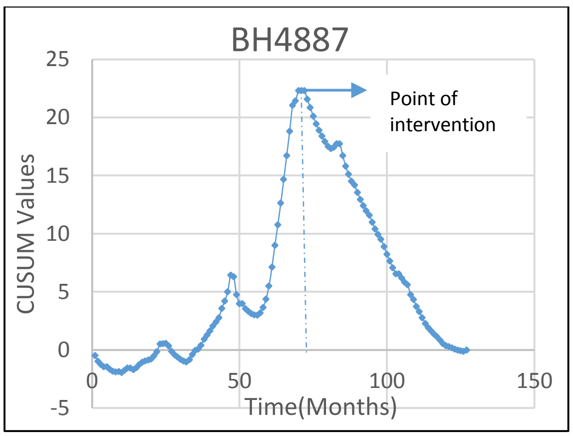


Fig A1: CUSUM Plot for BH4163 (2002-2012)   Fig A2: CUSUM Plot for BH4164 (2002-2012)



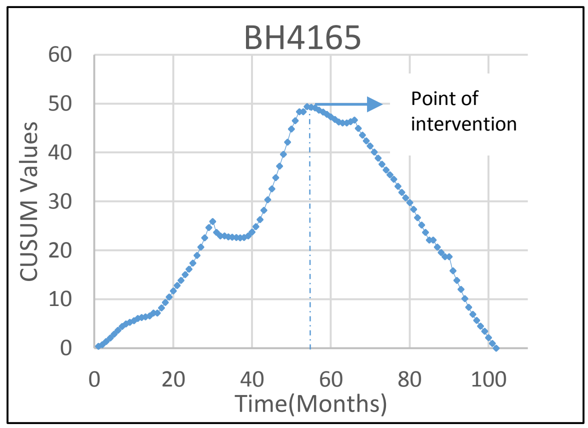
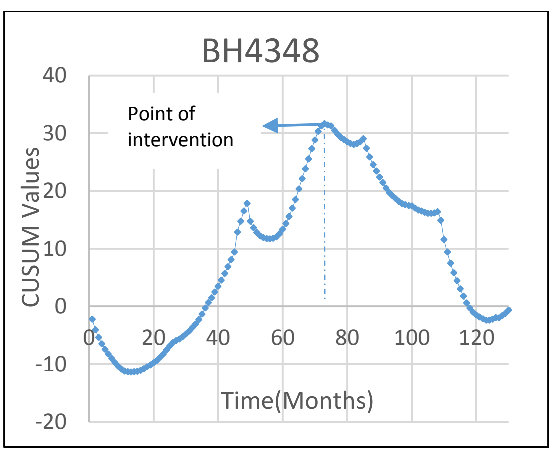


Fig A3: CUSUM Plot for BH4887 (2002-2012)   Fig A4: CUSUM Plot for BH4165 (2002-2012)



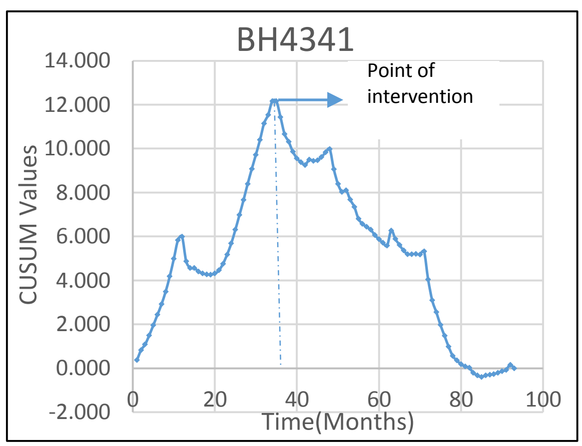
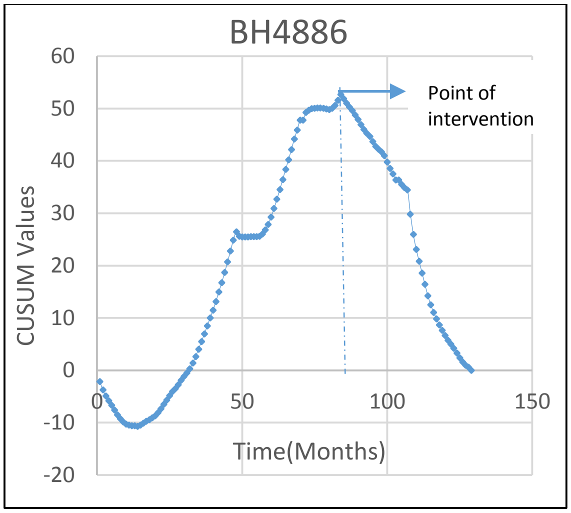


Fig A5: CUSUM Plot for BH4348 (2002-2012)   Fig A6: CUSUM Plot for BH4341 (2002-2012)



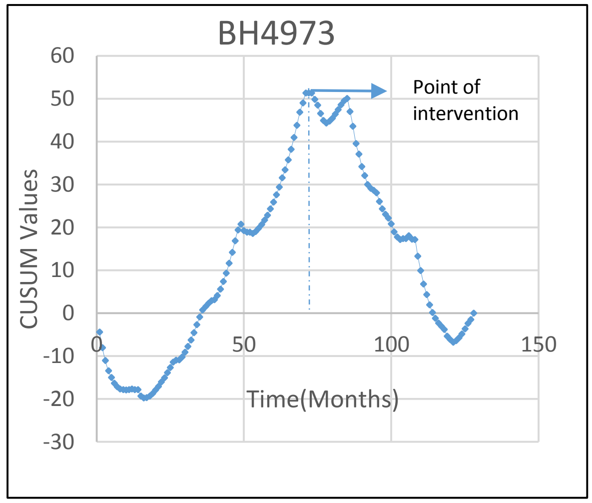
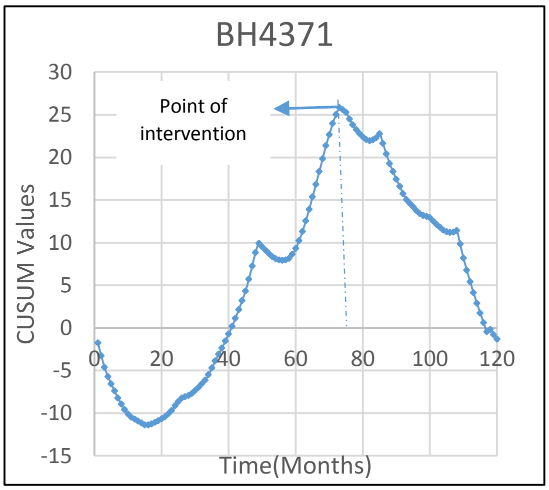


Fig A7: CUSUM Plot for BH4886 (2002-2012) Fig A8: CUSUM Plot for BH4973 (2002-2012)



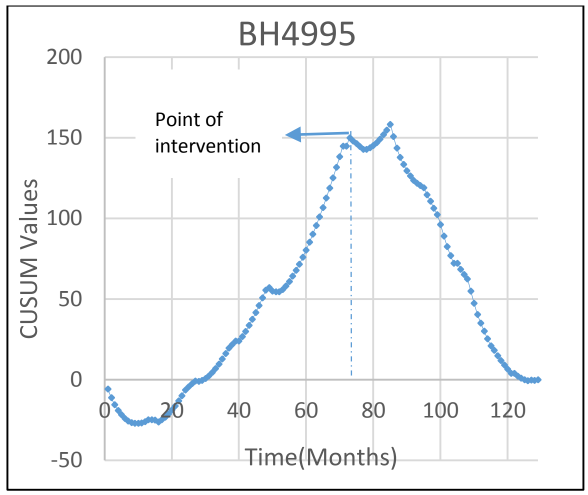
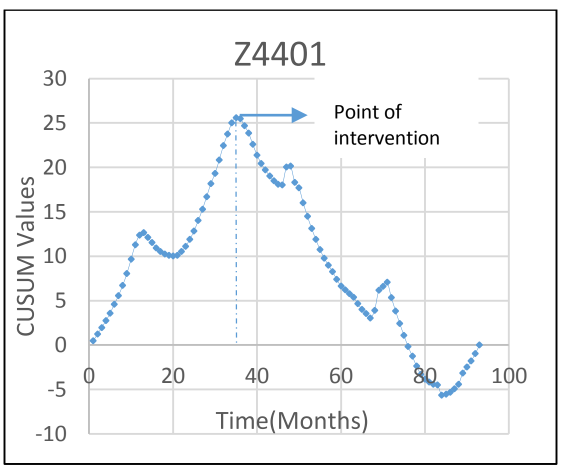


Fig A9: CUSUM Plot for BH4371 (2002-2012) Fig A10: CUSUM Plot for BH4995 (2002-2012)



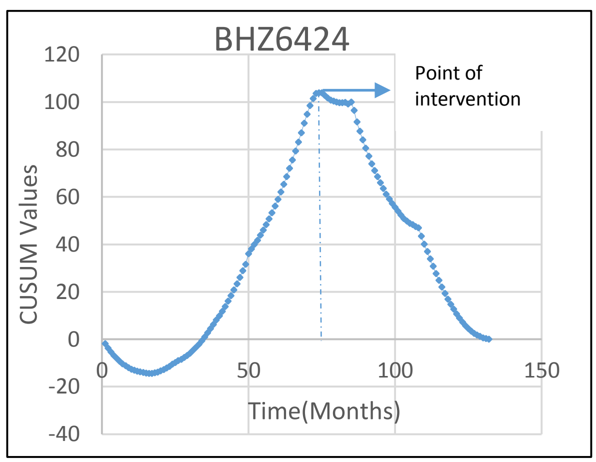


Fig A11: CUSUM Plot for Z4401 (2002-2012) Fig A12: CUSUM Plot for BHZ6424 (2002-2012)

APPENDIX B– Student t-test

BH 4341

t1(Months)	sample1		t2(months)	sample2		
1	2.53		1	1.43	40	1.60
2	2.61		2	1.38	41	1.59
3	2.42		3	1.81	42	1.58
4	2.57		4	1.71	43	1.66
5	2.62		5	1.85	44	1.66
6	2.63		6	1.98	45	1.74
7	2.64		7	2.03	46	1.95
8	2.73		8	2.40	47	1.99
9	2.85		9	2.10	48	2.05
10	2.96		10	2.19	49	2.10
11	3.00		11	2.30	50	1.93
12	2.32		12	2.37	51	2.03
13	1.03		13	2.31	52	2.09
14	1.85		14	1.24	53	2.23
15	2.16		15	1.49	54	2.18
16	2.00		16	1.79	55	2.19
17	2.07		17	2.23	56	2.22
18	2.11		18	1.73	57	2.23
19	2.15		19	1.83	58	2.21
20	2.21		20	1.62	59	2.40
21	2.30		21	1.92	60	1.99
22	2.46		22	2.02	**Av**	1.94
23	2.58		23	2.03	**S**	0.33
24	2.66		24	1.91	**S^2**	0.11
25	2.78		25	1.97	**S^2/n**	0.0019
26	2.83		26	1.99		
27	2.84		27	2.03	**Upper bound**	1.96
28	2.89		28	2.85	**Lower Bound**	-1.96
29	2.83		29	1.77		
30	2.81		30	1.89	**t**	7.09
31	2.83		31	1.90		
32	2.91		32	1.99		
33	2.55		33	2.16		
34	2.78		34	2.17		
35	2.16		35	2.13		
Av	2.50		36	2.30		
S	0.40		37	0.88		
S^2	0.16		38	1.20		
S^2/n	0.0045		39	1.62		

BH 4348							
t1(Months)	**sample1**	**t1(Months)**	**sample1**	**t2(months)**	**sample2**	**t2(months)**	**sample2**
1	3.45	38	6.51	1	6.03	38	2.37
2	3.81	39	6.67	2	5.47	39	3.45
3	4.34	40	6.65	3	5.45	40	3.73
4	4.52	41	6.71	4	4.93	41	4
5	4.73	42	6.79	5	4.97	42	4.28
6	4.82	43	6.86	6	5.12	43	4.26
7	4.88	44	6.87	7	5.25	44	4.38
8	4.95	45	6.99	8	5.26	45	4.49
9	5.04	46	9.1	9	5.36	46	4.75
10	5.17	47	7.5	10	5.53	47	4.97
11	5.32	48	7.47	11	5.8	48	5.14
12	5.53	49	6.95	12	5.95	49	5.32
13	5.62	50	2.55	13	6.17	50	5.32
14	5.68	51	4.57	14	4	51	5.46
15	5.74	52	4.815	15	4.16	52	5.62
16	5.83	53	5.06	16	4.35	53	5.84
17	5.92	54	5.35	17	4.53	54	5.95
18	5.99	55	5.5	18	4.61	55	5.59
19	5.99	56	5.62	19	4.7	56	6.05
20	6.05	57	5.73	20	4.72	57	6.05
21	6.04	58	5.88	21	4.87	58	6.17
22	6.2	59	6.23	22	5.1	59	6.33
23	6.23	60	6.42	23	5.17	Av	5.12
24	6.37	61	6.67	24	5.16	S	0.75
25	6.29	62	6.89	25	5.28	S^2	0.56
26	6.26	63	7.06	26	5.46	S^2/n	0.009
27	5.99	64	7.19	27	5.51		
28	5.96	65	7.45	28	5.57	Upper bound	1.96
29	6.04	66	7.46	29	5.31	Lower Bound	-1.96
30	6.16	67	7.35	30	5.3		
31	6.16	68	7.36	31	5.51	t	6.08
32	6.22	69	7.45	32	5.39		
33	6.29	70	7.14	33	5.51		
34	6.46	71	7.16	34	5.63		
35	6.6	72	6.6	35	5.74		
36	6.67			36	5.88		
37	6.59			37	4.13		
		Av	6.09				
		S	1.058				
		S^2	1.119				
		S^2/n	0.016				

BH4165							
t1(Months)	**sample1**	**t1(Months)**	**sample1**	**t2(months)**	**sample2**	**t2(months)**	**sample2**
1	4.89	36	4.44	1	7.15	36	2.96
2	4.86	37	4.46	2	6.23	37	4.5
3	5.12	38	4.58	3	6.31	38	3.06
4	5.25	39	4.83	4	4.5	39	3.39
5	5.27	40	5.29	5	5.55	40	3.62
6	5.29	41	5.57	6	4.42	41	4.5
7	5.28	42	5.98	7	4.37	42	1.63
8	5.07	43	6.39	8	4.05	43	2.56
9	4.78	44	6.64	9	4.07	44	2.64
10	4.83	45	6.73	10	4.06	45	2.66
11	5.01	46	6.79	11	3.97	46	2.74
12	4.68	47	6.89	12	3.96	47	3.05
13	4.62	48	6.93	13	4.02	48	3.27
14	4.71	49	7	14	4.27	49	3.29
15	5.1	50	7.15	15	4.55	50	3.45
16	4.5	**Av**	5.4	16	4.73	51	3.25
17	5.54	**S**	0.938	17	4.82	52	3.32
18	5.64	**S^2**	0.88	18	2.81	53	3.5
19	5.57	**S^2/n**	0.018	19	3.12	**Av**	3.71
20	5.75			20	3.37	**S**	0.99
21	5.65			21	3.45	**S^2**	0.99
22	5.52			22	3.26	**S^2/n**	0.02
23	5.68			23	3.26		
24	5.65			24	3.25	**Upper bound**	1.96
25	5.72			25	3.35	**Lower Bound**	-1.96
26	6.1			26	3.51		
27	6.22			27	3.57	**t**	8.89
28	6.37			28	3.1		
29	6.58			29	3.25		
30	5.74			30	3.39		
31	2.28			31	3.5		
32	3.83			32	3.15		
33	4.5			33	2.83		
34	4.28			34	2.94		
35	4.43			35	3.03		

BH4886					
t1(Months)	**sample1**				
1	5.13	36	8.72	71	7.29
2	5.72	37	8.74	72	8.75
3	6.1	38	8.83	73	7.74
4	6.36	39	8.8	74	7.59
5	6.44	40	8.79	75	7.33
6	6.36	41	8.95	76	7.35
7	6.38	42	9.11	77	7.32
8	6.53	43	9.07	78	7.18
9	6.7	44	9.21	79	7.17
10	6.87	45	9.31	80	7.2
11	7.07	46	9.37	81	7.5
12	7.2	47	9.36	82	7.83
13	7.29	48	8.85	83	8.29
14	7.16	49	6.46	84	8.37
15	7.5	50	7.15	**Av**	7.91
16	7.67	51	7.29	**S**	0.95
17	7.66	52	7.28	**S^2**	0.91
18	7.63	53	7.37	**S^2/n**	0.01
19	7.62	54	7.27		
20	7.68	55	7.29		
21	7.9	56	7.38		
22	7.97	57	7.68		
23	8.18	58	8.1		
24	8.09	59	8.35		
25	8.13	60	8.68		
26	8.05	61	8.9		
27	7.84	62	9.06		
28	8.07	63	9.12		
29	8.13	64	9.24		
30	8	65	9.19		
31	7.99	66	9.17		
32	8.07	67	9.2		
33	8.37	68	9.33		
34	8.52	69	9		
35	8.72	70	9.16		

t2(months)	**sample2**		
1	6.52	36	6.28
2	6.42	37	6.41
3	6.55	38	6.55
4	6.61	39	6.45
5	6.45	40	6.38
6	6.41	41	6.4
7	6.34	42	6.45
8	6.35	43	6.7
9	6.55	44	6.92
10	6.68	**Av**	6.1
11	6.31	**S**	0.88
12	6.4	**S^2**	0.78
13	6.7	**S^2/n**	0.02
14	6.81		
15	6.55	**Upper bound**	1.96
16	6.09	**Lower Bound**	-1.96
17	6.07		
18	6.23	**t**	10.72
19	6.09		
20	7.29		
21	6.45		
22	6.65		
23	6.88		
24	2.67		
25	3.4		
26	4.42		
27	5.04		
28	5.05		
29	5.09		
30	5.13		
31	5.55		
32	5.85		
33	6.02		
34	6.16		
35	6.25		

BH4164								
t1(Months)	**sample1**				**t2(months)**	**sample2**		
1	3.24	36	6.04		1	4.58	36	3.05
2	3.55	37	5.89		2	4.72	37	3.35
3	3.8	38	5.82		3	4.3	38	3.69
4	4.04	39	5.95		4	4.31	39	4.71
5	4.24	40	5.93		5	4.53	40	4.71
6	4.28	41	5.97		6	4.66	41	3.86
7	4.32	42	6.05		7	4.7	42	3.94
8	4.38	43	6.13		8	4.79	43	3.95
9	4.49	44	6.15		9	4.96	44	4.07
10	4.62	45	6.24		10	5.21	45	4.63
11	4.96	46	6.64		11	5.33	46	4.53
12	4.94	47	6.64		12	5.54	47	4.63
13	4.99	48	6.69		13	3.78	48	4.8
14	5.11	49	6.15		14	3.6	49	4.75
15	4.87	50	4.55		15	3.87	50	4.9
16	5.11	51	4.69		16	4.1	51	5.03
17	5.24	52	5.11		17	4.14	52	5.22
18	5.28	53	4.75		18	4.25	53	5.28
19	5.3	54	4.89		19	4.25	54	5.31
20	5.37	55	4.98		20	4.41	55	5.34
21	5.37	56	5.08		21	4.62	56	5.4
22	5.47	57	5.16		22	4.69	57	5.46
23	5.51	58	5.3		23	4.63	58	5.63
24	5.66	59	5.62		24	4.73	59	5.54
25	5.4	60	5.8		25	4.91	Av	4.66
26	5.51	61	5.99		26	4.96	S	0.568
27	5.28	62	6.18		27	4.93	S^2	0.322
28	5.26	63	6.35		28	4.7	S^2/n	0.005
29	5.34	64	6.43		29	4.72		
30	5.24	65	6.53		30	4.84	Upper bound	1.96
31	5.47	66	6.56		31	4.8	Lower Bound	-1.96
32	5.52	67	6.56		32	5.11		
33	5.59	68	6.58		33	5.04	t	6.75
34	5.75	69	6.66		34	5.18		
35	5.88	70	6.34		35	5.28		
		71	6.39					
		72	5.11					
		73	5.92					
		Av	5.45					
		S	0.78					
		S^2	0.61					
		S^2/n	0.01					

BH4371								
t1(Months)	**sample1**				**t2(months)**	**sample2**		
1	2.61	36	5.19		1	5.66	36	4.28
2	2.85	37	5.15		2	5.42	37	4.42
3	2.99	38	5.09		3	5.18	38	4.52
4	3.3	39	5.19		4	4.1	39	2.79
5	3.47	40	5.2		5	4.06	40	2.7
6	3.54	41	5.23		6	3.6	41	2.95
7	3.56	42	5.32		7	3.67	42	3.01
8	3.63	43	5.38		8	3.8	43	3.08
9	3.72	44	5.4		9	3.92	44	3.14
10	3.85	45	5.48		10	3.95	45	3.2
11	3.98	46	5.77		11	4.05	46	3.22
12	4.15	47	5.9		12	4.21	47	3.31
13	4.14	48	5.95		13	4.47	48	4.63
14	4.135	49	5.45		14	4.6	49	3.74
15	4.13	50	3.95		15	4.83	50	3.86
16	4.36	51	3.94		16	3.25	51	4.02
17	4.48	52	3.97		17	3.12	52	4.01
18	4.54	53	4		18	3.24	53	4.13
19	4.55	54	4.14		19	3.4	54	4.26
20	4.575	55	4.22		20	3.47	55	4.45
21	4.6	56	4.31		21	3.55	56	4.54
22	4.72	57	4.4		22	3.51	57	4.56
23	4.75	58	4.53		23	3.64	58	4.62
24	4.91	59	4.86		24	3.93	59	4.65
25	4.86	60	5.05		25	3.95	60	4.72
26	4.76	61	5.25		26	3.9	61	4.87
27	4.54	62	5.47		27	4.02	62	4.81
28	4.51	63	5.61		28	4.16	**Av**	4.04
29	4.59	64	5.71		29	4.23	**S**	0.57
30	4.69	65	5.8		30	4.22	**S^2**	0.33
31	4.73	66	5.85		31	3.97	**S^2/n**	0.01
32	4.77	67	5.84		32	3.98		
33	4.84	68	5.86		33	4	**Upper bound**	1.96
34	4.99	69	5.93		34	4.05	**Lower Bound**	-1.96
35	5.13	70	5.63		35	4.165		
		71	5.66				**t**	5.67
		Av	4.70					
		S	0.80					
		S^2	0.64					
		S^2/n	0.01					

BHZ6424								
t1(Months)	**sample1**				**t2(months)**	**sample2**		
1	23	36	26.6		1	25.21	36	21.37
2	23.04	37	26.72		2	24.85	37	21.57
3	23.23	38	26.63		3	23.77	38	21.75
4	23.42	39	26.73		4	23.81	39	21.78
5	23.64	40	26.71		5	24.03	40	21.82
6	23.74	41	26.8		6	24.41	41	21.85
7	23.83	42	26.94		7	24.47	42	22.02
8	24.02	43	27.13		8	24.69	43	22.06
9	24.15	44	27.17		9	24.84	44	22.25
10	24.26	45	27.33		10	25.07	45	22.46
11	24.4	46	27.46		11	24.26	46	22.67
12	24.55	47	27.62		12	25.61	47	22.83
13	24.67	48	27.69		13	21.49	48	22.99
14	24.62	49	27.64		14	19.96	49	23.03
15	24.56	50	29.39		15	21.01	50	23.21
16	24.77	51	26.82		16	21.23	51	23.4
17	24.96	52	26.78		17	21.36	52	23.66
18	25.08	53	26.74		18	21.58	53	23.87
19	25.21	54	26.94		19	21.71	54	23.97
20	25.4	55	27.1		20	22.02	55	24.21
21	25.39	56	27.22		21	22.3	56	24.29
22	25.6	57	27.36		22	22.35	57	24.24
23	25.65	58	27.46		23	22.43	58	24.62
24	25.78	59	27.64		24	22.56	59	24.62
25	25.6	60	27.76		25	22.89	**Av**	23.14
26	25.67	61	27.97		26	22.99	**S**	1.26
27	25.5	62	28.11		27	23.24	**S^2**	1.59
28	25.6	63	28.24		28	23.2	**S^2/n**	0.03
29	25.73	64	28.34		29	23.32		
30	25.93	65	28.46		30	23.5	**Upper bound**	1.96
31	25.99	66	28.6		31	23.74	**Lower Bound**	-1.96
32	26.15	67	28.7		32	24.01		
33	26.26	68	28.79		33	24.27	**t**	12.72
34	26.39	69	28.92		34	24.17		
35	26.5	70	28.69		35	24.3		
		71	28.54					
		72	27.83					
		73	27.11					
		Av	26.31					
		S	1.61					
		S^2	2.59					
		S^2/n	0.04					

BH4995

t1(Months)	sample1				t2(months)	sample2			
1	8.82	36	18		1	21.4	36		10.95
2	9.46	37	18.25		2	14.7	37		11.51
3	10.37	38	16.71		3	19.85	38		11.82
4	11.16	39	17.06		4	12.57	39		7.25
5	11.94	40	14.7		5	13.19	40		7.14
6	12.42	41	17.49		6	12.95	41		7.74
7	13.05	42	17.85		7	12.78	42		9.27
8	13.72	43	18.37		8	14.75	43		9.75
9	14.31	44	18.53		9	15.82	44		9.94
10	14.65	45	18.86		10	16.05	45		10.42
11	15.06	46	19.12		11	16.53	46		12
12	15.45	47	19.32		12	16.89	47		11.34
13	15.79	48	19.49		13	17.32	48		11.65
14	14.7	49	16.33		14	17.65	49		11.92
15	14.6	50	12.62		15	18.11	50		12.09
16	13.59	51	14.25		16	7.19	51		12.16
17	15.92	52	14.7		17	7.57	52		14.7
18	16.36	53	15.99		18	9.04	53		13
19	16.72	54	16.93		19	10.23	54		13.46
20	17.2	55	17.51		20	10.8	55		13.66
21	17.23	56	18.01		21	11.58	56		14.2
22	17.77	57	18.32		22	11.92	57		14.95
23	17.94	58	18.58		23	12.9	58		14.68
24	18.17	59	18.97		24	13.27	59		15.06
25	16.86	60	19.16		25	13.4	**Av**		12.36
26	16.54	61	19.53		26	10.25	**S**		3.21
27	16.32	62	19.76		27	10.67	**S^2**		10.31
28	14.51	63	19.96		28	10.5	**S^2/n**		0.17
29	15.26	64	20.21		29	10.69			
30	15.95	65	20.4		30	8.6	**Upper bound**		1.96
31	16.27	66	20.65		31	7.46	**Lower Bound**		-1.96
32	16.84	67	20.83		32	8.19			
33	17.18	68	21		33	9.04	**t**		8.19
34	17.54	69	21.21		34	10.08			
35	17.78	70	21.29		35	14.7			
		71	21.4						
		Av	16.74						
		S	2.81						
		S^2	7.92						
		S^2/n	0.11						

BH6423							
t1(Months)	**sample1**			**t2(months)**	**sample2**		
1	3.75	36	8.94	1	3.78	36	3.1
2	4.54	37	8.98	2	3.98	37	2.98
3	4.82	38	9.56	3	3.08	38	3.47
4	5.46	39	10.53	4	3.47	39	3.41
5	6.04	40	10.67	5	3.92	40	3.68
6	6.62	41	10.99	6	4.73	41	3.74
7	6.88	42	11.45	7	5.06	42	4.22
8	7.2	43	11.79	8	5.68	43	4.8
9	7.7	44	11.98	9	6.28	44	5.36
10	7.38	45	11.72	10	6.95	45	6.81
11	7.06	46	7.75	11	7.29	46	6.16
12	5.82	47	7.6	12	8.03	47	5.74
13	6.93	48	7.06	13	2.53	48	6.12
14	7.49	49	8.45	14	2.46	49	6.6
15	8.15	50	9.27	15	2.63	50	7.23
16	8.52	51	9.91	16	3.35	51	7.65
17	9.03	52	10.5	17	3.21	52	7.86
18	9.16	53	10.89	18	3.61	53	8.27
19	9.08	54	11.01	19	3.69	54	8.39
20	8.81	55	11.25	20	4.31	55	8.36
21	9.05	56	11.54	21	4.83	56	8.42
22	6.87	57	11.75	22	5.05	57	7.45
23	6.28	58	11.91	23	3.95		
24	5.58	59	12.02	24	4.21	**Av**	5.08
25	4.88	60	11.99	25	4.43	**S**	1.75
26	5.22	61	12.24	26	4.5	**S^2**	3.06
27	5.72	62	12.61	27	3.87	**S^2/n**	0.05
28	6.17	63	12.68	28	3.4		
29	7.17	64	12.94	29	3.63	**Upper bound**	1.96
30	7.73	65	13.27	30	4.04	**Lower Bound**	-1.96
31	8.4	66	12.19	31	4.66		
32	8.82	67	11.66	32	7.06	**t**	10.22
33	9.4	68	7.06176	33	5.58		
34	9.3	69	9.56	34	6.04		
35	8.36			35	6.27		
		Av	8.89				
		S	2.42				
		S^2	5.88				
		S^2/n	0.09				

Z4401						
t1(Months)	**sample1**		**t2(months)**	**sample2**		
1	6.45		1	5.86	36	6.42
2	6.7		2	5.17	37	4.25
3	6.7		3	5.11	38	4.46
4	6.74		4	4.7	39	4.54
5	6.8		5	4.76	40	4.62
6	6.98		6	5.01	41	4.7
7	6.93		7	5.24	42	4.88
8	7.12		8	5.29	43	4.85
9	7.29		9	5.41	44	5.05
10	7.57		10	5.59	45	5.39
11	7.6		11	5.86	46	5.65
12	7.04		12	8	47	5.7
13	6.23		13	6.06	48	5.9
14	5.43		14	4.13	49	4.8
15	5.39		15	5.35	50	6.07
16	5.34		16	4.27	51	6.21
17	5.57		17	4.45	52	6.34
18	5.68		18	4.6	53	6.47
19	5.83		19	4.73	54	7.22
20	5.9		20	4.8	55	6.64
21	6.03		21	5.01	56	6.65
22	6.4		22	5.17	57	6.8
23	6.51		23	5.21	58	6.92
24	6.78		24	5.1	**Av**	5.52
25	6.9		25	5.2	**S**	0.89
26	7.14		26	5.55	**S^2**	0.79
27	7.21		27	5.52	**S^2/n**	0.01
28	7.39		28	5.58		
29	7.41		29	5.25	**Upper bound**	**1.96**
30	7.14		30	5.31	**Lower Bound**	**-1.96**
31	7.47		31	5.5		
32	7.59		32	5.44	**t**	**35.37**
33	7.25		33	6.84		
34	7.23		34	8.23		
35	6.54		35	6.39		
Av	6.69					
S	0.68					
S^2	0.46					
S^2/n	0.01					

BH4887								
t1(Months)	**sample1**				**t2(months)**	**sample2**		
1	2.42	36	3.02		1	5.14	37	2.91
2	2.43	37	3.25		2	3.28	38	2.53
3	2.61	38	3.44		3	3.8	39	2.55
4	2.72	39	3.24		4	2.91	40	2.7
5	2.91	40	3.28		5	2.91	41	2.05
6	2.7	41	3.32		6	2.16	42	2.5
7	2.72	42	3.27		7	2.2	43	2.28
8	2.86	43	3.26		8	2.15	44	2.47
9	2.94	44	3.7		9	2.25	45	2.39
10	2.81	45	3.54		10	2.35	46	2.42
11	3.12	46	3.72		11	2.42	47	2.55
12	3.1	47	4.33		12	2.43	48	2.62
13	2.91	48	2.78		13	2.51	49	2.62
14	2.76	49	1.36		14	2.71	50	2.71
15	3.08	50	2.12		15	3.02	51	2.6
16	3.17	51	2.91		16	3.19	52	2.61
17	3.1	52	2.48		17	2.91	53	2.72
18	3.02	53	2.67		18	1.9	54	2.83
19	2.99	54	2.73		19	2	55	2.8
20	3	55	2.78		20	2.2	56	2.8
21	3.17	56	2.89		21	2.3	57	2.82
22	3.28	57	3.11		22	2.6	58	2.84
23	3.55	58	3.37		23	2.27	59	2.86
24	2.94	59	3.63		24	2.3	60	3.03
25	2.94	60	4.02		25	2.37	**Av**	2.593
26	2.68	61	4.54		26	2.48	**S**	0.472
27	2.43	62	4.78		27	2.51	**S^2**	0.222
28	2.62	63	4.67		28	2.3	**S^2/n**	0.004
29	2.72	64	4.78		29	2.34		
30	2.71	65	4.94		30	2.42	Upper bound	1.96
31	2.76	66	4.95		31	2.5	Lower Bound	-1.96
32	2.83	67	5.01		32	2.28		
33	3.1	**Av**	3.188		33	2.24	t	-5.64
34	3.35	**S**	0.705		34	2.32		
35	3.25	**S^2**	0.496		35	2.35		
		S^2/n	0.007					

APPENDIX C - ACF AND PACF for different borehole series

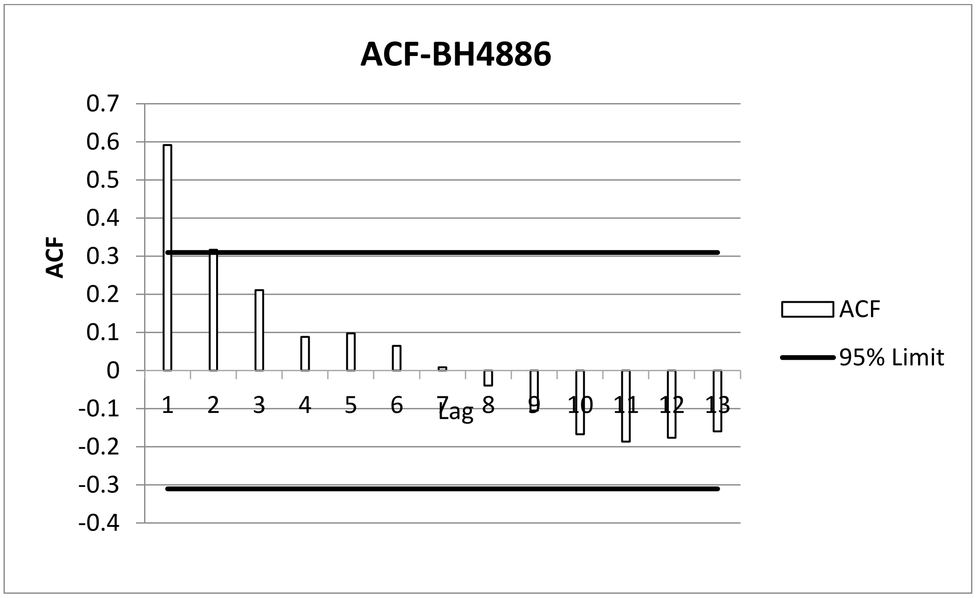


Fig B1: ACF Plot for BH4886

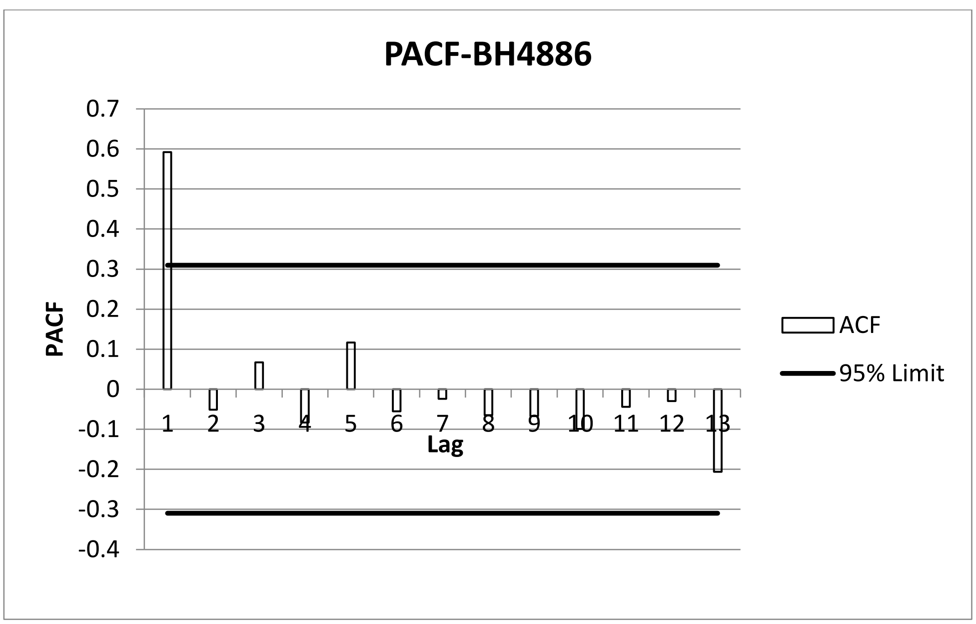


Fig B2: PACF Plot for BH4886

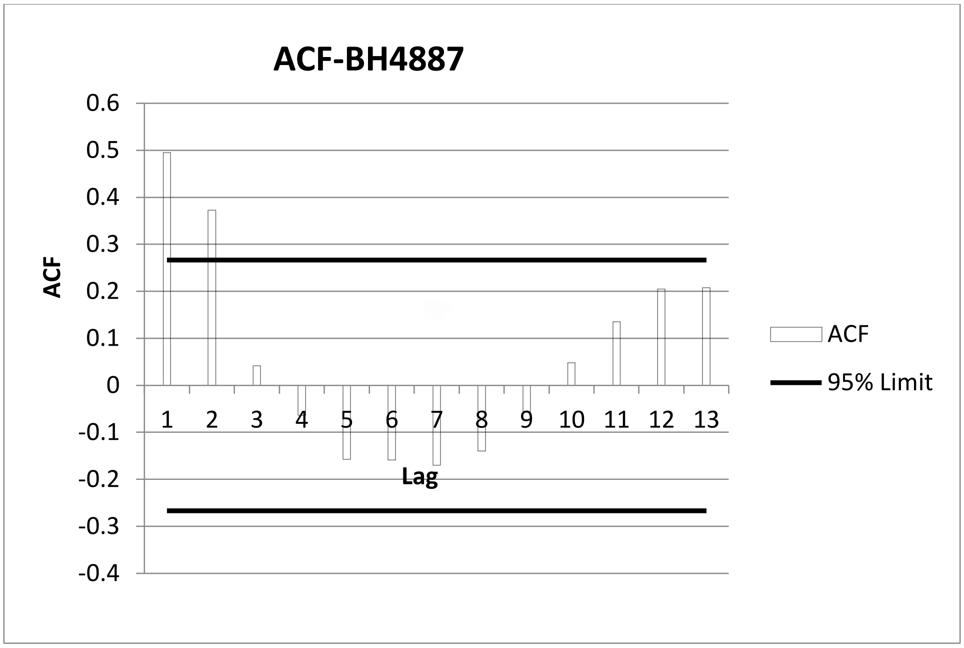


Fig B3: ACF Plot for BH4887

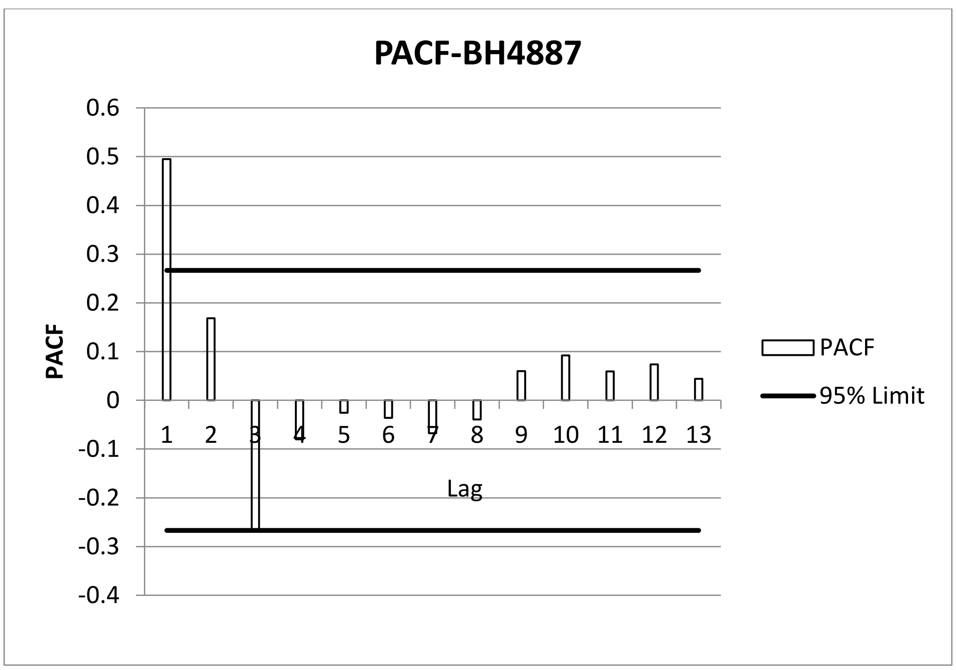


Fig B4: PACF Plot for BH4887

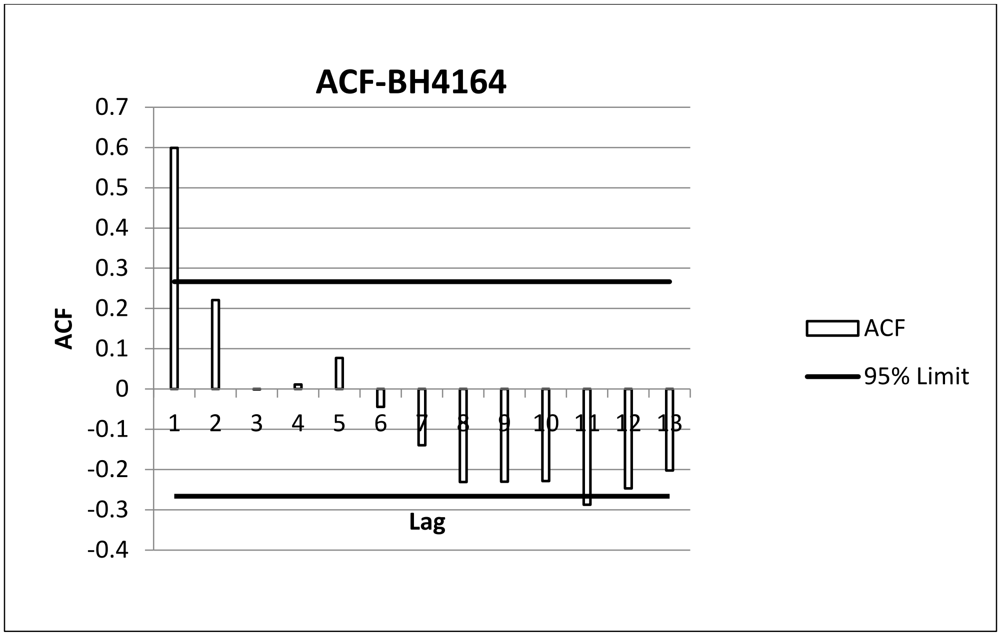


Fig B5: ACF Plot for BH4164

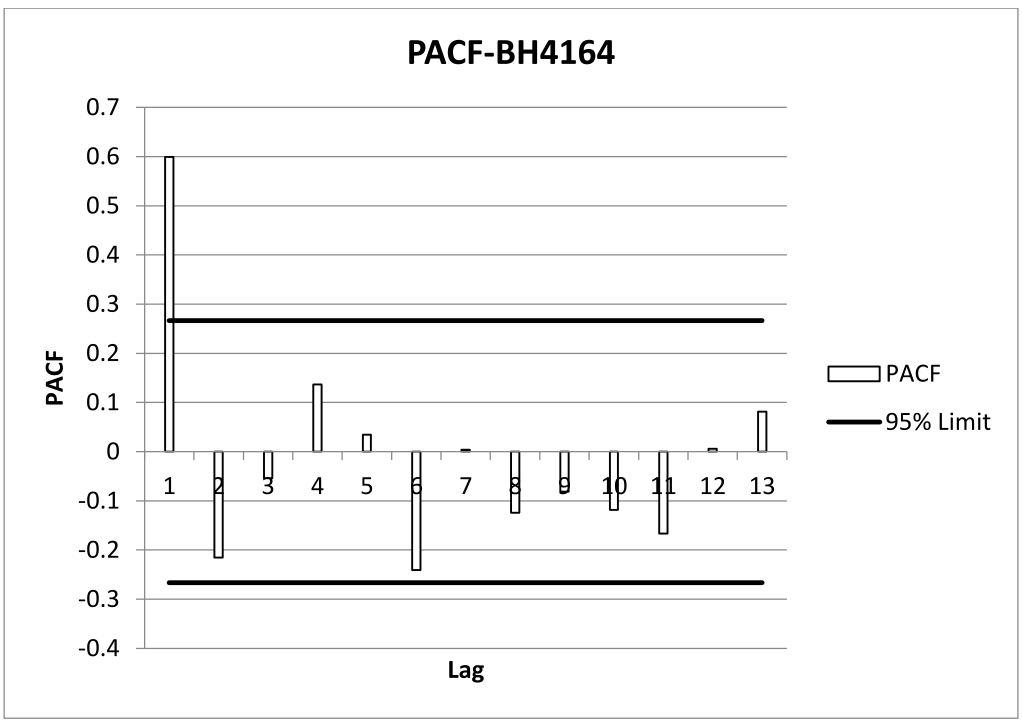


Fig B6: PACF Plot for BH4164

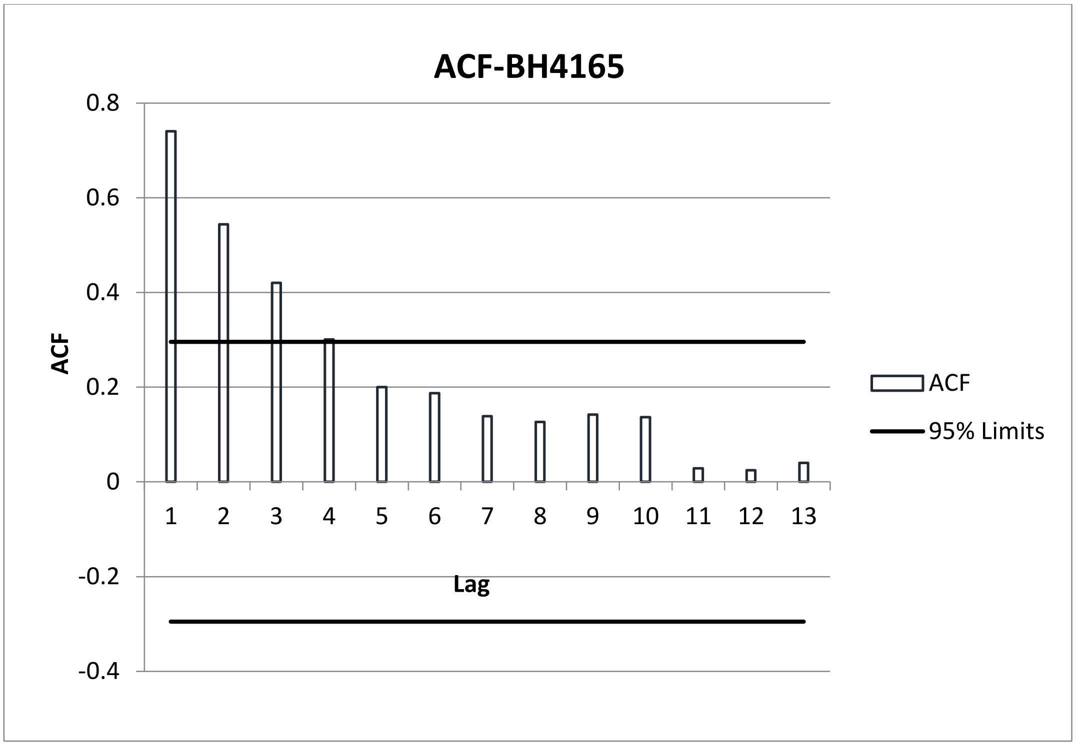


Fig B7: ACF Plot for BH4165

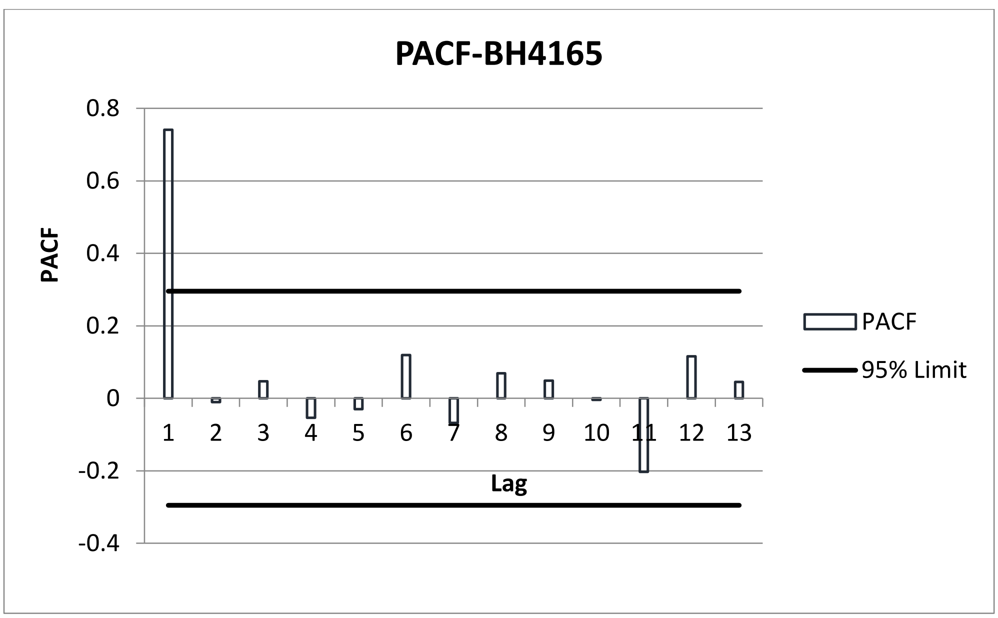


Fig B8: PACF Plot for BH4165

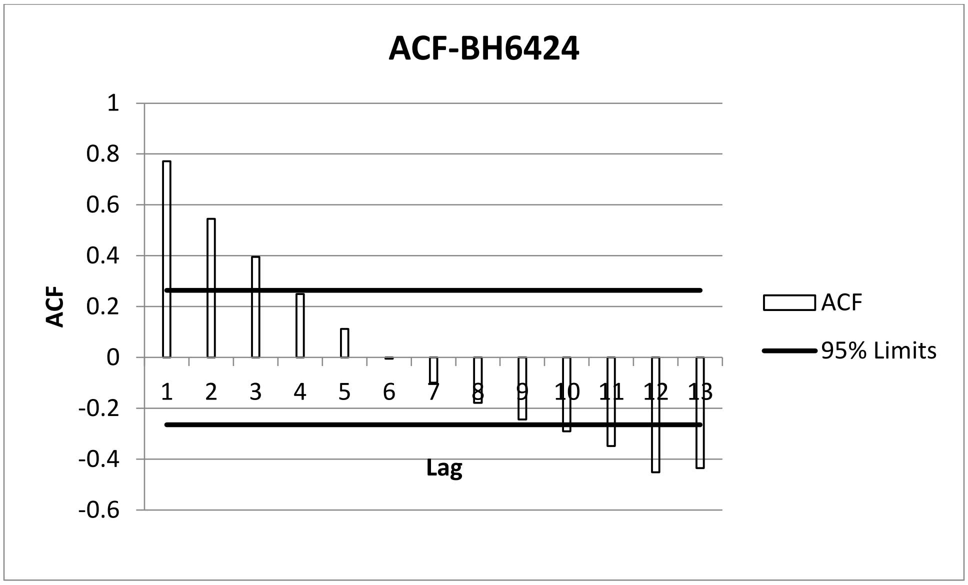


Fig B9: ACF Plot for BH6424

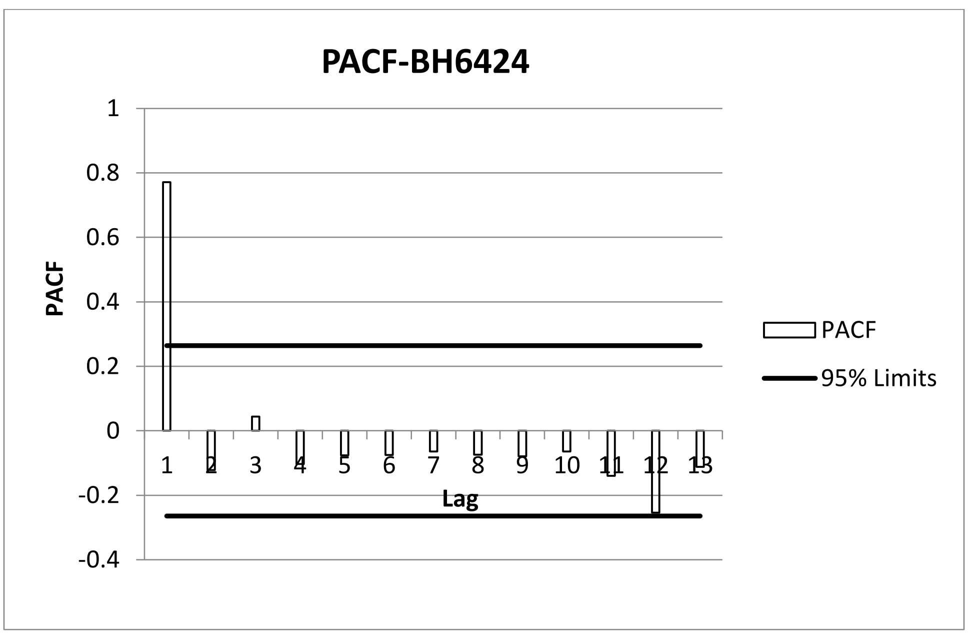


Fig B10: PACF Plot for BH6424

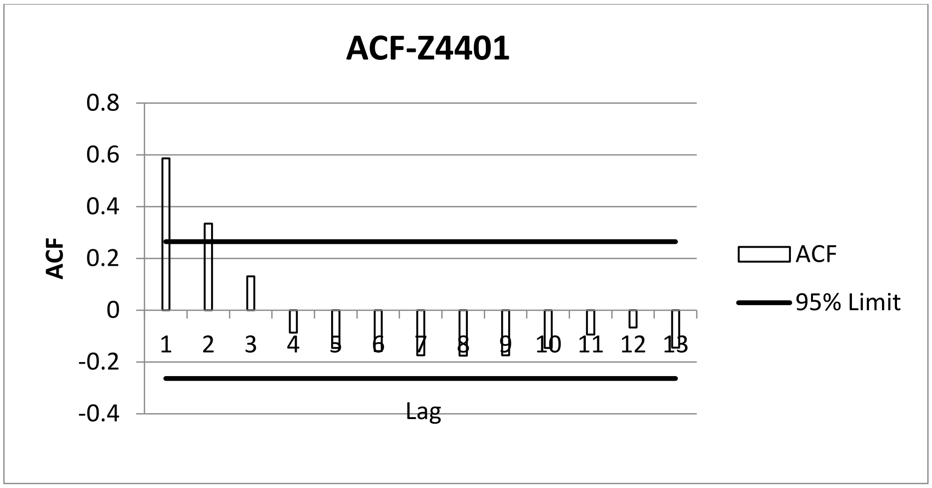


Fig B11: ACF Plot for Z4401

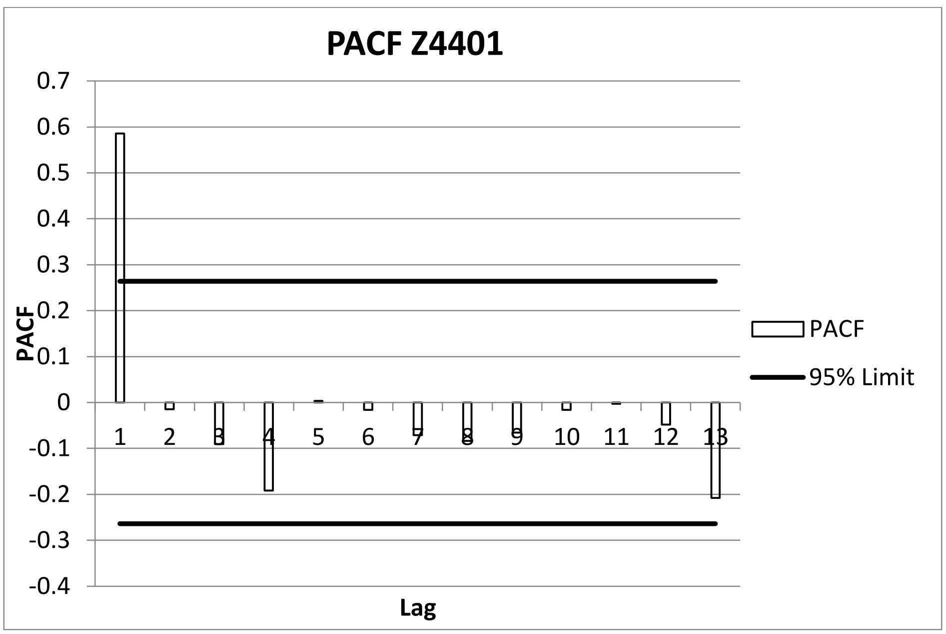


Fig B12: PACF Plot for Z4401

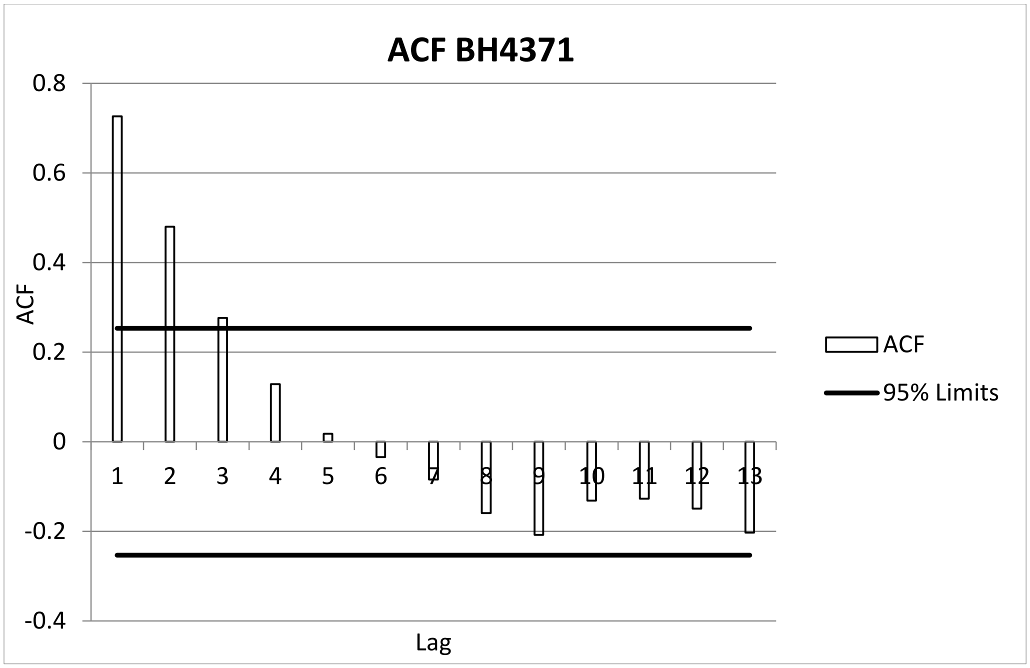


Fig B13: ACF Plot for BH4371

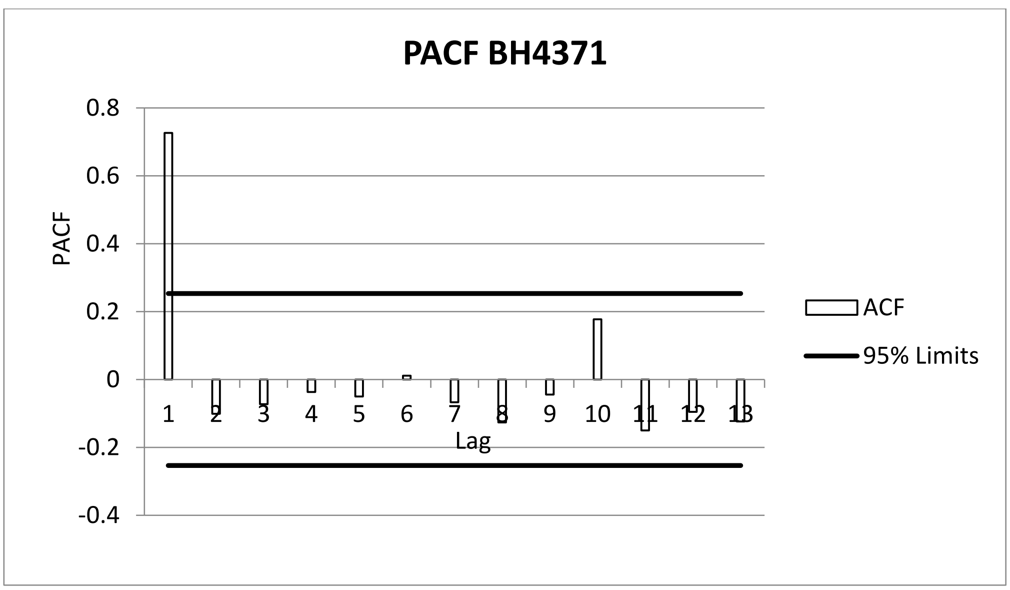


Fig B14: PACF Plot for BH4371

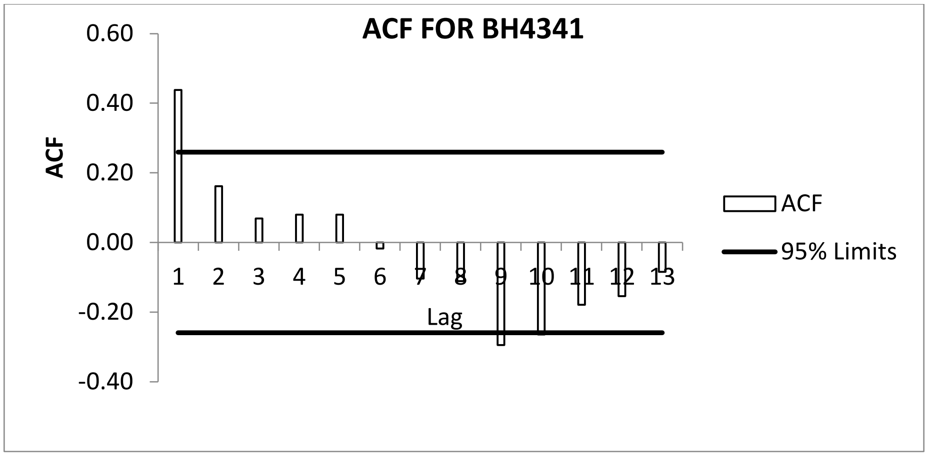


Fig B15: ACF Plot for BH4341

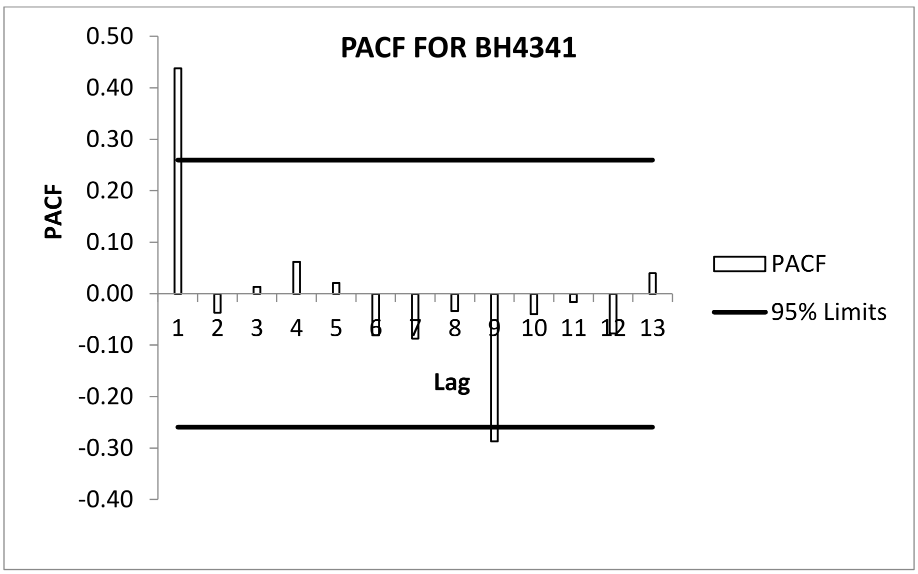


Fig B16: PACF Plot for BH4341

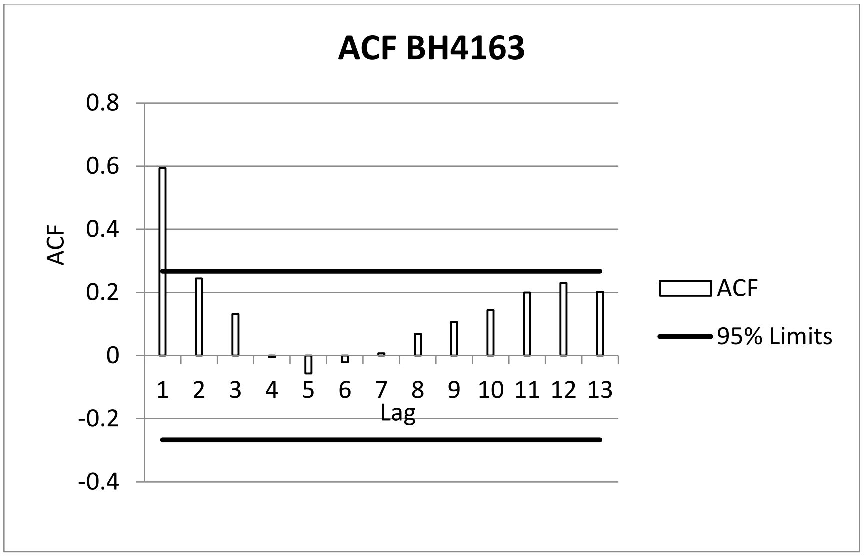


Fig B17: ACF Plot for BH4163

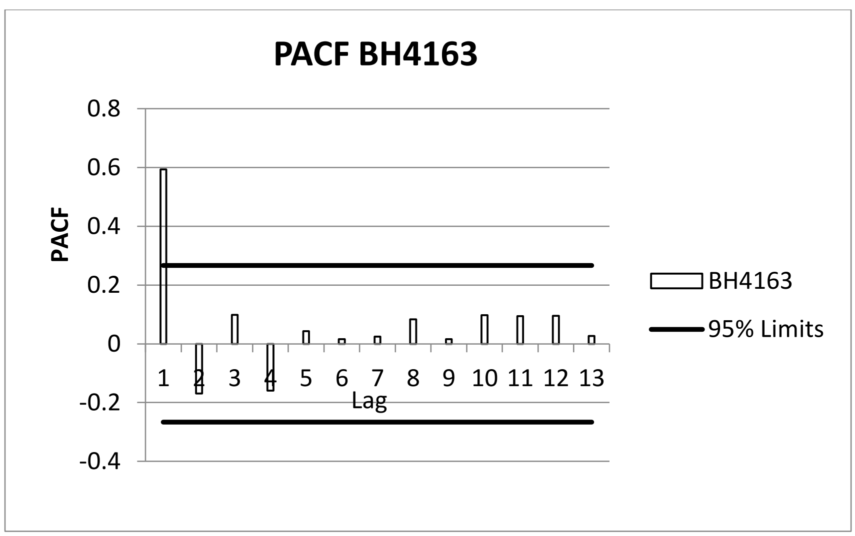


Fig B18: PACF Plot for BH4163

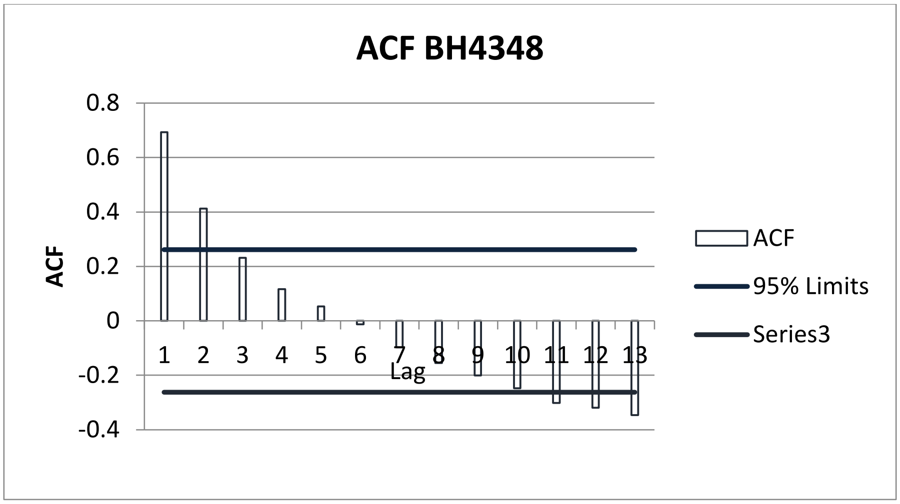


Fig B19: ACF Plot for BH4348

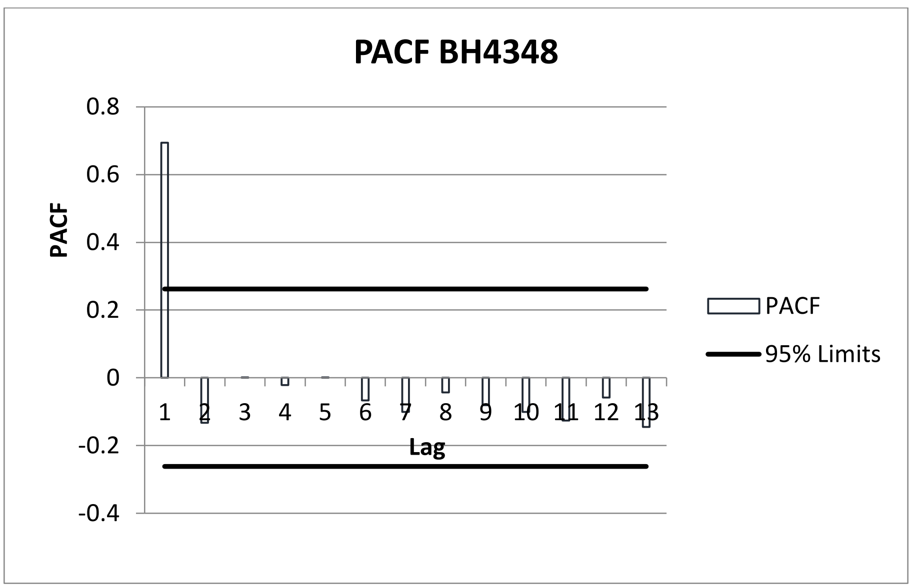


Fig B20: PACF Plot for BH4348

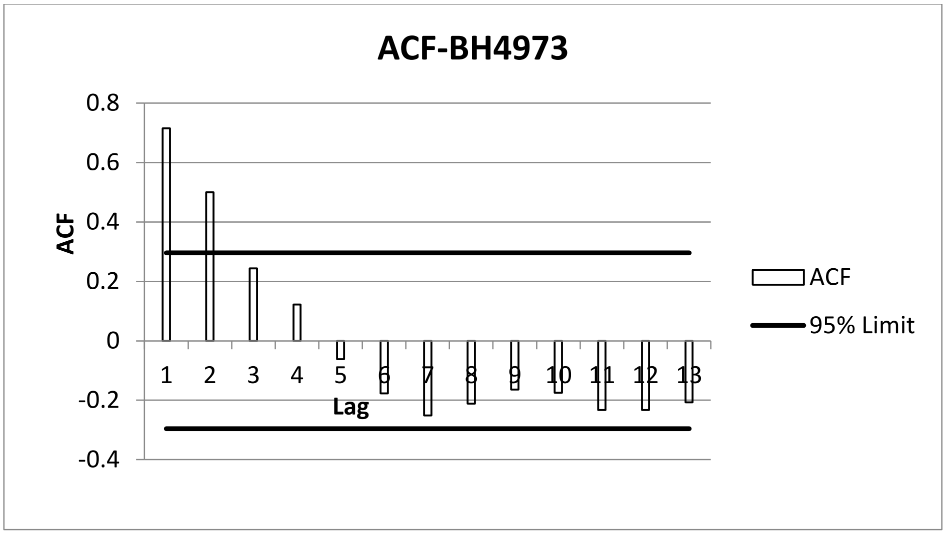


Fig B21: ACF Plot for BH4973

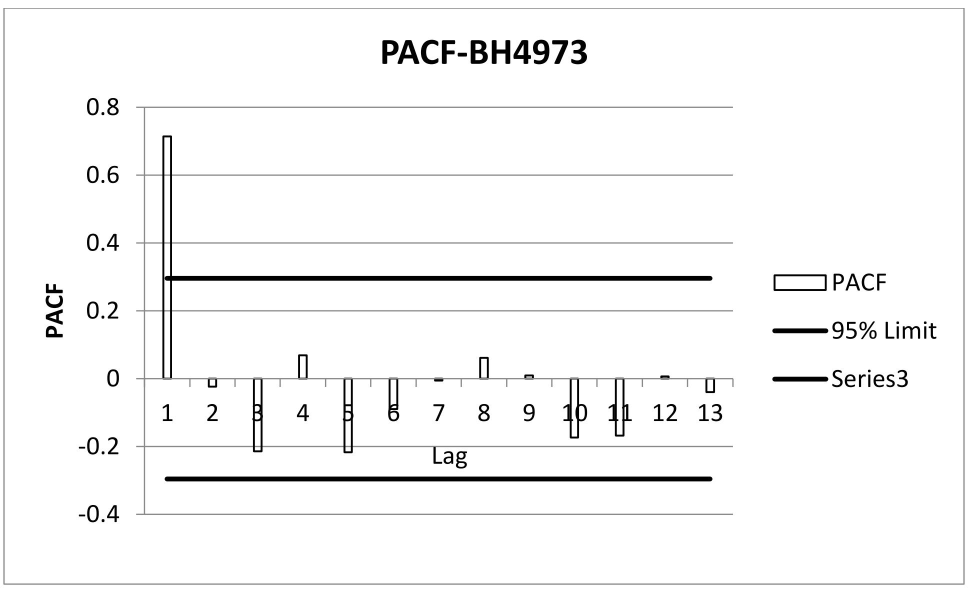


Fig B22: PACF Plot for BH4973

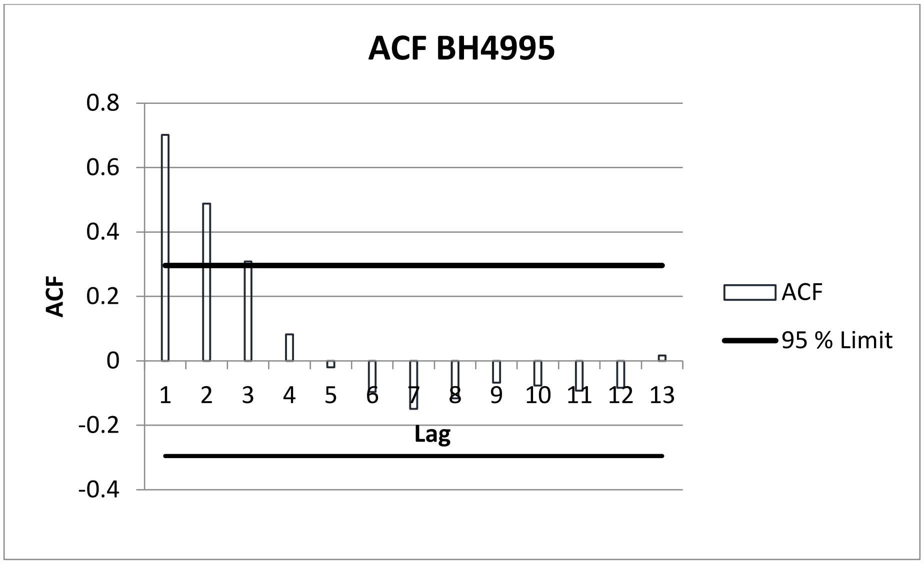


Fig B23: ACF Plot for BH4995

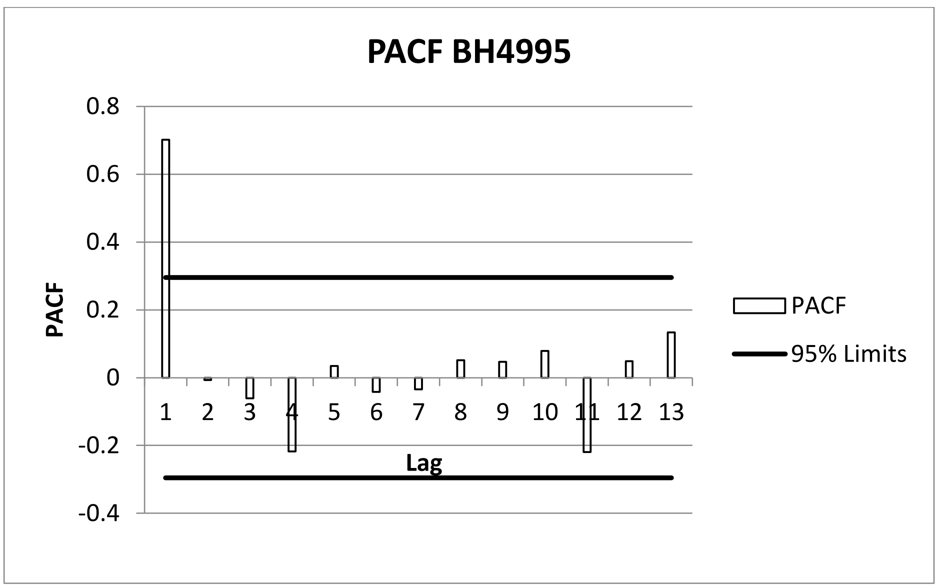


Fig B24: PACF Plot for BH4995

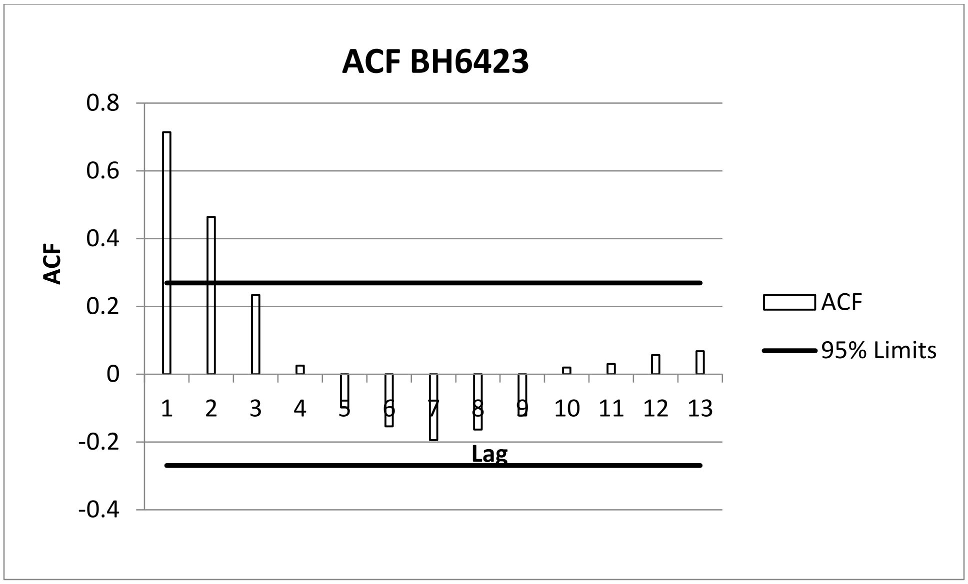


Fig B25: ACF Plot for BH6423

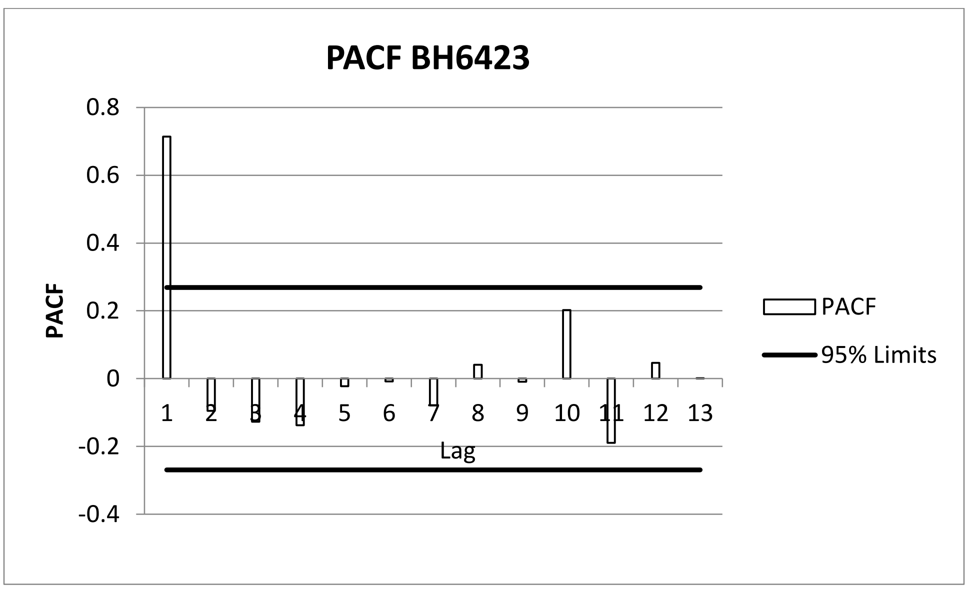


Fig B26: PACF Plot for BH6423

APPENDIX D – Thomas Fiering models fitted at different boreholes

Table D1: BH 4341

Month	**S**	**Mean**	**Z**	**Correlation**	**Thomas-Fiering Model**
January	0.63	1.65	0.43	-0.2	$X_{jan} = 1.65 - 0.6(X_{dec} - 2.16) + 0.98Z_1$
February	0.37	1.45	-2.12	0.35	$X_{Feb} = 1.45 + 0.21(X_{jan} - 1.65) + 0.34Z_{2}$
March	0.24	1.74	1.08	0.91	$X_{mar} = 1.74 + 0.59(X_{feb} - 1.45) + 0.234_{3}$
April	0.58	1.99	-0.27	0.79	$X_{Apr} = 1.99 + 1.91(X_{mar} - 1.74) + 0.35Z_{4}$
May	0.27	1.86	-2.17	-0.09	$X_{may} = 1.86 - 0.04(X_{Apr} - 1.99) + 0.27Z_{5}$
June	0.18	1.80	0.02	0.165	$X_{jun} = 1.8 + 0.11(X_{may} - 1.86) + 0.19Z_{6}$
July	0.15	1.86	-0.72	0.98	$X_{jul} = 1.86 + 0.82(X_{jun} - 1.8) + 0.03Z_{7}$
August	0.36	1.92	0.21	0.87	$X_{Aug} = 1.92 + 2.09(X_{jul} - 1.86) + 0.18Z_{8}$
September	0.19	1.98	-0.56	0.73	$X_{sept} = 1.98 + 0.39(X_{Aug} - 1.92) + 0.13Z_{9}$
October	0.12	2.08	0.46	0.96	$X_{oct} = 2.08 - 0.61(X_{sept} - 1.98) + 0.03Z_{10}$
November	0.14	2.11	-1.81	0.89	$X_{nov} = 2.11 + 1.04(X_{oct} - 2.08) + 0.06Z_{11}$
December	0.21	2.16	1.52	0.85	$X_{dec} = 2.16 + 1.28(X_{nov} - 2.11) + 0.11Z_{12}$

Table D2:BH6423

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
February	0.97	3.58	-0.39	-0.04	$X_{feb} = 3.58 - 0.02(X_{jan} - 5.84) + 0.97Z_2$
March	1.06	3.65	-0.23	0.99	$X_{mar} = 3.65 + 1.08(X_{feb} - 3.58) + 0.15Z_3$
April	0.63	3.19	-0.09	0.91	$X_{apr} = 3.19 + 0.54(X_{mar} - 3.65) + 0.26Z_4$
May	0.06	3.41	-0.46	0.27	$X_{may} = 3.41 + 0.03(X_{apr} - 3.19) + 0.06Z_5$
June	0.36	3.59	1.56	0.98	$X_{jun} = 3.59 + 5.88(X_{may} - 3.41) + 0.07Z_6$
July	0.57	4.13	-1.09	0.97	$X_{jul} = 4.13 + 1.54(X_{jun} - 3.59) + 0.14Z_7$
August	0.70	4.47	0.86	0.93	$X_{aug} = 4.47 + 1.14(X_{jul} - 4.13) + 0.26Z_8$
September	1.38	5.68	0.38	0.67	$X_{sept} = 5.68 + 1.32(X_{aug} - 4.47) + 1.02Z_9$
October	0.73	5.56	0.71	0.52	$X_{oct} = 5.56 + 0.28(X_{sept} - 5.68) + 0.62Z_{10}$
November	0.95	6.01	0.42	0.99	$X_{nov} = 6.01 + 1.29(X_{oct} - 5.56) + 0.13Z_{11}$
December	1.71	5.84	1.45	0.98	$X_{dec} = 5.84 + 1.76(X_{nov} - 6.01) + 0.34Z_{12}$
January	2.59	5.11	0.90	0.57	$X_{jan} = 5.11 - 0.86(X_{dec} - 5.84) + 2.13Z_1$

Table D3:BH4165

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	1.13	3.56	1.45	0.83	$X_{jan} = 3.57 + 1.13(X_{dec} - 3.82) + 0.63Z_1$
February	0.90	2.61	0.90	0.46	$X_{Feb} = 2.61 + 0.36(X_{jan} - 3.57) + 0.8Z_2$
March	0.64	3.66	0.44	-0.11	$X_{mar} = 3.66 - 0.08(X_{feb} - 2.61) + 0.64_3$
April	0.47	3.52	-2.12	0.86	$X_{Apr} = 3.52 + 0.63(X_{mar} - 3.66) + 0.24Z_4$
May	0.62	3.45	1.09	0.96	$X_{may} = 3.45 + 1.27(X_{Apr} - 3.52) + 0.17Z_5$
June	0.58	3.42	-0.28	0.97	$X_{jun} = 3.42 + 0.9(X_{may} - 3.45) + 0.14Z_6$
July	0.49	3.42	-2.17	0.99	$X_{jul} = 3.42 + 0.84(X_{jun} - 3.42) + 0.07Z_7$
August	0.51	3.39	0.02	0.99	$X_{Aug} = 3.39 + 1.04(X_{jul} - 3.42) + 0.08Z_8$
September	0.58	3.96	-0.72	-0.19	$X_{sept} = 3.96 - 0.21(X_{Aug} - 3.39) + 0.57Z_9$
October	0.61	3.61	0.21	-0.28	$X_{oct} = 3.61 - 0.3(X_{sept} - 3.96) + 0.59Z_{10}$
November	0.62	3.84	-0.56	0.97	$X_{nov} = 3.83 + 0.99(X_{oct} - 3.61) + 0.15Z_{11}$
December	0.83	3.82	0.47	0.9	$X_{dec} = 3.82 + 1.2(X_{nov} - 3.83) + 0.36Z_{12}$

Table D4:BH4164

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	1.05	4.53	0.904	0.95	$X_{jan} = 4.53 + 2.38(X_{dec} - 4.97) + 0.33Z_{1}$
February	0.68	4.16	0.437	0.51	$X_{Feb} = 4.16 + 0.33(X_{jan} - 4.53) + 0.58Z_{2}$
March	0.67	4.24	-2.12	0.97	$X_{mar} = 4.24 + 0.96(X_{feb} - 4.16) + 0.16Z_{3}$
April	0.48	4.45	1.08	0.03	$X_{Apr} = 4.45 + 0.02(X_{mar} - 4.24) + 0.48Z_{4}$
May	0.43	4.46	-0.27	0.99	$X_{may} = 4.46 + 0.89(X_{Apr} - 4.45) + 0.06Z_{5}$
June	0.55	4.31	-2.17	0.99	$X_{jun} = 4.31 + 1.27(X_{may} - 4.46) + 0.08Z_{6}$
July	0.53	4.42	0.02	0.98	$X_{jul} = 4.42 + 094(X_{jun} - 3.42) + 0.07Z_{7}$
August	0.53	4.43	-0.72	0.93	$X_{Aug} = 4.43 + 1.04(X_{jul} - 4.42) + 0.11Z_{8}$
September	0.53	4.59	0.21	0.96	$X_{sept} = 4.59 + 0.96(X_{Aug} - 4.43) + 0.15Z_{9}$
October	0.35	4.81	-0.55	0.98	$X_{oct} = 4.81 + 0.65(X_{sept} - 4.59) + 0.07Z_{10}$
November	0.44	4.90	0.465	-0.21	$X_{nov} = 4.90 - 0.26(X_{oct} - 4.81) + 0.43Z_{11}$
December	0.42	4.97	-1.81	-0.433	$X_{dec} = 4.97 - 0.41(X_{nov} - 4.90) + 0.38Z_{12}$

Table D5:BH4163

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	0.76	2.04	-0.09	0.15	$X_{jan} = 2.04 + 0.28(X_{dec} - 2.85) + 0.75Z_{1}$
February	0.34	1.63	-0.46	-0.49	$X_{Feb} = 1.63 - 0.22(X_{jan} - 2.04) + 0.3Z_{2}$
March	0.29	1.92	1.56	0.87	$X_{mar} = 1.92 + 0.74(X_{feb} - 1.63) + 0.14Z_{3}$
April	0.23	1.88	-1.08	0.21	$X_{Apr} = 1.88 + 0.17(X_{mar} - 1.92) + 0.22Z_{4}$
May	0.30	2.01	0.86	0.94	$X_{may} = 2.01 + 1.23(X_{Apr} - 1.88) + 0.1Z_{5}$
June	0.33	2.05	0.38	0.81	$X_{jun} = 2.05 + 0.89(X_{may} - 2.01) + 0.19Z_{6}$
July	0.32	2.25	0.71	0.98	$X_{jul} = 2.25 + 0.95(X_{jun} - 2.05) + 0.06Z_{7}$
August	0.42	2.32	0.421	0.95	$X_{Aug} = 2.32 + 1.25(X_{jul} - 2.25) + 0.11Z_{8}$
September	0.42	2.76	1.454	0.08	$X_{sept} = 2.76 + 0.08(X_{Aug} - 2.32) + 0.99Z_{9}$
October	0.33	2.67	0.904	0.44	$X_{oct} = 2.67 + 0.34(X_{sept} - 2.76) + 0.3Z_{10}$
November	0.88	3.07	0.437	0.33	$X_{nov} = 3.07 - 0.88(X_{oct} - 2.67) + 0.42Z_{11}$
December	0.41	2.85	-2.12	0.79	$X_{dec} = 2.85 - 0.37(X_{nov} - 3.07) + 0.25Z_{12}$

Table D6:BHZ6424

Month	**S**	**Mean**	**Z**	**Correlation**	**Thomas-Fiering Model**
January	1.79	23.13	1.454	0.23	$X_{jan} = 23.13 + 0.52(X_{dec} - 23.46) + 1.74Z_1$
February	1.74	22.79	0.904	0.72	$X_{Feb} = 22.79 + 0.70(X_{jan} - 23.13) + 1.21Z_2$
March	2.06	22.39	0.437	0.93	$X_{mar} = 22.39 + 1.10(X_{feb} - 22.79) + 0.14Z_3$
April	1.28	22.45	-2.12	0.97	$X_{Apr} = 22.45 + 0.6(X_{mar} - 22.39) + 0.31Z_4$
May	1.20	22.51	1.085	0.99	$X_{may} = 22.51 + 0.92(X_{Apr} - 22.45) + 0.17Z_5$
June	1.25	22.64	-0.277	0.99	$X_{jun} = 22.64 + 1.03(X_{may} - 22.51) + 0.18Z_6$
July	1.31	22.88	-2.17	0.99	$X_{jul} = 22.88 + 1.03(X_{jun} - 22.64) + 0.18Z_7$
August	1.32	23.00	0.018	0.997	$X_{Aug} = 23 + 1.00(X_{jul} - 22.88) + 0.10Z_8$
September	1.31	23.24	-0.722	0.99	$X_{sept} = 23.24 + 4.05(X_{Aug} - 23) + 0.18Z_9$
October	1.28	23.47	0.21	0.99	$X_{oct} = 23.47 + 0.97(X_{sept} - 23.24) + 0.18Z_{10}$
November	1.28	23.57	-0.556	0.993	$X_{nov} = 23.57 + 0.99(X_{oct} - 23.47) + 0.15Z_{11}$
December	0.97	23.46	0.465	0.95	$X_{dec} = 23.46 + 0.72(X_{nov} - 23.57) + 0.3Z_{12}$

Table D7:BH4886

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	2.09	5.08	0.904	-0.62	$X_{jan} = 5.08 - 2.95(X_{dec} - 6.4) + 1.64Z_1$
February	1.86	5.54	0.437	-0.34	$X_{Feb} = 5.54 - 0.3(X_{jan} - 5.08) + 1.75Z_2$
March	1.28	5.88	-2.12	0.994	$X_{mar} = 5.88 + 0.68(X_{feb} - 5.54) + 0.18Z_3$
April	1.11	5.91	1.085	0.985	$X_{Apr} = 5.91 + 0.85(X_{mar} - 5.88) + 0.19Z_4$
May	0.92	5.84	-0.277	0.95	$X_{may} = 5.84 + 0.79(X_{Apr} - 5.91) + 0.29Z_5$
June	0.73	5.85	-2.17	0.99	$X_{jun} = 5.85 + 0.79(X_{may} - 5.84) + 0.1Z_6$
July	0.74	5.90	0.018	0.99	$X_{jul} = 5.90 + 1.0(X_{jun} - 5.85) + 0.1Z_7$
August	0.66	5.84	-0.722	0.99	$X_{Aug} = 5.84 + 0.88(X_{jul} - 5.90) + 0.09Z_8$
September	1.08	6.26	0.21	0.993	$X_{sept} = 6.26 + 1.62(X_{Aug} - 5.84) + 0.13Z_9$
October	0.55	6.18	-0.556	0.997	$X_{oct} = 6.18 + 0.5(X_{sept} - 6.26) + 0.04Z_{10}$
November	0.47	6.39	0.465	0.996	$X_{nov} = 6.39 + 0.85(X_{oct} - 6.18) + 0.04Z_{11}$
December	0.44	6.40	-1.812	0.779	$X_{dec} = 6.40 + 0.73(X_{nov} - 6.39) + 0.28Z_{12}$

Table D8:Z4401

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	0.82	5.34	-0.179	0.14	$X_{jan} = 5.34 + 0.09(X_{dec} - 6.36) + 0.81Z_1$
February	0.65	4.83	-0.399	-0.02	$X_{Feb} = 4.83 - 0.02\left(X_{jan} - 5.34\right) + 0.65Z_2$
March	0.43	5.13	-0.235	0.4	$X_{mar} = 5.13 + 0.26\left(X_{feb} - 4.83\right) + 0.39Z_3$
April	0.56	4.79	-0.098	0.41	$X_{Apr} = 4.79 + 0.53(X_{mar} - 5.13) + 0.53Z_4$
May	0.33	4.79	-0.465	0.99	$X_{may} = 4.79 + 0.58\left(X_{Apr} - 4.49\right) + 0.05Z_5$
June	0.29	4.95	1.563	0.97	$X_{jun} = 4.95 + 0.85\left(X_{may} - 4.79\right) + 0.07Z_6$
July	0.35	5.08	-1.085	0.95	$X_{jul} = 5.08 + 1.15\left(X_{jun} - 4.95\right) + 0.11Z_7$
August	0.28	5.15	0.86	0.97	$X_{Aug} = 5.15 + 0.78\left(X_{jul} - 5.08\right) + 0.07Z_8$
September	0.80	5.66	0.388	0.83	$X_{sept} = 5.66 + 2.37\left(X_{Aug} - 5.15\right) + 0.45Z_9$
October	1.40	6.16	0.71	0.99	$X_{oct} = 6.16 + 1.73\left(X_{sept} - 5.66\right) + 0.2Z_{10}$
November	0.49	5.79	0.421	0.89	$X_{nov} = 5.79 + 0.311(X_{oct} - 6.16) + 0.22Z_{11}$
December	1.22	6.36	1.454	0.52	$X_{dec} = 6.36 + 1.29(X_{nov} - 5.79) + 1.04Z_{12}$

Table D9:BH4371

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	1.06	4.21	-0.179	0.497	$X_{jan} = 4.21 + 1.35(X_{dec} - 4.22) + 0.92Z_1$
February	0.70	3.55	-0.399	0.64	$X_{Feb} = 3.55 + 0.42(X_{jan} - 4.21) + 0.54Z_2$
March	0.65	3.59	-0.235	0.97	$X_{mar} = 3.59 + 0.9(X_{feb} - 3.55) + 0.16Z_3$
April	0.53	3.52	-0.098	0.92	$X_{Apr} = 3.52 + 0.75(X_{mar} - 3.59) + 0.2Z_4$
May	0.38	3.53	-0.465	0.97	$X_{may} = 3.53 + 0.70(X_{Apr} - 3.52) + 0.09Z_5$
June	0.37	3.60	1.563	0.99	$X_{jun} = 3.60 + 0.96(X_{may} - 3.53) + 0.05Z_6$
July	0.37	3.67	-1.085	0.99	$X_{jul} = 3.67 + 0.99(X_{jun} - 3.60) + 0.05Z_7$
August	0.39	3.68	0.86	0.99	$X_{Aug} = 3.68 + 1.04(X_{jul} - 3.67) + 0.06Z_8$
September	0.39	3.79	0.388	0.99	$X_{jan} = 4.21 + 1.35(X_{dec} - 4.22) + 0.92Z_9$
October	0.29	4.26	0.71	-0.39	$X_{Feb} = 3.55 + 0.42(X_{jan} - 4.21) + 0.54Z_{10}$
November	0.36	4.15	0.421	-0.31	$X_{mar} = 3.59 + 0.9(X_{feb} - 3.55) + 0.16Z_{11}$
December	0.39	4.22	1.454	0.98	$X_{Apr} = 3.52 + 0.75(X_{mar} - 3.59) + 0.2Z_{12}$

Table D10:BH4348

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	0.45	5.81	-0.179	-0.34	$X_{jan} = 5.81 - 0.33(X_{dec} - 5.46) + 0.06Z_1$
February	0.76	4.72	-0.399	-0.48	$X_{Feb} = 4.72 - 0.81(X_{jan} - 5.81) + 0.67Z_2$
March	1.46	4.36	-0.235	0.82	$X_{mar} = 4.36 + 0.58(X_{feb} - 4.72) + 0.84Z_3$
April	0.88	4.56	-0.098	0.96	$X_{Apr} = 4.56 + 0.75(X_{mar} - 4.36) + 0.25Z_4$
May	0.78	4.70	-0.465	0.998	$X_{may} = 4.70 + 0.88(X_{Apr} - 4.56) + 0.05Z_5$
June	0.59	4.76	1.563	0.98	$X_{jun} = 4.76 + 0.74(X_{may} - 4.70) + 0.11Z_6$
July	0.49	4.88	-1.085	0.99	$X_{jul} = 4.88 + 0.82(X_{jun} - 4.76) + 0.07Z_7$
August	0.56	4.94	0.86	0.99	$X_{Aug} = 4.94 + 1.13(X_{jul} - 4.88) + 0.08Z_8$
September	0.48	5.00	0.388	0.98	$X_{sept} = 5 + 0.84(X_{Aug} - 4.94) + 0.10Z_9$
October	0.49	5.16	0.71	0.99	$X_{oct} = 5.16 + 1.01(X_{sept} - 5) + 0.07Z_{10}$
November	0.47	5.34	0.421	0.97	$X_{nov} = 5.34 + 0.93(X_{oct} - 5.16) + 0.11Z_{11}$
December	0.46	5.46	1.454	0.98	$X_{dec} = 5.46 + 0.96(X_{nov} - 5.34) + 0.09Z_{12}$

Table D11:BH4995

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	2.42	9.95	0.904	-0.35	$X_{jan} = 9.95 - 0.98(X_{dec} - 11.24) + 2.27Z_1$
February	1.93	8.28	0.437	-0.96	$X_{Feb} = 8.28 - 0.77(X_{jan} - 9.95) + 0.54Z_2$
March	1.75	8.67	-2.12	0.99	$X_{mar} = 8.67 + 0.90(X_{feb} - 8.28) + 0.25Z_3$
April	0.40	8.63	1.085	-0.11	$X_{Apr} = 8.63 - 0.03(X_{mar} - 8.67) + 0.4Z_4$
May	1.39	8.82	-0.27	0.58	$X_{may} = 8.82 + 2(X_{Apr} - 8.63) + 1.13Z_5$
June	1.31	9.42	-2.17	0.99	$X_{jun} = 9.42 + 0.93(X_{may} - 8.82) + 0.18Z_6$
July	1.31	10.12	0.018	0.99	$X_{jul} = 10.12 + 0.99(X_{jun} - 9.42) + 0.18Z_7$
August	1.10	10.65	-0.72	0.94	$X_{Aug} = 10.65 + 0.79(X_{jul} - 10.12) + 0.38Z_8$
September	2.15	12.67	0.21	0.15	$X_{sept} = 12.67 + 0.29(X_{Aug} - 10.65) + 2.13Z_9$
October	1.16	12.07	-0.55	-0.37	$X_{oct} = 12.07 - 0.2(X_{sept} - 12.67) + 1.08Z_{10}$
November	1.14	12.08	0.465	0.86	$X_{nov} = 12.08 + 0.85(X_{oct} - 12.07) + 0.58Z_{11}$
December	0.86	11.24	-1.81	-0.99	$X_{dec} = 11.24 - 0.75(X_{nov} - 12.08) + 0.12Z_{12}$

Table D12:BH4887

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	0.43	2.55	1.563	0.69	$X_{jan} = 2.55 + 0.8(X_{dec} - 2.77) + 0.31Z_1$
February	0.27	2.25	-1.08	-0.91	$X_{Feb} = 2.25 - 0.57(X_{jan} - 2.55) + 0.11Z_2$
March	0.21	2.25	0.86	0.84	$X_{mar} = 2.25 + 0.65(X_{feb} - 2.25) + 0.11Z_3$
April	0.14	2.28	0.388	0.37	$X_{Apr} = 2.28 + 0.25(X_{mar} - 2.25) + 0.13Z_4$
May	0.07	2.30	0.71	0.81	$X_{may} = 2.30 - 0.41(X_{Apr} - 2.28) + 0.04Z_5$
June	0.13	2.42	0.421	0.37	$X_{jun} = 2.42 + 0.69(X_{may} - 2.30) + 0.12Z_6$
July	0.12	2.40	1.454	-0.44	$X_{jul} = 2.40 - 0.41(X_{jun} - 2.42) + 0.11Z_7$
August	0.14	2.43	0.904	0.99	$X_{Aug} = 2.43 + 1.16(X_{jul} - 2.40) + 0.02Z_8$
September	1.29	3.23	0.437	0.14	$X_{sept} = 3.23 + 1.29(X_{Aug} - 2.43) + 1.28Z_9$
October	0.37	2.75	-2.12	0.94	$X_{oct} = 2.75 + 0.27(X_{sept} - 3.23) + 0.13Z_{10}$
November	0.60	2.97	1.085	0.99	$X_{nov} = 2.97 + 1.61(X_{oct} - 2.75) + 0.08Z_{11}$
December	0.37	2.77	-0.27	0.56	$X_{dec} = 2.77 + 0.35(X_{nov} - 2.97) + 0.31Z_{12}$

Table D13:BH4973

Month	S	Mean	Z	Correlation	Thomas-Fiering Model
January	1.36	7.72	-0.098	0.7	$X_{jan} = 7.72 + 0.63(X_{dec} - 8.41) + 0.97Z_1$
February	1.34	6.06	-0.465	-0.1	$X_{Feb} = 6.06 - 0.1(X_{jan} - 7.72) + 1.34Z_2$
March	1.47	6.19	1.563	0.94	$X_{mar} = 6.19 + 1.03(X_{feb} - 6.06) + 0.5Z_3$
April	1.37	5.89	-1.085	0.97	$X_{Apr} = 5.89 + 0.9(X_{mar} - 6.19) + 0.33Z_4$
May	0.68	6.26	0.86	0.67	$X_{may} = 6.26 + 0.33(X_{Apr} - 5.89) + 0.5Z_5$
June	1.03	6.37	0.388	0.9	$X_{jun} = 6.37 + 0.36(X_{may} - 6.26) + 0.45Z_6$
July	1.02	7.09	0.71	0.99	$X_{jul} = 7.09 + 0.98(X_{jun} - 6.37) + 0.14Z_7$
August	1.42	7.35	0.421	0.99	$X_{Aug} = 7.35 + 1.38(X_{jul} - 7.09) + 0.2Z_8$
September	0.93	7.87	1.454	0.89	$X_{sept} = 7.87 + 0.58(X_{Aug} - 7.35) + 0.42Z_9$
October	0.97	8.42	0.904	0.99	$X_{oct} = 8.42 + 1.03(X_{sept} - 7.87) + 0.14Z_{10}$
November	1.04	8.65	0.437	0.02	$X_{nov} = 8.65 + 0.02(X_{oct} - 8.42) + 1.04Z_{11}$
December	1.52	8.41	-2.12	0.72	$X_{dec} = 8.41 + 1.05(X_{nov} - 8.65) + 1.05Z_{12}$

APPENDIX E –Comparison between forecasting results for ARIMA and T-F models

Borehole	Months of Forecasting	Forecasts (m) ARIMA	T-F	Observed (m) GWL	% Error ARIMA	T-F
BH4341	Sep-12	2.13	1.98	2.21	3.6	10.4
	Oct-12	2.09	1.96	2.4	12.9	18.3
	Nov-12	2.07	1.9	1.99	4.02	4.52
BH6423	Sep-12	8.25	3.91	8.39	1.7	53.4
	Oct-12	8.22	4.37	8.36	1.7	47.8
	Nov-12	8.2	5.55	8.42	2.6	34.1
BH4887	Sep-12	2.82	2.47	2.84	0.7	13
	12-Oct	2.82	2.29	2.86	1.4	20
	Nov-12	2.82	2.3	3.03	6.9	24
BH4973	Aug-12	10.07	8.36	9.64	4.5	13.3
	Sep-12	10.18	8.63	10	1.8	13.7
	Oct-12	10.35	9.15	10.01	3.4	8.6
BH4165	Dec-11	3.2	4.14	3.29	3	25.9
	Jan-12	3.13	4.14	3.45	9.3	19.9
	Feb-12	3.09	4.2	3.25	4.6	29.2
BH4164	Oct-12	4.77	5.04	5.46	12.6	7.8
	Nov-12	4.75	5.08	5.63	15.6	9.8
	Dec-12	4.72	4.58	5.54	14.8	17.3
BH4163	Oct-12	3.33	1.9	3.38	1.5	43.7
	Nov-12	3.18	2.39	3.49	8.9	31.7
	Dec-12	3.04	2.42	2.96	2.7	18.3
BHZ6424	Sep-12	24.26	23.45	24.29	0.1	3.5
	Oct-12	24.2	23.67	24.24	0.2	2.4
	Nov-12	24.22	23.7	24.62	1.6	3.7
BH4886	Aug-12	6.3	5.82	6.4	1.6	9
	Sep-12	6.23	6.24	6.45	3.4	3.2
	Oct-12	6.17	96.19	6.7	7.9	7.6
Z4401	Sep-12	6.74	5.12	6.65	1.4	23.1
	Oct-12	6.72	5.17	6.8	1.2	23.9
	Nov-12	6.73	5.55	6.92	2.7	19.7
BH4371	Oct-12	4.62	4.23	4.72	2.1	10.4
	Nov-12	4.58	4.15	4.87	6	14.9
	Dec-12	4.54	3.9	4.81	5.6	19
BH4348	Sep-12	6.02	5.36	6.05	0.5	11.4
	Oct-12	5.99	5.52	6.17	2.9	10.6
	Nov-12	5.95	5.74	6.33	6	9.4
BH4995	Oct -12	15.06	12.37	14.68	2.6	15.8
	Nov-12	15.17	12.26	15.06	0.7	18.6
	Dec-12	15.28	11.16	15.22	0.4	26.7

APPENDIX F

Table F1: Computation of R^2 for fitting Gaussian model (July 2005)

Distance	I	C	SemiVar	SemiCol_Dist	SemiCol_Avg	G_Gam_Sqr	G_Av_Sqr	R^2
0	0.08	0.1	0.01	0	0.006	7.97E-07	0.002	0.6
800	0.125	0.43	0.02	0.009	0.01	0	0.001	0.6
1600	-0.091	0.68	0.04	0.019	0.019	0	0	0.6
2400	0.16	0.92	0.05	0.032	0.031	0	8.44E-07	0.6
3200	-0.054	0.88	0.05	0.047	0.047	1.57E-05	1.59E-06	0.6
4000	0.034	1.61	0.09	0.06	0.061	0.001	0.002	0.6
4800	-0.393	1.73	0.1	0.072	0.071	0.001	0.002	0.6

Table F2: Computation of R^2 for fitting exponential model (July 2005)

Distance	I	C	SemiVar	SemiCol_Dist	SemiCol_Avg	G_Gam_Sqr	G_Av_Sqr	R^2
0	0.08	0.1	0.01	0	0.011	3.47E-05	0.003	0.7
800	0.125	0.43	0.02	0.022	0.023	1.73E-06	0.001	0.7
1600	-0.091	0.68	0.04	0.034	0.034	1.88E-05	0	0.7
2400	0.16	0.92	0.05	0.045	0.044	7.28E-05	1.54E-05	0.7
3200	-0.054	0.88	0.05	0.053	0.053	7.25E-06	3.73E-05	0.7
4000	0.034	1.61	0.09	0.06	0.06	0.001	0.001	0.7
4800	-0.393	1.73	0.1	0.066	0.066	0.001	0.002	0.7
5600	-0.497	1.58	0.09	0.071	0.07	0	0.001	0.7

Table F3: Computation of R^2 for fitting Exponential model (February 2006)

Distance	I	C	SemiVar	SemiCol_Dist	SemiCol_Avg	G_Gam_Sqr	G_Av_Sqr	R^2
0	-0.09	0.26	0.16	0	0.205	0.002	0.05	0.7
800	0.2	0.41	0.26	0.268	0.278	0	0.016	0.7
1600	-0.108	0.7	0.46	0.344	0.343	0.013	0.005	0.7
2400	0.243	0.87	0.49	0.401	0.399	0.009	0.011	0.7
3200	-0.121	0.93	0.56	0.445	0.446	0.014	0.031	0.7

Table F4: Computation of R^2 for fitting Gaussian model (February 2006)

Distance	**I**	**C**	**SemiVar**	**SemiCol_Dist**	**SemiCol_Avg**	**G_Gam_Sqr**	**G_Av_Sqr**	**R^2**
0	-0.09	0.26	0.16	0	0.173	0	0.05	0.4
800	0.2	0.41	0.26	0.199	0.205	0.003	0.016	0.4
1600	-0.108	0.7	0.46	0.274	0.273	0.034	0.005	0.4
2400	0.243	0.87	0.49	0.369	0.365	0.016	0.011	0.4
3200	-0.121	0.93	0.56	0.455	0.456	0.011	0.031	0.4

Table F5: Computation of R^2 for fitting Spherical model (July 2005)

Distance	**I**	**C**	**SemiVar**	**SemiCol_Dist**	**SemiCol_Avg**	**G_Gam_Sqr**	**G_Av_Sqr**	**R^2**
0	0.08	0.1	0.01	0	0.014	7.78E-05	0.002	0.7
800	0.125	0.43	0.02	0.031	0.034	9.48E-05	0.001	0.7
1600	-0.091	0.68	0.04	0.055	0.055	0	0	0.7
2400	0.16	0.92	0.05	0.074	0.073	0	8.44E-07	0.7
3200	-0.054	0.88	0.05	0.087	0.087	0.001	1.59E-06	0.7
4000	0.034	1.61	0.09	0.092	0.092	2.29E-07	0.002	0.7
4800	-0.393	1.73	0.1	0.092	0.092	4.97E-05	0.002	0.7