

Rainfall Forecasts of Yemen Based on Statistical and Probabilistic Approaches

Ammar Hameed S. Al-Jabobi

Submitted to the
Institute of Graduate Studies and Research
in partial fulfillment of the requirements for the degree of

Master of Science
in
Civil Engineering

Eastern Mediterranean University
May 2021
Gazimağusa, North Cyprus

Approval of the Institute of Graduate Studies and Research

Prof. Dr. Ali Hakan Ulusoy
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science in Civil Engineering.

Prof. Dr. Umut Türker
Chair, Department of Civil Engineering

We certify that we have read this thesis and that in our opinion it is fully adequate in scope and quality as a thesis for the degree of Master of Science in Civil Engineering.

Prof. Dr. Mustafa Ergil
Supervisor

Examining Committee

1. Prof. Dr. Mustafa Ergil

2. Prof. Dr. Umut Türker

3. Asst. Prof. Dr. Bertuğ Akıntuğ

ABSTRACT

Since reliable rainfall data are not available for the Republic of Yemen, this study analyzed hypothetically generated monthly rainfall data from January 1981 to December 2018, using POWER, 2019 information. Therefore, for each geomorphological basin, 3 representative hypothetical station locations are proposed with the help of Thiessen polygon approach. After determining the representative annual average rainfall datasets of four basins (Red Sea, Arab Sea, Gulf Aden and Rub 'Al Khali), five independent parametric and non-parametric data quality tests for each basin; Homogeneity, Consistency, Normality, Trend and Stationarity Unit-root are applied. To predict rainfall data for the next 3 years, 27 different ARIMA models were proposed for each basin and tested through Akaike Information Criteria (AIC). Among them, the best 3 representative models for each basin were selected and determined by suggested weighted average (1, 2 and 3). For this reason, the annual average dataset of each basin from 1981 to 2008 (28) was used to train these models, and the remaining annual average dataset from 2009 to 2018 (10) was used to test these trained datasets. Then, among the three selected models of each basin, the most suitable model was selected and used to predict the annual rainfall data for 3 consecutive years (2019, 2020 and 2021). These models are ARIMA (0,1,1) for the Red Sea Basin, ARIMA (0,1,1) for the Arab Sea Basin, ARIMA (2,1,2) for the Gulf Aden Basin, and ARIMA (0,1,2) for the Rub 'Al Khali Basin.

As a part of this study, three different frequency distributions among the commonly used ones; Normal, log-normal and Pearson Type III (Gamma) were selected and the

most representative frequency distribution function for each basin was determined by selecting that distribution having the closest p value to 1.00.

Also, based on Moving Averages with different time windows (2 to 9), the annual rainfall trend of each basin was determined. It has been determined that all basins showed a decreasing trend in the range of 1 - 2 mm/year.

Similarly, annual average rainfall data sets for each basin were systematically analyzed for wetness or dryness periods, and all basins were interpreted to be under the influence of the dry spell.

Keywords: rainfall, time-series models, trends, wet or dry spells, Yemen.

ÖZ

Yemen Cumhuriyeti için güvenilir yağış verileri bulunmadığından, bu çalışma, POWER, 2019 bilgilerini kullanarak, Ocak 1981'den Aralık 2018'e kadar varsayımsal olarak oluşturulan aylık yağış verilerini analiz etmiştir. Bu nedenle, her jeomorfolojik havza için, Theissen poligon yaklaşımı yardımıyla 3 temsili varsayımsal istasyon konumu önerilmiştir. Dört havzanın (Red Sea, Arab Sea, Gulf Aden ve Rub 'Al Khali) temsili yıllık ortalama yağış veri setlerini belirledikten sonra, her havza için, beş bağımsız parametrik ve parametrik olmayan veri kalitesi testi; Homojenlik, Tutarlılık, Normallik, Trend ve Durağanlık Birim-kökü uygulanmıştır. Sonraki 3 yıllık yağış verilerini tahmin etmek için, her bir havza için 27 farklı ARIMA modeli önerilmiş ve Akaike Bilgi Kriterleri (AIC) aracılığıyla test edilmiştir. Bunların arasından da, her havza için en iyi 3 temsili model, önerilen ağırlıklı ortalama (1, 2 ve 3) kullanılarak belirlenmiştir. Bu nedenle, bu modelleri eğitmek için, her havzanın 1981'den 2008'e (28) yıllık ortalama veri setleri, geriye kalan 2009'dan 2018'e (10) yıllık ortalama veri setleri de, bu eğitilen veri setlerini test etmek için kullanılmıştır. Ardından da, her havzanın seçilen bu üç modeli arasından, en uygun olan model belirlenip, birbirini izleyen 3 yılın yıllık yağış verileri (2019, 2020 ve 2021) tahmin edilmiştir. Bu modeller, Red Sea Havzası için ARIMA (0,1,1), Arab Sea Havzası için ARIMA (0,1,1), Gulf Aden Havzası için ARIMA (2,1,2) ve Rub' Al Khali Havzası için de ARIMA (0,1,2) dır.

Bu çalışmanın bir parçası olarak, yaygın olarak kullanılan sıklık dağılım fonksiyonlarından üçü; Normal, log-normal ve Pearson Tip III (Gamma) seçilmiş ve

her havzayı en iyi temsil eden sıklık dağılım fonksiyonu için, p değeri 1.00 en yakın olan dağılım fonksiyonu seçilip belirlenmiştir.

Ayrıca, farklı zaman aralıklarına (2 ila 9) sahip Hareketli Ortalamalara dayalı olarak, her havzanın yıllık yağış eğilimi belirlenmiştir. Tüm havzaların 1 - 2 mm/yıl aralığında azalma eğilimi gösterdiği tespit edilmiştir.

Benzer şekilde, her havza için yıllık ortalama yağış veri setleri, ıslaklık veya kuruluk dönemleri için sistematik olarak analiz edilmiş ve tüm havzaların kuruluk dönemi etkisi altında olduğu yorumlanmıştır.

Anahtar Kelimeler: yağış, zaman serisi modelleri, eğilimler, ıslaklık veya kuruluk dönemleri, Yemen.

ACKNOWLEDGEMENT

I would first like to thank my thesis supervisor Prof. Dr. Mustafa Ergil. The door of Prof. Ergil office was always open whenever I ran into a trouble spot or had a question about my thesis. He consistently allowed this research to be my own work but steered me in the right direction whenever he thought I needed it. Special thanks go to him for his countless hours of reflecting, reading, encouraging, and all patience throughout the entire process.

Thanks go to Civil Engineering Staff of Eastern Mediterranean University in North Cyprus, for their academic and scientific support throughout my study of M.Sc.

Special thanks, I must express my very profound gratitude to WAMY and to my friend Mahfood Zain, for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of research and even while writing this thesis. This accomplishment would not have been possible without them. Thank you all.

Lastly, I would like to thank my family for the continuous support they have given me throughout my studies.

TABLE OF CONTENT

ABSTRACT	iii
ÖZ	v
ACKNOWLEDGEMENT	vii
LIST OF TABLES.....	xiv
LIST OF FIGURES	xvii
1 INTRODUCTION	1
1.1 General	1
1.2 Thesis Overview	2
1.3 Objectives of This Study	2
2 LITERATURE REVIEW.....	3
3 STUDY AREA	9
3.1 Yemen.....	9
3.2 Climate of Yemen	9
3.3 Population of Yemen.....	10
3.4 Hydrology of Yemen.....	10
3.4.1 Infiltration and Runoff Areas	10
3.4.2 Main Watershed Areas	11
3.4.2.1 The Red Sea Basin.....	11
3.4.2.2 The Arab Sea Basin	11
3.4.2.3 The Gulf Aden Basin	11
3.4.2.4 The Rub' Al Khali Basin	12
3.5 Water Resources of Yemen	12
3.5.1 Conventional Water Resources	12

3.5.1.1 Rainfall.....	12
3.5.1.2 Surface Water	14
3.5.1.3 Groundwater.....	15
3.5.2 Unconventional Water Resources	16
3.5.2.1 Treated Wastewater	16
3.5.2.2 Seawater Desalination.....	16
3.6 Water Scarcity in Yemen.....	17
3.7 Rainfall Stations.....	17
3.8 Rainfall Data Details	19
4 THEORY AND METHODOLOGY.....	21
4.1 Introduction.....	21
4.2 Definitions	22
4.3 Time-Series.....	23
4.4 Time-Series Forecasting.....	24
4.5 The ARIMA Model.....	25
4.5.1 Comparison Between ARMA and ARIMA.....	28
4.6 How to Built an ARIMA Model	29
4.6.1 Unit Roots Concept.....	31
4.6.2 AutoCorrelation and Partial AutoCorrelation Functions.....	31
4.6.2.1 ACF and PACF of AR (p).....	31
4.6.2.2 ACF and PACF of MA (q).....	33
4.6.3 Alternative ARIMA Models	34
4.6.3.1 ARIMA(0,0,0).....	34
4.6.3.2 ARIMA(0,1,0).....	36
4.6.3.3 ARIMA(2,1,0).....	37

4.6.3.4 ARIMA(0,2,0)	39
4.6.3.5 ARIMA(0,2,1)	41
4.6.3.6 Comparison of ARIMA(2,1,0) and ARIMA(0,2,1).....	41
4.7 Akaike Information Criterion	43
4.8 Forecasting Criteria	45
5 APPLIED TESTS, FUNCTIONS, AND MODELS	49
5.1 Quality Tests	49
5.1.1 Adequacy of Sample Size	49
5.1.2 Homogeneity Test	49
5.1.3 Consistency Test.....	50
5.1.4 Normality Test	50
5.1.5 Trends Test.....	50
5.1.6 Stationarity Test	51
5.2 Models	51
5.2.1 Frequency Distribution Functions	51
5.2.1.1 Normal Frequency Distribution Equation.....	51
5.2.1.2 Log-Normal Frequency Distribution Equation	52
5.2.1.3 Pearson Type III (Gamma) Frequency Distribution Equation	52
5.2.2 Time-Series Models.....	52
5.3 Detecting a Trend.....	52
5.4 Detecting Wetness/Dryness	53
6 CALCULATIONS AND RESULTS	54
6.1 Introduction.....	54
6.2 Red Sea Basin	55
6.2.1 Quality Tests Results of Rainfall Datasets of Red Sea Basin	55

6.2.2 Generated Equations from the widely Used Frequency Distribution Functions of Red Sea Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018	56
6.2.3 Details of the Suggested ARIMA Models and Their AIC Scores for Red Sea Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018	56
6.2.4 Building a Forecast Model by ARIMA for Red Sea Basin	58
6.2.5 Detecting Appropriate MA Time Window with its Linear Trend Equation of Red Sea Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018.....	58
6.2.6 Detecting Wetness/Dryness of Red Sea Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018	59
6.3 Arab Sea Basin.....	60
6.3.1 Quality Tests Results of Rainfall Datasets of Arab Sea Basin	60
6.3.2 Generated Equations from the Widely used Frequency Distribution Functions of Arab Sea Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018.....	61
6.3.3 Details of the Suggested ARIMA Models and Their AIC Scores for Arab Sea Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018.....	61
6.3.4 Building a Forecast Model by ARIMA for Arab Sea Basin.....	62
6.3.5 Detecting Appropriate MA Time Window with its Linear Trend Equation of Arab Sea Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018.....	62
6.3.6 Detecting Wetness/Dryness of Arab Sea Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018	63
6.4 Gulf Aden Basin.....	64

6.4.1 Quality Tests Results of Rainfall Datasets of Gulf Aden Basin	64
6.4.2 Generated Equations from the Widely used Frequency Distribution Functions of Gulf Aden Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018.....	65
6.4.3 Details of the Suggested ARIMA Models and their AIC Scores for Gulf Aden Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018	65
6.4.4 Building a Forecast Model by ARIMA for Gulf Aden Basin.....	67
6.4.5 Detecting Appropriate MA Time Window with its Linear Trend Equation of Gulf Aden Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018.....	67
6.4.6 Detecting Wetness/Dryness of Gulf Aden Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018	68
6.5 Rub' Al Khali Basin	69
6.5.1 Quality Tests Results of Rainfall Datasets of Rub' Al Khali Basin	69
6.5.2 Generated Equations from the Widely used Frequency Distribution Functions of Rub' Al Khali Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 fo 2018.....	70
6.5.3 Details of the Suggested ARIMA Models and their AIC Scores for Rub' Al Khali Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018	70
6.5.4 Building a Forecast Model by ARIMA for Rub' Al Khali Basin.....	72
.....	72

6.5.5 Detecting Appropriate MA Time Window with its Linear Trend Equation of Rub' Al Khali Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018.....	72
6.5.6 Detecting Wetness/Dryness of Rub' Al Khali Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018	73
7 CONCLUSION AND RECOMMENDATIONS	74
7.1 Conclusion	74
7.2 Recommendation	77
REFERENCES	78
APPENDICES.....	83
Appendix1: Rainfall Data Details and Simple Statistical Measures of the Basins.	84
Appendix 2: Standard Normal Distribution Table.....	107
Appendix 3: χ^2 Distribution Table	109
Appendix 4: t-test Probability Values.....	110
Appendix 5: F-test Probability Values.....	112
Appendix 6 : Pearson Type III Distribution Table	115

LIST OF TABLES

Table 3. 1: The GPS coordinates of the selected representative rainfall stations for each basin of Yemen.....	19
Table 6. 1: Quality tests results of rainfall datasets of Red Sea basin	55
Table 6. 2: Generated equations from the widely used frequency distribution functions of red sea basin, based on the yearly averaged rainfall dataset from 1981 to 2018 ...	56
Table 6. 3: Details of the suggested ARIMA Models and their AIC scores for Red Sea Basin, based on the yearly averaged rainfall dataset from 1981 to 2018.....	56
Table 6. 4: Building a forecast model by ARIMA for Red Sea Basin	58
Table 6. 5: Detecting appropriate MA time window with its linear trend equation of Red Sea Basin based on the yearly averaged rainfall dataset from 1981 to 2018.....	58
Table 6. 6: Detecting wetness/dryness of Red Sea Basin based on the yearly averaged rainfall dataset from 1981 to 2018.....	59
Table 6. 7: Quality tests results of rainfall datasets of Arab Sea Basin	60
Table 6. 8: Generated equations from the widely used frequency distribution functions of Arab Sea Basin, based on the yearly averaged rainfall dataset from 1981 to 2018	61
Table 6. 9: Details of the suggested ARIMA models and their AIC scores for Arab Sea Basin, based on the yearly averaged rainfall dataset from 1981 to 2018.....	61
Table 6. 10: Building a forecast model by ARIMA for Arab Sea Basin	62
Table 6. 11: Detecting appropriate MA time window with its linear trend equation of Arab Sea Basin based on the yearly averaged rainfall dataset from 1981 to 2018	62
Table 6. 12: Detecting Wetness/Dryness of Arab Sea Basin based on the yearly averaged rainfall dataset from 1981 to 2018.....	63

Table 6. 13: Quality tests of rainfall datasets of Gulf Aden Basin	64
Table 6. 14: Generated equations from the widely used frequency distribution functions of Gulf Aden Basin, based on the yearly averaged rainfall dataset from 1981 to 2018.....	65
Table 6. 15: Details of the suggested ARIMA models and their AIC scores for Gulf Aden Basin, based on the yearly averaged rainfall dataset from 1981 to 2018	65
Table 6. 16: Building a forecast model by ARIMA for Gulf Aden Basin	67
Table 6. 17: Detecting appropriate MA time window with its linear trend equation of Gulf Aden Basin based on the yearly averaged rainfall dataset from 1981 to 2018 ..	67
Table 6. 18: Detecting Wetness/Dryness of Gulf Aden Basin based on the yearly averaged rainfall dataset from 1981 to 2018.	68
Table 6. 19: Quality tests of rainfall datasets of Rub' Al Khali Basin.....	69
Table 6. 20: Generated equations from the widely used frequency distribution functions of Rub' Al Khali Basin, based on the yearly averaged rainfall dataset from 1981 to 2018.....	70
Table 6. 21: Details of the suggested ARIMA models and their AIC scores for Rub' Al Khali Basin, based on the yearly averaged rainfall dataset from 1981 to 2018	70
Table 6. 22: Building a forecast model by ARIMA for Ruba' Al Khali Basin.....	72
Table 6. 23: Detecting appropriate MA time window with its linear trend equation of Rub' Al Khali Basin based on the yearly averaged rainfall dataset from 1981 to 2018	72
Table 6. 24: Detecting Wetness/Dryness of Rub' Al Khali Basin based on the yearly averaged rainfall dataset from 1981 to 2018.....	73
Table 7. 1: Forecasted rainfall values of 2019, 2020, and 2021 of each meteorological district based on the relevant most representative time series models	75

Table 7. 2: The Synopsis of the yearly averaged rainfall data test results with the suggested frequency distribution and time-series models for each meteorological district.....	76
---	----

LIST OF FIGURES

Figure 3. 1: Annual average rainfall of Yemen based on 1985 to 1991 (Gun, et al., 1996)	13
Figure 3. 2: Map of Yemen with four main basins and their main surface water systems (Gun, et al., 1996)	14
Figure 3. 3: Map of Yemen showing the main groundwater basins (Gun et al., 1996)	15
Figure 3. 4: Thiessen polygons based 12 representative rainfall stations of Yemen..	18
Figure 4. 1: (a) ACF of an AR (1) process $\beta_1 = 0.70$	32
Figure 4. 2: (a) ACF and (b) PACF of an AR (2) model with $\beta_1 = 0.3$ and $\beta_2 = 0.5$..	32
Figure 4. 3: (a) ACF and (b) PACF of an AR (2) model with $\beta_1 = 1.2$ and $\beta_2 = - 0.64$	33
Figure 4. 4: (a) ACF and (b) PACF of a MA(1) model with positive parameter θ	34
Figure 4. 5: (a) ACF and (b) PACF of a MA(1) model with negative parameter θ ...	34
Figure 4. 6: ACF of ARIMA(0,0,0).....	35
Figure 4. 7: PACF of ARIMA(0,0,0).....	35
Figure 4. 8: Residual ACF of ARIMA(0,1,0)	37
Figure 4. 9: Residual PACF of ARIMA(0,1,0)	37
Figure 4. 10: Residual ACF of ARIMA(2,1,0)	38
Figure 4. 11: Residual PACF of ARIMA(2,1,0)	38
Figure 4. 12: Residual plot of ARIMA(2,1,0).....	39
Figure 4. 13: Plot of ARIMA(2,1,0).....	39
Figure 4. 14: Residual ACF of ARIMA(0,2,0)	40
Figure 4. 15: Residual PACF of ARIMA(0,2,0)	40

Figure 4. 16: Plot of ARIMA(0,2,1)	43
Figure 4. 17: Flowchart illustrating how to build ARIMA model and forecast the unknown successive data	48

Chapter 1

INTRODUCTION

1.1 General

The hydro-climatological parameters are defined as the climatological factors that affect hydrology of the countries, among these parameters rainfall has the highest importance. Its amount should be measured accurately and has to be studied precisely. Although rainfall has a high positive effect on ecological sustainability of the living organisms, but can cause disasters like flooding or drying up of the existing reservoirs due to global warming. Predicting the rainfall of the following years, estimation of maximum and minimum rainfall amounts, spatial and temporal distribution of rainfall are the issues that always interest the engineers, planners, economists and other researchers. The amount of available water for estimating the water budget and scheduling demand pattern depending on precipitation and evaporation, the consumptive use of different crop patterns, the droughtiness are enforcing the researchers to carry out detailed studies of these hydro-climatological parameters. There is, therefore a need for more accurate climate model predictions that will provide meteorological information on national level and enable relevant climate change impact studies to assist adaptation strategies.

Weather forecasting plays an important role in our daily life. Especially in engineering, it shows itself more significantly. Hence, estimating the daily, monthly, seasonally and

even the yearly amount of rainfall values for different locations may guide the researchers to some extent, for their future strategies.

1.2 Thesis Overview

In chapter 2, some relevant studies in literature was detailed. Chapter 3, elaborates the study area, the hypothetical rainfall stations and the gathered data details. Chapter 4, details the applied methodology of the used methods in this study, whereas in Chapter 5 the results were presented and finally the conclusion is given in Chapter 6 based on the results and the findings of this study.

1.3 Objectives of This Study

1. By applying the Thiessen polygon, the hypothetical representative rainfall stations locations for each basin of Yemen will be determined.
2. Based on annual rainfall values, appropriate frequency equations will be determined for each basin of Yemen.
3. Based on proper Autoregressive Integrated Moving Average (ARIMA) model, the coming 3 years annual average rainfall values will be forecasted for each basin of Yemen.
4. Based on proper Moving Average (MA) with relevant time window, the yearly averaged rainfall dataset trend will be determined for each basin of Yemen.
5. Based on long-years monthly values wet/dry periods patterns will be determined for each basin of Yemen.

Chapter 2

LITERATURE REVIEW

From the relevant literature, it has been noted that, there isn't any study that has been conducted on rainfall trends and forecasts of Yemen as a whole. Majed, et al., (2019) studied Precipitation Analysis and Water Resources of Wadi Siham Basin (WSB), Yemen. In their study, they gathered the rainfall from 1979 to 2008 and statistically analyzed to assess the patterns of precipitation. They used Mann-Kendall and Sen slope analyses and concluded that, the annual precipitation at the Wallan and Al-Amir stations had substantial negative values (-4.72 mm/year and -6.11 mm/year respectively), while the Dhamar rainfall pattern was positive with 50.20 mm/year. The mean annual runoff was estimated to be 82.92 Mm³ or 23.94 per cent of the total annual rainfall in the WSB. The estimated runoff due precipitation was 4.85 per cent of the total rainfall, implying that the total deficit was 95.15 per cent.

Whereas Al-Falahi, et al., (2020) studied first time, to measure and evaluate the accuracy of several regular precipitation devices against the measurements available from the highland area of Yemen. Al Mahwit governorate, using the most commonly known methodological approaches at various time frames, attempted to analyze the precipitation in order to resolve the data constraint and to identify the most reliable grid interval for hydrological, regional and local climate modellings. The Statistical Downscaling Model (SDSM), was applied to estimate the potential impacts of the

climate change on water supply and the implementation of adaptation measures for Yemen (Wilby, 2002).

Farquharson, et al., (1996) using a water balance analysis of the mountainous regions in arid and semi-arid areas of Yemen with a diverse variety of annual average rainfall details, study on the methodology of how to construct mathematical models for regular rainfalls and for rainfall-runoff relationships. Their suggested models were not only suitable for those circumstances but even for different geomorphologic cases. Gun et al., (1996) provided a clear overview of Yemen's water-resource conditions, consistent maps, charts, tables and comprehensive data about the water resources in Yemen.

Nyatume, et al., (2018) applied AutoRegressive Integrated Moving Average (ARIMA) models to analyze and then forecast, the annual rainfall and maximum temperature over Tordzie watershed in the Volta Region of Ghana. Autocorrelation function and partial autocorrelation function were used to identify the models by visual inspection. The selected models were evaluated and validated using the Akaike Information Criterion (AIC). For the diagnostic analyses of the models, they checked for independence, normality, homoscedascity, p-p and q-q plots of the residuals. The best ARIMA model for rainfall of Kpetoe and Tordzinu districts of the Volta Region of Ghana were obtained by comparing the smallest AIC values, 190.07 and 178.23 respectively. The models efficiency was checked using the sum of the square error (SSE), the mean square error (MSE), the mean absolute percent error (MAPE) and the root mean square error (RMSE) measures respectively. The results of these analyses, they concluded that, the determined models were adequate and can guide the future water planning projections.

Shbary, et al., (2015) applied Box-Jenkins method for predicting long-term rainfalls of Sylhet station for Dhakka division of Bangladesh by establishing seasonal ARIMA models in their analyses, where the monthly rainfall data from 1980 to 2010 were used for training and validating the model. The predictive accuracy was verified by using the rainfall data from 1980 to 2006 for training and the data from 2007 to 2010 for validation. The projected monthly rainfall values were tested with real time-series as well as a second level validation.

Daniel, et al., (2015) similarly used ARIMA Modeling for forecasting of rainfall in Warri Town, Nigeria. The main objective of this study was to find a Seasonal ARIMA model that can accurately predict rainfall in Warri town. They collected available data on monthly average rainfall for Warri town from the National Meteorology Center in Oshodi, Nigeria, where the data of years 2003-2012 period used for modeling data and 2013 for prediction and validation of the data set. For at least one year, the identified Seasonal ARIMA (0,1,1) has proven to be satisfactory in forecasting that rainfall.

Hayek, et al., (2016) carried out a comparative review of monthly stream flow volume reaching to Al-Aroos River in the Syrian coast, so as to forecast the oncoming volumes. They adopted Box-Jenkins model to test the time-series data because of its high precision. In their study, they used monthly water volumes of for 15 years and performed the necessary tests on model residues. Hence, they commented on the best model that describes the data, was Seasonal ARIMA (1,2,1). In fact, they used the first 14 years data to establish the model and the remaining one year to validate it, based on the smallest weighted mean of the RMSE, MAP, MAE parameters. The best-predicted model is was ARIMA (1,1,0). Yuchuan et al., (2019), applied Autoregressive Integrated Moving Average (ARIMA) model to forecast near-term regional

temperature and precipitation, and they conclude that the ARIMA-based forecasting model is a quick, easy-to-understand, and dependable method for predicting regional temperature and precipitation for the next 2–20 years, which can be used in a variety of engineering applications. The ARIMA-based statistical time-series forecasting model, in combination with techniques for estimating confidence intervals for return periods and simulating future daily temperature and precipitation, offers a new way to get near-term regional precipitation and temperature data that is crucial for civil and environmental engineering applications.

Shiban, et al., (2019) worked on the Al-Hwaiz basin that is located eastern coast of the Mediterranean Sea at the northwestern corner of Syria. The rainfall data was gathered from three rain stations that cover the entire basin from 1959 to 2017. They applied ARIMA models and determined that ARIMA (1,1,3) model is a good representative of the data, and the ARIMA (2,1,0) model was the right model to forecast the future rainfall which was decreasing 6.13 mm per year during their study period.

Surajit, et al., (2010) worked on a multivariate regression model to predict summer monsoon (June–August) rainfall in India, based on data relating to the period 1871–1999. The trends and stability of the time-series have been examined through randomness and non-stationarity characters of these time-series. The suggested model was ARIMA (0,1,1). As a result, an autoregressive neural network (ARNN) model was applied. The neural network was having multilayer perceptron with detailed variable selection procedures. The non-linearity sigmoid was used when training the network. Finally, a three-three-one architecture of the ARNN model was obtained and, after a detailed statistical study, the superiority of ARNN was confirmed over ARIMA (0,1,1).

Moges, et al., (2020) studied on the stream flow predictions for the management of water supplies. Although there are several methods for predicting stream flows, they applied the process-based model (Soil and Water Assessment Tool-Variable Source Area Model-SWAT-VSA), the stochastic model (Artificial Neural Network-ANN), the Auto-Regressive Moving-Average (ARMA) model, and the Bayesian ensemble model. By using SWAT-VSA, ANN and ARMA, the stream flow data is projected from 1 to 8 days, according to quantitative precipitation prediction by the US National Weather Service. They determined that, SWAT-VSA and ANN models have improved forecasting of the total flow (Nash Sutcliffe Efficiency coefficient (NSE) (0.60 – 0.70) and the peak flow, but under-predicted low flows. During the forecast phase, the ANN had the highest predictive capacity (NSE 0.44 – 0.64), but all three models were under-predicted the peak flows. For the forecast period of 01/01/2017 to 02/28/2018, the fitted ARIMA time-series model with the QPF forecast data as covariates was used to forecast stream flow for 1- to 8-days lead times. Eight time-series of forecast stream flow data were extracted and compared to observed stream flow data, one for each forecast day.

Jan, et al., (2012) studied the daily forecasting of water demand that is an essential component of cost-effective and efficient maintenance and optimization of the urban water supply systems. In their study framework, they focused on the pairing of discrete wavelet transformations (WA) and artificial neural networks (ANNs) for urban water demand forecasting applications. They applied, multiple linear regression (MLR), multiple non-linear regression (MNLR), ARIMA, ANN and WA-ANN models for urban water demand forecasts of one-day lead times intervals of the summer months (May to August). Their relative performances were compared using the coefficient of

determination, the root mean square error, and the index efficiency. The main variables used to build and test the models were daily average precipitation, daily high temperatures and daily water demand data from 2001 to 2009 in Montreal, Canada. The MLR, MNLR, ARIMA, and ANN models have been found to deliver more precise urban demand predictions. The NumXL software program was used in their study to create ARIMA models for urban water demand forecasting. The autocorrelation function (ACF) was used to determine the stationarity of the input data series. The City of Montreal's urban water demand data series were detected to be non-stationary. The ARIMA models were trained using data from the training set (May 2001 to May 2008), then tested using data from the testing set (May 2008 to August 2009), and their goodness of fit was compared using the proper statistical measures. The model with the best forecasting performance among the numerous ARIMA models gave ARIMA (2,1,3) model to be the best fit.

Chapter 3

STUDY AREA

3.1 Yemen

Yemen is located in the southwestern part of the Arabian Peninsula between the latitudes 12.40° to 19.00° North, and between the longitudes 42.30° to 53.05° East. The total area of Yemen is about 527,970 km² with the exception of the Rub' Al Khali. In addition, having a coastline of 2000 km long that is overlooking to the Red Sea, the outskirts of the Gulf of Aden and the Arabian Sea. The elevation of the highest peak is about 3760 m above the sea level, which is the summit of Prophet Shoaib Mountain, in fact, it is the highest summit of the Arabian Peninsula. Kingdom of Saudi Arabia is located at its north, while the Gulf of Aden and the Arabian Sea are located at its south. The Sultanate of Oman is situated at the east, while the Red Sea is at its west. The Bab Al-Mandeb strait, that controls the crossing within the strait, is at the southwest, and it divides the Mayon Island into two parts. The Socotra Island, is the largest Yemeni island located at the Indian Ocean. Its area is estimated to be 3650 km² and is about 510 km away from the coast of the Gulf of Aden. In addition, there are more than 112 Yemeni islands scattered within the Arabian Sea and the Red Sea.

3.2 Climate of Yemen

The climate of Yemen is hot and humid in the coastal areas, moderate in the mountainous areas, and dry and hot in the eastern regions of the country:

- West Coast: Temperature reaches up to 54° C in summers and 35° C in winters.

- Southern Coast: Temperature reaches up to 37° C in summers and 25° C in winters. The amount of rain does not exceed 100 mm/year and does not exceed 10 days in a year.
- Highlands: It has a mild climate in most of the days of the year, it is hotter during the day time and colder at night, especially from October to February. The temperature reaches 5° C from November till January and it reaches up to 25° C in July.
- Average annual rainfall along the coastline and the desert is less than 50 mm, but its range is between 200 and 400 mm at hilly slopes and even more than 1000 mm over the western slopes of the mountains.

3.3 Population of Yemen

Yemen, based on 2018 years estimate, has a population of about 28 million, of which 46% under age 15 and 2.7% over age 65. The population was 4.3 million in 1950 and forecasted to be approximately 60 million in 2050 (World Prospects for Population, 2018).

3.4 Hydrology of Yemen

3.4.1 Infiltration and Runoff Areas

Van der Gun, 1996 identified the areas of infiltration and the areas due penetration of rainwater for Yemen so as to estimate the volume of escaped water to the sea. These specific areas are summarized as follows:

1. the areas where there is a little rainfall with limited terrain, where the environmental activity is relatively low, and tending to infiltrate rainwater. For example, the plains of Sana'a are naturally allowing the escape of water, but in fact, the process of escaping rainwater does not occur except after very intensive rainfall.

2. those areas where it rains a lot and is subjected to major changes along its terrain due to environmental activities tend to create conditions that help for escaping the water.

Despite this, the areas that create conditions for more water escape are found in the fast-flowing valleys, in which the water runs during the periods of very heavy rains, and even in these cases, the shallow flow enters and penetrates to the depths of the valley by feeding again (in principle) the groundwater system.

3.4.2 Main Watershed Areas

In Yemen, there are 78 huge watersheds, where most of the precipitation is in the form of rainfall. These watersheds can be grouped into 4 main drainage basins: the Red Sea Basin, the Arab Sea Basin, the Gulf Aden Basin and the Rub' Al Khali Basin.

3.4.2.1 The Red Sea Basin

There are three main valleys in the Red Sea basin: Wadi Sardoud, Wadi Siham, and Wadi More.

3.4.2.2 The Arab Sea Basin

The Arab Sea basin is complicated, as it includes the low valleys of Ghaydah (Wadi Haqat, Wadi Tanhalin, Wadi Giza, Wadi Fawra, Wadi Idna, Wadi Hadramout al-Kabir, and Wadi al-Jawf). Topographic conditions theoretically suggest that, the rainfall in the plains of these mountainous highlands near Sanaa, in fact drain water to the Arabian Sea, mainly by Wadi Masila.

3.4.2.3 The Gulf Aden Basin

There are seven main watersheds Hwaira, Hajar, Mayfa'a, Ahwar, Hassan, Bana and Thawban, in the Gulf Aden basin and their areas exceed 1000 km², occupying the land from the east that is ending at the west. These valleys drain the rainfall that are coming

from the slopes of the southern mountainous highlands having the similar characteristics of those valleys that drain the rainfall to the Red Sea. These catchments in fact, receive the highest rainfall within the country, and as the distance between them and the sea is short and even the slopes are so severe towards the coastal plains, facilitating the process of drainage of the rainwater quicker than the Red Sea catchments.

3.4.2.4 The Rub' Al Khali Basin

There are many bare valleys in the northern facing slopes of the mountainous highlands and plains at the eastern region. Among them, the most important ones are:

- Along Mountainous Highlands (from west to east), Wadi Najran, Atfan, Khub, Umrah and Ghamour.
- Along Eastern Plains (from west to east), Wadi Hadi, Iwat Asir, Mukhiar Al Khadra, Haradah, Qanab, Alwat, Harthouth, Ramah, Dehih Buwat, Arian, Rakhoot, Mathan, and Wadi Sheehan.

The flow of water in these sub-basins is scarce and in fact, the rainfall is assumed to be rapidly discharged to the sandy interior areas where the groundwater is believed to be fed. This is not certain and perhaps even this water may infiltrate to the Persian Gulf or may be evaporated in the depressing and rising sunrise which needs further scientific study.

3.5 Water Resources of Yemen

3.5.1 Conventional Water Resources

3.5.1.1 Rainfall

The total annual rainfall on Yemen is estimated to be about 1,500 Mm³/year (Yehya et al., 2005), and the average annual rainfall of Yemen between the years 1985-1991

ranges from 50 mm in the Rub' Al Khali Basin to more than 1000 mm at the mountains of the Red Sea Basin and the Arabian Sea Basin by Gun et al. (1996) as given in Fig.3.1.

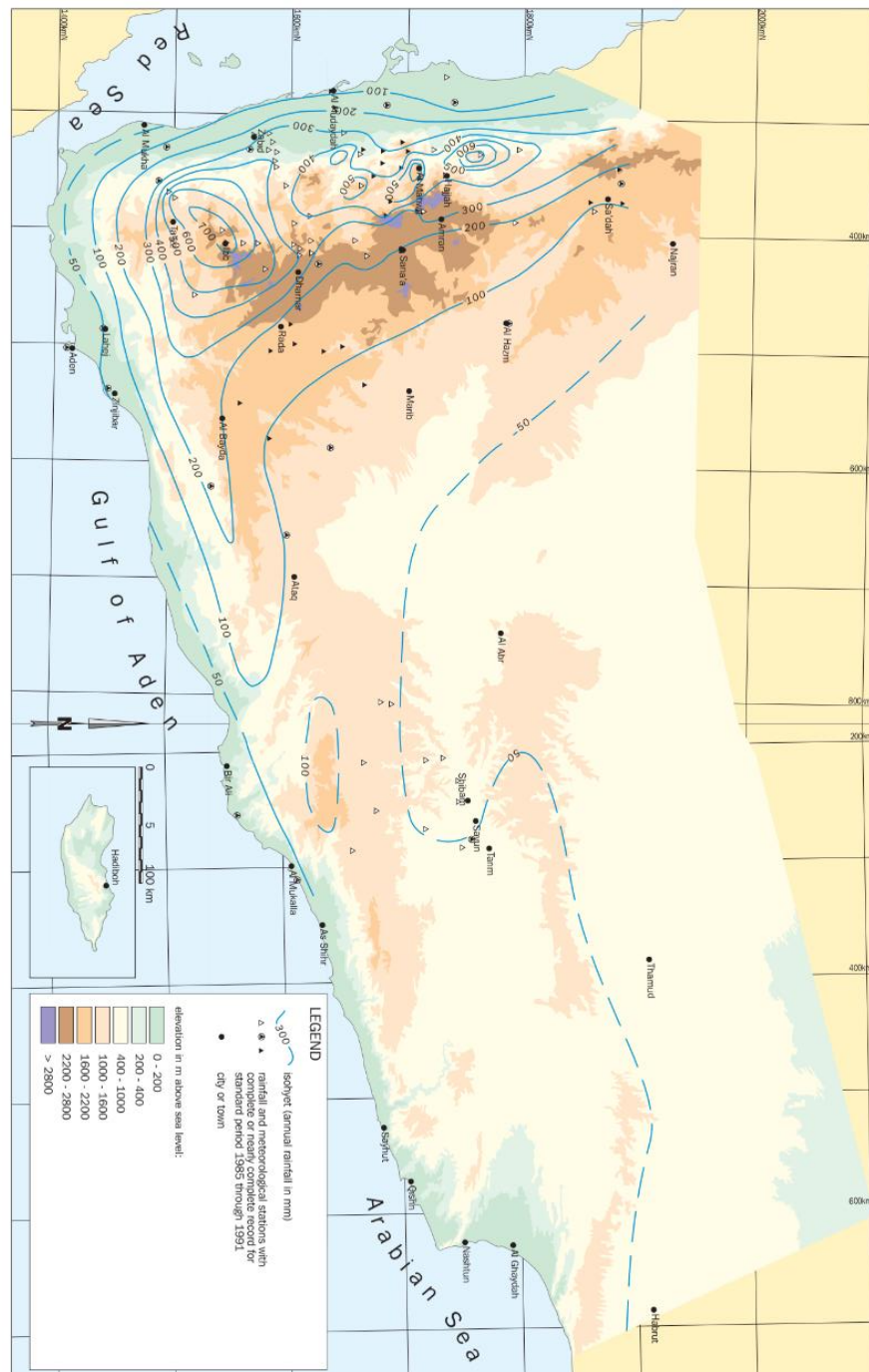


Figure 3. 1: Annual average rainfall of Yemen based on 1985 to 1991 (Gun, et al., 1996)

3.5.1.2 Surface Water

Yemen is divided into four main basins, each of which consists of a group of valleys and sub-surface water streams. Fig. 3.2 details these basins and the main water systems (Gun, et al., 1996).

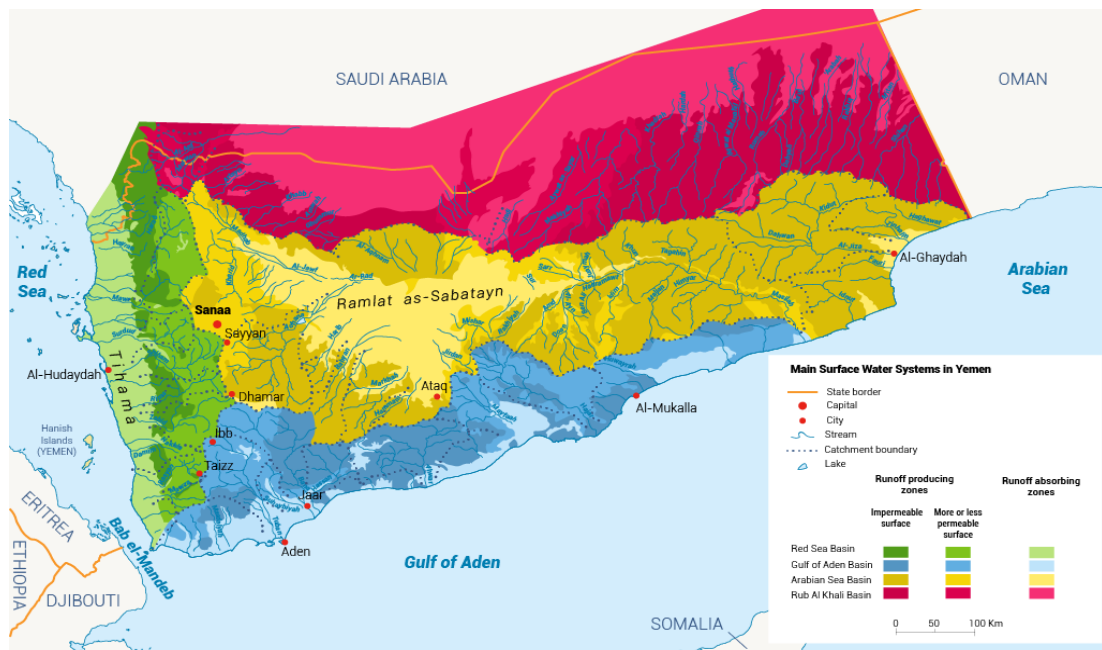


Figure 3. 2: Map of Yemen with four main basins and their main surface water systems (Gun, et al., 1996)

Their amounts are estimated nearly to be 2.5 billion m³/year and their basin wise distribution is:

1. The Red Sea Basin covering an area of 33,000 km², with an average annual rainfall of 135 mm and a total annual flow of 741 million m³.
2. The Gulf Aden Basin covering an area of 46,680 km², with an average annual rainfall of 51 mm and a total annual flow of 535 million m³.
3. The Rub' Al Khali Basin covering an area of 90,900 km², with an average annual rainfall of 28 mm and a total annual flow of 67 million m³.
4. The Arab Sea Basin is composed of three main sub-basins:

- i. Ramlat as Sabatayn: having an area of about 45,000 km², of the average annual rainfall of 30 mm with an annual water volume of 40 million m³.
- ii. Wadi Hadhramaut/Al-Messila branch: having an area of about 46075 km², of the average annual rainfall of 57 mm with an annual water volume of 18 million m³.
- iii. Al-Ghaydah sub-basin: having an area of about 115,375 km², of the average annual rainfall of 58 mm with an annual water volume of 77 million m³.

3.5.1.3 Groundwater

The total stored water is estimated to be nearly 10,370 billion m³, of which 1,525 million m³ is renewable water, (0.015% of the total groundwater). In Al-Mukalla (the area of the governorate) and Ramlet al-Sabaeen owns an estimate of 10 thousand billion cubic meters of water reserves, (96.4% of the total groundwater). Fig. 3.3, highlights the main groundwater basins of Yemen (Gun, et al., 1996).

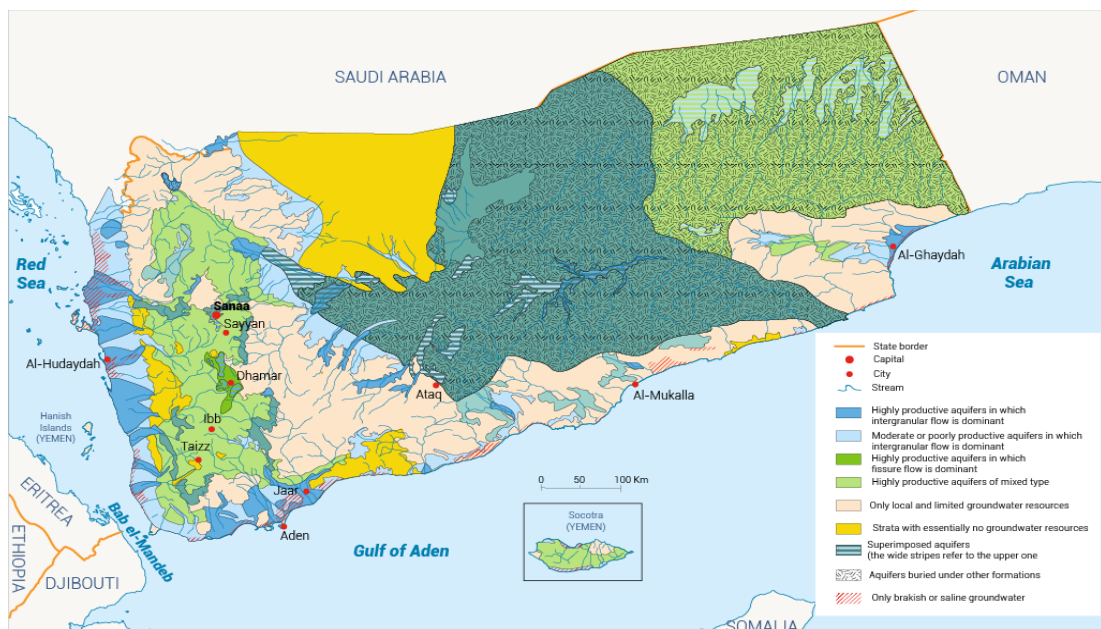


Figure 3. 3: Map of Yemen showing the main groundwater basins (Gun et al., 1996)

3.5.2 Unconventional Water Resources

3.5.2.1 Treated Wastewater

Different wastewater treatment technologies are implemented in Yemen, ranging from the most advanced kind of activated sludge process with ventilation to a simplest waste stability basin. More than 17 urban plants and 15 small rural wastewater treatment plants are operating where the total wastewater flow is nearly 300,000 cubic meters per day, (or about 100 million cubic meters per year). The treated volume quantity of the plants in the places like Sanaa and Ibb exceeded their design capacities, while the other treatment plants, as in Aden, are still below the required designed level. Biological Oxygen Demand (BOD) is used to characterize the treated water quality. A study conducted in 2005 showed that, Sanaa treatment plant treated water BOD was significantly increased from 550 mg/lit in 1985 to 800 mg/lit in 1992 and even reached to 1,100 mg/lit in 2000-2004.

3.5.2.2 Seawater Desalination

In Yemen, there are two desalination plants, the one in the city of Aden is called the Al Haswa Power Station. This plant is used to supply electricity to the city by heating the sea water and as a byproduct generates 69,000 cubic meters of freshwater per day. This quantity of water is added to the existing water supply network of Aden. The second desalination plant is still under construction for extension which is located 100 km to the south of Taiz Governorate in Al-Mokha area. The plant was constructed by a private sector (Hail Saeed Group) in 2002 with a total capacity for desalination (design capacity) volume 76,596 m³/day (or 28 million m³/year). The desalinated water production volume reached to 25,1 million cubic meters in 2006, (Wangnick Consulting, 2002).

3.6 Water Scarcity in Yemen

With only 125 cubic meters per person per year of renewable water resources, Yemen is among one of the most water-stressed countries in the world. This pattern is less than one-tenth of the threshold of water stress, which is 1,700 cubic meters per person per year (IRIN, 2008). The total demand for water exceeds 3,400 million cubic meters per year from the annual 2,500 million cubic meters of renewable resources, which results a drop in groundwater levels from one meter in Tuban Abyan to 6-8 meters per year in Sana'a Basin (Climate Change, 2001). It is estimated that, there are nearly 45,000 to 70,000 wells in Yemen of which most of them are under the private sector control where in fact, nobody can confirm the exact number since almost everyone was dug without a governmental permission or license (NWRA, 2007). Agriculture is the largest proportion of Yemen's water resources where it uses nearly 90% of which 37% is used just for irrigation (Yehya et al., 2005). Due to global changes, it is found that, the average rainfall in Sanaa has decreased by one sixth from 240 mm (between 1932 and 1968) to 200 mm (between 1969 to 1982) and to 180 mm (between 1983-2000) (Lichtenthäler, 2010).

3.7 Rainfall Stations

Yemen as mentioned above, can be grouped into four main basins. In this study, due to civil war, none of the existing meteorological stations were functioning. Hence, with the help of Thiessen polygon approach, 3 representative hypothetical rainfall station locations for each basin were determined and coded as shown in Fig. 3.4 as a part of this study.

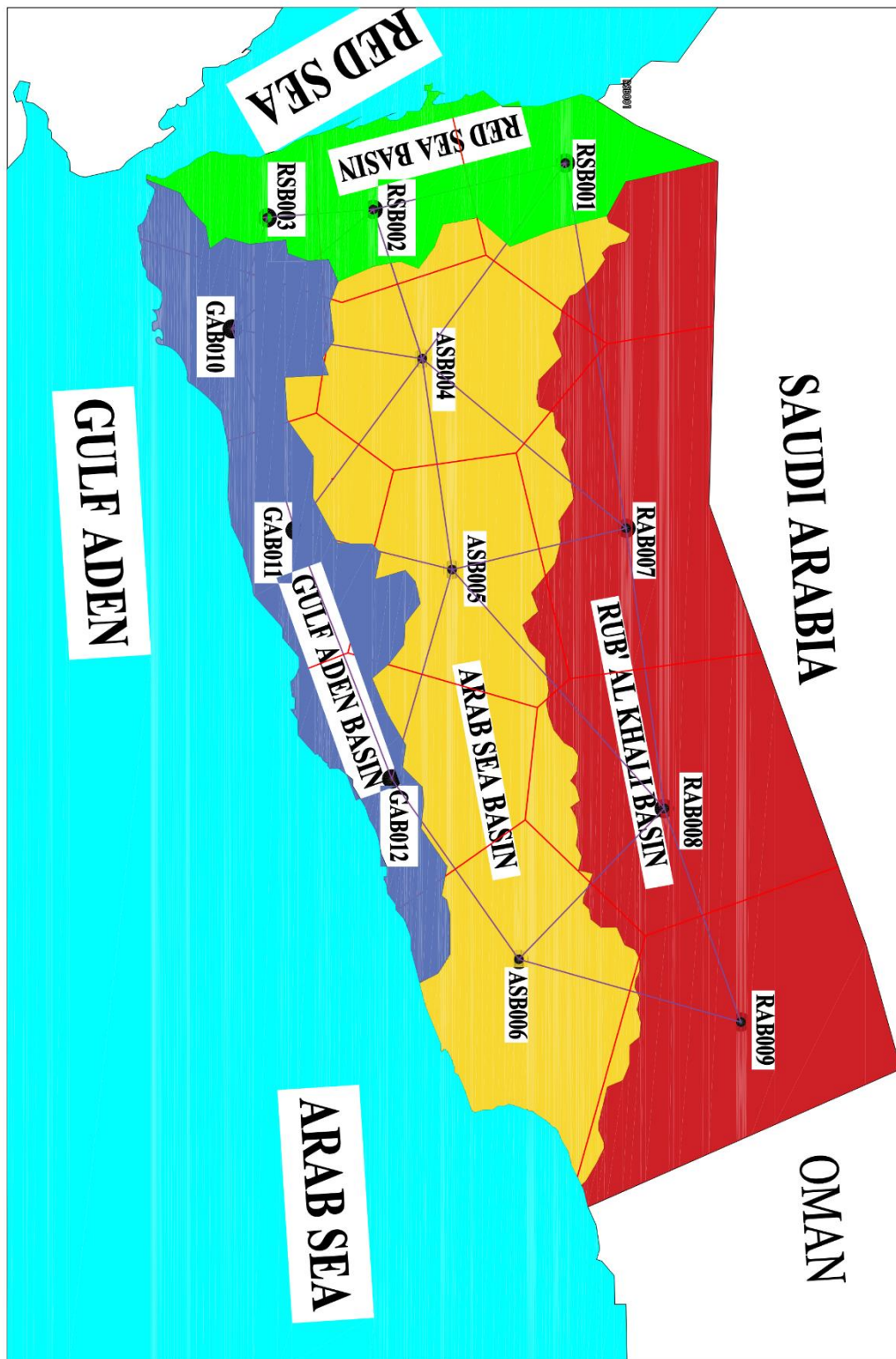


Figure 3. 4: Thiessen polygons based 12 representative rainfall stations of Yemen

3.8 Rainfall Data Details

With the help of Global Positioning System (GPS), the coordinates of the suggested hypothetical stations were detected as given in Table 3.1. Since no observed (measured) datasets are available, the monthly rainfall data values of these 12 hypothetical meteorological stations from January 1981 to December 2018 were gathered from the Prediction of Worldwide Energy Resources (POWER) archive. In fact, these datasets were obtained from the NASA Langley Research Center (LaRC) for POWER Project that was funded by the NASA Earth Science/Applied Science Program. These rainfall datasets, were derived from the NASA's GMAO MERRA-2 assimilation model GEOS 5.12.4 FP-IT. Note that, MERRA-2 is the recent version of NASA's Goddard Earth Observing System Data Assimilation System (GEOS) (Bosilovich et al. 2016). The GEOS version 5.12.4 has the same grid resolution as MERRA-2 (0.5 X 0.5 implying 50 km X 50 km). The GEOS 5.12.4 dataset is processed by the POWER project team on a daily basis and appended to the end of the MERRA-2 daily time-series to provide low latency products which are generally ready within about 2 days of real-time. The MERRA-2 values in the resulting daily time-series are typically updated every several months. Hence, the gathered rainfall data sets were the normalized rainfall roughly obtained values.

Table 3. 1: The GPS coordinates of the selected representative rainfall stations for each basin of Yemen

Basins	Stations		
	Codes	GPS Coordinates	
		Latitude	Longitude
Red Sea	RSB001	15.9692	43.5799
	RSB002	14.8041	43.8435
	RSB003	13.8354	44.0523
Arab Sea	ASB004	15.8513	44.419

	ASB005	16.6319	50.9888
	ASB006	16.3369	48.0884
Gulf Aden	RKB007	13.6272	45.1475
	RKB008	13.9153	46.2571
	RKB009	14.4266	48.0149
Rub' Al Khali	GAB010	17.0000	44.4519
	GAB011	17.7863	51.4282
	GAB012	17.0000	44.4519

Chapter 4

THEORY AND METHODOLOGY

4.1 Introduction

Statistics is a scientific knowledge that collects information regardless of its source, in most of the disciplines like agriculture, industry, medicine, economy, engineering, etc. It is also recognized to be one of the fundamental sciences in many areas due its widespread applicability. Definitely, it is one of most effective and vital resource in hydrological studies especially on climatic datasets where the predictions and forecasts are needed. It describes and explains the group of data with the help of basic scientific theories and tools that makes the presentation of the datasets easy, transparent, organized, more understandable, accurate, and complete in terms of assessment, interpretation and research that is analyzed with the help of some measures.

The hydrological data is a time-series dependent dataset, that may be stationary or non-stationary, since the nature of the hydrological processes is not strictly deterministic. So, the detection of any physical meaning in any time-series of hydrological variables is essential for any scientific study and even in practice. Especially, the data predictions and forecasts for the water resources systems, in fact, has been developed on the basis of statistical and probabilistic hydrology. Therefore, if for any design, the assumed (predicted or forecasted) value is incorrect, then the existing project or the plan has to be revised, due overhaul or under design, causing overhead cost and waste of time (Scholze, et al., 2006).

4.2 Definitions

Statistics: is a science that deals with collecting, analyzing, and presentation of the datasets:

Data: any quantity or quality that can be collected and used for any decision-making study;

Population: the complete group of all components of the study that includes all the subjects of that study;

Sample: is a representative subset of members selected from any specific population.

Usually in the statistical studies, if the size of the data in a dataset is less than or equal to 30, this dataset is referred as sample (Seyhan, 1994);

Parametric: any statistical property that is defined with the help of basic mathematical tools;

Non-parametric: any statistical property that measures without using even basic mathematical tools;

Confidence interval: in statistics, refers to the probability that a population parameter will fall between a set of values for a certain proportion of times. It is a measure the degree of uncertainty or certainty in a sampling method. Though can be any number of probability limits, with the most common being a 95% or 99% confidence level.

Degrees of freedom: is the number of independent coordinates of any statistical function;

Correlation: is a measure in statistics used to denote association between two quantitative variables;

Auto-covariance: is a coefficient that describes the degree of linear dependence of any time based organized successive data;

4.3 Time-Series

Time-series can be defined as collecting data points measured in a constant interval of time. In other words, a time-series is a sequence where the dataset is recorded over regular time intervals. It is simply a set of ordered data points with respect to time. Depending on the need and availability of the gathered data, a time-series can be of yearly, seasonally, monthly, weekly, daily, hourly, minutes, and even seconds wise.

Time-series is used in statistical methods for analyzing the dataset collected in successive time intervals so as to extract meaningful statistical results. This analysis is comprised of different algorithms or methods used to extract certain statistical information and characteristics of data, in order to predict the future values based on stored past time-series dataset. Hence, time-series helps in analyzing the past data, which then becomes an essential factor in forecasting the future data.

Any time-series is composed of three pattern types:

1- Trend;

A trend describes the time-series of the dataset without any overlaying or repeating in time base. The trend pattern exists, when there is a medium-term or long-term increase or decrease in the data. Sometimes, it is referred as changing direction, when it might go from an increasing trend to a decreasing trend where it does not have to be linear.

2- Seasonality;

A seasonal pattern exists, when a time-series is influenced by seasonal factors. It can be observed, if the same behavior of trend line is repeated in systematic intervals (periodically) over the time. Seasonality always has a fixed and known period.

3- Cyclic;

A cyclic pattern exists, when data exhibit rises and falls that are not of fixed period. The duration of these fluctuations is usually of at least 2 years. Many people confuse cyclic behavior with seasonal behavior, but they are really quite different. If the fluctuations are not of fixed period then they are cyclic; if the period is unchanging and associated with some aspect of the calendar, then the pattern is seasonal. In general, the average length of cycles is longer than the length of a seasonal pattern, and the magnitude of cycles tends to be more variable than the magnitude of seasonal patterns.

4.4 Time-Series Forecasting

Predicting the behavior of a variable over a time is a common problem that one encounters in many real case studies. Despite its importance, time-series forecasting is a topic often overlooked in Machine Learning. Hence, how to approach a time-series problem using Machine Learning techniques will be detailed below. Time-series analysis is particularly hard, because there is a difficulty that doesn't occur with other problems in Machine Learning, since the data has a particular order and it is highly correlated. This means that, if one takes two observations with the exact same attribute values, the outcome may be totally different due to the recent past measurements. This is mainly due to practical implications when one is attacking the problem to solve. For example, splitting the data between training and validation sets can't be done at random like one would do with typical Machine Learning problems because, the order of the data itself contains a lot of information.

On the other hand, forecasting which implies predicting the future values based on the time-series dataset one is going to take. Can be classified into two:

- 1- if only the previous values of the time series are used to predict its future values, it is called Uni-variate Time-series Forecasting whereas,
- 2- if the predictors other than the time-series are as well used to forecast, it is called Multi-variate Time-series Forecasting.

In any of the regression model, the response variable in the previous time period has become the new predictor, and errors have been associated in any simple linear regression model.

Note that, for a prediction of time t , the predicted data relies on $t-1$ and so on all the way, till $t-n$. This is called lagged prediction, since it relies on the data points that were in the previous period of time.

4.5 The ARIMA Model

ARIMA is an acronym that stands for AutoRegressive Integrated Moving Average. This acronym is descriptive, that is capturing the key aspects of the model itself. The ARIMA modeling approach offers a model-driven technique to time-series forecasting by using a theoretical framework developed by George E. P. Box (1919– 2013) and Gwilym M. Jenkins (1932– 1982) so usually referred as the Box–Jenkins method.

ARIMA forecasting algorithm based on the idea that, the information in the past values of the time-series can alone be used to predict the future values. It is actually a class of models that explains a given time-series based on its own past values, (i.e. its own lags and the lagged forecast errors), so that, equation can be used to forecast future values. The AutoRegressive Integrated Moving Average (ARIMA) is the go-to model for any time-series forecasting:

1- AutoRegressive Model (AR)

The AutoRegressive model is just a linear regression model that fits the present value based on previous values 'p', sometimes referred as lagged predictions. In other words, autoregression is a model that uses the dependent relationship between data and some number of lagged observations.

The notation AR(p) indicates an autoregressive model of order p where the AR(p) model is defined as:

$$y_t = c + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \varepsilon_t \quad \text{Eq. (4.1)}$$

where,

y_t is the variable of model,

c is the constant,

β_i is the parameter of the model varying from $\beta_1, \dots, \beta_p$ and

ε_t is the white noise error term.

2- Moving Average Model (MA)

The Moving Average model uses the dependency between a data and a residual error from a moving average model that is applied on the lagged data. This model proposes that, output is a linear combination of the current and various past values of a random variable.

The notation MA(q) indicates a moving average model of order q. The MA(q) model is defined as:

$$y_t = \mu + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad \text{Eq. (4.2)}$$

where,

y_t is the variable of the model,

μ is the constant term of the model,

θ_i is the parameter of the model varying from $\theta_1, \dots, \theta_q$ and

ε_t is the white noise error term.

3- Integrated (I)

The term Integrated, implies the use of differencing of raw data (e.g., subtracting a data from the data of the previous time step) in order to make the time-series stationary, In time-series forecasting, which is an implicit assumption, the model depends on time in some capacity. This seems pretty obvious. With that assumption out of the way, one needs to understand where on the spectrum of dependence, time falls in relation to the studied model.

So, instead of trying to forecast the value of the observed variable, it is easier to forecast how different the new value will be with respect to the last one. This means that, using the difference ‘d’, between consecutive steps as the target variable, instead of the observable variable itself. In other words, it is defining the number of differencing required to make the time-series stationary.

Each of these three components are explicitly specified in any ARIMA model as a parameter. A standard notation is, ARIMA(p,d,q) where by definition, the parameters have only positive integer values. Hence, to be able to define any ARIMA model, it is essential to characterize it by 3 terms (hyper-parameters) p, d, q. With the properly selection of these three hyper-parameters, it is expected that, one can get the best possible model. Note that, p defines the order of the Auto Regressive (AR) term (i.e. it refers to the number of lag data within the model which is used as predictors, which is also called the lag order), d defines the order of differencing term (i.e. it refers to the number of differencing required to make the time-series stationary, which is also called the degree of differencing), and q defines the order of the Moving Average (MA) term

(i.e. it refers to the number of lagged forecast errors, that defines the size of the moving average time window, which is also called the order of moving average) (Yurekli et al., 2005).

The mean and the constant terms are related by the below given equation as:

$$\text{Constant} = \text{Mean} * [1 - (\text{the sum of the AR coefficients})].$$

Hence, the mathematical presentation is given as:

$$\hat{y}_t = c + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} - \theta_1 \varepsilon_{t-1} - \dots - \theta_q \varepsilon_{t-q} \quad \text{Eq. (4.3)}$$

An ARIMA model is the one, where the time-series was differenced at least once to make it stationary and by combining the AR and the MA terms.

ARIMA model in words:

Predicted y_t = constant + linear combination lags of y (up to p lags) + linear combination of lagged forecast errors (up to q lags)

4.5.1 Comparison Between ARMA and ARIMA

The only difference between ARMA “Autoregressive Moving Average” and ARIMA “Autoregressive Integrated Moving Average” is the “integrated” part. The word Integrated refers to the number of times needed to difference a series in order to achieve stationarity, which is required for ARMA models to be valid. By differencing, it means forming a new series by subtracting observation 1 from 2, 2, from 3, etc. So an ARMA model is equivalent to an ARIMA model of the same MA and AR orders with no differencing.

The typical short-hand notation for ARMA is “ARMA(p,q)” where p is the AR order and q is the MA order. For ARIMA, the notation is “ARIMA(p,d,q)” where the added

d is the order of integration, or number of differences. So the following two statements sum it all up:

1. ARMA(p, q) is equivalent to ARIMA($p, 0, q$);
2. Given an ARIMA(p, d, q), if $d > 0$ one can model this as an ARMA by running an ARMA(p, q) after differencing the original series d times.

4.6 How to Built an ARIMA Model

A good way to generate an ARIMA model, is to simply propose some values for each hyper-parameter (any positive whole number from 0 up to say 5, where the upper limit depends on the sample size). A value of 0 can be used for a parameter, which indicates not to use that element within that model. The ARIMA model can be configured to perform the function of an ARMA model, and even a simple AR, I, or MA model.

Some special ARIMA(p, d, q) models:

- ARIMA(0, 0, 0) is a white noise model;
- ARIMA(1, 0, 0) is a first-order autoregressive model (AR(1));

$$\hat{y}_t = c + \beta_1 y_{t-1} + \varepsilon_t$$

- ARIMA(2, 0, 0) is a second-order autoregressive model (AR(2));

$$\hat{y}_t = c + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \varepsilon_t$$

- ARIMA(0, 0, 1) is a first-order moving average model (MA(1));

$$\hat{y}_t = \mu + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

- ARIMA(0, 0, 2) is a second-order moving average model (MA(2));

$$\hat{y}_t = \mu + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \varepsilon_t$$

- ARIMA(0, 1, 0) is simply a random walk model (or I(1) model);

$$\hat{y}_t = c + y_t - y_{t-1}$$

- ARIMA(0, 2, 0) is a random walk model of degree 2 (or I(2) model);

$$\hat{y}_t = c + y_t - 2y_{t-1} + y_{t-2}$$

- ARIMA(0, 1, 1) is a simple exponential smoothing model;

$$\hat{y}_t = y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

- ARIMA(1, 1, 0) is a differenced first-order autoregressive model;

$$\hat{y}_t = c + y_{t-1} + \beta_1 (y_{t-1} - y_{t-2})$$

- ARIMA(0, 2, 1) is a linear exponential smoothing model;

$$\hat{y}_t = 2y_{t-1} - y_{t-2} - \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

- ARIMA(1, 1, 1) is a simple mixed model;

$$\hat{y}_t = \mu + y_{t-1} + \beta_1 (y_{t-1} - y_{t-2}) - \theta_1 \varepsilon_{t-1}$$

- ARIMA(0, 1, 2) is a Damped (flatten over time) Holt's model;
- ARIMA(0, 2, 2) is double exponential smoothing model (Holt's linear method with additive errors);
- ARIMA(1, 1, 2) is a damped-trend linear exponential smoothing model;

$$\hat{y}_t = y_{t-1} + \beta_1 (y_{t-1} - y_{t-2}) - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2}.$$

One correct way to split the data would be, to keep the first 3/4 of the observations to train the model and the last 1/4 of the observation to validate and test the model's accuracy (i.e., approximately 75:25 ratio is reasonable proportion).

Hence, adjusting the model with the training data, and see how well each model performs by comparing the predicted and the tested (validation) dataset is important. This is called hyper-parameter optimization, and it is often done wrong. The score of each model with different parameters should be obtained against the validation set, not against the training set.

4.6.1 Unit Roots Concept

If a series is grossly under- or over-differenced (i.e., if a whole order of differencing needs to be added or cancelled), this is often signaled by a unit root in the estimated AR or MA coefficients of the model. Hence, if any time-series is non-stationary, the unit root in the AR coefficients needs a higher order differencing. Therefore,

- if the unit root in the AR part of the model has a sum of the AR coefficients almost exactly 1, one should reduce the number of AR terms by one and increase the order of differencing by one.
- if there is a unit root in the MA part of the model has a sum of the MA coefficients almost exactly 1, one should reduce the number of MA terms by one and reduce the order of differencing by one.

4.6.2 AutoCorrelation and Partial AutoCorrelation Functions

The ACF and the PACF are widely used in identifying ARMA models. ACF plot is merely a bar chart of the coefficients of ordinary correlation between a time-series and lags of itself. On the other hand, PACF plot is a plot of the partial correlation coefficients between the series and lags of itself.

4.6.2.1 ACF and PACF of AR (p)

The ACF of an AR (1) process is depicted in Figure 4.1. There is a decaying pattern in the ACF; the decay is exponential for $0 < \beta_1 < 1$ (Figure 4.1a), whereas, for $-1 < \beta_1 < 0$ (Figure 4.1b), the ACF is similar but alternates in sign. The PACF shows a single positive value at lag 1 of $0 < \beta_1 < 1$ and a negative spike at lag 1 of $-1 < \beta_1 < 0$.

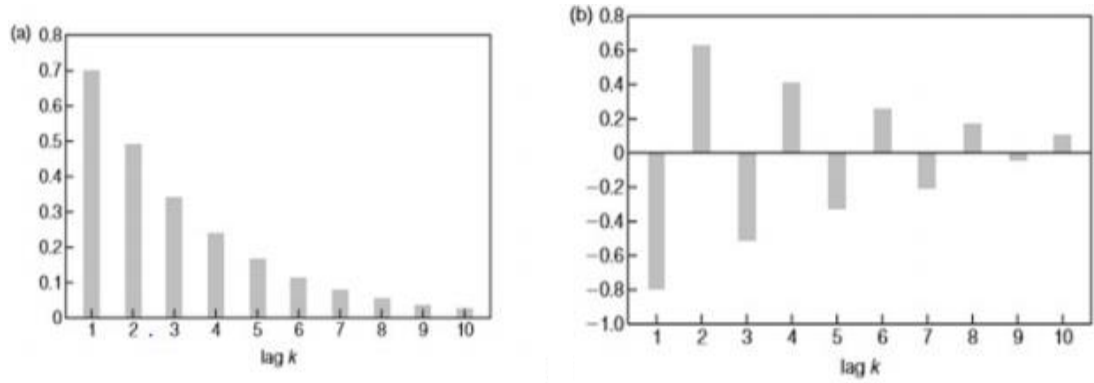


Figure 4. 1: (a) ACF of an AR (1) process $\beta_1 = 0.70$
(b) ACF of an AR (1) process $\beta_1 = -0.80$

The PACF is more complex to describe. It measures the correlation between y_t and y_{t-k} adjusted for the intermediate values $y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$ (or the correlation between y_t and y_{t-k} not accounted for by $y_{t-1}, y_{t-2}, \dots, y_{t-k+1}$). If one denotes β_{kj} of the j^{th} coefficient in an AR(k) model, then β_{kk} being the last coefficient, so it can be shown that, the β_{kj} will be nonzero for $k \leq p$ and zero for $k > p$, where p is the order of the autoregressive process.

Another basic process that occurs fairly often in practice is the AR(2) process. In this case, there are two autoregressive coefficients β_1 and β_2 . Figure 4.2 shows the ACF and the PACF of an AR (2) model with $\beta_1 = 0.3$ and $\beta_2 = 0.5$.

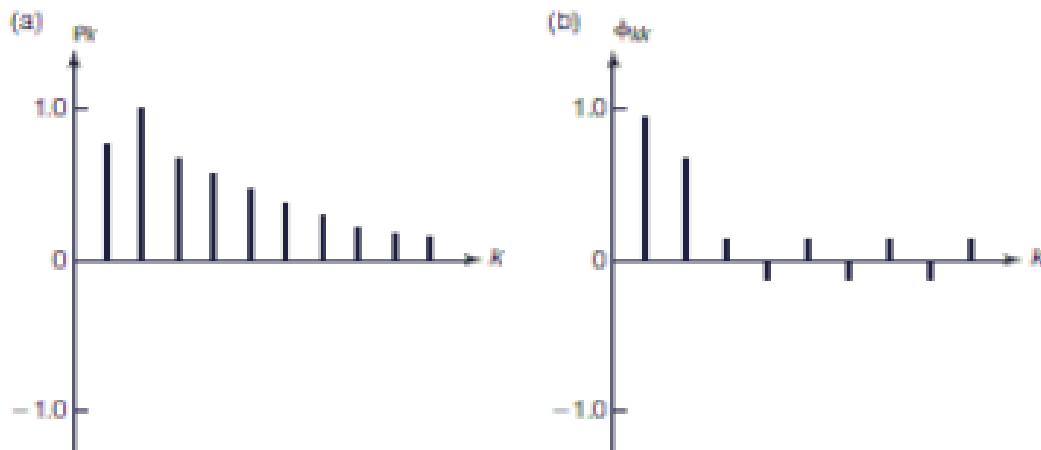


Figure 4. 2: (a) ACF and (b) PACF of an AR (2) model with $\beta_1 = 0.3$ and $\beta_2 = 0.5$

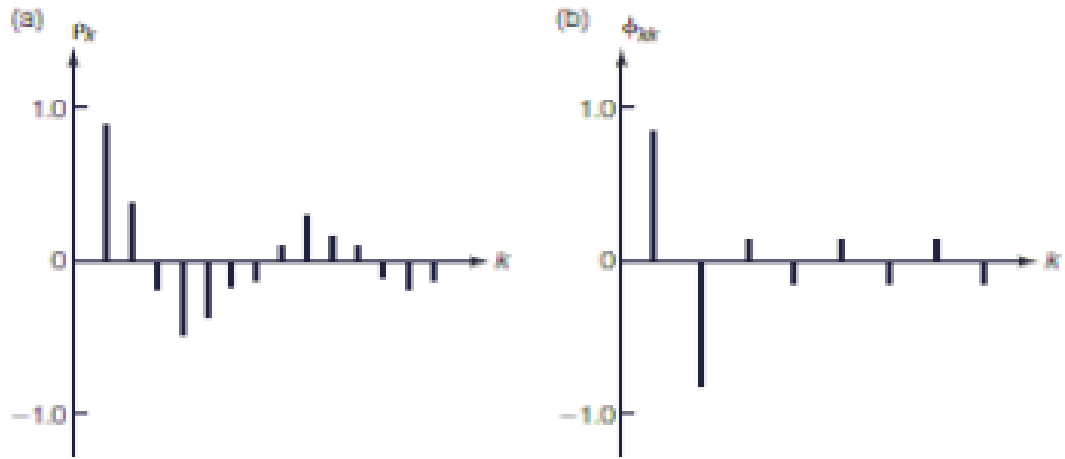


Figure 4. 3: (a) ACF and (b) PACF of an AR (2) model with $\beta_1 = 1.2$ and $\beta_2 = - 0.64$

Figure 4.3 shows ACF and ACF of an AR (2) model with $\beta_1 = 1.2$ and $\beta_2 = - 0.64$. The values in the ACF decay in a sinusoidal pattern, whereas the PACF has a positive value at lag 1 and a negative value at lag 2.

4.6.2.2 ACF and PACF of MA (q)

To identify if the model needs any MA terms, one can find out the required number of MA terms by inspecting mainly ACF plot. An MA term is technically, the error of the lagged forecast.

The ACF of a MA(q) process is 0, beyond the order q of the process (i.e., it has a cutoff after lag q). The ACF of a MA (1) process has one spike at lag 1, the others are 0. It has the value $\rho_1 = - \theta_1 / (1 + \theta_1^2)$ with $|\rho_1| \leq 1/2$.

The PACF of MA process is complicated, so in Figure 4.4 ACF and PACF of an MA (1) model with positive θ_1 is presented where there is a single negative spike at the lag 1 in the ACF and a decaying pattern in the PACF.

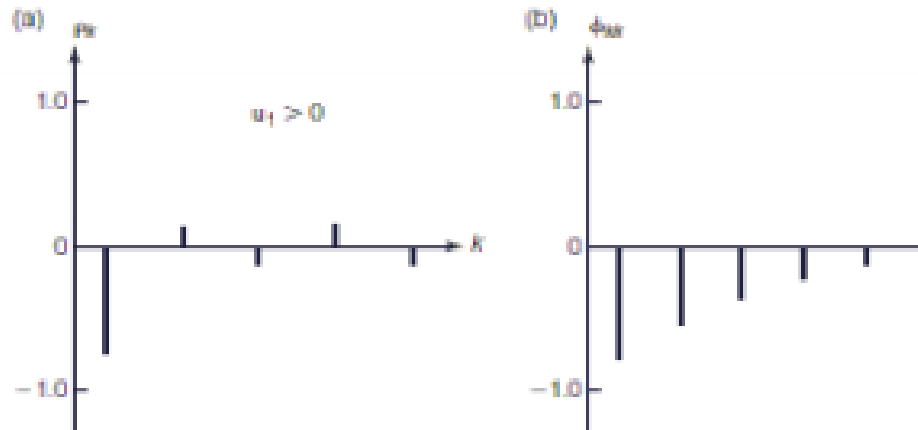


Figure 4. 4: (a) ACF and (b) PACF of a MA(1) model with positive parameter θ

The ACF of an MA(1) process with negative θ shown in Figure 4.5, where a single positive spike for ACF, but PACF shows a decaying pattern with spikes alternating above and below the zero line.

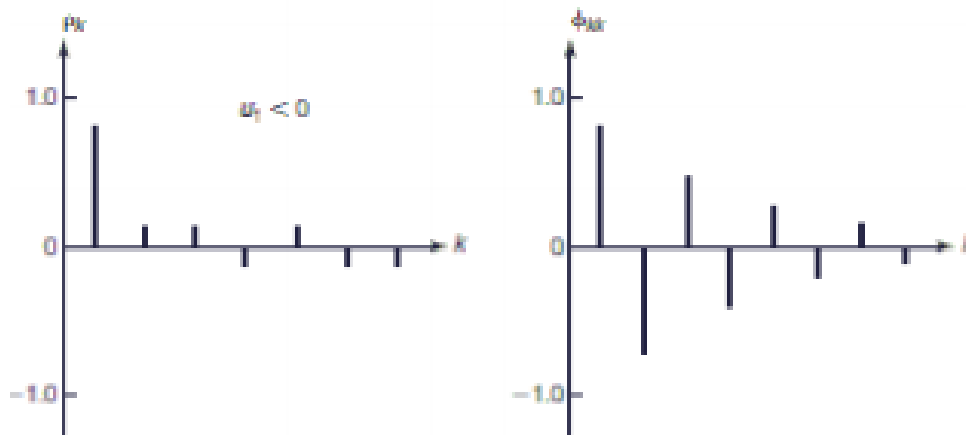


Figure 4. 5: (a) ACF and (b) PACF of a MA(1) model with negative parameter θ

4.6.3 Alternative ARIMA Models

4.6.3.1 ARIMA(0,0,0)

Its ACF is given in Figure 4.6.

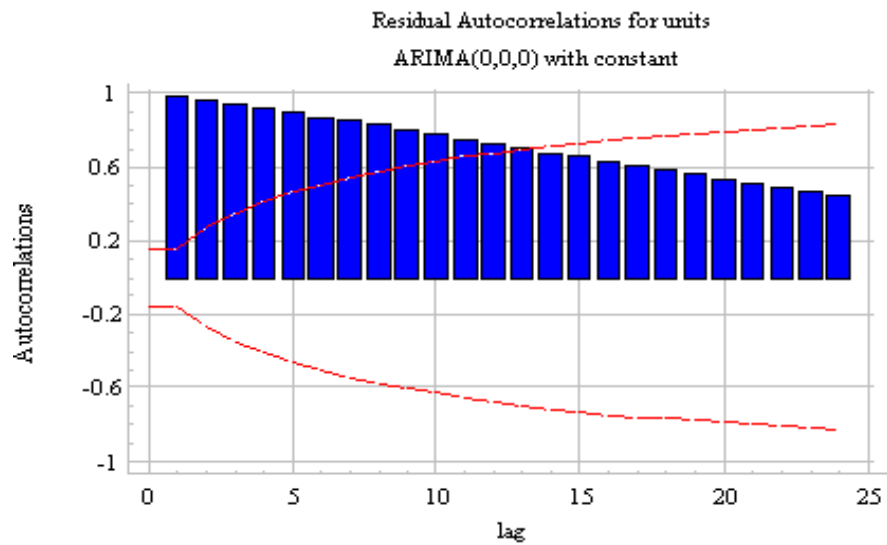


Figure 4. 6: ACF of ARIMA(0,0,0)

The autocorrelations are significant for a large number of lags, but perhaps the autocorrelations at lags 2 of PACF is given in Figure 4.7.

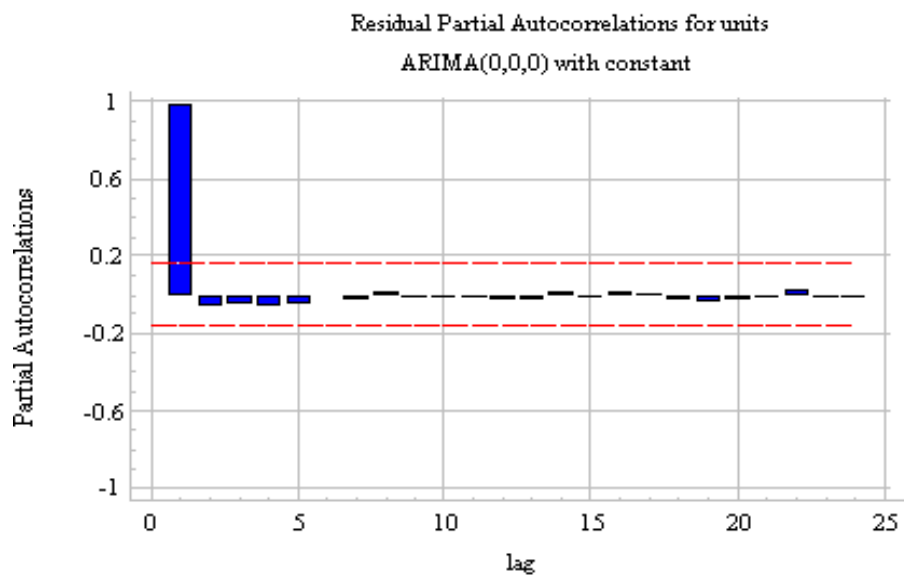


Figure 4. 7: PACF of ARIMA(0,0,0)

The forecasting equation for an AR(1) model for a series y with no orders of differencing is:

$$\hat{y}_t = \mu + \beta_1 y_{t-1}$$

If the AR(1) coefficient β_1 in this equation is equal to 1, it is equivalent as if it is predicting the first difference of y as constant (i.e. it is equivalent to the equation of the random walk model with growth):

$$\hat{y}_t = \mu + y_{t-1}$$

4.6.3.2 ARIMA(0,1,0)

In an AR(1) model, the AR term acts like a first difference where the autoregressive coefficient is equal to 1 where its ACF and PACF are given in Figure 4.8 and 4.9 respectively. Note that,

- ✦ if the PACF of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is positive, and;
- ✦ if the ACF of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is negative (i.e. if the series appears slightly over-differenced) then, consider adding an MA term to the model. The lag at which the ACF cuts off is the indicated number of MA terms.

The time-series needed (at least) one order of non-seasonal differencing to be stationarized as was determined. After taking one non-seasonal difference (i.e. fitting an ARIMA(0,1,0) model with constant) the ACF and PACF plots look like that is given in Figure 4.8 and 4.9 respectively.

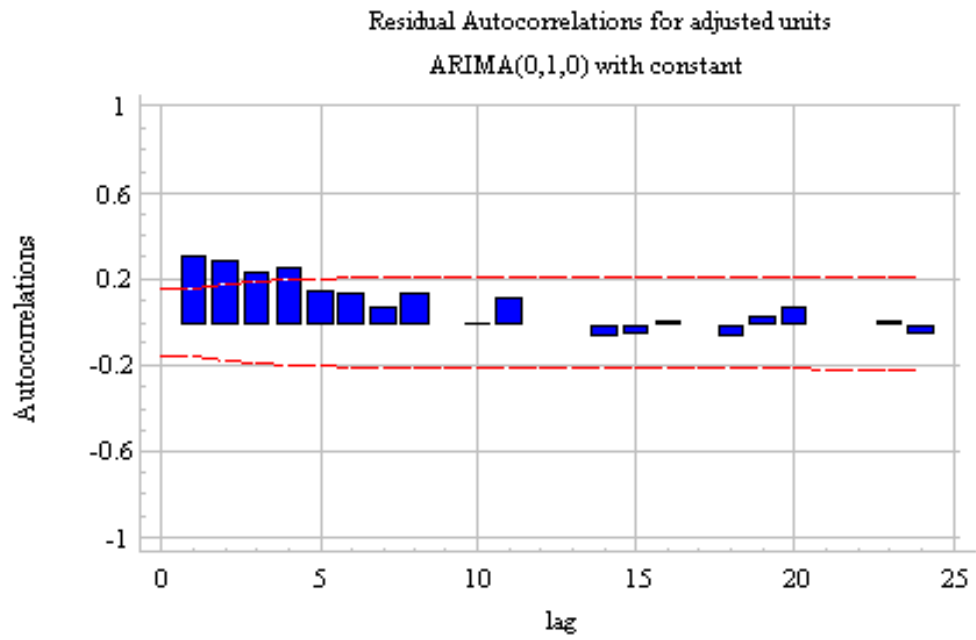


Figure 4. 8: Residual ACF of ARIMA(0,1,0)

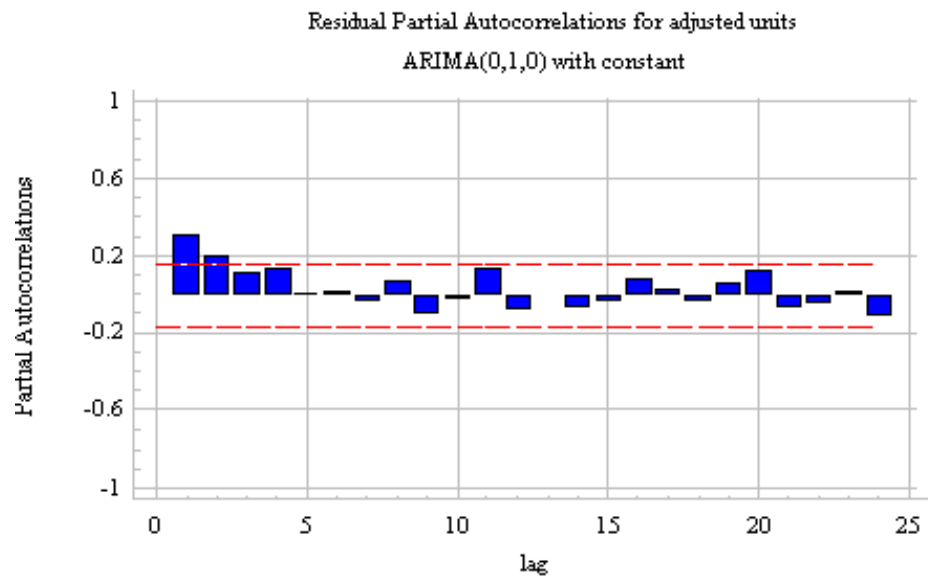


Figure 4. 9: Residual PACF of ARIMA(0,1,0)

4.6.3.3 ARIMA(2,1,0)

If one sets the order of the AR term to 2, (i.e. fit an ARIMA(2,1,0) model) can obtain, the following ACF and PACF plots for the residuals as given in Figure 4.10 and 4.11 respectively.

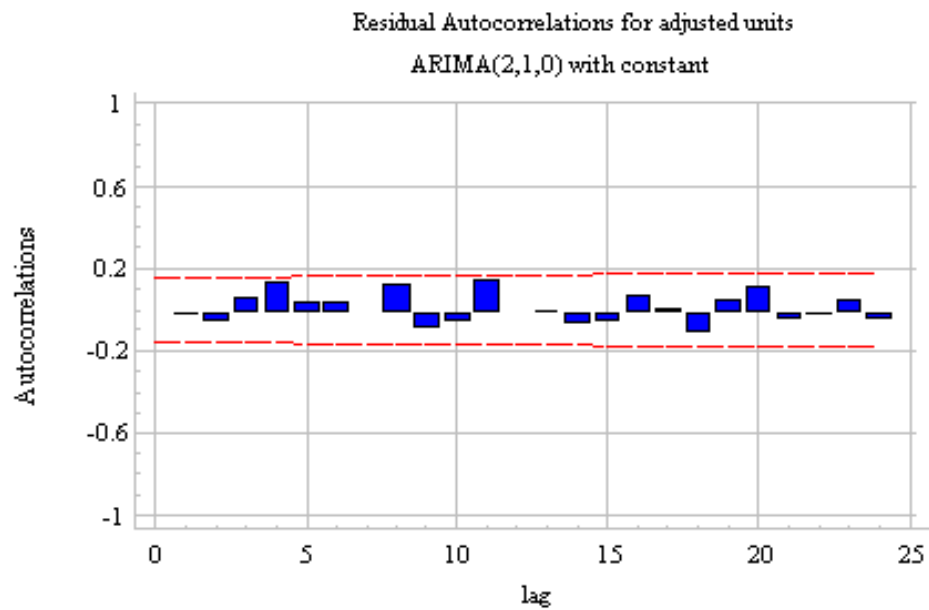


Figure 4. 10: Residual ACF of ARIMA(2,1,0)

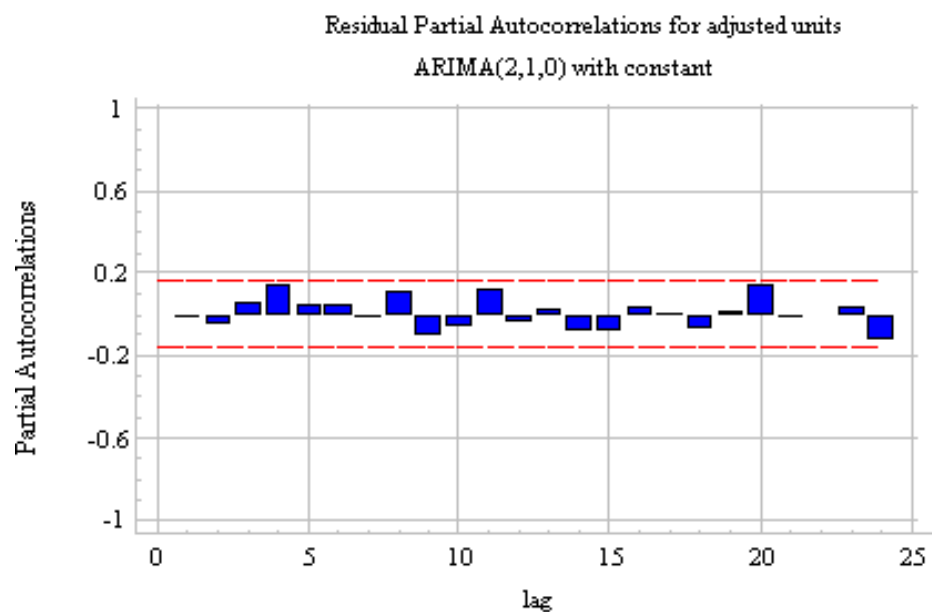


Figure 4. 11: Residual PACF of ARIMA(2,1,0)

The autocorrelation at the crucial lags; namely lags 1 and 2; has been eliminated, and there is no discernible pattern in higher-order lags. The time-series plot of the residuals shows a slightly worrisome tendency to wander away from the mean as given in Figure 4.12.

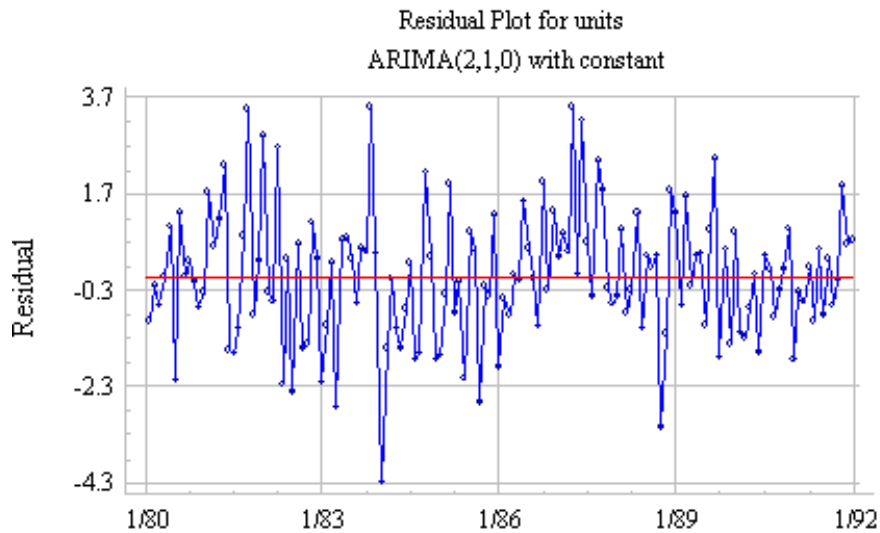


Figure 4. 12: Residual plot of ARIMA(2,1,0)

The (untransformed) forecasts for the model show a linear upward trend projected into the future as shown in Figure 4.13.

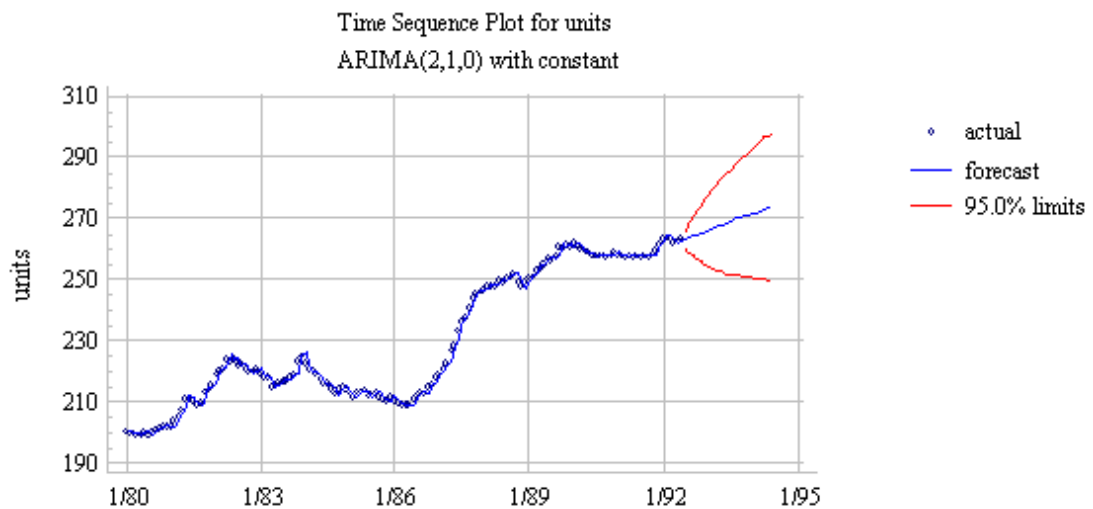


Figure 4. 13: Plot of ARIMA(2,1,0)

4.6.3.4 ARIMA(0,2,0)

One order of non-seasonal differencing yielded the lowest standard deviation (and a pattern of mild positive autocorrelation), while two orders of non-seasonal differencing yielded a more stationary-looking time-series plot (but with rather strong

negative autocorrelation). The ACF and PACF of the series (Figure 4.14 and 4.15) with two non-seasonal differences:

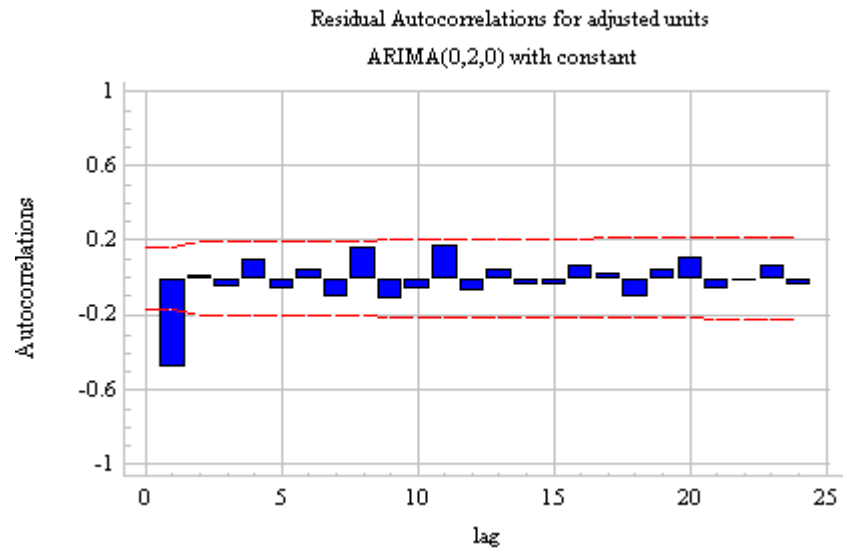


Figure 4. 14: Residual ACF of ARIMA(0,2,0)

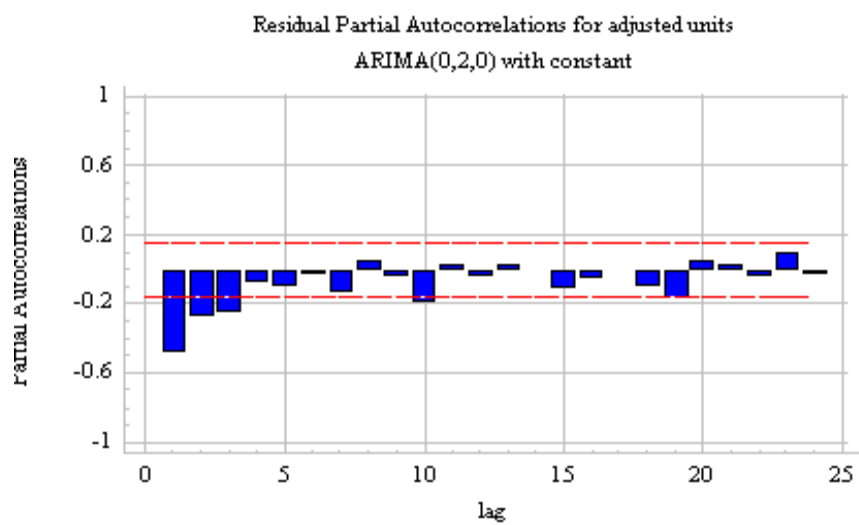


Figure 4. 15: Residual PACF of ARIMA(0,2,0)

4.6.3.5 ARIMA(0,2,1)

```
Analysis Summary
Data variable: units
Number of observations = 150
Start index = 1/80
Sampling interval = 1.0 month(s)

Forecast Summary
-----
Nonseasonal differencing of order: 2

Forecast model selected: ARIMA(0,2,1)
Number of forecasts generated: 24
Number of periods withheld for validation: 30
```

Statistic	Estimation Period	Validation Period
MSE	2.13793	0.856734
MAE	1.15376	0.771561
MAPE	0.518221	0.297298
ME	0.0267768	-0.038966
MPE	0.017097	-0.0148876

Parameter	Estimate	Std. Error	t	P-value
MA(1)	0.75856	0.0607947	12.4774	0.000000

```
Backforecasting: yes
Estimated white noise variance = 2.1404 with 147 degrees of freedom
Estimated white noise standard deviation = 1.46301
Number of iterations: 4
```

Notice that, the estimated white noise standard deviation (RMSE) is only very slightly higher for this model than the previous one (1.46301 here versus 1.45215 previously).

4.6.3.6 Comparison of ARIMA(2,1,0) and ARIMA(0,2,1)

Below is the model comparison report that shows the results of fitting the ARIMA(2,1,0) model and ARIMA(0,2,1) model

Model Comparison

Data variable: units
Number of observations = 150
Start index = 1/80
Sampling interval = 1.0 month(s)
Number of periods withheld for validation: 30

Models

(A) ARIMA(2,1,0) with constant
(B) ARIMA(0,2,1)

Estimation Period

Model	MSE	MAE	MAPE	ME	MPE
(A)	2.10757	1.11389	0.500834	0.0197748	0.0046833
(B)	2.13793	1.15376	0.518221	0.0267768	0.017097

Model	RMSE	RUNS	RUNM	AUTO	MEAN	VAR
(A)	1.45175	OK	OK	OK	OK	OK
(B)	1.46217	OK	OK	OK	OK	OK

Validation Period

Model	MSE	MAE	MAPE	ME	MPE
(A)	0.757818	0.726546	0.280088	-0.200813	-0.0778831
(B)	0.856734	0.771561	0.297298	-0.038966	-0.0148876

The two models, perform nearly identically in the estimation period. But on the basis of these statistical results alone, it would be hard to choose among the two models. However, if one plots the long-term forecasts made by the ARIMA(0,2,1) model as shown in Figure 4.16, can easily observe a significant difference from the earlier model as was detailed in Figure 4.13.

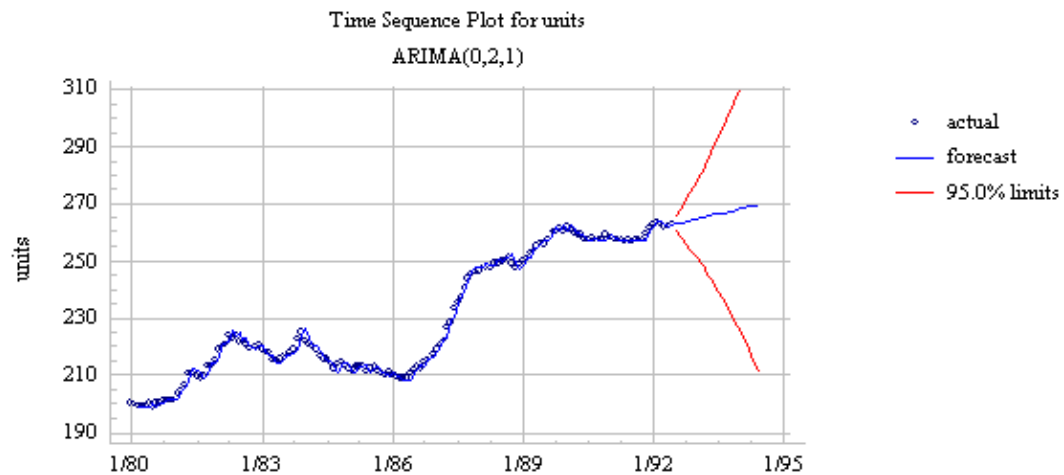


Figure 4. 16: Plot of ARIMA(0,2,1)

4.7 Akaike Information Criterion

Fitting a model refers to an examination of whether the statistical model employed in an application adequately explains the important features of the dataset at hand. On the other hand, selection a model refers to the choice of the statistical model that describes the dataset best, among several competing models. The model fit and model selection analysis for the linear models can be done by using residual analysis.

The most important part is to make sure that, the residuals of the model is random, and the estimated parameters are statistically significant. So, by determining the residuals of the chosen model from the plottings of ACF and PACF and doing several tests, the model is being checked. One of the popular residual analysis approach is Akaike Information Criterion (AIC).

The Akaike Information Criterion (AIC) is a mathematical method used for evaluating how well a model fits the dataset it was generated from. It is derived from frequency probability. In statistics, AIC is used to compare different possible models for

determining which one is the best fit for the dataset. The Akaike Information Criterion is a way of selecting a model from a set of models. It is a fined technique based on in-sample fit to estimate the likelihood of a model to predict/estimate the future values. So, by applying this criterion, the minimum AIC value is selected to be the best model. In other words, a good model is the one that has minimum AIC among all the other models.

Akaike's Information Criterion is useful in selecting predictors of regression especially for determining the order of ARIMA model. In literature, this criterion is expressed by different formulas. Among them, the most widely used ones are:

1- based on likelihood measure

- First order, applicable for large sample size,

$$AIC = -2\log(L) + 2K$$

where

L: the likelihood (probability) which is a measure of model fit. The higher the number the better is the fit and it is usually obtained from statistical output.

K: number of free parameters in the model plus the intercept. $K = p + q + k + 1$

p: autoregression order,

q: moving average order,

c: constant of the model, $k = 1$ for $c \neq 0$ and $k = 0$ if $c = 0$.

- Second orders, applicable for small sample size ($n/K < 40$),

$$AIC_c = -2\log(L) + 2K * (2K(K+1) / (n - K - 1))$$

where

n: the sample size.

2- based on least square regression measure

- First order, applicable for large sample size,

$$AIC = -2\log(\hat{\sigma}^2) + 2K$$

where

$\hat{\sigma}^2$: residual sum of the squares/n

- Second orders, applicable for small sample size ($n/K < 40$)

$$AIC_c = -2\log(\hat{\sigma}^2) + 2K * (2K(K+1) / (n - K - 1))$$

3- based on Chi-squared measure

- First order, applicable for large sample size,

$$AIC = -2\log(\chi^2) + 2K$$

where

χ^2 : chi-squared regression value

- Second orders, applicable for small sample size ($n/K < 40$)

$$AIC_c = -2\log(\chi^2) + 2K * (2K(K+1) / (n - K - 1))$$

Note that for this formula, the estimated variance must be included in the parameter count.

4.8 Forecasting Criteria

The commonly used accuracy metrics to judge the forecasted values are:

1. Mean Error (ME)
2. Mean Absolute Error (MAE)
3. Mean Absolute Percentage Error (MAPE)
4. Root Mean Squared Error (RMSE)
5. Correlation between the Actual and the Forecast (Corr)
6. Min-Max Error (minmax)

Typically, if it is required to compare forecasts of two different series, the MAPE, Correlation and Min-Max Error can be used since only these three are the percentage errors that vary between 0 and 1.

For the selection of an ARIMA model that adequately describes the data series, the values of the following metrics are used:

a) Root Mean Squared Error (RMSE),

It implies the deviation between the computed and the measured values through their squares so as to overcoming the effect of the negative signs.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{comp} - x_{meas})_i^2}{n}}$$

b) Mean Absolute Deviation (MAD),

It is the deviation between the computed and the measured values. Mean absolute deviation helps to get a sense of how "spread out" the values.

$$MAD = \frac{1}{n} \sum_{i=1}^n |x_{comp} - x_{meas}|_i$$

c) Mean Absolute Percentage Error (MAPE)

It is a measure of prediction accuracy of a forecasting method. It usually expresses this accuracy as a ratio.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_{comp} - x_{meas}}{x_{meas}} \right|_i * 100$$

where,

Σ : a fancy symbol that means ‘sum’

n : sample size

x_{comp} : the computed data value

x_{meas} : the measured data value.

Hence, the appropriate ARIMA model is being selected among the smallest sum value of the weighted means calculated by the below given formulas:

i- $\text{Weighted Mean 1} = (\text{RMSE} + 4 \text{ MAPE} + \text{MAD}) / 6$

ii- $\text{Weighted Mean 2} = (4 \text{ RMSE} + \text{MAPE} + \text{MAD}) / 6$

iii- $\text{Weighted Mean 3} = (\text{RMSE} + \text{MAPE} + 4 \text{ MAD}) / 6$

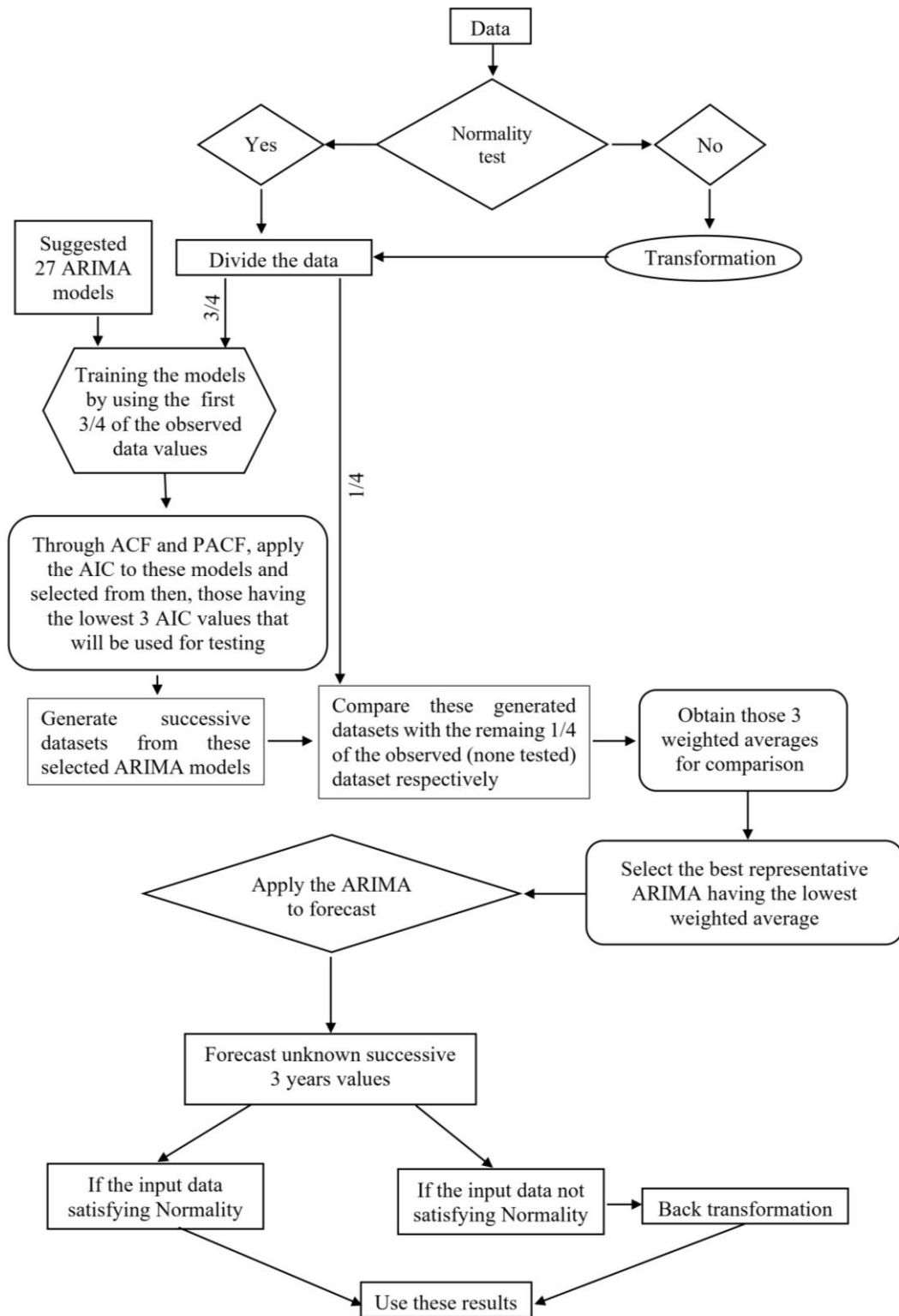


Figure 4. 17: Flowchart illustrating how to build ARIMA model and forecast the unknown successive data

Chapter 5

APPLIED TESTS, FUNCTIONS, AND MODELS

5.1 Quality Tests

To be able to propose or make robust sample inferences from any dataset under consideration, it is advisable to test statistically this sample inference by different test models. These test models are usually referred as quality tests and can both be parametric and non-parametric type.

5.1.1 Adequacy of Sample Size

The adequacy of sample size depends on,

- i- the coefficient of variance 'C_v' (the ratio of the existing samples standard deviation (s_x) to its arithmetic mean ($\bar{x}_{av}$),
 - ii- the expected (required) error (ϵ) for that study, like 10%, 15% etc.,
- and is given empirically as:

$$\text{Adequate size of the sample } n = \frac{C_v^2}{\epsilon^2} \quad (\text{Subramanya, K. 2013})$$

5.1.2 Homogeneity Test

Compares if two samples (populations) are from the same distribution type.

a- Parametric test

- t-test; to compare their means
- F-test; to compare their standard deviations

b- Non-parametric test

- χ^2 test for independence.

5.1.3 Consistency Test

Even if two datasets are derived from the same population, it could still be the case that, there are unaccounted differences between the two datasets, hence their consistency need to be tested.

a- Parametric test

- mean
- standard deviation
- skewness
- kurtosis

b- Non-parametric test

- Double Mass Curve ' R^2 '

5.1.4 Normality Test

Is used to determine whether the sample data has been drawn from a normally distributed population (within some tolerance).

a- Parametric test

- Jarqua-Bera test

b- Non-parametric test

- Kolmogorov-Smirnov test
- Shapiro-Wilk test

5.1.5 Trends Test

It is used to underlying the pattern of behavior within the time-series. It is also called Correlation or Randomness or Dependence.

a- Parametric test

- Moving average linear regression ' R^2 ' test

b- Non-parametric test

- Spearman's Rank 'ρ' test
- Mann-Kendall test with Theil-Sen trend line test.

5.1.6 Stationarity Test

A stationary time-series is one whose statistical properties such as mean, variance, auto-correlation, etc. are all constant over that time period.

a- Parametric test

- Dickey-Fuller test

b- Non-parametric test

- Runs test.

5.2 Models

5.2.1 Frequency Distribution Functions

The hydrologic records are short duration records. Statistical functions are used in order to obtain the maximum information from these short duration observations so as to evaluate the most probable nature of the corresponding populations (Usul, 2005). To predict a data, the standardized form of the very widely used frequency distribution functions are:

5.2.1.1 Normal Frequency Distribution Equation

$$x = \bar{x}_{av} + z s_x$$

where;

x: the required value,

$\bar{x}_{av}$: the arithmetic average of the values within the dataset,

s_x : the standard deviation of the values within the dataset,

z: z-score is measured in terms of standard deviations from the mean, given in Appendix 2.

5.2.1.2 Log-Normal Frequency Distribution Equation

$$\log(x) = \log(x)_{av} + z S_{\log(x)}$$

$\log(x)$: the logarithmic definition of the required value,

$\log(x)_{av}$: the arithmetic average of the dataset of which each data is defined by its logarithm,

$s_{\log(x)}$: the standard deviation of the dataset of which each data is defined by its logarithm,

z : z-score is measured in terms of standard deviations from the mean, given in Appendix 2.

5.2.1.3 Pearson Type III (Gamma) Frequency Distribution Equation

$$x = \bar{x}_{av} + K s_x$$

where,

x : the required value,

$\bar{x}_{av}$: the arithmetic average of the values within the dataset,

s_x : the standard deviation of the values within the dataset,

K : a value that is defined by the skewness of the given dataset, given in Appendix 6.

5.2.2 Time-Series Models

One of the popular and widely used statistical model for time-series analyzing and forecasting studies is the ARIMA model. In fact, it is a class of statistical models (Wang et al., 2014). In this study, ARIMA(p,d,q) model of different [(p = 0, 1, 2), (d = 0,1,2) and (q = 0,1,2)] combinations (i.e. $3^3 = 27$) were used.

5.3 Detecting a Trend

The Moving Average (MA) model is a parametric technique of smoothing and filtering the datasets. This method is used, since any of the time-series data may fluctuates in long run, and once the moving (running) average is used, simply these fluctuations are

get minimized and if there exists a trend then, it can easily be detected (Fukushima and Tanaka, 1990). In order for this model to be adopted, the time-series dataset was used where the linear lines of MA with different time windows from 2- up to 6- successive values based equations were generated separately for this study. Among them, for the representative MA order is determined from the slope of the line having the highest R^2 .

5.4 Detecting Wetness/Dryness

To detect wetness/dryness of the given dataset, the mean of the dataset is compared with each data and if the data is greater than its mean it is commented to be wet or as dry vice versa.

Chapter 6

CALCULATIONS AND RESULTS

6.1 Introduction

Yemen surface area is broadly grouped into 4 basins; Red Sea Basin (RSB), Arab Sea Basin (ASB), Gulf Aden Basin (GAB), Rub' Al Khali Basin (RKB), where in this study, 3 representative hypothetical rainfall stations were generated with the help of Thiessen polygon for each basin (RSB001, RSB002, RSB003, ASB004, ASB005, ASB006, GAB007, GAB008, GAB009, RKB010, RKB011, RKB012).

To satisfy the objectives of this study,

- 1- the five parametric and non-parametric data quality tests
 - a- Homogeneity,
 - b- Consistency,
 - c- Normality,
 - d- Trend, and
 - e- Stationarity were applied on the rainfall datasets from 1981 to 2018 (38 years' data) of each representative rainfall station and the three representative stations dataset based averaged rainfall dataset of each basin were determined.
- 2- one among the three frequency distribution equations
 - i- Normal,
 - ii- log-Normal, and
 - iii- Pearson Type III (Gamma) was suggested for each basin.

- 3- three among the ARIMA of 27 different combinations were selected to represent for each basin in which 3 successive years data values (2019, 2020 and 2021) were forecasted based on the most representative one from those three models.
- 4- detecting the Trend for each basin, using the most appropriate Moving Average (MA) model, and
- 5- commenting on the wet and the dry spells of each basin.

6.2 Red Sea Basin

6.2.1 Quality Tests Results of Rainfall Datasets of Red Sea Basin

Table 6. 1: Quality tests results of rainfall datasets of Red Sea basin

Sample size			Comment			
Available	Minimum required based on 10% error		Based on the other			
38	43		2 stations within the basin			
Quality Test	Type	Applied Test	RSB001	RSB002	RSB003	RSB-av.
Consistency	Parametric	Mean	46.5	215.7	215.7	135.0
		Std. Deviation	92.6	142.1	142.1	87.9
	Non-parametric	Double Mass Curve 'R ² '	0.962	0.998	0.996	0.985
		Comment	Accept	Accept	Accept	Accept
Homogeneity	Parametric	t-test	-6.45	0.50	4.74	9.39
		Comment	Reject	Accept	Reject	Reject
		F-test	0.16	1.13	4.87	9.23
		Comment	Accept	Accept	Reject	Reject
	Non-parametric	χ^2 -test	3.684	0.947	2.526	1.789
		Comment	> 1.2 Accept	> 1.0 Reject	> 1.0 Accept	> 1.1 Accept
Normality	Parametric	Jarque-Bera	0.06	0.06	0.06	0.06
		Comment	Accept	Accept	Accept	Accept
		Doornik Chi-Square	0.00	0.00	0.00	0.00
		Comment	Reject	Reject	Reject	Reject
	Non-parametric	Kolmogonov-Smirnov	0.023	0.595	0.293	0.580
		Comment	Accept	Accept	Accept	Accept
		Shapiro – Wilk	0.023	0.595	0.293	0.580
		Comment	Reject	Accept	Accept	Accept
Trend (Randomless) (Correlation) (Dependence)	Parametric	R ² (MA linear regression)	0.66	0.68	0.75	0.68
		Comment	Moderate	Moderate	Moderate	Moderate
	Non-parametric	ρ Spearman's Rank test	1.00	0.53	0.42	0.58
		Comment	V.Strong	Moderate	Moderate	Moderate
		Mann-Kendall with Theil-Sen trend line	1.00	0.39	0.31	0.51
		Comment	V.Strong	Weak	Weak	Moderate
Stationarity	Parametric	Dickey-Fuller	0.15	0.02	0.00	0.035
		Comment	Reject	Accept	Accept	Accept
	Non-parametric	Runs test	0.05	0.01	0.61	0.336
		Comment	Accept	Accept	Reject	Reject

6.2.2 Generated Equations from the widely Used Frequency Distribution Functions of Red Sea Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 2: Generated equations from the widely used frequency distribution functions

Name	Equation	p-value	Selected
Normal	$x = 135 + 87.9 z$	0.85	Normal
log-Normal	$\log(x) = 2.02 + 0.407z$	0.82	
Pearson Type-III	$x = 135 + 87.9 K$	0.22	

of red sea basin, based on the yearly averaged rainfall dataset from 1981 to 2018

6.2.3 Details of the Suggested ARIMA Models and Their AIC Scores for Red Sea Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 3: Details of the suggested ARIMA Models and their AIC scores for Red Sea Basin, based on the yearly averaged rainfall dataset from 1981 to 2018

Model Name		Test Results		AIC score	Comment
ARIMA	(p,q,r)	σ^2	m		
	(0,0,1)	8687.0	2	153.68	√
	(0,0,2)	8672.9	3	155.65	
	(0,1,1)	7629.1	3	153.53	√
	(0,1,2)	7756.5	4	155.81	
	(0,2,1)	12814.9	4	164.09	
	(0,2,2)	10513.9	5	162.83	
	(1,0,0)	8689.7	2	153.68	√
	(1,0,1)	8688.9	3	155.68	
	(1,0,2)	6951.9	4	154.00	
	(1,1,0)	11756.6	3	160.67	
	(1,1,1)	7632.0	4	155.54	
	(1,1,2)	6514.4	5	154.93	
	(1,2,0)	25951.9	4	175.74	

	(1,2,1)	11058.9	5	163.66	
	(2,0,0)	8689.8	3	155.68	
	(2,0,1)	8687.8	4	157.68	
	(2,0,2)	6666.9	5	155.31	
	(2,1,0)	10038.4	4	160.06	
	(2,1,1)	7672.0	5	157.63	
	(2,1,2)	6370.9	6	156.56	
	(2,2,0)	15637.14	5	169.38	
	(2,2,1)	10417	6	164.67	
	(2,2,2)	10309.14	7	166.50	

6.2.4 Building a Forecast Model by ARIMA for Red Sea Basin

Table 6. 4: Building a forecast model by ARIMA for Red Sea Basin

Years		2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Error Measures			Weighted Mean			Comment
Observed (mm)		82	202	92	142	120	167	74	102	42	1	RMSE	MAPE	MAD	1	2	3	
ARIMA Model	(0,1,1)	103.3	99.9	96.4	92.9	89.5	86	82.5	79.1	75.6	72.1	52.5	741.8	42.5	510.38	165.73	160.70	√
	(0,0,1)	139	147	147	147	147	147	147	147	147	147	70.9	1519.9	58.8	1034.86	310.36	304.35	
	(1,0,1)	139	146	147	147	147	147	147	147	147	147	71.0	1520.7	59.0	1035.44	310.59	304.59	

6.2.5 Detecting Appropriate MA Time Window with its Linear Trend Equation of Red Sea Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 5: Detecting appropriate MA time window with its linear trend equation of Red Sea Basin based on the yearly averaged rainfall dataset from 1981 to 2018

Moving Average (MA)	Time window					Comment
	2	3	4	5	6	
Linear Trend Equation	$y = -2.87x + 5875.1$	$y = -2.61x + 5345.2$	$y = -2.40x + 4933.2$	$y = -2.30x + 4730.4$	$y = -2.38x + 4894.9$	R ² = 0.37 selected MA(6)
R ²	0.22	0.26	0.27	0.31	0.37	

6.2.6 Detecting Wetness/Dryness of Red Sea Basin Based on the Yearly Averaged

Rainfall Dataset from 1981 to 2018

Table 6. 6: Detecting wetness/dryness of Red Sea Basin based on the yearly averaged rainfall dataset from 1981 to 2018.

The mean value of the yearly averaged rainfall dataset from 1981 to 2018 of Red Sea Basin, that is used for the comparison is 135 mm										
Year	Rainfall (mm)	Wet\Dry		Year	Rainfall (mm)	Wet\ Dry		Year	Rainfall (mm)	Wet\ Dry
1981	297	Wet		1994	187	wet		2007	122	dry
1982	213	Wet		1995	200	wet		2008	86	dry
1983	229	Wet		1996	150	wet		2009	82	dry
1984	105	Dry		1997	201	wet		2010	202	wet
1985	62	Dry		1998	139	wet		2011	92	dry
1986	523	Wet		1999	66	dry		2012	142	wet
1987	174	Wet		2000	91	dry		2013	120	dry
1988	150	Wet		2001	74	dry		2014	167	wet
1989	154	Wet		2002	131	dry		2015	74	dry
1990	37	Dry		2003	108	dry		2016	102	dry
1991	39	Dry		2004	107	dry		2017	42	dry
1992	72	Dry		2005	164	wet		2018	1	dry
1993	90	Dry		2006	136	wet				

Wetness: total number above the mean = 17	Dry Spell
Dryness: total number below the mean = 21	

6.3 Arab Sea Basin

6.3.1 Quality Tests Results of Rainfall Datasets of Arab Sea Basin

Table 6. 7: Quality tests results of rainfall datasets of Arab Sea Basin

Sample size			Comment			
Available	Minimum required		Based on the other			
38	62		2 stations within the basin			basins
Quality Test	Type	Applied Test	ASB004	ASB005	ASB006	ASB-av.
Consistency	Parametric	Mean	41.9	24.9	20.4	29.1
		Std. Deviation	40.6	18.0	22.2	22.8
	Non-parametric	Double Mass Curve 'R ² '	0.999	0.998	0.983	0.969
		Comment	Accept	Accept	Accept	Accept
Homogeneity	Parametric	t-test	5.52	3.48	4.51	1.51
		Comment	Reject	Reject	Reject	Accept
		F-test	8.37	0.25	0.44	1.43
		Comment	Reject	Accept	Accept	Accept
	Non-parametric	χ^2 -test	1.789	3.684	0.947	2.526
		Comment	> 1.1 Accept	> 1.2 Accept	> 1.0 Reject	> 1.1 Accept
Normality	Parametric	Jarque-Bera	0.058	0.058	0.058	0.058
		Comment	Accept	Accept	Accept	Accept
		Doornik Chi-Square	0.004	0.124	0.006	0.002
		Comment	Accept	Accept	Accept	Accept
	Non-parametric	Kolmogonov-Smirnov	0.004	0.124	0.006	0.002
		Comment	Accept	Accept	Accept	Accept
		Shapiro – Wilk	0.00	0.00	0.00	0.00
		Comment	Reject	Reject	Reject	Reject
Trend (Randomless) (Correlation) (Dependence)	Parametric	R ² (MA linear regression)	0.64	0.66	0.98	0.604
		Comment	Moderate	Moderate	V.Strong	Moderate
	Non-parametric	ρ Spearman's Rank test	0.288	0.464	0.441	0.398
		Comment	Weak	Moderate	Moderate	Moderate
		Mann-Kendall with Theil-Sen trend line	0.204	0.327	0.303	0.278
		Comment	Weak	Weak	Weak	Weak
Stationarity	Parametric	Dickey-Fuller	0.102	0.044	0.035	0.060
		Comment	Reject	Accept	Accept	Accept
	Non-parametric	Runs test	0.156	0.142	0.118	0.139
		Comment	Accept	Accept	Accept	Accept

6.3.2 Generated Equations from the Widely used Frequency Distribution Functions of Arab Sea Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 8: Generated equations from the widely used frequency distribution functions of Arab Sea Basin, based on the yearly averaged rainfall dataset from 1981 to 2018

Name	Equation	p-value	Selected
Normal	$x = 29.1 + 22.9 z$	0.244	Normal
log-Normal	$\log(x) = 1.342 + 0.358 z$	0.1923	
Pearson Type - III	$x = 29.1 + 22.9 K$	0.203	

6.3.3 Details of the Suggested ARIMA Models and Their AIC Scores for Arab Sea Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 9: Details of the suggested ARIMA models and their AIC scores for Arab Sea Basin, based on the yearly averaged rainfall dataset from 1981 to 2018.

Model Name		Test Results		AIC score	Comment
	(p,q,r)	σ^2	m		
ARIMA	(0,0,1)	4000.43	2	140.88	✓
	(0,0,2)	3990.61	3	142.84	
	(0,1,0)	3990.61	2	140.84	✓
	(0,1,1)	3561.44	3	140.96	
	(0,1,2)	3603.93	4	143.16	
	(0,2,0)	3603.93	3	141.16	
	(0,2,1)	6246.04	4	152.23	
	(0,2,2)	4593.43	5	149.16	
	(1,0,0)	4002.71	2	140.89	✓
	(1,0,1)	3988.79	3	142.83	
	(1,0,2)	3328.9	4	141.85	
	(1,1,0)	5497.64	3	148.13	
	(1,1,1)	3564.3	4	142.98	
	(1,1,2)	3173.56	5	143.06	
	(1,2,0)	12737.4	4	163.99	
	(1,2,1)	5282.14	5	151.47	
	(1,2,2)	6419.57	6	156.69	
	(2,0,0)	4002.79	3	142.89	
	(2,0,1)	3987.82	4	144.83	
	(2,0,2)	3729.54	5	145.72	
	(2,1,0)	4390.43	4	146.42	
	(2,1,1)	3549.02	5	144.90	
	(2,1,2)	3422.45	6	146.30	
	(2,2,0)	7384.82	5	157.00	
	(2,2,1)	4544.68	6	150.99	
	(2,2,2)	4561.32	7	153.05	

6.3.4 Building a Forecast Model by ARIMA for Arab Sea Basin

Table 6. 10: Building a forecast model by ARIMA for Arab Sea Basin

Years		2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Error Measures			Weighted Mean			Comment
Observed (mm)		51	126	64	110	75	110	50	75	36	1	RMSE	MAPE	MAD	1	2	3	
ARIMA MODEL	(0,0,1)	99	103	103	103	103	103	103	103	103	103	48	1071	40	729	217	213	
	(1,0,0)	99	102	103	103	103	103	103	103	103	103	48	1071	40	729	218	213	
	(0,1,1)	69	67	64	61	59	56	54	51	49	46	35	477	28	328	107	104	√

6.3.5 Detecting Appropriate MA Time Window with its Linear Trend Equation of Arab Sea Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 11: Detecting appropriate MA time window with its linear trend equation of Arab Sea Basin based on the yearly averaged rainfall dataset from 1981 to 2018

Moving Average (MA)	Time window					Comment
Linear Trend Equation	2	3	4	5	6	
	$y = -1.26x + 2556.1$	$y = -1.16x + 2358.8$	$y = -1.08x + 2197.7$	$y = -1.04x + 2113.6$	$y = -1.03x + 2085.6$	
	R^2	0.47	0.54	0.60	0.60	0.59

6.3.6 Detecting Wetness/Dryness of Arab Sea Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 12: Detecting Wetness/Dryness of Arab Sea Basin based on the yearly averaged rainfall dataset from 1981 to 2018

The mean value of the yearly averaged rainfall dataset from 1981 to 2018 of Arab Sea Basin, that is used for the comparison is 29.1 mm										
Year	Rainfall (mm)	Wet\Dry		Year	Rainfall (mm)	Wet\ Dry		Year	Rainfall (mm)	Wet\ Dry
1981	57	Wet		1994	29	dry		2007	19	dry
1982	111	Wet		1995	25	dry		2008	14	dry
1983	82	Wet		1996	37	wet		2009	8	dry
1984	29	Dry		1997	16	dry		2010	24	dry
1985	25	Dry		1998	32	wet		2011	21	dry
1986	86	Wet		1999	6	dry		2012	32	wet
1987	67	Wet		2000	13	dry		2013	24	dry
1988	19	Dry		2001	16	dry		2014	14	dry
1989	33	Wet		2002	32	wet		2015	10	dry
1990	43	Wet		2003	24	dry		2016	26	dry
1991	8	Dry		2004	29	dry		2017	8	dry
1992	23	Dry		2005	22	dry		2018	1	dry
1993	21	Dry		2006	21	dry				

Wetness: total number above the mean = 11	Dry Spell
Dryness: total number below the mean = 27	

6.4 Gulf Aden Basin

6.4.1 Quality Tests Results of Rainfall Datasets of Gulf Aden Basin

Table 6. 13: Quality tests of rainfall datasets of Gulf Aden Basin

Sample size			Comment			
Available	Minimum required		Based on the other			
38	65		2 stations within the basin		basins	
Quality Test	Type	Applied Test	GAB007	GAB008	GAB009	GAB-av.
Consistency	Parametric	Mean	74.0	24.0	24.0	49
		Std. Deviation	46.4	24.1	24.1	39.5
	Non-parametric	Double Mass Curve 'R ² '	0.999	0.998	0.983	0.958
		Comment	Accept	Accept	Accept	Accept
Homogeneity	Parametric	t-test	3.62	0.06	4.26	1.41
		Comment	Reject	Accept	Reject	Reject
		F-test	2.69	1.54	0.23	1.72
		Comment	Accept	Accept	Accept	Reject
	Non-parametric	χ^2 -test	1.789	3.684	0.947	2.526
		Comment	> 1.1 Accept	> 1.2 Accept	< 1.0 Reject	> 1.1 Accept
Normality	Parametric	Jarque-Bera	0.060	0.060	0.060	0.060
		Comment	Accept	Accept	Accept	Accept
		Doornik Chi-Square	0.238	0.238	0.238	0.238
		Comment	Accept	Accept	Accept	Accept
	Non-parametric	Kolmogonov-Smirnov	0.002	0.012	0.006	0.001
		Comment	Normal	Normal	Normal	Normal
		Shapiro – Wilk	0.00	0.00	0.00	0.00
		Comment	Reject	Reject	Reject	Reject
Trend (Randomness) (Correlation) (Dependence)	Parametric	R ² (MA linear regression)	0.865	0.963	0.598	0.809
		Comment	Strong	Strong	Moderate	Strong
	Non-parametric	ρ Spearman's Rank test	0.805	0.918	0.427	0.717
		Comment	Strong	Very Strong	Moderate	Strong
		Mann-Kendall with Theil-Sen trend line	0.655	0.790	0.312	0.586
		Comment	Moderate	Moderate	Weak	Moderate
Stationarity	Parametric	Dickey-Fuller	0.137	0.05	0.021	0.069
		Comment	Reject	Accept	Accept	Accept
	Non-parametric	Runs test	0.642	0.156	0.772	0.523
		Comment	Accept	Accept	Accept	Accept

6.4.2 Generated Equations from the Widely used Frequency Distribution Functions of Gulf Aden Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 14: Generated equations from the widely used frequency distribution functions of Gulf Aden Basin, based on the yearly averaged rainfall dataset from 1981 to 2018

Name	Equation	p-value	Selected
Normal	$x = 49 + 39.5 z$	0.544	Pearson Type-III
log-Normal	$\log(x) = 1.563 + 0.37 z$	0.414	
Pearson Type-III	$x = 49 + 39.5 K$	0.545	

6.4.3 Details of the Suggested ARIMA Models and their AIC Scores for Gulf Aden Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 15: Details of the suggested ARIMA models and their AIC scores for Gulf Aden Basin, based on the yearly averaged rainfall dataset from 1981 to 2018

Model Name		Test Results		AIC score	Comment
	(p,q,r)	σ^2	m		
ARIMA	(0,0,1)	1560.26	2	125.341	
	(0,0,2)	1415.92	3	125.739	
	(0,1,1)	1542.85	3	127.156	
	(0,1,2)	1159.1	4	124.437	
	(0,2,1)	1515.2	4	128.858	
	(0,2,2)	1240.87	5	127.562	
	(1,0,0)	1653.39	2	126.298	
	(1,0,1)	1516.21	3	126.869	
	(1,0,2)	965.393	4	121.419	√
	(1,1,0)	2154.39	3	132.666	

	(1,1,1)	1512.3	4	128.826	
	(1,1,2)	1129.61	5	126.011	
	(1,2,0)	4090.93	4	145.249	
	(1,2,1)	1284.14	5	128.127	
	(1,2,2)	1645.63	6	134.221	
	(2,0,0)	1091.66	3	121.447	√
	(2,0,1)	1323.29	4	126.623	
	(2,0,2)	1081.12	5	125.287	
	(2,1,0)	1442.47	4	128.046	
	(2,1,1)	2151.9	5	136.647	
	(2,1,2)	856.429	6	123.442	√
	(2,2,0)	2549.37	5	139.444	
	(2,2,1)	1172.1	6	128.621	
	(2,2,2)	1093.9	7	129.481	

6.4.4 Building a Forecast Model by ARIMA for Gulf Aden Basin

Table 6. 16: Building a forecast model by ARIMA for Gulf Aden Basin

Years		2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Error Measures			Weighted Mean			Comment
Observed (mm)		32	66	28	20	30	39	15	30	12	1	RMSE	MAPE	MAD	1	2	3	
ARIMA MODEL	(1,0,2)	50	73	68	64	61	59	57	56	55	54	658.0	32.1	449.8	138.3	136.9	658.0	
	(2,0,0)	50	74	64	42	42	58	65	55	46	50	600.2	27.3	409.8	125.0	123.3	600.2	
	(2,1,2)	15	24	28	19	13	15	15	11	6	6	81.9	13.1	59.9	28.0	25.5	81.9	√

6.4.5 Detecting Appropriate MA Time Window with its Linear Trend Equation of Gulf Aden Basin Based on the Yearly Averaged Rainfall

Dataset from 1981 to 2018

Table 6. 17: Detecting appropriate MA time window with its linear trend equation of Gulf Aden Basin based on the yearly averaged rainfall dataset from 1981 to 2018

Moving Average (MA)	Time window					Comment
	2	3	4	5	6	
Linear Trend Equation	$y = -1.83x + 3716.7$	$y = -1.55x + 3147.9$	$y = -1.36x + 2778.4$	$y = -1.28x + 2604.6$	$y = -1.26x + 2579.1$	$R^2 = 0.70$ selected MA(6)
R^2	0.36	0.45	0.60	0.65	0.70	

6.4.6 Detecting Wetness/Dryness of Gulf Aden Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 18: Detecting Wetness/Dryness of Gulf Aden Basin based on the yearly averaged rainfall dataset from 1981 to 2018.

The mean value of the yearly averaged rainfall dataset from 1981 to 2018 of Gulf Aden Basin, that is used for the comparison is 50.40 mm

Year	Rainfall (mm)	Wet\Dry		Year	Rainfall (mm)	Wet\ Dry		Year	Rainfall (mm)	Wet\ Dry
1981	96	Wet		1994	37	dry		2007	44	dry
1982	236	Wet		1995	68	wet		2008	25	dry
1983	94	Wet		1996	98	wet		2009	32	dry
1984	17	Dry		1997	46	dry		2010	66	wet
1985	15	Dry		1998	43	dry		2011	28	dry
1986	88	Wet		1999	42	dry		2012	20	dry
1987	60	Wet		2000	42	dry		2013	30	dry
1988	48	Dry		2001	31	dry		2014	39	dry
1989	90	Wet		2002	69	wet		2015	15	dry
1990	32	Dry		2003	18	dry		2016	30	dry
1991	22	Dry		2004	46	dry		2017	12	dry
1992	61	Wet		2005	67	wet		2018	1	dry
1993	35	Dry		2006	35	dry				

Wetness: total number above the mean=12	Dry Spell
Dryness: total number below the mean= 26	

6.5 Rub' Al Khali Basin

6.5.1 Quality Tests Results of Rainfall Datasets of Rub' Al Khali Basin

Table 6. 19: Quality tests of rainfall datasets of Rub' Al Khali Basin

Sample size			Comment			
Available	Minimum required		Based on the other			
38	59		2 stations within the basin		basins	
Quality Test	Type	Applied Test	RKB010	RKB011	RKB012	RKB-av.
Consistency	Parametric	Mean	38.5	30.6	21.2	30.1
		Std. Deviation	44.9	20.8	22.0	23.1
	Non-parametric	Double Mass Curve 'R ² '	0.945	0.940	0.972	0.952
		Comment	Accept	Accept	Accept	Accept
Homogeneity	Parametric	t-test	1.61	0.10	-2.34	-0.78
		Comment	Reject	Accept	Reject	Accept
		F-test	5.85	0.51	0.68	1.23
		Comment	Reject	Accept	Accept	Accept
	Non-parametric	χ^2 -test	5.789	6.263	7.684	2.526
		Comment	> 1.2 Accept	> 1.3 Accept	> 1.4 Accept	> 1.1 Accept
Normality	Parametric	Jarque-Bera	0.060	0.060	0.060	0.060
		Comment	Accept	Accept	Accept	Accept
		Doornik Chi-Square	0.238	0.238	0.238	0.238
		Comment	Accept	Accept	Accept	Accept
	Non-parametric	Kolmogonov-Smirnov	0.017	0.630	0.261	0.114
		Comment	Accept	Accept	Accept	Accept
		Shapiro – Wilk	0.00	0.00	0.00	0.00
		Comment	Reject	Reject	Reject	Reject
Trend (Randomless) (Correlation) (Dependence)	Parametric	R ² (MA linear regression)	0.45	0.36	0.60	0.47
		Comment	Moderate	Weak	Moderate	Moderate
	Non-parametric	ρ Spearman's Rank test	0.421	0.242	0.513	0.39
		Comment	Moderate	Weak	Moderate	Weak
		Mann-Kendall with Theil-Sen trend line	0.292	0.166	0.370	0.28
		Comment	Weak	Weak	Weak	Weak
Stationarity	Parametric	Dickey-Fuller	0.188	0.043	0.049	0.09
		Comment	Reject	Accept	Accept	Accept
	Non-parametric	Runs test	0.072	0.411	0.250	0.24
		Comment	Reject	Accept	Accept	Accept

6.5.2 Generated Equations from the Widely used Frequency Distribution Functions of Rub' Al Khali Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 20: Generated equations from the widely used frequency distribution functions of Rub' Al Khali Basin, based on the yearly averaged rainfall dataset from 1981 to 2018

Name	Equation	p-value	Selected
Normal	$x = 30.1 + 23.1 z$	0.26	Pearson Type-III
log-Normal	$\log(x) = 1.3665 + 0.3258 z$	0.19	
Pearson Type-III	$x = 30.1 + 23.1 K$	0.65	

6.5.3 Details of the Suggested ARIMA Models and their AIC Scores for Rub' Al Khali Basin, Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 21: Details of the suggested ARIMA models and their AIC scores for Rub' Al Khali Basin, based on the yearly averaged rainfall dataset from 1981 to 2018

Model Name		Test Results		AIC score	Comment
	(p,q,r)	σ^2	m		
ARIMA	(0,0,1)	483.29	2	106.00	
	(0,0,2)	474.23	3	107.69	
	(0,1,1)	383.55	3	104.19	✓
	(0,1,2)	309.02	4	102.62	✓
	(0,2,1)	623.70	4	114.21	
	(0,2,2)	555.96	5	114.31	
	(1,0,0)	464.94	2	105.36	✓
	(1,0,1)	437.97	3	106.37	
	(1,0,2)	428.51	4	108.01	
	(1,1,0)	550.05	3	110.14	
	(1,1,1)	375.29	4	105.83	

	(1,1,2)	354.39	5	106.88	
	(1,2,0)	1213.59	4	125.19	
	(1,2,1)	494.69	5	112.38	
	(1,2,2)	517.23	6	115.12	
	(2,0,0)	463.90	3	107.32	
	(2,0,2)	423.34	5	109.81	
	(2,1,0)	403.31	4	107.01	
	(2,1,1)	530.24	5	113.53	
	(2,1,2)	342.31	6	108.31	
	(2,2,0)	771.60	5	119.72	
	(2,2,1)	361.34	6	109.20	
	(2,2,2)	289.91	7	107.57	

6.5.4 Building a Forecast Model by ARIMA for Rub' Al Khali Basin

Table 6. 22: Building a forecast model by ARIMA for Ruba' Al Khali Basin

Years		2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Error Measures			Weighted Mean			Comment
Observed (mm)		13	22	24	23	21	7	10	36	14	2	RMSE	MAPE	MAD	1	2	3	
ARIMA MODEL	(0,1,2)	17	16	15	14	13	11	10	9	8	7	10.4	60.2	7.9	43.1	18.3	17.0	√
	(0,1,1)	10	8	6	4	2	0	0	0	0	0	17.1	84.4	14.4	61.5	27.9	26.5	
	(1,0,0)	29	33	35	36	36	36	37	37	37	37	20.4	290.7	17.9	200.2	65.0	63.8	

6.5.5 Detecting Appropriate MA Time Window with its Linear Trend Equation of Rub' Al Khali Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 23: Detecting appropriate MA time window with its linear trend equation of Rub' Al Khali Basin based on the yearly averaged rainfall dataset from 1981 to 2018

Moving Average (MA)	Time window					Comment
	2	3	4	5	6	
Linear Trend Equation	$y = -1.29x + 2619.2$	$y = -1.26x + 2551.4$	$y = -1.19x + 2417.6$	$y = -1.16x + 2357$	$y = -1.15x + 2325.2$	
R ²	0.49	0.57	0.62	0.63	0.64	R ² = 0.64 selected MA(6)

6.5.6 Detecting Wetness/Dryness of Rub' Al Khali Basin Based on the Yearly Averaged Rainfall Dataset from 1981 to 2018

Table 6. 24: Detecting Wetness/Dryness of Rub' Al Khali Basin based on the yearly averaged rainfall dataset from 1981 to 2018

The mean value of the yearly averaged rainfall dataset from 1981 to 2018 of Rub' Al Khali Basin, that is used for the comparison is 30.1 mm										
Year	Rainfall (mm)	Wet\Dry		Year	Rainfall (mm)	Wet\ Dry		Year	Rainfall (mm)	Wet\ Dry
1981	67	Wet		1994	30	Wet		2007	25	Dry
1982	73	Wet		1995	35	Wet		2008	17	Dry
1983	110	Wet		1996	33	Wet		2009	12	Dry
1984	32	Wet		1997	31	Dry		2010	21	Dry
1985	34	Wet		1998	22	Dry		2011	23	Dry
1986	83	Wet		1999	8	Dry		2012	23	Dry
1987	81	Wet		2000	11	Dry		2013	20	Dry
1988	10	Dry		2001	12	Dry		2014	8	Dry
1989	32	Wet		2002	30	Wet		2015	11	Dry
1990	23	Dry		2003	21	Dry		2016	31	Wet
1991	12	Dry		2004	38	Wet		2017	15	Dry
1992	40	Wet		2005	19	Dry		2018	2	Dry
1993	29	Dry		2006	20	Dry				

Wetness: total number above the mean=15	Dry Spell
Dryness: total number below the mean= 23	

Chapter 7

CONCLUSION AND RECOMMENDATIONS

7.1 Conclusion

In this study, Yemen as a whole is divided into four main geographical regions, with the help of the Thiessen polygon approach, where three representative stations for each region were determined hypothetically. For these stations 38-years (1981-2018) monthly rainfall data were gathered from POWER, 2019. Based on those three relevant stations, the arithmetic averages were used for establishing the yearly representative rainfall data-set for each region (basin).

- Based on five statistical data quality tests; the homogeneity, the consistency, the normality, the trend and the stationarity; each basin dataset was tested and all gave acceptable results.
- Among the three frequency distribution equations; Normal, log-Normal, and Pearson Type III, the best representative one was determined through the curve fitting approach for each basin.
- To predict oncoming 3-years, 2019, 2020, and 2021 rainfall value for each basin, 27 versions of ARIMA models were used. Dataset from 1981 to 2008 were used to train the models, and the last 10 successive years' data values from 2009 to 2018 to test them. Among these 27 ARIMA models, those three models having the lowest AIC values were selected and tested through the appropriate statistical measures (RMSE, MAPE, and MAD) and based on their

lowest Weighted Means 1, 2, and 3), the appropriate ARIMA model was determined for each basin and the forecasted values were generated and tabulated in Table 7.1.

Table 7. 1: Forecasted rainfall values of 2019, 2020, and 2021 of each meteorological district based on the relevant most representative time series models

Regions	Models	Years		
		2019	2020	2021
Red Sea Basin	ARIMA(0,1,1)	63.3	59.8	56.3
Arab Sea Basin	ARIMA(0,1,1)	43	40	38
Gulf Aden Basin	ARIMA(2,1,2)	16	22	14
Rub' Al Khali Basin	ARIMA(0,1,2)	4.2	4.1	2.7

- From the Moving Average approach test with different time windows, it was determined that, for all the basins the slope of the linear line is –ve , implying a rainfall decrease of nearly 2.4 mm/year, 1.0 mm/year, 1.3 mm/year, and 1.2 mm/year for the Red Sea Basin, the Arabian Sea Basin, the Gulf Aden Basin and the Rub' Al Khali Basin, respectively.
- For all four basins, their annual rainfall datasets were studied based on the comparison of each single data with the average value of that dataset and commented. In fact, it was determined that, all the basins were under severe dryness.

Table 7. 2: The Synopsis of the yearly averaged rainfall data test results with the suggested frequency distribution and time-series models for each meteorological district

Basins (Districts)	Quality Check Test					Suggested			Trend		Spell
						Frequency Distribution		Model			
	Homogeneity	Normality	Consistency	Trend	Stationarity	Name	Equation	ARIMA	Name	Equation	
Red Sea	Accept	Accept	Accept	Yes	Yes	Normal	$x = 135 + 87.9 z$	(0,1,1)	MA(6)	$y = -2.38x + 4894.9$	68% Dry
Arab Sea	Accept	Accept	Accept	Yes	No	Normal	$x = 29.1 + 22.9 z$	(0,1,1)	MA(5)	$y = -1.04x + 2113.6$	68% Dry
Gulf Aden	Accept	Accept	Accept	Yes	Yes	Pearson Type-III	$x = 49.0 + 39.5 K$	(2,1,2)	MA(6)	$y = -1.26x + 2579.1$	71% Dry
Rub' Al Khali	Accept	Accept	Accept	No	Yes	Pearson Type-III	$x = 30.1 + 23.1 K$	(0,1,2)	MA(6)	$y = -1.15x + 2325.2$	61% Dry

7.2 Recommendation

- Further investigation on rain harvesting with an appropriate technique might be a good start for overcoming the existing water shortage coupled with rapid increase in population.
- If there is a potential analysis in this field, it is better to take each basin separately by increasing the number of stations.
- Attention to integrated water resource management in Yemen and focus on it is very important to mitigate the effects of future challenges represented by the water imbalance between consumption and demand, and annual decrease rainfall.

REFERENCES

- Adamowski, J., H. Fung Chan, S. O., Prasher, B., Ozga-Zielinski, & Sliusarieva, A. (2012). Comparison of multiple linear and nonlinear regression, autoregressive integrated moving average, artificial neural network, and wavelet artificial neural network methods for urban water demand forecasting in Montreal, Canada, *Water Resources Res.*, 48, W01528, doi:10.1029/2010WR009945.
- Al-Falahi, A., Saddique, N., Spank, U., Gebrechorkos, H., & Bernhofe, C. (2020). Evaluation the Performance of Several Gridded Precipitation Products over the Highland Region of Yemen for Water Resources Management. *Article. Tharandt*. doi:10.3390/rs12182984.
- Al-Nozaily, F., Radhwan, F., & Salah, A. (2006). Evaluation of Sana'a wastewater treatment plant. *Journal of Science & Technology*, Vol. (11) No.(2).
- Al-Tawqi, A. (2018). *Reuse of wastewater from Sanaa WWTP through planting*. Water and Environment Centre, Science. www. Wec.edu.ye.
- Amjadi, M. (2015). *Statistic and Probabilistic Variations and Trends of Precipitation Data of TRNC*. MSc thesis. Gazimağusa. EMU.
- Batainah, A. (2019). *Statistic and Probabilistic Variations with Precipitation Predictions of Jordan*. MSc thesis. Gazimağusa. EMU.

Burlando, P. (2011). *Hydrology-Collection of formula*. Swiss Federal Institute of Technology. Zurich.

Daniel, E., & Fola, J. (2015). Seasonal ARIMA Modeling and Forecasting of Rainfall in Warri Town, Nigeria. *FUPR*. <http://dx.doi.org/10.4236/gep.2015.36015>.

ECMWF, *European Centre for Medium-Range Weather Forecasts*. (2019). Retrieved from [https://www.ecmwf.int/External geophysics, climate and environment](https://www.ecmwf.int/External/geophysics%2C%20climate%20and%20environment) (Climate)

Farquharson, F. A. K., Plinston, D. T., & Sutcliffe, J. V. (1996). Rainfall and runoff in Yemen, *Hydrological Sciences Journal*, 41:5, 797-811, doi: 10.1080/02626669609491546.

Ghatfan, A., Hayek, S., & Hamdan, N. (2016). Building a forecasting model for annual rainfall in Husn Suleiman station using mathematical modeling. *Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series* Vol. (38) No. (4) 2016.

Glass, N. (2010). The water crisis in Yemen: Causes, consequences and solutions. *Global Majority E-Journal*, Vol. 1, No. 1 (June 2010), pp. 17-30

Gun, A., & Ahmed, A. (1996). *The water resources of the Yemen: a summary digest of the available information*. (Report WRAY-35.) vii, [233] pp., 4 foldout maps. Sana'a, Yemen: Ministry of Oil and Mineral Resources; Institute of Applied Geoscience, Delft, Netherlands.

- Hayek, S., & Hammad, M.(2016). Predict of the monthly water volumes incoming in Al-Roos River in the Syrian Coast by using the time-series analysis. *Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series Vol. (38) No. (5) 2016.*
- Hyndman, R., & George, A. (2014). *Forecasting Principles and Practice*. ISBN 978-0-9875071-0-5
- Majed, A. A. W. & Wan Ruslan, I. (2019). Precipitation Analysis and Water Resource of Wadi Siham Basin, Yemen. *GEOGRAFI*, 7(2), 36-63.
- Moges, B., Dustin, G., Amy, S., Emily B., Daniel, R., Anthony, B. & Zachary, M. (2020). *Comparison of short-term streamflow forecasting using stochastic time-series, neural networks, process-based, and Bayesian models*. Department of Biological Systems Engineering, Virginia Tech, Blacksburg, VA, USA, doi.org/10.1016/j.envsoft.2020.104669.
- Nyatuame M., & Agodzo S.K. (2018). Stochastic ARIMA model for annual rainfall and maximum temperature forecasting over Tordzie watershed in Ghana. *Journal of Water and Land Development*. No. 37 p. 127–140. DOI: 10.2478/jwld-2018-0032.
- POWER, *Prediction of Worldwide Energy Resources*. (2019). Retrieved from <https://power.larc.nasa.gov/>

RRDP, Rural Road Development Program. (2009). *Road Sector Environmental Assesment in Yemen*. Report. MPOR. Sana'a.

Shbari, M., Hussain, M., & Sourav, R. (2015). Forecasting Monthly Precipitation in Sylhet City Using ARIMA Model. Sylhet. *Civil and Environmental research*. Vol. 7 No. 1.

Shiban,L., Alasaad, A., & Abdulrahman, A. (2019). Building a Forecasting Model for Annual Rainfall in Alhwaiz Basin Using Box-Jenkins Methodology. *Tishreen University. Lattakia. American Journal of Innovative Research and Applied Sciences*. ISSN 2429-5396 I.

Subramanya, K. 2013. (2013). *Engineering Hydrology*. Tata McGraw-Hill. New Delhi

Surajit, C. & Goutami, C. (2010). *Univariate modelling of summer-monsoon rainfall time-series: Comparison between ARIMA and ARNN*. Department of Computer Application, Pailan College of Management and Technology, 700 104 Kolkata, India. doi:10.1016/j.crte.2009.10.016.

Usul, N. *Engineering Hydrology*. Metu press, 2005.

Yehya, Q. & Al-Asbahi, A. (2005). *Water Resources Information in Yemen*. IWG-Env. Vienna: *International Work Session on Water Statistics*.

Yuchuan, L., & Dzombak, A. (2020). *Use of the Autoregressive Integrated Moving Average (ARIMA) Model to Forecast Near-Term Regional Temperature and*

Precipitation. Carnegie Mellon University, Pittsburgh, Pennsylvania. doi:
10.1175/WAF-D-19-0158.1 .

APPENDICES

Appendix1: Rainfall Data Details and Simple Statistical Measures of the Basins

1.1 Red Sea Basin

1.1.1 Red Sea Basin, Monthly Rainfall Data of the Representative Station RSB001 from 1981 to 2018 in (mm)

Table 1: Red Sea Basin, Monthly Rainfall Data of the Representative Station RSB001 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	0	0	40	3	28	0	3	14	1	1	0	0	90
1982	60	9	39	11	0	0	0	8	3	1	7	7	145
1983	2	27	7	18	9	10	9	8	1	8	4	3	106
1984	3	2	3	1	8	2	0	8	0	0	0	5	32
1985	3	1	0	2	12	1	11	5	1	2	0	4	42
1986	0	1	16	22	1	47	12	46	29	18	10	14	216
1987	88	3	16	11	28	14	5	14	0	1	0	0	180
1988	39	0	0	1	1	0	1	4	1	0	0	0	47
1989	0	7	5	13	0	1	0	2	0	0	0	3	31
1990	0	4	1	9	0	0	10	5	0	0	0	0	29
1991	2	4	7	0	0	0	0	0	0	0	0	0	13
1992	3	0	0	1	0	0	1	10	0	1	1	1	18
1993	0	1	0	2	6	0	0	1	0	0	0	0	10
1994	0	0	13	0	0	0	4	2	0	0	0	0	19
1995	0	0	1	0	0	0	0	3	0	1	0	0	5
1996	0	0	0	0	2	14	1	7	4	0	0	0	28
1997	1	0	2	0	14	1	0	0	3	4	5	0	30
1998	0	7	1	0	2	1	1	21	0	0	0	0	33
1999	0	0	1	0	0	0	1	2	0	4	0	1	9
2000	0	0	0	0	0	0	5	1	2	24	2	0	34
2001	2	0	1	1	2	0	4	11	0	0	0	0	21
2002	0	0	2	0	4	0	14	13	2	1	0	15	51
2003	0	0	5	32	2	0	0	8	0	0	0	1	48
2004	4	0	0	2	0	0	0	0	0	7	0	0	13
2005	10	0	20	17	3	0	1	1	0	0	0	0	52
2006	0	1	0	3	0	0	5	7	2	0	0	6	24

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
2007	3	1	2	2	1	0	5	6	0	0	0	0	20
2008	2	0	0	0	3	0	2	1	0	6	1	0	15
2009	1	0	4	3	1	0	3	8	0	0	0	0	20
2010	0	3	0	1	5	0	45	5	3	3	0	1	66
2011	0	0	0	2	6	2	0	18	3	0	11	0	42
2012	0	0	0	64	0	1	5	6	0	0	1	11	88
2013	7	0	1	3	0	0	5	28	5	0	2	18	69
2014	0	2	0	1	4	0	0	2	1	11	7	6	34
2015	2	0	1	0	0	0	0	0	0	0	3	30	36
2016	2	5	0	20	0	0	12	1	0	0	0	0	40
2017	0	6	0	0	0	0	0	1	0	0	2	0	9
2018	0	0	0	0	0	0	0	0	0	0	0	1	1
Mean	6.2	2.2	4.9	6.4	3.7	2.5	4.3	7.3	1.6	2.4	1.5	3.3	46.5
St.De v	17.5	4.7	9.5	12.1	6.7	8.1	7.8	8.9	4.7	5.1	2.8	6.3	46.4
Skew	3.6	3.9	2.6	3.1	2.6	4.6	3.8	2.5	5.2	2.8	2.1	2.5	2.1
Kurt osis	14.2	19.9	7.4	12.8	7.6	24.8	19.3	8.6	31.7	9.2	4.3	7.7	5.1

1.1.2 Red Sea Basin, Monthly Rainfall Data of the Representative Station RSB002 from 1981 to 2018 in (mm)

Table 2: Red Sea Basin, Monthly Rainfall Data of the Representative Station RSB002 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	0	0	20	23	128	19	70	39	29	10	0	0	338
1982	93	7	69	6	0	0	0	7	1	6	6	10	205
1983	2	57	3	45	14	10	14	13	1	64	2	5	230
1984	1	0	0	0	79	1	0	15	0	0	0	15	111
1985	2	0	0	1	24	1	28	3	1	0	0	1	61
1986	0	0	15	68	1	35	13	215	126	11	7	20	511
1987	46	2	24	70	24	7	2	8	2	10	0	0	195
1988	49	0	0	7	4	0	57	19	17	6	0	0	159
1989	8	27	11	88	0	3	3	37	2	0	0	7	186
1990	1	8	12	11	0	0	8	3	1	0	0	0	44

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1991	1	18	5	0	0	0	0	0	0	5	1	15	45
1992	9	2	3	8	0	0	4	33	5	3	1	1	69
1993	0	1	0	17	62	0	5	3	0	0	11	0	99
1994	0	0	15	2	3	0	49	5	62	0	68	1	205
1995	0	8	44	5	12	0	18	10	1	149	2	10	259
1996	5	1	4	9	10	17	17	14	16	122	3	0	218
1997	1	0	10	5	48	12	23	26	23	29	10	0	187
1998	0	3	0	3	22	2	6	68	30	24	0	0	158
1999	1	0	5	3	1	2	13	8	4	15	0	1	53
2000	0	0	1	2	1	4	9	19	33	23	5	0	97
2001	0	0	15	3	4	1	16	31	6	2	0	0	78
2002	3	0	1	1	9	20	13	19	46	8	0	14	134
2003	1	0	2	16	1	7	4	36	22	5	0	24	118
2004	5	0	0	19	3	18	7	7	8	34	4	7	112
2005	15	0	7	65	22	0	11	25	4	0	1	0	150
2006	1	3	2	7	0	3	5	83	17	11	1	8	141
2007	5	1	6	3	40	4	27	24	6	0	0	1	117
2008	3	0	0	1	9	16	20	13	5	7	1	0	75
2009	0	0	6	25	7	1	11	42	3	1	0	0	96
2010	0	13	6	3	34	58	79	25	2	5	0	1	226
2011	0	0	1	7	35	2	5	25	11	0	7	1	94
2012	0	0	2	45	13	2	21	12	4	1	0	3	103
2013	1	0	14	2	0	0	11	64	6	7	2	20	127
2014	1	1	0	6	13	10	4	26	39	61	4	4	169
2015	3	1	0	0	4	4	2	8	1	1	4	34	62
2016	2	6	0	17	5	1	57	22	1	2	0	0	113
2017	0	7	1	2	5	0	3	5	2	0	3	0	28
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean	6.8	4.4	8.0	15.7	16.8	6.8	16.7	26.6	14.1	16.4	3.8	5.3	141.4
St.Dev	17.7	10.3	13.3	22.5	25.8	11.4	19.6	36.1	23.4	32.0	10.9	8.1	92.6
Skew	3.7	3.8	3.0	1.9	2.6	2.8	1.8	3.8	3.1	2.9	5.3	1.8	1.8
Kurtosis	15.4	18.0	11.8	2.9	8.6	10.1	2.9	19.3	13.1	9.6	32.6	3.1	5.6

1.1.3 Red Sea Basin, Monthly Rainfall Data of the Representative Station RSB003 from 1981 to 2018 in (mm)

Table 3: Red Sea Basin, Monthly Rainfall Data of the Representative Station RSB003 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	1	0	16	36	99	49	157	50	45	5	0	1	459
1982	134	11	80	2	0	0	0	9	0	10	4	37	287
1983	5	130	1	78	17	18	39	27	1	29	1	6	352
1984	0	0	0	0	93	0	0	67	0	0	0	12	172
1985	1	0	0	2	22	1	51	2	1	0	0	1	81
1986	0	0	6	39	1	10	18	367	365	6	5	27	844
1987	16	1	24	55	18	5	2	6	7	11	0	0	145
1988	81	0	0	12	5	0	73	28	26	13	0	0	238
1989	12	34	24	92	0	4	5	50	3	0	2	14	240
1990	2	5	12	12	0	0	5	1	1	0	0	0	38
1991	0	25	6	0	0	0	0	1	0	5	2	17	56
1992	23	11	6	20	0	3	17	34	10	3	1	1	129
1993	1	2	0	32	64	0	15	6	0	0	44	0	164
1994	0	0	9	3	3	6	124	5	62	0	121	2	335
1995	0	39	82	11	23	0	42	12	2	100	7	16	334
1996	11	1	9	10	14	22	37	18	23	52	2	0	199
1997	4	0	23	7	55	41	90	97	30	29	8	0	384
1998	1	2	1	6	5	5	21	93	51	37	0	0	222
1999	3	1	17	2	7	6	41	22	5	29	1	1	135
2000	0	0	1	3	2	13	8	32	46	30	2	0	137
2001	0	0	27	4	5	2	33	34	10	5	0	0	120
2002	5	0	1	3	14	23	22	30	73	10	0	27	208
2003	2	0	0	4	2	18	17	52	33	4	0	27	159
2004	10	0	1	25	5	36	6	16	18	35	13	28	193
2005	14	0	1	146	42	1	23	50	7	0	3	0	287
2006	8	4	2	9	1	6	12	142	44	8	3	5	244
2007	6	0	21	5	78	9	48	36	12	1	4	1	221
2008	7	0	0	1	18	46	48	24	10	8	0	0	162
2009	0	0	5	21	13	5	15	60	9	3	0	0	131
2010	0	28	10	4	66	76	78	37	2	6	0	1	308
2011	0	1	2	10	50	15	13	28	15	2	6	0	142

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
2012	1	0	2	110	19	16	56	28	5	1	0	2	240
2013	0	0	26	2	0	0	26	78	11	16	1	2	162
2014	1	1	0	13	25	48	22	44	72	68	1	0	295
2015	2	2	0	0	15	19	5	45	2	1	5	29	125
2016	3	8	0	18	17	3	55	39	5	9	1	0	158
2017	0	5	4	7	30	2	13	11	8	2	5	0	87
2018	0	0	0	1	0	0	0	0	0	1	0	0	2
Mean	9.3	8.2	11.0	21.2	21.8	13.4	32.6	44.2	26.7	14.2	6.4	6.8	215.7
St.Dev	24.5	22.3	18.6	32.7	26.8	17.7	34.3	60.8	59.4	21.1	20.2	10.8	142.1
Skew	4.1	4.4	2.8	2.3	1.5	1.8	1.8	4.0	4.9	2.4	5.0	1.4	2.3
Kurtosis	19.2	24.0	8.9	5.8	1.6	3.2	4.1	21.0	28.6	6.8	28.4	0.9	9.0

1.1.4 Monthly Averaged Rainfall Data Based on 3 Representative Stations with Basic Statistical Measures of Red Sea Basin from 1981 to 2018 in (mm)

Table 4: Monthly Averaged Rainfall Data Based on 3 Representative Stations with Basic Statistical Measures of Red Sea Basin from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	0	0	26	21	85	23	77	34	25	6	0	0	297
1982	96	9	63	6	0	0	0	8	1	6	6	18	213
1983	3	71	4	47	13	13	21	16	1	34	2	4	229
1984	1	1	1	0	60	1	0	30	0	0	0	11	105
1985	2	0	0	2	20	1	30	3	1	1	0	2	62
1986	0	1	12	43	1	31	14	209	173	12	7	20	523
1987	50	2	22	45	23	9	3	10	3	7	0	0	174
1988	57	0	0	7	3	0	44	17	15	7	0	0	150
1989	7	23	13	64	0	3	3	30	2	0	1	8	154
1990	1	6	8	11	0	0	7	3	1	0	0	0	37
1991	1	16	6	0	0	0	0	1	0	3	1	11	39
1992	12	5	3	9	0	1	7	26	5	2	1	1	72
1993	0	1	0	17	44	0	7	3	0	0	18	0	90
1994	0	0	12	2	2	2	59	4	42	0	63	1	187

1995	0	16	42	6	12	0	20	8	1	83	3	9	200
1996	6	1	4	7	9	18	18	13	14	58	2	0	150
1997	2	0	12	4	39	18	38	41	19	21	7	0	201
1998	0	4	1	3	10	3	9	61	27	21	0	0	139
1999	1	1	8	2	3	3	18	10	3	16	0	1	66
2000	0	0	1	2	1	6	8	18	27	25	3	0	91
2001	1	0	15	3	4	1	18	25	5	2	0	0	74
2002	3	0	1	2	9	14	16	21	40	6	0	19	131
2003	1	0	2	18	2	8	7	32	18	3	0	17	108
2004	6	0	0	15	3	18	5	8	8	26	6	12	107
2005	13	0	9	76	22	1	12	26	4	0	1	0	164
2006	3	3	1	7	1	3	7	77	21	6	1	6	136
2007	5	1	10	4	40	5	27	22	6	0	1	1	122
2008	4	0	0	1	10	21	24	13	5	7	1	0	86
2009	1	0	5	16	7	2	10	36	4	1	0	0	82
2010	0	15	6	3	35	45	67	23	2	5	0	1	202
2011	0	0	1	6	30	6	6	24	10	1	8	0	92
2012	0	0	1	73	10	6	27	15	3	1	0	6	142
2013	3	0	14	2	0	0	14	57	7	8	2	13	120
2014	1	1	0	7	14	19	9	24	38	47	4	3	167
2015	2	1	0	0	6	8	2	18	1	1	4	31	74
2016	2	6	0	18	7	1	41	21	2	4	0	0	102
2017	0	6	2	3	12	1	5	6	3	1	3	0	42
2018	0	0	0	0	0	0	0	0	0	0	0	1	1
Mean	7.5	5.0	8.0	14.5	14.1	7.7	17.9	26.1	14.1	11.1	3.8	5.2	135.0
St.Dev.	18.8	12.1	12.5	20.3	18.6	10.0	18.5	34.4	28.6	17.7	10.3	7.5	87.9
Skew	3.5	4.3	2.8	1.9	2.0	1.8	1.6	4.0	4.5	2.5	5.0	1.6	2.2
Kurtosis	13.9	23.2	10.0	2.9	4.8	4.0	2.5	21.2	25.3	7.0	29.4	2.5	9.0

1.2 Arab Sea Basin

1.2.1 Arab Sea Basin, Monthly Rainfall Data of the Representative Station

ASB004 from 1981 to 2018 in (mm)

Table 5: Arab Sea Basin, Monthly Rainfall Data of the Representative Station ASB004 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	0	0	33	2	40	0	4	21	0	1	0	0	101

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
1982	57	13	33	16	0	0	0	9	3	1	8	8	148
1983	2	26	7	17	5	11	6	4	1	9	3	4	95
1984	2	1	3	0	8	1	0	5	0	0	0	5	25
1985	2	0	0	1	10	1	8	4	0	1	0	3	30
1986	0	1	14	18	1	37	6	40	25	15	9	13	179
1987	79	2	13	11	19	10	2	12	0	1	0	0	149
1988	34	0	0	1	1	0	2	2	1	0	0	0	41
1989	0	6	6	15	0	1	0	2	0	0	0	2	32
1990	0	5	1	7	0	0	12	6	0	0	0	0	31
1991	1	4	7	0	0	0	0	0	0	0	0	0	12
1992	2	0	0	1	0	0	0	15	0	2	1	2	23
1993	0	2	0	2	9	0	0	1	0	0	0	0	14
1994	0	0	16	0	0	0	7	3	0	0	0	0	26
1995	0	0	1	0	0	0	0	3	0	0	0	0	4
1996	0	0	0	0	2	17	1	6	4	0	0	0	30
1997	1	0	3	0	10	1	0	0	3	4	7	0	29
1998	0	4	0	0	3	1	2	26	0	0	0	0	36
1999	0	0	1	0	0	0	1	3	0	4	0	2	11
2000	0	0	0	0	0	1	7	1	1	18	2	1	31
2001	1	0	2	1	2	0	4	6	0	0	0	0	16
2002	0	0	2	1	7	0	12	12	2	0	0	9	45
2003	0	0	3	22	1	0	0	7	0	0	0	0	33
2004	2	0	0	2	0	0	0	0	0	8	1	0	13
2005	7	0	19	16	2	0	2	2	0	0	0	0	48
2006	0	0	0	4	0	0	4	10	2	0	0	3	23
2007	1	1	2	1	2	1	7	9	0	0	0	0	24
2008	1	0	0	0	4	0	3	2	0	9	1	0	20
2009	1	0	1	4	0	0	4	8	0	0	0	0	18
2010	0	3	0	2	6	0	32	6	2	3	0	1	55
2011	0	0	1	2	4	2	1	18	1	0	6	0	35
2012	0	0	0	58	0	1	6	6	0	0	0	5	76
2013	3	0	1	4	0	0	7	28	4	0	2	8	57
2014	0	1	0	1	6	0	0	2	2	6	5	4	27
2015	0	0	0	0	0	0	0	0	0	0	1	13	14
2016	1	2	0	20	0	0	6	4	0	0	0	0	33
2017	0	4	1	0	0	0	0	1	0	0	1	0	7

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean	5.2	2.0	4.5	6.0	3.7	2.2	3.8	7.5	1.3	2.2	1.2	2.2	41.9
St.Dev	16.0	4.7	8.2	10.8	7.2	6.7	5.7	8.7	4.1	4.2	2.4	3.6	40.6
Skew	3.6	3.9	2.4	3.1	3.6	4.0	3.2	2.0	5.2	2.3	2.1	1.8	2.0
Kurtosis	14.1	18.7	5.9	13.1	16.6	19.6	14.7	4.5	31.5	5.7	3.7	2.9	3.9

1.2.2 Arab Sea Basin, Monthly Rainfall Data of the Representative Station

ASB005 from 1981 to 2018 in (mm)

Table 6: Arab Sea Basin, Monthly Rainfall Data of the Representative Station ASB005 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	0	0	13	0	2	0	4	3	0	0	0	0	22
1982	23	12	1	1	0	2	7	13	3	0	12	0	74
1983	5	22	1	26	3	3	5	18	2	0	0	0	85
1984	0	0	0	0	0	0	10	12	2	0	10	0	34
1985	1	0	0	0	0	1	12	10	0	0	4	0	28
1986	0	7	0	2	0	1	13	14	1	0	0	1	39
1987	2	0	12	2	0	1	1	7	0	0	0	0	25
1988	0	0	0	3	0	0	7	3	1	0	0	0	14
1989	0	0	26	1	0	1	5	1	0	0	0	11	45
1990	1	39	0	0	0	0	1	1	0	0	0	0	42
1991	0	0	0	0	0	1	1	2	0	0	0	0	4
1992	0	0	0	10	0	0	5	4	0	1	0	0	20
1993	0	3	0	4	6	0	1	3	0	0	0	0	17
1994	0	0	5	0	0	0	35	3	0	0	0	0	43
1995	3	0	18	0	0	0	5	1	1	0	0	0	28
1996	2	0	1	2	1	22	8	9	0	0	0	4	49
1997	0	0	6	0	0	0	1	2	0	0	0	0	9
1998	0	16	12	0	0	0	0	1	0	0	0	0	29
1999	0	0	0	0	0	0	0	1	0	5	0	0	6
2000	0	0	0	0	0	0	2	5	2	0	0	0	9
2001	0	0	6	0	0	0	2	8	0	0	0	0	16
2002	0	0	0	12	2	0	10	8	0	0	0	1	33

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
2003	0	0	0	1	0	0	8	10	6	0	1	0	26
2004	2	0	0	3	0	1	7	11	2	0	7	0	33
2005	0	0	1	1	0	1	3	1	0	0	0	0	7
2006	0	0	0	6	0	0	17	9	0	0	0	0	32
2007	0	0	5	0	7	5	4	9	1	0	0	0	31
2008	0	0	0	0	2	1	2	3	0	8	2	0	18
2009	0	0	0	0	0	0	3	1	3	0	0	0	7
2010	3	1	0	0	0	0	6	1	0	0	4	0	15
2011	0	1	0	0	0	1	2	5	0	0	7	0	16
2012	0	2	0	0	0	0	3	7	0	0	0	0	12
2013	0	0	2	0	0	0	3	3	0	0	0	0	8
2014	0	0	0	0	0	0	2	4	0	0	0	0	6
2015	0	0	0	0	0	0	5	4	3	0	1	0	13
2016	0	0	0	2	3	6	13	10	1	0	0	0	35
2017	0	0	1	1	1	1	6	3	0	0	0	0	13
2018	0	0	0	0	1	0	1	0	2	0	0	0	4
Mean	1.1	2.7	2.9	2.0	0.7	1.3	5.8	5.5	0.8	0.4	1.3	0.4	24.9
St.Dev	3.8	7.6	5.8	4.7	1.6	3.7	6.3	4.4	1.3	1.5	2.9	1.9	18.0
Skew	5.2	3.5	2.4	3.7	2.6	4.9	2.8	0.8	2.0	4.3	2.4	5.0	1.4
Kurtosis	31.1	14.1	6.5	17.3	7.6	28.1	11.6	0.1	5.5	20.0	5.9	28.4	2.8

1.2.3 Arab Sea Basin, Monthly Rainfall Data of the Representative Station ASB006 from 1981 to 2018 in (mm)

Table 7: Arab Sea Basin, Monthly Rainfall Data of the Representative Station ASB006 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	0	0	9	1	21	0	1	13	0	0	0	2	47
1982	33	23	6	9	0	0	2	10	3	3	20	1	110
1983	4	27	4	11	1	9	4	4	1	0	0	0	65
1984	0	0	6	0	0	1	1	8	0	0	11	2	29
1985	0	0	0	0	0	0	7	5	0	0	4	0	16
1986	0	4	0	4	0	1	4	11	9	0	1	5	39

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1987	10	0	8	1	0	1	0	7	0	0	0	0	27
1988	1	0	0	0	0	0	0	0	0	0	0	0	1
1989	0	0	14	7	0	1	0	0	0	0	0	0	22
1990	0	50	0	5	0	0	1	0	0	0	0	0	56
1991	0	3	4	0	0	0	0	0	0	0	0	0	7
1992	0	0	0	19	0	0	0	1	0	5	0	1	26
1993	0	5	0	6	21	0	0	0	0	0	0	0	32
1994	0	0	10	0	0	0	7	0	0	0	0	0	17
1995	0	0	41	2	0	0	1	0	0	0	0	0	44
1996	1	0	0	3	0	18	7	4	0	0	0	0	33
1997	0	0	7	0	0	0	0	0	3	0	0	0	10
1998	0	20	10	0	0	0	0	1	0	0	0	0	31
1999	0	1	0	0	0	0	0	0	0	0	0	0	1
2000	0	0	0	0	0	0	0	0	0	0	0	0	0
2001	0	0	11	0	0	0	0	5	0	0	0	0	16
2002	0	0	0	9	0	0	1	8	0	0	0	0	18
2003	0	0	0	0	0	0	0	13	0	0	0	0	13
2004	0	0	0	5	0	0	0	0	0	0	36	0	41
2005	0	0	0	9	0	0	1	0	0	0	0	0	10
2006	0	0	0	9	0	0	0	0	0	0	0	0	9
2007	0	0	0	0	0	0	1	1	0	0	0	0	2
2008	0	0	0	0	1	0	0	0	0	2	0	0	3
2009	0	0	0	0	0	0	0	0	0	0	0	0	0
2010	0	0	0	1	0	0	0	0	0	0	0	0	1
2011	0	1	0	1	4	0	0	0	0	0	5	0	11
2012	0	0	0	1	0	0	0	7	0	0	0	0	8
2013	0	0	1	0	0	0	2	4	0	0	0	0	7
2014	0	0	0	0	0	0	0	5	3	0	0	0	8
2015	0	0	1	0	0	0	0	0	0	0	0	1	2
2016	0	0	0	3	0	0	1	7	0	0	0	0	11
2017	0	0	2	0	0	0	2	0	0	0	0	0	4
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean	1.3	3.5	3.5	2.8	1.3	0.8	1.1	3.0	0.5	0.3	2.0	0.3	20.4
St.Dev	5.5	9.9	7.3	4.3	4.7	3.2	2.0	4.0	1.6	1.0	6.7	0.9	22.2
Skew	5.2	3.3	3.6	1.9	3.9	4.6	2.1	1.1	4.1	3.9	4.0	3.8	2.0
Kurtosis	30.6	12.9	17.8	4.2	15.5	23.8	3.9	0.2	20.2	16.5	18.4	17.7	5.7

1.2.4 Monthly Averaged Rainfall Data Based on 3 Representative Stations with Basic Statistical Measures of Arab Sea Basin from 1981 to 2018 in (mm)

Table 8: Monthly Averaged Rainfall Data Based on 3 Representative Stations with Basic Statistical Measures of Arab Sea Basin from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981		0	21	13	47	16	55	25	15	2	0	0	57
1982	71	12	38	6	0	1	2	10	2	4	8	15	111
1983	4	59	3	40	8	11	17	16	1	13	1	3	82
1984	1	0	1	0	34	0	3	28	1	0	3	6	29
1985	1	0	0	1	11	1	24	5	0	0	1	1	25
1986	0	3	7	20	1	16	12	140	130	7	5	14	86
1987	32	1	16	23	12	5	2	8	2	4	0	0	67
1988	38	0	0	5	2	0	27	11	9	4	0	0	19
1989	4	13	19	36	0	2	3	18	1	0	1	9	33
1990	1	16	4	6	0	0	6	3	0	0	0	0	43
1991	0	10	4	0	0	0	0	1	0	2	1	6	8
1992	8	4	2	10	0	1	7	18	3	2	1	1	23
1993	0	2	0	13	26	0	5	3	0	0	15	0	21
1994	0	0	10	1	1	2	55	4	21	0	40	1	29
1995	1	13	34	4	8	0	16	5	1	33	2	5	25
1996	4	0	3	4	6	20	15	11	9	17	1	1	37
1997	2	0	11	2	22	14	30	33	11	11	5	0	16
1998	0	7	4	2	3	2	8	40	17	12	0	0	32
1999	1	0	6	1	2	2	14	9	2	13	0	1	6
2000	0	0	0	1	1	5	6	13	16	16	1	0	13
2001	0	0	12	2	2	1	13	16	3	2	0	0	16
2002	2	0	1	5	8	8	15	17	25	3	0	12	32
2003	1	0	1	9	1	6	8	23	13	1	0	9	24
2004	5	0	0	10	2	12	4	9	7	14	7	9	29
2005	7	0	7	54	15	1	9	18	2	0	1	0	22
2006	3	1	1	6	0	2	11	54	15	3	1	3	21
2007	2	0	9	2	29	5	20	18	4	0	1	0	19
2008	3	0	0	0	8	16	18	10	3	8	1	0	14
2009	0	0	2	8	4	2	7	23	4	1	0	0	8
2010	1	11	3	2	24	25	39	15	1	3	1	1	24
2011	0	1	1	4	18	6	5	17	5	1	6	0	21
2012	0	1	1	56	6	6	22	14	2	0	0	2	32

2013	1	0	10	2	0	0	12	36	5	5	1	3	24
2014	0	1	0	5	10	16	8	17	25	25	2	1	14
2015	1	1	0	0	5	6	3	16	2	0	2	14	10
2016	1	3	0	13	7	3	25	18	2	3	0	0	26
2017	0	3	2	3	10	1	6	5	3	1	2	0	8
2018	0	0	0	0	0	0	0	0	1	0	0	0	1
Mean	1.3	3.5	3.5	2.8	1.3	0.8	1.1	3.0	0.5	0.3	2.0	0.3	29.1
St.Dev	5.5	9.9	7.3	4.3	4.7	3.2	2.0	4.0	1.6	1.0	6.7	0.9	22.8
Skew	5.2	3.3	3.6	1.9	3.9	4.6	2.1	1.1	4.1	3.9	4.0	3.8	2.0
Kurtosis	30.6	12.9	17.8	4.2	15.5	23.8	3.9	0.2	20.2	16.5	18.4	17.7	5.7

1.3 Gulf Aden Basin

1.3.1 Gulf Aden Basin, Monthly Rainfall Data of the Representative Station

GAB007 from 1981 to 2018 in (mm)

Table 9: Gulf Aden Basin, Monthly Rainfall Data of the Representative Station GAB007 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	3	0	20	7	24	1	57	10	1	3	0	0	126
1982	206	20	17	7	0	0	0	26	0	17	8	8	309
1983	2	44	0	21	1	50	1	1	0	10	0	4	134
1984	0	0	0	0	9	0	0	5	0	0	0	5	19
1985	1	0	0	0	4	0	7	1	0	0	0	3	16
1986	0	0	10	17	0	11	1	34	33	10	9	13	138
1987	20	1	8	35	2	2	0	5	1	8	0	0	82
1988	15	0	0	8	5	0	20	1	26	3	0	0	78
1989	5	24	15	43	0	9	5	19	1	0	1	2	124
1990	2	7	5	12	0	0	3	0	0	0	0	0	29
1991	0	15	8	0	0	0	0	0	0	1	4	0	28
1992	22	8	8	28	1	0	5	7	5	3	2	2	91
1993	2	4	0	11	14	0	2	1	0	0	14	0	48
1994	0	0	14	3	0	1	12	1	5	0	20	0	56
1995	0	14	28	9	1	0	6	10	3	29	0	0	100
1996	9	1	5	5	16	29	4	18	36	12	1	0	136
1997	1	0	3	0	5	3	13	4	14	28	5	0	76
1998	1	0	1	2	0	1	3	19	5	2	0	0	34
1999	0	2	15	1	2	2	3	12	1	33	2	2	75

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
2000	0	0	0	6	3	20	18	5	5	17	0	1	75
2001	0	0	8	1	4	2	22	10	0	1	0	0	48
2002	5	0	0	1	16	3	12	15	51	1	0	9	113
2003	2	0	0	1	2	1	0	4	5	0	0	0	15
2004	1	0	0	7	4	4	1	0	1	17	14	0	49
2005	9	0	0	21	33	1	12	41	1	0	3	0	121
2006	1	3	1	7	0	2	10	26	12	1	0	3	66
2007	8	0	7	4	7	5	12	15	4	0	10	0	72
2008	3	0	0	0	8	3	11	5	4	7	0	0	41
2009	0	0	1	6	2	4	7	28	6	3	0	0	57
2010	0	13	4	2	22	58	12	14	0	1	0	1	127
2011	0	1	1	7	3	3	5	15	3	0	4	0	42
2012	0	0	1	14	1	2	10	8	1	1	0	5	43
2013	0	0	5	1	0	0	13	15	5	5	2	8	54
2014	1	2	0	4	4	9	5	10	18	17	1	4	75
2015	0	1	0	0	5	3	1	12	1	0	2	13	38
2016	2	4	0	12	4	1	11	8	2	4	0	0	48
2017	0	2	1	1	10	0	4	4	2	1	3	0	28
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean	8.4	4.4	4.9	8.0	5.6	6.1	8.1	10.8	6.6	6.2	2.8	2.2	74.0
St.Dev	32.9	8.8	6.7	9.9	7.5	12.7	9.9	9.9	11.4	8.8	4.7	3.6	54.0
Skew	5.7	2.9	1.6	1.9	2.0	3.0	3.1	1.2	2.4	1.6	2.1	1.8	2.1
Kurtosis	35.9	10.3	2.6	3.9	4.1	9.9	14.7	1.3	6.3	2.2	4.5	2.9	8.1

1.3.2 Gulf Aden Basin, Monthly Rainfall Data of the Representative Station

GAB008 from 1981 to 2018 in (mm)

Table 10: Gulf Aden Basin, Monthly Rainfall Data of the Representative Station GAB008 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	3	0	16	7	22	0	50	8	0	2	0	1	109
1982	173	20	12	8	0	0	0	23	0	8	14	13	271
1983	1	31	0	14	0	44	0	0	0	6	0	3	99

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1984	0	0	0	0	5	0	0	0	0	0	1	6	12
1985	1	0	0	0	0	0	9	1	0	0	0	0	11
1986	0	0	8	8	0	6	0	17	21	7	7	10	84
1987	12	0	6	30	1	2	0	5	1	8	0	0	65
1988	7	0	0	4	3	0	17	0	25	1	0	0	57
1989	4	18	9	39	0	9	4	11	0	0	1	8	103
1990	1	11	4	19	0	0	2	0	0	0	0	0	37
1991	0	11	7	0	0	0	0	0	0	0	3	4	25
1992	15	5	6	21	0	0	3	5	4	4	1	2	66
1993	1	5	0	9	10	0	1	0	0	0	7	0	33
1994	0	0	13	2	0	0	8	0	0	0	6	4	33
1995	0	8	19	6	0	0	5	8	3	14	0	12	75
1996	7	1	2	3	11	26	2	14	30	7	1	0	104
1997	0	0	1	0	3	1	8	1	7	26	3	0	50
1998	1	3	2	1	0	1	1	12	2	0	0	0	23
1999	0	4	9	1	2	1	1	7	0	16	2	2	45
2000	0	0	0	5	2	13	14	4	3	7	0	0	48
2001	0	0	8	0	2	1	16	6	0	0	0	0	33
2002	4	0	0	2	7	1	6	19	27	1	0	0	67
2003	1	0	0	0	4	0	0	4	3	0	0	15	27
2004	0	0	0	4	3	1	1	0	0	15	21	5	50
2005	7	0	0	6	17	0	4	30	0	0	0	0	64
2006	1	1	0	3	0	0	8	12	4	0	0	2	31
2007	11	0	1	1	2	3	8	9	1	0	12	0	48
2008	2	0	0	0	5	1	5	2	1	6	0	0	22
2009	0	0	0	2	1	2	3	18	6	4	0	0	36
2010	0	8	1	2	9	40	4	4	0	0	0	0	68
2011	0	2	0	6	0	1	3	9	1	0	2	0	24
2012	0	0	0	5	0	0	3	2	0	0	0	0	10
2013	0	0	2	0	0	0	9	5	2	2	1	0	21
2014	1	1	0	1	1	4	2	5	5	13	0	0	33
2015	0	0	0	0	1	1	0	1	0	0	0	3	6
2016	0	0	0	6	0	0	2	5	1	2	0	0	16
2017	0	0	2	0	1	0	1	2	0	0	0	0	6
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean	6.7	3.4	3.4	5.7	2.9	4.2	5.3	6.6	3.9	3.9	2.2	2.4	50.3

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
St.Dev	27.6	6.7	5.0	8.5	4.9	10.1	8.6	7.1	7.8	5.9	4.5	4.0	46.4
Skew	5.8	2.5	1.5	2.4	2.3	3.0	3.8	1.4	2.4	1.9	2.7	1.8	2.8
Kurtosis	36.5	7.5	1.9	6.7	6.3	9.5	19.3	2.0	5.3	4.0	8.3	2.8	12.6

1.3.3 Gulf Aden Basin, Monthly Rainfall Data of the Representative Station

GAB009 from 1981 to 2018 in (mm)

Table 11: Gulf Aden Basin, Monthly Rainfall Data of the Representative Station GAB009 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	1	0	3	10	19	0	7	7	1	1	0	5	54
1982	51	18	11	8	0	0	1	9	1	4	24	2	129
1983	2	26	0	7	0	10	2	3	0	0	0	0	50
1984	0	0	0	0	1	1	1	3	0	0	9	4	19
1985	2	0	0	0	0	0	9	4	0	0	2	0	17
1986	0	1	0	7	0	1	2	12	12	0	1	7	43
1987	7	0	3	12	0	3	0	6	0	2	0	0	33
1988	4	0	0	0	1	0	2	0	2	0	0	0	9
1989	0	2	16	17	0	3	1	2	0	1	0	0	42
1990	3	12	0	16	0	0	0	0	0	0	0	0	31
1991	0	6	6	0	0	0	0	0	0	0	0	0	12
1992	2	0	1	13	0	0	1	3	0	6	0	1	27
1993	0	5	0	8	10	0	0	0	0	0	2	0	25
1994	0	0	13	0	0	0	7	0	0	0	1	0	21
1995	0	1	22	2	0	0	1	0	0	2	0	1	29
1996	3	0	0	3	1	31	1	8	5	2	0	0	54
1997	0	0	7	1	0	0	0	0	4	1	0	0	13
1998	0	48	18	0	0	0	2	3	1	0	0	0	72
1999	0	4	0	0	0	0	3	0	0	0	0	0	7
2000	0	0	0	0	0	0	1	1	0	0	0	0	2
2001	0	0	8	0	0	0	0	2	0	0	0	3	13
2002	0	0	0	8	0	0	2	16	1	0	0	0	27
2003	0	0	0	2	1	0	0	6	0	0	0	2	11
2004	0	0	0	4	0	0	0	0	0	1	30	3	38

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
2005	1	0	0	11	1	0	1	1	0	0	0	0	15
2006	0	0	0	3	0	0	4	1	0	0	0	1	9
2007	1	0	0	0	0	1	6	3	0	0	1	0	12
2008	0	0	0	0	5	0	0	1	0	6	0	0	12
2009	0	0	0	0	0	0	0	1	1	1	0	0	3
2010	0	0	0	4	0	0	0	0	0	0	0	0	4
2011	0	5	0	2	0	0	1	0	0	0	10	0	18
2012	0	0	0	1	0	0	0	6	0	0	0	0	7
2013	0	0	12	0	0	0	2	1	0	0	0	0	15
2014	1	1	0	0	0	0	0	3	0	3	0	0	8
2015	0	0	0	0	0	0	0	0	0	0	1	1	2
2016	0	0	0	3	0	1	12	10	0	0	0	0	26
2017	0	0	2	0	0	0	0	0	0	0	0	0	2
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean	2.1	3.4	3.2	3.7	1.0	1.3	1.8	2.9	0.7	0.8	2.1	0.8	24.0
St.Dev	8.2	9.1	5.8	4.9	3.4	5.2	2.8	3.8	2.1	1.5	6.3	1.6	24.1
Skew	5.6	3.6	1.8	1.2	4.2	5.1	2.1	1.6	4.1	2.3	3.5	2.3	2.4
Kurtosis	35.5	15.9	2.6	0.6	20.5	29.7	4.6	2.6	20.7	5.4	13.0	6.0	8.4

1.3.4 Monthly Averaged Rainfall Data Based on 3 Representative Stations with Basic Statistical Measures of Gulf Aden Basin from 1981 to 2018 in (mm)

Table 12: Monthly Averaged Rainfall Data Based on 3 Representative Stations with Basic Statistical Measures of Gulf Aden Basin from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	2	0	13	8	22	0	38	8	1	2	0	2	96
1982	143	19	13	8	0	0	0	19	0	10	15	8	236
1983	2	34	0	14	0	35	1	1	0	5	0	2	94
1984	0	0	0	0	5	0	0	3	0	0	3	5	17
1985	1	0	0	0	1	0	8	2	0	0	1	1	15
1986	0	0	6	11	0	6	1	21	22	6	6	10	88
1987	13	0	6	26	1	2	0	5	1	6	0	0	60
1988	9	0	0	4	3	0	13	0	18	1	0	0	48

1989	3	15	13	33	0	7	3	11	0	0	1	3	90
1990	2	10	3	16	0	0	2	0	0	0	0	0	32
1991	0	11	7	0	0	0	0	0	0	0	2	1	22
1992	13	4	5	21	0	0	3	5	3	4	1	2	61
1993	1	5	0	9	11	0	1	0	0	0	8	0	35
1994	0	0	13	2	0	0	9	0	2	0	9	1	37
1995	0	8	23	6	0	0	4	6	2	15	0	4	68
1996	6	1	2	4	9	29	2	13	24	7	1	0	98
1997	0	0	4	0	3	1	7	2	8	18	3	0	46
1998	1	17	7	1	0	1	2	11	3	1	0	0	43
1999	0	3	8	1	1	1	2	6	0	16	1	1	42
2000	0	0	0	4	2	11	11	3	3	8	0	0	42
2001	0	0	8	0	2	1	13	6	0	0	0	1	31
2002	3	0	0	4	8	1	7	17	26	1	0	3	69
2003	1	0	0	1	2	0	0	5	3	0	0	6	18
2004	0	0	0	5	2	2	1	0	0	11	22	3	46
2005	6	0	0	13	17	0	6	24	0	0	1	0	67
2006	1	1	0	4	0	1	7	13	5	0	0	2	35
2007	7	0	3	2	3	3	9	9	2	0	8	0	44
2008	2	0	0	0	6	1	5	3	2	6	0	0	25
2009	0	0	0	3	1	2	3	16	4	3	0	0	32
2010	0	7	2	3	10	33	5	6	0	0	0	0	66
2011	0	3	0	5	1	1	3	8	1	0	5	0	28
2012	0	0	0	7	0	1	4	5	0	0	0	2	20
2013	0	0	6	0	0	0	8	7	2	2	1	3	30
2014	1	1	0	2	2	4	2	6	8	11	0	1	39
2015	0	0	0	0	2	1	0	4	0	0	1	6	15
2016	1	1	0	7	1	1	8	8	1	2	0	0	30
2017	0	1	2	0	4	0	2	2	1	0	1	0	12
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean	5.7	3.7	3.8	5.8	3.2	3.9	5.1	6.8	3.7	3.6	2.4	1.8	49
St.De v	22.9	7.0	5.3	7.4	4.8	8.6	6.5	6.1	6.8	5.0	4.5	2.3	39.5
Skew	5.7	2.6	1.7	2.0	2.3	2.8	3.4	1.1	2.3	1.5	2.7	1.7	2.8
Kurt osis	36.3	8.1	3.2	4.5	5.8	7.8	16.5	0.8	4.9	1.5	8.9	3.3	12.0 6

1.4 Rub' Al Khali Basin

1.4.1 Rub' Al Khali Basin, Monthly Rainfall Data of the Representative Station

RKB010 from 1981 to 2018 in (mm)

Table 13: Rub' Al Khali Basin, Monthly Rainfall Data of the Representative Station RKB010 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	0	0	63	5	36	0	17	23	0	1	0	0	145
1982	27	11	7	7	0	0	1	8	7	3	10	9	90
1983	2	45	14	21	9	9	6	4	1	7	6	1	125
1984	6	3	10	1	2	2	0	9	0	0	2	4	39
1985	1	2	0	4	12	4	14	6	1	3	1	6	54
1986	0	3	14	23	2	11	16	45	23	15	8	12	172
1987	81	3	18	9	35	14	4	19	0	0	0	0	183
1988	2	1	0	1	0	0	0	2	0	0	0	0	6
1989	0	0	3	10	0	0	0	0	0	0	0	1	14
1990	0	3	1	8	1	0	5	1	0	0	0	0	19
1991	1	3	14	1	2	0	0	0	0	3	0	0	24
1992	2	1	1	1	0	0	1	48	0	4	2	3	63
1993	1	4	0	14	16	0	0	1	0	0	0	0	36
1994	0	0	3	0	0	0	14	6	0	0	0	0	23
1995	2	0	1	0	0	0	3	2	0	0	0	0	8
1996	0	0	2	0	0	5	2	6	0	0	0	0	15
1997	1	0	29	0	8	5	0	0	1	2	13	0	59
1998	0	6	1	1	14	1	0	9	0	0	0	0	32
1999	2	0	2	0	0	0	1	1	4	0	0	0	10
2000	0	0	0	1	1	0	0	1	1	8	3	1	16
2001	1	0	3	0	0	0	2	1	0	0	0	0	7
2002	1	0	1	3	6	0	4	3	0	0	0	4	22
2003	0	0	1	5	0	0	0	7	0	0	0	0	13
2004	3	0	0	15	0	0	0	0	0	2	1	0	21
2005	5	0	7	7	1	0	0	1	0	0	0	0	21
2006	0	0	0	3	2	0	1	2	0	0	0	1	9
2007	2	0	0	3	0	0	4	3	0	0	0	0	12
2008	0	0	0	0	1	0	2	0	0	19	1	0	23
2009	4	0	1	1	1	0	2	13	0	0	0	0	22
2010	0	4	0	0	6	0	13	3	3	0	0	0	29

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
2011	4	0	0	0	3	0	1	7	0	0	3	0	18
2012	0	0	0	30	0	1	1	6	0	0	1	6	45
2013	5	0	3	2	4	0	1	4	1	0	2	5	27
2014	0	0	1	0	2	0	0	1	0	0	1	1	6
2015	0	0	1	0	0	0	0	0	0	0	0	5	6
2016	1	1	0	31	0	0	3	4	0	0	0	0	40
2017	0	8	1	0	0	0	0	1	0	0	0	0	10
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean	4.1	2.6	5.3	5.4	4.3	1.4	3.1	6.5	1.1	1.8	1.4	1.6	38.5
St.Dev	13.4	7.4	11.4	8.2	8.4	3.2	4.8	10.7	3.8	4.1	2.9	2.8	44.9
Skew	5.1	5.0	3.7	1.9	2.8	2.6	1.8	2.8	5.0	3.0	2.6	2.1	2.0
Kurtosis	29.8	29.5	17.5	3.1	8.5	7.1	2.5	8.9	28.6	9.9	7.2	4.5	3.9

1.4.2 Rub' Al Khali Basin, Monthly Rainfall Data of the Representative Station

RKB011 from 1981 to 2018 in (mm)

Table 14: Rub' Al Khali Basin, Monthly Rainfall Data of the Representative Station RKB011 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	1	0	14	0	0	0	3	2	0	0	0	0	20
1982	16	6	1	0	0	3	5	7	2	0	5	1	46
1983	6	24	1	48	2	0	4	26	2	0	0	0	113
1984	0	0	0	0	1	1	9	10	2	0	5	0	28
1985	0	0	0	1	0	1	11	14	0	0	3	0	30
1986	0	9	0	2	0	1	13	14	0	0	0	0	39
1987	0	0	16	4	0	0	1	7	0	0	0	1	29
1988	0	0	0	9	0	0	8	5	1	0	0	0	23
1989	0	0	33	0	0	8	8	1	0	0	0	19	69
1990	1	9	0	0	0	0	2	2	0	0	0	0	14
1991	0	0	0	0	0	2	1	3	1	0	0	0	7
1992	0	0	0	12	0	0	11	4	1	0	0	0	28
1993	0	2	0	1	2	0	1	4	0	0	0	0	10
1994	0	0	1	0	1	0	43	5	0	0	0	0	50

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1995	6	0	16	0	0	0	6	3	3	0	0	0	34
1996	1	0	2	1	2	26	12	3	0	0	0	6	53
1997	0	0	8	0	0	0	3	4	0	0	0	0	15
1998	0	4	3	0	0	0	1	2	1	0	0	1	12
1999	0	0	0	0	0	0	1	2	0	10	0	0	13
2000	0	0	0	0	0	0	4	8	4	0	0	0	16
2001	0	0	2	0	0	0	4	7	0	0	0	0	13
2002	0	0	0	14	6	1	21	10	0	0	0	3	55
2003	0	0	0	2	0	0	16	9	7	0	1	0	35
2004	4	0	0	1	0	2	9	16	2	0	2	1	37
2005	0	0	3	0	0	3	4	2	0	0	0	0	12
2006	0	0	0	2	0	0	24	13	1	0	0	0	40
2007	0	0	17	0	14	11	7	13	1	0	1	0	64
2008	0	0	0	0	2	1	3	5	0	15	3	0	29
2009	2	0	1	0	0	0	4	2	6	0	0	0	15
2010	6	2	0	0	0	0	9	2	1	0	12	2	34
2011	0	0	0	1	1	1	3	16	2	1	12	0	37
2012	0	3	0	0	0	0	8	5	0	0	0	0	16
2013	0	0	6	0	0	1	8	4	0	0	0	2	21
2014	0	0	0	0	0	0	6	2	0	0	0	0	8
2015	0	0	0	0	0	0	9	6	5	0	2	0	22
2016	0	0	0	0	5	8	16	11	2	0	0	0	42
2017	0	0	0	2	2	3	14	7	1	0	0	0	29
2018	0	0	0	0	1	0	1	1	3	0	0	0	6
Mean	1.1	1.6	3.3	2.6	1.0	1.9	8.2	6.8	1.3	0.7	1.2	0.9	30.6
St.Dev	3.0	4.3	6.9	8.1	2.5	4.7	8.0	5.4	1.7	2.8	2.9	3.2	20.8
Skew	3.6	3.9	2.7	4.7	3.9	3.9	2.4	1.4	1.7	4.3	2.9	4.9	1.7
Kurtosis	16.0	18.9	8.5	26.5	18.6	18.8	8.7	2.7	3.0	19.4	9.2	28.3	5.1

1.4.3 Rub' Al Khali Basin, Monthly Rainfall Data of the Representative Station

RKB012 from 1981 to 2018 in (mm)

Table 15: Rub' Al Khali Basin, Monthly Rainfall Data of the Representative Station RKB012 from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	0	0	7	0	15	0	1	9	0	0	0	4	36
1982	23	17	5	5	0	0	3	11	4	2	13	1	84
1983	4	26	7	31	2	8	4	8	1	0	0	0	91
1984	0	0	7	0	0	1	1	8	0	0	10	1	28
1985	2	0	0	0	0	0	7	6	0	0	4	0	19
1986	0	7	0	5	0	0	6	12	6	0	0	3	39
1987	10	0	11	1	0	1	0	7	0	0	0	0	30
1988	0	0	0	0	0	0	0	0	0	0	0	0	0
1989	0	0	12	2	0	0	0	0	0	0	0	0	14
1990	0	35	0	1	0	0	0	0	0	0	0	0	36
1991	0	1	5	0	0	0	0	0	0	0	0	0	6
1992	0	0	0	24	0	0	0	0	0	5	0	1	30
1993	0	5	0	6	29	0	0	0	0	0	0	0	40
1994	0	0	9	0	0	0	7	0	0	0	0	0	16
1995	0	0	61	3	0	0	0	0	0	0	0	0	64
1996	0	0	0	3	0	21	4	2	0	0	0	0	30
1997	0	0	14	0	0	0	0	0	6	0	0	0	20
1998	0	17	4	0	0	0	0	0	0	0	0	0	21
1999	0	0	0	0	0	0	0	0	0	0	0	0	0
2000	0	0	0	0	0	0	0	0	0	0	0	0	0
2001	0	0	13	0	0	0	0	3	0	0	0	0	16
2002	0	0	0	10	0	0	1	3	0	0	0	0	14
2003	0	0	0	0	0	0	0	16	0	0	0	0	16
2004	0	0	0	5	0	0	0	0	0	0	50	0	55
2005	0	0	0	22	0	0	1	0	0	0	0	0	23
2006	0	0	0	10	0	0	0	0	0	0	0	0	10
2007	0	0	0	0	0	0	0	0	0	0	0	0	0
2008	0	0	0	0	0	0	0	0	0	0	0	0	0
2009	0	0	0	0	0	0	0	0	0	0	0	0	0
2010	0	0	0	0	0	0	0	0	0	0	0	0	0
2011	0	0	0	1	11	0	0	0	0	0	1	0	13
2012	0	0	0	1	0	0	0	8	0	0	0	0	9
2013	0	0	1	0	1	0	4	6	0	0	0	0	12
2014	0	0	0	0	0	0	0	3	8	0	0	0	11
2015	0	0	3	0	0	0	0	0	0	0	0	1	4
2016	0	0	0	3	0	0	0	9	0	0	0	0	12

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
2017	0	0	2	0	0	0	3	0	0	0	0	0	5
2018	0	0	0	0	0	0	0	0	0	0	0	0	0
Mean	1.0	2.8	4.2	3.5	1.5	0.8	1.1	2.9	0.7	0.2	2.1	0.3	21.2
St.Dev	4.0	7.7	10.2	7.1	5.4	3.6	2.0	4.3	1.9	0.9	8.3	0.8	22.0
Skew	4.6	2.9	4.5	2.6	4.0	5.0	1.8	1.3	2.8	5.0	5.1	3.4	1.6
Kurtosis	24.9	9.1	25.4	7.2	18.6	28.4	2.5	0.9	7.6	27.8	30.2	12.7	2.7

1.4.4 Monthly Averaged Rainfall Data Based on 3 Representative Stations with Basic Statistical Measures of Rub' Al Khali Basin from 1981 to 2018 in (mm)

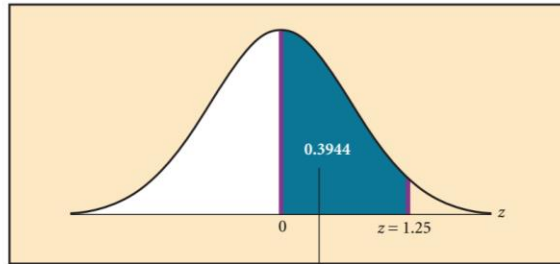
Table 16: Monthly Averaged Rainfall Data Based on 3 Representative Stations with Basic Statistical Measures of Rub' Al Khali Basin from 1981 to 2018 in (mm)

Month	January	February	March	April	May	June	July	August	September	October	November	December	Total
Year													
1981	0	0	28	2	17	0	7	11	0	0	0	1	67
1982	22	11	4	4	0	1	3	9	4	2	9	4	73
1983	4	32	7	33	4	6	5	13	1	2	2	0	110
1984	2	1	6	0	1	1	3	9	1	0	6	2	32
1985	1	1	0	2	4	2	11	9	0	1	3	2	34
1986	0	6	5	10	1	4	12	24	10	5	3	5	83
1987	30	1	15	5	12	5	2	11	0	0	0	0	81
1988	1	0	0	3	0	0	3	2	0	0	0	0	10
1989	0	0	16	4	0	3	3	0	0	0	0	7	32
1990	0	16	0	3	0	0	2	1	0	0	0	0	23
1991	0	1	6	0	1	1	0	1	0	1	0	0	12
1992	1	0	0	12	0	0	4	17	0	3	1	1	40
1993	0	4	0	7	16	0	0	2	0	0	0	0	29
1994	0	0	4	0	0	0	21	4	0	0	0	0	30
1995	3	0	26	1	0	0	3	2	1	0	0	0	35
1996	0	0	1	1	1	17	6	4	0	0	0	2	33
1997	0	0	17	0	3	2	1	1	2	1	4	0	31
1998	0	9	3	0	5	0	0	4	0	0	0	0	22
1999	1	0	1	0	0	0	1	1	1	3	0	0	8
2000	0	0	0	0	0	0	1	3	2	3	1	0	11

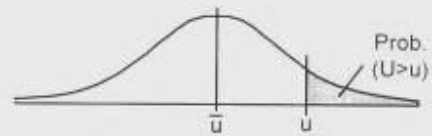
2001	0	0	6	0	0	0	2	4	0	0	0	0	12
2002	0	0	0	9	4	0	9	5	0	0	0	2	30
2003	0	0	0	2	0	0	5	11	2	0	0	0	21
2004	2	0	0	7	0	1	3	5	1	1	18	0	38
2005	2	0	3	10	0	1	2	1	0	0	0	0	19
2006	0	0	0	5	1	0	8	5	0	0	0	0	20
2007	1	0	6	1	5	4	4	5	0	0	0	0	25
2008	0	0	0	0	1	0	2	2	0	11	1	0	17
2009	2	0	1	0	0	0	2	5	2	0	0	0	12
2010	2	2	0	0	2	0	7	2	1	0	4	1	21
2011	1	0	0	1	5	0	1	8	1	0	5	0	23
2012	0	1	0	10	0	0	3	6	0	0	0	2	23
2013	2	0	3	1	2	0	4	5	0	0	1	2	20
2014	0	0	0	0	1	0	2	2	3	0	0	0	8
2015	0	0	1	0	0	0	3	2	2	0	1	2	11
2016	0	0	0	11	2	3	6	8	1	0	0	0	31
2017	0	3	1	1	1	1	6	3	0	0	0	0	15
2018	0	0	0	0	0	0	0	0	1	0	0	0	2
Mean	2.1	2.3	4.3	3.9	2.3	1.4	4.1	5.4	1.0	0.9	1.6	0.9	30.1
St.Dev	5.8	5.9	6.9	6.1	4.0	3.0	4.0	4.9	1.7	2.1	3.3	1.5	23.1
Skew	4.0	3.6	2.1	3.1	2.5	4.1	2.3	1.7	3.6	3.7	3.3	2.2	1.8
Kurtosis	17.5	16.2	4.6	13.4	6.8	21.2	7.9	3.9	17.0	17.0	14.0	5.5	3.4

Appendix 2: Standard Normal Distribution Table

Standard Normal
Distribution Table

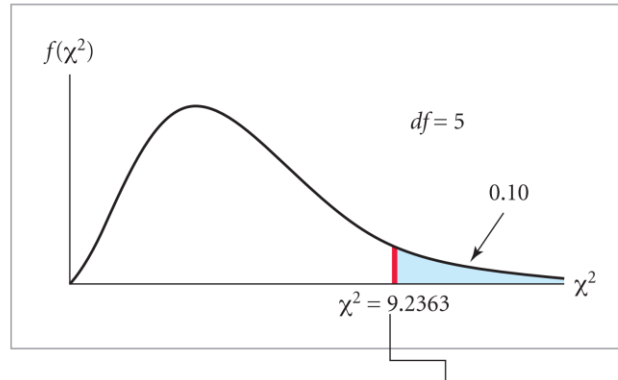


z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990



U	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	U
0.0	50000	49601	49202	48803	48405	48006	47608	47210	46812	46414	0.0
0.1	46017	45620	45224	44828	44433	44038	43644	43251	42858	42465	0.1
0.2	42074	41683	41294	40905	40517	40129	39743	39358	38974	38591	0.2
0.3	38209	37828	37448	37070	36693	36317	35942	35569	35197	34827	0.3
0.4	34458	34090	33724	33360	32997	32636	32276	31918	31561	31207	0.4
0.5	30854	30503	30153	29806	29460	29116	28774	28434	28096	27760	0.5
0.6	27425	27093	26763	26435	26109	25785	25463	25143	24825	24510	0.6
0.7	24196	23885	23576	23270	22965	22663	22363	22065	21770	21476	0.7
0.8	21186	20897	20611	20327	20045	19766	19489	19215	18943	18673	0.8
0.9	18406	18141	17879	17619	17361	17106	16853	16602	16354	16109	0.9
1.0	15866	15625	15386	15151	14917	14686	14457	14231	14007	13786	1.0
1.1	13567	13350	13136	12924	12714	12507	12302	12100	11900	11702	1.1
1.2	11507	11314	11123	10935	10749	10565	10383	10204	10027	98525	1.2
1.3	96800	95098	93418	91759	90123	88508	86915	85343	83793	82264	1.3
1.4	80757	79270	77804	76359	74934	73529	72145	70781	69437	68112	1.4
1.5	66807	65522	64255	63008	61780	60571	59380	58208	57053	55917	1.5
1.6	54799	53699	52616	51551	50503	49471	48457	47460	46479	45514	1.6
1.7	44565	43633	42716	41815	40930	40059	39204	38364	37538	36727	1.7
1.8	35930	35148	34380	33625	32884	32157	31443	30742	30054	29379	1.8
1.9	28717	28067	27429	26803	26190	25588	24998	24419	23852	23295	1.9
2.0	22750	22216	21692	21178	20675	20182	19699	19226	18763	18309	2.0
2.1	17864	17429	17003	16586	16177	15778	15386	15003	14629	14262	2.1
2.2	13903	13555	13209	12874	12548	12224	11911	11604	11304	11011	2.2
2.3	10724	10444	10170	99031	96419	93867	91375	88940	86563	84242	2.3
2.4	81975	79763	77603	75494	73436	71428	69469	67557	65691	63872	2.4
2.5	62097	60366	58677	57031	55426	53861	52336	50849	49400	47988	2.5
2.6	46612	45271	43965	42692	41453	40246	39070	37926	36811	35726	2.6
2.7	34070	32942	31841	30767	29720	28708	27729	26782	25867	24984	2.7
2.8	25551	24771	24012	23274	22557	21860	21182	20524	19884	19262	2.8
2.9	18658	18071	17502	16948	16411	15889	15382	14890	14412	13949	2.9
3.0	13499	13062	12639	12228	11829	11442	11067	10703	10350	10008	3.0
3.1	96760	93444	90226	87103	84074	81139	78298	75551	72898	70336	3.1
3.2	68714	66367	64093	61895	59765	57703	55706	53774	51904	50094	3.2
3.3	48342	46648	45009	43423	41889	40406	38974	37584	36243	34946	3.3
3.4	33693	32481	31311	30179	29086	28029	27009	26023	25071	24151	3.4
3.5	23263	22405	21577	20778	20006	19262	18543	17849	17180	16534	3.5
3.6	15911	15310	14730	14171	13632	13112	12611	12128	11662	11213	3.6
3.7	10780	10361	99611	95740	92010	88417	84957	81624	78414	75324	3.7
3.8	72348	69483	66726	64072	61517	59059	56694	54418	52228	50122	3.8
3.9	48096	46148	44274	42473	40741	39076	37475	35936	34458	33037	3.9
4.0	31671	30359	29099	27888	26726	25609	24536	23507	22518	21569	4.0
4.1	20658	19783	18944	18138	17365	16624	15912	15230	14575	13948	4.1
4.2	13346	12769	12215	11685	11176	10689	10221	97736	93447	89337	4.2
4.3	85399	81627	78015	74555	71241	68069	65031	62123	59340	56675	4.3
4.4	54125	51685	49350	47117	44979	42935	40980	39110	37322	35612	4.4
4.5	33977	32414	30920	29492	28127	26823	25577	24386	23249	22162	4.5
4.6	21125	20133	19187	18283	17420	16597	15810	15060	14344	13660	4.6
4.7	13008	12386	11792	11226	10686	10171	96796	92113	87648	83391	4.7
4.8	79313	75465	71779	68267	64920	61731	58693	55799	53043	50418	4.8
4.9	47918	45538	43272	41115	39061	37107	35247	33476	31792	30190	4.9
5	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	U

Appendix 3: χ^2 Distribution Table

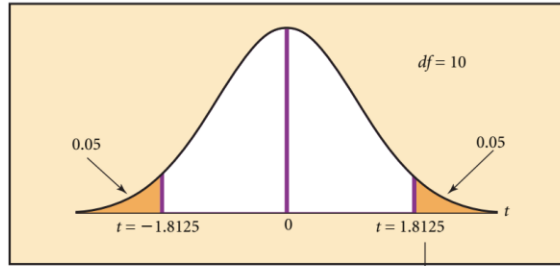


PROBABILITIES (OR AREAS UNDER CHI-SQUARE DISTRIBUTION CURVE
ABOVE GIVEN CHI-SQUARE VALUES)

	0.995	0.99	0.975	0.95	0.90	0.10	0.05	0.025	0.01	0.005
<i>df</i>	<i>Values of Chi-Squared</i>									
1	0.0000	0.0002	0.0010	0.0039	0.0158	2.7055	3.8415	5.0239	6.6349	7.8794
2	0.0100	0.0201	0.0506	0.1026	0.2107	4.6052	5.9915	7.3778	9.2104	10.5965
3	0.0717	0.1148	0.2158	0.3518	0.5844	6.2514	7.8147	9.3484	11.3449	12.8381
4	0.2070	0.2971	0.4844	0.7107	1.0636	7.7794	9.4877	11.1433	13.2767	14.8602
5	0.4118	0.5543	0.8312	1.1455	1.6103	9.2363	11.0705	12.8325	15.0863	16.7496
6	0.6757	0.8721	1.2373	1.6354	2.2041	10.6446	12.5916	14.4494	16.8119	18.5475
7	0.9893	1.2390	1.6899	2.1673	2.8331	12.0170	14.0671	16.0128	18.4753	20.2777
8	1.3444	1.6465	2.1797	2.7326	3.4895	13.3616	15.5073	17.5345	20.0902	21.9549
9	1.7349	2.0879	2.7004	3.3251	4.1682	14.6837	16.9190	19.0228	21.6660	23.5893
10	2.1558	2.5582	3.2470	3.9403	4.8652	15.9872	18.3070	20.4832	23.2093	25.1881
11	2.6032	3.0535	3.8157	4.5748	5.5778	17.2750	19.6752	21.9200	24.7250	26.7569
12	3.0738	3.5706	4.4038	5.2260	6.3038	18.5493	21.0261	23.3367	26.2170	28.2997
13	3.5650	4.1069	5.0087	5.8919	7.0415	19.8119	22.3620	24.7356	27.6882	29.8193
14	4.0747	4.6604	5.6287	6.5706	7.7895	21.0641	23.6848	26.1189	29.1412	31.3194
15	4.6009	5.2294	6.2621	7.2609	8.5468	22.3071	24.9958	27.4884	30.5780	32.8015
16	5.1422	5.8122	6.9077	7.9616	9.3122	23.5418	26.2962	28.8453	31.9999	34.2671
17	5.6973	6.4077	7.5642	8.6718	10.0852	24.7690	27.5871	30.1910	33.4087	35.7184
18	6.2648	7.0149	8.2307	9.3904	10.8649	25.9894	28.8693	31.5264	34.8052	37.1564
19	6.8439	7.6327	8.9065	10.1170	11.6509	27.2036	30.1435	32.8523	36.1908	38.5821
20	7.4338	8.2604	9.5908	10.8508	12.4426	28.4120	31.4104	34.1696	37.5663	39.9969
21	8.0336	8.8972	10.2829	11.5913	13.2396	29.6151	32.6706	35.4789	38.9322	41.4009
22	8.6427	9.5425	10.9823	12.3380	14.0415	30.8133	33.9245	36.7807	40.2894	42.7957
23	9.2604	10.1957	11.6885	13.0905	14.8480	32.0069	35.1725	38.0756	41.6383	44.1814
24	9.8862	10.8563	12.4011	13.8484	15.6587	33.1962	36.4150	39.3641	42.9798	45.5584
25	10.5196	11.5240	13.1197	14.6114	16.4734	34.3816	37.6525	40.6465	44.3140	46.9280
26	11.1602	12.1982	13.8439	15.3792	17.2919	35.5632	38.8851	41.9231	45.6416	48.2898
27	11.8077	12.8785	14.5734	16.1514	18.1139	36.7412	40.1133	43.1945	46.9628	49.6450
28	12.4613	13.5647	15.3079	16.9279	18.9392	37.9159	41.3372	44.4608	48.2782	50.9936
29	13.1211	14.2564	16.0471	17.7084	19.7677	39.0875	42.5569	45.7223	49.5878	52.3355
30	13.7867	14.9535	16.7908	18.4927	20.5992	40.2560	43.7730	46.9792	50.8922	53.6719

Appendix 4: t-test Probability Values

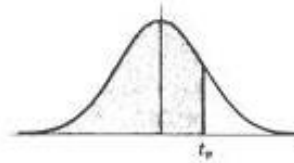
Values of t for
Selected Probabilities



Probabilities (Or Areas Under t -Distribution Curve)

Conf. Level One Tail Two Tails	0.1 0.45 0.9	0.3 0.35 0.7	0.5 0.25 0.5	0.7 0.15 0.3	0.8 0.1 0.2	0.9 0.05 0.1	0.95 0.025 0.05	0.98 0.01 0.02	0.99 0.005 0.01
$d.f.$	Values of t								
1	0.1584	0.5095	1.0000	1.9626	3.0777	6.3137	12.7062	31.8210	63.6559
2	0.1421	0.4447	0.8165	1.3862	1.8856	2.9200	4.3027	6.9645	9.9250
3	0.1366	0.4242	0.7649	1.2498	1.6377	2.3534	3.1824	4.5407	5.8408
4	0.1338	0.4142	0.7407	1.1896	1.5332	2.1318	2.7765	3.7469	4.6041
5	0.1322	0.4082	0.7267	1.1558	1.4759	2.0150	2.5706	3.3649	4.0321
6	0.1311	0.4043	0.7176	1.1342	1.4398	1.9432	2.4469	3.1427	3.7074
7	0.1303	0.4015	0.7111	1.1192	1.4149	1.8946	2.3646	2.9979	3.4995
8	0.1297	0.3995	0.7064	1.1081	1.3968	1.8595	2.3060	2.8965	3.3554
9	0.1293	0.3979	0.7027	1.0997	1.3830	1.8331	2.2622	2.8214	3.2498
10	0.1289	0.3966	0.6998	1.0931	1.3722	1.8125	2.2281	2.7638	3.1693
11	0.1286	0.3956	0.6974	1.0877	1.3634	1.7959	2.2010	2.7181	3.1058
12	0.1283	0.3947	0.6955	1.0832	1.3562	1.7823	2.1788	2.6810	3.0545
13	0.1281	0.3940	0.6938	1.0795	1.3502	1.7709	2.1604	2.6503	3.0123
14	0.1280	0.3933	0.6924	1.0763	1.3450	1.7613	2.1448	2.6245	2.9768
15	0.1278	0.3928	0.6912	1.0735	1.3406	1.7531	2.1315	2.6025	2.9467
16	0.1277	0.3923	0.6901	1.0711	1.3368	1.7459	2.1199	2.5835	2.9208
17	0.1276	0.3919	0.6892	1.0690	1.3334	1.7396	2.1098	2.5669	2.8982
18	0.1274	0.3915	0.6884	1.0672	1.3304	1.7341	2.1009	2.5524	2.8784
19	0.1274	0.3912	0.6876	1.0655	1.3277	1.7291	2.0930	2.5395	2.8609
20	0.1273	0.3909	0.6870	1.0640	1.3253	1.7247	2.0860	2.5280	2.8453
21	0.1272	0.3906	0.6864	1.0627	1.3232	1.7207	2.0796	2.5176	2.8314
22	0.1271	0.3904	0.6858	1.0614	1.3212	1.7171	2.0739	2.5083	2.8188
23	0.1271	0.3902	0.6853	1.0603	1.3195	1.7139	2.0687	2.4999	2.8073
24	0.1270	0.3900	0.6848	1.0593	1.3178	1.7109	2.0639	2.4922	2.7970
25	0.1269	0.3898	0.6844	1.0584	1.3163	1.7081	2.0595	2.4851	2.7874
26	0.1269	0.3896	0.6840	1.0575	1.3150	1.7056	2.0555	2.4786	2.7787
27	0.1268	0.3894	0.6837	1.0567	1.3137	1.7033	2.0518	2.4727	2.7707
28	0.1268	0.3893	0.6834	1.0560	1.3125	1.7011	2.0484	2.4671	2.7633
29	0.1268	0.3892	0.6830	1.0553	1.3114	1.6991	2.0452	2.4620	2.7564
30	0.1267	0.3890	0.6828	1.0547	1.3104	1.6973	2.0423	2.4573	2.7500
40	0.1265	0.3881	0.6807	1.0500	1.3031	1.6839	2.0211	2.4233	2.7045
50	0.1263	0.3875	0.6794	1.0473	1.2987	1.6759	2.0086	2.4033	2.6778
60	0.1262	0.3872	0.6786	1.0455	1.2958	1.6706	2.0003	2.3901	2.6603
70	0.1261	0.3869	0.6780	1.0442	1.2938	1.6669	1.9944	2.3808	2.6479
80	0.1261	0.3867	0.6776	1.0432	1.2922	1.6641	1.9901	2.3739	2.6387
90	0.1260	0.3866	0.6772	1.0424	1.2910	1.6620	1.9867	2.3685	2.6316
100	0.1260	0.3864	0.6770	1.0418	1.2901	1.6602	1.9840	2.3642	2.6259
250	0.1258	0.3858	0.6755	1.0386	1.2849	1.6510	1.9695	2.3414	2.5956
500	0.1257	0.3855	0.6750	1.0375	1.2832	1.6479	1.9647	2.3338	2.5857
∞	0.1257	0.3853	0.6745	1.0364	1.2816	1.6449	1.9600	2.3263	2.5758

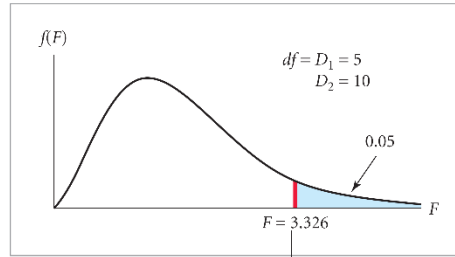
PERCENTILE VALUES (t_p)
for
STUDENT'S t DISTRIBUTION
with degrees of freedom
(shaded area = p)



df	$t_{0.995}$	$t_{0.99}$	$t_{0.975}$	$t_{0.95}$	$t_{0.90}$	$t_{0.80}$	$t_{0.75}$	$t_{0.70}$	$t_{0.60}$	$t_{0.55}$
1	63.66	31.82	12.71	6.31	3.08	1.376	1.000	0.727	0.325	0.158
2	9.92	6.96	4.30	2.92	1.89	1.061	0.816	0.617	0.289	0.142
3	5.84	4.54	3.18	2.35	1.64	0.978	0.765	0.584	0.277	0.137
4	4.60	3.75	2.78	2.13	1.53	0.941	0.741	0.569	0.271	0.134
5	4.03	3.36	2.57	2.02	1.48	0.920	0.727	0.559	0.267	0.132
6	3.71	3.14	2.45	1.94	1.44	0.906	0.718	0.553	0.265	0.131
7	3.50	3.00	2.36	1.90	1.42	0.896	0.711	0.549	0.263	0.130
8	3.36	2.90	2.31	1.86	1.40	0.889	0.706	0.546	0.262	0.130
9	3.25	2.82	2.26	1.83	1.38	0.883	0.703	0.543	0.261	0.129
10	3.17	2.76	2.23	1.81	1.37	0.879	0.700	0.542	0.260	0.129
11	3.11	2.72	2.20	1.80	1.36	0.876	0.697	0.540	0.260	0.129
12	3.06	2.68	2.18	1.78	1.36	0.873	0.695	0.539	0.259	0.128
13	3.01	2.65	2.16	1.77	1.35	0.870	0.694	0.538	0.259	0.128
14	2.98	2.62	2.14	1.76	1.34	0.868	0.692	0.537	0.258	0.128
15	2.95	2.60	2.13	1.75	1.34	0.866	0.691	0.536	0.258	0.128
16	2.92	2.58	2.12	1.75	1.34	0.865	0.690	0.535	0.258	0.128
17	2.90	2.57	2.11	1.74	1.33	0.863	0.689	0.534	0.257	0.128
18	2.88	2.55	2.10	1.73	1.33	0.862	0.688	0.534	0.257	0.127
19	2.86	2.54	2.09	1.73	1.33	0.861	0.688	0.533	0.257	0.127
20	2.84	2.53	2.09	1.72	1.32	0.860	0.687	0.533	0.257	0.127
21	2.83	2.52	2.08	1.72	1.32	0.859	0.686	0.532	0.257	0.127
22	2.82	2.51	2.07	1.72	1.32	0.858	0.686	0.532	0.256	0.127
23	2.81	2.50	2.07	1.71	1.32	0.858	0.685	0.532	0.256	0.127
24	2.80	2.49	2.06	1.71	1.32	0.857	0.685	0.531	0.256	0.127
25	2.79	2.48	2.06	1.71	1.32	0.856	0.684	0.531	0.256	0.127
26	2.78	2.48	2.06	1.71	1.32	0.856	0.684	0.531	0.256	0.127
27	2.77	2.47	2.05	1.70	1.31	0.855	0.684	0.531	0.256	0.127
28	2.76	2.47	2.05	1.70	1.31	0.855	0.683	0.530	0.256	0.127
29	2.76	2.46	2.04	1.70	1.31	0.854	0.683	0.530	0.256	0.127
30	2.75	2.46	2.04	1.70	1.31	0.854	0.683	0.530	0.256	0.127
40	2.70	2.42	2.02	1.68	1.30	0.851	0.681	0.529	0.255	0.126
60	2.66	2.39	2.00	1.67	1.30	0.848	0.679	0.527	0.254	0.126
120	2.62	2.36	1.98	1.66	1.29	0.845	0.677	0.526	0.254	0.126
∞	2.58	2.33	1.96	1.645	1.28	0.842	0.674	0.524	0.253	0.126

Appendix 5: F-test Probability Values

**F-Distribution
Table: Upper 5%
Probability (or 5%
Area) under
F-Distribution
Curve**



DENOMINATOR $df = D_2$		NUMERATOR $df = D_1$									
		1	2	3	4	5	6	7	8	9	10
1		161.446	199.499	215.707	224.583	230.160	233.988	236.767	238.884	240.543	241.882
2		18.513	19.000	19.164	19.247	19.296	19.329	19.353	19.371	19.385	19.396
3		10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845	8.812	8.785
4		7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041	5.999	5.964
5		6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735
6		5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147	4.099	4.060
7		5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.677	3.637
8		5.318	4.459	4.066	3.838	3.688	3.581	3.500	3.438	3.388	3.347
9		5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230	3.179	3.137
10		4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072	3.020	2.978
11		4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948	2.896	2.854
12		4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753
13		4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671
14		4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.646	2.602
15		4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641	2.588	2.544
16		4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.538	2.494
17		4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450
18		4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412
19		4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378
20		4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447	2.393	2.348
24		4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355	2.300	2.255
30		4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266	2.211	2.165
40		4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180	2.124	2.077
50		4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130	2.073	2.026
100		3.936	3.087	2.696	2.463	2.305	2.191	2.103	2.032	1.975	1.927
200		3.888	3.041	2.650	2.417	2.259	2.144	2.056	1.985	1.927	1.878
300		3.873	3.026	2.635	2.402	2.244	2.129	2.040	1.969	1.911	1.862

DENOMINATOR $df = D_2$		NUMERATOR $df = D_1$									
		11	12	13	14	15	16	17	18	19	20
1		242.981	243.905	244.690	245.363	245.949	246.466	246.917	247.324	247.688	248.016
2		19.405	19.412	19.419	19.424	19.429	19.433	19.437	19.440	19.443	19.446
3		8.763	8.745	8.729	8.715	8.703	8.692	8.683	8.675	8.667	8.660
4		5.936	5.912	5.891	5.873	5.858	5.844	5.832	5.821	5.811	5.803
5		4.704	4.678	4.655	4.636	4.619	4.604	4.590	4.579	4.568	4.558
6		4.027	4.000	3.976	3.956	3.938	3.922	3.908	3.896	3.884	3.874
7		3.603	3.575	3.550	3.529	3.511	3.494	3.480	3.467	3.455	3.445
8		3.313	3.284	3.259	3.237	3.218	3.202	3.187	3.173	3.161	3.150
9		3.102	3.073	3.048	3.025	3.006	2.989	2.974	2.960	2.948	2.936
10		2.943	2.913	2.887	2.865	2.845	2.828	2.812	2.798	2.785	2.774
11		2.818	2.788	2.761	2.739	2.719	2.701	2.685	2.671	2.658	2.646
12		2.717	2.687	2.660	2.637	2.617	2.599	2.583	2.568	2.555	2.544
13		2.635	2.604	2.577	2.554	2.533	2.515	2.499	2.484	2.471	2.459
14		2.565	2.534	2.507	2.484	2.463	2.445	2.428	2.413	2.400	2.388
15		2.507	2.475	2.448	2.424	2.403	2.385	2.368	2.353	2.340	2.328
16		2.456	2.425	2.397	2.373	2.352	2.333	2.317	2.302	2.288	2.276

Values of F Exceeded with Probabilities of 5 and 1 Percent

	df (numerator)																																																		
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞																											
df (denominator)	1	161	200	216	225	230	234	237	239	241	242	243	244	245	246	248	249	250	251	252	253	253	254	254	254	4.052	4.999	5.403	5.625	5.764	5.859	5.928	5.981	6.022	6.056	6.082	6.106	6.142	6.169	6.208	6.234	6.261	6.286	6.302	6.323	6.344	6.352	6.361	6.366		
	2	18.51	19.00	19.16	19.25	19.30	19.33	19.36	19.37	19.38	19.39	19.40	19.41	19.42	19.43	19.44	19.45	19.46	19.47	19.47	19.48	19.49	19.49	19.50	19.50	98.49	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.41	99.42	99.43	99.44	99.45	99.46	99.47	99.48	99.48	99.49	99.49	99.50	99.50	99.50	99.50	
	3	10.13	9.55	9.28	9.12	9.01	8.94	8.88	8.84	8.81	8.78	8.76	8.74	8.71	8.69	8.66	8.64	8.62	8.60	8.58	8.57	8.56	8.56	8.54	8.53	34.12	30.82	29.46	28.21	27.24	26.41	25.71	27.67	27.49	27.34	27.23	27.13	27.05	26.92	26.83	26.69	26.60	26.50	26.41	26.35	26.26	26.18	26.14	26.12	26.12	
	4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.93	5.91	5.87	5.84	5.80	5.77	5.74	5.71	5.70	5.68	5.66	5.65	5.64	5.63	21.20	18.90	16.69	15.58	15.22	15.01	14.98	14.80	14.66	14.54	14.45	14.37	14.24	14.15	14.02	13.93	13.83	13.74	13.69	13.61	13.57	13.52	13.48	13.46	13.46	
	5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.78	4.74	4.70	4.68	4.64	4.60	4.56	4.53	4.50	4.46	4.44	4.42	4.40	4.38	4.37	4.36	16.26	13.27	12.06	11.39	10.97	10.67	10.45	10.29	10.15	10.05	9.96	9.89	9.77	9.68	9.55	9.47	9.38	9.29	9.24	9.17	9.13	9.07	9.04	9.02	9.02	
	6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.03	4.00	3.96	3.92	3.87	3.84	3.81	3.77	3.75	3.72	3.71	3.69	3.68	3.67	13.74	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.79	7.72	7.60	7.52	7.39	7.31	7.23	7.14	7.09	7.02	6.99	6.94	6.90	6.88	6.88	
	7	5.59	4.74	4.34	4.12	3.97	3.87	3.79	3.73	3.68	3.63	3.60	3.57	3.52	3.49	3.44	3.41	3.38	3.34	3.32	3.29	3.28	3.25	3.24	3.23	12.25	9.55	8.45	7.85	7.46	7.19	7.00	6.84	6.71	6.62	6.54	6.47	6.35	6.27	6.15	6.07	5.96	5.90	5.85	5.78	5.75	5.70	5.67	5.65	5.65	
	8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.34	3.31	3.28	3.23	3.20	3.15	3.12	3.08	3.05	3.03	3.00	2.98	2.96	2.94	2.93	11.26	8.65	7.59	7.01	6.63	6.37	6.19	6.03	5.91	5.82	5.74	5.67	5.56	5.48	5.36	5.28	5.20	5.11	5.06	5.00	4.96	4.91	4.88	4.86	4.86	
	9	5.12	4.25	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.13	3.10	3.07	3.02	2.98	2.93	2.90	2.86	2.82	2.80	2.77	2.76	2.73	2.72	2.71	10.56	8.02	6.99	6.42	6.06	5.80	5.62	5.47	5.35	5.26	5.18	5.11	5.00	4.92	4.80	4.73	4.64	4.56	4.51	4.45	4.41	4.36	4.33	4.31	4.31	4.31
	10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.97	2.94	2.91	2.86	2.82	2.77	2.74	2.70	2.67	2.64	2.61	2.59	2.56	2.55	2.54	10.04	7.56	6.55	5.99	5.64	5.39	5.21	5.06	4.95	4.85	4.78	4.71	4.60	4.52	4.41	4.33	4.25	4.17	4.12	4.05	4.01	3.96	3.93	3.91	3.91	3.91
	11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.86	2.82	2.79	2.74	2.70	2.65	2.61	2.57	2.53	2.50	2.47	2.45	2.42	2.41	2.40	9.65	7.20	6.22	5.67	5.32	5.07	4.88	4.74	4.63	4.54	4.46	4.40	4.29	4.21	4.10	4.02	3.94	3.86	3.80	3.74	3.70	3.66	3.62	3.60	3.60	
	12	4.75	3.88	3.49	3.26	3.11	3.00	2.92	2.85	2.80	2.76	2.72	2.69	2.64	2.60	2.54	2.50	2.46	2.42	2.40	2.36	2.35	2.32	2.31	2.30	9.33	6.93	5.95	5.41	5.06	4.80	4.62	4.50	4.39	4.30	4.22	4.16	4.05	3.98	3.86	3.78	3.70	3.61	3.56	3.49	3.46	3.41	3.38	3.36	3.36	
	13	4.67	3.80	3.41	3.18	3.02	2.92	2.84	2.77	2.72	2.67	2.63	2.60	2.55	2.51	2.46	2.42	2.38	2.34	2.32	2.28	2.26	2.24	2.22	2.21	9.07	6.70	5.74	5.20	4.86	4.62	4.44	4.30	4.19	4.10	4.02	3.96	3.85	3.78	3.67	3.59	3.51	3.42	3.37	3.30	3.27	3.21	3.18	3.16	3.16	
	14	4.60	3.74	3.34	3.11	2.96	2.85	2.77	2.70	2.65	2.60	2.56	2.53	2.48	2.44	2.39	2.35	2.31	2.27	2.24	2.21	2.19	2.16	2.14	2.13	8.86	6.51	5.56	5.03	4.69	4.46	4.28	4.14	4.03	3.94	3.86	3.80	3.70	3.62	3.51	3.43	3.34	3.26	3.21	3.14	3.11	3.06	3.02	3.00	3.00	
	15	4.54	3.68	3.29	3.06	2.90	2.79	2.70	2.64	2.59	2.55	2.51	2.48	2.43	2.39	2.33	2.29	2.25	2.21	2.18	2.15	2.12	2.10	2.08	2.07	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.73	3.67	3.56	3.48	3.36	3.29	3.20	3.12	3.07	3.00	2.97	2.92	2.89	2.87	2.87	
	16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.45	2.42	2.37	2.33	2.28	2.24	2.20	2.16	2.13	2.09	2.07	2.04	2.02	2.01	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.61	3.55	3.45	3.37	3.25	3.18	3.10	3.01	2.96	2.98	2.86	2.80	2.77	2.75	2.75	
	17	4.45	3.59	3.20	2.96	2.81	2.70	2.62	2.55	2.50	2.45	2.41	2.38	2.33	2.29	2.23	2.19	2.15	2.11	2.08	2.04	2.02	1.99	1.97	1.96	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.52	3.45	3.35	3.27	3.16	3.08	3.00	2.92	2.86	2.79	2.76	2.70	2.67	2.65	2.65	
	18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.37	2.34	2.29	2.25	2.19	2.15	2.11	2.07	2.04	2.00	1.98	1.95	1.93	1.92	8.28	6.01	5.09	4.58	4.25	4.01	3.85	3.71	3.60	3.51	3.44	3.37	3.27	3.19	3.07	3.00	2.91	2.83	2.78	2.71	2.68	2.62	2.59	2.57	2.57	

(continued)

		df (numerator)																												
		1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞					
df (denominator)	19	4.38	3.52	3.13	2.90	2.74	2.63	2.55	2.48	2.43	2.38	2.34	2.31	2.26	2.21	2.15	2.11	2.07	2.02	2.00	1.96	1.94	1.91	1.90	1.88	1.88	1.88	1.88	1.88	1.88
	20	4.35	3.49	3.10	2.87	2.71	2.60	2.52	2.45	2.40	2.35	2.31	2.28	2.23	2.18	2.12	2.08	2.04	1.99	1.96	1.92	1.90	1.87	1.85	1.84	1.84	1.84	1.84	1.84	1.84
	21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.28	2.25	2.20	2.15	2.09	2.05	2.00	1.96	1.93	1.89	1.87	1.84	1.82	1.81	1.81	1.81	1.81	1.81	1.81
	22	4.30	3.44	3.05	2.82	2.66	2.55	2.47	2.40	2.35	2.30	2.26	2.23	2.18	2.13	2.07	2.03	1.98	1.93	1.91	1.87	1.84	1.81	1.80	1.78	1.78	1.78	1.78	1.78	1.78
	23	4.28	3.42	3.03	2.80	2.64	2.53	2.45	2.38	2.32	2.28	2.24	2.20	2.14	2.10	2.04	2.00	1.96	1.91	1.88	1.84	1.82	1.79	1.77	1.76	1.76	1.76	1.76	1.76	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.43	2.36	2.30	2.26	2.22	2.18	2.13	2.09	2.02	1.98	1.94	1.89	1.86	1.82	1.80	1.76	1.74	1.73	1.73	1.73	1.73	1.73	1.73	1.73
25	4.24	3.38	2.99	2.76	2.60	2.49	2.41	2.34	2.28	2.24	2.20	2.16	2.11	2.06	2.00	1.96	1.92	1.87	1.84	1.80	1.77	1.74	1.72	1.71	1.71	1.71	1.71	1.71	1.71	1.71
26	4.22	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15	2.10	2.05	1.99	1.95	1.90	1.85	1.82	1.78	1.76	1.72	1.70	1.69	1.69	1.69	1.69	1.69	1.69	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.30	2.25	2.20	2.16	2.13	2.08	2.03	1.97	1.93	1.88	1.84	1.80	1.76	1.74	1.71	1.69	1.68	1.68	1.68	1.68	1.68	1.68	1.68
28	4.19	3.34	2.95	2.72	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12	2.06	2.02	1.96	1.91	1.87	1.81	1.78	1.75	1.72	1.69	1.67	1.66	1.66	1.66	1.66	1.66	1.66	1.66
29	4.18	3.33	2.93	2.70	2.54	2.43	2.35	2.28	2.22	2.18	2.14	2.10	2.05	2.00	1.94	1.90	1.85	1.80	1.77	1.73	1.71	1.68	1.65	1.64	1.64	1.64	1.64	1.64	1.64	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.34	2.27	2.21	2.16	2.12	2.09	2.04	1.99	1.93	1.89	1.84	1.79	1.76	1.72	1.69	1.66	1.64	1.62	1.62	1.62	1.62	1.62	1.62	1.62
40	4.15	3.30	2.90	2.67	2.51	2.40	2.32	2.25	2.19	2.14	2.10	2.07	2.02	1.97	1.91	1.86	1.82	1.76	1.74	1.69	1.67	1.64	1.61	1.59	1.58	1.58	1.58	1.58	1.58	1.58
50	4.14	3.28	2.88	2.65	2.49	2.38	2.30	2.23	2.17	2.12	2.08	2.05	2.00	1.95	1.89	1.84	1.80	1.74	1.71	1.67	1.64	1.61	1.59	1.57	1.57	1.57	1.57	1.57	1.57	1.57
75	4.11	3.26	2.85	2.63	2.47	2.36	2.28	2.21	2.15	2.10	2.06	2.03	1.98	1.93	1.87	1.82	1.78	1.72	1.69	1.65	1.62	1.59	1.57	1.55	1.55	1.55	1.55	1.55	1.55	1.55
100	4.10	3.25	2.85	2.62	2.46	2.35	2.26	2.19	2.14	2.09	2.05	2.02	1.96	1.92	1.85	1.80	1.76	1.71	1.67	1.63	1.60	1.57	1.54	1.53	1.53	1.53	1.53	1.53	1.53	1.53
200	4.08	3.23	2.83	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00	1.95	1.90	1.84	1.79	1.74	1.69	1.65	1.61	1.58	1.55	1.53	1.51	1.51	1.51	1.51	1.51	1.51	1.51
500	4.07	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.07	2.04	2.00	1.95	1.90	1.84	1.79	1.74	1.69	1.65	1.61	1.58	1.55	1.53	1.51	1.51	1.51	1.51	1.51	1.51	1.51

	df (numerator)																									
	1	2	3	4	5	6	7	8	9	10	11	12	14	16	20	24	30	40	50	75	100	200	500	∞		
42	4.07	3.22	2.83	2.59	2.44	2.32	2.24	2.17	2.11	2.06	2.02	1.99	1.94	1.89	1.82	1.78	1.73	1.68	1.64	1.60	1.57	1.54	1.51	1.49		
	7.27	5.15	4.29	3.80	3.49	3.26	3.10	2.96	2.86	2.77	2.70	2.64	2.54	2.46	2.35	2.26	2.17	2.08	2.02	1.94	1.91	1.85	1.80	1.78		
44	4.06	3.21	2.82	2.58	2.43	2.31	2.23	2.16	2.10	2.05	2.01	1.98	1.92	1.88	1.81	1.75	1.72	1.66	1.63	1.58	1.56	1.52	1.50	1.48		
	7.24	5.12	4.26	3.78	3.48	3.24	3.07	2.94	2.84	2.75	2.68	2.62	2.52	2.44	2.32	2.24	2.15	2.06	2.00	1.92	1.88	1.82	1.78	1.75		
46	4.05	3.20	2.81	2.57	2.42	2.30	2.22	2.14	2.09	2.04	2.00	1.97	1.91	1.87	1.80	1.75	1.71	1.65	1.62	1.57	1.54	1.51	1.48	1.46		
	7.21	5.10	4.24	3.76	3.44	3.22	3.05	2.92	2.82	2.73	2.66	2.60	2.50	2.42	2.30	2.22	2.13	2.04	1.98	1.90	1.85	1.80	1.76	1.72		
48	4.04	3.19	2.80	2.56	2.41	2.30	2.21	2.14	2.08	2.03	1.99	1.96	1.90	1.86	1.79	1.74	1.70	1.64	1.61	1.56	1.53	1.50	1.47	1.45		
	7.19	5.08	4.22	3.74	3.42	3.20	3.04	2.90	2.80	2.71	2.64	2.58	2.48	2.40	2.28	2.20	2.11	2.02	1.96	1.88	1.84	1.78	1.73	1.70		
50	4.03	3.18	2.79	2.55	2.40	2.29	2.20	2.13	2.07	2.02	1.98	1.95	1.90	1.85	1.78	1.74	1.69	1.63	1.60	1.55	1.52	1.50	1.48	1.46		
	7.17	5.06	4.20	3.72	3.41	3.18	3.02	2.88	2.78	2.70	2.62	2.56	2.46	2.39	2.26	2.18	2.10	2.00	1.94	1.86	1.82	1.76	1.71	1.68		
55	4.02	3.17	2.78	2.54	2.38	2.27	2.18	2.11	2.05	2.00	1.97	1.93	1.88	1.83	1.76	1.72	1.67	1.61	1.58	1.52	1.50	1.46	1.43	1.41		
	7.12	5.01	4.16	3.68	3.37	3.15	2.98	2.85	2.75	2.66	2.59	2.53	2.43	2.35	2.23	2.15	2.06	1.96	1.90	1.82	1.78	1.71	1.66	1.64		
60	4.00	3.15	2.76	2.52	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92	1.86	1.81	1.75	1.70	1.65	1.59	1.56	1.50	1.48	1.44	1.41	1.39		
	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50	2.40	2.32	2.20	2.12	2.03	1.93	1.87	1.79	1.74	1.68	1.63	1.60		
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.02	1.98	1.94	1.90	1.85	1.80	1.73	1.68	1.63	1.57	1.54	1.49	1.46	1.42	1.39	1.37		
	7.04	4.95	4.10	3.62	3.31	3.09	2.93	2.80	2.70	2.61	2.54	2.47	2.37	2.30	2.18	2.09	2.00	1.90	1.84	1.76	1.71	1.64	1.60	1.56		
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.01	1.97	1.93	1.89	1.84	1.79	1.72	1.67	1.62	1.56	1.53	1.47	1.45	1.42	1.38	1.35		
	7.01	4.92	4.08	3.60	3.29	3.07	2.91	2.77	2.67	2.58	2.51	2.45	2.35	2.28	2.15	2.07	1.98	1.88	1.82	1.74	1.69	1.62	1.56	1.53		
80	3.96	3.11	2.72	2.48	2.33	2.21	2.12	2.05	1.99	1.95	1.91	1.86	1.82	1.77	1.70	1.65	1.60	1.54	1.51	1.45	1.42	1.38	1.35	1.32		
	6.96	4.88	4.04	3.56	3.25	3.04	2.87	2.74	2.64	2.55	2.48	2.41	2.32	2.24	2.11	2.03	1.94	1.84	1.78	1.70	1.65	1.57	1.52	1.49		
100	3.94	3.09	2.70	2.46	2.30	2.19	2.10	2.03	1.97	1.92	1.88	1.83	1.79	1.75	1.68	1.63	1.57	1.51	1.48	1.42	1.39	1.34	1.30	1.28		
	6.90	4.82	3.98	3.51	3.20	2.99	2.82	2.69	2.59	2.51	2.43	2.36	2.26	2.19	2.06	1.98	1.89	1.79	1.73	1.64	1.59	1.51	1.46	1.43		
125	3.92	3.07	2.68	2.44	2.29	2.17	2.08	2.01	1.95	1.90	1.86	1.83	1.77	1.72	1.66	1.60	1.55	1.49	1.45	1.39	1.36	1.31	1.27	1.25		
	6.84	4.78	3.94	3.47	3.17	2.96	2.79	2.65	2.56	2.47	2.40	2.33	2.23	2.15	2.03	1.94	1.85	1.75	1.68	1.59	1.54	1.46	1.40	1.37		
150	3.91	3.06	2.67	2.43	2.27	2.16	2.07	2.00	1.94	1.89	1.85	1.82	1.76	1.71	1.64	1.59	1.54	1.47	1.44	1.37	1.34	1.29	1.25	1.22		
	6.81	4.75	3.91	3.44	3.14	2.92	2.76	2.62	2.53	2.44	2.37	2.30	2.20	2.12	2.00	1.91	1.83	1.72	1.66	1.56	1.51	1.43	1.37	1.33		
200	3.89	3.04	2.65	2.41	2.25	2.14	2.05	1.98	1.92	1.87	1.83	1.80	1.74	1.69	1.62	1.57	1.52	1.45	1.42	1.35	1.32	1.26	1.22	1.19		
	6.76	4.71	3.88	3.41	3.11	2.90	2.73	2.60	2.50	2.41	2.34	2.28	2.17	2.09	1.97	1.88	1.79	1.69	1.62	1.52	1.48	1.39	1.33	1.28		
400	3.86	3.02	2.63	2.39	2.23	2.12	2.03	1.96	1.90	1.85	1.81	1.78	1.72	1.67	1.60	1.54	1.49	1.42	1.38	1.32	1.28	1.22	1.16	1.13		
	6.70	4.66	3.83	3.36	3.06	2.85	2.69	2.55	2.46	2.37	2.29	2.23	2.12	2.04	1.92	1.84	1.74	1.64	1.57	1.47	1.41	1.36	1.30	1.26		
1000	3.85	3.00	2.61	2.38	2.22	2.10	2.02	1.95	1.89	1.84	1.80	1.76	1.70	1.65	1.58	1.53	1.47	1.41	1.36	1.30	1.26	1.19	1.13	1.08		
	6.66	4.62	3.80	3.34	3.04	2.82	2.66	2.53	2.43	2.34	2.26	2.20	2.09	2.01	1.89	1.81	1.71	1.61	1.54	1.44	1.38	1.28	1.19	1.11		
∞	3.84	2.99	2.60	2.37	2.21	2.09	2.01	1.94	1.88	1.83	1.79	1.75	1.69	1.64	1.57	1.52	1.46	1.40	1.35	1.28	1.24	1.17	1.11	1.00		
	6.64	4.60	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.24	2.18	2.07	1.99	1.87	1.79	1.69	1.59	1.52	1.41	1.36	1.25	1.15	1.00		

df (denominator)

Appendix 6 : Pearson Type III Distribution Table

K value of Pearson Type III distribution

Skew Coeff (G)	Return period (T_r) in years									
	1.010	1.25	2	5	10	25	50	100	200	1000
	Percent chance (p)									
	99	80	50	20	10	4	2	1	0.5	0.1
3.0	-0.667	-0.636	-0.396	0.420	1.180	2.278	3.152	4.051	4.970	7.250
2.8	-0.714	-0.666	-0.384	0.460	1.210	2.275	3.114	3.973	4.847	6.915
2.6	-0.769	-0.696	-0.368	0.499	1.238	2.267	3.071	3.889	4.718	6.672
2.4	-0.832	-0.725	-0.351	0.537	1.262	2.256	3.023	3.800	4.584	6.423
2.2	-0.905	-0.752	-0.330	0.574	1.248	2.240	2.970	3.705	4.444	6.168
2.0	-0.990	-0.777	-0.307	0.609	1.302	2.219	2.912	3.605	4.298	5.908
1.8	-1.087	-0.799	-0.282	0.643	1.318	2.193	2.848	3.499	4.147	5.642
1.6	-1.197	-0.817	-0.254	0.675	1.329	2.163	2.780	3.388	3.990	5.371
1.4	-1.318	-0.832	-0.225	0.705	1.337	2.128	2.706	3.271	3.828	5.095
1.2	-1.449	-0.844	-0.195	0.732	1.340	2.087	2.626	3.149	3.661	4.815
1.0	-1.588	-0.852	-0.164	0.758	1.340	2.043	2.542	3.022	3.489	4.531
0.8	-1.733	-0.856	-0.132	0.780	1.336	1.993	2.453	2.891	3.312	4.244
0.6	-1.880	-0.857	-0.099	0.800	1.328	1.939	2.359	2.755	3.132	3.956
0.4	-2.029	-0.855	-0.066	0.816	1.317	1.880	2.261	2.615	2.949	3.666
0.2	-2.178	-0.850	-0.033	0.830	1.301	1.818	2.159	2.472	2.763	3.377
0.0	-2.326	-0.842	0	0.842	1.282	1.751	2.054	2.326	2.576	3.090
-0.2	-2.472	-0.830	0.033	0.850	1.258	1.680	1.945	2.178	2.388	2.808
-0.4	-2.615	-0.816	0.066	0.855	1.231	1.606	1.834	2.029	2.201	2.533
-0.6	-2.755	-0.800	0.099	0.857	1.200	1.528	1.720	1.880	2.016	2.268
-0.8	-2.891	-0.780	0.132	0.856	1.166	1.448	1.606	1.733	1.837	2.017
-1.0	-3.022	-0.758	0.164	0.852	1.128	1.366	1.492	1.588	1.664	1.786
-1.2	-3.149	-0.732	0.195	0.844	1.086	1.282	1.379	1.449	1.501	1.577
-1.4	-3.271	-0.705	0.225	0.832	1.041	1.198	1.270	1.318	1.351	1.394
-1.6	-3.388	-0.675	0.254	0.817	0.994	1.116	1.166	1.197	1.216	1.238
-1.8	-3.499	-0.643	0.282	0.799	0.945	1.035	1.069	1.087	1.097	1.107
-2.0	-3.605	-0.609	0.307	0.777	0.895	0.959	0.980	0.990	0.995	1.000
-2.2	-3.705	-0.574	0.330	0.752	0.844	0.888	0.900	0.905	0.907	0.909
-2.4	-3.800	-0.537	0.351	0.725	0.795	0.823	0.830	0.832	0.833	0.833
-2.6	-3.889	-0.499	0.368	0.696	0.747	0.764	0.768	0.769	0.769	0.769
-2.8	-3.973	-0.460	0.384	0.666	0.702	0.712	0.714	0.714	0.714	0.714
-3.0	-4.051	-0.420	0.396	0.636	0.660	0.666	0.666	0.667	0.667	0.668