Kwame Nkrumah University of Science and Technology

Estimation and Fitting Rainfall Pattern Distribution in Ghana using the Expectation-Maximization Algorithm with some Probability Models

By

Tuoyintir Alexis (Bsc.Statistics and Actuarial Science)

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF M.PHIL MATHEMATICAL STATISTICS

October 9, 2016

Declaration

I hereby declare that this submission is my own work towards the award of the M. Phil degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of the university, except where due acknowledgment had been made in the text.

Tuoyintir Alexis ....

Signature Date

Student(PG 2521414)
Certified by:

Mr. E Harris ....

Supervisor Signature Date

Certified by:

Dr. R. K Avuglah ....

Head of Department Signature Date

Dedication

This work is whole heartedly dedicated to the Almighty God and the entire Tuoyintir's family.

Acknowledgment

My utmost thanks go to the Almighty God for his divine protection, life and good health. I say glory be to thy name.

I wish to also express my sincere gratitude to my supervisor; Mr. Emmanuel Harris for the suggestions and time spent in supervising this work and all other lecturers in the Mathematics Department of Kwame Nkrumah University of Science and Technology for their guidance and support. I owe a special thanks to Mr. Asubonteng, the Kumasi Metro director of Ghana Meteorological Agency for his wonderful support.

I am again grateful to my senior brother; Mr. Peter Claver Tuoyintir and his wife for their support and prayers. I wish to also express my sincere gratitude to my one and only Madam Grace Kpetu for her warm support and prayers.

Finally, I wish to say thank you to all my course mates especially Ezekiel K. Gyimah for the support and encouragement and to any other persons who assisted me in diverse ways. God bless us all.

Abstract

The study investigated the rainfall pattern estimation involving fifteen selected rainfall stations across Ghana for the period of sixteen years (2000-2015) with the main objective of estimating the missing values and determining the annual average rainfall values for the period under study. The observed data was fitted with three probability distributions which were; Gamma, Lognormal and Normal. Goodness-of-fit tests were conducted in order to select the best model fit. These included Kolgmorov-Smirnov, Cramer-von Mises and Anderson-Darling tests together with the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The normal probability distribution was selected as the best model since it provided the minimum goodness-of-fit test statistic values. The EM algorithm which has the capability to deal with missing values was used to complete the missing data and together with the normal distribution, estimated the average rainfall values of the various stations. The estimates of the EM algorithm were observed to be better estimates for the data because they were smaller than the regular estimates and also provided the least log-likelihood values. Therefore, we recommend that the Expectation-Maximization algorithm (Normal EM algorithm) should be used to estimate the missing values as well as the annual average rainfall of the daily rainfall data recorded in Ghana.

Contents

Declaration . . . . **v**

Dedication . . . . **v**

Acknowledgment . . . . **v**

List of Tables . . . . **ix**

List of Figures . . . . **xii**

1 Introduction **1**

1.1 Background 3

1.1.1 Characteristics of Rainfall 4

1.2 Statement of Problem 5

1.3 Objectives of the study 5

1.3.1 Specific objectives 5

1.4 Structure of Methodology 5

1.5 Significance of the Study 6

1.6 Organization of the Study 6

2 Literature Review **8**

2.1 Introduction 8

2.1.1 The EM algorithm and applications 8

2.1.2 The Concept of Missing Values 11

2.1.3 Rainfall Pattern Estimation 12

3 Methodology **18**

3.1 Introduction 18

3.2 Fitting the Probability Distributions 18

3.2.1 The Gamma Distribution 19

3.2.2 The Log-normal Distribution 20

3.2.3 The Normal Distribution 20

3.3 Testing the Goodness-of-Fit of the Probability Distributions . . . 21

3.3.1 Kolmogorov-Smirnov Test 21

3.3.2 Cramer-von Mises Test 22

3.3.3 Anderson-Darling Test 22

3.3.4 Akaike Information Criterion 23

3.3.5 Bayesian Information Criterion 24

3.4 Maximum-Likelihood Estimation 24

3.5 The Expectation-Maximization (EM) Algorithm 25

3.6 Th Normal EM algorithm 29

4 Data Analysis and Results 31

4.1 Introduction 31

4.2 Data Description 31

4.3 Results and Discussions 33

5 Summary, Conclusion and Recommendations 46

5.1 Introduction 46

5.2 Summary 46

5.3 Conclusion 48

5.4 Recommendations 49

References 55

Appendix A 56

Appendix B 57

Appendix B1 58

Appendix C 59

Appendix C1 60

Appendix D 61

Appendix D1 62

Appendix E 63

Appendix E1 64

Appendix F 65

Appendix F1 66

Appendix G 67

Appendix G1 .... **68**
Appendix H **69**
Appendix H1 .... **70**
Appendix I **71**
Appendix I1 .... **72**
Appendix J **73**
Appendix J1 .... **74**
Appendix K **75**
Appendix K1 .... **76**
Appendix L **77**
Appendix L1 .... **78**
Appendix M **79**
Appendix M1 .... **80**
Appendix N **81**
Appendix N1 .... **82**
Appendix O **83**
Appendix O1 .... **84**
Appendix P **85**
Appendix P1 .... **86**

List of Tables

4.1 Descriptive statistics of daily rainfall data for the selected rainfall stations in Ghana (2000-2015) 32

4.2 Goodness of fit statistics........................ 34

4.3 Average Rainfall for the Stations Estimated using the regular Normal Distribution 35

4.4 Average Rainfall for the Stations Estimated using the Normal EM Algorithm 35

List of Figures

4.1 Bar Chart comparing the annual average rainfall and the EM estimates of all stations (2000-2015) 37

4.2 Bar Chart comparing the standard deviations of the observed values and the EM estimates of all stations (2000-2015) 38

4.3 Distribution plots of the annual average rainfall of Abetifi (2000-2015) 38

4.4 Distribution plots of the annual average rainfall of Ada (2000-2015) 39

4.5 Distribution plots of the annual average rainfall of Akrokerri (2000-2015) 39

4.6 Distribution plots of the annual average rainfall of Axim (2000-2015) 40

4.7 Distribution plots of the annual average rainfall of Bolgatanga (2000-2015) 40

4.8 Distribution plots of the annual average rainfall of Cape Coast (2000-2015) 41

4.9 Distribution plots of the annual average rainfall of Ejura (2000-2015) 41

4.10 Distribution plots of the annual average rainfall of Ho (2000-2015) 42

4.11 Distribution plots of the annual average rainfall of Koforidua (2000-

2015) 42
4.12 Distribution plots of the annual average rainfall of Krachi (2000-2015) 43
4.13 Distribution plots of the annual average rainfall of Salpon (2000-2015) 43
4.14 Distribution plots of the annual average rainfall of Sefwi (2000-2015) 44
4.15 Distribution plots of the annual average rainfall of Temale (2000-2015) 44
4.16 Distribution plots of the annual average rainfall of Wa (2000-2015) 45
4.17 Distribution plots of the annual average rainfall of Wenchi (2000-2015) 45
5.1 Annual Total Rainfall(mm)for all the Stations (2000-2015) 56
5.2 Sample Observed Data for Abetifi (200) 57
5.3 Sample Complete Data for Abetifi (2000) 58
5.4 Sample Observed Data for Ada (2000) 59
5.5 Sample Complete Data for Ada (2000) 60
5.6 Sample Observed Data for Akrokerri (2000) 61
5.7 Sample Complete Data for Akrokerri (2000) 62
5.8 Sample Observed Data for Axim (2000) 63
5.9 Sample Complete Data for Axim (2000) 64
5.10 Sample Observed Data for Bolgatanga (2000) 65
5.11 Sample Complete Data for Bolgatanga (2000) 66
5.12 Sample Observed Data for Cape Coast (2000) 67
5.13 Sample Complete Data for Cape Coast (2000) 68
5.14 Sample Complete Data for Ejura (2000) 69
5.15 Sample Complete Data for Ejura (2000) 70
5.16 Sample Complete Data for Ho (2000) 71
5.17 Sample Complete Data for Ho (2000) 72
5.18 Sample Observed Data for Koforidua (2000) 73
5.19 Sample Complete Data Koforidua (2000) 74

5.20 Sample Observed Data for Kete Krachi (2000) 75
5.21 Sample Complete Data for Kete Krachi (2000) 76
5.22 Sample Observed Data for Saltpond (2000) 77
5.23 Sample Complete Data for Saltpond (2000) 78
5.24 Sample Observed Data for Sefwi-Bekwai (2000) 79
5.25 Sample Complete Data for Sefwi-Bekwai (2000) 80
5.26 Sample Observed Data for Tamale (2000) 81
5.27 Sample Complete Data for Tamale (2000) 82
5.28 Sample Observed Data for Wa (2000) 83
5.29 Sample Complete Data for Wa (2000) 84
5.30 Sample Observed Data for Wenchi (2000) 85
5.31 Sample Complete Data for Wenchi (2000) 86

Chapter 1

Introduction

The Expectation-Maximization algorithm (EM) is a method of maximum likelihood or point estimation. It is an iterative method of estimation that can obtain the maximum likelihood (ML) estimates of data containing missing values or incomplete data. Maximum likelihood estimation is a statistical technique of determining the parameters of probability models. If it is used to estimate a set of data with a particular probability model, maximum likelihood estimation determines estimates for the distribution parameters.

The EM algorithm can be applied in two main ways. One application is when the observations have missing values because of shortfalls or mistakes during the process of observation. The second application is when the function's likelihood estimates can be calculated easily by taking into consideration that there exist extra or hidden parameters.

The Expectation-Maximization algorithm process also involves two stages or steps. That is the Expectation stage or step (E-step) and Maximization stage or step (M-step). The Expectation stage or step obtains the predicted value for the missing observations via an initial parameter estimate. The E-step of the algorithm fills the gaps in the missing or unobserved values with its expected value given a current value of the parameter and the original data. The Maximization step on the other hand calculates the maximum-likelihood estimated values (MLE's) of the parameter using the estimated data obtained from the E-step.

The M-step mostly to some extent is clear and relies on the implicit function or distribution. However, the E-step is complicated depending on the distribution. In case of a linear function or distribution, the predicted conditional probability estimates can be obtained easily and precisely but for complex probability distributions in which precise results do not exist, individuals could continue to perform some correct calculations such as that of Taylor's approximation or arithmetical approximations which include numerical integration and/or simulated form of approximation.

Several applications of the EM algorithm are on unobserved data or incomplete data. The unobserved or incomplete data could be lost information on the same random events which gives the original sample, for censored data; or the unobserved data could be from a different kind of variable that is connected in some way to the observed random event.

It is always assumed that the suggested procedures used in estimating data with missing information can be categorized into two procedures. These are the distribution-based procedure and data-based procedure.

The distribution-based procedure rewrites the statistical algorithm to enhance the estimation of the lost data and determines the estimates of the distribution all in one step. The data-based procedure in another breadth, deals with the lost data in a first stage and proceeds to complete the estimation of the distribution parameters in a second separate step.

An example of the application of distribution-based procedure is in the recent model prediction of crop structure which used the full information maximum likelihood (FIML) method in estimating lost observations. The commonest data-based procedure application is in the normal model multiple imputations (MI).

Meanwhile, with the Expectation-Maximization (EM) algorithm, this distinctive characteristic is not that clear. If the EM algorithm is modified to give parameter estimates in connection with a particular situation, then the EM is a distribution-based approach. Also, where the EM algorithm gives more standard outcomes like the variance and covariance matrix as well as vector of means which are then analyzed in a different step, then the algorithm becomes more or less a data-based approach.

1.1 Background

Rainfall is a factor of climate change which affects almost everybody in the world. Obtaining good estimates of the rainfall pattern within a year or period is very important.

Rainfall is a kind of climate change which takes place when water vapor within the atmosphere condenses into small droplets that can no longer be suspended in the air. The incidence of rain depends on so many factors. Events such as existing wind movement and direction, elevation of the ground, continental mass location, and positions with regards to the height of mountains all posed a major impact on the likelihood of rainfall. Condensation, and therefore precipitation, could take place at the time that water vapor in the presence of accumulated air is condensed. Rain is the simplest form or supply of freshwater for almost every area in the world, providing appropriate environment for different ecosystems and also supporting hydroelectric plants and crops irrigation.

Also,the amount of rainfall is measured by using a rain gauge. It can be calibrated in 100 millimeters (4-inches) for the plastic gauge and 200 millimeters (8-inches) for the metal gauge. The interior or middle cylinder is full up to 25 millimeters (0.98-inches) of rainfall, with the excess pouring into the exterior or external cylinder. Plastic gauges are calibrated on the interior cylinder down to 0.25 millimeters (0.0098-inches) resolution. The metal gauges also contain a standard stick designed with the suitable 0.25 millimeters (0.0098-inches) calibrations. When the internal cylinder is full, the quantity inside is poured out together with the left over rainfall in the external cylinder till every liquid in the internal cylinder move out and the external cylinder is drained. Rain gauges come in several types. Some of these include the wedge gauge, the tipping bucket rain gauge and the weighting rain gauge. Any of these can be used to measure rainfall accurately.

1.1.1 Characteristics of Rainfall

The amount of rainfall is determined by certain characteristics such as the intensity of rain, the duration and the frequency of the rain.

XIntensity: The intensity of rainfall is the degree or the amount of rain that have precipitated within a given period of time. Generally the amount of rainfall is recorded in millimeters per hour or inches per hour. An extremely strong storm may produce rain at a rate of 5 inches up to 10 inches per hour whilst a calm drizzle may produce rain at a rate of 0.2 inches up to 0.4 inches per hour. The rainfall intensity is one of the most significant factors engineers consider when designing highways and structures for controlling floods.

XDuration:The duration of rainfall is the number of minutes or hours that a rainfall period last. Usually, long period storms produces low average intensities. Higher intensities are normally related to shorter period storms.

XFrequency:The frequency of rainfall is the likelihood or chance of rainfall occurring. It is also described as the number of times or how regular a particular area experiences rainfall. With a certain storm period, the likelihood that a rainfall occurrence could be the same or exceeds an annual term is termed the re-occurrence interval. The reciprocal of the recurrence interval is also termed the return period.

In the southern part of Ghana, there exist two rainfall seasons: the first from the third month to the seventh month of the year and the second from the ninth month to the tenth month of the year. The southern part records the highest amount of rainfall, with most towns and villages' averagely experiencing rainfall of about 2000 millimeters every year. However, in the northern part of the country, the rainfall season is one period of wet and damp weather, beginning in the third month and ending somewhere in the ninth month of the year with most areas typically receiving rainfall of about 800 millimeters annually.

1.2 Statement of Problem

In most cases, we apply various tools in analyzing data without understanding the structure and pattern of the data. Sometimes, there exist a lot of missing values (latent) or incomplete information in the data subject to the analysis. Example; rainfall records always have some unrecorded or missing information or values due to continuous rain within two or more days, malfunctioning of rain gauge and officer absenteeism or office reshuffle. Records of survival data also encounter several missing values due to the demise of the individual involved.

Because these records may contain some missing but relevant information, ignoring may result in some biased results. Therefore, methods such as the EM algorithm which is capable of dealing with missing values and incomplete data would be appropriate for estimating these incomplete data parameters.

1.3 Objectives of the study

The main objective of this study is to estimates the missing values of the rainfall data in Ghana using the EM algorithm. It specifically seeks;

1.3.1 Specific objectives

1. To apply the EM algorithm to a probability distribution model to estimatethe annual average rainfall of some parts of Ghana.
2. To determine if a difference exist between the EM estimates and the true values of rainfall data containing missing values recorded by the Ghana Meteorological Agency.

1.4 Structure of Methodology

This research was carried out in Ghana. The main data used for this research is a secondary data consisting of daily rainfall figures recorded in millimeters (mm) for ten (10) year period from 1st January, 2006 to 31st December, 2015 was collected from the Ghana Meteorological Agency (GMET), Accra. This data was also sampled from twenty (20) rainfall stations all over the country out of which fifteen was used in the analysis.

A descriptive breakdown of the sampled data was carried out by calculating the means, medians, minimum and maximum values as well as the standard deviations. A goodness-of-fit test was conducted to select the best probability distribution fit for the data. The EM algorithm was applied on the normal distribution model to estimate the parameters.

The descriptive analysis and the EM estimates were performed using the R console and SPSS statistical software packages.

1.5 Significance of the Study

The results and findings of this research would be very useful to the Ghana Meteorological Service Agency and the general public. It would serve as a guide to the choice of the method to use in the analysis of data with missing values such as the rainfall pattern.

The results and findings would also explain the ideas of the EM algorithm, its application to the some probability models and how effective it is in the analysis of incomplete data. It would further contribute to the existing knowledge and literature in the field of academia and research.

Finally, the results and findings of this research would provide a foundation for further research work in similar areas or fields of study.

1.6 Organization of the Study

This research consists of five chapters. Chapter one considers a general overview of how the study was carried out and reported. It includes the background of the study, statement of the problem, objectives of the study, data collection procedure, and significance of the study. Chapter two reviews related literature based on the objective of the study and the methods used in achieving these objectives. Chapter three focuses on the various ideas with regards to the formulation of the statistical tool and model used in the analysis of the data. The fourth chapter considers the data sampling procedure, and analysis of the results. The five chapter then concludes the whole study by providing some recommendations to stakeholders based on the findings obtained from the study.

Chapter 2

Literature Review

2.1 Introduction

In this chapter, we discussed the various available literature related to the applications of the EM algorithm as a method of estimation. Under this section, we consider the following.

XThe history behind the formulation of EM algorithm.

XSome related applications of EM algorithm.

XThe concept of missing values in a data.

XThe rainfall pattern estimation.

The EM algorithm is an idea developed under point estimation to determine the estimates of data with missing values or incomplete data. It is applied to probability functions/mixtures to determine the estimates of these functions.

It is proven to be more efficient than the main maximum likelihood estimator (MLE).

The trend of rainfall pattern is the gradual change in the amount of rainfall recorded within a period of time. Rainfall is a very important factor with regards to climate change. The goodness of fit test is the degree of compatibility of a random sample with the hypothetical distribution models.

These three aspects in this chapter would summarize the most relevant literature with regards to this study.

2.1.1 The EM algorithm and applications

The expectation-maximization (EM) algorithm according to Dempster et al.

(1977) is a strong mechanism of maximum likelihood estimation of data with missing values or incomplete data. It is aimed at finding the values of a parameter which maximized a particular function given the real sample observations. The algorithm procedure is made up of two stages or steps. The Expectation stage (E-step) and Maximization stage (M-step). The Expectation stage estimates sufficient statistics in terms of complete data with regards to the observed data. The Maximization stage now continues with the estimated whole dataset to estimates the parameters using the method of maximum likelihood as if the re-estimated complete data were the original observed data.

The EM algorithm, just as other procedures of estimating missing values that neglects the pattern that causes the gaps within the data set, relies on the notion that the unobserved data in the data set are randomly missing or missing at random, such that the probability of a missing value is not dependent on the unobserved data (Rubin, 1976).

Just like the other methods of maximum likelihood estimation,Mclachlan and Krishnan (2008) stated that the EM algorithm is a statistical procedure for determining the zeros of the functions. It is a general algorithm that devised a step by step method of estimation of MLE's of data that contain missing values or where there exist missing values in a data or incomplete data. The expectation procedure is responsible for estimating data for the complete data problem taken into consideration the observed dataset of the incomplete data and the existing values of the parameters. The maximization step computes the estimates of the complete data from the E-step. This is the log-likelihood of the complete data problem that is ”manufactured” in the E-step. Suitable initial values of the parameters are chosen at the beginning and the E-step is perform alongside the M-step until the process converges. The sample data set is assumed to be incomplete and is taken to be the observable function of the so-called complete dataset. The idea of incomplete data considers the conventional sense of missing data and it is also applied to problems in which the complete data represents the data that would be available from some theoretical experiments.

According to Karlis (2005), the EM algorithm is a simple procedure for determining the maximum-likelihood (ML) estimation that uses the inherent latent structure of mixture models. He explained that the applications are appropriate for density models appearing as mixtures due to the missing operation, could be regarded as producing data with incomplete information. He also pointed out that the most relevant aspect of the EM algorithm is that it is not just a numerical procedure but also gives relevant statistical ideas. He however concluded that the E-step is not straightforward since at this step the expected likelihood of the incomplete observation distribution is determined where the expected values are taken in respect to the conditional models and that the M-step could be somewhat obvious because it maximizes the parameters of the log-likelihood obtain in the E-step in order to obtain updated estimates.

Redner and Walker (1984), and Jordan and Jacob (1994) described the EM algorithm as a procedure for determining the maximum-likelihood estimates (MLE's) or the parameter estimates of a particular distribution for a given dataset when the dataset is not complete or contains unobserved information. Bilmes (1998) also highlighted this description and further stated that there exists two ways of applying the EM algorithm. The first application is in a situation where the dataset actually contains missing values due to some difficulties or shortcomings with the observation procedures. The other application is the situation where maximizing the likelihood function is methodically difficult and when the likelihood can be broken down by taking into consideration the presence of and parameters for extra but unobserved (hidden) values. Wu (1983), Ghahramani and Hinton (1995) including several other authors stated that the parameter applications are most common in computing pattern recognition and it was used in their speech recognition experiment.

In applying the EM algorithm, the missing values are determined at the beginning by the parameter estimates of the distribution. Nelwamondo et al. (2007) performed a comparative research by applying artificial neural networks and the EM algorithm in completing missing values. Schneider (2001) applied the EM algorithm to complete and analyzed the unobserved figures in temperature change dataset. Firat (2011) in estimating missing data, applied the algorithm in the analysis of temperature changes. Kim and Ahn (2009) also used the algorithm to predict the unobserved values in daily rainfall observations. All of them gave the recommendation that the EM algorithm could be used effectively in analyzing data that contains missing values.

2.1.2 The Concept of Missing Values

Missing data occurs in different proportions and in different structures (Cohen and Cohen, 1983). The impact of missing data on the quality of research outcomes is dependent on the causes that led to the missing data, the mechanism of missing and the percentage of data that is missing (Tabachnick and Fidell, 2001).

It has been proven that the structure and pattern of missing data have serious effect on research outcomes than that of the quantity of unobserved values contain in a data (Tabachnick and Fidell, 2001). The two are important issues a researcher must solve before selecting a suitable method in dealing with missing data. As stated by Little and Rubin (1987), the mechanisms that cause the missing data can be categorized as missing completely at random, missing at random, and non-ignorable missing values.

Little and Rubin (1987), explained that, if the probability that a response is dependent on neither the observed nor the missing value that could have been recorded, the missing data are missing completely at random (MCAR). If the probability that a response is not dependent on the missing value itself but dependents on the original values or other completely observed variables, then the missing data is said to be missing at random (MAR).

From the perspective of Allison (2001), the missing data is taken to be ignorable if (a) the data are missing at random. (b) the parameters that determines the missing data process are not related to the parameters to be calculated. Ignoring simply means that there is no need taking into consideration the missing data as part of the estimation process.

However, if the missing data are non-ignorable, the proportion of missing data is dependent on the missing values themselves. In contrast to the ignorable situation, the missing data mechanism must be defined by the researcher and included in the data analysis in order to obtained unbiased parameter estimates.

Concerning the problem of how huge the percentage of missing data can be permitted by missing data procedures, there exist no firm procedure established by statisticians currently. In case of a few missing values in a random pattern from a larger dataset (the missing completely at random condition holds), the missing data problem is less dangerous and mostly every method for dealing with missing data produce almost the same results. Nevertheless, if a significant level of data is missing from a smaller to moderate size dataset, the problem can be very dangerous (Cohen and Cohen, 1983); (Cool, 2000) and (Tabachnick and Fidell, 2001).

2.1.3 Rainfall Pattern Estimation

Rainfall is considered among the most essential natural phenomena of climate change and its occurrence and distribution is unpredictable and temporal as well as spatially variant in nature. One of the serious problems of rain has to do with the interpretation of previous data on rainfall in relation to predicting rainfall patterns. Analysis of rainfall and estimation of annual maximum daily rainfall would improve the maintenance of water resources applications and also enable the proper usage of water resources (Subudhi, 2007). Probability and frequency analysis of rainfall data assist people to predict the expected rainfall at various intervals (Bhakar et al., 2006). Such information can be used to prevent floods and droughts and for planning and designing as well as construction of water resources such as reservoir design, flood control work and soil and water conservation planning (Agarwal et al., 1988) and (Debral et al., 2009).

Even though rainfall is unpredictable and changes with time and space, it is generally possible to forecast return periods using different types of probability distributions (Upadhaya and Singh, 1998). Therefore, probability analysis of rainfall is very relevant in solving various water management problems and to access the crop failure due to excess rainfall. Scientific forecast of rainfall and crops planning properly done could prove to be a vital instrument in the side of farmers for better economic gains (Bhakar et al., 2006). Frequency analysis of rainfall data has been attempted for various return periods by (Nemichandrappa et al., 2010); (Manikandan et al., 2011) and (Vivekanandan, 2012). Probability analysis of rainfall enables us to predict rainfall at various times. The probability distribution functions most commonly used to estimate the rainfall pattern include; normal, log-normal, log-Pearson, gamma and gumbel distributions. Kumar (2000)and Singh (2001) in their conclusions stated that the log-normal distribution is the best probability distribution model for forecasting annual maximum daily rainfall in a research conducted in Ranichavsi(Tehrin-Garhwal) and Tandong (Sikkim) respectively. Kumar et al. (2007), analyzed annual maximum rainfall and indicated that the log-Pearson type-III probability distribution model could be applied in designing hydraulic and soil and water conservation structures for Almora and similar places in Uttarakhand. Subudhi (2007) came out that the normal probability distribution model is the best fit for forecasting the average annual maximum daily rain of Chakapada block of Kandhamal district in Orissa.

In another hand, estimating meteorological time series with missing values is an important issue to many climatic analyses such as the study of rainfall patterns. Nearly all the influential precipitation data contain a fraction of incomplete entries. Deficient data may come as a result of misplaced record books due to conflicts and flames accidents and intermittent break down of permanent stations, faulty equipments and link failures. Solution to this problem is to eliminate periods with misplaced information from the analysis or neglect this difficulty when the quantity is not huge. Simolo et al. (2009) stated that some of these procedures, may neglect relevant information that can cause bias for several climatic studies.

In order to overcome these problems, several interpolation procedures are proposed several years ago which are intended to estimate incomplete data of rainfall data mostly on monthly or seasonal bases. Procedures for dealing with missing values in daily rainfall data, are not available and produce clear errors though those procedures execute better at minimum rainfall; for example (DeGaetano et al., 1995) and (Xia et al., 2001). The circumstance mostly turns to be complex when handling rainfall due to the huge length and occasional inconsistency. Furthermore, in this situation, the issue comes in two categories, once the location of time and the amount of rainfall for every day have to be re-estimated. That is determining precise parameters of incomplete data in daily rainfall data is still a problem particularly when long duration and foul-mouthed rain gauge networks are taking into consideration.

Between stations procedures for determining observations in climatic series are the easiest methods. These may be independent, because they use only the data obtain from the observation been estimated by imputing the unobserved data, e.g. with the expected value of earlier periods or by the data predicted outcome (Kemp et al., 1983). Despite being simple, these procedures are appropriate for data with autocorrelation to be high and for estimating future rainfall means and are therefore basically ineffective so far as average daily rainfall are considered.

Established procedures of completing incomplete rainfall data, on daily and monthly bases are often base on interpolations; these are replaced with values in an intended station determined by means of related values of closer stations. The reciprocal distance weighing procedure used by Cressman (1959) and Shepard (1968), is among the most widely applied distance-based schemes. Mostly, it is the same as calculating a weighted average by using inverse squared distance within intended and relatively near stations as weighting factors, with the aim that the present of positive association involving observations from close experiment stations. Also, since only the distance cant not be used to describe the connection condition of rainfall data and the choice of relatively close stations is very relevant to the adequacy of outcomes, several variants and adds-on to this technique are been developed. An example, inverse squared distance can be substituted by high power or by negative exponential expressions of distance or by pictorial associations that indicate correlative properties (Delay et al., 1994) and (Lloyd, 2005). Current concepts aim at improving the inverse distance weighing procedure and backward normal fractional procedure can be determine, in the case of (Teegavarapu and Chandramouli, 2005) and (Suhaila et al., 2008). Moreover, simplify methods together with the assume ”nearest station procedure” and ”single-best-estimator” example Wallis et al. (1991); Eischeid et al. (2000) and Xia et al. (2001) where frequently applied in calculating unobserved data of rainfall as well as other climatic data. Several complicated interpolation techniques to fill in gaps in climatic and rainfall data, obtain the functional correlation among intended and relatively near stations. These for instance, fitting spline-surface, e.g. Hutchinson and Gessler (1994), statistical techniques like simple interpolations and kriging (Creutin and Obled, 1982), regression-based techniques together with straight least squares or least absolute deviation condition, (Tabios and Sala, 1985); (Beauchamp, 1989) and

(Schneider, 2001). The efficiency of these procedures for prediction of rainfall data has been studied by so many comparatives procedures on daily base and on monthly base; example; Ashraf et al. (1997); Teegavarapu and Chandramouli (2005) and usually determine to be higher than that of weighting procedure. Particularly in Eischeid et al. (2000), determine that multiple-linear regression (MLR) performed better among several of the mostly used procedures involving missing data estimation in daily rainfall and other climatic data. Meanwhile, regression-based techniques, similar to the weighting procedures, experience the over-estimation of the total number of rainy days within a period. More so, the probability distribution of rainfall is not constant, since high levels of rainfall incidence are always analytically estimated bellow standards.

Several statistical characteristics of rainfall observations are applied in procedures developed by Karl and Knight (1995), Brunnetti et al. (2004) which is on the basis of the fit of two parameter Gamma distribution to individual stations daily observations. Consequently, to estimate the probability that rainfall occurred on any unobserved period, a random variable estimator is used together with probability distribution model assuming the experimental probability of rain for the day. Then, Gamma distribution is used alongside a random number generator to approximate the amount of precipitation of days classified to be wet-days. Despite the fact that the average statistical characteristics of rainfall (i.e. total number of days of rain and the quantity of rain on yearly and monthly bases) are not constant; this procedure is of less application on a daily time scale since it haphazardly identifies rainfall occurrences.

Lastly, together with the aforementioned methods described for estimating the amount of missing values in rainfall and other climatic observations, databased techniques based on neural network algorithms must be taking into consideration (Elshorbagy et al., 2000); (Khalil et al., 2001). That is, upon noticing that there is a correlation among data from intended and nearby stations. Current improvement within the section may be seen, for instance (Teegavarapu and

Chandramouli, 2005) and (Coulibaly and Evora, 2007). In Coulibaly and Evora (2007), particularly, various types of architectures on artificial neural networks have been determined and the efficiency of their estimation of unobserved daily rainfall observations appraised. The investigators concluded pretty precise outcomes for best performing models despite the fact that the experimental data information is not steadily constant.

Since the rainfall data recorded in Ghana also contain some amount of missing values which are not always considered in the analysis of the rainfall pattern but have a significant effect on the outcome, the aim is to apply the Expectation-Maximization (EM) algorithm to estimate the missing values and average annual one day precipitation pattern for sixteen selected stations using the best fit probability distribution model among the three considered for the study.(i.e. Gamma, Log-normal and Normal).

Chapter 3

Methodology

3.1 Introduction

In this chapter, we present the methodology of the study. It is categorized into three sections. The first Section looks at the basic principles behind fitting the theoretical distributions of the data. Section two focuses on the goodness-of-fit test for selecting the best theoretical probability distribution and the last section deals with the formulation and usage of the EM algorithm with regards to missing values analysis.

The sample data for this research is a secondary data obtained from the Ghana Meteorological Agency (GMET), Accra. It consists of daily rainfall figures recoded in millimeters (mm) for fifteen (15) selected rainfall stations all over Ghana. It covers a ten (10) year period starting from 1st January, 2000 to 31st December, 2015. The data analysis was performed using R x 64.3.1.0 and SPSS statistical software packages.

3.2 Fitting the Probability Distributions

Fitting distributions to data is a common task in statistics and it is based on choosing a probability distribution to model a random variable as well as determining parameter estimates for that distribution. There are so many theoretical probability distributions that are used to fit or estimate parameters of a given data. The distribution models commonly applied in estimating rainfall data are gamma, lognormal and the normal distribution among others. These are used to estimate the parameters of rainfall of various magnitudes within a given period.

The best probability distribution model is chosen by using the goodness-of-test such as the chi-square Kolmogorov-Smirnov, Cramer-von Mises etc.

3.2.1 The Gamma Distribution

The gamma probability distribution is one of the common distributions used statistically when unimodal and positive observations are accessible. In this case we outline shortly the sketch of fundamental characteristics of the gamma distribution and deliberate on the part relatively close in associated with length and rainfall pattern distributions.

The probability density function of a simple random variable that follows the gamma distribution is defined as:

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma\alpha} x^{\alpha-1} e^{-\frac{1}{\beta}x} \tag{3.1}$$

Where $\alpha,\beta > 0$, with α and β representing the shape and scale parameters respectively.

The function of the gamma distribution is defined such that the total area under the density function is a unit as:

$$\Gamma\alpha = \int_{0}^{\infty} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \tag{3.2}$$

The equations for the likelihood function for a random sample of size n are;

$$\sum_{i=1}^{n} log x_i - nlog(\hat{\beta}) - n\phi(\hat{\alpha}) = 0, \sum_{i=1}^{n} x_i - n(\hat{\alpha})(\hat{\beta}) = 0 \tag{3.3}$$

These are estimated step by step, and iterative techniques for estimating the parameters of the gamma distribution are currently inbuilt in several statistical software packages.

3.2.2 The Log-normal Distribution

If a random variable *x* follows the log-normal distribution, then the logarithm or natural logarithm of the variable is distributed normally. The probability distribution function of the variable is given by:

$$f(y) = \frac{1}{\sigma_y\sqrt{2\pi}} exp[\frac{-1}{2}(\frac{y-\mu_y}{\sigma_y})^2], for 0 \leq x \leq \infty \tag{3.4}$$

Where σ_y is the standard deviation and μ_y is the mean of $y = lnx$

The maximum likelihood estimates of m and s^2 for the parameter μ and parameter σ are

$$m = \frac{1}{n}\sum_{i=1}^{n} lny_i, s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(lny_i - m)^2 \tag{3.5}$$

These are estimated step by step, and certainly several techniques for estimating the log-normal parameters are currently available in several statistical software packages.

3.2.3 The Normal Distribution

The normal distribution which has two parameters is widely used as a relevant distribution for continuous variables which are evenly distributed. The probability density function is defined below as:

$$f(x,\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} exp[\frac{-1}{2}(\frac{x-\mu}{\sigma})^2], for -\infty \leq x \leq \infty \tag{3.6}$$

In which μ is the mean and σ is the standard deviation of random variable x. Where $N(\mu,\sigma) \sim (0,1)$.

The maximum-likelihood estimates of the parameters μ and σ^2 are;

$$\mu = \frac{1}{n}\sum_{i-1}^{n} x_i, \sigma^2 = \frac{1}{n-1}\sum_{i-1}^{n}(x_i - \mu)^2 \tag{3.7}$$

These can be solved by the maximum likelihood estimation, and several techniques of determining these parameters of the normal probability model are also accessible in several statistical software packages.

3.3 Testing the Goodness-of-Fit of the Probability Distributions

The goodness of fit test is the measure of how compatible a random sample is with regards to an existing probability distribution model. The goodness of fit test hypothesis is stated as follows.

H_0: the average annual rainfall data follows the selected distribution.

H_1: the average annual rainfall data does not follow the selected distribution

The following goodness of fit tests are commonly used in the case of continuous data with regards to the selection of the best probability distribution fit alongside a criteria called the goodness-of-fit criteria. The kolmogorovSmirnov test, Cramer-von Mises test and Anderson-Darling tests always go with the Akaike's information criterion and the Bayesian's information criterion.

3.3.1 Kolmogorov-Smirnov Test

The Kolmogorov-Smirnov test statics (D) is a statistical measure of how small or large the vertical variation exist between the theoretical and empirical cumulative distribution functions. This is designed as:

$$D = Max_{1 \leq i \leq n}[F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i)] \tag{3.8}$$

In which x_i = random sample, $i = 1,2,...,n$

The test statistic (D) is estimated for all the distributions in the group of the competing distributions and the distribution with the smallest statistic (D) value is selected as the best distribution.

3.3.2 Cramer-von Mises Test

The Cramer-von Mises test statistic is a method used to compare the fits of theoretical distributions (Cumulative distribution functions) with the empirical probability distribution models. It is also used for comparing the fit of two empirical distributions. The test statistic is defined as:

$$W^2 = \frac{1}{12n} + \sum_{i=1}[\frac{2i-1}{2n} - F(x_i)]^2 \tag{3.9}$$

Where x_i = random sample, i = 1,2,...,n

The test statistic W^2 is estimated for each and every distribution in the group of competing distribution and the best distribution fit is the one with the least W^2 value.

3.3.3 Anderson-Darling Test

The Anderson-Darling statistic A^2 is used for comparing the fit of the observed sample cumulative distribution function to the empirical distribution function. It is given by:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}(2i-1)[lnF(x_i) + ln(1 - F(x_{n-i-1}))] \tag{3.10}$$

Where x_i = the observed sample, i = 1,2,...,n

The test statistic A^2 is determined for all the competing distribution in the group of the fitting distribution and the one with the smallest test statistic A^2 is selected to be the best distribution fit.

3.3.4 Akaike Information Criterion

The Akaike Information Criterion is the number one distribution selection criterion to be globally accepted as a standard for statistical distribution fitting criteria. The Akaike Information Criterion is used to measure the comparative quality of a statistical distribution for an observed data. The AIC again, gives procedures for distribution selection because it is formulated on information theory. Furthermore, the AIC is an addition to the maximum idea and the maximum-likelihood idea applied to the estimation of model parameters once the category of the distribution has been classified. Mostly, the Akaike Information Criterion handles the trade-off between the goodness-of-fit of the distribution and the complications of the distribution. The test statistic of the AIC is given as:

$$AIC = 2k - 2ln(L) \tag{3.11}$$

In which k represents the number of parameters estimated in the distribution and L represents the optimal maximum value of the likelihood function for the distribution.

Assuming a group of competing distribution of different patterns, the maximum-likelihood estimation is applied in fitting the distribution and the AIC is determined in respect to all the distributions fit. The criterion for selecting the best distribution is done by taking the distribution which provides the smallest AIC test statistic. The first component of the AIC is to measure the goodness of fit of the distribution and the second component is termed the consequence function of the criterion because it penalizes the number of parameters used in competing model (Akaike, 1973).

The AIC is significance for; in sample and out of sample prediction efficiency of a particular distribution. In sample prediction indicates how the selected distribution fit the sample data for a particular observation whilst out of sample prediction is concentrated on showing how the fitted distribution best predict future values.

3.3.5 Bayesian Information Criterion

The Bayesian Information criterion (BIC) is another procedure very important for determining the more suitable distribution among the group of competing distributions. The BIC is determined by substituting the positive value $2k$ in the AIC expression by $-2L_m$. therefore, the BIC is expressed below as;

$$BIC = -2L_m + mln(n) \tag{3.12}$$

Where m is the number of parameters to be estimated in the model, n represents the size of the data and L_m represents the optimal maximum value of the model likelihood function.

The maximum-likelihood estimation is used in fitting the probability distribution (Schwarz., 1978). The BIC is determined for all of the distributions among the group of distributions competing and the one with the least BIC statistic is selected as the best distribution fit.

3.4 Maximum-Likelihood Estimation

This is a procedure for estimating the statistical model parameters. When applied to a data set and a given statistical model, maximum likelihood estimation obtains parameters estimates for the model.

Based on the hypothesis, the maximum-likelihood estimates of the model parameters are the estimate of μ and σ^2 which maximize the likelihood function of the distributions.

In using the method of maximum-likelihood, one have to first specify the joint density function for all observed data. With an identically independent distributed sample, the joint density function is provided as:

$$L = \Pi_{i=1} f(x_i|\theta) \tag{3.13}$$

The maximum-likelihood estimate (MLE) of θ is the parameter of θ that maximizes the equation (3.13) above. It is the parameter that makes the observed data "most likely". Instead of maximizing this product which can be quite difficult, one often have to use the fact that the algorithm is an increasing function so it will be equivalent to maximizing the log-likelihood. While the number of observations increased to infinity, sequences of maximum-likelihood estimators have the following characteristics.

XConsistency

XAsymptotic normality

XEfficiency

The lower bound of the Cramer-Rao is achieved when the observed sample size tends to infinity. With regards to this study, the maximum-likelihood estimation is used to determine the values of the parameters of the best fit probability distribution.

3.5 The Expectation-Maximization (EM) Algorithm

The EM algorithm is a statistical method devised by Dempster et al. (1977) to deal with difficulties encounter in the maximum-likelihood procedures with regards to the estimation of missing value problems. It is an iterative procedure use in estimating unknown parameters of incomplete data. Determining the parameter is simple if it is known that there are missing values. Similarly, with known given parameters, we are able to determine the estimates for the missing values. It has been formulated on the basis of the inverse dependence between the parameters of the model and the missing values. When the sample size is correctly selected, the EM algorithm can determine the missing values efficiently.

The EM algorithm develops an iterative principle for determining parameter estimates in which part of the data are missing. This iterative process can be outlined as follows;

XSubstitute missing values with estimated values

XEstimates parameters

XRepeat processes until they converge

These ideas have been consistently applied for a long period of time until Orchard and Woodbury (1972) in their information principle developed the hypothetical basis of the fundamental principle.

The EM algorithm, similar to other procedures for estimating incomplete data that neglects the process which causes the missing values in the data set, relies on the presumption that the unobserved data in the whole data set are missing at random, implying that the probability of the missing observation is independent of the missing value itself (Rubin, 1976).

The EM algorithm process have two part: the Expectation part (E-step) and the Maximization part (M-step). The E-step determines the predicted expectations of the missing data taking into consideration the observed data and estimates of the parameters of the distribution. The Maximization-step calculates the parameter estimates of the model to maximize the complete data loglikelihood function from the expectation-step. These two steps are performed step by step until the iteration converges (Schneider, 2001). The EM algorithm performs the expectation step and the maximization step alternatively to update the estimator θ_n, of the hidden parameter θ for n number of iterations.

The uncertain predictions of the unobserved values in the incomplete observed data and the parameter estimates of the model in the E-step are estimated using equation (3.14);

$$Q(\theta_0|\theta_n) = Ez|x,\theta_n[logL(\theta;x,z)] \quad (3.14)$$

In which, $L(\theta;x,z)$ represent the function of the likelihood, θ is the vector parameters, θ_n is the estimated parameter of the distribution, x represent the observed sample data and z represents the unobserved observations.

In the Maximizatin-step, the parameter estimates of the model can be estimated using equation (3.15) to maximize the complete-data log-likelihood function of the Expectation-step

$$\theta_{(n+1)} = arg_\theta maxQ(\theta|\theta_n) \quad (3.15)$$

For this research, we assume that there is a random vector y which joint density $f(y;\theta)$ is represented by a q-dimensional parameter. If the observed sample vector y are observed, it is of great concern to calculate the maximumlikelihood estimates of θ based on the probability function of y.

The function of log-likelihood of y

$$logL(\theta;y) = `(\theta;y) = logf(y;\theta) \quad (3.16)$$

Which is expected to be maximized.

For a data containing unobserved values, only an expression for the complete observation vector y, is observed. This is represented by denoting y as (y_{obs},y_{mis}), in which y_{obs} gives the original (partial observation) and y_{mis} gives the hidden or misplaced information. In trying to make the description simple, we presume that the unobserved data are lost at random Rubin (1976). Such that;

$$f(y;\theta) = f(y_{obs},y_{mis};\theta) = f_1(y_{obs},\theta).f_2(y_{mis}|y_{obs};\theta) \quad (3.17)$$

In which f_1 and f_2 gives the joint distribution functions of y_{obs} and y_{mis} of the observed data process. This implies,

$$\ell_{obs}(\theta,y_{obs}) = \ell(\theta;y) - logf_2(y_{mis}|y_{obs}) \quad (3.18)$$

In which, $\ell_{obs}(\theta,y_{obs})$ represents the log-likelihood of the observed data and $\ell(\theta,y)$ represents the complete data log-likelihood.

The EM algorithm becomes important in the maximization of ℓ_{obs} since it proves to be a bit complicated but maximizing the complete data log-likelihood ℓ is simple. On the other hand, because y is not observed, ℓ cannot be determined but maximized.

The EM algorithm attempts to maximize $\ell(\theta,y)$ step by step by substituting it with the conditional expectation given the observed data y_{obs}. This expectation is estimated with regards to the model of the complete data estimated at the currently estimated θ. In particular, if θ_0 is an initial value for θ, then on the first iteration it is supposed to calculate;

$$Q(\theta,\theta_0) = E\theta_0[\ell(\theta,y)|(y_{obs})] \quad (3.19)$$

$Q(\theta;\theta_0)$ is now maximized with respect to θ, that is, θ_1 is determine so that;

$$Q(\theta_1,\theta) \geq Q(\theta;\theta_0), forall \theta \in \Theta \quad (3.20)$$

Therefore, the two estimation procedures of the EM algorithm, thus the Expectation step (E-step) and Maximization step (M-step) can be expressed as below; The Expectation step (E-step) would estimate $Q(\theta;\theta_n)$, as in equation (3.14) where;

$$Q(\theta;\theta_n) = E\theta_n[`(\theta;y)|y_{obs}]$$

The Maximization step (M-step) then find θ_{n+1} as in equation (3.15) where;

$$\theta_{n+1} = arg_\theta maxQ(\theta|\theta_n)$$

such that

$$Q(\theta_{n+1},\theta_n) \geq Q(\theta,\theta_n)$$

The expectation and maximization steps are executed alternatively till $L(\theta) - L(\theta_n)$ is very negligible.

3.6 Th Normal EM algorithm

Let the complete-data $y = (y_i,y_n)$ be a random sample from $N(\mu,\sigma^2)$ as presented in equation (3.6)

Then;

$$f(y,\mu,\sigma_2) = (\frac{1}{2\pi\sigma^2})^{\frac{n}{2}} exp[\frac{-1}{2\sigma^2}(\sum y_i^2 - 2u\sum y_i + n\mu^2)] \quad (3.21)$$

Which implies that $(\sum y_i, \sum y_i^2)$ represent the sufficient statistics for $\theta = (\mu,\sigma^2)$. The complete data log-likelihood function is given as;

$$\ell(\mu,\sigma^2,y) = -\frac{n}{2}log(\sigma^2) - \frac{1}{2}\sum_{i=1}^{n}\frac{(y_i-\mu)^2}{\sigma^2} \quad (3.22)$$

which gives

$$-\frac{n}{2}log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} y_i^2 + \frac{\mu}{\sigma^2}\sum_{i=1}^{n} y_i - \frac{n\mu^2}{\sigma^2} \quad (3.23)$$

Suppose $y_i, i = 1,...,m$ are observed and $y_i, i = m+1,...,n$ are unobserved (at random) in which y_i are assumed to be *i.i.d* $N(\mu,\sigma^2)$.

Then the E-step computes,

$$E_\theta \sum_{i=1}^{n} (y_i | y_{obs}), E_\theta \sum_{i=1}^{n} (y_i^2 | y_{obs}) \tag{3.24}$$

From the complete data log-likelihood function above, the k^{th} iteration of the E-step

$$s_1^k = E\mu^{(k)}, \sigma^{2(k)} \sum_{i=1}^{n} (y_i | y_{obs}) = \sum_{i=1}^{n} y_i + (n-m)\mu^{(k)} \tag{3.25}$$

$$s_{2(k)} = E\mu^{(k)}, \sigma^{2(k)} \sum_{i=1}^{n} (y_i^2 | y_{obs}) = \sum_{i=1}^{n} y_i^2 + (n-m)[\sigma^{2(k)} + \mu^{(k)}] \tag{3.26}$$

The M-step can be determined by solving the expectations expressed in the E-step.

$$\mu^{(k)} = \frac{s_1^{(k)}}{n} \Longrightarrow \mu^{(k+1)} = \frac{s_1^{(k)}}{n} \tag{3.27}$$

$$\sigma^{2(k)} = \frac{s_{2(k)}}{n} - \mu^{2(k)} \Longrightarrow \sigma^{2(k+1)} = \frac{s_{2(k)}}{n} - \mu^{2(k+1)} \tag{3.28}$$

As the commonest algorithm available for complicated maximum likelihood estimations, the EM algorithm has so many interesting characteristics as compare to other iterative algorithms such as Newton-Raphson. First, it is usually simple to implement since it rest on complete dataset computations. The iteration of each E-step only requires taking expectations over complete-data conditional distributions. The iteration of each M-step of only involves complete dataset maximum likelihood estimation for which closed simple form formulas are already available. Secondly, it is numerically stable. Thus, each iteration required to increase the log-likelihood `$(\theta;y_{obs})$ in each iteration and if `$(\theta;y_{obs})$ is bounded, the sequence `$(\theta_n;y_{obs})$ converges to a fixed point. If the sequence θ_n

converges, it does so to a local maximum or saddle point of $\ell(\theta;y_{obs})$ and to a unique MLE if $\ell(\theta;y_{obs})$ is unimodal.

A major shortfall of EM is that its rate of convergence can be very time consuming if the data contains a lot of missing values. Dempster et al. (1977), proved that convergence is linear with the rate proportional to the percentage of information about θ in $\ell(\theta;y)$ that is observed.

Chapter 4

Data Analysis and Results

4.1 Introduction

This chapter displays, discusses and interprets the results from the study. It is generally organized into various sections such as preliminary analysis, the goodness-of-fit analysis and then the discussion of the accuracy of the EM estimates and results. The analysis of the data, plotting of graphs as well as the model fitting was carried out using the R x 64.3.1.1 and SPSS statistical software packages.

4.2 Data Description

The study has been considering rainfall figures for the period of sixteen years from January, 2000 to December, 2015 for fifteen (15) selected rainfall stations across the ten (10) regions of Ghana. Each station and its daily rainfall figures for the surrounding communities are analyzed separately in order to obtain accurate result.

Table 4.1 illustrates the descriptive statistics of the daily amount of rainfall of the selected stations for the period under discussion. These estimated statistics are the minimum amount of rain, the mean and the maximum amount of rain for each station. Also included in the table are the number of days of rain and the standard deviation

The results in Table 4.1 indicate that the daily minimum amount of rainfall recorded for the period is 0.1mm for all the stations except Akrokerri which recorded a daily minimum rainfall of 0.4mm. It is observed that, only three stations recorded maximum rainfall above 200mm with the highest rainfall of

219.9mm followed by 214.7mm and 205.7mm recorded for Koforidua, Saltpond and Ada respectively.

From the table, it is also observed that, among all the stations, the highest daily average rainfall of 16.09466 mm was recorded in Akrokerri in the Ashanti region of Ghana with standard deviation 15.89553 while the lowest average rainfall of 9.812865 mm was recorded in Koforidua in the Eastern region of Ghana with a standard deviation 15.06017.

Table 4.1: Descriptive statistics of daily rainfall data for the selected rainfall stations in Ghana (2000-2015)

Station	N	MINIMUM	MEAN	MAXIMUM	STD. DEV
Abetifi	1893	0.1	10.80481	145.6	14.68451
Ada	1157	0.1	11.00752	205.7	16.72671
Akrokerri	1310	0.4	16.09466	166.4	15.89553
Axim	2365	0.1	12.20351	173.5	20.40719
Bolgatanga	1098	0.1	13.81202	97.5	15.05656
Cape Coast	1356	0.1	10.93481	224	17.96238
Ejura	1448	0.1	14.3232	133.3	16.8232
Ho	1736	0.1	11.70132	154.2	15.73613
Koforidua	2052	0.1	9.812865	219.9	15.06017
Krachi	1564	0.1	13.43939	152.2	18.93294
Salpond	1458	0.1	10.72737	214.7	17.68883
Sefwi	2196	0.1	10.29877	118.4	14.1098
Tamale	1303	0.1	12.62832	120.3	15.48566
Wa	1393	0.1	11.81443	142.2	14.84487
Wenchi	1696	0.1	11.49322	118.4	14.65478

Source: Researcher's Computation Based on the Observed Data

It is clear in Table 4.1 that the observed daily rainfall figures of Axim produced the highest standard deviation of 20.40719 while the rainfall figures of Sefwi produced the least standard deviation of 14.1098.

We also observed that, only three stations recorded more than 2000 days of rain within the period under study with the highest number of rainy days of 2365 followed by 2196 and 2052 recorded for Axim, Sefwi and Koforidua respectively. Meanwhile, the least number of days of rain is 1098 recorded in Bolgatanga in the Upper East region followed by 1157 days recorded in Ada in the Greater Accra region.

The results in the Table 4.1 above are compared with the estimated parameters using the EM algorithm on the fitted model using the goodness of fit statistics.

4.3 Results and Discussions

The parameters of average annual rainfall estimates were computed. The maximum average values of the observation data (annual average rainfall) were fitted with three main probability distributions. The goodness of fit test was computed using Anderson-Darling, Cramer-von Mises and Kolmogorov-Smirnov tests. These were selected due to the nature of the observed data as well as the kind of probability distribution that is used for the analysis (continuous). The criteria used were the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC).

The data presented in table 4.2 found that the Cramer-von Mises gave the smallest test values followed by the Kolmogorov-Smirnov and the AndersonDarling test statistics.

The BIC statistics gave the highest test values followed by the AIC test statistics vales. It is also observed that the computed test of goodness-of-fit values for the three probability distributions; that is gamma, log-normal and normal, shows that the normal distribution provided the best fitted values followed by the log-normal distribution while the gamma distribution provided the worst fitted values among the three distributions.

This is because the goodness-of-fit statistics provided by the normal distribution are smaller than the others while the statistics provided by the gamma distribution are also the highest among the three probability distributions.

Table 4.2: Goodness of fit statistics.

Distribution	Kolmogorov	Cramer-von	Anderson Darling	AIC	BIC
Gama	2.14206	0.90270	5.80231	926.81055	949.98822
Lognormal	2.16107	0.87967	5.62686	924.52743	947.60508
Normal	2.13382	0.87580	5.61598	924.1781	947.36344

Source: Researcher's Computation Based on the Observed Data

These values were actually obtained from the sum of the goodness-of-fit statistics for the individual stations for the probability distributions. This indicated that the normal probability model is the most efficient distribution for modeling the data.

Therefore, the EM algorithm was applied on the normal probability distribution to estimate the missing values of the rainfall data as well as the parameters of the rainfall pattern for the selected rainfall stations. The results of the missing values of the stations estimated using the EM algorithm are presented in the appendix (B-P).

The NA's in the observed data presented in the appendix represent no rain or unrecorded rainfall (missing value) for that particular. These were successfully estimated and replaced by the normal EM algorithm. The results of the estimated missing values are closer to the true values of the observed data.

Table 4.3 below illustrates the estimated parameters of the daily amount of rainfall of the selected stations for the period under discussion. These estimated parameters include the mean, the standard deviation and the log-likelihood. These were estimated using the regular normal distribution (i.e. without EM
algoritm).

The results in Table 4.3 indicate that among all the stations, the highest average rainfall amount predicted is 16.09466 mm for Akrokerri in the Ashanti region of Ghana with standard 15.88946 and log-likelihood of -655.0 while the lowest average rainfall amount of 9.812865 mm was predicted in Koforidua in the Eastern region of Ghana with a standard deviation 15.0565 and log-likelihood of -1026.0.

Table 4.3: Average Rainfall for the Stations Estimated using the regular Normal Distribution

Station	Mean (Average Rainfall)	Standard Deviations	Log likelihood
Abetifi	10.80481	14.68063	-946.5
Ada	11.00752	16.71948	-578.5
Akrokerri	16.09466	15.88946	-655.0
Axim	12.20351	20.40287	-1182.5
Bolgatanga	13.81202	15.04971	-549.0
Cape Coast	10.93481	17.95576	-678.0
Ejura	14.3232	16.81739	-724.0
Ho	11.70132	15.73159	-868.0
Koforidua	9.812865	15.0565	-1026.0
Krachi	13.43939	18.92688	-782.0
Salpond	10.72737	17.68276	-729.0
Sefwi	10.29877	14.10659	-1098.0
Tamale	12.62832	15.47972	-651.5
Wa	11.81443	14.83954	-696.5
Wenchi	11.49322	14.65046	-848.0

Source: Researcher's Computation Based on the Observed Data

Meanwhile, we observed that the predicted daily rainfall figures of Axim still produced the highest standard deviation of 20.40287 while the predicted rainfall figures of Sefwi also produced the least standard deviation of 14.10659.

Table 4.4: Average Rainfall for the Stations Estimated using the Normal EM Algorithm

Station	Mean	Standard Deviations	No. of Iterations	Log likelihood
Abetifi	10.80266	14.68016	22	-946.5001
Ada	11.00089	16.71725	34	-578.5001
Akrokerri	16.08958	15.88983	32	-655.0001

Axim	12.20155	20.402	17	-1182.5001
Bolgatanga	13.80658	15.04962	38	-549.0001
Cape Coast	10.92916	17.95339	29	-678.0001
Ejura	14.31923	16.81701	29	-724.0001
Ho	11.69863	15.73101	24	-868.0001
Koforidua	9.810973	15.05586	20	-1026.0001
Krachi	13.43499	18.92568	26	-782.0001
Salpond	10.72231	17.68063	27	-729.0001
Sefwi	10.29643	14.10606	18	-1098.0001
Tamale	12.62407	15.47917	32	-651.5001
Wa	11.8108	14.839	30	-696.5001
Wenchi	11.48988	14.64992	24	-848.0001

Source: Researcher's Computation Based on the Observed Data

Table 4.4 above illustrates the estimated parameters of the daily amount of rainfall of the selected stations for the period under discussion. These estimated parameters include the mean, the standard deviation and the log-likelihood together with the number of iterations at which these estimates are obtained.

The results in Table 4.4 indicate that among all the stations, the highest average rainfall amount predicted is 16.08958 mm for Akrokerri in the Ashanti region of Ghana with standard 15.88983 and log-likelihood of -655.0001 while the lowest average rainfall amount of 9.810973 mm was predicted in Koforidua in the Eastern region of Ghana with a standard deviation 15.05586 and log-likelihood of -1026.0001. These estimates were obtained with 32 and 20 iterations for the two stations respectively.

Meanwhile, we observed that the predicted daily rainfall figures of Axim still produced the highest standard deviation of 20.402 while the predicted rainfall figures of Sefwi also produced the least standard deviation of 14.10606.

These results in Table 4.3 and Table 4.4 are compared using bar charts which give a clear difference between the estimates.

The Figure 4.1 above is a graphical comparison of the means of the observed data and means of the EM estimate for all the stations. From Figure 4.1, the blue bars indicating the graph of the means of the observed data are a bit high as compared to the green bars indicating the graph of the EM means. This plot therefore indicates a clear difference between the means of the observed data and the means of the EM estimates.

Figure 4.2 above also presents a graphical comparison of the standard deviations of the observed data and standard deviations of the EM estimates.

It is observed that the bar chart of the standard deviations of the observed data (blue bars) are slightly high as compared to the bars of the EM standard deviations (green bars). This also provided a different plot for the standard deviations of the observed data and the EM standard deviations.

The results in Table 4.4 shows that the estimated parameters using the normal EM algorithm are almost similar to the parameters estimated from the

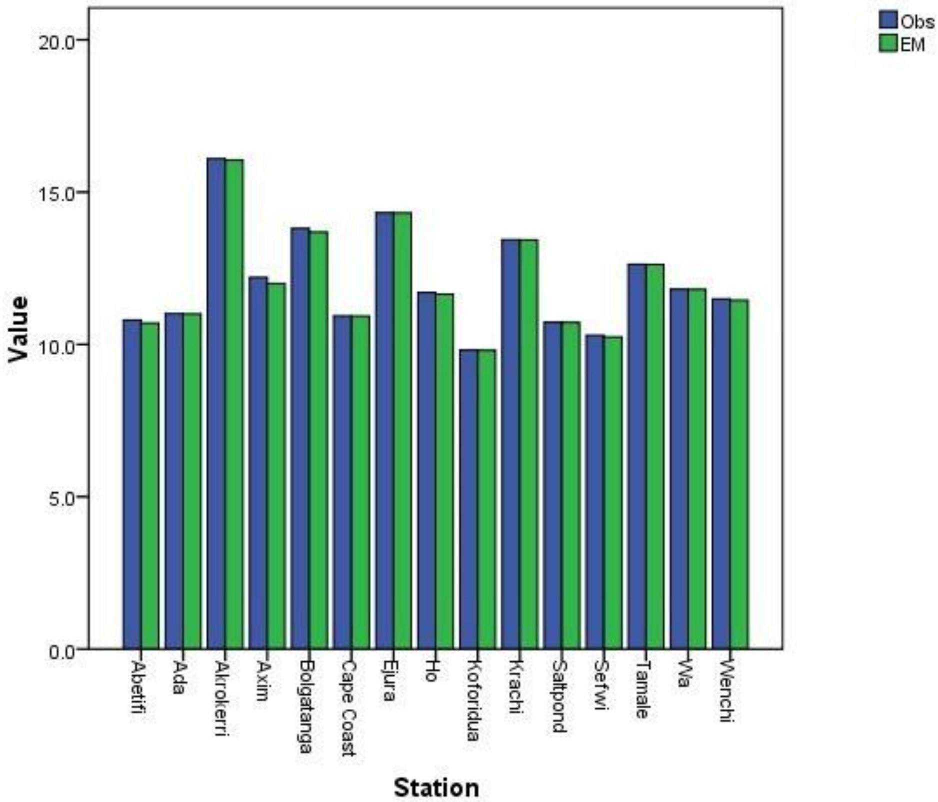


Figure 4.1: Bar Chart comparing the annual average rainfall and the EM estimates of all stations (2000-2015)

observed data presented in Table 4.3. The discrepancies are found to vary from 0.02 to 0.07 percent for the means and 0.01 to 0.08 percent for the standard deviations. The estimates are the same when corrected to two (2) decimal places but becomes different beyond two (2) decimal places. It is also observed that the log-likelihood estimates of the EM algorithm parameters gave the least values as compared to that of the observed data. This clearly indicated that the EM estimates are better for modeling the rainfall patterns of the selected stations.

The figures below show the density plot, cumulative distribution function plot (CDF), quantiles plot (Q-Q) and probability plots (P-P) of the annual average rainfall for the selected stations for the study. The density plot represents the probability function of the distribution fit along with the histogram of the empirical distribution. The cumulative distribution function plot represents the plot of both the empirical and fitted distributions.

The quantiles plots indicate the plot of the empirical quantiles (y-axis)

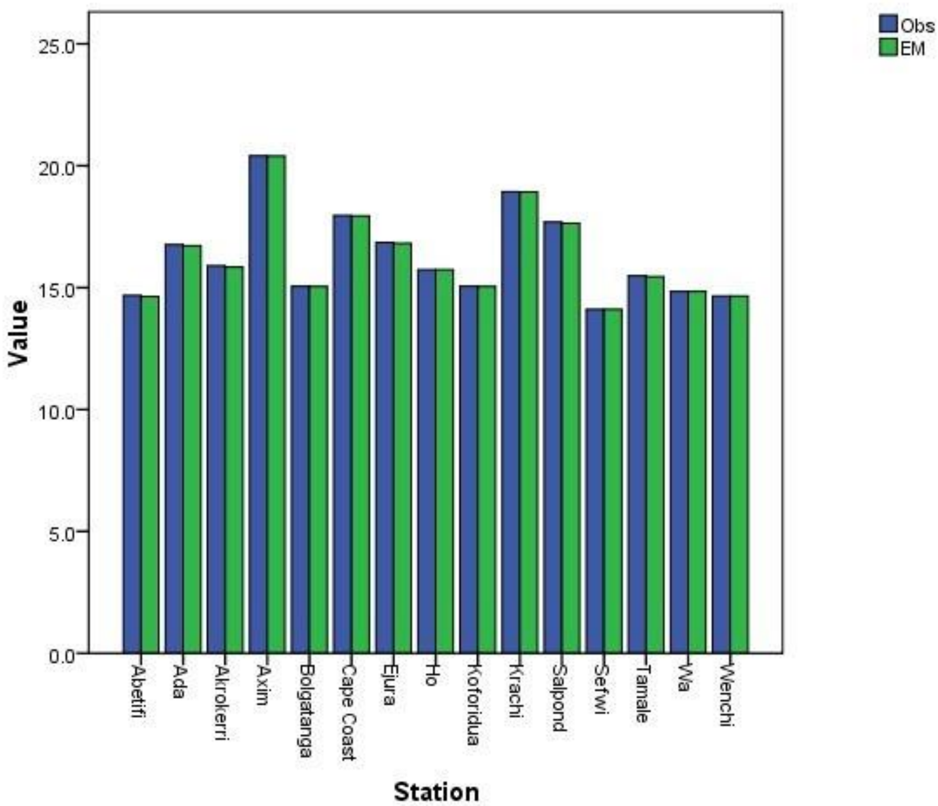


Figure 4.2: Bar Chart comparing the standard deviations of the observed values and the EM estimates of all stations (2000-2015)

and the theoretical quantiles (x-axis). It emphasizes the lack of fit at the distribution tails. The probability (P-P) plots also indicate the plot empirical distribution functions evaluated at each data point (y-axis) against the fitted distribution function (x-axis). It emphasizes the lack of fit at the center of the distribution.

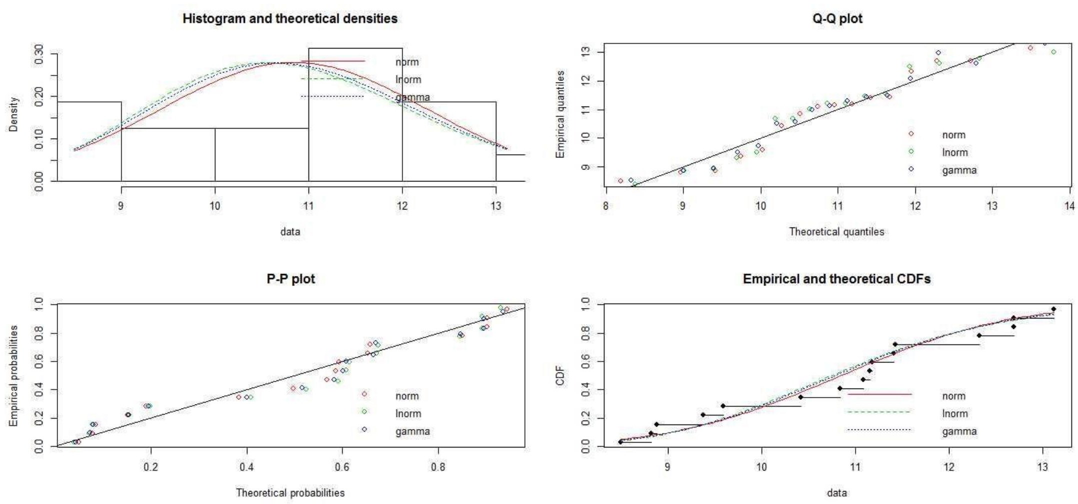


Figure 4.3: Distribution plots of the annual average rainfall of Abetifi (2000-2015)

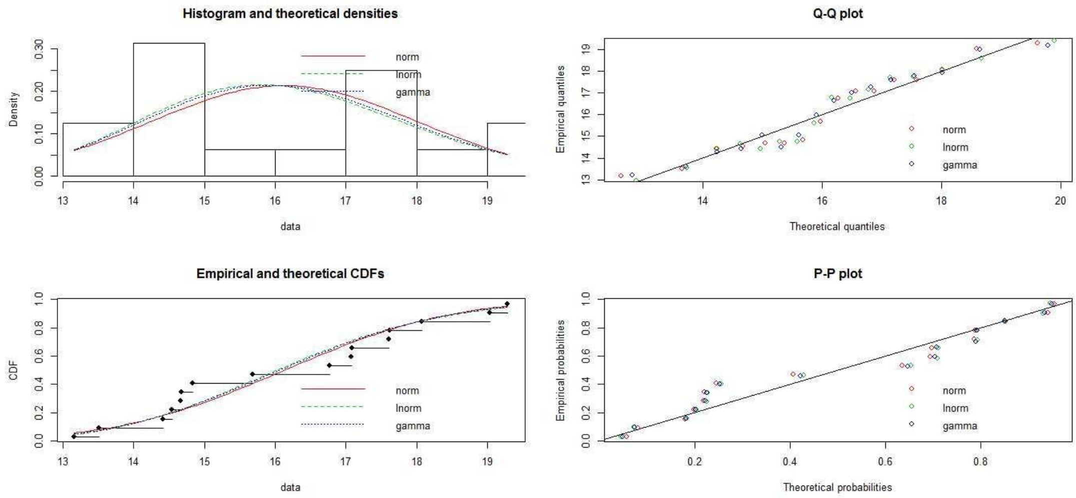


Figure 4.4: Distribution plots of the annual average rainfall of Ada (2000-2015)

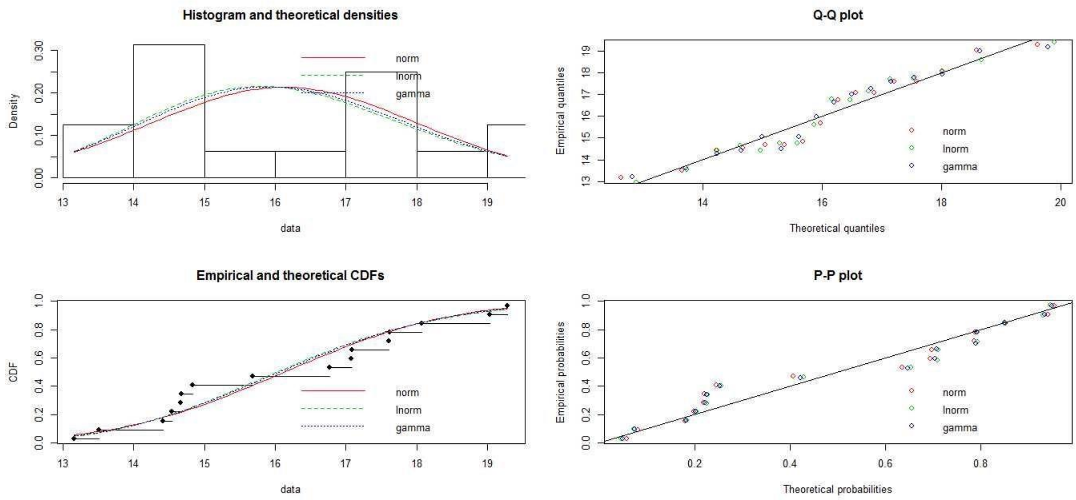


Figure 4.5: Distribution plots of the annual average rainfall of Akrokerri (2000-2015)

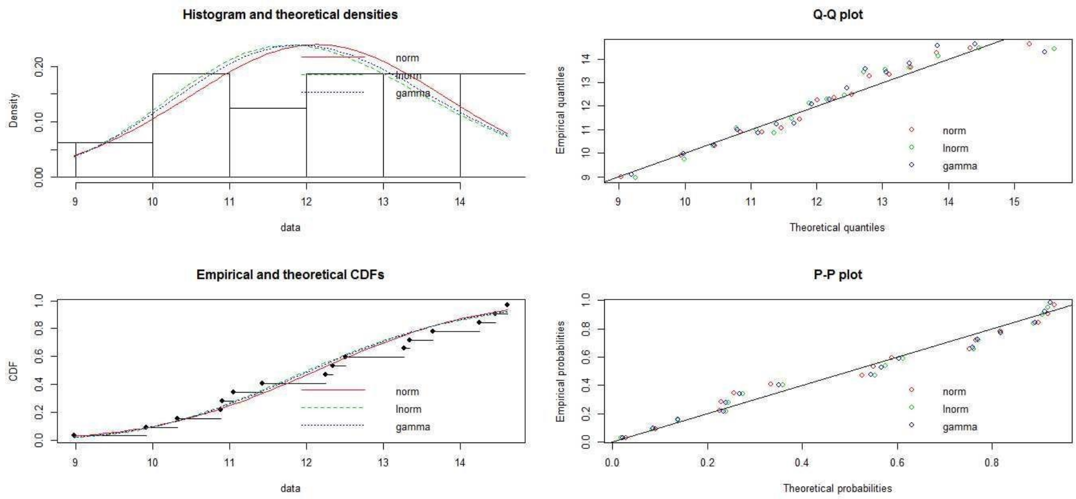
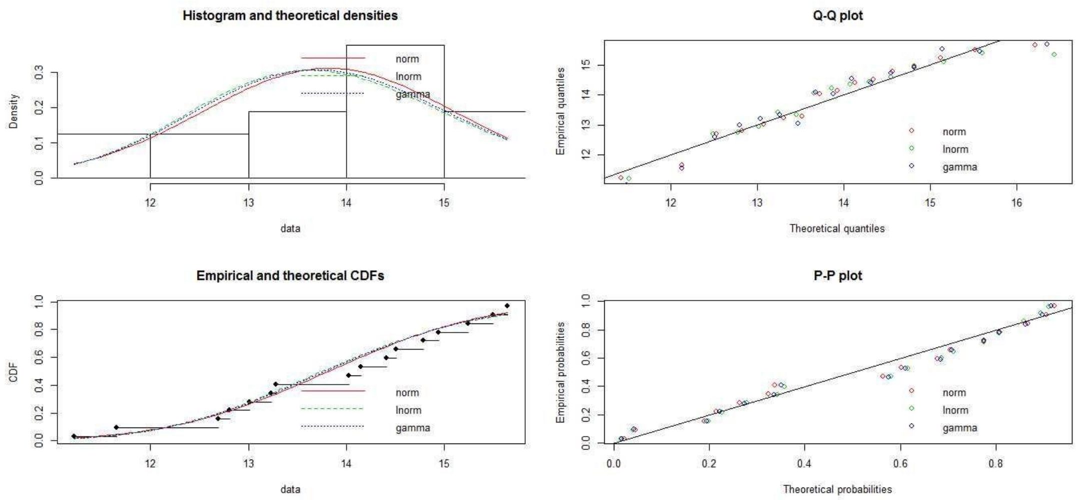


Figure 4.6: Distribution plots of the annual average rainfall of Axim (2000-2015)



2015)

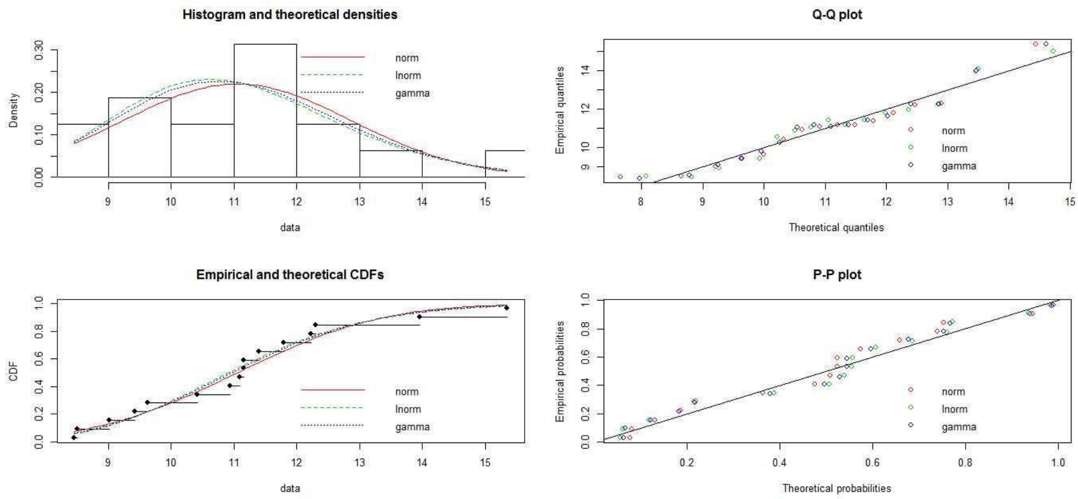
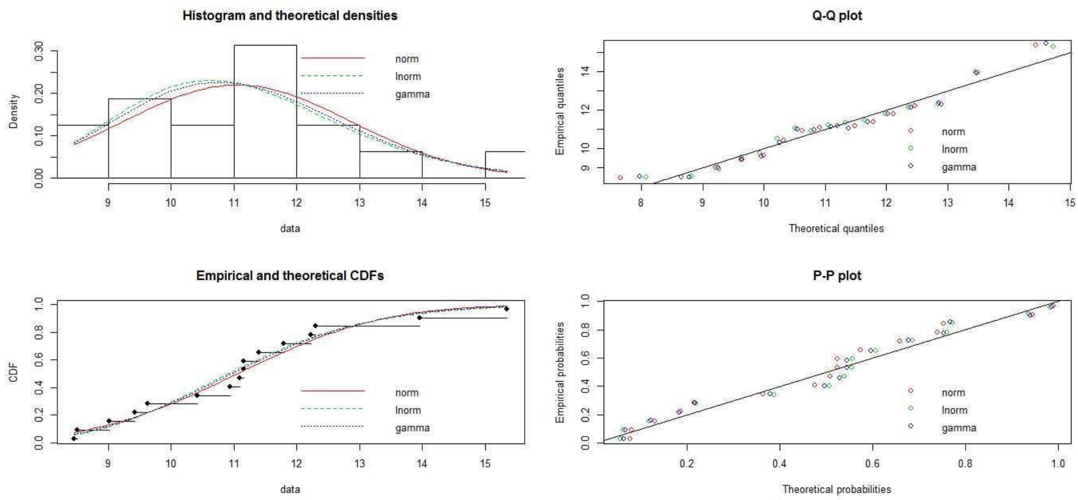


Figure 4.7: Distribution plots of the annual average rainfall of Bolgatanga (2000Figure 4.8: Distribution plots of the annual average rainfall of Cape Coast (20002015)



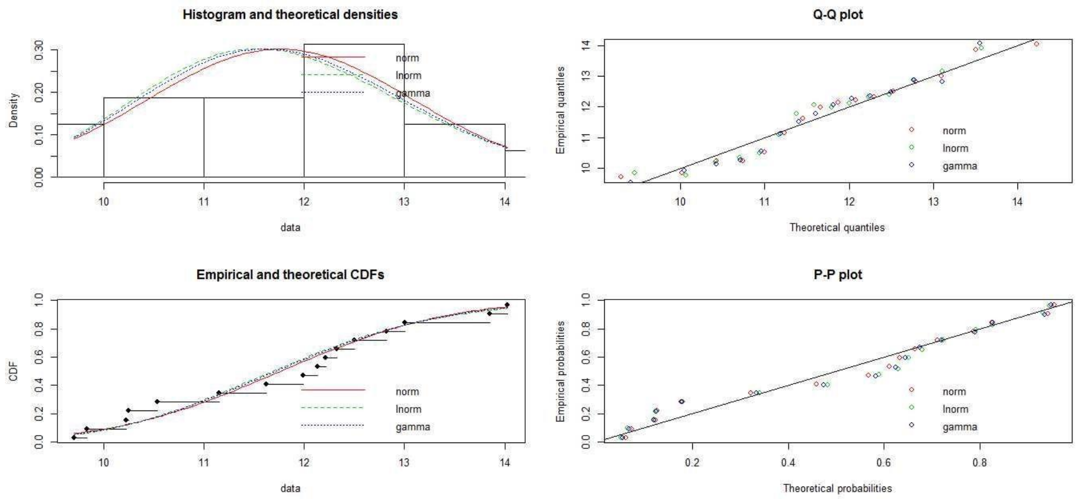
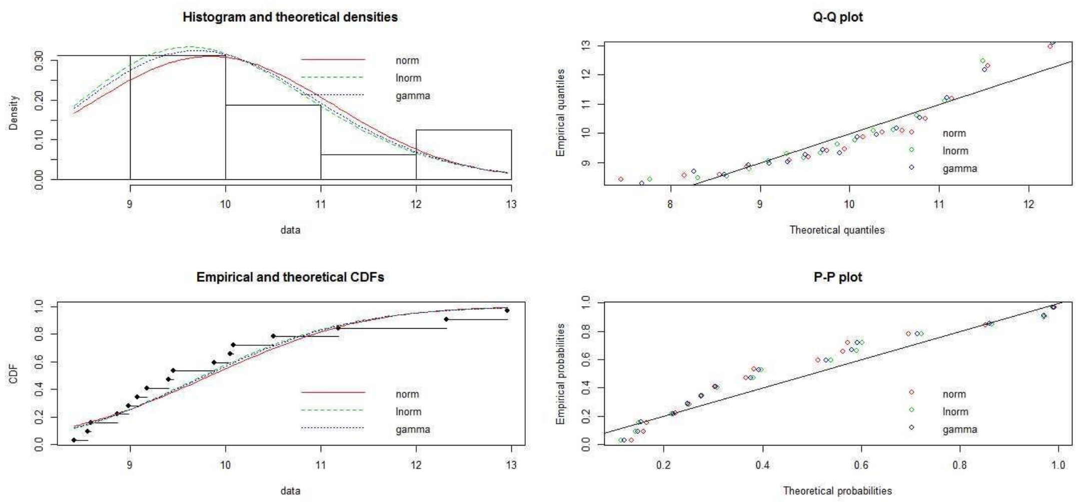


Figure 4.9: Distribution plots of the annual average rainfall of Ejura (2000-2015)

Figure 4.10: Distribution plots of the annual average rainfall of Ho (2000-2015)



2015)

Figure 4.11: Distribution plots of the annual average rainfall of Koforidua (2000-

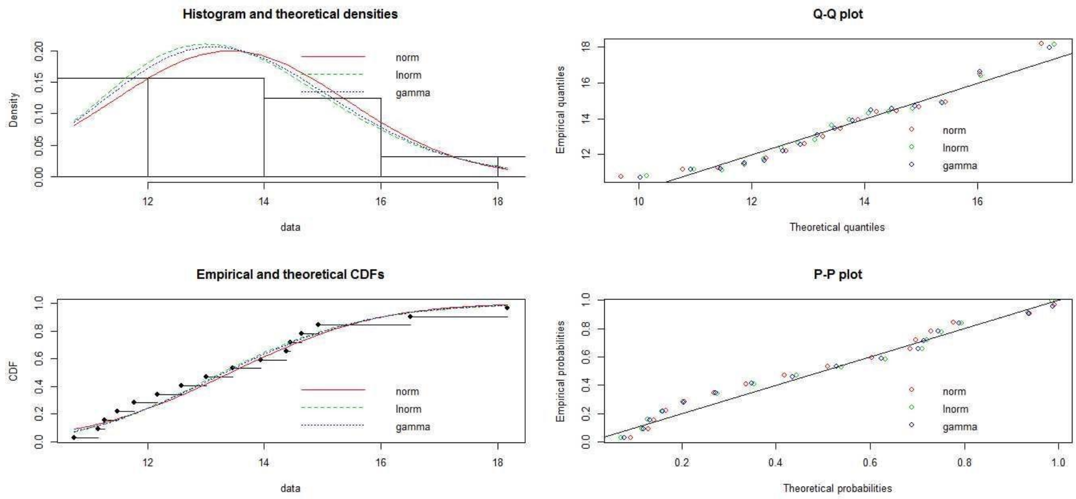
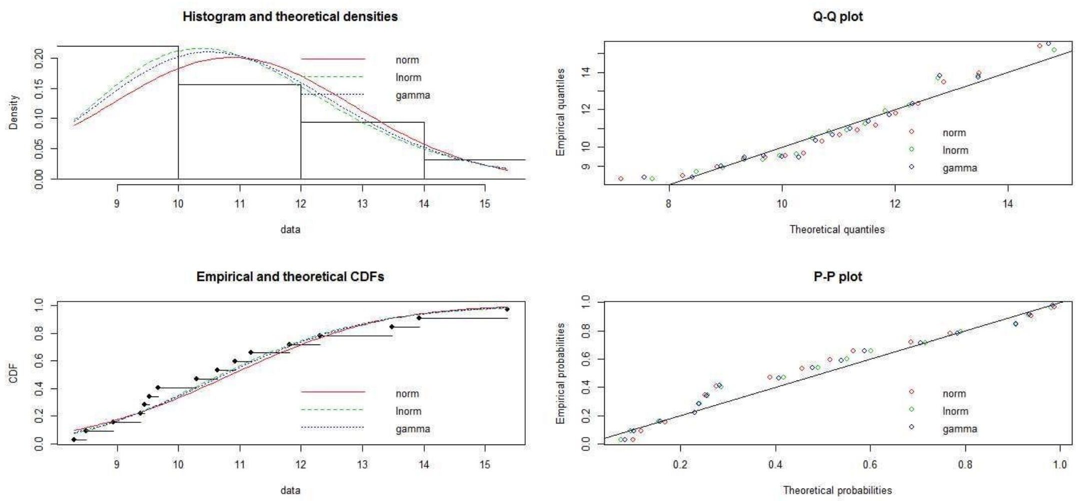


Figure 4.12: Distribution plots of the annual average rainfall of Krachi (20002015)



2015)

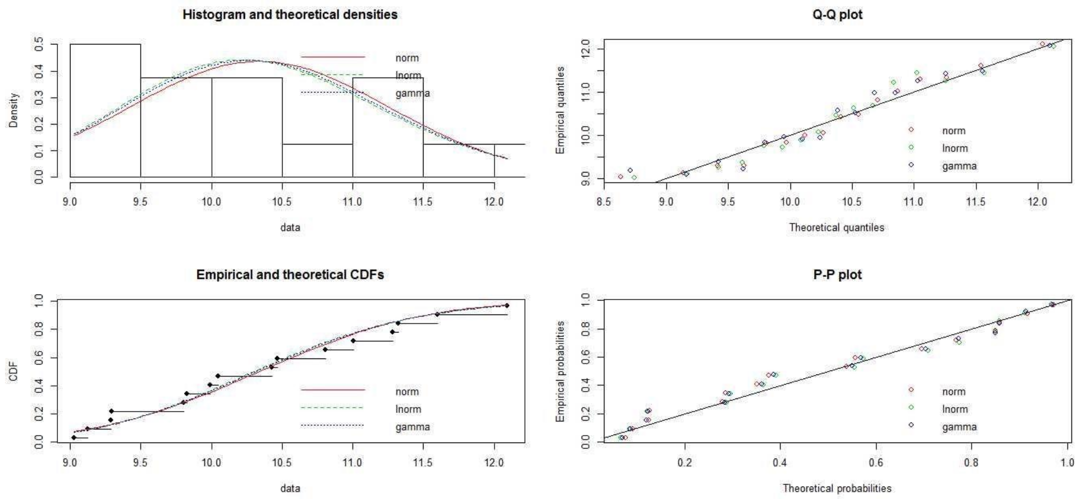
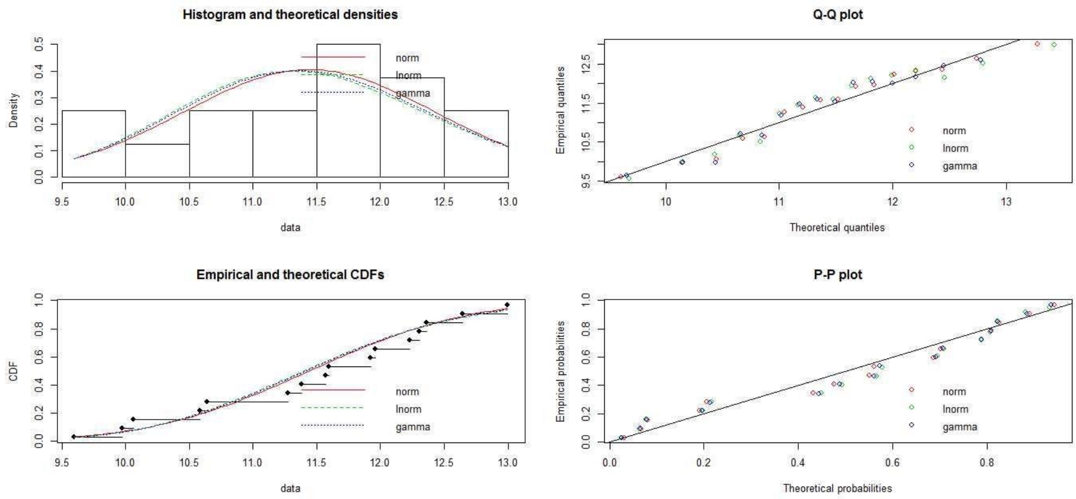


Figure 4.13: Distribution plots of the annual average rainfall of Salpon (2000Figure 4.14: Distribution plots of the annual average rainfall of Sefwi (2000-2015)

2015)



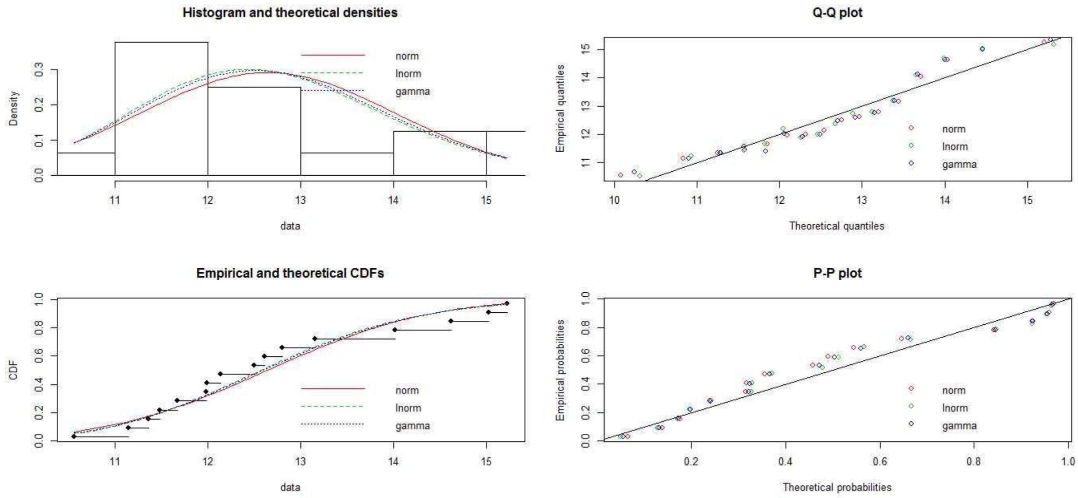


Figure 4.15: Distribution plots of the annual average rainfall of Temale (2000Figure 4.16: Distribution plots of the annual average rainfall of Wa (2000-2015)

2015)

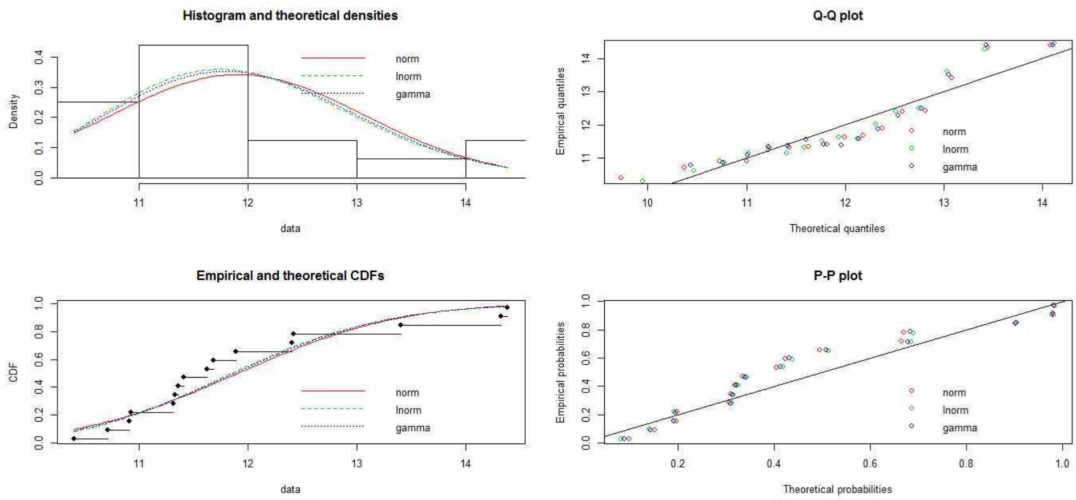


Figure 4.17: Distribution plots of the annual average rainfall of Wenchi (2000-

2015)

Chapter 5

Summary, Conclusion and Recommendations

5.1 Introduction

This chapter presents the summary of the outcome from the study together with the conclusions and the recommendations of some aspect of this work for further research. That is, the key results and findings of the efficiency of the normal EM algorithm in estimating the missing data, the goodness-of-fit statistics in selecting the appropriate distribution model and the accuracy of the estimated parameters.

5.2 Summary

Most empirical researches conducted to estimates the parameters of rainfall data have to assume certain probability distributions and patterns for estimating the parameters of the data. Some of which assume the absence of missing values in the data which goes a long way to affect the results. Others also used traditional methods of dealing with missing values, such as Listwise Deletion (LD), Pairwise Deletion (PD) and Mean Substitution (MS) which have proven not to provide good estimates for large amount of missing values in a data.

This research was carried out to determine the best estimates of rainfall data with special attention on the missing values in the data. It also sought to determine a best probability distribution model for modeling the rainfall pattern. The Expectation-Maximization (EM) algorithm was applied on the normal distribution to estimate the parameters of the data. In this research, we used a data for the period of sixteen years for fifteen selected rainfall stations across the country. The data for each station was analyzed separately in order to come out with accurate results.

From the results, it was observed that the highest number of rainy days experienced within the period was 40.47 percent recorded in Axim which predicted a daily average rainfall of 12.20155mm which was followed by 37.58 percent recorded in Sefwi with predicted daily average rainfall of 10.29643mm and 35.11 percent recorded in Koforidua with predicted daily average rainfall of 9.810973mm. This means that none of the stations experienced rainfall for up to 50 percent or more of the 5844 number of days under consideration for the study. The least number of rainy days experienced over the period was 18.79 percent recorded in Bolgatanga which predicted an average daily rainfall of 13.81202mm followed by 19.80 percent recorded in Ada with predicted average rainfall of 11. 00752mm.

The highest average daily rainfall predicted was 16.08958mm for Akrokerri with a standard deviation of 15.88983 and a log-likelihood of -655.0001 which are less than the observed average daily rainfall of 16.09466 with a standard deviation of 15.89553 and log-likelihood of -655.0000. The lowest average daily rainfall predicted was 9.810973mm recorded in Koforidua with a standard deviation of 15.05586 and a log-likelihood of -1026.0001 which are also smaller than the observed daily average rainfall of 9.812865mm with a standard deviation of 15.06017 and a log-likelihood of -1026.0000.

Also, the highest standard deviation observed in the study was 20.402 for the EM estimates and 20.40719 for the observed data both produced by the estimates of Axim. The lowest standard deviation observed was 14.10606 for the EM estimates and 14.1098 for the observed data produced by the estimates of Sefwi.

As observed from the results, the estimates exhibited a discrepancy of 0.02 percent and 0.06 percent for the lowest and highest values of the predicted average rainfall and the observed mean of the data. The standard deviations also show a discrepancy of 0.02 percent and 0.06 percent between the observed and predicted. This indicated that there exist a small level of difference between the normal EM estimates and the estimates of the observed which indicate the EM algorithm provided the best estimates of the data.

The values of the standard deviations indicated that there were some among of variability between the daily rainfalls figures recorded for the selected rainfall stations in Ghana over the period under study.

Therefore, the difference between the predicted estimates of the EM algorithm and the estimates of the observed data is very minimal when taken to a decimal place but the difference becomes significant when corrected to two (2) decimal places and beyond. It was also observed that the log-likelihood estimates of the EM algorithm parameters provided are smaller as compared to that of the observed estimates.

The compoleted data using the EM algorithm for all the stations for the year 2000 are presented in the Appendix

5.3 Conclusion

In order to achieve the set objectives, a theoretical basis was presented in chapter three (3) which outlined three probability distribution models which are gamma, log-normal and normal distribution. In arriving at the best model fit for the data, three goodness of fit tests; Anderson-Darling Cramer-von Mises, and KolmogorovSmirnov test were conducted together with two goodness-of-fit criteria; Akaike information criterion (AIC) and Bayesian information criterion (BIC) on the three distributions. The normal probability distribution model provided the lowest statistics among the three distributions and was therefore selected as the best fit model for the study.

The Expectation-Maximization (EM) algorithm was then applied on the normal probability model to estimate the average rainfall of the data. The results reveal that the EM algorithm successfully estimated the data since the predicted annual average rainfall values are closer to the true values of the means and standard deviations. As observed from the results, the EM estimates are relatively small which suggest that the EM algorithm estimated the missing values of the data correctly and provided better estimates.

Finally, considering the values of the log-likelihood, the normal EM estimates provided the lowest as compared to the log-likelihood values provided by the observed estimates. This clearly tells us that the normal EM estimates are the best for the data and we therefore conclude that the normal EM algorithm is better for estimating the rainfall patterns for the selected rainfall stations in Ghana. The EM complete data for all the stations for the year 2000 are presented in the Appendix (B-P).

5.4 Recommendations

From the basis of the summary and the conclusions obtained from the research, the following are recommended.

That the expectation-maximization (EM) algorithm together with the normal probability distribution should be used to estimate the missing values as well as the annual average rainfall of daily rainfall data recorded in Ghana.

That any other person including researchers interested in rainfall pattern modeling in general should consider the system of missing values in order to achieve accurate results since they have a significant effect on the outcome or the results. It is better to know the pattern of the data before applying a any statistical model to estimate the parameters.

REFERENCES

Agarwal, M. C., Katiyar, V. J., and Babu, R. (1988). Probability analysis of annual maximum daily rainfall of v .p aimalaya. *Indian Journal of soil conservation*.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *International Symposium on Information Theory*.

Allison, P. D. (2001). Missing data: Sage university paper series on quantitative application in the social sciences. *Thousand Oaks, CA: Sage*.

Ashraf, M., Loftis, J. C., and Hubbard, K. G. (1997). Application of geostatistics to evaluate partial weather station network. *Agricultural Forest Meteorology*.

Beauchamp, J. J. (1989). Comparison of regression and time series methods for synthesizing missing streamflow records. *Water Resource Bulletin*.

Bhakar, S. R., Bansal, A. N., Chajed, N., and Purohit, R. C. (2006). Frequency analysis of consecutive days maximum rainfall at banswara, rajasthan. *Indian ARPM Journal of Engineering and Applied Sciences*.

Bilmes, J. A. (1998). A gentile tutorial of the em algorithm and its application to parameter estimation for gaussian mixtures and hidden markov mdels. Technical report, University of Berkeley.

Brunnetti, M., Muageri, M., Monti, F., and Nanni, T. (2004). Charged in daily precipitation frequency and distribution in italy over the last 120 years. *Journal of Geophysical Research-Amosphere*.

Cohen, J. and Cohen, P. (1983). *Applied multiple regression /correlation analysis for bahavioural sciences, 2nd ed.* Hillsotale N J Lawrence Erlbaum Associates.

Cool, A. L. (2000). A review of methods for dealing with missing data: Paper presented at the annual meeting of the southwest educational research association. *Dallas*.

Coulibaly, P. and Evora, N. D. (2007). Comparison of neural network methods for infilling missing daily weather records. *Journal of Hydrology*.

Cressman, G. P. (1959). An operational objective analysis system. *Monthly Weather Review*.

Creutin, J. D. and Obled, C. (1982). Objective analysis and mapping techniques for rainfall fields; an objective comparison. *Water Resource Research*.

Debral, P. P., Mautushi, and Singh, R. P. (2009). Probability analysis for one day to seven consecutive days annual maximum rainfall for doimukh (itenagar), arunachal pradesh. *Journal of Indian Water Resources*.

DeGaetano, A. T., Eggleston, K. L., and Knapp, W. W. (1995). A method to estimate missing maximum and minimum temperature observations. *Journal of Aplied Meteorology*.

Delay, C., Neilson, R., and Philip, D. (1994). A statistical topograhic model for mapping climatological precipitation over mountainous terrain. *Journal of Aplied Meteorology*.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*.

Eischeid, J. K., Pasteris, P. A., Diaz, H. F., Plantico, M. S., and Lott, N. J. (2000). Creating a serially complete, national daily time series of temperature and precipitation for western united states. *Journal of Applied Meteorology*.

Elshorbagy, A. A., Panu, N. S., and Simonovic, S. P. (2000). Group-based estimation of missing hydrological data: Approach and general methodology i. *Journal of Hydrological Sciences*.

Firat, M. (2011). Analysis of temperature series: estimation of missing data and homogeneity test. *Meteorological Applicatios*.

Ghahramani, Z. and Hinton, G. E. (1995). Parameter estimation for linear dynamical systems. Technical report, Department of Computer Science University of Toronto Canada.

Hutchinson, M. F. and Gessler, P. E. (1994). Splines more than just a smooth interpolator. *Geoderma*.

Jordan and Jacob, R. A. (1994). Hierrarchical mixtures of experts and the em algorithm. *Neural Computations*.

Karl, T. R. and Knight, R. W. (1995). Trends in high frequency climate variability in the tewentieth century. *Nature*.

Karlis, D. (2005). Am em algorithm for mixed poison and other discrete distributions. *ASTIN Bulletin*.

Kemp, W. P. D., Burnell, D. G., Do, E., and Thomas, A. J. (1983). Estimating missing daily maximum and minimum temperatures. *Journal of Climate and Applied Meteorology*.

Khalil, M., Panu, V. S., and Lenox, W. C. (2001). Groups and neural networks based streamflow data infilling procedures. *Journal of Hydrology*.

Kim, T. W. and Ahn, H. (2009). Spatial rainfall model using a pattern classification for estimating missing daily rainfall data. *Stochastic Evironmental Research and Risk Assesment*.

Kumar, A. (2000). Predicting annual maximum daily rainfall of ranichavsi (tehrin-garhwal) based on probability analysis. *Indian Journal of Soil Conservation*.

Kumar, A., Kaushal, K. K., and Singh, R. D. (2007). Prediction of annual maximum daily rainfall of almora based on probability analysis. *Indian Journal of soil conservation*.

Little, R. J. and Rubin, D. B. (1987). *Statistical analysis with missing data*. New York, Wiley.

Lloyd, C. D. (2005). Assessing the effect of integrating evaluation data into the estimation of monthly precipitation in great britain. *Journal of Hydrology*.

Manikandan, M., Thiyagarajan, G., and Vijayakumar, G. (2011). Probability analysis for estimating annual one day maximum rainfall in tamil nadu agricultural united university. *Madras Agricultural Journal*.

Mclachlan, G. J. and Krishnan, T. (2008). *The EM algorithm and Extensions*. John Wiley and Son, 2nd ed.

Nelwamondo, F. V., Mohamed, S., and Marwala, T. (2007). Missing data: a comparison of neural network and expectation maximization techniques. *Current Science*.

Nemichandrappa, M. P., Ballakrishnan, and Senthilvel, S. (2010). Probability and confidence limit analysis of rainfall in raichur region. *Karnataka Journal of Agricultrual Sciences*.

Orchard, T. and Woodbury, M. A. (1972). *A missing information principle: theory and applications; Proceedings of the sixth Berkeley symposium on mathematical statistics and probability*. Theory of statistics.

Redner, R. A. and Walker, H. F. (1984). Mixture densities, maximum likelihood and the em algorithm. *SIAM Review*.

Rubin, D. B. (1976). Inference with missing data. *Biometrika*.

Schneider, T. (2001). Analysis of incomplete climate data; estimating mean values and covariance matrices and imputation of missing values. *Journal of climate*.

Schwarz., G. (1978). Estimating dimension of a model. *Annals of Statistics*.

Shepard, D. (1968). A two dimensional interpolation function for irrugularly spaced data. *Proceeding of the Twenty-third National conference of the Association for computing machinery: Washington D.C*.

Simolo, C., Brunetti, M., Maugeri, M., and Nanni, T. (2009). Improving estimation of missing values in daily precipitation series by a probability density function-preserving approach. *International Journal of Climatology*.

Singh, R. K. (2001). Probability analysis for prediction of annual maximum rainfall of eastern himalaya (sikkim mid hills). *Indian Journal of soil conservation*.

Subudhi, R. (2007). Probability analysis of prediction of annual maximum daily rainfall of chakpada block of kandharmal district in orissa. *Indian Journal of soil conservation*.

Suhaila, J., Sayang, M. D., and Jemain, A. A. (2008). Revised spatial weighting methods for estimation of missing rainfall data. *Asia Pacific Journal of Amospheric Sciences*.

Tabachnick, B. G. and Fidell, L. S. (2001). *Using Multivariate statistics*. Needham Heights, 4th ed. Allyn and Bacon.

Tabios, G. Q. and Sala, J. D. (1985). A comparative analysis of techniques for spatial interpolation of precipitation. *Water Resource Bulletin*.

Teegavarapu, R. S. V. and Chandramouli, V. (2005). Improved weighting methods, deterministic and stochastic data driven models for estimation of missing precipitation records. *Journal of Hydrology*.

Upadhaya, A. and Singh, S. R. (1998). Estimation of consecutive days maximum rainfall by various methods and their comparison. *Indian Journal of soil conservation*.

Vivekanandan, N. (2012). Intercomparison of extreme value distributions for estimation of apmr. *International Journal of Applied Engineering and Technology*.

Wallis, J. R., Letten-Mayer, D. P., and Word, E. F. (1991). A daily hydroclimatological data set for the continental united state. *Water Resources Research*.

Wu, C. F. J. (1983). On the convergence properties of the em algorithm. *The Annals of statistics*.

Xia, Y., Fabian, P., Winterhalter, M., and Zhao, M. (2001). Forest climatology estimation and use of daily climatological data for bavaria, germany. *Agricultural and Forest Meteorology*.

Appendix A

Figure 5.1: Annual Total Rainfall(mm)for all the Stations (2000-2015)

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Abetifi	1395.8	967.7	1405.4	1005.7	1375.7	977.5	1304.8	1426.1	1341.7	1209.3	1627.3	1300.4	1229.7	1268.9	1483.6	1133.9
Ada	391.3	779.2	842.3	1061.9	698.4	790.5	746	897.1	767	1070	957.4	884.3	669.7	635.5	722.9	822.2
Akrokerri	1081	1257.6	1418.4	1231.5	1338.9	1427.2	1485.2	1644.6	1446.3	1496.6	1482.9	1270.6	1119	1080.6	1320	983.6
Axim	1321	1851.4	2310.5	1897.5	1834	1841.4	1367.8	2109.5	1428.3	1740.9	2501.8	2084.6	1578.4	1428.6	2254.2	1311.4
Bolgatanga	1000.9	985.8	879.1	960.5	827.3	877.3	923.7	1085	969.7	1090.5	981.9	820.4	1158.6	774.5	824.9	1005.5
Cape Coast	991.1	820.8	937.9	888.9	930.1	904.1	1026.4	779.6	870.8	1090.1	947.9	887.3	1095.4	604.5	952.4	1100.3
Ejura	1040.3	1241.9	1371.5	1313.8	1430.5	1272.6	1437.1	1429.8	1232.5	1319.6	1573.5	1191.5	1124.1	1192.2	1313	1256.1
Ho	1220.8	1076.4	1300.5	1301.3	1468.5	1148.7	1165.3	1490.1	1551.5	1116.5	1318.6	1282.9	1229.2	1037.7	1375	1230.5
Koforidua	1135	1324.5	1454.4	1167.5	1220.2	942.8	1737	1213.7	1381.5	1295.9	1236.8	1313.2	1129.9	1161.3	1266.8	1155.5
Krachi	1274.4	915	1453.8	1423.5	1491.6	1387	1033.3	1149	1853.9	1617.5	1107.2	1229.8	1294.8	1418.8	1376.1	993.5
Saltpond	715	882.6	947.5	925.5	1005.6	1267.7	978.5	1025.6	980.9	972.6	1013.3	1028.5	802.6	730.7	1211.2	1152.7
Sefwi	1320	1252.9	1306.6	1471.6	1268.6	1268.9	1448.6	1366.2	1681.1	1392.1	1497.5	1459.1	1355.5	1291.4	1738.9	1497.1
Tamale	1023.9	791.3	922.5	1249	1102	1142.5	970.7	1046.6	1120.1	1103.3	1248	1082.4	1062.4	909	1026.3	654.7
Wa	1141.2	1005.8	939.3	1199.4	1080.2	1054.1	1010.8	996.8	1275.1	1131.7	1032.2	946	1078.8	1046.4	791.4	728.3
Wenchi	855.3	987.2	1351.7	1364.9	1331.7	1327.9	1193.1	1341	1310.3	1170.2	1577.8	1172.3	1122.2	1162	1394.7	806.4

Appendix B

Figure 5.2: Sample Observed Data for Abetifi (200)

Observed data for Abetif (2000)											
0	NA	NA	1	NA	NA	NA	17.9	NA	11.8	15.8	NA
NA	NA	NA	NA	NA	NA	NA	0.2	1	NA	14.1	NA
NA	NA	NA	NA	0.7	NA	NA	NA	NA	0.4	NA	NA
NA	NA	NA	2.2	15.6	NA	NA	4.6	NA	NA	16	NA
NA	NA	NA	NA	NA	NA	NA	7.1	8.4	1.1	NA	NA
NA	NA	NA	NA	11.6	NA	NA	NA	NA	6.1	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	1.5	0.2	NA	NA
NA	NA	NA	NA	NA	NA	NA	26.6	1.4	NA	NA	NA
NA	NA	NA	NA	1.7	NA	NA	NA	NA	2.6	3.4	NA
NA	NA	NA	32.1	NA	NA	NA	NA	5.7	NA	0.4	NA
NA	NA	NA	NA	NA	NA	NA	NA	15.6	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	4.2	NA	NA
NA	NA	NA	13.6	NA	NA	NA	NA	0.2	0.8	NA	NA
NA	NA	NA	NA	0.3	NA	NA	6.3	2.4	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	4.3	NA	NA	NA
NA	NA	NA	12.2	NA	NA	NA	NA	NA	11.9	NA	NA
NA	NA	NA	17.4	NA	NA	NA	NA	NA	11.6	NA	NA
NA	NA	NA	24.3	NA	29.8	NA	NA	1.2	17.8	NA	NA
NA	NA	9.4	NA	NA	NA	NA	NA	6.7	NA	NA	NA
9.1	NA	NA	NA	5.2	3	NA	NA	0.4	4.2	NA	NA
NA	NA	NA	6.9	13.2	13.8	NA	NA	NA	NA	NA	NA
NA	NA	NA	26.6	NA	8.8	NA	NA	NA	49.2	NA	NA
NA	NA	NA	NA	10.5	NA	NA	NA	NA	4.1	NA	NA
NA	NA	NA	7.7	NA	NA	NA	NA	10.1	NA	NA	NA
27.5	NA	NA	1.9	NA	47.6	NA	NA	NA	12.5	NA	NA
12	NA	4.9	7.3	12.4	11.6	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	1.6	NA	NA	NA	3.2	0.3	NA	NA
NA	NA	NA	NA	NA	NA	NA	0.5	20.9	NA	NA	NA
NA	NA	NA	NA	30.2	NA	NA	4.7	NA	NA	NA	NA
NA		NA	NA	4.5	8.9	NA	NA	7.4	NA	NA	NA
NA		17.5		28.4		NA	5.7		9.8		NA

Appendix B1

Figure 5.3: Sample Complete Data for Abetifi (2000)

Complete data for Abetifi (Year 2000) using EM algorithm

14.5734	28.31147	47.95053	35.25745	10.10	6.88312	32.09875	7.80000	1.7979
1.5000	2.60932	48.2634	11.90000	16.50000	29.30218	20.10370	37.08220	51.04
33.60918	21.47694	6.6058	0.13717	16.47385	8.89569	3.10000	12.6124243	2.3235
30.0697	3.31044	0.90000	4.00000	145.60000	2.000	0.400000	4.90000	7.0345
11.24519	40.82033	12.32938	3.96500	17.9000	44.800000	40.7044	40.6000	15.996
11.04596	0.79338	6.81578	51.1183	25.30000	11.3000	0.700	15.5643	28.8384
0.77212	26.45027	11.19931	9.00000	1.00000	0.100000	8.205	29.48065	1.8821
16.4560	7.52725	47.912091	5.520785	18.70000	13.5000	16.40000	7.84524	4.70833
10.789	20.610243	9.31091	15.975	25.8000	0.90000	5.60000	40.30000	15.0434
76.810	34.5676	14.07519	12.32821	3.10000	1.20000	3.300000	4.100	15.4514
16.436	46.42986	7.60000	13.24281	0.8000	0.10000	17.90000	10.30000	17.6485
10.791	2.65349	9.22319	4.95552	39.68	9.300000	9.92202	6.39631	1.15893
31.02506	14.64725	5.20000	0.20000	2.900000	0.80000	39.70000	16.16773	0.492
46.0000	9.50972	2.9388	17.99660	4.70000	0.8000	21.30000	42.1336	31.1037
32.49810	4.99029	10.487728	18.9287	3.11711	11.71	2.80000	18.7000	23.3349
8.9564981	12.170749	5.70085	35.99744	6.60000	23.5000	4.10000	0.2000	1.6162
.715	41.3288	19.30435	1.04539	6.10000	2.800000	1.200	33.78706	14.352
5.73610	7.6972	2.500507	13.10000	3.30000	0.20000	1.600	33.50000	17.9946
9.893164	23.564436	6.5691	4.5000	0.10000	38.40903	22.91540	3.30000	0.92506
19.776118	22.3223	25.221864	6.1221	1.00000	9.33640	0.500000	10.73815	28.7605
15.6293	0.139267	9.18834	55.60000	4.103098	14.3507	7.495606	50.28	0.29695
31.4050	42.04781	20.03823	23.33044	17.400	2.50000	5.1000	23.1649	6.36452
17.98467	16.52196	19.26081	16.5339	3.985	47.45473	2.000000	26.70689	29.020
18.1224	18.3703	43.7899	13.20000	17.59447	27.101184	10.76076	41.1833	35.89
5.932480	8.19415	13.74930	16.63170	63.90391	6.5442	10.0000	29.5889	27.6171
22.13734	17.17741	3.70000	16.3663	20.10000	8.243	12.4000	47.3211	15.1809
3216.84587	42.05968	10.79964	3.102506	20.3995	32.80064	4.79089	42.9252	21.6951
3.74031	1.32295	37.384111	8.200000	22.94249	0.20878	28.8000	8.330267	66.6303
25.093103	15.092101	1.7462	19.57506	13.3161	19.79277	23.60	20.3890	16.6821
19.30388	22.4038	4.97917	1.4445	2.10000	11.9000	5.30000	23.4968	0.80000
25.9597	22.666251	32.0000	1.800	48.70000	14.1213	11.200000	52.2605	3.75962
17.62858	45.23115	2.105688	24.67871	2.80000	19.39969	1.7000	3.100	4.3233
19.95108	22.5717	12.49345	7.7000	42.177	8.261842	32.456156	0.749502	38.6258
7.48421	6.363981	14.5255	39.60000	1.292	12.249707	14.42665	36.08637	12.14
20.27816	0.30000	41.69172	4.30000	13.85699	8.76062	7.78294	32.9644	22.956218
35.38939	3.91636	46.4461	29.8243	6.015302	0.30000	1.40000	3.6843	22.024759
0.53823	19.71693	34.800000	25.35626	15.4370	9.818	5.40000	3.9000	17.810084
23.79391	91.6000	5.400000	14.69448	17.29788	3.20000	46.2084	39.3547	6.236397
36.89810	19.700000	34.200	4.259677	26.5818	4.30000	5.97337	0.60000	
20.88132	9.59315	40.54554	0.60000	24.50158	14.50000	1.900	7.71219	
55.3476	23.99166	17.99367	11.10000	8.45928	35.20096	5.35588	14.64302	

Appendix C

Figure 5.4: Sample Observed Data for Ada (2000)

Observed data for Ada (2000)											
NA	NA	NA	NA	18.4	NA	6	26.2	NA	NA	NA	NA
NA	NA	NA	NA	6	7.3	0.2	0.8	NA	NA	NA	NA
NA	NA	NA	1.7	NA	NA	0.9	NA	NA	0.6	NA	NA
NA	NA	NA	NA	NA	NA	NA	6.3	NA	NA	NA	NA
NA	NA	NA	NA	NA	0.2	NA	3.2	NA	NA	NA	NA
NA	NA	NA	NA	NA	0.4	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	0.3	NA	NA	NA	0.3	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	2.8	26.3	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	14.7	NA	NA	NA	NA	4.9	NA	NA	NA
NA	NA	NA	NA	NA	4	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	2.8	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	25.5	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	12.4	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	12.4	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	0.8	NA	NA	NA	NA	NA	NA	NA
NA	NA	6.8	NA	28	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	0.3	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	15.7	NA	NA	0.3	NA	NA	1.6
NA	NA	NA	NA	9.6	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	1.8	NA	NA	NA	0.5	NA	NA	NA
NA	NA	NA	NA	11.4	NA	NA	NA	NA	NA	NA	NA
NA	NA	37.8	NA	1.8	NA	NA	0.5	NA	11.1	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	21.2	NA	NA
NA		NA	NA	NA	26.6	29.2	NA	NA	NA	NA	NA
NA		NA		1.7		NA	NA		NA		NA

Appendix C1

Figure 5.5: Sample Complete Data for Ada (2000)

Complete Data for Ada (Year 2000) using EM Algorithm

19.75559	8.84194	8.708349	18.4000	4.4377	6.8557	6.63224	13.57051	9.49
17.91815	0.589055	5.846006	6.00000	0.238570	7.404656	12.892	36.89827	1.8507724
3.530207	23.499208	8.121685	5.173722	4.000000	20.58196	14.5058	10.96806	9.54648
8.04207	1.376626	27.18084	5.985819	6.826384	18.971467	29.4383	4.25942	32.79870
1.90681	16.03910	2.52604	8.70774	35.48287	12.74433	2.69	4.55206	1.5646
0.850668	6.21537	37.80000	9.182552	6.97923	23.4390	21.04	22.359	14.35321
7.909150	8.39199	9.61468	0.26497	14.21807	6.5672683	19.4209	3.99562	1.55742
42.18326	20.25319	7.095800	0.3000	9.2206	7.45859	0.39497	14.22519	12.9869
10.84075	26.4113	0.78199	7.03514	24.2039	14.2040	0.30000	13.74163	11.1522
3.2945005	3.685188	7.7522719	26.300000	2.9788364	29.200000	22.7383	11.55362	5.9780
8.414072	9.911654	5.267199	12.1344	9.5014	6.1989	6.31137	5.5452	21.20
19.17878	7.244548	12.32935	20.9954	5.430447	26.2000	2 4.90	4.929	24.210
6.9573456	4.8983237	1.700000	2.850358	0.30000	0.8000	9.90946	5.00833	1.698
19.29140	3.434577	3.605768	5.485918	15.700000	3.278718	2.80	22.94	5.7081
0.67022	23.76319	5.717930	14.4195	5.706355	6.30	3.24	8.18	0.395145
7.44361	6.30362	15.40138	10.89112	11.2877	3.20	28.8041	11.100	15.0985
12.57426	14.5409	7.07770	10.94187	16.10740	4.79480	3.51390	21.20	20.66889
5.80618	13.89615	12.71034	12.4000	15.0275	19.15701	1.34426	14.8893	16.409756
18.6113	2.37995	12.3068	10.8041	6.69922	9.79714	8 5.997	1.62168	10.59909
11.6442528	24.136444	2.8000000	0.800000	5.8809704	6.34810	12.6636	13.5789	0.19601
11.71576	10.88489	14.7000	28.000	26.600	26.2964	3.28477	18.83512	9.48432
5.387450	11.84445	25.04085	3.91829	6.00	0.7934	4.75194	14.32768	20.5016
9.184778	1.9462806	9.445494	12.469892	0.20	5.0881	5.50	0.469331	0.4867911
11.05305	11.22517	3.298901	9.600000	0.900	21.5586	0.3000	15.0206	8.576530
4.249404	24.14261	25.50000	3.05759	17.519648	4.239990	10.3291	8.335782	4.472232
15.3345	2.994452	12.40000	1.80000	2.310479	10.38773	11.4697	21.2458	2.9315368
12.32911	10.1426	21.7153	11.40000	8.38755	17.6123	0.50	0.5222	2.4846504
15.7844	9.32312	1.40110	1.8000	22.24739	1.76515	2.9	5.102873	2.50036
11.4594	14.0747	0.27725	10.5041	7.15457	5.62824	2.56312	12.84262	9.205067
12.6650998	25.7890527	1.2144975	17.790708	9.75772	6.08571	8 18.455	18.45558	17.29754
1.57772	14.02559	4.1782	1.7000	26.9434054	24.18097	14.4218	7.701240	1.6000
8.219600	5.611461	5.23316	15.47115	16.377	19.8479	21.4686	22.3115	19.926419
20.679775	4.340799	23.951582	7.300000	1.089762 -	7.8823	5.14326	0.005233	5.890822
2.587069	12.543542	26.41979	9.936177	1.97181	28.41483	0.600	9.072175	5.73924
10.0448	2.67256	23.35655	4.51902	19.33585	9.718277	11.649	2.30072	14.055495
21.4630	10.1474	7.03362	0.200000	2.4160365	16.17232	2.78789	6.898	25.575748
13.14803	27.2008	20.61128	0.40000	12.5829	23.6876	17.2384	3.5379	10.08066
31.04095	27.3632	2.2364	10.8599	9.638642	17.66852	2.46181	1.934720	19.39689
17.0053	6.8000	0.249246	22.9079	7.7636	0.50000	5.61174	15.4678	11.2270
0.9530388	1.1217167	28.845337	3.92587400	6.59672	26.8740	4 17.373	3.63249	2.25374
20.9410	16.6449	5.57340	20.198	8.66698982	24.591958			

Appendix D

Figure 5.6: Sample Observed Data for Akrokerri (2000)

Observed data for Akrokerri (2000)											
NA	NA	NA	17.6	NA	NA	NA	24.4	NA	NA	NA	NA
NA	NA	NA	NA	8.2	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	27	2	11.2	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	10.2	NA	NA	NA
NA	NA	NA	NA	10.4	NA	NA	NA	NA	4.2	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	40	NA	5	36	NA	NA	NA
NA	NA	NA	16.2	NA	NA	36	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	9.6	NA	14.4	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	5.6	NA	NA	21.8	NA
NA	NA	NA	NA	NA	7.8	NA	1.8	4.2	NA	11	NA
NA	NA	NA	NA	22	5.4	NA	NA	2	NA	NA	NA
NA	NA	NA	NA	NA	11.6	NA	NA	NA	NA	NA	NA
NA	NA	53.6	NA	NA	NA	NA	3.6	16.4	NA	NA	NA
NA	NA	NA	8.2	8	8.2	NA	3	NA	10.6	NA	NA
NA	NA	NA	NA	NA	6.8	NA	NA	NA	NA	NA	5.2
NA	NA	5	NA	4	8.6	NA	2	NA	NA	NA	NA
NA	NA	NA	NA	NA	1.2	NA	NA	NA	17	NA	NA
NA	NA	NA	NA	67.4	NA	NA	NA	26.4	1.8	NA	NA
NA	NA	2	NA	NA	32.6	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	28.6	2.4	NA	NA	NA
NA	NA	NA	NA	13.4	4.6	NA	NA	NA	NA	NA	NA
NA	12	36.6	NA	NA	NA	18.4	NA	NA	NA	NA	NA
NA	NA	4.3	8.2	2.8	9.2	NA	NA	4.2	NA	NA	NA
NA	NA	NA	10	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	9.4	27.8	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	15	NA	NA	NA	5.4	NA	NA
NA	NA	4.5	NA	11	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	40.4	NA	NA	NA
NA	NA	5.2	NA	NA	NA	NA	31.6	NA	NA	NA	NA
NA		NA	29.2	NA	NA	10	1.4	20.6	53.6	NA	NA
NA		NA		4		NA	NA		NA		NA

Appendix D1

Figure 5.7: Sample Complete Data for Akrokerri (2000)

Complete Data for Akrokerri (year 2000) using EM Algorithm							
4.80518	4.11808	39.249	1.97580	9.20000	12.4650	26.4000	24.94076
29.1653	10.7766	5.20000	22.00000	21.645	5.0000	19.8447	5.23349
26.6699	9.06137	1.09490	7.2186	14.85799	21.6258	2.40000	20.76554
2.45865	29.8412	15.39318	18.8151	15.0000	10.75432	22.0357	52.4466
13.2575	38.2044	17.60000	8.000000	35.2065	5.60000	42.31531	17.2312
4.9253	7.34044	11.9723	30.849	6.38116	1.80000	4.20000	8.12030
17.0717	15.7965	3.39766	4.0000	15.23934	2.11711	5.998063	8.51774
13.0769	12.1744	12.8638	6.20668	9.71065	8.8474	30.9119	32.70235
59.6240	8.98798	9.48894	67.4000	5.41404	3.60000	28.7108	3.09072
17.0583	12.0000	19.07993	9.78598	17.66530	3.000000	2.87205	21.800
2.1385	7.0001	7.232	21.91848	27.0000	28.3524	40.4000	11.00000
27.3434	34.6080	16.2000	13.40000	24.2108	2.00000	33.2162	21.65464
28.3820	10.8965	10.10107	17.1267	22.7442	15.6409	20.6000	20.99793
11.7843	22.0834	23.25201	2.8000	11.4338	10.9569	10.9070	18.0264
28.5349	21.2077	11.94776	17.1956	10.3225	38.04836	15.79352	9.8665
1.4254	5.56782	19.5973	27.800	36.0000	28.6000	6.73726	4.35884
12.4446	35.1149	19.04924	17.0085	9.60000	1.2581	41.45522	9.13739
19.4125	17.1182	36.3432	11.00000	26.1287	9.24574	4.20000	33.49073
5.54962	18.4214	8.2000	7.6570	0.80216	31.6140	7.10783	13.43891
27.6113	0.30755	15.16341	19.27080	13.7266	3.422603	0.5100	22.5566
18.1495	17.5803	2.144	6.48811	32.5495	16.4431	5.8082	0.13328
18.2466	35.1233	31.8268	4.00000	12.0522	26.2546	14.86257	20.7770
9.65226	6.4023	4.23847	11.92979	15.5874	0.06156	6.79665	27.91530
10.1380	16.1101	1.959131	26.4969	38.92703	9.9792	8.78919	21.79384
17.3466	10.3259	3.98504	23.34676	24.577	31.60000	5.13656	2.9730
8.1066	21.450	3.33872	15.8298	0.85567	1.40000	16.36352	22.7349
23.1612	37.3593	4.77141	8.4729	0.3422	5.9292	17.91254	13.6563
19.0796	21.3835	8.2000	17.08433	28.59534	35.17542	10.6000	31.18915
23.7722	9.95648	10.0000	40.0000	5.616834	29.2908	6.28672	3.04484
17.8984	8.23082	9.40000	33.4466	19.42429	11.20000	5.8165	9.26578
19.5359	53.6000	34.863	7.667319	18.4000	10.2000	17.0000	19.7769
0.19298	19.3708	38.2159	8.36248	15.4257	13.04044	1.80000	27.3999
8.8272	1.2939	34.0558	7.80000	8.2078	40.9253	27.3989	12.79457
30.4205	5.0000	7.21658	5.4000	11.29454	36.000	21.9217	32.63662
5.8491	11.4453	29.2000	11.60000	14.1061	10.8625	4.33019	2.342
11.306	39.276	30.32748	2.65965	11.6462	14.4000	4.64933	41.81807
31.4842	2.00000	8.200000	8.2000	12.3917	24.2990	18.15657	5.46022
20.1917	39.4971	5.3729	6.80000	30.28765	4.20000	6.12184	25.28524
44.4918	0.81227	2.6741	8.6000	10.0000	2.00000	25.74692	2.469208
25.4303	36.6000	10.4000	1.20000	28.10047	34.50545	5.400	4.96318
3.62996	4.30000	41.510	6.93513	24.4000	16.400	5.67900	3.34222
9.67242	14.16231	9.3619	32.600000	19.6435	26.330964	9.95686	7.26889
3.13564	10.2750	5.79358	50.52432	2.0000	38.8328	25.92993	29.7663
34.2495	13.3655	14.16149	4.60000	34.16777	35.7335	53.6000	15.2262
4.20523	4.50000	14.8063	19.85210	11.25455	11.3427	30.7753	4.8491

Appendix E

Figure 5.8: Sample Observed Data for Axim (2000)

Observed data for Axim (2000)											
11.2	NA	NA	26	0.4	0.4	NA	1.9	NA	NA	0.1	NA
4.9	NA	NA	NA	NA	NA	1.1	0.8	6.8	NA	NA	NA
NA	NA	NA	0.3	3.5	NA	NA	NA	2.7	0.3	NA	NA
NA	NA	6.9	NA	5.3	NA	0.7	NA	NA	0.2	NA	NA
NA	NA	NA	NA	45.5	NA	NA	1	NA	1.8	NA	NA
NA	NA	NA	NA	NA	3.3	NA	0.8	NA	NA	NA	NA
NA	NA	12.5	41.2	NA	40.7	NA	NA	1.1	NA	NA	NA
NA	NA	NA	0.4	NA	NA	NA	NA	NA	NA	1	NA
NA	NA	0.2	NA	NA	NA	NA	NA	NA	NA	NA	0.7
NA	NA	NA	NA	NA	29.2	NA	NA	1.2	NA	NA	NA
NA	NA	NA	NA	NA	87.1	NA	NA	0.2	10	3.4	NA
NA	NA	NA	23.2	NA	1	NA	NA	0.4	NA	5.9	NA
NA	NA	2.8	NA	5	NA	NA	NA	0.2	NA	NA	NA
2.2	NA	1.3	NA	0.7	2.4	NA	NA	NA	0.1	NA	NA
2.9	NA	NA	18	7	9.2	NA	NA	0.7	NA	NA	NA
NA	NA	0.2	NA	2.1	44.5	NA	2.4	3	NA	NA	NA
NA	NA	NA	3	1.3	20.9	NA	0.3	7.3	NA	NA	NA
NA	NA	4	31.8	NA	NA	NA	1.1	0.9	35.8	1.6	18
NA	NA	NA	NA	14.4	6.8	NA	13.4	0.4	5.1	1.1	NA
NA	0.4	0.2	NA	NA	0.7	NA	3.5	NA	NA	NA	16.1
NA	NA	1.1	NA	12.6	2	NA	4.1	NA	NA	50.1	NA
NA	NA	9.6	NA	35.7	NA	19.6	NA	NA	NA	5.4	NA
NA	1.2	0.9	0.5	4.8	NA	NA	14.2	NA	NA	NA	NA
NA	NA	NA	NA	NA	5.4	NA	2.5	NA	NA	NA	NA
NA	NA	3.2	NA	NA	57.5	NA	NA	NA	NA	NA	NA
28	NA	NA	NA	NA	42.3	NA	NA	NA	NA	NA	NA
NA	9.7	NA	NA	120.1	25.1	NA	NA	0.3	NA	NA	NA
NA	1.2	NA	33.2	101.8	0.4	NA	NA	0.6	0.4	2.2	NA
NA	NA	NA	NA	2.1	NA	0.1	1.7	NA	3.2	NA	NA
NA		3.3	NA	5.8	NA	3.4	0.3	4.7	NA	NA	NA
NA		0.4	NA	1.7	NA	2	0.6	NA	NA	14.5	NA
NA		NA		NA		NA	3.4		NA		NA

Appendix E1

Figure 5.9: Sample Complete Data for Axim (2000)

Complete Data for Axim (year 2000) using EM Algorithm							
11.2000	4.5328	20.0129	11.3535	5.4000	0.8000	0.4000	0.1000
4.9000	22.8556	3.3000	12.487	57.5000	0.74329	17.2224	8.42927
16.15149	5.2137	0.4000	5.000	42.300	11.458	11.7148	6.83826
30.639961	37.6429	41.9264	0.7000	25.1000	3.842787	20.8384	24.494
27.202759	3.7404	26.000	7.0000	0.40000	16.990	6.02446	23.937
12.9193532	23.6877	19.9212	2.10000	9.9734	14.7996	9.8972	37.9954
8.7281954	0.4000	0.30000	1.3000	10.8616	23.8155	6.50884	26.7358
2.7486671	5.31108	4.18139	34.3059	6.8113	21.7953	7.53582	1.0000
13.982061	22.0140	1.80445	14.4000	19.4758	6.21620	0.30000	43.9560
8.47955	1.20000	17.1488	3.694651	1.10000	4.68555	0.60000	39.68704
72.594	31.5707	41.2000	12.6000	16.3834	2.40000	32.1858	3.40000
13.9635	43.0904	0.4000	35.7000	0.7000	0.3000	4.7000	5.9000
12.47843	0.57804	11.3150	4.8000	32.9592	1.10000	29.1731	6.09091
2.2000	9.700	25.2978	6.8342	0.9836	13.4000	17.5244	17.8014
2.9000	1.2000	44.567	4.0437	3.94655	3.50000	37.53030	20.8195
28.1304	12.2255	23.200	14.1315	20.65799	4.1000	0.3000	48.753
29.56096	7.2363	44.871	120.100	] 14.058	26.4573	0.2000	1.2709
6.6990	2.84737	8.41393	101.80	14.1527	14.2000	1.80000	1.60000
29.7716	0.1092	18.000	2.1000	13.8949	2.50000	5.96937	1.1000
7.5693	38.1367	9.97456	5.8000	1.01409	10.6378	7.63672	33.0456
7.60868	6.9000	3.00000	1.700000	17.01104	9.37446	20.2549	50.1000
17.206	5.476184	31.8000	16.1049	0.5959	35.3012	5.28035	5.40000
17.1768	20.8852	4.6202	0.40000	25.9650	7.0679	12.4489	30.014
28.4994	12.5000	8.87712	38.489	26.964	1.700	10.00	5.5768
15.4666	19.679	44.5299	43.1062	22.6253	0.300	2.653728	36.2196
28.000	0.2000	11.0409	37.3760	12.2713	0.600	29.52022	5.4907
15.6003	1.8636	0.5000	19.473	2.137852	3.4000	0.1000	12.221
3.76236	38.8349	11.6699	3.300000	13.999	11.937	12.011	2.20000
23.49701	14.0461	6.95806	40.700	36.5369	6.8000	5.55939	0.2527
14.3607	2.80000	4.852	32.2406	19.6000	2.70000	42.8755	47.5681
1.6334	1.30000	8.186082	2.1320	1.02828	44.0857	35.8000	14.5000
22.3698	15.8411	33.2000	29.2000	1.98579	24.3199	5.10000	0.25763
16.7477	0.20000	3.53741	87.10000	5.8693	8.3542	7.79974	8.8302
23.2114	9.95639	16.7481	1.00000	19.0853	1.1000	3.20241	17.533
15.1208	4.0000	0.42947	5.8494	60.0599	10.0042	34.012	30.004
17.3762	14.6826	0.4000	2.4000	17.8117	29.85	14.2471	0.17099
9.2669	0.2000	4.38056	9.2000	0.1000	1.200	13.1161	2.5736
21.6915	1.1000	3.5000	44.5000	3.40000	0.200	26.6306	16.608
32.3687	9.6000	5.3000	20.90000	2.0000	0.400	9.618	13.0065
1.4761	0.90000	45.50000	47.644	20.2812	0.2000	4.21279	0.7000
25.1058	38.8465	22.495	6.8000	1.90000	1.7961	0.40000	15.1402
33.8339	3.20000	6.92421	0.70000	0.8000	0.7000	3.2000	0.8734
18.2796	0.71408	17.46084	2.0000	10.93279	3.00000	17.6997	1.52093
51.7508	12.6575	16.7058	3.362549	38.9611	7.30000	38.91825	28.2069
25.4952	23.7558	40.5268	17.5129	1.0000	0.9000	30.8127	20.6624
3.5683	16.1000	18.000	1.10048	15.4762	15.9368		

Appendix F

Figure 5.10: Sample Observed Data for Bolgatanga (2000)

Observed data for Bolgatanga (2000)											
1.3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	24.3	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	8.7	14.4	39.7	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	14.9	3.6	NA	0.7	NA	NA	NA
NA	NA	NA	NA	15.8	NA	25.4	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	8.4	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	5.2	5.8	4.4	11.3	NA	30	NA	NA
NA	NA	NA	2.2	NA	10.2	35.4	6.4	NA	28.3	NA	NA
NA	NA	NA	NA	NA	NA	NA	1	20.1	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	2.4	6.5	NA	NA
NA	NA	NA	4.6	NA	NA	23.2	45.8	NA	NA	NA	NA
NA	NA	NA	NA	47.6	19.5	NA	37.3	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	6.8	NA	18.2	11.5	9.4	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	0.3	9.1	NA	NA
NA	NA	NA	NA	NA	54.5	4.7	12.6	0.2	NA	NA	NA
NA	NA	NA	NA	1.8	NA	NA	15.7	9.9	NA	NA	NA
3.5	NA	NA	NA	37.3	39.3	NA	NA	1.4	NA	NA	NA
NA	NA	NA	9.2	4	NA	37.6	NA	NA	NA	NA	NA
NA	NA	NA	3.8	NA	10.5	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	19	3	NA	NA	35.5	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	27.7	1.3	13.2	NA	NA	NA	NA
NA	NA	NA	NA	1.2	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	1.3	NA	NA	17.8	NA	NA	NA
NA	NA	NA	NA	NA	NA	22.9	NA	NA	NA	NA	NA
NA		NA	NA	NA	0.7	NA	2.2	NA	NA	NA	NA
NA		NA	NA	9	NA	13.9	17.3	NA	NA	NA	NA
NA		NA		NA		NA	9.2		NA		NA

Appendix F1

Figure 5.11: Sample Complete Data for Bolgatanga (2000)

Complete Data for Bolgatanga (year 2000) using EM Algorithm							
1.30000	7.536212	4.6563	29.79599	0.7000	45.80000	17.8000	26.90056
5.98524	29.48308	12.08269	30.18551	2.779884	37.30000	11.06832	1.55891
10.15841	15.29882	1.9059	15.28602	36.9615	25.5363	14.8145	15.77297
25.74053	5.17571	9.88226	13.42459	11.63051	9.400000	9.39526	16.92466
30.46728	15.37132	4.13575	1.8000	8.70000	31.61981	40.80947	5.92579
24.77556	46.1769	15.15433	37.3000	4.83853	12.6000	26.18202	22.58573
4.66017	4.79658	10.08204	4.00000	3.600000	15.7000	4.69502	37.43852
20.70254	21.7395	27.44816	13.7289	25.4000	13.49776	26.8877	2.39679
51.60816	4.78707	12.936208	19.00000	12.61967	10.72519	23.29136	26.70246
17.2547	19.929443	9.62625	8.69652	1.093678	26.9507	13.09301	3.94722
8.3668	17.49876	2.200	11.52362	4.4000	26.49808	17.2524	27.11074
8.75449	10.64482	39.12825 -	1.20000	35.40000	17.08399	1.814902	29.25144
32.3472	30.81094	15.3082	10.91358	10.6076	13.2000	30.0000	6.651697
2.569842	34.7989	4.60000	20.17358	18.14705	31.2622	28.3000	10.83976
21.56987	0.47375	5.89267	11.73908	23.2000	12.129557	5.622738	2.42462
20.92923	4.8386	21.48415	9.00000	29.32745	26.3979	6.500000	12.64827
18.03046	3.24723	6.80000	22.84047	4.628829	2.20000	2.09631	27.86958
10.07028	22.7269	8.375071	44.6452	11.5000	17.30000	15.81467	18.5407
3.80693	30.1068	18.50263	13.13499	15.39076	9.20000	45.61869	27.65077
3.50000	24.46408	21.30172	28.18535	4.700000	36.43729	33.298	0.08190
9.3589	16.76586	13.0232	3.319797	7.8809	25.25402	9.10000	19.64413
33.1163	2.464032	9.2000	14.9000	16.57023	39.7000	24.40655	43.36658
13.5552	15.288967	3.80000	30.42750	37.6000	19.82453	28.4479	29.11515
22.44977	29.8852	2.321012	10.57481	29.57137	0.70000	23.54952	23.17612
0.575255	2.07880	4.3967	29.56876	6.95644	4.18152	22.80898	22.6762
20.71374	14.0862	17.40909	5.80000	1.657285	8.40000	4.256899	7.041799
27.67731	8.6487	6.75724	10.20000	1.300000	8.52997	15.85656	14.56366
21.70566	6.60755	12.333398	12.0117	21.3266	28.40648	1.09211	8.33298
3.345514	6.0155	7.125867	33.74246	26.97548	34.58999	15.90026	34.0800
22.6237	0.58886	5.06271	9.84405	22.900	20.10000	25.5932	19.39823
13.7673	14.91899	8.8077	19.50000	29.9365	2.40000	8.54465	18.63400
30.8710	25.64027	7.32858	20.17492	13.90000	11.8575	33.62881	11.10807
3.4155	29.123110	24.3000	18.2000	23.2283	34.98012	20.94882	8.50935
9.48423	10.52814	0.685147	6.142668	21.68048	20.25534	53.86426	24.36479
19.738182	4.87987	28.8510	54.5000	3.51885	20.19625	21.07520	30.97846
27.17448	21.3451	5.841845	4.3085	14.40000	0.30000	5.620471	19.7389
12.92666	36.6076	15.8000	39.30000	1.263660	0.20000	5.074388	11.8860
32.28307	16.0790	14.613818	15.1764	4.26986	9.900000	25.7703	22.9163
2.73066	28.4216	13.3453	10.50000	16.80977	1.40000	12.06970	1.8519
41.23980	12.1503	5.20000	3.00000	17.15805	38.47976	11.0891	4.848300
5.7718	5.709577	11.6128	21.25325	8.65489	35.58523	15.14963	2.02780
25.1116	6.71755	10.73413	27.70000	11.3000	35.5000	14.1858	0.600531
1.9635	35.763662	30.3594	19.0537	6.40000	3.619423	31.77948	14.24255
5.28694	7.517886	16.95724	1.300000	1.000000	27.25301	28.04234	14.6029
23.216153	28.0323	47.6000	13.19141	11.4376	14.61228	14.8469	9.46008
15.52561	7.772609	1.23157	9.611207	20.9499	11.30628		

Appendix G

Figure 5.12: Sample Observed Data for Cape Coast (2000)

Observed data for Cape Coast (2000)											
18.6	NA	NA	NA	NA	NA	2.5	NA	NA	NA	NA	NA
2.2	NA	NA	NA	NA	48.2	6.5	3	NA	NA	NA	NA
NA	NA	NA	NA	NA	132.2	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	43.1	0.3	NA	NA	NA	NA
NA	NA	7.9	6.5	NA	NA	NA	0.9	NA	4.9	4.8	NA
NA	NA	1.4	NA	NA	20.6	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	0.8	NA	NA	NA	55.6
NA	NA	NA	NA	0.9	NA	NA	NA	0.4	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	0.3	NA	NA	NA	NA
NA	NA	NA	NA	26.6	NA	NA	2.1	NA	NA	NA	NA
NA	NA	NA	78.4	NA	NA	0.1	0.4	0.7	NA	NA	53.2
NA	NA	NA	NA	NA	5.1	NA	NA	NA	NA	NA	0.7
NA	NA	NA	NA	NA	4.3	NA	NA	0.9	NA	NA	NA
1.3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	1.3	NA	NA	NA	NA	NA	NA	NA	23.3	NA
NA	NA	NA	34.5	NA	1	NA	NA	NA	NA	0.3	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	2.7	8.7	0.7	NA	NA	NA	NA	NA	NA
NA	0.6	NA	2.2	28.9	NA	NA	NA	NA	NA	17.7	NA
NA	NA	55.9	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	9.4	0.3	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	6.5	9.5	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	20.9	NA	NA	1	NA	NA	NA
NA	NA	27	NA	16.2	15	NA	NA	NA	NA	NA	10.6
5.5	NA	0.3	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	24.5	NA	2.6	0.3	NA	NA	NA	NA	NA	NA
NA	NA	1.5	NA	30	NA	NA	NA	NA	NA	NA	NA
NA		NA	NA	NA	NA	NA	NA	NA	19.2	NA	NA
1.2		NA	NA	6.5	NA	1.6	2.7	NA	33.3	NA	NA
NA		NA		NA		9.5	NA		NA		22.8

Appendix G1

Figure 5.13: Sample Complete Data for Cape Coast (2000)

Complete Data for Cape Coast (year 2000) using EM Algorithm							
18.60000	3.811915	7.47301	28.9000	43.1000	22.2007	16.1340	44.0207
2.20000	45.21647	13.5636	14.5784	70.8618	45.0876	4.90000	12.4236
37.20503	2.12649	6.50000	9.40000	22.5294	8.98163	1.86341	26.7911
33.27284	29.25163	7.43811	6.50000	25.3546	10.8890	56.5710	8.54294
12.62725	0.60000	12.3081	5.63956	14.6598	25.3244	2.44735	23.9869
12.13776	8.228551	53.0952	16.2000	46.7244	8.19339	7.94928	35.2351
0.99189	23.0316	10.4782	20.1279	1.30228	12.0890	17.9053	25.5891
18.14823	38.26992	15.5032	2.60000	0.10000	5.1885	32.1730	4.06813
11.85330	51.44854	78.4000	30.0000	15.2607	35.9240	1.95700	27.0722
85.20088	2.81389	10.1127	20.8954	6.54880	15.8936	5.0968	12.7663
18.1270	16.13872	3.39903	6.50000	17.2842	2.70000	16.8470	40.3940
12.1228	10.43103	11.5175	39.8583	19.0835	8.51261	17.0322	3.95499
34.33413	5.410044	6.19945	1.02734	29.3978	51.2025	19.4731	5.84778
35.97064	2.277583	34.5000	48.2000	27.0866	6.77037	1.15347	22.4110
1.30000	45.78141	21.3126	132.200	9.26402	5.81619	0.41266	34.4229
9.816410	8.417424	2.70000	6.66751	7.51293	41.0634	34.4216	11.4083
36.21164	26.04553	2.20000	25.7855	32.4200	14.1462	25.7905	42.6749
6.506790	24.66564	2.64393	18.2348	10.0170	17.1575	1.92959	55.6000
10.85703	7.900000	7.16403	20.6000	12.8770	32.6183	16.0792	5.06130
21.8368	1.40000	27.8868	18.3434	42.5375	8.84997	19.8576	57.1427
17.4979	0.02065	10.0739	18.0485	10.2384	0.40000	0.89365	0.1489
34.7563	46.58017	22.1279	3.31274	15.8092	6.97209	31.8183	53.2000
19.84653	18.22148	21.2643	21.6134	52.5872	18.0960	0.19584	0.70000
19.99956	20.27496	48.5156	1.47081	29.9747	0.70000	6.93678	31.0908
6.45679	9.2376	15.1411	5.10000	7.40463	46.6754	19.2000	12.6438
5.500000	18.94967	12.1322	4.30000	1.60000	0.90000	33.3000	0.93213
24.72811	46.59337	41.3988	27.5516	9.50000	37.4026	32.1066	28.0290
18.58134	1.33570	2.07410	33.0001	9.29225	11.7957	39.7418	2.70114
4.021338	1.30000	5.66583	1.00000	3.00000	55.7360	30.5479	38.1520
1.20000	16.63294	2.47344	28.0362	36.3067	25.8699	16.9998	15.2401
27.74384	25.02426	14.0140	0.70000	0.09788	29.5367	23.9688	1.11180
21.31214	25.04766	16.2716	16.1911	0.30000	45.6197	73.8908	15.3572
28.70665	50.11680	46.1846	4.59834	0.90000	32.7386	4.80000	65.1176
19.45092	55.9000	51.4666	0.30000	21.8552	52.4387	18.3994	1.72421
22.0312	24.9426	44.9111	9.50000	0.80000	47.5549	4.04278	10.6000
8.44887	6.936178	0.90000	20.9000	15.5544	1.00000	4.66904	25.6437
22.66269	4.216926	20.1246	15.0000	0.30000	9.12068	42.7784	1.73956
39.18280	27.00000	26.6000	18.1680	2.10000	22.5176	13.6232	22.7200
0.463892	0.30000	39.0362	0.3000	0.40000	25.9703	25.3698	18.7937
26.56862	24.5000	0.28646	43.9511	21.6867	57.9263	24.3349	7.72249
40.85895	1.50000	4.53912	3.32897	9.04466	0.6986	19.6526	40.2969
23.06464	21.77103	56.6573	4.42438	13.4750	39.9571	6.79443	46.7388
61.35603	10.79187	5.99940	2.50000	9.59879	36.4887	23.3000	22.8000
31.31939	26.7883	17.8823	6.50000	10.7736	4.22727	0.30000	17.7000
3.03298	53.1379	13.5623	4.56196	38.9734	43.5882	15.6215	
23.9945	53.48556	8.70000	19.6812	35.5269	8.4340	5.6455	

Appendix H

Figure 5.14: Sample Complete Data for Ejura (2000)

Observed data for Ejura (2000)											
NA	NA	NA	NA	9.4	15.5	NA	54.1	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	9.7	NA	NA	NA	NA
NA	NA	NA	15	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	13.3	NA	NA
NA	NA	NA	19.1	NA	NA	NA	33.5	NA	NA	NA	NA
NA	NA	NA	NA	NA	67.1	NA	NA	6.6	NA	NA	NA
NA	NA	NA	NA	3.7	NA	NA	NA	12.2	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	4	4.6	NA	20.3	NA	NA	NA	NA
NA	NA	NA	NA	8.8	5	NA	NA	NA	9.7	NA	NA
NA	NA	NA	3.1	NA	NA	NA	NA	34	NA	NA	NA
NA	NA	NA	NA	1.1	NA	NA	NA	17.3	NA	NA	NA
NA	NA	NA	NA	NA	6.2	NA	NA	NA	NA	NA	NA
NA	NA	NA	24.6	1.4	NA	25.1	16	49	2.2	NA	NA
NA	NA	NA	NA	NA	3.1	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	18.8	NA	NA	NA	0.8	NA	NA
NA	NA	NA	25.9	NA	8.7	NA	NA	NA	1.6	NA	NA
NA	NA	NA	0.5	NA	16.7	NA	10.2	NA	NA	NA	NA
NA	NA	NA	5.8	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	46.2	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	4.1	NA	NA	NA
NA	NA	NA	3.3	21.1	NA	NA	NA	NA	NA	2.3	NA
NA	NA	NA	38.1	NA	9.7	NA	NA	6.9	0.8	NA	NA
NA	NA	NA	NA	63.5	NA	NA	NA	NA	4.8	NA	NA
3.3	NA	NA	NA	NA	23.4	NA	NA	NA	NA	2.5	NA
12.7	NA	15	NA	NA	46.2	NA	1.6	NA	NA	0.1	NA
NA	NA	5.3	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	7.9	9.4	NA	NA	NA
NA		NA	NA	NA	NA	3.1	28.7	NA	0.1	NA	NA
NA		NA	NA	8.1	25.4	50.8	NA	NA	NA	NA	NA
NA		12.2		NA		NA	NA		NA		NA

Appendix H1

Figure 5.15: Sample Complete Data for Ejura (2000)

Complete Data for Ejura (year 2000) using EM Algorithm							
32.4477	26.5802	15.0000	0.83994	16.2587	31.5027	13.3000	37.4759
29.5468	11.0708	2.49285	9.84858	9.83168	16.7257	46.7346	14.1658
4.31509	11.9907	19.1000	20.4156	7.75072	11.2804	6.80591	24.7650
13.9548	33.2333	13.4972	21.1000	19.0789	42.7741	0.86401	2.3000
4.26867	42.9556	9.57395	34.4051	26.6881	0.00571	8.20891	1.30197
18.3889	7.07632	20.7234	63.5000	24.9831	9.29121	18.7346	22.6963
13.7449	16.9064	6.95094	5.7583	11.8347	35.2942	9.70000	2.50000
67.8557	12.6957	10.2855	9.91925	10.5429	1.60000	6.44416	0.10000
18.3732	8.99158	3.1000	24.0232	28.9177	5.43568	8.76048	30.9945
3.94297	6.68066	25.5734	18.4528	2.38943	7.90000	7.42815	23.8783
30.3297	38.7748	12.4323	18.5329	25.1000	28.7000	2.20000	1.99923
31.5370	11.2102	24.6000	8.10000	14.5002	17.6581	17.5652	24.9724
12.2422	24.2150	21.3249	18.3154	36.3817	29.0639	0.80000	14.4185
31.7148	23.1970	20.6877	15.5000	12.5536	1.52848	1.60000	34.8004
0.20016	5.01566	25.9000	7.44433	16.6634	10.1439	19.3663	2.0827
13.0099	39.3640	0.50000	20.9453	43.7956	8.34963	5.85138	9.31451
21.1101	18.4429	5.80000	6.08549	27.1137	39.4343	5.30485	21.5337
7.90833	19.9578	40.7919	15.3253	0.46220	32.5935	30.3943	30.3953
30.6412	1.81445	16.1705	67.1000	1.8548	6.60000	24.0269	13.4167
19.6418	18.9801	3.30000	29.3456	31.7850	12.2000	0.80000	36.4830
19.7547	39.3738	38.1000	25.6836	5.07263	13.7025	4.80000	1.26653
9.76379	5.98581	3.94992	4.60000	21.1237	46.1186	3.57690	47.1564
13.2423	17.2710	35.5417	5.0000	16.4754	14.0846	6.86175	4.89057
18.7084	13.4607	3.47029	16.9452	10.9985	34.0000	19.6500	27.9371
3.30000	23.4788	0.82056	8.39276	11.6729	17.3000	5.65970	4.32737
12.7000	41.9731	3.17568	6.20000	3.1000	26.7906	0.10000	4.31276
7.96708	23.4014	5.33818	18.4035	50.8000	49.0000	28.4737	25.6783
25.4679	10.1174	7.00367	3.10000	14.9414	38.6555	5.14490	6.99314
20.7230	8.11138	9.4000	18.8000	54.1000	29.1527	10.1179	33.1463
26.1782	21.0615	39.0722	8.70000	9.70000	43.6861	28.6865	16.2435
19.3499	2.96108	42.9689	16.7000	12.0817	40.0832	34.3192	4.18021
21.2535	14.7622	38.1328	37.4245	12.9484	11.7290	27.5366	16.3299
1.23257	46.2000	9.84618	7.4563	33.5000	21.6124	7.54083	53.0397
11.7185	44.2016	33.7986	8.26443	33.7523	4.1000	22.6829	3.72842
33.9068	44.4584	3.70000	1.63492	31.2097	24.1595	59.5120	23.9186
5.34265	0.51265	4.78908	9.70000	21.3786	6.90000	18.5743	3.71709
14.6000	15.0067	4.00000	9.51899	20.3000	47.7344	7.98290	21.7616
35.1433	10.4877	8.80000	23.4000	38.2629	5.51580	8.44492	18.8651
22.0159	15.0000	1.65177	46.2000	11.6265	34.4780	36.5595	10.6975
50.2646	5.30000	1.1000	57.2774	13.0336	31.9193	5.04987	34.7287
28.1057	14.0805	46.7983	21.6210	23.6831	9.40000	23.7165	39.4810
2.76289	44.1704	1.40000	23.7052	16.0000	1.88183	22.9530	1.42293
12.7010	2.7297	9.42632	25.400	11.0449	37.1568	19.4987	7.75342
2.18825	12.2000	8.19193	15.8154	3.91805	11.2225	10.0128	1.88211
38.3580	16.4376	15.0057	39.4705	8.82815	16.9029	6.5240	
3.43164	12.4608	15.7553	5.96116	10.2000	6.37512	9.16526	

Appendix I

Figure 5.16: Sample Complete Data for Ho (2000)

Observed data for Ho (2000)											
NA	NA	NA	38.6	NA	29.7	NA	NA	NA	NA	1.4	NA
NA	NA	NA	NA	NA	NA	10.8	NA	NA	NA	NA	NA
NA	NA	NA	NA	2.1	NA	NA	6.7	NA	3.9	6.7	NA
NA	NA	NA	3.4	NA	NA	34.9	9.7	NA	12.4	7.8	NA
NA	NA	NA	NA	NA	13.3	1.6	NA	NA	NA	NA	NA
NA	NA	NA	NA	5.8	8.9	NA	NA	56.6	NA	NA	NA
NA	NA	NA	19.7	NA	NA	NA	NA	0.4	NA	NA	NA
NA	NA	NA	NA	NA	NA	34.9	10.7	NA	NA	NA	NA
NA	NA	NA	NA	34	NA	NA	NA	37.4	0.3	1	NA
6	NA	NA	NA	NA	NA	NA	NA	0.9	NA	NA	NA
8.6	NA	NA	NA	NA	3.7	NA	NA	14.6	NA	NA	NA
3.3	NA	NA	NA	NA	2.6	3.4	NA	1.3	1.9	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	4.3	NA	NA	NA
0.2	NA	NA	NA	NA	NA	NA	10.7	NA	1.2	1.2	NA
NA	NA	NA	8	56	NA	NA	NA	9.5	NA	NA	NA
NA	NA	NA	NA	4.1	15.2	NA	NA	12.1	NA	NA	NA
NA	NA	NA	NA	2.1	NA	NA	1.8	27.1	20	NA	NA
NA	10.6	NA	NA	NA	17.6	NA	NA	0.8	NA	1.5	NA
30	NA	17.1	1.1	NA	12	NA	NA	NA	NA	NA	NA
NA	NA	NA	15.4	0.3	4.9	NA	NA	13.4	NA	NA	NA
NA	NA	5	NA	4.8	1.4	NA	2.6	NA	NA	NA	0.4
NA	NA	NA	46.2	NA	0.5	NA	0.8	NA	0.5	NA	0.7
NA	NA	NA	NA	58.9	NA	NA	0.2	NA	NA	NA	34.5
NA	NA	1.7	NA	15.2	NA	NA	NA	2.4	1.2	NA	4.6
NA	NA	NA	NA	NA	1.3	NA	1.9	1.3	NA	NA	NA
NA	NA	21.3	NA	4.1	NA	NA	2.2	NA	NA	NA	NA
NA	NA	NA	NA	NA	0.4	13.1	10.5	5.6	NA	NA	1.6
NA	NA	NA	48.5	NA	NA	28.8	23	1.1	63.4	NA	NA
NA	NA	NA	6.8	NA	4.8	13.3	NA	NA	NA	NA	NA
NA		NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA		0.5		7.4		37.2	3.6		0.3		NA

Appendix I1

Figure 5.17: Sample Complete Data for Ho (2000)

Complete Data for Ho (year 2000) using EM Algorithm							
8.8549	14.04210	11.35915	0.884055	17.422892	23.0000	3.90000	1.50000
27.3482	0.39224	7.21636	40.50423	19.33359	9.47226	12.4000	9.92389
24.6888	10.6000	3.400000	0.30000	12.10055	28.5443	7.38151	19.5733
6.35421	32.76656	10.51005	4.8000	33.78645	3.60000	13.7898	3.50939
10.3948	0.7476485	38.09514	6.24334	3.06659	26.2133	24.2462	5.63288
1.5150	21.969	19.7000	58.90000	12.50696	17.2005	3.7996	9.20811
14.4598	7.75095	4.90083	15.20000	6.61490	32.6794	0.30000	13.7050
10.20243	13.39088	12.67092	9.90832	9.50384	8.26028	6.90119	15.3559
59.80876	28.06848	9.025243	4.10000	15.0924	9.5503	10.0528	2.9659
6.0000	36.98143	0.1130	11.35827	22.0681	56.6000	1.90000	2.46492
8.6000	4.08893	9.975381	12.04548	20.50506	0.40000	33.7533	25.4657
3.3000	13.10074	6.37863	1.628304	8.451260	19.3132	1.20000	19.6284
14.44546	9.24053	16.59998	11.42706	7.26697	31.9108	27.4819	0.88082
0.20000	5.84475	8.00000	7.4000	24.112104	34.9000	10.1635	8.68886
6.01306	3.7262049	3.973976	29.7000	13.1000	1.60000	20.000	15.6159
25.40664	33.1486	7.030996	16.31775	28.8000	4.43728	39.8812	2.79022
26.51344	7.8787	21.04625	29.142765	13.30000	5.17813	15.3105	23.7051
8.82485	19.80090	1.10000	2.88064	4.58885	34.9000	22.1620	2.31828
30.00000	18.86766	15.40000	13.3000	37.20000	0.89950	33.0393	6.8773
26.67643	2.1997986	8.9991	8.90000	10.8948	11.1249	0.50000	23.9002
2.214828	33.68886	46.20000	6.69519	30.95477	50.1110	24.3275	29.0639
9.5286	14.5094	17.1514	19.62508	6.7000	3.4000	1.20000	22.8459
16.95445	15.8982	16.56727	14.51837	9.700000	37.4000	37.6511	9.31142
9.64832	4.06173	34.99785	14.592	9.110301	0.90000	34.3481	18.3963
25.69213	15.001849	12.42605	3.70000	12.87795	14.6000	8.35431	52.1595
15.6084	33.69778	6.01943	2.600000	37.7515	1.30000	63.4000	14.6297
15.71190	3.0892	48.50000	14.392398	10.7000	4.30000	17.4149	4.92004
6.55268	13.43499	6.80000	4.42630	22.4583	7.72717	19.7500	5.34359
14.53830	14.73854	30.18467	16.8034	2.822056	9.50000	0.30000	31.1177
14.75273	19.126028	0.78308	15.2000	4.09869	12.1000	1.40000	7.02782
4.9055377	36.08075	1.64607	3.180575	26.74077	27.1000	41.36247	19.3439
20.9495	19.0550	2.10000	17.600000	2.25203	0.80000	6.700000	18.6440
16.5996	17.1000	0.51299	12.0000	10.7000	5.99022	7.80000	15.4772
21.60067	6.87689	7.29212	4.9000	16.96697	13.4000	2.65831	0.40000
15.3409	5.000000	5.800000	1.4000	12.70560	5.69491	29.2096	0.70000
17.0859	5.03782	8.81897	0.500000	1.80000	26.4819	26.8638	34.5000
3.528294	16.90998	33.42130	16.447828	12.4813	12.9349	0.673146	4.6000
13.14136	1.70000	34.00000	24.50442	8.30290	2.40000	1.00000	6.78103
28.68588	5.112907	36.99365	1.30000	11.29929	1.3000	31.66539	8.3793
2.49957	21.30000	32.5601	21.14724	2.60000	7.94307	7.88991	1.60000
15.783014	15.9316	11.42485	0.4000	0.800000	5.60000	13.09757	6.00396
29.819	38.12402	28.58675	13.136236	0.20000	1.10000	3.446098	31.9579
17.78488	38.3591	1.99209	4.800000	8.67767	36.8150	1.20000	10.5882
43.68204	2.868305	56.0000	5.29577	1.90000	2.39309	40.44585	20.3051
23.36771	0.5000000	4.10000	14.4732	2.20000	6.11944	3.8410	
0.13457	38.6000	2.10000	10.8000	10.50000	29.95782	3.19041	

Appendix J

Figure 5.18: Sample Observed Data for Koforidua (2000)

Observed data for Koforidua (2000)											
NA	NA	NA	21.6	NA	NA	12.2	13.8	0.1	1.4	3.3	NA
NA	NA	NA	4.2	6.2	68.6	NA	NA	0.8	0.1	18.5	NA
NA	NA	NA	NA	NA	NA	10.2	0.5	NA	1.4	NA	NA
NA	NA	NA	NA	NA	NA	1.7	NA	NA	3.7	4.8	NA
NA	NA	NA	1.3	NA	0.3	NA	2.6	NA	65.5	7.3	NA
NA	NA	NA	NA	NA	35.1	3.8	NA	0.3	NA	NA	NA
NA	NA	NA	NA	NA	0.8	0.1	0.6	0.4	NA	NA	NA
NA	NA	NA	NA	37	NA	NA	NA	NA	13.5	9.4	NA
NA	NA	NA	NA	NA	NA	7	26.1	NA	0.4	NA	NA
NA	NA	NA	NA	1.5	NA	NA	NA	1.6	NA	2.5	NA
NA	NA	NA	NA	NA	NA	0.6	NA	2.2	NA	NA	0.8
NA	NA	NA	NA	NA	NA	NA	22.4	1.5	NA	11.4	NA
NA	NA	NA	NA	NA	30	NA	NA	5.5	5.1	NA	NA
4	NA	NA	NA	NA	NA	13.4	1	NA	0.5	NA	NA
NA	NA	NA	NA	NA	NA	0.5	NA	1	1.6	NA	NA
NA	NA	3.6	19.9	NA	NA	NA	NA	1.3	0.3	36.4	NA
NA	2.3	NA	NA	1	12.4	NA	NA	15.3	NA	NA	NA
NA	NA	NA	25.1	NA	6	NA	7.5	8.2	NA	NA	NA
NA	NA	NA	NA	NA	4.2	NA	0.2	0.2	17.1	NA	NA
NA	NA	7.5	NA	2.4	NA	NA	NA	0.5	NA	10.1	NA
NA	NA	NA	0.8	0.9	11.3	NA	NA	3.7	NA	NA	NA
NA	NA	0.8	NA	1	28.7	NA	NA	0.2	2.3	0.5	NA
NA	NA	9	27.1	NA	21.4	NA	77	3.6	6.4	0.5	NA
NA	NA	NA	NA	14.7	0.5	NA	0.4	0.5	11.3	NA	3
NA	2.4	NA	NA	NA	NA	NA	1.6	NA	1.1	NA	NA
NA	15	NA	NA	NA	1.4	NA	7.4	4.3	21.4	NA	NA
NA	NA	0.4	NA	19.7	0.4	NA	0.2	NA	1.2	NA	NA
NA	NA	NA	NA	NA	4.4	16.6	NA	6.9	0.6	1.8	1
NA		NA	NA	0.3	NA	4.3	NA	NA	43.9	NA	NA
NA		NA	NA	4.9	0.9	73.6	NA	NA	NA	NA	0.5
NA		NA		NA		NA	NA		NA		NA

Appendix J1

Figure 5.19: Sample Complete Data Koforidua (2000)

Complete Data for Koforidua (year 2000) using EM Algorithm							
11.21400	19.32784	4.710101	4.05355	1.70000	0.20000	3.70000	21.6645
24.65759	2.30000	7.97361	2.40000	2.69053	0.20875	65.50000	10.1000
22.02254	5.23979	1.30000	0.90000	3.80000	14.3713	3.20259	34.8660
8.73616	15.70836	35.30600	1.00000	0.100000	10.1490	23.7992	0.50000
7.85944	25.37120	7.29609	13.76253	3.33142	77.0000	13.5000	0.5000
0.93904	34.20250	10.11469	14.70000	7.0000	0.40000	0.40000	31.5933
11.88719	1.611310	6.50241	13.72811	13.4631	1.60000	10.3763	5.8376
7.66882	10.54057	2.55211	26.4356	0.600000	7.40000	5.43015	14.8152
56.82071	6.71572	7.4438	19.70000	47.21179	0.20000	34.0376	17.1290
11.87297	2.40000	3.88004	0.41409	14.82313	14.8071	5.10000	1.80000
8.39814	15.00000	14.00775	0.30000	13.40000	5.78668	0.50000	38.5433
22.73374	3.35105	1.497414	4.90000	0.50000	8.75562	1.60000	0.19380
23.83040	1.25191	4.52642	4.19370	16.7163	6.15802	0.30000	26.5018
4.00000	30.4048	19.90000	17.00514	9.549545	0.10000	4.8113	24.1776
6.30385	5.36636	18.41329	68.60000	31.03678	0.80000	3.62321	3.10714
23.99190	17.179354	25.10000	11.94521	0.5983	6.94533	17.1000	28.9352
4.63469	16.25466	6.47646	12.01799	9.95223	25.8426	27.2432	5.37747
7.00119	0.26051	14.5541	0.30000	4.114153	23.5330	9.75405	10.5374
14.35898	30.94009	0.8000	35.10000	11.85693	0.30000	2.30000	0.97437
12.00009	11.93628	13.97534	0.80000	12.51396	0.40000	6.40000	37.6351
23.01661	13.3124	27.1000	11.82039	19.42578	14.6028	11.3000	1.36567
13.02526	6.4647	32.23708	1.9456	17.8771	29.9399	1.10000	5.60134
13.12781	12.42423	9.872061	14.20929	5.93368	1.60000	21.4000	0.80000
4.05249	30.94893	8.40444	0.71128	4.76024	2.20000	1.2000	12.2731
16.84528	0.62074	27.467999	18.73731	16.60000	1.50000	0.60000	21.8342
12.17743	10.871763	1.66425	30.00000	4.30000	5.50000	43.9000	1.03708
2.42044	17.04368	4.07115	21.83978	73.60000	5.74444	11.2233	3.14112
18.31743	16.51066	1.93186	18.5133	21.45106	1.00000	21.5839	11.5639
14.0074	3.60000	9.6655	10.57574	13.80000	1.30000	3.30000	11.1393
18.9626	33.31007	6.20000	12.40000	6.98697	15.3000	18.5000	12.7750
12.76016	16.4403	11.17834	6.0000	0.500000	8.20000	6.20492	0.49862
14.48924	4.37374	30.67498	4.20000	8.35488	0.20000	4.80000	0.00218
5.93613	7.5000	34.21460	2.807103	2.60000	0.50000	7.30000	22.7923
15.46114	2.551509	29.82166	11.30000	28.23105	3.70000	4.3978	17.0085
25.98294	0.80000	13.76035	28.7000	0.60000	0.20000	12.4009	1.56741
0.03651	9.000000	37.0000	21.40000	6.58668	3.60000	9.40000	3.00000
18.07859	14.314916	25.8847	0.500000	26.10000	0.50000	31.0039	11.0494
27.10617	7.50623	1.50000	11.9005	10.31982	7.02268	2.50000	13.0327
15.1818	18.2258	0.4663	1.40000	34.96550	4.30000	24.7900	0.32451
40.8417	0.400000	3.31611	0.400000	22.40000	16.6962	11.4000	1.00000
20.71349	35.3346	37.69303	4.4000	19.81243	6.90000	7.63028	21.0478
2.30682	35.56755	3.745989	29.17834	1.00000	5.21622	37.0756	0.50000
16.35362	5.2822	12.25771	0.90000	5.23636	8.37550	17.6104	0.14311
2.82880	8.814935	8.81407	12.200000	6.50130	1.40000	36.4000	4.37414
30.02624	21.6000	1.00000	1.95647	24.0556	0.10000	19.5188	
1.699359	4.200000	9.49498	10.20000	7.50000	1.40000	30.2964	

Appendix K

Figure 5.20: Sample Observed Data for Kete Krachi (2000)

Observed data for Krachi (2000)											
NA	NA	NA	NA	13.4	0.9	1.1	18.4	17.8	20.3	4.8	NA
0.7	NA	NA	NA	NA	NA	NA	25.2	NA	0.8	NA	NA
NA	NA	NA	NA	NA	NA	21	NA	NA	NA	NA	NA
NA	NA	NA	NA	2.6	0.1	NA	5.4	NA	NA	NA	NA
NA	NA	NA	5.7	NA	NA	2.2	3.6	0.8	NA	NA	NA
NA	NA	NA	NA	NA	21.6	9.2	2.9	47	NA	NA	NA
NA	NA	NA	NA	NA	NA	60	3.9	4	17	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	0.6	15	NA	NA
NA	NA	NA	2.2	NA	5.6	0.4	60.8	1.1	42.7	NA	NA
NA	NA	NA	NA	NA	NA	NA	2.6	23	2.2	NA	NA
NA	NA	NA	NA	NA	NA	0.4	NA	6.4	NA	NA	NA
NA	NA	NA	24.2	NA	NA	3.3	NA	9.4	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	10.7	2.2	NA	NA	NA
3.8	NA	NA	NA	NA	NA	37.4	1.2	1.2	NA	NA	NA
NA	NA	NA	0.7	NA	NA	4.2	9.1	NA	31.6	NA	NA
NA	NA	NA	NA	NA	5	NA	NA	46.2	NA	NA	NA
NA	NA	NA	NA	NA	22.6	NA	0.5	0.3	NA	NA	NA
NA	NA	NA	1.6	NA	NA	NA	0.8	29.5	NA	NA	NA
NA	NA	NA	20.5	NA	44.9	NA	4.5	12.9	NA	NA	NA
NA	NA	53.5	NA	NA	4.5	NA	NA	3.7	NA	NA	NA
NA	NA	8.9	10.8	NA	0.7	NA	NA	20.2	NA	NA	NA
NA	NA	NA	3.2	NA	NA	NA	NA	NA	0.6	NA	NA
NA	NA	NA	0.8	12.8	1.2	NA	65.1	2.7	NA	NA	NA
11.7	NA	NA	NA	45.2	NA	NA	9.1	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	14.6	12.3	4.2	NA	NA
NA	NA	5.5	2.4	NA	18.8	NA	0.7	NA	NA	NA	NA
NA	NA	16.6	NA	3.7	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	7.6	7.5	NA	NA	NA
NA		NA	NA	NA	NA	36.9	13.1	NA	NA	NA	NA
NA		NA	NA	1.1	66.8	0.3	NA	NA	NA	NA	NA
NA		NA		NA		5.9	NA		NA		NA

Appendix K1

Figure 5.21: Sample Complete Data for Kete Krachi (2000)

Complete Data for Krachi (year 2000) using EM Algorithm							
] 9.6471	36.38322	14.15895	12.492	3.53728	4.50000	14.45099	21.50081
0.70000	0.97271	10.12710	18.1921	2.200000	14.1973	8.93030	4.026983
30.39101	24.4422	5.70000	32.3756	9.200000	13.6575	40.86057	6.37541
27.44989	8.7178	0.020902	3.33163	60.00000	9.328250	17.0000	10.03763
6.88142	14.66345	11.17788	12.80000	13.9776	65.1000	15.0000	15.30259
11.64175	31.18749	7.20015	45.200000	0.4000	9.100000	42.70000	17.12828
1.82134	41.04455	2.20000	7.550244	7.46145	14.60000	2.20000	3.42598
16.13733	4.66791	18.5042	21.84973	0.40000	0.70000	2.50072	2.8718
11.42899	14.63430	4.54078	3.70000	3.30000	12.642027	6.91349	28.3091
66.28986	10.3652	24.20000	16.20208	10.364669	7.6000	33.27696	21.85349
16.1215	6.60971	7.921614	16.2833	37.4000	13.1000	8.01754	1.11998
6.50413	4.26676	23.42144	1.1000	4.200000	9.74272	31.6000	9.463367
28.24369	36.80577	0.70000	16.06277	16.836889	10.62147	15.39634	17.41589
3.80000	8.85910	10.09814	0.90000	24.55152	17.80000	26.96034	3.231647
29.46774	22.04417	19.11402	5.04102	22.82291	31.71372	4.05617	26.36197
9.90549	21.01208	1.60000	18.72914	9.4923	29.1359	7.77806	2.709709
29.64799	2.57867	20.5000	0.100000	8.18258	19.16842	10.97188	7.75165
2.30357	37.40321	18.4680	3.66334	26.81203	0.80000	37.47449	26.5777
10.68382	16.19212	10.8000	21.600000	4.92906	47.00000	0.6000	32.2884
18.89621	17.72803	3.20000	18.04419	12.19474	4.00000	30.53882	25.41181
10.5244	4.34611	0.800000	27.2459	34.3795	0.60000	11.38598	10.15187
28.5594	16.73678	38.850849	5.6000	10.22117	1.10000	4.2000	20.49086
17.40758	37.41308	13.8881	23.53312	14.38790	23.00000	44.25142	57.83043
11.70000	3.56227	2.400000	14.67355	41.89617	6.40000	16.78635	16.32524
17.52205	15.00395	6.51117	6.00259	24.98309	9.400000	24.65545	5.58706
7.39263	16.15385	33.52784	16.15214	36.9000	2.20000	36.68486	6.05547
15.93241	21.29781	1.01188	35.43684	0.30000	1.20000	27.05032	34.5597
16.46128	40.04846	1.6746	5.05316	5.90000	36.28687	41.78516	7.62637
5.57102	21.21928	13.4000	5.000000	18.40000	46.20000	4.8000	21.5388
23.31444	7.75119	0.71319	22.60000	25.20000	0.30000	38.13233	20.76475
18.50379	5.71731	7.918678	5.87248	2.97512	29.5000	9.385110	17.26251
24.03458	18.84703	2.600000	44.900	5.400000	12.90000	19.4055	7.645176
17.11169	53.5000	9.60725	4.50000	3.60000	3.700000	21.9880	9.12099
19.04160	8.90000	37.10731	0.70000	2.900000	20.20000	45.8896	6.785799
3.75617	5.508628	41.05806	0.84892	3.900000	9.28111	3.08576	35.4889
14.3875	17.47328	36.15488	1.200000	4.3869	2.70000	32.44947	11.85562
31.87029	42.30816	12.4892	12.15744	60.8000	10.70781	29.85525	22.6018
2.9102	42.5681	31.761	55.56487	2.60000	12.30000	0.598587	3.8265
17.3089	5.50000	2.34896	18.80000	29.71914	21.50485	35.1654	20.50441
33.12398	16.60000	0.83184	19.41427	2.63648	8.69153	8.871509	28.91765
19.81459	3.02627	44.940502	21.5274	10.70000	7.50000	14.63079	21.70282
48.45492	12.70823	7.050527	13.528	1.200000	6.47887	3.95699	0.47956
25.98879	8.12662	10.81199	66.8000	9.10000	6.444010	44.8759	22.81206
0.29469	11.7692	12.70726	1.10000	18.91005	20.3000	4.39374	12.11188
15.38364	42.27622	13.46726	37.51113	0.500000	0.80000	3.38249	
0.28792	5.27409	1.65492	21.00000	0.8000	29.43288	10.8292	

Appendix L

Figure 5.22: Sample Observed Data for Saltpond (2000)

Observed data for Saltpond (2000)											
8.6	NA	NA	NA	2.3	7.2	56.2	0.3	NA	NA	NA	NA
NA	NA	NA	NA	NA	26.8	NA	1	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	0.7	1.9	0.5	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	0.8	NA	NA	NA	NA
NA	NA	NA	6.8	NA	NA	1	NA	NA	2.3	9.9	NA
NA	NA	NA	NA	NA	NA	0.2	NA	0.8	NA	NA	NA
NA	NA	NA	NA	NA	0.7	NA	1.3	NA	NA	NA	NA
NA	NA	NA	NA	1.8	NA	NA	NA	NA	1.6	0.3	NA
NA	NA	NA	NA	NA	NA	0.3	NA	NA	NA	1.9	NA
NA	NA	NA	NA	40.1	NA	2	0.4	NA	NA	NA	NA
NA	NA	NA	61.6	NA	NA	NA	NA	NA	NA	NA	11.2
NA	NA	NA	NA	NA	14.1	NA	NA	0.7	NA	NA	NA
NA	NA	NA	NA	NA	5.7	NA	NA	1.2	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
2.3	NA	NA	NA	NA	6.8	2.2	NA	NA	NA	3.8	NA
NA	NA	1.6	16	NA	40.2	NA	NA	NA	NA	NA	NA
NA	NA	NA	0.4	NA	0.4	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	5.2	1.2	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	2.1	1.4	NA	NA	0.3	NA	40.9	NA
NA	NA	NA	NA	3.1	NA	NA	NA	NA	NA	NA	NA
NA	NA	67.4	NA	5.9	10	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	11	1.7	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	2.2	12.6	NA	NA	2.2	NA	NA	NA
NA	NA	NA	NA	41.3	2.1	NA	4	NA	NA	NA	NA
NA	NA	NA	NA	1.9	3	NA	NA	NA	NA	NA	NA
3.4	NA	24.6	NA	NA	2.1	NA	NA	1.3	NA	NA	NA
NA	NA	NA	NA	28.9	NA	NA	NA	NA	NA	NA	NA
NA	NA	28.7	NA	1.1	NA	NA	NA	NA	7.6	NA	NA
NA		1.8	1.7	NA	NA	NA	NA	NA	22.4	NA	NA
NA		NA	NA	NA	NA	25.1	NA	NA	NA	NA	NA
NA		NA		NA		NA	NA		NA		0.9

Appendix L1

Figure 5.23: Sample Complete Data for Saltpond (2000)

Complete Data for Saltpond (year 2000) using EM Algorithm							
8.60000	1.62382	10.60428	2.100000	13.79991	26.8835	27.56578	40.900
25.65689	20.14034	6.86540	3.100000	1.00000	24.49298	2.30000	9.598075
22.92948	5.55853	6.80000	5.90000	0.200000	15.24976	25.16006	17.44787
8.90725	16.12382	2.506464	11.00000	49.00158	31.12435	3.0809	16.7301
8.26999	26.3955	7.83983	2.200000	15.47775	6.080878	1.60000	13.482333
0.83685	35.53632	4.151128	41.3000	0.30000	4.00000	30.08437	4.56381
12.4389	1.80288	14.63379	1.900000	2.00000	7.40391	5.70104	10.98409
8.07269	11.045098	1.68499	0.563704	17.43729	17.41642	11.04184	3.7668
58.9473	7.086194	61.6000	28.9000	10.01934	5.5341	1.143617	30.3843
12.42419	3.60359	4.82017	1.1000	32.25966	8.53395	39.08931	8.46832
8.5574	1.43089	19.19374	4.47578	0.75441	3.44993	1.54863	18.4337
23.66561	31.60555	6.83854	17.73624	2.20000	24.76838	5.66256	6.07433
24.800708	5.6895	15.19932	12.4989	10.43614	10.87511	12.5681	16.48867
6.65989	17.91655	16.00000	7.20000	4.393443	5.75556	22.4643	24.29059
2.30000	16.9595	0.40000	26.80000	12.1374	35.36571	1.20853	17.600004
24.96786	0.13455	14.60025	12.57429	13.08764	0.50000	3.38631	2.97055
4.66204	32.15958	33.50203	12.36977	20.2417	4.84485	11.8341	18.6286
7.38167	12.48972	10.3531	2.14888	18.63871	3.885302	11.66483	8.70596
14.9973	13.91403	8.56391	14.84240	6.27674	0.80000	13.35786	27.8688
12.28558	6.55616	28.5658	0.70000	5.062175	28.3331	0.651196	2.89208
23.9584	12.9948	1.58749	0.87131	22.33798	9.9608	0.137356	3.90721
13.61687	32.16873	4.07874	19.25892	7.09675	11.75177	23.72627	15.39565
13.72301	0.7776	1.864476	22.7403	8.78280	22.47551	17.7397	11.2000
4.32962	11.38789	9.86914	19.2973	29.35559	6.28729	1.4872	23.72721
17.30058	17.5059	11.43502	14.10000	6.952633	0.70000	7.60000	7.764078
3.4000	17.22443	31.88518	5.70000	10.8166	1.20000	22.4000	29.45085
12.73933	34.61262	1.70000	11.08149	25.1000	4.68704	11.30159	3.65944
2.64036	17.15159	35.5489	6.80000	36.32607	12.70049	13.62458	39.48590
19.0945	1.60000	2.30000	40.2000	0.30000	32.22568	0.47098	0.25215
14.63343	4.66213	31.00195	0.40000	1.000000	25.7939	21.92061	21.41600
19.7623	2.77603	14.1075	1.20000	1.900000	8.03280	0.013031	8.91878
13.34248	14.9517	26.92702	1.40000	0.80000	0.30000	4.66255	0.7954
15.13216	7.6342	0.34756	3.04058	20.64192	38.51018	9.90000	19.2923
6.00908	67.4000	3.29724	10.0000	5.28479	18.09247	22.1206	1.724666
15.86792	18.7295	39.14919	1.70000	1.30000	0.338091	27.4165	26.313757
27.0287	36.70813	1.80000	12.60000	6.59406	2.20000	0.30000	10.42183
0.172892	36.9492	4.012375	2.100000	25.0338	31.49342	1.90000	0.9200
18.5771	5.332222	40.1000	3.000000	0.400000	22.55895	21.03949	10.5030
28.19129	24.60000	12.5522	2.10000	0.080979	1.30000	11.94006	45.0174
15.84898	9.25898	9.25808	12.45265	15.01017	36.2231	16.47611	1.34479
42.4082	28.7000	0.962859	30.3361	10.63985	32.8357	51.10252	17.63790
21.57455	1.800000	4.06052	2.16014	15.19099	6.17732	12.61317	1.35545
2.25257	5.01028	14.10979	2.91993	6.12459	15.4696	3.800000	15.6099
16.79169	8.38817	14.34435	56.2000	9.19759	17.864436	2.65524	0.90000
2.792848	36.67851	27.4973	3.31309	6.50895	40.02937	3.08962	
31.2137	7.41671	5.2000	0.70000	7.32385	0.33569	29.52268	

Appendix M

Figure 5.24: Sample Observed Data for Sefwi-Bekwai (2000)

Observed data for Sefwi (2000)											
NA	NA	NA	1.7	NA	11.2	43	8	10	NA	NA	NA
NA	NA	NA	4.7	NA	23	NA	NA	2	NA	7.9	NA
NA	NA	NA	NA	NA	NA	NA	1.1	0.9	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	0.2	21.3	NA
NA	NA	NA	60.1	28.9	NA	1.2	0.7	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	5	NA	3.5	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	16.8	0.9	NA	NA	1.7
NA	NA	NA	NA	NA	NA	NA	0.2	NA	0.7	NA	NA
NA	NA	NA	NA	NA	NA	0.3	20.5	2.2	NA	NA	NA
NA	NA	NA	NA	43	NA	2.9	NA	NA	NA	6	NA
NA	NA	NA	0.4	NA	2.9	NA	NA	9.7	NA	NA	NA
NA	NA	NA	NA	NA	0.6	0.2	NA	15.1	NA	NA	NA
NA	NA	NA	NA	5.2	16.6	NA	NA	0.6	NA	NA	NA
3	NA	NA	NA	NA	20	0.5	10	NA	4.8	0.5	NA
NA	NA	NA	NA	NA	4.8	29.8	0.4	NA	NA	NA	NA
NA	NA	NA	NA	NA	12.4	16.9	NA	6.5	0.7	NA	NA
NA	NA	0.3	NA	1.5	16.7	NA	NA	1.4	NA	13.5	NA
NA	NA	29	5.2	NA	19.3	NA	2	1.4	NA	NA	NA
4.8	NA	11.3	NA	NA	NA	NA	NA	13.1	NA	24.8	NA
NA	NA	NA	NA	6.1	28.2	NA	NA	1	NA	NA	NA
NA	NA	NA	NA	26.1	7.7	NA	NA	61.3	NA	3.4	NA
0.5	NA	NA	19.1	40	36.8	NA	2.4	NA	NA	0.7	NA
NA	NA	NA	0.5	NA	12.5	NA	35.1	2	1	NA	NA
NA	19.6	NA	NA	41.6	3.1	0.2	2.7	9.1	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	2.3	NA	NA
10.9	8.9	NA	30.2	NA	0.5	NA	NA	1.4	0.6	NA	NA
NA	NA	NA	41.6	11	15	NA	0.4	NA	NA	NA	NA
NA	NA	0.7	1.7	8	2.6	NA	1	NA	14.1	NA	NA
NA		93.6	36	NA	NA	0.5	9	3.8	NA	NA	NA
NA		NA	1	NA	NA	17.7	NA	5.6	NA	NA	NA
NA		5.2		NA		0.5	NA		NA		NA

Appendix M1

Figure 5.25: Sample Complete Data for Sefwi-Bekwai (2000)

Complete Data for Sefwi (year 2000) using EM Algorithm							
3.2282	26.5817	39.04315	38.728049	28.7936	8.339	0.4587	48.90617
17.06288	12.7837	2.209638	6.1000	1.2000	[232]	0.200000	24.8000
11.2994	5.46714	60.1000	26.1000	5.00000	14.270	4.7390	40.2396
8.3364	10.43597	20.628	40.000	12.739	1.12195	11.0093	3.4000
19.1432	36.2945	7.24273	43.144	36.1149	2.4000	7.1294	0.7000
10.6856	12.2429	9.4703	41.600	0.3000	35.10000	0.7000	26.16673
22.2519	2.50769	30.598	55.522	2.9000	2.7000	24.0378	11.1217
42.8163	6.9413	2.8312	32.79575	14.33075	10.34348	4.258640	49.0480
32.1395	19.600	0.40000	11.0000	0.200000	20.8408	18.6842	39.524
40.5033	0.23524	28.2919	8.00000	9.37086	0.4000	21.9277	4.4461
0.6661	8.90000	35.62470	29.10989	0.5000	1.0000	25.032	8.61864
8.9767	33.1713	34.04070	9.924171	29.8000	9.000000	4.8000	1.76240
3.543529	10.0446	13.40879	12.1798	16.900	0.232	21.89776	14.8815
3.00000	15.8461	10.92568	11.200	21.9934	0.31055	0.70000	5.75302
15.5544	13.08684	24.0061	23.0000	30.3351	10.0000	24.8108	10.478
1.6626	20.9005	5.2000	8.570	42.8753	2.0000	15.0438	0.6914
31.4923	33.2139	13.276	6.3849	0.7168	0.90000	18.1436	25.72159
25.4019	14.8885	22.673	23.2798	20.5605	3.59595	29.1555	1.4395
4.8000	4.1292	19.011	18.2289	13.0752	28.36316	35.2344	5.3435
35.62927	18.989	19.1000	13.94718	34.86075	3.50000	5.8482	1.70000
22.7922	38.96389	0.50000	5.34237	0.200000	0.9000	1.00000	15.3129
0.5000	29.1911	38.8952	6.78196	21.6113	18.9135	21.6377	31.8339
8.1511	1.8825	18.10673	5.14440	18.5646	2.20000	2.3000	18.5616
35.55266	9.68118	30.20000	2.9000	0.2265	18.1303	0.60000	5.57767
13.3758	41.2769	41.60000	0.60000	17.980	9.70000	10.84852	3.33403
10.9000	0.84359	1.70000	16.600	0.50000	15.1000	14.10000	9.4811
17.1143	7.0994	36.000	20.000	17.7000	0.6000	8.3502	17.5203
14.8238	0.44704	1.0000	4.8000	0.5000	2.0659	9.38895	20.01174
20.7230	7.07500	23.411	12.4000	8.0000	31.7689	26.4124	3.6174
13.5996	0.3000	3.3322	16.7000	14.63743	6.50000	34.6128	10.410
0.91102	29.000	7.41541	19.30000	1.10000	1.40000	7.9000	5.6296
0.7169	11.3000	22.1716	16.29054	8.51063	1.4000	21.7252	28.888
5.87779	15.72890	28.9000	28.2000	0.70000	13.1000	21.3000	45.79124
28.9254	16.546	28.0625	7.70000	10.22602	1.00000	10.596	20.28199
17.0503	4.55796	13.34622	36.8000	16.8000	61.300	2.96797	8.66734
36.7212	10.69322	12.97786	12.5000	0.200	14.7085	3.17264	22.4721
17.8255	18.65508	5.94780	3.1000	20.5000	2.00000	5.08167	18.1512
30.6981	10.658	43.0000	25.0958	7.9430	9.1000	22.1582	22.456
9.86591	15.6527	24.345	0.5000	24.86560	28.1216	6.0000	4.18406
13.9895	10.7039	3.279	15.0000	24.2838	1.4000	24.0838	14.0871
8.1756	0.7000	5.2000	2.6000	10.7903	6.7073	5.86620	10.914
1.44168	93.60	19.7687	33.8269	10.0000	20.39633	20.488	30.909
6.3559	18.5529	5.8218	12.406956	0.4000	3.8000	0.5000	29.1869
18.7646	5.20000	21.4769	43.0000	11.5438	5.6000	22.74743	19.7510
2.32003	1.70000	1.50000	32.5046	0.987219	1.26964	6.1118	36.3373
3.75370	4.70000	24.66720	16.0114	2.00000	20.6159	13.5000	
24.8000	21.3000	7.9000	21.7252	40.2396	3.4000	14.270	

Appendix N

Figure 5.26: Sample Observed Data for Tamale (2000)

Observed data for Tamale (2000)											
18.3	NA	NA	NA	NA	13.8	2.6	NA	NA	25	NA	NA
NA	NA	NA	NA	NA	0.2	NA	67.4	NA	0.2	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	12.1	NA	NA	15.1	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	0.6	NA	NA	NA	NA
NA	NA	NA	NA	NA	23.9	16.4	NA	10.6	NA	NA	NA
NA	NA	NA	NA	NA	4.4	7.1	NA	9.6	5.3	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	27.8	0.5	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	32.9	NA	NA	NA	0.2	NA	NA
NA	NA	NA	10	NA	0.7	NA	NA	0.5	24.3	NA	NA
NA	NA	NA	NA	11.9	1.6	NA	NA	9.4	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
NA	NA	NA	10.6	0.8	NA	6.7	NA	NA	21	NA	NA
NA	NA	NA	NA	43.5	NA	NA	7.9	14.8	0.3	NA	NA
NA	NA	NA	NA	0.3	NA	NA	NA	32.5	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	21.8	NA	NA	NA
NA	NA	NA	NA	NA	5.1	NA	NA	14.4	10.9	NA	NA
NA	NA	NA	1.3	NA	11.5	NA	4.8	14.2	NA	NA	NA
41.7	NA	37.4	NA	NA	NA	NA	0.6	19.1	NA	NA	NA
NA	NA	NA	NA	NA	14.5	NA	1.4	0.5	NA	NA	NA
NA	NA	NA	13.8	3.5	55.3	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	2.4	6.5	1.6	NA	NA
NA	NA	NA	NA	15.7	59.6	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	0.3	NA	NA	NA	5.4	NA	NA	NA
NA	NA	NA	NA	NA	12.5	NA	5.3	1.1	NA	NA	NA
NA	NA	1.1	NA	NA	NA	7.8	NA	NA	NA	NA	NA
NA	NA	NA	NA	NA	0.3	2	0.4	26.9	NA	NA	NA
NA		NA	10.3	NA	NA	11.3	NA	10.4	NA	NA	NA
NA		NA	NA	NA	13.2	NA	NA	5.7	NA	NA	NA
NA		NA		45.5		2.8	13.2		NA		NA

Appendix N1

Figure 5.27: Sample Complete Data for Tamale (2000)

Complete Data for Tamale (year 2000) using EM Algorithm							
24.60220	20.112258	1.679918	26.90000	9.300000	8.704778	6.882208	5.911906
22.38236	8.24406	10.10082	25.6000	10.2000	11.2059	8.2000	13.21400
3.52971	9.40331	7.09856	18.15555	6.158747	9.01760	26.78043	14.59196
10.45092	25.20337	15.63046	13.89291	14.3720	9.68085	4.38721	4.24995
3.03882	32.64313	5.09136	9.40000	8.80000	12.40000	13.28476	3.831735
13.8440	5.18731	7.64309	13.954216	5.20000	4.40000	22.0129	23.0308
10.29034	12.70959	19.34182	13.7878	19.10000	22.6000	1.39732	18.1583
51.69740	9.487428	9.28586	5.46893	4.60000	55.10000	7.534748	2.50946
13.83205	6.65292	16.0907	25.10000	20.19477	27.00000	6.617048	5.47848
3.24495	4.88455	1.00000	15.80024	18.89007	39.40000	29.9486	14.80905
22.98149	0.80000	15.60315	8.100000	10.0000	2.400000	24.71377	4.103275
23.90535	29.443848	30.987401	0.50000	8.82861	4.60000	0.8000	21.56123
9.14045	8.35068	12.14642	4.429101	7.840069	9.00000	10.25788	3.70934
24.04141	18.30231	3.25027	0.40000	0.70000	25.60055	35.06356	7.514814
0.07451	20.80000	26.9698	18.70000	21.9009	23.65489	11.0056	21.72403
9.72792	17.52332	0.10000	0.100000	2.05614	2.70000	20.27321	26.03432
15.92634	3.61043	2.42788	11.95499	28.80000	20.7000	29.35259	20.8441
6.27935	29.89478	0.40024	22.2284	26.2000	0.30000	22.08078	5.9981
23.21979	13.88539	2.5000	19.42611	5.30000	11.30000	33.20211	17.12993
14.40000	15.0446	2.202436	12.73922	10.86830	4.000000	7.2000	45.31253
0.10000	1.61614	40.40000	3.90000	27.61259	16.13179	30.44509	13.98586
14.80278	14.29647	4.312599	9.00000	9.37872	4.60000	8.74769	5.88105
14.88917	29.90222	5.58708	6.19468	0.700000	19.2000	16.31074	6.23460
7.24384	4.35282	29.67144	13.85521	12.52362	15.500000	1.90000	27.74859
10.36108	12.98860	18.8000	23.40000	5.6000	6.600000	18.25989	4.091977
14.08854	10.52822	32.65332	5.50000	16.30000	29.0521	36.30004	17.92084
5.86895	17.73898	28.95258	28.41062	33.28591	8.66919	12.40000	17.3366
19.26106	31.8913	7.7622	7.50000	2.20000	11.1000	2.80000	14.693
15.63016	17.67971	25.6359	5.478089	3.00000	25.0000	3.993163	7.43445
19.80460	7.51447	3.437057	3.80000	20.520499	9.7460193	26.15586	5.220065
14.57945	5.97937	1.03630	6.09648	3.100000	17.89525	4.20000	6.785826
16.03608	15.88922	35.583658	1.02340	28.4000	8.2242	38.5000	28.44992
1.17088	2.49358	6.98563	7.51187	3.70000	2.80000	24.19784	10.6123
9.19504	11.5241	22.30000	43.60256	0.58137	0.600000	1.21235	18.72322
25.71872	33.59686	6.49637	53.40000	1.64699	3.22588	5.40000	1.223988
3.86066	33.79309	11.25514	16.31735	24.09510	6.52785	28.20575	17.14014
11.40006	0.61998	11.82876	17.91224	5.200000	8.00000	8.360046	23.4902
26.66496	11.25587	0.4151	10.0000	3.654033	0.20000	12.70694	18.044
16.6195	7.797830	7.76408	11.874726	24.6000	5.10000	4.65073	1.30218
38.23622	10.54711	5.70000	29.9762	4.80000	23.87904	35.53492	18.881886
21.27957	33.5728	15.39488	4.33396	15.936788	12.57124	4.98038	10.80576
1.88656	2.31656	26.100087	11.60000	5.20000	8.40442	0.8888	26.40249
9.94689	12.35081	6.50000	12.21396	12.3797	11.40000	6.50936	1.36606
1.446829	13.3000	4.17874	0.10000	12.20000	32.50427	14.56392	6.900039
29.12492	9.30772	7.36280	7.295785	8.64408	0.60000	4.70356	
2.3983	4.400000	2.70000	53.3000	0.700000	0.22331	6.47607	

Appendix O

Figure 5.28: Sample Observed Data for Wa (2000)

Observed data for Wa (2000)											
28.8	NA	NA	NA	NA	3	47.8	NA	25.6	4.9	NA	NA
1.8	NA	NA	NA	1.7	NA	NA	13.5	0.8	5.1	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	11.1	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	0.1	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	14	1.8	NA	NA	NA	NA
NA	NA	NA	NA	NA	10.2	0.3	1.2	NA	0.3	NA	NA
NA	NA	NA	4.4	4.3	NA	30.5	NA	28.1	NA	NA	NA
NA	NA	NA	15.9	NA	NA	NA	1.2	50.9	NA	NA	NA
NA	NA	NA	NA	2.4	NA	NA	0.8	3.1	NA	NA	NA
NA	NA	NA	0.5	NA	33.6	NA	2.5	NA	3.2	NA	NA
NA	NA	NA	0.3	NA	NA	NA	NA	NA	17.9	NA	NA
NA	NA	NA	NA	NA	12.9	NA	2.1	41.8	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	5.2	NA	NA	NA
NA	NA	NA	0.9	9.4	NA	35.6	43.6	NA	8.4	NA	NA
NA	NA	NA	NA	14.3	21.3	1.6	18.1	2.2	NA	NA	NA
NA	NA	NA	16.3	NA	NA	NA	11.2	NA	NA	NA	NA
NA	NA	NA	1.8	NA	12	NA	22.4	5.3	NA	NA	NA
NA	NA	NA	NA	NA	0.7	4.8	0.1	24.7	22.4	NA	NA
NA	NA	NA	NA	NA	31.2	NA	NA	3.7	NA	NA	NA
36.9	NA	NA	NA	1.1	NA	NA	NA	8.3	8.9	NA	NA
NA	NA	NA	NA	NA	20	NA	12.7	1.9	NA	NA	NA
NA	NA	NA	29.2	11.5	18.7	2.5	NA	0.6	6.9	NA	NA
NA	NA	NA	NA	6.1	0.2	0.5	NA	16.3	NA	NA	NA
NA	NA	NA	NA	16.1	NA	NA	0.6	0.6	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	11.5	NA	NA	NA
NA	NA	NA	NA	NA	34.5	NA	28.1	0.2	NA	NA	NA
NA	NA	1.2	NA	NA	NA	NA	1.7	NA	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	9	5.7	NA	NA	NA
NA		NA	6.9	NA	NA	30.8	14.9	NA	NA	NA	NA
NA		NA	NA	NA	0.7	21.2	2.6	2.4	NA	NA	NA
NA		1.4		15.2		7.8	NA		0.2		NA

Appendix O1

Figure 5.29: Sample Complete Data for Wa (2000)

Complete Data for Wa (year 2000) using EM Algorithm							
28.80000	8.51382	14.5808	12.99063	8.1796	3.28700	8.69411	11.51549
1.80000	19.37192	23.12171	1.100000	14.000000	20.42186	6.06378	24.6303
34.34007	0.57417	0.46954	16.7962	0.30000	12.7000	0.300000	11.2349
11.83071	5.44248	13.04908	11.5000	30.5000	23.16771	16.171074	3.30696
11.95580	18.4478	4.40000	6.10000	20.58360	8.03194	12.7677	6.76011
4.23744	9.181449	15.90000	16.1000	7.043841	0.60000	29.06162	4.95675
6.42884	1.02267	21.88184	15.7271	12.6929	32.2851	3.2000	7.658768
0.20364	0.104258	0.50000	32.232209	11.3365	28.10000	17.90000	7.982144
3.66152	3.74476	0.30000	15.6742	13.24789	1.700000	7.08793	8.15359
9.301984	11.549402	18.02270	34.75428	1.72109	9.000000	13.47983	14.24720
18.8657	6.93440	9.32478	19.8666	35.600	14.9000	8.40000	0.59291
23.32532	30.16998	0.90000	2.867482	1.6000000	2.60000	7.16267	15.92289
5.10263	5.58599	6.89532	15.2000	14.57068	5.30074	16.70373	4.6380
0.59925	3.24096	16.30000	3.00000	5.67466	25.60000	5.515483	13.61729
27.83508	15.3824	1.8000	0.84118	4.80000	0.80000	22.4000	7.65897
29.25994	13.39609	11.32099	16.5722	10.14176	11.1000	14.66536	0.3332
24.68515	0.89885	3.906756	15.3558	9.61632	5.84138	8.900	15.617703
42.46698	0.060906	26.61813	12.4247435	11.21512	25.32979	21.91498	11.59797
20.8415	30.2350	27.4337	10.2000	2.5000	16.65015	6.90000	21.72621
36.90000	12.78039	29.2000	24.36054	0.500000	28.100000	1.843863	7.956916
27.2157	8.31481	5.54652	14.2053	0.25154	50.9000	6.81827	23.4317
3.35109	9.06647	11.2147	13.43423	8.18447	3.10000	5.60644	10.06742
16.51537	28.87062	17.8188	33.6000	2.763565	3.82209	11.261363	7.02861
20.35737	11.42586	2.674409	27.4933	19.52413	12.98048	11.41765	11.5432
10.40915	37.6292	25.3311	12.9000	25.00312	41.8000	14.3679	19.6922
28.51817	5.329739	8.85654	5.61226	30.80000	5.20000	17.34652	24.75573
7.52962	4.21685	6.900000	15.1826	21.2000	4.51839	39.4434	18.42760
43.11759	3.558894	9.831284	21.300000	7.80000	2.200000	0.2000	34.41325
3.85007	7.05843	11.8169	10.82922	6.67962	4.547997	17.6487	18.32347
14.59198	11.18312	1.70000	12.0000	13.50000	5.300000	20.16485	10.170524
13.9724	5.45055	18.89618	0.70000	10.00505	24.7000	20.80065	9.6440
22.71603	0.930583	1.2890	31.20000	0.1000	3.70000	9.60845	1.227067
4.16353	10.2912	40.1149	17.28577	1.80000	8.30000	28.01187	25.8933
19.09834	18.16766	7.76662	20.0000	1.200000	1.90000	3.936185	14.0450
21.97931	4.51446	4.3000	18.7000	0.668438	0.60000	9.83434	34.1988
22.59473	20.45387	5.37965	0.20000	1.20000	16.3000	3.19854	23.5968
6.03989	4.61037	2.400000	6.20045	0.80000	0.60000	13.1363	0.670264
38.89904	1.62974	26.188604	17.254436	2.50000	11.5000	17.31739	3.729134
20.2517	8.87988	31.6049	34.5000	9.87723	0.2000000	7.9794	23.51314
6.81260	1.20000	7.55173	19.26316	2.100000	24.371696	18.16007	18.186828
17.6178	3.77259	14.67670	23.2246	7.03772	5.7000	1.65330	29.9109
0.95027	10.161308	9.4000	20.49895	43.6000	4.715587	0.43095	0.155836
8.64042	15.9418	14.3000	0.70000	18.1000	2.4000	19.03589	28.1013
0.46852	1.40000	18.842856	47.8000	11.20000	4.90000	3.434231	27.62348
1.41567	6.52464	18.8156	11.70958	22.400000	5.10000	7.912523	
1.581969	13.75734	9.31097	14.0763	0.10000	29.36833	26.34322	

Appendix P

Figure 5.30: Sample Observed Data for Wenchi (2000)

Observed data for Wenchi (2000)											
NA	NA	NA	17.5	NA	28.4	8.9	NA	5.7	7.4	9.8	NA
NA	NA	NA	1	NA	NA	NA	17.9	NA	11.8	15.8	NA
NA	NA	NA	NA	NA	NA	NA	0.2	1	NA	14.1	NA
NA	NA	NA	NA	0.7	NA	NA	NA	NA	0.4	NA	NA
NA	NA	NA	2.2	15.6	NA	NA	4.6	NA	NA	16	NA
NA	NA	NA	NA	NA	NA	NA	7.1	8.4	1.1	NA	NA
NA	NA	NA	NA	11.6	NA	NA	NA	NA	6.1	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	1.5	0.2	NA	NA
NA	NA	NA	NA	NA	NA	NA	26.6	1.4	NA	NA	NA
NA	NA	NA	NA	1.7	NA	NA	NA	NA	2.6	3.4	NA
NA	NA	NA	32.1	NA	NA	NA	NA	5.7	NA	0.4	NA
NA	NA	NA	NA	NA	NA	NA	NA	15.6	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	NA	4.2	NA	NA
NA	NA	NA	13.6	NA	NA	NA	NA	0.2	0.8	NA	NA
NA	NA	NA	NA	0.3	NA	NA	6.3	2.4	NA	NA	NA
NA	NA	NA	NA	NA	NA	NA	NA	4.3	NA	NA	NA
NA	NA	NA	12.2	NA	NA	NA	NA	NA	11.9	NA	NA
NA	NA	NA	17.4	NA	NA	NA	NA	NA	11.6	NA	NA
NA	NA	NA	24.3	NA	29.8	NA	NA	1.2	17.8	NA	NA
9.1	NA	9.4	NA	NA	NA	NA	NA	6.7	NA	NA	NA
NA	NA	NA	NA	5.2	3	NA	NA	0.4	4.2	NA	NA
NA	NA	NA	6.9	13.2	13.8	NA	NA	NA	NA	NA	NA
NA	NA	NA	26.6	NA	8.8	NA	NA	NA	49.2	NA	NA
NA	NA	NA	NA	10.5	NA	NA	NA	NA	4.1	NA	NA
27.5	NA	NA	7.7	NA	NA	NA	NA	10.1	NA	NA	NA
12	NA	NA	1.9	NA	47.6	NA	NA	NA	12.5	NA	NA
NA	NA	4.9	7.3	12.4	11.6	NA	NA	NA	NA	NA	NA
NA	NA	NA	NA	1.6	NA	NA	NA	3.2	0.3	NA	NA
NA		NA	NA	NA	NA	NA	0.5	20.9	NA	NA	NA
NA		NA	NA	30.2	NA	NA	4.7	NA	NA	NA	NA
NA		NA		4.5		NA	NA		NA		NA

Appendix P1

Figure 5.31: Sample Complete Data for Wenchi (2000)

Complete Data for Wenchi (year 2000) using EM Algorithm							
10.40405	16.5582	11.976305	16.1264	17.546979	11.1687	0.400000	13.49448
10.50613	0.280853	12.648351	16.1042	4.99882	24.46560	14.88154	10.21412
4.207439	3.6920	2.20000	5.20000	11.10772	6.53363	1.1000	18.4794
5.995765	15.80405	19.6182435	13.20000	10.0008	11.749841	6.10000	5.743950
0.91559	8.24208	1.13259	8.347779	10.0617	5.09580	0.20000	19.87128
3.73745	1.583979	11.3983	10.50000	0.65511	14.380755	7.26116	7.466257
6.84161	0.834495	18.6064	11.35062	11.1412	5.25040	2.60000	6.48522
16.14503	3.805386	15.4571	14.45622	3.88148	12.71732	15.5690	10.1694
19.7844	10.174483	32.100	12.40000	9.025758	18.633	2.09862	16.8196
3.41466	4.90951	8.359047	1.60000	8.59696	2.25412	4.200000	20.9517
1.23844	25.3701	6.3764524	13.5838	9.9016839	0.50000	0.80000	15.78755
23.46466	5.30795	13.60000	30.2000	0.544139	4.70000	0.397728	28.8328
24.627441	1.8954	9.98808	4.50000	7.42848	4.81474	16.283	15.7025
20.894109	13.30249	3.9375835	28.4000	3.00466	5.70000	11.90000	9.0492277
35.4052	11.6815	12.2000	27.053006	16.6824	3.825809	11.6000	8.619575
17.7574	1.4829387	17.4000	13.540586	21.1536	1.00000	17.8000	0.251953
22.95922	0.69971	24.300	29.11118	4.70159	9.9394236	3.55197	21.8801
3.48413	25.423164	22.4715	16.96188	8.91419	8.56814	4.20000	12.2111
14.22704	11.179049	23.137	3.08946	1.294903	8.4000	7.20655	28.6579
9.1000	7.534849	6.90000	1.43587	8.809878	12.4746	49.20000	20.0060
17.3623	6.6494	26.600000	14.2734	6.492655	1.50000	4.10000	0.202434
9.243957	24.3097	3.7769098	13.2807	3.431822	1.40000	22.247208	3.79263
24.02211	10.0736	7.70000	10.8888	17.41498	14.90531	12.500	19.93767
6.8940749	31.4573	1.900000	20.62921	19.6557	5.70000	10.14681	15.59105
27.5000	3.60000	7.30000	12.34186	7.3040	15.60000	0.30000	25.158697
12.000	4.1906369	8.40258	11.712621	27.0962	32.93781	20.8493	0.62224
35.93619	3.65370	15.2907	23.18574	3.57633	0.200000	9.91784	23.6819
2.392500	5.010727	2.9319	3.830561	5.51636	2.40000	3.448112	23.29198
12.6574	9.875576	21.4210	13.13945	19.92136	4.3000	9.8000	17.1464
12.1518	5.1974194	7.97693	9.5867	17.90000	15.1519	15.80000	9.36038
19.2872	0.0100	7.273557	14.8557	0.200000	17.2052	14.10000	4.267657
2.648298	9.14771	0.700000	29.8000	14.33702	1.20000	6.266111	0.38116
16.33490	9.40000	15.6000	5.80939	4.60000	6.70000	16.0000	19.0329
18.68597	15.575	8.893994	3.00000	7.10000	0.40000	4.79445	4.86717
19.18819	4.43351	11.6000	13.80000	3.8684	17.72410	6.99947	3.878081
5.67836	17.441108	14.6711	8.800000	9.84351	8.590543	7.26337	19.809400
32.4935	3.012949	0.3025	14.830163	26.6000	23.60894	5.90445	5.57266
17.27609	0.580560	1.70000	16.46941	4.43672	10.1000	3.40000	30.8631
4.81011	7.995977	33.4858	47.6000	2.962049	3.9615	0.40000	20.00054
15.12666	4.9000000	5.58865	11.6000	20.63831	7.27605	10.87723	0.38512
0.02607	2.32927	3.640735	19.7022	3.09881	3.20000	1.233271	18.18768
7.800558	9.04170	22.121029	17.4779	24.715899	20.9000	13.74353	8.10121
0.367071	13.7589	0.30000	10.3052	6.30000	1.86080	4.53439	4.31082
1.904695	4.5751	6.541120	8.90000	6.34555	7.40000	11.8620	20.56852
0.541577	17.5000	6.91212	12.23660	5.6978	11.80000	6.99963	
7.69725	1.000000	12.72656	5.925675	13.9460	11.469505	0.47749	