UNIVERSITE JOSEPH KI-ZERBO

ECOLE DOCTORALE INFORMATIQUE ET CHANGEMENTS CLIMATIQUES

BURKINA FASO

Unité-Progrès-Justice

MASTER RESEARCH PROGRAM

SPECIALITY: INFORMATICS FOR CLIMATE CHANGE (ICC)

MASTER THESIS

Subject:

DESIGN AND IMPLEMENTATION OF AN OFFLINE DATA REPOSITORY FOR CLIMATE ANALYSIS

Presented on 5th. July, 2021 and by:

Prince ABOKYE

Major Supervisor

Prof. Harouna NAROUA

Co-Supervisor

Dr. Jacques Ulrich DIASSO

(Jury Members)

Academic year 2O20-2021

SUMMARY

DEDICATION

I dedicate this work to my grandmother, Miss Elizabeth Fosu for her unflinching support during my studies.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to the ALMIGHTY GOD for his love, wisdom, intelligence, knowledge and grace in the life He's given me. I am highly indebted to him for this opportunity.

My sincere gratitude also goes to the Director of ED-ICC, Prof. Tanga Pierre ZOUNGRANA and Deputy Director of ED-ICC, Dr. Ousmane COULIBALY for their support and supervision during my studies at the school.

My sincere appreciation also goes to my major supervisor, Prof. Harouna NAROUA and co-supervisor, Dr. Jacques Ulrich DIASSO under whose guidance this project was carried out.

I would like to also thank the Director of ANAM, Ouagadougou, Dr.Joel ZOUNGRANA for his approval for my internship, the availability of data for my research from ANAM and the entire staff of ANAM, Ouagadougou for their unwavering support during my internship.

I further extend my gratitude to the entire staff of ED-ICC and students of the school who contributed in diverse ways for the success of this project.

I also want to appreciate my entire family, especially my grandmother Miss Elizabeth FOSU for their unwavering support throughout my studies in this university.

DECLARATION

I certify that the work for this project was carried out under the supervision of Prof. Harouna Naroua and Dr. Jacques Ulrich Diasso. All care was taken in trying to make plain the subject and to put it into words that are easily understood.

Signed…………………………………………………………………………………......

PRINCE ABOKYE

(STUDENT)

Signed………………………………………………………………………………………

PROF. HAROUNA NAROUA

(MAJOR SUPERVISOR)

Signed…………………………………………………………………………………….

DR. JACQUES ULRICH DIASSO

(CO-SUPERVISOR)

ABSTRACT

There are quite a number of offline climate analysis tools for climate analysis but none has an inbuilt database to store outputs of analysis. Outputs are stored in a folder of the application or software used for the analysis. Quality control such as replacing white spaces with a value accepted by the software had to be done manually by the user. Climate indices are computed in bulk instead of allowing the user to compute specific indices needed at a particular time. For this reason, a data repository for climate analysis called PRIABO Data Repository was created through this research to help deal with these challenges. This data repository deals with the white spaces during the loading of the data into the data repository by replacing the white spaces with a value accepted by the software or data repository. This data repository also computes specific indices at a time depending on the needs of the user. There are 32 climate indices including core and non-core indices but this repository computes 16 climate indices with 2 non-core climate indices and 14 core climate indices. Outputs such as indices and graphs are stored in the database for each station.

Key words : Repository; Climate; Analysis; Data; Indices.

RESUME

Il existe un certain nombre d'outils d'analyse climatique hors ligne, mais aucun ne dispose d'une base de données intégrée pour stocker les résultats de l'analyse. Les résultats sont stockés dans un dossier de l'application ou du logiciel utilisé pour l'analyse. Le contrôle qualité tel que le remplacement des espaces blancs par une valeur acceptée par le logiciel doit être effectué manuellement par l'utilisateur. Les indices climatiques sont calculés en bloc au lieu de permettre à l'utilisateur de calculer des indices spécifiques nécessaires à un moment donné. Pour cette raison, un entrepôt de données pour l'analyse climatique appelé PRIABO Data Repository a été créé grâce à cette recherche pour aider à relever ces défis. Cet entrepôt de données traite les espaces blancs lors du chargement des données en les remplaçant par une valeur acceptée par le logiciel ou l'entrepôt de données. Cet entrepôt de données calcule également des indices spécifiques en fonction des besoins de l'utilisateur. Il existe 32 indices climatiques comprenant des indices principaux et secondaires, mais cet entrepôt calcule 16 indices climatiques dont 2 indices climatiques secondaires et 14 indices climatiques principaux. Les résultats tels que les indices et les graphiques sont stockés dans la base de données pour chaque station.

Mots clés : Entrepôt ; Climat ; Analyse ; Données ; Indices.

ACRONYMS AND ABBREVIATIONS

WISeREP : Weizmann Interactive Supernova Data Repository

PCIC : Pacific Climate Impacts Consortium

ETCCDI : Expert Team on Climate Change Detection and Indices

IPCC : Intergovernmental Panel on Climate Change

JDBC : Java Database Connectivity

API : Application Program Interface

ECMWF : European Centre for Medium-Range Weather Forecasts

CDO : Climate Data Online

NCDC : National Climatic Data Centre

PSL : Physical Sciences Laboratory

RAMADDA : Repository for Archiving, Managing and Accessing Diverse Data

NOAA : National Oceanic and Atmospheric Administration

CMIP5 : Coupled Modeling Intercomparison Projects

PMIP3 : Paleoclimate Model Intercomparison Projects

CHRS : Center for Hydrometeorology and Remote Sensing

PERSIANN : Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks

CCS : Cloud Classification System

CDR : Climate Data Record

ANNs : Artificial Neural Networks

CCl : Commission for Climatology

GEO : Geostationary Equatorial Orbit

PMW : Passive Microwave

LEO : Low Earth-Orbiting

COSMO : Consortium for Small-scale Modeling

STELLA : Stencil Loop Language

DSEL : Domain-Specific Embedded Language

CUDA : Compute Unified Device Architecture

CPU : Central Processing Unit

GPU : Graphics Processing Unit

SAGA : System for Automated Geo-scientific Analyses

GIS : Geographic Information System

GNU : GNU's Not Unix

GUI : Graphic User Interface

CCAFS : CGIAR Research Program on Climate Change, Agriculture and Food Security

CGIAR : Consultative Group on International Agricultural Research

CRAFT : CCAFS Regional Agricultural Forecasting Toolbox

DSSAT : Decision Support System for Agro Technology

APSIM : Agricultural Production System Simulator

SARRA-H : System for Regional Analysis of Agro-Climatic Risks

CPT : Climate Predictability Tool

iAIMS : Integrated Agriculture Information and Management System

CDAT : Climate Data Analysis Tool

SQL : Structured Query Language

GRIB : GRIdded Binary

NetCDF : Network Common Data Form

HDF : Hierarchial Data Format

NCAR : National Centre for Atmospheric Research

WMO : World Meteorological Organization

TABLES

Page

FIGURES

Page

CHAPTER ONE

INTRODUCTION

1.1 INTRODUCTION

A data repository can also be referred to as a data library or data archive. This is a general term to refer to a data set isolated to be mined for data reporting and analysis [1]. The data repository is a huge database infrastructure which is made up of several databases that collect, manage and store data sets for data analysis, sharing and reporting. The term data repository can be used to explain many ways to collect and store data. There are quite a number of examples of data repository namely data warehouse, data lake, data mart, data cube, metadata repository, etc. A data warehouse is a huge repository that accumulates data usually from many sources of a business, without the data being necessarily related. A data lake is a huge data repository that stores unstructured data classified and tagged with metadata. Data marts are subsets of the data repository. They are more aimed to what the data user needs and are quite easy to use. Metadata repositories store metadata and databases. Data cubes are lists of data with three or more dimensions stored as a table [1]. There are methods for promoting network storage of data in a data repository. The first aspect of these methods involves encrypting the data item using a key derived from its content, determining a digital fingerprint of the data item and storing the data item on a storage device at a location associated with the data fingerprint [8]. The second aspect involves determining the digital fingerprint of the data item, testing if the data item already exists in the data repository by comparing its digital fingerprint with the digital fingerprints of the data items already in the repository, etc. These types of data repositories for climate analysis have a backend and a frontend. The backend involves the process of getting the data into the data repository. It involves extracting the data from its source, integrating it programmatically, loading it into the data repository and performing quality control using MySQL and C++ or C sharp. The frontend involves the process of accessing the data from the data repository. This process involves extracting the data from the repository and displaying it to clients, computing climate indices and displaying the associated plots resulting from the analysis using MySQL and C++ or C sharp [9]. Most of the existing data repositories use SQL or other database programming tools, network storage of data and an interactive web-based graphical user interface. A typical example is the Weizmann Interactive Supernova Data Repository (WISeREP) which is an SQL-based database (DB) with an interactive Web-based graphical interface. This data repository serves as an archive of high-quality supernova (SN) spectra, including both historical (legacy) data and data that are accumulated by ongoing modern programs [10]. Concerning climate data analysis, the author considered the computation of climate indices. A climate index is here defined as a calculated value that can be used to describe the state and the changes in the climate system [11]. Climate change, in the Intergovernmental Panel on Climate Change (IPCC) usage, refers to a change in the state of the climate that can be identified (e.g. using statistical tests) by changes in the mean and/or the variability of its properties. That persists for an extended period, typically decades or longer [12]. Climate indices permit a statistical observation of the changes on the dependent climatological aspects, such as analysis and comparison of time series, means, extremes and trends. Statistical methods or techniques can be used for the analysis of climate extremes. Climate extreme events are hard to study and even harder to predict because they are by definition, rare and obey to different statistical laws [13]. These statistical techniques may involve estimation of parameters such as the mean (a measure of location), the variance (a measure of scale) and the correlation/regression coefficients (measures of linear association). In addition, it is often desirable to estimate the statistical significance of the difference between estimates of the mean from two different samples as well as the significance of the estimated measures of association [14]. To analyze the effect of the oceans and atmosphere on land climate, Earth Scientists have developed climate indices, which are time series that summarize the behavior of selected regions of the Earth's oceans and atmosphere [15]. The application of climate indices by definition reduces complex spaces and time variability into simple measures (packages of weather) [16]. These packages of weather are indices of large-scale climate fluctuations that act as proxies for the overall climate condition. Climate is the average weather of a geographical location for many years, normally 30 years or more. The complexity of the climate variability on all time scales requires the use of several refined tools to unravel its primary dynamics from observations [17].

For this research work, an Offline Data Repository for Climate Analysis will be created which will perform quality control, compute climate indices, give outputs like plots and store processed data (observation and satellite temperature and rainfall data) and outputs for future use.

1.2 RESEARCH PROBLEM

Main Research Problem

- Is it possible to have an offline data repository for climate analysis?

Secondary Research Problem

- Can a climate analysis tool be written with C++?
- Can a climate analysis tool be merged with a database?

1.3 HYPOTHESIS

Main Hypothesis

- An Offline Data Repository for Climate Analysis can be created using MySQL and C++ which will perform quality control, compute climate indices, give outputs like plots and store processed data (observation and satellite temperature and rainfall data) and outputs for future use.

Secondary/Specific Hypothesis

- A Climate Analysis Tool can be written in C++ using Text editor to create a class to search through the data and some functions to compute the climate indices.
- A Climate Analysis Tool written in C++ can be merged with a database created with MySQL using the MySQL Client Library, MySQL C++ Connector and a Text Editor to form an offline Data Repository for Climate Analysis.

1.4 RESEARCH OBJECTIVES

General Objective One

- Develop a C++ Tool for Climate Analysis

Specific Objectives for General Objective One

- Load Data
- Perform Quality Control
- Compute Climate Indices

General Objective Two

- Merge a Climate Analysis Tool with a Database

Specific Objectives for General Objective Two

- Create a Database
- Merge a C++ Application with a Database

1.5 JUSTIFICATION OF STUDY

With an ever-growing supply of climate data from satellites and environmental sensors, the magnitude of data and climate model output has begun to overwhelm the relatively simple tools currently used to analyze them. A computational approach will therefore be indispensable for the analysis. Informatics for Climate Change, which involves a collaboration between climate scientists and machine learning researchers in order to bridge the gap between data and understanding, helps to deal with this challenge [7]. However, dealing with this challenge will not be possible without a database. For this reason, I decided to create a data repository which will help conduct climate analysis. Moreover, most of the existing data repositories for climate analysis are online, meaning that they are not accessible without internet connection. The proposed data repository for climate analysis will be offline in order to make the data analysis accessible without internet connection. Another issue is the addition of new functionalities that have not been considered in the existing repositories such as loading and performing quality control on the climate data (rainfall and temperature), loading the activity data and emission factors of different locations, calculating the contribution of each greenhouse gas, calculating the total amount of greenhouse gas emitted in CO_2 equivalence, calculating the net present value of carbon credit benefit per year together with other benefit and net present values of opportunity cost per year and other costs given the required input, the Cost Benefit Ratio, etc. There are quite a number of offline climate analysis tools for climate analysis but none has an inbuilt database to store outputs of analysis. Outputs are stored in a folder of the application or software used for the analysis. Quality control such as replacing white spaces with a value accepted by the software had to be done manually by the user. Climate indices are computed in bulk instead of allowing the user to compute specific indices needed at a particular time. For this reason, a data repository for climate analysis called PRIABO Data Repository was created through this research to help deal with these challenges. This data repository deals with the white spaces during the loading of the data into the data repository by replacing the white spaces with a value accepted by the software or data repository. This data repository also computes specific indices at a time depending on the needs of the user. There are 32 climate indices including core and non-core indices but this repository computes 16 climate indices with 2 non-core climate indices and 14 core climate indices. Outputs such indices and graphs are stored in the database for each station.

1.6 SCOPE AND LIMITATION

This data repository is a console application that runs on Ubuntu or Linux operating systems. This is due to the use of some sudo commands during the coding which work on Linux or Ubuntu terminals but not on other platforms. Work is still in progress to develop a windows version and subsequently a platform independent version. Another limitation is that it runs on Linux or Ubuntu systems that have been configured with the MySQL Connector/C++ 1.1 which is a MySQL database connector for C++ applications or programs that connect to MySQL servers. Connector/C++ enables or allows the development of C++ applications that use the JDBC-based API (Java Database Connectivity-based Application Program Interface). This means that just installing MySQL on your system is not enough to run this application. Due to this limitation, a read me file has been added to the application folder to guide users on how to configure and use it.

1.7 LITERATURE REVIEW

There exists so many climate data repositories for climate analysis but none is offline. Another interesting thing about these repositories is that most of them do not have a column or space for quality control of data downloaded or have an already performed quality control on the data to be downloaded. Moreover, the climate analysis performed by these repositories are woefully inadequate. Furthermore, most of the climate analysis tools I came across are offline and perform more climate analysis than these repositories. The only issue with these climate tools is that they do not have in built databases to store processed data and outputs like plots after an analysis has been done. Below is a brief description of some of the climate data repositories for climate analysis I came across during my review:

- The Climate Data Store which is implemented by ECMWF (European Centre for Medium-Range Weather Forecasts) and supported by Copernicus [2]. This platform renders climate change services and helps dive into a wealth of information about the Earth's past, present and future climate. It is freely available and functions as a one stop shop to explore climate data. It also has a toolbox for performing analysis online and an Application Program Interface (API) that helps to download data from the repository. This platform exclusively maintains data for European nations and is a web application with a backend database.

- The Climate Data Online (CDO) which provides free access to NCDC's (National Climatic Data Centre) archive of global historical weather and climate data in addition to station history information [3]. These data include quality controlled daily, monthly, seasonal, and yearly measurements of temperature, precipitation, wind, and degree days as well as radar data and 30-year Climate Normals. It is possible to find and view global past weather and climate data by station name or identifier, zip code, city, country or state. This is also a web application with a backend database.
- The Physical Sciences Laboratory (PSL) Climate Data Repository which is a RAMADDA (Repository for Archiving, Managing and Accessing Diverse Data) server for storing and accessing information used in NOAA (National Oceanic and Atmospheric Administration) / PSL's Interpreting Climate Conditions research [4]. RAMADDA is a content management system that can provide access to data, metadata, presentations, images and analysis of geoscience data. It is a web-based application framework that provides a broad suite of services for content and data management, publishing and collaboration. It also brings together a number of concepts and technologies to provide an easy to use but powerful system for publishing, organizing, discovering. and accessing data and other holdings.

The review of the literature has also revealed many research works carried out on climate data analysis. Roland [5] has investigated on Big Climate Data Analytics, specifically on the Effective Knowledge Discovery from Colombia's Weather Data. The aim of his article was to develop a distributed machine learning application that classifies Colombian Climate Data and provides Decision Support to environmental designers seeking to understand the spatial and temporal use of low-energy design strategies. It gives an idea of how to analyze big climate data using machine learning and make informed decisions.

Stefan et al. [6] worked on Road Map to Climate Data Rescue Services. The aim of their article was to provide a summary of a survey that they have undertaken on several meteorological and climate data rescue projects, in order to identify the needs of climate data rescue services. They emphasized the need to transform hard copy climate data to soft copy climate data by storing them in data repositories.

Matheus et al. [18] also worked on a web repository called ecoClimate which is a free and open data repository developed to serve useful climate data to macro-ecologists and bio-geographers. This arose from the need for Climate Layers with which to build ecological niche models and test macro-ecological and bio-geographic hypotheses in the past, present and future. The data repository provides a package of processed, multi-temporal climate datasets from the most recent multi-model ensembles developed by the Coupled Modeling Intercomparison Projects (CMIP5) and the Paleoclimate Model Intercomparison Projects (PMIP3) across past, present and future time frames, at global extents and 0.5 degrees spatial resolution, in convenient formats for analysis and manipulation. It's main aim was to maintain consistency across diverse data and still retain information on uncertainties among model predictions. Some of the outputs or results of this data repository are maps of time series illustrating the global extent and the multi-temporal characteristics of climate data available in ecoClimate. This is also a web application with a backend database.

Phu et al. [19] have worked on a Data Portal called CHRS Data Portal which is an easily accessible public repository for PERSIANN global satellite precipitation data. The name of the data portal was derived from an organisation called Center for Hydrometeorology and Remote Sensing (CHRS). This organisation created the data portal with the aim of facilitating easy access to the three open data licensed satellite-based precipitation datasets generated by their Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) system : PERSIANN, PERSIANN-Cloud Classification System (CCS), and PERSIANN-Climate Data Record (CDR). As its name suggests, PERSIANN uses machine learning technique known as Artificial Neural

Networks (ANNs) to determine the relationship between remotely sensed cloud-top temperature, measured by long-wave infrared (IR) sensors on Geostationary Equatorial Orbit (GEO) satellites and rainfall rates with bias correction from passive microwave (PMW) measured by Low Earth-Orbiting (LEO) satellites. From the idea of the PERSIANN algorithm, researchers at CHRS created two additional hydrometeorological estimation algorithms: PERSIANN-CCS and PERSIANN-CDR. PERSIANN-CSS estimates global rainfall in near real-time and at a higher spatial resolution than PERSIANN through its cloud clustering algorithm and its use of IR (InfraRed) data as the sole input. PERSIANN-CDR uses the PERSIANN algorithm with historical IR data which began in 1983 to construct a consistent dataset for climate trend studies. The expected result of this data portal was an open, easy to use, online web system where users can quickly customize, visualize, compare and retrieve data from the three PERSIANN algorithms for analysis, modeling and public education.

The next part of the review focuses on the use of C++ or C sharp in designing applications in climate science such as climate models, etc. Oliver et al. [20] proposed a software implementation strategy for complex weather and climate models that produce performance portable, architecture diagnostic nodes etc. This relies on domain and data structure specific tools that are usable within common model development frameworks like FORTRAN, C++ and possibly Python in the future. Concerning the software implementation strategy, a refactoring project was performed on an atmospheric model called COSMO (Consortium for Small-scale Modeling). This strategy was done by rewriting the dynamical core of the model and refactoring the remaining FORTRAN code. This dynamical core was built on top of the domain specific Stencil Loop Language (STELLA) for stencil computations on structured grids. STELLA is a domain-specific embedded language (DSEL) that leverage on C++ and CUDA. CUDA means Compute Unified Device Architecture and it is a parallel computing platform. It enables softwares to perform calculations using both the CPU and GPU. Every tool mentioned earlier was implemented in C++ making extensive use of generic programming and template meta-programming. The results of this research shows that the refactored code outperformed the current production code and was performance portable to various hybrid CPU-GPU node architectures.

Conrad et al. [21] worked on System for Automated Geo-scientific Analyses (SAGA) v. 2.1.4 which is an open-source geographic information system (GIS), mainly licensed under the GNU General Public License. This software was initially designed for digital terrain analysis but it is now a comprehensive and globally established GIS platform for scientific analysis and modeling. SAGA was coded in C++ in an object-oriented design and runs under several operating systems including Windows and Linux with the help of the crossplatform wxWidgets GUI Library. The main functional features of this modular software architecture are an application programming interface for the development and implementation of new geoscientific methods, a user-friendly graphical interface with many visualization options, a command line interpreter and interfaces to interpreted languages like R and Python. The version 2.1.4 adds more than 600 tools which are implemented in dynamically loadable libraries and represent the broad scopes of SAGA in numerous fields of geoscientific endeavor and beyond. The results or outputs of this software are scientific applications such as digital terrain analysis, geomorphology, soil science, climatology and meteorology, as well as remote sensing.

Vakhtang et al. [22] worked on a framework for forecasting crop production, risk analysis and climate change impact studies since Regional crop production forecasting is growing in importance in both the public and private sectors. This growth is to ensure food security, optimize agricultural management practices and use of resources, and anticipate market fluctuations. They produced a model and data driven, easy to use forecasting and a risk assessment framework which will be an important tool for end users at different levels. A toolbox called CCAFS Regional Agricultural Forecasting Toolbox (CRAFT) for gridded crop modeling and yield forecasting along with risk analysis and climate impact studies was designed based on the Microsoft .NET Windows platform and includes a user-friendly client application developed in C# that provides the interface with crop models and a MySQL database implementation. As results, the CRAFT produces multiple simulation scenarios, maps and interactive visualizations using a crop engine that can run the pre-installed crop models DSSAT, APSIM and SARRA-H alongside the Climate Predictability Tool (CPT) for seasonal climate forecasts.

Yang et al. [23] developed a system called Integrated Agriculture Information and Management System (iAIMS) to fix a challenge facing scientists in their ability to dynamically access spatially referenced climatic, soil and cropland data. Their management system consisted of foundation class climatic, soil and cropland databases which served as a foundation to develop applications that address the different aspects of cropping systems performance and management. A Climate Data System was developed as a component of iAIMS to automatically fetch data from different web sources and consolidate them into a centralized database and deliver the data through a web-based interface. These climatic data are available through web pages and the steps to access them from different sources vary depending on how the data are rendered. The Climate Data System building process is divided into five major program modules corresponding to different phases of the process namely Data Requester, Data Fetcher, Data Parser, Data Filter and Data Explorer. All the program modules were implemented using Microsoft Visual Studio 2005 and the C#.NET Language. There are three types of data that are stored in the process namely original climatic data in file format, parsed climatic data in SQL Server database and filtered climatic data in SQL Server database. The result of this Climatic Data System component of the iAIMS is that it will help address water conservation, crop production and management, land use suitability analysis and bio-energy refinery site selection.

As stated earlier, there exist climate analysis tools which are offline but doesn't have in built database to store cleaned data and output of climate analysis after analysis have been done and have also been written in R programming language. On the contrary, the online climate analysis tools have in built databases to store data and output after analysis. A typical example is CDAT (Climate Data Analysis Tool) which was developed by Lawrence Livermore National Lab and written in Python [26]. The next section of the literature review will highlight two of these tools. This research work will be based on the concept of one of these tools. Another important factor which also influences the type of tool is the type of data format. Data format encountered in climate research generally fall into three categories namely GRIdded Binary (GRIB) format, Network Common Data Form (NetCDF) format and Hierarchial Data Format (HDF) [24]. For this research work, text file data format will be used and a compiled langage like C++ and C sharp will be used. According to Dennis Shea [25] of National Centre for Atmospheric Research

(NCAR), a climate data format called Binary are created from compiled languages such as FORTRAN or C/C++ and also text files data format are also used in climate research.

ClimPACT is the first offline climate analysis tool to be discussed. ClimPACT is a software package to calculate climate indices that are relevant for health, agriculture and water sectors and this software package allows you to calculate climate indices from your own daily weather data either in plain-text point data (eg. from a weather station) or across an entire gridded NetCDF file (eg., climate model output) [27]. Climpact was written in R programming langage. ClimPACT has an online version now. This research paper will be based on the ClimPACT methodology which includes loading data, performing quality control and computing climate indices but using C++ programming langage and including an in-built database using MySQL. The new features to be included will be saving data and output like graphs in the in-built database with new save button and opening existing data and outputs from a previous analysis with the use button. Before using ClimPACT2, developers usually recommend the use of RHTest software package to check for homogenity of the data before proceeding [29]. For this reason, the idea behind RHTest will be taken into consideration during the writing of the quality control program for the data repository [33]. 16 out of the climate indices will be computed in this research [32]. Two non-core indices namely Daily Mean Temperature (TM) and Diurnal Temperature Range (DTR) and fourteen core indices namely Summer days (SU25), Tropical nights (TR), Daily Mean Temperature above 5 degrees celsius (TM5a), Daily Mean Temperature below 5 degrees celsius (TM5b), Daily Mean Temperature above 10 degrees celsius (TM10a), Daily Mean Temperature below 10 degrees celsius (TM10b), Hot days (SU30), Very Hot days (SU35), Annual number of wet days (R1mm), Annual number of days with rainfall greater than or equal to 10mm (R10mm), Annual number of days with rainfall greater than or equal to 20mm (R20mm), Annual number of days with rainfall greater than or equal to 30mm (R30mm), Annual Total Rainfall from Wet Days (PRCPTOT) and Mean Rainfall Amount on a Wet Day or Simple Daily Intensity (SDII) will be computed [29]. Another version called ClimPACT2 was created but became deprecated in December 2020 [30]. ClimPACT2 updates ClimPACT which was based on RClimDex now called ClimDex [29]. Current version is Climpact and allows you to calculate climate indices from your own daily weather data either in plain-text point data (eg. from a weather station) online [31]. This was created with shiny apps and R programming langage.

RClimDex is the second offline climate analysis tool to be dicussed. RClimDex is also a software package for calculating climate indices [28]. It also has an online version and can also allow you to work with your own data just like ClimPACT. It was also written in R programming langage and was formally called RClimDex because of this. The RClimDex software was developed the World Meteorological Organization (WMO) Commission for Climatology (CCl) Expert Team on Climate Change Detection and Indices (ETCCDI). The majority of indices in ClimPACT2 are calculated using the code from the climdex.pcic R package which was developed by the Pacific Climate Impacts Consortium (PCIC) [29]. Due to this similarity, there wasn’t any need basing the research work on RClimDex also but it was important to make mention of it.

CHAPTER TWO
METHODOLOGY

2.1 DATA AND TOOLS

2.1.1 DATA

The study variables for the main and specific hypothesis are:

- Observation data for Daily Rainfall and Temperature for Sydney, Australia
- Observation data for Daily Rainfall and Temperature for Ouagadougou, Burkina Faso

There are two copies of each data, one for ClimPACT2 and the other for the PRIABO Data Repository.

2.1.2 TOOLS

The first tool used was **ClimPACT2**. It is a climate analysis tool used for computing climate indices. This was used first to load the first copy of the data, perform quality control, compute the climate indices and outputs stored in a folder in ClimPACT2. This was done for both the Sydney data and the Ouagadougou data.

The second tool used was the **PRIABO Data Repository** created using the following tools:

- **MySQL** – Allows the usage of MySQL Commands
- **MySQL Connector/C++ 1.1 Source Distribution** – Allows the connection of the C++ program to the database
- **C++ Programming Language** – Programming language used for the project
- **G++ or GNU C++ Compiler** – Allows compilation of the C++ Program
- **Text Editor** – Acts as an IDE for the C++ Program
- **Ubuntu/Linux Terminal Commands** – Sudo commands of the Ubuntu Terminal
- **GNU Plot** – Allows the plotting of graphs with C++ Program
- **CMake** – Allows the generation of native build tool files specific to your computer
- **MySQL Client Library** – Library through which the connection to the database is created
- **Boost C++ Libraries** – Provides support for task and structures of the C++ program

The second tool was also used to first load the second copy of the data, perform quality control, compute the climate indices and outputs stored in a folder in the PRIABO Data Repository folder and subsequently in the database of the data repository. This was done for both the Sydney data and the Ouagadougou data.

2.2 DEFINITION OF CLIMATE INDICES COMPUTED

The following table defines the different climate indices to be computed:

Table 2.1: Definition of Computed Climate Indices

NON CORE INDICES		
INDEX	**UNIT**	**INDEX DESCRIPTION**
TM	Degree Celsius (℃)	Daily Mean Temperature TM = (TX+TN)/2 where TX is Maximum Temperature and TN is Minimum Temperature
DTR	Degree Celsius (℃)	Diurnal Temperature Range DTR = TX - TN
CORE INDICES		
INDEX	**UNIT**	**INDEX DESCRIPTION**
SU25	Days	Summer Days which is Annual Number of Days when TX > 25℃
TR	Days	Tropical Night which is Annual Number of Days when TN > 20℃
TM5a	Days	Annual Number of Days when TM >= 5℃ where TM is Daily Mean Temperature
TM5b	Days	Annual Number of Days when TM <5℃
TM10a	Days	Annual Number of Days when TM >= 10℃
TM10b	Days	Annual Number of Days when TM <10℃
SU30	Days	Hot Days which is Annual Number of Days when TX

		>= 30℃
SU35	Days	Very Hot Days which is Annual Number of Days when TX > =35℃
R1mm	Days	Annual number of wet days or Annual number of days when PR >= 1mm where PR is Precipitation and mm is millimeters.
R10mm	Days	Annual number of days when daily PR >= 10mm
R20mm	Days	Annual number of days when daily PR >= 20mm
R30mm	Days	Annual number of days when daily PR >= 30mm
PRCPTOT	mm	Annual Total Rainfall from Wet Days (A Wet Day has PR >= 1mm) or Annual Sum of Daily PR >=1mm
SDII	mm/day	Mean Rainfall Amount on a Wet Day or Simple Daily Intensity (SDII)

2.3 DATABASE DESIGN

2.3.1 DATABASE DESIGN USING VISUAL PARADIGM

The database design using Visual Paradigm is shown in Figure 2.1.

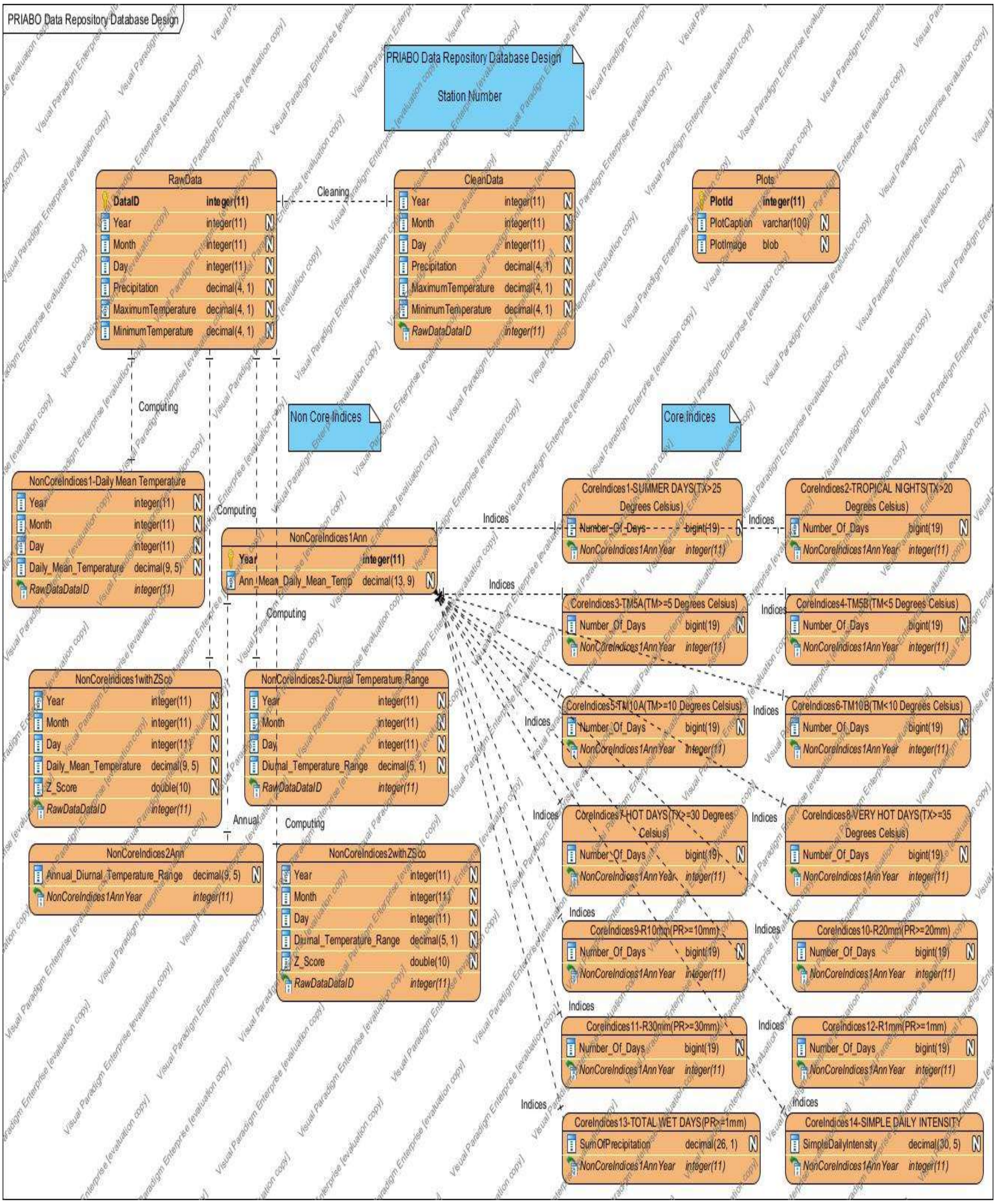

Figure 2.1: Database Design using Visual Paradigm

2.3.2 DATABASE DESIGN DESCRIPTION

There are twenty-three (23) tables in each database. Each station has its own database. For instance, station1 is a database for Sydney Observatory in Australia and Station2 is a database for ANAM, Ouagadougou in Burkina Faso. Apart from RawData table, CleanData table and Plots table, the rest of the tables were created automatically by the PRIABO Data Repository Application. Type in each table stands for the datatype. For instance, int stands for integer. Key in each table specifies if the column or field is a primary key or foreign key. For instance, PRI means the field is a primary key and MUL means multiple key which means that a particular key is a primary key in another table and a foreign key in the current table. Below is a list of tables and their properties or attributes.

Table 2.2: Tables in Station Number

Tables in Station Number
RawData
CleanData
Plots
NonCoreIndices1
NonCoreIndices1Ann
NonCoreIndices1withZSco
NonCoreIndices2
NonCoreIndices2Ann
NonCoreIndices2withZSco
CoreIndices1
CoreIndices2
CoreIndices3
CoreIndices4
CoreIndices5
CoreIndices6
CoreIndices7
CoreIndices8
CoreIndices9
CoreIndices10
CoreIndices11

CoreIndices12
CoreIndices13
CoreIndices14

Table 2.3: RawData Table

RawData		
Field	**Type**	**Key**
DataID	int(11)	PRI
Year	int(11)	
Month	int(11)	
Day	int(11)	
Precipitation	decimal(4,1)	
MaximumTemperature	decimal(4,1)	
MinimumTemperature	decimal(4,1)	

Table 2.4: CleanData Table

CleanData		
Field	**Type**	**Key**
DataID	int(11)	MUL
Year	int(11)	
Month	int(11)	
Day	int(11)	
Precipitation	decimal(4,1)	
MaximumTemperature	decimal(4,1)	
MinimumTemperature	decimal(4,1)	

Table 2.5: Plots Table

Plots		
Field	**Type**	**Key**
PlotId	int(11)	PRI
PlotCaption	varchar(100)	
PlotImage	blob	

Table 2.6: NonCoreIndices1 Table

NonCoreIndices1		
Field	**Type**	**Key**
DataID	int(11)	MUL
Year	int(11)	
Month	int(11)	
Day	int(11)	
Daily_Mean_Temperature	decimal(9,5)	

Table 2.7: NonCoreIndices1Ann Table

NonCoreIndices1Ann		
Field	**Type**	**Key**
Year	int(11)	PRI
Ann_Mean_Daily_Mean_Temp	decimal(13,9)	

Table 2.8: NonCoreIndices1withZSco Table

NonCoreIndices1withZSco		
Field	**Type**	**Key**
DataID	int(11)	MUL
Year	int(11)	
Month	int(11)	
Day	int(11)	
Daily_Mean_Temperature	decimal(9,5)	
Z_Score	double	

Table 2.9: NonCoreIndices2 Table

NonCoreIndices2		
Field	**Type**	**Key**
DataID	int(11)	MUL
Year	int(11)	
Month	int(11)	
Day	int(11)	
Diurnal_Temperature_Range	decimal(5,1)	

Table 2.10: NonCoreIndices2Ann Table

NonCoreIndices2Ann		
Field	**Type**	**Key**
Year	int(11)	MUL
|Annual_Diurnal_Temperature_Range	decimal(9,5)	

Table 2.11: NonCoreIndices2withZSco Table

NonCoreIndices2withZSco		
Field	**Type**	**Key**
DataID	int(11)	MUL
Year	int(11)	
Month	int(11)	
Day	int(11)	
Diurnal_Temperature_Range	decimal(5,1)	
Z_Score	double	

Table 2.12: CoreIndices1 Table

CoreIndices1		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.13: CoreIndices2 Table

CoreIndices2		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.14: CoreIndices3 Table

CoreIndices3		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.15: CoreIndices4 Table

CoreIndices4		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.16: CoreIndices5 Table

CoreIndices5		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.17: CoreIndices6 Table

CoreIndices6		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.18: CoreIndices7 Table

CoreIndices7		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.19: CoreIndices8 Table

CoreIndices8		
Field	**Type**	**Key**

Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.20: CoreIndices9 Table

CoreIndices9		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.21: CoreIndices10 Table

CoreIndices10		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.22: CoreIndices11 Table

CoreIndices11		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.23: CoreIndices12 Table

CoreIndices12		
Field	**Type**	**Key**
Year	int(11)	MUL
Number_Of_Days	bigint	

Table 2.24: CoreIndices13 Table

CoreIndices13		
Field	**Type**	**Key**
Year	int(11)	MUL
SumOfPrecipitation	decimal(26,1)	

Table 2.25: CoreIndices14 Table

CoreIndices14		
Field	**Type**	**Key**
Year	int(11)	MUL
SimpleDailyIntensity	decimal(30,5)	

2.4 PROGRAMMING APPROACH

2.4.1 CONFIGURATION APPROACH

The laptop or computer on which this program will run needs to be configured first. This configuration is Operating System (OS) dependent. There is a configuration approach for each OS and all this information is available on MySQL website. Linux or Ubuntu was used for development of this application hence the configuration description stated here is for Linux or Ubuntu.

- Install g++ compiler or GNU C++ Compiler
- Install MySQL
- Install some system prerequisites tools namely Build tools like CMake, MySQL Client Library and Boost C++ Libraries
- Obtain and unpack a connector/C++ Source Distribution from github
- Install the connector/C++ from Source on your Linux and Unix Like System like Ubuntu
- Test if the configuration was successful with a standalone program in the examples folder in the C++ Source Distribution file.
- If successful this should show on your terminal

“Connector/C++ standalone program example...

... running 'SELECT 'Welcome to Connector/C++' AS _message'

... MySQL replies: Welcome to Connector/C++
... say it again, MySQL
....MySQL replies: Welcome to Connector/C++

... find more at http://www.mysql.com, [34]"

2.4.2 PROGRAM STRUCTURE

The application is designed to be connected to the MySQL Database through MySQL Client API Library (libmysqlclient) and MySQL Connector/C++ 1.1 Source Distribution as illustrated in Figure 2.2.

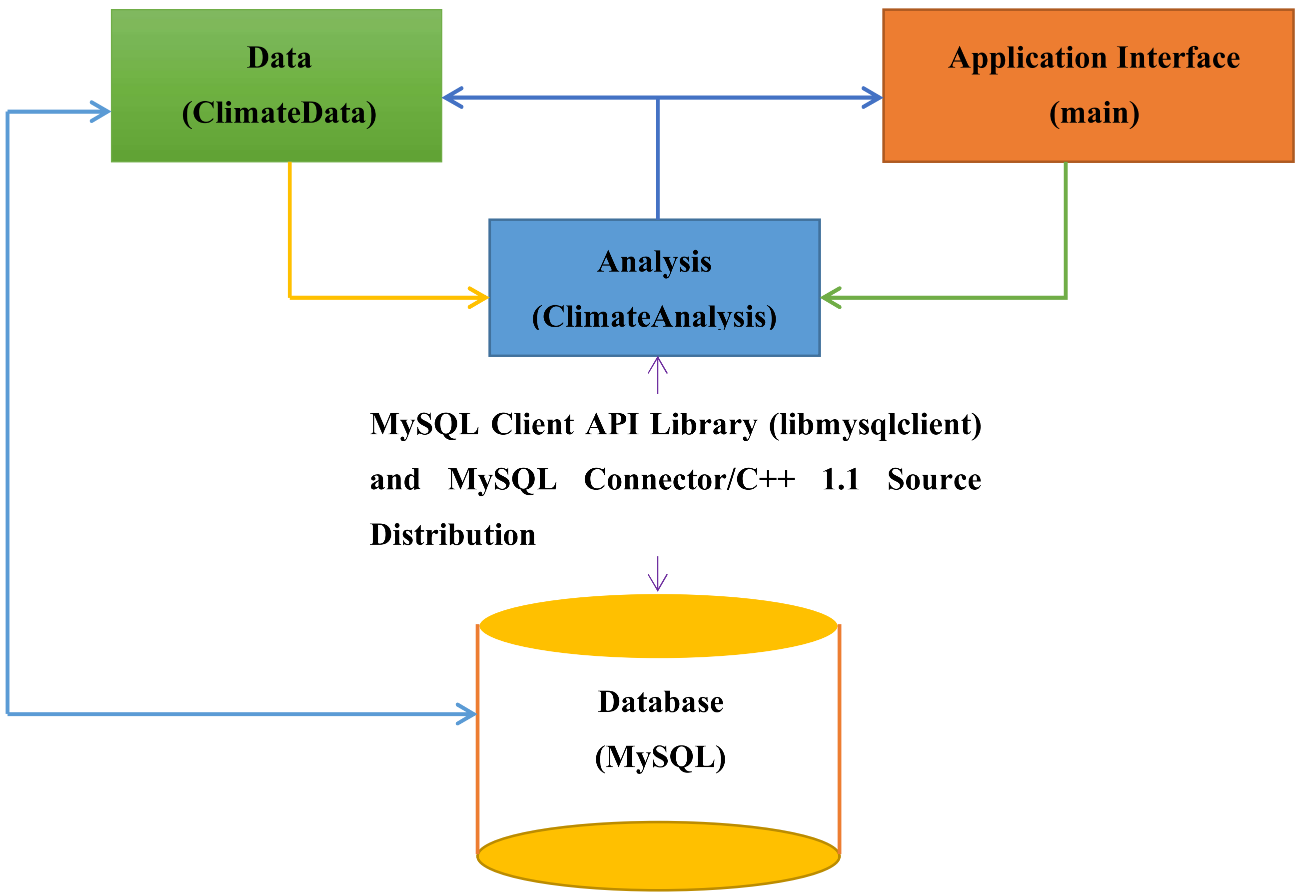

Figure 2.2: Program Structure

MySQL Client API Library (libmysqlclient) is a native C API library distributed with MySQL and implemented in the libmysqlclient library. It is normally installed after the installation of MySQL Server. It is also possible to install it separately. MySQL Connector/C++ 1.1 Source Distribution is an API based partially on the JDBC4.0 API standard. It has connectors for both C and C++. Its main function is to ensure database connectivity [35]. This also provides C++ and plain C interfaces for communicating with MySQL servers [34].

To fully understand this program, the codes are available at ED-ICC, Universite Joseph Ki-Zerbo, Ouagadougou, Burkina Faso.

2.5 IMPLEMENTATION AND USER GUIDE

To make things very easy, a make file was created for the compilation of this program. Also make sure you are using a Linux or Ubuntu Operating System machine as the program is platform dependent. Below are the steps to run the program;

- Open the PRIABO folder
- type **./priabo**
- The menu below will display on the terminal if successful

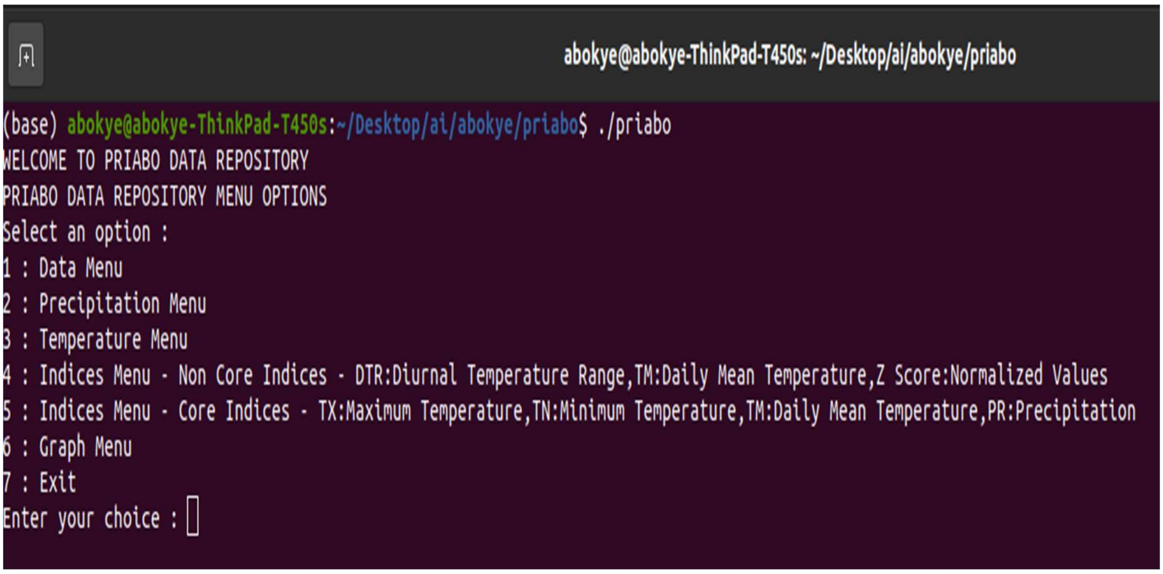

Figure 2.3 : Program Menu of PRIABO Data Repository

CHAPTER THREE
RESULTS AND DISCUSSION

3.1 RESULTS

The application was used to perform various calculations on the sixteen climate indices presented in chapter 2. Numerical and graphical results have been computed by both **ClimPACT2** and **PRIABO Data Repository**. To prevent any biasness, the results of one Core Index and one Non Core Index are selected from the sixteen climate indices computed. The one selected Non Core Index is Annual Mean Daily Mean Temperature (TM) and the one selected Core Index is Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) as shown in Tables 3.1, 3.3, 3.5 and 3.7.

Table 3.1: Annual Mean Daily Mean Temperature (TM) of ClimPACT2 and PRIABO Data Repository using Observation data for Daily Temperature for Sydney, Australia from 1936 to 2015.

CLIMPACT2		PRIABO DATA REPOSITORY	
Year	**Annual Mean TM**	**Year**	**Annual Mean TM**
1936	17.49203	1936	17.49203
1937	17.69794	1937	17.69794
1938	17.92849	1938	17.92849
1939	17.74258	1939	17.74258
1940	18.01639	1940	18.01639
1941	17.36685	1941	17.36685
1942	17.7589	1942	17.7589
1943	16.74274	1943	16.74274
1944	17.41052	1944	17.41052
1945	17.40205	1945	17.40205
1946	17.55247	1946	17.55247
1947	17.54918	1947	17.54918
1948	17.11772	1948	17.11772
1949	17.36676	1949	17.36676

CLIMPACT2		PRIABO DATA REPOSITORY	
Year	**Annual Mean TM**	**Year**	**Annual Mean TM**
1950	17.82904	1950	17.82904
1951	17.66425	1951	17.66425
1952	17.60027	1952	17.60027
1953	17.60589	1953	17.60589
1954	17.72014	1954	17.72014
1955	17.8037	1955	17.8037
1956	17.61434	1956	17.61434
1957	17.83159	1957	17.83159
1958	18.30317	1958	18.30317
1959	17.99548	1959	17.99548
1960	17.58634	1960	17.58634
1961	17.77616	1961	17.77616
1962	17.7874	1962	17.7874
1963	17.80342	1963	17.80342
1964	18.11161	1964	18.11161
1965	17.94807	1965	17.94807
1966	17.54437	1966	17.54437
1967	17.85083	1967	17.85083
1968	18.31671	1968	18.31671
1969	18.05822	1969	18.05822
1970	17.71484	1970	17.71484
1971	17.85481	1971	17.85481
1972	17.99301	1972	17.99301
1973	18.78981	1973	18.78981
1974	17.83475	1974	17.83475
1975	18.38278	1975	18.38278
1976	18.09686	1976	18.09686
1977	18.44342	1977	18.44342
1978	17.97088	1978	17.97088

CLIMPACT2		PRIABO DATA REPOSITORY	
Year	**Annual Mean TM**	**Year**	**Annual Mean TM**
1979	18.41901	1979	18.41901
1980	18.85589	1980	18.85589
1981	18.41538	1981	18.41538
1982	18.1589	1982	18.1589
1983	18.34821	1983	18.34821
1984	17.89144	1984	17.89144
1985	18.11438	1985	18.11438
1986	18.12686	1986	18.12686
1987	18.29423	1987	18.29423
1988	18.97486	1988	18.97486
1989	18.33945	1989	18.33945
1990	18.3789	1990	18.3789
1991	18.7105	1991	18.7105
1992	17.69438	1992	17.69438
1993	18.32176	1993	18.32176
1994	18.28819	1994	18.28819
1995	17.88777	1995	17.88777
1996	**-99.9**	**1996**	**18.10606**
1997	18.36192	1997	18.36192
1998	18.80932	1998	18.80932
1999	18.31178	1999	18.31178
2000	18.61885	2000	18.61885
2001	18.89068	2001	18.89068
2002	18.81497	2002	18.81497
2003	18.56699	2003	18.56699
2004	19.03128	2004	19.03128
2005	19.06342	2005	19.06342
2006	18.86329	2006	18.86329
2007	18.91192	2007	18.91192

CLIMPACT2		PRIABO DATA REPOSITORY	
Year	Annual Mean TM	Year	Annual Mean TM
2008	18.17486	2008	18.17486
2009	18.99325	2009	18.99325
2010	18.78678	2010	18.78678
2011	18.66575	2011	18.66575
2012	18.51421	2012	18.51421
2013	19.37041	2013	19.37041
2014	19.27912	2014	19.27912
2015	19.03712	2015	19.03712

Figures 3.1 and 3.2 give the profiles of the Annual Mean Daily Mean Temperature using Observation data for Daily Temperature for Sydney, Australia from 1936 to 2015.

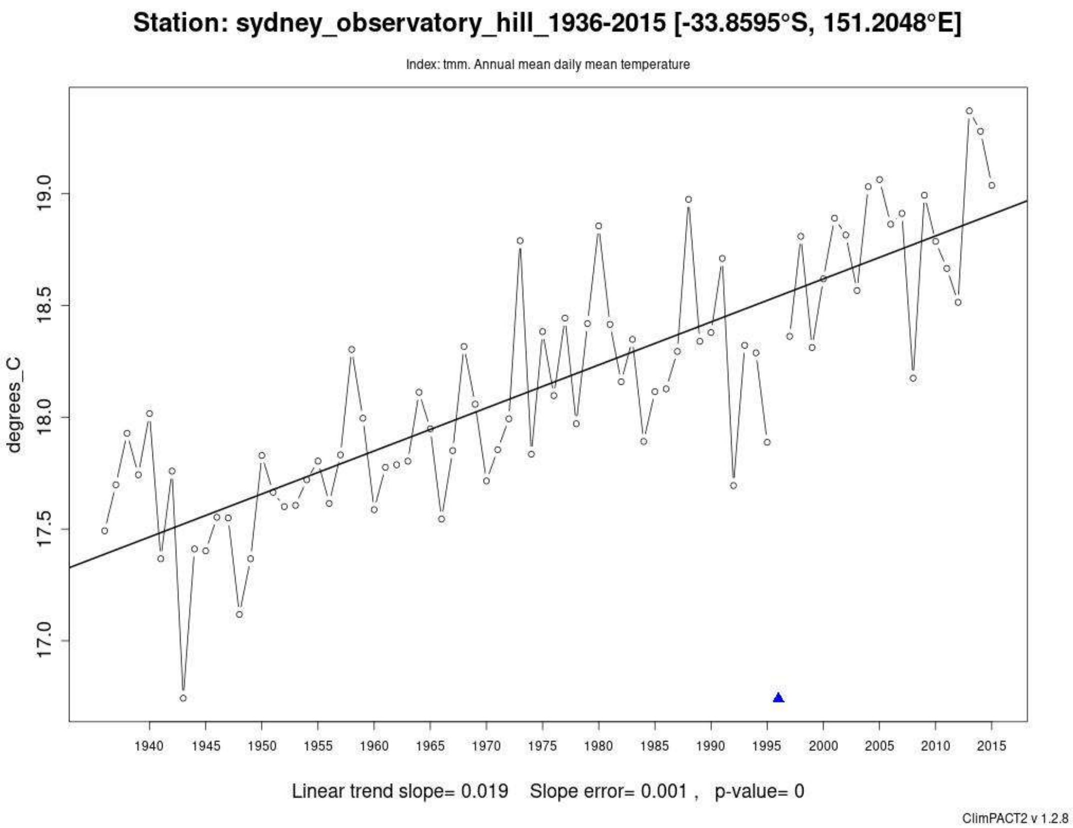

Figure 3.1 : Graph of Annual Mean Daily Mean Temperature of ClimPACT2 using Observation data for Daily Temperature for Sydney, Australia from 1936 to 2015.

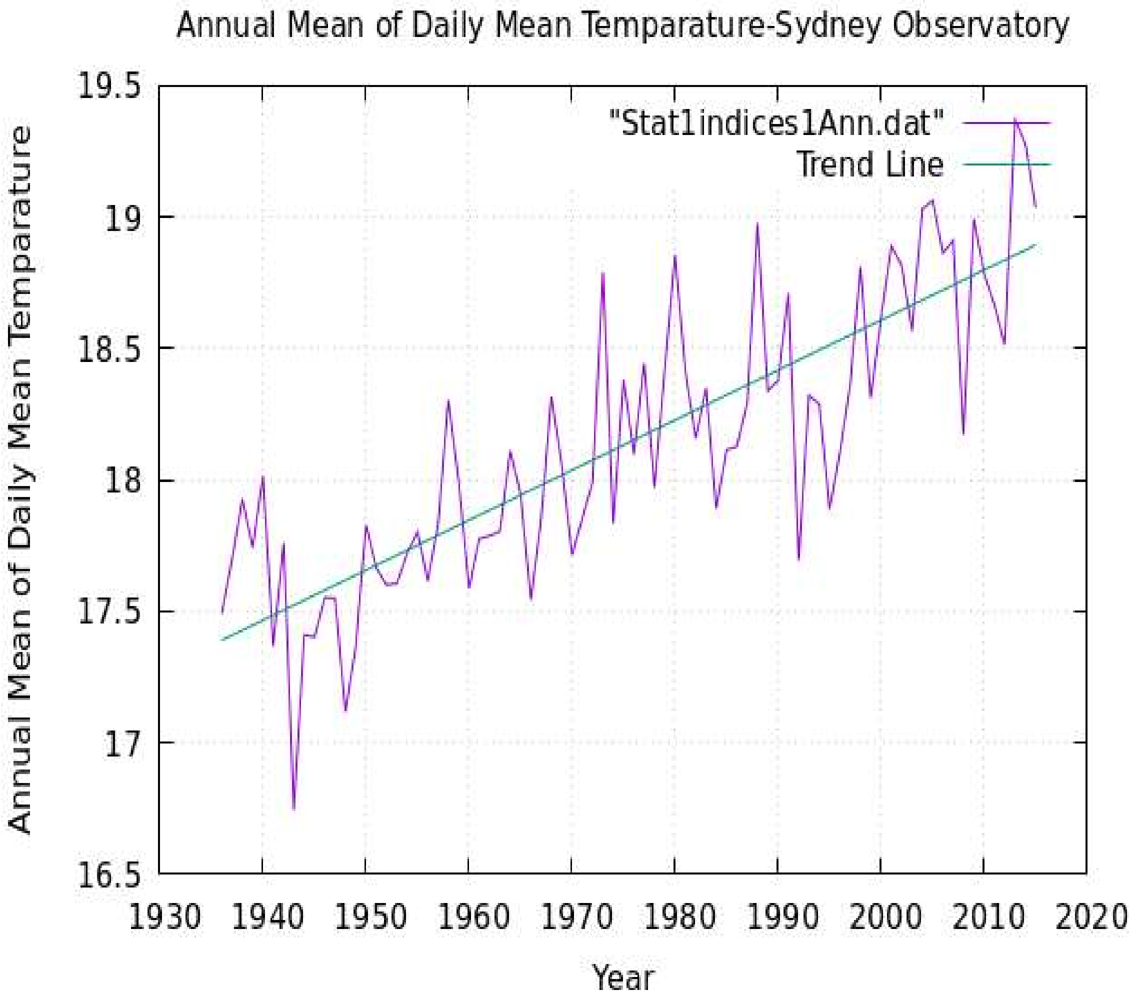

Figure 3.2 : Graph of Annual Mean Daily Mean Temperature of PRIABO Data Repository using Observation data for Daily Temperature for Sydney, Australia from 1936 to 2015.

The slope results of the Linear Trend Line of the above two Graphs of **ClimPACT2** and **PRIABO Data Repository** using Observation data for Daily Temperature for Sydney (Australia) from 1936 to 2015 are shown in Table 3.2.

Table 3.2: Slope Results of Linear Trend Line of Graphs of Annual Mean Daily Mean Temperature of ClimPACT2 and PRIABO Data Repository using Observation data for Daily Temperature for Sydney, Australia from 1936 to 2015.

-------------------------------------- Slope --------------------------------------			
ClimPACT2		**PRIABO Data Repository**	
Slope	**Slope Error**	**Slope**	**Slope Error**
0.019	0.001	0.0190303	0.001492

As stated earlier, the selected Core Index is the Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) whose numerical results are shown in Table 3.3.

Table 3.3 : Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of ClimPACT2 and PRIABO Data Repository using Observation data for Daily Rainfall for Sydney, Australia from 1936 to 2015.

CLIMPACT2		**PRIABO DATA REPOSITORY**	
Year	**PRCPTOT**	**Year**	**PRCPTOT**
1936	746.9	1936	746.9
1937	1298.9	1937	1298.9
1938	974.8	1938	974.8
1939	831	1939	831
1940	980.1	1940	980.1
1941	660.9	1941	660.9
1942	1229.3	1942	1229.3
1943	1270.3	1943	1270.3
1944	763.5	1944	763.5
1945	1153.6	1945	1153.6
1946	895.1	1946	895.1
1947	1028.5	1947	1028.5
1948	965.9	1948	965.9
1949	1662.8	1949	1662.8

CLIMPACT2		PRIABO DATA REPOSITORY	
Year	**PRCPTOT**	**Year**	**PRCPTOT**
1950	2174.3	1950	2174.3
1951	1332.1	1951	1332.1
1952	1492.2	1952	1492.2
1953	1025.1	1953	1025.1
1954	1027.1	1954	1027.1
1955	1816	1955	1816
1956	1691.7	1956	1691.7
1957	665.3	1957	665.3
1958	1481.7	1958	1481.7
1959	1491.9	1959	1491.9
1960	1276.2	1960	1276.2
1961	1426.9	1961	1426.9
1962	1123.7	1962	1123.7
1963	2015.7	1963	2015.7
1964	1085.4	1964	1085.4
1965	900.6	1965	900.6
1966	1212.1	1966	1212.1
1967	1323.4	1967	1323.4
1968	607.9	1968	607.9
1969	1429.8	1969	1429.8
1970	1092.2	1970	1092.2
1971	1090.6	1971	1090.6
1972	1257.9	1972	1257.9
1973	1478.8	1973	1478.8
1974	1584.2	1974	1584.2
1975	1277.6	1975	1277.6
1976	1769.2	1976	1769.2
1977	907.4	1977	907.4
1978	1486	1978	1486

CLIMPACT2		PRIABO DATA REPOSITORY	
Year	**PRCPTOT**	**Year**	**PRCPTOT**
1979	793.6	1979	793.6
1980	723	1980	723
1981	1025	1981	1025
1982	820.7	1982	820.7
1983	1330.9	1983	1330.9
1984	1785.6	1984	1785.6
1985	1201.8	1985	1201.8
1986	1210.7	1986	1210.7
1987	1294.8	1987	1294.8
1988	1845.2	1988	1845.2
1989	1509.6	1989	1509.6
1990	1956.8	1990	1956.8
1991	1104.4	1991	1104.4
1992	1263.6	1992	1263.6
1993	896	1993	896
1994	895.8	1994	895.8
1995	1215.2	1995	1215.2
1996	1132	1996	1132
1997	1064.1	1997	1064.1
1998	1632.6	1998	1632.6
1999	1450.8	1999	1450.8
2000	807.6	2000	807.6
2001	1344.8	2001	1344.8
2002	845.2	2002	845.2
2003	1188.2	2003	1188.2
2004	982.6	2004	982.6
2005	802.4	2005	802.4
2006	975.2	2006	975.2
2007	1485	2007	1485

CLIMPACT2		PRIABO DATA REPOSITORY	
Year	PRCPTOT	Year	PRCPTOT
2008	1069.2	2008	1069.2
2009	942.2	2009	942.2
2010	1136.8	2010	1136.8
2011	1343.4	2011	1343.4
2012	1197	2012	1197
2013	1331.2	2013	1331.2
2014	884.2	2014	884.2
2015	1317	2015	1317

Figures 3.3 and 3.4 show the profiles of the Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) using Observation data for Daily Rainfall for Sydney, Australia from 1936 to 2015.

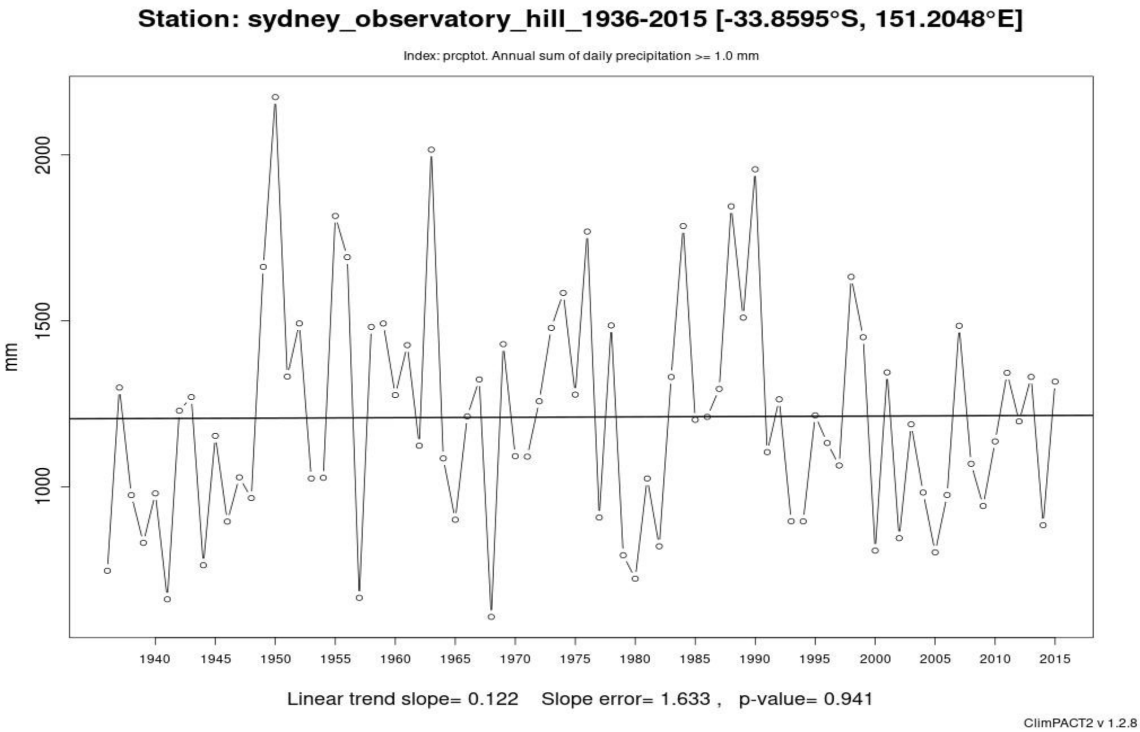

Figure 3.3 : Graph of Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of ClimPACT2 using Observation data for Daily Rainfall for Sydney, Australia from 1936 to 2015.

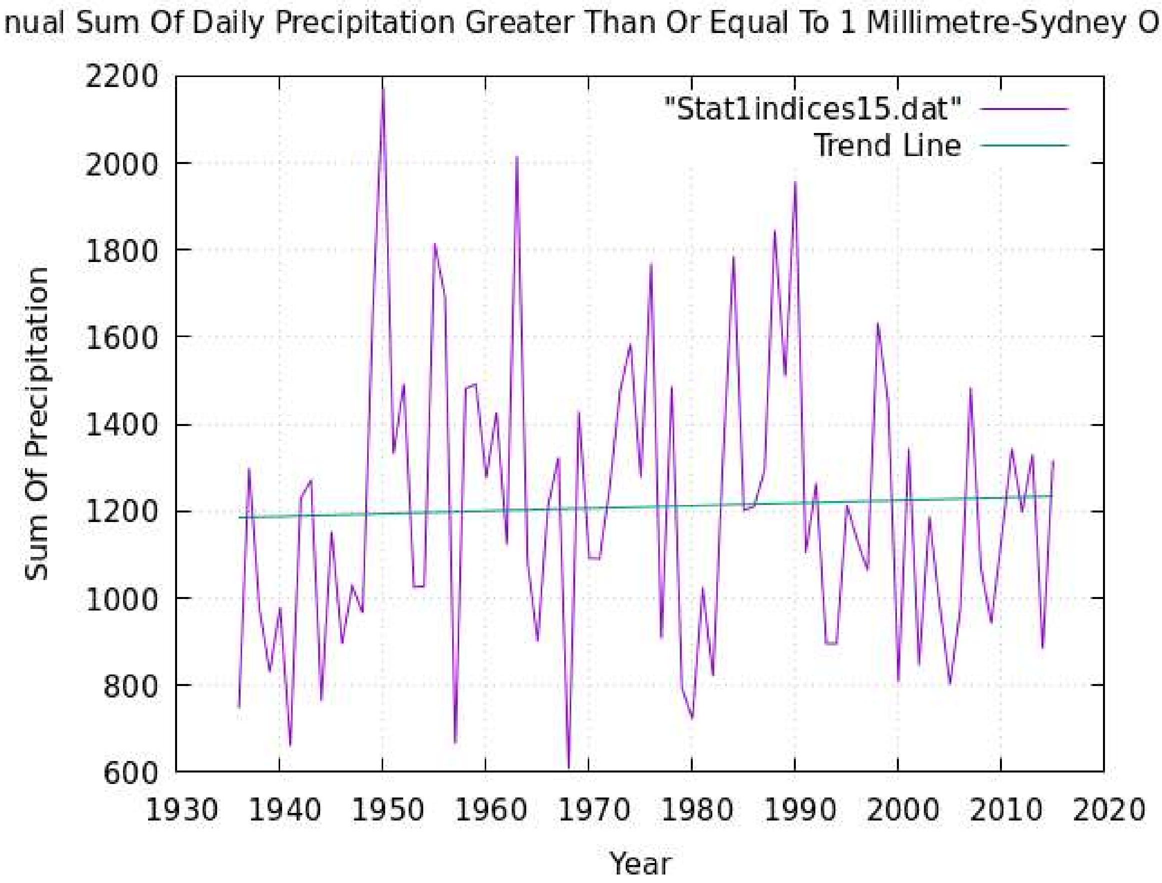

Figure 3.4 : Graph of Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of PRIABO Data Repository using Observation data for Daily Rainfall for Sydney, Australia from 1936 to 2015.

The slope results of the Linear Trend Line of the graph of the Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of **ClimPACT2** and **PRIABO Data Repository** using Observation data for Daily Rainfall for Sydney (Australia) from 1936 to 2015 are shown in Table 3.4.

Table 3.4 : Slope Results of Linear Trend Line of Graph of Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of ClimPACT2 and PRIABO Data Repository using Observation data for Daily Rainfall for Sydney, Australia from 1936 to 2015.

--------------------------------------- Slope ---------------------------------------			
ClimPACT2		**PRIABO Data Repository**	
Slope	**Slope Error**	**Slope**	**Slope Error**
0.122	1.633	**0.611958**	1.634

The results of the Annual Mean Daily Mean Temperature (TM) of the station of ANAM in Ouagadougou (Burkina Faso) have also been computed as shown in Table 3.5.

Table 3.5 : Annual Mean Daily Mean Temperature (TM) of ClimPACT2 and PRIABO Data Repository using Observation data for Daily Temperature for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

CLIMPACT2		**PRIABO DATA REPOSITORY**	
Year	**Annual Mean TM**	**Year**	**Annual Mean TM**
1970	28.34137	1970	28.34137
1971	28.19973	1971	28.19973
1972	27.9179	1972	27.9179
1973	28.50658	1973	28.50658
1974	27.90192	1974	27.90192
1975	27.85096	1975	27.85096
1976	27.58402	1976	27.58402
1977	**-99.9**	**1977**	**28.14164**
1978	28.36548	1978	28.36548
1979	**-99.9**	**1979**	**28.42356**
1980	**-99.9**	**1980**	**28.70836**
1981	28.50589	1981	28.50589
1982	28.04192	1982	28.04192
1983	28.43384	1983	28.43384

CLIMPACT2		PRIABO DATA REPOSITORY	
Year	**Annual Mean TM**	**Year**	**Annual Mean TM**
1984	28.45027	1984	28.45027
1985	28.40356	1985	28.40356
1986	28.0789	1986	28.0789
1987	28.84931	1987	28.84931
1988	28.12951	1988	28.12951
1989	28.17808	1989	28.17808
1990	28.79274	1990	28.79274
1991	28.36904	1991	28.36904
1992	28.01393	1992	28.01393
1993	28.85479	1993	28.85479
1994	28.36986	1994	28.36986
1995	28.68151	1995	28.68151
1996	29.06735	1996	29.06735
1997	28.91863	1997	28.91863
1998	29.24685	1998	29.24685
1999	28.71082	1999	28.71082
2000	28.86257	2000	28.86257
2001	28.98822	2001	28.98822
2002	29.16082	2002	29.16082
2003	29.05055	2003	29.05055
2004	29.18702	2004	29.18702
2005	29.48877	2005	29.48877
2006	29.2211	2006	29.2211
2007	29.15452	2007	29.15452
2008	28.51967	2008	28.51967
2009	29.41753	2009	29.41753
2010	29.2674	2010	29.2674

Figures 3.5 and 3.6 show the profiles of the Annual Mean Daily Mean Temperature (TM) using Observation data for Daily Temperature for ANAM, Ouagadougou (Burkina Faso) from 1970 to 2010.

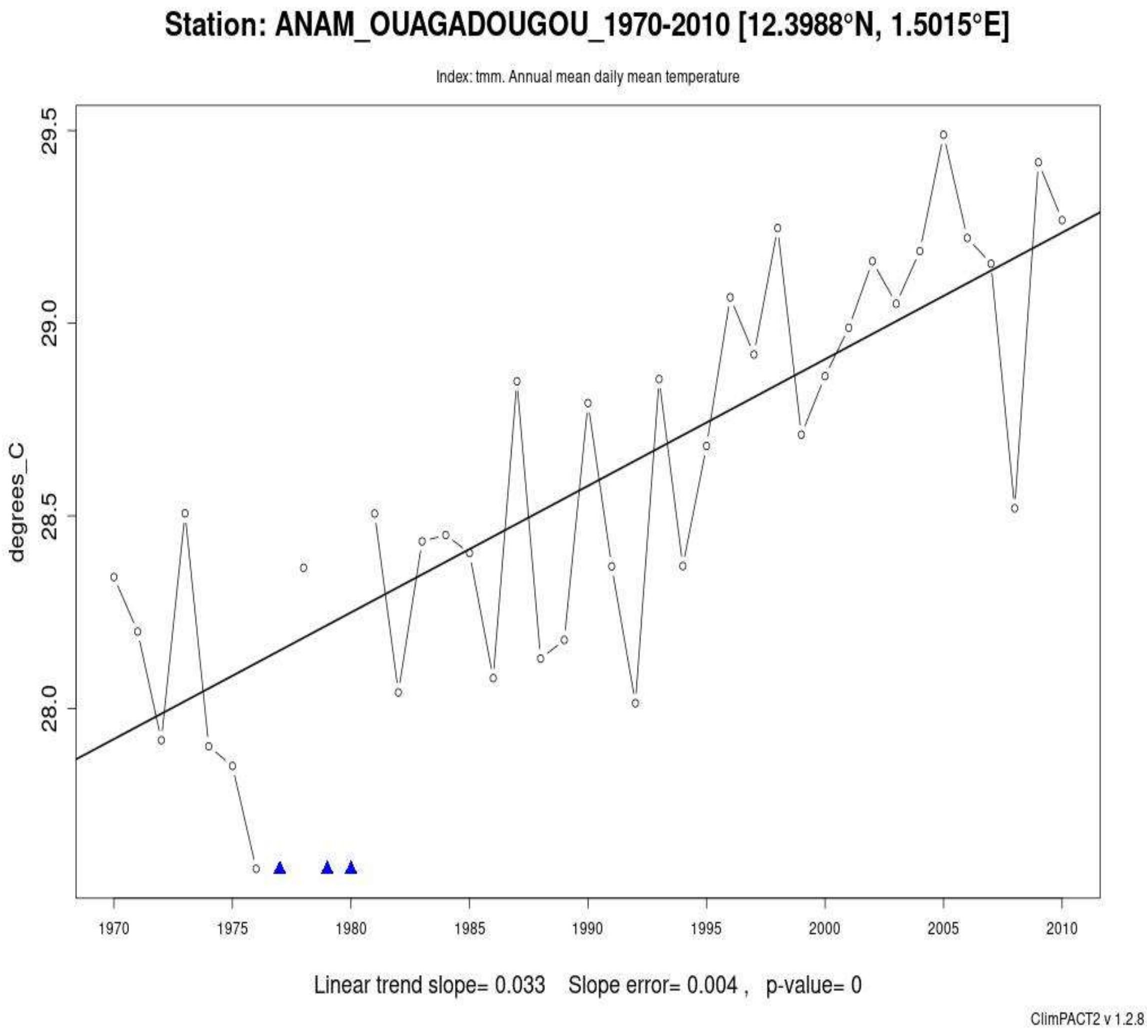

Figure 3.5 : Graph of Annual Mean Daily Mean Temperature (TM) of ClimPACT2 using Observation data for Daily Temperature for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

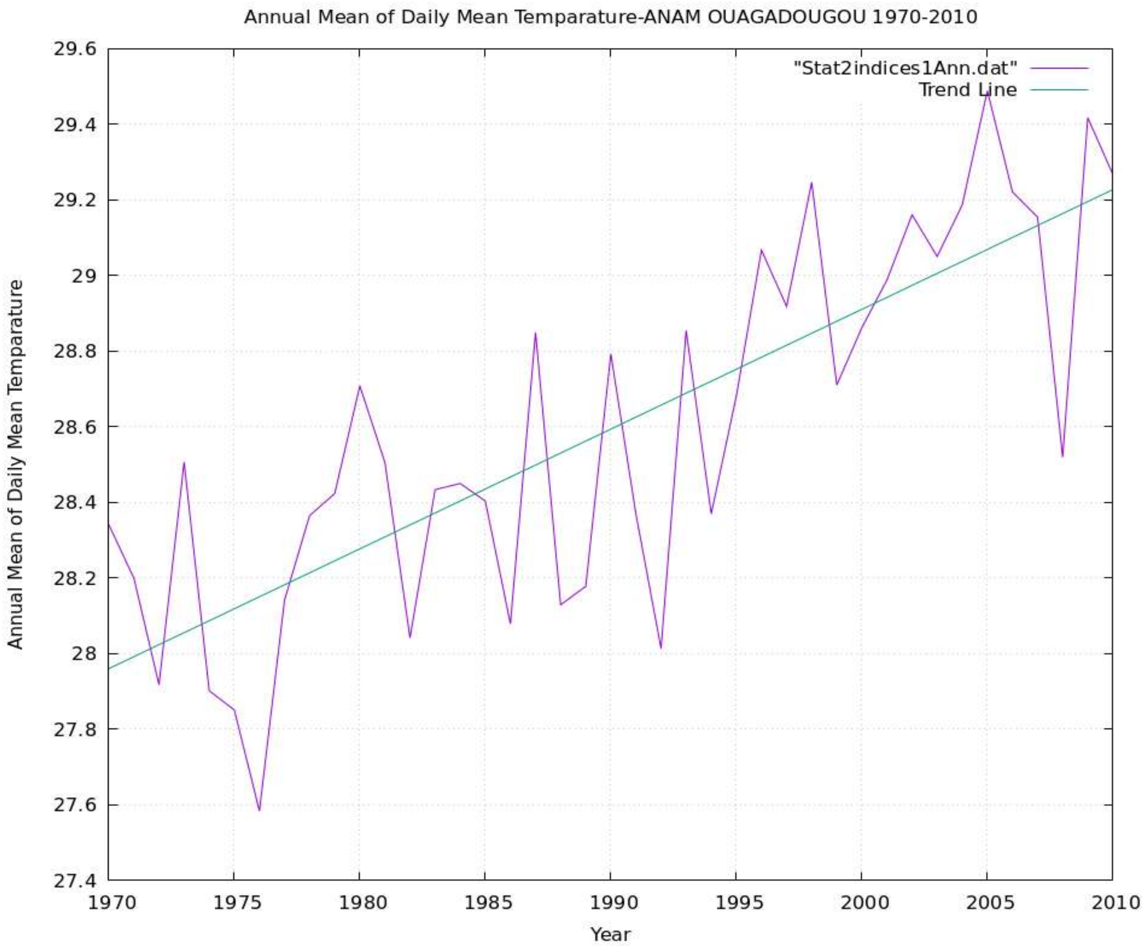

Figure 3.6 : Graph of Annual Mean Daily Mean Temperature (TM) of PRIABO Data Repository using Observation data for Daily Temperature for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

The slope results of the Linear Trend Line of Graphs of Annual Mean Daily Mean Temperature (TM) of **ClimPACT2** and **PRIABO Data Repository** using Observation data for Daily Temperature for ANAM, Ouagadougou (Burkina Faso) from 1970 to 2010 are shown in Table 3.6.

Table 3.6 : Slope Results of Linear Trend Line of Graphs of Annual Mean Daily Mean Temperature (TM) of ClimPACT2 and PRIABO Data Repository using Observation data for Daily Temperature for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

------------------------------------ Slope ------------------------------------			
ClimPACT2		**PRIABO Data Repository**	
Slope	**Slope Error**	**Slope**	**Slope Error**
0.033	0.004	0.0316803	0.003933

The numerical results of the Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of the station of ANAM in Ouagadougou (Burkina Faso) have been computed and shown in Table 3.7.

Table 3.7 : Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of ClimPACT2 and PRIABO Data Repository using Observation data for Daily Rainfall for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

CLIMPACT2		**PRIABO DATA REPOSITORY**	
Year	**PRCPTOT**	**Year**	**PRCPTOT**
1970	726	1970	726
1971	722.2	1971	722.2
1972	1055.5	1972	1055.5
1973	742.6	1973	742.6
1974	922	1974	922
1975	749.8	1975	749.8
1976	1103.4	1976	1103.4
1977	**-99.9**	**1977**	**631.3**
1978	757	1978	757
1979	725.5	1979	725.5
1980	588.7	1980	588.7
1981	709.6	1981	709.6
1982	629.3	1982	629.3
1983	670.4	1983	670.4

CLIMPACT2		PRIABO DATA REPOSITORY	
Year	**PRCPTOT**	**Year**	**PRCPTOT**
1984	565.9	1984	565.9
1985	670.2	1985	670.2
1986	790.2	1986	790.2
1987	781.1	1987	781.1
1988	729.9	1988	729.9
1989	792.5	1989	792.5
1990	673.5	1990	673.5
1991	891.9	1991	891.9
1992	692.7	1992	692.7
1993	746.3	1993	746.3
1994	719.3	1994	719.3
1995	692.7	1995	692.7
1996	672.9	1996	672.9
1997	578.1	1997	578.1
1998	663.5	1998	663.5
1999	792	1999	792
2000	587.1	2000	587.1
2001	613.7	2001	613.7
2002	652.7	2002	652.7
2003	842.1	2003	842.1
2004	769.9	2004	769.9
2005	834.1	2005	834.1
2006	592.4	2006	592.4
2007	708.6	2007	708.6
2008	759.7	2008	759.7
2009	916.7	2009	916.7
2010	834.2	2010	834.2

Figures 3.7 and 3.8 show the profiles of the Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) using Observation data for Daily Temperature for ANAM, Ouagadougou (Burkina Faso) from 1970 to 2010.

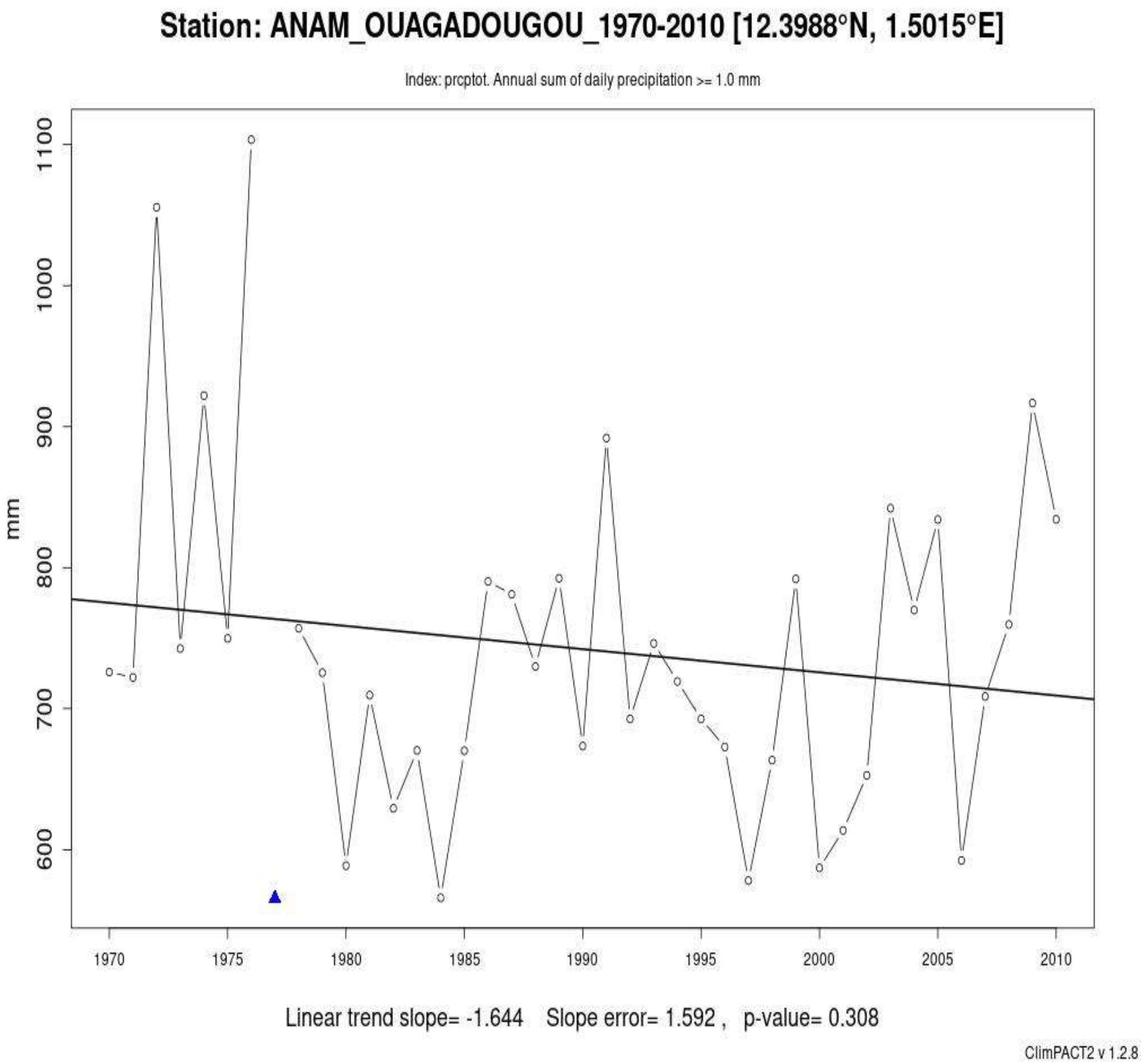

Figure 3.7 : Graph of Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of ClimPACT2 using Observation data for Daily Rainfall for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

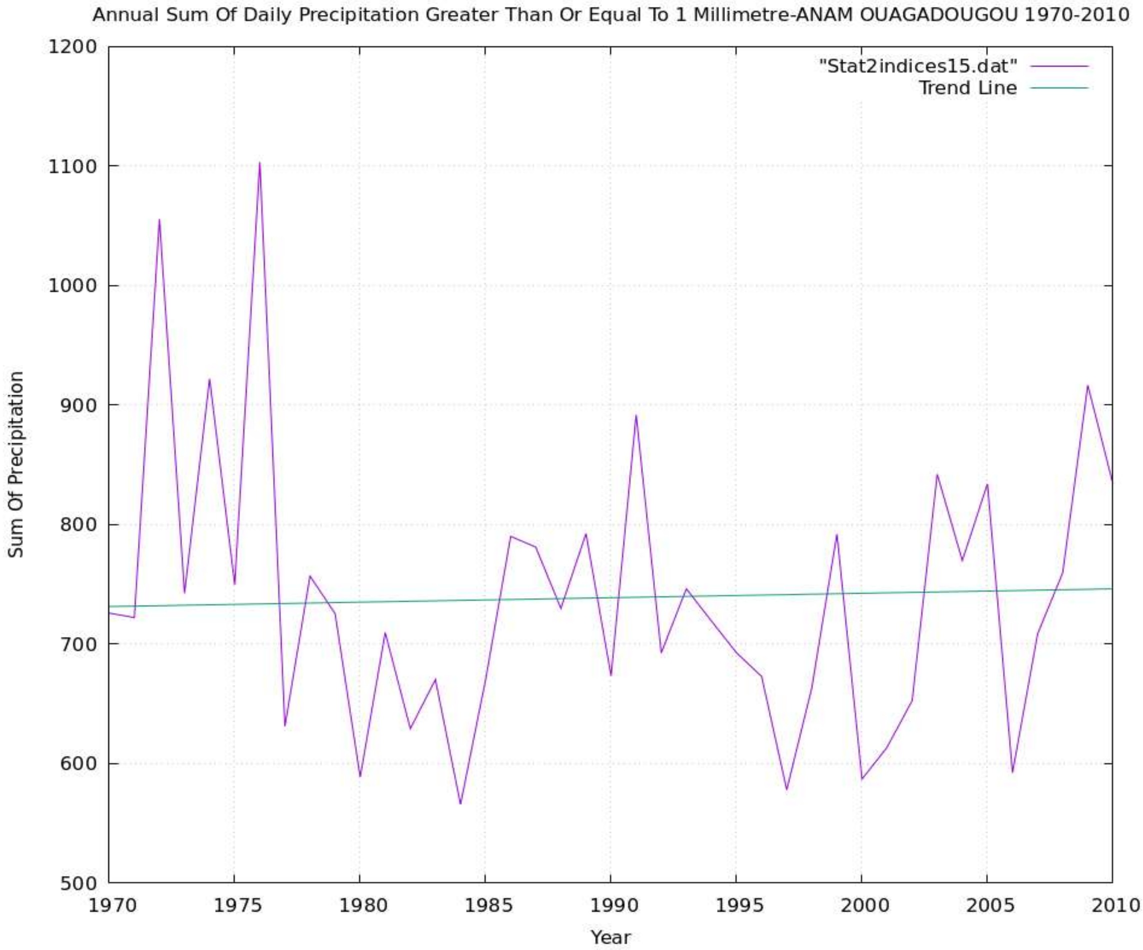

Figure 3.8 : Graph of Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of PRIABO Data Repository using Observation data for Daily Rainfall for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

Table 3.8 shows the slope results of the Linear Trend Line of Graphs of Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of **ClimPACT2** and **PRIABO Data Repository** using Observation data for Daily Temperature for ANAM, Ouagadougou (Burkina Faso) from 1970 to 2010.

Table 3.8 : Slope Results of Linear Trend Line of Graphs of Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of ClimPACT2 and PRIABO Data Repository using Observation data for Daily Rainfall for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

-------------------------------------- Slope --------------------------------------			
ClimPACT2		**PRIABO Data Repository**	
Slope	**Slope Error**	**Slope**	**Slope Error**
-1.644	**1.592**	**0.370759**	**1.595**

3.2 ANALYSIS AND DISCUSSION

From Table 3.1 which is the Annual Mean Daily Mean Temperature (TM) of ClimPACT2 and PRIABO Data Repository using Observation data for Daily Temperature for Sydney (Australia) from 1936 to 2015, it can be observed that the Annual Mean Daily Mean Temperature (TM) of ClimPACT2 for 1996 was not computed. To explain why, lets first take a look at the number of missing data in the CleanData table in the PRIABO Data Repository (Table 3.9) which is given a database name Station1 representing Sydney Observatory. The CleanData table contains all the data from 1936 to 2015 making a total of 29,220 records. In the missing data table, TX stands for maximum temperature and TN stands for minimum temperature. It is named cleaned data because of the first quality control applied to it which involves replacing all white spaces with -99.9.

Table 3.9 : Missing Data in CleanData Table from PRIABO Data Repository using Observation data for Daily Temperature for Sydney, Australia from 1936 to 2015.

Missing Data in CleanData Table from PRIABO Data Repository					
DataID	**Year**	**Month**	**Day**	**TX**	**TN**
13	1936	1	13	-99.9	-99.9
29	1936	1	29	-99.9	-99.9
619	1937	9	10	-99.9	-99.9
1111	1939	1	15	-99.9	-99.9
4694	1948	11	6	-99.9	-99.9
4695	1948	11	7	-99.9	-99.9

Missing Data in CleanData Table from PRIABO Data Repository					
DataID	**Year**	**Month**	**Day**	**TX**	**TN**
5074	1949	11	21	-99.9	-99.9
7860	1957	7	8	-99.9	-99.9
8278	1958	8	30	-99.9	-99.9
8279	1958	8	31	-99.9	-99.9
10795	1965	7	21	-99.9	-99.9
10883	1965	10	17	-99.9	-99.9
11005	1966	2	16	-99.9	-99.9
11370	1967	2	16	-99.9	-99.9
11390	1967	3	8	-99.9	-99.9
11822	1968	5	13	-99.9	-99.9
12440	1970	1	21	-99.9	-99.9
12803	1971	1	19	-99.9	-99.9
13162	1972	1	13	-99.9	-99.9
13735	1973	8	8	-99.9	-99.9
13789	1973	10	1	-99.9	-99.9
13887	1974	1	7	-99.9	-99.9
14290	1975	2	14	-99.9	-99.9
14583	1975	12	4	-99.9	-99.9
15642	1978	10	28	-99.9	-99.9
15751	1979	2	14	-99.9	-99.9
15819	1979	4	23	-99.9	-99.9
16105	1980	2	3	-99.9	-99.9
16575	1981	5	18	-99.9	-99.9
17194	1983	1	27	-99.9	-99.9
17533	**1984**	**1**	**1**	**-99.9**	**-99.9**
17587	**1984**	**2**	**24**	**-99.9**	**-99.9**
17642	**1984**	**4**	**19**	**-99.9**	**-99.9**
17658	**1984**	**5**	**5**	**-99.9**	**-99.9**
18534	1986	9	28	-99.9	-99.9

Missing Data in CleanData Table from PRIABO Data Repository					
DataID	**Year**	**Month**	**Day**	**TX**	**TN**
18574	1986	11	7	-99.9	-99.9
18688	1987	3	1	-99.9	-99.9
19086	1988	4	2	19.1	-99.9
19229	1988	8	23	-99.9	-99.9
20105	**1991**	**1**	**16**	**-99.9**	**-99.9**
20147	**1991**	**2**	**27**	**-99.9**	**-99.9**
20280	**1991**	**7**	**10**	**-99.9**	**-99.9**
20509	1992	2	24	-99.9	-99.9
20856	1993	2	5	-99.9	-99.9
20972	1993	6	1	-99.9	-99.9
21344	1994	6	8	-99.9	-99.9
21873	1995	11	19	-99.9	-99.9
22096	**1996**	**6**	**29**	**-99.9**	**7.6**
22097	**1996**	**6**	**30**	**18.4**	**-99.9**
22179	**1996**	**9**	**20**	**-99.9**	**-99.9**
22191	**1996**	**10**	**2**	**-99.9**	**10.0**
22192	**1996**	**10**	**3**	**-99.9**	**12.7**
22194	**1996**	**10**	**5**	**-99.9**	**-99.9**
22206	**1996**	**10**	**17**	**-99.9**	**-99.9**
22211	**1996**	**10**	**22**	**-99.9**	**9.8**
22212	**1996**	**10**	**23**	**-99.9**	**-99.9**
22213	**1996**	**10**	**24**	**-99.9**	**-99.9**
22214	**1996**	**10**	**25**	**-99.9**	**14.2**
24260	2002	6	2	14.3	-99.9
27021	2009	12	23	-99.9	19.7
27022	2009	12	24	-99.9	-99.9
27137	2010	4	18	25.0	-99.9
27287	2010	9	15	20.4	-99.9
28822	2014	11	28	23.3	-99.9

Other important observations are made as follows:

- From Table 3.9, it can be observed that the year 1996 has 11 records of missing data followed by the year 1984 with 4 records of missing data and 1991 with 3 records of missing data. The rest of the years in the table have either 1 or 2 missing data records out of 366 or 365 records. The surprising thing is that ClimPACT2 computed the climate indices for all these years except 1996 which is not right. Taking 11 records from 365 will give 354 records or 366 which will give 355 records. The 354 or 355 records can still be used to compute the climate index for that year and that is exactly what PRIABO Data Repository did by ignoring the 11 records and computing the climate index with the 355 or 354 records for 1996. In fact, 366 records are used for a leap year and 365 records for a non-leap year. This was equally done for 1984, 1991 and the remaining years.
- From Table 3.1, we observe that the results of both ClimPACT2 and PRIABO Data Repository are the same except for the years where ClimPACT2 didn't compute the climate index and instead replaced the value with -99.9.
- From Table 3.2, we observe that the slope and slope error of the linear trend line of both ClimPACT2 and PRIABO Data Repository are the same.
- Figure 3.1 was not used for the analysis because of the break in the graph due to the missing climate index data in 1996. This break has the power to change the direction of the linear trend line which will lead to wrong observations hence, the decision to ignore it. A similar thing occurs with the Annual Diurnal Temperature Range with the climate index for 1996 replaced with missing data due to 11 missing data records for that year in the CleanData table from the PRIABO Data Repository.
- From Figure 3.2, we observe that as we move from 1936 to 2015, the Annual Mean Daily Mean Temperature is increasing causing a change in the climate system. It also indicates that it's getting warmer hence, global warming.
- From Table 3.3, we can see that the Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of both ClimPACT2 and PRIABO Data Repository are the same.
- Figure 3.3 was not used for the analysis because of the break in the graph due to missing climate indices for reasons similar to those of Figure 3.1.
- From the Figure 3.4, it can be seen that there is a steady or slow increase in rainfall from 1936 to 2015 and this is due to the change in the climate system.
- From Table 3.4, it can be observed that the slope of PRIABO Data Repository is greater than ClimPACT2 with a margin of roughly 0.5 but the slope error is the same for both. This is very strange looking at the fact that both graphs were plotted with the same set of values. This will be investigated further as a perspective research after this as the cause for this difference can't be explained now.

- From Table 3.5, it can be observed that the Annual Mean Daily Mean Temperature (TM) of ClimPACT2 for 1977, 1979 and 1980 was not computed and to explain why lets first take a look at the number of missing data in the CleanData table (Table 3.10) in the PRIABO Data Repository which is given a database name Station2 representing ANAM, Ouagadougou, Burkina Faso. The CleanData table contains all the data from 1970 to 2010 making a total of 14,975 records.

Table 3.10 : Missing Data in CleanData Table from PRIABO Data Repository using Observation data for Daily Temperature for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

Missing Data in CleanData Table from PRIABO Data Repository					
DataID	**Year**	**Month**	**Day**	**TX**	**TN**
2749	**1977**	**7**	**11**	**-99.9**	**-99.9**
2750	**1977**	**7**	**12**	**-99.9**	**-99.9**
2751	**1977**	**7**	**13**	**-99.9**	**-99.9**
2752	**1977**	**7**	**14**	**-99.9**	**-99.9**
2753	**1977**	**7**	**15**	**-99.9**	**-99.9**
2754	**1977**	**7**	**16**	**-99.9**	**-99.9**
2755	**1977**	**7**	**17**	**-99.9**	**-99.9**
2756	**1977**	**7**	**18**	**-99.9**	**-99.9**
2757	**1977**	**7**	**19**	**-99.9**	**-99.9**
2758	**1977**	**7**	**20**	**-99.9**	**-99.9**
2899	**1977**	**12**	**8**	**-99.9**	**-99.9**
2900	**1977**	**12**	**9**	**-99.9**	**-99.9**
3612	**1979**	**11**	**21**	**-99.9**	**-99.9**
3613	**1979**	**11**	**22**	**-99.9**	**-99.9**
3614	**1979**	**11**	**23**	**-99.9**	**-99.9**
3615	**1979**	**11**	**24**	**-99.9**	**-99.9**
3616	**1979**	**11**	**25**	**-99.9**	**-99.9**
3617	**1979**	**11**	**26**	**-99.9**	**-99.9**
3618	**1979**	**11**	**27**	**-99.9**	**-99.9**

Missing Data in CleanData Table from PRIABO Data Repository					
DataID	**Year**	**Month**	**Day**	**TX**	**TN**
3619	1979	11	28	-99.9	-99.9
3620	1979	11	29	-99.9	-99.9
3621	1979	11	30	-99.9	-99.9
3622	1979	12	1	-99.9	-99.9
3623	1979	12	2	-99.9	-99.9
3624	1979	12	3	-99.9	-99.9
3625	1979	12	4	-99.9	-99.9
3626	1979	12	5	-99.9	-99.9
3627	1979	12	6	-99.9	-99.9
3628	1979	12	7	-99.9	-99.9
3659	1980	1	7	-99.9	-99.9
3771	1980	4	28	-99.9	-99.9
3772	1980	4	29	-99.9	-99.9
3773	1980	4	30	-99.9	-99.9
3774	1980	5	1	-99.9	-99.9
3775	1980	5	2	-99.9	-99.9
3776	1980	5	3	-99.9	-99.9
3777	1980	5	4	-99.9	-99.9
3778	1980	5	5	-99.9	-99.9
3779	1980	5	6	-99.9	-99.9
3780	1980	5	7	-99.9	-99.9
3781	1980	5	8	-99.9	-99.9
3782	1980	5	9	-99.9	-99.9
3783	1980	5	10	38.6	-99.9
3870	1980	8	5	-99.9	-99.9
3871	1980	8	6	-99.9	-99.9
3872	1980	8	7	-99.9	-99.9
3873	1980	8	8	32.2	-99.9
3940	1980	10	14	37.0	-99.9

Missing Data in CleanData Table from PRIABO Data Repository					
DataID	**Year**	**Month**	**Day**	**TX**	**TN**
4970	1983	8	10	-99.9	23.4
4973	1983	8	13	-99.9	23.5
4979	1983	8	19	-99.9	23.7
6302	1987	4	3	35.6	-99.9

- From Table 3.10, it can be observed that the year 1980 has 19 records of missing data followed by the year 1979 with 17 records of missing data and 1977 with 12 records of missing data. The rest of the years in the table have either 3 or 1 missing data records out of the 366 or 365 records. The surprising thing is that ClimPACT2 computed the climate indices for all these years except 1980, 1979 and 1977 which is debatable but normal. Taking 19 records from 365 will give 346 records or from 366 which will give 347 records. The 346 or 347 records can still be used to compute the climate index for that year and that is exactly what PRIABO Data Repository did by ignoring the 19 records and computing the climate index with the 346 or 347 records for 1980. This was equally done for 1979, 1977 and the remaining years.
- It can also be seen in Table 3.5 that the results of both ClimPACT2 and PRIABO Data Repository are the same except for the years where ClimPACT2 didn't compute the climate index and instead replaced the value with -99.9.
- It is observed from Table 3.6 that the slope and slope error of the linear trend line of both ClimPACT2 and PRIABO Data Repository are the same.
- Figure 3.5 was not used for the analysis due to the break in the graph due to the missing climate indices data in 1980, 1979 and 1977. This break has the power to change the direction of the linear trend line which will lead to wrong observation hence, the decision to ignore it.
- It can be observed from Figure 3.6 that as we move from 1970 to 2010, the Annual Mean Daily Mean Temperature is also increasing causing a change in the climate system. It also indicates that it's getting warmer hence, global warming.
- From Table 3.7, we can see that the Annual Sum of Daily Precipitation greater than 1mm (PRCPTOT) of both ClimPACT2 and PRIABO Data Repository are the

same except for the year 1977 for which ClimPACT2 didn't compute the climate index. To explain the cause of this missing data by ClimPACT2 in the year 1977, we looked into the CleanData table for station2 database (Table 3.11) to check the number of missing data records. PR in the table below stands for Precipitation.

Table 3.11 : Missing Data in CleanData Table from PRIABO Data Repository using Observation data for Daily Rainfall for ANAM, Ouagadougou, Burkina Faso from 1970 to 2010.

Missing Data in CleanData Table from PRIABO Data Repository				
DataID	**Year**	**Month**	**Day**	**PR**
2749	1977	7	11	-99.9
2750	1977	7	12	-99.9
2751	1977	7	13	-99.9
2752	1977	7	14	-99.9
2753	1977	7	15	-99.9
2754	1977	7	16	-99.9
2755	1977	7	17	-99.9
2756	1977	7	18	-99.9
2757	1977	7	19	-99.9
2758	1977	7	20	-99.9

- From Table 3.11, it can be observed that the year 1977 has 10 records of missing data out of the 366 or 365 records. This is why ClimPACT2 didn't compute the climate index for that year. Taking 10 records from 365 or 366 will give 355 or 356 records respectively. The 355 or 356 records can still be used to compute the climate index for that year and that is exactly what PRIABO Data Repository did by ignoring the 10 records and computing the climate index.
- Figure 3.7 was not used because of the break in the graph due to the missing data in 1977 and this turned the trend line showing a steady decrease in precipitation from 1970 to 2010 which is false.

- From Figure 3.8, it can be seen that there is a steady or slow increase in rainfall from 1970 to 2010 and this is due to the change in the climate system.
- From Table 3.8, it can be seen that the slope of PRIABO Data Repository is greater than that of ClimPACT2 with a margin of roughly 3.0 but the slope error is the same for both. This is not surprising at all due to the fact that ClimPACT2 replaced the value of the climate index for 1977 with missing data value of -99.9 and this broke the graph influencing the trend line to point in the opposite direction. The slope of ClimPACT2 is therefore not right even though the slope error is right.

CHAPTER FOUR

CONCLUSION AND THESIS PERSPECTIVES

4.1 CONCLUSION

In this thesis, a model Offline Data Repository (PRIABO Data Repository) for Climate Analysis has been created. It was designed in order to perform quality control, compute climate indices, give outputs like plots and store processed data (observation and satellite temperature and rainfall data) and outputs for future use. The numerical and graphical results of PRIABO Data Repository compared very well with those of ClimPACT2.

4.2 THESIS PERSPECTIVES

This research was a success as all the objectives are achieved and the hypothesis confirmed. It's time to broaden the scope as this will serve as a solid foundation. Below are the rest of the objectives to be accomplished in future research under this same topic:

- Load text format data and run quality control (MASTERS)
- Save raw and clean data in the inbuilt database (MASTERS)
- Calculate some of the climate indices (MASTERS)
- Plot the graphs for these indices (MASTERS)
- Save indices and plots in an inbuilt database
- Exit (MASTERS)
- Develop a GUI for the console application using QT in ubuntu (PHD)
- Develop a windows version of this application using Microsoft Visual Studio, MYSQL and C Sharp (PHD)
- Develop a platform independent version of this application (PHD)
- Add more functionalities to the application such as
    - ✓ Load activity data and emission factors and run quality control (PHD)
    - ✓ Save data (PHD)
    - ✓ Calculate the contribution of each greenhouse gas (PHD)
    - ✓ Calculate the total amount of greenhouse gas emitted in CO2E (PHD)
    - ✓ Load data for net present value and run quality control (PHD)

  - ✓ Save data (PHD)
  - ✓ Calculate the net present value (PHD)
  - ✓ Calculate the total net present value benefit (PHD)
  - ✓ Calculate the total net present value cost (PHD)
  - ✓ Calculate the cost benefit ratio (PHD)
  - ✓ Exit (PHD)
- Make room in the application for other data formats like csv, netcdf, grib etc. (POST-DOC)
- Compute the remaining climate indices (POST-DOC)
- Develop an online version of this application to make it accessible everywhere (POST-DOC)

BIBLIOGRAPHY

1. Digital Gaurdian, (2018). *What is a Data Repository?* https://digitalguardian.com/blog/what-data-repository
2. Copernicus. *Welcome to the Climate Data Store*. https://cds.climate.copernicus.eu/cdsapp#!/home
3. National Oceanic and Atmospheric Administration. *Climate Data Online*. https://www.ncdc.noaa.gov/cdo-web/
4. Physical Sciences Laboratory. *Welcome to the PSL Climate Data Repository*. https://psl.noaa.gov/repository
5. Roland Hudson (2018). *Big Climate Data Analytics:Effective Knowledge-Discovery from Colombia's Weather Data.* Research Gate. https://www.researchgate.net/publication/334523436
6. Stefan Bronnimann, Yuri Brugnara, Rob J. Allan, Manola Brunet, Gilbert P. Compo, Richard I. Crouthamel, Philip D. Jones, Sylvie Jourdain, Jürg Luterbacher, Peter Siegmund, Maria Antónia Valente, Clive W. Wilkinson (2018). *A Road Map to Climate Data Rescue Services*. Wiley Geoscience Data Journal. pp. 2-3. DOI: 10.1002/gdj3.56
7. Ting Yu, Nitesh V. Chawla, Simeon Simoff (2013). *Computational Intelligent Data Analysis for Sustainable Development*, Chapman and Hall/CRC. pp. 82-83.
8. Norman H. Margolus, Thomas F. Knight, Jr., Jered J. Floyed, Sam Hartman and George E. Homsy (2008). *Data repository and method for promoting network storage of data.* Red Hat Inc,pp.1 https://patentimages.storage.googleapis.com/be/a5/88/e5e7737e7678e6/US7412462.pdf
9. David Marco (2000). *Building and Managing the Metadata Repository,* Wiley Computer Publishing, pp. 146-148, ISBN: 0471355232.
10. Ofer Yaron and Avishay Gal-Yam (2012). *WISeREP-An Interactive Supernova Data Repository,* The Astronomical Society of the Pacific, Volume 124, pp. 668.
11. Universität Hamburg (2017). *Integrated Climate Data Center – ICDC*, Retrieved on 16th. February, 2021, https://icdc.cen.uni-hamburg.de/en/climate-indices.html

12. United Nations Framework Convention on Climate Change (2011), *Fact sheet: Climate change science - the status of climate change science today.* Retrieved on 18th. February, 2021, https://unfccc.int/files/press/backgrounders/application/pdf/press_factsh_science.pdf
13. Philippe Naveau, Marta Nogaj, Caspar Ammann, Pascal Yiou, Daniel Cooley, Vincent Jomelli (2005). *Statistical Methods for the Analysis of Climate Extremes*, Research Gate, https://www.researchgate.net/publication/228337559, pp. 1013. DOI: 10.1016/j.crte.2005.04.015
14. John R. Lanzante (1996). *Resistant, Robust And Non-Parametric Techniques For The Analysis Of Climate Data: Theory And Examples, Including Applications To Historical Radiosonde Station Data*, International Journal of Climatology, Volume 16, pp. 1197. https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/(SICI)1097-0088(199611)16:11%3C1197::AID-JOC89%3E3.0.CO;2-L
15. Michael Steinbach, Pang-Ning Tan, Vipin Kumar, Steven Klooster, Christopher Potter (2003). *Discovery of Climate Indices using Clustering*, ACM Digital Library, pp. 446, https://dl.acm.org/doi/abs/10.1145/956750.956801
16. Nils Chr. Stenseth, Geir Ottersen, James W. Hurrell, Atle Mysterud, Mauricio Lima, Kung-Sik Chan, Nigel G. Yoccoz, Bjorn Adlandsvik (2003). *Studying Climate Effects on Ecology through the use of Climate Indices : the North Atlantic Oscillation, El Nino Southern Oscillation and beyond*, The Royal Society, pp. 2089, DOI 10.1098/rspb.2003.2415
17. P. Yiou, E. Baert, M. F. Loutre (1996). Spectral Analysis of Climate Data, Springer Link, pp. 619, https://link.springer.com/article/10.1007%2FBF01931784
18. Matheus S. Lima-Ribeiro, Sara Varela, Javier Gonzalez-Hernandez, Guilherme De Oliveira, Jose Alexandre F. Diniz-Filho, Levi Carina Terribile (2015). *EcoClimate : A Database of Climate Data from Multiple Models for Past, Present and Future for Macro-Ecologist and Bio-Geographers*, Biodiversity Informatics, Volume 10, pp. 1-7, DOI: https://doi.org/10.17161/bi.v10i0.4955
19. Phu Nguyen, Eric J. Shearer, Hoang Tran, Mohammed Omabadi, Negin Hayatbini, Thanh Palacios, Phat Huynh, Dan Braithwaite, Garr Updegraff, Kuolin Hsu, Bob Kuligowski, Will S. Logan, Sorooshian (2019). *The CHRS Data Portal, An Easily Accessible Public Repository For Persiann Global Satellite Precipitation Data,*

Scientific Data, pp. 1-2, DOI: 10.1038/sdata.2018.296

20. Oliver Fuhrer, Carlos Osuna, Xavier Lapillonne, Tobias Gysi, Ben Cumming, Mauro Bianco, Andrea Arteaga, Thomas C. Schulthess (2014). *Towards A Performance Portable, Architecture Agnostic Implementation Strategy For Weather And Climate Models*, Volume 1, pp. 44-48, DOI: 10.14529/jsfi140103
21. O. Conrad, B. Bechtel, M. Bock, H. Dietrich, E. Fischer, L. Gerlitz, J. Wehberg, V. Wichmann, J. Bohner (2015). System for Automated Geoscientific Analyses (SAGA) v. 2.1.4, Geo-scientific Model Development, pp. 1991-1993, doi:10.5194/gmd-8-1991-2015
22. Vakhtang Shelia, James Hansen, Vaishali Sharda, Cheryl Porter, Pramod Aggarwal, Carol J. Wilkerson, Gerrit Hoogenboom (2019). A Multi-Scale And Multi-Model Gridded Framework For Forecasting Crop Production, Risk Analysis And Climate Change Impact Studies, Environmental Modeling and Software, Volume 115, pp. 144-147, https://doi.org/10.1016/j.envsoft.2019.02.006
23. Y. Yang, L.T. Wilson, J. Wang (2009). *Development Of An Automated Climatic Data Scraping, Filtering And Display System*, Computers and Electronics in Agriculture, pp. 77-81, Volume 71, https://doi.org/10.1016/j.compag.2009.12.006
24. S. Dhamodharavadhani, R. Rathipriya (2019). A Pilot Study on Climate Data Analysis Tools and Software, ResearchGate, https://www.researchgate.net/publication/316721783, pp. 1, DOI: 10.1109/GET.2016.7916863.
25. Dennis Shea (2021). Climate Data Guide, *Climate Data Analysis Tools and Methods*, https://climatedataguide.ucar.edu/climate-data-tools-and-analysis
26. Dean N. Williams, Robert S. Drach, Paul F. Dubois, Charles Doutriaux, Charles J. O'Connor, Krishna M. AchutaRao, Michael Fiorino (2001). *Climate Data Analysis Tool : An Open Software System Approach*, Program for Climate Model Diagnosis and Intercomparison (PCMDI), Lawrence Livermore National Lab (LLNL), pp. 1, www-pcmdi.llnl.gov/software

    UCRL-ID145189
27. Climpact (2012). Climpact, Retrieved on 1st. March, 2021, https://climpact-sci.org/
28. Climdex (2012). Climdex, Retrieved on 2nd. March, 2021, https://www.climdex.org/

29. Lisa Alexander, Nicholas Herold (2016). Climpact2 Manual, pp. 3.

30. Nicholas Herold (2020). ARCCSS-extremes/climpact2Retrieved on 3rd. March, 2021, https://github.com/ARCCSS-extremes/climpact2

31. Nicholas Herod, Jonathan McComb (2021). ARCCSS-extremes/climpact, Retrieved on 4th. March, 2021, https://github.com/ARCCSS-extremes/climpact

32. J. Sillmann, V. V. Kharin, X. Zhang, F. W. Zwiers, D. Bronaugh (2013). *Climate extremes indices in the CMIP5 multimodel ensemble : Part 1. Model evaluation in the present climate*, Journal of Geophysical Research: Atmospheres, Vol. 118, 1718–1719, doi:10.1002/jgrd.50203.

33. Wang, X. L., Y. Feng (2013). RHtestsV4 User Manual,
Climate Research Division, Atmospheric Science and Technology Directorate,
Science and Technology Branch, Environment Canada. 28 pp. [Available online at http://etccdi.pacificclimate.org/software.shtml]

34. Oracle (2021). *MySQL Connector/C++ 8.0 Developer Guide*. Retrieved on 20th. January, 2021 from https://dev.mysql.com/doc/connector-cpp/8.0/en/.

35. Debnath, M. (2016). *Database Programming with C/C++*. Retrieved on 11th. November, 2020 from https://www.codeguru.com/cpp/data/database-programming-with-cc.html.

TABLE OF CONTENTS