Development of a software application for statistical analysis of photovoltaic plant performance

by

Sven Fast

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Engineering at the Stellenbosch University

Supervisor: Prof. H.J. Vermeulen
Department of Electrical and Electronics Engineering

December 2015

Stellenbosch University https://scholar.sun.ac.za

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: ……………………………………

Copyright © 2015 Stellenbosch University
All rights reserved

Stellenbosch University https://scholar.sun.ac.za

Abstract

Economic and environmental concerns together with increasing fossil fuel prices are giving rise to the incorporation of increased amounts of renewable energy sources into the power grid. Furthermore, international policies such as the Kyoto Protocol and government endorsed financial support mechanisms aid significantly in making headway in this direction.

Amongst the numerous renewable energy technologies available, solar power is attracting a great deal of attention as it is a non-depletable and non-polluting source of energy. However, solar power has the drawbacks of being site dependant and intermittent in nature. For this reason, energy service providers and independent energy producers require accurate systems to forecast the power output of solar plants. Furthermore, time of use based energy generation statistics and forecasting models, i.e. with respect to the time when energy is being generated or consumed, are important in the context of small solar plants operating in conjunction with a local load. Generated energy forecasts and statistics are particularly useful in determining the return on investment of solar plants and conducting a financial analysis on feed-in tariffs and time of use tariff structures.

This project focusses on the development and software implementation of a long term forecasting methodology for the energy output of a solar plant. Forecasting models are derived using a statistical approach based on measured historical generation data and takes place in the time of use context. The project aims at determining whether it is possible to model the energy output of a solar plant, in the time of use context, with probability distributions commonly used to model solar radiation.

The implementation of the forecasting methodology includes the development of a relational database structure together with a forecasting software application. The relational database provides persistent storage for both historical generation data and time of use structure data, while the software application implements statistical theory to derive long term forecasting models.

Finally, a case study is conducted for an operational solar plant to test and evaluate the implemented forecasting methodology and software application. The case study is conducted with respect to time of use structures for seasonal and monthly datasets. It is found that the energy output of the solar plant can be successfully modelled and forecasted in the time of use context using monthly datasets. Furthermore, generation statistics are used to conduct a financial analysis on renewable energy feed-in tariffs and to determine the annual monetary savings from generated energy for the solar plant.

Stellenbosch University https://scholar.sun.ac.za

Opsomming

Ekonomiese en omgewingskwessies, tesame met toenemende fossiel brandstof pryse gee aanleiding tot die inlywing van verhoogde bedrae van hernubare energie bronne in die kragnetwerk. Internasionale beleide soos die Kyoto Protokol en regering onderskryfde finansiële steun meganismes bied aansienlik hulp in die vordering van hierdie rigting.

Onder die talle hernubare energie tegnologie tot ons biskikking, lok sonkrag 'n groot deel van die aandag, want dit is 'n onuitputbaar en nie- besoedelende bron van energie. Sonkrag het egter die nadele van gebieds afhanklikheid en hortend in natuur te wees. Om hierdie rede, benodig energie diensverskaffers en onafhanklike energie produsente akkurate stelsels om die kraglewering van sonkrag aanlegte te voorspel. Tyd van gebruik gebaseerde kragopwekking statistieke en voorspelling modelle, dws met betrekking tot die tyd wanneer energie gegenereer of verbruik word, is belangrik in die konteks van 'n klein sonkragte aanleg in samewerking met plaaslike laste. Gegenereerde energie voorspellings en statistieke is veral nuttig in die bepaling van die opbrengs op belegging van sonkrag aanlegte en die uitvoer van 'n finansiële ontleding op in - voer tariewe en tyd van gebruik tarief strukture.

Hierdie projek fokus op die ontwikkeling en sagteware implementering van 'n lang termyn vooruitskatting metode vir die energie-uitset van 'n sonkrag aanleg. Voorspellingsmodelle is afgelei deur 'n statistiese benadering wat gebaseer is op historiese data en vind plaas in die tyd van gebruik konteks. Die doel van die projek is om te bepaal of dit moontlik is om die energie-uitset van 'n sonkrag stasie te modelleer in die tyd van gebruik konteks , met waarskynlikheidsverdelings wat gebruik word om sonstraling te modelleer.

Die implementering van die vooruitskatting metode sluit in die ontwikkeling van 'n relasionele databasis struktuur tesame met 'n vooruitskatting sagteware program. Die relasionele databasis bied aanhoudende stoorplek vir beide historiese data en tyd van gebruik struktuur data, terwyl die sagteware program statistiese teorie implementer om langtermyn voorspelling modelle af te lei.

Laastens word 'n gevallestudie gedoen vir 'n operasionele sonkrag aanleg om die vooruitskatting metode en sagteware program te toets en evalueer. Die gevallestudie is uitgevoer met betrekking tot tyd van gebruik strukture vir seisoenale en maandelikse datastelle. Dit is bevind dat die energie-uitset van sonkrag aanlegte kan suksesvol gemodelleer en voorspel word in die tyd van gebruik konteks met bettrekking tot maandelikse datastelle. Verder word gegenereerde energie statistieke gebruik om 'n finansiële ontleding van hernubare energie in-voer tariewe uit te voer en om die jaarlikse monetêre besparing van gegenereerde energie vir die sonkrag aanleg te bepaal.

Stellenbosch University https://scholar.sun.ac.za

Acknowledgments

I would like to thank Prof HJ Vermeulen for his valuable contribution, guidance and efforts with this project. I would also like to thank Nelius Bekker and Adiel Jakoef for their support during this project. Finally, I would like to thank my loved ones for their constant motivation and encouragement.

Stellenbosch University https://scholar.sun.ac.za

Table of Contents

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

List of Figures

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

List of Tables

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

Stellenbosch University https://scholar.sun.ac.za

List of Abbreviations and Symbols

A	Ampere
AC	Alternating Current
CDF	Cumulative Distribution Function
COM	Common Object Model
CSV	Comma Separated Value
COV	Coefficient Of Variation
DBMS	Database Management System
DC	Direct Current
DLL	Dynamic Linked Library
ESP	Energy Service Provider
EP	Exceedance Probability
EXE	Executable
FK	Foreign Key
GUI	Graphical User Interface
IDE	Integrated Development Environment
IEP	Independent Energy Producer
kWh	Kilo Watt Hour
MA	Main Application
MPPT	Maximum Power Point Tracker
NB	Number of Bins
NERSA	National Energy Regulator of South Africa
OLE	Object Linking and Embedding
PAE	Profile Analysis Engine
PDF	Probability Density Function
PK	Primary Key
PTC	PVUSA Testing Conditions
PV	Photovoltaic
PVUSA	Photovoltaics for Scale Utility Applications
RDBMS	Relational Database Management System
REFIT	Renewable Energy Feed-in Tariffs
RMSE	Root Mean Square Error
ROI	Return On Investment
SD	Standard Deviation
SQL	Structured Query Language
STC	Standard Testing Conditions

Stellenbosch University https://scholar.sun.ac.za

TOU	Time Of Use
UML	Unified Modelling Language
V	Volt
WAMP	Windows Apache MySQL Php
ZAR	South African Rand

Stellenbosch University https://scholar.sun.ac.za

1 Project Overview

1.1 Introduction

Economic and environmental concerns [1] together with increasing fossil fuel prices [2] are giving rise to the incorporation of increased amounts of renewable energy sources into the power grid [3]. Amongst the numerous renewable energy technologies available, solar power is attracting a great deal of attention due to its potential of contributing to sustainable future energy supplies [4] [5].

The development of solar power is strongly connected to government endorsed financial support mechanisms such as capital subsidies and feed in tariffs [4]. In some countries solar power has expanded exponentially as a result of these financial support mechanisms, especially due to high feed in tariffs [6]. Furthermore, international policies such as the Kyoto Protocol aid significantly in the incorporation and development of renewable energy sources [1].

Solar power has the advantage of being a non-depletable and non-polluting source of energy [7]. However, solar power is site dependant and intermittent in nature [7] as it depends significantly on factors such as solar radiation, ambient temperature, pollution and cloud cover [8]. The intermittent nature of solar power poses a significant challenge to large scale grid integration [5] [9]. Unexpected variations in the power output of a solar plant may incur increased operational costs and jeopardise the reliability of energy supply [10].

Distributed generation [11] using solar power spread across different locations is becoming increasingly significant and is regarded as vital towards achieving carbon reduction goals [1]. The use of distributed generation reduces the need for expensive transmission systems and significantly reduces transmission losses [12]. However, finding a balance between energy generated and consumed across different locations is essential to maintaining grid stability.

The effective utilisation of solar power, while maintaining grid stability, requires the intelligent optimization and scheduling of energy generation and demand [1] [13]. For this reason, Energy Service Providers (ESPs) and Independent Energy Producers (IEPs) require accurate systems to forecast the power output of their solar plants [10]. As a result, solar power forecasting has become an active field in recent years [14] and is reputed to be very valuable [15].

Stellenbosch University https://scholar.sun.ac.za

1.2 Project Motivation

Modern prediction systems generally use numerical prediction with a forecast horizon of one to two days [5]. However, ESPs and IEPs are interested in a range of prediction horizons to manage power plants and forecast their energy production [14].

Solar power forecasting methodologies are classified into either a numerical prediction approach or a statistical approach [14]. The numerical approach incorporates predicted weather variables, such as solar radiation and temperature, together with PV power output models. The statistical approach of forecasting energy output is based on measured historical generation data and requires less input data and computational efforts [14].

Interconnecting geographically distant renewable energy sources such as solar power to a common power grid significantly stabilises the supply of energy [16]. However, finding an optimal balance and mix of geographically distant renewable energy sources requires accurate long term energy forecasts.

In June 2007 the National Energy Regulator of South Africa (NERSA) commissioned the study of Renewable Energy Feed-In Tariffs (REFITs), which culminated in the approval of REFIT guidelines in March 2009 [17]. Feed-in tariffs are the price paid by an ESP to a energy producer per kWh of renewable energy exported to the grid [4]. REFITs were set at fixed rates of South African Rand (ZAR) per kWh for each respective renewable energy technology [17].

Time Of Use (TOU) based forecasting models, i.e. forecasting models with respect to the time the energy is being consumed or generated, are particularly important in the context of an industrial consumer which also has onsite solar generation. These forecasting models are useful for applications such as the following:

- Conducting a financial analysis on REFITs and TOU tariff structures.
- Calculating the solar plant’s future savings, payback period and Return On Investment (ROI) for different TOU tariffs.

REFIT rates and TOU tariffs are subject to change and therefore affect the financial profitability of an industrial consumer which also has onsite solar generation. Increasing the REFIT rates results in an increase in financial profitability as the generated energy is sold to the ESP at a higher monetary value per kWh produced. Similarly, a decrease in TOU tariffs also results in and increase in financial profitability as the industrial consumer is using energy at a lower monetary value per kWh consumed.
As the incorporation of solar power reaches economic feasibility, REFIT rates paid to renewable energy producers are in fact being lowered [6]. This gives rise to the situation where it may be more

Stellenbosch University https://scholar.sun.ac.za

financially profitable for an industrial consumer with onsite solar generation to consume generated solar energy during expensive TOU tariffs rather than selling the energy to the ESP at lower REFIT rates. Therefore, TOU based energy generation forecasts enable an industrial consumer to determine the most financially profitable approach to using onsite generated renewable energy. This represents a major focus point for this project.

Calculating the payback period and ROI for a solar plant depends on an accurate estimate of future monetary savings from generated renewable energy. The future savings of an industrial consumer, which uses onsite solar generation to displace energy drawn from the supply grid, depends on the tariffs paid for energy by the consumer. TOU based forecasting models are therefore useful in calculating the pay-back period and ROI of a solar plant against different TOU tariffs.

1.3 Project Description

1.3.1 Overview

In view of the above considerations, this project aims to design and implement a long term forecasting methodology for the energy output of a solar plant. This methodology must consider the following:

- Be based on measured historical generation data.
- Utilise statistical theory and methods to derive long term forecasting models.
- Incorporate TOU structures such as the following:
  - TOU tariff structures.
  - Seasons and months of the year.
  - Hours of the day.
- Be supported by the development of a software package with database capabilities.

The implementation of the forecasting methodology includes the development of a relational database together with a forecasting software application. The relational database provides persistent storage for both historical generation data and TOU structures, while the software application implements statistical calculations to derive long term forecasting models.

This project aims to create long term forecasting models for a solar plant by analysing and processing historical generation data. The methodology will attempt to fit historical generation data to proposed probability distributions, commonly used to model solar radiation, by using goodness of fit tests. Therefore, the expected outputs of the forecasting methodology are probability distributions, within the TOU context, that describe the energy output of the solar plant.

Stellenbosch University https://scholar.sun.ac.za

1.3.2 Key Research Questions

The following key questions concerning the long term forecasting methodology are identified:

- Is it possible to create long term statistical forecasting models for the energy output of a solar plant by analysing historical generation data?
- Is it possible to statistically forecast the energy output of a PV system in the context of TOU structures?
- Is it possible to model the energy output of a PV system using probability distributions which are commonly used to model solar radiation?
- Is it possible to develop a database structure which can store historical generation data together with TOU structures?
    - Can the database structure incorporate changing TOU tariffs and structures?
- Can the long term forecasting methodology and database be implemented in a software application?

1.3.3 Research Objectives

The project involves the development of a database driven software application aimed at forecasting long term energy generation in the TOU context. The following research objectives have been formulated:

- Investigate long term energy output forecasting based on historical generation data sets.
- Develop a forecasting methodology that utilises statistical theory and methods.
- Research statistical theory and methods:
    - Hypothesis testing.
    - Parameter estimation.
    - Frequency distributions.
    - Goodness of fit testing.
- Investigate TOU tariff structures with focus on those available in South Africa.
- Investigate probability distributions commonly used to model solar radiation.
- Research database concepts:
    - Database models, design, packages and languages.
- Research the development of software applications:
    - Suitable software development environment.
    - Software design framework and software modelling language.
- Conducting a case study for an operational solar plant to achieve the following:
    - To test and evaluate implemented forecasting methodology and software application.

Stellenbosch University https://scholar.sun.ac.za

1.3.4 Research Tasks

The project consists of a number of components as shown in figure 1.1. These components involve the following tasks:

- Develop a relational database to store historical generation data together with TOU structures.
- Develop an analysis software application with database connectivity.
- Perform case study for an operational solar plant.
- Analyse results and derive conclusions and future recommendations.

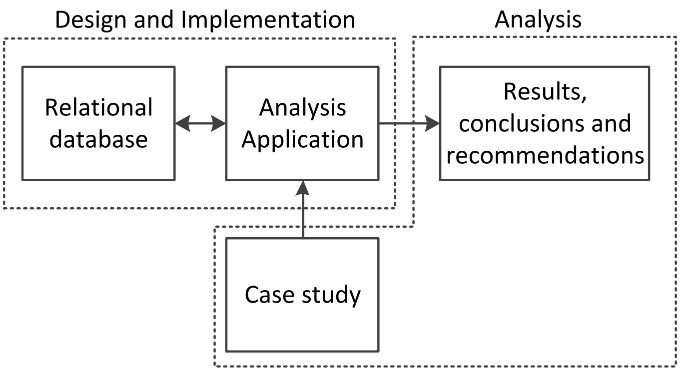

Figure 1.1: Main components of project.

The development of the software application involves the integration of all analytical components required to derive statistical models from historical generation data stored on a database. These analytical components include the following:

- Parameter estimation from historical generation data.
- Determining observed and expected frequency distributions from historical generation data and probability distributions.
- Goodness of fit testing to determine whether various probability distributions fit historical generation data.

The analysis software application incorporates six different probability distributions commonly used to model solar radiation which include the following [18] [19] [20] [21]:

- Normal distribution.
- Weibull distribution.
- Gamma distribution.
- Beta distribution.
- Logistic distribution.
- Exponential distribution.

Stellenbosch University https://scholar.sun.ac.za

A case study is conducted for an operational solar plant located in the Western Cape province of South Africa. The solar plant is implemented by an IEP as a supplementary energy source to mitigate energy consumed from the power grid. The case study involves the following tasks:

- Metering and logging of energy generation at half-hourly intervals.
- Importing and storing historical generation data in a database.
- Deriving statistical parameters and models from historical generation data with respect to the following TOU structures and datasets:
    - Seasonal datasets:
        - Half-hourly generation profile.
        - HomeFlex tariff structure.
        - MegaFlex tariff structure.
    - Monthly datasets:
        - Half-hourly generation profile.
        - HomeFlex tariff structure.
- Using the derived models and statistical parameters to forecast the generation of energy with respect to TOU structures.
- Testing and evaluating the forecasted energy against historical generation data.

1.4 Thesis Structure

This thesis document is structured into six chapters and three appendices. This structure can be summarised as follows:

- Chapter 1 presents the project overview, project motivation and project description.
- Chapter 2 presents a literature review focusing on the following:
    - Database concepts and platforms.
    - Software development platforms.
    - Software design and modelling framework.
    - Statistical inference, hypothesis testing and goodness of fit testing.
    - South African TOU tariff structures.
    - Solar power systems and models.
- Chapter 3 describes the design and implementation of a relational database.
- Chapter 4 describes the software application design and implementation along with the relevant use case models and activity diagrams.
- Chapter 5 presents the results of the case study.
- Chapter 6 presents the conclusions and recommendations for future work.

Stellenbosch University https://scholar.sun.ac.za

2 Literature Review

2.1 Overview

This literature review focuses on the development and software implementation of a methodology to forecast and model the long term energy output of a solar plant. The following aspects are investigated and discussed:

- Database concepts and platforms:
    - Relational database model.
    - Database management systems and languages.
    - Database applications.
- Software development platforms.
- Software design and modelling framework.
- Statistical inference and modelling methodologies:
    - Parameter estimation.
    - Frequency distributions.
    - Hypothesis testing and goodness of fit testing.
    - Probability distributions used to model solar radiation.
- South African time of use tariff structures.
- Solar power systems and models:
    - Solar radiation models.
    - PV system configurations.

2.2 Database System Concepts

2.2.1 Overview

Databases are designed and populated for specific purposes, with an intended group of users interested in some specific application [22]. This section briefly deals with the relational data model, database management systems and database applications.

2.2.2 Relational Data Models

A data model is defined as a collection of concepts used to describe the structure of a given database. Three of the most widely used higher-level data models are the relational data model, network data model and the hierarchal data model [22]. The network and hierarchal data models precede the relational data model and are therefore referred to as legacy database systems [22].The relational data model represents data as a collection of relations where each relation is a table of values. Each row in the relation is again a collection of related values which typically represents a real-world entity or relationship [22]. Relations (tables) consist of tuples and attributes [22] [23] where a tuple is defined as a row (record) and an attribute is defined as a column header (field). Figure 2.1 illustrates the relationship between tuples, attributes and relations.

Relation			
	Attribute 1	Attribute 2	Attribute 3
Tuple 1			
Tuple 2			
Tuple 3			
Tuple 4			

Figure 2.1: Relationship between tuples, attributes and relations.

All tuples in a relation must be distinct, meaning no tuples may have the same combination of values for all their attributes. A superkey is defined as a set of attributes which specifies a unique constraint for which no two tuples may have the same value. Every relation has a minimum of at least one superkey. It is common to designate one of the keys of a relation as the Primary Key (PK). A PK is used to identify tuples in a relation and may not be null or duplicated [22] [23].

Attributes of tuples in one relation may refer to tuples in another relation, thus linking the two relations in some way. The attribute of a relation that refers to a tuple in another relation is called a Foreign Key (FK), i.e. a FK in one relation refers to a PK in another relation. This allows for three categories of relationships to exist namely one-to-one, one-to-many and many-to-many [23]. Note that the FK in a relation must refer to the PK of a tuple that exists in another relation to maintain referential integrity [22] [23].

Below follows a brief summary of key concepts concerning a relational database [22]:

- A *table* or *relation* contains the actual data.
- A *row, record* or *tuple* presents a distinct entry in a table.
- A *field* or *attribute* presents a column in a table.
- A *value* represents the data in a field of a district row.
- A *primary key* is used to identify rows in a table.
- A *foreign key* establishes relationships between relations

Stellenbosch University https://scholar.sun.ac.za

There are several categories of constraints on the values in a database such as implicit constraints, schema based constraints, application-based constraints and data dependencies. Data dependencies are mainly used to test the goodness of the database design in a process called normalization. The normalisation process is aimed at preserving information and minimising redundancy.

2.2.3 Database Management Systems

Databases may be created, managed and maintained manually or by a group of applications specifically designed for that purpose [22]. A Database Management System (DBMS) is a set of applications tasked with constructing, defining and manipulating databases [22]. A DBMS has the following advantages [22] [23]:

- Enables the sharing and viewing of data between multiple users.
- Redundancy control.
- Restriction of unauthorised access.
- Persistent storage for program objects.
- Search techniques for efficient querying of data.

A transaction on a database by an executing program includes database operations such as inserting records, deleting records, reading records and applying updates. Transactions on a database are done by sending queries or requests to the DBMS [22].

There are several popular Relational DMBS (RDBMS) available such as Oracle, SQLServer, PostgreSQL and MySQL [22] [23]. Below follows a description of each RDBMS [24]:

- *Oracle:* Is the leading RDBMS in the commercial sector. It is scalable, reliable and runs on numerous operating systems. However, it requires a well-trained database administrator.
- *MySQL:* Is a very popular open source RDMS. It is well known for its performance and runs on numerous operating systems. Furthermore, it has a slimmer feature set for improved performance.
- *SQL Server:* Is a popular RDBMS that runs only in Windows. It delivers high performance at a low cost to the user.
- *PostgreSQL:* Is one of the most feature rich open source RDMSs which runs on numerous operating systems.

Stellenbosch University https://scholar.sun.ac.za

MySQL is chosen as a RDBMS for this project as it is open source and delivers the following [25]:

- A fast, scalable and reliable database server.
- A fast multithreaded Structured Query Language (SQL) database server developed for heavy-load production systems.
- The storing of data in separate tables as files which are optimised for speed.

The standard language for the relational database is the Structured Query Language (SQL) language, which provides a higher level declarative interface [22]. SQL has become standard language used by commercial DBMSs and all SQL standards from 1999 onward have a core specification that all SQL compliant RDBMS vendors are required to implement [22].

2.2.4 Database Applications

A server side implementation of a RDBMS is required in order to host a relational database on a computer. WAMPServer (Windows, Apache, MySQL and Php) is a server package which hosts MySQL locally on the computer it runs on. WAMPServer is selected for this project for the following reasons [26]:

- It runs in the Microsoft Windows environment.
- It is available for free.
- Incorporates MySQL.
- All server configuration settings are already set up.

To develop and test the relational database in this project, an established third party software application is required. Workbench is selected as a third party software application for the following reasons [27]:

- It has a graphical user interface for working with MySQL servers and databases.
- Creates and manages user defined connections.
- Has a built in SQL editor to execute queries on the database.
- Has a built in table editor to manage database tables.
- Allows the backup and recovery of databases.
- Supports database migration.
- It is available for free.

2.3 Software Development Platform

The Integrated Development Environment (IDE) considered for the development of the forecasting software application has the following requirements:

- Develop software applications for the Microsoft Windows environment.
- Support database connectivity.
- Support the development of Graphical User Interfaces (GUIs).
- Support modular and extensible software development.

The Embarcadero® Delphi™ IDE is chosen to develop the software application as it meets all these requirements. Delphi™ is a component based development platform which delivers fast development of GUI applications and database-driven multi-tier applications. Delphi™ is built on an excellent IDE framework with an integrated debugger and implements the Object Pascal language [28]. Furthermore, it generates standalone Windows executables which significantly simplifies application distribution and testing.

Delphi™ has built in support for several database implementations such as the Borland Database Engine, dbExpress and dbGo [28]. The dbExpress data driver architecture is employed as it provides high performance database connectivity to the following databases [29]:

- Oracle.
- SQL Server.
- MySQL.
- PostgreSQL.

The Delphi™ IDE has integrated support for creating Dynamic Linked Libraries (DLLs) [28] which are required for modular and extensible software development. DLLs are program modules that contain code which could be shared between Windows applications. DLLs are used to modularise and reuse code and to simplify the development and updating of software applications [30]. Furthermore, Delphi™ supports the use of the Common Object Model (COM) and Object Linking and Embedding (OLE) technology [30]. COM forms the basis of OLE and defines an application programming interface for communication between objects [30].

Stellenbosch University https://scholar.sun.ac.za

2.4 Software Modelling and Design

2.4.1 Unified Modelling Language

This project includes the design and implementation of a software application. Therefore, a standardised modelling language is required to visualise and document the software development. The Unified Modelling Language (UML) currently represents the "de facto" standard in software engineering [31] [32].

UML was formed through the unification of three object orientated methods namely the Booch Method, the Objectory Method and the Object Modelling Technique [31]. UML is described as a number of models that collectively describe a whole system, where each model comprises of a number of diagrams and documentation. Therefore each model is a complete description of the system from a certain perspective [32]. UML offers a framework for the integration of several types of diagrams including the following [32] [33]:

- *Use case diagrams*: Illustrate the interactions between any type of user and the system, thereby highlighting the primary functionality of a system.
- *Activity diagrams*: Illustrate the flow of tasks or activities within operations.

It is important to note that UML is purely a notation for visualizing, describing and documenting a software system and is not a design method.

2.4.2 Unified Process

The Unified Process is a framework used for design, which guides all the constituents of the design process. The Unified Process provides the inputs and outputs of each individual activity without constricting the way in which the activity should perform. The primary aim of the Unified Process is to define who does what, when do they do it and how to reach a specific goal [32]. The four key elements of the Unified Process are listed below [32]:

- It is iterative and incremental.
- It is use case driven.
- It is architecture centric.
- It acknowledges risk.

Stellenbosch University https://scholar.sun.ac.za

The Unified Process does not attempt to complete an entire design in the first attempt. It rather focuses on iterations which address different design aspects to move the design forward. This leads to a system being designed incrementally and identifying possible risks early on. The iterative approach can be divided into four basic steps [32]:

1. The first step is to plan.
2. Specify, design and then implement.
3. Integrate, test and run.
4. Finally feedback is obtained and used in the following iteration.

Use case diagrams present the interactions between the user and the system, i.e. highlighting the primary functionality of a system. Therefore, use case diagrams assist in identifying the main requirements of a system and act as a consistent thread throughout the entire development process. The roles of use case diagrams are given below [32]:

- Identify users of a system and their requirements.
- Assist in the creation and validation of system architecture.
- Direct the deployment of the system and the planning of the iterations.
- Leads to creating user documentation.
- Synchronises the content of the different models and drives traceability throughout the models.

The challenge of an iterative system development approach is that the situation could arise where a group of developers may be working on part of the implementation while another is working on part of the design. Therefore, a system architecture is required to ensure that all the components fit together seamlessly. An architecture can be thought of as a skeleton of the system and should be resistant to change and the evolving system design [32].

The Unified Process acknowledges risk in software design and development by highlighting the unknown aspects of the system being designed. This approach tries to implement and design the riskiest aspects of the system early on as it is usually the aspects which are not understood that have the biggest impact on the architecture of the final system [32].

The Unified Process is only a framework and there exists no universal process which is always applicable in a real-world project [32]. The Unified Process is flexible and extensible and it defines when activities should be performed and by which worker. Elements that do not fit the current project can be omitted and in turn additional elements can also be added to expand on some other aspect of the design [32].

Stellenbosch University https://scholar.sun.ac.za

2.4.2.1 Life Cycle Phases

The Unified Process consists of four phases namely Inception, Elaboration, Construction and Transition. The main roles and milestones of each individual phase is summarised below [32] [34]:

- *Inception*: The scope of the project is defined in the inception phase. The feasibility of the system is also established. The final output for this phase is the vision for the system including a very simplified use case model, the significant risks and a provisional architecture.
- *Elaboration*: Functional and non-functional requirements of the system are captured in this phase, as well as the creation of the final architecture to be used. The main output is the architecture, a detailed use case model and plans for the construction stage.
- *Construction*: The majority of the system is designed and implemented in this phase, as well as the final analysis of the system. Essentially, this is the phase where the system is built. The output of this phase is the implemented system along with its software, design and models. In this phase the product may not be without flaws.
- *Transition*: During this phase the system is moved to the user's environment. This includes deploying and maintaining the system. This is the final phase of a cycle therefore the output is the final release of the system.

2.4.2.2 Disciplines

One way to view disciplines in the Unified Process, is that they are the steps actually followed in the phases. Multiple disciplines can be active simultaneously in a life cycle phase. However, the emphasis at that time will be on the aim and milestones of the phase. There are five disciplines in the Unified Process as summarised below [34]:

- *Requirements:* This discipline focuses on activities allowing all functional and non- functional requirements of a system to be identified. It produces the use-case model and prototype user interface.
- *Analysis:* This discipline focuses on the restructuring of all requirements in terms of software to be created. It includes analysis of architecture and use cases.
- *Design:* This discipline focuses on the detailed design to be implemented. It includes architectural designs and design packages.
- *Implementation:* This discipline focusses on the actual coding and construction of the designed system as well as the compilation and deployment of the software. It includes testing and system integration.

Stellenbosch University https://scholar.sun.ac.za

- *Test:* This discipline focuses on activities that test the implemented software ensuring it meets the set requirements. It includes the designing, implementation and evaluation of tests.

The Unified Process is iterative and incremental and therefore all five disciplines are involved in each of the four life cycle phases. [32].

2.5 Statistical Inference

2.5.1 Overview

Statistical inference consists of methods used to draw conclusions about a population of values, based on samples or observations taken from the population [35]. It is possible to hypothesise the underlying probability distribution of an observed population of values and then to test whether the hypothesis should be rejected or accepted [35]. This section briefly deals with the following statistical theory and methods:

- Hypothesis testing.
- Parameter estimation.
- Frequency distribution and bin width estimators:
  - Sturges' rule.
  - Scott's rule.
- Goodness of fit tests:
  - Root Mean Square Error.
  - Chi-squared test.
- Probability distributions and approximations.

2.5.2 Hypothesis Testing

A statistical hypothesis is a statement about some parameter or probability distribution of a population of values [35]. The statement about the parameter or probability distribution is called the null hypothesis and is denoted by H_o. Hypothesis testing relies on using sample data from a random variable to compute a test statistic and then using the test statistic to evaluate the null hypothesis [35]. Sample data can take on any value and it is therefore necessary to define boundaries where a hypothesis about the sample data is accepted or rejected.

Stellenbosch University https://scholar.sun.ac.za

All values within the defined boundaries constitute the acceptance region and all values outside the defined boundaries constitute the critical region [35]. Values that define the boundaries are called critical values. The result of a hypothesis test is said to be significant if the calculated test statistic value falls within the critical region [35]. Therefore, the null hypothesis about a population will be rejected for an alternate hypothesis if the test statistic falls within the critical region [35].

This procedure allows for two types of erroneous conclusions to be drawn. The first type of error is rejecting the null hypothesis when it is true and the second type of error is failing to reject the null hypothesis when it is false [35]. The probability of rejecting the null hypothesis when it is true is denoted by α and is called the level of significance. The probability of failing to rejecting a hypothesis when it is false is denoted by β, and is called the β-error [35].

2.5.3 Parameter Estimation

A sample is defined as any subset of the elements of a population of measurements [35]. The sample and population mean (average) $\bar{Y}$ of a set of measurements $Y_1,\ldots, Y_N$ is given by the following relationship [36]:

$$\bar{Y} = \frac{1}{N}\sum_{i=1}^{N} Y_i \tag{2.1}$$

where Y_i denotes the *i*th measurement and *N* denotes the number of measurements. The sample variance s^2 and population variance σ^2 of a set of measurements $Y_1,\ldots, Y_N$ are given by the following relationships [36]:

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i - \bar{Y})^2 \tag{2.2}$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \bar{Y})^2 \tag{2.3}$$

where Y_i denotes the *i*th measurement and *N* denotes the number of measurements.

Stellenbosch University https://scholar.sun.ac.za

2.5.4 Frequency Distribution

The frequency distribution of a population of values is defined as an arrangement of the frequencies of observations in the population according to the values the observations take on [35]. The frequency distribution is obtained by dividing the observed data into mutually exclusive class intervals called bins [35] and counting the number of observations or occurrences that fall in each of the respective bins. The chosen bin width therefore has a significant impact on the resulting frequency distribution as small bin widths lead to under smoothing and large bin widths lead to over smoothing [37].

It is important to determine the optimal bin width which presents the essential structure of the observed data [37]. There are numerous ways of bin width selection [37] and is at the disposal of the investigator [38]. Two bin width estimators are considered in this study namely Sturges' rule and Scott's rule.

Sturges' rule is one of the earliest published rules [37] which is commonly used in practice [20] and in statistical packages [37]. Sturges' rule for the bin width *h* is given by the following relationship [37]:

$$h = \frac{Range\ of\ data\ values}{1+Log_2\, N} \tag{2.4}$$

where *N* denotes the number of observed data points. Sturges rule may lead to over-smoothed histograms especially for large data samples [37]. This could lead to a histogram lacking in important features of the data set.

Scott's rule asymptotically minimizes the integrated mean squared error [39] and is based on the optimal rate of decay of the bin width [37]. Scott's rule, which uses the Gaussian density as reference standard, represents a data based choice of bin width *h* and is given by the following relationship [37] [39]:

$$h = 3.49\sigma N^{-1/3} \tag{2.5}$$

where σ denotes an estimate of the standard deviation and *N* denotes the sample size. The number of bins, i.e. mutually exclusive class intervals, of the frequency distribution is determined by dividing the range of the observed data (maximum observed value – minimum observed value) by the determined bin width *h* and rounding the result up to the nearest integer. Once the number of bins is determined, the frequency distribution bin intervals are determined by dividing the range of the observed data points by the number of bins.

Stellenbosch University https://scholar.sun.ac.za

2.5.5 Goodness of Fit Testing

In order to determine how well a hypothesised probability distribution fits observed data, a judgment criterion is required. Two goodness of fit tests, namely the Root Mean Squared Error (RMSE) test and the Chi-squared test are considered as judgement criterion.

2.5.5.1 Root Mean Squared Error

The RMSE test is regularly employed in studies evaluating the performance of models [40]. The RMSE is given by the following relationship [19]:

$$RMSE = \left[\frac{1}{N}\sum_{i=1}^{N}(y_i - y_{ic})^2\right]^{\frac{1}{2}} \tag{2.6}$$

where y_i denotes the *i*th observed value, y_{ic} denotes the i_{th} computed (expected) value from proposed models and *N* denotes the sample size. Comparing the RMSE of different probability distribution on the same dataset indicates which one best fits the observed data.

2.5.5.2 Chi-Squared Test

In most statistical problems the distribution from which the samples are drawn is unknown. To test whether the samples were drawn from an underlying probability distribution, the Chi-squared test is commonly employed [41]. The Chi-squared test is used to determine the measure of the probability of a complex system of *N* errors occurring at least as frequently as the observed system [42].

In standard applications of the Chi-squared test the observations from a population are grouped into k mutually exclusive classes [38] and the number of observed occurrences in each class is obtained, i.e. the observed frequency distribution is determined. There is some null hypothesis that determines the probability of an observation falling in each respective class [38], i.e. the expected frequency distribution. The observed frequency distribution is then compared to expected frequency distribution and evaluated using the Chi-squared test. The Chi-squared goodness of fit test criterion is defined by the following relationship [38]:

$$X^2 = \sum \frac{(x_i - m_i)^2}{m_i} \tag{2.7}$$

where x_i denotes the *i*th observed class frequency and m_i denotes the *i*th expected class frequency.

The expected frequency m_i of a given bin or class is given by the following relationship [38]:

$$m_i = Np_i \tag{2.8}$$

where N denotes the number of observations and p_i denotes the ith expected class probability computed from the null hypothesis. If the magnitudes of the expected frequencies are too small, the test will not reflect the departure of the observed from the expected [35]. Some writers suggest that values of 1 and 2 can be regarded as the minimal value of the expected frequency in a class on the condition that most values exceed 5 [35]. Should the expected frequency of a class be too low, a class could be joined with an adjacent class [35] or the number of bins can simply be reduced until all expected frequencies are at least 1 or 2.

If the observed data follows the hypothesised probability distribution the X^2 statistic has approximately a Chi-square distribution with k-p-1 Degrees Of Freedom (DOF), where k denotes the number of exclusive classes and p denotes the number of estimated parameters of the hypothesised distribution [35] [38]. Therefore, the DOF for two parameter probability distributions such as the Normal, Weibull, Gamma and Beta are determined by subtracting 3 from the number of bins. Likewise, the DOF for single parameter probability distributions such as the Exponential and Logistic are determined by subtracting 2 from the number of bins. Note that the resulting DOF, from the number of bins and probability distribution is required to be one or more to be valid. If the resulting DOF is less than one, the result is inconclusive and could not be used to draw any statistical inference about the observed data.

$X^2_{\alpha,k-p-1}$ is defined as the percentage point of the chi-square random variable with *k-p-1* DOF, such that the probability that the X^2 statistic exceeds said value is the level of significance α [35]. Once the test statistic X^2 is calculated, it could be compared to the percentage point $X^2_{\alpha,k-p-1}$ to determine whether the null hypothesis should be accepted or rejected [35]. The hypothesised distribution (null hypothesis) is rejected when the following relationship is true [35]:

$$X^2 > X^2_{\alpha,k-p-1} \tag{2.9}$$

$X^2_{\alpha,k-p-1}$ is obtained from a percentage points table of the Chi-squared distribution for a chosen level of significance and determined DOF [35]. The percentage points for several DOF and levels of significance of the Chi-squared distribution are provided in appendix A.

Stellenbosch University https://scholar.sun.ac.za

2.5.6 Probability Distributions

The Weibull, Gamma, Normal, Logistic, Exponential and Beta [18] [19] [20] [21] probability distributions are considered to model the energy output of solar plants, as these are the most common distributions used to model solar radiation. It is assumed that the underlying probability distributions of the power output of a solar plant might follow that of solar radiation. In this section the Probability Density Function (PDF) and Cumulative Distribution Function (CDF) of each distribution is given along with their respective parameters and numerical implementation.

2.5.6.1 Weibull Probability Distribution

The Weibull PDF $f(k,c,x)$ and CDF $F(k,c,x)$ are given by the following relationships [19] [43]:

$$f(k,c,x) = \frac{k}{c}\left(\frac{x}{c}\right)^{k-1} e^{-\left(\frac{x}{c}\right)^{k}} \tag{2.10}$$

$$F(k,c,x) = \int_0^{\infty} f(x)dx = 1 - e^{-\left(\frac{x}{c}\right)^{k}} \tag{2.11}$$

where *k* denotes the shape parameter and *c* denotes the scale parameter. Parameters *c* and *k* are given by the following relationships [43] [44] [45]:

$$c = \frac{\bar{x}}{\Gamma\left(1+\frac{1}{k}\right)} \tag{2.12}$$

$$k = \left(\frac{\sigma}{\bar{x}}\right)^{-1.086} \tag{2.13}$$

where $\bar{x}$ denotes the mean, σ denotes the standard deviation and Γ denotes the Gamma function given by the following relationship [46]:

$$\Gamma(x) = \int_0^{\infty} \zeta^{x-1} e^{-\zeta} d\zeta \tag{2.14}$$

There are several methods used to calculate the Gamma function numerically with the Lanczos approximation being the simplest [47]. The Lanczos approximation for certain choices of integer *N*, rational γ and coefficients C_1, C2….C_N is given by the following relationship [47]:

$$\Gamma(z+1) = \left(z+\gamma+\frac{1}{2}\right)^{\left(z+\frac{1}{2}\right)} e^{-\left(z+\gamma+\frac{1}{2}\right)}\sqrt{2\pi}\left[C_0 + \frac{C_1}{z+1} + \frac{C_2}{z+2} + \cdots + \frac{C_N}{z+N}\right] \tag{2.15}$$

Using an N of 2 and a γ of 1.5 the Lanczos approximation has a relative error of $2.4 \cdot 10^{-4}$ everywhere in the right half of the complex plane and is given by the following relationship [48]:

$$\Gamma(z+1) = (z+2)^{\left(z+\frac{1}{2}\right)} e^{-(z+2)} \sqrt{2\pi} \left[0.999779 + \frac{1.084635}{z+1} \right] \qquad (2.16)$$

The Lanczos approximation given in equation 2.16 is very accurate and simple to implement numerically. Figures 2.2 and 2.3 illustrate the Weibull PDF and CDF. Figure 2.4 illustrates the Gamma function in the right half of the complex plane.

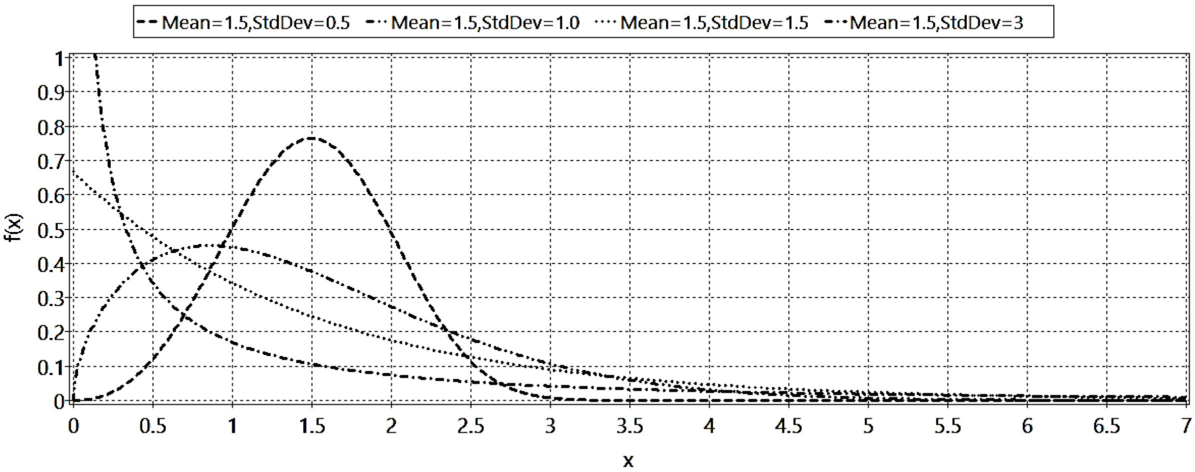

Figure 2.2: Weibull probability density function.

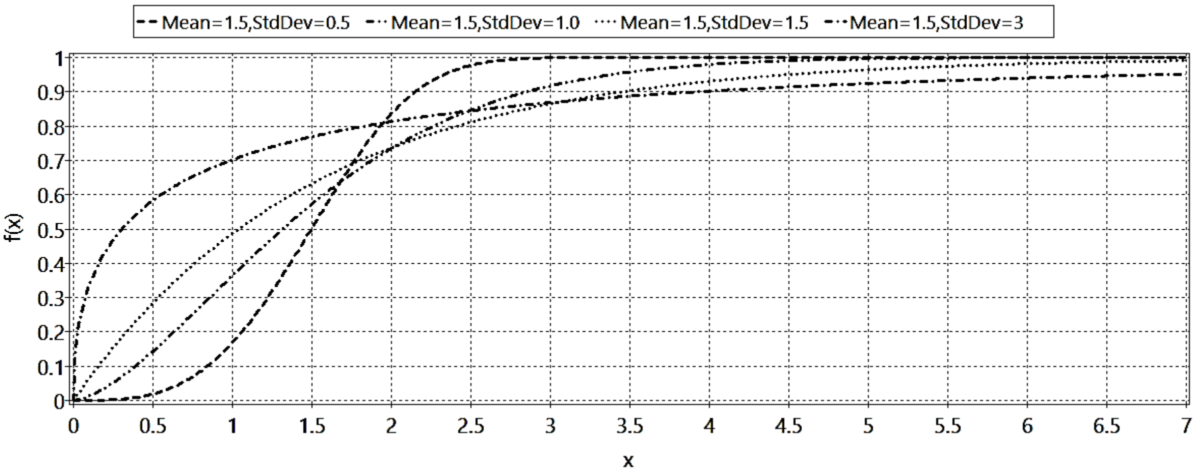

Figure 2.3: Weibull cumulative distribution function.

Stellenbosch University https://scholar.sun.ac.za

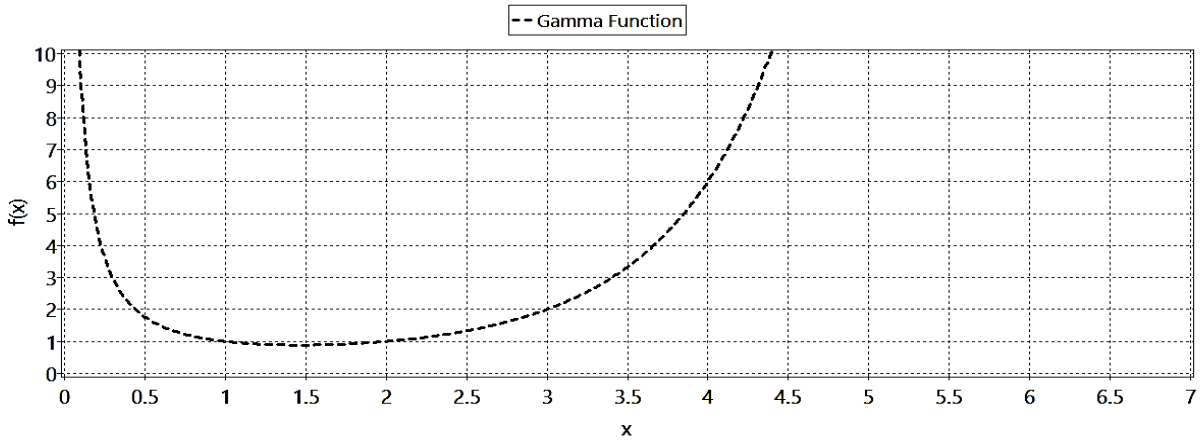

Figure 2.4: Gamma function.

2.5.6.2 Gamma Probability Distribution

The Gamma PDF $f(\alpha,\beta,x)$ and CDF $F(\alpha,\beta,x)$ are given by the following relationships [46]:

$$f(\alpha,\beta,x) = \frac{\alpha^{\beta}x^{\beta-1}e^{-\alpha x}}{\Gamma(\beta)} \tag{2.17}$$

$$F(\alpha,\beta,x) = I\left(\frac{\alpha x}{\sqrt{\beta}},\beta-1\right) = \frac{1}{\Gamma(\beta)}\int_0^{\alpha x}\zeta^{\beta-1}e^{-\zeta}d\zeta \tag{2.18}$$

where Γ denotes the Gamma function and I denotes Pearson's form of incomplete Gamma function given by the following relationship [46]:

$$I(u,p) = \frac{1}{\Gamma(p+1)}\int_0^{u\sqrt{p+1}}\zeta^{p}e^{-\zeta}\ d\zeta \tag{2.19}$$

Parameters α and β are given by the following relationships [46]:

$$\alpha = \frac{\bar{x}}{\sigma^2} \tag{2.20}$$

$$\beta = \frac{\bar{x}^2}{\sigma^2} \tag{2.21}$$

where $\bar{x}$ denotes the mean and σ denotes the standard deviation. The Gamma CDF could be implemented numerically by using the incomplete Gamma function [46] [47], which in turn can be implemented by using a combination of its series representation and continued fraction methods [47].

Stellenbosch University https://scholar.sun.ac.za

The incomplete Gamma function given by the following relationship [46] [47]:

$$P(u,p) = \frac{1}{\Gamma(u)} \int_0^p \zeta^{u-1} e^{-\zeta} d\zeta \tag{2.22}$$

where Γ denotes the Gamma function. Therefore, the Gamma CDF can be implemented by using the following relationship:

$$F(\alpha,\beta,x) = P(\beta,\alpha x) \tag{2.23}$$

Figures 2.5 and 2.6 illustrate the Gamma PDF and CDF. Figure 2.7 illustrates the incomplete Gamma function.

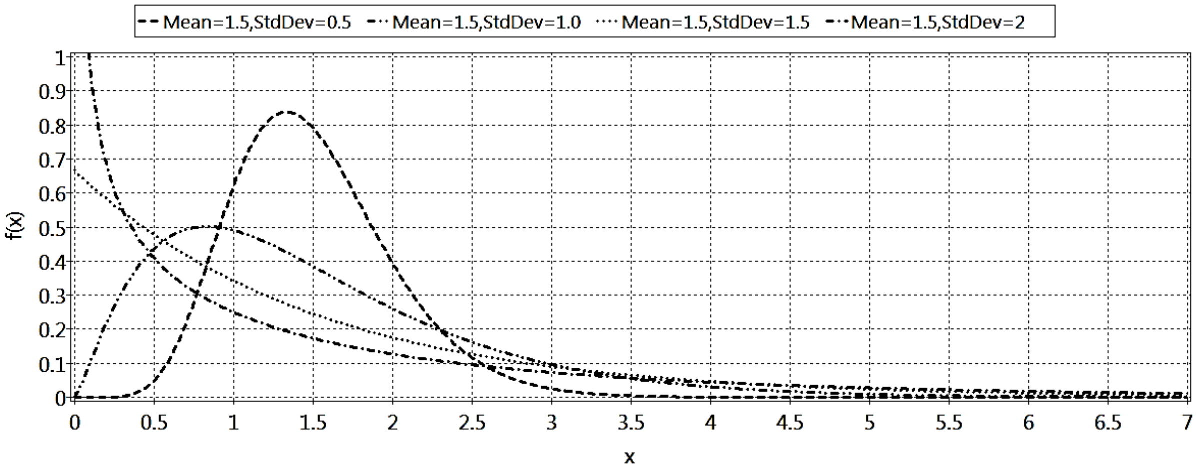

Figure 2.5: Gamma probability density function.

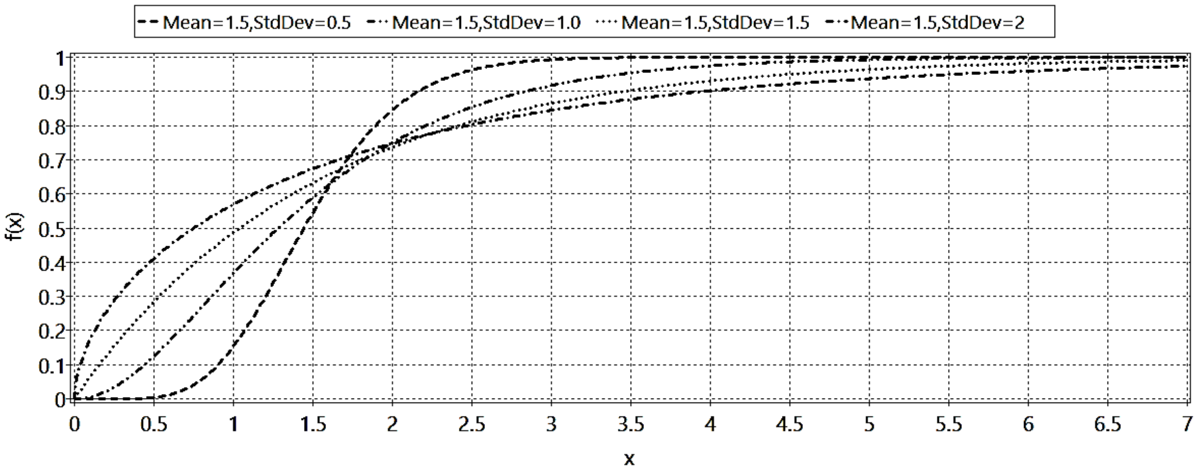

Figure 2.6: Gamma cumulative distribution function.

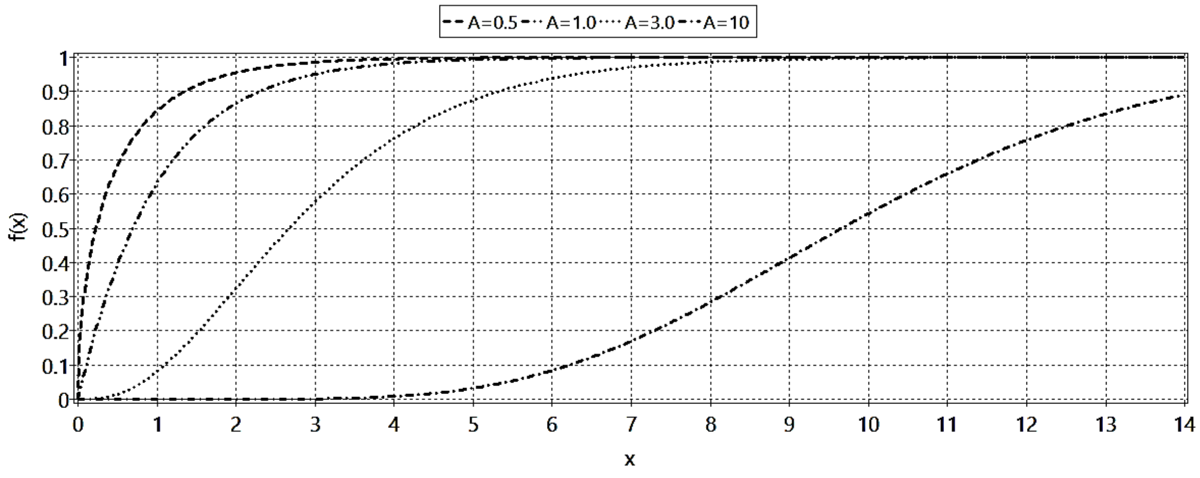

Figure 2.7: Incomplete Gama function

2.5.6.3 Normal Probability Distribution

The normal PDF $f(\bar{x}, \sigma, x)$ [15] [49] and CDF $F(\bar{x}, \sigma, x)$ [19] are given by the following relationships:

$$f(\bar{x}, \sigma, x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{\frac{-(x-\bar{x})^2}{2\sigma^2}} \tag{2.24}$$

$$F(\bar{x}, \sigma, x) = \frac{1}{2} + \frac{1}{2} Erf\left(\frac{x-\bar{x}}{\sigma\sqrt{2}}\right) \tag{2.25}$$

where $\bar{x}$ denotes the mean, σ denotes the standard deviation and Erf is the error function given by the following relationship [19] [47]:

$$\mathrm{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{2.26}$$

The error function is a special case of the incomplete Gamma function [47]. Therefore, it can be determined numerically using following relationship [47]:

$$\mathrm{Erf}(x) = P\left(\frac{1}{2}, x^2\right) \tag{2.27}$$

Figures 2.8 and 2.9 illustrate the Normal PDF and CDF. Figure 2.10 illustrates the error function.

Stellenbosch University https://scholar.sun.ac.za

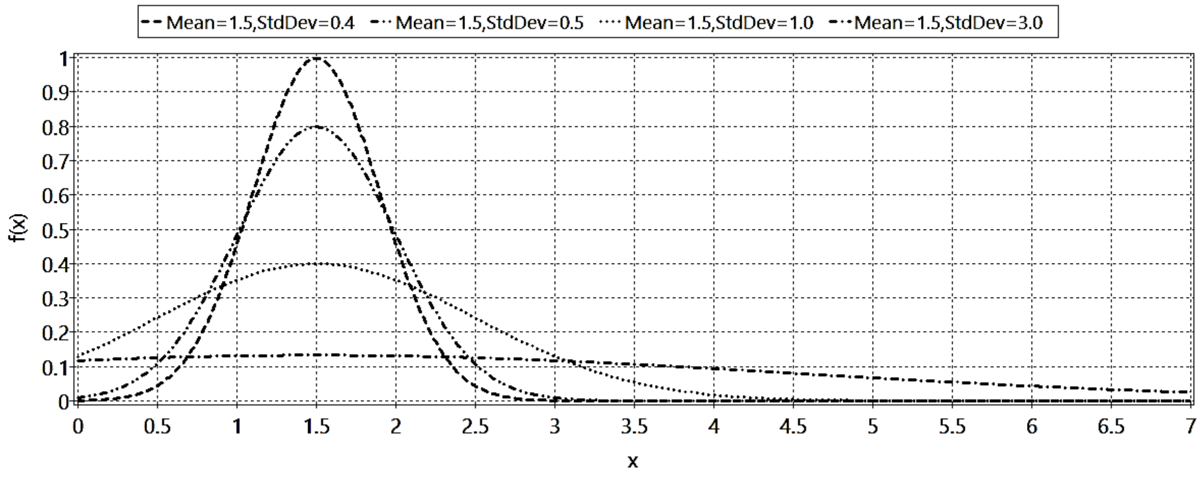

Figure 2.8: Normal probability density function.

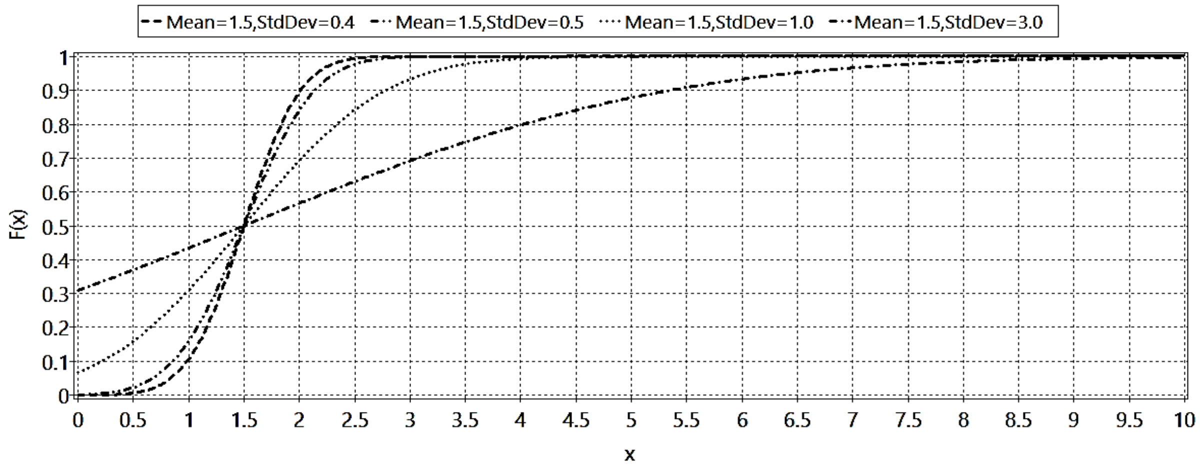

Figure 2.9: Normal cumulative distribution function.

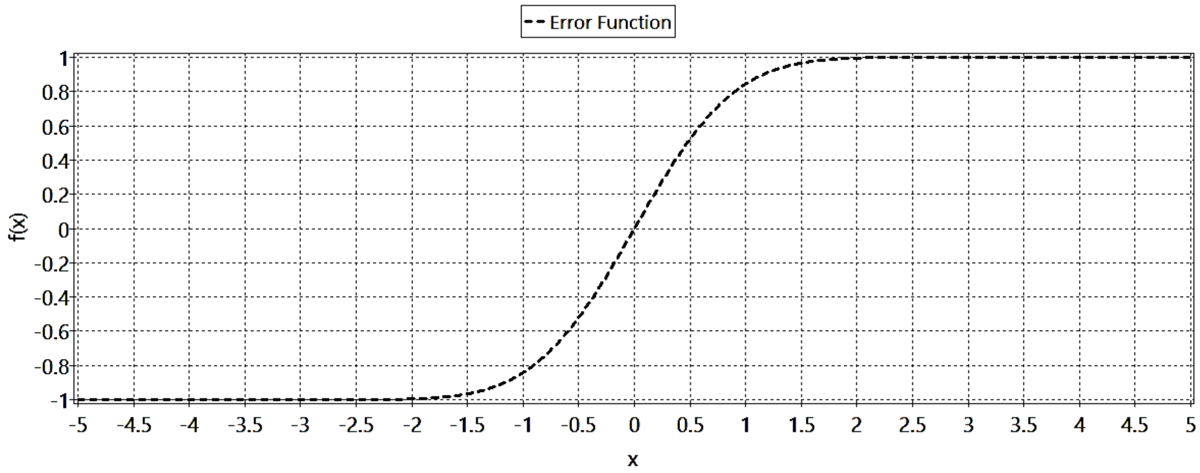

Figure 2.10: Error function.

Stellenbosch University https://scholar.sun.ac.za

2.5.6.4 Logistic Probability Distribution

The Logistic PDF $f(\bar{x},\alpha,x)$ and CDF $F(\bar{x},x)$ are given by the following relationships [19]:

$$f(\bar{x},\alpha,x) = \frac{e^{\frac{-(x-\bar{x})}{\alpha}}}{\alpha\left(1+e^{\frac{-(x-\bar{x})}{\alpha}}\right)^2} \tag{2.28}$$

$$F(\bar{x},x) = \frac{1}{1+e^{\frac{-(x-\bar{x})}{\alpha}}} \tag{2.29}$$

where $\bar{x}$ denotes the mean and α denotes the scale parameter given by the following relationship [19]:

$$\alpha = \frac{\sqrt{3}\sigma}{\pi} \tag{2.30}$$

where σ denotes the standard deviation. Figures 2.11 and 2.12 illustrate the logistic PDF and CDF.

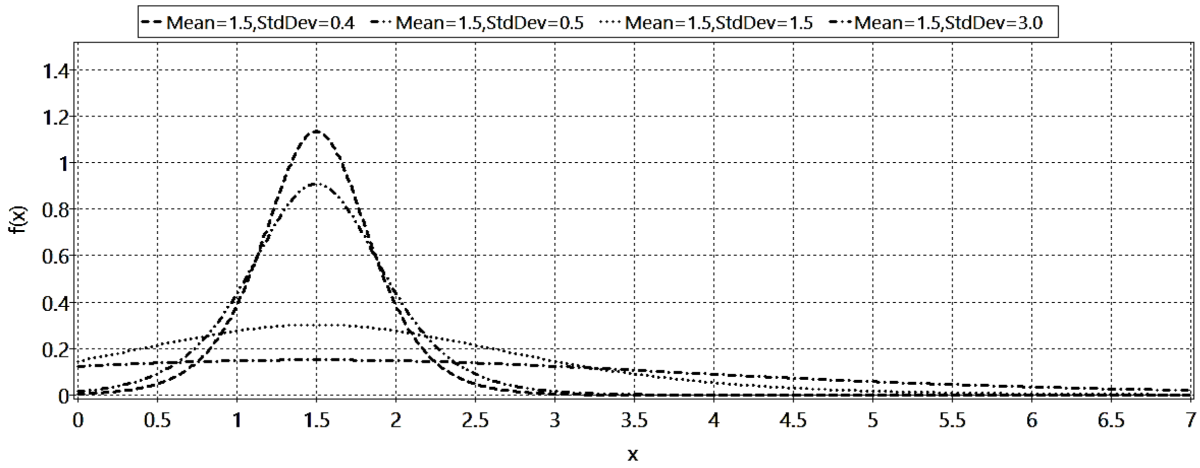

Figure 2.11: Logistic probability density function.

Stellenbosch University https://scholar.sun.ac.za

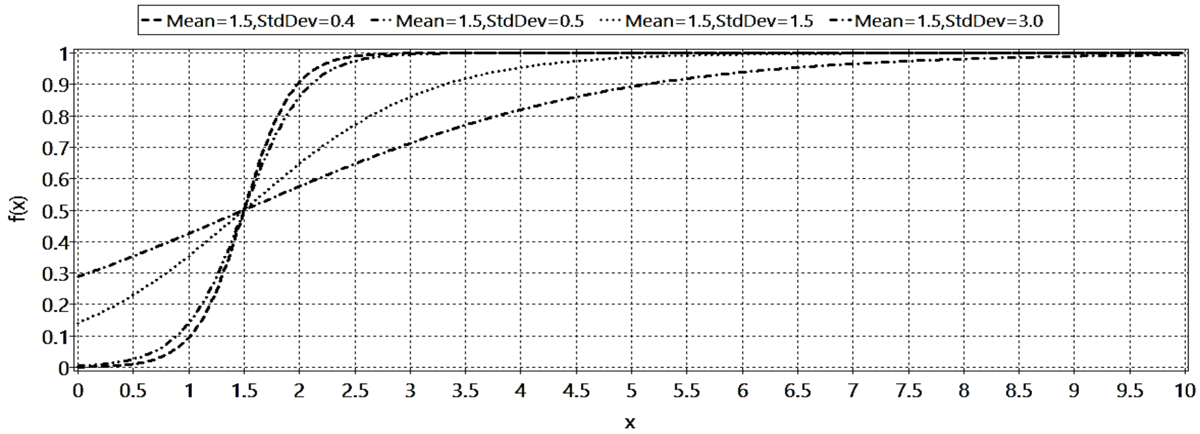

Figure 2.12: Logistic cumulative distribution function.

2.5.6.5 Exponential Probability Distribution

The Exponential PDF $f(\alpha, x)$ and CDF $F(\alpha, x)$ are given by the following relationships [46]:

$$f(\alpha, x) = \alpha e^{-\alpha x} \tag{2.31}$$

$$F(\alpha, x) = 1 - e^{-\alpha x} \tag{2.32}$$

where α denotes the rate parameter given by the following relationship [46]:

$$\alpha = \frac{1}{\bar{x}} = \frac{1}{\sigma} \tag{2.33}$$

Figures 2.13 and 2.14 illustrate the Exponential PDF and CDF respectively.

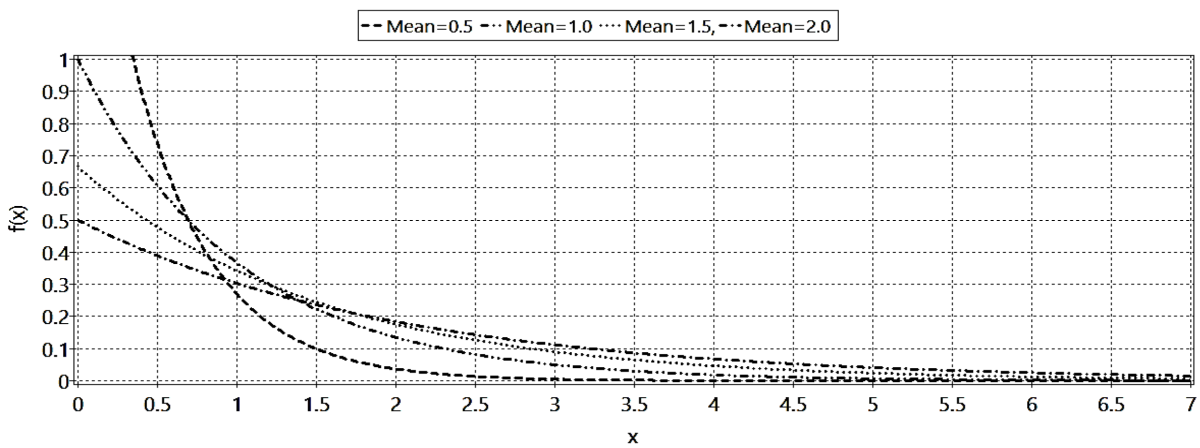

Figure 2.13: Exponential probability density function.

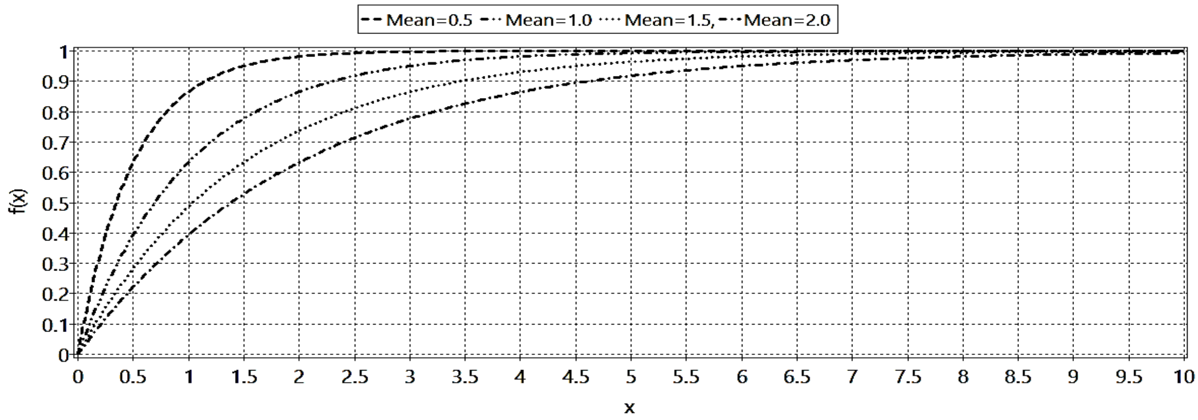

Figure 2.14: Exponential cumulative distribution function.

2.5.6.6 Beta Probability Distribution

The Beta PDF $f(\alpha,\beta,x)$ [46] [18] and CDF $F(\alpha,\beta,x)$ [46] [50] are given by the following relationships:

$$f(\alpha,\beta,x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{2.34}$$

$$F(\alpha,\beta,x) = \begin{cases} I_x(\alpha,\beta) & x < 1 \\ 1 & x \geq 1 \end{cases} \tag{2.35}$$

where I_x ( ) denotes the incomplete Beta function given by the following relationship [46] [50]:

$$I_x(\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^x \zeta^{\alpha-1}(1-\zeta)^{\beta-1} d\zeta \tag{2.36}$$

The parameters α and β are given by the following relationships respectively [18]:

$$\alpha = \left(\frac{\bar{x}^2 - \bar{x}^3}{\sigma^2}\right) - \bar{x} \tag{2.37}$$

$$\beta = \left(\frac{\bar{x}^3 - 2\bar{x}^2 + \bar{x}}{\sigma^2}\right) + \bar{x} - 1 \tag{2.38}$$

Implementing the incomplete Beta function numerically can be achieved by evaluating its continued fraction using the modified Lentz's method [47]. Figures 2.15 and 2.16 illustrate the Beta PDF and CDF respectively. Figure 2.17 illustrates the incomplete Beta function.

Stellenbosch University https://scholar.sun.ac.za

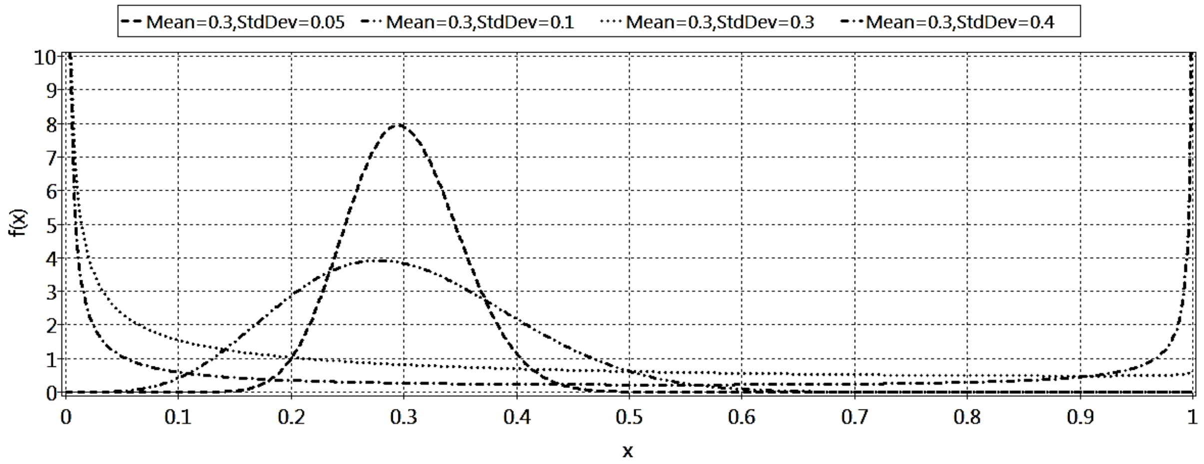

Figure 2.15: Beta probability density function.

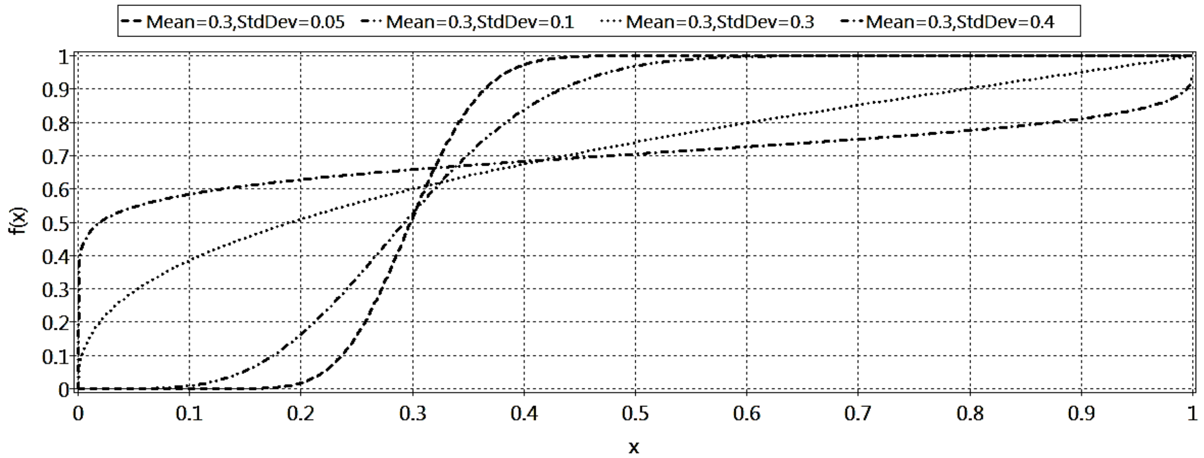

Figure 2.16: Beta cumulative distribution function.

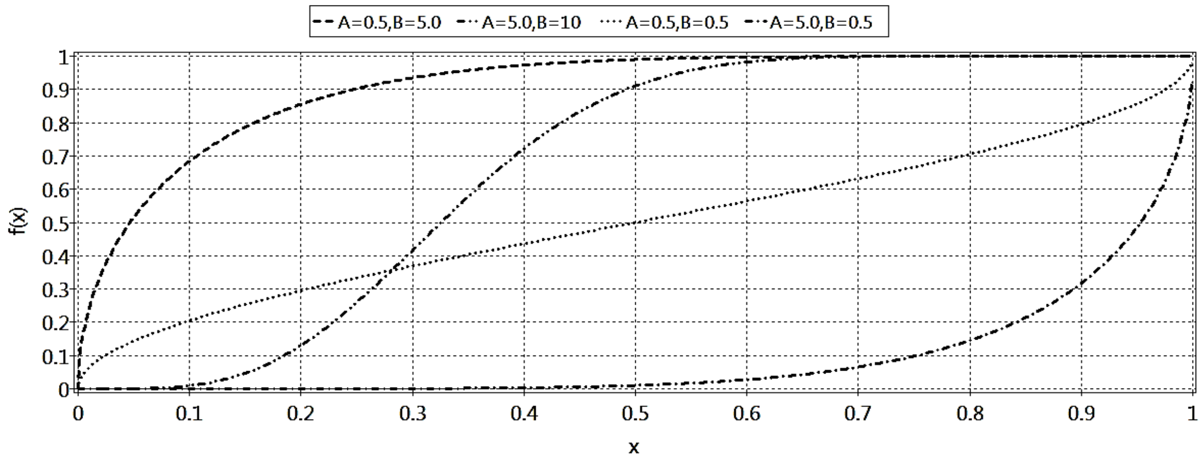

Figure 2.17: Incomplete Beta function.

Stellenbosch University https://scholar.sun.ac.za

2.6 Time of Use Tariff Structures

2.6.1 Introduction

In Time Of Use (TOU) tariff structures a set of different electricity tariffs are defined for different times of the day and seasons of the year [51]. In the 1980's the South African utility (Eskom) started to gradually introduce TOU tariff structures to South African consumers [52]. Currently, Eskom provides TOU tariff structures aimed at numerous types of consumers [53].This study however only considers the HomeFlex and MegaFlex TOU tariff structures.

Both these TOU tariff structures consist of two tariff seasons, namely the High Demand and Low Demand season. The calendar months of the year are grouped together according to the level of energy demand. The High Demand season consists of winter months June to August when the energy demand is at its highest, while the Low Demand consists of the remainder of the year when energy demand is lower [53]. These seasons are chosen by the Eskom and the TOU tariff rates are set accordingly, i.e. the tariff rates of the same tariff period differ between these two seasons. The High Demand season tariffs are significantly higher than the Low Demand season tariffs. Therefore, the financial profitability of an industrial consumer will be greatly affected by these two seasons.

2.6.2 MegaFlex Tariff

MegaFlex represents the main TOU tariff structure used by large industrial consumers, local authorities and municipalities. The MegaFlex tariff structure consists of three different tariff periods that vary with respect to hours of the day and days of the week. Table 2.1 summarises the MegaFlex tariff periods [53].

Table 2.1: MegaFlex tariff period hours.

Day of Week	Tariff Period	Period hours
Weekday	Evening Off-peak	22:00 - 06:00
	Morning Standard	06:00 - 07:00
	Morning Peak	07:00 - 10:00
	Afternoon Standard	10:00 - 18:00
	Evening Peak	18:00 - 20:00
	Evening Standard	20:00 - 22:00
Saturday	Evening Off-peak	20:00 - 07:00
	Morning Standard	07:00 - 12:00
	Afternoon Off-peak	12:00 - 18:00
	Evening Standard	18:00 - 20:00
Sunday	Off-peak	00:00 - 24:00

Stellenbosch University https://scholar.sun.ac.za

2.6.3 HomeFlex Tariff

The HomeFlex tariff is initially implemented on a voluntary basis to residential customers [54]. The HomeFlex tariff structure consists of two different tariff periods that vary only with respect to hours of the day. Table 2.2 summarises the HomeFlex tariff periods [54].

Table 2.2: HomeFlex tariff period hours.

Day of Week	Tariff Period	Period hours
Everyday	Evening Off-peak	20:00 - 07:00
	Morning Peak	07:00 - 10:00
	Afternoon Off-peak	10:00 - 18:00
	Evening Peak	18:00 - 20:00

2.7 Solar Power

2.7.1 Introduction

This section gives a brief overview the main aspects of solar power as an energy source. These aspects include the following:

- The fluctuating and intermittent nature of solar radiation.
- The modelling of solar radiation.
- Photovoltaic (PV) systems configurations and efficiency.

As solar radiation passes through the atmosphere it is scattered and absorbed by particles. Therefore, the radiation that reaches the surface depends significantly on the length of the path taken through the atmosphere [55]. Atmospheric parameters that attenuate the radiation include ozone, aerosol, dry air and water vapour [56]. However, the main factor that affects the difference in solar radiation between the outside of the atmosphere and the earth’s surface is cloud cover [7].

The solar radiation on a collector on the earth’s surface will be the sum of direct-beam radiation, diffuse radiation and reflected radiation [55]. Direct-beam radiation passes straight through the atmosphere to the collector and diffuse radiation is reflected off particles in the atmosphere. Furthermore, reflected radiation is bounced off the ground and other surfaces near the collector [55].

Stellenbosch University https://scholar.sun.ac.za

2.7.2 Solar Radiation

The extra-terrestrial solar radiation just outside the earth's atmosphere is the starting point to modelling the clear sky solar radiation. The extra-terrestrial solar radiation I_0 depends on the distance between the earth and the sun and is given by the following relationship [57] [58]:

$$I_0 = I_{SC}\left[1 + 0.033cos\left(\frac{360n}{365.25}\right)\right] \tag{2.39}$$

where I_0 is in Watts per square meter, I_{SC} is the solar constant and n is the day number starting from the first of January. The solar constant is the estimated average annual extra-terrestrial solar radiation which varies from 1.367 kW/ m^2 [58] to 1.377 kW/ m^2 [55].

Models discussed in this study deal with the radiation on horizontal planes. However, horizontal plane radiation can be transposed to any other plane [59].

2.7.2.1 Clear Sky Direct Beam Radiation

A simple model used to characterise the transmittance of beam radiation through a clear atmosphere is that of Hottel [57] [58] [60]. Hottel's model for the atmospheric transmittance of beam radiation τ_B takes into account the zenith angle and altitude and is given by the following relationship [58] [60]:

$$\tau_B = a_0 + a_1 e^{-k/\cos\theta_z} \tag{2.40}$$

where Θ_Z is the zenith angle and constants a_0, a_1, k are for the standard atmosphere with 23 km visibility and an altitude below 2.5 km. Constants a_0, a_1 and k are determine from the relationships a_0^*, a_1^* and k^* [57] [60] given by the following relationships respectively [58] [60]:

$$a_0^* = 0.4237 - 0.00821(6 - A)^2 \tag{2.41}$$

$$a_1^* = 0.5055 - 0.00595(6.5 - A)^2 \tag{2.42}$$

$$k^* = 0.2711 - 0.01858(2.5 - A)^2 \tag{2.43}$$

where A is the altitude of the observer in kilometres. [58].

Correction factors r_0, r_1 and r_k are applied to these constants to allow for different climates and are given by the following relationships [58] [60]:

$$r_0 = \frac{a_0}{a_0^*} \tag{2.44}$$

$$r_1 = \frac{a_1}{a_1^*} \tag{2.45}$$

$$r_k = \frac{a_k}{a_k^*} \tag{2.46}$$

Table 2.3 summarises different correction factors for Hottel's model for different climates [58].

Table 2.3: Correction factors for different climates for Hottel's model.

Climate Type	r_0	r_1	r_k
Tropical	0.95	0.98	1.02
Midlatitude Summer	0.97	0.99	1.02
Subarctic Summer	0.99	0.99	1.01
Midlatitude Winter	1.03	1.01	1

The clear sky horizontal beam radiation I_B is given by the following relationship [58]:

$$I_B = I_0 \tau_B cos\theta_Z \tag{2.47}$$

where I_0 is the extra-terrestrial solar radiation and τ_B is atmospheric transmittance for beam radiation [57].

2.7.2.2 Diffuse Radiation

There are several models available to estimate the diffuse radiation on a surface [55]. One of the models used to estimate the clear sky diffuse radiation on a horizontal surface is that of Liu and Jordan [58] [61]. Liu and Jordan developed an empirical relationship between the beam and diffuse radiation for clear days given by the relationship [58] [61]:

$$\tau_d = 0.271 - 0.294\tau_B \tag{2.48}$$

where τ_d is the ratio of diffuse radiation to the beam radiation on the horizontal plane.

Stellenbosch University https://scholar.sun.ac.za

The clear sky diffuse radiation I_D is given by the following relationship [58]:

$$I_D = I_0 \tau_d cos\theta_Z \tag{2.49}$$

2.7.2.3 Reflected Radiation

Reflected radiation is the component of solar radiation that is reflected by the surfaces in front of a collector [55]. Ineichen et al define albedo as the ratio between the ground reflected radiation and the global radiation incident on the ground [62]. Ineichen et al concluded that accurate results could be acquired under the assumption that the ground-reflected radiation is isotropic and using a constant averaged measured albedo for the site [62].

2.7.3 Photovoltaic System Configurations

There are several configurations of Photovoltaic (PV) systems that consist of a range of components such as inverters, maximum power point trackers, batteries and charge controllers [55]. For most applications the power of one individual panel is not enough, therefore the panels are grouped together into arrays to achieve the desired voltage and current output [63]. Typical PV Panels are comprised of 30 to 36 series connected solar cells with open-circuit voltages of about 20 Volts and short circuit currents of about 3 to 4 Amperes [63]. The three most common PV systems are grid connected systems, stand- alone systems and directly connected load systems [55].

2.7.3.1 Grid Connected Photovoltaic Systems

Grid connected PV systems feed power directly into the power grid through a power conditioning unit [55]. The power conditioning unit converts the DC power from the PV panels into AC power to be fed into the power grid or to supply a load. If a load draws more power at any instant than what can be supplied by the PV panels, the power conditioning unit draws the required power from the power grid to satisfy the demand [55]. When the PV panels provide more power than that being used by the load, the excess power is sent to the power grid. Furthermore, the power conditioning unit also has the function of keeping the PV panels operating at their highest efficiency by utilising a Maximum Power Point Tracker (MPPT) [55]. Tracking the maximum power point of a PV panel array is essential and many methods have therefore been developed and implemented [64].

Stellenbosch University https://scholar.sun.ac.za

Grid connected PV systems have many advantages such as simplicity, reliability and high operating efficiencies. Initial grid connected systems consisted of many series and parallel PV panels connected to a large central inverter. However, grid connected system configurations have since then progressed towards the implementing of string technology [63]. With string technology all the PV panels are arranged and configured into a number of groups. Each group consists of a number of series connected PV panels and an inverter [63]. This allows a PV system to be extended and scaled by simply adding and removing groups of PV panels and inverters.

2.7.3.2 Stand-alone PV Systems

Stand-alone PV systems generally consist of a PV array, energy storage such as batteries, a power processor and a maximum power point tracker [64]. These PV systems may also include an optional generator as a backup supply of power [55]. Stand-alone PV systems are well suited for remote locations. However, they have the drawbacks of significant battery losses [55] and the fact that battery storage is generally very expensive.

2.7.3.3 Directly Connected Load PV System

Directly connected load PV systems are very simple, reliable and cost effective. They have no power conditioning units or batteries to store energy. Examples of such systems are PV water pumps that pump water when the sun is shining [55]. These type of PV systems need to be carefully designed to be efficient [55].

2.7.4 PV System Efficiency

The efficiency of a PV system depends on the PV panels and the inverters. Many factors affect the power output of PV panels, with cell temperature [55] and the tilt angle [65] being the most significant. An increase in PV panel temperature is accompanied by a significant decrease in power output [55] [65]. Inverter efficiencies may vary depending on the load connected to it, with efficiencies of above 90% at high loads [55]. Inverter and panel efficiencies can add up to a significant decrease in the efficiency of a PV system. Losses of up to 25% have been reported by University of Tokyo [66].

The rated DC power output of PV panels under Standard Test Conditions (STC) can be used to estimate the performance of PV systems [55]. Standard test conditions as defined by standard (IEC 60904-3) are 1000 W/m^2 irradiance, 25°C cell temperature and an air mass of 1.5 [55].

Stellenbosch University https://scholar.sun.ac.za

Note that the DC power of a PV panel array is determined by simply adding the individual panel ratings under STC together [55]. The estimated AC power output P_{AC} can then be determined using the following relationship [55]:

$$P_{AC} = P_{DC,\ STC} \times (Conversion\ Efficiency) \tag{2.50}$$

where $P_{DC,\ STC}$ is the DC power of the PV panels under standard test conditions and the conversion efficiency accounts for the inverter efficiency, dirt on the collectors and ambient temperature [55].

In practice the temperature of a PV panel is likely to vary from the STC 25°C and therefore better test conditions are required. The PVUSA is a monitoring program that developed a rating system based on conducted field tests. The PVUSA Test Conditions (PTC) are defined as 1000 W/m^2 irradiance, 20°C cell temperature and a wind speed of 1 m/s [55] [67].

Stellenbosch University https://scholar.sun.ac.za

3 Database Design and Implementation

3.1 Overview

This chapter presents the design and implementation of a relational database structure used for storing all historical generation data together with all Time of Use (TOU) structure information. The database is hosted using WAMPServer as discussed in section 2.2.

A custom relational database topology is developed to store historical generation data from several meters and projects on the same database. This requires a generic approach to organising and referencing the data stored on the database. Therefore, the concept of profiles and profile sets is developed to relate the time-stamped values in the database.

As mentioned in section 2.4.2, the Unified Process acknowledges risk in design and development by highlighting the unknown aspects of the system being designed. The most significant risk identified for the database structure is a structure that restricts future changes. A non-generic database structure which does not allow for the changing and rearranging of data and TOU structures may lead to a situation where the software application must be altered in the future in order to accommodate a certain TOU structure.

TOU tariffs and structures are subject to change depending on the requirements of the analysis. Utility providers such as Eskom can change TOU tariffs as the cost of producing electricity changes. Furthermore, the defined time intervals of tariff structures are also subject to change depending on the utility provider. Therefore, the data contained in the database may nee to be rearranged or changed in the future. The database must therefore accommodate the changes and incorporate a generic approach to storing and accessing TOU structures.

3.1.1 Database Topology

A profile is defined as the historical time-stamped values of a measured parameter, while a profile set is defined as a set of profiles that are related in some way. Therefore, a metering device that measures multiple parameters has multiple profiles associated with it and is regarded as a profile set. The measuring of parameters is usually done as part of a project using a number of meters. All profile sets (meters) that are used to measure parameters in a project are related and therefore linked together.

Managing and accessing all stored profile data, i.e. historical generation data, and TOU structure data requires a fixed referencing system which is applicable across the entire database. Therefore, a set of main fields are required which allows the software application to navigate the data on the database. There are five main fields used in the majority of the database tables namely:

- *ID* is an integer field and is by default the Primary Key (PK) of all tables.
- *Designation* contains the unique name of a record (row).
- *Description* contains a brief description of a record.
- *Comments* contains additional notes on a record.
- Some *Foreign Key* (FK) or keys that reference a field in another table.

Note all primary keys and foreign keys are integer fields and are required to adhere to foreign key constraints. The *ID* and *Designation* fields together form a superkey as discussed in section 2.2.

3.1.2 Case Study

This project involves the measuring and logging of a solar plant's energy output using a number of power meters to measure the active power output. The active power output measurements are used to determine the historical generated energy data. All time-stamped historical generation data is stored on a relational database and linked to the power meters and the project.

Figure 3.1 illustrates the case study database structure were each profile set represents one of the N number of power meters used in the project. Each profile represents the generated energy measured by its respective meter. The time-stamped measured values are represented by each individual profile's generation data. The profile data of all profiles is stored together in a single table as indicated in figure 3.1. Likewise, all profiles are stored in a single table and all profile sets are stored in a single table.

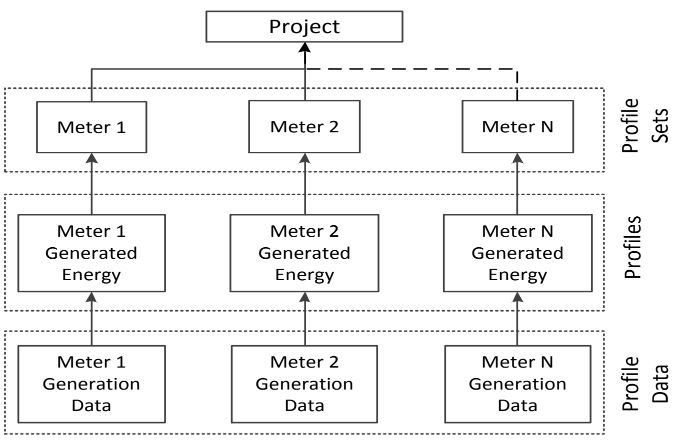

Figure 3.1: Case study database structure.

Stellenbosch University https://scholar.sun.ac.za

Note figure 3.1 only presents the main topology of the database and excludes several supplementary lookup-tables. A full and detailed description of all tables is given in the following section.

3.2 Database Tables

The data contained in the database consists of the historical generation data together with the TOU structure data. The generation data is stored separately from the TOU structure data as these tables have different structures and relations. However, the table structure for both the historical generation data and the TOU structure data are designed to be generic to allow for possible changes in the future. The database tables are divided into two different sets of tables namely:

- *Profile Tables:* These tables relate to the categories, units and relationships of the measured historical generation data as well as the projects they belong to.
- *Time Of Use Tables*-These tables relate to the relevant information for TOU structures.

3.2.1 Profile Tables

3.2.1.1 Overview

Figure 3.2 presents all the profile tables in the database and the relationships between them. Arrows indicate that a foreign key in the table refers to a primary key in the table that the arrow points to.

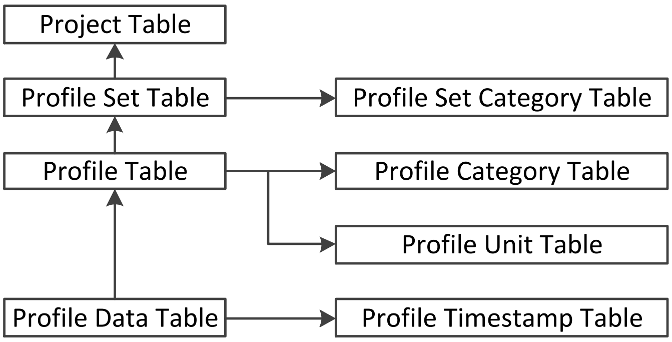

Figure 3.2: Profile tables in the database and the relationship between them.

As can be seen from figure 3.2 the profile data is stored in a set of tables. This is necessary to ensure that the database structure is generic and accommodates future changes. Using the table relation presented in figure 3.2 it is possible to define profiles and profile sets of any kind by allowing the user to set and alter the profile and profile set categories. Furthermore, the units can also be defined by the user in the Profile Unit table. This structure therefore allows for any type of profile to be saved with

Stellenbosch University https://scholar.sun.ac.za

any type of unit and linked to any type of profile set. For example, a project that measures the temperatures and power output of a number of gas turbines can easily be set up using this structure. The profile sets will be defined as the names of each individual gas turbine. The temperature and power output measurements of each respective gas turbine will be linked as profiles to the name of said gas turbine. The units of degrees Celsius and Watts will be added to the units table and each unit will be linked to the temperature and power output profiles of the gas turbines respectively.

3.2.1.2 Project Table

As mentioned, the collecting and storing of measured data is usually conducted as part of a project and it may involve multiple meters. The measurement data from multiple projects may need to be stored on the same database and therefore a top level table is required. The Project table contains the names of all projects on a database and has no Foreign Key (FK). Figure 3.3 presents the design of the Project table fields with the ID field as the Primary Key (PK).

Project	
PK	**ID**
	Title Description Comments Registration Date

Figure 3.3: Design of Project table.

The Project table is chosen as the top level table in order to distinguish between meters of different projects and different sites. Furthermore, this allows the user to group and link certain meters together. The Registration Date field is for administrative purposes and allows the user to set the date at which a project commenced.

3.2.1.3 Profile Set and Profile Set Category Tables

Profile sets are used to relate different profiles. In the case of this project they represent the individual power meters. Profile sets could describe different types of profile combinations and therefore a category look up table is required. Figure 3.4 presents the design of the Profile Set and Profile Set Category table fields with the ID fields as the PKs.

Stellenbosch University https://scholar.sun.ac.za

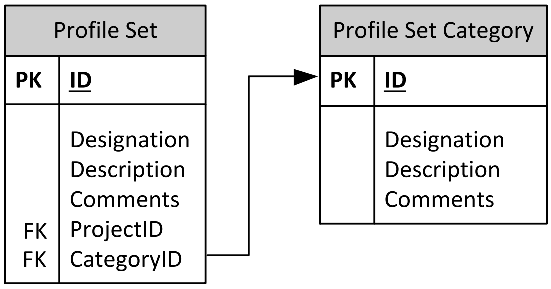

Figure 3.4: Design of Profile Set and Profile Set Category tables.

ProjectID is a FK which references the *ID* field in the Project table. *CategoryID* is a FK which references the *ID* field in the Profile Set Category table.

The Profile Set Category table allows the user to easily link profile sets on the database to any defined category or type of profile set. A new category can be defined and added to the Profile Set Category table by the user. The link between profile sets on the database and profile set categories can also be changed if necessary. Therefore, the category of a profile set can be changed to a newly defined and added category in the Profile Set Category table if required. This allows the database to be altered and rearranged in the future if necessary.

3.2.1.4 Profile, Profile Category and Profile Unit Table

Profiles represent the different types of parameters being measured and therefore can have different units. Therefore, a Profile Category and Profile Unit look-up table is required. Figure 3.5 presents the design of the Profile, Profile Category and Profile Unit table fields with the ID fields as the PKs.

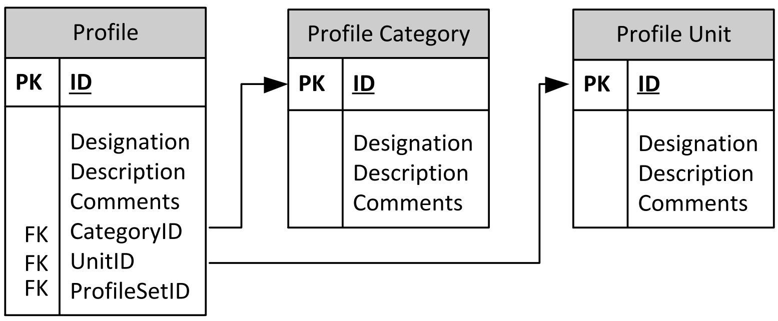

Figure 3.5: Design of Profile, Profile Category and Profile Unit tables.

ProfileSetID references the *ID* filed in the Profile Set table. *CategoryID* is a FK which references the *ID* field of the Profile Category table. *UnitID* is a FK which references the *ID* field in the Profile Unit table.

The Profile Category table allows the user to easily link profiles on the database to any defined profile category or type. A new category can be defined and added to the Profile Category table by the user. The link between profiles on the database and profile categories can also be changed if necessary. Therefore, the category of a profile can be changed to a newly defined and added category in the Profile Category table. This allows the database to be altered and rearranged in the future if necessary.

Furthermore, the Profile Unit table allows the user to easily link profiles on the database to any defined unit. A new unit can be defined and added to the Profile Unit table by the user. The profiles on the database and profile unit can then be linked. For example, if the profile consists of power measurements the unit of Watts will be added to the Profile Unit table and the respective profile will be linked to that unit. Therefore, any unit can be added to the database and therefore a profile can consist of any type of measurements. This allows the database to be generic in the type of data measurements stored and allows for the database to be used on a variety of applications.

3.2.1.5 Profile Data and Profile Timestamp Table

The Profile Data and Profile Timestamp tables contain the actual measured time-stamped values, i.e. the historical generation data. The measured values are stored separate from their respective timestamps. The values are stored in the Profile Data table and the timestamps are stored in the Profile Timestamp table. Figure 3.6 presents the design of the Profile Data and Profile Timestamp table fields with the ID fields as the PKs.

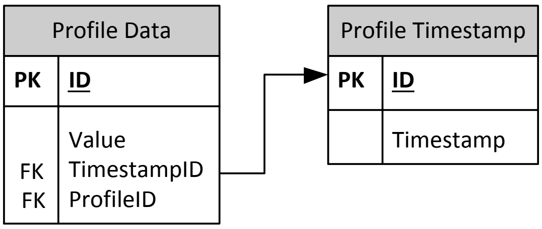

Figure 3.6: Design of Profile Data and Profile Timestamp tables.

TimestampID is a FK which references the *ID* field in the Profile Timestamp table and *ProfileID* is a FK which references the *ID* field in the Profile table.

The structure of the Profile Data and Profile Timestamp also allows for future changes made to the database as it allows the link between the data and the timestamps to be altered if necessary.

Stellenbosch University https://scholar.sun.ac.za

3.2.2 Time Of Use Tables

3.2.2.1 Overview

TOU structures differ from one another with respect to seasons of the year, days of the week and hours of the day. Therefore, a generic database structure is required which accommodates the storing of different TOU structures efficiently. Figure 3.7 presents all the TOU tables and the relationships between them. Again arrows indicate that a foreign key in the table refers to a primary key in the table that the arrow points to.

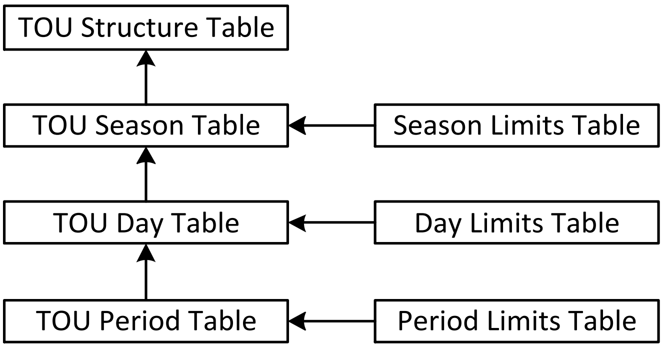

Figure 3.7: TOU tables and the relationships between them.

It is essential to split the TOU structures into seasons of the year, days of week and hours of the day as it enables the use of any type of TOU period. This way the user could manipulate the TOU structure to produce any number of periods over any number of seasons, days and hours. This database structure also allows for alterations to the TOU structures and tariffs.

The TOU seasons can be defined as any number or range of calendar months. This is particularly useful for defined TOU tariff structures provide by utilities such as Eskom. The TOU tariffs depend and change with regard to certain calendar months of the year as indicated in section 2.6. The database structure accommodates for these defined seasons by providing the user with a generic way to define seasons using the calendar months of the year. The Season Limits Table stores the range of calendar months defined for each season. Each user defined TOU season is stored in the TOU Season Table and is then linked to a range of defined months in the Season Limits Table.

TOU structures are generally defined on a day of the week basis. The TOU tariffs depend and change with regard to certain days of the week as indicated in section 2.6. The database structure accommodates for this by providing the user with a generic way to define tariff days using the days, or a range of days, of the week. The Day Limits Table stores the range of days of the week defined for

Stellenbosch University https://scholar.sun.ac.za

each tariff day. For example, a tariff day may be defined by the utility as all the weekdays Monday to Friday. The weekday tariff day is added to the Day Table and the range of days Monday to Friday is added to the Day Limits table.

The tariff period can be defined as any range of time during a day. The TOU tariffs depend and change with regard to certain times of the day as indicated in section 2.6. The database structure accommodates for these defined time intervals by providing the user with a generic way to define tariff periods using user defined time intervals. The Period Limits Table stores the range of times defined for each tariff period. Each user defined TOU tariff period is stored in the TOU Tariff Period Table and is then linked to a range of defined time intervals in the Period Limits Table.

This approach of connecting tariff periods to days and tariff days to seasons provides a generic way of defining any type of TOU tariff structure for any time of day, any day of the week and any month of the year.

3.2.2.2 TOU Structure Table

The TOU Structure table contains all the different TOU structure names. Figure 3.8 presents the design of the TOU Structure table fields with the ID field as the PK.

TOU Structure	
PK	**ID**
	Designation Description Comments

Figure 3.8: Design of TOU Structure table.

Note that due to the fact that this is a top level table it has no foreign keys. The TOU structure table is chosen as the top level table in order to distinguish between different TOU structures. The user can add any TOU structure name to this table. The tariff seasons are then linked to the user added TOU structure.

3.2.2.3 TOU Season and Season Limits Tables

Figure 3.9 presents the design of the TOU Season and Season Limits table fields with the ID fields as the PKs.

Stellenbosch University https://scholar.sun.ac.za

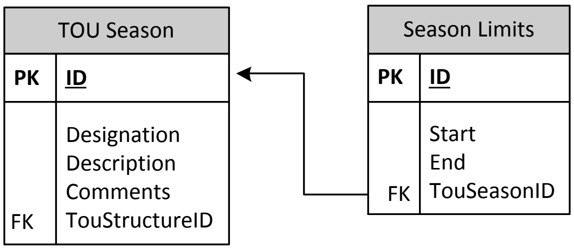

Figure 3.9: Design of TOU Season and Season Limits.

TouStructureID is a FK which references the *ID* field in the TOU Structure table. *TouSeasonID* is a FK which references the *ID* field in the TOU Season table. The Start and End fields in the Season Limits table represent the start and end of a TOU season in months of the year.

The TOU Seasons table contains the user defined TOU season names, while their limits (range of calendar months) are stored in the Season Limits table. The Start and End fields in the Season Limits table refer to the starting and ending calendar months of a user defined season.

3.2.2.4 TOU Day and Day Limits Tables

Figure 3.10 presents the design of the TOU Day and Day Limits table fields with the ID fields as the PKs.

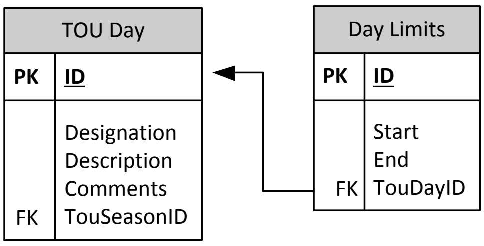

Figure 3.10: Design of TOU Day and Day Limits tables.

TouSeasonID is a FK which references the *ID* field in the TOU Season table. *TouDayID* is FK which references the *ID* field in the TOU Day table. The Start and End fields in the Day Limits table represent the start and end of a TOU day in days of the week with Sunday being the first day.

Stellenbosch University https://scholar.sun.ac.za

The TOU Day table contains the user defined TOU day names, while their limits (range of days of the week) are stored in the Day Limits table. The Start and End fields in the Day Limits table refer to the starting and ending days of the week of a user defined tariff day.

3.2.2.5 TOU Period and Period Limits Tables

Figure 3.11 presents the design of the TOU Period and Period Limits table fields with the ID fields as the PKs.

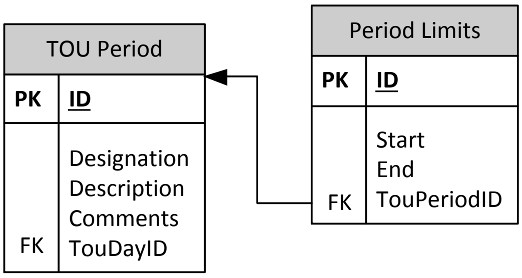

Figure 3.11: Design of TOU Period and Period Limits tables.

TouDayID is a FK which references the *ID* field in the TOU Day table. *TouPeriodID* is a FK which references the *ID* field in the TOU Period table. The Start and End fields in the Period Limits table represent the start and end of a TOU periods in hours of the day.

The TOU Period table contains the user defined TOU period names, while their limits (user defined time of day intervals) are stored in the Period Limits table. The Start and End fields in the Period Limits table refer to the starting and ending times of the interval of a user defined tariff period.

3.2.3 Testing of Database Structure

This project aimed at developing a database structure that could store historical generation data together with TOU structures. The database structure is required to be generic by design and consider that changes may need to be made to it in the future.

The database was tested by adding test profile sets, profiles and profile data, i.e. historical generation data. The database successfully stores historical generation data in a generic way which allows for future changes to be made.

Furthermore, the database structure is required to incorporate TOU structures and allow for future change. This was tested by adding a range of TOU structure and tariff structures to the database.

Stellenbosch University https://scholar.sun.ac.za

These added TOU structures and tariffs were then altered and rearranged. The database structure successfully stores and incorporates TOU structures and allows for future changes to be made.

Stellenbosch University https://scholar.sun.ac.za

4 Software Application Design and Implementation

4.1 Overview

This chapter provides an overview of the design and implementation of the forecasting software application. The software application is developed using disciplines and design strategies of the Unified Process outlined in section 2.4. The design and implementation of the software application is discussed with respect to the four phases of the Unified Process namely:

- Inception
- Elaboration
- Construction
- Transition

This section presents the use case and activity diagrams of the system design. Use case diagrams illustrate the interactions between the user and the system, while activity diagrams illustrate the flow of tasks or activities within operations.

4.2 Inception Phase

This phase defines the scope and feasibility of the project. The final output is the vision for the system, a very simplified use case model, the significant risks and a provisional system architecture.

4.2.1 Scope and Vision

The envisioned goal of this project is a software application that is capable of forecasting and modelling the long term energy output of a solar plant. The software application is required to incorporate historical generation data and the relational database structure presented in Chapter 3. The following is required from the envisioned software application:

- Connect to user selected relational database.
- Import historical data into selected database and check data integrity.
- Access and manipulate generation data stored on a database.
- Implement statistical methods to derive models from data stored on a database.
- Incorporate Time Of Use (TOU) structures.
- Implement an intuitive Graphical User Interface (GUI).
- Implement a modular and extensible software system design.

Stellenbosch University https://scholar.sun.ac.za

In view of the above requirements, the use case model of the forecasting software application is defined and presented in figure 4.1.

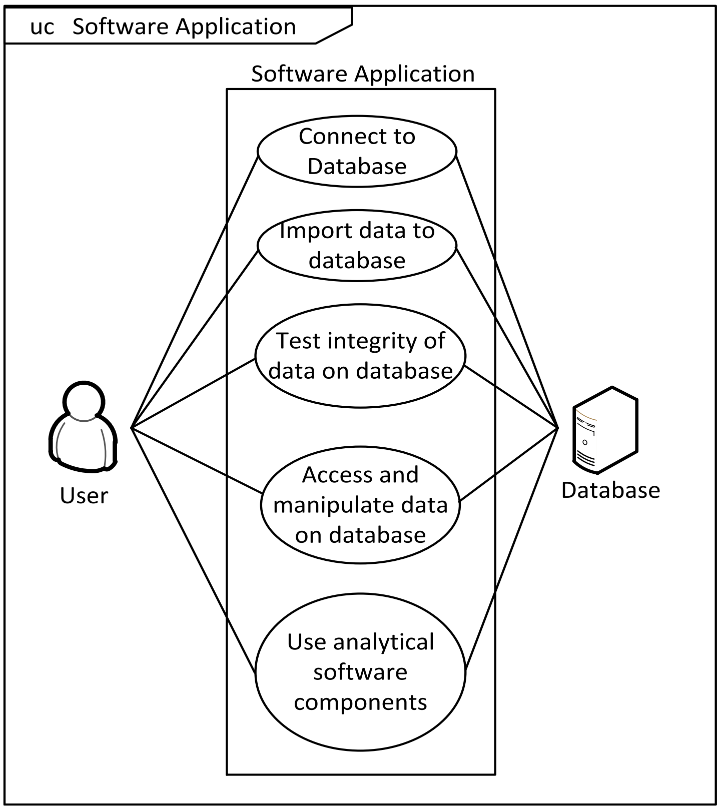

Figure 4.1: Use case diagram of software application.

4.2.2 Significant Risks

The most significant risk identified during this phase is the system design and architecture. A non-modular and non-extensible design may lead to code duplication and having to restructure the system design repeatedly. Designing the system as one unit could lead to the situation where earlier stages of the software implementation have to be altered to accommodate design limitations in later stages of the software implementation.

4.2.3 Provisional Architecture and Feasibility

Two system architectures are considered for the development of the software application. The first architecture is based on a modular design with a central hub, while the second architecture is based on a single unit which contains all the required functionality. The modular architecture is more complex to develop as it requires the system to communicate and transfer data between individual modules.

Stellenbosch University https://scholar.sun.ac.za

However, modular system architecture has the advantage of being extensible, i.e. functional modules can be added as required. Furthermore, a modular design allows for the reuse of code and allows for the implementing, debugging, testing and updating of each software module individually.

The system architecture based on a single unit containing all functionality is less complex to develop. However, this architecture is very limited in terms of adding functionality at a later stage. Furthermore, singular unit system architecture has the disadvantage of having to recompile and distribute the entire application when a certain function is altered or updated.

The modular system architecture with a central hub is chosen as provisional architecture for its advantage of extensibility, reuse of code and modular updating and debugging. The central hub links and manages all the different software components (modules) connected to it. This central hub is responsible for calling each software component as needed and providing it with all required parameters. Due to the fact that the software application is database driven, a database connection plays an essential role in the software system. The central hub is responsible for creating a single database connection and passing it to each of the software components as required. Figure 4.2 presents the provisional system architecture for the software application with *N* number of software modules.

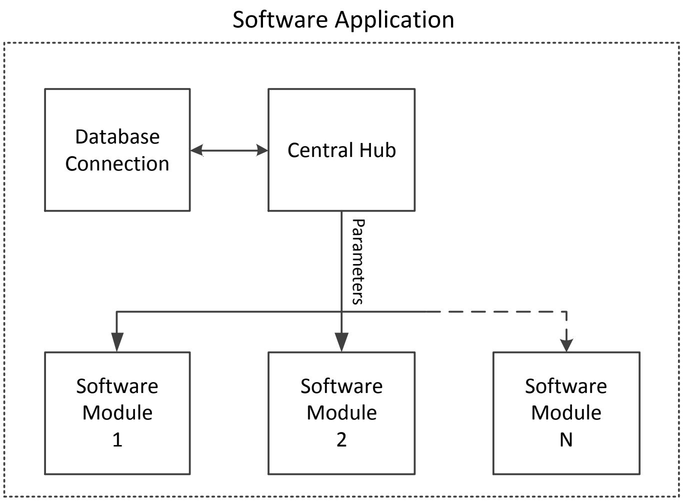

Figure 4.2: Provisional system architecture of software application.

4.2.3.1 Feasibility

The provisional system architecture is both modular and extensible and allows for any number of software components to be added to the system. This system design has the advantage of requiring only one database connection to be created and passed between the central hub and all software components. Therefore, this architecture is both a feasible and a practical approach to system design.

Stellenbosch University https://scholar.sun.ac.za

4.3 Elaboration Phase

Functional requirements of the system are captured in this phase as well as the creation of the final system architecture to be used. The main output is the architecture, a detailed use case model and plans for the construction stage.

4.3.1 Functional Requirements

The main functional requirements of the software application are as follows:

- A unique GUI for each software component.
- A MySQL database connection capable of connecting to databases on local or remote servers.
- A software component capable of importing time-stamped generation data from Comma Separated Value (CSV) files.
- A timeline integrity analysis component to check for duplicated, extra or missing generation data on a database.
- A Profile Administration System (PAS) which allows for the adding, removing and editing of profiles and profile sets on the database.
- A Profile Analysis Engine (PAE) which analyses and processes data on the database.
- The functionality to export results to Excel worksheets for further processing if required.

Furthermore the software application is required to have a GUI based multi-select filter which allows the user to select any combination of profiles from various profile sets and use them collectively. The multi-select filter software component has the following functional requirements:

- Connect to the database and provide the user with all the available projects, profiles and profile sets on the database.
- Create a list of all user selected profiles which is passed to other software components as a parameter.

4.3.2 Architecture

The final architecture used for the software application is based on the provisional architecture and functional requirements of the system. All software components are required to be grouped together according to functionality and developed as separate Dynamic Linked Libraries (DLLs). DLLs are used to modularise and reuse code which could be shared between Windows applications [30].

Stellenbosch University https://scholar.sun.ac.za

All software components are grouped into subsystems and then further divided into different categories. Each category of software components is developed as a separate DLL. The software application consists of four main subsystems namely:

- *Main Application*: The central hub of the system which manages all components and passes parameters such as the database connection and the multi-select list of profiles.
- *Database Connection*: The actual database connection which facilitates all database access.
- *Profile Administration System*: This subsystem consists of all software components aimed at managing and maintaining the profile sets, profiles and profile data on a database.
- *Profile Analysis Engine*: This subsystem consists of all software components aimed at the analysis or processing of generation data to derive forecasting models.

Each software component requires its own custom GUI which allows the user to select and specify settings unique to its functionality. Due to the fact that all software components are contained in DLLs, it is the responsibility of the Main Application to call the desired software component GUIs when required. Figure 4.3 presents the final architecture of the software application.

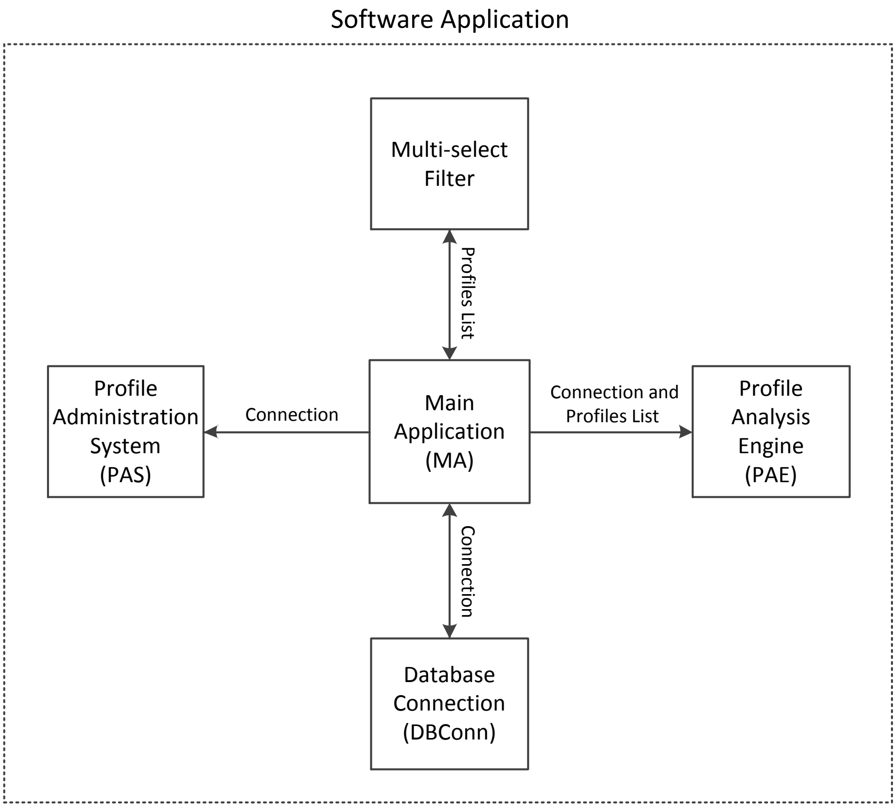

Figure 4.3: Final software application architecture.

Stellenbosch University https://scholar.sun.ac.za

In order to share a database connection and a list of profiles between the Main Application (MA) and the software components, the use of the Common Object Model (COM) and Object Linking and Embedding (OLE) is implemented. COM forms the basis of OLE and defines an Application Programming Interface (API) for communication between objects. COM objects consist of one or more interfaces used to call their methods [30]. These interfaces are used to instantiate objects and use methods inside the separate DLLs.

4.3.2.1 Profile Administration System

The Profile Administration System (PAS) is responsible for managing and maintaining of profile sets, profiles and profile data on the database. The PAS therefore requires the following software components:

- *Profile Manager*: Used for adding, editing and removing profiles from the connected database. Also used to remove profile data from database.
- *Profile Set Manager*: Used for adding, editing and removing profiles sets from the database
- *Profile data Importer*: Used for importing profile data into database from CSV files and link imported data to profiles.

4.3.2.2 Profile Analysis Engine

The Profile analysis Engine (PAE) is responsible for all analysis and processing of profile data and requires the following software Components:

- *Statistical TOU Analysis*: Used to derive TOU statistical parameters and models from historical generation data.
- *Timeline Integrity Analysis*: Used for investigating the timeline integrity of profile data to determine whether data is missing, extra or duplicated.

4.3.3 Detailed Use Case Model

This section presents the use case model of the software application. Figures 4.4 to 4.11 present the use case diagrams of all software components and subsystems that make up the software application.

Stellenbosch University https://scholar.sun.ac.za

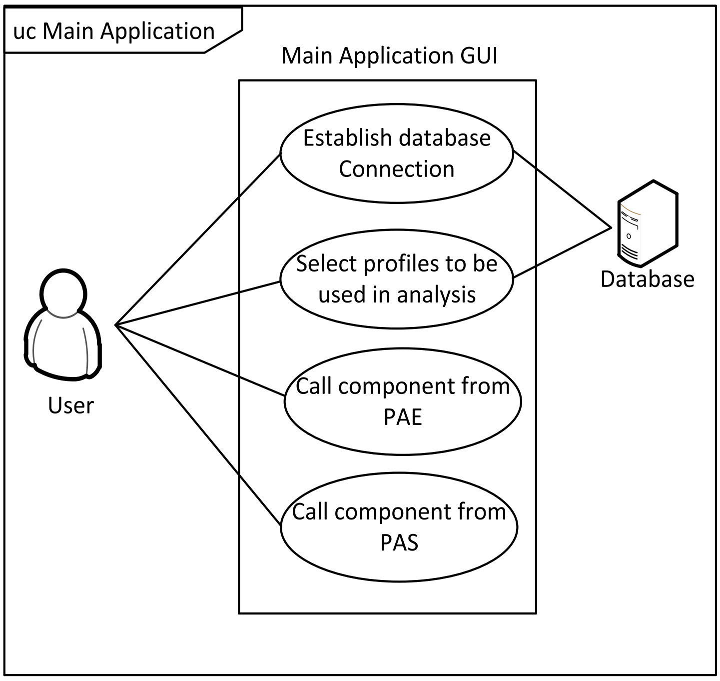

Figure 4.4: Use case diagram of main application.

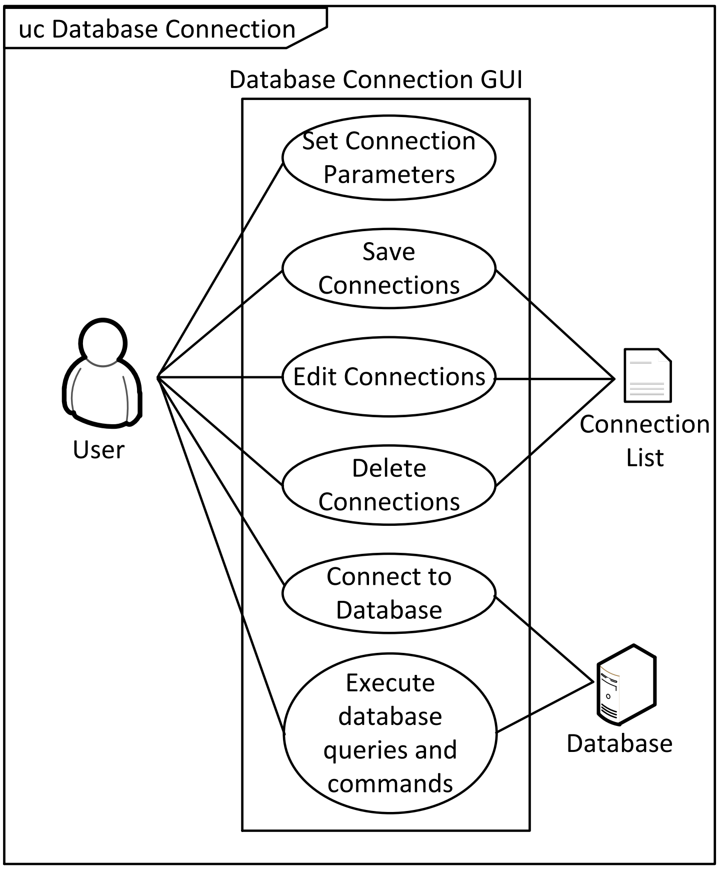

Figure 4.5: Use case diagram of database connection.

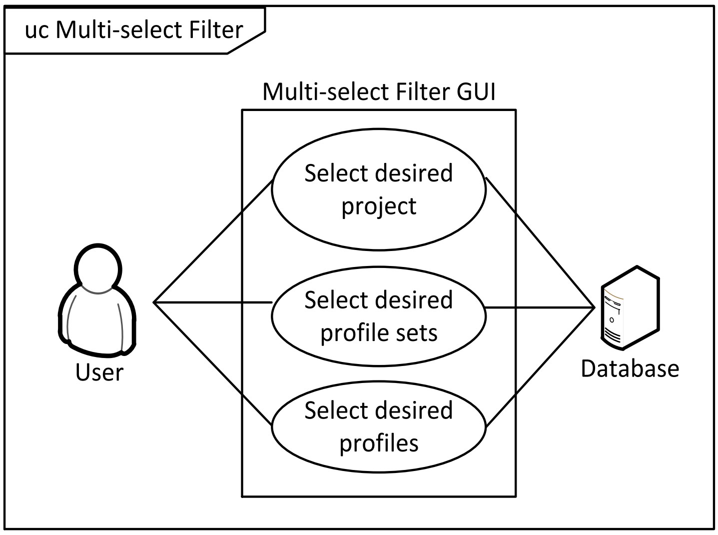

Figure 4.6: Use case diagram of multi-select filter software component.

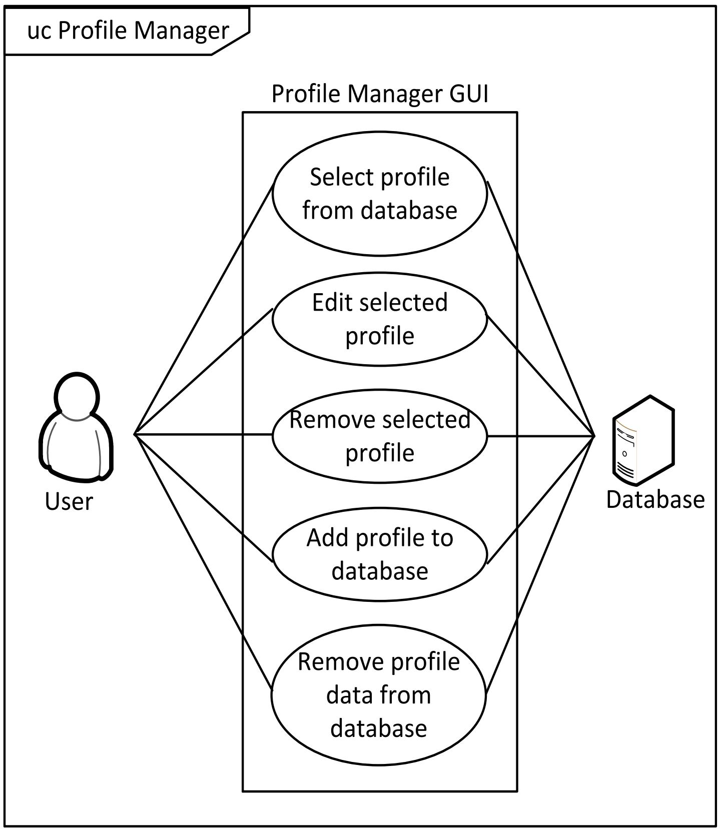

Figure 4.7: Use case diagram of profile manager software component.

Stellenbosch University https://scholar.sun.ac.za

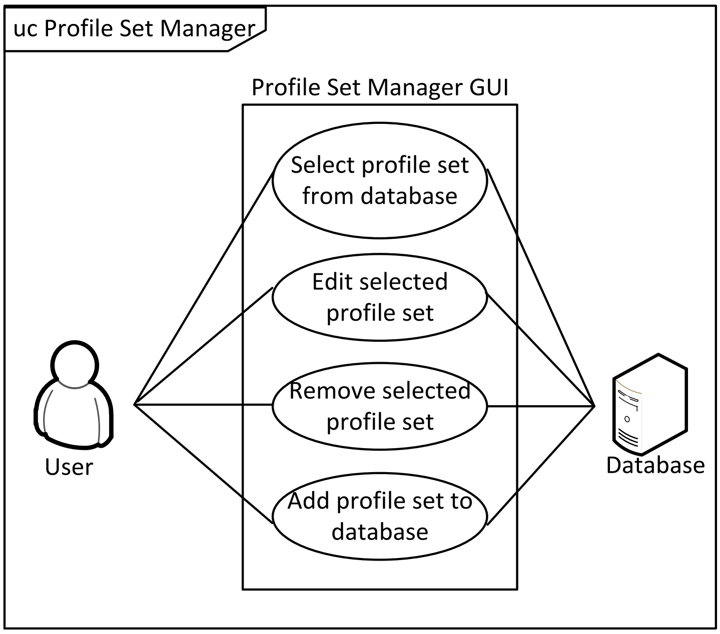

Figure 4.8: Use case diagram of profile set manager software component.

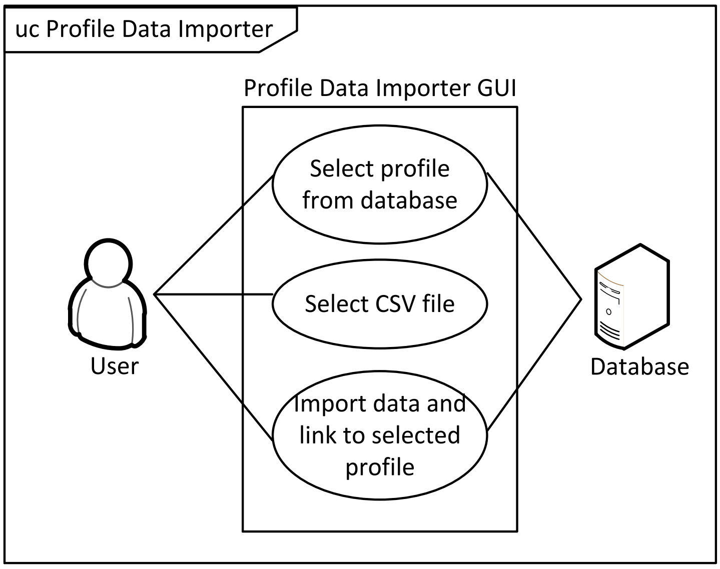

Figure 4.9: Use case diagram of profile data importer software component.

Stellenbosch University https://scholar.sun.ac.za

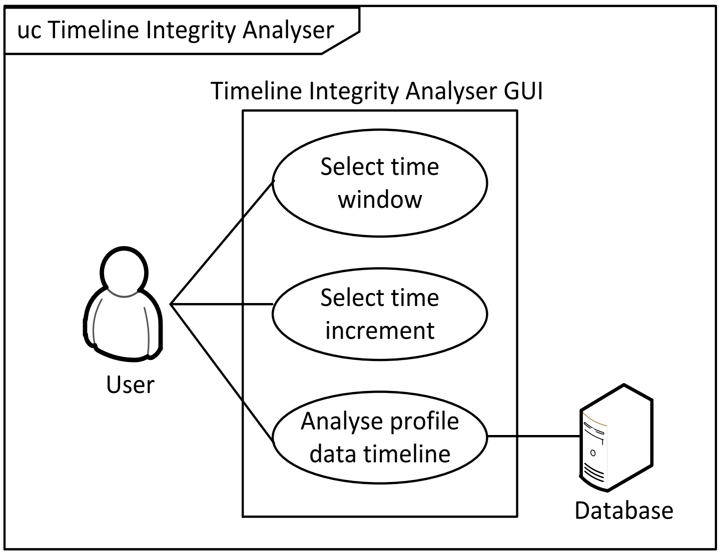

Figure 4.10: Use case diagram of timeline integrity analyser software component.

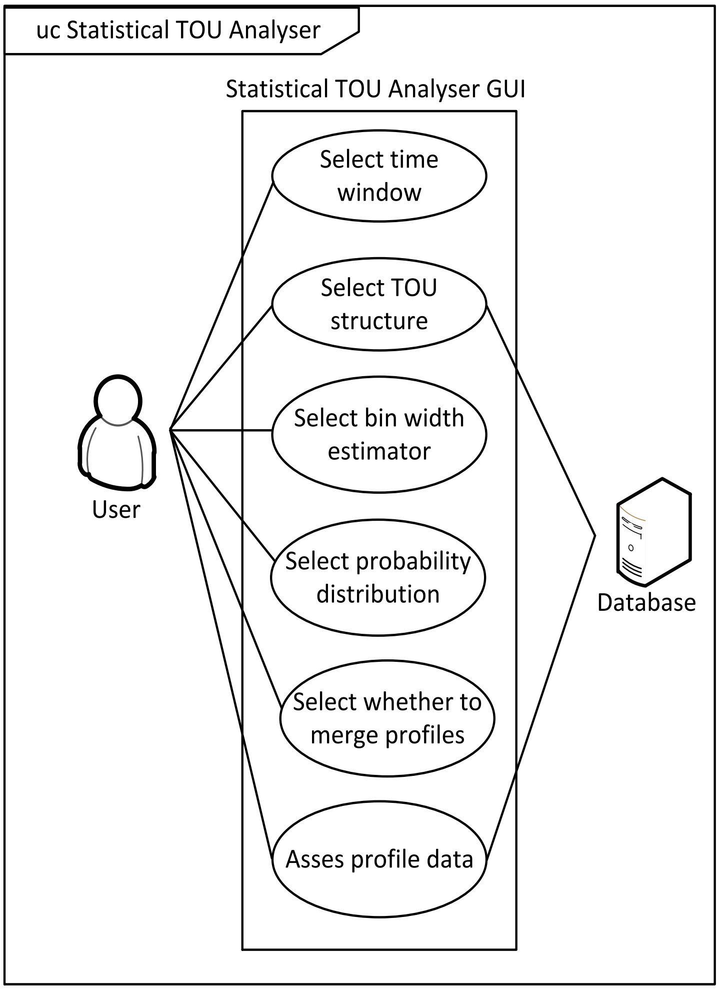

Figure 4.11: Use case diagram of statistical TOU analysis module software component.

Stellenbosch University https://scholar.sun.ac.za

4.4 Construction Phase

The majority of the system is designed and implemented in this phase as well as the final analysis of the system. Essentially this is the phase where the system is implemented. The output of this phase is the implemented system software, design and models.

4.4.1 Design and Implementation

All software components are implemented in Embarcadero's Delphi™ Integrated Development Environment (IDE) for the Microsoft Windows environment as discussed in section 2.3. The system architecture is designed to be modular and extensible with all software components implemented in DLLs. This section deals with the implementation of each software module and its GUI.

4.4.1.1 Main Application

The Main Application (MA) is designed and implemented to be the central hub of the software application and to meet the following requirements:

- Allow the user to establish a persistent database connection.
- Allow the user to create a list of profiles to be used with the multi-select filter.
- Pass the established database connection and user selected list of profiles to PAE modules.
- Pass the established database connection to PAS modules.
- Allow the user to select and use any of the components contained in either the PAE or PAS.

In order for the MA to pass an active database connection between the different software components in different DLLs, a COM interface was designed in implemented in the MA. This interface includes the class definitions and methods of the database connection which allows the MA to instantiate a database connection as an object even when it is not defined in the MA.

To establish a database connection the MA first calls a create method from the database connection interface which instantiates a database connection object in memory. However the connection is not connected to any database until the connection parameters are set. To set the connection parameters the MA calls a method to display the GUI of the database connection. The user then uses the database connection GUI to set the connection parameters and establish a connection to a database. This connection persists as long as the MA is running and could be changed to connect to a different database if desired. This way only one instance of the database connection exists at all times.

Stellenbosch University https://scholar.sun.ac.za

When a user wants to perform an analysis on generation data stored on the database, it is required to first select the data to use. The multi-select filter is designed and implemented to provide the user with an interactive GUI to select any combination of the profiles available on the database. The MA first instantiates and calls the multi-select filter GUI from its respective DLL. Once a user has selected the desired profiles, the multi-select filter creates a list containing the selected profiles and passes it back to the MA as a parameter.

Software components from the PAE and PAS are all database driven and therefore could only be called and used if a database connection is established. Therefore, the MA checks whether an active database connection is defined and exists with each call to a software component. Figure 4.12 presents the activity diagram of a call made to a software component in either the PAS or PAE.

Note that the solid black circle denotes the start of an activity and the encircled black circle denotes the end of an activity. Furthermore, a diamond shape denotes a decision and a rounded rectangle denotes an action.

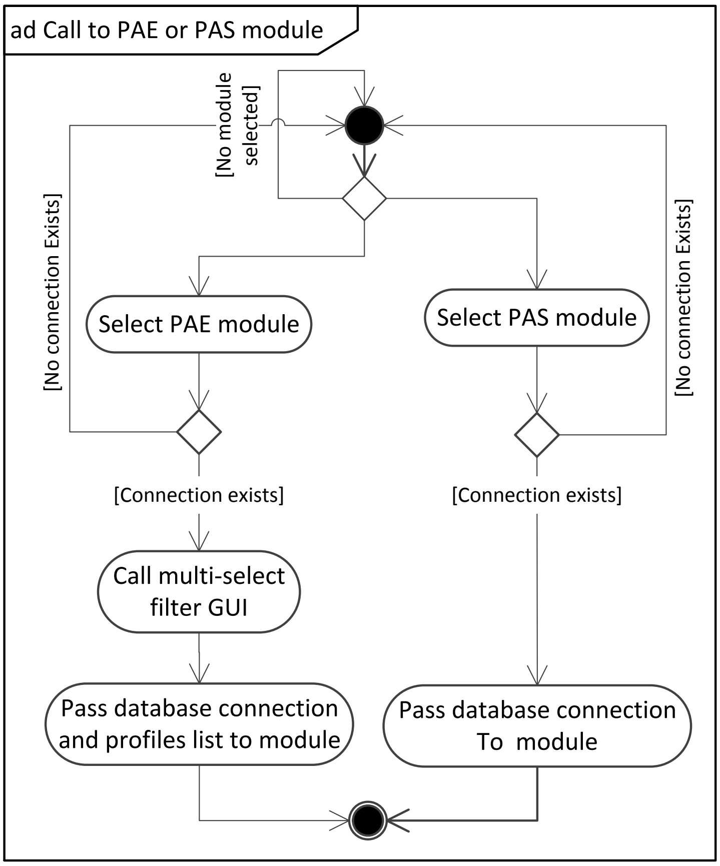

Figure 4.12: Activity diagram of a call made to a software module in the PAS or PAE.

Stellenbosch University https://scholar.sun.ac.za

4.4.1.2 Database Connection

The database connection is designed and implemented to meet the following requirements:

- User defined connections could be stored for reconnection at a later time.
- Stored connections could be edited or deleted.
- Establish a connection to a database using login credentials from user defined connections.
- Execute database queries and commands using a structured query language.

To save the user defined connections and make them available every time the application runs, the connections are saved on the hard drive of the computer running the application. The database connection GUI was designed and implemented to store all user connection information in an INI file format in the directory of the software application EXE. This allows the database connection to read and edit all the previously stored connection information every time the application runs. The database connection GUI also provides the user with the ability to select any of the previously stored connection in the INI file and establish a connection to a database.

A database connection is only established if the server address and database exists and the username and password is correct. The connection parameters include the server IP address, database name, username and user password. Figure 4.13 presents the activity diagram of connecting to the database.

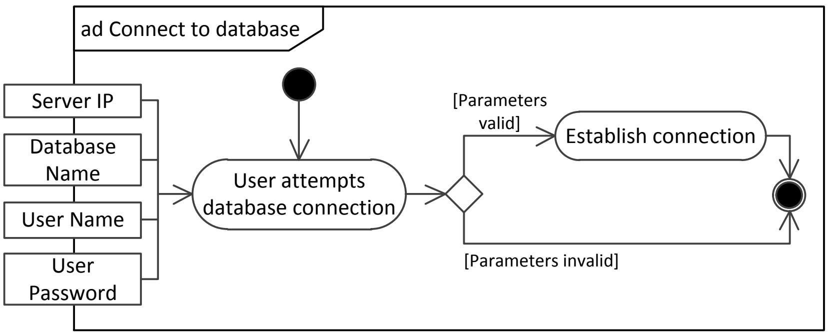

Figure 4.13: Activity diagram of connecting to a database.

The database connection is implemented in a DLL with all methods and class definition included in an interface as mentioned before. The different software components each have different functionalities and therefore different requirements from the database connection. This means that the database connection is required to be very versatile and accommodate all the components' needs. Therefore, the design of the database connection and interface was continually changed during the design of the software application until all requirements were met. The requirements of the database connection and interface were not adequately identified in the elaboration stage as the development

Stellenbosch University https://scholar.sun.ac.za

and actual implementation of all the software components, using the connection, revealed factors that were not considered, i.e. as the software modules were developed and tested, new functionalities were required from the database connection and therefore added. The final database connection provides the software components with the following available database operations:

- Execute database queries and commands.
- Return field values for given field names.
- Return date and time values for given field names.
- Step through records in a resultant dataset.
- Set record number in a resultant dataset.
- Edit records in resultant dataset.
- Add records to resultant dataset.
- Delete records from resultant dataset.
- Apply updates of resultant dataset to database.
- Check connection status.
- Return number of records in a resultant dataset.

Using these database operations all software components are able to achieve their required functionality. The database connection is implemented in the Delphi™ IDE using four dbExpress components namely:

- SQL Connection.
- SQL Dataset.
- Dataset Provider.
- Client Dataset.

The SQL Connection component establishes the actual connection to a database through a database managing system [30]. All queries and commands on a database are executed using the SQL Connection component and all data sent to or received from the database passes through it. The SQL Dataset component is a unidirectional dataset that simply captures the data received from the SQL Connection component in response to a query on the database [30]. The Dataset Provider component simply copies the data contained in the SQL Dataset component to the internal dataset of the Client Dataset component. Therefore, it provides the queried dataset to the Client Dataset component to be used by the software application. The Client Dataset component is a dataset with bidirectional navigation and has the ability to edit data [30]. This is the dataset used by the software application for all access to the database. Therefore, when the software application queries data from the database using the SQL Connection component, the resultant data set is copied to the SQL Dataset and then copied to the Client Dataset via the dataset provider.

Stellenbosch University https://scholar.sun.ac.za

4.4.1.3 Multi-select Filter

The multi-select filter is responsible for creating a list of user selected profiles from a database to be used by the PAE. Therefore, the multi-select filter requires a database connection to query and display all available profiles and profile sets on the database. The multi-select filter GUI is designed to meet the following requirements:

- Select desired project from database.
- Select desired profile sets from database.
- Select desired combination of profiles from database.

When the user has finished selecting all the profiles for an analysis, the multi-select filter compiles a list containing the selected profiles and passes it to the MA as a parameter. The multi-select filter GUI allows the user to select any number and combination of profiles from the database by dynamically creating a custom SQL query according to the filter settings.

There are five fields of the multi-select filter GUI that together allows the user to select any combination of profiles. These fields include the following:

- *Project field*: Projects in the Project table on the database.
- *Profile Set Category*: Profile set categories in the Profile Set Category table on the database.
- *Multi-selection of Profile Sets*: User selected profile sets which results from the Project and Profile Set Category fields of the filter.
- *Profile Category*: Profiles in the Profile table on the database.
- *Multi-selection of Profiles*: User selected profiles which results from all the filter fields.

The project field and profile set selection is mandatory, while profile set category and profile category are optional fields. If the optional profile set category is not selected, the profile sets of all categories are returned. Similarly if the profile category is not selected, profiles of all categories are returned. Figure 4.14 presents the activity diagram of a user using the multi-select filter. Note that the diamond shapes represent the different decisions a user can make using the GUI of the considered software module.

Stellenbosch University https://scholar.sun.ac.za

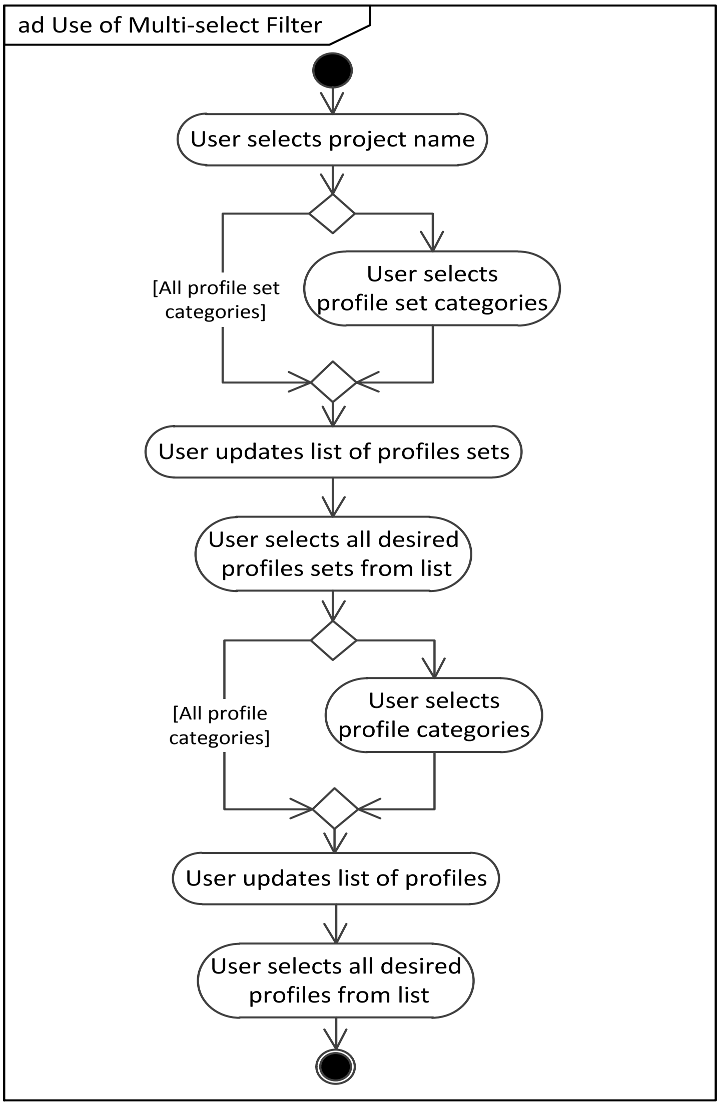

Figure 4.14: Activity diagram of the use of the multi-select filter.

4.4.1.4 Profile Set Manager

The profile set manager GUI is designed to meet the following requirements:

- Select profile set from a database.
- Edit selected profile set.
- Remove selected profile set from a database.
- Add profile set to a database.

The profile and profile set manager modules are developed together in a single DLL with a database connection interface. Although both modules are in a single DLL, they are still called and used individually.

Stellenbosch University https://scholar.sun.ac.za

When a profile set is selected to be edited or when a new profile set is to be added, a simple GUI is displayed which allows the user to enter all profile set data. These changes are then saved on the database using the database connection. Figure 4.15 presents the activity diagram of the use of the profile set manager.

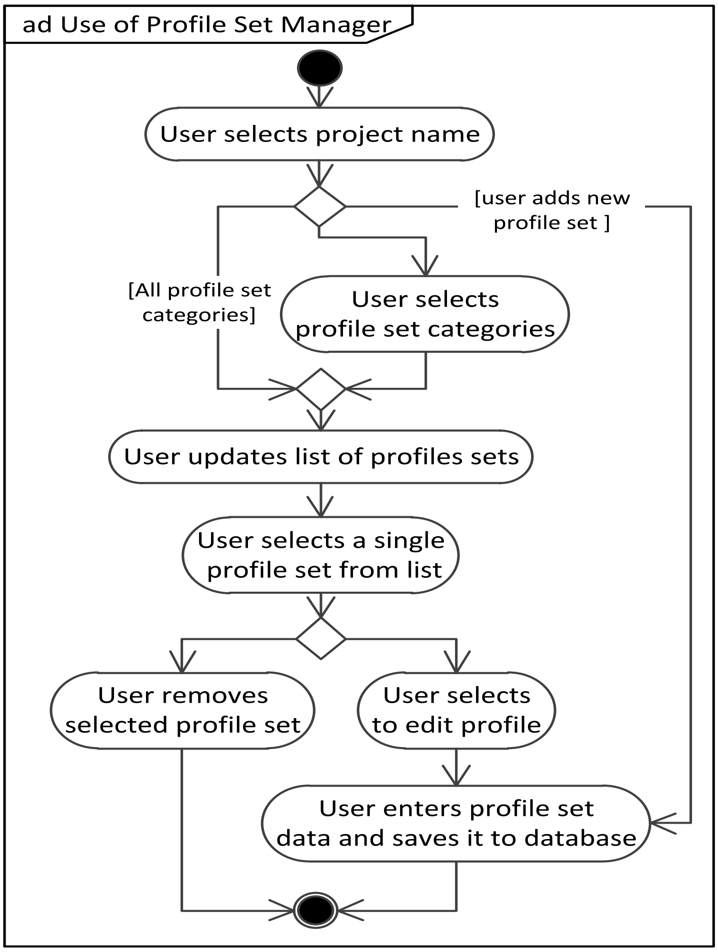

Figure 4.15: Activity diagram of the use of the profile set manager.

4.4.1.5 Profile Manager

The profile manager GUI is designed to meet the following requirements:

- Select profile from a database.
- Edit selected profile.
- Remove selected profile from a database.
- Add profile to a database.
- Remove profile data from a database.

Stellenbosch University https://scholar.sun.ac.za

When a profile is removed from the database the user can select whether the profile data associated with the profile should be removed as well. When a profile is selected to be edited or when a new profile is to be added, a simple GUI is displayed which allows the user to enter all the profile data. These changes are then saved on the database using the database connection. Figure 4.16 presents the activity diagram of the use of the profile manager.

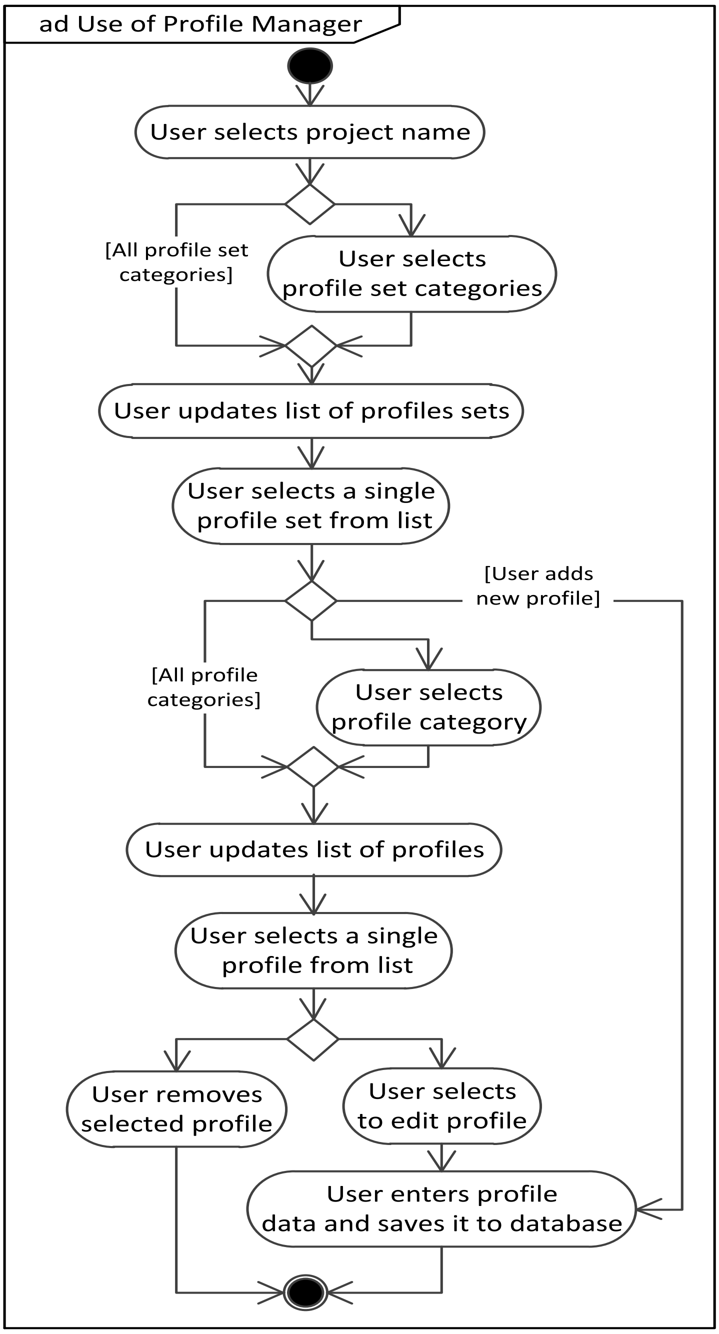

Figure 4.16: Activity diagram of the use of the profile manager.

Stellenbosch University https://scholar.sun.ac.za

4.4.1.6 Profile Data Importer

The profile data importer is developed as a DLL on its own with a database connection interface. The Profile Data Importer is designed to meet the following requirements:

- Select a profile from a database to which imported data should be linked.
- Select a CSV file from the computer running the software application.
- Import data and link it to selected profile.

The Profile Data Importer GUI is designed and implemented to allow the user to select any CSV file to be imported. The GUI allows the user to select the profile to which the data should be linked and to choose which delimiter is used in the CSV file. The format of the CSV files to be imported is very simple and consists of only two columns. The first column represents the timestamps and the second column the respective values.

4.4.1.7 Timeline Integrity Analyser

Profile data in CSV files may have some extra, missing or duplicated timestamps as a result of faults in the logging system. This gives rise to a broken timeline in the generation data which affects the results of any analysis performed on such data. Therefore, it is required to assess the timeline integrity of data to be used in an analysis and determine whether data is usable and intact. The timeline integrity analyser is a component of the PAE subsystem and is developed as a DLL on its own with a database connection interface. The timeline integrity analyser GUI is designed to meet the following requirements:

- Set the timeline of the analysis.
- Set time increment between successive timestamps.
- Analyse profile data timeline.

To analyse a timeline's integrity, a reference timeline must first be created which is then compared to the generation data on the database. To create a reference timeline the user must select a timeline on which the analysis is conducted, as well as the incremental time step between each successive timestamp. The timeline integrity analyser GUI allows the user to select the start date and time, end date and time and the time increments between each successive timestamp. The software component then creates a reference timeline for the timeline selected by the user. Once a reference timeline is created the software module compares it to the generation data on the database that falls in the analysis timeline. The software component then reports back the timestamps which are missing, extra or duplicated.

4.4.1.8 Statistical TOU Analyser

The statistical TOU analysis GUI is designed to meet the following requirements:

- Allow the user to set the timeline of the analysis.
- Allow the user to set the probability distribution used in the analysis.
- Allow the user to set the bin width estimator.
- Allow the user to set the TOU structure.
- Allow the user to set whether profiles should be merged or not.

The statistical TOU analyser GUI allows the user to select one of the following probability distributions:

- Normal distribution.
- Weibull distribution.
- Beta distribution.
- Gamma distribution.
- Exponential distribution.
- Logistic distribution.

The Probability Density Function (PDF) and Cumulative Distribution Function (CDF) of each of the probability distributions are implemented numerically as described in section 2.5.6. Generation data from all the selected profiles and the selected timeline is queried and analysed with the selected probability distribution. The analysis determines the following statistical parameters and goodness of fit test results of each TOU period:

- Cumulative value of generation data values.
- Maximum and minimum of generation data values.
- Mean of generation data values.
- Standard deviation and variance of generation data values.
- Timestamp count.
- Root mean square error of generation data values and selected probability distribution.
- Chi-squared test value of generation data values and selected probability distribution.

The Root Mean Square Error (RMSE) and Chi-squared test values are implemented according to theory presented in section 2.5.5.

Stellenbosch University https://scholar.sun.ac.za

The first step in the statistical TOU analysis process is to determine the total, minimum, maximum, mean, standard deviation, variance and timestamp count of the selected generation data values for the selected timeline.

To perform the goodness of fit tests, the profile data is first divided into a number of mutually exclusive classes called bins. The number of bins and bin widths are determined using either Scott's or Sturges rule as bin width estimators as discussed in section 2.5.4. Once the number of bins is determined, the bin intervals are determined by dividing the range of the selected generation data (maximum value – minimum value) into the number of bins.

Using these bin intervals the expected frequency (count) of each bin, i.e. the expected frequency distribution, is calculated using the CDF of the user selected probability distribution. To calculate the bin counts the bin probabilities must first be determined. This is achieved by using the standard deviation and mean of the selected generation data values as parameters in the selected probability distribution's CDF. If the selected distribution CDF uses parameters other than the standard deviation and mean they are calculated according to theory discussed in section 2.5.6. The bin probability is simply determined by subtracting the value obtained from the CDF for the bin start value from the value obtained from the CDF for the bin end value.

Once the bin probabilities have been determined, the bin counts of each class could be determined by multiplying the bin probability by the timestamp count of the generation data selected for the analysis. This produces the expected frequency distribution (bin count) of each bin. It is imperative that all bin counts be above two as discussed in section 2.5.5.2.

After the expected bin frequencies are calculated, a test is done to check whether any of the bins have an expected frequency (count) lower than two. If one of the bins has an expected bin count lower than two, the number of bins is decreased by one and the expected bin frequencies are recalculated and retested. This is repeated until all bin counts are at least two.

To compare the expected frequency counts against the user selected generation data, the observed count is determined, i.e. the observed frequency distribution. The observed count is simply the number of timestamps that fall in each of the bin intervals determined for the expected frequency. Using the observed and expected frequency distributions the RMSE and Chi-squared test values could be determined as described in section 2.5.5. The process is executed for each TOU period and all user selected profiles. Figure 4.17 presents the activity diagram of the statistical TOU analysis process for one TOU period and one profile.

Stellenbosch University https://scholar.sun.ac.za

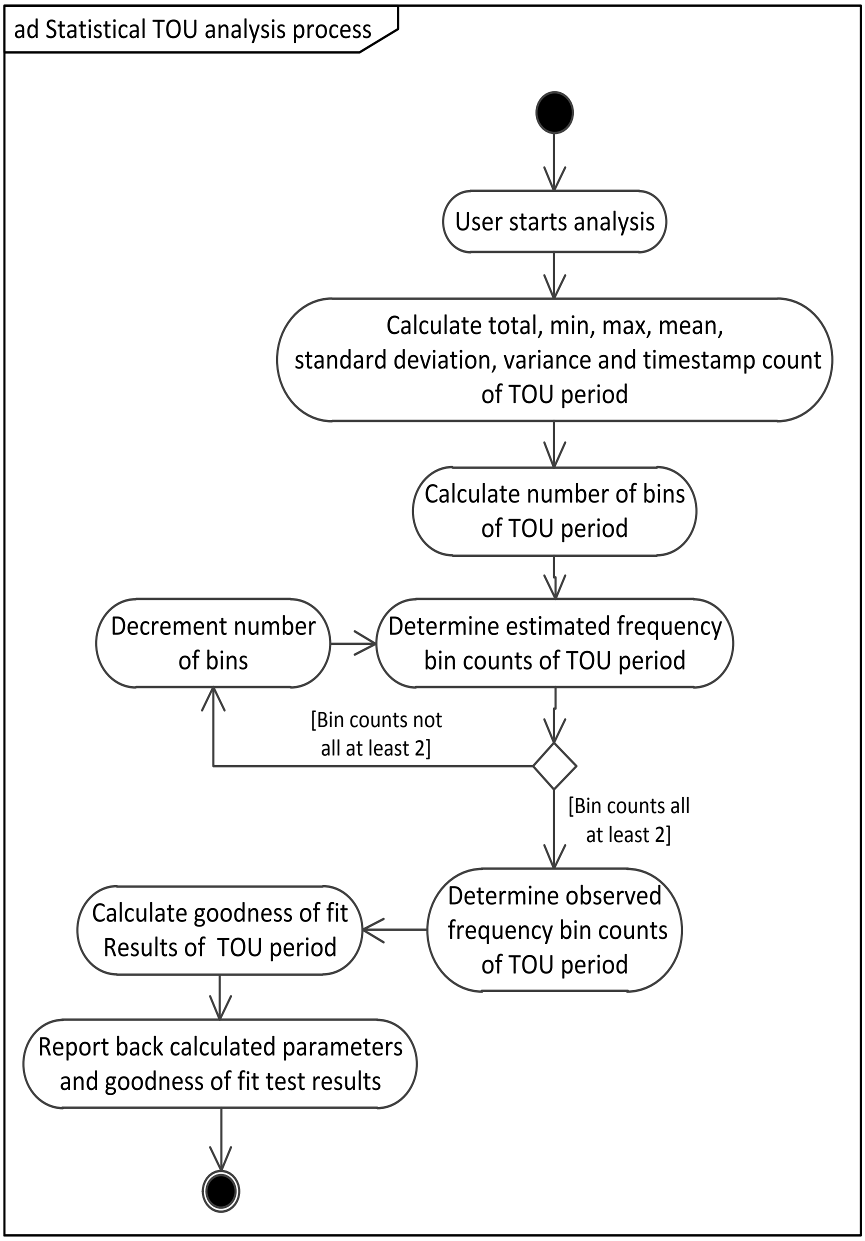

Figure 4.17: Activity diagram of statistical TOU analysis process.

4.4.1.9 Excel Functionality

All PAE software modules were implemented with Excel functionality to export results for further processing. The software modules create new Excel workbooks and populate them with analysis results. The Excel exporting component is implemented using the built in Excel support of Delphi™ IDE.

Stellenbosch University https://scholar.sun.ac.za

4.4.2 System Analysis and Testing

Testing implemented software is important to ensure that the software application is reliable. The testing discipline of the Unified Process involves the designing, implementing and the evaluation of tests on implemented software. The implementation and testing of each module was an iterative process which continued until the software components met their requirements.

4.4.2.1 Test Setup

In order to test the implemented software a test database was created according to the database structure presented in Chapter 3 and implemented in WAMPServer. A test TOU structure was created and implemented in the database. To test database queries and commands a third party MySQL Query Browser application is used which provides a graphical presentation of queried data. This application is used to query datasets from the test database and manually check whether results obtained from the application is correct.

4.4.2.2 Main Application and Database connection

The first two functions of the Main Application (MA) to be tested is its ability to call methods from an external DLL and pass parameters to an external DLL. A simple test DLL which only displays a string of text which was passed to it as a parameter from an external application is developed. The MA is then used to call the method and pass a string of text to be displayed.

To test the MA's ability to create and pass a database connection, the database connection DLL is first designed and implemented. The MA is then used to call and create a database connection from its DLL. The created database connection is passed back to the MA and used to execute a simple query on the test database.

The MA is responsible for calling the Multi-select Filter module in order to select the profile data to be analysed. The MA's ability to call the Multi-select Filter module was tested by adding test profile sets, profiles and profile data to the database and calling the Multi-select filter using the MA. The test data is then selected using the Multi-select filter and passed back to the MA.

Furthermore, the MA is responsible for calling all PAE and PAS modules and pass on the required parameters. The MA's ability to call these modules was tested as the modules were developed by calling the GUIs of modules from the MA.

Stellenbosch University https://scholar.sun.ac.za

4.4.2.3 Profile Set and Profile Manager

The profile set manager and profile manager are required to allow the user to select or add profile sets and profiles from a database. Furthermore these software modules are required to allow the user to edit or remove the profile sets and profiles stored on the database.

To test whether the profile set and profile manager met the set requirements, an established database connection is passed to the two modules and then used to populate the test database with profiles and profile sets. The software modules were then used to edit and remove the profile sets and profiles added to the test database. Results are checked using the third party MySQL query browser application.

4.4.2.4 Profile Data Importer

The profile data importer is required to allow the user to select a profile from the database and a CSV file from a computer. The profile data importer is the required to import all data in the selected CSV file into the selected profile stored on the database.

To test the profile data importer a test CSV file is created which is populated with known time-stamped values. The profile data importer is then used to select a profile from the database and import the test CSV file into the database and link it to one of the test profiles. Results are checked using the third party MySQL query browser application.

4.4.2.5 Timeline Integrity Analyser

The timeline integrity analyser is required to allow the user to select a time window in which the analysis should take place as well as a time increment between successive time stamps. The software module is then required to analyse the integrity of all profile data from all profiles passed to it from the multi-select filter.

To test the timeline integrity analyser, a number of test CSV files are created. The CSV files are populated with certain timelines. The timelines are altered in different ways in each separate CSV file to have intact, missing, duplicated and extra timestamps. These CSV files are then imported to the test database and connected to different profiles. The timeline integrity analyser is then used set a time window and increment and to test each different set of timestamps to determine whether correct results are obtained.

Stellenbosch University https://scholar.sun.ac.za

4.4.2.6 Multi-select Filter

The multi-select filter is required to allow the user to select a project stored on the database. All profile sets linked to the selected project is then provided to the user. The multi-select filter is then required to allow the user to select any number and combination of profile sets and provide the user with all profiles connected to the selected profile sets. The multi-select filter is then required to allow the user to select any number and combination of profiles and pass the selection on to the next software module.

To test the multi-select filter, it is simply passed an established database connection. All available profiles and profile sets on the multi-select filter are compared to the profiles and profile sets that were previously added to the test database. The list of profiles returned by the multi-select filter is then compared to the profiles selected by the user.

4.4.2.7 Statistical TOU Analyser

The statistical TOU analyser is required to allow the user to select the time window, TOU structure, bin width estimator and probability distribution to use in the analysis. The profile data from the selected profile sets(received from multi-select filter) are then analysed according to the user selected parameters.

To test the statistical TOU analyser, a CSV file with test data is imported into the test database. The statistical TOU analyser is then used to select the desired time window, TOU structure, bin width estimator and probability distribution to use in the analysis. All selections are checked against the database and software application. The third party MySQL query browser application is used to query all relevant generation data and the statistical indicators are calculated using Excel. These results are then compared to the statistical TOU analyser results.

4.4.3 System Analysis

The implemented system proves to be successful and meets all requirements. The implemented system architecture is well suited for developing and testing functional modules. Furthermore, the extensible and modular nature of the software application proves to be a code and time efficient development approach.

Stellenbosch University https://scholar.sun.ac.za

4.5 Transition Phase

During this phase the system is moved to the user's environment. This includes deploying and maintaining the system. This is the final phase of a cycle therefore the output is the final release of the system.

The final stage of the development process is deploying the software application and using it on the case study. The software application and all required DLLs are deployed on a Microsoft Windows computer running WAMPServer.

A new database is created for the case study according to the structure discussed in Chapter 3. The case study project name, profiles and profile sets are added to the case study database along with all relevant TOU structures. The historical generation data is compiled and imported into the case study database. All imported generation data is analysed using the timeline integrity analyser to ensure data integrity.

Stellenbosch University https://scholar.sun.ac.za

5 Solar Plant Case Study Results

5.1 Overview

A case study is conducted to test and evaluate the implemented forecasting methodology and software application. The main objective is to determine whether it is possible to fit historical generation data to the proposed probability distributions commonly used to model solar irradiation. Furthermore, the case study aims to illustrate how the analysis results are also useful in assessing the performance of a solar plant for varying TOU structures. The historical generation data available at the time of this project is very limited. Therefore, this case study does not focus on the statistical analysis of the available data, but rather to test whether the forecasting methodology is successful in the TOU context. The case study is conducted for an operational solar plant which provides supplementary energy to a cold storage facility in the Western Cape province of South Africa. The energy supplied by the solar plant is used to mitigate the cold storage's energy usage from the utility provider (Eskom) in order to reduce electricity expenses.

5.2 Solar Plant System Configuration

5.2.1 Panel and Inverter Configuration

The cold storage facility consists of two separate buildings each with a subsystem of PV panels, meters and inverters. Both buildings have dual pitched roofs which are painted white to reflect incoming solar radiation. All PV panels are mounted flush against the roofs of the buildings at fixed positions and orientations. The buildings are several stories high and provide unobstructed and un-shaded solar radiation to the PV panels.

The solar plant is comprised of a grid connected Photovoltaic (PV) system which implements string technology, as discussed in section 2.7.3.1. The solar plant supplies only a fraction of the cold storage facility's total energy demand while the remainder is satisfied by the utility provider. The two subsystems of the solar plant are summarised in in table 5.1 while figure 5.1 presents the PV system configuration.

Table 5.1: Solar plant subsystems summary.

	Subsystem 1	**Subsystem 2**	**Total**
Number of PV panels	1387	730	2117
Number of inverters	19	10	29

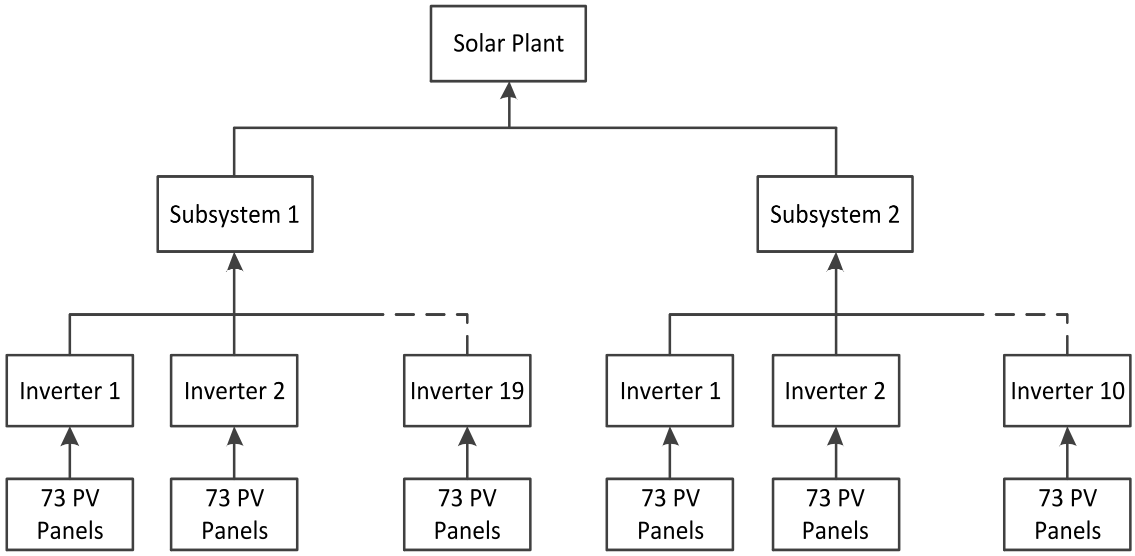

Figure 5.1: PV system configuration.

The inverters convert DC power from the PV panels into AC power to be used by the facility. Maximum Power Point Trackers (MPPTs) incorporated into the inverters keep the PV panels at an efficient operation point. All PV panels are connected in a number of series configurations to obtain the optimal voltage supply for the MPPTs. The inverter model used is SMA's STP-17000TL model with technical specifications given in table 5.2 [68].

Table 5.2: Inverter technical specifications.

DC Input	
Maximum DC Power at cos φ =1	17 410 W
Rated Input Voltage	600 V
AC Output	
Rated AC Power at 230V and 50 Hz	17 000 W
Maximum Apparent AC Power	17 000 VA
Maximum Output Current	24.6 A
Maximum Efficiency	98.2 %

5.2.2 PV Panel Array Rated Energy Output

The solar plant's PV panel array consists of 2117 PV panels rated at 240 Watt (W) peak under standard test conditions. The rated power output of the solar plant is 508.08 kW. Therefore, the rated energy output of the PV panel array during any given half-hour interval is 254.04 kWh.

Stellenbosch University https://scholar.sun.ac.za

5.3 Analysis Methodology

5.3.1 Data Acquisition

The software application developed in this project was done in conjunction with historical generation data which was made available by an industrial consumer with onsite solar generation. The industrial consumer measured and logged the power output of the PV system at half-hourly intervals with a number of power meters. The historical power output data provided in kW was used to determine the energy output in kWh for each timestamp over the analysis timeline.

A relational database is implemented according to the structure presented in Chapter 3 and populated with all profiles, profile sets and historical energy output data provided by the industrial consumer. The serial numbers of the power meters are added to the database as the different profile sets. The historical energy generation data of each individual meter is added as separate profiles and linked to their respective profile sets. The timeline integrity of historical generation data of each meter is checked as it is added to the database.

5.3.2 Data Analysis

Due to the fact that this research project has to be completed within a designated timeframe, only historical generation data from the start of February 2013 to the end of June 2014 is considered for the case study analysis. All individual profile sets (meters) are merged together to represent the solar plant as one complete system. The historical generation data is analysed on a tariff seasonal and monthly basis as depicted in figure 5.2.

The main objective of the seasonal analysis is to determine whether it is indeed possible to create long term energy output models within TOU tariff structure seasons. Energy output models within TOU structure seasons are particularly useful in determining the profitability of a solar plant during the utility defined seasons as the tariffs, i.e. the price paid per kWh for energy consumed, differ for these seasons. The seasonal analysis is conducted with respect to two Eskom tariff seasons, namely the High Demand and Low Demand season. The High Demand season consists of the calendar months of June to August, while the Low Demand season consists of the remaining calendar months of the year. The historical generation data for the entire analysis timeline is divided into two datasets, with a dataset for each seasonal interval. Each dataset is analysed with respect to a half-hourly generation profile, the MegaFlex and the HomeFlex TOU tariff structures.

The monthly analysis is conducted to determine whether it is possible to create long term energy output models of a solar plant with respect to the calendar months of a year. This is done due to the fact that the weather patterns differ for each month. The monthly analysis is conducted with respect to the calendar months of February and June to. The historical generation data of these two months for the entire analysis timeline is divided into two datasets, with a dataset for each monthly interval. Each dataset is analysed with respect to a half-hourly generation profile and the HomeFlex TOU structure.

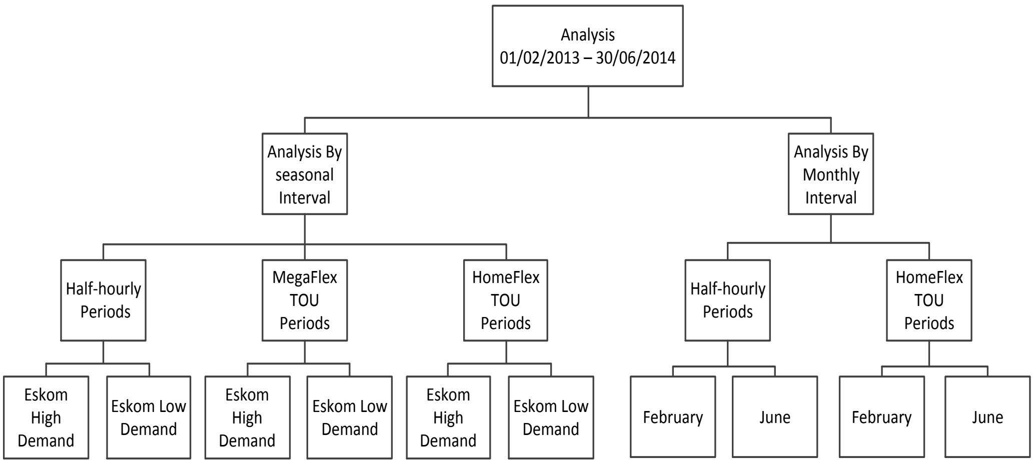

Figure 5.2: Case study analysis hierarchy.

Each individual analysis is divided into two sections. The first section presents the statistical parameters of the relative analysis datasets, while the second section aims at deriving statistical models.

The Chi-squared and Root Mean Square Error (RMSE) goodness of fit tests are employed to determine whether a hypothesised probability distribution fits the historical generation data for a given TOU period. The Chi-squared test with a 1% level of significance is regarded as the primary test criterion for this case study. The RMSE test acts as a supplementary indication of model performance, i.e. how well a probability distribution fits historical generation data. The goodness of fit tests are implemented according to theory discussed in section 2.5.5.

5.4 Analysis Results

It is decided to normalise all analysis results in order to evaluate and compare the solar plant’s performance between different seasons and months of the year. Analysis results are provided in per unit [p.u.] values and are normalised to the rated energy output of the entire solar plant.

Stellenbosch University https://scholar.sun.ac.za

5.4.1 Seasonal Analysis

This analysis is conducted with respect to the High and Low Demand seasons from the start of February 2013 to end of June 2014. During the analysis timeline a total of 122 days fall in the High Demand season and a total of 393 days fall in the Low Demand season.

5.4.1.1 Daily Half-hourly Generation Profile

This analysis is conducted for the daily half-hourly generated energy profile during the High and Low Demand seasons. All night time half-hourly intervals are excluded from the presented results.

5.4.1.1.1 Statistical Parameters

Tables 5.3 and 5.4 summarise the statistical parameters of the daily half-hourly generated energy during the High and Low Demand seasons for the entire analysis timeline. The maximum, average and standard deviation of the daily generated energy is normalised to the rated energy output of 254.04 kWh per half-hour interval.

Table 5.3: Statistical parameters of daily generated energy for half-hourly profile during the High Demand season.

Period start	Maximum energy[p.u.]	Average energy [p.u.]	Standard deviation of energy[p.u.]
07:00:00	0.013	0.001	0.003
07:30:00	0.105	0.012	0.016
08:00:00	0.228	0.045	0.036
08:30:00	0.250	0.095	0.057
09:00:00	0.352	0.152	0.079
09:30:00	0.443	0.218	0.102
10:00:00	0.517	0.281	0.119
10:30:00	0.572	0.328	0.141
11:00:00	0.616	0.376	0.142
11:30:00	0.658	0.405	0.146
12:00:00	0.662	0.413	0.147
12:30:00	0.670	0.411	0.154
13:00:00	0.756	0.399	0.159
13:30:00	0.642	0.384	0.151
14:00:00	0.646	0.353	0.151
14:30:00	0.621	0.318	0.142
15:00:00	0.551	0.277	0.126
15:30:00	0.437	0.206	0.107
16:00:00	0.346	0.142	0.082
16:30:00	0.261	0.078	0.057
17:00:00	0.139	0.027	0.032
17:30:00	0.026	0.004	0.006

Stellenbosch University https://scholar.sun.ac.za

Table 5.4: Statistical parameters of daily generated energy for half-hourly profile during the Low Demand season.

Period start	Maximum energy[p.u.]	Average energy [p.u.]	Standard deviation of energy[p.u.]
05:30:00	0.047	0.003	0.007
06:00:00	0.123	0.014	0.023
06:30:00	0.189	0.040	0.052
07:00:00	0.294	0.089	0.086
07:30:00	0.402	0.160	0.116
08:00:00	0.511	0.241	0.139
08:30:00	0.601	0.322	0.154
09:00:00	0.684	0.400	0.164
09:30:00	0.770	0.469	0.171
10:00:00	0.804	0.528	0.178
10:30:00	0.852	0.573	0.181
11:00:00	0.880	0.610	0.184
11:30:00	0.909	0.627	0.191
12:00:00	0.907	0.635	0.196
12:30:00	0.904	0.638	0.201
13:00:00	0.919	0.629	0.202
13:30:00	0.889	0.623	0.198
14:00:00	0.865	0.589	0.200
14:30:00	0.816	0.544	0.199
15:00:00	0.785	0.497	0.194
15:30:00	0.750	0.437	0.186
16:00:00	0.651	0.366	0.177
16:30:00	0.602	0.288	0.163
17:00:00	0.499	0.206	0.143
17:30:00	0.399	0.127	0.113
18:00:00	0.239	0.064	0.075
18:30:00	0.126	0.022	0.033
19:00:00	0.041	0.003	0.006

Figures 5.3 and 5.4 present the maximum and average daily generated energy from tables 5.3 and 5.4 in graphic format. The results indicate a greater availability of solar power during the Low demand season than that of the High Demand season. This is to be expected, as the Low Demand season consists of summer months while the High Demand season consists of winter months.

The results show the effect of different sunrise and sunset times for winter and summer months. Figures 5.3 and 5.4 indicate a one and a half hour shift in the solar profile between the High and Low Demand seasons. This lack of energy availability for morning and evening hours during winter months is one of the drawbacks of solar power, as these hours coincide with times of peak energy demand.

Stellenbosch University https://scholar.sun.ac.za

The results shown in figures 5.3 and 5.4 indicate a considerable difference in the solar plant's performance during the High and Low Demand seasons. The solar plant performs significantly better during the Low Demand season than during the High Demand season. This significant difference between the High and Low demand seasons is attributed to the path taken by the earth around the sun, weather conditions and the fact that the solar panels have fixed orientations and tilts. Fixed orientations and tilts of solar panels are inefficient and lead to decrease generation of energy during certain times of the year.

The average daily generated energy during the Low Demand season is about the same as the maximum daily generated energy of the High Demand Season. Furthermore, the peak average generated energy reaches a per unit value of about 64% during the Low Demand season, while reaching a per unit value of about only 41% during the High Demand Season.

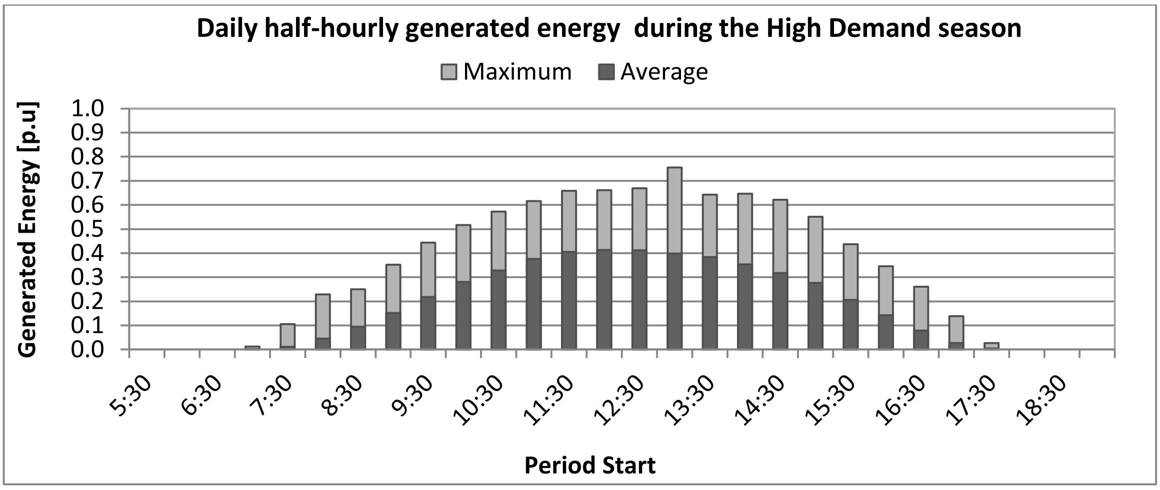

Figure 5.3: Daily average and maximum half-hourly generated energy for the High Demand season.

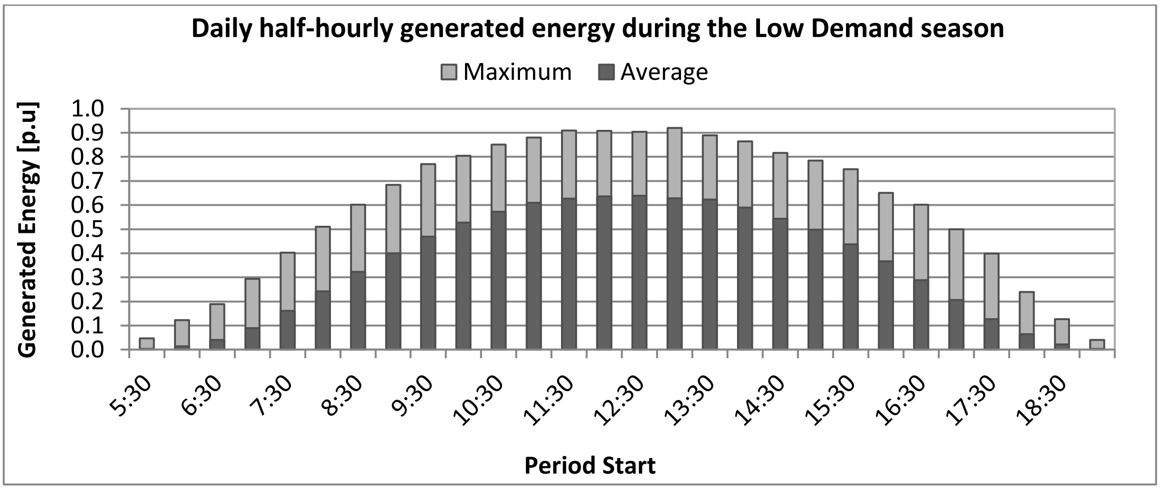

Figure 5.4: Daily average and maximum half-hourly generated energy for the Low Demand season.

Figure 5.5 presents the Coefficients Of Variation (COV) of the generated energy for each half-hour interval during the High and Low Demand seasons. The COV is a normalised measure of variation and is determined by dividing each half-hour interval's standard deviation by the average daily generated energy.

Figure 5.5 indicates that the daily generated energy varies more during the High Demand season than during the Low demand season. The results also show that the generated energy varies significantly for morning and evening hours with COV well over a 100%, and evens out towards mid-day hours with COV between 30% and 40%.

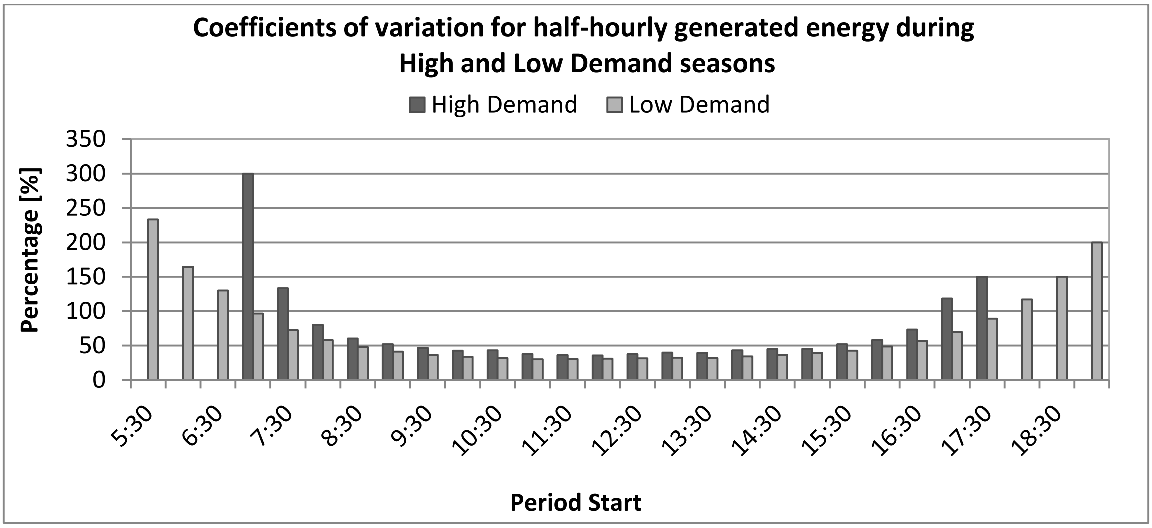

Figure 5.5: Coefficients of variation of daily half-hourly generated energy for High and Low Demand seasons.

5.4.1.1.2 Statistical Model

The Chi-squared and RMSE goodness of fit tests are employed to determine whether a hypothesised probability distribution could model a given half-hour interval. The Chi-squared test is regarded as the primary goodness of fit test criterion. The RMSE test is used as a supplementary indication of model performance. The Chi-squared test and RMSE values for the daily half-hourly analysis are provided in appendix B.

Tables B.1 and B.2 in appendix B summarise the results of the Chi-squared test for the daily half-hourly analysis during the High and Low Demand seasons. The Chi-squared test results for each distinct half-hour interval are given with respect to all the considered probability distributions. Each probability distribution's Chi-squared results are given as a value and a number of bins, with the

Stellenbosch University https://scholar.sun.ac.za

number of bins denoted by NB. The Chi-squared goodness of fit test involves determining the Degrees Of Freedom (DOF) from the number of bins as discussed in section 2.5.5.2.

Each distinct half-hour interval is best described by the probability distribution with the lowest Chi-squared test value. However, the resulting DOF calculated from the number of bins NB are required to be at least one in order for the result to be useful and conclusive. If the probability distribution with the lowest Chi-squared test value has resultant DOF of less than one, the probability distribution with the second lowest Chi-squared value is considered and so on.

To determine whether a hypothesised probability distribution should be accepted or rejected, the percentage points table for the Chi-squared distribution provided in Appendix A is used. The resultant DOF from the analysis and a level of significance of 1 % are used to read off the corresponding percentage point value of the Chi-squared distribution. If the analysis Chi-squared test value is lower than the corresponding percentage point value, the hypothesised probability distribution is accepted. If Chi-squared test value is larger the than the corresponding percentage point value, the probability distribution is rejected.

The probability distribution with the lowest Chi-squared test value and resultant DOF of at least one is chosen as the best performing conclusive result for this analysis. If all probability distributions result in DOF of less than one for a specific result, the result is regarded as inconclusive. Note that the best performing conclusive probability distribution, i.e. the probability distribution with the lowest Chi-squared test value and DOF of at least one, does not necessarily represent the best fitting probability distribution.

Tables 5.5 and 5.6 summarise the best performing conclusive models for the half-hourly analysis during the High and Low Demand seasons together with the model conclusions, i.e. whether the hypothesised probability distribution fits the observed historical generation data (accepted) or not (rejected). Chi-squared values for the best performing conclusive models (probability distribution) are provided together with the determined degrees of freedom denoted by DOF.

The results presented in table 5.5 show that a large portion of the half-hourly models for the High Demand season are rejected. The Chi-squared test and RMSE results given in tables B.1 and B.3 indicate that the Beta probability distribution performs the best of all considered probability distributions for the majority of the half-hourly intervals from 08:30:00 to 15:30:00. The Chi-squared test results given in table B.1 indicate that all the considered probability distributions have high resultant DOF and high Chi-squared test values for these half-hour intervals and are therefore conclusively rejected.

Stellenbosch University https://scholar.sun.ac.za

Table 5.5 indicates that only four intervals are successfully modelled for the daily half-hourly generated energy during the High Demand season. Furthermore, the Chi-squared test is inconclusive for two of the half-hour intervals, i.e. none of the considered probability distributions resulted in DOF of more than one.

Table 5.5: Best performing conclusive models for daily half-hourly profile during the High Demand season.

Period Start	Chi-squared value	DOF	Probability distribution	Average energy[p.u.]	Standard deviation of energy [p.u.]	Model conclusion
07:00:00	1.160	3	Beta	0.001	0.003	Accept
07:30:00	0	< 1	Inconclusive	0.012	0.016	Inconclusive
08:00:00	0	< 1	Inconclusive	0.045	0.036	Inconclusive
08:30:00	16.816	4	Weibull	0.095	0.057	Reject
09:00:00	22.311	4	Beta	0.152	0.079	Reject
09:30:00	32.113	4	Beta	0.218	0.102	Reject
10:00:00	35.534	3	Beta	0.281	0.119	Reject
10:30:00	37.256	3	Beta	0.328	0.141	Reject
11:00:00	29.511	4	Beta	0.376	0.142	Reject
11:30:00	22.566	4	Beta	0.405	0.146	Reject
12:00:00	26.389	4	Beta	0.413	0.147	Reject
12:30:00	31.019	4	Beta	0.411	0.154	Reject
13:00:00	46.616	4	Normal	0.399	0.159	Reject
13:30:00	31.762	4	Beta	0.384	0.151	Reject
14:00:00	13.196	3	Beta	0.353	0.151	Reject
14:30:00	47.782	4	Beta	0.318	0.142	Reject
15:00:00	22.810	4	Beta	0.277	0.126	Reject
15:30:00	19.243	3	Beta	0.206	0.107	Reject
16:00:00	10.620	3	Beta	0.142	0.082	Accept
16:30:00	9.627	5	Exponential	0.078	0.057	Accept
17:00:00	7.466	1	Exponential	0.027	0.032	Reject
17:30:00	5.570	4	Beta	0.004	0.006	Accept

The results presented in table 5.6 show that a large portion of the half-hourly models for the Low Demand season are rejected. The Chi-squared test and RMSE results given in tables B.2 and B.4 indicate that the Beta probability distribution performs the best, i.e. has the lowest Chi-squared test values, of all considered probability distributions for the majority of the half-hourly intervals from 09:30:00 to 18:00:00. The Chi-squared test results given in table B.2 indicate that all the probability distributions have high resultant DOF and high Chi-squared test values for these half-hour intervals and are therefore conclusively rejected. Furthermore, table 5.6 indicates that only seven half-hour intervals are successfully modelled for the daily half-hourly generated energy during the Low Demand season.

Table 5.6: Best performing conclusive models for daily half-hourly profile during the Low Demand season.

Period Start	Chi-squared value	DOF	Probability distribution	Average energy[p.u.]	Standard deviation of energy [p.u.]	Model conclusion
05:30:00	25.700	1	Beta	0.003	0.007	Reject
06:00:00	27.192	4	Beta	0.014	0.023	Reject
06:30:00	11.674	5	Beta	0.040	0.052	Accept
07:00:00	3.679	5	Beta	0.089	0.086	Accept
07:30:00	5.184	5	Beta	0.160	0.116	Accept
08:00:00	11.476	5	Beta	0.241	0.139	Accept
08:30:00	2.459	6	Beta	0.322	0.154	Accept
09:00:00	12.205	6	Beta	0.400	0.164	Accept
09:30:00	42.590	7	Beta	0.469	0.171	Reject
10:00:00	56.123	7	Beta	0.528	0.178	Reject
10:30:00	58.626	5	Beta	0.573	0.181	Reject
11:00:00	72.423	5	Beta	0.610	0.184	Reject
11:30:00	68.621	5	Beta	0.627	0.191	Reject
12:00:00	39.772	5	Beta	0.635	0.196	Reject
12:30:00	45.641	6	Beta	0.638	0.201	Reject
13:00:00	67.064	6	Beta	0.629	0.202	Reject
13:30:00	68.869	7	Beta	0.623	0.198	Reject
14:00:00	56.673	7	Beta	0.589	0.200	Reject
14:30:00	46.177	6	Beta	0.544	0.199	Reject
15:00:00	35.640	6	Beta	0.497	0.194	Reject
15:30:00	55.578	6	Beta	0.437	0.186	Reject
16:00:00	22.008	5	Beta	0.366	0.177	Reject
16:30:00	42.200	5	Beta	0.288	0.163	Reject
17:00:00	33.404	5	Beta	0.206	0.143	Reject
17:30:00	33.912	5	Beta	0.127	0.113	Reject
18:00:00	21.406	4	Beta	0.064	0.075	Reject
18:30:00	10.370	5	Beta	0.022	0.033	Accept
19:00:00	29.990	1	Weibull	0.003	0.006	Reject

5.4.1.2 MegaFlex Tariff Structure

This analysis uses the historical generation data of the entire analysis timeline for each respective tariff day of the MegaFlex tariff structure, i.e. the historical generation data of every day of the week is analysed with respect to the tariff periods of each distinct tariff day of the MegaFlex tariff structure. The MegaFlex tariff structure is discussed in section 2.6.

5.4.1.2.1 Statistical Parameters

Tables 5.7 and 5.8 summarise the statistical parameters of the daily generated energy for the MegaFlex tariff structure during the High and Low Demand season. The maximum, average and standard deviation of the daily generated energy is normalised to the rated energy output of each tariff period.

Stellenbosch University https://scholar.sun.ac.za

Table 5.7: Statistical parameters of daily generated energy for MegaFlex during the High Demand season.

Tariff day	Tariff period	Maximum energy [p.u.]	Average energy [p.u.]	Standard deviation of energy [p.u.]
Weekdays	Evening Off-peak	0	0	0
	Morning Standard	0	0	0
	Morning Peak	0.210	0.087	0.045
	Afternoon Standard	0.489	0.275	0.099
	Evening Peak	4.960E-04	2.703E-05	8.922E-05
	Evening Standard	0	0	0
Saturday	Evening Off-peak	0	0	0
	Morning Standard	0.358	0.191	0.075
	Afternoon Off-peak	0.457	0.251	0.097
	Evening Standard	0	2.703E-05	8.922E-05
Sunday	Off-peak	0.188	0.103	0.038

Table 5.8: Statistical parameters of daily generated energy for MegaFlex during the Low Demand season.

Tariff day	Tariff period	Maximum energy [p.u.]	Average energy [p.u.]	Standard deviation of energy [p.u.]
Weekdays	Evening Off-peak	0	0	0
	Morning Standard	0.156	0.027	0.037
	Morning Peak	0.534	0.280	0.131
	Afternoon Standard	0.752	0.495	0.161
	Evening Peak	0.095	0.022	0.028
	Evening Standard	0	0	0
Saturday	Evening Off-peak	0.016	0.003	0.004
	Morning Standard	0.662	0.402	0.142
	Afternoon Off-peak	0.724	0.465	0.162
	Evening Standard	0.095	0.022	0.028
Sunday	Off-peak	0.319	0.203	0.070

Figures 5.6 and 5.7 present the maximum and average daily generated energy from tables 5.7 and 5.8 in graphic format. Figure 5.6 clearly shows the effect of the late sunrise and early sunset times of the winter months during the High Demand season. The results indicate that no energy is generated during the weekday evening off-peak, morning standard or evening peak tariff periods. Furthermore, the results show that no energy is generated during the Saturday evening off-peak and standard tariff periods. This is not optimal as peak and standard tariff period charges are the most expensive and therefore offer the greatest opportunity for monetary savings.

The results shown in figure 5.7 indicate that energy is generated in the majority of the tariff periods during the Low Demand season. A large amount of energy is generated during the morning peak, morning standard and afternoon standard tariff periods. Therefore, the MegaFlex tariff structure offers a great opportunity for monetary savings during the Low Demand season.

Figures 5.6 and 5.7 indicate that the greatest amount of energy is generated in the morning peak and afternoon standard tariff periods for weekdays. The average daily generated energy in these tariff periods vary between about 9% and 28% of the rated tariff period energy for the High demand season, while varying between about 28% and 50% for the Low Demand season. Furthermore, the results show that the greatest amount of energy is generated in the morning standard and afternoon off-peak tariff periods for Saturdays. The average daily generated energy in these tariff periods vary between about 19% and 25% of the rated tariff period energy for the High demand season, while varying between about 40% and 47% for the Low Demand season.

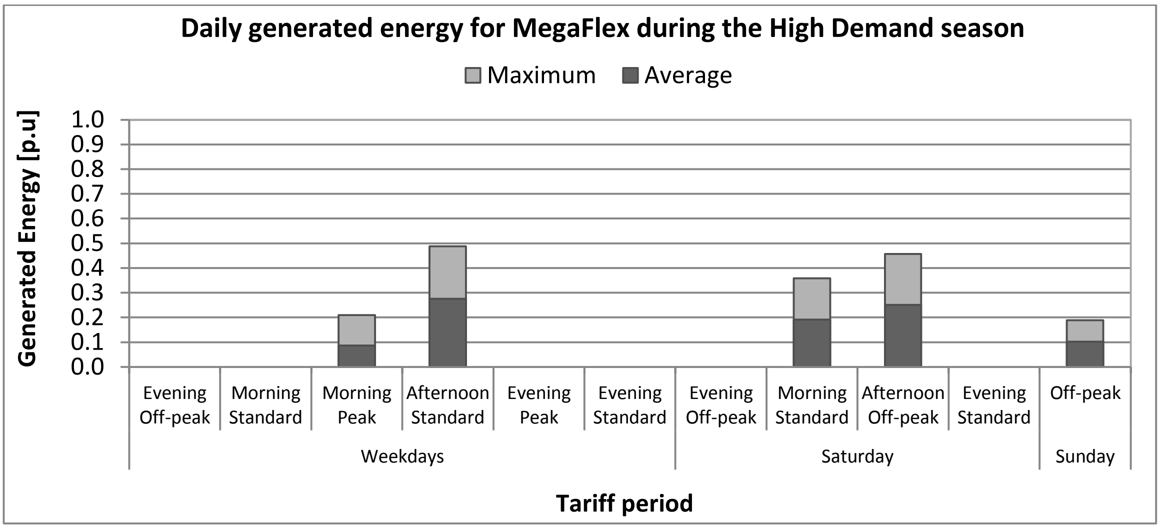

Figure 5.6: Daily average and maximum generated energy for MegaFlex during the High Demand season.

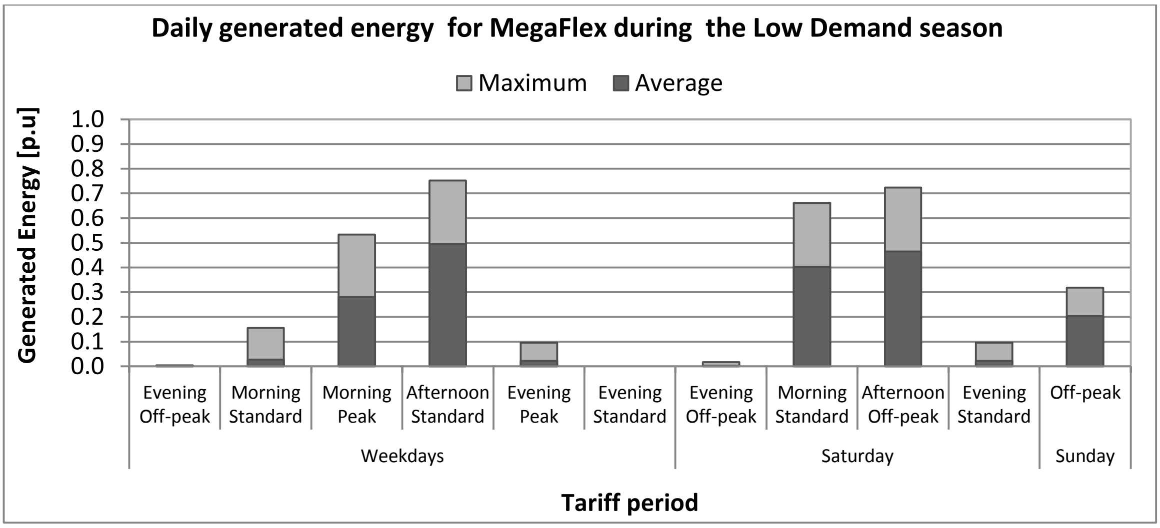

Figure 5.7: Daily average and maximum generated energy for MegaFlex during Low Demand season.

Note that Sundays consist of only off-peak tariff periods, therefore the generated energy is normalised to the entire 48 half-hour intervals of the day. This gives rise to a significantly smaller per unit value for this tariff period as it is normalised to several half-hour intervals where no energy is being generated. Figures 5.6 and 5.7 indicate that the average daily generated energy for Sundays is about 10% of the rated tariff period energy for the High Demand season, while being about 20% for the Low Demand season.

Figure 5.8 presents the COV of the generated energy for each MegaFlex tariff period during the High and Low Demand seasons. The COV is determined by dividing each tariff period's standard deviation by the average daily generated energy. The results show that the variation in generated energy is at its largest during early morning and late afternoon hours and at its lowest during mid-day hours. The generated energy varies significantly for morning and evening tariff periods during the Low Demand season with COV well over a 100% and evens out towards mid-day tariff periods with COV between 30% and 40%.

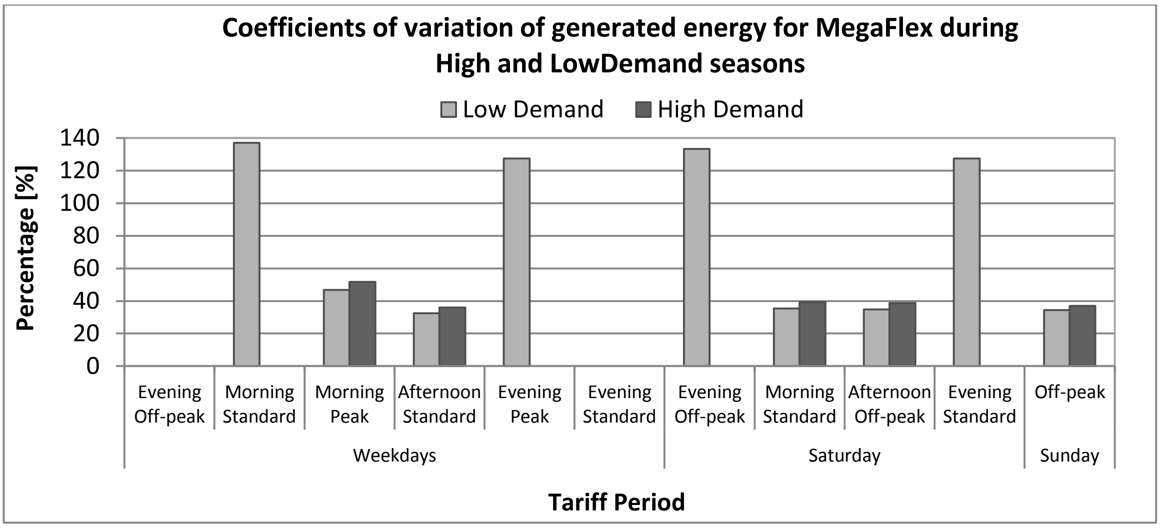

Figure 5.8: Coefficients of variation of generated energy for MegaFlex during High and Low Demand seasons.

5.4.1.2.2 Statistical Model

Tables B.5 and B.6 in appendix B summarise the results of the Chi-squared test for the MegaFlex tariff structure during the High and Low Demand seasons. The Chi-squared test results for each distinct tariff period are given with respect to all hypothesised probability distributions. Each probability distribution's chi-squared results are given as a value and a number of bins, with the number of bins denoted by NB.

Tables 5.9 and 5.10 summarise the best performing conclusive models, i.e. the probability distributions with the lowest Chi-squared test value and resultant DOF of at least one, for each tariff period during the High and Low Demand seasons together with the model conclusions.

From the results presented in table 5.9 it can be seen that only one of the weekday tariff periods is successfully modelled for MegaFlex during the High Demand season. Likewise, only one of the Saturday tariff periods is successfully modelled for MegaFlex during the High Demand season. These tariff periods represent times where little energy is generated during the High Demand season. The Chi-squared test and RMSE results given in tables B.5 and B.7 indicate that the best performing probability distributions, i.e. probability distributions with the lowest Chi-squared test results, are rejected for the remainder of tariff periods. Therefore, all considered probability distributions are conclusively rejected, i.e. do not fit historical generation data, for these tariff periods.

Table 5.9: Best performing conclusive models for MegaFlex during High Demand season.

Tariff day	**Tariff period**	**Chi-squared value**	**DOF**	**Probability distribution**	**Average energy [p.u.]**	**Standard deviation of energy [p.u.]**	**Model conclusion**
Weekday	Evening Off-peak	0	< 1	Inconclusive	0	0	No Energy
	Morning Standard	0	< 1	Inconclusive	0	0	No Energy
	Morning Peak	16.105	3	Normal	0.087	0.045	Reject
	Afternoon Standard	28.477	3	Gamma	0.275	0.099	Reject
	Evening Peak	0.146	1	Beta	2.70E-05	8.922E-05	Accept
	Evening Standard	0	< 1	Inconclusive	0	0	No Energy
Saturday	Evening Off-peak	0	< 1	Inconclusive	0	0	No Energy
	Morning Standard	37.644	4	Normal	0.191	0.075	Reject
	Afternoon Off-peak	28.812	4	Beta	0.251	0.097	Reject
	Evening Standard	0.146	1	Beta	2.70E-05	8.922E-05	Accept
Sunday	Off-peak	32.935	4	Normal	0.103	0.038	Reject

The results in table 5.10 show that about half of the tariff periods are successfully modelled for MegaFlex during the Low Demand season. The weekday morning and evening peak tariff periods are successfully modelled for the High demand season. However, the weekday afternoon standard tariff period is rejected while it represents the time period with the greatest amount of generated energy. The results also indicate that all Saturday tariff periods are successfully modelled with the exception of the evening off-peak. However, little energy is generated in this tariff period. Furthermore, the Sunday off-peak tariff period cannot be modelled for the Low demand season which is undesirable as this tariff period represents the entire day.

Table 5.10: Best performing conclusive models for MegaFlex during Low Demand season.

Tariff day	Tariff period	Chi-squared value	DOF	Probability distribution	Average energy [p.u.]	Standard deviation of energy [p.u.]	Model conclusion
Weekday	Evening Off-peak	26.546	1	Beta	1.99E-04	4.440E-04	Reject
	Morning Standard	20.802	6	Beta	0.027	0.037	Reject
	Morning Peak	4.573	6	Beta	0.280	0.131	Accept
	Afternoon Standard	31.460	6	Beta	0.495	0.161	Reject
	Evening Peak	14.757	5	Beta	0.022	0.028	Accept
	Evening Standard	0	< 1	Inconclusive	0	0	No Energy
Saturday	Evening Off-peak	20.831	7	Beta	0.003	0.004	Reject
	Morning Standard	10.449	7	Beta	0.402	0.142	Accept
	Afternoon Off-peak	17.753	7	Beta	0.465	0.162	Accept
	Evening Standard	14.757	5	Beta	0.022	0.028	Accept
Sunday	Off-peak	22.693	7	Beta	0.203	0.070	Reject

5.4.1.3 HomeFlex Tariff Structure

5.4.1.3.1 Statistical Parameters

Tables 5.11 and 5.12 summarise the statistical parameters of the daily generated energy for the HomeFlex tariff during High and Low Demand. The maximum, average and standard deviation of the daily generated energy is normalised to the rated energy output of each tariff period.

Table 5.11: Statistical parameters of generated energy for HomeFlex during High Demand season.

Tariff period	Maximum energy [p.u.]	Average energy [p.u.]	Standard Deviation of energy[p.u.]
Evening Off-peak	0	0	0
Morning Peak	0.210	0.087	0.045
Afternoon Off-peak	0.489	0.275	0.099
Evening Peak	4.960E-04	2.703E-05	8.922E-05

Table 5.12: Statistical parameters of generated energy for HomeFlex during Low Demand season.

Tariff period	Maximum energy [p.u.]	Average energy [p.u.]	Standard Deviation of energy[p.u.]
Evening Off-peak	0.016	0.003	0.004
Morning Peak	0.534	0.280	0.131
Afternoon Off-peak	0.752	0.495	0.161
Evening Peak	0.095	0.022	0.028

Figures 5.9 and 5.10 present the maximum and average daily generated energy from tables 5.11 and 5.12 in graphic format.

The results indicate that the greatest amount of energy is generated in the morning peak and afternoon off-peak tariff periods. The average daily generated energy in these tariff periods vary between about 9% and 28% of the rated tariff period energy for the High demand season ,while varying between about 28% and 50% for the Low Demand season

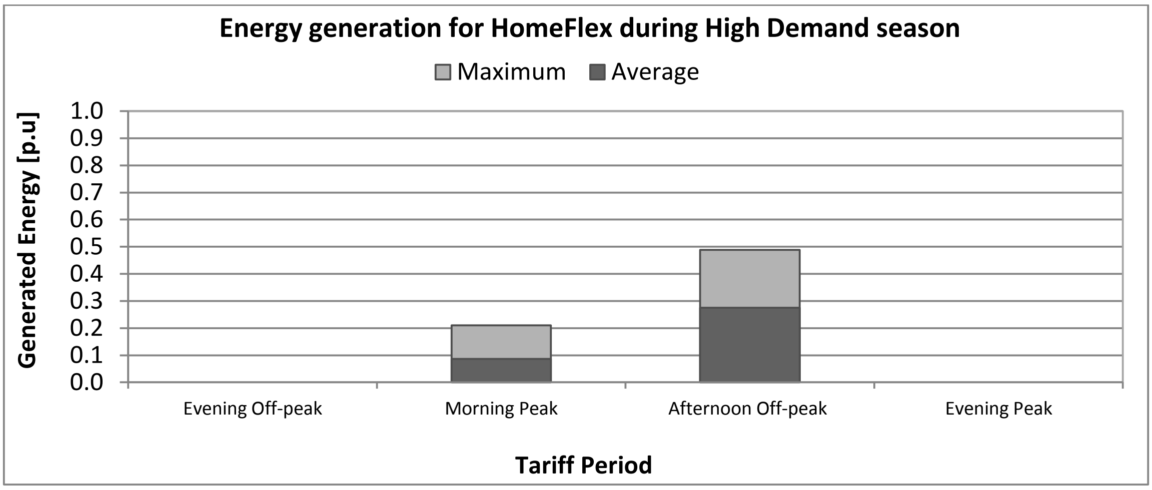

Figure 5.9: Daily average and maximum generated energy for HomeFlex during High Demand season.

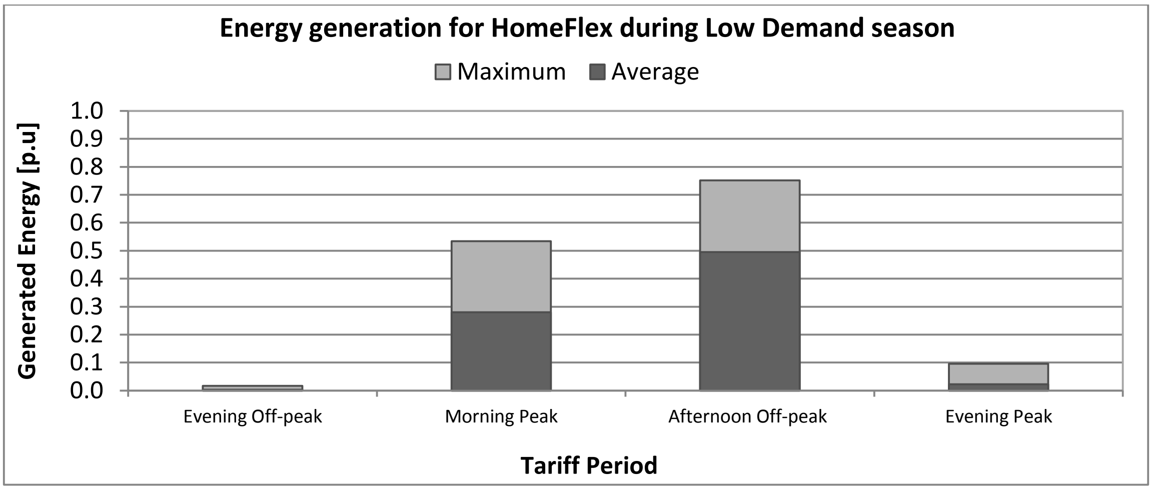

Figure 5.10: Daily average and maximum generated energy for HomeFlex during Low Demand season.

Stellenbosch University https://scholar.sun.ac.za

Figure 5.11 presents the COV of the generated energy for each HomeFlex tariff period during the High and Low Demand seasons. The COV is a normalised measure of variation and is determined by dividing each tariff period's standard deviation by the average daily generated energy.

The results show that the variation in generated energy is at its largest during early morning and late afternoon hours and at its lowest during mid-day hours. The generated energy varies significantly for morning and evening tariff periods during the Low Demand season with COV well over a 100% and evens out towards mid-day tariff periods with COV between 30% and 40%.

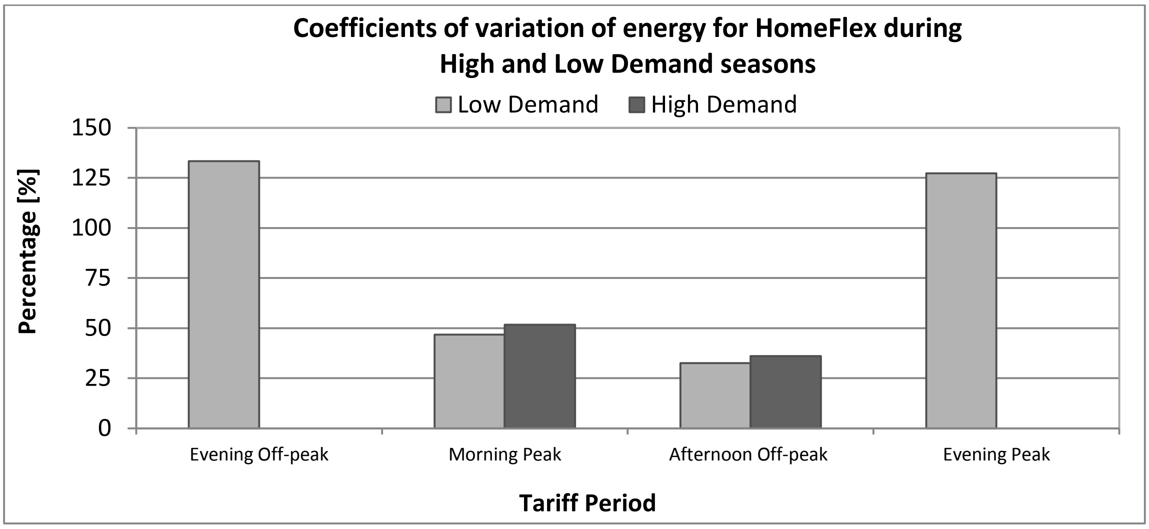

Figure 5.11: Coefficients of variation of generated energy for HomeFlex during High and Low Demand seasons.

5.4.1.3.2 Statistical Model

Tables B.9 and B.10 in appendix B summarise the results of the Chi-squared test for the HomeFlex tariff during the High and Low Demand seasons. The Chi-squared test results for each distinct tariff period are given with respect to all hypothesised probability distributions. Each probability distribution's chi-squared results are given as a value and a number of bins, with the number of bins denoted by NB.

Tables 5.13 and 5.14 summarise the best performing conclusive models, i.e. the probability distributions with the lowest Chi-squared test value and resultant DOF of at least one, for each tariff period during the High and Low Demand seasons. Chi-squared values for each probability distribution are provided together with the degrees of freedom denoted by DOF.

Stellenbosch University https://scholar.sun.ac.za

From results presented in table 5.13 it can be seen that none of the tariff periods can be modelled for HomeFlex during the High Demand season. The Chi-squared test and RMSE results given in tables B.9 and B.11 indicate that the best performing probability distributions are rejected for both tariff periods. Therefore, all considered probability distributions are conclusively rejected, i.e. do not fit historical generation data, for these two tariff periods.

Table 5.13: Best performing conclusive models for HomeFlex during High Demand season

Tariff period	Chi-squared value	DOF	Probability distribution	Average [p.u.]	Standard Deviation [p.u.]	Model conclusion
Evening Off-peak	0	< 1	Inconclusive	0	0	No Energy
Morning Peak	16.105	3	Normal	0.087	0.045	Reject
Afternoon Off-peak	28.477	4	Beta	0.275	0.099	Reject
Evening Peak	0	< 1	Inconclusive	0	0	No Energy

Table 5.14 shows that only the morning and evening peak tariff periods are successfully modelled for HomeFlex during the Low Demand season. Furthermore, the afternoon off-peak tariff period is rejected while it represents a time period with the greatest amount of daily generated energy. The Chi-squared test and RMSE results given in tables B.10 and B.12 indicate that the Beta probability distribution performs the best of all the considered distributions during the Low Demand season. However, the Beta distribution is rejected and therefore all considered probability distributions are rejected as models for the evening off-peak and afternoon off-peak tariff periods.

Table 5.14: Best performing conclusive models for HomeFlex during Low Demand season.

Tariff period	Chi-squared value	DOF	Probability distribution	Average [p.u.]	Standard Deviation [p.u.]	Model conclusion
Evening Off-peak	20.831	7	Beta	0.003	0.004	Reject
Morning Peak	4.573	6	Beta	0.280	0.131	Accept
Afternoon Off-peak	31.460	6	Beta	0.495	0.161	Reject
Evening Peak	14.757	5	Beta	0.022	0.028	Accept

5.4.2 Monthly Analysis

This analysis is conducted with respect to the monthly intervals of a year from the start of February 2013 to end of June 2014. The results of the summer month February and the winter month June are presented. The statistical parameters, Chi-squared test values and RMSEs for all calendar months of the year are provided in Appendix C.

Stellenbosch University https://scholar.sun.ac.za

5.4.2.1 Daily Half-hourly Generation Profile

During the analysis timeline a total of 56 days fall in the February month interval and a total of 60 days fall in the June month interval. All night time half-hourly intervals are excluded from the presented results.

5.4.2.1.1 Statistical Parameters

Tables 5.15 and 5.16 summarise the statistical parameters of the daily half-hourly generated energy during the calendar months February and June. The maximum, average and standard deviation of the daily generated energy is normalised to the rated energy output of 254.04 kWh per half-hour interval.

Table 5.15: Statistical parameters of daily generated energy for half-hourly profile during February.

Period start	Maximum energy [p.u.]	Average energy [p.u.]	Standard Deviation of energy[p.u.]
06:00:00	0.017	0.005	0.005
06:30:00	0.076	0.032	0.016
07:00:00	0.159	0.093	0.032
07:30:00	0.266	0.185	0.051
08:00:00	0.388	0.284	0.076
08:30:00	0.484	0.380	0.087
09:00:00	0.571	0.464	0.104
09:30:00	0.648	0.533	0.118
10:00:00	0.736	0.608	0.114
10:30:00	0.771	0.656	0.128
11:00:00	0.816	0.707	0.105
11:30:00	0.909	0.735	0.098
12:00:00	0.862	0.744	0.123
12:30:00	0.882	0.755	0.135
13:00:00	0.906	0.754	0.142
13:30:00	0.882	0.756	0.122
14:00:00	0.844	0.732	0.133
14:30:00	0.813	0.690	0.132
15:00:00	0.760	0.660	0.113
15:30:00	0.750	0.605	0.114
16:00:00	0.627	0.536	0.089
16:30:00	0.589	0.454	0.097
17:00:00	0.477	0.359	0.085
17:30:00	0.370	0.253	0.074
18:00:00	0.230	0.153	0.051
18:30:00	0.124	0.049	0.030
19:00:00	0.041	0.007	0.006

Stellenbosch University https://scholar.sun.ac.za

Table 5.16: Statistical parameters of daily generated energy for half-hourly profile during June.

Period start	Maximum [p.u.]	Average [p.u.]	Standard Deviation [p.u.]
07:30:00	0.022	0.006	0.005
08:00:00	0.152	0.035	0.022
08:30:00	0.194	0.079	0.040
09:00:00	0.246	0.132	0.063
09:30:00	0.336	0.194	0.092
10:00:00	0.419	0.251	0.115
10:30:00	0.468	0.292	0.138
11:00:00	0.574	0.335	0.141
11:30:00	0.546	0.375	0.144
12:00:00	0.569	0.383	0.147
12:30:00	0.576	0.386	0.151
13:00:00	0.565	0.368	0.156
13:30:00	0.495	0.351	0.148
14:00:00	0.514	0.326	0.137
14:30:00	0.422	0.285	0.121
15:00:00	0.359	0.243	0.102
15:30:00	0.270	0.169	0.082
16:00:00	0.190	0.110	0.058
16:30:00	0.104	0.049	0.030
17:00:00	0.018	0.007	0.004

Figures 5.12 and 5.13 present the maximum and average daily generated energy from tables 5.15 and 5.16 in graphic format. The results indicate a greater availability of solar power during the calendar month of February than that of June. This is to be expected, as February is a summer month and June is a winter month.

Figures 5.12 and 5.13 clearly show the effect of different sunrise and sunset times for winter and summer months. The results indicate a one and a half hour shift in the solar for mornings and a two hour shift for evenings. This lack of energy availability for morning and evening hours during the calendar month June is a drawback of solar power, as these hours coincide with times of peak energy demand.

The average generated energy shown in figures 5.12 and 5.13 indicate a considerable difference in the solar plant’s performance during the calendar months February and June. This significant difference between February and June is attributed to the path taken by the earth around the sun, weather conditions and the fact that the solar panels have fixed orientations and tilts. As mentioned before, fixed orientations and tilts of solar panels are inefficient and lead to decrease generation of energy during certain times of the year.

The solar plant performs significantly better during February than during June. The results indicated that the average generated energy during February is greater than the maximum generated energy during June. Furthermore, the peak average daily generated energy reaches a per unit value of about 76% during February, while reaching a per unit value of only about 39% during June.

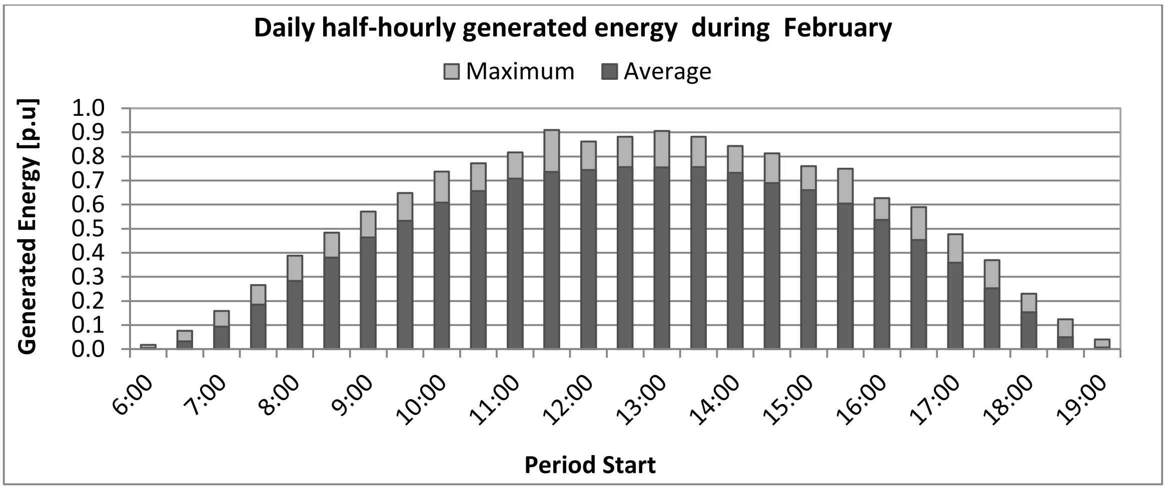

Figure 5.12: Daily average and maximum generated energy for half-hourly profile during February.

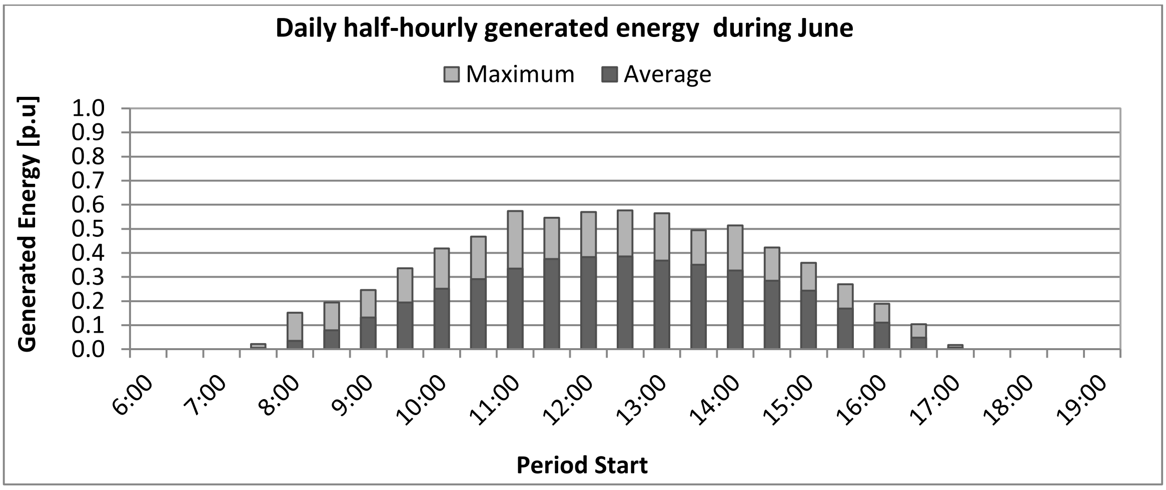

Figure 5.13: Daily average and maximum generated energy for half-hourly profile during June.

Figure 5.14 presents the Coefficients Of Variation (COV) of the generated energy for each half-hour interval during the calendar months February and June. The COV is a normalised measure of variation and is determined by dividing each half-hour interval's standard deviation by the average daily generated energy.

Figure 5.14 indicates that the generated energy varies more during June than during February. The results also show that the variation in generated energy is at its largest during early morning and late afternoon hours and at its lowest during mid-day hours. The COV for mid-day half-hour intervals are about 19% during February and about 40% during June.

From figures 5.5 and 5.14 it can be seen that the COV for mid-day half-hour intervals during the calendar month February are significantly lower than that of the Low Demand season. It can also be seen that the COV for mid-day half-hour intervals are roughly the same for the calendar month June and the High Demand season. The difference in the COV during February and the Low Demand season is attributed to the difference in the time intervals associated with each. The Low Demand season consists of nine different calendar months while February only represents one calendar month. Similarly, the results of the June month and the High demand season are roughly the same as the Low Demand season consists of only three winter months.

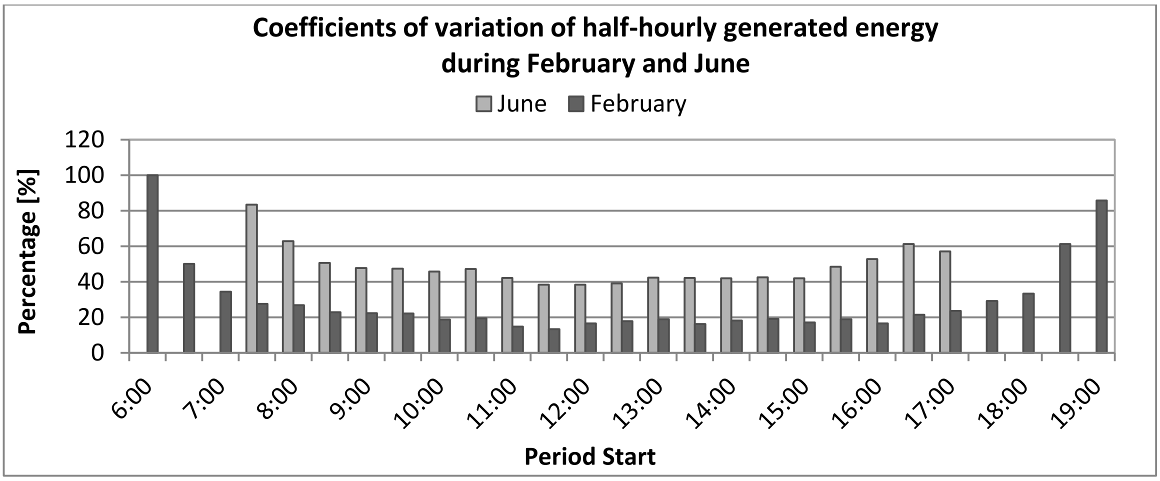

Figure 5.14: Coefficients of variation of half-hourly generated energy during February and June.

5.4.2.1.2 Statistical Model

Tables C.14 and C.18 in appendix C summarise the results of the Chi-squared test for the half-hourly profile during the calendar months of February and June. The Chi-squared test results for each distinct half-hour interval are given with respect to all hypothesised probability distributions. Each probability distribution's chi-squared results are given as a value and a number of bins, with the number of bins denoted by NB.

Tables 5.17 and 5.18 summarise the best performing conclusive models, i.e. the probability distributions with the lowest Chi-squared test value and resultant DOF of at least one, for each half-hour interval during the calendar months of February and June. The model conclusions, i.e. whether the hypothesised probability distribution fits the observed historical generation data (accepted) or not (rejected), are also provided. Chi-squared values for the best performing conclusive models are provided together with the determined degrees of freedom denoted by DOF.

The results presented in table 5.17 show that a large portion of the half-hourly models for the calendar month February are rejected. The Chi-squared test and RMSE results given in tables C.14 and C.26 indicate that the Beta and Logistic probability distributions performs the best in the half-hourly intervals from 08:30:00 to 17:00:00. However, these probability distributions results in DOF of less than one and are therefore inconclusive. Similarly, all the other probability distributions except the Exponential distribution have DOF less than one. Therefore, the Exponential distribution is chosen as the best performing conclusive model. Conducting the analysis on more historical generation data as it becomes available will result in higher numbers of bins and therefore DOF. With sufficient DOF the Beta and Logistic distributions can be conclusively accepted or rejected for these half-hour intervals.

Table 5.17 Best performing conclusive models for daily half-hourly profile during February.

Period start	Chi-squared value	DOF	Probability distribution	Average [p.u.]	Standard Deviation [p.u.]	Model conclusion
06:00:00	1.909	2	Beta	0.005	0.005	Accept
06:30:00	0.750	2	Weibull	0.032	0.016	Accept
07:00:00	2.361	2	Normal	0.093	0.032	Accept
07:30:00	2.812	1	Logistic	0.185	0.051	Accept
08:00:00	4.001	1	Beta	0.284	0.076	Accept
08:30:00	265.291	4	Exponential	0.380	0.087	Reject
09:00:00	312.278	4	Exponential	0.464	0.104	Reject
09:30:00	406.985	4	Exponential	0.533	0.118	Reject
10:00:00	339.354	5	Exponential	0.608	0.114	Reject
10:30:00	495.013	5	Exponential	0.656	0.128	Reject
11:00:00	488.954	6	Exponential	0.707	0.105	Reject
11:30:00	20.122	1	Weibull	0.735	0.098	Reject
12:00:00	479.603	4	Exponential	0.744	0.123	Reject
12:30:00	472.446	4	Exponential	0.755	0.135	Reject
13:00:00	346.619	4	Exponential	0.754	0.142	Reject
13:30:00	407.383	4	Exponential	0.756	0.122	Reject
14:00:00	535.681	4	Exponential	0.732	0.133	Reject
14:30:00	458.463	4	Exponential	0.690	0.132	Reject
15:00:00	587.226	4	Exponential	0.660	0.113	Reject
15:30:00	304.592	4	Exponential	0.605	0.114	Reject
16:00:00	401.531	4	Exponential	0.536	0.089	Reject
16:30:00	265.403	4	Exponential	0.454	0.097	Reject
17:00:00	6.859	1	Beta	0.359	0.085	Reject
17:30:00	7.309	2	Beta	0.253	0.074	Accept
18:00:00	2.082	2	Beta	0.153	0.051	Accept
18:30:00	0.919	2	Weibull	0.049	0.030	Accept
19:00:00	0	< 1	Inconclusive	0.007	0.006	Inconclusive

Stellenbosch University https://scholar.sun.ac.za

Table 5.18 indicates that that a large portion of the half-hourly models for the calendar month June are successfully modelled and accepted. The results show that the 11:30:00 half-hour interval is rejected for the calendar month June. However, the resultant Chi-squared value of this half-hour interval is very near to the limit of 9.21 as can be seen in the percentage points table provided in appendix A. Conducting the analysis on more historical generation data as it becomes available may result in improved Chi-squared test results and the acceptance of the Beta probability distribution for the 11:30:00 half-hour interval.

Table 5.18: Best performing conclusive models for daily half-hourly profile during June.

Period start	Chi-squared value	DOF	Probability distribution	Average [p.u.]	Standard Deviation [p.u.]	Model conclusion
07:30:00	1.101	1	Logistic	0.006	0.005	Accept
08:00:00	4.801	1	Exponential	0.035	0.022	Accept
08:30:00	0.738	1	Beta	0.079	0.040	Accept
09:00:00	6.086	2	Beta	0.132	0.063	Accept
09:30:00	7.814	2	Beta	0.194	0.092	Accept
10:00:00	5.347	1	Beta	0.251	0.115	Accept
10:30:00	3.845	1	Beta	0.292	0.138	Accept
11:00:00	8.212	2	Beta	0.335	0.141	Accept
11:30:00	9.459	2	Beta	0.375	0.144	Reject
12:00:00	3.780	2	Beta	0.383	0.147	Accept
12:30:00	4.78	2	Beta	0.386	0.151	Accept
13:00:00	4.430	2	Beta	0.368	0.156	Accept
13:30:00	4.0756	1	Beta	0.351	0.148	Accept
14:00:00	5.946	1	Beta	0.326	0.137	Accept
14:30:00	6.454	1	Beta	0.285	0.121	Accept
15:00:00	4.399	1	Beta	0.243	0.102	Accept
15:30:00	1.707	1	Beta	0.169	0.082	Accept
16:00:00	0.107	1	Beta	0.110	0.058	Accept
16:30:00	0.959	1	Beta	0.049	0.030	Accept
17:00:00	0.171	2	Logistic	0.007	0.004	Accept

5.4.2.2 HomeFlex Tariff Structure

5.4.2.2.1 Statistical Parameters

Tables 5.19 and 5.20 summarise the statistical parameters of the daily generated energy for the HomeFlex tariff during the calendar months of February and June. The maximum, average and standard deviation of the daily generated energy is normalised to the rated energy output of each tariff period.

Stellenbosch University https://scholar.sun.ac.za

Table 5.19: Statistical parameters of daily generated energy for HomeFlex during February.

Tariff period	Maximum energy [p.u.]	Average energy [p.u.]	Standard Deviation of energy [p.u.]
Evening off-peak	0.004	0.002	0.001
Morning peak	0.414	0.323	0.071
Afternoon off-peak	0.710	0.625	0.091
Evening peak	0.090	0.053	0.020

Table 5.20: Statistical parameters of daily generated energy for HomeFlex during June.

Tariff period	Maximum energy [p.u.]	Average energy [p.u.]	Standard Deviation of energy [p.u.]
Evening off-peak	0	0	0
Morning peak	0.145	0.074	0.034
Afternoon off-peak	0.347	0.246	0.092
Evening peak	0	0	0

Figures 5.15 and 5.16 present the maximum and average daily generated energy from tables 5.19 and 5.20 in graphic format.

Figures 5.15 and 5.10 show that average daily generated energy for the morning peak and afternoon off-peak tariff periods is more during the calendar month of February than for the Low Demand season. Furthermore, figures 5.16 and 5.9 indicate that average daily generated energy for the morning peak and afternoon off peak tariff periods is roughly the same for the calendar month of June and the High Demand season.

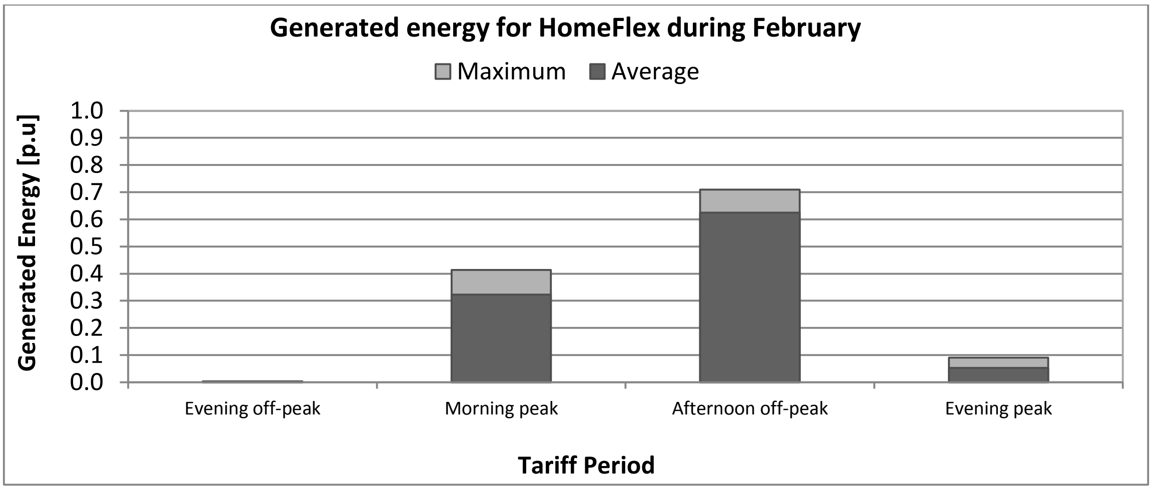

Figure 5.15: Daily average and maximum generated energy for HomeFlex during February.

Stellenbosch University https://scholar.sun.ac.za

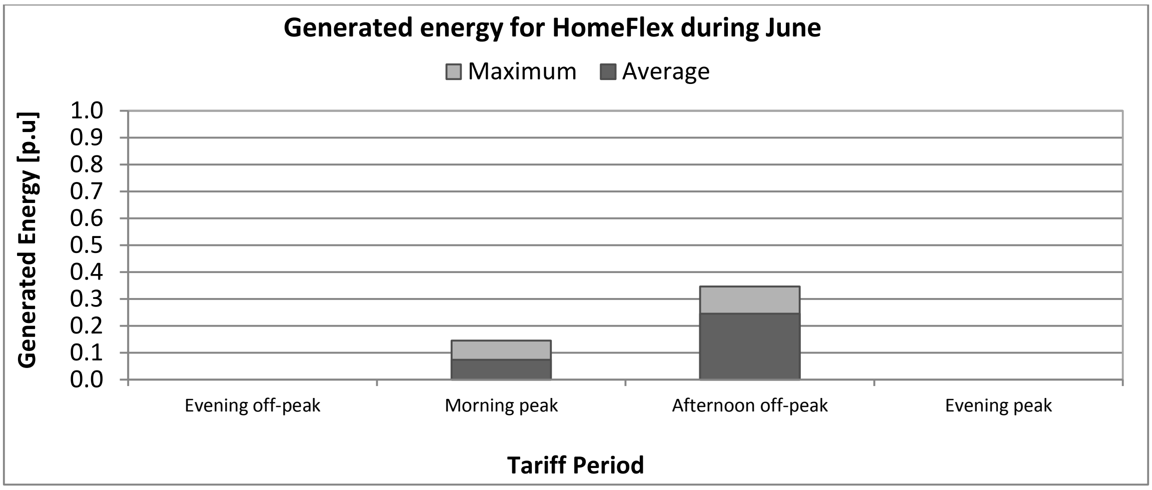

Figure 5.16: Daily average and maximum generated energy for HomeFlex during June.

Figure 5.17 presents the COV of the half-hourly generated energy for the calendar months February and June. From figures 5.17 and 5.11 it can be seen that the COV is significantly lower during June than during the High Demand season. It can also be seen that the COV is significantly lower during February than during the Low Demand season. The difference in the COV for the monthly and seasonal analysis is attributed to the difference in the time intervals associated with each. The seasons consist of several different calendar months with varying weather patterns, while the monthly intervals consist of a single month.

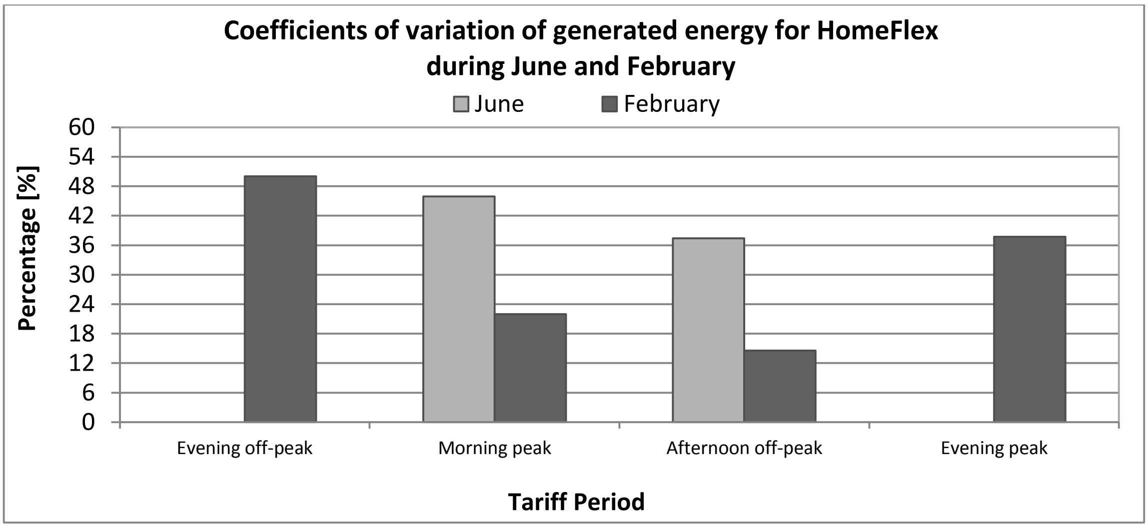

Figure 5.17: Coefficients of variation of generated energy for HomeFlex February and June.

Stellenbosch University https://scholar.sun.ac.za

5.4.2.2.2 Statistical Model

Table C.38 in appendix C summarise the results of the Chi-squared test for the HomeFlex tariff during the calendar months of February and June. Tables 5.21 and 5.22 summarise the best performing conclusive models, i.e. the probability distributions with the lowest Chi-squared test value and resultant DOF of at least one, for each tariff period during the calendar months of February and June.

From results presented in table 5.21 it can be seen that only the evening off-peak and evening peak tariff periods are successfully modelled for the HomeFlex tariff during February. The Chi-squared test results given in table C.38 indicate that the Logistic probability distribution performs the best in the morning peak and afternoon off-peak tariff periods. However, the Logistic probability distribution results in DOF of less than one for these tariff periods and is therefore inconclusive. Similarly, all the other probability distributions except the Exponential distribution have DOF less than one. Therefore, the Exponential distribution is chosen as the best performing conclusive model for the morning peak and afternoon off-peak tariff periods.

Table 5.21: Best performing conclusive models for HomeFlex during February.

Period Start	Chi-squared value	DOF	Probability distribution	Average energy [p.u.]	Standard Deviation of energy [p.u.]	Model conclusion
Evening off-peak	1.529	2	Weibull	0.002	0.001	Accept
Morning peak	270.358	4	Exponential	0.323	0.071	Reject
Afternoon off-peak	543.093	4	Exponential	0.625	0.091	Reject
Evening peak	0.992	2	Beta	0.053	0.020	Accept

Table 5.22 indicates that only the afternoon off-peak tariff period is successfully modelled for the HomeFlex tariff during June. The Chi-squared test results given in table C.38 indicate that the Beta distribution performs the best for the morning peak tariff period. However, the Beta distribution is rejected and therefore all probability distributions are rejected for the morning peak tariff period.

Table 5.22: Best performing conclusive models for HomeFlex during June.

Period Start	Chi-squared value	DOF	Probability distribution	Average energy [p.u.]	Standard Deviation of energy [p.u.]	Model conclusion
Evening off-peak	0	< 1	Inconclusive	0	0	No energy
Morning peak	14.009	2	Beta	0.074	0.034	Reject
Afternoon off-peak	2.017	1	Beta	0.246	0.092	Accept
Evening peak	0	< 1	Inconclusive	0	0	No Energy

Stellenbosch University https://scholar.sun.ac.za

5.4.3 Daily Energy Generation Forecast

5.4.3.1 Overview

This section deals with the forecasted daily generated energy for the half-hourly generation profile and the HomeFlex tariff structure during the calendar month June. The forecasted energy generation values are compared to the historical generation data for the analysis timeline to evaluate the forecasting models' performance.

The daily generated energy is forecasted by using the accepted statistical models, i.e. probability distributions that fit the historical generation data together with statistical parameters, derived from the historical generation data. The statistical models are used to predict the energy generated, in each half-hour interval and tariff period, with a specified exceedance probability.

The Exceedance Probability (EP) is the likelihood that a designated value will be exceeded [69]. The 90% EP value for a solar plant's energy output is the specific energy output value that will be exceeded 90% of the time, i.e. there is a 90% likelihood that the solar plant's output will be greater than the 90% EP value.

The EP values are determined from the Cumulative Distribution Functions (CDFs) of the accepted statistical models. The CDF of a probability distribution is used with derived statistical parameters to determine the value that results in a specified EP. For example, the 90% EP value occurs when the probability distribution's CDF is equal to 10% (0.1) [69].

5.4.3.2 Daily Half-hourly Generation Profile

Table 5.23 summarises the daily generated energy forecast models and EP values for the half-hourly profile during the calendar month of June. The 90 %, 80 % and 70 % EP values of the solar plant's energy output are provided and are normalised to the rated energy output of 254.04 kWh per half-hour interval. The historical generation data for June consists of 60 measurements for the analysis time-line.

Stellenbosch University https://scholar.sun.ac.za

Table 5.23: Daily generated energy forecast for half-hourly profile during June.

Period Start	Forecasting models			Forecasted total daily energy [p.u.]		
	Probability distribution	**Average energy [p.u.]**	**Standard Deviation of energy[p.u.]**	**90% EP value**	**80% EP value**	**70% EP value**
07:30:00	Logistic	0.006	0.005	0.001	0.003	0.004
08:00:00	Exponential	0.035	0.022	0.004	0.008	0.013
08:30:00	Beta	0.079	0.040	0.027	0.041	0.054
09:00:00	Beta	0.132	0.063	0.044	0.070	0.093
09:30:00	Beta	0.194	0.092	0.060	0.102	0.138
10:00:00	Beta	0.251	0.115	0.080	0.135	0.183
10:30:00	Beta	0.292	0.138	0.080	0.149	0.211
11:00:00	Beta	0.335	0.141	0.133	0.199	0.254
11:30:00	Beta	0.375	0.144	Rejected	Rejected	Rejected
12:00:00	Beta	0.383	0.147	0.157	0.241	0.309
12:30:00	Beta	0.386	0.151	0.155	0.241	0.309
13:00:00	Beta	0.368	0.156	0.128	0.213	0.284
13:30:00	Beta	0.351	0.148	0.104	0.203	0.287
14:00:00	Beta	0.326	0.137	0.120	0.191	0.251
14:30:00	Beta	0.285	0.121	0.092	0.163	0.223
15:00:00	Beta	0.243	0.102	0.082	0.142	0.192
15:30:00	Beta	0.169	0.082	0.041	0.081	0.120
16:00:00	Beta	0.110	0.058	0.024	0.049	0.073
16:30:00	Beta	0.049	0.030	0.009	0.018	0.028
17:00:00	Logistic	0.007	0.004	0.002	0.004	0.005

To evaluate the performance of the statistical models, the predicted EP values are compared to the historical generation data for the analysis timeline. The percentage of the total historical half-hour intervals that have a measured amount of generated energy above the forecasted EP value is determined for each interval, i.e. the proportion of the historical measurements above the predicted EP values. An accurate model for a given half-hour interval and EP value will result in a percentage of historical half-hour intervals above the forecasted EP value of at least the EP, i.e. a 90 % EP value must be exceeded by 90% of the historical half-hour intervals to be accurate.

Table 5.24 summarises the results for the percentage of the total historical half-hour intervals that have a measured amount of generated energy above the forecasted EP value while figure 5.18 presents the results in graphic format.

The results indicate that a large portion of historical half-hourly generated energy is slightly below the forecasted EP values. However, the historical generation data is limited to only two years. As more generation data becomes available the results will improve and provide a more precise indication of forecasting accuracy.

Stellenbosch University https://scholar.sun.ac.za

Table 5.24: Model performance for half-hourly forecasted energy generation for the calendar month of June.

Period start	Probability distribution	Percentage of historical half-hour intervals with measured generated energy above EP value [%]		
		90% EP value	**80% EP value**	**70% EP value**
07:30:00	Logistic	91.67	80	55
08:00:00	Exponential	93.33	90	85
08:30:00	Beta	85	76.67	68.33
09:00:00	Beta	86.67	80	66.67
09:30:00	Beta	86.67	80	70
10:00:00	Beta	85	76.67	70
10:30:00	Beta	88.33	78.33	71.67
11:00:00	Beta	86.67	80	73.33
11:30:00	Beta	Rejected	Rejected	Rejected
12:00:00	Beta	88.33	78.33	68.33
12:30:00	Beta	88.33	78.33	70
13:00:00	Beta	85	80	70
13:30:00	Beta	88.33	76.67	70
14:00:00	Beta	88.33	78.33	68.33
14:30:00	Beta	86.67	76.67	70
15:00:00	Beta	91.67	76.67	68.33
15:30:00	Beta	91.67	80	65
16:00:00	Beta	91.67	80	65
16:30:00	Beta	90	85	68.33
17:00:00	Logistic	85	80	68.33

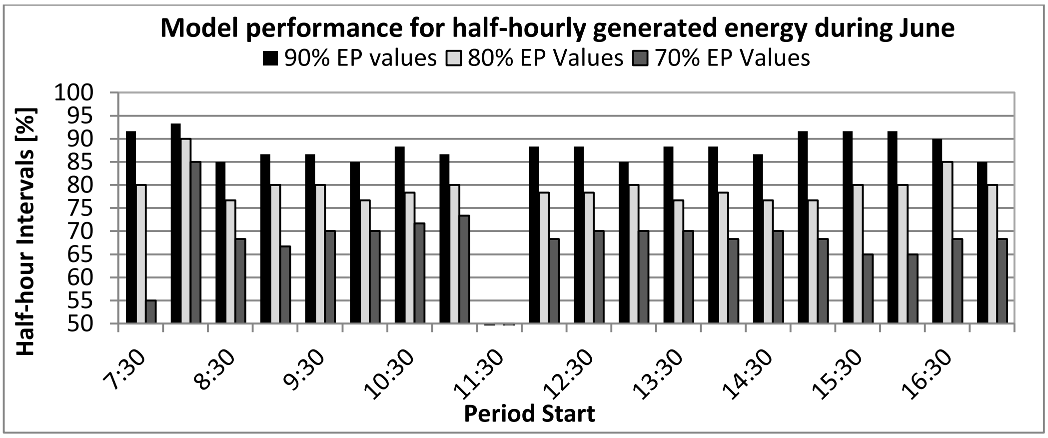

Figure 5.18: Model performance for half-hourly forecasted energy generation for the calendar month of June.

5.4.3.3 Home Flex Tariff Structure

Table 5.25 summarises the daily generated energy forecast models and EP values for the HomeFlex tariff structure during the calendar month of June. The 90 %, 80 % and 70 % EP values of the solar plant's energy output are provided and are normalised to the rated energy output of each tariff period.

From table 5.25 it can be seen that only the afternoon off-peak tariff period can be predicted. This tariff period presents the time period with the greatest amount of daily generated energy during the calendar month June.

Table 5.25: Daily generated energy forecast during for HomeFlex during June.

Tariff Period	Forecasting models			Forecasted daily energy [p.u.]		
	Probability distribution	Average energy [p.u.]	Standard Deviation of energy [p.u.]	90% EP values	80% EP values	70% EP values
Evening off-peak	Inconclusive	0	0	0	0	0
Morning peak	Beta	0.074	0.034	Reject	Reject	Reject
Afternoon off-peak	Beta	0.246	0.092	0.100	0.159	0.205
Evening peak	Inconclusive	0	0	0		0

To evaluate the performance of the statistical models, the predicted EP values are compared to the historical generation data for the analysis timeline. The percentage of the total historical tariff periods that have a measured amount of generated energy above the forecasted EP value is determined for each period.

Table 5.26 summarises the results for the percentage of the total historical tariff periods that have a measured amount of generated energy above the forecasted EP value. The results indicate that a large portion of historical half-hourly generated energy is slightly below the forecasted EP values.

Table 5.26: Model validation for forecasted energy generation for HomeFlex during the calendar month of June

Tariff Period	Probability distribution	Percentage of historical measurements above forecasted energy [%]		
		90% Certainty	80% Certainty	70% Certainty
Evening off-peak	Inconclusive	0	0	0
Morning peak	Beta	Reject	Reject	Reject
Afternoon off-peak	Beta	90	76.67	68.33
Evening peak	Inconclusive	0	0	0

5.4.4 Financial Analysis

The case study results show that not all the TOU structure intervals and periods could be successfully modelled with the limited historical generation data available. Therefore, complete models and forecasts for an entire year could not be derived using the available data. However, the historical generation data can be used to calculate the average monetary and annual savings. The generation data stored on the database could be analysed against the TOU tariff structures stored on the database to yield the total generated energy in each respective TOU period. This could be used to calculate the total monetary savings for an analysis timeline using the respective tariff rates of each TOU period.

5.4.4.1 Average Monetary Savings for Generated Energy

The solar plant's average monetary savings from generated energy is calculated using TOU based generation statistics and tariff period charges. The average savings from generated energy can be compared to REFIT rates to determine whether it is more profitable to consume or sell generated energy. Table 5.27 summarises the total generated energy and monetary savings for the entire analysis timeline. Note that this analysis distinguishes between the days that energy was generated historically, i.e. only energy generated on weekdays is used for weekday calculations and so on. From the total generated energy and monetary savings given in table 5.27 it can be calculated that the average savings for the generated energy is 66.04 cents/kWh. Therefore, selling generated energy back to the utility is only profitable at REFIT rates greater than 66.04 cents/kWh.

Table 5.27: Analysis timeline generated energy and monetary savings.

Tariff season	Tariff day	Tariff period	Tariff period charge [cents/kWh]	Generated energy [kWh]	Monetary savings [ZAR]
High Demand	Weekdays	Evening Off-peak	54.72	0	0
		Morning Standard	82.08	0	0
		Morning Peak	298.68	10810.552	32288.96
		Afternoon Standard	82.08	95695.412	78546.79
		Evening Peak	298.68	2.109	6.30
		Evening Standard	82.08	0	0
	Saturday	Evening Off-peak	54.72	0	0
		Morning Standard	82.08	8839.604	7255.55
		Afternoon Off-peak	54.72	13059.859	7146.35
		Evening Standard	82.08	0.56	0.46
	Sunday	Off-peak	54.72	24246.007	13267.42
Low Demand	Weekdays	Evening Off-peak	49.02	233.256	114.34
		Morning Standard	59.28	3878.972	2299.45
		Morning Peak	99.18	120503.741	119515.61
		Afternoon Standard	59.28	566842.345	336024.14
		Evening Peak	99.18	6202.788	6151.93
		Evening Standard	59.28	0	0
	Saturday	Evening Off-peak	49.02	789.917	387.22
		Morning Standard	59.28	55864.729	33116.61
		Afternoon Off-peak	49.02	79369.429	38906.89
		Evening Standard	59.28	1296.882	768.79
	Sunday	Off-peak	49.02	138174.438	67733.11
				Total	
				1125810.600	743529.92

5.4.4.2 Average Annual Savings

The solar plant's average annual savings from generated energy is calculated using TOU based generation statistics and tariff period charges. The average annual savings from generated energy can be used to calculate the payback period of the initial investment on a solar plant.

Table 5.28 summarises the estimated average generated energy and the monetary savings for the year 2014. The average generated energy per half-hour interval is determined for each tariff period from the analysis timeline and used to estimate the generated energy for the year 2014. The generated energy is estimated by multiplying the average generated energy per half-hour interval by the total amount of half-hour intervals in each tariff period for the year 2014. The payback period of the investment is determined by dividing the monetary value of the initial investment by the annual monetary savings.

Table 5.28: Average annual savings from generated energy.

Season	Day	Period	Average half-hourly energy [kWh]	Tariff period charge [cents/kWh]	Generated energy [kWh]	Monetary savings [ZAR]
High Demand	Weekdays	Evening Off-peak	0	54.72	0	0
		Morning Standard	0	82.08	0	0
		Morning Peak	22.115	298.68	8624.804	25760.56
		Afternoon Standard	69.909	82.08	72705.515	59676.69
		Evening Peak	0.007	298.68	1.785	5.33
		Evening Standard	0	82.08	0	0
	Saturday	Evening Off-peak	0	54.72	0	0
		Morning Standard	48.577	82.08	6314.976	5183.33
		Afternoon Off-peak	63.789	54.72	9951.088	5445.24
		Evening Standard	0.007	82.08	0.357	0.29
	Sunday	Off-peak	26.068	54.72	17517.684	9585.68
Low Demand	Weekdays	Evening Off-peak	0.051	49.02	158.889	77.89
		Morning Standard	6.854	59.28	2686.790	1592.73
		Morning Peak	71.244	99.18	83783.454	83096.43
		Afternoon Standard	125.728	59.28	394282.124	233730.44
		Evening Peak	5.649	99.18	4428.879	4392.56
		Evening Standard	0	59.28	0	0
	Saturday	Evening Off-peak	0.660	49.02	566.232	277.57
		Morning Standard	102.133	59.28	39831.962	23612.39
		Afternoon Off-peak	118.148	49.02	55293.332	27104.79
		Evening Standard	5.649	59.28	881.257	522.41
	Sunday	Off-peak	51.588	49.02	94096.557	46126.13
					Total	
					791125.684	526190.46

Stellenbosch University https://scholar.sun.ac.za

6 Conclusions and Recommendations

6.1 Overview

Solar power is attracting considerable attention due to its potential of contributing to sustainable future energy supplies [4] [5]. However, solar power has the drawbacks of being site dependant and intermittent in nature. For this reason, energy producers require accurate forecasting systems for the energy output of their solar plants [10]. Modern prediction systems generally have a forecast horizon of one to two days [5]. However, energy producers are interested in a range of prediction horizons, including long term horizons, to manage power plants and forecast their energy production [14].

Solar power forecasting methodologies are classified into either a numerical prediction approach or a statistical approach [14]. The numerical approach incorporates predicted weather variables, such as solar radiation and temperature, together with PV power output models. The statistical approach of forecasting energy output is based on measured historical generation data and requires less input data and computational efforts [14].

Time Of Use (TOU) based energy generation statistics and forecasting models, i.e. with respect to the time when energy is being generated or consumed, are important in the context of small solar plants operating in conjunction with a local load. Generated energy forecasts and statistics are particularly useful in determining the return on investment of solar plants and conducting a financial analysis on renewable energy feed-in tariffs and TOU tariff structures.

This project aims to develop a long term energy generation forecasting methodology based on measured historical datasets. The methodology is implemented in a software application and a case study is conducted to answer the following key questions:

- Is it possible to forecast the energy output of a PV system in the TOU context using a statistical approach?
    - Is it possible to forecast the energy output with respect to TOU tariff structures, tariff seasons, months of the year, days of the week and hours of the day?
- Is it possible to model the energy output of a PV system using probability distributions which are commonly used to model solar radiation?
    - Which of the probability distributions are suitable and which perform the best?

Stellenbosch University https://scholar.sun.ac.za

The literature review focuses on the development and software implementation of a methodology to forecast and model the long term energy output of a solar plant. This includes software design and modelling processes together with database concepts. Furthermore, the review includes statistical inference methods such as hypothesis testing and goodness of fit testing. The mathematical representation and numerical implementation of six different probability distributions are considered and investigated in depth. The literature review concludes in a brief overview of solar radiation modelling together with PV system configuration and efficiency.

The project consisted of the development and software implementation of a long term forecasting methodology and the main components presented in figure 1.1. This is achieved by accomplishing the following objectives:

- Investigate the feasibility of a long term TOU based forecasting methodology based on historical generation data: It is found that the forecasting methodology can successfully model the energy output of a solar plant within the TOU context.
- Design and implement a relational database: A relational database structure is developed which successfully incorporates all generation data and TOU structure data.
- Design and implement a software application: A long term forecasting software application with database connectivity is successfully designed and implemented.
- Investigate South African TOU tariff structures: Two TOU tariff structures are investigated and implemented in this project.
- Utilize statistical theory and methods in the long term forecasting methodology: Statistical methods such as parameter estimation, frequency distributions and goodness of fit tests are successfully investigated and implemented in the software application.
- Investigate probability distributions commonly used to model solar radiation: Six probability distributions are investigated and successfully implemented in the forecasting software application.
- Conduct a case study for an operational solar plant to achieve the following:
    - Investigate and substantiate the energy output in the TOU context: The generated energy of an operational solar plant is metered and logged from the start of February 2013 to the end of June 2014. Generation data is imported into the developed database and analysed using the forecasting software application.
    - To test and evaluate implemented forecasting methodology and software application: Results show that the implemented forecasting methodology and software application can successfully model the energy output of a solar plant when using monthly generation datasets. However, limited generation data results in a large number of inconclusive models.

Stellenbosch University https://scholar.sun.ac.za

6.2 Results and Conclusions

6.2.1 Design and Development

6.2.1.1 Relational Database

A custom relational database topology is developed to store historical generation data along with all relevant TOU structures. Managing and accessing all stored generation and TOU structure data with a software application requires a fixed referencing system across the entire database, i.e. storing historical generation data as profiles and profile sets.

The developed relational database structure successfully incorporates historical generation data and TOU structures. The database structure enables the software application to seamlessly create custom queries and access all user specified generation data as needed.

6.2.1.2 Profile Analysis Application

The developed software application is required to implement the long term forecasting methodology. This is achieved by satisfying the following requirements:

- Connect to a user selected relational database: A robust database connection is implemented which successfully accesses and manipulates historical generation data on a user selected database.
- Implement statistical methods to derive models from historical generation data stored on a user selected database: Statistical methods of deriving models from historical generation data are successfully implemented and utilised in the developed software application.
- Incorporate TOU structures: TOU structures are implemented in the relational database and are successfully utilised by the software application to analyse the generation data in the TOU context.
- Implement an intuitive graphical user interface: The implemented graphical user interface allows the user to create database connections, select the desired historical generation data from a selected database and utilise any of the analysis modules presented in section 4.3.
- Implement a modular and extensible system design: The implemented software architecture presented in figure 4.3 proves to be a time and code efficient approach to a modular and extensible system design.

Stellenbosch University https://scholar.sun.ac.za

6.2.2 Long Term Forecasting Methodology

The long term forecasting methodology investigated in this project is based on drawing statistical inferences from historical generation data. The methodology involves using parameter estimation, hypothesis testing and goodness of fit tests to determine whether proposed probability distributions fit historical generation data. Two goodness of fit tests are implemented, namely the Chi-squared test and Root Mean Square Error test (RMSE). The Chi-squared test is regarded as the primary goodness of fit test criterion. The RMSE test is used as a supplementary indication of model performance. Statistical inference takes place in the TOU context, i.e. the specific time of day as well as months and seasons of the year during which the energy is generated. This includes TOU tariff structures and half-hourly generation profiles.

The implemented long term forecasting methodology proves to be successful and feasible. Historical generation data is successfully modelled and used to forecast the energy output of the considered solar plant within the TOU context. However, a considerable number of models are inconclusive, i.e. degrees of freedom less than one in the Chi-squared test, due to the limited timespan of the historical data. Furthermore, TOU based historical generation statistics are successfully used to conduct a financial analysis on the monetary savings from generated energy of the solar plant.

6.3 Case Study and Analysis

The overall objective of the case study is to investigate and substantiate the energy output of a solar plant in the TOU context and to test the software implemented forecasting methodology. The case study is conducted as a seasonal and monthly analysis for an operational solar plant from the start of February 2013 to the end of June 2014.

The seasonal analysis is conducted for a half-hourly generation profile, the MegaFlex tariff and the HomeFlex tariff structure with respect to the High Demand and Low Demand seasons. The seasonal analysis results indicate that the variation in daily generated energy is at its smallest during early morning and late afternoon hours and at its largest during mid-day hours. A large portion of the seasonal forecasting models do not fit the historical generation data, i.e. have high Chi-squared values together with degrees of freedom above one, and are therefore conclusively rejected. Furthermore, the Beta probability distribution performs the best of all considered distributions for the largest portion of the seasonal models.

The monthly assessment is conducted for a half-hourly generation profile and the HomeFlex tariff structure with respect to the summer month February and the winter month June. The largest portion of the half-hourly models for the calendar month of June is accepted using the Beta probability distribution. Only one half-hour interval is rejected, however the resultant Chi-squared value is very close to the acceptable percentage point value.

The monthly analysis results also indicate that the Beta and Logistic probability distributions perform the best of all considered probability distributions during the calendar month February. However, the Chi-squared test results for these probability distributions are inconclusive, i.e. have DOF less than one, for a large portion of the models. Therefore, the Exponential distribution is chosen as the best performing conclusive model, i.e. the model with the lowest Chi-squared test value and DOF of one and above. However, the Exponential probability distribution does not fit the generation data and is therefore rejected.

From the case study it is concluded that it is possible to forecast the energy output of a PV system in the TOU context. Results indicate that forecasting the generated energy using seasonal data sets leads to the rejection of a large portion of models. However, the generated energy can be successfully forecasted in the TOU context using monthly datasets. The case study results also indicate that it is possible to model the energy output of a solar plant with probability distributions commonly used to model solar radiation.

6.4 Recommendations

Recommendations for the future development of the long term forecasting methodology are made with respect to the following areas:

- Storing data on a relational database.
- Software application.
- Historical generation data analysis timespan.
- Alternative approach to deriving models from historical generation data.

In the case study the energy output of the solar plant was measured and logged for every single half-hour interval of the day. Therefore, a substantial number of measurements are taken during night-time hours when no solar power is available, i.e. measurements with a measured energy output of zero. This takes up unnecessary space in the relation database and increases the computational time of an analysis. It is therefore recommended to incorporate the functionality of ignoring measurements with a measured energy outputs of zero when importing data into the database.

Stellenbosch University https://scholar.sun.ac.za

The implemented software application is capable of analysing historical generation data against only one probability distribution at a time. This means that the user is required run the same analysis for each different probability distribution, which is a time consuming process. It is therefore recommended to incorporate the functionality of allowing the user to run an analysis against all selected probability distributions simultaneously.

For the case study, only two years' worth of historical data is available for the calendar months of February to June. The half-hourly analysis results for the calendar month of February indicate that the Beta and Logistic distributions perform the best of all the considered probability distributions, but are inconclusive due to the limited timespan of historical generation data. It is therefore recommended to conduct the analysis on more historical generation data as it becomes available. This will result in higher numbers of bins and therefore DOF. With sufficient DOF the Beta and Logistic distributions can be conclusively accepted or rejected as a model for the calendar month February.

It is recommended to implement an empirical approach to creating cumulative distribution functions from historical generation data. The empirical cumulative distribution functions can be used to calculate the exceedance probability values without the generation data fitting a specific probability distribution. This is useful for the cases when the historical generation data cannot be modelled using a known probability distribution.

Stellenbosch University https://scholar.sun.ac.za

7 Works Cited

[1] B. Narayanaswamy, V. K. Garg, and T. S. Jayram, "Prediction Based Storage Management in the Smart Grid," in *IEEE Third International Conference on Smart Grid Communications (SmartGridComm)2012*, 2012, pp. 498 - 503.

[2] O. Pechak, G. Mavrotas, and D. Diakoulaki, "Investigating the Role of the Clean Development Mechanism Within the Global Green Energy Market: The Case of Wind Energy," in *7th International Conference on the European Energy Market (EEM) 2010*, 2010, pp. 1-6.

[3] D. Singh, N. K. Sharma, Y. R. Sood, and R. K. Jarial, "Global Status of Renewable Energy and Market: Future Prospectus and Target," in *International Conference on Sustainable Energy and Intelligent Systems (SEISCON 2011)*, 2011, pp. 171 - 176.

[4] A. Campoccia, L. Dusonchet, E. Telaretti, and G. Zizzo, "Feed-in Tariffs for Grid-connected PV Systems: The Situation in the European Community," in *IEEE Lausanne Power Tech*, 2007, pp. 1981 - 1986.

[5] B. Ernst, F. Reyer, and J. Vanzetta, "Wind Power and Photovoltaic Prediction Tools for Balancing and Grid Operation," in *CIGRE/IEEE PES Joint Symposium Integration of Wide-Scale Renewable Resources Into the Power Delivery System*, 2009, pp. 1 - 9.

[6] W. Mara´nda and M. Piotrowicz, "Sizing of Photovoltaic Array for Low Feed-in Tariffs," in *Proceedings of the 21st International Conference Mixed Design of Integrated Circuits & Systems (MIXDES)*, 2014, pp. 405 - 408.

[7] G. Tina and S.Gagliano, "Probability Analysis of Weather Data for Energy Assessment of Hybrid Solar/Wind Power System," in *4th IASME/WSEAS International Conference on ENERGY, ENVIRONMENT, ECOSYSTEMS and SUSTAINABLE DEVELOPMENT*, Algarve, 2008, pp. 217-223.

[8] S. Vergura and F. Vacca, "Bootstrap Technique for Analysing Energy Data from PV Plant," in *International Conference on Clean Electrical Power* , 2009, pp. 268-275.

[9] L. Xie and M.D. Ilić, "Model Predictive Dispatch in Electric Energy Systems with Intermittent Resources," in *IEEE International Conference on Systems, Man and Cybernetics* , 2008, pp. 42 - 47.

[10] R. Hossain, A. M. T. Oo, and A. B. M. Shawkat Ali, "Hybrid Prediction Method of Solar Power Using Different Computational Intelligence Algorithms," in *22nd Australasian Universities Power Engineering Conference (AUPEC)*, 2012, pp. 1 - 6.

[11] T. Ackermann, G. Andersson, and L. Soder, "Distributed Generation: a Definition," *Electric Power Systems Research*, vol. 57, no. 3, pp. 195–204, 2001.

Stellenbosch University https://scholar.sun.ac.za

[12] M. Gujar, A. Datta, and P. Mohanty, "Smart Mini Grid: An Innovative Distributed Generation Based Energy System," in *IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia)*, 2013, pp. 1 - 5.

[13] B. Narayanaswamy, T.S. Jayram, and V. Nyuk, "Hedging Strategies for Renewable Resource Integration and Uncertainty Management in the Smart Grid," in *3rd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies (ISGT Europe)*, 2012, pp. 1 - 8.

[14] M. G. De Giorgi, P. M. Congedo, and M. Malvoni, "Photovoltaic Power Forecasting Using Statistical Methods: Impact of Weather Data," *IET Science, Measurement and Technology*, vol. 8, no. 3, pp. 90-97, 2014.

[15] A. Di Piazza, M. C. Di Piazza, and G. Vitale, "Statistical Processing of Data Coming from a Photovoltaic Plant for Accurate Energy Planning," ISSIA-CNR, Palermo,.

[16] M.A. Delucchi and M.Z. Jacobson, "Providing All Global Energy with Wind,Water,and Solar Power, Part II: Reliability ,System and Transmission Costs, and Policies," *Energy Policy*, vol. 39, pp. 1170–1190, 2011.

[17] National Energy Regulator of South Africa, "Renewable Energy Feed-In Tariff Phase 2 July 2009," NERSA, Consultation Paper.

[18] H. F. Assuncao, J. F. Escobedo, and A. P. Oliveira, "Modelling Frequency Distributions of 5 Minute-Averaged Solar Radiation Indexes Using Beta Probability Functions," *Theor. Appl. Climatol*, vol. 75, pp. 213-224, 2003.

[19] T.P. Chang, "Investigation on Frequency Distribution of Global Radiation Using Different Probability Density Functions," *IJASE*, vol. 8, no. 2, pp. 99-107.

[20] Y. D. Arthur, K. B. Gyamfi, and S.K. Appiah, "Probability Distributional Analysis of Hourly Solar Irradiation in Kumasi-Ghana," *IJBSR*, vol. 3, no. 3, pp. 63-75, 2013.

[21] M. Y. Sulaiman, W.M.H. Oo, M. A. Wahab, and A. Zakaria, "Application of Beta Distribution Model to Malaysian Sunshine Data," *Renewable Energy*, vol. 18, pp. 573-579, 1999.

[22] R. Elmasri and S.B. Navathe, *Fundamentals of Database Systems*, 6th ed.: Addison-Wesley.

[23] P. Rob and C. Coronel, *Database Systems: Design, Implementation, and Management*, 8th ed. United States of America: Thomson Course Technology, 2009.

[24] K. Kline and D. Kline, *SQL In a Nutshell: A Desktop Quick Reference*. United States of America: O'Reilly & Associates, 2001.

[25] MySQL. (2014) MySQL 5.0 Reference Manual. [Online]. http://dev.mysql.com/doc/refman/5.0/en/index.html

[26] R. Bourdon. (2014) WampServer. [Online]. http://www.wampserver.com/en/

Stellenbosch University https://scholar.sun.ac.za

[27] MySQL. MySQL Workbench. [Online]. http://dev.mysql.com/doc/workbench/en/

[28] Embarcadero, Delphi® XE Product Reviewers Guide, 2010.

[29] Devart. (2014, July) Devart. [Online]. http://www.devart.com/dbx/

[30] S. Teixeira and X. Pacheco, Borland® Delphi™ 6 Developer's Guide, 2001.

[31] M. Petre, "UML in Practice," in *ICSE*, San Francisco, 2013, pp. 722-731.

[32] J. Hunt, *Guide to the Unified Process featuring UML, Java and Design Patterns*.: Springer.

[33] S.W. Amber, *The Elements of UML 2.0 Style*.: Cambridge University Press.

[34] Rational , "Rational Unified Process Best Practices for Software Development Teams," Rational Software White Paper.

[35] D.C. Montgomery and G.C. Runger, *Applied Statistics and Probability for Engineers*, 4th ed.: John Wiley and Sons Inc., 2007.

[36] W. Mendenhall and T.Sincich, *Statistics for Engineering and the Science*. New York, USA: Macmillan Publishing Company, 1992.

[37] M.P. Wand, "Statistical Computing and Graphics," *The American Statistician*, vol. 51, no. 1, pp. 59-64, 1997.

[38] W.G. Cochran, "The $\chi 2$ Test of Goodness of Fit," *The Annals of Mathematical Statistics*, vol. 23, no. 3, pp. 315-345, 1952.

[39] D.W. Scott, "On Optimal and Data-Based Histograms," *Biometrika*, vol. 66, no. 3, pp. 605-610, 1979.

[40] T. Chai and R.R. Draxler, "Root Mean Square Error (RMSE) or Mean Absolute Error(MAE)?," *Geosci. Model Dev. Discuss*, vol. 7, pp. 1525–1534, 2014.

[41] C. A. Williams, "On the Choice of the Number and Width of Classes for the Chi-Square Test of Goodness of Fit," *Journal of the American Statistical Association*, vol. 45, no. 249, pp. 77-86, March 1950.

[42] K. Pearson, "On the Criterion that a given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling," *Philosophical Magazine Series 5*, vol. 50, pp. 157-175, 1900.

[43] T.P. Chang, "Performance Comparison of Six Numerical Methods in Estimating Weibull Parameters for Wind energy Application," *Applied Energy*, vol. 88, no. 1, pp. 272-282, 2011.

[44] S.A. Ahmed, "Comparative Study of Four Methods for Estimating Weibull Parameters for Halabja, Iraq," *IJPS*, vol. 8, no. 5, pp. 186-192.

[45] F. C. Odo, S. U. Offiah, and P. E. Ugwuoke, "Weibull Distribution-Based Model for Prediction of Wind Potential in Enugu, Nigeria," *AASRFC*, vol. 3, no. 2, pp. 1202-1208, 2012.

Stellenbosch University https://scholar.sun.ac.za

[46] P.Z. Peebles Jr, *Probability, Random Variables, and Random Signal Principles*, 4th ed. New York, USA: McGraw-Hill, 2001.

[47] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Special Function," in *Numerical Recipes in C++*. New York, USA: Cambridge University Press, 2002, ch. 6, pp. 217-275.

[48] C.Lanczos, "A Precision Approximation of the Gamma Function," *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, vol. 1, no. 1, pp. 86-96, 1964.

[49] K. Ashenayi and I. Hoballah, "Application of Normal Distribution in Modeling Global Irradiation," in *Twentieth Southeastern Symposium on System Theory*, Charlotte, 1988, pp. 470-474.

[50] R. L. Scheaffer, W. Mendenhall, and D. Wackerly, *Mathematical Statistics with Applications*, 3rd ed. USA: PWS Publishers.

[51] R. de Sa Ferreira, "Time-of-Use Tariff Design Under Uncertainty in Price-Elasticities of Electricity Demand: A Stochastic Optimization Approach," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2285 - 2295, 2013.

[52] K. Campbell, "Implementing Time-of-Use Tariffs in RSA: The Eskom Experience," in *Seventh International Conference on Metering Apparatus and Tariffs for Electricity Supply*, Glasgow, 1992, pp. 42-46.

[53] Eskom, Tariff & Charges Booklet 2014/2015, 2014.

[54] Eskom, HomeFlex Introduction Booklet.

[55] G.M. Masters, *Renewable and Efficient Electric Power Systems*. New Jersey: John Wiley& Sons Inc., 2004.

[56] J. E. Sherry and C. G. Justus, "A Simple Hourly Clear-Sky Solar Radiation Model Based on Meteorological Paramaeters," *Solar Energy*, vol. 30, no. 5, pp. 425-431, 1983.

[57] A.O. Edeoja and C.A. Eloka-Eboka, "Experimental Validation of Hottel's Transmittance Model for Estimating Beam Radiation In Makurdi Location," *American Journal of Engineering Research*, vol. 2, no. 8, pp. 51-57.

[58] M.M. Khan and M.J.Ahmad, "Estimation of Global Solar Radiation Using Clear Sky Radiation in Yemen," *Journal of Engineering Science and Technology Review*, vol. 5, no. 2, pp. 12-19, 2012.

[59] J.E. Hay and D.C. McKay, "Calculation of Solar Irradiance for Inclined Surfaces: Verification of Models which use Hourly and Daily Data," International Energy Agency, IEA Task IX Final Report 1988.

[60] H.C. Hottel, "A Simple Model for Estimating the Transmittance of Direct Solar Radiation

Stellenbosch University https://scholar.sun.ac.za

Through Clear Atmospheres," *Solar Energy*, vol. 18, pp. 129-134, 1976.

[61] B.Y.H. Liu and R.C. Jordan, "The Interrelationship and Characteristic Distribution of Direct, Diffuse and Total Solar Radiation," *Solar Energy*, vol. 4, no. 3, pp. 1-19, 1960.

[62] P. Ineichen, O. Guisan, and R. Perez, "Ground-reflected Radiation and Albedo," *Solar Energy*, vol. 44, no. 4, pp. 207-214, 1990.

[63] E. Román, R. Alonso, P. Ibañez, S. Elorduizapatarietxe, and D. Goitia, "Intelligent PV Module for Grid-Connected PV Systems," *IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS*, vol. 53, no. 4, pp. 1066-1073, 2006.

[64] A. B. Raju, S. Ramesh Karnik, and Rohini Jyoti, "Maximum Efficiency Operation of A Single Stage Inverter Fed Induction Motor PV Water Pumping System," in *First International Conference on Emerging Trends in Engineering and Technology*, 2008, pp. 905-910.

[65] M. Hanif, M. Ramzan, M. Rahman, M. Khan, M. Amin *et al.*, "Studying Power Output of PV Solar Panels at Different Temperatures and Tilt Angles," *ISESCO Journal of Science and Technology*, vol. 8, no. 14, pp. 9-12, 2012.

[66] K. Kurokawa, H. Sugiyama, D. Uchida, K. Sakamoto, and T. Ohshiro, "Sophisticated Verification of Simple Monitored Data for Japanese Field Test Program," in *2nd World Conference and Exhibition on Photovoltaic Solar Energy Conversion*, Vienna, 1998, pp. 1941-1946.

[67] A. Kimber, T. Dierauf, L. Mitchell, C. Whitaker, T. Townsend *et al*, "Improved Test Method to Verify the Power Rating of a Photovoltaic (PV) Project," in *34th IEEE Photovoltaic Specialists Conference (PVSC)*, 2009, pp. 000316-000321.

[68] SMA, PV Inverter SUNNY TRIPOWER 8000TL / 10000TL / 12000TL / 15000TL / 17000TL Installation Manual.

[69] A. Dobos, P. Gilman, and M. Kasberg, "P50/P90 Analysis for Solar Energy Systems Using the System Advisor Model," National Renewable Energy Laboratory, Technical Report CP-6A20-54488, 2012.

Stellenbosch University https://scholar.sun.ac.za

Appendix A.

A.1 Chi-squared Distribution Percentage Points Table

The percentage points of the Chi-squared distribution are summarised in the table below [35].

Table A.1: Percentage points of Chi-squared distribution.

Degree of Freedom	**Percentage Points Chi-squared Distribution**			
	10 %	**5 %**	**2.5 %**	**1%**
1	2.71	3.84	5.02	6.63
2	4.61	5.99	7.38	9.21
3	6.25	7.81	9.35	11.34
4	7.78	9.49	11.14	13.28
5	9.24	11.07	12.83	15.09
6	10.65	12.59	14.45	16.81
7	12.02	14.07	16.01	18.48
8	13.36	15.51	17.53	20.09
9	14.68	16.92	19.02	21.67
10	15.99	18.31	20.48	23.21
11	17.28	19.68	21.92	24.72
12	18.55	21.03	23.34	26.22
13	19.81	22.36	24.74	27.69
14	21.06	23.68	26.12	29.14
15	22.31	25	27.49	30.58
16	23.54	26.30	28.85	32.00
17	24.77	27.59	30.19	33.41
18	25.99	28.87	31.53	34.81
19	27.20	30.14	32.58	36.19
20	28.41	31.41	34.17	37.57
21	29.62	32.67	35.48	38.93
22	30.81	33.92	36.78	40.29
23	32.01	35.17	38.08	41.64
24	33.20	36.42	39.36	42.98
25	34.28	37.65	40.65	44.31

Appendix B.

This appendix presents tables containing the goodness of fit test values for the conducted case study. The Chi-squared test values and Root Mean Square Errors (RMSE) are presented for the High Demand and Low Demand seasons with respect to the following time of use structures:

- Half-hourly generation profile.
- MegaFlex tariff structure.
- HomeFlex tariff Structure.

B.1 Half-hourly Generation Profile

B.1.1 Chi-squared Test Values

Tables B.1 to B.2 summarise the Chi-squared test values for the half-hourly generation profile. The Chi-squared test values are denoted by Value, and the numbers of bins are denoted by NB.

Table B.1: Chi-squared test values for half-hourly generation profile during High Demand.

Period Start	Probability Distribution											
	Normal		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
07:00:00	28.490	2	13.787	2	24.789	2	5.411	2	0	1	1.159	6
07:30:00	8.461	1	1.618	2	6.696	1	1.625	2	0	1	0.569	2
08:00:00	5.498	2	0.190	2	4.219	2	0.152	2	4.587	3	0.061	2
08:30:00	30.722	6	16.816	7	32.815	6	18.389	7	38.538	7	21.768	7
09:00:00	22.953	7	25.931	7	28.440	7	42.803	7	82.527	7	22.311	7
09:30:00	34.736	7	47.379	7	42.005	7	98.159	7	133.210	7	32.113	7
10:00:00	42.550	6	46.818	6	49.953	6	90.509	6	144.155	6	35.534	6
10:30:00	49.684	6	63.241	6	56.894	6	96.758	5	214.936	6	37.256	6
11:00:00	46.236	7	60.130	6	51.351	7	35.545	4	228.372	7	29.511	7
11:30:00	53.916	6	48.657	5	52.741	6	59.400	4	215.865	7	22.566	7
12:00:00	42.940	6	36.574	5	44.677	6	52.386	4	251.503	7	26.389	7
12:30:00	51.468	7	34.454	5	44.581	6	41.883	4	238.706	7	31.019	7
13:00:00	46.616	7	58.242	7	47.952	7	64.332	5	231.394	7	47.980	7
13:30:00	46.654	7	72.281	6	54.366	7	34.821	4	234.515	7	31.762	7
14:00:00	20.192	6	23.346	6	27.969	6	51.774	6	111.911	6	13.196	6
14:30:00	52.963	7	58.887	7	61.191	7	93.260	7	177.428	7	47.782	7
15:00:00	23.473	7	27.352	7	28.555	7	47.911	7	130.591	7	22.810	7
15:30:00	21.175	6	25.353	6	27.835	6	40.347	6	88.538	6	19.243	6
16:00:00	20.556	6	16.755	6	29.826	6	25.963	6	43.111	6	10.620	6
16:30:00	22.003	4	11.647	6	21.276	4	11.484	6	9.627	7	10.537	5
17:00:00	37.750	3	9.110	4	27.109	3	9.134	4	7.466	3	9.393	6
17:30:00	11.885	2	3.436	3	9.833	2	8.430	4	10.933	2	5.572	7

Stellenbosch University https://scholar.sun.ac.za

Table B.2: Chi-squared test values for half-hourly generation profile during Low Demand.

Period Start	Probability Distribution											
	Normal		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
05:30:00	61.092	1	63.661	4	52.268	1	33.606	4	0	1	25.7	4
06:00:00	104.381	2	72.060	6	79.081	2	53.896	6	166.642	2	27.192	7
06:30:00	552.028	8	90.957	8	535.878	8	80.167	8	156.925	8	11.674	8
07:00:00	294.970	8	69.826	8	333.126	8	71.499	8	64.800	8	3.679	8
07:30:00	139.913	8	75.169	8	200.263	8	100.00	8	73.072	8	5.184	8
08:00:00	65.988	8	55.987	8	112.423	8	96.645	8	145.990	8	11.476	8
08:30:00	32.663	9	49.829	9	66.163	9	180.82	9	248.551	9	2.459	9
09:00:00	39.594	9	75.927	9	64.358	9	221.15	6	386.976	9	12.205	9
09:30:00	64.307	10	96.317	8	74.302	10	220.31	5	547.777	10	42.59	10
10:00:00	93.695	9	70.041	6	113.068	10	293.53	4	756.626	10	56.123	10
10:30:00	83.222	7	69.114	6	102.729	8	134.88	3	897.520	10	58.626	8
11:00:00	93.417	6	97.500	5	110.219	7	178.82	3	1038.04	11	72.423	8
11:30:00	75.872	6	95.153	5	86.192	7	144.55	3	1000.79	10	68.621	8
12:00:00	103.203	6	103.45	5	111.414	8	171.36	3	1002.39	10	39.772	8
12:30:00	119.050	6	132.02	5	124.153	8	147.41	3	1003.37	10	45.641	9
13:00:00	88.354	7	103.71	6	154.089	9	352.3	4	966.456	10	67.064	9
13:30:00	110.836	7	137.11	5	118.294	8	137.67	3	1027.76	10	68.869	10
14:00:00	141.695	9	137.07	6	155.473	10	358.37	4	896.998	10	56.673	10
14:30:00	136.031	9	151.17	7	157.741	9	284.91	4	784.688	9	46.177	9
15:00:00	103.392	9	151.15	9	132.387	9	261.38	5	618.460	9	35.64	9
15:30:00	103.668	9	148.96	9	135.718	9	357.23	7	527.223	9	55.578	9
16:00:00	99.705	8	131.57	8	150.919	8	300.81	8	369.496	8	22.008	8
16:30:00	111.919	8	132.54	8	168.850	8	223.26	8	248.166	8	42.20	8
17:00:00	171.449	8	135.30	8	242.923	8	180.9	8	148.041	8	33.404	8
17:30:00	290.796	8	128.58	8	353.024	8	138.30	8	118.081	8	33.912	8
18:00:00	528.911	7	148.75	7	548.363	7	137.51	7	182.231	7	21.406	7
18:30:00	508.760	6	86.091	8	548.982	7	69.691	8	166.642	5	10.37	8
19:00:00	46.497	1	29.990	4	38.584	1	0.062	3	0	1	1.007	3

Stellenbosch University https://scholar.sun.ac.za

B.1.2 Root Mean Square Errors

Tables B.3 to B.4 summarise the RMSE for the half-hourly generation profile.

Table B.3: Root mean square errors for half-hourly generation profile during High Demand.

Period start	Probability Distribution					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
07:00:00	27.908	4.870	26.673	3.728	0.000	0.662
07:30:00	28.175	2.177	25.430	2.263	0.016	1.561
08:00:00	7.164	0.751	6.319	0.702	7.492	1.092
08:30:00	9.088	5.213	8.775	5.977	9.232	6.108
09:00:00	6.919	8.414	7.270	10.231	13.061	7.730
09:30:00	8.973	10.414	9.395	12.286	14.967	9.355
10:00:00	10.953	12.195	10.786	14.571	17.079	11.179
10:30:00	13.736	14.644	14.627	19.208	19.951	12.458
11:00:00	10.630	13.433	11.208	11.980	17.585	9.186
11:30:00	13.208	13.911	13.303	14.355	17.400	7.758
12:00:00	14.107	13.907	14.640	13.489	18.452	9.514
12:30:00	12.212	13.392	14.297	11.837	17.785	9.979
13:00:00	12.571	13.258	12.852	17.089	19.241	12.540
13:30:00	11.923	15.409	12.767	12.140	17.933	10.185
14:00:00	8.071	9.048	8.947	11.666	14.776	6.970
14:30:00	12.246	13.003	12.738	14.794	16.112	12.036
15:00:00	8.537	9.342	9.121	11.152	15.426	8.487
15:30:00	9.165	10.352	10.041	12.244	14.959	8.920
16:00:00	6.724	8.062	7.611	10.190	11.407	6.287
16:30:00	11.116	2.627	11.730	2.844	5.875	6.211
17:00:00	21.333	2.927	17.796	3.255	3.243	4.640
17:30:00	21.437	2.295	19.180	3.356	6.222	1.779

Stellenbosch University https://scholar.sun.ac.za

Table B.4: Root mean square errors for half-hourly generation profile during Low Demand.

Period start	Probability Distribution					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
05:30:00	127.385	20.387	119.551	16.892	0.000	14.068
06:00:00	61.969	12.569	56.746	10.582	26.657	7.706
06:30:00	60.997	14.474	60.978	14.058	27.587	5.152
07:00:00	39.396	15.319	41.002	15.569	13.643	5.070
07:30:00	25.040	17.974	28.639	20.998	13.796	5.560
08:00:00	17.613	15.655	21.998	19.567	25.525	8.519
08:30:00	10.736	12.788	14.876	18.219	29.023	3.498
09:00:00	11.651	13.051	15.124	26.604	35.028	6.229
09:30:00	13.864	17.971	15.887	30.714	36.366	10.481
10:00:00	19.202	23.168	19.491	50.000	40.558	14.824
10:30:00	27.166	23.849	27.101	68.184	43.584	20.286
11:00:00	25.669	32.109	29.243	69.439	41.776	23.239
11:30:00	23.839	32.519	26.811	64.392	45.156	23.007
12:00:00	29.594	37.284	27.953	66.223	44.768	18.124
12:30:00	30.998	43.324	26.872	66.718	44.435	17.032
13:00:00	26.079	26.169	29.386	61.726	44.101	21.627
13:30:00	26.742	43.975	27.355	62.571	44.857	19.337
14:00:00	25.436	30.197	24.882	63.319	42.183	17.811
14:30:00	25.608	29.108	27.766	64.349	43.778	18.235
15:00:00	22.571	23.568	25.294	42.355	39.975	15.203
15:30:00	21.871	23.619	24.552	33.955	37.768	17.697
16:00:00	23.108	25.003	27.416	30.133	34.918	12.502
16:30:00	23.352	25.635	27.753	30.452	28.878	15.449
17:00:00	27.623	26.863	31.539	30.744	21.349	12.978
17:30:00	40.265	26.727	42.982	27.823	22.023	11.767
18:00:00	64.737	25.448	65.845	24.880	34.447	9.424
18:30:00	80.937	13.481	75.524	13.696	27.311	4.568
19:00:00	113.915	15.310	105.350	0.613	0.001	3.820

B.2 MegaFlex Tariff Structure

B.2.1 Chi-squared Test Values

Tables B.5 to B.6 summarise the Chi-squared test values for the MegaFlex tariff structure. The Chi-squared test values are denoted by Value, and the numbers of bins are denoted by NB.

Table B.5: Chi-squared test results for MegaFlex during High Demand.

Day of Week	Tariff Period	Chi-squared test results											
		Normal		Weibull		Logistic		Gamma		Exponential		Beta	
		Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
Weekdays	Evening Off-peak	0	1	0	0	0	0	0	0	0	0	0	0
	Morning Standard	0	0	0	0	0	0	0	0	0	0	0	0
	Morning Peak	16.105	6	23.839	7	17.949	6	30.32223	7	79.141	7	28.321	7
	Afternoon Standard	28.800	7	31.441	7	29.301	7	28.4772	6	209.360	7	28.716	7
	Evening Peak	28.601	1	0.012	1	25.785	1	4.835	2	1.408E-14	1	0.146	4
	Evening Standard	0	0	0	0	0	0	0	0	0	0	0	0
Saturday	Evening Off-peak	0	0	0	0	0	0	0	0	0	0	0	0
	Morning Standard	37.644	7	45.959	7	37.957	7	56.22632	6	213.069	7	39.295	7
	Afternoon Off-peak	35.416	7	36.923	7	42.116	7	68.73606	7	171.035	7	28.812	7
	Evening Standard	28.601	1	0.012	1	25.785	1	4.835	2	1.408E-14	1	0.146	4
Sunday	Off-peak	32.935	7	36.204	7	33.717	7	41.70055	6	223.087	7	33.126	7

Table B.6: Chi-squared test results for MegaFlex during Low Demand.

Day of Week	Tariff Period	Chi-squared test results											
		Normal		Weibull		Logistic		Gamma		Exponential		Beta	
		Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
Weekdays	Evening Off-peak	62.276	1	66.187	4	53.401	1	20.22027	3	5.447E-12	1	26.546	4
	Morning Standard	184.342	4	84.285	9	266.876	5	72.16775	9	135.772	5	20.802	9
	Morning Peak	25.317	9	29.892	9	55.022	9	123.0391	9	226.854	9	4.573	9
	Afternoon Standard	63.755	8	49.790	6	105.962	9	176.9818	4	716.862	10	31.460	9
	Evening Peak	625.434	8	123.370	8	622.810	8	110.5304	8	185.673	8	14.757	8
	Evening Standard	0	0	0	0	0	0	0	0	0	0	0	0
Saturday	Evening Off-peak	103.552	3	93.040	9	147.215	4	79.73743	9	94.389	4	20.831	10
	Morning Standard	27.797	10	23.006	8	47.376	10	135.4237	5	479.777	10	10.449	10
	Afternoon Off-peak	63.981	10	59.986	7	92.221	10	145.4412	4	594.539	10	17.753	10
	Evening Standard	625.434	8	123.370	8	622.810	8	110.5304	8	185.673	8	14.757	8
Sunday	Off-peak	58.946	10	47.414	7	86.358	10	136.1726	4	602.025	10	22.693	10

B.2.2 Root Mean Square Errors

Tables B.7 to B.8 summarise the RMSE for the MegaFlex tariff structure.

Table B.7: Root mean square errors for MegaFlex during High Demand.

Day of week	Tariff period	Root Mean Square Error					
		Normal	Weibull	Logistic	Gamma	Exponential	Beta
Weekdays	Evening Off-peak	0	0	0	0	0	0
	Morning Standard	0	0	0	0	0	0
	Morning Peak	6.209	7.588	5.912	9.107	12.289	7.692
	Afternoon Standard	9.818	10.191	9.777	9.459	17.028	9.986
	Evening Peak	46.477	1.221	44.657	3.030	1.310E-06	0.336
	Evening Standard	0	0	0	0	0	0
Saturday	Evening Off-peak	0	0	0	0	0	0
	Morning Standard	10.735	11.489	10.668	12.485	18.101	11.190
	Afternoon Off-peak	9.962	10.290	10.587	12.296	14.625	9.372
	Evening Standard	46.477	1.221	44.657	3.030	1.310E-06	0.336
Sunday	Off-peak	10.484	10.891	10.582	11.564	17.672	10.607

Table B.8: Root mean square errors for MegaFlex during Low Demand.

Day of week	Tariff period	Root Mean Square Error					
		Normal	Weibull	Logistic	Gamma	Exponential	Beta
Weekdays	Evening Off-peak	128.374	20.100	120.607	12.704	4.627E-05	13.984
	Morning Standard	75.084	12.920	71.562	12.637	22.194	5.201
	Morning Peak	9.929	9.548	14.201	14.005	29.950	4.756
	Afternoon Standard	21.398	23.373	25.436	51.920	39.878	14.754
	Evening Peak	63.892	21.558	64.411	20.935	34.033	6.825
	Evening Standard	0	0	0	0	0	0.000
Saturday	Evening Off-peak	59.415	12.003	66.683	11.898	20.801	5.079
	Morning Standard	10.519	11.886	13.999	25.297	35.597	5.623
	Afternoon Off-peak	16.955	19.986	20.083	48.415	36.764	9.375
	Evening Standard	63.892	21.558	64.411	20.935	34.033	6.825
Sunday	Off-peak	16.781	20.852	20.133	48.996	37.612	10.398

B.3 HomeFlex Tariff Structure

B.3.1 Chi-squared Test Values

Tables B.9 to B.10 summarise the Chi-squared test values for the HomeFlex tariff structure. The Chi-squared test values are denoted by Value, and the numbers of bins are denoted by NB.

Stellenbosch University https://scholar.sun.ac.za

Table B.9: Chi-squared test values for HomeFlex during High Demand.

Tariff period	Chi-squared Test Results											
	Normal		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
Evening Off-peak	0	1	0	0	0	0	0	0	0	0	0	0
Morning Peak	16.105	6	23.839	7	17.949	6	30.322	7	79.141	7	28.321	7
Afternoon Off-peak	28.800	7	31.441	7	29.301	7	28.477	6	209.360	7	28.716	7
Evening Peak	28.601	1	0.012	1	25.785	1	4.835	2	0	1	0.146	4

Table B.10: Chi-squared test values for HomeFlex during Low Demand.

Tariff period	Chi-squared test results Low Demand											
	Normal		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
Evening Off-peak	103.552	3	93.040	9	147.215	4	79.737	9	94.389	4	20.831	10
Morning Peak	25.317	9	29.892	9	55.022	9	123.039	9	226.854	9	4.573	9
Afternoon Off-peak	63.755	8	49.790	6	105.962	9	176.982	4	716.862	10	31.460	9
Evening Peak	625.434	8	123.370	8	622.810	8	110.530	8	185.673	8	14.757	8

Stellenbosch University https://scholar.sun.ac.za

B.3.2 Root Mean Square Errors

Tables B.11 to B.12 summarise the RMSE for the HomeFlex tariff structure.

Table B.11: Root mean square errors for HomeFlex during High Demand.

Tariff period	Root Mean Square Error					
	Normal	**Weibull**	**Logistic**	**Gamma**	**Exponential**	**Beta**
Evening Off-peak	0	0	0	0	0	0
Morning Peak	6.209	7.588	5.912	9.107	12.289	7.692
Afternoon Off-peak	9.818	10.191	9.777	9.459	17.028	9.986
Evening Peak	46.477	1.221	44.657	3.030	0.000	0.336

Table B.12: Root mean square errors for HomeFlex during Low Demand.

Period	Root Mean Square Error					
	Normal	**Weibull**	**Logistic**	**Gamma**	**Exponential**	**Beta**
Evening Off-peak	59.415	12.003	66.683	11.898	20.801	5.079
Morning Peak	9.929	9.548	14.201	14.005	29.950	4.756
Afternoon Off-peak	21.398	23.373	25.436	51.920	39.878	14.754
Evening Peak	63.892	21.558	64.411	20.935	34.033	6.825

Stellenbosch University https://scholar.sun.ac.za

Appendix C.

This appendix presents tables containing the goodness of fit test values for the conducted case study. The Chi-squared test values and Root Mean Square Errors (RMSE) are presented for the calendar months of a year.

C.1 Half-hourly Generation Profile

C.1.1 Statistical Parameters

Tables C.1 to C.12 summarise the half-hourly generation profile statistical parameters for the months of January to December. All night-time and null half-hourly intervals are excluded.

Table C.1: Statistical parameters for half-hourly generation profile during January.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
05:00:00	0.032	0.021	0.001	0.004
05:30:00	34.564	3.484	1.115	0.921
06:00:00	232.986	17.963	7.516	3.143
06:30:00	660.542	31.553	21.308	6.389
07:00:00	1316.69	56.019	42.474	10.640
07:30:00	2197.992	89.686	70.903	12.651
08:00:00	2913.595	114.684	93.987	17.413
08:30:00	3600.815	139.78	116.155	18.238
09:00:00	4098.906	165.849	132.223	29.828
09:30:00	4521.14	175.111	145.843	29.540
10:00:00	5188.579	189.671	167.374	24.366
10:30:00	5488.319	211.663	177.043	29.223
11:00:00	5525.552	222.767	178.244	39.474
11:30:00	5615.129	227.743	181.133	45.889
12:00:00	5623.333	230.538	181.398	44.546
12:30:00	5756.89	229.766	185.706	42.609
13:00:00	5930.566	226.853	191.309	35.721
13:30:00	5857.345	223.368	188.947	35.507
14:00:00	5554.778	215.611	179.186	43.029
14:30:00	5135.526	205.85	165.662	43.654
15:00:00	4743.197	199.332	153.006	47.440
15:30:00	4359.11	179.651	140.616	45.827
16:00:00	3882.688	161.75	125.248	41.966
16:30:00	3243.117	139.17	104.617	37.314
17:00:00	2719.059	115.389	87.712	29.560
17:30:00	2087.74	101.302	67.346	23.482
18:00:00	1386.616	60.841	44.730	15.946
18:30:00	649.441	32.066	20.950	7.958
19:00:00	104.979	7.534	3.386	1.704
19:30:00	9.626	1.069	0.311	0.223

Stellenbosch University https://scholar.sun.ac.za

Table C.2: Statistical parameters for half-hourly generation profile during February.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
06:00:00	69.019	4.424	1.232	1.205
06:30:00	461.27	19.375	8.237	3.990
07:00:00	1328.672	40.429	23.726	8.126
07:30:00	2633.159	67.617	47.021	13.065
08:00:00	4034.06	98.537	72.037	19.310
08:30:00	5404.671	122.849	96.512	22.173
09:00:00	6597.232	145.137	117.808	26.327
09:30:00	7585.176	164.718	135.450	30.097
10:00:00	8648.65	187.097	154.440	29.050
10:30:00	9335.242	195.99	166.701	32.507
11:00:00	10062.59	207.379	179.689	26.575
11:30:00	10459.54	231.033	186.778	24.828
12:00:00	10583.84	218.911	188.997	31.293
12:30:00	10744.23	224.023	191.861	34.205
13:00:00	10727.05	230.134	191.555	36.069
13:30:00	10753.59	224.001	192.028	31.093
14:00:00	10415.87	214.319	185.998	33.694
14:30:00	9812.465	206.416	175.223	33.512
15:00:00	9386.831	193.115	167.622	28.822
15:30:00	8600.168	190.41	153.574	28.922
16:00:00	7627.057	159.265	136.197	22.553
16:30:00	6452.275	149.755	115.219	24.569
17:00:00	5107.331	121.105	91.202	21.618
17:30:00	3595.768	93.899	64.210	18.706
18:00:00	2179.132	58.392	38.913	12.872
18:30:00	700.733	31.386	12.513	7.713
19:00:00	104.42	10.405	1.865	1.609

Table C.3: Statistical parameters for half-hourly generation profile during March.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
06:00:00	1.227	0.297	0.020	0.057
06:30:00	133.253	5.811	2.149	1.568
07:00:00	699.708	20.432	11.286	4.829
07:30:00	1765.256	45.441	28.472	10.811
08:00:00	3286.336	79.101	53.005	18.100
08:30:00	4705.531	101.534	75.896	24.055
09:00:00	6106.171	125.981	98.487	27.435
09:30:00	7291.35	150.59	117.602	31.962
10:00:00	8084.32	165.076	130.392	39.497
10:30:00	8950.065	184.309	144.356	42.594
11:00:00	9614.202	190.758	155.068	46.417
11:30:00	9958.566	199.142	160.622	48.230
12:00:00	10335.14	205.348	166.696	48.398
12:30:00	10477.87	208.135	168.998	47.187
13:00:00	10481.9	206.777	169.063	44.779
13:30:00	10260.23	203.393	165.488	45.053
14:00:00	9700.382	193.086	156.458	44.045

Stellenbosch University https://scholar.sun.ac.za

14:30:00	9184.652	207.312	148.140	43.780
15:00:00	8514.979	194.591	137.338	43.634
15:30:00	7390.51	165.652	119.202	41.696
16:00:00	6207.638	144.073	100.123	37.792
16:30:00	4890.325	138.869	78.876	32.894
17:00:00	3560.089	107.777	57.421	24.848
17:30:00	2055.671	71.235	33.156	17.730
18:00:00	684.408	38.291	11.039	9.496
18:30:00	85.683	6.862	1.382	1.574
19:00:00	3.514	0.665	0.057	0.141

Table C.4: Statistical parameters for half-hourly generation profile during April.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
06:30:00	9.605	1.365	0.160	0.287
07:00:00	241.793	11.299	4.030	2.510
07:30:00	934.822	29.391	15.580	6.028
08:00:00	2032.463	63.156	33.874	11.845
08:30:00	3397.576	81.484	56.626	17.140
09:00:00	4785.973	107.492	79.766	22.077
09:30:00	5933.603	130.353	98.893	25.547
10:00:00	6843.374	146.735	114.056	28.978
10:30:00	7636.144	160.822	127.269	27.860
11:00:00	8363.932	172.812	139.399	28.292
11:30:00	8479.337	180.418	141.322	33.006
12:00:00	8700.114	194.102	145.002	33.827
12:30:00	8749.546	190.822	145.826	34.127
13:00:00	8478.868	186.587	141.314	35.877
13:30:00	8351.656	178.851	139.194	32.100
14:00:00	7984.154	174.491	133.069	32.185
14:30:00	7013.922	166.701	116.899	34.092
15:00:00	6194.851	146.609	103.248	29.150
15:30:00	5220.793	132.435	87.013	29.590
16:00:00	3915.968	108.662	65.266	25.593
16:30:00	2705.339	81.676	45.089	20.041
17:00:00	1469.233	54.553	24.487	12.947
17:30:00	384.286	25.145	6.405	5.784
18:00:00	36.483	4.404	0.608	0.946
18:30:00	0.292	0.128	0.005	0.020

Stellenbosch University https://scholar.sun.ac.za

Table C.5: Statistical parameters for half-hourly generation profile during May.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
07:00:00	36.882	2.845	0.595	0.776
07:30:00	403.844	17.539	6.514	4.028
08:00:00	1129.292	41.465	18.214	9.329
08:30:00	1981.812	55.512	31.965	15.519
09:00:00	3025.851	80.192	48.804	22.268
09:30:00	4096.444	103.695	66.072	27.880
10:00:00	5017.583	117.659	80.929	28.867
10:30:00	5848.092	130.221	94.324	30.503
11:00:00	6388.89	144.438	103.047	34.261
11:30:00	6812.036	155.447	109.872	34.540
12:00:00	7056.458	160.301	113.814	36.276
12:30:00	7064.231	164.713	113.939	38.185
13:00:00	6775.27	174.18	109.279	37.506
13:30:00	6630.662	159.437	106.946	35.474
14:00:00	6192.4	170.907	99.877	35.591
14:30:00	5293.74	149.158	85.383	34.371
15:00:00	4409.057	130.507	71.114	29.847
15:30:00	3509.327	105.001	56.602	23.610
16:00:00	2394.123	67.631	38.615	16.181
16:30:00	1239.763	43.805	19.996	11.217
17:00:00	362.787	21.899	5.851	4.568
17:30:00	28.438	3.728	0.459	0.673
18:00:00	0.008	0.008	0.000	0.001

Table C.6: Statistical parameters for half-hourly generation profile during June.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
07:30:00	96.195	5.695	1.603	1.207
08:00:00	537.241	38.528	8.954	5.604
08:30:00	1209.843	49.217	20.164	10.222
09:00:00	2007.521	62.614	33.459	15.913
09:30:00	2953.949	85.383	49.232	23.276
10:00:00	3824.609	106.509	63.743	29.285
10:30:00	4448.867	118.791	74.148	34.991
11:00:00	5113.274	145.873	85.221	35.865
11:30:00	5711.417	138.707	95.190	36.496
12:00:00	5833.128	144.637	97.219	37.436
12:30:00	5876.79	146.453	97.947	38.250
13:00:00	5603.316	143.442	93.389	39.542
13:30:00	5357.16	125.68	89.286	37.668
14:00:00	4975.346	130.678	82.922	34.847
14:30:00	4341.123	107.319	72.352	30.801
15:00:00	3702.292	91.128	61.705	25.849
15:30:00	2570.505	68.7	42.842	20.946
16:00:00	1684.146	48.192	28.069	14.670
16:30:00	744.425	26.304	12.407	7.546
17:00:00	109.436	4.665	1.824	1.043

Table C.7: Statistical parameters for half-hourly generation profile during July.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
07:00:00	0.002	0.002	0.000	0.000
07:30:00	38.975	4.273	1.257	1.076
08:00:00	268.455	16.586	8.660	4.293
08:30:00	664.339	36.383	21.430	9.072
09:00:00	1184.601	61.003	38.213	15.299
09:30:00	1769.677	84.401	57.086	21.275
10:00:00	2338.316	103.7	75.430	22.287
10:30:00	2803.777	120.54	90.444	25.653
11:00:00	3128.561	131.811	100.921	26.622
11:30:00	3247.517	144.238	104.759	32.029
12:00:00	3356.806	150.255	108.284	34.553
12:30:00	3429.784	166.991	110.638	35.084
13:00:00	3259.345	140.368	105.140	34.522
13:30:00	3257.932	137.279	105.095	31.748
14:00:00	2877.967	129.063	92.838	35.451
14:30:00	2573.52	120.761	83.017	33.659
15:00:00	2135.556	105.063	68.889	31.133
15:30:00	1757.03	85.895	56.678	25.285
16:00:00	1256.511	66.033	40.533	17.600
16:30:00	715.36	41.315	23.076	10.987
17:00:00	226.534	16.922	7.308	4.130
17:30:00	20.577	1.779	0.664	0.447
18:00:00	0.009	0.004	0.000	0.001

Table C.8: Statistical parameters for half-hourly generation profile during August.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
07:00:00	30.3	3.262	0.977	1.116
07:30:00	228.692	26.592	7.377	5.892
08:00:00	598.241	57.936	19.298	12.741
08:30:00	1062.975	63.424	34.290	20.028
09:00:00	1506.212	89.401	48.587	26.592
09:30:00	2030.688	112.657	65.506	31.366
10:00:00	2546.381	131.302	82.141	34.480
10:30:00	2918.778	145.273	94.154	40.768
11:00:00	3412.734	156.575	110.088	38.495
11:30:00	3581.298	167.285	115.526	39.049
12:00:00	3622.573	168.115	116.857	36.404
12:30:00	3446.89	170.139	111.190	42.436
13:00:00	3497.952	191.978	112.837	44.291
13:30:00	3287.889	163.213	106.061	42.214
14:00:00	3085.648	164.123	99.537	44.849
14:30:00	2929.581	157.817	94.503	42.463
15:00:00	2749.371	139.92	88.689	35.729
15:30:00	2068.042	111.127	66.711	32.029
16:00:00	1471.493	87.82	47.468	26.724
16:30:00	969.44	66.287	31.272	18.693
17:00:00	495.67	35.226	15.989	10.268
17:30:00	95	6.723	3.065	1.806
18:00:00	3.342	0.504	0.108	0.154

Stellenbosch University https://scholar.sun.ac.za

Table C.9: Statistical parameters for half-hourly generation profile during September.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
06:00:00	4.349	1.022	0.145	0.267
06:30:00	113.403	15.387	3.780	3.997
07:00:00	503.675	35.788	16.789	9.928
07:30:00	1074.92	68.012	35.831	16.525
08:00:00	1598.927	106.692	53.298	23.028
08:30:00	2160.69	129.11	72.023	30.504
09:00:00	2792.649	138.072	93.088	35.887
09:30:00	3344.415	184.624	111.481	42.430
10:00:00	3853.28	187.964	128.443	41.158
10:30:00	4165.61	198.557	138.854	42.512
11:00:00	4434.418	195.33	147.814	45.351
11:30:00	4472.954	199.699	149.098	46.003
12:00:00	4199.599	202.19	139.987	56.495
12:30:00	4272.456	219.403	142.415	58.093
13:00:00	4079.541	197.989	135.985	57.975
13:30:00	4228.024	209.608	140.934	52.740
14:00:00	3786.964	185.08	126.232	50.927
14:30:00	3569.765	178.13	118.992	49.108
15:00:00	3296.12	170.431	109.871	45.163
15:30:00	2919.66	139.666	97.322	36.107
16:00:00	2325.921	116.436	77.531	32.493
16:30:00	1753.97	89.499	58.466	25.969
17:00:00	985.048	59.833	32.835	16.565
17:30:00	366.967	28.461	12.232	7.052
18:00:00	43.056	2.555	1.435	0.648
18:30:00	0.493	0.139	0.016	0.034

Table C.10: Statistical parameters for half-hourly generation profile during October.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
05:30:00	11.578	2.057	0.373	0.533
06:00:00	152.482	14.092	4.919	3.002
06:30:00	536.3	30.045	17.300	6.838
07:00:00	1191.851	54.977	38.447	11.874
07:30:00	1971.939	86.869	63.611	17.205
08:00:00	2759.466	129.704	89.015	25.255
08:30:00	3353.759	133.811	108.186	28.161
09:00:00	3941.5	154.517	127.145	29.945
09:30:00	4546.462	195.687	146.660	33.115
10:00:00	5004.392	185.412	161.432	37.831
10:30:00	5343.259	200.315	172.363	31.449
11:00:00	5579.973	207.957	179.999	33.310
11:30:00	5555.184	220.519	179.199	42.080
12:00:00	5483.673	216.399	176.893	39.130
12:30:00	5452.322	216.657	175.881	44.804
13:00:00	5389.049	214.135	173.840	40.621
13:30:00	5167.091	207.95	166.680	54.650
14:00:00	4810.457	208.827	155.176	56.441

14:30:00	4563.97	190.115	147.225	47.774
15:00:00	4200.038	175.204	135.485	42.364
15:30:00	3765.404	178.235	121.465	40.303
16:00:00	3092.992	131.076	99.774	34.920
16:30:00	2382.093	103.178	76.842	27.542
17:00:00	1660.844	76.898	53.576	20.927
17:30:00	912.968	48.518	29.451	12.497
18:00:00	200.089	12.375	6.454	3.063
18:30:00	23.438	2.125	0.756	0.538
19:00:00	0.01	0.007	0.000	0.001

Table C.11: Statistical parameters for half-hourly generation profile during November.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
05:00:00	3.884	0.419	0.129	0.145
05:30:00	113.65	6.397	3.788	1.458
06:00:00	414.362	19.347	13.812	4.763
06:30:00	992.738	43.994	33.091	9.549
07:00:00	1709.202	73.073	56.973	16.436
07:30:00	2513.953	102.211	83.798	23.051
08:00:00	3149.574	129.297	104.986	31.406
08:30:00	3793.1	152.69	126.437	35.695
09:00:00	4296.639	173.68	143.221	38.168
09:30:00	4801.349	195.315	160.045	34.905
10:00:00	5037.359	204.363	167.912	38.871
10:30:00	5111.293	216.118	170.376	49.713
11:00:00	5411.776	220.158	180.393	47.009
11:30:00	5575.893	222.82	185.863	47.489
12:00:00	5746.571	225.316	191.552	47.427
12:30:00	5598.35	222.989	186.612	48.174
13:00:00	5531.692	233.569	184.390	47.682
13:30:00	5666.301	225.816	188.877	42.425
14:00:00	5172.468	219.63	172.416	47.039
14:30:00	4964.37	197.636	165.479	41.820
15:00:00	4422.105	178.583	147.404	42.697
15:30:00	3925.396	161.722	130.847	38.137
16:00:00	3473.614	147.175	115.787	34.000
16:30:00	2847.211	122.775	94.907	30.581
17:00:00	2127.97	97.236	70.932	26.146
17:30:00	1359.479	67.933	45.316	17.114
18:00:00	664.882	39.922	22.163	10.553
18:30:00	142.59	12.573	4.753	3.161
19:00:00	15.748	2.334	0.525	0.619
19:30:00	0.3	0.287	0.010	0.051

Stellenbosch University https://scholar.sun.ac.za

Table C.12: Statistical parameters for half-hourly generation profile during December.

Period Start	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
05:00:00	9.325	1.084	0.301	0.260
05:30:00	145.289	11.85	4.687	2.001
06:00:00	487.667	31.19	15.731	4.492
06:30:00	1118.085	47.94	36.067	5.772
07:00:00	1833.219	74.565	59.136	12.769
07:30:00	2523.119	101.929	81.391	14.334
08:00:00	3205.736	127.946	103.411	25.451
08:30:00	3793.713	151.465	122.378	29.771
09:00:00	4305.125	171.319	138.875	34.263
09:30:00	4742.578	187.153	152.986	37.912
10:00:00	5009.322	201.57	161.591	43.856
10:30:00	5281.949	216.318	170.385	46.350
11:00:00	5519.452	223.636	178.047	46.293
11:30:00	5712.985	228.313	184.290	48.583
12:00:00	5717.219	229.86	184.426	48.731
12:30:00	5591.77	229.348	180.380	54.347
13:00:00	5361.883	227.438	172.964	55.959
13:30:00	5319.617	219.223	171.601	50.763
14:00:00	5206.769	212.327	167.960	49.360
14:30:00	4745.881	198.632	153.093	48.735
15:00:00	4500.049	183.194	145.163	44.465
15:30:00	3983.15	166.329	128.489	42.782
16:00:00	3646.906	165.308	117.642	40.065
16:30:00	3229.449	152.904	104.176	26.543
17:00:00	2619.92	126.852	84.514	22.198
17:30:00	1879.385	81.978	60.625	16.006
18:00:00	1189.715	51.454	38.378	11.771
18:30:00	564.651	25.396	18.215	5.834
19:00:00	80.028	4.589	2.582	0.808
19:30:00	4.904	0.434	0.158	0.132

C.1.2 Chi-squared Test Results

Tables C.13 to C.24 summarise the results of the Chi-squared test on the half-hourly generation profile for the months of January to December. The Chi-squared test results for each distinct half-hour interval are given with respect to all hypothesised distribution functions. All night-time and null half-hourly intervals are excluded.

Table C.13: Chi-squared test values for half-hourly generation profile during January.

Period Start	Chi-squared test values January											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
05:00:00	8.037	1	0.003	1	7.345	1	0.002	1	0.000	1	0	1
05:30:00	5.056	3	1.001	3	5.256	3	1.181	3	1.079	3	3.192	4
06:00:00	1.243	2	0.910	3	0.697	2	1.741	3	24.016	4	0.981	3
06:30:00	2.790	4	2.037	4	5.161	4	4.578	4	37.564	4	0.779	4
07:00:00	2.179	3	1.432	3	0.375	2	1.061	2	82.019	4	0.186	3
07:30:00	1.678	2	1.879	3	0.960	2	1.942	2	121.917	4	0.402	3
08:00:00	0.885	2	0.890	2	0.462	2	0.875	2	149.846	4	0.106	2
08:30:00	17.438	3	12.609	3	0.767	2	1.766	2	191.178	3	7.191	3
09:00:00	0.605	2	0.624	2	0.648	2	0.630	2	189.076	4	0.034	2
09:30:00	16.261	3	13.546	3	15.796	3	19.738	3	174.175	3	4.159	3
10:00:00	0.826	1	1.435	2	0.606	1	0.794	1	266.252	4	0.106	2
10:30:00	0.502	2	0.437	2	0.500	2	0.641	2	216.124	4	0.011	2
11:00:00	0.754	2	0.822	2	0.492	2	0.760	2	167.322	4	1.192	3
11:30:00	15.829	3	14.322	3	14.736	3	20.692	3	174.031	4	3.120	3
12:00:00	0.705	2	0.774	2	0.615	2	0.721	2	128.867	4	2.696	3
12:30:00	0.842	2	0.925	2	0.786	2	0.819	2	191.110	4	1.411	3
13:00:00	1.787	2	1.897	2	1.011	2	1.726	2	187.200	3	2.536	3
13:30:00	1.320	2	1.604	2	0.853	2	1.125	2	238.315	4	0.135	2
14:00:00	1.699	2	2.116	2	1.218	2	1.404	2	169.538	4	0.140	2
14:30:00	15.991	3	14.719	3	15.386	3	20.763	3	195.466	4	3.289	3
15:00:00	16.016	3	15.695	3	15.613	3	22.344	3	145.878	4	4.794	4
15:30:00	19.131	3	19.297	3	1.500	2	2.303	2	147.994	4	8.383	4
16:00:00	1.665	2	1.977	2	1.171	2	1.368	2	144.805	4	2.744	4
16:30:00	15.885	3	16.838	3	1.100	2	1.232	2	124.083	4	3.813	4
17:00:00	10.165	3	11.058	3	0.997	2	1.560	2	115.433	4	1.125	4
17:30:00	16.064	3	16.362	3	17.472	3	24.268	3	50.262	4	1.017	4
18:00:00	19.193	3	20.268	3	18.726	3	1.982	2	125.845	4	4.661	4
18:30:00	6.672	3	7.700	3	7.295	3	12.413	3	38.196	4	2.525	4
19:00:00	4.355	4	4.021	4	3.153	4	2.637	4	24.111	4	6.953	4
19:30:00	2.960	2	1.632	2	2.076	2	1.385	2	5.229	3	2.687	2

Table C.14: Chi-squared test values for half-hourly generation profile during February.

Period Start	Chi-squared test values February											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
06:00:00	21.680	4	6.619	4	12.843	3	6.693	4	6.475	4	1.909	5
06:30:00	0.951	4	0.750	5	1.396	4	1.823	5	32.250	6	1.173	5
07:00:00	2.361	5	2.635	5	2.639	5	3.524	4	62.155	5	3.736	5
07:30:00	8.073	4	8.184	4	2.812	3	6.156	3	107.054	5	9.266	5
08:00:00	13.395	3	11.350	3	12.981	3	22.226	3	167.356	5	4.001	4
08:30:00	1.419	2	1.730	2	0.790	2	1.227	2	265.291	6	7.297	3
09:00:00	1.910	2	1.571	2	1.904	2	3.635	2	312.278	6	0.005	2
09:30:00	2.111	2	1.952	2	1.833	2	1.773	1	406.985	6	0.106	2
10:00:00	0.817	1	2.317	2	0.601	1	0.847	1	339.354	7	0.267	2
10:30:00	2.318	1	3.386	2	1.788	1	2.200	1	495.013	7	0.068	2
11:00:00	1.127	1	1.037	1	0.829	1	1.129	1	488.954	8	0.018	1
11:30:00	24.481	4	20.122	4	21.972	4	12.260	3	482.325	5	21.201	4
12:00:00	2.127	2	1.994	2	1.974	2	2.752	2	479.603	6	0.001	2
12:30:00	2.311	2	2.169	2	2.485	2	2.778	2	472.446	6	4.713	3
13:00:00	1.690	2	2.163	2	1.121	2	1.397	2	346.619	6	4.208	3
13:30:00	1.748	2	1.478	2	1.691	2	2.410	2	407.383	6	0	2
14:00:00	2.834	2	3.513	2	2.210	2	2.783	2	535.681	6	8.604	3
14:30:00	2.254	2	2.869	2	1.646	2	2.027	2	458.463	6	6.282	3
15:00:00	3.744	2	2.814	2	3.532	2	2.327	1	587.226	6	0.091	2
15:30:00	1.329	2	1.666	2	0.789	2	1.041	2	304.592	6	8.841	3
16:00:00	1.571	2	1.844	2	1.206	2	1.624	2	401.531	6	0.248	2
16:30:00	1.594	2	1.843	2	0.856	2	1.414	2	265.403	6	7.513	3
17:00:00	16.272	3	11.793	3	15.783	3	3.216	2	180.741	6	6.859	4
17:30:00	22.414	4	11.858	5	26.545	4	13.292	3	93.199	5	7.309	5
18:00:00	6.431	5	5.795	5	9.222	5	13.925	5	80.248	5	2.082	5
18:30:00	4.062	5	0.919	5	6.081	5	1.550	5	11.819	5	2.085	5
19:00:00	0.970	1	0.074	2	0.746	1	0.085	2	0.507	2	0.157	2

Table C.15: Chi-squared test values for half-hourly generation profile during March.

Period Start	Chi-squared test values March											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
06:00:00	13.032	1	0.009	1	11.589	1	0.005	1	5.717E-12	1	0.148	3
06:30:00	7.699	5	3.704	5	11.524	5	5.133	5	9.244	5	1.356	5
07:00:00	4.888	5	6.999	5	7.562	5	10.273	4	48.433	5	3.417	5
07:30:00	15.319	5	10.239	4	10.577	4	11.378	3	85.712	5	8.166	5
08:00:00	19.710	3	20.286	3	20.991	3	37.100	3	119.664	5	13.031	5
08:30:00	32.603	3	33.454	3	29.630	3	1.475	2	244.420	5	5.579	4
09:00:00	1.853	2	2.125	2	1.574	2	2.015	2	311.638	6	7.347	3
09:30:00	1.964	2	1.836	2	2.291	2	3.056	2	336.446	6	2.503	3
10:00:00	2.838	2	3.272	2	2.295	2	2.574	2	349.839	5	12.565	4
10:30:00	2.551	2	3.079	2	1.799	2	2.121	2	355.878	5	9.102	4
11:00:00	4.044	2	4.664	2	3.156	2	3.469	2	419.295	5	8.047	5
11:30:00	3.562	2	3.955	2	3.122	2	3.522	2	401.399	5	8.475	5
12:00:00	4.130	2	4.049	2	4.515	2	5.620	2	418.909	5	12.128	4
12:30:00	3.673	2	3.629	2	3.845	2	5.237	2	445.890	6	12.027	4
13:00:00	4.634	2	4.072	2	5.118	2	8.102	2	463.704	6	3.956	3
13:30:00	3.305	2	3.411	2	3.185	2	4.560	2	469.331	6	12.966	3
14:00:00	3.220	2	3.544	2	2.863	2	3.671	2	382.967	5	9.167	4
14:30:00	28.355	3	27.021	3	26.229	3	0.846	2	218.092	6	10.692	3
15:00:00	31.833	3	31.913	3	30.675	3	2.029	2	212.215	6	10.847	4
15:30:00	33.832	3	36.296	3	32.320	3	56.733	3	193.833	5	12.743	5
16:00:00	37.423	4	43.661	4	39.945	4	57.570	3	155.408	5	9.506	5
16:30:00	23.853	5	29.674	5	25.407	5	30.311	4	105.168	5	21.441	5
17:00:00	9.098	5	12.802	5	10.586	5	25.589	5	71.448	5	8.842	5
17:30:00	6.012	5	7.423	5	10.445	5	13.264	5	30.680	5	3.038	5
18:00:00	9.991	3	4.464	4	9.150	3	4.696	4	2.768	4	0.331	5
18:30:00	2.798	2	1.493	3	2.284	2	1.368	3	1.129	2	2.909	4
19:00:00	11.158	1	0.013	1	9.746	1	4.297	2	3.992E-09	1	1.749	4

Table C.16: Chi-squared test values for half-hourly generation profile during April.

Period Start	Chi-squared test values April											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
06:30:00	9.257	2	3.260	2	8.102	2	1.952	2	2.355E-06	1	0.720	3
07:00:00	3.032	4	0.270	4	2.687	3	0.839	4	10.294	6	1.328	5
07:30:00	4.188	6	3.976	6	6.384	6	8.530	6	50.741	6	3.627	6
08:00:00	5.609	6	6.230	6	5.904	6	6.399	5	78.495	6	5.638	6
08:30:00	17.626	4	16.236	4	18.620	4	19.711	3	124.861	5	7.774	5
09:00:00	1.166	2	1.502	2	0.634	2	1.009	2	217.472	6	8.639	3
09:30:00	0.899	2	28.267	3	0.679	2	0.892	2	257.349	6	11.825	3
10:00:00	29.969	3	26.475	3	27.134	3	3.069	2	237.166	5	7.947	4
10:30:00	22.807	3	17.772	3	21.077	3	2.702	2	228.795	5	5.247	4
11:00:00	1.429	2	1.792	2	0.862	2	1.168	2	290.276	6	7.518	3
11:30:00	26.652	4	20.442	4	26.399	3	36.896	3	223.210	5	16.973	5
12:00:00	15.757	3	11.775	3	15.095	3	0.589	2	205.770	6	1.933	4
12:30:00	29.364	3	24.159	3	27.825	3	1.602	2	292.016	6	2.866	4
13:00:00	25.041	3	16.528	4	24.474	3	36.170	3	181.098	5	13.865	5
13:30:00	16.364	3	13.082	3	14.034	3	1.666	2	233.327	6	1.254	4
14:00:00	33.108	3	27.817	3	31.938	3	1.319	2	199.470	5	3.642	4
14:30:00	16.933	4	14.717	4	28.696	3	41.351	3	150.946	5	15.663	5
15:00:00	9.574	5	7.757	5	7.298	4	13.967	4	124.783	5	3.795	5
15:30:00	14.857	5	14.330	5	18.609	5	23.845	4	97.633	5	5.580	5
16:00:00	7.416	5	7.672	5	11.307	5	16.959	5	59.143	5	2.932	5
16:30:00	8.987	5	8.246	5	14.743	5	15.199	5	37.065	5	3.142	5
17:00:00	9.796	5	4.522	5	15.685	5	6.575	5	10.869	5	2.173	5
17:30:00	2.695	2	1.237	3	1.644	2	1.101	3	1.402	3	7.689	4
18:00:00	5.523	2	0.451	2	4.580	2	0.182	2	0.000	1	1.164	3
18:30:00	16.519	1	0.003	1	15.213	1	0.002	1	8.572E-22	1	0.002	2

Table C.17: Chi-squared test values for half-hourly generation profile during May.

Period Start	Chi-squared test values May											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
07:00:00	19.709	3	2.540	3	15.256	3	1.917	3	4.454	3	2.893	5
07:30:00	8.176	4	1.296	5	9.498	4	2.170	5	15.935	5	2.199	5
08:00:00	7.286	6	9.159	6	5.310	5	14.937	6	39.960	6	7.316	6
08:30:00	11.837	5	14.257	5	19.199	5	28.792	5	43.801	5	0.290	5
09:00:00	25.465	5	32.387	5	32.804	5	35.367	4	78.405	5	8.763	5
09:30:00	25.239	5	30.953	5	30.931	5	46.365	4	108.250	5	7.064	5
10:00:00	16.532	4	19.179	3	18.086	3	33.967	3	127.967	5	1.617	5
10:30:00	8.972	3	9.547	3	7.691	3	1.992	2	174.184	5	2.263	4
11:00:00	15.319	3	16.279	3	14.161	3	1.438	2	161.138	5	1.610	5
11:30:00	13.425	3	13.511	3	12.243	3	0.712	2	162.857	6	3.082	4
12:00:00	15.107	3	15.673	3	13.250	3	1.151	2	179.245	6	8.337	4
12:30:00	25.210	3	25.896	3	25.539	3	46.888	3	135.701	5	4.927	5
13:00:00	10.238	4	11.252	4	12.594	4	36.435	3	192.591	6	4.010	4
13:30:00	26.095	3	25.960	3	28.012	3	48.518	3	165.358	6	4.205	4
14:00:00	7.420	4	9.087	4	5.207	4	11.455	3	131.266	6	31.979	5
14:30:00	21.363	5	26.298	5	22.929	5	22.145	4	103.737	5	18.401	5
15:00:00	21.380	5	25.411	5	25.227	5	11.963	4	87.984	5	16.793	5
15:30:00	8.900	5	9.336	5	13.110	5	18.123	5	52.972	5	5.055	5
16:00:00	4.880	5	5.707	5	8.625	5	15.166	5	49.833	5	1.474	5
16:30:00	13.623	5	6.929	5	19.451	5	8.601	5	14.751	5	5.247	5
17:00:00	0.998	2	1.372	3	1.220	2	1.056	3	0.881	4	2.635	3
17:30:00	5.056	1	0.081	2	4.057	1	0.139	2	0.000	1	0.870	2
18:00:00	22.693	1	0.000	1	21.735	1	0.001	1	0	1	0	1

Table C.18: Chi-squared test values for half-hourly generation profile during June.

Period Start	Chi-squared test values June											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
07:00:00	15.772	1	0.003	1	14.433	1	0.002	1	5.774E-21	1	9.840E-05	2
07:30:00	1.577	3	1.865	3	1.101	3	2.647	3	2.849	4	0.968	3
08:00:00	0.129	1	1.381	2	0.107	1	1.558	2	4.801	3	1.582	2
08:30:00	3.711	4	11.754	5	7.133	4	15.942	5	50.336	6	0.738	4
09:00:00	13.502	5	16.087	5	19.703	5	30.260	5	46.926	5	6.086	5
09:30:00	14.796	5	18.800	5	20.020	5	35.874	5	58.516	5	7.814	5
10:00:00	26.021	4	27.751	4	35.147	4	39.958	4	74.710	4	5.347	4
10:30:00	32.290	4	38.711	4	40.201	4	58.760	4	98.140	4	3.845	4
11:00:00	9.873	5	13.428	5	12.043	5	37.464	4	63.460	5	8.212	5
11:30:00	39.381	4	45.469	4	43.793	4	34.944	3	163.354	5	9.459	5
12:00:00	32.462	4	36.290	4	36.711	4	32.451	3	149.077	5	3.780	5
12:30:00	28.749	4	32.887	4	33.332	4	39.094	3	161.454	5	4.708	5
13:00:00	41.890	5	51.304	5	50.832	5	67.835	4	126.306	5	4.430	5
13:30:00	51.024	4	60.398	4	53.478	4	37.673	3	182.216	4	4.076	4
14:00:00	33.749	4	36.617	4	42.351	4	55.549	4	111.948	4	5.946	4
14:30:00	46.810	4	52.809	4	52.532	4	75.972	4	162.680	4	6.454	4
15:00:00	39.221	4	46.199	4	43.340	4	33.813	3	153.415	4	4.399	4
15:30:00	26.144	4	32.658	4	32.304	4	46.850	4	94.219	4	1.707	4
16:00:00	16.175	4	20.771	4	22.839	4	30.825	4	58.626	4	0.107	4
16:30:00	10.014	4	8.773	4	16.253	4	12.633	4	19.959	4	0.959	4
17:00:00	0.549	4	1.136	5	0.171	4	0.477	5	30.275	6	3.724	5
17:30:00	19.280	3	12.737	3	17.388	3	10.173	3	16.142	3	2.082	4

Stellenbosch University https://scholar.sun.ac.za

Table C.19: Chi-squared test values for half-hourly generation profile during July.

Period Start	Chi-squared test values July											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
07:00:00	9.900	1	0.001	1	9.307	1	0.002	1	3.664E-26	1	0	1
07:30:00	8.094	3	0.968	3	0.464	2	0.426	2	0.442	3	0.374	4
08:00:00	1.131	4	1.301	4	2.810	4	3.003	4	14.716	4	0.289	4
08:30:00	4.949	4	4.873	4	8.016	4	9.185	4	23.972	4	0.843	4
09:00:00	7.484	3	8.971	3	8.423	3	13.963	3	32.677	4	0.989	4
09:30:00	5.516	3	6.576	3	5.336	3	3.221	2	72.628	4	2.753	4
10:00:00	0.662	2	0.848	2	0.329	2	0.747	2	102.887	4	2.330	3
10:30:00	0.541	2	0.539	2	0.764	2	0.578	2	142.803	4	2.700	3
11:00:00	0.543	2	0.563	2	0.646	2	0.604	2	137.972	5	0.006	2
11:30:00	6.395	3	6.183	3	0.438	2	0.577	2	85.869	4	0.529	3
12:00:00	6.825	3	6.682	3	6.959	3	11.750	3	100.273	4	2.406	4
12:30:00	5.830	3	5.557	3	6.518	3	10.475	3	48.925	4	1.546	3
13:00:00	10.705	3	10.594	3	10.938	3	1.191	2	97.530	4	0.731	4
13:30:00	13.251	3	12.693	3	13.140	3	18.827	3	150.386	4	5.957	4
14:00:00	8.471	3	9.495	3	8.668	3	14.687	3	98.904	4	2.484	4
14:30:00	14.075	3	26.943	4	14.887	3	21.833	3	100.439	4	4.885	4
15:00:00	14.537	3	16.399	3	16.426	3	21.361	3	55.541	3	2.841	3
15:30:00	21.729	3	24.128	3	23.631	3	29.659	3	74.508	3	6.990	3
16:00:00	5.689	4	5.985	4	8.458	4	10.541	4	28.589	4	1.323	4
16:30:00	0.654	3	0.609	3	1.272	3	1.408	3	16.033	3	1.824	3
17:00:00	9.394	3	10.148	4	9.819	3	4.035	3	9.678	3	12.364	4
17:30:00	6.272	3	2.045	3	7.486	3	2.229	3	1.450	3	1.887	4
18:00:00	7.049	1	0.006	1	6.356	1	0.006	1	3.343E-11	1	0.038	2

Table C.20: Chi-squared test values for half-hourly generation profile during August.

Period Start	Chi-squared test values August											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
07:00:00	14.113	3	11.370	3	14.599	3	10.628	3	11.248	3	0.146	3
07:30:00	0.529	2	0.879	2	0.896	2	1.128	2	0.758	2	1.355	3
08:00:00	0.453	2	0.883	3	0.186	2	1.038	3	4.615	3	1.331	3
08:30:00	5.702	3	6.112	3	8.122	3	8.166	3	16.431	3	0.525	3
09:00:00	1.159	3	1.462	3	2.362	3	2.615	3	13.356	3	1.090	3
09:30:00	4.859	4	4.882	4	6.861	4	7.620	4	22.464	4	3.949	4
10:00:00	2.228	4	2.300	4	4.030	4	5.308	4	25.141	4	0.977	4
10:30:00	7.816	4	8.837	4	10.545	4	15.294	4	38.134	4	0.317	4
11:00:00	9.020	3	9.266	3	9.745	3	15.473	3	59.167	4	1.794	4
11:30:00	8.784	4	7.913	4	11.846	4	13.235	4	49.750	4	2.582	4
12:00:00	4.675	4	3.841	4	6.913	4	6.879	4	43.328	4	3.067	4
12:30:00	7.837	3	7.463	3	10.462	3	10.939	3	38.214	3	2.545	3
13:00:00	8.132	4	7.213	4	12.149	4	10.109	4	28.411	4	3.621	4
13:30:00	9.573	3	8.924	3	12.415	3	11.580	3	37.281	3	5.203	3
14:00:00	6.031	4	5.759	4	9.060	4	9.610	4	23.772	4	2.214	4
14:30:00	1.091	3	1.014	3	2.127	3	2.142	3	19.036	3	1.631	3
15:00:00	4.326	3	4.004	3	6.502	3	6.409	3	27.324	3	1.345	3
15:30:00	5.560	3	5.866	3	7.735	3	8.348	3	24.550	3	1.505	3
16:00:00	9.719	3	10.434	3	12.750	3	13.189	3	22.863	3	2.549	3
16:30:00	8.904	4	6.500	4	13.071	4	8.337	4	9.495	4	1.483	4
17:00:00	4.291	3	2.947	3	6.109	3	3.953	3	6.485	3	0.807	3
17:30:00	7.850	4	4.938	4	10.887	4	5.827	4	8.001	4	2.921	4
18:00:00	2.741	2	4.404	2	2.105	2	3.279	2	4.338	2	0.174	3

Stellenbosch University https://scholar.sun.ac.za

Table C.21: Chi-squared test values for half-hourly generation profile during September.

Period Start	Chi-squared test values September											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
06:00:00	4.547	2	0.015	1	4.972	2	0.012	1	2.260E-05	1	0.227	3
06:30:00	1.420	2	4.130	2	2.058	2	3.961	2	4.152	2	0.495	3
07:00:00	6.373	4	5.659	4	10.048	4	7.634	4	10.576	4	0.618	4
07:30:00	9.710	4	8.945	4	11.581	4	9.242	4	27.163	4	8.907	4
08:00:00	1.273	4	0.593	4	2.927	4	1.681	4	12.100	4	0.198	4
08:30:00	3.260	4	3.641	4	5.266	4	7.250	4	24.343	4	1.240	4
09:00:00	13.852	4	13.771	4	13.234	3	16.856	3	49.285	4	4.539	4
09:30:00	3.900	4	3.895	4	5.431	4	7.592	4	31.069	4	2.438	4
10:00:00	9.571	4	8.756	4	11.716	4	14.613	4	57.093	4	4.841	4
10:30:00	5.862	4	4.982	4	7.980	4	10.081	4	54.195	4	0.486	4
11:00:00	16.625	3	16.310	3	16.280	3	22.165	3	110.349	3	3.412	3
11:30:00	6.328	3	6.169	3	6.143	3	10.147	3	100.778	4	2.778	4
12:00:00	4.267	3	5.507	3	4.315	3	9.360	3	59.936	4	1.969	4
12:30:00	5.435	3	6.898	3	5.892	3	11.035	3	55.887	4	0.800	4
13:00:00	8.039	3	9.956	3	8.357	3	14.598	3	67.670	4	0.417	4
13:30:00	4.739	3	5.469	3	5.249	3	0.509	2	55.711	4	1.550	4
14:00:00	6.205	3	7.754	3	6.232	3	12.129	3	69.078	4	2.757	4
14:30:00	16.738	3	16.845	3	19.493	3	21.018	3	56.283	3	7.381	3
15:00:00	20.608	3	20.824	3	24.020	3	25.665	3	62.786	3	10.499	3
15:30:00	11.107	3	11.238	3	12.836	3	15.377	3	58.656	3	3.218	3
16:00:00	8.960	3	10.177	3	10.311	3	14.229	3	49.122	3	1.051	3
16:30:00	20.140	3	21.427	3	22.834	3	26.921	3	61.686	3	6.887	3
17:00:00	2.415	3	1.913	3	3.946	3	3.081	3	12.892	3	1.362	3
17:30:00	7.884	4	3.809	4	5.091	3	2.920	3	6.263	3	4.652	4
18:00:00	1.307	4	1.286	4	3.039	4	3.109	4	17.092	4	0.291	4
18:30:00	4.401	1	0.012	1	3.782	1	0.009	1	1.350E-06	1	0.002	2

Stellenbosch University https://scholar.sun.ac.za

Table C.22: Chi-squared test values for half-hourly generation profile during October.

Period Start	Chi-squared test values October											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
05:30:00	2.512	2	3.817	2	2.585	2	2.874	2	0.001	1	0.555	3
06:00:00	0.197	2	0.213	3	0.345	2	0.570	3	4.639	3	0.264	3
06:30:00	1.610	4	1.380	4	2.457	4	1.573	4	25.242	4	2.316	4
07:00:00	1.843	2	2.192	2	1.056	2	2.222	2	70.806	5	0.043	3
07:30:00	1.827	2	2.092	2	1.092	2	1.864	2	90.417	5	0.014	3
08:00:00	0.719	3	0.385	3	1.105	3	4.127	2	86.973	5	0.206	3
08:30:00	1.831	2	2.229	2	1.118	2	1.568	2	177.405	4	1.791	3
09:00:00	1.265	2	1.374	2	1.364	2	1.183	2	182.722	4	0.905	3
09:30:00	2.320	2	9.595	3	1.450	2	2.811	2	89.528	4	4.472	3
10:00:00	2.945	2	3.354	2	2.392	1	2.525	1	235.508	4	0.005	2
10:30:00	22.659	3	19.400	3	21.461	3	26.110	3	234.406	3	7.321	3
11:00:00	1.966	2	2.331	2	1.306	2	1.733	2	225.487	3	0.022	2
11:30:00	1.229	2	1.594	2	0.846	2	0.608	1	190.645	5	0.558	2
12:00:00	15.660	3	13.364	3	0.835	2	1.050	2	177.310	4	3.207	3
12:30:00	15.301	3	14.050	3	2.196	2	3.713	2	174.800	4	2.347	3
13:00:00	0.982	2	1.109	2	0.858	2	0.927	2	157.591	4	0.411	3
13:30:00	2.050	2	2.201	2	1.902	2	1.863	2	139.497	4	2.048	4
14:00:00	1.281	2	1.561	2	0.933	2	1.166	2	115.303	4	3.764	4
14:30:00	1.716	2	2.073	2	1.094	2	1.342	2	116.218	4	0.688	3
15:00:00	1.159	2	1.246	2	1.192	2	1.193	2	115.057	4	0.080	3
15:30:00	2.105	2	2.565	2	1.257	2	2.615	2	99.970	4	4.062	3
16:00:00	1.342	2	1.483	2	1.263	2	1.165	2	127.854	4	2.769	4
16:30:00	1.200	2	1.440	2	0.934	2	1.078	2	115.227	4	2.324	4
17:00:00	20.117	3	22.006	3	20.978	3	0.810	2	93.335	4	3.681	4
17:30:00	8.607	4	10.240	4	5.272	3	7.906	3	37.856	4	3.626	4
18:00:00	6.184	4	5.572	4	8.025	4	6.284	4	22.810	4	5.206	4
18:30:00	1.938	3	0.450	3	2.476	3	0.757	3	2.620	3	0.339	4
19:00:00	8.497	1	0.002	1	7.825	1	0.002	1	4.395E-18	1	0	1

Table C.23: Chi-squared test values for half-hourly generation profile during November.

Period Start	Chi-squared test values November											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
05:00:00	13.667	3	16.008	3	15.267	3	15.161	3	15.522	3	0.459	3
05:30:00	4.568	4	4.587	4	6.802	4	8.642	4	30.690	4	1.741	4
06:00:00	4.925	3	5.498	3	4.484	3	8.986	3	68.975	4	0.255	4
06:30:00	6.151	3	5.575	3	6.146	3	9.828	3	87.592	4	0.777	4
07:00:00	1.001	2	1.133	2	0.940	2	0.929	2	130.217	4	2.225	3
07:30:00	1.795	2	1.889	2	1.896	2	1.831	2	168.848	4	1.004	3
08:00:00	2.626	2	21.387	3	1.697	2	2.292	2	204.952	4	3.891	3
08:30:00	3.510	2	4.218	2	2.311	2	2.748	2	176.113	4	3.028	3
09:00:00	1.770	2	2.044	2	1.527	2	1.532	2	185.163	4	2.071	3
09:30:00	1.197	2	1.394	2	0.725	2	1.090	2	193.388	4	0.034	2
10:00:00	1.810	2	2.122	2	1.063	2	1.659	2	176.120	4	4.303	3
10:30:00	14.767	3	14.382	3	1.027	2	1.625	2	174.966	4	2.056	3
11:00:00	1.744	2	2.122	2	1.206	2	1.443	2	165.214	4	1.295	3
11:30:00	1.835	2	2.071	2	1.576	2	1.146	1	172.337	4	0.034	2
12:00:00	1.655	1	1.883	1	1.284	1	1.423	1	245.516	5	0.004	2
12:30:00	1.965	2	2.259	2	1.648	2	1.232	1	173.121	4	0.043	2
13:00:00	1.019	2	1.244	2	0.650	2	0.897	2	159.065	4	4.128	3
13:30:00	1.369	1	1.607	2	1.066	1	1.273	1	219.574	5	0.017	2
14:00:00	2.282	2	2.756	2	1.483	2	1.760	2	145.271	4	1.331	2
14:30:00	1.944	2	2.094	2	1.887	2	2.127	2	206.113	4	0.073	2
15:00:00	2.163	2	2.478	2	1.791	2	1.866	2	158.906	4	1.487	3
15:30:00	2.275	2	2.102	2	2.750	2	3.208	2	148.910	4	0.156	3
16:00:00	1.170	2	1.284	2	1.165	2	1.130	2	159.643	4	3.083	3
16:30:00	1.689	2	2.073	2	1.045	2	1.529	2	133.689	4	1.852	3
17:00:00	15.198	3	16.774	3	14.194	3	1.792	2	114.883	4	4.109	4
17:30:00	9.389	4	9.871	4	3.494	3	6.924	3	52.493	4	0.522	4
18:00:00	2.231	4	2.691	4	4.405	4	5.417	4	18.471	4	0.053	4
18:30:00	2.898	3	0.077	3	3.222	3	0.022	3	2.045	3	0.492	4
19:00:00	1.583	2	1.513	2	1.246	2	1.259	2	1.643	2	0.315	4
19:30:00	9.303	1	0.001	1	8.711	1	0.002	1	3.538E-24	1	0	1

Table C.24: Chi-squared test values for half-hourly generation profile during December.

Period Start	Chi-squared test values December											
	Gaussian		Weibull		Logistic		Gamma		Exponential		Beta	
	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
05:00:00	1.701	2	0.047	2	0.916	2	0.025	2	0.065	2	0.701	3
05:30:00	0.597	2	0.498	2	0.242	2	0.184	2	26.085	4	0.701	2
06:00:00	1.274	3	1.303	3	2.368	2	2.419	3	74.834	4	1.633	3
06:30:00	3.146	2	2.887	2	2.391	2	3.313	2	148.095	4	2.954	2
07:00:00	2.184	2	2.301	2	1.313	2	2.211	2	117.614	4	2.885	3
07:30:00	1.639	2	1.514	2	0.973	2	1.708	2	128.444	4	1.004	2
08:00:00	1.045	1	1.332	2	0.807	1	1.000	1	154.993	5	0.382	2
08:30:00	1.001	1	1.235	2	0.772	1	0.964	1	163.362	5	0.338	2
09:00:00	1.106	1	1.401	2	0.855	1	1.050	1	167.043	5	0.383	2
09:30:00	1.283	1	1.595	2	0.996	1	1.195	1	195.762	5	0.376	2
10:00:00	1.249	2	1.448	2	1.007	2	1.151	1	194.211	5	0.144	2
10:30:00	0.996	2	1.195	2	0.752	2	0.976	2	164.517	5	0.201	2
11:00:00	0.991	2	1.065	2	0.893	2	0.939	1	163.829	5	0.067	2
11:30:00	1.584	2	1.989	2	1.104	2	1.172	1	178.064	5	0.692	2
12:00:00	1.527	2	1.924	2	1.048	2	1.088	1	155.195	5	0.731	2
12:30:00	1.318	2	1.524	2	1.068	2	1.210	2	154.465	4	3.098	3
13:00:00	1.059	2	1.237	2	0.886	2	0.952	2	127.875	4	0.233	3
13:30:00	1.151	2	1.317	2	0.980	2	1.134	2	127.765	4	1.891	3
14:00:00	1.297	2	1.445	2	1.165	2	1.359	2	169.292	4	4.886	3
14:30:00	1.477	2	1.797	2	0.948	2	1.174	2	155.217	4	3.419	3
15:00:00	2.128	2	2.561	2	1.376	2	1.538	2	141.824	4	4.340	3
15:30:00	1.976	2	2.384	2	1.220	2	1.555	2	128.241	4	0.616	3
16:00:00	20.927	3	21.431	3	21.049	3	28.966	3	101.130	4	6.571	4
16:30:00	8.987	3	6.874	3	1.090	2	2.419	2	144.514	5	3.921	3
17:00:00	1.996	3	1.278	3	2.883	3	4.176	3	138.733	5	0.317	3
17:30:00	7.815	3	6.495	3	8.248	3	11.844	3	85.184	4	3.615	4
18:00:00	18.171	3	17.613	3	18.546	3	24.000	3	112.034	3	4.669	3
18:30:00	10.200	3	9.853	3	11.213	3	15.022	3	76.123	3	1.626	3
19:00:00	0.968	4	1.137	4	1.228	4	0.398	4	32.634	4	0.777	5
19:30:00	1.821	3	1.553	3	1.597	3	1.791	3	3.823	3	3.393	3

C.1.3 Root Mean Square Errors

Tables C.25 to C.36 summarise the results of RMSE test on the half-hourly generation profile for the months of January to December. RMSE test results for each distinct half-hour interval are given with respect to all hypothesised distribution functions. All night-time and null half-hourly intervals are excluded.

Table C.25: Root mean square errors January.

Period Start	Root Mean Square Errors January					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
05:00:00	12.269	0.283	11.858	0.266	0.000	0
05:30:00	4.385	0.957	4.276	1.013	1.187	2.058
06:00:00	2.868	1.525	2.196	2.426	6.093	1.553
06:30:00	1.982	1.736	2.599	2.504	4.995	1.002
07:00:00	2.896	2.462	2.031	3.156	7.342	0.834
07:30:00	3.507	2.606	2.746	3.793	8.119	0.898
08:00:00	3.052	3.027	2.339	3.056	8.993	0.717
08:30:00	7.796	6.554	2.729	3.844	12.008	4.317
09:00:00	2.437	2.573	1.907	2.386	10.978	0.320
09:30:00	7.637	6.980	7.334	8.391	11.892	2.821
10:00:00	4.583	4.076	3.970	4.500	11.626	0.535
10:30:00	2.264	2.312	1.850	2.164	11.044	0.196
11:00:00	2.876	2.986	2.178	2.886	9.944	2.164
11:30:00	7.778	7.450	7.382	8.762	9.970	2.991
12:00:00	2.676	2.827	2.049	2.722	8.965	2.484
12:30:00	2.903	3.112	2.286	2.786	10.905	2.248
13:00:00	4.196	4.309	3.342	4.131	12.919	2.579
13:30:00	3.834	4.126	3.155	3.590	11.507	0.737
14:00:00	4.312	4.730	3.634	3.855	10.962	0.732
14:30:00	7.429	7.203	7.016	8.376	10.530	2.355
15:00:00	7.486	7.503	7.127	8.685	9.820	2.812
15:30:00	7.781	7.974	3.912	4.660	10.428	2.806
16:00:00	4.092	4.418	3.290	3.732	11.023	2.488
16:30:00	7.257	7.592	2.751	3.411	9.818	2.509
17:00:00	6.395	6.649	3.069	3.890	9.354	1.461
17:30:00	8.004	8.099	8.239	9.598	6.736	1.282
18:00:00	8.265	8.576	7.920	4.227	9.876	2.984
18:30:00	5.170	5.493	5.375	6.806	6.370	2.478
19:00:00	3.340	3.119	2.812	2.420	7.405	4.090
19:30:00	4.355	2.564	3.641	2.342	3.614	3.482

Stellenbosch University https://scholar.sun.ac.za

Table C.26: Root mean square errors February.

Period Start	Root Mean Square Errors February					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
06:00:00	7.897	2.139	7.759	2.153	2.196	1.442
06:30:00	1.771	0.702	1.517	1.731	7.793	1.341
07:00:00	2.367	2.656	1.835	2.786	9.276	3.376
07:30:00	5.365	5.711	4.458	6.482	10.494	3.907
08:00:00	10.109	9.438	9.787	12.400	12.613	3.517
08:30:00	5.312	5.659	4.173	5.082	12.745	6.700
09:00:00	4.860	5.313	4.109	4.452	13.971	0.157
09:30:00	5.675	6.243	4.894	9.118	16.019	0.673
10:00:00	6.306	3.918	5.455	6.415	12.670	0.885
10:30:00	10.292	6.351	9.152	10.054	14.795	0.474
11:00:00	7.328	7.052	6.352	7.336	12.391	0.999
11:30:00	9.037	8.362	7.792	9.584	15.417	8.845
12:00:00	5.756	6.292	4.882	5.352	14.496	0.093
12:30:00	5.589	6.075	4.654	5.229	14.275	4.256
13:00:00	5.984	6.452	5.016	5.539	13.109	4.594
13:30:00	5.077	5.445	4.277	4.765	13.478	0.066
14:00:00	7.531	8.390	6.539	6.866	15.790	5.282
14:30:00	6.896	7.600	5.893	6.290	14.961	4.637
15:00:00	6.005	6.655	5.196	10.310	16.650	0.562
15:30:00	4.942	5.135	4.067	4.680	12.721	7.356
16:00:00	5.791	6.263	4.917	5.381	13.497	1.116
16:30:00	4.883	4.999	3.864	4.824	12.734	6.925
17:00:00	9.532	8.459	8.974	6.689	10.387	3.486
17:30:00	7.205	3.641	7.302	9.073	9.324	3.348
18:00:00	3.266	3.267	3.572	4.823	8.282	2.092
18:30:00	2.514	1.026	2.856	1.642	5.548	2.201
19:00:00	6.900	0.436	6.100	0.480	1.338	0.628

Table C.27: Root mean square error March.

Period Start	Root Mean Square Errors March					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
06:00:00	22.646	0.741	21.630	0.577	0.000	0.567
06:30:00	3.848	2.662	4.575	3.187	4.442	1.901
07:00:00	2.470	3.388	2.689	6.855	9.390	2.862
07:30:00	4.426	5.054	4.434	8.624	11.236	3.899
08:00:00	12.247	12.545	12.242	15.629	12.802	5.235
08:30:00	14.572	14.948	13.312	5.408	16.488	4.393
09:00:00	6.309	6.900	5.142	5.607	15.293	5.458
09:30:00	5.279	5.761	4.354	4.833	15.788	3.632
10:00:00	7.652	8.282	6.289	6.707	19.411	5.953
10:30:00	7.516	8.163	6.145	6.636	19.606	6.052
11:00:00	9.109	9.780	7.728	7.827	21.133	4.642
11:30:00	8.238	8.871	6.917	7.119	20.689	4.519
12:00:00	7.384	7.966	6.283	6.485	21.149	6.524
12:30:00	7.234	7.855	6.133	6.351	17.993	6.493
13:00:00	6.724	7.310	5.773	6.040	18.322	3.515
13:30:00	7.416	8.080	6.290	6.473	18.478	6.451
14:00:00	7.835	8.507	6.615	6.792	20.253	4.562
14:30:00	14.129	14.014	13.295	4.432	14.006	8.732

Stellenbosch University https://scholar.sun.ac.za

15:00:00	15.421	15.582	14.760	6.191	13.761	6.604
15:30:00	15.806	16.464	14.981	18.854	15.012	4.873
16:00:00	10.768	11.390	10.868	19.670	13.635	4.546
16:30:00	7.696	8.466	7.706	11.488	12.412	7.778
17:00:00	4.653	5.691	4.760	7.134	10.998	4.879
17:30:00	3.459	4.535	4.285	5.961	7.519	2.838
18:00:00	8.847	2.960	8.080	2.906	1.776	0.748
18:30:00	7.647	1.209	6.365	1.110	2.049	2.088
19:00:00	21.308	0.902	20.187	2.614	0.000	1.229

Table C.28: Root mean square errors April.

Period Start	**Root Mean Square Errors April**					
	Gaussian	**Weibull**	**Logistic**	**Gamma**	**Exponential**	**Beta**
06:30:00	10.382	2.323	9.517	1.990	0.012	1.183
07:00:00	3.373	0.886	4.593	1.499	4.184	1.590
07:30:00	3.026	3.089	3.500	4.147	7.198	2.873
08:00:00	3.006	3.298	2.874	3.728	8.802	3.245
08:30:00	6.683	6.694	6.429	12.253	11.410	3.802
09:00:00	5.029	5.497	3.863	4.818	13.228	7.958
09:30:00	4.520	14.386	3.449	4.381	13.820	8.883
10:00:00	14.731	14.150	13.658	7.456	14.315	4.633
10:30:00	12.966	11.729	12.113	6.951	13.072	2.755
11:00:00	5.614	5.966	4.570	5.264	13.157	6.520
11:30:00	8.869	8.073	14.085	16.593	12.955	5.593
12:00:00	11.159	9.885	10.729	3.593	11.372	1.453
12:30:00	14.957	13.814	14.289	5.535	13.381	2.614
13:00:00	13.845	7.750	13.386	16.100	12.490	5.610
13:30:00	11.599	10.605	10.630	5.786	11.778	1.714
14:00:00	15.484	14.465	14.827	5.183	13.479	2.125
14:30:00	6.076	5.936	14.016	16.655	12.595	6.265
15:00:00	4.908	4.562	3.925	5.971	10.599	3.433
15:30:00	5.644	5.681	5.998	7.641	9.963	3.767
16:00:00	3.538	3.947	3.993	5.570	8.131	2.545
16:30:00	3.649	4.003	4.257	5.428	6.648	2.607
17:00:00	4.528	3.211	5.490	3.764	3.906	2.423
17:30:00	7.484	3.020	6.024	2.652	3.785	5.684
18:00:00	10.206	0.830	9.214	0.550	0.043	2.066
18:30:00	24.289	0.409	23.549	0.361	0.000	0.062

Stellenbosch University https://scholar.sun.ac.za

Table C.29: Root mean square errors May.

Period Start	Root Mean Square Errors May					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
07:00:00	12.477	1.789	10.854	1.461	1.986	2.135
07:30:00	4.575	1.846	4.413	2.554	6.736	2.356
08:00:00	3.403	4.186	3.036	5.113	7.904	3.915
08:30:00	4.359	4.810	5.360	6.109	7.555	0.782
09:00:00	6.488	7.497	6.892	9.807	10.002	4.423
09:30:00	7.660	8.266	8.168	11.141	11.520	4.164
10:00:00	7.935	12.057	11.135	14.811	12.600	2.017
10:30:00	8.312	8.671	7.330	6.405	14.283	2.219
11:00:00	10.865	11.285	10.074	5.541	13.646	1.877
11:30:00	10.404	10.535	9.694	3.737	11.586	3.493
12:00:00	11.565	11.787	10.784	3.325	12.287	6.476
12:30:00	13.539	13.868	13.149	16.783	12.783	2.862
13:00:00	4.545	4.806	5.050	16.234	13.329	3.108
13:30:00	13.544	13.689	13.541	16.882	12.350	3.175
14:00:00	5.955	6.666	4.528	9.114	11.443	10.122
14:30:00	7.678	8.267	7.858	9.114	12.320	7.381
15:00:00	7.814	8.289	8.445	6.562	11.323	7.019
15:30:00	4.396	4.814	5.114	6.421	8.216	3.493
16:00:00	3.234	3.669	4.083	5.275	8.345	2.049
16:30:00	5.235	4.316	5.990	4.880	4.510	3.705
17:00:00	4.276	2.516	3.391	1.776	2.239	4.839
17:30:00	15.356	0.691	13.960	0.749	0.018	1.751
18:00:00	27.842	0.080	27.417	0.231	0	0

Table C.30: Root mean square error June.

Period Start	Root Mean Square Errors June					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
07:00:00	23.871	0.428	23.083	0.330	0.000	0.014
07:30:00	3.415	3.724	2.612	4.356	2.990	2.522
08:00:00	2.696	2.061	2.457	2.189	5.924	2.272
08:30:00	3.981	6.444	5.521	7.658	7.834	1.802
09:00:00	5.332	6.006	6.173	7.406	8.128	3.971
09:30:00	5.516	6.349	6.076	7.701	8.869	4.572
10:00:00	8.668	9.173	9.626	10.736	10.362	4.058
10:30:00	9.780	10.804	10.482	12.390	12.861	3.451
11:00:00	4.316	5.057	4.428	9.996	9.727	4.713
11:30:00	10.581	11.159	10.875	14.439	13.316	3.986
12:00:00	11.599	11.864	12.401	14.177	12.765	3.509
12:30:00	10.387	10.743	11.156	15.628	13.155	3.976
13:00:00	8.902	9.355	9.554	13.203	11.433	3.169
13:30:00	13.473	14.224	13.651	15.439	17.311	3.720
14:00:00	10.451	10.858	11.367	12.479	12.784	4.768
14:30:00	12.883	13.443	13.504	14.913	15.628	5.024
15:00:00	12.156	12.767	12.721	14.253	15.981	4.345
15:30:00	9.442	10.149	10.338	11.399	12.542	2.532
16:00:00	6.822	7.744	7.810	9.113	10.121	0.585
16:30:00	5.376	5.386	6.571	6.457	6.460	1.897
17:00:00	1.842	1.895	0.762	0.893	6.745	3.449
17:30:00	10.873	4.760	9.531	4.278	4.472	2.436

Table C.31: Root mean square errors July.

Period Start	Root Mean Square Errors July					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
07:00:00	13.255	0.144	12.958	0.227	0.000	0
07:30:00	5.774	1.477	2.189	1.100	0.863	0.991
08:00:00	1.182	1.439	1.777	2.243	4.221	0.757
08:30:00	2.591	2.769	3.091	3.734	4.556	1.122
09:00:00	5.492	5.916	5.811	7.209	6.231	1.569
09:30:00	4.814	5.207	4.701	4.993	8.303	2.124
10:00:00	2.644	2.907	1.937	2.794	9.090	2.949
10:30:00	2.075	2.253	1.702	2.099	10.455	3.032
11:00:00	2.248	2.437	1.772	2.192	8.170	0.148
11:30:00	5.141	5.101	1.628	2.434	8.084	1.488
12:00:00	5.134	5.136	5.015	6.527	8.489	2.365
12:30:00	5.002	4.879	5.261	6.476	6.640	2.609
13:00:00	5.893	6.014	5.566	3.382	8.469	1.131
13:30:00	6.814	6.761	6.520	8.001	9.807	3.325
14:00:00	5.372	5.786	5.128	6.880	8.803	2.176
14:30:00	7.020	7.382	7.019	8.671	9.010	3.280
15:00:00	6.883	7.377	7.198	8.355	8.861	2.737
15:30:00	8.318	8.638	8.733	9.441	9.656	4.258
16:00:00	2.781	3.082	3.166	4.001	4.876	1.457
16:30:00	1.316	1.220	1.767	1.668	5.043	2.419
17:00:00	5.912	4.618	5.937	3.520	5.507	5.095
17:30:00	4.857	2.228	5.145	2.016	1.293	1.831
18:00:00	11.673	0.436	11.214	0.424	0.000	0.597

Table C.32: Root mean square errors August.

Period Start	Root Mean Square Errors August					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
07:00:00	6.575	3.102	6.431	3.054	3.230	0.510
07:30:00	1.782	1.610	1.583	1.813	1.280	1.611
08:00:00	2.122	1.139	1.455	1.426	3.071	1.366
08:30:00	4.038	4.007	4.829	4.488	4.818	1.256
09:00:00	1.702	1.850	2.444	2.385	5.028	1.931
09:30:00	2.686	3.080	2.889	3.859	4.648	2.800
10:00:00	1.649	1.850	2.100	2.741	4.678	1.481
10:30:00	3.511	3.757	3.953	4.549	5.555	0.631
11:00:00	5.624	5.810	5.566	7.177	6.931	1.355
11:30:00	3.539	3.474	3.881	4.431	5.633	1.602
12:00:00	2.444	2.308	2.696	3.251	5.325	2.065
12:30:00	4.953	4.902	5.636	5.864	6.720	2.416
13:00:00	3.502	3.338	4.119	3.882	4.437	2.398
13:30:00	5.299	5.152	6.027	5.765	6.584	3.355
14:00:00	2.766	3.018	3.160	3.882	4.423	1.906
14:30:00	1.645	1.584	2.297	2.155	5.236	2.268
15:00:00	3.567	3.502	4.316	4.344	5.816	1.706
15:30:00	3.999	4.037	4.722	4.638	5.524	1.865
16:00:00	5.373	5.204	6.184	5.600	5.354	2.710
16:30:00	3.589	2.884	4.178	3.171	2.488	1.700
17:00:00	3.648	2.356	4.302	2.523	2.861	1.720
17:30:00	3.379	2.241	3.726	2.180	2.592	2.350
18:00:00	4.911	2.879	4.090	2.600	3.154	0.666

Stellenbosch University https://scholar.sun.ac.za

Table C.33: Root mean square errors September.

Period Start	Root Mean Square Errors September					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
06:00:00	4.829	0.671	4.336	0.586	0.026	0.435
06:30:00	2.652	3.290	2.406	3.246	3.346	0.807
07:00:00	2.855	2.387	3.421	2.716	2.815	1.028
07:30:00	4.476	4.013	4.724	3.593	6.219	4.354
08:00:00	1.263	0.891	1.759	1.651	4.108	0.638
08:30:00	2.278	2.642	2.578	3.650	5.080	1.638
09:00:00	4.201	4.419	6.362	7.443	6.043	2.250
09:30:00	2.496	2.724	2.572	3.713	5.328	2.206
10:00:00	3.802	3.768	3.864	4.796	6.282	2.661
10:30:00	3.054	2.882	3.420	3.873	5.771	0.702
11:00:00	7.683	7.610	7.570	8.794	10.905	3.094
11:30:00	4.842	4.804	4.679	5.988	7.833	2.062
12:00:00	3.898	4.440	3.691	5.477	7.384	1.520
12:30:00	4.551	5.064	4.672	6.244	7.176	1.286
13:00:00	5.219	5.851	5.096	6.873	7.649	0.868
13:30:00	4.090	4.454	4.069	2.193	7.396	1.357
14:00:00	4.798	5.342	4.669	6.443	7.853	2.109
14:30:00	7.030	7.104	7.550	7.888	8.065	3.949
15:00:00	7.876	7.830	8.556	8.469	8.407	5.198
15:30:00	5.830	5.879	6.214	6.801	7.974	2.480
16:00:00	5.345	5.657	5.684	6.575	7.773	1.565
16:30:00	7.922	8.287	8.319	9.245	9.074	4.194
17:00:00	2.537	2.075	3.263	2.409	4.169	1.930
17:30:00	3.459	2.177	3.467	1.759	3.307	2.948
18:00:00	1.288	1.270	1.890	2.005	4.198	0.751
18:30:00	9.499	0.591	8.927	0.504	0.006	0.074

Table C.34: Root mean square error October.

Period Start	Root Mean Square Errors October					
	Gaussian	Weibull	Logistic	Gamma	Exponential	Beta
05:30:00	4.323	2.504	3.705	2.293	0.126	1.215
06:00:00	1.346	0.862	1.144	1.480	3.169	0.820
06:30:00	1.775	1.486	2.232	1.351	5.940	1.991
07:00:00	3.749	4.053	2.947	4.124	6.597	0.274
07:30:00	3.692	3.901	2.956	3.785	7.115	0.240
08:00:00	1.547	1.202	1.737	5.387	7.074	0.956
08:30:00	4.331	4.681	3.501	4.062	10.909	2.136
09:00:00	3.225	3.481	2.633	3.063	10.376	1.401
09:30:00	4.076	6.497	3.277	4.509	7.886	4.389
10:00:00	5.134	5.666	7.498	7.675	13.006	0.129
10:30:00	8.886	8.210	8.413	9.513	13.218	3.588
11:00:00	4.501	4.836	3.699	4.263	14.143	0.266
11:30:00	3.769	4.129	3.214	3.978	9.634	1.285
12:00:00	7.710	7.181	2.418	3.185	9.489	2.896
12:30:00	7.485	7.258	4.561	5.577	10.077	2.241
13:00:00	3.149	3.394	2.490	3.000	9.911	1.243
13:30:00	4.229	4.498	3.545	3.662	10.966	1.640
14:00:00	3.552	3.900	2.781	3.417	9.962	2.425
14:30:00	4.235	4.566	3.435	3.806	10.155	1.354

Stellenbosch University https://scholar.sun.ac.za

15:00:00	3.197	3.474	2.594	2.880	10.065	0.372
15:30:00	3.959	4.345	3.146	4.422	9.527	4.175
16:00:00	3.469	3.751	2.798	3.179	10.564	2.358
16:30:00	3.429	3.762	2.690	3.275	9.997	2.193
17:00:00	8.227	8.731	8.059	2.775	8.946	2.791
17:30:00	3.463	3.993	4.064	5.157	6.394	2.512
18:00:00	3.585	3.070	3.935	2.811	6.154	3.333
18:30:00	2.708	0.867	2.888	1.174	2.733	0.703
19:00:00	12.528	0.259	12.146	0.274	0.000	0

Table C.35: Root mean square errors November.

Period Start	**Root Mean Square Errors November**					
	Gaussian	**Weibull**	**Logistic**	**Gamma**	**Exponential**	**Beta**
05:00:00	6.028	3.768	6.018	3.716	3.800	1.059
05:30:00	2.719	2.903	2.994	3.936	5.171	1.854
06:00:00	4.420	4.622	4.219	5.753	7.146	0.595
06:30:00	4.770	4.595	4.623	5.897	7.346	0.882
07:00:00	2.998	3.262	2.374	2.907	9.407	2.639
07:30:00	3.659	3.978	3.069	3.275	10.942	1.249
08:00:00	4.872	8.604	3.992	4.593	11.486	2.730
08:30:00	5.638	6.045	4.797	5.115	10.994	2.511
09:00:00	3.982	4.331	3.281	3.626	10.911	2.104
09:30:00	3.527	3.750	2.786	3.393	10.475	0.461
10:00:00	4.189	4.467	3.360	4.036	9.999	2.772
10:30:00	7.244	7.215	3.035	3.871	10.403	2.184
11:00:00	4.184	4.549	3.429	3.825	10.423	1.541
11:30:00	4.145	4.541	3.566	5.219	11.562	0.340
12:00:00	6.149	6.508	5.495	5.751	11.024	0.088
12:30:00	4.334	4.740	3.739	5.393	11.519	0.388
13:00:00	3.294	3.575	2.594	3.109	10.454	3.361
13:30:00	5.760	3.631	5.147	5.575	10.102	0.200
14:00:00	4.441	4.782	3.728	4.060	10.727	2.199
14:30:00	3.950	4.334	3.356	3.498	12.060	0.406
15:00:00	4.481	4.833	3.797	3.903	11.027	1.589
15:30:00	3.329	3.574	2.871	2.998	11.133	0.655
16:00:00	3.192	3.477	2.585	2.913	11.004	2.852
16:30:00	4.032	4.379	3.196	3.856	9.918	2.013
17:00:00	7.472	7.856	7.107	3.975	9.429	2.719
17:30:00	4.016	4.112	3.556	4.731	6.124	1.001
18:00:00	1.713	1.926	2.307	2.711	4.459	0.321
18:30:00	3.263	0.567	3.507	0.323	2.334	0.956
19:00:00	3.867	1.748	3.133	1.628	1.936	0.725
19:30:00	12.690	0.161	12.387	0.225	0.000	0

Stellenbosch University https://scholar.sun.ac.za

Table C.36: Root mean square errors December.

Period Start	Root Mean Square Errors December					
	Gaussian	**Weibull**	**Logistic**	**Gamma**	**Exponential**	**Beta**
05:00:00	3.911	0.600	3.062	0.501	0.811	1.593
05:30:00	1.971	1.668	1.352	0.978	6.701	1.989
06:00:00	2.379	2.404	3.936	3.304	9.898	2.698
06:30:00	4.012	3.707	3.426	4.197	10.136	3.712
07:00:00	4.178	4.247	3.370	4.224	8.610	3.024
07:30:00	3.424	3.198	2.731	3.543	8.520	2.093
08:00:00	5.194	3.930	4.615	5.090	9.299	1.021
08:30:00	5.093	3.807	4.522	5.005	9.476	0.950
09:00:00	5.328	4.024	4.738	5.204	9.616	1.022
09:30:00	5.697	4.271	5.080	5.519	10.317	1.011
10:00:00	3.674	4.010	3.083	5.426	10.241	0.684
10:30:00	3.375	3.688	2.787	2.975	9.504	0.821
11:00:00	3.142	3.443	2.622	4.945	9.494	0.453
11:30:00	4.222	4.591	3.627	5.470	9.834	1.458
12:00:00	4.130	4.490	3.533	5.290	9.308	1.509
12:30:00	3.702	4.014	3.028	3.258	11.594	2.768
13:00:00	3.283	3.585	2.599	3.011	10.583	0.761
13:30:00	3.445	3.750	2.804	3.050	10.600	2.456
14:00:00	3.560	3.868	2.931	3.128	12.117	3.476
14:30:00	3.968	4.296	3.195	3.578	11.595	3.164
15:00:00	4.696	5.040	3.929	4.128	11.147	3.326
15:30:00	4.480	4.817	3.642	4.058	10.589	1.334
16:00:00	8.980	9.089	8.957	10.428	8.511	3.110
16:30:00	5.898	5.264	2.747	4.123	8.888	3.983
17:00:00	2.899	2.337	3.417	4.038	8.627	1.138
17:30:00	5.602	5.153	5.677	6.759	7.303	1.968
18:00:00	8.062	7.927	8.143	9.160	10.899	3.675
18:30:00	6.027	5.920	6.293	7.150	9.122	2.196
19:00:00	1.346	1.454	1.421	0.538	6.423	0.802
19:30:00	2.527	2.293	2.399	2.435	3.395	2.979

Stellenbosch University https://scholar.sun.ac.za

C.2 HomeFlex Tariff Structure

C.2.1 Statistical Parameters

Table C.37 summarises the statistical parameters of generated energy for the calendar months January to December.

Table C.37: Statistical parameters of generated energy for HomeFlex during calendar months January to December.

Month	Tariff period	Total generated energy [kWh]	Maximum generated energy [kWh]	Average generated energy [kWh]	Standard deviation of generated energy [kWh]
January	Evening off-peak	928.124	51.833	29.939	9.868
	Morning peak	18649.138	721.622	601.585	98.757
	Afternoon off-peak	76710.928	3057.285	2474.546	490.897
	Evening peak	2150.662	96.683	69.376	23.730
February	Evening off-peak	530.556	22.781	9.474	5.117
	Morning peak	27582.97	631.131	492.553	108.674
	Afternoon off-peak	142312.499	2886.213	2541.295	368.705
	Evening peak	2989.4	91.787	53.382	20.354
March	Evening off-peak	134.48	6.014	2.169	1.608
	Morning peak	23854.352	512.847	384.748	111.673
	Afternoon off-peak	129666.53	2642.864	2091.396	575.381
	Evening peak	773.605	44.913	12.478	10.957
April	Evening off-peak	9.605	1.365	0.160	0.287
	Morning peak	17326.23	419.768	288.771	78.623
	Afternoon off-peak	100491.517	2206.121	1674.859	360.117
	Evening peak	36.775	4.532	0.613	0.961
May	Evening off-peak	0	0	0	0
	Morning peak	10674.125	287.089	172.163	74.071
	Afternoon off-peak	75022.857	1715.226	1210.046	362.590
	Evening peak	0.008	0.008	0.000	0.001
June	Evening off-peak	0	0	0	0
	Morning peak	6804.936	221.033	113.416	52.440
	Afternoon off-peak	59898.826	1409.039	998.314	373.001
	Evening peak	0	0	0	0
July	Evening off-peak	0	0	0	0
	Morning peak	3926.049	202.648	126.647	47.525
	Afternoon off-peak	36385.093	1613.468	1173.713	335.971
	Evening peak	0.009	0.004	0.000	0.001
August	Evening off-peak	0	0	0	0
	Morning peak	5457.108	319.989	176.036	89.844
	Afternoon off-peak	40178.74	1985.81	1296.088	445.171
	Evening peak	3.342	0.504	0.108	0.154
September	Evening off-peak	117.752	15.931	3.925	4.203
	Morning peak	11475.276	646.949	382.509	136.537
	Afternoon off-peak	52710.297	2483.96	1757.010	529.242
	Evening peak	43.549	2.689	1.452	0.672
October	Evening off-peak	700.36	39.801	22.592	9.673
	Morning peak	17764.977	698.719	573.064	125.986
	Afternoon off-peak	68363.709	2663.233	2205.281	480.641
	Evening peak	223.537	14.5	7.211	3.540

Stellenbosch University https://scholar.sun.ac.za

November	Evening off-peak	1524.634	68.347	50.821	15.391
	Morning peak	20263.817	813.77	675.461	170.760
	Afternoon off-peak	71971.848	2895.946	2399.062	572.430
	Evening peak	823.52	52.765	27.451	13.532
December	Evening off-peak	1760.366	92.064	56.786	11.676
	Morning peak	20403.49	811.489	658.177	137.119
	Afternoon off-peak	73325.706	2946.675	2365.345	576.335
	Evening peak	1839.298	78.237	59.332	17.609

C.2.2 Chi-squared Test Results

Table C.38 summarises the results of the Chi-squared test for HomeFlex during the calendar months of January to December. The Chi-squared test results for each tariff period are given with respect to all hypothesised distribution functions.

Table C.38: Chi-squared test values for HomeFlex during calendar months January to December

Month	Tariff Period	Chi-squared test results											
		Normal		Weibull		Logistic		Gamma		Exponential		Beta	
		Value	NB	Value	NB	Value	NB	Value	NB	Value	NB	Value	NB
January	Evening off-peak	3.030	4	3.118	4	3.522	4	2.349	4	33.076	4	3.955	4
	Morning peak	9.707	4	7.466	4	15.782	3	17.397	3	141.553	2	7.078	4
	Afternoon off-peak	10.741	3	7.942	3	1.049	2	14.251	3	141.523	3	2.145	3
	Evening peak	16.205	3	16.960	3	15.639	3	1.559	2	73.803	4	0.867	4
February	Evening off-peak	5.516	5	1.529	5	4.537	4	1.702	5	13.760	5	2.455	5
	Morning peak	0.671	2	17.903	3	0.613	2	0.717	2	270.358	6	7.236	3
	Afternoon off-peak	2.150	2	2.491	2	1.623	2	2.016	2	543.093	6	6.145	3
	Evening peak	3.389	5	2.729	5	6.871	5	6.906	5	42.501	5	0.992	5
March	Evening off-peak	7.909	5	3.131	5	10.808	5	4.079	5	8.623	5	2.644	5
	Morning peak	1.128	2	1.231	2	1.224	2	1.318	2	222.942	6	9.735	3
	Afternoon off-peak	2.450	2	2.363	2	2.715	2	3.641	2	279.311	6	5.526	3
	Evening peak	11.190	3	4.279	4	10.143	3	4.370	4	2.645	4	0.244	5
April	Evening off-peak	9.257	2	3.260	2	8.102	2	1.952	2	0.000	1	0.720	3
	Morning peak	25.772	4	22.572	4	10.154	3	15.832	3	171.250	6	14.710	5
	Afternoon off-peak	8.736	3	7.535	4	9.343	3	14.417	3	194.247	5	6.464	5
	Evening peak	5.691	2	0.535	2	4.756	2	0.235	2	2.27E-05	1	1.094	3
May	Evening off-peak	0	1	0	0	0	0	0	0	0	0	0	0
	Morning peak	12.561	5	15.749	5	15.999	5	33.655	5	70.590	5	8.103	5
	Afternoon off-peak	14.326	3	11.685	4	14.854	3	26.456	3	130.722	5	2.959	5
	Evening peak	22.693	1	0.000	1	21.735	1	0.001	1	0.000	1	0	1
June	Evening off-peak	0	0	0	0	0.000	0	0	0	0	0	0	0
	Morning peak	17.965	5	20.891	5	22.546	5	30.136	5	71.113	5	14.009	5
	Afternoon off-peak	29.242	4	31.308	4	32.083	4	48.507	4	152.786	4	2.017	4
	Evening peak	0	0	0	0	0	0	0	0	0	0	0	0
July	Evening off-peak	0	0	0	0	0	0	0	0	0	0	0	0
	Morning peak	1.489	4	1.976	4	7.860	3	11.372	3	37.160	4	2.251	4
	Afternoon off-peak	3.430	3	2.896	3	0.777	2	0.369	2	89.304	4	0.017	3
	Evening peak	7.049	1	0.006	1	6.356	1	0.006	1	3.34E-11	1	0.038	2

Stellenbosch University https://scholar.sun.ac.za

August	Evening off-peak	0	0	0	0	0	0	0	0	0	0	0	0
	Morning peak	7.827	4	8.012	4	11.126	4	11.604	4	22.139	4	3.306	4
	Afternoon off-peak	2.409	4	1.756	4	4.661	4	3.642	4	28.650	4	1.328	4
	Evening peak	2.741	2	4.404	2	2.105	2	3.279	2	4.338	2	0.174	3
September	Evening off-peak	1.371	2	4.087	2	1.914	2	3.880	2	4.092	2	3.519	4
	Morning peak	2.943	4	2.940	4	4.413	4	6.194	4	33.857	4	1.275	4
	Afternoon off-peak	12.528	4	11.024	4	16.082	4	17.855	4	70.074	4	2.597	4
	Evening peak	2.620	4	2.303	4	4.536	4	3.600	4	15.987	4	1.403	4
October	Evening off-peak	1.960	4	2.064	4	3.341	4	3.401	4	23.435	4	1.535	4
	Morning peak	2.182	2	2.398	2	1.236	2	2.164	2	141.752	4	2.552	3
	Afternoon off-peak	10.105	3	8.288	3	0.993	2	1.759	2	137.193	3	1.080	3
	Evening peak	1.917	4	1.470	4	3.608	4	2.542	4	13.888	4	0.955	4
November	Evening off-peak	4.822	3	4.739	3	4.328	3	8.190	3	78.866	4	0.340	4
	Morning peak	2.549	2	3.066	2	1.596	2	2.206	2	203.385	4	1.073	3
	Afternoon off-peak	2.057	2	2.615	2	1.398	2	1.514	2	158.843	4	0.693	2
	Evening peak	2.768	4	2.490	4	4.807	4	3.938	4	14.761	4	1.280	4
December	Evening off-peak	2.162	3	2.452	3	1.188	3	1.966	3	70.234	4	2.707	3
	Morning peak	1.032	2	1.341	2	0.662	2	0.777	2	177.993	5	0.640	2
	Afternoon off-peak	1.855	2	2.313	2	1.207	2	1.385	2	147.276	5	1.021	2
	Evening peak	14.209	3	13.429	3	14.693	3	18.999	3	100.056	3	3.190	3

Stellenbosch University https://scholar.sun.ac.za

C.2.3 Root Mean Square Errors

Table C.39 summarises the results of the RMSE test for HomeFlex during the calendar months of January to December.

Table C.39: Root mean square errors for HomeFlex during calendar months January to December.

Month	Tariff period	Root Mean Square Error					
		Normal	Weibull	Logistic	Gamma	Exponential	Beta
January	Evening off-peak	2.547	2.616	2.450	1.844	5.864	3.090
	Morning peak	3.520	2.902	7.453	7.951	14.898	2.803
	Afternoon off-peak	6.405	5.566	3.183	7.266	11.207	2.416
	Evening peak	7.974	8.185	7.713	3.761	7.805	0.989
February	Evening off-peak	3.353	1.599	2.620	1.442	6.204	2.477
	Morning peak	3.734	11.547	2.829	3.737	12.591	6.918
	Afternoon off-peak	6.636	7.130	5.570	6.245	14.049	4.705
	Evening peak	2.569	2.450	3.514	3.742	6.539	1.401
March	Evening off-peak	3.718	2.099	4.196	2.485	4.543	2.605
	Morning peak	4.721	5.196	3.714	4.452	13.477	7.528
	Afternoon off-peak	5.962	6.504	4.952	5.360	14.628	5.043
	Evening peak	9.094	2.331	8.175	2.238	1.356	0.659
April	Evening off-peak	10.382	2.323	9.517	1.990	0.012	1.183
	Morning peak	7.751	7.742	8.510	10.523	11.208	5.710
	Afternoon off-peak	8.296	4.379	8.314	10.296	12.410	4.575
	Evening peak	10.148	0.908	9.183	0.631	0.037	1.933
May	Evening off-peak	0	0	0	0	0	0
	Morning peak	5.079	5.887	5.324	7.369	9.735	4.783
	Afternoon off-peak	10.521	6.489	10.343	13.423	11.808	1.846
	Evening peak	27.842	0.080	27.417	0.231	0	0
June	Evening off-peak	0	0	0	0	0	0
	Morning peak	6.373	6.972	7.004	8.225	9.154	5.743
	Afternoon off-peak	10.617	10.853	11.116	12.389	14.730	2.739
	Evening peak	0	0	0	0	0	0
July	Evening off-peak	0	0	0	0	0	0
	Morning peak	1.758	2.114	5.453	6.321	6.255	2.346
	Afternoon off-peak	3.700	3.488	1.571	1.708	8.031	0.285
	Evening peak	11.673	0.436	11.214	0.424	0.000	0.597
August	Evening off-peak	0	0	0	0	0	0
	Morning peak	3.468	3.888	3.872	4.679	4.596	2.462
	Afternoon off-peak	1.713	1.513	2.307	2.224	4.521	1.331
	Evening peak	4.911	2.879	4.090	2.600	3.154	0.666
September	Evening off-peak	2.737	3.263	2.430	3.208	3.327	1.966
	Morning peak	2.352	2.466	2.535	3.545	5.520	1.721
	Afternoon off-peak	4.530	4.317	5.001	5.150	6.190	2.102
	Evening peak	2.092	1.738	2.553	1.912	4.259	1.713
October	Evening off-peak	1.848	1.689	2.354	1.973	5.396	1.640
	Morning peak	4.508	4.693	3.604	4.475	8.911	2.774
	Afternoon off-peak	6.078	5.601	3.129	4.080	11.167	1.153
	Evening peak	1.817	1.317	2.228	1.605	4.920	1.476
November	Evening off-peak	4.372	4.330	4.139	5.512	7.138	0.524
	Morning peak	4.886	5.258	4.024	4.596	10.957	1.454
	Afternoon off-peak	4.558	4.976	3.907	4.082	10.617	1.491
	Evening peak	2.152	1.763	2.628	2.014	4.442	1.636

Stellenbosch University https://scholar.sun.ac.za

December	Evening off-peak	2.568	2.774	1.783	2.431	8.050	2.940
	Morning peak	3.331	3.593	2.797	3.066	8.891	1.400
	Afternoon off-peak	4.230	4.581	3.565	3.851	8.451	1.872
	Evening peak	7.056	6.888	7.047	8.144	10.293	2.655