# Modelling Optimum Water Release for Agricultural Lands Using Goal Programming Approach

A Thesis Submitted in Partial Fulfillment of the Requirements for the

Master of Science in Operational Research

By

S.A.S Rasangika

199524N

2022

# Modelling Optimum Water Release for Agricultural Lands Using Goal Programming Approach

Master of Science in Operational Research

By

S.A.S Rasangika

199524N

External Supervisor:

Dr. D.M Samarathunga

Internal Supervisor:

Dr. P.M. Edirisinghe

Department of Mathematics

University of Moratuwa

Moratuwa.

# Declaration

I declare that this dissertation has not been previously accepted in substance for any degree in the University or any other Institution. I state that this dissertation is the result of my own independent work, except where otherwise indicated or acknowledged under the supervision of Dr. D.M. Samarathunga, ( Senior Lecturer, Faculty of Science, University of Ruhuna) and Dr. Pasan Edirisinghe, ( Senior Lecturer, Faculty of Engineering, University of Moratuwa).

...............................
S.A.S.Rasangika
199524N
Department of Mathematics
University of Moratuwa
Moratuwa
22 / 03/ 2023

External Supervisor :

...............................
Dr. D.M Samarathunga
Senior Lecturer
Department of Mathematics
University of Ruhuna
Matara
22 / 03 / 2023

Internal Supervisor :

...............................
Dr. P.M. Edirisinghe
Senior Lecturer
Department of Mathematics
University of Moratuwa
Moratuwa
22 / 03 / 2023

# Acknowledgment

First of all I would like to express deep gratitude to my guide Dr. D.M. Samarathunga, Senior lecturer in department of Mathematics for his guidance, encouragement, and gracious support throughout the course of my work, for his expertise in the field that motivated me to work in this area and for his faith in me at every stage of this research.

I would like to thank Dr. Pasan Edirisinghe, Senior Lecture and the coordinator of the Master degree programmer for his timely guidance and coordination in the event of being success of my project

Also, I am grateful to the staff members of Technical services division at water management unit in Walawa who help me to supply the data for the research.

Finally I would like to thank for my parents and friends who help me to success this research and I wish to grant my heartful gratitude for anyone whom I could not mention in particular.

# Abstract

The agricultural sector is one of the main sectors which benefits from several sources of water supply. This sector faces several challenges in securing water supply for its respective crops. Since the demand for water is increasing due to several factors, including population growth and an increased level of water dependent activities, and changes to crop selection influenced by various factors, including ever changing demand, the need for optimal use of water is identified as a priority for the rural agriculture sector in Sri Lanka. Past inflow data was used to obtain inflow values from the Udawalawa reservoir by using a forecasting model and the seasonal ARIMA model which was a statistical forecasting model used to forecast the inflow using RStudio software. In this respect water reservoir management is a crucial factor for allocating water for agricultural crops for study, a Goal programming model was formulated to determine the optimum water release for agricultural lands in Udawalawa region. Six main agricultural land divisions of the Udawalawa irrigation scheme were selected, and two of them received water from the right bank canal while the other four received water from the left bank canal of the Udawalawa reservoir. A Goal programming model with a 246 number of variables, 72 equality constraints, and 138 inequality constraints is solved using the MATLAB programming language and results show that water allocation from each reservoir can be used to fulfill the water demand throughout the year by using the water flow and allocation for each crop according to the priorities. In this case, it was assumed that 60% of the farming land used to cultivate paddy while the remaining 40% of the farming land used to cultivate other crops. These findings can be used by the stakeholders when making decisions on water allocation, not just based on the demand but also meeting a balance of crop selection.

**Key words: Water Release Optimization, Goal Programming, Forecasting**

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

The earth is made lively, green, happy, and growing because of water. In fact, it is one of the most precious and important resources that sustain life. We as humans require water every day of our day-to-day activities as well as it is important for agriculture and for most industries. Natural sources of water include rainwater, ponds, rivers, streams, lakes, oceans and springs, whereas man-made sources encompass dams, wells, tube wells, tanks, reservoirs, canals, and similar infrastructure.

The agricultural sector is one of the main sectors which benefits from man-made water sources. Even though there are many of them available, people still have to face a lot of challenges. Water scarcity has a significant impact on water security because many human activities and processes rely on water. Food security is one of the most challenging and it has been involved in several socioeconomic downfalls. Increasing the population rapidly often results in a drive-up demand for food, which results in additional use of livable land. Another challenge that farmers have to face is climate change, which is causing more severe droughts and floods. Changes in rainfall patterns have impacted on traditional farmers as they are based on regulated systems that have been in practice for a number of years. As an example, in Sri Lanka, people used to do paddy cultivation in two different seasons per year, called the Maha and Yala seasons. With climatic changes, farmers are struggling to get the maximum harvest out of their crops since they have to rely on both natural and man-made water sources.

The reservoir is one of the man-made water sources that is widely used around the world for many purposes. A man-made body of water used for storing water is called a reservoir.

Typically, dams are built across rivers to form reservoirs. It is also possible to create a reservoir by damming the outlet of a natural lake to regulate water levels. The amount of water that flows out of the reservoir is controlled by the dam. Reservoirs come in various sizes and complexities, ranging from small single-purpose impoundments to large and intricate multi-purpose impoundments. Reservoirs are a significant source of water that is utilized worldwide. There are single-purpose reservoirs that are designed for a specific function, such as irrigation, hydroelectric power generation, navigation, flood control, water supply, recreation, or low-flow regulation. However, the trend has shifted towards constructing multi-purpose reservoirs that can serve at least two primary functions.

Today, most developed countries use statistical analysis data to regulate the control of water release from reservoirs to get the maximum results and to maintain good environmental quality. As multi-purpose reservoirs have been introduced, data management and statistical analysis are playing a dominant role at the highest professional level. In Sri Lanka, it is hard to find such automated systems based on statistics to regulate the water release even though there is complexity involved because of rice being the major crop and some other crops being cultivated simultaneously. Because of not having a proper control system in water releasing, the chance of having a higher probability of either getting more than required or getting less than the required amount of water for a particular crop has led to not obtaining an optimum harvest. To guarantee food security and sustainable water management in agriculture, it is critical to increase crop production while using less water. This is necessary to improve water usage efficiency in the agricultural sector without adversely affecting downstream water quality and quantity. This study is based on a set of reservoirs located in the Sabaragamuwa province in Sri Lanka, and the Udawalawe reservoir is the major one. Furthermore, the study discusses how the water is stored in those reservoirs and how that water is used optimally for the agricultural activities across the territory. Rice production or the paddy production is one of the major productions in Sri Lanka and majority of the land area and water quantity is consumed by the paddy cultivations. Apart from paddy there are several types of other crops are cultivated in the considered area such as sugarcane, Banana, etc.

## 1.1 Background of the study

Water is a crucial but limited resource for both human communities and ecological systems. Therefore, it is necessary to make the best use of water resources. This study employs a mathematical programming model to achieve the optimal distribution of water resources between the agricultural and environmental sectors.

## 1.2 Problem Statement

Although agriculture is the primary user of renewable water resources in the country, the yield from these resources is low. Therefore, it is imperative to conserve water consumption. Sri Lanka's existing data and research indicate that water is the scarcest factor in agricultural production, and the development of the agricultural sector depends on the quantity and quality of water resources and how they are managed and utilized. Thus, proper water management and efficient use of water resources are crucial for the sustainable development of the country and should be implemented with a consistent program. Water scarcity, population growth, economic development, and increasing demand across various sectors have led to a critical situation for water resources. Mathematical models are employed as tools for achieving the objectives of reservoir systems planners and managers. This study focuses on the optimization of water use in water reservoirs to meet agricultural needs.

## 1.3 Research Question

The following research question was considered in this study:

- what is the optimum water release plan for each month while ensuring the irrigation demands for the cultivation of several crops?

## 1.4 Research Objectives

The objectives of this research are study are:

- develop a mathematical programming model for determine the optimum amount of release water according to the irrigation demand.
- determine the percentage of lands for crop types.
- study and forecast monthly inflow to the Udawalawa reservior.

## 1.5 Significance of the Research

The production of agricultural goods depends on various resources, including land, water, labor hours, machine hours, and capital. In addition, factors such as irrigation methods, soil characteristics, crop patterns, cropping intensity, fertilizer use, climatic conditions, and socioeconomic factors must be considered. While utilizing these resources, it is essential to maximize gross returns, minimize total costs, and increase total production. However, these goals are inherently contradictory, and it is impossible to maximize or minimize all of them simultaneously. Achieving certain goals may come at the expense of others, and therefore, a compromise among the goals is necessary to arrive at a satisfactory solution in the decision-making process. Thus, planning for agricultural production is a critical factor.

In the field of agricultural planning, mathematical programming models have been utilized extensively over the past few decades. The most commonly employed optimization technique is linear programming (LP), which is often used to optimize a single objective, such as maximizing crop production by allocating cultivated land or minimizing costs. However, LP models assume that the parameters are well-defined and precise, which is not always the case in real-world situations where input data may be ambiguous, approximated, based on human observations, or derived from prior knowledge. Moreover, agricultural systems are inherently uncertain and ambiguous, making it impractical to find an optimal solution to production planning problems using LP models under this assumption. Therefore, the solution technique employed to address such issues must be capable of handling imprecise data.

## 1.6 Limitations and Assumptions

The demand for water is considered to be fixed and known for each crop and region. The model assumes that the cropping pattern does not change during the planning horizon. The model considers the maximization of the total cropped area and the minimization of the deviation from the target irrigation water demand for each crop and region. The model also considers the carryover storage from the previous season as initial storage. The irrigation system efficiency is assumed to be constant for each crop and region. The model assumes that there is no evaporation loss from the reservoirs. Finally, the model assumes that the decision variables are continuous and non-negative.

## 1.7 Outline of the thesis

This thesis comprises from 5 chapters. In chapter 1, we have already discussed about the background of study, problem statement, Research question, Research objectives, significance of this study and limitations and assumptions.

Chapter 2 presents a detailed literature review related to optimum water release for agricultural lands based on different modeling approaches and finally it will be discussed about the research gap in previous studies.

Chapter 3 demonstrates the research design and methods related to this study. Firstly, select cultivating areas, and then select a general model. After that, collection of data using data. After that, develop the model, and then analyze the collected data as well. At the final stage, determining the area of cultivated land.

Chapter 4 describes the analysis, results, and findings related to time series forecasting and optimization techniques.
Chapter 5 summarizes the discussion and conclusion of the thesis and highlights the areas for future research.

# Chapter 2

# Literature Review

The agriculture sector in Sri Lanka is a critical part of the local economy and significantly contributes to the overall fresh food supply, including several staple food products for the growing population and increasing demand. A large rural agriculture industry with numerous rural growers characterizes it. Among several rural agriculture sectors in Sri Lanka, the north-western area of Sri Lanka is one of the largest arrears, with several growers producing staple foods such as paddy for rice, sugar cane, and bananas [1]. Several studies have reported on rural farming in Sri Lanka from several perspectives, including food security challenges of agri-food supply from rural growers to city regions, mitigating the environmental impact of fresh food supply. Some studies have considered the agri-food value chain from food production to consumption from various perspectives, including value chain approaches for meeting dynamic economic objectives, environmental conditions, and social factors[2], the need for a high degree of coordination between several stakeholders [6], and the increasing level of food security issues in Sri Lanka due to several factors [4]. While these studies have considered several aspects of the agri-food sector in Sri Lanka from food supply to the consumer, food production from the perspective of irrigation of several crops in the rural agri-sector is considered one of the key areas that need urgent attention, particularly subject to changing climate conditions and increasing demand [3].

Irrigation of rural agri-food production is important in water management, especially when water supply is limited by changing climate conditions and varying demand for water across multiple crops during different seasons. Thus, managing irrigation in the rural agri-sector is crucial for the survival of growers and meeting increasing demand by consumers. Some research studies have reported on different methods of managing irrigation water in rural agriculture sectors in emerging economies [6]. Several studies have considered the optimization of reservoir operation, from the perspective of water allocation to different users, including the irrigation of requirements by growers [7]. Most of these studies have adopted optimization methods such as linear programming (LP) and integer programming(IP) to model water allocation and release subject to several technical and environmental constraints. Recently, B.C. Kumar Raju & Chandre Gowla C, Karthika B.S (2020)[9] proposed a mathematical model to derive the optimal water releases monthly for a single-purpose reservoir. This proposed method was applied to the Hemavathy reservoir, Hassan District, Karnataka, India. Findings based on the results from the mathematical mode conclude that the downstream amount of water was conserved from reduced spills.

Ajudiya et al. (2018) [12] developed an LP model to determine the maximum revenue from water supply release for the Aji-1 reservoir, based on equivalent deterministic inflows projected from the monthly probability distribution of observed stochastic inflows. The models were solved for nine different dependable inflow levels, with infeasible solutions for three levels. The study analyzed results for maximum revenue generation, optimum release of water supply, loss due to evaporation, and initial storage across six different dependability levels of inflow.

Multi-Objective Optimization is considered superior to Single Objective optimization. M. Janda Reddy and D. Nagesh utilized Multi-Objective Evolutionary (MOEA) [17] algorithms to derive a set of optimal operation policies for a multipurpose reservoir system. They utilized Evolutionary algorithms and Multi-Objective Genetic Algorithms (MOGA) to address the challenges associated with classical methods for Multi-Objective Optimization problems (MOOP). This approach was applied to a realistic reservoir system in Bhadra India, which is utilized for various purposes including irrigation, hydropower generation, and downstream water quality requirements. The proposal's results included many alternative policies, allowing the reservoir operator to choose the most suitable one.

A study by Ajudiya et al. (2018)[12] proposed a linear programming model for maximizing revenue from water release in the Aji-1 reservoir. The model took into account equivalent deterministic inflows, projected from the monthly probability distribution of observed stochastic inflows. The authors prepared and solved the model for nine different dependable inflow levels but found it infeasible for three levels. The results were analyzed for various factors such as maximum revenue generation, optimal water release, loss due to evaporation, and initial storage, for six different dependability levels of inflow. The study demonstrated the effectiveness of linear programming in optimizing water release from reservoirs for revenue generation.

The optimal utilization of available water resources through modelling techniques is a highly researched topic in the literature. Many books have been published on this topic, including "Water Resources Systems Modelling Technique and Analysis" by S. Vedula and P. P. Mujumdar (2005) [11]. This book provides several case studies of model development, applications, and techniques, including recently developed tools such as Artificial Neural Networks (ANN) and Fuzzy sets. The book illustrates how linear programming applications can be used to overcome several objectives, such as optimizing crop yield, multi-objective analysis of multi-reservoir systems, short-term reservoir operation, and reservoir operation for hydropower optimization. Additionally, the book covers the basics of recent modelling tools such as ANNs and Fuzzy sets. The applications illustrated in the book include inflow forecasting using ANN, Fuzzy rule-based reservoir operation, and Fuzzy LP for river water quality management and reservoir operation.

The book "Water Resource Systems Planning and Analysis" by D. P. Loucks, J. R. Stedinger, and D. A. Haith[13] provides a comprehensive guide to developing and applying various modeling approaches for water resource planning and management. The book aims to enable readers to not only model and analyze but also gain practical experience in using modeling approaches. It contains 13 chapters, and the fourth chapter discusses the use of optimization methods for defining infrastructure design and operating policies. The chapter illustrates the advantages and limitations of different optimization and simulation approaches using simple water allocation, reservoir operation, and water quality management problems. The fifth chapter focuses on optimization problems characterized by qualitative objectives and/or constraints and introduces recently developed statistical modeling methods, such as artificial neural networks, genetic algorithms, and genetic programming. The book acknowledges that precise quantification of system performance criteria, parameters, and decision variables is not always possible, especially when the values of variables are uncertain or fuzzy. Probability distributions may be used to quantify such uncertainties. Overall, this book is a valuable resource for researchers, practitioners, and students interested in water resource planning and management.

Apart from the watser storage/system infrastructure and national level policy frameworks discussed above, other area of consideration in water management for irrigation is the water quality. The water quality is critical not only for quality fresh produce, maintaining required environmental and safe farming conditions, but also the effect of irrigation on household water quality. Several studies have considered irrigation water quality and provided useful guideleines and stanadards in different contexts and demographical backgrounds. Among several factors considered by scholars, the impact of irrigation water quality on human health and indirect wastewater reuse are recognized as critical factors in recent times (Srinivasan and Reddy, 2009, Jeong et al[18]. The need for consideration of water quality for the purpose of irrigation and the effect of irrigation on household water quality is exacerbated in recent times , due to the impact of climate change and its impact on agriculture condirions. Furthermore, scarecity of freshwater resources, particularly in rural areas is driving to use low quality water for agriculture.

Addressing the issue of water quality in irrigation, several studie have reported on managing water quality using optimization methods. K. Sasikumar and P.P. Majumdar[9] have presented a proposal to manage the water quality in a river by using fuzzy optimization techniques. In fact, they have used fuzzy multiple optimizations to manage the water quality of the river system by using a fuzzy waste load allocation model (FWLAM). This model has a significant ability to integrate the competing objectives and aspirations of both the pollution control agency and dischargers. The goals of these parties are presented through the use of fuzzy sets, with the membership functions indicating their respective levels of satisfaction in achieving their goals.

According to, some studies have suggested the use of statistical methods based on past data to manage irrigation in rural agriculture. In hydrology, statistical analysis involves estimating the probability or frequency of future hydrologic events based on historical hydrologic records. Naresh Vijayakumar and Soorya Vennila conducted a study titled "A Comparative Analysis of Forecasting Reservoir Inflow using the ARMA Model Holt-Winters Exponential Smoothing Technique" to apply time series analysis techniques, such as the Autoregressive Moving Average (ARMA) model and Holt-Winters exponential smoothing, for forecasting reservoir inflow. This would enable better management of reservoir inflow. The study focused on the Krishnagiri Reservoir in Tamilnadu, India[13]. Although the ARMA model is considered the best fit for time series analysis in hydrology, it has a limitation in reproducing the trend, seasonality, and periodicity component of a time series. The study found that the exponential smoothing technique provides a better match between observed and computed values compared to the ARMA model.

It is evident from the above discussion that irrigation in rural agriculture is a complex process and requires careful consideration of various factors, particularly when operating under changing conditions. In the case of Sri Lankan rural agriculture of selected region, the problem of irrigation management with a limited water supply/stirage using a system of reservoirs is challenging one. Therefore, this research is aimed at addressing the broader research question: what is the optimum water release plan for each month while ensuring the irrigation demands for cultivation of several crops.? This research adopted a quantitative approach with a mathematical modelling and solving the problem using the past data.

# Chapter 3

# Methodology

The research methodology adopted for this research is a quantitative approach and involves a mathematical model for optimum release of water for a set of reservoirs, subject to several constraints. The mathematical formulation is considered LP model for optimal release of water from one storage with limited capacity to a set of reservoirs monthly, meeting the required demands of six catchment areas for a combination of crops, including paddy. The key constraints of this study can be divided into two categories: hard constraints and soft constraints. One of the major hard constraints is the continuity equation, which governs the mass balance of the reservoirs, while the limitation constraint indicates the maximum and minimum storage limits of reservoirs. The main soft constraint is the demand, which illustrates the amount of water volume required for each and every catchment, while the land allocation constraint is used as the deciding factor which allocates a proportion of land for the different types of crops being cultivated in each catchment. The LP model was solved using the Optimization Toolbox of MATLAB and the solutions of key decision variables are analyzed and discussed for decision making and recommendations.

Since the problem is very specific to a selected area of rural agriculture, study area is outlined next.

## 3.1 Linear Programming (LP)

When it comes to planning water resources, Linear programming is one of the most dominant optimization techniques, which is commonly used worldwide. Most of the analysis related to linear programming is done using various software available on the market, and it is done on both small and large scales. According to the system analysts, this is an extremely useful tool to solve large-scale systems. One of the main advantages of the linear programming technique is that it can be used to solve larger optimization problems that contain many variables and constraints.

**General Formulation**

Linear programming (LP) is a mathematical method used to solve optimization problems where the objective function and constraints can be represented as linear functions of decision variables. A linear programming model has a general structure, which includes the objective function and constraints expressed as linear equations or inequalities.
**Objective Function**

$$Maximize \quad or \quad Minimize = \sum_{j=1}^{n} (P_j x_j) \tag{3.1}$$

**Subject to**

$$\sum_{j=1} a_{ij} x_j \leq \quad or \quad \geq b_i \quad i = 1, 2, 3, ..., m \tag{3.2}$$

**Non-negativity**

$$x_j \geq 0 \quad for \quad i = 1, 2, 3, ..., n \tag{3.3}$$

Linear programming is a mathematical approach to solving optimization problems that involve linear objective functions and constraints. The structure of a linear programming model involves decision variables, an objective function, and constraints, which can be expressed as equalities (=) or inequalities (≥ or ≤). A range of linear programming software is available to find optimal solutions for unknown decision variables xj. Commercial linear programming programs typically enforce variable non-negativity within the solution algorithms, so there is no need to specify these conditions in an application. While users of linear programming software may not need to understand all the mathematical details of the solution procedure, they should understand how the procedure works and what the output means. Graphical methods and simplex methods are the two solution methods used for linear programming.

## 3.2 Goal Programming (GP)

The Goal programming technique was first utilized by Charnes, Cooper and Ferguson in 1955. However, the term ”Goal programming” was coined by Charnes and Cooper in their later work. Several articles before 1955 focused on the concept of Goal programming, which were compiled in a bibliography by Schniederjans and Jones. Additionally, Tamiz compiled an annotated bibliography covering the period from 1990-2000.

The problems associated with Goal programming can be classified based on the type of mathematical programming model being used, such as linear programming, integer programming, or non-linear programming. Goal programming is an extension of linear programming that enables modelling of multiple objectives simultaneously. The techniques employed in goal programming expand the modelling capabilities, allowing for the handling of multiple objectives.

In decision-making, it is common for multiple viewpoints and objectives to be involved, leading to competing interests and confrontations. Goal Programming (GP) belongs to the field of multi-objective optimization, which is a subset of multi-criteria decision analysis, also known as multiple-criteria decision making. GP is an extension of linear programming that can handle multiple objective measures, often with conflicting goals. However, it cannot provide a complete solution as it is difficult to conduct an entirely objective analysis when there are competing objectives. The approach of GP is to establish specific numeric goals for each objective and then seek a solution that satisfies all the given constraints while minimizing the sum of deviations from the stated goals. The goal programming approach finds a compromise solution based on the relative importance of each objective.

**General Formulation**

To implement goal programming, specific numerical goals are established for each objective, followed by the formulation of objective functions for each one. The solution then involves minimizing the total deviation of these objective functions from their corresponding goals. There are three possible types goals.

(1) A lower, one-sided goal sets a lower limit that we do not want to fall under (but exceeding the limit is ne).

(2) An upper, one-sided goal sets an upper limit that we do not want to exceed (but falling under the limit is ne).

(3) A two-sided goal sets a specific target that we do not want to miss on either side.

### Goals as Constraints

The central construct in goal programming is the deviation variable. Let

- $d_i^+$ = Positive deviation or the amount by which the $i^{th}$ goals is exceed.
- $d_i^-$ = Negative deviation or the amount by which the $i^{th}$ goals is underachieved.

### General form of the model

### Decision variable

Let $x_1, x_2, ..., x_n$ are the decision variables.

### Objective Function

$$min = \sum_{i=1}^{k}(d_i^+ + d_i^-) \tag{3.4}$$

**subject to**

$$f_i(x) - d_i^+ + d_i^- = T_i \tag{3.5}$$

### Non-Negativity

$$AX \leq b \tag{3.6}$$

and

$X \geq 0, \quad d^+ \geq 0; \quad d_i^- \geq 0,$
for $i = 1, 2, ..., k$

## 3.3 Study Area

In this research we consider Udawalawa irrigation scheme is considered in the research. It is shown in Figure No 3.1. Supplying water to agricultural activities is the major usage of the Udawalawa Reservoir. The reservoir consists of two major canals (Left bank canal and Right bank canal). In this study, four major tanks, which are located on the left and right banks, have been considered. By using the canal on the right bank, a land area of 12127.53*ha* and a land area of 10532.21*ha* are cultivated. Paddy (rice), sugar cane, and banana are the crops that can be seen in this area, and paddy (rice) is the major crop.

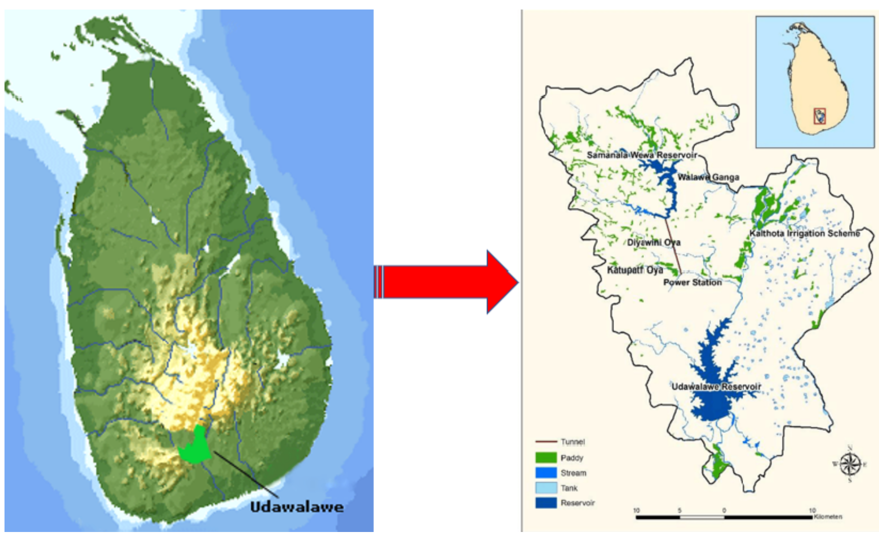


Figure 3.1: Overview of the study area

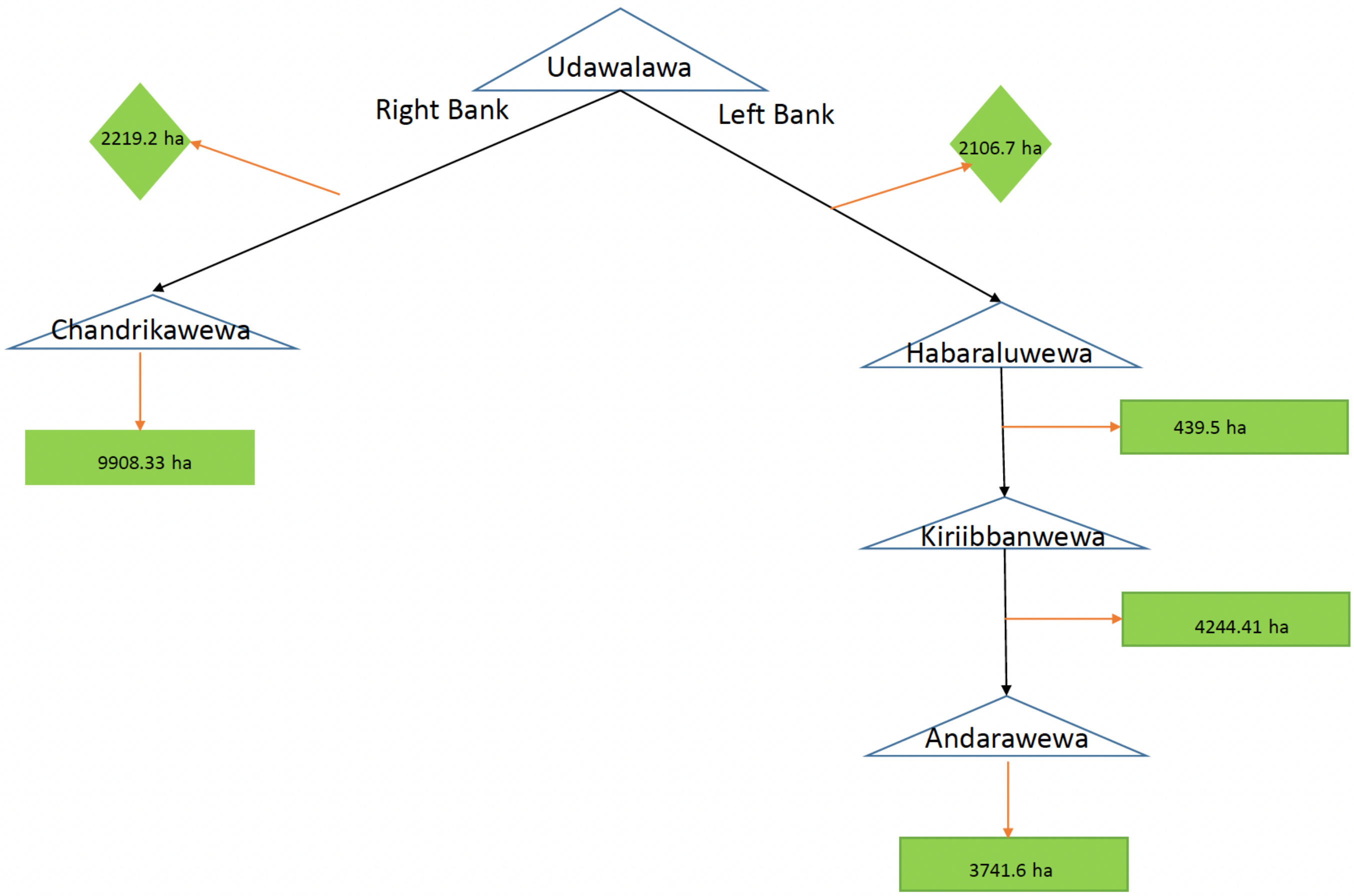


Figure 3.2: Land allocation of the study area

As shown in Figure No 3.2, the proposed irrigation scheme of Udawalawa reservoir includes five major storage facilities/tanks and provides irrigation for six areas, including two areas directly supplied by the main Udawalawa reservoir. Four sub-areas in the downstream are supplied by the main storage into their respective reservoirs and consist of demands for several crops, including paddy cultivation.

There is a total land area of 2219.2 ha, which is in between Udawalawa reservoir and Chandrikawewa reservoir on the right bank, and water is fed to it directly by the main reservoir. Chandrikawewa provides water to cultivate 9908.33 ha of land. On the other hand, the main reservoir, Udawalawe, directly feeds 2106 ha of land on the left bank while passing water to the Habaraluwewa, which feeds 439.5 ha. The next reservoir lined on the left bank is Kiribbanwewa, which provides water to cultivate 4244.1 ha of land. Finally, Andarawewa is the one which is lined farthest away on the left bank of Udawalawa reservoir, and it provides water to a land of 3741.6 ha to cultivate.

## 3.4 Data Collection

The technical services division of the water management unit in Walawa was used to gather information for the study. Inflow, storage of the reservoirs, and demands were the main three collected data. In addition, the water requirement for each and every single crop was collected.

### 3.4.1 Inflow

The amount of water entering the reservoir is known as the inflow. Inflow data from 1996 to 2020 has been used to do the study. Data was received in the form of daily inflow to the reservoir and then converted into monthly data.These Inflow measurements were made in Million Cubic Meters (MCM).

### 3.4.2 Storage of the reservoirs

The capacity of the reservoir is its storage. To do the study, we needed to know the maximum and minimum capacities of the reservoirs. In addition, initial storage of the reservoirs for each month should be considered.

### 3.4.3 Demand

As described in the introduction, water is required for different purposes, such as domestic use, agricultural activities, and hydroelectric power generation. This study is basically focused on agriculture. While paddy (rice) is being considered as the main crop, some other crops such as banana, sugarcane, etc., benefit from the reservoirs which are located around the Udawalawa region. All the required information was collected from those agricultural areas.

## 3.5 Problem formulation

We formulate the research problem under two models.

- Model 1 : Linear programming model
- Model 2 : Goal programming model

### 3.5.1 Model 1 : Linear programming model

As the first model using a linear programming model to calculate the release considering only the land used for paddy cultivation:

**Indices**

t = month $t = 1, 2, 3, ..., 12$

i= area $i = 1, 2, 3, ..., 6$

j= reservoir $j = 1, 2, 3, ..., 5$

**Inputs**

$I_{j,t}$ = Inflow to the reservoir or tank $j$ during the month $t$.

**Variables**

$RL_t$ = The amount of water should be released for the left bank from Udawalawa reservoir during the month $t$.

$RR_t$ = The amount of water should be released for the right bank from Udawalawa reservoir during the month $t$.

$RR_{j,t}$ = The amount of water should be released from $j^{th}$ reservoir during the month $t$.

$RP_{i,t}$ = The amount of water should be released to the $i^{th}$ region during the month $t$.

$S_{j,t}$ = $j^{th}$ reservoir storage at the beginning of month $t$.

$A_i$ = Proportion of land allocated for paddy.

$B_i$ =Proportion of land allocated to crop type $c(c = 1, 2, ..., l)$

$P_i$= Cultivated land ratio to the available land $(0 \leq Pi \leq 1)$

**Objective function**

In this model we consider the objective as maximizing the cultivated area land ratio.

$$Maximize \quad Z = \sum_{i=1}^{6} P_i \tag{3.7}$$

**Constraints**

The constraints can be divide into two parts.

- Hard constraints.
- Soft constraints.

- **Hard constraints**.

**01. Continuity Equations**

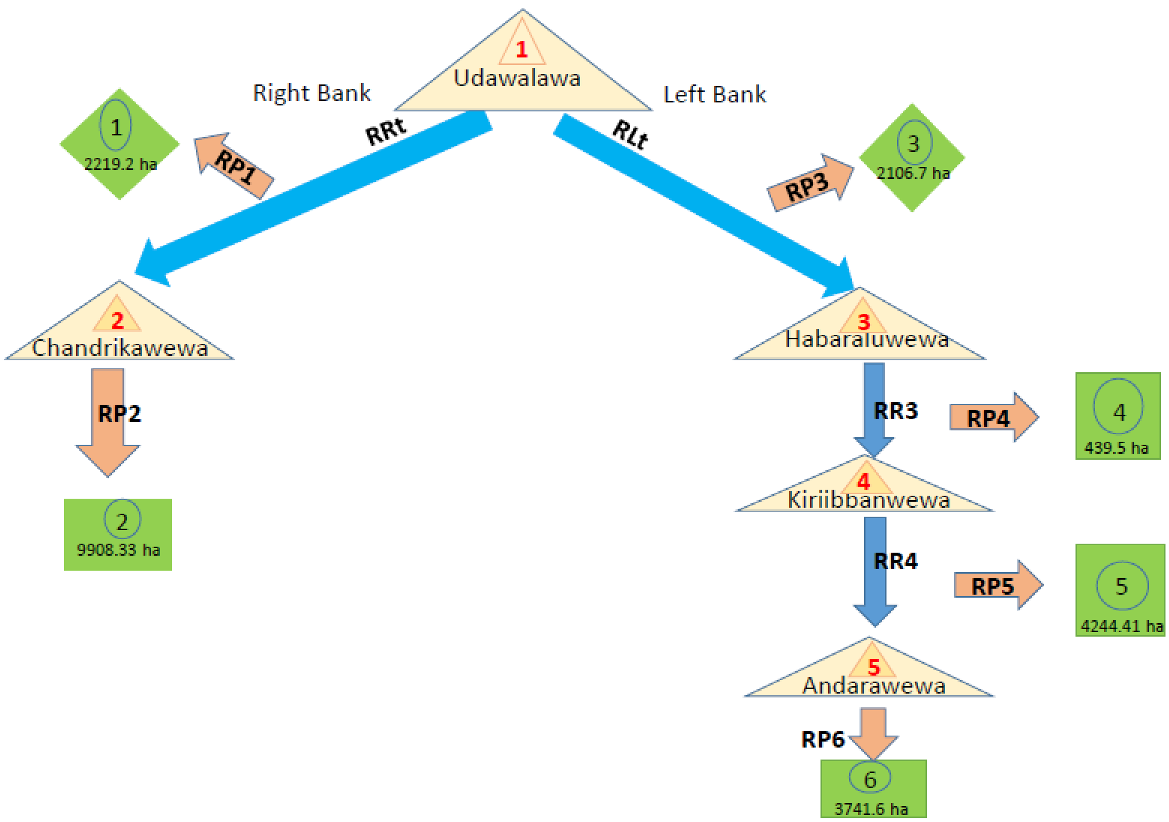


Figure 3.3: Plan for Water releasing

i . For Udawalawa Reservoir (j=1):

$$S_{j,t+1} = S_{j,t} + I_t - \sum_{j=1}^{n}(RR_t + LR_t) \qquad ;\forall \text{ t=1,2,3,...,12} \quad \text{and} \quad \text{j=1} \qquad (3.8)$$

ii . For Chandrikawewa Reservoir (j=2):

$$S_{j,t+1} = S_{j,t}+RR_t-RP_{i,t}-RP_{(i+1),t} \qquad ;\forall \quad \text{t=1,2,3,...,12,} \quad \text{i=1} \quad \text{and} \quad \text{j=2} \qquad (3.9)$$

iii . For Habaraluwewa Reservoir (j=3):

$$S_{j,t+1} = S_{j,t}+RL_t-RP_{i,t}-RR_{j,t} \quad ;\forall \quad t=1,2,3,...,12, \quad i=3 \quad \text{and} \quad j=3 \tag{3.10}$$

iv . For Kiriibbanwewa Reservoir (j=4):

$$S_{j,t+1} = S_{j,t}+RR_{j,t}-RP_{i,t}-RR_{j,t} \quad ;\forall \quad t=1,2,3,...,12, \quad i=4 \quad \text{and} \quad j=4 \tag{3.11}$$

v . For Andarawewa Reservoir (j=5):

$$S_{j,t+1} = S_{j,t}+RR_{j,t}-RP_{i,t}-RP_{(i+1),t} \quad ;\forall \quad t=1,2,3,...,12, \quad i=5 \text{ and} \quad j=5 \tag{3.12}$$

**02. Storage limits**

$$S_{j,min} \leq S_{j,t} \leq S_{j,max} \quad ;\forall \quad t=1,2,3,...,12, \quad j=1,2,...,5 \tag{3.13}$$

where, $S_{j,min}$ and $S_{j,max}$ are the minimum and maximum active storage of the reservoir.

- **Soft constraints**.

**01. Demands**

$$RP_{i,t} \geq \alpha_i H_i A_i + \sum_{c=1}^{l} \beta_c H_i B_{ic} \tag{3.14}$$

Where, $H_i$= Total area of the $i^{th}$ land

$\alpha_i$= The amount of water needed for paddy

$\beta_c$= The amount of water needed for crop $c$

$\sum_{i=1}^{6} P_i = \sum_{i=1}^{6} \left(A_i + B_i\right)$

**02. Limits for $A_i$**

In here land allocations for paddy cultivation at least 60% of the total land area.

$$\sum_{i=1}^{6} A_i \geq 0.6 \tag{3.15}$$

**03. Limits for $P_i$**

The ratio of cultivated land is between zero and one.

$$0 \leq \sum_{i=1}^{6} P_i \leq 1 \tag{3.16}$$

### 3.5.2 Model 2 : Goal programming model

Using Goal Programming model to calculate the release and proportion of land allocated for paddy and other crops as a second model:

**Indices**

t = month $t = 1, 2, 3, ..., 12$

i= area $i = 1, 2, 3, ..., 6$

j= reservoir $j = 1, 2, 3, ..., 5$

**Inputs**

$I_{j,t}$ = Inflow to the reservoir or tank $j$ during the month $t$.

**Variables**

$RL_t$ = The amount of water should be released for the left bank from Udawalawa reservoir during the month $t$.

$RR_t$ = The amount of water should be released for the right bank from Udawalawa reservoir during the month $t$.

$RR_{j,t}$ = The amount of water should be released from $j^{th}$ reservoir during the month $t$.

$RP_{i,t}$ = The amount of water should be released to the $i^{th}$ region during the month $t$.

$S_{j,t}$ = $j^{th}$ reservoir storage at the beginning of month $t$.

$A_i$ = Proportion of land allocated for paddy.

$B_i$ =Proportion of land allocated to crop type $c(c = 1, 2, ..., l)$

$P_i$= Cultivated land ratio to the available land $(0 \leq Pi \leq 1)$

$d_i^+$ = Positive deviation for proportion for paddy.

$d_i^-$ = Negative deviation for proportion for paddy.

$c_i^+$ = Positive deviation for proportion for other crops.

$c_i^-$ = Negative deviation for proportion for others crops.

**Objective function**

The objective of maximize the release of water for cultivated area.

$$Min \quad Z = \sum_{t=1}^{12}(d_i^- + c_i^-) \tag{3.17}$$

**Constraints**

The constraints can be divide into two parts.

- Hard constraints.
- Soft constraints.

**Hard constraints**

**01. Continuity Equations**

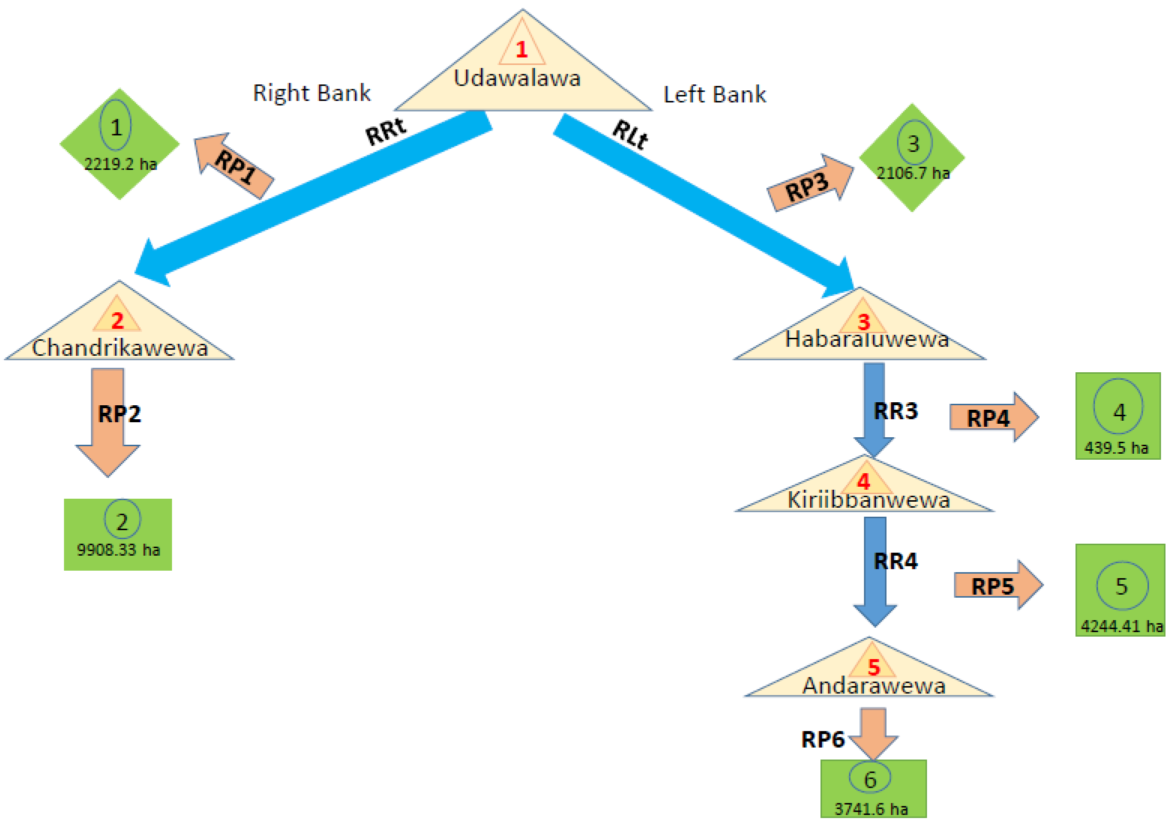


Figure 3.4: Plan for Water releasing

i . For Udawalawa Reservoir (j=1):

$$S_{j,t+1} = S_{j,t} + I_t - \sum_{j=1}^{n}(RR_t + LR_t) \qquad ;\forall \text{ t=1,2,3,...,12} \quad \text{and} \quad \text{j=1} \tag{3.18}$$

ii . For Chandrikawewa Reservoir (j=2):

$$S_{j,t+1} = S_{j,t} + RR_t - RP_{i,t} - RP_{(i+1),t} \qquad ;\forall \quad \text{t=1,2,3,...,12,} \quad \text{i=1} \quad \text{and} \quad \text{j=2} \tag{3.19}$$

iii . For Habaraluwewa Reservoir (j=3):

$$S_{j,t+1} = S_{j,t}+RL_t-RP_{i,t}-RR_{j,t} \qquad ;\forall \quad t=1,2,3,...,12, \quad i=3 \quad and \quad j=3 \tag{3.20}$$

iv . For Kiriibbanwewa Reservoir (j=4):

$$S_{j,t+1} = S_{j,t}+RR_{j,t}-RP_{i,t}-RR_{j,t} \qquad ;\forall \quad t=1,2,3,...,12, \quad i=4 \quad and \quad j=4 \tag{3.21}$$

v . For Andarawewa Reservoir (j=5):

$$S_{j,t+1} = S_{j,t} + RR_{j,t} - RP_{i,t} - RP_{(i+1),t} \qquad ;\forall \quad t=1,2,3,...,12, \quad i=5 \text{ and } \quad j=5 \tag{3.22}$$

**02. Storage limits**

$$S_{j,min} \leq S_{j,t} \leq S_{j,max} \qquad ;\forall \quad t=1,2,3,...,12, \quad j=1,2,...,5 \tag{3.23}$$

where, $S_{j,min}$ and $S_{j,max}$ are the minimum and maximum active storage of the reservoir.

**03. Goals constraints**

i . The first goal constraint is the proportion of the land allocated for paddy, which should be at least 60% of the total land.

$$\sum_{i=1}^{6}\left(A_i + d_i^- - d_i^+\right) = 0.6 \tag{3.24}$$

ii . The second goal constraint is the proportion of the land allocated for paddy, which should be at least 40% of the total land.

$$\sum_{i=1}^{6}\left(B_i + c_i^- - c_i^+\right) = 0.4 \tag{3.25}$$

**Soft constraints**

**01. Demands**

$$RP_{i,t} \geq \alpha_i H_i A_i + \sum_{c=1}^{l} \beta_c H_i B_{ic} \tag{3.26}$$

Where, $H_i$= Total area of the $i^{th}$ land

$\alpha_i$= The amount of water needed for paddy

$\beta_c$= The amount of water needed for crop $c$

$\sum_{i=1}^{6} P_i = \sum_{i=1}^{6} \left(A_i + B_i\right)$

**02. Limits for $P_i$**

The ratio of cultivated land is between zero and one.

$$0 \leq \sum_{i=1}^{6} P_i \leq 1 \tag{3.27}$$

**Relationship between the amount of water required for farming and the amount of water required for other crops**

Amount of water required for other crops = 0.6 * Amount of water required for farming. (3.8)

A Linear programming model with number of variables 222, 144 of inequality constraints and of 60 equality constraints is solved using MATLAB programming language.

A Goal programming model with number of variables 246, 138 of inequality constraints and of 72 equality constraints is solved using MATLAB programming language.

## 3.6 Forecasting the inflow to the Udawalawa reservoir

In this study, the only input variable that has been used is the inflow data of the Udawalawa reservoir. Inflow data for the coming years was predicted by using past data from 1996 to 2020. The aim is to build the most suitable and accurate model related to the agricultural sector. Time series forecasting is used to obtain the predicted values for the model.

### 3.6.1 Time series

Reservoir monthly inflow data can exhibit time-varying, non-stationary, and significant outliers due to the influence of unstable factors. These characteristics can alter the correlation between past and future inflow, and add noise to different time-series regions, making it challenging to build accurate forecasting models[26]. As a result, a single time-series forecasting model may struggle to capture the dynamic changing processes and features, and could face issues of local under-fitting or over-fitting.

Time series (TS) data refers to a set of observations that are collected sequentially over time. These observations are arranged in chronological order and can have varying time intervals such as annual, monthly, daily, hourly, and so on. Examples of time series data include monthly rainfall, weekly retail sales, daily confirmed COVID-19 cases and deaths, and hourly electrical demand.

A time series is a sequential arrangement of observations made over time, which can have different frequencies such as monthly, weekly, daily or hourly. It is a particular realization of a stochastic process, denoted by $\{X(t), t \in T\}$ where T is an index set and X(t) represents a random variable with reference to time. In Stochastic Dynamic Programming (SDP), reservoir inflow is considered as a stochastic process while developing recursive equations. Forecasting involves predicting the future accurately using historical data and knowledge of future events that could affect the forecast. However, the accuracy of a single forecast method is limited, which has led researchers to develop new technologies and methods for better hydrological predictions. In recent years, many hybrid approaches have emerged that combine multiple forecasting methods to improve reservoir inflow research and engineering practice[15].

#### 3.6.1.1 Time series components

The major components of a time series include trend pattern, seasonal pattern, cyclic pattern, and irregularity. Trend pattern represents a long-term increase or decrease in the data. Seasonal pattern occurs due to seasonal factors such as the quarter of the year, the month, or day of the week. Cyclic pattern shows rises and falls that are not of fixed frequency and have a duration of at least two years. Irregularity refers to variations that occur due to unpredictable factors that do not follow any particular pattern. Forecasting becomes easier if.[15],

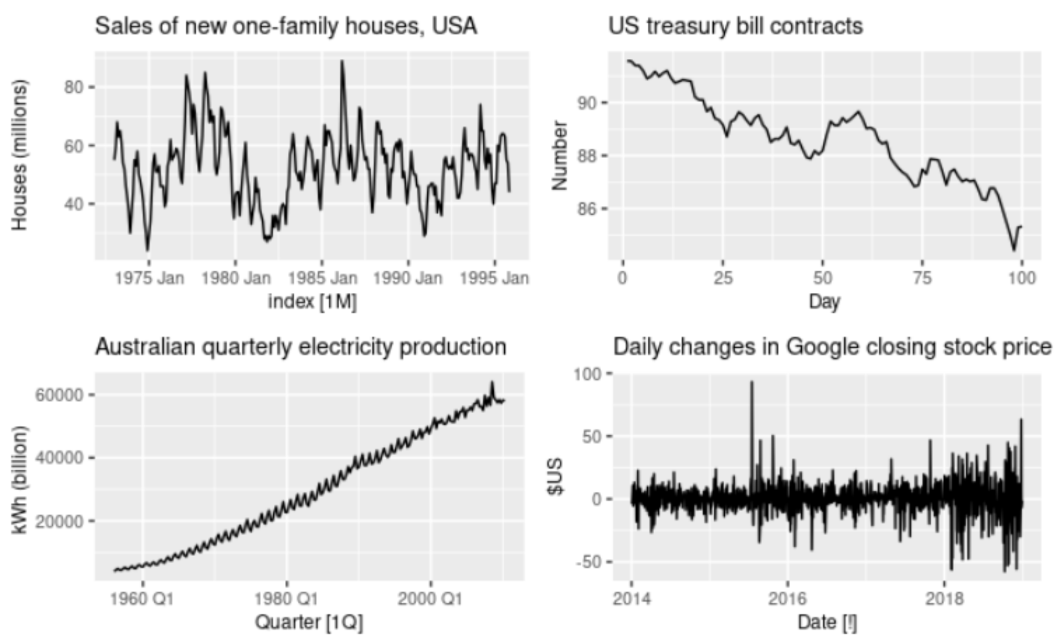


Figure 3.5: Four examples of time series showing different patterns.

The example of time series components is illustrated in Figure No. 3.5. The top left graph shows monthly housing sales exhibiting strong seasonality and cyclic behavior occurring over approximately a 610-year period. There is no discernible trend during this period. On the other hand, the top right graph of the US Treasury bill contracts displays no seasonality but a clear downward trend observed over 100 consecutive trading days in 1981. Although the data suggests a downward trend, it may be part of a longer cycle, which can only be established by examining a much longer series. The bottom left graph shows the Australian quarterly electricity production, with a strong increasing trend and seasonality. However, there is no evidence of cyclic behavior. Lastly, the bottom right graph displays the daily changes in the Google closing stock price, with no discernible trend, seasonality or cyclic behavior. The fluctuations appear to be random, and there are no strong patterns to assist with forecasting models.

#### 3.6.1.2 Time plots

When dealing with time series data, the most straight forward graph to create is a time plot. This plot displays the observations on the y-axis and the time of observation on the x-axis, with the points connected by straight lines. Time plots are useful for identifying trends, cycles, and irregularities in the data over time. Additionally, seasonal plots can be used to compare shifts in different series at specific time points.

**Time series plot:**

A time series plot is an initial way of visualizing time series data. It involves plotting observations against their corresponding time of observation. This type of plot can help identify trends, seasonality, and cyclic patterns within the data. A trend refers to the long-term increase or decrease in the data. When a series is affected by seasonal factors, it is considered to be seasonal. A cyclic pattern is present when the data exhibits rises and falls that occur at irregular intervals, usually lasting at least 2 years. An example of a time series plot is shown in Figure No. 3.6, which displays tea production in kg from January 2000 to June 2021.

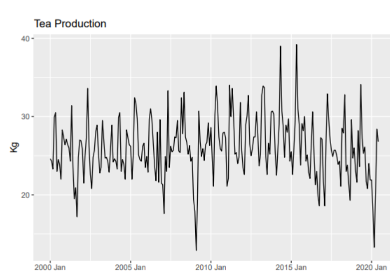


Figure 3.6: Example for time series plot

**Seasonal plot :**

A seasonal plot is a type of graphical representation that plots the data against the individual "seasons" in which the observations were made. It is similar to a time plot, but allows the underlying seasonal pattern to be seen more clearly, making it useful in identifying changes in the pattern over time. In the case of tea production, a seasonal plot reveals a similar pattern of fluctuation in production each year, with production ranging from 15kg to 80kg. May consistently shows the highest tea production each year, while November also shows a significant increase in total tea production. Figure No. 3.7 displays the monthly tea production in kilograms from January 2000 to June 2021.

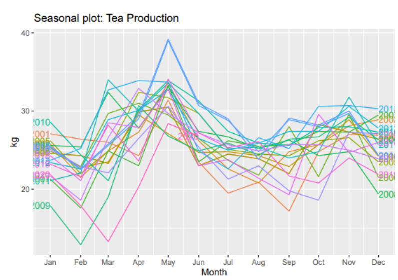


Figure 3.7: Example for seasonal plot of monthly

#### 3.6.1.3 Time series Features

Several time series features can be employed to investigate the properties of various time series. The following list comprises some of the useful time series features identified in this study.

**Auto-Correlation (ACF):-** All the autocorrelations of a series can be considered features of that series. Autocorrelation measures the strength of the linear relationship between lagged values of a time series and also computes autocorrelations of the changes in the series between periods. That is, the "difference" of the data and creating a new time series consisting of the differences between consecutive observations[28].

$$r_k = \frac{c_k}{c_0} = \frac{\sum_{t=k+1}^{T}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2} \tag{3.28}$$

#### Time series Decomposition

The process of breaking down a time series into its various components typically involves combining the trend and cycle components into a single trend-cycle component, which is often referred to as simply ”the trend.” Time series themselves are composed of three main components: a trend-cycle component, a seasonal component, and a remainder component that encompasses any other elements within the time series. In certain instances, particularly with time series that are observed daily or more frequently, there may be multiple seasonal components that correspond to different seasonal periods.

Various methods for extracting these components from a time series have been explored in the research, with the main goal being to enhance the accuracy of forecasting. In order to simplify the subsequent analysis and decomposition process, it can be beneficial to transform or adjust the time series beforehand[27].

**STL- Decomposition** STL is a robust and versatile technique used for decomposing time series. The acronym stands for ”Seasonal and Trend Decomposition Using Loess,” with loess being a non-linear method used for estimating relationships.

To illustrate the approach, Figure No. 3.8 showcases a monthly tea production dataset from Sri Lanka. The first panel depicts the original data, while the second panel represents the trend component the consistent frequency variation in the dataset. The third panel illustrates the seasonal component, showing variation in the data occurring at or near the seasonal frequency, resulting in strong fluctuations in this particular dataset. Finally, the fourth panel shows the remaining component, representing the residual variation in the dataset beyond what is accounted for by the trend and seasonal components[28].

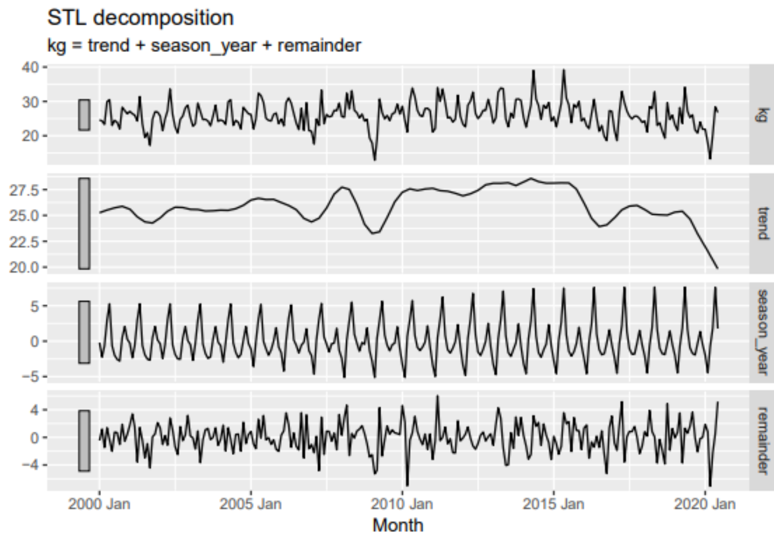


Figure 3.8: Example for STL plot

#### Trend and Seasonality in ACF plots

A time series decomposition can be used to measure the strength of a trend and seasonality in a time series[28]. Recall that the decomposition is written as,

$$y_t = S_t + T_t + R_t \tag{3.29}$$

where $T_t$ is the smoothed trend component, $S_t$ is the seasonal component, and $R_t$ is a remainder component. For strongly trended data, the seasonally adjusted data should have much more variation than the remainder component.

Therefore $var(R_t)/var(T_t + R_t)$ should be relatively small. But for data with little or no trend, the two variances should be approximately the same. So we define the strength of trend as:

$$F_T = max\Big(0, 1 - \frac{var(R_t)}{var(T_t + R_t)}\Big)$$

The strength of seasonality is defined similarly, but with respect to the de-trended data rather than the seasonally adjusted data:

$$F_S = max \quad \Big(0, 1 - \frac{var(R_t)}{var(S_t + R_t)}\Big)$$

A series with seasonal strength $F_S$ close to 0 exhibits almost no seasonality, while a series with strong seasonality will have $F_S$ close to 1 because $var(S_t + R_t)$ will be much smaller than.

### 3.6.2 Forecasting Methods

The process of producing forecasts can be split up into a few fundamental steps in RStudio [26]. It is shown in Figure No 3.9

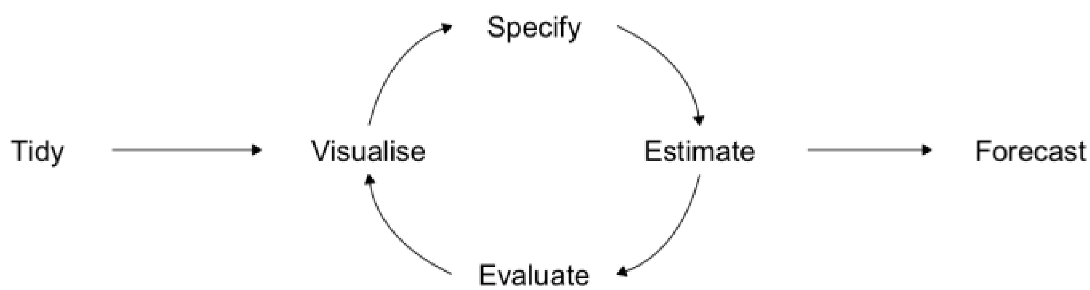


Figure 3.9: A tidy forecasting workflow

Preparing data is the first step of forecasting. Then that prepared data needs to be visualized. A model needs to be specified as the next step, and it is followed by the model estimation. Then an accuracy and performance evaluation need to be carried out, and finally, the production of the forecasts is to be done.

The most common approach in forecasting is to individually model the time series. These individual models are referred to as base forecasts. This study considered the Naive method, the Seasonal naive method (Snaive), the Average method, Drift method, the Auto Regressive Integrated Moving Average (ARIMA), and the Seasonal Auto-Regressive Integrated Moving Average (SARIMA) approach for generating base forecasts.

#### 3.6.2.1 MEAN(y): Average Method

In this approach, all future values are forecasted as equal to the historical data's average or mean. If we represent the historical data as y1, ..., yT, then this method assumes that all future values will be equal to $\{y_1, ....., y_T\}$.

Forecasts :

$$\hat{y}_{T+h|T} = \bar{y} = (y_1 + ..... + y_T)/T \tag{3.30}$$

#### 3.6.2.2 NAIVE(y): Naive Method

To generate Naive forecasts, we set all forecasts to the value of the most recent observation. In other words, there is no change in forecast values, and they remain constant at the last observed value.

Forecasts :

$$\hat{y}_{T+h|T} = y_T \tag{3.31}$$

#### 3.6.2.3 SNAIVE(y)∼ lag(m): Seasonal naive Method

A comparable technique is beneficial when dealing with highly seasonal data. In such cases, each forecast is set equal to the most recent observed value from the corresponding season (e.g., the same month of the previous year). The forecast for time $(T + h)$ is expressed formally as:

Forecasts :

$$\hat{y}_{T+h|T} = y_{T+h-m(k+1)} \tag{3.32}$$

where $m$ = seasonal period and $k$ is the integer part of $(h - 1)/m$

#### 3.6.2.4 Drift Method RW (y)∼ drift())

A modified version of the naive method involves allowing the forecasts to increase or decrease over time, with the amount of change, also referred to as "drift," set to the average change observed in the historical data. In this approach, the forecast for time $(T + h)$ is calculated as follows:

Forecasts :

$$\hat{y}_{T+h|T} = y_T + \frac{h}{T-1} \sum_{t=2}^{T}(y_t - y_{t-1})$$

$$= y_T + \frac{h}{T-1}(y_t - y_{t-1}) \tag{3.33}$$

#### 3.6.2.5 Auto-Regressive Integrated Moving Average Model - ARIMA(p,d,q)

ARIMA models offer an alternative approach to time series forecasting. These models seek to characterize the auto-correlations present in the data.

The initial step in this approach is to plot the data and identify any outliers, followed by transforming the data using a Box-Cox transformation if necessary. If the data is non-stationary, the first difference between data points must be computed until the data becomes stationary. Subsequently, the most suitable model is selected based on the AIC values. Next, the residuals of the chosen model are checked by plotting the ACF of the residuals and conducting a portmanteau test. If the residuals do not resemble white noise, a modified model is attempted. Finally, once the residuals resemble white noise, forecasts can be generated.

#### 3.6.2.6 Seasonal Autoregressive Integrated Moving Average Model SARIMA(p,d,q)(P,D,Q)$_m$

ARIMA models can effectively model a broad range of seasonal data, and to achieve this, a variant of ARIMA models known as SARIMA models are utilized. The term "SARIMA" refers to a seasonal ARIMA model, which involves the incorporation of supplementary seasonal terms in the ARIMA models used.

$$\text{SARIMA} \quad \underbrace{(p, d, q)} \quad \underbrace{(P, D, Q)_m}$$

Where $(p, d, q)$ and $(P, D, Q)$ are non-seasonal and seasonal components, respectively with a seasonality's. SARIMA model [30] was defined at Equation no (3.9),

$$\Phi(B^S)\phi(B)(1 - B^S)^D(1 - B)^d yt = \Theta(B^S)\theta(B)\epsilon_t \tag{3.34}$$

The seasonal ARIMA model, denoted as SARIMA, incorporates several parameters, including autoregressive (AR) parameters for seasonal and non-seasonal components denoted by $\Phi$ and $\phi$ respectively. The model also includes moving average (MA) parameters for seasonal and non-seasonal components denoted by $\Phi$ and $\phi$ respectively. Additionally, the model utilizes a backward operator represented by $B =$ and $B(y_t) = y_{t1}; (1 - B^S)^D = D^{th}$ which respectively represent the seasonal and non-seasonal differences. Furthermore, $B(y_t) = y_{t1}; (1 - B^S)^D = D^{th}$ represents an independently distributed random variable, while $P$ and $p =$ denote the orders of the AR components and $Q$ and $q =$ denote the orders of the MA components. $D$ and $d$ refer to the difference terms of the seasonal and non-seasonal components, respectively.

**Model identification.** To identify the appropriate SARIMA model for river flows, various methods were employed to check the stationarity of the data. These included graphical methods such as the sample autocorrelation function (ACF) and sample partial autocorrelation function (PACF), as well as unit root tests such as the Augmented Dickey-Fuller (ADF), Phillips-Perron (PP), and Kwiatkowski-Philips-Schmidt-Shin (KPSS)[32] tests at a significance level of ($\alpha$= 0.05). Additionally, the Mann-Kendall trend test was used to check for stationarity[33].

Based on the ACF diagram, the type (seasonal or non-seasonal) and order ($p, q, P, Q$) of model parameters were determined during model identification. The SARIMA parameters were optimized using conditional maximum likelihood, also known as the conditional sum of squares.

### 3.6.3 Fitted values and Residuals

**Fitted values**: To obtain fitted values in a time series, each observation can be forecasted by using all the preceding observations. These fitted values are represented as $\hat{y}_{t|t-1}$ and indicate the forecast of $y_t$ based on $y_1, y_2, y_3, ..., y_{t-1}$ and so on, up to $\hat{y}_{t|t-1}$. These fitted values are frequently used, and we may write them simply as $\hat{y}_t$. Typically, fitted values are based on one-step forecasts[26].

However, fitted values may not always be accurate forecasts since any parameters employed in the forecasting technique are estimated using all accessible observations in the time series, including future observations.

**Residuals**

In time series modeling, residuals refer to the leftover values after a model has been fitted. These residuals are computed as the difference between the actual observations and their corresponding fitted values, as shown by the equation[26]:

$$e_t = y_t - \hat{y}_t \tag{3.35}$$

When a transformation is applied in the model, it is sometimes beneficial to examine the residuals on the transformed scale. These residuals are referred to as **innovation residuals**.

**Residual Diagnostics**

To ensure a good forecasting method, the residuals should exhibit certain properties. First, they should be uncorrelated, as any correlations may suggest that there is remaining information that could be used to improve the forecasts. Second, the residuals should have a mean of zero, as a non-zero mean could indicate that the forecasts are biased. Third, the residuals should have constant variance, which is also known as homoscedasticity. Finally, the residuals should be normally distributed. Meeting these properties can help ensure accurate and reliable forecasts.

### 3.6.4 Statistical Test

To ensure the accuracy of a time series forecasting model, it is important to conduct **Diagnostic checkS** on the residuals. Residual correlograms, including the autocorrelation function (ACF) and partial autocorrelation function (PACF), as well as Ljung-Box Q tests[34], were employed to determine if the residuals were white noise with no autocorrelation. The ACF and PACF of the residuals were assessed at various lags to ensure that they were within the tolerance interval at 95% confidence limits. Additionally, a test was developed to evaluate whether a whole set of $r_k$ values was significantly different from a zero set.

There are two versions

(1) Box-Pierce test.

$$Q = T \sum_{k=1}^{h} r_k^2 \tag{3.36}$$

where $h$ is max lag being considered and $T$ is number of observations.

If each $r_k$ close to zero, $Q$ will be small.

If some $r_k$ values large (positive or negative), $Q$ will be large.

(2) Ljung-Box test.

$$Q^* = T(T+2) \sum_{k=1}^{h} (T-k)^{-1} r_k^2 \tag{3.37}$$

where $h$ is max lag being considered and $T$ is number of observations.

Modify the Box-Pierce test

Better performance, especially in small samples.

**Forecasting and performance evaluation of models**

To evaluate forecast accuracy, it is necessary to use actual forecasts. In the final step of modeling, forecast precision was assessed using performance measures. It should be noted that the size of residuals does not accurately indicate the magnitude of actual forecast errors. To determine the accuracy of forecasts, the model's performance on new data that was not included during model development must be evaluated. Akaike Information (AI)[35], Schwarz Bayes (SB)[36], and Hannan-Quinn (HQ)[37] criteria were used for this purpose. Additionally, the model performances were evaluated using the "root mean square error (RMSE) and mean absolute error (MAE). The model with the lowest AI, SB, and HQ values was selected for river flow forecasting".

$$AI = 2k - \log L \tag{3.38}$$

$$SB = 2k \log n - \log L \tag{3.39}$$

$$HQ = -2 \log L + 2k \log \log n \tag{3.40}$$

Where: $k$ = the amount of estimable parameters; $L$ = maximum likelihood; $n$ = number of sample.

# Chapter 4

# Analysis, Results and Findings

All the Results obtained from the study of fulfilling the needs of the farmers in Udawalawa region by using Linear Programming and Goal Programming has been discussed in this chapter.

To observe the amount of water inflow to the Udawalawa reservoir, predicted values obtained by using time series have been used. It is discussed under section 4.1. Predicted values have been used by using different forecasting values.

Data and the graphs related to release the water from the Udawalawa reservoir are presented in section 4.2. In addition the particular discussion helps to understand what is the demand, the amount of water currently releasing to do the agricultural activities and what is the actual required amount of water.

One of the important data obtained to do the study is the proportion of land utilized by paddy is to other crops in the udawalawa area. By using those data, the land can be optimally utilized with the available amount of water.

Final approach is to check the effect of data used for the study on obtaining optimum solutions. Goal programming has been used to monitor that.

## 4.1 Forecasting the inflow to the Udawalawa reservoir

The inflow to Udawalawa reservoir over a period of 1996 to 2020 is outlined using three segments; inflow over the entire period, Monthly Inflow, and Seasonal Inflow.

### 4.1.1 Time series analysis for inflow data to Udawalawa reservoir

The only input data used for this study is the inflow coming to the Udawalawe reservoir. For that, data from 1996 to 2020 has been used.These Inflow measurements were made in Million Cubic Meters(MCM). The following Figure No 4.1 shows the plot of the inflow pattern vs. time When observing the time series plot, there is a long-term fluctuation after 1996. It is continuous to be in the same value range until 2005, and from then there is an increasing time pattern until 2020.

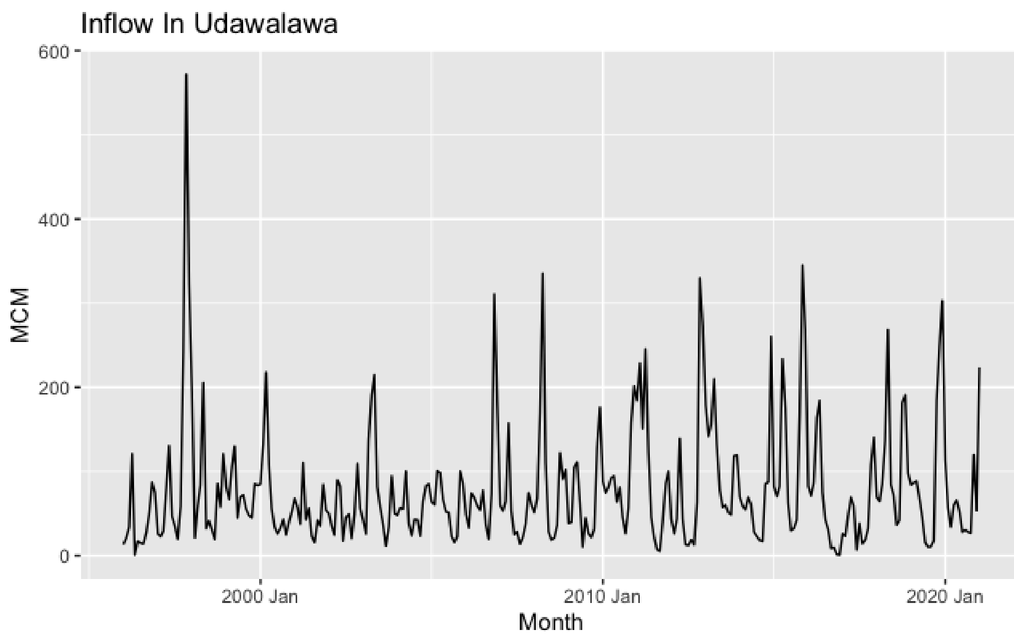


Figure 4.1: Inflow to the Udawalawa from 1996 to 2020

### Seasonal inflow pattern to Udawalawa

The following Figure No 4.2 shows the inflow pattern of the seasons. The main objective of the Udawalawa reservoir is to supply water to agricultural activities. The majority of water is required for the Yala and Maha seasons. Considering the seasonal plots for each year, it can be clearly seen that the highest inflows have been allocated to the months of April-July (yala season) and October-January (maha season).

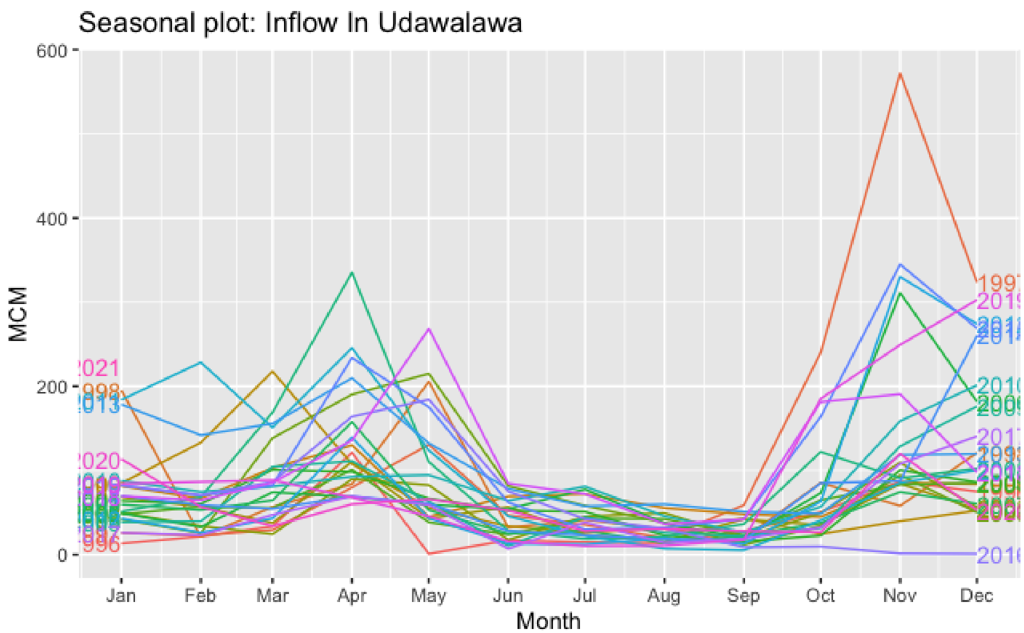


Figure 4.2: Seasonal plot for inflow to Udawalawa

**The seasonal sub series**

This Figure No 4.3 visualized the variation of the seasonal component. This includes twelve multiple panels. Those twelve vertical panels represent the months from January to February. The values for each month from 1996 – 2020 are plotted in relevant panels. By observing the sub-seasonal plot, the highest values can be seen in the months that belong to two seasons. Apart from that, there is a horizontal blue line as well. It illustrates the estimated values for each month. This is one of the methods to calculate the seasonal component.

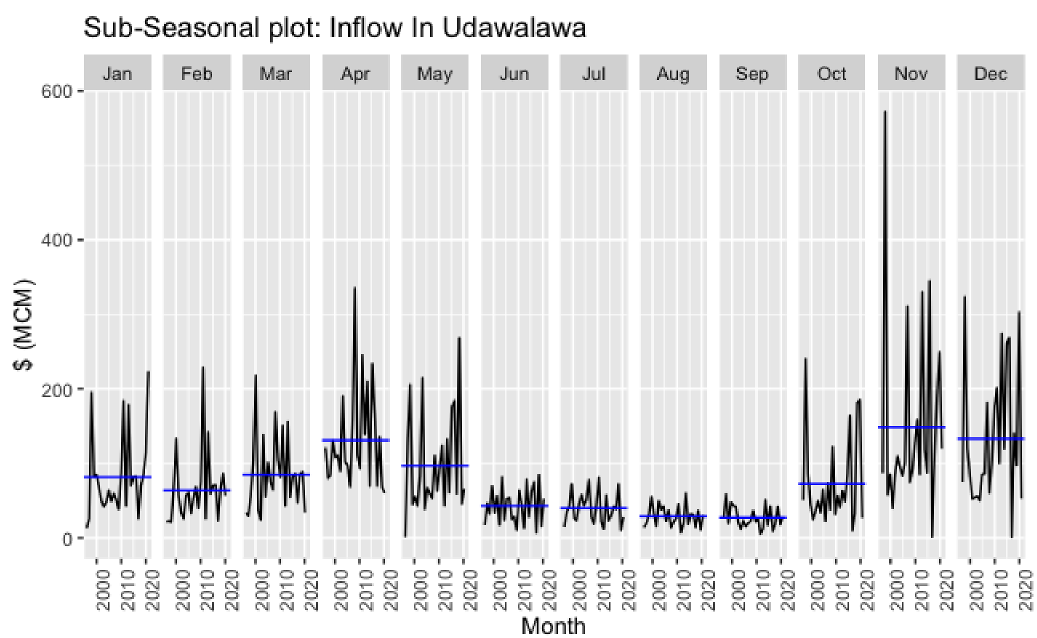


Figure 4.3: The seasonal sub series plot

**STL-Decomposition**

Under the decomposition method, it is trying to further express the original time series. Trend component, Seasonal component and residual are considered here. As per the additive decomposition,
By $Y_t = T_t + S_t + R_t$, Summation of trend, seasonal and residual can be obtained. There cant be seen a trend when observing the following STL decomposition Figure No 4.4. Therefore, it is necessary to consider about the seasonal capture. This seasonal component illustrates with a repetition pattern.The grey colour bar which is located left of the STL Decomposition Figure No 4.4 illustrates relative scales for each panel. Last panel shows a large grey bar. It says that there is a big variance in the remainder component.

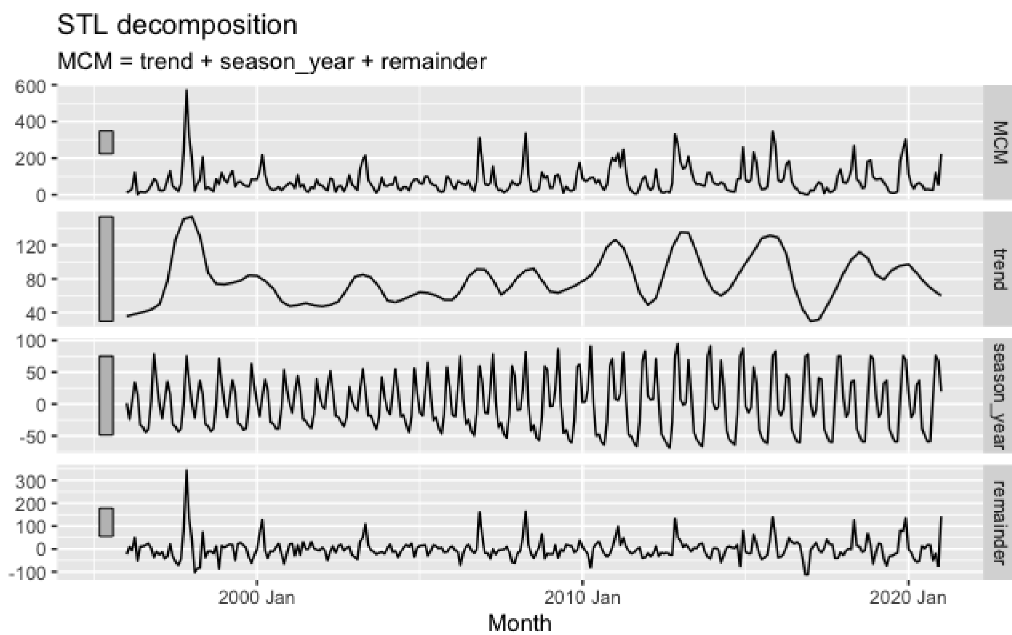


Figure 4.4: STL- Decomposition

### 4.1.2 Forecasting Methods

There are four major simple forecasting methods. They are **Mean method, Naive method, Snaive method and Drift method**. In mean method forecast of all future values is equal to mean of past data while forecast equal to last observed values in naive method. In snaive method, forecast to last value from same season and in the final drift method, forecasts are shown in the following Figure No 4.5.

#### Mean Method

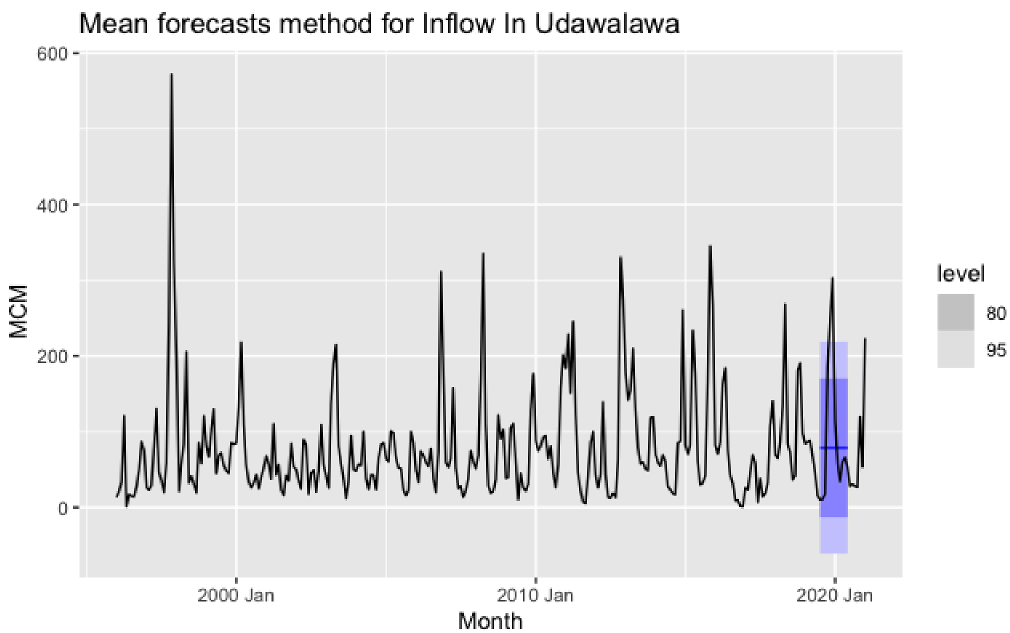


Figure 4.5: Mean method forecast appiled to Inflow in Udawalawa

The point forecast corresponds to the forecast from the mean technique, while the series in Figure No 4.5 is produced from the mean model.The straight line that runs parallel to the x-axis and represents the point forecast. It is important to note that the method relies on the supposition that the weights of the in-sample observations are equal, that is, that the first observation has the same exact weight of $1/T$ as the last one (being as important as the last one). Suppose the series exhibits some changes in level over time. The mean will not be appropriate in that situation since it will generate a forecast that is averaged out and takes into account values for both the changed and unchanged sections. However, Mean works well in many demand forecasting tasks and is a decent benchmark on intermittent data.

### Naive Method

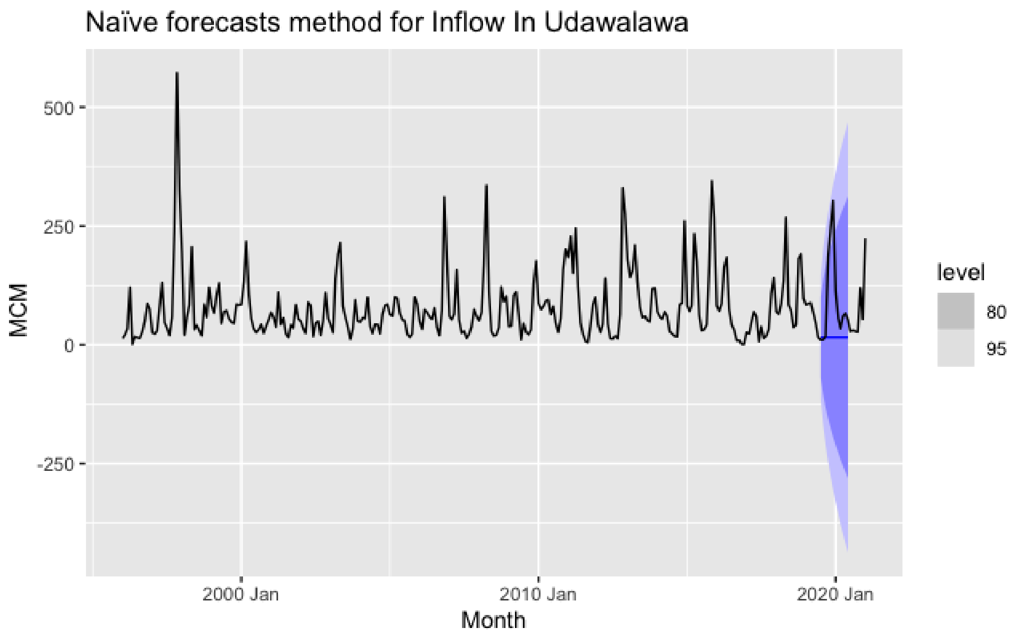


Figure 4.6: Naive method forecast appiled to Inflow in Udawalawa

Because Naive is built using Equation No 3.10, it lags the actual time series by one observation, as can be seen from the plot in Figure No 4.6. The straight line that runs parallel to the x-axis and represents the point forecast. The point forecast displayed in Figure No 4.6, despite displaying fast-level changes, is the best forecast for the time series given that the data was produced using Random Walk.

Note that Naive will update quickly and instantly reach the new level if the time series displays level shifts or other type of unexpected changes in dynamics. It won't, however, filter out the noise in the data because it only remembers one (last) observation; instead, it will copy it into the future. Consequently, its value in demand forecasting is limited (although it has applications in financial analysis). But because it is the most simple forecasting technique, it is regarded as one of the basic forecasting benchmark.

### Seasonal naive Method

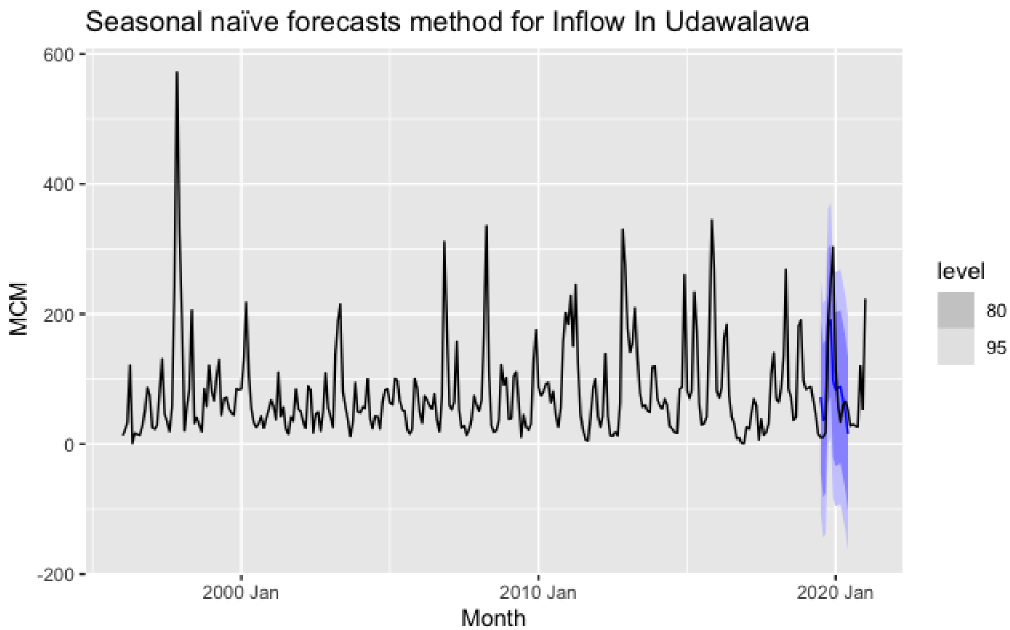


Figure 4.7: Seasonal naive method forecast appiled to Inflow in Udawalawa

Seasonal Naive is one of the widely used benchmarks for usage with seasonal data because it doesn't require the estimation of any parameters. Figure No 4.7 shows how the seasonal Random Walk-generated data appears and how the seasonal Naive point forecast applied to it performs.

The seasonal naive method is using the value from the january of 2020 in order to predict the january of 2021. The prediction of the febuary of 2021 is done with the value of the februaty of 2020, and so on. Nevertheless, although this method is more accurate and complex that the methods shown before, the trend has not been recognized yet.

### Drift Method

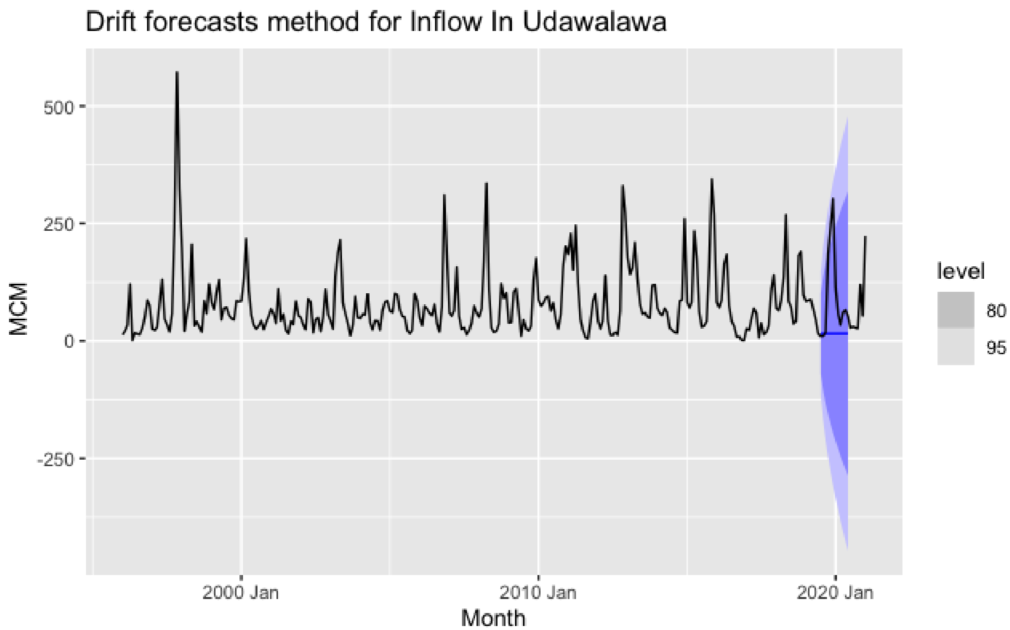


Figure 4.8: Drift method forecast appiled to Inflow in Udawalawa

Figure No 4.8 shown how the data generated from Drift method. This prediction understands the trend but lacks the seasonality pattern.

#### 4.1.2.1 Visualized representation

Figure No 4.9 shows visualization representation for the forecasting.

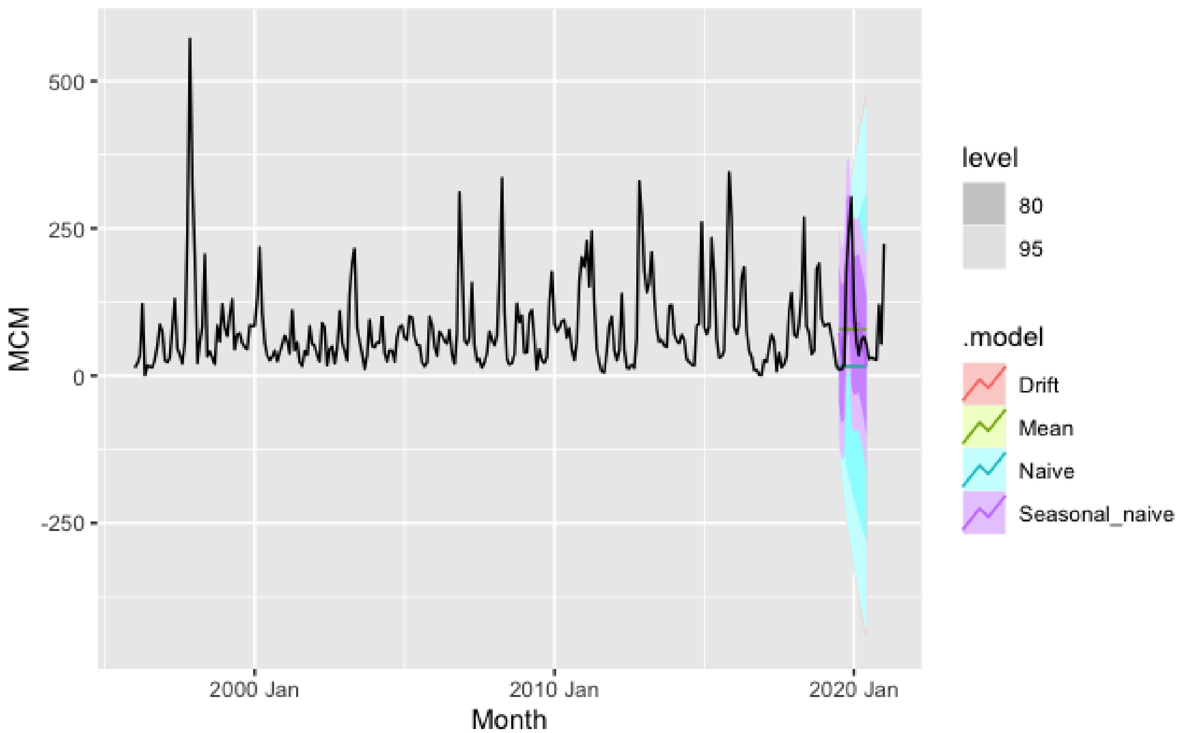


Figure 4.9: Visualized representation

This coloured area is used to do the seasonal estimate by looking at the behavior of the past data. But from this it is hard to conclude that which is the suitable method by this. For that residual diagnostics need to be used after model fitting.

#### 4.1.2.2 Residual diagnostics

As previously discussed in Section 3.6, residuals that fulfill four key properties are indicative of a good forecasting method. In Figure 4.10, we can observe the **innovation** residuals generated by the simple forecasting method. Each panel includes a time plot, a histogram, and an ACF plot of the residuals for the inflow in Udawalawa Reservoir. These graphs demonstrate that all forecasting methods produce forecasts that effectively consider all available information.

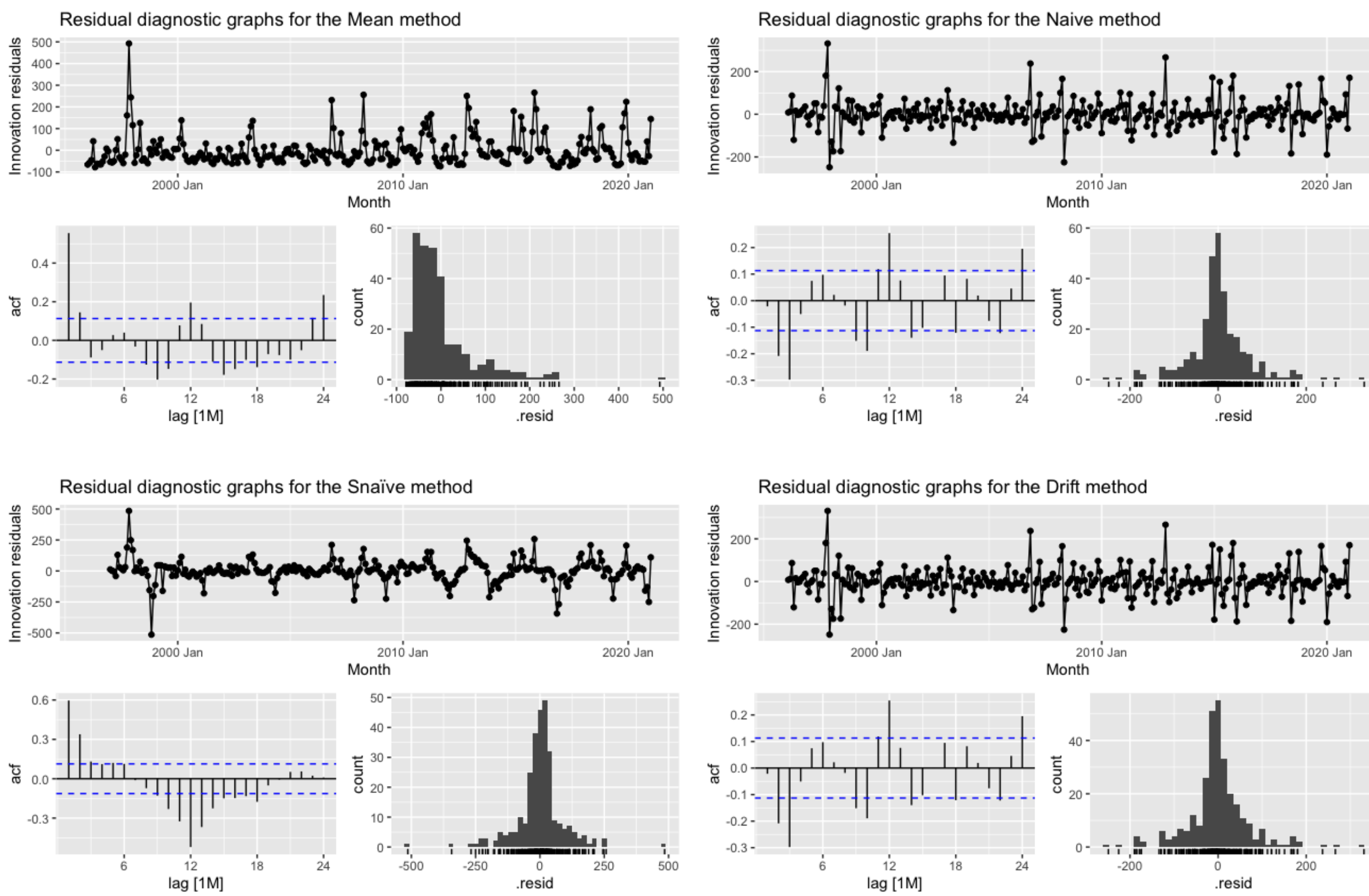


Figure 4.10: Residual diagnostic graphs for simple forecasting methods

In the top left of Figure No 4.10, the residual diagnostic graphs for the inflow in Udawalawa Reservoir using the mean method are presented. The time plot of the residuals indicates that the residual variation remains consistent across the historical data, and the mean of the residuals is not near zero. There is also a correlation in the residual series, which is evident in the histogram of the residuals. The histogram implies that the residuals may not follow a normal distribution since the right tail seems to be longer than expected. To ensure that the properties mentioned in Section 3.4 are fulfilled, we examine the ACF plot to verify whether each spike is within the critical band. However, in the mean method, many spikes do not fall within the critical band. This indicates that this method is not suitable for forecasting.

The residual diagnostic graphs for the Naive method applied to forecast the inflow in Udawalawa Reservoir are shown in the top right of Figure No 4.10. The time plot of the residuals indicates that the variation of the residuals remains consistent across the historical data. The histogram of residuals shows that the mean of the residuals is approximately zero, with a left tail that appears slightly longer. The properties discussed in Section 3.4 have been satisfied by examining the ACF plot to determine whether each spike is within the critical band. However, most of the spikes in this Naive method are outside the critical band, indicating that this method may not be effective for forecasting.

The residual diagnostic graphs for the Snaive and Drift methods used for forecasting the inflow in Udawalawa Reservoir are shown in the bottom left and bottom right of Figure No 4.10, respectively. In both cases, the time plot of the residuals shows that the residual variance can be assumed to be constant because the variation of the residuals remains relatively constant across the historical data, with a few outliers. The residuals histogram for both methods shows that the mean of the residuals is close to zero. However, many outliers are visible in the ACF plot, indicating no significant correlation.

In summary, the simple forecasting methods used do not fulfill certain properties. There are correlations between the innovation residuals, indicating that there is information in the residuals that can be utilized to improve the forecasts. If the residuals have a non-zero mean, the forecasts will be biased. Moreover, the variance of the innovation residuals is not consistent, which is referred to as **Heteroscedasticity**.

Ultimately, it can be inferred that these simple forecasting methods are not capable of capturing enough information from the given data.

When inspecting the ACF plot to verify if each spike is within the required limits, we are unintentionally conducting numerous hypothesis tests. The probability of obtaining a false positive for each test is low. However, with multiple tests, there is a chance of getting at least one false positive, which can mislead us into believing that the residuals have autocorrelation when they do not.

#### 4.1.2.3 Statistical approach

The Portmanteau test is used to see whether is there a significant different from a zero set, by looking out size the critical boundaries. Further this test is used to check the accuracy. There are two versions of the portmanteau test can be seen. They are **box-pierce test** and **Ljung-box test**. After applying portmateau test for autocorrelation,the results are shown in Table No 4.1.

| Simple forecasting Methods | Box-pierce Test | | Ljung-Box Test | |
|---|---|---|---|---|
| | stat-value | p-value | stat-value | p-value |
| Mean Method | 140 | 0 | 143 | 0 |
| Naive Method | 86.5 | $2.3\times10^{-13}$ | 89.1 | $7.2\times10^{-14}$ |
| Seasonal Naive Method | 283 | 0 | 291 | 0 |
| Drift Method | 86.5 | $2.3\times10^{-13}$ | 89.1 | $7.2\times10^{-14}$ |

Table 4.1: The results of the portmanteau test for simple forecasting methods

When considering the two output of the table, the p-value here is close to 0, and less than 0.05, which means that we would reject the null hypothesis. Thus, we can conclude that the residuals are distinguishable from a white noise series.

#### 4.1.2.4 Forecast accuracy in simple forecasting methods

According to previous subsections the summary of the simple forecasting methods includes in following Table No 4.2.

| Model | ME | RMSE | MAE | MPE | MAPE | MASE | RMSSE | ACF1 |
|---|---|---|---|---|---|---|---|---|
| Drift | 80.76 | 123.52 | 82.54 | 51.68 | 68.68 | 1.42 | 1.35 | 0.60 |
| Mean | 17.90 | 95.15 | 71.47 | -141.07 | 179.52 | 1.23 | 1.03 | 0.59 |
| Naive | 80.82 | 123.55 | 82.59 | 1.79 | 68.75 | 1.42 | 1.34 | 0.59 |
| SNaive | 12.51 | 69.05 | 46.53 | -82.50 | 119.20 | 0.80 | 0.75 | 0.36 |

Table 4.2: Summary of the basic forecasting methods

It is a standard practice to split the available data into two parts, namely training data and test data, when selecting models. The purpose of the training data is to estimate any parameters of a forecasting method, whereas the test data is used to evaluate the accuracy of the model. To match the forecasts, the accuracy function will extract the relevant periods from the data automatically when computing various accuracy measures.

Table No 4.2 demonstrates that the seasonal naive method performs the best among the data by comparing the values of Root Mean Square error (RMSE) and mean absolute percentage error (MAPE). Occasionally, different accuracy measures can lead to different results regarding the best forecast method. Nevertheless, in this case, all results have similarly high values. As mentioned in the statistical approach, we can infer that no simple forecasting method can be employed. Therefore, a significant amount of information is left unexplored using these forecasting methods.

#### 4.1.2.5 ARIMA method and Seasonal ARIMA method

The monthly historical time series plot of the Udawalawa Reservoir flow, spanning from January 1996 to December 2021, is presented in Figure No 4.1. The figure shows no apparent change in the mean of the series. However, the figure's periodic peaks and lows, observed during the wet and dry seasons, respectively, indicate the presence of seasonality in the time series. Thus, according to Figure No 4.1, the reservoir flow series appears to be stationary, except for its seasonality.

Moreover, the stationary property of the reservoir flow series was confirmed through the Augmented Dickey-Fuller (ADF) test, Phillips Person test (PP), and Kwiatkowski Philips Schmidt Shin unit root test (KPSS) at $\alpha$ = 0.05, as shown in Table No 4.3.

| Mann-Kendall trend test | | | Unit root and stationary tests | | | |
|---|---|---|---|---|---|---|
| Parameter | Mann-Kendall test | Seasonal Mann-Kendall test/period=12 | Parameter | ADF | PP | KPSS |
| p- value** | 0.5796 | 0.5810 | p-value * | ≤ 0.0001 | ≤0.0001 | 0.149 |
| $\alpha$ | 0.05 | 0.05 | $\alpha$ | 0.05 | 0.05 | 0.05 |

* one tailed; ** two tailed.
Explanations: ADF = Augmented Dickey-Fuller test, PP =Phillips-Perron test,
KPSS =Kwiatkowski-Philips-Schmidt-Shin test.

Table 4.3: Mann-Kendall trend test and unit root test for observed mean monthly Inflows

According to Table No 4.3, the p-value for the Mann-Kendall evaluation is 0.5796, indicating that there is no trend in the Udawalawa Reservoir inflow time series. Additionally, the $p$-value of 0.5810 for the seasonal Mann-Kendall test shows that there is no trend due to seasonality when considering 12 months of seasonality. As a result, the Udawalawa Reservoir is not trending based on the Mann-Kendall trend tests.

The computed p-values for the DF and PP tests were ≤ 0.0001, which is lower than the significance level ($\alpha$ = 0.05). This indicates that the non-stationary null hypothesis (i.e., there is unit root) is rejected at their level, proving the stationarity of the river flow. Furthermore, the KPSS test confirmed the stationarity of the time series since the computed p-value of 0.149 was greater than the significance level ($\alpha$ = 0.05). Based on all of the stationary analyses, it can be concluded that the observed mean monthly flow time series is stationary.

#### 4.1.2.6 Model Identification

The required AR() part and values of the MA() part which are required for the ARIMA model were obtained using ACF plot and PACF seen in Figure No 4.11. Hence, no need of non-seasonal differencing (d = 0). Seasonal differencing, on the other hand, is required for seasonal stationary. Moreover, as it is seen in Figure No 4.11, significant spikes were observed at intervals of every 12 months (12th, 24th lags, etc.) at the ACF plot and at lag 11th of the PACF plot. These significant spikes revealed that the observed mean monthly flow series is seasonally non stationary.

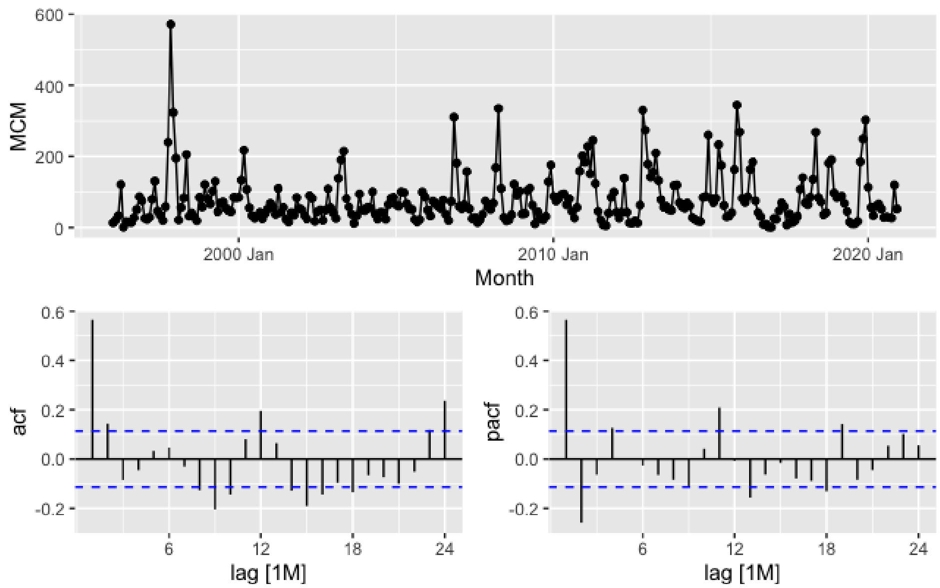


Figure 4.11: Autocorrelation function (ACF) and partial autocorrelation functions (PACF) for Udawalawa Reservoir inflows

Therefore, in order to make the time series seasonally stationary, seasonal differencing (D = 0) was performed on the observed inflow data. To further explore the data, SARIMA $(p, 0, q)(P, 0, Q)_{12}$ models were recommended, with the initial values of parameters p, q, and P determined based on the characteristics of seasonally differenced ACF and PACF plots (shown in Figure No 4.11).

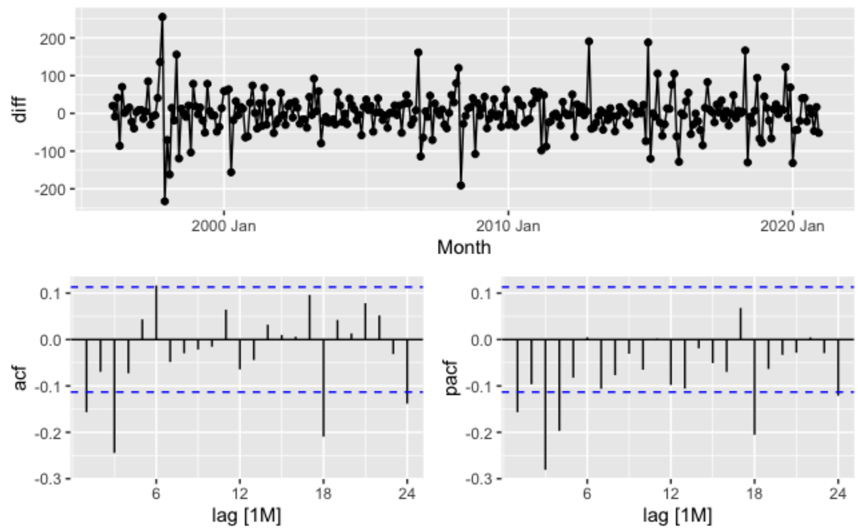


Figure 4.12: Seasonally differenced autocorrelation function (ACF) and partial autocorrelation functions (PACF) inflows water of Udawalawa Reservoir

PACF shows(Figure No 4.12) a decaying sinusoidal pattern, and ACF shows the last significant spike at lag 3. It suggests the availability of MA(3) components. Another significant spike of lag 18 can be seen by looking at the ACF plot. It suggests that there is another MA(2) component. **The SARIMA model can then be written as** $(0, 0, 3)(0, 0, 2)_{12}$.

Thus, the initial SARIMA model was identified as SARIMA $(0, 0, 3)(0, 0, 2)_{12}$. However, since model identification involves multiple attempts, the maximum values of 3 were set for $p, q$, and $P$ and $Q$ in the SARIMA model combinations. Therefore, more than two SARIMA models were identified for further evaluation based on the given conditions.

#### 4.1.2.7 Parameter Estimation

The Table No. 4.4 provides a summary of fit statistics for the out-of-sample forecast from 1996 to 2020 of the three best-fit SARIMA models selected from various models evaluated by RStudio software. In order to determine if the obtained model was the best or not, AI values were obtained by fixing the model for different values, based on the global minimization criteria. The resulting values for the models and their AI values are presented in Table No. 4.4. SARIMA $(0, 0, 3)(0, 0, 2)_{12}$ showed the best performance in forecasting the mean monthly flows of Udawalawa Reservoir, as indicated by the minimum performance measures, particularly the minimum value of AI and HQ statistics.

The least AI value obtained by the model ARIMA $(0, 0, 3)(0, 0, 2)_{[12]}$. This best model was become Seasonal ARIMA (SARIMA) model.Therefore, it was selected as the best model. In this selected model, residuals must be white noise. The following Figure No 4.13 shows residual plot values related to the selected model.

| Parameter | SARIMA | | |
|---|---|---|---|
| | (0,0,3)(0,0,2) | (0,0,1)(0,0,2) | (0,0,2)(2,0,0) |
| AI | 3178.14 | 3305.26 | 3276.91 |
| SB | 3178.52 | 3305.47 | 3277.2 |
| HQ | 3204.06 | 3323.78 | 3299.13 |
| RMSE | 55.69 | 58.67 | 55.75 |
| MAE | 36.73 | 39.48 | 36.90 |

Explanations: AI = Akaike information, SB = Schwarz Bayes criterion,
HQ = Hannan-Quinn information criterion, RMSE = root mean square error, MAE = mean absolute error.

Table 4.4: Performance values of the three best models

**Parameters of selected models**

The optimized model parameter values (Table No 4.5) obtained through the conditional maximum likelihood method fell within the significance intervals' minimum and maximum range.

| Parameter | MA1 | MA2 | MA3 | SMA1 | SMA2 | Constant |
|---|---|---|---|---|---|---|
| | 0.6809 | 0.3620 | 0.0432 | 0.1498 | 0.1058 | 77.8484 |
| s.e | 0.0599 | 0.0752 | 0.0588 | 0.0619 | 0.0622 | 8.3129 |

Table 4.5: Parameter estimates for SARIMA (0, 0, 3) (0, 0, 2) model

#### 4.1.2.8 Diagnostic Checking

**Autocorrelation test**

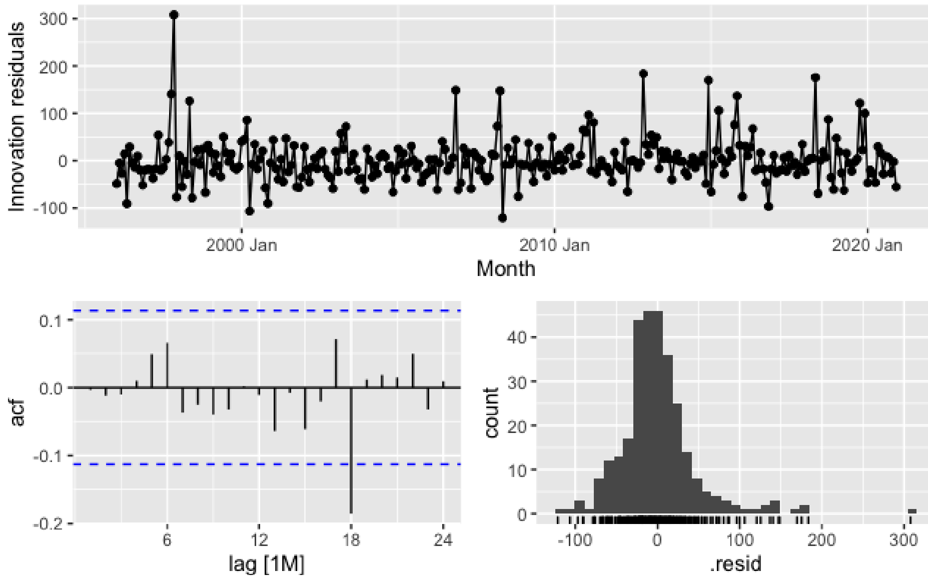


Figure 4.13: The residuals for the best model

The ACF and PACF plots of residuals (Figure No 4.13) indicate that all residuals at different lags fall within the 95% confidence interval, implying that each residual is relatively small compared to its standard error. One significant spike (at lag 18) among the 24 lags is consistent with white noise. The histogram of residuals shows that the mean is zero. These observations indicate that the errors obtained from the model are independent or white noise.

The LjungBox Q and Durbin-Watson auto-correlation tests (Table No 4.6) also support the graphical white noise test. The LjungBox Q test for auto-correlation up to the maximum lag order of 24 yielded a value of 20.4 with a p-value of 0.433. Similarly, the p-value of Durbin-Watson test for autocorrelation was 0.326. The p-values in these tests are higher than the significance level ($\alpha = 0.05$), indicating the absence of auto-correlations among residuals. Therefore, the selected model is appropriate as it can eliminate the residual dependency from the data.

| Parameter | Autocorrelation Test | | Heteroscedasticity Test | |
|---|---|---|---|---|
| | Ljung-Box Q' | Durbin-Watson | Breusch-Pagan | White |
| p | 0.433 | 0.326 | 0.806 | 0.985 |
| $\alpha$ | 0.05 | 0.05 | 0.05 | 0.05 |

Table 4.6: Autocorrelation and Heteroscedasticity tests

The following Table No 4.6 shows the results obtained from the tests.

**Heteroscedasticity Tests**

Table No 4.6 presents the outcomes of the heteroscedasticity test on residuals. The computed p-values for the Breusch-Pagans and White's tests were both greater than the significance level ($\alpha = 0.05$), at 0.806 and 0.985, respectively. These results indicate that the residuals are homoscedastic rather than heteroscedastic, meaning that the variance of residuals is constant across observations. This confirms the appropriateness of the linear SARIMA model in predicting inflows at Udawalawa reservoir. Furthermore, the Jarque-Bera normality test on residuals yields a value of 4.86452 with a p-value of 0.0878, indicating that the residuals are normally distributed at ($\alpha = 0.05$).

**4.1.2.9 SARIMA Forecasting**

The SARIMA $(0, 0, 3)(0, 0, 2)_{12}$ model was used to analyze the monthly inflow data of the Udawalawa reservoir. Validation and prediction time series graphs are shown in Figure No. 4.14. The forecast graph displays a comparable pattern to that of the historical and validation graphs and demonstrates the best performance. However, it is unable to accurately capture extremely high flows. Overall, the chosen model is deemed suitable for forecasting the inflow data of the Udawalawa reservoir.

The SARIMA $(0, 0, 3)(0, 0, 2)_{12}$ model provides insight into the order of dependencies within the time series data. A p-value of 0 indicates that the current data point, $y_t$, is dependent solely on its immediate preceding data point, $y_{t-1}$. Additionally, a P-value of 0 signifies that $y_t$ is dependent on the data from the same month in previous years, specifically $y_{t-12}$. The q-value of 3 indicates that the current data point is impacted by the residual shocks from the previous three data points, $\epsilon_{t-1}$,

$\epsilon_{t-2}$, and $\epsilon_{t-3}$. The Q-value of 2 in the second position $y_t$ signifies that the model is also impacted by the residual shocks from 12 months ago, $\epsilon_{t-12}$, and the 24 months before that, $\epsilon_{t-24}$. The non-seasonal component of the model was stationary, so no non-seasonal differencing (d = 0) was applied. Only one seasonal differencing was used (D = 0), indicating the absence of seasonality in the model.

Now the model can be used to obtain the future forecast values. The following Figure No 4.14 shows visualized results of future forecast values.

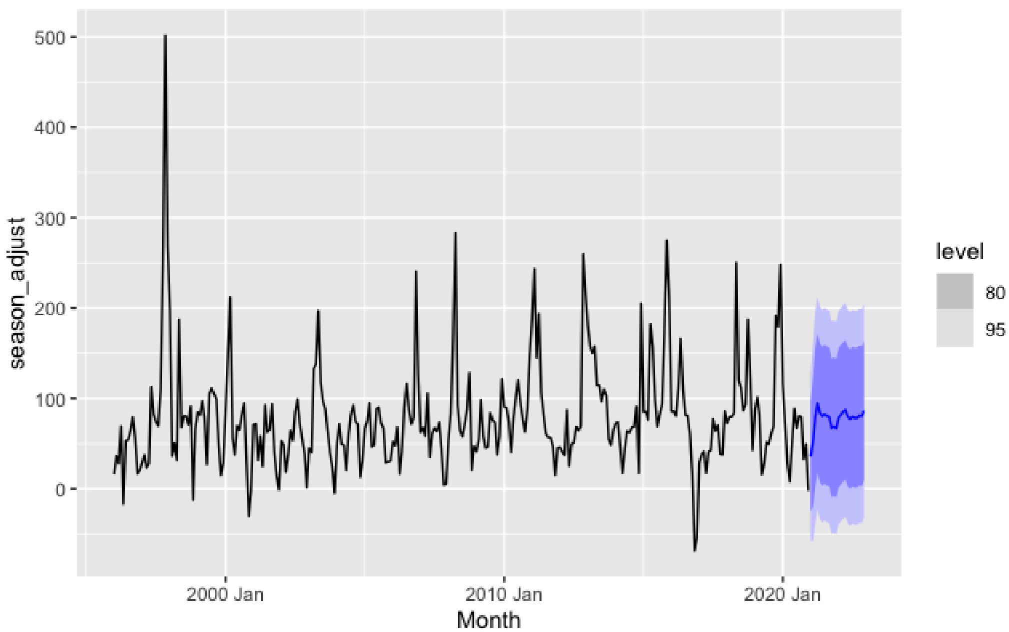


Figure 4.14: Forecasts of monthly Inflow data in Udawalawa using the SARIMA $(0,0,3)(0,0,2)_{[12]}$ model

In the graph, the past data is denoted by a black line, while the future forecast values are represented by a shaded area. A blue line is used to depict the point forecast. The strength of color used in the shaded area corresponds to the associated probability interval, which includes both 80% and 95% intervals.

**Forecasting values Obtain by SARIMA**$(0,0,3)(0,0,2)_{[12]}$ **model**

Table No. 4.7 displays the forecasting values for the year 2021, which were derived using the SARIMA method. The point forecasting values and two prediction interval values were also calculated. Prediction intervals provide a range in which we anticipate the value of $y_t$ to fall with a specific probability. Additionally, the forecasting values were obtained within the range of 80% to 95%. It is common for prediction intervals to widen as the forecast horizon increases. This is due to the fact that as we predict further into the future, there is a higher level of uncertainty associated with the forecast, leading to wider prediction intervals.

The Table No 4.7 shows that storage of the reservoirs increase in yala and maha seasons due to heavy rainfall and hence the amount of water released from the reservoirs tends to be higher. Forecasting value,which is also measured in Million Cubic Meters (MCM), can yeild results.

| | Point Forecasting | Lo80 | Hi80 | Lo95 | Hi95 |
|---|---|---|---|---|---|
| January | 35.7 | -25.358055 | 96.66967 | -57.65684 | 128.9685 |
| February | 52.9 | -19.097871 | 124.91384 | -57.21547 | 163.0314 |
| March | 78.6 | 1.982833 | 155.17561 | -38.56484 | 195.7233 |
| April | 94.5 | 17.649028 | 171.31536 | -23.02399 | 211.9884 |
| May | 83.9 | 7.070381 | 160.73671 | -33.60264 | 201.4097 |
| June | 80.2 | 3.391105 | 157.05743 | -37.28192 | 197.7305 |
| July | 82.7 | 5.833197 | 159.49953 | -34.83982 | 200.1725 |
| August | 80.7 | 3.836627 | 157.50296 | -36.83639 | 198.1760 |
| September | 79.5 | 2.633377 | 156.29971 | -38.03964 | 196.9727 |
| October | 67.3 | -9.500524 | 144.16581 | -50.17354 | 184.8388 |
| November | 69.0 | -7.857540 | 145.80879 | -48.53056 | 186.4818 |
| December | 67.0 | -9.816376 | 143.84995 | -50.48940 | 184.5230 |

Table 4.7: Forecasting values for 2021 In SARIMA

## 4.2 Results for Optimization methods

As previously discussed, the results of the study were obtained under two categories. The results obtained through the Linear Programming and Goal Programming technique are shown in the section 4.1. This section shows the amount of water needed to be released to each of the six catchments, obtained by solving the model.

### 4.2.1 Results for model I

Results of the model I will be discussed here including results obtained by linear programming technique. Land ratio of the cultivated area was maximized as the main objective. A linear programming model was developed for six different dependable cultivated areas and solved by using the MATLAB (R2015a) optimizer. The LP technique was used to obtain results using 204 hard and soft constraints based mainly on 222 variables.

#### 4.2.1.1 Release from Udawalawa reservoirs

Table No. 4.8 provides data on the amount of water discharged from the Udawalawa reservoir, measured in hectares meter (ha m). The table indicates that there are two methods for releasing water, one from the right bank and the other from the left bank.

| **Month** | Release from Left bank (ha m) | Release from Right bank (ha m) |
|---|---|---|
| **January** | 13000.94 | 6331.38 |
| **February** | 6563.41 | 4155.17 |
| **March** | 5486.1 | 3338.69 |
| **April** | 6709.6 | 4013.37 |
| **May** | 5478.54 | 3118.1 |
| **June** | 4057.47 | 2371.42 |
| **July** | 3597.57 | 1978.42 |
| **August** | 2678.61 | 1442.92 |
| **September** | 2613.12 | 1472.38 |
| **October** | 4130.28 | 2402.33 |
| **November** | 3844.55 | 2378.74 |
| **December** | 1453.6 | 1884.9 |

Table 4.8: Release from Udawalawa reservoir measured by hectare meter

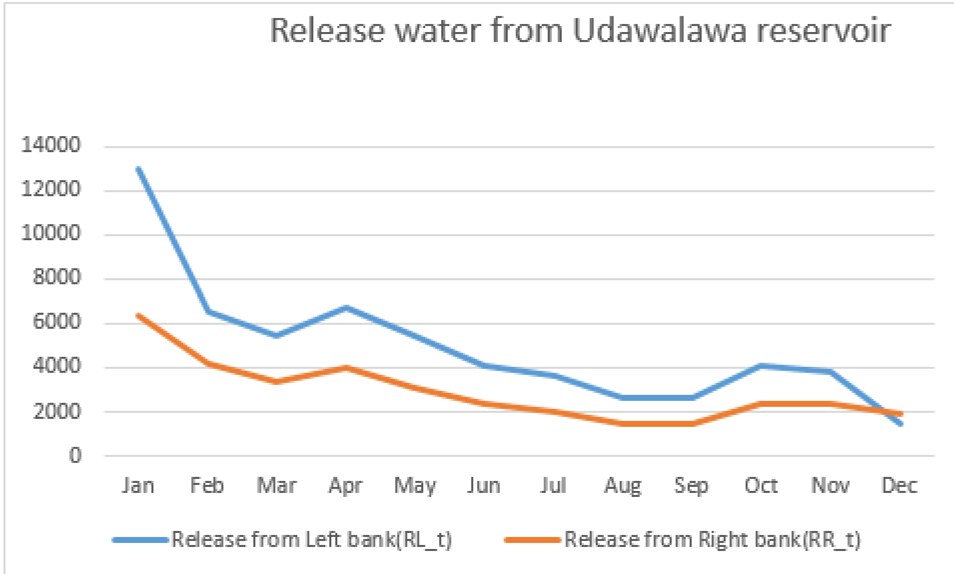


Figure 4.15: Release water from Udawalawa reservoir

The data from the Table No 4.8 were used to plot the line graph which shows in the Figure No 4.15. Even through the line graph shows data of 12 months(1 year) it is a cyclic process. The most amount of water has been released from the left bank of Udawalawa reservoir in the month of January and that amount has been gradually decrease throughout the year. Water release from right bank in Udawalawa reservoirs follows the same pattern as release data of left bank. But changes have been happened slightly which look almost constant through out the year.

#### 4.2.1.2 Release water from other Reservoir

For this research, four more reservoirs were taken into consideration in addition to the Udawalawa reservoir. The amount of water that has to be released from four different reservoirs is shown in Table 4.9 below. The amount of water released from each reservoir is measured in hectare meter (ha m).

| **Month** | **Habaraluwewa Reservoir( (RR3)** | **Kiriibbanwewa Reservoir(RR4)** | **Andarawewa Reservoir(RR5)** | **Chandrikawewa Reservoir(RR2)** |
|---|---|---|---|---|
| January | 4517.51 | 3021.16 | 1156.3 | 3717.51 |
| February | 3512.82 | 2365.96 | 1185.03 | 2055.01 |
| March | 3196.63 | 2258.49 | 1132.35 | 1654.53 |
| April | 4160.14 | 2739.43 | 987.13 | 2011.81 |
| May | 3700.65 | 2432.74 | 872.91 | 1437.42 |
| June | 3409.86 | 2284.85 | 777.58 | 894.36 |
| July | 5195.72 | 4298.17 | 195.53 | 717.32 |
| August | 2304.34 | 1681.02 | 840.5 | 719.74 |
| September | 2351.06 | 1721.96 | 863.02 | 735.73 |
| October | 3419.23 | 2291.66 | 790 | 922.53 |
| November | 3520.37 | 2336.16 | 832.81 | 945.53 |
| December | 3936.58 | 2456.55 | 910.73 | 638.47 |

Table 4.9: Release water from other reservoirs

Figure 4.16 shows the graphical representation of the data shown in the above Table No 4.9. There are 3 reservoirs located on the left bank of udawalawa reservoir. One of them is Habaraluwewa it has been denoted by $RR3$. Others are Kiriibbanwewa and Andarawewa. They are denoted by $RR4$ and $RR5$ respectively. The only reservoir located on the right bank of the udawalawa reservoir is Chandrikawewa and it is denoted by $RR2$.

It is evident that by looking at the graph shows in Figure No 4.16, there is a significant amount of water released from the Habaraluwewa and the Kiriibbanwewa throughout the year. There are two major reasons why the amount of water released from these two reservoirs are much higher than the other two. One of them is that the total cultivated area belonging to these two reservoirs is high and the next reason is to maintain the quantity of flow while feeding crops and other reservoirs simultaneously.

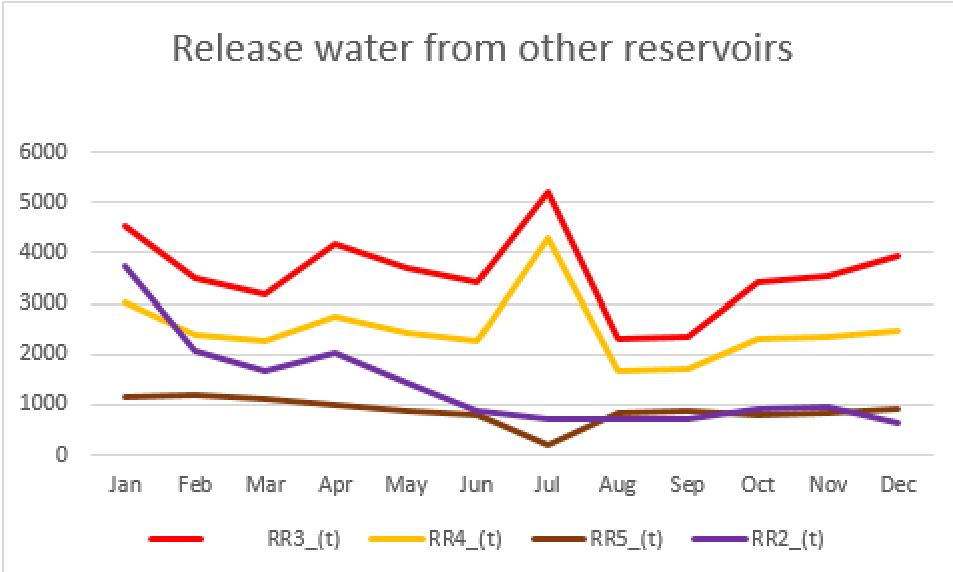


Figure 4.16: Release water from Other reservoirs

It is clear that the amount of water released in Yala and Maha paddy cultivating seasons is much higher compared to the rest of the months in a year.

#### 4.2.1.3 Releasing water for the cultivated land

Another set of results obtained through the LP model is shown in the Table No 4.10. It shows Figure No 4.17 what is the total amount of water that needs to be released to each and every 6 cultivated areas which are considered in the study.

| **Month** | **Region 1 (ha m)** | **Region 2 (ha m)** | **Region 3 (ha m)** | **Region 4 (ha m)** | **Region 5 (ha m)** | **Region 6(ha m)** |
|---|---|---|---|---|---|---|
| January | 2757.75 | 9304.24 | 2358.67 | 1682.62 | 1825.4 | 9003.72 |
| February | 2055.01 | 0 | 1913.95 | 1105.13 | 1185.03 | 0 |
| March | 1654.53 | 0 | 1497.82 | 941.64 | 1132.35 | 0 |
| April | 2040.13 | 9599.82 | 1763.64 | 1407.52 | 1739.75 | 9075.6 |
| May | 1690.43 | 9304.24 | 1576.19 | 1260.34 | 1560.52 | 9003.72 |
| June | 1460.26 | 9599.82 | 1233.75 | 1110.12 | 1504.67 | 9075.6 |
| July | 1256.1 | 9304.24 | 1043.95 | 918.19 | 4112.69 | 9003.72 |
| August | 719.74 | 0 | 544.04 | 632.26 | 840.5 | 0 |
| September | 735.73 | 0 | 547.93 | 652.27 | 863.02 | 0 |
| October | 1483.23 | 9599.82 | 1200.49 | 1114.95 | 1500.09 | 9075.6 |
| November | 1468.6 | 9304.24 | 1241.1 | 1194.65 | 1501.53 | 9003.72 |
| December | 1116.34 | 9599.82 | 1467.33 | 1317.09 | 1578.71 | 9075.6 |

Table 4.10: Release water for the paddy cultivated land

When considering the 6 cultivating areas, water has to be released to them throughout the year except lastly located areas in each left and right bank. In each year, compared to other months the amount of water released in the yala(October-January) and maha(April-July) seasons (months) are much higher. There is no water release to region 2 and region 6 from February to March and from August to September where farmers are not engaged in paddy cultivation. Therefore it can be concluded that other crops can be cultivated in those seasons in particular areas by using retained water. It is shows Figure No 4.17.

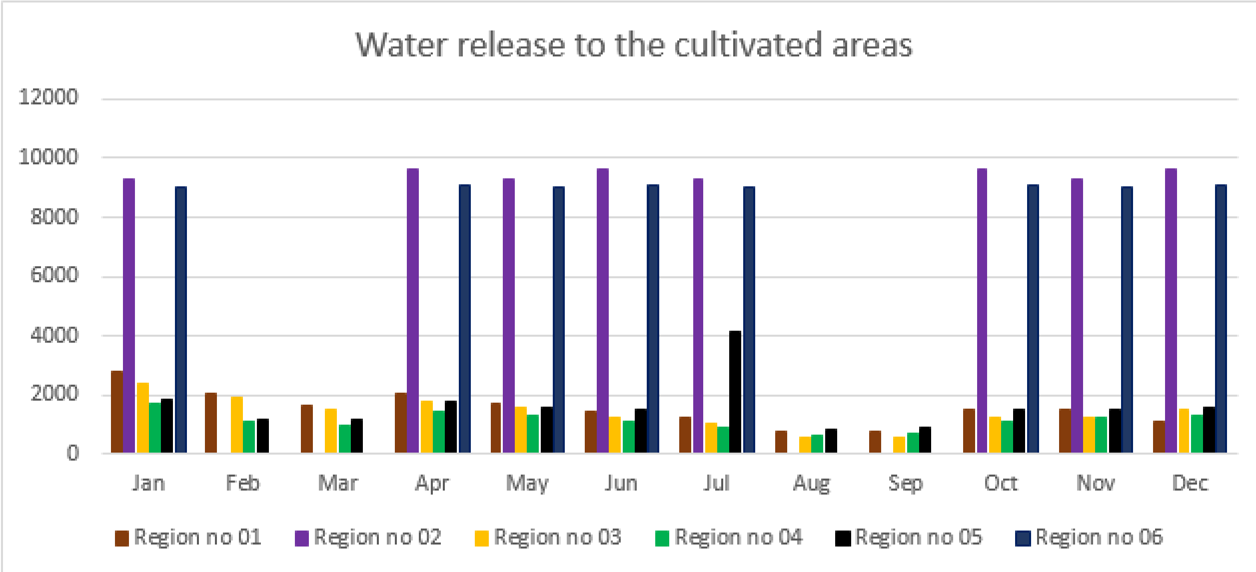


Figure 4.17: Release water to the cultivated areas

#### 4.2.1.4 Releasing water for the paddy and other crops cultivated by regions

The main objective of the study is to optimally use released water from the reservoirs by finding out the proportions of lands to be cultivated for each crop in considered areas.

Table No 4.11 shows the ratios between the area of land utilized to cultivate paddy and other crops. Main land area is divided as per those ratios.

| Regions | Releasing Ratio for paddy | Releasing Ratio for Other crops | Total Area(ha) | paddy(ha) | other (ha) |
|---|---|---|---|---|---|
| **Region 1** | 0.24 | 0.76 | 2219.2 | 532.608 | 1686.592 |
| **Region 2** | 0.29 | 0.71 | 9908.33 | 2873.4157 | 7034.9143 |
| **Region 3** | 0.21 | 0.79 | 2106.7 | 442.407 | 1664.293 |
| **Region 4** | 0.28 | 0.72 | 439.5 | 123.06 | 316.44 |
| **Region 5** | 0.2 | 0.8 | 4244.41 | 848.882 | 3395.528 |
| **Region 6** | 0.3 | 0.7 | 3741.6 | 1122.48 | 2619.12 |

Table 4.11: Ratio for paddy and other crops

Figure No 4.18 shows the graphical representation of the results shown in the Table No 4.11 which obtained by solving the linear programming model.Figure No 4.18 graphically represents the total area of land cultivated in 6 regions and proportion of land utilized by paddy and other crops.

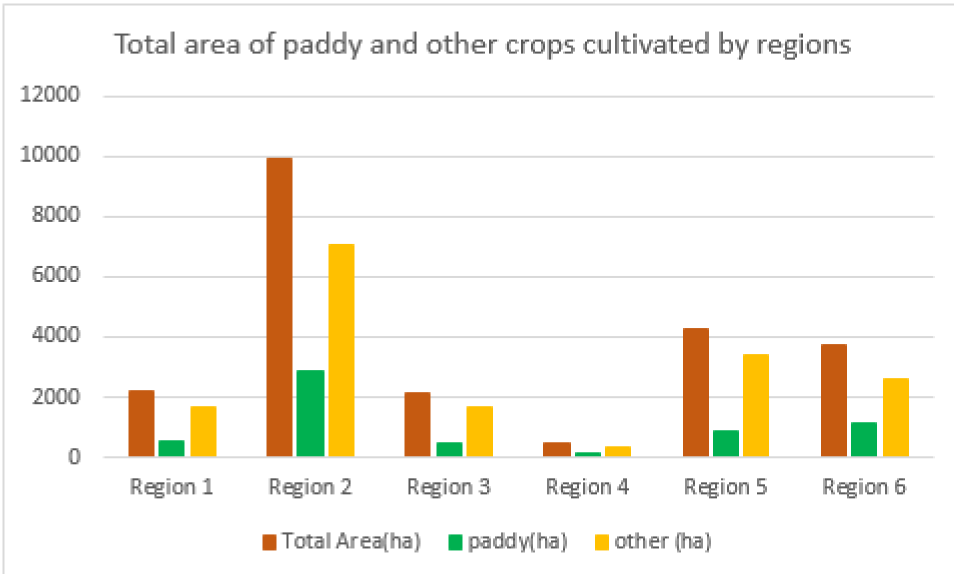


Figure 4.18: Release water to the cultivated areas

Figure No 4.19 shows the proportion of land utilized to cultivate paddy and other crops in each region separately. Even Though it was planned to cultivate paddy in 60% of the land in each area, the results obtained by linear programming model suggest to reduce the proportion of land utilized to cultivate paddy to less than 60%. Therefore it was decided to use a goal programming model instead of a linear programming model in order to optimize the results much further.

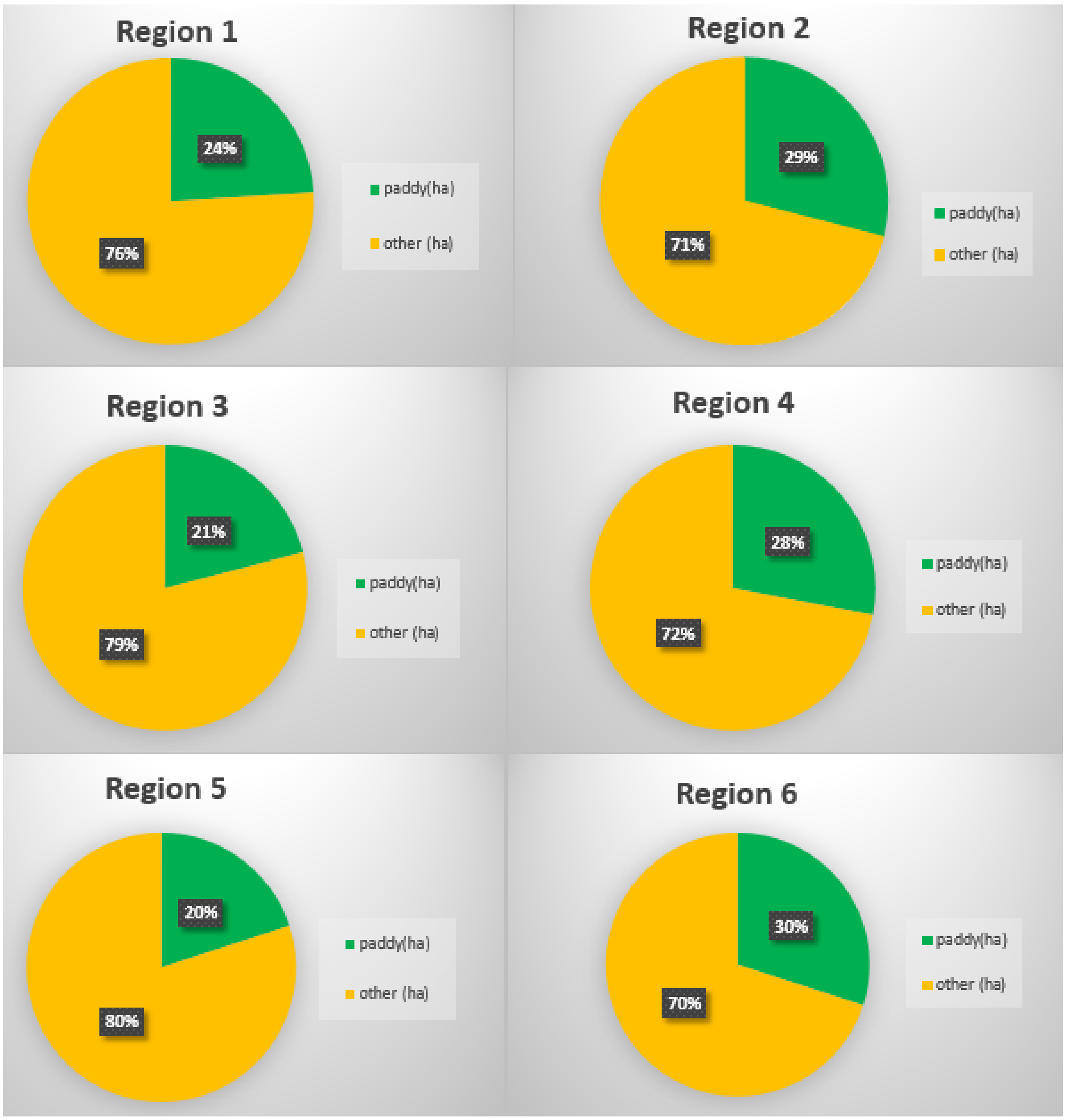


Figure 4.19: Land allocation optain by Linear programming technique

### 4.2.2 Results for Model II

**Model II** was developed in order to optimize the solutions which were obtained by solving the **model I**. All the results obtained by solving the **model II** will be discussed in this section.

#### 4.2.2.1 Release from Udawalawa reservoirs

The main objective of the study is to calculate the water quantity that should be released from the Udawalawa reservoir measured in hectares meter(ha m). Releasing is done in two ways, and they are from the left bank and the right bank. The following Table No 4.12 shows the amount of water that should be released from each bank in the 12 months of the year.

| **Month** | **Release from Left bank(RL_t)** | **Release from Right bank(RR_t)** |
|---|---|---|
| January | 2089.59 | 1660.14 |
| February | 2352.68 | 2616.4 |
| March | 0 | 0 |
| April | 6663.4 | 6414.81 |
| May | 4442.27 | 4276.54 |
| June | 6342.44 | 5901.18 |
| July | 2410.55 | 4790.17 |
| August | 0 | 0 |
| September | 0 | 0 |
| October | 9016.08 | 6414.81 |
| November | 2089.59 | 4276.54 |
| December | 9016.08 | 6414.81 |

Table 4.12: Release from Udawalawa reservoir

The data from the Table No 4.12 were used to plot the line graph which shows in the Figure No 4.20. Even through the line graph shows data of 12 months(1 year) it is a cyclic process.

The Figure No 4.20 shows water release from the left and right banks of the Udawalawa reservoir. When looking at the line graph, it can be seen that the amount of water released from both right and left banks in each month of the year are similar.

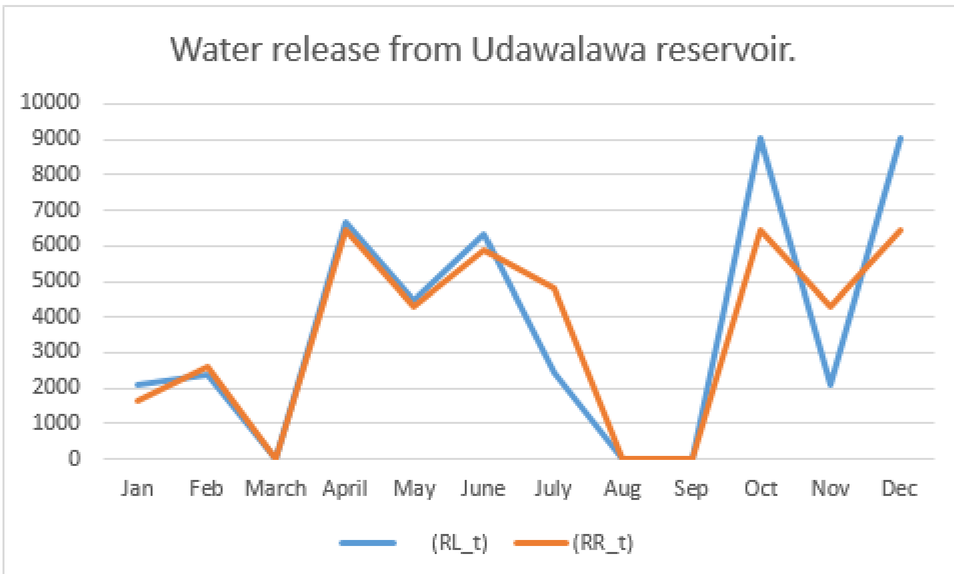


Figure 4.20: Release from Udawalawa reservoir

#### 4.2.2.2 Release water from one reservoir to other reservoir

Apart from the Udawalawa reservoir, four other different reservoirs were considered for this study. The following Table No. 4.13 shows the amount of water that needs to be released from Habaraluwewa reservoir to Kiriibbanwewa ($RR3_{(t)}$) and Kiriibbanwewa reservoir to Andarawewa reservoir ($RR4_{(t)}$). The amount of water released from each reservoir is measured in hectare meter (ha m).

| **Month** | **Habaraluwewa Reservoir $RR3_{(t)}$** | **Kiriibbanwewa Reservoir $RR4_{(t)}$** |
|---|---|---|
| January | 1579.98 | 3047.37 |
| February | 1973.73 | 320.96 |
| March | 0 | 0 |
| April | 5330.56 | 5052.50 |
| May | 3553.71 | 3368.33 |
| June | 5009.61 | 4731.54 |
| July | 1900.94 | 3368.33 |
| August | 0 | 0 |
| September | 0 | 0 |
| October | 7304.30 | 5373.46 |
| November | 1579.97 | 3047.37 |
| December | 7304.30 | 5373.46 |

Table 4.13: Release water from one reservoir to other

Figure No 4.21 shows the amount of water needed to be released from other reservoirs throughout the 12 months of the year. When considering the solutions, it is clear that the majority of the water has to be released during Yala(April-July) and Maha (October-January) seasons from the Habaraluwewa and Kiribanwewa reservoirs.

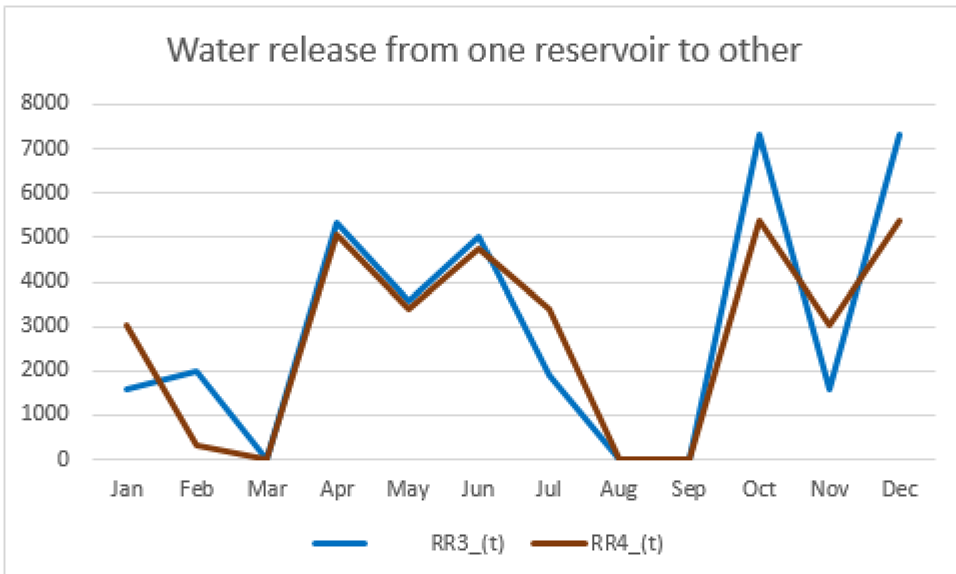


Figure 4.21: Release water from other reservoirs

#### 4.2.2.3. Releasing water for the cultivated land

Six main catchments were considered for this study. Table No 4.14 shows the amount of water needed to be released to each catchment in order to cultivate paddy and other crops. Region 1 and Region 2 are located on the right bank of the Udawalawa reservoir, and the rest of the regions are located on the left bank of the Udawalawa reservoir.

| **Month** | **Region 1 RP1(t)** | **Region 2 RP2(t)** | **Region 3 RP3(t)** | **Region 4 RP4(t)** | **Region 5 RP5(t)** | **Region 6 RP6(t)** |
|---|---|---|---|---|---|---|
| January | 936.01 | 3340.53 | 888.56 | 185.372 | 1790.2 | 1578.13 |
| February | 0 | 0 | 0 | 0 | 0 | 0 |
| March | 0 | 0 | 0 | 0 | 0 | 0 |
| April | 1404.02 | 5010.79 | 1332.84 | 278.06 | 2685.31 | 2367.19 |
| May | 936.01 | 3340.52 | 888.56 | 185.37 | 1790.21 | 1578.13 |
| June | 1404.02 | 5010.79 | 1332.84 | 278.06 | 2685.31 | 2367.19 |
| July | 936.01 | 3340.52 | 888.56 | 185.37 | 1790.2 | 1578.13 |
| August | 0 | 0 | 0 | 0 | 0 | 0 |
| September | 0 | 0 | 0 | 0 | 0 | 0 |
| October | 1404.02 | 5010.79 | 1332.84 | 278.06 | 2685.31 | 2367.2 |
| November | 936.02 | 3340.55 | 888.56 | 185.37 | 1790.2 | 1578.13 |
| December | 1404.02 | 5010.79 | 1332.84 | 278.06 | 2685.31 | 2367.2 |

Table 4.14: Release for the Cultivated area

Figure 4.22 shows Amount of water to be released to the cultivating areas throughout the year. As per the model, there is no water release will be happened during the months where paddy will not be cultivated. Excess amount of water remaining from the previous months, ground water and rainwater will be used to cultivate other crops in those months.

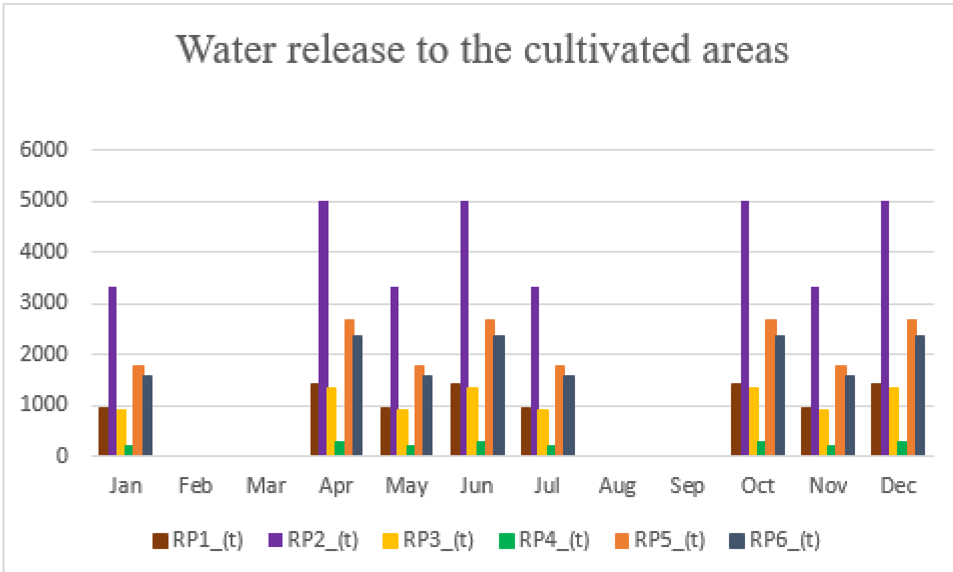


Figure 4.22: Release water to the cultivated areas

**Region 1- Chandrikawewa area**

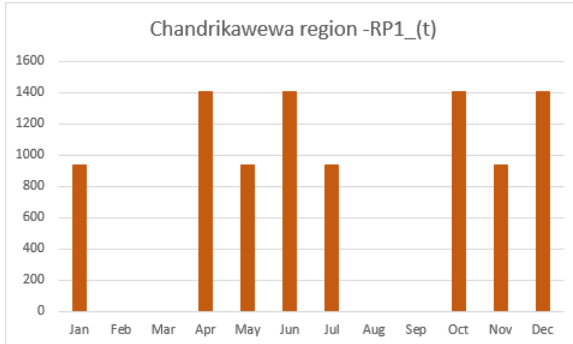


Figure 4.23: Release water to the Chandrikawewa area

A land area of 2219.2ha is used for agricultural purposes around the Chandrikawewa area. Figure No 4.23 shows amount of water to be released to the Chandrikawewa area in order to optimize the paddy and other crops cultivation throughout the year. These results given in the Table No 4.15.

| **Month** | **Region no 01 RP1(t)** |
|---|---|
| January | 936.01 |
| February | 0 |
| March | 0 |
| April | 1404.02 |
| May | 936.01 |
| June | 1404.02 |
| July | 936.01 |
| August | 0 |
| September | 0 |
| October | 1404.02 |
| November | 936.02 |
| December | 1404.02 |

Table 4.15: Release for the Chandrikawewa region

**Region 2 - Angunukolapelessa area**

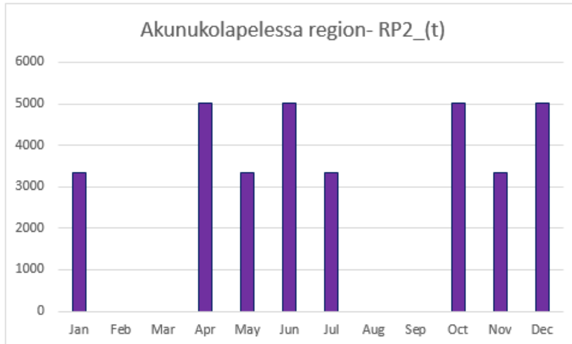


Figure 4.24: Release water to the Angunukolapelessa area

A total land area of 9908.33ha is used for agricultural purposes in Angunukolapelassa area. As the solution from the model, it has shown in Figure No 4.24 that amount of water needs to be released for the months except February-March and August-September. Those values are summarized in the Table No 4.16 below.

| **Month** | **Region no 02 RP2(t)** |
|---|---|
| January | 3340.53 |
| February | 0 |
| March | 0 |
| April | 5010.79 |
| May | 3340.52 |
| June | 5010.79 |
| July | 3340.52 |
| August | 0 |
| September | 0 |
| October | 5010.79 |
| November | 3340.55 |
| December | 5010.79 |

Table 4.16: Release for the Angunukolapelessa region

**Region 3 - Sevanagala area**

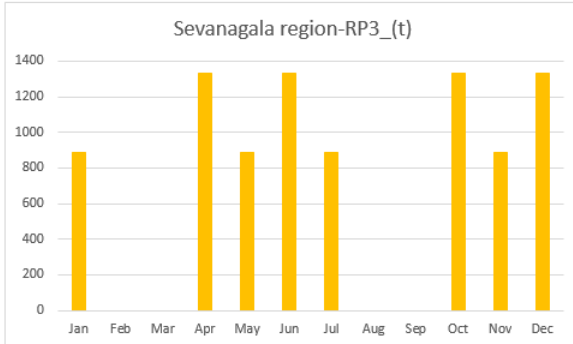


Figure 4.25: Release water to the Sevanagala area

Sevanagala area is located around the Habaraluwewa reservoir which is one of the reservoirs in the left bank of the Udawalawa reservoir. A land area of 2106.7*ha* is utilized for agricultural activities in this area. Figure No 4.25 shows the amount of water to be released to the Sevanagala area in order to optimize the paddy and other crop cultivation throughout the year. These results are given in Table No 4.17

| **Month** | **Region no 03 RP3(t)** |
|---|---|
| January | 888.56 |
| February | 0 |
| March | 0 |
| April | 1332.84 |
| May | 888.56 |
| June | 1332.84 |
| July | 888.56 |
| August | 0 |
| September | 0 |
| October | 1332.84 |
| November | 888.56 |
| December | 1332.84 |

Table 4.17: Release for the Sevanagala region

### Region 4 - Habaraluwewa area

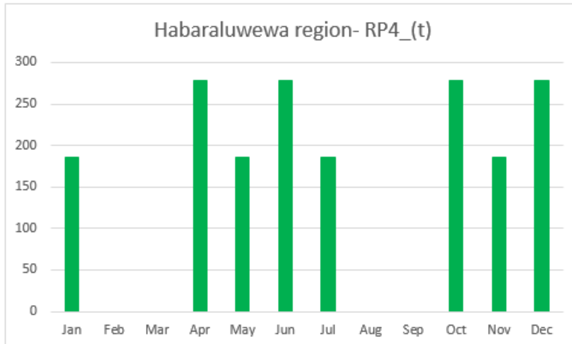


Figure 4.26: Release water to the Habaraluwewa area

Figure No 4.26 shows the amount of water to be released to the Habaraluwewa area in order to optimize the paddy and other crop cultivation throughout the year. Table No 4.18 the solutions obtained by the model II on amount of water releasing requirement for the Habaraluwewa area.

| **Month** | **Region no 04 RP4(t)** |
|---|---|
| January | 185.372 |
| February | 0 |
| March | 0 |
| April | 278.06 |
| May | 185.37 |
| June | 278.06 |
| July | 185.37 |
| August | 0 |
| September | 0 |
| October | 278.06 |
| November | 185.37 |
| December | 278.06 |

Table 4.18: Release for the Habaraluwewa region

**Region 5 - Kiriibbanwewa area**

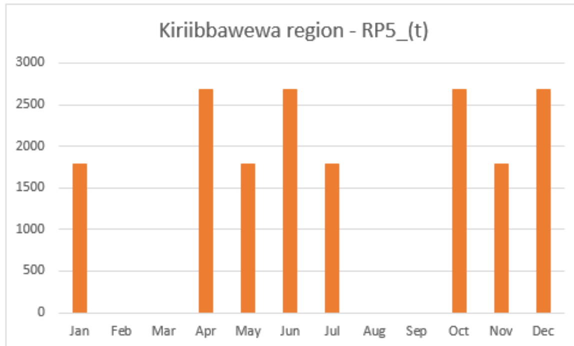


Figure 4.27: Release water to the Kiriibbanwewa area

| **Month** | **Region no 05 RP5(t)** |
|---|---|
| January | 1790.2 |
| February | 0 |
| March | 0 |
| April | 2685.31 |
| May | 1790.21 |
| June | 2685.31 |
| July | 1790.2 |
| August | 0 |
| September | 0 |
| October | 2685.31 |
| November | 1790.2 |
| December | 2685.31 |

Table 4.19: Release for the Kiriibanwewa region

Total land area of 4244.41$ha$ is used for agricultural purposes around the Kiriibbanwewa reservoir. Figure No 4.27 shows the amount of water to be released to the Kiriibbanwewa area and Table No 4.19 summarizes the statistics of the solutions obtained by the model on the amount of water-releasing requirement for the Kiriibbanwewa area.

**Region 6: Andarawewa area**

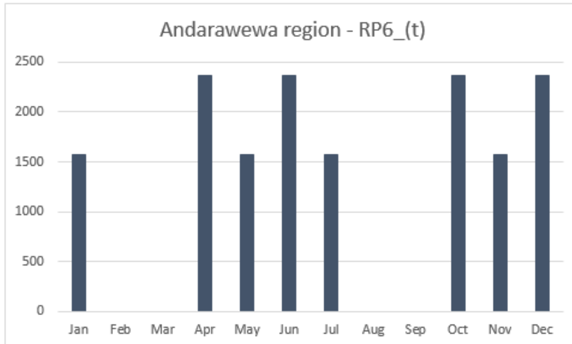


Figure 4.28: Release water to the Andarawewa area

A land area of 3741.6ha is used for agricultural purposes around the Andarawewa area. Figure No 4.28 shows the amount of water to be released to the Kiriibbanwewa area. It is noted that the water supply has decresed during the few months when paddy cultivation is not supply out. The results obtained by the model is shown in the Table No 4.20

| **Month** | **Region no 06 RP6(t)** |
|---|---|
| January | 1578.13 |
| February | 0 |
| March | 0 |
| April | 2367.19 |
| May | 1578.13 |
| June | 2367.19 |
| July | 1578.13 |
| August | 0 |
| September | 0 |
| October | 2367.2 |
| November | 1578.13 |
| December | 2367.2 |

Table 4.20: Release for the Andarawewa region

#### 4.2.2.4. Proportion of land allocation for paddy and other crops

The main objective of the study is to optimally use the released water from the reservoirs by finding out the proportions of lands to be cultivated for each crop in the considered areas.

The proportion of land allocation for paddy and other crops in each of the six catchments are illustrated in Table No 4.21

| **Regions** | **Ratio for paddy (A$_i$)** | **Ratio for other crops (B$_i$)** | **Cultivated land ratio to the available land (P$_i$)** |
|---|---|---|---|
| Region 1 | 0.6 | 0.4 | 1 |
| Region 2 | 0.18 | 0.82 | 1 |
| Region 3 | 0.6 | 0.4 | 1 |
| Region 4 | 0.6 | 0.4 | 1 |
| Region 5 | 0.6 | 0.4 | 1 |
| Region 6 | 0.6 | 0.4 | 1 |

Table 4.21: Proportion of land allocation for paddy and other crops

| Regions | Total Area(ha) | paddy(ha) | other (ha) |
|---|---|---|---|
| Region 1 | 2219.2 | 1331.52 | 887.68 |
| Region 2 | 9908.33 | 1769.67 | 8138.67 |
| Region 3 | 2106.7 | 1264.02 | 842.68 |
| Region 4 | 439.5 | 263.7 | 175.8 |
| Region 5 | 4244.41 | 2546.65 | 1697.76 |
| Region 6 | 3741.6 | 2244.96 | 1496.64 |

Table 4.22: Areas of land allocation for paddy and other crops

Figure No 4.29 shows the graphical representation of the results shown in the Table No 4.22 which obtained by solving the Goal programming model.This figure the total area of land cultivated in 6 regions and proportion of land utilized by paddy and other crops.

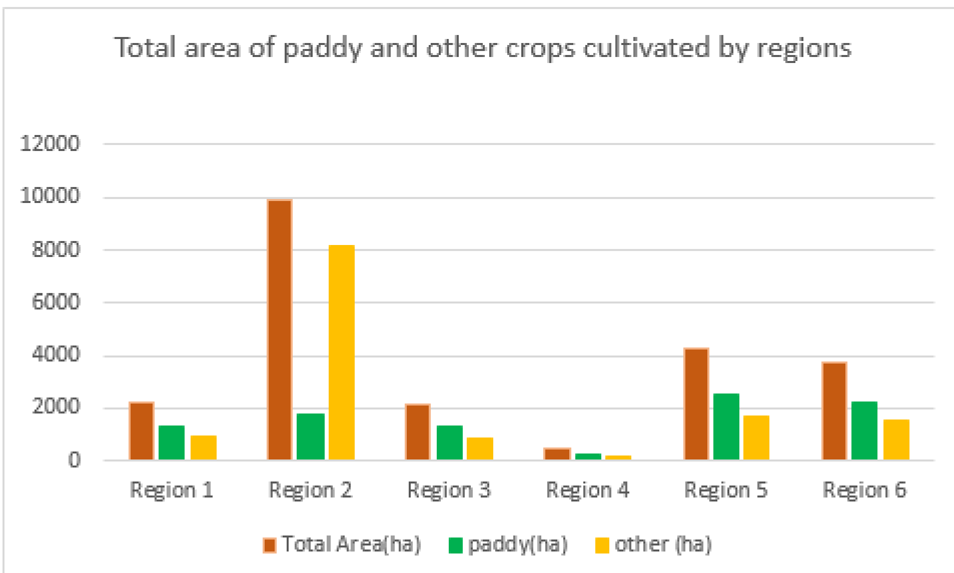


Figure 4.29: Release water to the cultivated areas

The following Figure No 4.30 shows land allocation for paddy cultivation and other crops cultivation in region 6.

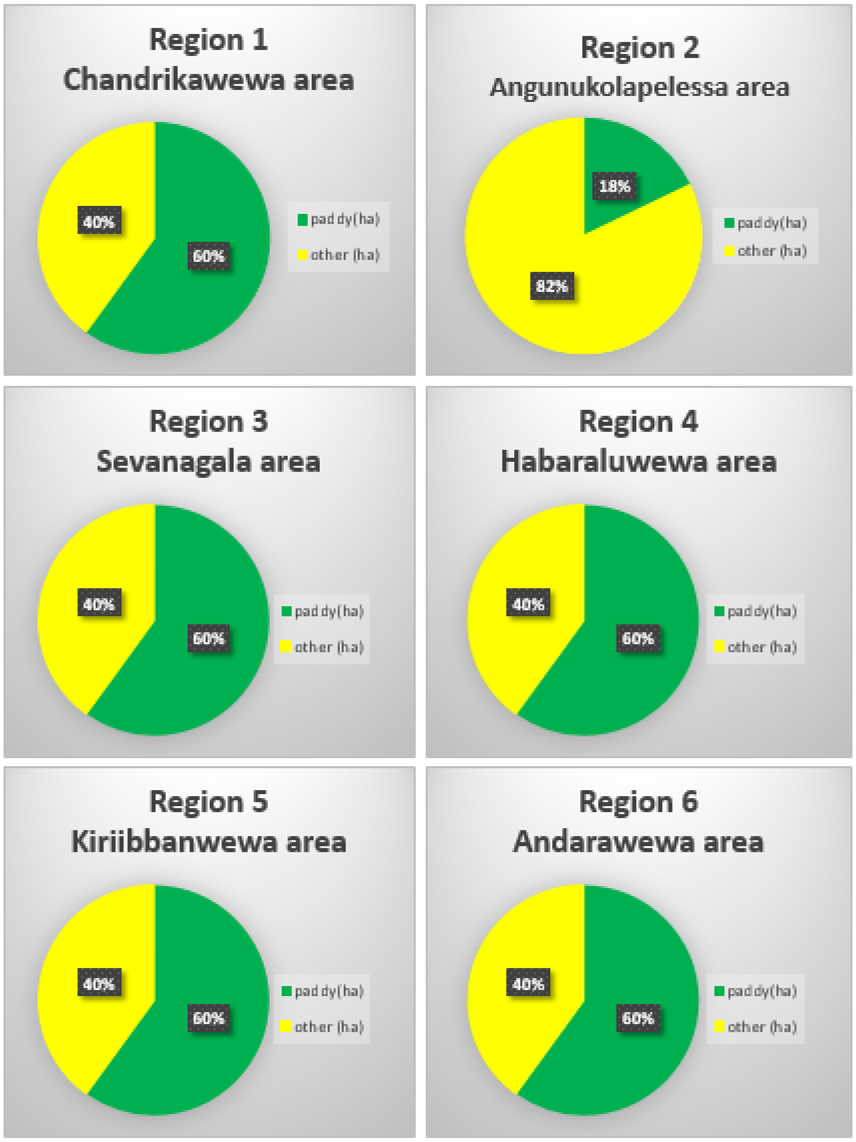


Figure 4.30: Land allocation optain by Linear programming technique

In order to optimize land and water use, the model determined that 5 of 6 regions should be used to grow paddy in 60% of their area, with the remaining 40% used to grow other crops. However, in Region 2, 82% of the area will be cultivated with other crops, while the remaining 18% will be cultivated with paddy.

### 4.2.3 Deviation Values for Goals

When obtaining the results by using Goal Programming technique, objective was minimized with the effect of two deviations.

#### First and Second Goal Deviations

Goal programming was incorporated to optimize the solutions obtained through linear programming. Apart from that, it was used to confirm the optimum status of other crops. The results of goal programming are summarized in the following Tables No 4.23 and 4.24. According to the results, all the deviations are equal to zero except the first goal negative deviation and the second goal positive deviation. Therefore to achieve the main goal an excess amount of other crops needs to be cultivated in region 2. The following tables shows the amount of deviations.

| **Paddy** | |
|---|---|
| Positive Deviation | Negative Deviation |
| $d1_2 = 0$ | $d1_1 = 0$ |
| $d2_2 = 0$ | $d2_1 = 0.4213$ |
| $d3_2 = 0$ | $d3_1 = 0$ |
| $d4_2 = 0$ | $d4_1 = 0$ |
| $d5_2 = 0$ | $d5_1 = 0$ |

Table 4.23: Deviation Values for First Goal

#### Second Goal Deviations

| **Other Crops** | |
|---|---|
| Positive Deviation | Negative Deviation |
| $c1_2 = 0$ | $c1_1 = 0$ |
| $c2_2 = 0.4213$ | $c2_1 = 0$ |
| $c3_2 = 0$ | $c3_1 = 0$ |
| $c4_2 = 0$ | $c4_1 = 0$ |
| $c5_2 = 0$ | $c5_1 = 0$ |

Table 4.24: Deviation Values for Second Goal

# Chapter 5

# Conclusion, Discussion and Future Directions

## 5.1 Conclusion

Failure to fulfill the water demand of the agricultural sector has been a problem for a longer time. Especially the areas which don't have enough natural sources of water have to face a lot of difficulties. Therefore, it is necessary to have a proper plan of releasing water from man-made sources in a scientific way. The main objective of this study was to develop a model which divres an optimal water release plan by meeting the agricultural demands. Mathematical techniques related to Operational research are being used regularly to optimize a lot of practical problems in day-to-day life including irrigation systems all over the world.

By using the input parameters such as inflow to the main reservoir, initial storage and demand of the agricultural areas, the model can be used to find the optimal release and the proportions of land that can be used to cultivate paddy and other crops. Results can be used to fulfill the water demand throughout the year by using water in an optimum manner as they were obtained according to the priorities, 60% of the farming land must be used to cultivate paddy, 40% of the farming land must be used to cultivate other crops and inflow values obtain through statistical forecasting have to be higher in number than the actual releasing inflow values. The highest water demand was observed in the Yala and Maha season, where the maximum amount of land is being used to do paddy cultivation.

## 5.2 Discussion

In this study a goal programming model was constructed to find the optimal water release as per the demand. What proportion of the considered each and every region(lands) needs to be utilized to cultivate different kinds of crops, mainly paddy vs. other crops and what is the optimum amount of water that needs to be released to each and every region are the main outcomes gained by solving the model. A cultivation plan for which proportions of land need to be used for paddy and other crops under the given requirements and the conditions was provided as results of the study. A case study of the Udawalawe irrigation scheme is considered for the research. Maximizing the allocation of available land for the cultivation of paddy and other crops was considered as the objective function and used other requirements and conditions to develop a Goal programming model in this study.

In order to find out water requirements for the paddy cultivation, two models were built as shown in section 3. The result obtained by the model 1 was not a practical solution because water allocation for the paddy cultivation was not enough to meet the requirements of the farmers. Therefore Model 2 was built by developing the model 1 by using the goal programming technique. To achieve the main objective, three goals were set as constraints during the development of model 1 into model 2. First one was that there must be a minimum 60% of the land needed to be utilized to grow paddy. Second one was that there cannot be land remaining without cultivating any of the crops. Third and the last one was all the amount of water which was released from the Udawalawa reservoir must be utilized for the cultivation.

As per the results, water can be released from the udawalawa reservoir to cultivate paddy in 60% of the land in 5 out of 6 regions. Therefore it is concluded that there is no sufficient amount of water in the udawalawa reservoir to feed 60% of paddy plantation in each and every region. As per the optimum solution obtained by the model, only 18% of the land in region 2 can be utilized to cultivate paddy and remaining part of the land can be utilized to grow other crops.

Since paddy is only cultivated in yala and maha seasons in the considered area, there was no water release to the land in non seasonal months from the reservoirs which are located in left and right banks of the udawalawa reservoir. Even Though farmers cultivate other crops in those non seasonal months, this study hasn't counted the amount of water required to cultivate other crops during that period and it is assumed that other sources such as rainfall, groundwater will be used in that particular period. But in seasonal months, the amount of water required from the reservoirs for the other crops has been counted in terms of a fraction of what paddy required. This is one of the drawbacks in the study due to lack of scientific data availability.

The only input data used in the study is the inflow received to the Udawalawa reservoir. Data of the amount of water received to the udawalawa reservoir from its upper area from 1996-2020 were taken from the Mahaweli Authority as a report. The predicted inflow for the year 2021 was predicted by using those data and Time series forecasting method. Since the basic forecasting method was not enough to obtain the amount of informative capture, an advanced forecast method was used to continue the study. Therefore seasoned ARIMA $\left(SARIMA(0,0,3)(0,0,2)_{[12]}\right)$ was picked up as the appropriate model.

## 5.3 Future Direction

The research will be further expanded by changing the constraints and assumptions. Water release to the agricultural lands will be considered not only in Yala and Maha seasons but also in all 12 months of the year. Other crops will be differentiated into few recognized crops and demand for them also will be considered when optimizing the amount of water released to each and every region throughout the year.

# Bibliography

[1] Chandrasiri, C, Dharmapriya, S, Jayawardana, J, Kulatunga, AK, Weerasinghe, AN, Aluwihare, CP & Hettiarachchi, D 2022, Mitigating Environmental Impact of Perishable Food Supply Chain by a Novel Configuration: Simulating Banana Supply Chain in Sri Lanka, Sustainability, vol. 14, no. 19, p. 12060.

[2] D. Ricketts, K, G. Turvey, C I. Gmez, M 2014, Value chain approaches to development, Journal of Agribusiness in Developing and Emerging Economies, vol. 4, no. 1, pp. 222.

[3] Fernando, S., Semasinghe, C., Jayathilake, N., Wijayamunie, R., Wickramasinghe, N. and Dissanayake, S. (2016), City region food system situational analysis, Colombo, Sri Lanka, available at: https://cgspace.cgiar.org/handle/10568/78818 (accessed 12 March 2017).

[4] Mayadunne, G Romeshun, K 2013, Estimation of Prevalence of Food Insecurity in Sri Lanka, Sri Lankan Journal of Applied Statistics, vol. 14, no. 1, p. 27.

[5] Jeong, H, Kim, H Jang, T 2016, Irrigation Water Quality Standards for Indirect Wastewater Reuse in Agriculture: A Contribution toward Sustainable Wastewater Reuse in South Korea, Water, vol. 8, no. 4, p. 169.

[6] Srinivasan, JT Reddy, VR 2009, Impact of irrigation water quality on human health: A case study in India, Ecological Economics, vol. 68, no. 11, pp. 2800 - 2807.

[7] Usman, MA Gerber, N 2019, Assessing the effect of irrigation on household water quality and health: A case study in rural Ethiopia, International Journal of Environmental Health Research, vol. 31, no. 4, pp. 433452.

[8] Willey, RG, Smith, DJ Duke, JH 1996, Modeling Water-Resource Systems for Water-Quality Management, Journal of Water Resources Planning and Management, vol. 122, no. 3, pp. 171179.

[9] Raju*, BCK, Gowda C, C B S, K 2020, Optimization of Reservoir Operation using Linear Programming, International Journal of Recent Technology and Engineering (IJRTE), vol. 8, no. 5, pp. 10281032.

[10] D.P. Locus, E.V.Beek; *Water resource systems planning and management*, Data-Fitting,Evolutionary ,and qualitative Modeling, Gewerbestrasse, Switzerland, **2017** , 137–200.

[11] S.Vedula, P.P Mujumdar; *Water resources systems:Modelling techniques and analysis*, **2005**, 235-240.

[12] Bhavann K.Ajudiya,S.M. Yadav,P.K.Mujumdar; **2018-June** *Optimization of water supply release using linear programming* , Journal of applied science and computations, Vol 5, Issues 6, p.251.

[13] Louks,D.P kindler,J.and Fedra,K, *Interactive water resources modeling and model use:an overview*, Water Resource Planning and Management, AGU,21(2) pp. 95-102,1985

[14] S. Busenberg, K. Cook; *Vertically Transmitted Diseases, Models and Dynamics, Biomathematics 23*, Springer-Verlag, Berlin, 1993.

[15] Hyndman,R.J, Athanasopoulos,G; *Forecasting: Principles and practice.OTexts*,2018

[16] Mathai,A.M,Haubold,H.J *Application to Stochastic Process and Time Series.Special Functions for Applied Scientists*,2008

[17] Janga Reddy, M Nagesh Kumar, D 2007, An efficient multi-objective optimization algorithm based on swarm intelligence for engineering design, Engineering Optimization, vol. 39, no. 1, pp. 4968.

[18] Fisher, B, Turner, RK Morling, P 2009, Defining and classifying ecosystem services for decision making, Ecological Economics, vol. 68, no. 3, pp. 643653.

[19] Cinzia Colapinto1Raja Jayaraman2Simone Marsiglio; *Multi-criteria decision analysis with goal programming in engineering, management and social sciences: a state-of-the art review.23*, .Manag Sci 2:138151, 1955

[20] Gabbrielli, E 2004, The use of modelling and reuse techniques in the development of water management systems in basins with limited water resources, Water Science and Technology, vol. 49, no. 7, pp. 3338.

[21] Fang, G, Guo, Y, Huang, X, Rutten, M Yuan, Y 2018, Correction: Fang, G., et al. Combining Grey Relational Analysis and a Bayesian Model Averaging Method

to Derive Monthly Optimal Operating Rules for a Hydropower Reservoir. Water 2018, 10, 1099, Water, vol. 10, no. 12, p. 1780.

[22] Mann, HB 1945, Nonparametric Tests Against Trend, Econometrica, vol. 13, no. 3, p. 245.

[23] Kaplan, EL Meier, P 1958, Nonparametric Estimation from Incomplete Observations, Journal of the American Statistical Association, vol. 53, no. 282, pp. 457 - 481.

[24] Izenman, A. J. (2008). Modern multivariate statistical techniques: Regression, classification and manifold learning. Springer

[25] Armstrong, J. S. (1978). Long-range forecasting: From crystal ball to computer. John Wiley Sons

[26] Hyndman, R. J., Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679688.

[27] Bickel, P. J., Doksum, K. A. (1981). An analysis of transformations revisited. Journal of the American Statistical Association, 76(374), 296311.

[28] Box, G. E. P., Cox, D. R. (1964). An analysis of transformations. Journal of the Royal Statistical Society. Series B, Statistical Methodology, 26(2), 211252.

[29] Guerrero, V. M. (1993). Time-series analysis supported by power transformations. Journal of Forecasting, 12(1), 3748.

[30] Wilson, GT 2016, Time Series Analysis: Forecasting and Control, 5th Edition, by George E. P.Box, Gwilym M.Jenkins, Gregory C.Reinsel and Greta M.Ljung, 2015. Published by John Wiley and Sons Inc., Hoboken, New Jersey, pp. 712. ISBN: 978-1-118-67502-1, Journal of Time Series Analysis, vol. 37, no. 5, pp. 709711.

[31] Dickey, DA Fuller, WA 1979, Distribution of the Estimators for Autoregressive Time Series With a Unit Root, Journal of the American Statistical Association, vol. 74, no. 366, p. 427.

[32] Kwiatkowski, D, Phillips, PCB, Schmidt, P Shin, Y 1992, Testing the null hypothesis of stationarity against the alternative of a unit root, Journal of Econometrics, vol. 54, no. 13, pp. 159178.

[33] Weber, E 1977, Kendall, M.: Multivariate Analysis. Charles Griffin b Co. LTD. London, High Wycombe 1975. 210 s., 9 Abb., 27 Tab., 1 Anhang, £ 6,80, Biometrical Journal, vol. 19, no. 4, pp. 309309.

[34] LJUNG, GM BOX, GEP 1978, On a measure of lack of fit in time series models, Biometrika, vol. 65, no. 2, pp. 297303.

[35] Akaike, H 1974, A new look at the statistical model identification, IEEE Transactions on Automatic Control, vol. 19, no. 6, pp. 716723.

[36] Schwarz, G 1978, Estimating the Dimension of a Model, The Annals of Statistics, vol. 6, no. 2.

[37] Hannan, EJ Quinn, BG 1979, The Determination of the Order of an Autoregression, Journal of the Royal Statistical Society: Series B (Methodological), vol. 41, no. 2, pp. 190195.