A Decision Support System For Sugarcane Irrigation Supply And Demand Management

Zubair Patel
Supervisor: Dr Jonas Stray
Co-Supervisor: Prof Theodor Stewart

Minor Dissertation Submission For Completion Of Masters Degree By Coursework And Dissertation

Submitted September 2016

Declaration

This dissertation is a presentation of my research work done for the completion of my Master Degree in Advanced Analytics and Decision Sciences at the University of Cape Town offered by the Statistical Science Department. Wherever contributions of others are involved, including research results, discussions, and recommendations, every effort is made to indicate this clearly by reference to the literature.

Abstract

Commercial sugarcane farming requires large quantities of water to be delivered to the fields. Ideal irrigation schedules are produced indicating how much water to be supplied to fields considering multiple objectives in the farming process. Software packages do not fully account for the fact that the ideal irrigation schedule may not be met due to limitations in the water distribution network. This dissertation proposes the use of mathematical modelling to better understand water supply and demand management on a commercial sugarcane farm. Due to the complex nature of water stress on sugarcane, non-linearities occur in the model. A piecewise linear approximation is used to handle the non-linearity in the water allocation model and is solved in a commercial optimisation software package. A test data set is first used to exercise and evaluate the model performance, then to illustrate the practical applicability of the model, a commercial sized data set is used and analysed.

Acknowledgements

This research was supported by the National Research Foundation of South Africa and the Harry Crossley Foundation. I thank the sponsors for the funds they made available ensuring peace of mind while I completed the Masters Degree.

I thank my dissertation supervisor Dr Jonas Stray from the Statistical Science Department at the University of Cape Town for his assistance in the model development and guiding me during the process of the completion of this document. I would also like to thank my dissertation co-supervisor Prof Theodor Stewart from the Statistical Science Department at the University of Cape Town for his invaluable comments and recommendations.

I would also like to express my gratitude towards the industry representatives from Illovo, Mr Brent Griffiths and Mr Bruce Meikle, that provided the data for this dissertation and answered our questions pertaining to the operations on the farm.

Contents

Chapter 1

Introduction

Global food production is facing pressure from an increasing human population with its need for fresh water. Free expenditure of available resources is no longer considered an option by responsible societies, and our planet carries—for all intents and purposes—a relatively fixed amount of fresh water in rivers and lakes ready to be utilised in agriculture, manufacturing, mining, transport, energy, to name just a few. Therefore it is not only desirable to prioritise certain activities in terms of water allocation but also to investigate how efficiently water that has already been allocated is distributed and utilised.

A commercial sugarcane farmer has approached South African universities providing research opportunities. These research opportunities come in light of communication with the industry representative revealing that there is a need for a holistic approach to understand the farm's operations and planning to improve the overall performance of the farm. The complexity of the operations at a commercial sugarcane farm level makes it difficult to study all of them in detail. This dissertation serves as a pilot to the possible future projects required to understand the entire sugarcane farm. Particularly, this dissertation will focus on the water distribution network's limitations by modelling the components to understand water supply and demand management.

This dissertation contains seven chapters. Chapter 1 introducing the problem and discusses the background information of the problem. Chapter 2 contains a literature survey on the theory and methodology required to understand and model the problem. Chapter 3 gives a formal description of the problem that will be used in the modelling process. Chapter 4 describes the mathematical formulation that will be used in the program and the program implementation. Chapter 5 is devoted to applying the model and analysing the results of the program. Chapter 6 and 7 collects the dissertation conclusions and future research possibilities.

1.1 Background and Brief Problem Description

Sugarcane is a species of the grass family, Poaceae, and is a perennial crop. The cultivation cycle of sugarcane begins with the planting of a single cultivar in a field that has been carefully prepared for growing seed cane. This seed cane is harvested just before the commencement of planting. The seed cane is cut into short pieces and placed in equidistant furrows at the chosen planting depth and are subsequently covered with soil. For four weeks, the newly planted field must be kept sufficiently wet by rainfall and/or the water distribution network for the seed cane to germinate, which essentially means that the water demand is equal to the evapotranspiration.

Depending on agroclimatic factors the crop is harvested after nine to twenty-four months. At this stage no replanting should be necessary, the crop will regerminate and grow into another crop. The water demand for the regerminating crop is the same as for the planted crop. A crop that originates directly from the seed cane is generally referred to as a plant crop while the subsequent crops are called ratoons. The water demand of a crop depends mainly on the cultivar, amount of sunlight, ambient temperature and the age of the crop (time elapsed since planting or last harvest). This water demand should be fairly well established for any large sugarcane farm and summarised using the concept of the crop factor. The crop factor is a real-valued constant that may be multiplied by known quantities to generate an estimated water demand for a field of a given age at a given time of the year. Generally, the crop factor is larger for older crops and warmer parts of the season. If the water demand cannot fully be met, yield losses are incurred while serious shortages may lead to a total loss of the crop.

In this dissertation, the water demand of a crop/field refers to the ideal water requirement for a crop/field to grow at a particular point in time or for a period of time. The water demand for every field for a period of time, such as year, is known as the irrigation schedule. The type of water distribution network that this dissertation focuses on is assumed to take water from some kind of large source of water such as a river, large canal, dam or lake. The water is fed through a network of pipes, canals and dams by means of mechanical pumps or gravity or both. The maximum water now allowed through any pipe, canal or pump is assumed to be a fixed value, as well as the maximum capacity of any dam. This water is then applied to a field via an irrigation system to satisfy the fields water demand. An irrigation system, such as a pivot sprinkler or furrow system, is a mechanism situated at the field to evenly distribute water across the field. The maximum flows and storage volumes present in the water distribution network, essentially means that the available water supply to any field will be limited.

An estate can be arranged into a hierarchy showing the dependencies among the components of the water distribution network and fields. Since the water demand is dependent on the time of year and the age of the crop, a sub-component of the water distribution network may at times be able to supply the water demand while unable at other times. If a large number of fields have reached maturity, or have just been harvested, at the same time, the water demand may easily exceed the water supply, resulting in an irrigation deficit. Some alleviation to such temporary shortages may be provided by a planned increase in dam levels, but the dams are also constrained by size and depth. However, due to the dependencies in the water distribution network, it is difficult to know the exact water supply available to each field. Additionally, the time of year can obviously influence water demand, exacerbating the problem during periods of low rainfall and/or high temperatures.

The need for incorporating the water distribution network's limitations has been noted as early as 1982 by Yaron & Dinar [60]. Software packages and decision support systems (DSS) that incorporate surrounding climate conditions and crop information produce accurate irrigation schedules and yield predictions are available to farmers. The software usually simulates a single crop or a homogeneous field's growth cycle to obtain its water demand. These software packages to a certain extent consider the limited water supply brought about by the limitations of a water distribution network.

1.2 Research Objectives

The objective of this dissertation is to improve irrigation water supply and water demand management using mathematical modelling, specifically by modelling the commercial sugarcane farmer's water distribution network. The mathematical model will be implemented in a programming language to simulate the water usage between the water distribution network and the fields. To obtain a better understanding of the objective, the following sub-objectives will be investigated.

1. Determine the severity of irrigation deficits by comparing a field's water demand and the actual water supplied to the field.
2. Identify points in time when the water distribution network may benefit from structural intervention to reduce the irrigation deficit.
3. Identify changes in the harvest schedule to reduce irrigation deficits.

1.3 Scope

This dissertation will assume that the harvest schedule is an input for the program, thus providing the irrigation schedule as well. In doing so, the analysis of the water distribution network becomes a water allocation problem or more generally, a supply and demand network. The intended use of the program is an aid for decision making. Mysiak *et al* [32] encourage the development of a malleable DSS. They discuss that a high number of failed DSS are due to the lack of handling real life unstructured problems. Thus, the developed model should have flexibility to be applied to more than one farm setting. The provided harvest schedule should be allowed to vary and the consequences on the water distribution network should be observable in the program results. The program should calculate the inflows and outflows of the water distribution network for a period of time and represent the results in an intuitive format. The program should be able to handle a large water distribution network (large farm) in a fair amount of computational time to be applicable for real time use [5]. The program should allow some flexibility such that a user may customise the setting to adopt to his/her farm layout. Numerous research finding have stressed the low absorption of models in water management such as in Sulis & Sechi [54]. Mysiak *et al* [32] also discuss the importance of developing a DSS with a friendly user interface for easy implementation. Due to time constraints, a user interface will not be created, and the program will be run in the programming language's environment.

1.4 Data

A commercial sugarcane farmer with farming estates in multiple countries has provided data on one of its farms. The commercial sugarcane farmer is a large-scale sugar producer that provides sugar locally and abroad. The commercial sugarcane farmer has provided a comprehensive description of the farms water distribution network and its service relations. This includes a schematic representation of the water distribution network, flow rates, and dam storage capacities. The data also includes climate history, harvest schedules, and crop factors. Since revealing the data provided may be detrimental to the commercial sugarcane farmer, the data will not be listed in its entirety. The author of this dissertation has signed the non-disclosure agreement requested by the commercial sugarcane farmer.

Chapter 2

Literature Review

Sections 2.1 to 2.2 surveys the literature to determine the consequences of not meeting the sugarcane water demand which will aid in achieving sub-objective 1. Section 2.3 briefly discusses irrigation systems. Section 2.4 surveys the available water allocation models in the literature to aid in achieving sub-objectives 2 and 3.

2.1 Sugarcane Water Demand

To apply a water allocation model in the context of a commercial sugarcane farm, the irrigation schedule is required. The irrigation schedule refers to the water demand for each field growing sugarcane to be supplied by the water distribution network for the harvest cycle or a particular period of interest. The granularity of the irrigation schedule may be at the monthly, weekly or daily level. The commercial sugarcane farmer has not explicitly provided the irrigation schedule but rather a data set with measurements that together with mathematical relationships, the sugarcane water demand can be calculated.

A crop's water demand is described by the Food and Agricultural Organisation (FAO) as the depth of water needed to meet the water loss through evapotranspiration (ET). Evapotranspiration is the simultaneous loss of water through evaporation and plant transpiration. It is the amount of water required by the crop to grow optimally, i.e. the amount of water required for the crop to reach its full production capacity. Theoretical relationships quantify the positive relationship between the amount of evapotranspiration and cane yield. Thompson [55] found the following relationship between cane yield and evapotranspiration:

$$Y_c = 0.969 \times \text{Total}(ET) - 2.5 \tag{2.1.1}$$

where Y_c is the cane yield in tons per hectare (t.ha^{-1}) and Total(ET) the total evapotranspiration in centimetres per hectare (cm.ha^{-1}) for the duration of the crops life cycle. Shih [48] used linear regression to relate the amount of evapotranspiration to cane yield, sugar yield and dry biomass. The cane yield and evapotranspiration relationship found in Florida was

$$Y_c = 1985 \times \text{Total}(ET) - 118574 \tag{2.1.2}$$

where Y_c is measured in kilograms per hectare (kg.ha^{-1}) and Total(ET) the total evapotranspiration measured in centimetres per hectare (cm.ha^{-1}) for the duration of the crops life cycle. Both equations illustrate that increased evapotranspiration results in greater cane yield. The latter equation, which was developed at a later stage, has a wider spread. The former could be a result of similar data points when used to estimate the coefficients.

Since evapotranspiration directly effects cane yield, evapotranspiration prediction is a key component in designing crop irrigation schedules. Most irrigation scheduling software rely on accurate crop evapotranspiration prediction by simulating climate, plant growth and soil moisture content. Evapotranspiration, usually measured in millimetres per day (mm.day^{-1}), can be directly measured with an apparatus known as a lysimeter. These lysimeters are expensive and can only be used for small areas [15]. Additionally, they do not provide the ability to forecast evapotranspiration if the grower is interested in developing an irrigation schedule. Two methods are available for calculating crop evapotranspiration: (1) direct crop evapotranspiration and (2) reference crop factor based calculations. Direct crop evapotranspiration calculations using theoretical relationships provide a sophisticated method so that for any given climatic condition, the crop evapotranspiration can be calculated. Allen *et al* [2] provides a guideline for computing crop evapotranspiration based on the Penman-Montieth equation

$$LE = \lambda ET = \frac{\Delta(R_n - G) + \rho_a c_p \left(\frac{e_s - e_a}{r_a}\right)}{\Delta + \gamma\left(1 + \frac{r_s}{r_a}\right)} \tag{2.1.3}$$

where $LE = \lambda ET$ is the latent heat flux (MJ.m^{-2}.day^{-1}), λ the latent heat of vaporisation (MJ.kg^{-1}), R_n is the net radiation (MJ.m^{-2}.day^{-1}), G the soil heat flux (MJ.m^{-2}.day^{-1}), e_s the saturation vapour pressure (kPa), e_a the actual vapaour pressure (kPa), ρ_a the mean air density (kg.m^{-3}), c_p the specific heat of the air (°C), Δ the slope of the saturation vapour pressure temperature relationship (kPa.°C^{-1}), γ the psychrometric constant (kPa.°C^{-1}), r_s the surface resistance and r_a the aerodynamic resistance (s.m^{-1}). The equations to calculate the parameters in Equation (2.1.3) can be found in [2]. Equally well performing predictor equations such as the Eddy Covariance and Bowen Ratio Energy Balance are available as well, but suffer the same pitfalls [47].

The measurements required to compute the crop evapotranspiration using these equations are extensive and cumbersome to determine. They are different for each crop and need to be measured at multiple time instances. Additionally, the equipment is environmentally sensitive as described by Shi *et at* [47]. Modern techniques of estimating evapotranspiration include the use of satellite imagery to obtain a spatially accommodated evapotranspiration estimate. See [59, 3, 15] for detailed explanation on the methodology. The commercial sugarcane farmer has provided climate data, but their data are insufficient to calculate the evapotranspiration using Equation (2.1.3). Equation (2.1.3) however, will be more appropriate for developing a model to use in multiple settings since using provided weather forecasts produces more accurate prediction.

The provided data enables calculation of evapotranspiration based on the crop factor method. The crop factor method is a simpler method developed by Allen *et al* [2] where the amount of measurements is reduced by introducing a reference crop, which together with its reference evapotranspiration and the crop factor, the crop evapotranspiration can be determined. The reference crop evapotranspiration is described by the FAO [17] as, *"ET_0 is the rate of evapotranspiration from a large area, covered by green grass, 8 to 15 cm tall, which grows actively, completely shades the ground and is not short of water"*. The interested crop evapotranspiration is then given by

$$ET = K \times ET_0 \tag{2.1.4}$$

where K is a specific crop factor dependent on the crop type, age and climate and ET_0 is dependent on climate conditions. Crop factors are provided by the FAO for various crops growing in different climatic conditions. Watanabe *et al* [56] encourages that crop factors be determined locally for more accurate estimation. However, they found that locally determined factors were not much different than the factors provided by the FAO in sub-humid climates. Inman-Bamber *et al* [23] in 2003 confirmed the FAO initial and mid stages of the sugarcane crop factors but not the end stage. Equation (2.1.4) depends on an evapotranspiration measurement. The reference crop evapotranspiration can be measured using the lysimeter as well, but a cheaper reliable alternative using evaporation pans is available. There are different designs of pans, such as the Class A Evaporation Pan and the various versions of the Sunken Pan. The Class A Evaporation Pan appears to be the popular pan of choice to measure the reference crop evapotranspiration. It was also used to determine the reference crop evapotranspiration measurements in the provided data. The pan is essentially a large-based low-walled cylinder (petri pan shaped). Determining the reference crop evapotranspiration involves measuring the rate of evaporation in the pan and using the pan factor (k_{pan}) to calculate the reference crop evapotranspiration, i.e.

$$ET_0 = k_{pan} E_{pan} \tag{2.1.5}$$

where E_{pan} is the evaporation measured from the pan. The pan factors depend on climate and is obtained in a table from the pan manufacturer or from theoretical relationships

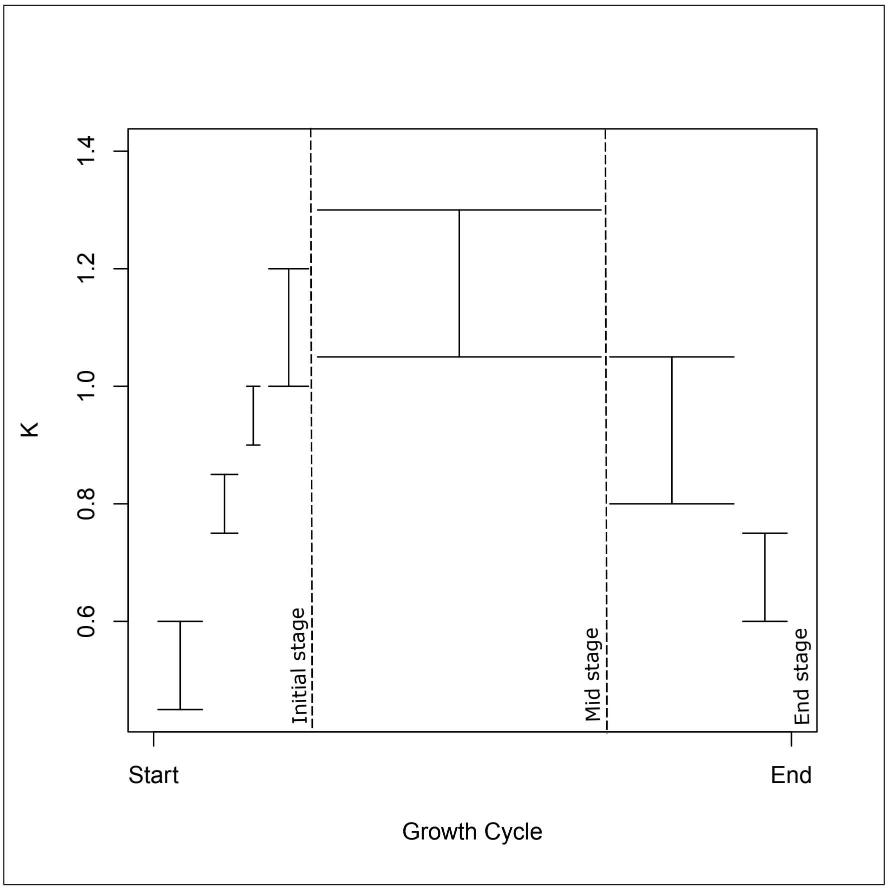


Figure 2.1.1: The possible crop factors over a sugarcane's life-cycle [18].

[19, 11, 51]. Figure 2.1.1 plots the possible crop factor range for the sugarcane life-cycle. The trend is approximately quadratic, indicating that the crop's water demand is low during the initial stage, high during the mid stage and low during the end stage relative to the reference crop.

Sugarcane water demand determination relies on accurate evapotranspiration determination. Various techniques presented in the literature provides a method for calculating evapotranspiration. It has also been shown that these techniques all produce coherent results. The evopotranspiration calculation depends on many factors and becomes complex when trying to determine from a purely theoretical approach such as Equation (2.1.3). The number of measurements are too much and expensive to obtain. Using the crop factor method, many of these measurements are summarised by the reference crop evapotranspiration and the specific crop factor. These factors take into account the weather, crop type, soil type etc. since they were determined on site. Given these values by the commercial sugarcane farmer, the sugarcane water demand calculation becomes easy. However, only averages have been provided for the entire farm and some accuracy will be lost.

2.2 Effects of Irrigation Deficits on Sugarcane

Growers have noticed that in certain situations, subjecting sugarcane to water stress (when a crop for some reason does not receive its water demand) has a beneficial effect on the economic gains of sugarcane. Equations (2.1.2) and (2.1.1) do not convey this and do not account for severe irrigation deficits during the crops lifecycle. Together with increasing water restrictions and the possibility of increased yields from water stressed sugarcane, research has gone into understanding the effects thereof. The effect of water stressed sugarcane needs to be quantified in some way to be used in the mathematical model.

Research conducted as early as 1985 by Ellis *et al* [16] found great economic benefit from subjecting sugarcane to water stress to reduce production costs. Their revised irrigation schedule produced a 20% saving on water supply. The effect of water stressed sugarcane in the literature can be summarised in its three stages, namely, (1) germination and tillering stage (initial), (2) stalk growth stage (mid), and (3) maturation stage (end).

1. **Germination and tillering**, a period of time approximately between 60 and 180 days long. In the early study of Ellis *et al* [16] conducted over the harvest seasons 1981-1983, irrigation was applied directly after harvest and then withheld. They report that in the most severe case where water was withheld for 10 weeks, the yield difference compared to the well-watered crop was insignificant. In a similar more detailed study by Robertson *et al* [42] experiments were conducted on the effect of water stressed sugarcane under different conditions. In two of the experiments sugarcane was subjected to water stress in the early season (germination and tillering stage) of a crop cycle under cooler and warmer conditions using the same cultivar. In the former experiment, no significant difference was found in the total biomass, sucrose yield or sucrose quality (concentration). In the latter experiment, total biomass and sucrose yield was reduced by 20% and 22% respectively compared to the well-watered field. However, the sucrose quality was unaffected by the water stress. This illustrated that caution is required when depriving crops of irrigation during periods of high temperature. These results corresponded to the experiments and findings conducted by Inman-Bamber [24] and Inman-Bamber & Smith [25]. Thus, there appears to be potential water savings and economic gains in subjecting sugarcane to water stress at the early stage without accumulating too many negative effects.

2. **Stalk growth**, a period approximately between 170 and 400 days long. Robertson *et al* [42] showed that in two experiments the effect of mid season (stalk growth stage) water stress during two different cycles, one stressed in cooler and one in warmer conditions, had deleterious effect in yields on both occasions. In the former experiment, a reported 26% and 37% decrease in total biomass and sucrose yield was found at

final harvest compared to well-watered fields respectively. No significant difference was found in the sucrose quality. In the latter experiment, a reported 32% and 43% decrease in total biomass and sucrose yield was found at final harvest compared to the well-watered field respectively. However, in this case, sucrose quality was significantly lower relative to the well-watered field. Subjecting sugarcane to water stress in this stage had a more severe effect in terms of yields.

3. **Maturation**, a period approximately between 50 and 70 days long. During the maturation stage, it is widely accepted to perform the practice known as "drying-off" to increase sucrose yield. The drying-off period may be reduced irrigation or complete termination towards harvesting. Ellis *et al* [16], in an experiment, ceased irrigation completely for three months without reducing the sucrose yield. Robertson & Donaldson [37] showed that in a variety of drying-off experiments the reduced irrigation usually increased or left the sucrose yield unchanged. Robertson *et al* [43] recommended that maximum economic benefit for drying-off is achieved when 4% to 8% of cane yield is reduced relative to a well-watered crop. Inman-Bamber [24] and Inman-Bamber & Smith [25] recommended a drying-off period that would result in the crop having three to four leaves less than the well-watered crop. Donaldson & Bezuidenhout [13] reports that local growers would schedule drying-off that would allow for twice the water holding capacity of the soil to dissipate from the evaporation pan. Donaldson & Bezuidenhout [13] verify this rule and provide guidelines for adjustments in different climates. According to Olivier *et al* [34], these adjustments are necessary in low water capacity soil to achieve positive benefits. Although drying-off results in small increases (maximum of 14% by Robertson & Donaldson [37]; maximum of 21% by Hagos *et al* [21] at 65 days drying with a 15 month aged crop) in sucrose yield, there is economic benefit from reduced or ceased irrigation from drying-off [37].

The three summaries indicate that there is opportunity for controlled irrigation deficit without a severe effect on the crop yield. The summaries point to at least a ranking among the three stages in water stressed sugarcane. The stalk growth stage can be classified as the most negatively affected by water stress, followed by the germination and tillering stage and lastly the maturation stage. Kirda [28] discusses the importance of knowing crop yield response to water stress since cultivar, soil, and climate plays a critical role in accurately quantifying tolerance. The effect of water stress on the different cultivars can be found in [50, 44, 26, 49]. The provided data set does not include the cultivars used on each field or the crop yield history so that quantifying the effect of water stress for the commercial sugarcane farmer is not possible. However, some general guidelines do exist for quantifying the irrigation deficit, and these guidelines may be used in the water allocation model. According to Kirda [28], sugarcane under a 25% irrigation deficit during the tillering stage has an expected cane yield of 90% of the well-watered crop. Berhe [6] classified irrigation deficits in three categories: a severe deficit at a 50% irrigation deficit, a mild deficit at a 30% irrigation deficit and a very mild deficit at a 10% irrigation deficit. Thus, despite the limited data on the sugarcanes' yield responses to irrigation deficits, it is possible to prioritise limited water supply from the water distribution network.

2.3 Irrigation Systems

There are three irrigation systems used in sugarcane water distribution networks: furrow, centre pivot and drip (trickle or micro) irrigation system. Furrow irrigation is the oldest irrigation system and as the name suggest, furrows are dug along and within the fields to be irrigated. The entry of water is required to be positioned on a slope such that together with gravity and the directed furrows, the water is able to reach the entire field. These furrows may flow into other fields to obtain a larger area of application. Overhead irrigation refers to irrigation systems situated above the crop. Within the overhead system designs, there are a few specific designs which include sprinklers (drag-lines), centre-pivots and rain guns. Rain guns have largely been replaced by centre-pivot and sprinklers due to the potential damaging large droplets and high pressure usage (Carr and Knox [10]). Drip irrigation systems applies drops of water directly above the crops root system from a network of small piping. The amount of piping is vast in comparison to the other systems since there is a pipe present at each crop row.

The differences among the irrigation systems are capital start-up costs, maintenance costs, simplicity, robustness, and efficiency. For this dissertation, only the irrigation efficiency will be discussed. An irrigation systems efficiency is defined as the ratio of the absorbed water to the applied water. Qureshi *et al* [35] reports potential irrigation efficiency as 30% to 90% for furrow irrigation, more than 90% for centre pivot, and 50% to 80% for drip irrigation. Furrow irrigation is wasteful. Large drainage and runoffs are present in furrow irrigation and it is difficult to adjust the furrows once crops are planted [16]. Hill & Keller [22] showed that cane yield increases when moving from furrow to sprinkle, and from sprinkle to drip irrigation. The increase in cane yield from sprinkler to drip irrigation is confirmed in an analysis conducted in 2003 by Merry [30] which resulted in a 15% increased sucrose yield and a 22% water cost reduction. It is evident that increased yields can be obtained from irrigation systems that distribute water more uniformly.

The irrigation system used for each field was not provided in the data set (Appendix A.2). However, an average irrigation efficiency ratio was provided. The irrigation efficiency is crucial when determining how much water the water distribution network should deliver to the field. Since the irrigation systems are not efficient, a field's water demand usually has to be adjusted to account for the inefficiency.

2.4 Water Allocation Models

A water allocation problem can be described as the allocation of limited water among competing users. Each user has a purpose for the water and the consumption leads to a benefit. Since each user may have a different purpose and different benefit from the consumption, determining the optimal or a near optimal allocation can be laborious using heuristics or exhaustive searches when the number of users are vast. Additionally, there may be constraints or conflicting objectives further complicating the allocation. Constraints refers to physical restrictions that the allocation has to adhere to. A water allocation model is appropriate to model the irrigation of the sugarcane fields since the process is an allocation of limited water among the fields supplied by the water distribution network.

The water allocation problem is a special case of the resource allocation problems belonging to the family of network flow problems. To solve the water allocation problem, a water allocation model is needed. In the model, benefits and constraints are contained in the form of mathematical equations to be implemented in a simulation or optimisation technique. Simulation-driven water allocation models are algorithmic in their nature. This means that the model is solved by sequential time steps adhering to a sequence of rules. Optimisation-driven water allocation models are solved by optimising a real function (the objective function) using an underlying mathematical search technique. The two techniques are similar in that they both "simulate" the system of interest [57]. While each technique has its advantages and disadvantages, the context of the water allocation problem will indicate the most suitable technique [29].

An advantage of a simulation model is the ability to simulate management decisions very accurately [29] and thus evaluate them accurately. Simulations are created by modelling the process or system of interest and incorporating the management decisions as a sequential set of rules. This type of modelling belongs to the class of agent based modelling (ABM). The simulation becomes easy to implement because the rules are derived from intuitive human decision making processes or checks. These decision making processes are captured by the simulation's logical statements such as the "if-then-else" statements to achieve the desired performance. This implementation of rules does however require the programmer to handle all instances that can arise and can thus become coding intensive [29]. The realism achieved by simulation models has made them applicable to water allocation problems in practical settings.

Wurbs [58] evaluated river/reservoir practice using the Water Rights Analysis Package (WRAP) simulation software for water resource planning and management activities. The model allocated limited water to competing users by prioritising water demands and allocating the water sequentially through time. The simulation's flexibility allowed for practical application in water management and was required to evaluate water permits before implementation. The applicability comes from long collaborative development with relevant professional authorities, indicating the importance of industry input. Prioritising the allocation for competing water use is an important trait that will be required in this model of the water distribution network. As discussed in Section 2.2, the different behaviours of the sugarcane crop subjected to water stress indicates that a ranking system will be appropriate for the prioritising the fields water demand. Berhe *et al* [6] evaluated a simulation model where in each time step a network flow optimisation was performed. This eliminated the explicit rule coding for each time step. Sequential time steps that only consider the current time point may not be appropriated since storage facilities are present. It would be desirable to have a model optimally decide when to fill, hold or release the storage.

Sulis & Sechi [54] reviewed strict simulation based models and optimisation based models and found that strict simulation based models provided equal or better results compared to those of optimisation based models. An important concept central to these models was the linking of water flow using continuity equations (also known as mass balance, water balance, flow continuity or function balance) coded into the model, i.e.

$$\sum_i \text{inflows}_{p,i} = \sum_j \text{outflows}_{p,j} \tag{2.4.1}$$

The equation states that all inflows at the pump station p must leave the pump station. If storage is available, the equation is easily modified to accommodate this property. The links are mapped out as a network of arcs and nodes. The nodes are water users, water supply, reservoirs or pump stations and the arcs are rivers, pipes, canals. The continuity equations and system constraints are responsible for transferring the limitations throughout the water distribution network. The constraints must be satisfied at each continuity equation. This mapping can accurately incorporate the water distribution network present on the commercial sugarcane farm.

Optimisation techniques are different to simulation techniques. They do not seek to simulate management decisions, but can incorporate them to an extent and the results may coincide with management decisions. Optimisation models seek the best solution by optimising a mathematically defined objective function. The solution may not always be practical since coding complex rules into an optimisation model is difficult [29]. Optimisation techniques have been widely used in water allocation models. The underlying mathematical search routine differs among optimisation techniques. Two classes of optimisation techniques are evident in water allocation problem solving literature, (1) exact optimisation and (2) heuristic search based.

Bras & Cordova [8] used stochastic dynamic programing (DP) to solve an intra-seasonal water allocation problem for a single crop. The model was used to capture the stochastic nature of the soil water content. Yaron & Dinar [60] used a systems analysis approach containing a linear programming (LP) and dynamic programming component in a two subsystem program to find optimal water allocation for farming irrigation. The LP model was used to maximise farm income by finding the most profitable farm activities and then using the shadow prices from the LP model and the dynamic program to generate new irrigation schedules. Shangguan *et al* [46] used dynamic programming in a three-layered program for optimal regional water allocation for irrigation. The first layer optimises irrigation scheduling. The second layer optimises water allocation among crops and the last layer optimally allocates water among the regions. Dynamic programming optimisation techniques seek optimal solutions by sequentially computing the recurrence relation. These models are well suited for optimising multi-stage decision processes [14]. The nature of dynamic programming makes it particularly suitable for water allocation models where the decision of when to apply the limited water occurs at each time point. However, solving practical sized dynamic programming problems may be a challenging task since there are fewer general algorithms and software packages available compared to the well-established solvers available for linear programming [4].

Central to these models was the separation of optimisation tasks. Separating the optimisation tasks reduces the computational time and power requirement. In the case of the commercial sugarcane farmer, it would be desirable to develop the irrigation schedule together with the water distribution network's limitations. However, limited data is available to develop this type of model which requires soil water content modelling and is beyond the scope of this dissertation. The focus of water allocation for irrigation in the farming literature has been on regional or collective water allocation rather than at the multiple field level. Regions may be considered fields to model the water distribution network with the appropriate upscaling.

Reca *et al* [38, 39] used a multi-stage optimisation scheme to first determine the crop irrigation schedule and then to optimise the water allocation to the different sectors of the Bembézar Irrigation District in Spain. They motivate the need to incorporate non-linearity of economic returns and opted for a piecewise linear approximated objective function so that LP solvers could be used. Section 2.2 is indicative that water stressed sugarcane does not produce linear losses so that incorporating non-linearity in the water allocation model would be desirable. Using a piecewise linear approximation function to capture the effect of water stress on sugarcane may be appropriate and will allow the use of LP solvers. Schlüter *et at* [45] applied a water allocation model in a water management context to optimally distribute water and maintain water quality for the users along the Amudarya River basin. The optimisation used multiple objective function construction with a simple weighted sum of the objectives to drive the optimisation. For this dissertation, multiple objectives may not be appropriate for the water distribution network model since the objective is only to minimise the irrigation deficit or maximise the irrigation application. However, once more operational aspects of the commercial sugarcane farm are considered, multiple objective optimisation becomes desirable.

Non-linearities pose potential problems for certain optimisation solvers. Although these problems can be sometimes easily mathematically formulated, the required computation time and power can grow drastically. Too large or complex models with non-linearities have been handled by recently new techniques such as heuristic and metaheuristic algorithms. These algorithms have been developed with no underlying mathematical assumptions and seek good solutions without guaranteeing optimality. There has been a shift in the research community towards heuristic based techniques due to problem complexities [5]. Two classes of metaheuristics are trajectory and population based techniques. Trajectory techniques seek to continuously improve an initial solution whereas population based techniques seek to improve a population of solutions. Cunha & Sousa [12] used simulated annealing to obtain an optimised looped-water distribution network. Cai *et al* [9] used a combination of a genetic algorithm (GA) and linear programming to solve a water management problem. The linear programming models were implemented at each stage of the genetic algorithm. Noory *et al* [33] evaluated a particle swarm optimisation algorithm for irrigation water allocation in multi-crop planning and obtained good results. Karterakis *et al* [27] compared linear programming to a differential evolution (DE) algorithm and found that the solutions were very similar. However, the computation time of the DE algorithm took considerably longer to achieve a good solution (75 hours). Azamathulla *et al* [5] compared a genetic algorithm and linear programming model in reservoir management for crop irrigation and found that both models produced practical results.

Water allocation models for the farming industry can be found in the literature, but not directly for sugarcane irrigation modelling at the multiple field level. From the literature it is evident that practical results have been achieved using either simulation or optimisation techniques. Simulations have been useful to evaluate management decisions before implementation. Optimisation techniques have been useful in proposing new irrigation schedules and water allocations. Optimisation techniques may be more appropriate due to limited management practice knowledge. The LP modelling technique is able to capture non-linearities in a system using integer variables. Together with the findings in Section 2.2, the effect of water stress can be incorporated using integer variables. The constraints in LP, such as continuity equations, makes it easy to capture the dependencies in a network of flows. The optimisation process determines the best flow capacities and storage releases to mitigate bottlenecks and under-worked components in the water distribution network. This is a key desirable trait of optimisation methods that will be used in this dissertation. Using linear programming as a technique to solve the water allocation will allow the best usage of limited water to be observed according to an objective function.

Chapter 3

Formal Problem Description

This chapter describes the commercial sugarcane farmers equipment used in operations to formulate the mathematical model. The commercial sugarcane farmer has more than 500 fields occupying more than 10,000 hectares of land with an average field spanning 20 hectares. The commercial sugarcane farmer requires large quantities of water to irrigate the fields. The continuous placing of the fields further distances the fields from water sources so that long pipes, canals or furrows are required to transport the water to the fields. Additionally, powerful electric pump stations are needed to drive the water in the pipes, canals or furrows to the irrigation systems.

On the commercial sugarcane farm, a river serves as one large water source. A main pump station situated at the water source is solely responsible for extracting water from the source and delivering it to the subsequent components of the water distribution network. The irrigation system, pipes, canals, furrows, pump station, storage facilities are collectively known as the water distribution network. The extracted water which leaves the main pump station via pipes, canals or furrows is delivered to subsequent pump stations or to storage facilities. The storage facilities are situated at pump stations or along the pipes, canals or furrows. The irrigation systems at the fields are fed by closely situated pump stations, storage facilities or by the network of pipes, canals or furrows which passes alongside the fields. Pump stations act as relay units when handling the water; the inflow at a pump station has to leave the pump station or be stored at its storage facility. This means that a pump station may not operate at its capacity if preceding pump stations are not operating at a large enough capacity. This property is known as the flow continuity or water balance, and limitations in the water distribution network components are transmitted by this property.

Farmers are able to control (1) the amount of water pumped at each pump station, (2) the amount of water stored at the storage facility, and (3) the amount of water applied to each field. The annual date at which a field is cut is known as the harvest date. The collection of harvest dates for all the fields of the farm is known as the harvest schedule.

If the harvest schedule is known, the ideal irrigation schedule can be produced using the crop factors. This is the ideal amount of water the field should receive, i.e. its water demand. However, due to the limited water supply the ideal water demand may not be met. The complex water distribution network makes it difficult to know the exact amount of water available to the fields during the year. A field whose water demand is not met experiences an irrigation deficit. The harvest schedule is designed with some consideration of this limited water supply so the farmer actively avoids similar harvesting dates for fields who share a feeding source. The harvest schedule has other factors to consider as well so that constructing a harvest schedule to accommodate only the water distribution network's water availability is impractical. Two such factors are the crop yields and the harvest fronts. These factors are in conflict with a harvest schedule that only accommodates the water distribution network's limitations. These factors exert strain on the water distribution network by favouring close harvest dates for neighbouring fields who share a pump station, therefore increasing the likelihood of overlapping high water demand periods. These two factors may be summarised as follows.

- Crop yield: Commercial sugarcane farmers aim to maximise the yields of their grown crops. The commercial sugarcane farmer revealed that the minimum time taken for typical sugarcane to grow to a point of worthy harvest was approximately 39 weeks and that a field is harvested once a year. This age requirement adds strain on the water distribution network because arbitrary harvest schedules cannot be used to accommodate the water distribution networks limitations.

- Harvest fronts: A harvest front is a fleet of heavy machinery and personnel which is responsible for cutting and collecting the ripe cane. Harvest dates need to be spread out to ease the workload on the harvest fronts. Operational costs of using the fleet are high. Since there is a limited number of fleets that service the farm and the distances between fields and mills are large, the harvest schedule also has to achieve a low travelling distance for the harvest front by assuring that harvested fields are situated close to each other. However, it is likely the case that neighbouring fields share a water source so that constructing a harvest schedule to accommodate harvest fleets will result in neighbouring fields with similar harvesting dates, adding strain on the water distribution network.

The simplest solution is to upgrade the water distribution network to handle the large demand, but a potential large capital cost calls for an attempt to first improve current practice. Azamathulla *et al* [5] notes that water management in reservoir systems are met with poor operation and management even upon the completion of a new projects. This indicates a need for evaluation first, rather than purchasing new equipment. A step in understanding the limitations of the water distribution network and evaluating the current system implementation, is to model the optimal implementation of the existing system, i.e. a model that would optimally assign limited water to achieve a maximum distribution of the limited water. It is ideal to construct a model that can optimally assign water, accommodates yield and harvest front optimisation, but such a task is out of the scope of this dissertation. A simpler model will be to isolate the components of the sugarcane farm and concentrate solely on the water distribution network. The following simplifying assumptions are made and limitations of the data noted to construct the model:

1. Time periods: The operational activities (irrigating, harvesting, and planting etc.) occur at the daily level, possibly hourly level. For this dissertation the time periods of interest will be restricted to weekly operations. The weekly level will be used because of computation and data limitations. The computation time will be shorter at the weekly level since fewer variables will be used. Shorter computation time is desirable, since in a practical setting, at any time during operational planning, it may be necessary to evaluate multiple harvest schedules. This may be tens or hundreds of harvest schedules. If the water allocation model is to be used as a sub-program, say, a genetic algorithm where computation times accumulate due to multiple runs, the computation time can increase exponentially. The data does not provide daily measurements but rather weekly and monthly recordings. Bras & Cordova [8] used eight-day time steps noting that irrigation cycles are usually seven to ten days.

2. Water demand: No distinction will be made between the different cultivars or seed/ratoon crop when using the crop factor method to determine the crop water demand. The data provides only an average weekly crop factor which loses the distinction between the cultivars and seed/ratoon crop. This makes the modelling simpler but the water requirement less accurate.

3. Water distribution network: To reduce the complexity of coding a general shaped network, the water distribution network will be simplified. The number of pump stations between the main pump station and a field will be restricted to be exactly two. Each pump station will be restricted to have exactly one storage facility. No field will have a storage facility. If there are exactly two pumps between the source and field, then the restriction is met. If there are strictly less than two pumps, a dummy pump station is added. If there are strictly more than two pump stations, pump stations are amalgamated to be reduced to two pump stations. Storage facilities along the network will be moved to the upstream pump station which services common fields and then amalgamated to one storage facility. The number of pump stations linked to the main pump stations may be arbitrary, and the number of pump stations

linking to these subsequent pump stations may be arbitrary as well, each with an arbitrary number of fields which they service.

4. Storage facilities: To simplify the modelling, the storage facilities will be unaffected by weather conditions. This means that rainfall does not increase storage facility levels and heat does not decrease the levels by evaporation.

5. Sugarcane yields and harvest fronts: Yield and harvest front data have not been provided by the commercial sugarcane farmer. The model will not incorporate the explicit quantification of sugarcane yield responses to irrigation deficits or the harvest front travel distance brought about by the harvest schedule. The harvest schedule will be taken as provided by the commercial sugarcane farmer. The effect of irrigation deficits will be incorporated in the model using the findings in Section 2.2.

Thus isolating the components of the commercial sugarcane farm and taking the harvest schedule as given, the model should optimally assign limited water to fields by adhering to the water distribution network's constraints and the simplifications mentioned above. The model should identify time periods in the system where there are irrigation deficits and which components of the water distribution network are responsible for the irrigation deficits. The acquisition of such a model can be used to evaluate changes in the system parameters such as flow capacities, storage capacities, harvest schedules or the behaviour during a dry season (low rainfall).

Chapter 4

Mathematical Problem and Program Description

This chapter formulates the mathematical model using the problem description in Chapter 3. The chapter describes the components and architecture of the mathematical model and program. Section 2.4 motivated the use of an optimisation method to solve the water allocation problem. Section 4.1 to 4.3 describes the mathematical model. Section 4.4 discusses the implementation of the mathematical model in a program and Section 4.5 demonstrates a numerical example.

4.1 Set Definitions

The network of flows can be represented as a directed acyclic graph. The pump stations and fields can be seen as the nodes and the pipes, canals and furrows the arcs of the graph. Let $\mathcal{F}$ denote the set of fields and let $\mathcal{P}$ be the set of pump stations. Using the simplified network of pump stations, storage facilities and pipes in Chapter 3, the graph can be divided into five levels:

- Level 1, the source of the large water supply.
- Level 2, the main pump station situated at the source of all water.
- Level 3, the cluster level. The set of pump stations that receive water from the main pump station. The cluster level pump stations are contained in the subset $\mathcal{C}$ of $\mathcal{P}$ ($\mathcal{C} \subset \mathcal{P}$).
- Level 4, the sub-cluster level. The set of pump stations that receive water from level 3. The sub-cluster level pump stations are contained in the subset $\mathcal{S}$ of $\mathcal{P}$ ($\mathcal{S} \subset \mathcal{P}$). Each pump station in the sub-cluster level is fed by a single pump station from the cluster level.

- Level 5, the field level. The fields receive water from level 4. Each field is fed by a single pump station from the sub-cluster level.

The set of fields belonging to a particular sub-cluster is denoted as $D(p)$ with $p \in \mathcal{S}$ and $D(p) \subset \mathcal{F}$. The set of sub-cluster pump stations belonging to a particular cluster pump station is denoted $E(p)$ with $p \in \mathcal{C}$ and $E(p) \subset \mathcal{S}$. A simple example of such a system is illustrated in Figure 4.1.1. In this example there are eight fields, four sub-cluster pump station and three cluster pump stations. The set contents in the example are as follows:

- The set of pump stations $\mathcal{P} = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8\}$
- The set of cluster and sub-cluster pump stations $\mathcal{C} = \{p_2, p_3, p_4\}$ and $\mathcal{S} = \{p_5, p_6, p_7, p_8\}$ respectively.
- The set of fields $\mathcal{F} = \{f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8\}$.
- The set of sub-cluster pump stations grouped by their particular cluster pump station $E(p_2) = \{p_5, p_6\}$, $E(p_3) = \{p_7\}$ and $E(p_4) = \{p_8\}$.
- The set of fields grouped by their particular sub-cluster pump stations $D(p_5) = \{f_1, f_2\}$, $D(p_6) = \{f_3\}$, $D(p_7) = \{f_4, f_5, f_6\}$ and $D(p_8) = \{f_7, f_8\}$

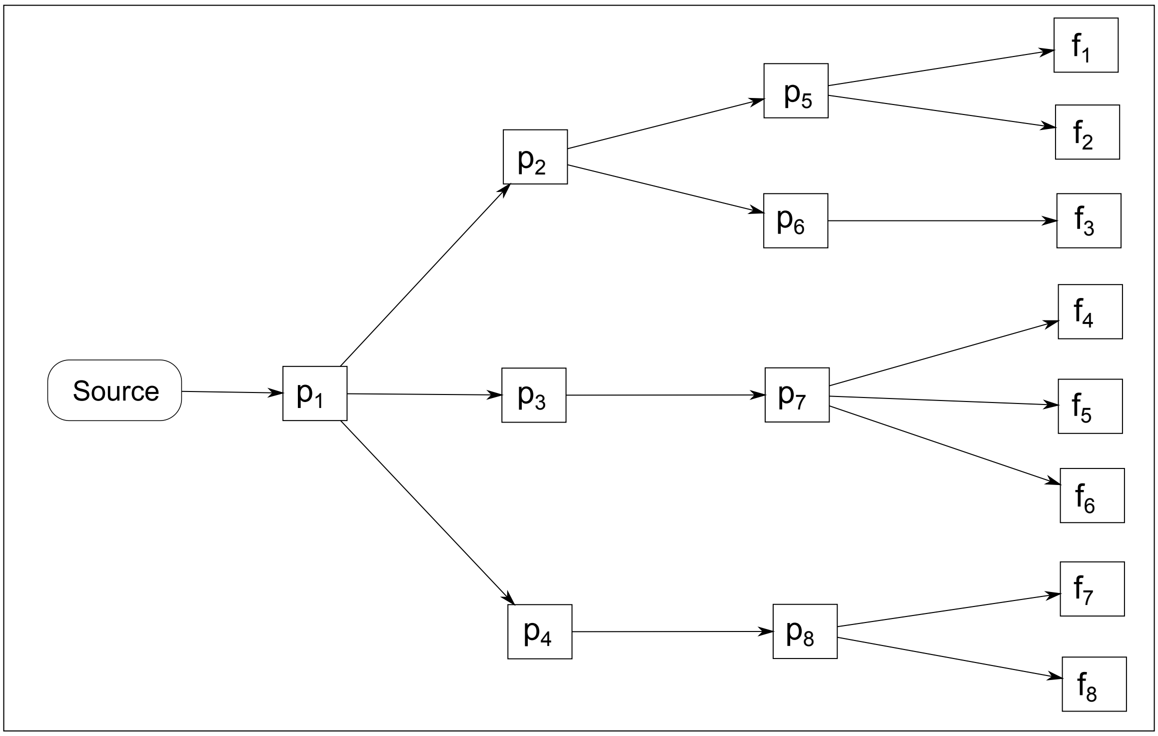


Figure 4.1.1: Eight-field water distribution network example.

Since the inflow at each pump station and field is unique, let the decision variable w_{pt} and w_{ft} be the inflow in cubic metres (m^3) at pump station p and field f during time t where $t \in T = \{t_1, t_2, ..., \}$ and $|T|$ is finite. The storage decision variable recording the amount of water in cubic metres (m^3) stored at pump station $p \in P$ at the end of time $t \in T$ will be denoted as s_{pt}.

4.2 Model Parameters and Hard Constraints

Each field $f \in \mathcal{F}$ has a time dependent water demand Ψ_{ft} to be supplied by the water distribution network at time $t \in T$. This quantity is determined using the equation

$$\Psi_{ft} = 0.001 \times \max\{0, (K_{ft}(ET_0)_t - r_t)\} \times A_f \times E^{-1} \tag{4.2.1}$$

where Ψ_{ft} is in cubic metres (m^3), K_{ft} is the crop factor for time period t and age of the crops at field f, $(ET_0)_t$ the reference crop evapotranspiration during time period t in millimetres (mm), r_t the predicted rain during time period t in millimetres (mm), A_f the field area in square metres (m^2) and E the average irrigation efficiency. The irrigation efficiency may be different at each field, but the data provides only an average quantity measured in the fraction of absorbed water to pumped water. The product $K_{ft}(ET_0)_t$ is the field water demand in millimetres before adjusting for the predicted rainfall. The maximum operator ensures that no water is to be supplied when the predicted rain exceeds the water demand.

The flow and storage decision variables have to adhere to the following constraints:

- Flow constraints: At any point in time there is a maximum inflow Γ_p of water that may enter a pump station, due to pipe and pump capacity.

$$w_{pt} \leq \Gamma_p \quad \text{for } p \in \mathcal{P} \text{ and } t \in \mathcal{T} \tag{4.2.2}$$

- Storage facility constraints: At any point in time the storage facility present at a pump station holds an amount of water less than the maximum water capacity Δ_p.

$$s_{pt} \leq \Delta_p \quad \text{for } p \in \mathcal{P} \text{ and } t \in \mathcal{T} \tag{4.2.3}$$

- Flow continuity constraints: These equations capture the balance of inflows and outflows at the pump stations. The flow continuity constraints are time dependent and the flow continuity constraint during the first period depends on the zero$^{\text{th}}$ time period. The zero$^{\text{th}}$ time period is used to capture the initial storage $s_{p,0}$. The flow continuity at the main pump station (indexed as pump station number 1 in example 4.1.1 and later referred to as "mc" without a subscript) is

$$s_{p_1,t-1} + w_{p_1,t} = s_{p_1,t} + \sum_{p \in \mathcal{C}} w_{pt} \quad \text{for } t \in \mathcal{T} \tag{4.2.4}$$

The cluster pump stations flow continuity equations are:

$$s_{p,t-1} + w_{pt} = s_{pt} + \sum_{p' \in E(p)} w_{p't} \quad \text{for } p \in \mathcal{C} \text{ and } t \in \mathcal{T} \tag{4.2.5}$$

The sub-cluster pump stations flow continuity equations are:

$$s_{p,t-1} + w_{pt} = s_{pt} + \sum_{f \in D(p)} w_{ft} \quad \text{for } p \in \mathcal{S} \text{ and } t \in \mathcal{T} \tag{4.2.6}$$

4.3 Objective Function

The purpose of the objective function is threefold:

1. Assign a strictly increasing value (benefit) for irrigating fields up to their water demand volume Ψ_{ft}.
2. Assign an increasing penalty for exceeding the water demand volume.
3. Assign priority to the fields that would suffer more adverse effects if under-irrigated.

The introduction of these characteristics requires a non-linear benefit function for each field's supplied irrigation volume at time period t, say

$$g_{ft}(w_{ft}) \quad \text{for } t \in \mathcal{T} \text{ and } f \in \mathcal{F} \tag{4.3.1}$$

At this point, the exact form and function values are not specified. Since the model does not take into account yields and harvesting fronts, we require only that the benefit function portray the relative differences and benefits as explained by points (1), (2) and (3) in the paragraph above. To use reliable and stable LP solvers, the non-linearity will be handled by a piecewise linear approximation, i.e. approximate the non-linear function by a sequence of linear functions. The piecewise linear approximation requires the original field inflow variable to be split into L variables corresponding to the number of linear functions that the non-linear function will be approximated by, i.e.

$$w_{ft} = w_{ft}^{(1)} + w_{ft}^{(2)} + ... + w_{ft}^{(L)} \quad \text{for } t \in \mathcal{T} \text{ and } f \in \mathcal{F} \tag{4.3.2}$$

with constraints on each segment

$$w_{ft}^{(l)} \leq \psi_{ft}^{(l)} \quad \text{for } l = 1, 2, ..., L \text{ and } t \in \mathcal{T} \text{ and } f \in \mathcal{F} \tag{4.3.3}$$

Constraint (4.3.3) corresponds to the maximum length a segment can attain. The idea is to begin utilizing the next segment if and only if the previous segment reaches its maximum value, i.e. in ascending order of superscript. To ensure that the segments' location increase in order of the superscript, the following additional constraints and binary variables are required for each field:

$$\begin{aligned}
\delta_{ft}^{(1)}\psi_{ft}^{(1)} &\leq w_{ft}^{(1)} \leq \psi_{f,t}^{(1)} \\
\delta_{ft}^{(2)}\psi_{ft}^{(2)} &\leq w_{ft}^{(2)} \leq \psi_{ft}^{(2)}\delta_{ft}^{(1)} \\
&\vdots \\
0 &\leq w_{ft}^{(L)} \leq \psi_{ft}^{(L)}\delta_{ft}^{(L-1)} \quad \text{for } t \in \mathcal{T} \text{ and } f \in \mathcal{F}
\end{aligned} \tag{4.3.4}$$

where $\delta_{ft}^{(l)} \in \{0, 1\}$. See [53] for a detailed explanation on piecewise linear approximations. The optimisation problem is now a mixed integer linear program (MILP). The downfall of this approach is that shadow prices (duality) becomes uninterpretable when binary or integer variables are included. The new benefit function $G_{ft}()$ for each field now depends on L variables subject to constraint sets (4.3.2), (4.3.3) and (4.3.4), i.e.

$$G_{ft}(w_{ft}^{(1)}, w_{ft}^{(2)}, ..., w_{ft}^{(L)}) = c_{ft}^{(1)}w_{ft}^{(1)} + c_{ft}^{(2)}w_{ft}^{(2)} + ... + c_{ft}^{(L)}w_{ft}^{(L)} \tag{4.3.5}$$

This means that each field $f \in \mathcal{F}$ at time $t \in \mathcal{T}$ has a non-linear benefit function dependent on L variables ($w_{ft}^{(l)}$) and L coefficients ($c_{ft}^{(l)}$). The coefficients $c_{ft}^{(l)}$ depends on the field's age a_{ft} at time t. Since we only need a relative difference in benefit functions to be portrayed, for a given vector of baseline coefficients $\mathbf{c} = [c^{(1)}, c^{(2)}, ..., c^{(L)}]'$, scalar multiply the vector according to the age category of the field, i.e.

$$\mathbf{c}_{ft} = [c_{ft}^{(1)}, c_{ft}^{(2)}, ..., c_{ft}^{(L)}]' = \begin{cases} \mathbf{c} & \text{if } a_{ft} \text{ is in germination or tillering stage} \\ \mathbf{c}e_1 & \text{if } a_{ft} \text{ is in stalk growth stage} \\ \mathbf{c}e_2 & \text{if } a_{ft} \text{ is in maturation stage} \end{cases} \tag{4.3.6}$$

with $e_1 > 1 > e_2$. The relationship between e_1 and e_2 is supported by the literature discussed in Section 2.2. Indicating that the stalk growth stage is most negatively impacted by an irrigation deficit, followed by the germination and tillering stage, and lastly the maturation stage. This is by no means the best way to handle the different aged crops responses to irrigation deficits. The exact values of $\mathbf{c}, e_1$ and e_2 are not specified but left as input parameters specified by a user. Given a benefit function for each field, a possible objective function is the sum of all the fields' benefit functions across the entire time period horizon T. Since different fields have different benefit functions, maximising the sum of the benefit functions will apply irrigation to the fields that yield the highest benefit. The objective function is thus the sum of all field benefits that needs to be maximised, i.e.

$$\max \sum_{f\in\mathcal{F},t\in\mathcal{T}} G_{ft}(w_{ft}^{(1)}, w_{ft}^{(2)}, ..., w_{ft}^{(L)}) \tag{4.3.7}$$

subject to constraint sets (4.2.2), (4.2.3), (4.2.4), (4.2.5), (4.2.6), (4.3.2), (4.3.3) and (4.3.4).

4.4 Program

This section describes how the MILP will be solved in a program. The program consists of two modules or sub-programs and is implemented in R [36], a free programming language for statistical computing, and the Gurobi optimiser R-package [40] provided by Gurobi commercial optimisation software [20]. R is used for visual representation, data analysis and preparing the MILP while the Gurobi optimiser package is used solely for optimising the MILP constructed in R. The Gurobi optimiser package sends the MILP formulation in R to the stand-alone Gurobi optimiser installed on the computer. Figure 4.4.1 represents the flowchart of the program. The inputs and outputs are depicted as circles whereas the program sub-programs are depicted as boxes.

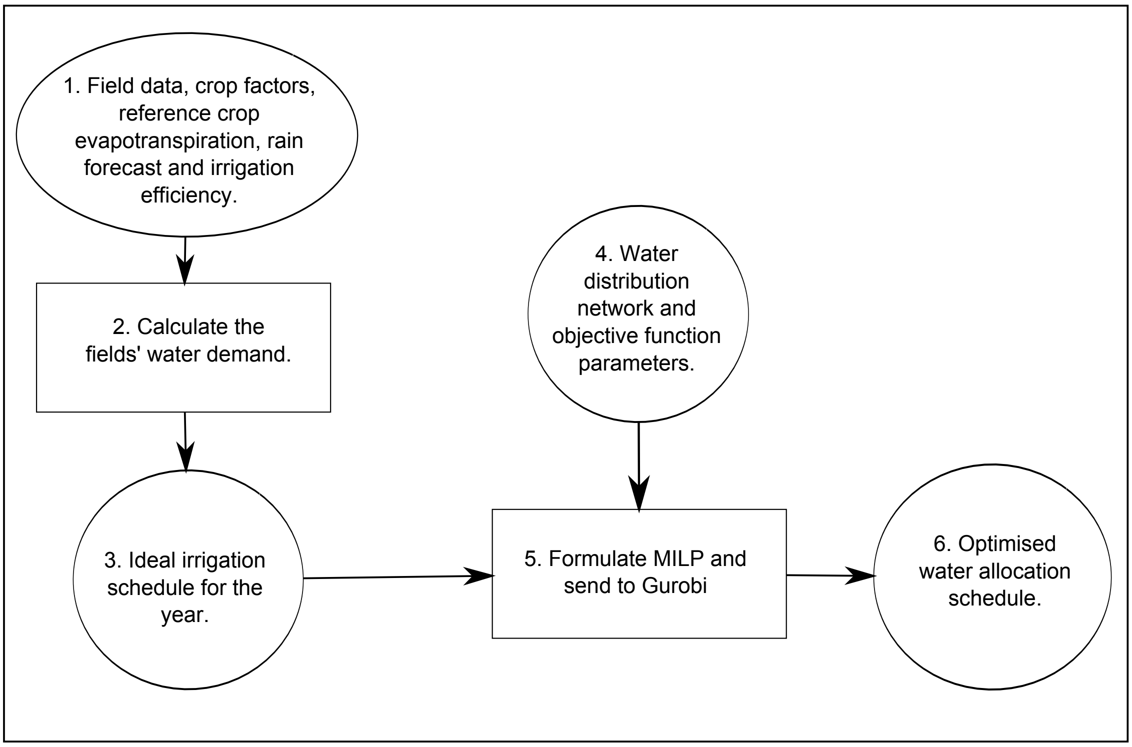


Figure 4.4.1: Program flowchart.

The inputs for the first sub-program (the square labelled 2) are the field data, reference crop evapotranspiration, crop factors and the rain forecast. The field's data consists of the growth start date, dry-off date, harvest date and the fields area measurement in hectares (ha). The dry-off date indicates the time after which no irrigation should be applied. The average weekly crop factors should be provided for each growth start week and spanning the longest harvest cycle in the field data (maximum length of growth to harvest date). The reference crop evapotranspiration and rain forecast should be provided at the weekly level. The irrigation efficiency is a scalar value on the real line $[0, 1]$. The first sub-program calculates the water demand volume to be supplied by the irrigation system (Ψ_{ft}) and the age of the fields during the year (a_{ft}) using the inputs in circle 1.

The inputs for the second sub-program are circles 3 and 4. Circle 4 contains the water distribution network and objective function parameters. The water distribution networks' parameters are the pump inflow capacities, field inflow capacities, and storage facility capacities. The inflow capacities are required in volume per time period t, in this case cubic metres per week. The storage capacity is the volume of water that the storage facility can hold at any time period t. Storage facilities may have a zero storage capacity indicating that a storage facility is not present at the pump station. The initial storage levels are required for use in the 0-th time period. The first task of the second sub-program is to formulate the MILP. The objective function parameters are the baseline coefficient vector $\mathbf{c}$, the length of the segments controlled by $\mathbf{b}$ and the age adjustment coefficients values (e_1 and e_2). The number of piecewise linear approximations used for each benefit function is four, i.e. $L = 4$. The last of the four linear functions is the penalty assignment. The penalty component of the benefit function requires a negative coefficient. The coefficients for the remaining three should be constructed so that the maximum of the benefit function is attained at Ψ_{ft}. This is assured if the first three coefficients are positive.

Figure 4.4.2 illustrates a benefit function for some field $f \in \mathcal{F}$ at time $t \in \mathcal{F}$. The x-axis is the amount of irrigation and the y-axis the dimensionless benefit. A dotted line is drawn at the water demand volume $\Psi_{ft} = 25,000$. In this example $\psi_{ft}^{(1)} = 10,000$, $\psi_{ft}^{(2)} = 7,500$, $\psi_{ft}^{(3)} = 7,500$ and $\psi_{ft}^{(4)} = 5,000$. Since the benefit function starts decreasing beyond Ψ_{ft} and the optimisation seeks to maximise the sum of benefit functions, the supplied water $w_{ft} = w_{ft}^{(1)} + w_{ft}^{(2)} + w_{ft}^{(3)} + w_{ft}^{(4)}$ should ideally be equal to 25,000 m^3, i.e. $w_{ft}^4 = 0$. The piecewise linear approximations are labelled on Figure 4.4.2 as numbers 1 to 4. The function for each approximation is

1. $c_{ft}^{(1)} w_{ft}^{(1)}$
2. $c_{ft}^{(1)} \psi_{ft}^{(1)} + c_{ft}^{(2)} w_{ft}^{(2)}$
3. $c_{ft}^{(1)} \psi_{ft}^{(1)} + c_{ft}^{(2)} \psi_{ft}^{(2)} + c_{ft}^{(3)} w_{ft}^{(3)}$
4. $c_{ft}^{(1)} \psi_{ft}^{(1)} + c_{ft}^{(2)} \psi_{ft}^{(2)} + c_{ft}^{(3)} \psi_{ft}^{(3)} + c_{ft}^{(4)} w_{ft}^{(4)}$

Note that $c_{ft}^{(4)}$ is negative, so that the last piecewise linear approximation is a decreasing function, and the variables $w_{ft}^{(l)}$ are not necessarily defined on the entire real line.

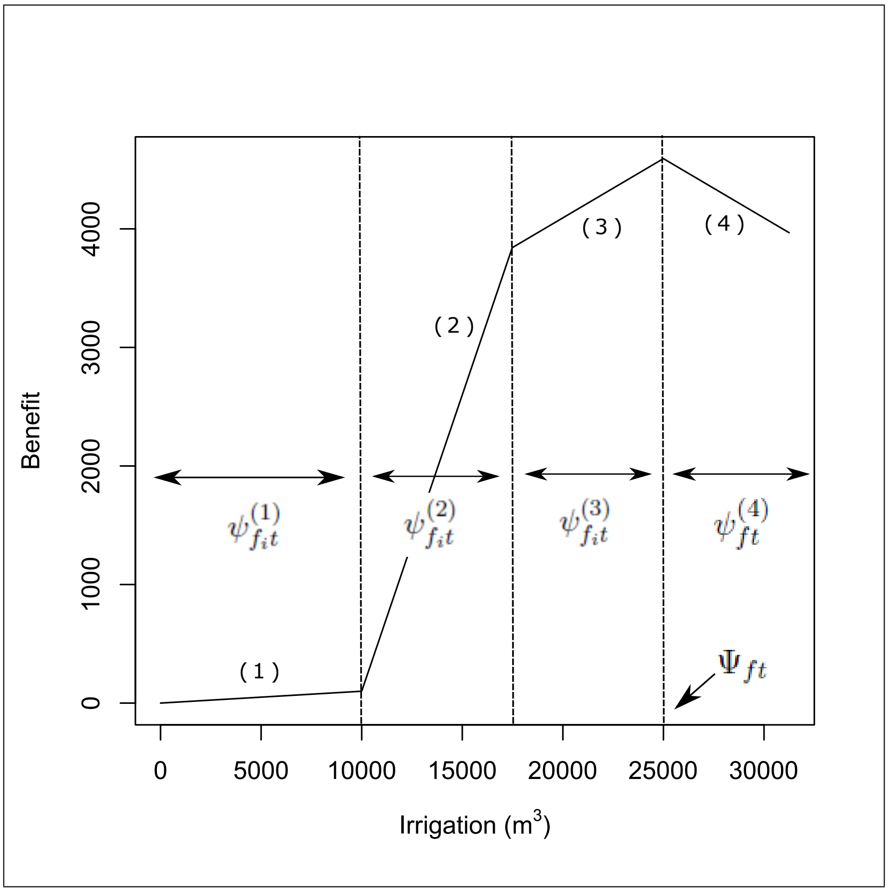


Figure 4.4.2: Illustration of a field benefit function.

After the MILP is formulated, the second task of the sub-program is to send the MILP to the Gurobi optimiser for solving and to summarise the results. Once the code has been run, the following command line performs all the necessary instructions (illustrated in Appendix A).

```
simIrr(idle, fieldData, flowData, storageData, cfData, refData, rainData,
eff, pnf, slopes, cutsAt, e1, e2, startWeek, multi, SOLVE)
```

The format of the data is illustrated in Appendix A and the parameter explanation is as follows:

- `idle`: A non-negative integer defined as the number of time periods between harvest and replant dates, i.e. the harvest date plus the idle becomes the following field's growth start date. The default value is zero, indicating a replant immediately after the harvest.

- `fieldData`: The field data including a column for field name, field size in hectares, the plant date, dry-off start date, harvest date, and pump station service relations.
- `flowData`: The inflow capacity matrix of size $(|\mathcal{P}|+|\mathcal{F}|)\times(|\mathcal{P}|+|\mathcal{F}|)$ measured in cubic meters (m^3). The row index containing the source node and column index containing the destination node.
- `storageData`: The initial and maximum storage levels data at the pump stations.
- `cfData`: The matrix of crop factors for each growth start week of the year and crop ages. The row index containing the start week (52 weeks) and the column index containing the crop's age (as many columns needed to calculate the the water demand for the longest crop cycle).
- `refData`: A single column matrix with 52 rows containing the weekly reference crop evapotranspiration measured in millimetres (mm).
- `rainData`: A single column matrix with 52 rows containing the rain forecast for each week of the year measured in millimetres (mm).
- `eff`: The irrigation efficiency. A scalar value on the real line $[0, 1]$.
- `pnf`: A vector of integers indicating which field's irrigation volume demand should be plot. Useful when the data set contains a large number of fields.
- `slopes`: The four element vector $\mathbf{c}$.
- `cutsAt`: The lengths of the segments will be controlled by a two element vector $\mathbf{b} = [b_1, b_2]'$ with values on the real line $[0, 1]$. The values correspond to proportions of the field's ideal irrigation volume. The first value indicates the proportional length of the first segment in terms of the ideal irrigation demand, i.e.

$$\psi_{ft}^{(1)} = b_1\Psi_{ft} \tag{4.4.1}$$

Similarly for the second segment, i.e.

$$\psi_{ft}^{(2)} = b_2\Psi_{ft} - b_1\Psi_{ft} \tag{4.4.2}$$

The length of the third segment is the remaining volume to make the ideal irrigation volume demand equal to the first three segments. The length of the last segment is only to carry the penalty. Since the coefficient of the last segment is negative, the benefit function, in theory, should never lie beyond the third segment, otherwise it would mean over-irrigating a field.

- `e1` and `e2`: The scalar values e_1 and e_2 on the real line $[0, 1]$ for the age adjustment factor.

- `startWeek`: The start week of the year for optimising the water allocation problem.

- `multi`: The number of weeks to solve the water allocation problem for, after the `startWeek`. If the MILP takes too long to solve this number can be reduced at the expense of accuracy.

- `SOLVE`: A logical value `TRUE/FALSE` or 0/1 indicating whether the MILP should be solved. Set to `FALSE` if the formulation of the MILP needs to be analysed.

The program outputs the decision variables (w_{pt}, w_{ft} and s_{pt}) in a matrix form with the time period's decision variables in each column. Since the number of decision variables can be vast, a visual summary of the decision variables is provided as well.

4.5 Numerical Example

A small example is presented to clarify the parameter usage and purpose of the non-linear benefit functions. For simplification, consider two fields f_1 and f_2 at a single time period t. Field f_1 is in its germination stage with $\Psi_{f_1t} = 100$ m^3 and field f_2 is in its maturation stage with $\Psi_{f_2t} = 100$ m^3. In this example there is only one pump station with a capacity of 130 m^3 and no storage facility. Assume a user specifies the following parameter values:

- The segments lengths controlled by $\mathbf{b} = [b_1 = 0.5, b_2 = 0.7]'$.

- The baseline coefficients $\mathbf{c} = [0.01, 0.5, 0.1, -0.1]'$.

- The age adjustment factors $e_1 = 1.1$ and $e_2 = 0.9$.

The segment lengths and constraints are:

- The segment lengths for field f_1 are $\psi^{(1)}_{f_1t} = b_1 100 = 50$, $\psi^{(2)}_{f_1t} = (b_2 - b_1) \times 100 = 20$, $\psi^{(3)}_{f_1t} = 30$ and $\psi^{(4)}_{f_1t} = 10$ (the arbitrary length segment to prevent over-irrigating). The segment lengths are the same for field f_2.

- The constraints on the split inflow variables for field f_1 are $w^{(1)}_{f_1t} \leq 50$, $w^{(2)}_{f_1t} \leq 20$, $w^{(3)}_{f_1t} \leq 30$ and $w^{(4)}_{f_1t} \leq 10$ (the split inflow variable that controls the penalty for over-irrigating). The constraints on the split inflow variables are the same for field f_2.

The coefficients for the benefit functions are calculated as follows:

- Since field f_1 is in the germination stage, its benefit function coefficient is just the baseline coefficients, i.e. $\mathbf{c}_{f_1t} = [c^{(1)}_{f_1t} = 0.01, c^{(2)}_{f_1t} = 0.5, c^{(3)}_{f_1t} = 0.1, c^{(4)}_{f_1t} = -0.1]'$.

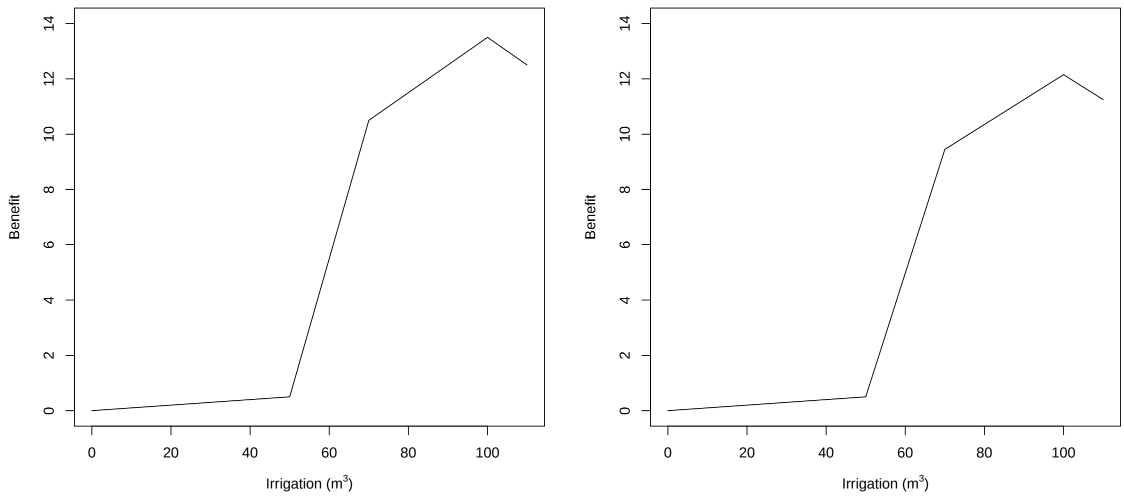
- Field f_2 is in the maturation stage. The benefit function coefficients is the baseline coefficient multiplied by the appropriate age adjustment factor, i.e. $\mathbf{c}_{f_2t} = \mathbf{c} \times e_2 = [0.009, 0.45, 0.09, -0.09]$.

The benefit functions are:

$$G_{f_1t}(w_{f_1t}^{(1)}, w_{f_1t}^{(2)}, w_{f_1t}^{(3)}, w_{f_1t}^{(4)}) = 0.01w_{f_1t}^{(1)} + 0.5w_{f_1t}^{(2)} + 0.1w_{f_1t}^{(3)} - 0.1w_{f_1t}^{(4)} \quad (4.5.1)$$

$$G_{f_2t}(w_{f_2t}^{(1)}, w_{f_2t}^{(2)}, w_{f_2t}^{(3)}, w_{f_2t}^{(4)}) = 0.009w_{f_2t}^{(1)} + 0.45w_{f_2t}^{(2)} + 0.09w_{f_2t}^{(3)} - 0.09w_{f_2t}^{(4)} \quad (4.5.2)$$

Figure 4.5.1 plots the benefit functions for fields f_1 and f_2.



(a) Field f_1 benefit function.

(b) Field f_2 benefit function.

Figure 4.5.1: Benefit functions.

The objective function is the sum of the benefit functions for field f_1 and f_2, i.e.

$$G_{f_1t}(w_{f_1t}^{(1)}, w_{f_1t}^{(2)}, w_{f_1t}^{(3)}, w_{f_1t}^{(4)}) + G_{f_1t}(w_{f_2t}^{(1)}, w_{f_2t}^{(2)}, w_{f_2t}^{(3)}, w_{f_2t}^{(4)}) = 0.01w_{f_1t}^{(1)} + 0.5w_{f_1t}^{(2)} + 0.1w_{f_1t}^{(3)} - 0.1w_{f_1t}^{(4)} + 0.009w_{f_2t}^{(1)} + 0.45w_{f_2t}^{(2)} + 0.09w_{f_2t}^{(3)} - 0.09w_{f_2t}^{(4)} \quad (4.5.3)$$

The $\delta_{ft}^{(l)}$'s needed are $\delta_{ft}^{(1)}, \delta_{ft}^{(2)}$ and $\delta_{ft}^{(3)}$ for $f \in \{f_1, f_2\}$ and $\delta_{ft}^{(l)} \in \{0, 1\}$. The constraints to split the inflow variable and behave in the desired manner for field f_1 are:

$$
\begin{aligned}
\delta_{f_1t}^{(1)}50 \le w_{f_1t}^{(1)} &\le 50 \\
\delta_{f_1t}^{(2)}20 \le w_{f_1t}^{(2)} &\le 20\delta_{f_1t}^{(1)} \\
\delta_{f_1t}^{(3)}30 \le w_{f_1t}^{(3)} &\le 30\delta_{f_1t}^{(2)} \\
0 \le w_{f_1t}^{(4)} &\le 10\delta_{f_1t}^{(3)}
\end{aligned} \tag{4.5.4}
$$

Constraint set 4.5.4 can be constructed similarly for field f_2. At this point, 130 m^3 needs to be distributed among the two fields by maximising the objective function. In an ideal situation, 200 m^3 would give the objective function maximum. The values for $w_{ft}^{(l)}$ needs to be determined that will maximise the benefit functions, thus minimising the irrigation deficit. Since the example is very small, it is easy to see that the values for $w_{ft}^{(l)}$ that maximise the objective function are as follows:

- For field f_1, the values of the split variables are $w_{f_1t}^{(1)} = 50$, $w_{f_1t}^{(2)} = 20$, $w_{f_1t}^{(3)} = 0$ and $w_{f_1t}^{(4)} = 0$. The total inflow at field f_1 is therefore $w_{f_1t} = 50 + 20 + 0 + 0 = 70$ m^3.

  To achieve the total inflow of $w_{f_1t} = 70$, the sum of the variables $w_{f_1t}^{(1)}$ and $w_{f_1t}^{(2)}$ had to be equal to 70. The total inflow of $w_{f_1t} = 70$ would not be possible from $w_{f_1t}^{(1)} = 50$ and $w_{f_1t}^{(3)} = 20$. For $w_{f_1t}^{(3)}$ to be equal to 20 it would require that $\delta_{f_1t}^{(2)} = 1$ and $\delta_{f_1t}^{(3)} = 0$. However, if $\delta_{f_1t}^{(2)} = 1$ then $\delta_{f_1t}^{(1)}$ would be forced to be equal to 1 and the total inflow would exceed 70. This illustrates how the constraints work to ensure that the split variables behave in the desired manner as explained in Section 4.3. The benefit function for field f_1 is $G_{f_1t}(...) = (0.01 \times 50) + (0.5 \times 20) = 10.5$.

- For field f_2, the values of the split variables are $w_{f_2t}^{(1)} = 50$, $w_{f_2t}^{(2)} = 10$, $w_{f_2t}^{(3)} = 0$ and $w_{f_2t}^{(4)} = 0$. The total inflow at field f_2 is therefore $w_{f_2t} = 50 + 10 + 0 + 0 = 60$ m^3. The benefit function for field f_2 is $G_{f_2t}(...) = (0.009 \times 50) + (0.45 \times 10) = 4.95$.

The objective function value is thus 15.45. Notice that to achieve the maximum objective function value, the available water had to be distributed between the two fields with field f_1 receiving more water.

Chapter 5

Model Implementation

This chapter illustrates the program implementation and the conclusions that can be drawn from an optimised water allocation model. Section 5.1 implements the program using a small data set. Section 5.2 implements the model using the larger data set provided by the commercial sugarcane farmer. Section 5.3 further test the developed model for completeness.

5.1 Test Data Set

A small data set is used so that the desired performance of the model can easily be verified at the field level. The test data set will be an extract from the larger data set with capacities adjusted to force an irrigation deficit. The test data set is tabulated in Tables 5.1.1 and 5.1.2. The data in Table 5.1.1 are the required harvest schedule, field areas and the pump service relationships. The water distribution network service relations in this data set is the network in Figure 4.1.1. The inflow capacities in Table 5.1.2 is a summary of the flow capacity matrix.

In this data set, the main pump station p_1 is denoted as mc ("main cluster") which draws the water from the source. The cluster pump stations are $p_2 = c_2$, $p_3 = c_3$ and $p_4 = p_4$. The sub-cluster pump stations are $p_5 = sc_5$, $p_6 = sc_6$, $p_7 = sc_7$ and $p_8 = sc_8$. The maximum rate at which a field is allowed to draw water from the pump station has been fixed to 30,000 m^3 per week. This value represents the irrigation system (furrow, centre-pivot or drip) capacity. The weekly average crop factor, weekly reference crop evapotranspiration and rain forecast data are not shown. To obtain the weekly rain forecast, the monthly value was divided by the number of weeks in the month and assigned the average for each week of that month.

At this stage, the data is in the required form and ready to be used in the program. The next step is to specify the parameters for the program. The objective function parameters need to be chosen to achieve the desired outcome as described in Chapter 4. The functional form illustrated in Figure 4.4.2 is desirable for a field's benefit function when the aim of the

Table 5.1.1: Field data and service relationships.

Field	Area (ha)	Growth Start	Harvest Date	Dry-off	Cluster	Sub-Cluster
f_1	27.3	2013-05-04	2014-05-04	2014-02-18	c_1	sc_1
f_2	16.4	2013-10-31	2014-11-24	2014-10-10	c_1	sc_1
f_3	9.3	2013-05-05	2014-04-08	2014-01-20	c_1	sc_2
f_4	24.1	2013-05-10	2014-04-11	2014-01-23	c_2	sc_3
f_5	18.1	2013-07-17	2014-06-29	2014-03-24	c_2	sc_3
f_6	9.8	2013-09-19	2014-10-02	2014-07-31	c_2	sc_3
f_7	15	2013-06-17	2014-05-29	2014-02-24	c_3	sc_4
f_8	16	2013-08-19	2014-09-02	2014-06-28	c_3	sc_4

Table 5.1.2: Capacity in 1,000 cubic metres (m^3).

Pump Station	Initial Storage	Storage Capacity	Inflow Capacity Per week
mc	75	90	50
c_1	0	45	30
c_2	0	50	30
c_3	0	50	30
sc_1	0	30	30
sc_2	0	20	30
sc_3	0	35	30
sc_4	0	25	30

optimisation is to distribute limited water. The slow increase for the first segment of the applied irrigation provides little benefit. Thereafter the benefit function starts increasing drastically. This encourages the optimisation process to pass the first segment. After the second segment, the increase in benefit slows down. This implies that it is more beneficial for the optimisation process to supply water to all fields at least until the second segment when there is limited water, i.e. each field's supplied water should lie on at least the second segment. This will also protect smaller fields from being swamped by larger fields. Using such a functional form does not mean that a minimum irrigation requirement exists, however as discussed in the literature review, evidence exists that certain irrigation deficits do not have extreme deleterious effects on yields but depends on the climate, cultivars and soil as well. The objective function parameters choice and motivations provided hereafter are to illustrate the program results and do not necessarily represent the commercial sugarcane farmer's ideal parameter choice.

- The length of the segments are chosen according the classification of irrigation deficits found in Berhe *et al* [6]. An irrigation deficit more than 50% will be considered severe. An irrigation deficit of less than 30% will be considered mild. Thus, the `cutsAt` parameter will be set as $\mathbf{b} = [0.5, 0.7]$.

- The first segment's slope should be small to convey the low benefit obtained for applying a severe irrigation deficit. The second segment slope should be high to encourage the optimisation to lie at least on the third segment. The third segment indicating mild irritation deficit, should have a relatively smaller slope than the second segment. The benefit function should decrease after the third segment requiring a negative slope. The baseline coefficient vector $\mathbf{c}$ will be set as $\mathbf{c} = [0.001, 0.5, 0.1, -0.1]$.

- The age adjustment coefficients need to convey the relative importance of the different growth stages. The age adjustment coefficients will be arbitrarily set as $e_1 = 1.1$ and $e_2 = 0.9$. Thus, the coefficient of the benefit function will be higher for the stalk growth stage and lower for the maturation stage relative to the germination and tillering stage.

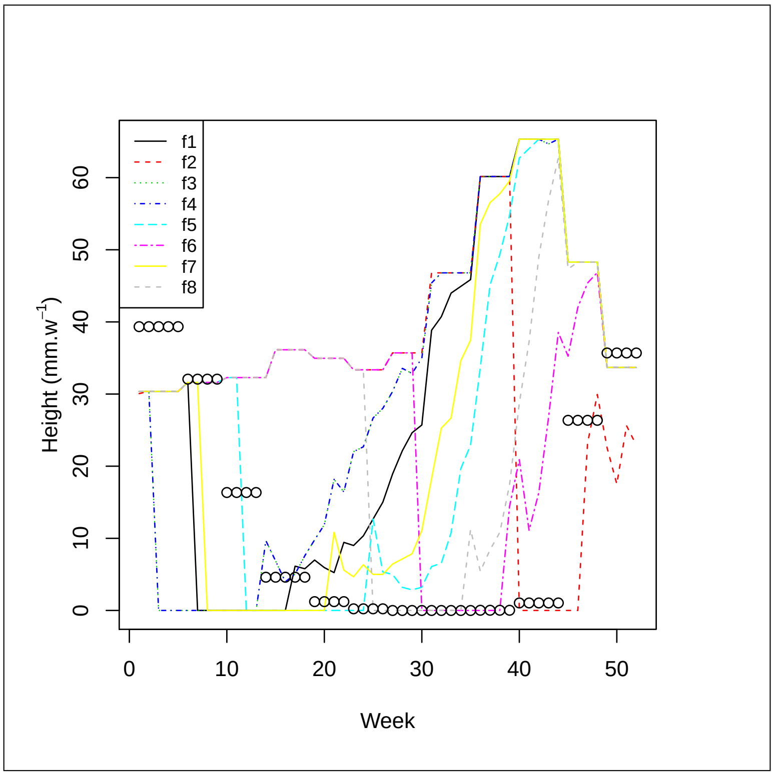


Figure 5.1.1: Weekly field irrigation height demand for the year.

The model was run for the 52 weeks of the year 2014. Figure 5.1.1 plots the field's irrigation height demand and the forecasted rain height. The rain forecast amount is plotted as small circles. Fields with zero irrigation height in Figure 5.1.1 indicate the field's dry-off

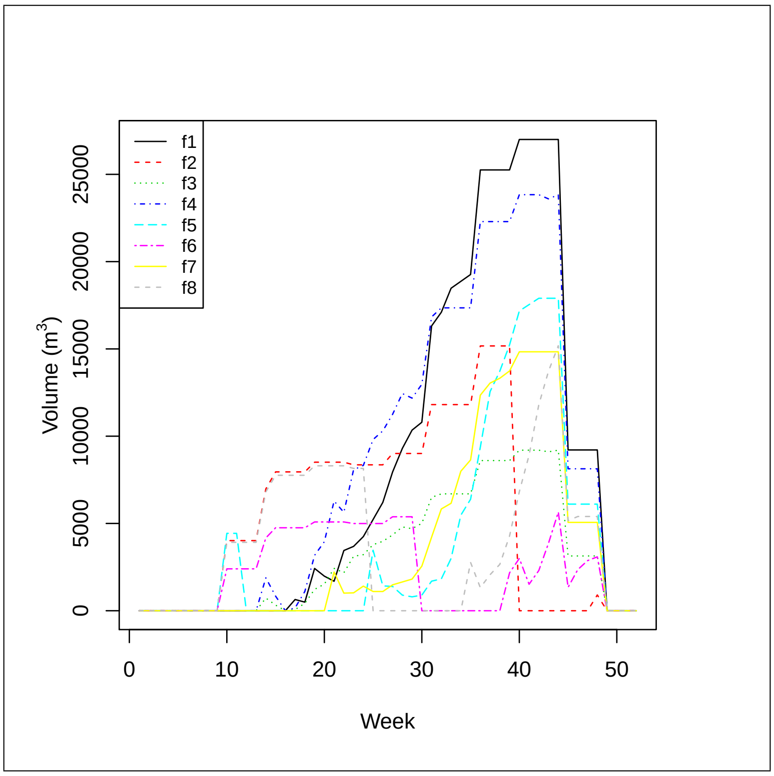


Figure 5.1.2: Weekly field irrigation volume demand for the year.

period. Weeks 31 to 44 are the high irrigation height demand periods. The high irrigation height demand in this period is due to the large reference crop evapotranspiration caused by the warmer climate. The irrigation volume to be supplied by the water distribution network should be zero when the rain height exceeds the irrigation height demand. These periods occur during weeks 1 to 9 and weeks 49 to 52 as seen in Figure 5.1.2. Figure 5.1.2 is thus the plot of $\Psi_{f,t}$ for each field during the year. Once the rain and area components have been considered, the water demand is positive between weeks 10 to 48 with weeks 31 to 44 remaining the high demand period. The fact that there is little to no rain during the high demand period exacerbates the strain on the water distribution network.

5.1.1 Supply and Demand Analysis

After the data required for the optimisation have been prepared (steps 1 to 4 of the program flowchart in Figure 4.4.1), the optimisation stage is carried out. Figure 5.1.3 plots the main pump station inflow, outflow and storage facility level for each week of the year. The inflow line is the plot of $w_{\text{mc},t}$ for each week. The outflow line is a plot of the sum of the cluster inflows, i.e. a plot of $\sum_{c\in\mathcal{C}} w_{c,t}$ for each week. The main pump station's storage facility level in Figure 5.1.3 is a plot of the storage facility's level for each week, i.e. a plot of $s_{\text{mc},t}$. The plot starts at week 1 and showing the storage level at the end of the week accounting for outflow. The labels on the right axis are the inflow (marked by "in") and outflow (marked by "out") capacities. The inflow at the main pump station starts at week 8 and ends at week 48, which corresponds to the end of the fields' irrigation volume demands. The early outflow at the main pump station is the distribution of the initial storage volume to the cluster level pump stations and storage facilities. The outflow ends at week 48 as well. The main pump station inflow is at capacity during the weeks 16 to 47 indicating high usage and possibly a bottleneck. The optimisation begins filling the storage facility to capacity before the high demand period. After the high demand period the storage facility level decreases and remains at zero for the remaining weeks of the year.

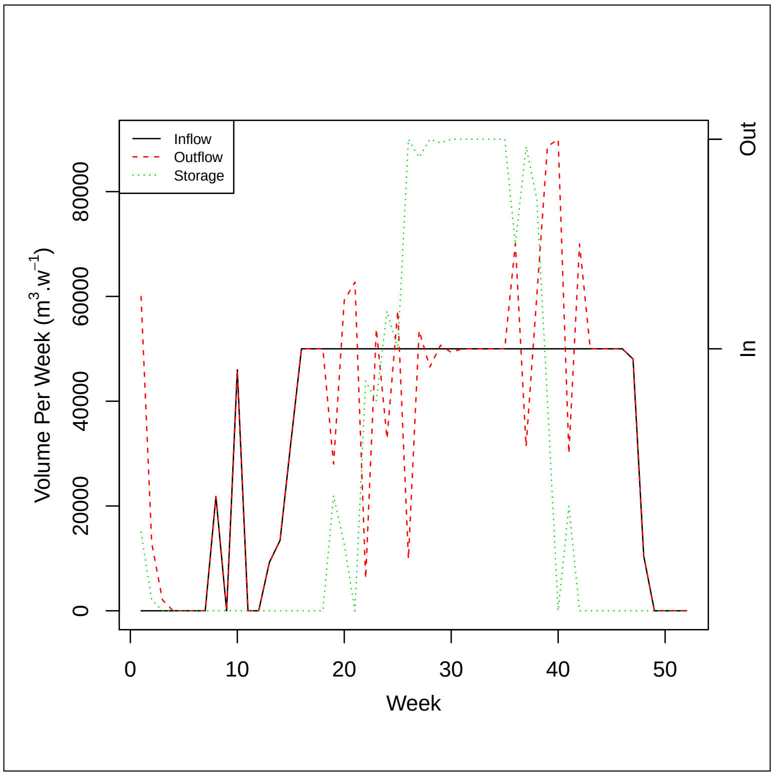


Figure 5.1.3: Main pump station usage.

Figure 5.1.4 is a plot of the cluster pump stations inflows for each week of the year, i.e. $w_{c,t}$ for $c \in \mathcal{C}$. Figure 5.1.5 plots the cluster pump stations outflow for each week of the year which is the sum of the inflows of the sub-cluster pump station which shares a common cluster pump station, i.e. $\sum_{sc \in E(c)} w_{sc,t}$ for $c \in \mathcal{C}$. The maximum inflows and outflows per week is marked on the right axis using the index of the cluster pump station as the marker. This maximum inflow corresponds to the inflow capacity listed in Table 5.1.2. Since all cluster pump stations have the same maximum inflow capacity it is only marked once. The maximum outflow at each cluster pump station is the sum of the maximum inflows of connected sub-clusters pump stations. Figure 5.1.6 plots the cluster pump stations storage facility levels for each week of the year, i.e. $s_{c,t}$ for $c \in \mathcal{C}$. The storage capacity at the cluster pump station is marked on the right axis with the cluster index as a marker. The cluster inflows may seem erratic, but needs to be read in conjunction with cluster outflows and storage levels. These inflows and outflows start earlier than the main pump station's. The large early inflow at cluster c_1 immediately feeds the subsequent sub-clusters, whereas the early inflow at c_2 and c_3 are sent to their storage facilities. The inflows take turns in peak usage to fill the storage facility for the high demand period. The storage facilities levels decreases drastically during the high demand period. The outflows of cluster c_1 does not reach peak usage indicating that it has unused capacity. As with the main pump station, the inflows, outflows and storage facility levels are at peak usage during the high demand period and start decreasing towards week 48.

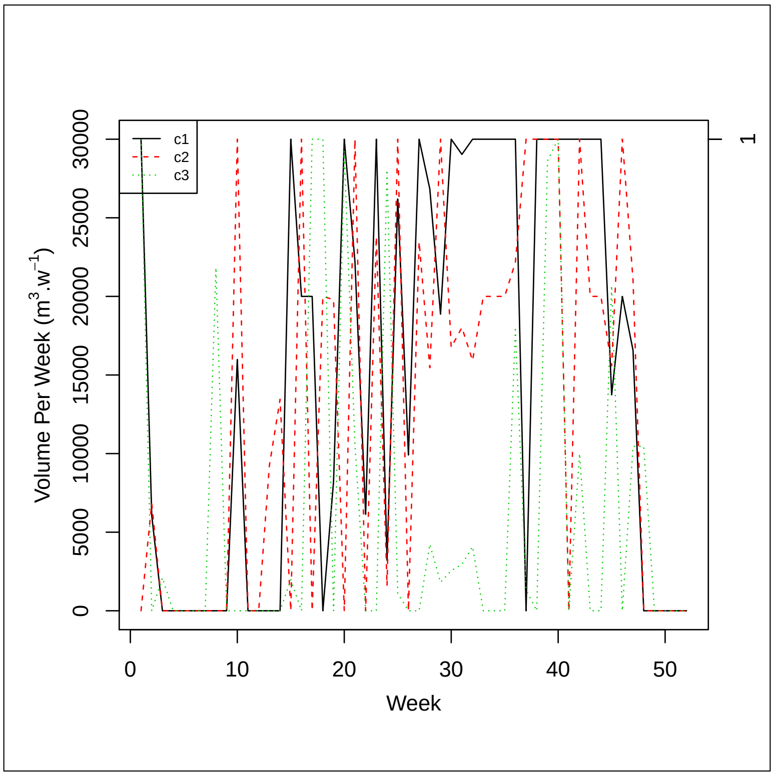


Figure 5.1.4: Cluster pump stations inflows for the year.

Figure 5.1.7 plots the sub-cluster pump station inflows for each week of the year, i.e. $w_{sc,t}$ for $sc \in \mathcal{S}$. Figure 5.1.8 plots the sub-cluster pump station outflows for each week of the year which is the sum of the fields' inflows which share a connected sub-cluster pump station, i.e. $\sum_{f \in D(sc)} w_{ft}$ for $sc \in \mathcal{S}$. In Figures 5.1.7 and 5.1.8, the maximum inflows and outflows per week for each sub-cluster pump station is marked on the right axis using the sub-cluster pump station index as the marker. Since each sub-cluster has the same inflow capacity per week the maximum inflow per week is marked only once. These maximum inflows corresponds to the pump inflow capacities in Table 5.1.2. Figure 5.1.9 plots the sub-cluster pump stations storage levels for each week of the year with the storage capacities

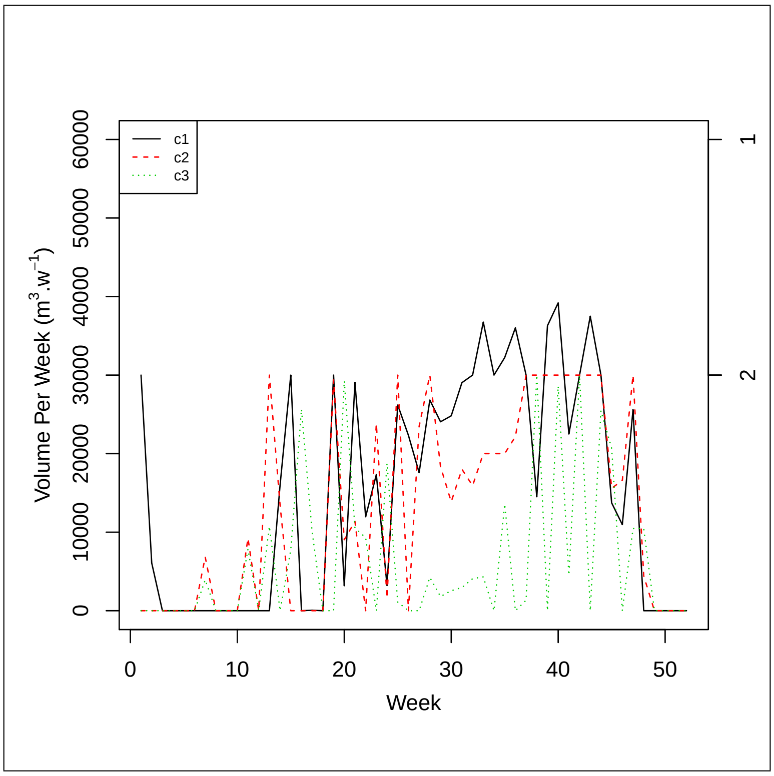


Figure 5.1.5: Cluster pump stations outflows for the year.

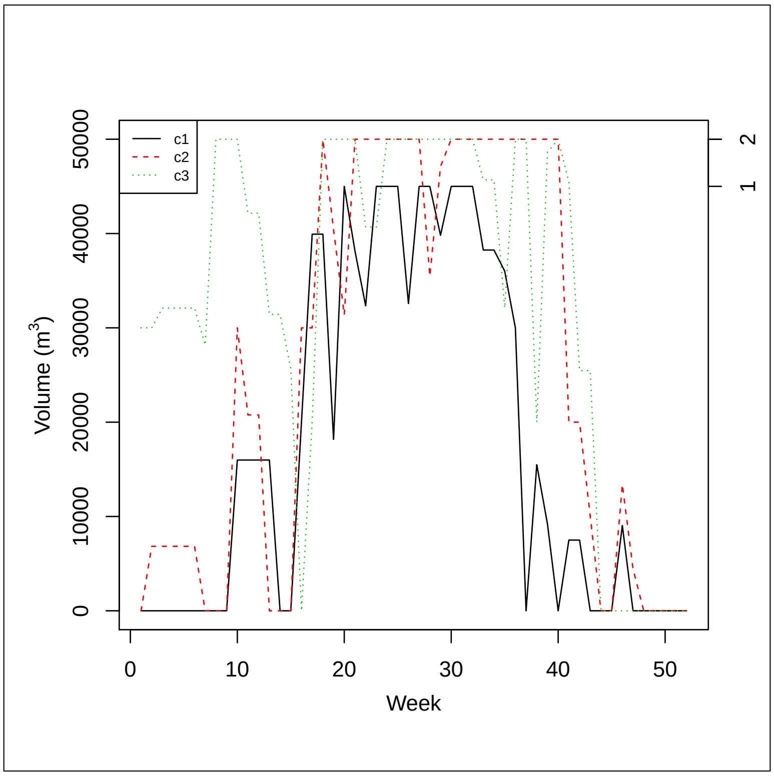


Figure 5.1.6: Cluster storage facility levels for the year.

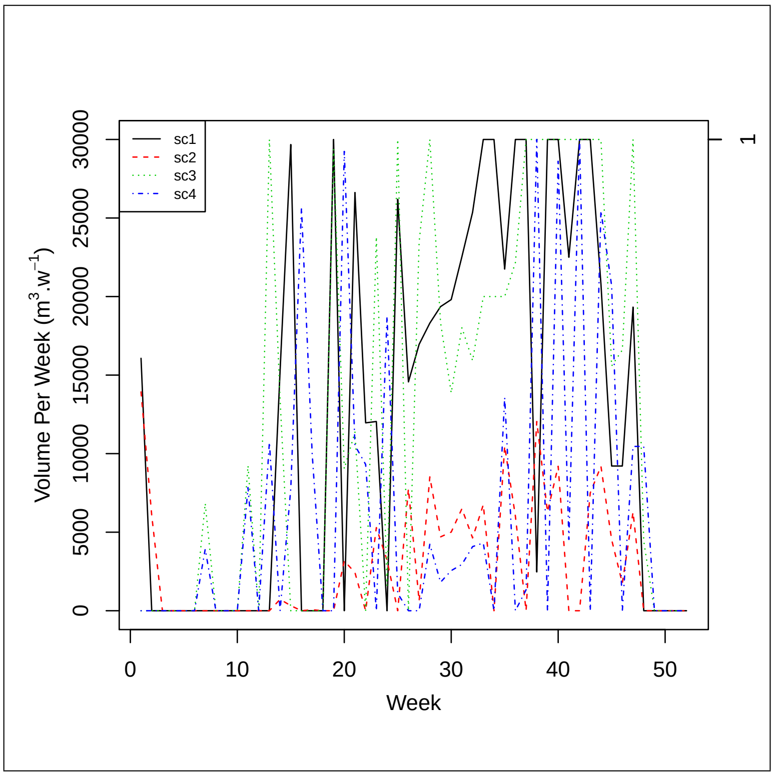


Figure 5.1.7: Sub-cluster pump stations inflows for the year.

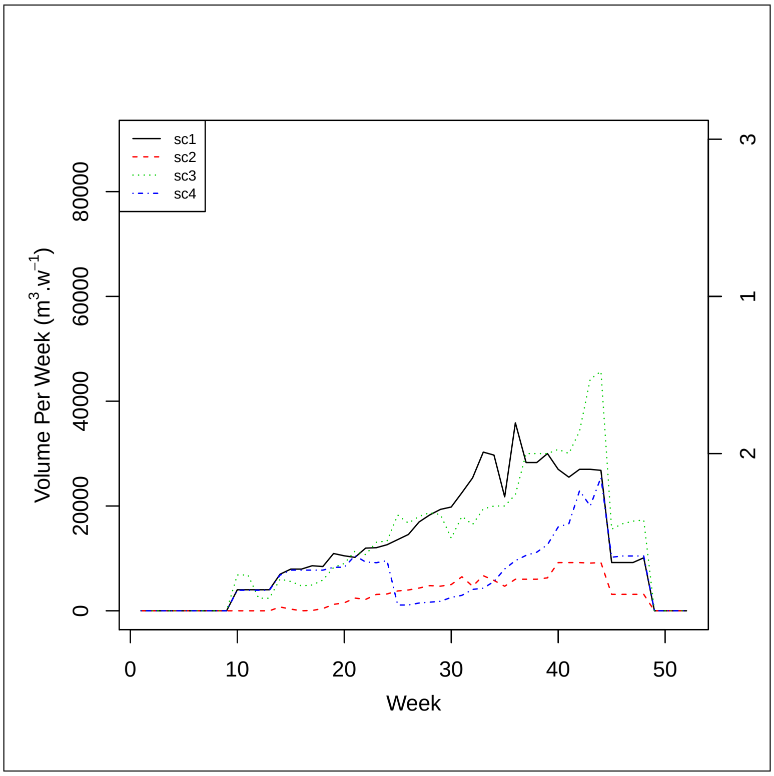


Figure 5.1.8: Sub-cluster pump stations outflows for the year.

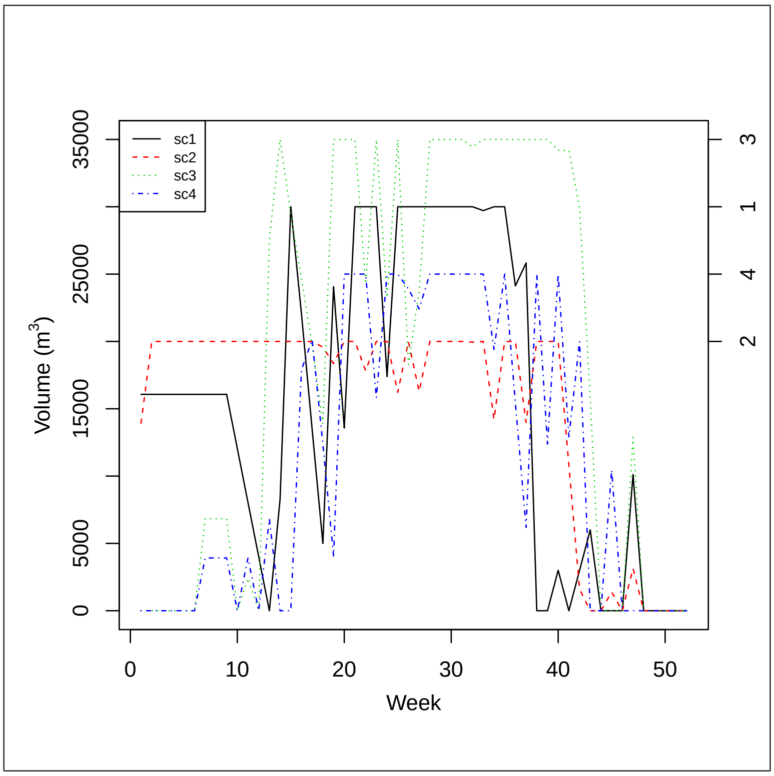


Figure 5.1.9: Sub-cluster storage facility levels for the year.

marked on the right axis using the sub-cluster pump stations index as a marker.

The inflows for week 1 at sub-cluster pump station sc_1 and sc_2 are the flows being received from cluster c_1 which was relaying the initial storage of the main pump station to the sub-clusters' pump stations storage facilities. The inflows at sub-cluster sc_2 are low since sc_2 only services one field. The sub-cluster pump stations plot outflows should look very similar to the plot in Figure 5.1.2 since it is the available water being supplied to the fields. Thus, sub-cluster pump stations' outflows begin only when the fields water demand is positive and ends when the fields' irrigation demand (Ψ_{ft}) is zero. As with the other pump stations' storage facilities, the sub-cluster storage facilities levels are at peak before the high demand-period, then drastically decrease and remain at zero till the end of the year.

Figure 5.1.10 plots the supplied irrigation volume for each field during the year, i.e. Figure 5.1.10 is a plot of w_{ft} for each field during the year. There is no difference between the demand and supplied irrigation during the low-demand periods (weeks 1 to 30 and weeks 45 to 52). The differences occur during the high demand period (week 31 to 44). The differences are due to the limited water supply during that period. The size of the deficit at each field is mitigated by the objective function in that the objective function controls at which fields the deficits occur. No field is being completely ignored in the optimisation during the high demand period, which demonstrates that the design of the objective function seems to have the intended effect as mentioned in Section 4.3.

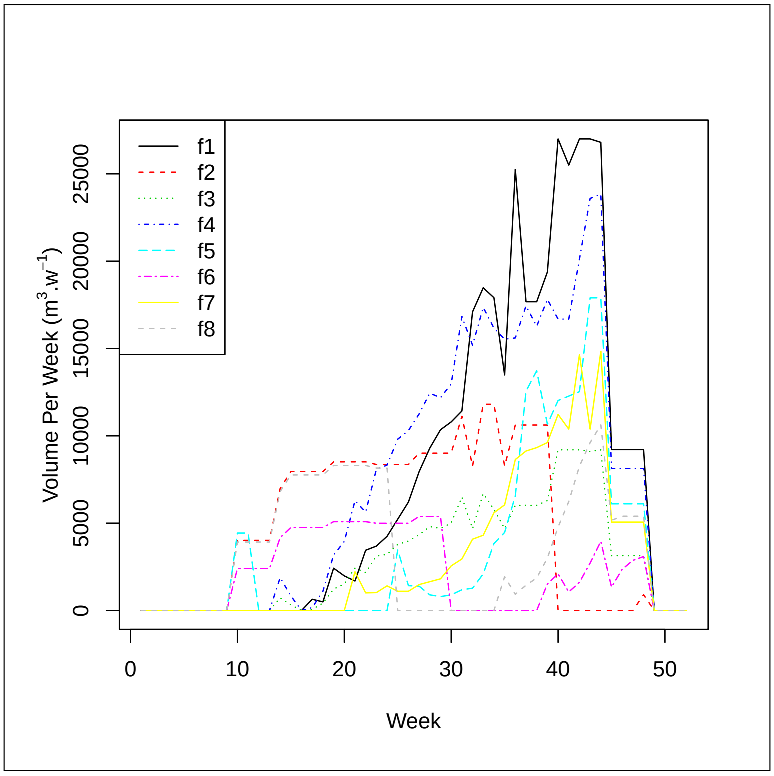


Figure 5.1.10: Supplied irrigation volume by the water distribution network

5.1.2 Structural Intervention

A potential large benefit of close inspection of the optimised water allocation model solution is to identify components of the water distribution network that are responsible for the irrigation deficits and at which times severe deficits occur. Identifying the water distribution network's components that operate at continuous peak usage may reveal the cause of irrigation deficits. Temporary or permanent improvements to these components can decrease the irrigation deficit. To this end, we observe that the main pump station's inflows is at peak usage for a long period of the 52 weeks. Cluster pump station c_1 operates almost continuously at its maximum inflow rate during the high demand period. Sub-cluster pump station sc_3's inflows operate at capacity for a shorter continuous period compared to cluster pump station c_1. All pump stations' storage facilities are at capacity before the high-demand period.

Figures 5.1.11 and 5.1.12 are plots of the irrigation deficits attributed to the cluster and sub-cluster pump stations respectively. Each plot sums the deficits of the fields that share the common cluster or sub-cluster. Cluster pump station c_1 is responsible for a larger portion of the irrigation deficit. A further breakdown at the sub-cluster pump station level

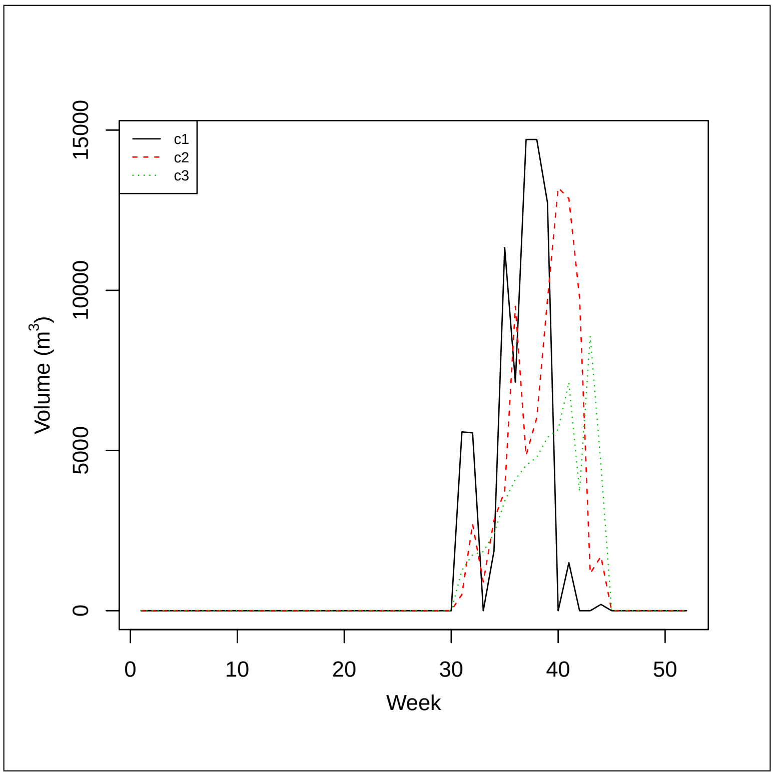


Figure 5.1.11: Irrigation deficits by clusters.

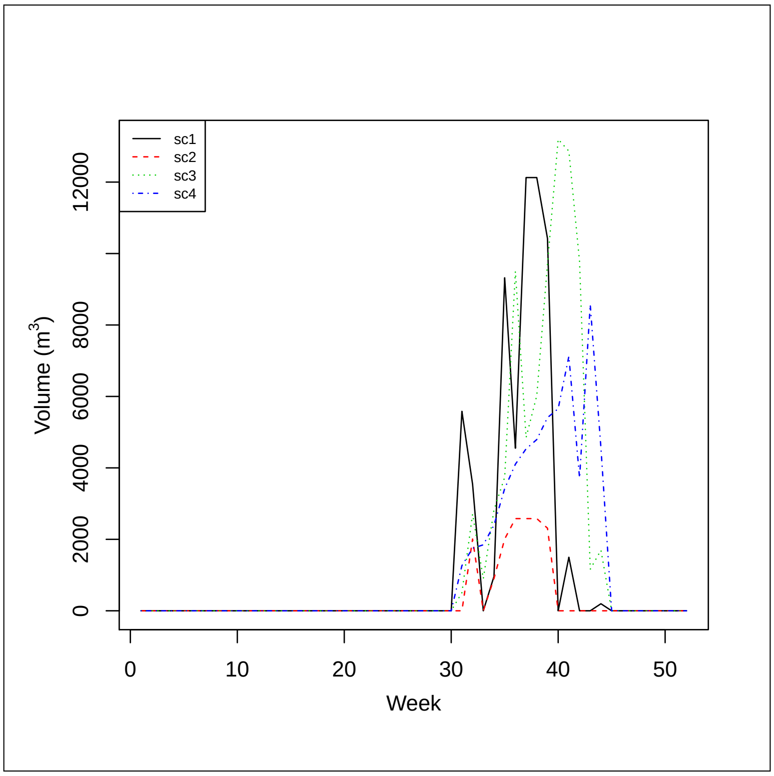


Figure 5.1.12: Irrigation deficits by sub-clusters.

indicates that sub-cluster sc_1 is the cause of the large irrigation deficit. This methodology may be conducted for the other components and at the field level as well. The total irrigation deficit for this optimisation is 213,897 m^3 for the year. The flexibility of an optimisation model allows parameters to be altered and the effect observed in the results, i.e. a scenario. As an example, increasing the pump stations' capacities who are continuously operating at maximum inflows should decrease the irrigation deficit. The main pump station satisfies this criteria and thus increasing the inflow capacity should decrease the irrigation deficit. Cluster c_1 and sub-cluster sc_3 are also possible candidates. Table 5.1.3 tabulates the different scenarios with respect to the test data in Table 5.1.2 including each scenario's total annual irrigation deficit and computation time.

Table 5.1.3: Different scenarios with their respective irrigation deficit in m^3 per year.

Scenario	Annual irrigation deficit in cubic metres	Computation time in seconds	Percentage change relative to base case scenario
1. Base Case	213,897	0.36	
2. mc at 110% flow capacity.	146,158	0.36	-32
3. mc with unlimited flow.	105,280	0.35	-51
4. mc with unlimited flow, c_1 and sc_3 at 110% inflow capacity.	49,000	0.35	-77
5. Ignoring rain	459,121	0.38	115

Scenario 1 is the optimisation described and analysed in Section 5.1.1. Scenario 2 uses the same parameter inputs as Scenario 1, but the main pump station inflow capacity is increased by 10%. The increased inflow capacity reduces the irrigation deficit as expected. Scenario 3 uses the same input parameters as Scenario 1 but has an unlimited inflow capacity at main pump station. Scenario 3 is included to illustrate that the irrigation deficit can be the result of multiple sources within the water distribution network. With an unlimited inflow at the main pump station, there is still an irrigation deficit, indicating that subsequent pump stations or storage facilities are responsible for the irrigation deficit. Scenario 4 uses the same parameter inputs as Scenario 1 but the inflow at the main pump station is unlimited and cluster pump station c_1 and sub-cluster pump station sc_3 inflow capacities are increased by 10%. The parameter setting in Scenario 4 further reduces the irrigation deficit. Similar changes to the water distribution network parameters can eliminate the irrigation deficit completely, i.e. by changing pump station inflow capacities or storage facility capacities. Scenario 5 uses the same parameter settings as Scenario 1, but the rain forecast is ignored, i.e. the rain forecast vector is set to zero. The irrigation deficit in Scenario 5 is 115% greater than in Scenario 1. The effect of incorporating the rain forecast is thus important.

Section 5.1.1 illustrates that close comparisons of the fields water demand and supplied water can be made. The inflow figures quickly identifies fields that deviate far from the desired inflows (weekly water demand plot verses the weekly supplied water). The methodology in Section 5.1.2 illustrates an approach to identify components responsible for irrigation deficits. Components are identified by finding pump stations who have long continuous capacity usage. Once these components are identified, a positive structural change can be applied by changing the parameters or input data. These structural changes correspond to increased pump station inflow or storage volume capacities. The effect of the structural changes are summarised by the effect on the annual irrigation deficit. Together Section 5.1.1 and 5.1.2 shows that a mathematical model can be used for strategic planning to identify issues that might arise during the year.

5.2 Commercial Data Set

To determine whether the program can be used in a practical setting, an industry sized data set needs to be tested in the program. An industry sized data set means that there will be more fields, pump stations and storage facilities resulting in a larger MILP formulation, i.e. more decision variables and constraints. Since the mathematical model was developed for a general case, the program will be able to solve a larger problem. However, a larger MILP requires greater computational power and time to solve. To protect the commercial sugarcane farmer's data, the actual harvest schedule, flow capacities and storage facilities data will not be listed in its entirety. It is important to note that since this is a mathematical model, the results from the optimisation does not necessarily represent the actual flow rates on the commercial sugarcane farm. The data set contains 540 fields with pump service relations listed in Table 5.2.1.

The industrial scenario was run using the same objective function parameters as in Section 5.1. The fields water demand plot will be suppressed since the number of fields will make the plot unreadable. Figure 5.2.1 is the weekly irrigation volume demand, i.e. the plot of $\sum_f \Psi_{f,t}$ for each week. Figure 5.2.2 plots the main pump usage for each week during the year. The maximum weekly inflow and outflow is marked on the right axis of Figure 5.2.2. The weekly storage capacity is also marked on the right axis of Figure 5.2.2 and is exactly zero, i.e. there is no storage facility at the main pump station. The weekly volume demand is positive between weeks 9 to 49 with the high-demand period being weeks 36 to 44. The zero demand is due to the forecast rain height being greater than the required irrigation height demand. Since there is no storage capacity at the main pump station, the outflow for the week is exactly the inflow for that week and since the inflow capacity is greater than the outflow capacity, the inflow will be limited by the outflow capacity. The inflow at the main pump station begins one week earlier than the irrigation volume demand. The onset of early inflow is most likely to start filling the storage facilities.

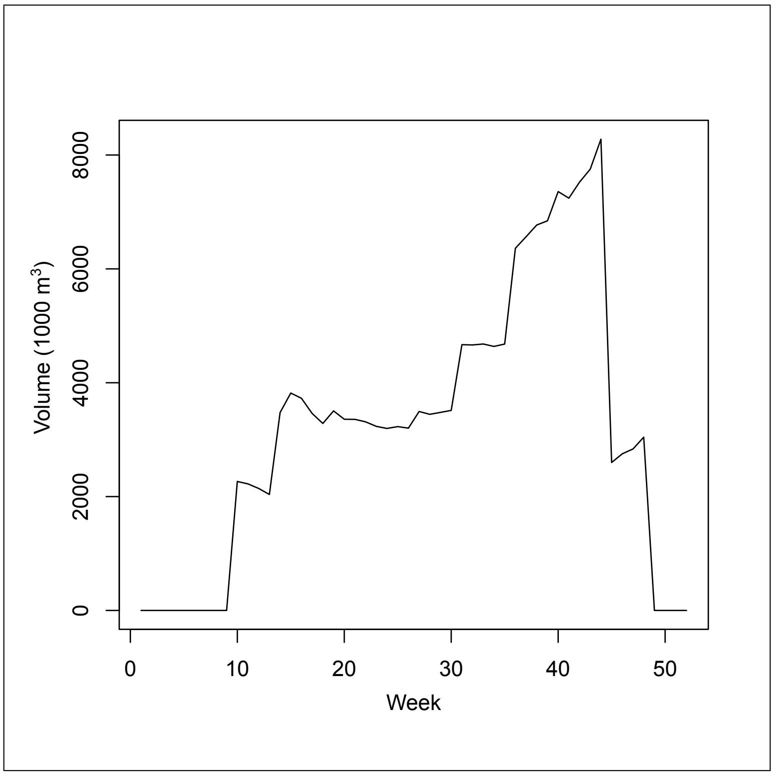


Figure 5.2.1: Weekly irrigation volume demand for the year.

Figures 5.2.3 and 5.2.4 plot the cluster pump station inflows and outflows respectively. Figure 5.2.5 plots the cluster storage facility levels for each week of the year. The cluster pump stations storage facilities' are filled and emptied throughout the year, except the cluster storage facility at cluster pump station c_4 since cluster pump station c_4 has the highest inflow and outflow capacity. Only cluster pump station c_3 is at peak outflow during the high demand period. Cluster pump stations c_1, c_2 and c_4 storage levels are at capacity at the end of the year. This represents a characteristic of an optimisation method, where it is possible for certain decision variables to be non-zero if it satisfies the constraints and does not interfere with the objective function optimisation albeit not intuitive. It occurs when the solver sets values during the search for the optimal solution and arrives at a solution with non-zero decisions variables without affecting the objective function optimisation. It is also typical of what is known as an end-of-horizon effect. This illustrates a limitation of the model developed but is a limitation of optimisation models in general. The end-of-horizon

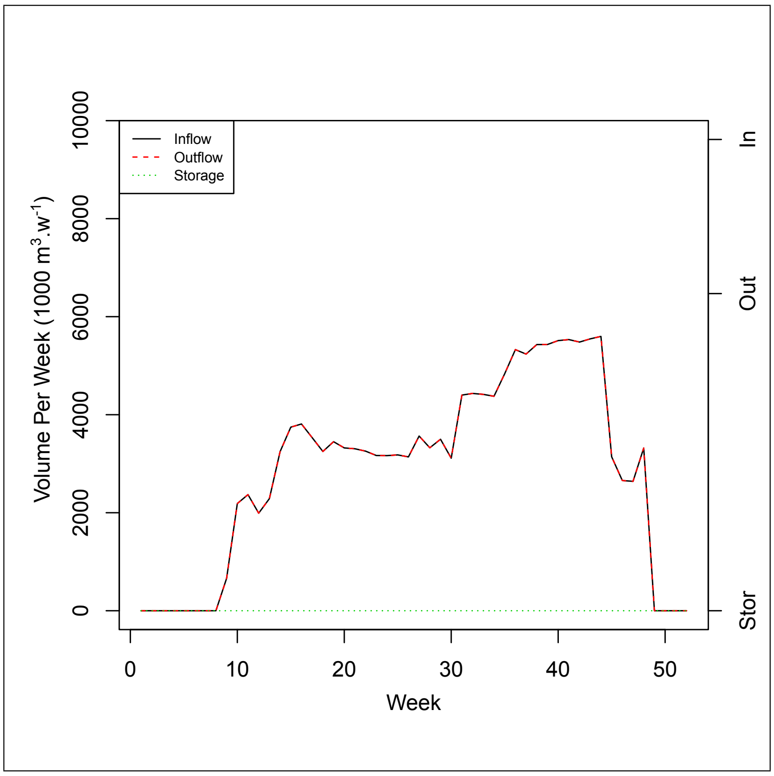


Figure 5.2.2: Main pump station weekly usage for the year.

effect can be controlled by adding additional constraints.

Table 5.2.1: Cluster sub-cluster service relationship.

Cluster	Serviced Sub-Clusters
c_1	sc_1, sc_2, sc_3
c_2	sc_4, sc_5
c_3	sc_6
c_4	sc_7, sc_8

Figures 5.2.6 and 5.2.7 plot the sub-cluster pump station inflows and outflows respectively. Figure 5.2.8 plots the sub-cluster storage facility levels for each week of the year. Sub-cluster pump stations sc_2 and sc_6 are operating at maximum inflow capacity during the high demand period. The sub-cluster pump station inflows and outflows end exactly when the fields' water demands end. The early inflow from the main pump station fills the sub-cluster pump station storage facility sc_7.

As in Section 5.1, the model can be rerun using different parameters and inputs. Table 5.2.2 (page 50) tabulates the different scenarios and their respective results. The scenarios were evaluated using R version 3.0.1 and Gurobi 6.0.2 on an Intel® Core™ i5-2400 CPU at 3.10GHz with 6GB of ddr3-1333 of memory and required 250MB of RAM to run. Each

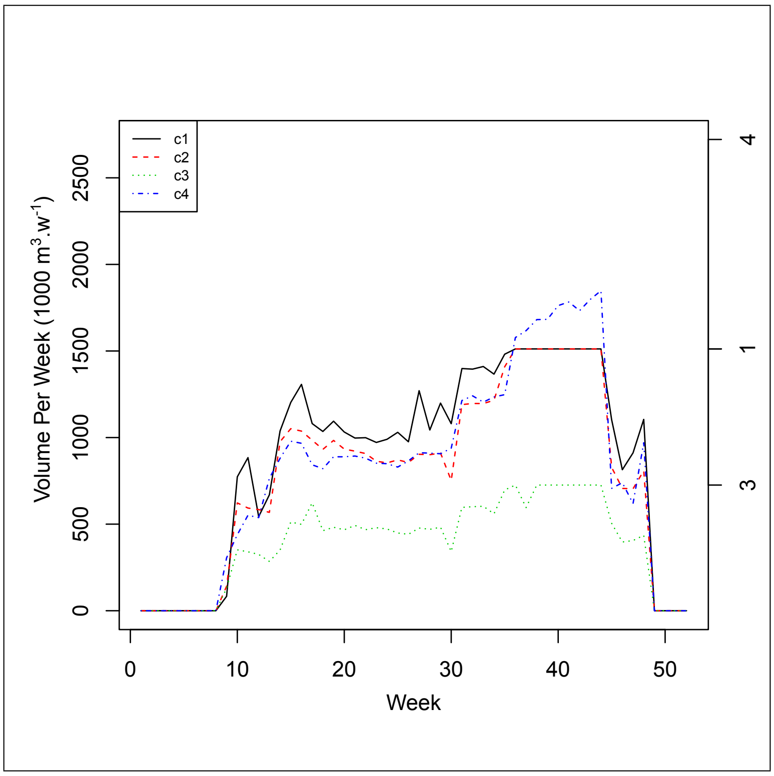


Figure 5.2.3: Cluster pump station inflows for the year.

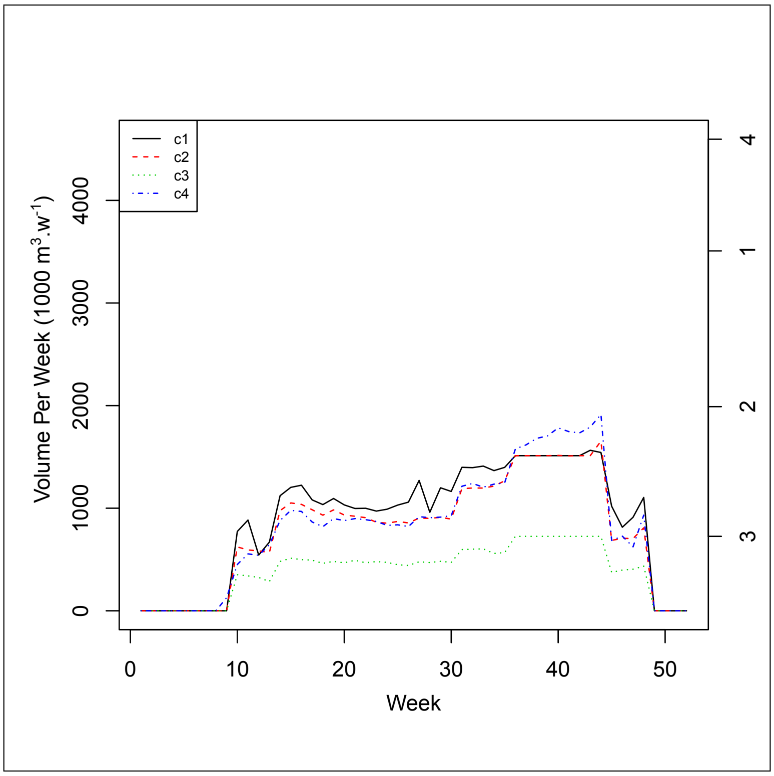


Figure 5.2.4: Cluster pump station outflows for the year.

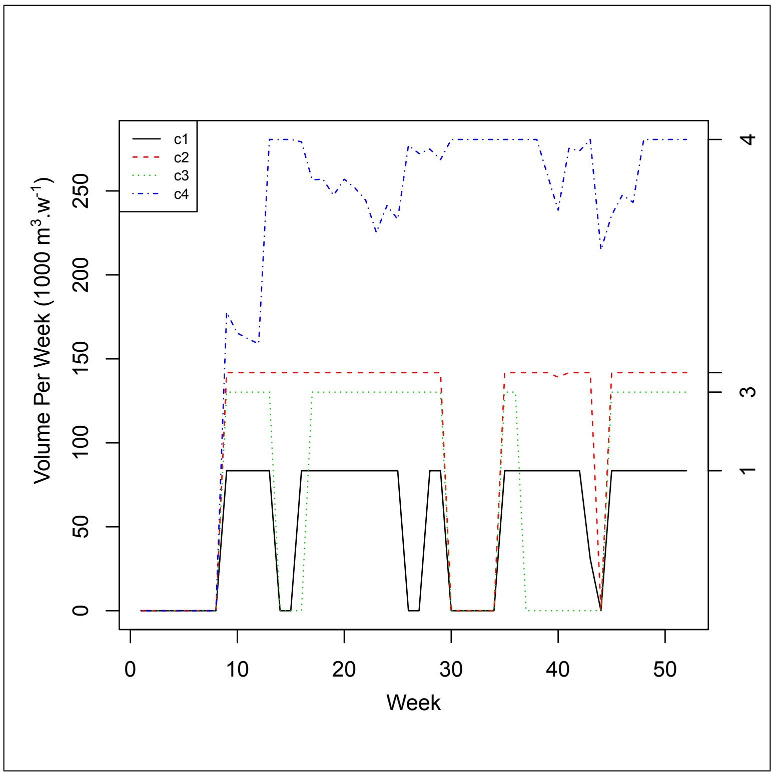


Figure 5.2.5: Cluster storage facility levels over the year.

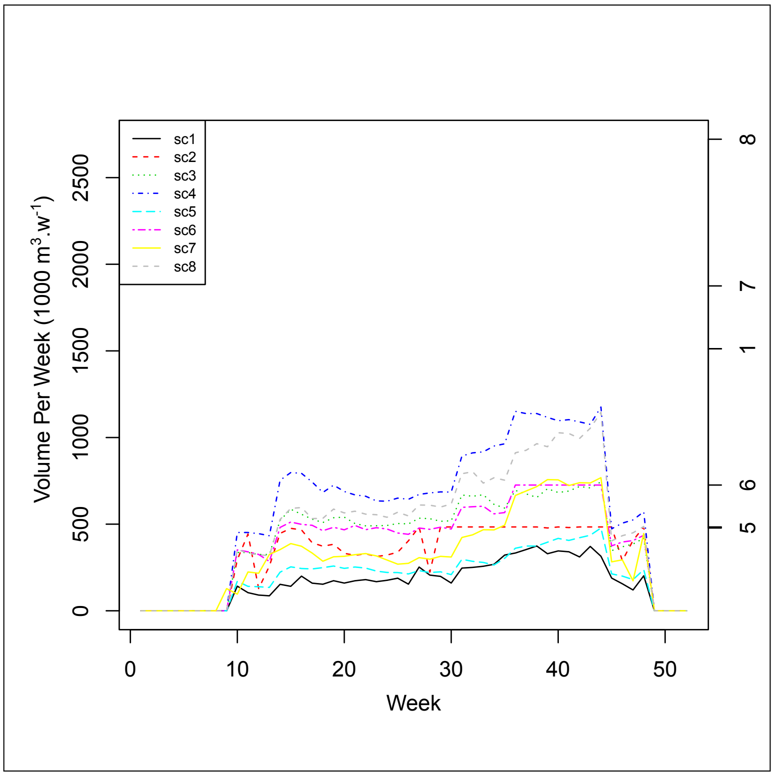


Figure 5.2.6: Sub-cluster inflows for the year.

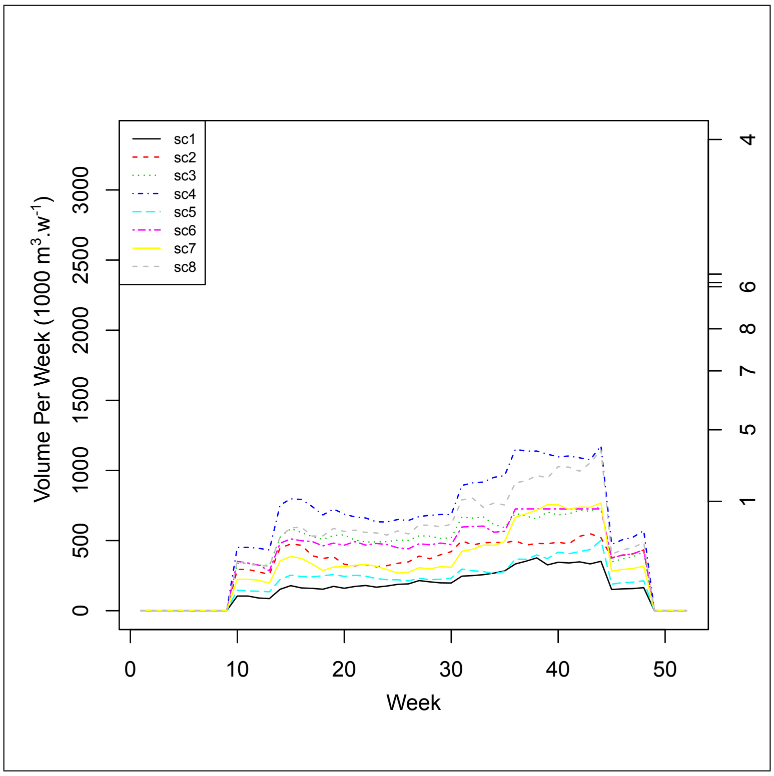


Figure 5.2.7: Sub-cluster outflows for the year.

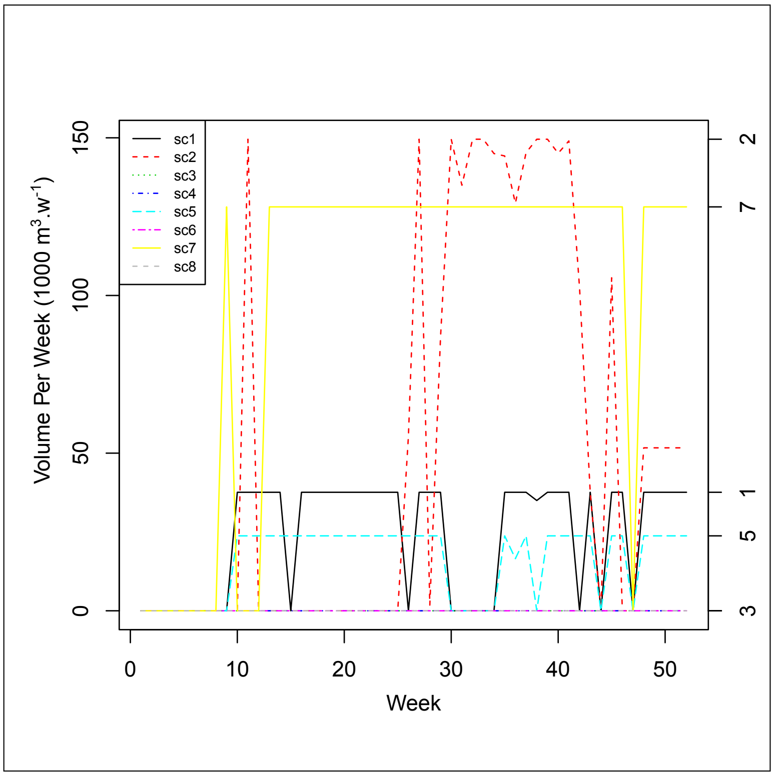


Figure 5.2.8: Sub-cluster storage facility levels over the year.

scenario consisted of 4,346 decision variables for each week (225,992 decision variables in total). The base case scenario represents the data set used above. Output of the scenario will be suppressed since only the total irrigation deficit will be analysed here. The computation time for a large farm is significantly greater than compared to the small data set containing eight fields as indicated by the times in Table 5.2.2.

Scenario 7 uses the same parameter inputs as Scenario 6 but the harvest schedule is now random. The provided harvest schedule designed by the commercial sugarcane farmer incorporates limited water availability (to an unknown extent), yield maximisation and harvest fleet travel distance minimisation. A random harvest schedule has been included to show that well-constructed harvest schedules, such as the one provided by the commercial sugarcane farmer, can reduce the irrigation deficit. A random harvest schedule will affect the water demand for the year because changing the harvest date also changes the crop water demand since the crop factors and reference crop evapotranspiration differs throughout the year (Equation (4.2.1) page 22). The field's growth and harvest dates are randomised while keeping the minimum age of the harvested fields at 39 weeks. The irrigation deficit in Scenario 7 is slightly more than double (106% greater) the irrigation deficit as in Scenario 6. This illustrates the importance of a well-constructed harvest schedule. The computation time for this Scenario is also double that of Scenario 6. The increased computation time could be attributed to a longer search in the solution space due to the harvest schedule not considering the water distribution networks limitations. Scenario 8 uses the same parameter inputs as Scenario 6 but the rain forecast is ignored. The scenario optimisation's time is similar to Scenario 6, but the irrigation deficit is 60% greater illustrating the importance of incorporating rain.

Table 5.2.2: Different Scenarios with respective irrigation deficit.

Scenario	Annual irrigation deficit in 1000 cubic metres	Computation time in seconds	Percentage change relative to base case scenario
6. Base case	16,990	10.47	
7. Random harvest schedule	34,916	22.50	106
8. Ignoring rain	27,203	11.89	60

5.2.1 Harvest Schedule Analysis

Two harvest schedules were used in the scenarios, namely, the provided harvest schedule and a randomised harvest schedule. The provided harvest schedule gave the lower irrigation deficit of the two. It would be desirable to know if it is possible to improve on the provided harvest schedule to reduce the annual irrigation deficit. To find an optimal harvest schedule accommodating the irrigation system, yield response and harvest fronts will require a new program or large modifications to the current program, however, it is possible to find out whether the provided harvest schedule can be improved by small adjustments.

Since the program and the data provided are at the weekly level, consider the possibility of decreasing the annual irrigation deficit by moving the harvest date of a field forward or backward one week in time. The experimental strategy of doing this for one field at a time is known in experimental design literature as the “one-factor-at-a-time” strategy. The preferred approach for initial experimental investigations more often involves factorial designs, fractional factorial designs or other efficient experimental designs found in mainstream literature (see, for example, [7, 31]). For this analysis, the fractional factorial design will be used to analyse the effects of shifting a subset of fields harvest dates. The method is as follows:

1. Consider a subset of 20 fields. This value corresponds approximately to all fields found within an actual planning period of two to three weeks.
2. Let the allowed manipulation of the harvest date be no more than 1 week forward (the high level) and no more than 1 week backward (low level). No shift in the harvest date will be the 0-th level.
3. Statease's Design-Expert [52] is used to estimate the effects of the 1 week shifts.

To get a better representation of the effect of the shifts, consider analysing 4 subsets of fields whose harvest dates fall into the following scenarios:

1. The early-season represented by fields harvest date nearest to week 15.
2. The early mid-season represented by fields harvest date nearest to week 25.
3. The late mid-season represented by fields harvest date nearest to week 35.
4. The late season represented by fields harvest date nearest to week 45.

The outcome of the experiment yields a regression model. The dependent variable y is the scaled annual irrigation deficit in 1000 cubic metres (1000 m^3) while the independent variables x_i for $i \in \{1, 2, ..., 20\}$ are the shifting values for the harvesting dates of the twenty fields. An x_i-value of -1 indicates a one week move backward in time of harvest, an x_i-value of 0 indicates no change in harvesting time compared to the original harvest schedule and an x_i-value of 1 indicates a move forward in time of harvest. Inherent to this approach, we expect to fail to account for the fact that moving a field forward early during the season will make it possible to move a field backward later in the season, since the early and late fields are not included in the same scenario. The regression model will be blind to that option, but we also do not expect fields to be moved too much in time due to factors other than irrigation system consideration, so this blindness seems to be a moot problem.

The regression coefficients for each scenario are listed in Table 5.2.3. The non-zero coefficients are statistically significant (p-values less than 0.001). The coefficients listed as zero are statistically insignificant according to the p-value. The outcome of the experiment is interpreted as follows:

Table 5.2.3: Regression coefficients for main effects for each scenario.

	Early-Season	Early Mid-Season	Late Mid-Season	Late-Season
Intercept	16996.33	17006.80	17005.80	17033.16
1	0.00	-2.92	-9.45	8.06
2	0.00	-3.20	-11.02	12.22
3	0.00	-5.42	-11.67	5.78
4	0.00	-7.77	-9.64	8.37
5	0.00	-3.80	-26.20	5.12
6	0.00	-3.02	-3.11	20.84
7	0.00	-2.58	-7.36	5.31
8	0.00	0.00	-8.14	7.47
9	0.00	0.00	-3.80	24.84
10	0.00	-2.30	-10.80	0.00
11	0.00	-2.83	-13.27	20.72
12	0.00	-5.92	-16.23	22.03
13	0.00	-2.83	-10.67	9.62
14	0.00	-4.11	-8.61	5.37
15	0.00	-1.83	-14.08	33.50
16	0.00	-4.80	-5.33	0.00
17	0.00	0.00	-5.11	0.00
18	0.00	0.00	0.00	0.00
19	0.00	-7.27	-8.02	0.00
20	0.00	0.00	-4.89	7.19

- The first scenario occurs during the beginning of the season. There were no statistically significant effects of moving any of the twenty fields forward or backward in time.
- The second scenario occurs during the early mid-season. Fifteen of the twenty fields showed significant effects if moved forward or backward in time. The full regression

model is

$$y = 17006.8 - 2.92x_1 - 3.2x_2 - 5.42x_3 - 7.77x_4 - 3.8x_5 - 3.02x_6 - 2.58x_7 - 0x_8 - 0x_9 - 2.3x_{10} - 2.83x_{11} - 5.92x_{12} - 2.83x_{13} - 4.11x_{14} - 1.83x_{15} - 4.8x_{16} - 0x_{17} - 0x_{18} - 7.27x_{19} - 0x_{20} \quad (5.2.1)$$

where a coefficient of 0 indicates that the corresponding field can be moved without causing any change in irrigation deficit. The regression equation says that the annual irrigation deficit is 17,006,000 m^3 when none of the harvest dates are shifted (the base case) whereas the optimisation base case annual irrigation deficit was 16,990,000 m^3. Shifting fifteen of the twenty fields one week forward results in an annual irrigation deficit of 16,945,000 m^3 according to the regression model and 16,954,000 m^3 in the optimisation. The regression model predicts the results of the optimisation model well. The weekly demand in the shifted harvest schedule is very similar to the base case, but differences can be seen when the scale is changed by focusing on the weekly demand between weeks 15 and 35 (Figure 5.2.9). The recommended harvest date shifts, from the regression analysis, thus redistributes the total weekly irrigation demand to reduce the annual irrigation deficit.

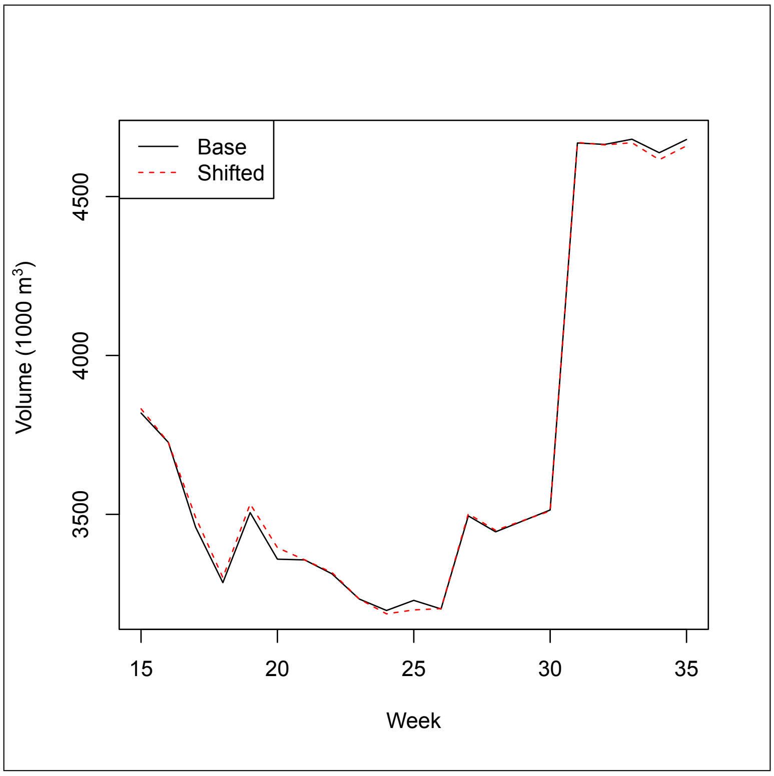


Figure 5.2.9: Total weekly demand between week 15 and 35.

It may not be appropriate to shift all the twenty fields harvest dates since the amount cut by the harvest fleet will be less at week 25 if all the fields are moved one week forward. The largest effect belongs to x_4 which in reality is field 76 (20.4 hectares) which has an original harvest date of 2014-04-11. Moving this field one week forward in time to 2014-04-18 decreases water deficits for the whole year by 7,770 m^3. To balance the tonnage to be cut for the week, one could move field 179 of 9.6 hectares (x_8) and field 400 of 10.55 hectares (x_{17}) backward in time one week which would have no additional effect on deficit, due to their zero-valued regression model coefficients. In this way, and considering only the five fields with the largest regression model coefficients within this scenario, there is potential for saving $7,770 + 7,270 + 5,920 + 5,420 + 4,800 = 31,180$ m^3 in reduced deficit over the 52 week horizon. A way to do the above more rigorously might employ some kind of constrained optimisation to maximise the savings based on maintained tonnage balance. This is not pursued in this dissertation.

- With the same approach to the late mid-season scenario the following regression model was achieved (note that the x_i's represent a totally different set of fields and are not comparable to the above scenario regression model):

$$\begin{aligned} y = 17005.8 - 9.45x_1 - 11.02x_2 - 11.67x_3 - 9.64x_4 - 26.2x_5 - 3.11x_6 - 7.36x_7 - \\ 8.14x_8 - 3.8x_9 - 10.8x_{10} - 13.27x_{11} - 16.23x_{12} - 10.67x_{13} - 8.61x_{14} - \\ 14.08x_{15} - 5.33x_{16} - 5.11x_{17} - 0x_{18} - 8.02x_{19} - 4.89x_{20} \quad (5.2.2) \end{aligned}$$

  We would of course like to move all the fields forward in time, but the tonnage balance must also be achieved. We would investigate all the possible swaps and find a combination to minimise the irrigation deficit.

- The last scenario shows similar potential as the early and late mid-season scenarios but with positive signs in the regression model coefficients.

The above methodology has illustrated that the provided harvest schedule can be changed to bring about a reduced irrigation deficit. However it is important to note that the reduction in the irrigation deficit by considering 20 fields in this illustration was marginal. The percentage decrease considering only twenty fields allowing a one week shift in the harvest date was less than 1% $(31,180/16,990,000)$ in one of the scenarios. However, by considering all fields and allowing more than one week shifts, a well-designed search algorithm may find improvements in the entire harvest schedule to significantly reduce the irrigation deficit. Reducing the annual irrigation deficit can only bring positive gains to the commercial sugarcane farmer. If monetary value can be attached to a cubic metre of water saved, small improvements as illustrated above can bring about large monetary gains.

5.3 Analysing Model Behaviour in Adverse Conditions

Section 5.2 illustrated the model behaviour using an industry sized data set under ideal conditions. It is desirable to know if the model behaves intuitively during adverse conditions, i.e. does the model distribute the limited water to the fields in a logical way, or does the model break down and perform erratically. Two important adverse conditions to consider in irrigation scheduling are extreme water shortages (drought) and faults in the water distribution network.

When the water supply is dependent on catchments from natural water sources such as lakes or streams, dry years have detrimental effects on available water and analysing the effect of a possible drought becomes crucial in irrigation scheduling. Faults in the system correspond to pump stations not operating at capacity due to damaged equipment such as leaking pipes, damaged motors, or damaged storage facilities etc. During a drought, a farmer would most likely reduce the number of grown crops to concentrate the limited water on a smaller number of crops, rather than irrigating thinly. When implementing a drought or a pump station malfunction in the model, it is expected to see fields under-irrigated and more dependent on the usage of the storage facilities. This section illustrates two adverse conditions implemented in the model to determine if the model is applicable in adverse conditions or if adjustments to the model parameters are required to obtain the desired outcome.

The two conditions have to be translated into the model. A drought would imply a shortage of water supply to the entire farm and would thus be captured by reducing the inflow capacity at the main pump station as well as a decrease in the rain forecast in the model. As an illustration, the scenario will be run restricting the inflow capacity at the main pump station to 50% of the base case inflow capacity and reducing the rain forecast by 50%. Faults in the components of the irrigation system is captured by changing the respective pump station inflow or storage facility capacity. As an illustration, consider a fault at the first cluster pump station causing the pump station to operate at 25% of the inflow capacity of the base case. To analyse the effect of a severe water shortage and faults in the system, the test data set will be used so that all the fields may be observed. The two conditions described above are listed in Table 5.3.1 together with their respective annual irrigation deficit.

Scenario 9 indicates that reducing the inflow capacity at the main pump station by 50% does not result in a 50% increase in the irrigation deficit but rather a large 407.915% increase. Due to the decrease in the rain forecast, the irrigation volume is now positive

Table 5.3.1: Scenarios with respective annual irrigation deficit.

Scenario	Annual irrigation deficit in cubic metres	Computation time in seconds	Percentage change relative to base case scenario
9. Main pump station inflow capacity and rain forecast reduced by 50%	1,086,415	0.98	408
10. Main pump station inflow capacity and rain forecast reduced by 50% using $\mathbf{c} = [0.5, 0.005, 0.0001, -0.1]$	1,086,407	0.98	408
11. Cluster c_1 pump station inflow at 25%	579,245	0.54	171

during the entire year as indicated by Figure 5.3.1. Figure 5.3.2 indicates that the inflows at the main pump station begins earlier than in the base case scenario and ends later. The main pump station inflows and its storage facility are at a longer capacity usage than compared to the base case. Since the base case has a higher inflow capacity at the main pump station it had more flexibility when irrigating the fields in the early period and could thus delay the filling of the storage facilities. The optimisation hedges by filling the storage facilities earlier than compared to the base case as shown in Figures 5.3.3 and 5.3.4.

Figure 5.3.5 plots the supplied irrigation volume to the fields. It is clear that irrigation deficits occur before and during the high demand period. Certain fields are being completely ignored during the year, i.e. not receiving any of its irrigation volume demand. This behaviour is not intuitive. The objective function parameters were not chosen with the intention to achieve this outcome but rather to distribute the limited water among the competing fields. In this case of a field continuously receiving none of its required irrigation volume, such as field f_2, it may be desirable to remove such a field from the crop plan since it may die off during extreme water stress. Despite an undesirable application of limited water, the program still indicates that irrigation deficits will occur during the year.

It is possible to choose different benefit function parameters (controlled by the vector $\mathbf{c}$ and $\mathbf{b}$) to obtain the desired result. Scenario 10 was run using the same conditions as Scenario 9 but the vector $\mathbf{c}$ is set to $\mathbf{c} = [0.5, 0.005, 0.0001, -0.1]$. Figure 5.3.6 plots the supplied irrigation for the year. This baseline coefficient vector was chosen to avoid any field being classified as having a severe irrigation deficit. Figure 5.3.6 indicates that this was indeed achieved and that benefit function parameters can be changed to achieve a more intuitive distribution of limited water. Note that the annual irrigation deficit of Scenarios 9 and 10 are the same, but the distribution of the limited water is different. Since the benefit functions are strictly increasing and the optimisation always seeks to supply all the available water, it is possible to observe the same or similar irrigation deficits.

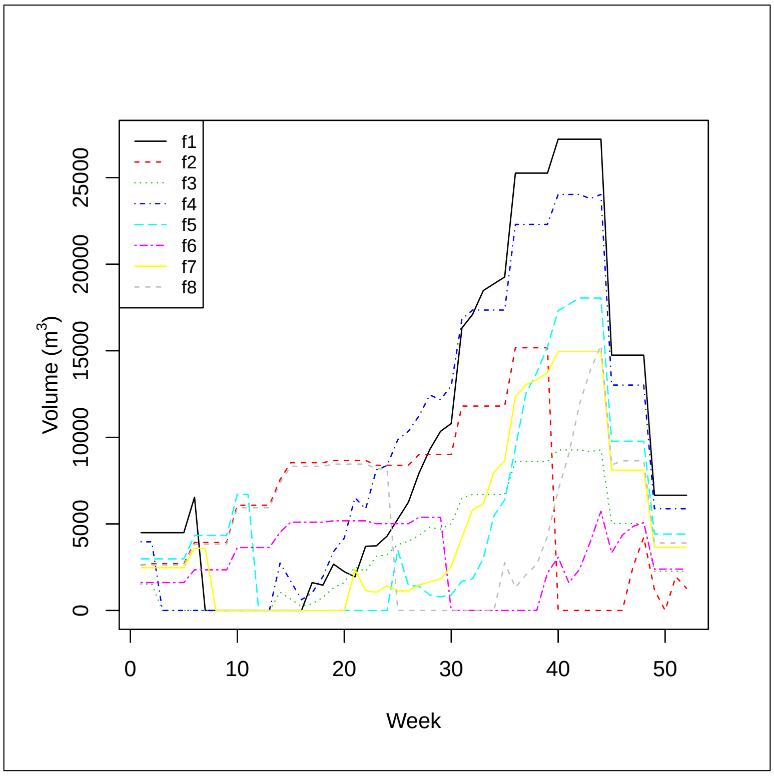


Figure 5.3.1: Fields irrigation volume demand for Scenario 9.

Scenario 11 is the condition with a fault at cluster pump station c_1 which services fields f_1, f_2 and f_3. These fields should be the only fields negatively affected by the fault and the excess capacity now available at the main pump station should be used at the other fields where deficits had existed in the base case. Figure 5.3.7 plots the main pump station usage. The inflows before the high-demand period are smaller than in the base case scenario since the inflow capacity for cluster pump station c_1 is smaller. Since the inflow capacity at cluster c_1 has been reduced, the would-have inflow delivered to cluster pump station c_1 from the main pump station is diverted to the remaining clusters and storage facilities. Cluster pump station c_1's inflow is at capacity for the entire water demand period as indicated by the long horizontal line in Figure 5.3.8. Cluster pump station c_1's storage facility is filled to capacity earlier than the base case. Figure 5.3.10 plots the supplied irrigation to the fields for the year with the 25% inflow capacity at cluster pump station c_1. The fields serviced by cluster pump station c_1 are indeed affected by the fault and the fields with no relation to cluster pump station c_1 are unaffected. Field f_2 receives none of its required irrigation

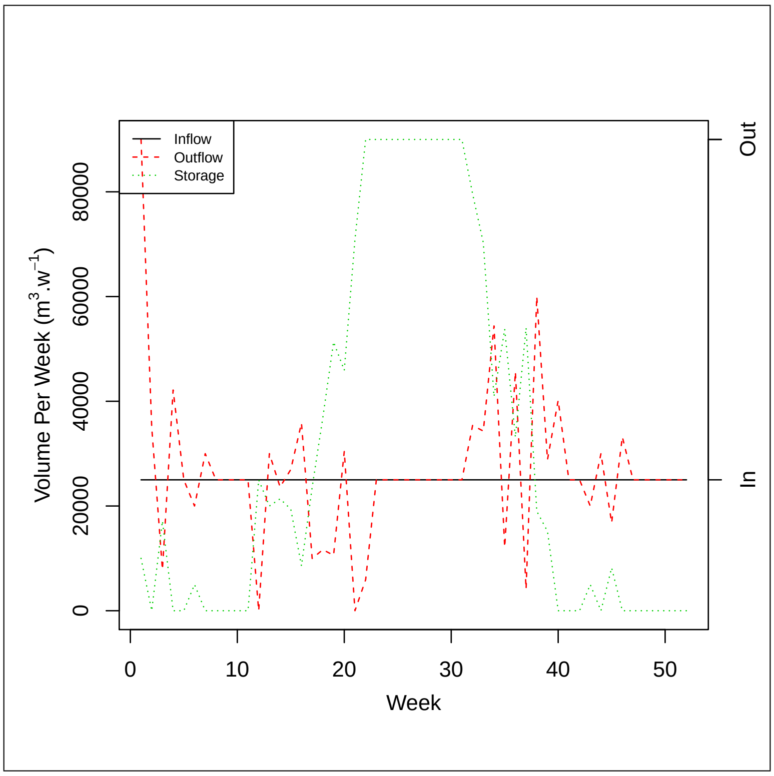


Figure 5.3.2: Main pump usage at 50% inflow capacity of Scenario 9.

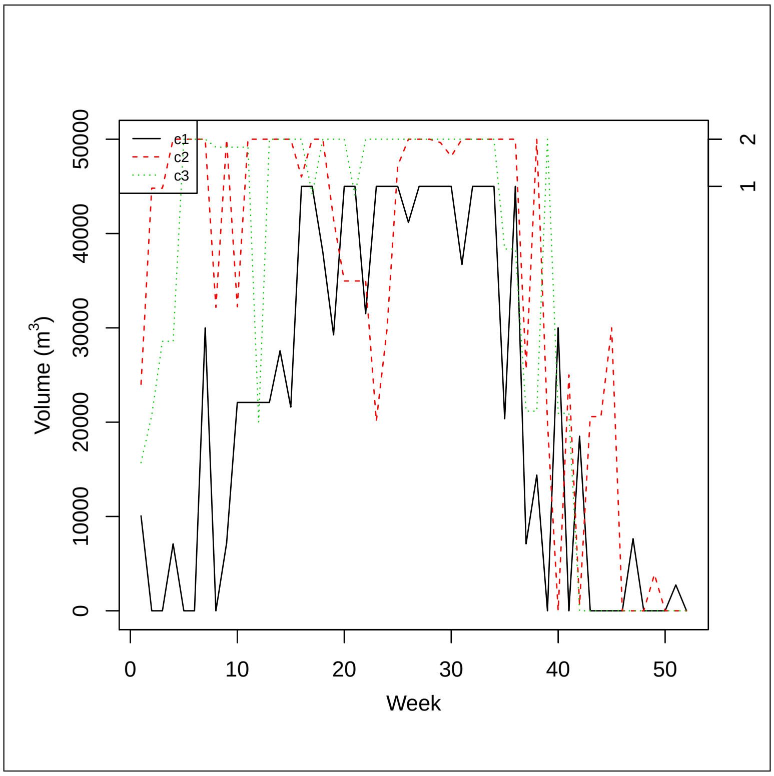


Figure 5.3.3: Cluster pump station storage capacities of Scenario 9.

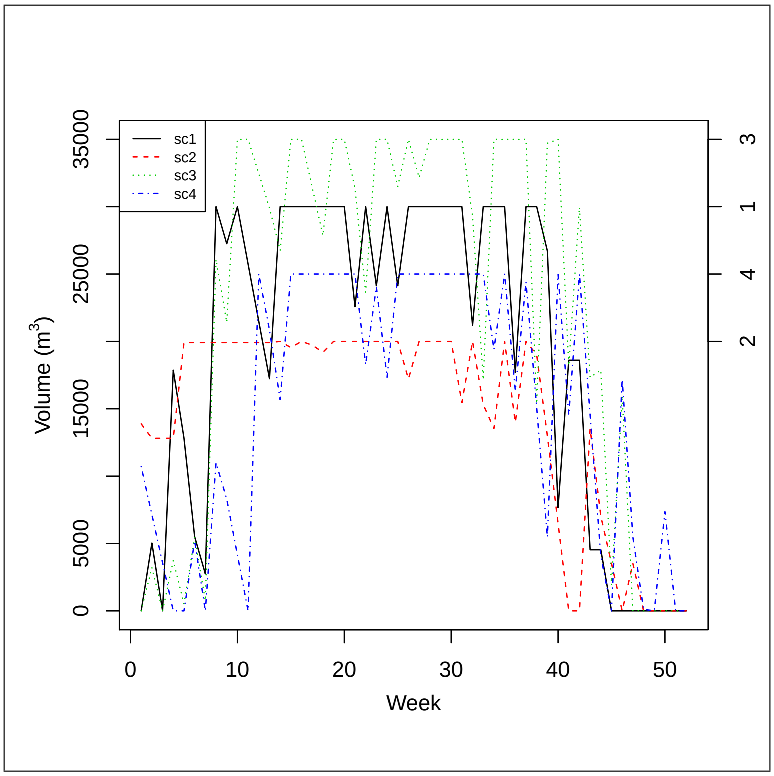


Figure 5.3.4: Sub-cluster pump station storage capacities of Scenario 9.

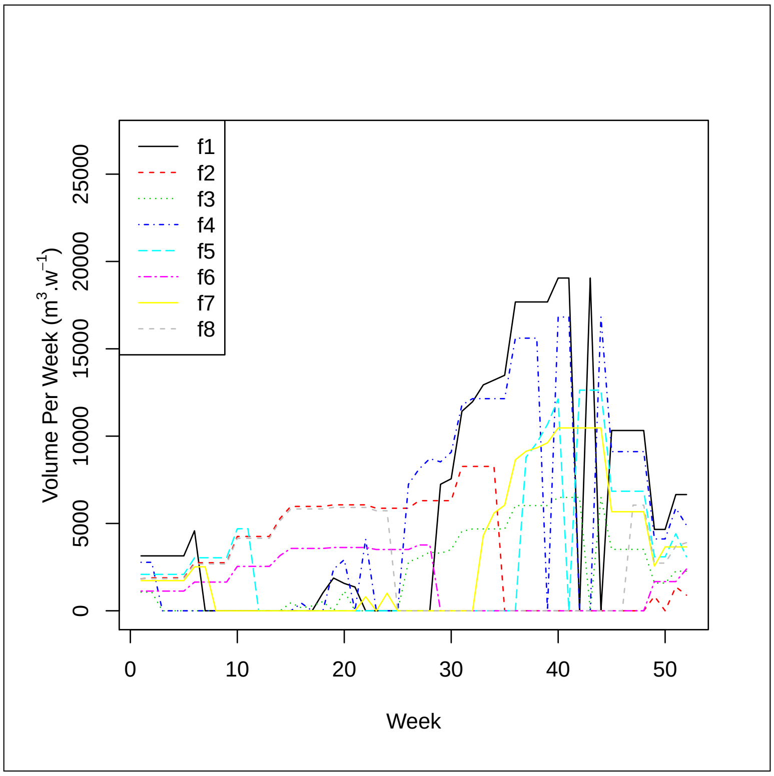


Figure 5.3.5: The supplied irrigation for each field of Scenario 9.

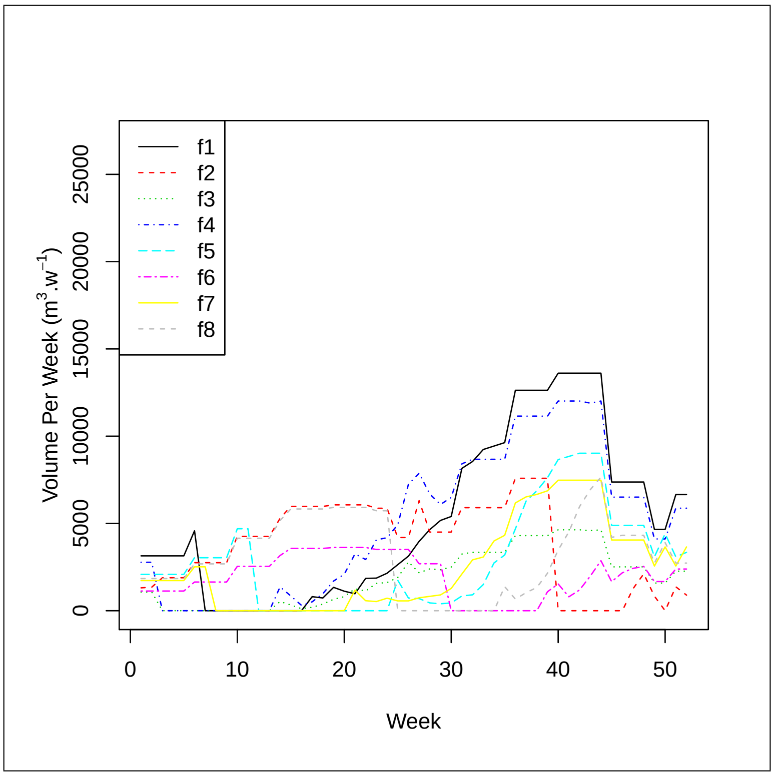


Figure 5.3.6: The supplied irrigation for each field of Scenario 10.

volume after week 35. Fields f_1 and f_3 experience an on-off irrigation application. In this case where field f_1, f_2 and f_3 are affected by the fault, it would be desirable to remove field f_2 from the harvest schedule since it may die off due to prolonged stress and distribute the limited water among the remaining fields. The distribution of limited water to the fields serviced by cluster pump station c_1 is thus counter-intuitive in this case.

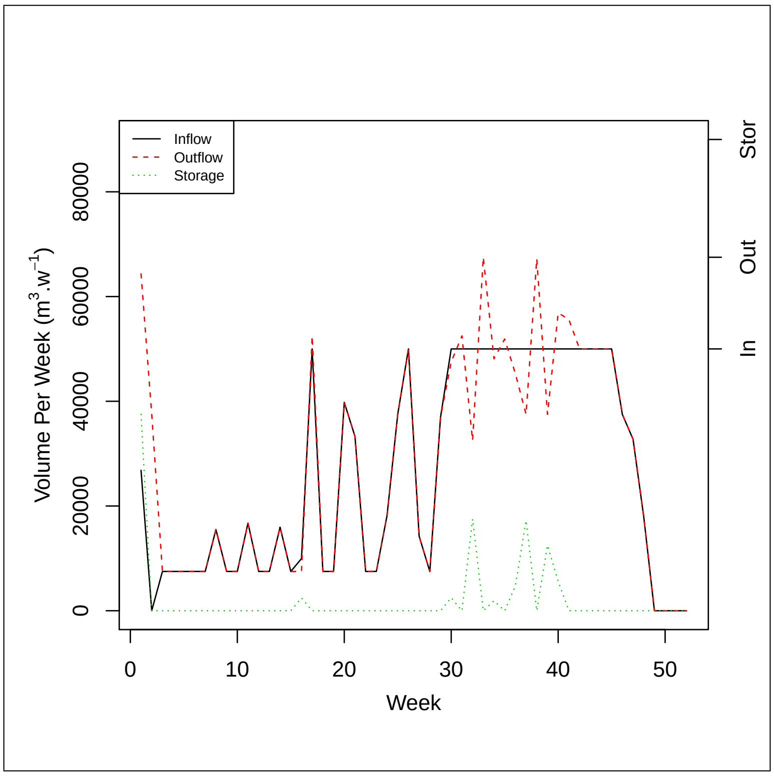


Figure 5.3.7: Main pump station usage for Scenario 11

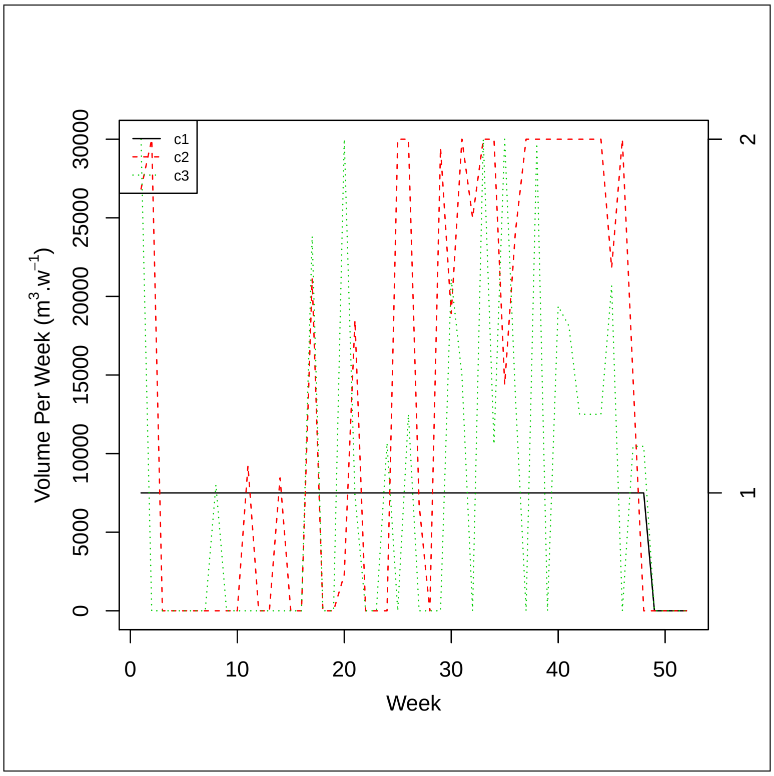


Figure 5.3.8: Cluster pump stations inflows for Scenario 11

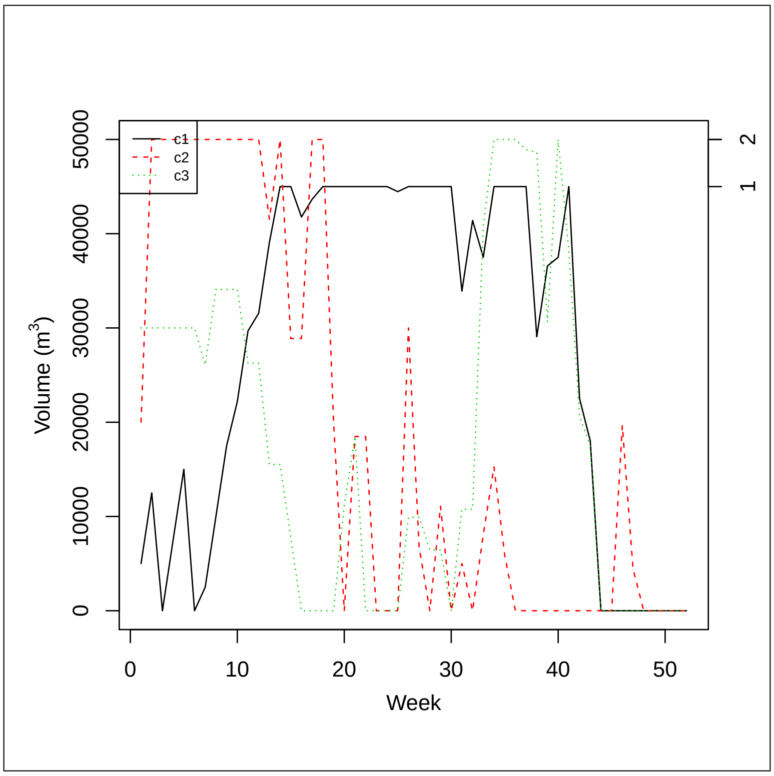


Figure 5.3.9: Cluster pump station storage facilities level's for Scenario 11

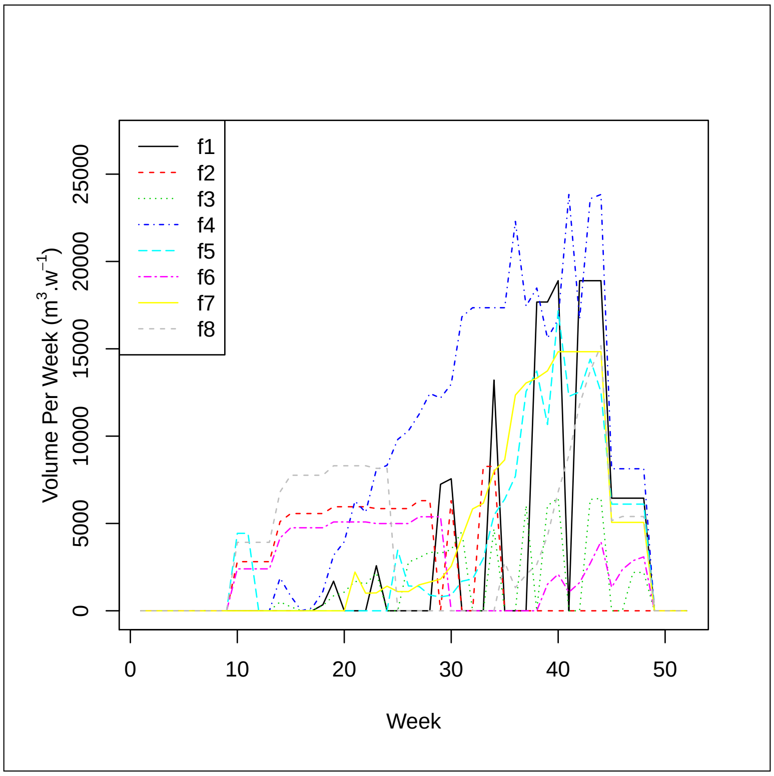


Figure 5.3.10: Supplied irrigation volumes for each field for Scenario 11

Chapter 6

Program Limitations

Relaxing the assumptions that were made in Chapter 3 enables desirable extensions to improve the model accuracy. This section discusses the extensions that can be made to improve the mathematical model and program.

1. Time periods. The time periods in the model has been generalised to time periods t of the time horizon $\mathcal{T}$. The program uses weekly time periods due to data limitations and to reduce computation time. Although the literature supports time periods of seven to ten days operational planning time intervals, the program can be easily adapted to a general operational planning time interval if daily data are supplied. To incorporate arbitrary time periods a positive integer or a string variable can be included as a required parameter of the program. Together with this parameter and the supplied data, the format of the data can be exploited to obtain daily or average daily measurements/capacities values. If the data is not supplied in the time periods required then an average can be determined from the data at the cost of losing the specific peculiarities that occur in that time interval. Once the data has been converted to a common time period, the MILP can be formulated and solved using these time periods.

2. Water demand. The accuracy lost by not differentiating between cultivar and plant/ratoon crop can be easily retrieved by requiring two additional columns in the field data set, one for cultivar type and another indicating whether plant or ratoon crop is used. This will require separate data matrices for each crop factor and plant/ratoon crop. The cultivar and seed/ratoon type then indicates which matrix the program should use when determining the fields' water demands. A simpler solution is to use modifiers on a base cultivar type together with a plant/ratoon crop data matrix to adjust the factors.

   Additionally, to incorporate the crop type specifications, a more in-depth determination of the irrigation demand would include the characteristics of the environment for each particular field. These characteristics could include the field position and soil type. The position of the field has an effect on the settling of the supplied irrigation. Fields situated at the bottom of hills may encounter the runoffs from neighbouring fields. The soil type quantifies the soil holding capacity and drainage characteristics. These characteristics may also be captured by including an indicator for each field in the field data set and adjusting the field water demand.

3. Water distribution network. The program can be made more general by allowing arbitrary water distribution network architecture. This will allow the model to accurately capture the farms network of pump station, pipelines, canals, furrows and storage facilities. The challenge involved to allow a general irrigation network architecture is to develop the code for formulating the continuity equations. In the program, the continuity equations are formulated at defined locations due to the assumptions made on the irrigation network architecture. To incorporate a general water distribution network the field data set would require additional columns to capture all the pump station service relations.

   The homogeneous pump inflow capacities throughout the year can also be relaxed by including the pump inflow capacities for each time period, which can also serve as a means to incorporate pump station faults during the year. The pump station inflow capacities are required in matrix format. Using multiple data matrices can be memory intensive but can be significantly reduced by using sparse matrices. Having flow matrices makes formulating continuity equations simpler since service relations are preserved in such a format. To this end, mathematical programming modelling languages, such as GAMS, AMPL, LINGO, AIMMS, may be of use.

4. Storage facilities. A storage facility's water volume is only determined by the amount reserved at the pump station. Depending on the construction of the storage facility, it may be the case that rain increases the water volume or that heat causes the water levels to decrease through evaporation. By including an additional column in the storage facility data to indicate which storage facilities are open or closed, the water level can be affected by external means. Since the program uses optimisation methods to simulate the water distribution network and not Monte Carlo simulation, the storage facility water volume cannot easily be decreased or increased during the optimisation. The effect will have to be incorporated by adjusting the storage facility's capacity for each time period or by adjusting the pump station inflow capacity.

5. Sucrose yields. The benefit function was designed to protect the fields from being under-irrigated. An objective function which aims to maximise the actual crop su-

crose yield would be desirable, i.e. a benefit function which quantifies the amount of sucrose yield in relation to the irrigation volume supplied for each week of the year. The literature surveyed does not provide this type of theoretical relationship. A possible solution may be to conduct a regression analysis on the sucrose yields, where the supplied irrigation volume for each time period serves as the explanatory variables. This requires the actual data on the sucrose yields and the supplied irrigation volume to the crop for several irrigation schedules. Such as the research conducted in Thompson [55] and Shih [48], but a more specific form to be used in this optimisation algorithm. A possible form may include a volume coefficient for each amount of water supplied during the week of the year to quantify the sucrose yield for a crops life cycle.

6. Harvest fronts. As discussed in Chapter 3, the travelling distance the harvest fronts make are an important aspect of operational planning so that incorporating the travel distance is desirable. A direct application of the model would be to use the program in a genetic algorithm, such as in Cai *et al* [9]. This would mean evaluating the travel distance of the proposed harvest schedule by assigning coordinates to the fields and calculating the travel distance made by the harvest fronts adhering to the harvest schedule. The objective of the optimisation would be to minimise the harvest fronts travel distance adhering to the water distribution network limitations. The developed program would thus be a sub-program of the genetic algorithm that will need to be called multiple times to find a good harvest schedule. In such an implementation the running time of a single call of the optimisation becomes extremely important.

7. The stochastic nature of the weather that is brought about in the reference evapotranspiration and rain has been incorporated in the model by using long-term averages. The program allows one data set estimate for each of the rain forecast and the reference crop evapotranspiration. These values representing forecasts for the following year may be point estimates or determined as a long-term average. It may be desirable to obtain a distributional form of these occurrences and allow the data set to vary according to the distribution when a single run is performed instead of using averages. Pump station faults can also be treated as occurring randomly based on previous occurrences. This can be achieved by choosing the number of faults and position of the faults according to well-chosen discrete distributions.

8. The program requires extensive data of the crop factors and reference crop evapotranspiration for field irrigation requirement calculation. The crop factor is required for each start week till the harvest point. If these data sets are not available, the crop factor estimates provided from the FAO may be used and the estimation method for reference crop evapotranspiration described in Allen [2] can be used. This leads to

the desirable aspect of including a default data set to be used in the program if no data is supplied.

As discussed by Mysiak *et al* [32], developing a DSS driven by a user friendly interface is an important feature for a DSS to be absorbed into practice. The developed model is implemented at the console level with all model parameter controls at the final command. The user interface should use the final command to design the user interface. The user interface should be intuitive for someone with little to no knowledge of programming. The user should be able to import data and set the desired parameters of the model without any difficulty. The results of the program should guide the user to adjust the model parameters to achieve better results. The output of the model should be clearly summarised and the consequences highlighted. Only the most important options should be available not to distract or confuse the user. The R-package `gWidgets` is one package available for creating a graphical user interface (GUI) in R.

The ideas discussed above have been on extending the developed program. A different approach to solving the water allocation problem and incorporating the above mentioned extensions would be combining the developed model with a simulation model such as in Berhe *et al* [6]. Instead of running the optimisation for the entire year, the optimisation can be run for smaller time periods, say for one week. After each optimisation the necessary rules to execute the program extension can be implemented. One such rule would be to link the the next simulation to the previous simulation by relaying the remaining storage volume. This approach, although more intuitive to implement, requires strict programming but will allow that all the extensions to be included. However, this will not implement the optimal usage of the limited water because it does not consider the entire problem at once but rather subsets of the problem.

Chapter 7

Conclusion

The objective of this dissertation has been to develop a mathematical model to aid in irrigation water supply and demand management. The dissertation has presented a mathematical model and shown that the model could be used in practice.

Chapter 2 provided an overview of the literature to develop a water allocation model. The literature focus has largely been on regional water allocation rather than at the field level. The regional water allocation literature contained desirable aspects for a water allocation model at the farm level. Both simulation and optimisation models have been used for water allocation models, but optimisation models were chosen because of the ability to hedge against dry periods by considering the best times to release and store water. Among the available optimisation methods, linear programming with integer variables was chosen to model the water distribution network. Previous research in network flow optimisation have been successful in modelling water allocation problems using linear programming. The MILP modelling technique and the widely available software packages made it possible to incorporate the non-linearities of the sugarcane responses to water stress.

Sugarcane water demand determination is well defined in the literature. Many reliable methods have been verified in the literature and are available for different farming environments. In the model formulation, the crop factor method was used together with the commercial sugarcane farmer's crop factors to construct the irrigation schedule.

To use an optimisation water allocation model, the objective function to drive the water distribution had to be defined. The effects of water stress on sugarcane provided a bases for constructing the objective function. Conveniently, the effects of water stress on sugarcane has been well documented in the literature detailing sucrose yields and general crop health indicators in response to the amount of water supplied. The documentation, at the simplest level, suggest that a priority rule exists between competing crops for limited water. The priority rule and the effect of the amount delivered to a field at any point in time was used to create the field's benefit function. The benefit functions were non-linear, and since the objective function was a sum of the benefit functions, the objective function was non-liner as well.

Chapter 4 described the mathematical model and program in detail. The literature provided a foundation for the mathematical model of the water distribution network. Continuity equations used in regional water allocation models allowed the limited water and flow dependences to be captured realistically by the model. The developed model was solved by a commercial solver and implemented in a free programming language. A description of the program data inputs and parameters, as well as the format of the data illustrated in Appendix A and B has been provided so that the model can be recreated by any interested person. The program was made as intuitive as possible so that the program has a higher chance of industry absorption and the possibility of further development. The program was also created with a degree of flexibility so that it can be applied to different water distribution networks.

Three sub-objectives have been created to achieve the objective of the dissertation, namely:

1. Determine the severity of irrigation deficits by comparing a field's water demand and the actual water supplied to the field.

2. Identify points in time when the water distribution network may benefit from structural intervention to reduce the irrigation deficit.

3. Identify changes in the harvest schedule to reduce irrigation deficits.

Chapter 5 served the purpose of exploring these sub-objectives to ultimately aid in improving irrigation water supply and demand management by mathematical modelling.

1. Section 5.1.1 (sub-objective 1): A small test data set was used to compare the supplied irrigation volume to the crop's volume demand. A small data set was used so that the analysis could be presented and analysed in its entirety. The program results were plotted as the optimised flows for each pump station, field, and storage facility. At this level, it was easy to identify pump stations operating at capacities and storage or pump stations being under-worked. At the field level, the fields water demand was compared to the supplied demand. The benefit function performed as intended; the limited water was distributed amongst the fields giving priority to certain fields depending on their age. During the high demand period, there were irrigation deficits but, the irrigation deficits were spread amongst the fields. The program results allowed for close inspection at the field or pump station level. Because the optimisation is done for an entire year, strategic planning can be conducted for the year. Fields of concern may be revealed from a run of the model and procedures to mitigate the irrigation deficits may be discussed.

2. Section 5.1.2 (sub-objective 2): By identifying periods of continuous operation at capacity levels, it was possible to determine whether the water distribution network could benefit from structural intervention. Continuous horizontal lines in the flow figures of Section 5 indicated possible components that could benefit from intervention. Given the freedom of an optimisation model, it was easy to change the parameters and input data. It was shown that increasing the relevant inflow capacities and storage capacities had positive effects. The results of the changes were clearly evident in the flow figures. The methodology shows that simple analysis can be made during strategic planning to determine where and when issues will arise and whether structural intervention would be beneficial.

3. Section 5.2 (sub-objective 3): Section 5.2 served the dual purpose of illustrating that the program was able to handle an industry sized problem and that the provided harvest schedule could be improved. The program was able to handle the data set without any issues. Similar analysis as in Section 5.1 was carried out illustrating that the program could be used for strategic planning in an industry setting.

   A natural question arises given the ability to vary a harvest schedule in the program; is it possible to change the harvest schedule to reduce the irrigation deficit? Instead of modifying the program and model, experimental design techniques was used to identify fields whose harvest dates could be shifted to reduce the annual irrigation deficit. The regression results indicated that there were field harvest dates that could be shifted to reduce the irrigation deficit. This reduction turned out to be insignificant when considering only a subset of fields. However, the methodology hints that if more fields are considered with more flexibility in the date shifting, a greater reduction in the irrigation deficit may be obtained.

Section 5.3 was included for completeness; to further assess the model applicability in a practical setting by analysing sensitivity to drastic scenarios using data inputs. In some cases, fields affected by the severe changes received an unintuitive irrigation volume supply during the year. This illustrated some issues with the model, but using different benefit functions, the problem was corrected.

The research in this dissertation should in future be extended to understand the problem in more detail and model the farm operations more accurately. The quantification of the irrigation deficit in the model, although motivated by research findings, needs to be addressed to achieve a more accurate consequence of not meeting the required irrigation demand. A desirable quantification would be in terms of real monetary value or actual yields. The sucrose yield in terms of the supplied water during each week of the crop life cycle was not available in the literature. Determining such a relationship and incorporating it into a model such as in this dissertation will provide a model that would more accurately distribute limited water. By considering more aspects of the commercial sugarcane farmer's problems it is inevitable that future research pursues multiple objective optimisation techniques in order to achieve a more accurate model of the commercial sugarcane farm operations.

Bibliography

[1] Ahuja RK, Magnanti TL & Orlin JB, 1988, *Network flows*, Massachusetts Institute of Technology.

[2] Allen RG, Pereira LS, Raes D & Smith M, 1998, *Crop evapotranspiration - Guidelines for computing crop water requirements*, FAO Irrigation and Drainage, paper 56.

[3] Allen RG, tasumi M, Morse A & Trezza R, 2005, *A Landsat-based energy balance and evapotranspiration model in Western US water rights regulation and planning*, Irrigation and Drainage Systems, vol. 19, pp. 251-268.

[4] Auspitz JL, Gasparski WW, Mlicki MK & Szaniawski, , *Praxiologies and the philosophy of economics*

[5] Azamuthulla HMd, Wu FC, Ghani AAb, Narulkar SM, Zakaria NA & Chang CK, 2008, *Comparison between genetic algorithm and linear programming approach for real time operation*, Hydro-environment Research, vol. 2, pp. 172-181.

[6] Berhe FT, Melesse AM, Hailu D & Sileshi Y, 2013, *MODSIM-based water allocation modeling of Awash River Basin, Ethiopia*, Catena, vol. 109, pp. 118128.

[7] Box G, Hunter WG & Hunter JS, 1978, *Statistics for Experimenters*, John Wiley and Sons.

[8] Bras RL & Cordova JR, 1981, *Intraseasonal Water Allocation in Deficit Irrigation*, Water Research, vol, 17, pp.886-874.

[9] Cai X, McKinney DC & Lasdon LS, 2001, *Solving nonlinear water management models using a combined genetic algorithm and linear programming approach*, Advances in Water Resources, vol. 24, pp 667-676.

[10] Carr MKV Knox JW, 2011, *The water relations and irrigation requirements of sugar cane (saccharum officinarum): A review*, Experimental Agriculture, vol. 47, pp. 1-25.

[11] Cuenca RH, 1989, *Irrigation system design: An engineering approach*, Prentice-Hall, Englewood Cliffs, N.J.

[12] Cunha Ma C & Sousa J, 1999, *Water Distribution Network Design Optimization: Simulated Annealing Approach*, Water Resource Planning and Management, vol. 125, pp. 215-221.

[13] Donaldson RA & Bezuidenhout CN, 2000, *Determining the maximum drying off periods for sugarcane grown in different regions of the South African Industry*, Procedures from the South African Sugar Technology Association vol. 74, pp. 163-166.

[14] Dreyfus SE, 1956, A comparison of linear programming and dynamic programming.

[15] Elhaddad A & Garcia LA, 2008, *Surface energy balance-based model for estimating evapotranspiration taking into account spatial variability in weather*, Irrigation and Drainage Engineering, vol. 134, pp. 681-689.

[16] Ellis RD, Wilson JH & Spies PM, 1985, *Development of an irrigation policy to optimise sugar production during seasons of water shortage*, Proceedings of the South African Sugar Technologists' Association, June 1985.

[17] Food and Agriculture Organization of the United Nations, 2014, *Chapter 3: Crop water needs*, [online], [Cited November 15$^{\text{th}}$, 2014], Available from `http://www.fao.org/docrep/S2022E/s2022e07.html`.

[18] Food and Agriculture Organization of the United Nations, 2014, *Crop Water Information: Sugarcane*, [online], [Cited November 15$^{\text{th}}$, 2014], Available from `http://www.fao.org/nr/water/cropinfo_sugarcane.html`.

[19] Frevert DK, Hill RW, Hill RW & Braaten BC, 1983, *Estimation of FAO evapotranspiration coefficients*, Journal of Irrigation and Drainage Engineering vol. 109, pp. 265-270.

[20] Gurobi Optimization and Incorporation, 2015, *Gurobi Optimizer Reference Manual*, `http://www.gurobi.com`.

[21] Hagos H, Worku W & Takele A, 2014, *Effect of drying off period and harvest age on quality and yield of ratoon cane(saccharium officinarium l.)*, Advances in Crop Science and Technology, vol. 2, issue 3, 1000133.

[22] Hill RW & Keller J, 1980, *Irrigation system selection for maximum crop profit*, WCS 1978 proceding of the 10$^{\text{th}}$ conference on winter simulation, vol. 2, pp. 495-505.

[23] Inman-Bamber NG & McGlinchey MG, 2003, *Crop coefficient and water-use estimates for sugarcane based on long-term Bowen ratio energy balance measurements*, Field Crops Research, vol. 83, pp. 125-138.

[24] Inman-Bamber NG, 2004, *Sugarcane water stress criteria for irrigation drying off*, Field Crops Research vol. 89, pp. 107-122.

[25] Inman-Bamber NG & Smith DM, 2005, *Water relations in sugarcane and response to water deficits*, Field Crops Research vol.92, pp. 185-202.

[26] Inman-Bamber NG, Lakshmanan P & Park S, 2012, *Sugarcane for water-limited environments: Theoretical assessment of suitable traits*, Field Crops Research vol. 134, pp. 95-104.

[27] Karterakis SM, Karatzas GP, Nikolos IK & Papadopoulou, 2007, *Application of linear programming and differential evolutionary optimization methodologies for the solution of coastal subsurface water management problems subject to environmental criteria*, Journal of Hydrology, vol. 342, pp. 270-282.

[28] Kirda C, *Deficit irrigation scheduling based on plant growth stages showing water stress tolerance*, Food and Agriculture Organization.

[29] Matrosov ES, Harou JJ & Loucks DP, 2008, *A computationally efficient open-source water resource system simulator e Application to London and the Thames Basin*, Environmental Modelling & Software, vol. 26, pp. 1599-1610.

[30] Merry RE, 2003, *"Dripping with success": The challenges of an irrigation redevelopment project*, Irrigation and Drainage Engineering, vol. 52, pp. 71-83.

[31] Montgomery DC, 2009, *Design and Analysis of Experiments 7th Edition*, John Wiley and Sons (Asia).

[32] Mysiaka J, Giupponib C & Rosatoc P, 2005, *Towards the development of a decision support system for water resource management*, Environmental Modelling & Software, vol. 20, pp. 203-214.

[33] Noory H, Liaghat AM, Parsinejad M & Haddad OB, 2012, *Optimizing Irrigation Water Allocation and Multicrop Planning Using Dicrete PSO Algorithm*, Jornal of Irrigation and Drainage Engineering, vol. 138, pp. 437-444.

[34] Olivier FC, Donaldson RA & Singels A, 2006, *Drying off sugarcane on soil with low water holding capacity*, Procedures from the South African Sugar Technology Association vol. 80, pp. 183-187.

[35] Qureshi ME, Wegener MK, Harrison SR & Bristow KL, 2001, *Economic evaluation of alternative irrigation systems for sugarcane in the Burdekin delta in north Queensland, Australia*, Water Resources Management, pp. 47-57.

[36] R Core Team, 2014, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical computing, Vienna, Austria, http://wwww.R-project.org.

[37] Robertson MJ & Donaldson RA, 1998 dunno which one, *Changes in the components of cane sugar sucrose yield in response to drying-off og sugarcane before harvest*, Field Crops Research vol. 55, pp. 201-208.

[38] Reca J, Roldán J, Alcaide M, López R & Camacho E, 2001, *Optimisation model for water allocation in deficit irrigation systems I. Description of the model*, Agricultural Water Management, vol. 48, pp. 103-116.

[39] Reca J, Roldán J, Alcaide M, López R & Camacho E, 2001, *Optimisation model for water allocation in deficit irrigation systems II. Application to the Bémbezar irrigation system*, Agricultural Water Management, vol. 48, pp. 117-132.

[40] Gurobi Optimization and Incorporation, 2015, *gurobi: Gurobi Optimizer 6.0 interface*, R package version 6.0-2, http://www.gurobi.com.

[41] Bates D & Maechler M, 2013, *Matrix: Sparse and Dense Matrix Classes and Methods*, R package version 1.0-12, http://CRAN.R-project.org/package=Matrix.

[42] Robertson MJ, Inman-Bamber NG, Muchow RC & Wood AW, 1999, *Physiology and productivity of sugarcane with early and mid-season water deficit*, Field Crops Research vol. 64, pp. 211-227.

[43] Robertson MJ, Muchow RC, Donaldson RA, Inman-Bamber NG & Wood AW, 1999, *Estimating the risk associated with drying-off strategies for irrigated sugarcane before harvest*, Australian Journal of Agricultural Research vol. 50, pp.65-77.

[44] Rodrigues FA, de Laia ML & Zingaretti SM, 2009, *Analysis of gene expression profiles under water stress in tolerant and sensitive sugarcane plants*, Plant Science vol. 176, pp. 286302.

[45] Schlüter M, Savitsky AG, McKinney DC & Leith H, 2005, *Optimizing long term water allocation in the Amudarya River delta: a water management model for ecological impact assessment*, Environmental Modelling & Software, vol. 20, pp. 529-545.

[46] Shangguan Z, Shao M, Horton R, Lei T, Qin L & Ma J, 2002, *A model for regional optmal allcoation of irrigation water resources under deficit and its application*, Agricultural Water Management, vol. 52, pp. 139-154.

[47] Shi T-T, Guan D-X, Wu J-B, Wang A-Z, Jin C-J & Han S-J, 2008, *Comparison of methods for estimating evapotranspiration rate of dry forest canopy: Eddy covariance,*

Bowen ratio energy balance, and Penman-Montieth equation, Geophysical research vol. 113.

[48] Shih SF, 1983, *Using Sugar Cane Yields and Evapotranspiration Relations for Water Management*, University of Florida, IFAS, Agricultural Research and Education Center, Belle Glade, Florida 33430, USA.

[49] Silva MdeA, Jifon JL, Silva JAG, Santos CM & Sharma V, 2012, *Relationships between physiological traits and productivity of sugarcane in response to water deficit*, Journal of Agricultural Science vol. 152, pp. 104-118.

[50] Smit MA & Singels A, 2006, *The response of sugarcane canopy development to water stress*, Field Crops Research vol. 98, pp. 91-97.

[51] Snyder RL, 1992, *Equation for evaporation pan to evapotranspiration conversions*, Irrigation and Drainage Engineering, vol. 118, pp. 977-980.

[52] Stat-Ease Incorporation, 2015, *Design-Expert Software Version 9*, 2021 East Hennepin Avenue, Suite 480, Minneapolis, MN 55413, `www.statease.com`.

[53] Stewart TJ, 2012, *Principles and Practice of Linear Programming*, Statistics Honours Course Notes, Department of Statistical Sciences, University of Cape Town.

[54] Sulis A & Sechi GM, 2013, *Comparison of generic simulation models for water resource systems*, Environmental Modelling and Software, vol. 40, pp. 213-225.

[55] Thompson GD & Boyce JP, 1967. *Daily measurements of potential evapotranspiration from fully canopied sugarcane*. Agriculture Meteorology. vol. 4, pp. 267279.

[56] Watanabe K, Yamamoto T,Yamada T, Sakuratani T, Nawata E & Noichana C, 2004, *Changes in seasonal evapotranspiration, soil water content, crop coefficients in sugarcane, cassava, and maize fields in Northeast Thailand*, Agricultural Water Management, vol. 67, pp. 133-143.

[57] Wurbs RA, 1993, *Reserviour-system simulation and optimization models*, Water resource planning and management, vol. 119, pp. 455-472.

[58] Ralph A. Wurbs RA, 2005, *Modeling river/reservoir system management, water allocation, and supply reliability*, Journal of Hydrology, vol. 300, pp. 100113.

[59] Yang x, Zhou Q & Melville M, 1997, *Estimating local sugarcane evapotranspiration using Landsat TM image and a VIT concept*, Remote Sensing, vol. 18, pp. 453-459.

[60] Yaron D & Dinar A, 1982, *Optimal Allocation of farm irrigation water during peak seasons*, American Agricultural Economics Association.

Appendices

Appendix A

Screenshots

A.1 R Screenshot

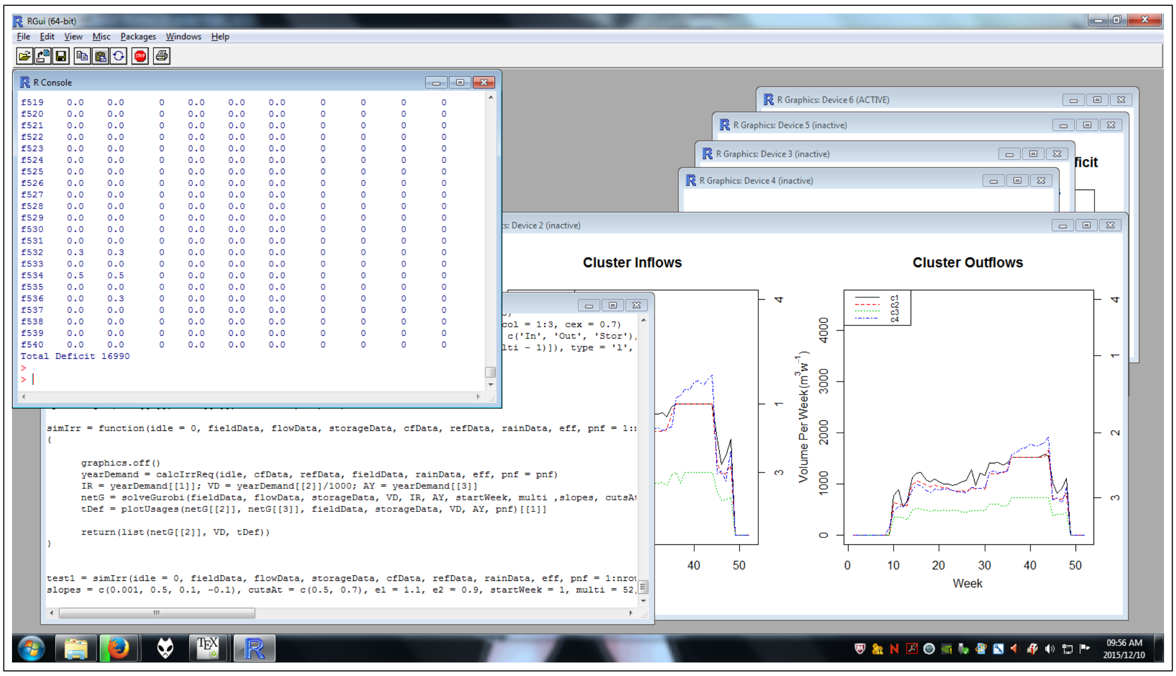
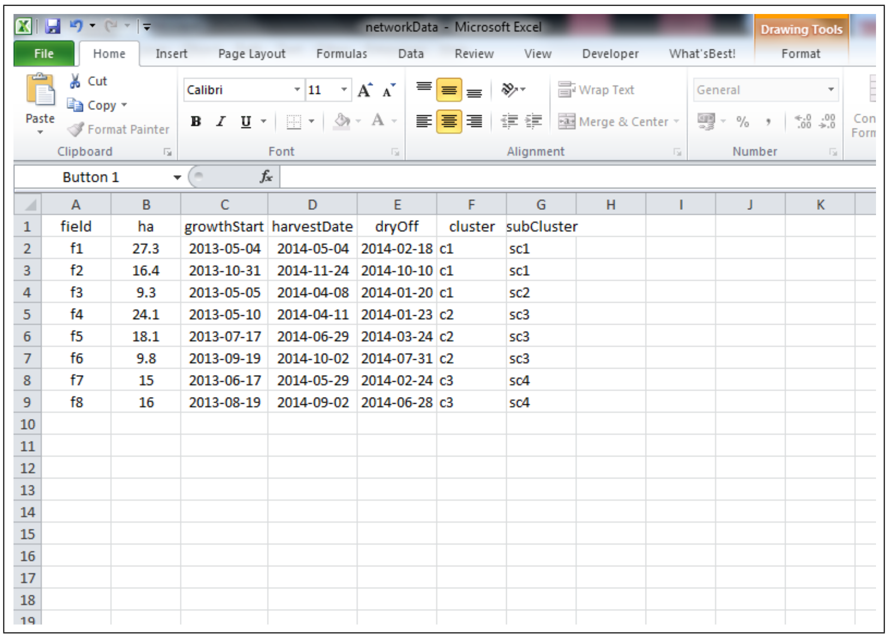


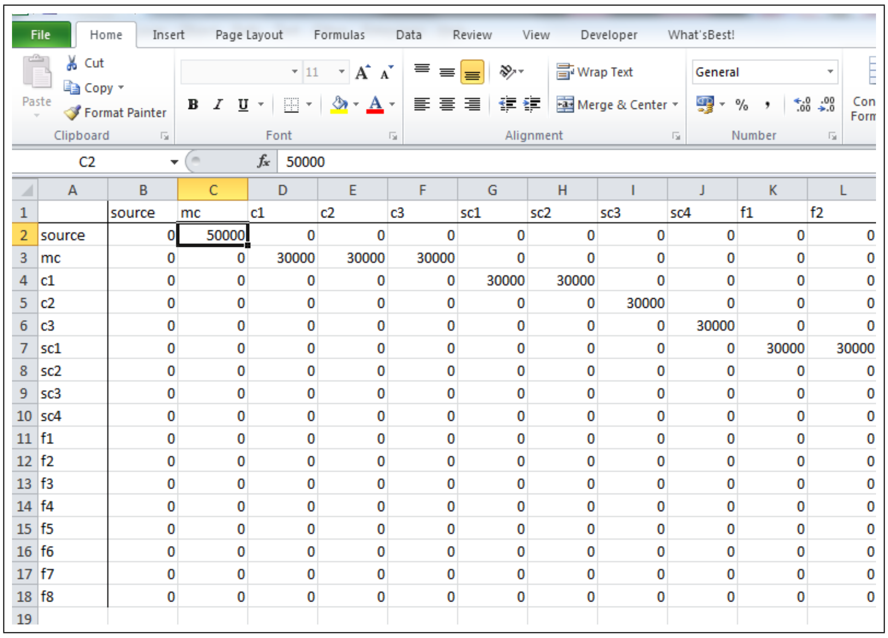
Figure A.1.1: A screenshot of the R environment.

A.2 Data Format



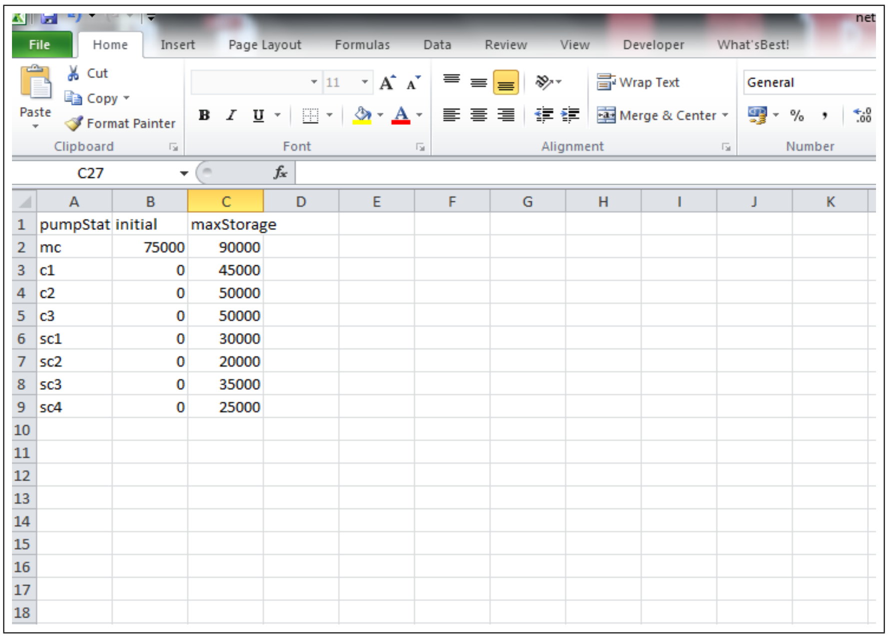
field	ha	growthStart	harvestDate	dryOff	cluster	subCluster
f1	27.3	2013-05-04	2014-05-04	2014-02-18	c1	sc1
f2	16.4	2013-10-31	2014-11-24	2014-10-10	c1	sc1
f3	9.3	2013-05-05	2014-04-08	2014-01-20	c1	sc2
f4	24.1	2013-05-10	2014-04-11	2014-01-23	c2	sc3
f5	18.1	2013-07-17	2014-06-29	2014-03-24	c2	sc3
f6	9.8	2013-09-19	2014-10-02	2014-07-31	c2	sc3
f7	15	2013-06-17	2014-05-29	2014-02-24	c3	sc4
f8	16	2013-08-19	2014-09-02	2014-06-28	c3	sc4

Figure A.2.1: Field data format as a ".csv" file type.



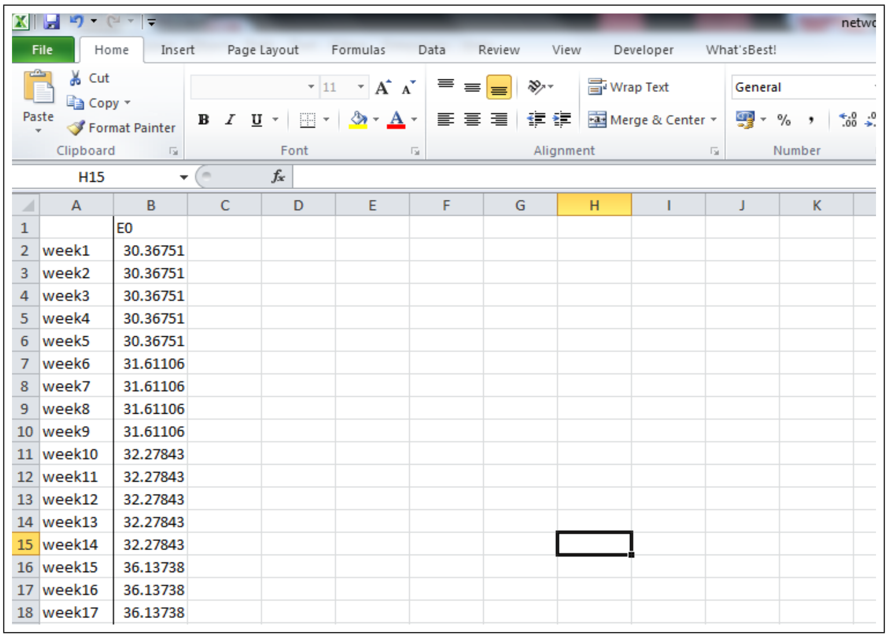
	source	mc	c1	c2	c3	sc1	sc2	sc3	sc4	f1	f2
source	0	50000	0	0	0	0	0	0	0	0	0
mc	0	0	30000	30000	30000	0	0	0	0	0	0
c1	0	0	0	0	0	30000	30000	0	0	0	0
c2	0	0	0	0	0	0	0	30000	0	0	0
c3	0	0	0	0	0	0	0	0	30000	0	0
sc1	0	0	0	0	0	0	0	0	0	30000	30000
sc2	0	0	0	0	0	0	0	0	0	0	0
sc3	0	0	0	0	0	0	0	0	0	0	0
sc4	0	0	0	0	0	0	0	0	0	0	0
f1	0	0	0	0	0	0	0	0	0	0	0
f2	0	0	0	0	0	0	0	0	0	0	0
f3	0	0	0	0	0	0	0	0	0	0	0
f4	0	0	0	0	0	0	0	0	0	0	0
f5	0	0	0	0	0	0	0	0	0	0	0
f6	0	0	0	0	0	0	0	0	0	0	0
f7	0	0	0	0	0	0	0	0	0	0	0
f8	0	0	0	0	0	0	0	0	0	0	0

Figure A.2.2: Flow matrix format as a ".csv" file type.



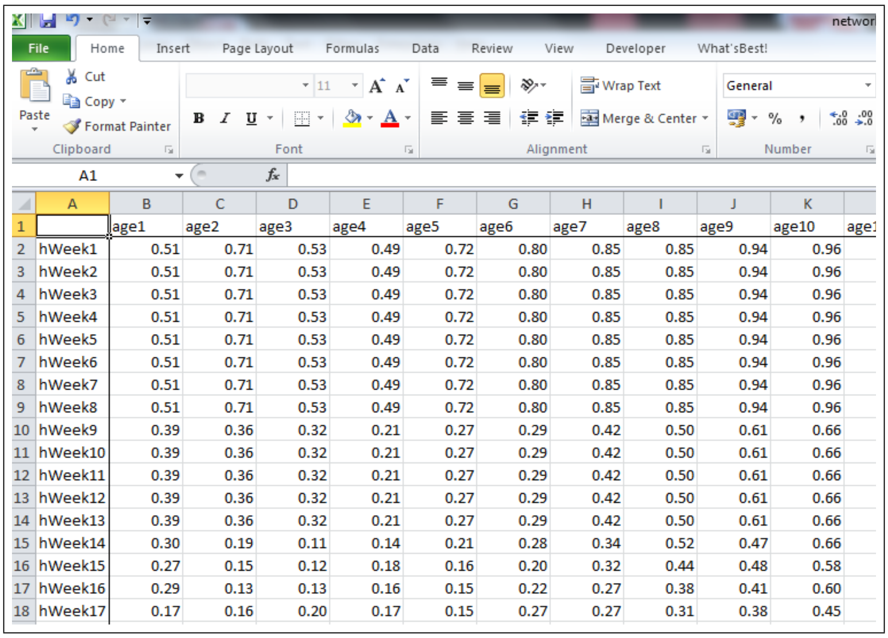
	A	B	C
1	pumpStat	initial	maxStorage
2	mc	75000	90000
3	c1	0	45000
4	c2	0	50000
5	c3	0	50000
6	sc1	0	30000
7	sc2	0	20000
8	sc3	0	35000
9	sc4	0	25000

Figure A.2.3: Pump station storage data format as a ".csv" file type.



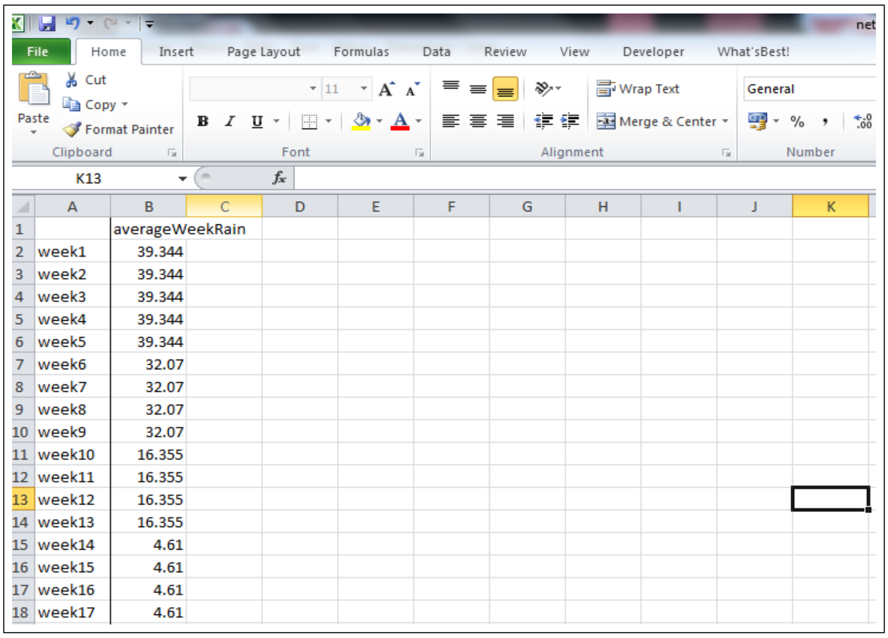
	A	B
1		E0
2	week1	30.36751
3	week2	30.36751
4	week3	30.36751
5	week4	30.36751
6	week5	30.36751
7	week6	31.61106
8	week7	31.61106
9	week8	31.61106
10	week9	31.61106
11	week10	32.27843
12	week11	32.27843
13	week12	32.27843
14	week13	32.27843
15	week14	32.27843
16	week15	36.13738
17	week16	36.13738
18	week17	36.13738

Figure A.2.4: Reference evapotranspiration data format as a ".csv" file type.



	age1	age2	age3	age4	age5	age6	age7	age8	age9	age10
hWeek1	0.51	0.71	0.53	0.49	0.72	0.80	0.85	0.85	0.94	0.96
hWeek2	0.51	0.71	0.53	0.49	0.72	0.80	0.85	0.85	0.94	0.96
hWeek3	0.51	0.71	0.53	0.49	0.72	0.80	0.85	0.85	0.94	0.96
hWeek4	0.51	0.71	0.53	0.49	0.72	0.80	0.85	0.85	0.94	0.96
hWeek5	0.51	0.71	0.53	0.49	0.72	0.80	0.85	0.85	0.94	0.96
hWeek6	0.51	0.71	0.53	0.49	0.72	0.80	0.85	0.85	0.94	0.96
hWeek7	0.51	0.71	0.53	0.49	0.72	0.80	0.85	0.85	0.94	0.96
hWeek8	0.51	0.71	0.53	0.49	0.72	0.80	0.85	0.85	0.94	0.96
hWeek9	0.39	0.36	0.32	0.21	0.27	0.29	0.42	0.50	0.61	0.66
hWeek10	0.39	0.36	0.32	0.21	0.27	0.29	0.42	0.50	0.61	0.66
hWeek11	0.39	0.36	0.32	0.21	0.27	0.29	0.42	0.50	0.61	0.66
hWeek12	0.39	0.36	0.32	0.21	0.27	0.29	0.42	0.50	0.61	0.66
hWeek13	0.39	0.36	0.32	0.21	0.27	0.29	0.42	0.50	0.61	0.66
hWeek14	0.30	0.19	0.11	0.14	0.21	0.28	0.34	0.52	0.47	0.66
hWeek15	0.27	0.15	0.12	0.18	0.16	0.20	0.32	0.44	0.48	0.58
hWeek16	0.29	0.13	0.13	0.16	0.15	0.22	0.27	0.38	0.41	0.60
hWeek17	0.17	0.16	0.20	0.17	0.15	0.27	0.27	0.31	0.38	0.45

Figure A.2.5: Crop factor matrix format as a ".csv" file type.



	averageWeekRain
week1	39.344
week2	39.344
week3	39.344
week4	39.344
week5	39.344
week6	32.07
week7	32.07
week8	32.07
week9	32.07
week10	16.355
week11	16.355
week12	16.355
week13	16.355
week14	4.61
week15	4.61
week16	4.61
week17	4.61

Figure A.2.6: Rain forecast matrix format as a ".csv" file type.

Appendix B

Program Code

```
library(stringr)
library(Matrix)
library(gurobi)

setwd(dir = 'I:/UCT/Dissertation/Code/RRR4')

#monday of week marks new week
getYear = function(yyyy_mm_dd){return(as.numeric(str_sub(yyyy_mm_dd, start = 1, end =
    4)))}
getMonth = function(yyyy_mm_dd){return(as.numeric(str_sub(yyyy_mm_dd, start = 6, end
    = 7)))}
getDay = function(yyyy_mm_dd){return(as.numeric(str_sub(yyyy_mm_dd, start = 9, end =
    10)))}
getWeek = function(yyyy_mm_dd){x = as.POSIXlt(yyyy_mm_dd);return(as.numeric(strftime(
    x, format = '%W')))}

fieldData = read.csv("networkData_sheet1.csv", header = T)
flowData = as.matrix(read.csv("networkData_sheet2.csv", header = T, row.names = 1))
storageData = read.csv("networkData_sheet3.csv", header = T)
refData = as.matrix(read.csv("networkData_sheet4.csv", header = T, row.names = 1))
cfData = as.matrix(read.csv("networkData_sheet5.csv", header = T, row.names = 1))
rainData = read.csv("networkData_sheet6.csv", header = T, row.names = 1)

fieldData[,3] = as.Date(fieldData[,3])
fieldData[,4] = as.Date(fieldData[,4])
fieldData[,5] = as.Date(fieldData[,5])
fieldData[,6] = as.numeric(fieldData[,6])
fieldData[,7] = as.numeric(fieldData[,7])

eff = 0.65
pnf = 1:nrow(fieldData)

calcIrrReq = function(idle = 0, cfData, refData, fieldData, rainData, eff, pnf)
{
        nf = nrow(fieldData)

        #growh and harvest period check
        if(sum(getYear(fieldData[,4]) == getYear(fieldData[1,4])) != nf &sum(getYear(
```

```
    fieldData[,3]) == getYear(fieldData[1,3])) != nf)
{
        cat('Fields not harvested in one year (inconsistent harvest year)!\n'
            )
        cat('Abort\n')
}else{
        harvestYear = getYear(fieldData[1,4])
        growthStartYear = getYear(fieldData[1,3])
}

#harvest and dryoff check
if(sum(getWeek(fieldData[,5]) < getWeek(fieldData[,4])) != nf)
{
        cat('A dryoff date proceeds a harvest date!\n')
        fieldData[getWeek(fieldData[,5]) > getWeek(fieldData[,4]),1]
}

irrReqYear = matrix(NA, nrow = nf, ncol = 52)
volReqYear = matrix(0, nrow = nf, ncol = 52)
ageYear = matrix(0, nrow = nf, ncol = 52)

for(i in 1:nf)
{
        startWeek = getWeek(fieldData[i,3])
        harvestWeek = getWeek(fieldData[i,4])
        dryWeek = getWeek(fieldData[i,5])
        growthLength = 52 - startWeek + harvestWeek
        waterLength = growthLength - (harvestWeek - dryWeek)

        fieldReq = as.vector(cfData[startWeek, 1:growthLength])*c(refData[
            startWeek:52,1], refData[1:dryWeek-1,1], rep(0, harvestWeek -
            dryWeek))
        irrReqYear[i, 1:(harvestWeek-1)] = fieldReq[(52-startWeek+2):(
            growthLength)]

        nextStart = harvestWeek + idle
        at = 52 - startWeek + 1
        if(nextStart>52)
        {
                irrReqYear[i, harvestWeek:52] = 0
                age = c((at):(at + harvestWeek - 1), rep(0, times = 52 -
                    length(at:(at+harvestWeek - 1))))
        }else{
                irrReqYear[i, nextStart:52] = as.vector(cfData[nextStart,
                    1:(52-harvestWeek-idle + 1)])*refData[nextStart:52,1]
                age = c(at:(at+harvestWeek - 1), rep(0, times = idle), 1:52)
                age = age[1:52]
        }

        ageYear[i,] = age
}

irrReqYear = irrReqYear - matrix(1,nrow = nf)%*%rainData[,1]
irrReqYear[irrReqYear<0] = 0
```

```r
        volReqYear = 10*irrReqYear*(fieldData[,2]%*%matrix(rep(1,52), nrow = 1))
        volReqYear = (1/eff)*volReqYear

        par(mar=c(5, 4, 4, 4) + 0.1)
        matplot(x = 1:52, y = t(volReqYear[pnf,]), type = 'l', ylab = '', xlab = 'Week', main = paste('Irrigation Volume Requirement for', harvestYear), col = pnf, lty = pnf)
        mtext(2,text=expression(paste("Volume Per Week ", '(' , m^3, w^-1, ')')),line=2)
        legend('topleft', legend = paste('f', pnf, sep = ''), lty = pnf, col = pnf, cex = 0.8)
        par(new=T)
        plot(x = 1:52,y = rainData[,1], axes=F, type='p', ylab = '', xlab = '', pch = 21)
        axis(4, ylim=c(0,max(rainData[,1])),lwd=1,line=0)
        mtext(4,text=expression(paste("Rain Height Per Week ", '(' , mm, w^-1, ')')), line=2)
        par(mar=c(5, 4, 4, 2) + 0.1)
        return(list(irrReqYear,volReqYear, ageYear))
}

#coeffs of benefit
cF = function(demandFields, ageFields, slopes = c(0.01, 0.5, 0.1, -0.1), cutsAt = c(0.2, 0.70), e1, e2)
{

        piecesMat = matrix(c(cutsAt, 1, 1.25), ncol = 1)%*%matrix(demandFields, nrow = 1)
        cutsMat = piecesMat - rbind(0, piecesMat[-4,] )

        #increse benefit for mid age, decrease benefit for mature
        ageEffect = rep(0, length(ageFields))
        ageEffect[ageFields <= 11] = 1
        ageEffect[ageFields <= 43 & ageFields >11] = e1
        ageEffect[ageFields >43] = e2

        slopesMat = matrix(slopes, ncol = 1)%*%matrix(ageEffect, nrow = 1)
        return(list(cutsMat, slopesMat))
}


#function to preparing MIL and solve
solveGurobi = function(fieldData, flowData, storageData, VD, ID, AY, startWeek, multi, slopes = c(0.01, 0.5, 0.1, -0.1), cutsAt = c(0.4, 0.70), e1, e2, SOLVE = FALSE)
{

        nfields = nrow(fieldData)
        nclust = max(fieldData[,6])
        nsub = max(fieldData[,7])
        clusters = 1:nclust
        subClusters = 1:nsub
        nnodes = 1 + nclust + nsub + nfields
        npumps = nrow(storageData)
```

```r
relMat = as.matrix(cbind(1,fieldData[,c(6,7)], (1 + nclust + nsub + 1) :
    nnodes))
relMat[,2] = relMat[,2] + 1
relMat[,3] = relMat[,3] + nclust + 1

flows = unique(rbind(relMat[,1:2], relMat[,2:3], relMat[,3:4])) #to be used
    for water balance
flows = rbind(c(0,1),flows)
nflows = nrow(flows)
colnames(flows) = c('fromNode', 'toNode')

#rows is constraints, number of rows = number of constraints. Colums variable
     coefficients and number of columns = n0. of varaibles
ncon = npumps + nflows + npumps + nfields + 7*nfields
nvar = npumps + nflows + 4*nfields + 3*nfields
spaac = cbind(c(),c(),c())
naam = c(paste('s', 1:npumps, sep = ''), paste('v', 1:nflows, sep = ''),
paste('v', (nflows-nfields+1):nflows, '.1', sep = ''), paste('v', (nflows-
    nfields+1):nflows, '.2', sep = ''),
paste('v', (nflows-nfields+1):nflows, '.3', sep = ''), paste('v', (nflows-
    nfields+1):nflows, '.4', sep = ''),
paste(rep(c('b1','b2','b3'), times = nfields), '.', sort(rep(1:6, 3)), sep =
    ''))


#storage Constraints
spaac = rbind(spaac, cbind(1:npumps, 1:npumps, 1))
b = storageData[,3]
ss = rep('<=', times = npumps)
here = npumps

#flow contraints
spaac = rbind(spaac, cbind((npumps + 1):(npumps + nflows),(npumps + 1):(
    npumps + nflows), 1))
b = c(b, matrix(1, ncol = nflows + 1)%*%flowData[,-1])
ss = c(ss, rep('<=', times = nflows))
here  = npumps + nflows

#splitting variable
spaac = rbind(spaac, cbind((here + 1):(here + nfields), (npumps + (nflows -
    nfields) + 1):(npumps + nflows),1))
spaac = rbind(spaac, cbind((here + 1):(here + nfields), nfields + ((npumps +
    (nflows - nfields) + 1):(npumps + nflows)),-1))
spaac = rbind(spaac, cbind((here + 1):(here + nfields), 2*nfields + ((npumps
    + (nflows - nfields) + 1):(npumps + nflows)),-1))
spaac = rbind(spaac, cbind((here + 1):(here + nfields), 3*nfields + ((npumps
    + (nflows - nfields) + 1):(npumps + nflows)),-1))
spaac = rbind(spaac, cbind((here + 1):(here + nfields), 4*nfields + ((npumps
    + (nflows - nfields) + 1):(npumps + nflows)),-1))

b = c(b, rep(0, times = nfields))
ss = c(ss, rep('=', times = nfields))
here = npumps + nflows + nfields
```

```
#continuity equation now
spaac = rbind(spaac, cbind((here+1):(here+npumps), 1:npumps, -1))
spaac = rbind(spaac, cbind((here+1):(here+npumps), npumps + (1:npumps),1))


#outflows
for(i in 1:npumps)
{
        fromCol = npumps + i
        toCol = npumps + flows[flows[,1]==i,2]
        fromCol
        toCol
        spaac = rbind(spaac, cbind(here + i, c(toCol),-1))
}
b = c(b, -storageData[,2])
ss = c(ss, rep('=', times = npumps))
here = npumps + nflows + nfields + npumps
fixedPart = nrow(spaac)

# the split variables constraints (the pieces sizes)
#fields on columns, time and pieces on rows
boundMat = c(); coeffMat = c()

#calculating bounds and slopes to use immediately and later
for(i in 1:multi)
{
        hh = cF(VD[,startWeek + i - 1], AY[,startWeek + i - 1], slopes,
            cutsAt, e1, e2)
        boundMat = rbind(boundMat, hh[[1]])
        coeffMat = rbind(coeffMat, hh[[2]])
}

pv = npumps + nflows
pb = pv + 4*nfields
#bounds and constraints on new variables now

#less than bounds
#the first less than inequalities that does not have a binary multiplied but
    a 'PARAMETER' in b vector
spaac = rbind(spaac, cbind((here+1):(here+nfields),(pv+1):(pv+nfields),1))
b = c(b, t(boundMat)[1:nfields])
ss = c(ss, rep('<=', times = nfields))
here = here + nfields

#the other three sets now
spaac = rbind(spaac, cbind((here+1):(here+3*nfields),(pv + nfields + 1):(pv+
    4*nfields),1))
spaac = rbind(spaac, cbind((here+1):(here+3*nfields),c(seq(pb+1 ,by = 3,
    length = nfields), seq(pb+2 ,by = 3, length = nfields), seq(pb+3 ,by = 3,
     length = nfields)), -1*t(boundMat)[(nfields+1):(4*nfields)]))
here = here + 3*nfields

#greater than bounds
```

```
spaac = rbind(spaac, cbind((here+1):(here+3*nfields),(pv + 1):(pv+ 3*nfields)
    ,1))
spaac = rbind(spaac, cbind((here+1):(here+3*nfields),c(seq(pb+1 ,by = 3,
    length = nfields), seq(pb+2 ,by = 3, length = nfields), seq(pb+3 ,by = 3,
     length = nfields)), -1*t(boundMat)[(1):(3*nfields)]))
here = here + 3*nfields

b = c(b,rep(0, times = 2*3*nfields))
ss = c(ss, rep('<=', times = 3*nfields), rep('>=', times = 3*nfields))
binInd = (pb+1):nvar

#objective function
obj = c(rep(0, times = pv),t(coeffMat)[1:(4*nfields)], rep(0, times = 3*
    nfields))
length(obj)
names(obj) = naam

#------------------------------------------------------------now multiple weeks
    -------------------------------------------

#one week formulatedm, now use first week to create entire year
vasseb = b
vassess = ss
vasseobj = obj
vassebinInd = binInd
vassespaac = spaac
vassenaam = naam

length(b);length(ss)
for(i in 1:(multi-1))
{
        #adding new variables
        naam2 = paste(vassenaam, '.t',i+1, sep = '')
        naam = c(naam,naam2)

        #storage, flow, spitting constraints and part water balance
        spaac = rbind(spaac, vassespaac[1:fixedPart,] + cbind(i*ncon,i*nvar,
            rep(0,fixedPart)))
        #previous water stoprage relating
        spaac = rbind(spaac, cbind((i*ncon + npumps + nflows + nfields + 1):(
            i*ncon + npumps + nflows + nfields + npumps), ((i-1)*nvar + 1):((
            i-1)*nvar + npumps),1))
        here = here + npumps + nflows + nfields + npumps

        #split variable bounds
        #less than bounds
        #the first less tha inequalities that does not haave a binary
            multiplied BUT A 'PARAMETER' in b vector
        spaac = rbind(spaac, cbind((here+1):(here+nfields),((i*nvar) + pv+1)
            :((i*nvar) + pv+nfields),1))
        here = here + nfields

        #the other three sets now
        spaac = rbind(spaac, cbind((here+1):(here+3*nfields),((i*nvar) + pv +
```

```
        nfields + 1):((i*nvar) + pv+ 4*nfields),1))
        spaac = rbind(spaac, cbind((here+1):(here+3*nfields),c(seq((i*nvar) +
            pb + 1 ,by = 3, length = nfields), seq((i*nvar) + pb+2 ,by = 3,
            length = nfields), seq((i*nvar) + pb+3 ,by = 3, length = nfields)
            ), -1*t(boundMat)[((4*i*nfields) + nfields+1):((4*i*nfields) + 4*
            nfields)]))
        here = here + 3*nfields

        #greater than bounds
        spaac = rbind(spaac, cbind((here+1):(here+3*nfields),((i*nvar) + pv +
            1):((i*nvar) + pv+ 3*nfields), 1))
        spaac = rbind(spaac, cbind((here+1):(here+3*nfields),c(seq((i*nvar) +
            pb+1 ,by = 3, length = nfields), seq((i*nvar) + pb+2 ,by = 3,
            length = nfields), seq((i*nvar) + pb+3 ,by = 3, length = nfields)
            ), -1*t(boundMat)[((4*i*nfields) +1):((4*i*nfields) + 3*nfields)
            ]))
        here = here + 3*nfields

        #creating the b vector and constraint vector sign
        b = c(b, vasseb[1:(npumps+nflows+nfields)], rep(0, times = npumps), t
            (boundMat)[((4*i*nfields) +1):((4*i*nfields) ++nfields)], rep(0,
            times = 6*nfields))
        ss = c(ss, vassess)
        obj = c(obj, rep(0, times = pv),t(coeffMat)[((4*i*nfields) +1):((4*i*
            nfields) +4*nfields)], rep(0, times = 3*nfields))
        length(obj)
        names(obj) = naam

        binInd = c(binInd, (i*nvar) + vassebinInd)
}

if(SOLVE == TRUE)
{
        #gurobi optimiser part
        model = list()
        sA = spMatrix(nrow = multi*ncon, ncol = multi*nvar, i = spaac[,1], j
            = spaac[,2], x = spaac[,3])

        model$A = sA
        model$obj = obj
        model$modelsense = "max"
        model$rhs = b
        model$sense = ss
        vt = rep('C', times = length(obj))
        vt[binInd] = 'B'
        model$vtype = vt

        result <- gurobi(model)
        decVec = result$x
        decMat = matrix(decVec, ncol = multi)
        rownames(decMat) = vassenaam
        colnames(decMat) = paste('week', startWeek:(startWeek + multi - 1),
            sep = '')
        appliedVol = decMat[(npumps+nflows - nfields + 1):(npumps+nflows), ]
```

```
            reqVol = VD[ ,startWeek:(startWeek + multi - 1)]
            defProp = (reqVol-appliedVol)/reqVol
            defProp[is.nan(defProp)] = 0
            print(round(defProp,1))
        }
        return(list(result$objval, decMat, c(npumps,nflows,nfields, startWeek), sA, b
            , ss, obj, spaac, constraints = multi*ncon, variables = multi*nvar))
}


#netG = solveGurobi(fieldData, flowData, storageData, VD, IR, AY, startWeek = 1,
    multi = 52, slopes = c(0.001, 0.5, 0.1, -0.1), cutsAt = c(0.5, 0.70), e1 = 1.1,
    e2= 0.9, SOLVE = T)
#decMat = netG[[2]]
#cc = netG[[3]]

plotUsages = function(decMat, cc, fieldData , storageData, VD, AY, pnf)
{
        multi = ncol(decMat)
        nc = max(fieldData[,6]) #number of Clusters
        cMaxIn = flowData[2,3:(2+ nc)]
        cMaxOut = rowSums(flowData[3:(2+nc),])
        cStorage = cbind(initial = storageData[2:(1+nc),2],decMat[2:(1+nc),])
        cInflows = decMat[(2+cc[1]):(1 + cc[1]+ nc),]
        cOutflows = cInflows + cStorage[,-(multi+1)] - cStorage[,-1]

        deficit = round(VD[,cc[4]:(cc[4]+multi-1)] - decMat[(cc[1]+cc[2]-cc[3]+1):(cc
            [1]+cc[2]),], 0)
        cDeficit = matrix(0, nrow = nc, ncol = multi)
        for(i in 1:nc)
        {
                cDeficit[i,] = colSums(deficit[fieldData[,6]==i,])
        }

        par(mfrow = c(1,2))
        matplot(x = (cc[4]):(cc[4] + multi - 1), y = t(cInflows), xlab = 'Week', type
             = 'l', col = 1:nc, main = 'Cluster Inflows', ylab = expression(paste("
            Volume Per Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3, 1, 0), ylim = c
            (0, max(cMaxIn)))
        legend('topleft', legend = paste('c', 1:nc, sep = ''), lty = 1:nc, col = 1:nc
            , cex = 0.7)
        axis(side=4, at=cMaxIn, labels = 1:nc, line=0)
        matplot(x = (cc[4]):(cc[4] + multi - 1), y = t(cOutflows), xlab = 'Week',
            type = 'l', col = 1:nc, main = 'Cluster Outflows', ylab = expression(
            paste("Volume Per Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3, 1, 0),
            ylim = c(0, max(cMaxOut)))
        legend('topleft', legend = paste('c', 1:nc, sep = ''), lty = 1:nc, col = 1:nc
            , cex = 0.7)
        axis(side=4, at=cMaxOut, labels = 1:nc, line=0)

        nsc = max(fieldData[,7]) #number of SUBClusters
        scMaxIn = matrix(colSums(flowData[3:(2+nc),(3+nc):(2+nc+nsc)]), ncol = nsc)
        scMaxOut = rowSums(flowData[(3+nc):(2+nc+nsc),])
        scStorage = cbind(initial = storageData[(2+nc):(1+ nc + nsc),2], decMat[(2+nc
```

```
    ):(1+ nc + nsc),])
scInflows = decMat[(2+cc[1] + nc ):(1+cc[1] + nc + nsc),]
scOutflows = scInflows + scStorage[,-(multi+1)] - scStorage[,-1]

scDeficit = matrix(0, nrow = nsc, ncol = multi)
for(i in 1:nsc)
{
        scDeficit[i,] = colSums(matrix(deficit[fieldData[,7] == i,],ncol=
            multi))
}

dev.new()
par(mfrow = c(1,2))
matplot(x = (cc[4]):(cc[4] + multi - 1), y = t(scInflows), xlab = 'Week',
    type = 'l', col = 1:nsc, main = 'Sub-Cluster Inflows', ylab = expression(
    paste("Volume Per Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3, 1, 0),
    ylim = c(0, max(scMaxIn)))
legend('topleft', legend = paste('sc', 1:nsc, sep = ''), lty = 1:nsc, col =
    1:nsc, cex = 0.7)
axis(side=4, at=scMaxIn, labels = 1:nsc, line=0)
matplot(x = (cc[4]):(cc[4] + multi - 1), y = t(scOutflows), xlab = 'Week',
    type = 'l', col = 1:nsc, main = 'Sub-Cluster Outflows', ylab = expression
    (paste("Volume Per Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3, 1, 0),
    ylim = c(0, max(scMaxOut)))
legend('topleft', legend = paste('sc', 1:nsc, sep = ''), lty = 1:nsc, col =
    1:nsc, cex = 0.7)
axis(side=4, at=scMaxOut, labels = 1:nsc, line=0)

dev.new()
par(mfrow = c(1,2))
matplot(x = (cc[4]):(cc[4] + multi - 1), y = t(cStorage[,-1]), xlab = 'Week',
     type = 'l', col = 1:nc, main = 'Cluster Storage', ylab = expression(
    paste("Volume Per Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3, 1, 0))
legend('topleft', legend = paste('c', 1:nc, sep = ''), lty = 1:nc, col = 1:nc
    , cex = 0.7)
axis(side=4, at=storageData[2:(1+nc),3], labels = 1:nc, line=0)
matplot(x = (cc[4]):(cc[4] + multi - 1), y = t(scStorage[,-1]), xlab = 'Week'
    , type = 'l', col = 1:nsc, main = 'Sub-Cluster Storage', ylab =
    expression(paste("Volume Per Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3,
     1, 0))
legend('topleft', legend = paste('sc', 1:nsc, sep = ''), lty = 1:nsc, col =
    1:nsc, cex = 0.7)
axis(side=4, at=storageData[(2+nc):(1+nc+nsc),3], labels = 1:nsc, line=0)
cat('Total Deficit' , sum(cDeficit), '\n')

dev.new()
par(mfrow = c(1,2))
matplot(x = (cc[4]):(cc[4] + multi - 1), y = t(cDeficit), xlab = 'Week', type
     = 'l', col = 1:nc, main = 'Cluster Deficit', ylab = expression(paste("
    Volume Per Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3, 1, 0))
legend('topleft', legend = paste('c', 1:nc, sep = ''), lty = 1:nc, col = 1:nc
    , cex = 0.7)
matplot(x = (cc[4]):(cc[4] + multi - 1), y = t(scDeficit), xlab = 'Week',
    type = 'l', col = 1:nsc, main = 'Sub-Cluster Deficit', ylab = expression(
```

```
        paste("Volume Per Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3, 1, 0))
    legend('topleft', legend = paste('sc', 1:nsc, sep = ''), lty = 1:nsc, col =
        1:nsc, cex = 0.7)


    #main pump station
    dev.new()
    par(mfrow = c(1,2))
    mpStorage = c(initial = storageData[1,2], decMat[1,])
    mpOutflow = decMat[cc[1] + 1,] + mpStorage[-(multi+1)] - mpStorage[-1]
    plot(x = (cc[4]):(cc[4] + multi - 1), y = decMat[cc[1] + 1,], main = 'Main
        Pump Station Usage', xlab = 'Week', ylab = expression(paste("Volume Per
        Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3, 1, 0), type = 'l', lty = 1,
        ylim = c(0, max(storageData[1,3], c(flowData[1,2], sum(cMaxIn)))) )
    lines(x = (cc[4]):(cc[4] + multi - 1), y = mpOutflow, col = 2, lty = 2)
    lines(x = (cc[4]):(cc[4] + multi - 1), y = mpStorage[-1], col = 3, lty = 3)
    legend('topleft', legend = c('Inflow', 'Outflow', 'Storage'), lty = 1:3, col
        = 1:3, cex = 0.7)
    axis(side=4, at=c(flowData[1,2], sum(cMaxIn), storageData[1,3]), labels = c('
        In', 'Out', 'Stor'), line=0)
    plot(x = (cc[4]):(cc[4] + multi - 1), y = colSums(VD[,(cc[4]):(cc[4] + multi
        - 1)]), type = 'l', xlab = 'Week', ylab = expression(paste("Volume Per
        Week ", '(' , m^3, w^-1, ')')), mgp = c(2.3, 1, 0), main = 'Total Weekly
        Demand')

    return(list(sum(cDeficit)))
}

#a function to call all the required functions
simIrr = function(idle = 0, fieldData, flowData, storageData, cfData, refData,
    rainData, eff, pnf = 1:nrow(fieldData), slopes, cutsAt, e1, e2, startWeek = 1,
    multi = 52, SOLVE = TRUE)
{

    graphics.off()
    yearDemand = calcIrrReq(idle, cfData, refData, fieldData, rainData, eff, pnf
        = pnf)
    IR = yearDemand[[1]]; VD = yearDemand[[2]]; AY = yearDemand[[3]]
    netG = solveGurobi(fieldData, flowData, storageData, VD, IR, AY, startWeek,
        multi ,slopes, cutsAt, e1, e2, SOLVE)
    tDef = plotUsages(netG[[2]], netG[[3]], fieldData, storageData, VD, AY, pnf)
        [[1]]

    return(list(netG[[2]], VD, tDef))
}


#running the function
test1 = simIrr(idle = 0, fieldData, flowData, storageData, cfData, refData, rainData,
     eff, pnf = 1:nrow(fieldData),
slopes = c(0.001, 0.5, 0.1, -0.1), cutsAt = c(0.5, 0.7), e1 = 1.1, e2 = 0.9,
    startWeek = 1, multi = 52, SOLVE = TRUE)
```

```
#timing a call of the function
system.time(expr = simIrr(idle = 0, fieldData, flowData, storageData, cfData, refData
    , rainData, eff, pnf = 1:nrow(fieldData),
slopes = c(0.001, 0.5, 0.1, -0.1), cutsAt = c(0.5, 0.7), e1 = 1.1, e2 = 0.9,
    startWeek = 1, multi = 52, SOLVE = TRUE))
rainData[,1] = 0

#supplied irrigation plot
dev.new()
app = test1[[1]][17:24,]
matplot(t(app), type = 'l', lty = 1:8, xlab = 'Week', col = 1:8, ylab = '', main = '
    Irrigation Volume Applied for 2014', ylim = c(0,max(VD)))
mtext(2,text=expression(paste("Volume Per Week ", '(' , m^3, w^-1, ')')),line=2)
legend('topleft', legend = paste('f', 1:8, sep = ''), lty = 1:8, col = 1:8)
```