Old Dominion University

ODU Digital Commons

Winter 2015

A Framework and System for a Multi-Model Decision Aid for Sustainable Farming Practices

Kasi Bharath Vegesana
Old Dominion University

Follow this and additional works at: https://digitalcommons.odu.edu/msve_etds

Part of the Agricultural Education Commons, Environmental Education Commons, and the Systems Engineering Commons

Recommended Citation

A FRAMEWORK AND SYSTEM FOR A MULTI-MODEL DECISION AID FOR SUSTAINABLE FARMING PRACTICES

by

Kasi Bharath Vegesana
B.E. March 2008, Andhra University
M.E. May 2010, Old Dominion University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

MODELING AND SIMULATION

OLD DOMINION UNIVERSITY
December 2015

Approved by:

Frederic D. McKenzie (Advisor)

Yuzhong Shen (Member)

Andrew J. Collins (Member)
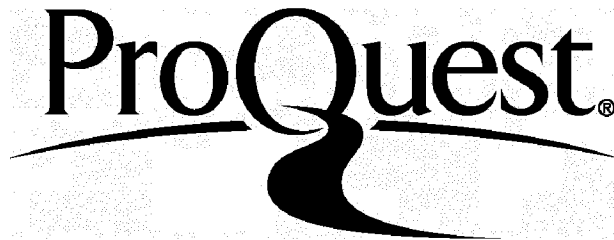
Michael J. Seifer (Member)

ProQuest Number: 10128806

ProQuest 10128806

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

A FRAMEWORK AND SYSTEM FOR A MULTI-MODEL DECISION AID FOR SUSTAINABLE FARMING PRACTICES

Kasi Bharath Vegesana
Old Dominion University, 2015
Director: Frederic McKenzie

Decision support systems (DSS) for farmers address the need for modeling multiple processes and scenarios that affect farmer decision making. Existing DSS have various drawbacks that stop them from being deployed as decision support tools. This research proposes a multi-model simulation framework that can be used to analyze farm management practices at the crop level, individual farm level and at the community level to show the impact and alternatives for smallholder farming practices. A generic crop growth model is proposed, based on existing equations. We run sensitivity analysis on the model to identify important variables. The outputs from the crop model are utilized in a series of linear programming models to estimate the optimal scheduling of crops. In addition to these models we build a rule-based fuzzy system to replicate existing trends among farmers. Predicting these trends help us in identifying the decision patterns of farmers and help us in conducting scenario analysis to gauge the farmers reactions to external stimuli. The various limitations and assumptions of the models are described, and we conclude with suggestions for improving these models.

This dissertation is dedicated to my family. You believed in me long before I learnt to believe in myself.

ACKNOWLEDGMENTS

I would like to thank my advisors Dr. Frederic McKenzie, Dr. Yuzhong Shen, Dr. Andrew Collins and Dr. Michael Seiler for their guidance and support during these five years. They have always encouraged me to pursue my ideas irrespective of the results. This has helped me grow from a student to a researcher. I am particularly grateful for their patience as they watched me take awkward steps on my PhD journey. Their timely comments and criticisms have provided much needed structure and direction to this dissertation. I would also like to thank the MSVE department for supporting me during my study and making my time in the department a smooth journey.

Without the constant support and encouragement of my parents, I would not have had the courage to pursue my own path in life. For this, I will be eternally grateful. This dissertation is also a testament to the patience and encouragement of my wife. Her support during these testing times gave me the strength to reach the finish line. No journey in life is complete without the company of friends, and I would like to believe that I have the greatest friends in the world. Their friendship has made ODU feel like a home away from home.

I do not yet know what life has in store for me, but I am happy to have followed the path of a PhD student. Whatever I may choose to do in life, I firmly believe that the experience of the last five years will always have a profound effect on me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The definition of sustainability in agriculture has been subject to multiple revisions. Various [1] interpretations and objectives have been prescribed to define sustainability. The defining objective of sustainability ranges from food sufficiency and proper allocation and management of resources to economic profitability and land productivity [1]. Since, the conditions dictating agricultural activities vary between different regions, varying evaluation criteria are required to judge sustainability. As such, it is hard to find one universal definition of sustainability that can be applied to all agricultural regions of the world [2][1].

These definitions of sustainable agriculture have progressed from a purely economic objective to now include ecological and social considerations [3]. Sustainable agriculture is now considered as, methods or practices that facilitate the development of social, economic and environmental objectives by finding a common ground between the various conflicting options that these objectives present. For example, Rigby et al [2], have developed indicators that measure sustainability as increase in a farmers' quality of life and self-reliance while minimizing off-farm inputs and the promotion of environmental quality.

An advance in seed production technology, fertilizers and availability of electricity has enabled farmers to increase their yield per unit area during the past few decades. The downside to all this growth has been the marked increase in soil erosion, soil leeching,

---

[1] IEEE Transactions and Journals style is used in this thesis for formatting figures, tables, and references.

water contamination due to fertilizers and pesticides [4]. This has created a situation where, the decisions made for temporary economic gains are affecting the long-term sustainability of farms [5]. It is necessary for policy makers and farmers to understand the decisions being made at farm level, and their consequences on the immediate environment, in order to create long-term plans for sustainable agriculture. Existing literature ([6], [7]) has shown the demand for systems that can incorporate scientific knowledge into tools that can act as decision aides for farmers. Various authors like Allen et al., [8] and Lowrance et al., [1] have proposed hierarchical approaches that can model farm management practices at field level, farm level, regional and national level using different models and constraints. This approach serves to show the farmers the impact of their decisions and the alternatives available to them to achieve different and sometimes conflicting objectives.

1.1 Motivation

This need for modeling various scenarios and decisions at various levels, and analyzing their impact, has resulted in the demand for expert systems that can aid farmers and decision makers in making decisions that meet the objective of sustainability [9]. Such a system would need to combine the various aspects of farm level procedures from crop growth dynamics to community based decision models [10]. It would require quantifying various decision alternatives and scenarios through data analysis and a review of previous work. The designers of such a system would also need to identify those areas of farm processes that have a significant impact on the farm level decision process, and eliminate excessive complexity in the system [11]. In order to give decision makers access to all these various aspects of decision making, we first need to understand the various

socioeconomic and environmental issues faced by farmers, and derive the criteria to measure sustainability.

Typically farmers are profit maximizers [12]. Their primary objective is to maximize their profits for each cropping season. Social and environmental welfare are generally treated as secondary objectives that are contingent upon the completion of the primary objective [13]. Not surprisingly, the actions taken during the pursuit of the primary objective can cause a significant impact on the secondary objectives. The farmers often face various issues in achieving their primary objective. These issues include, but are not limited to:

- Planting of high reward, high risk crops to balance farm expenditure.

- Prioritizing cash crops over subsistence crops for family and community use.

- Lack of educational and technical assistance, especially for smallholder farms.

- Lack of an economic buffer discourages adoption of new technologies.

- Lack of responsibility towards natural resources.

- Pollution of immediate environment due to irresponsible use of pesticides, herbicides, and artificial fertilizers.

- Decrease in biodiversity due to clearance of forest land for farming purposes.

- Soil erosion due to improper farming methods.

- Adverse effects on water sources due to improper irrigation, and chemical usage.

Using the various challenges faced by farmers as a base, organizations like the

World Bank [14] use a list of indicators to help monitor socioeconomic and environmental

sustainability of Agricultural and Rural development programs. Table 1 shows a subset of

the indicators, the assessment methods and tools.

Table 1 Indicators, assessment methods, and tools to evaluate sustainability

Impact indicator	Assessment	Tools
Stability of productivity and returns	Farm level cost benefit analysis	Sample surveys
Maintenance of soil fertility and vegetative cover	Soil analysis	Case studies and expert interviews
Loss or maintenance of habitats	No. of new settlements in forest area	Participatory environmental modeling
Extent of eutrophication of water courses	Agrochemicals used per area	Agrochemical statistics
Changes to the way of life of the indigenous people	Practice of traditional livelihood sources	Case studies
Socioeconomic situation of resettled groups or communities in terms of livelihoods, health, and education	In comparison to previous Situation	Sample surveys
Dissemination of indigenous or traditional practices that promote sustainable resource use	Uptake of traditional practices	Case studies

Our research effort stems from this need to educate farmers about the various

aspects of sustainable farming. There is a clear need for an expert system that can present

the farmers with the impacts of their decisions on the sustainability of their farming

practices. The proposed system should be a gateway for the farmers to evaluate their

decisions from multiple points of view. Such a system should aim to aid the farmers'

decision process to achieve sustainability. It should be noted that the proposed system aims

to be an aid to decision making, and not a predictive tool to suggest the best possible decisions.

1.2 Problem

The main problem associated with creating expert systems, which present multiple points of view of the problem, is the multi-disciplinary nature of those views. These systems need to incorporate knowledge of bio physical processes, farm management practices, resource utilization strategies and knowledge of farmer decision patterns. A detailed review of existing expert systems was conducted by [15]. The authors highlight the complexity involved in incorporating these multiple disciplines in one expert system. A tool that is aimed at helping farmers make decisions needs to be able to present the problem to the user from multiple perspectives and provide solutions for each of those perspectives. It should also be able to help the decision makers compare the results of the different perspectives, and provide a measure for computing the best decision or sequence of decisions.

Decision support systems for farmers, fall under this category of expert systems. Due to the multiple methods of formulation of a farming problem, no single modeling methodology can answer all the questions a decision maker might ask. The various decision modeling methods can only address specific sets of scenarios. For example, the authors in [16]–[18] treat farming problems as resource and policy optimization problems. They do not address the motivations behind decision making processes explicitly.

On the other hand, the researchers in [19] address farming problems as purely theoretical decision making problems. This causes the models to stress the importance of

marginal variables, like technical know-how, and assign them a weightage that is equal to prominent variables like market price of a product.

Additional problems arise when most models do not integrate crop growth models into their decision support systems. This problem stems from research groups that concentrate on specific problems of specific areas. Though this gives the research groups the flexibility to use historic yield data while formulating their problem, it becomes hard to apply their conclusions to other regions and crops. There exist a handful of systems like Farm Systems Simulator(FSSIM) [10] from the European Union (EU) that have successfully used crop growth models as a part of their systems. The problem with a system like this is the sheer complexity of the input data required, and the region specific constraints required. This effectively makes the tool very complicated to use, save for people who are well versed in intricate governmental policies and have extensive knowledge of crop specific phenology.

Thus the main problem can be summarized as, the unavailability of a decision support system that:

1. Integrates multiple modeling methodologies with a crop growth mode into a single system.

2. Does not require unrealistic amount of inputs from users who have a limited knowledge of the various methodologies.

3. Can address the same problem from economic, environmental and social perspectives and produce meaningful results.

These problems associated with existing decision support system for farmers are by no means exhaustive. However, they can serve as a template to answer specific

questions related to the design of a successful decision support system for farmers. These questions can be summarized as follows:

- What are the necessary processes and variables required for implementing a crop growth model to simulate the yields of a wide variety of crops?

- How can these yields, and other relevant outputs from the crop growth model, be used to make recommendations to the farmer for optimizing their economic and environmental gains?

- How can existing decision patterns of farmers be incorporated into the decision support system to provide a contrast between the status quo and optimal decisions?

The discussion of the background and the proposed research questions associated with the various solutions has provided us with the context for formulating the thesis statement for this research.

Thesis Statement: *A multi-model simulation framework can be used to analyze farm management practices at the crop level, individual farm level and at the community level to show the impact and alternatives for smallholder farming practices. Such a multi-model approach can be used as a decision making aide for farmers to educate themselves on sustainable farming practices, and support the decisions that result in improved economic and environmental conditions. An improvement in these conditions may also lead to better social situations and status for the farmers.*

1.3 Proposed System

In this section the proposed system and its components will be presented. The design requirements are listed and broken down into sub-requirements. These requirements have been derived from the various problems identified in the previous section. The design requirements for the final system can be listed as follows:

1. *Identify and implement a crop growth model that can initially model a specific set of crops.*

   A mathematical model necessary for simulating crop growth needs to be described and implemented. The necessary parameters that identify the different crop types and the data associated with those parameters will be collected. Sensitivity analysis should be conducted to identify the parameters that have a direct impact on plant growth. We also need to investigate other possible outputs that a crop growth model can provide in addition to the crop yields.

2. *Identify and implement a mathematical programming (MP)/Linear programming(LP) model for studying crop rotations and resource optimization.*
   We need to identify the variables necessary to formulate an MP/LP problem that utilizes crop yields from the crop growth program to select the best possible rotation of crops. We will have to develop objective functions to model different objectives for the farmers, i.e. optimizing yields vs optimizing revenue vs minimizing environmental impact.

3. *Identify and implement a decision making methodology to simulate the farmer decision process.*

A mathematical model to simulate the decision making processes of farmers needs to be developed. We have to identify the constraints and factors affecting the various decision making processes. Utility functions relating to the economic, environmental, and social objectives of farmers needs to be developed. The developed model needs to be evaluated using some real world data. Various scenarios for implementing these models need to be created.

4. *Create an input method for the users to evaluate the model.*

The system should contain flexible input requirements. Users with limited technical knowledge of the system should not be challenged while using it. Users with more knowledge of the system should be able to customize their model to a higher degree.

Requirements	Problem
1. Implement a crop growth model.	1. Integrate modeling methodologies with crop growth program.
2. Implement a MP/LP model for crop scheduling.	2. Does not require unrealistic input.
3. Implement a decision making methodology.	3. Address problem from economic, environment and social perspective.
4. Input method to evaluate model.	
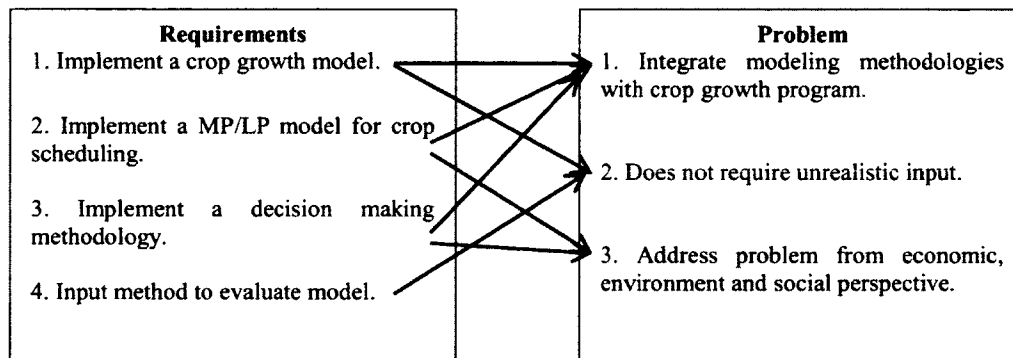
Fig. 1. Mapping system requirements to the problem.

This list of design requirements in Fig. 1 should be able to help us develop a system that can model the farmer decision making process from various perspectives. Using these requirements as a basic structure for the system, Fig. 2 can be considered as a conceptual model of the system. The various layers of the model are:

Fig. 2. Proposed Architecture of the system.

1. **Crop Growth model**: This model is the base component of the architecture. It consists of the necessary equations to model the growth of crops. Users can interact with this layer by specifying the type of crop they would like to plant, the season of planting, soil attributes and weather conditions. The crop growth model is capable of giving the yield and also the height of the crop. The output of this layer is used in both the optimizer and the individual decision model.

2. **Optimizer**: The optimizer is the second model of the architecture. Its purpose is to compute the optimum crop/combination of crops that the farmer can plant in order to maximize his profits, while maintaining a certain level of environmental friendliness. We can also generate the utilities associated with producing the optimal crops. These utilities can then be compared to the utilities from the previous layer.

3. **The individual decision model**: This is the third model of the system. It represents the decisions available to each individual farmer, based on the resources available to him. The model then computes the utility of the decisions made by the farmer.

4. **The community decision model**: This is the highest level of the model. It deals with the decisions that a community can make as a whole, in order to increase the standard of their agricultural practices through higher profits and environment friendly policies.

1.3.1 Proposed Contributions

From this research effort, there are a few contributions that I hope to make to the study of decision support systems for farmers.

1. An integrated framework of multiple models that can act as a decision aide for farmers.

2. Identifying and evaluating the important variables of a generic crop growth program that can be used in a decision support system and the development of the generic model.

3. Creating a model that can represent farmer decision processes.

4. Develop a combined model that can compare the decisions of a farming community and the economic impact to that of an optimal set of decisions.

In the next chapter we will take a look at the various decision support systems available in existing literature. The important features of these systems will be discussed along with the relevant background concepts.

Chapter 2

BACKGROUND AND PREVIOUS WORK

The general body of literature surrounding the use of simulation in decision making aides for farmers, can be split into four broad groups:

1. Using crop growth programs to simulate actual crop yield for various crops and soil regions. Additionally soil erosion, water run-off and nitrogen leeching are also estimated ([20], [21]).

2. Using optimization techniques or mathematical programming techniques to achieve the most profitable cropping sequences, farm policies, or resource management objectives ([16], [22]).

3. Using decision making methodologies like influence diagrams (ID), fuzzy logic systems, and agent based modeling (ABM), to simulate the various farm level and community level decisions that farmers might have to make ([19], [23]).

4. Applying game based methods and participatory simulations to teach farmers the importance of co-operative decision making ([24], [25]).

These groups of research present unique perspectives to the role of simulation in farmer decision making processes. These unique perspectives allow for research groups to apply a variety of methodologies to the same problem. It also gives birth to various research questions that forces research groups to analyze the problem in a variety of contexts. In order to understand the full capabilities of these various methodologies, we will need to take a closer look at the existing literature for these methods.

2.1 Crop Growth programs

The proposed system has a crop growth program as the starting layer of the simulator. In this section we will summarize various crop growth models that are widely used. Understanding these models can help us during the design and implementation of a generic crop growth model for our purposes. Crop growth programs are used to simulate crop yields, bio mass, plant height, CO_2 emissions, nitrogen content and soil carbon dynamics. Though, the chemical equations governing various crop physiological processes have been formulated for some time, it is only in the past few decades that robust mathematical model of crop growth have become prominent.

Crop growth programs are mostly implemented as a system of differential equations. Differential equations are used to simulate the dynamic behavior of complex systems. Systems are often represented as a collection of individual components. These components contain state variables and processes, and are described using mathematical equations. The structure of these components and the communication between them drives the behavior of the system as a whole.

Crop growth programs use these concepts to simulate the various physiological processes that aid in crop growth. The crop system is described using various state variables. These state variables change based on mathematical representations of the physiological processes that affect them. In software implementations of the crop growth programs, these processes are called either sequentially or are based on a discrete time based flow chart.

The general flow of physiological processes between different crop species is assumed to be similar. All species of plants need to perform Photosynthesis to convert

solar radiation into useful carbohydrates. The rate of photosynthesis varies between crops. It is based on crop specific coefficients, the available radiation, carbon dioxide levels, and available water [26]. The chemical equation of photosynthesis is written as:

$$6CO_2 + 12H_2O \xrightarrow{Light} (CH_2O)_6 + 6O_2 + 6H_2O \qquad (1)$$
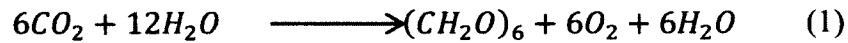
where, $(CH_2O)_6$ is the carbohydrate that has been produced from CO_2 and water through the use of light. Some of these carbohydrates are partitioned to the various organs of the crop, while the remaining is used for daily maintenance. The partitioning rates for the different organs are a crop specific parameter. The mathematical model for generic crop growth and the various variables involved will be explained in the next chapter.

The literature review for crop growth models a number of successful research efforts that have used the models as decision support systems. The research group in [27] have developed CropSyst, which is one of the most accepted tools for simulating crop growth. It uses existing mathematical models of crop growth to create a process oriented approach with attention to detail on software design to enhance interoperability. This has led to the adoption of CropSyst in many large scale simulations using high performance computing.

CropSyst has been used by various research groups to study multiple cropping phenomena. The availability of auxiliary tools like ClimeGen by [27], which simulates climate data, has increased the popularity of this tool. Researchers [20] have used this program as a decision support system to create an irrigation schedule in a pear orchard with a high yield prediction accuracy. Research groups [28] have also successfully applied this method to predict water use in Japanese plum trees. These simulations configure the

CropSyst program to the local conditions by combining it with field and experimental data. Fig. 3 shows a high-level diagram of the CropSyst model.
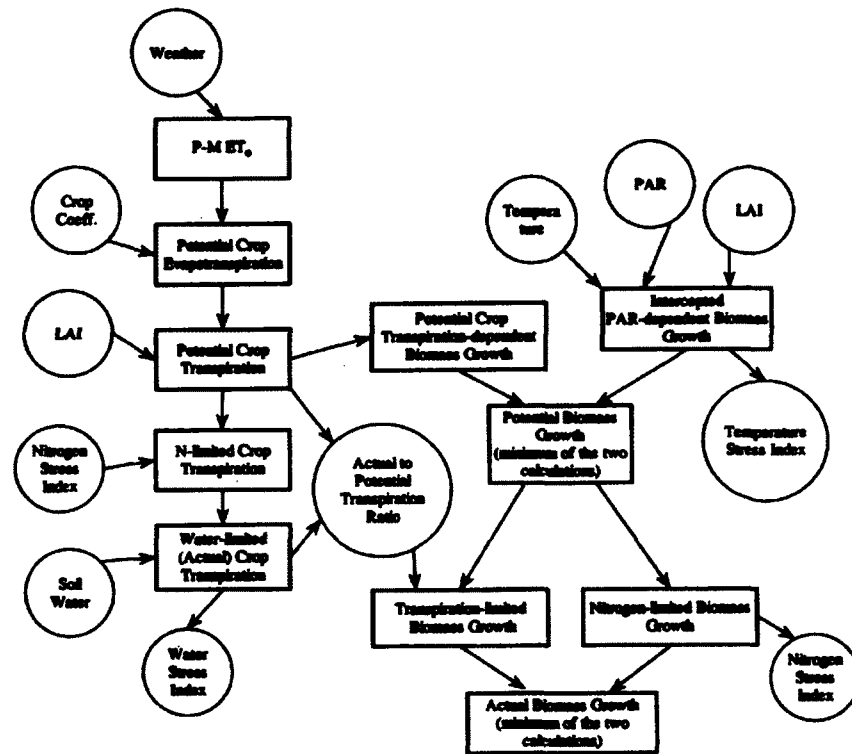
Fig. 3. CropSyst Model [27].

The main drawback of this tool has been the huge amount of data required, to get any meaningful results. The program requires detailed quantification of soil types, cropping procedures, and crop phenology. In the design of generic decision support systems, this data requirement can provide a significant roadblock.

The research team in [21] have developed another widely accepted crop growth model called Agricultural Production Systems simulator (APSIM), that can simulate the physiological process in farming systems. The following diagram shows an overview of the APSIM simulation framework.

Researchers in [26] have used this model to simulate the crop growth of German winter wheat, maize and field-pea varieties. They have modified existing formula to account for canopy conditions, and used it to study mono-cropping and intercropping effects of these crops. Additionally, the authors in [29] have used APSIM to model crop improvement strategies by combining it with genetic and breeding information. This has resulted in identifying interesting traits and genes that can increase crop yield.

Though, the APSIM model is useful in studying crop management, water balance, cropping systems, and inter-species interaction, its drawback lies in the limited number of crops available for simulation, and similar to CropSyst, has significant data requirements.

The authors in [30] provide an exhaustive study of the ecological and geochemical factors that affect plant growth. They provide a set of the most relevant equations necessary to model crop biomass growth. The main strength of the model lies in the fact that the authors have used the model in real time by conducting experiments in various parts of China and the United States. This model provides the basic science necessary to understand crop growth. In keeping with the trend of the previous crop growth models, the unrealistic data requirements have limited its use in generic decision support systems.

In addition to these multi-crop growth models, there exist crop and environment specific models, which are used to predict yield of certain crops. The authors in [31] have used a complex crop growth model, coupled with geographic information system land use maps, digital elevation models, soil texture, and radar data, to model rice growth on a large scale in the North China Plain.

Similarly, the authors in [32] have developed a crop growth model call PIXGRO to simulate the carbon dioxide exchange in the environment and its impact on the growth of

spring barley. It combines plant physiology, remote sensing, and crop growth modeling to simulate yield and carbon levels at the landscape, regional, and continental levels.

Though the crop growth models available in the literature have unique characteristics and methods of implementations, they are primarily based on a common structure. We will be using this structure while selecting the equations for our generic model, as described in Chapter 3.

2.2 Linear programming methods

The second layer of our proposed architecture is an optimizer. As such it is responsible for providing the correct selection of crops and their respective rotations for a given set of objectives. This leads us to consider optimization techniques for implementing this layer. Mathematical/Linear programming (MP, LP) models have a rich history of usage in modeling farming practices. Empirical models have been used to study the relationships between farming practices and related economic and environmental impacts by utilizing historical data. Regression models have been constructed based on this data and have been used to analyze the co-relation between existing farming procedures and the economic gains that they produce.

LP models first came to prominence in the book Mathematical Method of Production Management and Planning [33], detailing LP models and their solutions. It was followed by F.L.Hitchcocks original paper [34] on transportation which also used LP models. The late 1940s and 50s saw a growth in these models due to the introduction of the Simplex method for solving LP problems.

Mathematically, an LP problem can be written as:

Maximize $z = \sum_j c_j x_j$ (2)

Subject to $\sum_j a_{ij} x_j \leq b_i$ $(i = 1,2, \ldots\ldots, m)$ (3)

$x_j \geq 0$ $(j = 1,2, \ldots\ldots, n)$ (4)

In the problem statement above, we seek a set of values for the continuous variables

$x_1, x_2, x_3 \ldots\ldots\ldots x_n$ to maximize a linear objective function z, and satisfy a set of linear

constraints. An Integer Linear Programming (ILP) problem is one where at least one of the

variables takes only integer values. A Mixed Integer Programming problem (MIP), is an

LP where some, or all of the variables in the problem have an integer restriction.

Mathematically, a MIP can be written as:

Maximize $z = \sum_j c_j x_j + \sum_k d_k y_k$ (5)

Subject to $\sum_j a_{ij} x_j + \sum_k g_{ik} y_k \leq b_i$ $(i = 1,2, \ldots\ldots, m)$ (6)

$x_j \geq 0$ $(j = 1,2, \ldots\ldots, n)$ (7)
$y_k \in 0, 1, 2, \ldots$ $(k = 1,2, \ldots\ldots, p)$
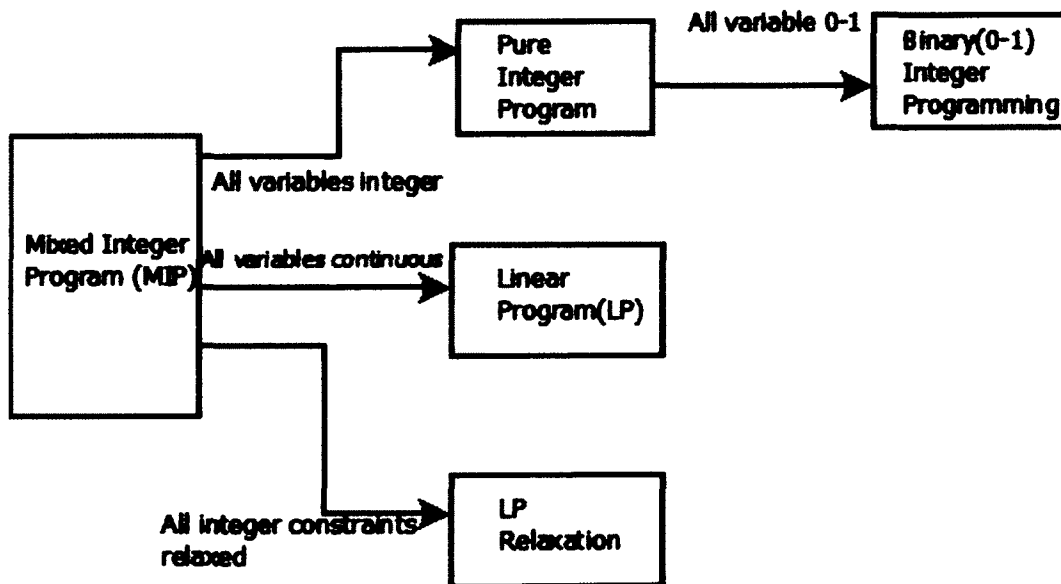
Fig. 4. A simple classification of the various LP problems.

Fig. 4 shows a simple classification of the different MP/LP models. These models are then used to predict future economic gains by extrapolating the results for the time-series data. The major drawback of this method is the model rigidity, which does not allow for the inclusion of new farming practices into the model, and can hinder the model from producing relevant and accurate results. The main contributions of these regression models are the help that they provide to researchers in identifying the necessary indicators that have a direct impact on the economic and environmental gains.

The analysis of regression models is able to produce the coefficients and weights required while formulating linear programming models. These coefficients and attribute weights can then be used to draw inferences on the constraints, and objective functions that make up an LP problem. Depending on the nature of the problem being studied, the usage of LP models can be divided into two sub-groups:

1.   Studying the effects of government policy on the socio-economic and environmental gains for farmers and farms.

2.   Analyzing various configurations of farm resource utilizations in order to maximize expected economic or environmental utility. It can also be extended to study the application of new farming technologies, by introducing them as constraints to the problem, or by eliminating existing constraints.

2.2.1 Linear Programming and Government Policy

The literature shows a widespread acceptance of the use of LP models to study the impact of governmental policies, on farm level activities. The authors in [22] have used

this approach to model the impact of European Union (EU) common agricultural policy (CAP) reforms on farm level economics, through the use of linear programming. The authors simulated various agricultural, technological and market scenarios for, rice, fruit and vegetable farming systems. The authors were able to show a correlation between water pricing and economic impact on farms. This economic impact was then used to show a co-relation to negative environmental impact on the surrounding farming areas.

Research groups [16] have conducted a study of the EU water policy, water framework directive (WFD), on its impact on irrigated farming in the EU. WFD aimed to use the pricing of water as a tool to regulate water use, and water pollution. The authors created various scenarios for multiple agricultural scenarios, to describe the application of these governmental policies. The LP models for these scenarios showed conflicting results for different agricultural systems. For example, in the cereal system, WFD was shown to have caused a drop in farm income and an increase in unemployment, while promoting a positive impact on the environment. On the other hand, fruit based agricultural systems showed a lot of resistance to the impact of WFD on farm incomes with negligible improvement of the environmental indicators.

The authors in [17] has extended the study of the impact of governmental policies on farms, to include dairy farms in the Dutch region. The authors selected two governmental policies, namely, Mineral Accounting System (MINAS), and Manure Transfer Agreement System (MTAS), which compel dairy farmers to reduce nitrogen and phosphate loss to ground water. The author found that, strict adherence to the government environment policies caused a decrease in surplus for the farmers. It was also found that, in high intensity farms, the cost of complying with the environment policies was much

greater than the penalty that the farms would have to pay for breaching the said policies.

[35] extended the research on MINAS and MTAS to study the environmental impacts of

such policies in dairy farms. The authors found that the enforcement of the environmental

policies caused a significant positive impact on the various environmental indicators.

Government policies regarding pesticide and fertilizer taxation were studied by

[36]. The authors aimed to develop economic incentives for farmers who had reduced

pesticide application levels. Linear programming models were created to simulate a base

scenario for farm level activities that followed governmental policies. The authors then

introduced additional constraints on the pesticide usage by providing economic incentives

to farmers for reduced pesticide use. Using an environmental based objective function, the

authors were able to optimize various scenarios that produced a maximum environmental

utility. The authors then extended this approach to include a trade-off between

environmental and economic objectives. This was useful to suggest environmental policies

that could potentially be beneficial to both the farmers, and the environment.

The authors in [37] applied a similar research procedure to study the impact of

government taxation policy for fertilizers on NO3 pollution in water. The authors were able

to conclude that a combination of the various policies, rather than implementing each

policy individually, would be able to decrease the overall pollution by 20%.

2.2.2 Linear Programming and Resource Utilization

One of the reasons for employing LP models at farm level is the ease with which

various farm activities can be represented as resources. Crop rotation, fertilizer and

pesticide usage, water usage, labor requirements, and available capital can all be modeled

as resources in a LP model. The access to various optimized LP tools on the market and the increased processing power of computers has enabled researchers to build more complex problems without worrying about processing constraints. A quick literature review on this topic shows us the use and acceptance for LP in farmer resource utilization.

The authors in [18] use LP to formulate a crop rotation scheme with no explicit, pre-determined rotations. The coefficients and weights were generated using regression analysis of historic cropping data records. The main drawback of the paper is that it does consider the validity of the rotations that are obtained as outputs. The crop rotations that are suggested by the model might be financially rewarding options, but they are not necessarily feasible. This could be due to the soil exhaustion from planting the same set of crops over and over again, financial constraints, or weather constraints. The absence of these factors, gives us outputs that are mathematically accurate but practically inaccurate. Though the model has a few drawbacks, it does provide a basic framework to create more complex crop rotation mechanisms by introducing more constraints and factors to the model.

The researchers in [38] also follow a similar approach where not all possible crop rotation schemes are enumerated. This creates a more efficient method for evaluating crop rotation schemes by solving only a subset of the possible crop rotations. This rotation model provides an in-depth approach into crop rotation by considering factors like money spent on each crop in both the pre and post cropping stages. It also provides a measure of judging the effects of the crops on each other.

The main drawbacks of such an approach can be classified into two parts. First, this procedure advocates eliminating the crop which has the least value for the maximizing

function. The authors do not mention the role of the dropped crop on the overall effectiveness of the soil quality and its effect in the yield of the remaining crops. Due to the absence of a variable that considers the effects of the chosen crops on the soil, the crop rotations that are yielded by this model might give a less than satisfactory result if applied practically. Second, the authors do not note the effects of the market on the crop production. Since these crops are planted in cycles over a period of time, it is possible that a change in the market value of the crops would change the values of the maximizing functions, thus affecting the crop rotations.

Similar research methodologies are applied by various research groups like [39], [40]. In these LP models, the objective is the effective utilization of external resources like warehouses, labor, and forest resources. The main contribution of these research groups is in the identification of external resources which can have a direct or in-direct impact on farming policies. This contribution can also be considered as a drawback for these models since; the excessive importance given to these external objectives might produce results that do not simulate real world results meaningfully.

From the literature review we can see that there are multiple ways of using linear programming methods to define farming problems. These various uses are mostly based upon the research question defined by a research group. This leads to the possibility of having a large number of problem formulations to cover every possible farm specific scenario. Since it is not possible to cover this exhaustive list of scenarios in our model, we will use a simple, yet flexible, set of formulations that can be expanded or simplified as necessary.

2.3 Decision Making Methodologies

The top two layers of our system are an individual and a community decision model respectively. We will be using these models to simulate the farmer decision process at an individual and a community level. This process requires the selection of a decision simulating methodology. In this section we will review popular methodologies used to simulate decision making and understand their strengths and drawbacks.

Decision theory is the study of decision making processes, where the ultimate goal is to derive the best decision that can maximize the expected utility of the decision maker's final objective. This decision is made with either full or partial information regarding a certain scenario. Though, the general approach of LP and decision theory appear to be similar, i.e. maximization of expected utility, they have fundamental differences in their use. LP and mathematical programming is used when various scenarios can be expressed as functions of variables. Decision theory is used when the effects, and frequencies, of various decisions and corresponding constraints can be quantified.

Since decision theory is used to model the actual decision making processes, it is a prime candidate to use in farmer decision making processes. Some important decision making processes are: Bayesian trees and influence diagrams (ID), decision trees, fuzzy set theory, and Agent based systems (ABS). Decision theory is especially useful when modeling farmer decision making under uncertainty and risk averse conditions.

2.3.1 Influence Diagrams

Influence diagrams, also called relevance diagrams, are acyclic directed graphs representing decision problems. The goal of influence diagram modeling is choosing such a decision alternative that will lead to the highest expected gain (utility).

Similar to Bayesian networks, influence diagrams are very useful in showing the structure of the domain, i.e., the structure of the decision problem. Influence diagrams contain four types of nodes (Decision, Chance, Deterministic, and Value) and two types of arcs (influences and informational arcs).
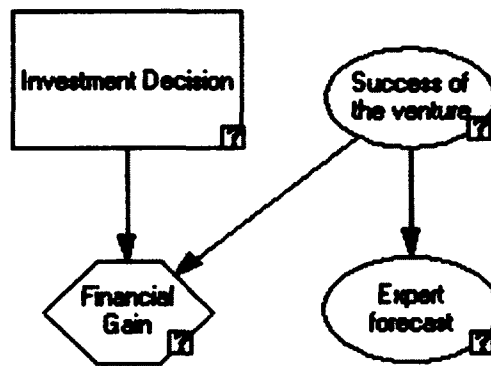
Fig. 5. Example of an Influence diagram.

In Fig. 5 the ovals represent the chance nodes. These nodes represent uncertainty in the system. The uncertainty associated with each chance node in the system is represented by a conditional probability potential:

$$P(C|\mathbb{I}\!\mathbb{I}_C) : sp\left(\{C\}\bigcup \mathbb{I}\!\mathbb{I}_C\right) \to [0; 1] \qquad (8)$$

where, $\mathbb{I}\!\mathbb{I}_C$ denotes the parents of C in the influence diagram. The rectangles represent the decision nodes. They contain the list of decisions available in the system at that point. The

hexagons represent the utility nodes. These nodes are used to quantify the impact of the decisions for a given set of chance nodes.

The purpose of the influence diagrams is to calculate the decision alternative, or the set of decision alternatives, that maximizes the expected utility of the system. For a set of chance variables **C**, utility variables **V**, and decision variables **D**, the optimal sequence of decisions that maximize the expected utility is given by:

$$\delta_{D_n}(C_0, D_1, \ldots \ldots, D_{n-1}, C_{n-1}) \tag{9}$$

$$= arg\,max_{D_n} \sum_{C_n} P(C_n | C_0, D_1, \ldots \ldots, D_{n-1}, C_{n-1}, D_n) \sum_V \varphi_V$$

where, δ_{D_n} is the optimal policy for the decision variable D_n and φ_V is the set of utility values for the utility node V.

The authors in [41] have demonstrated the use of IDs to study the motivation for farmer decisions related to farming methods and organic farming and the differences between conventional and organic farmers. Using this model he was able to classify organic farmers into 4 different categories. He was also able to identify the constraints that prevent conventional farmers to adopt organic farming techniques. This analysis has helped the author to suggest effective government policies to encourage organic farming. These include policies to encourage positive attitude towards organic farming, policies to promote research, and policies regarding economic incentives for organic farmers.

Influence diagrams have also been used to model farmer decision making process to stem the losses to biodiversity. The researchers in [23] have used this method to study the fencing practices of land holders in Australia. They were able to identify and understand the impact of causal relationships of the different variables like farm income, long term vision government support, and landholders' knowledge.
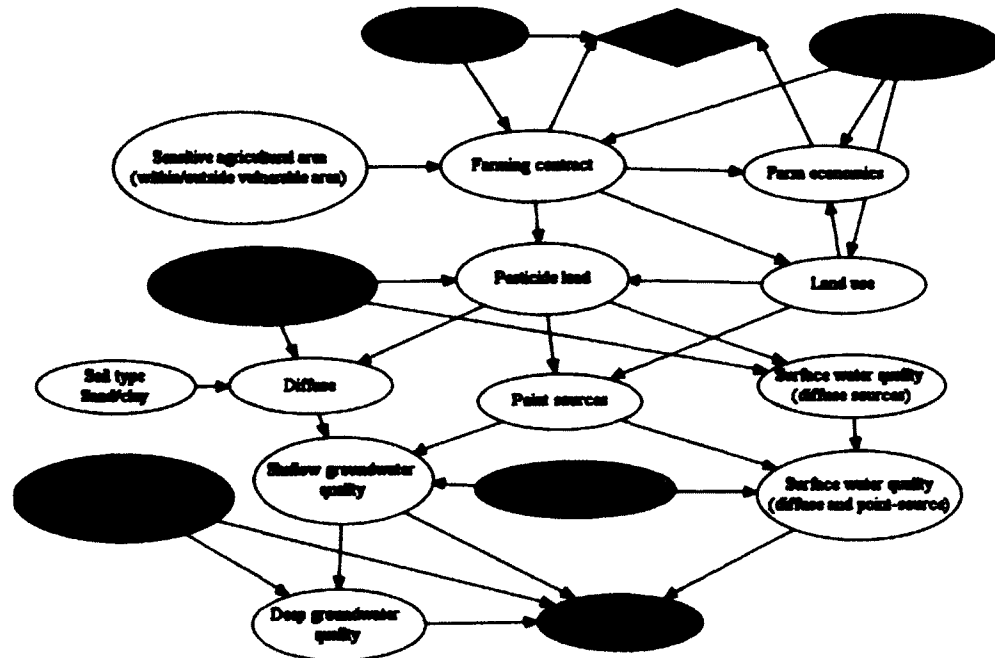
Fig. 6. ID to evaluate ground water management.

Bayesian networks have also been used to evaluate resource utilization by farmers. The authors of [42] have used Bayesian/Influence diagrams to model farmer decision process in the use of ground water management. The authors were able to use existing data to create an Influence diagram that analyzes the tradeoff between different objectives, and propose management options that can balance contradictory objectives. Fig. 6 shows the Influence diagram employed by the authors to evaluate the groundwater management.

2.3.2 Fuzzy Logic

Type-1 fuzzy set (T1 FS) theory was first introduced by Zadeh in 1965 and has been successfully applied in many areas, including modeling and control, data mining time-series prediction etc. A fuzzy set F is defined as a pair (F, μ_F). For every element $x \in F$,

the value $\mu_F(x) \in [0,1]$ is called the membership of x in (F, μ_F), and μ_F is called a fuzzy

membership function.
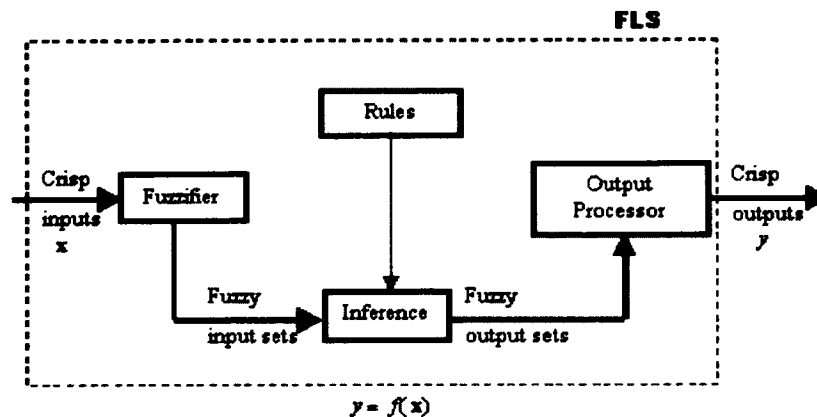
Fig. 7. Type 1 FLS

Though fuzzy sets have a wide range of applications, their relation to rule based

fuzzy logic systems have made them a popular tool in the engineering community. A rule

based fuzzy logic system (FLS) is shown in Fig. 7. It consists of a fuzzifier that maps the

incoming crisp input values onto the relevant fuzzy sets. A rule base that contains the rules

describing a FLS that is described completely in terms of type-1 fuzzy sets is called a type-

1 FLS.

Fuzzy membership functions are often represented using simple functions. They

are useful in graphically representing a fuzzy set. Piece-wise functions can also be used to

represent membership functions, but most applications use mathematical functions. Fig. 8,

Fig. 9 show a triangular and a Gaussian membership function.
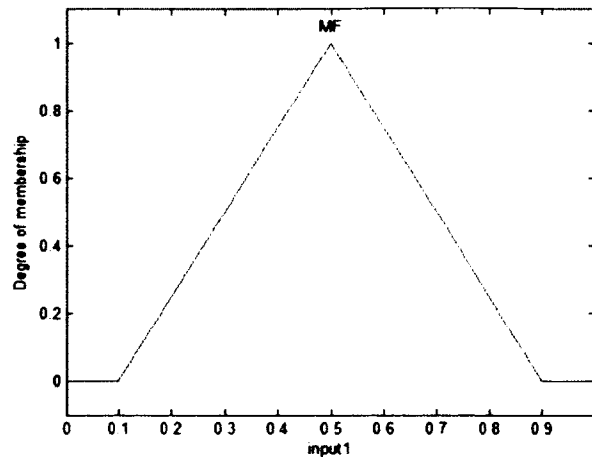
Fig. 8. Triangular membership function.
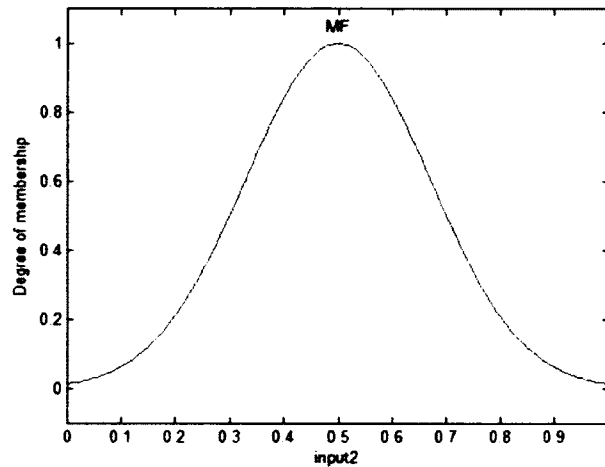
Fig. 9. Gaussian membership function.

When a data point is given as an input to an FLS, the inputs are mapped to fuzzy values based on the membership functions. Theses fuzzy values are then used as an input to an IF-THEN rule base to calculate the outputs. These rules are designed based on either existing data or expert knowledge.

The research group in [19] has used fuzzy methodology to study the motivations of farmers to diversify their farms in terms of cropping practices, crops used, and other technological improvements. Using linguistic variables based on rural farmer knowledge, the authors were able to build a relatively complex system that was able to model individual motives of farmers, rather than simple utility maximization. Though the use of this model for individual farms can be argued, it can help in creating a hierarchical model for groups of farms to study and implement government policy changes.

Fuzzy systems have also been used to study the environmental impact of farming systems. The authors of [43] have used existing data to create an FLS to propose strategies to increase the positive environmental impact of sugarcane farms in Iran. The authors were able to validate their proposed solutions through interviews with experts, and existing scientific knowledge.

Other researchers [44] have used a similar methodology to evaluate the risk to environmental biodiversity due to farming practices. The authors used biodiversity indicators to successfully identify farming practices that are suitable to maintain biodiversity. Organic farming practices like organic fertilization and reduced tillage were evaluated and were found to have positive effects on the environment while maintaining acceptable profitability levels for farmers.

2.3.3 Agent Based Systems

Agent based systems/models (ABS/ABMs) are helpful in simulating the behavior of individual and group entities/agents in the context of their environment. The flexibility to express the behavior of the agents in relation to the changes in their environment makes

ABS a powerful tool to simulate farmer decision processes. It can be used at both the individual farm level, and also a large community level.

Agent based models are simulations of groups of agents, who interact based on predetermined rules, set in a virtual environment. The agents are represented by attributes that describe their motivations or characteristics. The interactions between agents need to be explicitly defined. These interactions lead to emergent behaviors and changes in the environment that are later analyzed. The rules of interaction are often simple IF-THEN-ELSE rules that are developed by the modelers or domain experts.

Various research efforts have successfully used ABMS as a modelling tool to predict, and analyze, the impact of the farmer behavior at the farm, county, and country level. A review of the use of ABMs in farmer decision modeling was performed by [45]. The author found that ABMs were fast becoming the preferred tools for simulating farmer-environmental interactions.

An ABS has been used by [46] to model the impact of farmer behavior in relation to water and resource usage. It employs a simulated hydrological system and social dynamics to analyze the outcomes of various scenarios. The authors combined the model with biophysical models to simulate the hydrological system. They were able to extend the model to study the response of farmers to unforeseen disasters, and the impacts that they might have on farmer income.

A similar land use change model was built by [47] to simulate the cropping patterns of farmers in northeastern Thailand. The authors were able to incorporate both economic dynamics and social interactions between the farmers while building the system. Through this data the authors evaluated the effects of climate change on the change in land use. This

model was applied on multiple villages and the authors have identified the long term response of each village to the negative effects of climate change.

ABS has also been used as a training tool [48] by combining it with a participatory role playing game. This approach uses an ABS as a bridge between researchers and farmers by letting the researchers help the farmers in understanding and analyzing the results and limitations of the system. ABS can also be used for making policy recommendations using livelihood indicators as demonstrated by [49]. This approach shows the effectiveness of policy recommendations based on the use of ABS, despite data limitations.

This section on decision making methodologies has provided us with a large amount of information on the various decision making methodologies used to address agricultural problems and the context of their usage. Influence diagrams are effective tools to capture the structure and causality of the problem while attaching appropriate utility values to decisions. This is very useful during problem formulation as it helps us to identify the various factors and the relations between them. However, the definition of point probabilities for the various chance nodes requires large amounts of data or reliable input from subject matter experts. Increasing the variables also increases the number of probabilities to be calculated exponentially.

Fuzzy logic systems seem appropriate and flexible for addressing farmer problems. They allow the user to translate the uncertainty of the decision process through the use of fuzzy rules and membership functions. Defining these rules and membership functions can be considered a drawback of this method as a large chunk of the decision process is quite subjective and qualitative. Agent based methods can be considered to be a perfect fit for simulating decision process on a large scale. However, similar to the fuzzy systems

defining the necessary rules for the agents to act upon is not always accurate. This can lead to simulation models that might not be a close enough representation of the problem being addressed. In the decision model for the proposed architecture we will use these various decision making methodologies and address the problems associated with their use.

2.4 Participatory and Game Based Learning

In the previous sections we covered the various methodologies associated with the different layers of the proposed architecture. In this section we will take a brief look at decision support systems that have been developed to act as teaching tools for uneducated/technologically illiterate farmers. This review can be useful in making recommendations for implementing the proposed models for a target audience of this type.

Farmer decision support systems, which started out as purely mathematical and quantitative exercises, have recently started incorporating qualitative learning aspects into their methods. The previous decade has seen an emergence of game based technologies and participatory methods to educate farmers about the different options available to them w.r.t. new and improved farming practices. These research efforts aim to target the farmers thought process by presenting problems to them from multiple points of view. Game based approaches have tried to inspire farmers to adopt sustainable practices by presenting visual consequences of their decisions. These game based and participatory efforts are usually targeted towards smallholder and undereducated farmers who tend to resist new farming practices because of societal restrictions and risk averse behavior.

FARMSCAPE, developed by [24] is a participatory research targeted towards farmers in northeast Australia. It involved a 10 year study to evaluate the benefits of

simulation and modeling in educating farmers and their advisors. This research stemmed from a need to connect researcher and farmers in order to develop appropriate technologies for helping resource poor farmers. The authors provide performance indicators of different farming practices evaluated to create a learning experience. This co-operative learning experience for both the farmers and researchers has led to the creation of better farming practices by the researchers and an improvement in the analytical ability of the farmers.

The researchers in [25] have used a participatory research methodology to study the social and ecological effects on land use change. They further developed this model test alternative scenarios and farming practices. From their discussions and model simulations, the researchers were able to observe a decrease in soil fertility due to the overuse of fertilizers and hybrid crop varieties. They were also able to conclude that adoption of alternative, and more responsible cropping methodologies, would not be able to counter the degradation of soil fertility. Fig. 10 shows a causal diagram of the participatory sessions used to collect data on farmer behavior.
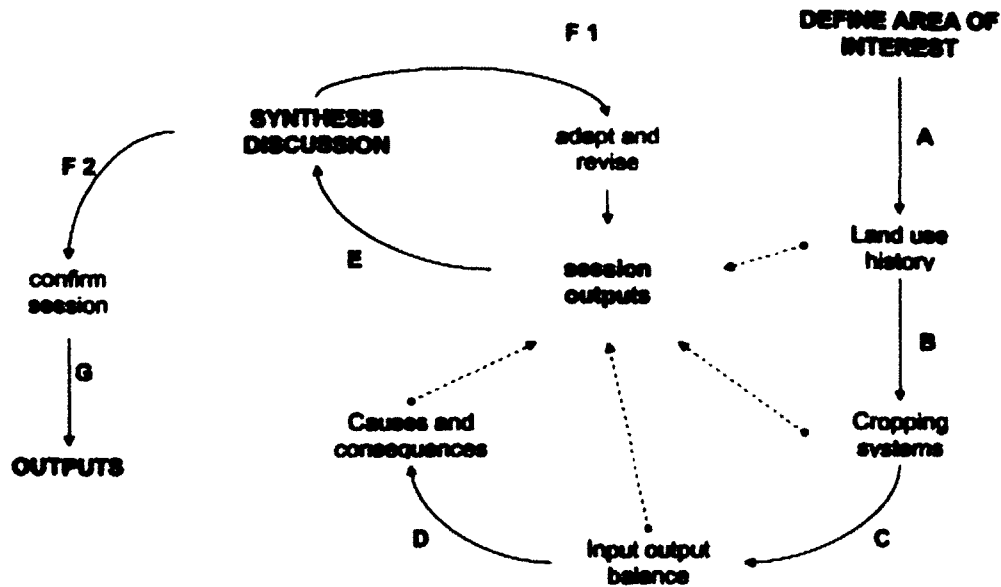
Fig. 10. Causal loop diagram of participatory methods [25].

Participatory simulation has also been used to identify solutions to water resource issues in watersheds. The research group in [50] applied a participatory approach to address phosphorous control options in St. Albans Watershed, Vermont. Through the discussions with various stakeholders, they were able to collect data and evaluate scenarios to reduce phosphorous load on the watershed. They were also able to provide a neutral ground for the various parties involved, to discuss the environmental effects of their actions on the watershed.

A similar approach was used by [51], to bring together farmers, lakefront property owners and city residents to discuss the environmental effects of eutrophication, nutrient runoff, water quality, and water management issues. It resulted in educating the stakeholders in alternative solutions to the water pollution problems. Data collected in this process was useful in building quantitative models for application in local decision making. It was instrumental in creating a dialogue between the various conflicting stakeholders to

achieve a common ground in terms of sustainable practices. The authors were also able to develop politically feasible solutions that could be implemented by lawmakers.

Participatory approaches are often combined with role playing games (RPGs) in multi actor systems to arrive at acceptable compromise solutions. These approaches are often in the form of board games or use computer interfaces. These games provide a sense of legitimacy to the alternative approaches proposed by researchers, since the farmers get a chance to visually perceive the changes to the environment, and their livelihood, based on the outcomes of these games.

The authors in [52] have used an RPG to educate farmers about effective actions to combat agricultural runoff. The actors included farmers, mayors, and advisors. They were tasked with finding a solution to disastrous runoff scenarios in the watershed by utilizing a game. The stakeholders were able to arrive at effective solutions to reduce the runoff by 50% after engaging in discussions about intercropping periods, and storage tanks.

Similarly, the researchers in [53] have used RPGs to study the decision making abilities of farmers when placed in a virtual environment with simulated scenarios to tackle problems of resource sharing. Their goal was to observe if farmers adhered to the game rules and if they were able to use the knowledge gained from their farming experiences to make decisions in the game. A similar approach was used by [54] to study how genetically modified crops and conventional crops could coexist in oilseed and maize farming. The RPG sessions were useful to understand the decision making processes of farmers and grain merchants. They were also able to evaluate the amount of co-operation between the actors and their response to advice from technical advisers.

2.5 Conclusions

In this chapter, the existing literature for decision support systems, and their respective concepts have been reviewed. We were able to look at the broad spectrum of research methodologies being used to address the problems faced by the farmers. Some of the approaches were quantitative, while others were purely qualitative. The selection of appropriate methodologies for modeling the various problems is a research problem in itself. It is also important to note that while some quantitative models were built and analyzed efficiently, the results are not intuitively understandable to farmers. Qualitative models can also suffer the same problem by not providing a realistic context to the farmers while simulating various scenarios.

A decision support system aimed at farmers should be able to provide results in understandable and contextually relevant ways. The review of the various methodologies was helpful in identifying possible directions for our research and the potential pitfalls of certain methods. In the following sections we will take a detailed look at the various methods used in implementing the different layers of the proposed architecture.

CHAPTER 3

CROP GROWTH MODEL

The mathematical model necessary for crop growth has been developed from existing resources. Several mathematical models are available to simulate the growth pattern of various crops [55], [56]. The drawback of these models is that they are crop specific. Since our project did not need the complexity of the various crop specific models, we have attempted to use a generic crop growth model to simulate the crop bio mass yields and plant height. In this section, we will look at a form of the crop biomass equation. The individual variables in the equation will be explored to see how we have arrived at the final form of the equation.

3.1 Generic Crop Growth Equation

A generic equation for plant biomass growth [56] can be written as:

$$\frac{dW}{dt_i} = \varepsilon \times 0.5 \, Q_0 \, [1 - e^{(-k*LAI_{i-1})}] * 0.0001 \tag{10}$$

where, $\frac{dW}{dt_i}$ is the daily increment in biomass weight for the crops, in tonnes/hectare (ta/ha). Q_0 is the daily solar radiation in MJ m^{-2} d^{-1}. ε is the radiation use efficiency that converts the daily radiation into photosynthetically active radiation that is used by the plants. This coefficient is crop specific. k is the extinction coefficient. It is generally assumed to have a value of 0.65 for all crops.

LAI_{i-1} is the Leaf Area Index for the previous day. LAI is a dimensional quantity that represents the one sided green leaf area per unit ground surface. In order to evaluate LAI, we need to calculate heat units, heat unit index, and heat unit factor for each day. A

crop starts growing once the daily average temperature exceeds the base temperature for the crop. Daily heat unit is the difference between the daily average temperature and the base temperature required for germination. Heat unit(HU) is given by [57] as:

$$HU = T_{Avg} - T_{base} \qquad (11)$$

Each day, if the value of the heat unit is greater than zero, it is accumulated as part of the total heat units HU_{tot}. These accumulated heat units are divided by the potential heat units for a crop to arrive at the heat unit index (HUI).

$$HU_{tot} = HU_{tot} + HU \text{ for } HU > 0 \qquad (12)$$

$$HUI = \frac{HU_{tot}}{HU_{pot}} \qquad (13)$$

The potential heat units for a crop are calculated by multiplying the difference between the optimal and base temperatures, T_{opt} and T_{base}, for a crop with the total number of growing days.

$$HU_{pot} = Planting\ duration * (T_{opt} - T_{base}) \qquad (14)$$

The heat unit index, HUI, is a value between 0 and 1 that is used to measure the progress of a crops growth as a function of the daily temperature. It is also used to calculate the heat unit factor (HUF), which indicates the fraction of the maximum leaf area index for the current heat unit index.

$$HUF_i = \frac{HUI}{HUI + \exp(ah_1 - ah_2 * HUI)} \qquad (15)$$

$$ah_2 = \frac{\ln(\frac{frp_1}{frl_1} - frp_1) - \ln(\frac{frp_2}{frl_2} - frp_2)}{frp_2 - frp_1} \qquad (16)$$

$$(17)$$

$$ah_1 = \ln(\frac{frp_1}{frl_1} - frp_1) + ah_2 * frp_1$$

where, $frp_1, frl_1, frp_2, frl_2$ are crop specific parameters that provide the fraction of the maximum leaf area index reached for a specific period in the growing stages. These values are regression coefficients that researchers have determined experimentally to fit the leaf development curve.

Finally, the leaf area index for each day is given by [58] as:

$$LAI_i = \begin{cases} LAI_{i-1} + dH_{F,i} * LAI_{max} * (1 - \exp(5 * (LAI_{i-1} - LAI_{max}))), & i < decline\ period \\ LAI_{i-1} * \frac{1 - HUI}{1 - HUI_{sen}}, & i \geq decline\ period \end{cases} \quad (18)$$

$$dH_{F,i} = HUF_i - HUF_{i-1} \quad (19)$$

where, HUI_{sen} is the heat unit index when the crop enters its decline stage. During the growth stages the LAI is an exponential function of the LAI from the previous day and the maximum leaf area index LAI_{max}. Once the crop starts declining, the leaf area also starts declining as a function of the heat unit index.

3.2 Evapotranspiration

Evapotranspiration is the combined process of plant transpiration and soil evaporation. Plants lose almost 99% of the water they take up due to evaporation. This process is called transpiration. Simultaneously, the soil surface also undergoes evaporation and loses water to the atmosphere.

Evapotranspiration is used as a means to calculate the water requirement of a crop for each day during its life cycle. Evapotranspiration is heavily influenced by the climate conditions. It is high in hot and dry conditions, and low in cloudy and cool areas. Crop

evapotranspiration for each day is calculated by first calculating the potential evapotranspiration. Potential evapotranspiration is defined as the evapotranspiration that would occur from a large area uniformly covered with green vegetation with an unconstrained access to water.

Various methods have been developed to calculate the potential evapotranspiration on any given day. The Penman model, the Penman-Monteith model, the Priestly-Taylor model, and the Hargreaves model have all been successfully used to calculate daily evapotranspiration. The current crop growth model implements the Penman model to calculate the evapotranspiration.

The Penman model calculates the evapotranspiration for a short green crop, like grass, that uniformly covers the surface of the land and has unconstrained water supply. The equation for the penman model is given by [57] as:

$$E_O = \frac{\Delta * R_N + psychro * R_C * VPD}{H_V * (\Delta + psychro)} \qquad (20)$$

where, E_O is the potential evapotranspiration for any given day, measured in mm/day. Δ is the slope of the saturation vapor pressure curve in kPa/°C. Vapor pressure is the amount of pressure exerted by vapor in a closed container. It is an indication of the rate of evaporation of water from the soil surface. The slope of the vapor pressure curve indicates the speed with which the surface water content of the soil is evaporating. It is an exponential function of the daily average temperature in °C, given by the formula:

$$\Delta = 25029.4 * \frac{exp\left(\frac{T_{Avg} * 17.269}{T_{Avg} + 237.3}\right)}{(T_{Avg} + 237.3)^2} \qquad (21)$$

R_C is the aerodynamic conductance of air in mm/kPa*day and a function of the wind speed W_S in m/s. It is calculated using the formula:

$$R_C = 2.7 + 1.63 * W_S \qquad (22)$$

VPD is the vapor pressure deficit in kPa. It is used to measure the difference in the actual water vapor pressure E_a, and the vapor pressure at saturation E_s, for the daily average temperature T_{Avg} and relative humidity R_h expressed as a fraction.

$$VPD = E_s - E_a \qquad (23)$$

$$E_s = 6.1078 * exp \left(\frac{T_{Avg} * 17.269}{T_{Avg} + 237.3} \right) \qquad (24)$$

$$E_a = E_s * R_h \qquad (25)$$

The psychrometric constant is given by the variable *psychro*. It is useful in relating pressure P_B, in kPa/°C, of water in air to a specific temperature. It is given by the formula:

$$psychro = P_B * 7.2063 * 10^{-4} \qquad (26)$$

The latent heat of vaporization of water, H_V, is defined at the daily average temperature T_{Avg} using the equation:

$$H_V = 2.501 - 0.0022 * T_{Avg} \qquad (27)$$

The potential evapotranspiration calculated in the previous step is for a reference crop like grass or alfalfa. To scale this value to a specific crop, and to calculate its daily water use, we need to multiply the potential evapotranspiration, E_O, value with the crop coefficient K_C.

$$E_S = E_O * K_C \qquad (28)$$

The crop coefficient K_C depends upon the crop type, the growth stages of the crop and the climate. The general crop coefficient encompasses the evaporation from both the crop and the soil. General values of the coefficient are available, and can be used to

calculate the daily water requirement. If we need to calculate the daily crop coefficient by taking the soil type into account, we will need to split the coefficient into the crop specific coefficient and the soil coefficient. The crop coefficient K_c is given by [59] as:

$$K_c = K_s * K_{cb} + K_e \qquad (29)$$

where, K_{cb} is the basal crop coefficient. For every crop, this value is defined for the different crop growth stages: initial, development, middle, and decline. It is important to know the duration of each of these stages for each crop and its respective coefficient. Table 2 shows a sample of the basal crop coefficient for different crops from [57] and [59], at the different growth stages.

Table 2 Crop coefficients and duration of stages

Crop	Kcini	Kcmid	Kclate	Initial duration	Development duration	Mid-stage duration	Decline duration
Broccoli	0.15	0.95	0.85	135	35	45	40
Lettuce	0.15	0.9	0.9	140	25	30	65
Onions	0.15	0.95	0.65	150	30	40	60

The soil coefficient K_e is calculated using the formula:

$$K_e = K_r * (1.21 - K_{cb}) \qquad (30)$$

The values K_r, K_s are evaporation reduction coefficients that are dependent on the depth of the water depleted from the top soil for the crops. These coefficients are given by the following formulae [59]:

$$K_r = \begin{cases} \dfrac{TEW - D_{e,i-1}}{TEW - REW}, & D_{e,i-1} > REW \\ 1, & D_{e,i-1} < REW \end{cases} \qquad (31)$$

$$K_s = \begin{cases} \dfrac{TAW - D_{e,i-1}}{TAW - RAW}, & D_{e,i-1} > RAW \\ 1, & D_{e,i-1} < RAW \end{cases} \tag{32}$$

The values TEW and REW are the total and readily evaporable water levels respectively, in mm, for different soils. TAW and RAW are total and readily available water levels each day, in mm, for a given crop-soil combination. TEW and REW values are readily available for major soil types. TAW, and RAW are given by:

$$TAW = 1000(\theta_{FC} - 0.5 * \theta_{WP}) * Z_r \tag{33}$$

$$RAW = p * TAW \tag{34}$$

The parameters θ_{FC}, and θ_{WP} are the water content of each soil at field capacity and wilting point respectively. These values are constants for each soil. Z_r is the root depth of the crop at each day. p is a crop specific constant that is used to calculate RAW from TAW. Table 3 lists θ_{FC} and θ_{WP} values for all the major soil types.

Table 3 Soil attributes

Soil type	θ_{FC}	θ_{WP}	REW	TEW
Sand	0.12	0.04	5	10
Loamy sand	0.16	0.06	6	13
Sandy loam	0.24	0.11	8	18.5
Loam	0.26	0.12	9	20
Silt loam	0.3	0.14	10	23
Silt	0.33	0.17	10	24.5
Silt clayloam	0.32	0.2	10	22
Silty clay	0.37	0.23	11	25.5
Clay	0.37	0.22	11	26

3.3 Nutrient Requirements

Crops require nitrogen and phosphorous for proper growth. The model calculates the potential nitrogen and phosphorous content of the crop for each day. The nutrient demand is then calculated by subtracting the actual content from the potential content. This nutrient demand is the amount of fertilizer required for a stress free growth. The potential content for each day is given by the formula [57]:

$$N_{pot} = W_i * (bn_1 + bn_2 * \exp(-bn_3 * HUI)) \tag{35}$$

$$P_{pot} = W_i * (bp_1 + bp_2 * \exp(-bp_3 * HUI)) \tag{36}$$

The parameters N_{pot}, and P_{pot} are the potential content for a given day. bn_1, bn_2, bn_3, bp_1, bp_2 and bp_3 are crop specific parameters that express the optimal N and P concentrations as a function of the heat unit index.

3.4 Stress Factors

Under ideal conditions the crop growth is stress free and the crop is able to achieve its maximum possible growth for each day. However, actual crop growth suffers from multiple forms of stress. Lack of sufficient water, sub-optimal temperature, and a lack of nutrients inhibit daily crop growth. This is modeled in the equations by multiplying the daily biomass with a stress factor. The new daily biomass is given by

$$\frac{dW}{dt_i} = Stress * \frac{dW}{dt_i} \tag{37}$$

Stress is a value between 0 and 1 that scales down the daily biomass to actual values. There are various kinds of stress acting on the crop. These are water stress, temperature stress, nitrogen stress and phosphorous stress. Stress is given by:

$$Stress = \min(water, temperature, nitrogen, phosphorous\ stress) \quad (38)$$

Water stress is the ratio of the available water content to the actual water necessary. It is given by:

$$Water\ Stress = \frac{RAW}{Required\ water\ content} \quad (39)$$

Temperature stress is a sinusoidal function of the daily average temperature, optimal temperature and the base temperature of the crop [58].

$$Temperature\ Stress = \sin(1.5707 * \frac{(T_{Avg} - T_{base})}{(T_{Opt} - T_{base})}) \quad (40)$$

Nutrient stress for both phosphorous and nitrogen is expressed as a function of the ratio of the actual nutrient content to the optimal nutrient content.

$$Nitrogen\ stresss = 200 * \frac{\frac{Actual\ nitrogen\ content}{Optimal\ Nitrogen\ content}}{(\frac{Actual\ nitrogen\ content}{Optimal\ Nitrogen\ content} + \exp\left(4.065 - 0.0535 * \frac{Actual\ nitrogen\ content}{Optimal\ Nitrogen\ content}\right))} \quad (41)$$

$$Phosphorous\ stresss = 200 * \frac{\frac{Actual\ phosphorous\ content}{Optimal\ phosphorous\ content}}{(\frac{Actual\ phosphorous\ content}{Optimal\ phosphorous\ content} + \exp\left(4.065 - 0.0535 * \frac{Actual\ phosphorous\ content}{Optimal\ phosphorous\ content}\right))} \quad (42)$$

3.5 Program Flow and Input Files

The computer model requires 3 input files to run the simulation. The first file is the crop parameter file. This file contains tab separated values that describe the unique values associated with a crop. Table 4 lists all the parameters in the crop input file

Table 4 Crop parameter file description

Parameter	Description
Crop Name	This parameter is used to id the crops
kcini	Initial value of the crop parameter. This value is used crop the beginning of the crop cycle till the beginning of the development cycle. Dimensionless quantity.
kcmid	The crop parameter during the middle part of crop cycle. It is also used, along with the kcini value, to estimate the crop parameter for the development stage
kclate	The crop parameter during the crop decline stage.
root	Maximum root depth of the crop in meters.
ht	Maximum crop height in meters.
p	Depletion fraction. A dimensionless quantity used to calculate the readily available water content.
rue	Radiation use efficiency. This is the potential unstressed growth rate per unit of intercepted photosynthetically active radiation. $(kg\ ha^{-1})/(MJ\ m^{-2})$
hi	Harvest index. A dimensionless quantity that represents the fraction of the total biomass that is considered as yield.
top	Optimal temperature for crop growth in °C
tbs	Minimum temperature for crop growth in °C
dmla	Maximum leaf area index is a dimensionless quantity.

frp1	First known point on the leaf area curve for a crop. This parameter signifies the % of the growing season at which the fraction of LAI is known.
frl1	The fraction of the maximum LAI associated with frp1.
Frp2	Second known point of the leaf area curve.
frl2	The fraction of the maximum LAI associated with frp2.
Bn1,bn2,bn3	Nitrogen uptake parameters at emergence, 0.5 maturity and complete maturity, respectively
bp1, bp2, bp3	Phosphorous uptake parameters at emergence, 0.5 maturity and complete maturity, respectively
tday	Total number of planting days for a crop. Used to calculate the potential heat unit index.
ini	Number of days for emergence of the crop.
dev	Duration of development stage of the crop. This period coincides with the exponential growth of biomass.
med	Duration of the middle stage of the growth cycle
late	Duration of crop decline

The second input file is the soil parameter file. This file contains the parameters to describe 9 different types of soils. In reality soil profiles are more detailed and require extra parameters to describe them. However, these parameters are area specific and their values are usually determined by on-field experiments. Table 5 lists these parameters along with their descriptions.

Table 5 Soil parameter file description

Parameter	Description
Soil Name	This parameter ids the soil type
θ_{FC}	This parameter signifies the water content of the field at field capacity. It has units of $m^3 \, m^{-3}$
θ_{WP}	This parameter is the soil water content at wilting point. It has units of $m^3 \, m^{-3}$
TEW	Total evaporable water content. It is the maximum depth of water that can be evaporated from the soil when the topsoil has been initially completely wetted. It has units of mm. This parameter can also be calculated using the formula $1000(\theta_{FC} - 0.5 * \theta_{WP}) * 0.10$. The value 0.10 signifies a depth of 10 cm. The units of this parameter are in mm.
REW	Readily evaporable water content. It is the maximum depth of water that can be evaporated from the topsoil layer without restriction. The units are in mm

The final input file that is required to run the simulation is the weather file. This file contains daily values required to calculate the potential and actual evapotranspiration for the crop. Field specific models incorporate extra attributes like stochastic models and the height of measuring instruments in the weather files. However, for the generic model the following 6 parameters have been found to be sufficient.

Table 6 Weather file description

Parameter	Description
Max Temp	Daily maximum temperature in °C
Min Temp	Daily minimum temperature in °C
Precipitation	Daily precipitation levels in mm/day
Wind	Daily wind speed in m/s
Relative humidity	A fraction that signifies the amount of water vapor in the air.
Solar	Daily solar radiation in MJ/m^2

Once the parameter files have been prepared, the model reads them and prepares the simulation. The following flowchart in Fig. 11 provides an overview of the steps performed by the model after specific crop, soil, and weather options have been prepared by the user.
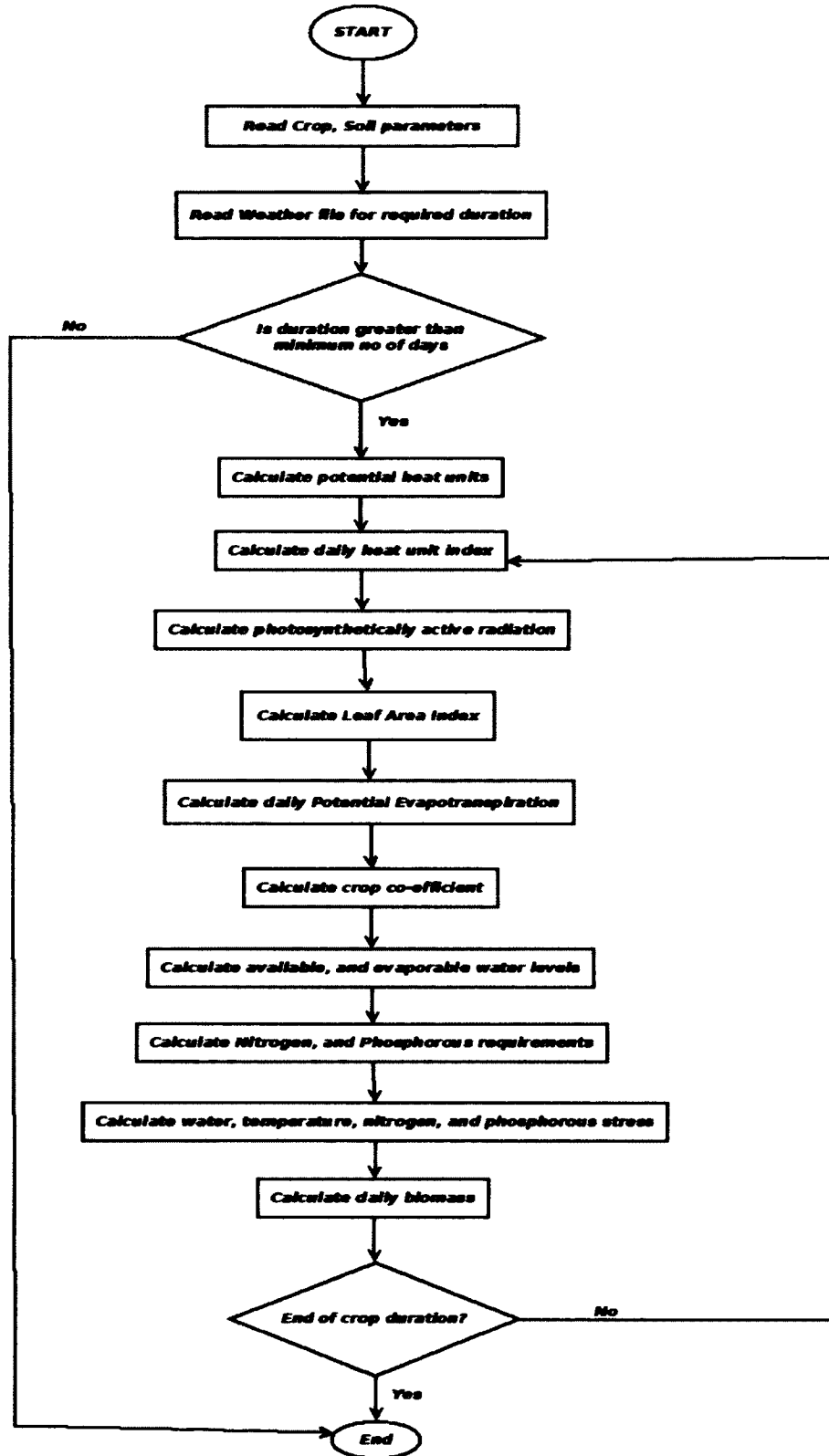
Fig. 11: Flowchart of Crop Growth program.

3.6 Model Applications

The crop growth model was implemented in MATLAB. Parameters for 25 crops are available for simulation. Weather files in the proper format can be downloaded from http://globalweather.tamu.edu/ . This allows us to simulate crop growth for different climatic regions. In this section we will look at the various applications of the model.

3.6.1 Simulating Crop Yields

The main application of the crop growth program is to simulate the yields of crops for various weather conditions. In addition to the yield at the end of the growing period, the model also tabulates the biomass growth at the end of each day. This allows us to see the progression of the crop biomass accumulation. Fig. 12 shows the biomass accumulation of various crops over a similar growing period.
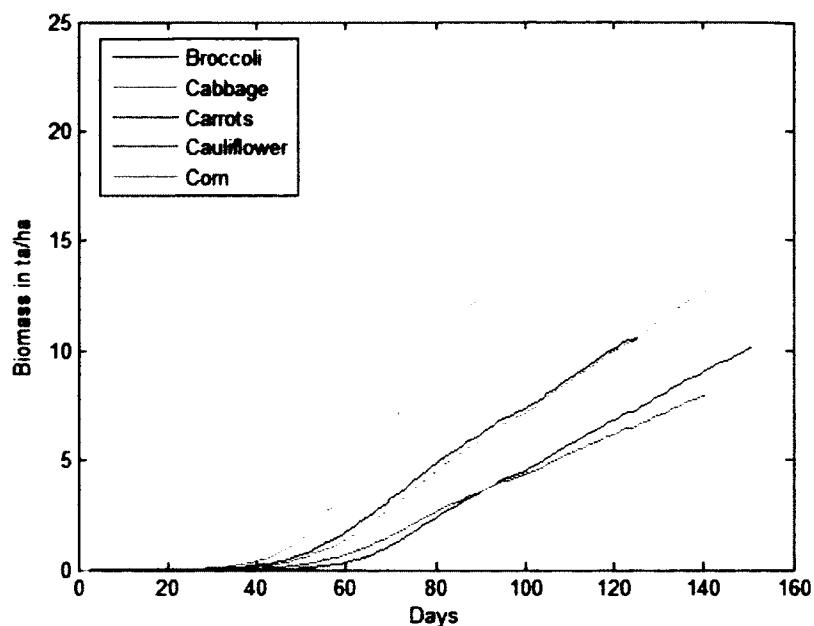
Fig. 12. Biomass growth for various crops.

The model can also be used to observe the biomass growth over multiple growing periods. For a given weather file for 2 years, Fig. 13 displays the biomass accumulation over 4 growing periods for Broccoli. Each period is equal to the total number of growing days for the crop. Period 1 is Jan-May, period 2 is May-Sept, period 3 Sept-Jan and period 4 is Jan-May.
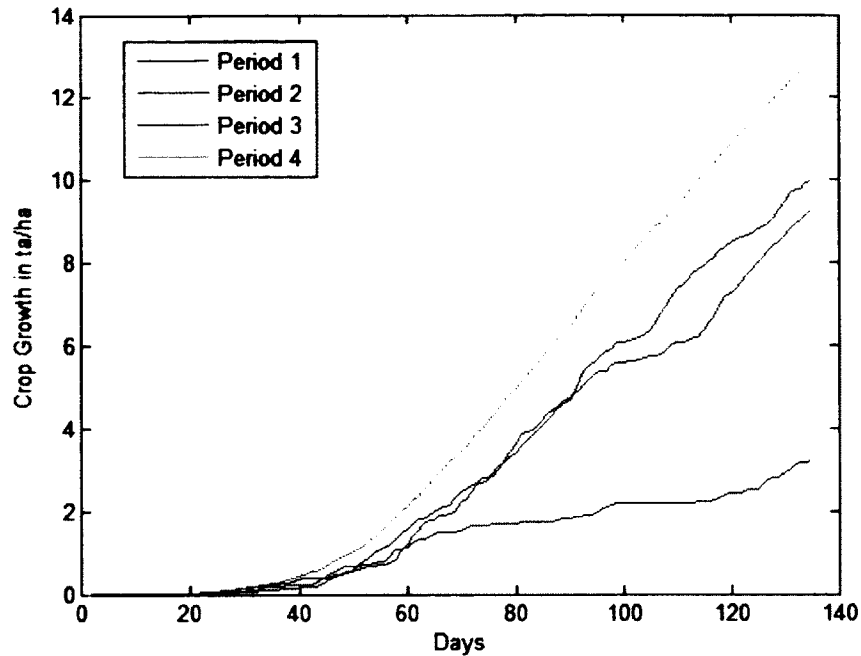
Fig. 13. Biomass growth for Broccoli over 4 different periods of growth.

Similarly, we can observe the total yield for a crop for different planting dates. Over

2 periods the model was used to simulate the total yield for all the possible planting dates.

Each day is treated as the beginning of a new cropping period and the total yield at the end

of the cycle was calculated. Fig. 14 shows the yield progression of corn for all the possible
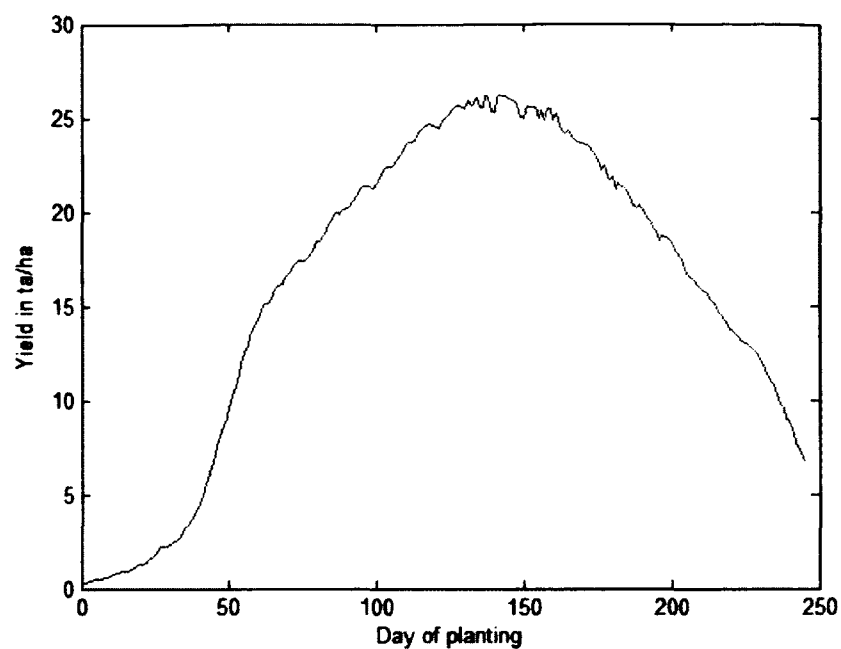
planting dates.

Fig. 14. Yield progression for Corn over all possible planting dates.

In addition to different planting dates the model can also be used to evaluate the variation in biomass for all the different soil types in the model. Fig. 15, Fig. 16 show the biomass growth for broccoli for different soil types.
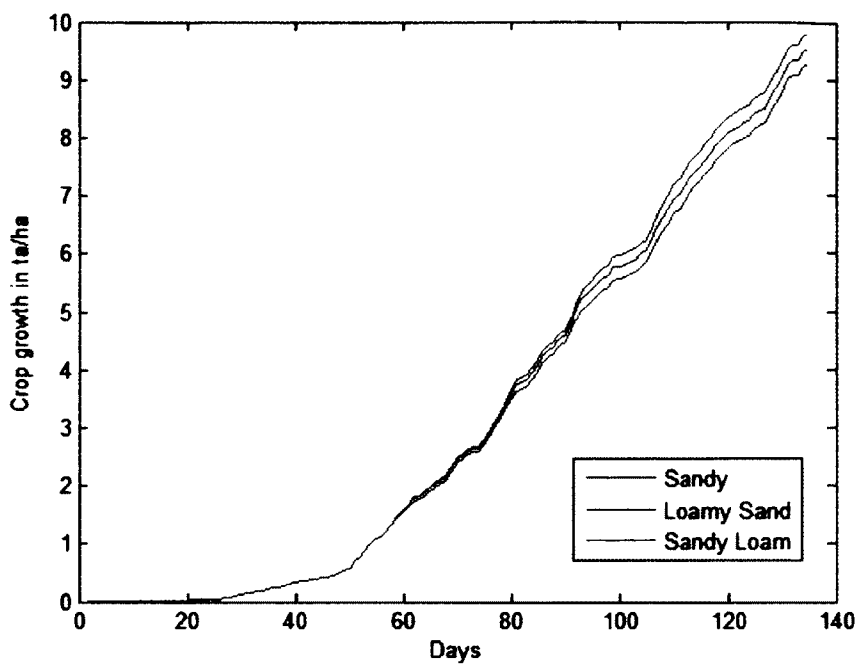
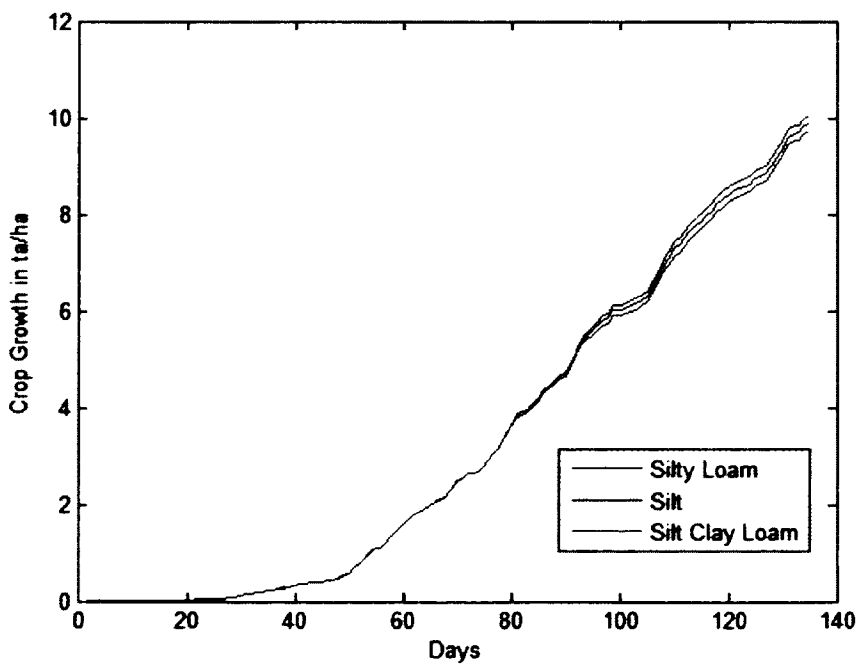Fig. 15. Biomass over 3 different soil types.

Fig. 16. Biomass over 3 different soil types.

Combining the previous two approaches, the model can be evaluated to observe the change in final yields of broccoli for a combination of all the possible planting dates, and all the available soils in Fig. 17, Fig. 18.
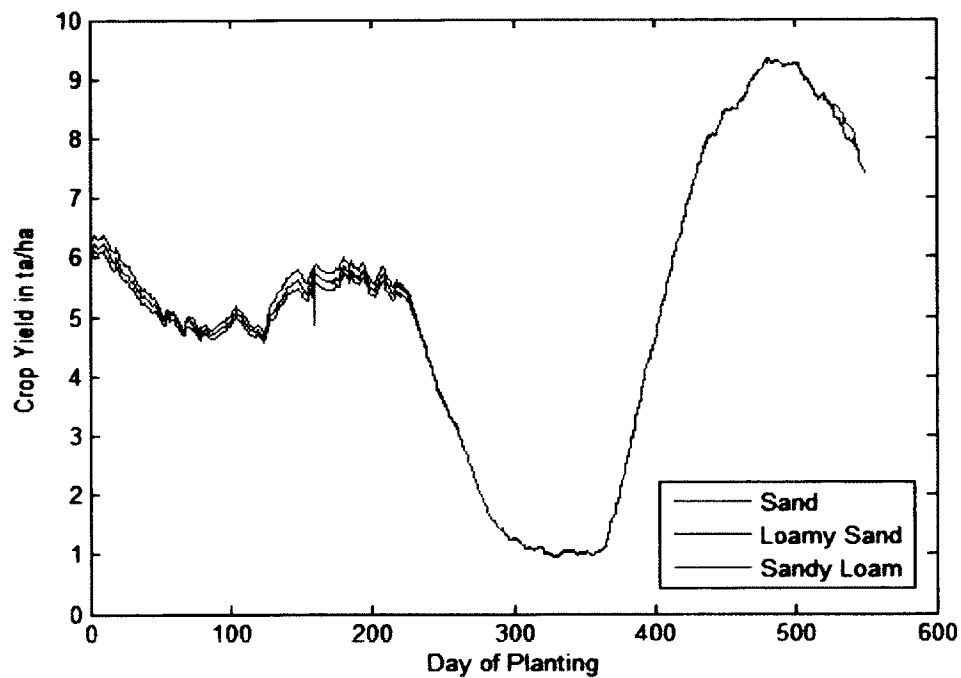

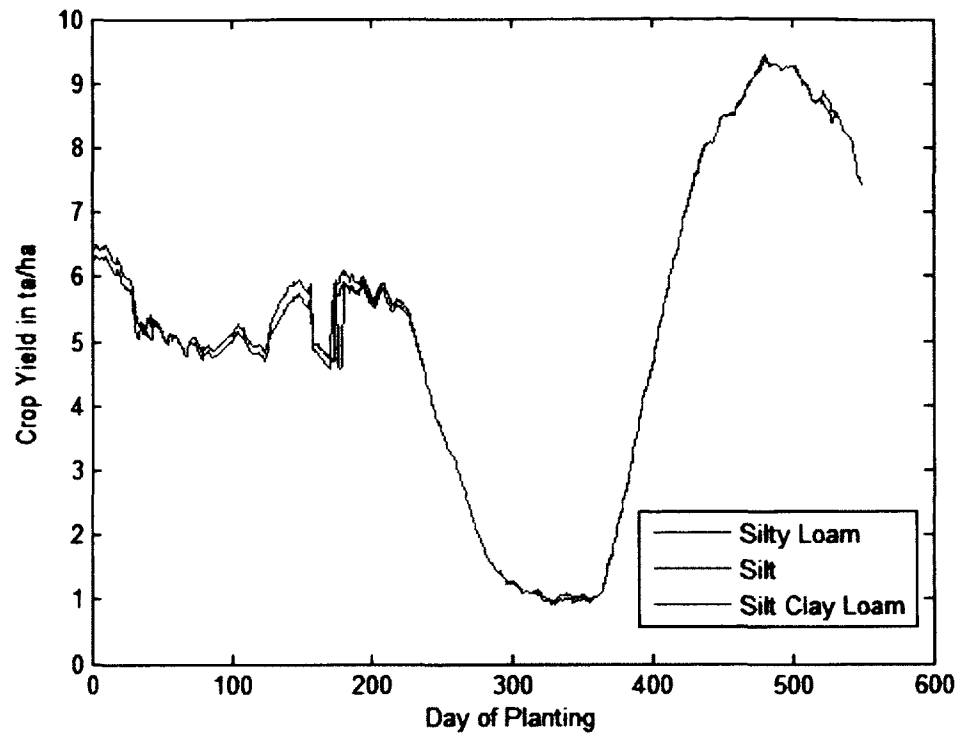Fig. 17. Yield progression over various soil types.

Fig. 18. Yield progression over various soil types.

The two graphs above represent simulated yields in a tropical climate without water stress. From the graphs we can observe that in stress free conditions the differences between the various soils are not significant.

3.6.2 Calculating Water and Nutrient Requirements

The model can be used to calculate the daily irrigation requirements for different crops. The daily water requirement is the volume of water necessary per hectare to keep the water stress value to 1. The evapotranspiration model calculates the amount of water lost by the crop each day. Fig. 19 shows the potential and actual evapotranspiration for the model over a 5 month period in a tropical climate generated by the model. Daily

precipitation is responsible for making up for this lost water. In the absence of precipitation

there needs to be water supplied through irrigation to make up for this water deficit. The

model calculates the amount of water required each day, in liters, to make sure there is a
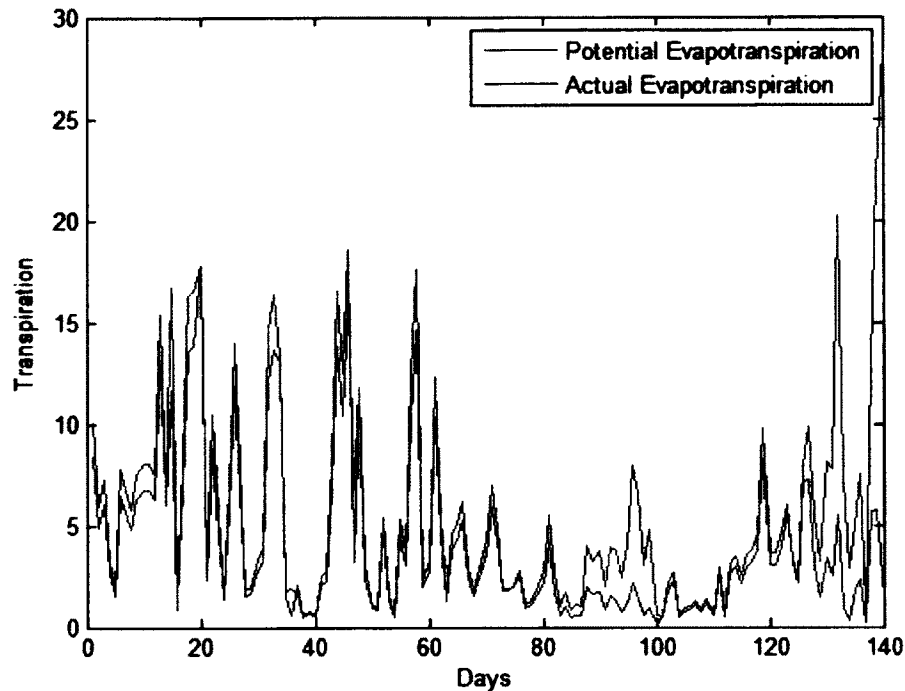
stress free growth.

Fig. 19. Potential and Actual Evapotranspiration over a 3 month period in mm/day.

Fig. 20 shows the amount of water required for different amounts of rainfall. The

model calculates the water required to make sure that there is no water stress. The graph

shows that the water required decreases as the amount of rainfall increases. A similar

analysis is done to show the water requirements for different soil types under the same

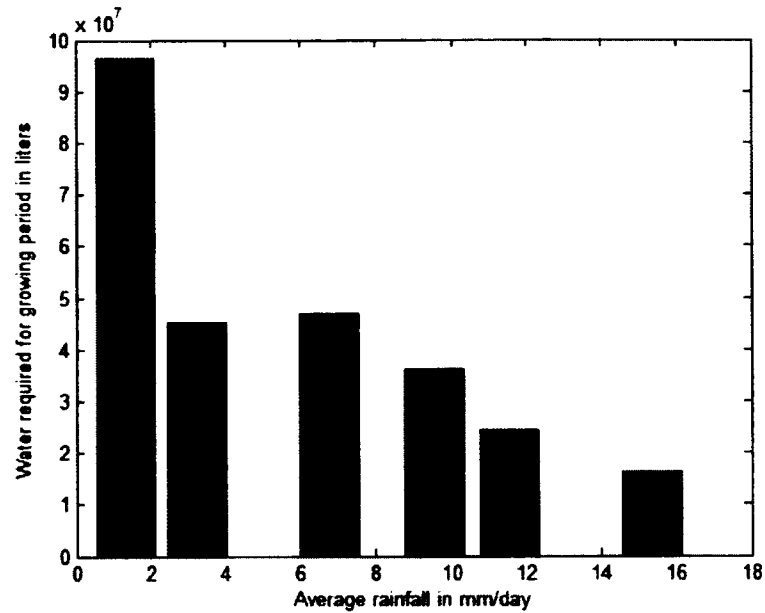weather conditions in Fig. 21.

Fig. 20. Water requirement for various average rainfalls.
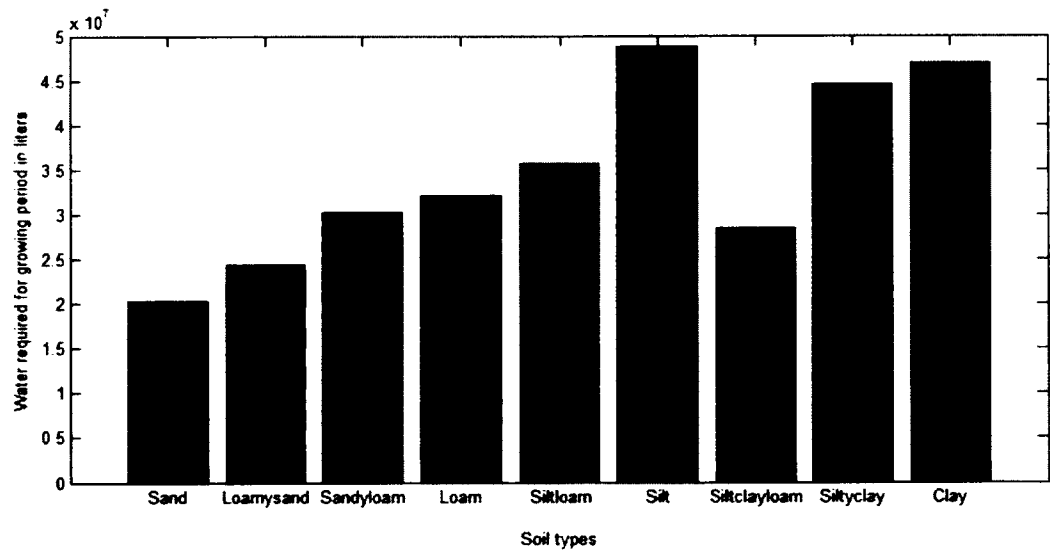

Fig. 21. Water requirement for various soil types.

Similar to water use, the model also calculates the required fertilizer, in tonnes, to make sure that there is no nitrogen and phosphorous stress on the crops. Fertilizers are

defined by their rating, which is the percentage of nitrogen and phosphorous content in the given fertilizer. For example, a fertilizer with a 35-40 rating contains 35% of nitrogen and 40% of P_2O_5. For a 100 pound bag, this would mean a nitrogen content of 35 pounds, and 40 pounds of P_2O_5. To calculate the amount of fertilizer required, we simply have to divide the amount of nitrogen, or phosphorous required by the percentage rating. If a crop requires 10 pounds of phosphorous per day, the farmer would need to apply $10/0.4 = 25$ pounds of fertilizer. However, fertilizers are not applied to crops daily. There are usually 3-4 applications of fertilizers over a growing period. This frequency can be adjusted in the model. Fig. 22 shows the fertilizer required for broccoli for, in kg/ha, for the various planting start dates.
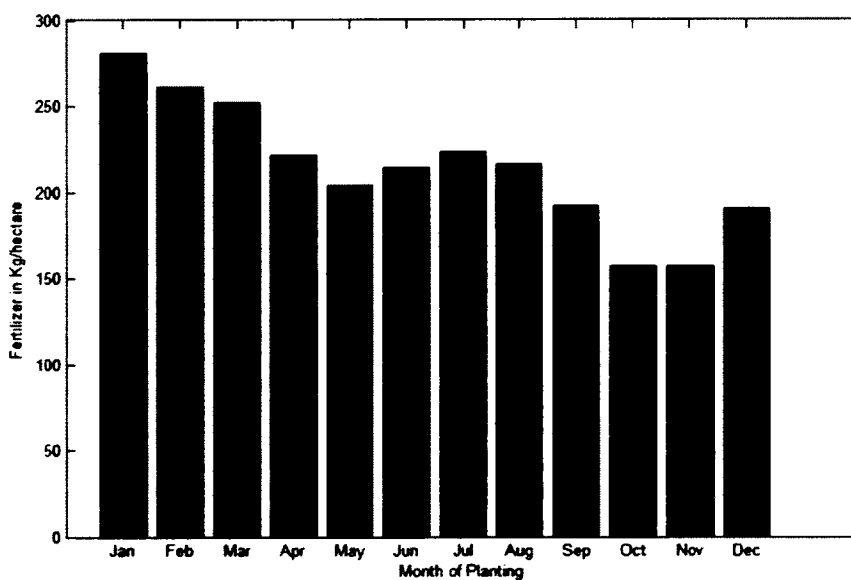

Fig. 22. Fertilizer required in Kg/hectare for various planting dates.

3.7 Sensitivity Analysis

The crop growth model has various crop specific coefficients and values. These values are unique to the crops. However, due to the variety of strains for each crop, there is variability in the values of these coefficients. In this section we will conduct a sensitivity analysis on the different crop and soil parameters. We will vary these values in increments and observe their effect on the crop yield or evapotranspiration values. If the end user decides to use a different value for the crop related coefficients, this analysis will give us a good estimate of expected variability in the output. The following variables have been used for the sensitivity analysis:

1. **Radiation Use Efficiency**: Radiation use efficiency in this model is assumed to be a constant for each different crop. In reality the value of radiation use efficiency can change due to tillage and irrigation practices. Researchers have conducted experiments on various crops using varying irrigation and tillage practices to collect empirical evidence of this variability. The following figures show the change in the yield for different changes in RUE. Fig 23, Fig. 24 shows the change in the output over the complete growth period. The second graph shoes the percentage change in the final crop yield.
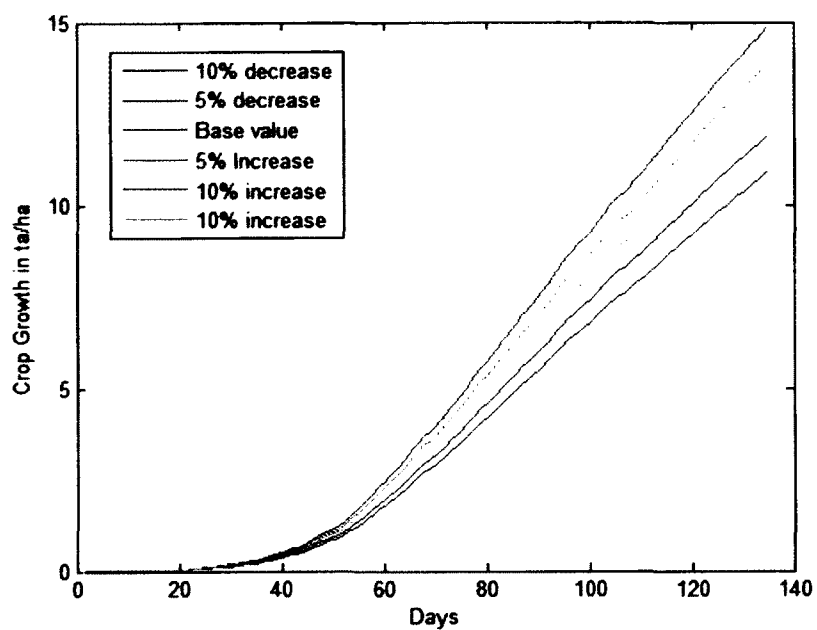
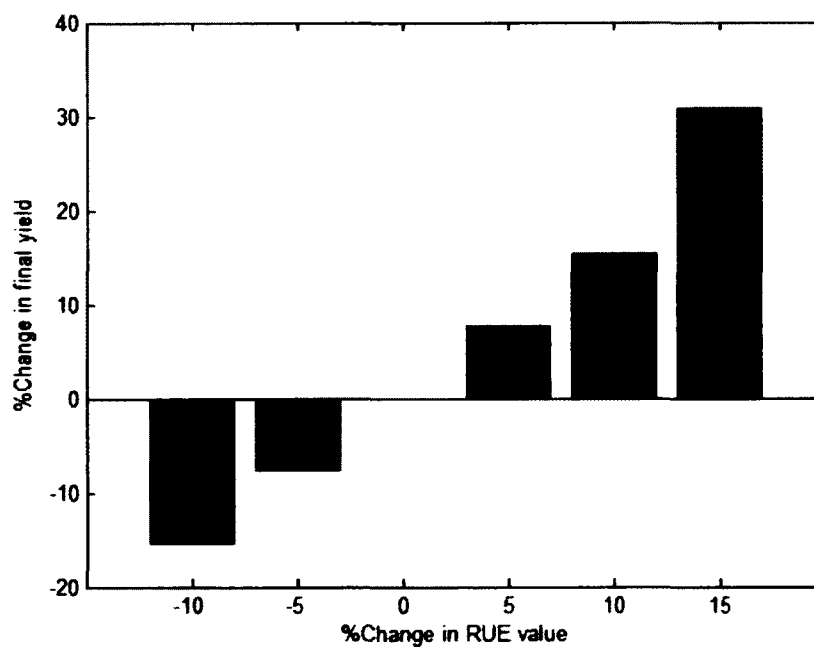Fig 23. Sensitivity of biomass to Radiation Use Efficiency.

Fig. 24. Sensitivity of yield to Radiation Use Efficiency.

2. **Crop optimal temperature**: All crops have an optimal temperature at which there is

no temperature stress acting on the crop. This optimal temperature is used to calculate

the potential heat units that determine the heat unit index for each day. In cold climates,

if the optimal temperature of a crop is low, the temperature stress on the yield is lower.

Most crop growth programs use an optimal temperature for a crop within 2°C of each

other. For a tropical weather data set, the optimal temperature was varied within 3°C of

the base value. The average yield for the growing season was observed to vary by 3-

6% as shown in Fig. 25, Fig 26. When the weather data is from a colder climate, the

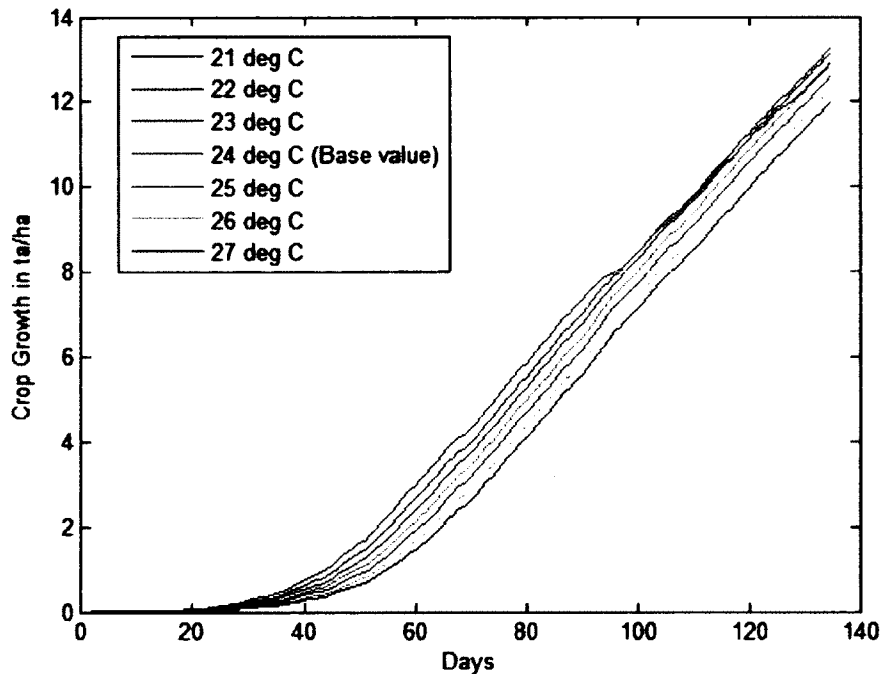fluctuation in the yield was found to be higher.

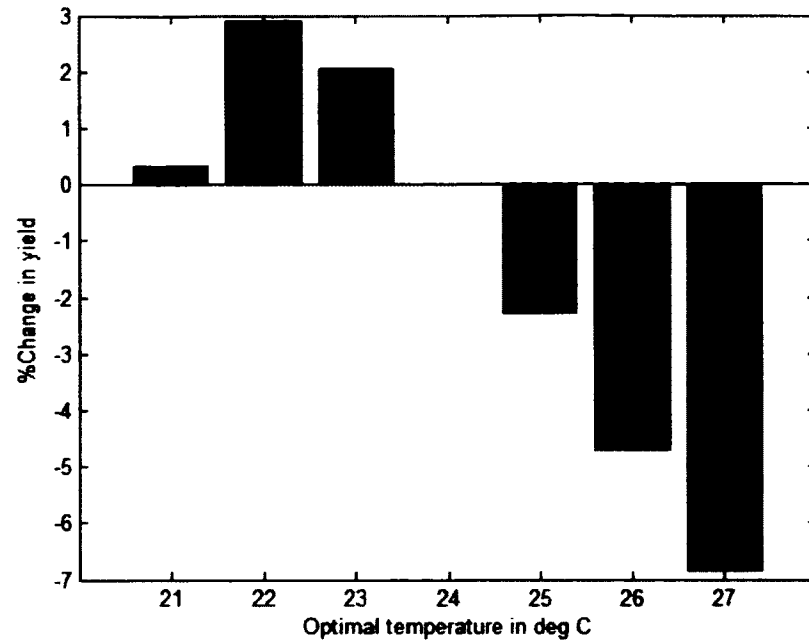Fig. 25. Sensitivity of biomass to Optimal Temperature.

Fig 26. Sensitivity of yield to Optimal Temperature.

3. **Maximum Leaf Area Index**: Leaf width or Leaf Area Index (LAI), signifies the amount of one sided area of photosynthetic tissue per unit ground surface. This index is an indication of the growth and quality of vegetation. It demonstrates the importance of the role of leaves in vegetation dynamics through photosynthesis, transpiration, rain interception, and respiration [60]. The maximum LAI usually varies by 10% between different models. In our analysis, the max LAI was varied by 5-10%, and the change in yield was observed. Lower values of LAI resulted in a decrease of 1.5-3.5% of the yield. Higher values of max LAI saw an increase of 2-5% in the yield
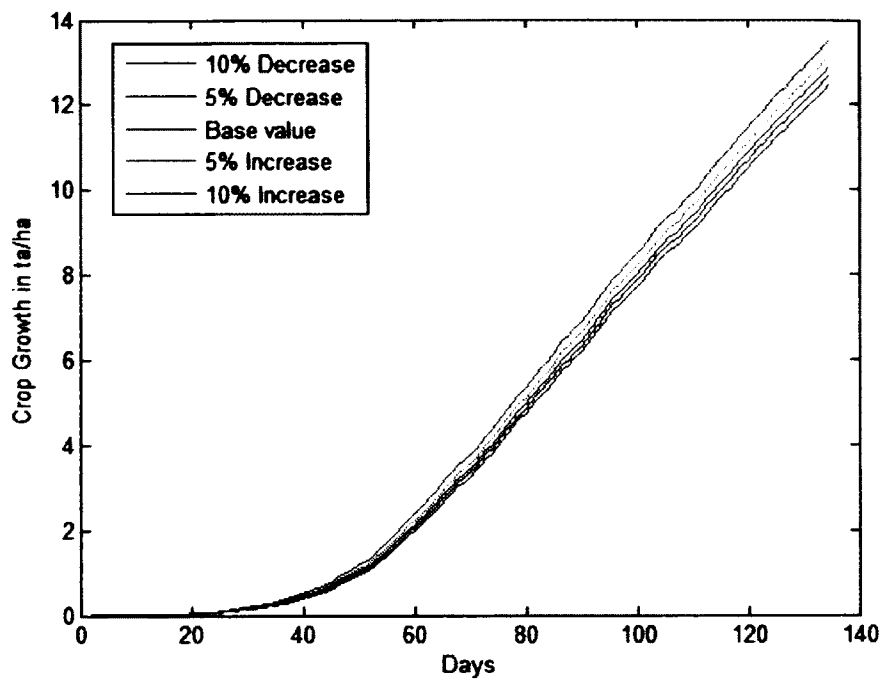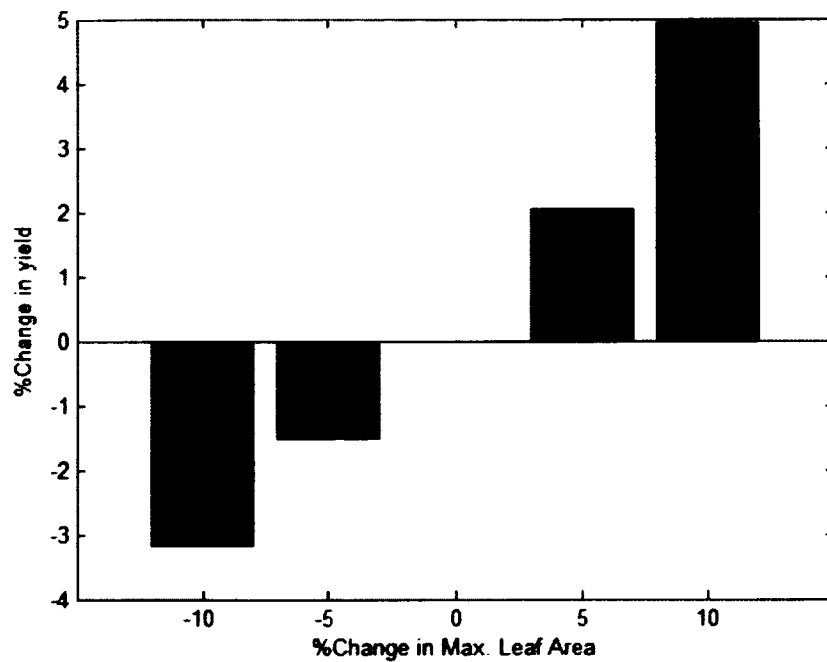
Fig. 27 Sensitivity of biomass to Max LAI.

Fig. 28. Sensitivity of yield to Max LAI.

3.8 Limitations and Conclusions

The proposed crop growth model is a generic approach to simulate crop growth. Due to advances in crop sciences there exist various strains and hybrids of crops that are suitable for growth under different circumstances. Since our model simulates the yield of common varieties of a crop, yield comparisons to various varieties might not always be accurate.

When crop models are applied to predict real world scenarios, researchers generally measure, and calibrate crop specific parameters. Soil parameters are also more extensive and calibrated accordingly. Crop rotations, tillage methods, and irrigation types also produce a lot of variation in the yields. These variations are not available in our crop growth model.

In addition to these limitations, the proposed model also has some differences in methodology when compared to established crop growth models like CropSyst [27]. In the proposed model evapotranspiration is calculated using the Penman equation as it requires fewer parameters. CropSyst uses a version of the Penman equation, called the Penman Monteith equation, to calculate evapotranspiration. In addition to this equation, the Priestly-Taylor model is also a popular equation. Since the data required for this equation is easier to obtain than the data for the Penman-Monteith equation, the CropSyst model uses the Priestly-Taylor equation whenever there is insufficient data.

Soil water evaporation and transpiration from plant organs are two ways in which water content is lost by the crops. Additionally, surface water runoff is another method by which the water content available to plants can be depleted. It is caused by the inability of

the soil to retain the water. The CropSyst model addresses this type of water loss by expressing the daily runoff as a function of the difference between, the daily water availability, and the surface retention factor. The surface retention factor is a dynamic parameter that is updated after every time step by calculating the water retained by the plant on the previous time step. The proposed model assumes that the crops are being grown on flat lands. The water loss due to runoff is a significant phenomenon only if the land has a sufficiently high slope. Hence, in this model the water loss due to surface runoff has been assumed to be zero.

In our proposed model the rooting depth is calculated using the Monte Carlo integration. The rooting depth for each day is calculated as the product of the root depth on the previous day, a constant elongation rate, and the time step used for integration. This time step is usually one day. This method continues until the root depth reaches its maximum value or until the simulation is stopped; whichever happens first. It is also assumed that there is no root death. The CropSyst model calculates root death as a function of the maximum root depth, the maximum LAI, and the root depth of the previous day. The model performs this calculation until the root depth reaches its maximum value, or until the simulation is stopped; whichever happens first.

The equations used in the prediction of plant biomass have already been discussed in the beginning of this section. The plant biomass was expressed as a function of solar radiation, leaf area index and the plant biomass. In the CropSyst model, the biomass growth is broken down into two stages; radiation dependent growth and water dependent growth. The radiation dependent growth is similar in methodology to the equations described by us in the previous section. The water dependent growth is expressed as a function of the actual

transpiration and the biomass transpiration coefficient. This coefficient signifies the amount of biomass produced per meter of transpiration.

The crop model was used to simulate the generic yield of various popular crops. However, we do not present a validation analysis for the yields of these different crops. The main reason for this is that in real world scenarios crops come in different varieties. They are usually genetically optimized to perform efficiently in a specific environment. In addition to this, there are a large number of enhancements available to the farmers in terms of fertilizers and irrigation. All these factors have made the average yield of crops a significantly large interval [61]. This makes yield validity tests seem meaningless since every simulated output in our model tends to lie somewhere in this rather large interval. A more accurate validity analysis would involve a concentrated study in a specific agricultural region where the variance in the average yield is small.

The proposed crop growth model was developed to understand the change in crop yields for various weather and soil conditions. It allows the user to calculate potential water and fertilizer requirements for different crops under various weather conditions. The model should not be used to validate crop yields since it does not include all the variables required for an accurate estimate of yield. Future versions of this model can be modified to estimate the effects on yield from crop rotations, companion planting, irrigation schedules, and tillage practices.

CHAPTER 4

CROP SELECTION AND SCHEDULING

In the previous chapter we developed a crop growth model that is able to produce yields, water requirements, and nutrient requirements for selected crops. In this chapter we will combine the crop growth model with a mathematical programming model to select the optimal crops and their schedules to maximize various objectives. We will start by simulating the schedule that selects the best performing crops. Two sets of crop rotations will be used; one with crop rotation and another without crop rotation. The second objective will be to maximize economic output of the crops. Real world data from the United States Department of Agriculture [62] will be used to calculate the best schedule. The third objective will be minimizing the environmental impact of the crops. The water and fertilizer requirements of the crops will be used as indicators. Finally a multi objective model will be used to maximize economic output while decreasing environmental impact.

4.1 Crop Scheduling for Maximizing Performance

The objective of this model is to select the list of crops and their planting dates to maximize the performance. Since each crop has a yield on a different scale, we will need to use a more normalized measure to measure the yield. For example, a yield of 10 tonnes/ha might be a poor return for a potato crop, while a yield of 5 tonnes/ha might be very good for an eggplant crop. This problem can be overcome by first simulating the crop yields for all the possible planting dates. The yield at each planting date is then divided by

the maximum yield for that crop. This serves to provide an accurate measure of performance for the crop by showing its proximity to the maximum potential yield.

The only constraint on the model is to ensure that at any given time more than one crop cannot be planted. The model is also setup to ensure that there is a 1 month fallow period after planting each crop. There is no crop rotation constraint, and the same crop can be planted in succession. The model is written as a binary integer program, where each decision variable can only take the values 0 or 1. The yields have been calculated for the beginning of each month during a 4 year simulation cycle. The list of variables is as follows:

- Y_{tj} : Normalized yield of crop j when planted on planting date i.

- x_{tj} : Decision variable for crop j on planting date i.

- j is the total number of crops

- t is the total time in months

- totaltime : Length of the simulation in months.

- numcrops : Total number of crops available for simulation.

- $cropduration_j$: Duration of crop cycle for crop j

$$maximize: \sum_{j=1}^{number\ of\ crops} \sum_{t=1}^{totaltime} Y_{tj} x_{tj} \tag{43}$$

$Constraint:$ \hfill (44)

$$\sum_{t=1}^{total\ time} \sum_{j=1}^{number\ of\ crops} \left(x_{tj} + \sum_{k=1}^{numcrops} \sum_{l=0}^{cropduration_j} x_{(t+l+1)k} \right) \leq 1$$

The model was simulated for 48 months using weather data from [63]. The weather data was for a selected tropical region with average temperatures of around 25^0C, and an

average rainfall of 13cm. The model was solved using the bintprog solver form MATLAB. The model can currently optimize 28 crops. To visualize the results of the simulation, we will use a Gantt chart. A Gantt chart is a form of bar graph that is used to visualize project schedules. This chart is useful to display the start date and the duration of the crops. An initial run for all the 28 crops was made to determine the most optimal sequence of crops. Fig. 29 displays the results from this simulation. The figure shows that broccoli, potatoes, barley, oats and millet produced the most optimal yield.

In practical applications farmers tend to select between a limited set of crops rather than all the possible crops that can be planted. to demonstrate the model behavior the simulation is executed again, by selecting sets of three crops, to observe the best sequence of plantings. The first set of crops is sugar beets, lettuce and carrots. The second set is peas, sweet peppers and eggplant. Fig. 30 and Fig. 31display the yields of the respective sets for a period of 48 months.
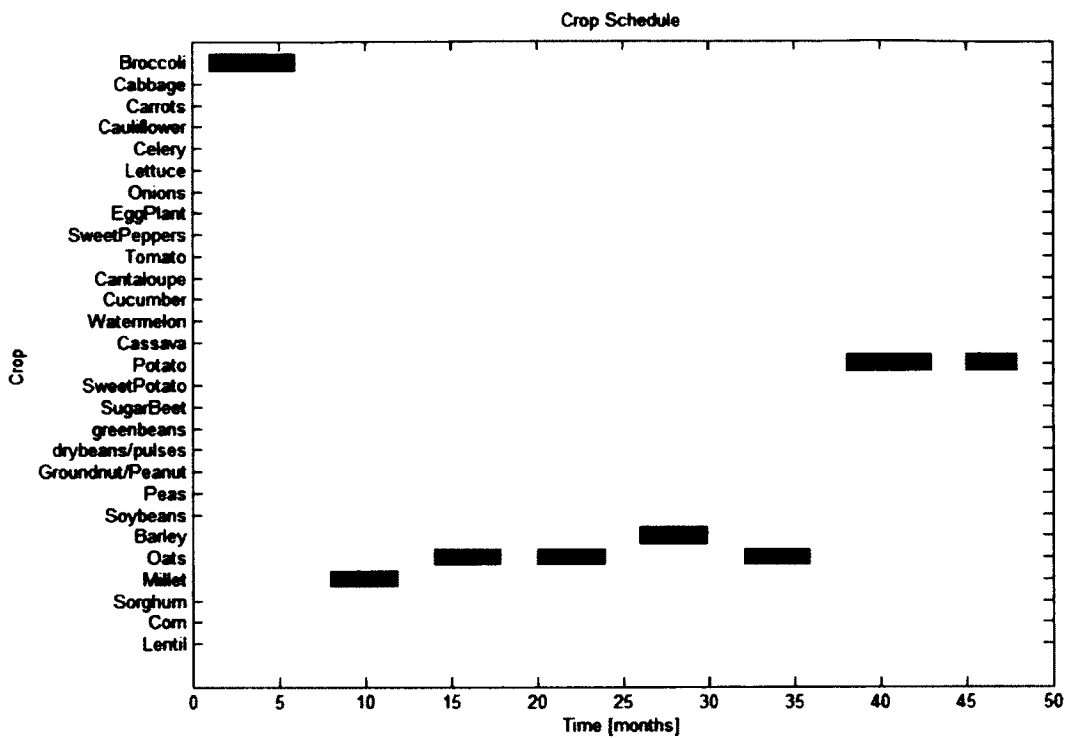
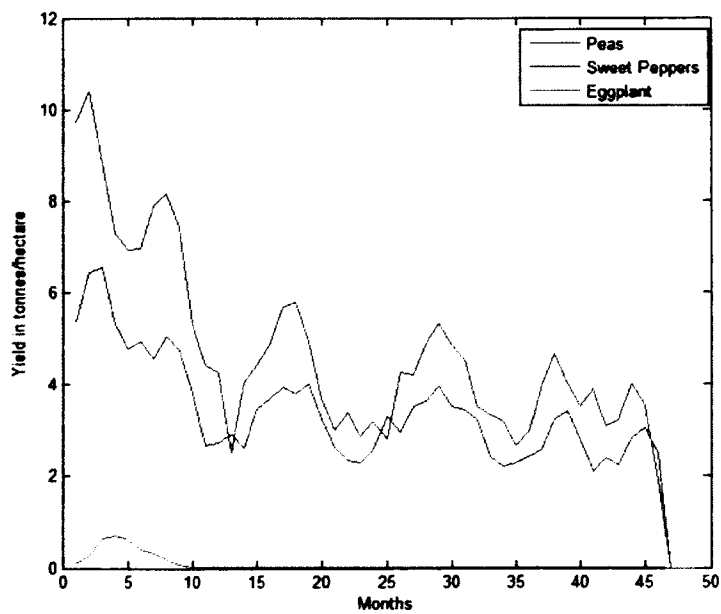Fig. 29. Crop Scheduling with now rotation for full set of crops.

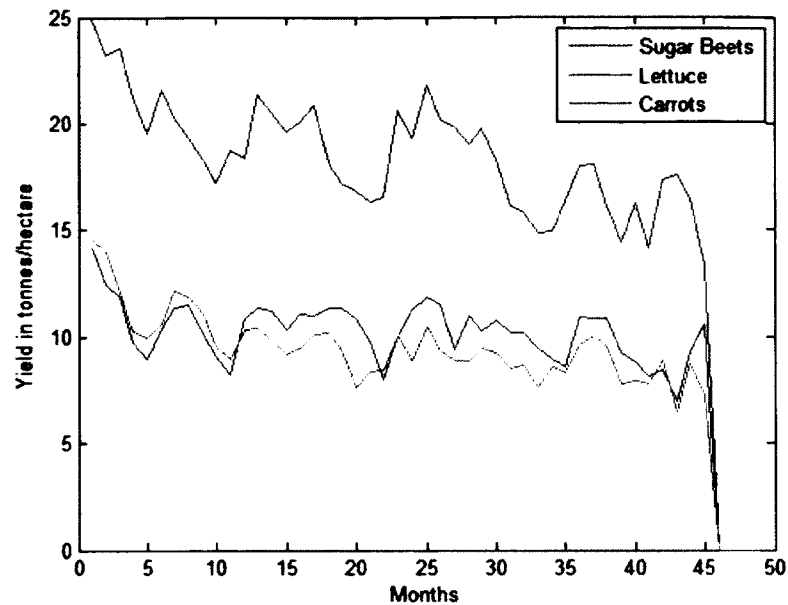Fig. 30. Yield/hectare of first set of crops.

Fig. 31. Yield/hectare of second set of crops.

The two sets of crops are used in the simulation to calculate the schedule for their cropping. Fig. 32, Fig. 33 show the results. For the first set of crops the crops schedule is a mix of all three crops even though there is no crop rotation in enabled. For the second set of crops the schedule is a mix of peas and sweet peppers while eggplant is never selected. This is due to the extremely low yield of eggplant.
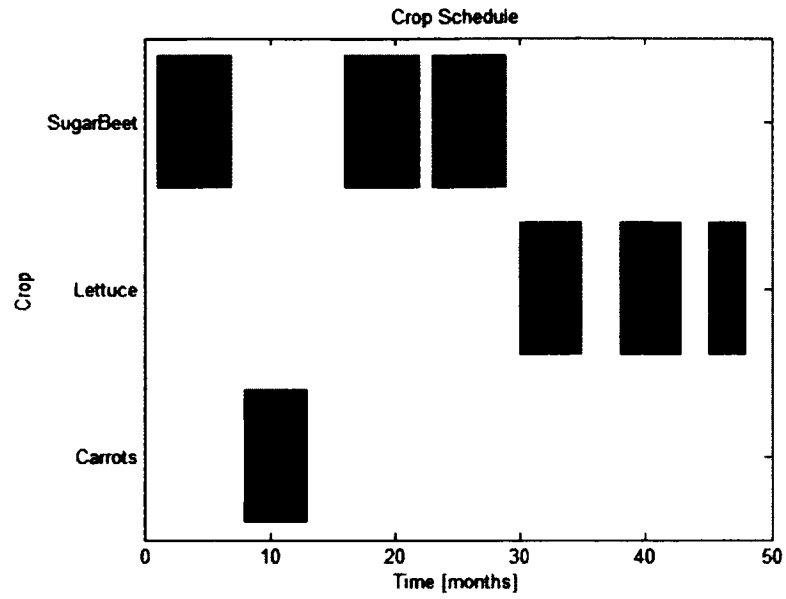
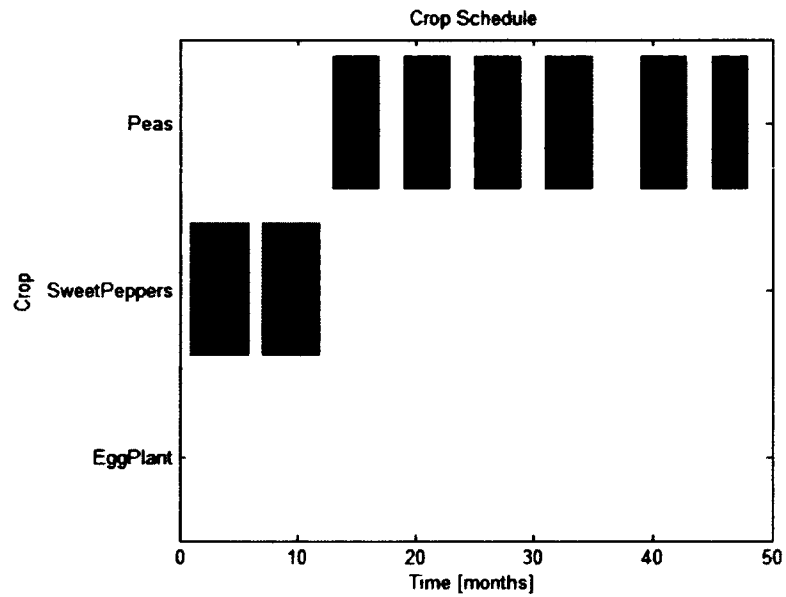Fig. 32. Crop scheduling for set of crops without crop rotation.

Fig. 33. Crop scheduling for second set of crops without crop rotation.

This model can now be extended to account for crop rotation. Crop rotation is implemented by specifying that the same crop not be re-planted for a minimum of three months. If the other crops do not fit the solution, the same crop can be planted again after waiting for 3 months. The rotation period of 3 months has been used for demonstration purposes and the actual length can be varied. The model does not take into account the economic/environmental benefit of leaving the land fallow between crop plantings. A new variable called *rotation* is added to the model. This variable specifies the duration for which a crop cannot be reused on the same field. This value can take a value between 1-3 months. Using the same notation as above, the model can be written as:

$$\textit{maximize:} \quad \sum_{j=1}^{number\ of\ crops} \sum_{t=1}^{totaltime} Y_{tj} x_{tj} \tag{45}$$

Constraint: (46)

$$\sum_{t=1}^{total\ time} \sum_{j=1}^{number\ of\ crops} \left(\sum_{m=t}^{t+cropduration_j+rotation} x_{mj} + \sum_{k=1}^{numcrops} \sum_{l=0}^{cropduration_j} x_{(t+l)k} \right) \le 1$$

The model is simulated using the same conditions and time period as before. From the Fig. 34 we can see that the best sequence is to plant mostly lettuce and one iteration each of sugar beet and carrot. Fig. 35 shows the simulation for peas, sweet peppers and eggplant. The model suggests that planting peas most of the time with one planting of sweet peppers has the highest yield. The reason for the selection of peas is the relatively low crop growth duration of 3 months. The remaining two crops have a 5 month growing period
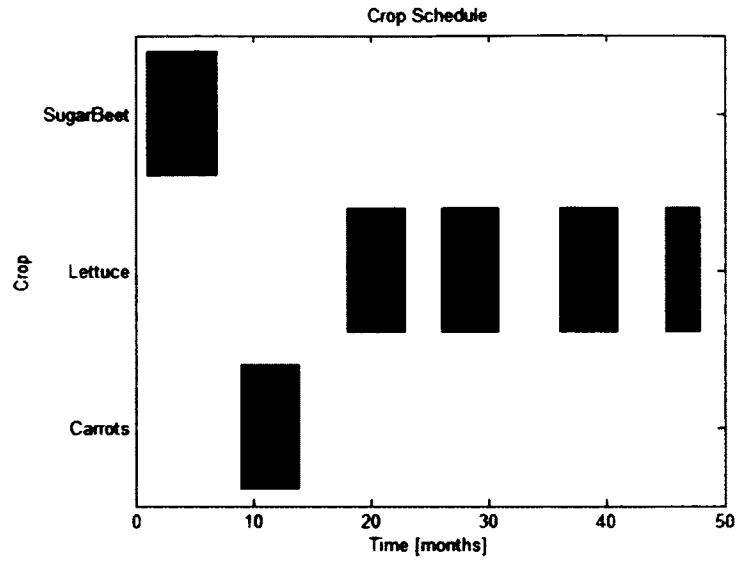
Fig. 34. Crop selection with explicit crop rotation for first set of crops.
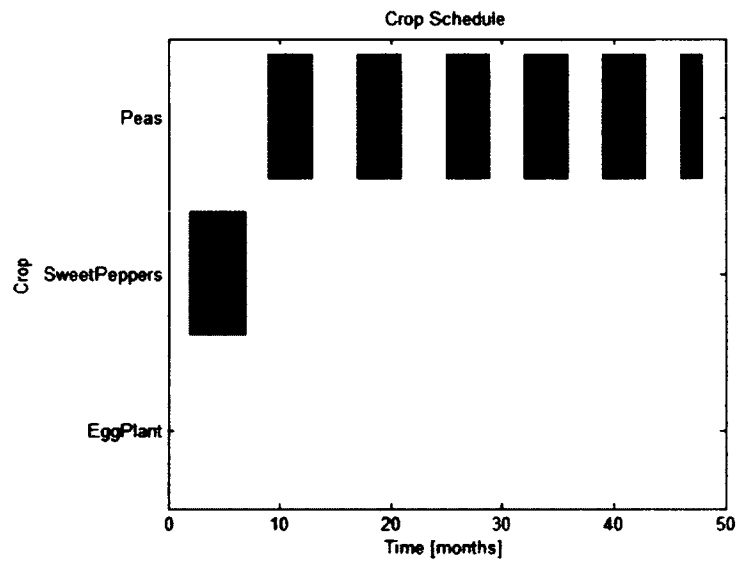

Fig. 35. Crop selection with explicit rotation.

4.2 Crop Scheduling for Maximizing Economic Profit

We will now extend the existing models to include the market prices of the crops. The objective function is now a profit maximizing function. The 48 month data for the

different vegetable prices has been obtained from the USDA website [62]. Fig. 36, Fig. 37

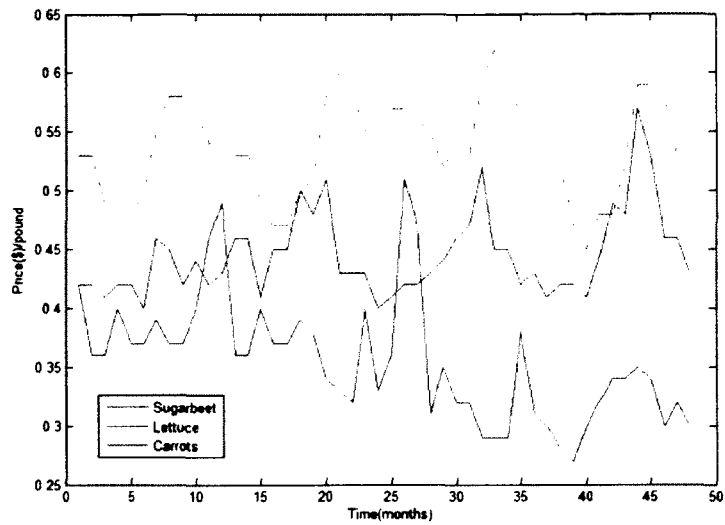provide the price/pound trends for the vegetables we will use in our examples.

Fig. 36. Price($)/pound for first set of crops obtained from USDA.

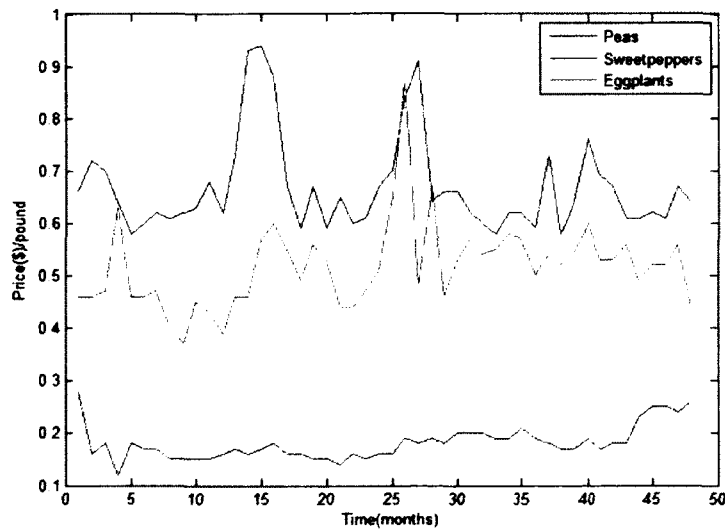
Fig. 37. Price($)/pound for first set of crops obtained from USDA.

To develop the necessary model, we only need to change the objective functions by including the market prices, C_{tj}, of the crop j for every month t. The constraints remain the same as before. The objective function for models without rotation and with rotation is the same. The objective function is given by equation:

$$maximize: \sum_{j=1}^{number\ of\ crops} \sum_{t=1}^{totaltime} C_{tj}Y_{tj}x_{tj} \qquad (47)$$

Fig. 38, Fig. 39 show us the results of the model when there is no rotation. We observe that only sugar beet and sweet pepper have been planted. From the price/pound graphs we know that these two crops had the highest price.
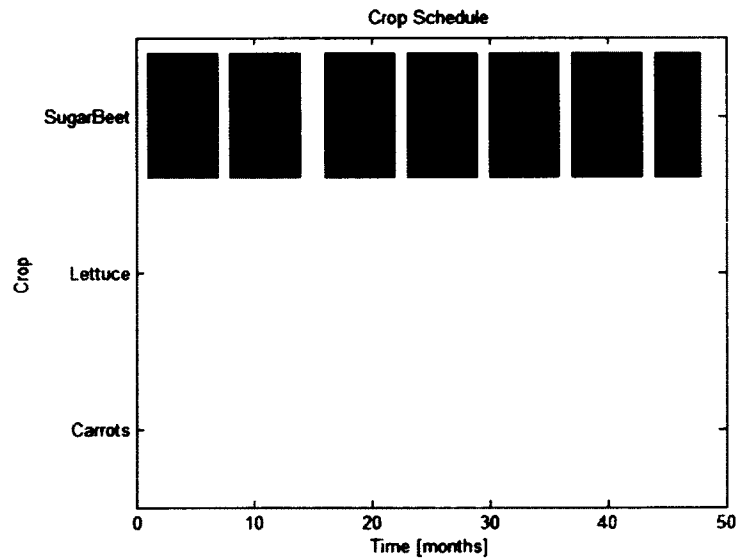

Fig. 38. Crop schedule when maximizing economic objective and no rotation for first set of crops.
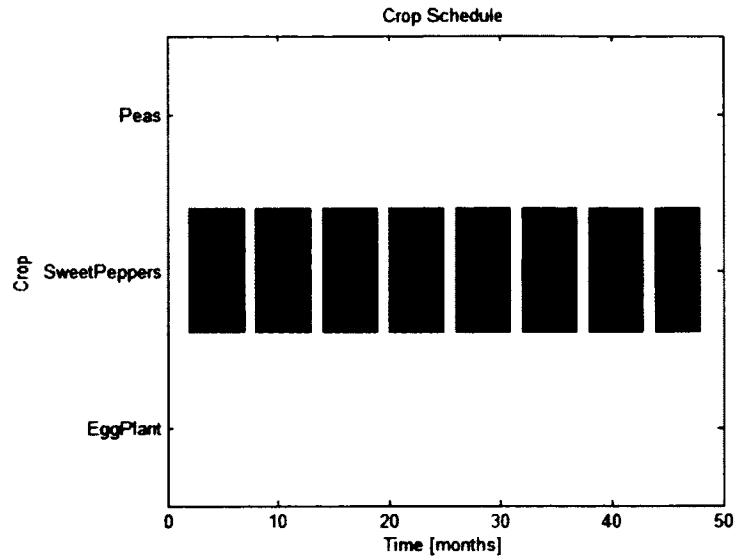
Fig. 39. Crop schedule when maximizing economic objective and no rotation for second

set of crops.

The simulation is repeated with explicit crop rotations. A crop cannot be planted for 3 months after it has been planted once. The results of this simulation for two different sets of crops are available in Fig. 40, Fig. 41. For the first set of crops, we can observe that only sugar beets and carrots have been selected. The price/pound graph for this set of crops shows us that the price of lettuce is quite low when compared to the other two crops. The model result also follows the same trend. In the second set of crops we observe that the model solution suggests sweet peppers for all the cycles. Peas and eggplant are never selected despite eggplant having a high price/pound value. The reason for this is that the yield price product of sweet peppers is higher than that of the other two crops.
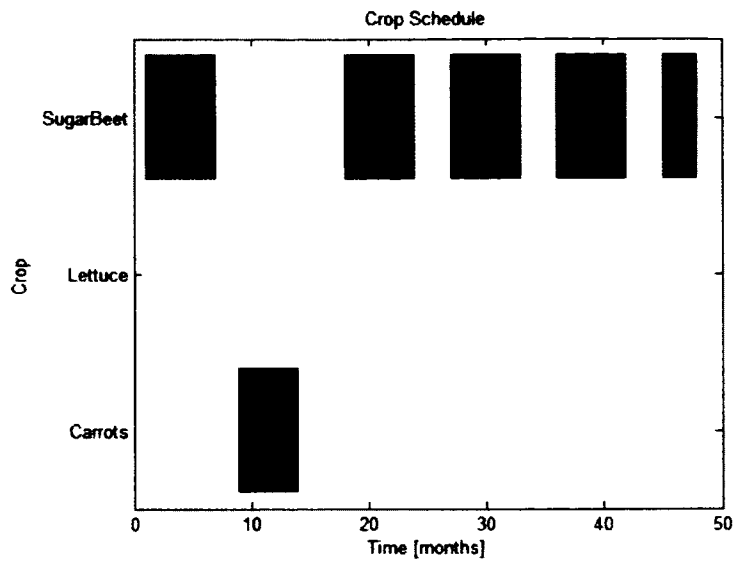
Fig. 40. Maximizing economic objective with explicit crop rotation for first set of crops.
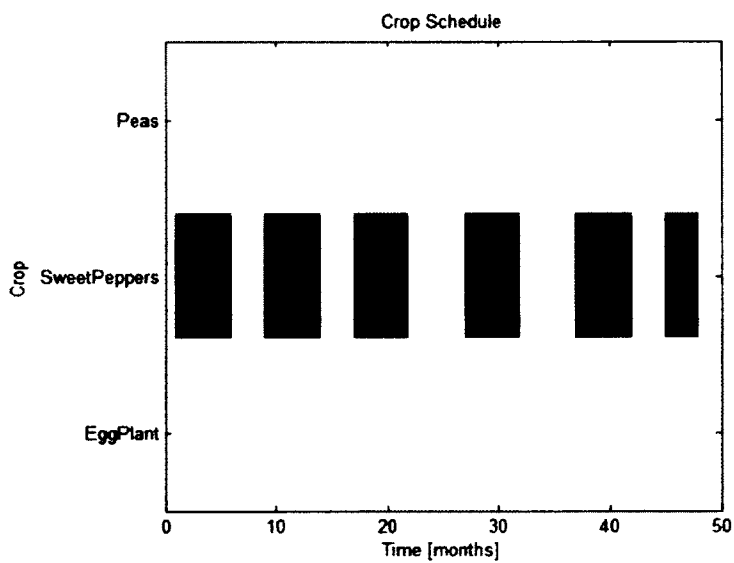
Fig. 41. Maximizing economic objective with explicit crop rotation for second set of

crops.

4.3 Crop Scheduling for Minimizing Environment Impact

In this section we will evaluate different models to study crop scheduling by minimizing the environmental impact of the crops being planted. The crop growth model discussed in the previous chapter is capable of producing the amount of water required by a crop to undergo stress free growth. We will use these values to determine the proper crop scheduling to plant crops that use the least amount of water. The water usage in liters for two sets of crops, for different planting dates, is given in Fig. 42, Fig. 43.

Fig. 42. Water usage in liters for the first set of crops for the different planting dates.
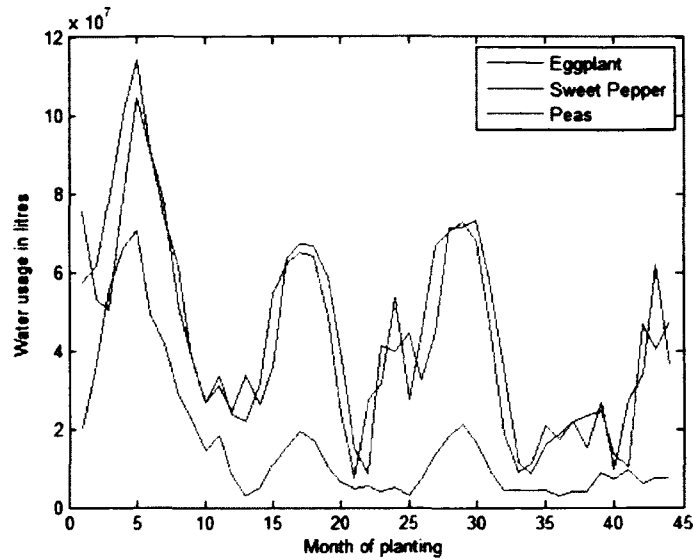
Fig. 43. Water usage in liters for the second set of crops for the different planting dates.

The models for minimizing environmental impact have the same constraints as the previous models. The objective functions are changed to include the water usage, W_{ij}, of each crop for different planting dates. The values for the water usage were obtained from the output of the Crop Growth Model discussed in the previous chapter. The objective function required for the model is given below:

$$minimize \sum_{t=1}^{totaltime} W_{tj}x_{tj} \qquad \forall j = 1,2, \dots.number\ of\ crops \qquad (48)$$

The model is first evaluated to determine the crop scheduling when there is no explicit crop rotation. The results for this are in the figure below. For the first set of crops, we observe that lettuce has been selected the most number of times and in the second set of crops only eggplant has been selected. From the water use graphs we know that these crops had the lowest water consumptions in their respective sets and the model results follow a similar trend.

It is also interesting to note that both these crops were among the least selected when the objective was increasing potential yield or economic impact.

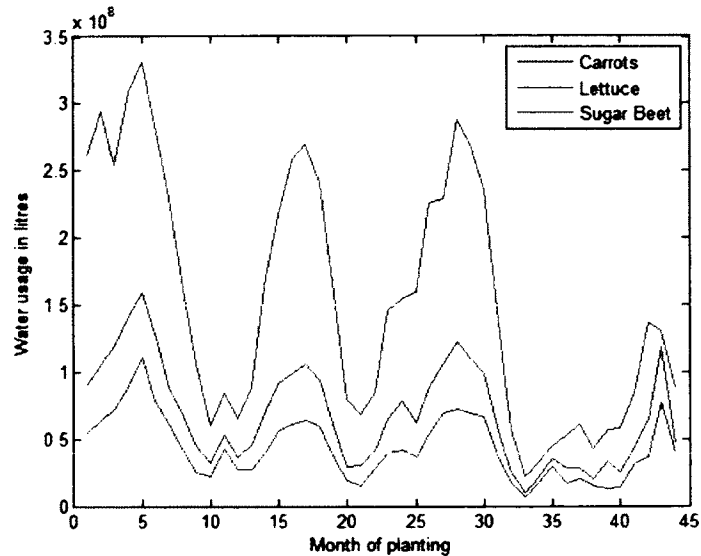

Fig. 44. Minimizing environmental impact with no crop rotation for first set of crops.
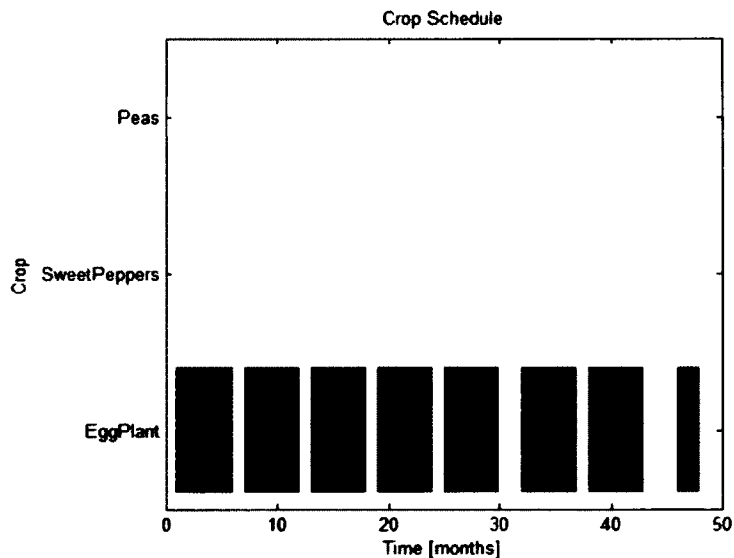
Fig. 45. Minimizing environmental impact with no crop rotation for second set of crops.

The model is re-evaluated to include crop rotation constrains. From the results for the two sets of crops we can see that lettuce and eggplant are still the preferred choice for planting as seen in Fig. 46, Fig. 47. It is important to note that the economic output of these selections is significantly low and would not be a helpful suggestion to the farmer.
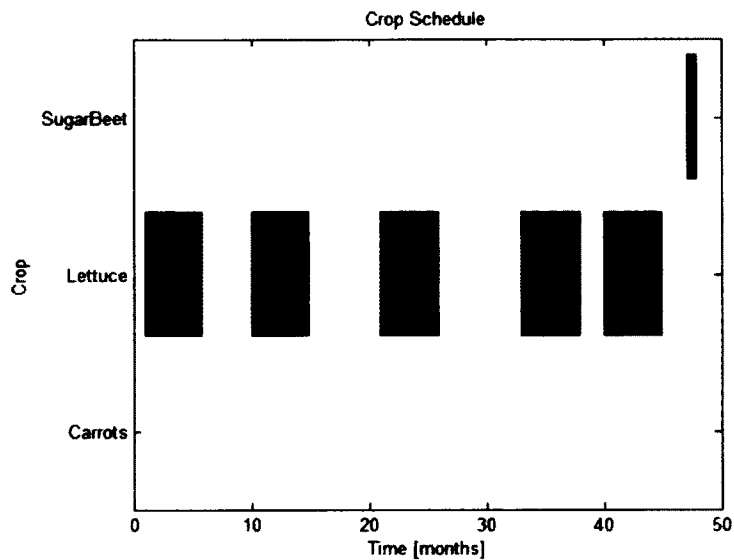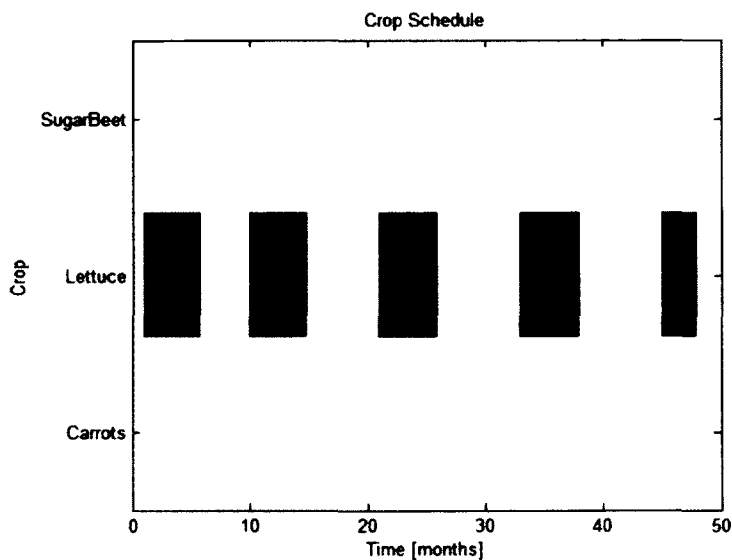

Fig. 46. Minimizing environmental impact with explicit crop rotation for first set of crops.
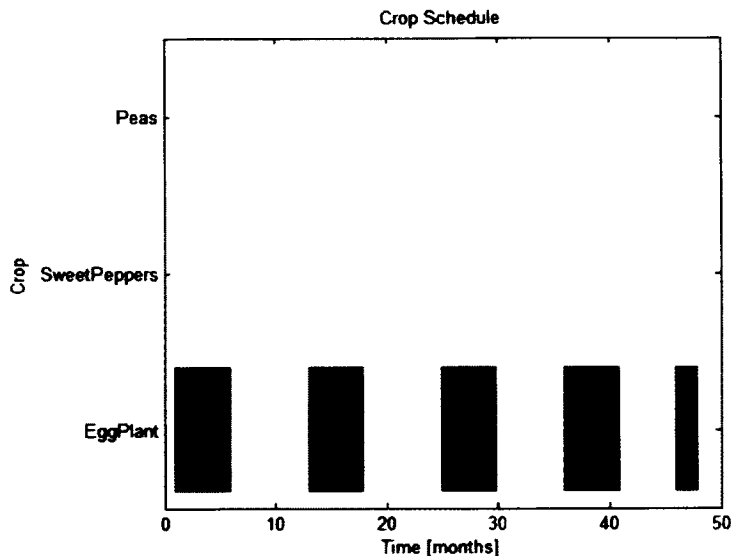
Fig. 47. Minimizing environmental impact with explicit crop rotation for second set of

crops.

The second method of assessing environmental impact is through the use of

fertilizer. The crop growth model has been used to calculate the fertilizer requirement for

the two sets of crops. Fig. 48, Fig. 49display the fertilizer in kg/hectare required for the

complete duration of the growing period. These values will be used to select the crop

schedule that minimizes the environment impact. For the first set of crops carrot has the

least fertilizer requirements for the given conditions. Eggplant has the least fertilizer

requirement for the second set of crops. These two crops also had the least water

requirements.
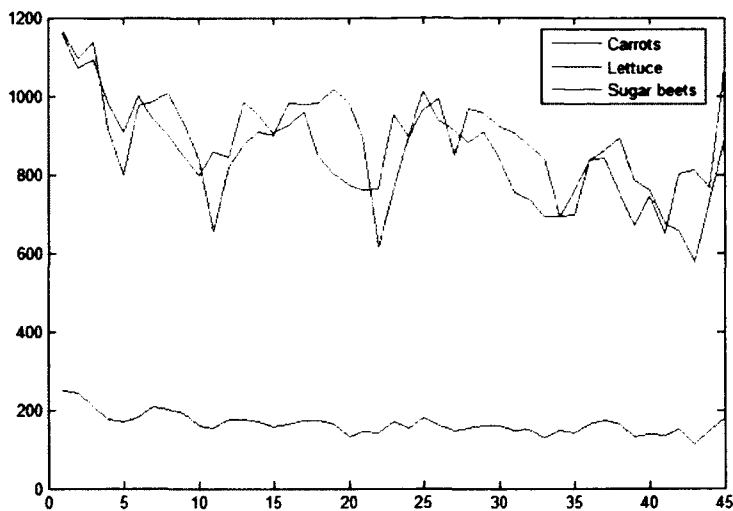
Fig. 48. Fertilizer requirements for the first set of crops.
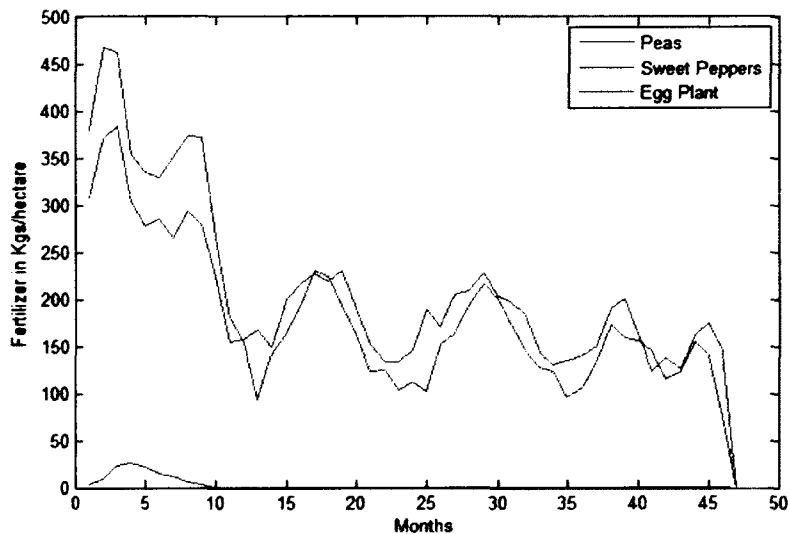
Fig. 49. Fertilizer requirements for the second set of crops.

The model has been run to minimize the environmental impact objective. A crop rotation period of 3 months has been used in the simulation. Fig. 50, Fig. 51show the schedule for the two sets of crops. Carrots and eggplants have the least amount of fertilizer

use and these crops have been picked the most in the schedule. Similar to the water use

objective these schedules produce significantly low yield and economic outputs.
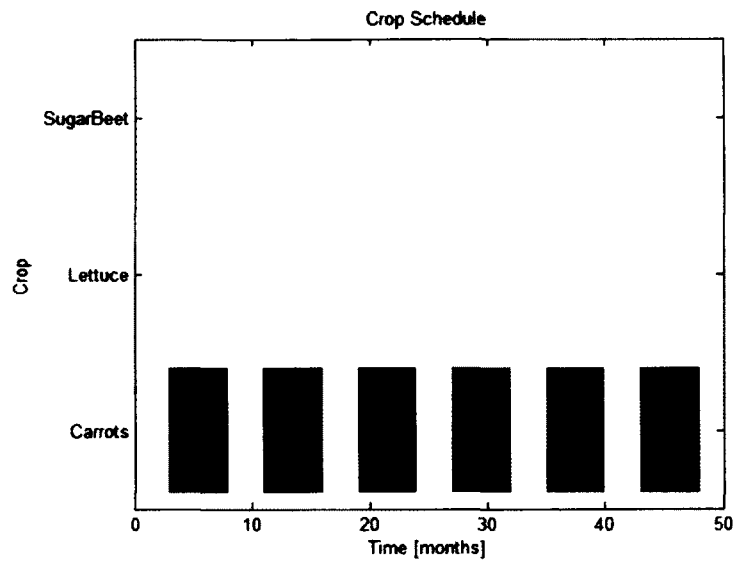
Fig. 50. Crop schedule for reducing fertilizer use for first set of crops.
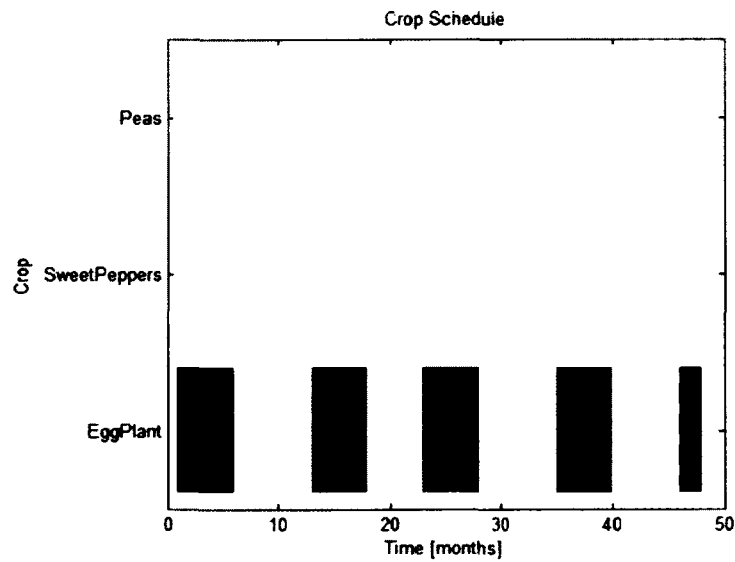
Fig. 51. Crop schedule for reducing fertilizer use for second of crops.

4.4 Crop Scheduling for Multiple Objectives

In the previous models we solved the problem for a single objective for a given set of constrains. In this section we will select two objectives that need to be satisfied while using the same constraints as before. The objective for this model is to schedule the crops that increase the economic yield, while reducing the amount of water consumed.

There are several methods to solve multiple objective problems. Evolutionary algorithms are widely used to solve multi objective problems by finding Pareto optimal solutions [64]. Scalarization techniques also exist where all the objective functions are combined into one objective function and weights are assigned to the different objective functions [65].

In our current model the two objectives are minimizing water usage and maximizing economic yield. This multi objective problem can be turned into a single objective problem by combining the two objective functions. This can be done by multiplying the potential water usage with a tentative water cost. Water costs for irrigation in USA range anywhere between $0.0005-0.001 per liter. The objective functions can now be combined to maximize the economic yield while reducing the water costs. The two objective functions that we use in the model are:

$$maximize: \quad \sum_{j=1}^{number\ of\ crops} \sum_{t=1}^{totaltime} \left(C_{tj} Y_{tj} x_{tj} - P_t W_{tj} x_{tj} \right) \tag{49}$$

C_{tj}, t_{ij} and W_{tj} are the cost, yield and water usage of crop j when planted in month t respectively. P_t is the price of water for the month t. The values for the water price are uniformly distributed between $0.0005-0.001 per liter. The constraints for this model are the same as the previous models. Fig. 52, Fig. 53show the scheduling for the two sets of

crops when there is no explicit crop rotation. In both the figures we can observe that the scheduling is a mix of multiple crops. However there is still not enough variety in the schedule.
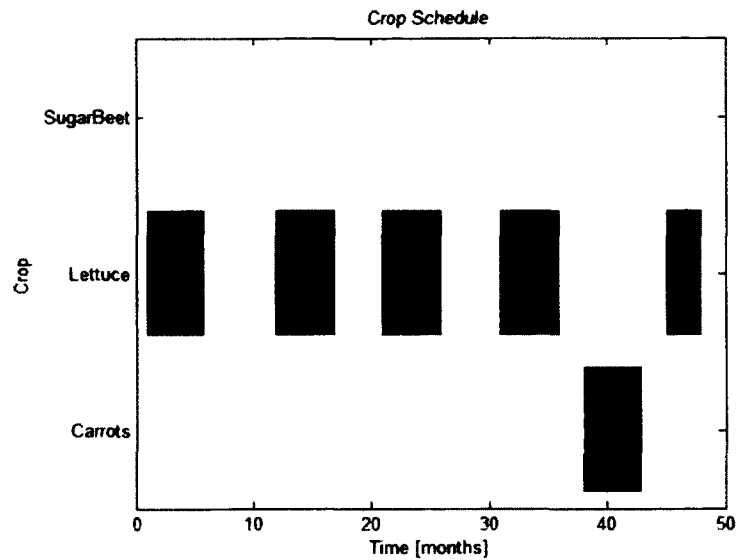
Fig. 52. Maximizing crop rotations for multiple objectives for first set of crops.
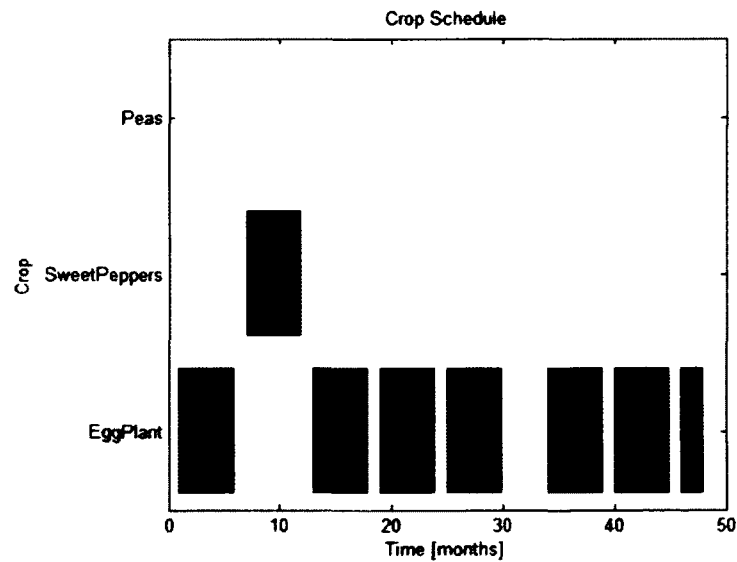
Fig. 53. Maximizing crop rotations for multiple objectives for second set of crops.

The model is evaluated again with the second set of constraints where crop rotation is enforced. The results for both sets of crops show that the selections are a mix of all the three crops, compared to the single crop selections that the model produced when the objective functions were exclusively economic or environmental benefit as seen in Fig. 54, Fig. 55. Surprisingly, eggplant still finds preference in the schedule despite its low yield. This can be attributed to its equally low environmental impact.
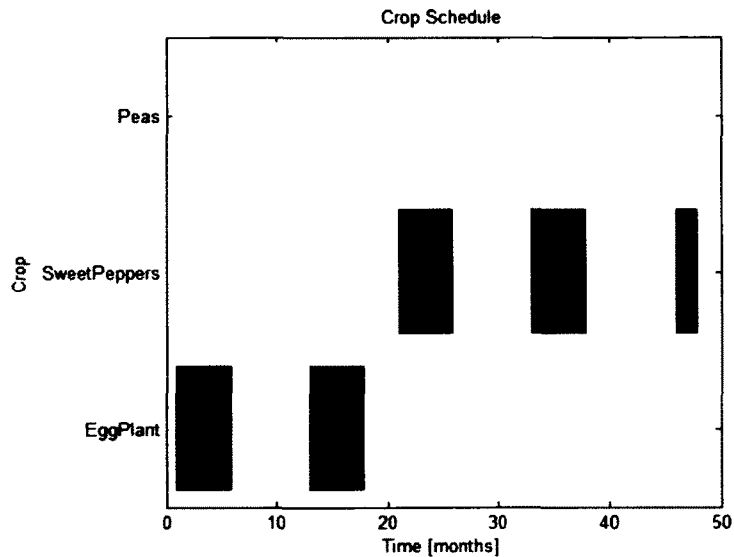
Fig. 54. Maximizing multiple objectives with rotation for first set of crops.

Crop Schedule

SugarBeet

Crop

Lettuce

Carrots

0   10   20   30   40   50

Time [months]
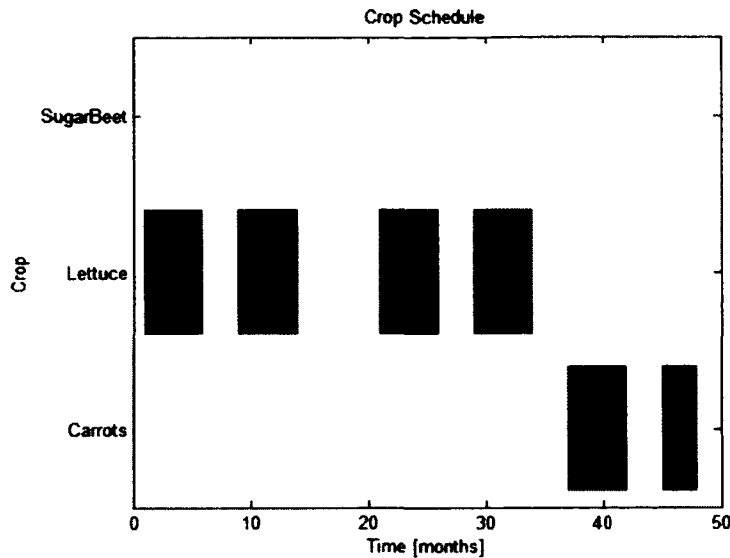
Fig. 55. Maximizing multiple objectives with explicit crop rotation for second set of

crops.

4.5 Limitations and Conclusions

In this chapter we have studied the behavior of various mathematical programming models for selecting and scheduling the planting of crops to fulfill various objectives. These objectives cover issues like performance, economic viability, and environmental impact but are not exhaustive. The decision variables in the developed model only simulate crop selection and scheduling. Decisions like resource utilization, borrowing money, farming practices and abandonment of farms need to be implemented to simulate real world situations.

The economic viability objective needs to include real world features like risk, market fluctuations and resource contentions. Similarly, the environmental impact objective only considers water usage. Future iterations of these models should also model impact of fertilizers, condition of soil and contamination of immediate environment.

The models used in this chapter are binary integer programming models. This makes the model very time consuming to run. When we modeled the performance objective using all the crops, the model ran for over five hours to produce a solution. This issue can be addressed by turning the problem into a quadratic objective programming, or using alternative algorithms to find a solution. We can also use linear programming models to implement mixed cropping schemes.

The results of this model follow directly from the results of the crop growth program discussed in the previous chapter. Hence, the validity of the models are tied directly to the validity of the crop growth model. This limitation hinders us from conducting a validation study of the different optimizers used by us, since we would have to first validate the crop model for a target region before using those results in the crop scheduling model. If the requirements to validate the crop growth model are too high, the optimizer can still be validated by using the historic yield data for a given region. We could then run the optimizer model using this historic data and compare the results to existing scenarios to validate the optimizer.

The multi objective model that was developed needs to include other objectives like social factors and farmer satisfaction. Proper weights need to be applied to the different objectives to get the more feasible results. Evolutionary algorithms also need to be studied to implement multi objective optimization. If extensive real world data is available, the use of Influence diagrams is recommended to find the best sequence of decisions.

CHAPTER 5

DECISION MODULES

The previous chapters dealt with deterministic models for evaluating crop yields and selecting optimal crops for planting. In an ideal scenario farmers would follow an optimal strategy of selecting crops and performing farm related activities. However, in reality farmers do not always make the most optimal choices. This could be the result of external factors like social, environmental or economic pressures. A simulator aimed at improving farmer decision processes first needs to understand the decision making patterns of the farmer and the factors affecting those patterns. The decision modules in the proposed simulator serve this purpose. We will study the methodology to develop an individual decision model from real world data. The various steps involved will be discussed in detail with examples. This individual model will then be extended to develop the community model. We will conduct multiple simulations and analyze the results. Finally, the results from the community model will be compared to the results from the combined crop-growth and crop selection models discussed in chapters 3 and 4.

5.1 Selecting a Modeling Methodology

In Chapter 2 we studied existing decision making methodologies used in current literature. Each methodology had its strengths and weaknesses. Influence diagrams act as great tools to express model structure, but calculating and updating point probabilities is a tough task requiring large amounts of data. Fuzzy logic can capture the uncertain nature of a decision making problem, but accurately deriving the rules to build an inference system

is not straightforward. Agent based modeling is useful in simulating the dynamic behavior of large groups of agents, but similar to fuzzy logic the rules governing agent behavior are not easy to derive. The study of the different modeling methodologies prompted us to develop a methodology to represent the farmer decision process that would incorporate the unique features of the different methods.

The first step in developing the methodology was to find a way to represent the structure of the decision process. The different decisions available to the farmers and the external factors affecting these decisions should be represented to show the causality of the decision process. An influence diagram is a valuable tool to show this process. The chance nodes can be used to depict the external factors. The decision nodes are used to show the available decisions and the utility nodes are used to the utility of the various decisions for the given factors. Fig. 56 is a representation of a sample decision process for a farmer. The expected crop, choice of fertilizer and required irrigation are all important factors that are affected by external factors like capital, soil condition and the availability of water. The decision nodes represent the different crops available to the farmer, the types of fertilizer and the choice of irrigation.
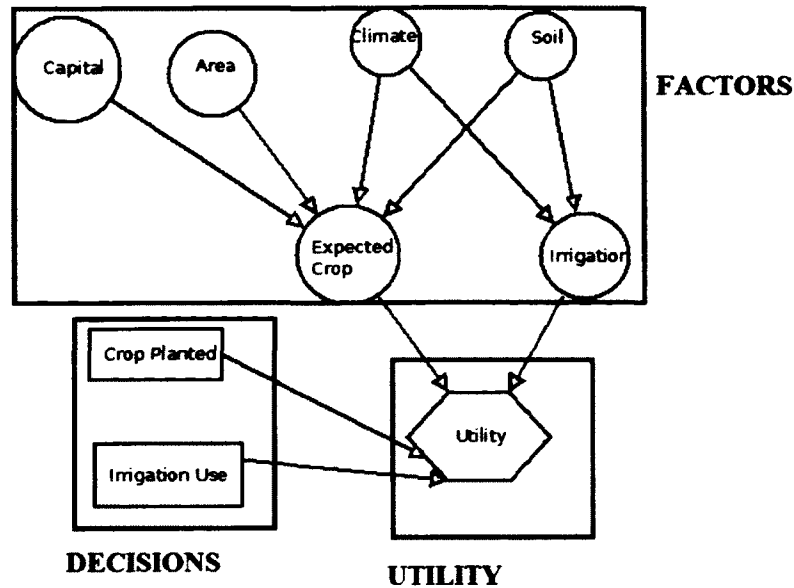
Fig. 56. Sample model of a decision process.

The second step in the modeling methodology is to calculate the impact of the factors. We need to use the data available to create a mathematical model that can output the impact of the factors for different input conditions. Influence diagrams use point based probabilities to represent this impact. The drawback to this method is that all the impacts are represented as probabilities [66]. Data constraints usually prevent us from calculating all the relevant probabilities. For example, two chance nodes with three states each would require nine probability values to be calculated. To overcome this problem we will use a fuzzy logic system to calculate the impact of the factors. Fuzzy logic uses a linguistic rule based inference system to calculate the outputs for a given set of inputs. This allows us to combat the effects of a large state space by only using rules that are pertinent to the training data set. Fuzzy rules are often derived by subject matter experts who have experience in the area of study being modeled. In this chapter we will demonstrate a quantitative method of deriving linguistic rules from the available data set.

In order to demonstrate the proposed modeling process we will use real world data that is publicly available. The data for this model was based on two sources: the Farm Accountancy and Data Network (FADN) [67], and EUROSTAT Database [68]. FADN is a European system of surveys that collects economic and land uses data for farms across Europe. It presents regional data on income and agricultural land use for all participating members of the European Union. Eurostat is the statistical office of the European Union situated in Luxembourg. Its task is to provide the European Union with statistics at European level that enable comparisons between countries and regions.

The main source of data for the presented model is from the FADN. The data was downloaded in the form of a csv and read into MATLAB. The various fields of the data are:

1. Year

2. Farm Size

3. Economic Size Units(ESU)

4. Number of farms of this particular size

5. Total utilized agricultural area

6. Area utilized and economic output for production of Cereals

7. Area utilized and economic output for production of potatoes.

8. Area utilized and economic output for production of vegetables

9. Area utilized and economic output for production of fruits

10. Area utilized and economic output for production of wine

11. Area utilized and economic output for production of forage crops:

The data was downloaded for the Piedmonte region in Italy. Piedmonte is one of the 20 regions of Italy with Turin as its capital. It is an extremely fertile agricultural region that grows a wide variety of crops. It is one of the largest producers of cereal (rice, maize, etc.) in the country. It is also famous for its vineyards and other fruits. The data was selected for a time period of 20 years, between 1989 and 2009. This data was used to infer the rules required to build the fuzzy logic system.

The use of the fuzzy logic system is twofold. First, we would like to determine the amount of money a farmer might spend on crops for a particular year. Once we obtain this value, we will then use other fuzzy logic systems that determine the amount of capital the farmer might allocate to each crop/hectare. The outputs of these fuzzy systems will then be used to calculate the area allocated to each crop, using an integer programming model.

From an initial analysis of the data, and a study of other research efforts like [69], we were able to select certain input factors that might have an effect on the output. For the initial investment, we selected Economic size, Area, Financial output from the previous year, and return on investment. The scatter plot for the relation between economic size and the initial capital, and the relation between area and initial capital are shown in Fig. 57, Fig. 58.
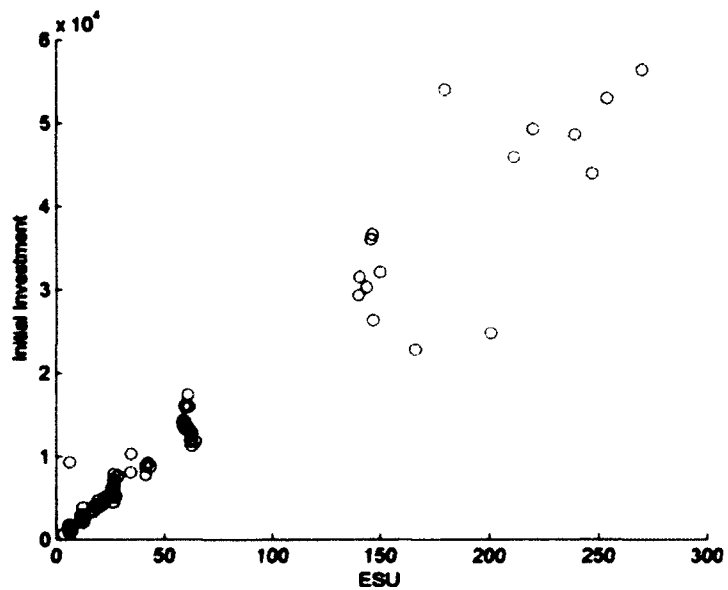
Fig. 57. Scatter plot of ESU vs Initial Capital.
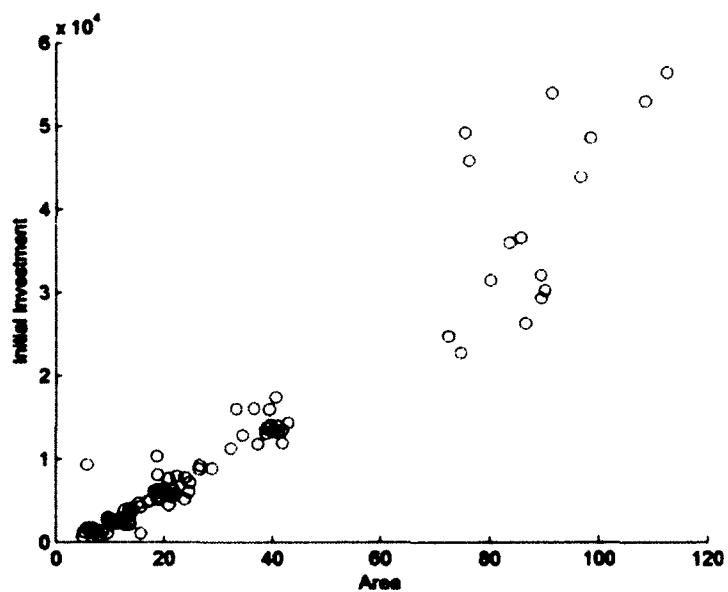
Fig. 58. Scatterplot of Area vs Initial Capital.

Table 7 shows the correlation factors for the various inputs. Using a similar approach, we calculated the correlation coefficients to measure the relationships between the various input factors, and the amount of capital spent on each crop/hectare. Based on

these values, we were able to select input factors that we felt had a strong enough effect on the capital/hectare for each crop. Table 8 lists these factors and the associated correlation values.

Table 7 Correlation factors for input capital

Factors	Initial Investment on Crops
Economic Size	0.9712
Area	0.9562
Output from previous year	0.9367
Return of investment	0.4258

Table 8 Correlation factors for the different crops

Factors	Cereals	Potatoes/Sugar beets	Vegetables	Fruits	Wine	Forage Crops
Economic Size	0.6106	0.4508	-0.2029	0.6560	0.3480	0.5850
Area	0.6678	0.4058	-0.1647	0.6734	0.3137	0.6487
Market value	0.6183	0.3734	0.4923	0.6506	0.3699	0.7751
Production Cost	0.6017	0.4099	0.4671	0.6138	0.2775	0.7376
Return of Investment	0.3776	0.2039	-0.3657	0.4082	0.2425	0.3717

5.2 Building the Individual Model

Based on our discussion about the various input factors and the outputs being simulated, Fig. 59 is the desired structure of the individual model. The ovals at the top represent the various input factors. These ovals are represented in the fuzzy logic system using appropriate membership functions. The ovals at the bottom are the outputs that are also represented using appropriate fuzzy membership functions. The output capital/hectare

is represented for each our 6 types of crops. The arrows represent the rules that explain the relation between the input and the output. The two outputs, capital/hectare, and initial capital then act as an input to an integer programming model. This model solves for the amount of area allocated to each crop by maximizing the farmers' potential income.
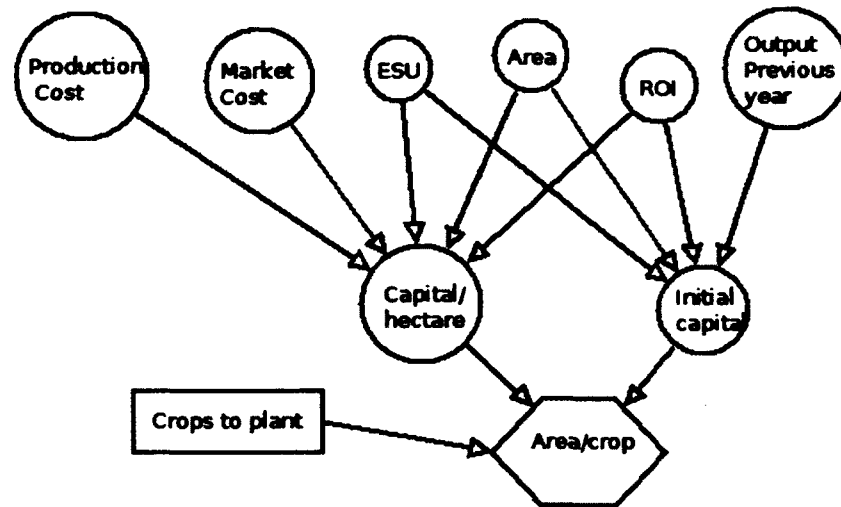
Fig. 59. Graphical representation of the model.

5.2.1 Selecting Membership Functions

To build appropriate membership functions to represent the variables, we need to understand the statistical behavior of the variables. Histograms are an easy method to observe the ranges of the various variables. The histograms of the various variables were observed to determine an appropriate shape and range for the membership functions. For example, Fig. 60, shows the histogram for the production costs of vegetables.

We can further split the data into smaller samples to fit the appropriate linguistic labels of the membership functions. For example, the label 'Small' for the production cost of vegetables can be obtained by splitting the data as shown in Fig. 61. This information

can now be used to draw the membership function for the associated 'Small' label. Fig. 62

shows the membership functions for all the inputs associated.

Fig. 60. Histogram of the variable Production cost.

Fig. 61. Histogram for the label 'Small' of Production cost.
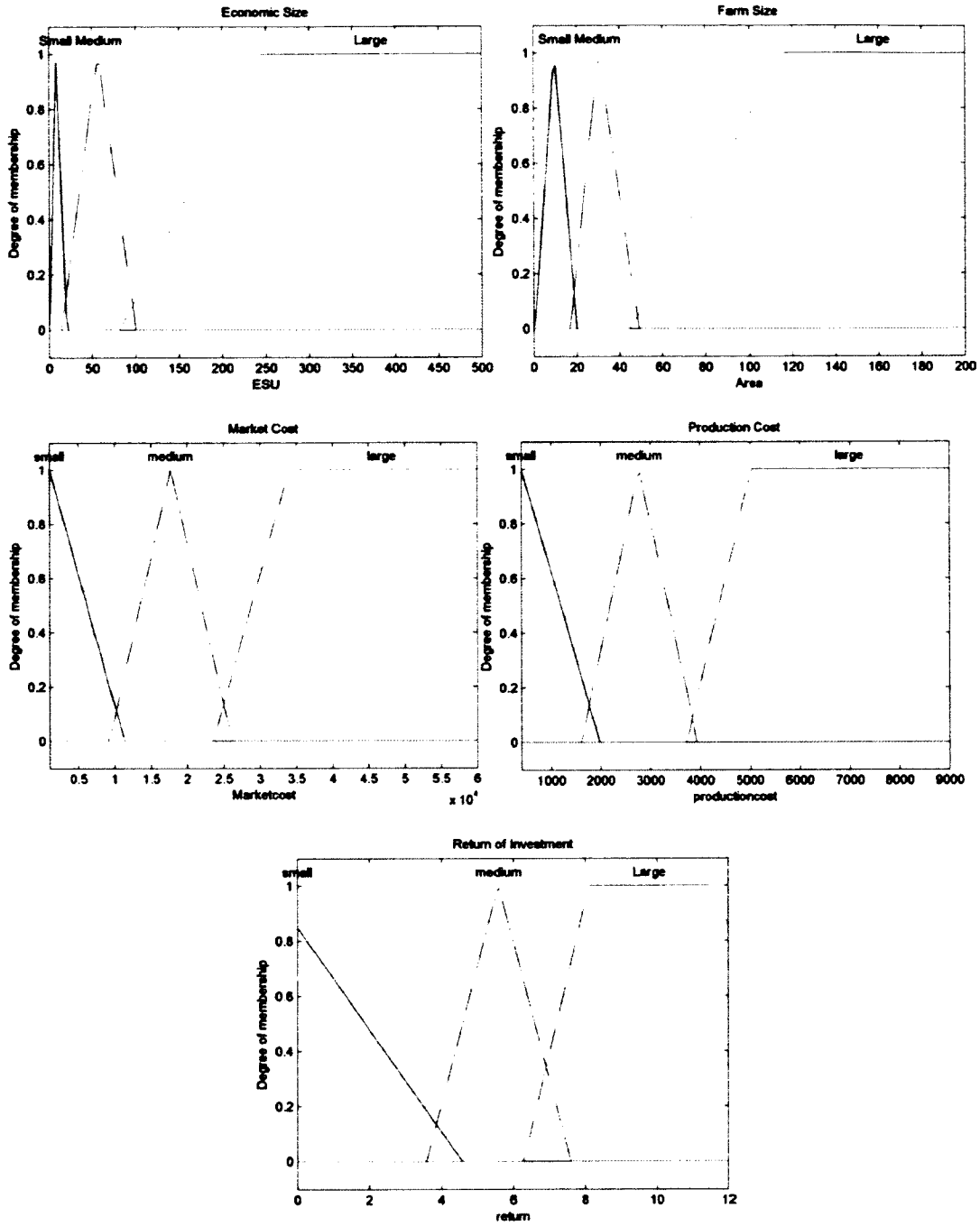
Fig. 62. Membership functions for the various input variables.

5.2.2 Rule Generation and Evaluation

The rule base of a fuzzy logic system is built using 'IF-THEN' conditions that assign an output label for various combinations of input labels. There are various methods to design the fuzzy rule base. Neural networks, genetic algorithms, and other fuzzy systems have been used in the past to derive the rule base of a fuzzy logic system.

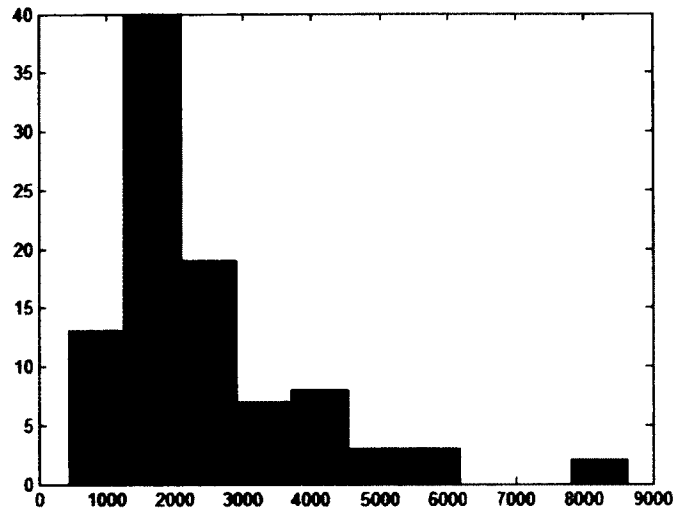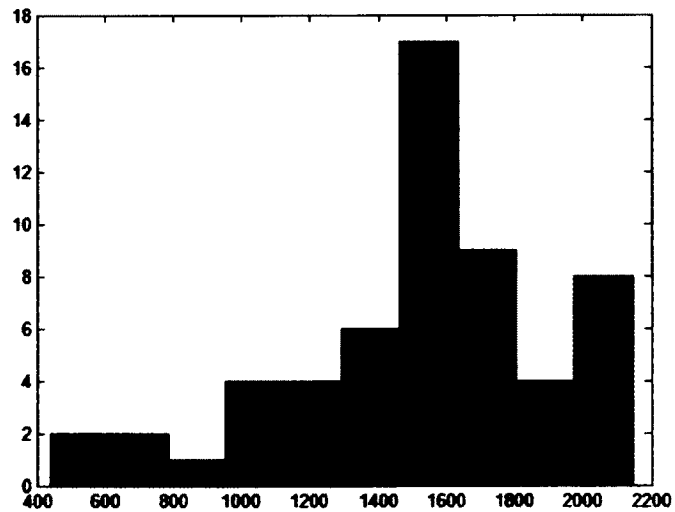In our approach we will use a modified version of the heuristic method proposed by [65]. A fuzzy rule R_j for an n-dimensional problem is written as:

$$\text{Rule } R_j: \text{If } x_1 \text{ is } A_{j1} \text{ and } \ldots\ldots\ldots \text{ and } x_n \text{ is } A_{jn} \text{ then Y is } C_j \text{ with } CF_j \qquad (50)$$
$$j=1,2,\ldots.N$$

where R_j is the jth fuzzy rule. $X_P = (x_{p1}, x_{p2}, \ldots\ldots x_{pn})$ is an n-dimensional pattern vector. Y is the desired output and A_{ji} and C_j are linguistic classifiers such as small, medium and large. CF_j is the certainty grade, or the weight of the fuzzy rule. The value of CF_j lies in the interval [0,1].

Calculate the compatibility grade $\mu_j(X_P)$ for each training pattern X_P with the fuzzy rule R_j using:

$$\mu_{R_j}(X_P) = \mu_{j1}(x_{p1}) \text{ x}\ldots.\mu_{jn}(x_{pn}) \quad \text{p=1,2,}\ldots\text{m} \qquad (51)$$
$$j=1,2,\ldots.N$$

where $\mu_{ji}(x_{pi})$ is the membership of the data point x_{pi} in the fuzzy set A_{ji}. Let us assume that the fuzzy set representing the output Y has h linguistic classifiers. For each class of the output variable calculate the sum of the compatibility grades for the training patterns for the rule R_j.

$$\beta_{class\,h}(R_j) = \sum_{X_P \,\in Class\,h} \mu_{R_j}(X_P) \qquad (52)$$

$$h = 1,2,\ldots c$$

For each class h in the above equation calculate the certainty grade CF_{hj}. This grade is the weight associated with the fuzzy rule. The certainty grade for each class of the rule R_j is given by:

$$CF_{hj} = \frac{\beta_{class\,h}(R_j)}{\sum_{h=1}^{c} \beta_{class\,h}(R_j)} \tag{53}$$

$$h = 1,2,\ldots c$$

If CF_{hj} is zero, or below a certain threshold, the rule R_j with consequent class h is not generated. The threshold can be set based on the number of training patterns and the number of the input variables. Table 9 shows a sample of the generated rules for the vegetable crop.

Table 9 Sample of rules for Vegetable FLS

1. If (ESU is Small) and (Area is Small) and (Marketcost is medium) and (Productioncost is small) and (return is Large) then (Vegetables is small) (0.66667)
2. If (ESU is Small) and (Area is Small) and (Marketcost is medium) and (Productioncost is medium) and (return is medium) then (Vegetables is small) (0.375)
3. If (ESU is Small) and (Area is Small) and (Marketcost is medium) and (Productioncost is medium) and (return is medium) then (Vegetables is large) (0.25)
4. If (ESU is Medium) and (Area is Small) and (Marketcost is small) and (Productioncost is small) and (return is medium) then (Vegetables is medium) (0.66667)
5. If (ESU is Medium) and (Area is Medium) and (Marketcost is small) and (Productioncost is small) and (return is medium) then (Vegetables is small) (0.5)
6. If (ESU is Medium) and (Area is Medium) and (Marketcost is medium) and (Productioncost is large) and (return is medium) then (Vegetables is large) (0.33333)
7. If (ESU is Large) and (Area is Large) and (Marketcost is small) and (Productioncost is small) and (return is medium) then (Vegetables is large) (0.083333)
8. If (ESU is Large) and (Area is Large) and (Marketcost is medium) and (Productioncost is medium) and (return is medium) then (Vegetables is small) (1)
9. If (ESU is Large) and (Area is Large) and (Marketcost is medium) and (Productioncost is large) and (return is medium) then (Vegetables is small) (1)

The rules generated are only effective in covering the various input combinations of the training dataset. In practical applications of the fuzzy logic system, we are often faced with data that are not always explained by the rules. For instance, the fuzzy logic system for cereals has 5 input variables, each described by three membership functions. This would require us to have 3^5, i.e., 243 rules to explain all possible combinations of the input space. However, we describe our fuzzy system for cereals using less than 40 rules. In such instances we will need to use an effective method to calculate the output using the available rules.

To overcome this problem we have adapted a variation of the degree of weighted convenience method proposed by [70]. The degree of weighted convenience is a method to calculate the importance of each fuzzy rule in a list of rules for a given data point. It is often used in scenarios where no single rule can explain the given data point. In such cases this method is used to select the rule that best explains the given data point. This selected rule is then evaluated to get the output. The definition of the degree of weighted conveniences, as proposed by Chen et al. is as follows:

Definition: Assume that there exists a testing datum $T = (x_1, x_2 \ldots \ldots x_n)$ and a fuzzy rule $R = ((L_1, L_2 \ldots, L_n), y_k)$. The degree of the weighted convenience is the summation of the multiplications of the membership value $\mu_{L_i}(x_i)$ and the individual weight w_i of the input variable X_i, i.e.,

Degree of weighted convenience of fuzzy rule R

$$R = \sum_{i=1}^{n} \mu_{L_i}(x_i) * w_i \tag{54}$$

where $\mu_{L_i}(x_i)$ denotes the degree of membership of x_i in the label L_i.

Chen et al.,[70] calculate the weight w_i using the formula:

$$w_i = \left(\frac{v_i}{\max(v_1, v_2 \ldots, v_n)}\right)^2 \tag{55}$$

$$v_i = \frac{|PD|}{|WD|} \tag{56}$$

where PD is the individual domain of the variable X_i for each type of output. In our approach, we replaced the weight with the absolute value of the correlation coefficient for the variable X_i with the output. For example, consider the following sample of three rules that represent the amount of capital allocated to vegetables per hectare:

R_o: If (ESU is Small) and (Area is Small) and (Marketcost is medium) and (Productioncost is medium) and (Return is Large) then (Vegetables is large)

R_1: If (ESU is Medium) and (Area is Small) and (Marketcost is medium) and (productioncost is medium) and (Return is medium) then

(Vegetables is medium)

R_2: If (ESU is Medium) and (Area is Small) and (Marketcost is small) and (productioncost is small) and (Return is medium) then (Vegetables is medium)

Now, assume we have to evaluate the output of the following instance of input: [50, 18, 10000, 2500, 7.2]. From the table, we already know the absolute correlation values that act as weights for our calculation. From the fuzzy membership functions in Fig. 62, we now calculate the degree of weighted convenience for each rule:

R_o: 0*0.2029 + 0.191*0.1647 + 0.0868*0.4923 +0.7468*0.4671 + 0.5017*0.3657 = **0.6065**

R_1: 0.8372*0.2029 + 0.191*0.1647 + 0.0868*0.4923 + 0.7468*0.4671 + 0.1937*0.3657 = **0.6637**

R_2: 0.8372*0.2029 + 0.191*0.1647 +0.1329*0.4923 + 0*0.4671 + 0.1937*0.3657 =

0.3376

From the above values we can see that rule R_1 has the highest chance of evaluating

the data. The amount of capital spent on vegetables per hectare is medium. Evaluating R_1

for the given input gives us an expected output of €3428.

5.2.3 Creating an Integer Programming Model

The integer programming model is used to calculate the amount of area that can be

allocated to each farm based on the total money being spent on the crops, and the

capital/hectare being spent on each crop. These values are obtained as outputs from the

fuzzy logic systems.

To create the objective function for the farmer, we will be assuming that he is a

profit maximizer and risk minimizer. The profit maximizer part aims to find the right mix

of areas to allocate to each crop, in order to maximize farmers earning potential. This

potential is based on the market prices of the crops in the previous year.

The risk minimizer part of the farmer will act to decrease the amount of risk that a

farmer is willing to take. This is manifested in the linear constraints of the model. To model

this, we will assume that the farmer does not spend more than a reasonable amount

compared to the previous years. This constraint is used to make sure that the farmer does

not use all his resources to plant only one type of crop. The general trend in the data shows

a change of approximately 5-10% between the investments. This variation in the

investment on crops is modeled using a variable called *rate*.

The linear programming model can be written as follows:

$$\text{maximize:} \quad \sum_{j=1}^{total \; number \; of \; crops} \sum_{t=1}^{total time} P_{tj} \, x_{tj} \tag{57}$$

Constraint 1: (58)

$$\sum_{t=1}^{total \; time} \sum_{j=1}^{number \; of \; crops} x_{tj} \leq Total \; Area_t$$

Constraint 2: (59)

$$\sum_{t=1}^{total \; time} \sum_{j=1}^{number \; of \; crops} C_{tj} x_{tj} \leq Total \; Capital_t$$

Constraint 3: (60)

$$\sum_{t=1}^{total \; time} \sum_{j=1}^{number \; of \; crops} C_{tj} x_{tj} \leq Investment_{(t-1)j}(1 + rate)$$

Y_{tj}:Crop yield of crop j during the time t

P_{tj}: Market price of crop j during time t

C_{tj}: Production cost of crop j during time t

$Total \; Area_t$: Total area available to the farmer at time t

$Total \; Capital_t$: Total capital available to the farmer at time t

$Investment_{(t-1)j}$: Investment on crop j for the previous year

$rate$: Percentage of increase in investment from the previous year is usually a

value between 5-10% [71]

t : The time period

j : The number of crops.

5.2.4 Individual Model Results

The model was run using 30 testing samples from the data. The model calculated the area that each farmer would allocate to the various crops, based on the input data. The output was then compared to the known allocated area. Fig. 63, contains the comparison for the simulated and actual values for the area allocated to cereals. A similar comparison was performed for the forage crop in Fig. 64. For the given testing data set, the percentage error for cereal and forage crops was less than 25% when compared to the actual values from [67].
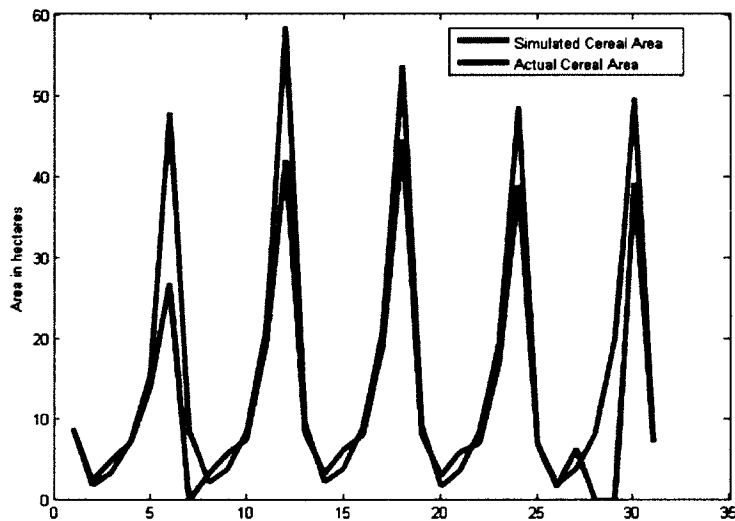

Fig. 63. Comparison of simulated vs Actual output for Cereal area.
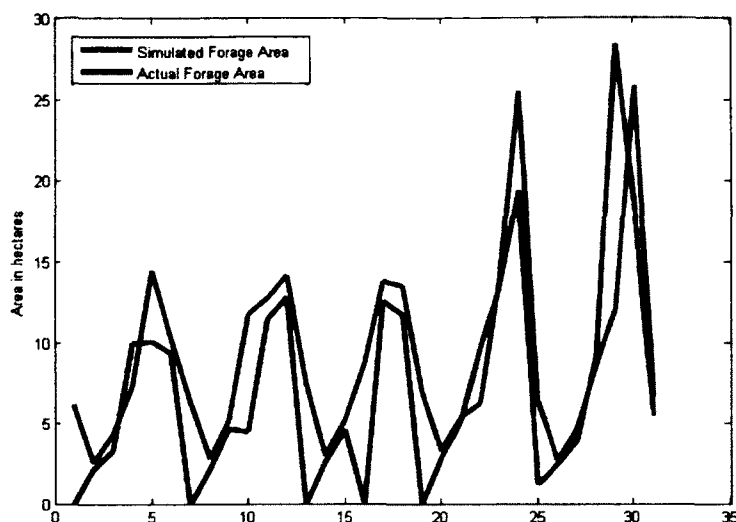
Fig. 64. Comparison of simulated vs Actual output for Forage area.

5.3 Community Model

The community model is considered as an extension to the individual model. The individual model computes the area allocated to the crops by an individual farmer. In the community model we will try to evaluate the aggregate behavior of the farmers to see if the cropping trends on a regional scale can be replicated. We will use the individual model to simulate the behavior of different types of farmers for each year. The mean of the resulting cropping trends will calculated and compared to the yearly averages for the Piedmont region. This model is also useful in conducting 'what-if' analyses. Once a base scenario is established the various parameters of the model can be varied to observe the change in the cropping trends for a region.

5.3.1 Results

The individual models were evaluated for the different farmer types over a period of 10 years between 2000 and 2009. Fig. 65, Fig. 66, Fig. 67, Fig. 68, Fig. 69 show the results of the yearly averages calculated from the model results. The X-axis represents the time period and the Y-axis represents the average area in hectares. The green colored curves are the actual values from the FADN database [67], while the blue colored curves are the results from the simulated model.
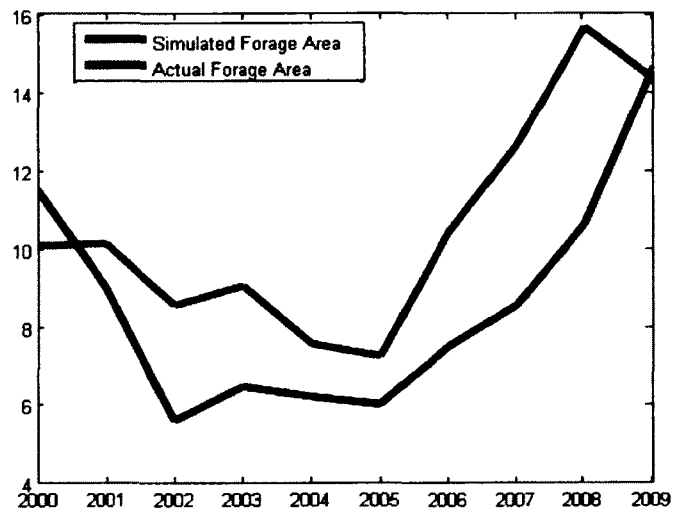
Fig. 65. Simulated yearly average for forage crop.
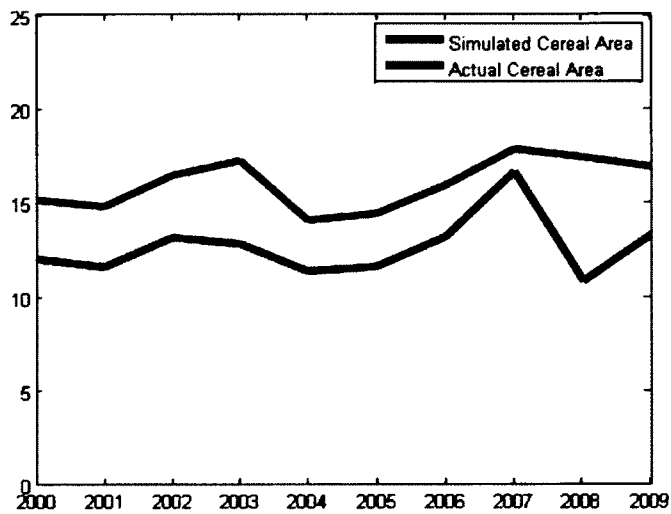
Fig. 66. Simulated yearly average for Cereals.
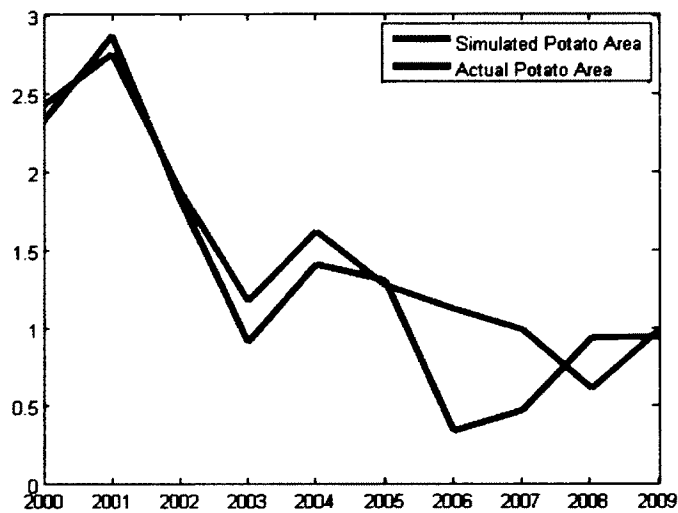
Fig. 67. Simulated yearly average for Potato crop.
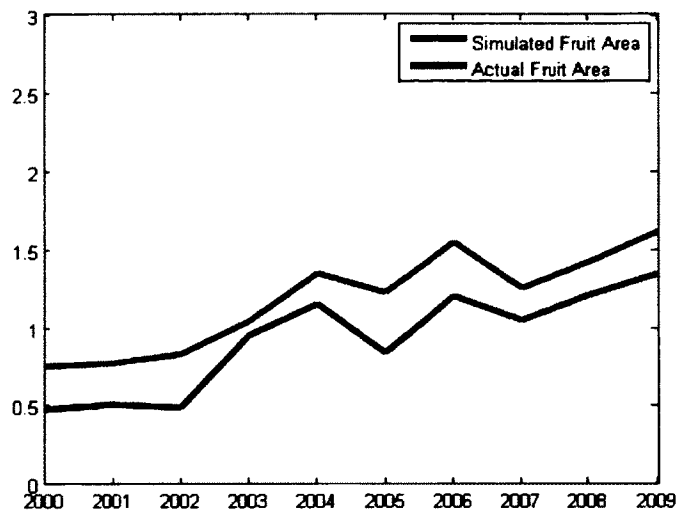
Fig. 68. Simulated yearly average for Fruits.
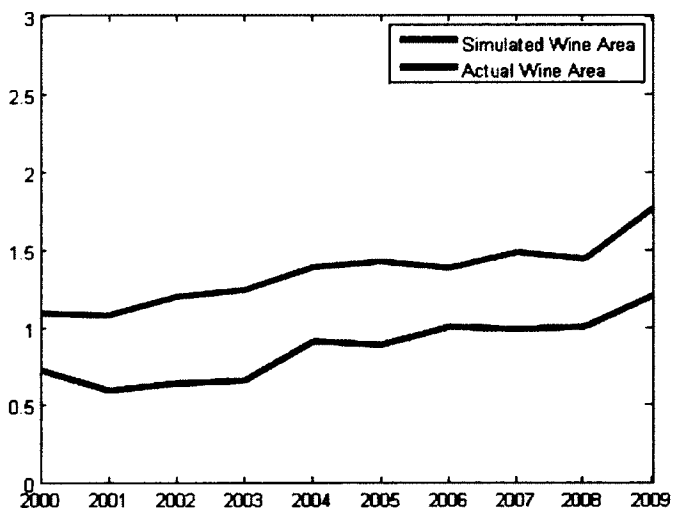
Fig. 69. Yearly average for Wine Crop.

From the figures we can observe that the yearly trends have been reflected faithfully while the actual yearly average values are not an exact match. One of the main reasons is

the limited number of training samples available to derive the rules for the system. External factors that are not reflected in the data are another reason for the underestimation of the values. The results for the fruit crop and the wine crop show the highest error rate. The major cause for this is the average area allocated to these types of crops. Both these crops are allocated less than two hectares on an average each year. This causes the result to be extremely sensitive to the values of the membership functions of the individual model.

5.3.2 Scenario Analysis

In addition to replicating the aggregate behavior of large groups of farmers, the community model can also be used to analyze hypothetical scenarios to observe changes in cropping trends. In this section we will simulate three 'what-if' scenarios to observe the response of the model. The results of the base scenario presented in the previous section will be displayed along with the actual cropping areas and the results from the new scenario. The green colored curves are the actual values from the FADN database [67], while the blue colored curves are the results from the base simulated model. The red curves are the new scenarios simulated using the model.
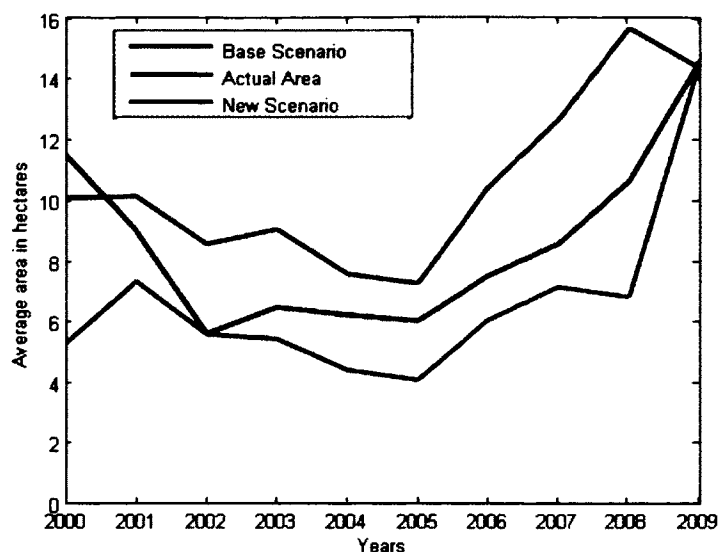
Fig. 70. Change in area for Forage crop.

In the first scenario the market price of forage crops was reduced by 20%. The simulation results are displayed in the figure above. We can observe that the change in cropping area ranges between 0-60% when compared to the base scenario. This shows the non-linear nature of the model. The price of the remaining crops was kept constant and the cereal crop showed the highest increase in cropping area for a decrease in the market price of the forage crop. A similar analysis can be done for the different crops to observe the model behavior.

For the second scenario we observe the behavior of the wine crop. Traditionally the wine crop is allocated the least amount of area in Piedmont. The reason for this is the high production cost and resources required. This scenario was used to find the amount of change required in the input conditions to cause a significant change in the cropping trends for the wine crop. The results are displayed in the Fig. 71.
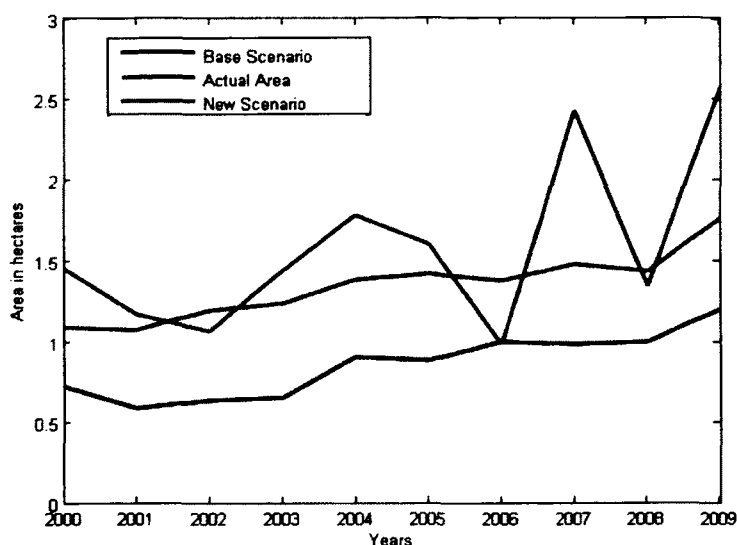
Fig. 71. Change in area for Wine.

The market price and production cost of wine were varied in increments of 5%. It was observed that once the market price was increased by 20% and the production cost decreased by 30% the average area allocated to the wine crop started changing. Though the change in the allocated area was only a maximum of two hectares, it was twice the value of the base case scenario.

In the last scenario we will observe the changes to the cropping patterns due to the changes in input parameters for selected farmers. The selected area has a high number of farmers belonging to the lower economic group. The model was modified to give a 20% hike to all farmers with an initial capital less than 100,000 euros. The model also provides a 30% subsidy to these farmers on all production costs. The model results showed that there was a significant impact on the wine and vegetable cropping pattern. Fig. 72, Fig. 73 show the results for the two crops.
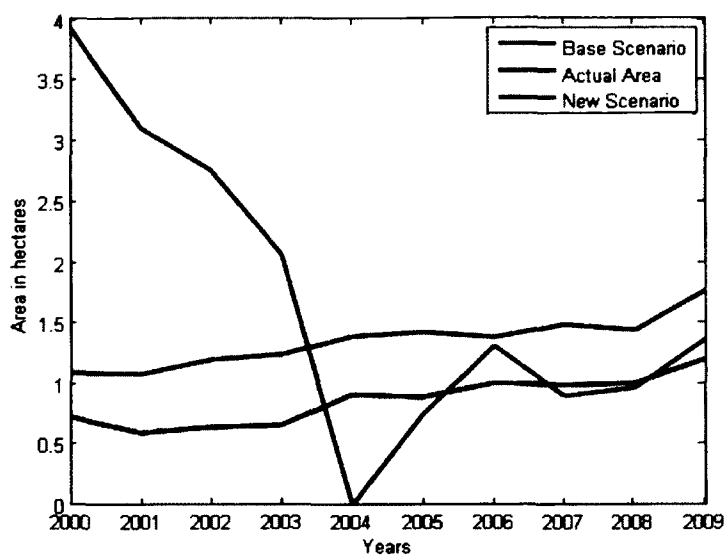
Fig. 72. Change in area of Wine for Scenario 3.
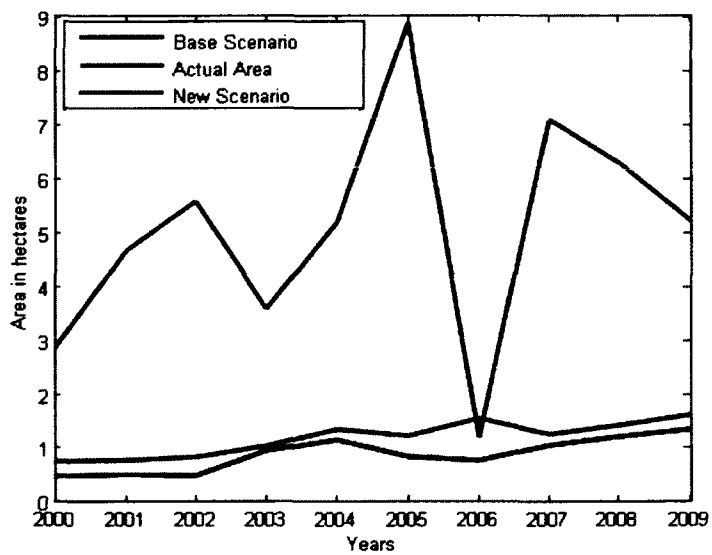
Fig. 73. Change in are of Vegetable Crop for Scenario 3.

The results show that the wine and vegetable crops had a maximum of 300% and 800% increase in their average yearly cropping area. This spike was not permanent and the values did not change by a large margin for some of the years.

5.4 Combined Model

In the previous section we evaluated a community model to simulate land use on a regional scale. The model was built to simulate farmer decision trends based on external inputs. These decisions reflect the farmer behavior and are not always the most optimal decisions. In this section we will compare the results of the community model to the results from the combined crop growth and crop selection model that was described in Chapter 3 and Chapter 4. The crop growth model is used to calculate the yield for the different crops and the crop selection model provides us with the best possible cropping decisions for the given set of crops and yearly conditions.

The first step is to calculate the yields of the different crops for a ten year period. We selected one crop for each family of crop in the original data. For example, barley was selected to represent cereal crops, tomato was selected to represent vegetables and soybeans were selected to represent forage crops. Fig. 74, Fig. 75 represent all the different yields for the 10 year period between 2000 and 2009.
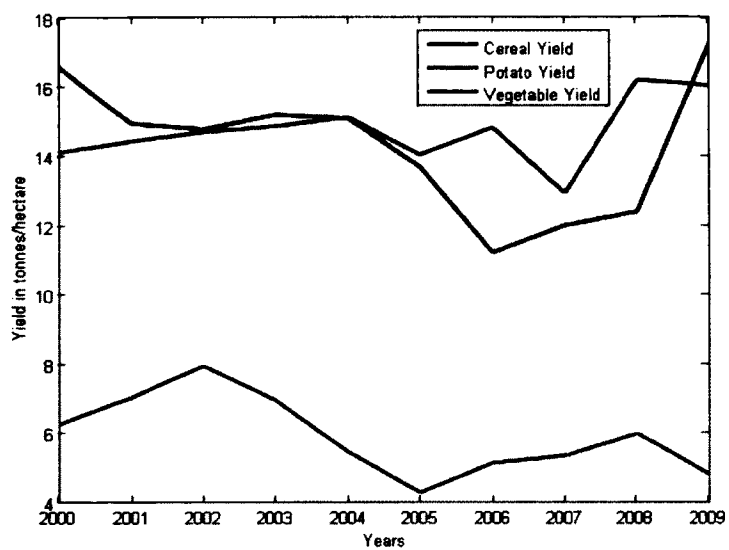
Fig. 74. Cereal, Potato and Vegetable yield.
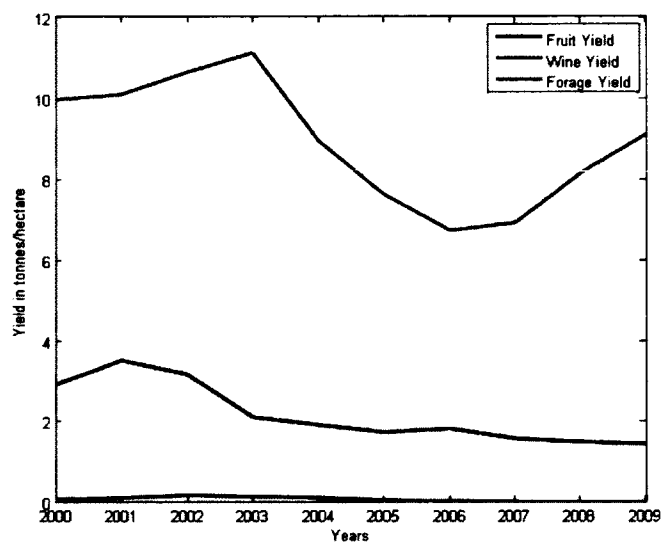
Fig. 75. Fruit, Wine and Forage yield.

5.4.1 Results

The crop selection model used to select the best possible cropping pattern is similar to the model used in Chapter 4. The first objective that is evaluated is the maximizing of the economic objective. The average market prices of the different crops have been calculated from the data and are displayed in Fig. 76, Fig. 77.
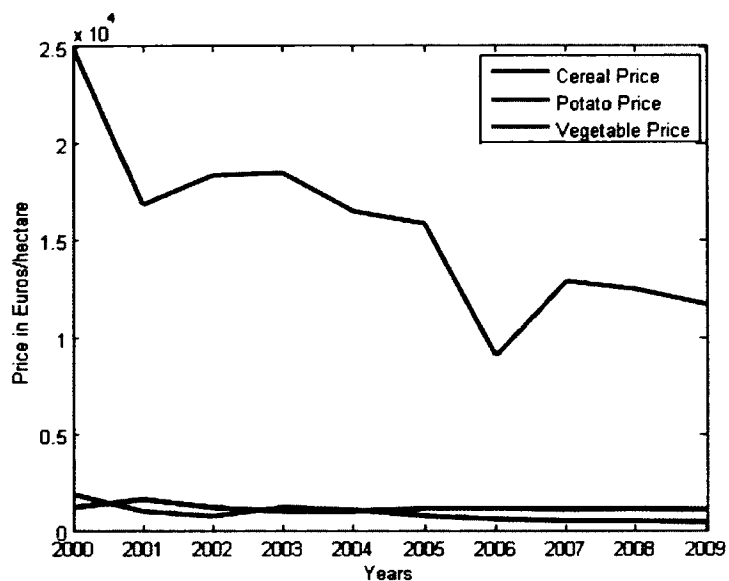
Fig. 76. Market price of cereal, potato and vegetable.
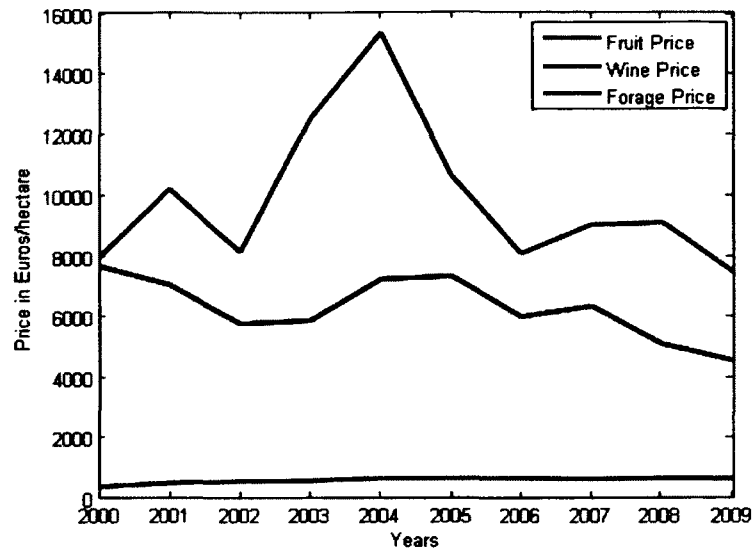
Fig. 77. Market price of fruit, wine and forage crops.

Given the yield, crop price and production cost, the model tries to maximize the monetary benefit of the crops while decreasing the production cost. There are two constraints to the model. The area allocated to each crop cannot exceed the total area. The production cost allocated to each crop cannot exceed the total capital available to the farmer. The objective model is written as:

$$maximize: \sum_{j=1}^{total\ number\ of\ crops} \sum_{t=1}^{totaltime} Y_{tj}(P_{tj}-C_{tj})\, x_{tj} \tag{61}$$

$$Constraint\ 1: \tag{62}$$

$$\sum_{t=1}^{total\ time} \sum_{j=1}^{number\ of\ crops} x_{tj} \leq Total\ Area_t$$

$$\textit{Constraint 2:} \tag{63}$$

$$\sum_{t=1}^{\textit{total time}} \sum_{j=1}^{\textit{number of crops}} C_{tj} x_{tj} \leq \textit{Total Capital}_t$$

Y_{tj}:Crop yield of crop j during the time t

P_{tj}: Market price of crop j during time t

C_{tj}: Production cost of crop j during time t

$\textit{Total Area}_t$: Total area available to the farmer at time t

$\textit{Total Capital}_t$: Total capital available to the farmer at time t

t : The 10 year time period between 2000 and 2009

j : The number of crops. The value is 6 in this scenario.

The model was run for the ten year time period between 2000 and 2009. The yearly average area allocated to the different crops was calculated and compared to the base scenario developed in the community model. The results are shown in Fig. 78, Fig. 79, Fig. 80, Fig. 81, Fig. 82, Fig. 83.
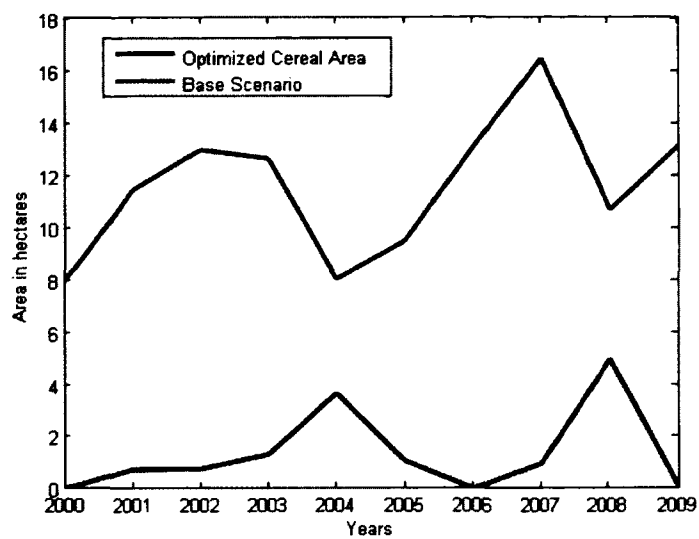
Fig. 78. Change in area for Cereal crop.
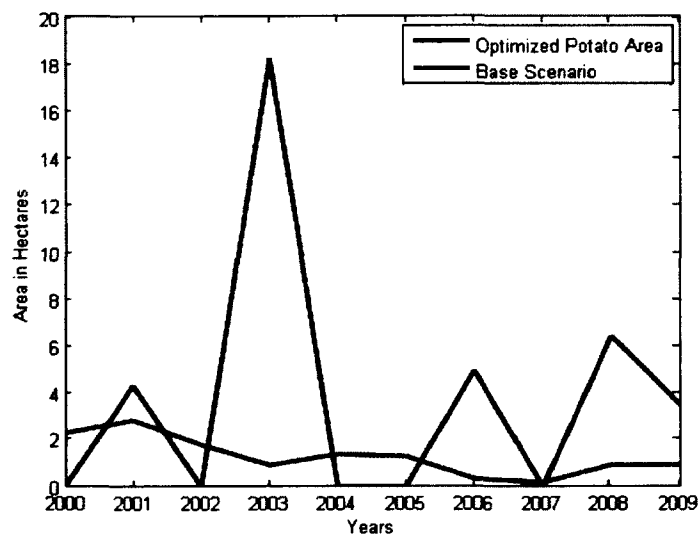
Fig. 79. Change in area for potato crop.
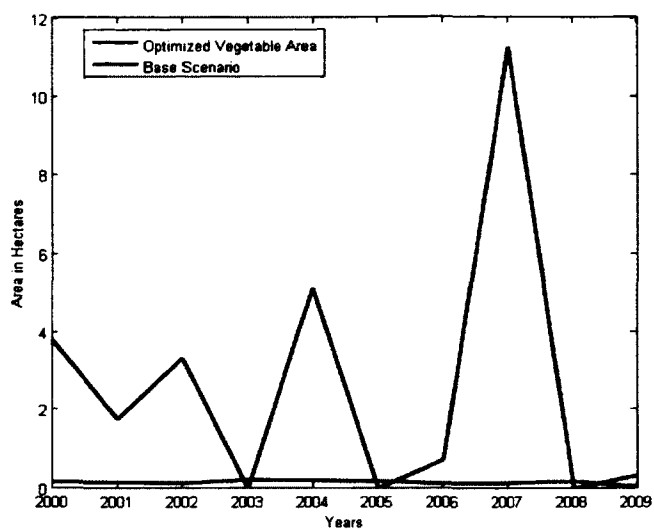
Fig. 80. Change in area for Vegetables.
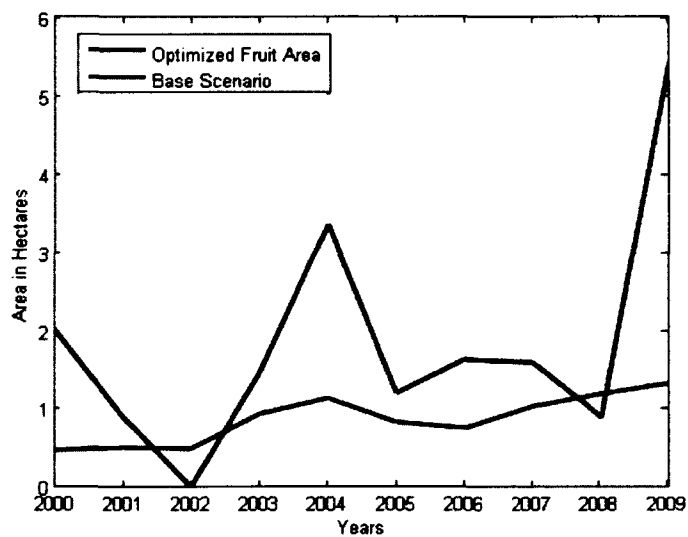
Fig. 81. Change in area for Fruits.
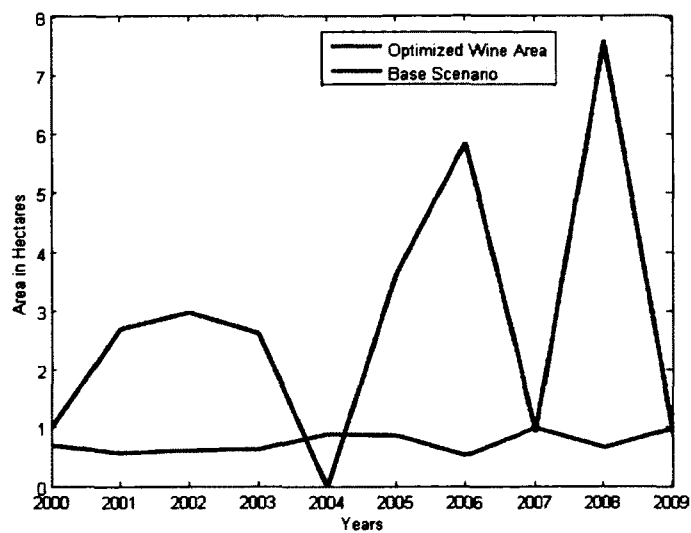
Fig. 82. Change in area for wine crops.
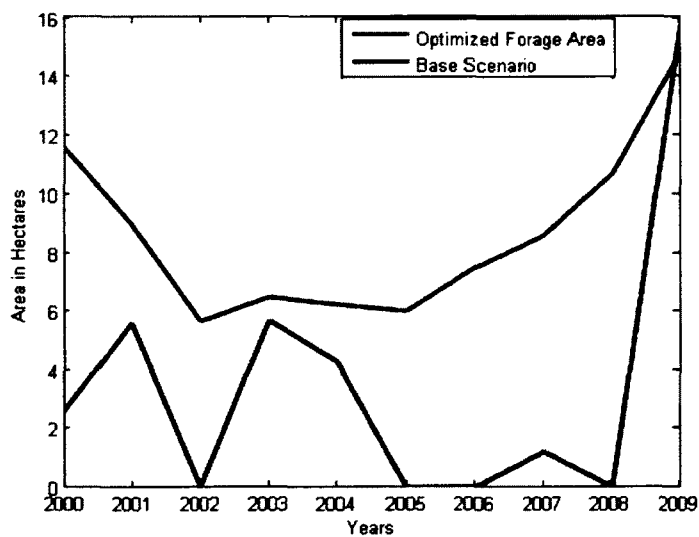
Fig. 83. Change in area for forage crops.

The simulation shows that the cereal and forage crops have a significant drop in the overall area allocated to them. This is a result of the low price of these crops despite the

high yield. On the other hand vegetable, fruit and wine area show dramatic spikes of increased land allocation. This could be a reflection of the prices of these crops despite their relatively lower yields. It is important to note that the model results reflect the best possible crop combinations without taking the risk associated into consideration. In the base scenario we can observe that farmers allocate large areas to forage and cereal crops because they consistently produce high yields and can be considered to be low risk crops. Fruit and wine crops have higher market price but the relatively low yields mean that there is a higher risk involved.

5.5 Limitations and Conclusions

In this chapter we studied the process of defining and implementing an individual decision model. This model was then extrapolated to study the cropping patterns of farmers on a regional scale. With the use of real world data we were able to see the process of deriving the membership functions and rules for the system and the results of implementing the model.

The biggest limitation of the current model is the lack of available data sources for thorough validation. This stops us from fine tuning the various aspects of the model or identifying any required additions. Data limitations also stop us from addressing other farmer objectives like social and environmental goals that farmers might have. The heuristic process for deriving the rules can also be improved by using genetic algorithms to select only the best possible rules.

The current community model is implemented as a collection of individual models. Future iterations of this model need to define a communication process between the

different agents to create a proper agent based model of the farming process. External agents like farm managers or village heads play an important role in the farming process of developing countries. With the help of subject matter experts the community model should implement such agents as it is useful in capturing the social aspects of the farmer decision process. The farmer agent behavior can be updated to become more dynamic. This can include the option to change the main objective of a farmer from economic gain to environmental responsibility or social well-being, depending on the farmers socio-economic status.

The combined model was helpful in showing the contrast between the farmer behavior and the ideal behavior. This combined model can also be extended to implement various objectives and case scenarios. If detailed data is available for a region the combined model can be useful in showing the appropriate crop rotations and contrast it with existing farmer behavior.

This chapter has covered model generation and evaluation for an available set of data. We were also able to conduct some initial scenario analysis using the base model. We believe that the actual methodology used to implement the model is quite robust, however, the accuracy of these scenarios and the results generated by this model are yet to be validated. As discussed before, the biggest hindrance to scenario validation is the lack of detailed data sets. If the individual and community models are to be used as actual decision support tools, a target study in a specific agricultural region needs to be conducted.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

This dissertation presents a simulation framework that can be used as a decision support system for farmers and researchers. Multiple models are presented that can be used either individually or together to evaluate farm specific scenarios. Chapters 1 and 2 introduce the need for this research and the existing research on this topic. The methods and associated research efforts available in the literature are explained in detail in chapter 2.

Chapter 3 introduced a generic crop growth program that can be used in our research. The mathematical equations and the important variables have been discussed. The model has been evaluated to calculate the yield, water use, and fertilizer requirements for various crops. Sensitivity analysis is also performed to identify the variables with the highest impact on the output. The yields generated by the model were compared to real world yield data available online [62]. We observed that the existing data sources show a huge variability in yield values for the same crop. This could be due to differing seed technology, farming and weather conditions. We observed that the values predicted by the model were within the observed ranges for all the crops. This shows us that the though the crop model results might not be completely accurate, the yield values generated by the model are a valid representation of the real world results.

Chapter 4 presents a scheduling model that can be used to calculate crop planting schedules for performance, economic and environmental objectives. We were able to demonstrate the changes in the cropping schedules based on the objective being maximized. A multi-objective approach has also been presented to study the model

behavior when more than one objective needs to be evaluated. Since this model works on the results of the crop model, the yield values used for scheduling fall in an acceptable range for the given weather conditions. The economic objectives were developed using real world price data from the USDA website [62] as seen in Fig. 36, Fig. 37. The optimizer output is a suggestion for the most optimal schedule given weather, yield and pricing conditions. As such it is hard to validate since we are not actually simulating a real world scenario. However, the actual data used in building the optimizer are based on real world data and valid crop model data.

Chapter 5 presents decision making methodologies to evaluate and predict farmer behavior. We developed an individual model using fuzzy logic and linear programming to simulate farmer behavior. This individual model was extended to develop the community model that can simulate farmer behavior on an aggregate scale. Real world data from FADN was used to run the model and discuss its limitations. Additionally the community model was combined with the crop growth model and the scheduling model to compare and contrast the farmer decision process to an optimal decision process. The data used to build these models was obtained from the FADN database. We compared the results of the model to the actual land use patterns in the data. This validation is displayed in Fig. 63-Fig. 69. We can observe that the change in land use was captured accurately, even if the actual land use values were not accurate.

This research advances the current research efforts by combining the different models into one simulation framework. This can allow farmers and other researchers to understand the factors and impact of existing decision trends. By combining the decision

modules with the crop growth and scheduling modules farmers can be advised on the best practices for their farms to fulfill necessary objectives.

Though multiple models have been developed to evaluate farming scenarios the simulation framework is not yet comprehensive. The crop growth program is still generic in nature and cannot simulate all possible environmental scenarios. Modern seed technology tailors crops for the region in which they are being planted. For example, carrots generally produce a yield of 10-20 tonnes/hectare. However, commercial productions of carrots have been shown to produce 70-80 tonnes/hectare. The crop parameters necessary to simulate such varying crop types are not yet available in the current literature. Further iterations of the crop growth model would benefit from adding these parameters.

Companion planting of crops is another interesting phenomenon that would greatly benefit our crop growth model. Existing research on this topic is still limited to experimentation and no substantial mathematical models exist on this topic. These factors also affect our ability to validate the model accurately. When the crop model output is compared to real world data the actual values might not always match up since the model is based on general values of the input parameters. Specific strains of a crop might have different parameter values when compared with the generic model. Farming practices and conditions also impact the overall yield of crops. In the absence of this information the crop model output cannot be used for accurate validation.

The crop scheduling methods works off of the output received from the crop growth program to evaluate various objectives as mentioned before. Future improvements to this model should try and implement social objectives to the model. Social objectives are highly

location specific and are generally qualitative. Quantifying and generalizing such objectives would improve the range and flexibility of the model. In addition to the LP methods, Multi Criteria Decision Making(MCDM) methods would also be a good match for the scheduling module. Since this optimizer model is directly related to the outputs from the crop growth model, the validation of the optimizer is heavily dependent on the validity of the crop model.

The decision making module is an adequate starting step to simulate farmer decision process. The motivation to develop the methods in that section was driven by existing data availability. The large number of variables involved in the farmers' decision process coupled with the low sample size of the available data led to the development of the existing methods. A better approach to refine and streamlines these models would be to conduct a study on a specific geographic location to identify the factors that had the highest impact on a farmer's decision process. This can help in creating a generic decision model that can be adapted to various regions. It would also be helpful to receive input from subject matter experts to assess the weights of the different variables used in the model.

The validation aspects of the individual and community models have room for improvement. As discussed previously, the rule generation methods are limited by the granularity of the data. The results from the models show that we were able to capture the change in decision trends while the actual values were not accurate. We feel that the accuracy of the models can be significantly increased through the access to a larger dataset. Since these models are also used for the scenario analysis, the validity of these simulations are directly tied to the validity of the models. Finally, the combined model discussed at the end of Chapter 5 uses the outputs from all the layers of the framework. Since we used the

combined model to simulate a hypothetical scenario the validation of the combined model is directly tied to the validation of the individual models. A focused case study could help us in evaluating the validity of the combined model.

The current framework is still not a cohesive multi-model unit. The information exchanged between the various levels of the model needs to be streamlined. This process will be benefitted by the development of an ontology for the proposed framework. The data types, classes and interfaces need to be fleshed out for advancing the composability and interoperability of the system. This process will be especially helpful when adapting the framework to different geographic regions as it can help us in identifying the exact data types and input required to run the simulations.

The proposed framework needs input from a researcher or an expert to run the simulations and explain the results. Considering that our target audience is uneducated farmers this approach is not yet viable for their use. The different game based and participatory approaches discussed in chapter 2 should serve as inspiration make our framework more approachable to the target audience. We will need to identify user interface components and visual representations that can make the models, and their results an intuitive learning experience for the farmers. We will have to find a middle ground between complexity and usability to provide an interactive experience for the farmer while trying to enhance their decision making process.

REFERENCES

[1]   R. Lowrance, P. F. Hendrix, and E. P. Odum, "A hierarchical approach to sustainable agriculture," *Am. J. Altern. Agric.*, vol. 1, no. 04, pp. 169–173, 1986.

[2]   D. Rigby, P. Woodhouse, T. Young, and M. Burton, "Constructing a farm level indicator of sustainable agricultural practice," *Ecol. Econ.*, vol. 39, no. 3, pp. 463–478, Dec. 2001.

[3]   S. Von Wirén-Lehr, "Sustainability in agriculture - An evaluation of principal goal-oriented concepts to close the gap between theory and practice," *Agric. Ecosyst. Environ.*, vol. 84, no. 2, pp. 115–129, 2001.

[4]   D. Tilman, K. G. Cassman, P. a Matson, R. Naylor, and S. Polasky, "Agricultural sustainability and intensive production practices.," *Nature*, vol. 418, no. 6898, pp. 671–677, 2002.

[5]   B. A. Keating and P. S. Carberry, "Emerging opportunities and challenges for Australian broadacre agriculture," *Crop and Pasture Science*, vol. 61 , no. 4. pp. 269–278, 2010.

[6]   E. Jakku and P. J. Thorburn, "A conceptual framework for guiding the participatory development of agricultural decision support systems," *Agric. Syst.*, vol. 103, no. 9, pp. 675–682, Nov. 2010.

[7]   M. Matthies, C. Giupponi, and B. Ostendorf, "Environmental decision support systems: Current issues, methods and tools," *Environ. Model. Softw.*, vol. 22, no. 2, pp. 123–127, Feb. 2007.

[8]   P. Allen, D. van Dusen, J. Lundy, and S. Gliessman, "Expanding the Definition of Sustainable Agriculture." 1991.

[9]   R. L. McCown, "Locating agricultural decision support systems in the troubled past and socio-technical complexity of 'models for management,'" *Agric. Syst.*, vol. 74, no. 1, pp. 11–25, 2002.

[10]   K. Louhichi, A. Kanellopoulos, S. Janssen, G. Flichman, M. Blanco, H. Hengsdijk, T. Heckelei, P. Berentsen, A. O. Lansink, and M. Van Ittersum, "FSSIM, a bio-economic farm model for simulating the response of EU farming systems to agricultural and environmental policies," *Agric. Syst.*, vol. 103, no. 8, pp. 585–597, 2010.

[11]   D. M. Oliver, R. D. Fish, M. Winter, C. J. Hodgson, a. L. Heathwaite, and D. R. Chadwick, "Valuing local knowledge as a source of expert data: Farmer engagement

and the design of decision support systems," *Environ. Model. Softw.*, vol. 36, pp. 76–85, 2012.

[12] H. H. Chouinard, T. Paterson, P. R. Wandschneider, and A. M. Ohler, "Will Farmers Trade Profits for Stewardship? Heterogeneous Motivations for Farm Practice Selection," *L. Econ.* , vol. 84 , no. 1 , pp. 66–82, Feb. 2008.

[13] I. R. Cooke, E. H. a Mattison, E. Audsley, a P. Bailey, R. P. Freckleton, a R. Graves, J. Morris, S. a Queenborough, D. L. Sandars, G. M. Siriwardena, P. Trawick, a R. Watkinson, and W. J. Sutherland, "Empirical test of an agricultural landscape model: The importance of farmer preference for risk aversion and crop complexity," *SAGE Open*, vol. 3, no. 2, pp. 1–16, 2013.

[14] M. Punkari and P. White, "Social and Environmental Sustainability of Agriculture and Rural Development Investments: A Monitoring and Evaluation Toolkit," *Development*, 2007.

[15] P. Y. Le Gal, P. Dugué, G. Faure, and S. Novak, "How does research address the design of innovative agricultural production systems at the farm level? A review," *Agric. Syst.*, vol. 104, no. 9, pp. 714–728, 2011.

[16] G. M. Bazzani, S. Di Pasquale, V. Gallerani, S. Morganti, M. Raggi, and D. Viaggi, "The sustainability of irrigated agricultural systems under the Water Framework Directive: first results," *Environ. Model. Softw.*, vol. 20, no. 2, pp. 165–175, 2005.

[17] P. B. M. Berentsen, "Effects of animal productivity on the costs of complying with environmental legislation in Dutch dairy farming," *Livest. Prod. Sci.*, vol. 84, no. 2, pp. 183–194, 2003.

[18] T. El-Nazer and B. A. McCarl, "The Choice of Crop Rotation: A Modeling Approach and Case Study," *Am. J. Agric. Econ.*, vol. 68, no. 1, pp. 127–136, 1986.

[19] R. Bosma, U. Kaymak, J. van den Berg, H. Udo, and J. Verreth, "Using fuzzy logic modelling to simulate farmers' decision-making on diversification and integration in the Mekong Delta, Vietnam," *Soft Comput.*, vol. 15, no. 2, pp. 295–310, 2011.

[20] J. Marsal and C. O. Stöckle, "Use of CropSyst as a decision support system for scheduling regulated deficit irrigation in a pear orchard," *Irrig. Sci.*, vol. 30, no. 2, pp. 139–147, 2012.

[21] B. A. Keating, P. S. Carberry, G. L. Hammer, M. E. Probert, M. J. Robertson, D. Holzworth, N. I. Huth, J. N. G. Hargreaves, H. Meinke, Z. Hochman, G. McLean, K. Verburg, V. Snow, J. P. Dimes, M. Silburn, E. Wang, S. Brown, K. L. Bristow, S. Asseng, S. Chapman, R. L. McCown, D. M. Freebairn, and C. J. Smith, "An overview of APSIM, a model designed for farming systems simulation," *Eur. J. Agron.*, vol. 18, no. 3–4, pp. 267–288, 2003.

[22] F. Bartolini, G. M. Bazzani, V. Gallerani, M. Raggi, and D. Viaggi, "The impact of water and agriculture policy scenarios on irrigated farming systems in Italy: An analysis based on farm level multi-attribute linear programming models," *Agric. Syst.*, vol. 93, no. 1–3, pp. 90–114, 2007.

[23] J. L. Ticehurst, A. Curtis, and W. S. Merritt, "Using Bayesian Networks to complement conventional analyses to explore landholder management of native vegetation," *Environ. Model. Softw.*, vol. 26, no. 1, pp. 52–65, 2011.

[24] P. S. Carberry, Z. Hochman, R. L. McCown, N. P. Dalgliesh, M. A. Foale, P. L. Poulton, J. N. G. Hargreaves, D. M. G. Hargreaves, S. Cawthray, N. Hillcoat, and M. J. Robertson, "The FARMSCAPE approach to decision support: farmers', advisers', researchers' monitoring, simulation, communication and performance evaluation," *Agric. Syst.*, vol. 74, no. 1, pp. 141–177, 2002.

[25] M. Lippe, T. Thai Minh, A. Neef, T. Hilger, V. Hoffmann, N. T. Lam, and G. Cadisch, "Building on qualitative datasets and participatory processes to simulate land use change in a mountain watershed of Northwest Vietnam," *Environ. Model. Softw.*, vol. 26, no. 12, pp. 1454–1466, 2011.

[26] C. Knörzer, Robertson, Hönninger, "Evaluation and Performance of the APSIM Crop Growth Model for German Winter Wheat, Maize and Fieldpea Varieties in Monocropping and Intercropping Systems," *J. Agric. Sci. Technol. B;9/2/2011, p698*, 2011.

[27] C. O. Stockle, S. A. Martin, and G. S. Campbell, "CropSyst, a cropping systems simulation model: Water/nitrogen budgets and crop yield," *Agric. Syst.*, vol. 46, no. 3, pp. 335–359, 1994.

[28] A. Samperio, M. J. Moñino, J. Marsal, M. H. Prieto, and C. Stöckle, "Use of CropSyst as a tool to predict water use and crop coefficient in Japanese plum trees," *Agric. Water Manag.*, vol. 146, no. 0, pp. 57–68, 2014.

[29] K. H. Chenu Graeme, van Oosterom Erik, Christopher Jack, McLean Greg, Doherty Al, Chapman Scott, "Modelling crop physiology and genetics to simulate Genotype x Management x Environment (GxMxE) interactions," *APSIM Users and Developers Forum*. Canberra, Australia, 2012.

[30] Y. Zhang, C. Li, Z. Xiuji, and M. I. Berrien, "A simulation model linking crop growth and soil biogeochemistry for sustainable agriculture," *Ecol. Modell.*, vol. 151, no. 1, pp. 75–108, 2002.

[31] X. Mo, S. Liu, Z. Lin, Y. Xu, Y. Xiang, and T. R. McVicar, "Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain," *Ecol. Modell.*, vol. 183, no. 2–3, pp. 301–322, 2005.

[32] S. G. K. Adiku, M. Reichstein, A. Lohila, N. Q. Dinh, M. Aurela, T. Laurila, J. Lueers, and J. D. Tenhunen, "PIXGRO: A model for simulating the ecosystem CO2 exchange and growth of spring barley," *Ecol. Modell.*, vol. 190, no. 3–4, pp. 260–276, 2006.

[33] L. V Kantorovich, "Mathematical Methods of Organizing and Planning Production," *Manage. Sci.*, vol. 6, no. 4, pp. 366–422, Jul. 1960.

[34] F. L. Hitchcock, "The Distribution of a Product from Several Sources to Numerous Localities," *J. Math. Phys.*, vol. 20, pp. 224–230, 1941.

[35] K. J. van Calker, P. B. M. Berentsen, I. M. J. de Boer, G. W. J. Giesen, and R. B. M. Huirne, "An LP-model to analyse economic and ecological sustainability on Dutch dairy farms: model presentation and application for experimental farm 'de Marke,'" *Agric. Syst.*, vol. 82, no. 2, pp. 139–160, 2004.

[36] K. Falconer and I. Hodge, "Pesticide taxation and multi-objective policy-making: farm modelling to evaluate profit/environment trade-offs," *Ecol. Econ.*, vol. 36, no. 2, pp. 263–279, 2001.

[37] N. B. . B. Fradj Cyril.; Clodic, Melissa.; Jayet, Pierre-Alain., "Limiting the Nitrogen Losses by N-tax and Bioenergy Support: A Quantitative Analysis of Environmental Policy Mix Impacts in the North of France," *European Association of Agricultural Economists 2011 International Congress*. Zurich, Switzerland, p. 12, 2011.

[38] J. K. Jolayemi and J. O. Olaomi, "A mathematical programming procedure for selecting crops for mixed-cropping schemes," *Ecol. Modell.*, vol. 79, no. 1–3, pp. 1–9, 1995.

[39] J. K. Jolayemi and F. O. Olorunniwo, "A deterministic model for planning production quantities in a multi-plant, multi-warehouse environment with extensible capacities," *Int. J. Prod. Econ.*, vol. 87, no. 2, pp. 99–113, 2004.

[40] F. Dolisca, J. M. McDaniel, D. A. Shannon, and C. M. Jolly, "Modeling farm households for estimating the efficiency of policy instruments on sustainable land use in Haiti," *Land use policy*, vol. 26, no. 1, pp. 130–138, 2009.

[41] J. Fairweather, "Understanding how farmers choose between organic and conventional production: Results from New Zealand and policy implications," *Agric. Human Values*, vol. 16, no. 1, pp. 51–63, 1999.

[42] R. Farmani, H. J. Henriksen, D. Savic, and D. Butler, "An evolutionary Bayesian belief network methodology for participatory decision making under uncertainty: An application to groundwater management," *Integr. Environ. Assess. Manag.*, vol. 8, no. 3, pp. 456–461, 2012.

[43] M. Sami, M. J. Shiekhdavoodi, M. Pazhohanniya, and F. Pazhohanniya, "Environmental comprehensive assessment of agricultural systems at the farm level using fuzzy logic: A case study in cane farms in Iran," *Environ. Model. Softw.*, vol. 58, no. 0, pp. 95–108, 2014.

[44] C. Sattler, U. Stachow, and G. Berger, "Expert knowledge-based assessment of farming practices for different biotic indicators using fuzzy logic," *J. Environ. Manage.*, vol. 95, no. 1, pp. 132–143, 2012.

[45] L. An, "Modeling human decisions in coupled human and natural systems: Review of agent-based models," *Ecol. Modell.*, vol. 229, no. 0, pp. 25–36, 2012.

[46] N. Becu, P. Perez, a. Walker, O. Barreteau, and C. Le Page, "Agent based simulation of a small catchment water management in northern Thailand Description of the CATCHSCAPE model," *Ecol. Modell.*, vol. 170, no. 2–3, pp. 319–331, 2003.

[47] G. P. Malanson, A. M. Verdery, S. J. Walsh, Y. Sawangdee, B. W. Heumann, P. M. McDaniel, B. G. Frizzelle, N. E. Williams, X. Yao, B. Entwisle, and R. R. Rindfuss, "Changing crops in response to climate: Virtual Nang Rong, Thailand in an agent based simulation," *Appl. Geogr.*, vol. 53, no. 0, pp. 202–212, 2014.

[48] N. Becu, A. Neef, P. Schreinemachers, and C. Sangkapitux, "Participatory computer simulation to support collective decision-making: Potential and limits of stakeholder involvement," *Land use policy*, vol. 25, no. 4, pp. 498–509, 2008.

[49] A. Smajgl, A. P. N. House, and J. R. a Butler, "Implications of ecological data constraints for integrated policy and livelihoods modelling: An example from East Kalimantan, Indonesia," *Ecol. Modell.*, vol. 222, no. 3, pp. 888–896, 2011.

[50] E. J. B. Gaddis, H. H. Falk, C. Ginger, and A. Voinov, "Effectiveness of a participatory modeling effort to identify and advance community water resource goals in St. Albans, Vermont," *Environ. Model. Softw.*, vol. 25, no. 11, pp. 1428–1438, 2010.

[51] A. Voinov and E. J. B. Gaddis, "Lessons for successful participatory watershed modeling: A perspective from modeling practitioners," *Ecol. Modell.*, vol. 216, no. 2, pp. 197–207, 2008.

[52] V. Souchère, L. Millair, J. Echeverria, F. Bousquet, C. Le Page, and M. Etienne, "Co-constructing with stakeholders a role-playing game to initiate collective management of erosive runoff risks at the watershed scale," *Environ. Model. Softw.*, vol. 25, no. 11, pp. 1359–1370, 2010.

[53] W. D. Barreteau, "A role-playing game in irrigated system negotiation: between play and reality," *J. Artif. Soc. Soc. Simul.*, vol. 6, no. 3, 2003.

[54] C. Sausse, M. Le Bail, B. Lecroart, B. Remy, and A. Messéan, "How to manage the coexistence between genetically modified and conventional crops in grain and oilseed collection areas? Elaboration of scenarios using role playing games," *Land use policy*, vol. 30, no. 1, pp. 719–729, 2013.

[55] R. L. McCown, "Changing systems for supporting farmers' decisions: problems, paradigms, and prospects," *Agric. Syst.*, vol. 74, no. 1, pp. 179–220, 2002.

[56] C. Teh, *Introduction to Mathematical Modeling of Crop Growth: How the Equations are Derived and Assembled into a Computer Program*. Boca Raton: Brown Walker Press, 2006.

[57] J. W. Williams, R. C. Izaurralde, and E. M. Steglich, "Agricultural Policy/Environmental eXtender theoretical documentation," no. June, p. 131, 2008.

[58] S. L. Neitsch, J. G. Arnold, J. R. Kiniry, and J. R. Williams., "Soil and Water Assessment Tool User's Manual Version 2005.," p. 494, 2005.

[59] R. Allen, L. S. Pereira, D. Raes, and M. Smith, "Crop evapotranspiration: Guidelines for computing crop requirements," *Irrig. Drain. Pap. No. 56, FAO*, no. 56, p. 300, 1998.

[60] F. Blanco Flávio and V. Folegatti Marcos, "A new method for estimating the leaf area index of cucumber and tomato plants." Sociedade de Olericultura do Brasil, 2003.

[61] B. K. Goodwin, M. Marra, N. Piggott, and S. Mueller, "Is Yield Endogenous to Price? An Empirical Evaluation of Inter-and Intra-Seasonal Corn Yield Response," *North Carolina State Univ.*, 2012.

[62] "USDA ERS -Vegetables and Pulses Data.", 2015. [Online]. Available: http://www.ers.usda.gov/data-products/vegetables-and-pulses-data.aspx.

[63] "Global Weather Data for SWAT.", 2015. [Online]. Available: http://globalweather.tamu.edu/.

[64] M. Fadaee and M. a. M. Radzi, "Multi-objective optimization of a stand-alone hybrid renewable energy system by using evolutionary algorithms: A review," *Renew. Sustain. Energy Rev.*, vol. 16, no. 5, pp. 3364–3369, 2012.

[65] H. Ishibuchi, T. Nakashima, and T. Murata, "Three-objective genetic-based machine learning for linguistic rule estraction," *Inf. Sci. (Ny).*, vol. 136, pp. 109–133, 2001.

[66] R. Qi, L. Zhang, and D. Poole, "Solving asymmetric decision problems with influence diagrams," ... *Tenth Int. Conf.* ..., pp. 491–497, 1994.

[67] "FADN Public Database." [Online]. Available: http://ec.europa.eu/agriculture/rica/database/database_en.cfm.

[68] "EUROSTAT Database." [Online]. Available: http://epp.eurostat.ec.europa.eu/portal/page/portal/agriculture/data/database.

[69] C. Sahrbacher and M. Graubner, "Adaptation of the agent-based model AgriPoliS to 11 study regions in the enlarged European Union," *Plant Prod.*, vol. 2005, no. 10, 2005.

[70] S.-M. Chen and F.-M. Tsai, "Generating fuzzy rules from training instances for fuzzy classification systems," *Expert Syst. Appl.*, vol. 35, no. 3, pp. 611–621, 2008.

[71] J. G. Arbuckle, L. W. Morton, and J. Hobbs, "Farmer beliefs and concerns about climate change and attitudes toward adaptation and mitigation: Evidence from Iowa," *Clim. Change*, vol. 118, no. 3–4, pp. 551–563, 2013.

VITA

Name : Kasi Bharath Vegesana

Education:

1. PhD in Modeling, Simulation and Visualization Engineering at Old Dominion University, 2015.

2. M.E. in Computer Engineering at Old Dominion University. 2010

3. B.E. in Electronics & Communications Engineering in Andhra University, Andhra Pradesh, India, 2008

PUBLICATIONS

1. Kasi Bharath Vegesana, Frederic D. McKenzie, "Analysis of generic crop growth model for use in decision support systems for farmers". Proc. SPIE 8762, PIAGENG 2013: Intelligent Information, Control, and Communication Technology for Agricultural Engineering, 87620O (March 19, 2013); doi:10.1117/12.2019738.
2. Collins, A., Vegesana, K., Seiler, M., O'Shea, P., Hettiarachchi, P., & McKenzie, F. (2013). "Simulation and mathematical programming decision-making support for smallholder farming." Environment Systems and Decisions, 33(3), 427–439. http://doi.org/10.1007/s10669-013-9460-7.
3. Vegesana, K. B., & McKenzie, F. D. (2014). "A mathematical model for representing farmer decision processes". Proceedings of the 2014 Summer Simulation Multiconference. Monterey, California: Society for Computer Simulation International.
4. ***Best Paper Award*** Vegesana, K. B., & McKenzie, F. D. (2015). "Optimal crop selection and scheduling using a generic crop growth model and integer programming model". 14th International Conference on Modeling and Applied Simulation 2015, Bergeggi, Italy.
5. Vegesana, K. B., & McKenzie, F. D. (2015). "An Agent Based Model to Simulate Farmer Decision Process". Spring Simulation Multiconference. Alexandria, Virginia.