Journal of Agriculture and Environmental Sciences
March 2014, Vol. 3, No. 1, pp. 209-237
ISSN: 2334-2404 (Print), 2334-2412 (Online)
Copyright © The Author(s). 2014. All Rights Reserved.
American Research Institute for Policy Development
42 Monticello Street, New York, NY 12701, USA.
Phone: 1.347.757.4901 Website: www.aripd.org/jaes

Modeling and Optimization for Crop Portfolio Management Under Limited Irrigation Strategies

J. B. Bokhiria[1], K.R. Fowler[2] and E.W. Jenkins[3]

Abstract

Managing existing water resources has become critically important as overuse, in conjunction with extreme droughts, has placed aquifers in jeopardy. Our goal in this work is to develop a flexible modeling and optimization framework to aid farmers in selecting crop portfolios which offer the best outcomes, under sustainable water usage limitations, over specified time frames. The flexibility is emphasized through incorporation of multi-objective algorithms, allowing farmers to define "best" individually. We then demonstrate the modeling and optimization approach on a three-crop farm over a two year planning horizon and with a case study, the Pajaro Valley of California, known for berry farming. We consider a ten year planning horizon under the presence of increasing water and sale prices to demonstrate how the modeling tool can be used for predictive purposes.

Keywords: mathematical modeling, multi-objective optimization, sustainability, water resources, farming

1. Introduction

Managing existing water resources has become critically important as overuse, in conjunction with extreme droughts, has placed aquifers in critical conditions of overdraft.

[1] Clarkson University, Department of Mathematics, Potsdam, NY 13699, USA.
E-mail: bokhirjb@clarkson.edu)
[2] Clarkson University, Department of Mathematics, Potsdam, NY 13699, USA. Corresponding author.
E-mail: kfowler@clarkson.edu
[3] Clemson University, Department of Mathematical Sciences, Clemson, SC 29634-0975, USA.
E-mail: lea@clemson.edu

Recent news articles even point to assessments that the High Plains Aquifer, located beneath agriculture intensive states Kansas and Nebraska, is in danger of being completely depleted [**Error! Reference source not found.**]. Farmers in these regions have agreed to reduce their irrigation needs by 20% [**Error! Reference source not found.**], but the overall short- and long-term impacts on their crop management strategies has yet to be determined. Water shortages in other areas of the country also plague farmers [**Error! Reference source not found.**, **Error! Reference source not found.**, **Error! Reference source not found.**].

Local water management agencies have placed withdrawal restrictions on aquifers to allow them to recover[**Error! Reference source not found.**]. These restrictions have a significant impact on regional farmers, as water use graphics indicate farmers use a disproportionate amount of groundwater to irrigate their crops [**Error! Reference source not found.**, **Error! Reference source not found.**]. These shortages are not likely to ease soon, meaning farmers must develop strategies for operating under these guidelines.

Our goal in this work is to develop a flexible modeling and optimization framework to aid farmers in selecting crop portfolios which offer the best outcomes, under sustainable water usage limitations, over specified time frames. The flexibility is emphasized through incorporation of multi-objective algorithms, allowing farmers to define “best” individually. The three objectives we consider are to minimize water usage, maximize profit, and minimize deviation from the current demand, but the framework allows for any metric of farm performance related to the underlying model. These are often competing objectives; trade-off curves allow decision makers, i.e. farmers, a chance to consider multiple scenarios. Planting decisions are made by tracking the availability of farm plots and allowing an optimization algorithm to select the crops over a specified time horizon. Although detailed crop information is required, the modeling framework can be adapted to any set of crops with varying growing seasons. Crop planting and harvesting schedules are enforced through a set of linear constraints.

This paper is outlined as follows. In Section

2, we describe the process by which we formulate the problem. We enumerate and explain the different steps we believe are imperative in both capturing the variables under consideration by a farmer and improving the performance of the optimization algorithm.

In Section

3, we demonstrate the idea on an example problem using three different crops with properties leading to competing multi-objectives. In Section 4, we consider a case study on the Pajaro Valley of California.

This region is known for water intensive berry farming and has agencies actively seeking sustainable agricultural practices. The crops under consideration, lettuce, strawberries, raspberries, and blackberries, have varied growing seasons, with raspberries in particular having dynamic model parameters, demonstrating the adaptability of our approach. We consider a 10-year planning horizon under the presence of increasing water and sales prices, to show the predictive capability of the modeling tool. We end with conclusions and directions for future work.

2. Modeling Framework

Ultimately, our deliverable for this project is a software tool providing useful analysis for a farmer. Throughout our development, we have considered how a generic user (farmer) might best adapt our framework for their own needs. In this section, we describe our general strategy for simulating the operation of a farm over a specified time frame and offer insight into our modeling choices.

The model is established by viewing the optimizer as a virtual farmer. This viewpoint guides the definition of the decision variables, as the optimization tool needs to be able to make decisions on the same time frame as the farmer. Thus, as an initial step, the planting schedule for the crops under consideration must be determined. The information should include the number of crops for which decisions should be made, along with the planting and harvesting schedules for the crops. We generically denote this as a *preprocessing* step.

Planting decisions are dynamic and depend on the calendar month and the growing seasons of the different crops. The preprocessing step aims to identify the minimum number of decision variables required for simulation of a multi-year planning window. Minimizing the number of decision variables improves the performance of the optimization algorithm, as the size of the search space is significantly reduced. We choose the base number of decision variables on the planning horizon for the farmer and on the initial farm state. The planning horizon determines the number of times a planting decision must be made, while the initial farm state determines the types of crops available for planting.

We assume the crop portfolio available to the farmer will not deviate significantly from the existing crop state.

The optimization problem, acting as the farmer, needs to allow for decisions based on combinations of demand, profit, and water usage considerations. The problem must also consider constraints, possibly dynamic, to enforce planting and harvesting schedules. Our modeling strategy strongly connects the crop selection and optimization formulation through a multistep process requiring the preprocessing step, the constraints formulation, an objective function formulation, an optimization step, and an analysis, or postprocessing, step. A flowchart depicting the strategy used to construct the problem is given in Figure 1.

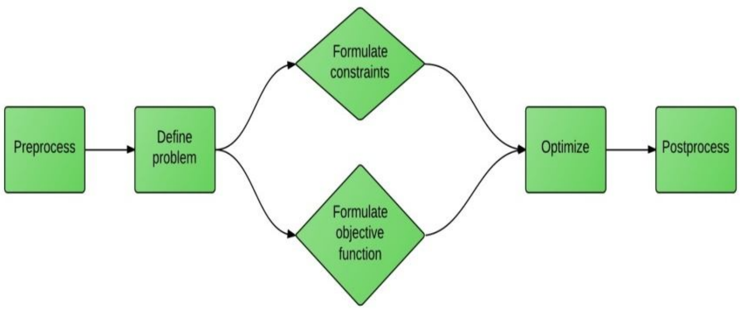

Figure 1: Flowchart outlining sequencing of process. Note the components of the problem formulation can be performed simultaneously.

2.1 Preprocessing

The preprocessing step is an analysis of the growing and harvesting schedules for the crops under consideration through the outline of a calendar that tracks when land becomes available and what is being harvested. This step leads to the determination of the decision variables for the optimization approach as well as the constraints which enforce the details of the growing season. The goals of this step are to minimize the number of decision variables (thereby reducing the size of the search space) and to account for the variety of planting scenarios possible under a given crop portfolio.

Let N_{crop} be the number of different crops under consideration. Each time a crop is harvested, the optimizer (i.e., the farmer) is allowed to plant something new. As part of the preprocessing analysis, we must determine, for each crop, the number of opportunities it could be planted, N_i. We then incorporate a decision variable to the planning model for that crop for each planting opportunity. In our case, we define as our decision variable the percentage of land allocated to crop *i* at each of the N_i planting opportunities. The total number of decision variables is $\sum_{i=1}^{N_{crop}} N_i$ and our vector of decision variables is denoted as

$$x = \left(x_1^1, x_2^1, \ldots, x_{N_1}^1, x_1^2, x_2^2, \ldots, x_{N_2}^2, \ldots, x_1^{N_{crop}}, x_2^{N_{crop}}, \ldots, x_{N_{N_{crop}}}^{N_{crop}}\right) \quad (1)$$

We also recommend the creation of a planning calendar during the preprocessing step for the purpose of visualizing decision-making and harvesting time points. This provides an overview of possible land allocation and guides the constraint formulations. We demonstrate this on a three-crop farm in the next section. Note this model formulation can account for multiyear crops, i.e., crops that remain in the ground during a planting opportunity while new crops of the same type are planted. In those cases, values for constraints or objective functions must be carefully computed, as the parameter values for the crop may change depending on the year.

2.2 Constraint Formulation

At each decision point, a linear inequality constraint must be enforced to ensure the total percentage of farm allocation is at most 100%. Land may be left fallow, which is beneficial in scenarios involving water restrictions or soil treatments used to increase nutrients (and thereby yields) in subsequent planting periods. Although we do not address that here, the incorporation of time dependent crop parameters could easily account for this.

Constraints are also used to model the planting and harvesting schedule using specific model parameters for each crop. For each crop, it is necessary to know the month they are planted and the number of months they remain in the ground.

For example, suppose x_1^1 percent of the farm is allocated to crop 1 at the first planting opportunity and more land becomes available for crop 1 at the next planting opportunity. More of crop 1 can be planted, even though the existing crop has not been harvested. We then simply enforce $x_2^1 \geq x_1^1$. The preprocessing step will have identified harvesting times for crops; thus, as crops are harvested, similar constraints can be enforced at each planting time if necessary. This provides flexibility in the modeling in the sense that crops with varying growing seasons, in terms of what month they are planted and how long they remain in the ground, can easily be incorporated. The dynamic planning, tied to the calendar year, is necessary to account for realistic decision-making. We illustrate this in the Case Study in Section

3 below.

2.3 Objective Functions

The objective functions must incorporate the information a farmer uses to make crop planning decisions. Examples of such information include profitability, limited changes in crop portfolios from year to year, meeting consumer demand on an annual basis, and minimizing the use of costly resources. The objective function can be defined as a single target, or it can include multiple, often competing, targets. Note that minimizing the use of costly resources does not necessarily increase profitability. Many crops have minimum resource requirements for successful harvest, so limiting the use of these resources often competes with profitability objectives.

We list the model parameters used in this work for each crop in Table 1. These parameters are used to describe the revenue generated by the crop, the deviation from the current demand, and the operational costs, although these definitions are adaptable to a variety of agricultural goals.

C_y^i	yield from one harvest (boxes/acre)
C_W^i	water usage (acre-ft/acre)
C_P^i	sales price (\$/box)
C_d^i	demand (% crop/year)
C_c^i	operational planting cost (\$/acre)

Table 1: Model parameters for crop type $(i = 1, 2, ..., N_{crop})$

Profit models can be as complex or as simple as needed once the farming model is in place. That is, once the decision variables are defined and the constraints are developed to describe the planting and harvesting schedules, a variety of metrics of interest to a farmer can be analyzed for a suite of feasible crop portfolios. If, for instance, A^i denotes the number of acres from crop *i*, a simple representation of profit could be

$$\text{Profit}^i = A^i (C_y^i C_P^i - P_W C_W^i - C_C^i) \qquad (2)$$

where P_W (\$/acre-feet) is the current price of water. Similarly, the amount of water used over the entire growing season can also be calculated for crop *i* using

$$\text{Water}^i = (C_W^i) A^i. \qquad (3)$$

To account for demand, we consider minimizing the deviation from a demand vector $D = (d^1, d^2, ..., d^{N_{crop}})$ containing the annual percentages of land allocated to each crop. Mathematically, this could be achieved using several metrics, including the absolute deviation or the maximum deviation. We use the $L2$-norm of the deviation for its smoothness properties. Although these are (mathematically) simple objective functions, we reiterate that the modeling framework allows the user to incorporate any agricultural metric of interest.

2.4 Optimization

The analysis of competing goals requires trade-off curves to provide information required for sound decision making. As the mathematical modeling described above was implemented in MATLAB, we determined the trade-off curves using their multi-objective genetic algorithm. A genetic algorithm is a gradient-free, global-search optimization algorithm.

A gradient-free algorithm was chosen to allow a user the ability to inorporate more sophisticated, mathematically challenging metrics without having to reconsider the optimization algorithm. Our profit and water objectives for this work are linear, and the demand-based objective is quadratic. However, the genetic algorithm can utilize any form of objective function, since it is designed to work only with function values and not explicitly with the function. Related works which consider multiobjective linear programming in the context of crop rotation and environmental farm planning include works by El-Nazer et al.[**Error! Reference source not found.**], Beneke et al.[**Error! Reference source not found.**], Sahoo et al.[**Error! Reference source not found.**], and Annetts et al.[**Error! Reference source not found.**].

Genetic algorithms (GAs) are part of a larger class of evolutionary algorithms and are classified as population based, global search heuristic methods [**Error! Reference source not found.**]. Genetic algorithms are based on biological processes such as survival of the fittest, natural selection, inheritance, mutation, and reproduction. Design points are coded as “individuals" or “chromosomes", typically as binary strings, in a population. Through the above biological processes, the population evolves through a user specified number of generations towards a smaller fitness value. We can define the following simple GA below;

- Require Population size n_p, Number of Generations n_g
- Generate initial population, determine fitness, and rank : $P_1 = p_1, ..., p_{n_p}$
- For $k = 1, ..., n_g$
    1. $P_{k+1} = \text{select}\left(P_k\right)$
    2. $P_{k+1} = \text{crossover}\left(P_{k+1}\right)$
    3. $P_{k+1} = \text{mutate}\left(P_{k+1}\right)$
    4. Determine fitness for P_{k+1}

During the selection phase, better points with smaller function values are arranged randomly to form a mating pool on which further operations are performed. Crossover attempts to exchange information between two design points to produce a new point that preserves the best features of both 'parent points'. Mutation is used to prevent the algorithm from terminating prematurely to a suboptimal point and is used as a means to explore the design space. Termination of the algorithm is based on a prescribed number of generations or when the highest ranked individual's fitness has reached a plateau. Genetic algorithms are often criticized for their computational complexity and dependence on optimization parameter settings, which are not known a priori. However, if the user is willing to exhaust a large number of function evaluations, the GA can help gain insight into the design space and locate initial points for fast, local single search methods. When analyzing several metrics, one single solution likely doesn't exist and it is more meaningful to consider trade-offs. A multi-objective GA, which we use here, evolves to generate a set of Pareto optimal solutions. In this set of points, a solution is considered non-dominated, and a member of the Pareto curve, if improving one objective function results in degrading another.

3. Three Crop Farm Example

We proceed by describing the modeling and optimization formulations on a generic farming scenario. For this demonstration, we consider three crops, *A*,*B*, and *C*, over a two year time frame. We let crops *A* and *B* have a four month growing season and crop *C* have an eight month growing season. We assume any of these crops can be planted in any month.

The preprocessing step requires we chart the decision points for the farmer over the growing period. Recall decision points occur when land becomes available, meaning the farmer must make an allocation decision for the available plot. For simplicity, we assume we begin with an open farm; thus, initially we can plant $x = (x_1^A, x_1^B, x_1^C)$ where the superscript identifies the crop and the subscript indicates the cardinality of the planting decision. Land next becomes available in four months, at which point any three of these can be planted again. Keep in mind, however, the acreage initially dedicated to crop *C* will still be occupied, leading to a constraint. The preprocessing step ultimately leads to the planting opportunities shown in Table 2, yielding 18 decision variables.

We also demonstrate in Table 2 the use of the decision variables, which are the total percentages of the farm allocated to a given crop, in calculating quantities for harvested and planted crops.

Year 1			
4Month Period	1	2	3
Crop *A*	x_1^A	x_2^A	x_3^A
Crop *B*	x_1^B	x_2^B	x_3^B
Crop *C*	x_1^C	x_2^C	x_3^C
Harvested		x_1^A, x_1^B	x_2^A, x_2^B, x_1^C
Planted		x_2^A, x_2^B $x_2^C - x_1^C$	x_3^A, x_3^B $x_3^C - (x_2^C - x_1^C)$
Year 2			
4 Month Period	4	5	6
Crop *A*	x_4^A	x_5^A	x_6^A
Crop *B*	x_4^B	x_5^B	x_6^B
Crop *C*	x_4^C	x_5^C	x_6^C
Harvested	x_3^A, x_3^B , $x_2^C - x_1^C$	x_4^A, x_4^B $x_3^C - (x_2^C - x_1^C)$	$x_4^A, x_4^B,$ $x_4^C - (x_3^C - (x_2^C - x_1^C))$
Planted	x_4^A, x_4^B $x_4^C - (x_3^C - (x_2^C - x_1^C))$	x_5^A, x_5^B $x_5^C - (x_4^C - ...$ $(x_3^C - (x_2^C - x_1^C)))$	x_6^A, x_6^B $x_6^C - (x_5^C - ...$ $(x_4^C - (x_3^C - ...$ $(x_2^C - x_1^C))))$

Table 2: Example planting schedule

The definition of the decision variables x_i^A, x_i^B, x_i^C (i=1,…,6) also means we must require, for each planting period,

$$x_i^A + x_i^B + x_i^C \leq 100 \tag{4}$$

Additional constraints are used to enforce harvesting schedules. Note that at period 2, since the amount of crop *C* that was initially planted in still in the ground, $x_2^C \geq x_1^C$. At period 3, the amount of crop *C* planted in period 1 will be harvested, leaving $x_2^C - x_1^C$ as the percent of acreage dedicated to crop *C* prior to any planting in period 3. As x_3^C represents the percent of acreage dedicated to crop *C* in period 3, we require $x_3^C \geq x_2^C - x_1^C$. As crop *C* is the only crop that can remain in the ground over successive planting period, we have the following five constraints associated with crop *C*:

$$\begin{aligned} x_2^C &\geq x_1^C \\ x_3^C &\geq x_2^C - x_1^C \\ x_4^C &\geq x_3^C - (x_2^C - x_1^C) \\ x_5^C &\geq x_4^C - (x_3^C - (x_2^C - x_1^C)) \end{aligned} \tag{5}$$

We will assume the three crops have varying properties to demonstrate the trade-off analysis for decision making. We will consider crop*A* to be high in demand and profitable, but water intensive. Crop *C*will have the lowest demand and lowest profit, but use the least amount of water. Crop *B*will be in the middle for all three properties. The model parameters describing the attributes are shown in Table 3.

Parameters	*A*	*B*	*C*
Water Usage(W_i) acre-ft / acre / year	3	2	1
Yield(Y_i) “boxes" / acre	1000	1000	4000
Sales Price(P_i) Box	\$4.00	\$3.00	\$2.00

Table 3: Model parameters for simplified three crop model.

Water usage is a straight forward calculation based on crops currently in the ground. We adjust the annual water requirement appropriately, as our decision variables represent the percentage of crops over a 4 month time step. For crop A, we have

$$\text{Water}^A = \sum_{i=1}^{6} A^A (C_W^A)/3 \tag{6}$$

where A^A denotes the total acreage of crop A over the two years. We scale the price of water by 1/3 as the price is quoted on an annual basis. We have similar computations for water prices for crops *B* and *C*.

For this example, we consider a simple profit model which is dynamically updated when crops are harvested. To compute the profit for crop A, we note it is harvested every period, giving

$$\text{Profit}^A = \sum_{i=1}^{6} (C_y^A C_P^A - P_W C_W^A / 3) A_i^A \tag{7}$$

We would have a similar profit model for crop *B*. However, for crop *C*, we must account for what is being harvested since plantings may overlap.

For this two year model, this gives a profit calculation for crop *C* as

$$\begin{aligned} \text{Profit}^C = (C_P^C C_P^C - P_W C_W^C)[& \underbrace{(x_2^C - x_1^C)}_{\text{harvested in year 1}} + \\ & (x_3^C - (x_2^C - x_1^C)) + (x_4^C - (x_3^C - (x_2^C - x_1^C))) \\ & \underbrace{+ (x_5^C - (x_4^C - (x_3^C - (x_2^C - x_1^C))))]}_{\text{harvested in year 2}} . \end{aligned} \tag{8}$$

This information can be extracted from the preprocessing step (see Table 2) and simplified prior to implementation. The entire expression was shown here to demonstrate the concept.

Our last objective accounts for the current demand for each crop. We assume to meet the current demand, this farm typically allocates 50% of its acreage to crop *A*, 35% to crop *B*, and 15% to crop *C*. For this example, we calculated the average yield of each crop per year and minimized the L_2-norm of the deviation from the demand for both years. For year one, if Y_1^i is the average yield of crop *i*, we measure the deviation from demand can be measured as

$$\text{Demand}_1 = \sqrt{(Y_1^A - 50)^2 + (Y_1^B - 35)^2 + (Y_1^C - 15)^2}. \qquad (9)$$

3.1 Numerical Results

We ran the multi-objective genetic algorithm from the MATLAB Global Optimization Toolbox with the default algorithmic parameters, except we seeded the initial population. The GA typical starts with a random population of design points, which in this case led to poor performance, often times failing to find feasible designs at all. In practice, it is common to use expert knowledge to provide reasonable initial iterates for the optimization algorithm. Thus, we seeded the initial population of the GA with five farm scenarios that satisfied the linear constraint that the total acreage had to be at most 100%. These can be found in Table 4 below. Each row is the percentage of each crop which was set to be uniform over the entire planting horizon. The GA has random aspects to its search algorithm and thus can give different results for each optimization run. We performed multiple optimizations and observed the Pareto fronts had similar shapes.

We show some of the representative trade-off curves in Figures 2 - 4. Here, each star in the Pareto set corresponds to a design point and the horizontal and vertical axes are values of the competing objectives. We see that, based on the models used here, the profit and water usage objectives are not necessarily competing, which makes sense since the cost of water is included in the profit model. The other two sets of objectives clearly are competing since moving towards a better value of one objective results in degrading the other.

As seen in Figure 3, profit is maximized for lower values of water usage. This indicates the cost of water significantly impacts any monetary gains realized by sales of the crop. The trade-off in Figure 2 shows the demand is associated with crops requiring large amounts of water.

If a farmer wished to lower his water usage, he would be required to plant crops not in demand. This, in turn, affects the prices of the crops, which has not been directly incorporated into this model. Taken together, the graphs show the difficult decisions farmers must manage to maintain their livelihood. In addition, more rigorous analysis of the connections between profitability and demand should be incorporated to better understand the effects of changing crop portfolios on return of investment.

Initial Seed	% crop *A*	% crop *B*	% crop *C*
1	40	40	20
2	10	10	80
3	33	33	33
4	25	25	50
5	25	50	25

Table 4: Initial design points seeded in GA for three-crop farm, set over the entire planting horizon

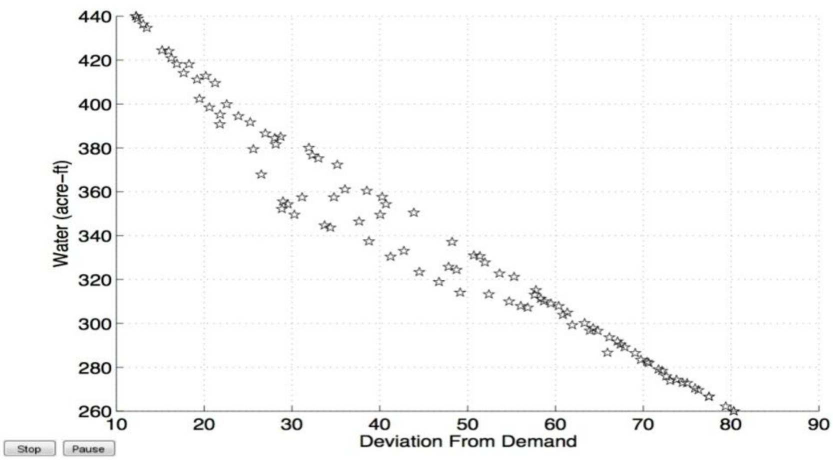

Figure 2: Trade-off curve for objectives minimizing water usage and deviation from demand. Note the current demand requires the most water

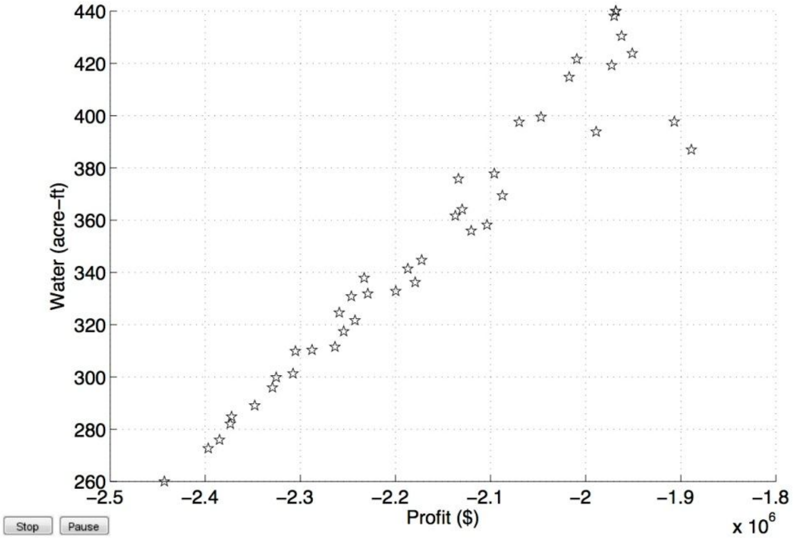

Figure 3: Trade-off curves for objectives minimizing water usage and maximizing profit. The negative values on the profit axis indicate we actually minimize the negative of profit. Note the maximum profit corresponds to minimal water usage.

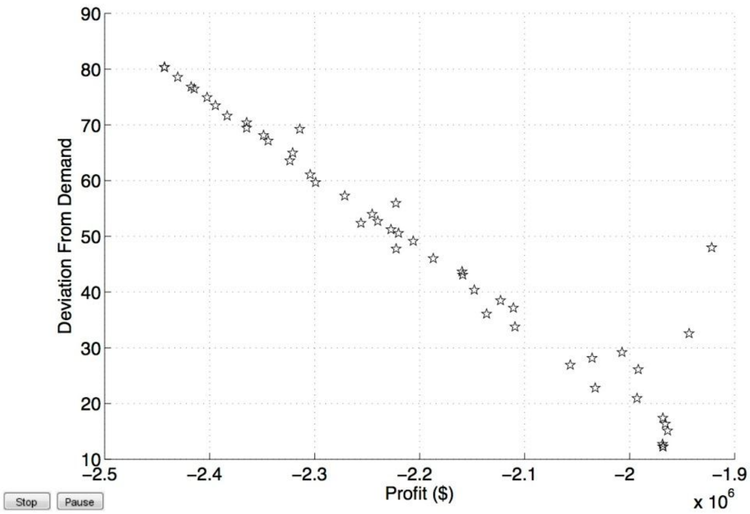

Figure 4: Trade-off curve for objectives maximizing profit and minimizing deviation from demand. The most profitable strategy requires the largest deviation from demand

4. Case Study: Pajaro Valley, CA

We now apply our modeling and optimization framework to the Pajaro Valley region of central California, which was developed for agriculture in the late 1800s. In the decades since its establishment, continued stresses on the groundwater basin from significant pumping for primarily (84%) agricultural use have resulted in the basin being in a "critical condition of overdraft" [**Error! Reference source not found.**, **Error! Reference source not found.**, **Error! Reference source not found.**]. This designation has caused the entire region to operate under water use restrictions.

Stakeholders in the Pajaro Valley, including farmers, land owners, environmental agencies, and residents, have been working on a multifaceted solution to the overdraft problem. The community holds quarterly dialogues, along with monthly subgroup meetings, to discuss possible long-term, and short-term, resolutions to the problem.

A task force involving the Pajaro Valley Water Management Agency (PVWMA) has also been established, and members of the task force asked for help on resolving the underlying optimization problem, which seeks to balance the competing interests of the different parties.

Urban water use is estimated at 10,000 acre-ft/yr [**Error! Reference source not found.**]. According to the same USGS report, the PVWMA service area encompasses about 70,000 acres, 40% of which is used for agriculture. The sustainable water yield of the basin is estimated between 24,000 acre-ft/yr and 48,000 acre-ft/yr, the higher yield being possible if pumping at the coast is eliminated and replaced by water from a different source (cf. [**Error! Reference source not found.**]). Thus, after taking into account the urban water use, the sustainable amount of water available is between 0.5 acre-ft/yr and 1.36 acre-ft/year per acre of agricultural land.

Berry farmers in the Pajaro Valley are particularly affected, as they must maintain profitability under reduced irrigation strategies. California grows more berries than any other region in the world [**Error! Reference source not found.**], providing roughly 87% of the country's strawberries in 2007. Other berry crops produced in the region include raspberries, blackberries, and blueberries. Strawberries are typically a high water usage crop, requiring 2.67 acre/ft of water per year, while raspberries and blackberries use only 2 acre/ft (or less) of water per year (see Table 6).

4.1 Preprocessing

Planting strategies are modeled over a ten year period using historical and extrapolated data for the sales prices of the different berries and the price of water. A planting year runs from September to the following August. We consider the percentage of each crop that is in the ground at a possible planting time for that crop, denoted by *S*, *R*, *B*, *L*, and *C* for strawberries, raspberries, blackberries, lettuce, and cover crops (we drop the x_i^L notation for simplicity). As previously, we identify the planting period for that crop with a subscript. For example, L_5 is the total percentage of lettuce in the ground at the 5th planting opportunity for lettuce. Thus, L_5 accounts for both unharvested lettuce and newly planted lettuce. We consider 2 month "time steps" as all crops have even life cycles. The growing periods for each crop are described below.

- Strawberries are a 14 month crop. Land is assigned to strawberries in September and the plot is occupied until the following November.
- Blackberries are a 60 month crop, planted in September.
- Raspberries are a 24 month year crop, again occupying land beginning in September. They yield twice during this period.
- Lettuce is a four month crop (including preparation), which can be planted at any month.

Over a 10 year time horizon, all berries can be planted each September, giving 10 variables for each berry. Lettuce is tracked more frequently. In year 1, everything is planted in September and only lettuce can be removed (in 4 month intervals). In year 2, strawberries will come out of the ground in November and the associated land is released for planting. Thus, after year 2, lettuce can be planted every 2 months, starting in November. For example, in January of year 2, the lettuce planted in September of year 2 would come out but lettuce planted in November would remain. New lettuce may also be planted at this decision point. We summarize the planting possibilities through the first three planting periods for berries in Table 5. The notation in the first row gives "month/year" , with month 9 referring to September, month 11 referring to November, etc. As noted earlier, the subscripts on the crop designation refer to the successive decision points for that particular crop.

Month	9/1	1/2	5/2	9/2	11/2	1/3	3/3	5/3	7/3	9/3	…
Strawberries	S_1			S_2						S_3	…
Raspberries	R_1			R_2						R_3	…
Blackberries	B_1			B_2						B_3	…
Lettuce	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	…

Table 5: Preprocessing information for the Pajaro Valley example. We outline the calendar planting schedule for each of the crops.

4.2 Constraints

We require each variable be nonnegative and the total percentage of land (including fallow land) sum to 100. We define the amount of fallow land as the percentage of land left after accounting for all the crops in the ground (i.e. 100 minus the total percentage of acreage allocated to "for profit" crops).

These requirements impose the following inequality constraint in September of the first year and the inequality constraintsin September of years 2 through 10.

$$S_1 + R_1 + B_1 + L_1 \le 100$$

$$S_j + R_j + B_j + L_{(4+6(j-2))} \le 100, \quad j = 2,\ldots,10 \quad (10)$$

As strawberries are a 14-month crop, land allocated to strawberries in November of year 2 through September of year 3 is $S_2 - S_1$. S_2 tracks strawberries planted in September of year 2 along with those strawberries planted in September of year 1, as those berries have not been removed from the ground. Similar adjustments are needed for the following years. Therefore, the constraints guiding the amount of land available in years 2 through 10 for planting lettuce are given by

$$L_j \le 100 - (S_k - S_{k-1}) - R_k - B_k, \; j = 5,\ldots,60, \; j \neq 10,16,\ldots,52, \; k = 2,\ldots 10 \quad (11)$$

The amount of land available for planting lettuce in year 1 is slightly different, as only S_1 is needed to account for acreage dedicated to strawberries and potential lettuce planting only occurs every 4 months.

Thus, in year 1, the constraint associated with planting lettuce is given by

$$L_j \leq 100 - S_1 - R_1 - B_1, \quad j = 2, 3. \tag{12}$$

Starting September, year 2, additional constraints on lettuce are needed as decision points for lettuce begin to overlap. Decision points now occur every two months, meaning the percentage of lettuce in the ground at these decision points includes both newly planted lettuce and lettuce planted at the previous decision point.

For example, the 5^{th} decision point occurs in November of year 2, two months after the 4^{th} decision point in September of year 2. Lettuce may have been planted in September, but it must stay in the ground 4 months. Thus, we must have

$$L_4 \leq L_5; \tag{13}$$

that is, we cannot yet harvest the lettuce planted in September. As this is year 2, we note that the lettuce in the ground in September is newly planted, as any lettuce planted at the previous decision point in year 1 would have completed its 4-month harvesting window.

However, at subsequent decision points, the decision variables for lettuce must telescope appropriately to define minimum values for percentages of acreage dedicated to lettuce. At decision point six, the amount of lettuce in the middle of its four month planting requirement is $L_5 - L_4$, the amount of new lettuce planted at decision point 5. Similar lower bounds are defined at future decision points, giving the set of constraints

$$\begin{aligned} &L_6 \geq L_5 - L_4; \\ &L_7 \geq L_6 - (L_5 - L_4); \\ &L_8 \geq L_7 - \big(L_6 - (L_5 - L_4)\big); \\ &\vdots \end{aligned} \tag{14}$$

The inequality constraints for raspberries follow the same trend, as R_2 accounts for raspberries planted in year 1 as well as raspberries planted in year 2. The amount of new raspberries in September of year 2 is thus $R_2 - R_1$, meaning the raspberries in the ground in September of year 3 must be at least this amount.

New raspberries in September of year 3 are R_3 less what is put in the ground in September of year 2, or $R_2 - R_1$. This leads to the set of inequality constraints associated with planting raspberries described using

$$\begin{aligned} &R_2 \geq R_1; \\ &R_3 \geq R_2 - R_1; \\ &R_4 \geq R_3 - (R_2 - R_1); \\ &R_5 \geq R_4 - \big(R_3 - (R_2 - R_1)\big); \\ &\vdots \end{aligned} \tag{15}$$

As with raspberries, strawberries from the previous September are in the ground during the current September. The inequality constraints imposed on strawberries thus match the form for raspberries. Thus, for strawberries, we have

$$\begin{aligned} &S_2 \geq S_1; \\ &S_3 \geq S_2 - S_1; \\ &S_4 \geq S_3 - (S_2 - S_1); \\ &S_5 \geq S_4 - \big(S_3 - (S_2 - S_1)\big); \\ &\vdots \end{aligned} \tag{16}$$

Overall we have a total of 144 linear inequality constraints.

4.3 Objective Functions

The specific operation cost, water usage, yield, and sale price for each crop is given in Table 6. It is important to note the values for lettuce are per crop (i.e. 4 months), while the remaining parameters are annual. Also, note the model parameters for raspberries change the second year they are in the ground, emphasizing the need to accurately track the newly planted versus the existing raspberries. Applying an averaged value to all raspberries currently planted will give infeasible solutions.

Parameters	Strawberry	Blackberry	Raspberry year 1	Raspberry year 2	Lettuce (per crop)	Cover
Operational $(S_i + L_i)$ Acre	$22000	$22500	$27000	$12000	$2200	$1850
Water Usage(W_i) acre-ft/acre	2.67	2	2	1.5	1	0
Yield(Y_i) "boxes"/acre	7000	3500	4800	5000	1	0

Table 6: Model parameters for the Pajaro Valley farm example. Note the differences in model parameters for different harvest years for raspberries.

Table 7 contains the changing prices of water and crops over a ten year period. These were based on market values from 2003-2008 and extrapolated with increasing water prices. This scenario could, for example, guide farmers to make decisions or analyze profit forecasts under the changing drought conditions in that region.

Year	1	2	3	4	5	6	7	8	9	10
Water ($/acre-ft)	120	120	160	160	171	175	177	179	185	190
Strawberry ($)	5.50	6.15	6.15	6.20	7.75	6.50	7.75	7.00	6.65	10.00
Raspberry ($)	5.40	5.30	5.50	5.65	11.50	10.75	10.50	10.25	10.00	10.00
Blackberry ($)	4.80	4.80	4.90	4.80	13.50	12.25	10.75	10.00	10.25	12.00

Table 7: Water cost and sales prices over 10-year horizon.

For this problem, the objective function associated with profit is the sum of the profit obtained for each crop, as calculated in the previous example. We need to calculate the acreage of each crop based on the variables *S*, *R*, *B*, and *L* , which will not necessarily give the complete picture of what is in the ground. These values represent the total acreage of a crop at specified planting times, but the fallow (unplanted) land changes in November when strawberries from the previous planting period are removed. This adjustment is not made in the current model.

Using this information, profit is accumulated by crop and year. For example, the total profit associated with strawberries P_S is given by

$$
\begin{aligned}
P_S &= c_S S_1 \text{(year 1)} \\
&+ \frac{c_S}{6} S_1 + c_S(S_2 S_1)\text{(year 2)} \\
&+ \frac{c_S}{6}(S_2 S_1) + c_S\big(S_3(S_2 S_1)\big)\text{(year 3)} \\
&+ \frac{c_S}{6}\big(S_3(S_2 S_1)\big) + c_S\Big(S_4\big(S_3(S_2 S_1)\big)\Big)\text{ (year 4)} \\
&+ \frac{c_S}{6}\Big(S_4\big(S_3(S_2 S_1)\big)\Big) + c_S\Big(S_5\Big(S_4\big(S_3(S_2 S_1)\big)\Big)\Big)\text{ (year 5).}
\end{aligned} \tag{17}
$$

where c_S represents the annual per acre profit for strawberries, calculated as $c_S = \text{price per box} * 7000 - 22000$.

The annual profit is adjusted starting in year 2 in Equation **Error! Reference source not found.** to account for the two months from September to November when strawberries planted during the previous year are still in the ground (and assumed to be still producing fruit).

The profit for raspberries differs depending on the age of the crop. The operational cost for the second production year of raspberries is lower than the first production year, and the yield changes as well.

$$
\begin{aligned}
p_r &= c_{r1}R_1 \text{ (year 1)} \\
&+ c_{r2}R_1 + c_{r1}(R_2 - R_1)\text{(year 2)} \\
&+ c_{r2}(R_2 - R_1) + c_{r1}(R_3 - (R_2 - R_1))\text{(year 3)} \\
&+ c_{r2}(R_3 - (R_2 - R_1)) + c_{r1}(R_4 - (R_3 - (R_2 - R_1)))\text{(year 4)} \\
&+ c_2(R_4 - (R_3 - (R_2 - R_1))) + c_{r1}(R_5 - (R_4 - (R_3 - (R_2 - R_1))))\text{(year 5).}
\end{aligned} \tag{18}
$$

Note we have

c_{r1}=price per box*3500 - 27000 and c_{r2}=price per box*4800–12000.

Profit calculations for the remaining crops are more straightforward. Fallow land has a negative contribution to profit, as nothing is produced but rent must still be paid.

Our second objective, to minimize the amount of irrigation, is calculated as in the previous example. Our final objective uses the current berry production for the Pajaro Valley as the value for demand. A typical planting year includes 30% strawberries, 10% raspberries, 10% blackberries, and 40% cover crops (i.e., lettuce) with the remaining 10% fallow. We minimize the deviation from this demand in the standard L_2-norm at each two month decision point.

4.4 Numerical Results

Although we could begin our farming model with any intitial farm configuration, we present results as if we were begining our planting season in September with an empty 100 acre farm. The numerical results highlight the decision difficulties currently faced by farmers operating under limited irrigation strategies. Using the more realistic crop parameters generates trade-off curves with much less linearity than the simplified example. However, we see the same general trends in Figures 5 - 7. As seen in Figure 5, current production in the valley actually requires significant irrigation. Reducing irrigation requirements will mean the farmers must alter their current portfolios.

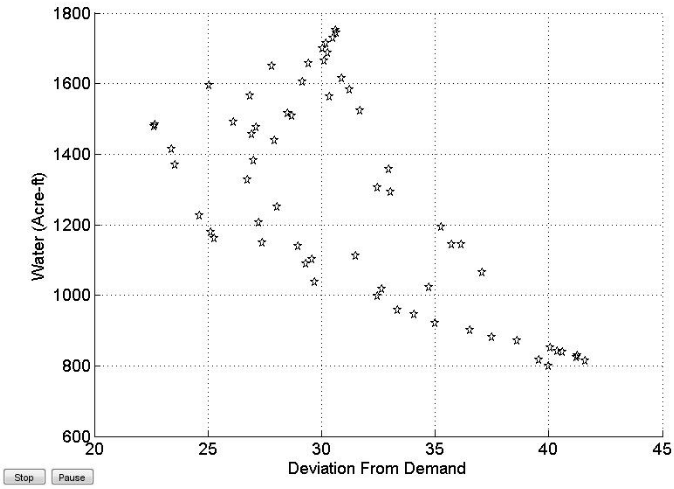

Figure 5: Trade-off curve for objectives minimizing water and deviation from demand. Note the current crop portfolio is the most water intensive. Also, the feasible solutions provide a variety of choices to balance the competing objectives.

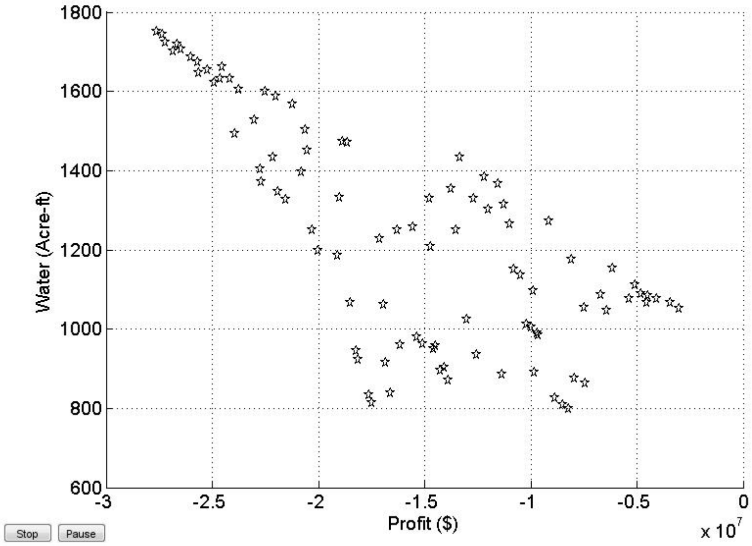

Figure 6: Trade-off curve for objectives minimizing water usage and maximizing profit. The more profitable feasible solutions are the most water intensive. With these model parameters, the less irrigation required,the less profitable the portfolio.

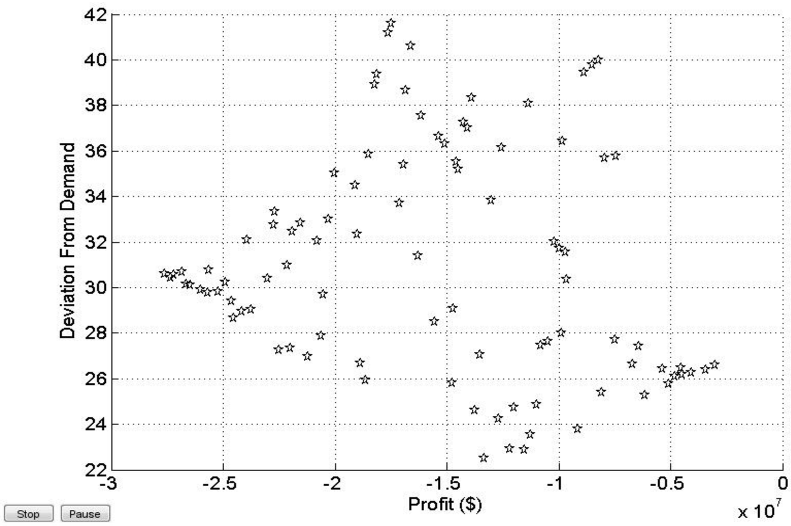

Figure 7: Trade-off curve for objectives maximizing profit and minimizing deviation from demand. No obvious trend is noticeable in these results, perhaps indicated more clearly defined objectives should be considered. The feasible solutions do not provide clear analysis to improve the decision-making capability of the farmer.

Table 8 contains the values of the three objectives for the best point found for each, given in bold. Furthermore, Figures 8, 9, and 10 show the amount of each crop planted for each objective. Note that the water usage solution requires most of the region remains fallow. Previous studies [**Error! Reference source not found.**] incorporated the sustainable yield as a hard constraint and resulted in similar solutions with just under 25% of the land allocated to raspberries, so these findings are consistent but allow growers to consider farming strategies that are less restrictive.

Water (acre-ft)	Profit ($)	Demand
7.989e+02	8.2456e+06	40%
1.7523e+03	**2.7638e+07**	30%
1.4781e+03	1.3627e+07	**22%**

Table 8: Best function values found for each objective. This information indicates a farmer loses an order of magnitude in profit for an order of magnitude reduction in water usage. Current crop portfolios are profitable yet water intensive.

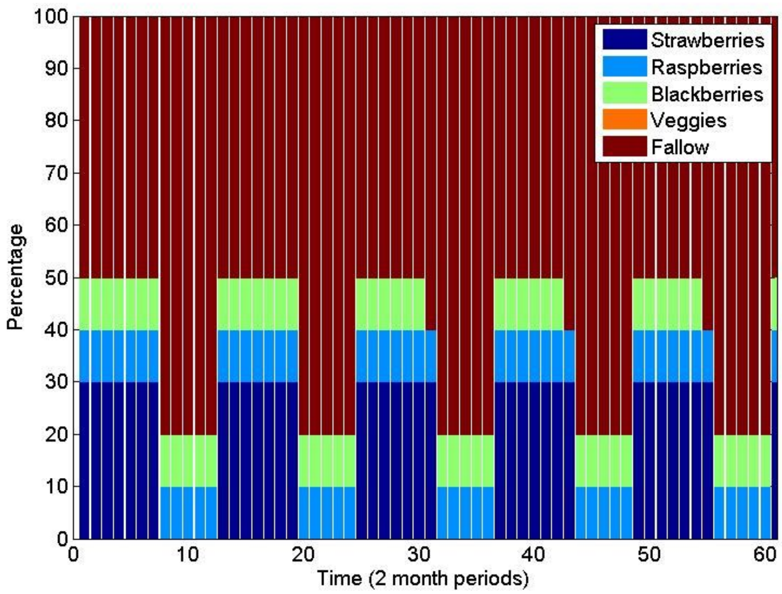

Figure 8: Crop portfolio for minimal water usage. Note the prevalence of unfarmed (i.e., fallow) land. Raspberries and blackberries remain in production throughout the planting cycle. However, once strawberries are planted in September of one year, they are not replanted until September two years removed.

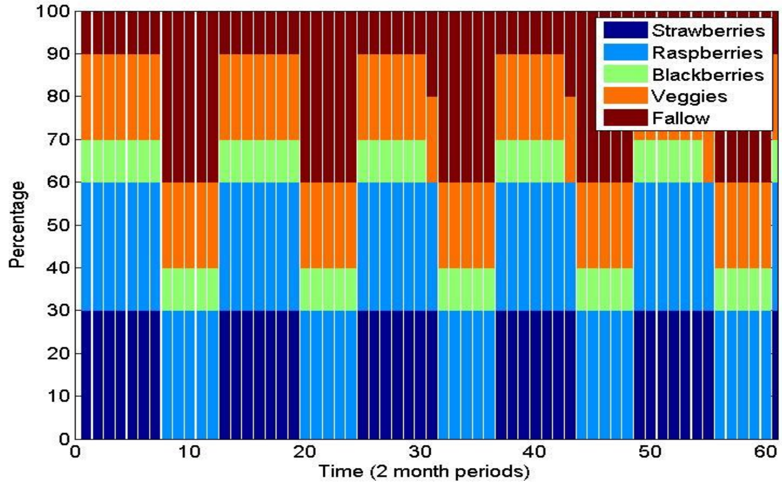

Figure 9: Crop portfolio for maximized profit. Note vegetables now appear on the farm, as well as increased levels of both strawberries and raspberries.

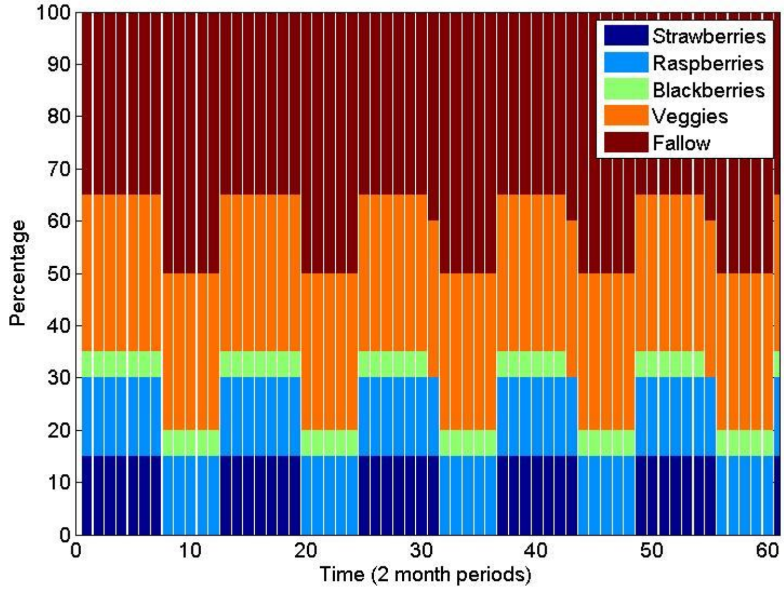

Figure 10: Crop portfolio for minimizing deviation from demand. The optimization algorithm is able to keep a consistent farm allocation throughout the planting cycle.

5. Conclusions

We have used a relatively simple model of a farming operation coupled with multi-objective optimization algorithms to aid a farmer in crop portfolio selection while operating under a variety of constraints. The use of a genetic algorithm provides a derivative-free, global search of the design space. The preprocessing step is an essential part of the optimization strategy, as effective management of the design space significantly improves the performance of the optimizer. Significant effort should be spent in this stage to ensure the optimization algorithm can locate potentially optimal points in a short time frame. The objective functions should be carefully formulated, incorporating any decisions a farmer is likely to make over the time horizon. We recommend the use of multiple objectives to best capture the decision making process of the farmer.

The results from the more specific case study for the Pajaro Valley reiterate the benefits of our modeling and optimization strategy. The framework put in place was able to handle the varied parameters associated with four distinct crops, each with different planting rules and yields. One of the crops, raspberries, even had different parameters associated with distinct production cycles.

Future work will address model development and optimization using more complicated realizations of a farming process and water delivery system. The basic strategy outlined in this paper, however, will remain the same. The flexibility of the overall strategy allows for a “plug-and-play" environment, giving users the capability to emphasize selected aspects of a given farming operation.

6. Acknowledgements

This work was sponsored, in part, by an American Institute of Mathematics SQuaRE grant. The authors are grateful to Dr. Estelle Basor and farmers in the Pajaro Valley region of California for their support and many useful discussions.

References

Pajaro Valley Water Management Agency. http://www.pvwma.dst.ca.us/.

J. E. Annetts and E. Audsley. Multiple objective linear programming for environmental farm planning. *J Oper. Res. Soc.*, 53(9):933–943, 2002.

N.L. Barber. Summary of estimated water use in the United States in 2005. Technical Report Fact Sheet 2009-3098, U.S. Department of the Interior, U.S. Geological Survey, October 2009. 2 p.

R.R. Beneke and R. Winterboer. *Linear programming. Applications to agriculture*. Aedos, 1984.

D. Charles. Kansas farmers commit to taking less water from the ground. http://www.npr.org/blogs/thesalt/2013/10/22/230702453/in-kansas, October 2013. [Online; posted 22-October-2013].

J.C. Chrispell, K.R. Fowler, E.W. Jenkins, T. Sendova, and S.E. Howington. Meeting a sustainable water yield for agricultural practices through crop rotation and catchment basin design. In *Proceedings of the 2012 SC Water Resources Conference*, Columbia, SC, 2012.

R.L. Cook. Three part series: Fundamental forces affecting growers and marketers. Part 2: The dynamic berry industry. *Blueprints*, Apr/May/June 2012.

T. El Nazer and B.A. McCarl. The choice of crop rotation: modeling approach. *Amer. J. Agr. Econ.*, 68(1):127–136, 1986.

D.E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, 1st edition, 1989.

R.T. Hanson. Geohydrologic framework of recharge and seawater intrusion in the Pajaro. Technical Report Water-Resources Investigations Report 03-4096, U.S. Department of the Interior, U.S. Geological Survey, Sacramento, CA, 2003.

R.T. Hanson. Geohydrologic framework of recharge and seawater intrusion in the Pajaro Valley, Santa Cruz, and Monterey Counties, California. Water-Resources Iinvestigations Report 03-4096, U.S. Department of the Interior, U.S. Geological Survey, 2003. 100 p.

M.J Johnson. Groundwater in North Monterey count, California, 1980. Technical Report Water-Resources Investigations Report 83-4023, U.S. Department of the Interior, U.S. Geological Survey, Sacramento, CA, 1983.

M.J. Johnson, C.J. Londquist, J. Laudon, and H.T. MItten. Geohydrology and mathematical simulation of the Pajaro Valley aquifer system, Santa Cruz, and Monterey counties, California. Technical Report Water-Resources Investigations Report 87-4281, U.S. Department of the Interior, U.S. Geological Survey, Sacramento, CA, 1988.

M.A. Maupin and N.L. Barber. Estimated withdrawals from principal aquifers in the United States, 2000. Technical report, U.S. Department of the Interior, U.S. Geological Survey, July 2005. 46 p.

F. Morris. Western Kansas farmers face dwindling water supply. http://hereandnow.wbur.org/2013/08/19/kansas-farmers-water, August 2013. [Online; posted 19-August-2013].

California Department of Natural Resources. California's Groundwater. Bulletin 118 - Update 2003. http://www.water.ca.gov/groundwater/bulletin118/bulletin118update2003.cfm.

B. Sahoo, A.K. Lohani, and R.K. Sahu. Fuzzy multiobjective and linear programming based management models for optimal land-water-crop system planning. *Water Resour. Manag.*, 20(6):931–948, 2006.