Agriculture and Food

http://www.aimspress.com/journal/agriculture

AIMS Agriculture and Food, 1(2): 208–226
DOI: 10.3934/agrfood.2016.2.208
Received: 31 January 2016
Accepted: 18 May 2016
Published: 23 May 2016

Research article

Development and Use of Mathematical Models and Software Frameworks for Integrated Analysis of Agricultural Systems and Associated Water Use Impacts

K.R. Fowler [1] E.W. Jenkins [2,*] M. Parno [3] J.C. Chrispell [4] A.I. Colón [1] and R.T. Hanson [5]

[1] Dept. of Mathematics, Clarkson University, 8 Clarkson Avenue, Potsdam, NY 13699, USA

[2] Dept. of Mathematical Sciences, Clemson University, O105 Martin Hall, Clemson, SC 29634, USA

[3] Dept. of Aeronautics and Astronautics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

[4] Dept. of Mathematics, Indiana University of Pennsylvania, 210 South 10th St., Indiana, PA 15705 USA

[5] California Water Science Center, U.S. Geological Survey, 4165 Spruance Road, San Diego, CA 92140, USA

[*] **Correspondence:** lea@clemson.edu; Tel: +1-864-656-6907; Fax: +1-864-656-5230.

Abstract: The development of appropriate water management strategies requires, in part, a methodology for quantifying and evaluating the impact of water policy decisions on regional stakeholders. In this work, we describe the framework we are developing to enhance the body of resources available to policy makers, farmers, and other community members in their efforts to understand, quantify, and assess the often competing objectives water consumers have with respect to usage. The foundation for the framework is the construction of a simulation-based optimization software tool using two existing software packages. In particular, we couple a robust optimization software suite (DAKOTA) with the USGS MF-OWHM water management simulation tool to provide a flexible software environment that will enable the evaluation of one or multiple (possibly competing) user-defined (or stakeholder) objectives. We introduce the individual software components and outline the communication strategy we defined for the coupled development. We present numerical results for case studies related to crop portfolio management with several defined objectives. The objectives are not optimally satisfied for any single user class, demonstrating the capability of the software tool to aid in the evaluation of a variety of competing interests.

Keywords: multi-objective optimization; water management; crop planning; code coupling

1. Introduction

Groundwater resources in the U.S. are withdrawn and used for a variety of purposes. However, based on data reported by the U.S. Geological Survey (USGS) [1], the majority of groundwater withdrawals, especially in states with a heavy agriculture sector, are used for agricultural irrigation. Recent studies of aquifer levels in these heavily agricultural states show a marked decrease primarily due to over pumping [2–9]. In several regions, the precipitous drop in levels has been exacerbated by extended periods of arguably historic droughts [10–12]. Farming entities in these regions have been particularly stressed as they are limited with regards to water availability options they can pursue. Many farmers must now operate under restrictions on their water use, either voluntarily or by law [13, 14]. The water restrictions imposed on California farmers have made national headlines; the impact of these restrictions reverberates throughout the country as California is one of the leading producers, both in terms of quantity and selection, of fresh fruits and vegetables in the U.S. California also has some of the largest acreage dedicated to agricultural production in the U.S., and more than half of the vegetables are grown on irrigated land [15]. While recent economic data suggest California farmers have had a record year in terms of profitability [16], the penalty has been the increased use of water resources for the high-revenue, high-water consumption crops that have fed this revenue stream [17].

Our objective in this work is to aid in the decision-making process of farmers and water management agencies via the coupling of mathematical modeling and simulation software with optimization algorithms. Towards this end, we have been developing a flexible software framework to enable farmers and water management agencies to better evaluate the effectiveness of water management strategies relative to objectives connected with stakeholders in an agricultural region.

Gomaa, et.al., provide a brief survey of related work [18], where many of the referenced studies occurred outside the U.S. and covered a range of objective considerations. A more extensive review of works related to the development of models supporting crop management decisions is given in the work by Dury, et.al. [19]. This review paper includes summaries of works where the crop decisions are made based on overall acreage devoted to a single crop versus those works where the crop decisions incorporate spatial considerations, including information on properties such as soil nutrient levels.

More closely related work to this include those that consider evolutionary algorithms and multi-objective approaches to quantify relationships between often competing objectives. Evolutionary algorithms were used to understand trade-offs between sustainability and crop planning in deVoil, et.al., [20], where sustainability was used in both an economic and resource allocation context. Multi-objective optimization of farming practices, including objectives based on crop rotation and environmental farm planning, has also been the focus of several related works [21–25].

Our approach is the first to incorporate the USGS One-Water Hydrologic Flow Model (MF-OWHM, [26]) as a simulation tool coupled to external optimization algorithms. Our previous analysis utilized the predecessor to MF-OWHM, the MODFLOW Farm Process Model (MF-FMP2, [27]). MF-FMP2 and MF-OWHM are agriculturally-focused water management softwares, with the latter offering extended support for the analysis of a wide-range of conjunctive-use issues within a given region. Details on MF-OWHM can be found in Section 3. We chose this software as our simulation workhorse for several reasons. First, the USGS MODFLOW water simulation software is widely used and well respected. Water management was the primary concern of our farming partners in California who were responsible for the genesis of our study. Second, both MF-OWHM and MF-FMP2 have been used extensively in a variety of contexts to study water resource management in heavily farmed areas, where conjunctive use analysis is required to represent the interests of all of the stakeholders in the region [28–35]. Finally, MF-OWHM supports a range of mechanisms for predicting water usage based on climate and plant growth characteristics, all of which enables us to easily evaluate increasingly complicated objective functions with physically realistic parameter spaces.

We utilize the suite of optimization tools available within the DAKOTA optimization package [36], developed and maintained by researchers at the U.S. Department of Energy Sandia National Laboratory. DAKOTA was chosen because of its capabilities in handling simulation-based optimization problems. Users only need to supply subroutines to evaluate the objective functions and constraints without providing any gradient information. Moreover, DAKOTA has a variety of different optimization algorithms allowing for future studies to analyze algorithm performance and suitability for these types of problems.

The objective functions used in this work use output from MF-OWHM. That is, each time the objective function is evaluated for a set of possible crop distributions, new data files must automatically be generated during the course of the optimization using the points the algorithm has chosen. A major contribution of this work is a framework consisting of a set of wrappers, written in Python, that facilitate communication between the simulator and the optimizer. We demonstrate the capability of the framework to choose appropriate crop distributions for planting over a two year time horizon using only ground water as the source of irrigation. This work is a significant improvement from previous efforts in that planting decisions can be made dynamically once a crop is harvested within that two year period [37]. We explain the decision making model, which is governed by the optimization algorithm, in detail below.

We proceed by describing the modeling approach to represent the stakeholders in the agricultural setting, the software framework including the hydrological simulator and the optimization suite, and then provide numerical results. We end with conclusions and describe some future work.

2. Virtual Farming Model

We use mathematical modeling approaches to represent various stakeholders in an agricultural context [20, 23, 37–39]. In this work, our modeling focus is on an individual "virtual" farmer who needs to evaluate a potential crop portfolio based on his need to generate a profit as well as meet a demand specified from industrial partners or membership in a farming cooperative. We emphasize, however, this is a particular choice made for this presentation and the framework can easily be manipulated to consider decisions from other perspectives (e.g., industrial partners, residential community members, water management agencies, or policy makers). The only requirement is the ability to quantify and appropriately model the primary goal of the represented stakeholder. The modeling effort requires identification of appropriate decision variables, objective functions, and constraints. One must also determine the number of crops for which decisions should be made along with the times at which the crops are planted and harvested. Note planting decisions are seasonal, depending on the calendar month and the growing cycle, making it necessary to find the minimum number of decision variables required to simulate a multi-year planting horizon. From a farming perspective, the purpose of the optimization problem may be to select crops based on trade-offs, e.g., between revenue generation and demand or between revenue generation and water usage. In this case, the optimization tool must be able to make decisions in the same time frame as the farmer.

The objective functions aim to provide a farmer with a suite of feasible farm states that allow for crop selection based on individual farming goals. To this end, let N_{crop} be the number of different crops under consideration. We consider changing the crop distribution after a harvest, which drives the number of decision variables in the model in that each time a crop is harvested and land thus becomes available for planting, a decision must be made as to what to do with that land.

The simulation tool used in this work evolves the farming environment over time using *stress periods*. The key idea is that over a given stress period, model parameters are kept constant, inflows and outflows are specified, and assignments for land use, such as cropping, are also held constant. The actual use and movement of water are changing over smaller time periods. We define the variable $x_i^t \in [0, 1]$ to be the fraction of the farm area devoted to crop i during stress period t. This variable is used in our framework to create input files needed by MF-OWHM (specifically the CID files).

We formulate constraints based on the amount of each crop planted at the beginning of each stress period. Let $p_i^t \in [0, 1]$ be the fraction of each farm's area *planted* with crop i at the start of stress period t. Let L_i be the time between planting and harvesting for each crop. Thus, if crop i is planted at the beginning of stress period t, it will be harvested at the beginning of stress period $t + L_i$. We assume that the crops can be planted and harvested instantaneously, so it is possible to both harvest and plant a crop at the beginning of a stress period.

The planting variables p_i^t are related to the fraction of each crop in the ground x_i^t by

$$x_i^t = x_i^{t-1} + p_i^t - p_i^{t-L_i}. \tag{1}$$

Assume that the simulation starts with $t = 1$ and that $x_i^t = 0$ for $t < 1$ and $p_i^t = 0$ for $t < 1$. Thus, given p_i^t for each stress period, x_i^t can be calculated for each stress period. The advantage of framing the problem in terms of p_i^t instead of x_i^t lies in the constraint definitions.

Note that each crop has a specified month it can be planted. Let $P_i \subseteq \{1, 2, 3, \ldots, 12\}$ be the months that crop i can be planted. For example, if crop i can only be planted in May or June, $P_i = \{5, 6\}$. These sets define constraints on the planting variables p_i^t. In particular, we have

$$p_i^t = 0 \quad \text{for} \quad t \notin P_i. \tag{2}$$

The constraints in both (1) and (2) can be generated automatically for all i and t if P_i and L_i are known for each of the N_c crops. Thus, in the framework presented here, these variables are included in the crop class.

Our objectives are defined in terms of farm revenue and an industrial demand for specific crops. We seek to maximize the total revenue over the entire time horizon as we simultaneously minimize deviation from industry demand. The demand objective can be in competition with the revenue objective in scenarios where high-demand crops are water-intensive. The simultaneous optimization means each objective function is considered separately, instead of combining them in a weighted, single-objective approach. As the objectives are competing, there is no single solution that simultaneously optimizes each objective. Our multi-objective approach instead provides a set of points, each optimal in a given context, giving stakeholders the ability to analyze trade-offs between possible solutions.

We calculate the revenue generated from a crop portfolio as the sales price times the yield of each crop minus the cost of water required for irrigation. For simplicity we drop the super scripts and use x_i

to denote the fraction of crop i planted at a given time period. The revenue generated from that planting can be calculated using

$$\text{Maximize } R_i = [x_i \times Y_i \times C_i \times A] - C_w \times W_{gw}, \tag{3}$$

where Y_i is the total yield for crop i (Weight/L^2), C_i is the sales price of crop i ($/Weight), A is the acreage of the farm (L^2), C_w is the cost of groundwater pumping ($/$L^3$) based on volume only, and W_{gw} is the volume of water extracted from the aquifer (L^3). Alternative revenue models could incorporate seed and labor costs for the distribution of different crops, varying water prices [40], the lift cost associated with groundwater pumping from shallow wells, or linear components that represent penalties or incentives that reflect policy or governance.

We model the demand objective using the l_2 norm of the deviation from a specified demand

$$\text{Minimize } D = \|Y_a - Y_d\|_2 = \sqrt{\sum_{i=1}^{N_c} ((Y_a)_i - (Y_d)_i)^2}, \tag{4}$$

where $(Y_a)_i$ is the actual yield (weight) and $(Y_d)_i$ is the demand yield (weight) for crop i dictated, for example, by an industrial partner or market trend. For any crop, the yield, Y_a, is not calculated explicitly by the simulation tool for this work but can be approximated based on the actual evapotranspiration data provided as output by from the MF-OWHM hydrologic model. To derive Y_a, we use a model given by the Food and Agriculture Organization of the United Nations (FAO) [41]

$$\left(1 - \frac{Y_a}{Y_m}\right) = K_y\left(1 - \frac{ET_a}{ET_m}\right), \tag{5}$$

where Y_m is the maximum yield in unstressed conditions (weight), ET_a is the actual crop evapotranspiration (L/T), K_y is the crop water production response coefficient, and ET_m is the maximum crop evapotranspiration in unstressed conditions (L/T). The values of Y_m for most crops are available in the literature.

We can find the maximum crop evapotranspiration ET_m given a known reference evapotranspiration (ET_0) and crop coefficient (K_c) using

$$ET_m = ET_0 \times K_c. \tag{6}$$

In practice, given a reasonable estimate on the unstressed yield for a given crop and the actual crop evapotranspiration, we can estimate the actual yield for a given crop in both stressed and unstressed conditions, extending the robustness of our results to account for drought scenarios.

Since our decision variables are the fraction of each type of crop, we require

$$\sum_{i=1}^{N_c} x_i \leq 1 \tag{7}$$

at each planting decision. Note this also allows land to go fallow.

3. Computational Framework

The main contribution of this study is the implementation of optimization algorithms in conjunction with hydrological models and mathematical representations of stakeholders to guide decisions in an agricultural setting. We proceed by describing the USGS MF-OWHM simulation code used to represent the agricultural setting. We then describe the optimization software suite and the framework developed to create the virtual farming model, which can account for dynamical planting decisions. This is achieved by creating a suite of subroutines that assigns the appropriate input (decision variables) and directs the output from the simulation tool to the objective functions during the course of the optimization.

3.1. Hydrological and Farm Processing Simulation

The One-Water Hydrologic Flow Model (MF-OWHM, [26]) is a MODFLOW-based (MF-05, [42]) integrated hydrologic flow model that is the most complete version, to date, of the MODFLOW family of hydrologic simulators needed for the analysis of a broad range of conjunctive-use issues. MF-OWHM fully links the movement and use of groundwater, surface water, and imported water for consumption by agriculture and natural vegetation on the landscape, and for potable and other uses within a supply-and-demand framework. MF-OWHM is based on the Farm Process for MODFLOW (MF-FMP) [27, 43] combined with local grid refinement, streamflow routing, surface-water routing process, seawater intrusion, and riparian evapotranspiration.

MF-OWHM allows not only for head-dependent flows of a traditional groundwater model but also flow-dependent and deformation-flows for a more complete coupling within the hydrosphere. By retaining and tracking the water within the hydrosphere, MF-OWHM accounts for "all of the water everywhere and all of the time." This approach provides more complete water accounting and provides a platform needed to address wider classes of problems such as evaluation of conjunctive-use alternatives, including sustainability analysis, potential adaptation and mitigation strategies, and development of best management practices [44]. MF-OWHM's broader ability to simulate more of the hydrosphere has served as a valuable tool for multiple research and applied modeling projects.

As research tools, both MF-FMP and MF-OWHM have been modified to investigate mathematical techniques, including subsidence feedback on conjunctive use [45], effects of climate change [35, 46], crop optimization [37, 43, 47], water-rights driven surface water allocations [48], and proper orthogonal decomposition model reduction [49–51]. MF-FMP and MF-OWHM have also been used to evaluate many applied projects within the U.S. Geological Survey and the private sector [28, 29, 31, 34, 49, 52–56].

The allocation of crops is input for an MF-OWHM simulation and evapotranspiration and water usage values are extracted at the end of a simulation. These values are then used to calculate revenue and yield using Equations (3) and (4) above. The details of the crops and the physical description of the farm for this work are described later.

3.2. Optimization

DAKOTA provides a suite of optimization strategies for a range of simulation-based scenarios and is an ideal framework for this study since it is open-source and flexible [36]. Understanding the applicability of optimization algorithms, in particular evolutionary algorithms, to real-world problems is an active area of research [57]. Since the objectives are competing, there is no single solution that simultaneously optimizes each objective. A multi-objective approach instead provides a set of points, giving stakeholders the ability to analyze trade-offs between points. In general, genetic algorithms move through "generations" of evaluation points by assessing the fitness of members of the generation and selecting members to continue to the next generation (through mutation or cloning), parent offspring for the next generation, or die (i.e. removing those points from the population). In a multi-objective setting the population evolves towards a set in which the points are non-dominated, known as a Pareto set. A non-dominated point has the property that its fitness cannot improve with respect to one objective without degrading the value of another. The basic steps of the algorithm are to initialize a population, evaluate the objective function and constraints, perform crossover and mutation, evaluate the new population members and assess the fitness of each population. Population members are then replaced to continue to the next generation. Termination of the optimization can be based on a maximum number of function evaluations (or iterations) or on performance metrics. The performance metrics track changes in the population from generation to generation. Finally, a post processing step reduces the final solution set so that a minimum distance exists between any two design points.

We use the multi-objective genetic algorithm (MOGA) developed by Eddy [58] as implemented in DAKOTA. This method was shown to perform well on a simplified farming model [37].

3.3. Implementation

For this work, DAKOTA and MF-OWHM are linked via objective function evaluations facilitated by a framework of Python wrappers. Figure 1 illustrates the configuration of the framework. Initial set up of the physical setting, generation of the data files for the simulator, and assignment of objective function parameters are done with an IPython notebook. We list the specific model parameters in Figure 1 in the upper left box. The objective functions and constraints are defined via a setup script and DAKOTA is used to solve the optimization problem with repeated calls to MF-OWHM for each objective function evaluation. Each simulation requires a new crop distribution, thereby requiring new data files which are created automatically by additional Python scripts. This is demonstrated in the upper right box in Figure 1.

4. Description of Example Problem

We demonstrate our modeling approach by considering a farm originally implemented in MF-FMP2 and modified for use in MF-OWHM and the virtual farmer as modeled by Equations (3) and (4) above. We consider three crops with competing properties and varying planting and growing schedules to demonstrate how the optimization algorithm can guide the planting decisions.

4.1. Agricultural Setting

The problem formulation uses a dataset to represent a farm originally included as a test problem with the suite of MF-FMP2 test problems and later modified for use with MF-OWHM. This farm model is a hypothetical realization of a region with multi-layered hydrology and distinct regions with different water usage requirements [27]. The model includes a 10 km by 11.5 km region with multiple

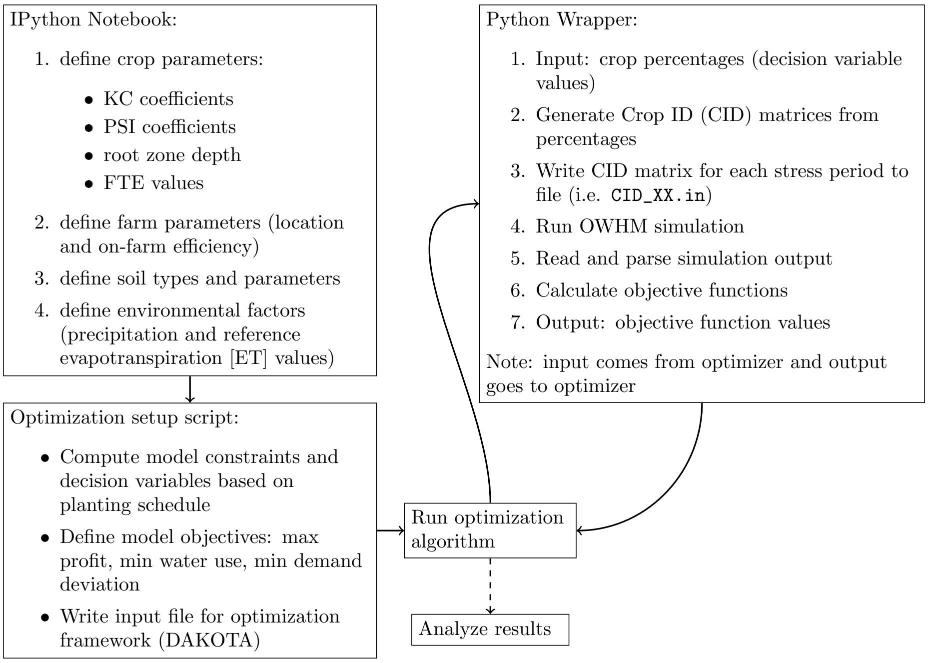

Figure 1. Graphical depiction of the workflow for the software coupling. Preprocessing and communication are handled using Python based tools.

farms, urban zones, riparian zones (interfaces between land and streams), and areas of natural vegetation. The water management system includes multiple wells, stream inflows and outflows, and natural precipitation.

The topography slopes downward from west to east and converges from the north and south toward a riparian region along the eastern edge. The underlying geology contains four aquifer layers separated by three layers of confining material [27]. The aquifer nearest the surface is unconfined with varying depth. The remaining (confined) layers are uniformly 60 m thick, with each layer of confining material between 5 m and 15 m thick. The saturated hydraulic conductivity (a measure of the ability of the aquifer to transmit fluid) varies from 10 m/day in the aquifer nearest the surface to 0.15 m/day in the aquifer furthest from the surface.

The example region is divided by a stream flowing west to east. The stream flow into the domain is prescribed at 50,000-100,000 m^3/day. No fluid flow is allowed into the region through the northern and southern boundaries. The eastern and western boundaries are general head boundaries. These specify a head-dependent flux designed to mimic known groundwater head at a specified distance. The topography and boundary conditions dictate a west-to-east directional groundwater flow.

The model domain consists of a 20x23 grid, where each cell within the grid is associated with a specific farm type. The production farm is modeled using a 10x10 block of cells in the upper left corner of the domain. The riparian vegetation zone is comprised of a block of 25 cells lying along the east boundary, and the remaining cells are associated with native vegetation landscape. A schematic of the model domain is given in Figure 2, with the different regions represented by distinct colors. The orange cells denote the production farm, the green cells denote the riparian region, and the dark gray cells denote the native vegetation. The blue circles show the locations of the wells, while the blue line represents a stream. The production farm is the focus of the simulation, but the properties associated with the surrounding landscape also affect the dynamics of the simulation [40].

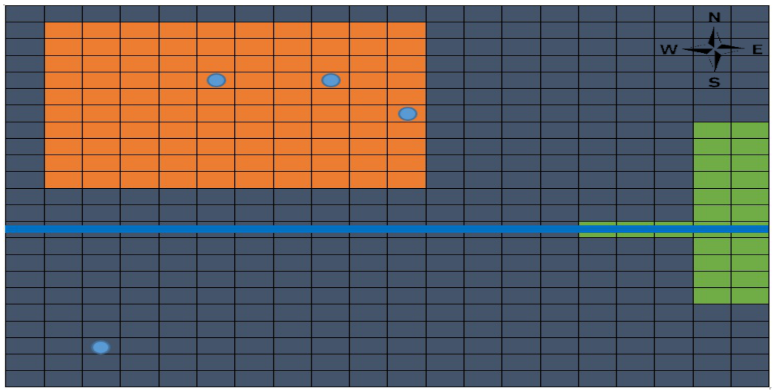

Figure 2. Diagram of test problem farm configuration. Orange cells are farm regions, dark gray cells are native vegetation, and green cells are riparian vegetation. The blue line represents a river, and blue circles denote wells.

In practice, the physical composition of the subsurface impacts the optimization as soil properties govern runoff and inefficiencies in the irrigation system. Our test problem consists of three different soil types which are predefined in MF-FMP and MF-OWHM. The soil types are silt, sandy loam, and silty clay. Their distribution is shown in Figure 3.

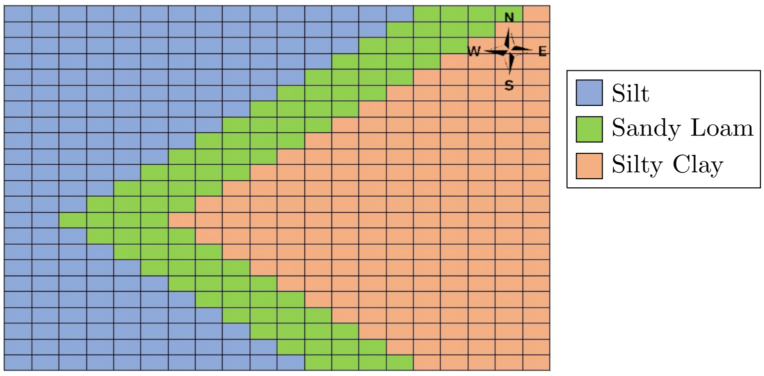

Figure 3. Distribution of soils throughout the model domain.

Several adjustments were made to allow for a more definitive assessment of the performance of the optimization algorithm and the coupling strategy. In particular, we reduced the number of farms to three; an agricultural farm, a riparian zone, and an area of native vegetation. We consider five crops: three are agricultural products of the farm (our decision variables), one is the vegetation in the riparian zone, and the final is the native vegetation. In this model, crops receive water through groundwater irrigation or precipitation only, with a uniform distribution across all crops. Four extraction wells are included in the model and all other water delivery options were not used. We incorporate precipitation data taken from a weather station in Watsonville, CA from January 2013 – December 2014. The volumetric flow rates for the precipitation are shown in Figure 6 (the dashed line) which varies significantly over the two years and includes two periods of little to no precipitation.

4.2. *Crops*

For this demonstration, we consider three crops over a two year time frame. We allow land to go fallow as an implicit fourth crop. That is, the production crops are not forced to occupy the total farm area, leaving any unoccupied acreage fallow. We assume that the water and sales prices do not vary significantly, but varying model parameters are a focus of future work. Further, we chose the three crops with differing properties to demonstrate the trade-off analysis for decision making. Crop 1 is a water intensive but high profit crop with low demand. Crop 1 can be planted in April or May and is in the ground for 4 months. Crop 2 uses the least amount of water, has medium demand, and is in the middle with regards to profitability. Crop 2 can be planted in January or June stays in the ground for 2 months. Finally, Crop 3 is the highest demand crop, has medium water usage, but is low profit in comparison with the other crops. Crop 3 can be planted only in June and is in the ground for 5 months. We show the planting possibilities over the two year horizon in Figure 4. Here p_i^t refers to the fraction of farmland planted with crop i at stress period t.

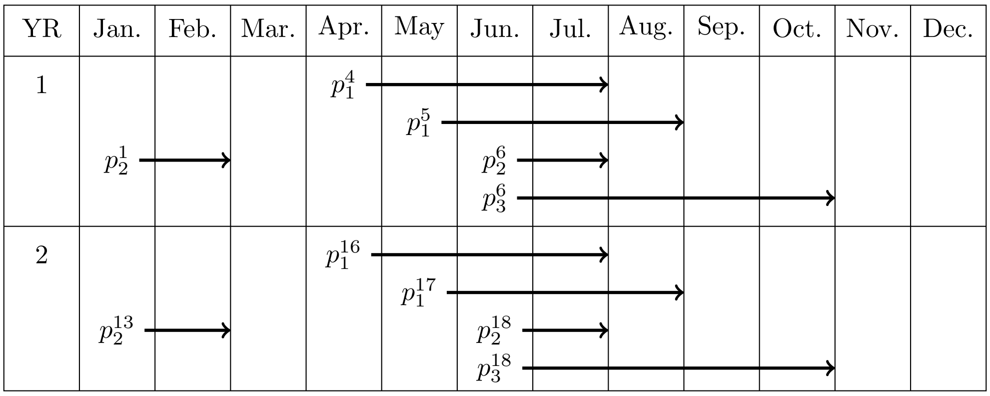

Figure 4. Planting schedule for crops 1,2,3 over the 2 year period.

Specific model parameters for the crops used in our simulation are provided in a database downloaded with MF-OWHM and [59–61] (We loosely base them on vegetable row crops, strawberries, and artichokes). The model parameters describing the attributes are shown in Table 1.

Table 1. Model parameters for each crop. The relative values between crops for each parameter give each an advantage with respect to a distinct objective.

Crop	1	2	3
Max Yield (Y: Metric Ton/acre)	8.0	24.5	8.8
Sales Price (C_i: \$/Metric Ton)	150.47	10.16	5.58
Demand (Y_d: Metric Tons)	3,300	19,700	53,750

5. Numerical Results

MOGA is a heuristic algorithm with randomness in its search phase; therefore, it is ideal to run the optimization multiple times. We performed 13 optimization trials using the default algorithm parameters including a mutation rate of 0.1 and a crossover rate of 0.8. The numerical experiments were done on a machine running Ubuntu 14.04.3 LTS with 16x AMD Opteron (TM) Processor 6212 and 32G of Memory. We use DAKOTA version 6.3 released November 15, 2015, and OWHM 1.0 Version 1.0.11 released May 5, 2015 for all simulations. A single optimization experiment took roughly two hours.

Figure 5 shows a typical trade-off curve generated by the optimizer. Each point on the curve corresponds to a specified planting portfolio so that a farmer could select a distribution of crops based on a personal preference in balancing the demand and profit objectives. In particular, notice that the objectives are indeed competing in that deviating from the demand leads to an increase in profit. We note the demand deviation measures the difference between what the farm yields and what we have specified as a demand for the corresponding crop (i.e., $Y_a - Y_d$ in Equation (4)). Recall with competing objectives there is no single solution that will optimize all the objective functions. DAKOTA does identify a "best" point in the Pareto set defined in terms of distance from the so-called utopia point (labelled above with a diamond and a star). The utopia point is defined as the point of extreme best values for each objective. The best point is then the point on the Pareto front closest to the utopia point. An example from the DAKOTA manual [36] highlights this concept. The best point as well as the one that gave a higher profit for one of the optimization runs are given in Table 2. The second point, (indicated with a triangle) was chosen by scaling the objective functions to both be between zero and one and then selecting the point closest to the Utopia point. A benefit of using multi-objective optimization is that a user can consider the trade-offs between the solutions on the Pareto front.

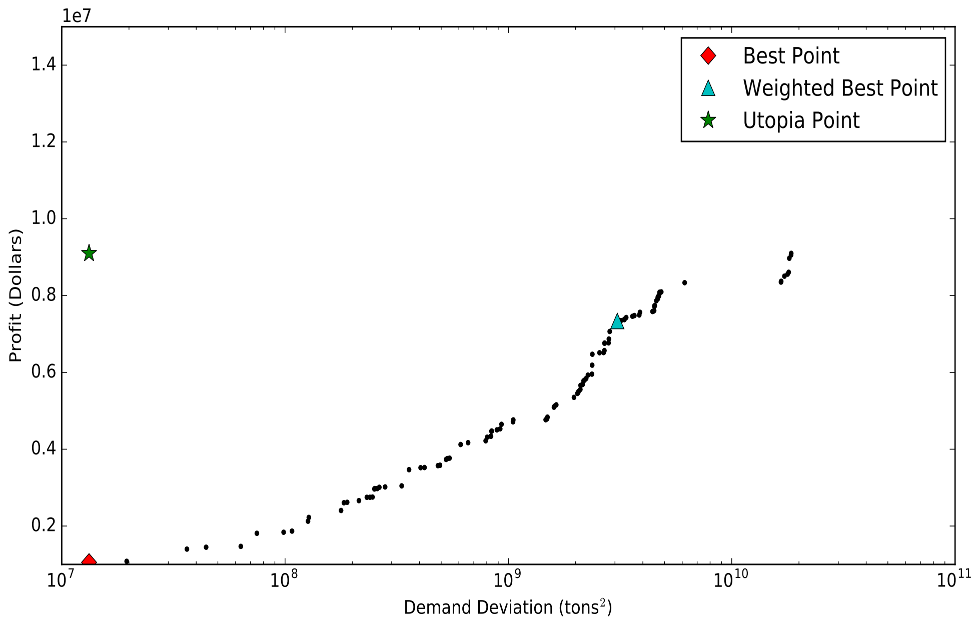

Figure 5. Representative trade-off curve from data in Table 2.

The crop portfolios are the same in terms of the fraction of Crop 1 for the third and fourth planting as well as the first planting of Crop 2. Note Crop 1 was a low demand, high water use crop. The precipitation during the time for the April and May plantings of Crop 1 in the second year was low (see Figure 6), requiring increased groundwater pumping and hence increased cost. Thus, the optimizer selected low percentages of this crop to appropriately balance the competing objectives.

The portfolios differ most in the amount of Crop 2 for its third planting. Crop 2 was a low water, medium demand, and high profit crop. Crop 2 is planted in January and in year 2 at that time, there is a significant precipitation event (see Figure 6). They also differ significantly in Crop 3 (the high demand crop) during its first planting. Since the best point planting portfolio is also trying to satisfy the demand objective and since the only possible planting for January is Crop 2 (leading to a profit), these differences make sense.

Table 2. Fractions of Crops 1,2,3 and the resulting profit, deviation from demand, and water usage. The superscript on p_i denotes the stress period.

	Example Best Point	Weighted Best Point	Mean (Best Point)	Std. Dev. (Best Point)
p_1^4 (April Yr 1)	0.06	0.41	0.08	0.0617
p_1^5 (May Yr 1)	0.46	0.43	0.11	0.1284
p_1^{16} (April Yr 2)	0.00	0.00	0.09	0.0929
p_1^{17} (May Yr 2)	0.14	0.14	0.09	0.0548
p_2^1 (Jan Yr 1)	0.37	0.37	0.33	0.0900
p_2^6 (June Yr 1)	0.08	0.06	0.10	0.1066
p_2^{13} (Jan Yr 2)	0.02	0.74	0.56	0.3360
p_2^{18} (June Yr 2)	0.21	0.14	0.17	0.1129
p_3^6 (June Yr 1)	0.37	0.09	0.36	0.1478
p_3^{18} (June Yr 2)	0.60	0.69	0.34	0.2016
Profit ($)	1,053,365.84	7,337,207.88	786,983.68	184,460.12
Demand Deviation (Tons)	3,640	55,439	2654	1922
Water Use (m^3)	6,376,770.85	6,723,169.86	6,452,116.95	543,924.99

For any given crop portfolio, a post-processing simulation can be run to analyze water usage. Figure 6 shows the volumetric flow rate for the precipitation events over the two year period and the extraction wells for the farm. The figure highlights that during rainfall events, the wells shut down; otherwise, they are operating at their maximum capacity. The crop fractions used to generate this plot were from Column 2 in Table 2 (the high profit point).

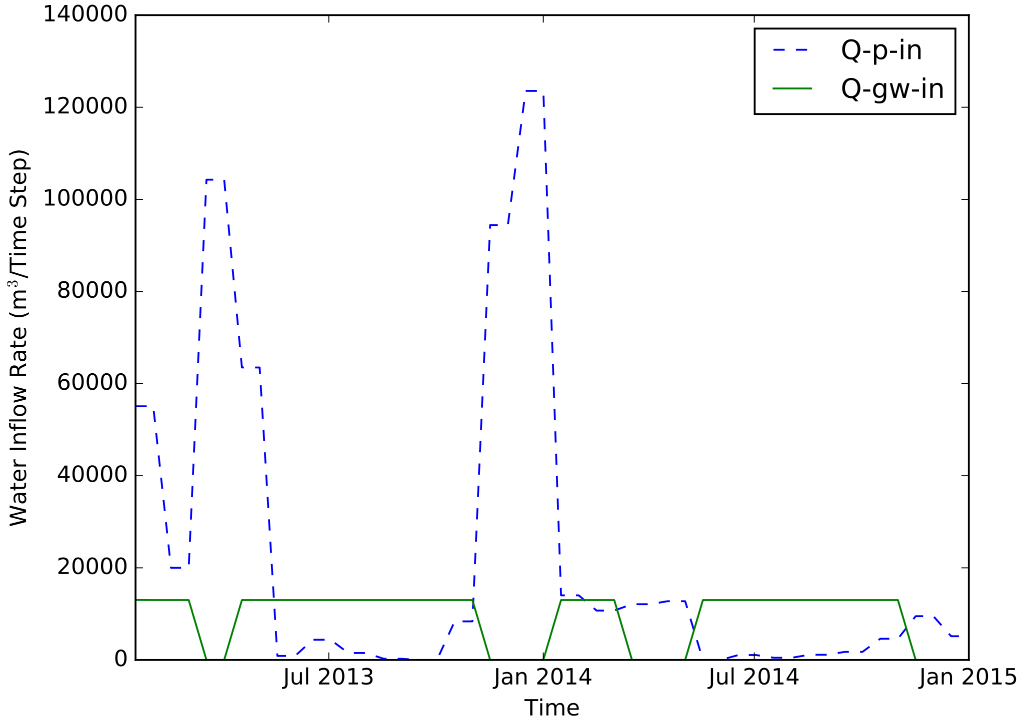

Figure 6. Volumetric water usage from wells and rainfall events over the two year horizon. Q_p denotes precipitation and Q_{gw} denotes groundwater usage.

6. Conclusions

This work is motivated by a need to understand, in a quantifiable way, how decisions made by stakeholders in an agricultural community impact water usage and revenue considerations for both individuals farmers and the entities purchasing their harvests. We have proposed a computational framework which uses dynamic planting rules and a simulation-based optimization approach to create a virtual farmer. The framework uses Python wrappers to manage the communication between the conjunctive-use hydrological simulator MF-OWHM and the DAKOTA optimization suite. We considered two competing objectives, minimizing deviation from industry demand and maximizing profit, and applied the MOGA to search the design space and present feasible solutions. We demonstrated the capabilities of this approach on a test problem with a realistic farming scenario and three crops with competing properties. Sets of solutions in the form of trade-off curves provide stakeholders with options when selecting a crop planning portfolio to balance their personal preferences. In particular, this work is an improvement over the initial use of DAKOTA coupled with MF-FMP2 proposed in Fowler, et.al., [37] in that a crop distribution could only be determined once at the beginning of a simulation. In that work, for example, the distribution of crops was defined at the beginning of the two-year horizon and the same crops were always planted when land became available. The approach described here is more realistic and allows the optimization algorithm to select new crops to be planted once land opens up from a harvest.

More important, however, is the ability of the framework to facilitate a wide range of future studies. More realistic objective functions that represent additional stakeholders, such as citizens or policy makers can easily be implemented. For example, the crop that is in highest demand could be weighted more heavily in that objective. Varying water prices and sales prices would also impact the solutions. In addition, DAKOTA has a suite of optimization algorithms and sensitivity analysis tools that can be applied to better understand the nature of the problems.

Finally, each component is open source, and the communication mechanism is defined in an object-oriented context. These communication tools have allowed us to take advantage of the significant research contributions from the USGS and Sandia National Laboratories. The object-oriented nature of the tools allows us to easily add more components to a crop class and incorporate new parameters into a given study. Future work includes development of graphical user-interfaces enabling users to easily define distinct objectives or new crop information, and studies on larger regions with several farmers and urban entities accessing the same water sources for their needs. We also note we only use groundwater sources for irrigation in this work; consideration of all possible water delivery systems is a direction of future research.

Table 3. Table of notation. Units used in this work are in parentheses.

Notation	Description	Units
x_i^t	fraction of farm area devoted to crop *i* during stress period *t*	no units
p_i^t	fraction of farm area planted with crop *i* at stress period *t*	no units
R_i	revenue from crop *i*	\$
Y_i	total yield for crop *i*	weight/L^2 (tons/acre)
C_i	sales price of crop *i*	\$/weight (\$/tons)
A	acreage of the farm	L^2 (m^2)
C_W	cost of groundwater pumping	\$/L^3 (\$/m^3)
W_{gw}	volume of extracted water	L^3 (m^3)
Y_a	actual yield	weight (tons)
Y_d	demand yield	weight (tons)
K_y	crop water production response coefficient	no units
ET_a	actual crop evapotranspiration	L/T (m/day)
ET_m	maximum crop evapotranspiration	L/T (m/day)
ET_0	reference evapotranspiration	L/T (m/day)
K_C	crop coefficient	no units
N_C	number of crops	no units

Acknowledgments

We would like to thank the American Institute of Mathematics for facilitating and supporting this collaboration. We would also like to thank the McNair Scholars program for their support of this work. We thank Corey Ostrove of the Applied Research Laboratories, The University of Texas at Austin, and Matthew Farthing of the Engineer Research and Development Center, U.S. Army Corps of Engineers, for their initial efforts on this framework. We would also like to thank USGS internal reviewers and the anonymous referee for the journal for suggestions that improved this manuscript, and we thank Dr. Alexander Schreiber for his assistance in generating figures.

Conflict of Interest

All authors declare no conflict of interest in this paper.

References

1 M. Maupin, J. Kenny, S. Hutson, J. Lovelace, N. Barber and K. Linsey, (2014), *Estimated use of water in the United States in 2010*, Circular 1405, U.S. Geological Survey, Available from: `http://dx.doi.org/10.3133/cir1405`.

2 I. James and S. Reilly, (2015), Pumped beyond limits, many U.S. aquifers in decline, *The Desert Sun*.

3 J. Thomas, G. Stanton, J. Bumgarner, D. Pearson, A. Teeple, N. Houston, J. Payne and M. Musgrove, (2013), *A conceptual hydrogeologic model for the hydrogeologic framework, geochemistry, and groundwater-flow system of the Edwards-Trinity and related aquifers in the Pecos County region, Texas*, Fact Sheet 2013-3024, U.S. Geological Survey.

4 F. Morris, (2013), Western Kansas farmers face dwindling water supply, Available from: `http://hereandnow.wbur.org/2013/08/19/kansas-farmers-water`.

5 D. Steward, P. Bruss, X. Yang, S. Staggenborg, S. Welch and M. Apley, (2013), Tapping unsustainable groundwater stores for agricultural production in the High Plains aquifer of Kansas, projections to 2110, *Proceedings of the National Academy of Sciences*, **110**: E3477–E3486.

6 J. Peterson and Y. Ding, (2004), Economic adjustments to groundwater depletion in the High Plains: Do water-saving irrigation systems save water?, *Am. J. Agr. Econ.*, **87**: 147–159.

7 V. McGuire, M. Johnson, R. Schieffer, J. Stanton, S. Sebree and I. Verstraeten, (2002), *Water in storage and groundwater management approaches, High Plains Aquifer, 2000*, Circular, U.S. Geological Survey.

8 J. Musick, F. Pringle, W. Harman and B. Stewart, (1990), Long-term irrigation trends – Texas High Plains, *Appl. Eng. Agric.*, **6**: 717–724.

9 M. Sophocleous, (2010), Review: groundwater management practices, challenges, and innovations in the High Plains Aquifer, USA – lessons and recommended actions, *Hydrogeol. J.*, **18**: 559–575.

10 California Water Science Center, (2016), The California drought, Available from: `http://ca.water.usgs.gov/data/drought/drought-impact.html`.

11 V. McGuire, (2014), *Water-level changes and change in water in storage in the High Plains Aquifer, Predevelopment to 2013 and 2011–2013*, Scientific Investigations Report 2014–5218, U.S. Geological Survey.

12 B. Scanlon, C. Faunt, L. Longuevergne, R. Reedy, W. Alley, V. McGuire and P. McMahon, (2012), Groundwater depletion and sustainability of irrigation in the US High Plains and Central Valley, *PNAS*, **109**: 9320–9325.

13 J. Medina, (2015), California cuts farmers' share of scant water, *The New York Times*, Available from: `http://www.nytimes.com/2015/06/13/us/california-announces-restrictions-on-water-use-by-farmers.html?_r=0`.

14 D. Charles, (2013), Kansas farmers commit to taking less water from the ground, *All Things Considered*, Available from: `http://www.npr.org/blogs/thesalt/2013/10/22/230702453/in-kansas`.

15 H. Wells, (2015), Vegetables and pulses, Available from: `http:www.ers.usda.gov/topics/crops/vegetables-pulses.aspx`.

16 D. Goolsby, (2015), Coachella Valley agriculture industry continues growing, *The Desert Sun*.

17 H. Cooley, K. Connelly, R. Phurisamban and M. Subramanian, (2015), *Impacts of California's Ongoing Drought: Agriculture*, Technical report, Pacific Institute, Oakland, CA, Available from: `http://pacinst.org/publication/impacts-of-californias-ongoing-drought-agriculture/`.

18 W. Gomaa, N. Harraz and A. el Tawil, (2011), Crop planning and water management: A survey, in *Proceedings of the* 41^{st} *International Conference on Computers & Industrial Engineering*, Los Angeles, CA, 319–324.

19 J. Dury, N. Schaller, F. Garcia, A. Reynaud and J. Bergez, (2012), Models to support cropping plan and crop rotation decisions: a review, *Agronomy Sust. Developm.*, **32**: 567–580.

20 P. deVoil, W. Rossing and G. Hammer, (2006), Exploring profit – sustainability trade-offs in cropping systems using evolutionary algorithms, *Environ. Modell. Softw.*, **21**: 1368–1374.

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

22 R. Beneke and R. Winterboer, (1984), *Linear Programming. Applications to Agriculture*, Aedos.

23 J. Groot, G. Oomen and W. Rossing, (2012), Multi-objective optimization and design of farming systems, *Agr. Syst.*, **110**: 63–77.

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

25 R. Sarker and T. Ray, (2009), An improved evolutionary algorithm for solving multi-objective crop planning models, *Comput. Electron. Agric.*, **68**: 191–199.

26 R. Hanson, S. Boyce, W. Schmid, J. Hughes, S. Mehl, S. Leake, T. Maddock III and R. Niswonger, (2014), *One-Water Hydrologic Flow Model (MODFLOW–OWHM)*, Techniques and Methods 6-A51, U.S. Geological Survey, Available from: `http://dx.doi.org/10.3133/tm6A51`.

27 W. Schmid and R. Hanson, (2009), *The Farm Process Version 2 (FMP2) for MODFLOW-2005 — Modifications and Upgrades to FMP1*, Techniques in Water Resources Investigations 6-A32, U.S. Geological Survey.

28 C. Faunt, R. Hanson, K. Belitz and L. Rogers, (2009), *California's Central Valley Groundwater Study: A powerful new tool to assess water resources in California's Central Valley*, Fact Sheet 2009-3057, U.S. Geological Survey.

29 C. Faunt, C. Stamos, L. Flint, M. Wright, M. Burgess, M. Sneed, J. Brandt, A. Coes and P. Martin, (2015), *Hydrogeology, hydrologic effect of development, and simulation of groundwater flow in the Borrego Valley, San Diego County, California*, Scientific Investigations Report 2015–5150, U.S. Geological Survey.

30 R. Hanson, W. Schmid, J. Lear and C. Faunt, (2008), Simulation of an aquifer-storage-and-recovery (ASR) system for agricultural water supply using the farm process in MODFLOW for the Pajaro Valley, Monterey Bay, California, in *MODFLOW and More 2008: Groundwater and Public Policy*, International Ground Water Modeling Center, Colorado School of Mines, 501–505.

31 R. Hanson, W. Schmid, J. Knight and T. Maddock III, (2013), Integrated hydrologic modeling of a transboundary aquifer system - Lower Rio Grande, in *MODFLOW and More 2013: Translating Science into Practice*, International Ground Water Modeling Center, Colorado School of Mines, Golden, CO, 5 p.

32 R. Hanson, W. Schmid, C. Faunt and B. Lockwood, (2010), Simulation and analysis of conjunctive use with MODFLOW's farm process, *Ground Water*, **48**: 674–689.

33 W. Schmid, J. P. King and T. Maddock III, (2009), *Conjunctive Surface-water/groundwater Model in the Southern Rincon Valley Using MODFLOW-2005 with the Farm Process*, Technical report, New Mexico Water Resources Research Institute.

34 R. Hanson, W. Schmid, C. Faunt, J. Lear and B. Lockwood, (2014), *Integrated hydrologic model of Pajaro Valley, Santa Cruz and Monterery Counties, California*, Scientific Investigations Report 2014-5111, U.S. Geological Survey.

35 R. Hanson, L. Flint, A. Flint, M. Dettinger, C. Faunt, D. Cayan and W. Schmid, (2012), A method for physically based model analysis of conjunctive use in response to potential climate changes, *Water Resour. Res.*, **48**: 23.

36 B. Adams, L. Bauman, W. Bohnhoff, K. Dalbey, M. Ebeida, J. Eddy, M. Eldred, P. Hough, K. Hu, J. Jakeman, L. Swiler and D. Vigil, (2009), *DAKOTA: A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 5.4 User's Manual*, Technical report SAND2010-2183, Sandia National Laboratories (updated April 2013).

37 K. Fowler, E. Jenkins, C. Ostrove, J. Chrispell, M. Farthing and M. Parno, (2014), A decision making framework with MODFLOW-FMP2 via optimization: Determining trade-offs in crop selection, *Environ. Modell. Softw.*, **69**: 280–291.

38 N. Lehmann, R. Finger, T. Klein, P. Calanca and A. Walter, (2013), Adapting crop management practices to climate change: Modeling optimal solutions at the field scale, *Agr. Syst.*, **117**: 55–65.

39 S. Lautenbach, M. Volk, M. Stauch and G. Whittaker, (2013), Optimization-based trade-off analysis of biodiesel crop production for managing agricultural catchment, *Environ. Modell. Softw.*, **48**: 98–112.

40 J. Bokhiria, K. Fowler and E. Jenkins, (2014), Modelling and optimization for crop portfolio management under limited irrigation strategies, *Journal of Agriculture and Environmental Sciences*, **2**: 1–13.

41 P. Steduto, T. Hsiao, E. Fereres and D. Raes, (2012), *Crop Yield Response to Water*, FAO irrigation and drainage, Paper 66, Food and Agriculture Organization of the United Nations.

42 A. Harbaugh, (2005), *MODFLOW-2005: The U.S. Geological Survey modular ground-water model: The groundwater flow process*, Techniques and Methods 6-A16, U.S. Geological Survey.

43 W. Schmid, R. Hanson, T. M. III and S. Leake, (2006), *User guide for the farm process (FMP1) for the U.S. Geological Survey's modular three-dimensional finite-difference ground-water flow model, MODFLOW-2000*, Techniques and Methods 6-A17, U.S. Geological Survey.

44 R. Hanson and W. Schmid, (2013), *Economic Resilience through "One-Water" Management*, Open File Report 2013-1175, U. S. Geological Survey.

45 W. Schmid, R. Hanson, S. Leake, J. Hughes and R. Niswonger, (2014), Feedback of land subsidence on the movement and conjunctive use of water resources, *Environ. Modell. Softw.*, **62**: 253–270.

46 I. Ferguson and D. Llewellyn, (2015), *Simulation of Rio Grande project operations in the Rincon and Mesilla Basins: Summary of model configuration and results*, Technical Memorandum 86–68210–2015–05, U.S. Bureau of Reclamation.

47 G. Schoups, C. Addams, J. Minjares and S. Gorelick, (2006), Sustainable conjunctive water management in irrigated agriculture: Model formulation and application to the Yaqui Valley, Mexico, *Water Resour. Res.*, **42**: 19.

48 W. Schmid and R. Hanson, (2007), Simulation of intra- or trans-boundary water-rights hierarchies using the farm process for MODFLOW-2000, *J. Water Res. Pl. – ASCE*, **133**: 166–178.

49 S. Boyce and R. Hanson, (2015), An integrated approach to conjunctive–use analysis with the one–water hydrologic flow model, MODFLOW-OWHM, in *MODFLOW and More 2015: Modeling a Complex World – Integrated Modeling to Understand and Manage Water Supply, Water Quality, and Ecology*, Colorado School of Mines, Golden, CO, 6–10.

50 S. Boyce, (2015), *Model Reduction via Proper Orthogonal Decomposition of Transient Confined and Unconfined Groundwater Flow*, PhD thesis, University of California Los Angeles.

51 S. Boyce, T. Nishikawa and W. Yeh, (2015), Reduced order modeling of the Newton formulation of MODFLOW to solve unconfined groundwater flow, *Adv. Water Res.*, **83**: 250–262.

52 C. Faunt, (2009), *Groundwater availability of the Central Valley aquifer, California*, Professional Paper 1766, U. S. Geological Survey.

53 R. Hanson, B. Lockwood and W. Schmid, (2014), Analysis of projected water availability with current basin management plan, Pajaro Valley, California, *J. Hydrol.*, **519**: 131–147.

54 R. Hanson, L. Flint, C. Faunt, D. Gibbs and W. Schmid, (2014), *Hydrologic models and analysis of water availability in Cuyama Valley, California*, Science Investigations Report 2014-5150, U.S. Geological Survey, Available from: `http://dx.doi.org/10.3133/sir20145150`.

55 R. Hanson and D. Sweetkind, (2014), *Water availability in Cuyama Valley, California*, Fact Sheet FS2014-3075, U.S. Geological Survey.

56 T. Russo, A. Fisher and B. Lockwood, (2014), Assessment of managed aquifer recharge site suitability using a GIS and modeling, *Ground Water*, **53**: 1–12.

57 H. Maier, Z. Kapelan, J. Kasprzyk, J. Kollat, L. Matott, M. Cunha, G. Dandy, M. Gibbs, E. Keedwell, A. Marchi, A. Ostfeld, D. Savic, D. Solomatine, J. Vrugt, A. Zecchin, B. Minsker, E. Barbour, G. Kuczera and F. Pasha, (2014), Evolutionary algorithms and other metaheuristics in water resources: Current status, research challenges and future directions, *Environ. Modell. Softw.*, **62**: 271–299.

58 J. Eddy and K. Lewis, (2001), Effective generation of Pareto sets using genetic programming, in *Proceedings DETC '01: ASME 2001 Design Engineering Technical Conferences*, Pittsburgh, PA, 1–9.

59 K. Demchak, J. Harper and L. Klime, Strawberry production, Available from: `http://extension.psu.edu/business/ag-alternatives/horticulture/fruits/strawberry-production#section-4`.

60 R. Smith, A. Baameur, M. Bari, M. Cahn, D. Giraud, E. Natwick and E. Takele, Artichoke production in California, Available from: `http://anrcatalog.ucanr.edu/pdf/7221.pdf`.

61 M. LeStrange, M. Cahn, S. Koike, R. Smith, O. Daugovish, S. Fennimore, E. Natwick, S. Dara, E. Takele and M. Cantwell, Artichoke production in California, Available from: `http://anrcatalog.ucanr.edu/pdf/7211.pdf`.

AIMS Press