OPTIMISING FRUIT CULTIVARS IRRIGATION VIA A HIERARCHICAL PARTITIONING METHOD

C. Esther Van Cauwenberghe and J. Alberto Bandoni

{cvancauwen, abandoni}@plapiqui.edu.ar

Planta Piloto de Ingeniería Química (PLAPIQUI - CONICET)

Cno. La Carrindanga Km 7, Bahía Blanca, 8000, Argentina

Tel.: +54-291-4861700 / FAX: +54-291-4861600

ABSTRACT

The urgent current necessities in modern agriculture has afforded in a growing application of computer science and electronics to this field. Precision and sustainable agriculture, involve not only new production methods, but also complex systems that integrates biological, technological and economical factors in a flexible environment, to cope with the uncertainty of nature. This form of agriculture aims to optimise proficiency and environmental protection, by assessing together forecasts for risk, damage and profit. One aspect to consider in order to reach that purpose, is the availability of good irrigation systems (Walker, 2002). Improving water efficiency in irrigated agriculture is a priority for better environmental and economic performance reducing the required amount of water and optimising the timing of application (Howell, 2001). Among numerous methods, hybrid biophysical/decisional simulation models are effective tools for strategies under different weather conditions. Nevertheless, optimising irrigation strategies for some specified agro-environmental expected criteria represents a computationally hard problem.

In this paper we propose a new approach for optimising the parameters of seed-fruit irrigation strategies, represented by a simulation-optimisation partitioning method, SOP-κP, which is designed to completely explore by sampling the domain of the strategy parameters. This exploration is based on a hierarchical decomposition of the domain that heuristically guides the search toward optimal zones, allowing considering high dimensional optimisation problems. Firstly, we evaluate a irrigation strategy by comparison with a systematic grid search on a simple two-parameter problem in cultivars areas located at North-Patagonian region -Southern Argentine- in order to optimise the expectation of the seed-fruit gain margin, and using 20 years of weather records as input data. The similar results confirm the sound behaviour of the partitioning method. Then, we apply our approach to a more complex optimisation problem that involves an eight-parameter strategy, for which the systematic grid search method is not effective. The best solution-strategy we obtain shows 2527.43 U$S ha^{-1} increase in the gain margin compared to a current strategy. A numerical model for simulation of infiltration and groundwater flow in a different porous media with variable saturation level is also incorporated. The different optimal parameters obtained for each context are in good agreement with the expert knowledge of irrigation advisors. Finally, we conclude discussing some limitations and possible improvements of the presented work.

1. INTRODUCTION

Modern agriculture has afforded in a growing application of computer science and electronics in order to reach precision and sustainability in its practice. This activity is an important consumer of water and irrigation, and consequent, essential in some areas to allow acceptable seed-fruit yield and quality. Improved water use efficiency in irrigated agriculture is therefore a priority for better environmental and economic performance (Howell, 2001). Irrigation is necessary because of the non-uniform distribution of rainfall, the very limited water-holding capacities of typical sandy soils, and the extreme sensitivity of many specialty crops to water stress. These factors and the economic losses from under-or-over-irrigation require that irrigations be scheduled as efficiently as possible.

Throughout the irrigated world, water is applied to fields unevenly and excessively, leading to wastage and to pollution of surface and ground-waters. Computer solution of the governing equations with given irrigation conditions allows rapid evaluation of physical layout and operation. Systematic, repeated simulation can lead to design parameters yielding optimum uniformity of infiltrated water and minimum deep percolation and runoff from the end of the field. Concerning seed-fruit crops in North-Patagonian region, southern Argentine, regardless the physiological stage of crop and soil water deficit, farms are flooded responding to the implicit irrigation calendar (see Figure 1), based on rotating farmers' rounds. This practices address to soils lixiviation and its consequent nutrient content loss. Thus, irrigation scheduling has been an important topic in agricultural research for several decades and continues to be so today. There is a vast amount of literature on irrigation scheduling methods (Botes et al., 1996; Bergez et al., 2002). Biophysical models linked with decisional models can be a helpful tool to solve irrigation scheduling problems. They constitute and important aid in understanding complex systems and help problem solvers to capture and reason about the essential features and dynamics of such systems (Keppens, 2001).

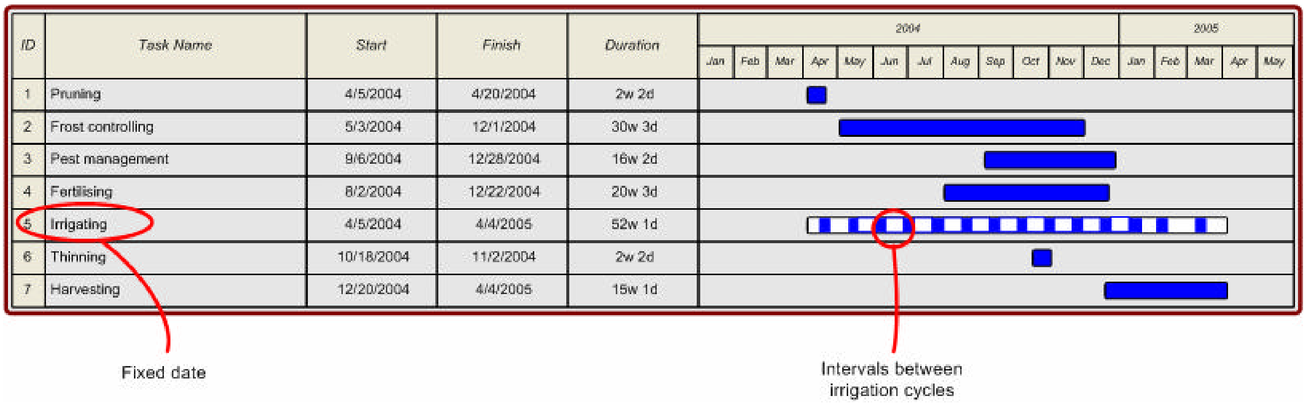

Figure 1: Description of the current strategy applied in some seed-fruit cultivar areas located Noth-Patagonian region (AR).

Current models of surface irrigation are insufficiently flexible to allow theoretical study of innovative irrigation techniques, and not always able to complete a simulation. Most of them are limited to single furrows, or border-strips with zero cross-slope and a uniformly distributed inflow at the upstream end. But large basins are currently being irrigated from a single inlet. The flow spreads out in all possible directions, and any one-dimensional simulation of the distribution of infiltrated water must be viewed as a very coarse approximation. Also, departures in a basin surface from a theoretical plane influence the flow; with the irrigation stream concentrated in the lower-lying areas, this can significantly affect infiltration uniformity. Only a two-dimensional model can simulate these factors with any accuracy. While a one-dimensional approach is suitable for furrows, in real situations flows in neighbouring furrows of a set are often coupled through common head and tail-water ditches. In some cases, the tail-water from a fast furrow will enter a slower furrow from its tail end and modify its ultimate infiltration profile. To fully appreciate the effects of such coupling, simulation of interconnected multiple furrows is necessary. On the other hand, developing tools by integrating knowledge in models may be of some help for planners. However, modelling water management is complex because it concerns different scales (scale of decision; scale of action; scale of policy; scale of management) and different actors (water users including farmers, factory managers and general public, policy makers and water manager). Crop models which simulate the dynamic of plant growth and water demand of one of several crops can provide quantitative contributions to the environmental impact assessment and be very useful for water management (Bergez et al., 2002). However they do not explicitly represent farmers' decision.

The objective of current development is to provide a new approach for optimising the parameters of seed-fruit irrigation strategies, represented by a simulation-optimisation partitioning method, SOP-κP, and validated simulation models capable of providing quick results for a wide variety of test combinations of design and management parameters.

One of the keys to improving agricultural water management is therefore to better understand the way farmers manage their irrigation and to model it (Cox, 1996; Bergez, 2002). Decisional models have to be based on decision rules in order to integrate how farmers adapt management to context. For the farmers, combining a decision model with a crop model makes it possible to calculate the best strategy that optimises water consumption and maximises revenue. For the planner, such models can be used to mimic farmer strategy (optimal or not) in order to anticipate water demand. Thus, irrigation strategies can be represented as sets of decision rules with lots of parameters (thresholds, quantities, dates, etc.). The design of innovative strategies that perform well for some specified agro-environmental criteria thus leads to large optimisation problems. Thus, we propose a new simulation-based optimisation approach that handles this difficulty using a generative partitioning clustering algorithm called SOP-κP.

In the first part of this work, we introduce the irrigation model and sub-models associated. In a second part, the solution procedures are illustrated. Then, model validation and testing is presented. Finally, results, conclusions and further researches are discussed.

2 MODELLING SEED-FRUIT IRRIGATION SYSTEMS

Modelling the whole system comprises complex relationships among different sectors and factors. We could think our system just like a complex hierarchical model. Each sub-system has a series of variables that could be treated considering Principal Components Analysis (PCA), sub-systems variables could be related to each other by means of Multiple Correspondence Analysis (MCA) and, systems and sub-systems should be analysed using Hierarchical Ascendant Clustering (HAC)), before to start to modelling tasks. Furthermore, the differential equations describing the biological process were discretised using a Global Hybrid Mixed Finite Element method (GHMFE) proposed by authors. All this results can not be included in this paper, since it represents a complete work itself. However, we present some structural description about the origin of the involved variables and their incorporation in a 'grey box' model. This analysis was carried out using MATLAB tools.

Biological simulation models, underlying bioeconomic models, vary considerably in their complexity and depend critically on the purpose for which they are constructed. An agricultural model representing pip fruit orchards is a particularly useful means of evaluating the effects on yield and profitability.

In order to obtain a rigurous model an infiltration, groundwater flow and nutrient uptake in variably-saturated porous media are developed. A numerical model for their simulation is presented. The algorithm consists in a discretisation of Richards' equation that combines a temporal linearisation using a Picard iteration with spatial approximation employing a hybridised mixed finite element procedure. The method is computationally efficient and mass conservative. Some relevant features of the associated algebraic problem and a numerical example of infiltration in North-Patagonian (AR) region are also included. The uptake of a single nutrient for root of crops is studied through a moving boundary model, which differ of previous models solving the problem in fixed domains, and then discretised via finite elements too.

2.1 MODEL DATA SETS

We faced management of the water resources for agricultural activities in North-Patagonian region. The rivers Neuquén and Limay converge forming the Black river. The river's water supply is utilised for irrigation, municipal, hydropower, industry, mining, recreation and environmental purposes. Its operation is governed by the Basins Inter-jurisdictional Authority (In Spanish: Autoridad Interjurisdiccional de Cuencas: AIC) which generally includes an inter-jurisdictional treaties, interstate legislation, etc.

Although the current state of the art does not allow accurate long range prediction of the climatic extremes in this region, stochastic hydrologic modelling can help managers get a better understanding and appreciation of the types of climate conditions they may face in the future and update scheduling irrigation systems. Operational studies of the North-Patagonian region require the consideration of statistical variability of the stream flow data. For this purpose, a number of techniques have been suggested and used in the past ranging from empirical procedures based on the historical record alone and the so called index sequential algorithm, introducing a stochastic simulation-optimisation partitioning method, SOP-κP, to refined techniques based on stochastic methods.

For this analysis, we began with 20 years of historically observed monthly data at 23 sites in the basin. This number of different sites has been taken considering that furrow layout varied considerably between field sites. The data have been naturalised in order to remove the effect of regulation or diversions. Then the 23 site system was partitioned into a system comprised of key steps, sub-steps and subsequent steps. Some infiltration characteristics for the irrigation events were estimated using the volume balance optimisation method from advance data and also with small area site infiltration tests on the high water-table site. Single site and multi-site models and aggregation and disaggregation techniques were utilised in order to determine stochastic daily stream flows at all sites. Like the original record, the stochastic traces were 20 years in length.

2.2 MODEL IDENTIFICATION

The basic state-space models are the following ones:

Discrete-Time Innovations Form:

$$\begin{aligned} x(kT+T) &= Ax(kT) + Bu(kT) + Ke(kT) \\ y(kT) &= Cx(kT) + Du(kT) + e(kT) \\ x(0) &= x_0 \end{aligned} \tag{1}$$

where T is the sampling interval, $u(kT)$ is the input at time instant kT, and $y(kT)$ is the output at time kT. (Ljung ,1999).

System Dynamics Expressed in Continuous Time:

$$\dot{x}(t) = Fx(t) + Gu(t) + \tilde{K}w(t)$$
$$y(t) = Hx(t) + Du(t) + w(t) \tag{2}$$
$$x(0) = x_0$$

We define a parameterised state-space model in continuous time due to physical laws are described in terms of differential equations. The matrices F, G, H, and D contain elements with physical significance. The numerical values of these are not be known. To estimate unknown parameters based on sampled data (assuming T is the sampling interval) we first transform (1) to (2) using the formulas

$$A = e^{FT} \qquad B = \int_0^T e^{Ft} G dt \qquad C = H \tag{3}$$

The value of the Kalman gain matrix K in (1) or in (2) depends on the assumed character of the additive noises and in

$$x(t+1) = Ax(t) + Bu(t) + w(t)$$
$$y(t) = Cx(t) + Du(t) + e(t) \tag{4}$$

and its continuous-time counterpart, where $w(t)$ and $e(t)$ are stochastic processes with certain covariance properties. This gives the directly parametrised innovations form. (Ljung, 1999). Taking into account the internal noise structure is important, we define a grey-box structure using the discretised physical laws contained in the infiltration and nutrient uptake models.

2.3 BIOPHYSICAL MODEL

2.3.1 INFILTRATION AND GROUNDWATER MODEL

Prediction of water movement in variably-saturated porous media is an important problem in many branches of science and engineering. The water motion is assumed to obey Richards' equation. This equation may be written in terms of pressure head (π-based form) or water content (ω-based form) as the dependent variable. Only the π-based form of the equation can be used for simulating water flow in soils with saturated regions, but unfortunately these models are inherently non-mass conserving. A greatly improved of the performance of π-based models can be made by using an appropiate temporal discretisation of a mixed form of Richards' equation. The approximations that are usually applied to the spatial domain are finite differences and finite element standard methods. In this work we use a numerical model to solve the mixed form of Richards' equation based on a global hybridised mixed finite element procedure. The algorithm produces perfectly mass conservative numerical solutions and it is computationally efficient.

We will consider the numerical simulation of underground water flow in a porous media domain $\Theta_i = (0,1)$, $i = 1,\ldots,p$, with boundary $\partial\Theta_i = \mathrm{M}^B \cup \mathrm{M}^T$, where $\mathrm{M}^T = \{z_i = 1\}$, $i = 1,\ldots,p$, $\mathrm{M}^T :\rightarrow \Re^p$. It will be assumed that water flow obeys the Richards' equation stated in the form:

$$i) \qquad \frac{\partial w(p)}{\partial t} + \nabla \cdot \vec{f} = 0 \qquad z_i \in \Theta_i \tag{5}$$

$$ii) \qquad \vec{f} = -\sum_{i=1}^{p} K_i(p) \cdot \nabla(p + z_i) \cdot d_i \qquad z_i \in \Theta_i$$

where ω and π are water content and pressure head, respectively; K_i is the hydraulic conductivity, which is assumed independent of p for saturated soils but varies strongly with p in unsaturated soils; z_i denotes the spatial dimensions; and t is time.

Equation (5.i) states conservation of mass for the water phase and (5.ii) defines the water flux $\vec{f}$ in terms of Darcy's law. Equations (5) are valid under the following assumptions: the porous media is undeformable; the water density remains constant; and the air mobility is much greater than the water mobility so that the air remains at essentially atmospheric pressure. We will consider solving (5) with the following boundary conditions:

B.C.1: $\vec{f}\cdot\vec{v} = f_{in}(t)$ on M^T (6)

B.C.2: $\vec{f}\cdot\vec{v} = f_{out}(t)$ on M^B

The function $f_{in}(t)$ represents the rainfall data, while the term $f_{out}(t)$ is used to represent the effect of the regional flow.

To solve the differential problem (5)-(6) we also need additional relations between the dependent variables w and p. We will use the following water retention and hydraulic conductivity models proposed by van Genutchen (Guarracino and Santos, 1997):

$$w(p) = \frac{w_s - w_r}{\left[1+(a|p|)^n\right]^m} + w_r \tag{7}$$

$$K_i(p) = K_{i,s}\frac{\left\{1-(a|p|)^{n-1}\left[1+(a|p|)^n\right]^{-m}\right\}^2}{\left[1+(a|p|)^n\right]^{m/2}}$$

where $m = 1 - 1/n$; w_r and w_s are the residual and saturated water contents, respectively; $K_{i,s} = K_{i,s}(z_i)$ is the saturated hydraulic conductivity; a and n are model parameters related to soil properties.

2.3.2 NUTRIENT UPTAKE MODEL

The nutrient uptake could be estimated through soil transport models coupled with absorption kinetics Michaelis-Menten-like on the root surface. The equations involved could be solved over fixed domains. The objective of this work consists in estimate the nutrient uptake considering rizospheric variable domains –variable root density- implementing a moving boundary model.

Now, we will consider the numerical simulation of the nutrient uptake flow in a porous media domain $\Theta_i = (0,1)$, $i = 1,\ldots,p$, with boundary $\partial\Theta_i = \mathrm{M}^B \cup \mathrm{M}^T$, where $\mathrm{M}^T = \{z_i = 1\}$, $i = 1,\ldots,p$, $\mathrm{M}^T :\rightarrow \Re^p$. The main differential equations can be stated in the form:

$$i) \qquad \frac{Dx(g+z_i,t)}{Dt} - D_i\nabla^2 x(g+z_i,t) - D_i\left(1+\frac{\vec{v}(t)s_0}{D_i b}\right)\nabla x(g+z_i,t) = 0 \qquad z_i \in \Theta_i \tag{8}$$

$$ii) \qquad D_i\nabla x((s(t),t)+\vec{v}\cdot x(s(t),t)) = \frac{k[x(s(t),t)-x_{th}]}{1+\frac{k[x(s(t),t)-x_{th}]}{J_{\max}}} = c\cdot x(s(t),t)\frac{Ds(t)}{Dt} \qquad z_i \in \Theta_i$$

where x is the nutrient concentration; r is the radial distance from the root axis; t is time; T is the maximum time which the equations have solution for; $s(t)$ represents the moving boundary; x_{th} is the threshold concentration under which the absorption is nule; $\vec{v}$ is the mean effective velocity of soil-solution over the root surface; b is the buffer potential related to a given nutrient; $J_{\max}$ is the maximum influx of nutrient; $k = J_{\max}/K_{\max}$, where $K_{\max}$ is the Michaelis constant, represents the root absorption potential; R is the rizospheric radium; j is the initial profile concentration; D_i is a diffusivity coefficient in the direction i; c is a dimensionless stoicheometric coefficient; and s_0 is the initial radium.

These equations can be re-written introducing the nutrient flux vector $\vec{f}(s(t))$ as follows:

$$\vec{f}(s(t)) = \frac{k[x(s(t),t)-x_{th}]}{1+\frac{k[x(s(t),t)-x_{th}]}{J_{\max}}} = c\cdot x(s(t),t)\frac{Ds(t)}{Dt}$$

Then

$$i) \qquad \frac{D\vec{f}(s(t))}{Dt} - D_i\nabla\vec{f}(s(t)) - D_i\left(1+\frac{\vec{v}(t)s_0}{D_i b}\right)\vec{f}(s(t)) = 0 \qquad z_i \in \Theta_i \tag{9}$$

$$ii) \qquad \vec{f}(s(t)) = D_i\nabla x((s(t),t)+\vec{v}\cdot x(s(t),t)) \qquad z_i \in \Theta_i$$

Equation (9.i) represents the mass and diffusive nutrient transport, (9.ii) show a mass balance over the root where the ions are incorporated considering a kinetic Michaelis-Menten type. Equations (9) are valid under the assumptions taken for the infiltration model. We will consider solving (9) with the following boundary conditions:

B.C.1: $x(g+z_i,0) = j(g+z_i) \qquad z_{i,0} \le z_i \le Z_i \qquad 0 < t < T \qquad z_i \in \Theta_i \qquad$ on M^T (10)

B.C.2: $-D_i \nabla x(\vec{s}(t)+Z_i - z_0, t) + \vec{v}(t) x(\vec{s}(t)+Z_i - z_0, t) = 0 \qquad z_i \in \Theta_i \qquad$ on M^B

while (10.i) is the initial profile concentration, and (10.ii) is the border condition considering a nule flux (it can be enter water but no nutrient). This equation also represents a moving boundary of constant thick. Furthermore

$$s(t) = s_0 \sqrt{e^{kt}} \qquad\qquad z_i \in \Theta_i$$

representing a growth radial law. For fixed soil volumes, the root length $l(t)$ is linked with the root radium by means of the expresion:

$$l(t) = l_0 \left(\frac{s(t)}{s_0} \right)^2$$

being l_0 the initial root length.

Thus, to solve the differential problem (9)-(10) we also need additional relations between the dependent variables x and g. We will use the following water retention and hydraulic conductivity models proposed by van Genutchen:

$$x(g(t)) = \frac{x_s - x_r}{\left[1 + (a|g|)^n\right]^m} + x_r \tag{11}$$

$$D_i(g) = D_{i,s} \frac{\left\{1 - (a|g|)^{n-1}\left[1 + (a|g|)^n\right]^{-m}\right\}^2}{\left[1 + (a|g|)^n\right]^{m/2}}$$

where $m = 1 - 1/n$; x_r and x_s are the residual and saturated water contents, respectively; $D_{i,s} = D_{i,s}(z_i)$ is the saturated hydraulic conductivity; a and n are model parameters related to soil properties.

2.4 MODEL CONSTRAINTS

The irrigation model is subject to the following constraints:

Hydrology model:

$$g_i(v_j, t'_j) = \sum_{i=1}^{I} \sum_{j=1}^{J} t'_{i,j} - t'_{i,0} \le 0 \qquad \forall i \in I \qquad \text{weather constraints}$$

(12)

Production model:

$$g_q\left(v_j, h_j\right) = \sum_{q=1}^{Q}\sum_{j=1}^{J} c_q\left(v_j, h_j\right)h_j - c_{q,0} \le 0 \qquad \forall q \in Q \qquad \text{equipment constraints} \tag{13}$$

$$g_r\left(v_j, h_j\right) = \sum_{r=1}^{R}\sum_{j=1}^{J} c_r\left(v_j, h_j\right)h_j - c_{r,0} \le 0 \qquad \forall r \in R \qquad \text{resource constraints} \tag{14}$$

$$g_l\left(v_j, h_j\right) = L\left(v_j, h_j\right) + \sum_{l=1}^{L}\sum_{j=1}^{J} c_l\left(v_j, h_j\right)h_j - c_{l,0} \le 0 \qquad \forall l \in L \qquad \text{legal constraints} \tag{15}$$

Biophysical model:

$$g_e\left(v_j, t'_j\right) = \sum_{h=1}^{H}\sum_{e=1}^{E}\sum_{j=1}^{J} p_h(w,p,x,g)\,b_e\left(v_j, t'_j\right)t'_{e,j} - t'_{e,0} \le 0 \quad \forall e \in E \qquad \text{ecologic constraints} \tag{16}$$

where c_q, c_r, c_l are cost functions, w, p, x, g are the biophysical model variables, and b_e represent an ecological index related to climate τ'_j and using the strategy ϖ_j. The efficiency under climate τ'_j and using the strategy ϖ_j could be defined in several ways; we chose an efficiency given by

$$h_j = c_j \cdot \frac{Yd\left(v_j, t'_j\right)}{Q\left(v_j, t'_j\right)} \tag{17}$$

where χ_j represents a production activity coefficient, Yd(ϖ_j , τ'_j) is the seed-fruit yield obtained under climate τ'_j and using the strategy ϖ_j , and Q(ϖ_j , τ'_j) is the amount of water used under climate τ'_j and using the strategy ϖ_j. It can be observed that if this coefficient take the value one, a direct relation between yield and water applied under climate τ'_j and using the strategy ϖ_j is expected.

2.5 BIOECONOMIC MODEL

The objective function to be maximised is the expectation of the direct margin (i.e., the gross margin minus specific costs for a given activity, here irrigation) over all the years of weather recording. The direct margin regarding irrigation may be written as:

$$OF :\rightarrow Z\left(v_j, t'_j\right) = Yd\left(v_j, t'_j\right)\cdot Pc - \left[Cp + Q\left(v_j, t'_j\right)\cdot c + t\left(v_j, t'_j\right)\cdot p\right] \tag{18}$$

The whole bioeconomic model consists in maximising objective function (18) subject to model equations (1) - (11), (17) and model constraints (12) – (16).

where Z(ϖ_j, τ'_j) is the direct margin for climate τ'_j and using the strategy ϖ_j, that is, the total revenue for climate τ'_j and using the strategy ϖ_j, Z_T, minus the base revenue using the current strategy, Z_0.

$$Z\left(v_j, t'_j\right) = Z_T\left(v_j, t'_j\right) - Z_0\left(v_{CS}, t'_0\right) \tag{19}$$

Yd(ϖ_j , τ'_j) is the seed-fruit yield obtained under climate τ'_j and using the strategy ϖ_j, Pc is the selling price for seed-fruit, Cp is the operational cost for seed-fruit production, Q(ϖ_j , τ'_j) is the amount of water used under climate τ'_j and using the strategy ϖ_j, c is the cost of irrigation water,

$t(\varpi_j, \tau'_j)$ the number of irrigation cycles performed and π is the cost of carrying out a new irrigation cycle.

3. SOLUTION METHODS

3.1 GLOBAL HYBRIDISED MIXED FINITE ELEMENTS PROCEDURE

We solved the biophysical model described in section 2.3 in a rigurous way by means of a global hybrid mixed finite element procedure. A temporal discretisation of the physical laws equations using a backward Euler method coupled with a Picard iteration scheme was done. Then we defined a spatial approximation using a global hybridised mixed finite element procedure. It can be demostrated that the problem obtained has a unique solution. Thus, we solved the algebraic associated problem using Lagrange multipliers.

3.2 OPTIMISATION PROCEDURES

It is generally acknowledged that there are two main families of clustering (unsupervised classification) methods: hierarchical and partitioning. The former ones create a tree structure splitting (reuniting) the initial set of objects in smaller and smaller subsets, all the way to singletons (and reverse), while the latter ones construct a partition of the initial set of objects into a certain number of classes, with the target number usually part of the input, along with the objects themselves. Most partitioning methods proposed for data mining can be divided into: discriminative (or similarity-based) approaches and generative (or model-based) approaches. In similarity-based approaches, one optimises an objective function involving the pair-wise data similarities, aiming to maximise the average similarities within clusters and minimise the average similarities between clusters. A fundamentally different approach is the model based approach which attempts to optimise the fit (global likelihood optimisation) between the data and some mathematical model, and most researchers do not consider them as clustering methods. Similarity-based partitioning clustering is also closely related to a number of operations research problems such as facility location problems, which minimise some empirical loss function (performance measure). There are no efficient exact solutions known to any of these problems for general number of clusters *m*, and some formulations are NP-hard. Given the difficulty of exact solving, it is natural to consider approximation, either through polynomial-time approximation algorithms, which provide guarantees on the quality of their results, or heuristics, which make no guarantees.

Thus, our work develops a Generative Partitioning Clustering (GPC) method. The irrigation management strategies can be defined as a stochastic optimisation problem, such as

Find $f^{*} = \arg\max\{\mathrm{E}(Z(f, t'))\}$ with $h \in H$ (20)

subject to irrigation model and model constraints, and where H is the value domain of ϕ, and E(Z) is the expected value of the gain margin Z, that depends on the irrigation strategy ϕ and on some random exogenous variable τ' like the weather. This problem has been faced by means of two different optimisation approaches: control and simulation statements.

3.2.1 CONTROL STATEMENTS

At first, we considered optimisation of just a single decision rule that determines the start of an irrigation campaign. We compare stochastic dynamic programming and reinforcement learning methods for identifying the optimal decision rules. For both optimisation methods the consequences of a strategy were calculated using the biophysical model, coupled with a stochastic weather generator.

Optimal decision rules were derived using two different procedures. The first relies on a discretisation of the domains d and t of the state variables Đ and T, where d and t are the two state-variables, being d the soil water deficit, and t the accumulated thermal units, and the range of d and t are the intervals Đ and T, respectively.. The discrete transition probabilities are then estimated by simulation, and an approximated numerical solution of the optimal solution is obtained by dynamic programming (Kennedy, 1986) reinforcement learning, which does not require an a priori estimation of the transition probabilities.

3.2.2 SIMULATION STATEMENTS

A more flexible and realistic formulation of the optimisation problem consists of searching by means of simulation for the best values for the parameters of a predefined irrigation strategy. In this case, the simulation model is considered as a black box function where the vector of strategy parameters $\varpi = (\varpi_1, \varpi_2, \ldots, \varpi_q)$ is taken as the input variable and where the output is the objective function Z. In most of the simulation models of agricultural production systems, the objective function Z also depends on the weather, which has to be considered as an unknown and uncontrollable random variable τ'. Optimising a strategy thus consists of searching for the set of parameters ϖ^* that maximises the expected value of the objective function according to the random weather series:

$$\text{Find} \quad v^* = \arg\max\{\mathrm{E}(Z(v, t'))\} \quad \text{with} \quad v \in G \tag{21}$$

subject to irrigation model and model constraints, and where G is the value domain of ϖ, and using the criterion $E(Z(\varpi))$ for estimation by averaging the objective function $Z(\varpi, \tau')$ over a large number of sampled variables τ'_j:

$$E(Z(v, t')) \approx \frac{\sum_j Z(v, t'_j)}{N} \tag{22}$$

Several efficient methods (stochastic approximation, random search, stochastic branch-and-bound, etc.) have been recently developed for solving this stochastic simulation optimisation problem, which is one of the most difficult problems of mathematical programming (Azadivar, 1999; Fu, 2001; Sutton and Baro, 1998). Stochastic iterative methods for deterministic optimisation problems like genetic algorithms, simulated annealing or tabu-search have also been adapted for the stochastic environment associated with simulation.

We develop a new method for stochastic simulation optimisation problems (SOP-κP) with continuous variable space $G \in R^{\kappa}$, where we assume that parameters τ'_j are bound-constrained, which means that the domain G is a hyper-cube on R^{κ}. This algorithm is designed to completely explore (by sampling) the domain G, in order to find local and global optimal values for $\varpi \in G$. for small values of κ, the simplest method for such problems is a systematic grid-search (SGS), which consists of estimating the objective function on all points located on finer and finer grids over G, until a maximum precision is reached for each variable. For large dimension problems ($\kappa > 2$) this method is not efficient since the number of grid points grows exponentially with κ. A more useful approach is therefore to prioritise the evaluation of points ϖ within promising cells of G that are supposed to contain optimal solutions. Its main advantage is to maximise, for a given budget of time or simulation runs, the chance of finding a good solution.

Therefore, SOP-κP is used to optimise the set of the eight parameters using weather records of North-Patagonian region in Southern Argentine from 1984 until 2004.

4 MODEL VALIDATION AND TESTING

Once stated, the model must be validated using a portion of field data. The analysis was carry out in MATLAB environment. Figure 2 shows a comparison between simulated model and field data sets. The fit percentage parameter is 99.91 %; a fitting quite satisfactory.

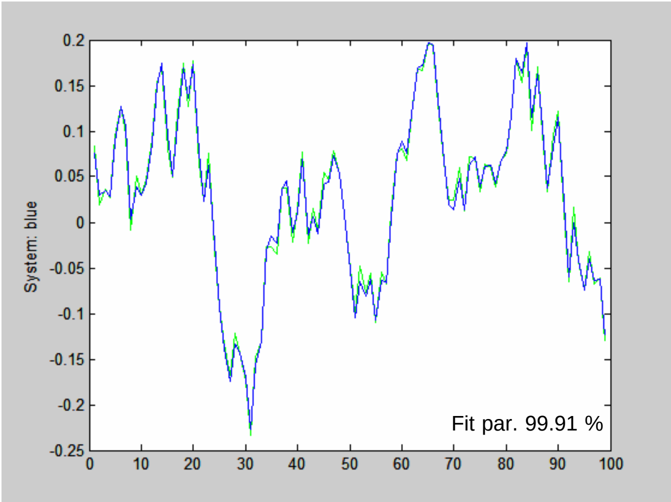

Figure 2: Comparison between simulated model (blue) and field data (green). Satisfactory agreement is achieved.

To demonstrate the usefulness of the developed algorithm we will simulate a basic strategy (BS). Then, a comparison between the systematic grid-search method (SGS) and the SOP-κP algorithm will be carried out on two parameters on the BS: the soil water deficit (δ) and the time (expressed as accumulated thermal units) to start the main irrigation campaign (λ). Water applied per cycle is defined as α. Thus we have λ_s (accumulated thermal unit to start the irrigation campaign), δ_s (soil water deficit to start the irrigation), α_s (irrigation applied at the first irrigation), δ_n (soil water deficit to start new irrigation cycle), α_d (irrigation depth applied after the first irrigation round), λ_p (accumulate thermal unit to stop the irrigation), δ_p (soil water deficit to stop the irrigation), α_p (irrigation applied at the last irrigation round).

4. RESULTS

Table 1 gives the ranges of the eight parameters that lead to the maximum expectation of the objective function. Irrigation has to start at 426 °C days and a 61 mm soil water deficit. The amount applied is then 59 mm. The next irrigation is due when the soil water deficit reaches 73 mm; 52 mm of irrigation water is then applied. Finally, the last irrigation cycle is performed if at 1307 °C day the soil water deficit is more than 116 mm; 59 mm of water irrigation is then applied.

Table 1: Eight-parameters values that maximise the gain margin. L.R.: Lower Range, U.R.: Upper Range.

	λ_s	λ_p	δ_s	δ_n	δ_p	α_s	α_n	α_p
L.R.	426 °C day	1307 °C day	61 mm	73 mm	116 mm	59 mm	52 mm	59 mm
U.R.	467 °C day	1366 °C day	73 mm	82 mm	134 mm	71 mm	59 mm	71 mm

The simulation of this strategy leads to a seed-fruit yield of 7.49 Tons ha^{-1}, and the margin is 4866.79 U$S ha^{-1}. Water used 138 mm.

Table 2 shows the results corresponding to different strategies and parameters utilised in our analysis. It comprises mean and variance data in order to get a useful comparison among them. The average selling price for seed-fruit is assumed to be 697.08 U$S Tons^{-1}. The seasonal operational costs (seed or plant-tree, weeding, fertiliser, insurance) are assumed to be 253.81 U$S ha^{-1}. The cost of irrigation water is assumed to be 0.59 U$S mm^{-1} and setting up a new irrigation cycle 6.37 U$S.

Table 2. Results from the simulation comparing different strategies. CS: current strategy, BS: basic strategy, 2-SGS: systematic grid-search with two parameters, SOP-2P/SOP-8P: simulation optimisation partitioning method for two and eight parameters, respectively.

(ϖ_j , τ'_j)	(μ, σ^2)	$Yd(\varpi_j , \tau'_j)$ [Tons ha^{-1}]	$Q(\varpi_j , \tau'_j)$ [mm]	$t(\varpi_j , \tau'_j)$ [mm]	$Z(\varpi_j, \tau'_j)$ [U$S ha^{-1}]	η_j [Tons ha^{-1} mm^{-1}]
CS	μ	3.89	190	1	2339.36	0.020
	σ^2	0.06	2.19	0	2.31	0.002
BS	μ	5.59	174	2	3527.47	0.032
	σ^2	0.09	1.46	0.01	3.01	0.002
2-SGS	μ	6.60	162	3	4232.23	0.041
	σ^2	0.08	2.07	0.02	3.28	0.002
SOP-2P	μ	6.67	153	3	4286.33	0.044
	σ^2	0.06	2.08	0.02	3.47	0.002
SOP-8P	μ	7.49	138	3	4866.79	0.054
	σ^2	0.09	1.75	0.02	3.25	0.002

As we can see, the highest gain margin is obtained with SOP-κP using eight parameters. In the Figure 3 a scheduling scheme is presented, showing duration and intensity of the irrigation.

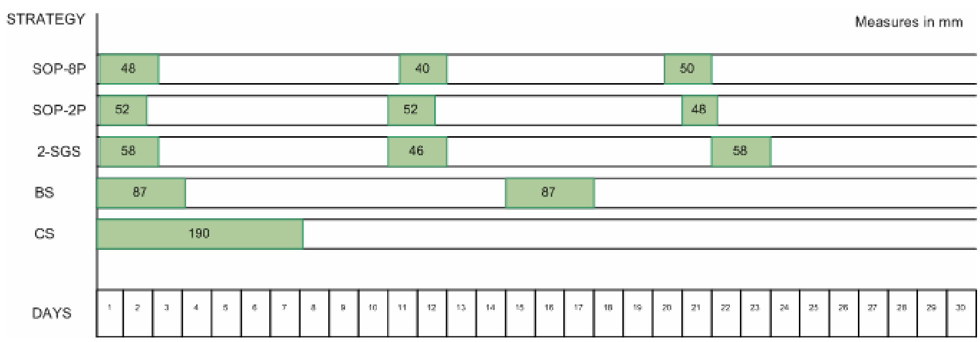

Figure 3: Monthly scheduling comparing different strategies.

Figure 4 allows a useful visualisation of the eight-parameter strategy. Intermediate cut planes covering the range of the eight-parameters that maximise the expectation of the objective function calculated with SOP-κP are illustrated. Note that this time-parameterised figures correspond to a three-dimensional graphic where the intersections between planes and surfaces give the moment in which water should be applied.

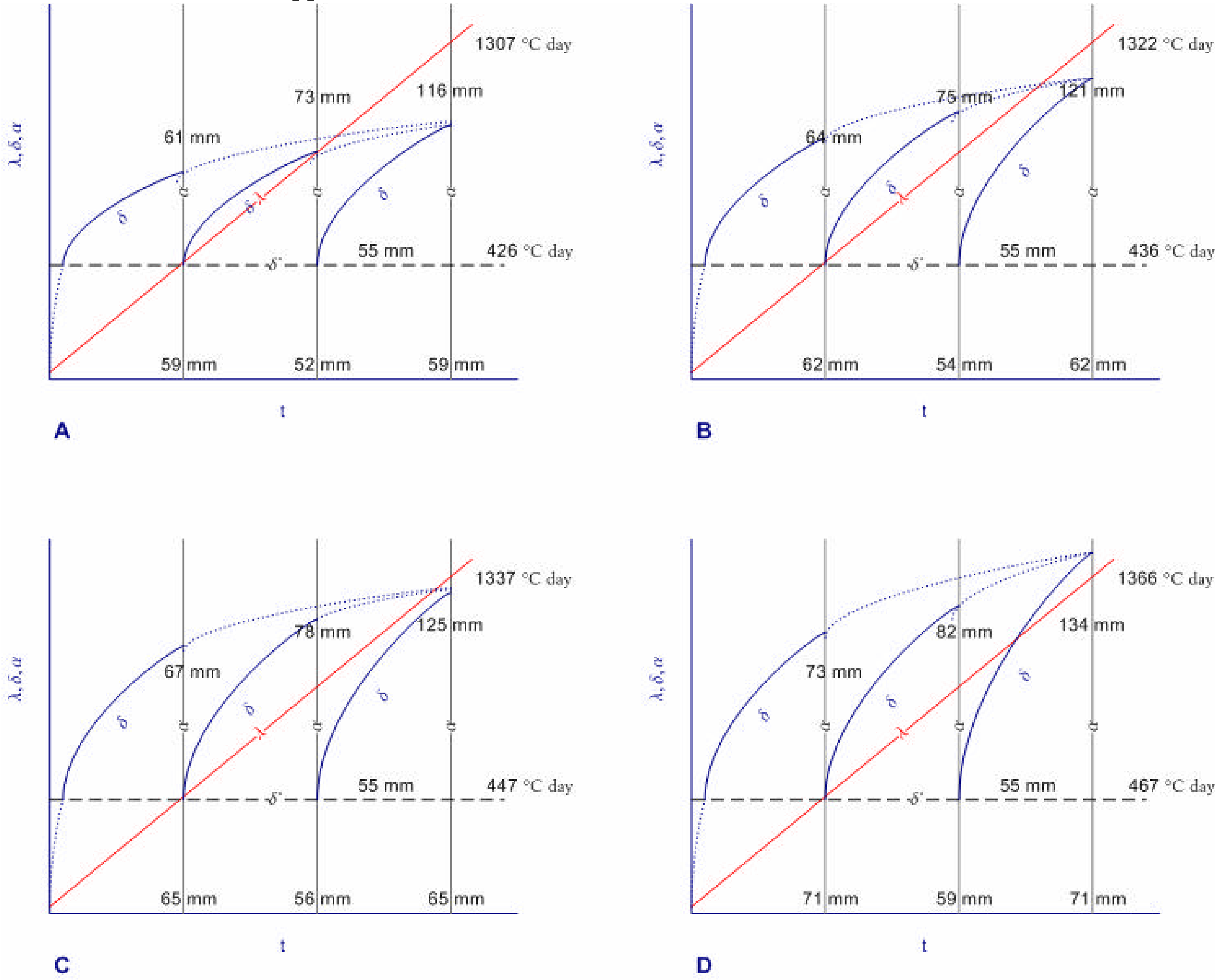

Figure 4: Cut planes covering the range of the eight-parameters maximising the expectation of the objective function calculated with SOP-κP. A: lower values, B-C: middle values, D: upper values.

6 DISCUSSION AND FURTHER RESEARCH

The system was characterised as a 'grey-box' model, getting the opportunity to specify biophysical parameters in a practical way exploring the underlying behaviour of the biophysical systems (Cox, 1996). Thus, possible uses of the model through simulation are to explore:

- the effect of the soil physical properties on soil water dynamics;
- crop growth response to different climatic conditions;
- crop growth response to different irrigation strategies;
- the effect that root depth has on the soil water dynamics;

- how different factors influence the amount of water infiltrating below the root horizon, so affecting leaching of nutrients or the rise of the water table;
- sustainable irrigation and cropping strategies to reduce run-off and through drainage.

Comparing different strategies, we can observe that the use of SOP-κP to parameterise the decision rules improved the direct gain margin. From the average 2339.36 U$S ha^{-1} obtained with the current strategy, we reach 4866.79 U$S ha^{-1} using SOP-8P.

Factors considering by the development of this work was the bio-physical model used to describe the system behaviour, the decisional model and the variables used to trigger the decision, soil and weather and, constraints. Taking into account that, by definition, a model is an imperfect representation of reality (Whisler et al., 1986; Boote et al., 1996) is quite important that the bio-physical model be both robust and sensitive. On the other hand, it is necessary to describe the parameterised strategy as a set of decision rules in order to use SOP-κP, leading to an interdisciplinary work among different sectors, such as farmers' cooperatives, irrigation advisors, pertinent authorities, etc. In this paper, this algorithm was used to calculate optimal strategies for the direct gain margin criterion, but it also be utilised as an improvement tool, re-optimising each set of parameters.

Considering that a number of assumptions made by the model were not always met in the field, the results indicate that in most cases the model simulates irrigation events reasonably well on such irrigated systems. This provides evidence that the model can be used with confidence as a robust predictive tool for investigating system parameter interactions and their effects on application efficiencies and water distribution uniformities.

ACKNOWLEDGEMENTS

We would like to thanks to UNS - CONICET for their financial support and to the members of the INTA - Alto Valle Experimental Station and AIC Authorities, for providing us reliable information about climate statistical reports and basin features.

REFERENCES

Azadivar, F. 1999. Simulation optimization methodologies. In: Proceedings of the 1999 Winter Simulation Conference, December 5-8, 1999. Squaw Peak, Phoenix, AZ, USA, pp. 93-100.

Bergez, J.E., Deumier, J.M., Lacroix,B., Leroy, P, Wallach, D. 2002. improving irrigation schedules by using a biophysical and a decisional model. Eur. J. Agron. 16, 123-135.

Boote, K.J., Jones, J.W., Pickering, N.B. 1996. Potential uses and limitations of crop models. Agron. J. 88, 704-716.

Botes, J.H.F., Bosch, D.J., Oosthuizen, L.K. 1996. A simulation and optimization approach for evaluating irrigation information. Agric. Syst. 51, 165-183.

Cox, P.G. 1996. Some issues in the design of agricultural decision support systems. Agric. Syst. 52,355-381.

Cros, M.J., Garcia, F., Martin-Clouaire, R., Duru, M. 2001. Simulation optimization of grazing management strategies. In: Proceedings of the 3rd European Conference on Information Technology in Agriculture (EFITA'01), Montpellier (FR).

Fu, M.C. 2001. Simulation Optimization. In: Proceedings of the 2001 Winter Simulation Conference, December 9-12, 2001, Crystal Gateway Marriott, Arlington, VA, USA, pp. 53-61.

Guarracino, L., Santos, J. E. 1997. Finite element in flow modelling. Comp. Mech. 13: 597-603.

Howell, T.A. 2001. Enhancing water use efficiency in irrigated agriculture. Agron. J. 93, 281-289.

Huyer, W., Neumaier, A. 1999. Global optimization by multilevel coordinate search. J. Global Optim., 331-355.

Keppens, J., Shen, Q. 2001. On compositional modelling. Knowledge Engineering Review 16(2): 157-200.

Ljung, L 1999. System Identification - Theory For the User, 2nd ed, PTR Prentice Hall, Upper Saddle River, N.J., pp 93-99.

Mayer, D.G., Belward, J.A., Burrage, K. 1998a. Optimizing simulation models of agricultural systems. Ann. Oper. Res. 82, 219-231.

Mayer, D.G., Belward, J.A., Burrage, K. 2001. Robust parameter settings of evolutionary algorithms for the optimisation of agricultural system models. Agric. Syst. 69, 199-213.

Salas, J.D., 1993. "Analysis and Modeling of Hydrologic Time Series", In Handbook of Hydrology, D.R. Maidment Editor, McGraw Hill Inc., New York.

Sutton, R.S., Barto, A.G. 1998. Reinforcement learning: an introduction. MIT Press. Cambridge.

Walker, D. 2002. Decision support, learning and rural resource management. Agric. Syst. 73, 113-127.