New Approaches for Teaching Irrigation Scheduling — Improving Understanding through Modeling Applications

Pau Martí, Jalal Shiri, María Gasque, and Pablo González-Altozano

Abstract—**This paper presents a teaching methodology for introducing specific modeling tools and their methods in the contents of the subject Irrigation Engineering, corresponding to the degree of Agricultural Engineering in Spanish universities. Modeling tools can be powerful educational resources allowing for a better understanding of the physical problems addressed in different subjects of scientific-technological careers. In particular, a Gene Expression Programming based approach is applied for estimating stem water potential, a key variable in irrigation scheduling, and, consequently, a crucial issue for Agricultural Engineers. This activity is scheduled in three practical sessions for the mentioned subject. Students learn a crucial topic in irrigation engineering and the fundamentals of a very robust modeling tool. The application of this methodology for solving a specific problem contributes to improving the understanding of the theoretical topic scheduled in the subject program. At the same time, students learn a new way of facing specific problems that might be encountered in their professional practice.**

Index Terms—**Modeling, GEP, FDR humidity probes, irrigation, agricultural engineering.**

I. INTRODUCTION

The study of crop water requirements and those technologies for an accurate and optimized dosage of irrigation water are key contents included in the curricular contents of Agricultural Engineering degree programs. Such contents are framed in subjects aiming at providing the necessary fundamentals for the design, project and management of the different irrigation systems, which are among the most important professional competencies of agricultural engineers. In practice, a suitable irrigation management is essential for an efficient water use. Indeed, a bad management might considerably worsen the water use efficiency of a well designed and equipped installation. In the frame of agricultural water management, optimized irrigation scheduling is essential to guarantee a sustainable water supply and an optimal crop growing.

In order to maximize productions, while keeping high quality standards, maintaining costs as low as possible, and preserving the environment, the development of better irrigation technologies is essential for addressing present and future water restrictions caused by an increasing water demand and limited water resources. Among them, the development of more accurate irrigation scheduling techniques and procedures is crucial for ensuring a sustainable water supply without reducing crop production and quality. Among irrigation scheduling techniques, the development of different sensors for monitoring climatic conditions, soil water content or plant water status has lead to great advances. In particular, soil water content is considered as a key parameter for planning irrigation doses [1]. The frequency domain reflectometry probes (FDR), with multiple depth capacitance sensors, have shown so far excellent performance among other techniques allowing for continuous measurement of soil water content. Indeed, they are currently widely applied in field applications. Nevertheless, the optimization of irrigation doses from FDR measurements requires accurate threshold values for individual crops, which are also site-specific. Moreover, there is no standard procedure for determining those thresholds. On the other hand, many plant physiological features respond directly to changes in the plant tissues, rather than to changes in soil water content [2]. This difficults the application of such soil moisture probes in irrigation scheduling.

Contrary to those techniques, there are other studies suggesting that plant based measurements are more reliable and accurate for irrigation scheduling [2], such as leaf water potential, which can be measured with a pressure chamber [3]. As an alternative to leaf water potential, stem water potential might be a more reliable indicator of water status and early water deficit in plants. It also presents less variability and is well related to tree and fruit growth and quality in a wide range of soils and under different irrigation systems. However, the measurement of stem water potential is destructive, as well as time and labor consuming. Further, it is not suited for automatic irrigation scheduling and control. So, the estimation of such plant-based variables turns into a task of great importance, and might allow relevant advances in irrigation scheduling methodologies.

Despite the key role played by modeling in the professional practice, it is often omitted or forgotten in the definition of the curricular contents of Spanish universities. Different proposals have been suggested in different Engineering Degrees aiming at introducing different modeling tools and their methods to overcome this educational lack. These modeling tools are not only very useful for straight professional application, but also indirectly, they are very powerful educational resources allowing for a better visualization and understanding of the physical problems addressed in the different subjects of scientific-technological careers.

Pau Martí is with the Department de Biologia, Universitat de les Illes Balears, 07122, Spain (e-mail: pau.marti@uib.cat).

Jalal Shiri is with the Department of Water Engineering, University of Tabriz, Iran (e-mail: j_shiri@yahoo.com).

María Gasque is with the Department of Applied Physics, Universitat Politècnica de València, 46022, Spain (e-mail: mgasque@fis.upv.es).

Pablo González-Altozano is with the Department of Rural and Agrifood Engineering, Universitat Politècnica de València, 46022, Spain (e-mail: pgaltozano@agf.upv.es).

This work aims at proposing a teaching methodology for introducing the application of Gene Expression Programming based models for estimating stem water potential as a complement for improving the learning of irrigation scheduling. This activity might be carried out in different subjects of the syllabus of the Agricultural Engineering Degree at Spanish Technical Universities, such as 'Irrigation Engineering'.

II. MODELING IN IRRIGATION SCHEDULING

So far, different authors have addressed different modeling problems in the field of irrigation scheduling. Focusing on stem and leaf water potential, among others, Steppe *et al.* [4] proposed a scheduling method based on modeling stem water potential and measurements of sap flow and stem diameter variation. Santos and Kaye [5] predicted leaf water potential in grapevines from near-infrared spectroscopy and multivariate analysis. Acevedo-Opazo *et al.* [6] suggested a methodology to extrapolate pre-dawn leaf water potential using linear combinations of ancillary data. Recently, Martí *et al.* [7] used artificial neural networks to estimate stem water potential from soil moisture measurements and standard climatic records.

III. GEP FUNDAMENTALS

Genetic Programming (GP) was first proposed by Koza [8] as a generalization of Genetic Algorithms (GAs) [9], which is particularly suitable where interrelationships among the relevant variables are poorly understood, a theoretical analysis is constrained by assumptions and therefore their solutions are of limited use, and there is a large amount of data in computer readable forms requiring tedious processing.

The GP algorithms firstly define an objective function as a qualitative criterion. Next, this function is used for measurement and evaluation of different solutions in a step by step manner of structural correction until GP leads to a suitable solution. GP is an evolutionary algorithm (EA) and is popular because of its high accuracy. Major advantages of GP are that it can be applied to areas where 1) the interrelationships among the relevant variables are poorly understood (or where it is suspected that the current understanding may well be less than satisfactory), 2) finding the ultimate solution is hard, 3) conventional mathematical analysis does not, or cannot, provide analytical solutions, 4) an approximate solution is acceptable (or is the only result that is ever likely to be obtained), 5) small improvements in the performance are routinely measured (or easily measurable) and highly valued, and 6) there is a large amount of data, in computer readable form, that requires examination, classification, and integration (such as satellite observations) [10].

Gene Expression Programming (GEP) is comparable to GP yet evolves computer programs of different sizes and shapes encoded in linear chromosomes of fixed lengths. The chromosomes are composed of multiple genes, each gene encoding a smaller subprogram. Furthermore, the structural and functional organization of the linear chromosomes allows the unconstrained operation of important genetic operators such as mutation, transposition and recombination. One strength of the GEP approach is that the creation of genetic diversity is extremely simplified as genetic operators work at the chromosome level. Another strength of GEP consists of its unique, multigenic nature which allows the evolution of more complex programs composed of several subprograms. As a result GEP surpasses the old GP system in 100-10,000 times [11], [12]. The advantages of GEP are [13]: 1) the chromosomes are simple entities: linear, compact, relatively small, easy to manipulate genetically (replicate, mutate, recombine, etc); 2) the expression trees are exclusively the expression of their respective chromosomes; they are entities upon which selection acts, and according to fitness, they are selected to reproduce with modification.

The procedure for modeling the target variable (stem water potential) is as follows: The first step is the selection of fitness function. For this problem, the fitness function, f_i, of an individual program, i, is expressed as [12]:

$$f_i = \sum_{j=1}^{n} \left(M - \left| C_{i,j} - T_j \right| \right);$$ in which M is the range of selection,

$C_{i,j}$ is the value predicted by individual program i for fitness case j, and T_j is the target value for fitness case j. For a perfect fit, $C_{i,j}=T_j$. The second step consists in choosing the set of terminals T and the set of functions F to create the chromosomes. In the current problem, the terminal set includes the considered inputs (i.e. standard climatic data and soil moisture measurements at different depths). Also, different mathematical functions were utilized ({+, -, *, /},{ $\sqrt{\ }$, $\sqrt[3]{\ }$, ln(x), e^x, x^2, x^3, sin(x), cos(x), arctg(x)}). The third step is to choose the chromosomal architecture. Length of head, h=8, and three genes per chromosomes were employed. The fourth step is to choose the linking function. The linking function must be chosen as "addition" or "multiplication" for algebraic sub trees [12]. Here, the sub trees were linked by addition. The fifth and final step is to choose the genetic operators. The parameters used per run are summarized as follows:

Number of chromosomes: 30, head size: 8, number of genes: 3, linking function: addition, fitness function error type: root relative squared error, mutation rate: 0.044, inversion rate: 0.1, one point recombination rate: 0.3, two point recombination rate: 0.3, gene recombination rate: 0.1, gene transposition rate: 0.1, insertion sequence transposition rate: 0.1, root insertion sequence transposition: 0.1.

Some examples of GEP operations are given as follows. The tree structure of the program $\left[-q + (\Pi^{0.5}/3p) \right]$ is shown in Fig. 1. A population of random trees representing the program is initially constructed and genetic operations are performed on these trees to generate individuals with the help of two distinct sets; the terminal set T and the function set F. For Fig. 1, $\{/,+,\sqrt{\ },/\} \subseteq F$ and $\{\Pi,3,p,q\} \subseteq T$. In order to generate a random tree one has to pick randomly from $T \cup F$, until all branches end up in terminals.

The crossover involves that two random nodes are selected from inside such program (parents) and thereafter the resultant sub-trees are swapped, generating two new programs

as in Fig. 2. In a mutation a sub tree is replaced by another one randomly (Fig. 3). Reproduction means an exact duplication of the program if it is found to be acceptable by the fitness criteria.
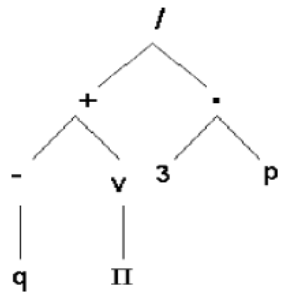

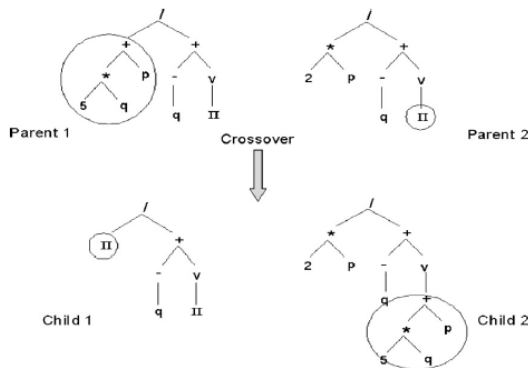Fig. 1. Tree structure of the program.

Fig. 2. Crossover.

Fig. 3. Mutation.

IV. EXPERIMENTAL PLOT PROVIDING THE DATA SET

The dataset used in this educational proposal is obtained from an experimental plot located in the nearby Senyera, Valencia, where the lecturers organized an experimental design for their research aims. In particular, the plot is planted with 'Navelina' orange trees (*Citrus sinensis* L. Osbeck) grafted on 'Cleopatra' mandarin (*Citrus reshni* Hort.), and FDR probes were used for monitoring soil moisture. A graphical representation of the experimental devices involved can be found in Martí *et al*. [7]. The FDR probe is placed 25 cm from the emitter's line inside a PVC access tube installed within the wetted area of the trees to record different dynamic variations and a wide range of soil water content within the zone covered by the FDR sensors. The probe, installed in the north side of the tree, has sensors at 10, 30, 50 and 70 cm depth (hereafter, h_1, h_2, h_3 and h_4, respectively). Data are stored in a data logger and are transmitted via GPRS to a central server. The store raw data are graphically displayed as volumetric water content. Midday stem water potential is measured around 12 h using a pressure chamber following the procedures described by Turner [14]. The leaves are wrapped in bags at least 2 hours previously to the measurement. A nearby automatic weather station of the Valencian Irrigation Technology Service (IVIA) provides additional meteorological inputs for the GEP models. The data set will be provided to the students in the practical sessions in the computer lab. This data set will be already previously preprocessed by the lecturers, ensuring data quality.

V. SCHEDULE AND CONTENTS OF EDUCATIONAL ACTIVITIES

The proposed educational activity is scheduled in at least 3 sessions: one practical session in the experimental plot and one at the computer lab. Between both, a third session introduces the GEP fundamentals to the students, and the basic commands of the software used.

A. Session 1: Preliminary Practical Session

The preliminary practical session is aimed at ensuring a suitable visualization of the experimental data acquisition. This session should allow a proper understanding of the modeling problem addressed and of the variables involved. This session would be focused on how to carry out the irrigation scheduling of the plot based on FDR soil moisture measurements and climatic inputs. Further, the measurement of stem water potential will be also explained and performed. The students will also visualize on-site how irrigation water is scheduled, i.e. how the calculated irrigation doses are automated and applied through the irrigation system. In this session, different soil water content profiles acquired with the FDR probes will be analyzed and interpreted (Fig. 4). Although the students will carry out some measurements, the data set for the application of the GEP models is based on lecturers' measurements corresponding to the last years.

B. Session 2: Preliminary Theoretical Session

In this second session the lecturer will introduce the fundamentals of GEP as simply as possible. The students should be able to understand the key methodological concepts and, at least, a qualitative approximate understanding of how GEP works and should be applied. This session also includes a brief introduction to the software used for the GEP application, Gene Xpro. Moreover, a brief introduction to modeling issues in irrigation scheduling is also planned, with a brief literature review of the most important existing approaches. Optionally, some scientific papers will be available for interested students. Some mention to important scientific journals will be made for those students who might want to look for further papers on their own [15]. A short guide with all the tackled contents will be provided, including theoretical foundations and important references. Further, the specific important steps that should be followed in the computer lab will be also summarized, including the definition of the statistical indicators that should be used to assess the model performance.

C. Session 3: Modeling Stem Water Potential with GEP

In this last session, the students will get a matrix with experimental measures of stem water potential, soil moisture measurements at different depths, and meteorological variables. Supervised by the lecturer, they will train GEP models based on different parametrizations using a toolbox for GEP application. So, there is no need for implementing complex algorithms. The students will be asked to find the

GEP models for 2 different input combinations, namely considering all inputs (GEP_1), and a selection of them (GEP_2).
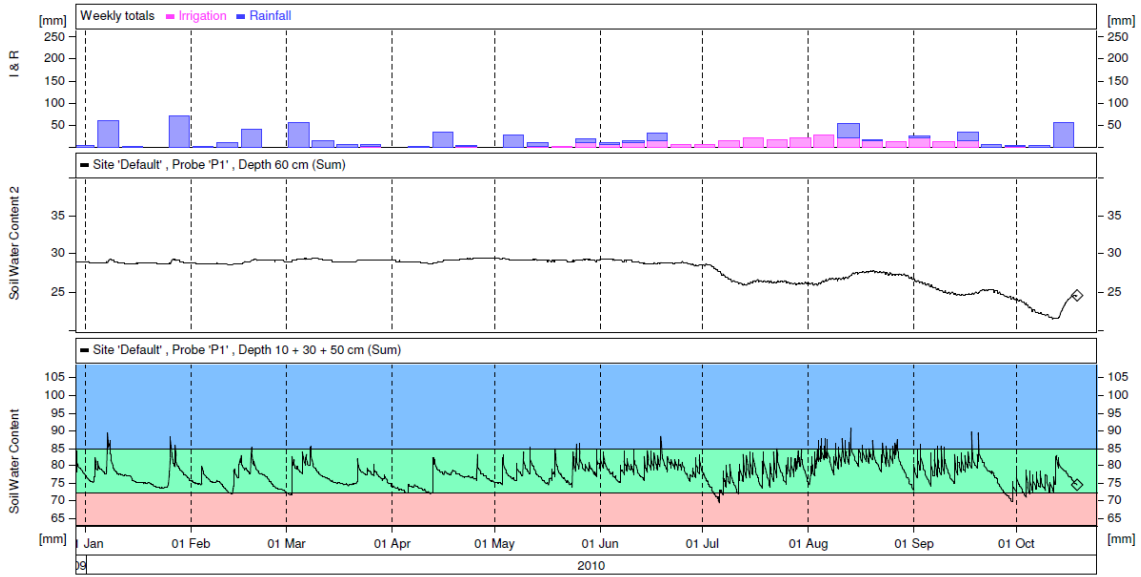

Fig. 4. Example of soil water content profile obtained with FDR probes.

In Spain, most of the subjects scheduled in the syllabus corresponding to the degree programs of engineering careers require an important basis of mathematical fundamentals. Consequently, all engineering degrees include a general training on Mathematics during the first years, incorporating practical sessions in a computer lab with mathematical software like Matlab [16]. This software uses a high-level technical computing language and includes an interactive environment for algorithm development, data visualization, data analysis and numerical computing. It allows a huge variety of computational possibilities and has become a very valued tool for all kind of engineers. In order to extend this general training and promote the use of this software within students with specific agricultural applications, the GEP estimations will be exported to Matlab, which will be applied to assess the model performance. Therefore, they will use this software for calculating different statistical indexes to assess the model performance, such as the mean squared error (MSE), the absolute average relative error (AARE), and the coefficient of determination (R^2). Moreover, a dispersion plot of the estimates will be provided, too.

The students' performance during this activity will be evaluated by the lecturer through a report, where the students should present a description of the contents studied and an assessment of the models developed, including the final GEP expressions and the performance indicators, and comment on them. The reports would be presented in groups of students, so that they might discuss the results and organize the work. This promotes sharing, discussing and clarifying ideas.

VI. RESULTS

The results should allow for a comparison between the two input combinations considered through their corresponding MSE, AARE and R^2 (Table I), whereas the dispersion plot is presented in Fig. 5.

The students should bear in mind some issues when analyzing the results. GEP_2 considers just h_3, instead of all depths, and the performance is even more accurate than GEP_1. So, a correlation between the different h_i measurements should be taken into account. A further analysis might be performed using principal component analysis. A deeper assessment should find the optimum depth for modeling stem water potential. A further issue that should be remembered is that these models were trained and tested with the same data set. Accordingly, these indicators cannot be used to evaluate their generalizability. A further additional step would be to apply independent validation procedures to assess the prediction ability of the developed model. This might include the application of cross-validation strategies, such as hold out validation, k-fold validation or leave-one-out validation.

TABLE I: PERFORMANCE INDICATORS OF GEP MODELS

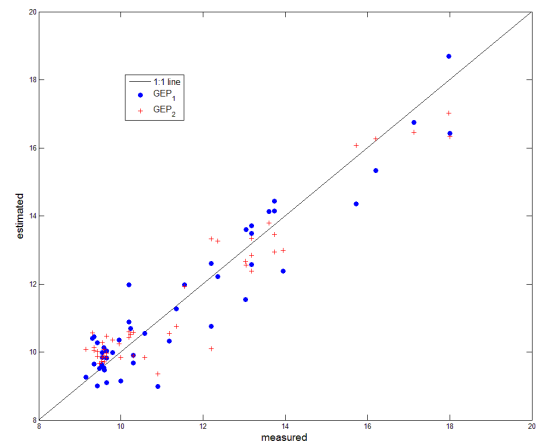| Model	MSE (MPa^2)	R^2 (-)	AARE (-)
GEP_1	0.629	0.894	0.054
GEP_2	0.529	0.918	0.052

Fig. 5. Dispersion plot of GEP estimations.

One of the advantages of using GEP over other data driven techniques is that it can produce explicit formulations of the relationship that rules the physical phenomenon. In this way, one may perform knowledge discovery using GEP, find some

confirmation of well-known physical relationships and evolve interesting new formulae, as well as upgrade particular cases of study. Accordingly, the encountered GEP expressions should be provided by the students in their report. In this case, GEP_1 would be expressed as:

$$\psi = Sin\left[\frac{\sin\left(h_3 HR^{-1}\right)}{\cos(h_3)+1.137268}\right] + \left[\left(T - \sqrt{2R_S} + V^2\right)h_4^{-1}\right] +$$
$$+ \left[\left(T - h_2 - h_1 + h_4 + 75.3391\right)h_4^{-1}\right]$$

While GEP_2 would be expressed as:

$$\psi = \sqrt{T + 9.5499V + 378.61} +$$
$$+ arctg\left[Exp\left(arctg(HR.V^{-1})\right) - \left(h_3 - T\right)\sqrt[3]{R_S}\right] +$$
$$+ arctg\left[T\left(v - h_3 - \cos(HR) + 14.8607\right)\right] - 7.2803$$

Nevertheless, it should be noted that these models rely on a limited data set and should not be used for prediction purposes in practice.

VII. CONCLUSIONS

A better understanding of the problems addressed in different subjects of scientific-technological careers can be achieved through the introduction of modeling tools and methodologies as complementary educational resources. These modeling tools can be obviously also very useful by themselves for the day-to-day practice of the future professionals. These contents are often not included in the programs of many subjects corresponding to scientific-technological degrees. This work presents an alternative way to teach the fundamentals of irrigation scheduling in the subject 'Irrigation and Drainage', corresponding to the Degree of Agricultural Engineering at the Universitat Politècnica de València. In particular, the traditional theoretical contents are complemented with a robust modeling approach, and, accordingly, a Gene Expression Programming based approach is applied for estimating stem water potential, a key variable in this field, in the computer lab.

Besides from learning a general crucial topic in irrigation engineering and the fundamentals of a very powerful modeling tool, the students can become aware of the great computational possibilities offered by specific numerical computing software when facing real specific problems that might be encountered in their professional practice. Further, they realize the applicability of the theoretical mathematical training acquired in other subjects through specific applications for addressing complex problems in agricultural engineering, fostering motivation within students, and supporting their concept formation, sense-making and experience.

This activity also shows the research activities in this field and, in particular, of the lecturers, promoting the interest for research as a possible professional activity. Finally, independent learning through scientific journals is also promoted.

REFERENCES

[1] G. S. Campbell and M. D. Campbell, "Irrigation scheduling using soil moisture measurements: Theory and practice," *Advances in Irrigation*, vol. 1, Academic Press, New York, 1982, pp. 25–42.

[2] H. G. Jones, "Irrigation scheduling: advantages and pitfalls of plant-based methods," *Journal of Experimental Botany,* vol. 55, no. 407, pp. 2427–2436, 2004.

[3] J. Peretz *et al.*, "Leaf water potential for management of high frequency irrigation on apples," *Transactions of the American Society of Agricultural Engineers,* vol. 27, pp. 437–442, 1984.

[4] K. Steppe *et al.*, "A step towards new irrigation scheduling strategies using plant-based measurements and mathematical modeling," *Irrigation Science,* vol. 26, pp. 505–517, 2008.

[5] A. O. Santos and O. Kaye, "Grapevine leaf water potential based upon near infrared spectroscopy," *Scientia Agricola*, vol. 66, no. 3, pp. 287–292, 2009.

[6] C. Acevedo-Opazo, B Tisseyre, J. A. Taylor, H. Ojeda, and S. Guillaume, "A model for the spatial prediction of water status in vines (*Vitis vinifera* L.) using high resolution ancillary information," *Precision Agriculture,* vol. 11, pp. 358–378, 2010.

[7] P. Martí, M. Gasque, and P. González-Altozano, "An artificial neural network approach to the estimation of stem water potential from frequency domain reflectometry soil moisture measurements and meteorological data," *Computers and Electronics in Agriculture,* vol. 91, pp.75-86, 2013.

[8] J. R. Koza, *Genetic Programming: On The Programming of Computers by Means of Natural Selection*, Cambridge, MA, USA: The MIT Press, p. 840, 1992.

[9] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison–Wesley, Reading, MA, USA, p. 432, 1989.

[10] W. Banzhaf *et al.*, *Genetic Programming*, Morgan Kaufmann, San Francisco, CA, USA, p. 512, 1998.

[11] C. Ferreira, "Gene expression programming in problem solving," presented at 6th Online World Conference on Soft Computing in Industrial Applications, 2001.

[12] C. Ferreira, "Gene expression programming: a new adaptive algorithm for solving problems," *Complex Systems,* vol. 13, no. 2, pp. 87-129, 2001.

[13] C. Ferreira, *Gene Expression Programming: Mathematical Modeling by an Artificial Intelligence*, Berlin, Heidelberg New York, NY, USA: Springer, p. 478, 2006.

[14] N. C. Turner, "Techniques and experimental approaches for the measurement of plant water status," *Plant Soil,* vol. 58, pp. 339–366, 1981.

[15] P. Martí, M. Gasque, and P. González-Altozano, "How do students of agricultural engineering perceive research?" in *Proc. the 9th International Technology, Education and Development Conference*, 2015, pp. 1619-1624.

[16] Matlab, *User's Manual Version 7.4.0 (R2007a)*, The MathWork Inc, Natick, Mass, 2007.

Pau Martí received his B.S., M.A. and Ph.D. degrees in agricultural engineering from the Universitat Politècnica de València (Spain). He is currently assistant professor at the Department of Biology of the Universitat de les Illes Balears. His main research interests focus on irrigation engineering, hydrology and agrometeorology.

Jalal Shiri received his B.S., M.A. and Ph.D. degrees in agricultural engineering from the University of Tabriz (Iran). He is currently assistant professor at the Water Department Engineering of the University of Tabriz. His main research interests focus on irrigation engineering, hydrology and agrometeorology.

Maria Gasque received her B.S., M.A. and Ph.D. degrees in agricultural engineering from the Universitat Politècnica de València (Spain). She is currently associate professor at the Dept. of Applied Physics of the Universitat Politècnica de València. Her main research interests focus on irrigation engineering and water management, and solar thermal energy storage.

Pablo González-Altozano received his B.S., M.A. and Ph.D. degrees in agricultural engineering from the Universitat Politècnica de València (Spain). He is currently associate professor at the Department of Rural and Agrifood Engineering of the Universitat Politècnica de València. His main research interests focus on irrigation engineering and water management, and solar thermal energy storage.