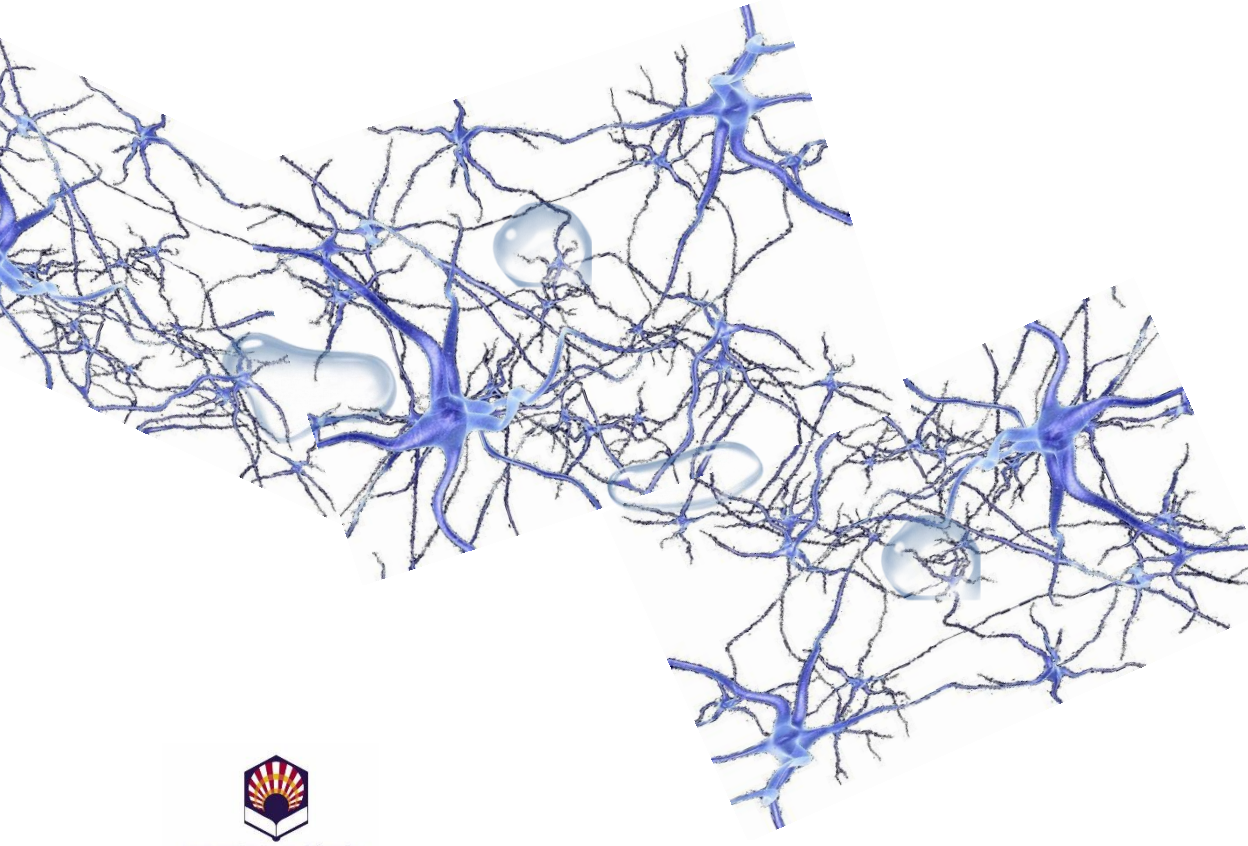
Optimum management of pressurized irrigation networks at different scales using Artificial Intelligent techniques

"Optimización de la gestión de redes de riego a presión a diferentes escalas mediante Inteligencia Artificial"



UNIVERSIDAD D CORDOBA

Departamento de Agronomía
Área de Ingeniería Hidráulica

Rafael González Perea

TITULO: *Optimización de la gestión de redes de riego a presión a diferentes escalas mediante Inteligencia Artificial.*

AUTOR: *Rafael González Perea*

Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

www.uco.es/publicaciones
publicaciones@uco.es

Departamento de Agronomía
Área de Ingeniería Hidráulica

Optimum management of pressurized irrigation networks at different scales using Artificial Intelligent techniques

"Optimización de la gestión de redes de riego a presión a diferentes escalas mediante Inteligencia Artificial"

Tesis doctoral presentada por

Rafael González Perea

para la obtención del título de

DOCTOR CON MENCIÓN INTERNACIONAL

POR LA UNIVERSIDAD DE CÓRDOBA

Directores

Dr. Emilio Camacho Poyato (Catedrático de Universidad)

Dr. Juan Antonio Rodríguez Díaz (Prof. Contratado doctor)

Mención de doctorado internacional

Esta tesis cumple con los requisitos establecidos por la Universidad de Córdoba para la obtención de la mención de doctorado internacional:

- Estancia de 3 meses realizada en Cranfield Water Science Institute, de la Universidad de Cranfield (Reino Unido), bajo la supervisión del Dr. Jerry W. Knox.

- Informe previo de dos doctores externos y con experiencia investigadora acreditada de alguna institución de educación superior o instituto de investigación de fuera de España.

- Un miembro del tribunal pertenece a un centro de investigación extranjero.

- Parte de la tesis está escrita en inglés y castellano.

Tesis como compendio de publicaciones

Esta tesis se presenta como compendio de publicaciones, cumpliendo con los requisitos establecidos por la Universidad de Córdoba para este fin. Cuatro capítulos de la presente tesis se corresponden con cuatro artículos científicos publicados en revistas incluidas en el primer cuartil según la última relación del Journal Citation Reports (2016) y dos de ellos se encuentran actualmente en fase de revisión:

1. González Perea R, Camacho Poyato E, Montesinos P, Rodríguez Díaz JA (2014) *Critical points: interactions between on-farm irrigation systems and water distribution network*. Irrigation Science 32: 255–265. Índice de impacto: 2.056. 1er cuartil en el área de agronomía, posición 18/ 81

2. González Perea R, Daccache A, Rodríguez Díaz JA, Camacho Poyato E and Knox JW (2017) *Modelling impacts of precision irrigation on crop yield and in-field water management*. Precision Agriculture. doi:10.1007/s11119-017-9535-4. Índice de impacto: 2.012. 1er cuartil en el área de agricultura, posición 9/ 56

3. González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA (2016) *Optimization of Irrigation Scheduling Using Soil Water Balance and Genetic Algorithms*. Water Resources Management 30:2815-2830. Índice de impacto: 2.848. 1er cuartil en el área de ingeniería civil, posición 16/ 125

4. González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA (2015) *Irrigation Demand Forecasting Using Artificial Neuro-Genetic Networks*. Water Resources Management 29: 5551-5567. Índice de impacto: 2.437. 1er cuartil en el área de ingeniería civil, posición 13/ 126

5. González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA (2017) *Prediction of Irrigation Event Occurrence at Farm Level using Optimal Decision Trees*. Water Resources Management (submitted). Índice de impacto: 2.848. 1er cuartil en el área de ingeniería civil, posición 16/ 126

6. González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA (2017) *Farmer's Behaviour Modelling by the Prediction of the Applied Irrigation Depth using Artificial Intelligence*. Water Resources Research (submitted). Índice de impacto: 4.397. 1er cuartil en el área de ciencias medioambientales, posición 28/ 229

TÍTULO DE LA TESIS: *Optimización de la gestión de redes de riego a presión a diferentes escalas mediante Inteligencia Artificial*

DOCTORANDO: Rafael González Perea

INFORME RAZONADO DEL/DE LOS DIRECTOR/ES DE LA TESIS

En los últimos años, el sector del regadío ha experimentado un proceso de modernización en el que las redes a presión han sustituido a las tradicionales redes de canales abiertos. Esto ha permitido aumentar considerablemente la eficiencia en el uso del agua, pero por el contrario, implica una mayor dificultad en la gestión, dado que entran en juego nuevos aspectos como la alta demanda energética de los sistemas a presión. Así, los gestores se ven en la necesidad de asegurar un suministro adecuado de agua, en términos de caudal y presión, pero minimizando el gasto en energía.

El desarrollo de nuevos sensores y sistemas de adquisición de dato y la mejora de las comunicaciones, cada vez a precios más reducidos, está cambiando la agricultura actual, la cual tiende a ser una actividad de precisión en la que se dispone de gran cantidad de información para la ayuda a la toma de decisiones

óptima. No obstante, es necesario desarrollar herramientas que, usando toda la información generada, ayuden a los gestores y técnicos de las zonas regables para usar los distintos recursos de la manera más eficiente posible. Así, técnicas que hasta ahora no eran comunes (inteligencia artificial, big data, lógica difusa, etc.) adquirirán cada vez más, un papel más importante en la gestión del regadío.

La presente Tesis Doctoral se organiza en dos grandes secciones que abordan los temas anteriormente citados.

En una primera sección se analiza el uso eficiente del agua y de la energía a escalas de red de distribución y de parcela. De esta forma, en el primer trabajo se analiza la función de los puntos críticos (hidrantes con elevados requerimientos de energía) en la gestión de la red de distribución de agua, estableciéndose estrategias para minimizar el impacto de los mismos en el consumo energético, pero sin afectar de una manera significativa a la productividad de los cultivos. Este aspecto se complementa con un segundo trabajo, más centrado a escala de parcela, pero que integra el modelo de crecimiento de cultivos Aquacrop y Sistemas de Información Geográfica para conocer con mayor precisión la variabilidad espacial en la productividad dentro de la finca y vincularla a la uniformidad del sistema de riego.

En el último trabajo de esta sección, se analiza la gestión de la zona regable como un todo, integrando la programación del

riego en cada una de las parcelas con la gestión de la red de riego, estableciendo estrategias de sectorización óptimas (turnos de riego con similares requerimientos de energía). Este trabajo representa un cambio en el concepto de gestión tradicional de las zonas regables, en el que la gestión de la red y la programación del riego en parcela no se hacen de forma conjunta.

La segunda sección de la Tesis se centra en la predicción de la demanda de riego a diferentes escalas. La predicción de la demanda es un aspecto cada vez más importante dado que, entre otras cosas, es esencial para una correcta contratación de la potencia y la energía.

En el primero de los trabajos de esta sección se usan algoritmos neuro-genéticos para predecir la demanda de agua diaria en la zona regable de la Zona Regable de la Margen Derecha del Río Bembézar, consiguiendo un error estándar de únicamente el 12.6 %.

Esta sección se complementa con otros dos trabajos en los que se usan técnicas de redes neuronales artificiales, lógica difusa y algoritmos genéticos para predecir los eventos de riego en cada parcela (cuando riega un agricultor en particular) y la demanda horaria de riego en toda la zona regable.

Por todo esto, consideramos que se trata de una Tesis de gran calidad y que aborda un problema real, de gran actualidad y

con gran aplicabilidad al sector. La Tesis se presenta como un compendio de seis artículos científicos, de los cuáles cuatro ya han sido publicados en algunas de las revistas más prestigiosas de su área de conocimiento, todas en el primer cuartil, y dos trabajos se encuentran actualmente en proceso de revisión:

1. González Perea R, Camacho Poyato E, Montesinos P, Rodríguez Díaz JA (2014) Critical points: interactions between on-farm irrigation systems and water distribution network. Irrigation Science 32: 255–265. Índice de impacto: 2.056. 1er cuartil en el área de agronomía, posición 18/ 81

2. González Perea R, Daccache A, Rodríguez Díaz JA, Camacho Poyato E and Knox JW (2017) Modelling impacts of precision irrigation on crop yield and in-field water management. Precision Agriculture 28 (4): 1095-1109. Índice de impacto: 1.675. 1er cuartil en el área de agricultura, posición 9/ 57

3. González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA (2016) Optimization of Irrigation Scheduling Using Soil Water Balance and Genetic Algorithms. Water Resources management 30:2815-2830. Índice de impacto: 2.848. 1er cuartil en el área de ingeniería civil, posición 16/ 125

4. González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA (2015) Irrigation Demand Forecasting Using Artificial Neuro-Genetic Networks. Water Resources

management 29: 5551-5567. Índice de impacto: 2.437. 1er cuartil en el área de ingeniería civil, posición 13/ 126

5. González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA (2017) Prediction of Irrigation Event Occurrence at Farm Level using Optimal Decision Trees. Water Resources Management (submitted). Índice de impacto: 2.848. 1er cuartil en el área de ingeniería civil, posición 16/ 126

6. González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA (2017) Farmer's Behaviour Modelling by the Prediction of the Applied Irrigation Depth using Artificial Intelligence. Water Resources Research (submitted). Índice de impacto: 4.397. 1er cuartil en el área de ciencias medioambientales, posición 28/ 229

Por todo ello, se autoriza la presentación de la tesis doctoral "Optimización de la gestión de redes de riego a presión a diferentes escalas mediante Inteligencia Artificial".

Córdoba, 25 de Septiembre de 2017

Firma de los directores

Fdo: Prof. Dr. Emilio Camacho Poyato

Fdo: Prof. Dr. Juan Antonio Rodríguez Díaz

Agradecimientos

"...tiempo al Tiempo le pido y tiempo el Tiempo me da ..."

Gracias a todas aquellas personas que han hecho posible la realización de esta tesis doctoral y me han regalado lo más preciado que se puede poseer, su *tiempo*:

A mis directores de tesis, Dr. Emilio Camacho Poyato y Dr. Juan Antonio Rodríguez Díaz. Gracias por toda la paciencia, dedicación, esfuerzo, confianza y guía que me habéis aportado en todo este viaje que comenzó hace ya más de 7 años como alumno de 4º de Ingeniería y espero que no finalice con esta tesis. Gracias de corazón por depositar en mí algo que jamás volveréis a recuperar, vuestro *tiempo*.

A la Dra. Pilar Montesinos Barrios, que para mí es la tercera directora de esta tesis. Gracias Pilar, por todas esas correcciones, ese pragmatismo que se refleja en todos y cada uno de los trabajos de esta tesis, por "jugar" y hacer que cada día sea un poco más divertido, pero sobre todo gracias por tu *tiempo*.

Agradecer al Ministerio de Educación, Cultura y Deporte por la concesión del contrato predoctoral de Formación del Profesorado Universitario (FPU), el cual ha hecho posible la realización de la presente tesis doctoral.

Al Dr. Jerry Knox, del Cranfield Water Science Institute (Universidad de Cranfield), por permitirme realizar mi estancia predoctoral junto a él y dedicarme parte de su valioso *tiempo*. Gracias a todos los compañeros de Cranfield, especialmente a Rocío y Carlos que gracias a ellos mi estancia allí fue inolvidable.

A mis compañeros del departamento de Ingeniería Hidráulica, especialmente a Irene Fernández y Teresa Carrillo, por estar siempre ahí, tanto en lo profesional como personal, y a Rafa Pérez compañero y amigo no sólo del departamento sino dentro y fuera de la pista. Gracias chicos, por todo este *tiempo* increíble y por ser en gran parte responsables de que yo pueda escribir estas palabras en este documento. A Manuel Martín, menor de cuatro hermanos, …, bueno ya sabes el resto. Gracias por hacer el día a día más divertido y enseñarme como se puede controlar algún '*as much as better*' que hemos encontrado por ahí. A Aída Mérida, gracias por toda tú dedicación y disponibilidad en este último año y como no por enseñarme nuevo vocabulario. A Carmen Alcaide, gracias por todos los buenos momentos que hemos pasado tanto dentro como fuera del trabajo. A Jorge García, gracias por el buen ambiente de trabajo.

A Puri y Fran por hacerme sentir en todo este *tiempo* como uno más de vuestra familia. Gracias chicos por incluirme cada semana en vuestra 'cofradía', con nuestro amigo el 'moro'. Espero y deseo que sea así durante un largo *tiempo*.

Un agradecimiento se me queda muy corto para ti. Gracias por demostrarme que el *tiempo* siempre hace su función, por recordarme que 'querer es poder', por hacerme creer que el "*el fate es el fate*", que los búhos existen…En definitiva, gracias por darme lo que la mayoría de las personas dedican la mayor parte de su *tiempo* en buscar, una ilusión.

Mis últimos y más especiales agradecimientos, como no podría ser de otra manera, van para mi familia. A mi hermana, que con ella el concepto *tiempo* se pierde, ya estábamos juntos desde antes de nacer, entendiéndonos cuando nadie más lo hacía. Gracias por apoyarme siempre, sigue así que lo conseguirás. Gracias A mis padres Rafa y María, que para vosotros hablar de *tiempo* carece de sentido. Gracias por toda la paciencia que habéis tenido, por toda la educación que me habéis dado, por enseñarme que sin esfuerzo no se consigue nada. Por demostrarme que siempre vais a estar ahí. Por inculcarme que haga lo que haga, cada noche debo acostarme con la conciencia tranquila. Simplemente gracias, porque soy lo que soy debido a vosotros.

Summary

Factors such as climate change, world population growth or the competition for the water resources make freshwater availability become an increasingly large and complex global challenge. Under this scenario of reduced water availability, increasing droughts frequency and uncertainties associated with a changing climate, the irrigated agriculture sector, particularly in the Mediterranean region, will need to be even more efficient in the use of the water resources. In Spain, many irrigation districts have been modernized in recent years, replacing the obsolete open channels by pressurized water distribution networks towards improvements in water use efficiency. Thanks to this, water use has reduced but the energy demand and the water costs have dramatically increased. Thus, strategies to reduce simultaneously water and energy uses in irrigation districts are required.

This thesis consists of nine chapters, which include several models to optimize the management of the irrigation districts and increase the efficiency of water and energy use. Chapter 1 provides the framework in which the present PhD thesis has been developed and the objectives are presented in Chapter 2.

The following chapters are organized in two thematic sections, with three chapters each. The first section includes Chapters 3, 4 and 5 and presents methodologies to increase the energy use efficiency considering both, water distribution networks and on-farm irrigation systems.

Chapter 3 shows a new approach of critical points control considering the interaction between the water distribution network and the irrigation system's performance in the critical field, which is the one supplied by hydrants with the highest energy requirements. Thus, the impacts of changes in the manometric regulation of the pump station has in the irrigation system of the critical field has been analyzed. The methodology has been applied to a real irrigation district, achieving average energy cost savings of 15 % with no significant yield losses.

The simplicity of the crop yield estimation in Chapter 3 leaded up to Chapter 4. Here a new methodology to assess the impacts of irrigation heterogeneity on crop yield using geospatial analysis and crop modelling techniques was developed. This chapter was done in collaboration with researchers from Cranfield University (UK) and the methodology developed was applied to a case study in Eastern England.

Chapter 5 describes a new methodology for water distribution networks sectoring but considering the soil water balance and

irrigation scheduling in each farm, using the multiobjetive genetic algorithm NSGA-II. The developed model determines the optimal threshold values of relative soil moisture in each field to identify the hydrants that make up the operating irrigation sector including farmer's profit and energetic criteria. The methodology has been applied to a real irrigation district achieving average energy cost savings of 27 %.

Section II is composed of Chapters 6 to 8 in which several models to predict the irrigation demand at different spatial and time scales have been developed. In chapter 6, a short-term forecasting model of daily irrigation water demand at irrigation district level using Artificial Neural Networks and the multiobjective algorithm NSGA-II has been developed. This predictive model was applied to a real irrigation district located in southern Spain. The results show that the model explains 93 % of the variability of the observed water demand with a standard error of 12.63 %.

Chapter 7 is aimed at modelling the farmer´s behavior and prediction of irrigation events. Thus, a new model combining Decision Trees and Genetic Algorithms is developed. The model was applied in a real irrigation district, classifying properly between 99.16 % and 100 % of the irrigation events.

Chapter 8 presents a model attempting to modelling the farmer's behavior and forecast the daily irrigation depth used by each

farmer using Artificial Neural Networks, Fuzzy Logic and the multiobjective algorithm NSGA-II. This model was tested in a real irrigation district achieving to explain 87 % of the variability observed with a standard error of 9.80 %.

The general conclusions drawn from this thesis and the avenues for future research are included in Chapter 9.

This thesis highlights the need for improving the management of the irrigation districts and presents several innovative strategies to optimize simultaneously water and energy use, while increasing farmer's profits at the same time.

Resumen

Factores tales como el cambio climático, el crecimiento de la población mundial o la competencia por los recursos hídricos hacen que la disponibilidad de agua se esté convirtiendo en un desafío global cada vez más grande y complejo. En este escenario de reducción de la disponibilidad de agua, aumento de la frecuencia de las sequías y de las incertidumbres asociadas a un cambio climático, el sector de la agricultura de regadío, en particular en la región mediterránea, tendrá que ser aún más eficiente en el uso de los recursos hídricos. En España, muchas comunidades de regantes se han modernizado en los últimos años, sustituyendo los obsoletos canales abiertos por redes de distribución de agua a presión con el objetivo de mejorar la eficiencia en el uso del agua. Gracias a esto, el uso del agua se ha reducido, pero la demanda de energía y los costos del agua se han incrementado drásticamente. Por lo tanto, se requieren estrategias para reducir simultáneamente el uso de agua y energía en las comunidades de regantes.

Esta tesis consta de nueve capítulos que incluyen varios modelos para optimizar la gestión de las comunidades de regantes y aumentar la eficiencia en el uso del agua y la energía. El capítulo

1 proporciona el marco en el que se ha desarrollado la presente tesis doctoral y los objetivos se presentan en el capítulo 2.

Los capítulos siguientes se han organizado en dos secciones, con tres capítulos cada una. La primera sección incluye los capítulos 3, 4 y 5 y presenta metodologías para aumentar la eficiencia en el uso de la energía considerando tanto la red de distribución de agua como los sistemas de riego en parcela.

El capítulo 3 muestra un nuevo enfoque de control de puntos críticos considerando la interacción entre la red de distribución de agua y el rendimiento del sistema de riego en la parcela crítica, aquella que se abastece por hidrantes con mayores necesidades energéticas. Así, se han analizado los impactos que los cambios en la regulación manométrica de la estación de bombeo tienen en el sistema de riego de la parcela crítica. La metodología se ha aplicado a una comunidad de regantes real, logrando un ahorro medio del coste energético del 15 % sin pérdidas significativas de producción.

La simplicidad de la estimación del rendimiento de los cultivos en el Capítulo 3 ha llevado a desarrollar un procedimiento más robusto en el Capítulo 4. La nueva metodología evalúa los impactos de la heterogeneidad del riego sobre el rendimiento de los cultivos usando el análisis geoespacial y las técnicas de modelización de cultivos. Este capítulo se realizó en colaboración

con investigadores de la Universidad de Cranfield (Reino Unido) y la metodología desarrollada se aplicó a un estudio de casos en el Este de Inglaterra.

El capítulo 5 describe una nueva metodología para la sectorización de redes de distribución de agua, pero considerando el balance hídrico del suelo y la programación del riego en cada parcela que compone la red de riego, utilizando el algoritmo genético multiobjetivo NSGA-II. El modelo desarrollado determina los umbrales óptimos de la humedad relativa del suelo en cada parcela de acuerdo a criterios energéticos y beneficio económico del agricultor. La metodología ha sido aplicada a una comunidad de regantes real logrando un ahorro medio del coste energético del 27 %.

La Sección II se compone de los Capítulos 6 a 8 en los que se han desarrollado varios modelos para predecir la demanda de riego a diferentes escalas espaciales y temporales. En el capítulo 6 se ha desarrollado un modelo de predicción a corto plazo de la demanda diaria de agua de riego a nivel comunidad de regantes utilizando Redes Neuronales Artificiales y el algoritmo multiobjetivo NSGA-II. Este modelo predictivo se aplicó a una comunidad de regantes del sur de España. Los resultados muestran que el modelo explica el 93 % de la variabilidad de la demanda de agua observada con un error estándar del 12,63 %.

El Capítulo 7 tiene como objetivo modelar el comportamiento del agricultor y predecir cuándo el agricultor decide regar. Así, se desarrolla un nuevo modelo combinando Árboles de Decisión y Algoritmos Genéticos. El modelo desarrollado se aplicó en una comunidad de regantes real, logrando clasificar correctamente entre el 99.16 % y el 100 % de los eventos de riego testados.

El Capítulo 8 presenta un modelo que predice la lámina de riego diaria aplicada por cada agricultor utilizando Redes Neuronales Artificiales, Lógica Difusa y el algoritmo multiobjetivo NSGA-II. Este modelo fue probado en una comunidad de regantes real logrando explicar el 87 % de la variabilidad observada con un error estándar del 9.80 %.

Las conclusiones generales extraídas de esta tesis y las líneas futuras de investigación se incluyen en el capítulo 9.

Esta tesis destaca la necesidad de mejorar la gestión integral de las comunidades de regantes y presenta varias estrategias innovadoras para optimizar simultáneamente el uso de agua y energía, al mismo tiempo que se incrementa el beneficio de los agricultores.

Table of Contents

List of Tables

List of Figures

List of symbols

ADI	Accumulated deep infiltration losses, mm
A_h	Irrigated area supplied by hydrant h, ha
A_T	Total irrigated area by all hydrants, ha
a, b	emitter spacing, m
bb	Parameter of the $\mu_{v,k}$
CatLevel	Number of categories or levels of a classification tree
C_d	Deficit coefficient
Chr	Chromosome
CU_c	Christiansen's uniformity coefficient, %
c_v	Fuzzy curve
CV_e	Coefficient of variation of the irrigation depths applied by all emitters
C_w	Water cost, € ha^{-1}
DaW	Daily amount of water applied by each farmer, mm
DAW_k	Daily amount of water applied by each farmer in the point k of the space PI_v – DaW
Demand_1	Water demand in the previous day, L s^{-1}
Demand_2	Water demand in the two previous days, L s^{-1}

DI_h	Deep infiltration losses corresponding to the hydrant h, mm
DM	Dry matter, t ha^{-1}
d	day
E(t)	Error measure that computes the performance of a node t
E_{ij}	Energy consumption per working day for each loading condition, kWh
ER_d	Effective rainfall of the day d, mm
$E_{T,h}$	Seasonal energy cost in the pumping station corresponding to the hydrant h, €
ET_0	Reference evapotranspiration, mm
$ET_{0,d}$	Reference crop evapotranspiration of the day d, mm day^{-1}
ET_0_1	Reference evapotranspiration in the previous day, mm
ET_h	Actual evapotranspiration for the crop of the hydrant h, mm day^{-1}
$ET_{max\ d,\ h}$	Evapotranspiration in no water stress conditions for the crop of the hydrant h in the days d of crop development, mm day^{-1}
$ET_{max,\ h}$	Evapotranspiration in no water stress conditions for the crop of the hydrant h, mm day^{-1}

ET_w	Evapotranspiration in so-called wilting point, mm day^{-1}
e	Total number of emitters
F1	Objective function 1 of the genetic algorithm
F2	Objective function 2 of the genetic algorithm
$F_{d,\,h}$	Demanded flow in the pumping station corresponding to the hydrant h on the day d,
$FIS_{chr\ I}$	FIS model belonging to chromosome *i* of the genetic algorithm
f	fraction of the field without water deficit
$fs_{v,j}$	Fuzzy Surface of the potential input variables *v* and *j*
g	Number of neuros of the input layer of the ANN
$\overline{H_e}$	Average of irrigation depths applied by all emitters, mm
$H_{d,h}$	Pressure head at the pumping station corresponding to irrigation sector of the hydrant h on the day d, mm
H_g	Total gross applied depth, mm
H_i	Applied irrigation depth for every emitter, mm
H_n	Net irrigation requirements, mm
HP	High pressure conditions, 40 m
H_r	Theoretical irrigation requirements, mm
Δh	Pressure decrease in each iteration i, m

h	Pressure head, m
h_c	Pressure at the critical hydrant, m
$h_{pi,\ max}$	Maximum pressure head, m
$h_{ps,\ i}$	Pressure head at the pumping station, m
h_t	Total number of hydrants of the irrigation network
I_1	Crop
I_2	Julian day.
I_3	Bank holiday (false or true)
I_4	Weekday
I_5	Daily maximum temperature, °C
I_6	Daily average temperature, °C
I_7	Daily average relative humidity, %
I_8	Precipitation event (false or true)
$I_{d,h}$	Applied irrigation depth to the crop of the hydrant h on the day d, mm
IE	Irrigation efficiency
IN	Irrigation needs, L ha^{-1} day^{-1}
i	Iteration index (it depends of each process)
J	Class of the classification tree
j	Number of random demand patterns
$K_{c\ d,h}$	Crop coefficient of the hydrant h taking into account days d of crop development

K_{sat}	Saturated hydraulic conductivity of the soil in the study area, mm day^{-1}
k_j	Sets of open hydrants
k_y	Yield response factor
$L_{d,h}$	Last day of the crop production period of the hydrant h
LP	Low pressure conditions, 10 m
M	Total number of points in the space PI_v – DaW m^3s^{-1}
MaxBranch	Maximum number of branch node splits per tree
MaxCatLevel	Maximum categories or levels of the classification tree
MinObsBranch	Minimum number of observations per branch node
MinObsLeaf	Minimum number of observations per leaf
$Missclass_{test,I}$	Farmer's decision i wrongly classified
MSE_{c_v}	Mean square error for the fuzzy curve c_v
$MSE_{fs_{v,j}}$	MSE of the fuzzy surface $fs_{v,j}$
m	Number of neurons of the second hidden layer of the ANN
mc	Momentum constant
mse	Mean sum of squares of the ANN errors
msereg	Mean sun of squares and mean of the sum squares of the ANN weights

msw	Mean sum of squares of the ANN weights
$Numclass_{test}$	Total number of farmer's decisions included in the test test
n	Number of neurons of the first hidden layer of the ANN
nDec	Number of decision variables of the algorithm genetic
nGEN	Number of generations of the algorithm genetic
nInput	Number of genes that makes up the chromosome in the optimization process of the ANFIS model
nm	Total number of the ANN weights
nPop	Number of chromosomes that makes up the initial population of the genetic algorithm
ns	Soil porosity
numNodes	Total number of nodes of the classification tree
OB	Total number of observations used for ANN training
OB_1	Total number of observations used for ANN training or validation sets
OB_{val}	Total number of observations used for ANN validation
P_c	Market price of the agricultural production, € kg^{-1}
PC	Perfect conditions of pressure, 25 m
P_{ij}	Power requirements at pumping station, kW

PI_v	Potential input variable v
$PI_{v,k}$	Value of the PI_v in the point k
Pr_h	Average market price of the crop of the hydrant h during irrigation season, € kg^{-1}
p_j	Probability that a case or data in a node belongs to class J
p_n	Open outlet probability
$\bar{\hat{Q}}$	Average of estimated water demand of training or validation sets, L s^{-1}
$\widehat{Q_t}$	Estimated water demand at the same time step t, L s^{-1}
$\overline{Q_v}$	Average of observed water demand of the validation set, L s^{-1}
$\bar{Q}$	Average of observed water demand of training or validation sets, L s^{-1}
Q_t	Observed water demand at the time step t, L s^{-1}
Q_{Tij}	Demanded flow in the pumping station, $m^3 s^{-1}$
q	Flow rate, L h^{-1}
q_{max}	Maximum flow allowed per hydrant, L s^{-1} ha^{-1}
q_n	Flow demanded for every open hydrant, L s^{-1}
R^2	Determination coefficient
R^2_{test}	Determination coefficient of the testing process of the ANFIS optimization
$R^2_{training}$	Determination coefficient of the training set

$R^2_{validation}$	Determination coefficient of the validation set
Rad	Solar radiation for the day to predict, MJ m^{-2}
Rad_1	Solar radiation in the previous day, MJ m^{-2}
$RSME_{test}$	RSME of the testing process of the ANFIS optimization
S_n	Irrigated area supplied by hydrant n, ha
SplitAlgorithm	Split Algorithm of the classification model
ss	Number of neuros of the output layer of the ANN
s*	Relative soil moisture from which the crops start to reduce transpiration
$s_{d, h}$	Relative soil moisture corresponding to the field of the hydrant h on the day d
$s_{d-1,h}$	Relative soil moisture of the hydrant h on the day d-1
s_{fc}	Relative soil moisture in field capacity
s_{hg}	Relative soil moisture in so-called hygroscopic point
$s_{irrigation}$	Relative soil moisture that determines the beginning of the irrigation
s_w	Relative soil moisture in the wilting point
T_{ave}	Average temperature for the day to predict, °C
t_a	Daily irrigation availability time
$t_{d,h}$	Irrigation time of the hydrant h on the day d, hours

t_n	Daily irrigation time in the peak month, h day^{-1}
$tree_{chr,i}$	Classification tree belonging to chromosome i
UC_E	Unit energy cost, € kWh^{-1}
w	Synaptic weight
X	Training set, %
x	Pressure head exponent
Y	Actual crop yield, kg ha^{-1}
Y_c	Crop yield, kg ha^{-1}
Y_h	Yield under actual conditions for the crop of the hydrant h, kg ha^{-1}
Y_{max}	Maximum crop yield without water stress, kg ha^{-1}
$Y_{max,\ h}$	Potential yield for the crop of the hydrant h when there are not limitations of water, kg ha^{-1}
$Y_{r,\ h}$	Relation between yield under actual conditions for the crop of the hydrant h and $Y_{max,\ h}$
$Z_{r\ d,h}$	Active soil depth (where most of crop roots associated to hydrant h on the day d are located), mm
α	Discharge coefficient, L h^{-1} m^{-x}
β	Coefficient which is used to fit the above expression to the power law
γ	Performance ratio
γ_w	Specific weight of water, 9.8 kN m^{-3}

$\mu_{v,k}$	Fuzzy membership function of the point k in the plot which relates the potential input variable v and the DaW
η	Pumping system efficiency
$\theta_{d\text{-}1,h}$	Volumetric soil moisture corresponding to the field of the hydrant h on the day d, $cm^3\ cm^{-3}$
θ_{hg}	Volumetric soil moisture at hygroscopic point, $cm^3\ cm^{-3}$
θ_w	Volumetric soil moisture at wilting point, $cm^3 cm^{-3}$
σ_e	Standard deviation of the irrigation depths applied by all emitters, mm
$\emptyset_g$	Gini diversity index

List of abbreviations

AI	Artificial Intelligence
ANFIS	Adaptive Neuro-Fuzzy Inference System
ANGN	Artificial Neuro-Genetic Network
ANN	Artificial Neural Network
BDD	Block Data Division
BFGS	Broyden–Fletcher–Goldfarb–Shanno algorithm
BGD	Batch Gradient Descent
BS	Brent's Search
BT	Backtracking
CNN	Computational Neural Network
CrossVal	Cross Validation
CS	Charalambous' Search
CZID	Canal del Zújar Irrigation District
DLL	Dynamic link library
dsigmf	Difference between two sigmoidal membership functions
DT	Decision Tree
ER	Evolutionary Robotic
FAO	Food and Agriculture Organizations of United Nations

FIS	Fuzzy Inference System
FL	Fuzzy Logic
FRU	Fletcher-Reeves Update
GA	Genetic Algorithm
gauss2mf	Asymmetric Gaussian membership function
gaussmf	Gaussian membership function
gbellmf	Generalized bell membership function
GDM	Gradient Descent with Momentum
GFSs	Genetic Fuzzy Systems
GIS	Geographic Information System
GS	Golden Section Search
GUI	Graphical User Interface
HBG	Hybrid-Bisection-Cubic Search
IDD	Index Data Division
IMZ	Irrigation Management Zone
INE	Instituto Nacional de Estadística
InterDD	Interleaved Data Division
LinST	Linear Transfer function
LM	Levenberg-Marquardt algorithm
LogST	Log-Sigmoid Transfer function
LSA	Line Search Algorithm
MAPA	Spanish Ministry of Agriculture, Fisheries and Food
MDB	M. D. Bembézar Irrigation District

MFs	Membership Functions
MLP	Multilayer Perceptron Network
MSE	Mean Square Error
NS	Network's Sectoring
NSGA-II	Non-dominated Sorting Genetic Algorithm
OSS	One Step Secant Algorithm
OVAC	One Versus All By Class algorithm
PBR	Powell-Beale Restarts
PCA	Principal Component-Based Partitioning
PF	Perform Function
PI	Precision Irrigation
pimf	Polynomial membership function
PLbyPurity	Pull Left By Purity algorithm
PRU	Polak-Ribiére Update
psigmf	Product of two sigmoidal membership functions
RB	Resilient Backpropagation
RDD	Random Data Division
RDP	Random Demand Patterns
RMSE	Root Mean Square Error
SCG	Scaled Conjugated Gradient
SEP	Standard Error Prediction
SWD	Soil Water Deficit, mm
TanST	Tan-Sigmoid Transfer function
TNF	Training Function

trapmf	Trapezoidal membership function
trimf	Triangular membership function
TS	Sugeno-Takagi
URI	Uniform Rate of Irrigation
USDA	United States Department of Agriculture
VLR	Variable Learning Rate
VLRM	Variable Learning Rate with Momentum
VRI	Variable Rate Irrigation
WAM	Water/irrigation Application Model
WCP	Water Content Pattern
WECP	Water and Energy optimization by Critical Points

1. Introduction

1.1. Background

Water is a key natural resource in the economic, social and political development of any region or country. Climate change, farming or industrial processes are some of the factors that make freshwater availability becomes an increasingly large and complex global challenge (Hunt 2004).

Furthermore, the Food and Agriculture Organization of the United Nations (FAO) expects that the world population will rise from the current 7 billions to 9 billions by 2050. Thus, an increase of the 70 % on food requirements is forecasted for the next thirty years (FAO 2011). Under this scenario and considering that there is not enough arable land to increase the food production in the estimated proportion, irrigated agriculture has a crucial role to play here.

In recent decades, the irrigated agriculture area has increased considerably covering a total area of 244M ha in 1989-1991 up to 287M ha in 2005-2007. Moreover, an additional increment of 32

M ha is predicted for 2050 to satisfy the growing food demand (Conforti 2011).

Consequently, irrigated agriculture is the most water demanding sector with around 70 % of the freshwater withdrawals at global level (Conforti 2011). Hence, improving the water use efficiency in the irrigation sector is essential to ensure global food supply in the coming years.

1.2. The Spanish irrigated agriculture

Water resources availability is a limiting factor for economic development in many water-stressed countries. One good example is the Mediterranean region where, due to its scarce and irregular rainfall, is one of the most water-scarce regions in the world (Daccache et al. 2014b). This is the case of Spain, where the expansion of irrigated production, coupled with tourism and urbanization has created significant water supply challenges (García-Ruiz et al. 2011).

Spain devotes 73 % of its national freshwater to irrigate 3.65M ha (INE 2016; MAGRAMA 2016) that represent one third of the irrigated land in the European Union (López-Gunn et al. 2012) . Since 2002, the Spanish government has developed a National Irrigation Plan and an Emergency Plan for Modernization of Irrigation with the aim of saving 3000 Mm^3 of water per year

(MARM, 2002 and 2006). These plans involved an investment of some 7400M €, affecting about 2M ha of the 3.5M ha of the current irrigated area (Lecina et al., 2010). The National Strategy for Sustainable Irrigation Modernization, Horizon 2015 continued with the improvement of the water management and promoted the sustainability of irrigation by pursuing energy efficiency (MARM, 2010).

As a consequence of this national aim to improve water use efficiency, pressurize networks have replaced the obsolete open-channel distribution systems (Plusquellec 2009) and many irrigated areas have migrated to more efficient irrigation systems such as drip and trickle (Playán and Mateos 2006). If in 1980, surface irrigation accounted for 80 % of the irrigated land, by 2009 it represented only 31 %. Drip irrigation changed over the same period from 2% to 46 % while the use of sprinkler irrigation has increased slightly (Rodríguez Díaz et al., 2011).

The upgrading of these water distribution systems allowed to reduce the water used by 21 % between 1950 and 2007 (Corominas 2010). However, these new water distribution systems require energy for their operation and consequently, the energy consumption per unit area had risen by 657 % in the same period (1950-2007) (Corominas 2010).

Several studies have been developed to assess the impact of the irrigation modernization process in Spain. Results showed a reduction of 23 % in the volume of water applied. However, the water costs were incremented by 52 % (Fernández García et al. 2014). Furthermore, after the liberalization of the electricity market in 2003 and the elimination of the electricity tariff for the irrigation sector in 2008, the electricity bill was increased by 120 % from 2008 to 2010 (Rodríguez-Díaz et al. 2011) and it has steadily risen up to now. This increase in the cost of the electricity tariff coupled with the intensification of the energy requirements have caused a significant increment of the production costs creating doubts concerning the profitability of irrigated agriculture in many cases.

Under this scenario of increasing energy demand, the European Union through the Directive 2006/32/CE on energy end-use efficiency and energy services tried to reduce the energy requirements by 9 %. With this aim, the Spanish government developed specific measures to improve the energy efficiency in the irrigation sector (IDAE 2005; IDAE 2007).

Afterwards, the European Union through Directive 2012/27/UE committed to achieve an energy saving of 20 % in 2020. Within this framework of energy savings in the horizon 2020, Spain launched its Action Plan 2011-2020 (IDAE 2011) which included

energy saving measures such as the promotion and the dissemination of irrigation techniques leading to higher energy efficiency, the migration to less energy demanding irrigation systems, such as drip or low pressure sprinkler systems, and energy audits. In 2010, the Spanish government proposed the first plan of the use efficiency both water and energy in irrigated agriculture (MARM 2010).

1.3. Reducing the energy dependence of the irrigation sector

In light of this increasing energy demand and the great rising of the electric tariff cost, many works have been focused on the development of techniques to improve the water use efficiency and reduce the energy consumption in pressurized networks. Network sectoring, where hydrants with similar energy requirements are grouped (Rodríguez Díaz et al. 2009; Carrillo Cobo et al. 2011; Fernández García et al. 2013), energy audits (Abadia et al. 2008); optimization of operation of the pumping station (Moreno et al. 2007) or the detection and control of critical points (Rodríguez Díaz et al. 2012; Fernández García et al. 2015), are some of the most effective techniques. Until now, all these techniques are applied at water distribution network level, without considering their impact on the on-farm irrigation systems.

Thus, the development of water and energy saving measures, that allow a comprehensive management of the irrigation networks including the on-farm irrigation systems, will provide a more accurate tool for irrigation districts managers.

Because of the modernization processes of the hydraulic infrastructures, most of the new irrigation systems has been designed to operate on-demand where water is continuously available for farmers who are free to decide how and when irrigate. Although the increased operation flexibility is a positive aspect for farmers, the management of the whole system is a more complex task for managers because the hydraulic systems have to work under a wide range of operating conditions, in terms of flow and pressure. Thus, new tools to accurately predict the real daily water demand either at network distribution level or at farm level are absolutely necessary to achieve an optimal management of the pressurized irrigation networks.

1.4. Artificial Intelligence in irrigation networks

In recent years, the development of new sensors capable to collect automatically thousands of data of the whole water-soil-plant-atmosphere system, offer new possibilities for the optimization of the water supply and irrigation processes. Thanks to the application of techniques such as Artificial Neural Networks, Fuzzy Logic or Genetic Algorithms, it is possible to develop tools

for the prediction of the daily or even hourly water demand. The seamless integration of these methodologies forms the core of Artificial Intelligent. The synergism of these techniques allows soft computing to incorporate human knowledge effectively, deal with imprecision and uncertainty, and design adaptation strategies for better performance.

Artificial Intelligent techniques have been applied to resolve many problems of water resources management and planning such as: modelling monthly, daily and hourly rainfall-runoff process (Anctil and Rat 2005; Agarwal et al. 2006), real-time river level and lake stage forecasting (Ondimu and Murase 2007) or the optimization of the multi-crop pattern plan using Fuzzy Logic and Genetic Algorithm (Rezaei et al. 2017). However, there are very little experiences about their application for irrigation water demand forecasting (Pulido-Calvo and Gutiérrez-Estrada, 2009).

In essence, the integration of artificial intelligence methodologies (Artificial Neural Networks, Decision Trees, Fuzzy Logic and Genetic Algorithm) with tools to optimize the operation of irrigation water distribution networks (such as sectoring and critical points control) including the interactions within the systems at different scales (distribution network and on-farm systems) will provide a useful decision support system for managers. This advanced decision support system will entail

better use of water and energy resources as well as greater economic benefits for farmers.

1.5. References

Abadia R, Rocamora C, Ruiz A, Puerto H (2008) Energy efficiency in irrigation distribution networks I: Theory. Biosyst Eng 101:21–27

Agarwal A, Mishra SK, Ram S, Singh JK (2006) Simulation of Runoff and Sediment Yield using Artificial Neural Networks. Biosyst Eng 94:597–613

Anctil F, Rat A (2005) Evaluation of neural network streamflow forecasting on 47 watersheds. J Hydrol Eng 10:85–88

Carrillo Cobo MT., Rodríguez Díaz JA, Montesinos P, et al (2011) Low energy consumption seasonal calendar for sectoring operation in pressurized irrigation networks. Irrig Sci 29:157–169

Conforti P (Editor) (2011) Looking Ahead in World Food and Agriculture: Perspectives to 2050. Food and Agriculture Organization, Rome

Corominas J (2010) Agua y energía en el riego en la época de la sostenibilidad. Ing del agua 17:219–233

Daccache A, Knox JW, Weatherhead EK, et al (2014) Implementing precision irrigation in a humid climate. Recent experiences and on-going challenges. Agric Water Manag 147:135–143

FAO (2011) The state of the world's land and water resources for food and agriculture (SOLAW). Managing systems at risk. Food and Agriculture Organization of the United Nations, Rome and Earthscan, London

Fernández García I, Montesinos P, Camacho Poyato E, Rodríguez Díaz JA (2015) Energy cost optimization in pressurized irrigation networks

Fernández García I, Rodríguez Díaz J a., Camacho Poyato E, Montesinos P (2013) Optimal Operation of Pressurized Irrigation Networks with Several Supply Sources. Water Resour Manag 27:2855–2869

Fernández García I, Rodríguez Díaz JA, Camacho Poyato E, et al (2014) Effects of modernization and medium term perspectives on water and energy use in irrigation districts. Agric Syst 131:56–63

García-Ruiz JM, López-Moreno JI, Vicente-Serrano SM, et al (2011) Mediterranean water resources in a global change scenario. Earth-Science Rev 105:121–139.

Hunt CE (2004) Thristy Plant-Strategies for Sustainable Water Management. New York

IDAE Spanish Institute for Energy Diversification and Savings (2005) Spanish Energy Efficiency Strategy. Action Plan 2005-2007

IDAE Spanish Institute for Energy Diversification and Savings (2007) Spanish Energy Efficiency Strategy. Action Plan 2008-2012

IDAE Spanish Institute for Energy Diversification and Savings (2011) Spanish Energy Efficiency Strategy. Action Plan 2011-2020

INE Instituto Nacional de Estadística (2016) Encuesta sobre el uso del agua en el sector agrario (año 2014). Madrid. Spain. http://www.ine.es/dyngs/INEbase/es/categoria.htm?c=Estadistica_P&cid=1254735976602. Accessed 4 Jun 2017

Lopez-Gunn E, Zorrilla P, Prieto F, Llamas MR (2012) Lost in translation? Water efficiency in Spanish agriculture. Agric Water Manag 108:83–95

MAGRAMA Spanish Ministry of Agriculture Food and Environment (2016) Encuesta sobre superficies y rendimientos de cultivo 2016. 158

MARM Ministry of Environment and Rural and Marine Affair (2010) National Strategy for Sustainable Modernization of Irrigation-Horizon 2015. Madrid, Spain

Moreno MA, Carrión PA, Planells P, et al (2007) Measurement and improvement of the energy efficiency at pumping stations. Biosyst Eng. doi: 10.1016/j.biosystemseng.2007.09.005

Ondimu S, Murase H (2007) Reservoir Level Forecasting using Neural Networks: Lake Naivasha. Biosyst Eng 96:135–138

Playán E, Mateos L (2006) Modernization and optimization of irrigation systems to increase water productivity. Agric Water Manag 80:100–116

Plusquellec H (2009) Modernization of large-scale irrigation systems: is it an achievable objective or a lost cause. Irrig Drain 58: S104–S120

Pulido-Calvo I, Gutiérrez-Estrada JC (2009) Improved irrigation water demand forecasting using a soft-computing hybrid model. Biosyst Eng 102:202–218

Rezaei F, Safavi HR, Zekri M (2017) A Hybrid Fuzzy-Based Multi-Objective PSO Algorithm for Conjunctive Water Use and Optimal Multi-Crop Pattern Planning. Water Resour Manag. doi: 10.1007/s11269-016-1567-4

Rodríguez-Díaz JA, Pérez L, Camacho E, Montesinos P (2011) The paradox of irrigation scheme modernization: more efficient water use linked to higher energy demand. Spanish J Agric Res 9:1000–1008

Rodríguez Díaz JA, López Luque R, Carrillo Cobo MT., et al (2009) Exploring energy saving scenarios for on-demand pressurised irrigation networks. Biosyst Eng 104:552–561

Rodríguez Díaz JA, Montesinos P, Camacho Poyato E (2012) Detecting Critical Points in On-Demand Irrigation Pressurized Networks - A New Methodology. Water Resour Manag 26:1693–1713

2. Objectives and thesis structure

2.1. Objectives

The overall objective of this thesis is to develop strategies to optimize water and energy use in Irrigation Districts both at water distribution network and at farm levels, considering the uncertainty in the irrigation scheduling caused by farmer's behavior.

With this aim, the following specific objectives have been formulated:

1. Development of a methodology to analyze the impacts of pump station's management on the irrigation system at the critical farms (farms supplied by critical points or hydrants with the highest energy requirements).

2. Development of a comprehensive model that integrates soil water balance and irrigation uniformity on-field scale to assess their impacts on crop yield. The model is developed within an innovative GIS framework to facilitate the evaluation of a wide range of impacts related to agricultural aspects and to water and energy use in irrigated agriculture.

3. Optimization of sectors operation in pressurized irrigation networks considering crop water requirements and soil water balance at farm level with the aim to maximize the farmer's profit and minimize the energy cost at pumping stations.

4. Development of a short-term forecasting model of daily irrigation water demand at irrigation district level using Artificial Neural Networks and Genetic Algorithms.

5. Modelling the farmer´s behavior and irrigation events forecast, combining Decision Trees and Genetic Algorithms into a single hybrid model.

6. Modelling the farmer´s behavior and daily water demand forecast at farm level, by a hybrid methodology, which combines Artificial Neural Networks, Fuzzy Logic and Genetic Algorithms.

2.2. Thesis structure

This thesis is organized in nine chapters. Following the introduction **(chapter 1)** and the objectives **(chapter 2)**, the Thesis is organized in two thematic sections, with three chapters each of them. The first section addresses the optimization of the water and energy efficiency in water distribution network and irrigation systems following a holistic approach. The second section focus on irrigation demand forecasting to improve the management of

water and energy resources. Finally, the conclusions are presented in chapter **9**.

Section I, titled "*Energy efficiency in water distribution and on-farm irrigation systems*", includes Chapters **3, 4** and **5** and presents methodologies to increase the energy use efficiency in both water distribution systems and on-farm irrigation systems. **Chapter 3** approaches the interactions between on-demand water distribution networks and irrigation systems' performance in critical points with the aim to enhance the overall efficiency of the irrigation infrastructure with minimal costs. This chapter has been published under the title "Critical points: interactions between on-farm irrigation systems and water distribution network" (2014) by González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA in *Irrigation Science.*

The main limitation of the work carried out in Chapter **3** is the simplistic approach followed to estimate crop yields. The aim of **Chapter 4** is to reduce this shortcoming thanks to a new methodology to assess in-field impacts of irrigation heterogeneity on crop yield using geospatial analysis and crop modeling techniques. This chapter was done in collaboration with researchers from Cranfield University (UK) and it has been published under the tittle "Modelling impacts of precision irrigation on crop yield and in-field water management" (2017) by González Perea R, Daccache A, Rodríguez Díaz JA, Camacho

Poyato E and Knox JW in *Precision Agriculture*. **Chapter 5** presents a new methodology for water distribution networks sectoring which considers the soil water balance and irrigation scheduling in each farm, using the multiobjetive genetic algorithm NSGA-II as optimization tool. This chapter has been published in *Water Resources Management* as "Optimization of Irrigation Scheduling Using Soil Water Balance and Genetic Algorithms" (2016) by González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA.

Section II, titled "*Irrigation demand forecasting models at different scales*", consists of **Chapters 6** to **8**. In this section, several models are developed to predict the irrigation demand at different spatial and time scales: at irrigation district and farm level, and at daily and hourly scale. In **Chapter 6** a short-term forecasting model of daily irrigation water demand at irrigation district level using Artificial Neuro-Genetic Networks is developed. This chapter has been published under the tittle "Irrigation Demand Forecasting Using Artificial Neuro-Genetic Networks" (2016) by González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA in *Water Resources Management*.

Chapter 7 is aimed at modelling the farmer´s behavior and forecast when the farmer decides to irrigate. This research is included in the article "Prediction of irrigation event occurrence

at farm level using optimal decision trees" (2017) by González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA, published in *Water Resources Management*. **Chapter 8** presents a model attempting to predict the hourly water demand at farm level using Artificial Neural Networks, Fuzzy Logic and Genetic Algorithms. This chapter corresponds to the article "Farmer's behaviour modelling by the prediction of the applied irrigation depth using artificial intelligence" (2017) by González Perea R, Camacho Poyato E, Montesinos P and Rodríguez Díaz JA in *Water Resources Research*.

Finally, **Chapter 9** gather the main conclusions reached in this thesis as well as the avenues for future research in the field of water-energy nexus in irrigated agriculture.

Section I

Energy efficiency in water distribution networks and on-farm irrigation systems

3. Critical Points: Interactions between On-Farm Irrigation Systems and Water Distribution Network

This chapter has been published entirely in the journal "Irrigation Science", González Perea R, Camacho Poyato E, Montesinos P, Rodríguez Díaz JA (2014)

Abstract. In this work, a new model useful to analyze interactions between the on-farm irrigation system supplied by critical points and the water supply network management was developed. The model evaluates the impacts of changes in the pressure head and demand simultaneity (number of open hydrants at a given time) on the irrigation system and evaluates emitter discharge and uniformity. It also estimates the potential reductions in crop yield due to decreased emission uniformity.

The methodology is applied in the Bembézar Irrigation District (Southern Spain). Results show that the additional cost required for providing maximum pressure to the critical field does not offset the increase in crop yield. Hence, an increment from 91.7 % to 92.1 % in yield in the critical field would represent increases in energy consumption from 0.15 to 0.17 kWh per m^{-3} of water. Also, the unit energy cost could be reduced by up to 0.11 kWh

per m^3 without implying significant reductions in crop yield. The importance of a good selection of emitters in the critical fields (fields that are supplied by the critical hydrants) was also evaluated.

Keywords. energy savings, pressurized irrigation, hydraulic modelling

3.1. Introduction

With the aim of increasing the irrigation efficiency and to give farmers maximum flexibility, many water distribution networks have been designed to supply pressurized water and arranged on demand. Thus, some of the obsolete open-channel hydraulic infrastructure has been replaced by new pressurized networks (Plusquellec 2009). This change increases the conveyance efficiency reducing water losses throughout the system. In addition, farmers get a much greater degree of flexibility, allowing the use of more efficient on-farm irrigation systems such as trickle or sprinkler, and therefore increasing uniformity and allowing more frequent irrigation (Lamaddalena et al. 2007; Rodríguez Díaz et al. 2007a; Pérez Urrestarazu et al. 2009).

However, these pressurized networks have significantly increased the energy demand. For example, in Spain, where an ambitious modernization plan of irrigation schemes has been carried out (MAPA, 2001), Corominas (2010) reported than while water use

has been reduced by 21 % from 1950 to 2007, the energy demand was subsequently increased by 657 %. For this reason, several authors have highlighted the necessity of reducing the energy requirements, improving the performance of both the water distribution and on-farm irrigation systems (Pulido-Calvo et al. 2003; ITRC 2005; Moreno et al. 2007; Rocamora et al. 2008; Vieira and Ramos 2009; Daccache et al. 2010).

There are several strategies for energy optimization in pressurized irrigation networks. Network sectoring, where hydrants with similar energy requirements are grouped, is one of the most effective measures (Rodríguez Díaz et al. 2009; Carrillo Cobo et al. 2011; Navarro Navajas et al. 2012). Another energy saving measure is the control of critical points, which are hydrants with high energy requirements (i.e., relatively high elevation and or outlet pressure).

Rodríguez Díaz et al. (2012) developed the WECP (Water and Energy optimization by Critical Point control) algorithm for detecting critical points in pressurized irrigation networks. It was applied in two pressurized irrigation networks in Southern Spain. The results showed that potential energy savings around 10–30 % were possible in each district when the theoretical irrigation requirements were modelled. However, the WECP offered energy saving measures at the distribution network level, not considering the possible on-farm irrigation implications in the fields supplied

by a critical hydrant. Reductions in the pressure head at the pumping station may drastically affect the distribution uniformity of the on-farm irrigation system and have significant negative impacts on crop yields (Smajstrla et al. 1990).

In water distribution systems that operate on demand, flows in pipes are subjected to fluctuations according to the simultaneity of the demand (Rodríguez Díaz et al. 2007a). However, when the water demand is high, the energy losses in pipes are increased and the pressure on hydrants is reduced. Related to this, several modelling approaches have been developed to assess the performance of on-demand systems. For example, the indexed characteristic curve approach (CTGREF Division Irrigation 1979; Bethery et al. 1981) evaluates the overall performance of the distribution system, while the AKLA model (Lamaddalena and Sagardoy 2000) provides more specific information about the percentage of hydrants with sub-optimal performance, their position and the magnitude of their pressure deficit. In the Apulia irrigation district (Italy), the critical hydrant showed a potential pressure variation ranging between 64 m and 24 m when the upstream system discharge fluctuates between 600 and 1,200 L s^{-1}. These fluctuations had important impacts on the on-farm irrigation system performance (Daccache et al. 2010).

In this work, a useful methodology to analyze the impacts of a pump station's management on the on-farm irrigation system is

developed and applied to a real irrigation network in Southern Spain. Thus, pressure head changes in the pump station affect the pressure at critical points and consequently the distribution uniformity of their on-farm irrigation systems. The impact of these pressure variations is evaluated in terms of crop yield.

3.2. Methodology

3.2.1. Study area

The M. D. Bembézar Irrigation District (MDB) (Southern Spain) has a total irrigated area of 11,950 ha (Fig. 3.1). The climate is

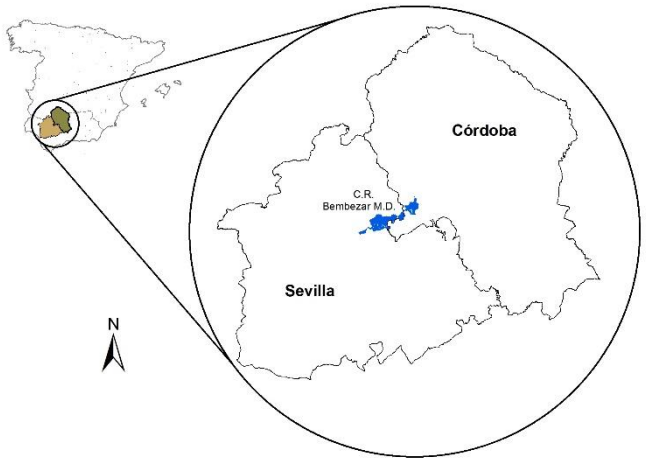


Fig. 3.1. Location of the BMD irrigation district, Spain.

Mediterranean with an annual average rainfall of 604 mm and an average annual temperature of 17.7 °C, with July being the hottest month (36.2 °C mean temperature). Under these circumstances, the average annual reference evapotranspiration (ET_0) is over

1,200 mm. The main crops in the irrigation district are: citrus, cotton, maize and fruit trees.

The water is conveyed from three different reservoirs (Bembézar, 342 Mm^3; Retortillo, 61 Mm^3; José Torán, 101 Mm^3) through a main channel of 40 km length and 12 $m^3\ s^{-1}$ of delivery capacity. Then, eleven pumping stations operate along the main channel to supply water to each sector. The network was designed to supply 1.2 $L\ s^{-1}\ ha^{-1}$ on demand at a minimum operational pressure head at the hydrant level of 35 m of water. Drip irrigation is the most common irrigation method. Sector VII, which covers a total irrigated area of 935 ha (Fig. 3.2), was analyzed in this work.

Critical field

The most critical point was identified in sector VII. Maize was grown in the field that is supplied by the most critical hydrant (critical field), and it has an irrigated area of 4.7 ha. A pressure regulation valve is placed downstream of the hydrant (which reduces the pressure head to 35 m) as well as a filter battery, whose friction losses were estimated to be 7 m.

The irrigation method is trickle, with an emitter spacing of a = 0.5 m and b = 1.8 m (spacing between emitters x spacing between

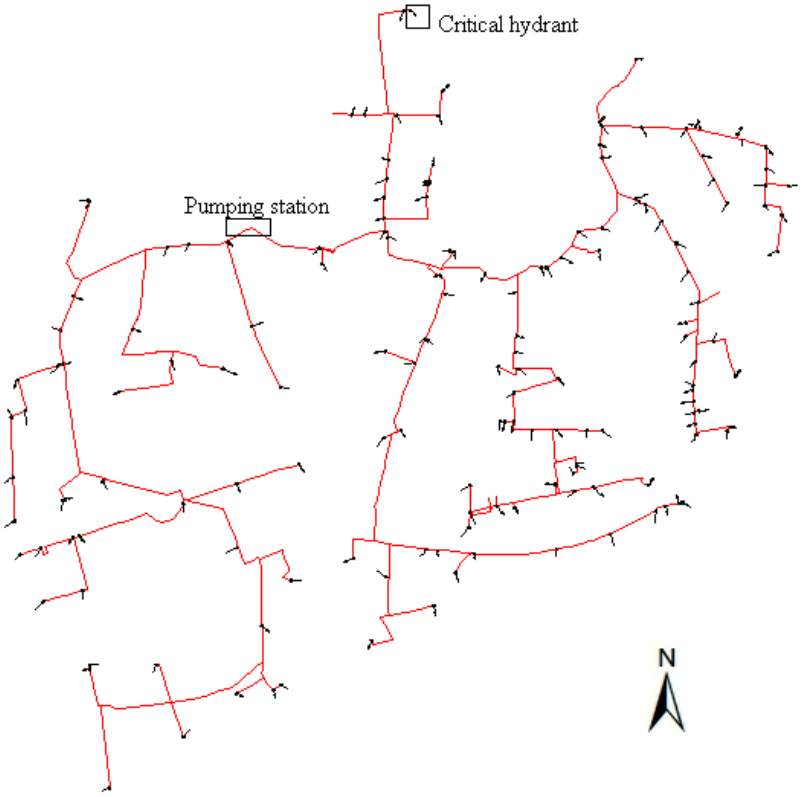


Fig. 3.2. Distribution network (sector VII, BMD) and location of the critical hydrant.

laterals). The nominal emitters flow is 2.2 L h^{-1}, and the pressure-compensation range varies from 10 m to 40 m. The emitter flow pressure head equation is:

$$q = \alpha \cdot h^{x} \qquad [3.1]$$

where q is the flow rate (L h^{-1}); h is the pressure head (m); α is the discharge coefficient (L h^{-1} m^{-x}); and x is the pressure head exponent. In this work, x = 0.04 and α = 1.79 L h^{-1} $m^{-0.04}$.

3.2.2. Problem formulation

Initially, the critical point detection was carried out using the WECP algorithm (Rodríguez Díaz et al. 2012). This algorithm identifies the critical points (most energy demanding hydrants) through several thousand network operation simulations under randomly generated loading conditions (different patterns of open/closed hydrants).

Then, a new model for analyzing the impacts of the network's management on critical fields was developed (Fig. 3.3).

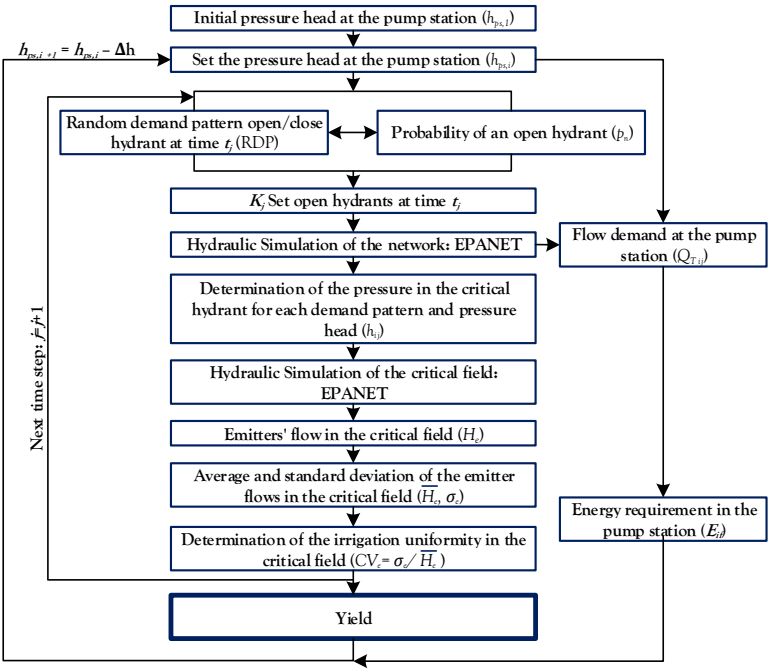


Fig. 3.3. Flowchart of the critical field evaluation model.

The model simulated the network's behavior during the peak water demand month for different pressure heads. Furthermore,

the model linked the simultaneity of the water demand and the pressure at hydrants, considering the probability of an open or closed hydrant as described in Carrillo Cobo et al. (2011). Thus, the theoretical daily average irrigation needs in the irrigation district per month, and hydrant (mm) was estimated as described in FAO 56 (Allen et al. 1998). This information was transformed into irrigation needs in the peak water demand period, *IN* (L ha^{-1} day^{-1}), and was used to estimate the daily irrigation time required in the peak month, t_n (h day^{-1}), per hydrant (n):

$$t_n = \frac{1}{3{,}600} \cdot \frac{IN}{q_{max}} \quad [3.2]$$

being q_{max} was the maximum flow allowed per hydrant (L s^{-1} ha^{-1}). Then, the open outlet probability, p_n, in the month of maximum demand (Clément 1966) was calculated for each hydrant (n) as follows:

$$p_n = \frac{t_n}{t_a} \quad [3.3]$$

where t_a is the daily irrigation availability time (24 h in on demand systems).

Finally, it analyzed the impacts of changes in pressure head on the on-farm irrigation system's behavior. The model was implemented in MATLAB, using the hydraulic simulator

EPANET (integrated through its dynamic link library, DLL; Rossman 2000).

Hydraulic behavior at different pressure heads at the pump station

Initially, the model fixed the pressure head at the pump station, $h_{ps,i}$, then j random demand patterns, *RDP*, (set of open/closed hydrants) were generated. A random demand pattern was generated for every iteration, j. The open hydrant probability value, p_n, smaller or equal than random numbers generated with random demand patterns, determine the k_j sets of open hydrants; otherwise, hydrants are considered closed with no water demand. The flow demanded for every open hydrant, q_n, was estimated as follow:

$$q_n = q_{max} \cdot S_n \qquad [3.4]$$

where S_n was the irrigated area supplied by hydrant n. Then, the network behavior under each random demand pattern and pressure head value, $h_{ps,i}$, was analyzed using EPANET.

All this process begins with a maximum pressure head ($h_{pi,max}$) and decreases in each iteration (i) in Δh. Thus, the effects that the pressure head at the pumping station and the simultaneity of the demand have in the pressure of the critical hydrant were analyzed. Limits of i and Δh depend on the type of emitters installed in the field.

Power and energy requirements

The demanded flow in the pumping station, Q_{Tij} in ($m^3\ s^{-1}$), was determined for each demand pattern. For each pair of demanded flow and pressure head (loading condition), the power requirements, P_{ij} (kW), at the pumping station were calculated according to the following equation:

$$P_{ij} = \frac{\gamma_w \cdot Q_{Tij} \cdot h_{ph,i}}{\eta} \quad [3.5]$$

where γ_w is the specific weight of water (9.8 kN m^{-3}) and η is the pump system efficiency (a pump efficiency of 0.8 was assumed). Then, the energy consumption in kWh per working day for each loading condition, E_{ij}, was estimated as follows:

$$E_{ij} = P_{ij} \cdot t_n \quad [3.6]$$

Pumped flow, pressure head, power and energy for each loading condition were averaged for the peak month.

Hydraulic behavior of the critical field

Based on a known pressure at the critical hydrant (h_c), the hydraulic behavior of the emitters in the critical field was analyzed and the possible reductions in yield due to variations in the distribution uniformity were estimated. Thus, the on-farm irrigation system was also modeled in EPANET.

Then, the pressure and irrigation depth distribution were calculated for all the emitters in the critical field. The descriptive statistics (mean, $\overline{H_e}$; standard deviation, σ_e; and coefficient of variation, CV_e) for the emitters were estimated. Considering that there are no runoff losses, the mean depth coincides with the total gross applied depth, H_g. The total number of emitters, e, in the critical field was calculated as the field area divided by the emitter spacing:

$$e = \frac{S_n}{a \cdot b} \qquad [3.7]$$

On-farm irrigation system evaluation

The on-farm distribution uniformity was evaluated using the ratio (H_g/H_r). This relationship was calculated according to Alabanda (2001):

$$\frac{H_g}{H_r} = \frac{1 - C_d}{IE} \qquad [3.8]$$

where IE is the irrigation efficiency. It is the ratio of the net irrigation requirements, (H_n) (mm) to the total gross applied depth, H_g (mm); C_d is the deficit coefficient, which is the ratio of the water deficit ($H_r - H_n$), and the theoretical irrigation requirements, (H_r). These coefficients were described by Losada (1996) and calculated according to Eqs. 3.9, 3.10 and 3.11.

$$IE = \frac{\overline{H_e} - \frac{1}{2} \cdot \frac{\left(\overline{H_e} + \sqrt{3}\sigma_e\right)^2 - {H_r}^2}{2\sqrt{3}\sigma_e} + f \cdot H_r}{\overline{H_e}} \quad [3.9]$$

where *f* is the fraction of the field without water deficit:

$$f = \frac{\overline{H_e} + \sqrt{3}\sigma_e - H_r}{2\sqrt{3}\sigma_e} \quad [3.10]$$

$$C_d = \frac{(1-f) \cdot H_r - \overline{H_e} + \frac{1}{2} \cdot \frac{(\overline{H_e} + \sqrt{3}\sigma_e)^2 - {H_r}^2}{2\sqrt{3}\sigma_e}}{H_r} \quad [3.11]$$

Additionally, the distribution uniformity of the flow is also evaluated, using Christiansen's uniformity coefficient (CU_c; Christiansen 1942):

$$CU_c = 100 \cdot \left[1 - \frac{\sum_{i=1}^{e} |H_i - \overline{H_e}|}{e \cdot H_g}\right] \quad [3.12]$$

where H_i was the applied irrigation depth for every emitter (mm), which is one of the EPANET outputs.

Crop yield estimation

The irrigation uniformity and crop yield reductions were linked using the following equation (Allen et al. 1998):

$$1 - \frac{Y}{Y_{max}} = K_y \cdot (1 - \frac{H_g}{H_r}) \quad [3.13]$$

where Y is the actual crop yield of the crop (kg ha^{-1}); Y_{max} is the maximum crop yield without water stress (kg ha^{-1}); and K_y is the yield response factor.

The farmer's benefit, in € ha^{-1}, was calculated according to Eq. 3.14. The crop production costs are independent of the network's management. Thus, the profit was calculated taking into account only water costs, which was calculated from the energy consumption per unit of irrigation water supplied, in kWh m^{-3} and the energy cost, in € kWh^{-1}.

$$Profit = (Y_c \cdot P_c) - C_w \quad [3.14]$$

where Yc is the crop yield (kg ha^{-1}); P_c is the market price of the agricultural production (€ kg^{-1}); and C_w is the water cost (€ ha^{-1}).

Alternative management scenarios

The analysis of alternative emitters can be easily carried out by changing the parameters in the emitter equation (Eq. 3.1). The influence of the irrigation system in the critical field can be easily evaluated using different pressure–flow equation parameters.

Thus, three alternative emitters according to Eqs. 3.15 (scenario a) and 16 (scenario B) were tested.

$$q = 0.73 \cdot h^{0.47} \quad [3.15]$$

$$q = 0.64 \cdot h^{0.53} \quad [3.16]$$

where q is in (L h^{-1}) and h is pressure head in (m).

3.3. Results and discussion

The model simulated the behavior of the network during the peak month (June). $h_{ps,max}$ was set to 55 m and Δh was 2 m. The i parameter ranged from 1 to 16. The number of iterations j was set to 2,250.

3.3.1. Water demand and pressure in the pump station

As the water distribution network is operated on demand, the supplied flows are subjected to fluctuations in the number of hydrants that are simultaneously open. Thus, flows ranged from 350 L s^{-1} (when low simultaneity occurs) to 840 L s^{-1} when most of the hydrants were open.

The influence of the simultaneity of the demand (set of open hydrants) in the pressure at hydrant level was very small. a linear relationship between the pressure head ($h_{ps,i}$) and the pressure at the critical hydrant (h_c) was identified ($h_c = h_{ps} - 13.9$; $r^2 = 1$). Due to the design criteria of the network (100 % of simultaneity), the energy losses are not too high even when all hydrants are open because the pipes were sized for the maximum demand. Wider ranges of pressure variation at a hydrant can be expected for other irrigation networks where some pipes may be undersized.

3.3.2. Irrigation uniformity in the critical field

The irrigation uniformity decreases when the pressure in the critical field drops below 35 m. Table 3.1 relates the average of the pressures for all the iterations at the critical hydrant ($\overline{h_c}$) and the pressure head at the pumping station with the CV_e and the CU_c in the critical field.

When the pressure in the critical hydrant was above 35 m, the pressure regulator was active. Then the maximum value of CU_c is 99.8 %, and when the pressure head at the hydrant is less than 35 m, the CU_c is slightly reduced from 99.8 to 98.0 %. as the pressure-compensating range of the emitters is from 10 m to 40 m, the CU_c and yield do not vary significantly in this range. Below 10 m of pressure, the emitters are outside their pressure-compensating range and significant reductions in flows are expected.

The spatial distribution of pressures and flows in the critical field is shown in Fig. 3.4 for three different pressure heads. When the pressure head at the pumping station is 53 m, the pressure in the critical hydrant is 39.2 m (Table 3.1) and the pressure regulator

Table 3.1. Relations of the pressure head at the pump station, average hydrant pressure, irrigation uniformity (σ_e, CV_e, CU_c) and crop yield in the critical field.

Pressure head (m)	$\overline{h_c}$ (m)	σ_e	CV_e (%)	CU_c (%)	Crop Yield (%)	Energy consumption per unit of water supplied (kWh m^{-3})
55	41.3	1.1	0.2	99.8	92.1	0.2
53	39.2	1.1	0.2	99.8	92.1	0.2
51	37.3	1.1	0.2	99.8	92.1	0.2
49	35.3	1.1	0.2	99.8	92.1	0.2
47	33.3	1.2	0.2	99.8	91.9	0.2
45	31.3	1.2	0.2	99.8	91.7	0.2
43	29.3	1.7	0.3	99.7	90.7	0.2
41	27.3	1.8	0.4	99.7	90.3	0.1
39	25.3	2.0	0.4	99.7	89.9	0.1
37	23.2	2.2	0.4	99.7	89.4	0.1
35	21.3	2.4	0.5	99.7	88.8	0.1
33	19.3	2.7	0.6	99.6	88.3	0.1
31	17.3	3.0	0.6	99.5	87.6	0.1
29	15.3	4.7	1.0	99.2	86.1	0.1
27	13.3	6.0	1.3	99.0	84.8	0.1
25	11.3	11.3	2.5	98.0	82.3	0.1

was active. As a consequence, all the emitters received adequate pressure, there are minimum variations in the pressure distribution due to the topography of the field, but all emitters operate within the pressure compensating range. In relation to

the flow distribution, all emitters supply the nominal flow (Fig. 3.4a).

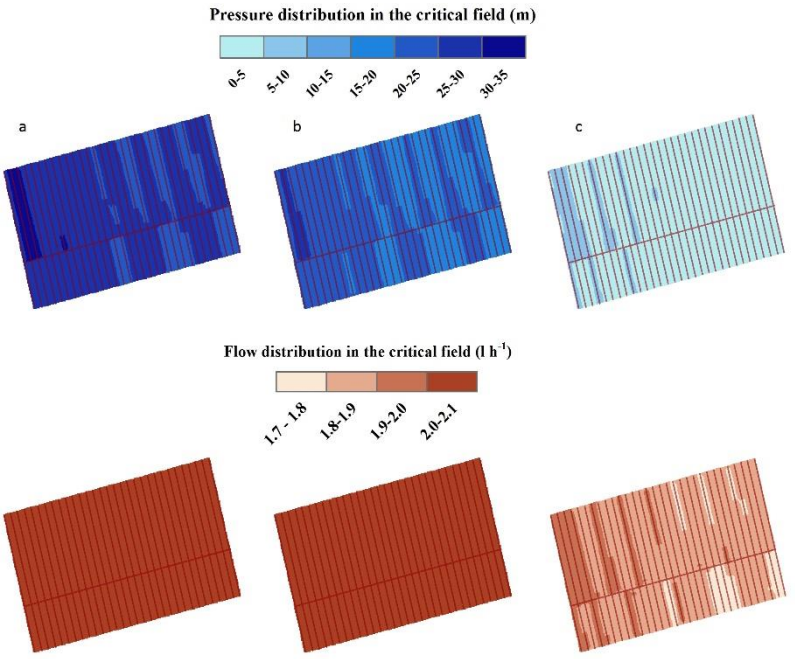


Fig. 3.4. Spatial distribution of pressure and flow in the critical field for **a** 53 m, **b** 43 m and **c** 25 m of pressure head at the pump station.

In Fig. 3.4b, the spatial pressure distribution for a pressure head at the pumping station of 43 m is shown. In this case, the pressure head at the critical hydrant was 29.3 m (Table 3.1), so the pressure regulator was inactive. However, all emitters operated within the auto-compensating range (10 m – 40 m) and the supplied flows are similar to those found in Fig. 3.4a (nominal flow). When the pressure head drops to 25 m, the pressure at the hydrant is 11.3

m and most of the emitters stop working properly, as they operate with less than 10 m of head.

Fig. 3.5 shows the evolution of the coefficient based on emitter flow variations at different pressures in the critical hydrant. When the pressure is higher than 35 m, the *CV* remains constant thanks to the pressure regulator. When the pressure is less than 35 m, the pressure regulator does not work and CV_e increases. Since the emitters are pressure compensating, the CV_e did not change much, while pressure was within of the pressure-compensation range. Given that the emitters are pressure compensating, large changes in CV_e are not expected (it changes from 0.2 % to 1.3 %).

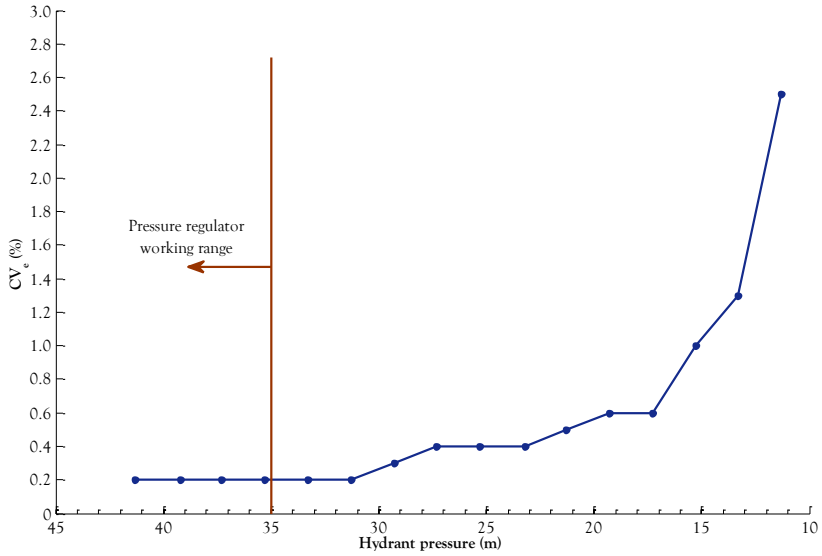


Fig. 3.5. Relation between pressure at a hydrant and the irrigation system's CV.

3.3.3. Yield in the critical field

The total number of emitters in the critical field is estimated from Eq. 3.7. In this case, the critical field had 20,175 emitters. Due to the large number of emitters, the calculation time required for the critical field was too high. Thus, the field was skeletonized, eliminating two out of three branches, and replaced by equivalent consumption points. Therefore, the skeletonized field had 6,725 emitters.

The irrigation time required in the peak month (t_n) was calculated from Eq. 3.2, and it was a constant value of 14.6 h.

Even when the pressure head at a hydrant is 35 m, the crop yield is 92.1 % of the potential yield (Table 3.1). According to the manufacturer, the emitter's nominal flow was 2.2 L h^{-1} and the irrigation events are scheduled according to this value. But in the hydraulic simulations, this nominal flow was not reached at any time even when the pressure was adequate. Furthermore, the emitters were not fully compensating because the pressure exponent of the emitter was, of course, not zero.

When the pressure at a hydrant drops to 11.3 m, the pressure of many emitters is lower than the minimum limit of admissible pressure and the yield is reduced to 82.3 % of the potential crop yield, because the discharge from those emitters is smaller.

Therefore, the crop receives less water than the calculated requirements. The spatial distribution in this case is shown in Fig. 3.4c. Fig. 3.6 shows the relationship between applied water and the theoretical irrigation requirements for each pressure at the hydrant level. When the pressure is reduced, the ratio $(H_g/\overline{H_r})$ is reduced too, so less water is available to the crop.

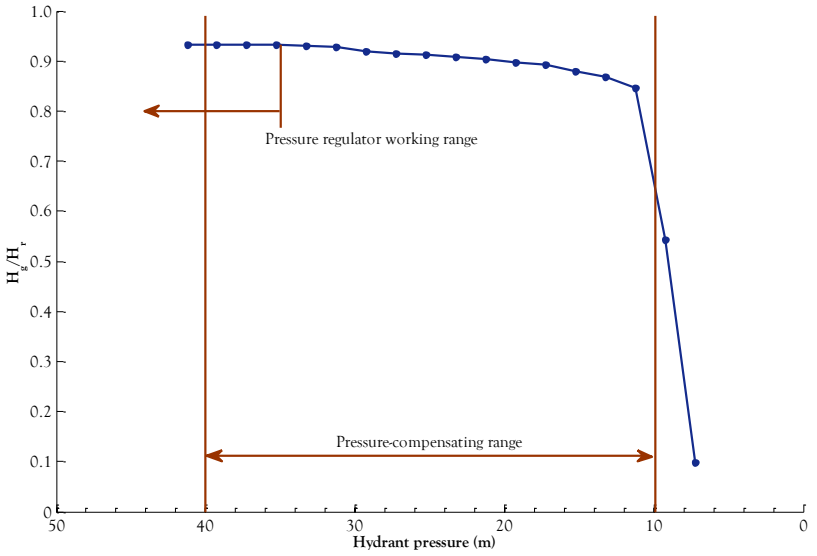


Fig. 3.6. Relationship between $(H_g/\overline{H_r})$ and pressure at a hydrant.

3.3.4. Energy use

The relationship between energy consumption per cubic meter in the peak month and the pressure head at the pumping station is shown in Table 3.1. The current operation of the pumping station is 45 m which provides around 32 m of pressure at the critical hydrant. The average unit energy consumption in the

current management is 0.15 kWh m^{-3}. In this case, the yield losses are a bit more than 8 %. However, if the pressure head was reduced to 33 m, the crop yield losses would be slightly smaller (12 %), but the unit energy consumption would be 0.1 kWh m^{-3} for all the water supplied by the pumping station in June. The system can even operate below this pressure.

The current water consumption in June was 972,486 m^3 (data provided by the irrigation district's staff) and, assuming a unit energy cost of 0.10 € kWh^{-1}, the energy cost, in the current condition, is 14,587 € (assuming similar irrigation scheduling for the whole irrigation season). If the system operated under 33 m at the pump station, the energy costs would be 10,697 €, which represents a savings of 27 % for the irrigation district in the peak month. This operation option does not cause significant losses in yield in the critical field. This fact means that the yield losses that may occur in the critical field are much lower than the increase in energy costs needed to provide more pressure at the critical hydrant. Finally, if the pressure at the pumping station was 49 m, the critical hydrant would receive the adequate pressure (35 m) and the energy costs would be 16,532 €. In this case, the maximum yield is achieved (92.1 %).

According to the annual statistics of the agriculture department of Andalucía (Spain; Agriculture department of Andalucía 2009),

maize had an average yield of 10,348 kg ha^{-1}, with an average market price of 0.14 € kg^{-1}. The cost of water for the critical field in June is shown in Table 3.2. Thus, profits of the critical field are 1,035.81 € ha^{-1}, for the current pressure at the pumping station. If the pressure head was 49 m, profits in the critical field decrease to 1,023.13 € ha^{-1}; that is, the increase in agricultural production value is less than the increase in the water cost. On the other hand, if the network operated at 33 m at the pumping station, profits in the critical field would be 1,080.10 € ha^{-1}, 25.90 € ha^{-1} more than the current condition (Table 3.2).

Table 3.2. Profit of the critical field.

Pressure head (m)	$\overline{h_c}$ (m)	Yield (%)	Y_c (kg ha^{-1})	P_c (€ ha^{-1})	C_w (€ m^3)	C_w (€ ha^{-1})	Profit (€ ha^{-1})
49	35.3	92.1	9.5	0.1	0.017	311.2	1,023.1
45	31.3	91.7	9.5	0.1	0.015	274.4	1,054.1
33	19.3	88.3	9.1	0.1	0.011	199.1	1,080.1

3.3.4. Sensitivity to other emitters

The effects of different irrigation emitters were tested in the model. Thus, two scenarios A and B were analyzed, with flow–pressure curves shown in Eqs. 3.15 and 3.16, respectively.

Scenarios A and B

The emitters in the scenarios A and B are not pressure compensating, so greater changes in flow are expected due to variations in pressure. The nominal flow (2.2 L h^{-1}) in the emitter is achieved when the pressure at the critical hydrant is 19.3 m and 33 m, respectively at the pump station.

The model was run for these two scenarios. The CV_e in scenario B varied from 8.1 % for the maximum simulated pressure head (33 m) to 47.0 % for the minimum pressure head (23 m). For the same pressure range, the CV_e in scenario B changes from 9.2 % to 46.9 % (Table 3.3). The CU_c ranged from 93.5 % to 62.5 % in scenario A and from 92.7 % to 62.6 % in scenario B. The sensitivity to changes in pressure head is higher than for the current emitters (Eq. 3.1), so they do not represent the best option for the critical field.

When yields are analyzed, both emitters achieve 100 % when the pressure head is 33 m, but it drops rapidly when the pressure is reduced (37.2 % and 34.6 %, respectively, when the pressure head drops to 23 m; Fig. 3.7). The closer relationship between flow and pressure lead to a poorer uniformity, so the higher spatial variability in the emitters' discharge (Fig. 3.8). Also, contrary to what happened when the current emitters where modeled, the ratio of applied and theoretical irrigation depths is very sensitive to pressure changes (Fig. 3.9).

Table 3.3. Average hydrant pressure, standard deviation (σ_e), CV_e, CU_c and yield for the scenarios A and B.

Pressure head (m)	Scenario A					Scenario B				
	$\overline{h_c}$ (m)	σ_e	CV_e (%)	CU_c (%)	Crop Yield (%)	$\overline{h_c}$ (m)	σ_e	CV_e (%)	CU_c (%)	Crop Yield (%)
33	19.3	44.1	8.1	93.5	100	19.2	50.1	9.2	92.7	100
31	17.3	46.7	9.4	92.5	90.9	17.3	52.3	10.5	91.6	90.9
29	15.2	50.5	11.2	91.1	80.7	15.3	55.6	12.5	90.1	79.7
27	13.3	56.1	14.2	88.7	69.2	13.3	60.5	15.6	87.5	67.3
25	11.3	66.1	19.9	84.2	55.7	11.3	69.4	21.7	82.7	53.0
23	9.3	115.5	47.0	62.5	37.2	9.3	109.4	46.9	62.6	34.6

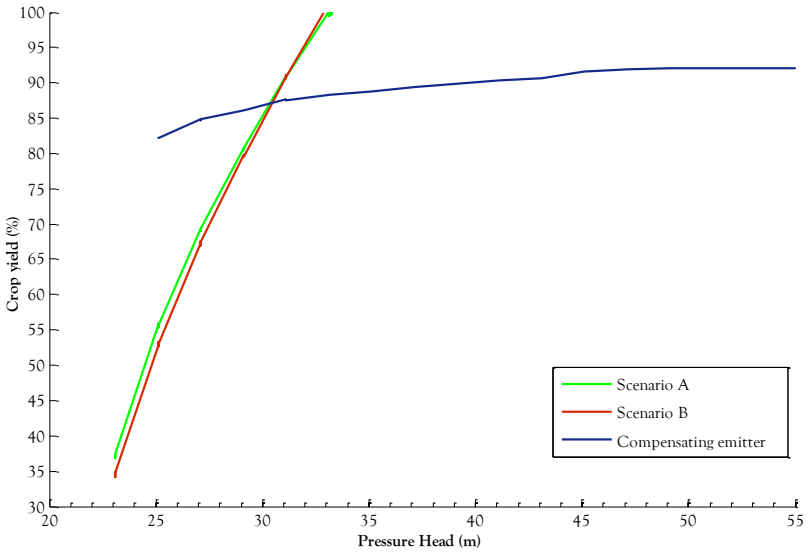


Fig. 3.7. Relationship between crop yield in the critical field and pressure head (m).

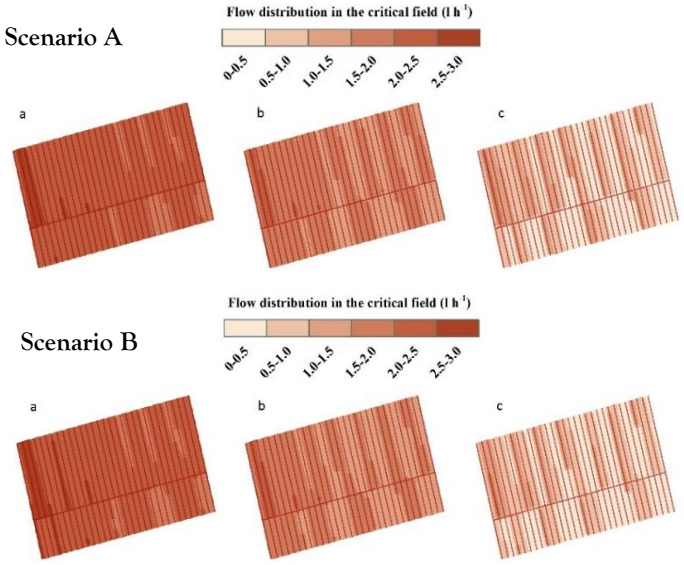


Fig. 3.8. Spatial flow distribution in the critical field of the non-compensating emitters in the critical field for scenarios A and B for 33 (a), 27 (b) and 21 (c) m, respectively, of pressure head at the pump station.

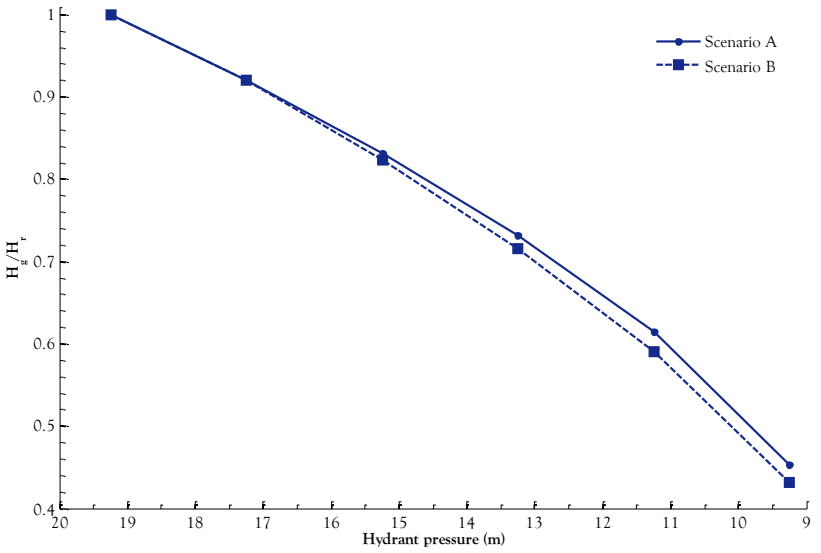


Fig. 3.9. Relationship between $(H_g/\overline{H_r})$ and hydrant pressure for scenarios A and B.

3.4. Conclusions

In this work, a new methodology to simulate the interactions between on-demand water distribution systems and irrigation performance in critical points was developed and applied in the BMD irrigation district. On-demand irrigation implies a significant expenditure in energy which is even higher when some critical points are responsible for a large percentage of the total pressure head. Thus, effective management of the critical points is necessary to enhance the overall efficiency of the irrigation infrastructure with minimal costs. However, detailed analysis at the water distribution and on-farm irrigation systems levels is needed before the adoption of improvement measures.

In this particular case, the results showed that the additional cost required for giving maximum pressure in the critical point does not offset the increase in yield. Here, an increment from 91.7 % to 92.1 % in yield in the critical field would represent increments in energy consumption from 0.15 kWh m^{-3} to 0.17 kWh m^{-3} and an increment of 8.5 % in the energy consumption in the peak demand month. This network management implies an increase in the cost of water in the critical field of 36.9 € ha^{-1} and a reduction in profits of 31.0 € ha^{-1}.

On the other hand, the unit energy cost could be reduced by up to 0.11 kWh m^{-3}, without causing significant reductions in yield, by setting the pressure head to 33 m. Under these conditions, the profit in the critical field would be 1,080 € ha^{-1} or 26 € ha^{-1} more than the current condition. an appropriate selection of emitters in the critical fields is essential to ensure optimal performance of the irrigation system.

Acknowledgments. This research is part of the AMERE project (AGL2011-30328-C02-02), funded by the Spanish Ministry of Economy and Competitiveness. Authors gratefully acknowledge Prof. Gary Merkley for his valuable comments.

3.5. References

Agriculture department of Andalucía (2009) Statistical annual report on agricultural and fisheries of Andalucía. Junta de Andalucía. Spain

Alabanda JS (2001) Evaluación y Manejo de Sistemas de Riego. Graduation thesis. University of Córdoba, Spain

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage, Rome, Italy

Bethery J, Meunier M, Puech C (1981) Analyse des défaillances et étude du renforcement des réseaux d'irrigation par aspersion. 297–324

Carrillo Cobo MT., Rodríguez Díaz JA, Montesinos P, López Luque R, Camacho Poyato E (2011) Low energy consumption seasonal calendar for sectoring operation in pressurized irrigation networks. Irrig Sci 29:157–169

Christiansen JE (1942) Irrigation by sprinkling. California agricultural experiment station bulletin, vol 670

Clément R (1966) Calcul des débits dans les réseaux d'irrigation fonctionant a la demande

Corominas J (2010) Agua y energía en el riego en la época de la sostenibilidad. Ing del agua 17:219–233

CTGREF Division Irrigation (1979) Programme ICARE. Calcul des caractéristiques indicées. NoteTechnique 6

Daccache A, Lamaddalena N, Fratino U (2010) On-demand pressurized water distribution system impacts on sprinkler network design and performance. Irrig Sci 28:331–339

ITRC (2005) CEC agricultural peak load reduction program. California Energy Commission, USA

Lamaddalena N, Fratino U, Daccache A (2007) On-farm Sprinkler Irrigation Performance as affected by the Distribution System. Biosyst Eng 96:99–109

Lamaddalena N, Sagardoy JA (2000) Performance analysis of on-demand pressurized irrigation systems. 132.

Losada A (1996) El Riego: Fundamentos Hidráulicos. Córdoba, Spain

MAPA, Ministry of Agriculture, Fisheries and Food (2001) National Irrigation Plan-Horizon 2008. Madrid, Spain

Moreno MA, Carrión PA, Planells P, Ortega JF, Tarjuelo JM (2007) Measurement and improvement of the energy efficiency at pumping stations. Biosyst Eng. doi: 10.1016/j.biosystemseng.2007.09.005

Navarro Navajas JM, Montesinos P, Poyato EC, Rodríguez Díaz J a. (2012) Impacts of irrigation network sectoring as an energy saving measure on olive grove production. J Environ Manage 111:1–9

Pérez Urrestarazu L, Rodríguez Díaz JA, Camacho Poyato E, López Luque R (2009) Quality of service in irrigation distribution networks: Case of Palos de la Frontera irrigation district (Spain)

Plusquellec H (2009) Modernization of large-scale irrigation systems: is it an achievable objective or a lost cause. Irrig Drain 58:S104–S120

Pulido-Calvo I, Roldán J, López-Luque R, Gutiérrez-Estrada JC (2003) Water delivery system planning considering irrigation simultaneity

Rocamora MC, Abadia R, Ruiz A (2008) Ahorro y Eficiencia energética en las Comunidades de Regantes

Rodríguez Díaz JA, Camacho Poyato E, López Luque R (2007) Model to Forecast Maximum Flows in On-Demand Irrigation Distribution Networks. J Irrig Drain Eng 133:222–231

Rodríguez Díaz JA, López Luque R, Carrillo Cobo MT, Montesinos P, Camacho Poyato E (2009) Exploring energy saving scenarios for on-demand pressurised irrigation networks. Biosyst Eng 104:552–561

Rodríguez Díaz JA, Montesinos P, Camacho Poyato E (2012) Detecting Critical Points in On-Demand Irrigation Pressurized Networks - A New Methodology. Water Resour Manag 26:1693–1713

Rossman L (2000) EPANET 2. User manual. US Environmental Protection Agency (EPA), USA

Smajstrla AG, Boman BJ, Clark GA, Haman DZ, Pitts DJ, Zazueta FS (1990) Field evaluation of irrigation systems: solid set or portable sprinkler systems

Vieira F, Ramos HM (2009) Optimization of operational planning for wind/hydro hybrid water supply systems. Renew Energy. doi: 10.1016/j.renene.2008.05.031

4. Modelling Impacts of Precision Irrigation on Crop Yield and in-field Water Management

This chapter has been published entirely in the journal "Precision Agriculture", González Perea R, Daccache A, Rodríguez Díaz JA, Camacho Poyato E, Knox JW (2014)

Abstract. Precision irrigation technologies are being widely promoted to resolve challenges regarding improving crop productivity under conditions of increasing water scarcity. In this paper, we describe the development of an integrated modelling approach involving the coupling of a water application model with a biophysical crop simulation model (Aquacrop) to evaluate the in-field impacts of precision irrigation on crop yield and soil water management. The approach allows for a comparison between conventional irrigation management practices against a range of alternate so-called 'precision irrigation' strategies (including variable rate irrigation, VRI). It also provides a valuable framework to evaluate the agronomic (yield), water resource (irrigation use and water efficiency), energy (consumption, costs, footprint) and environmental (nitrate leaching, drainage) impacts under contrasting irrigation management scenarios. The approach offers scope for including feedback loops to help define

appropriate irrigation management zones and refine application depths accordingly for scheduling irrigation. The methodology was applied to a case study in eastern England to demonstrate the utility of the framework and the impacts of precision irrigation in a humid climate on a high-value field crop (onions). For the case study, the simulations showed how VRI is a potentially useful approach for irrigation management even in a humid environment to save water and reduce deep percolation losses (drainage). It also helped to increase crop yield due to improved control of soil water in the root zone, especially during a dry season.

Keywords. AquaCrop, variable rate irrigation, onion, sprinklers, water resources

4.1. Introduction

In order to meet future food demands from a rising global population whilst minimizing any environmental impact, a commensurate increase in agricultural productivity (yield) coupled with improvements in water and nutrient efficiency will be necessary (Monaghan et al. 2013; Kumar et al. 2016). In this context, irrigated agriculture will play a critical role supporting increased production in arid and semi-arid regions, and enhancing crop quality through supplemental irrigation in temperate or humid regions (Daccache et al. 2014b; De Paz et al.

2015). However, freshwater availability and abstraction to support an expanding agricultural sector will need to be balanced against competing demands for domestic (household) water supply, water for industrial processing and to support environmental flows and protect ecosystems. A changing climate with greater rainfall uncertainty will exacerbate the situation and create severe challenges in managing and allocating freshwater supplies (Falloon and Betts 2010). The reliability of water resources is also a limiting factor for economic development in many water-stressed countries (Daccache et al. 2014a). With agriculture accounting for nearly three quarters (70 %) of all freshwater withdrawals and over 90 % of total consumptive water use (Siebert et al. 2010) this will inevitably lead to 'irrigation hotspots' where agricultural water demand exceeds available supplies (Knox et al. 2012). Taking into account current pressures on water resources and projected future increases in irrigated area, the agricultural sector needs to do more with less, increasing water productivity (t ha^{-1}) by improving water efficiency and producing more 'crop per drop' (Monaghan et al. 2013).

Various researchers (Fereres et al. 2011; González Perea et al. 2016) have developed decision support tools to help increase crop productivity and improve irrigation use efficiency. This intensification of agriculture will require growers to become more specialized and for many, investment in irrigation will likely be

justified on the basis of helping to increase productivity and profitability. Due to increasing demand from consumers for high quality fruit and vegetable products coupled to rising production (input) costs, farming businesses are starting to critically evaluate the impacts of irrigation non-uniformity on resource use, production and crop returns. In response, the research community have developed various definitions for 'precision irrigation' (PI). For example, Smith and Baillie (2009) defined PI to include the accurate and precise application of water to meet specific requirements of individual plants or management units to minimize adverse environmental impact or the application of water to a given site and timing to support optimum crop production, profitability or some other management objective. In this study, that definition by Smith and Baillie (2009) was similarly adopted with the PI concept representing a more holistic and adaptive approach to precision irrigation water management, rather than relating to only one method of application. Their definition also attempts to integrate the various factors influencing crop, soil and water management more closely with those that impact on irrigation engineering and hydraulic performance.

From a precision agriculture perspective, a number of questions emerge regarding how PI might be modelled and integrated with biophysical crop simulation to evaluate options to save water,

improve yield and support sustainable intensification. Reducing both energy use and the environmental impacts of irrigation abstraction, particularly in river basins or regions where irrigation demand is concentrated and/or where water resources are scarce, are also important drivers for change. For farming businesses involved in high-value crop production, where quality assurance is a major determinant of profitability, PI also offers potential to reduce crop variability and improve post-harvest quality. However, a number of fundamental questions remain. These include the importance of PI definition within a modelling framework, how PI relates to modelling approaches used to assess precision agriculture, and whether irrigation water distribution can be modelled at a field scale that is then geospatially compatible with biophysical crop modelling approaches.

In most studies, the spatial and temporal effects of irrigation heterogeneity on crop production are nearly always lumped together with management variables along with tillage, fertiliser management, seed rates and crop rotation. This is probably due to the spatial and temporal complexity that exists with irrigation, inherent spatial variability in soils and the lack of models capable of simulating the spatial distribution of irrigation without extensive model calibration. Thus, it is difficult to determine the in-field effects of irrigation management strategies including adoption of PI technologies on crop yield and other soil

management practices. In contrast, the use of crop simulation models in crop production, irrigation management, and climate change impacts has proven to be invaluable in improving our knowledge of the functioning of agricultural systems (Fraisse et al. 2006; Thorp et al. 2008; Casadesús et al. 2012). In this context, the Food and Agriculture Organization (FAO) has developed AquaCrop (Steduto et al. 2012), a crop water productivity model focused on simulating water-limited attainable crop yield. Depending on the objective and spatial and temporal scale of analysis, AquaCrop may be applied in different ways. Most studies to date have focused on its application at the operational and tactical scales, running the AquaCrop model at field scale, facilitated by its user-friendly interface (Raes et al. 2009), albeit designed for single runs. For applications at a more strategic level, the AquaCrop model can be applied over larger areas or for longer time periods, requiring a much larger number of model runs, a feature that is not possible with the standard AquaCrop version. However, the development of an AquaCrop plug-in program (FAO 2012) has facilitated the option of multiple simulations, by running a pre-defined list of projects in the standard user interface of the AquaCrop program.

Nevertheless, there is still the need to manually create project input files, requiring lengthy times to scale up AquaCrop applications from a few simulations to multiple runs. Thus, to

eliminate the time-consuming task of manually generating AquaCrop input and project files for multi-run simulation and to adapt it to be compatible with a GIS platform, Lorite et al. (2013) developed two independent tools (AquaData and AquaGIS). AquaData has been imbedded into AquaGIS generating a single package to facilitate file input to data visualization from the AquaCrop simulation. However, there is still a need to combine this improved functionality with an irrigation simulation model to take into account different irrigation strategies such as variable rate irrigation (VRI).

In this paper, a novel integrated modelling approach has been developed to assess in-field impacts of irrigation heterogeneity on crop yield and soil water management practices providing an innovate framework for evaluating wider agronomic and energy impacts. The study specifically considers how application non-uniformity typically observed under conventional overhead irrigation systems compares against so-called precision irrigation (PI) management, and from this, the consequent impacts on crop yield. The approach has been developed and applied to a case study field site in Eastern England representing an intensively managed farm production system involved in growing high value field scale vegetables for the premium retail (supermarket) sector. The research informs discussions regarding the justification of PI implementation in temperate environments where irrigation is

supplemental to rainfall. However, the approach developed is equally applicable to cropping systems in more arid or semi-arid agroclimatic environments.

4.2. Methodology

Various authors have recently investigated the potential for PI in outdoor agriculture based on field experimentation (García Morillo et al. 2015; Haghverdi et al. 2016; Mitchell et al. 2016). Here we develop an integrated modelling approach involving the coupling of a deterministic water/irrigation application model (WAM) with a biophysical crop model (AquaCrop) (Steduto et al. 2012) to simulate the impacts of irrigation heterogeneity caused, for example, due to wind drift, irrigation system pressure variation and/or sprinkler overlapping on crop growth and yield at the field scale. This allows for comparison between conventional irrigation versus alternative PI management strategies, and provides an innovative framework for evaluating wider agronomic, water resource (irrigation use and water efficiency), energy (consumption, cost, footprint) and environmental (nitrate leaching, drainage) impacts under different management scenarios. It also provides the potential for including feedback loops to help define irrigation management zones (IMZ) corresponding to areas within a single field which could be delimited for variable water and nutrient and water

management strategies. The approach was developed to evaluate overhead irrigation under a mobile hose-reel fitted with a boom, a system which is widely used in NW Europe and other humid climates where irrigation is supplemental to summer rainfall, on potatoes and field vegetables. They are the preferred method for irrigation on many high-value crops where high water uniformity is an essential component of production used to minimise variability in crop development (product size, shape, weight, appearance).

The modelling framework consisted of two components, a 'water/irrigation application' and a 'biophysical crop simulation' module. Water application module refers to the combined engineering, hydraulic and management components that are necessary to apply water via an overhead irrigation system. This component is dealt with by an irrigation simulation model; the biophysical crop simulation, development and yield aspects including soil water simulation is undertaken within the AquaCrop plug-in program. A brief description of the two modelling components, their integration and application are outlined below.

4.2.1. Water application model

This component was developed in Microsoft VB.Net and was also designed as a MapWindow plug-in (Ames 2007) to incorporate

spatial variability and mapping functionality. MapWindow is an open-source GIS software product originally developed at Utah State University. The model was designed to simulate dynamically and in real-time the water application of a hose- reel fitted with boom or linear move irrigation system. The model simulates the irrigation boom operating under variable conditions of pressure, wind speed and wind direction. The model can also be used to evaluate boom design (sprinkler spacing, sprinkler height above crop, pipe sizing) to achieve better irrigation uniformity or by irrigators to assess the implications of changing sprinkler type, the wind-in or pulling speed of the hose-reel or pumping pressure on system performance in terms of uniformity and volume of water applied. The boom model can also be used to assess system performance when operated either conventionally (uniform rate of irrigation, URI) or in a precision irrigation mode (variable rate irrigation, VRI). The latter can either be obtained by changing the wind-in speed of the hose-reel or by individually controlling each sprinkler (on/off) on the boom to provide differential wetting patterns. A flowchart summarizing the main components of the boom model is given in Fig. 4.1.

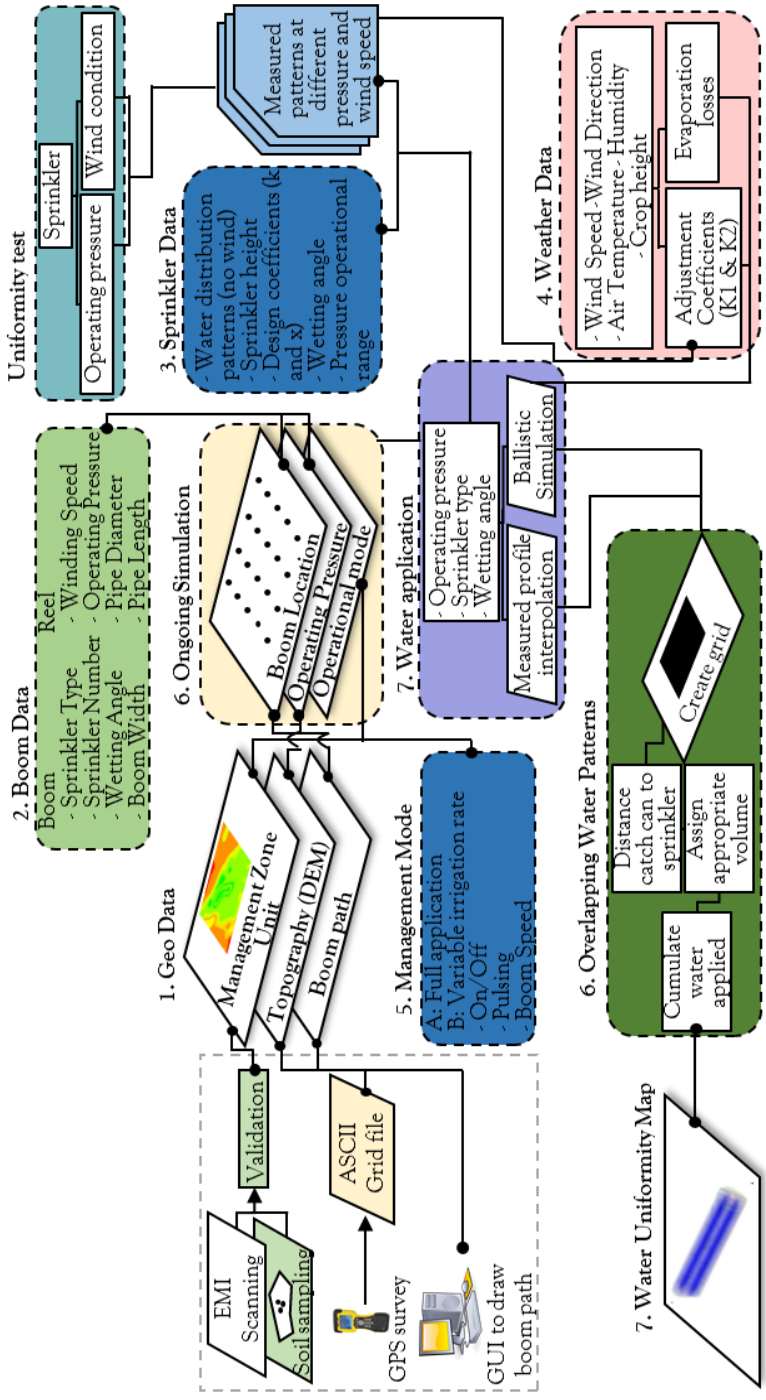


Fig. 4.1. Flowchart of the boom irrigation simulation GIS plugin.

4.2.2. AquaCrop yield model and plug-in program

The AquaCrop model simulates potential yields for herbaceous crops as a function of water consumption under different rainfed and irrigated regimes (Steduto et al. 2012). It directly links crop yield to water use and estimates biomass production from actual crop transpiration through a normalized water productivity parameter, which is the core of the AquaCrop growth engine. A detailed description of the AquaCrop model is reported in Steduto et al. (2012). The AquaCrop input files contain the growth development characteristics of the crop, and the local environment (climate, management practices, soil characteristics) in which the crop is cultivated. The input files are grouped into a 'project' with each project containing up to 11 input files. Input files can be created or modified using the user interface in AquaCrop (Raes et al. 2009). However, when multi-model simulations are planned, the generation of a large number of individual input files is a time-consuming and onerous task. The simulation results are recorded in output (text) files and can be aggregated into 10-day, monthly or annual summary data. The output consists of five files containing data regarding crop growth and production, the soil water balance, soil water content at different depths, and net irrigation requirements.

For a large number of model runs, the FAO has developed the AquaCrop plug-in program, which can perform identical calculation procedures to that in the AquaCrop standard program (Raes et al. 2012) but with the advantage that it facilitates inclusion of the AquaCrop modelling routines within external applications. However, due to an absence of a user interface, only simulation runs (single or multiple) previously defined within the AquaCrop model can be used in the plug-in program (Raes et al. 2012). The plug-in program runs the successive projects in batch mode with the intermediate (daily, 10-daily or monthly) and final (seasonal) simulation results of each project then being saved in an output file. This contains information on the simulation, including climatic and soil water balance parameters, stresses, biomass production, crop yield and water productivity (Raes et al. 2012). Further post processing is then required to analyse the individual output files for each simulation.

4.2.3. Model integration

The model developed in this study was implemented in Matlab (Pratap 2010) to facilitate its inclusion with other computation engines such as an optimization process using multi-objective genetic algorithms. The model consisted of three modules (i) setting (ii) computing and (iii) map building. Fig. 4.2 shows a

flowchart summarising the decision rules embedded within the model.

The 'setting' module is responsible for loading or creating the files necessary for the correct operation of the boom irrigation model. This module consists of two sub-modules termed 'project' and 'field'. Project is responsible for loading or creating the AquaCrop files needed to run the AquaCrop plugin. This sub-module provides the option of both loading an AquaCrop project (*.PRO extension) previously created or building a new AquaCrop project within the Matlab environment. The AquaCrop plugin also requires a climate file (*.CLI) consisting of temperature (*.TMP), reference evapotranspiration (*.ETO), rainfall (*.PLU) and atmospheric CO2 (*.CO2) files. Finally, a crop (*.CRO) file is required by AquaCrop. In addition, there are four optional files: management conditions (*.MAN), groundwater (*.GWT), initial conditions (*.SW0) and off-season conditions (*.OFF). The user has the option to create or modify each of these files or choose default options within the AquaCrop model (Table 4.1). The model also requires information relating to irrigation method, the irrigation schedule, soil depth and electrical conductivity of the soil. Finally, the user defines the crop cycle period (in growing degree days (GDD) or calendar day) and the cropping and simulation period. The AquaCrop project, irrigation and soil file are then automatically saved.

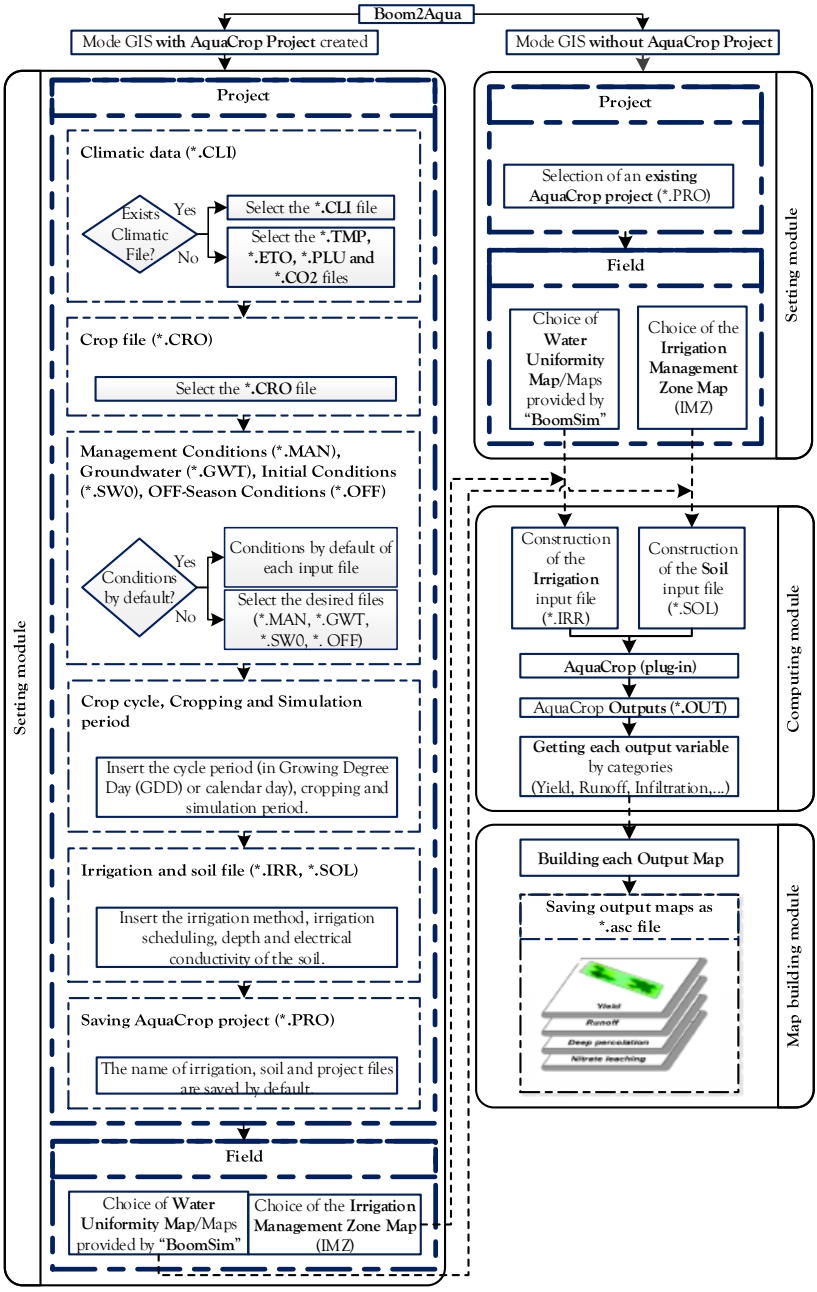


Fig. 4.2. Flowchart showing the decision rules for the boom irrigation simulation model.

Table 4.1. Default options of the management conditions, groundwater, initial conditions and off- season condition files.

File	Extension	Default options
Management conditions	*.MAN	In the absence of a field management file, no specific field management conditions are considered. It is assumed that soil fertility is unlimited, and that field surface practices do not affect soil evaporation or surface run-off.
Groundwater	*.GWT	In the absence of a groundwater file, no shallow groundwater table is assumed when running a simulation.
Initial conditions	*.SW0	In the absence of a file with initial conditions, it is assumed that in the soil profile the soil water content is at field capacity and salts are absent at the start of the simulation.
Off-Season conditions	*.OFF	In the absence of a file with off-season, no mulches and irrigation events are considered before and after growing cycle.

The second sub-module within the 'setting' module is 'Field'. In this sub-module, the user is required to provide maps of water uniformity generated by the WAM and the irrigation management zone map (IMZ map). The water uniformity map determines the irrigation depths applied in each zone and the IMZ map provides information on soil type variability. Thus, it is possible to assess the impact of water distribution patterns for a crop grown in multiple fields across a farm.

The computing module updates the irrigation and soil files with values provided by the water uniformity and IMZ maps, respectively. The AquaCrop plugin program is then launched and

output text files generated (*.OUT). The computing module is run for every pixel, with the pixel grid size determined by the resolution of the water uniformity and IMZ maps. With each iteration (each grid pixel), the output variables are read and stored by the software and then used to generate the output maps by the Map Building module. Once all the AquaCrop simulations (using the plugin program) have been completed, the Map Building module builds each output map. Output maps are saved as *.asc files ready for import and mapping in GIS software. The model provides nine output maps relating to: relative biomass (%), drainage (mm), harvest index (%), infiltration (mm), runoff (mm), transpiration (mm), relative transpiration (%), water productivity (kg m^{-3}) and yield (t ha^{-1}).

4.2.4. Case study model application

To demonstrate the application of the integrated WAM and AquaCrop modelling framework, a case study to assess the impacts of VRI on crop productivity was carried out for a field site in eastern England. Onion was chosen as the representative crop since it is considered to be one of the most important high value field vegetables grown in the UK, with c300,900 tonnes produced from 8,448 ha (DEFRA 2010). It is also highly sensitive to drought stress with irrigation needed to assure both crop yield and quality (Pérez-Ortolá et al. 2014). To calibrate the AquaCrop

model, the crop file (*.CRO) was parameterised using data from Pérez-Ortolá et al. (2014). A typical 'dry' (2010) and 'wet' (2011) year was chosen to assess the impacts of rainfall variability and VRI on crop yield. The annual reference evapotranspiration (ETo) and rainfall were 724 mm and 346 mm for the dry year, and 655 mm and 475 mm for the wet year, respectively. In eastern England, the onion crop is typically grown on light, low moisture retentive sandy loam soils. Most UK vegetable growers use hose-reel irrigation systems fitted with booms. In this study, the boom system had the following design configuration: 7 sprinklers with a sprinkler spacing (2.35 m), individual sprinkler height above the ground (1.35 m), hose-reel length (300 m), pipe diameter (110 mm), mini boom width (16.5 m) and a hose-reel wind-in speed which is a function of the scheduled irrigation depth. It should be noted that a boom with 7 sprinklers is not typical for field scale irrigation but rather a mini boom used in this study for irrigation evaluation and model development. However, the boom parameters were chosen to reflect typical operating settings found in field scale onion cropping in the UK (Perez Ortola 2013).

With the objective of assessing how an intelligent precision irrigation management system could improve water efficiency and productivity (yield), several scenarios were defined and simulated (Table 4.2). The first was an uniform scenario where the entire farm had a sandy loam texture (Fig. 4.3a). Under this scenario, a

uniform rate of irrigation (URI) was defined and scheduled, as might typically be practiced under conventional farming practice.

Table 4.2. Summary characteristics for each precision irrigation scenario.

Scenario	Soil type (% field area)	Irrigation scheduling approach	Proportion of scheduled irrigation applied (%)
1	Sandy loam (100%)	URI	100%
2	Sandy loam (65%) and clay loam (35%)	URI	100%
3	Sandy loam (65%) and clay loam (35%)	VRI- varying the wind in speed of the hosereel	Sandy loam (100%) and clay loam (40%)
4	Sandy loam (65%) and clay loam (35%)	VRI- individual control on each sprinkler	Sandy loam (100%) and clay loam (40%)

Notes: URI, uniform rate irrigation; VRI, variable rate irrigation

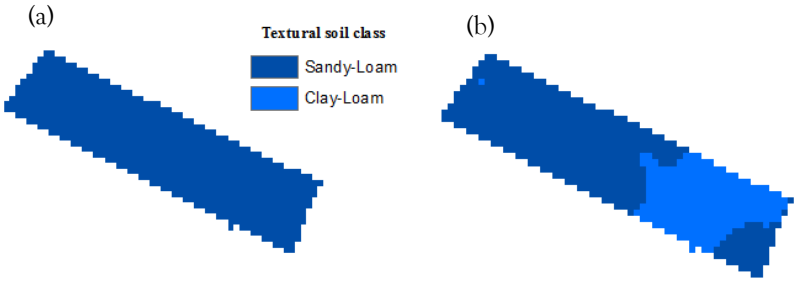


Fig. 4.3. Irrigation management zone (IMZ) maps for a conceptualised uniform farm (a) and a typical farm (b).

According to results from a farm business survey by Perez Ortola (2013), farmers typically irrigate at a soil water deficit (SWD) of

23 mm back to field capacity during canopy development and then allow a slightly larger SWD (29 mm) to accrue during bulb formation. This irrigation schedule was used in Scenario 1 and the resulting water uniformity map is shown in Fig. 4.4a.

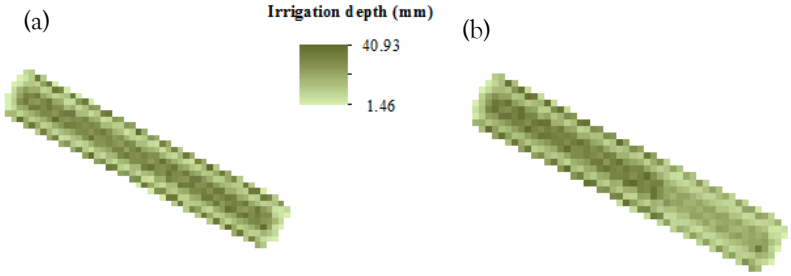


Fig. 4.4. Example water uniformity maps provided by the boom irrigation simulation model for URI (a) and VRI (b) and an average irrigation depth of 23 mm and a working pressure of 25 m (2.5 bar).

Under the second scenario, a typical farm was assumed where the predominant soil was a sandy loam but there were also some zones or areas with clay loam (Fig. 4.3b). This Scenario 2 reflected the situation observed in the case study region. The same irrigation schedule as used in Scenario 1 was used. In Scenarios 3 and 4, a precision irrigation management approach assuming VRI was defined. The VRI was achieved in Scenario 3 and 4 by changing the wind-in speed of the hose-reel and controlling each individual sprinkler on the boom, respectively.

The sprinklers used in the study were pressure compensating Nelson 3000 Rotator series which are widely used on both centre pivots and hose-reel boom systems in the UK and internationally.

For both scenarios, the boom was programmed to apply the full (100 %) irrigation need (23 mm and 29 mm) in the IMZs where there was sandy loam present and only 40 % of the scheduled irrigation in zones where clay loam was present; this was because a clay loam soil is typically able to store 60 % more water than a sandy loam soil. The derived water uniformity maps for each of these scenarios are shown in Fig. 4.4b. Pressure changes in the hose-reel can also have an effect on the depth of irrigation applied since the operating pressure will influence the droplet size, flow rate and hence discharge and wetted distribution pattern. In order to incorporate these pressure effects, the four scenarios were also modelled under three contrasting operating pressure conditions: ideal or perfect conditions (PC, 25 m [0.25 MPa]), high pressure (HP, 40 m [0.40 MPa]) and low pressure (LP, 10m [0.10 MPa]). These pressures were derived from previous experimental research by Knox et al. (2014) where the mini boom and sprinklers were evaluated to assess variations in sprinkler discharge, wetted areas and uniformity under 'no wind' and 'windy' operating conditions, across a range (15 to 40 m) of pressure conditions. A grid pixel resolution of 3 m was used for all scenario simulations.

4.3. Results and Discussion

4.3.1. Irrigation Management Scenarios

Onion yield, infiltration and drainage of irrigation water under the four scenarios described above and for two contrasting agroclimatic cropping seasons (2010 and 2011) were assessed. Box and whisker plots for each are shown in Fig. 4.5, Fig. 4.6 and Fig. 4.7. Each scenario is also evaluated under the three different working pressures (PC, LP and HP). It is also important to put modelled yields in context with typical farm yields. Pérez Ortolá and Knox (2014) reported that a yield of c10 t ha^{-1} dry matter (DM) corresponds to a green yield of c70 t ha^{-1}. In an average year, farmers in East Anglia typically achieve green yields of 50-60 t ha^{-1} (7 to 8.5 t ha^{-1} DM) but these can rise in dry years due to higher temperatures and increased radiation to 60 to 70 t ha^{-1} (8.5 to 10 t ha^{-1} DM). As expected these reported farm yields are lower than modelled yields due to various agronomic (pests/disease), water and nutrient (fertiliser) management factors.

Scenario 1: This scenario reflected uniform conditions on the farm in terms of soil texture and in-field variability. Thus, a URI was applied based on the irrigation schedule derived from the farmer survey (Perez Ortola 2013). The average onion yields in the wet season were 11.05, 11.07 and 11.03 t DM ha^{-1} for the three

working pressures, respectively, and 12.13, 12.14 and 12.13 t DM ha^{-1} for the dry season (Fig. 4.5).

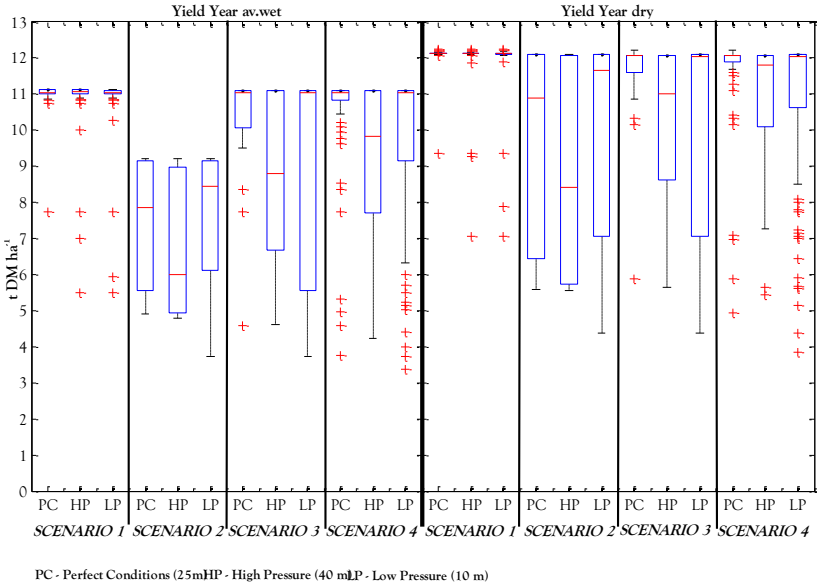


Fig. 4.5. Simulated onion yield (t ha^{-1}) under each scenario and for the two contrasting agroclimate seasons (wet and dry year).

In wet years, as expected, rainfall reduces the scheduled number of irrigation events, but increases the variability in soil moisture. In other words, the farmer has less control over one of the key variables that determines crop yield. In addition to rainfall, potential yield is also a function of other agoclimate conditions during the growing season, notably solar radiation and temperature. Indeed, inspection of the daily climate data and modelled output from the AquaCrop model confirmed that yield differences were also influenced by these parameters reducing the rate of crop development and growth. Excess water in the rooting

zone during wet years also delayed the timing and number of irrigation events and led to higher rates of deep percolation (drainage) which also contributed to increased nutrient (fertiliser) leaching. Thus, during a wet year, yield was reduced and variability increased even when the irrigation schedule and soil variability was optimal.

S*cenario 2*: This scenario reflected the management of a typical onion crop on a farm in the study area, with an irrigation schedule defined for a sandy loam soil. However, on most farms the soil is not uniform but includes parts of fields with different textural characteristics. This creates challenges in defining irrigation schedules for the driest part of a field whilst trying to limit any drainage losses. This scenario was therefore focused on the importance of managing different soil types to reduce both yield variability and the volume of water applied; the objective was thus to reduce drainage losses and increase the effective use of rainfall in the higher water holding capacity soils. For the three operating pressures (PC, HP and LP), the infiltration amounts in this scenario are the highest (Fig. 4.6) with most of the infiltrated water being lost as drainage (Fig. 4.7).

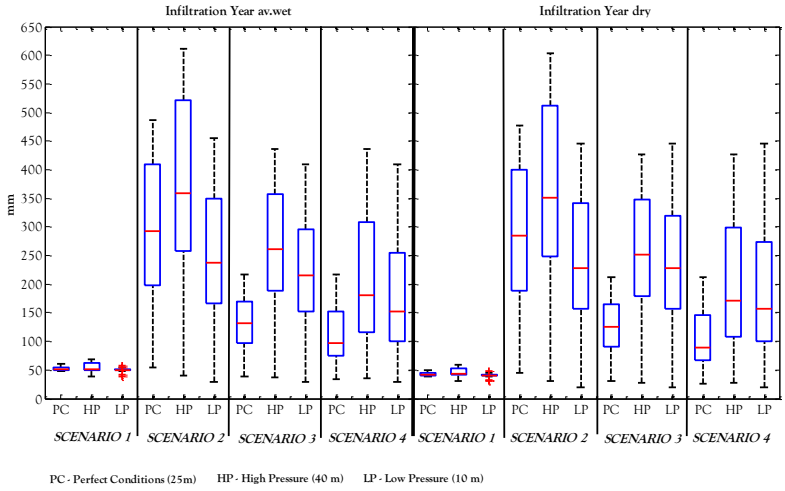


Fig. 4.6. Simulated infiltration from irrigation (mm) under the four scenarios for two contrasting agroclimatic seasons (wet and dry year).

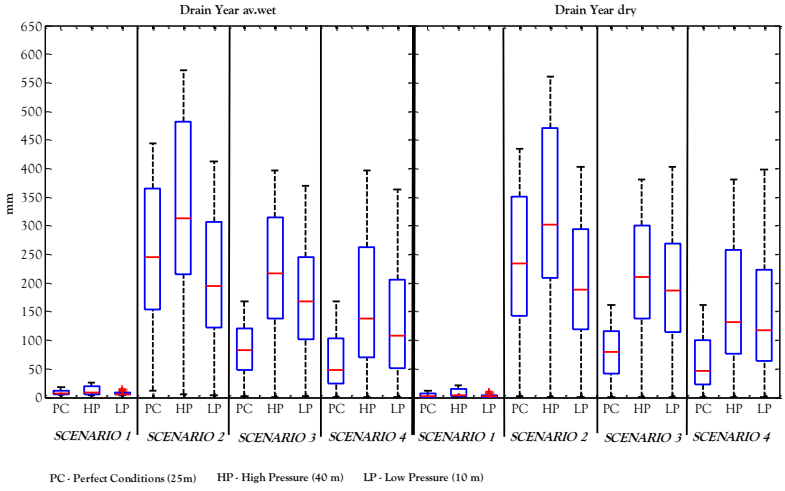


Fig. 4.7. Simulated drainage of irrigation water (mm) under each scenario for the two contrasting agroclimate seasons (wet and dry year).

The operating pressure affects both the volume of water discharged by the sprinkler as well as the droplet size distribution

pattern. Thus, larger water droplets created by a lower operating pressure would affect the sprinkler water distribution pattern and potentially damage the crop canopy and soil structure. Excess (high) operating pressure can be controlled through the use of pressure regulators fitted onto each sprinkler. However, if pressure regulators cannot be used then high pressure leads to a larger volume of water being concentrated around the sprinkler; this in turn leads to greater atomisation of small droplets which are more sensitive to wind drift. Hence any change in the operating (pumping) pressure of the system would affect not only the uniformity of the overlapping patterns but also the amount of water applied (scheduled depth) to the crop. For this reason, under Scenario 2, the infiltration and drainage is highest when the working pressure is high and lowest when working pressure is low (Fig. 4.6 and Fig. 4.7). A considerably lower onion yield compared to Scenario 1 is shown in Fig. 4.5. The average onion yield in the wet season was 7.86, 5.99 and 8.44 t DM ha^{-1} and 10.89, 8.41 and 11.65 t DM ha^{-1} for the dry season, for the three working pressures (PC, LP and HP), respectively. Thus, under Scenario 2, the yield was -28.9 %, -45.9 % and -23.5 % compared to Scenario 1, for the three working pressures in a wet season. In contrast, yield during a dry season was -10.2 %, -30.7 % and -4.0 % compared to Scenario 1. In wet years, rainfall reduces the scheduled number of irrigation events and buffers the irrigation

schedule for soils that are different to the scheduled sandy loam. Although the average yield is higher in a dry year compared to the wet year there is also much greater yield variability; this is largely due to the inappropriate irrigation schedule for field areas (35 %) that were assumed to be a clay loam in contrast to the 65 % area that was scheduled assuming a sandy loam soil. Under this scenario, yield variability is much higher compared to Scenario 1; in practice, this yield variability would also likely lead to greater variations in crop quality, which is an important determinant of crop price received by a farmer for quality assurance (Rey et al. 2016) particularly in high value crops such as onions and potatoes.

Scenarios 3 and 4: Under Scenario 3 and 4, the impacts of variable rate irrigation (VRI) implementation are modelled to take into account the spatial variability in soil type across the farm. This approach results in lower application depths being scheduled and applied in areas of the field where the soil has a higher available water holding capacity. The irrigation depth is thus lower to avoid runoff and drainage and increase efficiency of water use. In this study, the irrigation model was used to simulate VRI in two different ways. Firstly, VRI was achieved changing the wind-in speed of the hose-reel which varies the depth along the travel lane (Scenario 3) and secondly, by controlling each individual sprinkler along on the boom which varies the application depth

across the transect (Scenario 4). Since irrigation uniformity is achieved by overlapping the wetted patterns from adjacent sprinklers, the variable application with a boom system is constrained to a minimum spatial scale by the throw of the individual sprinklers. Under Scenario 4, the hose reel requires a controller to maintain a constant pull-in speed independently of the variable flow. Under current design, a minimum constant flow is needed to drive the hose-reel turbine needed to pull in the boom.

Infiltration was reduced in both scenarios relative to Scenario 2 (typical irrigation management) but drainage was also reduced (Fig. 4.6 and Fig. 4.7). Thus, the onion crop had a higher available water content in the root zone which contributed to the increase in final yield (Fig. 4.5).

The average onion yields in the wet season under Scenario 3 were 11.03, 8.80 and 11.03 t DM ha^{-1} for the three working pressures, and 12.07, 11.01 and 12.04 t DM ha^{-1} for the dry season, respectively (Fig. 4.5). These values are very close to those for Scenario 1 (uniform management) although the variability was markedly increased. A reduction in crop quality and hence price in the final product results when the yield variability increases. The average onion yield in the wet season under Scenario 4 was marginally higher than under Scenario 3, corresponding to 11.03,

9.84 and 11.03 t DM ha^{-1} for the three working pressures (wet season) and 12.07, 11.80 and 12.04 t DM ha^{-1} (dry season), respectively (Fig. 4.5). As in Scenario 1, during the dry season, Scenarios 3 and 4 achieved better yields than the wet season. When irrigation scheduling is close to optimal, the dry seasons achieved a higher onion yield due to improved control over the water content in the root zone. The results under these two scenarios show that onion yield values were similar but the variabilities in yield as well as infiltration and drainage were slightly higher under Scenario 3. Therefore, the most suitable way to implement VRI appears to be through individual control on each sprinkler along the boom, but this introduces a set of new hydraulic challenges. Not only it is more expensive because it is necessary to use individual remote control solenoid valves on each sprinkler, but the independent switching on/off sprinklers introduces a confounding problem with uniformity – sprinklers on a boom are designed to be operated simultaneously in order to generate the required overlapped pattern to maximise uniformity; however, by switching individual sprinklers on/off, the overlapping pattern is disturbed with consequent impacts on uniformity.

4.3.2. Methodological limitations

The approach developed has a number of methodological limitations which need to be recognised. These challenges include issues such as the availability of relevant geodata, developing a graphical user interface (GUI), facilitating its use for farmers and integrating these approaches with current modelling developments in precision agriculture and decision support systems. There is also a need to simulate each scenario under windy conditions. For all scenarios modelled here, there were 'no-wind' conditions, and hence no distortions in wetted pattern due to wind drift. For end users, there is also a need for careful documentation of modelling approaches and particularly how datasets are pre-processed prior to model input, and then how derived datasets are passed between individual models. Great care has to be taken when linking models, as errors in one are often propagated and may become exacerbated or attenuated through model integration. There is hence a risk of introducing additional modelling uncertainty, particularly where datasets of different provenance, scale and integrity are integrated. An uncertainty matrix could be used to identify sources of uncertainty both within the irrigation ballistics and crop modelling components, and then used to inform the interpretation of the crop modelling outputs.

4.4. Conclusions

The integrated modelling approach developed allows assessment of the spatial and temporal impacts of irrigation heterogeneity under conventional and precision irrigation management strategies on crop yield and soil water management at field scale. The development of this model for the automated multi-model operation of AquaCrop significantly improves its utility to simulate yield for numerous locations and conditions or for other applications that require this tool to be embedded into other computation engines. The case study results showed that VRI has potential to be a useful way in achieving water savings at the farm-scale due to reductions in infiltration and drainage. As a consequence, the final yield increased in the variable field because of higher water content in the root zone. Conversely, the results showed that the use of VRI in a dry season could improve crop yield due to improved control of water content in the root zone. Finally, the results also showed that the best way to apply VRI is by individually controlling each sprinkler on the boom although it is also more expensive due to the need for individually actuated (solenoid valves) on each sprinkler. It should be recognised that implementation of PI technologies and management approaches needs to be site and crop specific. PI approaches cannot be generalised across different farming systems and crop mixes,

highlighting the need for an integrated tool to assess potential benefits and trade-offs.

The approach described here provides a basis for evaluating the agronomic and economic impacts of PI implementation in other cop sectors to understand the impacts of irrigation heterogeneity on yield, but also more importantly on crop quality, and to identify strategies that can be used to reduce 'non-beneficial' water losses, to improve water and energy efficiency, and to reduce the environmental impacts associated with supplemental irrigation. Integrating biophysical and engineering models to advance our knowledge of these interactions will go some way to addressing these knowledge gaps.

Acknowledgments. The authors acknowledge Defra for funding this research through the Hortlink programme. This work forms part of HL0196. This research was supported by an FPU grant (Formación de Profesorado Universitario) from the Spanish Ministry of Education, Culture and Sports to Rafael González Perea. This work is also part of the TEMAER project (AGL2014-59747-C2-2-R), funded by the Spanish Ministry of Economy and Competitiveness.

4.5. References

Ames D (2007) Mapwin GIS reference manual. Idaho State University, Idaho Falls

Casadesús J, Mata M, Marsal J, Girona J (2012) A general algorithm for automated scheduling of drip irrigation in tree crops. Comput Electron Agric 83:11–20

Daccache A, Ciurana JS, Rodriguez Diaz JA, Knox JW (2014a) Water and energy footprint of irrigated agriculture in the Mediterranean region. Environmental Research Letters, 435 9(12), 124014. http://www.scopus.com/inward/record.url?eid=2-s-2.084919782106&partnerID=tZOtx3y1. Accessed 18 June 2015

Daccache A, Knox JW, Weatherhead EK, et al (2014b) Implementing precision irrigation in a humid climate - Recent experiences and on-going challenges. Agric Water Manag 147:135–143

De Paz JM, Albert C, Visconti F, et al (2015) A new methodology to assess the maximum irrigation rates at catchment scale using geostatistics and GIS. Precis Agric 16:505–531

DEFRA (2010) Agriculture in the United Kingdom 2010. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/208436/auk-2012-25jun13.pdf. Accessed 17 Mar 2014

Falloon P, Betts R (2010) Climate impacts on European agriculture and water management in the context of adaptation and mitigation–the importance of an integrated approach. Sci Total Environ 408:5667–87

FAO (2012) FAO Crop Model to Simulate Response to Water. Natural Resources and Environment Department. Rome, Italy: FAO. http://www.fao.org/nr/water/aquacrop.html. Date accessed 17 March 2016

Fereres E, Orgaz F, Gonzalez-Dugo V (2011) Reflections on food security under water scarcity. J Exp Bot 62:4079–86

Fraisse CW, Breuer NE, Zierden D, et al (2006) AgClimate: A climate forecast information system for agricultural risk management in the southeastern USA. Comput Electron Agric 53:13–27

García Morillo J, Martín M, Camacho E, et al (2015) Toward precision irrigation for intensive strawberry cultivation. Agric Water Manag 151:43–51

González Perea R, Fernández García I, Martin Arroyo M, et al (2016) Multiplatform application for precision irrigation scheduling in strawberries. Agric Water Manag 183:194–201

Haghverdi A, Leib BG, Washington-Allen RA, et al (2016) Studying uniform and variable rate center pivot irrigation

strategies with the aid of site-specific water production functions. Comput Electron Agric 123:327-340

Knox JW, Daccache A, Hess TM, et al (2014) Developing an intelligent overhead irrigation system for high quality horticultural field crops HL0196 Final Report September 2014 Table of Contents. 1–106

Knox JW, Kay MG, Weatherhead EK (2012) Water regulation, crop production, and agricultural water management—Understanding farmer perspectives on irrigation efficiency. Agric Water Manag 108:3–8

Kumar A, Sharma P, Joshi S (2016) Assessing the impacts of climate change on land productivity in indian crop agriculture: An evidence from panel data analysis. 18:1–13

Lorite IJ, García-Vila M, Santos C, et al (2013) AquaData and AquaGIS: Two computer utilities for temporal and spatial simulations of water-limited yield with AquaCrop. Comput Electron Agric 96:227–237

Mitchell JP, Shrestha A, Hollingsworth J, et al (2016) Precision overhead irrigation is suitable for several Central Valley crops. Calif Agric 70(2)

Monaghan JM, Daccache A, Vickers LH, et al (2013) More "crop per drop": constraints and opportunities for precision

irrigation in European agriculture. J Sci Food Agric 93:977–80

Pérez-Ortolá M, Daccache A, Hess TM, Knox JW (2014) Simulating impacts of irrigation heterogeneity on onion (Allium cepa L.) yield in a humid climate. Irrig Sci 33:1–14

Perez Ortola M (2013) Modelling the impacts of in-field soil and irrigation variability on onion yield. Cranfield University, UK

Pérez Ortolá M, Knox JW (2014) Water Relations and Irrigation Requirements of Onion (Allium Cepa L.): a Review of Yield and Quality Impacts. Exp Agric 51:210–231

Pratap R (2010) Getting started with Matlab. A quick introduction for scientist and engineers. Oxford University Press, USA

Raes D, Steduto P, Hsiao TC, Fereres E (2012) Reference Manual: AquaCrop Plug-in Program (Version 4.0). FAO, Rome, Italy 101(3):438

Raes D, Steduto P, Hsiao TC, Fereres E (2009) AquaCropThe FAO Crop Model to Simulate Yield Response to Water: II. Main Algorithms and Software Description. Agron J 101:438

Rey D, Holman IP, Daccache A, et al (2016) Modelling and mapping the economic value of supplemental irrigation in a humid climate. Agric Water Manag 173:13–22

Siebert S, Burke J, Faures JM, et al (2010) Groundwater use for irrigation – a global inventory. Hydrol Earth Syst Sci 14:1863–1880

Smith R, Baillie J (2009) Defining precision irrigation: a new approach to irrigation management. Swan Hill, Autralia: In: Irrigation Australia 2009

Steduto P, Hsiao TC, Fereres E, Raes D (2012) Crop yield response to water. FAO Irrigation and Drainage Paper No.66. Rome, Italy

Thorp KR, DeJonge KC, Kaleita AL, et al (2008) Methodology for the use of DSSAT models for precision agriculture decision support. Comput Electron Agric 64:276–285

5. Optimization of Irrigation Scheduling Using Soil Water Balance and Genetic Algorithms

This chapter has been published entirely in the journal "Water Resources Management", González Perea R, Camacho Poyato E, Montesinos P, Rodríguez Díaz JA (2016)

Abstract. In arid and semi-arid countries, the use of irrigation is essential to ensure agricultural production. Irrigation water use is expected to increase in the near future due to several factors such as the growing demand of food and biofuel under a probable climate change scenario. For this reason, the improvement of irrigation water use efficiency has been one of the main drivers of the upgrading process of irrigation systems in countries like Spain, where irrigation water use is around 70 % of its total water use. Pressurized networks have replaced the obsolete open-channel distribution systems and on farm irrigation systems have been also upgraded incorporating more efficient water emitters like drippers or sprinklers. Although pressurized networks have significant energy requirements, increasing operational costs. In these circumstances farmers may be unable to afford such expense if their production is devoted to low value crops. Thus, in this

work, a new approach of sustainable management of pressurized irrigation networks has been developed using multiobjective genetic algorithms. The model establishes the optimal sectoring operation during the irrigation season that maximize farmer's profit and minimize energy cost at the pumping station whilst satisfying water demand of crops at hydrant level taking into account the soil water balance at farm scale. This methodology has been applied to a real irrigation network in Southern Spain. The results show that it is possible to reduce energy cost and improve water use efficiency simultaneously by a comprehensive irrigation management leading, in the studied case, to energy cost savings close to 15 % without significant reduction of crop yield.

Keywords. Pressurized irrigation network, irrigation district management, soil water balance, genetic algorithm

5.1. Introduction

In arid and semi-arid countries, the use of irrigation water is essential to ensure agricultural production. Globally, irrigated agriculture is the primary user of freshwater, accounting for nearly 85 % of the total water consumption (Jury and Vaux 2007), and provides about 40 % of total food production (Fereres and Soriano 2007). Irrigation water demand is expected to increase in the near future due to foreseen alterations of rainfall regime caused by climate change (Rodríguez Díaz et al. 2007), and

increased food and biofuel demands. Irrigation schedule is part of the complex problem of optimal water resources management (English et al. 2002).

On farm irrigation systems have been upgraded with the aim of increasing irrigation efficiency incorporating more efficient water emitters like drippers or sprinklers (Playán and Mateos 2006). Consequently, pressurized networks have replaced the obsolete open-channel distribution systems (Plusquellec 2009). These changes increase the conveyance efficiency reducing water losses throughout the system. However, these pressurized networks have significant energy requirements, which may lead to additional costs for farmers that may be unable to afford such expense if their production is devoted to low-value crops. For this reason, several management strategies have been developed to reduce energy consumption in pressurized irrigation networks (Abadia et al. 2008; Daccache et al. 2010; Lamaddalena and Khila 2012). One of the most efficiency measures is network's sectoring, where farmers are organized in irrigation turns according to their energy demand. Previous works have shown that network sectoring can achieve potential energy savings between 20 % and 30 % (Rodríguez Díaz et al. 2009; Carrillo Cobo et al. 2011; Navarro Navajas et al. 2012). Another energy saving measure is the control of critical points, which are hydrants with high energy requirements. Rodríguez Díaz et al. (2012) developed the WECP

(Water and Energy optimization by Critical Point control) algorithm for detecting critical points in pressurized irrigation networks. It was applied in two pressurized irrigation networks in Southern Spain. The results showed that potential energy savings around 10 % and 30 % were possible in each network whilst satisfying the theoretical irrigation requirements. However most of these energy saving measures focus on reducing the energy demand without considering the irrigation scheduling at farm level.

When several objectives are considered in the operation of water networks, a more realistic approach of the problem is achieved and the decision-making process is significantly improved, as a wide range of alternatives are available. Heuristic approaches are useful when solving this sort of problems. Among heuristic techniques, the NSGA-II algorithm (Non dominated Sorting Genetic Algorithm) (Deb et al. 2002) has been successfully used to solve multiobjective problems related to design or management of water distribution networks (*e.g.* Siew and Tanyimboh 2012; Fernández García et al. 2013). Hence, this algorithm has been selected as optimization tool in this work.

There are several works in which irrigation scheduling was optimized using genetic algorithms with the aim of reducing drainage losses (Wardlaw and Bhaktikul 2004), or maximizing the

total farm income according to the operation rule of a reservoir (Sadati et al. 2014). We have no knowledge by far of any study focused on the simultaneous optimizations of irrigation scheduling at farm scale, water demand at hydrant level and energy consumption at the pumping station.

In this paper, a new approach of sustainable management of pressurized irrigation networks using a customized version of NSGA-II (Deb et al. 2002) has been developed. The model establishes the optimal sectoring operation during the irrigation season that maximize farmer's profit and minimize energy cost at the pumping station whilst satisfying water demand of crops at hydrant level taking into account the soil water balance at farm scale. This methodology has been applied to a real irrigation network in Southern Spain.

5.2. Methodology

5.2.1. Study area

TheM. D. Bembézar Irrigation District (MDB) is located in Andalusía (Southern Spain) with a total irrigated area of 11,950 ha. The climate is typically Mediterranean, with annual average rainfall of 604 mm concentrated in autumn and spring, and dry spells in summer. The average temperature in the area is 17.7 °C, being July the hottest month (mean temperature 36.2 °C). Under

these circumstances the average reference evapotranspiration is over 1,200 mm.

MDB was established in 1967. Initially the conveyance system was an open channel network that covered over 11,900 ha. In 2007, the hydraulic infrastructures were upgraded and the old network was replaced by a pressurized system arranged on-demand, so water is continuously available to farmers. The water is conveyed from three reservoirs (Bembézar, 342 Mm^3; Retortillo, 61 Mm^3 and José Torán, 101 Mm^3) through a main channel of 40 km length and 12 $m^3\ s^{-1}$ of delivery capacity. Then, eleven pumping stations operate along the main channel to supply water to each irrigation sector. The network was designed to supply 1.2 $L\ s^{-1}\ ha^{-1}$ on-demand with a service pressure at hydrants of 35 m. Drip irrigation is the most common irrigation system. The methodology described in this paper is applied to one of the eleven irrigation sectors, Sector VII, that irrigates 935 ha by 161 hydrants operate on demand. Its crop pattern is representative of the whole irrigation district: maize (38.00 %), Citrus trees (34.00 %), cotton (11.79 %), wheat (9.71 %), fruit trees (3.74 %), watermelon (1.46 %) and sunflower (1.30 %).

5.2.2. Problem Approach

This work was aimed at determining the optimal operation of the whole irrigation district from source to crop during the irrigation

season. The crop is watered using trickle systems. The network operation depends both on soil water content that determines the beginning of the irrigation ($s_{irrigation}$) at plot level and the network sectoring. The number of sectors is fixed according to the minimum number of sectors that allow every hydrant to satisfy the irrigation needs during the peak demand day of the season. These irrigation needs were calculated according to FAO (2009). A sector number was randomly assigned to each hydrant and was subsequently optimized during the optimization process. To achieve this aim two conflicting objectives were considered: the maximum farmers' profit and the minimal energy cost at the pumping station and minimal percolation losses.

The first objective function (F1) maximizes the total value of agricultural production value:

$$F1 = \left[\sum_{h=1}^{h_T} \frac{Y_{r,h} \cdot Y_{max,h} \cdot A_h \cdot Pr_h}{A_T} \right]_{norm} \quad [5.1]$$

where h_T, is the number of hydrants, $Y_{r,h}$ is the relation between yield under actual water stress conditions for the crop in the plot supplied by hydrant h and its maximal potential yield; $Y_{max,h}$ (kg ha^{-1}) is maximum potential yield for the crop irrigated by hydrant h when there are not limitations of water; A_h (ha) is the irrigated area supplied by hydrant h; Pr_h (€ kg^{-1}) is the average market price of the crop irrigated by h during the crop season and A_T (ha) the

total irrigated area by all hydrants (h_T). In order to compare the two objective functions, F1 was normalizes using a cumulative distribution function for the continuous uniform distribution on the interval [0, 2].

$Y_{r,h}$ was estimated with the following production curve included in the FAO 33 report (Doorenbos and Kassam 1979):

$$\left(1 - Y_{r,h}\right) = k_y \cdot \left(1 - \frac{ET_h}{ET_{max,h}}\right) \qquad [5.2]$$

where k_y is the yield response factor, ET_h (mm day^{-1}) the actual evapotranspiration for the crop irrigated by hydrant h and $ET_{max,h}$ (mm day^{-1}) the evapotranspiration without water stress conditions for the crop irrigated by hydrant h. The $ET_{max,h}$ was obtained by:

$$\mathrm{ET_{max,h}} = \sum_{\mathrm{d=1}}^{\mathrm{d_T}} \mathrm{ET_{max\ d,h}} = \sum_{\mathrm{d=1}}^{\mathrm{d_T}} \mathrm{K_{c\ d,h}} \cdot \mathrm{ET_{0,d}} \qquad [5.3]$$

where $ET_{max\ d,h}$ (mm day^{-1}) is the evapotranspiration without water stress conditions for the crop supplied by hydrant h in each of day, d, of crop growing season; $K_{c\ d,h}$ is the crop coefficient of the crop associated to hydrant h for day d of the crop development and $ET_{0,d}$ (mm day^{-1}) reference crop evapotranspiration of day d (Allen et al. 1998).

According to Laio et al. (2001), ET_h was calculated as follows:

$$ET_h = \sum_{d=1}^{d_T} \begin{cases} ET_w \cdot \dfrac{s_{d,h} - s_{hg}}{s_w - s_{hg}} & , s_{hg} < s_{d,h} \leq s_w \\ ET_w + \left(ET_{max\ d,h} - ET_w\right) \cdot \dfrac{s_{d,h} - s_w}{s^* - s_w} & , s_w < s_{d,h} \leq s^* \\ ET_{max\ d,h} & , s^* < s_{d,h} \leq 1 \end{cases} \quad [5.4]$$

where ET_w (mm day^{-1}) is the evapotranspiration in so-called wilting point; $s_{d,h}$ relative soil moisture at the plot associated to the hydrant h on the day d; s_{hg} relative soil moisture in so-called hygroscopic point; s_w relative soil moisture in the wilting point; s^* relative soil moisture from which the crops start to reduce transpiration.

The second objective function (F2) minimizes simultaneously the seasonal energy cost and the water losses by deep infiltration.

$$F2 = \sum_{h=1}^{h_T} \left[\left(E_{T,h}\right)_{norm} + \left(DI_h\right)_{norm} \right] \quad [5.5]$$

where $(E_{T,h})$norm (€) is the normalized seasonal energy cost in the pumping station corresponding to hydrant h and $(DI_h)norm$ (mm) is the normalized deep percolation losses corresponding to the plot irrigation system associated to hydrant h. Both $E_{T,h}$ and DI_h were normalized using a cumulative distribution function for a continuous uniform distribution on the interval [0, 1] to allow their summation. Therefore, F2 varies between 0 and 2.

$E_{T,h}$ were estimated as follows:

$$E_{T,h} = \sum_{d=1}^{d_T} \frac{\gamma_w \cdot F_{d,h} \cdot H_{d,h} \cdot t_{d,h}}{\eta} \cdot UC_e \qquad [5.6]$$

where γ_w (9.8 kN m^{-3}) is the water specific weight; $F_{d,h}$ (m^3 s^{-1}) the demanded flow in the pumping station corresponding to the hydrant h on day d; $H_{d,h}$ (m) the pressure head required at the pumping station to operate the irrigation sector of the hydrant h on day d; $t_{d,h}$ (hours) irrigation time of hydrant h on day d; η is the pumping system efficiency (in this work a pumping efficiency of 0.75 was assumed); d_T is the operation days of hydrant h and UC_e is the Unit energy cost (€ kWh^{-1}). Finally, $t_{d,h}$ was calculated by dividing the irrigation needs during the peak demand day by the design flow of the hydrant.

Daily $H_{d,h}$ values were obtained for each day of the irrigation season to satisfy a minimum service pressure in all hydrants in every sector using the hydraulic simulator EPANET (Rossman 2000). $F_{d,h}$ was also determined by EPANET according to hydrants that make up each irrigation sector.

According to (Laio et al. 2001), the percolation losses, DI_h were calculated as follows:

$$DI_h = \sum_{d=1}^{d_T} \left[\frac{K_{sat}}{e^{\beta \cdot (1 - s_{fc})}} \cdot \left[e^{\beta \cdot (s_{d,h} - s_{fc})} - 1 \right] \right], s_{fc} < s_{d,h} \leq 1 \qquad [5.7]$$

where K_{sat} (mm day^{-1}) is the average saturated hydraulic conductivity of the soil in the study area; β is coefficient which is used to fit the above expression to the power law and s_{fc} relative soil moisture at field capacity.

Energy cost depends on three factors: the unit energy cost, the amount of water applied and the pressure head at the pumping station. The amount of water applied is related to the soil moisture at the beginning of the irrigation period. Initial low soil water contents entail larger amounts of irrigation volume that may impact negatively on farmer's profits although entails lower deep percolation leakages and the reduction of pollutant entering into groundwater. Additionally, pressure head at the pumping stations is linked to the irrigation network sectoring. Then, factors will be considered to find optimum balance between the two objective functions.

5.2.3. Soil Water Balance

The daily value of soil water content affects ET_h in Eqs. 5.4 and 5.2. For that reason, the daily soil water balance for each hydrant in the irrigation network was calculated. Assuming negligible lateral soil moisture fluxes, the relative soil moisture in the plot associated to hydrant h on the day d was calculated by the following equation, $s_{d,h}$, assuming average climatic and soil values for the whole irrigated area:

$$s_{d,h} = s_{d-1,h} + \frac{ER_d + I_{d,h} - ET_{d,h} - DI_{d,h}}{ns \cdot Z_{r\,d,h}} \qquad [5.8]$$

where $s_{d-1,h}$ is the relative soil moisture in the plot associated to hydrant h on the day d-1; ER_d (mm) is the effective rainfall of the day d calculated by the USDA (Allan 1998); $I_{d,h}$ (mm) is the fixed applied irrigation depth to the crop associated to hydrant h on the day d; ns is soil porosity and $Z_{r\,d,h}$ (mm) the active soil depth (where most of crop roots associated to hydrant h on the day d are located).

According to Laio et al. (2001) $s_{d-1,h}$ was estimated as follows:

$$s_{d-1,h} = \frac{\theta_{d-1,h} - \theta_w}{ns} \qquad [5.9]$$

where $\theta_{d-1,h}$ (cm^3 cm^{-3}) is the volumetric soil moisture corresponding to the plot associated to hydrant h on the day d and θ_w (cm^3 cm^{-3}) is the volumetric soil moisture at wilting point.

5.2.4. Optimization Method. NSGA-II

The multi-objective algorithm NSGA-II (Deb et al. 2002) was implemented in MATLAB (Pratap 2010) to obtain the set of $s_{irrigation}$ value in each hydrant and the optimal network's sectoring according to their accomplishment of the objective functions stated above (Eqs. 5.1 and 5.5). The original NSGA-II algorithm was adapted to this problem (Fig. 5.1). Initially, a starting

population of *nPop* chromosomes ($s_{irrigation}$ values (soil Water Content Pattern, WCP) and network's sectoring, NS) was randomly generated. Each chromosome consisted of genes that represented the decision variables (*nDec*) of the problem: h_t values of $s_{irrigation}$ for each plot and h_t sector indexes linked to each hydrant. Therefore, *nDec* was $2 \cdot h_t$. The $s_{irrigation}$ value determines the beginning of irrigation at each plot. The soil moisture ranged from 10 % above wilting point to 90 % of the relative soil moisture in field capacity. The sector index pointed out the operating sector of each hydrant, which varies between 1 and the number of sectors in the network. The most energy-demanding hydrant in each sector determines the energy consumption of the pumping station. Real-coding was the coding procedure used to represent the decision variables (Elferchichi et al. 2009).

Hereafter chromosomes were modified using crossover and mutation operators to obtain successive generations of *nPop* improved chromosomes according to their objective function values. After several generations (*nGEN*) a set of *nPop* optimal chromosomes are obtained. These chromosomes define as the Pareto Front.

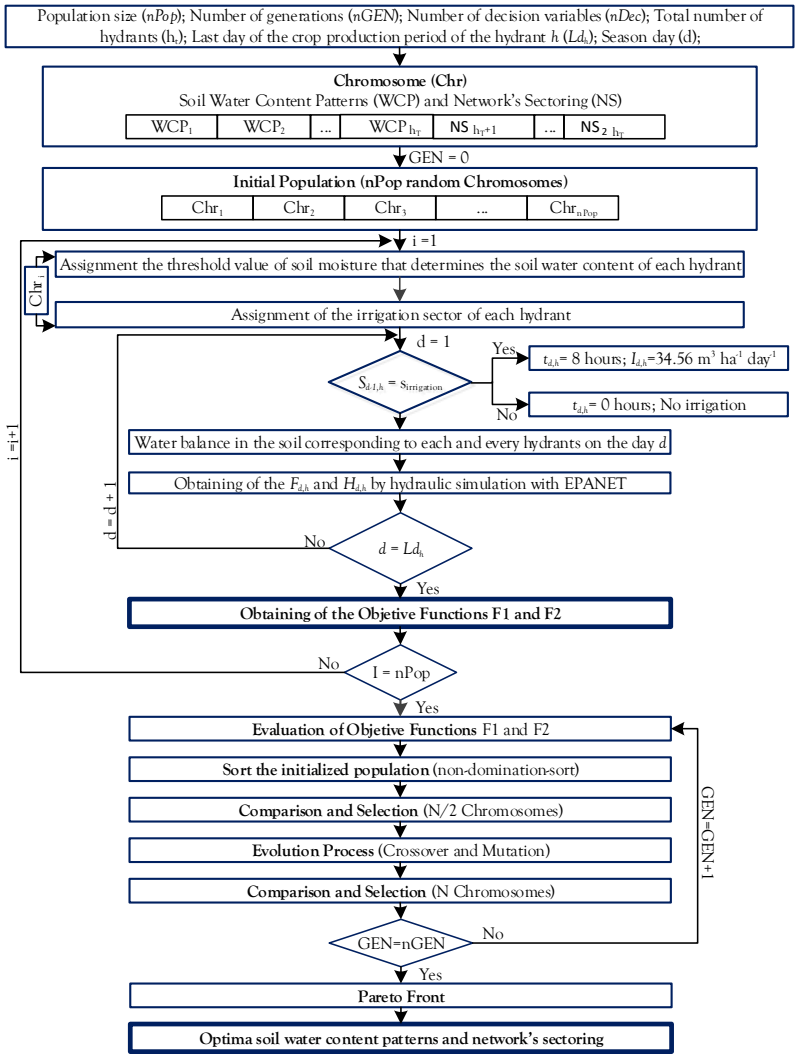


Fig. 5.1. Optimization process using NSGA-II.

5.3. Results and Discussion

The irrigation time required to satisfy average crop water needs on the peak water demand day was 7.84 h in Sector VII of BMD. Therefore, hydrants should be irrigated during 8 h turns. Thus,

hydrants were grouped in three irrigation sectors and the sector index associated to the genes responsible for the network's sectoring could only take integer values between 1 and 3. As the irrigation network was designed to supply 1.2 L s^{-1} ha^{-1} per hydrant, the applied irrigation depth ($I_{d,h}$) per hydrant was 34.56 m^3 ha^{-1} day^{-1}.

The methodology described above has been applied to find the optimal operation of Sector VII during 2009, year in which all the required data were available. The daily values of precipitation and reference evapotranspiration were obtained from the closest public weather station for this season. In this work, the evapotranspiration at wilting point (ET_w) was considered zero.

The crop coefficients K_c and the active soil depth $Z_{r\,d,h}$ values for each stage of the crop production cycle published in FAO 56 (Allen et al. 1998) were considered in this work. The average soil texture of the irrigation district was clay loam. Its hydraulic parameters, estimated according to the ROSETTA model (Schaap et al. 2001) and (Laio et al. 2001), are shown in Table 5.1. Potential yields (Y_{max}) and average market price (Pr) of crops cultivated in the irrigation district were obtained from the annual statistics of the Agriculture Department of Andalucía (Spain) (Consejería de Agricultura 2009) (Table 5.1).

Table 5.1. Values of yield, Y_{max}, crop price Pr, the number of hydrants associated with each crop, cropped area for the main crops and soil properties in the study area in BMD (Sector VII) for the 2009 season.

	Maize	Citrus trees	Cotton	Wheat	Fruit trees	Watermelon		Sunflower
Y_{max} (kg ha^{-1})	10,348	19,894	1,320	4,923	19,915	48,750		2,159
Pr (€ ha^{-1})	0.148	0.241	0.239	0.157	0.737	0.150		0.273
Number of hydrants	55	64	18	13	8	2		1
Cropped area (ha)	355.3	317.9	110.2	90.8	35.0	13.7		12.1
K_{sat} (mm day^{-1})	**s_{hg}**	**s_w**	**s_{fc}**	**s***	**θ_w (cm^3 cm^{-3})**	**n**	**β**	**Soil textural class**
150	0.35	0.38	0.83	0.68	0.20	0.48	20.8	Clay loam

5.3.1. Evolution of the objective functions in the optimization process

The model based on NSGA-II described in Fig. 5.1 was applied to MDB (Sector VII). The objective functions, F1 and F2, were optimized simultaneously. The total number of hydrants in the irrigation network (Sector VII, BMD) was 161. In Table 5.1 are shown the number of hydrants associated to each crop and their cropped area.

The random initial population consisted of 50 chromosomes comprised of 161 genes with $s_{irrigation}$ values and 161 genes with the sector index associated to each hydrant. The initial population evolved for 3,800 generations with probabilities for crossover and mutation of 90 % and 10 %, respectively.

Fig. 5.2 shows the Pareto front obtained in the optimization process (generation 3,800). This graph shows that both objectives (F1 and F2) are conflicting because improvements in one of them imply worsening the other (high F1 and low F2 values are desirable).

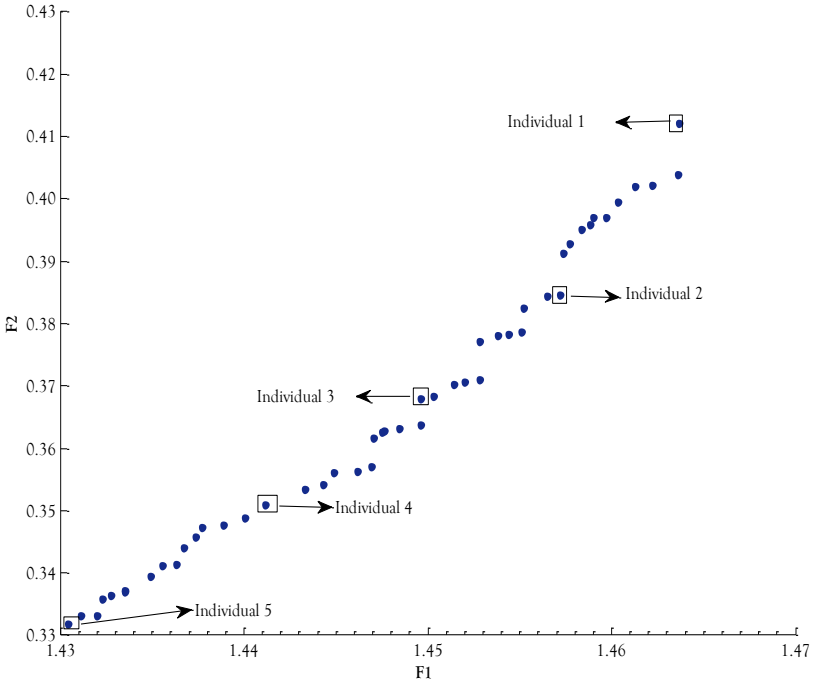


Fig. 5.2. Pareto Front in generation 3,800.

The total agricultural production value, total energy cost and total losses due to deep percolation for Sector VII, were obtained from F1 and F2 for every chromosome. The target values of these terms were: 1,878.60 € ha^{-1} for the value of the total production when there were not water availability restrictions so crops could reach their maximum potential yield ($Y_h = Y_{max,\,h}$) and zero for energy

cost and deep percolation losses (although these values were not realistic because they only could occur in rainfed conditions). Fig. 5.3a shows the evolution of the maximum values of production in Sector VII-MDB during 3,800 generations. The production value increased in the first 1,200 generations and slightly improved until generation number 2,914 when F1 was stabilized. Finally, after 3,800 generations, the production of Sector VII-MDB increased up to 1,375.14 € ha^{-1}. However, this solution may not be the best for the Irrigation District as entails increases of the total energy cost and deep percolation losses.

The minimum values of total energy cost and deep percolation losses in each generation are shown in Fig. 5.3b and 5.3c, respectively. After 3,289 generations, the total energy cost at the pumping station was stabilized. The total energy cost in this generation was 53,928 €. Although, this solution may not be the best, as it would imply lower crop yields. Thus, comparing the extreme values obtained in generations 1 and 3,800 and assuming an UC_e of 0.10 € kWh^{-1}, the energy cost in the first generation (GEN1) was 62,632 € (Fig. 5.3b).

After the optimization process a maximum reduction of energy consumption of 13.90 % was achieved in the last generation

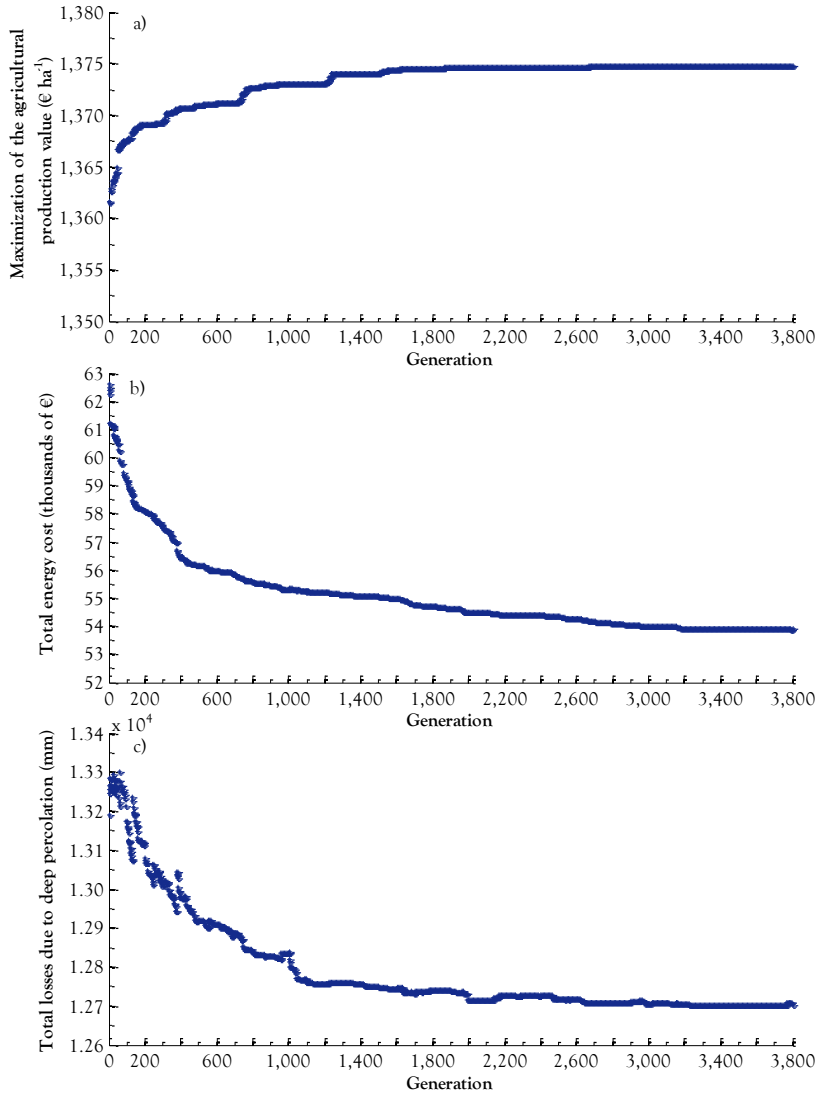


Fig. 5.3. F1 and F2 evaluation during the optimization process (3,800 generations). Maximization of the agricultural production value **(a)**, minimization of total energy cost **(b)** and minimization of total deep percolation loses **(c)**.

(GEN 3,800) with an associated cost of 53,928 €. This solution leads to a reduction in the total deep percolation losses and

production value of 3.79 % and 0.47 %, respectively (Fig. 5.3c and 5.3a). The total increase in the production value between the first and the last generation was 1.00 %, from 1,361.54 € ha^{-1} to 1,375.14 € ha^{-1} (Fig. 5.3a). This solution increased the energy cost and deep percolation in 1.26 % and 3.00 %, respectively. Thus, there is a wide range of intermediate solutions with different values for the key variables.

5.3.2. Optimal individuals

In order to facilitate the interpretation of the results, 5 individuals were selected from the 50 chromosomes in the Pareto front achieved last generation (3,800) and their values are shown in Table 5.2. Individual 1 was selected because it showed the best value of F1. For the same reason, Individual 5 achieved the best values for the terms of function F2 and therefore it was also selected. Individuals 2, 3 and 4 were randomly selected to represent the range of intermediate scenarios in the Pareto Front defined by combinations of the three key variables (production values, energy cost and deep percolation losses). The irrigation network was divided into three operating sectors which size (number of hydrants) varies in each individual Thus, if the irrigation network operates under conditions corresponding to Individual 1, the number of hydrants for the three sectors would be 58, 52 and 51, respectively. Under this scenario, the maximum production of Sector VII was achieved (1,375.14 € ha^{-1}) and the

Table 5.2. Production value, energy cost, energy consumption, water losses, hydrants per sector, hydrant elevation respect to pumping station for the 5 individuals selected from the Pareto Front.

		Individuals				
		1	2	3	4	5
Production value (€ ha^{-1})		1,375.14	1,368.56	1,362.00	1,353.53	1,343.57
Total Energy cost (€)		69,000	63,826	60,881	57,738	53,928
Total Losses due to Deep Percolation (mm) $x10^4$		1.38	1.34	1.30	1.27	1.27
Number of hydrants	Sector I	58	55	55	52	47
	Sector II	52	58	58	54	54
	Sector III	51	48	48	55	60

total energy cost in this scenario would be 69,000 €. The minor energy consumer was individual 5. Under this scenario, the total energy cost was 21.84 % lower than individual 1, that represents energy cost savings of 15,072 € while the production value was reduced by 2.30 % compared to individual 1 (which involves 31.57 € less incomes per hectare). The number of hydrants per sector, in this scenario, is more heterogeneous than in the other individuals (47, 54 and 60 for each sector respectively). Individuals 3 and 4 have similar values of the key variables. Production value was 0.96 % (individual 3) and 1.57 % (individual 4) lower than individual 1. Under these scenarios, the total energy cost was 11.77 % (individual 3) and 16.32 % (individual 4) less than the first individual. In individual 2, the production value is slightly reduced (0.48 % or 6.58 € ha^{-1} less than individual 1) and it has the same hydrant number per sector

than individual 3. However, the total energy cost is 7.50 % lower than the first individual with energy cost savings of 5,174 €.

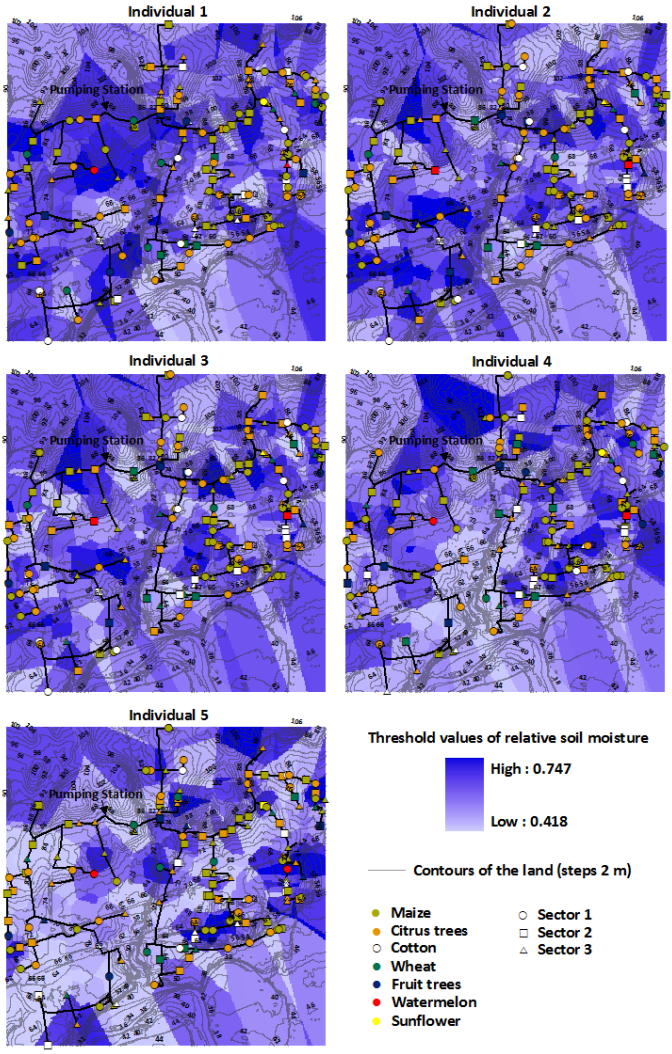


Fig. 5.4. Irrigation network sectoring and spatial distribution of crops and $s_{irrigation}$ values of each hydrant in the five individuals selected.

Fig. 5.4 shows spatial distribution of crops and $s_{irrigation}$ values for each hydrant as well as irrigation network sectoring for the five

individuals selected. This figure shows as individuals with higher crop production and therefore higher energy consumption have highest $s_{irrigation}$ values (darker areas). In all individuals, sectoring of the irrigation network does not show a clear relation to terrain elevation. So, the irrigation network was not sectored considering the topography. However, topography shows a slight correlation with $s_{irrigation}$ values. Overall, the highest areas have higher $s_{irrigation}$ values as it is not allowed that the soil moisture to drop too low to avoid a higher energy consumption. The $s_{irrigation}$ values are highly related to spatial distribution of crops in the irrigation network. Thus, hydrants which crops have high water needs or are more sensitive to water stress like maize or citrus trees, their $s_{irrigation}$ values are higher than less sensitive crops or with lower economic value.

Fig. 5.5 shows, variability of $s_{irrigation}$ values for each irrigation sector in the five individuals selected, respectively. The upper and lower bounds were 0.747 and 0.418, respectively, in all irrigation sectors for the 5 individuals. A lower $s_{irrigation}$ value means lower amount of water applied and lower total energy cost. Fig. 5.5 shows clearly that the variability of the $s_{irrigation}$ values are close to the upper limit in individual 1 and decreases to the lower boundary in individual 5 where energy cost and crop yield are minimal. The variability of these $s_{irrigation}$ values are also reduced from individual 1 to individual 5. Thus, in individual 1 most of the hydrants have an

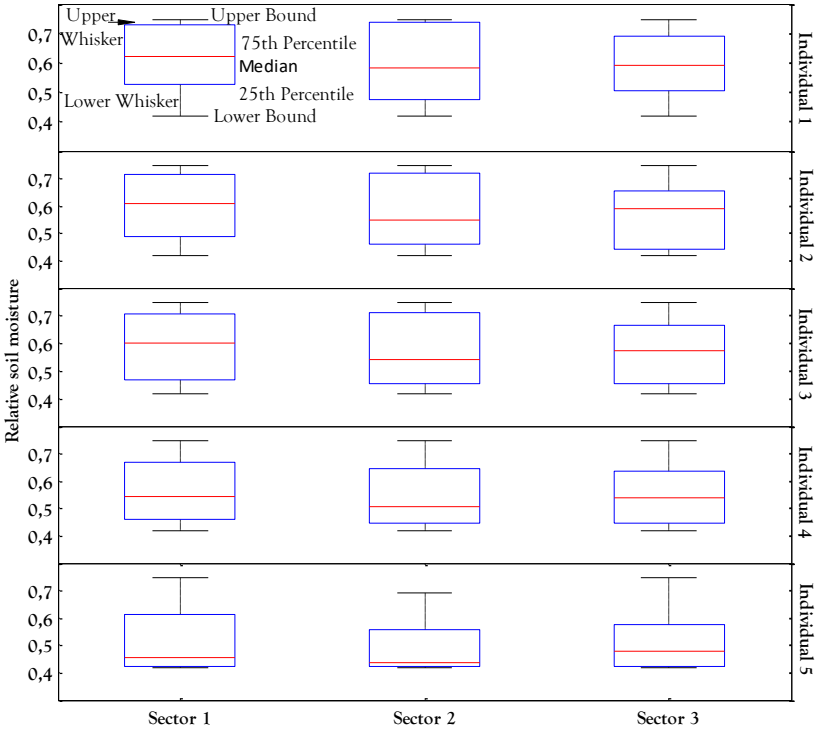
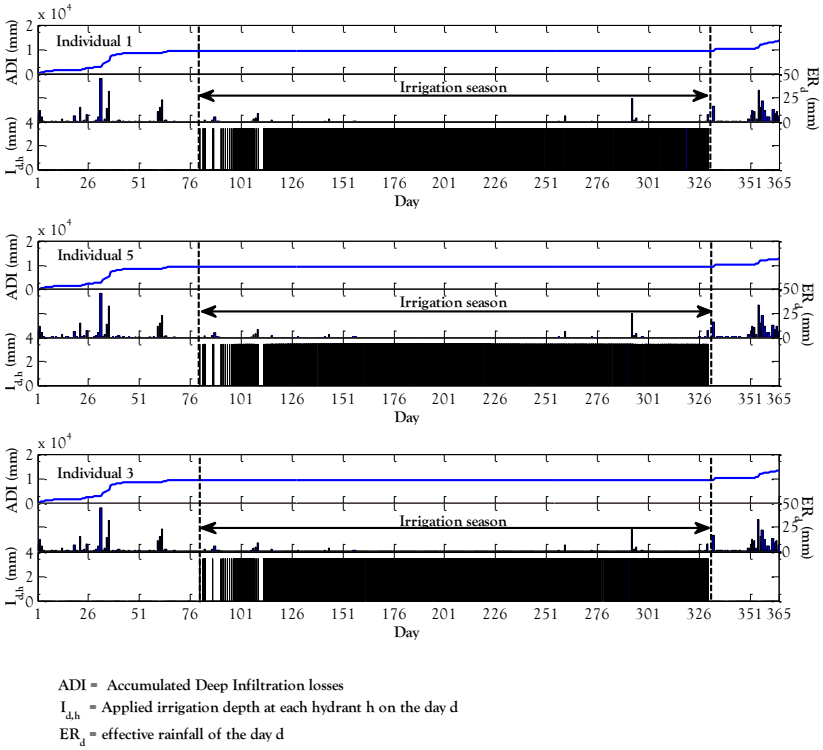


Fig. 5.5. Boxplot of variability of $s_{irrigation}$ values for each irrigation sector of the 5 individuals from Pareto Front.

associated high $s_{irrigation}$ value with median values for sector 1 of 0.622, 0.582 for sector 2 and 0.592 for sector 3, respectively. Consequently, under this scenario the total energy cost and the agricultural production reached the maximum value (69,000 € and 1,375.14 € ha^{-1} respectively). Individual 5 has the minimum value of the total energy cost (53,928 €) and its hydrants have the lowest $s_{irrigation}$ values (medians of 0.453, 0.435, 0.790 for sectors 1, 2 and 3, respectively). Intermediate $s_{irrigation}$ values are those obtained for Individuals 2, 3 and 4. Thus, under these scenarios,

the total energy cost was 63,826 € for individual 2, 60,881 € for individual 3 and 57,738 € for individual 4.

Fig. 5.6 shows cumulative deep infiltration losses, effective rainfall and the applied irrigation depth at each hydrant during the whole year.



ADI = Accumulated Deep Infiltration losses
$I_{d,h}$ = Applied irrigation depth at each hydrant h on the day d
ER_d = effective rainfall of the day d

Fig. 5.6. Cumulative deep percolation losses, applied irrigation depth and effective rainfall at plot scale in the 2009 season for individuals 1, 3 and 5.

This figure shows that water losses by deep percolation are caused by rainfall and not by over-irrigation. During the irrigation season, in the three individuals, the curve of accumulated deep

percolation losses remains constant. During the months with maximum irrigation needs (half of the irrigation season) all hydrants applied water.

Under the scenario corresponding to individual 1, the total energy cost and production of Sector VII, was 69,000 € and 1,375.14 € ha^{-1}, respectively. Knowing that the total irrigation area of Sector VII is 935 ha, the total income due to crop production was 1,285,756 €. Likewise, the energy costs and the total income due to crop production for individual 2, 3, 4 and 5 were 63,826 € and 1,279,604 €; 60,881 € and 1,273,470 €; 57,738 € and 1,265,551 €; 53,928 € and 1,256,238 €, respectively. Thus, considering only the crop production value and the energy cost, individual 1 had the highest profit out of the five individuals studied. Consequently, it can be considered the best individual and therefore the best management strategy for the studied area. According to the real data recorded in Sector VII of MDB Irrigation District, the total energy cost was 80,660 €. Thus, if the irrigation network had been operated under the conditions of the Individual 1, this irrigation sector would have obtained energy savings of 14.5 %.

5.4. Conclusions

The evolution of water distribution systems to pressurized networks has improved water use efficiency, but has increased

dramatically energy cost. To overcome this problem, we have presented a methodological approach to improve simultaneously energy and water use efficiency in irrigation networks.

The proposed methodology provides the optimal hydrant grouping in irrigation sectors and the optimal soil water content at the beginning of the season in each irrigated plot that, through a daily soil water balance, maximize the total economic value of crops production at irrigation district scale whilst minimize both the energy cost at the pumping station and the percolation losses for the whole irrigation district. A multi-objective optimization problem was stated with two objective functions and was solved using a customized version of the multi-objective genetic algorithm NSGA-II.

The methodology proposed has been applied to the BMD Irrigation District (Sector VII). The obtained results showed that integrated irrigation management would lead to energy cost savings close to 15 % in BMD Irrigation District without significant reduction of crop yield. This work shows that it is possible to reduce energy cost and improve water use efficiency. However, in order to achieve a global optimum use of water and energy, the operation of both the main water supply network and the irrigation network at farm level must be carried out jointly.

Acknowledgments. This research is part of the TEMAER project (AGL2014-59747-C2-2-R), funded by the Spanish Ministry of Economy and Competitiveness.

This research was supported by an FPU grant (Formación de Profesorado Universitario) from the Spanish Ministry of Education, Culture and Sports to Rafael González Perea.

5.5. References

Abadia R, Rocamora C, Ruiz A, Puerto H (2008) Energy efficiency in irrigation distribution networks I: Theory. Biosyst Eng 101:21–27

Allan JA (1998) Virtual water: A strategic resource global solutions to regional deficits. Ground Water 36:545–546

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage, Rome, Italy

Carrillo Cobo MT., Rodríguez Díaz JA, Montesinos P, et al (2011) Low energy consumption seasonal calendar for sectoring operation in pressurized irrigation networks. Irrig Sci 29:157–169

Consejería de Agricultura C de A (2009) Statistical annual report on agricultural and fisheries of Andalucía

Daccache A, Lamaddalena N, Fratino U (2010) On-demand pressurized water distribution system impacts on sprinkler network design and performance. Irrig Sci 28:331–339

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6:182–197

Doorenbos J, Kassam AH (1979) Yield response to water. FAO Irrigation and Drainage, Rome

Elferchichi A, Gharsallah O, Nouiri I, et al (2009) The genetic algorithm approach for identifying the optimal operation of a multi-reservoirs on-demand irrigation system. Biosyst Eng 102:334–344

English MJ, Solomon KH, Hoffman GJ (2002) A Paradigm Shift in Irrigation Management. J Irrig Drain Eng 128:267–277

FAO (2009) CROPWAT 8.0 Model

Fereres E, Soriano MA (2007) Deficit irrigation for reducing agricultural water use. J Exp Bot 58:147–59

Fernández García I, Rodríguez Díaz J a., Camacho Poyato E, Montesinos P (2013) Optimal Operation of Pressurized Irrigation Networks with Several Supply Sources. Water Resour Manag 27:2855–2869

Jury WA, Vaux HJ (2007) The Emerging Global Water Crisis: Managing Scarcity and Conflict Between Water Users. Adv. Agron. 95:1–76

Laio F, Porporato A, Fernandez-Illescas C., Rodriguez-Iturbe I (2001) Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress. Adv Water Resour 24:745–762

Lamaddalena N, Khila S (2012) Energy saving with variable speed pumps in on-demand irrigation systems. Irrig Sci 30:157–166

Navarro Navajas JM, Montesinos P, Poyato EC, Rodríguez Díaz J a. (2012) Impacts of irrigation network sectoring as an energy saving measure on olive grove production. J Environ Manage 111:1–9

Playán E, Mateos L (2006) Modernization and optimization of irrigation systems to increase water productivity. Agric Water Manag 80:100–116

Plusquellec H (2009) Modernization of large-scale irrigation systems: is it an achievable objective or a lost cause. Irrig Drain 58:S104–S120

Pratap R (2010) Getting started with Matlab. A quick introduction for scientist and engineers. Oxford University Press, USA

Rodríguez Díaz JA, Weatherhead EK, Knox JW, Camacho E (2007) Climate change impacts on irrigation water requirements in the Guadalquivir river basin in Spain. Reg Environ Chang 7:149–159

Rodríguez Díaz JA, López Luque R, Carrillo Cobo MT., et al (2009) Exploring energy saving scenarios for on-demand pressurised irrigation networks. Biosyst Eng 104:552–561

Rodríguez Díaz JA, Montesinos P, Camacho Poyato E (2012) Detecting Critical Points in On-Demand Irrigation Pressurized Networks - A New Methodology. Water Resour Manag 26:1693–1713

Rossman LA (2000) Epanet 2. Users Manual

Sadati S, Speelman S, Sabouhi M, et al (2014) Optimal Irrigation Water Allocation Using a Genetic Algorithm under Various Weather Conditions. Water 6:3068–3084

Schaap MG, Leij FJ, van Genuchten MT (2001) rosetta: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. J Hydrol 251:163–176

Siew C, Tanyimboh TT (2012) Penalty-Free Feasibility Boundary Convergent Multi-Objective Evolutionary Algorithm for the Optimization of Water Distribution Systems. Water Resour Manag 26:4485–4507

Wardlaw R, Bhaktikul K (2004) Application of genetic algorithm for irrigation water scheduling y. 414:397–414

Section II

Irrigation demand forecasting models at different scales

6. Irrigation Demand Forecasting Using Artificial Neuro-Genetic Networks

This chapter has been published entirely in the journal "Water Resources Management", González Perea R, Camacho Poyato E, Montesinos P, Rodríguez Díaz JA (2015)

Abstract. In recent years, a significant evolution of forecasting methods has been possible due to advances in artificial computational intelligence. The achievement of the optimal architecture of an ANN is a complex process. Thus, in this work, an Evolutionary Robotic (study of the evolution of an ANN using Genetic Algorithm) approach has been used to obtain an Artificial Neuro-Genetic Networks (ANGN) to the short-term forecasting of daily irrigation water demand that maximizes the accuracy of the predictions. The methodology is applied in the Bembézar Irrigation District (Southern Spain). An optimal ANGN architecture (ANGN (7, 29, 16, 1)) has achieved obtaining a Standard Error Prediction (SEP) value of the daily water demand of 12.63 % and explaining 93 % of the total variance observed during validation process. The developed model proved to be a powerful tool that, without long dataset and time requirements, can be very useful for the development of management strategies.

Keywords. Optimal forecasting models, Artificial intelligence, Seasonal model update, Evolutionary robotics

6.1. Introduction

As a result of the increasing competition for water resources, water is considered a scarce and valuable resource that requires a rigorous management and extreme caution to prevent its depletion. In countries like Spain, where 73 % of the National freshwater is devoted to irrigation sector (MAGRAMA 2013; INE 2014), many irrigated areas have been subjected to modernization processes with the aim of increasing the water use efficiency, through more efficient irrigation systems such as drip and trickle (Playán and Mateos 2006). Consequently, most of the new irrigation systems are operated on-demand where water is always available for farmers and they decide when to irrigate and the duration of each irrigation event. However, the increased operation flexibility hinders the prediction of water demand in irrigation districts. This circumstance causes problems for managers who would need this information for the day to day management tasks such as contracting the electric energy supply. Since the liberalization of the Spanish Electricity Market, on 1st January 2008, the special tariffs for irrigation disappeared and now, irrigation districts are subject to the general industrial tariffs. During the months of June and July, the peak of the irrigation

season, most of the irrigation time occurs in periods of expensive tariffs. Recently, in order to minimize energy cost, irrigation district managers have the possibility to hire electric energy through different modalities for adjusting the electrical power contracted to the real power absorbed at the irrigation network. One of these modalities is known as the electricity tariff indexed to pool (pass through), where the electrical energy is paid at a variable price based on the wholesale market. Another modality is known as superindexed electricity tariff where a central purchasing body has its own strategy of purchasing power (buying futures market, pass through and intraday Markets) and the users have to pay in advance their estimated electricity consumption. Thus, to achieve an optimal management, irrigation district managers need tools to estimate accurately the real daily water demand of the entire irrigation network.

Modelling techniques have been used to estimate the crops daily water requirements, from empirical or functional (Doorenbos and Pruitt 1977; Doorenbos and Kassam 1979; Allen et al. 1998) to mechanistic approaches (Van Aelst et al. 1988). However, water requirements calculated are not always suitable for predicting actual use (i.e., consumer demand) due to changes in the weather conditions and local farmer practices that affect the actual amounts of water applied.

In recent years, a significant evolution of forecasting methods has been possible due to advances in artificial computational intelligence, in particular the Artificial or Computational Neural Networks (ANNs or CNNS). A neural network is a system that allows for linear or nonlinear relationship between outputs and inputs. Its main features are inspired in the nervous system which gives them several advantages such as to have adaptive learning ability, to be self-organizing, to be able to operate in parallel in real time and to provide fault tolerance by redundant information coding.

Several specific applications of ANN to water resource management and planning include the modeling of monthly, daily and hourly rainfall–runoff processes (Hsu et al. 1995; Lorrai and Sechi 1995; Mason et al. 1996; Abrahart et al. 1999; Tokar and Johnson 1999; Thirumalaiah and Deo 2000; Tokar and Markus 2000; Chiang et al. 2004; Moradkhani et al. 2004; Anctil and Rat 2005; Agarwal et al. 2006), real- time river and lake stage forecasting (Thirumalaiah and Deo 1998; Abrahart and See 2000; See and Openshaw 2000; Thirumalaiah and Deo 2000; Abrahart and See 2002; Cameron et al. 2002; Nayebi et al. 2006; Ondimu and Murase 2007), rainfall forecasting (French et al. 1992; Zhang et al. 1997; Kuligowski and Barros 1998), groundwater modeling (Rogers and Dowla 1994; Yang et al. 1997), assessment of stream's hydrologic and ecological response to climate change (Poff et al.

1996), drought analysis (Shin and Salas 2000), etc. However, Pulido-Calvo and Gutiérrez-Estrada (2009) is the only application of ANN related with water demand forecasting in pressurized systems at irrigation district level. In that study, the performance of a hybrid methodology combining feed forward CNN, fuzzy logic and genetic algorithms to forecast one-day ahead daily water demands at irrigation districts were analyzed. The developed methodology was applied to a real irrigation district located in southern Spain. The forecast of the individual models was corrected via a fuzzy logic approach whose parameters were adjusted using a genetic algorithm in order to improve the forecasting accuracy. A major limitation in this model was the determination of the ANN architecture by trial and error.

The ANN and genetic algorithms are soft-computing technologies that can be very effective when used on their own. However, when combined together, the individual strengths of each approach can be exploited in a synergistic manner for the construction of powerful, hybrid and intelligent systems (See and Openshaw 2000). The discipline that studies the evolution of an ANN using Genetic Algorithm is known as Evolutionary Robotic (ER). The achievement of the optimal architecture of an ANN (minimum computational speed and maximum forecast accuracy) is a complex process. Thus, in this work, an ER approach has been used to obtain an ANN to the short-term forecasting of daily

irrigation water demand that maximizes the accuracy of the predictions. The developed model is tested with actual data recorded in the water distribution network of a real irrigation district.

6.2. Methodology

6.2.1. Study area and Data Source

The developed model was applied to Bembézar M.D. Irrigation District (BMD), located in Andalusía (Southern Spain). The BMD water distribution network irrigates 11,950 ha within the Bembézar River basin (Fig. 6.1). The climate in the area is Mediterranean with annual average rainfall and temperature of 604 mm and 17.7 °C, respectively. The mean annual potential evapotranspiration is over 1,200 mm.

BMD is made up of eleven working independently. All of them were designed to supply 1.2 L s^{-1} ha^{-1} on-demand (water is always available for farmers) with a service pressure of 35 m at hydrant level. Drip irrigation is the most common irrigation system.

Among the eleven sector networks, Sector VII was selected for this study. It covers a total irrigated area of 935 ha (Fig. 6.2) and includes the most representative crops in the region: maize, citruses, cotton, wheat, fruit trees, watermelon and sunflower.

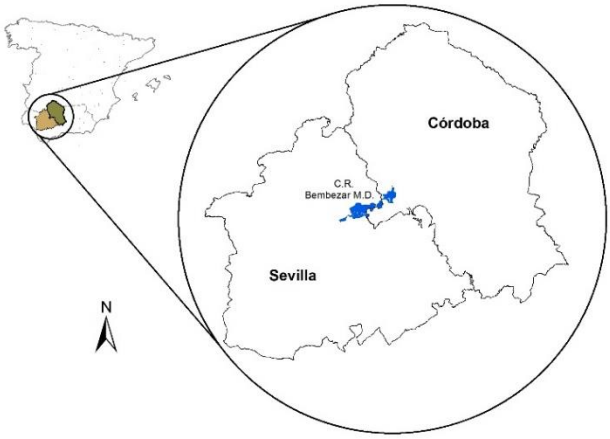


Fig. 6.1. Location of Bembézar M.D. irrigation district, Spain.

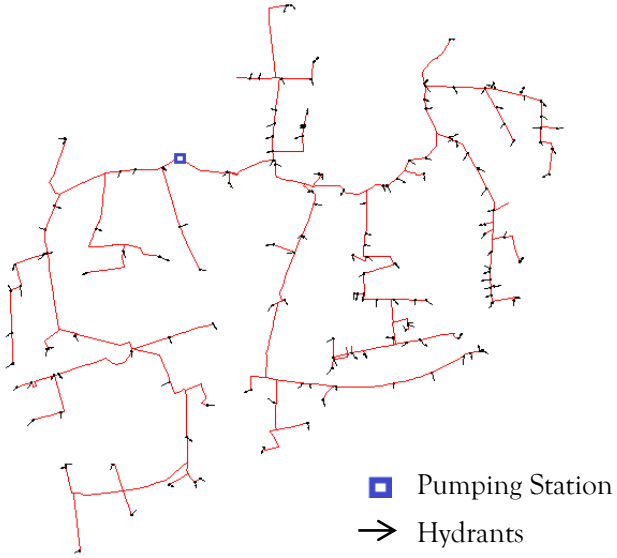


Fig. 6.2. Layout of the Sector VII irrigation scheme (BMD).

At the pumping station pressure heads and pumped flows every minute are recorded by a telemetry system. Water demand data

were aggregated at daily level for the 2010, 2012 and 2013 irrigation seasons. Due to errors in the recording process of the water demand values, the data series was not complete in any of the irrigation seasons studied. Further, then all water demand values were removed to facilitate the training process of ANN. The daily climatic data were obtained from the agroclimatic station placed in the irrigation district.

6.2.2. Optimal Artificial Neuro-Genetic Networks

The both the architecture and the training and validation processes of ANNs are critical to obtain accurate predictions from a set of input variables. A multiobjective genetic algorithm has been implemented to identify the best sets of architecture parameters as well as the best training and validation alternatives that characterized optimal ANNs, called Artificial Neuro-Genetic Networks, ANGNs. To facilitate the understanding of the procedure, a description of the decision variables required to create optimal ANGN and their position (gene) in the artificial chromosome required to apply the genetic algorithm are given next.

6.2.2.1. Architecture Parameter of ANNs

The most widely used ANN is the Multilayer Perceptron Network (MLP) (Rumelhart et al. 1986), whose structure is presented in

Fig. 6.3. A typical four-layer feed forward ANN (*g*, *n*, *m*, *ss*), has *g*, *n*, *m* and *ss* nodes or neurons in the input, first hidden, second hidden and output layers, respectively. Each layer consists of a number of neurons, which are connected to the next layer's neurons by synaptic weights (*w*). All the connections are feed forward, thus they are only allowed to transfer information from a previous layer to the next one. The number of neurons in the

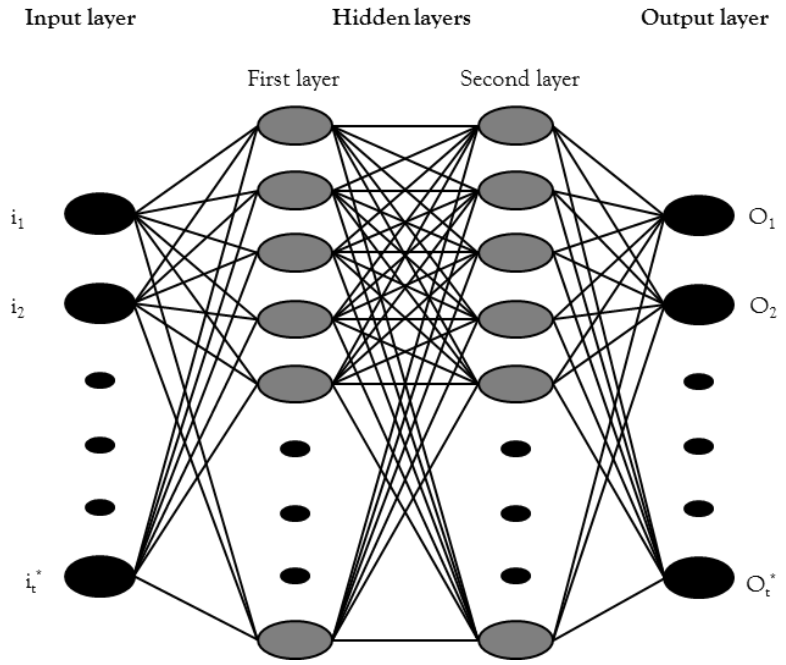


Fig. 6.3. Multi-layer perceptron neural network.
*t is total number of input variables of the neural network.

input and output layers are the number of input and output variables respectively. The numbers of neurons of the first, *n*, and second hidden layer, *m*, are two of the decision variables (structure variables) included the chromosome or set of variables

to be optimized by the genetic algorithm, in the genes located in positions 1 and 2. The transfer functions between layers (input, hidden and output layer) are the three transfer function variables located in genes 3, 4 and 5 respectively. The transfer functions used are shown in Table 6.1.

Table 6.1. Decision variables of each chromosome of the genetic algorithm (NSGA-II).

	Gene Number	Variable name	Value rate
ANN architecture	**1**	Neuron number of the first hidden layer	Integer value between 1 and 50.
	2	Neuron number of the second hidden layer	Integer value between 1 and 50.
	3	Transfer function of the input layer	Integer value between 1 and 3: (1) *Log-Sigmoid Transfer Function* (*LogST*); (2) *Tan-Sigmoid Transfer Function* (*TanST*); (3) *Linear Transfer Function* (*LinT*).
	4	Transfer function of the hidden layers	Integer value between 1 and 3: (1) *Log-Sigmoid Transfer Function* (*LogST*); (2) *Tan-Sigmoid Transfer Function* (*TanST*); (3) *Linear Transfer Function* (*LinT*).
	5	Transfer function of the output layer	Integer value between 1 and 3: (1) *Log-Sigmoid Transfer Function* (*LogST*); (2) *Tan-Sigmoid Transfer Function* (*TanST*); (3) *Linear Transfer Function* (*LinT*).

Table 6.1. Continuation.

	Gene Number	Variable name	Value rate
Training process of ANN	**6**	Training function, *TNF*	Integer value between 1 and 12: (1) *Batch Gradient Descent* (*BGD*) (Rumelhart et al. 1986); (2) *Gradient Descent with Momentum* (*GDM*) (Rumelhart et al. 1986); (3) *Variable Learning Rate with momentum* (*VLRM*) (Hagan et al. 1996); (4) Resilient Backpropagation (*RB*) (Riedmiller and Braun 1993); (5) *Fletcher-Reeves Update* (*FRU*) (Fletcher and Reeves 1964); (6) *Polak-Ribiére Update* (*PRU*) (Fletcher and Reeves 1964);(7) *Powell-Beale Restarts* (*PBR*) (Powell, 1977); (8) Scaled Conjugated Gradient (*SCG*) (Møller 1993); (9) *BFGS Algorithm* (Dennis and Schnabel, 1983); (10) *One Step Secant Algorithm* (*OSS*) (Battiti 1992); (11) Levenberg-Marquardt algorithm (LM) (Hagan and Menhaj 1994); (12) *Variable Learning Rate* (*VLR*) (Hagan et al. 1996).
	7	Momentum constant (*mc*)	When TNF is 2 or 12 mc gets a decimal value between 0 and 1. For the rest of TNF mc is 0.
	8	Line search algorithm, *LSA*	When *TNF* takes the values of 1, 2, 3, 4, 8, 11 or 12; *LSA* gets a values of 0, however when *TNF* is 5, 6, 7, 9 or 10; *LSA* gets an integer value between 1 and 5: (1) Golden Section Search (*GS*) (Hagan et al. 1996); (2) Brent's Search (*BS*) (Brent 1973); (3) Hybrid Bisection-Cubic Search (*HBC*) (Scales 1985); (4) Charalambous' Search (*CS*) (Charalambous 1992); (5) Backtracking (*BT*) (Dennis and Schnabel 1983).
	9	Learning function	Integer value 1 or 2: (1) Gradient Descent; (2) Gradient Descent with Momentum.

Table 6.1. Continuation.

	Gene Number	Variable name	Value rate
Improving generalization	**10**	Data division function	Integer value between 1 and 4: *(1) Index Data Division* (*IDD*); (2) *Random Data Division* (*RDD*); (3) Block Data Division (*BDD*); (4) *Interleaved Data Division* (*InterDD*).
	11	Value of training set	Integer value between 75% and 82%.
	12	Perform function, *PF*	Integer value between 1 and 2: (1) mean sum of squares (*mse*); (2) *msereg*.
	13	Performance ratio (γ)	When *PF* is 1, γ gets a value of 0 and when *PF* is 2, γ *gets a* decimal value between 0 and 1.

6.2.2.2. Training of an ANN

To determine the set of weights, a corrective-repetitive process called learning or training of the ANN is performed. Training aims to define the interconnections between neurons (weights) adjusting the weights through training patterns (known set of inputs and outputs). These interconnections are adjusted using an error convergence technique.

Twelve different training functions can be chosen by the GA for ANGN training. The training function is located in gene 6. The standard backpropagation learning algorithm is the most widely used supervised algorithm in ANNs. The simplest implementation of backpropagation learning updates the network weights in the direction in which the performance

function decreases most rapidly, the negative of the gradient. There are two different ways in which this gradient descent algorithm can be implemented: incremental mode and batch mode. In this work, the batch mode has been used. In this batch mode, all the inputs are applied to the network before the weights are updated. There are many variations of the back propagation algorithm. In this work, several of these training methods have been used to optimize the accuracy of the prediction model.

In Table 6.1 the 12 training algorithms used in this work are shown (field number 6). *Batch Gradient Descent* (BGD) is the simplest training method in which the weights are updated in the direction of the negative gradient of the performance function. A variant of the BGD was *Gradient Descent with Momentum* (GDM). GDM allows the network to respond not only to the local gradient, but also to recent trends in the error surface. GDM depends on two training parameters: *learning rate* that is similar to the simple gradient descendent *momentum constant* (*mc*) that defines the amount of momentum (Hagan et al. 1996). With momentum, a network can slide through such a minimum and avoid getting stuck in a shallow local minimum (Demuth et al. 2009). In this work, learning rate had a constant value of 0.05 and *mc* was a decision variable in the genetic algorithm (field number 7). These two (BGD and GDM) training methods are often too slow for practical problems but with the right combination of

the rest of parameters that form the ANGN it can be achieved acceptable performance. The rest of the training functions are aimed at improving the learning time and thus the performance of the ANGN. Some of these training algorithms can use different search algorithm (line searches routines) of the negative gradient of the performance function. It is often difficult to predict which of these line search routines provides the best results for any given problem. For this reason, line search routine was a decision variable in the genetic algorithm. In Table 6.1, field number 8 shows the search routines used in this work.

The learning process of the weights was another decision variable in the genetic algorithm (field number 9). To the backpropagation training algorithms described above, the main learning functions were: *Gradient Descent* and *Gradient Descent with Momentum*.

6.2.2.3. ANN generalization

During the neural network training, the error on the training set is driven to a very small value, but when new data is presented to the network the error is usually higher. This phenomenon is known as overfitting. The network has memorized the training example, but it did not learn to generalize to new situations. Two methods have been used to improve the network generalization: *Early Stopping* and *Regulation* (Demuth et al. 2009). In the first method, the available data is divided into two subsets. The first

subset is the training set, which is used for computing the gradient and updating the network weights. The second subset is the validation set. The validation error normally decreases during the initial phase of training, as the training set error does. However, when the network starts to overfit the data, the error on the validation set rises. When the validation error increases for a specified number of iterations, the training is stopped, and the network training is finalized. In this work, the function for dividing data into training and validation sets was a decision variable in the genetic algorithm. Four functions have been used for splitting data into the two subsets (Hagan et al. 1996) and are shown in Table 6.1 (field number 10). The first function (*Index Data Division*, IDD) divide the data according to its position in the data serial so that the registers of the data serial are assigned to the training set, and the validation set alternatively. The second function (*Random Data Division*, RDD) divide the input data randomly so that a percentage of the data are assigned to the training set and another percentage for the validation set. The third function (*Block Data Division*, BDD) divide the input data randomly so the first X % of the data are assigned to the training set, and (100-X) % to the validation set. Another way to divide the input data is to cycle samples between the training set and validation set according to percentages. The last function (*Interleaved Data Division*, InterDD) divide the input data in this

way. Gene located in position 11 represents the X % of data set that constitute the training set and (100-X) % represents the validation data set. The most authors take a fixed value of X of 80 %. In this work X ranges from 75 % to 82 % to the genetic algorithm optimize the data numbers of the training and validation sets.

The typical performance function used for training feed forward neural networks is the mean sum of squares (*mse*) of the network errors:

$$mse = \frac{1}{OB} \cdot \sum_{t=1}^{OB} (Q_t - \widehat{Q_t})^2 \qquad [6.1]$$

where *OB* is the total number of observations used for neural network training; Q_t is the observed water demand at the time step *t* (L s^{-1}) and $\widehat{Q_t}$ is the estimated water demand at the same time step *t* (L s^{-1}).

Adding the mean of the sum of squares of the network weights to Eq. 6.1, it is possible to improve the neural network generalization. Thus, a new performance function is shown in Eq. 6.2.

$$msereg = \gamma \cdot \text{mse} + (1 - \gamma) \cdot msw \qquad [6.2]$$

where γ is the performance ratio that forces the network response to be smoother and less likely to overfit. *msw* is defined by:

$$msw = \frac{1}{nm} \cdot \sum_{j=1}^{nm} w_j^2 \qquad [6.3]$$

where nm is the total number of the neural network weights.

Both, the performance function and the performance ratio, are decision variables located in 12 and 13 gene of the chromosome, respectively.

6.2.3. Optimizing the ANGN Model with a Multiobjective Genetic Algorithm

6.2.3.1. Problem Approach

A multiobjetive optimization problem with two objective functions was stated for the prediction of the daily water demand. The aim of the first objective function (*F1*) was to maximize the sum of the coefficients of determination (R^2) of training and validation sets. This coefficient describes the proportion of the total variance in the observed data that can be explained by the model.

$$F1 = \left[R^2_{training} + R^2_{validation}\right] \qquad [6.4]$$

R^2 was calculated according to the following equation:

$$R^2 = \left(\frac{\sum_{t=1}^{OB_1}(\widehat{Q_t} - \bar{\hat{Q}}) \cdot (Q_t - \bar{Q})}{\sqrt{\sum_{t=1}^{OB_1}(\widehat{Q_t} - \bar{\hat{Q}})^2 \cdot \sum_{t=1}^{N_1}(Q_t - \bar{Q})^2}} \right) \qquad [6.5]$$

where OB_1 is the total number of observations used for neural network training or validation sets; $\bar{\hat{Q}}$ is the average of estimated water demand of training or validation sets (L s^{-1}) and $\bar{Q}$ is the average of observed water demand of training or validation sets (L s^{-1}).

The aim of the second objective function (*F2*) was to minimize the average normalized root mean square error (*RMSE*) of the validation sets.

$$F2 = RMSE_{norm} = \left[\sqrt{\frac{1}{OB_{val}} \cdot \sum_{t=1}^{OB} (\widehat{Q_t} - Q_t)^2}\right]_{norm} \qquad [6.6]$$

where OB_{val} is the total number of observations used for neural network validation.

In order to compare the two objective functions, *F2* was normalized using a cumulative distribution function for the continuous uniform distribution on the interval [0, 2]. Thus, the minimum value of both objective functions was 0 and the maximum value 2.

With a view to be able to make comparisons between different models other measure of variance applied was the percent standard error of prediction (*SEP*) (Ventura et al. 1995). The *SEP* is defined by

$$SEP = \frac{100}{\overline{Q_v}} \cdot RMSE \qquad [6.7]$$

where $\overline{Q_v}$ is the average of the observed water demand of the validation set (L s^{-1}).

6.2.3.2. Optimization Method. NSGA-II

The multiobjetive algorithm NSGA (Deb et al. 2002) was implemented in MATLAB (Pratap 2010) to obtain an optimal prediction model of water demand in an irrigation network. The standard NSGA-II algorithm was adapted to solve the problem stated in section 6.2.3.1 (Fig. 6.4). In the first step, the initial population of *nPop* chromosomes, composed by 13 genes defined in sections 6.2.2.1, 6.2.2.2 and 6.2.2.3 was randomly generated. Every chromosome represents one model of ANGN.

Then, the objective functions, *F1* and *F2*, were calculated for each chromosome. In the remaining stages, the chromosomes were modified (crossover and mutation) and the top *nPop* were selected based on their objective function values. The process was repeated several generations (*nGEN*). Finally, the set of *nPop* optimal chromosomes obtained in the last generation define the Pareto Front.

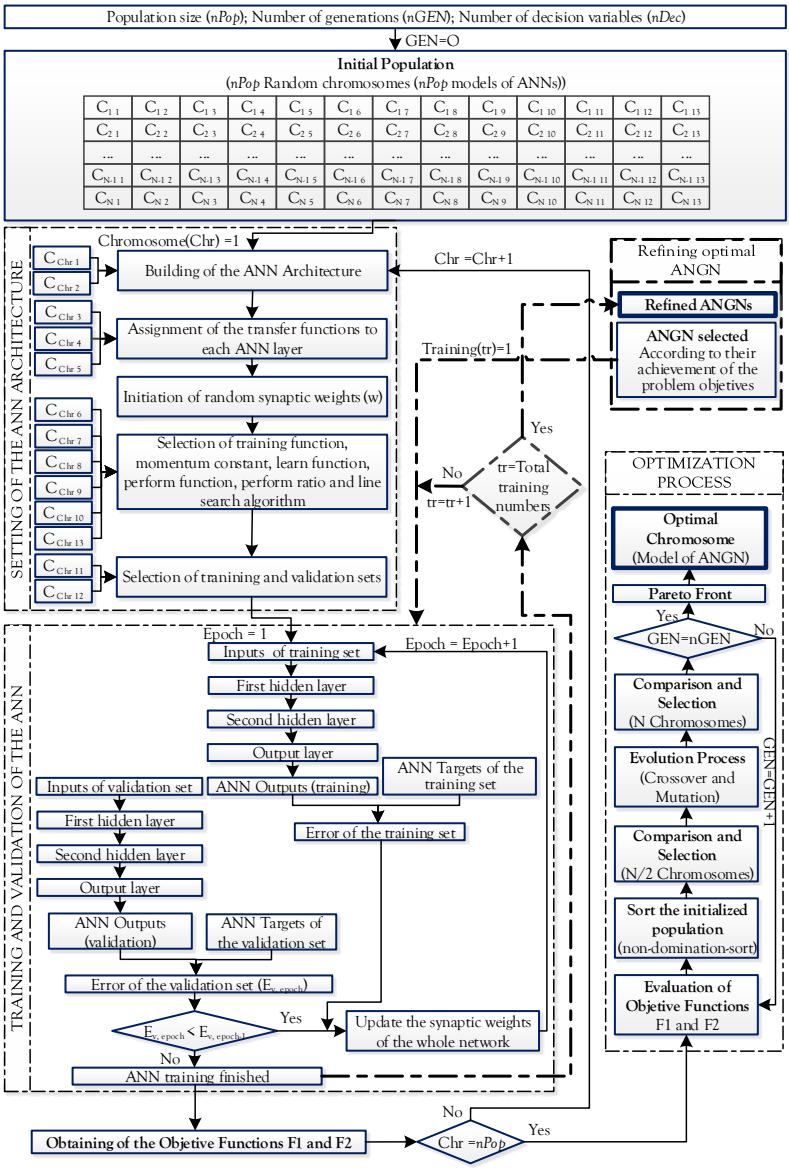


Fig. 6.4. Optimization process using NSGA-II.

A time constraint (maximum 60 s runtime) controlled the number of epoch in the training process. Therefore, there is no a fixed number of epochs in the training process. If the convergence

of the training algorithm is slow, the number of epochs in a period of 60 s is lower than if the convergence of the training algorithm is faster. Thus, an indirect way, the multiobjetive genetic algorithm kept those ANGN settings faster. To predict water demand in real time, speed training is a key factor to implement these models in real irrigation districts.

6.2.3.3. Refining Optimal ANGNs

After obtaining the Pareto front, some of the ANGN models were selected according to their achievement of the problem objectives. These ANGN models were trained with a random initial configuration of synaptic weights. In order to remove this random effect that can make the solutions of the perform function are positioned in an unwanted location of the error function, these ANGNs were trained again without time constraint. Thus, the accuracy of the predictions of water demand is lightly improved.

6.3. Results and Discussion

Twenty-one potential variables were evaluated in order to choose the most representatives for the prediction model. These potential variables are composed by weather variables provided by the nearest weather station to the study area and the values of registered water demand, all of them for the day to predict, the previous day and two previous days to predict. After performing

a correlation analysis for 21 potential variables, 7 input variables in the prediction model were selected: Water demand in the previous day (*Demand_1*) (L s^{-1});Water demand in the two previous days (*Demand_2*) (L s^{-1}); Average temperature for the day to predict (T_{ave}) (°C); Solar radiation for the day to predict (*Rad*) (MJ m^{-2}); Solar radiation in the previous day (*Rad_1*) (MJ m^{-2}); Reference evapotranspiration for the day to predict (ET_0) (mm day^{-1}); Reference evapotranspiration in the previous day (ET_0_1) (mm day^{-1}). Therefore, the neuron numbers of the input layer were 7 (*g*=7) and the neuron number of the output layer was 1 (water demand) (*s*=1). Thus, the ANGN architecture in all predictive models will be ANGN (7, *n*, *m*, 1).

6.3.1. The Pareto Front of *nPop* Artificial Neuro Genetic Networks

The developed model was applied to BMD (Sector VII) and both objective, *F1* and *F2*, were optimized. The random initial population consisted of 70 individuals (chromosomes) which were composed of 13 genes (Table 6.1). Every gene represents a different characteristic of the ANGN. The initial population was evolved for 130 generations and the probabilities for crossover and mutation were set to 90 % and 10 %, respectively.

The Pareto front (Fig. 6.5) was obtained in generation 130. This graph clearly shows that both objectives (*F1* and *F2*) are not

conflicting because most individuals are concentrated in the lower right corner of the graph, i.e., the zone of the graph where *F1* reached the maximum value and *F2* obtained the minimum value. A detailed view of this region of the Pareto front shows the three best individuals which are also highlighted (Fig. 6.5). The lowest *RMSE* value is represented by individual *I1*. The highest value of R^2 in the validation period is represented by individual *I3* and individual *I2* is in between the other two individuals.

This individual *I1* was the one the lowest associated error when predicting the daily water demand, so *I1* represents the most accurate predictive model. However, *I1* was unable to explain the maximum variability of the observed water demand. On the other hand, *I3* was the best predictive model when explaining the variability of the observed data but it was the less accurate predictive model. Depending on the objectives of the irrigation district manager, such as the procurement of daily electric energy, it might not be advisable to have a predictive model that explains the maximum variability but with a high prediction error or a predictive model that even with a low prediction error was not able to represent the variability observed. In that case, individual *I2* might be the best predictive model.

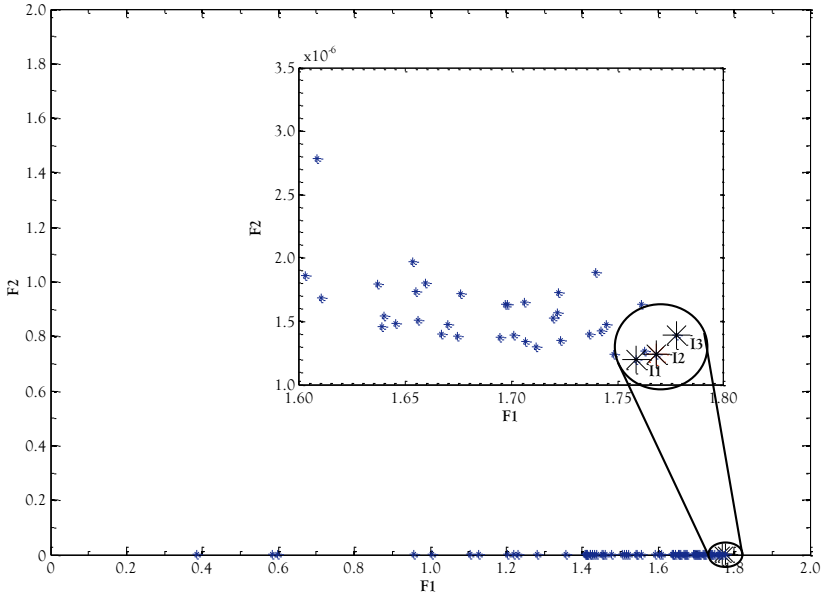


Fig. 6.5. Pareto front for generation 130.

6.3.2. Refined Optimal ANGN Models

The three best individuals were trained again (200 training each of them) to improve their accuracy of their predictions, as the initial values of weights were randomly generated. The three ANGN models are shown in Table 6.2. The R^2 and *RMSE* values in the validation period were 0.93, 0.90 and 0.92; 55.18 L s^{-1}, 58.10 L s^{-} 1 and 53.42 L s^{-1} for *I1*, *I2* and *I3*, respectively. The model *I2* reached the lowest R^2 value (0.90) and the highest *RMSE* value (58.10 L s^{-1}). Therefore, this network is not the best option to predict the daily water demand.

Table 6.2. The three best ANGN models.

Model		1	2	3
Neuron number of the first hidden layer		22	3	29
Neuron number of the second hidden layer		14	23	16
Transfer function of the input layer		LogST	TanST	LogST
Transfer function of the hidden layers		LogST	LogST	TanST
Transfer function of the output layer		TanST	LinT	TanST
Training function		RB	PRU	PRU
Momentum constant (mc)		-	-	0.87
Line search algorithm		CS	BS	CS
Learn function		Gradient Descent	Gradient Descent	Gradient Descent with Momentum
Data division function		InterDD	InterDD	RDD
Value of training set (%)		81	75	80
Perform function		mse	mse	mse
Performance ratio (γ)		-	-	-
Validation period	R^2	0.92	0.90	0.93
	RMSE (L s^{-1})	53.42	58.10	55.18
	SEP	12.72	13.48	12.63

(-) Parameter not used in this ANGN configuration

The model *I3* was trained with the 80 % of all available data. This ANGN model was trained with *PRU* algorithm (*Polak-Ribiére Update Algorithm*), achieved the highest R^2 value (0.93) in the validation period. The numbers of neurons in the two hidden layers were 29 and 16, respectively. Thus, the neural network configuration is represented as [ANGN3 (7, 29 16, 1)]. However, this increase of the explained variance level was not linked to the

error (*RMSE*) reduction. Model *I1* achieved the lowest *RMSE* value (53.42 L s^{-1}), which was trained with the *RB* algorithm (*Resilient Back propagation Algorithm*) and it is the most accurate to predict irrigation water daily demand according to *RMSE* value. The network configuration was [ANGN1 (7, 22, 14, 1)]. The *SEP* values of the three ANGN models were 12.72 %, 13.48 % and 12.63 %, respectively. The optimal characteristics of individuals *I1*, *I2* and *I3* are shown in Table 6.2.

Scatterplots for models *I1* and *I3* are presented in Fig. 6.6 which also includes the diagrams of the observed and forecasted water demands from the validation period. This figure shows that the model *I3* [ANGN3 (7, 29, 16, 1)] obtained the closest match between forecasted and observed water demands over the whole daily water demand range. The higher differences in both graphs occurred in some peak demand days where the model was not been able to predict accurately. There are some reasons why these models can not accurately predict some peak values, such as the variability of the training set. Thus, it is possible that the models *I1* and *I3* have not been properly trained to reproduce extreme values due to the lack of patterns of extreme events in the training set. This is the main reason for the difference between the predicted and observed values in some peak days. If new training patterns were added and the ANGN was trained again, the

predictive model would be probably able to predict the daily water demand in a wider range of values (Fig. 6.6).

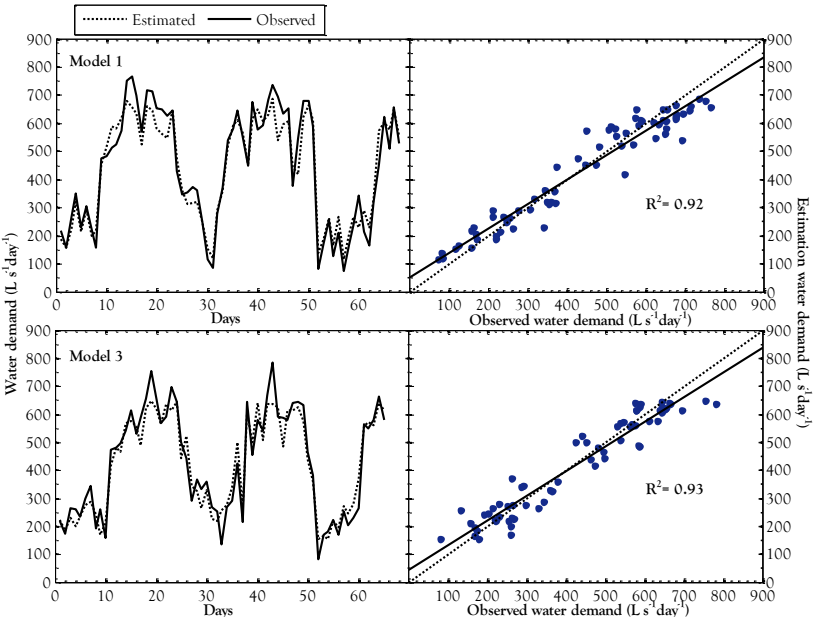


Fig. 6.6. Water demand prediction versus observed values (validation period).

In model *I1*, the highest difference between the predicted and observed values occurred in the days 14 and 44. In the first point (day 14), the observed value of water demand on that day was 751.54 L s^{-1} while the predicted value was 680.56 L s^{-1}. Thus, the difference between both values is slightly higher than 9 %. In the day 44, the difference between observed and predicted values was 22.5 % (the observed and predicted values were 694.42 L s^{-1} and 537.72 L s^{-1}, respectively). In model *I3*, the days 19, 33 and 43 had the largest difference between the observed and predicted

values. The values observed in those three days were 755.60 L s^{-1}, 133.18 L s^{-1} and 782.53 L s^{-1}, respectively and the values predicted were 646.53 L s^{-1}, 258.10 L s^{-1} and 634.43 L s^{-1}, respectively. These differences represent a 14.4 %, 48.4 % and 18.9 % for data numbers 19, 33 and 43, respectively. The observed values at day 33 is very unusual in the data set used, hence the great difference between the observed values and predicted values in this point. But even knowing these possible differences in the prediction of extreme values, which occurs in a very limited number of cases, this tool may help managers in the day to day network operation.

According to the R^2 and *SEP* values the most adequate neural network model to predict daily water demand in the Bembézar M.D. Irrigation District was the model *I3*. These values (R^2=0.93 and *SEP*=12.63 %) improves those obtained by Pulido-Calvo and Gutiérrez-Estrada (2009) where R^2 was 0.89 and *SEP* was 20.27 % in their hybrid model for daily water demand forecast. Consequently, the model developed in this paper explains better the variability of observed daily water demand and with smaller prediction error in the daily forecasts. There are two main factors for this improvement. On the one hand, the number of neurons in the first and second hidden layers were optimized by the genetic algorithm (NSGA-II) and it influences the accuracy and generalization of the predictive model. Thus, a small number of

neurons increases the computation speed of the ANGN but also increases the prediction error while more neurons increase accuracy and computation time of the ANGN and also increases the risk to overfitting the observed patterns. Therefore, to obtain an optimal generalization of the ANGN is necessary to determine the optimum number of neurons that provides the optimal ANGN architecture. On the other hand, the relationship between training time and accuracy of the obtained model determines the performance of the ANGN. There are training functions that can converge from ten to one hundred times faster than others. It is difficult to predict which training function will perform best on a given problem. Thus, the inclusion of the training function as a decision variable of the genetic algorithm has been decisive to obtain a good generalization of the predictive model.

This predictive model is fed by some climatic variables that are estimated by agroclimatic station. Consequently, uncertainty caused by the estimation of these climatic variables may influence the accuracy of the predictive model. This limitation should be taken into account in decision-making of the Irrigation District manager. However, the main limitation of the model developed in this work is the accuracy in peak water demand days. The incorporation of the predictive model in an irrigation district should be a dynamic process. Thus, every irrigation season the

ANGN should be updated with new data and trained again with a longer dataset. Then, the predictive model will probably explain better the observed variability and will obtain more accurate results even in peak days.

6.4. Conclusions

A forecasting model of the daily water demand of an irrigated that combines ANNs and genetic algorithm features has been developed. The model has been calibrated and validated using real data of pumped water. A genetic algorithm was used to find the optimal neural network settings to explain the maximum water demand variance with minimal estimation error.

The developed model was applied to predict water demand one-day ahead in the BMD Irrigation District, Southern Spain. The model has predicted 93 % of the variability of the observed water demand with a standard error of 12.63 %. Thus, the results show that this hybrid methodology improves the accuracy of the predictions of previous models with smaller errors. The addition of data from new irrigation seasons will improve the accuracy even in peak demand days.

This model can be useful for irrigation districts managers for different objectives. If the irrigation district has an electricity tariff indexed to pool where the energy price varies each day and each

hourly scheduling period, the developed model will allow managers to design strategies to distribute the water consumption throughout the day and therefore minimize energy costs. On the other hand, if the irrigation district has a superindexed electric tariff where energy consumption is paid in advance, the developed model can be used to establish a daily irrigation schedule that approaches the estimated electricity consumption. Thus, the developed model proved to be a powerful tool that, without long dataset and time requirements, can be very useful for the development of management strategies.

Acknowledgments. This research was supported by an FPU grant (Formación de Profesorado Universitario) from the Spanish Ministry of Education, Culture and Sports to Rafael González Perea. This work is part of the TEMAER project (AGL2014-59747-C2-2-R), funded by the Spanish Ministry of Economy and Competitiveness.

6.5. References

Abrahart RJ, See L (2000) Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. Hydrol Process. 14:2157–2172

Abrahart RJ, See L (2002) Multi-model data fusion for river flow forecasting: An evaluation of six alternative methods

based on two contrasting catchments. Hydrol Earth Syst Sci 6(4):655–670

Abrahart RJ, See L, Kneale PE (1999) Using pruning algorithms and genetic algorithms to optimise network architectures and forecasting inputs in a neural network rainfall-runoff model. J Hydroinf 1(2):103–114

Agarwal A, Mishra SK, Ram S, Singh JK (2006) Simulation of Runoff and Sediment Yield using Artificial Neural Networks. Biosyst Eng 94(4):597–613

Allen RG, Pereira LS, Raes D, Smith M (1998) Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage, Rome, Italy

Anctil F, Rat A (2005) Evaluation of neural network streamflow forecasting on 47 watersheds. J Hydrol Eng 10(1):85–88

Battiti R (1992) First- and Second-Order Methods for Learning: Between Steepest Descent and Newton's Method. Neural Comput 4(2):141–166

Brent RP (1973) Algorithms for Minimization Without Derivatives. Prentice-Hall, Englewood Cliffs, New Jersey

Cameron D, Kneale P, See L (2002) An evaluation of a traditional and a neural net modelling approach to flood forecasting for an upland catchment. Hydrol Process 16:1033–1046

Charalambous C (1992) Conjugate gradient algorithm for efficient training of artificial neural networks. IEE Proc 139(3):301–310

Chiang Y-M, Chang L-C, Chang F-J (2004) Comparison of static-feedforward and dynamic-feedback neural networks for rainfall–runoff modeling. J Hydrol 290:297–311

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6:182–197

Demuth HB, Hagan MT, Beale MH (2009) Neural Network ToolboxTM 6. User's Guide. The MathWorks Inc., Natick, Massachusetts, USA

Dennis JE, Schnabel R.B. (1983) Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall, Englewood Cliffs, New Jersey

Doorenbos J, Kassam AH (1979) Yield response to water. FAO Irrigation and Drainage Paper 33, FAO, Rome

Doorenbos J, Pruitt WO (1977) Guidelines for Predicting Crop Water Requirements. Irrigation and Drainage Paper 24, Rome

Fletcher R, Reeves CM (1964) Function minimization by conjugate gradients. Comput J 7:149–154

French MN, Krajewski WF, Cuykendall RR (1992) Rainfall forecasting in space and time using a neural network. J Hydrol 137:1–31

Hagan MT, Menhaj MB (1994) Training feedforward networks with the Marquardt algorithm. IEEE Trans Neural Netw 5(6):989–93

Hagan MT, Demuth HB, Beale MH (1996) Neural network design. PWS Publishing Co., Boston, Massachusetts

Hsu K, Gupta HV, Sorooshian S (1995) Artificial Neural Network Modeling of the Rainfall-Runoff Process. Water Resour Res 31(10):2517–2530

INE Instituto Nacional de Estadística (2014) Encuesta sobre el uso del agua en el sector agrario (año 2012). Nota de prensa. Madrid. Spain. http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft26%2Fp067%2Fp01&file=inebase&L=0. Accessed 1 Jan 2015

Kuligowski RJ, Barros AP (1998) Experiments in short-term precipitation forecasting using artificial neural networks. Mon Weather Rev 126(2):470–482

Lorrai M, Sechi GM (1995) Neural nets for modelling rainfall-runoff transformations. Water Resour Manag 9:299–313

MAGRAMA Spanish Ministry of Agriculture, Food and Environment (2013) Boletín Hidrológico Semanal 1.

Madrid. Spain. http://www.magrama.gob.es/es/agua/temas/evaluacion-de-los-recursos-hidricos/boletin-hidrologico/

Mason JC, Tem'me A, Price RK (1996) A neural network model of rainfall-runoff using radial basis functions. J Hydraul Res 34(4):537–548

Møller MF (1993) A scaled conjugate gradient algorithm for fast supervised learning. Neural Networks 6:525–533

Moradkhani H, Hsu K, Gupta H V., Sorooshian S (2004) Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. J Hydrol 295:246–262

Nayebi M, Khalili D, Amin S, Zand-Parsa S (2006) Daily Stream Flow Prediction Capability of Artificial Neural Networks as influenced by Minimum Air Temperature Data. Biosyst Eng 95(4):557–567

Ondimu S, Murase H (2007) Reservoir Level Forecasting using Neural Networks: Lake Naivasha. Biosyst Eng 96(1):135–138

Playán E, Mateos L (2006) Modernization and optimization of irrigation systems to increase water productivity. Agric Water Manag 80:100–116

Poff NL, Tokar S, Johnson P (1996) Stream hydrological and ecological responses to climate change assessed with an

artificial neural network. Limnol Oceanogr 41(5):857–863

Powell MJD (1977) Restart procedures for the conjugate gradient method. Math Program 12:241–254

Pratap R (2010) Getting started with Matlab. A quick introduction for scientist and engineers. Oxford University Press, USA

Pulido-Calvo I, Gutiérrez-Estrada JC (2009) Improved irrigation water demand forecasting using a soft-computing hybrid model. Biosyst Eng 102(1):202–218

Riedmiller M, Braun H (1993) Direct adaptive method for faster backpropagation learning: The RPROP algorithm. In: 1993 IEEE International Conference on Neural Networks. Publ by IEEE, 586–591

Rogers LL, Dowla FU (1994) Optimization of groundwater remediation using artificial neural networks with parallel solute transport modeling. Water Resour Res 30(2):457–481

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(9):533–536

Scales LE (1985) Introduction to non-linear optimization. New York : Springer-Verlag

See L, Openshaw S (2000) A hybrid multi-model approach to river level forecasting. Hydrol Sci J 45(4):523–536

Shin H-S, Salas JD (2000) Regional Drought Analysis Based on Neural Networks. J Hydrol Eng 5(2):145–155

Thirumalaiah K, Deo MC (1998) River stage forecasting using artificial neural networks. J Hydrol Eng 3(1):26–31

Thirumalaiah K, Deo MC (2000) Hydrological Forecasting Using Neural Networks. J Hydrol Eng 5(2):180–189

Tokar AS, Johnson PA (1999) Rainfall-Runoff Modeling Using Artificial Neural Networks. J Hydrol Eng 4(3):232–239

Tokar AS, Markus M (2000) Precipitation–runoff modeling using artificial neural networks and conceptual models. J Hydrol Eng 5(2):156–161

Van Aelst P, Ragab RA, Feyen J, Raes D (1988) Improving irrigation management by modelling the irrigation schedule. Agric Water Manag 13:113–125

Ventura S, Silva M, Pérez-Bendito D, Hervás C (1995) Artificial neural networks for estimation of kinetic analytical parameters. Anal Chem 67(9):1521–1525

Yang CC, Prasher SO, Lacroix R, et al (1997) Artificial neural network model for subsurface-drained farmlands. J Irrig Drain Eng 123(4):285–292

Zhang M, Fulcher J, Scofield RA (1997) Rainfall estimation using artificial neural network group. Neurocomputing 16:97–115

7. Prediction of Irrigation Event Occurrence at Farm Level using Optimal Decision Trees

This chapter is currently under review in the journal "Water Resources Management", González Perea R, Camacho Poyato E, Montesinos P, Rodríguez Díaz JA (2017)

Abstract. Irrigation water demand is highly variable and depends on farmers' decision about when to irrigate. Their decision affects the performance of the irrigation networks. An accurate daily prediction of irrigation events occurrence at farm scale is a key factor to improve the management of the irrigation districts and consequently the sustainability of the irrigated agriculture. In this work, a hybrid heuristic methodology that combines Decision Trees and Genetic Algorithm has been developed to find the optimal decision tree to model farmer's behaviour, predicting the occurrence of irrigation events. The methodology has been tested in a real irrigation district and results showed that the optimal models developed have been able to predict between 68 % and 100 % of the positive irrigation events and between 93 % and 100 % of the negative irrigation events.

Keywords. Artificial Intelligence, Multiobjective genetic algorithm, Irrigation scheduling

7.1. Introduction

Factors such as climate change, world population growth or the competition for water resources make freshwater availability a large and complex global challenge, mainly in those regions where rainfall is scarce and irregular. This is the case of Spain which nowadays devotes 73 % of its national freshwater to irrigate 3.65M ha (INE 2016). The expansion of irrigated land coupled with tourism and urbanization has created significant water supply pitfalls (García-Ruiz et al. 2011).

Therefore, improving water use efficiency is a key to maintaining the sustainability of the irrigated agriculture. Related to this, water demand forecasting could be one of the main tools to improve the management of the irrigation districts and help managers in the decision-making processes. Previous research works focused on the prediction of water demand at irrigation district level, using neuro-genetic algorithms (Pulido-Calvo and Gutiérrez-Estrada 2009; González Perea et al. 2015). However, forecasting water demand at individual farmer level is an extremely complex task.

Apparently, and mainly in on-demand irrigation networks where water is continuously available to farmers, water demand is highly variable and apparently follows a pure random process. However, it depends on climatic factors such as evapotranspiration, climate

and on other social and economic factors like local farmers' practices, crops value or energy prices. Due to the high number of factors that intervene in the irrigation scheduling, the prediction of water demand is complex and the water demand forecasting models must consider several variables at the same time. As irrigation scheduling is the process of deciding when and how much to irrigate, the prediction of the occurrence of irrigation events should be the first step to build a robust water demand-forecasting model.

Nowadays, the new telemetry and data acquisition systems provide new possibilities that were not available in the past. Usually the irrigation districts continuously collect lots of information aimed at billing and rarely for improving the decision-making processes. However, big-data and artificial intelligence techniques are the right tools to integrate all these datasets, and extract useful information for managers and give an additional value and usefulness to systems installed in the field.

Decision Tree methods, DT, have been widely used in machine learning, expert systems, and multivariate analysis. These methods are probably the most highly developed techniques for partitioning sample data into a collection of decision rules (Jang et al. 1997). However, these DT procedures have been applied to several fields of agriculture (Loureiro et al. 2016; Zhang et al.

2017), operating rules for reservoirs (Kumar et al. 2013) or urban water distribution systems (Loureiro et al. 2016) but no previous works have been developed in the field of irrigation forecasting.

The main limitation of DTs is the determination of either the best algorithm to find the best split of the predictive model and the best decision tree architecture or the cross-validation process. In most works, these variables are determined by trial and error so the achievement of an optimal solution is not warrantied. Aiming at overcoming this limitation, in this work a Genetic Algorithm (GA) has been used to optimize the different parameters that make up the Decision Tree. Therefore, a new methodology combining Decision Tress and Genetic Algorithm has been developed to model farmer's behaviour and forecast the occurrence of irrigation events. The non-sorting multi-objective genetic algorithm, NSGA-II (Deb et al. 2002) has been used as GA and the predictive model has been implemented in MATLAB (Pratap 2010). This methodology has been applied to a real irrigation district in Spain.

7.2. Methodology

7.2.1. Study area and data source

The predictive model was developed and tested in *Canal del Zujar Irrigation District* (CZID), located in southwest of Spain. CZID is

made up of ten independent hydraulic sectors and convers a total irrigated area of 21,141 ha. Among the ten sectorial networks, Sector II was selected for this study. This sector covers an irrigated area of 2,691 ha being the main crops tomato, maize and rice (90 % of the total irrigated area). Drip irrigation is the irrigation method used in tomato and maize crops while rice is flood irrigated.

A telemetry system operates in Sector II of the CZID to record hourly water consumption by means of flowmeters installed at hydrant level. This information is transmitted to the central offices using mobile communication technology. The water consumption records were aggregated at daily level in each hydrant for the 2015 irrigation season. In addition, at each hydrant, information about the crop type and the farm size were also recorded. The daily climatic data of maximum and average temperature (°C), average relative humidity (%), precipitation event (this input takes a value of 1 when occurs an event of precipitation), were obtained from the agroclimatic station placed in the irrigation sector. Thus, daily irrigation and daily climatic date from 1st January 2015 until 31st December 2015 were used in this work. Julian day, weekday and bank holidays were additional input recorded for the same period.

7.2.2. Problem approach

The development of a farmer's behaviour model that predicts the daily occurrence of irrigation events, the first step in the irrigation scheduling process, is addressed in this work by decision trees and genetic algorithms. The occurrence of irrigation events is a binary decision and so the prediction model is converted into a binary classification problem (irrigation or non irrigation events). On the other hand, the main parameters that make up the architecture and the training process of a decision trees which are generally fixed by trial and error, in this work, they are optimized by the multiobjective genetic algorithm NSGA-II.

7.2.2.1. Decision Trees

A decision tree is a tree structure composed of internal and external nodes connected by branches, which divides the input set into mutually exclusive regions (i.e. the Julian day may divide the decision trees into several branches). Each of these regions is assigned a label, a value, or an action that characterizes its data points. The internal nodes, known as decision-making unit, assess a decision function to determine which child node to visit next. The nodes, which are associated with the labels (e.g. irrigation or non irrigation events) that characterize the given data, are known as leaves or terminal nodes (external nodes) and they have no

child. The classical structure of a typical decision tree is shown in Fig. 7.1.

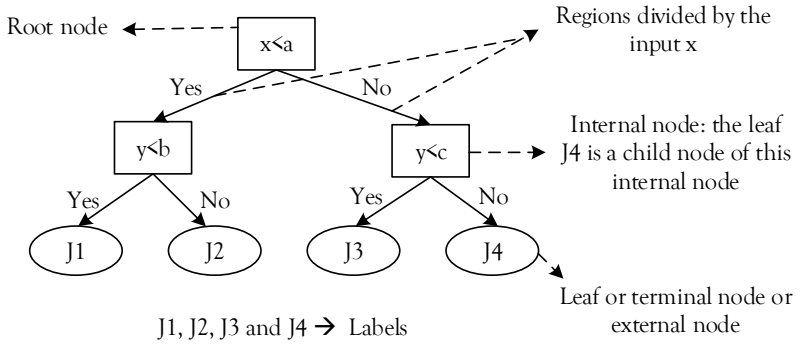


Fig. 7.1. Classical structure of a decision tree.

There are two main types of decision trees: regression and classification. The terminal node labels in regression trees are constants or equations that specify the forecasted output value of a given input vector. However, the leaf nodes of the decision trees in classification trees contain a label that indicates the group or class (*J*) (e.g. irrigation or non irrigation events) to which a given feature vector belongs. The farmer's decision about to irrigate or not is a classification problem with two classes (J=2), so a classification tree is used in this work. Thus, initially the vector composed of several attributes (the inputs set such as daily average temperature, Julian day or precipitation occurrence) is presented to the first node (root node) of the classification tree. Then, the branching pattern is defined according to the decision function

used. The offshoot continues until a leaf node is reached and a label (irrigation or non irrigation) is assigned to the given input data.

The split algorithm of the classification tree determines the ramification process. There are several algorithms with different search procedures to find the best split on a categorical prediction (classification) during the training process. In this work, the four most widely used split algorithms have been considered. The first one, known as *Exact* (Breiman et al. 1993), considers all $2^{CatLevel-1}$ – 1 combinations of tree splitting, being *CatLevel* the number of categories or levels of a classification tree. The second algorithm used is named *Pull Left By Purity* ('*PLbyPurity*') (Breiman et al. 1993). It starts with all *CatLevel* on the right branch. Then, the algorithm moves each category to the left branch to achieve the minimum measure of error for the J-classes among the remaining levels. Finally, from this sequence, the algorithm chooses the data division that has the lowest measure of error. *Principal Component-Based Partitioning* (*PCA*) (Coppersmith et al. 1999) is the third algorithm. This algorithm finds a close-to-optimal binary partition of the *CatLevel* levels by searching for a separating hyperplane that is perpendicular to the first principal component of the weighted covariance matrix of the centred class probability matrix. The last split algorithm considered is *One Versus All By Class* (*OVAC*) (Breiman et al. 15993). It starts with all *CatLevel* levels on the right

branch and for each of the J-classes, the algorithm orders the categories based on the probability of each class.

An optimization procedure, in this case a genetic algorithm, is required to choose the best split algorithm for each data set. The selected split algorithm will affect the accuracy of the classification tree that will be tested with new data set (test set) in the testing process.

7.2.2.2. Error measurement and Decision Tree improvement

An error measure, *E(t)*, that computes the performance of a node *t* in separating data from different classes is necessary to grow the classification tree. The error function is generally referred as the impurity function. This function tends to zero when all data belong to the same class. By contrast, the impurity function takes the maximum value when the data are uniformly distributed through all classes.

In this work, the *Gini diversity index*, $\emptyset_g$, (Breiman et al. 1984) has been used as impurity function and it is defined for J-classes as follow:

$$\emptyset_g(p_1, \dots, p_J) = \sum_{i \neq j} p_i p_j = 1 - \sum_{j=1}^{J} p_j^2 \qquad [7.1]$$

where p_1, ..., p_J are the probability that a case or data in a node belongs to class *j*.

The accuracy of the classification tree can be improved using cross validation. Cross validation randomly divides the training data into *k* parts. Then, *k* new trees are trained and the accuracy of the forecasting process is validated with the *k*-1 data set not included in each k training. Finally, the last training (k^{th}) gives the trained classification tree. However, this optimization technique is quite time consuming. Therefore, a balance between training time and model accuracy should be reached. Thus, in this work, the use or not of cross validation is one of the variables included in the optimization process.

7.2.2.3. Control of the Classification Tree depth

A balance between simplicity and forecasting power must be considered during the development of a classification tree. A classification tree with many leaves is often highly accurate during the training process. However, a deep tree tends to overfit in the training process and the model accuracy decreases significantly, when new data are presented (testing process). In contrast, shadow trees do not achieve the highest accuracy during the training process, which can be achieved in the testing process. To avoid overfitting during de training process, the depth of the classification trees must be controlled. There are four classification tree characteristics to control the classification tree depth. The first one is to fix the maximum number of branch

node splits per tree (*MaxBranch*), which is high in deep trees. Another characteristic is the minimum number of observations (e.g. days in the cropping season) per leaf (*MinObsLeaf*). Thus, the lower *MinObsLeaf* value, the deeper tree. The third characteristic is the minimum number of observations per branch node (*MinObsBranch*). This number is inversely proportional to the depth of the classification tree. The last classification tree characteristic to control the tree depth is the maximum categories or levels of the classification tree (*MaxCatLevel*). A large value of *MaxCatLevel* may increase the computation time and memory overload. However, a small value can cause a poor model accuracy. All these characteristics are often set up by default by the computer model used or they are stablished by trial and error. However, these values could compromise the accuracy and robustness of the predictive model. In this work, these four characteristics are variables considered in the optimization process.

7.2.2.4. Classification Tree Optimization with GA

Classification trees can stablish the different regions that the input space may be classified according to the assigned labels. However, there are several tree's characteristics that are often chosen by trial and error causing a loss of efficiency, robustness or precision of the predictive model. Here, these tree's

characteristics are automatically determined and optimized by the multiobjetive GA, NSGA-II (Deb et al. 2002). Thus, the optimization of the classification tree has been raised as a two-objective optimization problem. The first objective function, *F1* maximizes the farmer's decisions properly classified according to the Eq. 7.2:

$$F1 = 1 - \frac{\sum_{i=1}^{Numlass_{test}} Misclass_{test,i}}{Numclass_{test}} \quad [7.2]$$

where $Misclass_{test,i}$ equals 1 when the farmer's decision i is misclassified with respect to the observed farmer's decision within the test set and $Numclass_{test}$ is the total number of observations of the test set, i.e., the total number of farmer's decisions included in the test set.

The second objective function, *F2* (Eq. 7.3), minimizes the number of nodes that make up the classification tree. Thus, both computing time and the depth of the tree are minimized to maximize the efficiency and accuracy of the predictive model.

$$F2 = numNodes \quad [7.3]$$

where *numNodes* is the total number of nodes of the classification tree.

NSGA-II algorithm starts the optimization process by the random

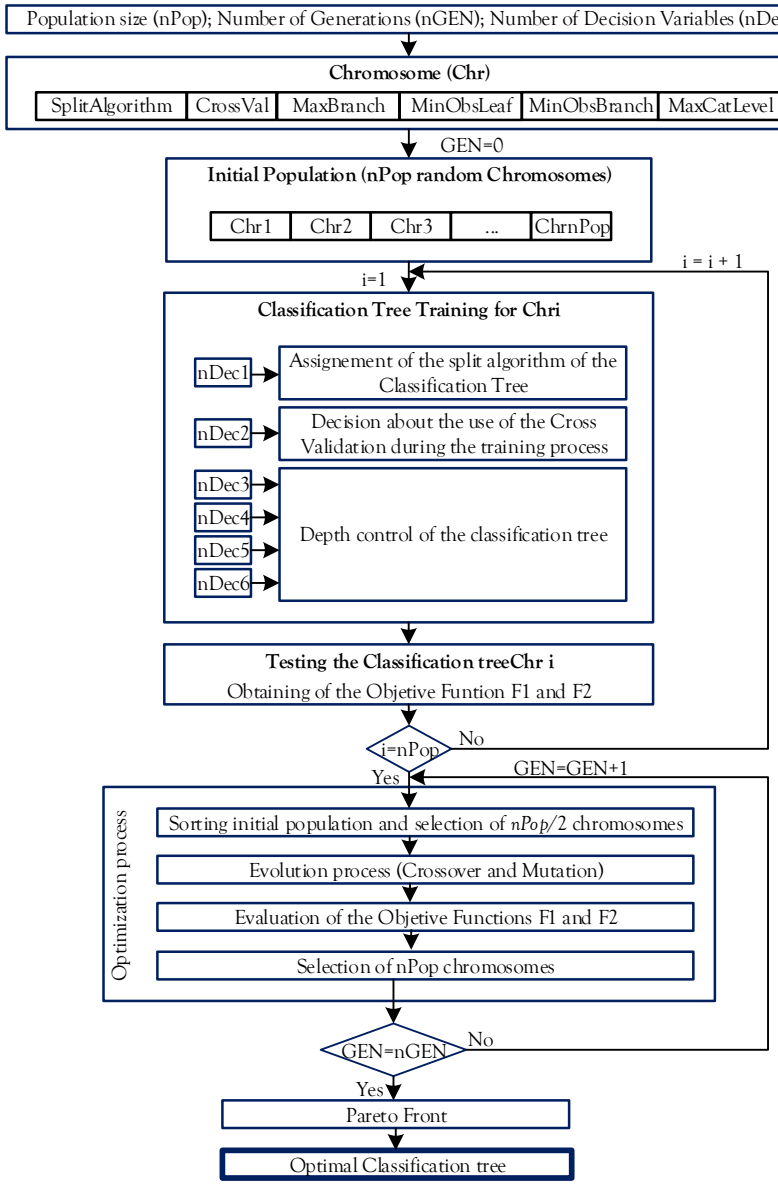


Fig. 7.2. Flow chart of the optimization process of the classification tree.

generation of an initial population of *nPop* size. Each chromosome of the initial population is composed of 6 genes or

decision variables (the split algorithm of the classification tree, the use or not of Cross Validation and the four parameters needed to control the tree depth). Once the initial population is created, a classification tree (*Classification tree*$_{Chr\ i}$) is generated and trained from each chromosome. After this, every *Classification tree*$_{Chr\ i}$ is tested by the test set and the objective functions *F1* and *F2* are calculated.

Then, the chromosomes are modified (crossover and mutation) and the fittest *nPop* decision trees are selected based on their objective function values. The process is repeated several generations (*nGEN*). Finally, the set of *nPop* optimal chromosomes obtained in the last generation define the Pareto Front. Table 7.1 shows the decision variables and its position (Gene) within the chromosome (*Chr*) as well as the ranges of values associated to each decision variables. A brief description of each decision variable is also shown in Table 7.1. Frequently, *MaxBranch* equals 5 but this parameter depends on the input data and so it is difficult to fix it in advance. Thus, in this work, *MaxBranch* varies within a wider range, from 5 to 20. Similarly, *MinObsLeaf*, *MinObsBranch* and *MaxCatLevel* ranged from 20 to 100.

Table 7.1. Decision variables of the NSGA-II GA.

Decision Variable	Gene	Range of values*	Description
SplitAlgorithm	1	1 to 4	These genes determine the split algorithm of the classification model: 1: Exact Algorithm (*Exact*). 2: Pull Left By Purity Algorithm (*PLbyPurity*). 3: Principal Component-Based Partitioning Algorithm (*PCA*). 4: One Versus All By Class Algorithm (*OVAC*).
CrossVal	2	0 to 1	Decision about the use of Cross Validation: 0: No Cross Validation. 1: Cross validation.
MaxBranch	3	5 to 20	Maximum number of branch node splits per tree.
MinObsLeaf	4	20 to 100	Minimum number of observations per leaf.
MinObsBranch	5	20 to 100	Minimum number of observations per branch node.
MaxCatLevel	6	20 to 100	Maximum categorical levels of the tree.

*Integer values between the Range of values

7.3. Results and Discussion

The total number of hydrants of CZID Sector II is 649. After data processing, 627 hydrants were selected to develop the model of the daily farmer's behaviour model (those hydrants without relevant information were removed from the analysis). Information about daily hydrant operation (whether farmer irrigates (1) o not (0)) each day of 2015 (365 days) was stored.

Thus, a total number of 228,855 observations (irrigation events) were used in the analysis. From this data set, 183,084 observations were randomly selected for the training process (80 % of the total observations) and the remaining 20 % (45,771 observations) were considered as test set.

7.3.1. Model inputs

8 input variables (Table 7.2) have been considered to define farmers' behaviour within the predictive model. These variables are related to the irrigation process in the study area. The farmer's

Table 7.2. Input variables of the classification tree.

Input	Description
I_1	Crop.
I_2	Julian day.
I_3	Bank holiday (false or true).
I_4	Weekday.
I_5	Daily maximum temperature (°C).
I_6	Daily average temperature (°C).
I_7	Daily average relative humidity (%).
I_8	Precipitation event (false or true).

decision about when it is necessary to apply water is related to the crop type (I1) and the Julian day (I2) that determines the phenological state of the plant and therefore its sensibility to the water stress. Bank holidays (I3) and the weekday (I4) are factors linked to social aspects of the study region. Daily maximum

temperature (I5), daily average temperature (I6) and daily average relative humidity (I7) are variables related to the farmer's warming sensation that also condition his/her decision. Finally, precipitation events (I8) has also been taken into, taking a false value when the precipitation in a day is null and vice versa.

7.3.2. The Pareto Front of the optimization process

Initially the trees' structure was optimized using the multiobjetive NSGA-II. 100 individuals (chromosomes) made up the random initial population that evolved for 500 generations evaluating F1 and F2. 90 % and 10 % were the probabilities considered for crossover and mutation.

The Pareto front (generation 500) obtained in the optimization process and the computing time requirements in the training process are shown in Fig. 7.3a and 7.3b. The Pareto font shows that objectives F1 and F2 are clearly conflicting. Thus, the higher depth of the classification trees (higher number of nodes), the higher the accuracy of the predictive model but the higher computing time requirements (Fig. 7.3b). The fastest classification tree (*CT3*), the most accurate classification tree (*CT1*) and one with intermediate results (*CT2*) were selected to analyse the results. The (*CT1*) correctly predicted 100 % of the irrigation events in the test set requiring 13 % more of computing time than (*CT3*) that could correctly forecast 99.16 % of the

irrigation events in the test set. The results obtained by *CT2* ranged between *CT1* and *CT3*.

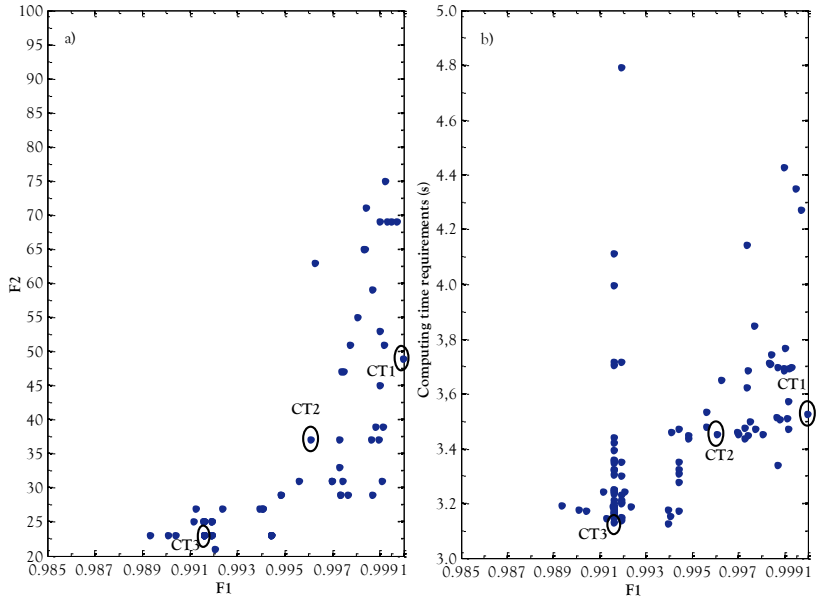


Fig. 7.3. a) Pareto front for generation 500; b) computing time requirements in training process.

All individuals of the Pareto front were trained with cross validation. Thus, the optimization process confirms that cross validation is essential to achieve the best results.

7.3.3. Optimal Classification Trees

The setting parameters (genes) and objective function values of the three selected configurations are shown in Table 7.3. *PCA* algorithm was used by *CT1* as split algorithm in the training process while *OVAC* and *PLbyPurity* were the split algorithms for

CT2 and *CT3*, respectively. The *Exact* algorithm, which was commonly used by other authors in this sort of analysis, was not selected in any classification tree. The configurations obtained in the Pareto front highlights that the optimum number of categorical levels of the classification tree is small. The variation range for *MaxCatLevel* was from 5 to 20 and the optimal *MaxCatLevel* for *CT1* and *CT2* was 5 and *CT3* had 9 maximum categorical levels. *MaxBranch* was similar for *CT1* (36) and *CT2* (35) and slightly lower for *CT3*(21). The results obtained for both *MinObsLeaf* and *MinObsBranch* parameters show that a minimum number of observations are necessary either by leaf or by branch. When the minimum number of observations per leaf is low, the minimum number of observation per branch is high. Thus, the values of *MinObsLeaf* and *MinObsBranch* were 20 and 78, respectively, for *CT1* while these parameters were 82 and 20, respectively, for *CT2* and for *CT3* they were 31 and 41, respectively.

Table 7.3. Setting parameters (genes) and objective function values of *CT1*, *CT2* and *CT3*.

Group	CT1	CT2	CT3
SplitAlgorithm	PCA	OVAC	PLbyPurity
CrossVal	yes	yes	yes
MaxBranch	36	35	21
MinObsLeaf	20	82	31
MinObsBranch	78	20	41
MaxCatLevel	5	5	9
F1	1	0.9961	0.9916
F2	49	37	23

With the aim to visualize the performance of each classification tree, Table 7.4a, 7.4b and 7.4c shows the confusion matrix for *CT1*, *CT2* and *CT3*, respectively. Confusion matrix is a specific table layout where each column of the matrix represents the instances in a predicted class while each row represents the instances in an actual class. Thus, *CT1* was the most accurate model with the 100 % of the successful irrigation events. Therefore, the number of false positive and false negative were 0. The global precision of *CT2* was 99.61 %. However, this index is sometime misleading and the dimensions of the classes should be analysed. The test set contained 9,250 positive irrigation events, in other words, the farmer decided 9,250 times to apply irrigation water. *CT2* classified as no irrigation 2,544 positive irrigation events (false negative) and 2,689 negative irrigation events as positive irrigation events (false positive). Hence, although the global accuracy of the classification model was 99.61 %, *CT2* rightly classified the 73 % of the positive irrigation events and the 93 % of the total negative irrigation events. Similarly, *CT3* classified the 32 % of the positive irrigation events as false negative and only the 7 % was classified as false positive. 80 % of the total irrigation events were negative irrigation events in both training and test data sets, consequently, *CT2* and *CT3* learnt better to forecast negative than positive irrigation events.

Table 7.4. Confusion matrix for *CT1* (a), *CT2* (b) and *CT3*(c).

a)

		Predicted values		Accuracy	
		Irrigation	No Irrigation	Positive (Irrigation)	Negative (No Irrigation)
Observed values	Irrigation	9,250	0	100 %	100 %
	No Irrigation	0	36,521		

b)

		Predicted values		Accuracy	
		Irrigation	No Irrigation	Positive (Irrigation)	Negative (No Irrigation)
Observed values	Irrigation	6,706	2,544	73 %	93 %
	No Irrigation	2,689	33,823		

c)

		Predicted values		Accuracy	
		Irrigation	No Irrigation	Positive (Irrigation)	Negative (No Irrigation)
Observed values	Irrigation	6,329	2,921	68 %	93 %
	No Irrigation	2,537	33,984		

This limitation of the classification tree could be overcome adding more observations to the training set or limiting the training set to the days of the effective irrigation season (set days with positive irrigation events), avoiding days without irrigation.

A scheme of the classification trees of *CT1*, *CT2* and *CT3* are shown in Fig. 7.4a, 7.4b and 7.4c, respectively. The highest tree depth was *CT1* with the highest number of nodes. The three figures show that the input variables I3 (Bank holiday), I7 (daily average humidity) and I8 (Precipitation occurrence) were not included in the classification trees. Therefore, despite of the input variables selected in this work were previously discussed with farmers, the predictive model shows that these three variables are not significant to forecast the occurrence of daily irrigation events. Probably, the input variables I2 (Julian day) and the weekday (I4) implicitly include the input variable I3 (Bank holiday). Similarly, the maximum and average temperature are connected to the relative humidity and the classification tree found this relationship. Finally, although I8 should be important in irrigation scheduling, the classification tree did not link it with the irrigation decision. It may be explained by the lack of rainfall in the central months of the year (typical in the Mediterranean climate), when most of the irrigation events occur so this variable becomes irrelevant.

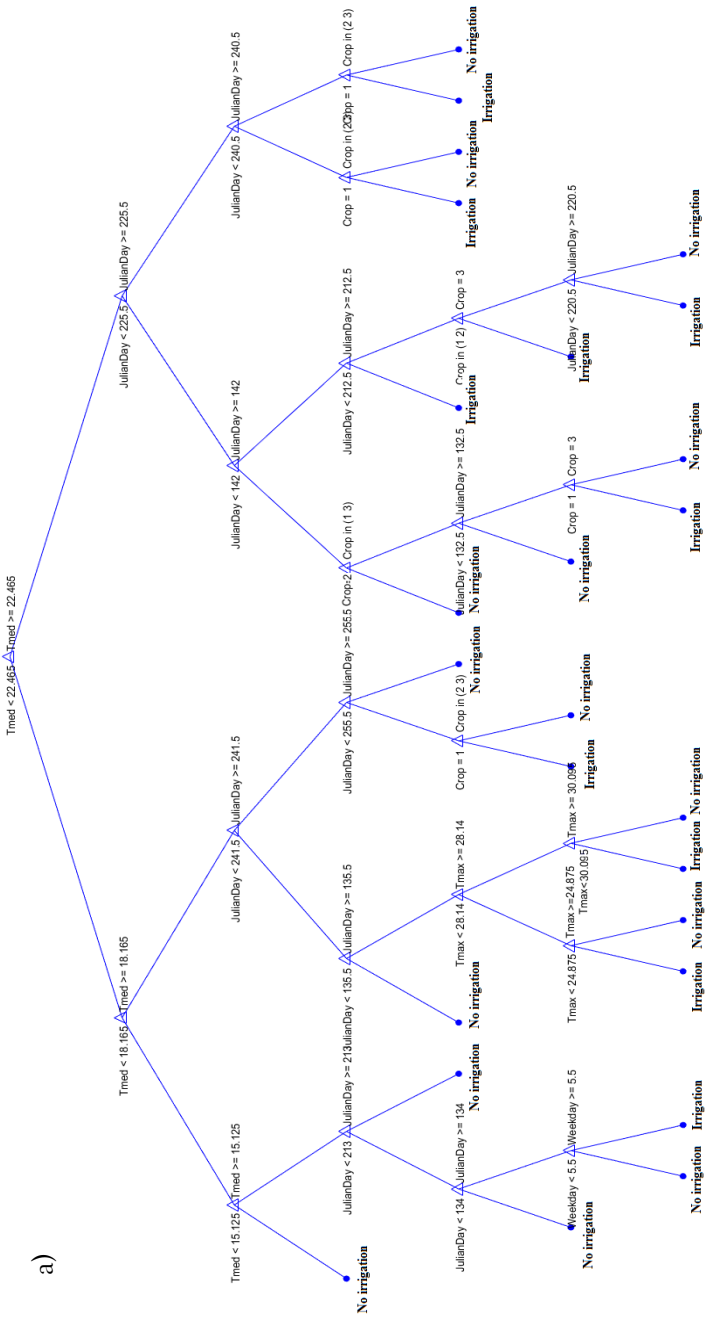


Fig. 7.4. Schemes of the classification trees a) *CT1*; b), *CT2* and c) *CT3*.

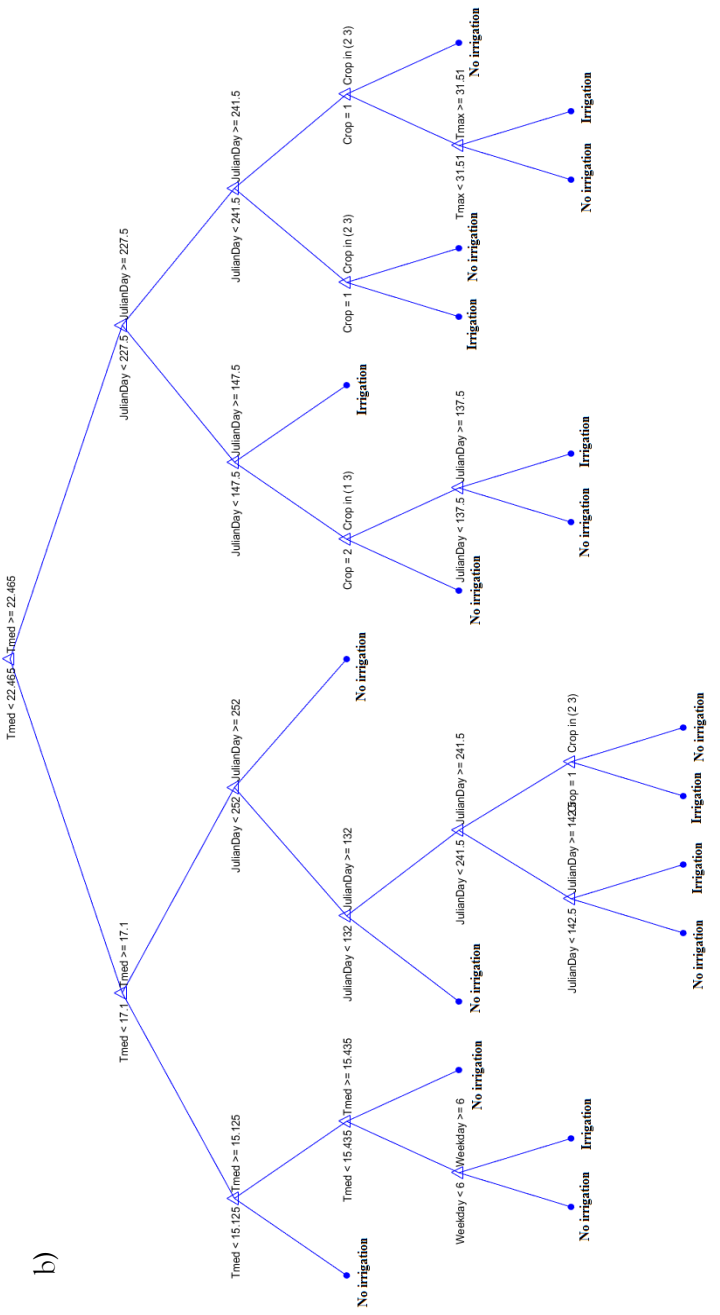


Fig. 7.4. Schemes of the classification trees a) *CT1*; b), *CT2* and c) *CT3*. (Continuation)

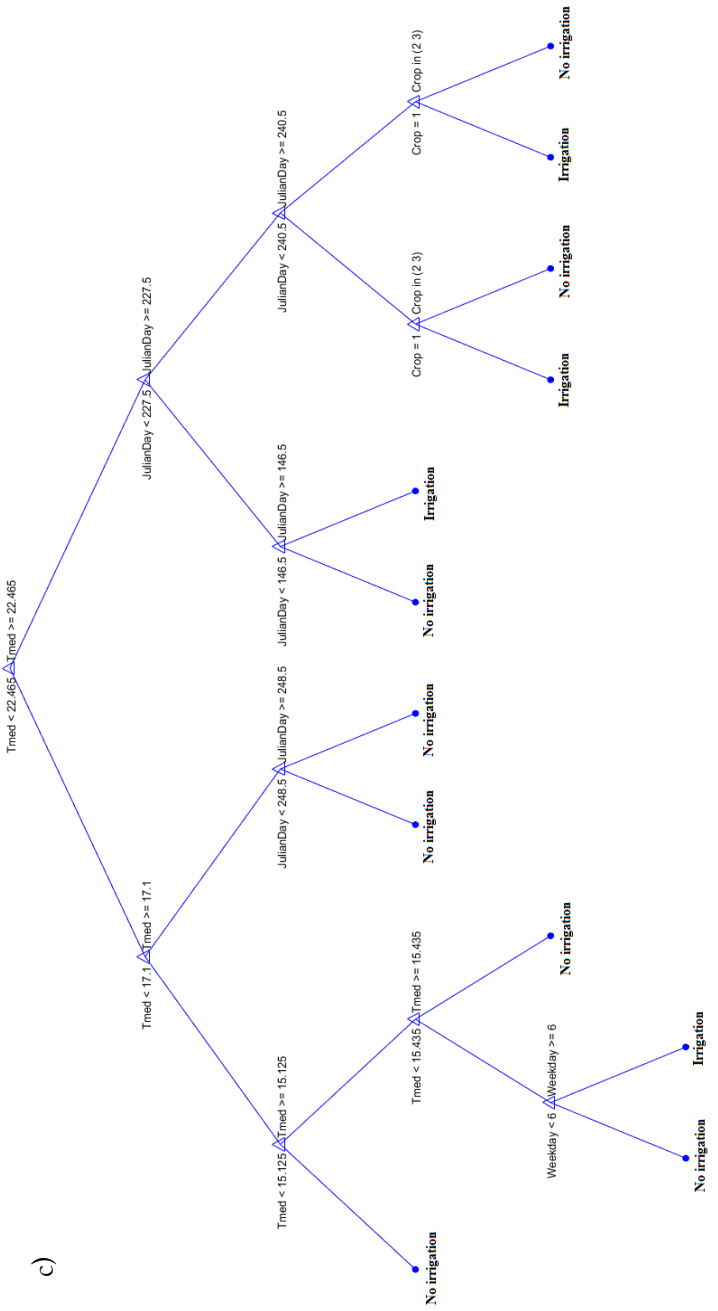


Fig. 7.4. Schemes of the classification trees a) *CT1*; b), *CT2* and c) *CT3*. (Continuation)

7.4. Conclusions

Irrigated agriculture needs new strategies and tools to improve water use efficiency. Modelling farmers' behaviour would help irrigation district managers to achieve an integrated management of the irrigation district based on the prediction of the occurrence of irrigation events, that will aid to set the optimum operational point of the pumping station as well as to hire the most convenient electrical tariff on futures markets.

In this work, decision trees were successfully used as classification models to forecast when farmers irrigate. The use of optimal decision trees obtained from a multi-objective genetic algorithm provides successful predictions of when farmers irrigate when they are applied to a real case study.

In the case study, the optimal classification models predicted properly between 99.16 % and 100 % of the given data test set. This global index of model accuracy can sometime be misleading and an assessment of the accuracy of each class should be analysed. The classification models predicted between 68 % and 100 % of the positive irrigation events and between 93 % and 100 % of the negative irrigation events.

This work represents the first step in the prediction of the irrigation scheduling at farm level, defined by when, how much

and for how long to irrigate. So, it focuses on the prediction of when irrigation events occur. Future research should be devoted to the development of models to fully predict the irrigation scheduling at farm level, including irrigation depths and timing.

Acknowledgments. This research was supported by an FPU grant (Formación de Profesorado Universitario) from the Spanish Ministry of Education, Culture and Sports to Rafael González Perea. This work is part of the TEMAER project (AGL2014-59747-C2-2-R), funded by the Spanish Ministry of Economy and Competitiveness.

7.5. References

Breiman L, Friedman JH, Olshen RA, Stone CJ (1993) Classification and Regression Trees, Chapman &. Boca Raton

Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and regression trees. Wadsworth, Inc., Belmont, California

Coppersmith D, Hong SJ, Hosking JRM (1999) Partitioning Nominal Attributes in Decision Trees. Data Min Knowl Discov 3:197–217

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6:182–197

García-Ruiz JM, López-Moreno JI, Vicente-Serrano SM, et al (2011) Mediterranean water resources in a global change scenario. Earth-Science Rev 105:121–139

González Perea R, Poyato EC, Montesinos P, Díaz J A. R (2015) Irrigation Demand Forecasting Using Artificial Neuro-Genetic Networks. Water Resour Manag 29:5551–5567.

INE IN de E (2016) Encuesta sobre el uso del agua en el sector agrario (año 2014). Madrid. Spain. http://www.ine.es/dyngs/INEbase/es/categoria.htm?c=Estadistica_P&cid=1254735976602. Accessed 4 Jun 2017

Jang J-SR, Sun C-T, Mizutani E (1997) Neuro-Fuzzy And Soft Computing Jang. Prentice-Hall, USA

Kumar ARS, Goyal MK, Ojha CSP, et al (2013) Application of ANN, Fuzzy Logic and Decision Tree Algorithms for the Development of Reservoir Operating Rules. Water Resour Manag. doi: 10.1007/s11269-012-0225-8

Loureiro D, Mamade A, Cabral M, et al (2016) A Comprehensive Approach for Spatial and Temporal Water Demand Profiling to Improve Management in Network Areas

Pratap R (2010) Getting started with Matlab. A quick introduction for scientist and engineers. Oxford University Press, USA

Pulido-Calvo I, Gutiérrez-Estrada JC (2009) Improved irrigation water demand forecasting using a soft-computing hybrid model. Biosyst Eng 102:202–218

Zhang H, Wang D, Gartung JL (2017) Influence of irrigation scheduling using thermometry on peach treewater status and yield under different irrigation systems

8. Farmer's Behaviour Modelling by the Prediction of the Applied Irrigation Depth using Artificial Intelligence

This chapter is currently under review in the journal "Water Resources Research", González Perea R, Camacho Poyato E, Montesinos P, Rodríguez Díaz JA (2017)

Abstract. Irrigation water demands are highly variable and depends on the behaviour of each farmer affecting the performance of the irrigation networks. The farmer's behaviour is influenced by precise variables and uncertain or imprecise variables that conditions the applied irrigation depth. The prediction of this farmer´s behaviour is essential for a right management of the irrigation districts and the design of the news irrigation networks. Hence, in this work a hybrid methodology combining Artificial Neural Networks, Fuzzy Logic and Genetic Algorithm has been developed with aim to modelling the farmer's behaviour and forecast the daily irrigation depth used by each farmer. The developed models have been tested in a real Irrigation District, located in Southwest of Spain. Three optimal models for the main crops that makes up the irrigation district have been achieved. The representability (R^2) and the accuracy of the predictions (Standard Error Prediction, SEP) were 0.72, 0.87

and 0.72; and 22.20 %, 9.80 % and 23.42 %, for the rice, maize and tomato crop models, respectively.

Keywords. Irrigation scheduling, Prediction, ANFIS, Genetic Algorithm

8.1. Introduction

Climate change and the growing water demand of some economic sectors such as industry or agriculture are reducing freshwater availability. Irrigated agriculture is the main water user, accounting for nearly 85 % of the total water consumption in the world (Jury and Vaux, 2007). The sustainability of the irrigated agriculture is strongly linked to the improvement of water use efficiency. Water demand forecasting could be one of the main tools to design accurately new irrigation systems and improve the management of older pressurized irrigation networks. Irrigation water demands are highly variable. They depend on the behaviour of each farmer that is affected by both measurable variables (e.g. agroclimatic variables or the size of irrigated area) and non-measurable variables (e.g. local traditional practices or the days of leave during the irrigation season).

Fuzzy Logic (FL) is an Artificial Intelligence (AI) technique initially developed by (Zadeh 1965) to explain the human thinking and decision system. FL can be applied as a Fuzzy

Inference System (FIS) designed to transform linguistic concepts into mathematical and computational structures for daily water demand forecasting. FIS is a rule-based system that consists of a rule base, a database with membership functions (MFs) which determine the membership grades of each input variable to each fuzzy set and the combination of fuzzy rules produces the system results (inference system). However, the FIS have two major limitations. The first restriction is to set the type of membership functions and their optimal number. In most works, these variables are determined by trial and error so finding an optimal solution is not guaranteed. Thus, one of the most popular approaches to overcome this constraint is the use of genetic fuzzy systems (GFSs), a hybrid combination of FL and Genetic Algorithms (GAs). GFSs have been already used for water demand forecasting at irrigation districts level (Pulido-Calvo and Gutiérrez-Estrada, 2009) but there is no previous work that uses GFSs to predict the farmer's behaviour. The second restriction is the inability of FIS to select automatically the MF parameters and design the fuzzy rules. However, the combination of Artificial Neural Networks (ANNs) and FL, known as Adaptive Neuro Fuzzy Inference System (ANFIS), overcomes this drawback. Thus, an ANFIS uses the learning ability of the ANN to define fuzzy rules. ANFIS has been used for several applications such as the intelligent allocation of water resources (Chang et al., 2016) or

the optimization of the reservoir operation (Safavi et al., 2013). Although this technique has not been applied yet to characterize farmer's behaviour. Therefore, in this work a hybrid methodology that combines GFSs and ANFIS has been developed to forecast the daily amount of water applied by each farmer. The non-sorting genetic algorithm, NSGA-II (Deb et al., 2002), is the multi-objective GA included in the GFS developed in this work. This methodology has been applied to a real irrigation district in Spain to predict farmers' behaviour during 2015 irrigation season.

8.2. Methodology

8.2.1. Study area and data source

The data recorded in *Canal del Zujar Irrigation District* (CZID) (in southwest of Spain) have been the base to develop and test the predictive model built in this work. CZID is made up of ten independent hydraulic sectors and covers a total irrigated area of 21,141 ha. Sector II was selected for this study. This sector covers an irrigated area of 2,691 ha where the main crops are tomato, maize, grapevine and rice.

The Sector II of the CZID has a telemetry system with flowmeters that records hourly flowrates at hydrants level. For 2015 irrigation season, hourly records were aggregated at daily level. In addition, information about crop types and sizes of the farms watered from

each hydrant was also available. The daily climatic data were obtained from the weather station placed in the irrigation sector.

8.2.2. Problem approach

Irrigation scheduling process consists of two main steps: occurrence of the irrigation event and the amount of water applied. In this work, a farmer's behaviour model that forecast the daily irrigation depth applied by each farmer is developed using GFS and ANFIS. Thus, the first phase of the model building process has been the identification of the main input variables. Then, a FIS is designed using an ANFIS model which is optimized by the NSGA-II GA.

8.2.3. Model Inputs Identification

Although the construction of forecasting methods requires huge amount of data, the first step in this process is reducing the dimension of the input space to identify the relevant input variables within the whole dataset. There are several techniques to do this, such as principal components analysis or partial least square cardinal components. However, when the selected variables are used in nonlinear models, model predictions are usually quite poor (Lin et al., 1996). Therefore, in this work, fuzzy curves and fuzzy surfaces have been used to easily select the independent significant inputs for the hybrid model according to

the methodology developed by Lin et al. (1996). Thus, they automatically identify the independent significant inputs for applying them in the model. Initially, for each potential input variable, a plot is created relating every potential input variable to target variable to be predicted (*DaW*, daily amount of water applied by each farmer). Then, for each point represented in each plot a fuzzy membership function is created according to the following expression:

$$\mu_{v,k}(PI_v) = \exp(-(\frac{PI_{v,k} - PI_v}{bb})^2) \qquad [8.1]$$

where $\mu_{v,k}$ represents the fuzzy membership function of the point *k* in the plot which relates the potential input variable *v* and the daily amount of water applied by each farmer; PI_v is the potential input variable *v*; $PI_{v,k}$ is the value of the PI_v in the point *k* and *bb* takes a value close to two (Lin et al., 1996).

Hereafter, each fuzzy membership function is defuzzied producing a fuzzy curve c_v for each potential input PI_v using:

$$c_v(PI_v) = \frac{\sum_{k=1}^{M} \mathrm{DaW}_k \cdot \mu_{v,k}(PI_v)}{\sum_{k=1}^{M} \mu_{v,k}(PI_v)} \qquad [8.2]$$

where M is the total number of points in the space PI_v – *DaW* and DaW_k is the daily amount of water applied by each farmer in the point *k* of the space PI_v – *DaW*.

Then, the Mean Square Error (*MSE*) is computed for each space $PI_v - c_v$:

$$MSE_{c_v} = \frac{1}{M}\sum_{k=1}^{M}(c_v \cdot (PI_{v,k}) - \mathrm{DaW}_k)^2 \qquad [8.3]$$

where MSE_{c_v} is the mean square error for the fuzzy curve c_v.

MSE values of each c_v are sorted in ascending order. If there is a completely random relationship between the PI and the daily amount of water applied by each farmer, the fuzzy curve is flat and MSE_c is large. On the contrary, if MSE_c value is small the relationship between PI and the daily amount of water applied by each farmer is more significant.

A fuzzy surface is a space with two-dimensional fuzzy curve. According to Lin et al. (1996) a fuzzy surface ($fs_{v,j}$) is defined as Eq. 8.4.

$$fs_{v,j}(PI_v, PI_j) = \frac{\sum_{k=1}^{M} \mathrm{DaW}_k \cdot \mu_{v,k}(PI_v) \cdot \mu_{j,k}(PI_j)}{\sum_{k=1}^{M} \mu_{v,k}(PI_v) \cdot \mu_{j,k}(PI_j)} \qquad [8.4]$$

where PI_v and PI_j are two potential input variables.

Then, similarly to Eq. 8.3 the *MSE* is computed for the fuzzy surfaces:

$$MSE_{fs_{v,j}} = \frac{1}{M}\sum_{k=1}^{M}(fs_{v,j}(PI_v, PI_j) - \mathrm{DaW}_k)^2 \qquad [8.5]$$

Fuzzy curves are initially used to rank all the potential input variables in ascending order. The potential input variable with the smallest MSE_c is the most important input variable. According to Lin et al. (1996), 20 % of the potential input variables with largest MSE_c are eliminated. Then, fuzzy surfaces are used to find the independent input variables and to eliminate the related input at each step. Thus, in each step new fuzzy surfaces are computed and 20 % of the potential input variable with largest MSE_{fs} are eliminated.

8.2.4. Fuzzy Inference System (FIS)

Due to their unique features in forecasting complex phenomena, FIS is one of the best tools for modelling human thinking (e.g. farmers' decisions). A fuzzy system is a nonlinear relationship between inputs and outputs based on a set of "IF-THEN" rules. While the antecedent of a rule defines a fuzzy region in the input space (e.g. crop, maximum daily temperature, weekday, ...), the consequent specifies the output in a fuzzy region. Fig. 8.1 shows a flow chart of a typical three step - FIS. The aim of the first step (Fuzzification) is to transfer the input vector into fuzzy If-Then rules through the MFs and linguistic variables, i.e. a vector with input variables (crisp values) is turned into linguistic variables (e.g.

the value of the variable temperature is 25 °C (crisp value) becomes to linguistic variable 'the temperature is HIGH'). The rule base and the MFs form the knowledge base (Fig. 8.1). Then, the optimal design of the knowledge base is established by an ANN (Section 8.2.5).

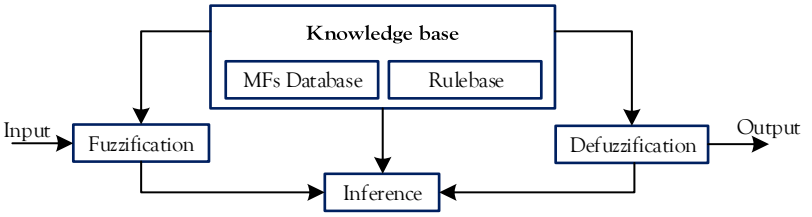


Fig. 8.1. Structure of a Fuzzy Inference System (FIS).

There are two types of FISs, Sugeno-Takagi (TS) FIS and Mamdani FIS, which main differences are the way that the outputs (Fig. 8.1) that are determined. TS FIS, due to its more compact and computationally efficient representation than Mamdani FIS, is selected for modelling the farmer's behaviour. The rule's consequent in TS FIS can be either a linear equation, called 'first -order TS FIS' or a constant coefficient, called 'zero-order TS FIS'. Due to the complex modelling of the farmer's behaviour, first-order TS FIS has been selected. Two examples of typical TS FIS rules are:

Rule 1: If x is A_1 *and y is* B_1 *Then* $f_1 = p_1 \cdot x + q_1 \cdot y + r_1$,
Rule 2: If x is A_2 *and x is* B_2 *Then* $f_2 = p_2 \cdot x + q_2 \cdot y + r_2$,

where x and y are inputs; A, B are linguistic variables; f is the consequence of each rule, and p, q and r, are parameters which will be determined by a ANN in the following section.

The inference step (step 2) uses these fuzzy If-Then rules to assign a map from fuzzy inputs to fuzzy outputs based on fuzzy composition rules (Li, 2006) (e.g. for maize, when temperature is HIGH (fuzzy input) the amount of water applied is HIGH (fuzzy output). The last step (Defuzzification) transfers fuzzy sets into crisp values (e.g. for maize, if temperature is HIGH then the amount of water applied is HIGH, fuzzy value, and so the applied irrigation depth is 15 mm, crisp value).

8.2.5. ANFIS

The main drawback to build up a FIS is the lack of systematic procedures to define both MFs parameters and the rule base. Nevertheless, the ANNs can learn its structure from the input-output sets. Thus, in this work, an Adaptive Neuro Fuzzy Inference System, ANFIS, resulting from the combination of an ANN and a FIS has been created to determine the MFs parameters and to find the rule base through the ANN learning ability to set the relationship between input and output. Then, resulting fuzzy rules depend on the input structure. The mathematical background about the ANFIS learning process is detailed in Jang et al. (1997). The typical structure of an ANFIS

is composed by five layers (Fig. 8.2) where the adaptive nodes represents different values depending on the input variables and the fixed nodes develop the same function independently of the input variables (e.g. product function) (in Fig. 8.2 squares are adaptive nodes and circles are fixed nodes). The number of nodes in the different layers depends on the number of rules considered. Fig. 8.2 is a simple example of a 5 layer-ANFIS with two inputs. Its operation is described next.

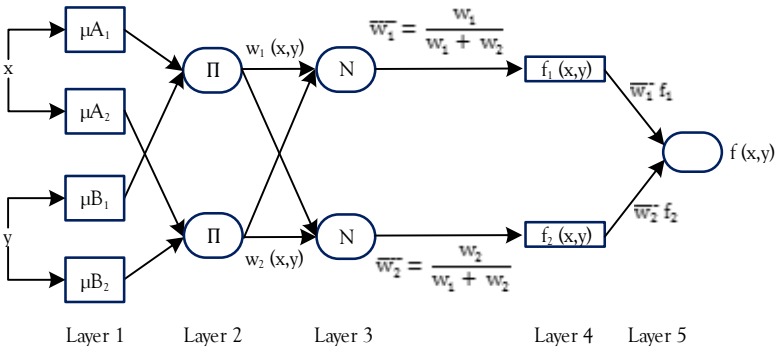


Fig. 8.2. Structure of an Adaptive Neural-Fuzzy Inference System (ANFIS).

The nodes of the first layer (fuzzification layer) use the MFs to get outputs (O_i^j) (i and j are the node and layer indexes respectively) which are calculated according to Eqs. 8.6 and 8.7.

$$O_i^1 = \mu_{Ai}(x) \; for \; i = 1,2 \qquad [8.6]$$

$$O_i^1 = \mu_{Bi}(y) \; for \; i = 3,4 \qquad [8.7]$$

where μ_{Ai} and μ_{Bi} are the membership functions whose optimal typology will be determined by a GA.

The layer 2 is composed of rule nodes that calculate the firing strength of each rule, w_i, as follows:

$$O_i^2 = w_i = \mu_{Ai}(x) \cdot \mu_{Bi}(y) \quad for\ i = 1, 2 \qquad [8.8]$$

The average nodes constitute the third layer 3 aimed at the calculation of the ratio of the i^{th} node firing strength to the sum of all firing strengths. This ration is computed according to the following equation:

$$O_i^3 = \bar{w}_i = \frac{w_i}{w_1 + w_2} \quad for\ i = 1, 2 \qquad [8.9]$$

The following layer (layer 4) is the defuzzification layer that consist of the consequent nodes. The output of each node represents the contribution of i^{th} rule to the output model and is computed as follow:

$$O_i^4 = \bar{w}_i \cdot f_i = \bar{w}_i \cdot (p_i x + q_i y + r_i) \quad for\ i = 1, 2 \qquad [8.10]$$

In this work, the optimal values of p, q and r are obtained using two possible learning algorithms. Thus, the learning method is a decision variable of the GA, that can choose either the backpropagation method (Hagan et al., 1996) which is used for all parameters of the MFs or the hybrid method (backpropagation method + least squares method) where backpropagation method

is used for the parameters associated with the input MFs, and least squares method is used for the parameters associated with the output MFs Backpropagation method is a steepest descent method which the parameters of the MFs are moved along the negative of the gradient of the performance function while least squares method uses the classical least squares problems resolution.

The final layer (layer 5) consists of a single fixed node, the output node. This layer computes the overall output as the summation of all incoming signals from the previous layer by the following equation:

$$O_i^5 = f(x,y) = \sum_i \overline{w}_i \cdot f_i = \frac{\sum_i w_i \cdot f_i}{\sum_i w_i} \quad for\ i = 1,2 \quad [8.11]$$

8.2.6. Optimal ANFIS

ANFIS can determine the rule base and the MF parameters which build the knowledge base of a FIS. However, the number of MFs which divides the universe of discourse of every input, the type of MFs that characterizes every linguistic label and the learning method must be previously defined. In most works, these variables are determined by trial and error. On the contrary, in this work, the optimal values of these variables are automatically determined by the multiobjetive GA NSGA-II (Deb et al., 2002). Thus, two objective functions, *F1* and *F2*, are defined to search

the values of the ANFIS parameters that better forecast the farmer's behaviour. While *F1* maximizes the determination coefficient of the testing process (R^2_{test}), *F2* minimizes the average normalized root mean square error in the same testing process ($RMSE_{test}$). The Standard Error Prediction (*SEP*) is also computed to compare different behavioural models. According to Ventura et al. (1995) *SEP* is calculated as follow:

$$SEP = \frac{100}{\bar{V}} RMSE_{test} \qquad [8.12]$$

where $RMSE_{test}$ is the *RMSE* in the test process (mm) and $\bar{V}$ is the average of observed daily water demand of the test set.

Fig. 8.3 shows the flow chart of the optimization process of the ANFIS model. Firstly, an initial population of *nPop* size is randomly generated by NSGA-II algorithm. Each chromosome of the initial population represents an ANFIS model and consists of *nInput* genes which divides the discourse universe of every input variable, *nInput* genes which define the number and type of MFs of every input variable and an additional gen for the learning method. Therefore, the size of each chromosome is $2 \cdot nInput+1$. After the initial population is created, for each chromosome an ANFIS model is generated ($FIS_{chr\ i}$) and the knowledge base is obtained according to the Eqs. 8.6 to 8.11. Every $FIS_{chr\ i}$ is trained with a data subset (training set) which were randomly obtained of

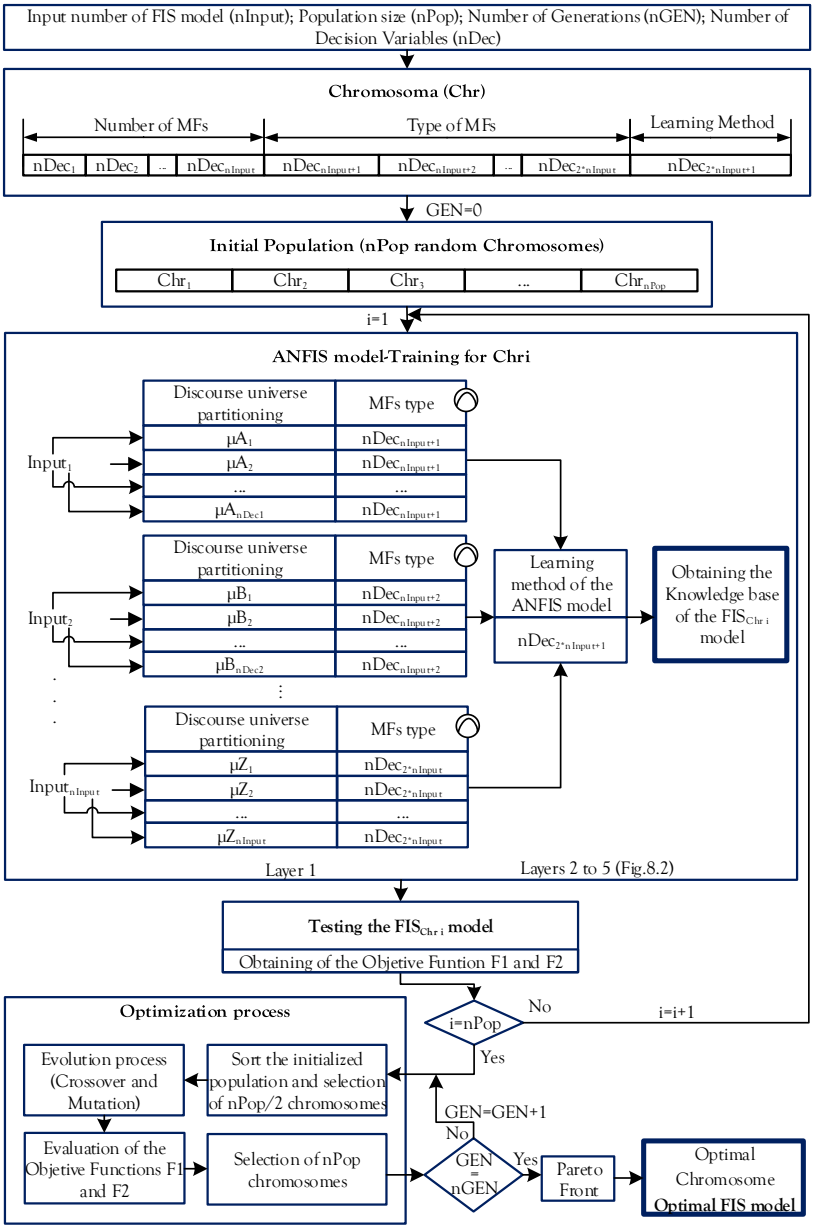


Fig. 8.3. Flow chart of the optimization process of the FIS model.

the total data set. Hereafter, with the aim to assess the reliability of the training process every $FIS_{chr\,i}$ is tested by the test set and the objective functions *F1* and *F2* are calculated.

Then, according to their objective function values the chromosomes are selected and modified (crossover and mutation) to generate a new set of *nPop* chromosomes. The process is repeated several generations (*nGEN*). Finally, the set of *nPop* optimal chromosomes obtained in the last generation defines the Pareto Front.

Table 8.1. Decision variables of the NSGA-II GA.

Decision Variable	Gene position	Range values**	Description
Number of MFs	From 1 to *nInput**	1 to 10	These genes determine the discourse universe division of each input variable.
Type of MFs	From *nInput* +1 to 2 *nInput*	1 to 8	These genes determine the shape of the MFs of each input variable: 1: Triangular MF (*trimf*). 2: Trapezoidal MF (*trapmf*). 3: Gaussian MF (*gaussmf*). 4: Asymmetric Gaussian MF (*gauss2mf*). 5: Generalized bell MF (*gbellmf*). 6: Difference between two sigmoidal MF (*dsigmf*). 7: Product of two sigmoidal MF (*psigmf*). 8: Polynomial MF (*pimf*).
Learning method	2 nInput +1	0 to 1	This gene determines the learning method of the ANFIS: 0: Backpropagation method. 1: Backpropagation method + least squares method.

*nInput: number of the input variables of the predictive model.
**Integer values between the Range of values.

Table 8.1 shows the decision variables and their positions (Gene) in the chromosome (Chr) of the GA as well as the value ranges that every decision variables can take for the nInput input variables of the predictive model. A brief description of each decision variable is also shown in Table 8.1.

8.3. Results and Discussion

8.3.1. Model inputs

The methodology developed above has been applied to *CZID* described in Section 8.2.1, during the 2015 irrigation season with the aim to predict the daily irrigation depth applied by each farmer from the following 18 potential inputs (PI) selected according to the methodology developed by Lin et al. (1996) which are shown in Table 8.2. As these authors suggest, 20 % of the input variables are reduced in each step of the identification process. Thus, a summary of the identification process of the model input variables is shown in Table 8.3. Finally, the following 5 inputs variables were identified: Applied irrigation depth in the previous day, mm (PI_{17}); Applied irrigation depth in the two previous days, mm (PI_{18}); Julian day (PI_{13}); Daily maximum relative humidity, % (PI_6); Daily average temperature, °C (PI_4).

Table 8.2. Potential model inputs.

Potential Input	Description
PI_1	Farm area (ha).
PI_2	Crop.
PI_3	Daily maximum temperature (°C).
PI_4	Daily average temperature (°C).
PI_5	Daily minimum temperature (°C).
PI_6	Daily maximum relative humidity (%).
PI_7	Daily minimum relative humidity (%).
PI_8	Daily average relative humidity (%).
PI_9	Daily maximum wind speed (m s-1).
PI_{10}	Daily average wind speed (m s-1).
PI_{11}	Daily rainfall (mm).
PI_{12}	Daily rainfall boolean (mm). This variable equals 0 when the daily precipitation is null. Otherwise, this variable is 1.
PI_{13}	Julian day.
PI_{14}	Weekday.
PI_{15}	Month.
PI_{16}	Boolean Holidays. This variable equals 1 for bank holidays and vacation days. Otherwise, this variable is 0.
PI_{17}	Applied irrigation depth in the previous day, mm.
PI_{18}	Applied irrigation depth in the two previous days, mm.

The volume of water that each farmer applies every day depends on, essentially, the applied irrigation depth in the previous and two previous days, which is considered by the variables PI_{17} and PI_{18}, respectively. There is direct relationship between the applied irrigation depth, crop type and its phenological stage (that varies along the irrigation season) which is considered by the variable PI_{13}. In addition, the farmer's decision about the amount of water to apply is frequently conditioned by the farmer's warming sensation and it is considered by variables P_{I6} and P_{I4}. Thus, the identified variables defined farmers' behaviour.

Table 8.3. Potential model inputs.

Steps	Remaining inputs	Significance ranked by ascending MSE	Identified input	Eliminated inputs
1	All inputs, PI_1 to PI_{18}	PI_{17}, PI_{18}, PI_1, PI_{13}, PI_{15}, PI_7, PI_3, PI_8, PI_6, PI_{14}, PI_9, PI_4, PI_{10}, PI_{11}, PI_5, I_{12}, PI_2	PI_{17}	PI_5, PI_{12}, PI_2
2	PI_{18}, PI_1, PI_{13}, PI_{15}, PI_7, PI_3, PI_8, PI_6, PI_{14}, PI_9, PI_4, PI_{10}, PI_{11}	PI_{18}, PI_{13}, PI_8, PI_{14}, PI_6, PI_7, PI_3, PI_4, PI_9, PI_{15}, PI_1, PI_{10}, PI_{11}	PI_{18}	PI_1, PI_{10}, PI_{11}
3	PI_{13}, PI_8, PI_{14}, PI_6, PI_7, PI_3, PI_4, PI_9, PI_{15}, PI_1	PI_{13}, PI_{14}, PI_4, PI_8, PI_7, PI_6, PI_3, PI_9, PI_{15}	PI_{13}	PI_9, PI_{15}, PI_3
4	PI_{14}, PI_4, PI_8, PI_7, PI_6	PI_6, PI_4, PI_7, PI_8, PI_{14}	PI_6	PI_8, PI_{14}
5	PI_4, PI_7	PI_4, PI_7	PI_4	PI_7

8.3.2. ANFIS optimization

The NSGA-II optimized the objective functions *F1* and *F2*. The random initial population consisted of 100 individuals (chromosomes) that evolved 40 generations. In this case study, each chromosome was composed by 11 ($2 \cdot nInput+1$) genes because of 5 input variables were considered.

Fig. 8.4 shows the Pareto front obtained in the 40^{th} generation of the optimization process. This figure shows two groups of solutions which are clearly distinct from one another. This difference is related to the last gene of the chromosome (learning method). The first group was trained with the backpropagation method obtaining a single value, 96.13 mm, for RMSE (*F2*).

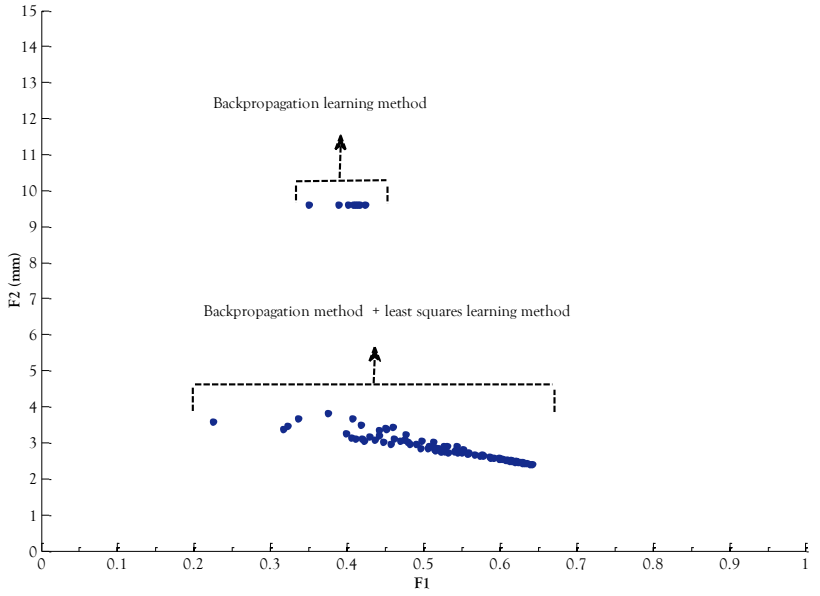


Fig. 8.4. Pareto front for generation 40.

Alternatively, the second group, trained by the backpropagation and least squares hybrid method, reached *F2* values between 2.41 mm and 3.84 mm. Furthermore, the *F1* values (R^2_{test}) ranged from 0.35 to 0.42 for the first group while the *F1* values for the second group ranged from 0.22 to 0.64. Therefore, the Pareto front highlights that the hybrid method is the best learning method for this case study.

The Pareto front shown in Fig. 8.4 was obtained by the GA (40th generation) using the training and testing sets which contain information about all hydrants independently of the crop associated with each hydrant. However, the irrigation method and the daily water demand for each crop is related with the

farmer's behaviour (Fig. 8.5). Fig. 8.5 shows that the irrigation depths applied and the temporal distribution of irrigation events for rice were fairly distinct from those applied to maize and tomato as well as the temporal distribution of irrigation events due to the use different irrigation methods. This fact hinders the forecasting process. Therefore, the accuracy of the model prediction can be improved grouping the training and testing sets according to the irrigation method, i.e., rice (surface irrigation) and maize and tomato (drip irrigation) and optimizing again the ANFIS models, considering the same number of chromosomes in the initial population and generations than for the previous ANFIS optimization.

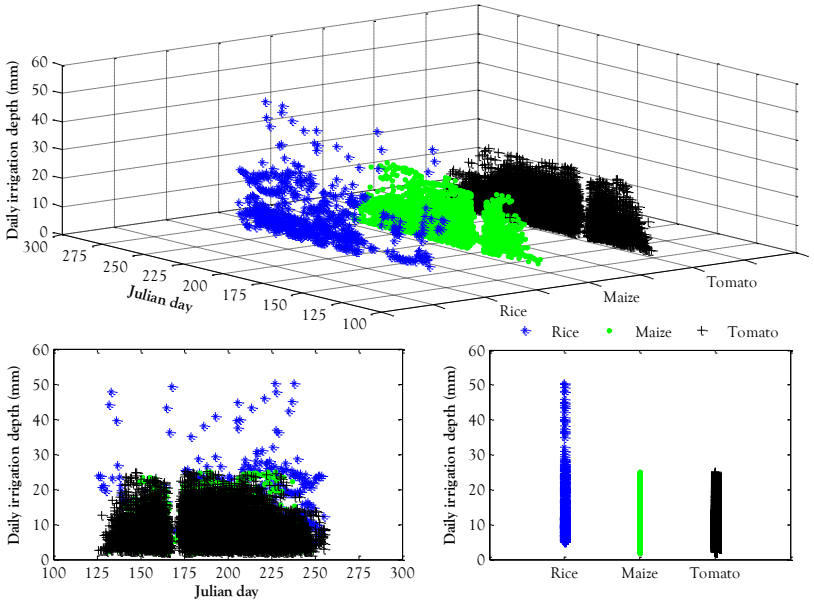


Fig. 8.5. Daily irrigation depth (mm) for each crop during the irrigation season.

Fig. 8.6a and 8.6b show the Pareto fronts for rice and for maize and tomato, respectively, for generation 40. Optimal *F1* and *F2* values were considerably improved with the new training and

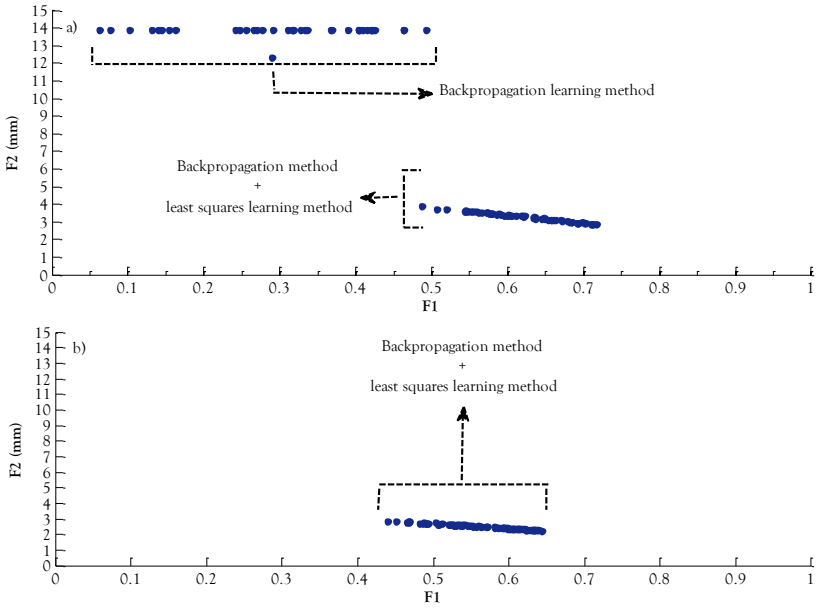


Fig. 8.6. Pareto fronts for generation 40: a) rice; b) maize and tomato.

testing sets. The Pareto front for rice (Fig. 8.6a) shows that there were more individuals with backpropagation learning method than the first optimization process where all crops were trained together. This increment of individuals which were trained with backpropagation learning method is an effect of the change in the size of the training and testing sets. In this case, the GA algorithm did not have enough irrigation records to get better results with this learning method. The *F1* and *F2* values ranged from 0.06 to 0.49 and 12.30 mm to 13.86 mm, respectively, for the training

with backpropagation learning method, and for the alternative method *F1* and *F2* values varied from 0.49 to 0.72 and from 2.85 mm to 3.88 mm, respectively.

The Pareto front for maize and tomato (Fig. 8.6b) shows that all individuals were trained with hybrid learning method. Consequently, although the representativeness of the predictive model (*F1*) did not improve significantly compared to the first model (Fig. 8.4), the accuracy of the predictions was considerably increased. Thus, the *F1* and *F2* values for maize-tomato model ranged from 0.44 to 0.64 and 2.25 mm to 2.82 mm, respectively. Therefore, the accuracy of the maize-tomato model obtained in the 40^{th} generation was 41.40 % better than the model considering the three crops together in a later generation (40).

Fig. 8.5 and Fig. 8.6a and 8.6b highlight that the irrigation method is affects the architecture and the results of the predictive model. Although the irrigation system was similar for maize and tomato, both the cultural practices for each crop and the farmer's behaviour were completely different what would probably have affected the model forecasts. Thus, to overcome this limitation new training, validation and testing sets were created separating maize and tomato, and a new optimization process was carried out considering the same GA parameters (generation number,

size of the initial population, etc.) than the previous optimization processes.

Fig. 8.7a and 8.7b show the Pareto front of maize and tomato, respectively. Like the rice forecast model, the size of the training, validation and testing sets of the maize model (Fig. 8.7a) allowed some individuals trained by backpropagation learning method in the Pareto front. The values of the objective functions evaluated by the GA, *F1* and *F2*, were clearly conditioned by the learning method. If during the random generation of the initial population, individuals were located far away from the optimal regions of the solution space, the GA was not able to shift solutions towards the optimum's surrounding region. This limitation in the generation of the initial population linked to the short size of the training, validation and testing sets resulted in a Pareto front with two clusters. For these reasons, the size of the initial population should be large enough to distribute evenly individuals throughout of solution space. Thus, for the maize model, the *F1* values ranged from 0.27 to 0.38 for backpropagation learning method and from 0.54 to 0.87 for hybrid learning method. The *F2* values ranged from 1.30 mm to 2.44 mm for the hybrid learning method and for backpropagation learning method was 13.68 mm.

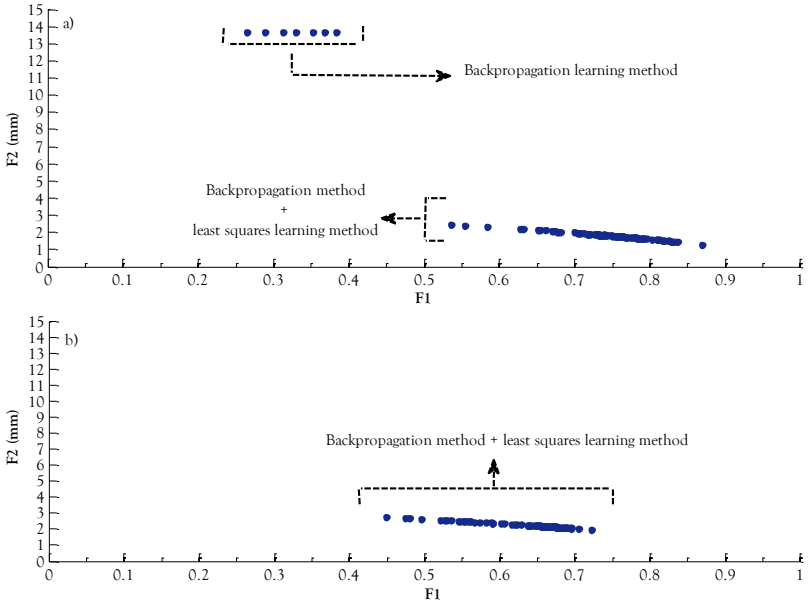


Fig. 8.7. Pareto fronts for generation 40: a) maize; b) tomato.

Fig. 8.7b shows that, for tomato, all individuals were trained with the hybrid learning method. In this case, the tomato forecast model had a 254 % more recorded irrigation events than the maize model (15,372 irrigation events for tomato against to 4,345 irrigation events for maize). For this reason, the GA was able to eliminate individuals with the backpropagation learning method, although during the generation the initial population some of them were located far from the optimal region. Thus, the *F1* and *F2* values for the tomato crop ranged from 0.45 to 0.72 and from 1.97 mm to 2.77 mm, respectively.

Table 8.4 shows the number of observations of the different training validation and testing sets, their optimization time

Table 8.4. Number of observations of the different training, validation and testing sets, their time requirements and the maximum and minimum SEP and R^2 values of the Pareto front for each optimization process.

	Optimization processes	Total recorded irrigation events	Training set	Validation set	Testing set	Time requirements (seconds)	SEP (%)		R^2	
							maximum	minimum	maximum	minimum
1	Rice+Maize+Tomato	20,346	16,276	2,035	2,035	$2{,}3{\cdot}10^5$	110.00	27.52	0.64	0.22
2	Rice	629	501	64	64	$4{,}3{\cdot}10^3$	108.16	22.20	0.72	0.063
	Maize+Tomato	19,717	15,773	1,972	1,972	$8{,}6{\cdot}10^4$	32.85	26.19	0.64	0.44
3	Maize	4,345	3,475	435	435	$2{,}0{\cdot}10^4$	103.61	9.80	0.87	0.27
	Tomato	15,372	12,298	1,537	1,537	$2{,}2{\cdot}10^5$	32.98	23.42	0.72	0.45

requirements and the maximum and minimum SEP and R^2 values of each Pareto front. The 80 % - 10 % - 10 % proportion for training, validation and testing set was always kept in each optimization process. While the maximum and minimum SEP values in the Pareto front of the first optimization process (Rice + Maize + Tomato) were 110 % and 27.52%, respectively, for rice were 108.16 % and 22.20 %, respectively and for maize-tomato models were 32.85 % and 26.19 %, respectively. The optimization time requirements are mainly related to the number of observations and so the time requirements for rice only were 95.3 % lower than for the first optimization process. The segregation of the training, validation and testing sets of the maize-tomato model into independent sets (maize and tomato) improved highly the SEP values. The minimum SEP value for each crop was 9.80 % (maize) and 23.42 % (tomato). However, the optimization time requirements were similar or higher than the maize + tomato and rice + maize + tomato models.

The maximum and minimum values of R^2 were also improved when each crop is independently trained, ranging from 0.062 to 0.72 in the rice model, from 0.27 to 0.87 in the maize model and from 0.45 to 0.72 for the tomato model.

8.3.3. Optimal ANFIS models

The best individuals (ANFIS models) in each optimization process have been selected. Table 8.5 shows the values of the genes that make up every ANFIS model as well as their RMSE, R^2 and SEP values. The number of membership functions of the five input variables are quite small taking into account that the maximum value was fixed by the GA in 10 and neither of them were higher than 4. This fact shows that the limits of the first five genes of the chromosome were rightly set and the universe of discourse of each input variables were rightly portioned. Table 8.6 shows the linguistic variables of each fuzzy set which partitioned each input variable for the best ANFIS models selected. The number of partitions of each universe of discourse depends on the degree of precision required for that variable. The higher number of partitions of an input variable, the higher actions based on this variable that can be carried out in the predictive process, i.e., the predicted variable is more conditioned by changes of this input variable. The irrigation system, the cultural practices and the growth phases of each crop that made up the ANFIS model 1 were completely different. Therefore, the input variable that relates these crop's characteristics (Julian day) is not a conditioning variable and hence it had only a single fuzzy set (the partition of the universe of discourse universe was 1). However, in the ANFIS models 2, 4 and 5 devoted to single crops,

Table 8.5. Gene values and RMSE, R^2 and SEP values of the test set of the best ANFIS models of each optimization process.

	ANFIS models				
	1	**2**	**3**	**4**	**5**
	Rice + Maize + Tomato	**Rice**	**Maize + Tomato**	**Maize**	**Tomato**
Number of MFs of gen 1	2	1	2	2	4
Number of MFs of gen 2	2	1	3	3	3
Number of MFs of gen 3	1	4	3	4	4
Number of MFs of gen 4	3	3	3	2	2
Number of MFs of gen 5	2	3	1	2	2
Type of MFs of gen 1	Product of two sigmoidal	Gaussian	Trapezoidal	Trapezoidal	Gaussian
Type of MFs of gen 2	Polynomial	Asymmetric Gaussian	Gaussian	Trapezoidal	Polynomial
Type of MFs of gen 3	Trapezoidal	Product of two sigmoidal	Triangular	Difference between two sigmoidal	Product of two sigmoidal
Type of MFs of gen 4	Polynomial	Difference between two sigmoidal	Gaussian	Asymmetric Gaussian	Generalized bell
Type of MFs of gen 5	Trapezoidal	Generalized bell	Asymmetric Gaussian	Difference between two sigmoidal	Generalized bell
Learning method	Hybrid method	Hybrid method	Hybrid method	Hybrid method	Hybrid method
RMSE (mm)	2.41	2.85	2.25	1.30	1.97
R^2	0.64	0.72	0.64	0.87	0.72
SEP (%)	27.52	22.20	26.19	9.80	23.42

Table 8.6. Linguistic variables of each fuzzy set to partition each input variable of the selected ANFIS models.

Input variables	ANFIS models				
	1 Rice + Maize + Tomato	2 Rice	3 Maize + Tomato	4 Maize	5 Tomato
Applied irrigation depth in the previous day	HI/LO	Single	HI/LO	HI/LO	VL/LO/HI/VH
Applied irrigation depth in the two previous days	HI/LO	Single	HI/M/LO	HI/M/LO	HI/M/LO
Julian day	Single	VL/LO/HI/VH	HI/M/LO	VL/LO/HI/VH	VL/LO/HI/VH
Daily maximum relative humidity	HI/M/LO	HI/M/LO	HI/M/LO	HI/LO	HI/LO
Daily average temperature	HI/LO	HI/M/LO	Single	HI/LO	HI/LO

VL: Very Low; LO: Low; M: Medium; HI: High; VH: Very High

the crop phenological state was determining to predict the amount of irrigation applied. Consequently, this input variable was divided into four fuzzy sets: Very Low, Low, High and Very High. These partitions match entirely the different phenological states of a crop. Except for the Julian day, the fuzzy sets for the variables of the ANFIS model 2 (rice) were completely different from those of ANFIS models 4 and 5 (maize and tomato, respectively). This highlights that the farmer's behaviour with these crops was completely different. While the daily maximum relative humidity and the daily average temperature take High, Medium or Low values for ANFIS model 2, these variables only take High or Low values in ANFIS models 4 and 5. ANFIS models 4 and 5 had similar fuzzy sets in each input variable. Only the first input variable (applied irrigation depth in the previous day) took High or Low values for ANFIS model 4 and Very Low, Low, High or Very High for ANFIS model 5. Taking into account the five ANFIS models, the input variable which had more influence on farmer's behaviour was the daily maximum relative humidity because of the universe of discourse was always divided into two or three fuzzy sets.

The most frequent membership functions of the input variables were trapezoidal and Gaussian due to the adaptability of their parameters. Because to its shape stiffness, Triangular is just used by the third input variables of ANFIS model 3. The five models

selected were trained under the hybrid learning method (Backpropagation method + least squares method) showing a better efficiency than backpropagation method.

The management of each crop was different even with the same irrigation method. Thus, when two or more crops were used together to train ANFIS models, its representativeness is much lower than if a single crop were used. ANFIS model 1 and ANFIS model 3 were trained with rice+maize+tomato and maize+tomato, respectively and their R^2 values were 0.64 in both models, while when a single crop was used like in ANFIS model 2 (rice), ANFIS model 4 (maize) and ANFIS model 5 (tomato) the R^2 values were 0.72, 0.87 and 0.72, respectively. The accuracy of the prediction was also strongly affected by the crop. Thus, when several crops were considered in the training process the RMSE values ranged from 2.25 mm (SEP value of 26.19 %) for ANFIS model 3 to 2.41 mm (SEP value of 27.52 %) in ANFIS model 1. However, as it is shown in the scatterplots in Fig. 8.8a, 8.8b and 8.8c, when the crops were trained independently the RMSE and SEP values were improved, being 2.85 mm and 22.20 % for ANFIS model 2 (Fig. 8.8a), 1.30 mm and 9.80 % for ANFIS model 4 (Fig. 8.8b) and 1.97 mm and 23.42 % for ANFIS model 5 (Fig. 8.8c).

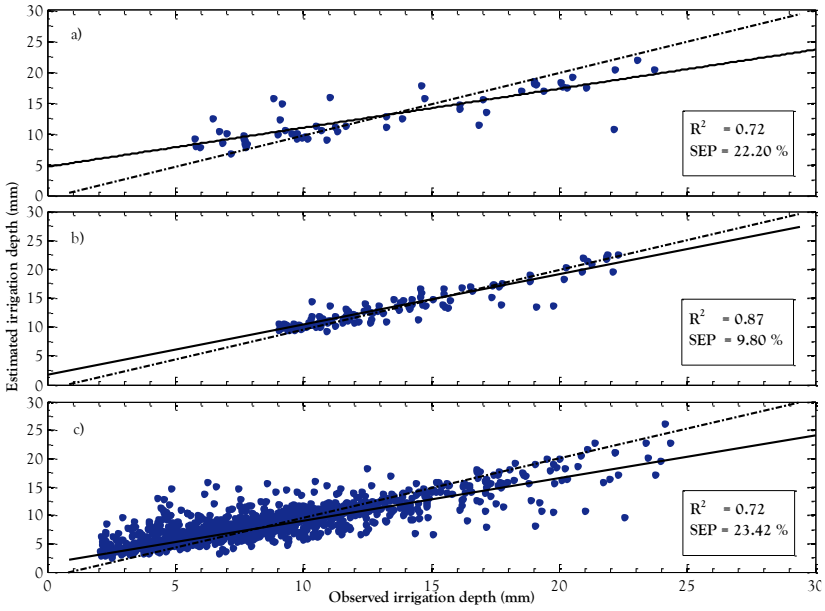


Fig. 8.8. Scatterplots between observed and estimated irrigation depths (testing period) for ANFIS models: a) rice, model 2; b) maize, model 4; (c) tomato, model 5.

Fig. 8.9 shows an example of the architecture of the best ANFIS model (ANFIS model 4), the MFs and the linguistic labels of the five input variables and three examples of the Sugeno-Takagi rules that make up the rule base of the FIS. The ANFIS model 4 was composed of 96 rules like R3, R27 and R48. The other ANFIS models, 2 and 5, had a similar architecture with 36 and 192 rules, respectively and the MFs and linguistic labels shown in the Table 8.5 and 8.6.

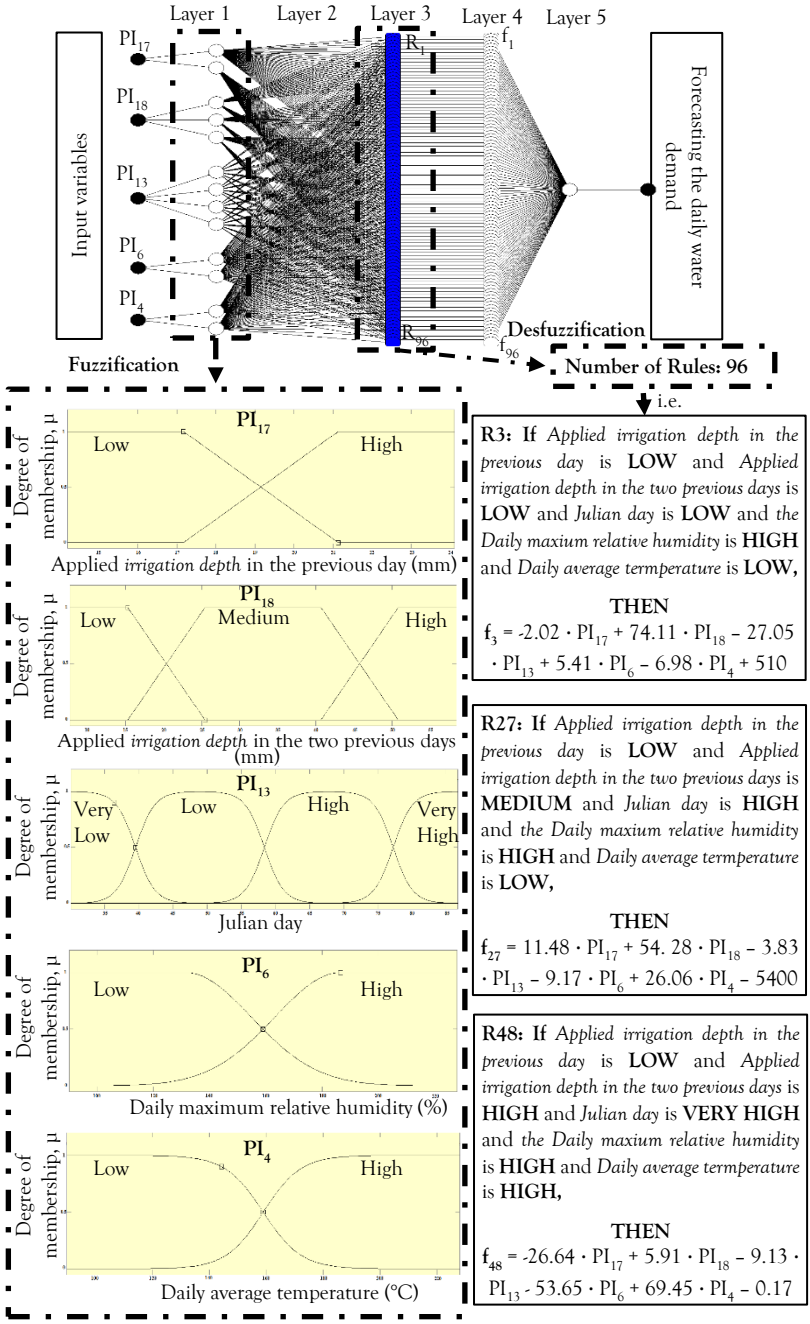


Fig. 8.9. Architecture of the ANFIS model 4, MF of each input variable and three examples of Sugeno-Takagi rules.

8.4. Conclusions

A model of farmer's behaviour to forecast the daily irrigation water used by each farmer has been developed combining Artificial Neural Networks, Fuzzy Logic and Genetic Algorithms. The predictive model was trained, validated and tested with the irrigation depths recorded for the three main crops (rice, maize and tomato) cultivated at the *Canal del Zújar Irrigation District* (southwestern Spain).

Results shown that the farmer's behaviour and the cultural practices are different for each crop even though the irrigation system was the same. Thus, when several crops are trained together, the representativeness of the model and the accuracy of the predictions were considerably worse than when each crop was trained independently. Hence, the R^2 values for rice, maize and tomato model were 0.72, 0.87 and 0.72, respectively and the SEP values were 22.20 %, 9.80 % and 23.42 % for these models, respectively. The irrigation systems and the size of the training, validation and testing sets conditioned the quality of the results. Thus, the rice model was lightly worse than maize and tomato models.

Both input variables selected in this work and the linguistic variables that portioned the universe of discourse of each of them give a great information about the farmer's management with

each crop. This information together with the developed models is a powerful tool for irrigation district managers to stablish strategies to save energy and water in the irrigation districts as well as is a useful tool to design of new irrigation systems, synchronizing the design flow rate and pressure rate on hydrant and pumping station with the real farmer´s behaviour and their real water consumption. Finally, combining these models with new models to predict the hourly distribution of the irrigation depth, new strategies of electrical energy hiring could be formulated.

Acknowledgments. This research was supported by an FPU grant (Formación de Profesorado Universitario) from the Spanish Ministry of Education, Culture and Sports to Rafael González Perea. This work is part of the TEMAER project (AGL2014-59747-C2-2-R), funded by the Spanish Ministry of Economy and Competitiveness.

8.5. References

Chang F-J, Wang Y-C, Tsai W-P (2016) Modelling Intelligent Water Resources Allocation for Multi-users. Water Resour Manag

Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 6:182–197

Hagan MT, Demuth HB, Beale MH (1996) Neural network design. PWS Publishing Co., Boston, Massachusetts

Jang J-SR, Sun C-T, Mizutani E (1997) Neuro-Fuzzy And Soft Computing Jang. Prentice-Hall, USA

Jury WA, Vaux HJ (2007) The Emerging Global Water Crisis: Managing Scarcity and Conflict Between Water Users. Adv. Agron. 95:1–76

Li Z (2006) Fuzzy chaotic systems: modeling, control, and applications. Springer-Verlag, Berlin, Heidelberg

Lin Y, Cunningham III GA, Coggeshall S V (1996) Input variable identification—fuzzy curves and fuzzy surfaces. Fuzzy sets Syst 82:65–71

Pulido-Calvo I, Gutiérrez-Estrada JC (2009) Improved irrigation water demand forecasting using a soft-computing hybrid model. Biosyst Eng 102:202–218

Safavi HR, Chakraei I, Kabiri-Samani A, Golmohammadi MH (2013) Optimal Reservoir Operation Based on Conjunctive Use of Surface Water and Groundwater Using Neuro-Fuzzy Systems. Water Resour Manag. doi: 10.1007/s11269-013-0405-1

Ventura S, Silva M, Pérez-Bendito D, Hervás C (1995) Artificial neural networks for estimation of kinetic analytical parameters. Anal Chem 67:1521–1525

Zadeh LA (1965) Fuzzy Sets. IEEE Trans Fuzzy Syst 8:416–424

9. Conclusions

9.1. General conclusions

- Irrigation districts are complex systems where there is a close relationship between the water distribution network and the on-farm irrigation systems. Thus, a holistic approach considering both systems simultaneously is essential to achieve an optimal use of water and energy resources.
- Traditional energy saving measures have been reformulated including the on-farm irrigation system in the optimization process. Consequently, energy cost savings between 15 % and 27 % were achieved.
- The joint use of crop modelling and precision irrigation offers new possibilities to irrigation managers to improve efficiency in the use of water and energy resources. In this context, a new integrated modelling approach that considers the spatial and temporal variability in the field and the uniformity of the irrigation system has been developed.
- Artificial Intelligence techniques are powerful tools for developing new water and energy saving strategies as well as for assisting irrigation district managers in their decision-

making process. In this research, a predictive model of daily water demand at distribution network level have been developed with an error in the estimation of the water demand of 12.6 %.

- A deeper knowledge of farmers' behavior in irrigation scheduling and water use is important to reduce the uncertainty that on-demand management causes on irrigation districts managers. In this context, two new models based on Decision Trees and fuzzy logic techniques have been developed and applied to a real case study. Thanks to this, the 100 % of the irrigation events in a real irrigation district has been properly classified and the daily water demand at farm level in a real case study has been predicted with an estimation error of 9.80 %.

9.2. Avenues for future research

A few future research lines derived from the outputs of this thesis are listed below:

- Integration of the electrical tariffs in the farmer's behavior model in order to estimate the energy costs and provide incentives to farmers to reduce their water use in peak energy price hours.

- Development of a "Universal Parent Model" which will able to classify any irrigation district and build automatically its own predictive model of daily water demand.
- Development of a decision support system, DSS, that would integrate all the models presented in this thesis. This DSS would be able to carry out a wide range of tasks from making water demand predictions, giving recommendations about networks sectoring, establishing the optimal operation of the pumping station and the optimum irrigation scheduling as well as searching for the best strategy to buy electricity based on previous experience and futures projections.
- Development of models to optimize the operation of the whole irrigation system that combines conventional and on-site renewable energy sources to reduce the energy supply dependence while ensuring the full satisfaction of crop irrigation requirements.
- Integration of all these methodologies in tools based on ICTs to provide optimal solutions within a user-friendly interface to be useful for farmers and managers.

9. Conclusiones

9.1. Conclusiones generales

- Las comunidades de regantes son sistemas complejos donde existe una estrecha relación entre la red de distribución de agua y los sistemas de riego en parcela. Así, para conseguir un uso óptimo de los recursos agua y energía es esencial aplicar un enfoque probabilístico que tenga en cuenta ambos sistemas conjuntamente.
- Las tradicionales medidas de ahorro energético han sido reformuladas incluyendo el sistema de riego en parcela en el proceso de optimización. Mediante la reformulación de estas medidas se ha conseguido un ahorro de energía entre el 25 % y el 27 %.
- El uso conjunto de la modelización de cultivos y el riego de precisión ofrece nuevas posibilidades a los gestores de riego para mejorar su actual eficiencia en el uso de agua y energía. En este contexto, se ha desarrollado un nuevo enfoque de modelo integral que considera la variabilidad espacial y temporal en la parcela y la uniformidad del sistema de riego.

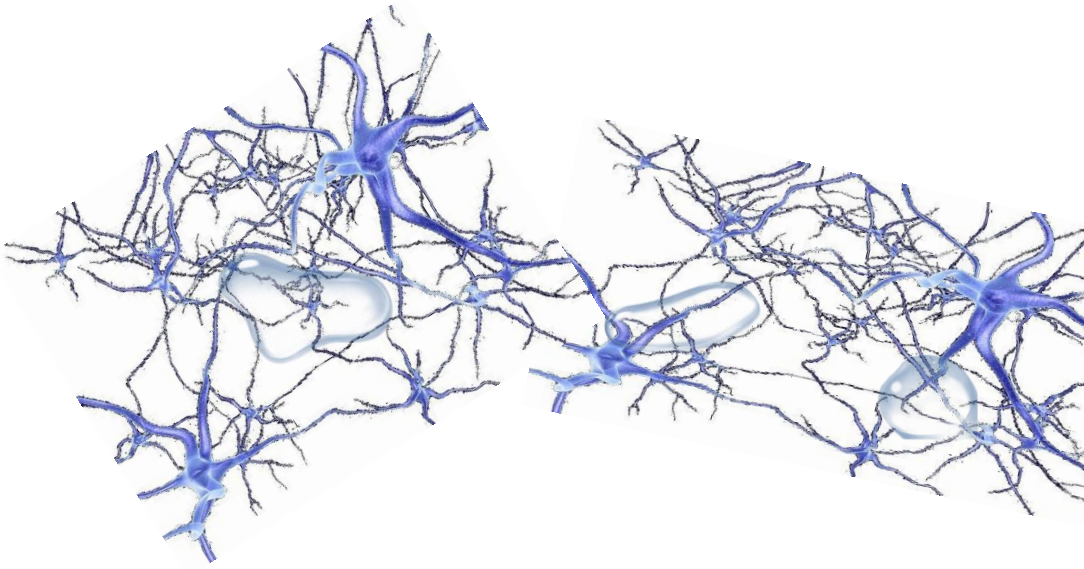
- La Inteligencia Artificial es una herramienta potente para el desarrollo de nuevas estrategias de ahorro de agua y energía y proporciona una útil herramienta de decisión para los gestores de las comunidades de regantes. En esta tesis, se ha desarrollado un modelo predictivo de demanda diaria de agua a escala de red de distribución con error en la estimación del 12.63 %.
- Conocer en profundidad el comportamiento del agricultor sobre el uso que hace del agua y la programación del riego es esencial para reducir la incertidumbre que los sistemas de riego organizados a la demanda ocasionan a los gestores de las comunidades de regantes. En este contexto, se han desarrollado dos nuevos modelos basados en Árboles de Decisión y técnicas de lógica difusa y han sido aplicados a un caso de estudio real. Gracias a la aplicación de estos modelos en un caso de estudio real, se ha podido clasificar correctamente el 100 % de los eventos de riego y se ha podido predecir la demanda de agua de riego de forma horaria con un error del 9.80 %.

9.2. Nuevas líneas de investigación derivadas de esta tesis

A continuación se enumeran algunas líneas de investigación futuras derivadas de los resultados de esta tesis:

- La integración de las tarifas eléctricas en el modelo del comportamiento del agricultor que estime los costes energéticos y proporcione incentivos al agricultor para reducir su uso de agua durante las horas donde el precio de la energía es mayor.
- El desarrollo de un "Modelo Universal Padre" que sea capaz de clasificar cualquier comunidad de regantes y construir de forma automática su propio modelo predictivo de demanda diaria de agua.
- Desarrollar un sistema de apoyo a la decisión que integre todos los modelos presentados en esta tesis y sea capaz no sólo de hacer predicciones sino de dar recomendaciones sobre la sectorización de las redes, el punto óptimo de funcionamiento de la estación de bombeo o la mejor estrategia de compra de energía basándose en experiencias previas y estableciendo proyecciones futuras.
- Desarrollar modelos de optimización que combinen la fuente de energía convencional con fuentes de energía renovables para reducir la dependencia energética y que asegure la total satisfacción de las necesidades de riego de los cultivos.
- Integración de todas estas metodologías en herramientas basadas en las nuevas Tecnologías de la Comunicación y de la Información (TICs) proporcionando soluciones óptimas bajo

una interface intuitiva y de fácil uso para los agricultores y gestores.



UNIVERSIDAD Ð CORDOBA