

Utah State University

DigitalCommons@USU

All Graduate Theses and Dissertations

Graduate Studies

8-2011

Bayesian Data-Driven Models for Irrigation Water Management

Alfonso F. Torres-Rua
Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/etd>

 Part of the [Civil and Environmental Engineering Commons](#)

Recommended Citation

Torres-Rua, Alfonso F., "Bayesian Data-Driven Models for Irrigation Water Management" (2011). *All Graduate Theses and Dissertations*. 979.
<https://digitalcommons.usu.edu/etd/979>

This Dissertation is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



BAYESIAN DATA-DRIVEN MODELS FOR
IRRIGATION WATER MANAGEMENT

by

Alfonso F. Torres Rua

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Civil and Environmental Engineering

Approved:

Wynn Walker
Co-Major Professor

Mac McKee
Co-Major Professor

Gary Merkley
Committee Member

David Stevens
Committee Member

Gilberto Urroz
Committee Member

Byron Burnham
Dean of Graduate Studies

UTAH STATE UNIVERSITY
Logan, Utah

2011

Copyright © Alfonso F. Torres Rúa 2011

All Rights Reserved

ABSTRACT

Bayesian Data-Driven Models for Irrigation Water Management

by

Alfonso F. Torres Rua, Doctor of Philosophy

Utah State University, 2011

Co-Major Professor: Dr. Wynn Walker
Co-Major Professor: Dr. Mac McKee
Department: Civil and Environmental Engineering

A crucial decision in the real-time management of today's irrigation systems involves the coordination of diversions and delivery of water to croplands. Since most irrigation systems experience significant lags between when water is diverted and when it should be delivered, an important technical innovation in the next few years will involve improvements in short-term irrigation demand forecasting.

The main objective of the researches presented was the development of these critically important models: (1) potential evapotranspiration forecasting; (2) hydraulic model error correction; and (3) estimation of aggregate water demands. These tools are based on statistical machine learning or data-driven modeling. These, of wide application in several areas of engineering analysis, can be used in irrigation and system management to provide improved and timely information to water managers. The development of such models is based on a Bayesian data-driven algorithm called the Relevance Vector Machine (RVM), and an extension of it, the Multivariate Relevance Vector Machine (MVRVM). The use of these types of learning machines has the advantage of avoidance

of model overfitting, high robustness in the presence of unseen data, and uncertainty estimation for the results (error bars).

The models were applied in an irrigation system located in the Lower Sevier River Basin near Delta, Utah.

For the first model, the proposed method allows for estimation of future crop water demand values up to four days in advance. The model uses only daily air temperatures and the MVRVM as mapping algorithm.

The second model minimizes the lumped error occurring in hydraulic simulation models. The RVM is applied as an error modeler, providing estimations of the occurring errors during the simulation runs.

The third model provides estimation of future water releases for an entire agricultural area based on local data and satellite imagery up to two days in advance.

The results obtained indicate the excellent adequacy in terms of accuracy, robustness, and stability, especially in the presence of unseen data. The comparison provided against another data-driven algorithm, of wide use in engineering, the Multilayer Perceptron, further validates the adequacy of use of the RVM and MVRVM for these types of processes.

(149 pages)

To my parents,
Faustino Torres and Luz Rua,
and my brothers and sisters,
Juan, Lisbet, Lucia, Jose, Luzdia and Daniel

ACKNOWLEDGMENTS

I would like to express my sincere appreciations to my advisors, Dr. Wynn Walker and Dr. Mac McKee, for their valuable advice and encouragement. I sincerely appreciate the words of wisdom of Dr. Walker and his role in directing my attention towards the practical utility of my present and future work. Dr. McKee allowed me to develop academically in a free environment, allowing me to explore my edge in research.

I also thank my committee, Dr. Gary Merkley, Dr. David Stevens, and Dr. Gilberto Urroz, for their guidance and opportune questions throughout the research. Special thanks to the Utah Water Research Laboratory staff for the financial and logistic support. Without them this work would not have been possible.

I sincerely appreciate Jim Walker, water commissioner of the Sevier River Basin, and Roger Hansen from the US Bureau of Reclamation for their enthusiasm and support to this research developed in Delta, UT.

I also would like to thank my friends and coworkers, Dr. Andres Ticlavilca and Dr. Omar Alminagorta, for their support and thoughtful questions.

And last but not least, I give special thanks to my family and friends in Peru and USA for their help, encouragement, and moral support, throughout this time at Utah State University.

Alfonso F. Torres Rua

CONTENTS

	Page
ABSTRACT.....	iii
ACKNOWLEDGMENTS	vi
LIST OF TABLES.....	x
LIST OF FIGURES	xi
CHAPTER	
1. INTRODUCTION	1
General introduction	1
Purpose and objectives.....	2
Purpose of the study.....	2
Objectives	3
Research motivation.....	3
Research contributions.....	4
Dissertation organization	4
2. FORECASTING DAILY POTENTIAL EVAPOTRANSPIRATION USING MACHINE LEARNING AND LIMITED CLIMATIC DATA....	6
Abstract	6
Introduction.....	7
Theoretical development.....	9
Potential evapotranspiration.....	9
Hargreaves ET_0 equation	10
Multi-layer perceptron	12
Multivariate relevance vector machine	15
Material and methods.....	19
Area of study.....	19
Data description	21
Methodology	21
Results.....	24

	Comparison of forecasted to estimated crop ET	32
	Conclusions and discussion	37
	References	39
3.	MACHINE LEARNING APPROACH FOR ERROR CORRECTION OF HYDRAULIC SIMULATION MODELS.....	42
	Abstract	42
	Introduction.....	43
	Theoretical development.....	45
	Saint Venant equations	45
	Numerical solution of the Saint Venant equations	47
	Model error	48
	Statistical learning machines.....	51
	Multi-layer perceptron	51
	Relevance vector machine	54
	Material and methods.....	57
	Site description.....	57
	Data acquisition	59
	Hydraulic simulation model.....	59
	Learning machines	61
	Results.....	63
	Hydraulic model performance	63
	Error modeling	66
	Conclusions and discussion	73
	References.....	76
4.	MULTIPLE-DAY IRRIGATION WATER DEMAND FORECAST USING MULTIVARIATE RELEVANCE VECTOR MACHINES	79
	Abstract	79
	Introduction.....	80
	Theoretical development.....	82

Irrigation water demand.....	82
Multi-layer perceptron	85
Multivariate relevance vector machine	88
Material and methods.....	92
Site description.....	92
Data description	94
Aggregate water demand forecasting model.....	94
Learning machines	97
Results.....	98
Available information	98
Water demand forecasting model	101
Conclusions and discussion	110
References.....	113
5. SUMMARY, CONCLUSIONS, AND RECOMENDATIONS	116
Summary and conclusions	116
Recommendations for future work	120
References.....	121
APPENDIX.....	123
CURRICULUM VITAE.....	135

LIST OF TABLES

Table		Page
2.1	Best learning machines configuration.....	26
2.2	Goodness-of-fit per approach	26
2.3	Crop distribution in Canal B for 2009	34
3.1	Hydraulic simulation model for Canal A.....	60
3.2	Variables tested for the aggregate error correction model.....	64
3.3	Variables included in ϵ_A correction model using MLP and goodness-of-fit obtained for test data (2009).....	67
3.4	Variables included in ϵ_A correction model using RVM and goodness-of-fit obtained for test data (2009).....	67
4.1	2008 and 2009 crop areas for Canal B.....	99
4.2	Selected soil moisture monitoring stations	102
4.3	Tested variables in the forward variable selection procedure.....	105
4.4	Goodness-of-fit and error bar values for variable selection procedure using MLP (2009 data)	105
4.5	Goodness-of-fit and error bar values for variable selection procedure using MVRVM (2009 data).....	106

LIST OF FIGURES

Figure	Page
2.1	Area of study, ACA Canal B in Delta, Utah.....20
2.2	Used ET_0 forecasting approaches23
2.3	Goodness-of-fit values for the evaluated approaches27
2.4	Peak season forecast for Direct Approach-MLP28
2.5	Peak season forecast for Direct Approach-MVRVM29
2.6	Peak season forecast for Indirect Approach-MLP30
2.7	Peak season forecast for Indirect Approach-MVRVM.....31
2.8	Bootstrapping results for the Direct Approach33
2.9	Bootstrapping results for the Indirect Approach.....34
2.10	Agricultural crops in Canal B – 200935
2.11	Best approach ET forecasting performance36
3.1	Simulation model and error sources50
3.2	Canal A location, Delta, Utah58
3.3	Simulation + error correction model.....62
3.4	Observed water levels, discharge and aggregate error (ϵ_A) for 2008.....64
3.5	Observed water levels, discharge and aggregate error (ϵ_A) for 2009.....65
3.6	ϵ_A statistics for 2009 irrigation season65
3.7	(a) Actual, simulated ϵ_A and (b) ϵ_A residuals obtained using the best MLP error correction model (2009).....69
3.8	Statistical characteristics of ϵ_A and ϵ_A residuals using the MLP model (2009).....69
3.9	(a) Actual, simulated ϵ_A and (b) ϵ_A residuals obtained using the best RVM error correction model (2009).....70
3.10	Statistical characteristics of ϵ_A and ϵ_A residuals using the RVM model

	(2009).....	70
3.11	Goodness-of-fit statistics for MLP and RVM error correction models from Bootstrap Analysis (2009 Data)	73
4.1	Area of study, ACA Canal B in Delta, Utah.....	93
4.2	Crop distribution for ACA Canal B, 2009	99
4.3	Local K_c and 2009 actual evapotranspiration values for main crops in ACA Canal B	101
4.4	One and two days water demand forecast for 2009 irrigation season using MLP	107
4.5	One and two days water demand forecast for 2009 irrigation season using MVRVM	107
4.6	η and RMSE Bootstrap results for the MLP (2009 irrigation season).....	111
4.7	η and RMSE Bootstrap results for the MVRVM (2009 irrigation season) ..	111

CHAPTER 1

INTRODUCTION

General introduction

Population growth worldwide and associated increase in demand for food, potable water and other services create the possibility of a future water shortage problem which will require methods to reduce the water use in activities that use large quantities, such as irrigated agriculture. This type of agriculture will be the principal source of water to supply increased urban and industrial demands. Still, changes toward reduced water use in irrigation will be slow, costly and disruptive. Among the many reasons that could be argued, one of the most apparent is the lack of adequate information or tools to support better decisions related to more efficient water management in irrigation.

Water related sources of information exist and increase every day in number and quantity, and government and private organizations expand efforts to collect, store and make available collected data. There still remains a need for models or tools that can provide information to manage water. Thus, the collected data do not necessarily translate into adequate information for water management, and in some cases could be impractical as inputs for hydraulic or hydrologic models.

The implications of this dilemma are vast, having significant influence in the control on the water supply, demand tradeoff, precise scheduling of future releases from water storages, water loss minimization and control of flow rate in canals while providing adequate amounts of water for irrigated lands. Therefore, it is important that research be done to address the issues regarding adequate information and tools necessary for decision makers.

Various hydraulic and hydrologic models have been used to address informational needs in irrigation for years with relative degrees of success, depending on the adequacy of the input data and the used simulation tools. Particular models employed are linked to the resources available to the organization responsible for water management. Thus, new ways of supplying adequate information and enhancement of models already in use is necessary to provide better support for management activities. For an operator of an irrigation canal, the critical information is related to expected near term (next days) values of water requirements for irrigation, e.g. crop evapotranspiration and agricultural command area (ACA) water requirements. In terms of enhancement of in-use simulation models, adequate correspondence between measured and simulated hydraulic parameters is important to precisely estimate the amount and timing of water deliveries, thereby offering better control over the allocated water for irrigation.

Also, a crucial decision in today's irrigation system management involves the coordination of water releases or diversions and the delivery timing of these flows to the croplands. Since most irrigation systems experience significant lags between when water is released or diverted in comparison to when it should be delivered, perhaps the most important technical innovation in the next few years will involve demand forecasting.

Purpose and objectives

Purpose of the study

The purpose of this study was to develop adaptable methods and tools that allow for better management of water destined for irrigation, using state-of-the-art supervised

learning machines, while measuring their possible performance when implemented for everyday use.

Objectives

The objectives of the research were to:

- Develop a method that quantifies short-term future crop water needs in irrigation lands in limited climatic data scenario.
- Develop a method that allows for error correction in simulation models that account for combined parameters, variables and structural error sources.
- Develop an approach to forecast near-term irrigation water demand for an agricultural command area to improve canal operations in large-scale irrigation systems.
- Present an adequate procedure that allows for the development and replication of the proposed methods by comparison of data-driven algorithms.
- Estimate future performance of the methods developed using supervised learning machines by measurement of goodness-of-fit parameters.

Research motivation

Recent research literature has shown some promising applications in a variety of water resources management problems through the use of Bayesian learning machine algorithms. This initiated the idea that these algorithms could be potentially applied for irrigation water demand management. Given the Bayesian theory imbedded in these algorithms, these can also provide additional information about the variability of the results obtained.

Research contributions

The proposed research has demonstrated the applicability of Bayesian data-driven algorithms to provide adequate solution to the objectives mentioned earlier. This study was the first attempt to use Bayesian learning machine algorithms for:

- Daily ET_0 forecasts based on limited weather data.
- Minimization of aggregated or lumped error from a physical-based simulation model.
- Near-term daily future estimations of water demand for an ACA based on local information.

Dissertation organization

The dissertation consists of five chapters. Chapter 1 is an introduction to this document which includes the motivation, description of the overall objectives, and motivations for the major contributions of the research, and outlines the conceptual framework for the developed models. Chapter 2 provides an insight of previous work on potential evapotranspiration forecasting as it appears in scientific literature; it describes in a detailed manner the procedure developed; and shows the results obtained and a comparison with a similar alternative method. Chapter 3, similar in structure to the previous one, describes the proposed methods in the literature to provide an error correction model for physical-based models, and details the proposed procedure developed in this study for a coupled physical- and statistical-based model to reduce the impact of lumped or aggregate error in the simulation results. This chapter also discusses the obtained results and provides a comparison with an alternative data-driven algorithm of wide use in the scientific literature. Chapter 4, structured after the previous two chapters, analyzes the current proposed methods to determine future water demand

forecasts for an irrigation command area and presents a new methodology to forecast water demand based on only SCADA and limited climatic data. Chapter 5 provides a summary of this work, draws the major conclusions that follow, and provides recommendations for further research.

The structure of this document is based on the paper dissertation format. As result, some redundancies and repetition of parts of the material presented, especially the description of the data-driven algorithms and area of study, occur.

CHAPTER 2
FORECASTING DAILY POTENTIAL EVAPOTRANSPIRATION USING MACHINE
LEARNING AND LIMITED CLIMATIC DATA¹

ABSTRACT

Anticipating or forecasting near-term irrigation demands is a requirement for improved management of conveyance and delivery systems. The most important component of a forecasting regime for irrigation is a simple, yet reliable, approach to estimate future crop water demands, which in this paper is represented by the reference or potential evapotranspiration (ET_0). In most cases, weather information for the irrigation system is limited to a reduced number of measured variables; therefore estimation of ET_0 values is restricted. This paper summarizes the results of two forecasting ET_0 approaches under the mentioned condition. The first or direct approach involved forecasting ET_0 directly using historically computed values. The second or indirect approach involved forecasting the required weather parameters for the ET_0 calculation based on historical data and then computing ET_0 . A statistical machine learning algorithm, the Multivariate Relevance Vector Machine algorithm (MVRVM) is applied for both of the forecasting approaches. The general ET_0 model used is the 1985 Hargreaves Equation which requires only minimum and maximum daily air temperatures and is thus well suited to regions lacking more comprehensive climatic data. The utility and practicality of the forecasting

¹ Reprinted from Agricultural Water Management Journal, Vol. 98/4, Alfonso F. Torres, Wynn R. Walker and Mac McKee, "Forecasting Daily Potential Evaporation Using Machine Learning and Limited Climatic Data," pages 553-562, Copyright (2011), with permission from Elsevier.

methodology is demonstrated with an application to an irrigation project in Central Utah. To determine the advantage and suitability of the applied algorithm, another learning machine, the Multilayer Perceptron (MLP), is tested in the present study.

Introduction

Population increases over the next decades will place a substantial emphasis on achieving higher irrigation efficiencies and greater production per unit of water. A key component of any strategy to improve irrigation water management will be related with improvement of water delivery strategies and efficiencies within the irrigation delivery networks. As Kumar et al. (2002) note, evapotranspiration (ET) is one of the most important components of the hydrologic cycle and its accurate estimation is of vital importance for such diverse areas as hydrologic water balance, irrigation system design and management, crop yield simulation and water resources planning and management. Likewise, achieving higher irrigation system performance will depend on reliable forecasts of cropland ET and will require that such forecasts be far enough in the future to compensate for lag time travel of the water supply. ET estimation is an important input to water management and irrigation scheduling because crop demands are generally the largest component of water diversions.

A number of computational methods have been developed to estimate potential evapotranspiration (ET_0) from climatic data. These methods vary in complexity from models that require only basic information, such as maximum and minimum air temperature (Hargreaves, 1974), to complex models that estimate ET_0 through energy balance models, such as the Penman - Monteith method (Allen et al., 1998). The advantage of simple models is their suitability in regions with minimal available weather

data. Their major disadvantage is that these may not reflect the effects of localized climatic and geographic variations such as narrow valleys, high ground elevations, extreme latitudes or strong winds. Also, simple methods are usually best suited to weekly or monthly ET_0 estimates than daily estimates.

In recent years there have been several attempts to estimate and forecast ET_0 with a higher degree of accuracy and over extended futures. Some of them involve numerical and statistical approaches that attempt to accurately simulate the random nature of the meteorological variables (Yamashita and Walker, 1994). The occurrence of difficulties related with these attempts forced researchers to look for other techniques using data-driven tools or statistical learning machines, such as Artificial Neural Networks (Kumar et al., 2002; Lai et al., 2004; Smith et al., 2006), Simple Bayes Classifier and k-Nearest Neighbors (Verdes et al., 2000), Support Vector Machines and Relevance Vector Machines (Gill et al., 2006). These newer approaches have been used primarily to forecast hourly ET_0 values up to 24 hours in advance. Forecasting of daily ET_0 beyond one day using data-driven algorithms have not been reported, even though these methods are known for having excellent modeling accuracy, particularly in representing complex nonlinear behavior (Lai et al., 2004).

The objective of this paper is to demonstrate the adequacy of two different approaches for forecasting daily ET_0 using statistical learning machines and limited climatic information. The learning machine algorithm used is the Multivariate Relevance Vector Machine (MVRVM). The potential crop evapotranspiration is estimated by the 1985 Hargreaves Equation. These results are then compared with ET values of the area under study to determine the accuracy of the forecasted estimates obtained. Also, for

comparison and benchmarking analysis, another learning machine was tested, the Multi-layer Perceptron, to determine the suitability of the proposed mapping algorithm.

Theoretical development

Potential evapotranspiration

The potential or reference evapotranspiration (ET_0) expresses the evaporating power of the atmosphere at a specific location and time of the year and does not consider crop characteristics or soil factors. As it is mentioned by Allen et al. (1998), the only factors affecting ET_0 are climatic parameters. Consequently, ET_0 is a climatic parameter and can be computed from weather data. Among the several methods to estimate ET_0 , the FAO Penman-Monteith method is recommended as the sole method for determining ET_0 . This method has been selected because it closely approximates grass ET_0 at the location evaluated, is physically based, and explicitly incorporates both physiological and aerodynamic parameters.

Situations might occur where data for some weather variables are missing. The use of an alternative ET_0 calculation procedure, requiring reduced meteorological parameters, should generally be avoided. It is recommended that ET_0 should be calculated using the standard FAO Penman-Monteith method after resolving the specific problem of the missing data (Allen et al., 1998). Despite of this, when climatic data as Net Radiation is not available, Allen et al. (1998) suggest the Hargreaves ET_0 equation for its use.

Hargreaves ET₀ equation

Hargreaves and Allen (2003) note that the current Hargreaves equation was developed in 1975 in an attempt to improve the ET₀ equation developed by Christiansen (1968). Using eight years of daily cool season grass in precision weighting lysimeters and weather data, Hargreaves performed regressions among measured ET₀ and temperature data using several ET₀ methods. Several posterior attempts to improve the resulting ET₀ equation led to the 1985 Hargreaves ET₀ Equation:

$$ET_0 = 0.0023 \cdot R_a (T_C + 17.8) \cdot T_R^{0.5} \quad (2.1)$$

where:

ET₀: potential evapotranspiration (mm/day) of a reference crop (grass),

T_{max}, and T_{min}: maximum and minimum daily air temperature (°C),

TC: 0.5 (T_{max} + T_{min}),

TR: T_{max} - T_{min} and;

R_a: extraterrestrial solar radiation (mm/day).

The following empirical simplifications allow R_a estimation using the latitude and the day of the year, as mentioned by Allen et al. (1998):

$$R_a = 37.6 * d_r (\omega_s \sin(\varphi_1) \sin(\delta) + \cos(\varphi_1) \sin(\omega_s)) / \lambda \quad (2.2)$$

$$\delta = 0.4093 * \sin(2\pi(284 + J) / 365) \quad (2.3)$$

$$d_r = 1 + 0.033 * \cos(2\pi J / 365) \quad (2.4)$$

$$\omega_s = \cos(\tan(\varphi_1) \tan(\delta)) \quad (2.5)$$

being:

d_r: relative distance from the earth to the sun,

J: day of the year,

ω_s : sunset hour angle (rad),

ϕ : latitude (rad),

δ : declination of the sun (rad) and

λ : latent heat of vaporization, $\lambda \approx 2.54$.

Hargreaves and Allen (2003) stated that the best use of Eq. 2.1 would be for ET estimation in regional planning and reservoir operation studies. The attractiveness of Hargreaves ET₀ model is its simplicity, reliability, minimal data requirement, and ease of computation. The viability of using Hargreaves instead of Penman-Monteith ET₀ equation is demonstrated in the study by Trajkovic and Kolakovic (2009) where the difference between Hargreaves and Penman-Monteith ET₀ equations is in the range of -4.7% to 6.9% for all the weather stations used.

The estimated reference or potential evapotranspiration is translated into the actual crop evapotranspiration, ET, by adjusting ET₀ for crop variety growing stage as follows:

$$ET = K_c * ET_0 \quad (2.6)$$

in which K_c is the crop growth stage factor (no dimensional) and ET has the same units as ET₀. Tables and information related with the crop stage factor have been developed for most of the agricultural crops around the world. Average values of K_c for various crops can be found in Allen et al. (1998), while location-specific K_c can be found in research publications (Wright, 1982).

Multi-layer perceptron

Among the large number of implementations of Artificial Neural Network (ANN) models, the Multi-Layer Perceptron (MLP) is one of the most widely used because of its ability to approximate any smooth function (Nabney, 2002). An interesting characteristic of this ANN is the inclusion of the Bayesian Inference Method to calibrate the MLP parameters. The Bayesian Inference Method also allows estimation of the variability related to the predicted outputs. The MLP architecture can be described as:

$$y^{(n)} = \mathbf{W}^{\text{II}*} \cdot \tanh\left(\mathbf{W}^{\text{I}*} x^{(n)} + \mathbf{b}^{\text{I}*}\right) + \mathbf{b}^{\text{II}*} \quad (2.7)$$

where:

$y^{(n)}$: MLP output vector, $y^{(n)}=[y_1, \dots, y_m, \dots, y_M]$,

$x^{(n)}$: input vector $x^{(n)}=[x_1, \dots, x_d, \dots, x_D]$,

$\mathbf{W}^{\text{I}*}, \mathbf{W}^{\text{II}*}$: optimized weights for the first and second layer respectively,

$\mathbf{W}^{\text{I}*} = [w_{1,1}^{\text{I}}, \mathbf{K}, w_{D,NN}^{\text{I}}]$, $\mathbf{W}^{\text{II}*} = [w_{1,1}^{\text{II}}, \mathbf{K}, w_{NN,M}^{\text{II}}]$

M: number of components of the output vector,

D: number of components in the input vector,

NN: number of hidden neurons,

$\mathbf{b}^{\text{I}*}, \mathbf{b}^{\text{II}*}$: bias vectors for the first and second layer, respectively.

Using a dataset $\Lambda = \{x^{(n)}, t^{(n)}\}_{n=1}^N$, where N is the number of training cases, the calibration of the MLP is performed by optimizing the network parameters $\mathbf{W} = \{\mathbf{W}^{\text{I}}, \mathbf{W}^{\text{II}}, \mathbf{b}^{\text{I}}, \mathbf{b}^{\text{II}}\}$ in order to minimize the Overall Error Function E (Bishop, 1995):

$$E = \frac{\beta}{2} \sum_{n=1}^N \left(t^{(n)} - y^{(n)} \right)^2 + \frac{\alpha}{2} \sum_{i=1}^W W_i^2 \quad (2.8)$$

$$E = \beta \cdot E_{\Lambda} + \alpha \cdot E_w$$

where:

E_{Λ} : data error function,

E_w : penalization term,

W : number of weights and biases in the neural network, and

α, β : Bayesian hyperparameters.

In Bayesian terms, the goal is to estimate the probability of the weights and bias of the MLP model, given the dataset Λ :

$$p\left(\mathbf{W} | t^{(n)}\right) = \frac{p\left(t^{(n)} | \mathbf{W}\right)p\left(\mathbf{W}\right)}{p\left(t^{(n)}\right)} \quad (2.9)$$

where, as explained by MacKay (1992):

$p(\mathbf{W}|t^{(n)})$: the posterior probability of the weights,

$p(t^{(n)}|\mathbf{W})$: the dataset likelihood function,

$p(\mathbf{W})$: the prior probability of the weights, and

$p(t^{(n)})$: the evidence for the dataset.

Assuming a Gaussian distribution for the error term $\xi^{(n)} = t^{(n)} - y^{(n)}$ and the weights \mathbf{W} , the likelihood and the prior probabilities can be expressed:

$$p\left(t^{(n)} | \mathbf{W}, \beta\right) = \left(2\pi\beta^{-1}\right)^{-N/2} \exp\left(-\beta E_{\Lambda}\right) \quad (2.10)$$

$$p\left(t^{(n)} | \mathbf{W}, \alpha\right) = \left(2\pi\alpha^{-1}\right)^{-N/2} \exp\left(-\alpha E_w\right) \quad (2.11)$$

E_Λ models the uncertainty (or error) of the target variables as Gaussian zero-mean noise and variance $\sigma^2 \equiv \beta^{-1}$. E_W defines the conditional probability of W with variance $\sigma_W^2 \equiv \alpha^{-1}$. Then Eq. 2.9 can be expressed as:

$$p(W | t^{(n)}, \alpha, \beta) = \frac{p(t^{(n)} | W, \beta) \cdot p(W | \alpha)}{p(t^{(n)} | \alpha, \beta)} \quad (2.12)$$

$$p(W | t^{(n)}, \alpha, \beta) = \frac{\exp\left(E(W^*) - \frac{1}{2} \Delta W^T H \Delta W\right)}{\exp\left(E(W^*) \cdot (2\pi)^{W/2} |H|^{1/2}\right)} \quad (2.13)$$

in which,

$E(W^*)$: expected optimized values for the weights and bias,

$H = \text{Hessian matrix } H = \beta \nabla \nabla E_\Lambda + \alpha I$, I is the identity matrix.

$\Delta W = W - W^*$.

Once the distribution of W has been estimated by maximizing the likelihood for α and β , the prediction $y^{(n)}$ and its standard deviation $\sigma_y^{(n)}$ can be estimated by integrating (marginalizing) over W and the regularization parameters α and β (Bishop, 1995):

$$p(y^{(n)} | x^{(n)}, t^{(n)}) = \int p(t^{(n)} | x^{(n)}, W^*) p(W^* | t^{(n)}) \cdot dW \quad (2.14)$$

This can be approximated by:

$$p(y^{(n)} | x^{(n)}, t^{(n)}) \propto \left(2\pi \sigma_y^{(n)2}\right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \sigma_y^{(n)2} \left(y^{(n)*} - t^{(n)}\right)^2\right) \quad (2.15)$$

where $y^{(n)}$ is the output and $\sigma_y^{(n)^2}$ is the output variance from the MLP. The output variance can be expressed as:

$$\sigma_y^{(n)^2} = \beta^{-1} + \mathbf{g}^T \mathbf{H}^{-1} \mathbf{g} \quad (2.16)$$

where \mathbf{g} denotes the gradient of $y^{(n)}$ with respect to the weights; $\mathbf{g} \equiv \nabla W y^{(n)} | \mathbf{W}^*$. The output variance has then two sources; the first arises from the intrinsic noise in the target data; and the second from the posterior distribution of the ANN weights (Pierce et al. 2008). The output standard deviation vector $\sigma_y^{(n)}$ can be interpreted as the error bar for confidence interval estimation (Bishop, 1995).

Multivariate relevance vector machine

The Multivariate Relevance Vector Machine (MVRVM), developed by Thayananthan et al. (2008), is a general Bayesian framework for obtaining multivariate sparse solutions to regression tasks. The MVRVM is based on the Relevance Vector Machines framework developed by Tipping (2001) and Tipping and Faul (2003) which was extended to handle multivariate outputs. This learning machine is particularly useful in hydrology and water resources because of the generalization properties and the probabilistic estimation, useful to estimate prediction uncertainty (Tripathi and Govindajaru, 2007). The mathematical formulation of the MVRVM is:

$$\mathbf{y}^{(n)} = \mathbf{W}^* \cdot \Phi[\mathbf{x}^{(n)}] \quad (2.17)$$

where:

$\mathbf{x}^{(n)}$ and $\mathbf{t}^{(n)}$: input and target vectors that belong to the dataset $\{\Lambda\}_{n=1}^N$, as defined for the MLP,

$y^{(n)}$: MVRVM output vector $y^{(n)}=[y_1, \dots, y_M]$; $m \in 1 \leq m \leq M$,

M: number of components in the target and MVRVM output vectors,

N: number of training cases,

$\overset{*}{W}$: optimized weight matrix, $\overset{*}{W} = [w_{1,1}, \dots, w_{m,rv}, \dots, w_{M,RV}]$,

RV: number of optimal cases or relevance vectors selected by the MVRVM from the N training cases, $RV \ll N$, $rv \in 1 \leq rv \leq RV$,

$\overset{*}{\Phi}[x^{(n)}]$: optimized design matrix or basis function (represented also by $\overset{*}{\Phi}$) that can be related with a kernel function $\overset{*}{\Phi} = K \left[x^{(n)}, \left\{ x^{(*)} \right\}_{rv=1}^{RV} \right]$.

The kernel function is a weighting function for the input vector $x^{(n)}$ used in non-parametric estimation techniques, e.g. kernel regression models. It provides an adjustment to the $x^{(n)}$ vector based on RV optimal cases or relevant vectors, $x^{(*)}$, which are selected automatically among the N training input vectors. For calibration of the MVRVM a variation of the Overall Error function (Eq. 2.8) is used and by means of the Bayesian Inference Method the distribution of the weights of the model (Eq. 2.17) is estimated, similar to the MLP calibration process. The MVRVM error term or residual $\xi^{(n)} = t^{(n)} - y^{(n)}$ is assumed to be probabilistic independent zero-mean Gaussian, with variance σ_{ξ}^2 . The detail of the MVRVM algorithm is as follows:

Assuming a Gaussian prior probability distribution for the weights (Tipping, 2001), and representing $A = \text{diag}(\alpha_1^{-2}, \dots, \alpha_N^{-2})$, and $B = \text{diag}(\beta_1, \dots, \beta_M)$, where each element α_n is a hyperparameter that determines the relevance of the associated basis function for every case in the training data. $\beta_m = \sigma_{\xi}^2$ represents the noise or error variance

in the m^{th} component of the target data (Thayananthan et al., 2008). The prior distribution over the weights is represented by:

$$p(\mathbf{W} | \mathbf{A}) = \prod_{m=1}^M \prod_{n=1}^N \mathcal{N}(w_{m,n} | 0, \alpha_n^{-2}) \quad (2.18)$$

where $w_{m,n}$ is the element at (m, n) of the weighting matrix, $\mathbf{W} = [w_{1,1}, \dots, w_{m,n}, \dots, w_{M,N}]$.

The likelihood distribution of \mathbf{W} can be expressed as:

$$p(\{t^{(n)}\}_{n=1}^N | \mathbf{W}, \mathbf{B}) = \prod_{n=1}^N \mathcal{N}(t^{(n)} | \mathbf{W} \cdot \Phi, \mathbf{B}) \quad (2.19)$$

with $\Phi = \mathbf{K} \left[\{x^{(n)}\}_{n=1}^N, \{x^{(n)}\}_{n=1}^N \right]$. The likelihood of the target $t^{(n)}$ can be written as:

$$p(\{t^{(n)}\}_{n=1}^N | \mathbf{W}, \mathbf{B}) = \prod_{m=1}^M \mathcal{N}(\tau_m | w_m \cdot \Phi, \beta_m) \quad (2.20)$$

τ_m is a vector with the m^{th} component of all the target data and w_m the weight vector of the m^{th} component of the output vector $t^{(n)}$. The prior distribution over the weights can be rewritten as:

$$p(\mathbf{W} | \mathbf{A}) = \prod_{m=1}^M \mathcal{N}(w_m | 0, \mathbf{A}) \quad (2.21)$$

The posterior probability of \mathbf{W} can be written as the product of separate Gaussians of the weights vectors of each output dimension:

$$p(\mathbf{W} | \{t^{(n)}\}_{n=1}^N, \mathbf{B}, \mathbf{A}) \propto \left(\{t^{(n)}\}_{n=1}^N | \mathbf{W}, \mathbf{B} \right) \cdot p(\mathbf{W} | \mathbf{A}) \quad (2.22)$$

$$p(\mathbf{W} | \{t^{(n)}\}_{n=1}^N, \mathbf{B}, \mathbf{A}) \propto \prod_{m=1}^M \mathcal{N}(w_m | \mu_m, \Sigma_m) \quad (2.23)$$

The terms $\mu_m = \beta_m^{-1} \Sigma_m \Phi^T \tau_m$ and $\Sigma_m = (\beta_m^{-1} \Phi^T \Phi + \mathbf{A})^{-1}$ are the mean and the variance of the weight matrix respectively. Marginalizing the data likelihood over the weights:

$$p\left(\left\{t^{(n)}\right\}_{n=1}^N \mid \mathbf{A}, \mathbf{B}\right) = \int p\left(\left\{t^{(n)}\right\}_{n=1}^N \mid \mathbf{W}, \mathbf{B}\right) \cdot p(\mathbf{W} \mid \mathbf{A}) \cdot d\mathbf{W} \quad (2.24)$$

$$p\left(\left\{t^{(n)}\right\}_{n=1}^N \mid \mathbf{A}, \mathbf{B}\right) = \prod_{m=1}^M \left| \mathbf{H}_m \right|^{\frac{1}{2}} \exp\left(-\frac{1}{2} \boldsymbol{\tau}_m^T \mathbf{H}_m^{-1} \boldsymbol{\tau}_m\right) \quad (2.25)$$

\mathbf{H}_m is the Hessian matrix for the m^{th} component of the target vector, $\mathbf{H}_m = \beta_m \mathbf{I} + \Phi^T \mathbf{A}^{-1} \Phi$. An optimized set of hyperparameters $\left\{ \alpha_{rv}^* \right\}_{rv=1}^{RV}$ and noise parameters $\left\{ \beta_m^* \right\}_{m=1}^M$ is obtained by maximizing the marginal likelihood as described by Tipping and Faul, (2003). The final hyperparameter values are:

$$\mathbf{A}^* = \text{diag}\left(\alpha_1^{*-2}, \mathbf{K}, \alpha_{RV}^{*-2}\right) \quad (2.26)$$

$$\boldsymbol{\Sigma}_m^* = \left(\beta_m^* \Phi^T \Phi + \mathbf{A}^* \right)^{-1} \quad (2.27)$$

The optimized mean vector and the weight matrix are:

$$\boldsymbol{\mu}_m^* = \beta_m^* \boldsymbol{\Sigma}_m^* \Phi^T \boldsymbol{\tau}_m \quad (2.28)$$

$$\mathbf{W}^* = \left(\mu_1^*, \dots, \mu_M^* \right)^T \quad (2.29)$$

The MVRVM output and output error bar vectors are:

$$\mathbf{y}^{(n)*} = \mathbf{W}^* \cdot \Phi \quad (2.30)$$

$$\boldsymbol{\sigma}_y^{(n)*} = \text{sqrt}\left(\mathbf{B}^{-1} + \Phi^T \cdot \boldsymbol{\Sigma} \cdot \Phi\right) \quad (2.31)$$

Materials and methods

Area of study

The water resources of the Sevier River Basin in Central Utah (Fig. 2.1) are among the most heavily utilized in the Western US. Substantial efforts to increase efficiency via canal lining and on-farm improvement such as conversion to sprinkler irrigation and laser land leveling were made during 1960 - 1990 period. From 1990 to the present, all reservoirs and stream offtakes have been equipped with SCADA technology and web-based data summaries (SRWUA, 2009). Canal automation was introduced in 1994 and shown to not only to result in substantial reduction in losses but also to considerably shorten the response time between farmer demands and system deliveries (Walker and Stringam, 1999, 2000).

Most recently, attention has been focused on improving the coordination between farmer demands, canal deliveries, and reservoir diversions, which depend to a large extent on forecasting the irrigation demand. In order to develop, test and implement the ET_0 forecasting approaches of this study, a subsystem at the lower end of the Sevier River was selected. The Canal B system as shown in Fig. 2.1 commands about 10,500 hectares in extent and is managed by the Delta Canal Company. The Canal B area is connected to the DMAD reservoir by a 9 km. canal (Canal A). The DMAD gates as well as the Canal B gates are automated and operated as a SCADA system by local water masters. The lag time from DMAD Reservoir to the Canal B headgates is about 3 hours. The lag time between the Canal B inlet and an individual farm within the Canal B area averages 9 hours.

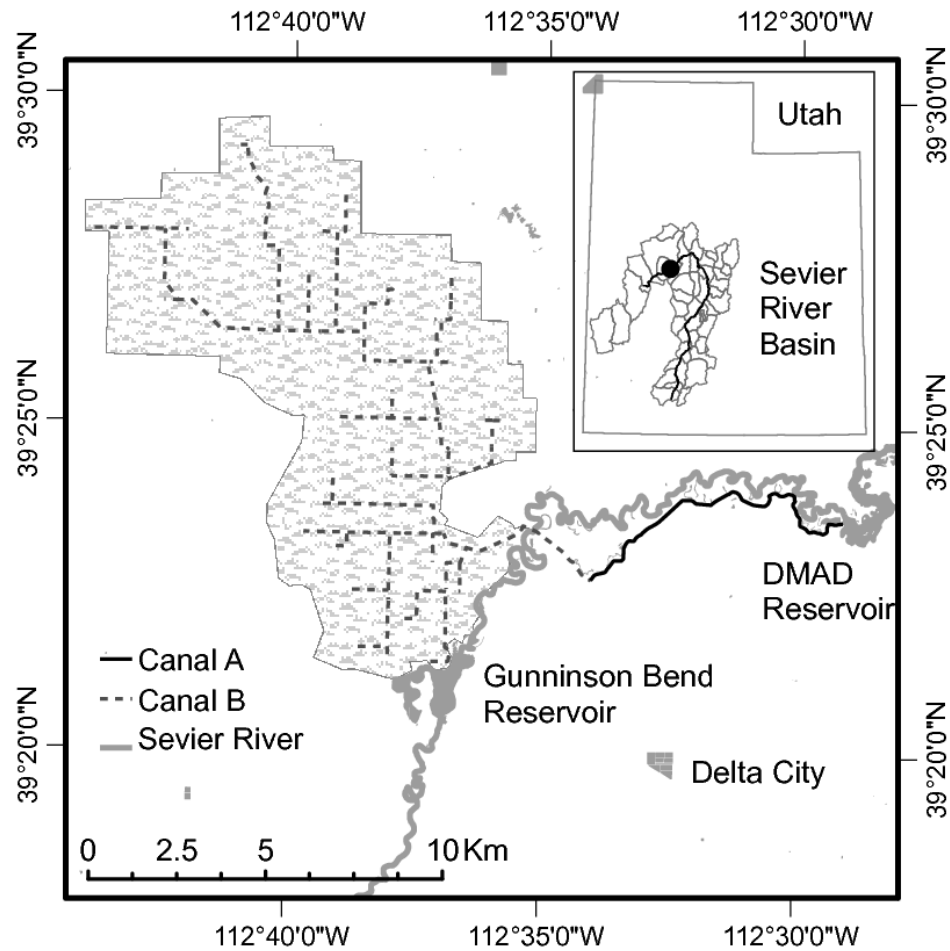


Fig. 2.1 Area of study, ACA Canal B in Delta, Utah.

The DMAD Reservoir is supplied water on a demand basis from Sevier River Bridge Reservoir upstream. The lag time from Sevier Bridge Reservoir to DMAD Reservoir is approximately 3 days. Thus, an emerging crop demand in the Canal B area can be supplied within about 12 hours if water is available in DMAD Reservoir, or 4 days if water must be conveyed from Sevier Bridge Reservoir. The goal of the entire system is to provide water to an individual farm within 12 hours of an order by the irrigator. This goal relies heavily on the SCADA system and the regulation capacity of DMAD Reservoir.

The water management goals over the next few years are to reduce the DMAD regulation capacity and improve the reliability of the 12 hour delivery interval period. It is expected that controlled DMAD Reservoir levels will reduce seepage, evaporation, and administration losses by about 25 to 50%. The most important capability needed to achieve this goal is to develop a reliable and accurate forecast of irrigation demand, which begins with the ET estimates.

Data description

All the weather data for this study were taken from the meteorological station located in Delta, Utah (WMO Station Number 72479), available at the NOAA - National Climatic Data Center website (2009). From this station, daily minimum and maximum air temperatures over the full period from January 2000 until December 2009 were available. For each of the 10 years, a subset was selected that only includes the daily air temperatures during the agricultural season (March to October, ~ 256 days). Information about crop coefficients (K_c) for the Lower Sevier River Basin was obtained from the study by Wright (1982). Information about crop distributions and effective area per crop for the years 2006 to 2009 was obtained from the LandSat Imagery Program website (2009).

Methodology

As noted in the introduction to this paper, two approaches were considered for forecasting ET_0 using the 1985 Hargreaves equation. The first approach (Direct Approach) involved the calculation of historical ET_0 from the daily minimum and maximum air temperatures and then applying the machine learning algorithm described

above to simulate directly the ET_0 time series, obtaining as result the forecasted values of ET_0 . The second approach (Indirect Approach) involved applying the learning machine to the daily minimum and maximum air temperatures. Then ET_0 is computed using the forecasted air temperatures. A schematic view of the two approaches is presented in Fig. 2.2.

For both of the approaches considered, the data collected from the NOAA website was divided into two groups or datasets; the first group was used for training the learning machines and the second group for testing or estimating the accuracy of the results provided by the calibrated learning machines. It was considered a training/testing dataset ratio of 1.5:1. This gives a training data size over 6 irrigation seasons (years 2000 to 2005) with $N = 1476$ cases. The testing dataset involved 4 irrigation seasons (years 2006 to 2009) with $N_* = 984$ cases.

Two testing criteria have been used to evaluate the results: (1) the Root Mean Square Error (RMSE); and (2) the Nash-Sutcliffe Efficiency Index (η). The Nash- Sutcliffe Efficiency Index is recommended for nonlinear modeling problems (McCuen et al., 2006).

$$RMSE = \sqrt{\sum_{n=1}^{N_*} (y_*^{(n)} - t_*^{(n)})^2 / N_*} \quad (2.32)$$

$$\eta = 1 - \frac{\sum_{n=1}^{N_*} (y_*^{(n)} - t_*^{(n)})^2}{\sum_{n=1}^{N_*} (t_*^{(n)} - \overline{t_*^{(n)}})^2} \quad (2.33)$$

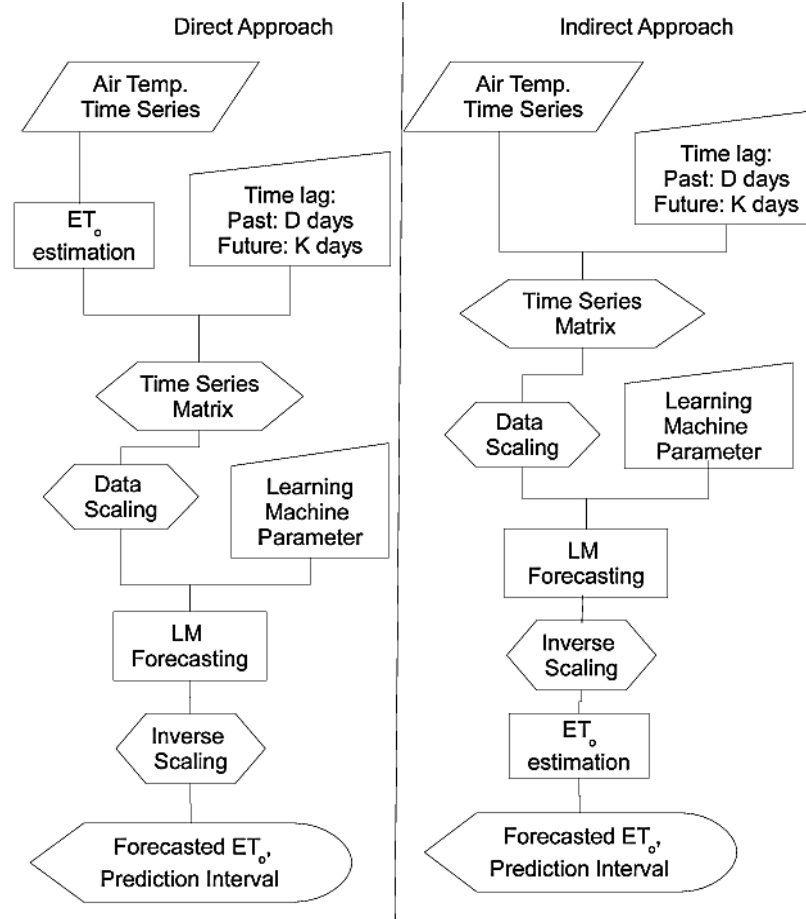


Fig. 2.2 Used ET_0 forecasting approaches.

where $t_*^{(n)}$: calculated ET_0 values for the testing data, $y_*^{(n)}$: forecasted values of ET_0 for the testing data, N_* : number of samples or cases in the testing data, and $\bar{t}_*^{(n)}$ average values of the calculated ET_0 .

The RMSE values allow to rank the performance of each learning machine, being large RMSE values an indication that the error between the calculated and predicted ET_0 values is large too. The η value measures the closure of the calculated vs. the predicted ET_0 values in a non-dimensional range (from $-\alpha$ to 1). A η value of 1 is an indication of perfect correspondence. A η value of 0 indicates that the forecasted ET_0 is not better than

the average of the calculated ET_0 values. As a reference of the optimal range for η , a study by Khan and Coulibaly (2005) suggests that an adequate performance of learning machines for forecasting flow rates should yield η values in the range of 0.8 to 1.0.

For both of the approaches, the main issue is to determine for each learning machine the adequate number of past values or inputs (D) of daily air temperatures or ET_0 values for the forecasting of multiple future values or outputs (K) of air daily temperatures or ET_0 respectively. About the learning machine algorithms, for the MLP the parameter to calibrate is the number of neurons in the hidden layer, and for the MVRVM the kernel width parameter σ_k . The optimal values of number of inputs D and the respective learning machine parameter was selected by trial procedure aimed at obtaining the best RMSE and η values.

To ensure good generalization of the learning machines tested under variation of the training data, a bootstrap analysis was built for each approach on the best calibration of the MLP and MVRVM, to evaluate the significance of the testing criteria and draw conclusion about model reliability (Khalil et al., 2006). Also, in order to compare the actual crop ET in the area under study vs. the best forecasted estimates from the used approaches, a graphical analysis is performed.

Results

Using the training and testing datasets described earlier, the calibration of the machine learning algorithms was made using both the Direct and Indirect Approaches. To determine the performance and accuracy of the learning machines for forecasting, a 7-day forecast horizon was solicited for the ET_0 analyses. This value is in fact larger than

the required 4-day ahead forecast for the area under study, nevertheless, additional information could be useful for extended daily water management operations.

Under this consideration, the search for the best number of required inputs - learning parameter value combination based on the approaches presented in Fig. 2.2 was performed. For the Direct Approach, the best calibrated MLP was obtained with 7 days of ET_0 values (inputs) in the past, using 9 hidden neurons. The best calibrated MVRVM was obtained with 10 past daily ET_0 values using the Laplace kernel and σ_k value of 10. The type of kernel for the MVRVM models was selected in a preliminary test.

For the Indirect Approach, the best calibrated MLP was found using 5 daily air temperatures (maximum and minimum) values in the past (10 inputs) and 13 hidden neurons. In the same manner for the best calibrated MVRVM, the number of past daily air temperatures was of 8 days (16 inputs), the type of kernel is Laplace and the σ_k value is 3. Table 2.1 shows the best configuration values of the learning machines used per approach and Table 2.2 shows the values of the goodness-of-fit parameters values for forecasted days 1, 3, 4, and 7 per approach. Fig. 2.3 also show the behavior of these statistical parameters in relation with the number of forecasted ET_0 days for each learning machine and approach.

The goodness-of-fit parameters (Table 2.2) indicate that, on average, the two considered approaches and the learning machines used were able to provide a reasonable ET_0 forecast up to 3 days ahead for the four irrigation seasons considered as test data (2006 -2009) considering a threshold $\eta \geq 0.8$ (Khan and Coulibaly, 2005). Beyond the fourth day, only the models from the Indirect Approach were able to provide higher η values than the threshold value considered. Also, after the fourth forecasted day, the

Table 2.1. Best learning machines configuration.

Description	MVRVM		MLP	
	Direct	Indirect	Direct	Indirect
Approach	Direct	Indirect	Direct	Indirect
kernel type/optim function	Laplace	Laplace	Secant grad	Secant grad.
kernel width σ_k / hidden neurons	10	3	9	13
Days in the past (inputs):	10	8	7	5
Forecasted days (outputs):	7	7	7	7

Table 2.2. Goodness-of-fit per approach.

Approach	Direct - BNN			
Day	1	3	4	7
RMSE (mm/day)	0.65	0.84	0.86	0.91
η	0.88	0.80	0.79	0.77

Approach	Direct - MVRVM			
RMSE (mm/day)	0.65	0.85	0.87	0.89
η	0.88	0.80	0.79	0.77

Approach	Indirect - BNN			
RMSE (mm/day)	0.65	0.84	0.84	0.85
η	0.88	0.80	0.80	0.79

Approach	Indirect - MVRVM			
RMSE (mm/day)	0.65	0.83	0.84	0.85
η	0.88	0.80	0.80	0.80

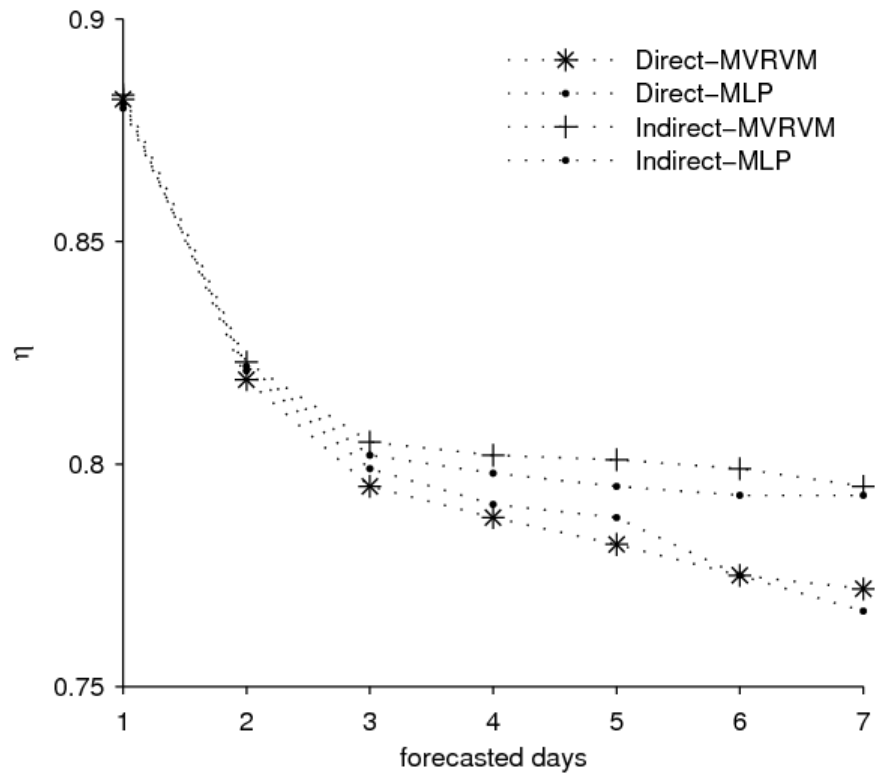


Fig. 2.3. Goodness-of-fit values for the evaluated approaches.

remaining forecasted ET_0 values can be considered as a broader reference of the actual ET_0 .

Figs. 2.4 to 2.7 show the calculated and forecasted ET_0 values for the 2009 irrigation season and also the correspondence among these values for forecasted 1, 4, and 7 days. For day 1, the learning machine models are able to estimate the future seasonal (long term), the mid-term trends of the ET_0 plus its daily variation. From day 2 to 7, the accuracy of the estimation of the daily trend decreases. The subplots (to the right), which show the 45° degree plot in the Figs. mentioned, provides insight of the relationship of the forecasted ET_0 values when compared with their respective calculated values. These figures indicate that there is small sub-estimation and over-estimation of the forecasted

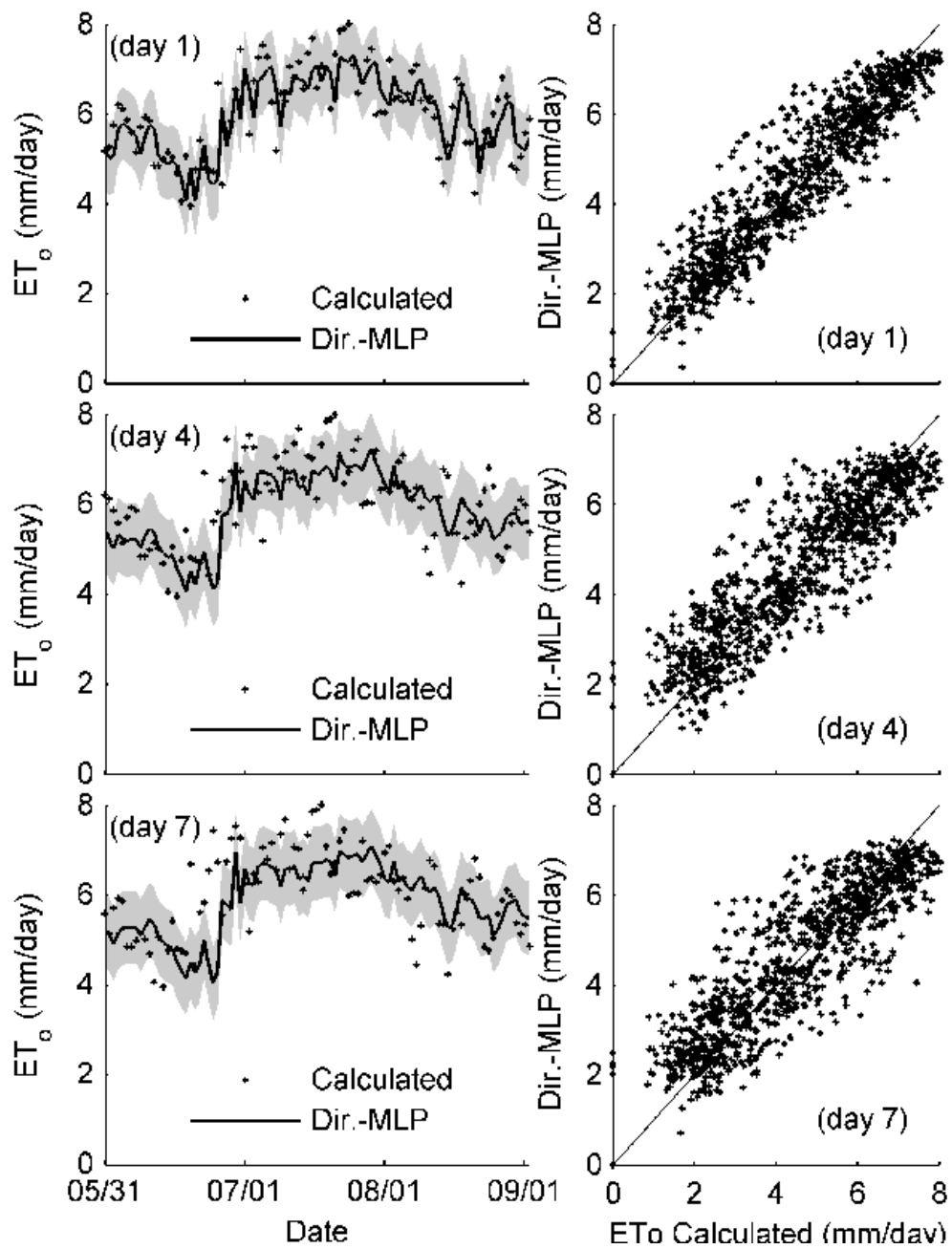


Fig. 2.4. Peak season forecast for Direct Approach-MLP.

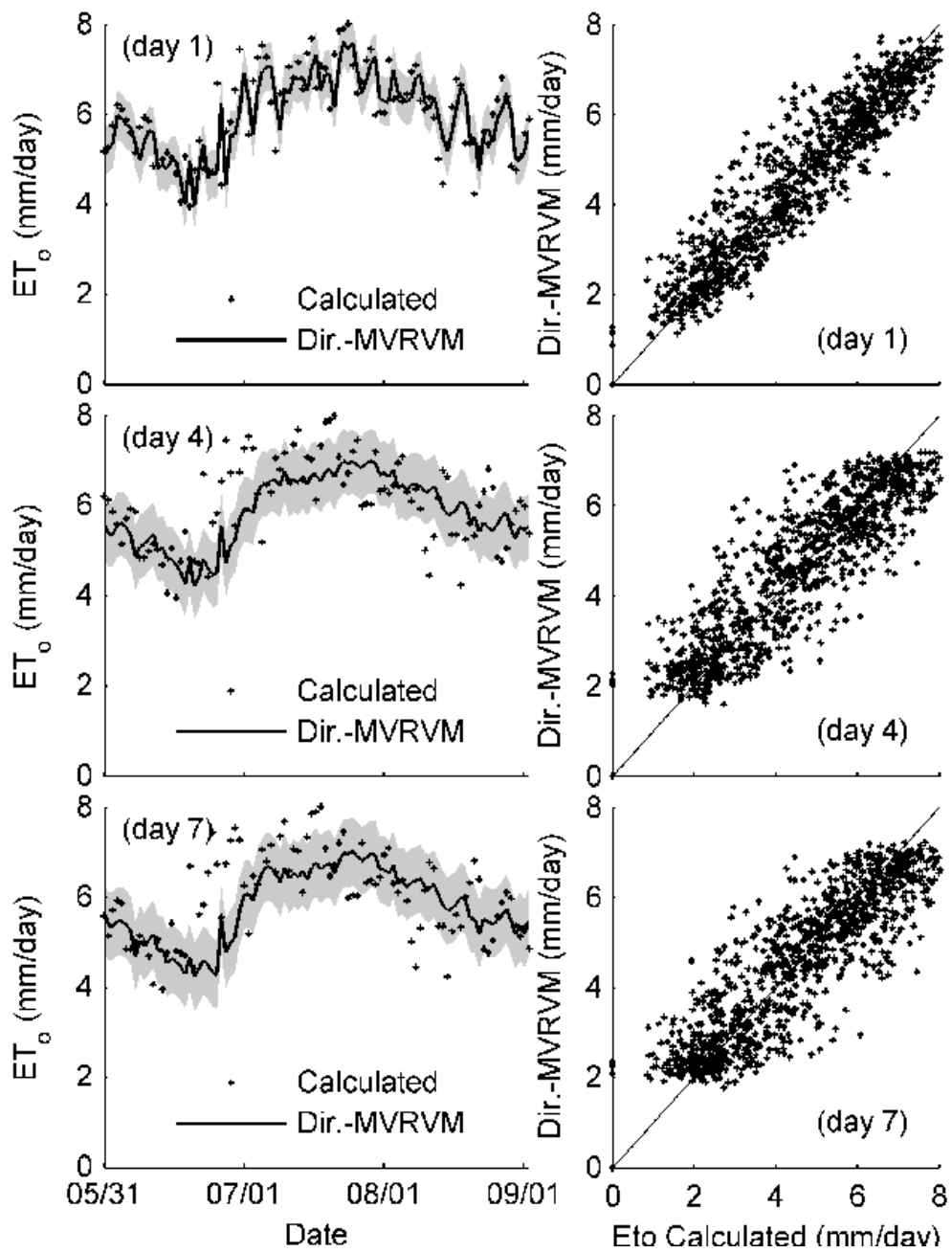


Fig. 2.5. Peak season forecast for Direct Approach-MVRVM.

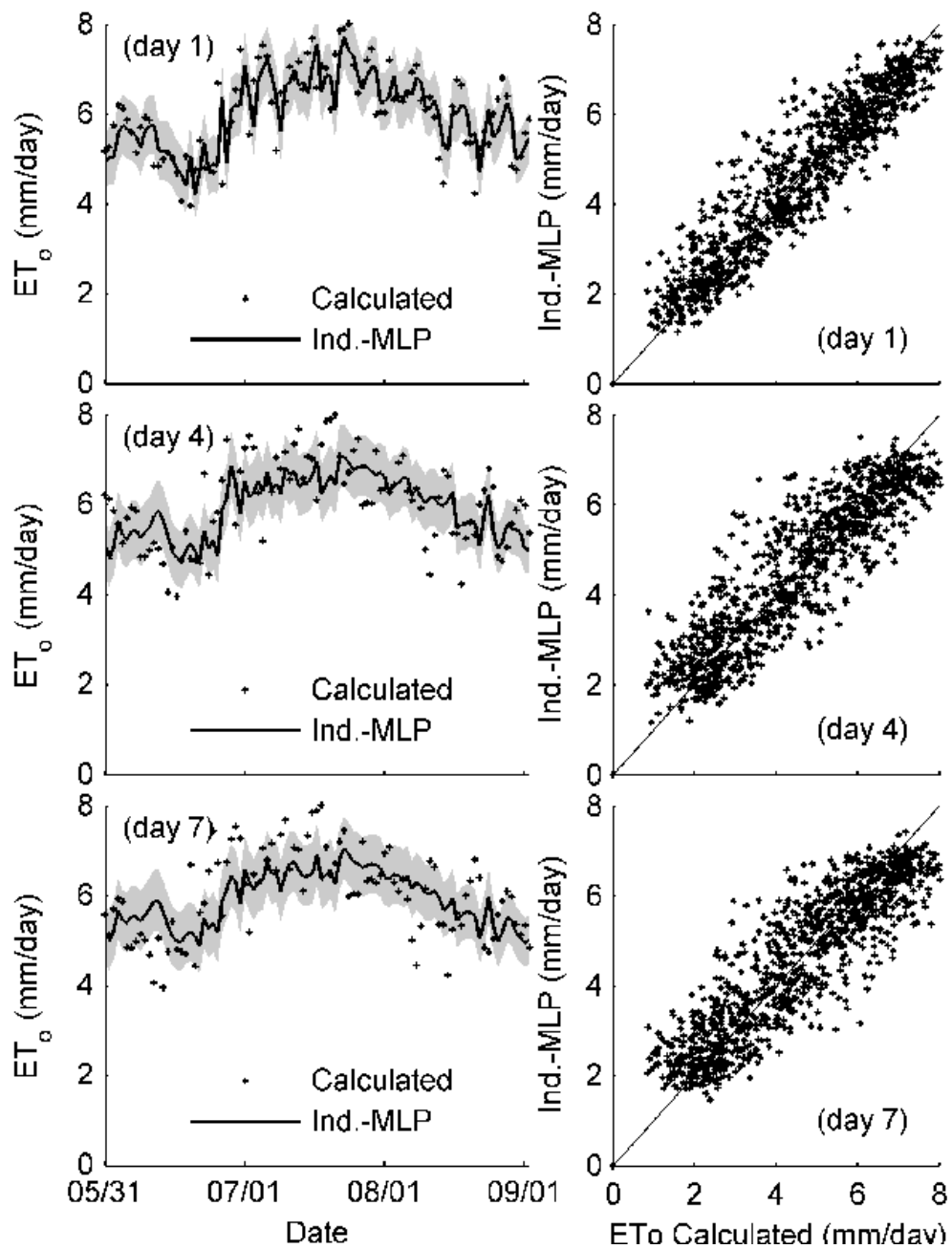


Fig. 2.6. Peak season forecast for Indirect Approach-MLP.

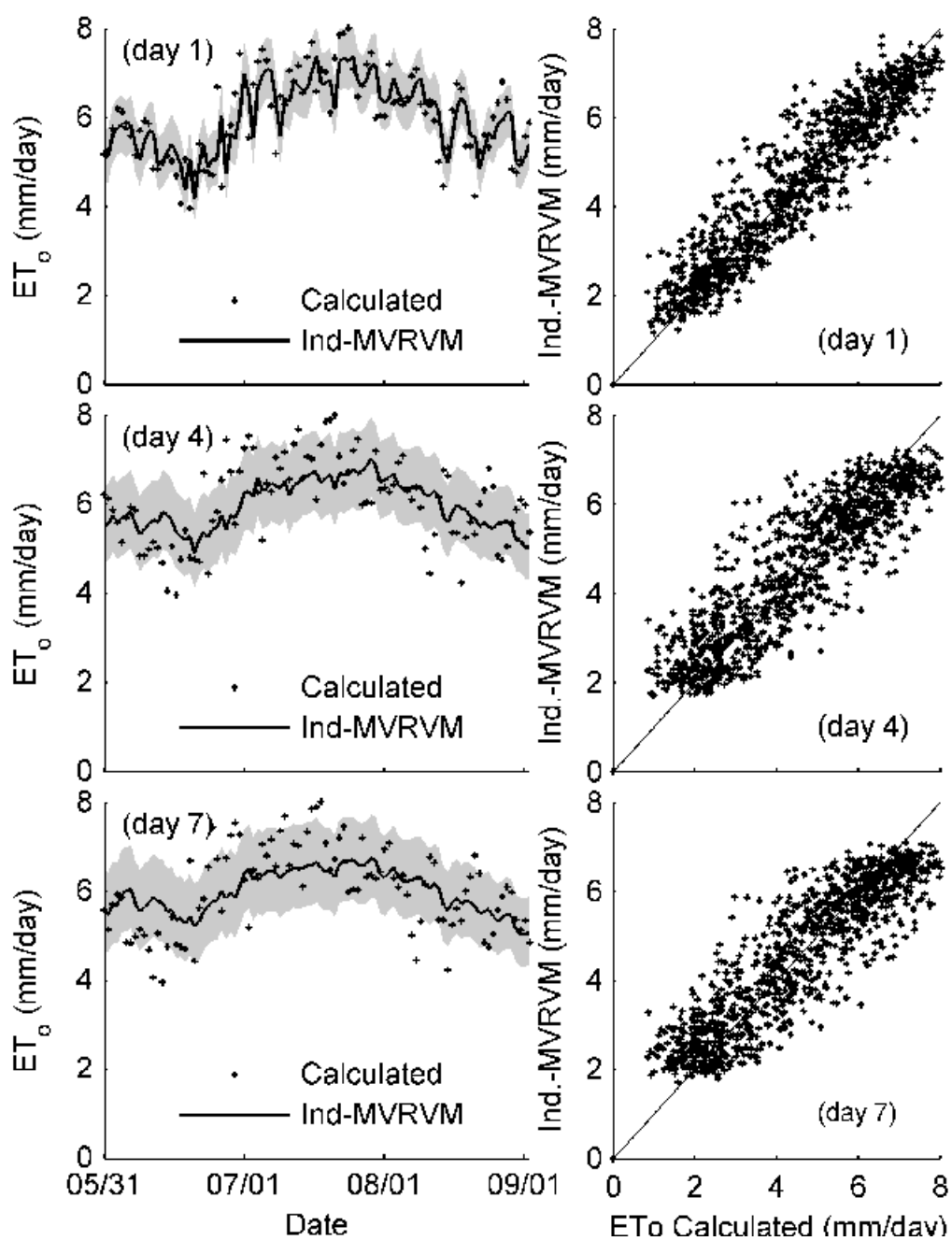


Fig. 2.7. Peak season forecast for Indirect Approach-MVRVM.

maximum and minimum values respectively. This characteristic of the results seems to increase along with the forecast time interval for all the approaches considered.

When comparing the Direct and Indirect Approaches, there is a small advantage of the Indirect Approach related with the error bar estimation for the forecasted ET_0 . The error bar of the Indirect Approach varies along the irrigation season, providing smaller error bar values for low ET_0 values and broader values for seasonal peak ET_0 values.

This is a result of forecasting the required weather variables for the 1985 Hargreaves ET_0 equation. As described by Eq. 2.1, the forecasted daily maximum and minimum air temperatures and their respective error bars are affected by the extraterrestrial radiation estimation value which is smaller at the beginning and end of the irrigation season and maximum at the peak season.

In terms of stability and robustness of the models, Figs. 2.8 and 2.9 show the performance of the η parameter for the learning machines used for both approaches. In general, MLP models for either approach proved to be less robust than the MVRVM, which is demonstrated by the wider distribution of the η histogram for MLP when compared with the distribution obtained for the η histogram of the MVRVM. When comparing Direct and Indirect Approaches, it is the latter approach that provides in average better goodness-of-fit values as also is demonstrated in Table 2.2.

Comparison of forecasted to estimated crop ET

In order to determine the practical adequacy of the best forecasting approach tested, a comparison among the forecasted and the actual crop ET in daily basis was performed for the year 2009 using actual data. For this purpose, forecasted ET_0 values

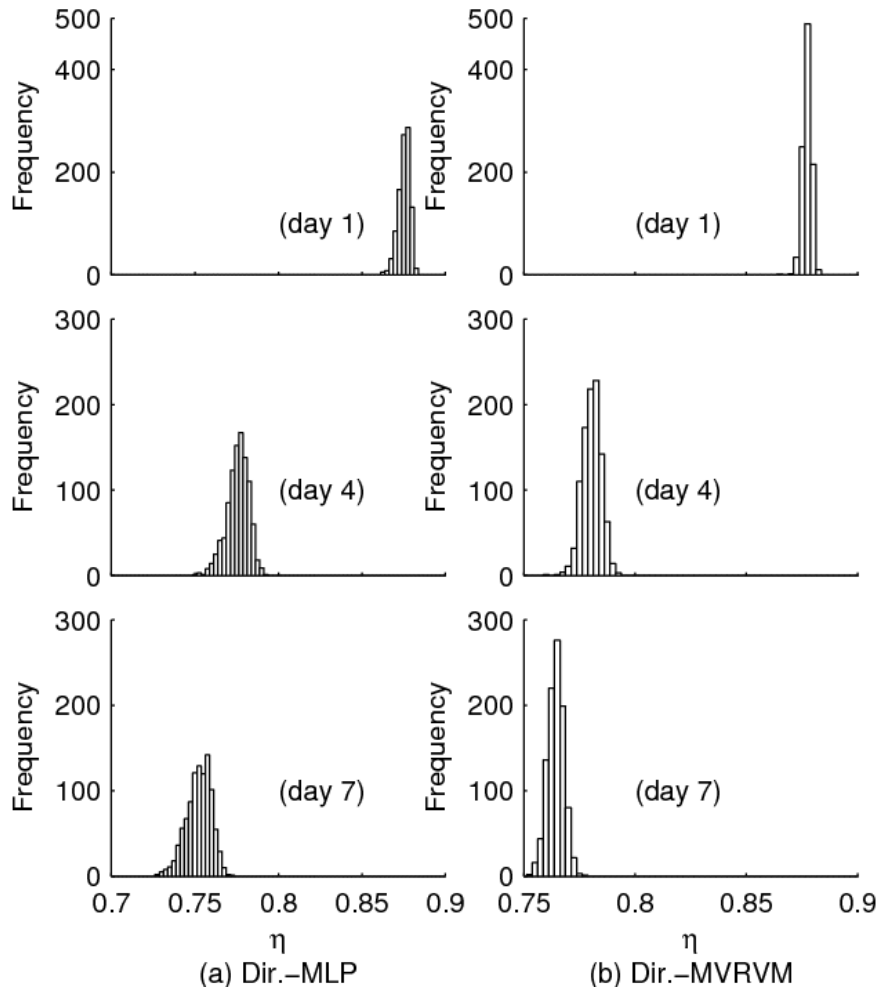


Fig. 2.8. Bootstrapping results for the Direct Approach.

from the Indirect Approach using the MVRVM model, information about local crop coefficients K_c by Wright (1982) and the distribution of the crops in the area under study were considered as described in the Material and Methods Section. Three main crop groups were considered accordingly to the agriculture of the area: alfalfa, corn and small grains. The crop area estimation and its relative percentage are presented in Table 2.3. Also the crop areas identification using a LandSat 5 TM satellite image for the year 2009 is presented in Fig. 2.10.

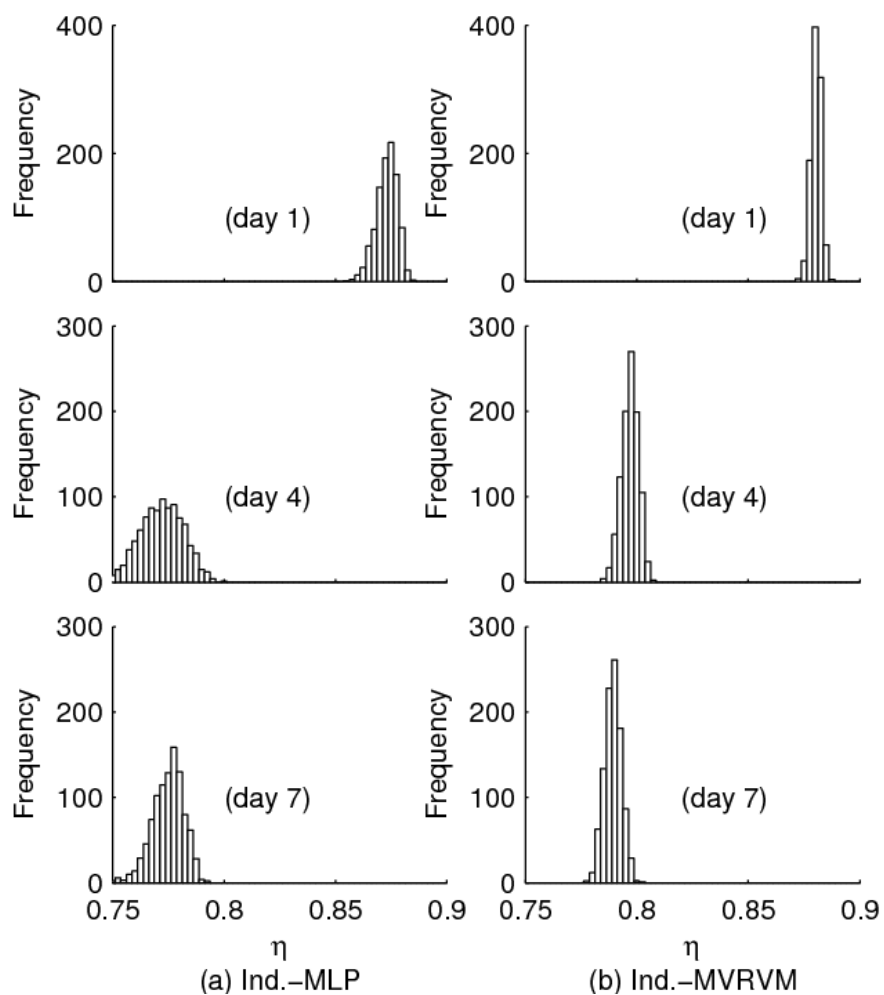


Fig. 2.9. Bootstrapping results for the Indirect Approach.

Table 2.3. Crop distribution in Canal B for 2009.

Crop	Area (ha)	%
Alfalfa	3369.2	32.0
Corn	723.6	7.0
Small Grains	323.3	3.0
Fallow	6105.2	58.0
Total	10521.2	100.0

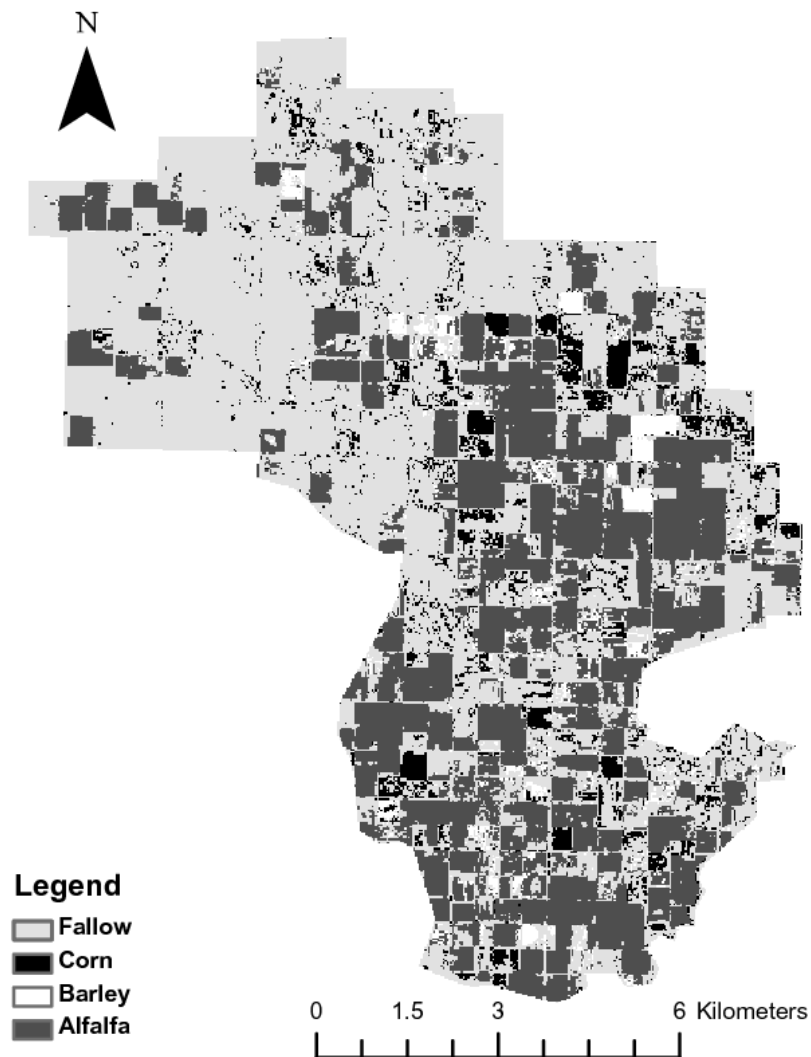


Fig. 2.10. Agricultural crops in Canal B – 2009.

With this additional information, a comparison of the best approach developed against the actual crop ET values for Canal B was performed. The results of this new comparison are presented in Fig. 2.11.

As it is shown, the results in the last figure indicate a small underestimation by the Indirect Approach using the MVRVM model during the peak season for 2009 when compared to the actual crop ET for the forecasting 4 days interval as in shown in the Fig.

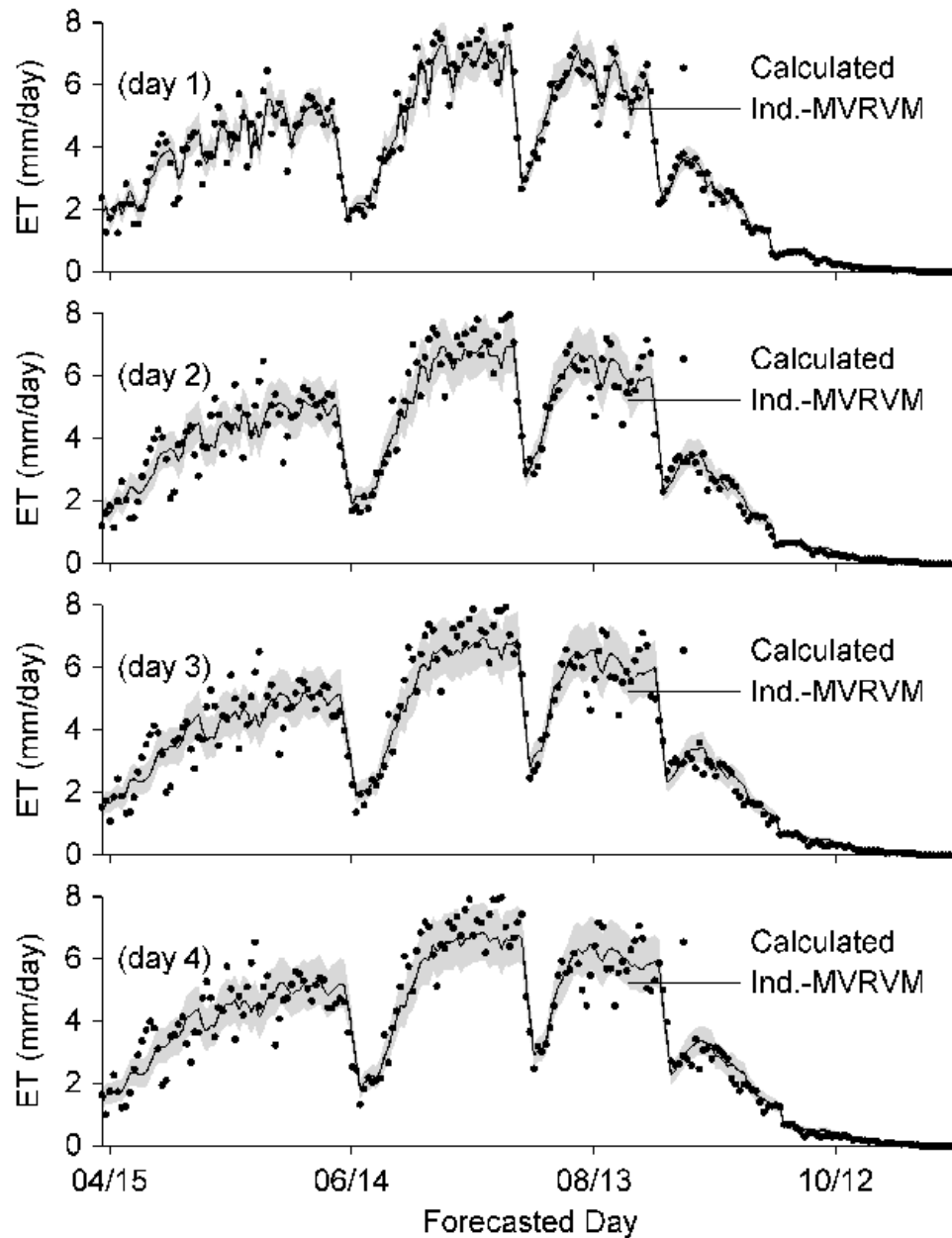


Fig. 2.11. Best approach ET forecasting performance.

2.11. Nevertheless, the good performance of the forecast results obtained by this Approach is demonstrated again by the estimates of forecasted ET along the irrigation season, given that an estimated reference for the short term forecasted water demand required for the crops in Canal B is not currently available. Therefore, the Indirect

Approach using the MVRVM provided better performance, being also the most robust and stable model among the others developed in this study for the time interval of 7 days considered.

Conclusions and discussion

The present study demonstrates the adequacy of forecasting near term daily ET_0 information necessary for water management purposes based on the 1985 Hargreaves ET_0 equation. Two approaches were tested using the Multivariate Relevance Vector Machine algorithm. The first approach, Direct Approach, involves the estimation of ET_0 time series from historical data. The second approach, Indirect Approach, considers forecasting the required climatic data for the 1985 Hargreaves ET_0 equation, daily maximum and minimum air temperatures using the learning machine mentioned and later, using these forecasted values, estimate the future ET_0 values. For performance comparison purposes, an Artificial Neural Network model, the Multilayer Perceptron was also applied in both of the proposed forecasting schemes.

The results indicates that using the approaches proposed in this study it is possible to forecast up to 4 days of daily ET_0 ahead in time within a reasonable range for the goodness of fit parameter $\eta \geq 0.8$. Also the specific use of these learning machines provides an additional estimation of the expected variability values for every forecasted day, thus giving an excellent estimation of the accuracy of the forecasted ET_0 .

When comparing the performance of the approaches and the learning machines used, the results obtained in this study indicate that despite the similar performance of the two approaches considered, based on the goodness-of-fit values obtained, the Indirect Approach provides better ET_0 forecasting capabilities for larger time intervals than the

Direct Approach. This outcome was also expected, since the learning machines in this mentioned approach are used to model and forecast only the behavior of the climatic parameters required for the 1985 Hargreaves ET_0 equation, while for the Direct Approach the learning machines are required to model and forecast the combined effect of the trend of the climatic variables plus the Extraterrestrial Radiation component of the Hargreaves equation. Therefore the Indirect Approach procedure can be extended to other ET_0 equation that requires a small number of climatic parameters. Nevertheless, for forecasting ET_0 values based on models that requires a high number of climatic parameters such as Penman-Monteith, the computational time required to perform the methodology used in Indirect Approach could be excessive, being Direct Approach a better and practical option.

The comparison of learning machines, MVRVM and MLP, also indicates that the former one provides more stable and robust results than the latter model, as is demonstrated by the bootstrapping results. Thus, the application of Indirect Approach using the MVRVM proves to be the best among the options considered in this study.

The forecast of several days ahead in time is affected by the level of relationship of the time series value with the past ones. Thus, the precision of the ET_0 forecasted decreases in time. Still, the used learning machines were able to find relationships among the previous past days with the forecasted future values, as demonstrated by the goodness-of-fit parameters (Table 2.2). Also the advantage of using learning machines that includes the Bayesian Inference Method is the additional information about the variability of the forecasted ET_0 values.

Finally, a comparison of the best approach (Indirect Approach) using the MVRVM with the calculated crop ET was performed considering the year 2009. These results confirm again the good performance of the MVRVM using the mentioned Approach, providing a very good approximation to the actual values of crop evapotranspiration for the Canal B location, indicating the usability of the method proposed in this study for water delivery planning purposes.

Futures studies on this topic are related with estimation of near term water balance for the irrigated lands and also with geospatial analysis of water requirements, which can provide information about future water demands to be delivered in the canal system.

References

- Allen, R. G., Pereira, L. S., Raes, D., Smith, M. 1998. Crop evapotranspiration guidelines for computing crop water requirements. *Irrig. and Drain.*, FAO, 56, 300.
- Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. Oxford University Press. Oxford.
- Christiansen, J. E. 1968. Pan evaporation and evapotranspiration from climatic data. *Proc., Amer. Soc. Civil Eng., J. Irrig. Drain. Eng.*, ASCE, 94, 243–265.
- Gill, M. K., Asefa, T., Kemblowski, M. W., McKee, M. 2006. Soil moisture prediction using support vector machines. *American Water Res. Assoc.*, 42, 1033—1046.
- Hargreaves, G. H. 1974. Moisture availability and crop production. *Trans., ASAE*, 18(5), 980–984.
- Hargreaves, G. H., Allen, R. G. 2003. History and evaluation of Hargreaves evapotranspiration equation. *J. Irrig. Drain. Eng.*, ASCE, 129(1), 53–63.
- Khalil, A. F., McKee, M., Kemblowski, M., Asefa, T., Bastidas, L. 2006. Multiobjective analysis of chaotic dynamic systems with sparse learning machines. *Advances in Water Res.*, 29(1), 72 – 88.

- Khan, M., Coulibaly, P. 2005. Streamflow forecasting with uncertainty estimate using bayesian learning for ann. Proc., Intl. Joint Conf. on Neural Networks, Vol. 5, IEEE. 2680–2685.
- Kumar, M., Raghuwanshi, N. S., Singh, R., Wallender, W. W., Pruitt, W. O. 2002. Estimating evapotranspiration using artificial neural network. J. Irrig. Drain. Eng., ASCE, 128, 224–233.
- Lai, L. L., Braun, H., Zhang, Q. P., Wu, Q., Ma, Y. N., Sun, W. C., Yang, L. 2004. Intelligent weather forecast. Proc., Int. Conf. Machine Learning and Cybernetics, 2004., Vol. 7.
- LandSat Imagery Program website, www.landsat.gsfc.nasa.gov, accessed November, 2009.
- MacKay, D. 1992. A practical bayesian framework for backpropagation networks. Neural Computation, 4(3), 448–472.
- McCuen, R. H., Knight, Z., Cutter, A. G. 2006. Evaluation of the Nash–Sutcliffe efficiency index. J. Hydrol. Eng., 11(6), 597–602.
- Nabney, I. T. 2002. NETLAB: Algorithms For Pattern Recognition. Springer-Verlag New York, Inc., New York.
- NOAA - National Climatic Data Center website, www.ncdc.noaa.gov/oa/ncdc.html, accessed November, 2009.
- Pierce, S. G., Worden, K., Bezazi, A. 2008. Uncertainty analysis of a neural network used for fatigue lifetime prediction. Mechanical Systems and Signal Processing, 22(6), 1395 – 1411. Special Issue: Mechatronics.
- Smith, B. A., McClendon, R. W., Hoogenboom, G. 2006. Improving air temperature prediction with artificial neural networks. Intl. J. Computational Intelligence, 3(3), 179–186.
- SRWUA - Sevier River Water Users Association website, www.sevierriver.org, accessed November, 2009.
- Thayananthan, A., Navaratnam, R., Stenger, B., Torr, P., Cipolla, R. 2008. Pose estimation and tracking using multivariate regression. Pattern Recognition Letters, 29(9), 1302–1310.
- Tipping, M., Faul, A. 2003. Fast marginal likelihood maximization for sparse bayesian models. Proc., 9th Intl. Workshop on Artificial Intelligence and Statistics. 3–6.

- Tipping, M. E. 2001. Sparse bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.*, 1, 211–244.
- Trajkovic, S., Kolakovic, S. 2009. Estimating reference evapotranspiration using limited weather data. *J. Irrig. Drain. Eng.*, 135(4), 443–449.
- Tripathi, S., Govindaraju, R. 2007. On selection of kernel parameters in relevance vector machines for hydrologic applications. *Stochastic Env. Res. and Risk Assessment*, 21, 747–764.
- Verdes, P. F., Granitto, P. M., Navone, H. D., Ceccatto, H. A. 2000. Frost prediction with machine learning techniques. *Proc., VI Argentina Congress on Computer Sci.*, 1423–1433.
- Walker, W. R., Stringam, B. L. 1999. Low cost adaptable canal automation for small canals. *ICIC Journal*, 48(3):39-46.
- Walker, W. R., Stringam, B. L. 2000. Canal automation for water conservation and improved flexibility. *Proc., 4th Decennial Nat. Irrig. Symposium*.
- Wright, J. L. 1982. New evaporation crop coefficients. *Irrig. and Drain. Div., ASCE Proc.*, 108, 57–74.
- Yamashita, S., Walker, W. R. 1994. Command area water demands I: validation and calibration of UCA model. *J. Irrig. Drain. Eng., ASCE*, 120(6), 1025–1042.

CHAPTER 3
MACHINE LEARNING APPROACH FOR ERROR CORRECTION OF HYDRAULIC
SIMULATION MODELS²

ABSTRACT

Modernization of today's irrigation conveyance systems typically employs supervisory control and data acquisition (SCADA) technologies to improve system efficiency and management effectiveness. Hydraulic simulation models have proven to be useful tools supporting SCADA systems, particularly when used to develop and test operating rules and detecting sensors malfunctions. Nevertheless the SCADA sensors, flow measurement structures and gate controls are not unconditionally accurate within the relatively harsh environment of the irrigation system. Also fluctuations in power to the sensors, hydraulic transients in the canal, and damped sensor locations create readings that can confuse both human and computer controllers. Also parameters used in simulation models are also equipped with some degree of uncertainty or distortion. One of the major sources of uncertainty is the spatial and temporal distribution of seepage flows. In order to maximize the effectiveness of the SCADA system, accurate and reliable measurement and simulation of discharges, water levels, and position of regulation structures are necessary. Achieving this goal depends on understanding and evaluating the errors and uncertainty associated with both the SCADA readings and the simulation model output. This paper outlines the theoretical combined application of a

² Coauthored by Alfonso F. Torres, Andres M. Ticlavilca, Wynn R. Walker and Mac McKee

statistical learning machine, the Relevance Vector Machine, and a hydraulic simulation model and demonstrates its practical application in an irrigation system in Central Utah.

Introduction

Historically, canal modernization meant rehabilitation to restore a canal to original constructed conditions and to reduce seepage. Rehabilitation generally improves the canal's capability to regulate and control flows with improved structures and water measurement devices. More recently, the concept of canal modernization has been enlarged to include the much wider goal of improving water management within the entire irrigation system. Under this concept rehabilitation may not be part of the project. Nevertheless, an inherent component of today's canal modernization is the mechanization and automation of canals' inlet, outlet, division, and regulation structures. Among the irrigation systems in the US, the most widely used form of canal automation is the supervisory control and data acquisition or SCADA system. Through sensors and telemetry a canal operator can determine the status of the canal in real time and where necessary, remotely actuate changes in the control structure settings to adjust the status to a revised or corrected condition.

Two of the questions that emerge regarding the feasibility and utility of canal modernization are: (1) Will the costs be justified by lower losses; and (2) How should the canal be operated within its real time capability? The primary tool for evaluating these questions is the hydraulic simulation model. Not surprisingly there are extensive investments to develop hydraulic models that can simulate water flow conditions in canals. The linkage between the SCADA system and the hydraulic model is an important factor in improving water management in canal-based irrigation systems. The data stream

from the SCADA system validates and refines the accuracy of the hydraulic model while the hydraulic model evaluates the effects of alternative decisions by the supervisory controller and forecasts system status. A major use of the hydraulic model is the development and testing of operating rules for the SCADA system to follow in managing the canal. The central issue in using hydraulic models is their accuracy.

Experience indicates that there are two main concerns related with the accuracy of hydraulic simulation models. The first is the accuracy of the data used by the model. It is generally assumed that the information describing the canal characteristics (model parameters) as well continuous observations of the system (model variables) contain minimal errors or deviations (Gaussian, white error or noise). For example, a modern SCADA system may be sensing, recording, and transmitting data describing water levels and gate positions every few minutes and may be impacted by sensor lag time or sensitivity of power fluctuations which may not be evaluated by the hydraulic simulation. The second concern is related with the model's numeric approximation to describe the actual physical environment. This is related also with the stability of the numeric approximation under a wide range of operating conditions in the canal and over time. Rosenberry (1990) reports that the accuracy of simulation models degrades over time as physical characteristics change and may not be reflected in the model input, thus the importance of SCADA data to recalibrate and refine the simulation model.

One of the more important conditions for effective hydraulic simulation is the determination of the magnitude and source of input and numerical errors associated with both sources model variables and parameters. Minimizing the impact of these errors on the simulation is not a simple or straightforward task and requires a relationship analysis

(Pebesma et al., 2005). In recent years, new tools have become available to perform mapping analyses that exploit the statistical characteristics of the data. Examples of these can be found in Pebesma et al. (2005), Zechman and Ranjithan (2007), and Thyer et al. (2009). These models are known as statistical data-driven tools or learning machines, which have been used to estimate relationships among complex multidimensional non-linear variables.

For purposes of this study a data-driven tool, the Relevance Vector Machines (RVM), has been used to simulate and reduce the aggregate error caused by model parameters, variables and numeric approximation of a hydraulic simulation model for canal flow control. Data from a SCADA system implemented to operate and manage a feeder canal in Central Utah is used to calibrate the hydraulic model and estimate the aggregate error. The performance of RVM to simulate the aggregate error is compared the results of the application of another learning machine, an artificial neural network-based model called the Multilayer Perceptron model (MLP). Also, the impact of future new data on the data-driven algorithms is tested by application of a bootstrap analysis.

Theoretical development

Saint Venant equations

The Saint-Venant equations express the laws of mass and momentum conservation for one-dimension analysis of open-channel flows can be written as:

$$\frac{\partial Q}{\partial x} + \frac{\partial A}{\partial t} + q(x, t, A, Q) = 0 \quad (3.1)$$

$$\frac{1}{g} \frac{\partial Q}{\partial t} + \frac{\partial}{\partial x} \left(P + \frac{Q^2}{A \cdot g} \right) - A \cdot S_o + D = 0 \quad (3.2)$$

In the continuity equation (Eq. 3.1) Q is the flow rate (m^3/s), A is the flow cross-section area (m^2), x is the longitudinal distance in the direction of flow (m), and t is the elapsed time. The parameter q is the lumped expression of seepage, evaporation and tributary inflows and aggregate model error. The conservation of momentum (Eq. 3.2) is expressed in terms of Q , A , and t , x , variables as defined for Eq. 3.1. Flow depth is represented by y (m) while P is defined as the net hydrostatic pressure acting on a fluid element per unit weight of water (m^2) and D as the drag force, or the product of friction slope and area (m^2). The canal slope is S_o , the top width of the flow cross-section is T (m), and g is the acceleration of gravity (9.807 m/s^2). The friction slope S_f is defined by the Manning Equation:

$$S_f = \frac{Q^2 \cdot n^2}{A \cdot R_h^{2/3}} \quad (3.3)$$

where n is the Manning roughness coefficient and R_h is the hydraulic radius (m). Eq. 3.1 assumes seepage losses, evaporation losses, ungauged inflows, measurement errors by the SCADA system, and volume balance errors associated with the numerical solution of Eqs. 3.1 and 3.2, which can be incorporated into the simulation model as a lumped error term. The largest components of this lumped or aggregate error term are seepage losses which are spatially varied along the canal reach. For instance, seepage depends on discharge, local water table elevations, and the permeability of channel bed materials all of which vary during the irrigation season. The total magnitude of seepage losses can generally be well estimated over large periods of time using inflow-outflow measurements. However, the rates at which these losses occur over the time step interval

used in a hydraulic simulation are difficult to quantify and are thus lumped with other sources of error.

Numerical solution of Saint-Venant equations

A numerical solution of Eqs. 3.1 and 3.2 can be accomplished with a first order, Lagrangian, deformable control volume (DCV) solution introduced by Strelkoff and Katapodes (1997). Haie (1984) converted the DCV to an Eulerian form which was then detailed for surface irrigation by Walker and Skogerboe (1987).

When Eq. 3.1 is integrated using the DCV solution the correspondent Eulerian terms can be written as:

$$\frac{\partial Q}{\partial x} = \frac{[\theta \cdot (Q_L - Q_R) + (1 - \theta) \cdot (Q_J - Q_M)]}{\partial x} \quad (3.4)$$

$$\frac{\partial A}{\partial t} = \frac{[\varphi \cdot (A_L - A_J) + (1 - \varphi) \cdot (A_R - A_M)]}{\partial t} \quad (3.5)$$

where θ and φ are temporal and spatial averaging coefficients respectively (no dimensional). The corresponding integration for Eq. 3.2 is:

$$\frac{\partial q}{\partial t} = \frac{[\varphi \cdot (q_L - q_J) + (1 - \varphi) \cdot (q_R - q_M)]}{\partial t} \quad (3.6)$$

$$\frac{1}{g} \cdot \frac{\partial Q}{\partial t} = \frac{1}{g} \cdot \left[\frac{\varphi \cdot (Q_L - Q_J) + (1 - \varphi) \cdot (Q_R - Q_M)}{\partial t} \right] \quad (3.7)$$

$$\begin{aligned} \frac{1}{\partial x} \left[P + \frac{Q^2}{A \cdot g} \right] &= \frac{\theta \cdot \left[\left(P + \frac{Q^2}{A \cdot g} \right)_R - \left(P + \frac{Q^2}{A \cdot g} \right)_L \right]}{\partial x} + K \\ &\frac{(1 - \theta) \cdot \left[\left(P + \frac{Q^2}{A \cdot g} \right)_M - \left(P + \frac{Q^2}{A \cdot g} \right)_J \right]}{\partial x} \end{aligned} \quad (3.8)$$

$$-A \cdot S_o = -S_o \{ \theta \cdot [\varphi \cdot A_L + (1 - \varphi) \cdot A_R] + (1 - \theta) \cdot [\varphi \cdot A_J + (1 - \varphi) \cdot A_M] \} \quad (3.9)$$

$$D = \theta \cdot [\varphi \cdot D_R + (1 - \varphi) \cdot D_R] + (1 - \theta) \cdot [\varphi \cdot D_J + (1 - \varphi) \cdot D_M] \quad (3.10)$$

The selection of θ and φ for the DCV solution is of special interest given that their values affect the stability of the hydraulic model. Fread (1974) indicates that the higher the φ and θ values, the more unstable the numerical solution, recommending using φ and θ values close to 0.55. Later, Chaudhry (1993) reduces the φ and θ range to 0.6 to 0.7, being the suggested value 0.6 for both of the parameters, which are also used in the present model.

Model error

Zechman and Ranjithan (2007) classify the errors that affect simulation models into two groups. The first group is the parameter error (ϵ_p) that occurs when only partial or incomplete information of the attributes of the real system is available. A typical error in canal modeling in this classification is the fixed parameters in the model. The second group is the structural error (ϵ_s) associated with the inaccuracies in the model due to non-modeled processes, for example incorrect hypotheses and simplifications. In modeling canal hydraulics using SCADA data, it is also necessary to consider a third error group called input or observation error (ϵ_o) which involves the distortion or noise in measured variables or observations required for the simulation model. This third error is also called measurement error (Chesner, 1991). The combined effect of these errors in the model results produces a simulation error (ϵ_{sim}) which can be approximated by comparing the simulated and measured model output. Inside the model, error sources can be aggregated into a single term, called aggregated error (ϵ_A) which in this study is the q variable in Eq.

3.1. A schematic description of a generic model plus the errors sources and the simulation error is shown in Fig. 3.1.

The combined effect of the error sources can mislead management decisions and provide wrong information to the computational models. For instances, Rosenberry (1990) discusses the effect of the sensor error on interpretation of long term water-level data for groundwater management purposes, claiming that the differences between the real and the measured value obtained from the sensors used to monitor water level in wells could be unpredictable in some cases and constant in others.

An obvious concern about the error sources is the unfeasibility to determine a priori their magnitude and effect on the model results to provide individual measurement corrections. Furthermore, it is the effect of the aggregate error the one noticeable on the simulation results and not the ones produced by each error source per se. This indicates the limitation to determine without further analysis of the aggregate error, the possible sources for its occurrence. Also the impact of the aggregate error on a certain model could be more or less noticeable on the results depending on factors mentioned before such as model numerical algorithm, error imbedded in the parameters and variables of the model, etc. Nevertheless, it is because of the differences among the simulation values and the actual results that the effect of the aggregate error is perceptible and correction of it may be required.

In irrigation canal-fed systems, certain steps are commonly followed to determine the seasonal performance of the system by evaluating maintenance deficiencies, improving operations practices, and others (Skogerboe and Merkley, 1996). These steps involve an assessment of the operations activities and system hydraulic performance.

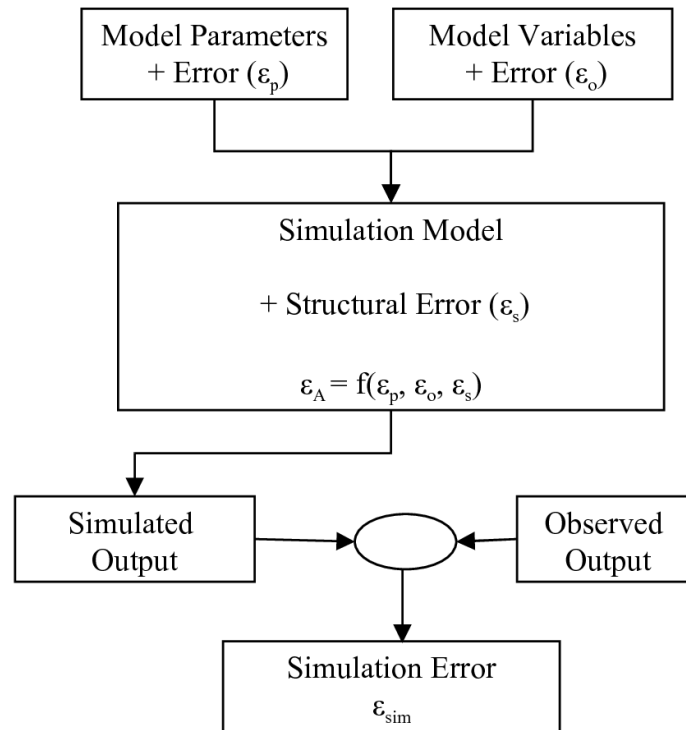


Fig. 3.1. Simulation model and error sources.

While these are very important and highly recommended steps to identify areas that interfere or reduce the seasonal performance of the system, still some error sources in smaller time intervals cannot be thoroughly addressed by these practices. Thus, it is important to develop approaches that allows for aggregate error minimization.

In order to minimize the impact of the aggregate error various approaches have considered each type of error in the model and introduced correction values for the structural error (Zechman and Ranjithan, 2007). Also, error analysis of the model performance to determine the possible error sources and error modeling can be applied as mentioned by Pebesma et al. (2005). A complete approach is limited by number of parameters and variables in the model, complexity of the model and targeted time interval

for the simulation. This limitation confines the use of these developed approaches on small time interval operations (hours).

Statistical learning machines

Statistical models appeared as new tools to identify and model processes on which mechanistic or physical-based approaches present difficulties (e.g. simulation models). These data-driven tools use the statistical properties of inputs and outputs of the process under study to define relationships among them. Since their inception, these statistical models have demonstrated their valuable use in several areas as hydrology, weather forecasting, remote sensing and others (Khalil et al., 2005, 2006; Asefa et al., 2006, Ticlavilca and McKee, 2010).

Multi-layer perceptron

Artificial Neural Networks (ANNs) have been used extensively for simulation and forecasting in such diverse areas as finances, power generation, water resources and environmental science (Maier and Dandy, 2000). The Multi-Layer Perceptron (MLP) is one of the most widely used ANNs (Nabney, 2002) because of its ability to approximate any smooth function. The MLP architecture can be described as:

$$y^{(n)} = W^{II} \cdot \tanh\left(W^I x^{(n)} + b^I\right) + b^{II} \quad (3.11)$$

where:

$y^{(n)}$: MLP output vector, $y^{(n)}=[y_1, \dots, y_m, \dots, y_M]$,

$x^{(n)}$: input vector $x^{(n)}=[x_1, \dots, x_d, \dots, x_D]$,

W^I, W^{II} : optimized weights for the first and second layer respectively,

$$W^I = [w_{1,1}^I, K, w_{D,NN}^I], \quad W^{II} = [w_{1,1}^{II}, K, w_{NN,M}^{II}]$$

M: number of components of the output vector,

D: number of components in the input vector,

NN: number of hidden neurons,

b^I, b^{II} : bias vectors for the first and second layer respectively.

Using a dataset $\Lambda = \{x^{(n)}, t^{(n)}\}_{n=1}^N$, where N is the number of training cases, the calibration of the MLP is performed by optimizing the network parameters $W = \{W^I, W^{II}, b^I, b^{II}\}$ in order to minimize the Overall Error Function E (Bishop, 1995):

$$E = \frac{\beta}{2} \sum_{n=1}^N \left(t^{(n)} - y^{(n)} \right)^2 + \frac{\alpha}{2} \sum_{i=1}^W W_i^2 \quad (3.12)$$

$$E = \beta \cdot E_{\Lambda} + \alpha \cdot E_W$$

where:

E_{Λ} : data error function,

E_W : penalization term,

W: number of weights and biases in the neural network, and

α and β : Bayesian hyperparameters.

In Bayesian terms, the goal is to estimate the probability of the weights and bias of the MLP model, given the dataset Λ :

$$p(W | t^{(n)}) = \frac{p(t^{(n)} | W) p(W)}{p(t^{(n)})} \quad (3.13)$$

where, as explained by MacKay (1992):

$p(\mathbf{W}|\mathbf{t}^{(n)})$: the posterior probability of the weights,

$p(\mathbf{t}^{(n)}|\mathbf{W})$: the dataset likelihood function,

$p(\mathbf{W})$: the prior probability of the weights, and

$p(\mathbf{t}^{(n)})$: the evidence for the dataset.

Assuming a Gaussian distribution for the error term $\xi^{(n)} = \mathbf{t}^{(n)} - \mathbf{y}^{(n)}$ and the weights \mathbf{W} , the likelihood and the prior probabilities can be expressed:

$$p(\mathbf{t}^{(n)} | \mathbf{W}, \beta) = (2\pi\beta^{-1})^{-N/2} \exp(-\beta E_{\Lambda}) \quad (3.14)$$

$$p(\mathbf{t}^{(n)} | \mathbf{W}, \alpha) = (2\pi\alpha^{-1})^{-N/2} \exp(-\alpha E_{\mathbf{W}}) \quad (3.15)$$

E_{Λ} models the uncertainty (or error) of the target variables as Gaussian zero-mean noise and variance $\sigma^2 \equiv \beta^{-1}$. $E_{\mathbf{W}}$ defines the conditional probability of \mathbf{W} with variance $\sigma_{\mathbf{W}}^2 \equiv \alpha^{-1}$. Then Eq. 3.13 can be expressed as:

$$p(\mathbf{W} | \mathbf{t}^{(n)}, \alpha, \beta) = \frac{p(\mathbf{t}^{(n)} | \mathbf{W}, \beta) \cdot p(\mathbf{W} | \alpha)}{p(\mathbf{t}^{(n)} | \alpha, \beta)} \quad (3.16)$$

$$p(\mathbf{W} | \mathbf{t}^{(n)}, \alpha, \beta) = \frac{\exp\left(\mathbf{E}\left(\mathbf{W}^*\right) - \frac{1}{2} \Delta \mathbf{W}^T \mathbf{H} \Delta \mathbf{W}\right)}{\exp\left(\mathbf{E}\left(\mathbf{W}^*\right) \cdot (2\pi)^{W/2} |\mathbf{H}|^{-1/2}\right)} \quad (3.17)$$

In which,

$\mathbf{E}\left(\mathbf{W}^*\right)$: expected optimized values for the weights and bias,

\mathbf{H} = Hessian matrix $\mathbf{H} = \beta \nabla \nabla E_{\Lambda} + \alpha \mathbf{I}$, \mathbf{I} is the identity matrix.

$\Delta \mathbf{W} = \mathbf{W} - \mathbf{W}^*$.

Once the distribution of W has been estimated by maximizing the likelihood for α and β , the prediction $y^{(n)}$ and its standard deviation $\sigma_y^{(n)}$ can be estimated by integrating (marginalizing) over W and the regularization parameters α and β (Bishop, 1995):

$$p(y^{(n)} | x^{(n)}, t^{(n)}) = \int p(t^{(n)} | x^{(n)}, \mathbf{W}^*) p(\mathbf{W}^* | t^{(n)}) \cdot d\mathbf{W} \quad (3.18)$$

This can be approximated by:

$$p(y^{(n)} | x^{(n)}, t^{(n)}) \propto \left(2\pi\sigma_y^{(n)2} \right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}\sigma_y^{(n)2} \left(y^{(n)*} - t^{(n)} \right)^2 \right) \quad (3.19)$$

where $y^{(n)}$ is the output and $\sigma_y^{(n)2}$ is the output variance from the MLP. The output variance can be expressed as:

$$\sigma_y^{(n)2} = \beta^{-1} + \mathbf{g}^T \mathbf{H}^{-1} \mathbf{g} \quad (3.20)$$

\mathbf{g} denotes the gradient of $y^{(n)}$ with respect to the weights; $\mathbf{g} \equiv \nabla W y^{(n)} | \mathbf{W}^*$. The output variance has then two sources; the first arises from the intrinsic noise in the target data; and the second from the posterior distribution of the ANN weights (Pierce et al., 2008). The output standard deviation vector $\sigma_y^{(n)}$ can be interpreted as the error bar for confidence interval estimation (Bishop, 1995).

Relevance vector machine

Tipping (2001) introduced the Relevance Vector Machine (RVM), a Bayesian approach for classification and regression models. As in the case for the MLP, its use in engineering topics is increasing over time (Ghosh and Mujumdar, 2008). The development of the RVM concept is developed as follows; given a training data set of input-target vector pairs $\{x_n, t_n\}_{n=1}^N$, where N is the number of observations; the model

has to learn the dependency between input and output target with the purpose of making accurate predictions of t for previously unseen values of x :

$$\begin{aligned} t &= y + \varepsilon \\ t &= \Phi(x) w + \varepsilon \end{aligned} \quad (3.21)$$

where w is a vector of weight parameters and $\Phi(x) = [1, K(x, x_1), \dots, K(x, x_N)]$ is a design matrix where $K(x, x_n)$ is a fixed kernel function. The error ε is conventionally assumed to be zero-mean Gaussian with variance σ^2 . A Gaussian likelihood distribution for the target vector can be written as:

$$p(t | w, \sigma^2) = (2\pi)^{-N/2} \sigma^{-N} \exp\left\{-\frac{\|t - y\|^2}{2\sigma^2}\right\} \quad (3.22)$$

Tipping (2001) proposed imposing an additional prior term to the likelihood or error function to avoid that the maximum likelihood estimation of w and σ^2 suffer from severe over-fitting from Eq. 3.22. This prior is added by applying a Bayesian perspective, and thereby constraining the selection of parameters by defining an explicit zero-mean Gaussian prior probability distribution over them:

$$p(w | \alpha) = (2\pi)^{-M/2} \prod_{m=1}^M \alpha_m^{1/2} \exp\left(-\frac{\alpha_m w_m^2}{2}\right) \quad (3.23)$$

where M is the number of independent hyperparameters $\alpha = (\alpha_1, \dots, \alpha_M)^T$. Each α is associated independently with every weight to moderate the strength of the prior (Tipping, 2001) and to control the generalization ability of the model (Khalil et al., 2006). Bayesian inference considers the posterior distribution of the model parameters, which is given by the combination of the likelihood and prior distributions:

$$p(w | t, \alpha, \sigma^2) = \frac{p(t | w, \sigma^2) p(w | \alpha)}{p(t | \alpha, \sigma^2)} \quad (3.24)$$

The posterior distribution is Gaussian $N(w|\mu,\Sigma)$ with covariance $\Sigma = (A + \sigma^{-2} \Phi^T \Phi)^{-1}$ and mean $\mu = \sigma^{-2} \Sigma \Phi^T t$; where A is defined as $\text{diag}(\alpha_1, \dots, \alpha_M)$. An optimal set of hyperparameters α^{opt} can be obtained by formulating the maximization of the marginal likelihood with respect to α (Tipping, 2001). The marginal likelihood is then given by its logarithm $L(\alpha)$:

$$L(\alpha) = \log p(t | \alpha, \sigma^2) = \log \int_{-\infty}^{\infty} p(t | w, \sigma^2) p(w | \alpha) dw ,$$

$$L(\alpha) = -\frac{1}{2} [N \log 2\pi + \log |C| + t^T C^{-1} t] \quad (3.25)$$

where $C = \sigma^2 I + \Phi A^{-1} \Phi^T$. The optimal set of hyperparameters α^{opt} and noise parameters $(\sigma^{\text{opt}})^2$ are obtained by maximizing the marginal likelihood using the fast marginal likelihood maximization algorithm proposed by Tipping and Faul (2003). During the optimization process many elements of α go to infinity, for which the posterior probability of the weight becomes zero. The few nonzero weights are the relevance vectors (RVs) which generate a sparse representation. The optimal parameters are used to obtain the optimal weight matrix with optimal covariance Σ^{opt} and mean μ^{opt} . Given a new input x^* , we can compute the predictive distribution for the corresponding target t^* (Tipping, 2001):

$$p(t^* | t, \alpha^{\text{opt}}, (\sigma^{\text{opt}})^2) = \int p(t^* | w, (\sigma^{\text{opt}})^2) \cdot p(w | t, \alpha^{\text{opt}}, (\sigma^{\text{opt}})^2) dw \quad (3.26)$$

Taking into consideration that both terms in the integrand are Gaussian, Eq. 3.26 is computed as:

$$p(t^* | t, \alpha^{\text{opt}}, (\sigma^{\text{opt}})^2) = N(t^* | y^*, (\sigma^*)^2) \quad (3.27)$$

where y^* is the predictive mean and $(\sigma^*)^2 = [(\sigma_1^*)^2, \dots, (\sigma_r^*)^2, \dots, (\sigma_M^*)^2]^T$ is the predictive variance with $(\sigma^*)^2 = (\sigma^{\text{opt}})^2 + \Phi(x^*)^T \Sigma^{\text{opt}} \Phi(x^*)$ which contains the sum of two variance

terms: the noise on the data and the uncertainty in the prediction of the weight parameters (Tipping, 2001). The standard deviation σ^* of the predictive distribution is defined as a predictive error bar of y^* (Bishop, 1995). Readers interested in greater detail regarding sparse Bayesian regression, its mathematical formulation and the optimization procedures of the model are referred to Tipping (2001) and Tipping and Faul (2003).

Material and methods

Site description

The water resources of the Sevier River Basin in Central Utah (Fig. 3.2) are among the most heavily utilized in the Western US. Substantial efforts to increase efficiency via canal lining and on-farm improvement such as conversion to sprinkler irrigation and laser land leveling were made during 1960 - 1990 period. From 1990 to the present, all reservoirs and stream offtakes have been equipped with SCADA technology and web-based data summaries (SRWUA, 2009). Canal automation was introduced in 1994 and shown not only to result in substantial reduction in losses but also to considerably shorten the response time between farmer demands and system deliveries (Walker and Stringam, 1999, 2000). Most recently, attention has been focused on improving the coordination between farmer demands, canal deliveries, and reservoir diversions.

In order to develop, test and implement the model proposed of this study, a subsystem at the lower end of the Sevier River was selected. The agricultural command areas (ACAs) are connected to the DMAD Reservoir by a 9 km. canal (Canal A). The

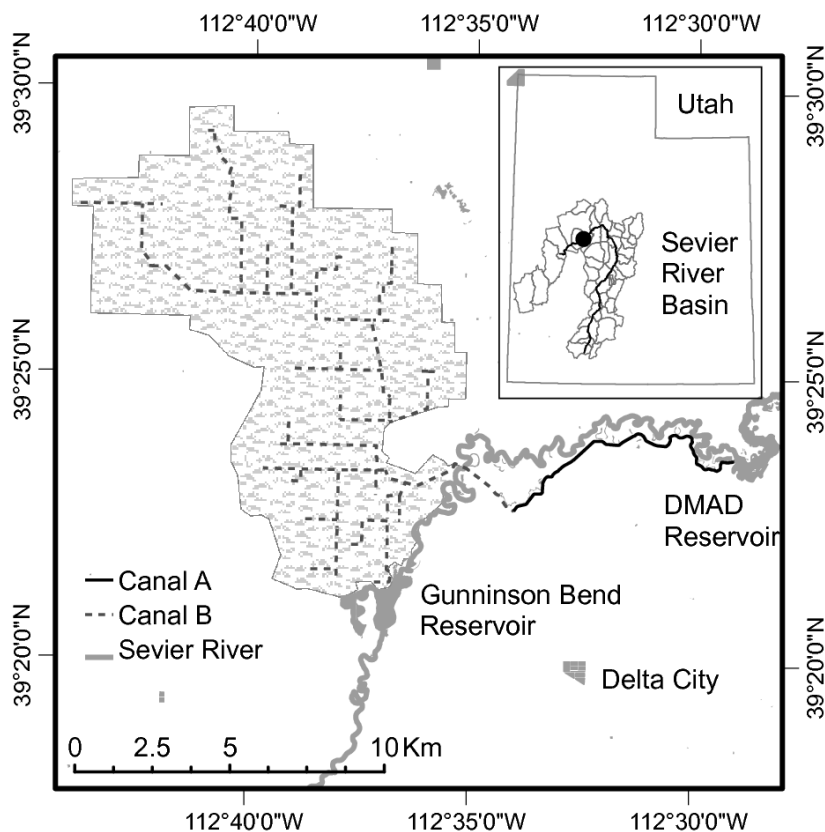


Fig. 3.2. Canal A location, Delta, Utah.

DMAD gates as well as the Canal A gates are automated and operated as a SCADA system by local water masters.

The DMAD Reservoir is supplied water on a demand basis from Sevier River Bridge Reservoir located upstream. The lag time from Sevier Bridge Reservoir to DMAD Reservoir is 3 days. Thus, an emerging water demand in any ACA can be supplied within conveyed from Sevier Bridge Reservoir. Thus, the goal of the entire system is to provide water to an individual farm within 12 hours of an order by the irrigator. This goal relies heavily on the SCADA system and the regulation capacity of DMAD Reservoir.

The water management goals over the next few years are to increase the DMAD regulation capacity and improve the reliability of the 12-hour delivery interval period. It

is expected that by better control of the DMAD Reservoir level will reduce seepage, evaporation and administration losses by about 25 to 50%. The most important capability needed to achieve this goal is to develop a reliable and accurate forecast of irrigation demand, which is related with better or improved models to manage water allocation.

Data acquisition

The information used for this study was collected from two sources. The Sevier River Water User Association website (SRWUA, 2009) provided data on water levels and discharge for years 2008 and 2009. Information about the hydraulic characteristics of Canal A such as Manning's roughness coefficient, channel slope and cross section dimensions were obtained from previous field work in the area under study (Walker and Stringam, 1999, 2000).

Hydraulic simulation model

A hydraulic model was developed from Eqs. 3.4 to 3.10 and applied to the hydraulic conditions in Canal A. The configuration of the hydraulic model is presented in Table 3.1.

As discussed above, the error sources in the hydraulic model can be grouped into: (1) parameter error sources (ϵ_p) such as spatial and time averaging coefficients, distance and time steps considered and canal roughness coefficient; (2) observation error sources (ϵ_o) such as measured inflow and outflow rates; and (3) the structural error source (ϵ_s) associated with the numerical approximation of Eqs. 3.1 and 3.2, and primarily attributed to the values of q determined by the model. To provide a solution that takes into account the lumped effect of the multiple error sources, the aggregate error (ϵ_a) in the hydraulic

Table 3.1. Hydraulic simulation model for Canal A.

Model Characteristics	Value
Approximation:	Saint- Venant - Deformable Control Volume
Parameters:	Canal length : 9 km Hydraulic Area $A=9.12y^{1.427}$ (m ²) Manning Eq. $A^2R=0.073A^{2.943}$ (m ³) Bed slope: 0.00011 Canal roughness coefficient: 0.018 Maximum flow : 12 m ³ /s Distance / time steps used: 1m/1hr Spatial averaging coefficient (θ): 0.6 Time averaging coefficient (ϕ): 0.6
Variables:	Inflow rate at canal head (Q_{in}) 1hr values from SCADA Outflow rate at canal end (Q_{out}) 1hr values from SCADA On-Demand Variation (OD_v) 1hr values from SCADA
Output:	Canal water level (h) 1hr values from SCADA

model is determined by forcing the actual water levels in Canal A to equal the simulated values and then equating the volume balance adjustment to the parameter q in Eq. 3.1.

Thus,

$$\varepsilon_A = q(x, t, A, Q) \quad (3.28)$$

where ε_A shares the same units as the other components of Equation 3.1. Here it is important to indicate that it is very difficult to determine the individual contribution,

order and magnitude of the canal seepage and the other error sources. As mentioned before, the effect of the aggregate error is the only one noticeable and feasible of being measured. Hence, it is very difficult to establish if any error source e.g. ε_o has the same magnitude and impact than the seepage losses in the canal for this study.

The data-driven algorithms were applied to determine the relationship of the aggregate error ε_A with other model variables available in order to minimize ε_A , and its impact in the simulation results. The main requirement of the application of these algorithms for the ε_A minimization problem is a real-time capability to work in a coupled mode with the hydraulic simulation model.

Learning machines

The data-driven algorithm selected for the ε_A correction model is the RVM. For performance comparison the MLP was also tested. The proposed error correction model considers a combination of the hydraulic simulation model to estimate the water levels in Canal A with a machine learning model to estimate the aggregated error that comprises the error sources. The proposed model approach is presented in Fig. 3.3.

Two testing criteria are used to evaluate the results of the ε_A modeling: the Root Mean Square Error (RMSE); and the Nash-Sutcliffe Efficiency Index (η). The Nash-Sutcliffe Efficiency Index use is recommended for non-linear modeling problems along with other statistical indicators (McCuen et al., 2006).

$$\text{RMSE} = \sqrt{\sum_{n=1}^{N_*} (y_* - t_*)^2 / N_*} \quad (3.29)$$

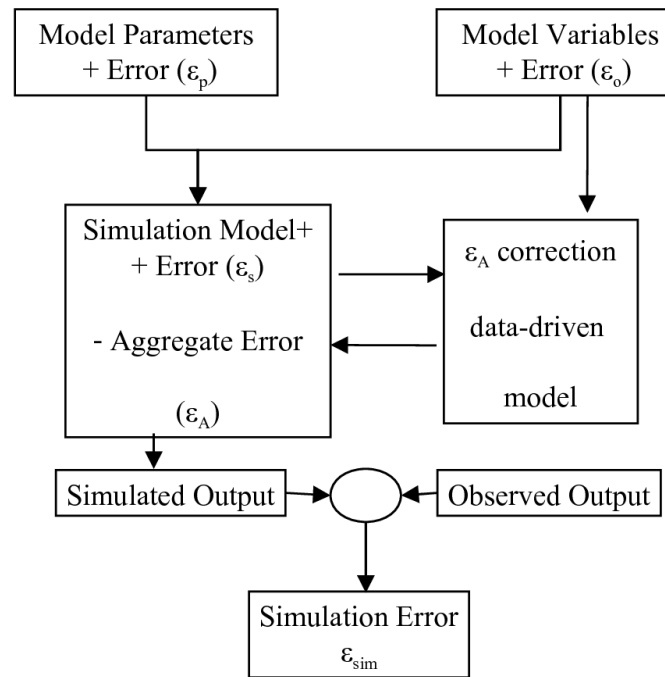


Fig. 3.3. Simulation + error correction model.

$$\eta = 1 - \frac{\sum_{n=1}^{N_*} (y_* - t_*)^2}{\sum_{n=1}^{N_*} (t_* - \bar{t}_*)^2} \quad (3.30)$$

where t_* actual values for the testing data, y_* forecasted values of for the testing data, N_* is the number of samples or cases in the testing data, and \bar{t} the average value. The RMSE value allows ranking the performance of each learning machine, being large RMSE value an indication that the error between the calculated and simulated aggregate error is large too. The value of η measures in a non-dimensional range (from $-\alpha$ to 1) the closure of the calculated vs. the simulated aggregate error values, being $\eta = 1$ an

indication of perfect correspondence. A η value of 0 indicates that the simulated aggregated error is not better than the average of the aggregate error values.

To determine the robustness of the learning machines in presence of unseen data, a bootstrapping analysis of the goodness-of-fit parameters allows a comparison and selection of the best data-driven model.

Results

Hydraulic model performance

The hydraulic model used to simulate the flow conditions in Canal A was calibrated with the data described in Table 3.2. A graphical description of the data used is presented in Figs. 3.4 to 3.6.

These figures show the ε_A behavior pattern along the irrigation seasons (2008, 2009) and the water levels and discharge in Canal A. ε_A presents a strong correlation with both Canal A variables, especially with the flow rate values, despite ε_A seems to be random as shown in Fig. 3.6c. Fig. 3.6 presents ε_A statistical characteristics, such statistical distribution, autocorrelation pattern and relationship with the Canal A flow rate.

Considering that ε_A values should comply with Gaussian or white error characteristics such as normal, independent and identically distributed (NIID) set, departure from this assumption indicates that ε_A contains an imbedded structure that is not accounted for in the hydraulic model. ε_A statistics resemble to a normal distribution with mean = $1.4 \cdot 10^{-3} \text{ m}^3/\text{m}/\text{hr}$ and standard deviation = $8.86 \cdot 10^{-2} \text{ m}^3/\text{m}/\text{hr}$ (Fig. 3.6 a). Nevertheless, the Partial Correlation analysis (Fig. 3.6 b) indicates that the values are highly correlated with the immediate past value. Identifying the structure imbedded in the

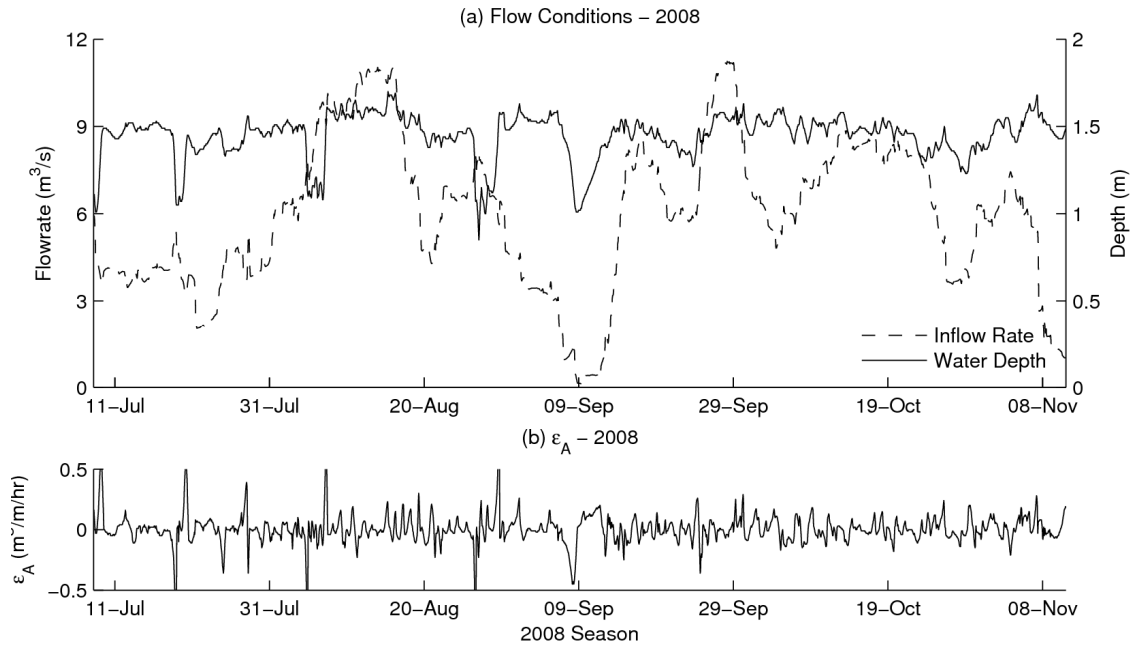


Fig. 3.4. Observed water levels, discharge and aggregate error (ϵ_A) for 2008.

Table 3.2. Variables tested for the aggregate error correction model.

Variable	Units	Symbol
Canal Inflow	cms	Q_{in}
Canal Water Depth	m	h_{in}
Water On-demand	cms	OD_v
Aggregate Error time series	cms	$\epsilon_{A(t-1)} : \epsilon_{A(t-n)}$; n: number of hours in the past

aggregated error and subtracting it to the hydraulic simulation model can lead to a better approximation of the simulation results. Hence, the incorporation of the error correction model pursues this objective, looking for a relationship between the inputs of the hydraulic model and the aggregate error present in the hydraulic simulation model.

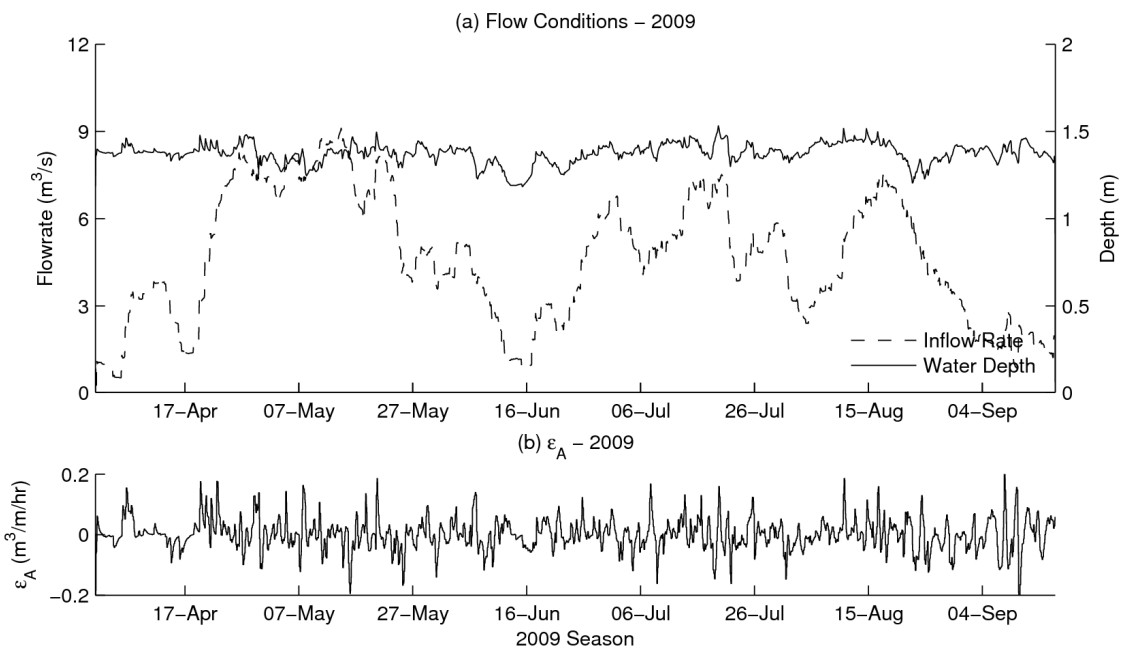


Fig. 3.5. Observed water levels, discharge and aggregate error (ϵ_A) for 2009.

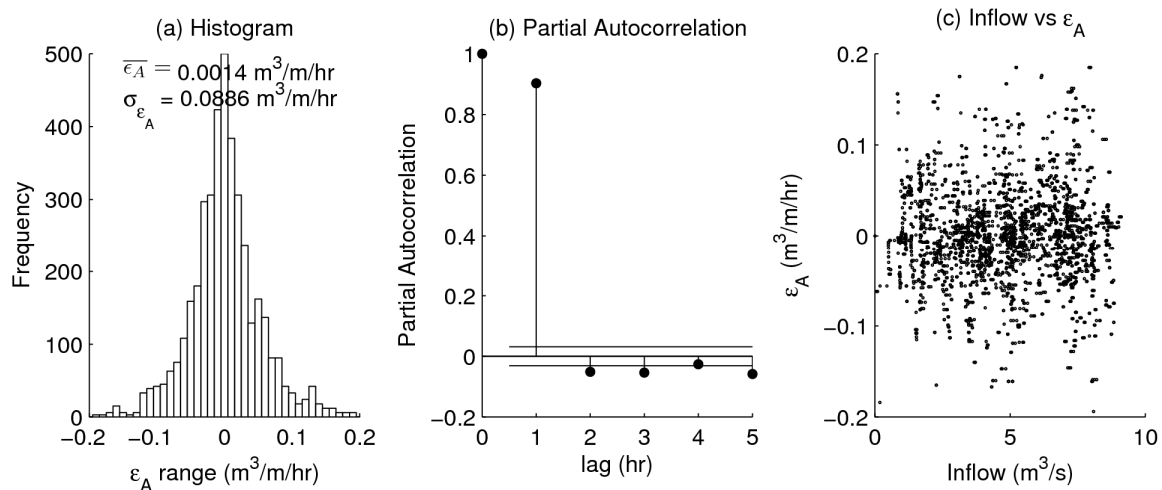


Fig. 3.6. ϵ_A statistics for 2009 irrigation season.

Error modeling

To develop an adequate ε_A correction model using data-driven algorithms, it is necessary to determine which variables required by the hydraulic model have the strongest relationship with ε_A . There is not a straightforward solution to this issue, given the possible synergy effect of two or more variables over ε_A . Therefore, several steps were followed to determine the best ε_A model: Given the limited data available for this study (2 irrigation seasons), the data was separated in two groups. The 2008 year data was used for training and calibrating the data-driven algorithms while the data for 2009 was used to verify the adequacy of the tuned algorithm with the 2008 data. There are four possible variables to be included in the ε_A correction model, so several variable combinations were tested for the MLP and RVM. To define the best variable - ε_A set a variable ranking procedure or stepwise forward variable selection was followed as recommended by Guyon and Elisseeff (2003), rating every tested variables combination by the goodness-of-fit parameters values including the data-driven error bar (σ_y) and visual analysis.

Defined the variables to be included in the ε_A correction model (Table 2), the data-driven algorithms were fine-tuned with the data available for this study (2008 and 2009). The parameter to calibrate for the MLP was the number of neurons in one hidden layer. For the RVM model, the parameter to calibrate was the kernel width (σ_m). The results of this procedure are presented in Table 3.3 and 3.4.

As shown in the mentioned tables, the variables from the hydraulic simulation model demonstrated that the strongest relationship in the error correction model includes $\varepsilon_{A(t-1)}$ and OD_v using MLP with 1 neuron located in the hidden layer. Similarly, for the

Table 3.3. Variables included in ε_A correction model using MLP and goodness-of-fit obtained for test data (2009).

Variable Combination	Hidden Neurons	RMS (m ³ /m/hr)	η	σ_y (m ³ /m/hr)
$\varepsilon_{A(t-1)}$	9	0.0249	0.813	0.035
$\varepsilon_{A(t-1)}, \varepsilon_{A(t-2)}$	8	0.0290	0.747	0.027
$\varepsilon_{A(t-1)}, \varepsilon_{A(t-2)}, \varepsilon_{A(t-3)}$	1	0.0312	0.706	0.027
$\varepsilon_{A(t-1)}, OD_v^*$	1	0.0248	0.815	0.034
$\varepsilon_{A(t-1)}, Q_{in}$	1	0.0249	0.813	0.035
$\varepsilon_{A(t-1)}, h_{in}$	1	0.0249	0.813	0.035
$\varepsilon_{A(t-1)}, OD_v, Q_{in}$	1	0.0248	0.815	0.034
$\varepsilon_{A(t-1)}, OD_v, h_{in}$	1	0.0248	0.814	0.034
$\varepsilon_{A(t-1)}, OD_v, h_{in}, Q_{in}$	1	0.0248	0.814	0.034

* Best variable combination obtained

Table 3.4. Variables included in ε_A correction model using RVM and goodness-of-fit obtained for test data (2009).

Variable Combination	σ_m	RMS (m ³ /m/hr)	η	σ_y (m ³ /m/hr)
$\varepsilon_{A(t-1)}$	0.4	0.0245	0.819	0.035
$\varepsilon_{A(t-1)}, \varepsilon_{A(t-2)}$	2.8	0.0256	0.802	0.036
$\varepsilon_{A(t-1)}, \varepsilon_{A(t-2)}, \varepsilon_{A(t-3)}$	3.4	0.025	0.812	0.046
$\varepsilon_{A(t-1)}, OD_v$	2.4	0.0248	0.815	0.035
$\varepsilon_{A(t-1)}, Q_{in}^*$	1	0.0245	0.820	0.035
$\varepsilon_{A(t-1)}, h_{in}$	0.8	0.0247	0.816	0.035
$\varepsilon_{A(t-1)}, OD_v, Q_{in}$	3	0.0246	0.818	0.035
$\varepsilon_{A(t-1)}, OD_v, h_{in}$	3.2	0.025	0.812	0.035
$\varepsilon_{A(t-1)}, OD_v, h_{in}, Q_{in}$	3.7	0.0247	0.817	0.035

* Best variable combination obtained

RVM, the variables from the hydraulic model for the error correction model were $\varepsilon_A(t-1)$ and Q_{in} , using a Gaussian kernel with a kernel width of 1.0 configuration for the data-driven algorithm (Table 3.4). The $\varepsilon_{A(t-1)}$ variable is common for both of the algorithms and has a great impact on modeling the $\varepsilon_{A(t)}$ pattern and the inclusion of other variables affects in positive or negative form the relationship found by the data-driven algorithms with ε_A .

In the results shown in Tables 3.3 and 3.4, for both mapping algorithms, the $\varepsilon_{A(t-1)}$ variable by itself (autoregressive model) provides good goodness-of-fit statistics for the model error correction. Nevertheless, in an internal test the performance of this single variable to model the aggregate error is less robust than the selected variable combination along the irrigation season.

Figs. 3.7 and 3.9 presents the temporal behavior of the actual and simulated ε_A and the ε_A residual obtained for the 2009 irrigation season using the best calibrated MLP and the RVM respectively. Figs. 3.8 and 3.10 show the statistical characteristics of ε_A also for MLP and RVM. The capability of the developed models to simulate ε_A either using the MLP or the RVM is demonstrated and the accuracy of the models is further detailed in the included subplots.

Both the MLP and RVM algorithms were able to identify and map the variables in the hydraulic simulation model that have the strongest influence and are enough to replicate ε_A behavior, proving that the methodology followed in this study to minimize ε_A is adequate. Thus the new ε_A from the coupled hydraulic simulation – error correction model is the ε_A residuals which as seen in the graphical results have strong differences with the original ε_A .

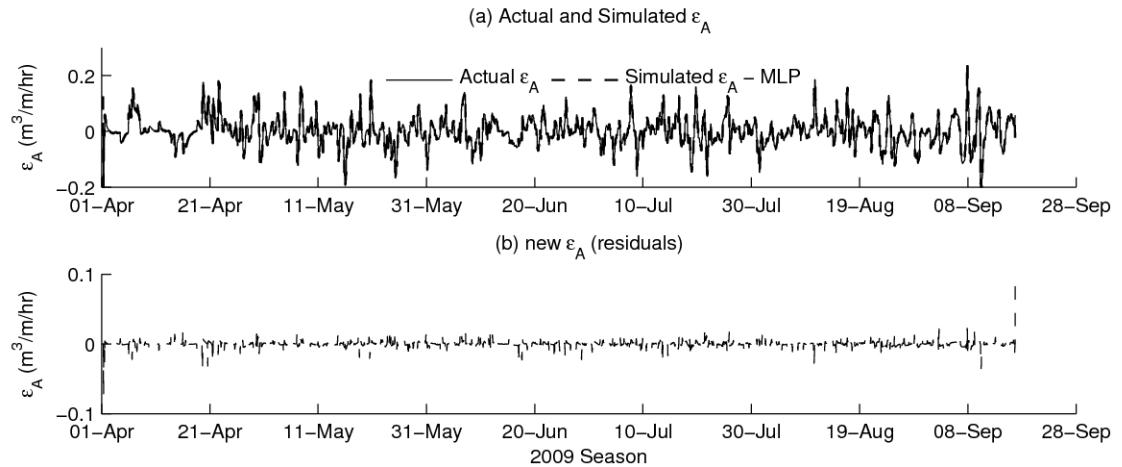


Fig. 3.7. (a) Actual, simulated ϵ_A and (b) ϵ_A residuals obtained using the best MLP error correction model (2009).

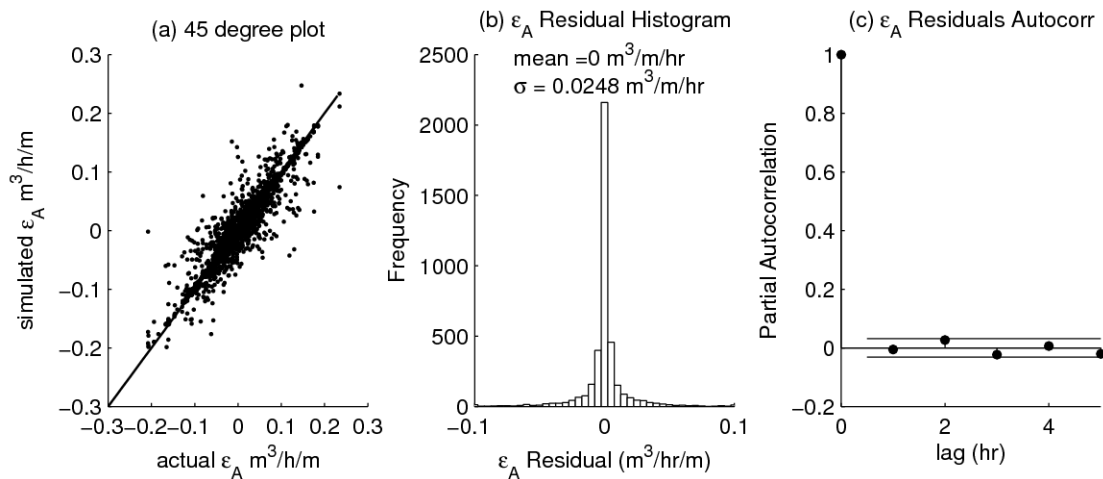


Fig. 3.8. Statistical characteristics of ϵ_A and ϵ_A residuals using the MLP model (2009).

The results obtained in terms of the hydraulic simulation model imply an effective reduction of the simulation error (ϵ_{sim}). The ϵ_A reduction model improves the simulation results, and allows a better correspondence among actual and simulated canal water levels, therefore reducing ϵ_{sim} values. In statistical terms the results demonstrate there is a strong correspondence between the actual and simulated ϵ_A as shown in Figs. 3.8 (a)

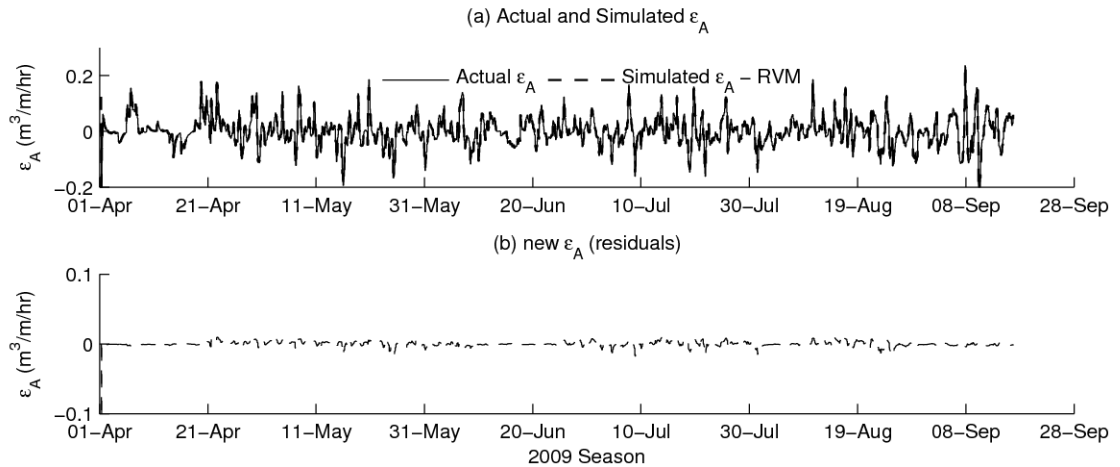


Fig. 3.9. (a) Actual, simulated ϵ_A and (b) ϵ_A residuals obtained using the best RVM error correction model (2009).

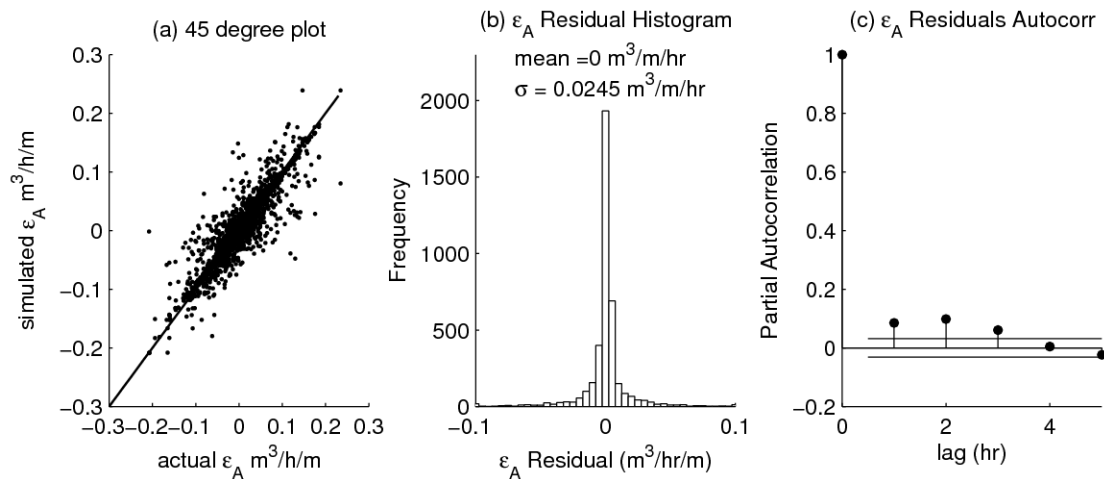


Fig. 3.10. Statistical characteristics of ϵ_A and ϵ_A residuals using the RVM model (2009).

and 3.10 (a), the distribution of the ϵ_A residuals seems to comply better with NIID characteristics. For example for the MLP model results it was obtained a mean of $0 \text{ m}^3/\text{hr}/\text{m}$ and a standard deviation of $2.48 \cdot 10^{-2} \text{ m}^3/\text{hr}/\text{m}$. For the RVM the mean obtained has a mean of $0 \text{ m}^3/\text{hr}/\text{m}$ and $2.45 \cdot 10^{-2} \text{ m}^3/\text{hr}/\text{m}$. These results and the goodness of parameters obtained: $\eta = 0.815$ and $\text{RMS} = 0.0248 \text{ m}^3/\text{hr}/\text{m}$ for the MLP and $\eta = 0.820$

and $0.0245 \text{ m}^3/\text{m}/\text{hr}$ for the RVM. These results indicate that the RVM has a slight advantage over the performance by the MLP model. Figs. 3.8 and 3.10 also give more insight of the ε_A residuals in terms of autocorrelation, which is reduced considerably.

A separate analysis is required for the ε_A residuals autocorrelation values obtained for the MLP and RVM models. As shown in Figs. 3.8 (c) and 3.10 (c) there is a strong reduction of the autocorrelation behavior of the ε_A values for both mapping algorithms. Nevertheless, the MLP model is able to capture in a better fashion the underlying behavior of the aggregate error when compared with the RVM model results. This translates into a better approximation of the ε_A residuals to the white or random noise characteristics mentioned before. This lower performance of the RVM can be explained by the amount of data used in this study (two irrigation seasons). For the error correction model, the RVM algorithm provides better performance than the MLP but requires more information to completely model the behavior of the aggregate error from the hydraulic simulation model.

Also it is important to mention that the ε_A correction model seems to lose precision at the beginning and end of the irrigation season (larger ε_A residual values). This could be related with the strong unsteady flow conditions that occur during the quick filling and drainage of the canal reach. It is in these situations where the largest ε_A residual values occur.

After completing the calibration and testing of the ε_A correction model for both of the data-driven algorithms used here, it is important to determine the suitability of each of the ε_A correction models under different irrigation conditions that the 2008 and 2009 seasons, looking for the stability and robustness of the results provided by the ε_A

correction models. Thus a bootstrapping analysis was performed on each best-configured data-driven algorithm in order to evaluate their predictive power and robustness, as well to estimate the properties of the goodness-of-fit. A 1000-iteration bootstrapping analysis was applied varying the training dataset (2008 data) by random sampling with replacement. After training the learning machine, the testing dataset (2009 data) is used to obtain the goodness-of-fit values as explained by Anguita et al. (2000). For each bootstrap iteration the goodness-of-fit values (RMS and η) for 2009 data were stored. The bootstrap analysis results are presented in Fig. 3.11.

Several conclusions can be drawn from the results of the bootstrapping analysis. First, the range of the goodness-of-fit parameters in the presented histograms is smaller for the RVM when compared with the MLP for both of the goodness-of-fit parameters. This implies that once calibrated, the RVM is less susceptible to the variation of the training data as could occur under actual working conditions. For the MLP, the statistical measures present a non-smooth distribution, indicating that the MLP performance is affected by training data variation. This is because MLP and ANN-based algorithms are more sensitive to the initial weights used by the algorithm. This is a strong limiting condition for practical applications of the ε_A correction error model along the hydraulic simulation model using MLP as mapping algorithm and its use is not recommended. Therefore, based on the results obtained in this study, the proposed coupled hydraulic simulation model – ε_A correction model using RVM as the data-driven algorithm is recommended.

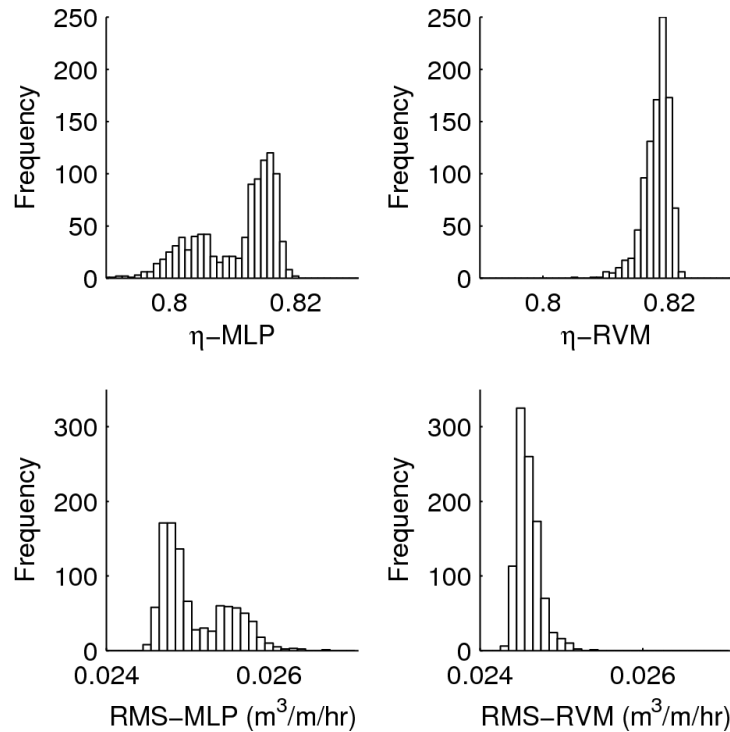


Fig. 3.11. Goodness-of-fit statistics for MLP and RVM error correction models from Bootstrap analysis (2009 data)

Conclusions and discussion

This study presents the findings of a combined application of a statistical learning machine and a hydraulic simulation model to minimize the lumped or aggregate error caused by uncertainties and errors presents in SCADA systems and hydraulic simulation models. The proposed application tests the developed application using hydraulic information from an irrigation canal fed in Central Utah for the years 2008 and 2009.

The aggregate error in this study comprises seepage and lateral flows in the canal reach in an hour basis, the uncertainty imbedded in the SCADA data and errors in the numerical approximation of the hydraulic simulation model. The aggregate error can affect the precision of the results obtained, the water levels in the canal reach, this affecting human and computer controllers.

For the aggregate error correction model, the chosen data-driven algorithm is the Relevance Vector Machine and its performance is compared against the results of another data-driven tool, the Multilayer Perceptron.

The results obtained indicate that the combination of the hydraulic simulation – learning machine model is capable to minimize the aggregate error adequately, capturing its behavior pattern along the irrigation season. This provides means to reduce the aggregate error that ultimately improves the performance of the hydraulic simulation model. The variables from the hydraulic simulation model required to estimate the aggregate error are the previous aggregate error ($\varepsilon_{A(t-1)}$) and the inflow rate (Q_{in}) values when using the RVM with fit statistics $RMS = 0.0245 \text{ m}^3/\text{m}/\text{hr}$ and $\eta = 0.820$ for the RVM. For the MLP the variables required were ($\varepsilon_{A(t-1)}$) and the on-demand hourly variation (OD_v) with $RMS = 0.0248 \text{ m}^3/\text{m}/\text{hr}$ and $\eta = 0.815$. Also the statistics calculated for the residuals indicates that these comply better with NIID characteristics.

It was found also that the RVM is affected by the amount of data available for training. While the RVM and MLP can perform in a similar fashion with the same amount of data. The RVM cannot capture completely the aggregate error pattern, being a small autocorrelation in the residuals of the model. This can be corrected by providing more information (one or more irrigation seasons) for the training data.

In general the two learning machine algorithms (RVM and MLP) performed in similar way mapping the relationship among the aggregate error and the variables from the hydraulic model using 2008 – 09 information. Nevertheless the MLP is more susceptible to be influenced by the characteristics of the data used to train the algorithm. This is an indication of the limited suitability of the MLP algorithm for the error

correction model and its use is not recommended. On the other hand, the RVM has better generalization properties, providing better results as demonstrated by the bootstrapping analysis. Therefore, the RVM is the data-driven algorithm recommended for modeling the aggregate error.

In hydraulic terms, the estimation of the aggregate error provides a new mean to estimate the aggregate error term that in other way would still remain not accounted in the hydraulic simulation model e.g. seepage and lateral flow in hourly basis and the error or noise in the SCADA system data for the present study. This is because the aggregate error comprises the error sources in the hydraulic simulation model.

It is very difficult to determine the individual contribution of the error sources and each error source impact on the model. The methodology followed in this study considers dealing with the lumped error produced by all the error sources and correct it, thus improving the simulation results.

Besides of the results obtained in this study, the proposed approach is not limited or restricted to minimize aggregate errors in hydraulic simulation models. Similar applications of coupled physical-based and data-driven models could be developed using the methodology explained in this study.

Future work on this area is related with the implementation of the developed methodology in the SCADA system in the Lower Sevier River Basin, Central Utah. Also, given the versatility of the developed approach and the learning machine algorithm, a multivariate error modeling approach for several outputs from a hydraulic simulation model or similar will be analyzed and tested.

References

- Anguita, D., Boni, A., Ridella, S. 2000. Evaluating the generalization ability of support vector machines through the bootstrap. *Neural Process* 11(1), 51-58.
- Asefa, T., Kemblowski, M. W., McKee, M., Khalil, A. 2006. Multi-time scale stream flow predictions: the support vector machines approach. *J. Hydrol.* 318 (1-4): 7-16.
- Bishop, C. M. 1995. *Neural Networks For Pattern recognition*. Oxford University Press, Oxford.
- Chesner, A. 1991. The effect of measurement error. *Biometrika*, 78(3), 451-462.
- Chaudhry, M. H. 1993. *Open-channel flow*. Prentice-Hall Inc., New York.
- Fread, D. L. 1974. Numerical properties of implicit four-point finite difference equations of unsteady flow. NOAA Technical Memorandum NWS, HYDRO-18, National Weather Serv., U.S. Department of Commerce. NOAA, Silver Springs, MD.
- Ghosh, S., Mujumdar, P. P. 2008, Statistical downscaling of GCM simulations to streamflow using relevance vector machine. *Adv. Water Res.*, 31(1), 132-146.
- Guyon, I., Elisseeff A. 2003. An introduction to variable and feature selection. *J. Mach. Learn. Res.*, 3(3), 1157 - 1182.
- Haie, N. 1984. Hydrodynamic simulation of continuous and surged surface flow. Ph.D Thesis, Utah State University, Logan.
- Khalil, A. F, McKee, M., Kemblowski, M. W., Asefa, T. 2005, Sparse bayesian learning machine for real-time management of reservoir releases. *Water Res. Res.*, 41, W11401.
- Khalil, A. F., McKee, M., Kemblowski, M. W., Asefa, T., Bastidas, L. 2006. Multi-objective analysis of chaotic dynamic Systems with sparse learning machines. *Adv. Water Res.*, 29 pp.72-88.
- Maier, R. M. Dandy, C. D. 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Env. Modeling and Software*, 15(1), 101 – 124.
- MacKay, D. 1992. A practical bayesian framework for backpropagation networks. *Neural Computation*, 4(3), 448–472.

- McCuen, R. H., Knight, Z., Cutter, A. G. 2006. Evaluation of the Nash–Sutcliffe efficiency index. *J. Hydrol. Eng.*, 11(6), 597 – 602.
- Nabney, I. 2002. *Netlab: algorithms for pattern recognition*. Springer-Verlag, London.
- Pierce, S. G., Worden, K., Bezazi, A. 2008. Uncertainty analysis of a neural network used for fatigue lifetime prediction. *Mechanical Systems and Signal Processing*, 22(6), 1395 – 1411. Special Issue: Mechatronics.
- Pebesma, E.J., Switzer, P., Loague, K. 2005. Error analysis for the evaluation of model performance: rainfall-runoff event time series data. *Hydrol. Proc.*, 19, 1529-1548.
- Rosenberry, D. O. 1990. Effect of sensor error on interpretation of long-term water-level data, *Ground Water* 28(6), 927 – 936.
- SRWUA - Sevier River Water Users Association website, www.sevierriver.org, accessed November, 2009.
- Strelkoff, T., Katopodes, N.D. 1977. Border-irrigation hydraulics with zero inertia. *J. Irrig. Drain. Div.*, 103(3), 325-342.
- Skogerboe, G. V., Merkley, G. P. 1996. *Irrigation Maintenance and Operations Learning Process*. Water Res. Publications LLC, Highlands Ranch.
- Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S. W., Srikanthan, S. 2009. Critical evaluation of parameters consistency and predictive uncertainty in hydrological modeling: a case study using bayesian total error analysis. *Water Res. Res.*, 45, W00B14.
- Ticlavilca, A., McKee, M. 2010. Multivariate bayesian regression approach to forecast releases from a system of multiple reservoirs. *Water Res. Manag.*, DOI: 10.1007/s11269-010-9712-y.
- Tipping, M. E. 2001. Sparse bayesian learning and the relevance vector machine. *J. Machine Learning Res.*, 1, 211–244.
- Tipping, M. E., Faul, A. C. 2003. Fast marginal likelihood maximization for sparse bayesian models. *Proc., Ninth Int. Workshop on Artificial Intelligence and Statistics*, Key West, FL, Jan 3-6.
- Walker, W. R., Skogerboe, G. V. 1987. *Surface Irrigation Theory and Practice*. Prentice-Hall, New Jersey.
- Walker, W. R., Stringam, B. L. 1999. Low cost adaptable canal automation for small canals. *ICIC Journal*, 48(3):39-46.

- Walker, W. R., Stringam, B. L. 2000. Canal automation for water conservation and improved flexibility. Proc., 4th Decennial Nat. Irrigation Symposium.
- Zechman, E. M., Ranjithan, S. 2007. Evolutionary computation-based approach for model error correction and calibration. Adv. Water Res., 30(5), 1360-1370.

CHAPTER 4
MULTIPLE-DAY IRRIGATION WATER DEMAND FORECAST USING
MULTIVARIATE RELEVANCE VECTOR MACHINES³

ABSTRACT

Characterization of future water demands in an agricultural command area (ACA) is affected by factors such as the crop type and stage, soil characteristics, weather pattern, water availability and distribution, farmers' and water managers' decisions among others. The estimation of future water deliveries is valuable information for water managers, canal operators, etc. This information is critical for irrigated areas where the water source is located at far distances. In many irrigation systems considerable investment has been done implementing SCADA systems to monitor the current conditions of canal systems. Nevertheless a missing component in irrigation water management is the information about future water deliveries for the next days to schedule the respective amounts from the water storage location. In engineering and science-related areas, data-driven tools or learning machines have proved to be very useful mapping relationships among inputs-outputs under incomplete or limited data scenarios. Therefore, these algorithms could be of use to develop models for water discharge estimations required for an ACA based on limited available data. This study presents a machine learning-based methodology that utilizes local available information (geospatial imagery, climatic data, soil moisture and historical water releases) of an ACA to anticipate required immediate daily future water deliveries. The data-driven tool chosen is the Multivariate Relevance Vector Machine.

³ Coauthored by Alfonso F. Torres, Andres M. Ticlavilca, Wynn R. Walker and Mac McKee

Aerial and satellite imagery can provide information of spatial variability in the system, actual evapotranspiration can relate with crop water needs, on-ground soil moisture sensors and collected water releases can supply historical information on water soil availability and water demand of the ACA for testing and validation purposes. The practical application of this methodology is demonstrated in an ACA located in Central Utah.

Introduction

As water becomes scarcer, competition intensifies and its value rises, especially in semiarid regions where irrigation is the largest water user in the basin (Svendsen, 2005). The core of irrigation water management of an agricultural command area (ACA) is based on implementation of structures and models that provide information about the state of the irrigation system and control over its storage and distribution processes (Pulido-Calvo and Gutierrez-Estrada, 2009). In modern irrigation systems these structures are accompanied by their automation being a widely used form the supervisory control and data acquisition or SCADA systems. Besides the SCADA implementation, hydraulic simulation models allow to assess in real time water flow conditions under different scenarios. Nevertheless, to manage and control successfully the irrigation system, future aggregate water requirement for the ACA is necessary to be accounted.

Estimation of required future water releases into the ACA is of critical importance for managing and planning activities by water managers and decision makers, especially in water-scarce areas (Bontemps and Couture, 2002). The value of future aggregated water demand information is related with: (a) water delivery efficiency management, especially in systems that involve large conveyance times (one to more days) from water

storage locations to agricultural areas, (b) water budgeting and distribution in the ACA canal internal network, and (c) expected water flow conditions from SCADA and hydraulic simulation models.

The characterization of present and future crop water demands in an ACA is affected by agronomical, hydrological and weather factors such as type and growth stage of crops, soil characteristics, weather pattern, water availability, and quality. Other factors that have a direct effect on water demand are farmers' and water managers' knowledge and behavior pattern during the irrigation season, as well as legal and institutional factors such as farmers' water rights. Most of these factors vary spatially and/or temporally, affecting the water demand estimation in its different time scales (daily, weekly, monthly, annually, etc.), especially in large irrigation systems.

Methodologies to estimate water demand for irrigation have been developed considering different points of view: spatial information systems (Herrero and Casterad, 1999; Ojeda-Bustamante et al., 2007), economical assessment of farmer behavior for water use (Bontemps and Couture, 2002), detailed characterization of the water system including spatial layout of crops (Lecina and Playan, 2006), and many others. From these methodologies two major approaches can be identified: conceptual or physical-based versus statistical-based or data-driven modeling (Pulido-Calvo and Gutierrez-Estrada, 2009).

Physical-based models can be used to estimate water requirements assuming external factors have a similar influence or low impact on the system such as weather variation, farmers' and water managers' behavior and knowledge patterns, water rights issues, etc. On the other hand, statistical models can provide a direct mapping among the

mentioned factors and the future water requirements with no detailed considerations about the internal structure of the physical processes that relate them (Pulido-Calvo and Gutierrez-Estrada, 2009). Another major difference among the two modeling approaches is that physical-based models requires a complete (or near complete) set of variables to estimate future water releases, whereas data-driven models only need to use available information to provide the same (or near similar) results.

The purpose of this paper is to assess the potential of a data-driven model, the Multivariate Relevance Vector Machine (MVRVM), for forecasting short-term irrigation water demand (up to two days in advance) using data from a SCADA system, weather information, soil moisture sensors and remote sensing data. The procedure to develop the model is outlined here, and results and considerations made are discussed. To demonstrate its practical application, the proposed methodology is applied to an ACA located in the Lower Sevier River Basin (Canal B). Assessment of stability and robustness of the methodology was performed, and a comparison of benchmarking performance is offered against the Multilayer Perceptron, a type of Artificial Neural Network.

Theoretical development

Irrigation water demand

Estimation of water needs in agriculture has always been a concern for farmers, water managers, and decision makers. Excessive water application to croplands is related to “water losses” such as deep percolation, runoff, soil surface water evaporation, and others. These losses, added to the consequent air volume reduction in the soil profile, root nutrient depletion, and water logging, negatively affect the crop dry mass and yield

production (Perry et al., 2009). On the opposite side, reduced water application causes reduced crop yields and in some cases, total loss (Sarwar and Perry, 2002).

These non-optimal water application consequences led to investigations to determine procedures that allow more precise estimation of water conveyance requirements from water sources to agricultural lands. Parameters such as evapotranspiration, crop coefficients, and water application ratios were developed, along with numerical and empirical models for crop water requirements and aggregate water demand estimation. This on-farm information combined with that obtained by the SCADA system is of great value when used for water control and distribution in an irrigated area.

There are major challenges in estimation of actual aggregate water needs within an ACA. First is the inclusion of spatial and temporal variability of critical parameters. Soil characteristics, crop type, coverage area and growing stage affect the accuracy of the aggregate water needs estimation. Also, the response of farmers to crop growing stage and irrigation timing is of high importance. Finally, water conveyance and distribution in the ACA which is related with water supply sources also affect the response of the water system to the farmers' water requirements.

Considering the spatial component, in many places of the western US, records are not generally available of the type of crops grown or the acreage covered by irrigation systems. This situation is somewhat different in other places such as some countries in Latin America and Asia where records of intended crops and their respective areas are acquired by the Water User Associations (WUAs) before the irrigation season starts. Nevertheless, the availability of aerial and satellite imagery in different spectra and

resolution formats now provide new means to estimate these variables. This spatial information can be useful to determine parameters such as crop, type, actual irrigated land areas, top soil moisture, and others (Herrero and Casterad, 1999). Spatial information has been proposed in previous studies for on-demand irrigation system management (D'Urso et al., 1995; Herrero and Casterad, 1999)

Soil moisture monitoring is another potentially important source of information of the quantity of water supplied to the cropland (Houser et al., 1998; Bellingham, 2009) in support of irrigation systems operation. The importance of soil moisture data is related to developing irrigation schedules (Jensen et al., 1970). Still, soil moisture is not information collected in most irrigation systems, but its implementation has become a trend among large WUAs which usually are well funded. One of the most interesting aspects of soil moisture data is its synthesis of several characteristics of the monitored area, such as crop stage, amount of water supplied, soil physical and water holding characteristics, agricultural labor availability and costs, farmers' irrigation pattern, water management, weather pattern, groundwater effect, and others.

The absence of information sources for most irrigation systems has led to several efforts to simulate them with water management models using physical- or statistical-based approaches. Some examples of these models were mentioned in the Introduction section. Focusing on statistical-based models, previous attempts for water demand forecasting using data-driven tools in the literature are limited. Some worth mentioning have used a Linear Regression – Artificial Neural Network approach (Pulido-Calvo et al., 2007), a Genetic Algorithm - Artificial Neural Networks (Kim et al., 2001; Pulido-Calvo and Gutierrez-Estrada, 2009) and Relevance Vector Machines (Flake et al., 2010). So

far, water demand forecasting has been limited to one day ahead, or a separate single model per forecasted day. This can be a constraint for their practical use by water managers and decision makers when water conveyance requires several days. Lack of short-term water demand information requires canal and reservoir managers to “guess” future releases from water storage and diversions into irrigation canals. This could negatively affect the adequate supply and distribution of irrigation water.

Multi-layer perceptron

Among the large number of implementation of Artificial Neural Network (ANN) models, the Multi-Layer Perceptron (MLP) is one of the most widely used because of its ability to approximate any smooth function (Nabney, 2002). An interesting characteristic of this type of ANN is the inclusion of a Bayesian Inference Method to calibrate the MLP parameters. The Bayesian Inference also allows estimation of the uncertainty related to the predicted outputs. The MLP architecture can be described as:

$$y^{(n)} = W^{II} \cdot \tanh\left(W^I x^{(n)} + b^I\right) + b^{II} \quad (4.1)$$

where:

$y^{(n)}$: MLP output vector, $y^{(n)} = [y_1, \dots, y_m, \dots, y_M]$,

$x^{(n)}$: input vector $x^{(n)} = [x_1, \dots, x_d, \dots, x_D]$,

W^I, W^{II} : optimized weights for the first and second layer, respectively,

$$W^I = [w_{1,1}^I, K, w_{D,NN}^I], \quad W^{II} = [w_{1,1}^{II}, K, w_{NN,M}^{II}]$$

M: number of components of the output vector,

D: number of components in the input vector,

NN: number of hidden neurons,

b^I, b^{II} : bias vectors for the first and second layer, respectively.

Using a dataset $\Lambda = \{x^{(n)}, t^{(n)}\}_{n=1}^N$, where N is the number of training cases, the calibration of the MLP is performed by optimizing the network parameters $W = \{W^I, W^{II}, b^I, b^{II}\}$ in order to minimize the Overall Error Function E (Bishop, 1995):

$$E = \frac{\beta}{2} \sum_{n=1}^N \left(t^{(n)} - y^{(n)} \right)^2 + \frac{\alpha}{2} \sum_{i=1}^W W_i^2 \quad (4.2)$$

$$E = \beta \cdot E_{\Lambda} + \alpha \cdot E_W$$

where:

E_{Λ} : data error function,

E_W : penalization term,

W: number of weights and biases in the neural network, and

α and β : Bayesian hyperparameters.

In Bayesian terms, the goal is to estimate the probability of the weights and bias of the MLP model, given the dataset Λ :

$$p(W | t^{(n)}) = \frac{p(t^{(n)} | W) p(W)}{p(t^{(n)})} \quad (4.3)$$

where, as explained by MacKay (1992):

$p(W | t^{(n)})$: the posterior probability of the weights,

$p(t^{(n)} | W)$: the dataset likelihood function,

$p(W)$: the prior probability of the weights, and

$p(\mathbf{t}^{(n)})$: the evidence for the dataset.

Assuming a Gaussian distribution for the error term $\xi^{(n)} = \mathbf{t}^{(n)} - \mathbf{y}^{(n)}$ and the weights \mathbf{W} , the likelihood and the prior probabilities can be expressed:

$$p(\mathbf{t}^{(n)} | \mathbf{W}, \beta) = (2\pi\beta^{-1})^{-N/2} \exp(-\beta E_{\Lambda}) \quad (4.4)$$

$$p(\mathbf{t}^{(n)} | \mathbf{W}, \alpha) = (2\pi\alpha^{-1})^{-N/2} \exp(-\alpha E_{\mathbf{W}}) \quad (4.5)$$

E_{Λ} models the uncertainty (or error) of the target variables as Gaussian zero-mean noise and variance $\sigma^2 \equiv \beta^{-1}$. $E_{\mathbf{W}}$ defines the conditional probability of \mathbf{W} with variance $\sigma_{\mathbf{W}}^2 \equiv \alpha^{-1}$. Then Eq. 3 can be expressed as:

$$p(\mathbf{W} | \mathbf{t}^{(n)}, \alpha, \beta) = \frac{p(\mathbf{t}^{(n)} | \mathbf{W}, \beta) \cdot p(\mathbf{W} | \alpha)}{p(\mathbf{t}^{(n)} | \alpha, \beta)} \quad (4.6)$$

$$p(\mathbf{W} | \mathbf{t}^{(n)}, \alpha, \beta) = \frac{\exp\left(\mathbf{E}\left(\mathbf{W}^*\right) - \frac{1}{2} \Delta \mathbf{W}^T \mathbf{H} \Delta \mathbf{W}\right)}{\exp\left(\mathbf{E}\left(\mathbf{W}^*\right) \cdot (2\pi)^{W/2} |\mathbf{H}|^{-1/2}\right)} \quad (4.7)$$

In which,

$\mathbf{E}\left(\mathbf{W}^*\right)$: expected optimized values for the weights and bias,

\mathbf{H} = Hessian matrix $\mathbf{H} = \beta \nabla \nabla^T E_{\Lambda} + \alpha \mathbf{I}$, \mathbf{I} is the identity matrix.

$\Delta \mathbf{W} = \mathbf{W} - \mathbf{W}^*$.

Once the distribution of \mathbf{W} has been estimated by maximizing the likelihood for α and β , the prediction $\mathbf{y}^{(n)}$ and its standard deviation $\sigma_y^{(n)}$ can be estimated by integrating (marginalizing) over \mathbf{W} and the regularization parameters α and β (Bishop, 1995):

$$p(y^{(n)} | x^{(n)}, t^{(n)}) = \int p(t^{(n)} | x^{(n)}, \mathbf{W}^*) p(\mathbf{W}^* | t^{(n)}) \cdot d\mathbf{W} \quad (4.8)$$

This can be approximated by:

$$p(y^{(n)} | x^{(n)}, t^{(n)}) \propto \left(2\pi \sigma_y^{(n)^2} \right)^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \sigma_y^{(n)^2} \left(y^{(n)*} - t^{(n)} \right)^2 \right) \quad (4.9)$$

where $y^{(n)}$ is the output and $\sigma_y^{(n)^2}$ is the output variance from the MLP. The output variance can be expressed as:

$$\sigma_y^{(n)^2} = \beta^{-1} + \mathbf{g}^T \mathbf{H}^{-1} \mathbf{g} \quad (4.10)$$

where \mathbf{g} denotes the gradient of $y^{(n)}$ with respect to the weights; $\mathbf{g} \equiv \nabla \mathbf{W} y^{(n)} | \mathbf{W}^*$. The output variance has then two sources; the first arises from the intrinsic noise in the target data; and the second from the posterior distribution of the ANN weights (Pierce et al., 2008). The output standard deviation vector $\sigma_y^{(n)}$ can be interpreted as the error bar for confidence interval estimation (Bishop, 1995).

Multivariate relevance vector machine

The Multivariate Relevance Vector Machine (MVRVM), developed by Thayananthan et al. (2008), is a general Bayesian framework for obtaining multivariate sparse solutions to regression tasks. The MVRVM is based on the Relevance Vector Machines framework developed by Tipping (2001) and Tipping and Faul (2003) which was extended to handle multivariate outputs. This learning machine is particularly useful in hydrology and water resources because of the generalization properties and the probabilistic estimation, useful to estimate prediction uncertainty (Tripathi and Govindajaru, 2007). The mathematical formulation of the MVRVM is:

$$y^{(n)} = \mathbf{W}^* \cdot \Phi[x^{(n)}] \quad (4.11)$$

where:

$x^{(n)}$ and $t^{(n)}$: input and target vectors that belong to the dataset $\{\Lambda\}_{n=1}^N$, as defined for MLP,

$y^{(n)}$: MVRVM output vector $y^{(n)} = [y_1, \dots, y_M]$; $m \in 1 \leq m \leq M$,

M: number of components in the target and MVRVM output vectors,

N: number of training cases,

\mathbf{W}^* : optimized weight matrix, $\mathbf{W}^* = [w_{1,1}, \dots, w_{m,rv}, \dots, w_{M,RV}]$,

RV: number of optimized cases or relevance vectors selected by the MVRVM from the N training cases, $RV \ll N$, $rv \in 1 \leq rv \leq RV$,

$\Phi[x^{(n)}]$: optimized design matrix or basis function (represented also by Φ^*) that can be related with a kernel function $\Phi^* = K[x^{(n)}, \{x^{(*)}\}_{rv=1}^{RV}]$.

The kernel function is a weighting function for the input vector ($x^{(n)}$) used in non-parametric estimation techniques, e.g. kernel regression models. It provides an adjustment to the $x^{(n)}$ vector based on RV optimal cases or “relevant vectors,” $x^{(*)}$, which are selected automatically among the N training input vectors. For calibration of the MVRVM a variation of the Overall Error function (Eq. 2) is used and, by using the Bayesian Inference Method the distribution of the weights of the model (Eq. 3) is estimated, similar to the MLP process. Also, the MVRVM error term or residual $\xi^{(n)} = t^{(n)} - y^{(n)}$ is assumed to be probabilistic independent zero-mean Gaussian, with variance σ_ξ^2 .

The detail of the MVRVM algorithm is as follows:

Assuming a Gaussian prior probability distribution for the weights (Tipping, 2001), and representing $A = \text{diag}(\alpha_1^{-2}, \dots, \alpha_N^{-2})$, and $B = \text{diag}(\beta_1, \dots, \beta_M)$, where each

element α_n is a hyperparameter that determines the relevance of the associated basis function for every case in the training data. $\beta_m = \sigma_\xi^2$ represents the noise or error variance in the m^{th} component of the target data (Thayananthan et al., 2008). The prior distribution over the weights is represented by:

$$p(\mathbf{W} | \mathbf{A}) = \prod_{m=1}^M \prod_{n=1}^N \mathcal{N}(w_{m,n} | 0, \alpha_n^{-2}) \quad (4.12)$$

where $w_{m,n}$ is the element at (m, n) of the weighting matrix, $\mathbf{W} = [w_{1,1}, \dots, w_{m,n}, \dots, w_{M,N}]$. The likelihood distribution of \mathbf{W} can be expressed as:

$$p\left(\left\{t^{(n)}\right\}_{n=1}^N \mid \mathbf{W}, \mathbf{B}\right) = \prod_{n=1}^N \mathcal{N}\left(t^{(n)} \mid \mathbf{W} \cdot \Phi, \mathbf{B}\right) \quad (4.13)$$

$\Phi = \mathbf{K} \left[\left\{ \mathbf{x}^{(n)} \right\}_{n=1}^N, \left\{ \mathbf{x}^{(n)} \right\}_{n=1}^N \right]$. The likelihood of the target $t^{(n)}$ can be written as:

$$p\left(\left\{t^{(n)}\right\}_{n=1}^N \mid \mathbf{W}, \mathbf{B}\right) = \prod_{m=1}^M \mathcal{N}\left(\tau_m \mid w_m \cdot \Phi, \beta_m\right) \quad (4.14)$$

where τ_m is a vector with the m^{th} component of all the target data and w_m the weight vector of the m^{th} component of the output vector $t^{(n)}$. The prior distribution over the weights can be rewritten as:

$$p(\mathbf{W} | \mathbf{A}) = \prod_{m=1}^M \mathcal{N}(w_m | 0, \mathbf{A}) \quad (4.15)$$

The posterior probability of \mathbf{W} can be written as the product of separate Gaussians of the weights vectors of each output dimension:

$$p\left(\mathbf{W} \mid \left\{t^{(n)}\right\}_{n=1}^N, \mathbf{B}, \mathbf{A}\right) \propto \left(\left\{t^{(n)}\right\}_{n=1}^N \mid \mathbf{W}, \mathbf{B}\right) \cdot p(\mathbf{W} | \mathbf{A}) \quad (4.16)$$

$$p\left(\mathbf{W} \mid \left\{t^{(n)}\right\}_{n=1}^N, \mathbf{B}, \mathbf{A}\right) \propto \prod_{m=1}^M \mathcal{N}(w_m \mid \mu_m, \Sigma_m) \quad (4.17)$$

where $\mu_m = \beta_m^{-1} \Sigma_m \Phi^T \tau_m$ and $\Sigma_m = (\beta_m^{-1} \Phi^T \Phi + A)^{-1}$ are the mean and the variance of the weight matrix respectively. Marginalizing the data likelihood over the weights:

$$p\left(\left\{t^{(n)}\right\}_{n=1}^N \mid \mathbf{A}, \mathbf{B}\right) = \int p\left(\left\{t^{(n)}\right\}_{n=1}^N \mid \mathbf{W}, \mathbf{B}\right) \cdot p(\mathbf{W} \mid \mathbf{A}) \cdot d\mathbf{W} \quad (4.18)$$

$$p\left(\left\{t^{(n)}\right\}_{n=1}^N \mid \mathbf{A}, \mathbf{B}\right) = \prod_{m=1}^M \left| \mathbf{H}_m \right|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \tau_m^T \mathbf{H}_m^{-1} \tau_m\right) \quad (4.19)$$

being \mathbf{H}_m the Hessian matrix for the m^{th} component of the target vector, $\mathbf{H}_m = \beta_m \mathbf{I} + \Phi^T \mathbf{A}^{-1} \Phi$. An optimized set of hyperparameters $\{\alpha_{rv}^*\}_{rv=1}^{RV}$ and noise parameters $\{\beta_m^*\}_{m=1}^M$ is obtained by maximizing the marginal likelihood as described by Tipping and Faul (2003). The final hyperparameter values are:

$$\mathbf{A}^* = \text{diag}\left(\alpha_1^{*-2}, \mathbf{K}, \alpha_{RV}^{*-2}\right) \quad (4.20)$$

$$\Sigma_m^* = \left(\beta_m^* \Phi^T \Phi + \mathbf{A}^* \right)^{-1} \quad (4.21)$$

The optimized mean vector and the weight matrix are:

$$\mu_m^* = \beta_m^* \Sigma_m^* \Phi^T \tau_m \quad (4.22)$$

$$\mathbf{W}^* = \left(\mu_1^*, \dots, \mu_M^* \right)^T \quad (4.23)$$

The MVRVM output and output error bar vectors are:

$$\mathbf{y}^{(n)} = \mathbf{W}^* \cdot \Phi \quad (4.24)$$

$$\sigma_y^{(n)} = \text{sqrt}\left(\mathbf{B}^{*-1} + \Phi^T \cdot \Sigma \cdot \Phi \right) \quad (4.25)$$

Material and methods

Site description

The water resources of the Sevier River Basin in Central Utah (Fig. 4.1) are among the most heavily utilized in the Western US. Substantial efforts to increase efficiency via canal lining and on-farm improvement such as conversion to sprinkler irrigation and laser land leveling were made during 1960 - 1990 period. From 1990 to the present, all reservoirs and stream offtakes have been equipped with SCADA technology and web-based data summaries (SRWUA, 2009). Canal automation was introduced in 1994 and shown not only to result in substantial reduction in losses but also to considerably shorten the response time between farmer demands and system deliveries (Walker and Stringam, 1999, 2000). Most recently, attention has been focused on improving the coordination between farmer demands, canal deliveries, and reservoir diversions.

In order to develop, test and implement the model proposed of this study, a subsystem at the lower end of the Sevier River Basin was selected. The agricultural command area (ACA) irrigated by Canal B is connected to the DMAD Reservoir by a 9 km. canal (Canal A). The DMAD gates as well as the Canal A gates are automated and operated as a SCADA system by local water masters.

The DMAD Reservoir is supplied water on a demand basis from Sevier River Bridge Reservoir located upstream. The lag time from Sevier Bridge Reservoir to DMAD Reservoir is 3 days. Thus, an emerging water demand in any ACA can be supplied within about 12 hours if water is available in DMAD Reservoir, or 4 days if water must be conveyed from Sevier Bridge Reservoir. Thus, the goal of the entire system is to provide

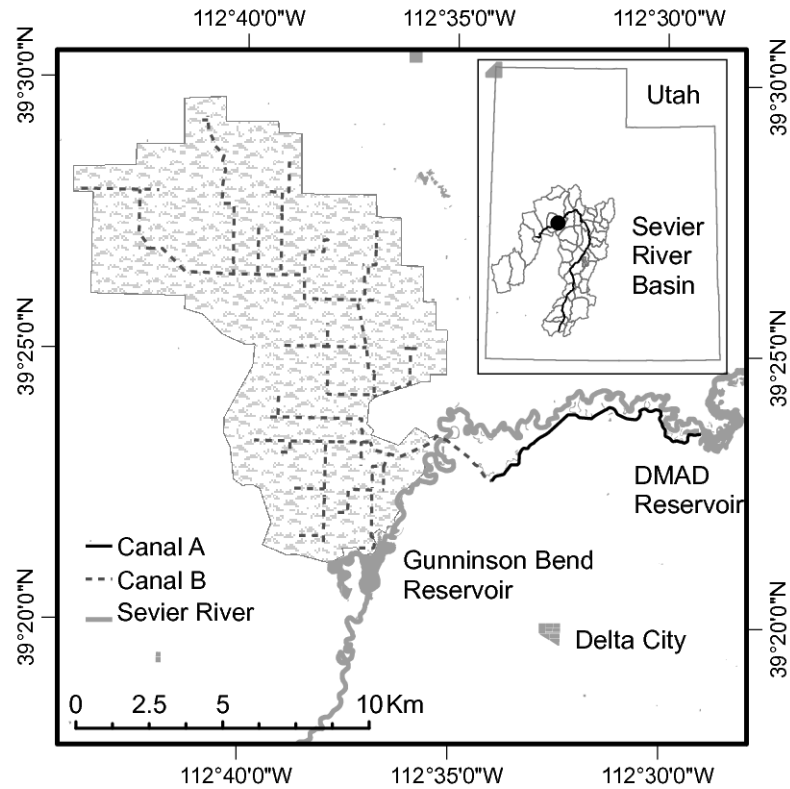


Fig. 4.1. Area of study, ACA Canal B in Delta, Utah.

water to an individual farm within 12 hours of an order by the irrigator. This goal relies heavily on the SCADA system and the regulation capacity of DMAD Reservoir.

The water management goals over the next few years are to increase the DMAD regulation capacity and improve the reliability of the 12-hour delivery interval period. It is expected that by better control of the DMAD Reservoir level will reduce seepage, evaporation and administration losses by about 25 to 50%. The most important capability needed to achieve this goal is to develop a reliable and accurate forecast of irrigation demand, which is related with available estimation of future aggregate water demand under the current irrigation system conditions.

Data description

For the area of study, information from several sources for years 2008 and 2009 was collected. Spatial imagery of the ACA Canal B was obtained from the NASA LandSat TM5 Program (2009) for the month of May. This month was chosen because by this time crops in the area under study are mature enough to allow identification from bare soil and fallow vegetation areas using remote sensing techniques. Weather data was collected from the local NOAA station located in Delta, Utah (Station Number 72479). Maximum and minimum daily air temperatures and precipitation records were available from water station. Local crop coefficients (K_c) values were obtained from the study done by Wright (1982). Past water discharges conveyed to Canal B records were obtained from the SCADA database accessible from the Sevier River WUA website (www.sevierriver.org). Soil moisture records were collected from monitoring stations located across the ACA Canal B being this information also accessible from the WUA website.

Aggregate water demand forecasting model

The proposed model for forecasting water requirements for the areas irrigated by Canal B is based on the relationship of these factors: a) conveyed water into Canal B, related with the current water management of the irrigation system by water masters, reservoir and canal managers, farmers' water orders and current irrigation system conveyance capacity, and b) agricultural information of the ACA, such as water requirements by type and stage of the crops. The data-driven algorithms, MVRVM or MLP are used to provide the required mapping among these factors. The proposed model was based on the water balance equation (4.26) on any field and developed as follows:

The water balance equation expresses the sum of inflows and outflows in a defined soil volume over a specific time interval. The net sum of inflows and outflows produces a variation in the soil moisture content.

$$I_t + Pp_t - ET_c - Ro_t - Dp_t = \Delta S_t \quad (4.26)$$

where:

I_t : Net Irrigation (mm/day)

Pp_t : Precipitation (mm/day)

ET_t : Crop Evapotranspiration (mm/day)

Ro_t : Runoff (mm/day)

Dp_t : Deep percolation (mm/day)

ΔS_t : Soil moisture variation (mm/day)

t : time (days)

Nevertheless, there are some components of Eq. 4.26 that are not measured or assumed small enough, such as Ro_t and Dp_t . On the other hand, Pp_t is extremely scarce and very sparse for the area under study. These conditions change Eq. 4.26 to:

$$I_t \propto \Delta S_t + ET_t \quad (4.27)$$

where the symbol \propto indicates proportionality. Considering the water conveyance and application efficiency (ϵ), the discharge required for a given field at the ACA inlet would be $Q_t = I_t/\epsilon$ (m^3/s). The ϵ is not necessarily uniform throughout the irrigation season, being mostly influenced by water management procedures. Therefore, modifying Eq. 4.27:

$$Q_t \propto \Delta S_t + ET_t \quad (4.28)$$

$$Q_t = f(ET_t, \Delta S_t) \quad (4.29)$$

To extend Eq. 4.29 for the entire ACA, it is necessary to take into consideration the implemented crops and water soil characteristics in the area. Thus, the type and acreage of crops in the ACA needs to be quantified for its inclusion in the model. About the water soil characteristics, rarely spatial distribution of soil moisture is available for any ACA. For the area under study only selected locations which correspond to the Soil Moisture Monitoring Sites (S_{site}) can be included in the model:

$$Q_{S_t} = \sum Q_t \propto f(ET_{crop,t}, S_{site,t}) \quad (4.30)$$

where Q_{S_t} is the total inflow for the ACA at time t (m^3/s), the subscript crop is related with each crop in the ACA and the subscript site is related with the soil moisture monitoring location. To include the implicit pattern of the human behavior in the operation of the irrigation system, historical water releases for Canal B are considered for the model (Flake et al., 2010). With these last modifications, it is possible to extend Eq. 4.30 for forecasting purposes:

$$Q_{S_{t+l:t+k}} = f(Q_{S_{t-t-r}}, ET_{crop,t-t-m}, S_{site,t-t-p}) \quad (4.31)$$

where the indexes r , m , n , and p refer to time steps in the past. The index k represents the number of forecasted days. The letter f represents the mapping algorithm, in this case MVRVM or MLP. In summary, the data-driven algorithms are required to provide estimations of future aggregate water releases with only the available data in the ACA: historical discharges, ET crop estimation and soil moisture records.

Learning machines

The data-driven algorithm selected for the aggregate water demand forecasting model is the MVRVM. For performance comparison the MLP was also tested. The model considers the use of historical water releases into ACA Canal B, agricultural and soil water information. As note in the Theoretical Development section, the data-driven algorithms require tuning of their own parameters, for the MVRVM, the type of kernel function and kernel width. For the MLP is the number of hidden layers, the number of neurons in these layers, and the training function.

To determine the accuracy of the forecasted results of the data-driven algorithms two goodness-of-fit criteria were used: (1) the Root Mean Square Error (RMSE); and (2) the Nash-Sutcliffe Efficiency Index (η). The Nash-Sutcliffe Efficiency Index is recommended for non-linear modeling problems (McCuen et al., 2006).

$$\text{RMSE} = \sqrt{\sum_{n=1}^{N_*} (y_*^{(n)} - t_*^{(n)})^2 / N_*} \quad (4.32)$$

$$\eta = 1 - \frac{\sum_{n=1}^{N_*} (y_*^{(n)} - t_*^{(n)})^2}{\sum_{n=1}^{N_*} (t_*^{(n)} - \overline{t_*^{(n)}})^2} \quad (4.33)$$

where:

$t_*^{(n)}$: historical discharge for the testing data,

$y_*^{(n)}$: forecasted discharge values for the testing data,

N_* : number of samples or cases in the testing data, and

$\overline{t_*^{(n)}}$: average values of the historical flow rates.

The RMSE values allow ranking of the performance of each learning machine. Large RMSE values indicate that the error between the historical and predicted discharge values is large. η measures in a non-dimensional range (from $-\alpha$ to 1) the closure of the historical vs. the predicted discharge values. A η value of 1 indicates perfect correspondence. A η value of 0 indicates that the forecasted flow rate is not better than the average of the historical water flow values.

Results

Available information

The initial information required to develop the water demand model is the identification of the types of crops and their area coverage in the ACA. LandSat 5 TM images from the GLOVIS USGS website (2009) were downloaded for the area under study (path/row: 38/33). Two images that correspond to the May month for 2008 and 2009 were obtained. Once processed, these images allowed identification of crops and other land cover in the Canal B area and quantification of the acreage covered by each. Three main crops were identified: alfalfa, corn, and small grains (barley, wheat). Fallow vegetation was also detected during the image processing. To verify the accuracy of the crop identification results a visual comparison was made during field trips to the ACA. The processed satellite image for 2009 is presented in Fig. 4.2 and the areas covered by each crop for each year is presented in Table 4.1.

During the field trips to the ACA, interviews with local farmers, water masters and managers was possible. Based on these conversations it was determine that water delivered in Canal B is used only for irrigation and not for other purposes, e.g. human

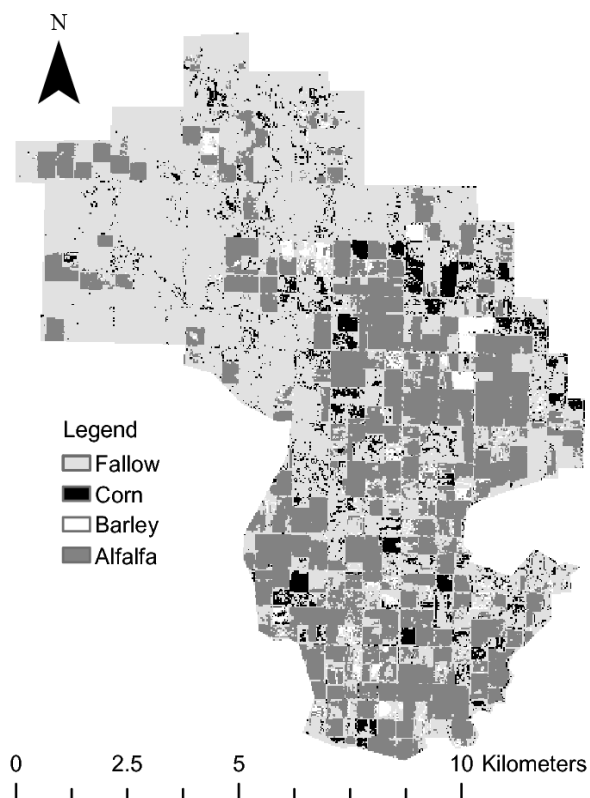


Fig. 4.2. Crop distribution for ACA Canal B, 2009.

Table 4.1. 2008 and 2009 crop areas for ACA Canal B.

Crop	2008		2009	
	(ha)	%	(ha)	%
Alfalfa	3183.9	30.0	3369.2	32.0
Corn	2378.5	23.0	723.6	7.0
Small Grains	305.6	3.0	323.3	3.0
Fallow	4653.2	44.0	6105.2	58.0
Total	10521.2	100.0	10521.2	100.0

consumption, industrial and animal production. These other types of uses only employ water from wells and other sources.

Based on the crop identification results, potential evapotranspiration rates were estimated for the 2008 and 2009 irrigation seasons using daily air temperatures and the 1985 Hargreaves ET_0 Equation. This equation is recommended for water planning purposes and requires minimal weather information to provide ET_0 estimations with good approximation (Allen et al., 1998). The weather station in Delta only provides daily maximum and minimum air temperatures, and precipitation records for the area under study, which is enough for the application of the 1985 Hargreaves Equation. Local crop coefficients for the Midwest area of the United States were obtained from Wright (1982) and used to estimate the actual crop evapotranspiration rates. Fig. 4.3 shows the behavior of, respectively, the local crop coefficients and the actual values of ET for the 2009 irrigation season.

The soil moisture data was obtained from monitoring sites located in the croplands irrigated by Canal B. The soil moisture data can be accessed from the Sevier River WUA website (www.sevierriver.org). These stations have provided records of soil moisture at depths of 1 and 2 ft in 44 selected farms since 2007 (88 sensors). For the 44 available soil moisture stations, a quality control analysis was performed, addressing the quality of soil moisture records, e.g. error or noise imbedded in the data, and data completeness. The results indicate data from some soil moisture sites in 2008 was affected by noise, given the harsh working environment to which the soil moisture probes and ancillary equipment are exposed. Therefore, after the quality control analysis, a final selection of 11 stations with 2008 and 2009 irrigation seasons was performed and is

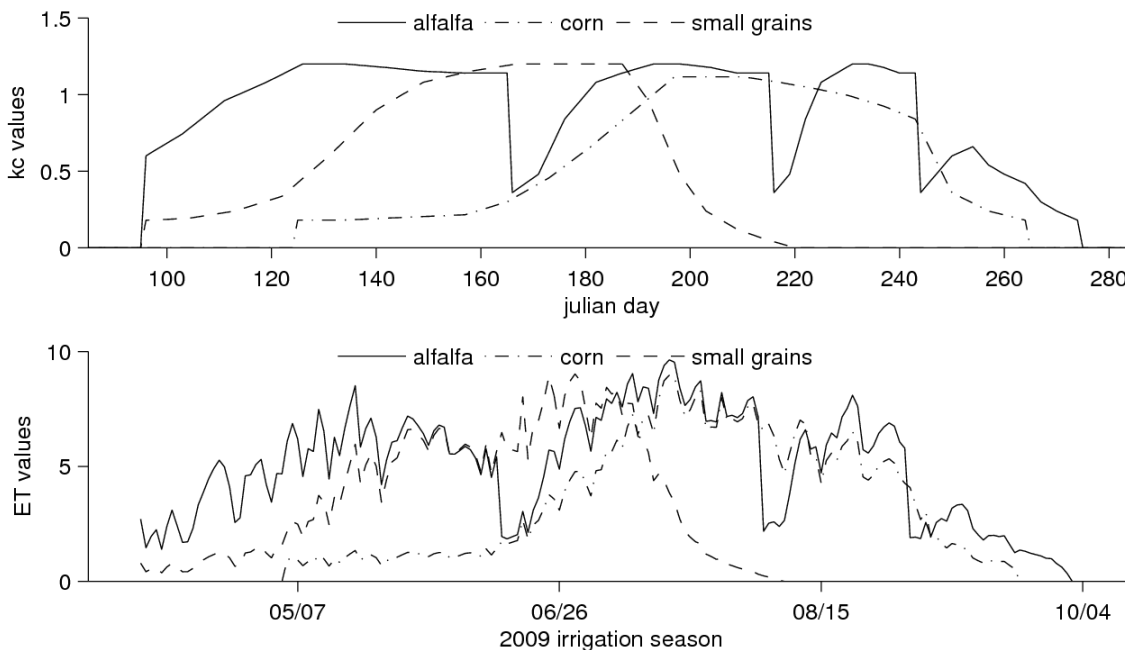


Fig. 4.3. Local Kc and 2009 actual evapotranspiration values for main crops in ACA Canal B

presented in Table 4.2. The soil moisture data from these stations will be tested for inclusion in the proposed model. Historical records of daily water diversions into Canal B were retrieved from the WUA database for 2008 and 2009 irrigation seasons. This information can be also accessed from the Sevier River WUA website.

Water demand forecasting model

After acquisition of available data for ACA Canal B, calibration of the mapping function for the water demand forecast model, MVRVM and the MLP was performed. These learning machines have an advantage over other learning models in terms of model calibration, which was the reason for their selection in this study. The Bayesian Inference Method used by these algorithms to estimate their parameters is a technique that avoids model overfitting issues. Thus, there is no need for other techniques, e.g. cross validation, to analyze the data-driven calibration. For the calibration of the learning machine

Table 4.2. Selected soil moisture monitoring stations.

Station	Crop 2008	Crop 2009
104-b	no crop	no crop
107-a	alfalfa	alfalfa
109-a	alfalfa	alfalfa
109-b	alfalfa	alfalfa
111-a	alfalfa	alfalfa
115-a	alfalfa	corn
115-b	alfalfa	corn
116-a	corn	barley
116-b	corn	barley
118-a	alfalfa	alfalfa
118-b	alfalfa	alfalfa

algorithms, two factors were taken into account: (a) parameter selection and tuning, and (b) the optimal information or variables required as inputs for the forecasting model.

To proceed with the calibration of the water demand forecasting model, it is necessary to distribute the available data into two groups: one for training and tuning the data-driven algorithms and a second group to verify the adequacy of the calibrated algorithm. As mentioned before, available data includes daily information for 2008 and 2009 irrigation seasons, covering the months of April to October (~256 days). For this study, data from the 2008 irrigation season was used to calibrate the forecasting model,

while 2009 irrigation season data was used for verification of the adequacy of the model (goodness-of-fit values).

In terms of dimensionality or best inputs for the forecasting model, an inclusion of all available data in the model is not recommended because this can result in reduction in the performance of the data-driven algorithm. Among the available data there is an optimal number of variables or sources of information that provides the best relationship with the future water demand. To identify the best input variables, a selection of the inputs for the data-driven models was conducted as advised by Guyon and Elisseeff (2003). In this selection (also called forward variable selection), every variable or variable combination is tested against the desired outputs using the selected learning machine and the goodness-of-fit values obtained are stored for ranking purposes. Variable or variable combinations that have the highest correspondence with the desired outputs (best-fit values) are kept as fixed inputs in the model while the other variables are being included in the learning machine in combination with the fixed ones. The process stops when the addition of any new variable in the data driven model does not improve the goodness-of-fit statistics and the visual analysis. This procedure ensures that optimal inputs are included in the data-driven algorithm while the synergy effect of these inputs is maximized.

The calibration of the parameters for the data-driven algorithms was performed during the execution of the variable selection methodology. During the test of each variable or variable combination, an internal tuning of the data-driven algorithm parameters was performed. The tuned parameter for the MVRVM was the kernel width (σ_m), while the type of kernel was fixed to Gaussian as recommended by Gill et al. (2006)

for hydrologic problems. For the MLP, the tuned parameter was the number of neurons in one single hidden layer using the secant gradient optimization function, as recommended by Nabney (2002).

For ACA Canal B, there are three main data sources to be included as inputs in the data-driven algorithms: past Canal B inflows, crop water requirements and soil moisture. As shown in Eq. 4.31, it is expected that past values of the available information are required for the forecasted water releases. These past values are identified in Eq. 4.31 by the indexes (m , p , and r). These indexes are assumed not to be the same for all the tested variables and their value shall be defined by the variable selection procedure. This is because the possible lags that can occur between a soil moisture site and the water releases for the Canal B area. The same criterion is applied to other variables like crop evapotranspiration. A general 5 days lag in the past was applied for every variable that could be included as input in the forecasting model. This makes a total of 130 variables tested during the variable selection procedure. A summary of the total number of input variables is presented in Table 4.3.

Results of the forward variable selection with their corresponding tuned data-driven parameter for the MLP and MVRVM are presented in Tables 4.4 and 4.5, respectively, with the best variable combination shown at the end of each table. The data-driven error bar (σ_y) is also included as a measurement of the approximation of the model to the actual water demand values. For both of the data-driven algorithms, the best variable combination includes data from the three main sources: historical Canal B inflow, alfalfa actual evapotranspiration, and soil moisture data from at least two monitoring sites. The number of steps back in time is different for each algorithm and

Table 4.3. Tested variables in the forward variable selection procedure.

Available Data	Variables	Notation Example	Details
Discharge	5	$Q_{S_t}, Q_{S_{t-1}}, \dots, Q_{S_{t-4}}$	Past Canal B inflows
Crop ET	15	ET Alfalfa _t , ..., ET Alfalfa _{t-4} , ET Corn _t , ..., ET Corn _{t-4} , ET Grain _t , ..., ET Corn _{t-4}	Past crop water requirements for crops in ACA.
Soil Moisture Sites	110	104-b-1ft _t , ..., 104-b-1ft _{t-4} , 104-b-2ft _t , ..., 104-b-2ft _{t-4} , 118-b-2ft _t , ..., 118-b-2ft _{t-4} .	11 stations, 2 sensors per station.

Table 4.4. Goodness-of-fit and error bar values for variable selection procedure using MLP (2009 data).

Tested Variables	Hidden	η	η	RMSE	RMSE	σ_y	σ_y
	Neurons	$Q_{S_{t+1}}$	$Q_{S_{t+2}}$	$Q_{S_{t+1}}$	$Q_{S_{t+2}}$	$Q_{S_{t+1}}$	$Q_{S_{t+2}}$
Q_{S_t}	8	0.92	0.78	0.40	0.68	0.94	0.95
$Q_{S_t}, Q_{S_{t-1}}$	4	0.93	0.78	0.36	0.69	0.80	0.80
$Q_{S_t}, Q_{S_{t-1}}, 109\text{-a-}2\text{ft}_{t-3}$	3	0.93	0.77	0.34	0.65	0.73	0.73
$Q_{S_t}, Q_{S_{t-1}}, 109\text{-a-}2\text{ft}_{t-3},$ ET Alfalfa _t	1	0.92	0.79	0.35	0.65	0.88	0.88
$Q_{S_t}, Q_{S_{t-1}}, 109\text{-a-}2\text{ft}_{t-3},$ ET Alfalfa _t , 107-a-1ft _t , 116-a-1ft _{t-2} , $Q_{S_{t-2}}$	1	0.92	0.81	0.44	0.62	0.87	0.87

Table 4.5. Goodness-of-fit and error bar values for variable selection procedure using MVRVM (2009 data)

Tested Variables	σ_m	η $Q_{S_{t+1}}$	η $Q_{S_{t+2}}$	RMSE $Q_{S_{t+1}}$	RMSE $Q_{S_{t+2}}$	σ_y $Q_{S_{t+1}}$	σ_y $Q_{S_{t+2}}$
Q_{S_t}	3.5	0.92	0.77	0.40	0.69	0.49	0.86
$Q_{S_t}, Q_{S_{t-1}}$	8.8	0.95	0.81	0.33	0.63	0.36	0.73
$Q_{S_t}, Q_{S_{t-1}}, 109\text{-a-}1ft_{t-2}$	8.5	0.94	0.79	0.34	0.66	0.34	0.64
$Q_{S_t}, Q_{S_{t-1}}, 109\text{-a-}1ft_{t-2}, ET\ Alfalfa_t$	7.5	0.95	0.82	0.32	0.60	0.35	0.70
$Q_{S_t}, Q_{S_{t-1}}, 109\text{-a-}1ft_{t-2}, ET\ Alfalfa_t,$ $109\text{-a-}2ft_{t-2}$	7.8	0.95	0.83	0.32	0.59	0.35	0.70
$Q_{S_t}, Q_{S_{t-1}}, 109\text{-a-}1ft_{t-2}, ET\ Alfalfa_t,$ $109\text{-a-}2ft_{t-2}, 115\text{-b-}2ft_t$	7.5	0.95	0.82	0.33	0.61	0.34	0.65

variable tested. From 130 available variables, only six were selected by the MLP and seven by the MVRVM to provide forecasting of aggregate water demand into ACA Canal B for one and two days into the future.

The graphical representation of the best calibration of the forecasting models is presented in Figs. 4.4 and 4.5 using 2009 data. These figures present the approximation of the learning machines to the estimation of future ACA Canal B water demands for two days in advance. The figures also include a 1-standard error bar ($\sigma_y^{(n)}$), provided by each data-driven algorithm.

It is interesting to note the variables selected by each data-driven algorithms from the 130 available ones. The MLP algorithm required at least three previous inflow rates

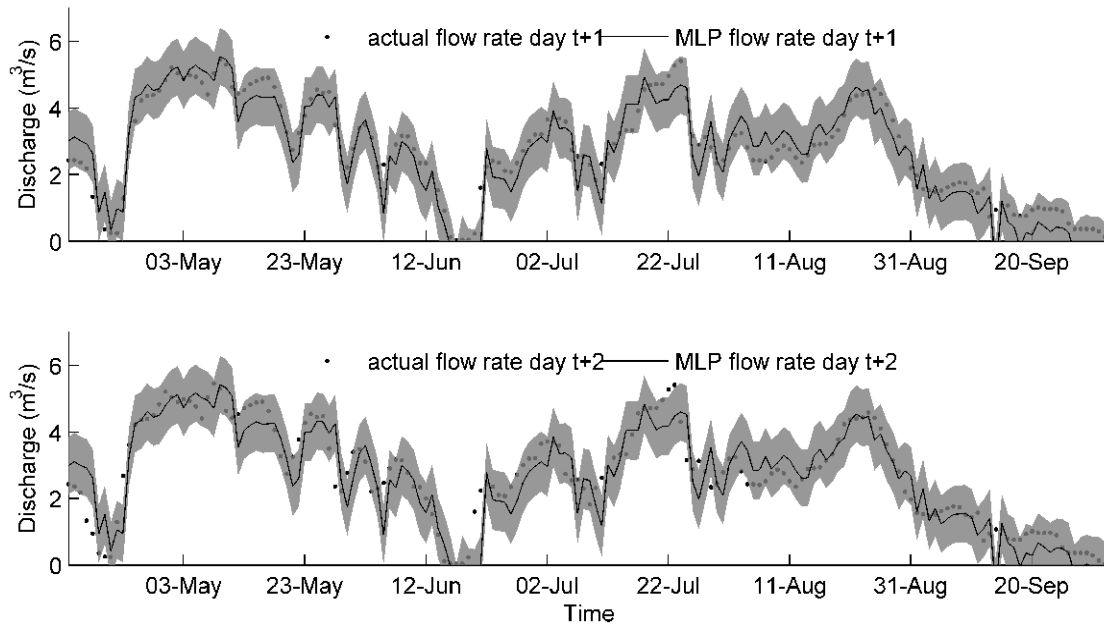


Fig. 4.4. One and two days water demand forecast for 2009 irrigation season using MLP.

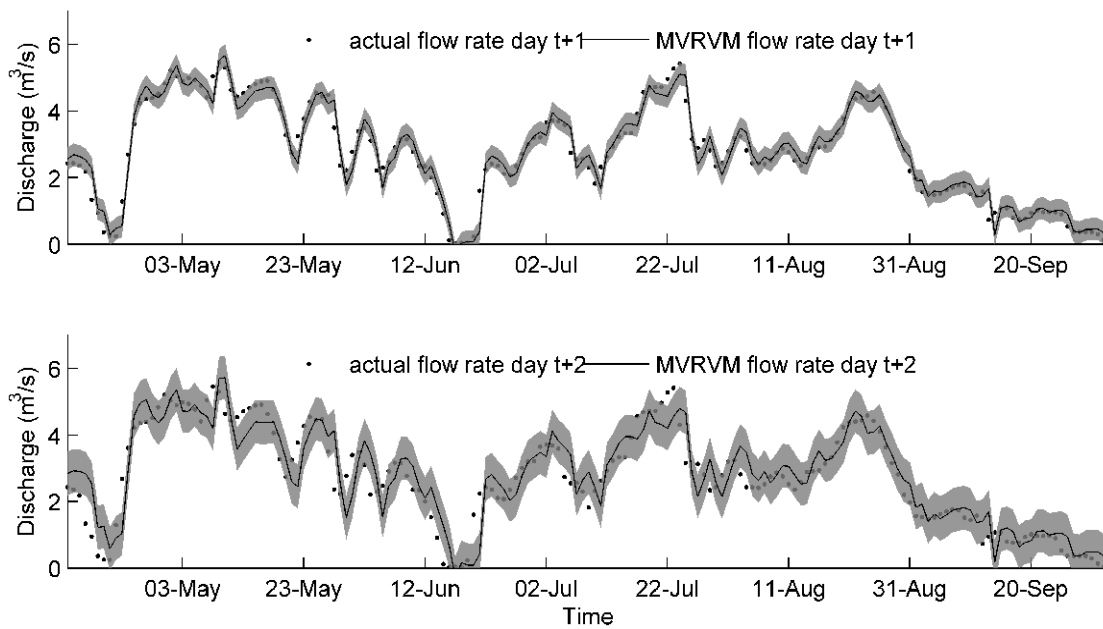


Fig. 4.5. One and two days water demand forecast for 2009 irrigation season using MVRVM.

of Canal B, plus the evapotranspiration requirements for alfalfa at time t and soil moisture information from sites 107-a, 109-a and 116-a at different depths and time steps. Similarly for the MVRVM, two previous inflow rates, alfalfa evapotranspiration requirements at time t , and soil moisture from monitoring sites 109-a and 115-b at different depths and time steps.

It is interesting to note the variables selected by each data-driven algorithms from the available ones mentioned in Table 4.3. The MLP algorithm required at least three previous inflow rates of Canal B, plus the evapotranspiration requirements for alfalfa at time t and soil moisture information from sites 107-a, 109-a and 116-a at different depths and time steps. Similarly for the MVRVM, two previous inflow rates, alfalfa evapotranspiration requirements at time t , and soil moisture from monitoring sites 109-a and 115-b at different depths and time steps.

The variables selected by the data-driven algorithms indicate their importance in forecasting future ACA Canal B water demand values. Alfalfa, as shown in Fig. 4.2 and Table 4.1, is by far the largest crop produced and the ACA water requirements are related to alfalfa water demand pattern. Soil moisture information from Site 109-a is required for both of the algorithms. Table 4.2 indicates this site has recorded soil moisture conditions for alfalfa for 2008 and 2009 irrigation seasons. This inclusion in both models might indicate that soil moisture from this site averages moisture conditions for alfalfa in the entire ACA. The other soil moisture sites included in the -forecasting model provide data on soil water conditions for corn and small grains in the ACA (107-a, 116-a, 115-b).

Now, when comparing the statistical (goodness-of-fit) results obtained for both data-driven algorithms for the best forecasting models, the MVRVM performs slightly

better than the MLP (Tables 4.3 and 4.4). The standard error bars from the best models can also provide insight about the performance of the models. These error bars are the sum of two effects: one from the error contained within the data, and the error from the data-driven algorithm itself. Therefore, the smaller the error bars the better is the approximation of the learning machine to the forecasted water demand values. Considering this concept, Tables 4.4 and 4.5, and Figs. 4.4 and 4.5 show clearly that MVRVM can provide a better correlation using available data for the ACA and forecasted water diversions than the MLP.

A consideration to keep in mind is the approximation of the forecasting results obtained for both mapping algorithms. From Figures 4.4 and 4.5 there is a lag between the forecasted and actual flow deliveries in Canal B. This is more evident for the second forecasted day. This is because of the available data, two irrigation seasons, one for training and testing. Additional information for training will reduce the time lag, enhancing the goodness-of-fit parameters at the same time.

Given the best variable combination and data-driven tuned parameters for both algorithms, it is important to assess the suitability of the models under different irrigation conditions than the 2008 and 2009 seasons. The assessment is conducted to determine the stability and robustness of the forecasting models when they are presented with previously unseen data. A 1000-fold bootstrap with replacement was applied to the training data (2008), while keeping the testing data (2009) constant. For each fold, the goodness-of-fit parameters were estimated and stored as explained by Anguita et al. (2000). To determine the characteristics of the bootstrap results, a graphical

representation of the goodness-of-fit statistics (Efron et al., 1993) is presented in Fig. 4.6 and 4.7.

Several conclusions can be drawn from the last figures. First, the range of the goodness-of-fit parameters in the presented histograms is smaller for the RVM when compared with the MLP for both of the goodness-of-fit parameters. This implies that once calibrated, the RVM is less susceptible of providing reduced performance due to the variation of the training data as could occur under actual working conditions. For the MLP, the statistical measures from the bootstrap show a more dispersed distribution, indicating that the MLP performance is affected by training data variation. This could be a strong limiting condition for practical applications to estimate future aggregate water demand using MLP as the mapping algorithm.

Conclusions and discussion

This study presents the findings of a proposed water demand forecasting model based on statistical learning machines and historical information of flow rates, crop water demand and soil moisture data for an agricultural command area (ACA). The practical application of the proposed model is tested in an irrigation system in Central Utah (Canal B).

For the forecasting model, the chosen data-driven algorithm is the Multivariate Relevance Vector Machine (MVRVM) and its performance is compared against the results of another data-driven tool, the Multilayer Perceptron (MLP). Over 130 possible inputs variables were tested to determine the most optimal combination of them to provide the forecasted flow rates.

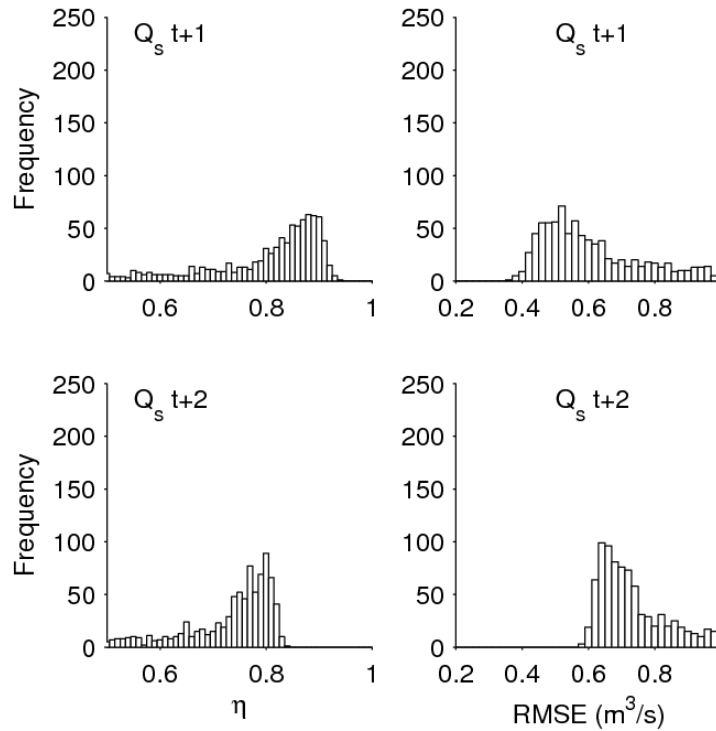


Fig. 4.6. η and RMSE Bootstrap results for the MLP (2009 irrigation season).

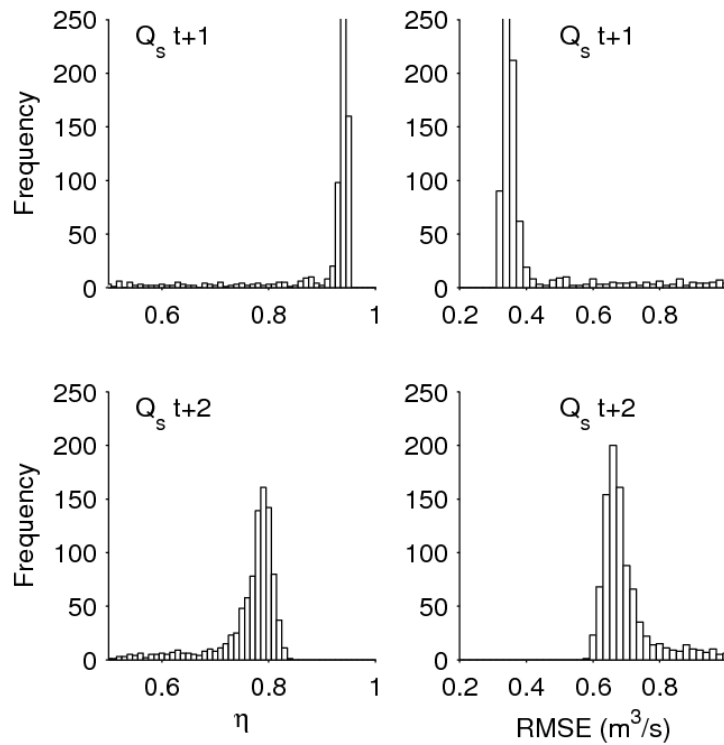


Fig. 4.7. η and RMSE Bootstrap results for the MVRVM (2009 irrigation season).

The results indicate that the proposed forecasting model is able to adequately estimate the future water deliveries for ACA Canal B up to two days in advance. For this, past inflow values, evapotranspiration of the crop with the largest planted acreage in the ACA (alfalfa), and soil moisture at one and two feet from soil moisture sites for the MVRVM algorithm was required. A similar set of variables was identified for the MLP algorithm.

The results also show that the past inflow rates are the most influential variables in the forecasting scheme. The Canal B inflow is related to the water management operations in the ACA (farmers' water orders, water masters and water manager). Thus, the human factor is an important component in the forecasting model.

When comparing the results obtained by the data-driven algorithms, the MVRVM performs better than the MLP as demonstrated by the goodness-of-fit values and the graphical analysis. Furthermore, when assessment of robustness and stability was performed by application of bootstrap analysis, again MVRVM was less affected by unseen new data than the MLP. This implies that the MLP (a type of neural network model) is less suitable for water demand forecasting tasks.

The approach to develop the forecasting model in this study allows for replication of the model under different scenarios and locations. It is advisable that data quality completeness of the available information is performed for the time intervals desired. Also, to reduce performance lags in the results it is advisable to include at least three irrigation seasons of daily data, two for training the MVRVM algorithm.

Future work on this topic will be related the extension of the model of two to four days by incorporation of detailed spatial information. Also, the MVRVM model will be implemented for use by water managers of the ACA Canal B.

References

- Anguita, D., Boni, A., Ridella, S. 2000. Evaluating the generalization ability of support vector machines through the bootstrap. *Neural Process Letters*, 11(1), 51- 58.
- Allen, R. G., Pereira, L. S., Raes, D., Smith, M. 1998. Crop evapotranspiration guidelines for computing crop water requirements. *Irrig. Drain., FAO*, 56, 300.
- Bellingham, B. K. 2009. Method for irrigation scheduling based on soil moisture data acquisition. 2009 Irrig. District Conf.
- Bishop, C. M. 1995. *Neural Networks For Pattern Recognition*. Oxford University Press, Oxford.
- Bontemps, C., Couture, S. 2002. Irrigation water demand for the decision maker. *Env. and Develop. Econ.*, 7, 643-657.
- D'Urso, G., Menenti, M., Santini, A. 1995. Remote sensing and simulation modeling for on-demand irrigation systems management. *Proc. CID/FAO Workshop on Irrig. Scheduling*, Rome.
- Efron, B., Tibshirani, R., Tibshirani, R.J. 1993. *An Introduction to The Bootstrap*. Chapman & Hall, Florida.
- Flake, J., Moon, T. K., McKee, M., Gunther, J. H. 2010. Application of the relevance vector machine to canal flow prediction in the Sevier River basin. *Agric. Water Manag.*, 97, 208-214.
- Gill, M. K., Asefa, T., Kemblowski, M. W., McKee, M. 2006. Soil moisture prediction using support vector machines. *Amer. Water Res. Assoc.*, 42, 1033—1046.
- Glovis USGS website, glovis.usgs.gov, accessed Nov. 2009.
- Guyon, I., Elisseeff A. 2003. An introduction to variable and feature selection. *J. Machine Learning Res.*, 3(3), 1157 - 1182.
- Herrero, J., Casterad, M. A. 1999. Using satellite and other data to estimate the annual water demand of an irrigation district. *Env. Monit. and Assessm.*, 55, 305-317.

- Houser, P. R., Shuttleworth, W. J., Famiglietti, J. S., Gupta, H. V., Syed, K. H., Goodrich, D. C. 1998. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Res. Res.*, 34(12), 3405-3420.
- Jensen, M. E., Robb, D. C. N., Franzoy, C. E. 1970. Scheduling irrigation using climate-crop-soil data. *Proc. Amer. Soc. of Civil Eng., J. Irrig. Drain. Div.*, 96, 25-38.
- Kim, J. H., Hwang, S. H., Shin, H. S. 2001. A neuro-genetic approach for daily water demand forecasting. *KSCE J. Civil Eng.*, 5(3), 281-288.
- Lecina, S., Playan, E. 2006. Model for the simulation of water flows in irrigation districts. i: description. *J. Irrig. Drain.*, 132, 310-321.
- MacKay, D. 1992. A practical bayesian framework for backpropagation networks. *Neural Computation*, 4(3), 448-472.
- McCuen, R. H., Knight, Z., Cutter, A. G. 2006. Evaluation of the Nash-Sutcliffe efficiency index. *J. Hydrol. Eng.*, 11(6), 597-602.
- Nabney, I. T. 2002. *NETLAB: Algorithms for pattern recognition*. Springer-Verlag New York, Inc., New York.
- Ojeda-Bustamante, W., Gonzales-Camacho, J. M., Sifuentes-Ibarra, E., Isidro, E., Rendon-Pimentel, L. 200). Using spatial information systems to improve water management in Mexico. *Agric. Water Manag.*, 89, 81-88.
- Perry, C., Steduto, P., Allen, R. A., Burt, C. M. 2009. Increasing productivity in irrigated agriculture: agronomic constraints and hydrological realities. *Agric. Water Manag.*, 96, 1517-1524.
- Pierce, S. G., Worden, K., Bezazi, A. 2008. Uncertainty analysis of a neural network used for fatigue lifetime prediction. *Mechanical Systems and Signal Processing*, 22(6), 1395 – 1411. Special Issue: Mechatronics.
- Pulido-Calvo, I., Montesinos, P., Roldan, J., Ruiz-Navarro, F. 2007. Linear regressions and neural networks approaches to water demand forecasting in irrigation districts with telemetry systems. *Biosystems Eng.*, 97, 283-293.
- Pulido-Calvo, I., Gutierrez-Estrada, J. C. 2009. Improved irrigation water demand forecasting using a soft-computing hybrid model. *Biosystems Eng.*, 102, 202-218.
- Sarwar, A., Perry, C. 2002. Increasing water productivity through deficit irrigation: evidence from the Indus Plains of Pakistan. *Irrig. and Drain.*, 51, 87-92.
- SRWUA - Sevier River Water Users Association website, www.sevierriver.org, accessed November, 2009.

- Svendsen, M. 2005. Irrigation and River Basin Management, opinions for Governance and Institutions. CABI Publishing, Colombo, Sri Lanka, 258 p.
- Thayananthan, A., Navaratnam, R., Stenger, B., Torr, P., Cipolla, R. 2008. Pose estimation and tracking using multivariate regression. *Pattern Recognition Letters*, 29(9), 1302–1310.
- Tipping, M. E. 2001. Sparse bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.*, 1, 211–244.
- Tipping, M. E., Faul, A. 2003. Fast marginal likelihood maximization for sparse bayesian models. *Proc., 9th Int. Workshop on Artificial Intelligence and Statistics*. 3–6.
- Tripathi, S., Govindaraju, R. 2007. On selection of kernel parameters in relevance vector machines for hydrologic applications. *Stochastic Env. Res. and Risk Assessment*, 21, 747–764.
- Walker, W. R., Stringam, B. L. 1999. Low cost adaptable canal automation for small canals. *ICIC Journal*, 48(3):39-46.
- Walker, W. R., Stringam, B. L. 2000. Canal automation for water conservation and improved flexibility. *Proc., 4th Decennial Nat. Irrig. Symposium*.
- Wright, J. L. 1982. New evaporation crop coefficients. *Irrig. and Drain. Div., ASCE Proc.*, 108, 57–74.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Summary and conclusions

The complexity of agricultural conditions and human decisions in an irrigation system require better, yet simple, approaches or methods to translate available, real-time data about the state of the system into valuable, decision-relevant information for water masters, water managers and decision-makers. This is the objective of this dissertation.

The methods presented here were developed based on a new statistical learning machine tool, the Relevance Vector Machines and its extended version, the Multivariate Relevance Vector Machine, which have been reported to be successful in many other fields related to water management problems.

Three real different issues in irrigation water management were analyzed, as shown in Chapters 2, 3 and 4. These issues are: 1) estimation of future crop water for water management purposes when climatic data is limited, 2) error correction or minimization in simulation models, and 3) generation of an aggregate water demand forecast based on actual agricultural conditions and irrigation system management.

To demonstrate the performance of the developed models a location in Central Utah in the Lower Sevier River Basin was selected, the agricultural command area (ACA) called Canal B. This ACA covers approximately 10,000 hectares. The main crops in the ACA are alfalfa, corn and barley. Water is conveyed to the ACA inlet by a 9 km canal, called Canal A, from a storage facility located upstream, the DMAD Reservoir. A SCADA system controls the irrigation system while the internal canal network flow condition in the ACA is operated by local water masters and water managers. The ACA

also has a network of on-farm soil moisture sensors that monitor soil water variation on an hourly basis.

For the three issues that are the central focus of this research, the analytic methods used here are strongly based on Bayesian learning machine models, the Relevance Vector Machine (RVM) and the Multivariate Relevance Vector Machine (MVRVM). The advantages of these models are their probabilistic approach to provide a mapping function among the available input-output data, while avoiding overfitting issues that could affect their performance, which have been often seen in previous engineering applications of data-driven models.

Chapter 2 presents the development of a method for estimation of future daily water crop demands, also called evapotranspiration (ET_0). The critical point here is the limited climatic information available, maximum and minimum daily air temperatures, recorded by the local weather station. Utilizing available data and a well-known ET_0 model, i.e., the 1985 Hargreaves Equation, future ET_0 values are mapped against past ones, using the MVRVM as the mapping function. Two questions are answered here: 1) How far in time can ET_0 be forecasted, and 2) is there any advantage of forecasting: the required weather variables for the ET_0 equation (daily air temperatures) or the already calculated ET_0 ?

The second issue addressed (Chapter 3) is the development of a method that allows the reduction of lumped errors that occur in hydraulic simulation models. These lumped or aggregate errors (ϵ_A) are the consequence of noise imbedded in the model parameters and inputs (parameter and observation errors respectively), and accuracy of the numerical approximation of the simulation model (system error) to the actual

phenomena. The objective for this analysis is to develop an error correction model that only uses information from the simulation model and that can provide a way to minimize the aggregate errors, thus reducing its impact on the simulated results. An additional requirement is that the error correction model should work under the same conditions as the simulation model (i.e., real-time conditions).

The third issue addressed in this research (Chapter 4) is the development of a model that estimates the irrigation water required for an ACA under its actual agricultural and water management conditions. There is not an easy answer for this question. Estimates of irrigation water requirements should be mostly driven by crop water needs and local soil characteristics. Nevertheless, on-farm management and the operation of the irrigation system are major components with a large impact on water that must be delivered to the ACA. The model should capture this information. Also the proposed model must be limited to only the available information in the ACA to provide the required forecasted information.

Additionally, for each of the models developed in Chapters 2 to 4, it is always important to determine their suitability for the issue addressed in terms of robustness, and accuracy in the presence of new, previously unseen data and in comparison with other widely used models in similar issues. For this reason, a performance comparison against a Bayesian Artificial Neural Network algorithm is proposed. This algorithm is the Multilayer Perceptron (MLP) which is also used as a mapping tool to develop models in a manner similar to the RVM and MVRVM. Thus, once calibrated, these mapping algorithms are subjected to a bootstrap analysis to determine their robustness when given new data, as if they were already implemented in the ACA.

The findings for the potential evapotranspiration forecast (Chapter 2) indicates that it is possible to estimate future crop water requirements in the Delta, Utah area for up to four days in advance using historical weather data and the MVRVM as mapping function. This method requires the mapping of weather variables required by the ET_0 model rather than calculated ET_0 values. The difference between using air temperatures or ET_0 in the mapping function is that the forecast of weather variables allows for larger forecast periods than the direct use of historical ET_0 records.

The results for the error correction model (Chapter 3) show that it is possible to develop a coupled physical- and statistical-based model that provides minimization of the lumped or aggregate error while performing in real-time. The mapping algorithm here is the RVM and the required information is a small set of inputs used by the simulation model. The mapping provided by the RVM among the lumped error and the inputs of the hydraulic simulation model can allow for an identification of possible error sources that could be examined later by water managers, such the conditions of water flow recording sensors, SCADA system, among others.

In Chapter 4, the results indicate the possibility of predicting short term water deliveries by making use of local and general information of an ACA. Here they key is the implementation of an SCADA system to provide a constant and reliable source of information. By making use of the MVRVM as a mapping function, water deliveries for several days in the future can be estimated.

An issue that has been denoted in this research is the influence of the amount of data for the calibration of these learning machines. As it has been demonstrated in this study, an irrigation season of daily or hourly data allows for general calibration and

variable selection. Nevertheless, to improve the accuracy of the models, more information is required (2 or more irrigation seasons) for training purposes. In this way the behavior of the variable to be modeled can be done with higher accuracy.

In general, the results for each of the methods developed were satisfactory. The proposed mapping function (RVM or MVRVM) is an excellent choice to perform the required mapping task. Also these algorithms have the advantage of being more robust and stable than the alternative, the MLP, as was found in each analysis.

An advantage of using these Bayesian-based algorithms is their reduced time for calibration (no crossvalidation techniques required), probabilistic approach and estimation of error bars for the model results and the weights. The error bar estimation has not been exploited at their full potential so far e.g. indication of adequacy of the suitability of the model to the data, outlier identification or measurement of the noise or variation in the data.

Recommendations for future work

The work presented here is focused on developing methods to provide adequate information to water managers and others. The information used comes from historical records from an existing SCADA system (temporal data). Thus there is a need to explore the use of a combination of temporal-spatial data to produce better information.

Also, the good performance of the methods developed here opens the doors to other questions. One is related to the future performance of the methods once implemented for everyday use by water managers and others. The bootstrap analysis provides some insight about this, allowing selecting the mapping algorithm that is more robust and stable. Nevertheless, an on field test of the proposed models is recommended.

This test is important for several reasons: 1) it allows for feedback from the final users, water managers and others, to the modeler to suit the model in a comprehensible manner for them, 2) it allows the modeler to refine the model, to adjust it for actual working conditions, and 3) it permits the users to adapt to the new sources of information.

Another question is related to the frequency of recalibration required for the learning machine once implemented for use. This is related with the frequency of data generation (hrs, days), agricultural season pattern, etc. Additional work during the field test period, as proposed above, can provide insight about this.

A final question is related to the black-box algorithm concept that is commonly applied to learning machine algorithms. The RVM and MVRVM, given their conceptualization, can be explained in better fashion than earlier types of learning machines (e.g., Artificial Neural Networks). Still, additional work is necessary to illustrate the internal concepts used like the prior likelihood and posterior probabilities, and others, and provide these as an outcome from the algorithm. In any application of RVM-type machines for modeling of hydrologic or hydraulic processes, the question of the physical meaning of the choice of relevance vectors always arises. This is sometimes easily answered, such as for groundwater forecasting or monitoring applications (see Ammar et al., 2008, and Asefa et al., 2005), but remains unclear for time series applications such as the ones in this research. Further research is needed here.

References

Ammar, K., Khalil A., McKee M., Kaluarachchi J. 2008, Bayesian deduction for redundancy detection in groundwater quality monitoring networks, *Water Res. Res.*, 44, W08412, doi:10.1029/2006WR005616.

Asefa, T., Kemblowski M., Lall U., Urroz G. 2005, Support vector machines for nonlinear state space reconstruction: Application to the Great Salt Lake time series, *Water Res. Res.*, 41, W12422, doi:10.1029/2004WR003785.

APPENDIX

ELSEVIER LICENSE

This is a License Agreement between Alfonso F Torres (You) and Elsevier (Elsevier). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

License number

Reference confirmation email for license number

License date

Jan 14, 2011

Licensed content publisher

Elsevier

Licensed content publication

Agricultural Water Management

Licensed content title

Forecasting daily potential evapotranspiration using machine learning and limited climatic data

Licensed content author

Alfonso F. Torres, Wynn R. Walker, Mac McKee

Licensed content date

15 December 2010

Licensed content volume number: n/a

Licensed content issue number: n/a

Number of pages: 1

Type of Use: reuse in a thesis/dissertation

Portion: full article

Format: both print and electronic

Are you the author of this Elsevier article? Yes

Will you be translating? No

Order reference number

Title of your thesis/dissertation:

BAYESIAN DATA-DRIVEN MODELS FOR IRRIGATION WATER
MANAGEMENT

Expected completion date: May 2011

Estimated size (number of pages): 150

Elsevier VAT number: GB 494 6272 12

Billing Type

Invoice

Billing address: 1600 Canyon Rd, Logan, UT 84321, United States

Customer reference info

Permissions price: 0.00 USD

Value added tax 0.0%

0.00 USD / GBP

Total: 0.00 USD

ELSEVIER LICENSE

TERMS AND CONDITIONS

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking accept in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. (CCC), at the time that you opened your Rightslink account and that are available at any time at <http://myaccount.copyright.com>).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE

SOCIETY COPYRIGHT OWNER]. Also Lancet special credit - Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier.

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier at permissions@elsevier.com)

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your

proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. Translation: This permission is granted for non-exclusive world English rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article. If this license is to re-use 1 or 2 figures then permission is granted for non-exclusive world rights in all languages.

16. Website: The following terms and conditions apply to electronic reserve and author websites:

Electronic reserve: If licensed material is to be posted to website, the web site is to be password-protected and made available only to bona fide students registered on a relevant course if: This license was made in connection with a course, This permission is granted for 1 year only. You may obtain a license for future website posting, All content posted to the web site must maintain the copyright information line on the bottom of each image, A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> or the Elsevier homepage for books at <http://www.elsevier.com> , and Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

17. Author website for journals with the following additional clauses:

All content posted to the web site must maintain the copyright information line on the bottom of each image, and the permission granted is limited to the personal version of your paper. You are not allowed to download and post the published electronic version of your article (whether PDF or HTML, proof or final version), nor may you scan the printed edition to create an electronic version,

A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx> , As part of our normal production process, you will receive an e-mail notice when your article appears on Elsevier's online service ScienceDirect (www.sciencedirect.com). That e-mail will include the article's Digital Object Identifier (DOI). This number provides the electronic link to the published article and should be included in the posting of your personal version. We ask that you wait until you receive this e-mail and have the DOI to do any posting. Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

18. Author website for books with the following additional clauses:

Authors are permitted to place a brief summary of their work online only. A hyper-text must be included to the Elsevier homepage at <http://www.elsevier.com> All content posted to the web site must maintain the copyright information line on the bottom of each image You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version.

Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

19. Website (regular and for author): A hyper-text must be included to the Homepage of the journal from which you are licensing at <http://www.sciencedirect.com/science/journal/xxxxx>. or for books to the Elsevier homepage at <http://www.elsevier.com>

20. Thesis/Dissertation: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission.

21. Other Conditions: None

v1.6

Gratis licenses (referencing \$0 in the Total field) are free. Please retain this printable license for your reference. No payment is required.

If you would like to pay for this license now, please remit this license along with your payment made payable to COPYRIGHT CLEARANCE CENTER otherwise you will be invoiced within 48 hours of the license date. Payment should be in the form of a check or

money order referencing your account number and this invoice number RLNK10781447.

Once you receive your invoice for this order, you may pay your invoice by credit card.

Please follow instructions provided at that time.

Make Payment To:

Copyright Clearance Center

Dept 001

P.O. Box 843006

Boston, MA 02284-3006

If you find copyrighted material related to this license will not be used and wish to cancel, please contact us referencing this license number 2425520035814 and noting the reason for cancellation.

Questions? customer@copyright.com or +1-877-622-5543 (toll free in the US) or +1-978-646-2777.

Permission to include papers in dissertation

Alfonso Torres <a.torres@aggiemail.usu.edu> Tue, Jan 18, 2011 at 12:56 AM

To: Andres Ticlavilca <Andres.T@aggiemail.usu.edu>

Dear Andres,

This mail is to request a permission to include in my dissertation the complete copy of the following papers in which you appear as a coauthor:

Machine Learning Approach for Error Correction of Hydraulic Simulation Models

Multiple Day Irrigation Water Demand Forecast Using Multivariate Relevance Vector
Machines

Thank you for your help,

Alfonso F. Torres

Andres Ticlavilca <andres.t@aggiemail.usu.edu> Tue, Jan 18, 2011 at 9:20 AM

To: Alfonso Torres <a.torres@aggiemail.usu.edu>

Dear Alfonso,

You have my permission

Andres

CURRICULUM VITAE

Alfonso Faustino Torres Rua

Ph.D Research Assistant

Department of Civil and Environmental Engineering

Utah State University

Utah Water Research Laboratory (UWRL), Logan Utah 84321

Tel: (435) 890 0196; email: a.torres@aggiemail.usu.edu

EDUCATION:

Doctor of Philosophy, Civil and Environmental Engineering, Aug 2011, Utah State University, Logan, Utah.

Major: Water Resources Management

Dissertation Title: Bayesian Data Driven Models for Irrigation Water Management

Master of Science, Biological and Irrigation Engineering, December 2006, Utah State University, Logan, Utah.

Major: Irrigation Engineering

Dissertation Title: Transition Submergence and Hysteresis Effects in Three-Foot Cutthroat Flumes

Bachelor of Science, Agricultural Engineering, December 2000, La Molina National Agrarian University, Lima, Peru.

Dissertation Title: Technical and Economical Assessment of Implemented Dripping Irrigation in the Agrocurumy Ranch, Piura

RESEARCH ASSISTANT II: (Jan 2007 – Aug 2011) Utah Water Resources Laboratory, Utah State University (USU)

- Installation and maintenance of soil moisture data collection system in Delta, UT (USBR).

- Lecturer for graduate level classes at USU.
- Analysis and improvement of water management models for Lake Fork River, UT (USBR, WUA).
- Team member of the USU Technology Transfer for Water Management Project (Peru).
- Organizer and Speaker of the USU's First Workshop on Application of Statistical Learning Machines to Water Management (2009).

RESEARCH ASSISTANT I: (Jan – Dec 2006) Biological and Irrigation Engineering, USU

- In charged of development and testing of new models to improve flow rate estimation in Cutthroat flumes.
- Assisted in evaluation and implementation of water control structures in the Logan River Basin.

TECHNICAL CONSULTANT II: (Jan 2004 – Aug 2005) General Agency of Multiannual Programming (DGPM - Spanish acronym), Ministry of Economy and Finances of Peru

- Responsible for orientation, integration, follow-up, and evaluation of multiannual plans at regional and local levels, including programming of the public investment and development in under-developed regions.
- Responsible of economic and social feasibility of government projects in agriculture, water, and others.
- Speaker at workshops related with implementation and execution of projects using public funds.
- Team member for the implementation of risk analysis in public-funded projects design and execution (climate change and others)

TECHNICAL CONSULTANT I: (Mar 2002 – Dec 2003) Office of Public Investment (ODI –Spanish acronym), Ministry of Economy and Finances of Peru

- Team member for the implementation of National Disaster Relief Programs for the 1998 El Niño Phenomenon and 2001 Earthquake in Peru. (World Bank and Inter-American Development Bank).
- Field visits to disaster areas to monitor progress of rebuilding damaged infrastructure.
- In charged of monitoring financial execution of projects to restore basic services (agriculture, education, roads, etc.), in disaster areas.

LANGUAGES: Spanish (Native), English

HONORS:

La Molina National Agrarian University, Class 2000-II (BS Valedictorian)

Spring Runoff Conference 2010, Utah State University – Presentation Award

PUBLICATIONS:

- Forecasting of Potential Evapotranspiration using Learning Machines, Journal of Hydrology, 2010
- Improved Calibration of Cutthroat Flumes, Irrigation Science, 2007
- Transition Submergence and Hysteresis Effects in Three-Foot Cutthroat Flumes, USCID Conference, 2007.
- Cutthroat Measurement Flume Calibration for Free- and Submerged-Flow Using a Single, ASCE, 2007.
- Analysis of Submergence and Hysteresis Phenomenon in Cutthroat Flumes. MS Thesis. USU, 2006.
- Technical and Economic Assessment of Pressurized Irrigation System in Agrocurumuy Field, Piura, Peru. BS Thesis. La Molina National Agrarian University, 2001.