ASSESSMENT OF DATA-DRIVEN MODELING STRATEGIES FOR WATER DELIVERY CANALS

Isaías Tavares [*,1] **José Borges** [**,1] **Mário Mendes** [***,1]
Miguel Ayala Botto [*,1]

[*] *IDMEC / Instituto Superior Técnico, Technical University of Lisbon, Av. Rovisco Pais, 1049-001 Lisboa, Portugal. {isaias.tavares@ist.utl.pt, ayalabotto@ist.utl.pt}*
[**] *IDMEC / Departamento de Ciências e Tecnologias de Engenharia da Academia Militar - Rua Gomes Freire, 1169-203 Lisboa, Portugal. {jose.borges@academiamilitar.pt}*
[***] *IDMEC/ ISEL - Instituto Superior de Engenharia de Lisboa, Polytechnic Institute of Lisbon, Rua Conselheiro Emídio Navarro 1, 1959-007 Lisboa. {mmendes@dem.isel.pt}*

Abstract: The aim of this work is to develop nonlinear dynamical models for the canal system of Núcleo de Hidráulica e Controlo de Canais. The canal is a nonlinear system and thus should be modeled to meet given operational requirements, while capturing all relevant system dynamics, such as the resonance waves created due to the movements of gates, and also contributing to the controller precision. The nonlinear modeling is based on data-driven methods, namely Composite Local Linear Models, Fuzzy Models and Artificial Neural Networks. These models are identified using data collected from the experimental facility, and their performance is assessed based on suitable validation criteria. The modeling results show the effectiveness of these models while capturing all significant dynamics for the canal system. (Copyright CONTROLO2012)

Keywords: Nonlinear Modeling, Water Canal System, Composite Local Linear Models, Fuzzy Models, Artificial Neural Networks

1. INTRODUCTION

New challenges, such as climate change and desertification in some regions, demand for the careful management of water resources. The use of formal systems control theory applied to water delivery canals appears as a powerful solution to handle the problem of water scarcity. A contribution towards the solution of such problem in the southern region of Portugal is proposed through the AQUANET research project.

[1] Research for this paper was partially supported by project *AQUANET - Decentralised and Reconfigurable Control for Water delivery Multipurpose Canal Systems*, PTDC/EEA-CRO/102102/2008, FCT, Portugal and Fundaçgo para a Ciência e a Tecnologia, through IDMEC under LAETA.

The aim of such project is to develop and implement modeling and control strategies, which are then tested and validated experimentally in a canal systems from Núcleo de Hidráulica e Controlo de Canais (NuHCC), Évora University.

The canals should be managed in such a way as to create flexible conformities on everything that depend on it. This requires water readiness for use at any time and foreseeable situation. To this end, an elaborate system of interconnections between various canal systems should be done in order to have optimal management of water. Since the development is supported by the efficiency that comes in the transport of water, minimal losses need to be guaranteed through canal control. The efficiency of a canal can be described as the ability to maintain a certain water level. The difficulties in predicting the moments of demand, as well as the amount to use at any given time has proved to be one of the main challenges from the control perspective. Hence the need for models with a good level of accuracy in estimating the canal's output in order to calculate the control actions optimal.

The modeling of canals, similarly to other complex processes, can be addressed using classical modeling approaches: first principles or data-driven system identification. Therefore, several authors have contributed to the modeling of canals systems. They developed different models with different identification tools. The first principles modeling using a complete mathematical model with the equations of Saint-Venant's boundary conditions for both internal and external supplement was proposed by (Coron *et al.*, 1999). Several authors have applied data-driven modeling approaches to canal systems, e.g. complete linearized model (Xu and Sallet, 1998). Simplifications to the Saint-Venant Equations resulted in the form of infinite order linear transfer functions (Baume *et al.*, 1998), finite order nonlinear models (Liu *et al.*, 1995) and the combination of finite order linear models and state space models (Kosuth, 1994).

Data-driven modeling based on the use of Artificial Neural Networks (ANN) (Toudeft, 1995) or Fuzzy Models (FM) (Voron and Bouillot, 1997) was also proposed. Composite Local Linear Models (CLLM) (Borges *et al.*, 2011) were successfully applied to model a first principles model for the same canal system that is being addressed in the current paper. The main advantages of these methods come from their ability to provide nonlinear models using well established tools.

The work presented in this paper addresses the modeling of the experimental NuHCC canal system located in Évora, Portugal using CLLM, FM and ANN. The resulting models are presented, as well as suggestions for how to overcome problems while using these tools. The experimental input/output data was collected from the NuHCC facility and the models are assessed using both training and validation datasets based on the *mean squared error* and *variance accounted for* performance criteria. The comparison of performance for the resulting data-driven models is also discussed.

2. DATA-DRIVEN TOOLS

Given the difficult analysis of nonlinear systems, often the solution to confront these systems consists in linearizing them, thus avoiding nonlinear aspects. This procedure may not be the best to completely describe the system dynamics, especially for cases where the dynamic behavior is nonlinear. In the case of the canal system, the modeling shows dispersive effects due to nonlinear propagation of long waves (Benjamin *et al.*, 1972). The design of model based controllers is highly influenced by this effect (Wagemaker, 2005). Therefore, it is required to address the data-driven modeling problem using nonlinear system identification tools.

2.1 *Composite local linear models*

Composite Local Linear Model (Verdult *et al.*, 2002) are defined as,

$$x(k+1) = \sum_{i=1}^{s} p_i\left(\phi(k)\right)\left(A_i x(k) + B_i u(k) + O_i\right) \tag{1}$$

$$y(k) = Cx(k) + v(k) \tag{2}$$

where s is the number of local models, O_i denotes the state offsets, $x(k) \in \mathbb{R}^n$ is the state vector, $u(k) \in \mathbb{R}^m$ is the input, $y(k) \in \mathbb{R}^\ell$ is the output, $v(k) \in \mathbb{R}^\ell$ is a white-noise sequence and $p_i\left(\phi(k)\right) \in \mathbb{R}^s$ are parameterized using normalized radial basis functions to schedule the combination of local models,

$$p_i\left(\phi_k\right) = \frac{r_i\left(\phi(k); c_i; w_i\right)}{\sum_{j=1}^{s} r_j\left(\phi(k); c_i; w_i\right)} \tag{3}$$

with the following radial basis functions,

$$r_i = \exp\left(-\left(\phi(k) - c_i\right)^T diag\left(w_i\right)^2 \left(\phi(k) - c_i\right)\right) \tag{4}$$

where c_i is the center and w_i the width of the $\imath$-th radial basis function, $\phi(k) \in \mathbb{R}^q$ is the scheduling vector, defined in an operating point of the system and depends on the input and the system state,

$$\phi(k) = \psi\left(x(k), u(k)\right) \tag{5}$$

with $\psi : \mathbb{R}^{n\times m} \rightarrow \mathbb{R}^q$, which is nonlinear and known. The goal is to determine, from a finite number of measurements of the input $u(k)$ and output $y(k)$, the matrices A_i, B_i, O_i, C and the centers c_i and widths w_i that describe the radial basis functions. These parameter result from the optimization of the cost function $J_N(\theta, c, w)$, which is derived from the minimization of an error vector.

The local model can be initialized using subspace identification (Verhaegen, 1994) and the initial values for radial basis function are distributed uniformly over the operating range.

The combination of separable least squares and projected gradient was proposed by (Borges *et al.*, 2004) to deal with non-uniqueness problems of state-space matrices along the optimization process.

2.2 *Fuzzy Models*

Fuzzy logic and *fuzzy sets* were introduced by Zadeh in 1965. The fuzzy models can be built by encoding expert knowledge into linguist rules. The modeling and control of dynamic systems in the presence of lack of precision, strong nonlinear dynamics and a high degree of uncertainty, can be addressed using *fuzzy logic* (Babuska, 1998). The use of fuzzy modeling in system identification and control, fault diagnostic, classification and decision-support system is addressed e.g. in (Roubos, 2002). These systems can be modeled using *fuzzy models*, where the relationships between variables are defined using *"if-then"* rules, such as,

$$\textbf{If} \quad antecedent \quad \textbf{then} \quad consequent$$

where the propositions are of type "x is A", where x is the linguistic variable and A is the linguistic constant.

The mathematical framework which makes use of fuzzy sets for both static and dynamic systems is known as *fuzzy systems* (Babuska, 1998). Fuzzy systems can process and evaluate different types of information such as crisp, interval, and fuzzy. When the output is crisp a fuzzifier is used.

Takagi-Sugeno fuzzy models approximate dynamic systems using state-transition functions with a NARX structure,
If $y(k)$ is A_{i1} **and** $y(k-1)$ is A_{i2} ...
and $y(k-n_y+1)$ is A_{in_y} **and** $u(k)$ is B_{i1}
and $u(k-1)$ is B_{i2} **and** ... $u(k-n_u+1)$ is B_{in_u}
Then $y(k+1) = \sum_{j=1}^{n_y} a_{ij} y(k-j+1) + \sum_{j=1}^{n_u} b_{ij} u(s-j+1) + c_i$
where $y(k-1), ..., y(k-n_y+1)$ denote past model outputs, and $u(k-1), ..., u(k-n_u+1)$ denote past model inputs.

2.3 *Artificial Neural Networks*

Artificial Neural Network can be seen as a functional imitation of biological neural networks. ANN are considered to be black-box models, where it is assumed that no previous knowledge about the process is needed to derive the mapping (either linear or nonlinear) between the system inputs and outputs. Rather than that, this kind of nonparametric models are derived solely form the process data. ANN have important properties for the approximation of continuous nonlinear relationship: they have self-learning ability and they are able to extract process features from data using a training algorithm. These properties provide the modeling of nonlinear processes with great flexibility (Haykin, 1999).

Figures 1 and 2 generically present two network architectures, *series-parallel* and *parallel*, for ANN-NARX models. The backward time-shift operator is represented as z^{-1}, i.e. $x(k-1) = z^{-1}x(k)$, while $\hat{y}_i$ are the outputs and x_k the inputs. In nonlinear modeling using ANN the choice of the network to represent a nonlinear physical process depends on the dynamics and complexity of the network that is best for representing the problem at hand. E.g. recurrent networks suffer from instability and sensitivity due to noise, on the other hand, feedforward networks might not be powerful enough to capture the dynamics of the nonlinear dynamical system (Mandic and Chambers, 2001).

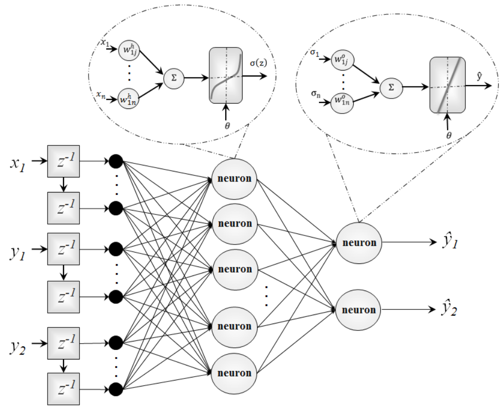

Fig. 1. Series-parallel ANN architecture.

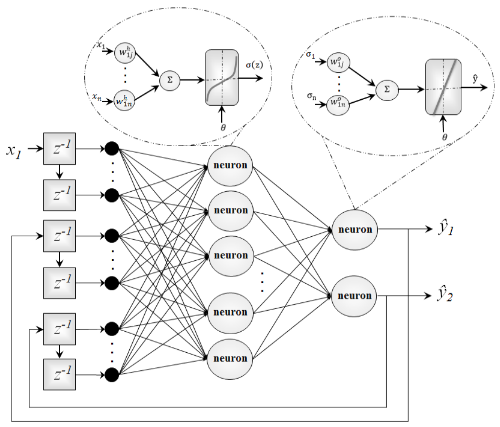

Fig. 2. Parallel ANN architecture.

The modeling process is based on finding the neural network that fits an ARX model equation,

$$\hat{y}(k) = \hat{f}(y(k-1), \ldots, y(k-n), u(k-1), \ldots, u(k-m)) \quad (6)$$

In order to obtain the ANN-NARX models some steps have to be made: process data collection; selection of a model structure; model estimation; model validation. The Levenberg-Marquardt algorithm was chosen to train the ANN-ARX networks because is the fastest algorithm, and with better results, on function approximation problems. In general, on function approximation problems, for networks containing up to a few hundred weights, the Levenberg-Marquardt algorithm will have the fastest convergence.

3. DESCRIPTION OF WATER CANAL SYSTEM

The NuHCC canal system consists of four elements: central station, automatic canal, traditional canal and two storage reservoirs. Figure 3 presents a schematic representation of both the automatic canal and the main reservoir. The traditional canal differs from the automatic canal because it is not instrumented, it is a return canal to ensure that it operates in closed-loop. The smaller reservoir provides flow to the automatic

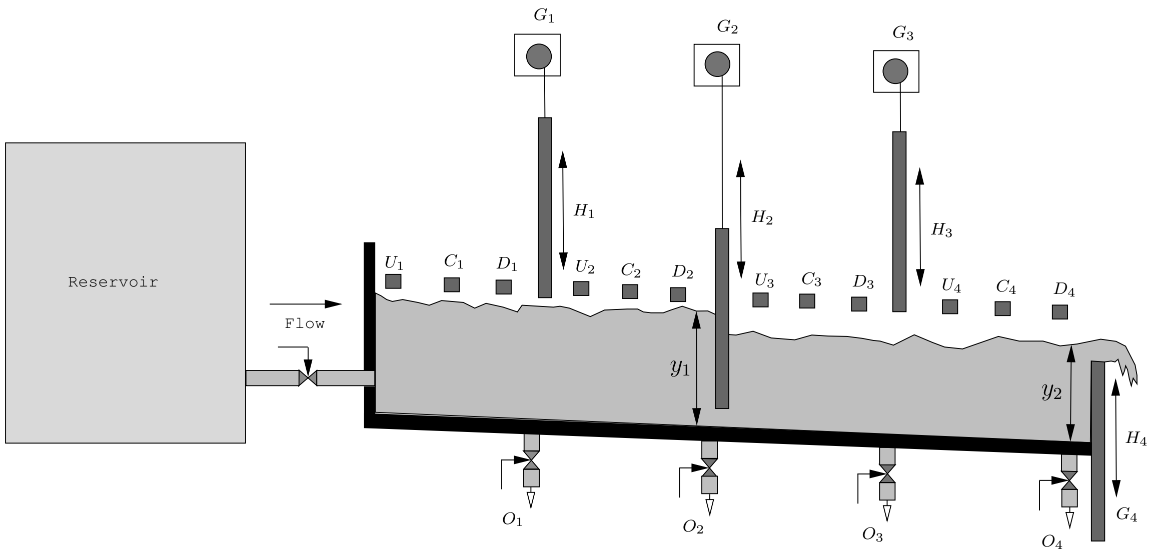

Fig. 3. Schematic drawing for the NuHCC canal system.

canal to ensure the functioning of two submersible pumps that increase the flow from the smaller reservoir. The central station performs the remote control and monitoring via a central server equipped with a *SCADA system*.

The automatic canal has four trapezoidal sections. The *orifice gate* in the first three sections has lengths of 35 m and the last *overshot gate* has 36 m. It can also highlight other features such as trapezoidal cross section with 0.15 m wide crawler, backrest slope of 1:0.15, height of 0.90 m, total length of 141 m, average longitudinal slope of 1.5×10^{-3}. The project has a nominal flow rate of 0.09 $m^3 \times s^{-1}$. Free surface is measured by a set of three sensors fitted in each section, a downstream (*D*), a center (*C*) and an upstream (*U*). The gates located at the end of each section are actuated by electric motors. The first three gates allow the flow under the gate *orifice gate*, and the last gate allows flow above the gate *overshot gate*. The gates were modeled taking into account the saturation of the actuators installed on them, thus opening and closing with limited speed. On the wall of the canal there is a hole upstream, which takes water inlets in order to discharge the flow into the return canal (called traditional canal). These water offtakes are controlled by a butterfly type valve, monitored by an electromagnetic meter. With 300 mm in diameter, one electric meter by regulation, operates a *MONOVAR valve* which allows the admission of the flow of the reservoir.

In terms of instrumentation, the canal of NuHCC is equipped with water level sensors, electromagnetic flowmeters and motorized butterfly valves and sluice gates. The control is done using electro-actuators and sensors which are connected to local PLCs (Programmable Logic Controllers). All local PLCs are connected through a MODBUS network.

For purposes of modeling, the excitation variables of the NuHCC open-canal (used as input) are: the flow from the reservoir, the flow of water intakes on the side of the canal (also designated offtake), lift heights of the gates (both overshoot and undershoot). In the canal is collected as output the water depth downstream of each existing pool.

Quality criteria
The models derived from channel with different modeling techniques will be evaluated according to three criteria of quality. These criteria are:

- VAF (Variance Account For):

$$VAF = 100\% \times \left(1 - \frac{var(y_i - \hat{y}_i)}{var(y_i)}\right) \quad (7)$$

- MSE (Mean Square Error)

$$MSE = \frac{\sum_i^N (y_i - \hat{y}_i)^2}{N} \quad (8)$$

- Computational Cost: the time for each modeling technique to estimate the model parameters.

VAF and MSE are ways to evaluate the difference between an estimator and the true value of the quantity being estimated. Where y_i is the process output, $\hat{y}_i$ is the estimated output (model output), N is the number of samples. When the signals are the same, the *VAF* is 100% and if they differ, it will be close to 0.

4. EXPERIMENTAL MODELING

The configuration used to collect data consists of two types of the undershoot sluice gates open (including the first and the third sluice gates), the second sluice gate (type undershoot) and the last sluice gate (type overshoot) closed to form two pools, offtakes at upstream and downstream pools (Figure 3). In this configuration eight tests were performed:

(1) Step sequence in upstream inflow.

(2) Step sequence in middle gate elevation.
(3) Step sequence in downstream gate elevation.
(4) Input sequence in upstream inflow, middle gate elevation and downstream gate elevation.
(5) Short sequence in upstream inflow.
(6) Short sequence in middle gate elevation.
(7) Short sequence in downstream gate elevation.
(8) Short sequence in upstream inflow, middle gate elevation and downstream gate elevation.

Regarding the modeling task, 70% of data were used for training, and 30% for validation. The model structure for CLLLM, FM and ANN models that have resulted from system identification are summarized in tables 1 to 4.

Table 1. CLLM model structure.

CLLM	Parameters
Number of local models	4
System order	2
Block size	20
Center limit	$[-0.25, 0.25]$

Table 2. FM model structure.

FM	Parameters
Number of cluster	$[2; 2]$
Delay input	$[1; 1; 1; 1]$
Delay output	$[1; 1; 1; 0; 1; 1; 1; 1; 0; 1]$
Transport delay	$[1; 1; 1; 0; 1; 1; 1; 1; 0; 1]$

Table 3. ANN-SP model structure.

ANN-SP	Parameters
Number of neurons	2
Delay input	1
Delay output	$[1, 2]$

Table 4. ANN-P model structure.

ANN-P	Parameters
Number of neurons	2
Delay input	1
Delay output	$[1, 2]$

It can be seen from Figure 4 that all the models, except ANN-SP, have initialization problems in case of both outputs, although there is a convergence towards the experimental output, except for FM in case of the output for the upstream pool. Although these models reveal that the series-parallel configuration performs better while fitting both outputs, in the parallel configuration, this adjustment is less accomplished. This result is due to the structure of neural network dynamics, including the procedures for computing the delay.

It can be seen from both figures 4 that the ANN model in the series-parallel configuration presents the best modelling result for pool 1, followed by CLLM.

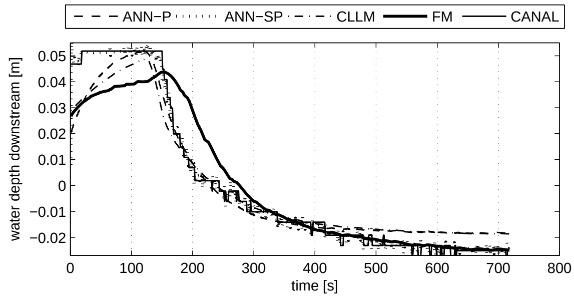

Fig. 4. Downstream water depth for pool 1.

In case of pool 2, the best fitting is again achieved by the ANN in the series-parallel configuration, followed by the ANN in a parallel configuration.

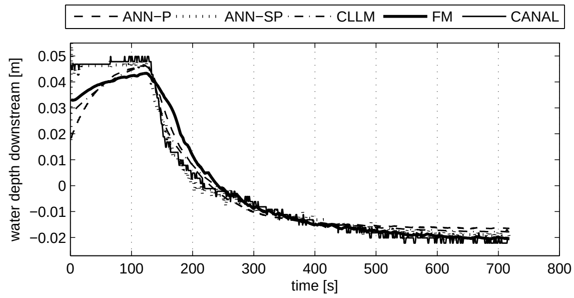

Fig. 5. Downstream water depth for pool 2.

The values for the quality criteria (VAF, MSE and Computational Time) are presented in tables 5 to 7.

Table 5. VAF and MSE using training data.

	VAF (%)		MSE	
	Pool 1	Pool 2	Pool 1	Pool 2
CLLM	98.88	99.21	0.38×10^{-4}	0.24×10^{-4}
FM	99.29	99.55	0.29×10^{-4}	0.18×10^{-4}
ANN-SP	99.77	99.76	0.77×10^{-5}	0.72×10^{-5}
ANN-P	97.44	97.94	0.87×10^{-4}	0.62×10^{-4}

Table 6. VAF and MSE using validation data.

	VAF (%)		MSE	
	Pool 1	Pool 2	Pool 1	Pool 2
CLLM	94.05	96.47	0.51×10^{-4}	0.23×10^{-4}
FM	91.68	95.70	0.36×10^{-3}	0.17×10^{-3}
ANN-SP	99.25	98.97	0.63×10^{-5}	0.71×10^{-5}
ANN-P	94.89	94.33	0.43×10^{-4}	0.66×10^{-4}

The results of the computational costs presented here are dependent on the configuration parameters. For neural networks models, this cost can be reduced if decrease the values of some criteria like: epoch (by

Table 7. Computational time.

Models	CPUtime in seconds
CLLM	5.1935×10^3
FM	113.65
ANN-SP	68.98
ANN-P	7.5806×10^3

default 1000), MSE's performance, by default 10^{-6}, and gradient, by default 10^{-12}. For composite local linear models the cost can be reduced by reducing the number of iterations (by default 400) or the error tolerance to be achieved 10^{-12}. For fuzzy models this cost can be reduced by reducing the tolerance of the stopping criterion, by default 10^{-2}. The configuration of these parameters affects both the computational cost and the quality of the models, thus defining them in terms of goal that you want.

5. CONCLUSIONS

The results of models derived using CLLM to the two pools configuration showed that the resulting models (VAF1$_{CLLM}$ $= 98.3093$, VAF2$_{CLLM}$ $= 97.8480$, VAF1$_{FM}$ $= 98.8007$, VAF2$_{FM}$ $= 99.1299$, VAF1$_{ANN}$ $= 99.8365$, VAF2$_{ANN}$ $= 99.8311$, where 1 - upstream pool, 2 - downstream pool) is a good approach to canal systems modeling.

The models derived using FM, for two a pools configuration, showed lesser results compared with other tools used, except for the ANN with parallel configuration. The reason might be due to the large number of parameters that need to be tuned.

The models derived from ANN with series-parallel configuration for two pools configuration showed better results, VAF and MSE, but the modeling process took longer and the model variability is higher. This problem can be addressed through the use of filtered data, or different training settings to modify the stopping criterion (MSE, number of epoch, gradient, etc). In the case of models derived form ANN with parallel configuration this problem is not verified.

The system identification tools considered in this work (CLLM, FM and ANN) were used to derive models for the NuHCC water canal system. The models were trained using experimental data collected in the canal where a two pools configuration is considered. The modeling results show a good performance of models derived from such data-driven modeling tools.

6. REFERENCES

Babuska, R. (1998). *Fuzzy Modeling for Control*. Kluwer Academic Publishers. USA.

Baume, J.-P., J. Sau and P.-O. Malaterre (1998). Modelling of irrigation channel dynamics for controller design. In: *Systems, Man, and Cybernetics, 1998. 1998 IEEE International Conference on*. Vol. 4. pp. 3856–3861.

Benjamin, T. B., J. L. Bona and J. J. Mahony (1972). Model equations for long waves in nonlinear dispersive systems. *Phil. Trans. Royal Society of London. Math. Phy. Sci.* **272**(1220), 47–78.

Borges, J., I. Tavares and M. Ayala Botto (2011). Modeling of a water canal system using a weigthed composition of local linear state-space models. *18th IFAC WC, Milano, Italy*.

Borges, J., V. Verdult, M. Verhaegen and M. Ayala Botto (2004). Separable least square for projected gradient identification of composite local linear state-space models. *16th MTNS*. Leuven, Belgium.

Coron, J. M., B. d'Andrea Novel and G. Bastin (1999). Lyapunov approach to control irrigation canals modeled by saint-venant's equations. *ECC 1999, Karlruhe, Germany*.

Haykin, S. (1999). *Neural Networks - A Comprehensive Foundation*. Prentice Hall Int.. USA.

Kosuth, P. (1994). Techniques de regulation automatique des systemes complexes : application aux systemes hydrauliques a surface libre. *Thése de Doctorat, Institut National Polytechnique de Toulouse - CEMAGREF - LAAS CNRS,*.

Liu, Fubo, Jan Feyen and Jean Berlamont (1995). Downstream control of multireach canal systems. *Journal of Irrigation and Drainage Engineering* **121**(2), 179–190.

Mandic, D. P. and J. A. Chambers (2001). *Recurrent Neural Networks for Prediction - learning algorithms, architectures and stability*. John Wiley & Sons, LTD. England.

Roubos, H. (2002). Bioprocess modeling and optimization, fed-batch clavulanic acid production by strepto-myces clavuligerus. PhD thesis. TUDelft. Delft, Netherlands.

Toudeft, A. (1995). Neural control of nonlinear non-minimum phase dynamical systems. *ICANNGA95*. Ales, France.

Verdult, V., L. Ljung and M. Verhaegen (2002). Identification of composite local linear state-space models using a projected gradient search. *Int. J. Cont.* **75**(16-17), 1385–1398.

Verhaegen, M. (1994). Identification of the deterministic part of mimo state space models given in innovations form from input-output data. *Automatica* **30**, 61–74.

Voron, B. and A. Bouillot (1997). Application of the fuzzy set theory to the control of a large canal. *Int Works. Regul. Irrig. Canals, RIC97* pp. pp. 317–331. Marrakech, Morocco.

Wagemaker, Robbert (2005). Model predictive control on irrigation canals, aplication of various internal models. Master's thesis. TUDelft. Delft, Netherlands.

Xu, C. Z. and G. Sallet (1998). Proportional and integral regulation of irrigation canal systems governed by the st venant equation. *IEEE SMC 1998*.