System Identification of Distributory Canals in the Indus Basin

Saad Abdul Aleem[1] **Abubakr Muhammad**[2] **Hasan Arshad Nasir**[3]

Department of Electrical Engineering, Lahore University of Management Sciences, Lahore, 54792 Pakistan (email: [1] aleems@seas.upenn.edu, [2] abubakr@lums.edu.pk, [3] hnasir@student.unimelb.edu.au)

Abstract: Water is becoming a scarce resource for many developing countries and its management has become an important issue. In this work, we derive channel models for regulation in irrigation canals using system identification techniques. We take inspiration from existing research on control of large-scale irrigation networks and perform experimental identification of distributory irrigation channels in Pakistan. System identification models are computationally inexpensive and accurate for simulating water profiles; thus qualifying for addressing efficient control design problems. The channel models are estimated and validated on operational data from Khaira Distributory off Main BRB Canal, Lahore. Our results show that system identification techniques capture the efficacy of canal models appropriately and will play a vital role in building accurate models for controlled irrigation canals in Pakistan.

Keywords: System Identification; Open Channel Flows; Controlled Irrigation Canals; Parameter Estimation

1. INTRODUCTION

According to Seckler et al. (1999), many developing countries are suffering from water scarcity due to the inefficiencies of their existing irrigation networks. Although water losses in irrigation networks are large, they can be substantially reduced by employing improved control design for irrigation systems (Weyer (2008)). Over the last decade, control of irrigation canals has attracted great interest from the research community (Cantoni et al. (2007); Ooi and Weyer (2007); Litrico and Fromion (2009); Negenborn et al. (2009)). With rapid advances in communication infrastructure and increased computational power, automatic control of irrigation channels will not only combat water scarcity but will also pave the way for efficient, demand-based, water delivery mechanisms. This is why, a large, efficient and well managed irrigation system is of immense importance for an agro-based economy like Pakistan.

To achieve efficient control design for irrigation canals, accurate models are required which capture essential dynamics of the system. These models can be derived either from *physical modelling* of water flow in open channels or from *system identification* experiments. In literature, Saint Venant equations have been traditionally used to describe open channel flows. These equations develop a mathematical model based on physical principles, involving mass and energy balances (Chaudhry (2008)). They model channel flows accurately and have been widely used in hydraulics engineering (Litrico and Fromion (2009)). Simulations from Saint Venant equations agree considerably with real data from open channels (Ki Ooi et al. (2005)), allowing them to serve as a starting point for simulating flow and water level profiles (called hydrographs) in irrigation canals (Ooi and Weyer (2007); Litrico et al. (2005); Nasir and Muhammad (2011)). However, Saint Venant equations are hyperbolic partial differential equations, i.e. they suffer from modelling complexity and lack parsimony; thus cannot be directly used for efficient control design. Additionally, in the context of control and prediction, Saint Venant equations are computationally intensive and, sometimes, infeasible to implement (Mareels et al. (2003)).

These constraints have led to development of simpler models for irrigation systems based on system identification techniques. These models are obtained directly from the observed data and have been widely used for control of irrigation networks because of their low computational overhead (see e.g. Weyer (2008)). In the presence of representative data, system identification processes can effectively capture the relevant dynamics of irrigation channels (Weyer (2008); Mareels et al. (2003)). Results of system identification for irrigation canals in Weyer (2001) and Eurén and Weyer (2007) confirm that these models are parsimonious, have low computational overhead and are fairly accurate for control design.

This paper attempts to address system identification aspects which play a crucial role in developing a controlled irrigation network. We take inspiration from existing lit-

$\star$ S. A. Aleem is presently with ESE Department at University of Pennsylvania, USA and H. A. Nasir is presently with EEE Department at University of Melbourne, Australia.

$\star\star$This work is co-sponsored by a partnership with International Water Management Institute (IWMI), a National Grassroots Research Initiative student grant by ICT R&D Fund, Pakistan; and via the logistical and technical support of PMIU, Punjab Irrigation Department, Pakistan.

erature on identifying system parameters and perform system identification of irrigation channels in Pakistan. We performed extensive experimentation on Khaira Distributory, which stems from Main BRB Canal, Lahore, to generate a representative dataset of *hydrographs*. In this paper, we present our complete system identification procedure, from experimental design to model validation, and we hope that our findings will help in building accurate models for controlled irrigation canals in Pakistan. These actual experiments are supplemented by the results of channel simulations which have allowed us to explore various effects on system response as a function of different physical parameters. The work reported in this paper runs in parallel to the efforts at LUMS to build a scalable hydrometry infrastructure for the Indus basin canal network (Ahmad et al. (2013)). In particular, the data recordings have been obtained using smart sensors developed by the authors' group.

The remainder of the paper is organized as follows. In Section 2, we present the modelling of open channels. In Section 3, we discuss our experimental results in identifying lumped parametric models of irrigation canals. The performance of the estimated model is also compared with the actual response. We conclude in Section 4.

2. OPEN CHANNEL MODELLING

In this section we present the modelling of open channels. We discuss both physical and data-based modelling.

2.1 Physical modelling: Saint Venant equations

In constructing physical models of an open channel, Saint Venant equations are used as a starting point for simulating flow and level profiles (Ooi and Weyer (2007); Litrico et al. (2005)). These equations develop mathematical models based on physical principles, involving mass and energy balances (Chaudhry (2008)), and are given by:

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = 0,$$
$$\frac{\partial Q}{\partial t} + \Big(\frac{gA}{B} - \frac{Q^2}{A^2}\Big)\frac{\partial A}{\partial x} + \frac{2Q}{A}\frac{\partial Q}{\partial x} + gAS_f - gAS_o = 0, \tag{1}$$

where x is the distance coordinate; t is time; A is the cross-sectional area of the channel; B is the width of water surface; Q is the flow (discharge); g is the gravitational constant (taken as $9.81m/s^2$); S_f and S_o are frictional and bed slope of the channel respectively. The frictional slope is modeled with the classical Manning formula i.e. $S_f = \frac{Q^2\eta^2}{A^2R^{\frac{4}{3}}}$ where, η is the Manning coefficient, R is the hydraulic radius, defined by $R = \frac{A}{P}$ with P as the wetted perimeter (Chaudhry (2008)). In this paper, we have employed Saint Venant equations in *simulating* data for testing our system identification technique. Saint Venant equations model the channel flows accurately (Ki Ooi et al. (2005)), if the channel geometry and operating conditions are known. The equations have been used extensively in hydraulic and civil engineering (Litrico and Fromion (2009)).

2.2 System identification based modelling

System identification is a data-based modelling technique, used to generate a viable transfer function between informative input and output data. Unlike Saint Venant equations, these irrigation models do not necessarily have a physical meaning. Methods in system identification provide validity for only certain input signals and around certain operating points. In order to use these techniques effectively, we must have informative input/output data.

Model structure and parameter estimation play an important part in system identification. We need to use some prior information about the physical process for modelling a mathematical relationship between representative input/output data. In our model, downstream water level is the controlled variable, and the gate position is the manipulated (input) variable. So, our model should have downstream water level as output and gate positions as input (Weyer (2001)).

In this procedure, we use fundamental physical insight about the channel and try to develop a representative parametric model of the system. The methodology in deriving the model structure for open channel pools has been explained in (Nasir and Muhammad (2011); Weyer (2001); Eurén and Weyer (2007)). To start with, consider a mass (or volume) balance equation

$$\frac{dV(t)}{dt} = Q_{in}(t) - Q_{out}(t), \tag{2}$$

where V is the volume in a pool, t is the time, Q_{in} is the inflow at the upstream gate and Q_{out} is the outflow from the downstream gate. For simplicity, we have assumed that there are no off-shoots along the length pool.

To develop a relationship between flow and water level, we need to study canal gates. Canal gates are essential hydraulic structures which play an important role in regulating the water flow in irrigation channels (Eurén and Weyer (2007)). Two such hydraulic structures, widely used in regulating the flow of irrigation canals, are *undershot gates* and *overshot gates* (as shown in Fig. 1). In the

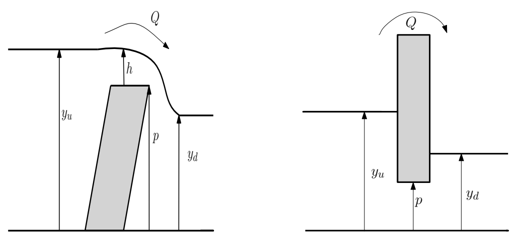

Fig. 1. Schematic side view of Overshot gate (left) in free flow and Undershot gate (right) in submerged flow. For an overshot gate, under free flow conditions, the flow over the gate is expressed as $Q \approx \theta h^{\frac{3}{2}}$ and for undershot gates, under submerged flow conditions, the flow is approximated by $Q \approx \theta p\sqrt{y_u - y_d}$; where Q - water flow, y_u - upstream water level, y_d - downstream water level, p - gate position and h - head over the gate.

literature related with hydraulics and flows (for example in Bos (1976)), the flow equations for both overshot and undershot gates can be found (see Fig. 1 for the flow relations of overshot and undershot gates). These flow expressions are, at best, approximations which model the flow well under *free flow* conditions, for overshot gates, and under *submerged flow* conditions, for undershot gates.

Substituting inflow and outflow relationships with respective gate models and assuming volume in the pool is proportional to the water level at the downstream (i.e. assuming uniform channel cross-section), we can write Eq.(2) as a non-linear differential equation in terms of downstream water level y_2. Additionally, we know that there is a time delay for the water at the upstream to reach downstream. Consider a case in which we have one overshot gate at the upstream and one overshot gate at the downstream (see Fig. 2). Assuming free flow conditions for the system, we can write Eq.(2) as

$$\frac{dy_2(t)}{dt} = \tilde{c_1} h_1^{\frac{3}{2}}(t - \tilde{\tau}) + \tilde{c_2}\Big(y_2(t) - p_2(t)\Big)^{\frac{3}{2}}, \qquad (3)$$

where $\tilde{\tau}$ is the time delay for the water at upstream gate to reach downstream water and $h_2(t)$ is substituted as $y_2(t) - p_2(t)$. In case of multiple hydraulic structures at the upstream/downstream, we will have to introduce indexing and summation to represent cumulative Q_{in} and Q_{out}.

In terms of implementation, we are interested in discretization of the continuous first order, non-linear, differential equation into a first order difference equation; yielding a discretized model for the open channel system. Using the Euler's approximations for the derivative term (as explained in Weyer (2001) and Eurén and Weyer (2007)), we obtain the following model structure

$$y_2[k+1] = y_2[k] + c_1 h_1^{\frac{3}{2}}[k - \tau] + c_2\Big(y_2[k] - p_2[k]\Big)^{\frac{3}{2}}, \qquad (4)$$

where c_1 (equivalent to $T\tilde{c_1}$)and c_2 (equivalent to $T\tilde{c_2}$) are the unknown system parameters and τ is the discretized time delay (which is obtained by dividing the actual time delay $\tilde{\tau}$ by the sampling time T). The system parameters $[c_1, c_2]^T$ are estimated using the least squares, where discretized time delay, τ, was approximated to the nearest integer to facilitate algorithm implementation.

3. SIMULATIONS & EXPERIMENTAL DESIGN

In this section we apply the estimation techniques in identifying the lumped parametric models of irrigation canals, using data from simulated channel flows, as well as from actual canal experimentation at Khaira Distributory of Main BRB Canal, Lahore.

3.1 Open Channel Simulator

We used an open channel simulator (outlined in Nasir and Muhammad (2011); Tariq et al. (2012)) to simulate water profile in an open pool, assuming uniform rectangular channel, subject to free-flow boundary conditions. In our simulation, we modelled overshot gates at the upstream and the downstream end (see Fig. 2) for various pools.

Channel Model: Based on previous discussion in Section III, we can develop the simplified, parametric, model for the open channel. The model is given by

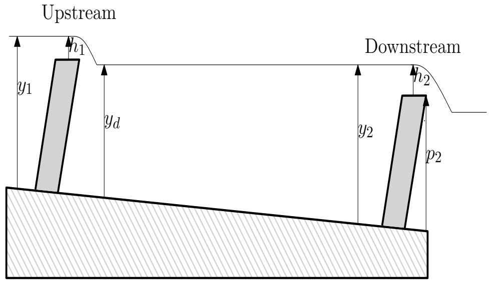

Fig. 2. Side-view sketch of a pool in open channel simulator with overshot gates at upstream and downstream.

$$y_2[k+1] = y_2[k] + \theta_1 h_1^{\frac{3}{2}}[k - \tau_k] + \theta_2\Big(y_2[k] - p_2[k]\Big)^{\frac{3}{2}}, \qquad (5)$$

where y_2 is the downstream water level, h_1 is the head over the upstream gate, τ_k is the discretized time delay, p_2 is the downstream gate position and θ_1 and θ_2 are the unknown parameters. To express this in *linear regression* form, we assume $\phi[k+1] = \Big[h_1^{\frac{3}{2}}[k - \tau_k], \Big(y_d[k] - p_2[k]\Big)^{\frac{3}{2}}\Big]^T$, $\theta = [\theta_1, \theta_2]^T$ and $y[k+1] = y_2[k+1] - y_2[k]$. Then we can express the system model in the form of linear regression as, $y[k] = \phi^T[k]\theta$.
In our simulations, we used *step test* (see Fig. 3) to determine the time delay τ_k for the system model (Weyer (2001), Nasir and Muhammad (2011)). Values of τ_k for different pools have been tabulated in Table 1.

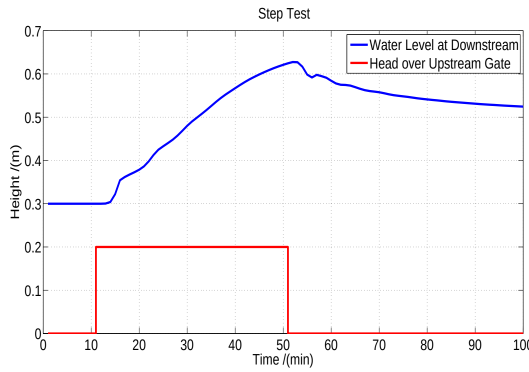

Fig. 3. Step Test for Open Channel.

Simulation Results & Parameter Estimation: We used our simulator, based on Saint Venant equations, on some test input signals to generate representative trajectories of an open channel. To excite different modes in the pool, head over the upstream gate and downstream gate position was randomly changed for a simulation time of 100 min. In addition to simulation criteria, a rectangular channel of uniform bedwidth of 5m was modelled and the canal was 'discretized' into 20 equally spaced measuring units to monitor water level at different locations in the pool. Afterwards we used our system identification techniques to extract system parameters of the simplified model for various pools, which has been tabulated in Table 1. In Fig. 4, one such hydrograph (at 500m downstream) and corresponding estimated water level is shown.

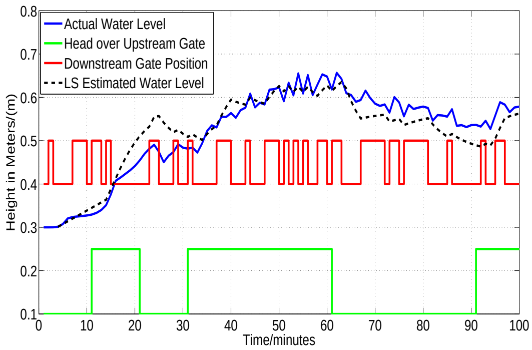

Fig. 4. Simulated & Estimated Response for Downstream Water Level.

Table 1. Estimation Results for Open Channel Simulator

Pool	1	2	3
Length (m)	360	500	800
τ_k (sec)	120	180	300
θ_1	0.3529	0.1893	0.0994
θ_2	-0.6381	-0.3312	-0.3032
Avg. Sq. Error (m^2)	0.003	0.0012	0.0011

3.2 Canal Experimentation: Khaira Distributory

We carried out extensive experiments and collected data points from Khaira distributory. It stems from Main BRB Canal and is approximately 18 km long, with channel width of 10 ft ($\approx$ 3.05m) and height 4 ft ($\approx$ 1.22m). In these experiments, we only had control over the upstream gate (which was undershot) and there was no physical (hydraulic) structure present at the downstream ends of the canal, which happened to be points of measurement. We modelled it as an *always open*, hypothetical, overshot gate (see Fig. 5).

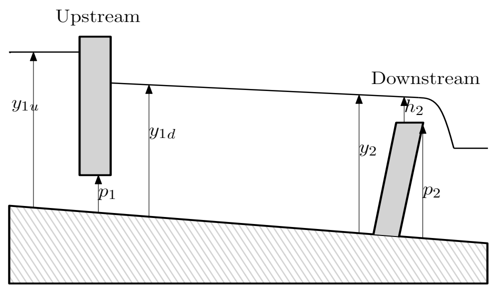

Fig. 5. Side-view sketch of Khaira Distributory Pool with undershot gate at upstream and 'hypothetical' overshot gate downstream (point of measurement). The value of p_2 is always taken to b 0 to model an *always open* overshot gate.

Channel Model: With an undershot gate at the upstream end and an overshot gate at the downstream the model of Khaira Distributory Pool is given by

$$y_2[k+1] = y_2[k] + \theta_1 p_1[k-\tau_k]\left[\sqrt{y_u - y_d[k-\tau_k]}\right] + \theta_2\Big(y_2[k] - p_2[k]\Big)^{\frac{3}{2}}, \tag{6}$$

where y_2 is the downstream water level, p_1 is the opening of the undershot gate at the upstream and p_2 is the position of, '*hypothetical*', overshot gate at the downstream. In addition, y_u is assumed to be constant (because distributory is drawing water from a large water body) and p_2 is the always taken to 0, to model an '*always open*' hypothetical gate. Using similar substitutions (as in the previous subsection), the system model was expressed in the form of linear regression to perform system identification.

Fig. 6. Upstream Gate at Khaira Distributory.

Experimental Procedure: Different phases of our canal experimentation can be summarized in the following steps.

- Sensors are placed at appropriate sites (usually at downstream bridges) along the length of canal (see Fig. 7). They record the height of water in the channel and communicate, through GSM module, to a cell phone, or through GPRS to a server.
- At the Upstream, gate is closed and then opened to model step input, while monitoring the gate position.
- The readings are recorded and in conjunction with the input are used to perform system identification.

Fig. 7. Experimental Procedure: Sensor Deployment.

We kept the sampling time for our experiment at 10 seconds. As our readings were transmitted through GSM module, it introduced non uniform delays between successive messages. To cater for this discrepancy, we used *data interpolation* to obtain a uniformly sampled data. The data was adjusted with a uniform sample time of 10 seconds (as shown in Fig. 8). It can be verified that this interpolation technique preserves the shape of the corresponding hydrograph. This pre-processing was necessary to generate a uniform data for linear regression.

Parameter Estimation: In our experiment we closed the upstream gate and then re-opened it. The water level was measured at 50m, 300m and 550m downstream lengths, after every 10s (see Fig. 9). For channel modelling, we used the data from 300m and 550m sensors. The corresponding data was processed and interpolated. For estimation process, the linear regression model was attempted to fit the observed response. In our estimation, y_u was assumed to be constant and $p_2[k]$ was taken to be zero to model

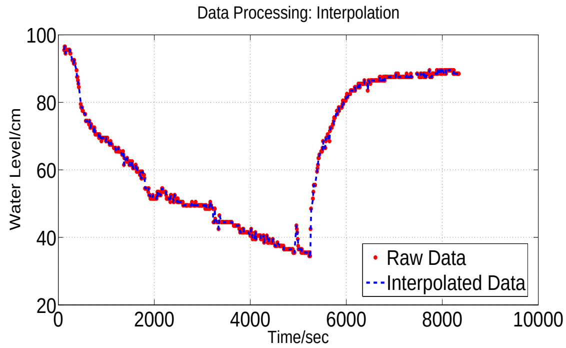

Fig. 8. Data Interpolation of Measurements.

an *'always opened'*, hypothetical, overshot gate. In addition, y_d was measured near the opening of the upstream gate. The response delay, τ_k, was inspected from the raw data. In order to obtain a better estimate of the time constant for the pool, the upstream gate was opened and closed suddenly (as if to model a sudden impulse). This accounts for the apparent 'jump' in the gate input and the corresponding 'blip' in hydrographs. The results of our experiments, along with the estimated parameters, have been tabulated in Table 2. In Fig. 10, the response of 550m sensor along with respective upstream gate input and estimated water level is shown.

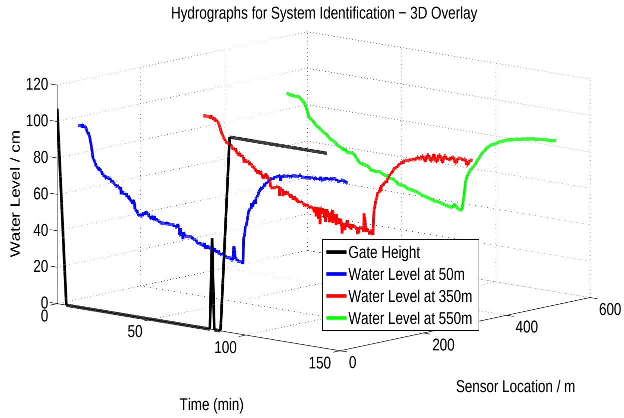

Fig. 9. Experimentation: 3D Overlay of Hydrographs.

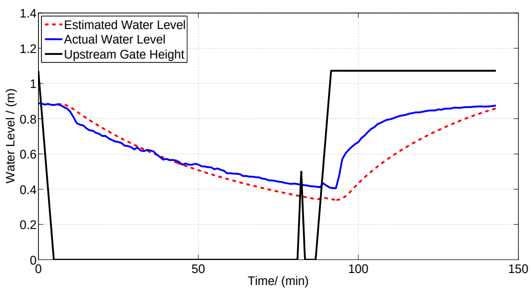

Fig. 10. Actual & Estimated Response for Khaira Distributory.

Our estimated parameters (for both the simulated and actual experimentation) make intuitive sense from a physical point of view. θ_1 is positive as it is related with the inflow of the water (from upstream), while θ_2 is negative as it is associated with the outflow of water. For canal experimentation, the difference in order of magnitude, for estimated

Table 2. Experimental & Estimation Results for Khaira Distributory

Pool	1	2
Length (m)	300	550
$\tilde{\tau}$ (sec)	200	350
θ_1	0.0161 x 10^{-3}	0.0152 x 10^{-3}
θ_2	-0.3271 x 10^{-3}	-0.2737 x 10^{-3}
Avg. Sq. Error (m^2)	0.0164	0.0155

inflow and outflow parameters, can be attributed to the absence of hydraulic structure at the downstream.

3.3 Model Validation

We performed model verification of our system identification for both the channel simulations and canal experimentation. For our open channel simulator, we stored the estimated parameters and generated another water level profile through some other randomly generated side information (head over the upstream gate and downstream gate position). The result of model verification for the open channel simulator are shown in Fig. 11. From the superposition of estimated profile and actual hydrograph, we can infer that under steady states, our model tracks the actual water level relatively well. It is indicating the presence of some distributed transient modes in the system (which the hypothesized model does not take into account in its simplified lumped construction) that eventually diminish under steady boundary conditions.

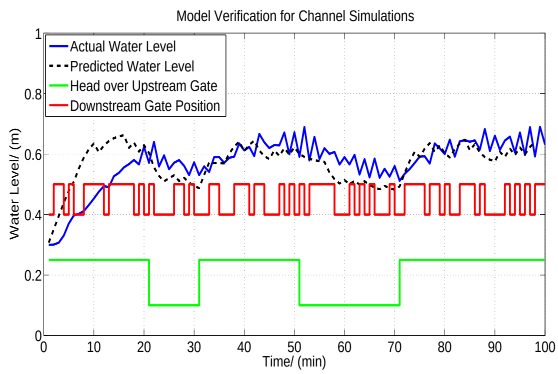

Fig. 11. Model Verification for Channel Simulations.

Similarly, for our actual experimentation, we carried out another set of experiments at the canal (with different set of input data). For the new side information, we compared the response of the estimated water level to that of the actual water level. This is shown in Fig. 12. Again, it can be seen that when we re-open the gate, there is some poor tracking exhibited by the estimated model. However, it again begins to track the actual level after some time.

We also used average squared prediction error, given by $\frac{1}{N}\sum_{i=1}^{N}\epsilon^2(i,\hat{\theta})$ where $\epsilon(i,\hat{\theta}) = y(i) - \hat{y}(i,\hat{\theta})$, as a performance metric for model validity. $\hat{\theta}$ is the vector of estimated parameters and $\hat{y}(i,\hat{\theta})$ is the predicted value given $\hat{\theta}$, obtained from an ARX type predictor (Ljung (1999)) defined from the model structure in Eq.(4). The results of our model verification have been tabulated in Table 3 and Table 4

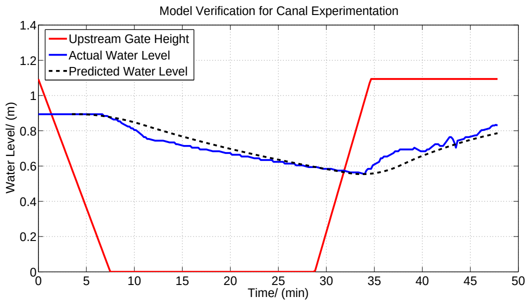

Fig. 12. Model Verification for Canal Experimentations.

Table 3. Errors in Channel Simulation

	Sim 1	Sim 2	Sim 3
Length (m)	360	500	800
Prediction Error(m^2)	0.0094	0.0054	0.0087

Table 4. Errors in Canal Experimentation

	Pool 1	Pool 2
Length (m)	300	550
Prediction Error(m^2)	0.0077	0.0014

It is important to notice that the choice of data for system identification is very important. One of the drawbacks of system identification is that (at times) it can become very specific to only certain kinds of input signals (thus undermining generalization). For this purpose, we excite the channel with many modal inputs to extract the most out of the distributed nature of the system and capture these behaviours as a lumped parameter in our model.

4. CONCLUSION

In this paper, steps have been taken towards realizing controlled irrigation systems in Pakistan. We have presented complete system identification procedure, from experimental design to model validation, for actual irrigation canals. It was found that a first-order nonlinear model, derived from prior knowledge about the physical system, gave accurate results in predicting downstream water level of an actual irrigation canal. In this study we have verified;

- Design of an accurate canal model can be obtained through elaborate experimentation, which captures dynamics of the underlying physical system.
- The proposed methodology of system identification is suitable for introducing controlled irrigation systems in developing countries.

Moreover, the estimated models are simple which makes them suitable for control design. These results illustrate that system identification, for control in irrigation canal, can play an important part in managing water resources for developing countries.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Asad Khalid, Hassan Zaheer and Sarmad Jabbar for assisting in field work. We express our sincere appreciation for Dr. Erik Weyer's feedback on our experimentation results. We thank Mr. Zahoor Ahmed for development of sensors used in our study. Finally, we are thankful Mr. Habibullah Bodhla from Punjab irrigation for granting us access to Khaira distributory.

REFERENCES

Ahmad, Z., Asad, E.U., Muhammad, A., Ahmad, W., and Anwar, A. (2013). Development of a low-power smart water meter for discharges in indus basin irrigation networks. In *Wireless Sensor Networks for Developing Countries*, 1–13. Springer.

Bos, M. (1976). Discharge measurement structures. *Publications International Inst. for Land Reclamation and Improvement.*

Cantoni, M., Weyer, E., Li, Y., Ooi, S.K., Mareels, I., and Ryan, M. (2007). Control of large-scale irrigation networks. *Proceedings of the IEEE*, 95(1), 75–91.

Chaudhry, M. (2008). *Open-Channel Flow.* Springer.

Eurén, K. and Weyer, E. (2007). System identification of open water channels with undershot and overshot gates. *Control Engineering Practice*, 15(7), 813–824.

Ki Ooi, S., Krutzen, M., and Weyer, E. (2005). On physical and data driven modelling of irrigation channels. *Control Engineering Practice*, 13(4), 461–471.

Litrico, X. and Fromion, V. (2009). *Modeling and control of hydrosystems.* Springer.

Litrico, X., Fromion, V., Baume, J.P., Arranja, C., and Rijo, M. (2005). Experimental validation of a methodology to control irrigation canals based on saint-venant equations. *Control Engineering Practice*, 13(11), 1425–1437.

Ljung, L. (1999). *System identification.* Wiley Online Library.

Mareels, I., Weyer, E., and Ooi, S.K. (2003). Irrigation networks: A systems engineering approach. In *Control and Automation, 2003. ICCA'03. Proceedings. 4th International Conference on*, 15–28. IEEE.

Nasir, H. and Muhammad, A. (2011). Control of very-large scale irrigation networks: A cps approach in a developing world setting. In *World Congress IFAC*, volume 18, 10739–10745.

Negenborn, R.R., van Overloop, P.J., Keviczky, T., and De Schutter, B. (2009). Distributed model predictive control of irrigation canals. *NHM*, 4(2), 359–380.

Ooi, S.K. and Weyer, E. (2007). Control design for an irrigation channel from physical data. In *Control Applications, 2007. CCA 2007. IEEE International Conference on*, 270–275. IEEE.

Seckler, D., Barker, R., and Amarasinghe, U. (1999). Water scarcity in the twenty-first century. *International Journal of Water Resources Development*, 15(1-2), 29–42.

Tariq, M.U., Nasir, H.A., Muhammad, A., and Wolf, M. (2012). Model-driven performance analysis of large scale irrigation networks. In *Cyber-Physical Systems (ICCPS), 2012 IEEE/ACM Third International Conference on*, 151–160. IEEE.

Weyer, E. (2001). System identification of an open water channel. *Control engineering practice*, 9(12), 1289–1299.

Weyer, E. (2008). Control of irrigation channels. *Control Systems Technology, IEEE Transactions on*, 16(4), 664–675.