

System Identification of Irrigation Channels with Overshot and Undershot Gates

Systemidentifiering av bevattningskanaler
med olika typer av luckor

Karin Eurén

Abstract

System Identification of Irrigation Channels with Overshot and Undershot gates

Karin Eurén, Department of Information Technology

Water resources in Australia are limited. For a farmer the access to water is crucial and due to the dry climate the farmers in Australia can not rely on precipitation. Irrigation is therefore a very important part of the farming industry. The Coleambally Irrigation Area is situated in the southern parts of New South Wales close to the border of Victoria. The Irrigation Network often supplies the irrigation channels with too much water to be sure that the demand of water is satisfied. Due to this over supply a great amount of water gets wasted. Design of a better control system would be able to reduce the water wastage.

A mathematical model describing the dynamics of the irrigation system can be used as a tool for the control system design. The aim of this project was to build a mathematical model with the system identification approach. The model should be able to describe the downstream water level of a single pool of an irrigation channel which has both undershot and overshot gates. A model was built by estimating unknown parameters of a chosen model structure from a set of experimental data. The data was collected from an experiment performed on the real irrigation system in Coleambally.

The result of the system identification was a first order output error grey box model. The model performs well on validation data and may therefore be used for design of a more efficient control system. The model gave such good results that it additionally may be used for various simulation purposes.

Keyword: System Identification, Irrigation Channels, Grey Box modelling, Mathematical Modelling

Referat

Systemidentifiering av bevattningskanaler med olika typer av luckor

Karin Eurén, Institutionen för informationsteknologi, Systemteknik

I Australien är vattenresurserna begränsade. För lantbrukare är tillgängligheten på vatten mycket viktig. På grund av det torra klimatet kan inte de Australiensiska bönderna förlita sig på nederbörden. Bevattningssystemen är därför en viktig del i jordbruksindustrin. Bevattningsområdet i Coleambally ligger i södra New South Wales nära gränsen till staten Victoria. Bevattningsnätet i Coleambally förser ofta bevattningskanalerna med för mycket vatten för att vara säker på att lantbrukarna får den mängd vatten de behöver. På grund av denna tillförsel av överskottsvatten går stor mängd av vatten förlorad. Design av ett bättre reglersystem skulle kunna minska den stora förlusten av vatten.

En matematisk modell beskrivande dynamiken av bevattningssystemet är ett bra redskap vid en design av ett bättre reglersystem. Syftet med det här projektet var att genom systemidentifiering bygga en matematisk modell av bevattningssystemet. Modellen syftade till att beskriva vattennivån i en sträcka av bevattningskanalerna, sträckan i kanalen skulle ha två olika typer av luckor, en typ där vattnet strömmar över luckan och en annan typ där vattnet strömmar under luckan. En modell byggdes genom att parametrar från en vald modellstruktur estimerades från experimentella data. Data samlades under ett experiment som utfördes på en bevattningskanal i Coleambally.

Resultatet från systemidentifieringen blev en första ordningens output error grey box modell. Modellen visar goda resultat vid validering och bör kunna användas vid design av ett bättre reglersystem. Modellen visar så god överensstämmelse med valideringsdata att den även kan användas för olika fall av simulering.

Nyckelord: System identifiering, Bevattningskanaler, Grey box modellering, matematisk modellering

PREFACE

This report is the result of a Master thesis done at the University of Melbourne, Australia. The thesis is part of an on going research project about optimization of irrigation systems by signal processing and control. The project is run by Dr. Erik Weyer at the CSSIP group, Centre for Sensor Signal and Information Processing in the Department of Electrical and Electronic Engineering.

Supervisor: Dr. Erik Weyer, Department of Electrical and Electronic Engineering, Melbourne University

Examiner: Bengt Carlsson, Department of Information Technology, Uppsala University

Copyright © Karin Eurén

Department of Information Technology, Uppsala University.
Printed at Department of Earth Science, Uppsala University, Uppsala 2004
UPTEC W 04 029, ISSN 1401-5765

TABLE OF CONTENT

1 INTRODUCTION.....	1
1.1 PROJECT DESCRIPTION.....	1
1.1.1 <i>Control of the water flow in the CIA</i>	2
1.1.2 <i>The use of mathematical models in control design</i>	2
1.1.3 <i>The aim of the project</i>	2
1.2 OUTLINE OF THE REPORT	3
2 THEORY	4
2.1 PHYSICAL MODELLING	4
2.2 SYSTEM IDENTIFICATION.....	4
2.2.1 <i>Black box modelling</i>	4
2.2.2 <i>Grey box modelling</i>	4
2.3 THE SYSTEM IDENTIFICATION PROCEDURE.....	4
2.3.1 <i>Prior knowledge</i>	5
2.3.2 <i>Experiment design</i>	5
2.3.3 <i>Design of input signal and data collection</i>	6
2.3.4 <i>Model structure selection</i>	8
2.3.5 <i>Choice of model order</i>	10
2.3.6 <i>Grey box model structure</i>	10
2.3.7 <i>Parameter estimation</i>	10
2.3.8 <i>Model validation</i>	12
2.3.9 <i>Model validity</i>	12
3 INFORMATION ABOUT THE IRRIGATION SYSTEM	13
3.1 SYSTEM DESCRIPTION.....	13
3.1.1 <i>History of the Area</i>	13
3.1.2 <i>Site description</i>	13
3.1.3 <i>Pool description</i>	14
3.1.4 <i>Earlier studies</i>	14
3.2 PHYSICAL KNOWLEDGE	14
3.2.1 <i>St Venant equations</i>	15
3.2.2 <i>Integrator</i>	16
3.2.3 <i>The flow through the undershot and overshot gates</i>	16
4 SYSTEM IDENTIFICATION PROCEDURE APPLIED ON THE CIA	19
4.1 THE EXPERIMENT	19
4.1.1 <i>Limitations of the experiment</i>	19
4.1.2 <i>The step test</i>	20
4.1.3 <i>The results from the step test</i>	20
4.1.4 <i>The Binary Signal</i>	22
4.1.5 <i>The result from the experiment</i>	22
4.1.6 <i>Discussions of the experiment results</i>	28
4.1.7 <i>Data processing</i>	28
4.2 MODEL STRUCTURE SELECTION	29
4.2.1 <i>The OE Simulation model of the downstream water level</i>	30
4.3 PARAMETER ESTIMATION AND MODEL VALIDATION	31
5 RESULTS	33
5.1 ESTIMATION AND VALIDATION RESULTS IN FIGURES	33
5.1.1 <i>Estimated parameters Horticultural – Coly 3</i>	33
5.1.2 <i>Estimated parameters Morundah – Grants</i>	34
5.2 RESULTS IN PLOTS.....	34
5.2.1 <i>Plots of Horticultural – Coly 3</i>	34
5.2.2 <i>Plots of Morundah – Grants</i>	35
5.2.3 <i>Validation of the HC models on a step test</i>	35
5.3 TIME DELAYS OF THE MODEL	35
6 DISCUSSION	41

6.1 DISCUSSION OF THE H-C MODEL RESULTS.....	41
6.1.1 <i>The separation of the two H-C models</i>	41
6.2 DISCUSSION OF THE M-G MODEL RESULTS.....	42
6.3 ADDED KNOWLEDGE.....	42
6.4 SUGGESTIONS FOR FUTURE WORK.....	43
7 CONCLUSIONS	44
8 ACKNOWLEDGEMENTS.....	45
9 REFERENCES.....	46
9.1 LITERATURE REFERENCES.....	46
9.2 INTERNET REFERENCES.....	46
9.3 PERSONAL COMMUNICATION	47

1 INTRODUCTION

Australia is the driest continent in the world, but yet it has one of the highest rates of water use per capita. This high rate is not sustainable for the future of the environment and the future welfare of Australia. Since the country now and then gets hit by the El Niño, it can suffer several years of droughts after each other. The time and length of the El Niño are difficult to predict. So it is very important that Australia manage its fresh water in the best possible way and save water for the future. At this point this is not the case (website 1, 2004).

Even though Australia is so dry, it has a warm climate that is propitious for growing crops. For a farmer access to water is crucial, and for obvious reasons the farmers in Australia can not only rely on precipitation. In Australia the farming industry is very big, so all together the farmers have a high demand of water. In fact of all the water that is used in the country 70-75% is used for irrigation (website 2, 2004).

It can lead to severe economic consequences for the farmers if water is not delivered on demand. To be sure that the demand is satisfied, the irrigation channels often operate with more water than necessary. The water losses due to this are approximately 15% (website 3, 2004). Since there is so much water used for irrigation, huge savings can be made if irrigation systems could be operated more restrictive. Not only the farmers but also the whole country would benefit from the water savings that could be made. The environmental health is of course the most important thing in this discussion, but the incitement to act is as in any other industry the economical issue. Water is expensive so waste of water is waste of money.

In many of the irrigation systems the control of water losses is very poor. An improvement of the control system can drastically reduce the losses. An earlier study of irrigation channels in Australia involving optimization and control, showed that, with a good control system installed, water losses in the irrigation channels could be reduced by 50 percent. The cost of the engineering infrastructure was paid back in less than a year with the savings made due to improved control (Mareels, 2003). Better knowledge about the dynamics of the irrigation channels is a helpful tool in the development of a better control system. Another study was made on the HMC, Haughton Main Channel, in Queensland. It involved building a mathematical model describing the dynamics of a pool in the irrigation system. The model was built by system identification methods and gave very good accuracy. Based on this model a much more sophisticated control system was built.

1.1 PROJECT DESCRIPTION

This project is a study of the Coleambally irrigation system in New South Wales, Australia. The CIA, Coleambally Irrigation Area, is very big, it supplies water to 452 farms and covers about 79 000 ha. It is situated 650 km south west of Sydney, not far from the border New South Wales - Victoria (website 4, 2004). In this irrigation system the problem with water losses is significant. The structures of the irrigation channels are similar to those in the HMC in Queensland. The basic idea was that, because the method of building a good control system with system identification methods gave such good results in the HMC, it could also work for the CIA. In the HMC the water flow is controlled by overshoot gates (the water flows over the gates). In the CIA the water flow is controlled by both overshoot and undershot gates (the water flows under the gates). The overshoot and undershot gates can be compared with gates that in literature are commonly called sharp crested weir and sluice gate, see Figure 1.1.

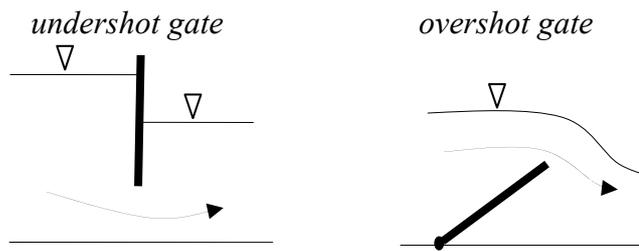


Figure 1.1 Schematic pictures of side view of an undershot gate and an overshot gate, respectively. The arrows represent the water flow.

1.1.1 Control of the water flow in the CIA

To be able to deliver the water demanded by the farmers the water levels are controlled to a certain set point. The gates control the water level immediate upstream the gates, i.e. the control is a simple feedback loop from upstream water level to gate position. This type of control makes the irrigation network loose huge amounts of water. The system does not consider what happens upstream or downstream a gate site. For example, if the water level at one gate site is too high and at the upstream gate site the water level is too low, water will be lost in the downstream pool. The pool will not be filled up again very fast because the upstream pool wants keep its water. To be able to keep all the set points the network has to be manually supervised. A better way of controlling the irrigation network would be to use distant downstream control, where a gate site controls the upstream water level of the downstream gate site. The water losses would be less with this type of control design, for the example above, a chain reaction upstream the irrigation channel would occur. And all the set points can eventually be satisfied. Mathematical models of the irrigation channels can optimize the control even better. The movements of the gates would consider the property of the pool dynamics and earlier gate movements both upstream and downstream in the pool.

1.1.2 The use of mathematical models in control design

In control theory mathematical models are used as tools for designing control systems. The model gives information about properties of the system such as step responses, time delays, time constants, corner frequencies etc. They are important things that need to be considered when designing control systems. Another benefit of using models is that when the control system is designed it can be tested and tuned on the model instead of the actual system. The last step can be very helpful, in many systems it can be expensive to test and tune the control parameters on the actual system and sometimes that could even be dangerous.

1.1.3 The aim of the project

The aim for this project was to build a mathematical model with system identification methods. The model should be able to describe the water level at the downstream end of a channel reach in the CIA. The channel reaches of interest is operated with both undershot and overshot gates. The project was aimed to finding a good model structure which could describe the flow through the undershot gate. (The physics behind the overshot gate was already roughly known and had in the HMC study found to give good results.) The model quality should at least be so high that it would be suitable as a tool for design of a good control

system, and hopefully also suitable for simulation purposes. A simulation model of the system should be able to simulate a channel reach with only a starting value of the water levels inside the pool and the gate positions for the simulation period.

1.2 OUTLINE OF THE REPORT

In Chapter two a guide through the theory of system identification is given. Different ways of building models will start the Chapter. In Section 2.2 system identification will be described briefly so the reader will get an idea of what system identification is about. Then in Section 2.3 the system identification procedure will be described in more detail.

In Chapter three a description of the studied system is given that is studied. The site with the Coleambally Main Channel will be described in Section 3.1.1-3.1.2 and in Section 3.1.3 the design of a single pool. A short guide of the earlier study at the HMC is given in Section 3.1.4. In Section 3.2 the physical knowledge of the system is found.

The system identification routine for this project is described in Chapter 4, including experiment, model structure selection, parameter estimation and model validation.

The result, discussion and conclusion are found in Chapter 5, 6 and 7 respectively.

2 THEORY

2.1 PHYSICAL MODELLING

Building a mathematical model of a system can be a very difficult and time consuming task. A very straightforward method, the most logical one, is to use the knowledge of physical laws that is acting on the system. See Section 3.2.1 for an example of physical modelling. This can be very complicated with iterative numerical solutions in both space and time.

2.2 SYSTEM IDENTIFICATION

System identification is a model building approach that does not interests in what actually happens within the system. The system identification engineer sees the system from the outside. He/She notice what is put into the system and what comes out from it. With these observations he/she moulds his/her model from an already existing model structure. A simple example of this technique is to adjust a straight line to a set of data. In this case the model structure is the common straight line expression,

$$y(x) = kx + m \tag{2.1}$$

The unknown parameters in equation (2.1) are k , and m . The task of the system identification engineer is to estimate the unknown parameters of the chosen model structure, in this case k and m . It is called system identification because the engineer is identifying the system dynamics from observing the system behaviour in the way mentioned above.

2.2.1 Black box modelling

Within system identification the demand on information about the system is small, which is a big advantage. The information needed about the system is the outputs interesting for modelling and the inputs affecting the outputs. Further, a set of empirical data from the actual system is needed. The data should describe what happens with the system under different circumstances. The data is the only information given from the system. In cases where the system has more complex dynamics this method could be more efficient. Black box modelling is the descriptive name for this approach. The advantage of this approach is that the model is concentrating on the essential part of the system and does not bother with the complicated mathematics behind the process. Some people call it the engineering approach. The disadvantage is that the model is limited by the quality of the data given by the system.

2.2.2 Grey box modelling

For some systems, only parts of the dynamics are known and sometimes only parts of the dynamics knowledge are worth including in the model. In this case the two methods described above can be combined. The name for this type of model is called grey box. Grey box modelling is the method used in this project.

2.3 THE SYSTEM IDENTIFICATION PROCEDURE

System Identification is a straight forward procedure with a few important steps, see Figure 2.1. To be able to create a model we need data, as said in the Section above, if data not already exist an *experiment* on the system to obtain data is the first step in the procedure, the next step is to *choose a model structure* suitable for the particular system. In both of these steps it is very helpful to use *prior information* about the system. Whether the system has a quick respond or not from input to output could be an example of prior information. When the data set is obtained and a model structure is chosen the *parameter estimation* can begin, the

parameters obtained will give the best model within the chosen model structure. The last step is the *model validation*, this step is to check if the model can reproduce the system behaviour and thus if the model is good enough for the purpose of the model. If the model satisfy the given model criteria the system identification is finished. If it does not satisfy them the next step is to go back in the procedure and choose another model order or model structure or, in the worst case, do another experiment if it shows that the data is not informative enough. This procedure is followed until a satisfying model is obtained. For an overview of the system identification procedure see Figure 2.1 (Söderström, Stoica, 1989).

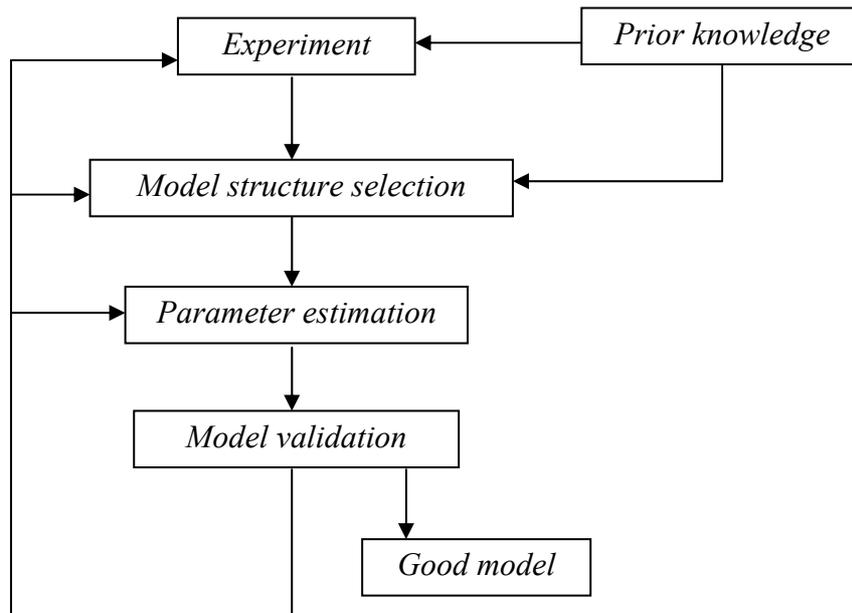


Figure 2.1 The system identification procedure.

2.3.1 Prior knowledge

It is important to use the prior information about the system so the system identification can be carried out in the optimal way. Prior information can be used when designing the experiment. For example, if the gain of the system is known it is easy to get an idea of how much the inputs needs to be changed to get a reasonable change in the output. Prior information can also give a hint on which model structure and model order to use. Time and money will not be wasted on non suitable model structures. Prior knowledge can help out in the analyzing of the result, by giving an idea on what to expect of the results.

2.3.2 Experiment design

The data is the weakest link for the system identification model building. The model is based on the experimental data and can not predict anything outside the conditions of the collected data. Informative data is therefore the key issue when designing an experiment.

The basic idea of the experiment is to attain a change in the output by making changes in the inputs. The output is the variable subject for modelling. Prior information is a useful tool to find out which variables that are inputs to the system, the needed changes in magnitude and frequency of the inputs. In a lot of systems there are many different inputs, but to build a

simple model as possible it is important to know which of the inputs that the output is more sensitive to (Ljung, 1999). As a complement to the prior information, or if there is no information at all it is very common to first perform a step test on the system. The step test will give an idea about the system dynamics, and thus how to design the experiment.

A common way to generate the inputs is to create Binary Signals (BS). The inputs changes between levels with different frequencies. See Figure 2.2.



Figure 2.2 An example of a binary signal

The model is validated on a separate set of data, the validation set. The model is thus only valid for conditions that were present during the data collection of the validation set. If control system design is the purpose of the model a wise choice is to perform the experiment during conditions that are common for the system (Ljung, 1999).

2.3.3 Design of input signal and data collection

The plot in Figure 2.3 gives an example of a bode plot of a system. To capture the dynamics of this system the inputs needs to be excited in the frequencies where the system has bends in the bode plot (Ljung, 1999). What is important is to see how the system responds for these frequencies. For example if the system does not excite the frequencies around ω in Figure 2.3, it is not possible to detect the resonance peak in the bode plot. Thus a very important part of the system properties will be lost. Further it is often important to excite the system with frequencies higher than the corner frequency ω_c . This is to be able to identify the slope in the bode plot. To exactly know which frequencies the system has to be excited with and what levels the inputs needs to jump between is difficult to know. This is often figured out by prior knowledge and experience, but also from step test applied to the system (Ljung, Glad, 1991). An idea of the time constant and ω_c can be given from a step test.

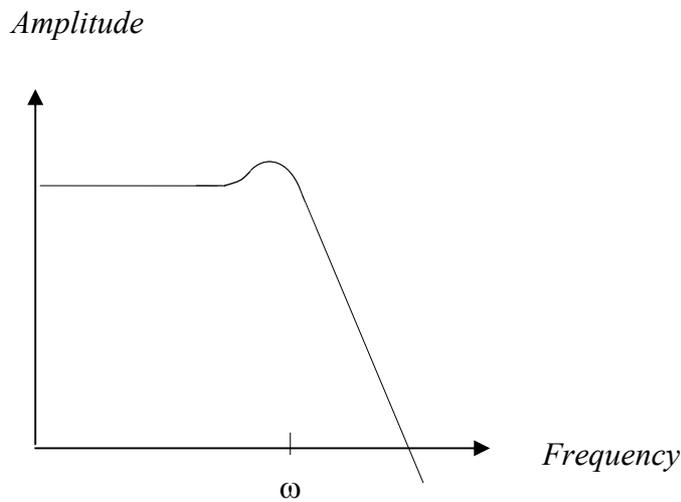


Figure 2.3 An example of a bode plot for a system.

Assume that the true system can be modelled by a certain model structure then the variance of the model becomes smaller if the estimation data set is increasing. It can be compared with a normal average. For example, the more data that is included in the search for the average age in a population the more precise the calculation of the average age will be. So as a conclusion for this Section the more data that can be collected during the experiment and the more variation in the data the better for the building of the model. It needs to be said that if the model has a bias, the bias will not get smaller the more data that is used in the estimation. Another thing that needs to be considered when collecting data is the sampling interval. In this project there is not a problem to sample fast enough. But in other situations it is important to know how fast the sampling rate needs to be so no information gets lost in the sampling. System properties with frequencies higher than the Nyquist frequency (half the sampling frequency) are not possible to detect. If the sampling rate is too slow important system properties might get lost in the sampling. The model estimation may be sensitive to noise if the data is sampled too frequent. So it is important to decide a sampling frequency that is suited for the system in question. A rule of thumb is to sample with a rate that is around 10 times the bandwidth of the system (Ljung, 1999).

For the experiment it is important to consider the limitations of the system. In theory it would be best to be able to design the experiment freely without any limitations at all, the more information about the system the better. But in reality it may not be allowed or not even possible to change the inputs as much as one could wish for. In some systems it could be very expensive and/or even dangerous to disturb the system, so often there is a limit on how big the changes can be.

2.3.4 Model structure selection

If a system is linear or can be linearised around a particular working point it can be modelled with various linear model structures. There are several common model structures that are used. They are all versions of the equation (2.2) below.

$$y(t) = G(q, \theta) * u(t) + H(q, \theta) * e(t) \quad (2.2)$$

where y is the output, u is the inputs and e is the white noise. G and H are the transfer functions. For system identification purposes it is practical to use a discrete time model since the sampling of the collected data is discrete. G and H can be expressed as rational functions of polynomials of the shift operator q , see equations. (2.3), and (2.4)

$$G(q, \theta) = \frac{B(q, \theta)}{F(q, \theta)} = \frac{b_1 q^{-nk} + b_2 q^{-nk-1} + \dots + b_{nb} q^{-nk-nb-1}}{1 + f_1 q^{-1} + \dots + f_{nf} q^{-nf}} \quad (2.3)$$

$$H(q, \theta) = \frac{C(q, \theta)}{D(q, \theta)} = \frac{1 + c_1 q^{-1} + \dots + c_{nc} q^{-nc}}{1 + d_1 q^{-1} + \dots + d_{nd} q^{-nd}} \quad (2.4)$$

where nk is the time delay, nb and nf is the order of the numerator respectively denominator of G , for H the orders are nc and nd . b_i, f_i, c_i, d_i are the parameters in the theta vector θ that is to be determined and q is the shift operator. By choosing different values on the parameters nb, nc, nd, nf and nk any linear system of interest can be described. There are several different models structures of this kind that is commonly used. Two of these model structures, ARX (*Auto Regression with External input*), and OE (*Output Error*) are described in this part, they were the two model structure that were used in this project. They are two special cases of the above model structure equation (2.2).

In the ARX model structure, as can be seen in equation (2.5), $C = I$ ($nc = 0$), $F = D = A$. The polynomial A is thus the common denominator for both the inputs and the noise.

$$A(q)y(t) = B(q)u(t) + e(t) \quad (2.5)$$

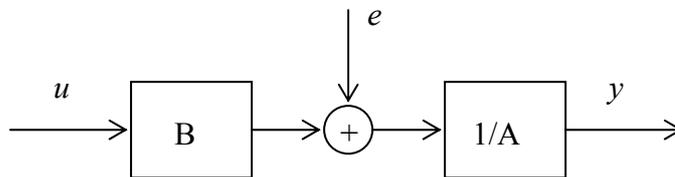


Figure 2.4 A picture of a system with ARX model structure, u is the input, e is the noise, y is the output. B and A are the numerator and denominator of the system.

As shown in Figure 2.4 the noise comes in early in the process and goes through the same dynamics as the inputs.

In the OE model structure $H = 1$ ($nc = nd = 0$) see equation (2.6), the noise comes in late in the process and basically only effect the output, see Figure 2.5.

$$y(t) = \frac{B(q)}{F(q)}u(t) + e(t) \quad (2.6)$$

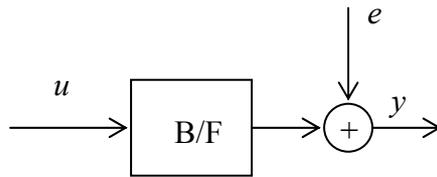


Figure 2.5 A picture of a system with OE model structure, u is the input, e is the noise, y is the output. B and F are the numerator and denominator of the system.

For the ARX structure the best prediction of y is given in equation (2.7). From equation (2.5) the optimal prediction can be found by neglecting the noise.

$$\hat{y}_{ARX}(t) = -a_1y(t-1) - a_2y(t-2) - \dots - a_nay(t-na) + b_1u(t-nk) + b_2u(t-nk-1) + \dots + b_nbu(t-nk-nb+1) \quad (2.7)$$

In the OE structure the noise can not be neglected in the same way. Equation 2.6 can be rewritten as follows.

$$y(t) = \frac{B(q)}{F(q)}u(t) + e(t) \quad \Rightarrow$$

$$y(t) + f_1y(t-1) + f_2y(t-2) \dots + f_nfy(t-na) = b_1u(t-nk) + b_2u(t-nk-1) + \dots + b_nbu(t-nk-nb+1) + e(t) + f_1e(t-1) + f_2e(t-2) \dots + f_nae(t-na) \quad \Rightarrow$$

$$y(t) = -f_1y(t-1) - f_2y(t-2) \dots - f_nfy(t-na) + b_1u(t-nk) + b_2u(t-nk-1) + \dots + b_nbu(t-nk-nb+1) + e(t) + f_1e(t-1) + f_2e(t-2) \dots + f_nae(t-na)$$

The best predictor of the expression above is when the noise $e(t)$ is neglected, but there is still old noise $e(t-1)$, $e(t-2)$... etc that makes the expression difficult to predict. To be able to predict $y(t)$ the old noises are considered by having old predicted values in the predictor instead of measured. This is explained with a longer proof in (Söderström, Stoica, 1989). The optimal predictor for the OE model structure is given in equation (2.8)

$$\hat{y}(t)_{OE} = -f_1\hat{y}(t-1) - f_2\hat{y}(t-2) - \dots - f_{n_f}\hat{y}(t-na) + b_1u(t-nk) + b_2u(t-nk-1) + \dots + b_{n_b}u(t-nk-nb+1) \quad (2.8)$$

where a_n are elements in the A vector and consequently b_n and f_n are the elements in B and F . For the A , B and F vector see equation (2.5) and (2.6) and Figures 2.4 and 2.5 y is the measured data, \hat{y} is the predicted values and e is the noise (Ljung and Glad,1991).

ARX is sensitive to high frequency properties and the OE structure is more sensitive to low frequencies. ARX can give good results in lower frequency areas if both the inputs and the outputs are filtered through a low pass filter. If the filter is the same for both the inputs and outputs and the system is linear there will be no change in the input-output relation. This is also the case for the OE model structure. If a high pass filter is applied, the OE can give better results in the higher frequency areas. In the case of modelling an irrigation channel reach, which has the interesting dynamics, for control purposes, in the low frequency area (changes due to dynamic properties are slow) it is more interesting with models that make more accurate models in the low frequency area (Ljung,1999).

2.3.5 Choice of model order

When the model structure is chosen an order to the model has to be determined. It is very important that a correct model order is chosen. A too low model order will not be able to describe the true dynamics. But on the other hand the model can be worse if too many parameters are included. The model may try to describe the noise that has nothing to do with the dynamics of the system. Too many parameters will only incorporate more uncertainty to the model. If there is knowledge about the system behaviour it can be an easy task to Figure out what model order that would give a good result, in other situations there are tools that can help to see if the model order chosen is the right one, see literature (Ljung, Glad, 1991) or (Ljung, 1999).

2.3.6 Grey box model structure.

In the case of this project there is prior knowledge that can be used for the model structure selection. To incorporate the knowledge into the model, the knowledge needs to be combined with the theory about model structure selection above. It will be better described how this was done in the Section about “model structure selection” in Chapter 4.

2.3.7 Parameter estimation

Within system identification there are different methods to use, but the method that is seen as the basic system identification approach is the prediction error method. The theory of Prediction error method is quite simple. The basic idea is to minimize the function $V(\theta)$, which is a function of the squared error between the measured output and the predicted/simulated output from the model. The θ vector contains the estimated parameters. See equation (2.9) and (2.10)

$$V(\theta) = \frac{1}{N} \sum_{t=1}^N \ell(\varepsilon(t, \theta))^2, \quad (2.9)$$

$$\varepsilon(t, \theta) = L(q)(y(t) - \hat{y}(t, \theta)) \quad (2.10)$$

where $\ell(\cdot)$ is a scalar valued function, $L(q)$ is a filter, $y(t)$ measured output, $\hat{y}(t, \theta)$ the predicted output and N is the number of measurement used in the estimation (Ljung, 1999). This method is an optimization problem that within the chosen model structure finds the model that gives the smallest average error between the model output and the measured output. Equation (2.9) needs to be minimized to find the θ vector that satisfies this optimization.

For some model structures it can be very straight forward and simple to find the θ vector with an analytical solution. Other model structures are more complicated and need an iterative searching method to find the optimal solution. Since ARX and OE are the model structures mentioned above, the methods for these two model structures are described below

In the case for the ARX model structure the minimization of (2.9) can be done analytically. The only thing that is needed for the parameter estimation is the measured data. The ARX-model structure can be written as a linear function. The ARX model from equation (2.5) can be written as

$$y(t) + a_1 y(t-1) + \dots + a_n y(t-n) = b_1 u(t) + b_2 u(t-1) + \dots + b_m u(t-m-1) + e(t) \quad (2.11)$$

The parameters a and b can be combined into a θ vector and $y(t)$ and $u(t)$ into $\varphi(t)$. If the $e(t)$ is neglected the ARX model can be written as equation (2.12)

$$y(t) = \varphi(t)\theta \quad (2.12)$$

$$\varphi(t) = \begin{bmatrix} y(t-1) \\ y(t-2) \\ \vdots \\ y(t-n) \\ u(t) \\ u(t-1) \\ \vdots \\ u(t-m-1) \end{bmatrix}, \quad \theta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \\ b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

where y is the output, φ is the vector with old y and inputs for the whole data set and θ is the unknown parameters. From this expression and the optimization conditions in expression (2.9) and (2.10) θ can be solved as follows.

$$\hat{\theta} = \left[\sum_{t=1}^N \varphi(t)\varphi^T(t) \right]^{-1} \left[\sum_{t=1}^N \varphi(t)y(t) \right] \quad (2.13)$$

where $\hat{\theta}$ is the estimate of the unknown parameters, $\varphi(t)$ is the input vector for t and $y(t)$ is the output, For the theory behind this expression see literature (Söderström, Stoica, 1989).

The OE model structure is using the value from the last prediction step to predict the next step. It is not possible to find an analytical expression for this optimisation problem. Finding

the θ vector is therefore more complex than for the ARX case. The optimization is solved by using an iterative search scheme. For every step in the search for θ the prediction for the estimation set is calculated. This can be very time consuming, but with a routine in a suitable programming language this can be made more efficient. The search starts with a given starting value for θ . The $V(\hat{\theta})$ is calculated and compared with the previous $V(\hat{\theta})$. When the minimum $V(\hat{\theta})$ is found the search for θ stops. There are different searching methods with different searching properties. The method used in this project is the Levenberg-Marquardt searching method. This method is supposed to be able to handle the search even though the model is over parameterized or if the data is not informative enough. The numerical search results in a minimum of $V(\hat{\theta})$. If this is a global or local minimum is difficult to know. If it is a local minimum it means that it is not the best model in the particular model structure. If the model shows satisfying results in the validation it shows that the model is good for its purpose even though it is not the best one (Ljung, 1999).

2.3.8 Model validation

To be able to judge the model quality there is a need for a measurement that can show how good the models are. The model validation is the tool for such a measurement.

When the unknown parameters are estimated the best model within that model structure, and on that particular estimation set, is built. The model is adapted to the estimation data set. The $V(\hat{\theta})$ does not give a true judgement of how good the model is to describe the system. It gives a very good judgement on how good it is to model the estimation set. But the model could have tried to describe the noise of that particular set. What happens when it tries to model the same system but with another set of noise? The purpose of the validation is to predict/simulate another data set and compare it with the measured values. To evaluate how well the model can describe the validation data set, $V(\hat{\theta})$ is calculated, see equation (2.9) and (2.10). $V(\hat{\theta})$ is never used as the only factor to evaluate a model. A common way to complement it is to study the plot of the prediction/simulation and the measured data (Ljung, Glad, 1991).

The use of the model decides whether the model is good enough or not. The demand on accuracy of the model can vary depending on the area of use. A model used for control does not demand as high accuracy as a model aimed for simulation. As long as the main trend in the data can be described the controller can easily handle the static error by having an integrating property. And by having low gain in the higher frequency area it does not matter too much if the model does not capture the high frequency properties. For simulation purposes there is a higher demand on model accuracy. Of course it depends on why the simulation is made how good the model has to be.

2.3.9 Model validity

When it comes to using the model for practical purposes, it is very important to know for what conditions the model is valid. The model is only valid for the conditions which were during the collections of the validation set. For example, the model is valid if the inputs are changed in similar manners and size as the validation set, and if the outputs are moving in the same range as in the validation set. Outside of these conditions the model is not valid at all and nothing can be said about the model performance (Ljung, 1999).

3 INFORMATION ABOUT THE IRRIGATION SYSTEM

As said earlier in this report, the prior knowledge of the system is an important part of the system identification. The known parts of the system need to be looked into and judged whether or not it is useful for the model building. In this Chapter the prior knowledge of the system is given, starting with the site description and earlier studies and ending with the physical knowledge.

3.1 SYSTEM DESCRIPTION

3.1.1 History of the Area

As a result of the Snowy Mountain Hydro Electrical Scheme, diverted water could be used for other purposes than hydro electric power. One idea was to divert the water for irrigation purposes. That was the main reason why the Coleambally Irrigation Area (CIA) was formed. The CIA irrigates an area of about 79 000 ha which is all together 452 farms. The Small town of Coleambally was built in the 1960's as a consequence of the irrigated agriculture in the area. The size of the town is about 1200 inhabitants (website 4, 2004).

3.1.2 Site description

The CIA is an open channel system, where the water flows in contact with the atmosphere. As described above the CIA is a very big irrigation system. The part of the system that was studied in this project was the Coleambally Main Channel (CMC). The Main Channel consists of a number of pools and the gate sites that separates the pools. For a schematic picture see Figure 3.1.

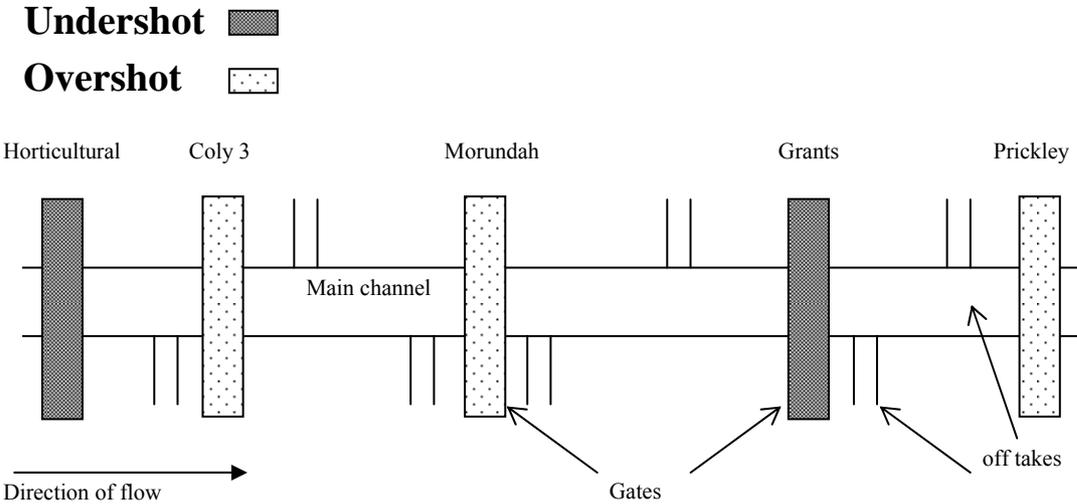


Figure 3.1 A top view of the Coleambally Main Channel Where Horticultural, Coly 3, Morundah, Grants and Prickley are the gate sites, the grey bars are representing undershot gates and the bars with dots are representing overshot gates.

The studied channel reach of the CMC starts with Horticultural gate site, with undershot gates, and ends with Prickley gate site with overshot gates. In between there are Coly 3,

overshot, Morundah, overshot, Grants, undershot. At each gate sites there are different number of gates. For an overview see Table 3.1

Table 3.1 The number of gates at each gate site.

Gate site	Horticult.	Coly 3	Morundah	Grants	Prickley
Type of gates	Undershot	Overshot	Overshot	Undershot	Overshot
Number of gates	4	6	6	4	5

In the rest of the report the gate sites will be referred to as the first letter of their names. For examples, Horticultural will be referred to as H.

A pool is considered as the channel reach between two gate sites. There are differences in the lengths of these pools. H-C is about 2.8 km, M-G about 5.4 km and G-P about 5-6 km.

3.1.3 Pool description,

Of the above described pools, the H-C and M-G were the ones used in this project. It is because they both have undershot and overshot gates, which was the main criterion for being studied in this project, see the aim of the project in Section 1.1. For H-C there are 4 undershot gates at the upstream end of the pool and 6 undershot gates at the downstream end. For the M-G it is the other way around. From each pool there are off takes. The off takes are considerable smaller than the flow from and to the pool.

3.1.4 Earlier studies

Experience from earlier work in the same field of study was a very useful tool during the project. The work used was “system identification of an open water channel”, (Weyer 2001), performed on the Haughton Main Channel (HMC) in Queensland, Australia. The HMC have similar design to the CIA. The difference is that the HMC is a smaller channel and it has only overshot gates and only maximum two gates in parallel. Weyers work gave very satisfying results. He managed to build a first and second order model that captured the trends in the data very well. With a third order model he could even capture the wave dynamics which is a high frequency property in the HMC. The appearance of wave dynamics was very clear when looking at the experiment data (Weyer, 2001). The physical theory about the system that Weyer used, is described in Section 3.2.2. Since the HMC and CIA have similar design, the same theory is used in this project.

3.2 PHYSICAL KNOWLEDGE

The pools that are studied in this project have the design as in Figure 3.2. The Figure shows the H-C pool where there are undershot gates upstream and overshot gates downstream.

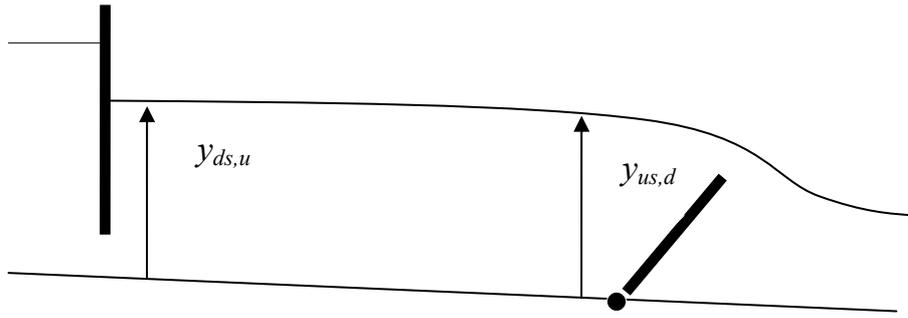


Figure 3.2 A side view of a single pool in the irrigation channel, where the inflow is controlled by an undershot gate and the outflow is controlled by an overshot gate. $y_{ds,u}$ is downstream water level of the upstream end gate site. $y_{us,d}$ is the upstream water level of the downstream end gate site. The subscripts ds,u stands for **d**ownstream at **u**pstream end.

3.2.1 St Venant equations

Below the physical knowledge about the system is described. This is the prior knowledge used in the model structure selection and experiment design.

In Section 2.1 it was said that a model could be built from physical knowledge. In this project the dynamics of an irrigation pool of this kind can be described by the Saint Venant equations. The Saint Venant equations are based on the conservation laws of mass and momentum (Baume et al.,1998). With various assumptions and simplifications the one dimensional flow in an open channel can be described by equations (3.1) and (3.2).

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

$$\frac{\partial Q}{\partial t} + \left(\frac{gA}{B} - \frac{Q^2}{A^2} \right) \frac{\partial A}{\partial x} + \frac{2Q}{A} \frac{\partial Q}{\partial x} + gA(S_f - \bar{S}) = 0 \quad (3.2)$$

where A is the cross Sectional area of the channel, B is the width of the water surface, Q is the flow, g is the gravity constant, S_f is the friction slope, which is a non linear function of the channel properties, and \bar{S} is the mean bed slope. These equations are partial differential equation and quite numerically demanding to simulate. The simulations are made by iterative numerical calculations in both space and time, see Figure 3.3.

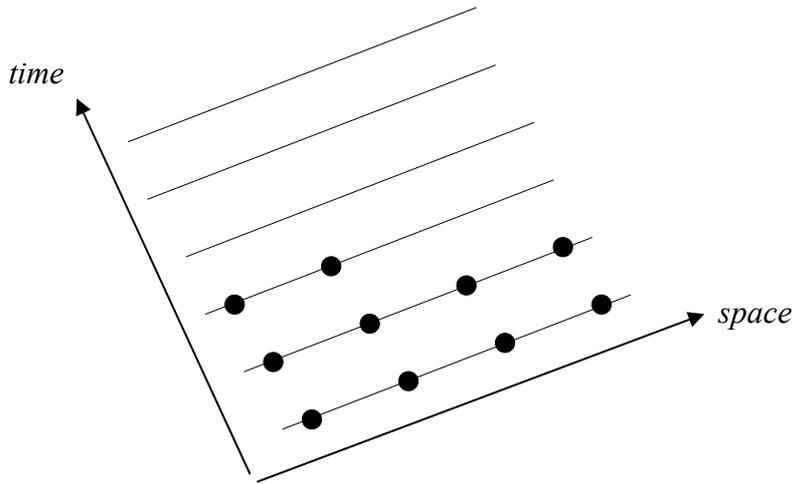


Figure 3.3 A mental picture of physical model iterations over space and time. The dots represents numerical solutions that are solved first in space and then for every time step.

To build this kind of model, many things about the system have to be known. It is not only the physical laws but also properties of the pool that can be difficult to determine. One example in this case is the friction slope. To start the simulations, values of a whole profile in space is needed.

This is however a time consuming and possibly an unnecessary task in this project. The interesting output is the downstream water level of the pool, and not all the information about the water profile upstream is needed. There is another way of describing the physics of the system in a simpler way, which is described below.

3.2.2 Integrator

A simple way of viewing an irrigation pool is to regard it as an integrator. A simple mass balance, volume of water that goes into the pool minus the volume that goes out from the pool is the change of volume inside the pool during a certain period of time. If it is assumed that the cross sectional area of the pool is constant and increases linearly with height, the change in water level is proportional to the change in volume, see equation (3.3).

$$\frac{dy}{dt} = K(Q_{in} - Q_{out}) \quad (3.3)$$

where y is the water level, K is a constant, Q_{in} and Q_{out} are the inflow and outflow.

3.2.3 The flow through the undershot and overshot gates

The basic theories of overshot and under shot gates are sketched in Figure 3.4. In the following Section the physical theory of this is described in more detail.

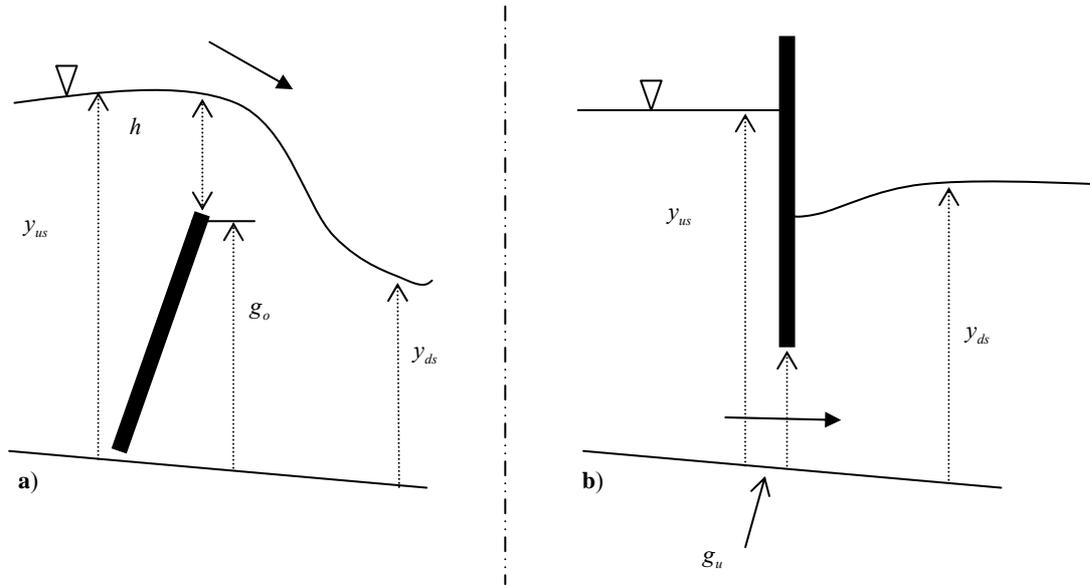


Figure 3.4 a) A detailed picture of an overshoot gate, g_o is the gate position (subscript o stands for overshoot gate), h is the head over gate, y_{us} is the upstream water level and y_{ds} is the downstream water level. b) A detailed picture of an undershot gate, g_u is the gate opening, y_{us} is the upstream water level and y_{ds} is the downstream water level.

The flow over an overshoot gate can be described by the expression in equation (3.4)

$$Q = CL\sqrt{2g}h_1^{3/2} \quad (3.4)$$

where C is a constant, L is the width of the gate, g is the gravitation constant and h_1 is the head over gate which can be written as upstream water level minus gate position (Bos, 1976) and (Fox et al, 1994). This expression was used by Weyer in the HMC study. This expression is valid considered that the gate is in free flow. If a gate is in free flow the downstream water level of that gate is lower than the gate position. In the CIA there is a possibility that drowned flow can occur, the downstream water level of a gate is above the gate position. When drowned flow occur the flow properties changes and thus the flow function, see equation (3.5) which is an empirical function (Webber, 1971)

$$Q = Q_1 \left[1 - \left(\frac{h_2}{h_1} \right)^{3/2} \right]^{0.385} \quad (3.5)$$

where Q is the flow through the gate, Q_1 is the flow expression for the gate in free flow equation (3.4), h_1 is the head over gate, distance between upstream water level and gate position. h_2 is the downstream head, distance between downstream water level and gate position. The total flow expression (3.6) is as follows

$$Q = \begin{cases} CL\sqrt{2g}h_1^{3/2} & \text{if } h_2 \leq 0 \\ CL\sqrt{2g}h_1^{3/2} \left[1 - \left(\frac{h_2}{h_1} \right)^{3/2} \right]^{0.385} & \text{if } h_2 > 0 \end{cases} \quad (3.6)$$

With a closer look at this function it can be seen that the function is continuous. When the immediate downstream water level equals the gate position the fraction h_2/h_1 becomes zero. The drowned flow expression becomes the free flow expression (Webber, 1971). In literature there are different ways of describing drowned flow. It has been found for a particular channel that the above expression gives the best results in modeling drowned water flow. This was discovered in a fourth year project, by Gordon Chessum, supervised by Weyer at the Department of Electrical and Electronic Engineering, Melbourne University. Therefore equation (3.5) was used in this project to model drowned flow.

In the CIA the undershot gates are submerged. It means that the immediate downstream water level, y_{ds} in Figure 3.4 b, is above the gate position. The flow under the undershot gate depends not only on the upstream water level and gate position, but also on the downstream water level. From the Bernoulli equation, the following expression can be found in literature,

$$Q = CLg_u \sqrt{2g(y_{us} - y_{ds})} \quad (3.7)$$

where C is a discharge coefficient, L is the width of the gate, g_u is the gate opening, g is the gravitational force, y_{us} is the upstream water level and y_{ds} is the downstream (Baume, 1998). But the literature do not totally agree, for example in (Fox, 1994) it says that it is impossible to analyze flow through an undershot gate with submerged flow. One particular expression (3.8) may be used with an appropriate value of the discharge coefficient.

$$Q = CLg_u \sqrt{2g(y_{us})} \quad (3.8)$$

The discharge coefficient C depends on the ratio (*gate position*)/(*upstream water level*). This means that C changes with conditions (Rouse, 1961). That is a hint of that the system behaves different depending on conditions and that a system identification model shall be used very carefully within the range of the models validity. The expression (3.7) above was the one used in this project. Since it was mentioned in a paper about modeling and control of open channel irrigation systems it seemed more suitable to start with (Baume et al., 1998).

4 SYSTEM IDENTIFICATION PROCEDURE APPLIED ON THE CIA

In this Chapter it is described how the system identification procedure was applied to the irrigation system.

4.1 THE EXPERIMENT

The experiment design is an important part of the system identification and it is essential that the experiment is well planned so the aim of the experiment is fulfilled.

The experiment was conducted, at the Rubicon office in Melbourne, over a period of 4 days, starting 2/2 2004 and ending 5/2 2004. The signals were sent via radio network from a remote desktop to the gates in the field.

The aim of the experiment was to get an informative data set where there is a significant variation in water level with a sample rate that also captures the higher frequency properties of the system. The plan for the experiment was to first put a step to the system to get an idea of the system properties such as time delay, time constant, gain etc. With this information a BS (Binary Signal) was created and then later applied to the system as the exciting signals.

The sampling period was close to every second minute. Every gate site from Horticultural to Prickley was sampled with the mentioned sampling rate. The properties that were sampled were the upstream water level, downstream water level, gate positions and flow. All the gate sites were manually sampled at a rate of one in every 60-75 minutes. This was done for later comparison with the automatically sampled data.

4.1.1 Limitations of the experiment

Disturbances can have severe affects on the irrigation system. The water delivery to the farmers can be considerable lowered if the water levels are sunken too much for a long time. It can have a large negative effect on the crop production if the water demand at the time is high. On the other hand, if the water level rises too high, there will be water wasted. This is not a minor problem. The cost of water per Ml per day is ~\$100 (Weyer, 2004). It is easy math to Figure out that the cost of the water wasted will not be small if the gate change is a couple of centimetres too big. 1mm gate change per gate gives a flow change of 1 Ml per day (Rubicon, 2004). Fortunately there was not a big demand of water at the time of the experiment so no particular lower bound for the water levels were set. There were however a couple of other limitations that had to be considered:

- The water level change should not vary much more than 10 cm.
- All the water levels had upper limits that they should not be raised above, see Table 4.1
- Only the two middle gates at each gate site were allowed to be manipulated.
- There were only four days to conduct the whole experiment procedure including the step test.
- Experiments could only be conducted during office hours. That constrains the experiment to be around 10 hours for one continuous experiment (1 day).
- The experiment had to be monitored the whole time.
- Information sent via the radio network from the irrigation channel was limited.
- The input signals were limited to be gate positions.

Table 4.1 The upper limits of the water levels in the experiment at each gate site.

Gate site	Upper limit [m]
H	2.12
C	2.10
M	2.05
G	2.02
P	2.15

4.1.2 The step test

Day one was used to get an idea of how much the gates could be opened to get a reasonable step within the limitations above. On day two the step test was applied to the system. The gates were moved at gate sites H and M, inducing a step on the water levels upstream C and G. The step in gate positions was 200 mm. For the overshoot gates there are nonlinear dynamics associated with the transfer function between gate position and water head (distance between upstream water level and gate position) the gate positions were changed during the step test to keep the step in head constant. The system is linear from flow to water level, constant head would give a good linear step response for the overshoot gate. It is not true for the undershoot gate, constant head over gate would not give a linear step response. The head over gate was accidentally held constant on the undershoot gates at H. During the step test, the closest gate sites were held constant so the step test would be as undisturbed as possible.

4.1.3 The results from the step test

Day one gave a hint on how big the gate openings could be for a reasonable step response. As said earlier the step in gate positions was chosen to be 200 mm. The step response turned out to be good but small. The system could have handled a slightly bigger step. The step gave information about the time delays of the different pools. In Figure 4.1 the step tests can be viewed, the time delay is the time between the step in input and the beginning of the step in output. The estimated time delays are shown in Table 4.2.

Table 4.2 Time delays found in the experiment.

Pool	Time delay [min]
H-C	~10
M-G	~19

The time constants for the different pools are estimated from the plots as the time from the start of the step until steady state, divided by five. The time constant is thus a measurement on how fast the system responds to a change in the inputs.

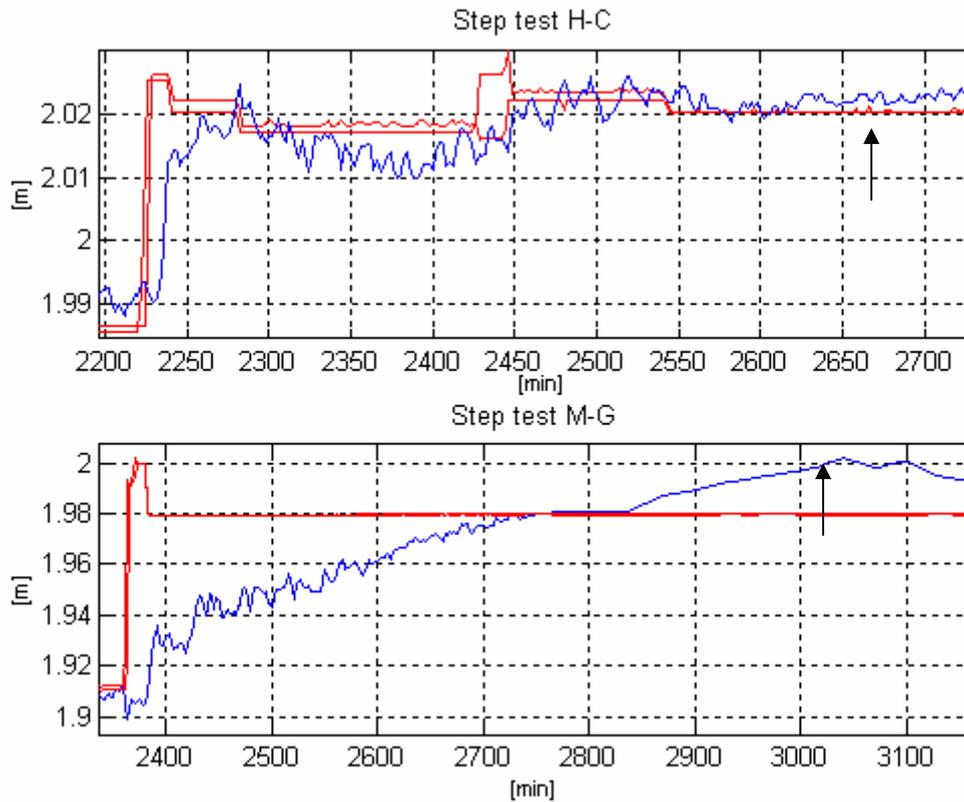


Figure 4.1 Step tests in the water levels at the downstream end of the H-C pool and M-G pool. The step in gate positions was 200mm. The gate positions are not in the same scale as the y-axis. The arrows represent the estimated steady state of the water level.

The time constants are represented in Table 4.3. It is worth mentioning that the determination of the time constants was not a simple task because the step was rather small compared to the noise. In Figure 4.1 arrows are representing where steady state was estimated to be.

Table 4.3 Time constants of the experiment

Pool	Time constant [min]
H-C	68
M-G	120

The size of the step responses are represented in Table 4.4. The step for the longer pool M-G was much bigger than for the shorter pool H-C. It was expected to get a bigger step in the shorter pool. An explanation to that could be that there was little time to wait for steady state before applying the step. Before the steps were applied gates had been moved in the direction of the step for the M-G pool and in the opposite direction for the H-C pool.

Table 4.4 Step response sizes of the experiment.

Pool	Step response [cm]
H-C	3.5
M-G	9

4.1.4 The Binary Signal

The BS was designed as signals jumping between two levels in gate positions. The gate positions are the inputs that are available for control. It is therefore interesting to know how the system behaves when a change occur in the gate positions. The energy was chosen to be close to the corner frequency, ω_c , very low frequencies were intentionally not chosen considering the limitation of time. It was important to get as much information as possible from the short time available. The signal was designed so that a significant but not too big variation in water level could occur and that the amount of water into the pools would be roughly the same as the amount water out from the pool averaged over the experiment. The time constants from the step test were used to design the frequencies of the binary signals. The time constants were 68 min and 120 min for the different pools. The Signals were chosen to change levels in intervals of 30 min to 320 min for the H-C pool, the pool gets information about the system around the bends in the bode plot, close to the corner frequency. For the M-G pool the intervals were chosen to be longer due to the longer time constant, between 60min to 200 min. The levels of which the gates were jumping between were ± 100 mm around set point for the first BS experiment. Set points are the gate positions which were during the start of the experiment. The levels were chosen with the information from the step test. Because of the limitations of the experiment rather small levels was the first choice so the limitations not were exceeded. It showed that the levels could be increased for three of the gate sites to get a more informative result for the last day of experiment. So the levels for day 4 were changed for H and G to set point ± 150 mm. The signals that were applied to the system can be studied in Figures 4.2-4.9, together with the system response to these signals.

To perform the experiments within the limitations, the system had to be adjusted before applying any signal. Water was let out from the two most downstream pools to lower the water levels so it was possible to stay within the limitation boundaries.

4.1.5 The result from the experiment

In Table 4.5 the range of the water levels from the experiment is represented

Table 4.5 The largest water level differences for the two days of experiment. DSwl.diff. stands for down stream water level difference and USwl.diff. stands for Upstream water level difference. These water levels refer to a specific gate site. Upstream and downstream the particular gate site.

Gate site	DSwl.diff.	USwl.diff.	DSwl.diff.	USwl.diff.
	day 1 [cm]	day 1 [cm]	day 2 [cm]	day 2 [cm]
H	12	8	13	15
C	6	9	12	9
M	8	17	11	13
G	18	14	15	13

In Table 4.6 the gate positions for all the gates are represented for both the binary signal experiments and for all the gate sites.

Table 4.6 The gate positions of the experiment of the different gates at each gate site. The gate position is the distance between the bottom of the channel to the opening of the gate. For the two middle gates the values for the two levels of the binary signals are given.

Data set	H		C		M		G	
	BS 1 [mm]	BS2 [mm]	BS1 [mm]	BS2 [mm]	BS1 [mm]	BS2 [mm]	BS1 [mm]	BS2 [mm]
Gp1	153	153	1450	1450	2010	2010	60	160
Gp2	980/1180	940/1235	1540	1540	1670	1570	665/865	540/735
Gp3	980/1180	940/1235	1400/1200	1375/1075	975/775	995/795	665/865	540/740
Gp4	103	103	1390/1190	1375/1075	970/770	1015/810	8	8
Gp5			1248	1135	1710	1710		
GP6			1780	1780	1710	1610		

The result of the experiment is given in Figure 4.2-4.9. In the Figures the upstream water level and the gate positions are moved in the plots to get a better view of the information. The data is still plotted in the same scale. In the Figure text it is given how much they are moved. For example in Figure 4.2 the upstream water level is moved -0.4 m, it means that the water level is not actually starting at 1.6 m. Its real water level is $1.6+0.4$ which is 2.0 m.

In Figures 4.2 and 4.3 there are arrows pointing at certain events in the experiment. These are just to show how the output behaves depending on how the BS is designed. For example in Figure 4.2, when the gate positions for the undershot gates increase (1) at H the flow through the gates increases and the water level upstream the gate decreases (2). After the time delay of the pool the upstream water level at C starts to increase due to the increased flow into the pool (3) in Figure 4.3. About an hour later, also in Figure 4.3, the gate positions increases at C. This means that the overshot gates rises and the outflow decreases (4). The upstream water level at C is increasing even more due to that (5). Arrow number 6 in Figure 4.3 is pointing at a time in the experiment where the gate is in free flow, the water level is below the gate positions. This shows that it is important to consider both free flow and drowned flow when modelling this irrigation channel.

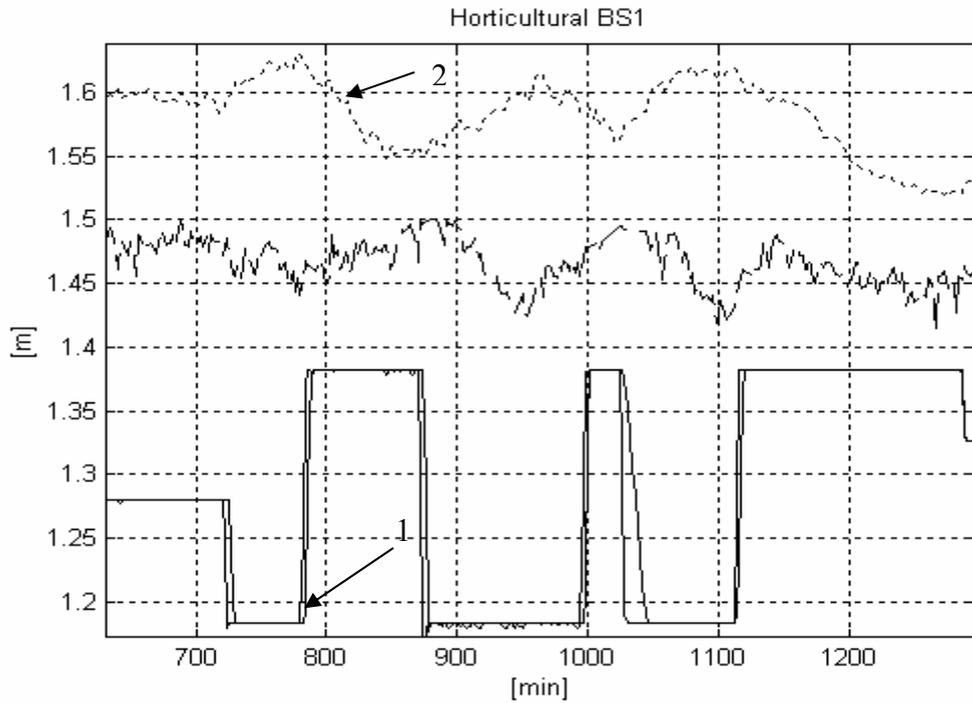


Figure 4.2 Experiment results from H day1. Dotted line: upstream water level, dashed line: downstream water level, solid lines: gate positions. In the plot upstream level is moved -0.4 m and gate pos. are moved +0.2m

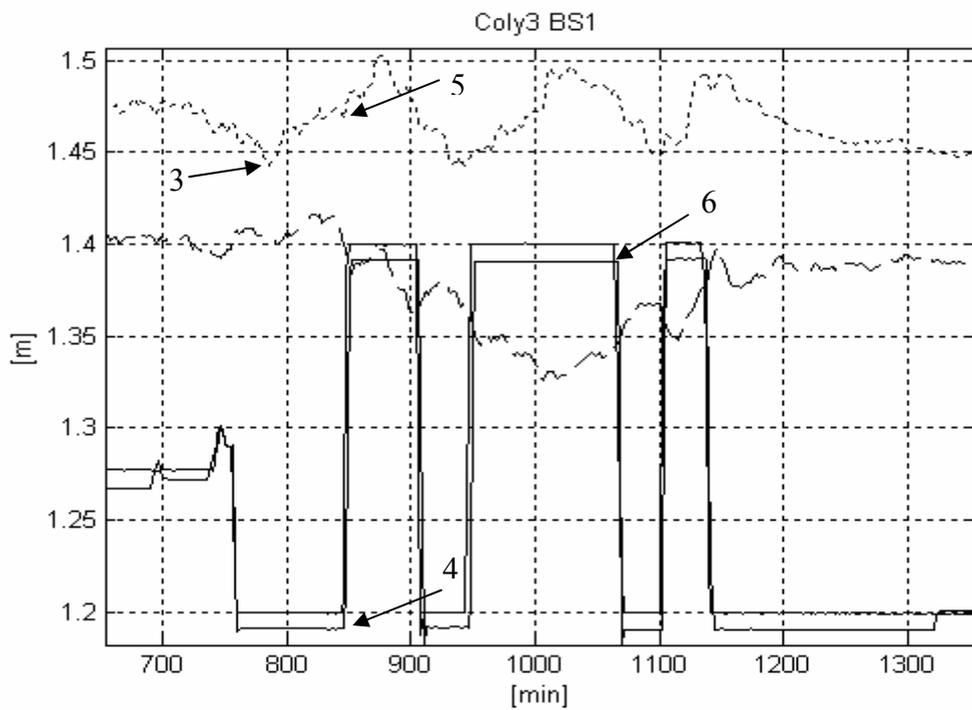


Figure 4.3 Experiment results from C day1. Dotted line: upstream water level, dashed line: downstream water level, solid lines: gate positions. In the plot only the upstream level is moved -0.55 m. This means that it is easy to see whether the flow is drowned or not by comparing the downstream water level with the gate positions

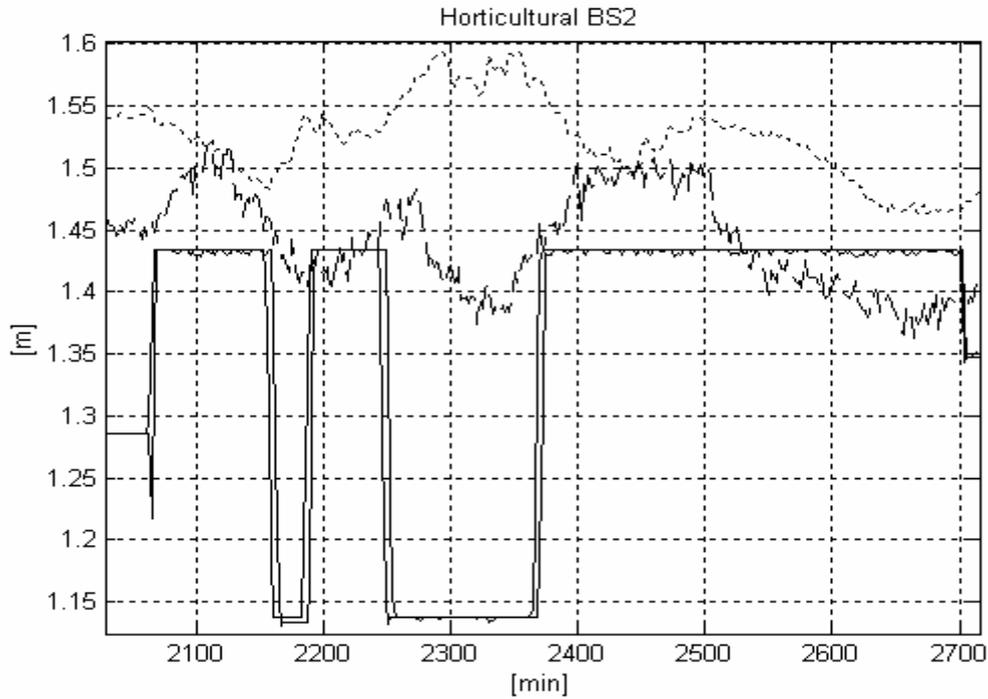


Figure 4.4 Experiment results from H day2. Dotted line: upstream water level, dashed line: downstream water level, solid lines: gate positions. In the plot upstream level is moved -0.4 m and gate pos. are moved 0.2m

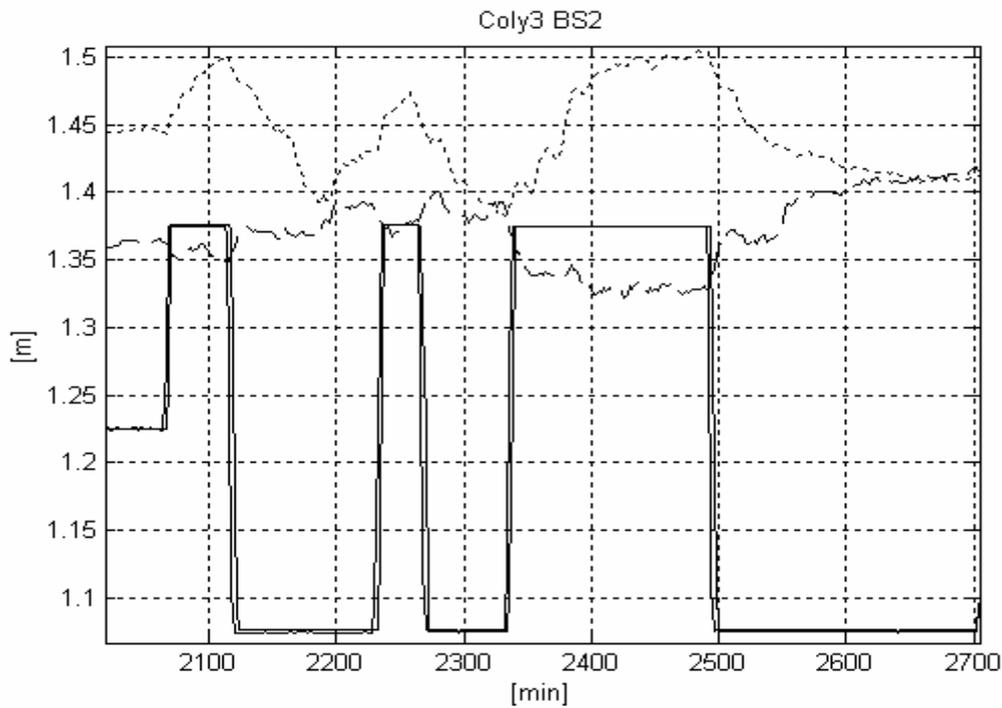


Figure 4.5 Experiment results from C day2. Dotted line: upstream water level, dashed line: downstream water level, solid lines: gate positions. In the plot upstream level is moved -0.55 m. This means that it is easy to see whether the flow is drowned or not by comparing the downstream water level with the gate positions.

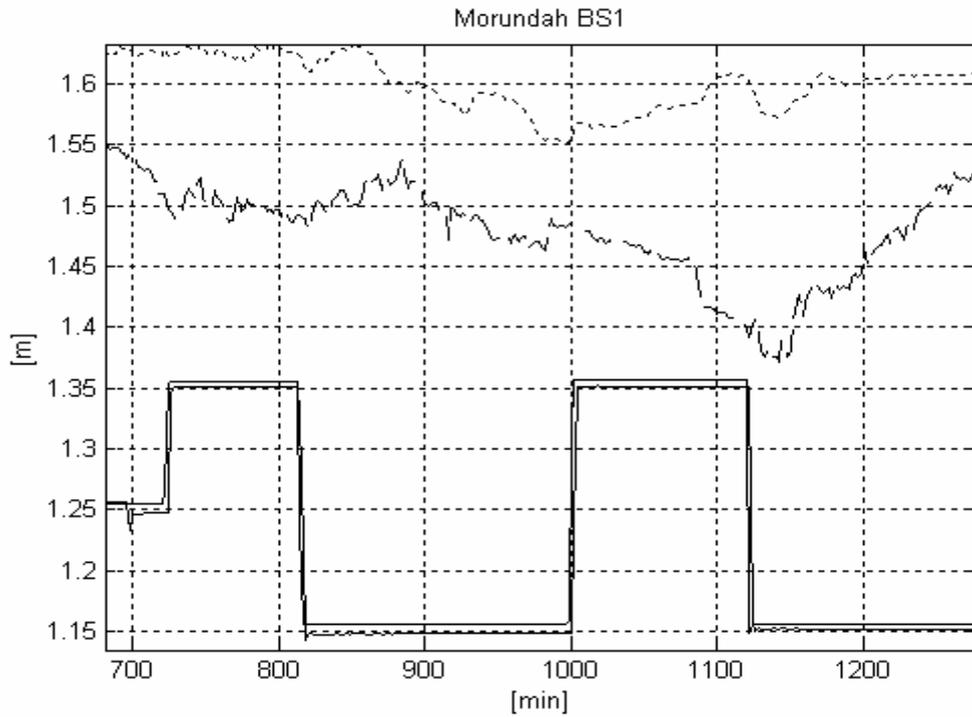


Figure 4.6 Experiment results from M day1. Dotted line: upstream water level, dashed line: downstream water level, solid lines: gate positions. In the plot upstream level is moved -0.4 m and gate pos. are moved 0.38m

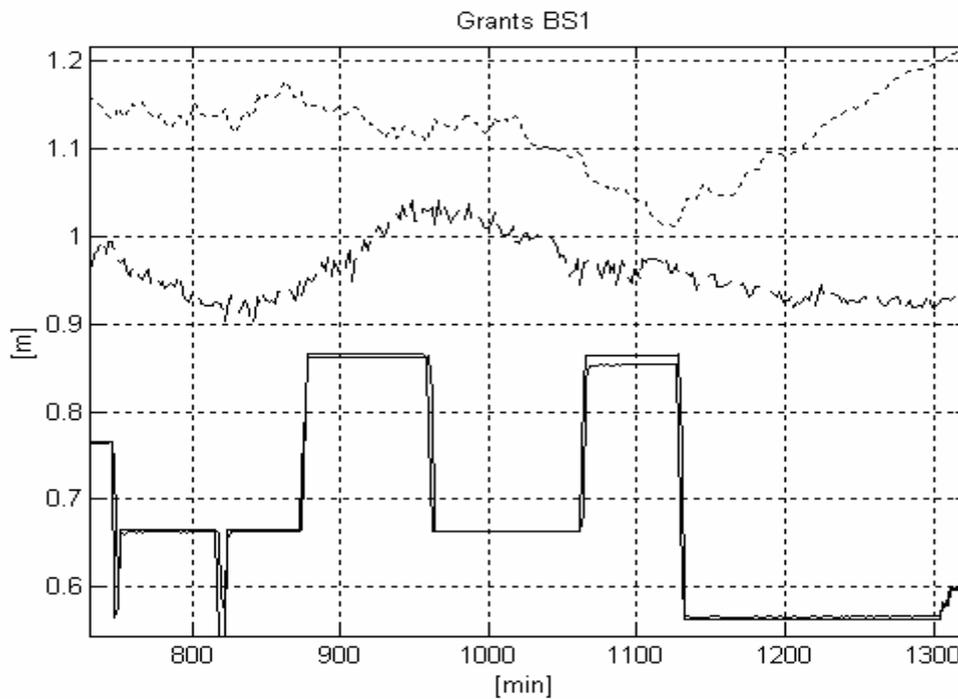


Figure 4.7 Experiment results from G day1. Dotted line: upstream water level, dashed line: downstream water level, solid lines: gate positions. In the plot upstream level is moved -0.8m. This means that it is easy to see whether the flow is drowned or not by comparing the downstream water level with the gate positions.

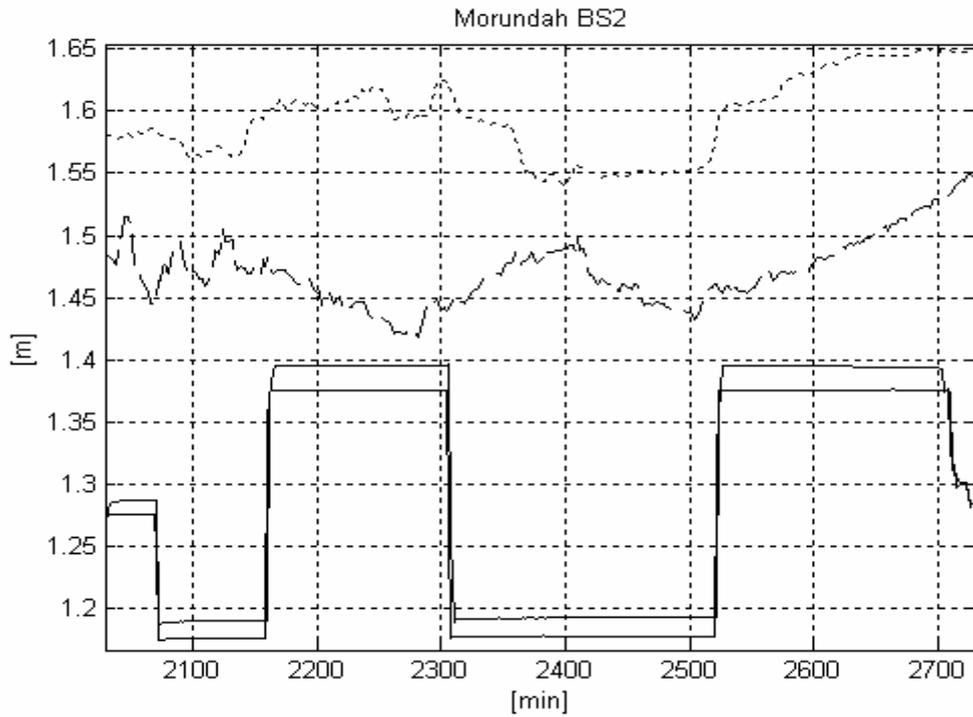


Figure 4.8 Experiment results from M day2. Dotted line: upstream water level, dashed line: downstream water level, solid lines: gate positions. In the plot upstream level is moved -0.4 m and gate pos. are moved 0.38m

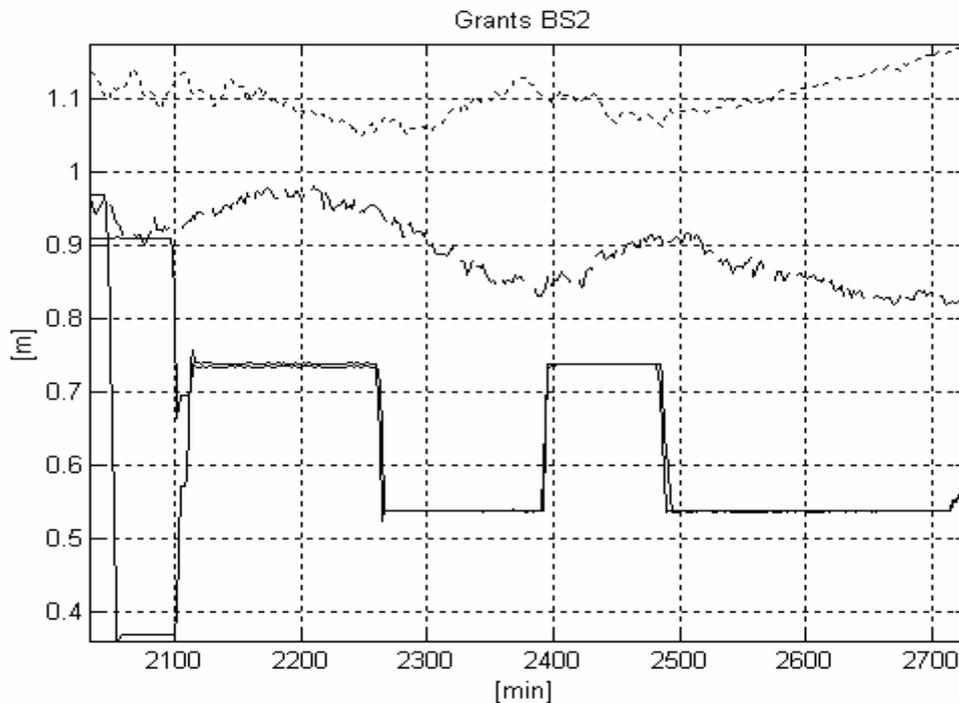


Figure 4.9 Experiment results from G day1. Dotted line: upstream water level, dashed line: downstream water level, solid lines: gate positions. In the plot upstream level is moved -0.8m. This means that it is easy to see whether the flow is drowned or not by comparing the downstream water level with the gate positions.

4.1.6 Discussions of the experiment results

In Figure 4.3 and 4.5 (C BS 1 and C BS 2, Binary Signals) gate positions and downstream water levels are not moved from their original values of the levels in the plots. It is easy to see where the flow is drowned or not. Where the gate positions are below the down stream water level the flow is drowned. The other gates which are not changed in the experiment are much less open and thus in free flow during the whole experiment. It is the same for the other overshoot gate, M, Figure 4.6 and 4.8. But the difference in this case is that the flow through the two changed gates in the middle are drowned during the whole experiment. The gates on the side that are not changed are not included in the plots see Table 4.6.

When gate positions are lowered for the undershot gates, H and G, it means that the flow decreases, see Figure 3.4. It is the opposite for the overshoot gates, C and M, where a lowered gate position gives increased flow. Thus the water level upstream an undershot gate will decrease when the gate position is raised and the water level downstream will increase. The water level upstream an overshoot gate will increase if the gate position is raised and downstream level will decrease and vice versa.

Take the example H-C, where undershot is upstream and overshoot is downstream. The water level at the downstream end of pool will increase when the upstream gate positions are raised but first after the time delay of the pool, and lowered when the downstream gate positions are lowered. To get a good picture of what happens in a pool during the experiment, study Figure 4.2 and 4.3 together as well as Figures (4.4, 4.5), (4.6, 4.7) and (4.8, 4.9)

4.1.7 Data processing

The data was loaded in to Matlab for analyzing. The data set was examined for outliers. No outliers were found in the interesting parts of the data set. It was also resampled, to get the data evenly distributed over time, with a sample rate of every second minute. A function in matlab, "INTERPL", interpolates between the data points and resamples with desired sample rate.

4.2 MODEL STRUCTURE SELECTION

With the rule Try Simple Thing First (Ljung, 1991) in mind a first order ARX model structure was first used.

The equation below, which is equation (3.1) from Chapter 3, is a very simple expression of the water level change.

$$\frac{dy}{dt} = K (Q_{in} - Q_{out}) \quad (4.1)$$

From that expression an ARX model structure can be built.

$$\hat{y}_{us,d}(t+1) = y_{us,d}(t) + KQ_{in}(t-\tau) - KQ_{out}(t) \quad (4.2)$$

where $y_{us,d}(t+1)$ is the water level at time $t+1$, $y_{us,d}(t)$ is the level at time t (the index us,d stands for upstream water level at downstream end of pool), τ is the time delay for the inflow to the water level y . Q can be replaced by either of the expressions (3.6) or (3.7), describing the flows through the different types of gates. If the pool H-C is going to be modelled the inflow shall be described by equation (3.7) (undershot gate) from Chapter 3, and vice versa for the M-G pool. K and the constants for the flow equations are combined together to the two constants c_1 and c_2 . See equation (4.3) for the example when H-C is modelled.

$$\hat{y}_{us,d}(t+1)_{(H-C)} = y_{us,d(H-C)}(t) + c_1 g_u(t-\tau) \sqrt{(y_{us,u}(t-\tau) - y_{ds,u}(t-\tau))} - c_2 (y_{us,d(H-C)}(t) - g_o(t))^{3/2} \quad (4.3)$$

where the index us,u and ds,u stand for upstream and downstream water level at upstream, end of pool. g_u and g_o are the gate openings for the undershot and overshot gates respectively. Equation (4.3) describes only the situation when the overshot gates are in free flow. In the actual model the expression has to alternate between the two cases whenever the gates changes in flow condition.

At each gate site there is more than one gate and the gates are not always moved together. The gates at one gate site have the same size and design. The total volume of water that flows through a gate site is the sum of the water that flows over the gates together. Equation (4.3) can be expanded as follows,

$$\hat{y}_{us,d}(t+1)_{(H-C)} = y_{us,d(H-C)}(t) + c_1 \left[(g_{u1}(t-\tau) + g_{u2}(t-\tau) \dots + g_{un}(t-\tau)) \sqrt{(y_{us,u}(t-\tau) - y_{ds,u}(t-\tau))} \right] + c_2 \left[(y_{us,d(H-C)}(t) - g_{o1}(t))^{3/2} + (y_{us,d(H-C)}(t) - g_{o2}(t))^{3/2} \dots + (y_{us,d(H-C)}(t) - g_{on}(t))^{3/2} \right] \quad (4.4)$$

where n is number of gates and the gate openings for the different gates are shown separately. For the undershot case the gates openings can be taken out from the flow equation and multiplied outside as in equation (4.4).

The ARX predictor is estimated to predict the system as a one step ahead prediction. The ARX model structure needs the measured value of the last step in the prediction, see equation (4.4). It is a useful model structure but if the model is meant for simulation it is better to build an OE model. And also in this case, where the important dynamics are in the low frequency area, it may be beneficial to use the OE model structure. In fact in this project the ARX model was used to give a starting value for the search of the parameters of the OE model and also to give a hint if the grey box model structure at all was working for this system, (Weyer, 2004).

The predictor for the OE case is given by

$$\hat{y}_{us,d}(t+1)_{(H-C)} = \hat{y}_{us,d(H-C)}(t) + c_1 \left[(g_{u1}(t-\tau) + g_{u2}(t-\tau) \dots + g_{un}(t-\tau)) \sqrt{(y_{us,u}(t-\tau) - y_{ds,u}(t-\tau))} \right] + c_2 \left[(\hat{y}_{us,d(H-C)}(t) - g_{o1}(t))^{3/2} + (\hat{y}_{us,d(H-C)}(t) - g_{o2}(t))^{3/2} \dots + (\hat{y}_{us,d(H-C)}(t) - g_{on}(t))^{3/2} \right] \quad (4.5)$$

The OE model structure makes it possible to simulate the water level for unlimited time, but at the same time it makes the parameter estimation more complex. It is not possible to get an analytical solution, iterative search is required, see Section 2.3.7.

4.2.1 The OE Simulation model of the downstream water level

In both cases, H-C and M-G, the immediate downstream level in the upstream part of the pool is needed to calculate the inflow of water to the pool. It is a paradox to use the OE model for simulation of the downstream end water level if the model needs measured values of the upstream end water level, $\hat{y}_{ds,u}$. This problem can only be solved by building another model, describing the immediate downstream water level at the upstream end. In equation (4.6) below the upstream end water level is replaced by the simulated upstream end water level. For the H-C pool the inflow to the pool is dependent on the immediate downstream water level of H because H is an undershot gate. For the M-G pool the water inflow is dependent on the immediate downstream water level of M if the M gates are in drowned flow.

$$\hat{y}_{us,d}(t+1)_{(H-C)} = \hat{y}_{us,d(H-C)}(t) + c_1 \left[(g_{u1}(t-\tau) + g_{u2}(t-\tau) \dots + g_{un}(t-\tau)) \sqrt{(y_{us,u}(t-\tau) - \hat{y}_{ds,u}(t-\tau))} \right] + c_2 \left[(\hat{y}_{us,d(H-C)}(t) - g_{o1}(t))^{3/2} + (\hat{y}_{us,d}(t) - g_{o2}(t))^{3/2} \dots + (\hat{y}_{us,d(H-C)}(t) - g_{on}(t))^{3/2} \right] \quad (4.6)$$

To simulate the downstream end water level a combination of the two models is needed, one describing the upstream end water level and the other describing the downstream end water level. The water level simulated in one model should be used as a water level in the other model simulating the next value for the first water level. For a descriptive picture see Figure 4.9.

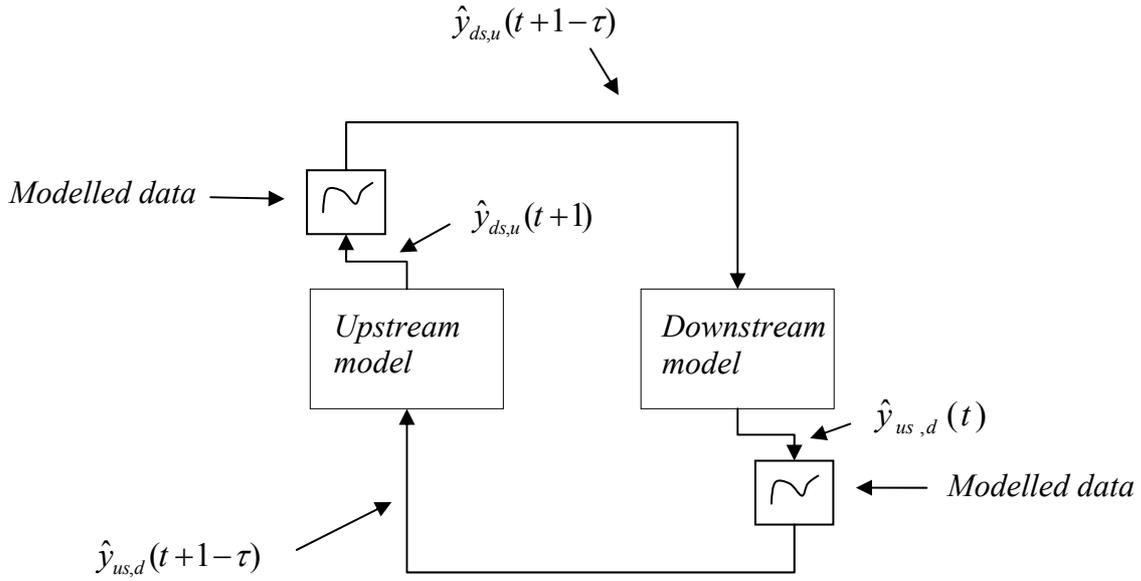


Figure 4.9 A schematic picture of the combined model for describing the downstream end water level. Both the models need the values of the output of the other model after a time delay specific to the models.

The model used for the upstream end water level was built based on the same theory as the downstream end water level model. The downstream and the upstream water level of the pool are both measurements of the volume change in the pool, (which is the theory base of the models, see Section 3.2.2). The difference between the models is that the time delay from inflow in the upstream end water level model is zero and for the outflow it is a time delay as for the inflow in the opposite model.

$$\begin{aligned} \hat{y}_{ds,u}(t+1)_{(H-C)} = & \hat{y}_{ds,u(H-C)}(t) + c_3 \left[(g_{u1}(t) + g_{u2}(t) \dots + g_{un}(t)) \sqrt{(y_{us,u}(t) - \hat{y}_{ds,u(H-C)}(t))} \right] + \\ & c_4 \left[(\hat{y}_{us,d}(t-\tau) - g_{o1}(t-\tau))^{3/2} + (\hat{y}_{us,d}(t-\tau) - g_{o2}(t-\tau))^{3/2} \dots + (\hat{y}_{us,d}(t-\tau) - g_{on}(t-\tau))^{3/2} \right] \end{aligned} \quad (4.7)$$

Equation (4.6) and (4.7) will alternate in calculating the water levels for the two models. The values calculated for one model will directly be put in the other model for a water level that is needed in the first model.

4.3 PARAMETER ESTIMATION AND MODEL VALIDATION

The data set was split into two sets. The first binary signal experiment (BS1) was chosen as the first data set and the second binary signal experiment (BS2) as the second data set. The split was necessary for validation of the model.

Even though it is not much to say about the actual parameter estimation it is a big and time consuming task. The parameter estimation included writing Matlab routines. For the ARX model the analytical solution was calculated, see equation 2.14 in Section 2.3.7, and for the OE model the iterative search for θ was performed. The Matlab command “LSQNONLIN” was used in the search for the optimal solution. “LSQNONLIN” is using the Levenberg

Marquard searching method, see the theory Section 2.3.7. The two models for upstream and downstream end of pool were estimated together. The optimization found the best pair of models that gave the total minimum error for both water levels. The $\hat{\theta}$ vector that was estimated

$$\hat{\theta} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{bmatrix}$$

is the $\hat{\theta}$ that minimizes the total error for both the models, where c_1, c_2 are the parameters for the downstream end water level model and c_3, c_4 are the parameters for the upstream end water level model.

Models were estimated from both data set 1, BS1, and data set two, BS2 to get as much information as possible from the two data sets. It is a way of analysing the model quality, to see if the model can be reproduced by a new set of data. It will also give a hint on the sensitivity of the model. The models were evaluated both against its estimation set and its validation set. As the result of the parameter estimation and the model validation the estimated parameters, the criterion function and plots were given. The plots are showing how well the model can simulate the measured data. The criterion function used in the evaluation was the basic version of the prediction error method.

$$V(\theta) = \frac{1}{N} \sum_{t=1}^N \varepsilon(t, \theta)^2, \quad (4.8)$$

Based on the results an extra validation of the H-C models was made against a step test. This was to distinguish the two models from the swapping of the data and see which one that performs the best

5 RESULTS

Below the results of the system identification are shown. The models of the downstream end water level of H-C are called HCd1 if the model is estimated on the data set BS1, and HCd2 if the data set is estimated on data set BS2. The models of the upstream end water level are called HCu1 and HCu2 where the u stands for the upstream end. M-G has the same labelling logic, MGd1, MGd2 etc. HC1, HC2, MG1 and MG2 are the combined models, with the parameters c_1 - c_4 for data set BS1 and BS2.

5.1 ESTIMATION AND VALIDATION RESULTS IN FIGURES

5.1.1 Estimated parameters Horticultural – Coly 3

In Table 5.1 below the result for pool H-C is shown. cf stands for criterion function, see theory in Section 4.3 equation (4.8). It is the mean squared error between simulation model and measured output for the data set that was used for estimation of the parameters. cfv is the same as cf but it is the mean squared error between simulation and measured output for the validation data set. c_1 , c_2 are the parameters for the HCd model. c_3 , and c_4 are the parameters for the HCu model, see the model structure selection Section in Chapter 4. The first two columns are the results of the identification with all the significant digits given from Matlab. The following two columns are the results when the original results are rounded off to two significant digits. The last column is when the results are rounded off to three significant digits. This was done to see how many significant digits that needs to be considered and still get a good result. For the HC2 case two significant digits was enough, for the HC1 case three significant digits had to be considered.

Table 5.1 The system identification results for the first order OE models, where cf , stands for the criterion function for the estimation set, cfv stands for the criterion function for the validation set, c_1 - c_4 are the estimated parameters, HC1 and HC2 are the models estimated on data set BS1 and data set BS2 respectively. The model structure used was a first order OE model structure.

Model	HC2	HC1	HC2	HC1	HC1
Cf [m ²]	8.185e-5	5.569e-5	6.039e-5	2.14127e-3	5.538e-5
cfv [m ²]	6.347e-5	7.353e-5	7.164e-5	4.00895e-3	9.283e-5
c_1	0.01262494	0.01054406	0.013	0.011	0.0105
c_2	-0.00795972	-0.00660941	-0.0080	-0.0066	-0.00661
c_3	0.01070372	0.00880606	0.011	0.0088	0.00881
c_4	-0.00680996	-0.00556130	-0.0068	-0.0056	-0.00556

The cfv of the H-C model says that the average error on the model is 0.85 and 0.96 cm for the models with fewer significant digits in the parameters (which still gave good results) column 3 and 5 in Table 5.1. As stressed earlier it is not possible to draw any strong conclusions only from the cfv .

The parameters c_1 - c_4 changes about 17% from model built on BS2 to model built on BS1, but the ratios, c_1/c_2 and c_3/c_4 do not change much at all. The ratios changes less than one percent. See Table 5.2

Table 5.2 Ratios for c_1/c_2 and c_3/c_4 for the models HC1 and HC2

Model	HC1	HC2
c_1/c_2	1.586	1.595
c_3/c_4	1.572	1.583

5.1.2 Estimated parameters Morundah – Grants

The labels in Table 5.3 have the same meaning as the labels in Table 5.1. The cfv for the M-G model says that the average error is 1 and 1.3 cm for the models with fewer significant digits in the parameters (which still gave good results) column 3 and 5 in Table 5.3. For the both cases BS1 and BS2, two significant digits were enough to get a result similar to the original. Though for the BS2 case three significant digits gave an even more similar result, see the cfv value in the table.

Table 5.3 Table over the system identification results for the first order OE models, where cf , stands for the criterion function for the estimation set, cfv stands for the criterion function for the validation set, c_1-c_4 are the estimated parameters, MG1 and MG2 are the models estimated on data set BS1 and data set BS2 respectively. The model structure used was a first order output error model structure.

Model	MG2	MG1	MG2	MG1	MG1
$cf [m^2]$	1.1809e-4	1.5309e-4	2.6831e-4	2.1695e-4	1.5297e-4
$cfv [m^2]$	1.8156e-4	1.0721e-4	1.7487e-4	2.5936e-4	1.0755e-4
c_1	0.00349498	0.00358925	0.0035	0.0036	0.00359
c_2	-0.00568037	-0.00592034	-0.0057	-0.0059	-0.00592
c_3	0.00247268	0.00319130	0.0025	0.0032	0.00319
c_4	-0.00394079	-0.00525445	-0.0039	-0.0053	-0.00525

Between the two data sets, BS1, BS2, the parameters c_1 and c_2 do not change much at all, only about 2.7% for c_1 and 4% for c_2 from model MG2 to MG1. For the upstream end water level model the parameters c_3 and c_4 changes much more, about 30% for c_1 and 33% for c_2 .

5.2 RESULTS IN PLOTS

5.2.1 Plots of Horticultural – Coly 3

In Figure 5.1 and 5.2 the simulation plots are represented, from the model validations of the HCd models. It can be seen that the models follow the validation data well for the whole simulation period, 11 and 13 hours. That is longer than for the HMC study which only had 8 hours of validation data. Important to note is that the simulation only uses the first measured values of the water level and the rest is simulated and that the model still give a good prediction after so long period of simulation. In the plots it is shown that it is more a “worst case” when the model differs from data with about 1 cm, other times the model is very well adapted to the data.

In Figure 5.5 and 5.6 the simulation plots for the HCu models can be studied. These plots also show that the models follow the validation data well. After 7-8 hours of simulation a drift on the model starts to become more distinct. The drift is not very big, it is about 2-3 cm when it is worst.

5.2.2 Plots of Morundah – Grants

In Figure 5.3 and 5.4, the plots of the MGd simulation models are represented. For these models the same thing can be said, the simulation follows the validation data well for the whole simulation period. For this case the difference between the simulation and validation data is also about 1 cm in the worst case.

For the MGu model in Figure 5.7 and 5.8 it is a clear difference in model quality compared to the HCd, HCu and MGd models. The model catches the dynamics well, but it is not capable of predicting the water level. It differs about 4-7 cm in the worst case.

Worth noting is that the change in water level during the experiment is between 10-20 cm and the total simulation time is 11 and 13 hours.

5.2.3 Validation of the HC models on a step test

Based on the results, a decision was made to validate the H-C models against a step test. This was done to see which one of the two models that was giving the best results when the gate positions had the same positions for a long time. In Figure 5.9 and 5.10 it can be seen that both the models are performing well. There is a slight difference but it is very small.

5.3 TIME DELAYS OF THE MODEL

Models with different time delays were created. The time delays that were used were values around the time delays estimated from the experiment results. The time delays that were found to give the best models were 10 min for H-C and 20 min for M-G, i.e. 5 and 10 samples respectively in the models. See Table 5.4.

Table 5.4 Values of the criterion function calculated for the validation data for models estimated with different time constants. The high lighted column is the time delays which gave the best results.

Time delay[min]	6	8	10	12	14
$cf_{v_{HC1}} [m^2]$	7.6719e-5	7.558e-5	7.353e-5	1.060e-4	7.9e-5
Time delay[min]	16	18	20	22	24
$cf_{v_{MG1}} [m^2]$	4.999e-4	3.878e-4	1.072e-4	2.802e-4	4.858e-4

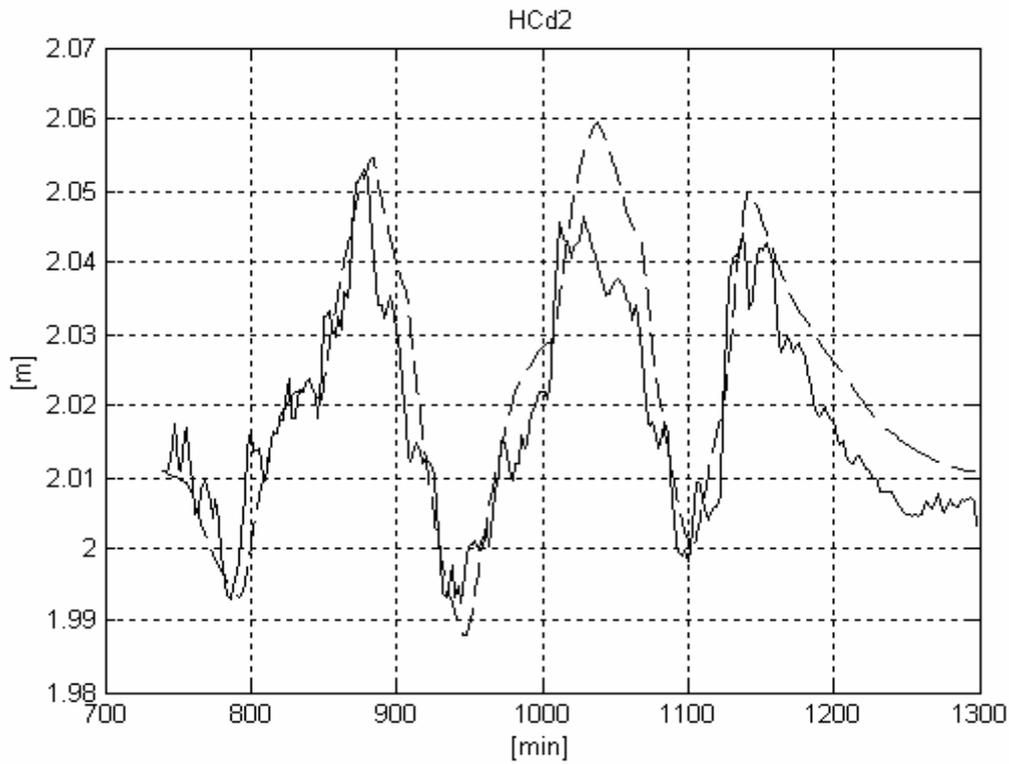


Figure 5.1 The simulated output (dashed) from the model HCd2 compared with the measured output from the validation data set, BS1 (solid). The estimated parameters of model HCd2 was rounded off to two significant digits.

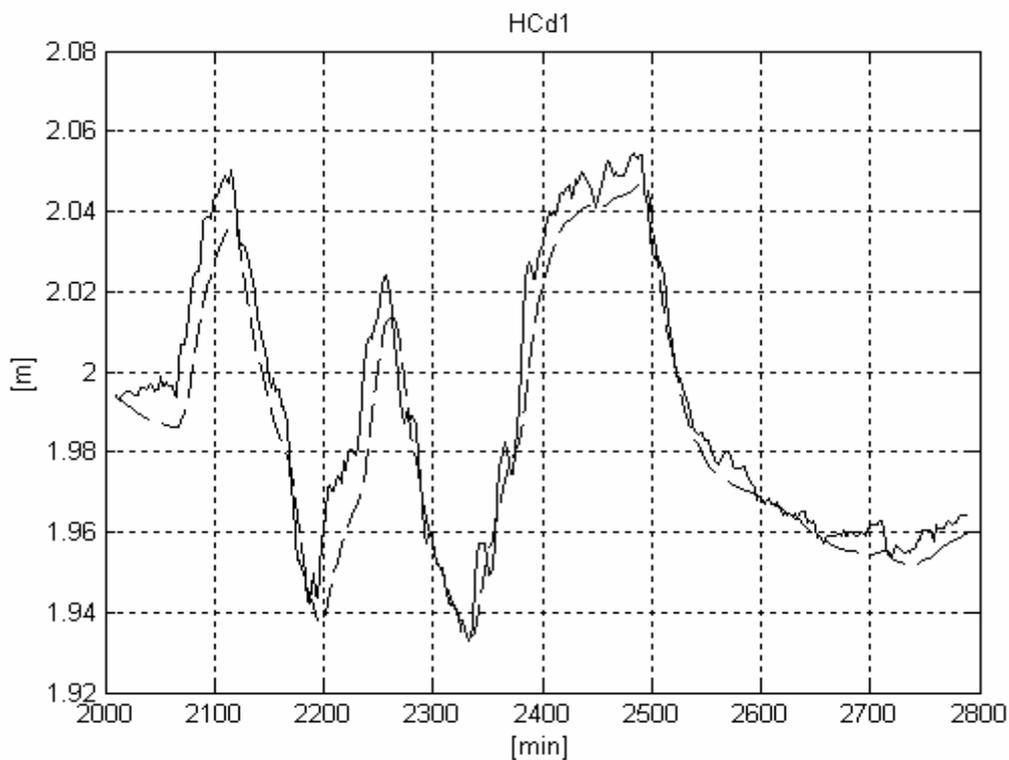


Figure 5.2 The simulated output (dashed) from the model HCd1 compared with the measured output from the validation data set, BS2 (solid). The estimated parameters of model HCd1 was rounded off to three significant digits.

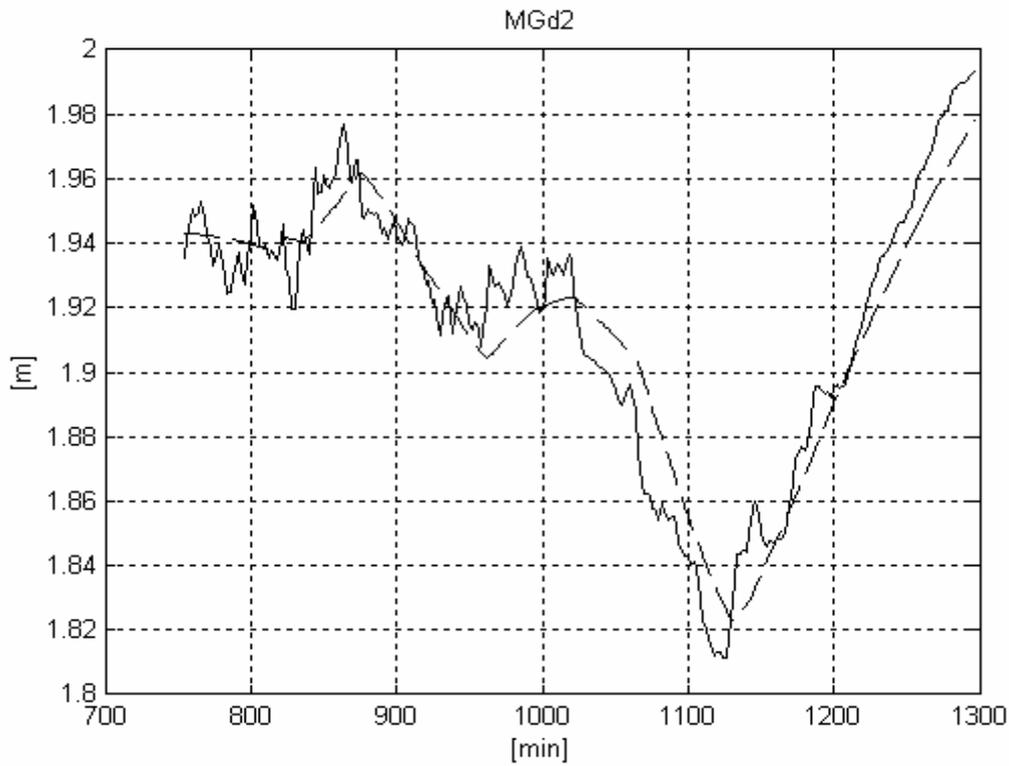


Figure 5.3 The simulated output (dashed) from the model MGd2 compared with the measured output from the validation data set, BS1 (solid). The estimated parameters of model MGd2 was rounded off to two significant digits.

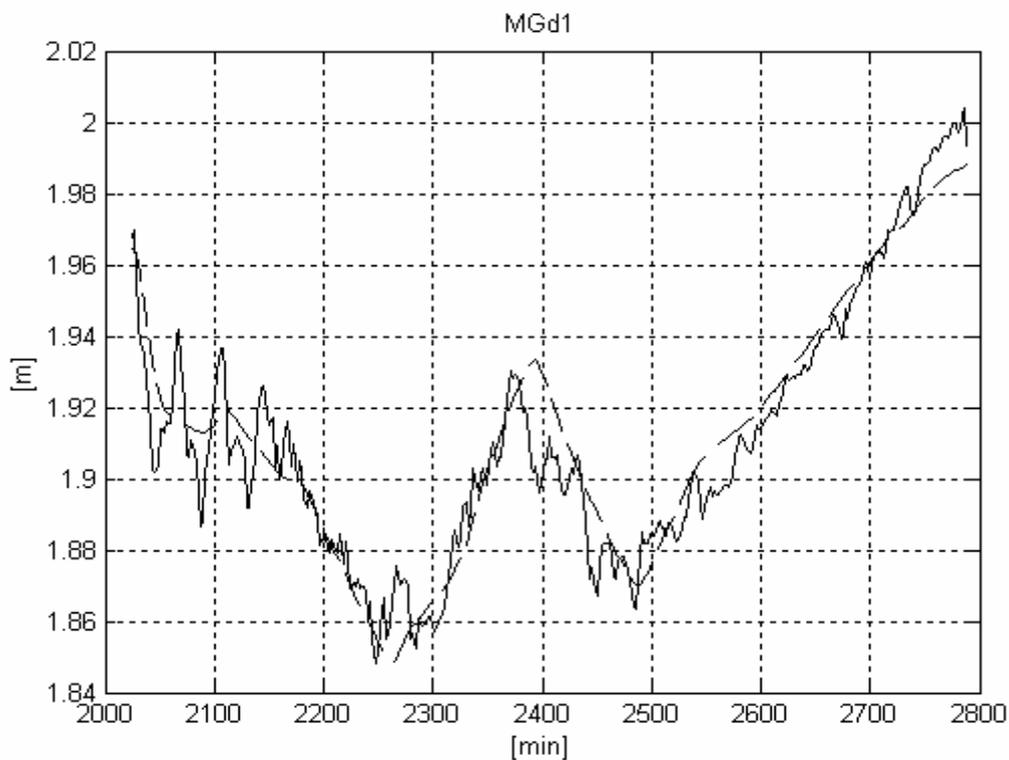


Figure 5.4 The simulated output (dashed) from the model MGd1 compared with the measured output from the validation data, BS2 set (solid). The estimated parameters of model MGd1 was rounded off to three significant digits.

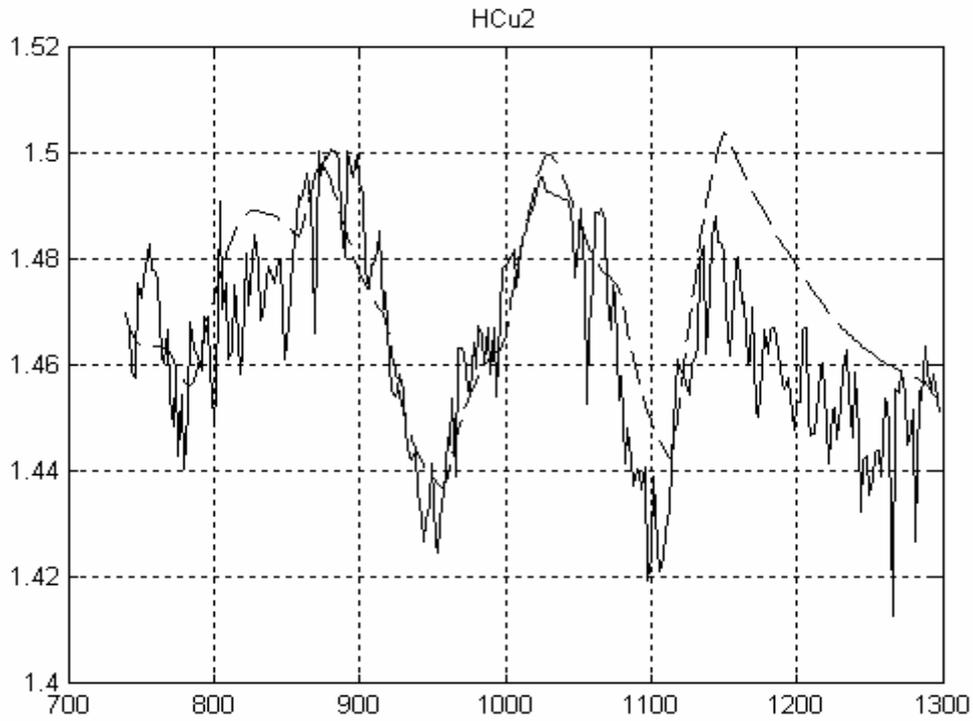


Figure 5.5 The simulated output (dashed) from the model HCu2 compared with the measured output from the validation data set, BS1 (solid). The estimated parameters of model HCu2 was rounded off to two significant digits.



Figure 5.6 The simulated output (dashed) from the model HCu1 compared with the measured output from the validation data set, BS2 (solid). The estimated parameters of model HCu1 was rounded off to three significant digits.

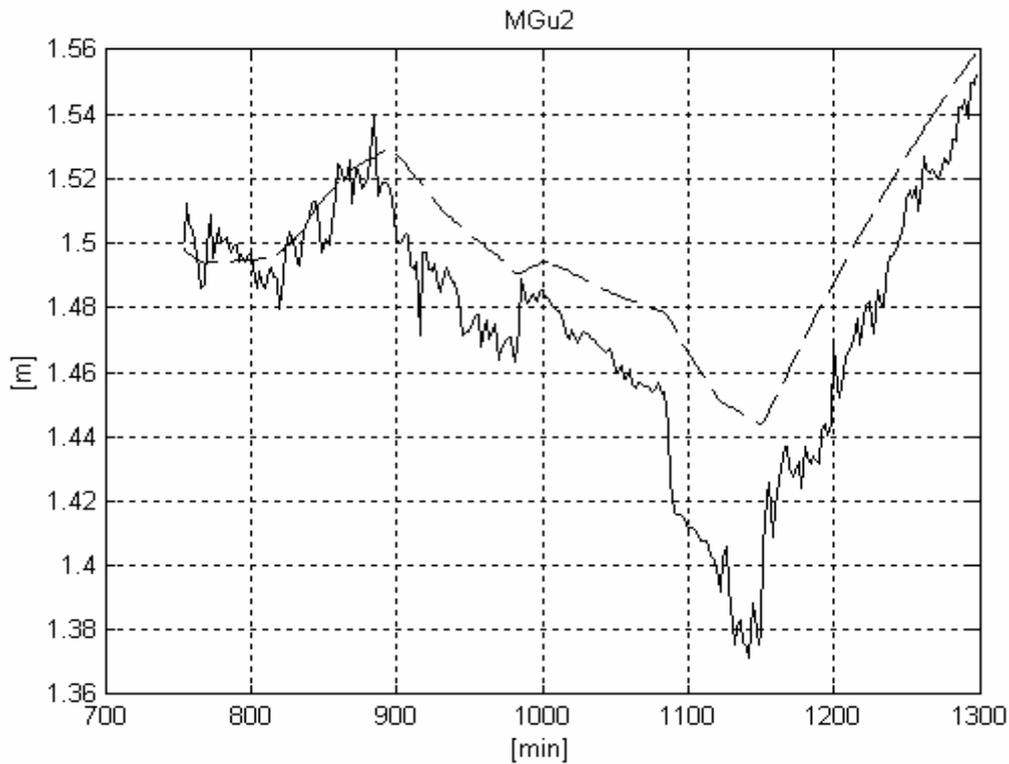


Figure 5.7 The simulated output (dashed) from the model MGu2 compared with the measured output from the validation data set, BS1 (solid). The estimated parameters of model MGu2 was rounded off to three significant digits.

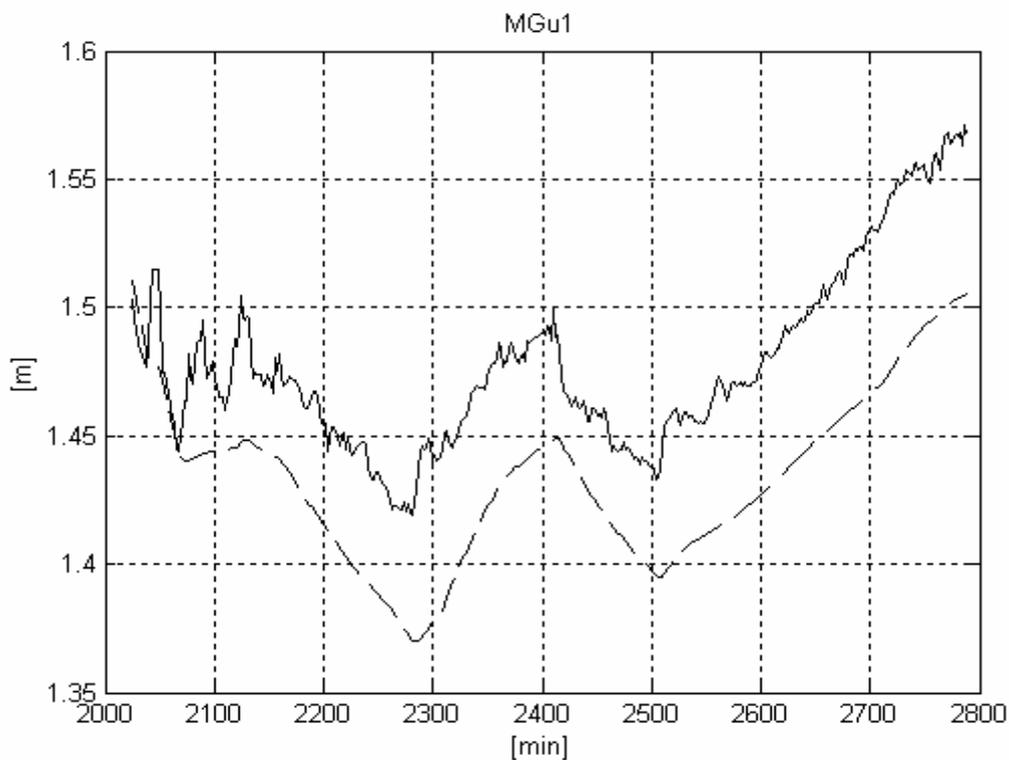


Figure 5.8 The simulated output (dashed) from the model MGu1 compared with the measured output from the validation data set, BS2 (solid). The estimated parameters of model MGu1 was rounded off to three significant digits.

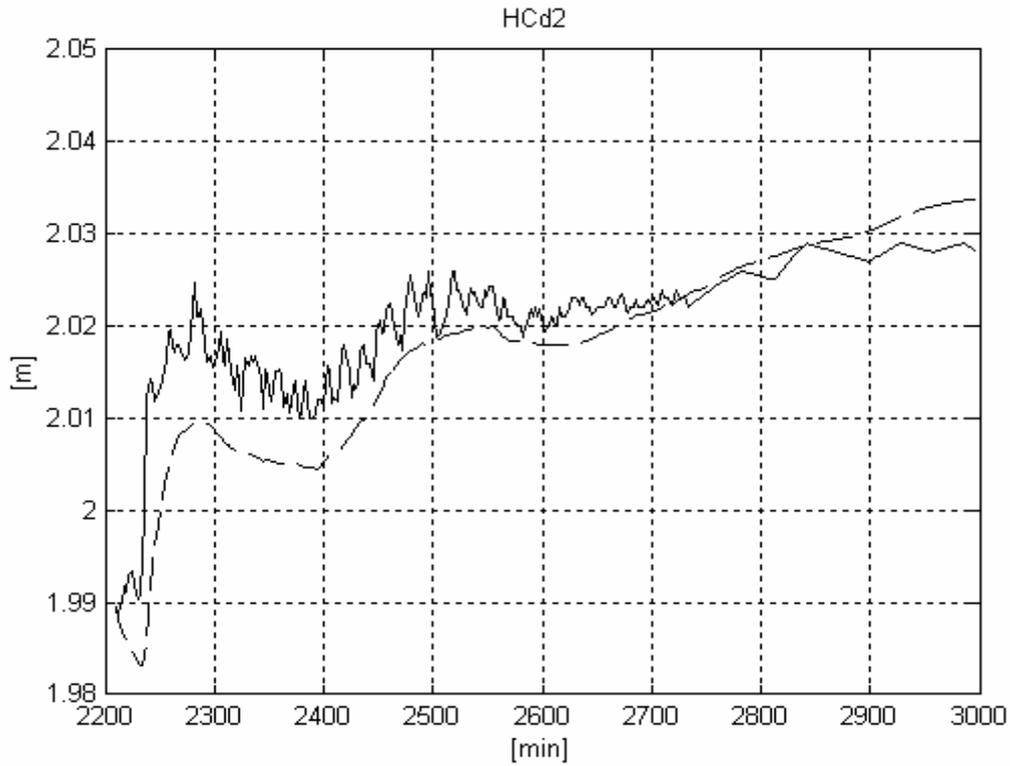


Figure 5.9 The simulated output (dashed) from the model HCd2 compared with the measured output from the validation data set, the step test (solid).

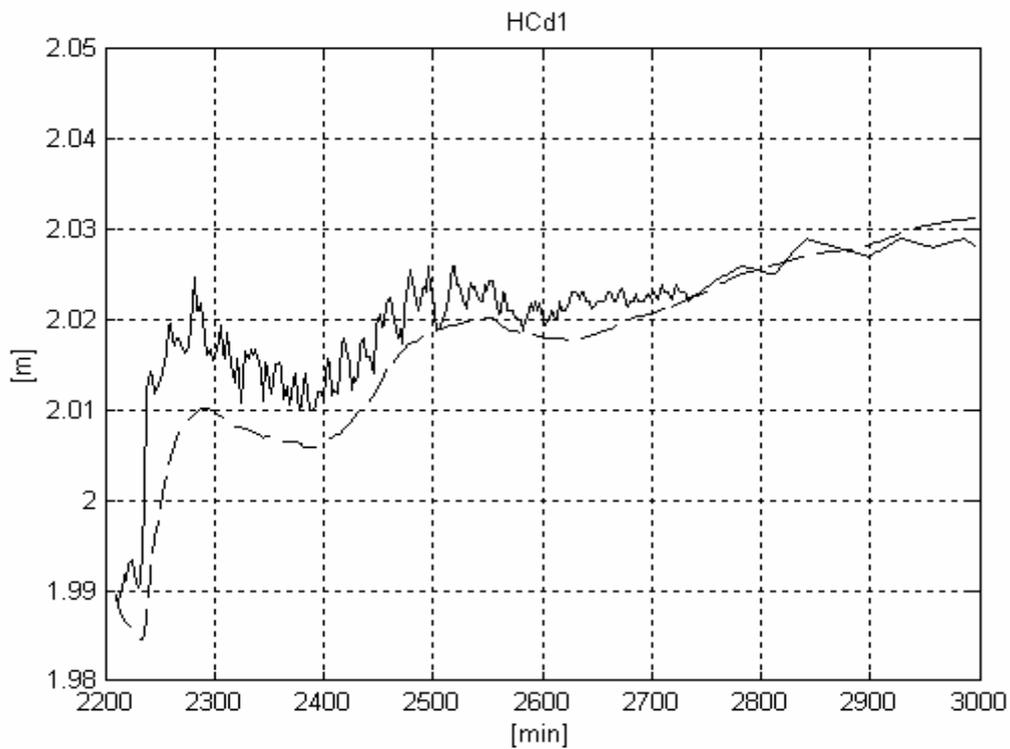


Figure 5.10 The simulated output (dashed) from the model HCd1 compared with the measured output from validation data set, the step test (solid).

6 DISCUSSION

6.1 DISCUSSION OF THE H-C MODEL RESULTS

The two models derived for the pool H-C are performing well. If the models are going to be used as tools for control design, the models are believed to be accurate enough. The quality demand is not very high for a model used for control purpose in a system with slow dynamics. As long as the model captures the main dynamics in the data, the rest can easily be handled by the controller. If the model is meant to be used for simulations it is not easy to say if the model is good enough, it all depends on the demand of simulation accuracy. What still can be said is that the models follow data very well, except for some extreme situations where the models differ from the data about 1 cm. The model is of first order so it can not describe the wave properties of the system. Another point to make is that the models perform very well throughout the whole simulation, 11 and 13 hours. It is up to the user to decide if this model quality is good enough.

The H-Cu models also give good results. They have a drift at the end of the simulation (after 7-8 hours) for both the models. It is difficult to see any strong evidence why this drift occurs. It needs to be mentioned that the upstream end models used the same time delay as the downstream end models. It should have been tested which time delay that was best. If a more suitable time delay had been found, the result of the upstream end model could have been better.

6.1.1 The separation of the two H-C models

For the H-C models there was a considerable difference between the c-parameters depending on what estimation set that was used. The parameters changed with 17 percent from model HCd2 to HCd1. But the ratios c_1/c_2 and c_3/c_4 did not change much at all. That could be the explanation why the two models are both doing so well but still have big differences in the parameters. As long as the ratio between the two parameters is the same, the models have no problem in predicting the water level. That is only valid if the gates are changing often and roughly with the same magnitude. The problem occurs when the gates are opened for longer times, then it is necessary to know the correct parameter values. For example, if the parameters are overestimated, both the inflow and outflow will be overestimated but big inflow will be compensated with a big outflow. If the changes are small and frequent as in the discussion above the problem may not show. The water level can still be simulated correctly.

To be able to separate HC1 and HC2 the models were validated against a step test. This was done to see which one that performed best when the gates openings only are changed once. In the results it was very difficult to separate the two models. Both the models performed well. A slightly better result was given for HC1, see Figure 5.9 and 5.10. The result indicates that the even for gate position changes with low frequency the model is not very sensitive to a change in the parameters as long as the ratios c_1/c_2 and c_3/c_4 are constant. The step test was only of 4 cm, a bigger step might have been able to separate them better. According to model validation theory a model is not valid outside the conditions the model has been validated for. This is very evident in the H-C modelling case. A recommendation could be to use the HC1 model. The recommendation is based on the step test result and the fact that the HC1 model is validated on a data set with broader water level range (about 10 cm) than the HC2 model.

HC1 can be rounded off to two significant digits, HC2 needs to consider three, to show similar result as the original parameters. Having the above discussion about the ratio in mind,

it can be seen that when rounding off HC1 to two significant digits changes the ratio with more than ratio changes for HC2. That can be the explanation why HC1 is more sensitive for rounding off the parameters.

6.2 DISCUSSION OF THE M-G MODEL RESULTS

The quality of the MGd models is also very high. The simulation follows the measured data well, except it makes small misses in time and therefore gives an unfair *cfv* value. The misses in time at the slopes will give big contributions to the *cfv* value as, for example, seen in the plot in Figure 5.3. But it does not necessary mean that the model is bad.

The MGu1 and MGu2 models do not give as good results as the other models. The models capture the main trend but do not give an accurate simulation of the water level. The question is why this model is so bad in comparison with the other models and what does this result mean for the MGd models? The answer can be that the downstream water level of the M gate site does not have that much influence on the inflow to the pool. The information of the influence from the water level is not good enough to estimate a model from. Even though the flow is submerged during the whole experiment for the two moving gates, this is not the case for the four gates (of the six in parallel) that are fixed. It should also be noted (this was also mentioned in Section 6.1) that the same time delay was used for both the upstream end model and the downstream end model. I.e., no particular search for an optimal time delay for the upstream end model was made. This is something that of course should have been done. The MGd models give very good result considering the bad MGu result. How can this be? The simulation of MGd is, as said before, based on the simulation of MGu. The above discussion said that the downstream water level does not have much influence on the inflow. If that is true, a small error in the water level will not make a too big difference in the downstream water level. The upstream water level does not need to be very accurate and still give a good simulation of the downstream water level.

Which model that is best to use, as said before, is of course up to the user to decide, but according to the plots a recommendation could be the MG1 model because it covers a broader range, about 15 cm. The model MG1 is thus valid for a broader range than MG2

In this irrigation channel the downstream water level is the important water level for control. For both the H-C and the M-G pools the downstream water level models give the best results.

6.3 ADDED KNOWLEDGE

The results of this project are believed to be interesting in the sense that it has never before, as far as the author knows, been built a system identification model on an irrigation channel, which:

- has both undershot and overshot gates
- is as long as 5-6 km
- has as much as 6 gates in parallel that operates individually

The study at the HMC only had valid results from channels where there were maximum two gates operated in parallel that always had the same position and that the channel lengths were up to 3 km. The overshot gates in the HMC were only operated in free flow. The new added experience is that the first order OE model can even describe drowned flow well. The validation set of this project is longer than what was used in the HMC study. Approximately 13h compared with approximately 8h for the HMC study (Weyer, 2001).

In this work it has been shown that a first order output error model is well suited to describe the system dynamics for a pool with the above properties.

6.4 SUGGESTIONS FOR FUTURE WORK

A better way of comparing the models would have been to validate the models on a third data set. It would then be easier to say which one of the models that performs better than the other. But a third data set was not available, and the project was done during a limited period.

The result of this project is satisfying. It gave a surprisingly good performance for such a low model order. This is very beneficial because the simpler the model is the less uncertainty and complexity is incorporated in the simulation. Even though the models have such a good performance, they are not able to describe the waves that occur in the pool. With the work on the HMC in mind, a third order model would have been interesting to build. The HMC project gave very good simulations with a third order model. The model captured the waves very well. Because of unpredicted delays in the project there was no time for such an investigation.

The simulation of the pools makes more sense if the model describes the whole main channel. The main channel starts with a reservoir with an approximately constant water level over time. The channel also ends with an approximately constant water level. If there are models in between these water levels that can describe the channel behaviour, there would be a possibility of modelling all the water levels of the whole Coleambally main channel, with starting values of the first point in time and the two water levels at the reservoirs.

For control purposes it is enough to have made the two models for a pool with undershot gates both upstream and downstream of a pool to get an idea of how the pools approximately behave. From the results of this project it can be seen that the time delay and time constant are linear with pool length. It needs to be said that they are not linear with big difference in flow and for big variations in conditions of water level. In the same irrigation system, properties of the system behaviour can be approximated with this knowledge without the need for building a proper model for every channel pool. One thing that still needs to be investigated is how the system behaves when there are undershot gates both upstream and downstream of the pool.

In summary, based on the discussion above, the following topics may be topics for future studies

- compare the models by validation against a third independent data set.
- build a third order model for the pools H-C and M-G.
- building models for a pool where there are undershot gates both upstream and downstream of the pool.
- building models of the whole channel so that the channel can be simulated with only starting values and the water levels of the reservoir and the gate positions for the simulation period.
- designing the control system of the channel.

7 CONCLUSIONS

In this Master thesis project, estimation of models from data with the system identification approach was the applied method. The system object for modelling was an irrigation channel with both undershot and overshoot gates. The H-C pool has undershot gates upstream and overshoot gates downstream. Vice versa for the M-G pool.

The models estimated for the H-C and M-G pools, both first order output error simulation models, describes the system dynamics very well for the validation data. The models can be expected to be well adapted for control purposes, which was the main aim for this project. The models have a good accuracy and may be used for various simulation purposes. It is of course up to the user to decide if the model is good enough for a particular simulation cause.

8 ACKNOWLEDGEMENTS

I would like to thank Dr. Eric Weyer, my supervisor at Melbourne University, for great support, help and expertise during my work. I would also like to thank the people at Rubicon system, especially Paul Jones and Tony Oaks, for letting me perform the experiment on the real system and all the help during the experimentation. Last but not least I want to thank Magnus Holmgren (former class mate) and Professor Bengt Carlsson, Uppsala University, who helped to make it possible for me to do my thesis work at the University of Melbourne.

9 REFERENCES

9.1 LITERATURE REFERENCES

Roberson, J.A., Crowe, C.T., 1993 Engineering fluid mechanics , 5th ed., Houghton Mifflin Company, Boston

Baume, J.P., Malaterre. P.O., Sau, j., 1998, "Modelling and regulation of irrigation canals: existing applications and ongoing researches", IEEE, 0-7803-4778 -1/98, pages 3850-3855

Bos, M.G., 1976, Discharge Measurement Structures, International Institute for Land Reclamation and Improvement/ILRI, Wageningen

Daily, J. W., Harleman, D.R.F., 1966, Fluid Dynamics, Addison -Wesley Publishing Company Inc., Massachusetts

Fox, R. W., McDonald, A.T., 1994, Introduction to Fluid Mechanics (SI version), 4th ed. John Wiley & Sons inc, New York

Ljung, L., 1999, System Identification - Theory for the User, 2nd ed., Prentice Hall PTR, New Jersey

Ljung, L., Glad, T., 1991 Modellbygge och simulering, Studentlitteratur, Lund

Mareels, I., Weyer, E., Ooi, S. K., 2003, "Irrigation Networks: A system Engineering Approach", International water & irrigation, Vol 23 No 4, pages 18-30

Roberson, J.A., Crowe, C.T., 1993 Engineering fluid mechanics , 5th ed., Houghton Mifflin Company, Boston

Rouse, H., 1961, Fluid Mechanics for Hydraulic Engineers, General Publishing Company Ltd, London

Söderström, T., Stoica, P., 1989 System Identification, Prentice Hall, Engelwood Cliffs

Webber, M.B., 1971, Fluid Mechanics for Civil Engineers (S.I. Edition), Chapman and Hall, Ltd, London

Weyer, E., 2001, "System identification of an open water channel", Control Engineering Practice 9, 1289-1299

9.2 INTERNET REFERENCES

1. Melbourne Water,
<http://conservewater.melbournewater.com.au/content/conservedriest.htm>, 2004-04-22
2. <http://www.farmweb.au.com/h2o/h2use.html>, 2004-04-22
3. UNESCO, http://www.unesco.org/water/wwap/targets/facts_and_Figures.pdf , 2004-04-22
4. <http://www.colyirr.com.au/AboutCICL/index.asp>, 2004-04-22

9.3 PERSONAL COMMUNICATION

Weyer, E., 2004

Rubicon System, 2004