Closed-loop Irrigation Scheduling and Control

by

Jannatun Nahar

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

Doctor of Philosophy

in

Process Control

Department of Chemical and Materials Engineering

University of Alberta

Abstract

The increasing population and adverse climate conditions are escalating fresh water scarcity globally. Since irrigation consumes a large portion of fresh water, it is very important to improve irrigation efficiency. One such method on improving irrigation efficiency is to use a closed-loop scheme instead of the traditional open-loop irrigation schemes. There are many challenges in implementing a closed-loop irrigation scheme. These challenges include soil moisture sensing, soil parameter and state estimation based on limited measurements, control-oriented model development, scheduler and controller designs that take into account various constraints in irrigation. In this thesis, rigorous methods are proposed to overcome some of these challenges.

First, a systematic approach based on system observability analysis and state estimation is developed to estimate the soil moisture inside an agro-hydrological system where measurements are not easily available. A discrete-time state-space model based on Richards' equation is used to describe the agro-hydrological system that considers water dynamics in the system and the interaction between the soil, the plant and the atmosphere. The nonlinear agro-hydrological system is linearized every sampling time and the observability of the overall system is determined based on locally linearized models at every sample instant. Based on the linearized models, we investigate how the number and location of output measurements affect the degree of observability of the system. To demonstrate the efficiency of the proposed approach, state estimation is performed using the extended Kalman filter on both simulated and real field data. The parameters of the model are estimated using prediction error method based on historical output measurements.

Next, using the information from estimated states and measurements, we have performed a comparative study between closed-loop and open-loop irrigation scheduling and control. In agriculture irrigation management, irrigation scheduling is typically

performed in an open-loop fashion and is done only once at the beginning of a growing season. In this work, we study whether closed-loop scheduling with closed-loop control can lead to improved performance in terms of crop yield and water conservation in agriculture irrigation. The interaction between the soil water, the crop (maize in this work) and the atmosphere is described by an agro-hydrological model, which is a partial differential equation. In the proposed scheduling and control scheme, both the scheduler and the controller are designed using model predictive control (MPC). The scheduler uses a long time horizon (with a sampling period of one day) that covers the entire crop growth season and the horizon shrinks as time moves. The primary objective of the scheduler is to maximize the crop yield. The controller uses a much shorter prediction horizon and a much finer sampling period. The primary objective of the controller is to track the soil moisture reference calculated by the scheduler. To alleviate the computational complexity of the scheduler and the controller, a linear parameter varying (LPV) model is identified for the scheduler and controller, respectively. The performance of the closed-loop scheduling scheme is evaluated against the traditional open-loop scheduling scheme under different scenarios.

Furthermore, this thesis has extended research on closed-loop irrigation scheduling to a special case where irrigation is performed using storm water to irrigate recreational turfs. In this work, a modeling and scheduling approach for an integrated storm water management and irrigation problem is presented. The primary objective is to simultaneously ensure that the green space is irrigated appropriately and the level of the storm water pond is maintained adequately. It is proposed to use closed-loop irrigation scheduling to achieve this objective. A steady-state model is developed to calculate the soil water storage for different irrigation amounts. To handle the uncertainties, real-time feedback from the pond is used to re-evaluate the scheduling optimization problem every week. Simulation results show that the proposed closed-loop scheduling gives much improved control performance.

Preface

The materials presented in this thesis are part of the research project under the supervision of Dr. Sirish L. Shah and Dr. Jinfeng Liu, which is funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Chapter 3 of this thesis is a revised version of two papers, the first titled, Parameter and state estimation of an agro-hydrological system based on system observability analysis, *Computers & Chemical Engineering*, 121:450-464, 2018 and the second titled, Observability analysis for soil moisture estimation, *Proceedings of Control Conference Africa (CCA)*, pages 110-114, Johannesburg, South Africa, 2017. For both papers, I performed the manuscript composition under the supervision of Professor Jinfeng Liu and Professor Sirish L. Shah.

Chapter 4 of this thesis is a revised version of a recently completed paper titled, Closed-loop schedule and control for precision irrigation, *Industrial & Engineering Chemistry Research*. I was responsible for the overall manuscript composition under the supervision of Professor Jinfeng Liu and Professor Sirish L. Shah. Yawen Mao assisted with the LPV model identification and Su Liu assisted with the optimization problem formulation.

Chapter 5 of this thesis is a revised version of a paper that is currently under review and titled as, Improved storm water management through irrigation rescheduling for city parks, *Control Engineering Practice*. I was responsible for the overall manuscript composition under the supervision of Professor Jinfeng Liu and Professor Sirish L. Shah. Su Liu assisted with the optimization problem formulation.

Acknowledgements

Foremost, I acknowledge my Lord Allah (SWT), who has blessed my life with great parents, sibling, spouse, child, in-laws and academic mentors. Without these blessings, I would not be able to achieve what I have achieved today.

I would like to express my sincere gratitude and appreciation to my academic supervisors Dr. Sirish L. Shah and Dr. Jinfeng Liu who have guided me with advice, knowledge and kindness, not only to be successful academically but also as a parent, *a mother*. The patience that Dr. Liu showed while I came back from maternity leave and was struggling to get back on course is exemplary. Without his belief in me and pushing me to accomplish my dream, this work would not have been possible. I am in debt to the kindness that my supervisors have shown me.

I would also like to thank all my group members from Process Systems and Control Engineering (PSACE) group, specifically Su Liu, Soumya Sahoo, Yawen Mao, Benjamin Decardi-Nelson (our IT man!), Song Bo and Tianrui An for their support and valuable opinions in my research.

My special thanks to Hope Moore for teaching me how to create illustration using Adobe and helping me to make figures such as Figure 1.4. I would also like to thank a few people in this cold city of Edmonton, who have brought warmth in my life with their friendship and support, Sahina, Samia and their families. I am grateful to my cousins Shaheen and Lucky who have been great friends, teachers and inspiration. My thanks to my uncles, aunts (specially Juthoni, Gullah) and cousins for their love and support. One thing that I realized while growing up is that relations are a function of time that may evolve or deteriorate, but the important aspect is how we learn from it. Every positive and negative events in our life shapes our way of life. I am really grateful that whatever good and bitterness I have experienced in my life shaped my path towards something good.

I must also thank my friends Hasin, Samira and Jannat for being there when I needed them most. I am extremely thankful to Shams, my brother for his constant poking and love and my father Musa for his wise guidance. I am also grateful to my mother-in-law, Zohora for helping me to continue my study.

Finally, I would like to mention two people Setaria Khanom (my mother) and Arifin (spouse), thank you for being in my life. I owe it all to you.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

The basic needs of humans for survival are food and water, where food production through agriculture also depends on water. Globally, the water consumption rate has grown to be more than twice the rate of population increase in the 20th century. As population growth continues, 60% more food will be needed to satisfy the demand of more than 9 billion people worldwide by 2050 [4], which will lead to great water stress on the available fresh water resources. It is clear that policies related to water conservation and new technologies for more efficient water consumption need to be developed; otherwise, water scarcity will become a global issue in the near future.

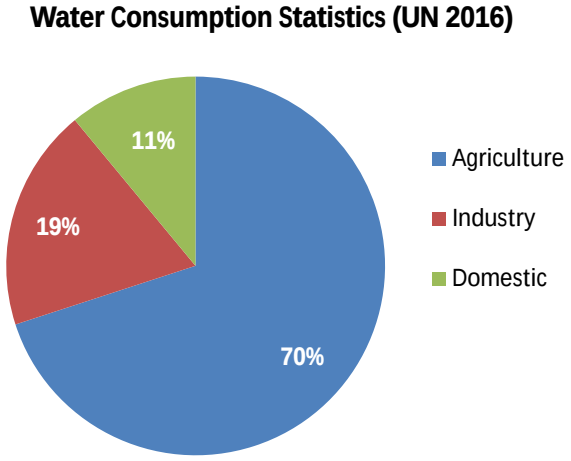


Figure 1.1: Worldwide water consumption statistics in 2016.

According to the statistics of the United Nations, agriculture consumes about 70% of the global fresh water [5]. It is clear that even an incremental improvement

in agriculture water usage efficiency may result in a substantial saving of fresh water. According to Alberta Irrigation [6], in the province of Alberta even a saving as small as 1% irrigated water would result in a saving of about 23 million cubic metres of water each year. The Food and Agriculture Organization (FAO) of the United Nations identified three aspects in improving agriculture water management [4]:

- Reduce water losses
- Increase water productivity
- Water reallocation

The first option is to increase water use efficiency by reducing water losses. The second option is to increase crop productivity with respect to water usage and the third option is to reallocate water to higher value applications. Among these three options, the most popular and effective one is the first option.

Aside from agriculture, a significant amount of water is used for maintaining lawns, parks, and sport fields. In Cristina et al. [7], it is reported that the water use for these outdoor green spaces ranges between 22% and 38% of the total residential water consumption in cool climate regions and between 59% and 67% in hot climate regions across North American cities. The total lawn area in US is about 1.5% of the overall country size, which is significant. Therefore, improving the irrigation efficiency for lawns, parks and sport fields is also important. The above considerations motivate this thesis.

1.2 Background

1.2.1 Current irrigation practice

Irrigation science dates back to the early history when people started to grow crops. While modern irrigation technologies have been adopted in many developed countries, traditional irrigation methods are still widely used in many parts of the world. Broadly speaking, the current irrigation methods can be classified into two categories: traditional and modern methods. The traditional irrigation methods mainly involve surface irrigation, where irrigation water is either flooded over the crop field (flooding

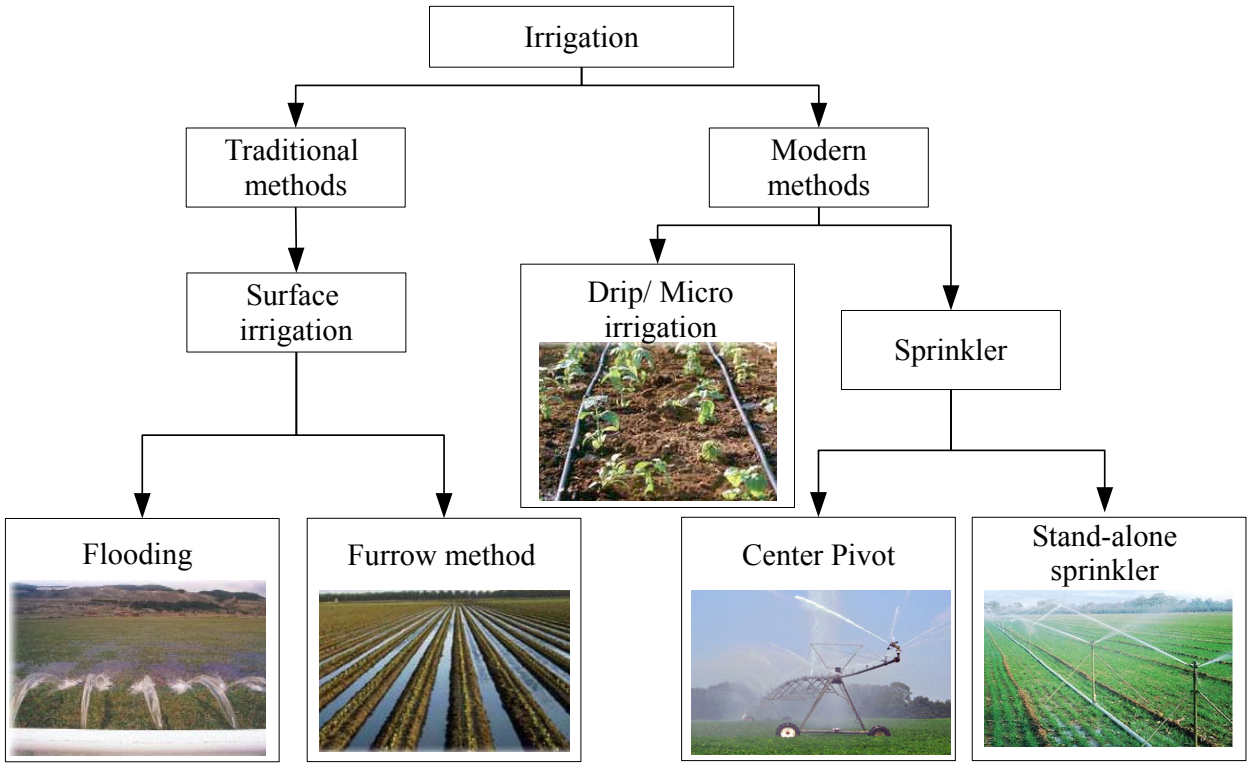


Figure 1.2: Classification of irrigation methods [1].

method) or water is passed through furrows (furrow method). On the other hand, modern irrigation methods can be further classified into two types. One is the drip or micro-irrigation where water is irrigated close to crop roots. The other one is the sprinkler-based irrigation where the irrigation water is sprayed over the crops through sprinklers. A sprinkler may be installed as a stand-alone unit or be mounted on a rotating pivot (center pivot). Figure 1.2 shows the classification of different irrigation methods. In Canada, the two most popular irrigation methods are central pivots for crops (around 93% in 2014) and drip irrigation for fruits (around 65% in 2014) [8]. Sports fields, lawns and parks are mainly irrigated using stand-alone sprinkler systems.

In current practice, the amount of water to be distributed to a farm is allocated according to the farmer's request which is determined primarily based on the farmer's experience. From a systems engineering's perspective, this kind of irrigation management, where irrigation decisions are primarily based on the empirical or heuristic knowledge of the farmer and no real-time feedback information (e.g., soil moisture)

from the field is used, is called open-loop irrigation.

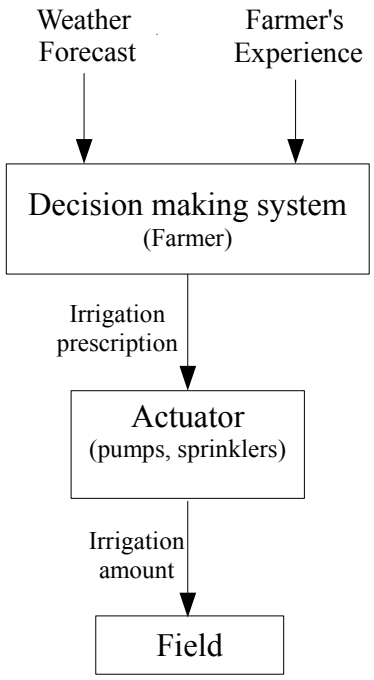


Figure 1.3: Open-loop irrigation.

Figure 1.3 shows the information flow of the open-loop irrigation strategy. The open-loop irrigation strategy often leads to either over irrigation or less irrigation.

1.2.2 A brief review of the agro-hydrological system

An agro-hydrological system characterizes the hydrological cycle between the soil, the atmosphere and the crop. Figure 1.4 gives an illustration of an agro-hydrological system. In the agro-hydrological system, water transportation takes place by means of rain, drainage, evaporation, root water extraction and irrigation. The roots of the crop act as water sinks that extract water from the soil. At the surface when their is rain or irrigation water may enter the soil depending on the soil moisture condition. When the soil is unsaturated, water infiltrates the soil; however, when the soil is saturated, the infiltration tends to cease and ponding starts to occur. After a certain ponding height, the water level breaks and runoff starts to occur. It is to be noted that some of the rain or irrigated water may not reach the soil surface and be intercepted by the crop canopy. Evaporation occurs at the bare soil surface and the

crop canopy. The evaporation and transpiration is usually calculated together for a cropped field and termed as evapo-transpiration. Since we mainly consider the water dynamics in the root zone inside soil, our work is restricted to the vadose zone. The vadose zone is the soil profile above the water table.

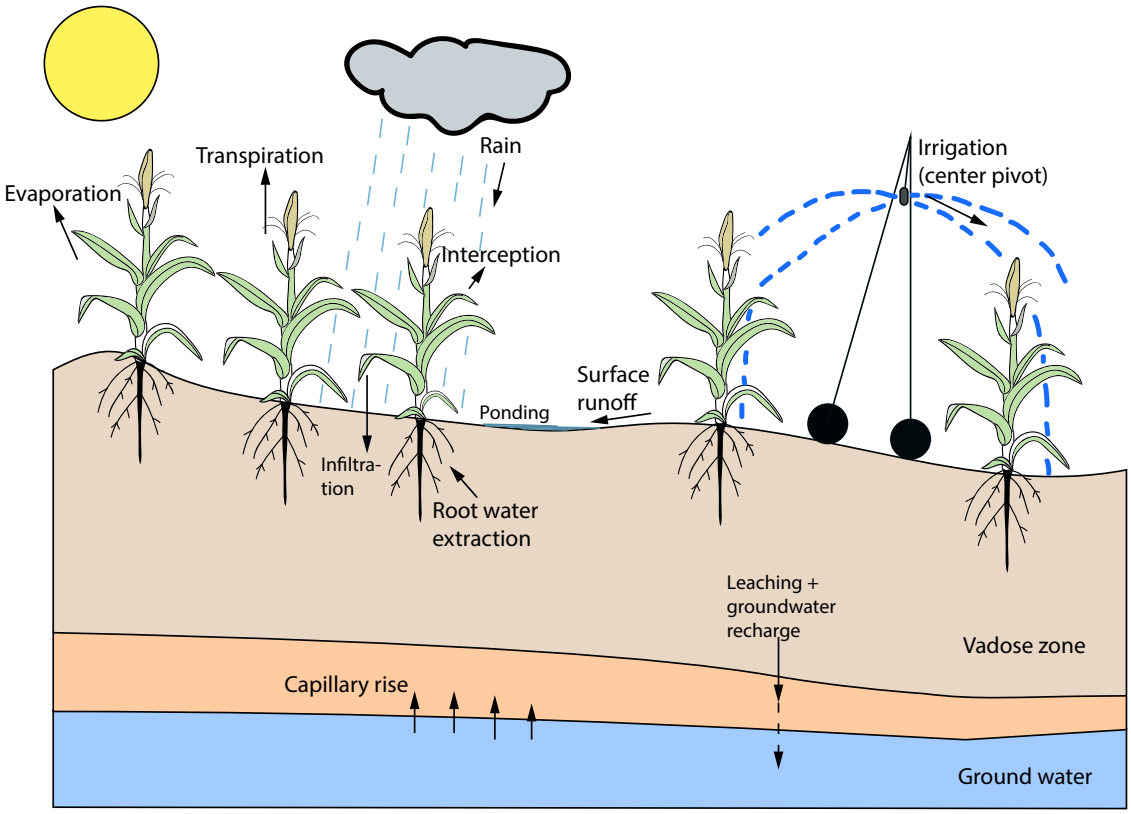


Figure 1.4: An agro-hydrological system.

The modeling of an agro-hydrological system will be described in Chapter 2.

1.3 Literature review

1.3.1 Soil moisture estimation and sensor placement

One major barrier in developing a closed-loop irrigation control system is the availability of proper soil moisture data. It is not feasible to install sensors at all depths or at every location in the farm. An alternative is to estimate the soil moisture using state estimation techniques. The estimation of soil moisture profile has been considered in many studies as a part of developing advanced climatology and weather

forecasting models [9]. Different algorithms for soil moisture estimation with synthetic data [10, 11, 12], remote sensing data [13, 14, 15, 16, 17, 18] and ground data [19, 9, 20, 21] were developed. Specifically, a simplified linear version of the Darcy-Buckingham equation was used to develop a Kalman filter in Walker et al. [22]. While the linear Kalman filter is easy to implement and works in certain conditions, the use of a simplified linear model limits its applicability in extreme weather conditions. In Reichle et al., a comparative analysis between the extended Kalman filter (EKF) and ensemble Kalman filter (EnKF) using synthetic remote sensing data was performed for soil moisture estimation [11]. The requirement of the Jacobian matrix calculation during linearization makes the EKF computationally expensive which can be avoided in the EnKF; but for good estimation performance, the EnKF may require a large number of ensembles. While the performance of both filters were satisfactory, the authors of the above paper suggested EnKF for soil moisture estimation due to its flexibility and robustness. Some studies used iterative EnKF algorithm to better handle nonlinear systems and to enforce constrains [23] and to improve the computational efficiency [24]. Since Kalman based filters assume Gaussian distribution for sensor noise and disturbances, the filter performance may drop when non-Gaussian distribution is present. For EnKF, this problem was addressed in the work of Erdal et al., where it was pointed out that the relation between pressure head and soil moisture in extreme weather conditions is non-unique so the ensemble mean may be affected by some extreme ensemble values and therefore give erroneous results [25]. In their work, they suggested the incorporation of a damping term using a normal score transform where the non-Gaussian variable is transferred into a Gaussian one. The use of particle filter to handle non-Gaussian distribution in soil data assimilation was also discussed in a few studies [26, 27, 28]. While this approach does not depend on the Gaussian distribution assumption, it may have difficulties in updating if the realizations are far from the true states [28].

The presence of bias in the soil moisture estimation was also considered in some studies [9, 29]. In [9], bias estimation was performed along with regular soil moisture estimation using EnKF. The approach was shown to have improved overall performance. However, it was also concluded that bias estimation in layers where no observations were available is impossible. In another study [30], it pointed out that

the persistent bias in the Kalman filter estimates evolves from non-Gaussian errors caused by incorrect soil hydraulic parameters. This confirms the importance of accurate soil parameters in soil moisture retrieval and signifies parameter estimation as an unavoidable part of state estimation. While in [30] dual parameter and state estimation were performed using EKF, the approach failed to give accurate soil moisture estimation over the entire profile due to the consideration of homogeneity in parameter estimation. In [31, 32], the dual parameter and state estimation approach was extended to consider soil profile heterogeneity. However, it is still unclear how the number of states or augmented parameters affects the estimation performance and what the minimum number of measurements is in order to have reliable state estimation. In [33], sensor failure analysis was performed where they discussed the impacts of the number of sensors or measurements and their placement in state estimation. Though this study aimed to determine the number of measurements required for a reliable observation of the system, they did not perform observability analysis of the system.

The above consideration motivates the work of Chapter 3 on the development of a systematic approach to soil moisture data assimilation based on system observability analysis.

1.3.2 Closed-loop irrigation for agricultural system

The incorporation of feedback information (e.g., soil moisture content) in the irrigation decision making can lead to significant water conservation. Different control algorithms based on real-time field feedback such as soil moisture measurement have been reported in the literature. For example, it was reported in the work of Belayneh et al. [34] that the closed-loop irrigation can save 63% water compared to the water irrigated by an experienced manager for a commercial pot-in-pot nursery, which corresponds to an annual saving of $5263. The study by Gutierrez et al. [35] showed 90% water saving compared to traditional approaches for a wireless distributed network using on-off controller. Saavoss et al. [36] reported a revenue increase of 62% per year for a sensor-based irrigation. Goodchild et al. proposed a modified proportional-integral-derivative (PID) control to improve the irrigation precision [37]. In Bahat et al., an irrigation controller based on the fuzzy-logic methodology was presented to

decide on how far to open the water valve and how much water to be added to the soil, by considering the temperature, air humidity, wind speed and water budget as the fuzzy variables [38]. In Cid-Garcia et al., a new methodology based on mathematical models of linear programming and site-specic management zones was presented to determine the crop pattern and the use of water in agricultural fields considering the crop requirements in real-time [39]. Kim et al. developed a closed-loop irrigation control system for a self-propelled lateral-move sprinkler irrigation system [40]. Besides the above results, closed-loop irrigation were also reported in Pawlowski et al. and Navarro-Helln et al. [41, 42].

Model predictive control (MPC) has also been investigated in the optimal control of agriculture irrigation. MPC is a very flexible optimal control framework based on solving constrained optimal control problem on-line repeatedly, and has been widely used in modern manufacturing industries due to its ability to handle multivariate processes and to address state and input constraints [43, 44]. In the MPC based studies, due to the use of a prediction model, it is possible to incorporate weather forecast along with other environmental and crop factors. Hence the irrigation amount can be controlled accurately without hampering crop yields. Park et al. used MPC to maintain the field at desired soil moisture and salt levels [45]. McCarthy et al. used crop production models in MPC to optimize irrigation time and amount [46]. Delgoda et al. implemented MPC to minimize root zone soil moisture deficit and irrigation amount with a limit on water supply [47]. Recently, Mao et al. developed a zone MPC algorithm to maintain soil moisture within a target zone instead of a set-point. This approach was shown to lead to further water conservation [48]. The above mentioned irrigation control studies all focus on short time (hourly or daily) irrigation management. A typical objective of these studies is to maintain the soil moisture or root zone soil moisture deficit at a pre-determined set-point or zone.

Irrigation scheduling is another important subject in agriculture irrigation management, which considers water management over a much longer time period (e.g., the entire growing season of a crop). In general, a water balance accounting for plant water consumption, evaporation and drainage loss, historical weather data and avaiable soil moisture is used in irrigation scheduling [49, 50, 51]. Different objectives were considerd in irrigation scheduling including, for example, profit maximization

[52], crop yield maximization [53] and irrigation uniformity optimization [54]. One common feature of the above scheduling studies are that the scheduling optimization problem is only solved once at the beginning of the crop growing season and the solution is used for the entire season.

In Chapter 4 of this thesis, we will introduce feedback into irrigation scheduling and integrate it with irrigation control to form a hierarchical closed-loop scheduling and control scheme. It is well recoganized in chemical process industries that the closed-loop scheduling and control can bring improved economic performance [55, 56, 57, 58, 59]. The primary objective of Chapter 4 is to study the closed-loop scheduling and control in agriculture irrigation management and investigate whether closed-loop scheduling brings any benefits to irrigation management.

1.3.3 Irrigation for recreational turf using storm water

Increased urbanization and reduced natural vegetation have led to significantly increased stormwater volumes. Stormwater is surface runoff that flows over roofs, driveways and roads rather than infiltrating into the ground during any rain or snow melt events [60]. Stormwater is typically collected using city drainage system and is either stored in stormwater ponds or discharged to downstream water bodies. In the literature, there are studies on improving stormwater quality [61, 62, 63, 64] and reducing storm water volumes by using permeabale pavements or contruction materials [65, 66]. The reuse of stormwater in irrigating city green space such as parks, sport fields and green roofs is advocated for sustainability and has been considered in different studies [67, 68, 69, 70, 71]. When stormwater is used for irrigating neighboring parks or sport fields, the stormwater is managed near to its source which reduces the chance of flooding in downstream areas. To increase the efficiency of the stormwater pond, overflow and dry-out events need to be reduced. Thus instead of using tradition open-loop scheduling, a closed-loop scheduling problem is studied to understand the interplay between storm water pond level and irrigation scheduling in Chapter 5.

1.4 Thesis contributions and outline

In this thesis, we consider closed-loop irrigation scheduling and control as illustrated in Figure 1.5.

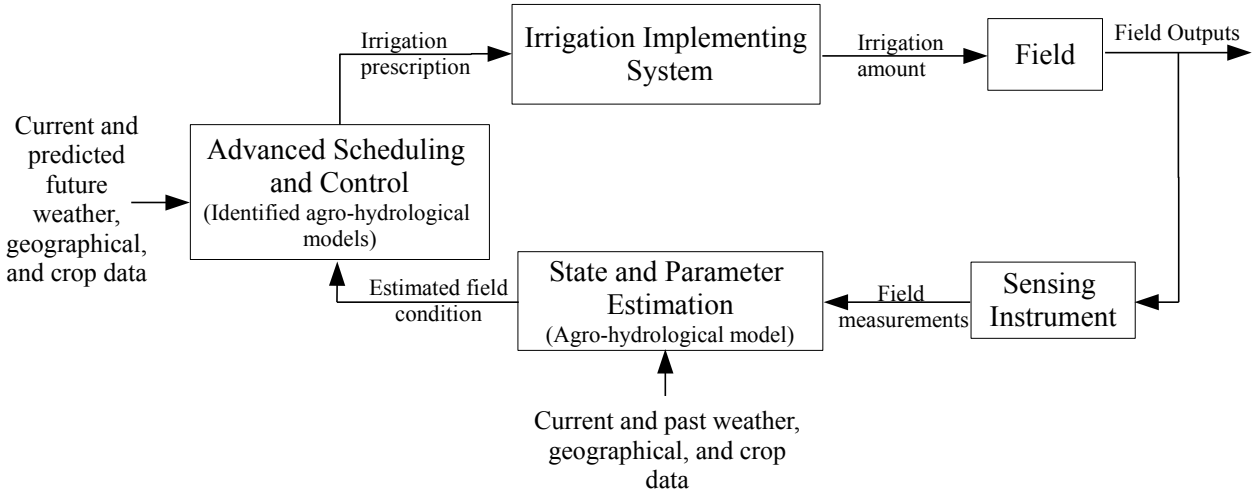


Figure 1.5: Closed-loop irrigation.

The rest of this thesis are organized as follows:

Chapter 2 describes an agro-hydrological model and the relevant preliminaries.

In Chapter 3, we propose to use the degree of observability analysis to determine the minimum number of sensors and to identify the optimal sensor location for state estimation. The main contributions of this chapter include:

- State and parameter estimation for an agro-hydrological system based on a thorough degree of observability analysis;
- An optimal sensor placement algorithm based on the degree of observability of the system.

In Chapter 4, we develop a closed-loop scheduling and control strategy for a crop field and explore whether this scheme results in water conservation. We have also developed a Linear Parameter Varying (LPV) model to represent the agro-hydrological system, which is computationally more efficient and may be used in the design of MPC. The main contributions of this chapter include:

- An LPV model identification method for the agro-hydrological system;

- Extensive simulations comparing the performance of the closed-loop irrigation scheduling with the traditional open-loop scheduling.

In Chapter 5, we explore a special case of irrigation scheduling. In this case, storm water is used to irrigate city park green space. The main contributions of this chapter include:

- The integrated control problem formulation and the corresponding closed-loop scheduler design;
- Extensive simulations illustrating the effectiveness of the proposed approach.

Finally, Chapter 6 ends with concluding remarks followed by proposal for a few future research directions.

Chapter 2

Preliminaries on modeling of agro-hydrological systems

2.1 Richards' equation and its discretization

The dynamics of water flow in saturated media can be described by the Darcy's law, which was developed by Henry Darcy about 1856. In Darcy's law, the flow flux is proportional to the pressure gradient and the proportional constant is the hydraulic conductivity. In 1907, Edgar Buckingham discovered that the water flux in unsaturated media is proportional to the gradient of the capillary potential in the media [72]. It was found later that the equation developed by Buckingham was a generalized version of the Darcy's law which was applicable to both saturated and unsaturated media. These two achievements in soil physics gave birth to the Richards' equation developed by L. A. Richards in 1931 [73]. Richards' equation is widely used today to describe the water flow in porous media. In Richards' equation, the capillary flow of liquids in porous media is characterized in terms of the impacts of the gravity and the pressure gradient.

2.1.1 Richards' equations

1D Richards' equation is expressed as follows:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z}[K(h)(\frac{\partial h}{\partial z} + 1)] - S(h) \tag{2.1}$$

where θ (cm^3/cm^3) is the moisture content, z (cm) is the vertical distance (positive upward), h (cm) is a pressure component that is known as the matric potential in

unsaturated medium and simply as pressure head for saturated medium [74], S (hr^{-1}) includes the source and sink terms such as soil water extraction rate by crop roots, and K (cm/hr) is the hydraulic conductivity. Note that the consideration of 3D Richards' equation would improve the overall simulation of the agricultural land but it also increases computational complexity. Thus for simplicity we considered 1D model with runoff and run on events at the soil surface.

The relation between the matric potential h and the soil moisture content θ is described by the Van Genuchten model [75] as shown in equation (2.2) below:

$$\theta(h) = \theta_{res} + (\theta_{sat} - \theta_{res})(1 + |\alpha h|^{\eta})^{-m} \tag{2.2}$$

where θ_{res} (cm/cm) is the residual moisture content, θ_{sat} (cm/cm) is the saturated moisture content, η and m are empirical shape factors, and α (cm^{-1}) is the inverse of air entry suction. In general, m is considered as a dependent variable of η such as $m = 1 - \frac{1}{\eta}$. Thus Richards' equation can be also written in terms of the matric potential, which is well known as $h-$based Richards' equation as shown below:

$$C(h)\frac{\partial h}{\partial t} = \frac{\partial}{\partial z}[K(h)(\frac{\partial h}{\partial z} + 1)] - S(h) \tag{2.3}$$

where $C(h) = \partial\theta/\partial h$ (cm^{-1}) is the specific moisture capacity also known as the hydraulic capacity.

The relation relating the hydraulic conductivity with the soil moisture content [76] is shown in equation (2.4).

$$K = K_{sat} S_e^{\lambda}[1 - (1 - S_e^{\frac{1}{m}})^m]^2 \tag{2.4}$$

where K_{sat} is the saturated hydraulic conductivity, $\lambda = 0.5$ is the shape parameter [77] and S_e is the relative saturation given by

$$S_e = \frac{\theta - \theta_{res}}{\theta_{sat} - \theta_{res}}. \tag{2.5}$$

2.1.2 Model discretization

Richards' equation is a nonlinear partial differential equation, which in general is difficult to handle analytically [78]. In Van Dam et al. [79], an implicit backward, finite difference method with explicit linearization of hydraulic conductivities was

developed to discretize Richards' equation. In our work, the same discretization scheme is used. We consider time $t \in [0, T]$ and the length from the top $z \in [0, L]$. The time is discretized with P nodes ($k = 1, 2, ..., P$) and the length is discretized with N unequally spaced nodes ($i = 1, 2, ..., N$). Note that an adaptive time scheme is considered in the time step discretization where the time steps can be reduced to much smaller steps if certain tolerance on the gradient is reached. The subscript $._i$ denotes a variable at discrete length node i. The discretized equation at node i is shown below:

$$C_i(h_i(k))\frac{h_i(k+1)-h_i(k)}{\triangle t} = \frac{1}{\triangle z_i}\Bigg(K_{i-1/2}(h(k))\frac{h_{i-1}(k+1)-h_i(k+1)}{\frac{1}{2}(\Delta z_{i-1}+\Delta z_i)} + K_{i-1/2}(h(k)) - K_{i+1/2}(h(k))\frac{h_i(k+1)-h_{i+1}(k+1)}{\frac{1}{2}(\Delta z_i+\Delta z_{i+1})} - K_{i+1/2}(h(k))\Bigg) - S_i(k) \tag{2.6}$$

where Δt is the time interval between two consecutive time instants, Δz_i is the thickness of compartment i and $K_{i-1/2}(h(k)) = \frac{1}{2}(K_{i-1}(h_{i-1}(k)) + K_i(h_i(k)))$, and $h = [h_1\ h_2\ ...\ h_N]^T$ is the entire state vector. Equation (2.6) can be rearranged into the following equation:

$$d_i(h(k))h_{i-1}(k+1) + b_i(h(k))h_i(k+1) + a_i(h(k))h_{i+1}(k+1) = e_i(h(k)) - u_i(k) \tag{2.7}$$

where $a_i(h) = -\frac{K_{i+1/2}(h)}{\frac{1}{2}(\Delta z_i+\Delta z_{i+1})}$, $d_i(h) = -\frac{K_{i-1/2}(h)}{\frac{1}{2}(\Delta z_{i-1}+\Delta z_i)}$, $b_i(h) = C_i(h_i)\frac{\Delta z_i}{\Delta t} - a_i(h) - d_i(h)$, $e_i(h) = C_i(h_i)\frac{\Delta z_i}{\Delta t}h_i + K_{i-1/2}(h) - K_{i+1/2}(h)$, and $u_i(k) = S_i(k)\Delta z_i$. Equation (2.7) describes the water content dynamics at node i. Defining vector $E = [e_1\, e_2\, ...e_N]^T$ and $u = [u_1\, u_2\, ...u_N]^T$, and combining the equations for the N nodes together, we obtain the following matrix equation:

$$A(h(k))h(k+1) = E(h(k)) + u(k) \quad \text{or} \quad h(k+1) = A(h(k))^{-1}E(h(k)) + A(h(k))^{-1}u(k) \tag{2.8}$$

where $A(h)$ is a $N \times N$ tridiagonal matrix as shown below:

$$A(h) = \begin{bmatrix} b_1(h) & a_1(h) & 0 & 0 & ... & ... & 0 \\ d_2(h) & b_2(h) & a_2(h) & 0 & ... & ... & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ 0 & ... & ... & 0 & d_{N-1}(h) & b_{N-1}(h) & a_{N-1}(h) \\ 0 & ... & ... & ... & 0 & d_N(h) & b_N(h) \end{bmatrix} \tag{2.9}$$

Note that in matrix A, we do not have d_1 and a_N. This is related to the boundary conditions of the system. The effects of the boundary conditions are considered as the source and sink terms in the model and are included in u and will be explained in Section 2.1.3. Equation (2.2) is also discretized for the N nodes as follows:

$$\theta(h) = \begin{bmatrix} \theta_1(h_1) \\ ... \\ \theta_N(h_N) \end{bmatrix} \tag{2.10}$$

Typically, in an agro-hydrological system, soil moisture contents at different depths (i.e.,nodes) are measured. That is, the measured outputs of the system may be described as follows:

$$y = D\theta \tag{2.11}$$

where D is a matrix indicating the nodes where the soil moisture contents are measured.

Equation (2.8) and equation (2.11) define the discretized system and its outputs. Taking into account additive process noise and measurement noise, the system model can be written in the following general form:

$$h(k+1) = f(h(k), u(k)) + w(k) \tag{2.12}$$

$$y(k) = g(h(k)) + v(k) \tag{2.13}$$

where $f(h,u) := A(h)^{-1}E(h) + A(h)^{-1}u$, $g(h) = D\theta(h)$, and w and v denote the process and measurement noise. The above discretized agro-hydrological model is used throughout this thesis. The model was validated against Hydrus 1D model [80] as shown in the appendix.

2.1.3 Boundary conditions

The two boundary conditions required for the solution of Richards' equation are reflected in u_1 and u_N terms in vector u. In equation (2.6), when $i = 1$, either the term $K_{i-1/2}\dfrac{h_{i-1} - h_i}{\frac{1}{2}(\Delta z_{i-1} + \Delta z_i)}$ is replaced by a flux term $q_{top}\,(cm/hr)$ known as flux boundary condition or the term h_{i-1} is replaced by pressure head term $h_o\,(cm)$ known as the head boundary condition. Specification of either of these boundary conditions at the top node or compartment removes the term d_1 and are reflected in variable e_1

and u_1 as shown below:

$$b_1(h^k)h_1^{k+1} + a_1(h^k)h_2^{k+1} = e_1(h^k) - u_1^k. \tag{2.14}$$

Flux boundary condition:

$$a_1(h) = -\frac{K_{1\frac{1}{2}}(h)}{\frac{1}{2}(\Delta z_1 + \Delta z_2)} \tag{2.15a}$$

$$b_1(h) = C_1(h_1)\frac{\Delta z_1}{\Delta t} - a_1(h) \tag{2.15b}$$

$$e_1(h) = C_1(h_1)\frac{\Delta z_1}{\Delta t}h_1 - K_{1\frac{1}{2}}(h) \tag{2.15c}$$

$$u_1(k) = q_{top} + S_1(k)\Delta z_1 \tag{2.15d}$$

Head boundary condition:

$$a_1(h) = -\frac{K_{1\frac{1}{2}}(h)}{\frac{1}{2}(\Delta z_1 + \Delta z_2)} \tag{2.16a}$$

$$b_1(h) = C_1(h_1)\frac{\Delta z_1}{\Delta t} - a_1(h) + \frac{K_{-1/2}(h)}{\frac{1}{2}(\Delta z_{-1} + \Delta z_1)} \tag{2.16b}$$

$$e_1(h) = C_1(h_1)\frac{\Delta z_1}{\Delta t}h_1 + K_{-1/2}(h) - K_{1\frac{1}{2}}(h) \tag{2.16c}$$

$$u_1(k) = \frac{K_{-1/2}(h)}{\frac{1}{2}(\Delta z_{-1} + \Delta z_1)}h_0 + S_1(k)\Delta z_1 \tag{2.16d}$$

It can be observed that in equations (2.15d) and (2.16d), the input is reflected in the term q_{top} or h_0 depending on the flux or head condition. The choice of flux or head boundary condition at the top node depends upon the saturation of soil. If the top node is saturated and ponding occurs then the head boundary condition is chosen whereas when the top node is unsaturated, flux boundary condition is chosen. The switching between these boundary conditions and the estimation of h_0 is implemented in this work following [81]. Similarly, for the bottom node or compartment $i = N$, either the term $K_{i+1/2}\frac{h_i - h_{i+1}}{\frac{1}{2}(\Delta z_i + \Delta z_{i+1})}$ is replaced by a flux term q_{bot} (cm/hr) or the term h_{i+1} is replaced by pressure head term h_{bot} (cm) thereby removing the term a_N from the A matrix and is reflected in e_N and u_N. The bottom boundary condition is described by the drainage condition of the soil profile.

2.2 Evapo-transpiration and crop growth models

Richards' equation describes the water flow inside the soil matrix. Depending on the boundary of the system under consideration, the complexity of the overall solution to the Richards' equation varies. Since crop grows on the soil surface and irrigation along with the rain all enters the soil from the surface, this makes the agro-hydrological system to have a temporal and spatially varying top boundary and makes the problem very challenging. The crop model is integrated with the Richards' equation through the sink term in equation (2.1) and the top boundary condition as explained in Section 2.1.3. In this section, the crop model and the crop water requirements will be explained.

2.2.1 Evapo-transpiration modeling

One of the most common approach for modeling the crop growth for determining irrigation water requirement is based on the idea of 'standard' or 'reference' crop as described in Allen et al. [82]. The reference crop is defined in [82] as "*a hypothetical reference crop with an assumed crop height of* $0.12\,m$*, a fixed surface resistance of* $70\,s/m$ *and an albedo of* 0.23." This description of the reference crop closely resembles the green grass surface which fully shades the soil.

As shown in Figure 2.1, the mass balance of water in the soil-atmosphere interface is given by:

$$Q_0 = R + I + RN - RF - \mathcal{I}n - ET \tag{2.17}$$

where $Q_0\,(cm/hr)$ is the water that reaches the soil surface, $R\,(cm/hr)$, $I\,(cm/hr)$, $RN\,(cm/hr)$, $RF\,(cm/hr)$, $\mathcal{I}n\,(cm/hr)$, and $ET\,(cm/hr)$ are rain, irrigation, run-on, run-off, interception and evapo-transpiration amount, respectively. Note that Q_0 is the amount of water that reaches the soil surface, not the water that infiltrates the soil surface. The infiltration rate is governed by Q_0 and the saturation condition of the system (the current pressure head under the soil surface). The interception is the amount of water that is intercepted by the leaves of the crop before it reaches the surface. The interception in the agro-hydrological model is designed following the study by Van Dam et al., [79]. Run-on and run-off events occur when the input flux becomes higher than the infiltration rate and when the surface is not flat.

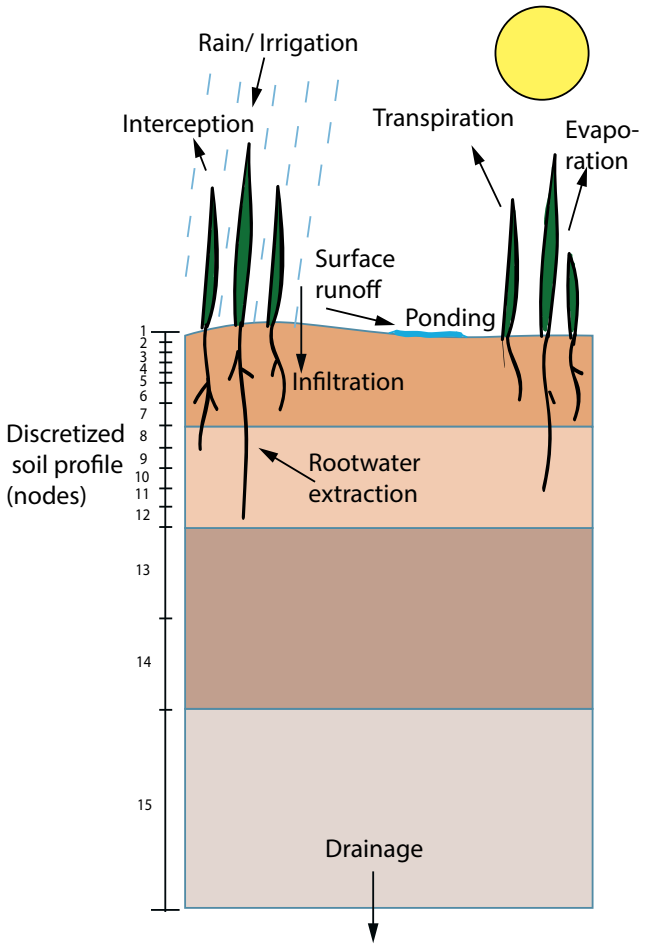


Figure 2.1: Modeling of a reference crop for irrigation water requirement.

The evapo-transpiration value (ET) accounts for the plant water requirement. It is a combined term for evaporation and transpiration. Evaporation accounts for the water loss beneath the crop canopy and surrounding exposed soil, where as transpiration is the process of water movement through the crop and its evaporation from the aerial parts, such as leaves, stems and flowers [48]. Ideally, the amount of water transpired is equal to the root water extraction (sink term).

Calculation of the potential evapo-transpiration

The evapo-transpiration (ET) is calculated using Penmon-Montieth equation [82] using hourly weather data, which is shown as follows:

$$ET_0 = \frac{\frac{\Delta_v}{\lambda_w}(R_n - G) + \frac{\rho_a c_p}{\lambda_w}\frac{(e_s - e_a)}{r_a}}{\Delta_v + \gamma(1 + \frac{r_s}{r_a})} \tag{2.18}$$

where $ET_0\,(cm/hr)$ is the potential evapo-transpiration for the reference crop (grass), $R_n\,(J/cm^2hr)$ is the net radiation, $G\,(J/cm^2hr)$ is the soil heat flux, $(e_s - e_a)\,(kpa)$ is the vapor pressure deficit of the air, $c_p(J/g^0C)$ is the air specific heat, $\rho_a(g/cm^3)$ is the mean air density, $\gamma(kPa/^0C)$ is the psychometric constant, $\Delta_v(kPa/^0C)$ is the slope of vapor pressure curve, and $r_s\,(hr/cm)$ and $r_a\,(hr/cm)$ are the surface and aerodynamic resistances, respectively. The potential and the actual evapo-transpiration of a crop are defined as:

$$ET_p = \mathcal{K}_c ET_0 \tag{2.19}$$

$$ET_a = \mathcal{K}_s(\theta) ET_p \tag{2.20}$$

where $\mathcal{K}_c$ is the crop factor, an empirical value that describes crop evapo-transpiration with respect to the reference crop and $\mathcal{K}_s(\theta)$ is the water stress factor which is a function of the soil moisture. It is to be noted that equation (2.20) is true when the stress on crops from other factors such as nutrients deficiency or diseases are considered to be negligible.

In general, the crop factor incorporates the crop characteristics and the averaged effects of evaporation of the soil [82] and is a function of the crop growth stage. Since this formulation averages out the effects of crop characters involving transpiration and evaporation from the crop canopy and soil, to have a more accurate estimation of the crop potential evapo-transpiration, the approach used in this thesis follows the work of Van Dam et al. [79]. Instead of determining the potential evapo-transpiration of the reference crop and then directly using literature values for $\mathcal{K}_c$, equation (2.18) is slightly modified for the crop which would preserve the same physical aspect of equation (2.19) but with more accuracy. The Penmon-Montieth equation as shown in equation (2.18) actually adds up the contribution from two events, the radiation

term $\frac{\frac{\Delta_v}{\lambda_w}(R_n - G)}{\Delta_v + \gamma(1 + \frac{r_s}{r_a})}$ and the aerodynamic term $\frac{\frac{\rho_a c_p}{\lambda_w}\frac{(e_s - e_a)}{r_a}}{\Delta_v + \gamma(1 + \frac{r_s}{r_a})}$. These terms can be directly updated for crops at various growth stage by updating the values of r_s, r_a and R_n of a crop at different stages. In the model these terms are calculated separately from three different scenarios; bare wet soil (ws), dry plant canopy considering full soil coverage (dp), wet plant canopy considering full soil coverage (wp), which refers to the calculation of potential evaporation rate from a wet bare soil, potential transpiration rate from a dry crop and potential transpiration rate from a wet crop, respectively. The contributions from all these scenarios are used to calculate the ET_p value with more accuracy. The steps of ET_p calculation are discussed below:

1 *Modify the psychometric constant:* The term $\gamma(1 + \frac{r_s}{r_a})$ is modified and divided in three terms to accommodate the three scenarios mentioned above.

$$\gamma_{ws} = \gamma(1 + \frac{r_s^{ws}}{r_a^{ws}}) \tag{2.21}$$

$$\gamma_{dp} = \gamma(1 + \frac{r_s^{dp}}{r_a^{dp}}) \tag{2.22}$$

$$\gamma_{ws} = \gamma(1 + \frac{r_s^{wp}}{r_a^{wp}}) \tag{2.23}$$

where r_s^{ws} (hr/cm), r_s^{dp} (hr/cm), r_s^{wp} (hr/cm) are surface resistance of wet soil, dry crop and wet crop, respectively. The values for r_s^{ws} and r_s^{wp} are generally taken as zero and the value of r_s^{dp} is specific to crop type. r_a^{ws} (hr/cm), r_a^{dp} (hr/cm), r_a^{wp} (hr/cm) are aerodynamic resistance of wet soil, dry crop and wet crop, respectively, which is a function of crop height. For soil, r_a^{ws} is considered fixed with the crop height of $0.1\,cm$.

2 *Modify the net radiation:* The radiation term R_n is the combined effect of short wave radiation and long wave radiation. The net short wave radiation is soil and crop type dependent while the net long wave is independent of crop or soil and is dependent on geographical location and day time. Thus the net short wave is divided in three groups as rn^{ws}, rn^{wp} and rn^{dp}. These three properties depend on soil and crop type only and do not vary with crop growth stage. Let us symbolize the net long wave by rl which is the same for the three scenarios.

3 *Contributions from the three scenarios:* The contributions from the three scenarios, potential evaporation from wet bare soil et_{ws}, potential transpiration from wet crop with full coverage et_{wp}, and potential transpiration from dry crop with full coverage et_{dp} are as shown below:

$$et_{ws} = \frac{\frac{\Delta_v}{\lambda_w}(rn^{ws} - rl - G) + \frac{\rho_a c_p}{\lambda_w}\frac{(e_s - e_a)}{r_a^{ws}}}{\Delta_v + \gamma_{ws}} \tag{2.24a}$$

$$et_{wp} = \frac{\frac{\Delta_v}{\lambda_w}(rn^{wsp} - rl - G) + \frac{\rho_a c_p}{\lambda_w}\frac{(e_s - e_a)}{r_a^{wp}}}{\Delta_v + \gamma_{wp}} \tag{2.24b}$$

$$et_{dp} = \frac{\frac{\Delta_v}{\lambda_w}(rn^{dp} - rl - G) + \frac{\rho_a c_p}{\lambda_w}\frac{(e_s - e_a)}{r_a^{dp}}}{\Delta_v + \gamma_{dp}} \tag{2.24c}$$

4 *For partly covered soil:* The calculation of the evapo-transpiration shown so far considers either bare soil condition or fully covered soil, which is not true for the crop which partially covers the soil at various growth stages. So this step adjusts the calculated values for partly covered soils.

- For partly covered soil, the potential evaporation reduces with respect to bare soil as follows:

$$W_{frac} = \frac{R_i}{et_{wp}} \; with \; W_{frac} \leq 1.0 \tag{2.25a}$$

$$E_p = et_{ws}(1 - W_{frac})e^{-\kappa_{gr} LAI} \tag{2.25b}$$

  where W_{frac} defines the fraction of the day the crop is wet, R_i is intercepted rain, E_p denotes potential evaporation of partially covered soil, κ_{gr} is the extinction co-efficient for solar radiation and LAI is the leaf area index which varies with crop type and their growth stage.

- For partly covered soil the potential transpiration is calculated using the et_{dp} value as follows:

$$T_p = et_{dp}(1 - W_{frac}) - E_p \tag{2.26}$$

Actual evaporation calculation

The potential evaporation amount may undergo more reduction as the surface gets drier, which accounts for the actual evaporation E_a. As the soil gets drier the soil

hydraulic conductivity decreases which affects the evaporation rate. The maximum evaporation rate the top soil can sustain is calculated according to the Darcy's law as follows:

$$E_{atm} = K_{1/2}\left(\frac{h_{atm} - h_1 - z_1}{z_1}\right) \tag{2.27}$$

where $K_{1/2}$ is the average hydraulic conductivity between soil surface and the first node, h_{atm} is the soil water pressure head in equilibrium with the air relative humidity. It is to be noted that E_{max} depends on the thickness of top soil layer and the hydraulic conductivity $K(\theta)$ of that layer. Since this layer undergoes various events such as splashing rain, dry crust, root extension and different cultivation practice, the true value of E_{max} is often limited to erroneous results due to inadequate understanding of this layer dynamics. Thus an empirical method of calculating the cumulative actual evaporation during a drying cycle using Black's method [79] is also considered as follows:

$$E_c = \sum E_a = \beta_l t_{dry}^{1/2} \tag{2.28}$$

where β_l is the specific soil parameter ($cmd^{-0.5}$), characterizing the evaporation process and t_{dry} is the time after a significant amount of rainfall. If the precipitation amount is greater than this significant amount t_{dry} is reset to zero.

In the model, the three evaporation values from equations (2.25b), (2.28) and (2.27) are compared depending on the rain amount, and the final evaporation amount is chosen as follows:

$$E_a = min(min(E_p, E_c), max(0, E_{max})) \tag{2.29a}$$

Actual transpiration and root water extraction

The calculation of the transpiration amount is the most important part in the crop model, as well as in this whole study. The transpiration amount mentioned in equation (2.26) is equal to the potential water requirement of the plant through the root zone. A plant or crop which does not endure any water stress would extract this amount of water using its roots. But since roots span over a zone in the soil, this potential transpired water requirement is distributed over the entire root zone as follows:

$$S_p(z) = \frac{l_{root}(z)}{\int_{-D_{root}}^{0} l_{root}(z)dz} T_p \tag{2.30}$$

where D_{root} is the root layer thickness, $l_{root}(z)$ is the length of root upto node $z(i)$ and $S_p(z)$ is the potential root water extraction for depth z corresponding to node i. This potential root water extraction will undergo reduction due to various stress as follows:

$$S_a(z) = \alpha_{wr}\alpha_{dr}\alpha_{sr}\alpha_{fr}S_p(z) \tag{2.31}$$

where $S_a(z)$ is the actual root water extraction at node i, α_{wr}, α_{dr}, α_{sr} and α_{fr} are the reduction factors due to wet stress, drought stress, salinity stress and frozen soil stress, respectively. In this study, α_{sr} and α_{fr} are not considered with the assumption that soil is not frozen during the simulated period and the soil salinity is at standard conditions. The actual transpiration amount is given as follows:

$$T_a = \int_{-D_{root}}^{0} S_a(z)dz \tag{2.32}$$

2.3 Crop growth modeling

In this thesis, we consider a simple crop growth model. The simple crop growth model uses the following relation to define the leaf area index (LAI), root depth (RD) and crop height (CH) all as a function of the crop development stage (DVS). The DVS value is zero for crop at the day of seeding and two at the day of harvest. In the simple crop model, one must specify the number of days the crop season includes (k_{CS}) and then develop a linear function of DVS (equation 2.33a) to describe different crop growth stages. The values of LAI, RD and CH at different crop growth stage is then expressed in terms of DVS as shown below:

$$DVS(\underline{k}) = \frac{2\underline{k}}{k_{CS}} \tag{2.33a}$$

$$\frac{dLAI}{d\underline{k}} = f_{c1}(DVS(\underline{k})) \tag{2.33b}$$

$$\frac{dRD}{d\underline{k}} = f_{c2}(DVS(\underline{k})) \tag{2.33c}$$

$$\frac{dCH}{d\underline{k}} = f_{c3}(DVS(\underline{k})) \tag{2.33d}$$

where $\underline{k}$ is the day number after the crop is sown. Thus the evaporation and the transpiration values are calculated using these $LAI(\underline{k})$, $RD(\underline{k})$ and $CH(\underline{k})$ values at

different crop growth stages. The equation (2.17) is slightly modified to accommodate the modified evaporation and transpiration calculation.

$$Q_0 = R + I + RN - RF - \mathcal{I}n - E_a \tag{2.34}$$

This Q_0 flux is the water that is actually present in the soil surface, and depending on the soil moisture condition this value is used to determine the top flux for flux boundary condition (Neuman BC) or the surface pressure head for head boundary condition (Drichilet BC). The transpiration amount is represented by the sink term in Richards' equation as follows:

$$S(i) = S_a(z(i)) = \alpha_{wr}(i)\alpha_{dr}(i)S_p(z(i)) \tag{2.35}$$

Note that the ratio of the actual to the potential transpiration is given by the stress factors $\alpha_{wr}\alpha_{dr}$ which is the $\mathcal{K}_s$ value in equation (2.20) and is an important parameter in irrigation decision making.

2.4 Conclusions

The agro-hydrological model includes a few elements including the Richards' equation that describes the water dynamics inside the soil, the evapo-transpiration model that describes how water is consumed by the roots of the crop and the surface, and the crop growth model that describes how water requirement of the crop changes over time. These models will be used in the later chapters to simulate the agro-hydrological systems.

Chapter 3

State and parameter estimation based on system observability

3.1 Introduction

In this chapter, a systematic approach based on system observability analysis and state estimation is developed to estimate the soil moisture inside an agro-hydrological system where measurements are not easily available. A discrete-time state-space model based on Richards' equation is used to describe the agro-hydrological system as explained in Chapter 2. The nonlinear agro-hydrological model is linearized every sampling time and the observability of the overall system is determined based on the locally linearized model every sampling time. Based on the linearized models, we investigate how the number and location of output measurements affect the degree of observability of the system.

In Section 3.2, the details on observability analysis for this nonlinear system is described, and an optimal sensor placement algorithm is proposed. Based on an observable system, parameter and state estimation is performed in Sections 3.3 and 3.4. Since linearization is required at each step for the observability analysis, and given that the system under study considers only vertical one dimensional flow of water, the filter chosen for state estimation in this study is the EKF. Finally the validation of optimal sensor placement algorithm is performed on both synthetic and real field data in Sections 3.5 and 3.6.

3.2 Observability analysis

A system is said to be observable, if the knowledge of the input(s) and the output(s) over a finite time interval $[t_0,\ t_f]$ is adequate to uniquely determine the initial state $h(t_0)$ [83]. Since the agro-hydrological system is a nonlinear process, in order to check observability, the system is linearized at different points of the state trajectory and the observability of these locally linearized models are investigated [84]. Therefore the nonlinear model was simulated over some soil moisture range which is within the operation trajectory and at each time step corresponding to the current states the system is linearized. The linearization of the system at $h(k)$ (equations (2.12) and (2.13)) can be described using the following form with the assumption of zero process and measurement noise (without the loss of generality):

$$h(k+1) = F(k)h(k) + B(k)u(k) \tag{3.1}$$

$$y(k) = G(k)h(k) \tag{3.2}$$

where $B(k)$, $F(k)$ and $G(k)$ are obtained numerically using the jacobian matrices as shown below:

$$F(k) = \begin{bmatrix} \frac{\partial f_1}{\partial h_1} & \cdots & \frac{\partial f_1}{\partial h_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial h_1} & \cdots & \frac{\partial f_N}{\partial h_N} \end{bmatrix}_{h=h(k)} \tag{3.3}$$

$$B(k) = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_p} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial u_1} & \cdots & \frac{\partial f_N}{\partial u_p} \end{bmatrix}_{u=u(k)} \tag{3.4}$$

$$G(k) = \begin{bmatrix} \frac{\partial y_1}{\partial h_1} & \cdots & \frac{\partial y_1}{\partial h_N} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_m}{\partial h_1} & \cdots & \frac{\partial y_m}{\partial h_N} \end{bmatrix}_{h=h(k)} \tag{3.5}$$

where N, p and m denote the number of state variables, the number of input variables and the number of measured outputs, respectively. The observability test can be

carried out using different approaches. In this study, the rank test is performed, in which the system is considered to be observable at time k if the observability matrix $\mathcal{O}(k)$ has a full column rank; that is,

$$\text{rank } (\mathcal{O}(k)) = n \tag{3.6}$$

where the observability matrix is given by:

$$\mathcal{O}(k) = \begin{bmatrix} G(k) \\ G(k)F(k) \\ ... \\ G(k)F(k)^{N-1} \end{bmatrix}. \tag{3.7}$$

The rank test provides a 'yes' or 'no' answer to the system observability and does not provide any information on how strongly or weakly observable a system is. Therefore we perform a degree of observability analysis of an observable system based on the observability matrix $\mathcal{O}(k)$ to determine to what extent the system is observable [85].

3.2.1 Degree of observability analysis

Following the work of Ham et al., it is considered that the lack of co-linearity of the rows of the observability matrix $\mathcal{O}(k)$ is an indication of the degree of observability [85, 86]. This implies that the degree of observability can be quantified by investigating the spatial relationships among all the rows of the observability matrix $\mathcal{O}(k)$ [87]. Given that only the angles between the row vectors of $\mathcal{O}(k)$ are important in analyzing the linear independence, the row vectors of $\mathcal{O}(k)$ are normalized into unit length vectors. Let the normalized observability matrix at time k be denoted as $\mathcal{N}$ and the rows of observability matrix $\mathcal{O}(k)$ be denoted as $\mathcal{O}^j(k)$ where $j = 1, 2, ..., mN$ are the rows. The normalized element at the j-th row and the i-th column in matrix $\mathcal{N}$ is given by:

$$\mathcal{N}_i^j = \frac{\mathcal{O}_i^j}{\sqrt{\left(\sum_{i=1}^{N} |\mathcal{O}_i^j|^2\right)}} \tag{3.8}$$

where $i = 1, 2, .., N$ denotes the elements inside each row. It was shown, that the smallest eigenvalue of $\mathcal{N}^T\mathcal{N}$ indicates the degree (or lack or weakness) of observability of a system and the corresponding eigenvector refers to the direction of weakest observability [85]. That is,

$$deg\mathcal{N} = \lambda_{min} \tag{3.9}$$

where $deg\mathcal{N}$ is the degree of observability and λ_{min} is the minimum among all the eigenvalues λ_i $(i = 1, \ldots, N)$. If we denote the eigenvector corresponding to λ_{min} as V_{min}, then V_{min} is the direction of weakest observability. The elements of V_{min} with large values correspond to states which contribute more towards unobservability. The observability of the system is highest when all its eigenvalues are equal to each other. On the other hand, the system is unobservable when one or more of its eigenvalue is equal to zero. The relative measure of system observability is given by the ratio of eigenvalues:

$$\mathcal{R} = \frac{\lambda_{max}}{\lambda_{min}} \tag{3.10}$$

where λ_{max} is the maximum eigenvalue of $\mathcal{N}^T\mathcal{N}$. The system having the value of the eigenvalue ratio one or closer to one is considered to be a highly observable system. Since the system under study is highly nonlinear so this degree of observability analysis needs to be performed at every instant over the entire time trajectory which should cover different nonlinear regions. The degree of observability analysis will help us to determine the minimum number of measurements required for the system to be observable and the optimal placement of the sensors for an improved degree of observability.

3.2.2 Optimal sensor placement

The aim of the optimal sensor placement algorithm is to find the smallest number of measurements that ensures the observability of the entire system and correspondingly to determine the best locations for measurements. The overall strategy is described below:

1. Generate input so that the soil moisture stays well between θ_{res} and θ_{sat}. It is better to avoid extreme soil conditions such as very low or very high soil moisture due to a high degree of nonlinearity.

2. Divide the simulation horizon by the sampling time to obtain l sampling points. Accordingly linearize the nonlinear model for each sample point using the jacobian in equation (3.5) for the state matrix $F(q = 1, 2, ..., l)$.

3. Intially, let us consider that all states N can be measured, so $m = N$. Thus,

the set of states which are measured (available measurements) can be defined as $\mathcal{M}^{(m)} = Z$ where $Z = \{1, 2, ..., N\}$ and m is the number of available measurements. Let us also denote the set of states which are not measured (unavailable measurements) by $\mathcal{U}^{(m)}$. Thus at initial step $\mathcal{U}^{(m=N)} = \emptyset$.

4. Use equation (3.7) to perform the rank test. If the system is observable then proceed to next step. If the system is unobservable, this algorithm stops.

5. Normalize the observability matrix using equation (3.8). Then perform the degree of observabilty analysis using equations ((3.9)-(3.10)) for $q = 1, 2, ..., l$. Identify the largest value of $\mathcal{R}$ which stands for the lowest degree of observability and let us denote the corresponding sample point by q_{max}. Now perform rest of the analysis for the sample point q_{max}.

6. Determine the lowest eigenvalue λ_{min} of the normalized observability matrix $\mathcal{N}^T\mathcal{N}$. Analyze the eigenvector V_{min} corresponding to λ_{min}. The elements of V_{min} with large values contribute towards unobservability and small values contribute towards observability. So identify the position or index of the smallest element in the V_{min} vector and let us denote this position by p which also refers to the state with relatively more observability. Thus remove the measurement corresponding to the position p. Therefore, $\mathcal{U}^{(m-1)} = \{p\} \cup \mathcal{U}^{(m)}$ and $\mathcal{M}^{(m-1)} = \{r : r \in Z \; and \; r \notin \mathcal{U}^{(m-1)}\}$. In this step of the algorithm, measurements with lest effect on the system observability is identified and removed.

7. Go to step 4. Continue until the system becomes unobservable in the rank test (step 4).

In the algorithm, the rank test at step 4 determines the minimum number of measurements required for the system to be observable. The most important step in this algorithm is step 6 which actually gives information on sensor placement. Since all the states in this system refer to the same type of variable, soil moisture, but at different depths, therefore identifying the right depth for sensor placement is challenging. At step 6 of this algorithm, measurements are reduced based on the degree of observability. Note that in the proposed algorithm, the aim is to find the smallest number of

measurements that ensures the observability and then accordingly to determine the best locations for measurements. The algorithm stops when the observability rank test indicates that the system is unobservable (step 4). The algorithm may be adjusted to take other factors into account (e.g., the minimum degree of observability) by revising the termination criterion in step 4 (e.g., in addition to the rank test, the degree of observability is also checked).

3.3 Model parameter estimation

The parameters of the model are uncertain and need to be estimated. Due to the heterogeneity of soil at different depths, several parameters (θ_{sat}, θ_{res}, K_{sat}, η and α) vary with depth. In this work, we consider that there are four different soil types according to the four layers shown in Figure 2.1 in the region under consideration. Within each layer, the soil is considered homogeneous. For each layer these five parameters are estimated.

In this work, we will use the prediction error method (PEM) to estimate the parameters. PEM is an optimization based approach and its objective is to minimize the model prediction error subject to the model parameters that are to be identified [88, 89]. Let us use Φ to denote the vector containing all the $p = 20$ parameters that need to be estimated; that is, $\Phi = [\phi_1,\ \phi_2,\ \phi_3,\ \phi_4]^T$ with $\phi_i = [\theta_{sat,i},\ \theta_{res,i},\ K_{sat,i},\ \eta_i,\ \alpha_i]^T$ and assume that at anytime k we have measurements $y(k)$ for $k = 0, 1, \ldots, t$ and we have all the information about the boundary conditions (i.e., u). The corresponding PEM optimization problem takes the following form:

$$\min_{\phi,\ \hat{h}(0)} J(\phi) = \text{trace}\left(\sum_{k=1}^{t} \epsilon(k)\epsilon(k)^T\right) \tag{3.11a}$$

$$\text{s.t.,}\ \hat{h}(k+1) = f(\hat{h}(k), u(k)),\ k = 0, \ldots, t-1 \tag{3.11b}$$

$$\hat{y}(k) = g(\hat{h}(k)),\ k = 1, \ldots, t \tag{3.11c}$$

$$\epsilon(k) = y(k) - \hat{y}(k),\ k = 1, \ldots, t \tag{3.11d}$$

$$0 < \theta_{sat,i} < 1, 0 <\ \theta_{res,i} < 1,\ K_{sat,i} \geq 0,\ i = 1, 2, 3, 4 \tag{3.11e}$$

where $\hat{h}(0)$ denotes the initial condition of the system state, $\hat{h}$ and $\hat{y}$ denote the predicted state and output vectors, y denotes the actual measurements and ϵ is the

prediction error at each time instant. In the PEM optimization problem, both the parameters as well as the system initial state are optimized/estimated. The objective of the PEM is to minimize the trace of the covariance of the prediction error as shown in equation (3.11a). The nominal model (3.11b)-(3.11c) of the system is used in PEM to predict the output of the system. The error between the predicted output and measurements is shown in equation (3.11d) and the constraints imposed on the parameters are described in equation (3.11e).

3.4 Design of the extended Kalman filter

To handle the nonlinearity of the agro-hydrological model, the state estimation of the system is performed using the extended Kalman filter. The algorithm consists of two steps. The first step is the prediction step which uses the model to generate the output and the second step is the update step where the predicted output is updated according to the difference between the actual measurement and the predicted output obtained in the first step. These steps are described as follows:

Prediction step

$$\hat{h}(k+1|k) = f(\hat{h}(k|k), u(k)) \tag{3.12a}$$

$$P(k+1|k) = F(k)P(k|k)F(k)^T + Q(k) \tag{3.12b}$$

Update step

$$\tilde{y}(k+1) = y(k+1) - g(\hat{h}(k+1|k)) \tag{3.13a}$$

$$S(k+1) = G(k+1)P(k+1|k)G(k+1)^T + R(k+1) \tag{3.13b}$$

$$K(k+1) = P(k+1|k)G(k+1)^T S(k+1)^{-1} \tag{3.13c}$$

$$\hat{h}(k+1|k+1) = \hat{h}(k+1|k) + K(k+1)\tilde{y}(k+1) \tag{3.13d}$$

$$P(k+1|k+1) = (I - K(k+1)G(k+1))P(k+1|k)) \tag{3.13e}$$

The prediction of states at time instant k based on the model $f(\hat{h}(k|k), u(k))$ is given by $\hat{h}(k+1|k)$. The error covariance matrix of $\hat{h}(k|k)$ is described by $P(k|k)$, and the predicted error covariance matrix at time $k+1$ based on the corresponding value at time k is denoted by $P(k+1|k)$. The matrices $Q(k)$ and $R(k+1)$ describe the covariance of process noise and measurement noise of the system, respectively.

Remark 1 *The goal of this study is to develop an optimal sensor placement algorithm for agro-hydrological systems. While EKF is used in this work, other state estimation algorithms (such as EnKF, iterative EnKF, and particle filters) may be used in the proposed framework. Depending on the objective and the degree of complexity of the specific problem, one method may excel the others.*

3.5 Model observability analysis and simulations

3.5.1 Soil profile and problem description

In this work, 110 cm of soil profile with four different homogeneous soil layers are considered. The first homogeneous layer is 15 cm deep, the second, third and fourth layers are 25 cm, 40 cm and 30 cm deep, respectively. These homogeneous layers are further discretized into small compartments to obtain the numerical solution to the Richards' equation as shown in Figure 2.1. The states of this system correspond to the soil moisture of the discretized soil compartments. Thus, the number of states in this agro-hydrological system varies according to the discretization of the soil profile. The higher the number of nodes (N), the better is the accuracy of the numerical solution of the Richards' equation. On the other hand, the high number of nodes may cause observability issues due to limited available measurements. Therefore, a trade-off exists between the number of states and system observability while performing the discretization of the soil profile. In this study, the soil profile is discretized into 15 compartments, so there are 15 states as shown in Table 3.1. The details on the choice of 15 states is presented in the appendix. Details on the use of degree of observability analysis to determine optimal sensor placement and how the observability improves the overall parameter and state estimation are addressed in this chapter.

3.5.2 Observability analysis and optimal measurement nodes

The algorithm mentioned in Section 3.2.2 is implemented to identify optimal sensor locations using observability. The weather conditions were simulated for 460 hours for six weather parameters including maximum and minimum temperatures, humidity, wind speed, solar radiation and the amount of precipitation on an hourly basis. Since precipitation has a direct impact on the soil moisture and is considered to be the

Table 3.1: The discretized soil profile with 15 nodes or compartments. Each of these nodes corresponds to the states of the system.

soil layer	soil sublayer	height of sublayer, cm	height of compartments, cm	number of compartments
1	1	10.00	2.00	5
1	2	5.00	2.50	2
2	3	25.00	5.00	5
3	4	40.00	20.0	2
4	5	30.00	30.0	1

input of the model, therefore among the weather parameters only the precipitation values are reported in this chapter.

Input generation: The input is generated with the range of $[0, 2]$ using 'idinput' in MATLAB as shown in Figure 3.1 with the simulation horizon of 460 hours. Therefore $l = 460$ in step 2 in the algorithm. It is noted that the hydraulic conductivity function shown in equation (2.4) makes the linearization challenging when the soil moisture is very small (highly negative h) or the soil moisture is near saturation (h close to zero). In these cases, the linearized model may not give appropriate degree of observability of the original nonlinear system. The input should be generated to avoid extreme soil conditions such as very low or very high soil moisture.

Optimal sensor placement algorithm: Initialy, all the states are measured ($m = N = 15$). At each sampling points the state matrix $F(k = 1, 2, ..., l)$ is linearized and observability matrix $\mathcal{O}(k)$ is generated. Here, measurement set $\mathcal{M}^{(15)} = Z$ where $Z = \{1, 2, ..., 15\}$. The system is fully observable in the rank test since the rank is equal to the number of states, that is 15, as shown in Figure 3.1.

For this observable system the degree of observability analysis (step 5) is performed. The largest $\mathcal{R}$ value was observed at few sampling points. We considered largest $\mathcal{R} = 172.7$ at $q_{max} = 460$. The lowest eigenvalue λ_{min} for the $\mathcal{N}^T\mathcal{N}$ matrix at sample point q_{max} is 1.00 and the corresponding eigenvector V_{min} is shown in Figure 3.2. In the figure, it is observed that for $\mathcal{M}^{(15)}$ there is no minima. Since there is no minima, so instead of choosing one measurement to be removed, we choose two points corresponding to the lowest values, states 14 and 15, for further analysis. Figure 3.3 is used to present the sensor placement results. The y-axis shows the states corresponding to different depths inside soil and the x-axis shows the number of to-

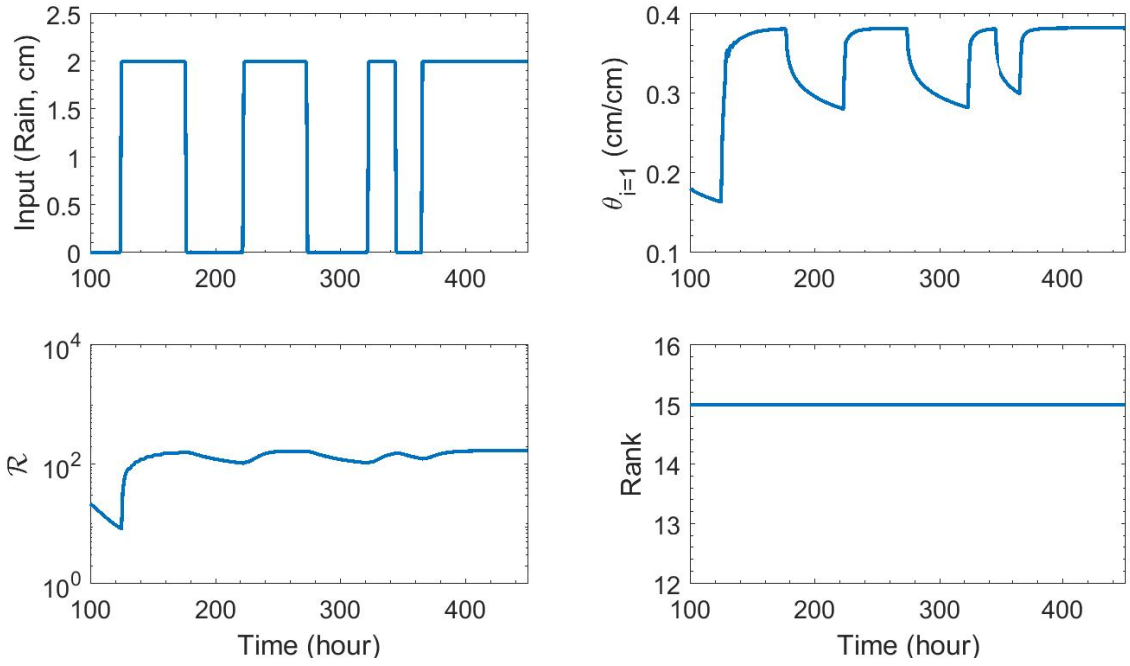


Figure 3.1: Input values for optimal sensor placement. The rank and $\mathcal{R}$ trajectory is shown only for measurements $m = 15$ and the corresponding simulated soil moisture value at node $i = 1$ is also shown

tal measurements used in the corresponding analysis. When all states are measured ($m = 15$), all the points in y-axis are represented using green circles. The two blue outer circles indicate the two states (15 and 14) which are further analyzed. It is observed that when measurement 14 is removed, the largest value of $\mathcal{R}$ is 7.498×10^8 whereas when measurement 15 is removed, the largest $\mathcal{R}$ value is 4.839×10^{10}. This implies that if measurement 15 is removed the system is relatively more unobservable. Therefore, at this step measurement 14 is removed. Thus, whenever the elements in V_{min} as shown in Figure 3.2 does not explicitly show one minimum trajectory, instead, it shows a set of elements with small values then we perform further analysis to identify the measurement to be removed. The rest of the steps in the algorithm are performed similarly and are summarized in Table 3.2

If there are few elements in V_{min} and all of them have relatively close values, then all of them must be analyzed further as explained before. Also in some cases, one must decide the locations based on experience. For example after removing state 4 ($\mathcal{M}^{(12)}$), the next state removed was state 1 ($\mathcal{M}^{(11)}$) though state 15 showed the lowest value among the elements of V_{min}. If we observe the discretization, then it can

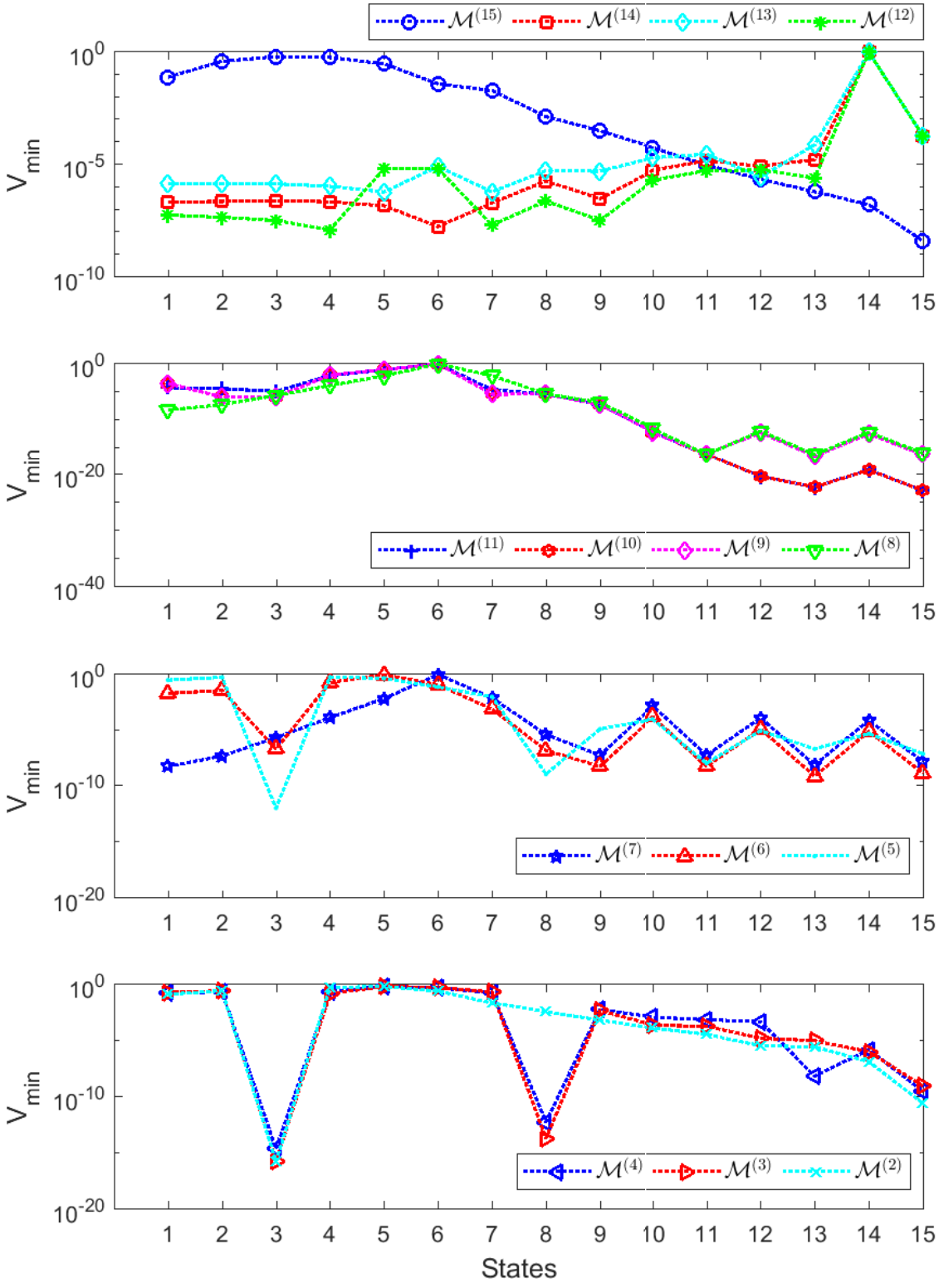


Figure 3.2: Eigenvector corresponding to λ_{min} for different measurement combinations for degree of observability analysis for optimal sensor placement.

Table 3.2: Optimal sensor placement algorithm for the simulated problem.

m	$\mathcal{U}^{(m)}$	$\mathcal{M}^{(m)}$	$rank(\mathcal{O})$	Maximum $\mathcal{R}$	λ_{min}	Location for small values in V_{min}	Measure -ment removed
15	$\{\}$	$\{1,2,3,4,5,6,$ $7,8,9,10,11,$ $12,13,14,15\}$	15	172.65	1.00	15,14	14
14	$\{14\}$	$\{1,2,3,4,5,6,$ $7,8,9,10,11,$ $12,13,15\}$	15	7.498×10^{8}	2.126×10^{-7}	6	6
13	$\{14,6\}$	$\{1,2,3,4,5,$ $7,8,9,10,11,$ $12,13,15\}$	15	5.734×10^{8}	2.5472×10^{-7}	5,7	5
12	$\{14,6,5\}$	$\{1,2,3,4,7$ $8,9,10,11,$ $12,13,15\}$	15	6.406×10^{9}	2.07×10^{-8}	4,7	4
11	$\{14,6,$ $5,4\}$	$\{1,2,3,7$ $8,9,10,11,$ $12,13,15\}$	15	2.49×10^{9}	2.07×10^{-8}	1,2,15	1
10	$\{14,6,$ $5,4,1\}$	$\{2,3,7,8,$ $9,10,11,$ $12,13,15\}$	15	9.56×10^{9}	3.94×10^{-9}	12,13,15	12
9	$\{14,6,5,$ $4,1,12\}$	$\{2,3,7,8,$ $9,10,11,$ $13,15\}$	15	9.56×10^{9}	3.94×10^{-9}	2,3,7,11	7
8	$\{14,6,5,$ $4,1,12,7\}$	$\{2,3,8,$ $9,10,11,$ $13,15\}$	15	1.98×10^{11}	3.94×10^{-9}	10,11,13,15	10
7	$\{14,6,5,$ $4,1,12,$ $7,10\}$	$\{2,3,8,$ $9,11,$ $13,15\}$	15	1.99×10^{11}	1.20×10^{-10}	2,11,13,15	2
6	$\{14,6,5,$ $4,1,12,$ $7,10,2\}$	$\{3,8,$ $9,11,$ $13,15\}$	15	1.89×10^{13}	7.93×10^{-13}	9,11,13,15	9
5	$\{14,6,5,$ $4,1,12,$ $7,10,2,9\}$	$\{3,8,$ $11,13,$ $15\}$	15	3.38×10^{16}	4.44×10^{-16}	3,8,11	11
4	$\{14,6,5,$ $4,1,12,7,$ $10,2,9,11\}$	$\{3,8,$ $13,15\}$	15^{*}	6.15×10^{15}	4.4×10^{-15}	3,8,13,15	13
3	$\{14,6,5,4,$ $1,12,7,10,$ $2,9,11,13\}$	$\{3,8,15\}$	15^{*}	8.31×10^{18}	4.4×10^{-15}	3,8,15	8
2	$\{14,6,5,4,1,$ $12,7,10,2,$ $9,11,13,8\}$	$\{3,15\}$	15^{*}	2.7436×10^{19}	6.27×10^{-19}		

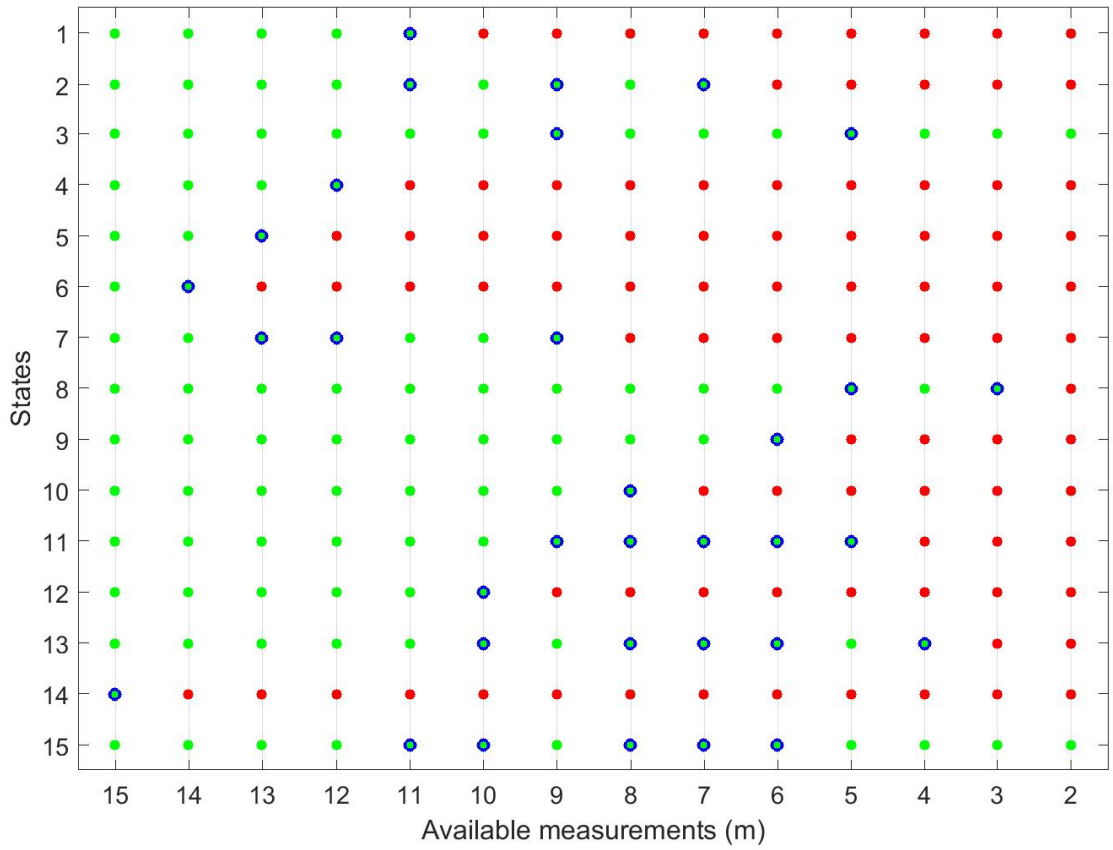


Figure 3.3: Degree of observability analysis for optimal sensor placement. Green points indicate the set of measured nodes $\mathcal{M}^{(m)}$ and red points shows the set of unmeasured nodes $\mathcal{U}^{(m)}$ for any number of measurements m. The blue outer circles represents the states which are analyzed in the next step.

be observed that state 15 corresponds to 30 cm inside the soil near the bottom. The measurement corresponding to state 14 is already removed so removing measurement from state 15 at this early stage may mislead the sensor placement process in which case the algorithm may end up giving a higher number of required measurements. We also observed some other locations including states 1 and 2 for further analysis. It was found that when state 1 is removed, $\mathcal{R} = 9.561 \times 10^9$; when state 2 is removed, $\mathcal{R} = 1.3156 \times 10^{10}$; but when state 15 is removed, $\mathcal{R} = 4.249 \times 10^{19}$. This confirms that removing state 15 will highly decrease overall observability. Therefore, based on this further analysis, measurement was removed from state 1. If we observe the $\mathcal{R}$ value in Table 3.2, then we can have an idea of how the observability is gradually reduced as the number of sensors decreases. The nonlinearity of the system can also be noticed. When all the states are measured, $\mathcal{R} = 172$; but even when one

measurement is removed, the system $\mathcal{R}$ becomes as high as 10^8. For $m = 4, 3, 2$, the rank for most of the trajectories was 15 but some rank deficient points were observed specially for low soil moisture. The ranks for these measurements are shown with an asterisk. When one measurement was used, the system became totally unobservable.

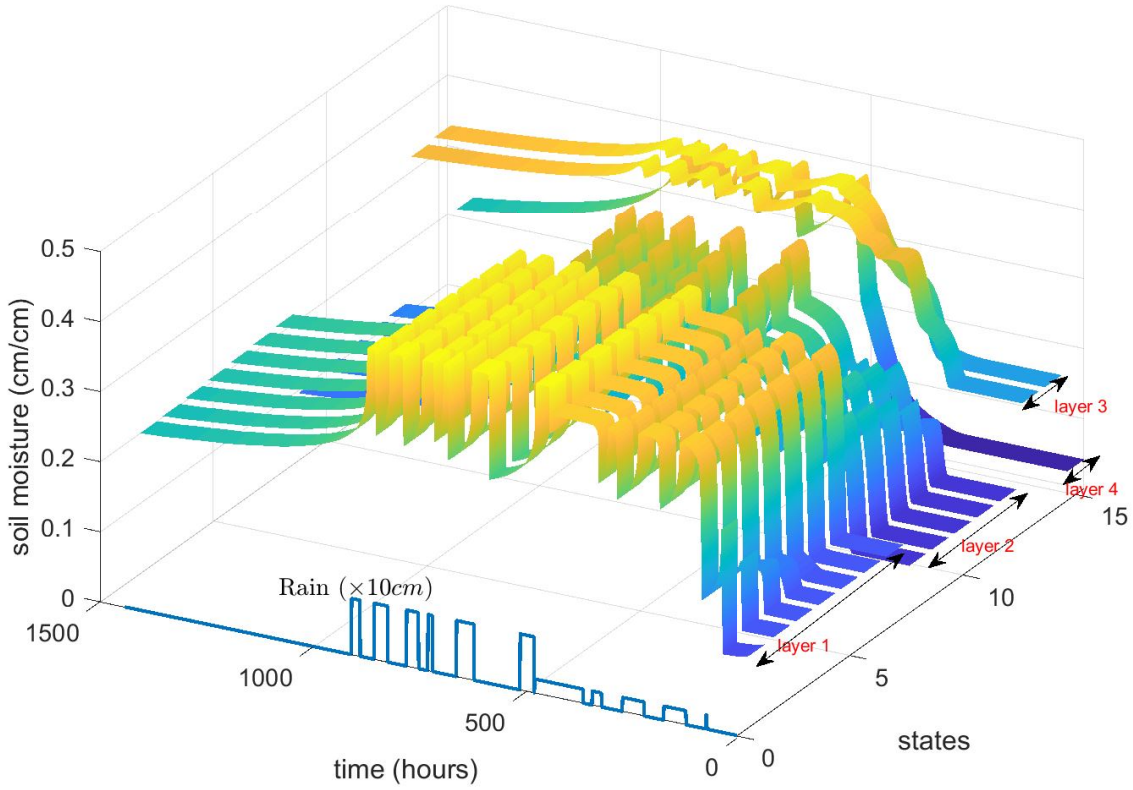
3.5.3 Parameter estimation based on simulated data

In this section, we evaluate the performance of PEM algorithm described in Section 3.3 based on the observable system identified in Section 3.5.2. The actual soil parameters for the system is described in Table 3.3 and the simulated soil moisture using the soil property is shown in Figure 3.4. To demonstrate the effects of data quality and data quantity in the parameter estimation process the simulated data was generated for the time period of 1440 hours with inputs varying between different ranges as shown in the time axis of Figure 3.4(a).

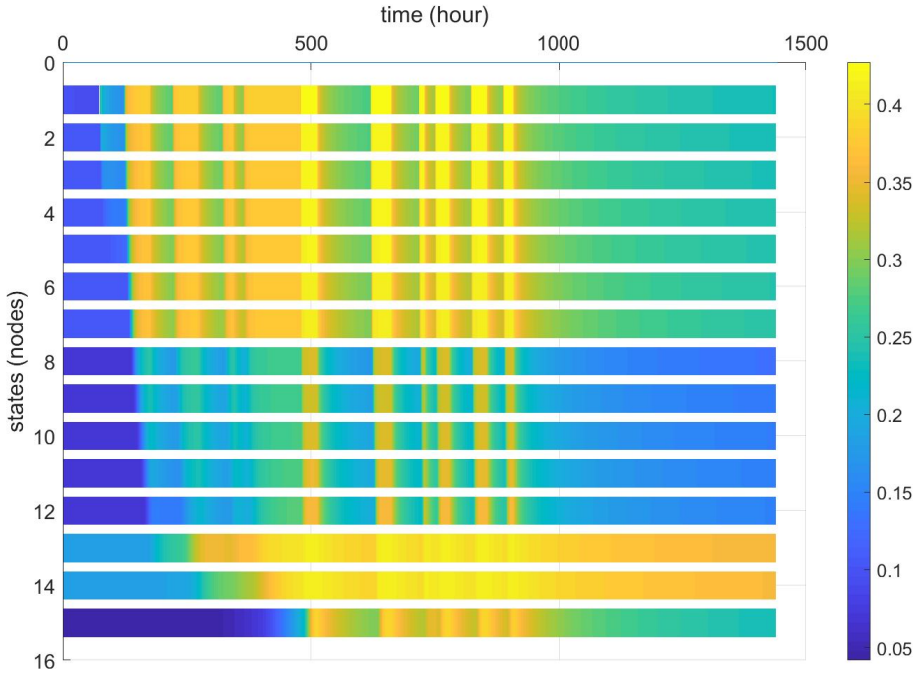
Table 3.3: Actual and estimated parameter values for simulated data set.

Parameter	Soil Layer	Soil type	Parameter value (actual)	Initial Parameter $P_{i=1}$	Initial Parameter $P_{i=2}$
Ksat (cm/h)	1	Loam	1.0400	1.2480	2.358
	2	Sandy loam	4.4208	3.9788	1.248
	3	Clay loam	0.2600	0.3640	1.248
	4	Sandy clay	1.3100	1.9650	1.703
θ_{res} (cm/cm)	1	Loam	0.078	0.0624	0.015
	2	Sandy loam	0.065	0.0780	0.0624
	3	Clay loam	0.095	0.0665	0.0624
	4	Sandy clay	0.01	0.0180	0.015
θ_{sat}(cm/cm)	1	loam	0.43	0.3354	0.273
	2	Sandy loam	0.41	0.2665	0.3354
	3	Clay loam	0.41	0.3690	0.3354
	4	Sandy clay	0.39	0.3120	0.234
α	1	Loam	0.036	0.0288	0.0354
	2	Sandy loam	0.075	0.0900	0.0288
	3	Clay loam	0.019	0.0152	0.0288
	4	Sandy clay	0.059	0.0442	0.0472
η	1	Loam	1.56	1.3	1.3
	2	Sandy loam	1.89	1.3	1.3
	3	Clay loam	1.31	1.3	1.3
	4	Sandy loam	1.48	1.3	1.3

The following property of the simulated data can be noticed from the figure.



(a) 3D plot for soil moisture profile. The rain input is shown along the time-axis.



(b) 2D plot for soil moisture profile. The color intensity represents the soil moisture values.

Figure 3.4: The time series soil moisture profile for different states.

- All four soil types can be distinguished based on the soil moisture pattern. The states that belong to the same soil layer have similar soil moisture profile. The four measurements belong to the four distinct layers.

- At the beginning all the states have low soil moisture and as precipitation increases the soil moisture starts to increase. The soil moisture near the top states increase rapidly. As water flows downward the bottom states tend to increase in soil moisture (Figure 3.4(b)).

- The saturated soil hydraulic conductivity of layers 1 to 4 are 1.04 cm/hr, 4.42 cm/hr, 0.26 cm/hr and 1.31 cm/hr, respectively. So when water enters the soil and flows through layer 1 into layer 2, the flow dynamics become faster since K_{sat} is much higher in layer 2. As water goes into layer 3 the flow dynamics slow down due to low K_{sat} value. Thus more water accumulates inside layer 3 increasing its soil moisture. Inside layer 4, the conductivity is higher and water again flows at a faster rate.

Since the actual parameters of the system are known in this simulation, this study enables us to investigate the requirements and the performance of the PEM algorithm. In the PEM algorithm the simulated soil moisture data at the four measurement locations corresponding to states $\mathcal{M}^{(4)} = \{5, 8, 13, 15\}$ as mentioned in Section 3.5.2 are considered to be known. The soil parameter values and all other states are unknown. We identify these unknown values and compare them with the known values to evaluate the performance of the PEM algorithm.

Three data sets were constructed from the simulated data.

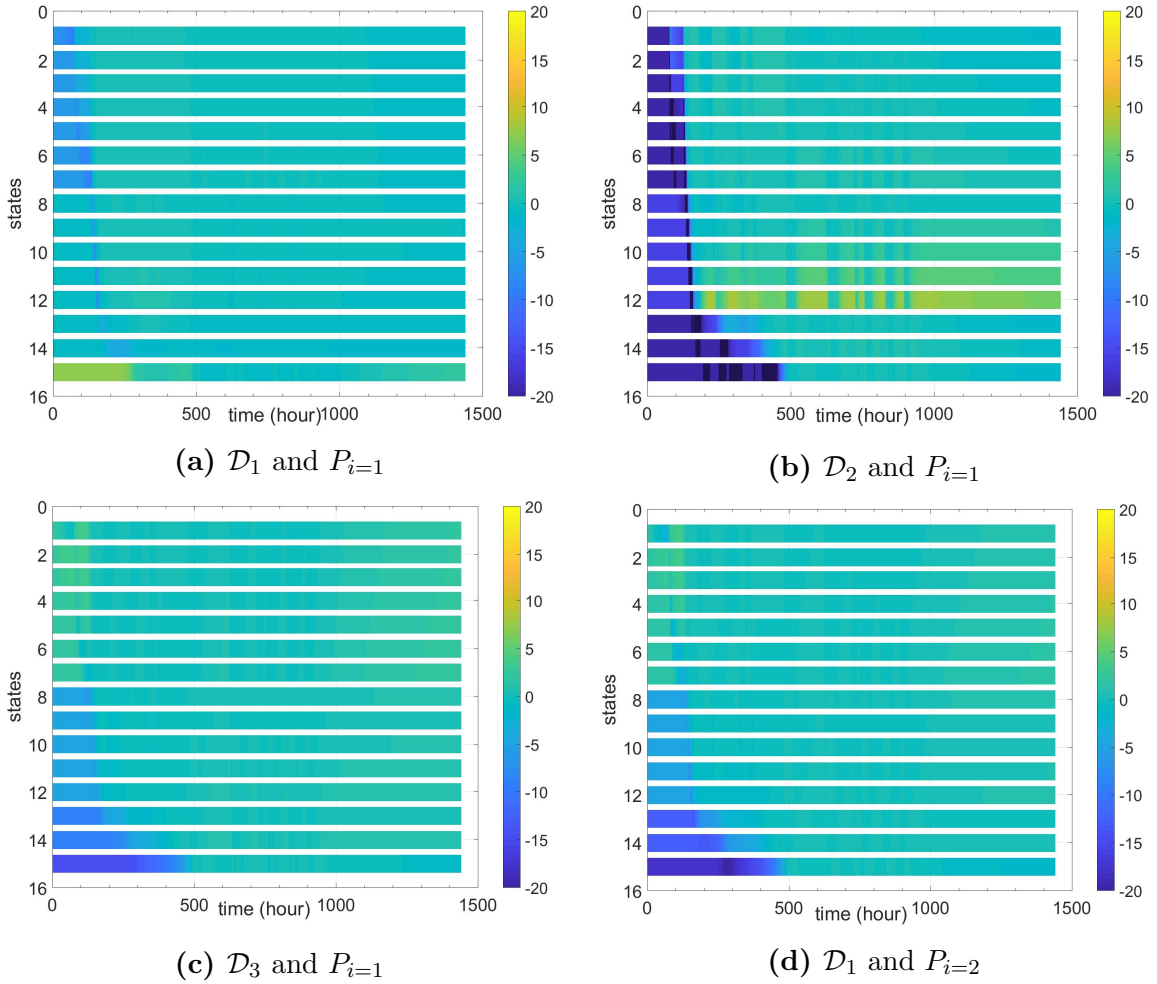
- Data set 1: $\mathcal{D}_1 = \{\mathcal{M}_k^{(4)} | k = 80, ..., 344\}$

- Data set 2: $\mathcal{D}_2 = \{\mathcal{M}_k^{(4)} | k = 500, ..., 764\}$

- Data set 3: $\mathcal{D}_3 = \{\mathcal{M}_k^{(4)} | k = 200, ..., 550\}$

where the simulation time (sample point) is denoted by k hours. Three cases are studied with the three data sets and initial parameter set $P_{i=1}$. The fourth case is studied with data set $\mathcal{D}_3$ and initial parameter set $P_{i=2}$. The estimated parameter values are shown in Figure 3.6. The percentage error between the simulated soil

moisture profile with true parameter values and estimated parameter values , $e_i = 100\dfrac{\bar{\theta}_i - \theta_i}{\theta_i}$ are shown in Figure 3.5, where $\bar{\theta}_i$ and θ_i are the true and estimated values at compartment $i = 1, \ldots, N$, respectively.



(a) $\mathcal{D}_1$ and $P_{i=1}$ **(b)** $\mathcal{D}_2$ and $P_{i=1}$

(c) $\mathcal{D}_3$ and $P_{i=1}$ **(d)** $\mathcal{D}_1$ and $P_{i=2}$

Figure 3.5: Error plot for the soil moisture profile generated by estimated parameter with respect to true parameter values. The color bar shows the error for each states at each sample time.

The estimated parameters for cases (a) $\mathcal{D}_1$ and $P_{i=1}$ (c) $\mathcal{D}_3$ and $P_{i=1}$ and (d) $\mathcal{D}_1$ and $P_{i=2}$ gave fair results since the percentage of error was usually low, less than 5%. For the first 500 hour, percentage of error was noticed to be higher at the bottom nodes. Since the actual simulated soil moisture in this region was very low (Figure 3.4) which corresponds to high nonlinearity, the error was higher. The error for case (b) $\mathcal{D}_2$ and $P_{i=1}$ was relatively higher (dark blue and yellow region in Figure 3.5) but was still within an acceptable range except the first 500 hours. The estimated parameter

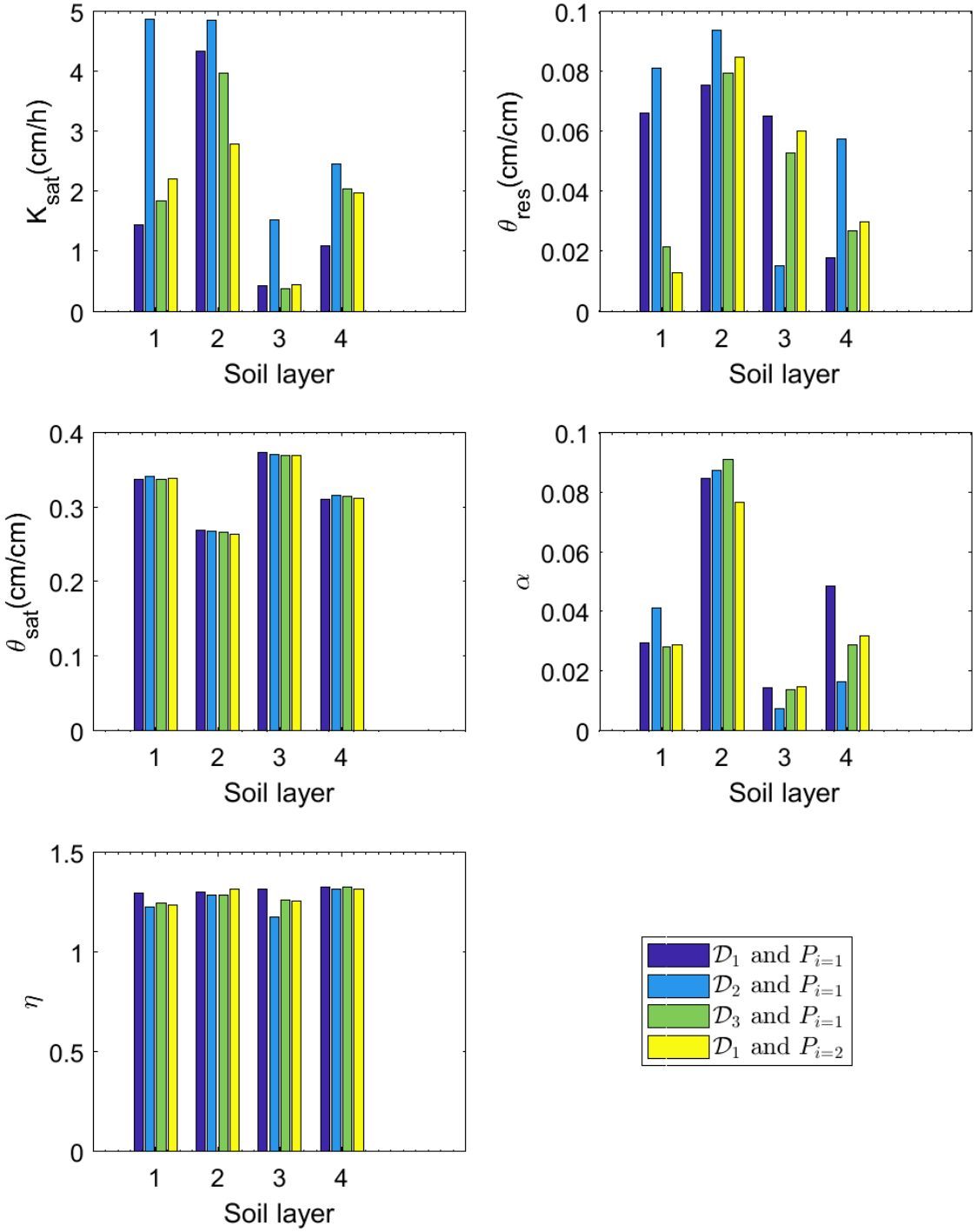


Figure 3.6: The estimated parameter values.

values are shown in Figure 3.6. It can be observed that the estimation of θ_{sat} and η showed less variance. The most sensitive parameters are K_{sat} and θ_{res}, and their estimation depends highly on the data quality. Among these estimated parameter sets the percentage fit for (a) $\mathcal{D}_1$ and $P_{i=1}$ was found maximum (around 99%), so this parameter set was selected for this system.

3.5.4 State estimation

Using the parameter set estimated with $\mathcal{D}_1$ and $P_{i=1}$, the EKF algorithm was executed to estimate the entire soil moisture profile based on four $\mathcal{M}^{(4)}$, three $\mathcal{M}^{(3)}$ and two $\mathcal{M}^{(2)}$ output measurements, respectively. For the measurement set with three measurements, two different combinations were used. One is the relatively more observable system given by the optimal sensor placement algorithm and denoted by $\mathcal{M}^{(3)} = \{3, 8, 15\}$. Another measurement combination denoted by $\mathcal{M}^{(3)*} = \{3, 13, 15\}$ is randomly chosen to demonstrate the importance of the optimal sensor placement algorithm. It can be observed from Figure 3.7 that $\mathcal{M}^{(3)*}$ performs very poorly with respect to $\mathcal{M}^{(3)}$. Also from the figure, it can be clearly observed that as the number of measurements are reduced the observability decreases and the estimated values deteriorate. It can be also noticed that some bias is present from the model plant mismatch for the unmeasured compartments.

3.6 Application to field data

3.6.1 Data description and problem formulation

The ground data used in this study was collected from the St.Albert Weather Station (Alberta Agriculture and Rural Development, 2014) located north of Edmonton, Alberta, Canada. According to the Canadian soil classification system the area under consideration has chernozenic soil comprised mainly of fine textured silt [90] as shown in Figure 3.8.

The hourly soil moistures at depths 5 cm, 20 cm, 50 cm and 100 cm were measured using Delta-t theta probes. The atmospheric forcing data was collected from the same weather station which includes precipitation, maximum and minimum air temperatures, relative humidity, incoming solar radiation and wind speed at 2 $meter$

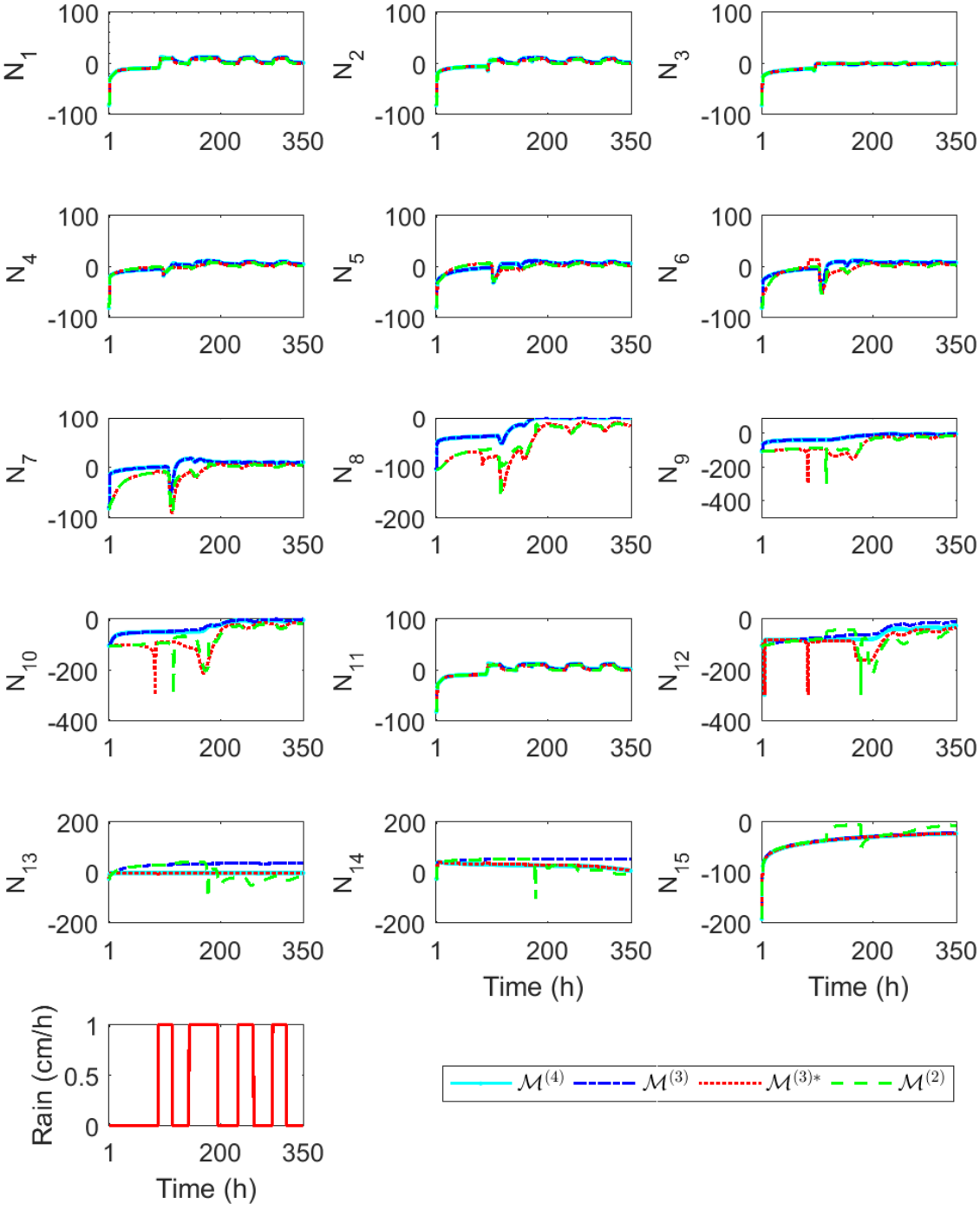


Figure 3.7: The error of state trajectory of true soil moisture and the estimated soil moisture at the 15 states given by EKF algorithm using the estimated parameters.

Figure 3.8: Black Chernozem soil. Photo taken from [2]

height on an hourly basis. The data was collected over a period of 5 months from 1st April, 2014 to 25 August, 2014 and is available online [91]. The data set is denoted by $\mathcal{R}_T$, where time $T = 1, 2, ..., 3505$ hours and is shown in Figure 3.9. It can be observed that the soil moisture at depth $100\,cm$ varied very slightly over the 4 months period. Only a small change in the soil moisture was observed around $k = 800$ hours. So instead of considering free drainage at the bottom, Dirichlet boundary condition was chosen at the bottom. The soil moisture value at $-100\,cm$ was chosen as the bottom boundary condition. Some irregular features can be observed in this data set. There was a gradual increase in soil moisture at the beginning though no rain was noticed. This could be due to the melting of snow during spring (April and May), so this part of data was not considered in parameter estimation. During the period $T = 1400 - 1700$ hours, there were sudden reduction of soil moisture at depth $-20\,cm$ and $-50\,cm$ right after rain. This indicates that there may be cracks inside soil and the water from these layers may have drained out of the system rapidly. Another pos-

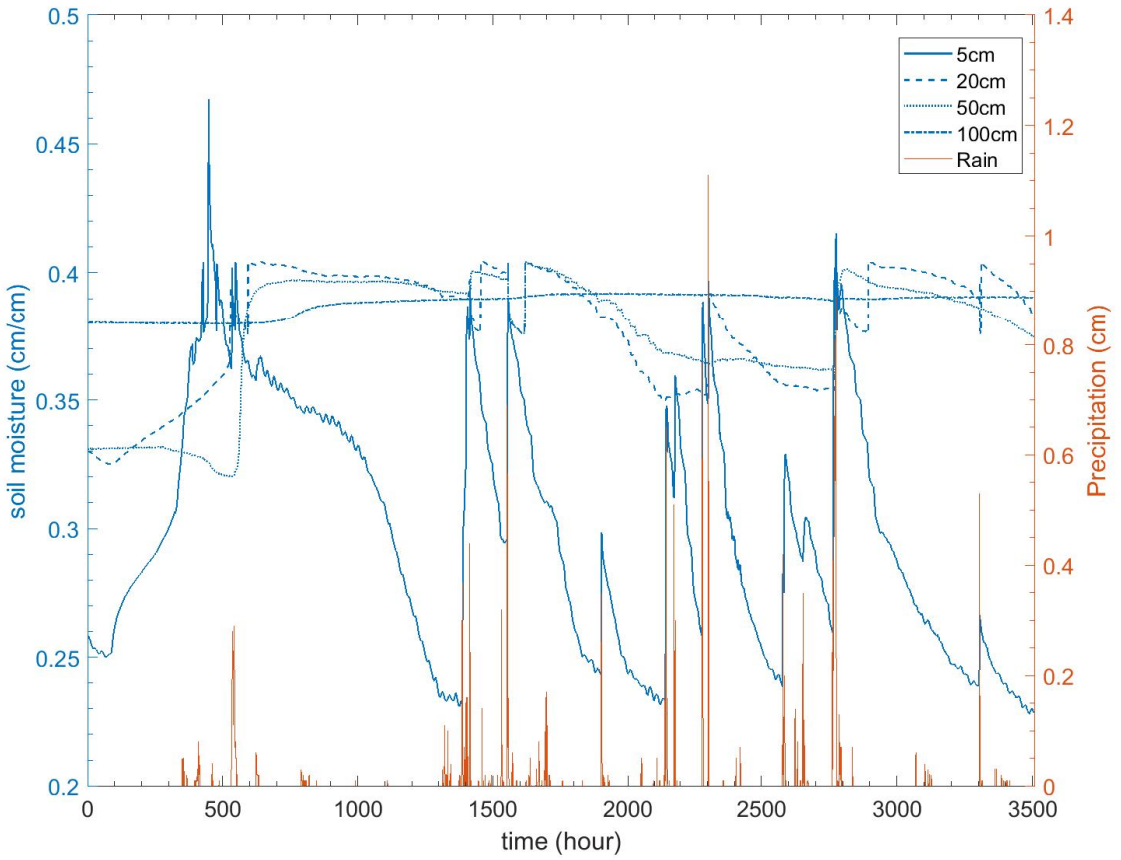


Figure 3.9: Real soil moisture profile for different rain inputs at St.Albert weather station.

sibility is that there may be significant horizontal flow of water which is not captured in the model. These events cause high model plant mismatch and makes the overall estimation process highly challenging.

The last measurement available is at depth 100 cm so we consider the overall system to be of 110 cm depth and use the simulated studies in the previous section to construct the soil profile for this real system. Thus, the system is discretized into 15 states similar to the simulated example as shown in Table 3.1. Though the soil type reported is only silt based on literature, to capture the variations among soil with depths and their dynamics, in this case we have also assumed four different homogeneous layers. Therefore the same observability analysis holds for this real system. With four, three and two measurements, the system is observable; but with one measurement the system becomes unobservable. Parameter identification is performed with the maximum number of available measurements that is all four measurements.

Then state estimation is performed with three and two measurements and the other measurements are used for validating the results.

3.6.2 Parameter and state estimation

First, we use the PEM algorithm to estimate the parameters of the model based on the collected data. A study by Carsel et al. [92] gives an estimate of Van Genuchten model parameters for different soil types, which is used to initialize the system. Though the soil type mentioned in the literature is silt type, the plot of real data shows different behaviors at different depths. The measurement at depth $-5\,cm$ has high θ_{sat} values and relatively faster dynamics. The measurements at depth $-20\,cm$ and $-50\,cm$ show slow dynamics and the θ_{sat} is at medium range. The bottom is not changing which can be due to shallow ground water, or impervious layers and horizontal flows. To capture this heterogeneity, we initialized the soil layers with soil property of silt, clay loam, silty clay loam and sandy clay loam as mentioned in [92]. The initial values as well as the estimated values are also shown in Table 3.4. The estimated parameters gave a percentage fit of about 50% for the top three layers which is considered to be acceptable. The estimated parameters are further used to estimate all the states using the EKF algorithm.

From the four available measurement data sets, measurement combinations $\mathcal{M}^{(4)}$, $\mathcal{M}^{(3)}$ and $\mathcal{M}^{(2)}$ are used in EKF to estimate the rest of the soil profile. The state estimation results are shown in Figure 3.10. It can be observed that when EKF is not used, the open-loop prediction of the states are prone to high error. This demonstrates the importance of simultaneous parameter and state estimation. Introduction of EKF with measurement set $\mathcal{M}^{(4)}$ reduced the error. The result also suggests that for the cases where we use three measurements ($\mathcal{M}^{(3)}$) for updates and estimate the remaining one, the EKF performed fairly well. In the case of two measurements ($\mathcal{M}^{(2)}$) the performance drops compared with the case with three measurements.

To further understand the results, an error analysis is performed. The percentage error is calculated only at the compartments where measurements are available. If at time k, the actual and estimated soil moisture at compartments corresponding to available measurements are given by $\bar{\theta}_M(k)$ and $\theta_M(k)$, respectively, where M denotes the compartments where the real measurements are available ($M = 3, 8, 13$ and 15),

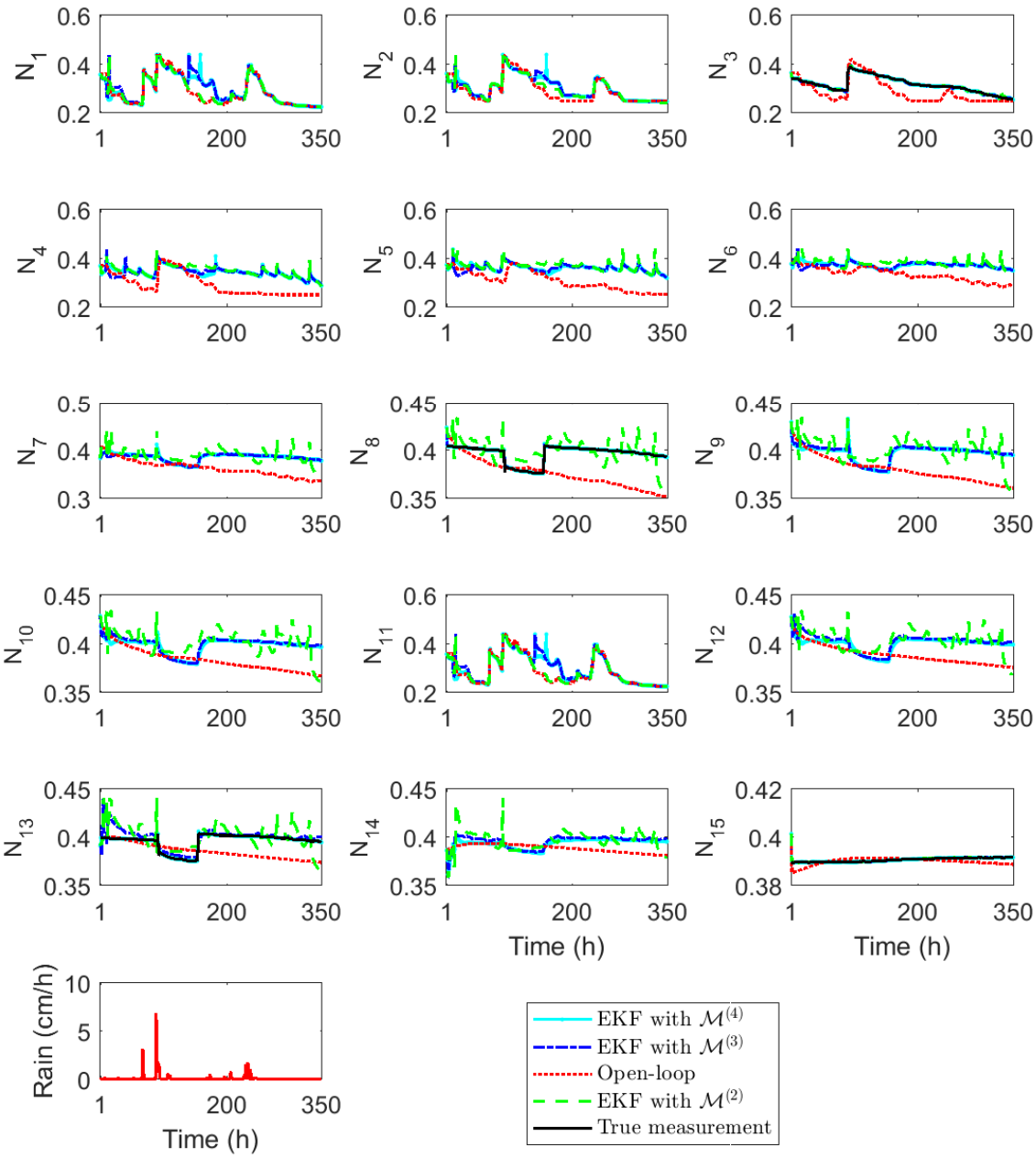


Figure 3.10: Estimated states and true states for real data.

Table 3.4: Parameters estimation for real scenario.

Parameter	Soil Layer	PEM estimation value	Initial Value
Ksat	1	0.238028	0.25
	2	0.389312	1.3
	3	0.683031	0.94
	4	0.185454	0.2
θ_R	1	0.0794364	0.1
	2	0.0959717	0.095
	3	0.0957004	0.089
	4	0.1	0.1
θ_S	1	0.439905	0.460
	2	0.4336	0.41
	3	0.441656	0.43
	4	0.401649	0.39
α	1	0.0322652	0.016
	2	0.0154362	0.019
	3	0.0219894	0.01
	4	0.00920637	0.015
η	1	1.13453	1.15
	2	1.17903	1.2
	3	1.18906	1.25
	4	1.09	1.09

the error at time k is given by:

$$e_M(k) = (\bar{\theta}_M(k) - \theta_M(k)) \tag{3.14}$$

The percentage error (%Error) at compartment M is calculated as follows:

$$\%Error_M = \frac{1}{n}\sum_{k=1}^{n} |\frac{e_M(k)}{\bar{\theta}_M(k)}|100 \tag{3.15}$$

where n is the total number of data samples. The results are shown in Table 3.5. In Figure 3.10, the true and estimated state trajectory for all compartments is shown. It can also be observed that the trajectory for the rest of compartments stays in agreement with the measured values.

3.7 Conclusions

A nonlinear state-space agro-hydrological model was used for parameter and state estimation. A detailed observability analysis based on rank test and degree of ob-

Table 3.5: The percentage error (%Error) observed at the measured compartments.

Measurement Set	% Error at depth -5 cm	% Error at depth -20 cm	% Error at depth -50 cm	% Error at depth -100 cm
$\mathcal{M}^{(4)}$	0.5272	0.0332	0.0130	0.0140
$\mathcal{M}^{(3)}$	0.5583	0.0294	1.0036	0.0155
$\mathcal{M}^{(2)}$	0.5879	2.5375	2.4590	0.0189
Open-loop	10.9373	5.5318	3.4634	5.0712

servability was performed to develop an observable system after discretization and to determine the optimum measurement locations. The results suggest that defining the 110 cm of soil profile with 15 states gives a fairly observable system with four measurements. This system remains observable even with fewer (three or two) measurements but becomes unobservable with only one measurement. The optimal sensor placement algorithm identifies the locations where the sensor should be placed in order to maintain an observable discrete system. This is important to ensure better parameter and state estimation as shown by the results.

Since the soil parameters were unknown, parameter estimation was performed using Prediction Error Method (PEM). The estimated parameter set was used to perform state estimation using the extended Kalman filter where the soil moisture values at unmeasured depths were estimated. The results have demonstrated a good estimation since the percentage error with respect to real data is less than 3%. Hence, the optimal sensor placement algorithm introduces a systematic observability analysis of the system using a model and thereby reduces the need for numerous experiment to identify optimal sensor locations. In summary, one can conclude that the proposed algorithm can be used for optimal sensor placement as well as the estimation of soil hydraulic parameters and soil moisture profile ensuring that the resulting system is observable and thereby produces reliable results.

Chapter 4

Closed-loop scheduling and control for precision irrigation

4.1 Introduction

In agriculture irrigation management, irrigation scheduling is typically performed in an open-loop fashion and is done only once at the beginning of a growing season. In this chapter, we study whether closed-loop scheduling with closed-loop control can lead to improved performance in terms of crop yield and water conservation in agriculture irrigation. To alleviate the computational complexity of the scheduler and the controller, instead of directly using the nonlinear model as discussed in Chapter 2, a linear parameter varying (LPV) model is identified for the scheduler and controller, respectively, which is discussed in Section 4.2. In the proposed scheduling and control scheme, both the scheduler and the controller are designed using model predictive control (MPC). The problem formulation and design is discussed in Section 4.3. The primary objective of the scheduler is to maximize the crop yield over a longer horizon, whereas the primary objective of the controller is to track the soil moisture reference calculated by the scheduler considering a relatively shorter horizon. The performance of the closed-loop scheduling scheme is evaluated against the traditional open-loop scheduling scheme under different scenarios.

Notation

Throughout this chapter, k is used to denote hourly sampled time instants, $\underline{k}$ is used to denote daily sampled time instants and $\bar{k}$ denotes weekly sampled time instants.

4.2 Scheduling and control model identification

The direct use of the agro-hydrological model, equations (2.12) and (2.13), in the scheduler and the controller is computational challenging due to the stiffness of the nonlinear system when the soil moisture is low, the high dimensions of the discretized system, and the large horizon covered by the scheduler. In order to overcome the computational challenge, we propose to identify an LPV model for the scheduler and the controller, respectively. The two LPV models can be identified based on the simulated input-output data from the agro-hydrological model. LPV models have been widely used to describe nonlinear processes due to its simple structure and its ability to approximate complex nonlinear systems [93, 94, 95, 96].

Scheduling model

The output of the scheduling model is the average root zone soil moisture content described as follows:

$$\theta_s = \frac{1}{N_r} \sum_{i=1}^{N_r} \theta_i \tag{4.1}$$

where θ_s is the root zone soil moisture, θ_i is determined by equation (2.13), and $N_r < N_D$ is the deepest node that contains the crop roots. Note that N_r changes over time as the length of roots grows (Figure 4.1).

The LPV model consists of a set of linear models that are connected with varying parameters which are functions of scheduling variables [97]. The scheduling variables are the ones that can be used to reflect the nonlinearity of the dynamics of the system. For the scheduling model, two scheduling variables are used. They are the soil moisture s_1 ($s_1 = \theta$) and the root depth s_2. The sampling period used in the scheduling model is one day.

For each scheduling variable, a few working points need to be determined and around those working points, linear models are identified. Let us assume that we have H working points for s_1 and G working points for s_2. These working points define a mesh with HG nodes. For each of these nodes, a linear model should be identified. Let us also use (i, j) ($i = 1, 2, ..., H,\ j = 1, 2, ..., G$) to denote the node defined by the i-th s_1 working point and the j-th s_2 working point. For each node

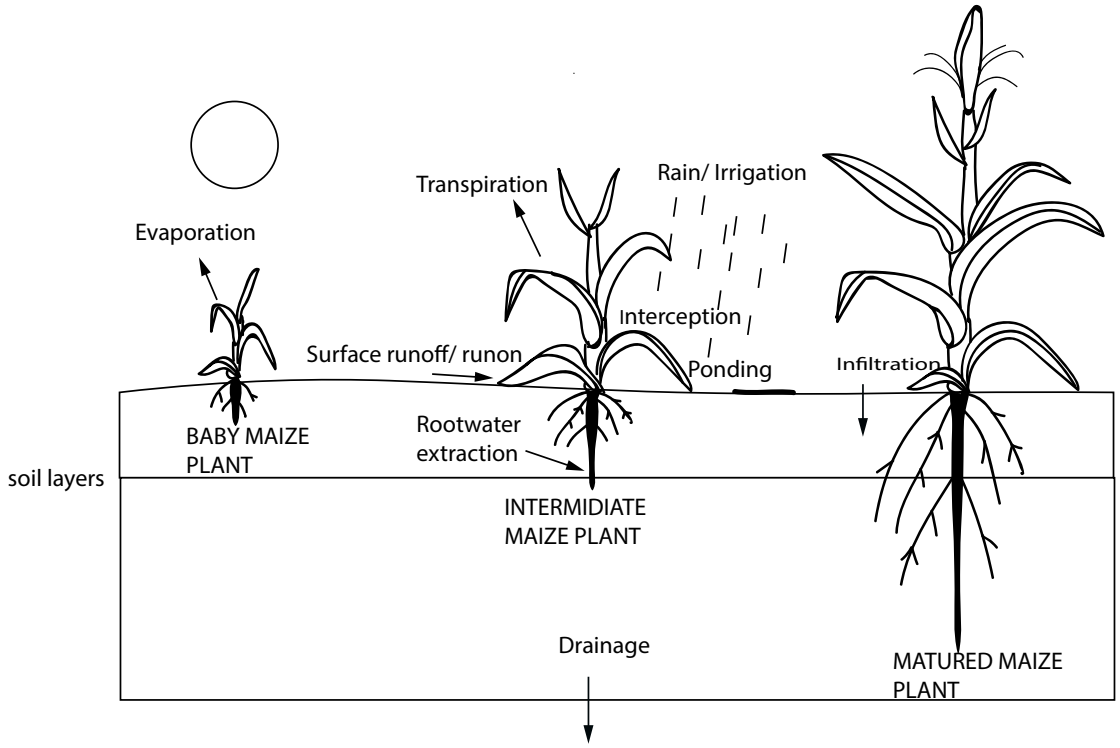


Figure 4.1: Maize at different growth stages.

(i, j), a linear model of the following form is identified:

$$x_{i,j}(\underline{k}) = -\sum_{p=1}^{n_a} a_p^{i,j} x_{i,j}(\underline{k} - p) + \sum_{q=1}^{n_b} b_q^{i,j} u(\underline{k} - q) + d^{i,j} \tag{4.2}$$

where $\underline{k}$ denotes the discrete time with a sampling period of a day, $x_{i,j}$ is the output of the linear model for (i, j), $a_p^{i,j}$ $(p = 1, 2, ..., n_a)$, $b_q^{i,j}$ $(q = 1, 2, ..., n_b)$ and $d^{i,j}$ are model parameters, u is the manipulated input. The model orders n_a and n_b are determined prior to the identification process. This model can also be expressed as follows:

$$x_{i,j}(\underline{k}) = \varphi_{i,j}^T(\underline{k})\beta_{i,j} \tag{4.3}$$

where $\varphi_{i,j}(\underline{k}) = [-x_{i,j}(\underline{k} - 1), ..., -x_{i,j}(\underline{k} - n_a), u(\underline{k} - 1), ..., u(\underline{k} - n_b), 1]^T$, and $\beta_{i,j} = [a_1^{i,j}, ..., a_{n_a}^{i,j}, b_1^{i,j}, ..., b_{n_b}^{i,j}, d^{i,j}]^T$.

With respect to s_1 (soil moisture), the linear models are combined together using Gaussian weighting functions as follows:

$$y_*^j(\underline{k}) = \sum_{i=1}^{H} \alpha_i(s_1(\underline{k} - 1)) x_{i,j}(\underline{k}) \tag{4.4}$$

where y_*^j is an intermediate variable and the weighting function α_i is given by:

$$\alpha_i(s_1(\underline{k})) = \frac{\omega_i(s_1(\underline{k}))}{\sum_{i=1}^{H} \omega_i(s_1(\underline{k}))} \tag{4.5}$$

where

$$\omega_i(s_1(\underline{k})) = \exp\left[-\frac{(-\text{sign}(s_1(\underline{k}) - s_{1,i}) + 1)}{2} \times \frac{-(s_1(\underline{k}) - s_{1,i})^2}{2\sigma_{i1}^2}\right] + \exp\left[-\frac{(-\text{sign}(s_1(\underline{k}) - s_{1,i}) + 1)}{2} \times \frac{-(s_1(\underline{k}) - s_{1,i})^2}{2\sigma_{i2}^2}\right] - 1$$

with $s_{1,i}$ the soil moisture value of the i-th pre-specified s_1 working point, σ_{i1} and σ_{i2} being two parameters that needs to be identified which are related to the validity width of the linear model [48].

With respect to s_2 (root length), the outputs obtained from equation (4.4) are further linearly interpolated to get the final output of the LPV model:

$$y(\underline{k}) = \frac{s_{2,j+1} - s_2(\underline{k})}{s_{2,j+1} - s_{2,j}} y_*^{j+1}(\underline{k}) + \frac{s_2(\underline{k}) - s_{2,j}}{s_{2,j+1} - s_{2,j}} y_*^j(\underline{k}) \tag{4.6}$$

given that the root length at time $\underline{k}$ lies between $s_{2,j}$ and $s_{2,j+1}$ (i.e., $s_{2,j} \leq s_2(\underline{k}) \leq s_{2,j+1}$). Note that the length of the roots grows according to the time after seeding and is considered as a function of time in this work. In equation (4.6), y is the average water content within the root zone given by the LPV model. Equation (4.6) is the model that will be used in the scheduler and may be written equivalently in the following state-space form:

$$x(\underline{k}+1) = Ax(\underline{k}) + Bu_s(\underline{k}) + d \tag{4.7}$$

$$y(\underline{k}) = C(s_1(\underline{k}-1), s_2(\underline{k}))x(\underline{k}) \tag{4.8}$$

where A, B are composed of elements in $\beta_{i,j}$ $(i = 1, 2, \ldots, H, j = 1, 2, \ldots, G)$, C depends on the two scheduling variables, d is composed of $d^{i,j}$, and u_s is composed of the current and previous daily irrigation amounts. The detailed identification of the above LPV model will be discussed in Section 4.4.

Control model

Since the controller uses a relatively much smaller prediction horizon (a week), we consider an LPV model with only the soil moisture (s_1) as the scheduling variable.

In this model, we consider the root length of the crop is constant within each control prediction horizon. The output of the control model is also the average soil water content within the root zone as defined in equation (4.1). The sampling period of the LPV model for control purpose is also a day. The LPV takes the following form:

$$\tilde{x}_i(\underline{k}) = -\sum_{p=1}^{n_a} a_p^i \tilde{x}_i(\underline{k}-p) + \sum_{q=1}^{n_b} b_q^i u(\underline{k}-q) + d^i \tag{4.9a}$$

$$\tilde{y}(\underline{k}) = \sum_{i=1}^{G} \alpha_i(\tilde{s}_1(\underline{k}-1))\tilde{x}_i(\underline{k}) \tag{4.9b}$$

where $\tilde{x}_i$ is the output of a sub-model and $\tilde{y}$ is the final output of the LPV model, a_p^i $(p = 1, 2, ..., n_a)$, b_q^i $(q = 1, 2, ..., n_b)$ and d^i are model parameters, α_i is a weighting factor determined following equation (4.5). At different crop growing stages, the above control model should be re-identified to account for the growth of the crop. The above model can also be written equivalently in the following state-space form:

$$\tilde{x}(\underline{k}+1) = \tilde{A}\tilde{x}(\underline{k}) + \tilde{B}u_c(\underline{k}) + \tilde{d} \tag{4.10}$$

$$\tilde{y}(\underline{k}) = \tilde{C}(s_1(\underline{k}-1))\tilde{x}(\underline{k}) \tag{4.11}$$

where $\tilde{A}$, $\tilde{B}$ are composed of a_p^i and b_q^i $(i = 1, 2, \ldots, H)$, $\tilde{C}$ depends on the scheduling variable, $\tilde{d}$ is composed of d^i, and u_c is composed of the current and previous daily irrigation amounts. The details will be discussed also in Section 4.4.

4.3 Problem formulation and proposed design

4.3.1 Closed-loop scheduling and control problem formulation

In this section, we describe the proposed closed-loop scheduling and control framework for precision irrigation. The scheduler and the controller form a hierarchical decision making system. In this work, both the scheduler and the controller will be designed in the framework of MPC. The primary objective of the scheduler is to maximize the crop yield at the end of the entire growing season considering a much longer horizon that covers the entire growing season and the horizon shrinks as time moves. It uses historical weather data, available weekly weather forecast and soil moisture measurement from the field to calculate the soil moisture reference trajectories for

the lower layer controller. The primary objective of the controller is to track the soil moisture reference calculated by the scheduler. It uses a fixed prediction horizon of a week and takes into account the soil moisture measurement and daily weather forecast in its irrigation command calculation. A schematic of the closed-loop scheduling and control framework is shown in Figure 4.2.

4.3.2 Design of scheduler and controller

Design of the scheduler

As explained earlier, the main objective of the scheduler is to minimize the crop yield deficiency while reducing the irrigation water usage. The crop yield is modeled as a function of the potential evapo-transpiration and the actual evapo-transpiration as follows [52]:

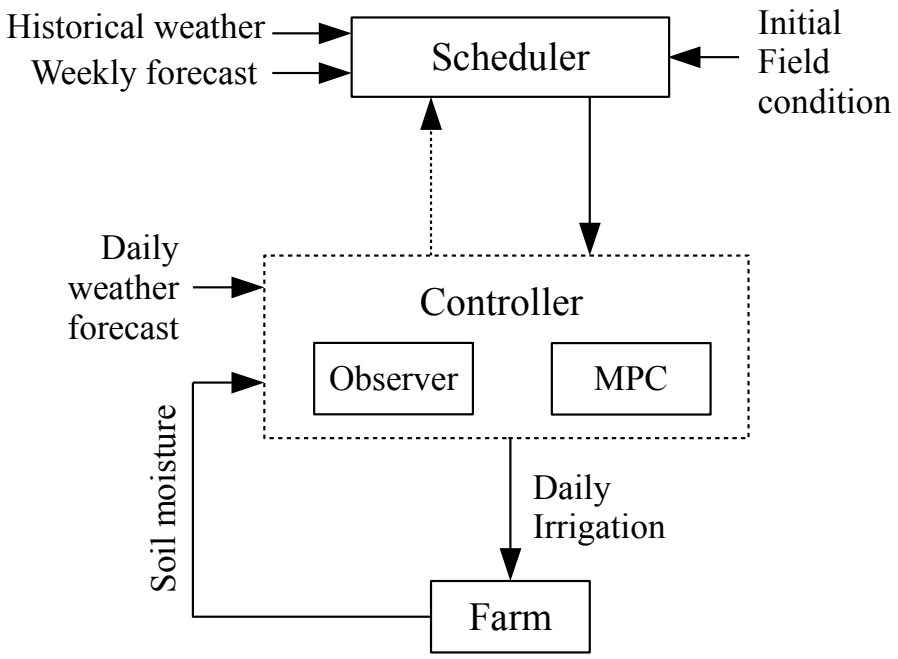
$$(1 - \frac{Y_a}{Y_p}) = \sum_{\bar{k}=1}^{T} \mathcal{K}_y(\bar{k})(1 - \frac{ET_a}{ET_p}) \tag{4.12}$$

where Y_a is the actual yield, Y_p is the potential maximum yield, ET_a is the actual evapo-transpiration and ET_p is the potential evapo-transpiration, and $\mathcal{K}_y(\bar{k})$ is the crop sensititivity for the growing period at time $\bar{k}$. When the actual yield is equal to the potential yield, equation (4.12) takes its minimum value zero. According to the definition of the potential and the actual evapo-transpirations of a crop as given in equations (2.19) and (2.20), the following relation can be obtained to describe the relation between the crop yield and soil moisture:

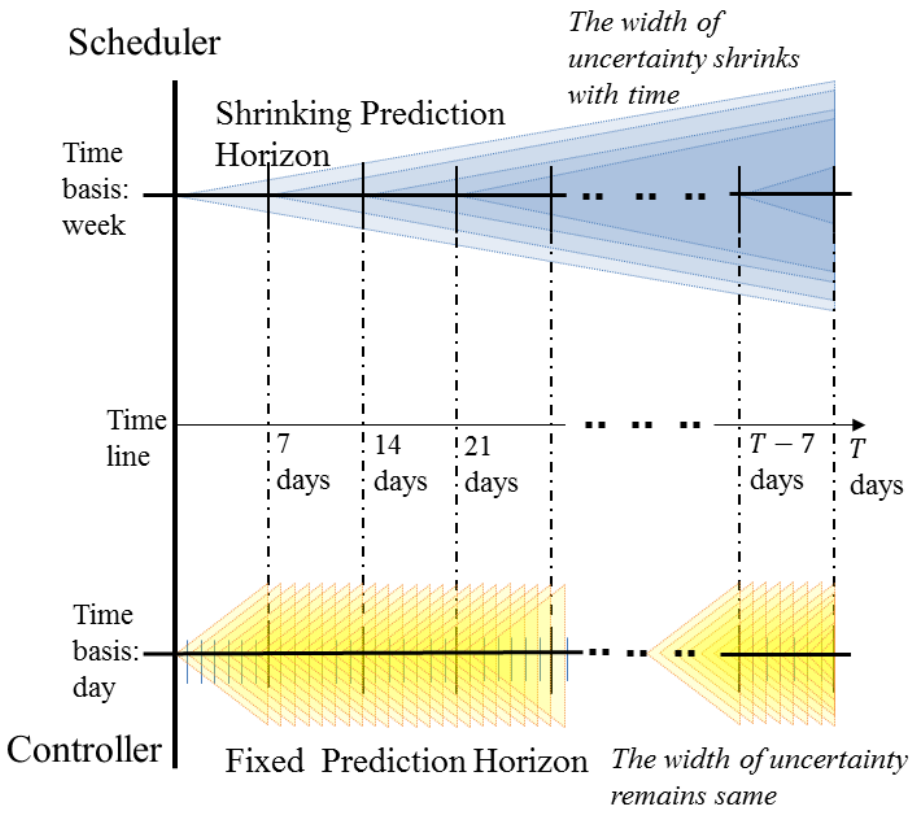
$$(1 - \frac{Y_a}{Y_p}) = \sum_{\bar{k}=1}^{T} \mathcal{K}_y(\bar{k})(1 - \mathcal{K}_s(\theta(\bar{k}))) \tag{4.13}$$

In this work, following the work of Van Dam et al. [79], a simple maize model is used to calculate the evapo-transpiration amounts at different growing stages.

From equation (4.13), it can be seen that crop yield relates to the water stress factor $\mathcal{K}_s$. The stress factor characterizes the suppression of root water update due to either drought or insufficient aeration. Figure 4.3 shows a typical relation of $\mathcal{K}_s$ and soil moisture [3]. When the soil moisture θ is between θ_L (limiting point) and θ_a (anaerobic point), there is no water stress ($\mathcal{K}_s = 1$). When the soil moisture θ is below θ_L but larger than θ_w (wilting point), water stress takes place due to drought.



(a) Process block diagram



(b) The design of prediction horizon.

Figure 4.2: Structure of the closed-loop scheduling and control system.

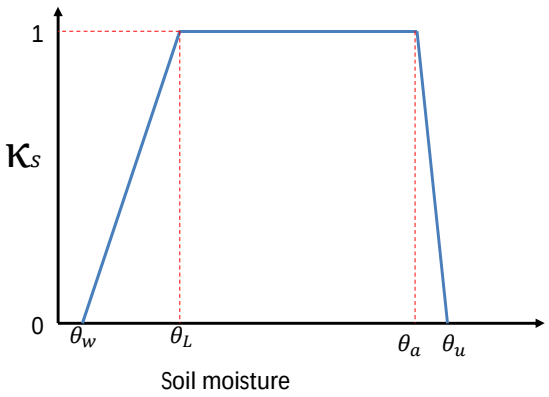


Figure 4.3: The relation of soil moisture content and water stress factor [3].

When the soil moisture θ is above θ_a but less than θ_u, there is crop water stress due to insufficient aeration; when is above θ_u, crop transpiration stops. For all other soil moisture region, root water extraction is completely suppressed. Therefore, it is also important to ensure that the soil moisture is always between θ_w and θ_u. Specifically, the value of the water stress factor $\mathcal{K}_s$ is determined as follows [3]:

$$\mathcal{K}_s = \begin{cases} \dfrac{\theta - \theta_w}{\theta_L - \theta_w}, & \text{for} \quad \theta_w \geq \theta \leq \theta_L \\ 1, & \text{for} \quad \theta_L < \theta < \theta_a \\ \dfrac{\theta - \theta_a}{\theta_U - \theta_a}, & \text{for} \quad \theta_a \geq \theta \leq \theta_U \\ 0, & \text{for} \quad \theta \leq \theta_w \text{ or } \theta \geq \theta_U \end{cases} \tag{4.14}$$

The scheduler optimization problem is evaluated every week. It uses a prediction horizon from the current week to the end of the growing season T. Within each week, the irrigation is considered to be constant. For week $\bar{k}$, the optimization problem is shown below:

$$\min_{u,\varepsilon_1,\varepsilon_2} \left\{ \left(1 - \frac{Y_a}{Y_p}\right)^2 Q_{s1} + \sum_{l=\bar{k}}^{T-1} u(l)^2 Q_{s2} + \sum_{l=\bar{k}}^{T} \left[\varepsilon_1(l)^2 Q_{s3} + \varepsilon_2(l)^2 Q_{s4}\right] \right\} \tag{4.15a}$$

$$\text{s.t.} (1 - \frac{Y_a}{Y_p}) = \sum_{l=1}^{\bar{k}} \mathcal{K}_y(l)(1 - \mathcal{K}_s(\theta_s(l))) + \sum_{l=\bar{k}+1}^{T} \mathcal{K}_y(l)(1 - \mathcal{K}_s(y(l))) \tag{4.15b}$$

$$x(l+1) = A^7 x(l) + \sum_{\underline{k}=0}^{6} A^{6-\underline{k}} B u(l) + \sum_{\underline{k}=0}^{6} A^{6-\underline{k}} d, \; l = \bar{k}, \ldots, T-1 \tag{4.15c}$$

$$y(l) = Cx(l), \; l = \bar{k}, \ldots, T \tag{4.15d}$$

$$x(\bar{k}) = \bar{x}(\bar{k}) \tag{4.15e}$$

$$0 \leq u(l) \leq u_{\max},\ l = k, \ldots, T-1, \tag{4.15f}$$

$$y(l) - \varepsilon_1 < \theta_u,\ y(l) + \varepsilon_2 > \theta_w,\ l = \bar{k}, \ldots, T \tag{4.15g}$$

$$\varepsilon_1(l) \geq 0,\ \varepsilon_2(l) \geq 0,\ l = \bar{k}, \ldots, T \tag{4.15h}$$

In the optimization problem (4.15), equation (4.15a) defines the cost function to be minimized. In equation (4.15a), the first term indicates the crop yield deficiency at the end of the growing season, the second term measures the irrigation amount in the remaining growing season, the last summation term includes two slack variables that are used to handle constraints on the soil moisture in equation (4.15g), and Q_{s1}, Q_{s2}, Q_{s3} are positive weighting factors. Equation (4.15b) is the model used to calculate the crop yield deficiency. Note that in equation (4.15b) the first part on the right-hand-side is calculated based on actual soil water content measurements and the second part is calculated based on soil water content predicted in the scheduler. Equations (4.15c)-(4.15d) are the LPV model for scheduling evaluated for a week with constant input. Note that the sampling period of model identified in the previous subsection is a day. The scheduler optimization problem, however, evaluates once every week. Within a week, the input is held constant. Equation (4.15c) reflects this type of consideration. Equation (4.15e) defines the initial condition of the scheduling optimization problem and $\bar{x}(\bar{k})$ denotes the current state measurement. Note that we assume that the state is measured. Equation (4.15f) requires that all the irrigation should be non-negative and is less than the maximum allowed limit $u_{\max}$. Equation (4.15g) imposes the constraint such that $\theta_w < y < \theta_u$ as soft constraints through the introduction of two slack variables ε_1 and ε_2. Equation (4.15h) implies that the two slack variables are non-negative. The weighting factors Q_{s3} and Q_{s4} should be tuned to be relatively higher to avoid constraint violation.

The solution to the above scheduling optimization problem is a sequence of optimal irrigation amount from week $\bar{k}$ to T. Let us denote the optimal weekly irrigation as $u^*(l|\bar{k}), l = \bar{k}, \ldots, T-1$. Let us also denote the corresponding weekly optimal soil moisture as $y^*(l|\bar{k}), l = \bar{k}, \ldots, T-1$. The first two values of the optimal irrigation values $u^*(\bar{k}|\bar{k}), u^*(\bar{k}+1|\bar{k})$ and the corresponding soil moisture values $y^*(\bar{k}|\bar{k}), y^*(\bar{k}+1|\bar{k})$ are sent to the controller as references. At the next week, the scheduling optimization problem is re-evaluated with new soil moisture measurement.

Design of the controller

In the controller design, the control LPV model is first augmented with an output disturbance to account for model-plant mismatch. Then, an observer is designed to estimate the model-plant mismatch. The augmented model is as follows:

$$\tilde{x}(\underline{k}+1) = \tilde{A}\tilde{x}(\underline{k}) + \tilde{B}u_c(\underline{k}) + \tilde{d} \tag{4.16}$$

$$p(\underline{k}+1) = p(\underline{k}) \tag{4.17}$$

$$\tilde{y}(\underline{k}) = \tilde{C}\tilde{x}(\underline{k}) + p(\underline{k}) \tag{4.18}$$

Based on the above augmented model, the following observer is designed to estimate the output disturbance:

$$\hat{x}(\underline{k}|\underline{k}) = \hat{x}(\underline{k}|\underline{k}-1) + L_x(\underline{k})(\bar{y}(\underline{k}) - \hat{y}(\underline{k}|\underline{k}-1)) \tag{4.19}$$

$$\hat{p}(\underline{k}|\underline{k}) = \hat{p}(\underline{k}|\underline{k}-1) + L_p(\underline{k})(\bar{y}(\underline{k}) - \hat{y}(\underline{k}|\underline{k}-1)) \tag{4.20}$$

$$\hat{x}(\underline{k}+1|\underline{k}) = \tilde{A}\hat{x}(\underline{k}|\underline{k}) + \tilde{B}u_c(\underline{k}) + \tilde{d} \tag{4.21}$$

$$\hat{p}(\underline{k}+1|\underline{k}) = \hat{p}(\underline{k}) \tag{4.22}$$

$$\hat{y}(\underline{k}|\underline{k}) = \tilde{C}\hat{x}(\underline{k}+1|\underline{k}) + \hat{p}(\underline{k}+1|\underline{k}) \tag{4.23}$$

where $\hat{x}$, $\hat{p}$, $\hat{y}$ and $\bar{y}$ denote the estimated state, estimated disturbance, estimated output, and output measurement, respectively. The observer gains L_x and L_p are determined as discussed in Mao et al. [48].

The primary objective of the controller is to track the soil moisture and total irrigation water consumption references calculated by the scheduler. The controller is also designed in the framework of MPC. The controller is evaluated every day with updated soil moisture measurements. The controller uses a fixed prediction horizon of a week (i.e., 7). At day $\underline{k}$, the controller is formulated as follows:

$$\min_{u,\xi_1,\xi_2} \sum_{l=\underline{k}}^{\underline{k}+7} \left[\hat{y}(l) - y^*(l)\right]^2 Q_c + \left[\sum_{l=\underline{k}}^{\underline{k}+6} u(l) - \sum_{l=\underline{k}}^{\underline{k}+6} u^*(l)\right]^2 R_c + \sum_{l=\underline{k}}^{\underline{k}+7} \left[\xi_1(l)^2 Q_{c1} + \xi_2(l)^2 Q_{c2}\right] \tag{4.24a}$$

$$\text{s.t.} \tilde{x}(l+1) = \tilde{A}\tilde{x}(l) + \tilde{B}u(l) + \tilde{d},\ l = \underline{k}, \ldots, \underline{k}+6 \tag{4.24b}$$

$$p(\underline{k}+1) = p(\underline{k}) \tag{4.24c}$$

$$\tilde{y}(l) = \tilde{C}\tilde{x}(l) + p(l),\ l = \underline{k}, \ldots, \underline{k}+7 \tag{4.24d}$$

$$\tilde{x}(\underline{k}) = \hat{x}(\underline{k}|\underline{k}), \quad p(\underline{k}) = \hat{p}(\underline{k}|\underline{k}) \tag{4.24e}$$

$$0 \leq u(l) \leq u_{\max}/7,\ l = \underline{k}, \ldots, \underline{k}+6 \tag{4.24f}$$

$$\tilde{y}(l) - \xi_2(l) \leq \theta_u,\ \tilde{y}(l) + \xi_1(l) \geq \theta_w,\ \ l = \underline{k}, \ldots, \underline{k}+7 \tag{4.24g}$$

$$\xi_1(l) \geq 0,\ \xi_2(l) \geq 0,\ l = \underline{k}, \ldots, \underline{k}+7 \tag{4.24h}$$

In the optimization problem (4.24), equation (4.24a) defines the cost function to be minimized. In equation (4.24a), y^* and u^* denote the references from the scheduler. The first term requires that the predicted average soil moisture content within the root zone should be close to the reference from the scheduler, the second term requires that the total water consumption over a week is close to the predicted consumption of the scheduler, and the last terms contain two slack variables. In equation (4.24a), Q_c, Q_{c1}, Q_{c2} and R are positive weighting factors. Equation (4.24b) is the LPV model for control. Equation (4.24c) implies that a constant output disturbance is considered in the optimization of the controller and the disturbance value is given by the observer as specified in equation (4.24e). This constant output disturbance is added to the output equation as specified in equation (4.24d). Equation (4.24e) defines the initial condition of the control optimization problem and $\hat{x}$ is from the observer. Equation (4.24f) requires that all the irrigation should be non-negative and is smaller than or equal to one seventh of the weekly upper limit. Equation (4.24g) imposes the constraint such that $\theta_w < \tilde{y} < \theta_u$ as soft constraints. Equation (4.24h) implies that the two slack variables are non-negative.

The solution to the optimization problem (4.24) is a sequence of optimal irrigation amount for time $l = \underline{k}, \ldots, \underline{k}+6$ denoted by $u^*(l|\underline{k})$, $l = \underline{k}, \ldots, \underline{k}+6$. At time $\underline{k}$, only the first element from the optimal daily irrigation input is applied to the field. The rest are discarded. At the next day $\underline{k}+1$, the optimization problem is solved again.

4.4 Results and discussion

4.4.1 LPV model identification and validation

In the identification of the scheduling LPV model, three working points are considered for each of the two scheduling variables ($H = 3$, $G = 3$). In the input-

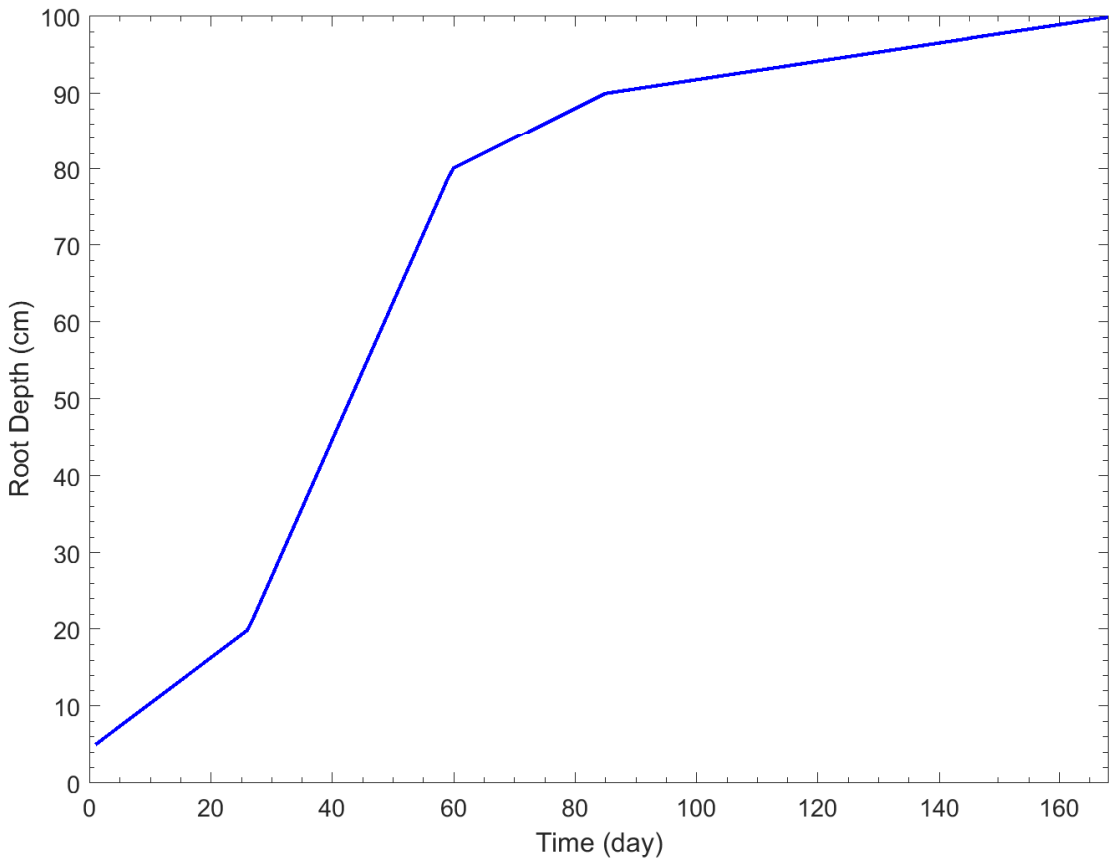


Figure 4.4: Root length of maize according to time after seeding.

output data generation, the three s_2 working points are $s_2 = [s_{2,1}, s_{2,2}, s_{2,3}]^T = [10.3571, 55.3571, 99.881]^T$ cm. These working points are picked according to the root length of maize at different growing stages (see Figure 4.4). Specifically, the first working point of s_2 corresponds to the root length about ten days after seeding and is picked to represent the baby plant. The second working point is the mean value of the root length of the rapid growing stage and the last working point is the root length at the fully matured stage. For each working point of s_2, three different s_1 working points are considered. For $s_{2,1}$, the three s_1 working points are $s_1^{\{1\}} = [s_{1,1}^{\{1\}}, s_{1,2}^{\{1\}}, s_{1,3}^{\{1\}}]^T = [0.0755, 0.0936, 0.1792]^T$. For $s_{2,2}$, the three s_1 working points are $s_1^{\{2\}} = [s_{1,1}^{\{2\}}, s_{1,2}^{\{2\}}, s_{1,3}^{\{2\}}]^T = [0.0789, 0.1211, 0.1988]^T$. For $s_{2,3}$, the three s_1 working points are $s_1^{\{3\}} = [s_{1,1}^{\{3\}}, s_{1,2}^{\{3\}}, s_{1,3}^{\{3\}}]^T = [0.0835, 0.1176, 0.1564]^T$. For each s_2 working point, the three s_1 working points are determined according to three (small,

medium, large) steady-state irrigation amounts, which are picked based on extensive simulations and trail-and-error. In determining these irrigation amounts, one consideration is the possible soil moisture range of the system (about 0.075 - 0.35) and another consideration is the nonlinearity feature of the agro-hydrological system. The orders of the local models n_a and n_b are both picked to be one based on the results of Mao et al. [48].

The input and output data for the LPV identification generated from the agro-hydrological model is shown in Figure 4.5. In generating the data, the daily weather

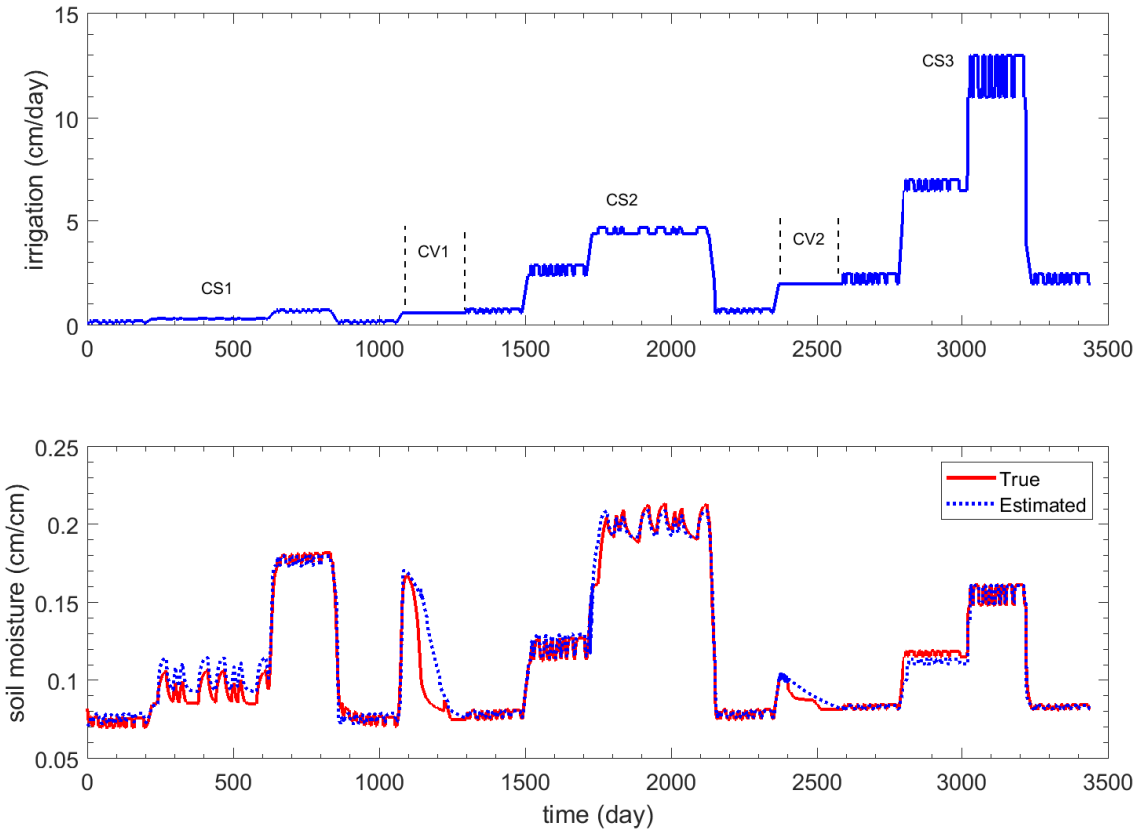


Figure 4.5: Input-output data for model identification with two scheduling variables.

condition is kept constant with no rain. Binary input signals are used around the working points. During the transition from one working point to another of s_1 (soil moisture), constant inputs are used. In the top plot of Figure 4.5, $CS1$, $CS2$, $CS3$ correspond to the three working points of s_2 (root depth). During these regions, binary inputs are used. Region $CV1$ and $CV2$ correspond to transition periods from $s_{2,1}$ to $s_{2,2}$ and from $s_{2,2}$ to $s_{2,3}$. In these two regions, constant inputs are used to

Table 4.1: Estimated local model parameters and weights for the scheduling LPV model.

Soil moisture \Root length		1	2	3
1	a_1	-0.7749	-0.6337	-0.4202
	b_1	0.0152	0.0074	0.0022
	d_1	0.0144	0.0235	0.0433
2	a_1	-0.8745	-0.7334	-0.1006
	b_1	0.0542	0.0083	0.0054
	d_1	-0.0045	0.0106	0.0643
3	a_1	-0.5701	-0.8873	-0.3665
	b_1	0.0264	0.0065	0.0036
	d_1	0.0575	-0.0072	0.0556
Gaussian weights	σ_1^2	0.0201	0.0193	0.0194
	σ_{21}^2	0.0094	0.0129	0.0157
	σ_{22}^2	0.0200	0.0210	0.0209
	σ_3^2	0.0205	0.0203	0.0186

show the effect of crop root zone on soil moisture. As crop root length increases (from $CS1$ to $CS3$), more irrigation is required to maintain the soil moisture as shown in the top plot of Figure 4.5. Following the approach in Mao et al. [48], the parameters of the LPV model are identified and shown in Table 4.1. The bottom plot of Figure 4.5 shows the predicted trajectory and the actual trajectory generated by the agro-hydrological model. The LPV model overall has a good fit over the entire period with the percentage difference between the predicted value of the LPV model and the actual value of the agro-hydrological model about 5% or less except in the region $CV1$ which corresponds to the rapid growing stage (whose percentage difference is about 18%). To further validate the developed LPV model over the crop growing season, an open-loop simulation of the agro-hydrological model is performed. Fixed weather condition is considered with no rain, and an input irrigation of 2.5 cm/day is considered. The result is shown in Figure 4.6. The overall percentage difference between the predicted value given by the LPV model and the value generated by the agro-hydrological model is about 27% with most of it contributed from the period between day 23 and day 45 which corresponds to the rapid growing stage. In the rest of the period, the percentage difference is around 15%. Given the complexity of the agro-hydrological system and the use of feedback, the mismatch between the LPV model and the agro-hydrological model is considered to be acceptable. Note that the

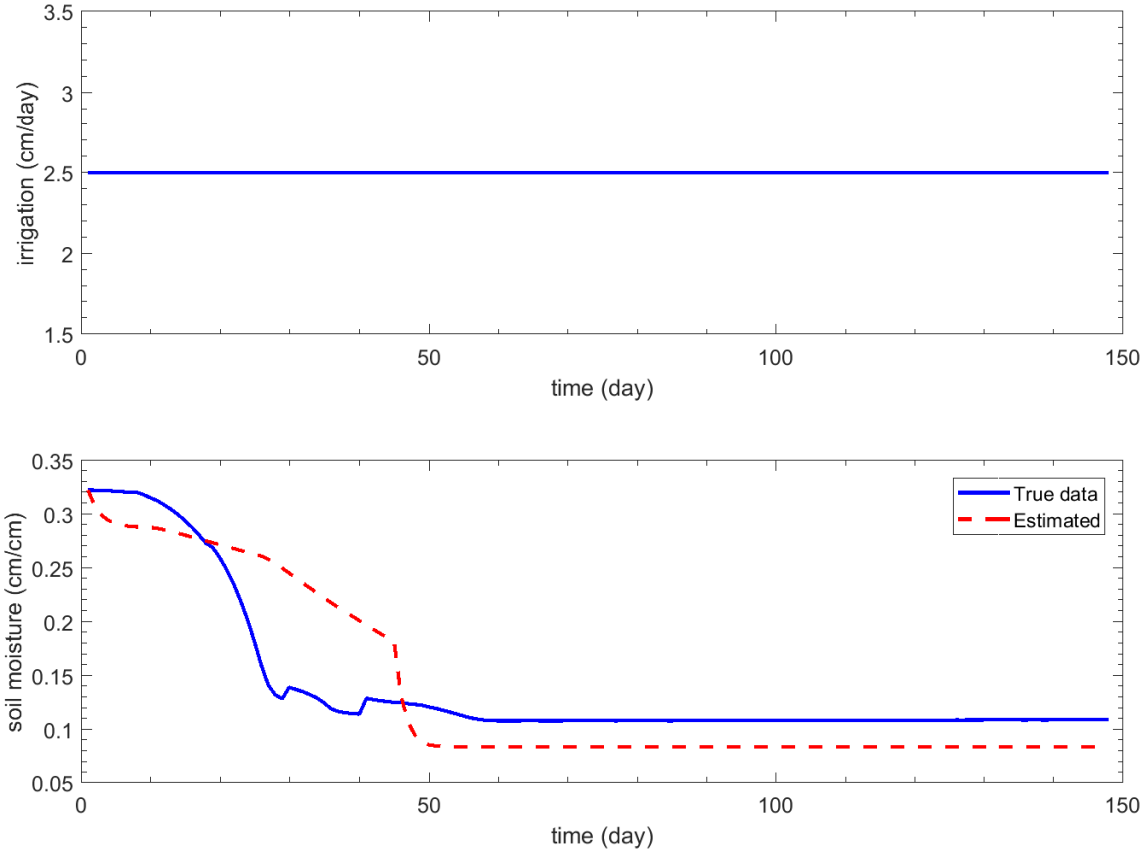


Figure 4.6: Validation of the scheduling LPV model.

performance of the LPV model may be improved by including more working points.

In the controller, the three local linear models of each root length working point are weighted using the corresponding Gaussian weights in Table 3.1 and then is used in the controller design. For example, in $CS1$ and upto the middle of $CV1$, the local models of root length working point 1 are weighted and used in the MPC controller. The switching between root length working points is based on the root length.

4.4.2 Scheduling and control simulation results

The proposed closed-loop scheduling and control is applied to the agro-hydrological system under different scenarios. The proposed closed-loop scheduling approach is compared with an open-loop scheduling scheme in which the scheduling optimization is only solved once at the beginning and applied throughout the growing season. In the open-loop scheduling scheme, the same closed-loop controller is used. The crop

Table 4.2: The values of crop sensitivity factor.

Crop sensitivity factor ($\mathcal{K}_y$)	Development stage (DVS)
0.4000	0
0.4000	0.2375
0.9000	0.2500
0.9000	0.5625
1.5000	0.5750
1.5000	0.7500
2.3000	0.7625
2.3000	1.3750
0.5000	1.3875
0.5000	1.6250
0.2000	1.6375
0.2000	2.0000

growing season is considered to be 23 weeks from seeding. The values of the crop sensitivity factor $\mathcal{K}_y$ at different crop growth stages (from 0 to 2 with 0 indicating initial seeding stage and 2 implying mature stage) for maize are shown in Table 4.2 [98]. The weekly irrigation upper limit is $u_{\max} = 70\ cm/week$, which implies that the daily irrigation limit is $10\ cm$.

Scenario 1: Favorable weather

In the first scenario, we consider that the amount of rain in the current crop season is similar to that of the historical weather data. In this scenario, we consider that the actual total rain in the crop growing season is $164.05\ cm$ but the historical weather data suggests the total rain over the season to be $154.50\ cm$.

In the closed-loop scheduling, the scheduler is updated at the end of each week with an update weather forecast for the next two weeks and the current soil moisture measurement. While in the open-loop scheduling, only the historical weather data is used at the beginning of the season. As for the controller, in both schemes, the field soil moisture is updated every day with field measurements and the weather data is also updated for the next seven days. Uncertainties are considered in weather forecast.

First, we consider a water conservation case. That is, in the scheduler design, more emphasis is put on water conservation. In this case, the weighting factors are

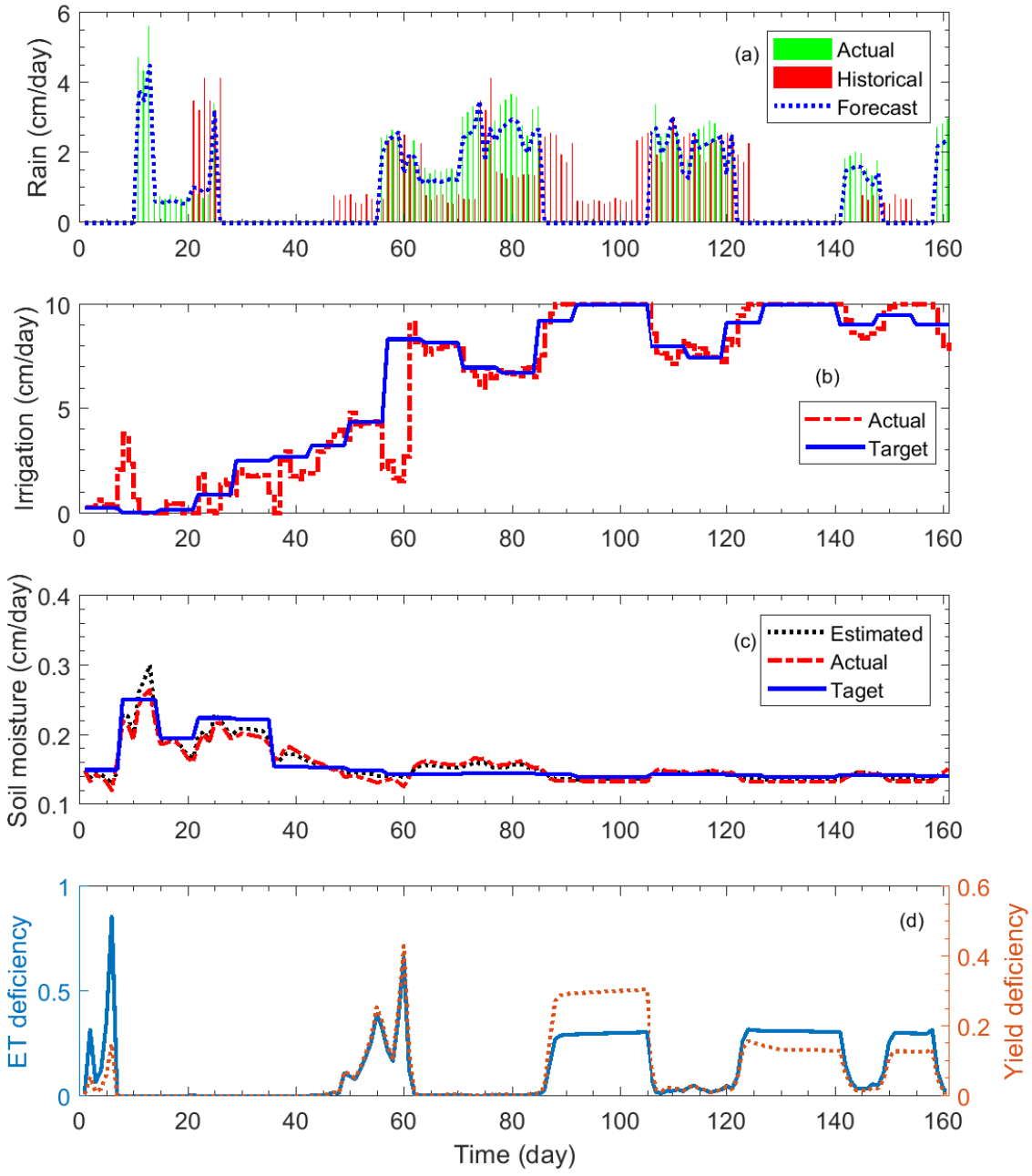


Figure 4.7: Simulation results of the closed-loop scheduling scheme under favorable weather scenario with more weighting on water conservation.

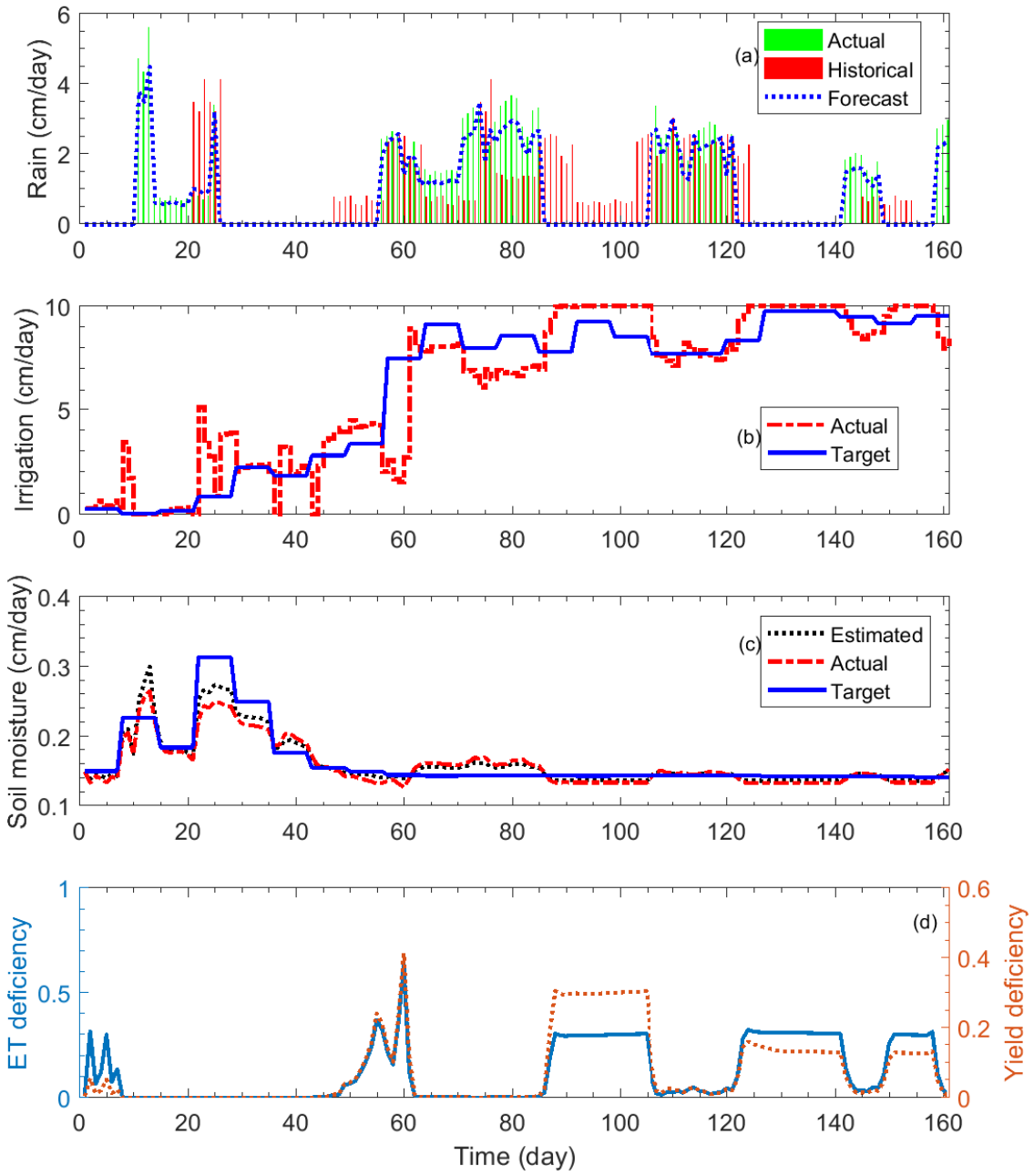


Figure 4.8: Simulation results of the open-loop scheduling scheme under favorable weather scenario with more weighting on water conservation.

chosen as $Q_{s1} = 1000^2$, $Q_{s2} = 100$, $Q_{s3} = 10^9$ and $Q_{s4} = 1$. The value of Q_{s3} is chosen very high to ensure that the constraint on the lower soil moisture limit is not violated. The weight on the upper soil moisture limit is relatively low since it is less important and can be controlled to some extent by the irrigation amount. Compared with the next case to be studied, $Q_{s2} = 100$ puts more weighting on water consumption. The values for θ_u and θ_w are taken as $0.41\, cm/cm$ and $0.15\, cm/cm$. The wilting value θ_w for the system is considered to be $0.09\, cm/cm$. Since there is model-plant mismatch, a lower limit that is higher than the wilting point is used. For the MPC controller, the weighting factors are $Q_c = 100$, $R_c = 0.001$, $Q_{c1} = 1000$ and $Q_{c2} = 0.01$. In the controller, a different lower limit $0.13\, cm/cm$ on the soil moisture is used since the controller gets daily feedback of soil moisture from the farm and it can handle uncertainty better. The simulation results are shown in Figure 4.7 for the closed-loop scheduling scheme and in Figure 4.8 for the open-loop scheduling scheme. For simplicity, only one day weather forecast is shown in the figures. In both schemes, the MPC controller overall is able to track the soil moisture targets well. The over prediction by the LPV model subplot (c) of Figures 4.7 and 4.8 is relatively significant in the crop rapid growing stage which is due to the relatively big mismatch of the LPV model from the actual system in the rapid growing stage. The peak at around day 15 is due to the sudden increase in the actual rain that is not captured in the weather forecast. Table 4.3 summarizes the results of this case. From the table, it can be seen that the open-loop scheme has a slightly (1.5%) better crop yield deficiency while the closed-loop scheme has a slightly reduced (1.9%) irrigation amount.

Second, we consider a case that we put less emphasis on water conservation (i.e., more on crop yield reduction). In the scheduler, the weights are $Q_{s1} = 1000^4$, $Q_{s2} = 1$, $Q_{s3} = 10^9$ and $Q_{s4} = 1$. The lower soil moisture constraint in the scheduler is decreased to $\theta_w = 0.13\, cm/cm$ whereas for the controller the lower limit is increased and set to be the set-point from the scheduler. The weights in the controller cost function are kept same as in the previous case. Thus the controller in this case tries to track the soil moisture set point from some higher values thereby ensuring less yield deficiency. A summary of the results is given in Table 4.3. In this case, we can see that the closed-loop and the open-loop schemes give very similar crop yield while the closed-loop scheme uses slightly (0.8%) less water.

Table 4.3: Simulation results under favorable weather scenario.

	More water conservation		Less water conservation	
	Closed-loop	Open-loop	Closed-loop	Open-loop
Yield deficiency	12.47	12.29	5.71	5.71
Irrigation implemented by controller (cm^3/cm^2)	990.10	1009.30	1222.40	1211.90
Irrigation prescribed by scheduler(cm^3/cm^2)	1020.50	993.44	1266.80	1242.00

Scenario 2: Dry weather

In this scenario, we consider that the amount of rain in the current crop season is significantly less than the average historical weather data. In this set of simulations, we consider that there is 79.4401 cm of rain in the crop growing season.

As in the previous scenario, we first consider a water saving mode in the scheduler. The scheduler and controller settings are the same as in the previous scenario. A summary of the results is given in Table 4.4. It can be observed that the closed-loop scheme saves more (2.3%) water relative to the open-loop scheme at the cost of (8.2%) less yield.

We also consider a similar case like in the previous scenario to put less weight on water conservation. The results are summarized in Table 4.4. From the table, a similar conclusion can be drawn. That is, the closed-loop scheme is able to use (2.5%) less water but at a cost of (5.9%) less yield.

To understand the effects of weights in the scheduling optimization, let us look at the results of the same scheduling scheme (closed-loop or open-loop) in different operating modes (more water conservation and less water conservation). From Table 4.3 (columns 2 & 4), it can be observed that when the objective in the closed-loop scheduler is changed to consider less water conservation, an increase of 54.24% in yield can be achieved with the cost of a 23.46% increase of water usage under the favorable weather condition. From Table 4.3 (columns 3 & 5), it can be seen that an increase of 53.52% in yield can be achieved with the cost of a 20.07% increase of water usage for the open-loop scheduler under the favorable weather condition. This may imply that the yield is more sensitive to the weights and a relatively smaller change in the consumption of irrigation water may result a much larger change in the yield.

A similar pattern can be observed from the results of Table 4.4 for the dry weather condition.

Table 4.4: Simulation results under dry weather scenario.

	More water conservation		Less water conservation	
	Closed-loop	Open-loop	Closed-loop	Open-loop
Yield deficiency	17.94	16.57	9.73	9.12
Irrigation implemented by controller (cm^3/cm^2)	1055.30	1079.60	1234.50	1266.80
Irrigation prescribed by scheduler(cm^3/cm^2)	1092.50	993.56	1273.70	1245.10

In addition to the above cases, we consider another case in which the water price varies. The water priced is increased in the relatively dry season between week 10 and 15 corresponding to day 70 and 105. The corresponding Q_{s2} value changes according to the price as follows:

$$Q_{s2}(\bar{k}) = \begin{cases} 1, & if\ \bar{k} < 10\ or\ \bar{k} > 15 \\ 100, & if\ 10 \leq \bar{k} \leq 15 \end{cases}$$

The results are shown in Figure 4.9 for the closed-loop integrated scheme and in Figure 4.10 for the open-loop irrigation scheme. A summary of the result is given in Table (4.5). The result shows that less water is used relatively to open-loop scheme. In the figures, it can be observed that a tendency to use more water when water price is low presents. The soil moisture set point before day 60 is around $0.3\,cm/cm$. This creates a water storage which later is used when the water price increases. It can be observed that the actual irrigation during day 59 drops to zero since the soil moisture set point gets reduced. In this case, the closed-loop scheme leads to a 5.3% water saving but at the cost of 5.2% crop yield reduction.

4.5 Conclusions

In this work, we investigated closed-loop scheduling together with closed-loop control for agriculture irrigation. In the proposed design, LPV models were identified and used in the design of the scheduler and the controller. The designed hierachical scheduling and control system is applied to a field with maize. In the simulations, a couple of different weather scenarios were considered. Within each scenario, a few

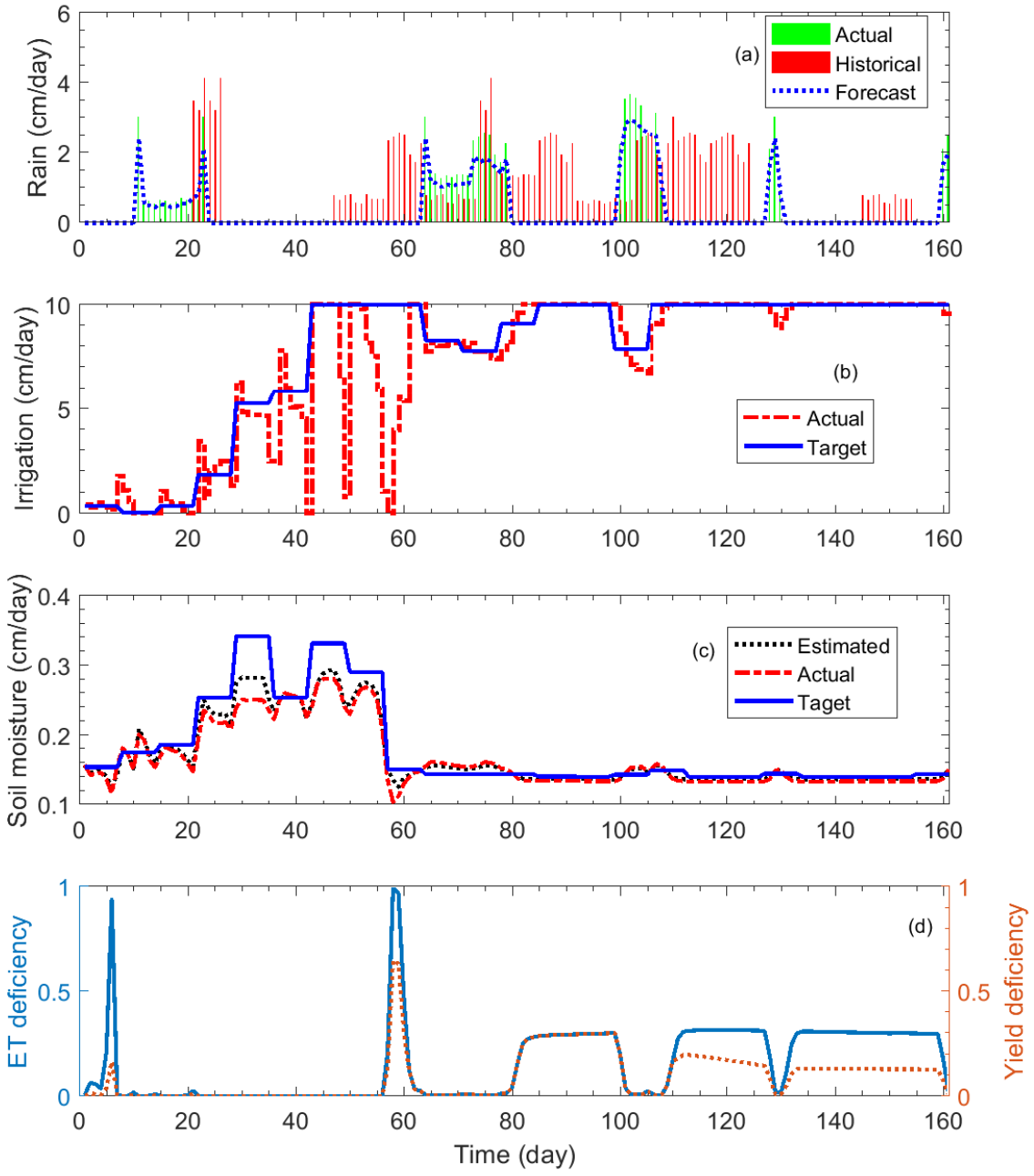


Figure 4.9: Simulation results of the closed-loop scheduling scheme under dry weather scenario with varying water price.

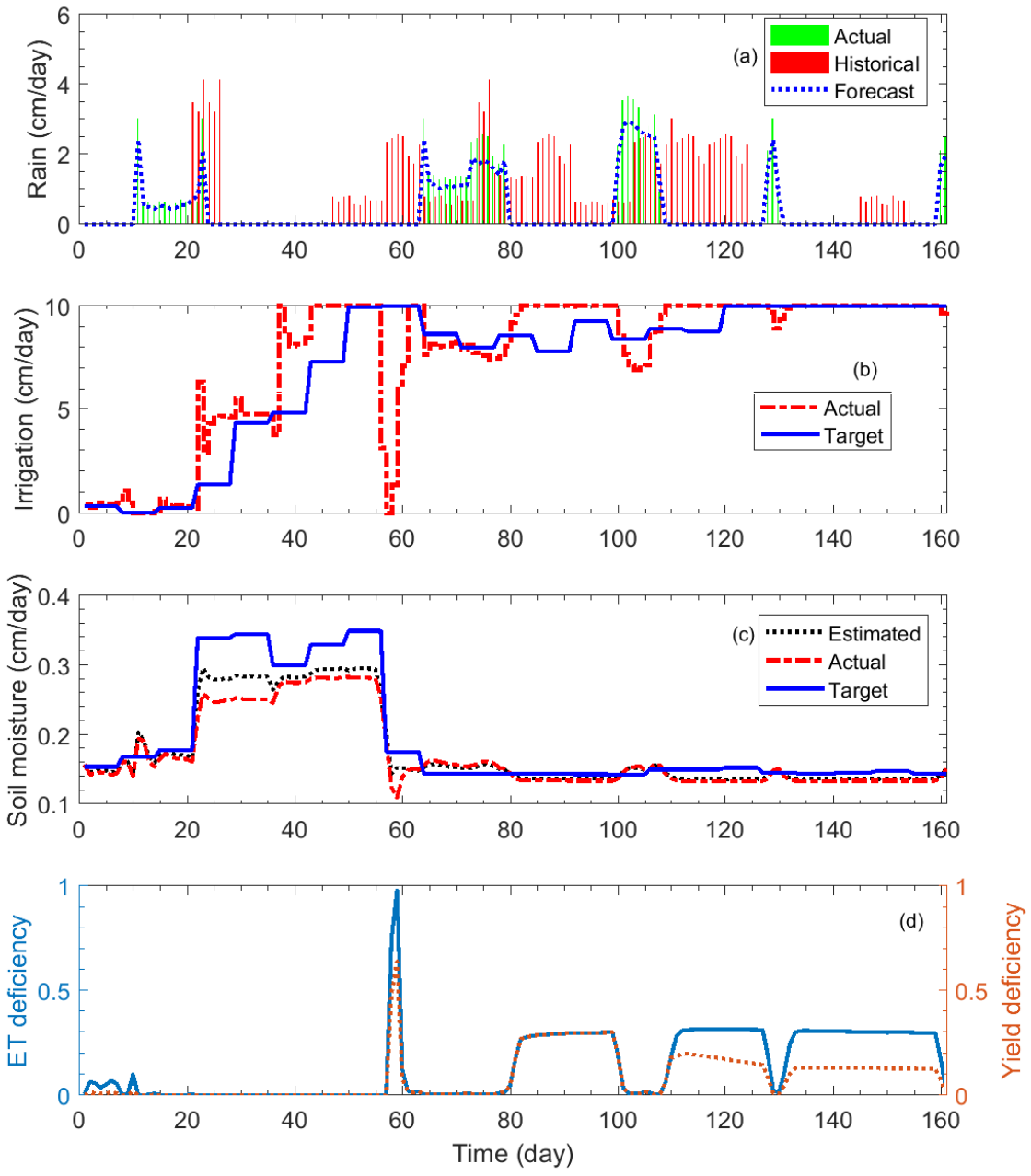


Figure 4.10: Simulation results of the open-loop scheduling scheme under dry weather scenario with varying water price.

Table 4.5: Simulation results under dry weather scenario with varying water price.

	Closed-loop	Open-loop
Yield deficiency	14.9129	14.1672
Irrigation implemented by controller (cm^3/cm^2)	1163.6	1228.9
Irrigation prescribed by scheduler (cm^3/cm^2)	1236.8	1167.7

different cases were studied including different weights on water conservation and varying water prices. From the simulation results, we can observe that the gain (either in water conservation or crop yield) in closed-loop scheduling in most of the cases comes with a cost (either lower crop yield or increased water consumption). The results in the simulations considered do not show a very clear benefit of using closed-loop scheduling. One possible reason for the results is that the crop maize considered in this work requires significant water in the growing season. It may have a much stronger impact on the irrigation pattern than the impacts of uncertainty in weather and model. This may have caused the irrigation to be high in most of the growing season as can be seen in the simulation results and thus makes the difference between closed-loop scheduling and open-loop scheduling not obvious. Moreover, in this work, the spatial heterogeneity is not considered and the field only includes maize. When spatial heterogeneity and more types of crops are considered, the benefits of a closed-loop scheduling may be more obvious. Further, in this work, we have only considered the soil moisture as a controlled variable, when other manipulated inputs (e.g., salinity) are considered, closed-loop scheduling may bring more benefits.

Chapter 5

Improved storm water management through irrigation rescheduling for city parks

5.1 Introduction

City storm water can be a good source for irrigating city green space such as parks and golf courses. In this chapter, a modeling and scheduling approach for an integrated storm water management and irrigation problem is presented. Specifically, we consider a park area with a storm water pond, which represents a typical setting in modern communities. The primary objective is to simultaneously ensure that the green space is irrigated appropriately and the level of the storm water pond is maintained adequately such that pond overflow or dry-out events are minimized. The overall system is discussed in Section 5.2. We propose to use closed-loop irrigation scheduling to achieve the objective. In the proposed approach, a steady-state model is also developed to calculate the soil water storage for different irrigation amounts which is developed in Section 5.3. Section 5.4 describes the formulation of an optimization problem to calculate the optimal irrigation amount. To handle the uncertainties (in weather forcast and modeling), real-time feedback from the pond is used to re-evaluate the scheduling optimization problem every week. Simulation results are presented in Section 5.5 which show that the proposed closed-loop scheduling gives much improved control performance in that it has less overflow or dry-out events from the storm water pond as well as maintains good green space soil moisture conditions. The results also imply that the proposed approach may be used to improve city storm

water management by integrating it with city park irrigation.

5.2 System description and modeling

5.2.1 System description

City storm water typically is from rain or snow and is collected using city drainage system and stored in wet ponds. Figure 1 shows the integrated storm water pond and

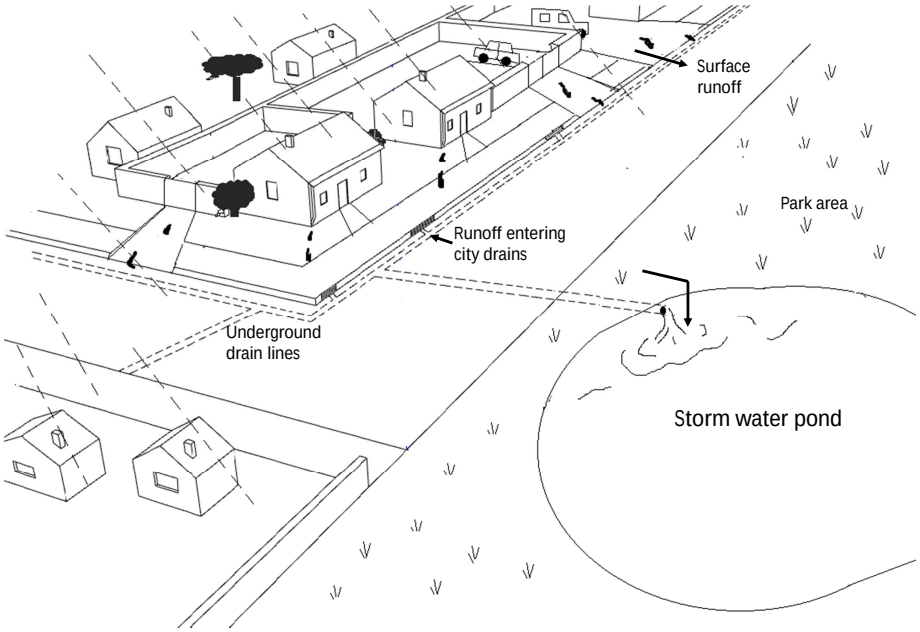


Figure 5.1: An illustration of the system considered.

city green space system considered in this work. In this system, storm water from the neighboring community is collected by the stormwater pond, which acts as a settling tank for the contaminants. Instead of discharging the stormwater directly to downstream water bodies, we consider using the less contaminated water at the top of the pond for irrigating the nearby green space. This helps reduce tap water consumption in irrigation and adds flexibility to stormwater management. The primary objective is to simultaneously ensure that the green space is irrigated appropriately and the level of the storm water pond is maintained adequately such that pond overflow or dry-out events are minimized.

5.2.2 Pond level model

The dynamics of the pond level are modeled based on mass balance where the inputs to the pond include the storm water from city drainage system $S\,(cm^3/hr)$ and pond precipitation volume $R_p\,(cm^3/hr)$, and the outputs include irrigation $I_p\,(cm^3/cm^2hr)$ and evaporation $E_p(cm^3/cm^2hr)$ from the pond. The following assumptions are made while developing the pond level model:

- The pond cross-sectional area is the same at all depths.
- The leakage from the pond is negligible.

The equation that describes the dynamics of the pond level is as follows:

$$A_p\frac{dP_L}{dt} = S(t) + R_p(t) - A_f I_p(t) - A_p E_P(t) \tag{5.1}$$

where, $P_L\,(cm)$ denotes the pond level, $A_p\,(cm^2)$ is the cross-sectional area of the pond and $A_f\,(cm^2)$ is the area of the green space that needs to be irrigated. The evaporation from the pond $E_p\,(cm/hr)$ is calculated following the methods described in Kohli et al. [99], where reference evapo-transpiration (ET) is used to model evaporation rate of open water similar to the pond considered in this work.

The storm water flow rate from city drainage system $S(t)$ is modeled as a function of the precipitation rate $R(t)\,(cm/hr)$ and it is assumed that at any time 50% precipitation on the neighborhood of the drainage, which has an area of $A_c\,(cm^2)$, enters the storm water collecting drainage system. Therefore $S(t) = 0.5A_cR(t)$. Also, the precipitation on the pond $R_p(t) = A_pR(t)$.

Remark 2 *In this work, the pond is modeled as a tank with the same cross-sectional area from the top to the bottom for brevity. In the proposed closed-loop scheduling framework, different types of pond models can be used in a straightforward manner as long as the model can represent the dynamics of the pond level adequately. When a different pond model is used, we should update (5.25c) in the scheduling optimization problem (5.25) and the scheduling implementation algorithm remains the same.*

Remark 3 *Note that in this work, we do not consider water run-off or run-on events due to slopes explicitly. However, we assume that water run-on or run-off affects the*

level of the storm water pond through city drainage system (i.e., $S(t)$ in equation (6)). It is assumed that at any time 50% precipitation on the neighborhood enters the pond through the storm water collecting drainage system.

5.2.3 Modeling of weather conditions

To model the weather conditions, two different weather datasets were generated. One weather dataset represents the historial weather data and the other weather dataset contains the actual weather data. To simulate realistic weather forecast an algorithm is developed which generates weather forecast either between or near to historical and actual weather data. The historical weather data contains weather data for 32 weeks for six weather parameters including maximum and minimum temperatures, humidity, wind speed, solar radiation and the amount of precipitation on hourly basis. Since only precipitation has a direct impact on irrigation decision making, therefore among the weather parameters only precipitation values are reported in the study. In the actual weather dataset, the precipitation amount may vary considerably from the historical weather data.

In the proposed scheduling approach, the historical weather data and two weeks' weather forecast are used. The following algorithm is used to generate the weather forecast, where the actual and the historical precipitation values are denoted by R_{act} and R_{his}, respectively, and their percentage deviation is denoted by ΔR.

a. Check whether the difference between the actual and the historical weather data ΔR is smaller than 20%. If so, the actual weather data is used as the forecast data.

b. If ΔR is greater than 80%, check if $R_{act} > R_{his}$. If true, then the forecasted precipitation $R_{for} = R_{act}(1 - N_r)$, otherwise $R_{for} = R_{act}(1 + N_r)$, where $N_r \in [0, 1]$. For both conditions, we have considered $N_r = 0.2$.

d. If ΔR is between 20% and 80%, check if $R_{act} > R_{his}$. If true, then forecasted precipitation $R_{for} = R_{act}(1 - N_r)$, otherwise $R_{for} = R_{his}(1 - N_r)$. Here, we have considered $N_r = 0.2$ when $R_{act} > R_{his}$ is true, otherwise N_r is taken as 0.3.

For example, if the historical and actual precipitation values of a week is given by $100\,mm$ and $300\,mm$, respectively, then the deviation ΔR is 66%. Further, R_{act} is smaller than R_{his}, so the forecast is given by $R_{his}(1 - N_r) = 0.7R_{his}$. Note that in the above algorithm, N_r is a tuning parameter. When N_r is small (close to 0), the generated weather forecast will be close to the actual weather; when N_r is big (close to 1), the generated forecast will be close to the historical data. In this work, we use $N_r = 0.2$ to make the forecast closer to the actual weather condition and at the same time to introduce some uncertainty in the forecast.

5.3 Semi-empirical steady-state model for scheduling

The models described in Chapter 2 will be used to simulate the green space and the pond, but this dynamic soil moisture model is not suitable for scheduling with large time horizon due to its high computational complexity. In this section, a semi-empirical steady-state model will be developed for scheduling. Here, we have used the same water stress co-efficient K_s as explained in Figure 4.3 following the work of Feddes et al. [3]. The desired stress factor K_s is one, which refers to maximum ET or maximum root water extraction. In addition to the soil moisture terms θ_w, θ_L θ_a and θ_u, another soil moisture value termed as field capacity θ_F is considered where $\theta_L < \theta_F < \theta_a$. In soil moisture deficit model [82], instead of θ_a, θ_F is considered as the upper limit since it is assumed that water above field capacity drains sufficiently fast and over-irrigation never happens. However, in the problem considered in this work, water conservation is not the primary objective and the green space may need to be controlled close to the saturation but below aeration point, to avoid pond overflows. Therefore, a model that is suitable for scheduling for the problem considered in this work needs to be developed. The modeling objective is to determine a steady-state input-output relation between the influx q_t and soil moisture θ. Typically, $q_t = I_p + I_s + R_p - ET$, where I_s is irrigation from tap water.

The steady-state model is developed based on the analytical solution to Richards' equation in Tracy et al. [100]. To derive the steady-state model, instead of Van Genuchten's model (equation (2.2)), the simplified Gardner's model as shown in

equations (5.2)-(5.3) is used to estimate the relative hydraulic conductivity and soil moisture content.

$$K(h) = K_{sat}K_r = K_{sat}e^{\alpha h} \tag{5.2}$$

$$\theta(h) = \theta_{res} + (\theta_{sat} - \theta_{res})S_e \tag{5.3}$$

where $S_e = K_r = e^{\alpha h}$ and α is some empirical parameter depending on soil properties.

5.3.1 Steady-state model development

Using Gardner's model equation (5.2) for estimating the relative hydraulic conductivity, Richards' equation (equation (2.1)) can be rewritten as:

$$\frac{1}{K_{sat}}\frac{d\theta}{dt} = \frac{d}{dz}(K_r\frac{dh}{dz}) + \frac{dK_r}{dz} \tag{5.4}$$

Now, let us define,

$$\bar{h} = e^{\alpha h} - \epsilon \tag{5.5}$$

where ϵ is a constant. To drive the system into steady-state the input needs to be kept fixed, so the top flux q_t (cm/hr) is fixed. Therefore, we consider $\epsilon = q_t$. Using equation (5.5) the gradient of $\bar{h}$ can be derived as follows:

$$\frac{d\bar{h}}{dz} = \alpha e^{\alpha h}\frac{dh}{dz} \tag{5.6}$$

Using the Gardner's model relation $K_r = e^{\alpha h}$ and equation (5.6), the gradient terms in R.H.S of equation (5.4) can be expressed in terms of the transformed $\frac{d\bar{h}}{dz}$ as follows:

$$K_r\frac{dh}{dz} = e^{\alpha h}(\frac{1}{\alpha}e^{-\alpha h}\frac{d\bar{h}}{dz}) = \frac{1}{\alpha}\frac{d\bar{h}}{dz} \tag{5.7a}$$

$$\frac{dK_r}{dz} = \alpha e^{\alpha h}\frac{dh}{dz} = \frac{d\bar{h}}{dz} \tag{5.7b}$$

Again, using equation (5.3) and $S_e = K_r$, the change of soil moisture is given by:

$$\frac{d\theta}{dt} = (\theta_{sat} - \theta_{res})\frac{dS_e}{dt} = (\theta_{sat} - \theta_{res})\frac{d\bar{h}}{dt} \tag{5.8}$$

Bashed on equations (5.7a), (5.7b) and (5.8), equation (5.4) can be rewritten as follows:

$$c_a\frac{d\bar{h}}{dt} = \frac{d^2\bar{h}}{dz^2} + \alpha\frac{d\bar{h}}{dz} \tag{5.9}$$

where $c_a = \dfrac{\alpha(\theta_{sat} - \theta_{res})}{K_{sat}}$. For steady-state $\dfrac{d\bar{h}}{dt} = 0$. Therefore, equation (5.9) can be written as:

$$\frac{d^2\bar{h}_{ss}}{dz^2} + \alpha\frac{d\bar{h}_{ss}}{dz} = 0 \tag{5.10}$$

and the general solution of which is given by:

$$\bar{h}_{ss} = A_1 + A_2 e^{-\alpha z}. \tag{5.11}$$

To solve this equation, the two boundary conditions need to be defined to get the two parameters A_1 and A_2. In the Green-Ampt problem [100], pressure head boundary conditions were considered for which A_1 and A_2 can be easily determined. For our problem, we need the relation between steady-state input flux and soil moisture, therefore flux boundary conditions are required. In order to consider flux boundary, equation (5.11) needs to be differentiated. So, if flux boundary is chosen both at the top and the bottom, the value of A_1 cannot be determined. Now to overcome this problem, a relation from the assumption of steady-state is used and an empirical term is added.

At steady-state, input is equal to output, which implies $q_t = -q_b$. Now for free drainage condition at the bottom, the gradient $\dfrac{dh}{dz}|_{z=L}$ is equal to zero. According to VanDam et al. [79] the flux is expressed in terms of pressure gradient as follows:

$$q_b = -K(h)(\frac{dh}{dz}|_{z=L} + 1) \tag{5.12}$$

Now using the relation $K(h) = K_r e^{\alpha h}$, we get the following relation between $h_{ss}|_{z=L}$ and q_t:

$$q_b = -K_{sat} e^{\alpha h}(\frac{dh}{dz} + 1) \tag{5.13a}$$

$$-q_t = -K_{sat} e^{\alpha h_{ss}|_{z=L}} \tag{5.13b}$$

$$h_{ss}|_{z=L} = \frac{1}{\alpha} ln(\frac{q_t}{K_{sat}}) \tag{5.13c}$$

From the above equation and equation (5.5) we get:

$$\bar{h}_{ss}|_{z=L} = e^{\alpha h_{ss}|_{z=L}} - q_t \tag{5.14a}$$

$$\bar{h}_{ss}|_{z=L} = \frac{q_t}{K_{sat}} - q_t. \tag{5.14b}$$

Again, using equation (5.12) and $K(h) = K_r e^{\alpha h}$ for the top boundary (z=0), we get:

$$q_t = -K_{sat} K_r (\frac{dh_{ss}}{dz}|_{z=0} + 1) \tag{5.15a}$$

$$q_t = -K_{sat} e^{\alpha h_{ss}|_{z=0}} (\frac{dh_{ss}}{dz}|_{z=0} + 1) \tag{5.15b}$$

$$\frac{q_t}{-K_{sat} e^{\alpha h_{ss}|_{z=0}}} - 1 = \frac{dh_{ss}}{dz}|_{z=0} \tag{5.15c}$$

Using equation (5.6), we can express the above equation in terms of $\bar{h}_{ss}$ as follows:

$$\frac{q_t}{-K_{sat} e^{\alpha h_{ss}|_{z=0}}} - 1 = \frac{1}{\alpha e^{\alpha h_{ss}|_{z=0}}} \frac{d\bar{h_{ss}}}{dz}|_{z=0} \tag{5.16a}$$

$$\frac{d\bar{h_{ss}}}{dz}|_{z=0} = -\alpha \frac{q_t}{K_{sat}} - c' \tag{5.16b}$$

where $c' = \alpha e^{\alpha h_{ss}|_{z=0}}$. Differentiating equation (5.11) at $z = 0$ and using equation (5.16b), we get:

$$\frac{d\bar{h_{ss}}}{dz}|_{z=0} = -\alpha A_2 e^{-\alpha z}|_{z=0} \tag{5.17a}$$

$$-\alpha A_2 = -\alpha \frac{q_t}{K_{sat}} - c' \tag{5.17b}$$

$$A_2 = \frac{q_t}{K_{sat}} + c \tag{5.17c}$$

where $c = \dfrac{c'}{\alpha}$. Again, using equation (5.11) for $z = L$ and equation (5.14b), we get:

$$\bar{h_{ss}}|_{z=L} = A_1 + A_2 e^{-\alpha L} \tag{5.18a}$$

$$\frac{q_t}{K_{sat}} - q_t = A_1 + A_2 e^{-\alpha L} \tag{5.18b}$$

$$A_1 = \frac{q_t}{K_{sat}} - q_t - (\frac{q_t}{K_{sat}} + c) e^{-\alpha L} \tag{5.18c}$$

Putting the values of A_1 and A_2 from equation (5.18c) and (5.17c) into equation (5.11), we get:

$$\bar{h}_{ss} = -q_t + \frac{q_t}{K_{sat}} - (\frac{q_t}{K_{sat}} + c) e^{-\alpha L} + (\frac{q_t}{K_{sat}} + c) e^{-\alpha z} \tag{5.19}$$

Finally, using the relation from equation (5.5), we get the following expressions:

$$h_{ss} = \frac{1}{\alpha} ln(\bar{h}_{ss} + q_t) \tag{5.20a}$$

$$h_{ss} = \frac{1}{\alpha} ln(\frac{q_t}{K_{sat}} - (\frac{q_t}{K_{sat}} + c) e^{-\alpha L} + (\frac{q_t}{K_{sat}} + c) e^{-\alpha z}) \tag{5.20b}$$

Equation (5.20b) shows the steady-state pressure h_{ss} for fixed input q_t as a function of depth z. The value of c is unknown in this equation since only the flux at $z = 0$

is known, not the value of h_{ss}. But to our knowledge, there is no analytical means to identify the value of c. It is assumed that at steady-state condition, $h_{ss}|_{z=0}$ will remain fixed and the terms in c will generate some constant value. Therefore an empirical method is used to identify the value of the c term. Since the output of the steady-state model is θ, so h_{ss} from equation (5.20b) is used to estimate θ_{ss} using Van Genuchten's model (equation (5.21)) with the additional term c_1 which reflects the role of the c term in equation (5.20b).

$$\theta_{ss}(h) = \theta_{res} + (\theta_{sat} - \theta_{res})(1 + |\alpha_v h_{ss}|^{\eta})^{-m} + c_1 \tag{5.21}$$

Since c_1 is unknown in the analytical solution, we propose to identify c_1 based on the data obtained from the system. The dynamic agro-hydrological model described in Chapter 2 is used to generate soil moisture data for 20 different values of the input flux q_t (cm/hr). In each case, the input flux q_t is kept constant for 200 hr. It is assumed that within 200 hours the soil moisture will reach steady-state and the soil moisture value at 200^{th} hour is recorded. Based on this simulated data, a steady-state relation between the input flux q_t and the steady-state soil moisture θ is generated.

It can be observed from Figure 5.2, that the relation between the flux q_t and soil moisture θ is similar to exponential growth curve. To capture this relation, c_1 is chosen as an exponential function of q_t. We considered $c_1 = -(0.03 + 0.15e^{(-5q_t)})$ in equation (5.21) where the numerical values were determined using curve fitting. The figure shows that equation (5.21) performs considerably well, so this approach is used to determine the output steady-state soil moisture values. Since Gardner's model simplifies relative saturation S_e to be equal to $e^{\alpha h}$ (equation (5.3)) instead of the complex term $(1+|\alpha_v h|^{\eta})^{-m}$ used in Van Genuchten's model, the use of Gardner's model to determine θ_{ss} from h_{ss} results poor performance and therefore is not used. Based on equations (5.20b) and (5.21) as well as the expression of q_t, the steady-state input-output model at depth z can be expressed as a compact function as follows:

$$\theta_{ss} = g(I_s, I_p, R, ET, z) \tag{5.22}$$

In the proposed scheduling algorithm, the water storage of the entire root zone will be considered. To address this, the entire soil profile under consideration is divided into N compartments as shown in Figure 5.3. In each compartment, the soil

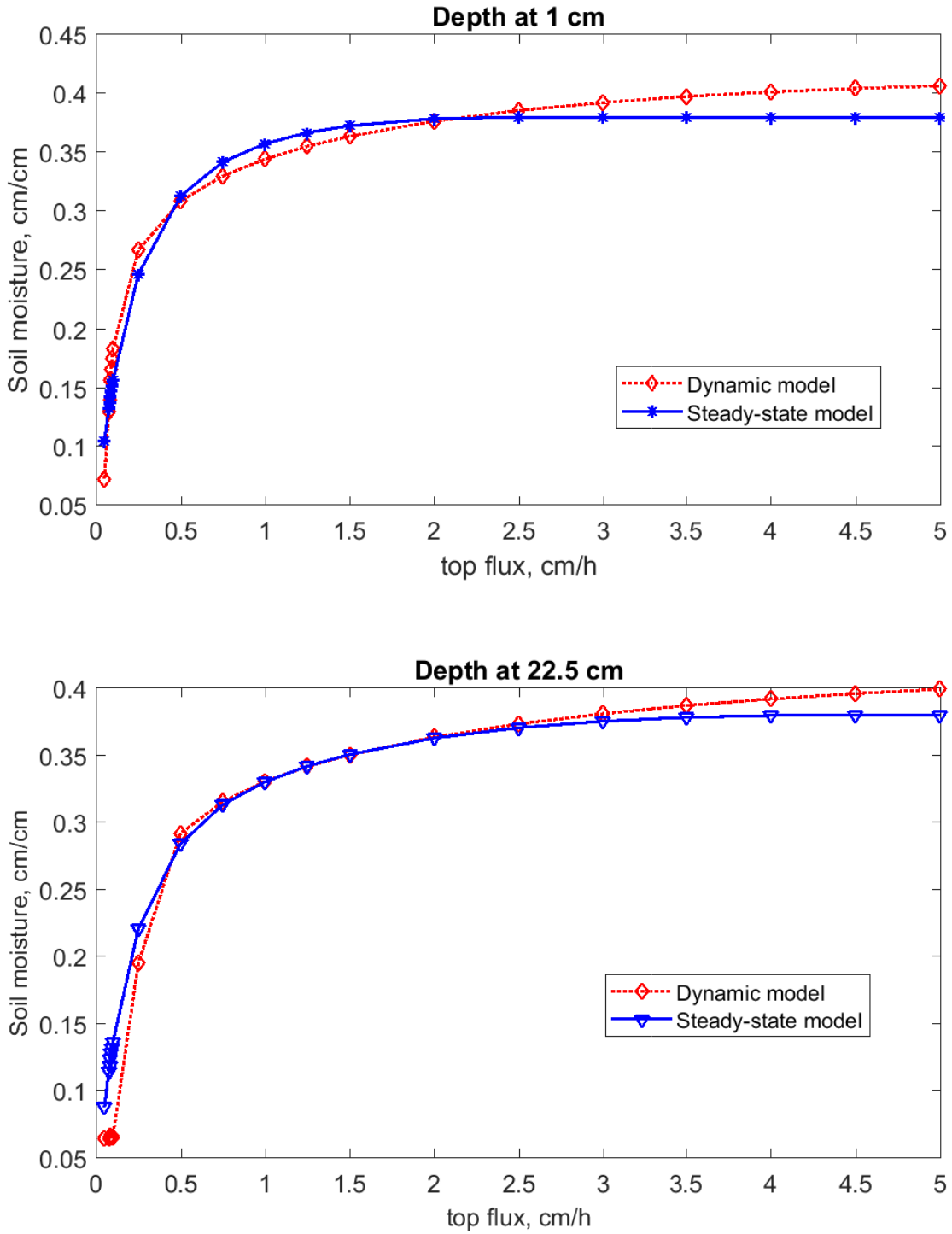


Figure 5.2: The steady-state model between input flux q_t and output soil moisture θ. Left figure is for depth 1 cm and right figure is for depth 22.5 cm. The intermediate depths are not shown but they also have similar results.

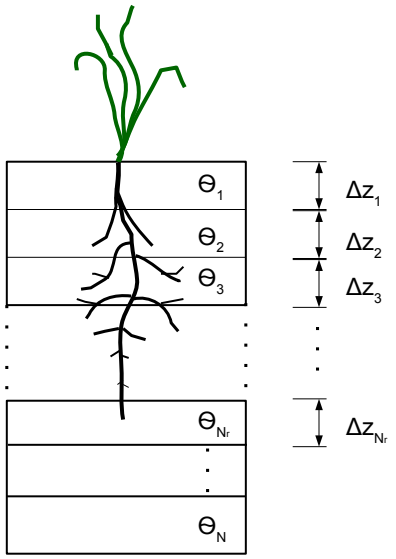


Figure 5.3: The soil water storage inside root zone.

moisture is assumed to be a constant over the week and will be estimated using the above steady-state model. The compartment where the root ends is denoted by N_r, where $N_r < N$.

The water storage within the entire root zone may be approximated as follows:

$$W_s = \sum_{i=1}^{r} \theta_i \Delta z_i \tag{5.23}$$

where $W_s\,(cm)$ is the soil water storage, Δz_i and θ_i are the height and soil moisture of compartment i, respectively. Based on equations (11) and (12), the relation between the irrigation inputs and the soil water storage can be expressed as follows:

$$W_s = \sum_{i=1}^{r} g_i(I_s, I_p, R, ET, z)\Delta z_i \tag{5.24a}$$

$$:= f(I_p, I_s, ET, R) \tag{5.24b}$$

5.4 Proposed scheduling algorithm

In the proposed approach, the scheduling problem is formulated as an optimization problem, which is re-evaluated every sampling time (a week) based on pond level feedback to accommodate uncertainty in modeling and weather conditions. The scheduler determines the (reference) amount of water used from the pond for irrigation and the

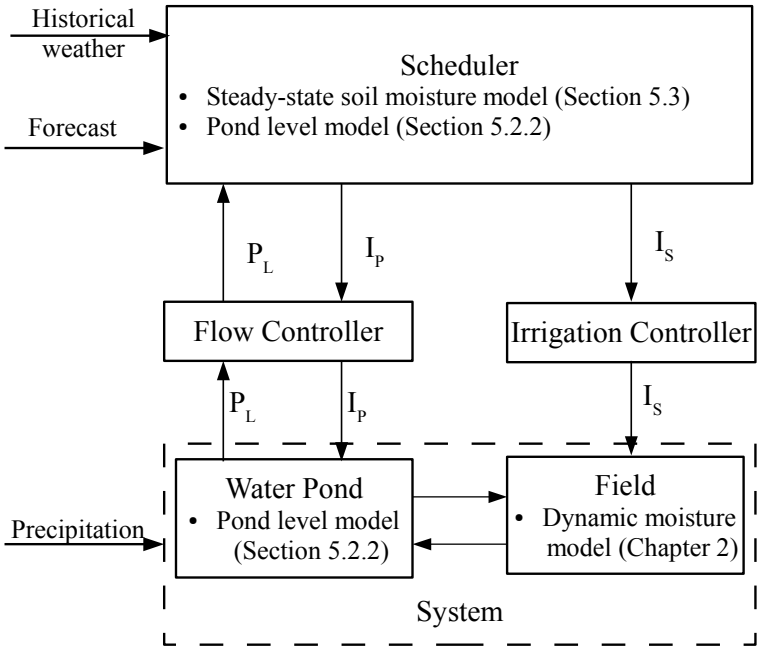


Figure 5.4: The hierarchy formed by the proposed scheduler and the lower-level set-point tracking controllers. The scheduler determines the (reference) amount of water used from the pond for irrigation and the (reference) amount of tap water used for irrigation. These reference amounts are sent to lower-level controllers for implementation.

(reference) amount of tap water used for irrigation based on the steady-state soil moisture model and the pond model as well as available weather information. These reference amounts are sent to lower-level controllers for implementation. The overall hierarchy of the decision making process is shown in Figure 5.4.

5.4.1 Formulation of the scheduling optimization

Scheduling for irrigation using storm water involves solving multiple prioritized objectives using one single objective function. For irrigation scheduling the most prioritized objective is to maintain soil water storage of the root zone within a desired range. Some secondary objectives are to minimize the overflow from the pond for wet seasons and to avoid pond drying out and to minimize the use of tap water for irrigation when there is no water available in the pond for irrigation. The formulation of the

scheduling optimization problem at the sampling time $\bar{k}$ is shown below:

$$\min_{I_p, I_s, \varepsilon_l, \varepsilon_u, \delta} \sum_{i=\bar{k}}^{\bar{k}+N} (C_1\varepsilon_l(i) + C_2\varepsilon_u(i) + C_3\delta(i) + C_4 I_s(i) + C_5 I_p(i)) \tag{5.25a}$$

$$\text{s.t.} \quad W_s(i+1) = f(I_p(i), I_s(i), ET(i), R(i)) \tag{5.25b}$$

$$P_L(i+1) = P_L(i) + \frac{0.5A_c + A_p}{A_p} R(i) - \frac{A_g}{A_p} I_p(i) - E_p(i) \tag{5.25c}$$

$$W_s + \varepsilon_l \geq L_s \tag{5.25d}$$

$$W_s - \varepsilon_u < U_s \tag{5.25e}$$

$$P_L - \delta < L_{max} \tag{5.25f}$$

$$\varepsilon_l \geq 0, \quad \varepsilon_u \geq 0, \quad \delta \geq 0, \quad I_s \geq 0, \quad I_p \geq 0, \quad P_L \geq 0 \tag{5.25g}$$

Equation (5.25a) describes the objective function that needs to be minimized over the horizon $i = 1...N$ to determine the optimal solution. The decision variables of this optimization are tap water usage I_s, pond water usage I_p, slack variables ε_l and ε_u for maintaining soil moisture within range, and slack variable δ for avoiding overflow. In the objective function the tap water usage is directly minimized with the cost C_4. The irrigation amount with pond water I_p is also penalized with a small cost C_5 to make the solution unique inside the feasible range. Equation (5.25b) is the steady-state model that predicts soil water storage W_s as described in Section 5.3. Equation (5.25c) is the discrete-time version of the pond level model as described in Section 5.2.2. Equation (5.25d)-(5.25g) are the constraints of this optimization problem. Equation (5.25d) is used to keep soil water storage W_s above some lower limit L_s. The variable ε_l is added to ensure a feasible solution is achieved. Since the violation of this constraint is not acceptable so the variable ε_l is associated with the highest cost C_1 in the objective function. Similarly, equation (5.25e) ensures that W_s stays below some upper limit U_s and ε_u is also associated with large cost C_2. Equation (5.25f) describes that the pond must not overflow and the slack variable δ is associated with the cost C_3. Equation (5.25g) implies all physical constraints. The irrigation amount, pond level, and the slack variables are all non-negative real numbers. In addition, the maximum value for the pond level is equal to the pond height L_{max}. In the design of (5.25), soft constraints are used to incorporate constraints on soil moisture and overflow to ensure that (5.25) is always feasible. This approach

ensures that a solution exists for the optimization problem.

5.4.2 Scheduling implementation algorithm

The proposed scheduling optimization problem does not explicitly take into account of the uncertainties. Two primary uncertainties include modeling uncertainty due to the use of a simplified steady-state model in the optimization problem and weather forecast uncertainty. Both of the uncertainties have a significant impact on the scheduling results. To deal with these uncertainties, the scheduling optimization problem is re-evaluated every week based on new pond level measurement and new weather forecast to form a closed-loop scheduling scheme as shown in Figure 5.4. At each sampling time,

- *Step 1:* A new measurement of the pond level $P_L(\bar{k})$ and the weather forecast for the following two weeks are obtained.
- *Step 2:* Based on the new measurement and new weather forecast, the optimization problem (14) is evaluated to generate the future trajectories of $I_p(i)$ and $I_s(i)$ for $i \in [i = \bar{k}, \bar{k} + N]$.
- *Step 3:* Only the first step values $I_p(\bar{k})$ and $I_s(\bar{k})$ are sent to the irrigation control system for implementation for the current week.
- *Step 4:* At the next sampling time $\bar{k} + 1$, go to Step 1 ($\bar{k} + 1 \rightarrow \bar{k}$).

5.4.3 Lower-level controllers

As shown in Fig. 5.4, the scheduler calculates the reference I_p and I_s and sends them to the lower-level controllers for tracking. In this work, it is assumed that ideal lower-level controllers are in place and can track the reference values immediately and perfectly for the simplicity of presentation. In practice, PID controllers may be used to track these references.

Remark 4 *Note that before implementing the proposed approach, a preliminary study of the field and the pond should be conducted to collect data for model parameter identification and uncertainty characterization purposes. Once adequate models of*

the pond and the field are identified, extensive simulations based on the models may be carried out for tuning of the weights, sampling time, prediction horizon in the proposed design.

5.5 Simulations

In this section, we carry out simulations to demonstrate the performance of the proposed closed-loop scheduling approach. In the simulations, the models described in Chapter 2 and Section 5.2.2 are used to simulate the dynamics of the pond level and the soil moisture of the green space. Specifically, the parameters used in the simulations are shown in Table 5.1.

Table 5.1: Parameters of the experimental setup.

$A_f = 3 \times 10^9\ cm^2$	Area of green space
$A_c = 4A_f$	Neighboring area
$A_p = 0.2A_f$	Storm water pond area
$L_{max} = 1500cm$	Maximum pond level
$L_{min} = 0$	Minimum pond level
$SPD = 110\ cm$	Soil profile depth
$RZD = 25\ cm$	Root zone depth
$\theta_{ini} = 0.15\ cm/cm$	Initial field condition
$\theta_{res} = 0.065\ cm/cm$, $\theta_{sat} = 0.41\ cm/cm$,	Van Genuchten model parameters
$\alpha_v = 0.075\ cm^{-1}$,$\eta = 1.589$, $K_{sat} = 4.4208\ cm/h$	for sandy loam soil

5.5.1 Closed-loop v.s. open-loop scheduling

First, we compare the performance of the proposed closed-loop scheduling and an open-loop scheduling method under two weather conditions over a time period of 16 weeks. In the open-loop scheduling method, the scheduling optimization is solved in the first week only (no re-evaluation) and the solution obtained in week 1 is used for the entire 16 weeks.

In the first case, we consider a wet weather condition. As mentioned in Section 5.2.3, the historical weather data has precipitation during week $1-6$ with an average of $240\,mm$, week $13-14$ with an average of $150\,mm$ and finally week $23-28$ with an average of $340\ mm$ in Figure 5.5(a). The actual weather data is shown for the first 16 weeks only. It is considered that the actual rain takes place between weeks

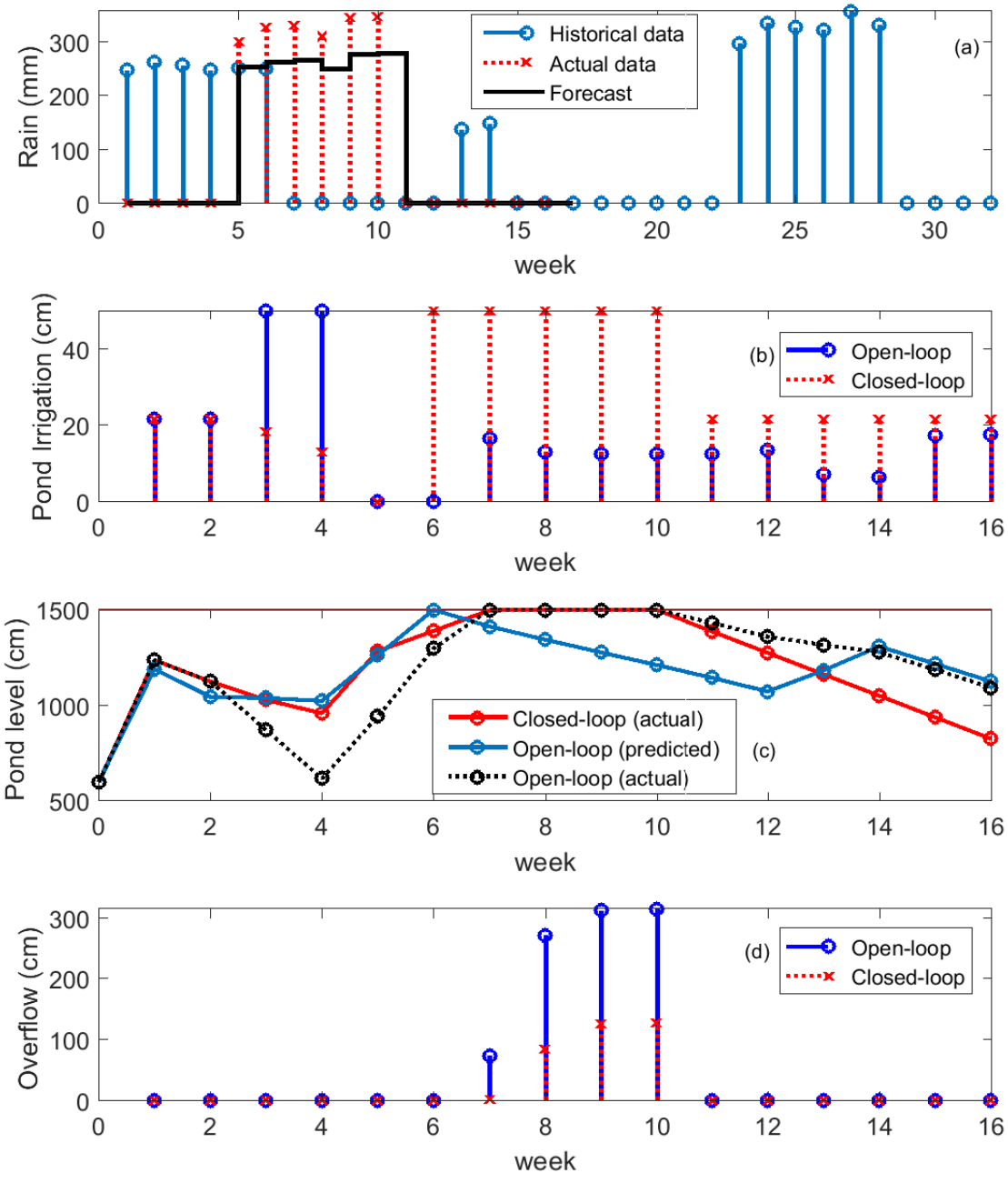


Figure 5.5: Simulated results on open-loop and closed-loop irrigation scheduling under wet weather conditions. (a) weather conditions, (b) weekly irrigation amount from the pond, (c) the pond level and (d) overflow incidents

$5-10$ with an average of $340\,mm$. The weather forecast is available real time for two coming weeks.

The weighting factors in the cost function of the scheduling optimization problem are selected as follows: $C_1 = 10000$, $C_2 = 10000$, $C_3 = 100$, $C_4 = 100$, $C_5 = 0.001$. The optimization horizon for both closed-loop and open-loop scheduling is taken as $N = 16$ weeks. In the optimization problem, the upper limit for soil water storage is considered to be $U_s = 5.00\,cm$ corresponding to $U_\theta = 0.25\,cm/cm$ of soil moisture and the lower limit for soil water storage is $L_s = 2.5\,cm$ corresponding to $L_\theta = 0.1\,cm/cm$ of soil moisture. Initial pond level is considered to be $L_{in} = 1350\,cm$. Soil water storage corresponding to saturation is $12.5\,cm$. Figure 5.5 shows the simulation results. Figure 5.5(b) shows the pond irrigation amount. In the open-loop scheduling, the historical weather data is used. Due to the mismatch of the historical weather data and the actual weather data, the prescribed pond irrigation is overall much lower in open-loop scheduling compared with closed-loop scheduling. This renders more (both time and amount) overflows under open-loop scheduling. In closed-loop rescheduling scheme, pond irrigation increases after week 6 following the heavy rain incident thereby reducing the chance of overflow as indicated by the pond level in Figure 5.5(c) and overflow incidents shown in Figure 6(d). In Figure 5.5(c), during open-loop scheduling the predicted overflow occurred at only week 6, but due to heavy rain the overflow took place for the entire period between week 7 and week 10. Though the period of overflow in the closed-loop rescheduling is similar to open-loop scheduling but the amount of overflow is much smaller in closed-loop rescheduling as shown in Figure 5.5(d). The soil moisture condition is shown in Figure 5.6 and the soil water storage values are shown in Table 5.2. It is clear from Table 5.2 that the soil water storage value at all time is less than the saturated soil water storage value of $12.5\,cm$ for both schemes. Except the rainy periods (week 5 to 9) the actual soil water storage value for the closed-loop scheduling scheme is within the upper U_s and lower limits L_s, but for open-loop scheduling the soil water storage value was much higher. Similar results were noticed in Figure 5.6. It shows that at the beginning (week 1), the soil moisture at the bottom of root zone was low for both closed-loop scheduling and open-loop scheduling scheme. During the actual rainy weeks (week 5 to 9) both schemes give higher soil moisture. But as soon as rain stops, the closed-

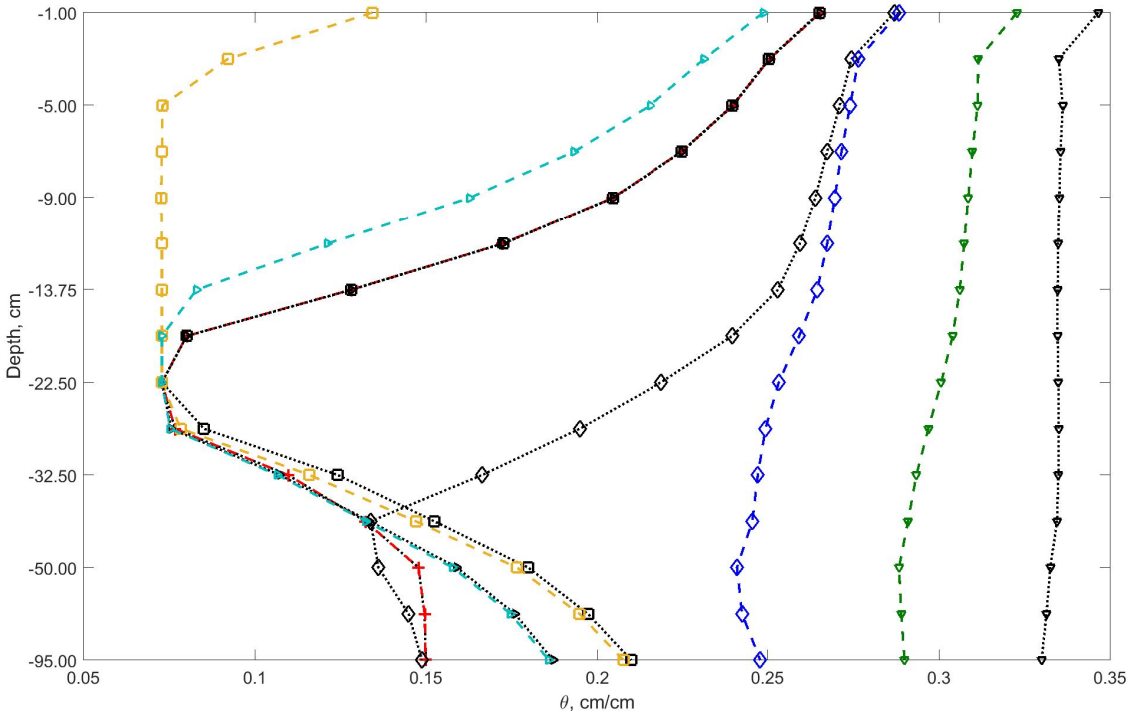


Figure 5.6: Soil moisture condition for wet weather conditions. Results for weeks 1, 5, 9, 13 and 16 are indicated by markers $+$, $\diamondsuit$, $\triangledown$, $\square$ and $\triangleright$, respectively. Dotted lines represent rescheduling scheme (closed-loop) and dashed lines stand for scheduling without feedback (open-loop).

loop scheduling shows better performance in terms of ground condition (week 13 and 16) since the soil moisture remains between field capacity and permanent wilting.

Table 5.2: Soil water storage in root zone for the case of wet weather conditions.

Week no.	Closed-loop scheduling	Open-loop scheduling
1	3.888	3.888
2	3.889	3.886
3	3.473	7.632
4	2.514	7.660
5	6.303	6.654
6	8.296	6.363
7	8.093	7.295
8	8.420	7.590
9	8.400	7.687
10	8.305	7.510
11	4.136	2.461
12	3.897	2.622
13	3.888	1.986
14	3.888	1.942
15	3.895	3.263
16	3.891	3.345

Table 5.3 shows the summary of the results. It is clear that when closed-loop scheduling scheme is used, the overflow of the pond can be reduced. It needs to be mentioned that though the soil moisture is above field capacity ($\theta_F = 0.25\, cm/cm$) at some periods but it is still below saturation ($\theta_{sat} = 0.41$) so there are no ponding events.

In the second case, we consider a dry weather condition. The dry weather data was generated using low precipitation amounts as shown in Figure 5.7(a). The historical weather data contain precipitations during weeks 1-6 with an average of 15 mm of rain, weeks 13-14 with an average of $14mm$ and weeks 23-28 with an average of $11mm$. The actual weather has rain between week 4 and week 12 with an average of 17 mm.

Table 5.3: Summary of results for the case of wet weather conditions.

	Open-loop scheduling	Closed-loop scheduling
Total pond water used (cm)	271.8299	453.3219
No of times pond overflows	3	3
Total amount of overflows (cm)	974.7851	341.7129

It is considered that the maximum tap water that can be used is $10\,cm/week$. Note that under significantly different weather conditions, some of the parameters in the scheduling optimization problem should be re-tuned to obtained good performance. In particular, the upper and lower limits for soil water storage need to be adjusted. In the previous wet weather condition, overflow was the main event that needed to be avoided and the the crop was under stress between θ_a (some aeration point near saturation) and θ_F (field capacity). In the dry weather condition, the pond is more likely to dry out than overflow. Further, as shown in Figure 4.3, the soil moisture may violate the lower limit of θ_W which may create long term damage on the crops. The cost of the violation of this lower limit is much higher. To accommodate the modeling mismatch, the lower limit of soil water storage in the optimization is considered to be $L_s = 5.0\,cm$ and the upper limit is considered as $U_s = 7.5\,cm$, which are significantly higher than the previous case. At the same time, the cost associated with the lower limit is increased. Thus the weighting factors in the cost function of the scheduling optimization problem are selected as follows: $C_1 = 1.0 \times 10^9$, $C_2 = 10000$, $C_3 = 100$, $C_4 = 100$, $C_5 = 0.001$. The initial pond level is considered as $L_{in} = 600\,cm$ and the prediction horizon is again $N = 16\,weeks$. The actual soil water storage values as simulated by the dynamic model are shown in Table 5.4.

Figure 5.7 shows the simulation results. Figures 5.7(b), (c) and (d) show pond irrigation amount, pond level and potable water irrigation amount, respectively. From Figure 8(b), it can be seen that in the closed-loop scheduling scheme, the pond irrigation is lower and the pond level is well maintained above zero for the entire period. While in the open-loop scheduling case, more pond water is used in irrigation and the pond level is very close to zero at the end of the period. In both schemes, tap water is used at its maximum allowable capacity.

The soil moisture condition is shown in Figure 5.8. It shows that both for closed-loop and open-loop scheduling the soil moisture at the bottom of the root zone reaches wilting point but the top root zone where most of the root water extraction takes place stays above the lower limit. In fact, it can be observed from Table 5.4 that the total soil water storage value stays above the lower limit $L_s = 2.5\,cm$ at all times. Table 5.5 shows the summary of the results for the case of dry events. It is clear that when closed-loop scheduling is used, the pond level can be maintained by reducing

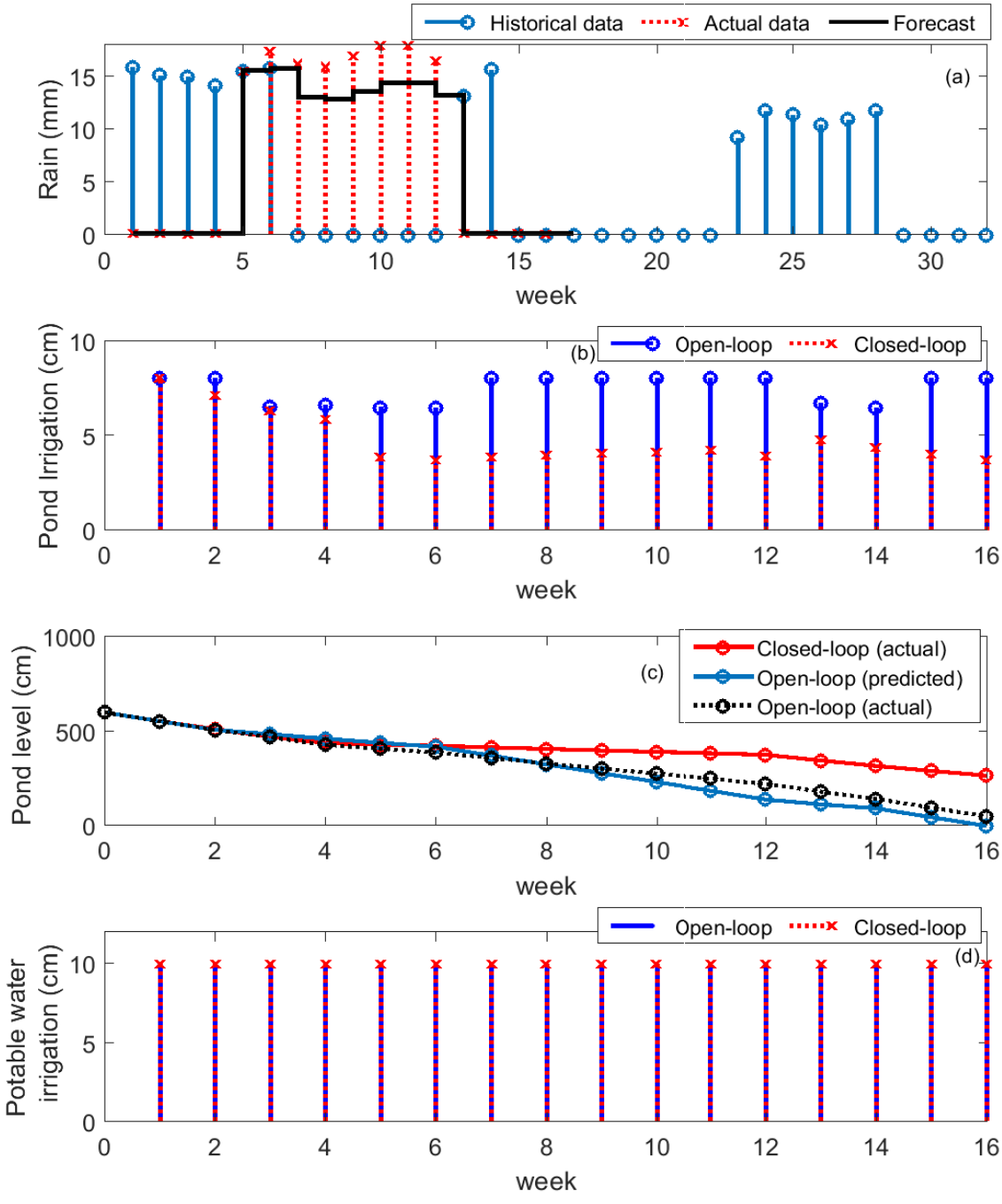


Figure 5.7: Simulated results on open-loop and closed-loop irrigation scheduling under dry weather conditions. (a) weather conditions, (b) weekly irrigation amount from the pond, (c) the pond level and (d) overflow incidents

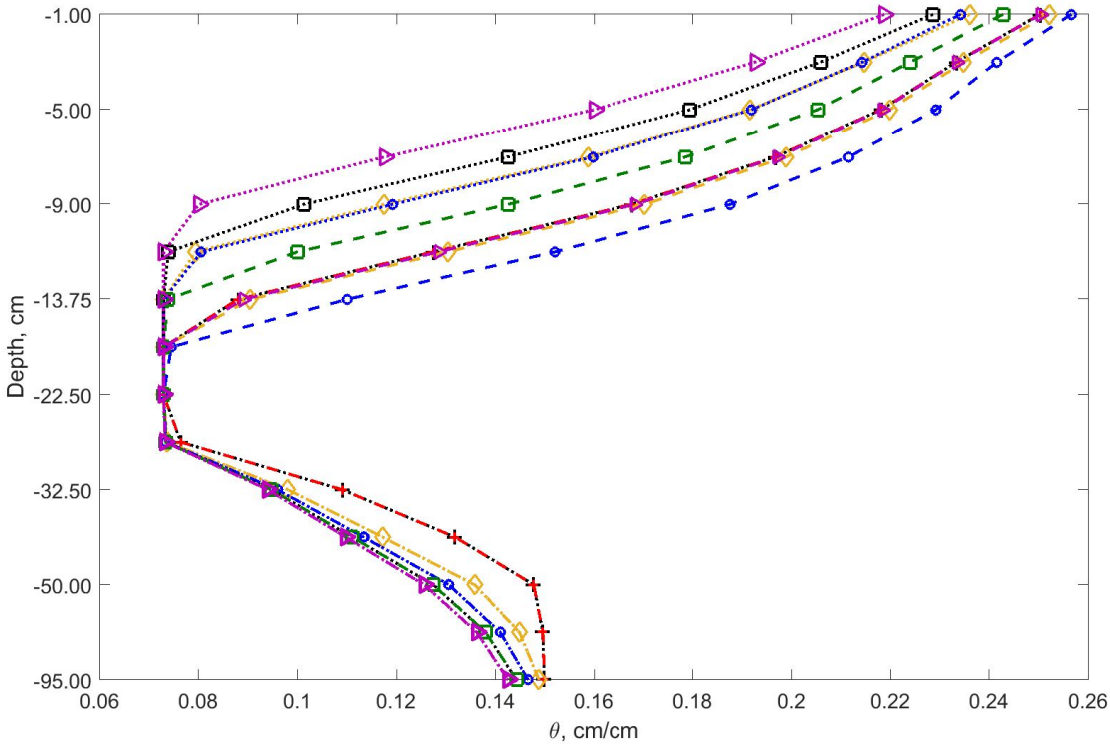


Figure 5.8: Soil moisture condition for dry weather conditions. Results for weeks 1, 5, 9, 13 and 16 is shown by markers $+$, $\diamondsuit$, 0, $\square$ and $\triangleright$, respectively. Dotted lines represents rescheduling scheme (closed-loop) and dashed lines stand for scheduling without feedback (open-loop).

Table 5.4: Soil water storage in root zone for dry weather condition.

Week no	Closed-loop scheduling	Open-loop scheduling
1	3.401	3.401
2	3.227	3.407
3	3.080	3.117
4	2.996	3.132
5	2.949	3.435
6	2.914	3.430
7	2.918	3.642
8	2.928	3.640
9	2.953	3.647
10	2.965	3.641
11	2.991	3.649
12	2.912	3.629
13	2.813	3.152
14	2.750	3.104
15	2.691	3.407
16	2.634	3.411

Table 5.5: Summary of results for dry weather condition.

	Open-loop scheduling	Closed-loop scheduling
Total pond water used (cm)	119.5474	75.6760
Pond dry outs	Close to drying at $16th$ week	No dry outs
Total amount of potable irrigation(cm)	160	160

Table 5.6: Summary of results.

	N=12 weeks	N=8 weeks
Total pond water used (cm)	436.8951	289.9045
No of times pond overflows	3	5
Total amount of overflows (cm)	331.2093	824.622

the use of pond water for irrigation. It should be noted that in closed-loop scheduling, the soil water storage is also maintained well. The improved pond level management is essentially due to the use of feedback from the pond and weather forecast.

Finally, to demonstrate the effect of prediction horizon in closed-loop scheduling a third case study is performed with the weather data and initial conditions similar to the first case, but with different prediction horizons $N = 8\,weeks$ and $N = 12\,weeks$. The results are shown in Figure 5.9. Figures 5.9(a), (b), (c) and (d) show weather condition, pond irrigation amount, pond level and potable water irrigation amount, respectively. It can be clearly observed that when prediction horizon $N = 12\,weeks$ is used, pond level (Figure 5.9(c)) is well maintained and less pond overflow occurred (Figure 5.9(d)). For $N = 8\,weeks$, though the predicted maximum pond level is reached at week 5 and 10 only, but due to weather uncertainty the overflow actually starts at week 6 and continues till week 10. The actual overflow which occurred after maximum pond level is reached is also shown in Figure 5.9(d). The soil water storage condition is similar to the first case and is not shown.

The summary of results for analyzing the effects of prediction horizon is presented in Table 5.6. It is clear that horizon $N = 12\,weeks$ performs considerably well relative to $N = 8\,weeks$. This suggests that smaller horizon significantly reduces the performance of closed-loop scheduling. Again, when this result is compared with the closed-loop results from the first case with $N = 16\,weeks$ (Table 5.3 closed-loop scheduling), the results for horizon $N = 12\,weeks$ performs slightly better. This confirms that there is some optimal horizon length beyond which the performance of closed-loop scheduling will not improve significantly anymore.

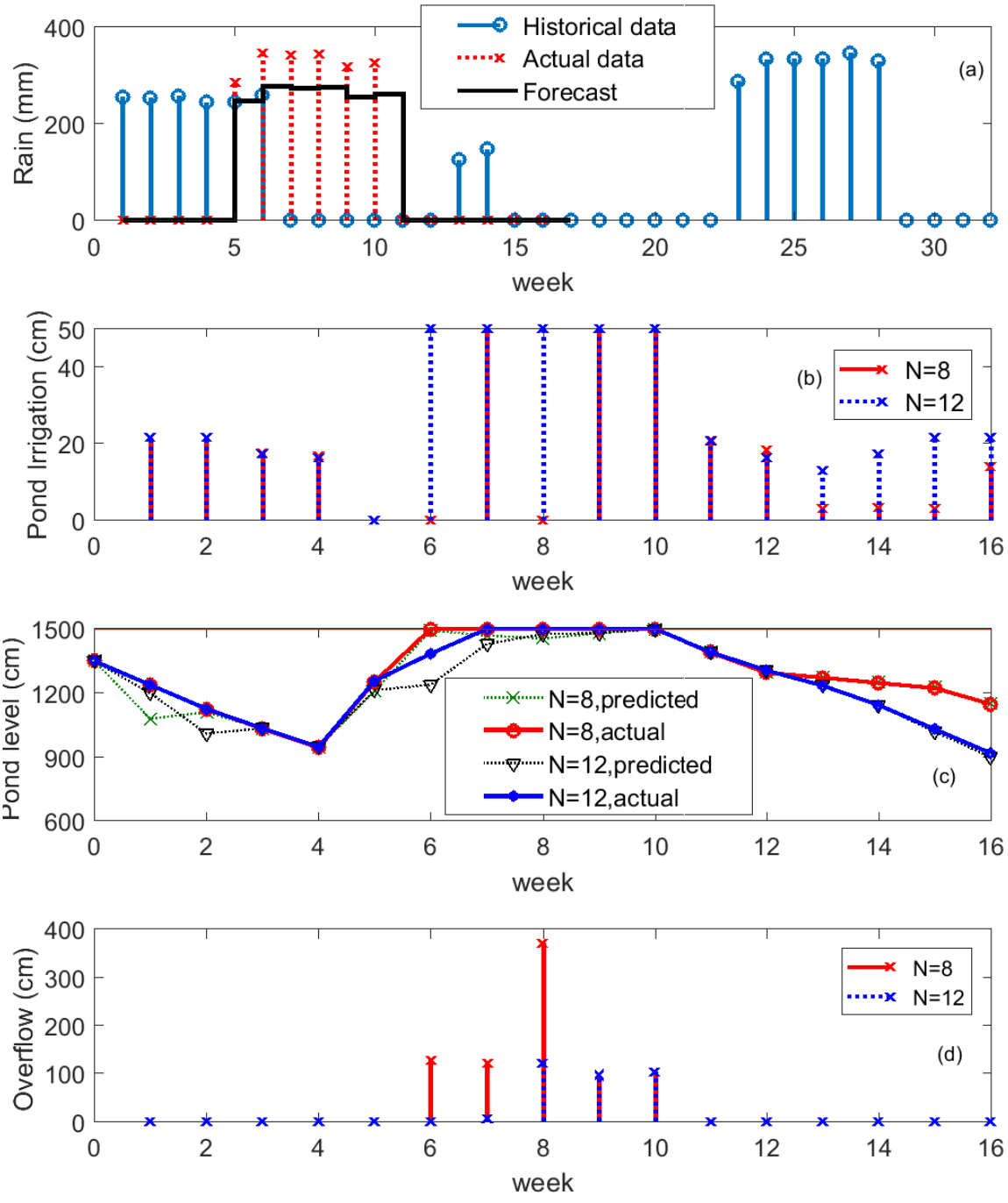


Figure 5.9: Results on effects of prediction horizon N for closed-loop irrigation scheduling for a park using city storm water.

5.6 Conclusions

Irrigation scheduling using city storm water poses new challenges in irrigation scheduling. There is a need of optimization to ensure better ground conditions and good pond levels. The results of closed-loop irrigation scheduling suggests that overflows could be reduced relative to regular open-loop irrigation scheduling. The selection of horizon for closed-loop irrigation scheduling is important since it affects the overall results. Depending on regional weather and associated uncertainties an optimum horizon needs to be identified for improved closed-loop irrigation scheduling.

The choice of costs in the scheduling problem corresponding to the two-fold objectives are very important and needs to be decided wisely. Depending on the choice, whether to emphasize field condition or pond level, the costs would be different and that would give different optimal solutions. In this study, the costs chosen for both wet and dry cases were such that the grass would not wilt, though it will face some water stress in some regions. This is done to minimize undesirable events associated with pond levels. The development and the use of steady-state model for scheduling allow us to explore the soil moisture region beyond field capacity and overcome the limitations of soil moisture deficit model.

Chapter 6

Conclusions and future work

6.1 Conclusions

In this thesis, we have identified existing challenges associated with precision irrigation and proposed some systematic approaches for closed-loop scheduling and control in irrigation.

Feedback is a key aspect of closed-loop scheduling and control, which depends highly on quality measurements. In agriculture, it is not possible to measure soil moisture at every location; therefore, state and parameter estimation is important. In Chapter 3, we designed state and parameter estimation based on system observability analysis. Specifically, the degree of observability analysis is performed by linearizing the nonlinear agro-hydrological model locally at different operating points and a novel optimal sensor placement algorithm was proposed.

In Chapter 4, we designed a closed-loop scheduler and controller to balance crop yield and irrigation water conservation. Both the scheduler and controller were designed using model predictive control (MPC). The scheduler determines the soil moisture set-points and the controller tracks the set-points given various weather and model uncertainties. An LPV model was designed to tackle the nonlinear time-varying (crop growth) agro-hydrological model and to reduce the computational complexity while optimization. The designed hierarchical scheduling and control system was applied to a field with maize. Simulation studies under different weather scenarios with different weights on water conservation were considered. It was found that the benefits of closed-loop scheduling (increased yield or reduced water usage) comes with a price (increased water usage or decreased yield). The results may be explained as

follows:

- Maize requires relatively huge amount of water in the growing season, which may have made the differences between closed-loop and open-loop scheduling less obvious.

- The study considered a farm with one crop Maize and spatial heterogeneity was not considered. This simplification may have made the benefits of closed-loop irrigation less obvious.

In Chapter 5, we extended the proposed closed-loop irrigation scheduling to a special case where irrigation is performed using city stormwater for sustainability. In this case, the objective of scheduler is two-fold; one is to maintain the pond level to avoid overflow or dry-out events and the second is to keep the soil moisture within the desired range. A steady-state model was developed and used to describe the agro-hydrological system in scheduler optimization. Based on the simulation results, the proposed approach can reduce the chance of overflow or dry-out events of the storm water pond and can adequately irrigate the field.

In conclusions, the benefits of closed-loop scheduling for irrigation depend on a few factors such as irrigation water source, weather uncertainty, field variability, the type of crop. In some cases, the closed-loop scheduling outperforms the open-loop scheduling while in other cases the benefits of closed-loop scheduling are not as obvious.

6.2 Future research directions

6.2.1 Observability and optimal sensor placement for 3D systems

Observability analysis for higher dimensional or complex systems has gained much attention in recent years due to the advancement of graph theories. Studies on degree of observability and optimal sensor placement have started to gain interests in various disciplines including health [101, 102, 103, 104], power and manufacturing industries [105, 106, 107] based on structural, dynamical and symbolic observability analysis. The implementation of these new techniques open the door to study the

observability and optimal sensor placement for soil moisture data assimilation in 3D spatial temporal systems where high number of states need to be considered.

6.2.2 Computationally efficient agro-hydrological models

The modeling of agro-hydrological systems has also evolved in the past years. While different models are available to serve different purposes, there is room to develop computationally efficient 3D models that consider the water dynamics and detailed crop biology for numeric optimization. More research may be devoted to LPV models.

6.2.3 Closed-loop scheduling and distributed MPC controller

To determine the benefits of closed-loop scheduling in agriculture irrigation, it would be interesting to perform more studies considering different crops other than maize, spatial heterogeneity and other manipulated inputs. It would be more interesting to study 3D cases with multiple crops. Distributed MPC [108, 109] may also be explored in these cases.

6.2.4 EMPC and closed-loop schedule and control

Economic MPC (EMPC) has also gained attention in recent years. In economic MPC, a general economic cost is used instead of the conventional quadratic cost [110, 111, 112]. It would be interesting to compare the performance of economic MPC with the hierarchical closed-loop scheduling and control in terms of, for example, crop yield, water usage efficiency and computational complexity.

Bibliography

[1] United States Geological Survey. Irrigation water use: Surface irrigation. Retrieved from https://water.usgs.gov/edu/irfurrow.html on June 11, 2018.

[2] Michael Pidwirny. Soil classification. Fundamentals of Physical Geography, 2nd Edition, 2006. Retrieved from http://www.physicalgeography.net/fundamentals/10v.html on June 11, 2018.

[3] Reinder A. Feddes. *Simulation of field water use and crop yield*. Pudoc, 1982.

[4] Food and Agriculture Organization of The United Nations. Coping with water scarcity - an action framework for agriculture and food security. *FAO Water Reports 38*, 2012.

[5] Food and Agriculture Organization of The United Nations. AQUASTAT 2016. Retrieved from http://www.fao.org/nr/water/aquastat/water_use/index.stm on June 8, 2018.

[6] Alberta Irrigation Projects Association. Irrigation sector conservation, efficiency, productivity plan 2005-2015. Retrieved from http://aipa.ca/wp-content/uploads/2013/11/AIPA-CEP-Final-Version-1.pdf on June 8, 2018.

[7] Cristina Milesi, Christopher D. Elvidge, and Ramakrishna R. Nemani. Assessing the extent of urban irrigated areas in the United States. *Remote Sensing of Global Croplands for Food Security*, 217, 2009.

[8] Statistics Canada, Environment, Energy and Transportation Statistics Division. Agricultural water survey. Retrieved from https://www150.statcan.gc.ca/n1/pub/16-508-x/16-508-x2016001-eng.htm on June 11, 2018.

[9] Gabriëlle J. M. De Lannoy, Paul R. Houser, Valentijn R. N. Pauwels, and Niko E. C. Verhoest. State and bias estimation for soil moisture profiles by an ensemble kalman filter: Effect of assimilation depth and frequency. *Water Resources Research*, 43(6), 2008.

[10] Jeffrey P. Walker, Garry R. Willgoose, and Jetse D. Kalma. One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: a comparison of retrieval algorithms. *Advances in Water Resources*, 24(6):631–650, 2001.

[11] Rolf H. Reichle, Jeffrey P. Walker, Randal D. Koster, and Paul R. Houser. Extended versus ensemble Kalman filtering for land data assimilation. *Journal of Hydrometeorology*, 3(6):728–740, 2002.

[12] Yixin Mao, Wade T. Crow, and Bart Nijssen. Dual state/rainfall correction via soil moisture assimilation for improved hydrologic prediction - A synthetic study using the VIC model in the Arkansas-Red River Basin. In *AGU Fall Meeting Abstracts*, 2016.

[13] Paul R. Houser, W. James Shuttleworth, James S. Famiglietti, Hoshin V. Gupta, Kamran H. Syed, and David C. Goodrich. Integration of soil moisture remote sensing and hydrologic modeling using data assimilation. *Water Resources Research*, 34(12):3405–3420, 1998.

[14] Binayak P. Mohanty, Michael Cosh, Venkat Lakshmi, and Carsten Montzka. Remote sensing for vadose zone hydrology - a synthesis from the vantage point. *Vadose Zone Journal*, 12(3), 2013.

[15] Javier Loizu, Christian Massari, Jesús Álvarez-Mozos, Javier Casalí, and Mikel Goñi. ASCAT soil moisture data assimilation through the Ensemble Kalman Filter for improving streamflow simulation in Mediterranean catchments. In *EGU General Assembly Conference Abstracts*, volume 18, 2016.

[16] Rolf H. Reichle, Gabriëlle J. M. De Lannoy, Randal D. Koster, John S. Kimball, Wade T. Crow, Qing Liu, et al. First results from the SMAP Level 4 surface

and root zone soil moisture (L4_SM) data product. In *American Meteorological Society Annual Meeting Abstracts*, 2016.

[17] Han Qiu, Narendra N. Das, Amor V. M. Ines, and Mantha S. Phanikumar. Impact of spatial resolution of SMAP soil moisture product on hydrologic modeling of soil moisture. In *AGU Fall Meeting Abstracts*, 2016.

[18] Douglas C. Baldwin, Salvatore Manfreda, Klaus M. Keller, and Erica A. H. Smithwick. Predicting root zone soil moisture with soil properties and satellite near-surface moisture data across the conterminous United States. *Journal of Hydrology*, 546:393–404, 2017.

[19] Rolf H. Reichle, Randal D. Koster, Jiarui Dong, and Aaron A. Berg. Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation. *Journal of Hydrometeorology*, 5(3):430–442, 2004.

[20] Wouter A. Dorigo, Wolfgang Wagner, Roland Hohensinn, Sebastian Hahn, Christoph Paulik, Angelika Xaver, et al. The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, 15(5):1675–1698, 2011.

[21] Hendricks Franssen, Harrrie-Jan Zhang, Wolfgang Kurtz, Stefan Kollet, and Harry Vereecken. Improved characterization of root zone soil moisture in land surface models by assimilation of groundwater level data. An example with TerrSysMP. In *EGU General Assembly Conference Abstracts*, volume 19, page 15653, April 2017.

[22] Jeffrey P. Walker, Garry R. Willgoose, and Jetse D. Kalma. One-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: A simplified soil moisture model and field application. *Journal of Hydrometeorology*, 2(4):356–373, 2001.

[23] Yaqing Gu and Dean S. Oliver. An iterative ensemble kalman filter for multiphase fluid flow data assimilation. *Spe Journal*, 12(04):438–446, 2007.

[24] Xuehang Song, Liangsheng Shi, Ming Ye, Jinzhong Yang, and I. Michael Navon. Numerical comparison of iterative ensemble Kalman filters for unsaturated flow inverse modeling. *Vadose Zone Journal*, 13(2), 2014.

[25] Daniel Erdal, Mostaquimur A. Rahman, and Insa Neuweiler. The importance of state transformations when using the ensemble Kalman filter for unsaturated flow modeling: Dealing with strong nonlinearities. *Advances in Water Resources*, 86:354–365, 2015.

[26] Hamid Moradkhani, Kuo-Lin Hsu, Hoshin Gupta, and Soroosh Sorooshian. Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resources Research*, 41(5), 2005.

[27] Jun Qin, Shunlin Liang, Kun Yang, Ichiro Kaihotsu, Ronggao Liu, and Toshio Koike. Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal. *Journal of Geophysical Research: Atmospheres*, 114(D15), 2009.

[28] Damiano Pasetto, Matteo Camporese, and Mario Putti. Ensemble Kalman filter versus particle filter for a physically-based coupled surface–subsurface model. *Advances in Water Resources*, 47:1–13, 2012.

[29] Daniel Erdal, Insa Neuweiler, and Ute Wollschläger. Using a bias aware EnKF to account for unresolved structure in an unsaturated zone model. *Water Resources Research*, 50(1):132–147, 2014.

[30] Haishen Lü, Zhongbo Yu, Yonghua Zhu, Sam Drake, Zhenchun Hao, and Edward A. Sudicky. Dual state-parameter estimation of root zone soil moisture by optimal parameter estimation and extended kalman filter data assimilation. *Advances in Water Resources*, 34(3):395–406, 2011.

[31] Che-Chuan Wu and Steven A. Margulis. Feasibility of real-time soil state and flux characterization for wastewater reuse using an embedded sensor network data assimilation approach. *Journal of Hydrology*, 399(3):313–325, 2011.

[32] Che-Chuan Wu and Steven A. Margulis. Real-time soil moisture and salinity profile estimation using assimilation of embedded sensor datastreams. *Vadose Zone Journal*, 12(1), 2013.

[33] Damiano Pasetto, Guo-Yue Niu, Luke Pangle, Claudio Paniconi, Mario Putti, and Peter A. Troch. Impact of sensor failure on the observability of flow dynamics at the Biosphere 2 LEO hillslopes. *Advances in Water Resources*, 86:327–339, 2015.

[34] Bruk E. Belayneh, John D. Lea-Cox, and Erik Lichtenberg. Costs and benefits of implementing sensor-controlled irrigation in a commercial pot-in-pot container nursery. *HortTechnology*, 23(6):760–769, 2013.

[35] Joaquín Gutiérrez, Juan F. Villa-Medina, Alejandra Nieto-Garibay, and Miguel A. Porta-Gándara. Automated irrigation system using a wireless sensor network and GPRS module. *IEEE Transactions on Instrumentation and Measurement*, 63(1):166–176, 2014.

[36] Monica Saavoss, John Majsztrik, Bruk Belayneh, John Lea-Cox, and Erik Lichtenberg. Yield, quality and profitability of sensor-controlled irrigation: a case study of snapdragon (*Antirrhinum majus L.*) production. *Irrigation Science*, 34(5):409–420, 2016.

[37] Martin S. Goodchild, K. D. Kühn, Malcolm D. Jenkins, K. J. Burek, and A. J. Dutton. A method for precision closed-loop irrigation using a modified PID control algorithm. *Sensors & Transducers*, 188(5):61, 2015.

[38] M. Bahat, Gideon F. Inbar, Oded Yaniv, and Moti Schneider. A fuzzy irrigation controller system. *Engineering Applications of Artificial Intelligence*, 13(2):137–145, 2000.

[39] Nestor M. Cid-Garcia, Angel G. Bravo-Lozano, and Yasmin A. Rios-Solis. A crop planning and real-time irrigation method based on site-specific management zones and linear programming. *Computers and Electronics in Agriculture*, 107:20–28, 2014.

[40] Yunseop Kim, Robert G. Evans, and William M. Iversen. Evaluation of closed-loop site-specific irrigation with wireless sensor network. *Journal of Irrigation and Drainage Engineering*, 135(1):25–31, 2009.

[41] Andrzej Pawlowski, Jorge A. Sánchez-Molina, Jose L. Guzmán, Francisco Rodríguez, and Sebastian Dormido. Evaluation of event-based irrigation system control scheme for tomato crops in greenhouses. *Agricultural Water Management*, 183:16–25, 2017.

[42] Honorio Navarro-Hellín, Jesús M. Rincon, Rafael Domingo-Miguel, Fulgencio Soto-Valles, and Roque Torres-Sánchez. A decision support system for managing irrigation in agriculture. *Computers and Electronics in Agriculture*, 124:121–131, 2016.

[43] David Q. Mayne, James B. Rawlings, Christopher V. Rao, and Pierre O. M. Scokaert. Constrained model predictive control: Stability and optimality. *Automatica*, 36(6):789–814, 2000.

[44] S. Joe Qin and Thomas A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11(7):733–764, 2003.

[45] Yeonjeong Park, Jeff S. Shamma, and Thomas C. Harmon. A Receding Horizon Control algorithm for adaptive management of soil moisture and chemical levels during irrigation. *Environmental Modelling & Software*, 24(9):1112–1121, 2009.

[46] Alison C. McCarthy, Nigel H. Hancock, and Steven R. Raine. Simulation of irrigation control strategies for cotton using Model Predictive Control within the VARIwise simulation framework. *Computers and Electronics in Agriculture*, 101:135–147, 2014.

[47] Dilini Delgoda, Hector Malano, Syed K. Saleem, and Malka N. Halgamuge. Irrigation control based on model predictive control (MPC): Formulation of theory and validation using weather forecast data and AQUACROP model. *Environmental Modelling & Software*, 78:40–53, 2016.

[48] Yawen Mao, Su Liu, Jannatun Nahar, Jinfeng Liu, and Feng Ding. Soil moisture regulation of agro-hydrological systems using zone model predictive control. *Computers and Electronics in Agriculture*, 154:239–247, 2018.

[49] Kelly R. Thorp, Douglas J. Hunsaker, Kevin F. Bronson, Pedro Andrade-Sanchez, Edward M. Barnes, et al. Cotton irrigation scheduling using a crop growth model and FAO-56 methods: Field and simulation studies. In *Transactions of the American Society of Agricultural and Biological Engineers*, volume 60(6), pages 2023–2039, 2017.

[50] J. Cahoon, J. Ferguson, D. Edwards, and P. Tacker. A microcomputer-based irrigation scheduler for the humid mid-south region. *Applied Engineering in Agriculture*, 6(3):289–295, 1990.

[51] Marvin E. Jensen, David C. N. Robb, and C. Eugene Franzoy. Scheduling irrigations using climate-crop-soil data. In *Proceedings of the American Society of Civil Engineers, Journal of the Irrigation and Drainage Division*, volume 96(IRI), pages 25–38, 1970.

[52] Rafael L. Bras and Jose R. Cordova. Intraseasonal water allocation in deficit irrigation. *Water Resources Research*, 17(4):866–874, 1981.

[53] Robin Wardlaw and Jonathan Barnes. Optimal allocation of irrigation water supplies in real time. *Journal of Irrigation and Drainage Engineering*, 125(6):345–354, 1999.

[54] Leila Hassan-Esfahani, Alfonso Torres-Rua, and Mac McKee. Assessment of optimal irrigation water allocation for pressurized irrigation system using water balance approach, learning machines, and remotely sensed data. *Agricultural Water Management*, 153:42–50, 2015.

[55] Sebastian Engell and Iiro Harjunkoski. Optimal operation: Scheduling, advanced control and their integration. *Computers & Chemical Engineering*, 47:121–133, 2012.

[56] Cara R. Touretzky and Michael Baldea. Integrating scheduling and control for economic MPC of buildings with energy storage. *Journal of Process Control*, 24(8):1292–1300, 2014.

[57] Vassilis M. Charitopoulos, Lazaros G. Papageorgiou, and Vivek Dua. Closed-loop integration of planning, scheduling and multi-parametric nonlinear control. *Computers & Chemical Engineering*, 2018.

[58] Jinjun Zhuge and Marianthi G. Ierapetritou. Integration of scheduling and control for batch processes using multi-parametric model predictive control. *AIChE Journal*, 60(9):3169–3183, 2014.

[59] Efstratios N. Pistikopoulos and Nikolaos A. Diangelakis. Towards the integration of process design, control and scheduling: Are we getting closer? *Computers & Chemical Engineering*, 91:85–92, 2016.

[60] Stormwater City of Guelph, Ontario, Canada 2017. Retrieved from `https://guelph.ca/living/environment/water/stormwater/` on July 8, 2018.

[61] Ben Urbonas and Peter Stahre. *Stormwater: Best management practices and detention for water quality, drainage, and CSO management*. 1993.

[62] Jianjun Chen, Richard C. Beeson Jr, Thomas H. Yeager, Robert H. Stamps, and Liz A. Felter. Potential of collected storm water and irrigation runoff for foliage and bedding plant production. *University of Florida, Florida*, 2002.

[63] Ernest O. Nnadi, Alan P. Newman, Stephen J. Coupe, and Fredrick U. Mbanaso. Stormwater harvesting for irrigation purposes: An investigation of chemical quality of water recycled in pervious pavement system. *Journal of Environmental Management*, 147:246–256, 2015.

[64] Ben R. Urbonas. Design of a sand filter for stormwater quality enhancement. *Water Environment Research*, 71(1):102–113, 1999.

[65] Christoper J. Pratt. Permeable pavements for stormwater quality enhancement. In *Urban stormwater quality enhancement: source control, retrofitting, and combined sewer technology*, pages 131–155. ASCE, 1990.

[66] Md A. Rahman, Monzur A. Imteaz, Arul Arulrajah, Mahdi M. Disfani, and Suksun Horpibulsuk. Engineering and environmental assessment of recycled construction and demolition materials used with geotextile for permeable pavements. *Journal of Environmental Engineering*, 141(9):04015019, 2015.

[67] David J. Sample and James P. Heaney. Integrated management of irrigation and urban storm-water infiltration. *Journal of Water Resources Planning and Management*, 132(5):362–373, 2006.

[68] Melanie Schwecke, Bruce Simmons, and Basant Maheshwari. Sustainable use of stormwater for irrigation case study: Manly golf course. *The Environmentalist*, 27(1):51–61, 2007.

[69] Nicholaus D. VanWoert, D. Bradley Rowe, Jeffrey A. Andresen, Clayton L. Rugh, R. Thomas Fernandez, and Lan Xiao. Green roof stormwater retention. *Journal of Environmental Quality*, 34(3):1036–1044, 2005.

[70] Guido Petrucci, José-Frédéric Deroubaix, Bernard De Gouvello, Jean-Claude Deutsch, Philippe Bompard, and Bruno Tassin. Rainwater harvesting to control stormwater runoff in suburban areas. An experimental case-study. *Urban Water Journal*, 9(1):45–55, 2012.

[71] Josefina Herrera, Carlos A. Bonilla, Lina Castro, Sergio Vera, Rodolfo Reyes, and Jorge Gironás. A model for simulating the performance and irrigation of green stormwater facilities at residential scales in semiarid and Mediterranean regions. *Environmental Modelling & Software*, 95:246–257, 2017.

[72] John R. Nimmo and Edward R. Landa. The soil physics contributions of Edgar Buckingham. *Soil Science Society of America Journal*, 69(2):328–342, 2005.

[73] Lorenzo A. Richards. Capillary conduction of liquids through porous mediums. *Journal of Applied Physics*, 1(5):318–333, 1931.

[74] Arthur W. Warrick. *Soil water dynamics*. Oxford University Press, 2003.

[75] Martinus T. Van Genuchten. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil science society of America Journal*, 44(5):892–898, 1980.

[76] Yechezkel Mualem. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resources Research*, 12(3):513–522, 1976.

[77] Marcel G. Schaap and Martinus T. Van Genuchten. A modified Mualem–van Genuchten formulation for improved description of the hydraulic conductivity near saturation. *Vadose Zone Journal*, 5(1):27–34, 2006.

[78] Jean Y. Parlange, R. Haverkamp, and J. Touma. Infiltration under ponded conditions: 1. optimal analytical solution and comparison with experimental observations. *Soil Science*, 139(4):305–311, 1985.

[79] Jos C. Van Dam, Piet Groenendijk, Rob F. A. Hendriks, and Cor M. J. Jacobs. *SWAP version 3.2: Theory description and user manual*. Wageningen UR, Alterra, Postbus 47, 6700 AA Wageningen, The Netherlands, 2 edition, 2008.

[80] Jannatun Nahar, Jinfeng Liu, and Sirish L. Shah. Observability analysis for soil moisture estimation. In *Proceedings of the Control Conference Africa CCA 2017, IFAC-PapersOnLine*, volume 50(2), pages 110–114. Elsevier, 2017.

[81] Jos C. Van Dam and Reinder A. Feddes. Numerical simulation of infiltration, evaporation and shallow groundwater levels with the Richards equation. *Journal of Hydrology*, 233(1):72–85, 2000.

[82] Richard G. Allen, Luis S. Pereira, D. Raes, and Martin Smith. FAO Irrigation and drainage paper No. 56. *Rome: Food and Agriculture Organization of the United Nations*, 56(97):e156, 1998.

[83] Chi-Tsong Chen. *Linear system theory and design*. Oxford University Press, Inc., 1995.

[84] Jing Zeng, Jinfeng Liu, Tao Zou, and Decheng Yuan. Distributed Extended Kalman Filtering for Wastewater Treatment Processes. *Industrial & Engineering Chemistry Research*, 55(28):7720–7729, 2016.

[85] Fredric M. Ham. *Determination of the degree of observability in linear control systems*. PhD thesis, Iowa State University, 1980.

[86] Fredric M. Ham and R. Grover Brown. Observability, eigenvalues, and Kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems*, AES-19(2):269–273, 1983.

[87] R. Grover Brown. Not just observable, but how observable. In *Proceedings of the National Electronics Conference, Chicago, ILL*, volume 22, pages 709–714, 1966.

[88] Lennart Ljung. *System identification*. Wiley Online Library, 1999.

[89] Biao Huang and Ramesh Kadali. *Dynamic modeling, predictive control and performance monitoring*. Springer-Verlag: Berlin, 2008.

[90] Government of Canada. Agriculture and Agri-Food Canada 2013. Retrieved from `http://sis.agr.gc.ca/cansis/soils/ab/STL/~~~~~/A/description.html` on July 8, 2016.

[91] Government of Canada. Current and historical Alberta weather station data viewer, AgroClimatic Information Service 2013. Retrieved from `http://agriculture.alberta.ca/acis/alberta-weather-data-viewer.jsp` on July 8, 2016.

[92] Robert F. Carsel and Rudolph S. Parrish. Developing joint probability distributions of soil water retention characteristics. *Water Resources Research*, 24(5):755–769, 1988.

[93] Xing Jin, Biao Huang, and David S. Shook. Multiple model LPV approach to nonlinear process identification with EM algorithm. *Journal of Process Control*, 21(1):182–193, 2011.

[94] Yolanda Bolea, Vicenç Puig, and Joaquim Blesa. Linear parameter varying modeling and identification for real-time control of open-flow irrigation canals. *Environmental Modelling & Software*, 53:87–97, 2014.

[95] Xianqiang Yang, Biao Huang, and Huijun Gao. A direct maximum likelihood optimization approach to identification of LPV time-delay systems. *Journal of the Franklin Institute*, 353(8):1862–1881, 2016.

[96] Zuhua Xu, Jun Zhao, Jixin Qian, and Yucai Zhu. Nonlinear MPC using an identified LPV model. *Industrial & Engineering Chemistry Research*, 48(6):3043–3051, 2009.

[97] Jiangyin Huang, Guoli Ji, Yucai Zhu, and Paul V. Bosch. Identification of multi-model LPV models with two scheduling variables. *Journal of Process Control*, 22(7):1198–1208, 2012.

[98] Pasquale Steduto, Theodore C. Hsiao, Elias Fereres, and Dirk Raes. FAO irrigation and drainage paper No. 66 Crop yield response to water. *FAO–Food and Agriculture Organization of the United Nations, Rome*, 2012.

[99] Amit Kohli and Karen Frenken. Evaporation from artificial lakes and reservoirs. Technical report, AQUASTAT Programme, Food and Agriculture Organization of the United Nations, 2015.

[100] Fred T. Tracy. Analytical and numerical solutions of Richards Equation with discussions on relative hydraulic conductivity. In *Hydraulic Conductivity-Issues, Determination and Applications*. InTech, 2011.

[101] Yang-Yu Liu, Jean-Jacques Slotine, and Albert-László Barabási. Observability of complex systems. *Proceedings of the National Academy of Sciences*, 110(7):2460–2465, 2013.

[102] Luis A. Aguirre, Leonardo L. Portes, and Christophe Letellier. Structural, dynamical and symbolic observability: From dynamical systems to networks. *PloS one*, 13(10):e0206180, 2018.

[103] J. D. Stigter, Daniel Joubert, and Jaap Molenaar. Observability of complex systems: Finding the gap. *Scientific Reports*, 7(1):16566, 2017.

[104] Dániel Leitold, Ágnes Vathy-Fogarassy, and János Abonyi. Controllability and observability in complex networks–the effect of connection types. *Scientific Reports*, 7(1):151, 2017.

[105] Yuanyuan Zhao, Peng Yuan, Qian Ai, and Tianguang Lv. Optimal PMU placement considering topology constraints. *International Journal of Electrical Power & Energy Systems*, 73:240–248, 2015.

[106] Siddharth Bhela, Vassilis Kekatos, and Sriharsha Veeramachaneni. Enhancing observability in distribution grids using smart meter data. *IEEE Transactions on Smart Grid*, 9(6):5953–5961, 2018.

[107] Kang He, Minping Jia, and Qingsong Xu. Optimal sensor deployment for manufacturing process monitoring based on quantitative cause-effect graph. *IEEE Transactions on Automation Science and Engineering*, 13(2):963–975, 2016.

[108] Panagiotis D. Christofides, Riccardo Scattolini, David M. de la Peña, and Jinfeng Liu. Distributed model predictive control: A tutorial review and future research directions. *Computers & Chemical Engineering*, 51:21–41, 2013.

[109] Jinfeng Liu, David M. de la Peña, and Panagiotis D. Christofides. Distributed model predictive control of nonlinear process systems. *AIChE Journal*, 55(5):1171–1184, 2009.

[110] Su Liu and Jinfeng Liu. Economic model predictive control with extended horizon. *Automatica*, 73:180–192, 2016.

[111] Moritz Diehl, Rishi Amrit, and James B. Rawlings. A Lyapunov function for economic optimizing model predictive control. *IEEE Transactions on Automatic Control*, 56(3):703–707, 2011.

[112] Matthias A. Müller, David Angeli, and Frank Allgöwer. On the performance of economic model predictive control with self-tuning terminal cost. *Journal of Process Control*, 24(8):1179–1186, 2014.

[113] Jirka Simunek, Martinus T. Van Genuchten, and Miroslav Sejna. The HYDRUS-1D software package for simulating the one-dimensional movement

of water, heat, and multiple solutes in variably-saturated media. *University of California-Riverside Research Reports*, 3:1–240, 2005.

Appendix A

Model validation

The nonlinear state-space model described in Chaper 2 was validated against HYDRUS [113] by simulating the soil moisture for a soil profile of 110 cm with four different soil layers. The first 15 cm is the first layer; the next 35 cm is the second layer; the third layer has 40 cm and the remaining 30 cm represents layer 4. In the

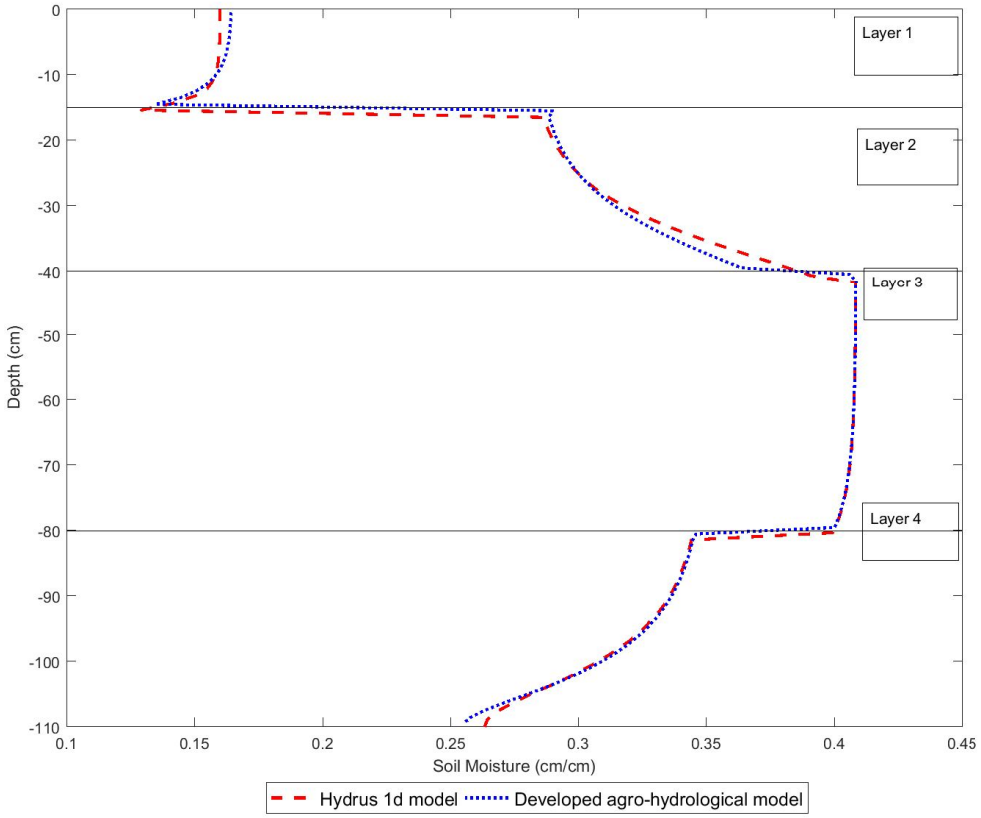


Figure A.1: Soil moisture profile at 48 hours given by HYDRUS model and the developed model.

simulation, a constant flux of 0.2 cm/h infiltration at the top boundary was considered. The bottom boundary condition was free drainage. Crops were not considered and the simulation length was 48 hours with the initial soil moisture 0.24 cm/cm. The final soil moisture is shown in Figure A.1. The figure clearly shows that the results from the developed model closely matches the results from HYDRUS.

Appendix B

More on the discretization of the agro-hydrological model

Throughout this thesis, we have used a system discretized into 15 compartments, which is decided based on the following analysis.

B.1 Data description

The study was conducted with synthetic data considering a soil profile of 110 cm with four different homogeneous soil layers. The synthetic data was generated using the agro-hydrological model. It is considered that only four measurements are available at depth 5 cm, 20 cm, 50 cm and 100 cm. The measurement data was collected on an hourly basis.

B.2 Analysis of the system

To perform the state estimation of this system, the soil profile needs to be discretized in a way that the observability of the system is ensured with the four measurements. In this study, the soil profile is discretized into 44, 33, 31, 23 and 15 nodes in order to create the system with 44, 33, 31, 23 and 15 states, respectively. The observability of the system was checked based on the linearized models along a typical state trajectory. It was found that for 44 states, the rank of the observability matrix is not full at any portion of the trajectory; therefore, the system is unobservable. The cases with 33 states and 31 states were unobservable at some parts of the trajectory as shown in Figure B.1. When the number of states is reduced to 23 states, the system becomes

observable in terms of rank test. In the figure, it can be seen that both systems with 15 states and 23 states are observable. Based on this, we further carry out the degree of observability analysis.

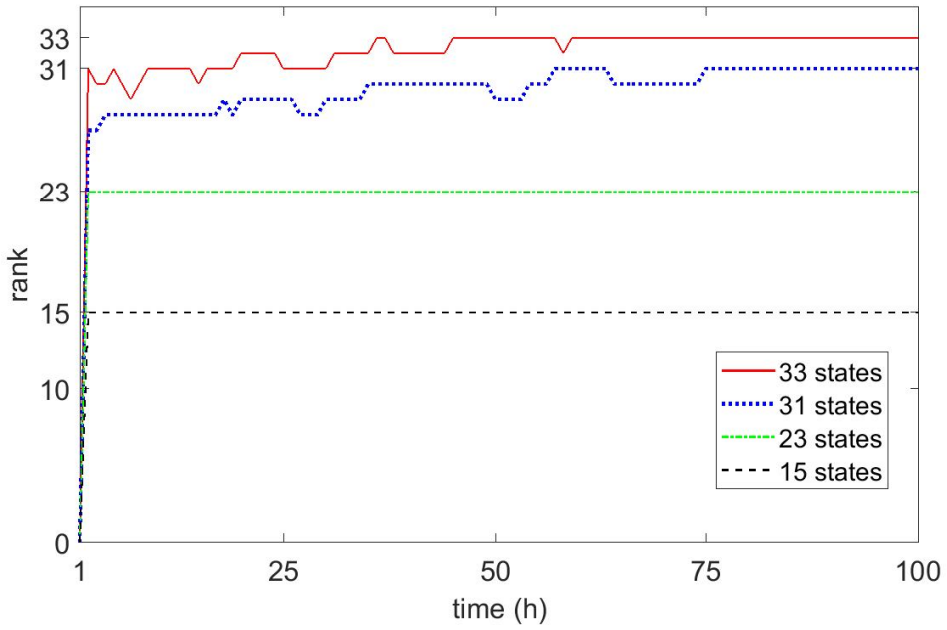


Figure B.1: The rank of the observability matrix of the system with different states.

The results of the degree of observability analysis is shown in Figure B.2. The figure shows that as we reduce the number of states, the ratio of the largest and the smallest eigenvalues $\lambda_{max}/\lambda_{min}$ decrease, which suggests that the system becomes more observable. This analysis suggests that a discretization with 15 states is a good option.

At the same time, we would like to point out that while reducing the number of states enhances the degree of observability of the system with four output measurements, we may have numerical precision issue or other numerical issues when the number of states is too few. In this work, the number of states was not reduced to be less than 15 because of these considerations.

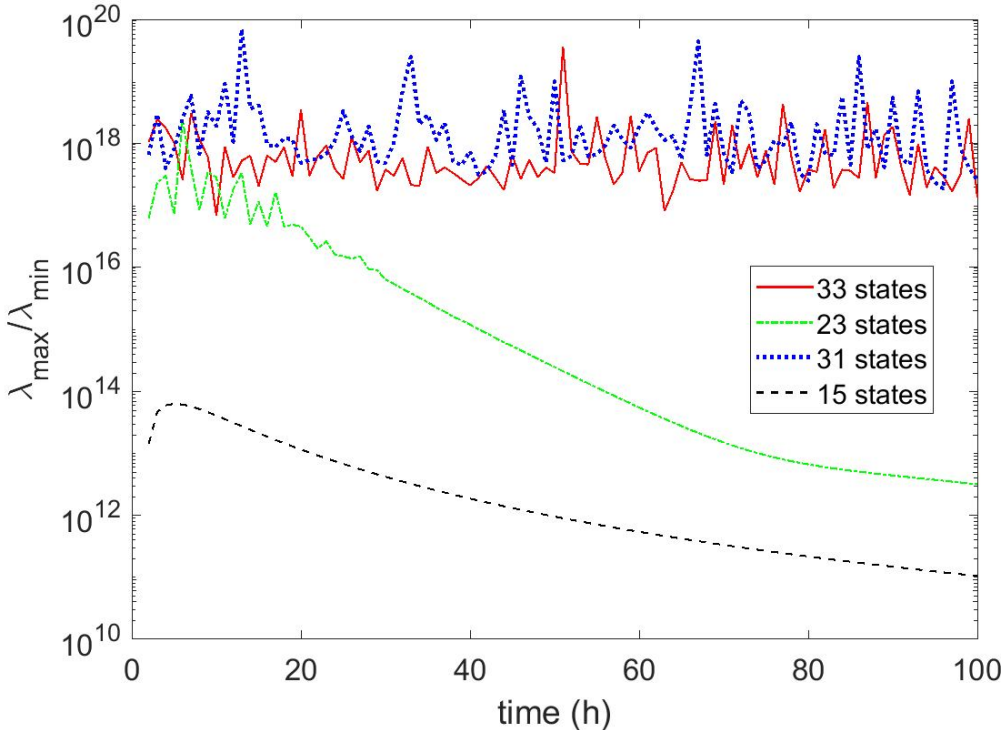


Figure B.2: The ratio of the largest and the smallest eigenvalues of matrix $\mathcal{N}^T\mathcal{N}$ for systems with 33, 31, 23 and 15 states along a typical state trajectory, where $\mathcal{N}$ is the normalized observability matrix as defined in Chapter 3.