1

Introduction

1.1 Challenges in Irrigation and a Need for Innovation

Irrigation constitutes the largest component of freshwater use within agriculture, which itself accounts for approximately 70% of global freshwater withdrawals [1]. This substantial demand is occurring alongside growing concerns over water scarcity, a challenge driven by both climate change and population growth. Agriculture is both susceptible to these pressures and partially responsible for them, particularly due to its continued reliance on inefficient irrigation practices. In many cases, irrigation is implemented as an open-loop system, where water is applied based on fixed schedules rather than actual soil moisture conditions. The absence of feedback in such systems often results in excessive water use and significant inefficiencies.

Precision irrigation has emerged as a means of addressing the inefficiencies that characterize many conventional irrigation methods. A fundamental aspect of this approach is the transition from open-loop to closed-loop irrigation systems [2]. Closed-loop systems utilize real-time feedback, most commonly from soil moisture sensors, to guide irrigation decisions, ensuring that water is applied only when it is required. Developing such systems involves the integration of optimal control strategies alongside machine learning techniques. Optimal control provides a structured approach for determining irrigation schedules that balance agronomic objectives with resource constraints. In parallel, machine learning contributes the ability to extract insights from data and support adaptive responses to dynamic field conditions.

Although both optimal control and machine learning have shown promise in irrigation applications, further advancements are needed to meet growing demands for precision and adaptability. A key priority is the refinement of existing control and learning methods to improve the accuracy of scheduling decisions within closed-loop frameworks. Moreover, current implementations often integrate these techniques in a limited or weakly coordinated manner. For example, many systems rely on a single machine learning paradigm or embed machine learning only superficially within optimal control-based irrigation schedulers. This fragmented integration has constrained the overall effectiveness of such systems. Overcoming these limitations requires a unified framework that combines optimal control and machine learning into a cohesive, closed-loop irrigation scheduling architecture.

Precision Irrigation for Agriculture: Integrating Machine Learning and Optimal Control Strategies,
First Edition. Bernard Twum Agyeman and Jinfeng Liu.

This book aims to advance the application of optimal control and machine learning in the development of closed-loop irrigation systems. Beyond refining these methods individually, it introduces a unified scheduling framework that integrates optimal control with all three major paradigms of machine learning: supervised, unsupervised, and reinforcement learning (RL). By embedding enhanced machine learning techniques and adapted control strategies within a single framework, the proposed approach seeks to improve water-use efficiency and support sustainable agricultural practices in the face of increasing global water constraints.

1.2 Current Advances and Unresolved Challenges

At its core, irrigation scheduling involves specifying the appropriate times and quantities of water application to support healthy crop development over the course of a growing season. These decisions are typically made at regular intervals, often daily or hourly, depending on the specific water requirements of the crop and the configuration of the irrigation system in use. For example, a daily scheduling approach identifies which days within a scheduling horizon require irrigation and assigns water volumes accordingly.

In settings where rainfall alone is insufficient to support plant development, effective scheduling is essential for maintaining improved crop yields. Both under- and over-irrigation present significant challenges. Inadequate water supply during key developmental stages can lead to physiological stress and diminished crop performance. On the other hand, excessive irrigation may result in surface runoff, deep percolation, or soil saturation—all of which contribute to inefficient water use and environmental degradation. These concerns highlight the importance of precision in irrigation scheduling and have motivated the development of a diverse set of tools and methodologies aimed at improving scheduling efficiency. The following section reviews major contributions, recent innovations, and emerging opportunities for advancement in this area.

In conventional approaches to irrigation scheduling, soil moisture status is commonly inferred. This inference is typically based on one of three sources: indicators of plant water stress [3–6], data from field sensor systems [7–10], or estimations of crop evapotranspiration [6, 11–13]. An irrigation event is typically triggered when the inferred soil moisture content is lower than a specified threshold. The applied water volume is then calculated with the goal of restoring the soil profile to field capacity, which is defined as the maximum moisture content the soil can retain after drainage. While these methods are appealing due to their simplicity and low computational requirements, they are inherently limited by the accuracy of the underlying estimates. Errors in soil moisture inference, whether due to model assumptions, measurement noise, or environmental variability, can lead to poorly timed or miscalculated irrigation volumes. Such inaccuracies increase the likelihood of over-irrigation, which may cause waterlogging and nutrient leaching, or under-irrigation, which can induce plant stress and reduce yields.

To improve the accuracy and reliability of irrigation schedules, recent applications have targeted process-based modeling methods [14–17]. These approaches represent interactions within the soil–crop–atmosphere system, as captured by the underlying model, to determine irrigation schedules. Mechanistic frameworks such as the soil water balance model [18], Richards' equation [16], and crop models like AQUACROP [17] are frequently employed for this purpose. While process-based models provide a more rigorous system representation compared to rule-based methods, their practical deployment is often constrained by the complexity and effort involved in calibration. In response, attention has shifted toward data-driven alternatives that infer system behavior from empirical data rather than physical principles. Black-box models, including machine learning methods such as feedforward neural networks [19–21] and recurrent neural networks [22], have been investigated as flexible tools for capturing the dynamics of crop water use and informing irrigation decisions. These models minimize the need for detailed parameterization and can generalize across field conditions when trained on high-quality data. For a comprehensive review of both conventional and model-based scheduling approaches, readers may consult [23, 24].

Approaches based on process models for irrigation scheduling can be broadly categorized into two types: those that rely on rule-based strategies and those that rely on optimization techniques. The first category, which can be referred to as nonoptimal control methods, combines process models with simple decision rules to guide irrigation [22, 23]. While these methods are relatively easy to implement and computationally inexpensive, they typically do not account for resource limitations and do not seek to optimize irrigation volumes. In contrast, optimal control-based methods formulate irrigation scheduling as an optimization problem, where decisions regarding the timing and quantity of water application are derived by maximizing or minimizing a specific performance criterion [16, 17, 25]. These approaches allow for the explicit inclusion of constraints on available water, soil moisture thresholds, and system dynamics.

Model predictive control (MPC) is a widely used optimal control technique that computes control actions by forecasting system behavior over a finite horizon, while explicitly incorporating system constraints. By anticipating future states of the Bernard Twum Agyeman system, MPC allows for irrigation plans that enhance crop performance while conserving inputs such as water and energy. A typical use case of MPC in irrigation is set-point tracking [17, 26], where the controller seeks to maintain soil moisture at a fixed target level while optimizing water use. However, recent studies have adopted a zone-control formulation [25], which seeks to determine irrigation schedules that keep soil moisture content within some optimal range. This perspective offers greater flexibility and more accurately reflects the practical objectives of irrigation scheduling—that is, to sustain soil moisture within bounds that support healthy crop development throughout the season.

Scheduling models typically incorporate discrete decisions, particularly when limited resources must be distributed across competing demands [27]. In the case of scheduling irrigation on a daily basis, deciding which specific days within a horizon should receive irrigation represents a binary choice—whether to irrigate or not on each day. As such, a full treatment of the scheduling problem within an optimal control framework (or MPC framework) must account for both discrete and continuous control variables: the former governing the timing of irrigation events, and the latter specifying the quantity of water to apply. Some studies [25, 26] address irrigation timing outside the predictive control loop using heuristic rules. While straightforward, such strategies cannot guarantee that the resulting schedule is optimal. Other implementations [17] implicitly deduce irrigation timing from the magnitude of the irrigation amount, interpreting a near-zero value as an indication to skip irrigation on a given day. Although this simplification reduces complexity, it fails to explicitly control the timing of irrigation and can misrepresent actual water needs. In agronomic cases where a small yet essential quantity of water is required, threshold-based approaches may erroneously suppress irrigation events if the required amount falls below the predefined threshold. Furthermore, empirical results indicate that these implicit strategies often produce schedules characterized by frequent low-volume irrigation events, which can drive up operational costs. While careful tuning of the MPC cost function could mitigate this issue by imposing a stronger trade-off between timing and volume, the tuning process itself is nontrivial and problem-dependent. Importantly, recent advances in computational power and optimization tools remain underutilized in this domain. There is considerable potential to incorporate discrete decision-making directly within the MPC framework to achieve more efficient and practical scheduling outcomes.

Recent progress in sensing technologies and meteorological forecasting, combined with the increased accessibility of computational resources, has created favorable conditions for the application of machine learning in irrigation scheduling. These methods are increasingly explored as alternatives to traditional rule-based systems, offering the ability to extract insights from large volumes of historical data. Machine learning provides a flexible foundation for modeling complex relationships and has been applied under three primary paradigms: supervised learning, which supports predictive modeling; unsupervised learning, which facilitates pattern discovery; and RL, which enables adaptive decision-making in uncertain environments. For an extensive overview of machine learning applications in this domain, the reader may consult [28, 29].

Methods such as k-means clustering [30, 31], principal component analysis [32], and fuzzy clustering [33] have been extensively applied to divide large agricultural fields into irrigation management zones (MZs). These zones are typically delineated using spatial attributes such as soil electrical conductivity, yield patterns, and topographic variation. Defined as field subregions with relatively uniform soil and crop characteristics, MZs provide a systematic means of addressing spatial variability in large-scale agricultural systems. Incorporating these zones into closed-loop irrigation schedulers enables localized control strategies, facilitating more precise water application and improved resource efficiency. However, despite the demonstrated value of unsupervised learning for zone delineation, significant challenges remain in developing zone-specific model-based irrigation controllers. A key limitation lies in the incomplete integration of essential model parameters, particularly soil hydraulic properties. Acquiring high-resolution spatial data on these parameters across large fields is often both labor-intensive and costly. One practical solution involves estimating them through inverse modeling approaches. This underscores the need for computationally efficient methods capable of inferring spatial distributions of soil hydraulic properties to support the development of MZs that are compatible with process-based irrigation models.

Soil moisture and hydraulic property estimation have been the subject of extensive research, with numerous techniques developed to infer these variables from observational data. Methods such as the extended Kalman filter (EKF) [34, 35], the ensemble Kalman filter (EnKF) [36–40], and the moving horizon estimator [41, 42] have been widely applied for this purpose. A common limitation of these approaches is their dependence on point-based soil moisture measurements, which constrains their ability to capture the spatial variability of soil hydraulic parameters across large fields. Remote sensing technologies, particularly those based on microwave sensing, offer a promising alternative by providing spatially distributed soil moisture data, enabling parameter estimation at scale. Furthermore, several studies developing estimation frameworks overlook the importance of assessing parameter estimability. Estimability, which evaluates whether specific model parameters can be uniquely determined from available measurements [43, 44], is a critical yet frequently neglected step. Its inclusion is especially important in large-scale agricultural settings, where the inability to uniquely estimate certain parameters can lead to unreliable estimates. Moreover, given the sparse and often noisy nature of soil moisture data, attempting to estimate all hydraulic parameters across the field is rarely feasible. A more practical strategy involves identifying a subset of parameters that can be estimated with sufficient accuracy. Integrating a parameter selection step into the estimation process not only enhances the robustness of the results but also reduces computational burden by lowering the dimensionality of the problem. In summary, improving current estimation methodologies requires the combined use of spatially distributed soil moisture observations, formal estimability analysis to identify reliably estimable parameters, and parameter selection techniques to focus on those that are both estimable and informative.

Reliable predictions of weather conditions and soil moisture are essential for effective irrigation scheduling, as they guide decisions about when to irrigate and how much water to apply. Supervised learning techniques have played a key role in this area by leveraging historical data to construct predictive models tailored to specific field conditions. A wide range of algorithms has been applied for this purpose, including support vector machines [45, 46], decision trees [47, 48], random forests [49, 50], k-nearest neighbors [51, 52], feed-forward neural networks [19–21], and long short-term memory networks [22]. Among these applications, one of the most critical challenges is accurately estimating soil moisture within the crop root zone, as this directly impacts plant water uptake and the precision of scheduling decisions. Models addressing this task often adopt ensemble learning strategies [19] or multi-input, multi-output architectures [53]. While these models are well-suited for rule-based scheduling systems, their computational complexity can hinder integration into optimization-based frameworks. This underscores the need for supervised models that are not only accurate but also computationally efficient and compatible with real-time, control-oriented irrigation scheduling.

RL has emerged as a promising tool for optimizing irrigation scheduling by enabling agents to learn effective water management strategies through interaction with their environment. Prior applications [54–57] have adopted model-free RL approaches, such as Q-learning, to develop policies that respond to changing field conditions. These approaches have demonstrated improvements in water-use efficiency and crop yield across a variety of settings. Recent developments in RL, such as hybrid actor-critic algorithms [58] and hierarchical RL architectures [59], provide mechanisms for jointly managing discrete and continuous action spaces. These methods offer a pathway to more effective control policies, where RL agents can simultaneously determine whether to irrigate and, if so, how much water to apply.

Scaling daily irrigation scheduling to large agricultural fields composed of multiple MZs introduces an added level of complexity. In this setting, scheduling must account for zone-specific decisions regarding both the timing and quantity of irrigation. One effective way to structure this problem is through a multi-agent system (MAS), in which each zone is assigned an autonomous agent responsible for making local irrigation decisions. MAS-based approaches have been explored in the irrigation domain [60], with agent-based modeling offering a common framework for simulating decentralized coordination [61, 62]. While such models support local decision autonomy, they often rely on predefined rules and lack mechanisms for adapting to dynamic environments. Recent progress in RL has demonstrated the potential to overcome this limitation by enabling agents to learn from interaction with their environment. When integrated into MAS frameworks, RL gives rise to multi-agent reinforcement learning (MARL), which supports both collaborative and independent learning among agents. This paradigm presents an opportunity for addressing the complexities of irrigation scheduling in large-scale fields that are delineated into MZs.

During the development of MARL frameworks, the initial approach in many settings is to adopt a fully observable Markov decision process formulation, where each agent is assumed to have access to the true underlying state of its environment. While this assumption simplifies the training process, it is often impractical in real-world applications such as agricultural systems. In these contexts, complete information about the system state (i.e., the spatial distribution of soil moisture in the root zone) is rarely available. Instead, agents must rely on partial observations obtained through sensor networks, which offer only limited insight into the full environmental state. As such, while it is reasonable to begin the development of MARL-based irrigation systems under the fully observable assumption, the insights gained from this simplified setup should inform the design of more realistic frameworks. These subsequent frameworks should relax the full observability assumption and incorporate mechanisms for decision-making under partial observability.

Recent advancements in sensor technology have made it feasible to obtain high-resolution weather data, including hourly forecasts. Until now, the discussion in this section has focused primarily on daily irrigation scheduling, where decisions are based on daily weather summaries and soil moisture estimates to determine both the timing and amount of irrigation. While daily data offer useful aggregated insights, they often mask important intraday fluctuations in weather variables that can significantly influence crop water requirements. To better capture these sub-daily dynamics, a more comprehensive scheduling approach would incorporate both daily and hourly weather information into the decision-making process. However, in the context of MPC, especially when the problem includes both discrete and continuous decision variables, naively combining multi-timescale data can result in a substantial increase in computational complexity. This motivates the need for computationally efficient strategies to manage and integrate information across different temporal resolutions. One promising direction comes from process systems engineering, where hierarchical control architectures are commonly used. Applied to irrigation, a similar structure could involve a scheduling layer that operates at a daily level—using daily weather forecasts to plan irrigation timing and volumes—alongside a lower-level control layer that responds to hourly weather fluctuations to ensure that the daily schedule is executed accurately. Such a hierarchical framework will allow for scheduling at longer timescales while retaining the flexibility to adapt to short-term environmental variability, thereby improving both the adaptability and efficiency of the irrigation system.

In conclusion, the current body of research on irrigation scheduling reveals several critical gaps. These include the lack of an integrated framework combining optimal control with all three machine learning paradigms, and the absence of strategies that jointly address discrete and continuous decisions in daily scheduling. Efficient, scalable methods for field-scale hydraulic parameter estimation are also underdeveloped. Moreover, supervised learning models that accurately represent soil moisture dynamics and align with control-based scheduling remain limited. RL methods capable of mixed-action control, as well as multi-agent approaches suited for zone-based field management, have yet to be fully explored. Finally, progress in hierarchical frameworks that couple scheduling with control to leverage multi-timescale information remains minimal.

1.3 Objectives and Organization of the Book

Motivated by the lack of a unified irrigation scheduling framework that integrates optimal control with the three machine learning paradigms, this book has the following broad objectives:

1. To design a computationally efficient approach for estimating soil hydraulic parameters at the field scale.
2. To develop an MPC-based irrigation scheduler that employs both discrete and continuous controls to address the daily irrigation scheduling problem in homogeneous fields and spatially variable fields characterized by multiple MZs.
3. To create a unified framework that integrates the MPC-based irrigation scheduler with unsupervised learning for management zone delineation (using estimated hydraulic parameters), a streamlined supervised learning model for soil moisture modeling within the delineated zones, and RL to enhance the computational efficiency of the overall scheduling framework.
4. To develop a MARL framework for daily irrigation scheduling in large-scale fields characterized by multiple MZs.
5. To develop a hierarchical framework that integrates scheduling and advanced control, enabling the use of multi-timescale information for adaptive and efficient irrigation decision-making.
6. To demonstrate the application of the unified scheduling framework in large-scale agricultural fields.

The book is organized as follows. Chapter 2 reviews fundamental concepts of soil moisture modeling in agricultural fields. This chapter introduces and derives the Richards' equation, which is used throughout this book to describe soil moisture dynamics in large-scale agricultural fields. Additionally, a numerical solution technique for the Richards' equation is presented.

In Chapter 3, machine learning and MPC are reviewed. For machine learning, the three main paradigms are introduced, followed by a detailed discussion of techniques such as feedforward neural networks, long short-term memory networks, physics-informed neural networks, k-means clustering, actor-critic RL methods, proximal policy optimization, and MARL. In the context of MPC, key aspects such as formulation, feasibility, and tuning are examined.

Chapter 4 presents estimation of soil moisture and hydraulic parameters in the context of large-scale agricultural fields. The chapter introduces an estimation framework that leverages soil moisture data obtained from microwave remote sensors mounted on a center pivot irrigation system. A key component of this framework is a tailored sensitivity analysis, adapted to account for the characteristics of the remotely sensed measurements, which is used to evaluate parameter estimability prior to estimation. To enhance reliability, an orthogonal projection method is applied to determine a subset of hydraulic parameters suitable for unique estimation. This is followed by a modified EKF designed to incorporate the results of the parameter selection step. The proposed approach is validated through both a simulated scenario and a real-world case study involving a segment of a large-scale field.

In Chapter 5, the computational challenges associated with using high-dimensional and complex models for soil moisture estimation in large-scale agricultural fields are addressed through a performance-triggered model reduction framework. This framework employs a structure-preserving reduction approach and dynamically adjusts the order of the reduced model based on prediction performance. This enables the system to capture evolving field dynamics while maintaining computational efficiency. The reduced-order modeling approach is integrated with a state estimation scheme based on the EKF. A simulated case study on a large-scale field is used to demonstrate the framework's ability to effectively balance estimation accuracy with computational cost.

Chapter 6 presents a mixed-integer MPC with zone control to solve the irrigation scheduling problem on a daily basis. The chapter provides tailored formulations for both uniform and spatially heterogeneous fields. To mitigate the computational demands associated with mixed-integer optimization, a data-driven modeling strategy is presented. Furthermore, the chapter introduces an alternative formulation that substitutes binary decision variables with a smooth approximation. The effectiveness of the proposed MPC approaches is illustrated through two case studies: one involving a uniform field and the other a spatially heterogeneous field.

Chapter 7 presents a framework that combines mixed-integer MPC with all three paradigms of machine learning to tackle irrigation scheduling in fields that exhibit spatial variability. The chapter begins by outlining a three-step management zone delineation strategy that employs estimated soil hydraulic parameters alongside k-means clustering to partition fields into zones suitable for process-based modeling. It then details the use of long short-term memory networks to construct surrogate models of soil moisture for each zone, capturing moisture variability within the root zone while accounting for changes in rooting depth over time. To solve the discrete decision component of the mixed-integer MPC, the framework incorporates decentralized hybrid proximal policy optimization agents. These agents, in tandem with a heuristic strategy, significantly reduce the computational burden of scheduling. The proposed framework is validated through an application to a section of a large-scale field and benchmarked against a standard threshold-based irrigation method.

Chapter 8 presents a semi-centralized MARL framework to address daily irrigation scheduling in fields exhibiting spatial variability. The framework adopts a hierarchical structure that separates decision-making across two levels: a centralized agent is responsible for determining the timing of irrigation events, while decentralized agents independently regulate the irrigation amounts within their respective zones. To address the non-stationarity often encountered in multi-agent settings, a state augmentation strategy is introduced. The effectiveness of the semi-centralized approach is demonstrated through its application to a spatially heterogeneous field and evaluated against the decentralized MARL method presented in Chapter 7.

Chapter 9 presents a hierarchical framework that combines semi-centralized MARL with an advanced control layer to address irrigation scheduling in large-scale, spatially variable fields. The scheduler, trained under a partially observable Markov decision process framework, constructs belief states using an EKF. Operating at the upper level of the hierarchy, the scheduler generates daily irrigation schedules by incorporating daily soil moisture estimates, weather forecasts, and crop information. The lower level consists of an MPC layer, which refines and executes these schedules using higher-frequency inputs such as hourly weather data, real-time crop conditions, and sensor feedback. The effectiveness of the proposed hierarchical approach is demonstrated through its application to a large-scale field scenario.

Finally, Chapter 10 outlines potential research directions aimed at advancing and refining the methodologies presented in the preceding chapters.