Efficient Forecasting of Crop Water Demand

A Thesis

submitted for the award of the degree of

Doctor of Philosophy

in

Computer Science and Engineering Department

Submitted by

Ravneet Kaur Sidhu
(Reg no: 951403006)

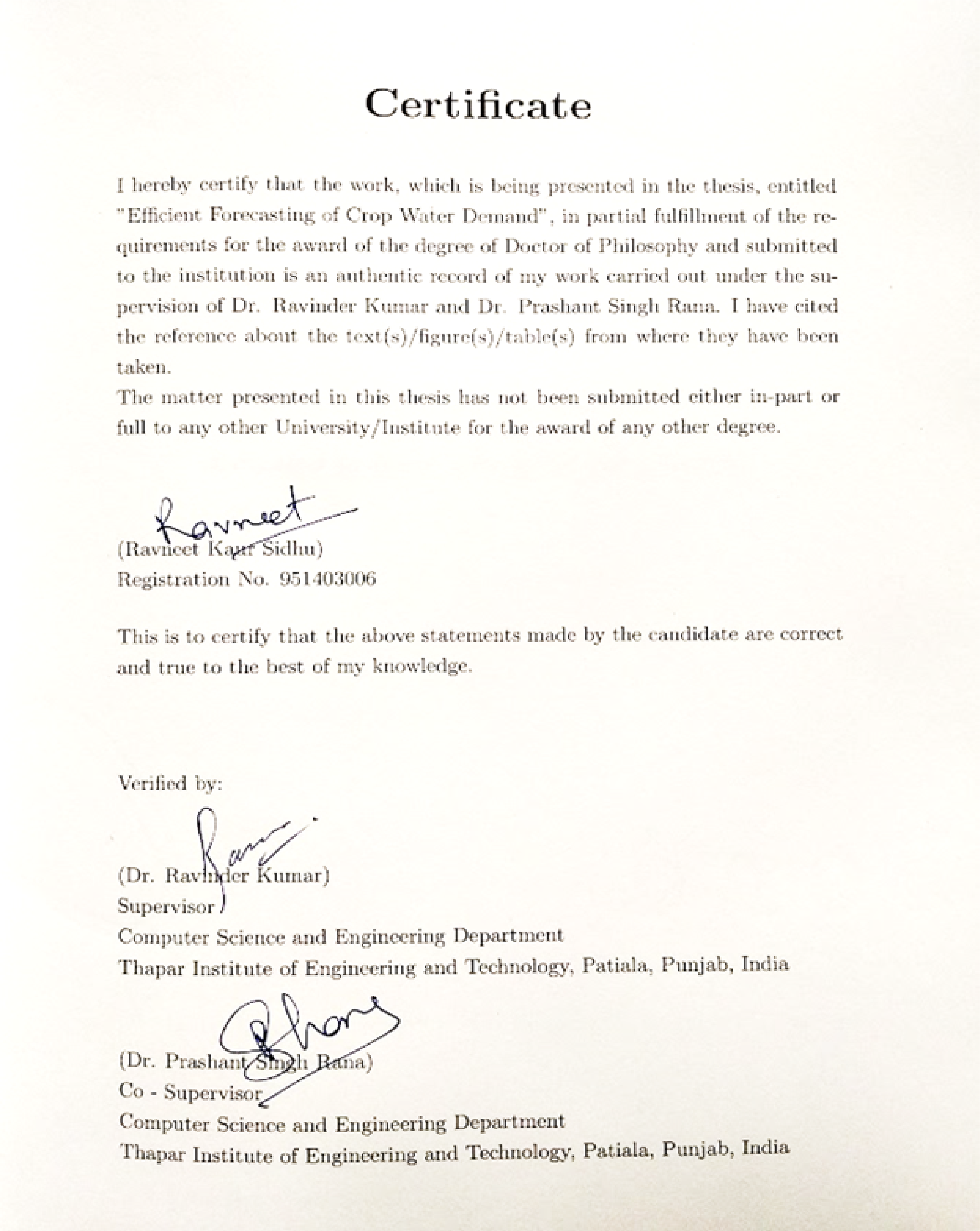
Under the Guidance of

Dr. Ravinder Kumar
Associate Professor

Dr. Prashant Singh Rana
Associate Professor

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Thapar Institute of Engineering and Technology,
Patiala, Punjab - 147004, India
June 2021



Certificate

I hereby certify that the work, which is being presented in the thesis, entitled "Efficient Forecasting of Crop Water Demand", in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy and submitted to the institution is an authentic record of my work carried out under the supervision of Dr. Ravinder Kumar and Dr. Prashant Singh Rana. I have cited the reference about the text(s)/figure(s)/table(s) from where they have been taken.
The matter presented in this thesis has not been submitted either in-part or full to any other University/Institute for the award of any other degree.

Ravneet
(Ravneet Kaur Sidhu)
Registration No. 951403006

This is to certify that the above statements made by the candidate are correct and true to the best of my knowledge.

Verified by:

(Dr. Ravinder Kumar)
Supervisor
Computer Science and Engineering Department
Thapar Institute of Engineering and Technology, Patiala, Punjab, India

(Dr. Prashant Singh Rana)
Co - Supervisor
Computer Science and Engineering Department
Thapar Institute of Engineering and Technology, Patiala, Punjab, India

dedicated to those who go to sleep hungry.

Acknowledgements

I thought this day would never arrive. Now that it has, my heart overflows with gratitude for those who have been with me throughout the ordeal.

I want to express my deep gratitude to Dr. Ravinder Kumar and Dr. Prashant Singh Rana, my research supervisors, for their patient guidance, enthusiastic encouragement, and useful critiques of this research work. Their guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better team to mentor my Ph.D. study.

I want to express my sincere gratitude to Dr. Kumar for accepting me as his student. I deeply thank him for his constant support and for giving me the freedom to explore my intellectual curiosity without objection. His advice and encouragement were always important guiding lights for when I lost my focus.

Sincere thanks must also go to Dr. Rana for his continuous support by always willing to answer my many questions. His motivation and enthusiasm are contagious. I can only hope that I have been able to absorb some of his magical intuition for doing interesting collaborative research.

I would also like to thank my former supervisor Dr. Damandeep Kaur who helped me to start my doctoral research and taught me how to find scientific references efficiently.

I want to thank the members of my thesis committee: Dr. Rajendra Kumar Sharma, Dr. Sushma Jain and Dr. Mukesh Singh. They generously gave their time to offer insightful comments towards improving my work.

The contribution of my friends, those from before, cannot be left unsaid. Their visits and telephonic conversations were always a source of immense joy. To those who I was fortunate enough to befriend during my stay here, I thank you for sharing this journey with me. Your presence and continued support has helped me sail through this. A heartfelt thank you to my seniors who were like elder siblings, usually showering love and care but also keeping us in check, as and when needed.

I thank my fellow labmates for not only their stimulating discussions and valuable suggestions; but also for all the fun we shared. I truly appreciate my colleagues, including those from other disciplines, for enriching me by sharing their experiences. We've all been there for one another and have taught ourselves and each other many tools. I know that I could always ask for advice and opinions on any issue that I may be dealing with. I'm thankful to the graduate and post-graduate students at TIET. My life would have been very dull without their fun interactions.

Numerous faculty members have always been very kind with their words of encouragement. Crossing paths with them while walking on campus has ever brought a smile to my face. I wish to give them my sincerest thanks. The office staff has always been available and ready for the assistance of any kind. The security guards, the hostel staff, and mess staff have not only been extremely diligent with their duties but have often accommodated requests with a smile. They have made my stay at the campus a pleasurable one.

I wish to thank my family for always having my back. Irrespective of what they were dealing with on an individual level, they were always available for a conversation. I owe a lot to my parents, who encouraged and helped me at every stage of my personal and academic life, and longed to see this achievement come true. My sister has been my best friend all my life; I love her dearly and thank her for her care and concern. Interacting with my niece has always allowed me to drop all baggage and enjoy the moment. I thank her for her constant concern and encouragement, the most effective I've ever come across so far. Not only was home a sanctuary, they made numerous trips to meet me at the campus, ensuring that my spirits were kept flying high. I would not have made it this far without them.

My research would have been impossible without the support of Dr. M L Jat, Dr. H S Jat, and Dr. Balwinder Singh at the International Maize and Wheat Improvement Center (CIMMYT) and Dr. H S Sidhu at Borlaug Institute for South Asia (BISA), Ludhiana. My wholehearted thanks to them for sharing their domain knowledge, without which this research would have been an uphill task.

Abstract

Water plays an important role in the creation of everything we produce. Around 70% of freshwater around the world and up to 95% in several developing countries is used for farming. Agriculture is the main sector for global water use. The growing water scarcity is one of the major challenges for agriculture. However, water resources need to be used more efficiently and sustainably in agricultural production as population growth, coupled with increasing urbanization and industrialization, will cause less and less water to become less available for cereal production. Traditional systems of irrigation and water use and management in agriculture are highly inefficient with low water productivity, and cannot ensure long-term sustainable food security. Many innovative technologies, such as micro-irrigation (low volume irrigation), offer an efficient alternative to traditional flood irrigation. Drip irrigation (surface and subsurface) systems provide water (and nutrients) to the crop root zone, where it can be utilized most efficiently. Recent advances in the science of sensor technologies and the internet of things (IoT) can be useful in the automation of irrigation systems, which can help in addressing the emerging challenges of labor shortages and inefficiency of water use in agriculture. The use of wireless sensor networks (WSNs), IoTs, and communication technology for the automation of irrigation in general and drip irrigation in cereal crops, in particular, is an entirely new and futuristic field. The purpose of this study has been to forecast the crop water demand efficiently. To sustain and improve agriculture production from the depleting freshwater resources. In this work, various parameters affecting the crop water demand have been identified. The problem of forecasting crop water demand has been formulated as classification, regression, and a deep learning problem. Hence, various machine learning models and two deep models have been implemented, and the most suitable for irrigation demand forecasting have been selected. Machine learning classification models have been found to predict the need for irrigation accurately. While the regression and deep learning models have predicted, the amount of water to be applied quite well. The best performing models were ensembled, and an algorithm defining the approach was proposed and implemented. We also used dynamic simulation model APSIM (Agriculture Production System Simulator) which is being used commercially for irrigation scheduling in many countries, however, it needs intensive data to calibrate and simulate the irrigation scheduling. We compare the APSIM

results to our results to see if there is any loss of prediction accuracy if we used data-driven models which are based on readily available datasets.

Table of Contents

Title **Page**
No.

List of Figures

Figure No. **Title** **Page No.**

List of Tables

Table No. **Title** **Page No.**

List of Abbreviations

ABR	Adaboost Regressor
ANNs	Artificial Neural Networks
BCM	Billion Cubic Meters
CA	Conservation Agriculture
CAGR	Compound Annual Growth Rate
CGWB	Central Groundwater Board
CPE	Cumulative Pan Evaporation
CRI	Crown Root Initiation
DA	Digital Agriculture
DI	Deficit Irrigation
DTR	Decision Tree Regressor
ET	Evapotranspiration
ETR	Extra Tree Regressor
EVS	Explained Variance Score
FIRB	Furrow-Irrigated Raised Beds
GBR	Gradient Boosting Regressor
GIA	Gross Irrigated Area
GPRS	General Packet Radio Service
HDPE	High-Density Polyethylene
IS	Irrigation Scheduling
IW	Irrigation Water
KNN	K-Nearest Neighbour
LR	Linear Regression
LSTM	Long Short Term Memory
M2M	Machine-To-Machine
MAE	Mean Absolute Error
MSE	Mean Square Error
Pan E	Open Pan Evaporation
PAWC	Plant Available Water Capacity
PRD	Root-Zone Drying
QR	Quick Response
RF	Radio-Frequency
RFID	Radio Frequency Identification
RFR	Random Forest Regressor

RNN	Recurrent Neural Network
SDD	Stress Degree Days
SMC	Soil Moisture Content
SMP	Soil Matric Potential
SMP	Soil Matric Potential
SSDI	Subsurface Drip Irrigation
SVR	Support Vector Regressor
TDR	Time-Domain Reflectometry
WIU	Wireless Information Unit
WP	Water Productivity
WSN	Wireless Sensor Network
WSUs	Wireless Sensor Units
WUE	Water Use Efficiency

Chapter 1

Introduction

By 2050, the world's population will increase to about 9.15 billion from the current level of 7.79 billion; the population of South Asia will form a large percentage (24.5%) of the global population, with India accounting for 17.9% of the total world population (Table 1.1). Cereals continue to be an essential part of the human diet [1]. Wheat, maize, rice, and, to some extent, millet and sorghum are major food items crucial to the survival of billions of people around the globe. The South-Asian region is one of the world's key breadbaskets, producing almost 20% of global wheat and 31% of rice. These are the staple diet of the majority of the world's population. The International Food Policy Research Institute's report [2] indicates that the total cereal production in the world and South Asia will have to be increased by 50.1% and 62.7% respectively by 2050 from the 2010 level under the no-climate-change scenario. Whereas, under the climate-change scenario, likely increase is estimated to be only 38.7 and 48.4% in the world and in South Asia, respectively, indicating the severe implications of the effects of climate change for cereal production and food security (Table 1.2).

Water is a crucial input of the agricultural production process, and its availability is also related to nearly all the socioeconomic and environmental impacts of climate and population changes and their implications. The imbalance between water availability and water demand leads to water scarcity, which is one of the most urgent challenges in the world [3]. Moreover, global water demand is projected to increase by 55% between 2000 and 2050, i.e., from 3500 to 5425 km^3. Evidence has shown that climate change hurts world water resources and food production with a high degree of regional variability and scarcity [4]. Meeting this projected need in South Asia is doubly challenging, considering that most of the land suitable for farming is already under cultivation and that almost half of it faces multiple climatic hazards such as water shortage and heat stress. Sustainably increasing agricultural production to meet the growing demand for food, especially with a growing scarcity of water, is a significant issue under the changing climate. South Asia is home to nearly 25% of the global population

Table 1.1: Population trend and projection for India, South-Asia and the world [7]

Population / Year	**1950**	**1960**	**1970**	**1980**	**1990**	**2000**	**2010**	**2015**	**2020**	**2030**	**2050**
World population (Billions)	2.54	3.03	3.7	4.46	5.33	6.14	6.96	7.38	7.79	8.31	9.15
South Asia population (Billions)	0.48	0.57	0.71	0.9	1.13	1.39	1.64	1.75	1.86	2.02	2.24
India population (Billions)	0.38	0.45	0.56	0.7	0.87	1.06	1.23	1.31	1.38	1.5	1.64
Share of India in world population (%)	14.84	14.85	15	15.68	16.39	17.2	17.74	17.75	17.7	18.1	17.91
Share of South Asia in world population (%)	18.78	18.87	19.29	20.2	21.28	22.64	23.56	23.7	23.82	24.26	24.5

Table 1.2: Impact of climate change on production and per capita consumption of cereals [2]

	Total cereal production (million tons)					**Per capita food consumption (kg per capita per year)**				
		Without climate change		With climate change			Without climate change		With climate change	
	2010	2030	2050	2030	2050	2010	2030	2050	2030	2050
World	2155	2746	3235	2622	2989	143.5	146.7	148.3	143.4	140.4
South Asia	279	384	454	362	414	148.5	150.7	154.1	147.5	145.9
Production in South Asia as % of world	12.95	13.98	14.03	13.81	13.85	-	-	-	-	-

but contains only about 4.6% of the global annual renewable water resources, which are very unevenly distributed between countries and river basins [5, 4]. Except for Bhutan and Nepal, per capita water availability is below the world average and continues to decline as populations increase. More than 90% of all groundwater extracted in South Asia is used by the agriculture sector, compared with 70% globally. The 60% of agricultural water use in South Asia is from surface water sources and 40% from groundwater. Cereal crops, especially Rice, are major consumers of water in the region's agriculture sector. Resource shortages, driven by climatic, institutional and social changes in many regions of Asia, combined with growing imperatives to increase food production whilst ensuring environmental sustainability, are driving research into modified agricultural practices [6]. Therefore, appropriate strategies and modern tools and technologies for cereal production, especially those aimed at producing more with less water and still minimizing the impact of environmental externalities due to projected climate change is crucial for sustainable food security in the future.

The average water availability in South Asia is low (1137 m^3/person/year), but varies widely between different South Asian countries [4], with the lowest availability in Pakistan (1306) and India (1458), the two major cereal-producing countries of the region (Table 1.3). India has 2.4% of the world's total geographical area, 18% of the world's population, and only 4% of its freshwater

Table 1.3: Water resources in South Asian countries [4, 5]

Country / Parameter	**Afghanistan**	**Bangladesh**	**Bhutan**	**India**	**Nepal**	**Pakistan**	**Sri Lanka**	**South Asia**
Rainfall/snowfall (km3/year)	213	396	84	3560	225	393	112	4980
Internal renewable water resources (km3/year)	47	105	78	1446	198	55	53	1982
Surface water (km3/year)	10	1122	0	635	12	265	0	-
Groundwater (km3/year)	10	25	7	363	19	47	7	-
Total renewable water resources (km3/year)	65	1227	78	1911	210	247	53	-
Per capita availability of total renewable water (m3/person/year)	2008	7622	100645	1458	7372	1306	2549	1137
Agricultural water uses as % of total water use	98	88	94	90	98	94	87	91

resources. Annual water availability in India was about 3000 cubic meters (m^3) per capita in 1951, which has declined to the current level of 1458 m^3 per capita due to an increase in population and increased water use in different sectors of the economy. As per the Falkenmark Index, a commonly used indicator of water scarcity, a country with renewable water availability below 1700 m^3 per capita per annum is categorized as water-stressed, although this index cannot be directly applied to the South Asian region due to variations in lifestyle and water usage as compared to developed nations. The declining per capita availability calls for more restraint and higher efficiency as declining per capita total renewable water availability makes India a water-stressed region in the coming near future.

The major issues related to the variability in available water resources in the region are:

1. Large temporal variability, leading to disasters such as floods and droughts.
2. A high regional mismatch between availability and rapidly increasing demand for water for various uses, while availability remains nearly the same.
3. Increasing and unsustainable use of both surface and groundwater resources to meet the growing demand.

High temporal variability in India is primarily due to the monsoon climate, where about 80% of the annual rainfall takes place in a limited span of four months, i.e., from June to October. Consequently, during this period, rivers carry 70-75% of their annual flow, which at times is beyond their safe carrying capacity. During the remaining eight-month period, river flows account for the residual 25-30%, and many rivers run dry during summer months. Groundwater levels also follow a somewhat similar pattern of rise and fall, but with some time-lag. Considerable variability in water availability gives rise to a host of problems, including floods and droughts. Besides, there are significant spatial variations in water availability that lead to scarcity in some regions and surpluses in others, usually occurring at the same time. Agricultural water use as a percentage of the total water use in the country is 90% (Table 1.3), and hence needs close attention and greater efforts in developing and deploying precision water management technologies and practices that may reduce the volume of water used by agriculture by improving water use efficiency.

1.1 Water in agriculture

Water is an essential input in any agricultural production system and plays a role not only as a critical constituent of plant food synthesis and translocation but also in reducing stress and vulnerability to climatic risks. Irrigated agriculture accounts for the major share of food grain production in the country as it accounts for 55% of the total food production, and about 74% of the production of major cereals (wheat and rice); is thus the fulcrum of country's food security. Since the area under cultivation has already reached a saturation point, and there are very limited further opportunities for expansion of the cultivated area, the growing food demand has to be met through productivity gains from the same land. A large proportion of agricultural land is under rainfed farming, which faces the growing complexities of climate change, and substantial productivity gains in these areas cannot be expected. Therefore, irrigated farming (both surface and groundwater) has to play a significant role in future production gains in the region through increased cropping intensity as well as yield enhancement. An increase in cropping intensity in irrigated areas provides an opportunity to improve the production of agricultural commodities but requires an increased amount of irrigation water as well.

1.1.1 Current status

Cereal-based systems in general and rice-wheat (13.5 m ha) and rice-rice (10.0 m ha) systems, in particular, form the basis for food security in South Asia [8]. Rice, a water-guzzling cereal crop, is grown on about 42 m ha in different rice-based cropping systems in India and consumes a significant share of irrigation water. *Total irrigation water required to produce 1kg rice is around 2500mm,* *however consumptive water use (evapotranspiration) of rice is around 700-800mm and rest is water loss from the system in form deep drainage.Government's irrigation support policies has resulted in the irrational use of irrigation water resources and waste of irrigation water resources which have created a large gap of about 22 m ha between gross irrigation potential created (118 m ha) and utilized (96 m ha) in India.* This gap can potentially be bridged to a large extent by improved irrigation efficiency and water productivity using innovative cropping systems and precision water-management practices. At the current level of utilized irrigation potential of 96 m ha, India utilizes approximately 650 billion cubic meters (BCM) of irrigation water, giving a delta of about 68 cm per hectare of gross irrigated area (GIA).There is a huge difference in utilization of irrigation water resources as we move from north-west to eastern India. In northwestern Indian states like Punjab and Haryana irrigation water resources (mainly groundwater) are over-exploited and resulted in groundwater table decline. However, in eastern India states like Bihar, Jharkhand, West Bengal irrigation water resources are not utilized fully and there is big potential to increase the irrigation water use. Considering that 70% of the average annual rainfall of 1,170 mm is sufficient for meeting the demand of crops for water, gross irrigation water use works out as 148 cm per hectare of GIA. Based on this data gross irrigation water use works out as 1480mm per hectare of GIA, however, we need only 820mm (which is 70% of the average annual rainfall of 1170mm) to meet the demand of consumptive use of water of crops. This is quite high in comparison to irrigation water use in developed countries (about 90 cm per hectare in the USA), and hence there is a high potential for bringing more area under irrigation through water-efficient technologies and production systems.

The development of groundwater has made a significant contribution to an increase in irrigated area, as about 63% of it depends on groundwater for irrigation. Highly subsidized or free power for pumping in many states is being provided to incentivize the use of groundwater for irrigation. However, it has also encouraged users to overexploit deeper aquifers, resulting in depletion of the water table and deterioration of water quality in many areas. The latest assessment carried out in India by the Central Groundwater Board (CGWB) reveals that out of 6,584 Development Blocks in the country, 1034 are over-exploited, 253 are critical, and 681 are semi-critical, while 96 are saline. Waterlogging is another challenge that has emerged both due to poor drainage of irrigated soils and over-use of irrigation water in canal commands. Moreover, the projected

Table 1.4: Sector-wise current and projected water demand (BCM) in India

Committee Name	**Standing Sub-Committee of Ministry of Water Resources**			**National Commission on Integrated Water Resources Development**		
Year/ Sector of economy	2010	2025	2050	2010	2025	2050
Agriculture	688	910	1072	557	611	807
Domestic	56	73	102	43	62	111
Industry	12	23	63	37	67	81
Others	57	87	210	73	103	181
Total	813	1093	1447	710	843	1180

impacts of climate change (rise in temperature and variability in rainfall due to a reduction in the number of rainy days and increasing intensity of rain) have significant negative consequences on the availability of water for agriculture [9, 10]. South Asia in general and India, in particular, are more vulnerable to climate change due to significantly high pressure on natural resources and their overexploitation because of a high density of population, the majority of which is dependent on agriculture.

1.1.2 Future trends

Future water use projections for the developing and developed regions of the world indicate that these will have increased irrigation water (IW) requirements of 50% and 16%, respectively, between 2000 and 2080 [11]. At the same time, with the growth of other sectors of the economy, water demand has increased, creating more competition for agriculture and a wider gap between demand for and supply of irrigation water in agriculture. Future demand for water by different sectors of the economy, such as urban centers, industries, and infrastructure in 2030 and 2050 scenarios, has been estimated to be significantly high (Table 1.4). It is estimated that by 2050, total water demand in India by all sub-sectors (1180 BCM) will surpass the total utilizable water resources of the country [12]. Agriculture's share will be reduced to 68% (Table 1.4); hence cereal systems that produce the major share of food, will have to be sustained with a much lesser quantity of water, and that too of poor quality water. Therefore, the productivity of cereal systems needs to be improved through the precision management of irrigation water while protecting groundwater resources.

1.2 Methods of Irrigation

By 2025, more than one-third of the global population will face a severe shortage of water, with availability becoming increasingly scarce [13], semi-arid regions of Asia, the Middle-East, and sub-Saharan Africa (where a substantial part of the world's population lives below the poverty line) being the worst affected areas. The demand for food and other agricultural commodities in these areas with fast-growing populations is increasing, while the availability of freshwater is continuously declining. As agriculture is a major consumer of water for irrigation in South Asia, there is an increasing need to re-allocate the water used for irrigation to domestic, industrial, and other purposes, but this may seriously impair food security in the region. Sustainable management of water in agriculture aims to match water availability and demand in quantity and quality, in time and space, at affordable cost, and with acceptable or no environmental implications. The sustainability and economic health of irrigated agriculture depend on the ability of producers to adapt to growing constraints on water, particularly by improving the efficiency of water use. Efficient water use not only reduces production costs but also increases profitability by reducing the loss of nutrients.

1.2.1 Surface irrigation

Managing irrigation water at field scale can be improved by quantifying the water balance and using advanced techniques for scheduling irrigation. Reduced availability of water for irrigation may result in lower productivity of irrigated crops, especially rice, that consumes about 120-180 cm per hectare of water [14]. However, over the years, on account of advancements in research, several tools and techniques have been developed for various production systems and ecologies that help in saving water and improving water productivity in agriculture.

For cereal-based systems, different methods of irrigation (flood, furrow, micro-irrigation) are being practiced by Farmers (Figure 1.1), the most common being flood irrigation [14]. Flood irrigation is the traditional system practiced in the country, providing protective irrigation to crops. In this method, fields can be irrigated in two ways: by using earthen (unlined) or brick-lined channels or by using underground cement, concrete or high-density polyethylene (HDPE) pipes. The conveyance losses are higher when using unlined channels for flood irrigation as compared to using cemented or HDPE underground pipes. The efficiency of flood irrigation largely depends upon stream size, field size, and soil type, as the most significant loss of water is due to leaching beyond the root zone. Flood irrigation in unlevel fields causes temporary waterlogging in low-lying areas and adversely affects the productivity of upland field crops (wheat, maize, and pulses). Waterlogging is more severe and prolonged in soils used for growing rice, as puddling leads to the development of a hardpan (having low permeability) at shallow depths in such soils. This hardpan can adversely impair the productivity of upland crops (e.g., wheat) that follow rice [15].

Furrow irrigation is generally used for raising vegetables and potatoes, but furrow-irrigated raised beds (FIRB) are also being used to a limited extent for growing wheat, maize, and other crops. Saving in the requirement for IW has also been reported for rice and many other crops, sown or planted on raised beds [16]. Several studies conducted across South Asia on the FIRB system, show a significant reduction (20-30%) in the use of irrigation water for different cereal crops and cropping systems [16, 17].

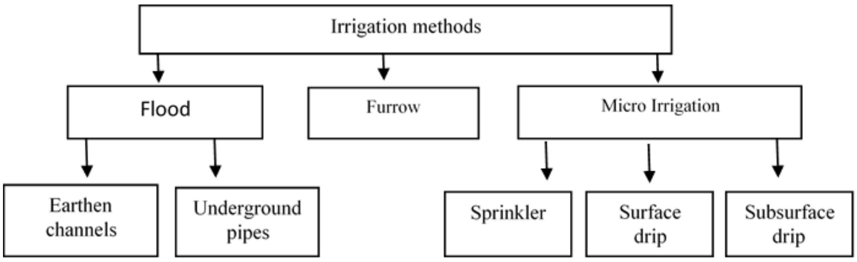


Figure 1.1: Irrigation methods for cereal-based systems

1.2.2 Pressurised irrigation

A pressured irrigation system is a network of pipes and fittings designed to supply irrigation water to the field under pressure. Micro-irrigation is a pressurized but low-volume irrigation system for irrigable areas. Precision water-management technologies (like a sprinkler, surface drip, and subsurface drip) help cultivators not only in saving the irrigation water but also in reducing the cost of production through efficient use of fertilizers and reduced cost of labor and other agricultural inputs. Micro-irrigation systems provide a 35-65% saving of water as compared to traditional flood irrigation systems, along with an appreciable increase in the productivity of crops and the income of farmers [14]. The advantages of micro-irrigation in terms of a significant increase in irrigation water use efficiency have already been demonstrated [18]. Despite multiple benefits and government support, the adoption of micro-irrigation in South Asian countries is very low compared to that in developed countries like Russia, the USA, Brazil, Israel, Spain, China (Figure 1.2). In South Asia, the largest acreage under micro-irrigation is in India, where it has seen steady growth over the years. Of the total acreage under micro-irrigation, sprinkler systems account for 56% and drip irrigation systems for 44%; drip irrigation is increasing at a faster rate, having a CAGR of 9.85% during 2012-2015, while sprinkler irrigation has grown at a pace of 6.60% in the same time. The current acreage under micro-irrigation in India is 7.73 m ha, whereas there is a potential of bringing about 42 m ha under it [19].

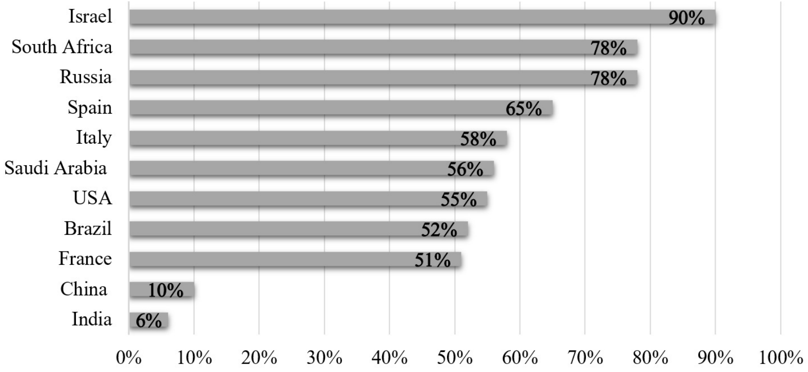


Figure 1.2: Penetration rate of micro-irrigation in major countries [12]

1.2.2.1 Sprinkler irrigation

Sprinkler irrigation technology is a pressurized micro-irrigation system, where irrigation water is generally applied over the crop canopy in an identical way to rainfall. It was introduced in the world about a century ago, but its increased use for larger areas started during the past few decades. In this system, water is pumped through a system of aluminum or HDPE pipes for distribution. Sprinkler systems eliminate water loss through leakage and seepage in conveyance channels. Water is more uniformly distributed even in an uneven field, thereby reducing wastage. Compared to conventional flood irrigation, by using the sprinkler system for cereals and other field crops, a significant saving in irrigation water (30-70%) for various crops grown in different soil types and agro-ecologies has been reported across the globe [20].

1.2.2.2 Surface drip irrigation

Surface drip irrigation is a micro-irrigation system that applies IW to plants by trickling water drop by drop at the soil surface near the plant roots. This irrigation system distributes water as well as nutrients through a network of pipes, valves, and emitters. Apart from highly efficient and uniform IW application and prevention of soil loss due to erosion, the fertigation also reduces nutrient loss due to leaching. However, the system requires a high initial capital investment and a high cost of labor for moving laterals during the cropping season. A large number of studies in different crops and especially cereal systems across the world [21, 22, 23, 24] have reported that drip irrigation systems help make a significant saving (up to 50%) of water in cereals (wheat, rice, corn) and other crops in comparison to flood irrigation.

1.2.2.3 Sub-surface drip irrigation

In contrast to the surface drip irrigation system, the subsurface drip irrigation (SSDI) system reduces the labor cost for anchoring laterals during the cropping season and ensures a longer life of laterals [14]. SSDI minimizes water loss from the soil surface due to evaporation [25] and allows direct application of nutrients along with water to the crop root zone, which ensures more efficient use of fertilizer, reduces the growth of weeds and cost of labor, and enables the performance of cultural operations with ease [26, 27, 28, 29, 30]. Conservation agriculture (CA) systems integrate three major interventions (viz. zero tillage, crop rotation, and soil surface cover, i.e., mulch), which can potentially improve irrigation efficiency by simultaneously reducing evaporation and runoff and improving soil properties that help in better water retention and storage in the soil profile. Some researchers [31, 24] have reported irrigation water savings of 48-53% and 42-53% in rice and wheat respectively under a layering of SSDI and CA compared to the traditional flood irrigated rice-wheat system.

1.3 Automation of Micro-irrigation System

Automation of the micro-irrigation system means the operation of an irrigation system with minimal or no human interventions. Automation of irrigation is desirable when the farm size is large and needs to be divided into small sections for efficient use of water, given the available stream size. Automation of micro-irrigation is gaining momentum in India due to the increasing scarcity

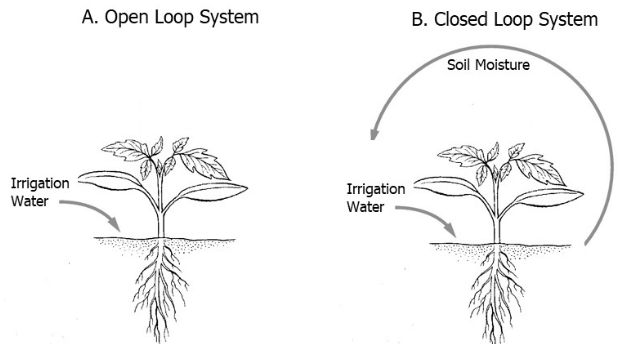


Figure 1.3: Open-Loop and Closed-Loop Irrigation

of water and the elimination of manual operation of control valves, thereby reducing labor requirements. It offers an opportunity to modify the frequency of irrigation and fertigation of crops to optimize water and fertilizer use efficiency. The operating time of irrigation systems can be increased, and the exact amount of irrigation can be applied by better control of pump operation, and thereby optimize energy use.

1.3.1 Types of Automation

It can be a *time-based system* wherein the basis of irrigation is time duration. The operating time of the system is worked out based on water stream size and volume of irrigation water required. In *volume-based system*, the pre-set amount of water is applied to different sections of the field segments with the use of automatic volume metering valves. It may be an *open-loop system* (Figure 1.3 (A)), where an operator decides the amount of water and time of application of an irrigation event. In a *closed loop system* (Figure1.3 (B), feedback is collected from a different type of sensors and decision about how much irrigation water to apply and when to apply is based on a general control strategy developed by the operator.

1.3.2 Computer Based Irrigation Control

These days the computer-based irrigation control systems are used to manage micro-irrigation systems, and these consist of a combination of hardware and software for managing irrigation, fertigation, and other related practices such as maintenance. These systems can be interactive that collect and process information from various points in the system, and allow manual control of the system from a central point by remote operation of valves or other control devices. Or, these can be fully automatic irrigation systems that automatically actuate the valves, pumps based on feedback received from sensors, and control the performance of the system. In such systems, the manual intervention is eliminated and replaced by a specially programmed computer to react appropriately to any changes in the parameters monitored by sensors. The automatic functions are activated by feedback from field units and corrections in the flow parameters by control of devices in the irrigation system until the desired performance level is attained. Automatic systems can also perform auxiliary functions such as stopping irrigation in case of rain, injecting acid to control pH, sounding alarms. Most control systems include protection in emergencies such as loss of the handled liquid due to pipe burst. They close the main valve of the whole system or branching when an unusually high flow rate or an unusual pressure drop is reported by the sensors.

1.4 Machine Learning Approaches

Agriculture is a vital component of the global economy, and Digital Agriculture (DA) involves the use of agro-technology in conjunction with precision farming. DA or Smart Farming is emerging as a new concept using ML-powered high-precision algorithms. These cutting-edge developments contribute to sustainable productivity growth in agriculture. ML is increasingly used to carry out different agricultural operations with better efficiency, which is getting better over time. Smart tractors equipped with sensors, GPS systems, and radars are used for land preparation, precision seeding and crop planting, and harvesting without human driving intervention.

Year after year, ML applies in increasingly scientific fields including, for example, meteorology, climatology, aquaculture, agriculture, economics, food safety, medicine, bioinformatics, ,and robotics. ML tasks are typically classified into different categories depending on the learning type, learning models, or the models employed to implement the given task. The introduction of learning models in ML is limited to those implemented in this work.

1.4.1 Regression

Regression is a supervised model of learning which aims to provide the prediction of an output variable according to the known input variables. Mostly used algorithms consist of linear regression, logistic regression, and step by step regression. More complex regression algorithms viz. ordinary least square regression, multiple linear regression, multi-variate adaptive regression splines, and local scatter-plot smoothing have also been developed.

1.4.2 Clustering

Clustering is an unsupervised learning model application, typically used for finding natural data groupings (clusters). Well-established clustering techniques are the hierarchical technique, the k-means technique, and the expectation maximisation technique.

1.4.3 Bayesian Models

Bayesian models (BM) are probabilistic graphic models that are used to analyse within the framework of Bayesian inference. This is a supervised learning model, used to solve classification or regression problems. Some of the most prominent Bayesian algorithms are Naive Bayes, Gaussian naive Bayes, Bayesian network, Multinomial naive Bayes, and Bayesian belief network.

1.4.4 Instance-based Models

Instance-based models (IBM) are memory-based models that compare new examples with training database instances. These generate hypotheses directly from the available data, and by utilising specific instances, create regression or classification predictions. But complexity of these models increases with increase in data. Most common learning models of this type are K-nearest neighbour, learning vector quantization, and locally weighted learning.

1.4.5 Decision Trees

Decision trees (DT) are regression or classification models with a tree-like architecture. With DT, the dataset is organized progressively into smaller homogeneous subsets, while a connected tree graph is generated at the same time. The leaf nodes reflect the final prediction made after following the root-to-leaf path (expressed as a classification rule). The common learning algorithms of DT are regression and classification trees, the iterative dichotomiser, and the chi-square automatic interaction detector.

1.4.6 Artificial Neural Networks (ANNs)

There are two categories of ANNs i.e., ”Traditional” and ”Deep” ANNs. These are inspired by the versatility of the human brain, emulating complex functions like pattern creation, learning, reasoning, and decision making. Millions of neurons constitute the human brain and these inter-communicate and process any given information. Likewise, an ANN, as a simplified model of the biological neural network structure, consists of interconnected processing units arranged in a specific topology. Numerous nodes are organized in multiple layers, viz. an input layer for feeding the data into the system; one or more hidden layers for learning; and an output layer to convey the prediction.

ANNs are models for supervised learning usually used for the classification and regression problems. The commonly used learning algorithms in ANNs include the perceptron algorithms, radial basis function networks, back-propagation and resilient back-propagation. Numerous learning algorithms based on ANN, viz. Counter-propagation algorithms, auto-encoders, XY-Fusion, and supervised Korhonen and Hopfield networks, adaptive-neuro fuzzy inference systems, self-organizing maps, multilayer perceptron, generalised neural regression networks, extreme learning machines, ensemble neural networks, and self-adaptive evolutionary extreme learning machines. Deep ANNs are generally known as Deep Neural Networks (DNNs) or Deep Learning (DL). These are a fairly new area of research that allow computational models comprising of several processing layers for learning complex representations of data using many levels of abstraction. The DL model can itself perform the feature extraction step in some cases and thus can enhance the state of the art in many areas including agriculture. An ANN having several hidden layers between the layers of input and output is a DNN and can be a supervised, partially supervised, or even uncontrolled model. A typical DL model is the coevolutionary neural network (CNN), in which convolutions are performed image domain to extract feature maps. Some common DL architectures are the deep Boltzmann machine, auto-encoders, and deep-belief network.

1.4.7 Support Vector Machines

Statistical learning theory laid the foundation of support vector machines (SVMs). Intrinsically, SVM is a binary classifier that classifies the data instances by constructing a separating linear hyperplane. The classification capabilities of conventional SVMs can be significantly improved by using the ”kernel trick” for turning the original feature space into a space of higher dimension. SVMs can be used for regression, classification, and clustering. Based on global optimization, SVMs can tackle overfitting problems that arise in high-dimensional spaces, making those engaging in different applications. The most commonly used SVM algorithms include, least squares SVM, support vector regression, and successive projection algorithm SVM.

1.4.8 Ensemble Learning

Ensemble learning (EL) models are designed to enhance the predictive performance of a specified statistical learning technique by constructing a linear combination of a simpler base learner. Assuming that a single hypothesis is represented by each trained ensemble, these multiple-classifier systems allow the hybridization of hypotheses that are not caused by the same base learner, resulting in better results in the case of substantial variability between the single models. Decision trees have usually been used as the primary learner in EL models, such as random forest. A large number of boosting and bagging techniques viz. adaboost, and bootstrap aggregating or bagging algorithms, etc. have also been suggested

By applying ML to sensor data, farming systems will progressively develop into real AI systems. This will provide deeper insight in management issues and richer recommendations for the subsequent decisions and actions, and thus improving the farm production. The use of ML models is, thus, expected to be much more common in the future, thereby allowing for the possibility of development of integrated and usable techniques. At the moment in agriculture, these techniques take into account individual strategies and solutions and are not properly correlated to the decision-making process, as seen in other applications. This incorporation of automated data collection, data processing, ML implementation, and decision-making or support would provide realistic tools that are compatible with so-called knowledge-based farming to increase the productivity and the quality of bio-products. Smart farming thus has real potential to deliver a more productive and sustainable agricultural production, based on a more accurate and resource-efficient approach. New smart farms will eventually realize humanity's everlasting dream and feed our rising population, which could exceed 9.6 billion by 2050.

1.4.9 ML for irrigation management

Intelligent systems are being developed to monitor crop conditions in the field with hyperspectral images and laser scanning. These can significantly improve diagnostic accuracy, contribute to a comprehensive multi-dimensional analysis of crop health, and improve its harvest potential. Precision farming systems, together with ML algorithms, use real-time data to make location-specific decisions for the application of pesticides, fertilizers or irrigation, rather than their application in the field as a whole, thereby reducing waste and costs while avoiding larger environmental footprints. Water is a crucial input in an agricultural production system and plays the role of a moderator in reducing the vulnerability to the climatic change. Santiago et.al stressed the importance of prior planning for development initiatives aimed at improving food security for small-scale farmers while maintaining social equity and offering a rigorous framework for ex ante assessment of the impact of agricultural interventions [32]. Irrigated agriculture accounts for about 55% of food-grain production in the country, and irrigated areas are the fulcrum of its food security, contributing about 74%of the production of major cereals, i.e., wheat and paddy of the country. In India, the net sown area is stagnating at around 140 million hectares (Mha), and the prospectus of its horizontal expansion is limited. Development of irrigation in this country has contributed to an increase of the gross irrigated area (GIA), i.e., from 22.56 Mha to 96.46 Mha from 1950-51 to 2015-16 [33]. Improvements in productivity of crops can be a strategic way forward to increase production. Moreover, another opportunity to improve the production of agricultural commodities is provided by the increase in cropping intensity in irrigated areas as there is a vast scope to improve water use efficiency therein. The protected cultivation systems, i.e., the use of greenhouse technology in crop production, also use Artificial Neural Networks (ANNs) and Fuzzy Logic Controllers to control the environment, i.e., the temperature and humidity of these structures. Irrigation systems that were inefficient and difficult-to-handle in open fields are already being replaced by automated systems capable of predicting and applying water as per crop needs. It not only improves the water use efficiency but also improves the productivity of crops as well as their quality.

1.5 Thesis Organization

The thesis is organized into seven chapters. A brief outline of these chapters is given below:

Chapter 1: This chapter describes the background of the emerging scarcity of freshwater resources and the need for their efficient use in agriculture. Agriculture, being the largest consumer of freshwater, has to produce more food from the decreasing quantity of available water and that too of marginal quality. Thus, there is a need to efficiently forecast the crop water demand to ensure the judicious use of available water and to improve water productivity. It briefly describes the methods of irrigation and the increasing use of micro-irrigation systems to conserve irrigation water. The automation of drip irrigation systems further economizes on water use, reducing human intervention. It briefly introduces the use of machine learning crop water demand forecasting.

Chapter 2: This chapter presents the literature survey which is related to our work. Various existing models for automation of irrigation and the application of machine learning in this automation process have been discussed to identify the research gaps.

Chapter 3: This chapter defines the problem being tackled in this research endeavor after the identification of significant research gaps based on the literature survey given in Chapter 2. Based on these gaps, the objectives of the present research have been defined. It also briefly describes the experimental set-up where rice is grown under a drip-irrigated system. Wether, soil, and crop-related parameters from this location of the experimental site have been collected for training and validation of models in this research work.

Chapter 4: This chapter describes the application of basic machine learning models for predicting the crop water demand as a regression problem. The data pre=processing, model selection, hyper-parameter tuning, and model validation stages have been described in detail. The results received after implementing these models have been discussed and analyzed. The best performing model and dataset have been identified.

Chapter 5: This chapter describes the prediction of crop water demand as a classification and regression problem. Classification predicts the presence/absence of an irrigation event, and regression quantifies the amount of water required. Various basic classification machine learning models have been applied, and the most suitable has been chosen to predict the first stage of a multi-level model. The use of an advance deep learning model has been described for predicting the water quantity in the second stage of the model. The performance of this proposed approach has been benchmarked against the industry standard (APSIM), and the best performing dataset has been identified.

Chapter 6: This chapter summarizes the key findings of this research and suggests possible directions for research in future.

Chapter 2

Literature Review

The need for judicious use of irrigation water in agriculture, i.e., the application of the right amount of water at the right time, is increasing due to growing competition for water and its increasing scarcity. The part of rainwater, which is utilized by the crop during different stages of its growth, is termed as 'effective rainfall' and is generally not enough to support crop growth. Irrigation supplies the remaining amount of water. So, irrigation water requirement (IWR) is the net depth of water (in millimeters) required to be applied to a crop to fully satisfy its specific water needs to achieve its full production potential in a particular growing environment.

2.1 Irrigation scheduling and approaches

Irrigation scheduling (IS) is the process of deciding when to irrigate, and for how long. Drip Irrigation (DI) influences the amount of water and nutrients applied to the crop and ensures that there is no wastage due to deep percolation or runoff. It combines the agronomic expertise with weather data to balance the effect of realities of the stage of growth of the crop, weather, and cultural operations.

Scheduling irrigation for any irrigation method determines the productivity and efficiency of any irrigation system. Irrigation scheduling is not necessarily based on the full crop water requirement but is designed to ensure the optimal use of allocated water. In irrigated areas, water productivity (WP) can be improved with better systems for water conveyance, allocation, and distribution. Irrigation water losses can be drastically reduced by using advanced irrigation methods, including drip irrigation systems that allow water to be delivered precisely when and where it is needed [10]. Saving irrigation water does not necessarily help in saving crop evapotranspiration (ET). Therefore, optimal use of irrigation water in time and space involves scheduling its application to conserve water and optimize crop yields based on scientific principles.

Table 2.1: Approaches for on-farm irrigation scheduling

Irrigation scheduling approaches	**Parameters**
Plant observation based	Visible changes in plant appearance, such as leaf curling and wilting. Plant moisture status can also be measured using sap flow sensors, infrared guns, and pressure bombs, which measure leaf water potential.
Weather data based	Crop water loss is estimated on the basis of either evaporation from an open-pan (PE) water surface or historical climate data, such as length of the day (sunshine hours), maximum and minimum temperatures during the day, wind speed, and relative humidity. Irrigation schedules based on 0.9 to 1.2 IW/cumulative PE have been observed to be optimal for different crops.
Soil moisture monitoring based	Soil moisture used to assess water stress in the plants is measured by gravimetric (volume basis) and soil metric potential method. Use of a tensiometer is a most practical and simple tool for farmers to use for scheduling irrigation.

Irrigation scheduling is a systematic process of estimating the IW requirement that avoids water stress in the crop over relatively short periods. Researchers have developed improved irrigation schedules for flood-irrigated cereal-based systems [34]. Various conventional approaches for scheduling irrigation using different parameters, viz. crop evapotranspiration, plant observation, canopy temperature, soil moisture monitoring are given in Table 2.1 and are described briefly in the following section.

2.1.1 Climate-based

Soil moisture availability and ET-based crop water demand are better scientific practices than fixed schedule irrigation (e.g., weekly supply from canal outlets) because the volume of the IW application is much lower. Climate-based approaches for irrigation scheduling involve the use of a measure of cumulative potential evaporation [35]. Potential evaporation is determined using open-pan evaporation, and reference ET are calculated from meteorological data in a variety of ways, but the modified Penman-Monteith method is generally preferred. Crop ET is calculated from potential ET using crop coefficients. Irrigation is scheduled after a pre-determined amount of ET has occurred, and this threshold amount varies with soil type plant available water capacity (PAWC), crop type (e.g., shallow versus deep-rooted crops) and stage of crop growth. The threshold is determined using information from past studies on crop water use. Optimal irrigation means the application of water precisely equal to the ET of the crop.

The CROPWAT software developed by the FAO Land and Water Development Division is widely used to determine crop water requirements and irrigation scheduling. This model is a simple water balance model that allows the simulation of crop water stress conditions and estimations of yield reductions based on well-established methodologies for the determination of crop evapotranspiration and yield responses to water.

2.1.2 Pan Evaporation-based

Irrigation scheduling based on evaporation involves the application of IW based on moisture deficit in the soil profile to a level such that it does not start to affect crop growth. Irrigation frequency is set to correspond to the time it takes for moisture depletion to drop about 50% of the field capacity. The estimated reduction in soil moisture due to loss through ET is made up by applying an irrigation [36]. Carefully managed deficit-irrigation of crops have high potential to significantly reduce IW demand, as they cover a significant portion of the area under various crops. State-of-the-art delivery systems based on advanced irrigation technologies are needed to implement deficit-irrigation strategies successfully. Prihar et al. [37] proposed a simple concept for scheduling wheat irrigation based on the ratio of the fixed depth of IW to the cumulative pan evaporation (CPE) since the previous irrigation (which equals to open pan evaporation (Pan_E) minus amount of rain). The quantity of IW is worked out depending on the permissible moisture deficit in the soil profile [38]. The deficit-irrigation practice propagates the use of moisture stored in the soil profile by extending the depth of the root zone through an increase in the length of roots along with the receding moisture front. Deficit irrigation has been reported to save two irrigations of wheat out of the six normally applied without any penalty on crop productivity [39].

2.1.3 Soil-based

Irrigation scheduling based on soil moisture includes assessing the soil moisture status (volumetric soil water content or matric potential) within the root zone and knowing the wilting point for the irrigation of a specific crop. Initially, the moisture level is raised to field capacity, and as water is consumed by the crop, water is applied to the crop as per requirement. Water budgeting is a quantitative method wherein current models of estimating ET based on temperature and crop water use is employed. If ET during the season exceeds effective rainfall, it is necessary to irrigate to sustain crops in the field. Once the ET rate of a location is known, the volume of IW determined may be applied at a known rate in a given time through a calibrated irrigation system, and 10% extra water may be applied to account for the losses in the delivery system. The available water in the root zone is the volumetric soil moisture content between permanent wilting point and field capacity of the soil, and the irrigation threshold is generally expressed as percentage depletion of total available water. Alternatively, irrigation can be scheduled based on soil matric potential (SMP), at an explicit depth in the soil profile [40].

Tensiometers and granular matrix sensors are most commonly used for determining SMP, and these can be logged or manually read. A tube tensiometer is easy to use, robust, and relatively inexpensive device. It consists of a plastic tube connected to a porous ceramic cup at one end and a vacuum gauge at the other. Tensiometers provide in situ soil moisture content in real-time and are accurate measures of SMP in the moisture range from 0 to about -80 kPa, which is the moisture range required for major crops. The threshold value of SMP for scheduling irrigation depends on the type of crop to be grown. Based on this criterion, Kukal et al. [41] recommended that puddled transplanted rice (PTR) should be irrigated at an SMP of -16 kPa measured at a soil depth of 20 cm. For scheduling irrigation under different tillage and mulch treatments, Gupta et al. [42] recommended an SMP of -35 kPa at a soil depth of 32.5 cm for wheat, and -15 kPa at a 17.5 cm soil depth for dry seeded rice.

2.1.4 Plant-based

Plant-based irrigation scheduling is based on the physiological and phenological conditions of the crop. The physiological condition (water stress level) can be judged from canopy temperature [43] depression relative to air temperature (measured by infrared thermometry). Then the calculation of the cumulative stress degree days (SDD) [44], and crop water stress index [45, 46] can be used for scheduling irrigation. Phenological stages can also be used to determine when to irrigate. For example, in wheat, critical growth stages for irrigation are crown root initiation (CRI), tillering, jointing, flowering, and the grain-filling stage. Water stress at any of these stages may result in loss of yield depending upon the severity of the stress.

Das et al. [47] reported that the canopy temperature minus the air temperature (Tc-Ta) for wheat under stressed conditions was higher than under unstressed conditions throughout the season. The Tc-Ta could explain the 55% variation in wheat yield in sandy loam soil, and this index could be used as a tool to assess crop water status and hence to schedule irrigation. Gontia and Tiwari [48] concluded that the crop water stress index based on Tc-Ta and vapor pressure deficit could be used for monitoring plant water status and planning irrigation schedules in central India. However, this technique is costly and may not be economically feasible for smallholder farmers. Moreover, all irrigation techniques require a better understanding of the sensitivity of crops to water stress and its ecological and physiological basis, which generally varies through the different stages of crop growth [49].

2.1.5 Deficit-irrigation

Deficit irrigation (DI) is a water-saving irrigation strategy used in many parts of the world [50] in which IW is applied at lower amounts than the full crop water requirement (i.e., ET), thereby increasing water use efficiency (WUE). It has been argued that the level of irrigation supplied under DI should be between 60 and 100% of ET.Evaporation (E) component of ET is non-beneficial loss of water from the soil surface, by reducing E component of ET, also increased water productivity (WP). Water productivity (WP) increases under DI, relative to its value under full irrigation, since small amounts of irrigation increase crop ET more or less linearly up to a point when yield reaches its maximum value, and additional amounts of irrigation do not increase it any further. The location of that point is not easily defined, and thus, when water is not limited, irrigation is applied in excess to avoid the risk of a yield penalty. The amount of IW saving and improved WP under DI strongly depends on the crop (and cultivar), planting date, soil, and site specifications, and is generally accompanied by no or minor yield loss [51]. In particular, many crops have different sensitivities to water stress at various stages of development, and the DI program must be designed to manage the stress, so that yield decline is minimized. Information regarding the crop response to DI is essential to achieve such objectives when water is limited.

In DI, irrigation water is applied during the drought-sensitive growth stages of a crop to maximize the water productivity, by allowing the crops to sustain some degree of water deficit and yield reduction [52]. For DI, it is necessary to identify the critical growth stages of the various crops for their water demand, and irrigation needs to be applied at critical growth stages to realize the maximum WUE. Ali et al. [53] suggested that crown root initiation (CRI) and booting to heading are the two most sensitive growth stages of wheat, and that drought stress should be avoided during these stages. To quantify the irrigation water requirement of DI, it is necessary to define the full ET requirements of the crop. When irrigation is applied at rates below the ET, the crop extracts water from the soil reservoir to compensate for the deficit. This results in two situations. One, if sufficient water is stored in the soil and transpiration is not limited by soil water, even though the volume of IW is reduced, the ET and yield are unaffected; in the other case, if the soil water supply is insufficient to meet the crop demand, growth and transpiration are reduced, and DI induces an ET reduction below its maximum potential [10]. If the stored soil water that was extracted is replenished by seasonal rainfall, the DI practice is sustainable and has the advantage of reducing IW use. In some situations, both water use and ET are reduced by DI, but yields may be negatively affected. In the rice-wheat (RW) system in Punjab, India, around four irrigations are recommended for the wheat crop at different growth stages, and the amount of these generally turns out to be lower than the actual amount of ET of the crop. It is perhaps the water stored in the profile during the previous rice-crop season that is used for supplying the extra water to be used for meeting the ET requirements of the wheat crop.

Research indicates that DI can increase the WP of wheat, maize, and rice by 10-42% [54]. Deficit irrigation (DI) can also save water by reducing the IW depth by watering only the plant root zone and increasing the interval between successive irrigations. Results of field experiments conducted in Dhenkanal, Odisha (India) showed that with two supplemental irrigations, the WP of maize, groundnut, sunflower, wheat, and potato was 0.55, 0.22, 0.23, 0.41, and 2.27 kg per m^3, and the application of three irrigations resulted in WP enhancement by 40, 14, 22, 38 and 7%, respectively [55]. It has been suggested that yields and WP could increase even more if DI were used in combination with water conservation practices (e.g., mulching) or rainwater harvesting techniques [56]. Studies conducted in China showed that ridge furrow planting combined with the DI system significantly increased wheat and maize yields and WP compared with the conventional flood irrigation method, and this system has been widely adopted in many parts of northwest China [57].

Another form of DI, partial root-zone drying (PRD), increases WP while margining the crop yield under limited water resources. PRD is a novel improvement of DI in which each half of the root zone is irrigated alternately in scheduled irrigation events [51]. From an exhaustive review, Sepaskhah and Ahmadi [58] concluded that in comparison to the DI strategy, PRD is a successful alternative irrigation technique compared to full irrigation that can save up to approximately 50% of IW without significant yield loss and thereby improve the WP. It is necessary with this technique to irrigate the two sides of the root system alternately to keep the roots in the dry soil alive and fully functional and sustain the supply of root signals [59].

Alternate or every-other-furrow irrigation is also considered as PRD irrigation. Alternate furrow irrigation has been successfully used as a water-saving irrigation method [60]. Partial root drying has been adopted for different crops by using alternate furrow irrigation, resulting in higher WP [51].

2.2 Scheduling for automated systems

Irrigation systems in the country are still manually operated. However, feeding the growing population under the projected scenario of limited water availability and increased demand due to climate change [61, 62], needs an infusion of modern tools and technologies primarily for making more precise use of water for irrigation [63].

The advent of Wireless Sensor Networks (WSNs) spurred a new direction of research in the agricultural and farming domain. Raghuvanshi et al. [63] have analyzed comprehensively the specific requirements, the devices, sensors, and communication techniques associated with WSNs in improving farming applications. Various case studies that thoroughly explore the proposed solutions according to their design and related implementation parameters have been reviewed, specifically with features like autonomous operation, low maintenance, and low initial investment. The network and node architectures of WSNs, the associated factors, and classification according to different applications have been presented along with the specific issues and challenges associated with deploying WSNs for simplified, low cost, and scalable systems. Finally, they have listed the prospects and problems associated with the existing applications. And the directions for future research with associated factors for improvement in using the new age technologies. Overall, futuristic pre-planning is required for the success of these applications, specifically to overcome the problems from a global perspective. The inclusion of WSNs is envisioned to be useful for advancing the agricultural and farming industries by introducing new dimensions. Semi-automated and automated field irrigation systems are now being introduced to replace conventional uncontrolled surface irrigation systems. An automatic irrigation system can irrigate fields according to available soil moisture, soil type, and the climatic requirement for the crop-water relationship. These devices are ideally suited for applying precise irrigation through the surface, sprinkler, and drip irrigation systems in the prevailing situation of water and labor shortages. Intelligent, sensor-based irrigation systems can help to achieve optimal water use, through significantly improving water use through better scheduling, and thereby significantly improving water use efficiency [63, 64].

2.2.1 Sensor-based system

Automated precision irrigation is a data-centric approach and requires real-time data on soil moisture, temperature, and weather parameters. The sensors are employed for real-time monitoring/sensing these parameters and are useful for planning irrigation. Besides, monitoring soil moisture content, the sensors precisely assess the crop water demand from weather data. Initially, tensiometers and granular matrix sensors were used for estimating soil moisture to automate the control of irrigation water. Electromagnetic-based sensors [65] were also used and showed promise for the real-time estimation of soil moisture. Information from the moisture sensor is provided as input to the micro-controller that transforms the data from analog to digital. The micro-controller transmits the data using a radio-frequency (RF) module. The Bluetooth and Wi-Fi-based wireless transmitters can transmit these data to the base station. The micro-controller displays the soil moisture and flow levels. This facilitates a remote assessment of the field conditions and real-time monitoring of the variable-rate irrigation controller.

An irrigation controller is employed to operate a solenoid valve that supplies water to crops when soil moisture content falls below a pre-defined critical point. When the soil moisture content is low, the power automatically switches on the pump, and water from the tank starts flowing through the pipe; when the soil moisture reaches a pre-defined, soil-specific saturation limit, the power turns off, and the water stops. A typical example of automation is a wireless sensor network (WSN) based irrigation system that comprises different types of sensors, viz. soil moisture sensors that are positioned over the entire farm, and wind speed and direction sensors. Multiple wireless sensor nodes and a gateway consisting of a sensor-based wireless monitoring system can provide a unique and secure solution with improved spatial and sequential resolutions for automated precision irrigation.

In some systems, irrigation scheduling is accomplished by using a tensiometer to measure soil metric potential (SMP) that determines the soil moisture content, for operating irrigation through an automated controller system. The fuzzy logic system, which works on the principle of assigning a particular output depending on the probability of the state of the input, can improve the reliability of the estimated value of the irrigation amount and facilitate decision making. Some researchers are also working on variable-rate sensors for improving the efficiency of irrigation systems.

Jain et al. [66] used a Radio Frequency (RF) module along with a remote monitoring unit for the automation of an irrigation system. Ojha et al. [63] undertook a critical review of the available evidence on the deployment of WSNs in international and national scenarios, and the potential use of these in the advancement of different applications in agriculture. WSNs and ICTS play a key role in enhancing the application of precision technologies in agriculture, viz. monitoring of soil moisture, weather, and crops in real-time [67, 68, 63, 69]. Table 2.2 lists various sensors and methods currently being used in automation o irrigation systems for precision application of irrigation water to crops.

2.2.2 WSN-based system

Communication without using wires reduces the cost. Wireless communication can be put to good use for managing agricultural land and monitoring irrigation operations. WSNs are recognized as a powerful tool for collecting and processing data for low-cost and low-energy precision agriculture [81]. Precision agriculture implies better allocation of resources in space and time, in some place producing more-with-less and others producing more-with-more [82]. WSNs provide high spatial and temporal resolution for monitoring soil and crop parameters via wirelessly-connected sensor nodes installed across the field [83, 84] that automatically transmit data in multi-hop mode. Technological advances in micro-controllers and wireless communication have led to the development of embedded devices (nodes) that are smaller and have a higher capability for

Table 2.2: Sensors and methodology used for precision irrigation

Sensing elements	**Methodology used**	**Ref**
Sensors for humidity, soil moisture and soil temperature, LDR sensor, CO2 sensor	Based on fuzzy logic plus neural network pattern classification technique	[70]
Air temperature sensor, air humidity sensor	Fuzzy logic controller is used for data acquisition. Well devised fuzzy rules produce appropriate time and duration for irrigation.	[71]
LDR sensor, soil moisture sensor, temperature sensor		[72]
Sensors for humidity, soil moisture and temperature	Ontology-based decision support system	[73]
Smart phone	Pixel differentiation process	[74]
Soil moisture sensor	Priestley-Taylor method to find reference evapotranspiration	[68]
Sensors for soil moisture and soil fertility	Threshold-based	[75]
Soil sensor array		[76]
Soil moisture sensor, humidity sensor temperature sensor		[77, 78]
Soil moisture sensor		[79]
Humidity sensor, soil temperature sensor		[80]

sensing, measuring, controlling, transmitting, processing, and storing information. WSNs can thus improve the monitoring of plants and soil and can improve productivity, efficiency, and profitability while minimizing risks due to climatic aberrations, water shortages, pest infestation, and other factors hostile to crop growth and development. WSNs are helping to achieve better response times due to real-time sensing and communication.

There are multiple ways to use wireless sensors for irrigation scheduling. Based on the threshold values of temperature and soil moisture content, the gateway allows automated actuation of the ir rigation system. Some of the sensors, modules, and valves that are commercially available for assembly in an irrigation network are quite complex and expensive to be deployed for managing fixed irrigation systems. The cost and complexity of installation, operation, and maintenance of these systems have been limiting factors in their adoption. Sensor-based automated systems for irrigation are of two types: (1) systems that assess soil moisture content based on the measurement of various weather and other parameters; and (2) systems that measure soil moisture content directly using sensors [85].

A WSN can be used in diverse environments and has advantages over a wired network in terms of size, flexibility, distributed intelligence, and cost. Certain other advantages of wireless transmission include its potential for deployment in areas where cabling is not feasible, as well as its more considerable simplification and significant cost savings over cabling. Soil moisture content and temperature are transmitted wirelessly, thereby permitting the adjustment of the water supply to daily weather fluctuations or variations in the volumetric soil profile moisture [86]. The adoption of wireless technologies eliminates 20-80% of cabling costs in the setting up of industries. Water distribution in a field for agricultural production can be monitored precisely through the use of wireless sensor networks.

2.2.2.1 Components of the system

The components of a WSN application are a sensor node, master node, database, and base station (BS). Wireless sensor networks based on communication technology and microcontrollers can improve existing monitoring methods to suitably support real-time response for a wide array of applications. Recent developments in wireless technology and microelectronics have led to the creation of low-power-consuming and low-cost components. In a sensor network, different communication technologies such as Zigbee, WI-FI, Bluetooth, RF, and General Packet Radio Service (GPRS) have been put to use for efficient communication between the elements of a network – the sensor nodes and data receiver – and the network itself.

Zigbee technology that carries small amounts of data over a short distance is suitable for a small farm, whereas, for automation of multiple sensors in larger areas, Wi-Fi technology can be used. Zigbee technology is the preferred choice among various wireless technologies due to its efficiency in low-power connectivity and its ability to connect a large number of items of equipment into a single network. Zigbee technology utilizes a license-free and globally available 2.4GHz frequency band. It allows wireless applications for personal area wireless networks, using a standardized set of high-level communication protocols sitting atop a cost-effective, low-power digital radio based on the IEEE 802.15.4 standard.

An alternate mode of communication is based on a cellular internet interface (GPRS protocol) using a mobile data service cellular global system. The GPRS permits mobile phones continuous connectivity to the network and transfers the desired data promptly. The GPRS technology is more practical and useful due to its ability to transfer data; it can both receive and send data during a call without disconnecting it. Internet connectivity allows for real-time data inspection on the website. Its low cost permits wider use of the technology in remote control applications; mesh networking offers better reliability, and broad range and low power use permit batteries to be smaller with longer life. The software, which works on a WSN and nodes, monitors the entire farm from a distant location using IoT. The network has two nodes: a camera node to monitor crops and collect images; and a sensor node that captures all the soil and environmental data, viz. temperature, hours of sunshine, soil moisture, air humidity.

2.2.2.2 Process of automation

Automation of the irrigation system is accomplished by linking two wireless sensor units (WSUs) through radio transceivers, which enable the transfer of data related to temperature, soil moisture, and sunshine using XBee software and a GPRS module to convey the information via the public mobile network to a web server. Tracking the information online is enabled through devices with internet access or a smartphone Wi-Fi network. Pathan and Hate [87] used a different set of sensors, viz. temperature sensor LM35, LDR, and soil moisture sensor. Avati and Patil [88] and Hussain et al. [89] have demonstrated a Zigbee-based wireless protocol for drip irrigation. A WSN application based on radio frequency identification (RFID) and quick response (QR) codes were proposed by Nam et al. [69] for the management of irrigation systems. An application of Wi-Fi modules and circuitry design of sensors for automation of irrigation is depicted in Figure 2.1.

Gutiérrez et al. [76] developed and tested a microcontroller and wireless communication-based WSN system for an experiment on irrigation automation of malt barley. The system comprised two components: WSUs in the field; and a wireless information unit (WIU), connected by radio transceivers that enabled the transmission of data on soil moisture content and temperature, using Zigbee technology to implement a WSN. It was operated by photovoltaics and comprised distributed networks of temperature- and soil-moisture sensors installed in the root zone of the crop. Every sensor node included a temperature probe, a soil moisture probe, a radio transceiver, a microcontroller for data acquisition, and a microcontroller-based receiver to which the sensed data were transmitted. This

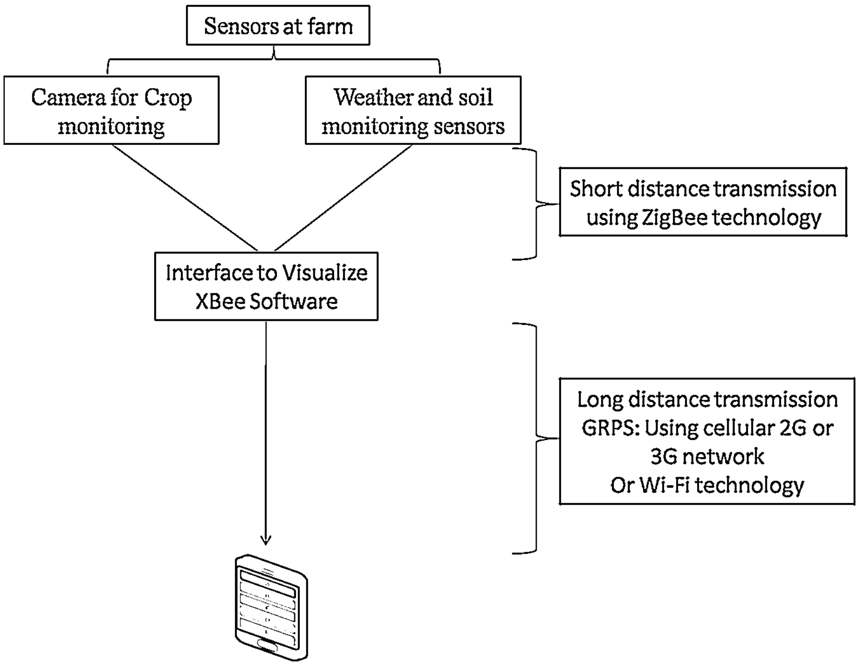


Figure 2.1: An application of Wi-Fi modules and circuitry design of sensors for irrigation automation

gateway allowed the automatic actuation of irrigation as soon as temperature and soil moisture values reached the threshold values. The gateway device also handled information from sensors, triggered actuators, and communicated the information to the web application. Threshold values of temperature and soil moisture were used to develop an algorithm that was programmed to control the quantity of water through a gateway based on microcontrollers. A cellular-Internet interface-based duplex communication connection enabled the system to inspect the data and accomplish irrigation scheduling through web-based programming.

Mafuta et al. [90] further improved this WSN-based, solar-powered, and robust automated irrigation system. The end nodes sense soil temperature and soil moisture content and transmit the information to the coordinator node. Based on this data, the coordinator node decides whether to irrigate the field or not. The coordinator node forwards the information on soil temperature, soil moisture, valve status, battery level, and performance of wireless link through an Xbee unit to the gateway node that transmits it to a remote station via a cellular network. The significant advantages of the system are adequate irrigation and higher energy efficiency due to power-saving modes. However, the primary issues with the system are that it ignores sunshine duration, air humidity, and wind speed when determining reference ET, which reduces the efficiency of irrigation, and it has a lower coverage area because of the Xbee devices [91].

2.2.2.3 IoT-based systems

Anbarasi et al. [92] proposed an IoT application-based system for a smart irrigation system to improve water use and crop yield, wherein the decision is taken based on real-time data from the soil. This automated irrigation system switches the pump ON and OFF based on the moisture content of the soil. A soil moisture sensor measures the exact soil moisture content, and this input operates the pumping motor through an operational amplifier. Farmers can see the soil moisture status in a mobile application or web page using a modem through IoT. They can check whether the water sprinklers are turned ON or OFF at any time. This model includes a sensor hub and a control hub. In a multi-cropping system, information about soil moisture content is required. This is because information about an assortment of yields is stored, and this is confirmed with information detected by sensors. This system reduces the physical contribution of the irrigator and improves efficiency and, thus, productivity. In rural areas with a severe rainfall deficiency, this model can be used effectively to achieve extraordinary outcomes in most of the soils. Sani Abba et al. [93] designed an IoT-based irrigation monitoring and control system, using different sensors and actuators, for facilitating a smart irrigation system to provide an autonomous supply of an adequate amount of water from a reservoir to crops. The main objective was to enable users to monitor and control the remote farm independently to improve crop productivity. The system employed a soil moisture sensor to estimate soil water content, a pump to supply the desired amount of irrigation water to the crop, and a Wi-Fi module to access sensed data via the Internet. The sensor readings that were communicated to the Internet in real-time were stored on a web server that served as the main base station. The flow chart for the process is given in Figure 2.2. Based on the analysis of these readings, a user can precisely monitor the soil moisture content in the field. The system ensures real-time notification about any inadequacy or excess in the supply of irrigation water as water pump is turned off or on depending on the moisture content of the soil. Experimental and simulation results validate the flexibility and applicability of the proposed system and are of the utmost importance for the users because through remote monitoring and control of irrigation supplies. The system reduces the cost of supervision.

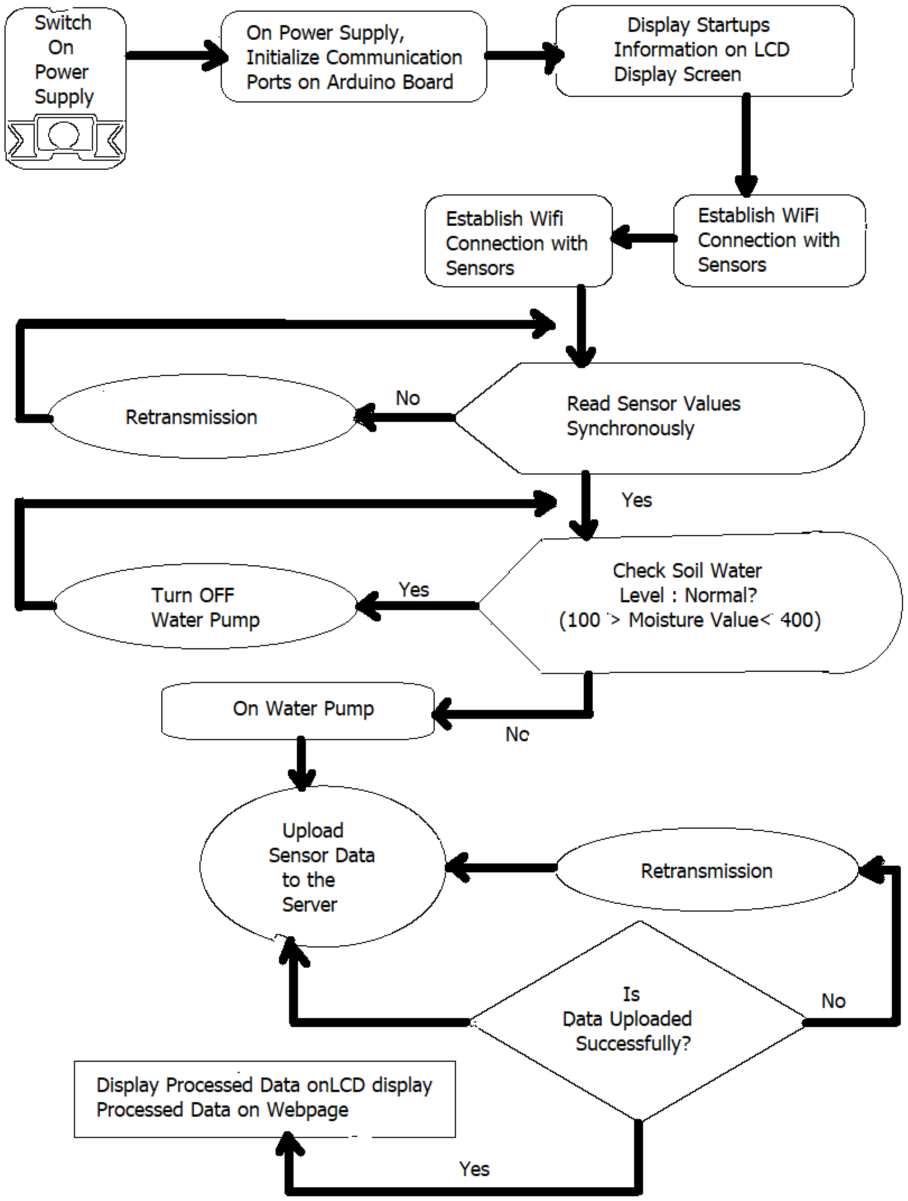


Figure 2.2: Flow chart of IoT-based irrigation automation process

2.2.2.4 GSM-based systems

A global system for mobile communications (GSM) based wireless network has been employed for irrigating crops in an automated irrigation system. Chaware et al. [94] proposed automation of irrigation for shallow-rooted crops like vegetables, where the moisture sensors were mounted near the root zone of the plants to sense and provide the output. These measured values were then fed to the controller that sent these to the user via the GSM modem. Information about time of sensing, status of water availability, and water requirement of the crop would be provided along with the starting of the motor. The system navigates through a controller that can optimally handle the sensors for soil moisture assessment with a low requirement of energy.

Tope and Patel [95] designed a system consisting of two units: a transmitter unit and a receiver unit, to determine the parameters effectively with the use of sensors for temperature, humidity, and soil moisture. A soil moisture probe comprising several moisture sensors measures the soil moisture content. 2.4 GHz ISM band, XBee modules are commonly employed for efficient wireless data transmission without a license. The solar-powered system has a duplex communication link based on a cellular-Internet interface that permits data inspection and programming of the irrigation schedule through a web page. In this irrigation system, the use of solar power is significantly appropriate in situations where the supply of electrical power is expensive. A real-time system based on GSM and a Zigbee module is incredibly versatile and cost-effective. A significant drawback of a GSM technology-based WSN is that the water level in this system is not continuously monitored.

2.3 Types of irrigation automation

2.3.1 Tensiometer and capacitance-based

Tensiometers and capacitance are used for detecting soil moisture for the automation of irrigation water application. Testezlaf et al. [96] designed and evaluated an automated irrigation system for real-time water management in greenhouses. The irrigation system consisted of a pressure transducer-mounted tensiometer to monitor the soil water content in pots. The pressure transducers were chosen due to their high-level analog performance, low cost, low range of operating pressure, and accuracy. The transducers' output voltages were interfaced to a PC via the input port of a digital/analog board responsible for the acquisition of data and the control system. The strategy for automated control utilized a closed-loop unit that supplied irrigation water on demand. This strategy helped the system to sustain constant soil moisture content (soil metric potential setpoint). The control interface paused for a preset time interval termed "sensor response time" to continue to monitor the tension in the container for the next irrigation activity to take place. This was necessary to account for the time taken after water application for change in SMP to be measured by the tensiometer. The pumps were controlled by solid-state relays. The system monitored the water status of plants in the container accurately, applied water according to the plants' demand, and kept a record of applied water. However, Testezlaf and colleagues suggested further testing of the system at the field level to assess its feasibility on a commercial scale. During testing at lower tension values, a significant variance was observed in the data recorded by sensors placed in similar substrate conditions in different containers. This variation in sensed data suggests that in managing an irrigation system relying on sensors that measure SMP, the number of plants being monitored, and whether these form a substantial sample of the total number of plants or the experimental plot in the irrigation station should be considered.

Lakshmiprasad et al. [97] developed an intelligent irrigation system consisting of a server equipped with a microcontroller processor to measure different parameters. It utilized only one parameter at a time, e.g., a soil moisture sensor to assess the moisture content. For instance, if we are using a soil moisture sensor for irrigation control in an automated drip irrigation system, then whenever the soil moisture level falls below the threshold value, and only then does the system direct the valve to change from the OFF to the ON position. When the soil moisture level reaches the optimum preset level, then the system is switched OFF automatically. To measure the soil moisture content, a coded soil moisture sensor (10 HS) was used. This sensor has a low power requirement and a much higher resolution, and uses a capacitance technique to estimate the volumetric moisture content of the soil by determining its dielectric constant; it enables measurement as many times as required over a long period with low battery usage. It has a GSM communication device that transmits the data to a farmer or receives it from a farmer who can take appropriate action. The system is also equipped with electrochemical sensors to assess the spatial variability of various chemical properties of the soil. The capacitance-based sensor consists of two metal plates separated by a non-conductive polymer film. This film gathers air moisture, which causes minor changes in voltage between these plates. These changes are converted into digital values indicating the moisture content of the air. The RS 232-level converter, which performs between -15V and +15 V, is used for serial communication.

Recently, an automatic subsurface drip irrigation system based on WSNs using tensiometer-mounted SMP sensors has been designed and is being evaluated in the rice-wheat system at the research farm of Borlaug Institute for South-Asia (BISA), Ladhowal (Ludhiana), Punjab, India. Results from initial observations reveal that the new automatic system performed well and delivered the desired quantity of irrigation water whenever SMP values fell below the prefixed threshold value [24]. The system is being evaluated on a large farm for its performance and further improvements if required. Stone et al. [85] developed a magnificent, microcomputer-based, low-cost data-acquisition system to assess soil moisture potential continuously. The system comprises a tensiometer-mounted pressure transducer that performs well under field conditions.

2.3.2 Smart phone-based

A recent development in irrigation automation technology includes the use of sensor-based smartphone apps [74]. On activation of the app, the lighting circuit is switched on, and the app takes a photograph of the soil; the lighting circuit is then switched off to save power. The RGB photo of the soil is transformed into a grayscale image wherein white pixels represent a dry area, and dark pixels represent a wet area. Next, the process of pixel segregation is started to determine the number of saturated pixels, which is the difference between the grey and light field matrices A number is computed from the relatively wet soil, and the router node communicates this to the gateway node that controls the time of application of the irrigation water. This system has the advantages of reduced complexity of hardware, wide-ranging network coverage due to the use of Wi-Fi, and more energy efficiency due to being able to tap the mobile to put it into sleep mode. More irrigation sensors have to be embedded to achieve wet soil areas in the smartphone, but that increases the cost of the system.

2.3.3 Solar power-based

Merlin et al. [75] designed and tested an automated irrigation system using WSNs that is solar-powered, reliable, and economical. As soon as wireless sensor nodes determine soil moisture, the measured value is compared with reference values utilizing an on-stream camera. This actuates the valve for irrigation. When irrigation is not required, the microcontroller is in sleep mode, and when irrigation is required, it shifts to active mode. The significant advantages of the system are increased energy efficiency by using the power-saving mode and dynamic irrigation. However, as air temperature, hours of sunshine, and wind speed are not considered for working out reference ET. In this case, the irrigation efficiency of the system is reduced.

2.3.4 Combination of soil and weather sensors-based

A combination of soil moisture sensors and weather sensors has also been advocated for efficient application of irrigation by several researchers. Gokul et al. [98] reviewed existing automated irrigation technologies. At present, mainly wireless LAN, Bluetooth, and Zigbee are used as transmission mediums from end nodes to the sink node, which results in coverage of a lesser area. Some of the current methods only use soil moisture content and humidity values for deciding when to irrigate. However, crop water demand is also dependent on the duration and intensity of sunshine, wind speed, and air temperature, which contributes to a flawed estimation of the irrigation schedule. Further, the use of an algorithm based on threshold values and fuzzy logic affects the precision of the output. Several factors – the algorithm for determining reference ET, the capabilities of the sensor node to communicate and process the information, and likely exceptions during the application of irrigation – need special attention for efficient automated irrigation systems. Therefore, to achieve high precision, the combination of soil moisture sensors and weather-based sensors is highly useful.

John et al. [99] described a WSN for sensing the relative humidity of the air, soil moisture, and temperature. The system comprises clusters of nodes, and the lifetime of the network node is improved through a sleep-wake up plan. A drip irrigation system automated with an effective routing protocol termed Equalized Cluster Head Election Routing Protocol (ECHERP) was evaluated by Nikolidakis et al. [68]. All data sensed by sensor nodes are forwarded to the base station that acts as a gateway to transmit these data to the server. This system employs the Priestley-Taylor method for determining the reference ET. The system also considers the historical data to make the irrigation automation process more useful by taking optimal irrigation decisions. This system has higher energy efficiency than other routing protocols like LEACH, TEEN, and PEGASIS, and decisions based on historical data help in achieving high-impact optimal irrigation. The system has limitations, as humidity and duration of sunshine are not considered for estimating reference ET; also, irrigation efficiency is low, and the overall coverage range of the Zigbee protocol is also low.

Similarly, Anurag et al. [100] developed a WSN for precision irrigation wherein weather, and environmental data are sensed and transmitted to a central repository in real-time. The network consisted of three separate sections: (i) sensor nodes; (ii) wireless mesh network; and (iii) actuation devices based on standards laid out in IEEE-802.15.4; the selection was based on four attributes suitable for commonly grown crops. They developed a new algorithm of static routing suitable for sensing applications. The tree-based routing can be most efficiently used as compared to other Zigbee-supported routing algorithms.

Chikankar et al. [101] used a Zigbee technology-based WSN with a tree-based structure to control soil moisture, air humidity, and temperature. They first prepared a pictorial topology with routers installed along the path at suitable distances. A basic algorithm was used to determine the address of various routers. The algorithm operated on an algorithm of the lowest depth-first, i.e., the first address was assigned to the deepest router, and then it "worked up" the topology. The tracking needs to permit us for the advanced setting of addresses to maintain the tree structure. The actuation was accomplished based on the information provided by the sensors. The system generated automatic warning messages on the console after the threshold value was exceeded so that the necessary action could be initiated. At the controller end, these valve-based actuators were operated by discrete electrical switches. An RFID device was attached to each actuator to identify and control it uniquely.

Another automated drip irrigation system, powered by photo photovoltaically and, consisting of a distributed wireless network of temperature and soil moisture sensors are deployed in the crop root zone. Every sensor node includes a soil moisture sensor, temperature sensor, data acquisition microcontroller, and radio transceiver; the sensed data is communicated to a receiver that is also based on a microcontroller. The automated drip irrigation system developed by Nalini Durga and Ramakrishna [80] has three parts, viz. a humidity-sensing part, a control unit, and an output segment for deciding irrigation. The moisture content of the soil is sensed through a YL-69 soil sensor (a resistance-type sensor). The control unit consists of an arduino microcontroller. The switching controls are of an on/off type that depends on soil moisture content. The soil moisture module consists of two parts, an amplifier circuit, and the probes. The analog output provides information regarding plant moisture in real-time, and this output is used by the system. A water pump connected to the relay module works when this module receives such instruction from the microcontroller. The Arduino board is energized using a 7V to 12V solar panel, a plug-in adaptor, or a wall wart. The smart irrigation system for automatic watering is depicted as a Block diagram in Figure 2.3. Whenever the soil moisture content is low, the sensors detect the moisture change and signal to the microcontroller to start the pump. Several studies carried out in different production systems and geographical areas across the world have demonstrated the tangible benefits of irrigation systems using different sensors (Table 2.3) over traditional water management practices.

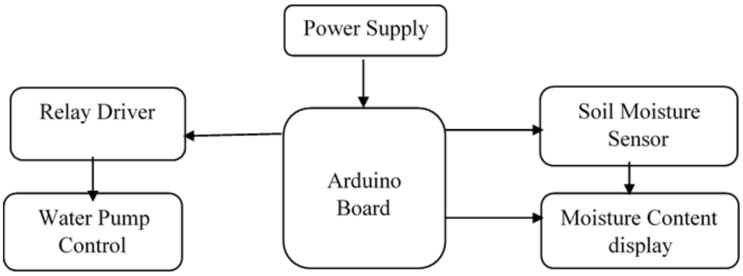


Figure 2.3: Block diagram of smart irrigation system

In another study, Soorya et al. (2013) compared a soil moisture sensor-based automated drip irrigation system with a non-automated drip system for okra (Abelmoschus esculents) and reported a substantial reduction in water use with the automated drip system. Using a microcontroller-based automated drip irrigation system for sweet corn that was irrigated when soil moisture content reached 70% of field capacity, Deekshithulu et al. (2018) reported 36% less water use as compared to the flood irrigation method, with respective water use of 332 and 520 mm.

Table 2.3: WSN architecture and effect on system efficiency

End node/ sensors and their mode of transmission	**Communication**	**Irrigation scheduling/application criteria**	**Irrigation automation procedure**	**System efficiency**	**Ref**
Temperature, soil moisture sensors and wireless camera	Cellular-internet interface	Algorithm based on soil moisture and temperature threshold values	Webpage controlled	Water savings up to 85% and reduced cost	[102]
Temperature, moisture, pressure, vapor pressure sensors with radio transceiver	Wired communication system (field bus)	Crop water- stress index (CWSI)-based	-	-	[66]
Soil moisture sensors	Bluetooth	Neural network based	-	-	[103]
Soil moisture content and temperature sensors	Cellular-internet interface GPRS	Threshold based	-	Aggregation of data improves the residual energy by 10 % and throughput by 13 %. Due to use of energy conservation techniques, this system achieves better energy efficiency.	[77]
Humidity, soil moisture, and soil temperature sensors	GPRS	Threshold based	operate the irrigation control valve	saving up to 25% in irrigation water use	[91]
WSN of temperature and soil-moisture sensors located in the root zone.	GPRS	Threshold values of soil moisture and temperature	Threshold-based gateway actuates	Needs better design of hardware and software	[104]
Soil-moisture and temperature sensors	CDMA	Threshold values of soil moisture content and temperature-based algorithm	Micro-controller-based gateway regulates the amount of water	-	[95]
Humidity, sunshine, wind speed and soil temperature sensors	SMS or GPRS	Threshold values of wind speed, RH, sunshine duration and soil moisture	Wireless interface unit (WIU) operated solenoid valve	Overlooking of weather parameters reduced irrigation efficiency	[80]

2.4 ML-based irrigation forecasting

There are several critical features for irrigation, such as water stress, gas-exchange rate, evapotranspiration rate, stomatal conductance, and closing of stomata. Cultivators can use information such as light, humidity, and temperature levels to modify irrigation schedules and avoid the risk of damaging crops. Machine learning is an upcoming approach for forecasting the irrigation water demand. New thermal imaging, in combination with other image processing and analytic data techniques, attempts to decrease crop water stress and provide irrigation scheduling. Thermal imaging is a noncontact and nonintrusive technique in which a much fewer number of sensors are required than in soil monitoring and measurement. These sensors can work in conjunction with drip irrigation methods and fertigation to avoid unnecessary waste of water and fertilizer.

2.4.1 Image-based

Several studies have demonstrated that information from thermal imaging is an appropriate approach to identifying key parameters to schedule irrigation. Dong et al. [105] collected and processed high-resolution images through a cloud-based smart irrigation system. These images are based on input from the human visual system, and a have density-based measure. The captured thermal image is decomposed into blocks, and the canopy temperature distribution is extracted for the considered blocks. *An optimization algorithm, such as the genetic algorithm or a recursive algorithm, is used to calculate a threshold value for the control.* It has been suggested that temperature distribution be introduced and utilized as a quality measure for automated intelligent CoT-based irrigation systems.

Remote sensing provides a fast and reliable alternative for traditional in situ water status measurement in vineyards. In this study the authors use multispectral sensors on board of a UAV platform for obtaining high-resolution images that were used to monitor the water status in the vineyard [106]. These aerial multispectral images were processed to derive several vegetation indices (VIs). Statistical methods and algorithms for machine learning were used to test the associations between the midday stem water potential (Ψ_{stem}) of grapevines and VIs. Results by simple regression showed no significant relationships between VIs individually and (Ψ_{stem}). A pattern recognition ANN model was developed for irrigation scheduling purposes. It uses (Ψ_{stem}) measured in the study site as inputs and multiple thresholds of (Ψ_{stem}) as outputs. This study confirmed that a UAV platform equipped with multispectral sensors and machine learning modeling strategies could be an efficient and affordable option to assess vine water status and to map vineyard spatial variability to optimize irrigation management. However, additional pixel by pixel data information available per plant may be useful for physiological studies and more in-depth irrigation practices, such as identification of localized irrigation problems (broken emitters or flooded furrows). Nevertheless, problems associated with irradiance and reflectance still need to be investigated for other potential applications.

2.4.2 ET-based

Irrigation demand can be indirectly estimated by forecasting the Evapotranspiration. Usually, reference Evapotranspiration (ETo) is estimated by using the well-known FAO-56 Pennman-Montieth method [CROPWAT 8.0], which requires reliable and good quality climatic data. In a poor data environment, it's application is restricted. The correctness of the ETo forecast remains a challenging computational task since inaccurate weather variables can alter the forecast accuracy. Mohamed Abdrabbo and Gamal El Afandi studied four reference evapotranspiration (ET) methods, namely: Blaney-Criddle, Hargreaves, Thornthwaite, and FAO-56 Penman-Monteith. The results revealed that the Blaney-Criddle and Thornthwaite equations had the lowest ET values comparing with Hargreaves and Penman-Monteith method [107]. Wang et al. [108] used artificial neural networks to determine real-time ETo forecast values for short-term use by using weather information from weather forecast messages for the daily public release. Four ANN learning algorithms including Generalized Feedforward (GFF), Linear Regression (LR), Multilayer Perceptron (MLP) and Probabilistic Neural Network (PNN) are implemented with a combination of three sets of inputs consisting of minimum (Tmin) and maximum (Tmax) daily air temperatures, extraterrestrial radiation (Ra), and net solar radiation (Rs) for ETo in Dallas. The model evaluation used the coefficient of correlation, mean square error, normalized mean square error, mean absolute error, and mean square error capacity score. Results showed that input data set Tmin, Tmax, and Rs yielded the highest performance with the Multilayer Perceptron (MLP) backpropagation network capable of reproducing the closest values to the observed FAO-56 PM. The results of this study can lead to the development of an automated customized interactive website to be interfaced directly with public weather forecast data for deriving short-term ETo forecast values for irrigation purposes in Texas.

Krupkar *et al.* [109] reviewed the then available computational and statistical methods used to determine Evapotranspiration. They discussed various machine learning methods, such as Logistic Regression, Support Vector Machine (SVM), Decision Tree Classifier, Systematic Forests, Artificial Neural Networks (ANNs), Fuzzy Logic techniques and ARIMA, which have been previously used to automate the water irrigation prediction process. These algorithms capture and interpret semantic relationships between the various parameters and hence can be used to make predictions. However, they propose the use of a recurrent neural network (RNN) model that uses the LSTM activation function to model the irrigation requirement. A typical LSTM circuit consists of an input gate (i), the forget gate (f), the output gate (o), the cell gate (c), and the hidden layer outputs (h). The RNN is found to be robust enough to map the intrinsic variations in observations that occur due to various factors like global warming, faulty equipment. It is recommended to explore the use of deep learning models to predict Evapotranspiration, as trained models require minimal space (as opposed to conventional memory AI models) and are computationally fast in testing.

In their paper [73], the authors implemented an autonomous closed-loop zone-specific irrigation by describing the design and deployment of an adaptable decision support system integrated with a wireless sensor/actuator network (WSAN) for zone-specific irrigation control in a greenhouse setting. The primary plant signals and environmental parameters explored are plant's leaf Temperature, Chlorophyll Fluorescence, Ambient Temperature, Humidity, and Soil Moisture. The learning goals drought stress and heat stress were categorized with values TRUE/FALSE. Machine learning algorithms such as probabilistic logic, fuzzy logic, and Bayesian networks have been used for inducing new rules and determining thresholds of plant-based parameters by analyzing logged datasets. Also, the machine learning process was applied to extend the ontology of the system to cope, for example, with a sensor type failure or to improve the accuracy of a plant state diagnosis. The effectiveness of the developed system was validated by comparing its agronomic performance to traditional agricultural practices. It is suggested that the integration of a chlorophyll content meter sensor would allow a system upgrade to incorporate fertigation. The platform may also be expanded to detect the presence of specific pathogens or pests. Infection alerts and pathogens or pest-specific responses may be issued by integrating sensors for the detection of plant-emitted volatile organic compounds (VOCs). Some such compounds are ethylene that signals general stress, jasmonic acid esters that signal pest attack, and salicylate esters that signal pathogen attack.

2.4.3 Soil-based

Khan *et al.* [110], the authors prepared a dataset containing suitable attributes such as Maximum and Minimum Temperature, Wind speed, Humidity, Rainfall, Solar Radiation, Soil Type, Crop Type, and Crop Water Usage for forecasting irrigation water demand. To predict an irrigation occurrence, they applied and compared the effectiveness and efficiency of several data mining classification methods such as logistic regression decision tree, SVM, SysFor, and ANN. Predictions of accuracy made by SVM, logistic regression, and traditional ETc method is found to be 78%, 75%, and 77%, respectively. Interestingly, ANN performs better than SVM by carefully forecasting the water demand to actual water used with 95% accuracy. SysFor 's proximity of prediction accuracy performs best at 97.5%, followed by decision tree 96 percent accuracy. Therefore, we suggest that SysFor, decision tree, and ANN techniques are ideally suited to forecast an irrigation event. They suggested that the performance of the prediction model can be further improved by adding more attributes that have a strong influence on water requirements such as seepage, soil moisture, and cropping stage.

Navarro *et al.* [111] proposes a Smart Irrigation Decision Support System, SIDSS, to manage irrigation in three commercial citrus tree orchids in southeastern Spain. They design and developed a closed-loop automatic decision support system that adapted to local disturbances and measurement errors. Two machine learning techniques, Partial Least Square Regression (PLSR) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) are proposed as the reasoning behind the adaptive SIDSS. The plantation weekly irrigation needs are estimated based on both soil and climatic variables gathered by several autonomous field nodes. A VWC sensor was used to record the soil variables of the site. Hence, a detailed analysis of the plantation's soil textures is required to extrapolate the sensor information to the specific soil texture where the DSS is to be deployed. Performance is measured against a human expert 's decisions. Although the use of a soil sensor resulted in a 22 percent lower weekly error compared to the performance of using only weather information, the authors suggested that soil sensors should be removed from the experiment due to their high dependence on site-specific soil properties.

Goap *et al.* [112] proposed an IoT based smart irrigation architecture along with a hybrid machine learning-based approach to predict the soil moisture. The soil moisture is affected by several ground parameters like soil temperature, and environmental variables, e.g., air temperature, Ultraviolet (UV) light radiation, and relative humidity of the crop field. SVR, coupled with K-means, was used to predict the upcoming soil moisture values based on which decision is made whether to irrigate the field or not. This paper proved the concept that IoT based smart irrigation management systems can help in achieving optimum water-resource utilization in precision farming.

2.4.4 ML-based

Goldstien *et al.* [113] aimed to use data collected from different sources to predict irrigation recommendations. In addition to actual irrigation records determined by the agronomist, the dataset captured various sensor-based features, weather features, and plot-describing features. Models were developed using eight different variables subsets to determine which variables consistently contributed to predictive accuracy. It was also shown that different feature sets were optimal for different plots, suggesting that developing unique models for different plots might be useful. Based on this, different irrigation recommendations were developed using the traditional linear regression, Gradient Boosted Regression Trees (GBRT), and Boosted Tree Classifier (BTC). Results showed that GBRT and BTC models were more accurate in predicting irrigation decisions than linear regression models, with 93 and 95 percent success rates, respectively. The resulting model can automate irrigation decision making and can significantly facilitate the irrigation planning process of the agronomist. The proposed solution was suggested to be applied to other crops as well as for new decision-making processes in agriculture.

Chapter 3

Problem Formulation and Experimental Design

The growing water crisis and its increasingly negative impact on the economy and environment call for judicious use of available water resources of the country. During 20^{th} century, water management mostly focused on a search for new water supplies, based on water storage and distribution from large projects. However, the focus at present is on-demand management by efficiently allocating available water resources amongst various competing uses. As agriculture is the largest consumer of water, efforts to efficiently use water in agriculture can play a crucial role in meeting its future needs as well as of other sectors of the economy. Sustainability of the country's agriculture is also crucial because of its direct impact on the livelihood of millions of farmers.

3.1 Problem formulation

Efficient forecasting of crop water demand is necessary to prepare an efficient irrigation schedule for optimal use of available irrigation supplies. Irrigation scheduling is the decision of how much irrigation water to apply to a field and, at what time, to improve irrigation efficiency. Over-irrigation leads to wastage of water, labor and energy; leaching of costly nutrients below the crop root zone; and reduces soil aeration and thus the crop yields. On the other hand, under-irrigation leads to moisture stress in the plant and reduces crop yield. To maximize the benefit, two critical elements in irrigation scheduling are the accurate measurement of the volume of water applied and its uniform distribution across the field. It also enables the irrigator to plan water application among the various fields to abate water stress and thus improve crop quality and yield as well as net returns by use of 'saved' water to irrigate additional area during the periods of water-shortages.

There is a need for a completely automated irrigation system so that the wastage with respect to water; that comes with the involvement of humans can be done away with. Hence, a system that can predict the schedule of the irrigation required by the crop(s) in the field needs to develop. Though various models have been developed for this purpose, yet the application of data mining methods and techniques to model the data collected from agricultural fields needs to be explored further. The models that exist are either linear or have just started to explore the relationship between various parameters such as as-soil, crop, field, and environment. With the use of WSNs for agricultural data collection, the volume of data collected has increased. Hence, machine learning needs to be utilized to discover new insights from the data. In the present times, agriculture needs data scientists who can process the data collected over the years to give well-researched recommendations.

The problem of irrigation scheduling consists of when to irrigate and how much to irrigate, i.e., efficient forecasting of crop water demand. To minimize the non-productive loss of water while maintaining the crop yield and hence increasing the water productivity. Forecasting can be based on soil matric potential while maintaining the soil matric potential in safe limits where crop yields do not decline and avoid non-productive water loss. The efficient use of water depends on the timely application of water in the right way and the right amount. The objective of irrigation scheduling is to minimize wastage of water and irrigation costs while maximizing yield and crop quality.

Daily demand needs to be forecast to schedule pumping effort and minimize cost. There is a need to examine methodologies for demand modeling and prediction in a real-time environment for an on-demand irrigation system. The focus has to be on the development of total daily volume demand models. The daily crop water requirements estimated by the rate of percolation and evaporation do not always meet the actual need.

3.2 Research gaps

From the detailed literature review, the following research gaps have been identified:

i. Machine learning algorithms are mainly being used to forecast weather, rainfall, and yield and irrigation water demand as well. However, the number of parameters being considered is limited to weather parameters. For better model development for irrigation scheduling, focus from ET needs to be shifted to soil parameters like available water content, soil texture, topology, salinity, porosity, etc.

ii. There is more than one approach to forecast irrigation water requirements such as a statistical approach, sensor network-based approach, and as a machine learning approach. There are instances in literature where one model has been compared to another model of the same approach. However, a comparison of the different models on one data set and analysis of the obtained results on specific evaluation parameters is yet to be done.

iii. There aren't many standardized and fully automated irrigation systems available. Either the automation is in parts that inter-operate to provide complete service or the development has not followed a specific standard. Hence, standards need to be developed.

iv. Farmers have resorted to indiscriminate pumping of groundwater so that the field irrigation demand can be met. This is further encouraged by subsidies given on installation of pumps in the field and free electricity supplied to the farm. Hence, there is a need for guidelines backed by analytical research that lays down the amount and frequency of irrigation water supplied to the crop.However with current irrigation support policies like free irrigation water and electricity, farmers have very little incentives to adopt efficient water management practices.

v. Given the present conditions, climate change has arrived and is there to stay. The 70-80% of annual rainfall that was received during monsoon, i.e., from June to September, has now reduced, and the amount of winter rains has increased. It not only spoils the crop in the field but also promotes pumping of groundwater to make up for the loss. There is a need to document and study its effect so that water can be used effectively and efficiently.

3.3 Objectives

The main objectives of the proposed research work are mentioned as follows:

i. To study the existing methods for estimating crop water requirements and collect real data based on factors affecting.

ii. Propose and validate an efficient prediction model.

iii. Compare the performance of the proposed model with the existing prediction methods.

3.4 Methodology

A conceptual model generally aims to formulate the daily water requirements for crop irrigation by the rates of percolation and evapotranspiration that have been predicted at the stage of irrigation planning. Many models have been used to simulate crop water requirements. These do not always match actual use due to changes in field conditions such as weather parameters and farmers' perceptions. It should be pointed out that data availability often determines the model choice. Continuous measurements of climatic data (precipitation, temperature, relative humidity, wind speed) can be quickly and cost-effectively obtained as compared to continuous measurements of soil characteristics, initial soil moisture, infiltration, and thus are more suitable for operational forecasting purposes. A list of few parameters and their brief description is given in Table 3.1.

The existing approaches for predicting water demand (time and amount of irrigation) for the crop. The factors affecting the crop water requirement have been identified and selected for further analysis. The data-set has been collected from the experimental site of Central Soil Salinity Research Institute (CSSRI)-International Centre for Improvement in Wheat and Maize (CIMMYT) Research Platform at Karnal, Haryana (India). Various machine learning models have been implemented using Python and its libraries. A comparative analysis of their performance has been carried out.

For this purpose, a multi-level model has been developed and implemented to predict the crop water requirement using machine learning techniques. The model has been benchmarked against APSIM (an industry standard for forecasting crop water demand for drip-irrigated fields).

3.5 Experimental Set-up

The observations related to the amount of irrigation water used in rice crop with subsurface drip irrigation (SDI) were collected from the experimental farm of Indian Council of Agricultural Research - Central Soil Salinity Research Institute (29° 70'N, 76° 95'E), Karnal, Haryana, India (Figure 3.1). The experimen-

Table 3.1: Parameters affecting CWR and their brief description.

Parameter	**Description**
Solar radiation	It is radiant energy emitted by the sun.
Air temperature	It is a measure of hotness or coldness of the air and is the most common weather parameter.
Air humidity	It is the amount of water vapor in the air and indicates the likelihood of precipitation, dew, or fog.
Wind speed	It is the measure of motion in the air with respect to the surface of the earth.
Wind direction	It is an indicator of the direction of the wind.
Crop parameters	Crop type, crop variety, growth stage, and crop height are some crop features that may be considered.
Soil salinity	It is the salt content of the soil.
Plant density	It is the number of individual plants that occur within a given unit of space.
Type of irrigation system	Different types of farm irrigation systems are flood, sprinkler, drip, and micro-irrigation.
Soil Texture	Based on particle size, soil can be categorized into clay (d<0.002 mm), silt (d:0.002-0.05 mm), sand (d> 0.05 mm), coarse, medium and fine.
Soil Structure	It is the arrangement of the solid parts of the soil and the pore space between them.
Soil Drainage	It is the process of water removal from the surface and sub-surface of an area.
Soil Fertility	It refers to the ability of soil to sustain plant growth.
Type of farm	A farm may be under a monoculture system or with a variety of crops. It may be combined with or separate from livestock.
Residue management	There is a difference in burning the residue or mulching it in the field.
Water quality	It refers to the chemical, physical, biological and radiological characteristics of water.
Evapotranspiration	It is the process of water transfer from the land to the atmosphere by evaporation and transpiration from plants.
Rainfall	It is the quantity of rain falling within a given area in a given time.

tal site represented the sub-tropical and semi-arid climate and characterized by three distinct seasons, i.e., Kharif (July-October), Rabi (November-March) and Zaid (April-June). The Kharif season (wet monsoon season) coincides with the South-West monsoon and receives 80% of total average annual rainfall (670 mm). The site was under a continuous rice-wheat system for ten years before the establishment of the experiment. The soil of the study site (0–15 cm layer)

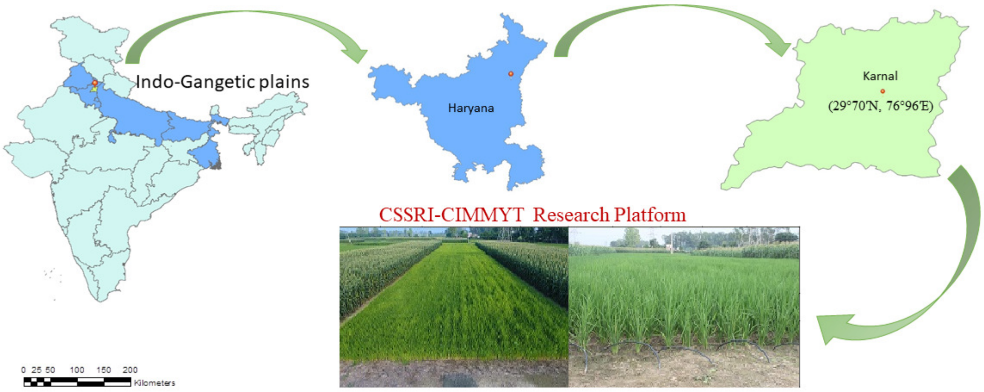


Figure 3.1: Location and layout of experiment

was loam in texture (34% sand, 46% silt, and 20% clay) with a pH of 8.0 (1:2 soil: water). The soil was low in major available nutrients (N, P, K) and organic carbon (0.45%).

The rice crop was sown under anchored wheat residue condition using the Happy Seeder machine. The drip irrigation (DI) system consisted of polyethylene laterals with an inside diameter of 16 mm and was laid parallel to sowing rows. The laterals were provided with in-line emitters of 2.0 liters per hour capacity at a pressure of 135 kPa and spaced at 30 cm so that the complete plot area was wetted. Each dripline served three rows (67.5 cm) of rice crops. Groundwater for the DI system was filtered using a hydro-cyclone filter coupled with a 100-micron screen filter fitted at the source of irrigation.

In the rice, irrigation was applied based on soil moisture potential (SMP) using a tensiometer (IRROMETER, River-side, California) installed at 15 cm soil depth mid-way between the two crop rows in the center of each plot. The irrigation in the crop was given at -20 to -30 kPa. In DI, the water meter is fitted after the venture injector in between the horizontal pipe just above the ground to make ease in taking readings. All three replications of a rice scenario received a similar amount of irrigation at the same time. The volume of water applied in SDI plots for rice crops in each irrigation was measured with the help of water meter (Dasmesh Mechanical Works, Punjab, India) fitted to the delivery pipe.

Chapter 4

ML-based crop water demand forecasting

Rice is one of the world's most popular food crops. Since its production is dependent on intensive water use, water management is critical to ensure the sustainability of water resources. However, minimal data is available on water use in rice irrigation. In this chapter, traditional machine learning methods have been used to predict the irrigation schedule of rice daily. The data of the year 2013-2015 is used to train the models and to optimize it further. The data of 2016-2017 is used for testing the models. Correlation thresholds are used for feature selection, which helps in reducing the number of input parameters from the initial 26 to final 11. The models estimated the crop water demand as a function of weather parameters. Results show that Adaboost performed consistently well with an average accuracy of 71% as compared to other models for predicting the irrigation schedule.

4.1 Introduction

Rice (Oryza Sativa L.) is one of the majorly consumed food-crop in Asia [114]. A significant share of available freshwater is used to irrigate rice [41]. Lowland irrigated rice is inefficient [115] because a large proportion of the irrigation water is lost to seepage and deep percolation [116]. The water-use efficiency of rice is lower than in other crops. On average, 2500 liters of water are used, ranging from 800 liters to more than 5000 liters to produce 1 kg of coarse rice [117]. An increase of 10% in irrigation efficiency can help extend irrigation to an additional 14 million ha. Solely reducing water use in puddled transplanted rice resulted in a proportional reduction in yield, hence various management practices of rice cultivation have to change simultaneously to improve water productivity, without compromising on the productivity of yield.

The soaring population is resulting in huge pressure on the limited freshwater resources of the world. The growing population is also contributing to an increased demand for water from domestic as well as industrial sectors. Consequently, increasing the scarcity of available water demands efficient use of irrigation water in agriculture. Since the available water sources are already almost completely exploited, there is a need for achieving increased water productivity through efficient irrigation. Usually, 15% of the water delivered is lost in conveyance, another 15% is lost during on-farm supply through field channels, and 25% is lost due to inefficient water use practices, thus leaving only 45% of the water supplied to be utilized by the crop [118]. The efficiency of irrigation water use depends on the timing, duration, and method of irrigation employed. Information from multiple sources such as soil, plant, and atmosphere is required. There are two different approaches for estimating the crop water requirement to manage irrigation requirement in crops [119] efficiently; one, a conceptual approach based on various factors viz. soil moisture, seepage, and evapotranspiration and the other, a *data-driven* approach approach based on training a model using the available data. The *data-driven approach* is more accurate than the conceptual approach [119]. Saleem et al. [120] developed a framework for controlling real-time irrigation scheduling that took into account some of the most common constraints, such as restricted water availability or the maximum or minimum quantity of water applied. It is desired for effective irrigation scheduling that the water supply matches the water demand as closely as possible. Rice has been generally grown under puddled transplanted conditions over large tracts in irrigated areas. However, for improving water use efficiency in irrigated rice, experiments are being conducted to raise the crop under a drip irrigation system. In this, the water slowly drips out of an emitter, either into the root zone or onto the surface. This minimizes fertilizer and water use. [121].

Data available on water use in drip irrigated rice, and its irrigation schedule is limited. The study [122] attempts to estimate the crop and irrigation water requirement of rice. In their work [73], the authors have integrated decision support systems and a wireless sensor network to achieve an autonomous, flexible zone-specific closed-loop irrigation system.

Crop water demand may be calculated using direct measurements of stem or leaf water potential of a plant [123]. More commonly, however, it is measured from indirect measurements such as moisture content of the root zone [124] and reflectance in the near-infrared bands [125]. Saturation, field capacity, or wilting point account for the soil water content. These then can be used to estimate the total available water and the readily available water. The Penman-Montieth method is a commonly used method for determining the crop water requirement [126, 127]. In this, evapotranspiration for a reference crop (ETo) is estimated using field climatic data. This ETo is then multiplied by a pre-determined crop-specific coefficient to obtain crop water use. This crop coefficient is dependant on many factors such as local crop characteristics, soil moisture, day of the year (for planting), climatic conditions, length of the growing season, irrigation regimes, and management [128, 35]. Unlike the one-step approach [127, 35], this two-step approach is a simpler and more reliable method to estimate actual evapotranspiration (ETa) [122].

In evapotranspiration based estimation of water demand, the reliability of the climatic data determines the accuracy of crop water demand estimates. In many cases, recorded weather data is reduced to the number of variables measured by a weather station in an area because of which there is a limit to the estimation of evapotranspiration (ET) [129]. Torres et al. [130] used the Multivariate Relevance Vector Machine algorithm and a Multi-layer Perceptron model to forecast ETo. In one of the two above mentioned approaches, the future ETo values are calculated as a times series forecast from historical data; or as (the other approach) estimates the ETo values using the forecast climatic data in a Hargreaves ETo equation. It was found that the latter approach that took into consideration the weather parameters had better goodness of fit results. The Pennman-Montieth method of the Food and Agriculture Organization of the United Nations (FAO) is better suited for estimating crop water requirements.

For sensor-based estimation of water demand, sensors may be used to collect information about the conditions of the field and environment. Climate parameters based models have an open-loop structure while the models using data from soil sensors form a closed-loop one [111]. However, closed-loop systems did not get wide acceptance due to the technological limitation of the sensors. This has changed in the last few years due to the availability of more suited and versatile sensors [63]. Variability in soil and crop characteristics, as well as their irrigation requirements overgrowth cycle [131] along with non-uniformity of irrigation methods [132], adds complexity which may make the model results deviate from ideal. Human experts who can comprehensively interpret the different variables are few, not readily available, and slow to analyze the data. Machine learning models can assist experts instead. Crop parameters

[133], volumetric water content of soil [134], ET [135], total available water in root zone [136] and stem water potential [137, 138] have been estimated in the past using machine learning techniques.

Most of the modeling techniques that have been developed to characterize the state of irrigation systems lack practical application. This is due to the unavailability of the required data and the overhead expense incurred during data acquisition. Machine learning approaches based on data-driven learning modeling such as artificial neural networks, support vector machines have been used [119] to overcome this drawback. The time-series nature of soil water content variations has been simulated using leaning machine models. However, since the objective was to estimate the values from historic SWC values, it was a somewhat restrictive approach to determine irrigation scheduling based on that [139].

Perea et al. [140] successfully predicted the presence or absence of an irrigation event using genetic algorithm based input parameters to build an optimized decision tree. Khan et al. [141] performed a comparative study checking for the effectiveness of decision tree, artificial neural network, systematically developed forest, support vector machine, and logistic regression in predicting the irrigation water requirement. They recommend that decision tree, systematically developed forests, and artificial neural networks were best suited for forecasting the irrigation water demand. Andriyas and McKee [142] used decision tree and random forest for making short-term water demand forecasts for alpha alpha.

The rest of the paper is organized as follows: a brief overview of the data set and its features is mentioned in Section 4.2. The methodology describing the experiments is covered in Section 4.3. Results and discussion are in Section 4.4, and the conclusion is in Section 4.5.

4.2 Materials and methods

4.2.1 Site description

The experimental site on (76 $^\circ N$ 96 $^\circ E$) is located at Karnal in Haryana, India, at a farm of the Central Soil Salinity Research Institute (CSSRI). The weather data at the site was recorded using an automated weather station. Daily weather data, including minimum and maximum temperature, humidity,

Table 4.1: Soil properties at experimental site

Soil Property	**Soil Sampling Depth) Mean ± SE)**	**Unit**
Clay	19.89 ± 0.50	(%)
Silt	46.07 ± 0.76	(%)
Sand	34.03 ± 0.77	(%)
pH	8.00 ± 0.02	(dS m-1) (1:1 soil:water)
EC	0.37 ± 0.02	(dS m-1) (1:1 soil:water)
Total Carbon	0.56 ± 0.01	(%)
Available P	5.74 ± 0.29	(mg kg-1)
Exchangeable K	130 ± 1.73	(mg kg-1)
TN	0.06 ± 0.002	(%)
Particle Density	2.57 ± 0.01	(g cm-3)

wind speed, rainfall and rate of evapotranspiration is measured The incorporated sensors are air temperature and humidity sensor, wind sensors, thermopile pyrometer, and automatic rain gauge. All variables are sampled every hour and recorded daily. The site is a reclaimed alkali loam soil. Table 4.1 lists the soil characteristics of the experimental site. Laser-leveler was used to level the field, which was then divided into 12 permanent plots. These plots were separated by 1.0 m wide and 0.20 m high earthen bunds. The crop stand was uniform across the entire area. Drip irrigation lines (67.5 cm apart with emitter spacing of 30 cm) were used for each row.

4.2.2 Feature measurement

Multiple parameters affect crop water demand. Some of these are weather-based, some crop-based, other soil-based, and some other. Soil parameters such as salinity, texture, structure, drainage, and fertility have not been accounted for because the data has been collected from subplots of a farm, hence not providing variation that may have otherwise been modeled. However, soil temperature at different levels and at different times during the day has been used to capture the effect of soil parameters somehow. Crop parameters such as crop type, crop variety, crop height, and plant density have also not been explicitly included. Crop growth stage and number of days since sowing have been engineered to reflect the crop parameters partially. Other parameters such as type of irrigation system, type of farm, residue management practice, and water quality have not been accounted for, as there was no source of variability that may have helped the model to learn better. Weather-based (such as solar radiation, rainfall, air humidity, wind direction, wind speed, evapotranspiration, and air temperature) parameters have been used in this study. They have been described in Table 4.2.

Table 4.2: Features of data set obtained from experimental site

Code	**Type**	**Name**	**Unit**	**Instrument**
V01	Derived from Date	No.of days after sowing		
V02	Derived from Date	Stage of growth		
V03	Temperature	Maximum temperature	$^{\circ}C$	Thermometer
V04	Temperature	Minimum temperature	$^{\circ}C$	Thermometer
V05	Temperature	Gross Minimum temperature	$^{\circ}C$	Thermometer
V06	Temperature	Dry bulb at 6:00 A.M.	$^{\circ}C$	Thermometer
V07	Temperature	Dry bulb at 2:00 P.M.	$^{\circ}C$	Thermometer
V08	Temperature	Wet bulb at 6:00 A.M.	$^{\circ}C$	Thermometer
V09	Temperature	Wet bulb at 2:00 P.M.	$^{\circ}C$	Thermometer
V10	Humidity	Humidity at 6:00 A.M.	Mm of Hg	Humidity probe
V11	Humidity	Humidity at 2:00 P.M.	Mm of Hg	Humidity probe
V12	Humidity	Relative humidity at 6:00 A.M.	Mm of Hg	Humidity probe
V13	Humidity	Relative humidity at 2:00 P.M.	Mm of Hg	Humidity probe
V14	Soil heat flux	Soil temperature 5 cm below surface at 6:00 A.M.	$^{\circ}C$	Thermometer
V15	Soil heat flux	Soil temperature 5 cm below surface at 2:00 P.M.	$^{\circ}C$	Thermometer
V16	Soil heat flux	Soil temperature 15 cm below surface at 6:00 A.M.	$^{\circ}C$	Thermometer
V17	Soil heat flux	Soil temperature 15 cm below surface at 2:00 P.M.	$^{\circ}C$	Thermometer
V18	Soil heat flux	Soil temperature 20 cm below surface at 6:00 A.M.	$^{\circ}C$	Thermometer
V19	Soil heat flux	Soil temperature 20 cm below surface at 2:00 P.M.	$^{\circ}C$	Thermometer
V20	Sunshine hours	Sunshine hours	Hrs/day	
V21	Wind speed	Wind speed	Km/hr	Anemometer
V22	Wind direction	Wind direction		
V23	ET	ET		
V24	Rain	Rainfall	Mm/day	
V25	Derived from Rain	No. of days since last rain		
V26	Derived from Irrigation	No. of days since last irrigation		
V27	Target	Amount of irrigation	Mm/day	

Where $^{\circ}C$ = Degree Celsius, Mm of Hg = Millimeter of Mercury, Km/hr = Kilometer per hour and Mm/day = Millimeter per day.

4.3 Methodology

4.3.1 Data collection and pre-processing

The first step in any machine learning approach is data collection and processing. As data about the weather and the data about irrigation came from two different sources, it had mismatched sampling intervals. This resulted in the need to compile the data from the weather station and the flow meter into one data set. This was done with the guidance of an agriculture expert. This resulted in a data set that contained a one-dimensional array containing all the values corresponding to the input variables for a day. Five data sub-sets (corresponding to the year 2013, 2014, 2015, 2016, and 2017) were created (having variable season length of 127, 98, 92, and 102 entries each) and compiled into one. Since the data set represents 3 replications of the experimental plot for the same set of input parameters, we obtained different quantities of water as

Table 4.3: Features incorporated in different data sets

Set No.	**Parameters Included**
Set 1	V01, *V02*, V03, V04, *V05*, V06, *V07*, V08, *V09*, *V10*, V11, V12, V13, V14, *V15*, *V16*, *V17*, *V18*, V19, V20, V21, V22, V23, V24, V25, V26
Set 2	V01, V03, *V04*, V06, *V08*, V11, V12, *V13*, V14, V19, V20, V21, V22, V23, V24, V25, V26
Set 3	V01, V03, V06, V11, *V12*, *V14*, V19, V20, V21, V22, V23, *V24*, V25, V26
Set 4	V01, V03, V06, V11, V19, V20, V21, V22, V23, V25, V26

the target value. The average was considered as the modified target and the standard deviation as the range. Also, the wind direction was a categorical variable, containing different wind directions. One-hot encoding was applied to the categorical variable wind direction, and it was converted to a Boolean one.

Irrigation was applied based on tensiometer readings. To monitor the soil matric potential (SMP), gauge-type soil tensiometers (IRROMETER, Riverside, California) were installed at 15-cm and 30-cm depths in all plots immediately after each crop was planted. Irrigation water applied to each plot was measured using a Woltman® turbine water meter. To measure the amount of water applied in each irrigation, the water meter reading (kiloliter, kL) was recorded at the start and end of the irrigation of each plot, each irrigation was applied until a 5-cm flooding depth was achieved. The amount of irrigation water applied was calculated as water depth (mm).

4.3.2 Feature selection

The next step applied was feature selection, which was accomplished using correlation. Since the variables are continuous, Pearson Correlation Coefficients were used to create a matrix depicting the dependencies in the data. A subset of features (Table 4.3) was obtained by dropping off the variables with values at fixed thresholds and discarding those that had a value above the threshold. Set 1 includes all the variables mentioned in Table 4.2. Set 2 discards those variables which had a correlation coefficient greater than that of 0.9. Set 3 and Set 4 further discarded those variables that had a correlation coefficient greater than 0.8 and 0.7, respectively.

4.3.3 Model selection

Models chosen were Linear Regression (LR), Support Vector Regressor (SVR), Decision Tree Regressor (DTR), Random Forest Regressor (RFR), Extra Tree Regressor (ETR), Adaboost Regressor (ABR), and Gradient Boosting Regressor (GBR). Sklearn was used to implement these models in Python. Along with the traditional machine learning methods, a Neural Network has also implemented on Tensorflow.

LR model is based on the assumption that the output variable can express as a linear combination of the input variables. Given the complex interaction between the input variables, the linear model was unable to capture the behavior of the input features for this regression exercise. Hence, it was found to be highly unsuitable for use. An SVR finds the hyper-plane that differentiates the data plotted as a point in n-dimensional space where the value on the co-ordinate is equal to the feature value. A DTR uses a recursively partitioning approach where the growth of the tree is achieved by splitting at each attribute iteratively. This model is computationally cheap, easy to understand-implement-use, straightforward to train, and easy to interpret. Overfitting is prevented by pruning the tree. Random Forest is an accurate and efficient method which grows out of many trees. Each tree returns a value for the target, and a decision is made based on the most votes. An RFR fits a number of decision trees on sub-samples of the data (the size of which is the same as that of the original input sample). An ETR fits numerous randomized Decision Trees on variable-sized sub-samples of the data. Both RFR and ETR improve accuracy by averaging the result. A GBR fits a regression tree on the negative gradient of a loss function. At every iteration, it optimizes an arbitrary differentiable of the loss function. An AdaBoost regressor (ABR) iteratively fits a regressor with adjusted weights on original data. Neural Network (NN) is a series of neurons that can adapt to changing input in such a way that the best possible results can be obtained as output.

4.3.4 Data splitting

In this study, regression estimators have been used to predict the irrigation schedule for rice. Here the data has been divided into 60% and 40% partitions for creating the training set and testing set, respectively. The data was kept in sequence, given the inherent time series-ness of the data. The data splits used for the 3 fold cross-validation are as shown in Table 4.4

Table 4.4: Splits of data set for validation

Iteration	**Years for Training**	**Year for Testing**
1	2013-2014	2015
2	2013-2015	2016
3	2013-20116	2017

Table 4.5: Models and their respective optimized hyper-parameter values

<table>
<tr><th>Model</th><th>Parameter</th><th>Value (Set1)</th><th>Value (Set2)</th><th>Value (Set3)</th><th>Value (Set4)</th></tr>
<tr><td rowspan="2">SVR</td><td>C</td><td>2</td><td colspan="2">256</td><td>128</td></tr>
<tr><td>Gamma</td><td>0.000244140625</td><td colspan="2">0.0039063</td><td>0.0078125</td></tr>
<tr><td rowspan="2">DTR</td><td>Maximum depth</td><td>2</td><td>7</td><td>3</td><td>3</td></tr>
<tr><td>Minimum samples per split</td><td>32</td><td>35</td><td colspan="2">60</td></tr>
<tr><td rowspan="3">ABR</td><td>Learning rate</td><td>0.001</td><td colspan="3">0.0001</td></tr>
<tr><td>Loss function</td><td>Square</td><td colspan="2">Exponential</td><td>Linear</td></tr>
<tr><td>Number of estimators</td><td>500</td><td colspan="2">400</td><td>800</td></tr>
<tr><td rowspan="3">RFR</td><td>Maximum depth</td><td>2</td><td colspan="2">3</td><td>4</td></tr>
<tr><td>Minimum samples per split</td><td>32</td><td>42</td><td colspan="2">49</td></tr>
<tr><td>Number of estimators</td><td colspan="2">125</td><td>25</td><td>10</td></tr>
<tr><td rowspan="3">ETR</td><td>Maximum depth</td><td>10</td><td>7</td><td>13</td><td>9</td></tr>
<tr><td>Minimum samples per split</td><td>35</td><td>41</td><td>35</td><td>33</td></tr>
<tr><td>Number of estimators</td><td>10</td><td>800</td><td>10</td><td></td></tr>
<tr><td rowspan="4">GBR</td><td>Maximum depth</td><td>3</td><td>2</td><td colspan="2">3</td></tr>
<tr><td>Minimum samples per split</td><td>111</td><td>68</td><td>108</td><td>112</td></tr>
<tr><td>Number of estimators</td><td>10</td><td>25</td><td colspan="2">10</td></tr>
<tr><td>Learning rate</td><td>0.4</td><td>0.2</td><td colspan="2">0.5</td></tr>
<tr><td rowspan="3">NN</td><td>Number of hidden layers</td><td>0</td><td>2</td><td>2</td><td>2</td></tr>
<tr><td>Number of epochs</td><td>3</td><td>13</td><td>6</td><td>7</td></tr>
<tr><td>Network architecture</td><td>(48,1)</td><td>(39,9,11,1)</td><td>(34,8,1,1)</td><td>(32,7,7,1)</td></tr>
</table>

4.3.5 Hyper-parameter optimization

Of the numerous parameters used for model building using sklearn package, only 6 were optimized using GridSearchCV viz: maximum depth (md), maximum features (mf), minimum samples per split (mss), minimum samples per leaf (msl), number of estimators (ne), learning rate (LR), gamma (G) and C, where applicable. GridSearchCV method has an in-built cross-validation functionality that was used for implicit cross-validation while optimizing the model parameters. Three-fold cross-validation was used to find the optimized value of various parameters (which are mentioned in Table 4.5). SVR, DTR, and ABR had a limited number of parameters that needed to be optimized. Hence, their optimization was done in one go. RFR, ETR, and GBR, however, have multiple parameters that can be a candidate for optimization. To reduce the time complexity of the exhaustive grid search on the limited computing resources available, this was done in a two-stage approach. Maximum depth, minimum samples per leaf, and minimum samples per split were optimized at stage one. The values received were then added to the values being sent with maximum features, learning rate, and the number of estimators. The parameter values received from this stage were used to train/test and validate the models. For Neural Network, the architecture of the network and the number of epochs were optimized. The number of layers and number of nodes in each layer was optimized together as these were interdependent, and after finalizing the architecture, an early stopping criterion was used to find the optimal number of epochs. To keep the time complexity minimum, optimization of the number of hidden layers was done up-to-the value of at maximum 2.

4.3.6 Performance evaluation

The performance of the model can be measured using multiple evaluators, some more suitable than others. The models used have been evaluated against mean square error (calculated using eq 4.1), coefficient of determination (calculated using eq 4.4), mean absolute error (calculated using eq 4.2), estimated variance explained (calculated using eq 4.3) and accuracy with respect to a range. 3-fold cross-validation is used to measure the performance of the predictive models.

4.3.6.1 Mean Square Error (MSE)

It is the average of the square of errors. This parameter is similar to MAE, but is sensitive to outliers.

$$MSE = \frac{\sum_{i=1}^{n}(p_i - a_i)^2}{n} \tag{4.1}$$

Where a is actual value of the target variable, p is predicted value of the target variable and n is no. of instances.

4.3.6.2 Mean Absolute Error (MAE)

It is the average of how much does the predicted value deviate from the real value.

$$MAE = \frac{\sum_{i=1}^{n}(p_i - a_i)}{n} \tag{4.2}$$

Where a is actual value of the target variable, p is predicted value of the target variable and n is no.of instances

4.3.6.3 Explained Variance Score (EVS)

It measures how well the model has explained the variations in the dataset. This gives the ratio between the variance of error and variance of true values. The highest value for this parameter can be 1.

$$EVS = 1 - \frac{Var(p - a)}{Var(a)} \tag{4.3}$$

Where a is the actual value of the target variable, and p is predicted value of the target variable.

4.3.6.4 Coefficient of Determination ($\mathbf{R^2}$ score)

It describes the amount of variance explained by the regression model. One is the desired value. R^2 equal to zero means that the model has failed.

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \tag{4.4}$$

Where :
Sum of Squares Total: SST = $\sum(p - \bar{p})^2$, Sum of Squares Regression: SSR = $\sum(\acute{p} - \bar{\acute{p}})^2$, Sum of Squares Error: SSE = $\sum(p - \bar{\acute{p}})^2$, p is predicted value of the target variable

4.3.6.5 Accuracy

It is reflective of the correct predictions made in comparison to the total predictions made. As explained in Section 5.3.3.1, accuracy was calculated with respect to the range.

4.3.7 Results

Table 4.6: Model Evaluation parameters of different models for predicting irrigation events

Set	Model Name	MSE	R^2	MAE	EVS	Acc
1	SVR	541.92	0.10	15.75	0.10	9.5
	DTR	**102.40**	**0.83**	**5.17**	**0.84**	**70.5**
	RFR	109.80	0.82	5.48	0.82	64
	ETR	166.23	0.72	8.71	0.73	12
	GBR	113.17	0.81	5.98	0.82	11
	ABR	115.49	0.81	5.56	0.81	**70.5**
	NN	407.79	0.29	16.40	0.21	27.8
2	SVR	537.45	0.10	18.35	0.10	9
	DTR	187.59	0.69	7.26	0.70	67
	RFR	**123.61**	0.79	5.90	**0.80**	66.5
	ETR	173.63	0.71	8.51	0.72	18
	GBR	130.56	0.78	6.34	0.79	10
	ABR	125.23	**0.79**	**5.86**	**0.80**	**71**
	NN	374.08	0.17	15.99	0.13	29.9
3	SVR	563.39	0.06	18.56	0.06	10.5
	DTR	142.00	0.76	6.35	0.77	65
	RFR	133.59	0.78	6.17	0.78	69
	ETR	159.00	0.73	8.21	0.75	9
	GBR	142.89	0.76	6.55	0.77	7
	ABR	**125.86**	**0.79**	**5.86**	**0.80**	**70.5**
	NN	275.47	0.05	13.67	0.11	28.7
4	SVR	521.49	0.13	17.18	0.13	9.5
	DTR	142.00	0.76	6.35	0.77	65
	RFR	133.20	0.78	6.07	0.78	68
	ETR	145.36	0.76	7.04	0.77	36
	GBR	146.36	0.76	6.73	0.77	9
	ABR	**125.79**	**0.79**	**5.83**	**0.80**	**71**
	NN	300.26	0.11	14.15	0.08	33.8

Traditional machine learning models like the Linear Model, Support Vector Machine, Decision Tree, and Random Forest was used to predict the irrigation in drip-irrigated rice. The performance of these models is shown in Table 4.6.

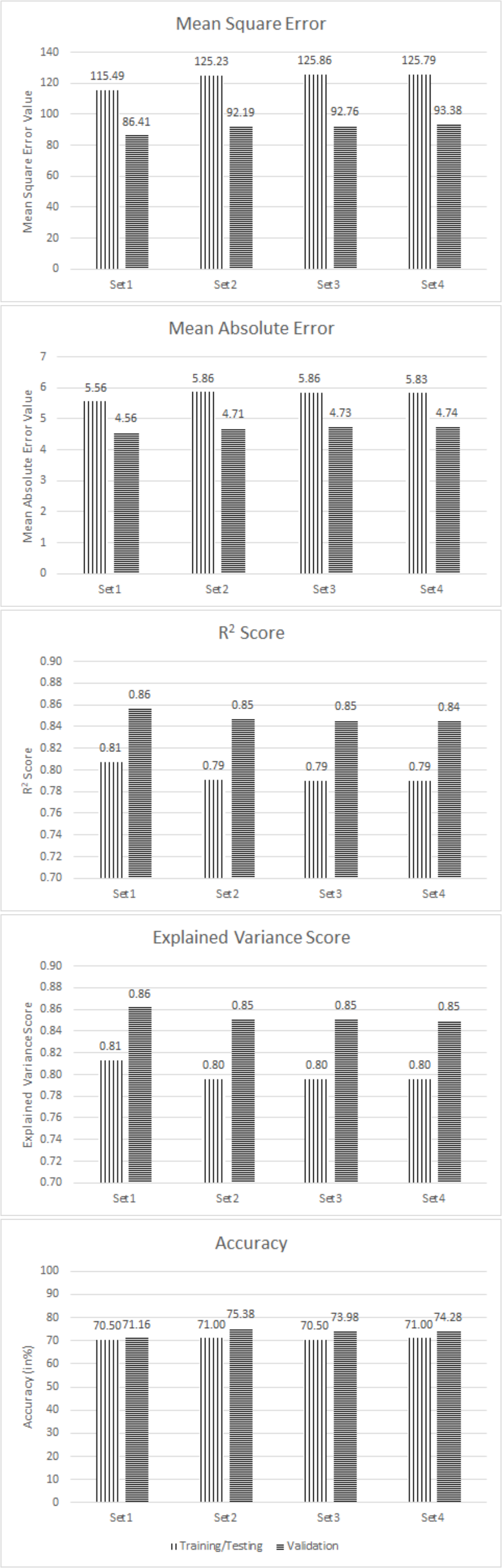


Figure 4.1: Comparison of results of Adaboost for the various data sets

The SVR was unable to give a decent prediction. This may be because the dynamics of the data were beyond its modeling capabilities. The NN also did not perform well which maybe because of the time-series nature of the data. Also, the black-box nature of the method reduces the explain-ability of the results.

DTR and RFR performed reasonably well. Of the four random forest regressors used, GBR and ETR did not perform well. The stage-wise optimization of the loss function in gradient boosting and the averaging the accuracy of trees constructed using a random subsample size less than that of the original input sample size in Extra-Tree Regressor did not fit well with the data being processed in this application. However, the averaging of the accuracy of trees with subsample size the same as the input sample size in Random Forest Regressor and creating additional error adjusted trees for Adaboost, both suited the application quite well. Where ever applicable, the value for the minimum number of samples per split was 2, and the value of maximum features (found through optimization) came out to be equal to 'auto.'

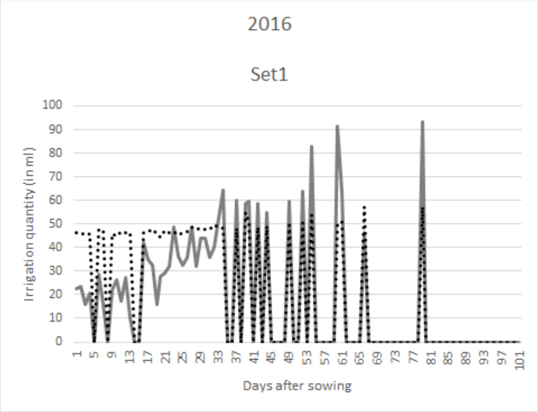
The performance of the estimators is analyzed over the four datasets. For Set 1, DTR gives the best performance. It has the least error and highest R^2 and explained variance score. It may be noted that the accuracy with respect to the range for both DTR and ABR is the same.

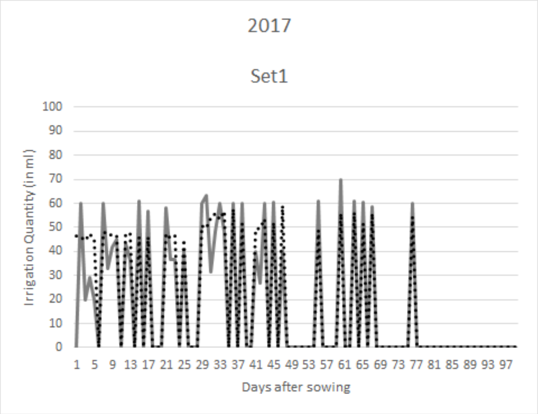
For Set 2, RFR has the highest value of explained variance score and R^2 while having the least error. Although accuracy with respect to the range is higher in the case of ABR, given the better performance of RFR on more than one evaluation parameter, it is deemed better performing.

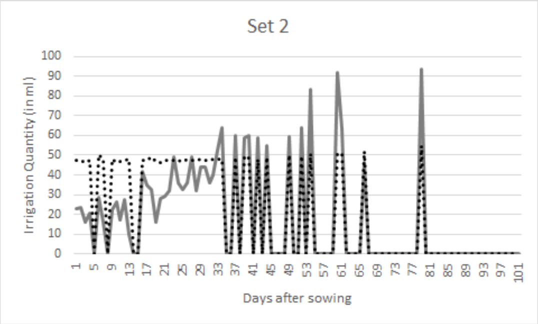
Observing the performance parameter values for Set 3, it can be seen that ABR has the best values for all the parameters. It has the least value for error and the largest value for R^2, explained variance score and accuracy with respect to range.

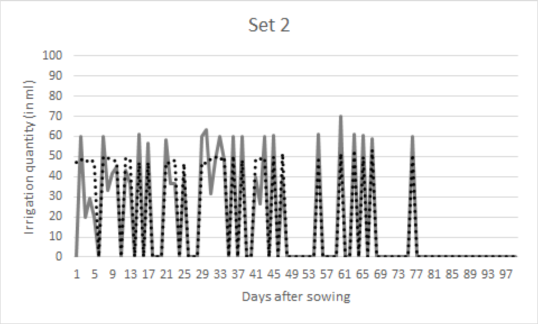
For Set 4 as well, it can be observed that with the largest values for R^2, explained variance score and accuracy with respect to range, Adaboost is the best performing model. It also has the smallest value for MSE and MAE.

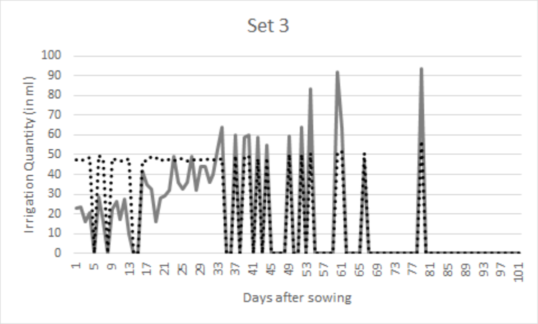
On observing the value of hyperparameters, it can be seen that ABR needed the least number of trees in the forest when the exponential loss function was being minimized during the boosting iteration (400 trees for Set 2 and Set 3); a little more when square loss function was being minimized (500 trees for Set 1)

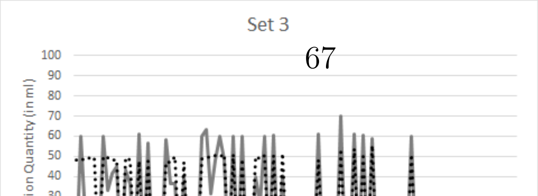

2016
Set1
Irrigation quantity (in ml)
0 10 20 30 40 50 60 70 80 90 100
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93 97 101
Days after sowing


2017
Set1
Irrigation Quantity (in ml)
0 10 20 30 40 50 60 70 80 90 100
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93 97
Days after sowing


Set 2
Irrigation Quantity (in ml)
0 10 20 30 40 50 60 70 80 90 100
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93 97 101
Days after sowing


Set 2
Irrigation quantity (in ml)
0 10 20 30 40 50 60 70 80 90 100
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93 97
Days after sowing


Set 3
Irrigation Quantity (in ml)
0 10 20 30 40 50 60 70 80 90 100
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89 93 97 101
Days after sowing


Set 3
ion Quantity (in ml)
30 40 50 60 70 80 90 100

and the most number of trees when the linear loss function was being minimized over the boosting iterations (800 trees for Set 4).

Given that only 3 hyper-parameters are to be tuned for ABR, 4 for DTR, and 5 for RFR, the most exhaustive and somewhat least time consuming (in comparison) parameter tuning can be performed for ABR. On observing the 3-fold validation test results in Figure 4.1, it can be observed that the principle of machine learning (Performance of the machine should improve as the amount of data increases over time) is validated. MSE and MAE have been minimized, and R^2 score, EVS and accuracy have improved.

4.3.8 Discussion

Figure 4.2 shows that the model can capture the behavior of the target variable sufficiently. The predicted values for the year 2016 and 2017, both reflect the actual value closely. For predictions made on Set 1, Set 2, Set 3 and Set 4 for both the years follow a similar trend. The model is better able to predict the event of irrigation than the required quantity. However, ABR has performed well over three sets. Hence it is safe to say that it is adaptive to change in the number of input parameters. Set 4 consists of all the most important factors affecting water demand. Also, it has the least number of input features as compared to the other three sets. Hence using it can be easy to acquire, store, process, and use. Given this ease without any considerable loss of performance, it can be concluded that the ABR algorithm on Set 4 is most suitable for forecasting farm-level irrigation.

Since all the sets have similar weather parameters, ABR performs equally well. However, Set 4 is said to be better because it has the least number of input parameters. Set 1 has 9 variables depicting the air temperature, whereas Set 2 has 5; Set 3 and Set 4, both have 3 variables. Similarly, for humidity, Set 1 has 4, Set 2 has 2, Set 3 and Set 4 both have 1 parameter respectively. Set 1 has 6 parameters for soil temperature; Set 2 has 2, Set 3 has 1, and Set 4 has none. Days after sowing, rainfall, ET, sunshine hours, wind speed, and days since last rain are the constant parameters across all Sets.

DTR has previously been used to classify irrigation events and RFR has been used to predict irrigation. However, our approach combines the two objectives and gives a reasonably good performance. Hence, improving on the previous research.

4.4 Conclusion

In this chapter, we explored traditional machine learning methods for estimating the irrigation schedule of drip-irrigated rice. The performance of the models is expressed using the standard regression performance parameters such as mean absolute error, median absolute error, explained variance score, coefficient of determination, mean square error, and accuracy. The models are validated using a 3 fold iterative process in which the previously used dataset becomes the training set, and the data of the next time quantum becomes a test set. The performance of the models has been improved by using GridSearchCV for fine-tuning hyperparameters. Given the widespread application of neural networks in the field of irrigation scheduling, we can extend this current work using the concepts of deep learning. In the coming future, water availability is likely to be a more significant constraint than land, so it is time to likely change mindset from improving agricultural productivity from per unit land to per unit of water. Hence, the use of a computationally intelligent approach becomes more relevant in the present scenario.

Chapter 5

LSTM-based multi-level model for forecasting

Rice is a staple food crop around the world, and its demand is likely to rise significantly with growth in population. Increasing rice productivity and production largely depends on the availability of irrigation water. Thus, the efficient application of irrigation water such that the crop doesn't experience moisture stress is of utmost importance. In this chapter, a Long Short Term Memory (LSTM) based neural network with Logistic Regression has been used to predict the daily irrigation-schedule of drip-irrigated rice. The correlation threshold of 0.75 was used for the selection of features, which helped in limiting the number of input parameters. Also, a dataset based on the recommendation of a domain expert, and another used by the tool Agricultural Production Systems Simulator (APSIM) was used for comparison. Field data comprising of weather station data and past irrigation schedules have been used to train the model. Grid search has been used to optimize the hyper-parameters of the model. Nested cross-validation has been used for validating the results. The results show that the correlation-based selected dataset is as useful as the domain expert-recommended dataset in predicting the water requirement using LSTM as the base model. The models were evaluated on different parameters, and a multi-criteria decision evaluation (TOPSIS) was used to find the best performing.

5.1 Introduction

Rice is a major food crop in the world. Despite the water scarcity, its production over the next two decades must increase significantly to meet the challenge of feeding the growing population. The increasing population is exerting enormous pressure on the available freshwater resources of the world. It is leading to an increased water demand from other sectors of economy, such as industrial as well as domestic. Consequently, the agriculture sector, which is the largest user of available water supplies, is expected to utilize it for irrigation. Thus, the future increase in productivity and production of rice largely depends on the availability of irrigation water and its efficient use. Usually, only about 45% of the irrigation water applied to crop is utilized by it. Remaining is lost in conveyance up to the field (15%), in field channels (15%), and due to inefficient irrigation practices (about 25%) [118]. However, land preparation, maintaining the water level in puddled rice fields, and the creation of anaerobic conditions through soil saturation require more water than the evapotranspiration needs of rice. So, for optimal use of depleting water resources, efficient use of water for irrigation to meet the crop water requirement in a manner that crop doesn't experience moisture stress is crucial.

Changing climate is not just a cause of concern, but a reality today. Also, available freshwater resources are scarce. Hence, for sustaining life on Earth, action must be taken to mitigate the impact of both the issues mentioned above. One method is to ensure the adequate irrigation of crops that is adjusted to the variations in the climate. Reducing wastage during irrigation applications can be ensured by using a drip irrigation system. However, to ensure the benefit of its use, an irrigation schedule for a drip irrigation system has to be different from the conventional schedule. Although APSIM is available for generating a drip irrigation schedule, it does not have the adaptability to learn from the schedules created by the domain experts. Hence, an LSTM based model has been used to learn from the pre-existing schedules that are adaptable to climate change. The adaptability comes from the data based learning that is used to develop the model.

In our field experiments, irrigation schedule was being managed by field manager and the total amount of irrigation applied to each plot were decided by fiend manager by visually seeing if irrigation water height is 5 cm or not this added the variability to irrigation amount being applied at each irrigation event and became a challenge for the data-driven model also.

The rest of the paper has been arranged as follows. Section 5.2 provides a brief overview of related work. Section 5.3 gives the introduction to the data set used and its features. Section 5.4 covers the methodology describing the experiments. Section 5 comprises of the results obtained and the related discussion, and Section 6 presents the conclusion, along with the scope of work in the future.

5.2 Related work

The traditionally grown rice in puddled transplanted conditions has a very high irrigation water demand. However, its yield are significantly higher in both lowland and upland conditions as compared to direct seeded rice [143]. The improvement in agronomic practices for cultivation and change in the methods of irrigation can play a crucial role in curtailing the demand. Efforts are being made to make efficient use of available rainfall through the development of short-duration varieties and change in crop calendar, which means a shift in date of transplanting and thereby reduce the water used for rice production. Alternate methods of water application such as alternate wetting and drying, water recycling, conjunctive use, and growing irrigated rice under aerobic conditions have also been recommended for irrigated rice crops [144].

The method, timing, and duration of irrigation have a significant bearing on the efficiency of the irrigation water use in a crop. The information regarding various parameters related to weather, soil, and plant is required for efficient management of crop water requirements. The crop water requirement is a function of evapotranspiration (ET), which is the process of water transfer to the atmosphere by evaporation from land and transpiration from plants. It can be estimated using two different approaches [119]; first, a conceptual approach which is based on various parameters such as soil moisture content (SMC), its depletion, evapotranspiration and the second, a theoretical approach which is based on developing a mathematical model and training it using the available time-series data. As compared to the conceptual approach, the crop water demand estimated using the theoretical approach gives more precise outcomes [119]. Saleem *et al.* [120] developed a mathematical model for predicting the crop irrigation schedules in real-time. They took into account some of the standard operational constraints, such as the practical limits on minimum or maximum amount of water applied and limitations on water availability. In an optimum irrigation schedule, the crop water demand very closely matches the available water supplies.

In irrigated environments, rice is mostly planted in puddled transplanted conditions. Research trials are being conducted by planting rice under a drip irrigation system to improve irrigation efficiency, which permits the raising of the crop following optimal irrigation and fertigation schedules. In drip irrigation system, water is applied drop by drop in the crop root zone, through a network of pipes, valves, and emitters, either laid on the soil surface or below the surface of soil [121].

Crop water requirement may be estimated using direct measurements of leaf water potential of a plant [123]. It is determined by measurement of SMC in root zone of the crop [124] and the reflectance in the near-infrared bands [125].

The SMC or the readily available water can be estimated by calculating the difference between its value at saturation (field capacity) and the wilting point. Sensors may be used to collect information about SMC from the field. The models based on weather parameters have an open-loop structure, whereas the models based on data gathered by soil sensors form a closed-loop [111]. However, the closed-loop systems are now better suited for irrigation scheduling due to the availability of more versatile sensors [63]. Variability in crop characteristics, soil type, change in water requirement of plants over the growth cycle [131], and dissimilarity in methods of irrigation [132] makes the development of models more complex. All these reasons make the results deviate from the ideal. Machine learning techniques can assist the agriculture experts as these have been used in the past to calculate crop parameter values [133], the volumetric soil water content [134], ET [135], total water available in the crop root zone [136], and stem water potential [137]. The modeling techniques developed for such irrigation systems are not practically used due to the non-availability of requisite data and the cost involved in its acquisition. Machine learning models based on data-driven learning, such as artificial neural networks and support vector machines, have been used to overcome these issues[119]. Machine learning techniques are being used for simulation of the time-series nature of variations in SMC. However, as the goal was to determine the SMC values based on historical data, the approach to schedule irrigation based on that was somewhat restrictive [139]. Further, only limited data on the use of irrigation water and its scheduling in drip irrigated rice is available.

Soil Matric Potential (SMP) is a measure of the force with which the soil particles and the pore space hold the water. It indicates the soil moisture availability to plants and can be used for scheduling irrigation. Jimenez *et. al.* collected data from two different soil types were fed into an LSTM neural network model to predict the irrigation schedule [145]. Capraro *et. al,* [146] designed and developed a soil moisture controller to facilitate automatic irrigation that regulates the moisture level in the root zone. It controls the opening and closing of valves based on soil characteristics, the amount of irrigation water supplied, and the consumption by the crop. Multiple linear regressions and NN were trained using the water demand and climatic variables of two years for forecasting the irrigation demand. However, in the existing methods, a significant drawback was the presence of noise in the data. An effective way to reduce this noise content was to pre-process the data before training the model. Thus, Calvo *et. al.* [147] employed the data pre-processing using a smoothening process that significantly improved the results. Along with an extensive survey of various machine learning approaches being used in agriculture, Krupakar *et. al.* [109] also proposed the use of a recurrent neural network (RNN) to model the irrigation requirement of a crop. The advantage of using such an approach is that the model inherently learns from sequential data. Hence, the long input streams do not loose information, and the sequence information gets incorporated in the hidden layer. Furthermore, Internet-of-things (IoT) is coming up as a new sphere of application, and F. Bu and X Wang [148] developed an IoT based irrigation scheduling system using deep reinforced learning and cloud computing and recommended to explore the implementation of LSTM for modeling the irrigation schedules.

5.3 Materials and methods

5.3.1 Description of site

The trial site (29.6857°N and 76.9905 °E) is located at the Central Soil Salinity Research Institute (CSSRI) farm in Karnal in Haryana, India. The site climate is somewhat dry (semi-arid), with an average of 700 mm of precipitation per year. June to September are the months when most of the precipitation (75-80%) is received. The daily minimum temperature in January varies from 0-4$^{\circ}C$, while the highest daily temperature in June ranges from 41-44$^{\circ}C$. The relative humidity throughout the year is in the range of 50-90%. Weather information on site was logged using an automated weather station. All variables were sampled and recorded daily.

5.3.2 Feature measurement

Crop water demand is affected by various parameters. Some of these are soil-based, while others are weather-based, and there are still others that are crop-based. Soil parameters, such as soil structure, texture, fertility, salinity, and drainage, have not been considered as the data were collected from sub-plots of the same farm. Hence, the effect of soil variation remains unmodeled. However, effect of soil parameters has been captured with use of soil temperature at different levels and at different times during the day. Crop parameters such as type of crop, its variety, its height, and the plant density have not been included explicitly. However, number of days since sowing and growth stage of the crop has been used to provide for the effect of cropping stage. There is no other source of variability that may have helped the model to learn better. Thus, other parameters such as type of farm, type of irrigation system, water quality, and crop residue management practices have not been considered. Weather parameters such as rainfall, solar radiation, air temperature, wind speed and direction, evapotranspiration, and air humidity have been used in this study. Their detailed description is as follows:

5.3.2.1 Weather parameters

The weather has a tremendous effect on the water demand of the crop. Various parameters have been considered, some of which are described below.

1. **Temperature**: The temperature of the air affects its water vapor holding capacity and, thereby, the ET. Lower the temperature, lower is the crop water demand. The temperature is measured at a standard height of 2m above ground level, *i.e.* at the canopy level of the crop.
2. **Humidity**: This measures the quantity of water vapor in the air. The degree of air saturation, which is a ratio between the amount of water the ambient air holds and the total amount it could hold at the same temperature, is termed as Relative Humidity (RH). It is measured as the difference between the dry and the wet bulb temperatures.
3. **Wind speed and direction**: These affect the water vapor removal over the evaporating surface. It is recorded at 2m or 3m above the surface *i.e.* at canopy level.
4. **Evapotranspiration**: Process of water transfer to the atmosphere by evaporation from land and transpiration from plants is called evapotranspiration. Higher the evapotranspiration, the higher is the crop water demand.
5. **Rainfall**: It is the amount of rain falling within a given area in a given time. The intensity and amount of rainfall affect the irrigation water

requirement of the crop.

6. **Solar radiation**: More solar radiation means more available energy to vaporize water. Sunshine hours, which reflect this component in our data, affect the incoming solar radiation and hence the crop water demand.

5.3.2.2 Soil parameters

1. **Soil heat flux**: Soil Heat flux is the energy utilized in heating the soil and is usually estimated by measuring the temperature of the soil at three different depths of 5, 10, and 20m below the surface. It is measured at two times during the day, *i.e.* morning (6 A.M.) and afternoon (2 P.M.).
2. **Texture**: The soil type at the experimental site is a reclaimed alkali loam and is characterized by a homogeneous and deep sandy clay loam texture.

5.3.2.3 Crop parameters

1. **Type**: Rice is considered as a cereal grain in developing countries. Because it is a major food supplement, it is commercially grown in many parts of the world as a cash crop.
2. **Variety**: Oryza Sativa, commonly referred to as Asian rice, is the variety under consideration.
3. **Growth stage**: Rice varieties have five main crop stages: germination, panicle initiation, tillering, maturity and harvest. These have been engineered from the date of sowing.

5.3.2.4 Type of irrigation System

In these experiments, the drip irrigation system has been laid with a drip line spacing of 67.5 cm and emitter spacing of 30 cm for each row.

5.3.2.5 Residue management

The burning of Rice crop residue is a major source of environmental pollution, so it was used as a mulch in these experimental plots. The heat from burning causes the soil properties to change and hence alters the composition of soil over time, causing variations in CWD of the crop.

5.3.3 Methodology

The workflow for the methodology is showcased in Figure 5.1. Pseudo-code for the same is written in Algorithm , and the steps are explained in detail in the following subsections.

5.3.3.1 Data collection and pre-processing

In any machine learning approach, the first step is data collection and pre-processing. As weather and irrigation information came from two different sources, sampling intervals did not match such that the information from the weather station and the flow meter had to be compiled into one data collection. It was performed with the advice of a domain expert/agronomist. The information set comprising of a one-dimensional array with all the values corresponding to the one-day input variables was prepared. Five sub-sets of information were developed (corresponding to 2014, 2015, 2016, 2017, and 2018, with a fixed season duration of 112 days each) and compiled in one. The data set contained three trial plot replications for the same set of input parameters with different amounts of water as CWR for each plot. The average of these 3 values was taken as target and standard deviation as the range of acceptable error. Wind direction was a categorical variable with distinct wind directions. One-hot encoding has been implemented to transform the categorical variable into a boolean one. The input features used in the datasets are described in Table 5.1. In this table, Julian date is the count of the number of days since January 1, 4713 BC. It has been used to reflect the shift in the dates of the crop cycle.

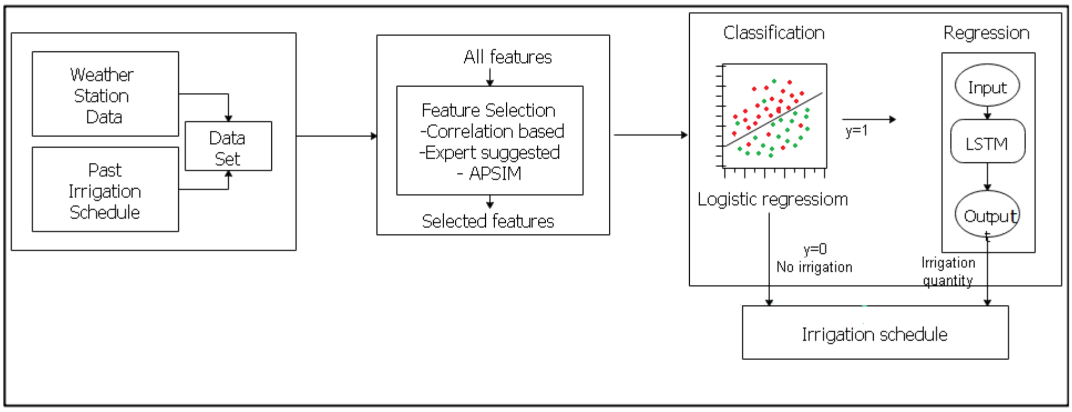


Figure 5.1: Methodology proposed for development of a multi-level model

Algorithm 5.1 Methodology

```
Input :
$f_w$ = [ $T_{Min}$, $T_{Max}$, $Dry_B$, $RH_{Max}$, $RH_{Min}$, $RH_{Avg}$, $T_{Soil}$, $S_{Hrs}$, $W_S$, $W_D$, ET, R, $Y_r$, E]
Output :
S [day, status, quanity]
Step 1: Data set:
    i. Past weather data {$W_P$}
   ii. Past irrigation schedule {$S_P$}
  iii. Current weather data {$W_C$}
Step 2: $f_w \Longleftrightarrow$ Mapping$S$
Step 3: Pre-process the data to make it suitable for machine learning;
(i) Feature engineering based on auxiliary information available in the dataset
    • {$J_D$, $D_{SS}$, $S_G$ } ← Date of sowing
    • $D_{LR}$ ← R
          $D_{LR} = \begin{cases} 0, & if R \neq 0 \\ Count of days since R \neq 0, & otherwise. \end{cases}$
    • {$Y_B$, $D_{LI}$ } ← $Y_R$
          $Y_B = \begin{cases} 0, & if Y_R \neq 0 \\ 1 = 0, & otherwise. \end{cases}$
          $D_{LI} = \begin{cases} 0, & if Y_B = 1 \\ Count of days since Y_B \neq 1, & otherwise. \end{cases}$
(ii) One-hot encoding on $W_D$
(iii) Normalise the data between {0,1}
Step 4: Train/Test a machine learning model
for each day of the cropping season: do
    Predict irrigation event
    if irrigation is required then
        Predict amount of water
    else
        No irrigation required
    end if
end for
Step 5: Repeat Steps 1 - 4, k times (k-fold cross validation)
```

* Refer to Table 5.1

5.3.3.2 Feature selection

The next stage was the selection of features that were performed using correlation. As mentioned in Section 4.3.2 of the previous chapter, Pearson Correlation Coefficients were used to create a matrix representing the data dependencies. A threshold of 0.75 was used to obtain a subset of features by dropping those

Table 5.1: Data set features and its subsets

Code	**Symbol**	**Name**	**Parameters selected based on correlation**	**Parameters suggested by domain expert**	**Parameters used by APSIM**
V01	J_d	Julian date	Yes	Yes	Yes
V02	D_{SS}	Days since start of season	Yes	Yes	No
V03	S_G	Stage of growth	No	Yes	No
V04	T_{Min}	Maximum temperature	Yes	Yes	Yes
V05	T_{Max}	Minimum temperature	No	Yes	Yes
V06	Dry_B	Dry bulb at 2:00 P.M.	Yes	No	No
V07	RH_{Max}	Relative humidity at 6:00 A.M.	Yes	Yes	No
V08	RH_{Min}	Relative humidity at 2:00 P.M.	No	Yes	No
V09	RH_{Avg}	Average Relative humidity	No	No	Yes
V10	T_{Soil}	Soil temperature	Yes	No	No
V11	S_{Hrs}	Sunshine hours	Yes	Yes	Yes
V21	W_S	Wind speed	Yes	Yes	Yes
V13	W_D	Wind direction	Yes	Yes	No
V14	ET	Evapotranspiration	Yes	Yes	No
V15	R	Rainfall	Yes	Yes	Yes
V16	D_{LR}	No. of days since last rain	Yes	Yes	No
V17	D_{LI}	No. of days since last irrigation	Yes	Yes	No
V18	Y_b	Event of irrigation	Yes	Yes	No
V19	Y_r	Amount of irrigation	Yes	Yes	Yes
V20	E	Range of acceptable error	Yes	Yes	No

variables with values above it to ignore those having very positive and strong co-relations (correlation-based dataset). Also, a data set (Expert suggested dataset) containing features recommended by a domain expert (agricultural scientist) was used.

5.3.3.3 Model selection

Classification: At first, the need for an irrigation event was assessed using some simple classification algorithms. Here, F18 was considered to be an independent variable, and remaining (except F19 and F20) were considered as dependent variables. Different classification models were used during the model training, and their details are given below.

1. Logistic regression (LoR): It is a predictive analysis algorithm. It is used to describe and explain data as a function of one or more independent variables that can be a nominal, ordinal, interval, or ratio. It can manage input that can either be dense or be sparse.

2. Support Vector Machine (SVM): An SVM is a discriminatory classifier that defines a hyperplane. For a given labeled data set, an optimal hyperplane is produced by the algorithm, which is later used for new samples.

3. Decision Tree (DT): A DT is a flowchart in which a feature is represented by an internal node, a decision rule by a branch, and a result by each leaf node. The default values for the parameters lead to fully grown and unpruned trees. Some appropriately specific values of the model parameters

should be set to reduce the complexity and size of the trees.

4. Random Forest (RF): A large number of individual decision trees that operate effectively as an ensemble comprises of an RF. Each tree predicts a class forecast; the class with the maximum number of votes becomes the prediction.

LSTM for regression: A time-series compatible non-linear regression model was chosen to predict the amount of water required to meet the crop demand. LSTM is a deep learning neural network that is appropriate for use in learning from sequential information as it is robust against long-term dependency issues. The historical aspect of data is stored inside the input vector. An input layer, a hidden layer, and an output layer usually comprise of the architecture. The input layer contains nodes equal to the number of input variables, and the output layer contains only one output node. The weight vectors of the layer connections are recalculated at every iteration of backpropagation. Here, feature F18 was disregarded, and F19 was considered as the independent variable. Keeping the dependent variables the same, F20 gave the acceptable range of error. Data collected by sensors is configured as an input vector containing 7-day data. The objective of the training is to correctly predict the amount of water to be applied to the plant. Precision and loss were accumulated to monitor the progress of the training. It has was experimentally determined that 1000 iterations were sufficient to obtain acceptable precision. The model matched the target's actual values with those predicted to help optimize the parameters.

5.3.3.4 Data splitting

In this work, both classification and regression estimators have been used to predict the irrigation schedule for rice. The training set and testing set have been created here by dividing the data here into 60% and 40% partitions, respectively. Given the inherent time series-ness of the data, the training set of data was kept in a sequence; the nested cross-validation technique was used to perform the cross-validation.

5.3.4 Hyper-parameter optimization

Out of the various parameters used for LSTM, only five were optimized such as number of epochs (number of iterations), batch size (number of samples in each iteration), activation function (defines the transformation on the input), optimizer (responsible for governing the weights and learning rate), and dropout rate (controls overfitting). The time complexity of an exhaustive grid search on the limited computing resources available was reduced by adopting a stage-wise approach. Number of epochs and batch size were optimized in one stage due to their interdependence while the rest of the parameters were done one by one.

5.3.5 Performance evaluation

The performance of the model can be measured in numerous ways. Their relative suitability depends upon the considered application. The models have been evaluated against error, correlation, and accuracy with respect to a range. The performance of the predictive models was measured using three-fold nested cross-validation.

5.3.5.1 For classification

1. Sensitivity: Also known as recall or true positive rate, it is the proportion of actual positive cases that are correctly identified. In simpler words, it is a measure of how often true were predicted as true.

   $$Sensitivity = \frac{TP}{TP + FN} \tag{5.1}$$

   In the above equation, TP refers to True Positive, and FN refers to False Negative.

2. Specificity: It is the proportion of actual false cases that are correctly identified.

   $$Specificity = \frac{TN}{TN + FP} \tag{5.2}$$

   In the above equation, TN refers to True Negative, and FP refers to False Positive.

3. Area Under the Curve (AUC): The Receiver Operating Characteristic (ROC) curve is the plot between sensitivity and 1-specificity. To change the curve to a number, AUC is calculated. The ranges and their equivalent interpretations are as follows: 1-0.90 = Excellent; 0.90-0.80 = Good; 0.80-0.70 = Fair; 0.70-0.60 = Poor; 0.60-0.50 = Fail.

4. Accuracy: The most common evaluation metric, accuracy is the number of correct predictions made to the ratio of all the predictions made.

$$Accuracy = \frac{TP + TN}{TotalData} * 100 \tag{5.3}$$

   In the above equation, Total Data is the sum of TP, TN, FP, and FN.

5. F1 Score: It is a measure of accuracy that takes into account the balance between precision and recall.

$$F1Score = \frac{Precision * Recall}{Precision + Recall} * 2 \tag{5.4}$$

5.3.5.2 For regression

Of the evaluation parameters described in Subsection 6, Section 3 of Chapter 4, Mean Absolute Error, Coefficient of Co-relation and accuracy w.r.t. a range were used.

The Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS): To make a multi-criteria decision based on the multiple evaluation parameters, TOPSIS was used to help make the correct choice. A cumulative score and ranking of the various approaches were obtained, and the highest-ranking was chosen as the most suitable.

5.4 Results and discussion

5.4.1 Results

In this study, a two-level ensemble model has been developed. At first, only an LSTM was used for prediction. However, the performance of the model was quite low. Hence, a hybrid model with an additional stage of classification to predict the occurrence of an irrigation event was added. Once the need for irrigation is established at the first level, then the second level method uses regression to predict the amount of water needed to be applied to the crop. Since all the classification models performed very quite well, the models were evaluated against precision, recall, AUC, accuracy, and F1 score. The LSTM has been evaluated against Mean Absolute Error, coefficient of correlation between the actual and predicted value, and accuracy with respect to a range. Three-fold cross-validation was used to measure the robustness of the model.

Table 5.2: Performance of LSTM neural network model

Model		**LSTM**									
Data set		**Expert suggested**					**Correlation based selected**				
Batch Size	**Epochs**	**MAE**	R^2	**Accuracy**	**Topsis Score**	**Rank**	**MAE**	R^2	**Accuracy**	**Topsis Score**	**Rank**
112	1000	20.79	0.16	53.27	0.04	12	20.2	0.16	52.23	0.33	11
	750	20.84	0.16	53.57	0.31	8	20.63	0.18	50.86	0.48	9
	500	21.48	0.17	54.46	0.60	6	21.52	0.18	47.32	0.62	5
	250	22.45	0.17	52.68	0.76	3	21.56	0.19	45.98	0.74	3
7	1000	17.47	0.23	57.79	0.28	10	18.46	0.2	57.59	0.42	10
	750	18.1	0.21	59.38	0.40	7	18.21	0.22	54.91	0.57	6
	500	17.46	0.22	61.61	0.69	4	19.88	0.2	54.02	0.69	4
	250	***18.63***	***0.2***	***56.7***	***0.86***	***1***	***18.08***	***0.2***	***56.25***	***0.80***	***1***
1	1000	19.56	0.22	56.7	0.26	11	17.26	0.26	60.71	0.52	7
	750	19.94	0.15	54.02	0.30	9	22.13	0.05	50.54	0.22	12
	500	18.96	0.17	59.82	0.61	5	22.29	0.12	51.34	0.49	8
	250	19.13	0.17	59.38	0.77	2	20.25	0.18	53.13	0.74	2

Table 5.3: Performance of classification models

Data set	**Evaluation Parameter**	**Logistic Regression**	**Support Vector**	**Decision Tree**	**Random Forest**
Expert selected	f1 score	**0.97**	0	0.96	0.93
	Precision	**0.96**	0	0.95	0.88
	Recall	**0.85**	0	0.9	0.87
	AUC	**0.91**	0.5	0.88	0.94
	Accuracy	**100**	71.88	100	96.43
	Topsis score	**0.61**	0.4	0.6	0.59
	Rank	**1**	4	2	3
Correlation based selected	f1 score	**0.93**	0	0.98	0.97
	Precision	**0.82**	0	0.81	0.82
	Recall	**0.89**	0	0.97	0.94
	AUC	**0.82**	0.5	0.98	0.97
	Accuracy	**100**	71.88	99.11	98.21
	Topsis score	**0.6**	0.4	0.56	0.58
	Rank	**1**	4	3	2

Table 5.2 shows the performance of an LSTM on an Expert selected and correlation-based selected data set. Different batch sizes (Daily - value count 1, Weekly - value count 7 and Season length - value count 112) and different values of epochs (*i.e.* 250, 500, 750 and 1000) were also optimized to find the most satisfactory result. It was found that for both the data set, the optimum batch size was found to be 7 with epoch size equal to 250.

In Table 5.3, we show the performance of classical machine learning models for the classification stage of the model on two different data sets *i.e.* Expert selected and correlation-based selected. Four models, namely, Logistic regression, Support Vector, Decision Tree, and Random Forest), were evaluated on five evaluation parameters,*i.e.* F1score, Precision, Recall, AUC and Accuracy. Furthermore, TOPSIS was used to make a multi-criteria decision. Based on the

TOPSIS score, it can be seen that for Expert selected data set, logistic regression performs best, followed by decision Tree and random forest. Whereas for the correlation-based dataset as well, logistic Regression performs well, followed by random forest and decision tree.

Table 5.4: Performance of Multi-level model

Model		**Logistic Regression with LSTM**									
Data set		**Expert selected**					**Correlation based selected**				
Batch Size	**Epochs**	**MAE**	R^2	**Accuracy**	**Topsis Score**	**Rank**	**MAE**	R^2	**Accuracy**	**Topsis Score**	**Rank**
112	1000	9.39	0.67	74.55	0.37	7	9.78	0.62	75.45	0.35	7
	750	9.11	0.67	75.45	0.68	4	9.48	0.64	75.89	0.67	4
	500	**9.02**	**0.68**	**74.55**	**0.98**	**1**	**9.7**	**0.63**	**75**	**0.97**	**1**
	250	10.43	0.6	75.45	0.10	12	10.13	0.63	76.79	0.13	10
7	1000	10.36	0.58	73.21	0.33	9	9.97	0.61	75	0.35	8
	750	10.63	0.58	75.45	0.65	6	9.61	0.63	76.34	0.67	5
	500	10.49	0.57	74.55	0.86	2	10.61	0.59	73.66	0.92	3
	250	*9.85*	*0.63*	*74.55*	*0.13*	*11*	*10.1*	*0.61*	*75.45*	*0.11*	*11.5*
1	1000	10.09	0.63	73.21	0.35	8	10.8	0.53	75.45	0.33	9
	750	10.15	0.58	76.34	0.65	5	10.37	0.55	76.34	0.65	6
	500	10.72	0.53	74.55	0.82	3	9.66	0.62	75.89	0.97	2
	250	19.13	0.17	59.38	0.77	2	20.25	0.18	53.13	0.74	2

Combining the classification and LSTM as a multi-level model, Table 5.4 shows the results obtained. As stated before, different batch sizes with different epochs were used to find the optimal performance value. It was found to be a batch size of 112 and 500 epochs. Both the data sets performed equally well.

On close examination, it can be observed that both the datasets *i.e.* expert suggested as well as selection based on correlation performed equally well and are comparable in performance. This can be indicative that both the datasets are interchangeably usable. However, We chose the correlation-based selected dataset for validation as the choice of input parameters is in-line with the data-based approach.

Figure 5.2 shows that the hybrid LSTM performs well during validation. Figure 5.2(a) shows that R^2 increases when the model is trained on larger time-series data. Figure 5.2(b) shows that during three-fold cross-validation, the MAE reduces from 12.6 to 8.0 as additional data is used for training the model. Figure 5.2(c) indicates that accuracy also improves as the model is trained on a larger dataset. Hence, it can be concluded that as the amount of data being used to train the model increases, the performance of the proposed approach improves.

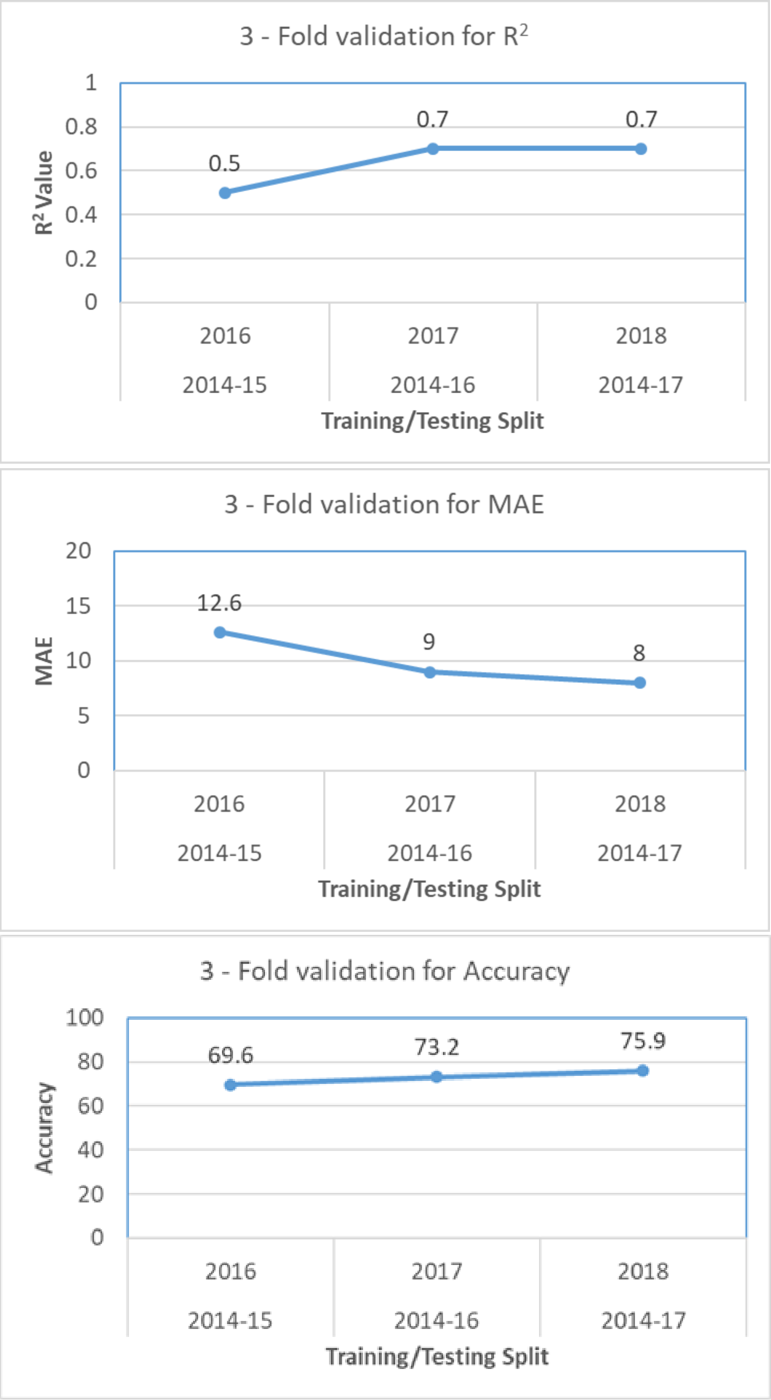


Figure 5.2: Validation of Multi-level model on the Expert-selected data set

5.4.2 Comparison with APSIM

The APSIM is a tool for modeling and simulating agricultural systems. It contains a suite of modules that allow the simulation of systems for a variety of plant, soil, climate, and management interactions. The APSIM irrigation module allows the user to specify a multi-year irrigation schedule and to configure an automatic irrigation schedule calculated based on soil moisture.

Table 5.5: Comparison of Multi-level model with APSIM for predicting the occurrence of an irrigation event

Evaluation Parameter	**Proposed approach**	**APSIM**
Senstivity	0.79	0.43
Specificity	1	0.77
Accuracy	94.19	67.41
F1 Score	0.88	0.43

Table 5.6: Comparison of Multi-level model with APSIM for predicting the amount of irrigation

Evaluation Parameter	**Proposed approach**	**APSIM**
Correlation	0.59	0.05
MAE	10.27	19.75
Accuracy	75.9	58.93

Table 5.5 and Table 5.6 show the performance of our proposed approach as compared with APSIM. From Table 5.5, it can be observed that our proposed approach performs better than APSIM on all parameters. Our approach can correctly predict the absence of an irrigation event every single time while predicting reasonably well the need for an irrigation event as well. Also, Table 5.6 shows similar results. Our proposed technique is better at predicting the irrigation quantities closer to the original value than APSIM. This may be because APSIM, while capable of predicting irrigation schedules for subsurface drip-irrigated crops, does not self learn from the data. It instead runs on predefined set rules of irrigation scheduling. Since our approach is based on learning from field data, it is capable of capturing the interaction between the independent variables better and hence perform well in predicting the irrigation schedule.

5.4.3 Discussion

For machine learning approaches, it is important to compare the results of one with the other to find the most appropriate approach. There are various ways to do this, both for classification and regression problems. Some standard classification evaluation parameters are the confusion matrix, F1 Score, Sensitivity, Specificity, and Accuracy. For regression, the standard evaluation parameters used are correlation between Actual and predicted values, accuracy

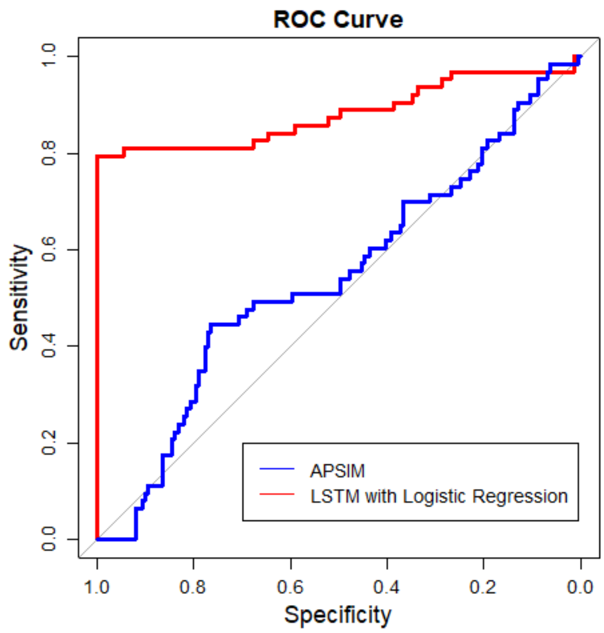


Figure 5.3: ROC curve for Multi-level model and APSIM

with respect to an acceptable range, and some measure of error. These have been stated in the preceding sub-section on the comparison. However, a graphical representation of the results makes it easier to understand and process the results.

Receiver Operating Characteristic (ROC) curve is one of the most popular graphical representation used for the evaluation of a two-class classification problem. It is a plot between true positive rate and false positive rate as the threshold varies from $-\infty$ to ∞. We can observe in Figure 5.3 that the ROC curve of the hybridized LSTM lies above that of APSIM. Hence, it can be concluded that for any arbitrary value of the threshold, hybrid LSTM performed better for this dataset in terms of both false negatives and false positives than APSIM. The diagonal line represents the performance of a trivial classifier, that randomly allocates classes to the objects.

Figure 5.4 shows how well the respective models fit the regression values. A diagonal line is desirable as it shows a high correlation between the actual and the predicted values. As can be seen, LSTM with Logistic Regression performs better than APSIM. This means that our model can predict better, not only the frequency but also the amount of irrigation to a nearer value of the actual.

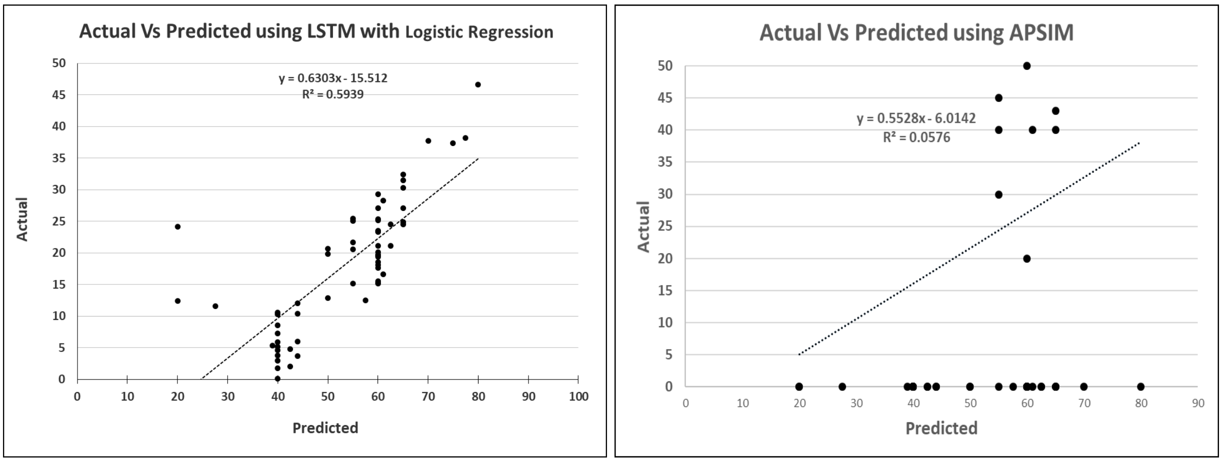


Figure 5.4: Comparison of actual and predicted value for Multi-level model and APSIM

5.5 Conclusion

The proposed approach was used to correctly predict the timing and amount of water for an irrigation event. The classification allows for the model to know to weather the crop requires water or not, while the regression stage calculates the amount of water required. The logistic regression was able to predict whether the irrigation is required or not correctly. The LSTM can estimate the required amount of water by capturing the intrinsic variations in the observations. However, the availability of more data that not only captures the dynamic interaction of the input variables but also contains more number of affecting input variables can help improve the performance of the approach.

Chapter 6

Conclusions and scope for future work

With the increasing water scarcity in the region, particularly Indo-Gangetic plains (IGP), agricultural production, even at the current level, might not be sustainable. So, significant efforts are needed to conserve water by making the best use of every drop to produce more with it. Bundling new cutting-edge technologies such as micro-irrigation, and especially drip irrigation, coupled with automation in water application can help to increase the application efficiency and water productivity, especially in high water-consuming, cereal-based systems.

The present study has been carried out to develop a multi-level model for forecasting crop water demand, which may be used as an input to an automated irrigation system to plan the efficient use of water in agriculture.

6.1 Conclusion

Earlier studies have shown that the recent advances in the development of WSNs have introduced a new dimension into the automation of drip irrigation, promoting the precision application of water and nutrients simultaneously. The introduction of WSNs and Cloud-based technologies has led to an opportunity where on-field data can be captured for analysis and learning. The previously existing Threshold-based algorithms and fuzzy logic-based irrigation scheduling techniques are limited to perform well in the context of the current existing scenario. However, these are not adaptive to the rapidly changing climate conditions. Hence, machine learning and its aides can be explored to not only calculate the irrigation demands of various crops better but also help design a forecasting system that is inherently adaptive to climate change. This is ensured, as the models learn from the changing climatic parameter values and are hence make better predictions, in this case. The availability of more data that not only captures the dynamic interaction of the input variables but also contains a higher number of affecting input variables can help improve the performance of the approach.

The performance of the experimental models is stuck at a bottleneck of about 70-75% accuracy because the enormous amount of data that is required to train/test/validate the models is often not available in the required quantity. Also, in many cases, the data for some influential parameters are not recorded. Add to this the limitations of the experimental site and experiment design. This is a field-level experiment; it has been upscaled by removing the dependence on soil characteristics. An agronomist's irrigation schedule is based on tensiometer readings, which are a measure of soil metric potential. However, this very parameter has been excluded during model development and has been indirectly inference from air and soil temperature. It can be concluded that soil parameters are a vital input to assess the crop water demand. Ignoring soil parameters can make the experiment widely applicable as weather data is more readily available from public weather stations as compared to soil parameters, but it reduces the forecasting accuracy.

6.2 Scope for future work

Research is an iterative and continuous procedure. The work presented in this thesis focused on the efficient prediction of crop water demand primarily using climatological data. Some of the suggestions for future work are:

Weather parameters are only a partial component of calculating crop water demand. On a field level, crop parameters such as crop variety and date of sowing are factors that influence the demand. Whereas on a broader scale, the dependence on soil parameters and the effect of soil composition and type cannot be ignored. Hence, data for such parameters need to be included in the dataset for holistic model development. As future work, it is recommended that additional information for other influencing parameters be modeled into the approach.

Further, to optimize the permissible benefits to the users and address their legitimate concerns, additional research and development efforts are required to standardize WSN protocols and communication frequencies. The calibration of sensors and equipment, as well as their maintenance, needs regular monitoring. An approach to automate the process and alert the need for the same is required.

New technological advances, viz. sensor-cloud computing, big-data analytics, and the Internet of Things (IoT), can provide flexible control mechanisms in real-time for on-field parameters. These may be explored to develop a standardized framework for futuristic applications in automated drip irrigation systems.

Subsurface drip is laid below the surface of the soil. The evapotranspiration (ET) component of the CWD gets reduced in this case. This research may be extended to forecast for the crop water requirement in the subsurface drip. This work may be extended not only to other crops but to encompass various crop rotations as well.

References

[1] J Diouf. How to feed the world in 2050. *Population & Development Review*, 35(1):837–839, 2009.

[2] Global food policy report. Technical report, International Food Policy Research Institute, 2019.

[3] SK Sun, C Li, PT Wu, XN Zhao, and YB Wang. Evaluation of agricultural water demand under future climate change scenarios in the loess plateau of northern shaanxi, china. *Ecological indicators*, 84(1):811–819, 2018.

[4] G Lacombe, P Chinnasamy, and A Nicol. Review of climate change science, knowledge and impacts on water resources in south asia. *Journal of Irrigation and Drainage Engineering*, 61(2):22–31, 2019.

[5] FAO. Aquastat database, 2016. Available at `www.faodata`.

[6] DS Gaydon, E Wang, PL Poulton, B Ahmad, F Ahmed, S Akhter, I Ali, R Amarasingha, AK Chaki, C Chen, et al. Evaluation of the apsim model in cropping systems of asia. *Field Crops Research*, 204:52–75, 2017.

[7] UN. World population prospects 2019. 2019.

[8] RK Jat, TB Sapkota, RG Singh, ML Jat, M Kumar, and RK Gupta. Seven years of conservation agriculture in a rice–wheat rotation of eastern gangetic plains of south asia: Yield trends and economic profitability. *Field Crops Research*, 164(1):199–210, 2014.

[9] ML Jat, JC Dagar, TB Sapkota, B Govaerts, SL Ridaura, YS Saharawat, RK Sharma, JP Tetarwal, RK Jat, and H Hobbs. Climate change and agriculture: Adaptation strategies and mitigation opportunities for food security in south asia and latin america. In *Advances in Agronomy*, volume 137, pages 127–235. Elsevier, 2016.

[10] SS Kukal, ML Jat, and HS Sidhu. Improving water productivity of wheat-based cropping systems in south asia for sustained productivity. In *Advances in Agronomy*, volume 127, pages 157–258. Elsevier, 2014.

[11] G Fischer, FN Tubiello, HV Velthuizen, and DA Wiberg. Climate change impacts on irrigation water requirements: Effects of mitigation, 1990 - 2080. *Technological Forecasting and Social Change*, 74(7):1083–1107, 2007.

[12] Intended nationally determined contributions: Template for vulnerability

assessment, mitigation and adaptation. indc- vulnerability report. Technical report, National Commission for Integrated Water Resource Development (NCIWRD). Government of India Ministry of Water Resources, River Development and Ganga Rejuvenation National Water Mission, 2015.

[13] X Cai and MW Rosegrant. Global water demand and supply projections: Part 1. a modeling approach. *Water International*, 27(2):159–169, 2002.

[14] HS Jat, PC Sharma, A Datta, M Choudhary, SK Kakraliya, HS Sidhu, B Gerard, and ML Jat. Re-designing irrigated intensive cereal systems through bundling precision agronomic innovations for transitioning towards agricultural sustainability in north-west india. *Scientific reports*, 9(1):1–14, 2019.

[15] SS Kukal and GC Aggarwal. Puddling depth and intensity effects in rice–wheat system on a sandy loam soil: I. development of subsurface compaction. *Soil and Tillage Research*, 72(1):1–8, 2003.

[16] KM Choudhary, HS Jat, DP Nandal, DK Bishnoi, JM Sutaliya, M Choudhary, PC Sharma, and ML Jat. Evaluating alternatives to rice-wheat system in western indo-gangetic plains: Crop yields, water productivity and economic profitability. *Field Crops Research*, 218(1):1–10, 2018.

[17] KK Sarker, A Hossain, J Timsina, SK Biswas, SL Malone, MK Alam, HW Loescher, and M Bazzaz. Alternate furrow irrigation can maintain grain yield and nutrient content, and increase crop water productivity in dry season maize in sub-tropical climate of south asia. *Agricultural Water Management*, 238(1):106229, 2020.

[18] P Soman, MS Prasad, VR Balasubramaniam, S Singh, C Dhavarajan, VB Patil, and S Jha. Effect of drip irrigation and fertigation on the performance of several rice cultivars in different rice ecosystems in india. *International Journal of Agriculture Sciences*, 1(1):975–3710, 2018.

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

[20] KK Bandyopadhyay, AK Misra, PK Ghosh, KM Hati, KG Mandal, and M Moahnty. Effect of irrigation and nitrogen application methods on input use efficiency of wheat under limited water supply in a vertisol of central india. *Irrigation Science*, 28(4):285–299, 2010.

[21] T Parthasarathi, K Vanitha, S Mohandass, S Senthilvel, and E Vered. Effects of impulse drip irrigation systems on physiology of aerobic rice.

Indian Journal of Plant Physiology, 20(1):50–56, 2015.

[22] R Sharda, G Mahajan, M Siag, A Singh, and BSh Chauhan. Performance of drip-irrigated dry-seeded rice (oryza sativa l.) in south asia. *Paddy and water environment*, 15(1):93–100, 2017.

[23] T Parthasarathi, K Vanitha, S Mohandass, and E Vered. Evaluation of drip irrigation system for water productivity and yield of rice. *Agronomy Journal*, 110(6):2378–2389, 2018.

[24] HS Sidhu, ML Jat, Y Singh, RK Sidhu, N Gupta, P Singh, P Singh, HS Jat, and B Gerard. Sub-surface drip fertigation with conservation agriculture in a rice-wheat system: A breakthrough for addressing water and nitrogen use efficiency. *Agricultural Water Management*, 216(1):273–283, 2019.

[25] JE Ayars, CJ Phene, RB Hutmacher, KR Davis, RA Schoneman, SS Vail, and RM Mead. Subsurface drip irrigation of row crops: A review of 15 years of research at the water management research laboratory. *Agricultural water management*, 42(1):1–27, 1999.

[26] CJ Phene, KR Davis, RB Hutmacher, BB Yosef, DW Meek, and J Misaki. Effect of high frequency surface and subsurface drip irrigation on root distribution of sweet corn. *Irrigation Science*, 12(3):135–140, 1991.

[27] CR Camp, PJ Bauer, and PG Hunt. Subsurface drip irrigation lateral spacing and management for cotton in the southeastern coastal plain. *Transactions of the ASAE*, 40(4):993–999, 1997.

[28] CR Camp. Subsurface drip irrigation: A review. *Transactions of the ASAE*, 41(5):1353, 1998.

[29] FR Lamm and TP Trooien. Subsurface drip irrigation for corn production: A review of 10 years of research in kansas. *Irrigation Science*, 22(4):195–200, 2003.

[30] FR Lamm and CR Camp. Subsurface drip irrigation. In *Developments in Agricultural Engineering*, volume 13, pages 473–551. Elsevier, 2007.

[31] CR Camp, PJ Bauer, and WJ Busscher. Evaluation of no-tillage crop production with subsurface drip irrigation on soils with compacted layers. *Transactions of the ASAE*, 42(4):911, 1999.

[32] Santiago Lopez-Ridaura, Romain Frelat, Mark T van Wijk, Diego Valbuena, Timothy J Krupnik, and ML Jat. Climate smart agriculture, farm household typologies and food security: an ex-ante assessment from eastern india. *Agricultural systems*, 159:57–68, 2018.

[33] DOAC. Agricultural statistics at a glance, 2017.

[34] SS Prihar and BS Sandhu. Irrigation of field crops. principles and practices. *Irrigation of field crops: Principles and Practices.*, 1987.

[35] RG Allen, LS Pereira, D Raes, and M Smith. Crop evapotranspiration: Fao irrigation and drainage paper 56. *FAO, Rome, Italy*, 20(1), 1998.

[36] SL Davis and MD Dukes. Irrigation scheduling performance by evapotranspiration-based controllers. *Agricultural Water Management*, 98(1):19–28, 2010.

[37] SS Prihar. Scheduling irrigations to wheat, using pan evaporation. *Indian Journal of Agricultural Science*, 44(1):567–571, 1974.

[38] SS Prihar, PR Gajri, and RS Narang. Permissible profile-water depletions for optimum wheat production on two different soils. *Journal of the Indian Society of Soil Science*, 26(1):7–11, 1978.

[39] SS Prihar, KL Khera, KS Sandhu, and BS Sandhu. Comparison of irrigation schedules based on pan evaporation and growth stages in winter wheat 1. *Agronomy Journal*, 68(4):650–653, 1976.

[40] SK Kakraliya, HS Jat, I Singh, TB Sapkota, LK Singh, JM Sutaliya, PC Sharma, RD Jat, and M Choudhary. Performance of portfolios of climate smart agriculture practices in a rice-wheat system of western indo-gangetic plains. *Agricultural Water Management*, 202(1):122–133, 2018.

[41] SS Kukal, GS Hira, and AS Sidhu. Soil matric potential-based irrigation scheduling to rice (oryza sativa). *Irrigation Science*, 23(4):153–159, 2005.

[42] E Humphreys, SS Kukal, and PL Eberbach. Effects of tillage and mulch on the growth, yield and irrigation water productivity of a dry seeded rice-wheat cropping system in north-west india. *Field Crops Research*, 196(1):219–236, 2016.

[43] SA O'Shaughnessy and SR Evett. Canopy temperature based system effectively schedules and controls center pivot irrigation of cotton. *Agricultural Water Management*, 97(9):1310–1316, 2010.

[44] SB Idso, RD Jackson, PJ Pinter, RJ Reginato, and JL Hatfield. Normalizing the stress-degree-day parameter for environmental variability. *Agricultural Meteorology*, 24(1):45–55, 1981.

[45] RD Jackson, WP Kustas, and BJ Choudhury. A reexamination of the crop water stress index. *Irrigation Science*, 9(4):309–317, 1988.

[46] X Wang, W Yang, A Wheaton, N Cooley, and B Moran. Efficient registration of optical and ir images for automatic plant water stress assessment. *Computers and Electronics in Agriculture*, 74(2):230–237, 2010.

[47] DK Das, G Singh, and AK Sutradhar. Remote sensing of wheat grown

under differential irrigation, row spacings and nitrogen levels. In *6th Asian Conference on Remote Sensing, Hyderabad, India*, pages 400–405, 1986.

[48] N K Gontia and K N Tiwari. Development of crop water stress index of wheat crop for scheduling irrigation using infrared thermometry. *Agricultural water management*, 95(10):1144–1152, 2008.

[49] H Zhang, X Wang, M You, and C Liu. Water-yield relations and water-use efficiency of winter wheat in the north china plain. *Irrigation Science*, 19(1):37–45, 1999.

[50] JE Fernandez, AP Martin, JMT Ruiz, MV Cuevas, CMR Dominguez, SE Farag, AM Sillero, JM Garcia, VH Santana, and AD Espejo. A regulated deficit irrigation strategy for hedgerow olive orchards with high plant density. *Plant and Soil*, 372(2):279–295, 2013.

[51] SH Ahmadi, MN Andersen, F Plauborg, RT Poulsen, CR Jensen, AR Sepaskhah, and S Hansen. Effects of irrigation strategies and soils on field grown potatoes: Yield and water productivity. *Agricultural Water Management*, 97(11):1923–1930, 2010.

[52] E Fereres and MA Soriano. Deficit irrigation for reducing agricultural water use. *Journal of experimental botany*, 58(2):147–159, 2007.

[53] MH Ali, MR Hoque, AA Hassan, and A Khair. Effects of deficit irrigation on yield, water productivity, and economic returns of wheat. *Agricultural water management*, 92(3):151–161, 2007.

[54] BS Soundharajan and KP Sudheer. Deficit irrigation management for rice using crop growth simulation model in an optimization framework. *Paddy and Water Environment*, 7(2):135–149, 2009.

[55] J Kar, H Bremer, JR Drummond, YJ Rochon, DBA Jones, F Nichitiu, J Zou, J Liu, JC Gille, and DP Edwards. Evidence of vertical transport of carbon monoxide from measurements of pollution in the troposphere (mopitt). *Geophysical Research Letters*, 31(23), 2004.

[56] MH Ali and MSU Talukder. Increasing water productivity in crop production - a synthesis. *Agricultural Water Management*, 95(11):1201–1213, 2008.

[57] JB Zhou, CY Wang, H Zhang, F Dong, XF Zheng, W Gale, and SX Li. Effect of water saving management practices and nitrogen fertilizer rate on crop yield and water use efficiency in a winter wheat - summer maize cropping system. *Field Crops Research*, 122(2):157–163, 2011.

[58] AR Sepaskhah and SH Ahmadi. A review on partial root-zone drying

irrigation. *International Journal of Plant Production*, 4(4):241–258, 2012.

[59] SZ Kang, XT Hu, HJ Cai, and SY Feng. New ideas and development tendency of theory for water saving in modern agriculture and ecology. *Journal of Hydraulic Engineering*, 35(12):1–7, 2004.

[60] D Grimes, V Walhood, and W Dickens. Alternate-furrow irrigation for san joaquin valley cotton. *California Agriculture*, 22(5):4–6, 1968.

[61] AK Misra. Climate change and challenges of water and food security. *International Journal of Sustainable Built Environment*, 3(1):153–165, 2014.

[62] B Kayatz, F Harris, J Hillier, T Adhya, C Dalin, D Nayak, RF Green, P Smith, and AD Dangour. "more crop per drop": Exploring india's cereal water use since. *Science of the total environment*, 673(1):207–217, 2019.

[63] T Ojha, S Misra, and NS Raghuwanshi. Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges. *Computers and Electronics in Agriculture*, 118(1):66–84, 2015.

[64] PK Jha, SN Kumar, and AVM Ines. Responses of soybean to water stress and supplemental irrigation in upper indo-gangetic plain: Field experiment and modeling approach. *Field Crops Research*, 219(1):76–86, 2018.

[65] JM Blonquist, SB Jones, and DA Robinson. Precise irrigation scheduling for turfgrass using a subsurface electromagnetic soil moisture sensor. *Agricultural Water Management*, 84(2):153–165, 2006.

[66] R Jain, S Kulkarni, A Shaikh, and A Sood. Automatic irrigation system for agriculture field using wireless sensor network (wsn). *International Research Journal of Engineering and Technology (IRJET)*, 3(4):1602–1605, 2016.

[67] CJ Watras, M Morrow, K Morrison, S Scannell, S Yaziciaglu, JS Read, YH Hu, PC Hanson, and T Kratz. Evaluation of wireless sensor networks (wsns) for remote wetland monitoring: Design and initial results. *Environmental monitoring and assessment*, 186(2):919–934, 2014.

[68] SA Nikolidakis, D Kandris, DD Vergados, and C Douligeris. Energy efficient automated control of irrigation in agriculture by using wireless sensor networks. *Computers and Electronics in Agriculture*, 113(1):154–163, 2015.

[69] WH Nam, T Kim, EM Hong, JY Choi, and JT Kim. A wireless sensor network (wsn) application for irrigation facilities management based

on information and communication technologies (icts). *Computers and Electronics in Agriculture*, 143(1):185–192, 2017.

[70] AG Mohapatra and SK Lenka. Neural network pattern classification and weather dependent fuzzy logic model for irrigation control in wsn based precision agriculture. *Procedia Computer Science*, 78(1):499–506, 2016.

[71] V Rahangadale and D Choudhary. On fuzzy logic based model for irrigation controller using penman-monteith equation. *MIET*, 2011.

[72] F Touati, M Al-Hitmi, K Benhmed, and R Tabish. A fuzzy logic based irrigation system enhanced with wireless data logging applied to the state of qatar. *Computers and electronics in agriculture*, 98(1):233–241, 2013.

[73] C Goumopoulos, B O'Flynn, and A Kameas. Automated zone-specific irrigation with wireless sensor/actuator network and adaptable decision support. *Computers and Electronics in Agriculture*, 105(1):20–33, 2014.

[74] JG Jaguey, JF Medina, AL Guzman, and MA Gandara. Smartphone irrigation sensor. *IEEE Sensors journal*, 15(9):5122–5127, 2015.

[75] GM Suba, YM Jagadeesh, S Karthik, and ER Sampath. Smart irrigation system through wireless sensor networks. *ARPN Journal of Engineering and Applied Sciences*, 10(17):7452–7455, 2015.

[76] J Gutierrez, JF Medina, A Garibay, and MA Gandara. Automated irrigation system using a wireless sensor network and gprs module. *IEEE Transactions on Instrumentation and Measurement*, 63(1):166–176, 2013.

[77] MN Sudha, ML Valarmathi, and AS Babu. Energy efficient data transmission in automatic irrigation system using wireless sensor networks. *Computers and Electronics in Agriculture*, 78(2):215–221, 2011.

[78] RW Coates, MJ Delwiche, A Broad, and M Holler. Wireless sensor network with irrigation valve control. *Computers and electronics in agriculture*, 96(1):13–22, 2013.

[79] X Dong, MC Vuran, and S Irmak. Autonomous precision agriculture through integration of wireless underground sensor networks with center pivot irrigation systems. *Ad Hoc Networks*, 11(7):1975–1987, 2013.

[80] S Nallani and VB Hency. Low power cost effective automatic irrigation system. *Indian Journal of Science and Technology*, 8(23):1, 2015.

[81] N Wang, N Zhang, and M Wang. Wireless sensors in agriculture and food industry - recent development and future perspective. *Computers and Electronics in Agriculture*, 50(1):1–14, 2006.

[82] Bruno Gérard, John Hellin, Andrew Mcdonald, and Timothy Krupnik. Precision agriculture for smallholder farmers: An option? 12 2013.

[83] IF Akyildiz and X Wang. A survey on wireless mesh networks. *IEEE Communications Magazine*, 43(9):S23–S30, 2005.

[84] YD Kim, YM Yang, WS Kang, and DK Kim. On the design of beacon based wireless sensor network for agricultural emergency monitoring systems. *Computer standards & interfaces*, 36(2):288–299, 2014.

[85] KC Stone, AG Smajstrla, and FS Zazueta. Microcomputer-based data acquisition system for continuous soil water potential measurements. 1985.

[86] OM Grant, MJ Davies, H Longbottom, and CJ Atkinson. Irrigation scheduling and irrigation systems: Optimising irrigation efficiency for container ornamental shrubs. *Irrigation Science*, 27(2):139–153, 2009.

[87] SA Pathan and SG Hate. Automated irrigation system using wireless sensor network. *Int J Eng Tech Res*, 5(6):6–9, 2016.

[88] JS Awati and VS Patil. Automatic irrigation control by using wireless sensor networks. *Journal of Exclusive Management Science*, 1(6):1–7, 2012.

[89] R Hussain, JL Sahgal, and Md Riyaj. Control of irrigation automatically by using wireless sensor network. 2013.

[90] M Mafuta, M Zennaro, A Bagula, G Ault, H Gombachika, and T Chadza. Successful deployment of a wireless sensor network for precision agriculture in malawi. *International Journal of Distributed Sensor Networks*, 9(5):150703, 2013.

[91] X Yu, W Han, and Z Zhang. Remote monitoring system for intelligent irrigation in hybrid wireless sensor networks. *International Journal of Control and Automation*, 8(3):185–196, 2015.

[92] M Anbarasi, T Karthikeyan, L Ramanathan, S Ramani, and N Nalini. Smart multi-crop irrigation system using iot. *Int. J. Innov. Technol. Explor. Eng*, 8(1):153–156, 2019.

[93] S Abba, JW Namkusong, JA Lee, and ML Crespo. Design and performance evaluation of a low-cost autonomous sensor interface for a smart iot-based irrigation monitoring and control system. *Sensors*, 19(17):3643, 2019.

[94] A Raut, M Panse, D Chaware, and A Koparkar. Sensor based automated irrigation system. *International Journal of Engineering Research & Technology*, 4(5), 2015.

[95] G Tope and A Patel. Smart irrigation system. *Technical Research Organization India*, 3(3):23–26, 2016.

[96] R Testezlaf, FS Zazueta, and TH Yeager. A real-time irrigation control

system for greenhouses. *Applied Engineering in Agriculture*, 13(3):329–332, 1997.

[97] CN Lakshmiprasad, R Aashish, and JS Muzaffar. Smart irrigation in agriculture.

[98] V Gokul, R Ezhilarasie, and A Umamakeswari. A review of wireless sensor networks in automated irrigation. *Indian Journal of Science and Technology*, 9(1):48, 2016.

[99] J John, VS Palaparthy, S Sarik, MS Baghini, and GS Kasbekar. Design and implementation of a soil moisture wireless sensor network. In *Twenty First National Conference on Communications (NCC)*, pages 1–6. IEEE, 2015.

[100] D Anurag, S Roy, and S Bandyopadhyay. Agro-sense: Precision agriculture using sensor-based wireless mesh networks. In *First ITU-T Kaleidoscope Academic Conference-Innovations in NGN: Future Network and Services*, pages 383–388. IEEE, 2008.

[101] PB Chikankar, D Mehetre, and S Das. An automatic irrigation system using zigbee in wireless sensor network. In *International Conference on Pervasive Computing (ICPC)*, pages 1–5. IEEE, 2015.

[102] G Nisha and J Megala. Wireless sensor network based automated irrigation and crop field monitoring system. In *2014 Sixth International Conference on Advanced Computing (ICoAC)*, pages 189–194. IEEE, 2014.

[103] AAS Ammri and S Ridah. Smart irrigation system using wireless sensor network. *International Journal of Engineering Research & Technology (IJERT)*, 3(1):2278–0181, 2014.

[104] S Jadhav and S Hambarde. Automated irrigation system using wireless sensor network and raspberry pi. *International Journal of science and research (IJSR)*, 4(12):2056–2058, 2015.

[105] M Roopaei, P Rad, and Kwang RK Choo. Cloud of things in smart agriculture: Intelligent irrigation monitoring by thermal imaging. *IEEE Cloud Computing*, 4(1):10–15, 2017.

[106] Maria Romero, Yuchen Luo, Baofeng Su, and Sigfredo Fuentes. Vineyard water status estimation using multispectral imagery from an uav platform and machine learning algorithms for irrigation scheduling management. *Computers and electronics in agriculture*, 147(1):109–117, 2018.

[107] Mohamed Abdrabbo and Gamal El Afandi. Comparison of different evapotranspiration equations final, 04 2015.

[108] S Traore, Y Luo, and G Fipps. Deployment of artificial neural network for

short-term forecasting of evapotranspiration using public weather forecast restricted messages. *Agricultural Water Management*, 163(1):363–379, 2016.

[109] H Krupakar and A Jayakumar. A review of intelligent practices for irrigation prediction. *arXiv preprint arXiv:1612.02893*, 2016.

[110] MA Khan, MZ Islam, and M Hafeez. Data pre-processing and evaluating the performance of several data mining methods for predicting irrigation water requirement. *arXiv preprint arXiv:2003.00411*, 2020.

[111] NH Hellin, JM Rincon, RD Miguel, FS Valles, and RT Sanchez. A decision support system for managing irrigation in agriculture. *Computers and Electronics in Agriculture*, 124(1):121–131, 2016.

[112] A Goap, D Sharma, AK Shukla, and CR Krishna. An iot based smart irrigation management system using machine learning and open source technologies. *Computers and Electronics in Agriculture*, 155(1):41–49, 2018.

[113] A Goldstein, L Fink, A Meitin, S Bohadana, O Lutenberg, and G Ravid. Applying machine learning on sensor data for irrigation recommendations: revealing the agronomist's tacit knowledge. *Precision agriculture*, 19(3):421–444, 2018.

[114] J Timsina and DJ Connor. Productivity and management of rice–wheat cropping systems: Issues and challenges. *Field Crops Research*, 69(2):93–132, 2001.

[115] KK Garg, BS Das, M Safeeq, and PBS Bhadoria. Measurement and modeling of soil water regime in a lowland paddy field showing preferential transport. *Agricultural Water Management*, 96(12):1705–1714, 2009.

[116] T Li, E Humphreys, G Gill, and SS Kukal. Evaluation and application of oryza2000 for irrigation scheduling of puddled transplanted rice in north west india. *Field Crops Research*, 122(2):104–117, 2011.

[117] Bas Bouman. How much water does rice use. *Management*, 69(1):115–133, 2009.

[118] Z Mimi and M Smith. Statistical domestic water demand model for the west bank. *Water International*, 25(3):464–468, 2000.

[119] IP Calvo and JCG Estrada. Improved irrigation water demand forecasting using a soft-computing hybrid model. *Biosystems Engineering*, 102(2):202–218, 2009.

[120] SK Saleem, DK Delgoda, SK Ooi, KB Dassanayake, L Liu, MN Halgamuge, and H Malano. Model predictive control for real-time irrigation

scheduling. *IFAC Proceedings Volumes*, 46(18):299–304, 2013.

[121] IPS Ahlawat and RP Sharma. Agronomic terminology, 1993.

[122] K Djaman, VC Mel, BV Bado, B Manneh, L Diop, D Mutiibwa, DR Rudnick, S Irmak, and K Futakuchi. Evapotranspiration, irrigation water requirement, and water productivity of rice (oryza sativa l.) in the sahelian environment. *Paddy and Water Environment*, 15(3):469–482, 2017.

[123] K Steppe, DJW Pauw, and R Lemeur. A step towards new irrigation scheduling strategies using plant-based measurements and mathematical modelling. *Irrigation Science*, 26(6):505, 2008.

[124] P Steduto, TC Hsiao, D Raes, and E Fereres. Aquacrop - the fao crop model to simulate yield response to water: I. concepts and underlying principles. *Agronomy Journal*, 101(3):426–437, 2009.

[125] HG Jones. Irrigation scheduling: Advantages and pitfalls of plant-based methods. *Journal of Experimental Botany*, 55(407):2427–2436, 2004.

[126] HL Penman. Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 193(1032):120–145, 1948.

[127] JL Monteith. Evaporation and environment. the stage and movement of water in living organisms. In *19th Symp. Soc. Exp. Biol.* Cambridge University Press, 1965.

[128] GH Hargreaves and ZA Samani. Reference crop evapotranspiration from temperature. *Applied engineering in agriculture*, 1(2):96–99, 1985.

[129] IA El-Magd and T Tanton. Remote sensing and gis for estimation of irrigation crop water demand. *International Journal of Remote Sensing*, 26(11):2359–2370, 2005.

[130] AF Torres, WR Walker, and M McKee. Forecasting daily potential evapotranspiration using machine learning and limited climatic data. *Agricultural Water Management*, 98(4):553–562, 2011.

[131] L Maton, D Leenhardt, M Goulard, and JE Bergez. Assessing the irrigation strategies over a wide geographical area from structural data about farming systems. *Agricultural systems*, 86(3):293–311, 2005.

[132] KH Amer, AA Samak, and JL Hatfield. Effect of irrigation method and non-uniformity of irrigation on potato performance and quality. *Journal of Water Resource and Protection*, 8(3):277, 2016.

[133] MS Sreekanth, R Rajesh, and J Satheeshku. Extreme learning machine for the classification of rainfall and thunderstorm. *Journal of Applied Science*, 15(1):153–156, 2015.

[134] E Giusti and SM Libelli. A fuzzy decision support system for irrigation and water conservation in agriculture. *Environmental Modelling & Software*, 63(1):73–86, 2015.

[135] SS Abdullah, MA Malek, NS Abdullah, O Kisi, and KS Yap. Extreme learning machines: A new approach for prediction of reference evapotranspiration. *Journal of Hydrology*, 527(1):184–195, 2015.

[136] I Campos, C Balbontin, JG Piqueras, MP Dugo, CMU Neale, and A Calera. Combining a water balance model with evapotranspiration measurements to estimate total available soil water in irrigated and rainfed vineyards. *Agricultural Water Management*, 165(1):141–152, 2016.

[137] MV Vela, I Abrisqueta, W Conejero, J Vera, and MCR Sanchez. Soft computing applied to stem water potential estimation: A fuzzy rule based approach. *Computers and Electronics in Agriculture*, 115(1):150–160, 2015.

[138] I Abrisqueta, W Conejero, MV Vela, J Vera, MF Ortuno, and MRC Sanchez. Stem water potential estimation of drip-irrigated early-maturing peach trees under mediterranean conditions. *Computers and Electronics in Agriculture*, 114(1):7–13, 2015.

[139] F Karandish and J Simnek. A comparison of numerical and machine-learning modeling of soil water content with limited input data.

[140] RG Perea, EC Poyato, P Montesinos, and JAR Diaz. Prediction of irrigation event occurrence at farm level using optimal decision trees. *Computers and Electronics in Agriculture*, 157(1):173–180, 2019.

[141] MA Khan, MZ Islam, and M Hafeez. Irrigation water requirement prediction through various data mining techniques applied on a carefully pre-processed dataset. *Journal of Research and Practice in Information Technology*, 43(22):1–17, 2011.

[142] S Andriyas and M McKee. Recursive partitioning techniques for modeling irrigation behavior. *Environmental Modelling & Software*, 47(1):207–217, 2013.

[143] AJ McDonald, SJ Riha, JM Duxbury, TS Steenhuis, and JG Lauren. Water balance and rice growth responses to direct seeding, deep tillage, and landscape placement: findings from a valley terrace in nepal. *Field crops research*, 95(2-3):367–382, 2006.

[144] R Kharif. Package of pratices for the crops of punjab. *Punjab Agricultural University Ludhiana*, 2019.

[145] A Jimenez, BV Ortiz, L Bondesan, G Morata, and D Damianidis. Artificial neural networks for irrigation management: A case study from

southern alabama, usa. In *Precision agriculture'19*, pages 918–929. Wageningen Academic Publishers, 2019.

[146] F Capraro, D Patino, S Tosetti, and C Schugurensky. Neural network-based irrigation control for precision agriculture. In *IEEE International Conference on Networking, Sensing and Control*, pages 357–362. IEEE, 2008.

[147] IP Calvo, P Montesinos, J Roldan, and FR Navarro. Linear regressions and neural approaches to water demand forecasting in irrigation districts with telemetry systems. *Biosystems Engineering*, 97(2):283–293, 2007.

[148] F Bu and X Wang. A smart agriculture iot system based on deep reinforcement learning. *Future Generation Computer Systems*, 99(1):500–507, 2019.

List of Publications

1. Ravneet Kaur Sidhu, Ravinder Kumar, Prashant Singh Rana *"Machine Learning Based Crop Water Demand Forecasting using Minimum Climatological Data"*, Multimedia Tools and Applications, Springer, 79(1):13109-13124, 2020. [SCI, IF 2.1]
2. Ravneet Kaur Sidhu, Ravinder Kumar, Prashant Singh Rana, *"Long Short Term Memory Neural Network Based Multi-level Model for Smart Irrigation"*, Modern Physics Letters B, World Scientific Publishing, 2020 [In Press]. [SCIE, IF 0.73]
3. Ravneet Kaur Sidhu, Ravinder Kumar, Prashant Singh Rana, Mangi Lal Jat *"Automation in drip irrigation for enhancing water use efficiency in cereal systems of South Asia: Status and Prospects"*, Advances in Agronomy, Elsvier. [Under Review, IF 5.8]