UNIVERSITY OF SOUTHAMPTON
Faculty of Physical Sciences and Engineering
Department of Electronics and Computer Science

Thesis submitted for Mphil

Supervisors: Dr. Nick R. Harris, Dr. Geoff V. Merrett

Optimal Reservoir and Back Runoff Channels based Two Farms Irrigation Discharge Prediction System

By Marwan Khan

15 December 2020

UNIVERSITY OF SOUTHAMPTON

Faculty of Physical Sciences and Engineering

Department of Electronics and Computer Science

Abstract

By Marwan Khan

Most parts of world are currently facing an acute water shortage that is likely to become worse in the coming years. Climate change and global warming has a significant impact on the hydrological cycle. Both these factors climate change and global warming effect on the rainfall patterns and temperature. As the temperature increase from 2 Celsius to 4 Celsius it will rise the evaporation from the land and sea. The rainfall will be in higher intensity in higher latitudes and decrease in mid latitudes. The areas of the world which has scarce water will become drier and hotter. Global water withdrawal for agricultural sector is approximately 70%. However, most of this fresh water approximately 50 % is wasted due to inefficient and poorly managed irrigation system. The farming community in under develop countries of the world is wasting a huge amount of fresh water by using outdated and poorly managed flood irrigation (surface irrigation).Runoff estimation/prediction can be very valuable in water management and irrigation scheduling management. In this research an optimal reservoir precision irrigation system based on runoff estimation between two farms (farm1 and farm2) has been proposed to reduce water waste and to utilize the runoff water in nearby farm i.e farm2 or divert it back to reservoir through back runoff channels from both the farms in case of surplus amount of water left from either irrigation or there is an excessive rainfall. NRCS (Natural Resources Conservation Service), ANN (Artificial Neural Network),DT (Decision Tree),SVR (Support Vector Regression) and MLR (Multiple Liner Regression) are used to predict discharge, peak discharge and time to peak at farm1 and farm2 outlets. The performance of these algorithms is evaluated using different performance metrics. Overall, ANN show good performance for different datasets and scenarios while MLR show worse performance. Beside this an IOT (Internet of Things) based model is developed which remotely retrieved data from different environmental and agricultural based sensors such as temperature sensor, soil moisture sensor and crop stage sensor. The current conditions of farms is retrieved from sensors on mobile application, the end user has to only enter the precipitation depth/irrigation depth and the predication results are displayed in form of table showing NRCS predication, and other machine learning algorithms predication for total discharge, peak discharge and time peak, their comparison and also their respective hydrographs are displayed for different farm conditions.

Table of Contents

List of Figures

List of Tables

Declaration of Authorship

I, Marwan Khan, declare that this thesis entitled 'Optimal Reservoir and Back Runoff Channels based Two Farms Irrigation Discharge Prediction System' and the work presented it is entirely my own. All direct or indirect sources used are acknowledged as references. This work was done wholly or mainly while in candidature for a research degree at this University.

Marwan Khan

Dated: 14-12-2020

Acknowledgement

I cannot express enough thanks to my supervisory team for their continued support and encouragement: Dr Nick Harris my main supervisor and Dr Geoff Merrett my co-supervisor. Thank you for providing me a learning opportunity.

My completion of this MPhil thesis could not have been accomplished without support of my caring , loving and supportive wife Sanam Noor and my children Aleena khan, Numair khan and Zunair khan . Your encouragement when the times got rough are much appreciated and duly noted.

Finally My heartfelt thanks to my parents Mr. Sher Muhammad khan and Mrs: Abida for your limitless prayers for my success on every step and particularly for my this journey of studying from abroad.

Chapter 1: Introduction

Climate change and rise in the world population have tremendous pressure on the supplies of global fresh water. Water is critical for every aspect of life. To achieve food security, water play a key role in both supporting the ecosystem such as agriculture and related economic activities. In the production of energy and food, water is the key factor. There are around 1.2 billion people in the world who lives in the water stressed areas. One third of the global population will live in the regions of water scarcity by 2040. [1].

Food, energy and water production and consumption forms a nexus. Alterations in any one factor have significant impact on the others. [1], [2]. Both in the development planning and in planning for adjustment, they are treated as an independent issue. In the maximization of the collaboration and minimization of trade-offs by the systematic use of land, water, energy, and other important assets to bring effective adjustment for the climate change. [1].

In an early and middle twentieth century the world's irrigated area increased very rapidly due to rise in population growth and food demand. Under the medium projections the global population rise is estimated to be more than 9 billion people by 2050. Future global demand for food is predicated to be 70 % by 2050 and approximately double for developing nations. On the global perspective, Irrigation provides 40 percent of the global food from around 20 percent of agricultural land or around 300 million hectares. 11% increase is estimated for the amount of water withdrawal for irrigated agriculture to meet the demand for biomass production [3].

Particularly In the developing countries the climate change is intimidating fundamental factors of life such as water, energy, food production and the environment as well as on the global scale. Farming water management and agriculture has been impacted by climate change as it is able to be seen from the precipitation, aqueduct and aquifers from which the water is to be collected. Considerable adjustment will be required for the supply and best usage of any of these demolishing assets. The tracing of the upper limit of the model scenario for ambience carbon dioxide (CO_2), temperature and sea level rise has been developed in the fourth assessment (AR4) by the international panel on climate change (IPCC). The evidence of climate change is now unambiguous. Current negotiations focus not to increase the global temperature beyond 2 degrees Celsius to avoid any negative impacts. The global climate model (GCM) simulation modelling has been used to predict these impacts on global scale while to downscale predications at national, regional level and river basin, regional climate modelling (RCM) has been utilized. By 2080 the global atmospheric temperature will rise approximately 4 degree Celsius, which is consistent with doubling of carbon dioxide ($CO_{2)}$ concentration [3].

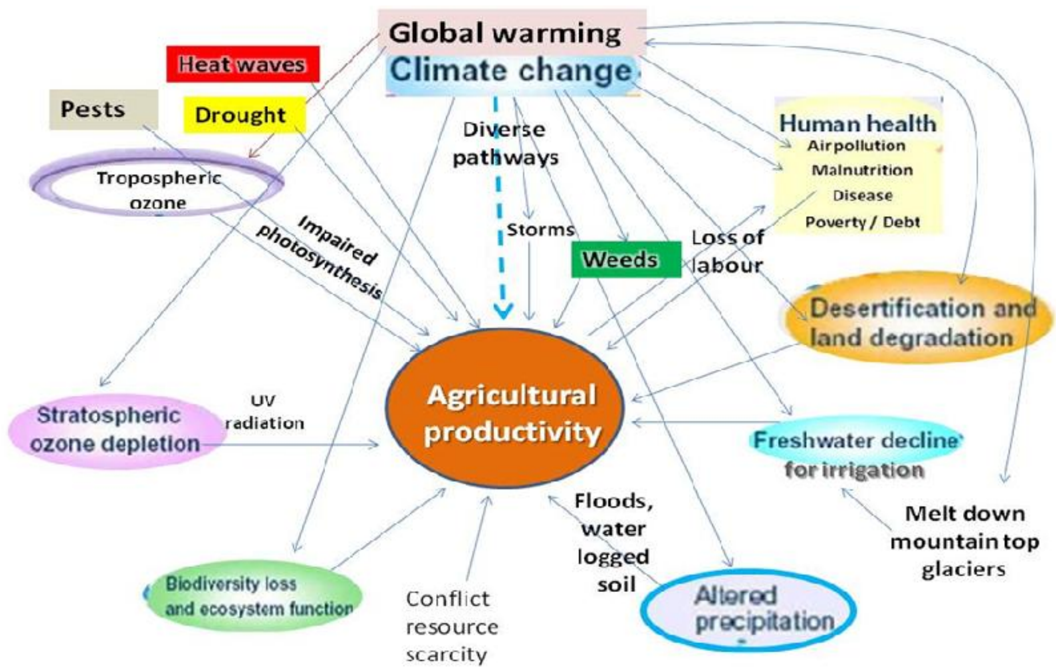


Figure 1.1 Multiple impact of global warming and climate disruption on agriculture [4]

The hydrological cycle will fasten due to global warming and it will rise the temperature which in turn will increase evaporation rate from the ground and sea. The precipitation is predicated to rise in higher latitudes and decline in mid latitudes already dry, semi-arid to arid areas because the temperature will quickly rise in higher latitudes than mid or lower latitudes. From the sea level the mean temperature is excepted to be higher at the altitude which will be responsible for snow melt and glacier retreat. The areas of the world which has water scarcity will become drier and hotter. Temperature and precipitation are predicated to be more dynamic which will lead into floods and droughts more likely in the same area.

In the Indian monsoon for example , there should be dams to gather the huge amount of flood flow that usually occurred due to steamy tropics and intens Rather than drier parts of southern Europe and north America which dry to semi-parched and have lower precipitation, the spill over and groundwater both are probably going to decrease significantly. [3].

Overall water withdrawal proportions are 19 % manufacturing, 11 % civil and 70 % for farming, anyway these values are one-sided and change from one landmasses to different mainlands of the world. The water withdrawal proportions are profoundly reliant on both atmosphere and spot of farming. [4].

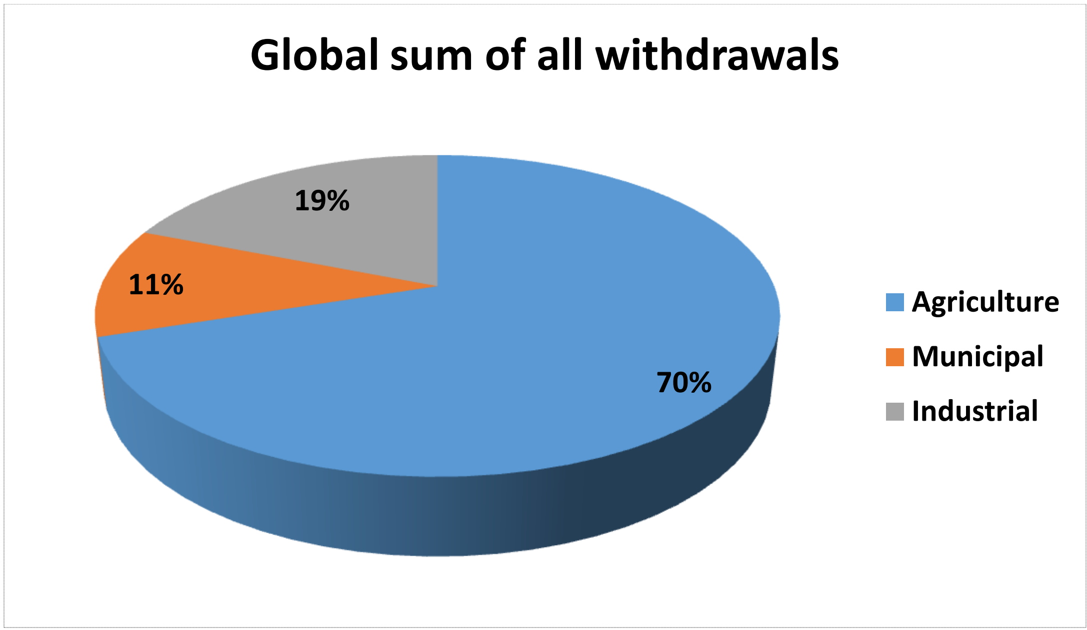


Figure 1.2: Global sum of all water withdrawals [4]

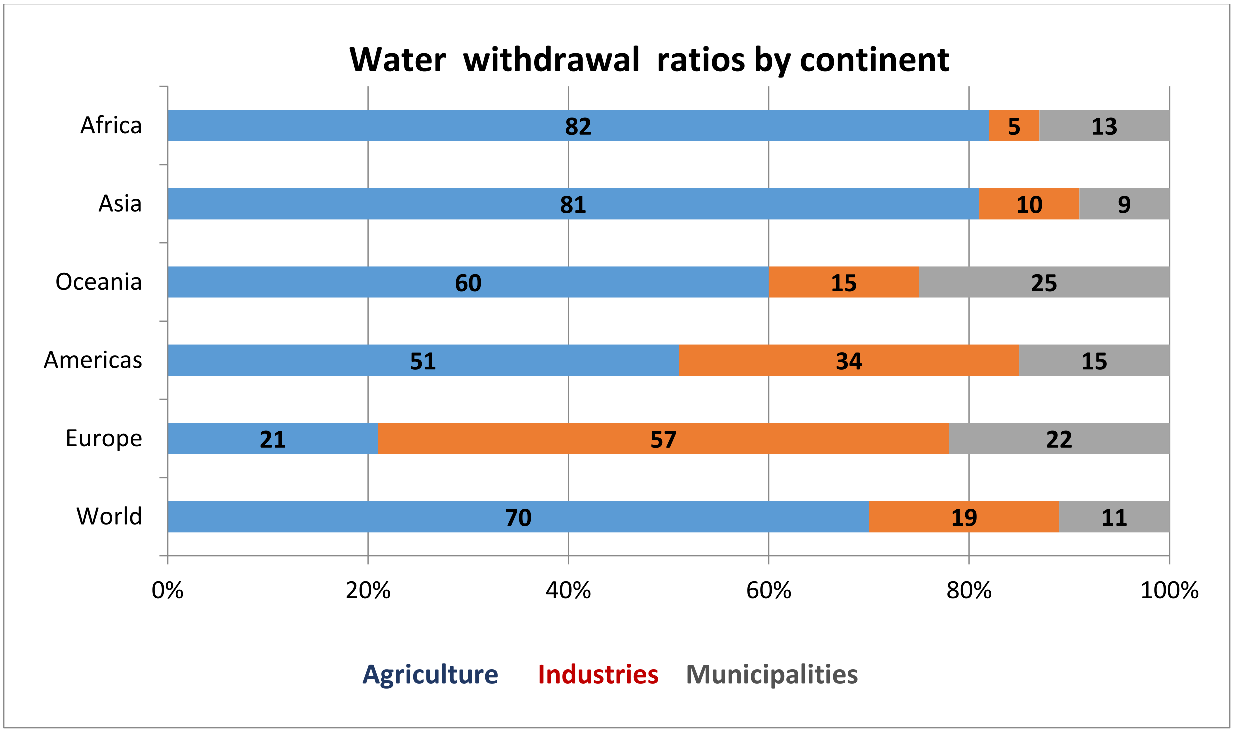


Figure 1.3 : Water withdrawal ratios by continent [4]

Well established methods of agriculture are rainfed agriculture and irrigated agriculture. More than 80% of the world agriculture crop land is based on rainfed. Rainfed agriculture output 60 %

of the world food. Rainfed farming is particularly influenced by the environmental change. while on the other side, irrigated farming gives greater efficiency in developed nations of the world with efficient water system frameworks.

Table 1.1 Net changes in major land use (million ha) [5]

	1961	2009	Net increase 1961-2009
Cultivated land	1368	1527	12%
Rainfed agriculture	1229	1229	-0.2%
Irrigated agriculture	139	301	117%

Rainfed farming rules land use in numerous nations of the world and subsequently it is the significant determinant of hydrology and overflow in a stream basin.

However, net increase is 117% in major land use which is due to irrigated agriculture from 1961 to 2009 while there is no increase in land use by rainfed agriculture. The cultivated land increase is 12 % from 1961 to 2009 [3].

1.1 Precision Irrigation

Precision irrigation is turning into a fundamental segment of cultivating in numerous zones of the world since it is an apparatus for guaranteeing nourishment security. This outcomes in developing rivalry for accessible new water supplies between agribusiness, industry and household use. A marker of this competition is that during the most recent couple of decades, ground water is draining at a disturbing rate in numerous agrarian zones. Additionally, farming should create more nourishment to address the necessities of a developing populace. Whenever irrigated farming is to extend so as to satisfy developing needs for nourishment, at that point new water system practices and instruments must be created for progressively proficient water use. Precision water system is one potential methodology. [6].

Water mishandling and crude techniques for development have prompted the sharp lessening in water assets. Because of poor proficiency of primitive water system technique, for example, "Surface Irrigation" or "Flood Irrigation" frameworks cause a 40 to 60 percent loss of water as spillover. [7], [8]. Traditional water system practice includes applying water consistently over all aspects of the field without considering the spatial inconstancy in soil and yield water needs; this subsequently prompts over water system in certain zones of the field while different zones of the field are under inundated. [9]. The dangers related with over flooding incorporate surface spillover, profound permeation and draining of nitrates and supplements. Those related with under watering are increasingly subjective and incorporate decrease for crop yields and quality, just as inefficient utilization of manure and other supplemental contributions for crop yield [10],

[11]. There is developing logical enthusiasm for the potential task that precision water system can make towards improving harvest profitability, and expanding water and vitality productivity in irrigated farming. [11]. To keep up the production capacity, the need is to switch towards more effective methods for utilizing the water, which can augment the economics and social gains while utilizing the minimum resources. This is conceivable through unique endeavors coordinated at improving value in access and optimum allocation of water assets [12]. Water use efficiency, which can be characterized as "yield generated per unit precipitation and/or irrigation of water applied", is the primary determinant of the crop yield under conditions of water shortage [13].One potential arrangement (and now broadly utilized) is the way toward reusing the wastewater for farming, household and manufacturing use. A large number of researches are presently devoted to the recycling water strategies. Another solution is the precision irrigation system, implemented at micro level, which looks after the use of water on individual farms using Wireless Sensor Networks. As reported by Pacific Institute, 'precision irrigation' can save water in irrigation from 11 percent to 50 percent [14].

Different precision water system frameworks are being used, which are sprinkler (focus rotate) water system framework, under surface (sprinkler) water system framework and small scale water system or subsurface (trickle) water system frameworks. Every exactness water system framework has its potential application efficiency. "The water application efficiency is how well an irrigation system performs its primary task of delivering water from the conveyance system to the crop. The objective is to apply the water and store it in the crop root zone to meet the crop water requirement. The irrigation efficiency is defined from three point of view (1) irrigation system performance (2) the uniformity of water application and (3) the response of crop to irrigation" [15] .PAE are discussed in table below.

Table 1.2 "Potential" application efficiencies for well designed and well managed irrigation systems [15]

Irrigation systems	**"Potential" application efficiency (%)**
Sprinkler irrigation systems	
LEFA	80-90
Linear move	75-85
Center pivot	75-85
Travelling gun	65-75
Side roll	65-85
Hand move	65-85
Solid set	70-85
Surface irrigation systems	
Furrow (conventional)	45-65
Furrow (surge)	55-75
Furrow (with tail water reuse)	60-80
Basin (with or with out furrow)	60-75
Basin(paddy)	40-60

Precision level Basin	65-80
Micro irrigation system	
Bubbler (lower head)	80-90
Micro spray	85-90
Micro-point source	85-90
Micro-line source	85-90
Subsurface drip	>95
Surface drip	85-90

Water system planning is critical to successfully oversee water assets and advance gainfulness of an irrigation activity. Water system planning is essentially knowing when and how much irrigation water to apply. Successful water system planning amplifies benefit while limiting sources of inputs, for example, water and energy. There are a multiple factor that influence water system planning. Sort of yield, phase of harvest improvement, soil properties, soil-water connections, accessibility of water supply, lastly climate conditions (temperature, wind, precipitation, and others) have a basic task while deciding viable water system planning. Evapotranspiration (normally known as ET) is the term used to portray the aggregate of evaporation and plant transpiration starting from the earliest stage from the earth to the climate. Corn ET values during the sweltering summer months can run as high as a half inch for each day [16].

The parity of a few center viewpoints is anyway significant for the successful execution of a effective precision water system framework. Executing an exactness water system framework includes endeavors on ongoing checking of yield and soil conditions, planning water system and control of the water system application equipment. Research has been predominantly centered around the detecting and control parts of precision water system with much headways in the most recent decade. [17], [11].

1.2 Motivation

The motivation of my research work is based on the below three broad aspects.

- Currently Global Problem

The mismanagement of water and the use of primitive techniques for irrigation has led to waste about 40 to 60 percent of water in form of flooding irrigation system. On the global scale there is a sharp rise such as 70 percent predicated in food growth projected till 2050 which is linked to increase in the population growth of around 9 billion in 2050. There is a need for the use of efficient irrigation techniques and water waste management techniques to be utilized in agricultural sector in order to increase food and its security for people living in water stress regions of the world. One reason for the motivation is to focus my research work on the recent global issue to address.

- Machine learning model utility in hydrological modeling

Initially Physical based models are used for understanding, simulation and estimation of hydrological processes as well as the use of physical based equations which explains these processes. The statistical models, including time series analysis forms another popular modeling technique for hydrological processes, climate change and studying earth system. However both the physical based models and statistical models has shortcomings in terms of accuracy, the need for a lot of data, computational cost and weakness in uncertainty analysis. Therefore well-established machine learning models are an ideal choice to be utilized for hydrological processes,climate change and earth systems understanding , predication and to overcome the shortcomings of physical and statistical models through efficient computation and intelligence as an evidence from literature . The other reason for the motivation is to solve the hydrological problem in the domain of agriculture water management from irrigation or rainfall through machine learning techniques and explore the capabilities of these ML models through new results obtained in this research work and also to provide a comprehensive review in these Machine learning models in various hydrological processes modeling from the already existing .literature.

- Previous literature most relevant work in the area and its short comings.

There are very rare work in the literature regarding two Farms Irrigation Discharge Prediction System and the utilization of the surplus water in the agricultural lands. The motivation of this research is based on research “Enabling proactive agricultural drainage reuse for improved water quality through collaborative networks and low complexity data-driven modeling” [18]**.** In this research Thesis the author proposed a novel framework which manage the agriculture drainage and nutrient losses at proactive way at field scale. The complicated models at field scale are replaced by in-situ sensing, communication and low complexity predicative models which are best suited for autonomous operation. Local field scale Wireless sensor networks are combined through information exchange mechanism through the utilization of Water Quality Management using Collaborative Monitoring (WQMCM) framework. The components of the WQMCM includes 1) neighbour learning and linking 2) low complexity predicative models for drainage dynamics 3) low complexity model for nitrate losses and 4) decision support model for drainage and nitrate losses. The concept is taken from the Natural Resource Conversation Model (NRCS) and De-Nitrification ,De-composition (DNDC) models to develop drainage dynamics and nitrate losses predicative models respectively. Machine learning models which are used in this thesis are 1) multiple linear regression 2) M5 model tree 3) artificial neural network 4) C4.5 and 5) Naïve Bayes algorithms. The author described the thesis contribution as follow 1) Architecture development and implementation of WQMCM for networked catchment 2) 50 % less parameterization by developing drainage dynamics and nitrate losses predicative models. 3) M5 tree algorithm is used for validation of predicative models for drainage dynamics and nitrate losses, the data for 12 months long event dataset is used from a catchment in Ireland. The modelling results are compared with existing models and then further tested with other machine learning models. 4) The use of Naïve Bayes model for the development of decision support model for reusability of drainage and nitrate losses.

However, there are gaps in the author research work [18].Which are described as follow.

1) The dataset that is used for Q – predicative model is for stream flow forecasting from one of the catchment location in the Ireland , The author mentioned that the dataset is publicly available from the Environmental Protection Agency (EPA) website. However for the agricultural farms discharge predicative model there is no interrelated data available which take into consideration all the parameters of agriculture farms such as the underlying process that how farm soil parameters , crop parameters , irrigation parameters (water related) and environmental parameters are available. Most of these dataset are available independently from each other such as USGS Water Data for the nation, International Soil Moisture Network, European Soil Data Center (ESDAC),sustaining the earth watershed's ,Agricultural Research Data System (STEWARDS) .These datasets provide valuable data for managing water and soil resources, however none provide a detailed dataset that links soil , water movement and crop growth at various stages with different agriculture system [19].
2) The other limitation in the author work is no real world deployment of the system nor real world data is acquired from the use of appropriate sensors for soil moisture , crop stages and their interlinking so that the sensory data are to be feeded into the machine learning models .
3) There is no user interface for the end users such as the use of agriculture -internet of things (AG-IOT) application in form of mobile and desktop based applications. As mobile application is of core importance due to the 5G mobile computing capabilities emergence.
4) The author used a mathematical equation for the estimation of time (t1 and td) in which according to the author there is no mathematical equation directly exists for tp which is a term used in the td mathematical model .The author also mentioned that there is no mathematical expression for t1 as well . According to the author own understanding has extracted these two parameters which is not mentioned clearly that how ? Thus, it is very difficult for other author to interpret.
5) In the author thesis M5 tree models are considered as an efficient machine learning algorithms in the hydrological modelling , however from the latest research it is clear that the multiple model ANN perform well than M5 tree model in water resource management [20]. In the latest literature review it is cleared that Artificial neural networks such as Multiple model ANN, Ensemble , Deep learning and Extreme learning machines are gaining popularity than other machine learning models in hydrology and agriculture [21], [22], [23].

To tackle the shortcomings in the author work, The concept of Optimal Reservoir and Back Runoff Channels based Two Farms Irrigation Discharge Prediction System is developed which provides the following features development and implementation in the motivation of this research work.

1. Back runoff channels based concept for water saving in the reservoir storage
2. Irrigation and rainfall based runoff modelling
3. Mobile and desktop based development and implementation of the concept as currently there is a boom of 5G technology emergence in mobile computing which could be utilized in the agriculture domain.
4. Use of Internet of things (IOT) such as Arduino and Android based application in the digital agriculture domain (AG-IOT).
5. A simplified decision support system and recommendation for the reservoir and two farms flow paths opening closing and diversion of surplus water back to reservoir through back runoff channels or divert surplus water to stream if the reservoir storage capacity threshold exceeds the specified limit. An early warning system for stream or early alarming systems for farms are established.
6. Use of the available most relevant sensors for data accumulation from the real world
7. Machine learning modelling in hydrological and agriculture modelling which reflect on that ANN is better than rest of well-established machine learning models and in the future deep learning, multiple models ANN, ensemble is to be utilized.
8. Simulation based on SCS- NRCS is performed and then compared with other machine learning models both numerically and graphically.

1.3 Problem Statement

The precision irrigation system has long been used on individual farms to minimize water waste. However, the best precision irrigation system cannot guarantee 100 % water utilization. Very rare work has been done so far on two farms system to utilize water waste. The precision irrigation system based on wireless sensor network can be extended to two farms in such a way that the surplus amount of water/ runoff of a farm can be utilized in the second nearby farm for an irrigation purpose. But the question arise here is how to predict runoff proactively so that the nearby farm can be able to adjust its irrigation system accordingly and the runoff water can be utilized for the second farm's irrigation need. The runoff hydrograph is the instantaneous runoff/discharge rate against time. It is composed of total discharge, peak discharge, total time and time to peak. In order to estimate the runoff, the local environment of the farm need to be sensed. There are some important variables for the runoff estimation such as soil moisture, crop stages, precipitation / irrigation depth and ambient temperature. there are various well established and low cost sensors developed for soil moisture estimation, temperature estimation ,precipitation/irrigation depth estimation and crop stages estimation, however to acquire real data from some of the sensor is very complex and tedious task such as crop sensor . Machine learning application can be utilized to learn and predict the runoff of the farms environment with minimum prior knowledge. The goal is how the Artificial intelligence specifically Machine learning and internet of things (IOT) can be integrated to be utilized in the reservoir based two farms agriculture domain for the smart irrigation and runoff estimation? Beside that how decision support system and back runoff channels [24] helps to save the surplus amount of water in the farms either from the excessive rainfall or from the reservoir irrigation system?.

1.4 Problem solution

- To solve the problem, an Unified Modeling Language (UML) diagram has been developed for the conceptual modeling in the Methodology chapter of this thesis. With the help of UML diagrams different irrigation scenarios are established such as reservoir based irrigation and rainfed- irrigation, the farms scenarios when the farms are dry and wet. UML also elaborates on how the wireless sensor network based on the two farms and reservoir can communicate with each other through packet transmission and how the farms and reservoir inlets and outlets sensors works.
- To achieve the conceptual modeling to maximum extent, there are certain tools such as USDA-NRCS simulator and MATLAB used for the simulation purpose for the Reservoir –two farms runoff/discharge predication. The NRCS scripting has been done in detail for “Reservoir and back runoff channels based two farms irrigation discharge system”. The curve number entity in the NRCS has been constituted and scripted for the three conditions of soil moisture (dry , wet and average) and an additional condition extreme wet on the basis of the mathematical equations which exists and for the three different crop stages such as (crop stage 1, crop stage2 and crop stage 3) taken on the basis of NRCS & TR 55 documentation on urban hydrology for small watersheds. A separate desktop based graphical user interface has been developed for reservoir and back runoff channels based two farms irrigation total discharge and peak discharge system on the basis of relevant mathematical equations.
- A decision support system has been developed for the Reservoir and back runoff channels based two farms irrigation discharge predication system in which different scenarios are developed for both the farms such as soil moisture conditions are dry , average wet , wet and extremely wet conditions , crop stages =1,2,3 (however crop stage 1 is selected in this thesis), crop stage 1= 0% surface cover(fallow land),crop stage 2= < 50% surface cover (small grains) and crop stage 3= >75 % surface cover (small grains), irrigation/perception depth is also dynamically selected and the rest of variables selected are fixed . The detail is mentioned in decision support system for runoff estimation in chapter 4 as well as chapter 5 of this thesis.
- For the real world modeling to learn and predict reservoir and back runoff channels based two farms irrigation total discharge /peak discharge system , an Arduino based setup has been made to accumulate data for various input parameters such as soil moisture through Arduino soil moisture sensor , temperature from Arduino temperature sensor, crop stages through Arduino digital camera (however the values are

not extract through image processing feature extraction rather they are assigned as proxy values from the NRCS and TR 55 documentation for urban hydrology on small watersheds) , precipitation depth or irrigation depth values can be entered directly by the end user such as (farmer, hydrologist or machine expert) from nearby online meteorological site available to the mobile app developed in the android studio, which directly rendered the data from sensors and display it in the mobile application the NRCS based total discharge predication values and its bar graphs for comparison between NRCS predication value and other machine learning algorithms predication values and peak discharge and time to peak hydrographs as well . The tools that has been used for mobile application building is android studio, Arduino sensors script, PHP script , my sql as database and MATLAB integration.

1.5 Objectives achieved

- Reservoir and back runoff channels based two farms total discharge system (in unit of time) are developed through UML based conceptual modeling.
- Developed a desktop based graphical user interface for reservoir and back runoff channels based two farms total discharge system and peak discharge system and generate results for predications values and their hydrographs.
- Optimal reservoir based water distribution among multiple farms has been achieved for the heterogeneous irrigation purpose on two farms.
- Save water waste and energy through back runoff channels based system when the surplus water is left from either irrigation or an excessive rainfall to fill in the reservoir.
- Hydrographs generation on the basis of peak discharge and time to peak on farm 1 and farm 2 outlet reflecting an early alarming system for the other farm to the adjust the water in the other farm sprinkler based precision irrigation system and an early warning system for river basin to adjust the water that are to be travelled to the river basin arrived from farm2.
- An Internet of things (IOT) based android mobile application has been developed for the end user. The end user (farmer, hydrologist , machine expert) can very conveniently find out the two farms soil moisture , crop stage and temperature status by entering the precipitation depth values from his mobile app and any end user have installed the mobile app will also get the total discharge or peak discharge predication values for both farm1 and farm2 on the basis of NRCS and sophisticated machine learning algorithms such as Multiple linear regression , support vector regression , regression trees and artificial neural network , the comparison is shown in tabulated form and bar graphs as well as in hydrographs form to show peak discharge and time to peak .

1.6 Research contributions

Research contributions of this thesis are as follow:

- Very limited work has been done in irrigation and precipitation runoff modeling on multiple farms scale, most of that limited work done is based on too much assumptions based on their empirical or mathematical equations. However on my this research work both the irrigation and precipitation runoff modeling is carried out on the basis of logical conceptual modeling in the form of unified modeling diagrams and the relevant mathematical or empirical equations based on the existing literature.
- In the previous work done from the literature , the data is only utilized from the simulator for precipitation runoff modeling , In my this thesis report the contribution is that for modeling purpose of Irrigation/precipitation the simulation is performed in the appropriate simulator and utilize with correct and relevant mathematical equations for total discharge , Curve number for dry , wet and average etc , as well as their appropriate real world data is acquired from different sensors in a lab-farm setup. In the previous work in literature there is no real world data acquired for this concept.
- To show the water distribution and irrigation or precipitation runoff modeling done through hydrological simulator and machine learning models among two farms a sophisticated desktop based graphical user interface is developed for the ease of understanding to the users.
- The other contribution of the work is development of Andriod –Ardunio based mobile application for the irrigation /precipitation runoff modeling which has not previously carried out for other work that has been done in this area.
- Another interesting feature of this research work is the utilization of back runoff concept to save water waste and energy to fill in the extra runoff water into the reservoir and then reuse that water for farms irrigation needs.
- The contribution of this work is integration of Reservoir and Two farms for irrigation and precipitation runoff modeling and the utilization of back runoff channel concept , a simplified decision support system is built for decision making for water inflow and outflow.UML concept of how overall system works in logical manner and wireless sensor networks integration.

1.7 Published papers

1. Optimal Reservoir And Back Runoff Channels Based Two Farms Irrigation total Discharge Predication System, International Journal of Advanced Research in Engineering and Technology (IJARET), Volume 11, Issue 10, October 2020, PP. 837-848 , ISSN Online: 0976-6499,Scopus indexed ,impact factor 10.94

2. Irrigation Runoff Volume Prediction using Machine Learning Algorithms, Journal : European international journal of science and technology (EIJST) , ISSN :2304-9693 . Volume 8 issue 1, January 2019

3. Performance Analysis of Regression-Machine Learning Algorithms for Predication of Runoff Time , Journal :Agrotechnology , ISSN: 2168-9881, Volume 8 issue 1 , February 2019, impact factor 1.04

1.8 Published book

1. Irrigation Runoff modeling-Regression Based System, published on 11-4-2019, ISBN -13:978-620-0-00476-5, ISBN-10: 6200004765, EAN: 9786200004765, Published by Lambert academic publishing, number of pages: 184. Author by Marwan khan

Chapter 2: Literature Review

As discussed in Chapter 1 that demand for freshwater is increasing day by day due to the rapid growth of the world's population. The effects of global warming and climate change also pose a serious concern to water consumption and food security. As a result, water for irrigation purposes are intensely used by many farmers around the world, with the increasing concern about their inefficient and wasteful consumption. Due to the scarcity of water resources around the world, the need for optimum utilization of water resources has increased. The irrigation system requires more consideration to improve the optimum consumption of water in agriculture. The literature gives a brief insight to application of WSNs and IOT in precision irrigation, hydro-agricultural simulator and machine learning algorithms that are being in hydrology.

2.1 WSN and IOT in Precision Irrigation

The innovative concept of WSN has emerged with the advancement of wireless communication, sensing devices, and low power hardware. WSN comprises of tiny devices called nodes. These devices are distributed over a region and work autonomously to monitor the environment to gather the information. These nodes are further categorized into two types, a source node that gathers the information, and the other is sink or gateway node, which gets information from the source node. A sink node has more computational power contrasted with a source node.

WSNs have arisen as a fundamental driver for automation in precision agriculture, hydrological management, and monitoring. WSN gives the establishment to IoT frameworks and supports in monitoring and sending the conditions of the environment. Internet of Things (IoT) - the term originally begat by K. Ashton in 1999. The IOT broadens the web and extend internet beyond PCs and cell phones to an entire scope of different things, and environment. IOT has arisen to depict an organization of interconnected gadgets - sensors, actuators, cell phones, among others - which communicate and work together with one another to achieve basic targets. IoT will turn into the most unavoidable innovation overall [25].

A DSS for irrigation management uses different strategies for site-specific irrigation decisions. An ‘open-loop’ strategy based DSS schedule irrigation at predefined intervals with predefined irrigation volumes. Open-loop approach is an inefficient technique as they do not consider any sort of sensor feedback on plant water status, soil moisture condition and climatic parameters. This strategy is based on historical data and heuristics which often result in overwatering of plants and wastage of fertilizer .In contrast to open loop strategy, closed-loop irrigation strategies aim to irrigate when the soil moisture content reaches a specific threshold or when plant sensors indicate a specific stress threshold. These closed-loop irrigation strategies have been shown to boost water use efficiency as compare to open loop strategy [26].

To make optimal usage of water resources, the precision irrigation integrate information, communication, and control technologies in the irrigation process [27]. The integration of IoT for data acquisition and monitoring, control theory, and decision support technologies in irrigation management are very vital for efficient precision irrigation system . The precision irrigation must have control technology to reallocate inputs and adjust irrigation management according to the plants response to ensure optimal water-saving.

WSN integrate conventional soil matric potential monitoring with wireless communication, to transfer real-time data and irrigation management [28]**,** [29].IoT and advanced control strategies are being utilized to achieve improved monitoring and control of irrigation farming [30]. IOT based solutions are aiding in data acquisition and intelligent processing, while bridging the gaps between the digital and physical worlds. IoT based smart irrigation decision systems can help in accomplishing optimum water-resource utilization in the precision irrigation. Innovate smart system is used to predict the irrigation requirements of a field using the sensing of parameter like soil moisture, soil temperature, and environmental conditions as well as the weather forecast data from the web. Efficient monitoring system for various parameters that affect the plant growth and development is extremely imperative towards planning an effective and efficient irrigation system to improve food production with minimum water loss. WSN and IOT technology has made it possible to monitor the environment and collect data that reflect real-time status of soil, plant, and weather of the irrigation areas of the plants. With rapid success of IOT, it has become possible to develop a real-time monitoring system, by utilizing low-cost sensors and communication technologies for the irrigation process [31].

IOT-based WSN has been used in agriculture to monitor the field condition by using various sensors. These sensors are deployed in the agricultural environment to improve production yields through intelligent farming decisions and to gather information regarding crops, plants, temperature measurement, humidity, and irrigation systems [32].A smart irrigation system was presented in [33] in which a Raspberry Pi was used alongside with soil moisture , temperature and humidity sensor. These sensors and the Raspberry Pi were interconnected with irrigation system. A mobile application was developed empowering both manual and automatic water flow control. In automatic mode, water flow was automatically turned ON/OFF based on the status of the soil moisture. In manual mode, the client had the option to screen the soil water status. An alarm was produced when the water level of soil was getting under a particular threshold, and the client turned it ON/OFF utilizing a portable application.

In [34], an IoT-based irrigation system framework was introduced utilizing soil moisture sensors obliged by ATMEGA 328P on an Arduino UNO board alongside a GPRS module. The data assembled from the sensors were delivered off the cloud i.e., Things Speak, where diagrams were created to visualize the information patterns. A web-based interface was additionally developed where the farmer had the option to check the status of water, in the event that it was ON/OFF. Real Time model for irrigation system framework was introduced in [35] in which soil moisture sensors and soil temperature sensors were utilized to evaluate the water status of the soil. RFID was utilized to send information to the cloud for additional information examination. Utilizing ATMEGA 328, a water sprinkler framework for irrigation system was introduced in

[36] utilizing temperature, humidity, and soil moisture sensors. The water sprinkler was controlled based on the soil moisture level to spare water and lessen human efforts. In [37] a cost-effective drip irrigation system framework for a house was proposed in which a Raspberry Pi, Arduino, electronic water control valve and hand-off were utilized. ZigBee were utilized for correspondence. The user turned ON/OFF the water valve by sending commands to the Raspberry Pi, which further processed the commands through the Arduino. For real time irrigation system framework, complete equipment and programming necessities, issues and challenges and focal points were inspected in [38] where a big picture of the total framework was given.

Machine-learning and AI techniques have incredible potential to unlock the worth of big data for irrigated agriculture [26]. Present state-of-the-art irrigation DSS schedule the irrigation process by integrating predictive process-based crop models with pre-defined triggers (e.g. soil moisture targets). Machine-learning methods have the ability to learn from historical data, this capability of ML enable them to boost the effectiveness of those irrigation control approaches. specifically, ML could be utilized to develop optimal adaptive spatial and temporal real time irrigation system [39].

ML techniques e.g. ANN,MLR,SVR,DT trained on historical datasets also offer opportunities for real-time prediction of optimal irrigation decisions based solely on observation data from sensors. Deep machine-learning approaches have been applied successfully for prediction of relevant hydrological processes such as soil moisture and groundwater levels [40] [41]. ML based an automated DSS [42] was proposed to manage the irrigation on a certain crop field, considering both climatic and soil factors gave by climate stations and soil sensors. The work emphasizes the significance of soil sensor data, the utilization of which achieved a 22% decrease in weekly error contrasted with utilizing just climate data.

2.2 Agricultural (CROP, SOIL) and Hydrological Simulators

DSSAT stands for decision support system for agrotechnology transfer. The researchers from across the globe has used it since last 15 years. It supports a software for the evaluation and application of 16 different crop models for different purposes. The newly re-design DSSAT cropping system (CSM) has adopted a modular approach . It has soil, crop, weather modules, and a module related to light ,water and their interactions with agricultural environment. The primary focus of the DSST-CSM simulator is the simulation of crop growth and yield over a chunk or an area of land and also the soil water, carbon and nitrogen modifications that occur over time beneath the DSSAT-CSM system. CropSyst stands for cropping simulation model . It follows system based approach . CropSyst is a cropping machine simulation version which is based on multi -year, multi- crop and each day time step. CropSyst consist of a set of applications that is as comply with:Cropping system simulator (CropSyst),A climate generator (ClimGen), GIS CropSyst co-operator software (ARC-CS),A watershed model (CropSyst watershed) and other numerous miscellaneous applications. The gain of CropSyst over the DSSAT is as, one DSSAT method has been gradual in adaptation to more general simulation platform that might permit the consumer to without problems concatenate these models CropGro and Ceres and many others and allow to simulate crop rotations. The second gain of CropSyst over DSSAT and EPIC model is a more potent emphasis on software layout. In short, CropSyst has been used extensively to predict soils, weather , agricultural control impact on crop yield, drought version , nitrogen stability , water and lots of different problems related to cropping system in distinct parts of the world [43]. APSIM [44] stands for Agricultural Production Systems Simulator. It is basically a model developed for farming system simulation. It follows modular approach. It has been developed by the Agricultural Production Systems Research Unit in Australia. The APSIM has been developed for the simulation purpose of biophysical processes in the faming or agricultural systems keeping in view the important factors such as ecological and economic outcomes from the agricultural management practice keeping in view the climate risk factor as well. The APSIM advantage over the DSSAT and other models is the more detailed assessment of the future management strategies for the farmers. The performance of the real farmers to effectively simulate there is a need of a model, as from year to year basis there is changes took place to unfold in the climatic and environmental conditions needs an effective decision tree and management model. None of the above models has the ability to do this. The other Simulator is Aquacrop [45] which is water driven crop simulator model for a farm scale which is distinct from other well known crop models such as BOSFOST which are radiation driven [46] .HYDRUS 1D,2/3D this simulator aim is the analysis of flow of water in a porous media at plot scale [47].SWAT is an hydrological and erosion model for large scale catchment [48] ,WATEM/SEDEM this is another simulator model for water , tillage erosion model at small catchment scale [49] while CSLE is another simulator for Chinese soil loss equation model for field scale [50], [51]. These simulators are either related to crop growth or yield predication , bio physical processes modelling , erosion modelling and water flow modelling inside soil. The simulator uses alot of inputs variables to simulate the output variables and are complex in operation.

The most relevant simulator for agricultural runoff predication from irrigation or rainfall for this research problem , which uses less input and output variables to acquire the simulated dataset is the USDA-NRCS simulator/model which stands for United States Department of Agriculture - Natural Resources Conservation Service. NRCS is described in next section,

2.3 SCS-NRCS

It is an agency of United States of America which help to assists farmers on agricultural technical aspects. SCS –NRCS stands for soil conservation service-natural resources conservation service. The SCS-CN model is simple, empirical model with clearly stated assumptions and few data requirements. Therefore, it has been widely used for water resource management, storm water modelling and runoff estimation for single rainfall events in small agricultural or urban watersheds.

SCS-CN soil conservation service-curve number method consists of water balance equation. the water balance equation consists of following terms 1) (P-Ia-Q) actual moisture retention by the soil 2) {SM, where SM≥ (P- Ia- Q)}; the potential maximum retention 3) (Q) the actual runoff taking place 4) {P-Ia where (P-Ia) ≥ Q)} and the potential maximum runoff [52].

The first equation is as follows:

$$\frac{(P-I_a-Q)}{SM} = \frac{Q}{P-I_a} \qquad (1)$$

(since the ratios between the actual and potential moisture retained and the actual and potential surface runoff should be equal). The second equation is as follows:

$$Q = \frac{(P-\mathrm{I}_a)^2}{(P-\mathrm{I}_a)+\mathrm{S}_\mathrm{M}} \qquad (2)$$

The second equation is as follows, which relates the initial abstraction (Ia) to the potential maximum retention (S). The third equation is as follows:

$$Q = \frac{(P-0.2S_M)^2}{(P-0.2S_M)+S_M} \qquad (3)$$

The fourth equation is as follows:

$$Q = \frac{(P-0.2SM)^2}{(P+0.8SM)} \quad (for\ p \geq 0.2S) \qquad (4)$$

The fifth equation is as follows:

$$Q = 0\ (for\ P \leq 0.2S) \qquad (5)$$

The sixth and final equation is as follows:

$$S_m = 5[P + 2Q - (4Q^2 + 5PQ)^{1/2}\ ] \qquad (6)$$

Curve number derivation

In the National Engineering Handbook, section 4 (NEH-4) it has been discussed in detail about how the different curve numbers values have been constituted for a watershed under various conditions under different surface conditions the CN values has been obtained for different land use types which are linked with a particular antecedent moisture condition [53], [53] (AMC) antecedent moisture condition which is a synonym for (ARC) antecedent rainfall condition. In the technical release documentation (TR- 55) AMC is referred as ARC[24].There are three classes of AMC I,II,III which have been given for corresponding dry season, moderate season and very wet season in relation to AMC based on 5 days antecedent rainfall before storm under consideration during a particular dormant or growing season . The TR-55 explains the application of SCS-CN method for small urban watersheds. CN values for average (ARC-II),dry (ARC-I) and wet (ARC-III) are 12.7-27.9, <12.7 and >27.9 for dormant season , 35.6-53.3, <35.6 and >53.3 for growing season in correspondence with total 5 days antecedent rainfall (mm)

Derivation of Runoff

In the TR-55 documentation the CN value for AMC II has been mentioned .AMC I and AMC III has been obtained from AMC II conversion.[26]

Equation 7 is as follow.

$$\boldsymbol{CN_1} = \frac{\mathbf{4.2} * \boldsymbol{CN(II)}}{\mathbf{10} - (\mathbf{0.058} * \boldsymbol{CN(II)})} \qquad \textbf{(7)}$$

Equation 8 is as follow .

$$CN_{III} = \frac{23 * CN(II)}{10 + (0.13 * CN(II))} \qquad (8)$$

Here CN I is curve number for dry condition , CN II is curve number for normal condition and CN III is curve number for wet condition.

Equation 9 is as follow.

$$S = \frac{25400}{CN} - 254 \qquad (9)$$

$$S = \frac{25400}{CN} - 254 \qquad (10)$$

The CN value is substituted for S value with respect to each condition and the runoff (Q) is obtained from equation 4

The runoff (mm) for each landuse element was computed using the following formula; Equation 10 is as follow:

$$Q = \frac{Q_t * A_t}{A} \qquad (11)$$

Hydrologic soil groups

Musgrave in 1955 documented about hydrologic soil groups (HSG) in a handbook of agriculture (USDA), The soil are classified into four groups A,B,C and D. The four groups are defined by SCS soil scientists .Group A soil have low runoff potential and high infiltration rates even when thoroughly wetted. Group B soils have moderate infiltration rates when thoroughly wetted and consist chiefly of moderately deep to deep, moderately well to well drained soils with moderately fine to moderately coarse textures. Group C soils have low infiltration rates when thoroughly wetted and consist chiefly of soils with a layer that impedes downward movement of water and soils with moderately fine to fine texture. Group D soils have low infiltration rate and high runoff potential [54], [55].

The SCS NRCS model has been used in various hydrological applications such as 1) mountainous watershed [56]. 2) SCS predication of Rainfall-Runoff using SCS-CN Method with Remote Sensing and GIS Techniques [54]. SCS -CN and Geographic information system-based method for identifying potential water harvesting sites in the Kali Watershed [57].4) Surface Runoff Depth by SCS- CN approach combined with Satellite Image and Geographic information Techniques [55]. Simulation using a continuous SCS CN method- based hybrid hydrologic model [58].

2.4 Hydrological and Data Driven Model

Hydrological models give us a wide scope of huge applications in the multi-disciplinary water assets management and planning. Soil water distribution and variation are helpful in predicting and understanding various hydrologic processes, such as, rainfall/runoff process, irrigation scheduling and climatic changes. Soil water content forecasting is essential to the development of irrigation systems. In hydrology the fundamental and essential variable is soil moisture. Soil moisture data is fundamentally significant for many application areas such as irrigation scheduling, rainfall/runoff generation processes, climate investigations, reservoir management, crop yield forecasting, meteorology, , and natural hazards predictions. Prior knowledge on soil water content conduct cannot just assistance in better management and comprehension of hydrological frameworks but it also helps in improved forecasting, particularly in precision irrigation.

Physically based models have showed tremendous capabilities to predict a various range of hydrological events, like storm, rainfall/runoff, hydraulic models of flow, floods prediction. However, there's still a gap in short-term prediction capability of physical models. They need different forms of hydro-geomorphological monitoring datasets, requiring intensive computation, which are main constraint in short-term prediction. Furthermore, physically based models also require in-depth knowledge of hydrology, which is quiet challenging. data-driven models, e.g., machine learning (ML) have been used as an alternative to physical models in order to overcome the shortcoming associated with physically based model. ML models numerically formulate the nonlinearity, solely supported historical data without requiring knowledge about the underlying physical processes [59]. Data-driven prediction models using ML are significant tools as they're simple to develop and require minimum inputs. ML models are comparatively less complex than physical model with fast training, validation, testing and evaluation with high performance. The continual advancement of ML methods over the last twenty years demonstrated their suitability for hydrological processes with a suitable rate of outperforming conventional approaches.

Hydrology is the study of water which investigates the life cycle of water and its movement in term of geographical features. The quantity of runoff produced is very vital for managing and coping with water sources [60]. One of the essential elements in water resource management is runoff. it’s important to have correct expertise of watershed and its components to properly model runoff from precipitation process [61]. Hydrologic models are normally utilized for runoff assessment. The overflow cycle is started with precipitation on a watershed. There are certain processes which take place before the water can run all the way down to the channel stream and closer to downstream. Some of the rainfall water evaporate from land, plants etc and return to atmosphere. A segment of it penetrate into the soil depending on soil type, floor cover, predecessor moisture and characteristics of watershed [62].

As a result of overirrigation, excessive water is stored within the trees’ root zone, which result in greater tailwater runoff problems. There are numerous factors that, directly or indirectly, influence the irrigation runoff, which make it a very complex process to analyze [63]. The irrigation system water application rate is a main consideration that influence runoff [64]. The properties of the irrigation system and properties of field are the principle factors that influence the surface runoff. The irrigation system factors determine water application rate and depth, which if it's not viable with the soil penetration limit can bring about runoff. The soil penetration limit will depend upon various soil parameters and along these the properties of the irrigated field that decide its surface storage limit. Other factors, like meteorological factors (wind and air temperature) [65] and crop canopy [66] can likewise influence the runoff. Saturation excess and infiltration excess lead toward runoff generation. In saturation excess the soil turns out to be completely soaked with water, surpassing the water holding limit of the soil; when the excess precipitation can not, at this point be held in the soil, then the surplus rainfall can no longer be held in the soil, the water is directed to another location through overland flow [67]. Infiltration excess occurs when precipitation intensity surpasses the maximum rate that water can infiltrate into the soil, and water must flow over land to a different area [68].

A review of the literature by the authors showed that, despite the development and advancement of precision irrigation system, there is very rare attempt to use precision irrigation system at two or multi form level.

After literature review it is cleared that previous studies focused only on individual farm monitoring, a gap exists in determining how precision irrigation system can be established between two farms. Runoff is the water from rain or irrigation that is not absorbed and held by the soil but run over the ground and through loose soil. Runoff is usually associated with negative implication such as erosion, water loss etc. It can however be used for irrigation of crops. Efficient and intelligent use of runoff water for irrigation in nearby farm could help to save water. Modeling runoff can help to understand, control, and monitor the quantity of water resources. Hydrological information such as precipitation, temperature (for evapotranspiration estimates), (runoff), water storage can be evaluated to design optimal decision support system between two farms.

Over the past several years, various attempts have been made to produce soil water content estimates by using different machine learning algorithm, such as MLR, ANNs and Decision Tree. In the field of hydrology various research have been conducted for runoff forecasting. In next section, the four machine learning algorithms i.e MLR, SVR, DT and ANN which have been used in this thesis for runoff prediction are discussed.

2.4.1 MLR (Multiple Linear Regression)

MLR is the expansion of Simple Linear Regression to the case of multiple explanatory variables. MLR is the method of setting up relationship between a dependent variable "y"& set of independent variables x1 , x2 , x3 ...xn, governing a phenomenon.

In hydrological domain, runoff is viewed as subject to precipitation at various stations [69] MLR was utilized to show the connection between independent variables (Temperature, Precipitation) and a dependant variable (Runoff) in the Litani River in Lebanaon [70]. Another study [71] tested MLR model for rainfall runoff modeling on data sets for the river Jhelum catchment (J&K, India).The performance was checked by using different It was observed that the MLR model got simulated very well with a small value of MSE, RMSE a high value of R2, revealing that the model is quite efficient in predicting the discharge of river Jhelum.

Another investigation [71] tried MLR model for precipitation runoff modeling on dataset for the Jhelum catchment (J&K, India). The performance was checked by utilizing diverse statistical assessment measurements. It was seen that the MLR model got mimicked very well with a little estimation of MSE, RMSE a high estimation of R2, uncovering that the model is very proficient in anticipating the discharge of stream Jhelum. Another study made a correlation among various model to make expectation about monthly flow in river. MLR as a statistical method, ANN and adaptive neuro-fuzzy inference system (ANFIS) as non-linear ones and K-nearest neighbors (KNN) as a non-parametric regression method were utilized to forecast the monthly flow in the St. Clair River between the US and Canada. Various scenarios for input combinations were characterized to contemplate the impact of various information on the results. Performances of the models are assessed utilizing statistical metrics as the performance criteria. Results obtained showed that adding lag times of flow, temperature, and precipitation to the to the information sources improve the precision of the predictions significantly. Further the performance of models was improved by using wavelet transform [72].

Reference evapotranspiration is a significant factor in hydrological cycle as it can affect the amount of runoff and irrigation water needs.Least square based method of MLR and Penman-Monteith model were used for the estimation of reference evapotranspiration in the Megecha catchment. Multiple climatic and environmental variables were given as inputs to the model which can assess the impact of each variable on ETo. The inputs variables were air temperature, solar radiation, humidity and wind etc. The strongly positive co related variables were maximum temperature, wind speed and sun hour for reference evapotranspiration. It is showed that relative humidity is negatively correlated. The MLR gave coefficient of determination is =0.92 and residual error of 0.26 mm/day for the meteorological station considered. It clearly showed that MLR represent a linear trend and the input variables of the model were fitting. Thus, MLR could be utilized in the estimation of monthly reference evapotranspiration successfully [73]. Wind drift and evaporation losses (WDEL) played significant role in building water conservation strategies in sprinkler irrigation. In research [74] MLR and ANN model were used to predict WDEL. The data were collected from several published work on WDEL estimation and its relevant operational, design and meteorological conditions of parameters in sprinkler irrigation. Overall five combinations of input variables were used which were air temperature , relative humidity , wind speed (WS), water discharge by auxiliary nozzles, water discharge by main nozzle, auxiliary nozzle diameter(da), riser height and operating pressure for WDEL model building and its estimation . 70% and 30% data split were used for model training and testing. The benchmark metrics used were coefficient of correlation (r),MAE,RMSE and overall indices of model performance (OI). The results of modeling showed that (r) values for ANN was higher and MLR was lower the values were 0.84-0.95 and 0.79-0.86. The RMSE were 2.662%-4.886% for ANN was lower and higher for the MLR which was 4.56% and 5.51 %. MAE values were ranged 2.19%-3.72% and 3.51% -4.41% for ANN and MLR respectively. OI (0.79-0.90) and (0.74 and0.81) for ANN and MLR respectively. Two input variables design parameters (da) and climatic parameter (WS) had significant impact of WDEL estimation. Thus, ANN is superior to MLR in predication of WDEL from sprinkler irrigation.

2.4.2 SVR (Support Vector Machine)

SVM is Kernal based machine learning model which is used for regression and classification. SVR is the altered version of SVM, where the dependent variable is numerical in instead of categorical. SVR is a non-parametric technique and allows the creation of nonlinear models. The SVR method utilizes kernel functions to generate the model. Some of the frequently used kernel functions are Polynomial, Linear, Radial Basis and Sigmoid.

SVR has been effectively applied in the fields of water resource engineering and hydrology for purposes, for example, runoff prediction, flood estimating, lake water level forecast. SVRs are able to learn more effectively when using scarce and incomplete hydrologic data. This advantage is because of two outstanding features of SVRs: their excellent capability in generalization of the unseen data (testing phase) and their proficiency for application in large scale problems using only a small number of support vectors [75].

Precipitation modeling is very important in various disaster management such as flood and drought. As heavy rain could lead towards flooding while low rain could lead to drought. Different regression based machine learning algorithms such as MLR, Lasso Regression (LR) and SVR were used for precipitation predication for rainfall runoff modeling. SVR showed better performance than the rest of models on the basis of benchmark metrics which were MAE and R^2. The MAE and R^2 values were 10.9 and 0.99 for MLR, 4.3 and 0.99 for SVR and 11.7 and 0.99 for LR machine learning models [76].

Research study [22] proposed a runoff prediction method that combinedly use ANN and SVR to predict runoff. The strength of ANN is its high accuracy on predicting runoff when the amount of rainfall is high. The decision criteria for choosing either ANN or SVR model was the amount of rainfall on the previous month. If this amount was higher than the threshold value, then ANN model was chosen; otherwise, the SVR model was used. The performance of proposed combined model was compared on basis of R and RMSE metrics with ANN, SVM and LR. From the experimental results, the proposed method shows good efficiency to predict runoff against other models.

Another work [77] presented a comparative study of rainfall-runoff modeling between a SVM-based approach and the Storm Water Management Model (SWMM). The performance was evaluated through RMSE and R metrics. The two models show comparable performance. Both models properly model the hydrograph shape and the time to peak. In the simulated events, the results of SVR were slightly better as it showed higher values of the coefficient determination R and lower values of RMSEs as compared to the results of SWMM. However, SVR algorithm tends to underestimate the peak discharge by 10%, while SWMM tends to overestimate peak flow rate by 20%. Both models generally tend to overestimate the total runoff as well.

In research work [78] a comparative study has been done for various machine learning algorithms i.e. MLR, SVR, ANN and Extreme Learning Machine (ELM). The focus of machine learning models is deriving the hydropower Reservoir operational rule. operational rule helps in management of hydropower reservoir scientifically and also provide help to operators to take appropriate decision with limited runoff predication information. The data is taken from Hongjiadu city in china from 1952 to 2015 are selected for this modeling. The RBF kernel is selected for the SVM, while sigmoid function is selected for ANN and ELM. The modeling shows that SVM, ANN and ELM perform better than conventional ML. This study [79] described the design of automatic irrigation scheduling and the use of machine learning models for an efficient decision support system. Nine orchards were tested during 2018.The machine learning models used were Linear regression (LS), Random forest regression (RFR) and support vector regression (SVR) for Irrigation decision support system (IDSS). In the LR (stepwise, forward, and backward) were used for final estimated model. The results showed that regression models are substantial for designing automatic irrigation scheduling system.

2.4.3 Decision Tree

A decision tree is a tree structure consisting linked internal and external nodes dividing the input set into mutually exclusive regions. A mark,a value,or an action which characterizes its input is assigned to each of these regions. The internal nodes, known as decision-making unit, assess a decision function to determine which child node to visit next. Nodes associated with labels (irrigation vs non irrigation) that categorized the question are called leaves or terminal nodes (external node) and have no children. There are two main types of decision trees: regression and classification. The leaf node labels in regression trees are constants or equations that specify the forecasted output value of a given input vector. However, the leaf nodes of the decision trees in classification trees contain a label that indicates the group or class (J) (e.g. irrigation or non-irrigation events) to which a given feature vector belongs [80].

M5 model tree is the is the most acclaimed algorithm of decision tree which have the qualities of the classification and the regression methods. Numerous studies have been done with respect to the effectiveness of M5 model tree in simulation of hydrological measures. The accuracy of M5 model tree is comparable to the classic ANN model. Estimation of discharge in a river is very important in flood management. Rating curve is used to establish relationship between the water level (also called the stage) and discharge. ANN and M5 model trees were used in [81] to predict discharge in an Indian river. Historic data for the period 1990 to 1998 was used to establish a relationship between the water level (stage) and discharge. The prediction accuracy ANN and M5 model trees were superior as compare to traditional rating curves.

Research study [82] make a similar investigation of two data-driven models ANN and model trees (MT), in rainfall–runoff transformation. The outcome showed that with short lead time (1 hour) both models performed very well for runoff estimation. While both models struggle to show good result for runoff predication with higher lead-time (6 hours).The presentation of ANN is somewhat in a way that Was better than MT for higher lead times. The disadvantage of ANN is that they are not effectively interpretable.Statistical model and three distinct data driven model strategies were utilized to stream level estimating [83]. The outcomes demonstrated that data driven methodologies performed better than statistical methodology. The performance was assessed through RMSE, MAE, Coefficient of Efficiency (CofE), and R^2. The outcome presumed that M5 model trees are competent for the development of transparent stream/river level gauging models.

MT based approach is very simple, very fast in training and its result is simple and easily interpretable. In this research M5 tree is used .M5 tree-based models have multivariate linear models in the leaves .M5 trees are more suitable for high dimension data. They can tackle tasks up to hundreds of attributes. Model trees have an advantage over regression trees in terms of predictive accuracy. Furthermore, the model trees are able to make predictions lying outside the range observed in the training cases, which is not the case with regression trees. M5 model trees also used as modular models forming committee machine. The reuse and recycling of nutrient-rich drainage water can be a valuable strategy to gain economic-environmental benefits. Another research [84] proposed a simplified data driven discharge (Q) prediction model and response time predicative models (t1 and td) by employing M5 trees. The proposed model can work with resource-constrained system thus making it more suitable for Wireless Sensor Network. They propose systems that proactively control irrigation strategies and reuse drainage water among multiple farms in a catchment. The proposed model uses minimum parameters derived from existing NRCS model. The significance of the proposed model can be judged from the fact that it gives higher accuracy for Q and td models that is 94 % while t1 model gives comparatively better results with 84 % of accuracy.

The GBRT (Gradient Boosted Regression Trees) is a type of additive model that makes predictions by combining decisions from a sequence of base regression tree models. Irrigation recommendations system based on machine learning algorithm with support of agronomist's encysted knowledge was proposed in [85]. Different regression and classification algorithms were applied on dataset to develop models that were able to predict the weekly irrigation plan as recommended by the agronomist. By comparing the resulting models, it was found that the best regression model in terms of RSME was GBRT with 93% accuracy in prediction of irrigation plan/recommendation., and the best classification model was the Boosted Tree Classifier, with 95% accuracy (on the test-set). The developed model is helpful to the agronomist's irrigation management.

This research thesis [86] address the issue of forecasting of the water waste, measure of depleted water and estimation of the basic times of seepage occasions in a horticultural field by utilizing machine learning and data mining techniques. The model utilized various parameters like crop growth stages, day of the period, slope of the field, precipitation, temperature, overflow, water waste to quantify water seepage releases from fields. ML models are used to predict the amount of drained water from a field. The ML approach includes predictive models that represent the obtained knowledge in patterns such as model trees and regression trees, ensembles, and polynomial induced equations. Regression and model trees patterns were used because they express the gained knowledge in the most understandable way. Furthermore, the possibility of model trees to include within a rule a linear regression model additionally improves the accuracy of the learned model. While ensembles technique aims to enhance the predictive performance of their base classifier. The proposed system resulted in highly accurate predictive models for prediction of the amount of drained water from a field.

Evaporation is a significant component of the hydrological cycle, and estimating evaporation loss is primarily essential for water resources management, evaluation of irrigation schedule as well as agricultural modeling. The potential of M5 model tree and artificial neural network (ANN) was investigated in [87] for estimating ET_0 in California, USA using MODIS products. The coefficient of determination values of the ANN and M5 tree models were over 0.79 and 0.80, respectively. The results suggested that the M5 Tree model could be successfully applied in modeling $ET_{0.}$

2.4.4 ANNs (Artificial Neural Networks)

ANNs are proficient numerical modeling frameworks with efficient parallel processing, empowering them to mimic the biological neural network using inter-connected neuron units. Among all ML strategies, ANNs are the most mainstream learning algorithms, known to be adaptable and proficient and efficient in modeling complex processes with a high fault tolerance and accurate approximation. ANN approach is utilized for prediction with greater precision as compare to traditional statistical models. ANNs are considered one the most reliable data-driven model as they predict from historical data instead of taking into consideration the physical characteristics of catchment. ANNs are able to construct black-box models of complex and nonlinear relationships of rainfall and flood, as well as river flow and discharge forecasting [88]

The main research challenge in hydrology is to build up the models that can simulate the catchment feedback, so that such models are fit to estimate future river discharge, flood prediction. The advancement and progression of ANN strategy has added another measurement to mimic such frameworks and has been applied as an effective method to tackle different issues related with water resources management. The water resources applications using ANNs which include the simulation rainfall runoff event, climate change, evapotranspiration process, river flow forecasting, reservoir inflow modeling, ground water quality prediction. The water resource applications utilizing ANNs which incorporate the simulation of precipitation runoff process, evapotranspiration measure, stream flow anticipating, climate change repository ground water quality forecast.

The model developed using ANN is easy to implement and result runoff in close association with real values [89]. In research [90] the effects of climate change in the runoff process in the Three-River Headwater Region (TRHR) on the Qinghai-Tibet Plateau were analyzed, ANN models, one with three input parameters (previous runoff, air temperature, and precipitation) and another with two input parameters (air temperature and precipitation only), were designed to simulate and forecast the runoff variation in the TRHR. The ANN model with three input parameters has a significantly superior real-time prediction capability and produces a high-accuracy performance in the simulation and forecasting of the runoff dynamics. At the point when no field perceptions of the runoff were accessible, the ANN model was created utilizing just two parameters (precipitation and air temperature). Two parameters ANNs models also has a good accuracy for simulation and prediction of the variations in runoff**.**

The study [91] developed an empirical Rainfall Runoff model for Rajsamand India Catchment using monthly rainfall and runoff data in mm received during the month of July, August, September and monsoon period respectively for the past twenty years (1996-2015) using conventional regression and ANN approach. 20 years were divided into four parts each having five years' data. 20 years were categorized into four sections each having five years' information. The 15 year data was used for training and remaining data for 5 years was used for validating the trained model. A three layer feed-forward neural network has been used comprising of four input neurons and four output neurons representing the rainfall and runoff. A total 12 laboratory experiments were conducted utilizing rainfall simulator to generate runoff hydrograph using various slope and rainfall intensity over the catchment [92]. For the validation of noticed runoff hydrograph information were reproduce utilizing ANN. The correlation of noticed and anticipated runoff hydrograph uncovered that the ANN predicts the runoff information sensibly well in noticed hydrograph. The outcomes and relative investigation showed that the ANN was more suitable to predict river runoff of a catchment than other classical regression model.

The paper [75] developed another model which joins the SVR model with a geomorphologic ANN model (GANN) to replicate the daily runoff hydrograph in a watershed. The performance effectiveness of the proposed model (SVR-GANN) for simulating daily runoff was contrasted with distinctive ANN-based models including ANN with embedded geomorphologic characteristics (GANN), ANN with genetic algorithm (ANN-GA), ANN adopted with the fuzzy inference system (adapted neuro-fuzzy inference system – ANFIS) and also traditional SVR model from the viewpoints of simplicity (parsimony), equifinality, robustness, reliability, computational time, hydrograph ordinates and saving the main statistics of the observed data.

The proposed model endeavor the effectiveness of the two models at the same time in a unit model. Consolidating the geomorphologic attributes of watershed straightforwardly in the GANN designs advances it from a pure black-box ANN model to a more proficient model with oriented structure mimicking the physical properties of the watershed for daily runoff simulation. This model reduces the trial and error efforts in determining the best ANN architecture which can be time consuming. The GANN model has the minimum parameters which make more simple model. The results showed that prediction accuracy of the SVR-GANN model was in a manner that is in a way that is better than those of ANN-based models and the proposed model can be applied as a promising, vigorous prediction for precipitation runoff modeling.

The soil moisture predication is very important parameter for the agriculturists to monitor plant status and growth. In research study [93] regression based three machine learning models 1) Shallow neural network 2) SVR and 3) MLR has been used for the predication of soil moisture in advance for 1 day, 2 days and 7 days ahead. The research study showed that shallow neural network outperformed SVR and MLR. The benchmark metrics used were mean square error (MSE) and coefficient of determination (R^2). Three data sets were used two data sets from online repository and one from Sensenut device (wireless sensor network). The shallow neural network is basically the use of one hidden layer and one output layer. Rain as an additional parameter improved the results drastically for the predication of soil moisture.

The effectiveness of some data-driven models such as SVR and ANN and combination of them with wavelet transforms (WSVR and WANN) were examined for predicting evaporation rates at Tabriz (Iran) and Antalya (Turkey) stations [94]. For evaluating the performances of these models, four different statistical metrics the RMSE, the MAE, the R, and NSE were utilized. Experimental results showed that ANN was the best model for predictions of PE in both Tabriz and Antalya stations. However, WT did not have a positive influence in increasing the precision of ANN and SVR predictions. Similarly, in another study [20] several machine learning models were compared to assess and simulate pan evaporation at monthly scale at two stations in India. The machine learning models were multiple model –artificial neural network (MM-ANN), support vector machine (SVM), multi-gene genetic programming (MGGP),M5Tree and multivariate adaptive regression spline (MARS) . The bench mark metrics were mean absolute percentage error (MAPE), Willmott's Index of agreement (WI), root mean square error, Nash-Sutcliffe efficiency (NSE),Legate and McCabe's Index (LM) and visual inspection . The MM-ANN and MGGP models were better than rest of models during testing phase to simulate monthly pan evaporation with six inputs parameters more accurately. The results clearly showed that MM-ANN and MGGP Models NSE,WI,LM,RMSE,MAPE are 0.95,0.98,0.80,0.53 mm/month 9.988% at Pantnagarstation and 0.91,0.97,0.72 and 0.36 mm/month,12.2% at Ranichauri station respectively. Thus, these two models will help the agriculturist and hydrologist in the water resource management.

Deep learning techniques are based on ANN. In survey paper [95] a comprehensive study has undertaken, in which different machine learning methods and deep learning method in the area of hydrological processes, climate change and earth systems are described. Deep learning research work is still progressing, and the machine learning methods such as ANN, SVR, DT are already well established in these areas. While hybridization and ensemble are also now gaining popularity for higher performance.

2.5 Summary

The summary of literature review clearly showed that regression which is a supervised learning method of machine learning has been used in hydrology (flood forecasting, stream flow forecasting, rainfall –runoff modelling , drought forecasting) and agriculture water management /irrigation water management(reference evapotranspiration predication , wind drift evaporation losses predication, automatic irrigation scheduling and decision support system and soil predication). The machine learning models that were used multiple linear regression, artificial neural network, support vector regression and regression tree (and many variations of these algorithms and ensembles such as multiple model –ANN and M5 tree, random forest). Multiple linear regression has been used in hydrology , agriculture water management and its modelling , however Multiple linear regression has shown limitation in nonlinear modelling , mostly the hydrological processes modelling has shown nonlinear behaviour or relationship among the different input and output variables. The artificial neural network, regression tree and support vector regression has shown good results in comparison with multiple linear regression in the literature for handling nonlinear data in hydrological modelling. However the support vector regression is a complex modelling approach, in which different parameters are to be selected such as various kernels (RBF, poly, Gaussian, sigmoid etc) and regularization parameter C for an efficient modelling The selection of high dimensional kernel for hydrological modelling is a complex task and it could generate too many support vectors which could reduce the training speed . SVR is also not suited for large datasets and noisy data. unlike regression tree SVR cannot be interpreted easily. Decision tree such as Regression tree,M5 tree and its ensemble such as random forest has been used in hydrology and it has shown good results , However artificial neural network has also shown excellent results in hydrology and particularly Multiple model –ANN has shown better results in hydrology than M5 tree (decision tree). Further Recently Deep learning (DL) and Extreme learning machines (ELM) based on artificial neural networks has been utilized in the hydrological processes modelling which has shown promising results than other machine learning models. In the deep learning the research is still progressing.

Chapter 3:

This chapter presents the conceptual models of proposed system. Section 3.1 presents the conceptual models designed in UML (Unified Modeling Language) which is partially used in proposed decision based system. While Section 3.2 presents the proposed decision based system, which is further described in Chapter 4:

3.1 Conceptual modeling in UML

There are further three sequence diagrams.

1) In absence of rainfall *(UML sequence diagram)*

In the absence of rainfall, there is no rainfall considered in this scenario to be built-up, the only way to water the farms are through reservoir water, which is the irrigation water.

2) In presence of rainfall

In the presence of rainfall, there is rainfall considered in this scenario to be built-up , to water the farms are through the irrigation water (reservoir water) as well as the utilization of rainfall –runoff water. There are further two sub –scenario considered in presence of rainfall.

- Presence of rainfall and farms needs water (dry)*(UML sequence diagram)*

In this sub -scenario there is a continuous rainfall in both of the farms that is farm 1 and farm 2 at the same time , both the farms are dry initially and both of the farms needs water to meet its irrigation requirements.

- Presence of rainfall and farms are already saturated (wet)*(UML sequence diagram*)

In this sub-scenario there are rainfall in both of the farms that is farm 1 and farm 2 at the same time and both the farms are wet initially, Thus require no further water.

1) In the absence of rainfall

In absence of rainfall, the sequence diagram for irrigating the farms from reservoir water is as follows.

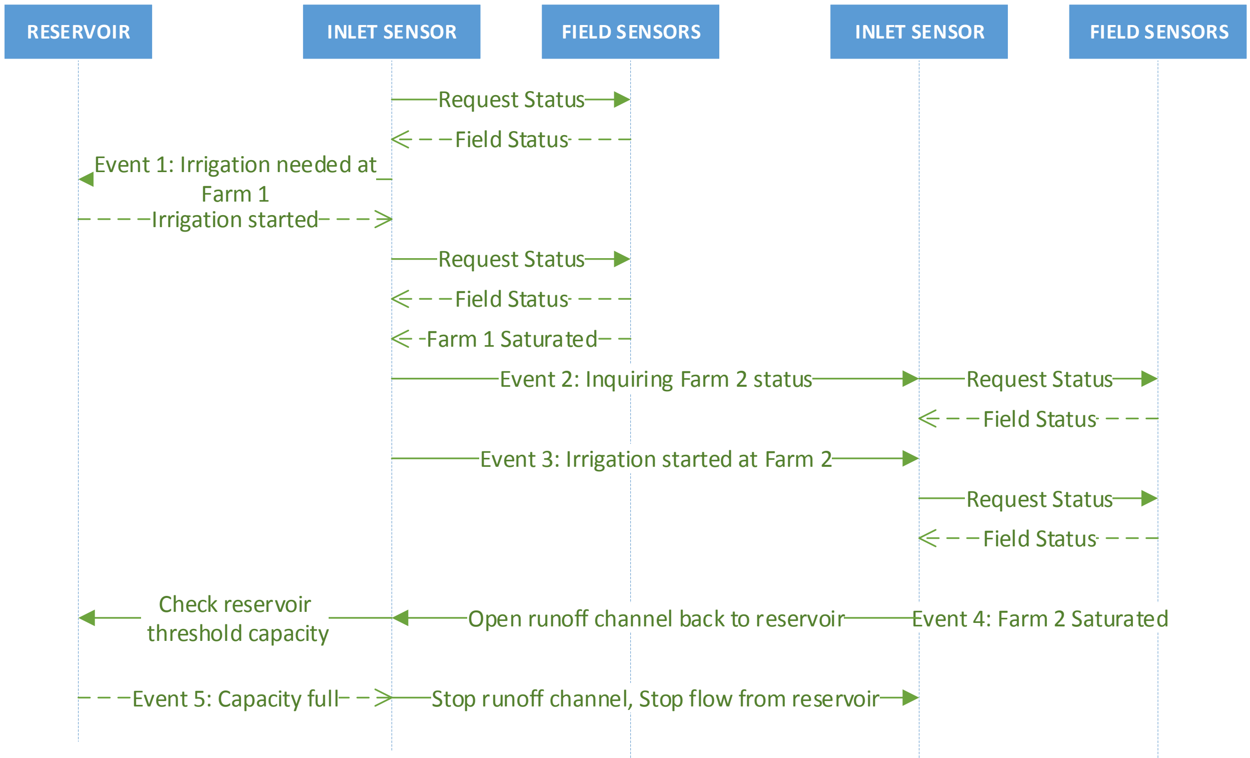


Figure 3.1 UML sequence diagram 1

Step 1: Inlet sensor of farm 1 inquire the fresh hydrological status of the farm, to check if it need irrigation.

Step 2: The reservoir's outlet is opened as a result of event 1, which starts irrigation of farm 1. Meanwhile the inlet sensor of Farm 1 periodically checks the farm is fully saturated.

Step 3: If farm 1 doesn't need any further watering, event 2 will occur to inquire about the farm 2 updated hydrological status.

Step 4: Event 3 opens water to the farm 2 from the outlet of farm 1.

Step 5: When the hydrological status of farm 2 is met up, event 4 will cause opening back-runoff channels to the reservoir from both farm 1 and 2.

Step 6: The threshold capacity of the reservoir is periodically checked. When reached, event 5 will cease back-runoff channels to the reservoir.

2) Presence of rainfall when the farms needs water

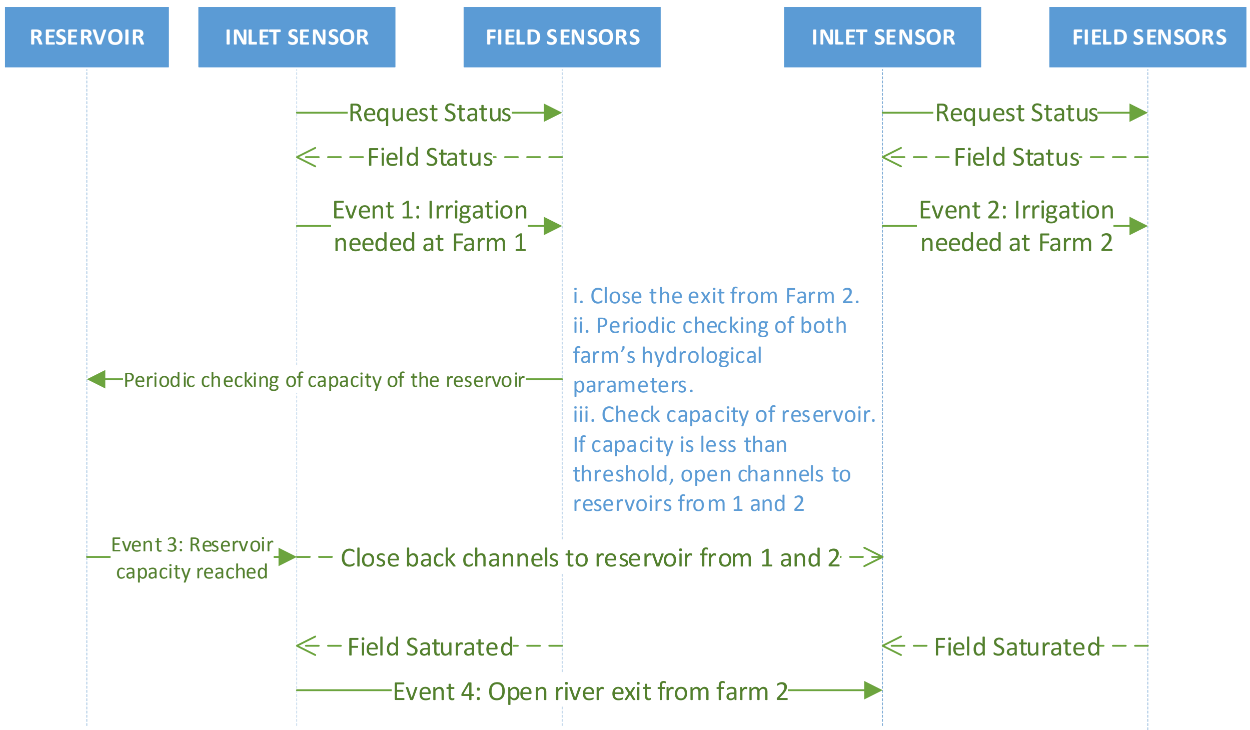


Figure 3.2 UML sequence diagram 2

When rainfall is occurring in the farms under consideration are in a low moisture state, the following sequence of events will be performed.

Step 1: The inlet sensors of both Farm 1 and 2 will inquire about the updated field status.

Step 2: *Event 1* generated at inlet sensor of farm 1, while *event 2* at farm 2. These will cause the river runoff from Farm 2 to be closed.

Step 3: Both farms are periodically checked by their own inlet sensors for determining if they have been saturated with water.

Step 4: At the same time, the reservoir capacity is also checked if it is less than the threshold value. If found less the back-runoff channels from both Farm 1 and 2 are let open.

Step 5: When the reservoir capacity is reached, *event 3* will cause the back-runoff channels from farms 1 and 2 to be closed.

Step 6: *Event 4* will trigger up when the hydrological conditions of Farm 1 marks the farm as fully saturated. Similarly, *event 5* will occur when Farm 2 gets fully saturated.

Step 7: As both farms now don't need any further water, the river channel from farm 2 is opened.

3) Presence of rainfall when farms are already saturated

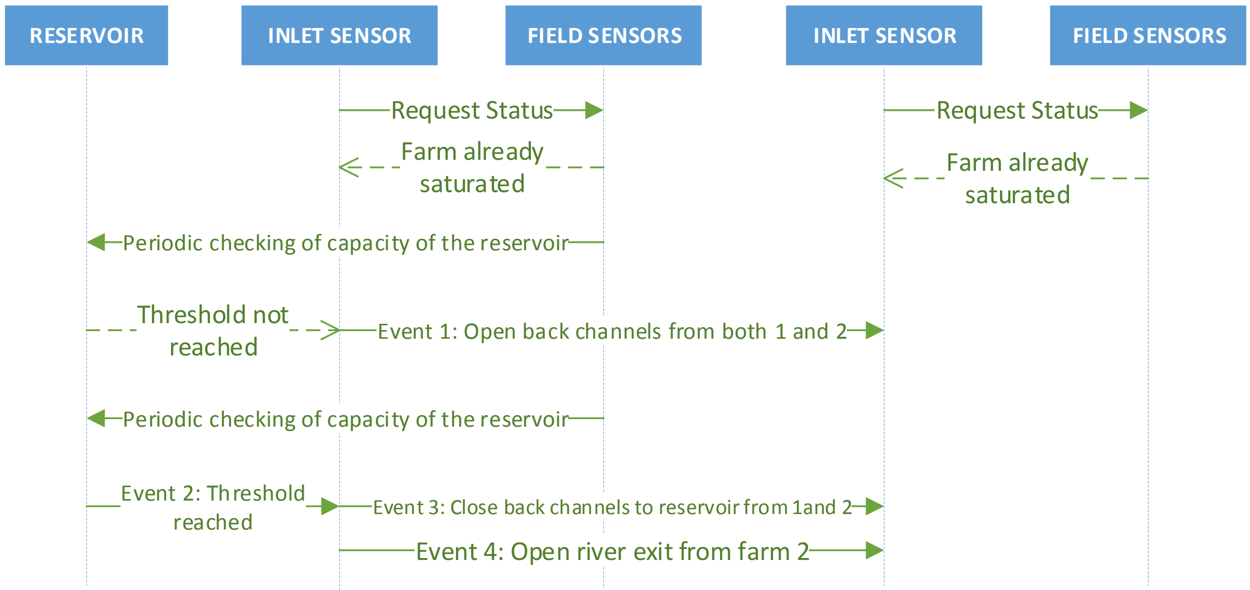


Figure 3.3 UML Sequence diagram 3

In presence of rainfall, when the farms are already saturated with irrigation, the following steps will occur.

Step 1: Inlet sensor of Farm 1 inquires about the hydrological condition of the farm. At the same time the inlet sensor of Farm 2 also gets the fresh status of the farm. Under this scenario, the farms are having a saturated condition.

Step 2: Then the reservoir capacity is checked if it is less than the threshold value.

Step 3: Upon *event 1*, both the back-runoff channels from Farm 1 and 2 are opened into the reservoir, if the reservoir has a low capacity.

Step 4: Once the threshold capacity of the reservoir is reached, *event 2* is triggered to stop the back-runoff channels to the reservoir from Farm 1 and 2.

Step 5: In order to prevent the rainfall water from over-irrigating the farms, the river runoff channel from Farm 2 is opened during *event 3*.

3.2 Two Farms Decision Based System

Figure 3.4 shows visualization of two farm system.

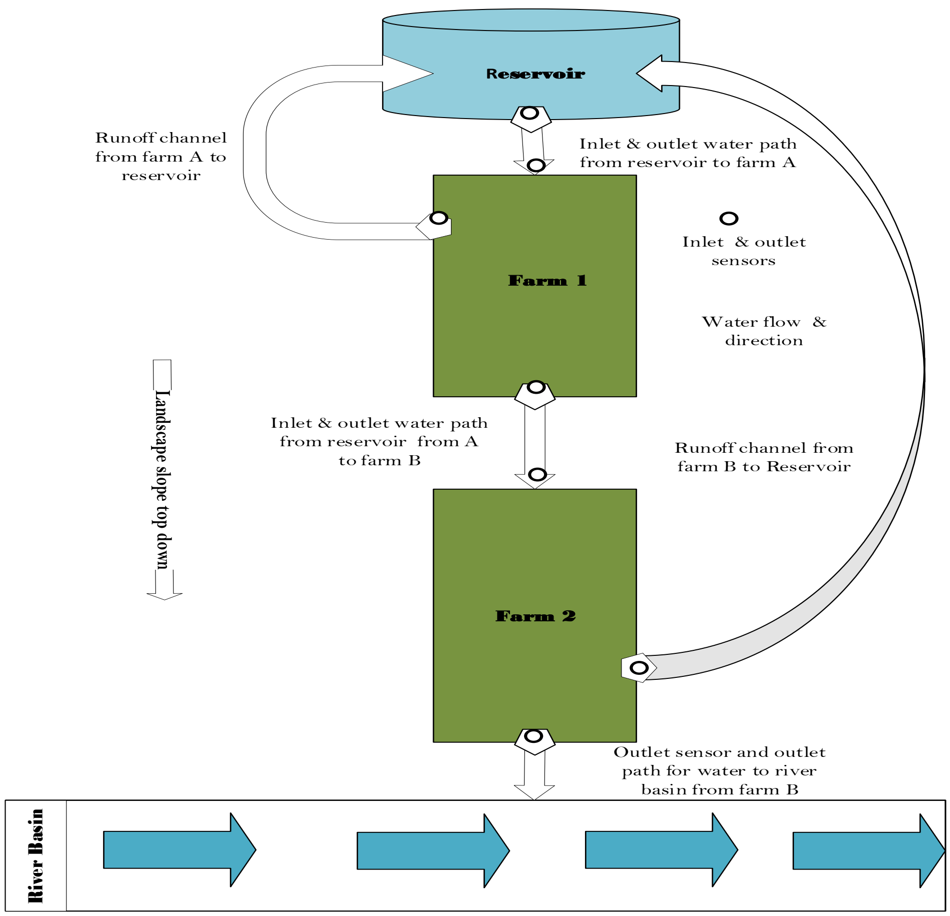


Figure 3.4 Visualization of two Farms

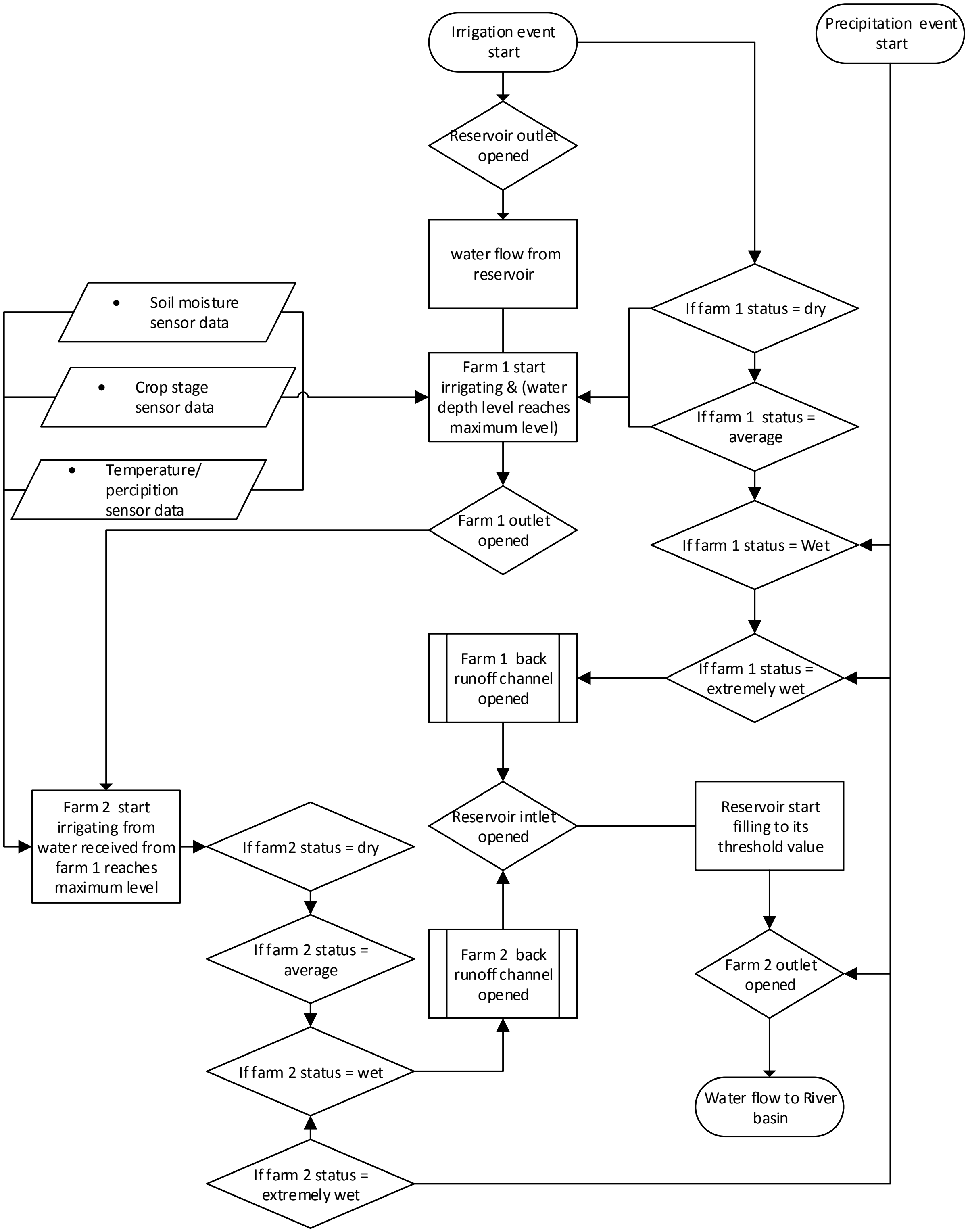


Figure 3.5 System diagram of decision based system

Figure 3.5 show the proposed two farms decision based system . Which is described as follow.

- Both farm 1 and farm 2 have soil moisture sensor, crop stage sensor, temperature and precipitation depth sensors/ irrigation depth sensor installed.
- The sensor will collect the relevant information and if the conditions on both fields are dry and average wet the irrigation event will be triggered on the fields.
- The water from the reservoir will flow to the precision irrigation system on farm 1 and farm 1 will start irrigating uniformly by sprinklers until it get to desired wet state. Then farm 1 outlet will be opened for any surplus amount to discharge.
- The farm 2 inlet sensor will receive the surplus amount of water and the local sprinkler based precision irrigation system will utilize it and an uniform irrigation water will be sprinkled on the farm 2 if the conditions of the farm 2 are dry and average wet .
- If both the farm 1 and farm 2 are wet or extremely wet due to irrigation event happened or there is continuous rainfall on both the fields, In that case the farm1 and farm 2 back runoff channel to the reservoir will be opened to fill the capacity of reservoir but if reservoir reaches its threshold capacity in that case the farm 1 and farm 2 outlet will be opened to pass away the surplus water to the river basin.

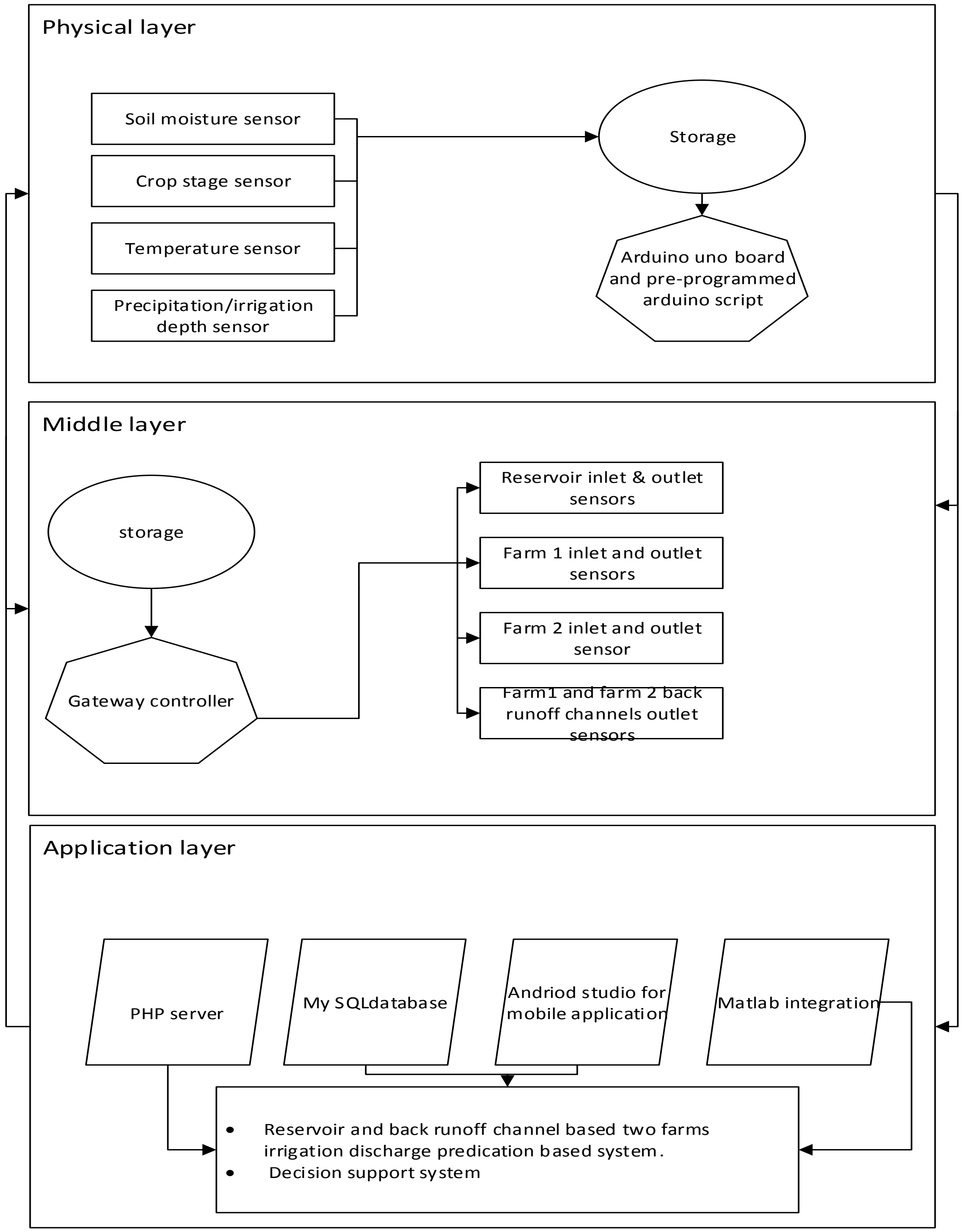


Figure 3.6 Layered architecture of proposed model

Figure 3.6 shows the layer architecture of the model. On the physical layer there are sensors for data acquisition regarding current conditions of the farms , The precision based sprinkler system installed on the farms which are linked directly to the farm agricultural and environmental sensors to apply irrigation on the farms . The WSN based precision irrigation systems of the farms are linked through direct information exchange between the WSNs in the form of packet transmission and acknowledgment.

In the middle layer there is gateway controller which will be responsible for the bidirectional communication between the physical and application layer and also for the reservoir, both farm1 and farm 2, back runoff channels inlets and outlets sensors of farm 1 and farm 2 to open and close the flow paths. In case the amount of water is surplus or there is an extensive rainfall took place on the farms the water will be directed towards the reservoir from the farm 1 and farm 2 back runoff channels to save the water that could be further utilized in the irrigation purpose. If the amount of water exceeds the threshold capacity of the reservoir and there is still continuous rainfall on the farms in that case the water will be diverted to the river basin. In this way an early alarming system and an early warning system will be presented in the form of hydrographs peak discharge and time to peak so that the other farm and river basin know the amount of water discharge and time of discharge to adjust its local irrigation system or alert to authorities if the river basin receives a higher total discharge, high peak discharge and time to peak predication and its hydrographs well in proactive manner.

In the application layer there is various tools that runs on the base station/ (server station) such as PhP server for client and server request handling, my sql for database , arduino studio and scripting for rendering of the data from the sensors and matlab integration so that NRCS based predication is carried out for total discharge , peak discharge and time to peak in form of predication values and hydrographs. the machine learning algorithms predication for total discharge and peak discharge and its pie graphs representation for total discharge and hydrographs for peak discharge and time to peak in tabulated form and displayed to the end user (landowner, hydrologist and machine expert) as well .Optimal reservoir and back runoff channels based two farm irrigation discharge predication system provides the facility of water distribution efficiently to farms and that how much is required by the farms on the basis of farms conditions and also save water waste in reservoir before diverting it to the river basin .

3.3 Curve Numbers Selection

NRCS simulator has been trained and tested against appropriate antecedent soil moisture values and crop stages values. From the below Table 3.1 the curve numbers are selected and converted to its appropriate average, dry and wet antecedent conditions through the mathematical equations 7 and 8 which are described in Chapter 2. Table 3.2 shows the selected curve numbers for simulation. NRCS is trained and tested on selected curve numbers that is shown in Table 3.2.

Table 3.1 Runoff curve numbers for cultivated agricultural lands

Cover description			Curve numbers for hydrologic soil group			
Cover type	Treatment 2/	Hydrologic condition 3/	A	B	C	D
Fallow	Bare soil	—	77	86	91	94
	Crop residue cover (CR)	Poor	76	85	90	93
		Good	74	83	88	90
Row crops	Straight row (SR)	Poor	72	81	88	91
		Good	67	78	85	89
	SR + CR	Poor	71	80	87	90
		Good	64	75	82	85
	Contoured (C)	Poor	70	79	84	88
		Good	65	75	82	86
	C + CR	Poor	69	78	83	87
		Good	64	74	81	85
	Contoured & terraced (C&T)	Poor	66	74	80	82
		Good	62	71	78	81
	C&T+ CR	Poor	65	73	79	81
		Good	61	70	77	80
Small grain	SR	Poor	65	76	84	88
		Good	63	75	83	87
	SR + CR	Poor	64	75	83	86
		Good	60	72	80	84
	C	Poor	63	74	82	85
		Good	61	73	81	84
	C + CR	Poor	62	73	81	84
		Good	60	72	80	83
	C&T	Poor	61	72	79	82
		Good	59	70	78	81
	C&T+ CR	Poor	60	71	78	81
		Good	58	69	77	80
Close-seeded or broadcast legumes or rotation meadow	SR	Poor	66	77	85	89
		Good	58	72	81	85
	C	Poor	64	75	83	85
		Good	55	69	78	83
	C&T	Poor	63	73	80	83
		Good	51	67	76	80

Table 3.2 Selected curve numbers for simulation

Cover type	treatment	Hydrologic condition	Soil group c (CN)
fallow	Bare soil		91
Small grains	Striaght row (SR)	poor	84
Small grains	SR	good	83

The soil group C is selected, CN is dependent on cover type, treatment, hydrologic condition and hydrologic soil group. On the basis of the above curve numbers. The AMC mathematical equations have been used while for crop stages the Cover types of fallow and small grains have been used. the small grain poor hydrologic condition refer increase runoff as the surface is covered with less than 50 percent ground cover while good hydrologic condition refer decrease in runoff as the ground surface covered with greater than 75 percent while fair means with 50 percent to 75 percent . Poor means increase runoff and good means low runoff .Another curve number extreme wet is also considered and assigned CN=100.

3.4 Model evaluation criteria

Model assessment procedures are based on the following three step methods.

1) Random sampling and cross validation

2) Performance measuring parameters /regression error metrics
3) Comparative assessment.

3.4.1 Random Sampling and Cross Validation

“For the strength of the assessment of the model execution to be guaranteed, the dataset was randomly divided into two sets , one is the training dataset and the other one is the testing dataset” [96].

“Cross validation is a strategy to assess models and gives preferable outcomes over residual evaluation. That is the reason that the residual strategy doesn't give a transparent sign that how accurate the learner will foresee the information when exposed to new sets. This issue is conquered when the whole information isn't exposed to the training stage as a portion of the information can be expelled before the training begins thus, this information can be utilized for the testing reason. This is the basic fundamental philosophy of the cross validation” [97].

The least complex type of the cross validation is the holdout technique where the dataset is separated into two sets , one for training and one for testing. The function approximator utilizes the training set and it is then utilize to foresee the yield esteems for the information in the testing set. The quantity of wrong presumptions made is included request to give the mean absolute test set error, which is utilized to assess the model. The advantage of this technique is that it doesn't require some time to process the outcomes though it is relied upon to give high variance. The assessment or evaluation is for the most part detailed upon the information/data focuses that are utilized in training and those that are utilized in testing.

The k-fold cross validation is an approach to improve the holdout strategy where the dataset is sorted into k subsets where the holdout technique is rehashed k number of times. The initial sample is classified into k similarly measured sub-samples in the k-fold cross validation. Out of the sub samples made, just one lot of the subsample information is held as the testing set information while the others k – 1 subsamples are utilized for the training purposes. This validation procedure is repeated k number of times utilizing an alternate k sub sample set every moment. The outcomes can be averaged to result in single forecast. The advantage of utilizing cross validation is that it utilizes all the observations for both the training and the testing purpose and each set is utilized just a single time for the approval. For the most part 10-fold cross validation is for the most part used however k stays an unfixed parameter. Like for example in the event that k is set equivalent to 2 (k=2) at that point it would mean a 2-fold cross validation, in which the dataset is arbitrarily rearranged into two sets d0 and d1, to level the sets and afterward the training is performed on d0 and testing on d1, and the other way around.

In case $k = n$ (number of observations), the k-fold cross validation is the replica of leaving-one-out-cross-validation. In the stratified k-fold cross validation, the folds are selected in order to equalize the mean response value in all the folds while in dichotomous classification, each fold is composed of rough numbers of the same proportions of the two class labels.

3.4.2 Performance Evaluation Parameters (regression error metrics)

The following metrics are used for the predicative accuracy (performance) of trained models.

- Mean square error (MSE)

It is stated that "The mean squared error determines how close a set of points is to a regression line. This is done by taking the distances from the set of points to the regression line and squaring them. These distances are the "errors". The squaring at the end is essential to remove any negative signs and helps give weight to larger differences. It is referred to as the mean squared error as you are evaluating the average of a set of error" [98].

"The smaller the value of the MSE, the nearer you are to determining the line of best fit. This depends on the data as well; it might not be possible to get a very small value for the mean squared error. The MSE includes both the variance of the estimator and its bias and is also known as the second moment (about the origin) of the error. MSE is the variance of the estimator for an unbiased estimator. Similar to the variance, MSE uses the same units of measurement as the square of the quantity being estimated. The MSE is a measure of the quality of an estimator—it is always non-negative, and values closer to zero are better" [99].

$$\mathbf{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{Y}_i^{\wedge} - \mathbf{Y}_i)^2 \qquad (12)$$

- $Y_i^{\wedge}$ Predicted value'
- Y_i observed value
- n = the number of errors,
- Σ = summation symbol (which means "add them all up"),

- Root Mean Square Error (RMSE)

It is stated that "The root mean square error (RMSE) is calculated using the square root of the residuals. It depicts the absolute fit of the model to the data telling how close the observed data points presently are to the model's predicted values. The R-squared is a relative measure of the fit while RMSE is an absolute measure of the fit. RMSE can also be interpreted as the unexplained variance's standard deviation, and possesses the property of being in the same units as the response variable. The lower values of RMSE depict a better fit. RMSE is a good measure of how accurately the model predicts the response". [100]

The RMSE with respect to the estimated variable X_{model} is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}} \qquad (13)$$

- X_{obs} : observed values

- X_{model} : modelled values at time/place i.

- Relative Root Mean Square Error (RRMSE)

It is stated that " The RRMSE is the ratio of the variance of the residuals to the variance of the target values themselves. Values of RRMSE can range between 0 and 1, where 0 means perfect forecasting. The value is normally multiplied by 100 to show a percentage of relative error.

Or alternatively can be defined as The Relative Root Mean Square Error (RRMSE) is denoted by dividing the RMSE by the mean observed data": [101] [102]

$$RRMSE = {}^{RMSE}/_{mean(Pred\_vals)} \qquad (14)$$

- Coefficient of Determination (R^2)

It is stated in this article [103], that "R^2, is the coefficient of determination which is used to detect how the distinct values in one variable can be used to explain the difference in a second variable. R-squared has a very crucial functionality that its scale is intuitive which means that it ranges from zero to one, with zero illustrating the fact that the proposed model does not improve the prediction over the mean model and one means that it has a perfect prediction. Improvement in the regression model concludes in the proportional rises in R-squared".

"R^2 represents the variability that can be explained by the model in terms of goodness of fit. R^2 ranges between 1 and 0. It is equal to 1 if the predictions are perfect, i.e. a linear relationship exist between the predicted and measured values represented by a straight line. It thus provides a way to quantify the accuracy of the model to predict the dependent variable" [104].

$$R^2 = \frac{SSR}{SST} = \frac{\sum(y_i^{\wedge} - y^{-})^2}{\sum(y_i - y^{-})^2} \qquad (15)$$

- SSR stand for Sum of square regression
- SST stand for total sum of squares
- $y_i^{\wedge}$ Predicated value
- y_i Observed value
- y^{-} Mean value

- Mean absolute error (MAE)

It is stated in this article [105] that "The mean absolute error (MAE) is the measure of the distinguished values between two continuous variables. In a set of forecast the MAE measures the average magnitude of errors, without considering their direction. The amount of error in a measurement is known as absolute error while the MAE is the mean over a validation sample of the absolute values of the differences between forecast and its corresponding observation. The range of MAE can be from zero to infinity. All the individual differences are equally weighted in the average as contrast to RMSE".

$$MAE = \frac{\sum_{i=1}^{n}|y_i - x_i|}{n} = \frac{\sum_{i=1}^{n}|e_i|}{n} \qquad (16)$$

- n = the number of errors,
- Σ = summation symbol (which means "add them all up"),
- e_i = absolute errors

- Normalize root mean square error NRMSE

It is stated in this research article that "The dimensionless forms of the RMSE are quite handy, as the user often wants to compare RMSE with the distinct units. There are two methods for that [106]".

1. Normalize the RMSE to the range of the observed data

$$NRMSE = \frac{RMSE}{X_{obs,max} - X_{obs,min}} \qquad (17)$$

2. Normalize to the mean of the observed data.

$$NRMSE = \frac{RMSE}{X_{obs}} \qquad (18)$$

3.4.3 Comparative Assessment

The comparative assessment has been done by comparing the following:

- The developed desktop based graphical user interface (GUI) from the mathematical equations and NRCS script with machine learning algorithms for reservoir-two farm based irrigation- total discharge system (here the dataset to the machine learning algorithms are given from the generated and trained dataset from the GUI based on mathematical equations and NRCS based script) and results are collected for some important scenarios as well for the comparison on NRCS predication with other machine learning algorithms predications for total discharge.

- Different machine learning algorithms are compared with each other on the basis regression error metrics for a separate dataset collected from single farm system based on a watershed from NRCS simulator to find out the peak discharge (mm) and time to peak (hr) on two different models. However the results are omitted and not shown in this thesis report due to space limit.
- Different machine learning algorithms are compared with each other on the basis of regression error metrics for dataset collected for reservoir based two farm irrigation discharge/runoff system based on a developed desktop based GUI for “reservoir and back runoff channels based two farm peak discharge (cfs) and time to peak in (mins)” .
- An IOT project has been developed in which through a mobile app data are retrieved from the sensors to the server, stored on sever, preprocessed by Matlab and then send back to user displaying all NRCS and machine learning predication for total discharge , peak discharge and time to peak and their results in table , pie graphs and hydrographs display .

Chapter 4: Decision based system on estimation of runoff

Chapter 4 presents decision support system at two farms levels based on estimation of runoff. NRCS and machine learning are used for calculation of runoff. Runoff estimation can be very valuable in water management and irrigation scheduling. Correct estimation of runoff help in utilization of wasted water. Different scenarios are discussed in detail. Figure 4.1 illustrates the graphical user interface (GUI) of proposed decision based system.

4.1 Total Discharge Predication at Farm1 outlet and Farm2 Outlet

The discharge predication has been done at farm1 outlet and farm 2 outlet and shown for individual samples and different scenarios.

4.1.1 Scenario 1 Dry Condition

The water reservoir connected with farm 1 and farm 2 through inlets and outlets, both farms are simulated here in the NRCS having soil moisture condition = dry on field 1 and field 2 as well. the crop stages =1 on both farm 1 and farm2, the irrigation depth and duration that take place on farm 1 is 58 mm and 1 hour respectively . There are other variables which are fixed such as initial abstraction Ia, area of the farms, elevation difference, maximum length and time of concentration. the reservoir capacity to full is 70000 liter however its intial value is 40000 liters. Here the data is trained & tested on the NRCS and other machine learning algorithms such as ANN, DT , SVR and MLR to show the predication of total discharge at farm 1 outlet 1 and farm 2 outlet 2 .

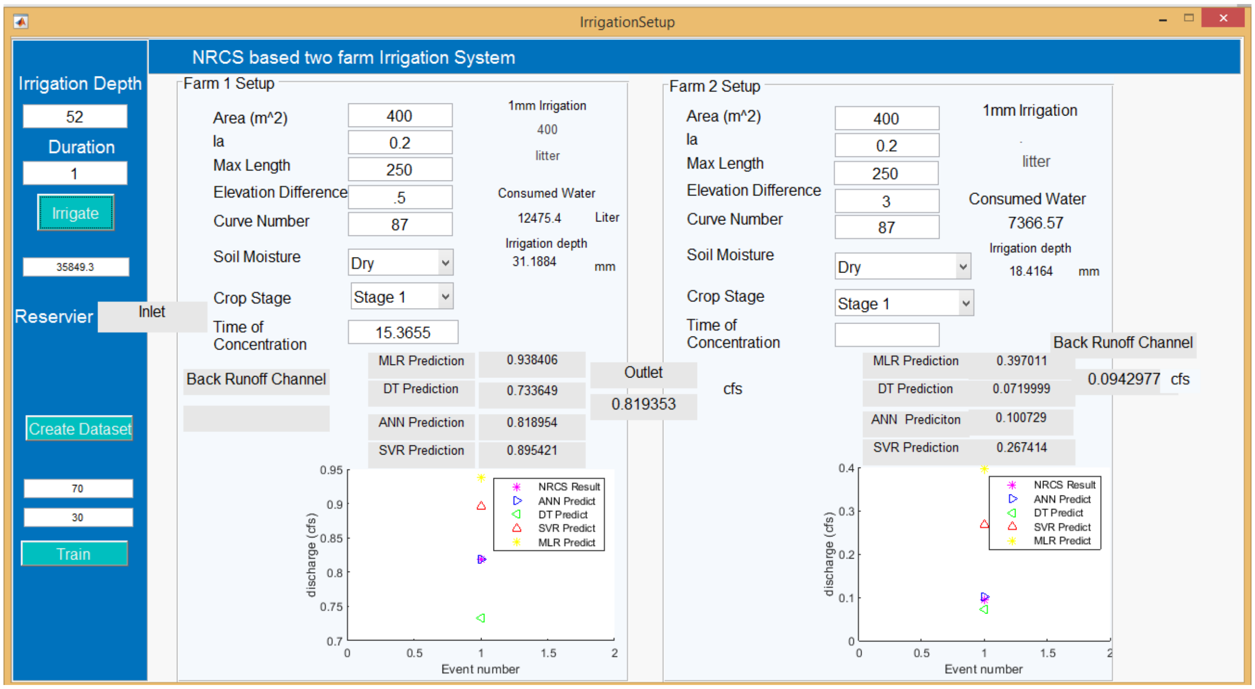


Figure 4.1 Discharge prediction at farm1 and farm2 for dry condition

In the above scenario1 in figure 4.1 the predicted value at outlet 1 is 0.81 while at on outlet 2 it is =0.09 . The GUI also shows MLR, DT , ANN and SVR predications.

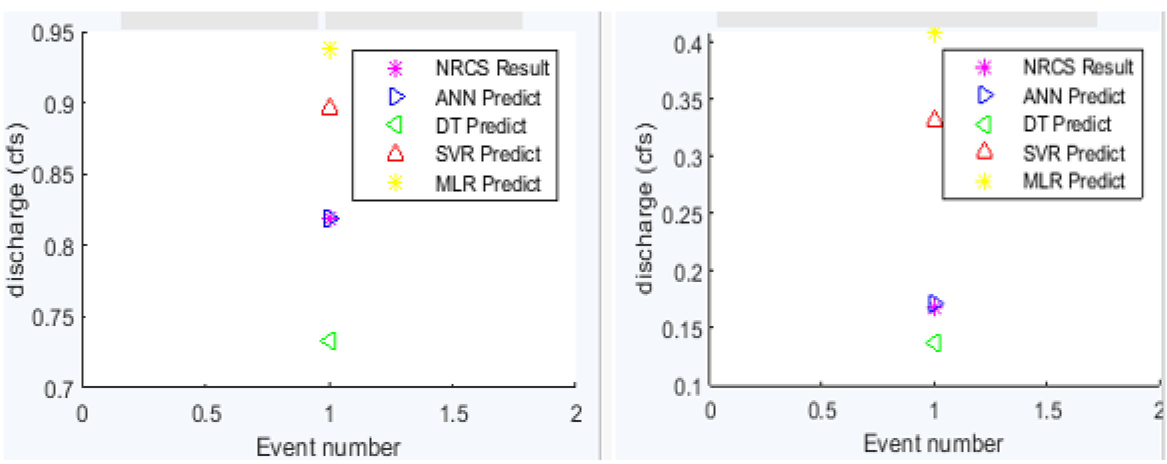


Figure 4.2 Discharge prediction at farm1 and farm2 for dry condition

Figure 4.2 shows discharge (cfs) vs Event number for outlet1 and outlet 2 shows that NRCS trained based system is very well predicated by ANN ,followed by DT , then followed by SVR and at last MLR for graph1 and for graph two ANN is very close , followed by DT , MLR and

SVR . the predicated values and NRCS predicated result are also listed on developed simulator output.

4.1.2 Scenario 2 Wet Condition

Similar scenario 2 like scenario 1 only the soil moisture is changed to wet at field 1 and field 2 as well and irrigation depth or precipitation depth =52mm. the rest of variables values are the same. When both field 1 and field 2 are wet . then five conditions will be checked and a decision making will be carried out to either use back runoff channel or not from field1 and field 2 to the reservoir?

- If total discharge at outlet 1, =< 0 back runoff channel will not use.
- If total discharge at outlet 1 , > 0 the back runoff channel will be used to divert water to the reservoir back.
- If total discharge at outlet 2, =< 0 back runoff channel will not use.
- If total discharge at outlet 2 , > 0 the back runoff channel will be used to divert water to the reservoir back.
- If the reservoir threshold capacity reach to maximum limit (70000 liters) then back runoff channels will be stopped and water will be diverted to the river basin.

Figure 4.3 Discharge prediction at farm1 and farm2 for wet condition

In the figure 4.3 the scenario 2 is implemented and displayed for discharge predication for farm1 and farm2 outlets.

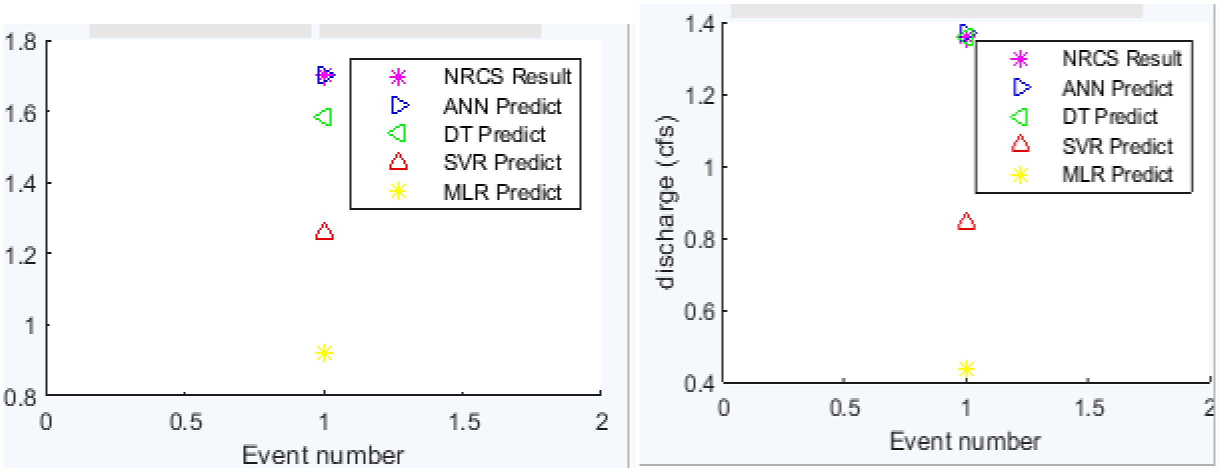


Figure 4.4 Discharge prediction at farm1 and farm2 for wet condition

The other machine learning algorithms predications are also listed, as can be seen in the above figure 4.4. The pink asterick shape represents NRCS predicated value. The yellow asterick shape represents MLR predicated value, The blue triangle shape represents ANN predicated value , the red triangle shape represents SVR predicated value and green triangle shape represents DT predicated value .

Here the above Figure 4.4 for outlet 1 and outlet 2 shows that NRCS discharge value has been predicated closely by ANN followed by DT , SVR and finally the MLR predication is less close to NRCS predicated value for farm 1.

4.1.3 Scenario 3 Average Wet

If scenario 1 is referred for all other input variables only the soil moisture conditions at field 1 and field 2 both are average wet and irrigation depth =52 mm for irrigation duration =1 hour , then two conditions will be checked and a decision making will be carried out to either use back runoff channel or not from field 2 to reservoir?

- If total discharge at outlet 2, =< 0 back runoff channel will not use.
- If total discharge at outlet 2 , > 0 the back runoff channel will be used to divert water to the reservoir back.

In this case the water will flow from farm 1 (when its status changes from average wet to wet) to farm 2 as an overflow then the NRCS GUI , the reservoir field box shows ‘inlet’ text representing that the water from field 2 will be diverted to back runoff channel of field 2 towards the reservoir inlet which is open. (when the status of farm 2 changes from average wet to wet). It is shown in figure 4.5 below.

Figure 4.5 Discharge prediction at farm1 and farm2 for average wet condition

In the figure 4.5 shows the desktop based application for discharge predication at both farm1 and farm2 outlets for average wet condition. The results shows that the back runoff channel will be used for farm 2 to reservoir as the value at the farm 2 is 0.74368 which is greater than zero .

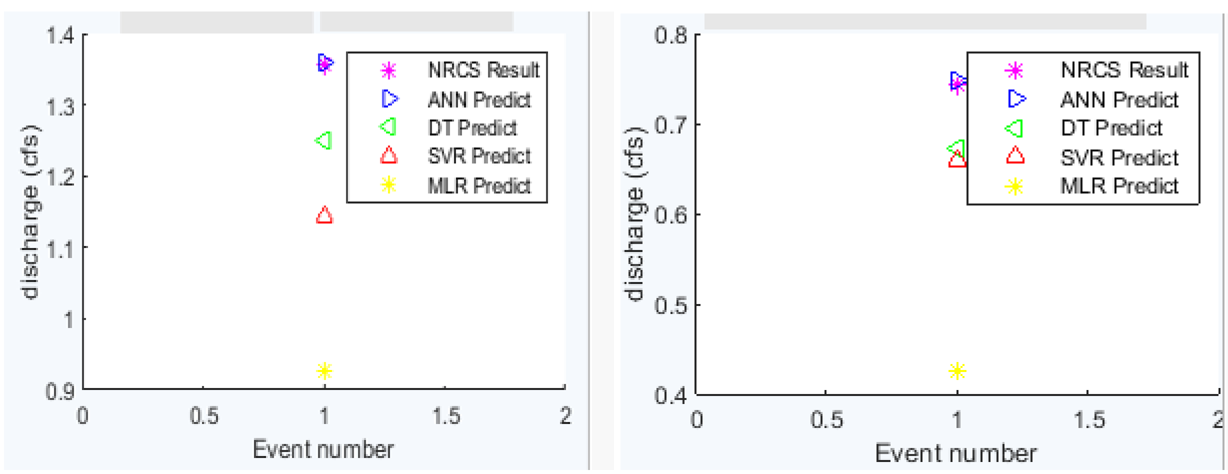


Figure 4.6 Discharge prediction at farm1 and farm2 for average wet condition

The results at the above Figure 4.6 shows that ANN values is very close, followed by the DT , and then followed by SVR ,however MLR predicated value is not close to NRCS predicated value for both farm1 and farm 2 .

4.1.4 Scenario 4 Extremely Wet

Both the field 1 and field 2 soil moisture condition are extremely wet and the crop stage on field 1 and field 2 is crop stage 1. The rest of variables values are the same like scenario 1. the reservoir has 40000 litre of water and there was a past precipitation happened for 1 hour and precipitation depth is 52 mm, and in this case the irrigation will not take place as decision making process and the reservoir textbox shows Extreme wet value . It has been shown in the below GUI figure 4.7 and its results graphs will be blank due to zero runoff as there is no current irrigation event happened.

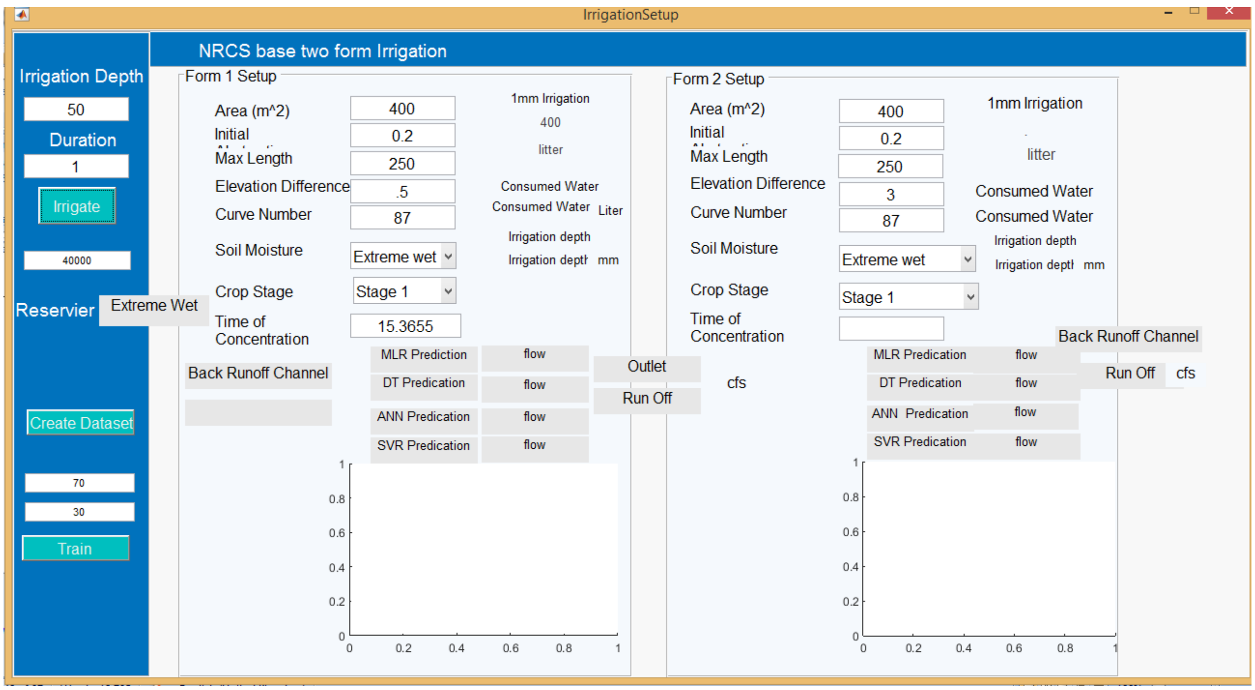


Figure 4.7 Extreme wet condition

4.1.5 Scenario 5 Average Wet at Farm 1and Extreme Wet at Farm 2

In this scenario two fields 1 and field 2 is average wet and extremely wet respectively. The crop stages is selected = crop stage 1 . the rest of variables are feeded the same like first scenario. here the decision making process will be taken and farm 1 outlet field box shows extreme wet , the water is diverted by the back runoff channel of field 1 to the reservoir to store.

- If total discharge at outlet 1, =< 0 back runoff channel will not use.
- If total discharge at outlet 1 , > 0 the back runoff channel will be used to divert water to the reservoir back.

Figure 4.8 Discharge prediction for average wet at farm 1 and extreme wet condition at farm 2

Figure 4.8 show the discharge vs Event number for the field 1 back runoff channel. The NRCS has been predicated very well by ANN and DT , however MLR and SVR has shown less accurate results . The NRCS predicated value is 1.35786 which is predicated by ANN also the closer 1.3584, and DT predicated value is =1.24972 against the NRCS predicated value for peak discharge (cfs) while SVR predicated value=1.14292 and MLR predicated value=0.925649. however, there is no flow of water to farm 2 therefore no graph generated for farm 2.

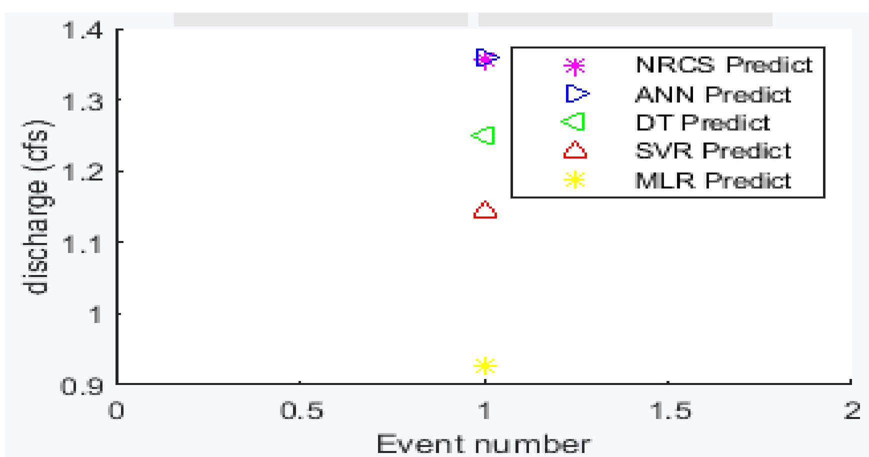


Figure 4.9 Farm1 discharge prediction for average wet condition

The figure 4.9 shows farm1 discharge predication for average wet condition , the predicated values are for ANN, DT , SVR , MLR and NRCS.

Table 4.1 Algorithms performance against NRCS for discharge prediction

Algorithms performance on Farm1 against NRCS	Predication performance on outlet1	R^2	RMSE	Algorithms performance on Farm2 Against NRCS	Predication performance on outlet 2	R^2	RMSE
DT	Very good	0.9998	0.03208	DT	very good	0.98	0.19382
SVR	good	0.9576	0.50786	SVR	satisfactory	0.61	0.77183
MLR	satisfactory	0.9243	0.65652	MLR	good	0.56	0.87271
ANN	excellent	0.9999	0.01393	ANN	excellent	0.99	0.00882

In the above figures from 4.1 upto 4.9 the graphical result display is shown for an individual events for a particular single samples on x –axis and its corresponding total discharges (cfs) predication on y axis.

However table 4.1 shows different algorithms results in terms of RMSE and R^2 on the basis of overall data samples that are 1134 , which are trained and tested on 70:30 random sample dataset split .

4.2 Peak Discharge Predication at Farm1 outlet and Farm2 outlet

In this section peak discharge is predicted at farm1 and farm2 outlet. Their respective hydrographs in relation to total time and time to peak are also generated. The peak discharge predication has been done at farm1 outlet and farm 2 outlet and shown for individual samples and different scenarios. The hydrographs for each individual scenario is also developed for both farm 1 and farm 2. The composite hydrographs are also generated for 90 mm irrigation depth.

Note: it must be clear in prior that the reservoir and back runoff channels -two farm based irrigation discharge system is developed to reflect on the real irrigation runoff modeling, for that purpose irrigation duration is taken for an hour with continuous irrigation depth after intial abstraction which could range from 1 mm to 100 mm (however the actual watershed is for 24 hour duration with 8 inches of rainfall depth to generate the hydrographs).

As There are no real data for all parameters for the actual farms to generate the hydrographs at the moment. For this purpose the watershed hydrographs are taken and generated for an hour duration with the discharge values for the irrigation distribution in the range of 1mm to 100 mm on the NRCS simulator.

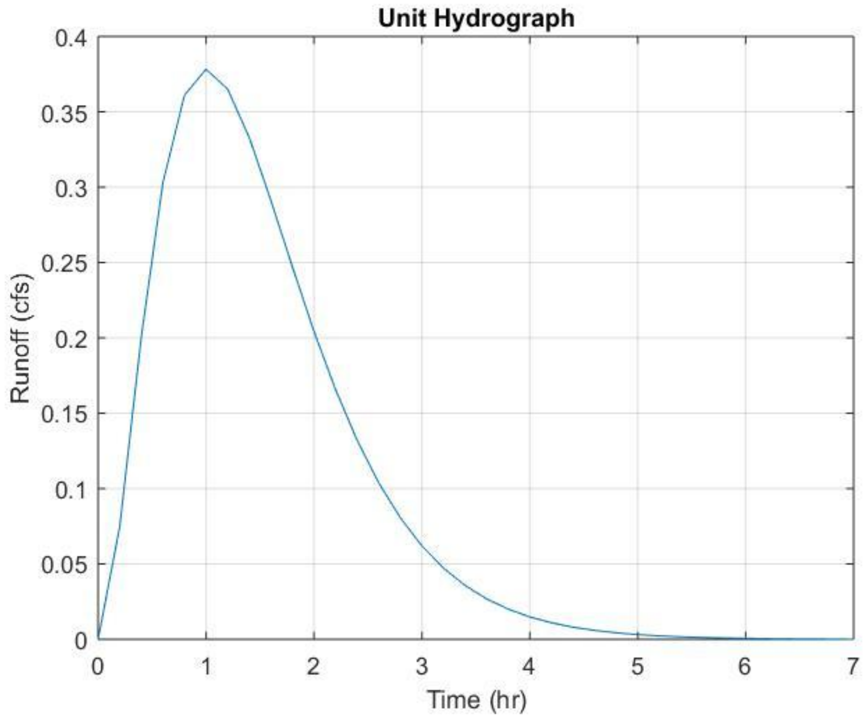


Figure 4.10 The unit hydrograph for two farm irrigation on basis of NRCS simulator

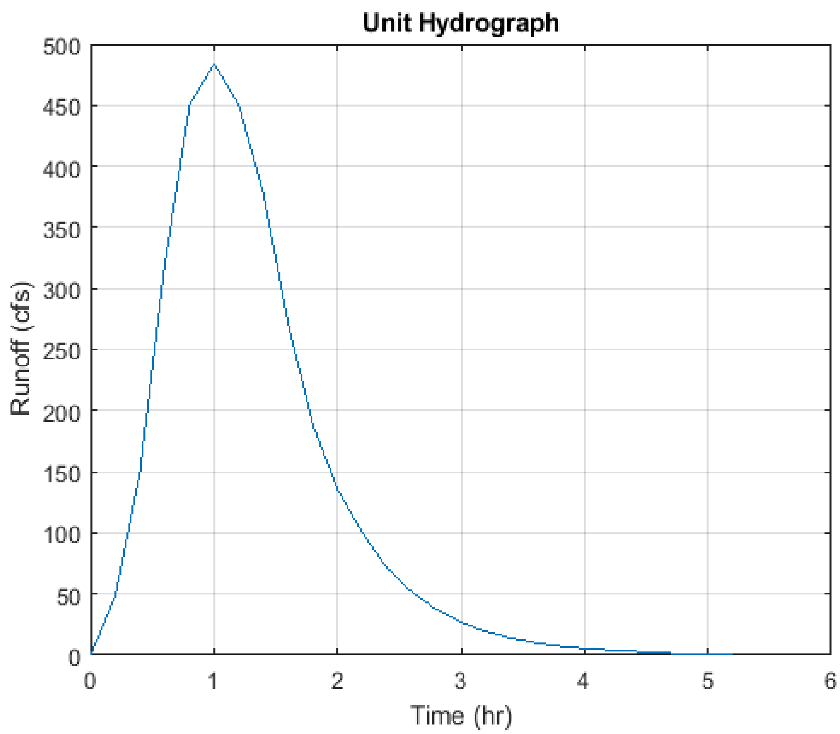


Figure 4.11 The unit hydrograph for the whole watershed

4.2.1 Scenario 1 Dry Conditions and Crop Stage 1

In this scenario as shown in figure 4.12 both fields soil moisture conditions are dry and the crop stage 1 is selected on them , the irrigation duration is 50 mm and area of the farm is 400 meter square with certain other variables which are fixed values.

Figure 4.12 Peak discharge prediction for dry condition

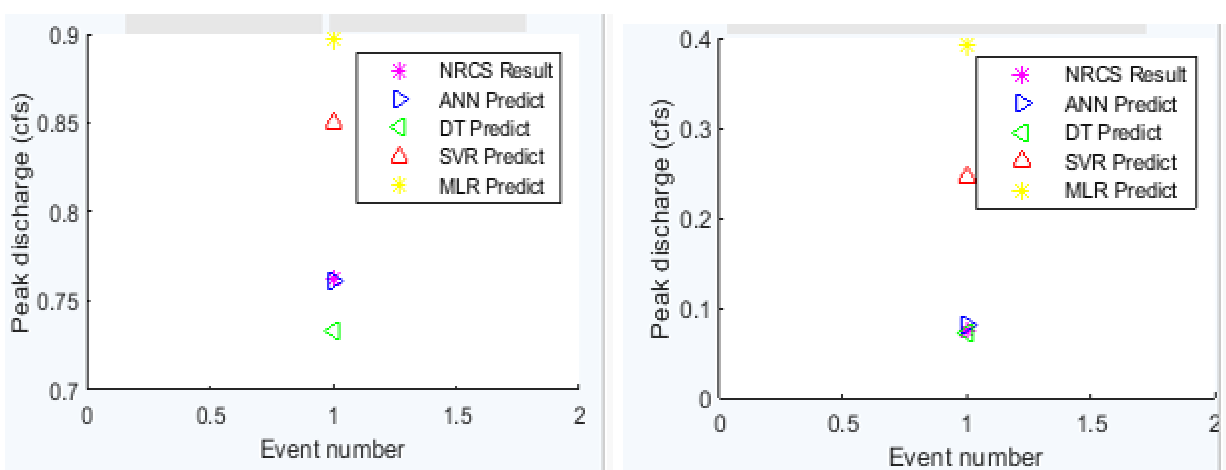


Figure 4.13 Peak discharge prediction for dry condition at farm1 and farm 2 outlet

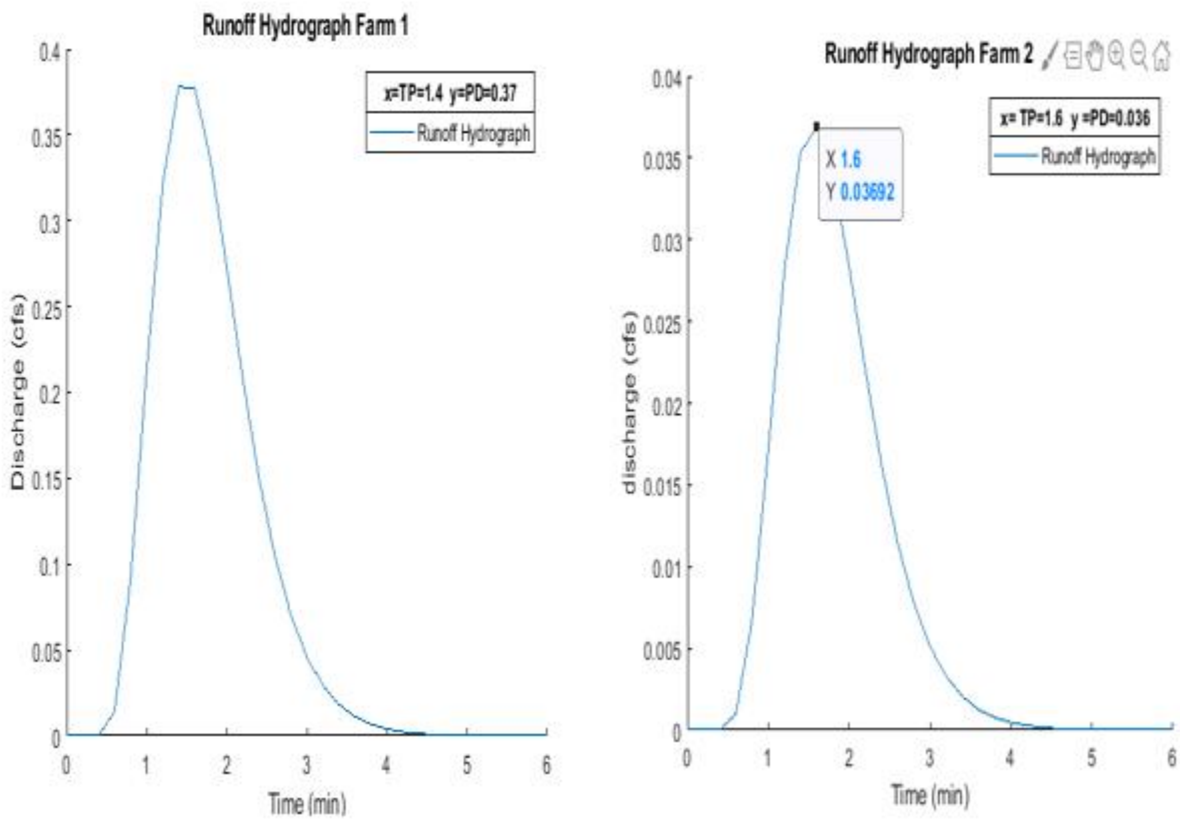


Figure 4.14 Peak discharge prediction against time to peak on basis of NRCS at farm1 and farm2

The above Figure 4.14 shows hydrographs integrated from the NRCS for peak Discharge against time to peak value for farm 1 and farm 2 outlets . These hydrographs are generated on the basis of an hour and the other conditions such as irrigation distribution for an hour which is 50 mm related to irrigation peak discharge model . The hydrographs show that peak discharge at outlet 1 is higher than the peak discharge at outlet 2.however the hydrographs are not exactly the same value as predicated by the reservoir –two farms irrigation discharge model but its just closer representation on the NRCS based hydrographs .The dry soil moisture condition is represented by blue hydrographs curve.

4.2.2 Scenario 2 Wet Conditions and Crop Stage 1

The farm 1 and farm 2 both are wet and crop stage selected is 1, the irrigation depth 50 mm and the rest of variables all are fixed/constant value as shown in below figure 4.15..

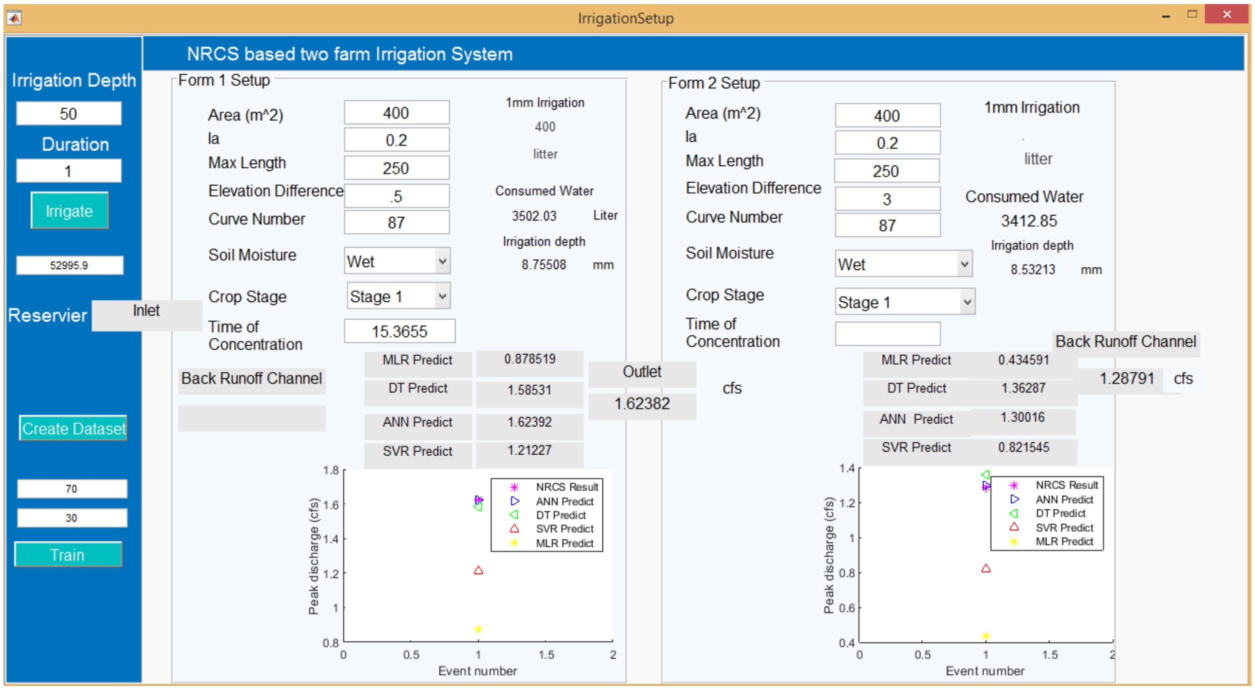


Figure 4.15 Peak discharge prediction for wet condition at farm1 and farm2

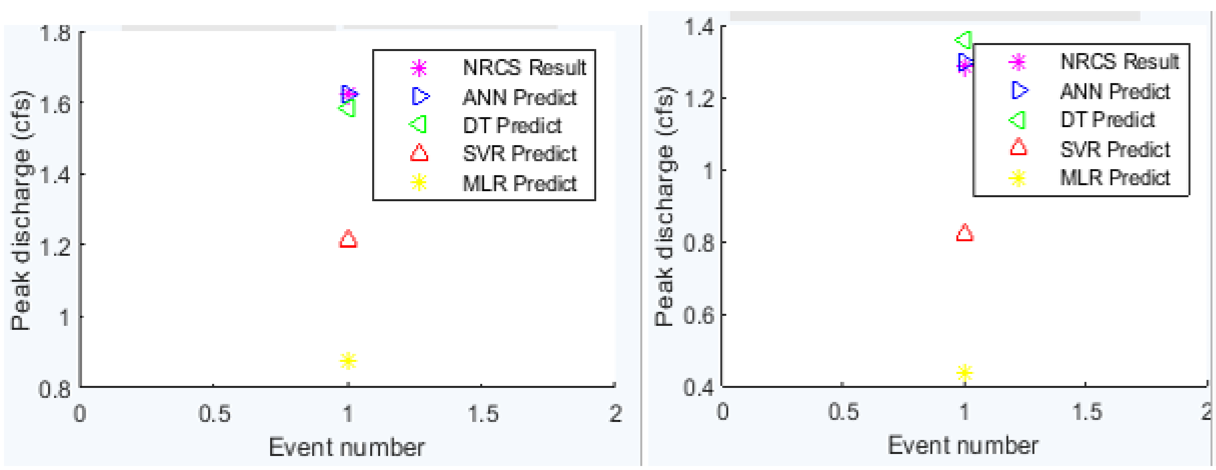


Figure 4.16 Peak discharge prediction at farm1 and farm2 for wet condition

In the above figure 4.16 x-axis represents the peak discharge against an event number or sample number , both the graphs clearly shows that NRCS peak discharge predication at outlet 1 and outlet 2 is predicated very well by ANN and then DT , however MLR and SVR predication are not that accurate.

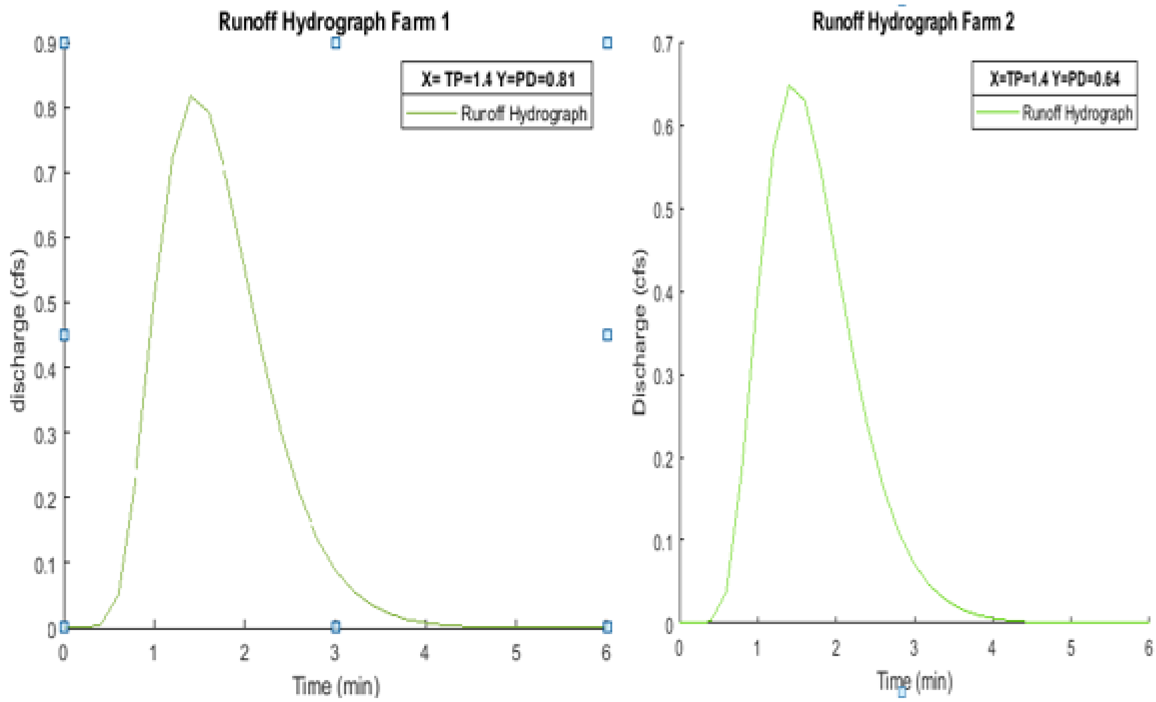


Figure 4.17 Peak discharge prediction on basis of NRCS for wet condition

The hydrographs are based on NRCS but its for an hour duration and discharge values for 50 mm depth of irrigation to represent our reservoir based two farm system whose real hydrographs are not available due to the lack of real world data for all the parameters values for the actual two farms system. Here the hydrographs is for crop stage 1 and wet conditions of soil moisture on both the farms 1 and 2. Figure 4.17 above shows that both the hydrographs are clearly showing higher values than considered for dry and crop stage =1 conditions on both the farms . The wet soil moisture condition is represented by green hydrographs curve.

4.2.3 Scenario 3 Average Conditions and Crop Stage1

In the below figure 4.18, the scenario 3 is shown in whcih Both farms are having crop stage 1 and soil moisture average conditions . 50 mm of irrigation depth for an hour and the rest of variables are taken as constant. The predication is done on NRCS and then ANN , DT , MLR and SVR . The ANN predication is very accurate followed by DT , how ever the MLR and SVR predications are not very accurate.

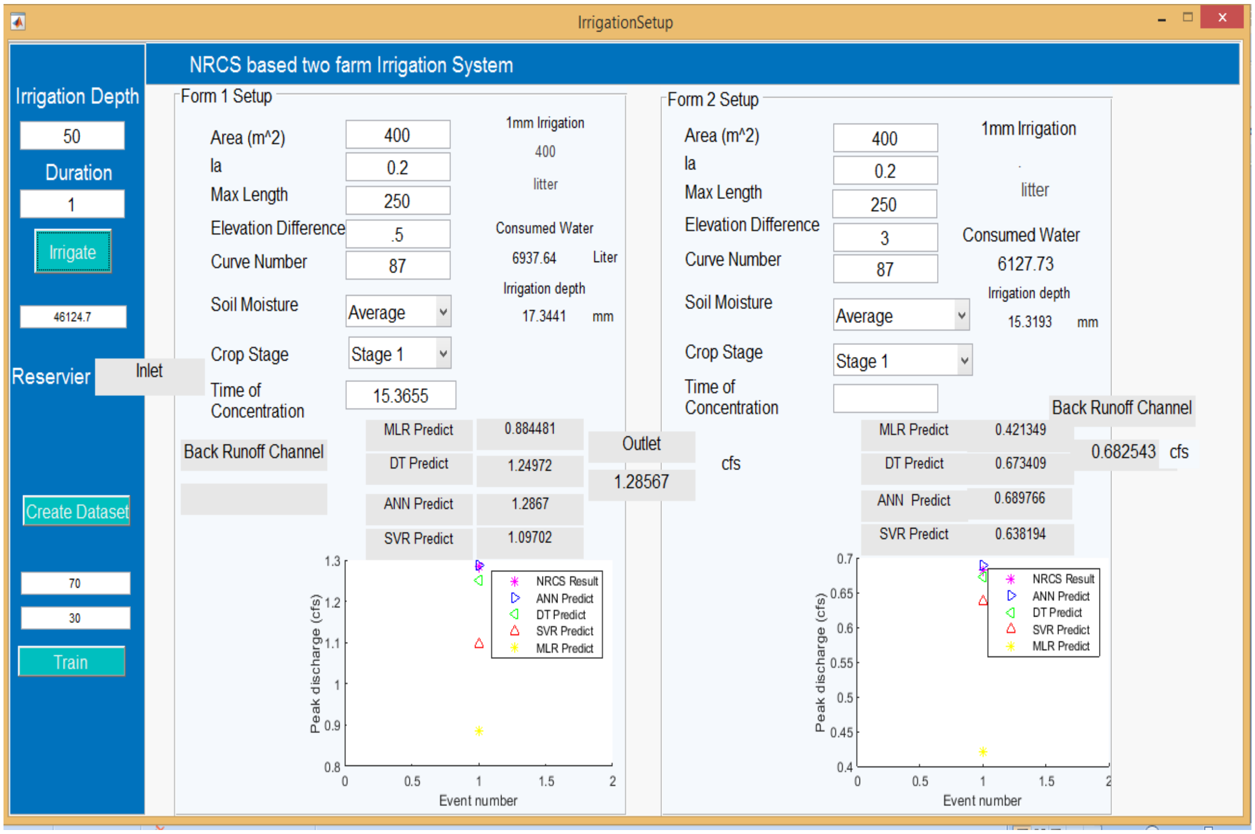


Figure 4.18 Peak discharge prediction for average condition

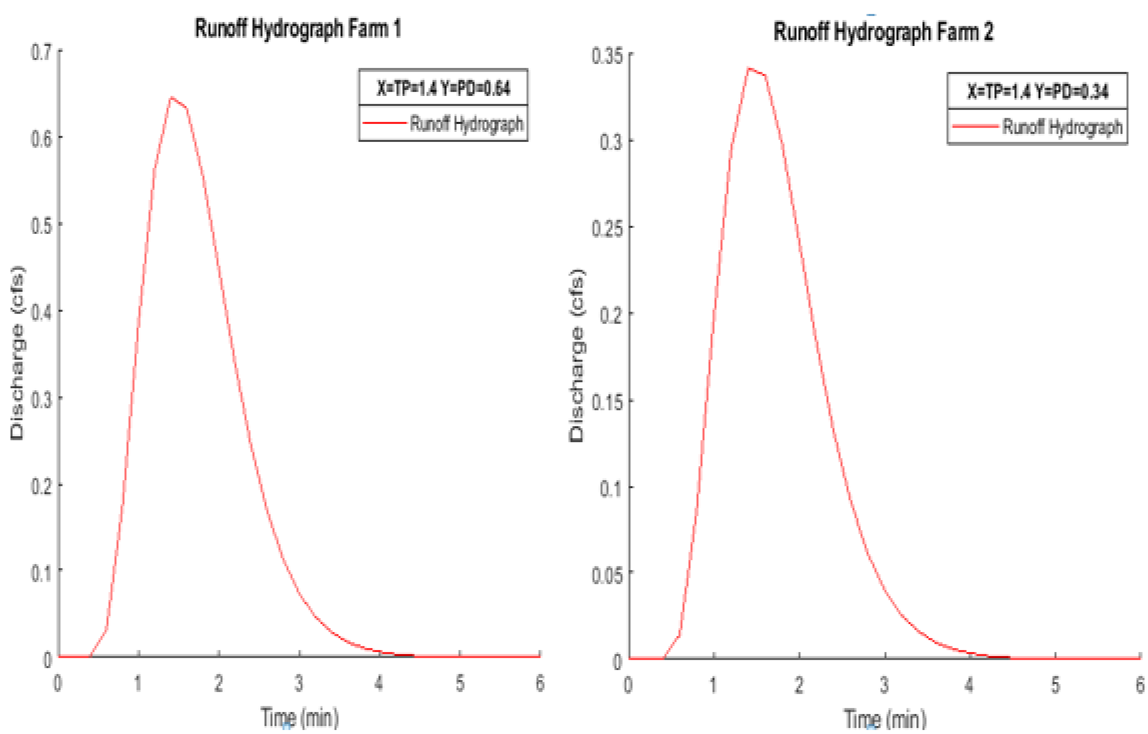


Figure 4.19 Peak discharge prediction for average condition based on NRCS

In the above figure 4.19, The X –axis represents time to peak and Y-axis represents peak discharge , the hydrographs values are average as compared to wet and dry conditions and best reflect to represents farm1 outlet and farm2 outlet of reservoir –two farms irrigation peak discharge based system .

Table 4.2 Algorithms performance for peak discharge prediction against NRCS

Algorithms performance on Farm1 against NRCS	Predication performance on outlet1	R^2	RMSE	Algorithms performance on Farm2 Against NRCS	Predication performance on outlet 2	R^2	RMSE
DT	Very good	1	0.00000	DT	very good	0.98	0.0255
SVR	good	0.95	0.059275	SVR	Satisfactory	0.67	0.109
MLR	satisfactory	0.18	0.18016	MLR	Good	0	0.193
ANN	excellent	1	0.00000	ANN	Excellent	0.99	0.00057

In the above figures from 4.12 upto 4.19 shows the graphical results for an events on a particular single samples on x-axis and its corresponding peak discharges on y-axis. The figures also show the hydrographs for their respective samples in multiple colours blue for dry moisture condition, green for wet soil moisture condition and red for average wet soil moisture condition.

The above table 4.2 shows overall results on the basis of RMSE and R^2 for data samples equal to 891.The above table show results that are acquired on the basis of 10 fold cross validation for peak discharge at outlet 1 and peak discharge at outlet 2 .

4.3 Composite Hydrographs

The following Figure 4.20 shows composite hydrographs for peak discharge (cfs) and time to peak (mins) for dry , wet , average Soil Moisture Conditions and crop stage = 1 for 90 mm irrigation depth at farm1 outlet.

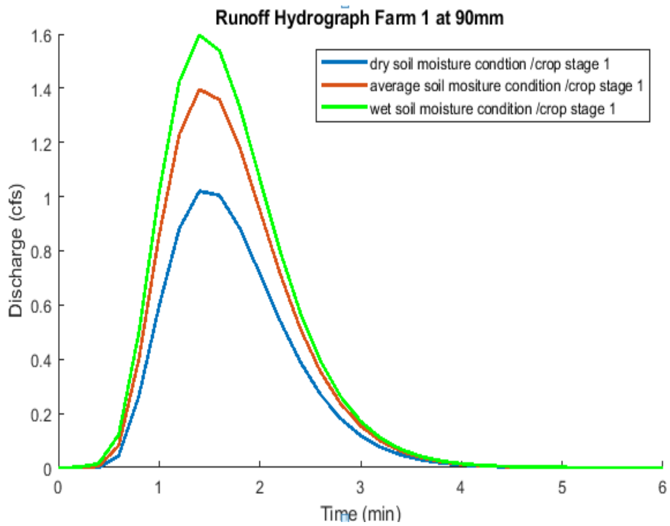


Figure 4.20 Composite hydrographs for dry,wet , average condition and crop stage 1 at farm1

The following Figure 4.21 shows composite hydrographs for peak discharge (cfs) and time to peak (mins) for dry , wet , average Soil Moisture Conditions and crop stage= 1 for 90 mm irrigation depth at farm2 outlet.

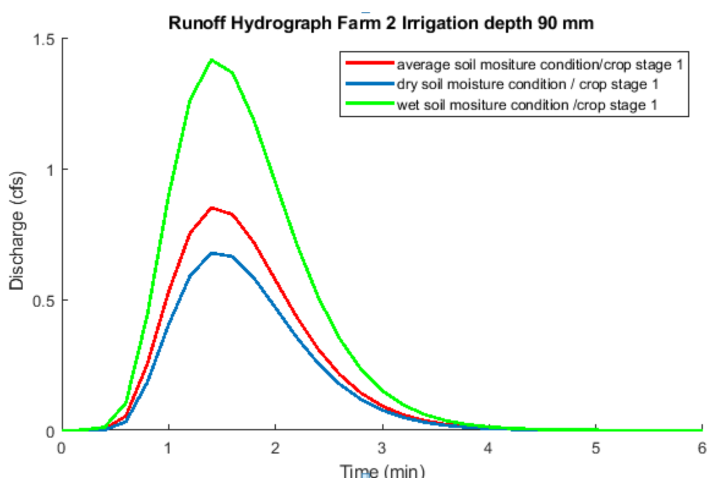


Figure 4.21 Composite hydrographs for dry,wet , average condition and crop stage 1 at farm2

Figure 4.22 Composite hydrographs for dry,wet , average condition and crop stage 2 at farm1 with 90 mm irrigation depth at the farm1 outlet , the rest of variables are selected as fixed values.

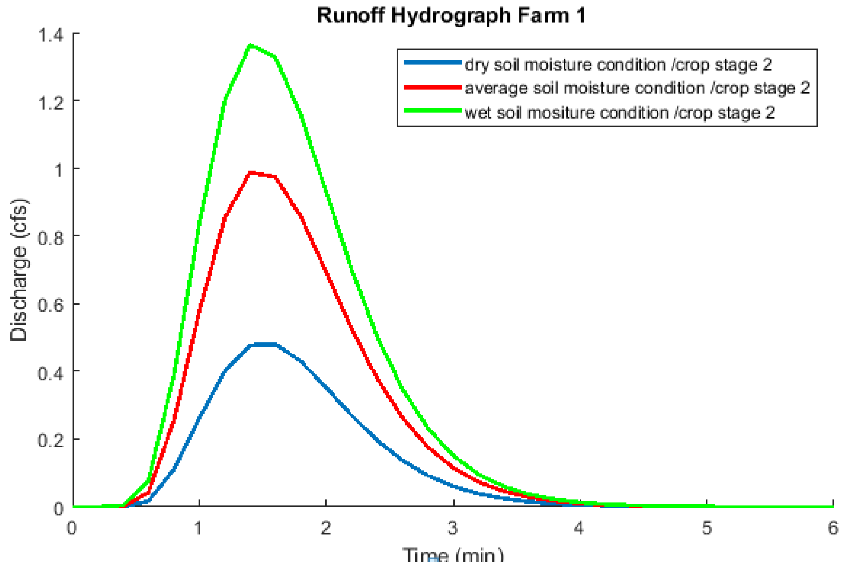


Figure 4.22 Composite hydrographs for dry,wet , average condition and crop stage 2 at farm1

Similarly,Figure 4.23 Composite hydrographs for dry,wet , average condition and crop stage 2 at farm2.

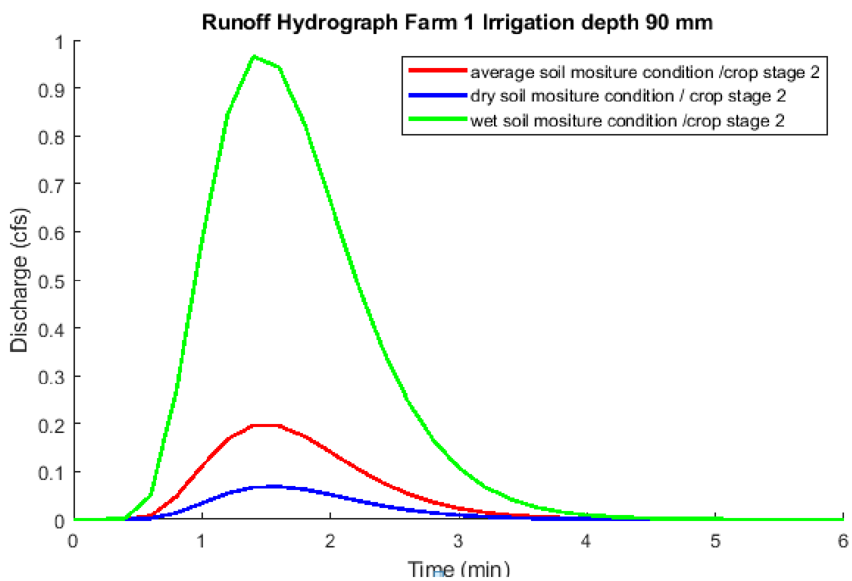


Figure 4.23 Composite hydrographs for dry,wet , average condition and crop stage 2 at farm2

Chapter 5: WSN deployment in Real world (IOT based model)

Due to some limitation the proposed system could not be fully implemented in real world. However, to practically demonstrate the usability of the proposed system, the proposed system is partially tested in real world. IOT (Internet of Things) based app is developed which help the user to retrieve the status of farms on mobile device. Data is retrieved from arduino soil moisture sensor and temperature sensor (LM35 arduino sensor).while for crop stages proxy values are assigned to the model which is constituted on NRCS and TR 55 document on small watershed for urban hydrology to reflect on real world values for crop stages such as fallow land, small grain less than 50 percent cover and small grain with 75 percent surface cover. The data is further analyzed and the current condition of farms are being displayed on mobile. On the physical layer WSN nodes are deployed on two farms for data elicitation from the real world, While Machine learning is used on application layer for the decision making that how much amount of water is Predicated (peak discharge) and how much time it takes to reaches to the peak (time to peak) at the farms outlet level. The predication for total discharge (cfs) is also developed at farms outlets. While in the middle layer such as gateway is responsible for inlet and outlets sensors of reservoir , farms and farms back runoff channels to open and close the flow paths after the decision making on the application layer.

The real challenge is that how precision irrigation system and WSN deployed on the two farms save water waste , energy and also is cost effective.

There are various climatic, soil moisture and crop growth development parameters to be measured on these sensor deployed on the farms for onward processing and decision making to be taken on machine learning at the application layer.

5.1 Description of the hardware (laboratory setup)

The device consists of three components:

- Wireless sensor module
- Base station
- Software

5.1.1 Wireless sensor module

The wireless sensor module or device consistsof ardunio Uno board, which is a microcontroller board based on ATmega328P. It has 14 digital input/output pins (six can be used for PWM outputs and six can be used for analog inputs, USB conncection and 16 MHz quartz crystal.It can powered by a USB cable or by an external 9 volt battery. Arduino Uno is developed by arduino.cc. Arduino Uno device can be configured with ardunio software (IDE).Uno means one in italian and was chosen to mark the release of arduino software IDE 1.0. The ATmega328 on the ardunio uno comes preprogrammed with a bootloader that allows uploading the new code to it with out the use of an external hardware programmer. The communication is possible through original STK500 protocol.

Communication

The arduinoUno can make communication with other arduino uno board, microcontrollers or with a computer.UART TTL (5V) serial communication is provided by ATmega328, which is available on digital pins 1(TX) and 0 (RX). An ATmega16U2 on the board channels this serial communication over USB and appears as a virtual com port to software on the computer.

There is also wireless communication between arduino uno board to another arduino uno board possible through NRF24L01+PA+LNA wireless transceiver module. It provides 1100 meters range [107] [108] .

The connected nodes to the arduino Uno device

Large number of nodes are scattered through a geographic agricultural location to monitor environmental and soil parameters that can affect the irrigation process. Some important nodes of our developed system are discussed as follow:

- **Soil moisture sensor**
- **Temperature sensor**

Soil moisture sensor (arduino)

The fundamental node to monitor and estimate the soil moisture of a farm is mandatory because it will help to know that how much water is required by the plants to be provided by the irrigation system. When the soil moisture of the farm is known in prior a specific amount of water should be provided to save water waste.

The soil moisture sensor that has been used with the arduino Uno board is the arduino soil moisture sensor . it is shown as in the below pictures.

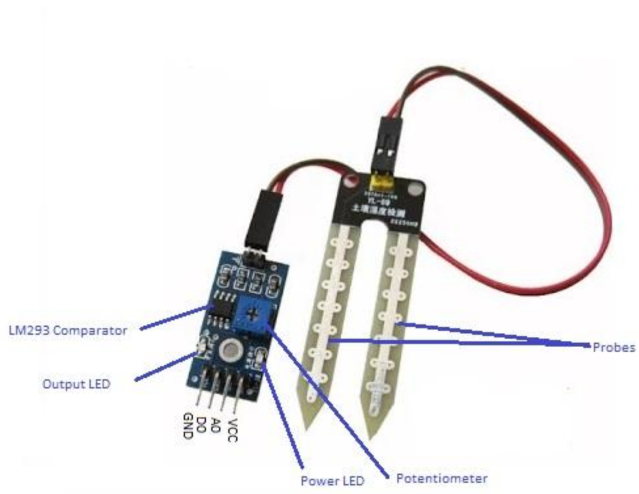


Figure 5.1 Arduino Soil moisture sensor [109]

The above figure 5.1 shows arduino soil moisture sensor in which the soil moisture sensor consists of two legged lead that can be plugged inside anywhere in soil on the farm, the basic aim is to measure the soil water content. This has two header pins cable that has to be connected with an amplifier /A-D circuit which in turns connected with the arduino Uno. The data can be collected both in digital and analog form.

How soil moisture sensor works

Soil moisture sensor estimate volumetric water content of the soil based on the dielectric constant of the soil. Soil's ability to transmit electricity is called dielectric constant. When the soil water content increases the soil's dielectric constant also increases. The measurement of dielectric constant gives a predictable estimation of the water content of the soil.

Temperature sensor (arduino) / weather node

The weather node is responsible for monitoring the environment, there are various types of sensors to monitor the environments that are humidity, temperature, windspeed, solar radiation etc.

However, Air temperature to monitor has a crucial impact on the crops growth, low air temperature there is less water absorption and movement of water in plants.

"The LM35 series are precision integrated-circuit temperature devices with an output voltage linearly proportional to the Centigrade temperature. The LM35 device has an advantage over linear temperature sensors calibrated in Kelvin, as the user is not required to subtract a large constant voltage from the output to obtain convenient Centigrade scaling. The LM35 device does not require any external calibration or trimming to provide typical accuracies of ±¼°C at room temperature and ±¾°C over a full −55°C to 150°C temperature range."

(LM35 Precision Centigrade Temperature Sensors)

LM35 is three terminal linear temperature sensor from National semiconductors.

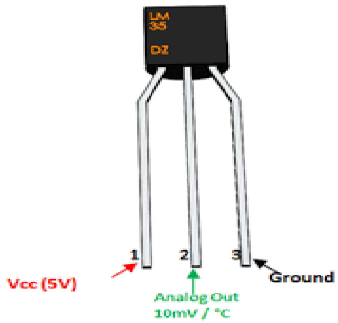


Figure 5.2 LM35 Precison Centigrade temperature sensor [110]

Pole mounted Multispectral/hyperspctral camera connected with arduino

The pole mounted camera can be placed horizontally to arduino uno or pole mounted vertically to take crop images .

5.1.2 Base station

Base station is referred as a centralized component that is specifically used for the data elicitation from the different sensor nodes attached to arduino uno, In the experimental setup the base station is composed of two further components that is a laptop and a gateway mote. The laptop runs on windows operating system. whenever the data is received by the base station /laptop from the gateway mote which in turns receive data from different nodes of arduino uno boards, the laptop has various WSN based software (arduino IDE) and machine learning algorithms Script for the real world data processing and execution to provide intelligent decision support system by predicating peak discharge and time to peak at the farms outlets level.

The below schematic view represents the system from three layers.

- Physical layer
- Middle layer
- Application layer

5.2 IOT based model practical setup on two neighbouring farms

A practical setup on the farms has been carried out by the implementation of different sensor nodes for soil moisture , temperature and crop camera on the two nearby adjacent farms.

Topology of the two neighbouring farms

The point-to-point topology of the two neighbouring farms A and B which has been considered in this experimental setup . For fallow land (crop stage 1) mean 0 % surface cover , when soil moisture condition is dry . the data from the soil moisture , crop stage and temperature is rendered from sensors installed in the fields and precipitation depth value is given from the local meteorology data available on mobile app by the end user as shown below in figure 5.3.

Figure 5.3 Fallow land (crop stage 1)

Figure 5.4 shows small grain (crop stage 2) means that the land is covered with < 50 surface cover and here soil moisture condition is dry as well.

Figure 5.4 small grain (crop stage 2)

Figure 5.5 shows small grains with = > 75 % surface cover (crop stage 3) with dry soil moisture condition . The sensors for temperature , soil moisture and crop camera ardiuno is installed in the fields and are able to have Ethernet connection as well as wifi connection between ardiuno uno boards installed on two fields 1 and 2.

Figure 5.5 small grain (crop stage 3)

IOT design diagram

The android application has been developed in JAVA v8 for the mobile application. The working of android application is described in the following steps. Figure 5.66 shows systematic diagram of IOT based model.

Step 1 : When the user entered value for precipitation/irrigation depth , the request for the soil moisture, crop stage and temperature data for two farms has been made to the server /base station. There must be an IP address for internet connection .

Step 2: The Xampp server is installed which can act as local server and live server , developed apache friends , consists of apache http server, mariaDB data base (derived from mysql) and read scripts written in php and perl . The data request is then forwarded to the arduino uno which has been configured by ardiuno IDE V1.8.6 and relevant arduino sensor for temperature , crop stage and soil moisture through a router/gatway .

Step 3: Once the data has been rendered from the sensors it is stored on the local server /live server in the database.

Step 4: The matlab has been connected with local server database, data has been imported to matlab which preprocess the data according to NRCS train and then run several sophisticated machine learning algorithms ANN, MLR, DT and SVR.

Step 5: Finally the data has been met by the local server and response for data handling is sent to the user mobile app, which display the farm 1 and farm 2 based results for NRCS , ANN, MLR, DT and SVR predication and also generates the relevant graphs showing discharge values in bar graphs. The mobile app usage is used as an emulator.

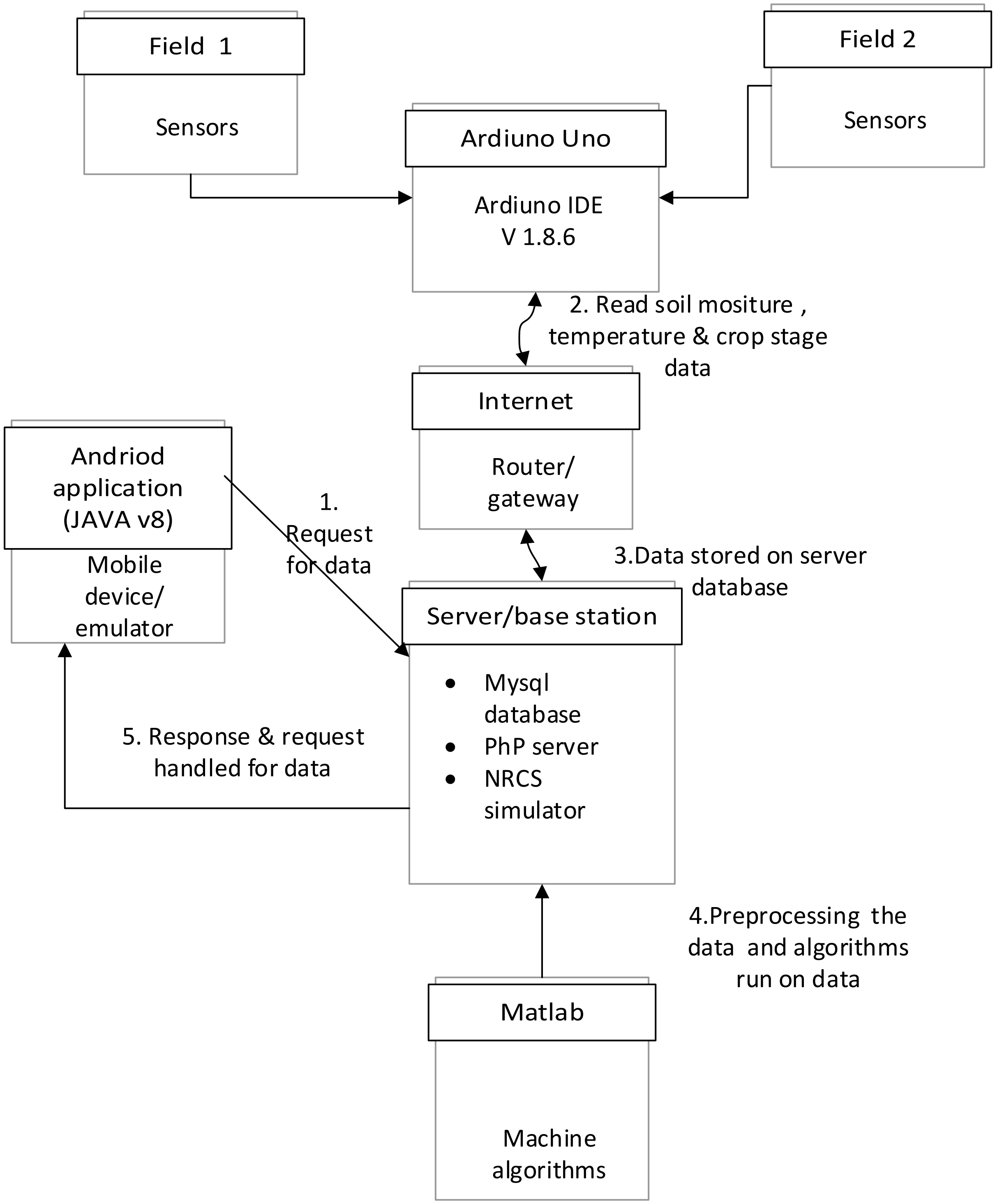


Figure 5.6 IOT based app systematic diagram

5.3 Mobile Emulator Results from Android studio

5.3.1 Results of Dry conditions

In this emulator results the data is retrieved from soil moisture sensor, crop stage sensor and temperature sensor, both the fields 1 and 2 have dry conditions and crop stage one. Here crop stage means fallow land (0 % surface cover).

The values for precipitation or irrigation can be given in millimeter in the mobile emulator. The other soil moisture conditions are dry, average wet, wet and extreme wet while the crop stages are crop stage 1 (0% surface cover) , crop stage 2 (50 % surface cover) and crop stage 3 (>=75 % surface cover).

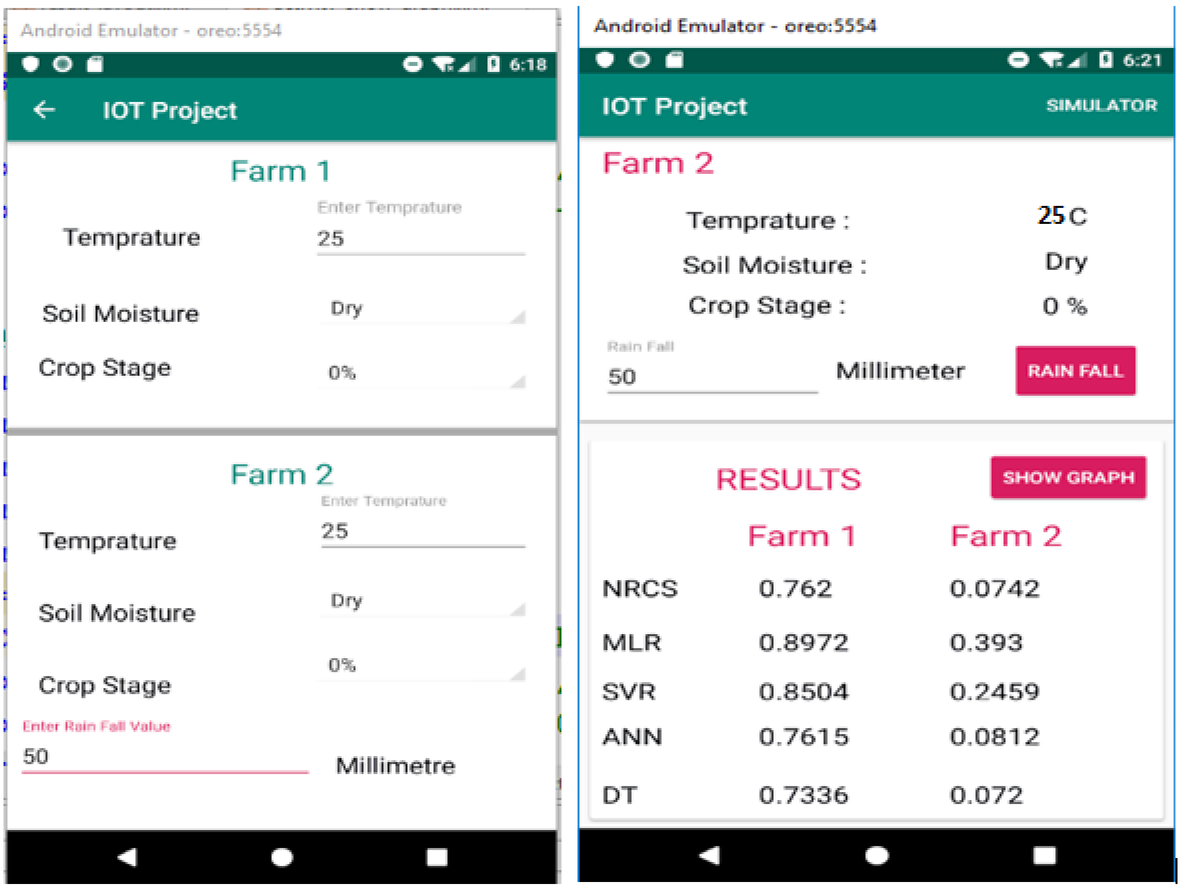


Figure 5.7 Parameters values and discharge prediction at farm1 and farm2 outlet

Figure 5.7 shows farm 1 and farm2 initial variables such as temperature , soil moisture , crop stage and precipitation or irrigation value variable in millimeter for both field 1 and field 2 and also shows the results for total discharge predication in the numeric values for farm 1 outlet on the basis of NRCS, MLR, SVR , ANN and DT and similarly for farm 2.

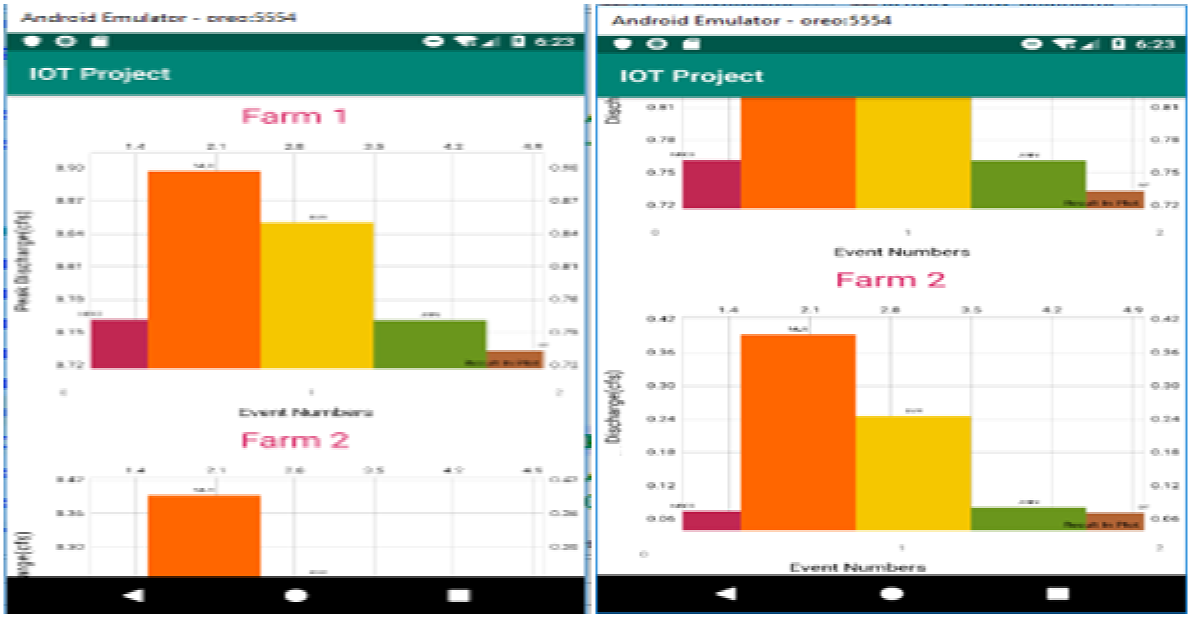


Figure 5.8 Total discharge from farm 1 and farm 2 against event numbers

Here Figure 5.8 shows the total discharge values against event numbers for farm 1 and farm 2 in bar graphs form.

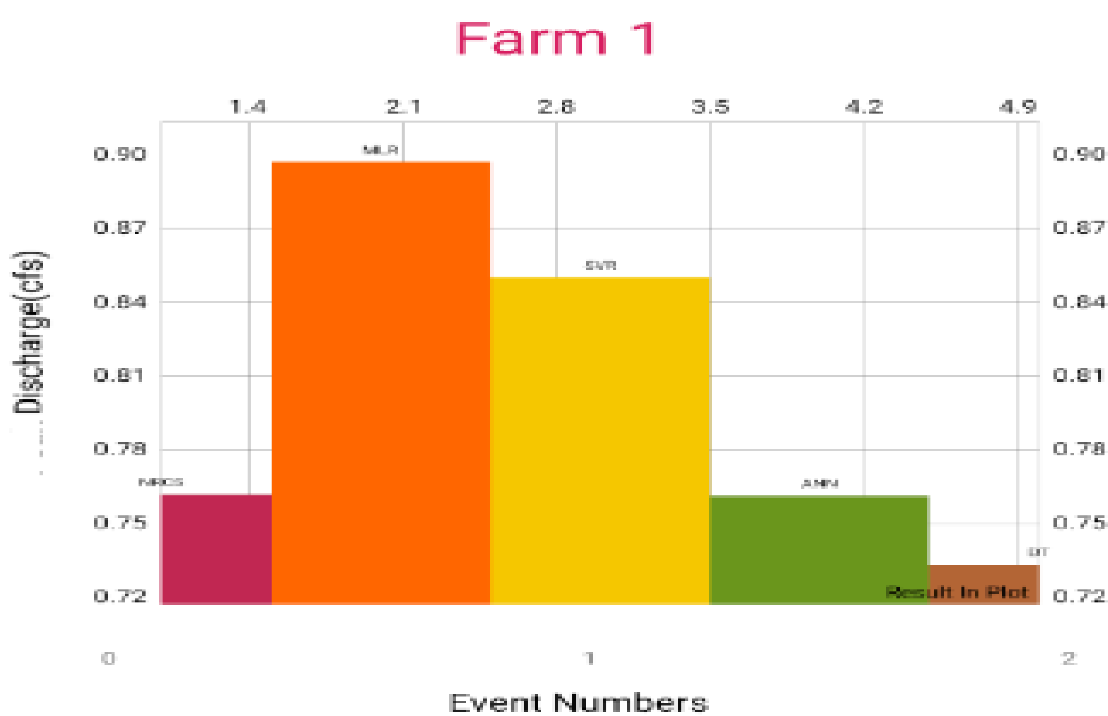


Figure 5.9 Discharge prediction by algorithms at farm1

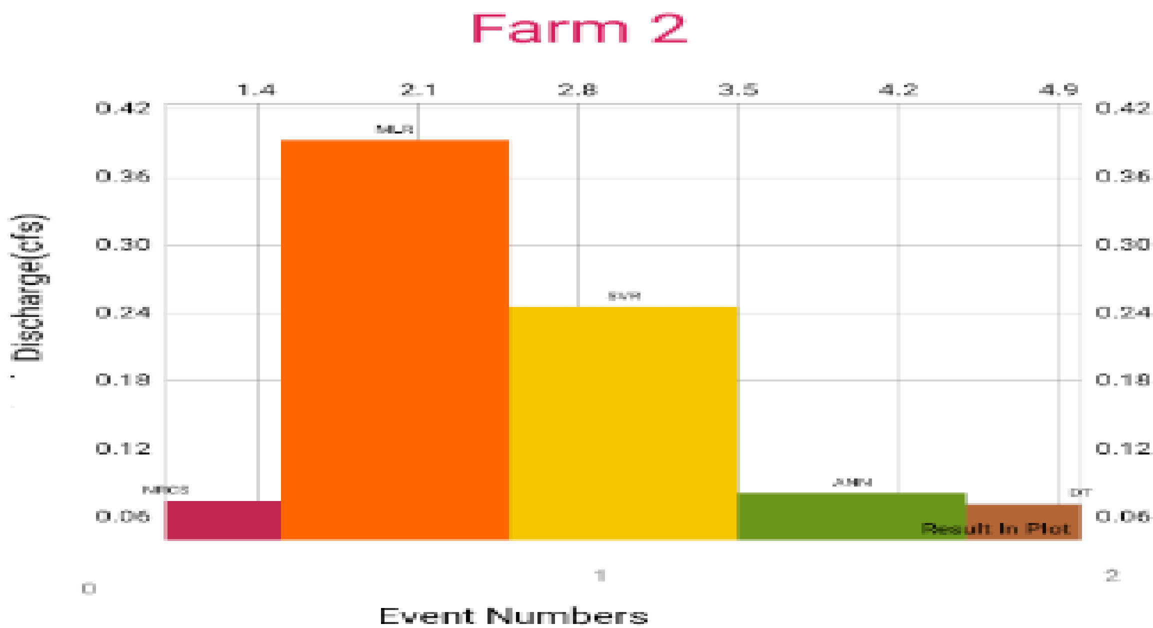


Figure 5.10 Discharge prediction by algorithms at farm2

The X–axis represented in the above bar graph for farm1 Event numbers (sample numbers/ tuple numbers which is in the range of 0 and 2 that is equal to 1) while the Y-axis shows the discharge in cubic feet per second (cfs).Figure 5.9 shows that the NRCS value for discharge predication is closely predicated by ANN , then followed by DT however SVR and MLR predication is not that closer to the NRCS predication value for discharge.

Similary Figure 5.10 shows that the NRCS predicated value for discharge at farm2 outlet is closely predicated by ANN, followed by DT, however SVR and MLR predication is not that closer to the NRCS value. . Both SVR and MLR shows a high rise away from the NRCS predicated value in the bar graph.

5.3.2 Results of Wet condition

The android emulator Figure 5.11 below for mobile application shows farm 1 and farm2 sensor variables such as temperature , soil moisture , crop stage and precipitation depth (or irrigation depth) variable , here the crop stage 1 is selected on both fields, temperature value is 25 degree Celsius for both farm 1 , farm2 and soil moisture is equal to wet for both fields 1 and 2. Irrigation or precipitation depth value = 50 mm.

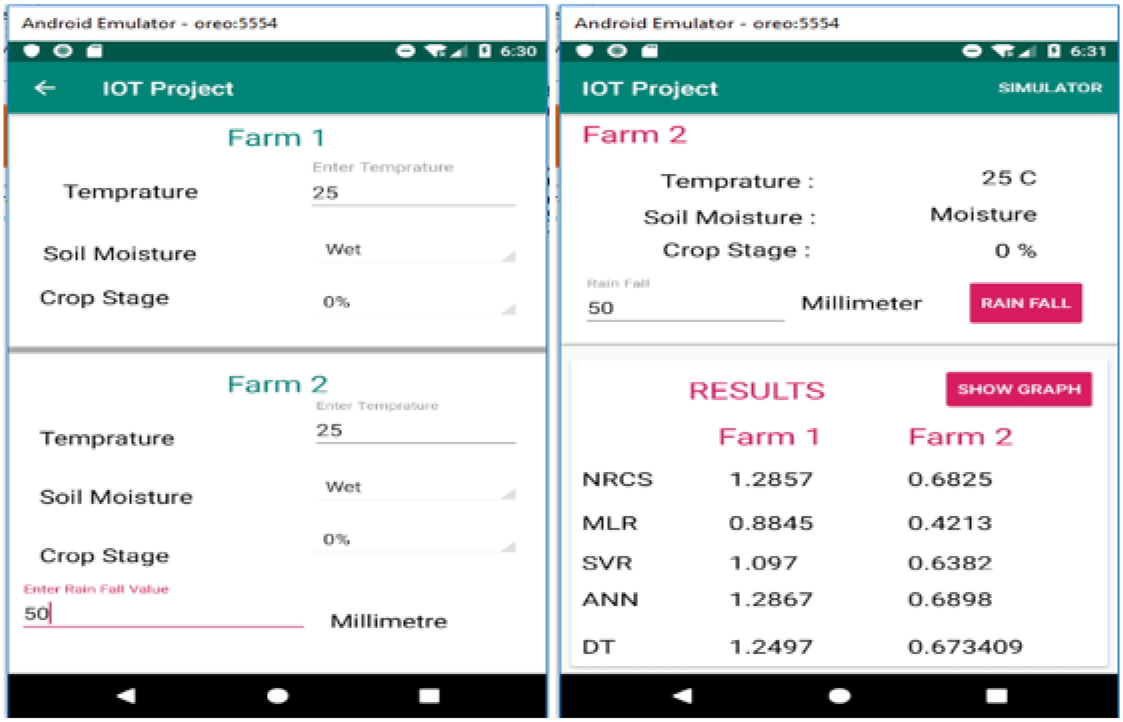


Figure 5.11 Parameters values and discharge prediction at farm1 and farm2 outlet for wet condition

Figure 5.11 shows farm 1 and farm 2 results for the predication of total discharge. ANN, DT show very closer predication values as compare to NRCS predicated value however SVR and MLR are far away from the NRCS predicated value . Here NRCS predicated value is closely predicted by ANN (ist rank), then followed by DT 2^{nd} rank) , followed by SVR(3^{rd} rank) and at last by MLR (4rth rank) on the basis of predication value accuracy.

Figure 5.12 results of different algorithms for wet condition

The above android emulator Figure 5.12 shows the results of different algorithms for ANN, DT and SVR , MLR in bar graphs form.

The bar graph shows discharge in cubic feet per second for both farm 1 and farm 2 on x-axis, while the bar graph on y-axis shows event numbers which is equal to 1 and is in the range of 0 and 2 for both farm 1 and farm 2. Figure 5.12 farm1 bar graph show that the NRCS predicated value for total discharge is equal to 1.28 cfs on farm 1 outlet which is very closely predicated by ANN discharge value =1.28 ,followed by DT =1.24 and however SVR =1.09 and MLR =0.88 which is not that closer to the predicated discharge value for the NRCS .The farm 2 results Figure 5.12 farm 2 show similar tendency like farm 1 , here the NRCS predicated discharge value =0.68 cfs which is closely predicated by ANN=0.68 and DT =0.67, however SVR=0.63 and MLR=0.42 is not very well/ closely predicated the discharge.

Similary for average wet soil moisture conditions on both farms 1 and 2 and extreme wet soil moisture conditions are also tested on the emulator.

5.4 Predication of peak discharge two farm system and its Hydrographs.

Here the predication of the peak discharge is also carried out similar like total discharge predication on the mobile app emulator, However the results are displayed in the hydrographs form for the farm1 and farm 2 outlets for various soil moisture and crop stage conditions. Here only the android studio based mobile emulator are displayed for different scenarios.

5.4.1 Dry conditions & crop stage 1

When both fields 1 and 2 have dry conditions and crop stage 1 , then the hydrographs displayed by IOT based mobile app emulator developed in android studio is as follow.

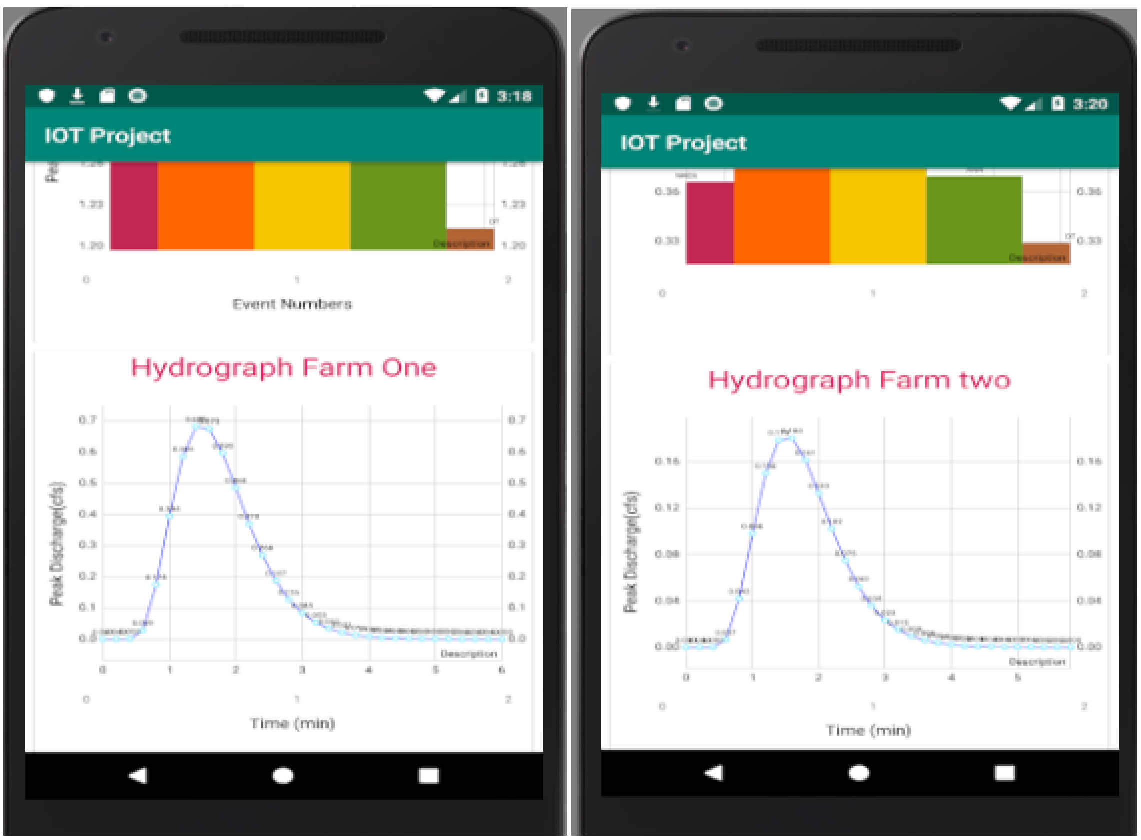


Figure 5.13 Hydrograph farm 1 & farm 2 for dry condition

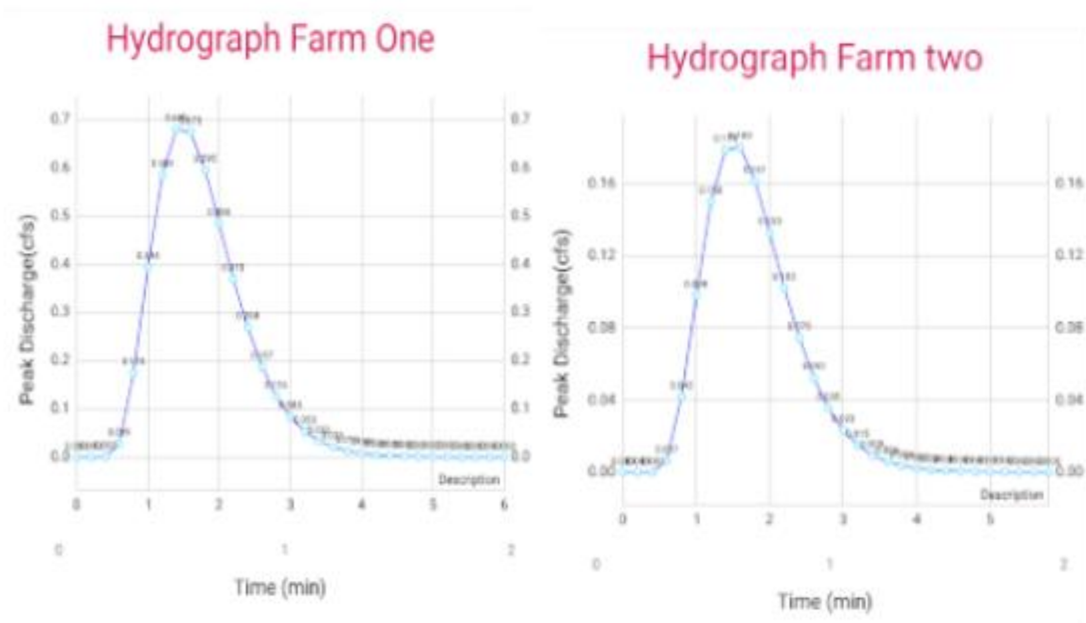


Figure 5.14 Hydrograph farm 1 & farm 2 for dry condition

The above hydrographs figure 5.14 show peak discharge (cfs) in x-axis and time to peak (mins) in y-axis for both farm 1 outlet and farm 2 outlet , here the farm1 outlet hydrographs is higher than the farm 2 outlet.

5.4.2 Wet conditions and crop stage 1

Both the farms 1 and farm 2 conditions are wet and crop stage 1 is selected for this scenario . the hydrographs are generated for both farms 1 and 2 outlet which shows peak discharge (cfs) and time to peak (mins).

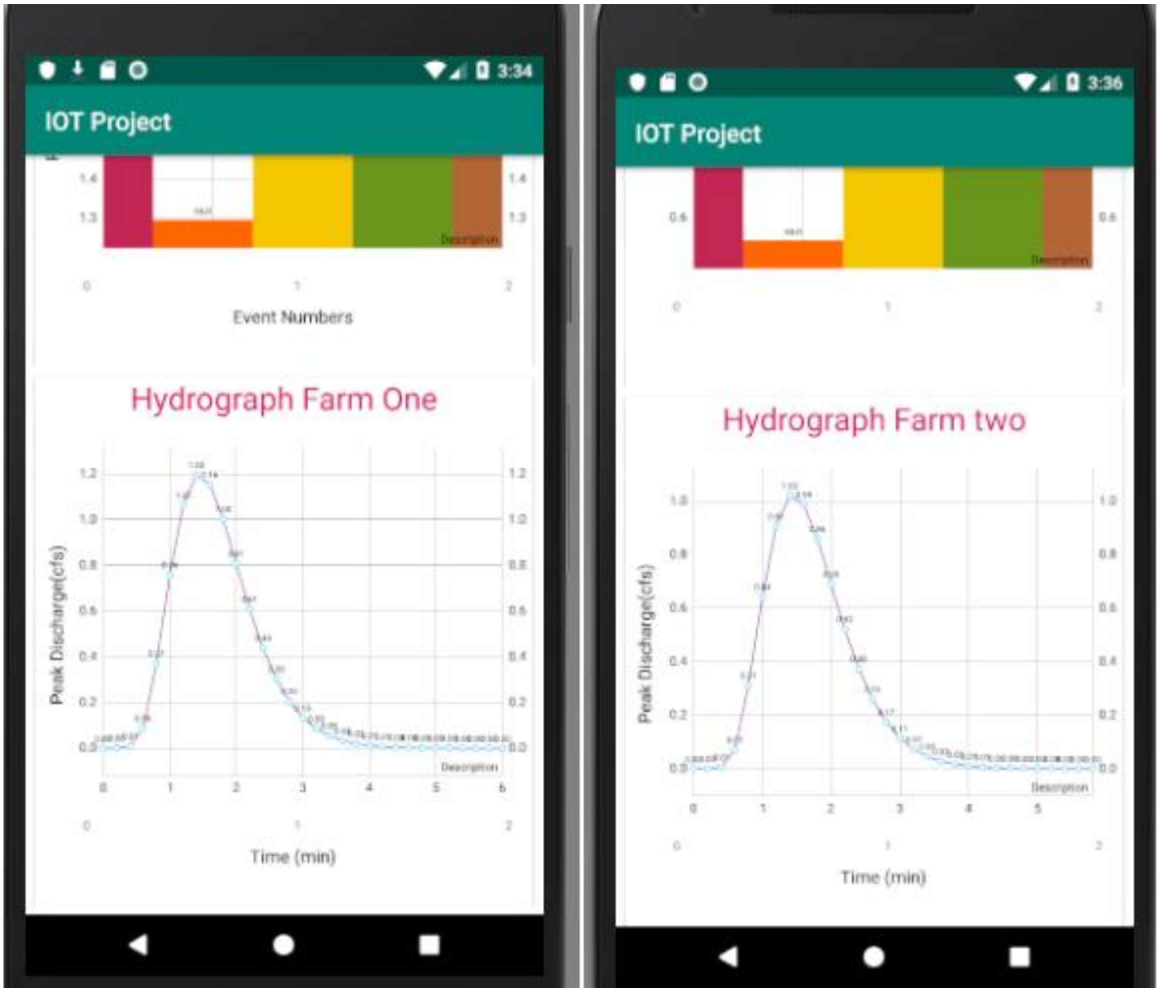


Figure 5.15 hydrograph for farm 1 and farm2 for wet condition

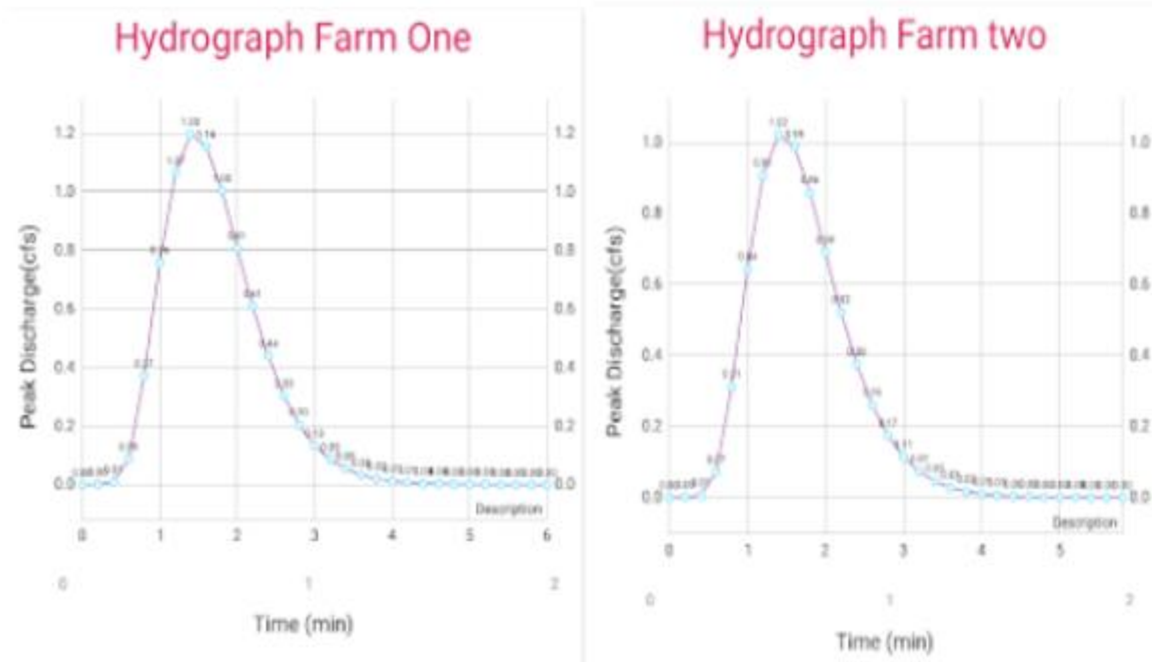


Figure 5.16 hydrograph for farm 1 and farm2 for wet condition

The x-axis of the figure 5.16 shows peak discharge (cfs) while y-axis time to peak (mins. the hydrograph of farm 1 shows higher peak discharge than farm 2 hydrographs, if overall the soil moisture conditions dry, average and wet and crop stage 1 on both farms are considered , then the results shows that for dry soil moisture conditions and crop stage 1 its both farm 1 and farm 2 hydrographs have lowest values for the peak discharge. The average soil moisture conditions and crop stage 1 on both the farms 1 and farm 2 and its outlets hydrographs are higher than dry conditions based fields, however the hydrographs shows the highest values for the wet conditions and crop stage 1 on both farm 1 and farm 2 outlets.

Similary for average wet conditions and crop stage 1 on boths fields are also trained and tested however its graphical results are not included due to space limit in this thesis , however its results are included in below table 5.1.

Table 5.1 Hydrograph peak discharge comparison

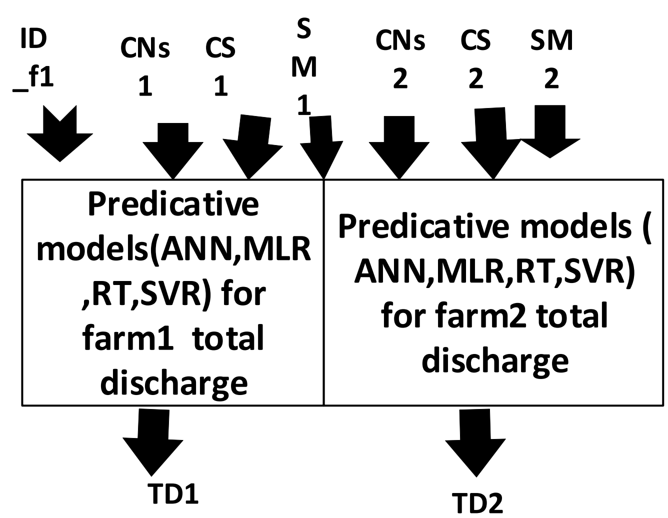
Wet conditions and crop stage 1 on both fields	Highest hydrograph at farm 1 outlet	Peak discharge value =1.2 cfs	lowest hydrograph at farm 2 outlet	Peak discharge value= 1 cfs
Average conditions and crop stage 1 on both fields	Averagely high hydrograph at farm1 outlet	Peak discharge value =1.0 cfs	Average low hydrograph value at farm 2 outlet	Peak discharge value =0.7 cfs
Dry conditions and crop stage 1 on both fields	Lowest hydrograph at farm 1 outlet	Peak discharge value =0.7 cfs	Very low hydrograph at farm 2 outlet	Peak discharge value =0.16

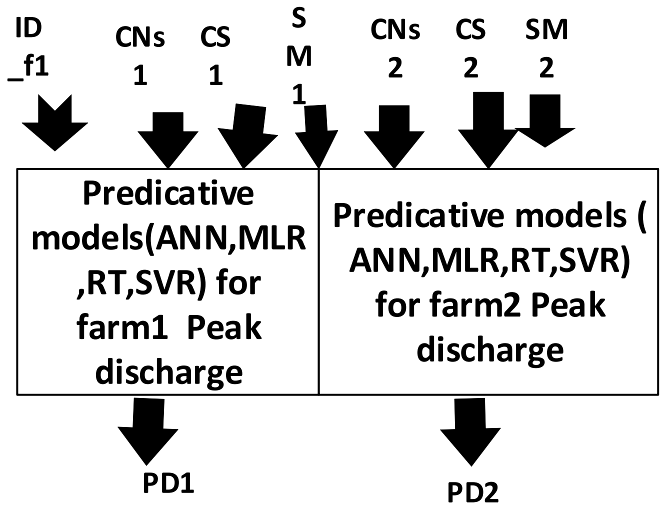
Chapter 6: Results and Discussion

6.1 Development of two Farm Prediction Model for total Discharge

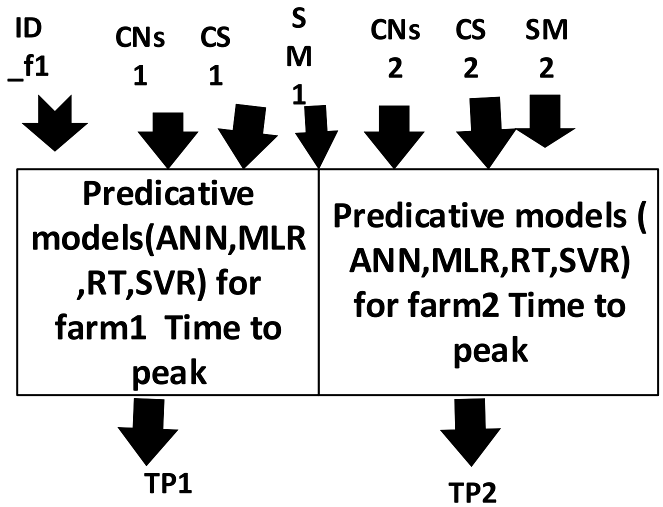
Two farm predicative irrigation system uses different parameters from the fields. i.e. from field 1, the parameter are the crop stage and the soil moisture of that field, while from the second field the parameters are the crop stage and soil moisture of the field 2. The irrigation depth is also input to the model from the farm1. Temperature input variable values= 25 degree Celsius are constant and only used in mobile application rather than desktop based application and below mentioned modelling. Using these parameters, Three different machine models are developed on the basis of different output parameters that are, 1) for the predication of total discharges at farm 1 and farm 2 outlets at unit cubic feet per second.In the second modeling approach 2), peak discharges are predicated at farm 1 and farm 2 outlets. In the third modeling approach 3), time to peak are predicated at farm1 and farm 2 outlets. Different machine learning models are trained and tested on randomized different split of data and on the basis of cross validation to evaluate the models . Figure 6.1 shows systematic diagram of two farms.

- ID_f1 stands for irrigation depth input parameter at farm 1.
- CNs 1 stands for Curve numbers input parameter selected at farm1 .
- SM1 stands for soil moisture input parameter conditions selected at farm1.
- CS 1 stand for crop stage input parameter selected at farm1.
- CNs 2 stands for Curve numbers input parameter selected at farm 2.
- SM2 stands for soil moisture input parameter selected at farm 2.
- CS2 stands for Crop stage input parameter selected at farm 2.
- TD1 stands for an output parameter total discharge from farm 1.
- TD2 stands for an output parameter total discharge from farm 2.





- PD1 stands for peak discharge output parameter at farm1 .
- PD2 stands for peak discharge output parameter at farm2.



- TP1 stands for time to peak output parameter at farm 1.
- TP2 stands for time to peak output parameter at farm2.

Figure 6.1 Systematic diagram of two farms

Table 6.1 Data set for two farms

twoFormDataset.xlsx

	A	B	C	D	E	F	G	H	I	J
					twoFormDataset					
	CS_1	SM_1	CS_2	SM_2	CN_1	CN_2	ID_1	CFS_1	MM_1	CFS_2
	Number	Number	Number	Number	Number	Number	Number	Number	Number	Number
1	1	1	1	1	84.8024	84.8024	1	0	0	0
2	1	1	1	1	84.8024	84.8024	9	0	0	0
3	1	1	1	1	84.8024	84.8024	17	0.4669	0.0460	0
4	1	1	1	1	84.8024	84.8024	25	1.6457	0.1620	0
5	1	1	1	1	84.8024	84.8024	33	3.2904	0.3239	0
6	1	1	1	1	84.8024	84.8024	41	5.2566	0.5174	0
7	1	1	1	1	84.8024	84.8024	49	7.4539	0.7336	0
8	1	1	1	1	84.8024	84.8024	57	9.8229	0.9668	4.4010e-04
9	1	1	1	1	84.8024	84.8024	65	12.3230	1.2129	0.0084
10	1	1	1	2	84.8024	84.8024	1	0	0	0
11	1	1	1	2	84.8024	84.8024	9	0	0	0
12	1	1	1	2	84.8024	84.8024	17	0.4669	0.0460	0
13	1	1	1	2	84.8024	84.8024	25	1.6457	0.1620	0
14	1	1	1	2	84.8024	84.8024	33	3.2904	0.3239	0

The total number of dataset samples or tuples are 1134 that has been acquired from the NRCS scripting through NRCS simulator that include different curve numbers for farm 1 and different curve numbers for farm 2 and irrigation depth at farm 1.

The data is also acquired from sensors related to Curve numbers for farm 1 and curve numbers for farm 2 which are input data such as soil moisture , crop stage for farm 1 and soil moisture ,crop stage for farm 2.

while the output data is the total discharge at farm1 and farm 2 outlet which are represented in the above dataset table as CFS_1 (cubic feet per second) and CFS_2 respectively .However MM_1 is a column in the table where CFS_1 is converted into Cubic millimeter per second but this column is not used in the modeling .

It must be noted here that the data samples are independent (that is IID) , independent and identically distributed. This is not a time series forecasting rather it is a predication.

Multiple linear equations for discharge models.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \qquad (19)$$

Here $X_{1=}$ irrigation depth , $X_{2=}$ Curve numbers, X_3=Soil moisture, X_4 = Crop stage while

β_o= is the value of Y when all the independent variables is equal to zero and is called coefficient ,β_1 ,β_2 , β_3,β_4 are also called regression coefficients and each coefficient shows change in Y dependent variable corresponding to one unit change in independent variable. While holding all the variables as a constant. Statistical experiments tests and assess that the value of each variable coefficeint is how much different from zero.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 \qquad (20)$$

Here X_5=Curve numbers for farm2, X_6= soil moisture of farm 2 , X_7=crop stage of farm 2.while Y = Total discharge . Similar equations are developed for Peak discharge and Time to peak at farm 1 and farm 2.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \qquad (21)$$

Y = peak discharge at farm 1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 \qquad (22)$$

Y= peak discharge at farm 2.

Similarly,

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \qquad (23)$$

Y = Time to Peak at farm 1.

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 \qquad (24)$$

Y= Time to peak at farm 2.

Regarding the Neural network hidden structure, the number of hidden neurons that are utilized are equal to 10. On 70:30 ratio random data split 70 percent data is used for training, 15 percent for validation and 15 for testing the model. the neural network Levenberg Marquardt (ANN-LMA) back propagation algorithm is used for training and testing. LMA is like quasi-Netwon method and the approach that LMA uses is second order training speed.LMA does not compute Hessian matrix and to compute gradient it uses Jacobian matrix. the Back propagation simply means backward propagation of errors. In the appendix its Matlab based hidden structure is shown. 'Trainlm' is the neural network training function. The purpose of 'trainlm' is to modify the weight and bias values on the basis of Levenberg Marquardt optimization. The reason for its selection is due to its fastest training capability in Matlab and also very highly recommended supervised learning algorithm. By default 1000 epochs are used to train the neural network. The validation in the neural network is the process in which validation vectors are used to quit the training process earlier if the network start degrade or remain the same while the testing in the neural network uses test vectors which ensures that the network generalize well and its effect on training is null. The downside is too much memory required than other algorithms. Neural network range from single layer to multiple layer and recently the work is in progress in deep learning. Support vector machine can be used as for classification or regression. Support vector machine creates a hyperplane that maximize the margin of separation between two classes of data. The type of SVM based on kernels are Gaussian or Radial basis function (RBF), linear , polynomial and Sigmoid. In the case of Support vector regression , linear support vector regression is used and its Matlab code or formula is shown in the appendix . To create RegressionSVM model in Matlab 'fitrsvm' is used. The kernel used is linear . Regression SVM model store data, parameters values, algorithmatic implementation and support vectors. ‘Predict’ keyword is used to predict values for new data in Matlab. Regression SVM also computes the mean square error and epsilon-insensitive loss through ‘loss’ function. Decision tree is binary split for regression (regression tree) is utilized for training and testing from the Matlab toolbox. ‘fitrtree’ is used to create regression tree object. fitrtree accepts the predictors and response data and other input parameters. ’Predict’ is used to make predication for new data. The object function ‘loss’ is used to trace errors. ’Prune’ function is used to create sequence of subtrees and ‘view’ function is used to view regression tree.

6.2 Discharge at the farm 1 outlet Predicative Model Results

These results are predicated for the total discharge in cfs for the farm 1 outlet .

6.2.1 Case A Results for Discharge Prediction on farm1 outlet

The data is divided into 70:30 dataset randomly, out of which 70 percent is used for training and 30 percent is used for testing.

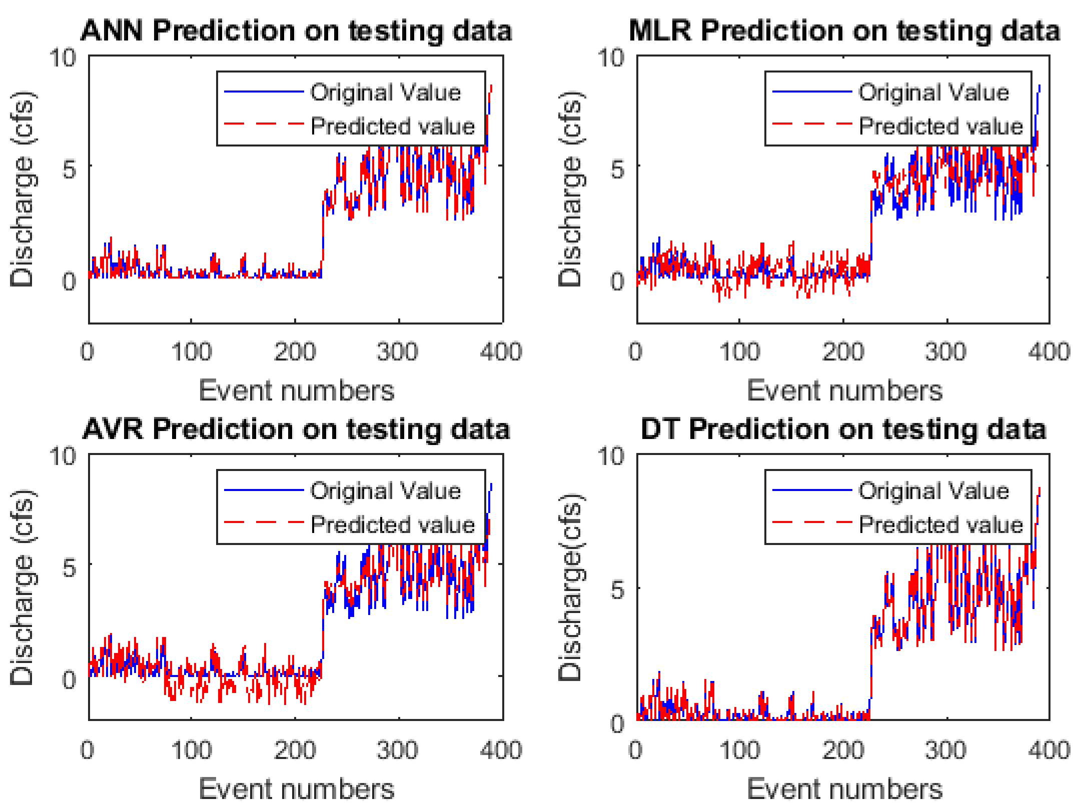


Figure 6.2 Discharge prediction for case A at farm outlet

Table 6.2 Algorithms performance for discharge prediction for case A at farm1 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	8.1297	0.00019411	0.0055351	0.99997	0.013932	0.0063004	0.0003311
MLR	0.24899	0.43102	0.26083	0.92432	0.65652	0.29193	0.03788
SVR	3.7918	0.25792	0.20177	0.95763	0.50786	0.23947	-0.090265
Decision Trees	0.58828	0.0009501	0.012246	0.99985	0.030824	0.013926	0.0023858

Figure 6.2 shows that DT and ANN show excellent predication and excellent fit . However SVR and MLR show not very good fit as compare to the other two algorithms . From Table 6.29 above it is also clear that R –square of the ANN and DT are outstanding 0.99997 and 0.99985 respectively , however the SVR and MLR R-square are 0.957 and 0.924 respectively shows intermediate fit when compare to other algorithms R-square . The benchmark metrics MSE,NRMSE, RMSE and RRMSE and MAE show that ANN shows minimum error RMSE= 0.013 , followed by decision tree with RMSE=0.030 , however the MLR shows highest error MSE=0.43 than SVR with MSE=0.25 and then compare to the other two machine learning algorithms. The reason is that ANN is a complex algorithm and it work well when data provided to it is scarce as compare to other algorithms.

6.2.2 Case B Results for Discharge Prediction at farm1 outlet

The dataset is split into 30:70 percent, here 30 percent divided is used for training and 70 percent data is used for testing.

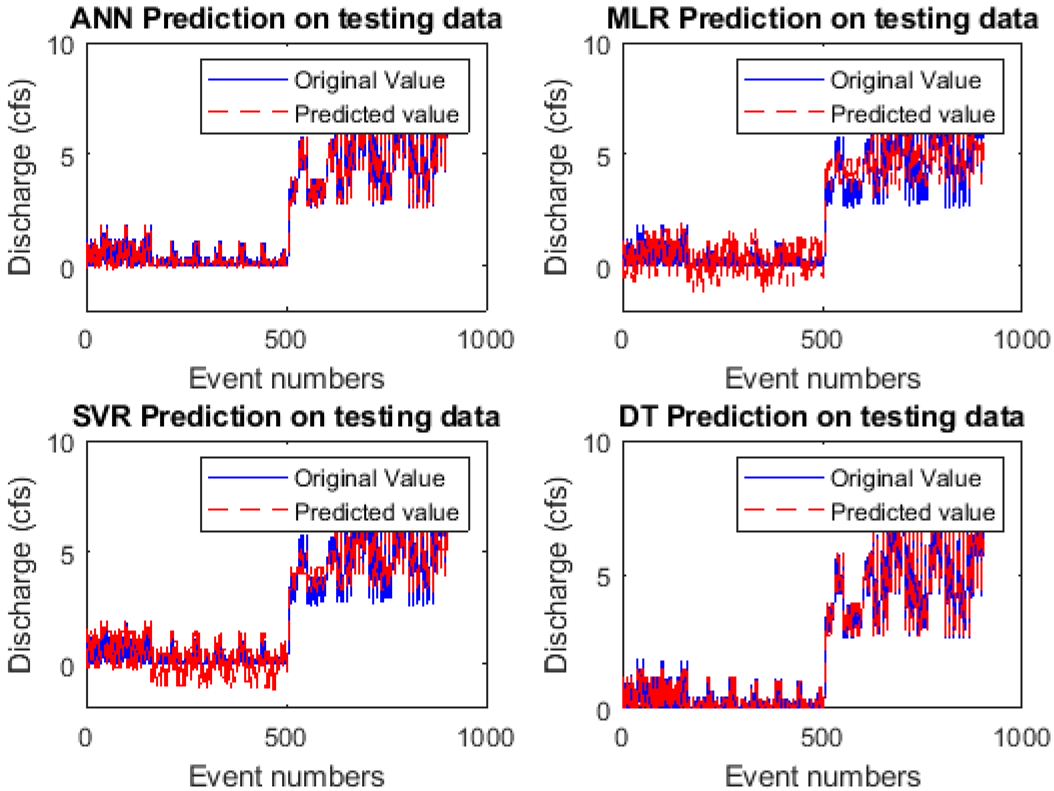


Figure 6.3 Case B results for discharge prediction at farm1 outlet

Table 6.3 Algorithms performance for Case B at farm1 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	0.51658	0.0013791	0.014811	0.99978	0.037137	0.015923	0.0035713
MLR	0.0019706	0.40937	0.25518	0.92895	0.63982	0.27082	0.033898
SVR	0.14594	0.23751	0.19436	0.95943	0.48735	0.21207	-0.030521
Decision Trees	0.062854	0.010539	0.040943	0.99833	0.10266	0.044059	0.0014498

Here Figure 6.3 shows the results of case B. It is clear that ANN and DT the original and predicated lines blue and red dashed clearly overlaps each other while there is difference between original and predicated lines when the above figures for MLR and SVR are retrieved. In Table 6.3 R-square clearly shows that ANN and DT is 0.99978 and 0.99833 shows good predication and fit while the SVR and MLR is 0.95 and 0.92 respectively shows less good fit and predication . on the basis of MSE,NRMSE,RMSE,RRMSE and MAE the ANN show lowest error MSE= 0.001 , followed by DT with MSE=0.010, followed by SVR with MSE=0.23 and then MLR with highest MSE value =0.40. However the training time of ANN is highest, followed by Decision tree, followed by SVR while MLR is the minimum training time/execution time. The reason is that ANN is a complex algorithm and it work well when data provided to it is scarce as compare to other algorithms.

6.2.3 Case C Results for Discharge Prediction on a single farm

Case C use k=10 fold cross validation on the dataset.

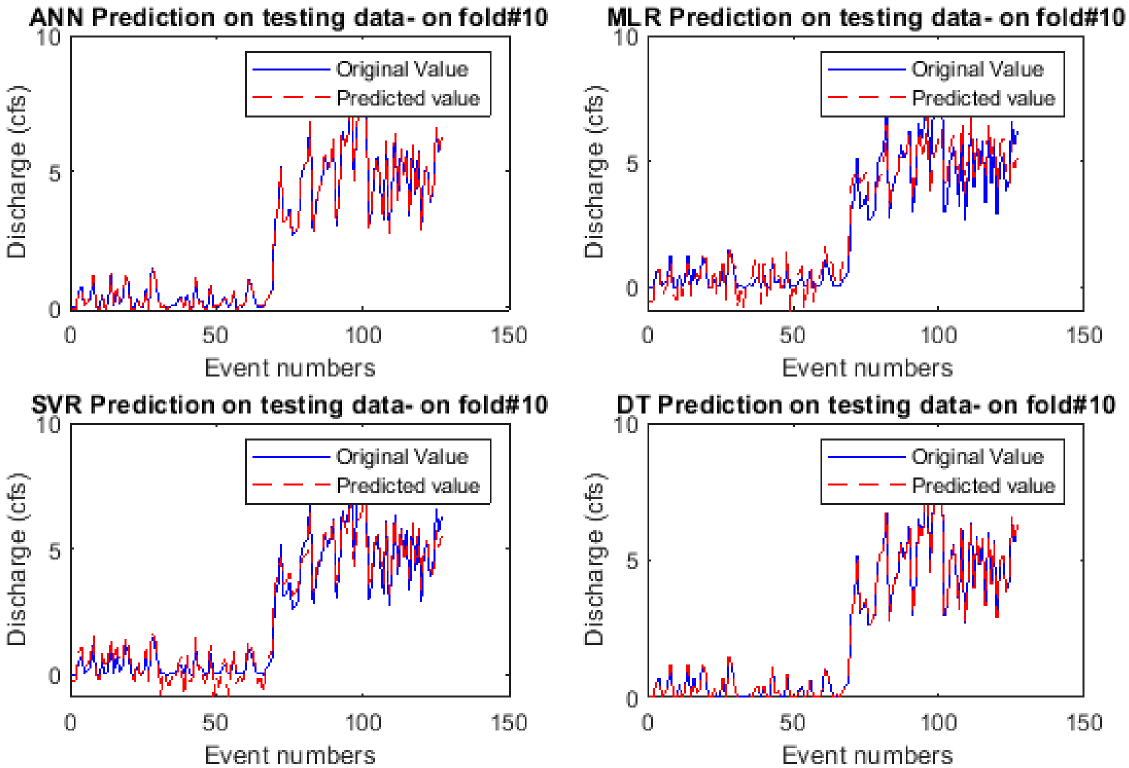


Figure 6.4 Case C results for discharge prediction on farm1 outlet

Table 6.4 Algorithms performance for case C at farm1 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	3.8938	5.2845e-05	0.0024214	0.99999	0.006067	0.0026289	-4.4287e-05
MLR	0.0087067	0.40387	0.25089	0.93048	0.63449	0.26925	0.03237
SVR	0.55819	0.25139	0.89171	0.95795	0.50062	0.22308	-0.077328
Decision Trees	0.094024	0.00055733	0.0093113	0.99991	0.023544	0.010128	-3.1274e-05

The above Figure 6.4 clearly show that for the 10 fold cross validation the ANN and DT the predication are quiet good as well. Table 6.4 shows that the R-square value of ANN and DT is maximum and near to one (0.99999 and 0.99991) however the SVR and MLR R square values = 0.95 and 0.93 respectively . The MSE values of SVR and MLR is below 0.5 clearly show that MLR and SVR does not perform very well during 10 fold cross validation's predication and goodness of fit. The MSE, NRMSE, RMSE, RRMSE and MAE of ANN is minimum followed by Decision tree while the SVR and MLR is higher than the other two algorithms such that SVR is second last while MLR is the last in the algorithms table list on the basis of maximum error. On the basis of training time on 10 fold cross validation ANN takes a lot of time than all other algorithms while MLR shows minimum time to train/execute .

Table 6.5 Overall performance of algorithms for discharge prediction at farm1 outlet

Algorithms performance	Training time	70:30 dataset	30:70 dataset	K=10 fold cross validation
ANN	satisfactory	Excellent	excellent	Excellent
DT	very good	very good	very good	very good
SVR	good	Good	Good	Good
MLR	excellent	Satisfactory	satisfactory	Satisfactory

6.3 Predicative Model Results at Farm 2 outlet

6.3.1 Case A Results for Discharge Prediction on farm2 outlet

In case A the dataset is randomly divided into 70:30 in which 70 percent data samples are trained for training while the remaining 30 percent is used for testing.

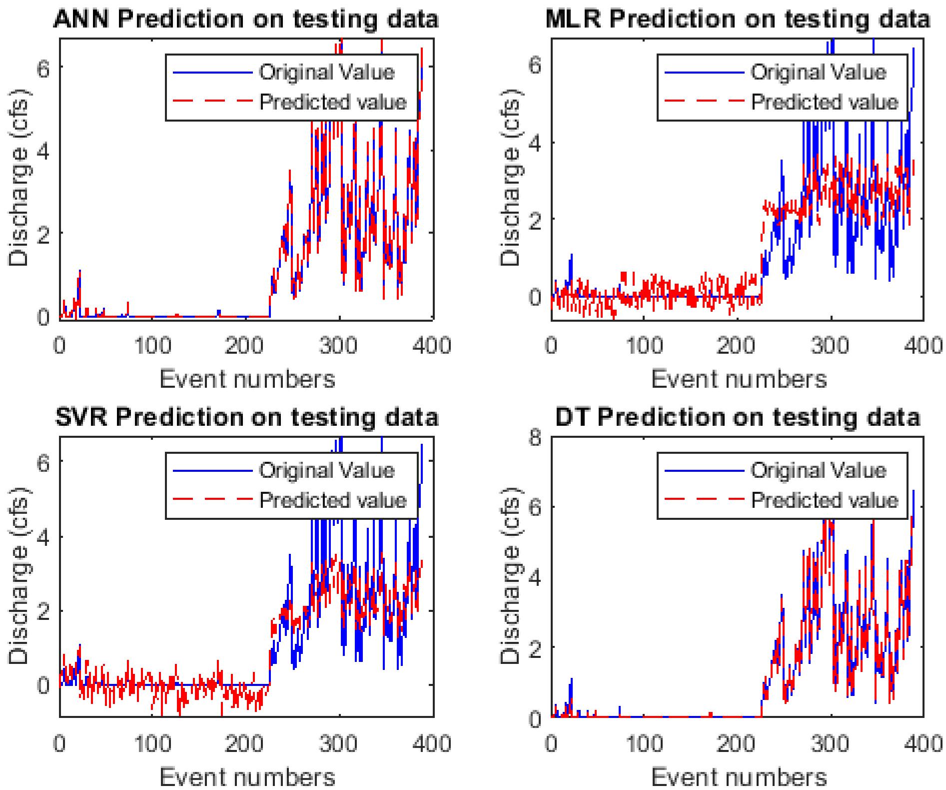


Figure 6.5 Discharge prediction at farm2 outlet for case A

Table 6.6 Algorithms performance for discharge prediction for Case A

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	1.5841	7.7891e-05	0.005367	0.99997	0.0088256	0.0080219	0.00090677
MLR	0.0015999	0.76163	0.53071	0.56283	0.87271	0.78632	0.01058
SVR	0.3975	0.59573	0.46936	0.61229	0.77183	0.84063	-0.18112
Decision Trees	0.16799	0.037567	0.11787	0.98576	0.19382	0.1775	-0.0073252

From the visual analysis of the above Figure 6.5 the peak discharge at the field 2 outlet show that ANN predication is excellent , followed by the decision tree the predicated and original lines clearly overlaps each other , however the MLR and SVR shows that the predicated and original values line does not overlap each other perfectly . In Table 6.6 the R-square shows 0.999 for the ANN and the rest of regression metrics MSE, NRMSE, RMSE , RRMSE and MAE is almost negligible and minimum while decision tree shows the minimum error RMSE value and R square value that is 0.008 and 0.985 respectively, followed by SVR with RMSE is equal to 0.771 and MLR with RMSE is equal to 0.872 .In this case 70:30 ratio the ANN perform well on minimum testing data .

6.3.2 Case B Results for Discharge Prediction on farm2 outlet

The dataset is divided into 30:70 randomly in which 30 percent of dataset samples used for training and 70 percent of dataset sample is used for testing.

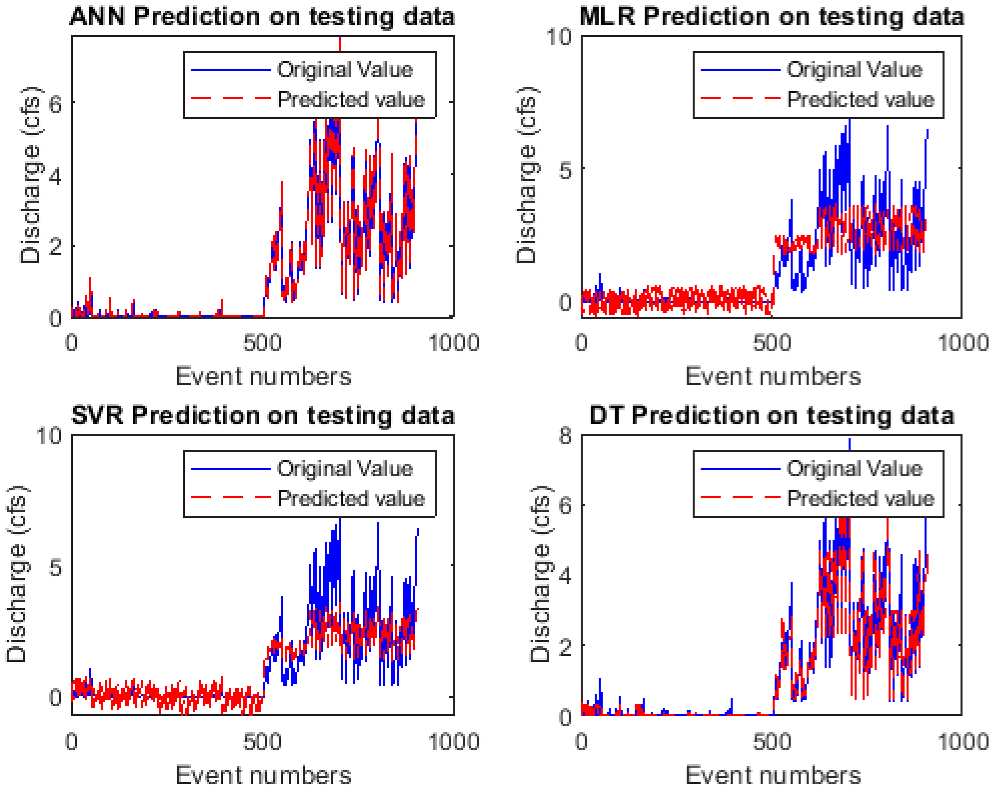


Figure 6.6 Case B results for discharge prediction at farm2 outlet

Table 6.7 Algorithms performance for discharge prediction for case B at farm2 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	1.76	2.2908e-05	0.0029336	0.99999	0.0047862	0.0041799	-0.00021942
MLR	0.0068121	0.75822	0.5337	0.55893	0.87076	0.74962	0.016308
SVR	0.25081	0.63082	0.4868	0.55415	0.79424	0.80643	-0.1604
Decision Trees	0.14583	0.14139	0.23046	0.94051	0.37601	0.34145	-0.044063

Similarly from the visual Figure 6.6 above as the tested data size is increased to 70 % of dataset while training dataset is 30% of total selected database. The ANN shows quiet good predication and fit again followed by DT while MLR and SVR shows a poor fit . Table 6.7 indicates the MSE = 0.00005 , NRMSE =0.002 , RRMSE=0.004 and MAE=-0.0002 values of ANN is less as compare to all other algorithms. The second algorithm in the list which shows minimum error on the regression metrics is DT with MSE=0.141 and r-square value=0.949, followed by SVR with MSE=0.653 and then finally MLR with MSE=0.758. However the training time is higher for ANN followed by SVR , while MLR shows the minimum algorithm execution time for the results generated followed by decision tree.

6.3.3 Case C Results for Discharge Prediction on farm2 outlet

The following Figure 6.7 shows results of case C at farm2 outlet.

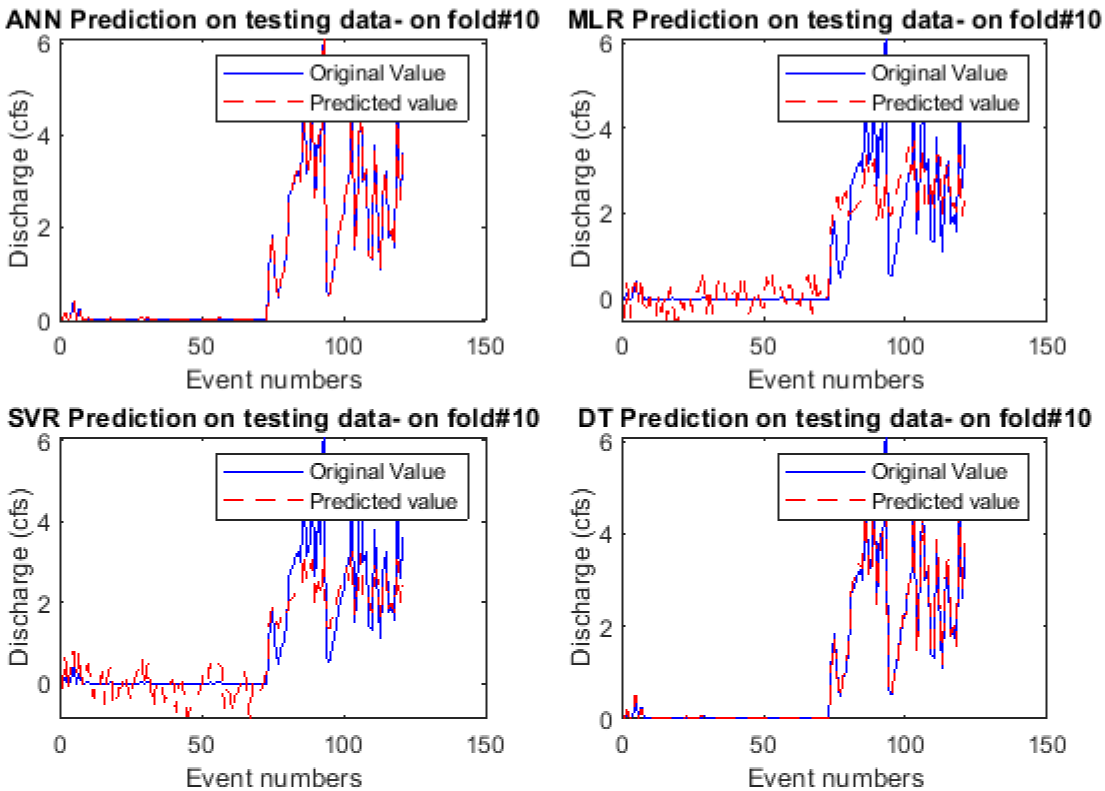


Figure 6.7 Case C Results for Discharge Prediction on farm2 outlet

Table 6.8 Algorithms perform for Case C for Discharge Prediction on farm2 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	2.7828	0.00013726	0.006162	0.99994	0.0097493	0.0088039	-5.5546e-05
MLR	0.00034903	0.73248	0.52904	0.58971	0.85251	0.72688	0.044187
SVR	0.21288	0.57026	1.4078	0.62373	0.75078	0.75602	-0.13688
Decision Trees	0.021866	0.022374	0.089516	0.9914	0.14548	0.1289	-0.0035537

From the visual display from the above figure 6.7 the ANN predication is accurate and fit is good followed by DT on k=10 fold cross validation , How ever the SVR and MLR does not show a good fit , the original value line and predicated value line does not clearly overlap each other .if the above table 6.8 is rendered here such as MSE, NRMSE, RMSE, RRMSE and MAE it shows that ANN on the cross validation perform excellent again with the minimum error values such as MSE of ANN=0.0001 and R square =0.99994 , followed by DT with MSE =0.0223 and R square =0.9914 however the MSE of SVR = 0.5702 and R square= 0.6237 and MLR with MSE=0.732 and R square =0.589.

Table 6.9 Overall performance of algorithms at farm2 outlet

Algorithms performance	Training time	70:30 dataset	30:70 dataset	K=10 fold cross validation
ANN	satisfactory	excellent	excellent	excellent
DT	very good	very good	very good	very good
SVR	good	good	Good	good
MLR	excellent	satisfactory	satisfactory	Satisfactory

6.4 Peak Discharge and Time to Peak Discharge Predication Model

In these predicative models for peak discharge and time to peak at farm 1 and farm 2 outlets, a total of four regression based models have been developed , trained and tested . The inputs parameters to these models are crop stage , soil moisture of farm1 , crop stage , soil moisture of farm2 (CN1 and CN2) and irrigation depth of farm 1 while the outputs parameters of the models are peak discharge at outlet 1 and time to peak at outlet 1 (farm1) and peak discharge at outlet 2 and time to peak at outlet 2 (farm 2) .

The following Table 6.10 show dataset for Peak Discharge and Time to Peak Discharge Predication Model. The dataset consists of 891 samples.

The data nature is independent that is independent and identically distributed (IID). That is a row of data sample is independent from other row of data sample in the dataset/database.

Table 6.10 Dataset for Peak Discharge and Time to Peak Discharge Predication Model

	A	B	C	D	E	F	G	H	I	J	K	L
			PeakDischargeAndTimeToPeakForFarm1OutletAndFarm2outletDataset									
	CS_1	SM_1	CS_2	SM_2	CN1	CN2	ID_1	...	PD_outlet1	TP_outlet1	PD_outlet2	TP_outlet2
	Number	Number	Number	Number	Number	Number	Number	.	Number	Number	Number	Number
1	1	1	1	1	84.8024	84.8024	50	...	0.3753	1.4000	0.0366	1.6000
2	1	1	1	1	84.8024	84.8024	52	...	0.4039	1.4000	0.0466	1.6000
3	1	1	1	1	84.8024	84.8024	54	...	0.4329	1.4000	0.0576	1.6000
4	1	1	1	1	84.8024	84.8024	56	...	0.4625	1.4000	0.0696	1.6000
5	1	1	1	1	84.8024	84.8024	58	...	0.4924	1.4000	0.0827	1.6000
6	1	1	1	1	84.8024	84.8024	60	...	0.5228	1.4000	0.0967	1.6000
7	1	1	1	1	84.8024	84.8024	62	...	0.5535	1.4000	0.1117	1.6000
8	1	1	1	1	84.8024	84.8024	64	...	0.5845	1.4000	0.1275	1.6000
9	1	1	1	1	84.8024	84.8024	66	...	0.6159	1.4000	0.1443	1.6000
10	1	1	1	1	84.8024	84.8024	68	...	0.6476	1.4000	0.1618	1.6000
11	1	1	1	1	84.8024	84.8024	70	...	0.6796	1.4000	0.1801	1.6000
12	1	1	1	2	84.8024	93	50	...	0.3753	1.4000	0.1347	1.4000
13	1	1	1	2	84.8024	93	52	...	0.4039	1.4000	0.1550	1.4000
14	1	1	1	2	84.8024	93	54	...	0.4329	1.4000	0.1762	1.4000
15	1	1	1	2	84.8024	93	56	...	0.4625	1.4000	0.1984	1.4000
16	1	1	1	2	84.8024	93	58	...	0.4924	1.4000	0.2214	1.4000
17	1	1	1	2	84.8024	93	60	...	0.5228	1.4000	0.2453	1.4000
18	1	1	1	2	84.8024	93	62	...	0.5535	1.4000	0.2699	1.4000
19	1	1	1	2	84.8024	93	64	...	0.5845	1.4000	0.2953	1.4000
20	1	1	1	2	84.8024	93	66	...	0.6159	1.4000	0.3213	1.4000

Here in the above table 6.10 , CS_1 stands for crop stage at farm1 , SM_1 stands for soil moisture at farm 1, CS_2 stands for crop stage at farm 2 , SM_2 stands for soil moisture at farm2, CN1 Stands for curve numbers selected at farm 1 , CN2 stands for curve numbers selected at farm 2, ID_1 stands for irrigation depth , PD_outlet 1 stands for peak discharge at farm 1, TP_outlet1 stands for time to peak at farm 1,PD_outlet 2 stands for peak discharge at farm2 and TP_outlet2 stands for time to peak at farm 2.

6.5 Peak Discharge Predication at Farm1 outlet

6.5.1 Case A Results for Peak Discharge at Farm1 outlet

The ratio of dataset split randomly is 70:30 where 70 percent of data is reserved for training and the remaining data for testing.

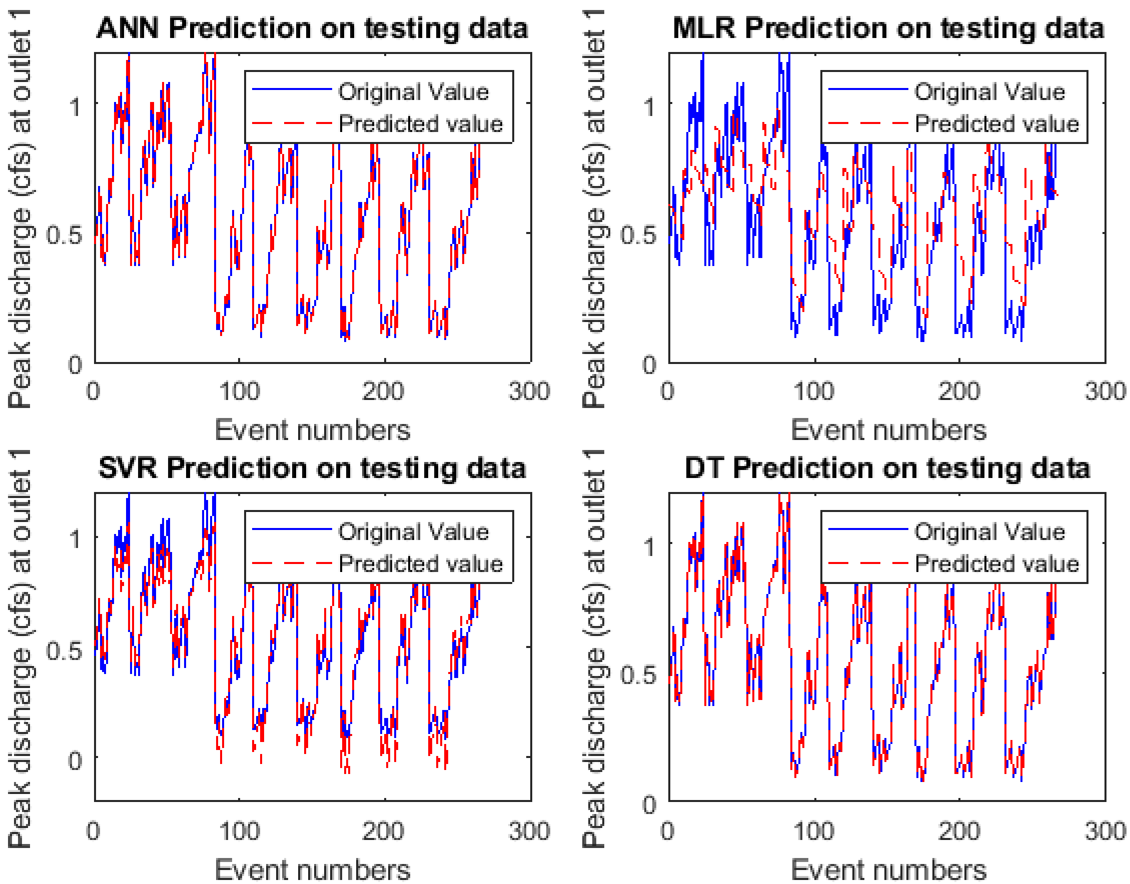


Figure 6.8 Case A Results for Peak Discharge at farm1 outlet

Table 6.11 Algorithms performance for Peak Discharge for case A at farm1 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	2.8543	4.1774e-09	0.00021777	1	6.4633e-05	0.00011684	-1.8676e-06
MLR	0.0014067	0.034249	0.62354	0.10263	0.18507	0.31236	0.039296
SVR	0.50518	0.0036048	0.20229	0.95974	0.06004	0.11175	-0.015893
Decision Trees	0.29481	1.758e-05	0.014127	0.9998	0.0041928	0.0075722	0.00053825

The above Figure 6.8 and Table 6.11 clearly shows that ANN mean square error and other regression error metrics are minimum that is a very small number RMSE= 0.0001 the r-square is perfectly 1 , shows a good fit , followed by DT and then SVR while MLR shows worse fit in terms of r –square value 0.10 and MSE is higher than other bench mark performance algorithms. ANN shows good performance when trained by a lot of numbers of samples in case of total discharge or peak discharge for two farms scenario.

6.5.2 Case C Results for Peak Discharge at Farm1 outlet

Case C k=10 fold cross validation where dataset is divided into equal 10 folds in which one fold is used for testing and remaining folds for training ,then the next fold is used for testing and the rest for training and so on .

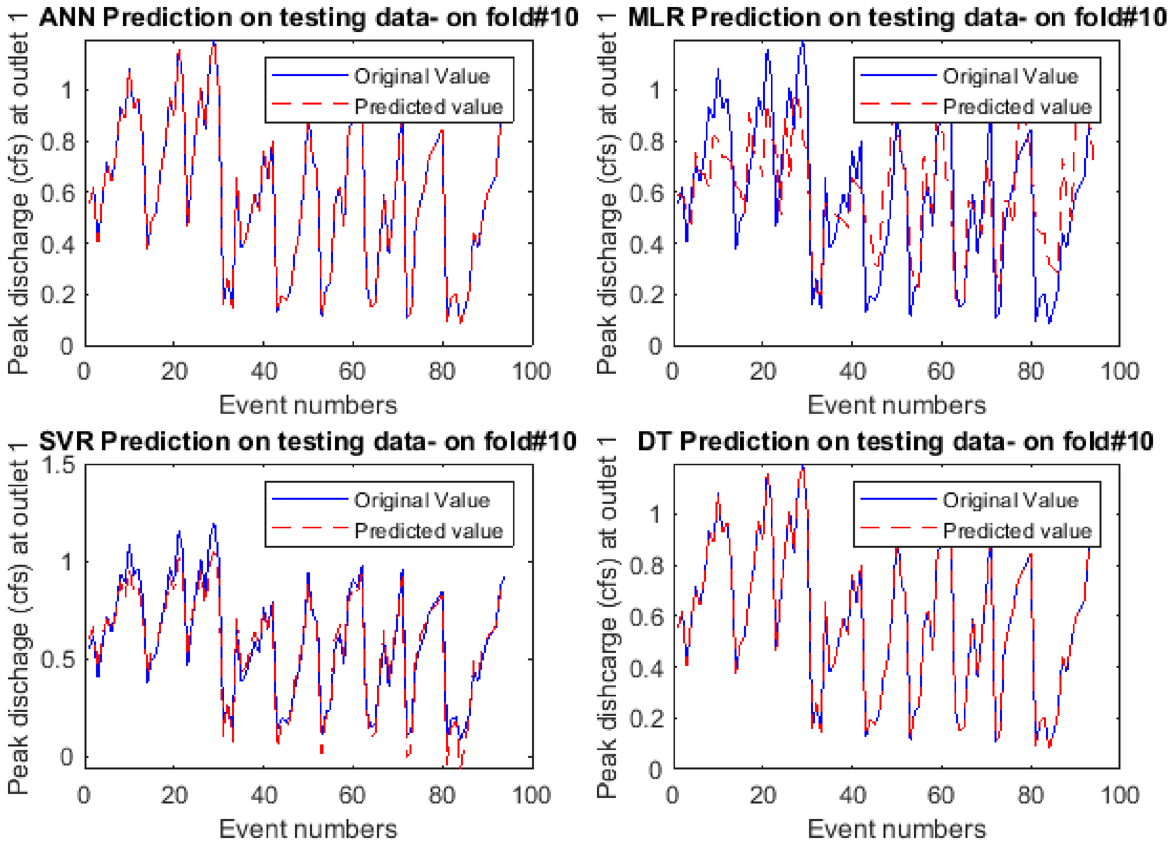


Figure 6.9 Case C results for Peak discharge prediction at farm1 outlet

Table 6.12 Algorithms performance for Case C for Peak discharge prediction at farm1 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	3.5271	3.969e-09	0.00019107	1	5.6261e-05	9.6725e-05	-8.7237e-07
MLR	0.0036512	0.032538	0.61126	0.18504	0.18016	0.30475	0.011032
SVR	0.31373	0.0035157	7.6137	0.95788	0.059275	0.10498	-0.015671
Decision Trees	0.047324	1.2363e-32	3.6082e-16	1	1.0612e-16	1.8258e-16	2.5373e-17

As can be seen from figure 6.9 and table 6.12, in K =10 fold cross validation, the DT improves its accuracy performance in terms of R -square and regression error metrics such as MSE ,RMSE etc than the case a 70:30 ratio. This is because decision tree perform well when used with cross validation for peak discharge estimation on outlet 1 of the farm 1. However ANN performed well both with 70:30 ratio of data set as well as with cross validation , While SVR and MLR has shown no big difference in its performance on this dataset for estimation of peak discharge at outlet1. The SVR and ANN training time slightly increases in the cross validation.

Table 6.13 Overall performance of algorithms for peak discharge prediction at farm1 outlet

Algorithms performance	Training time	70:30 dataset	K=10 fold cross validation
ANN	satisfactory	Excellent	excellent
DT	very good	very good	very good
SVR	good	Good	good
MLR	Excellent	Satisfactory	Satisfactory

6.6 Time to Peak Predication at Farm1 outlet

Time to peak is the time on which the water flows at maximum on the outlet 1.

6.6.1 Case A Results for Time to Peak Predication at Farm1 outlet

The dataset is divided into randomly split of 70:30 ratio. 70 percent data is used for training and remaining 30 percent for testing .

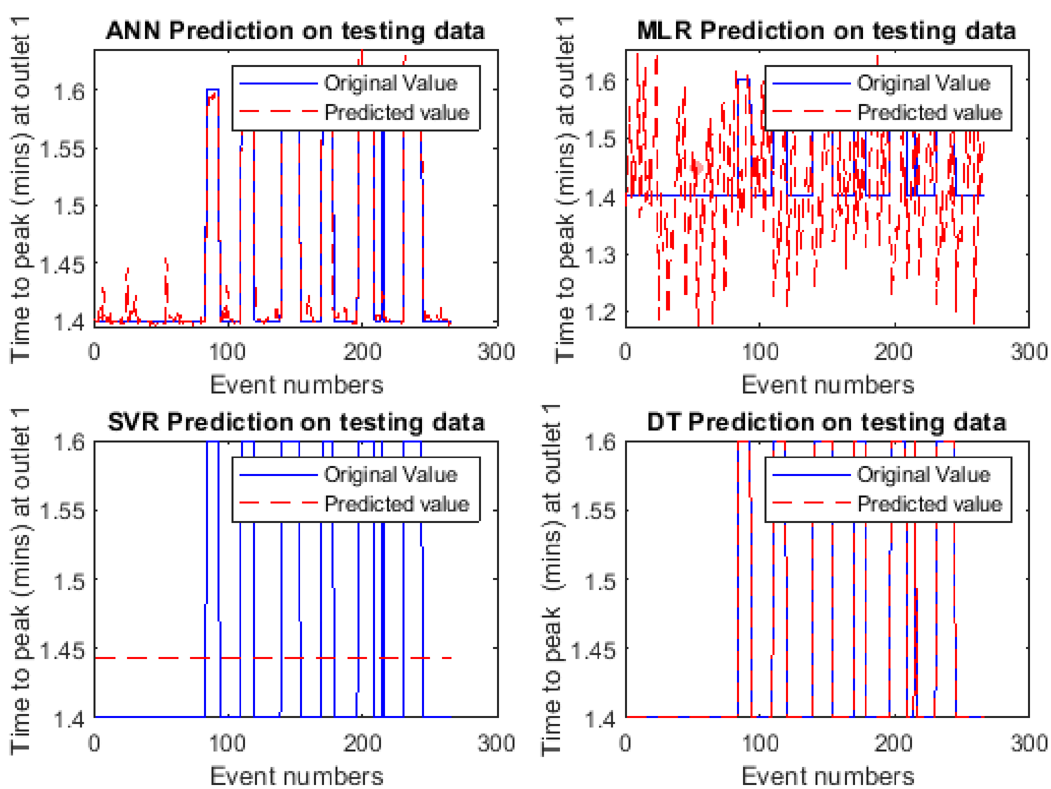


Figure 6.10 Case A Results for Time to Peak Predication at Farm1 outlet

Table 6.14 Case A Results for Time to Peak Predication at Farm1 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	0.4709	0.00034686	0.21133	0.95445	0.018624	0.012812	0.001162
MLR	0.0013681	0.022057	1.6852	0	0.14852	0.10457	-0.032119
SVR	0.17884	0.0078102	1.0028	0	0.088375	0.061205	-0.0085242
Decision Trees	0.08841	1.4939e-31	4.3857e-15	1	3.8651e-16	2.6611e-16	3.0521e-16

From the above Figure 6.10 and Table 6.14 it is clear that decision tree R-square is perfectly 1 and a very minimum MSE value for time to peak at outlet 1(farm 1) , the reason is that for most of inputs samples in the dataset for time to peak predication the output values are mostly the same which shows the nature of data is not very much different or varied . ANN closely followed the DT in terms of goodness of fit and on the basis of error metrics. How ever MLR and SVR shows a poor fit of the model and higher error metrics values than DT and ANN .

6.6.2 Case C Results for Time to Peak Predication at Farm1 outlet

Figure 6.11 shows Results of case C dataset. Case C: here k= 10 fold cross validation used on the dataset for training and testing .

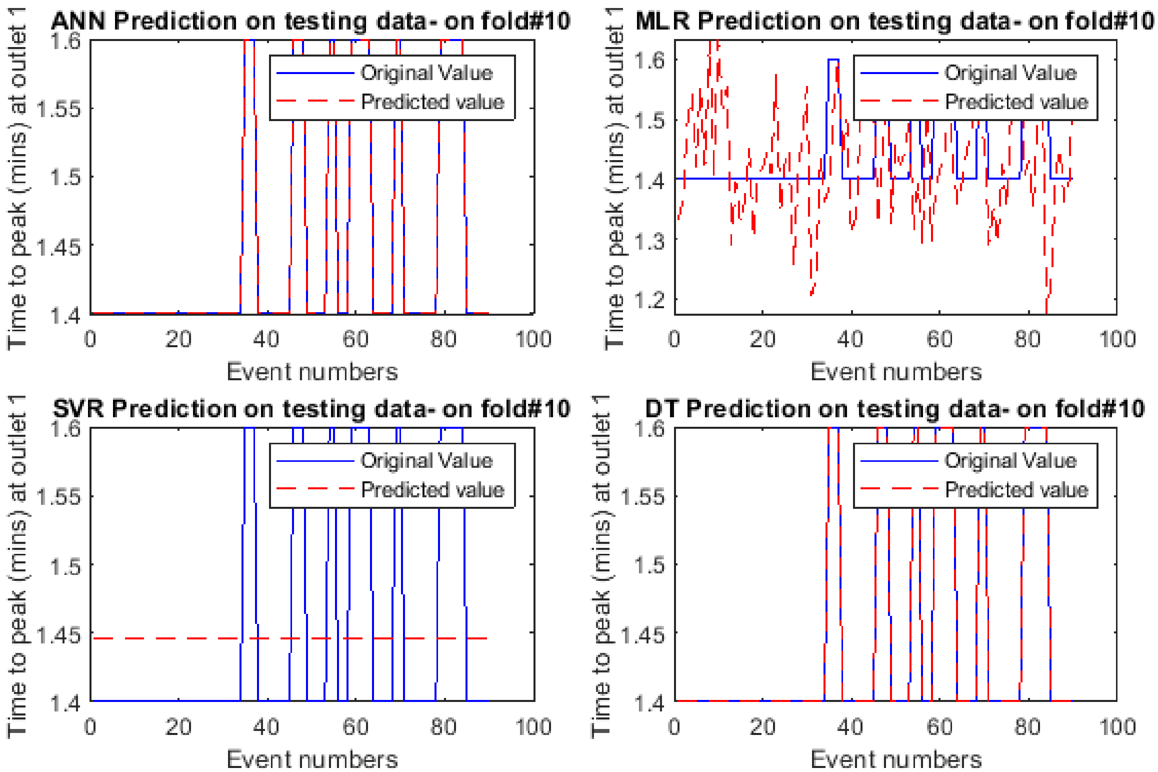


Figure 6.11 Case C Results for Time to Peak Predication at Farm1 outlet

Table 6.15 Algorithms performance for Case C for Time to Peak Predication at Farm1 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	1.4054	5.9297e-05	0.028661	0.99024	0.0024476	0.0016943	-0.00026301
MLR	0.0057715	0.020489	1.6834	0	0.14298	0.099462	-0.0089803
SVR	0.26189	0.007134	26.478	0	0.084461	0.058391	1.5695e-06
Decision Trees	0.091649	3.598e-28	2.1812e-13	1	1.8526e-14	1.2808e-14	1.2703e-14

In the cross validation results Figure 6.11 and Table 6.15 show that ANN improves its performance as well when used for time to peak value predication on outlet 1 and also minimize the mean square error , and other bench mark regression error metrics such as RMSE, RRMSE , NRMSE, MAE .While both the SVR and MLR does not show a good fit on data , their R-square = 0.

Table 6.16 overall performance of algorithms for time to peak prediction at farm1 outlet

Algorithms performance	Training time	70:30 dataset	K=10 fold cross validation
ANN	satisfactory	very good	very good
DT	very good	Excellent	excellent
SVR	good	Good	good
MLR	Excellent	Satisfactory	Satisfactory

6.7 Peak Discharge at Farm2 Outlet

This is peak discharge at farm 2 outlet which shows the highest discharge at the outlet 2.

6.7.1 Case A Results for Peak Discharge at Farm 2 Outlet

In case A the data is randomly divided into 70 /30 in which 70 percent data is used for training and 30 percent data for testing.

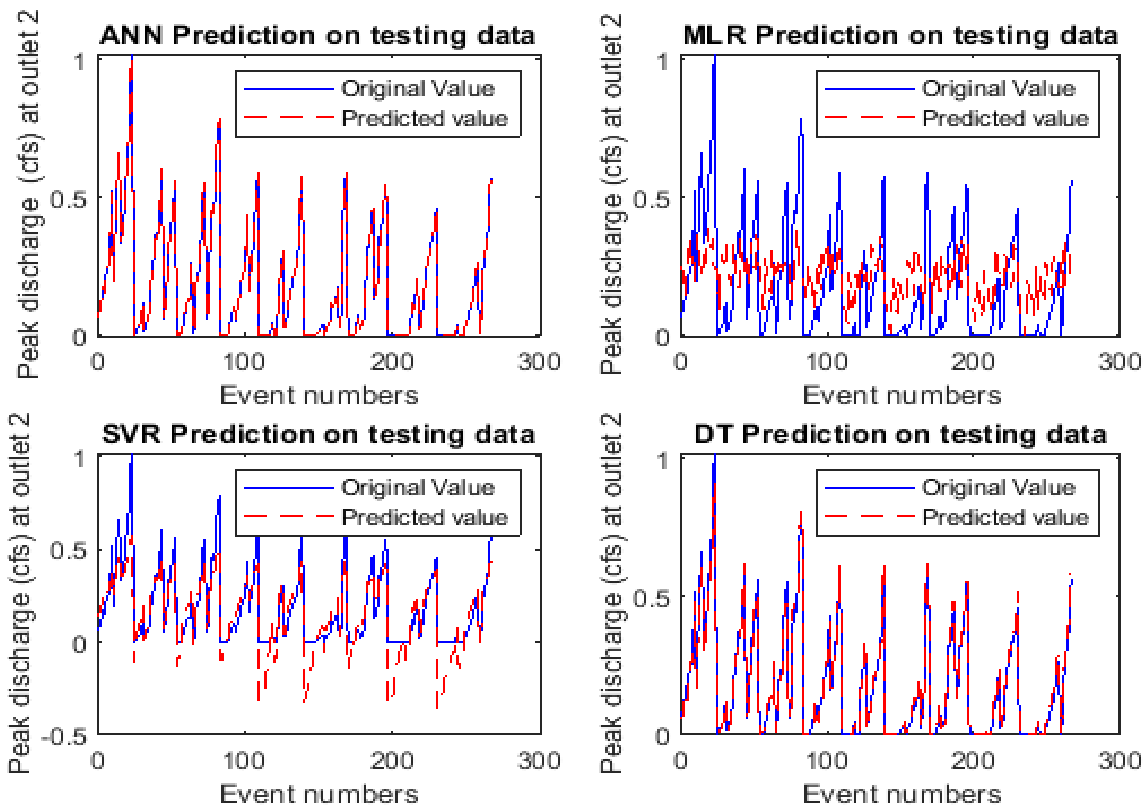


Figure 6.12 Case A results for Peak Discharge at Farm 2 Outlet

Table 6.17 Algorithms performance for Peak Discharge at Farm 2 Outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	11.223	2.3151e-07	0.0023255	0.99999	0.00048116	0.002616	-4.94e-05
MLR	0.31636	0.034746	0.90093	0	0.1864	0.85511	0.034009
SVR	4.1403	0.0134	0.55947	0.69468	0.11576	0.79076	-0.037593
Decision Trees	0.61808	0.00062951	0.12127	0.98497	0.02509	0.1351	0.0017351

From the visual analysis of the above Figure 6.12 and above Table 6.17 Algorithms performance for Peak Discharge at Farm 2 Outlet results it is worth to be noticed that ANN perform excellent to predict peak discharge (cfs) at the outlet 2, The R-square value is equal to 0.9999 almost perfect fit while the MSE value is very minimum , RMSE value is equal to 0.000. The decision tree after the ANN perform well on R-square value =0.984 while RMSE value =0.250. However MLR and SVR does not perform very well the R-square of SVR=0.694 and MLR =0 , both MLR RMSE= 0.186 and SVR RMSE = 0.115 and other bench mark performance error metrics are higher than ANN and DT .

6.7.2 Case C Results for Peak Discharge at Farm2 outlet

Case C : k fold cross validation , in which k=10 one fold is used for testing and the other is used for training and so on.

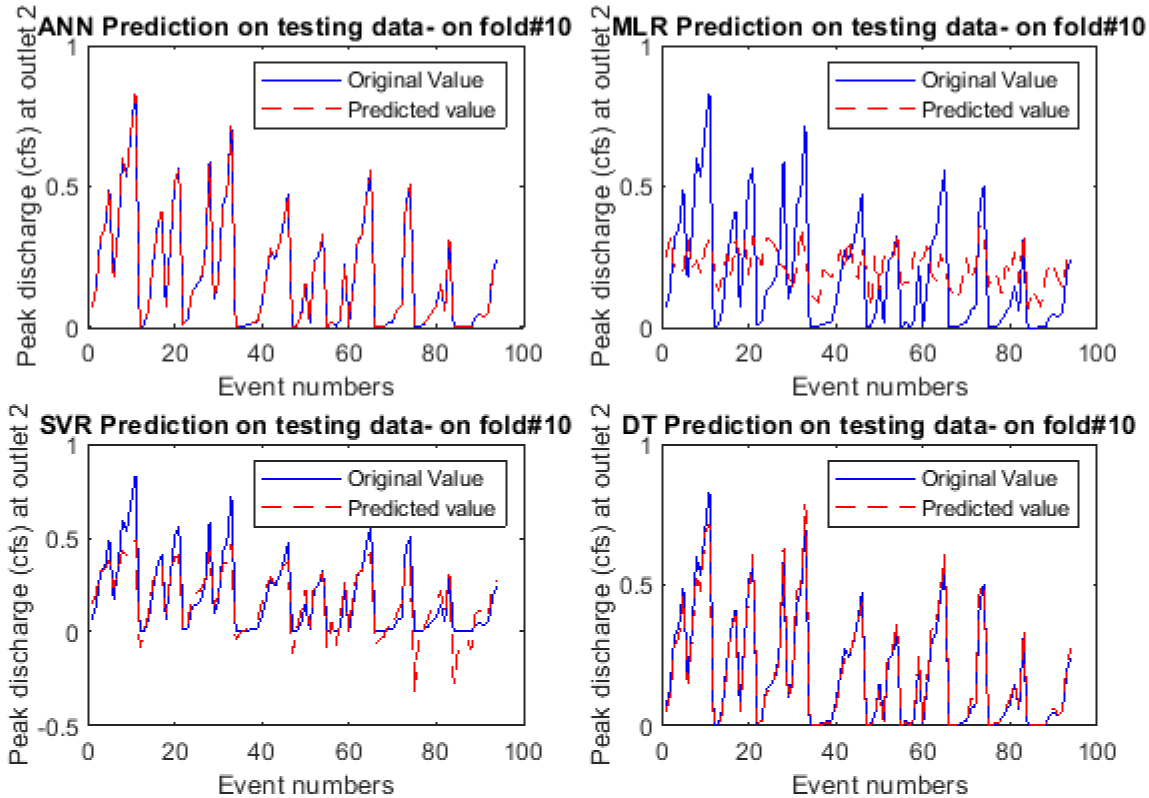


Figure 6.13 Case C Results for Peak discharge at Farm2 outlet

Table 6.18 Case C Results for Peak discharge at Farm2 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	2.5373	3.9061e-07	0.0026712	0.99999	0.0005763	0.0028378	-1.8411e-06
MLR	0.0019258	0.037706	0.8908	0	0.19335	0.90919	0.0093993
SVR	0.4985	0.012084	10.329	0.67148	0.10916	0.6387	-0.033122
Decision Trees	0.080381	0.00065557	0.11813	0.98558	0.025543	0.12626	-0.00072414

The k= 10 fold cross validation for peak discharge (cfs) at the outlet 2 shows from the above Figure 6.13 and Table 6.18 that ANN is excellent again in terms of goodness of fit (r-square =0.9999) and the error metrics such as MSE , NRMSE ,RMSE , RRMSE and MAE is minimum value is equal to zero. Then followed by DT whose R –square value is =0.985 while MSE =0.0006 . The SVR and MLR does not show goodness of fit and error metrics MSE, RRMSE , NRMSE and MAE values are higher than ANN and DT .

Table 6.19 Overall performance of algorithms at farm2 outlet for peak discharge

Algorithms performance	Training time	70:30 dataset	K=10 fold cross validation
ANN	satisfactory	excellent	Excellent
DT	very good	very good	very good
SVR	good	good	Good
MLR	Excellent	satisfactory	Satisfactory

6.8 Time to Peak Prediction at Farm 2 outlet

6.8.1 Class A Results Time to Peak Prediction at Farm 2 outlet

Case A the data is randomly partitioned into 70/30 , 70 percent data used for training and 30 percent for testing .

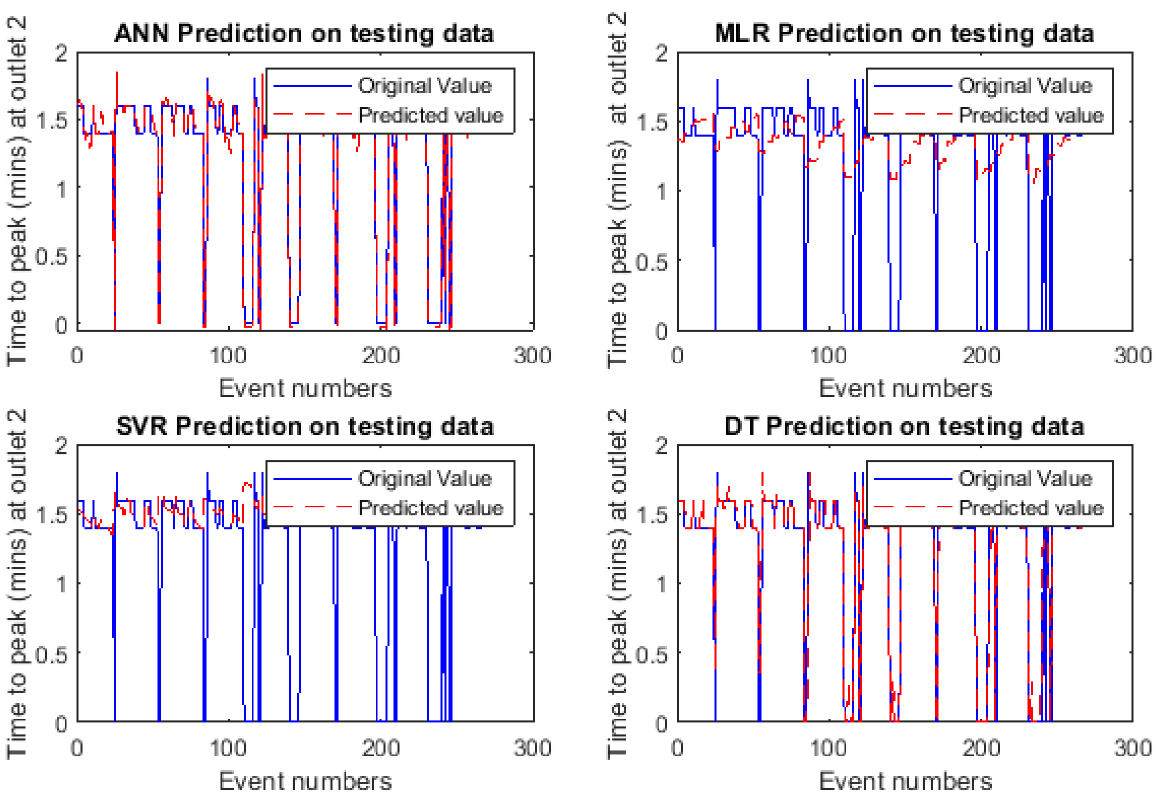


Figure 6.14 Case A results for Time to Peak Prediction at Farm 2 outlet

Table 6.20 Algorithms performance for Time to Peak Prediction at Farm 2 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	0.29324	0.0030744	0.094009	0.99117	0.055447	0.04391	-0.0039135
MLR	0.0012178	0.28052	0.89799	0	0.52964	0.39284	0.081571
SVR	0.20866	0.48929	1.186	0	0.69949	0.45665	0.26511
Decision Trees	0.02237	0.047795	0.37067	0.8445	0.21862	0.17046	0.015898

From the above Figure 6.14 and Table 6.20 it is clear that ANN predication and fit on the dataset is excellent followed by DT while SVR and MLR predication and fit is poor. Above graphs show that ANN and DT the blue original line clearly overlaps by dotted red predicated line. However the blue line (original value) does not overlaps by dotted red line (predicated value). Looking to the above table , the ANN predication and fit is excellent R-square value is =0.991 followed by DT R-square value =0.844 while MLR and SVR both R –square is equal to zero shows a poor fit of data. The regression error metric for ANN that is RMSE =0.055 and DT that is RMSE=0.218 is lower than other machine learning algothims MLR and SVR .

6.8.2 Case C Results for Time to Peak Prediction at farm2 outlet

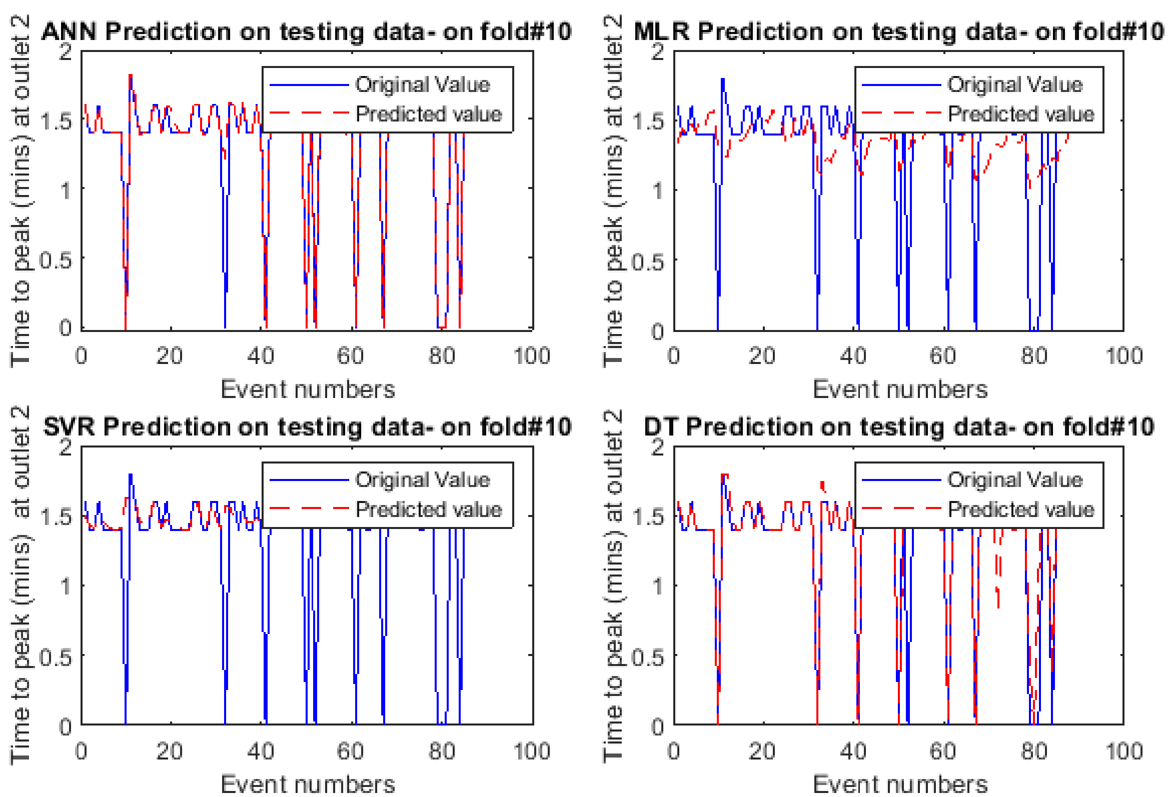


Figure 6.15 Case C Results for Time to Peak Prediction at farm2 outlet

Table 6.21 Algorithms performance for Case C for Time to Peak Prediction at farm2 outlet

	training_time	MSE	NRMS	R2	RMSE	RRMSE	MAE
ANN	0.89244	0.0091032	0.14639	0.96581	0.076594	0.057378	0.0014645
MLR	0.00041345	0.22511	0.90068	0	0.47437	0.35391	0.010941
SVR	0.26493	0.35046	4.2678	0	0.59181	0.39274	0.17719
Decision Trees	0.070032	0.048784	0.39472	0.80226	0.2076	0.1557	0.0021326

The above Figure 6.15 Case C Results for Time to Peak Prediction at farm2 outlet Figure 6.15 shows that ANN and DT blue line is overlapped by red dotted line shows a goodness of fit but it does not overlapped for SVR and MLR . then rendering the above Table 6.21 the R-square of ANN and DT is equal to 0.965 and 0.802 respectively. However both MLR and SVR R-square is equal to zero . The MSE of ANN , DT are equal to 0.009 and 0.04 while MLR and SVR are equal to 0.225 and 0.350. The RMSE of ANN, DT , SVR and MLR are equal to 0.076, 0.474,0.591 and 0.207. The regression error metrics shows that ANN and DT are good as compare to MLR and SVR. However the Training time of ANN and SVR are higher as compared decision tree and MLR.

Table 6.22 Overall performance of algorithms for time to peak prediction at farm2 outlet

Algorithms performance	Training time	70:30 dataset	K=10 fold cross validation
ANN	satisfactory	Excellent	Excellent
DT	very good	very good	very good
SVR	good	Satisfactory	Satisfactory
MLR	Excellent	Good	Good

6.9 Conclusion

This work is focused on optimal Reservoir and back runoff channels based two farms irrigation discharge based system. The aim is that how different hydrographs based on total discharge , total time , peak discharge and time to peak are to be learned and predicated on the farm1 outlet and farm 2 outlet when farm1 and farm 2 have different field and environment conditions such as temperature , soil moisture , crop stage , irrigation /precipitation depth related to different scenarios setup and samples generation . Obviously the concept is new and very little work has been done so far there is no exact real world data found initially to train and test the Reservoir based two farms irrigation system. Therefore for this purpose NRCS based simulator is picked up to generate the samples for different soil moisture , crop stages and irrigation / precipitation depth to represent as much as possible the real world scenario setups. A detail scripting has been done in NRCS simulator to acquire all possible samples for total discharge , peak discharge and time to peak as an output parameters . Initially the training and testing is carried out on a one watershed based one farm system on the NRCS to learn and predict peak discharge and time to peak so that the predication is possible to be carried out, however as mentioned earlier that the NRCS simulator uses a 24 hour duration based rainfall distribution and a unit hydrograph which is specific to the behaviour of a particular watershed and does not represent the actual farms and their unit hydrographs as there is no real world data. For reservoir based two farms irrigation system an one hour duration based irrigation distribution is used in a specific range to best fit irrigation based two farms system and a separate desktop based graphical user interface (GUI) has been developed for Reservoir and back runoff channels based two farm irrigation discharge system . then the total discharge as an output is learned and predicated for farm 1 and farm 2 outlet , here the inputs to the two farms irrigation system are crop stages (1 , 2 , 3) , soil moisture conditions (dry, average, wet and extremely wet) and irrigation / precipitation depth in a specific range 1 mm to 100mm , these all information is extracted from the TR 55 documentation urban hydrology for small watersheds and already existing equations for the soil moisture for dry , wet and average conditions in the literature and used in the NRCS simulator. The datasets are collected for peak discharges and time to peak values and total discharges at farm 1 and farm 2 outlets for the inputs variables which are irrigation depth, curve numbers for farm1 and farm 2 , soil moistures and crop stages at farm1 and farm2 acquired from sensors . it is worth to be noticed that sophisticated machine learning algorithms which are artificial neural network (LMA) , decision trees (regression trees) , support vector regression (linear and least square) and multiple linear regression are used in Matlab tool to learn and predict peak discharge , time to peak and total discharge . From the results which are available in graphical and tabulations form it is very much clear that artificial neural network work excellently well for most of the predication results for discharge , time to peak and peak discharge , however it's training time is always very much high as compared with other algorithms , the decision tree after ANN work very good on the bench mark regression error metrics , The SVR (LS) are put on the scale of good while MLR satisfactory because their R-square values are almost far away from 1 and their mean square are higher than artificial neural network and decision tree. In some cases the SVR and MLR has shown worse results . overall the non linear machine learning models such as artificial neural network and decision tree are very accurate and shows best fit of original and predicated data for different cases a (70:30) , case b (30:70) and case c (k= 10 fold cross validation). But the down side of ANN is its higher execution or training time while for decision tree its accuracy is lower than artificial neural network in most of the cases . In the future its best idea to use hybrid modelling of the various machine learning algorithms to improve its accuracy and training time . The IOT based mobile app has also been developed in the android studio for the optimal reservoir –back runoff channels based two farms irrigation discharge system for both discharge estimation , peak discharge predication, time to peak predication and the machine learning tabulated results , bar graphs and hydrographs for different scenario to be shown to the end user that could be farm owner ,hydrologist ,machine learning expert. The user has to enter only the current precipitation depth /irrigation depth value manually or from a meteorological site and IOT based mobile app will rendered the data for soil moisture condition , crop stage and temperature from these sensor in the field and display the results predication of NRCS, ANN , DT, SVR and MLR for both farm 1 and farm 2 outlet discharge , peak discharge values and the hydrographs showing peak discharge and time to peak . The actual data from the sensors from two different fields are rendered from the arduino based sensors for soil moisture conditions dry , average, wet and extremely wet (percentage) , temperature (Celsius) and crop stages 1 (fallow less than 0% surface cover), 2 (less than 50% surface cover) and 3 (greater than 75 % surface cover small grains) .The optimal reservoir and back runoff channel based two farms irrigation total discharge or peak discharge predication system helps to efficiently utilize the water waste and also saves the water waste by utilizing back runoff channels on both farm1 and farm2 to fill the reservoir storage capacity from the extra surplus amount of water either generated from the continuous rainfall on both of the farms or from an excessive irrigation .The total discharge , peak discharge and time to peak predications and hydrographs generation at field 1 and field 2 provide the information to let know the field 2 that how much amount of total water (discharge) , what is its peak discharge and time to peak discharge to be better utilize in the field 2 precision irrigation system and is sort of an early alarming system to field 2 , similarly the river basin knows that how much total discharge from field 2 is expected what is its peak discharge and time to peak and is sort of an early warning system to the river basin . Thus the Artificial intelligence specifically machine learning and IOT integration plays a key role in order to save water waste and energy.

Regarding the sensitivity analysis tells about the uncertainty of the output of an empirical or machine learning model which can be divided and allocated to different sources of uncertainty in its inputs. It shows how dependent an output variable is on the input variable. This is already been done on previous report on previous data set collected for runoff volume. the results are only highlighted as follow:

- When the single input irrigation depth is given to the model of runoff volume , the trend shows the linear relationship however Mean square error value is high and co-efficient of determination R^2 value is not 1 but 0.It shows irregular fit of predicated line and not well synchronize with the original line.
- When Curve number is given to the model of runoff volume , the same situation like irrigation depth as above.
- However when the irrigation depth and curve number is given as inputs to the model , the model shows an interesting results and drastically improves its results both in terms of minimum Mean square error and R^2. And original line clearly overlapped by predicated line.

It could be concluded for total discharge model as well , if the same input parameters irrigation depth and curve number are given to the model and in the similar pattern the data samples it will show the exact behavior and pattern. Other variables when dynamically selected such as temperature can also be utilized for sensitivity analysis in future.

Chapter 7: Future Work

In this report the following directions are emphasised on which the future work can be carried out.

7.1 Simulating real world data

a) Bridging the gap between real data and simulated data by providing highly realistic dataset for evaluating the reservoir- two farms irrigation total discharge model system. It will be tried to develop proper farm unit hydrograph on basis of different parameters of farms.
b) In the future work the temperature values will be taken dynamically and its sensitivity analysis will be carried out that how its impact the model will be seen.
c) For the crop stages the data will be tried to acquire from the real world through an image capturing device and an image processing will be carried out to extract useful features from those acquired images for fallow land stage1, small grains with 50 % surface cover stage 2 and small grains with 75% surface cover.

7.2 Dynamic modeling of farms

In this report, the two farms decision-based system used in Chapter 4 have fixed area of 400 m^2 area. The current model can be modified for dynamic allocation of values for an area and time of concentration and other important variables such as maximum length / distance and elevation difference to develop model to predict the runoff/ discharge.

.

7.3 Multiple farms system architecture (approach 1)

One of extension of current work is to extend it to multiple farms. As shown in figure 7.1 there are multiple farms system, farm1, farm2 and farm3 with the inflows and outflow path from main reservoir, each farm could have different length and width. The whole catchment is divided into multiple farms. The multiple farms have different soil moisture conditions such as dry, average and wet while these farms will have heterogeneous crop stages such as fallow land, small grains etc.

There would be different sensors on the farms for an integrated network that can acquire data for agricultural and environmental parameters and the similar structure already mentioned in this thesis in brief. The use of gateway , server, decision system for inflow and outflow of water from the reservoir and farms as well as predication of discharges.

- Optimal distribution of water to the farms based on its soil moisture, crop stage and temperature requirement.
- Back runoff channel based system save water waste from the irrigation or rainfall as the water can be diverted to the main reservoir.
- Multiple farms hydrographs predication for total discharge, peak discharge and time to peak.

- An early alarming system for each farm outlet and An early warning system for the river basin.
- The support for end user to know the current status of the farms from mobile application as well on desktop based application .

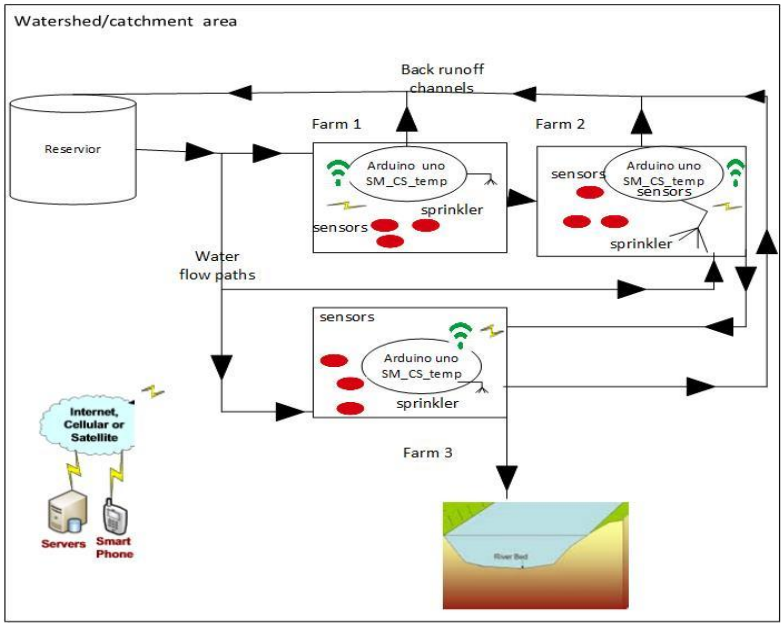


Figure 7.1 Multiple farms precision irrigation system

7.4 Ensemble Modeling:

The ensemble methods could be utilized to develop multiple learning algorithms models to give a better predication accuracy than the individual machine learning algorithm. Ensemble methods are meta-algorithms that combine several machine learning algorithms to decrease variance (bagging), bias (booting) or improve predications (stacking).In this report, ANN predicative accuracy is good than regression tree, however its training time is higher, while regression tree predicative accuracy is lower than ANN though its training time is lower than ANN. SVR and MLR does not show not very good accuracy. These models can be improved by bagging and boosting techniques or through stacking and blending techniques.

7.5 Zone management on a farm (approach 2)

- The second approach is zone management on an individual farm. The problem with the existing system is that both farm1 and farm2 have homogeneous crop stages and similar soil moisture condition on the whole farms. There could be the possibility of heterogeneous soil moisture and crop stages on the whole farms.
- The considered deployed precision irrigation systems are sprinklers such as linear move and center pivothowever it can waste water upto 25 % as an application

efficiency . The water pours from sprinkler generates runoff and transfer the runoff to the reservoir through electric pumps uses high energy and is costly.

- **The solution to improve the existing system**
  - One of the solution is to divide the onsite farm into multiple regions based on their soil moisture and crop stage conditions either from the sensors cluster formation dynamically or fixed sizes regions on the farm and then use variable rate irrigation , VRI brings irrigation efficiency and water productivity. Then estimate the runoff of each region and reuse the accumulative runoff.
  - The best irrigation system to be deployed would be drip irrigation such as surface drip or subsurface drip irrigation . The application efficiency of subsurface drip irrigation is 95 %.
  - The subsurface system also controls the runoff and reduces it . The subsurface utilizes the rain water harvesting as well. 1) one method is to collect rain water in tank from a farm house located near a farm 2) other method is to create a pond near a farm and collect rain water and then use for irrigation as subsurface drip , however this collection of rain water is location specific , so the rain pours on the other regions nearby will be wasted and not reuse. 3) In this approach the whole farm will be taken as rain water collection location , the subsurface drip will have laterals pipes and two pipes , one is influx from a reservoir and the other is outflux from the farm system . the lateral pipes takes the rain water from soil , keep the soil away and maintain farm saturation to desired level and if surplus that water is either kept in the outflux main pipe or drain away. The water table is 3 to 5 feet below the main pipe. Well managed subsurface drip irrigation system does not drain all the water from the soil, it maintains the water table constant. In this approach the rainfall surface water runoff is saved and controlled on the whole farm as farm save more quantity of rainfall water. In contrast on the same farm if this well managed water management system does not exit then once the farm is saturated it cannot absorb water anymore and the water will surface runoff on the entire farm causing the erosion of the surface soil and nutrients loss to the ditch or stream. In the well managed subsurface drip irrigation system there will be the use of inline water control structure in the main outflow pipe and is installed near the outlet. The control structure will have stoplogs which can be added or removed if the conditions are dry or wet. If dry conditions is anticipated for longer period of time the stop logs will maintain a level and water will be present in the pipe to fulfil the needs of plants else the stoplogs will be removed and water will be drain out of the field outlet. The structure is shown in figure 7.2.

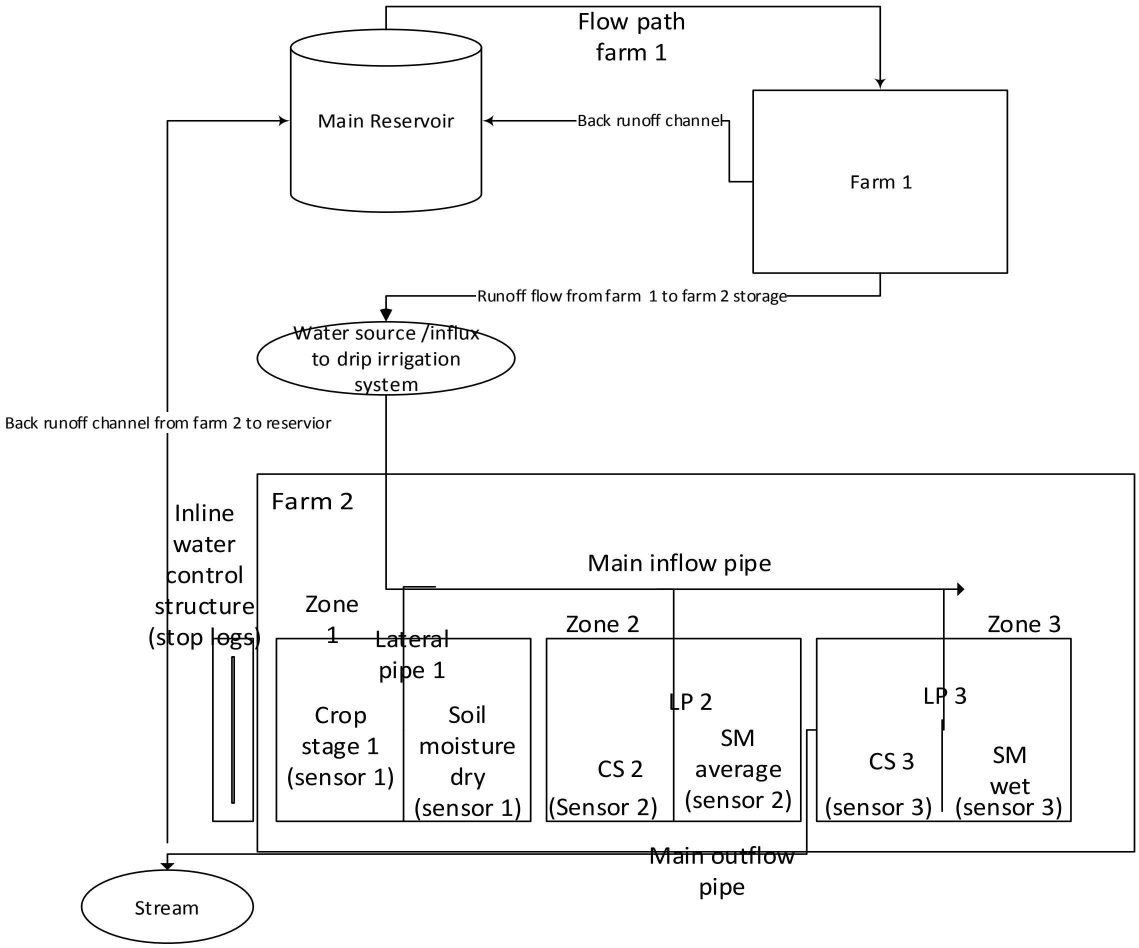


Figure 7.2 Heterogeneous irrigation system on farm 2

Bibliography

[1] G. Rasul and B. Sharma, "Water, Food, and Energy Nexus in South Asia: Implications for Adaption to climate change," in *Handbook of Climate Change Adaptation.*, Springer, Berlin, Heidelberg, 2015, pp. 1329-1350.

[2] P. Hall, N. Harris and L. Ang, "Water, Energy and Food Security: integrated technology solutions for sustainability," in *IEEE Conference on Technologies for Sustainability (SusTech2014)*, Portland, 2014.

[3] H. Turral, J. Burke and J. Faurès, "Climate change, water and food security," FAO, Rome, 2011.

[4] Food and Agriculture organization of Unitied Nations, "AQUASTAT - FAO's Global Information System on Water and Agriculture," FAO, 2016. [Online]. Available: http://www.fao.org/nr/water/aquastat/water_use/index.stm. [Accessed 29 04 2020].

[5] Food and agriculture organization of the unitied nations, "http://faostat.fao.org," FAO, 2010. [Online]. Available: http://faostat.fao.org. [Accessed 29 04 2020].

[6] G. Vellidis, V. Liako, W. Porter, M. Tucker and X. Liang, "A Dynamic Variable Rate Irrigation Control System," in *13th International Conference on Precision Agriculture*, St. Louis, Missouri, USA, 2016.

[7] C. Sanchez, D. Zerihun, T. Strelkoff, A. Clemmens and K. Farrell-Poe, "Development of management guidelines for efficient irrigation of basins on sandy soils," *Applied Engineering in Agriculture,* vol. XXIV, no. 2, pp. 215-224, March 2008.

[8] H. Zia, N. Harris, M. Rivers, N. Coles and R. Yates, "how could sensor networks helps with agricultural water managment issues? Optimizing irrigation scheduling through networked soil- mositure sensors," in *Sensor Application Symposium (SAS)*, Zadar ,Croatia, 2015.

[9] A. Daccache, J. Knox, A. Daneshkhah and T. Hess, "Implementing precision irrigation in a humid climate – Recent experiences and on-going challenges," *Agricultural Water Management,* vol. 147, no. 1, pp. 135-143, January 2015.

[10] E. AL-KARADSHEH, E. SOURELL and R. KRAUSE, "Precision Irrigation: New strategy irrigation water management," in *Conference on International Agricultural Research for*, Witzenhausen, 2002.

[11] O. Adeyemi, I. Grove, S. Peets and T. Norton, "Advanced Monitoring and Management Systems for Improving Sustainability in Precision Irrigation," *Sustainability,* vol. IX, no. 3, p. 353, 2017.

[12] UN water thematic intiative, "coping with water scarcity : A strategic issue and priority for system-wide action," IATP, Washington, D.C, 2006.

[13] A. Blum, "Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress," *Field Crops Research,* vol. CXII, no. 2-3, pp. 119-123, june 2009.

[14] G. Marks, "A method to save water and energy while increasing crop yield, a targeted approach for california agriculture," in *Precision irrigation*, 2010, p. 17.

[15] S. Irmak, L. Odhiambo, W. Kranz and D. Eisenhauer, "Irrigation Efficiency and Uniformity, and Crop Water Use Efficiency," *Biological Systems Engineering,* pp. 2-8, 2011.

[16] Y. Bradley, "Irrigation Scheduling Crucial During a Drought," 2012. [Online]. Available: https://www.nrcs.usda.gov/wps/portal/nrcs/detail/ks/newsroom/?cid=nrcs142p2_033543. [Accessed 07 March 2019].

[17] N. Shah and I. Das, "Precision Irrigation: Sensor Network Based Irrigation," in *Problems, Perspectives and Challenges of Agricultural Water Management*, London, Intech Open, 2012, pp. 218-232.

[18] h. Zia and N. Harris, *Enabling proactive agricultural drainage reuse for improved water quality through collaborative networks and low complexity data-driven modeling,* 2015.

[19] J. Gómez, A. Ben-Gal, j. Alarcón and G. De Lannoy, "SHui, an EU-Chinese cooperative project to optimize soil and water management in agricultural areas in the XXI century," *International Soil and Water Conservation Research,* vol. 8, no. 1, pp. 1-14, 2020.

[20] A. Malik, A. Kumar, S. Kim, M. Kashani, V. Karimi, A. Sharafati, M. Ghorbani, N. Al-Ansari, S. . Salih, Z. Yaseen and K. Chau, "Modeling monthly pan evaporation process over the Indian central Himalayas: application of multiple learning artificial intelligence model)," *Engineering Applications of Computational Fluid Mechanics,* vol. 14, no. 1, 2020.

[21] D. Zhanga, J. Lin, Q. Penga and D. Wang, "Modeling and simulating of reservoir operation using the artificial neural network, support vector regression, deep learning algorithm," *Journal of Hydrology,* vol. 565, pp. 720-736, 2018.

[22] R. Chanklan, N. Kaoungku, K. Suksut, K. Kerdprasop and N. Kerdprasop, "Runoff Prediction with a Combined Artificial Neural Network and Support Vector Regression," *International Journal of Machine Learning and Computing,* vol. 8, no. 1, 2018.

[23] X. Li, J. Sha and Z. Wang, "Comparison of daily streamflow forecasts using extreme learning machines and the random forest method," *Hydrological Sciences Journal,* pp.

1857-1866, 2018.

[24] H. Lee, "Radionuclides: Their Biogeochemical Cycles and the Impacts on the Biosphere," in *Encyclopedia of Ecology*, Elsevier, 2008, pp. 2966-2975.

[25] J. Jagannatha, N. Polosky, A. Jagannath, F. Restuccia and T. Melodia, "Machine learning for wireless communications in the Internet of Things: A comprehensive survey," *Ad Hoc Networks*, vol. 39, 2019.

[26] O. Adeyemi, I. Grove, S. Peets and T. Norton, "Advanced Monitoring and Management Systems for Improving Sustainability in Precision Irrigation," *Sustainability*, vol. 9, no. 3, 2017.

[27] S. Shibusawa, "PRECISION FARMING APPROACHES FOR SMALL SCALE FARMS," *IFAC Proceedings Volumes*, vol. 2001.

[28] M. Chappell, S. Dove, M. VanIersel, P. Thomas and J. Ruter, "Implementation of Wireless Sensor Networks for Irrigation Control in Three Container Nurseries," *HortTechnology*, vol. 23, no. 6, pp. 747-753, 2013.

[29] J. Lea-Cox, W. Bauerle, M. VanIersel, G. Kantor, T. Bauerle, E. Lichtenberg, D. King and L. Crawford, "Advancing Wireless Sensor Networks for Irrigation Management of Ornamental Crops: An Overview," *HortTechnology*, vol. 23, no. 6, pp. 717-714, 2013.

[30] E. Abioye, M. Abidina, A. MahmudS, S. Buyamina, M. Ishaka, M. Otuoze, P. Onotu and M. Ramlid, "A review on monitoring and advanced control strategies for precision irrigation)," *Computers and Electronics in Agriculture*, vol. 173, 2020.

[31] F. Ferrández-Pastor, J. García-Chamizo, M. Nieto-Hidalgo and J. Mora-Martínez, "Precision agriculture design method using a distributed computing architecture on internet of things context," *Sensors*, vol. 18, no. 6, pp. 1710-1731, 2018.

[32] K. Haseeb, I. Din and A. Mogren, "An Energy Efficient and Secure IoT-based WSN Framework: An application to Smart Agriculture," *Sensors*, vol. 20, no. 7, 2020.

[33] G. Prasanna, S. Parvatham and S. Krishna, "Web Based Automatic Irrigation System Using Raspberry Pi Processor on Embedded Linux," *International Journal & Magazine of Engineering, Technology, Management and Research*, vol. 3, no. 10, 2016.

[34] S. Rawal, "IOT based Smart Irrigation System," *International Journal of Computer Applications*, vol. 159, no. 8, pp. 17-11, 2017.

[35] T. Vellidis, C. .Perry, C. Kvien and C. Bednarz, "A real-time wireless smart sensor array for scheduling irrigation," *Computers and Electronics in Agriculture*, vol. 61, no. 1, pp. 44-50, 2008.

[36] ,. KumarB, P. Srivastava, R. Agrawal and V. Tiwari, "MICROCONTROLLER BASED AUTOMATIC PLANT IRRIGATION SYSTEM," *International Research Journal of Engineering and Technology (IRJET),* vol. 4, no. 5, 2017.

[37] N. Agrawal and S. Singhal, "Smart drip irrigation system using raspberry pi and arduino," in *International Conference on Computing, Communication & Automation*, Noida, 2015.

[38] M. Yousif, K. Ghafar, R. Zahari and T. Lim, "A rule-based smart automated fertilization and irrigation systems," in *Ninth International Conference on Graphic and Image Processing (ICGIP*, Qingdao, 2017.

[39] A. McCarthy, N. Hancock and S. Raine, "VARIwise: A general-purpose adaptive control simulation framework for spatially and temporally varied irrigation at sub-field scale," *Computers and Electronics in Agriculture,* vol. 70, no. 1, 2010.

[40] F. Karandish and J. Šimůnek, "A comparison of numerical and machine-learning modeling of soil water content with limited input data.," *Journal of Hydrology,* vol. 542, pp. 892-909, 2016.

[41] S. Sahoo, T. Russo, J. Elliott and I. Foster, "Machine learning algorithms for modeling groundwater level changes in agricultural regions of the US," *Water Resource Research,* 2017.

[42] H. Navarro-Hellín, J. Martinez-del-Rincon, R. Domingo, F. Soto and R. Torres, "A decision support system for managing irrigation in agriculture," *Computers and Electronics in Agriculture,* vol. 124, pp. 121-131, 2016.

[43] C. Stockle, M. Donatell and R. Nelson, "CropSyst, a cropping systems simulation model," *European Journal of Agronomy,* vol. XVIII, no. 3-4, pp. 289-307, January 2003.

[44] B. Keating, P. Carberry and M. Probert, "An overview of APSIM, a model designed for farming systems simulation," *European Journal of Agronomy,* vol. XVIII, no. 3-4, pp. 267-288, January 2003.

[45] P. Steduto, T. Steduto, D. Raes and E. Fereres, "AquaCrop – the FAO crop model to simulate yield response to water: I. Concepts and underlying principles," *Agronomy Journal,* vol. 101, no. 3, pp. 426-437, 2009.

[46] B. e Wita, SanderJanssenaF, K. D. KraalingenaI, I. Supita, R. der Wijngaarta and K. Diepena, "25 years of the WOFOST cropping systems model," *Agricultural Systems,* vol. 168, pp. 154-167, 2019.

[47] J. Šimůnek and V. Genuchten, "Modeling nonequilibrium flow and transport with HYDRUS," *Vadose Zone Journal,* vol. 7, pp. 782-797, 2008.

[48] "https://swat.tamu.edu/docs/," [Online]. [Accessed 1 August 2020].

[49] K. Van Oost, G. Govers and P. Desmet, "Evaluating the effects of changes in landscape structure on soil erosion by water and tillage," *Landscape ecology,* vol. 15, p. 577–589, 2000.

[50] L. Baoyuan, Z. Keli and X. Yun, "An empirical soil loss equation," in *process of soil erosion and its environment effect (Vol. II), 12th international soil conservation organization conference*, Beijing , 2002.

[51] G. Chen, Z. Zhang, Q. Guo, X. Wang and Q. Wen, ". Quantitative assessment ofsoil erosion based on CSLE and the 2010 national soil erosion survey at regionalscale in Yunnan Province of China.Sustainability,," *Sustainability,* vol. 11, no. 12, 2019.

[52] T. Nonglait and B. Tiwari, "Application of SCS-CN method for estimation of runoff in a humid microwatershed," *International journal of current agricultural sciences,* vol. VI, no. 10, pp. 121-127, october 2016.

[53] Hydraulic Engineer, "NATIONAL ENGINEERING HANDBOOK," in *HYDROLOGY*, Washington DC, USDA NRCS, 1972, pp. 4-102.

[54] R. Raju, G. Raju and M. Rajasekhar, "Estimation of rainfall runoff using SCS-CN Method wih RS and GIS Techniques for Mandavi Basin in YSR Kadapa District of Andhra Pradesh,India," *Hydrospatial Analysis,* vol. II, no. 1, pp. 1-15, Feburary 2018.

[55] T. Devi and K. , Surface Runoff Depth by SCS Curve Number Method Integrated with Satellite Image and GIS Techniques, S. V. K. S. B. R. (. Sarma A., Ed., springer, 2016.

[56] M. Ebrahimian, A. Nuruddin, M. Soom and A. Sood, "Application of NRCS-curve number method for runoff estimation in a mountainous watershed," *Caspin journal of environmental sciences (CJES),* vol. X, no. 1, pp. 103-114, Janurary 2012.

[57] D. Ramakrishnan, A. Bandyopadhyay and K. Kusuma, "SCS-CN and GIS-based approach for identifying potential water harvesting sites in the Kali Watershed, Mahi River Basin, India," *Journal of Earth System Science,* vol. CXVIII, no. 4, pp. 355-368, August 2009.

[58] Y. Cho and B. Angel, "Spatially distributed long-term hydrologic simulation using a continuous SCS CN method-based hybrid hydrologic model," *Hydrological Processes,* vol. XXXII, no. 7, pp. 904-922, Feburary 2018.

[59] A. Mosavi, P. Ozturk and K. Chau, "Flood Prediction Using Machine Learning Models: Literature Review," *Water,* vol. 10, no. 11, 2018.

[60] P. Zhang and S. Ariaratnam, "Life cycle cost savings analysis on traditional drainage systems from low impact development strategies," *Frontiers of Engineering Management* ,

pp. 1-10, 2020.

[61] P. Saha and K. Zeleke, "Rainfall-Runoff Modelling for Sustainable Water Resources Management: SWAT Model Review in Australia," in *Sustainability of Integrated Water Resources Management*, Springer, 2015, pp. 563-578.

[62] "Hydrologic Modelling System HEC HMS Technical Reference Manual," 2000.

[63] L. Silva, Runoff under sprinkler irrigation Affecting Factors and Control, New york: Nova Science, 2010.

[64] L. Silva, "Are basin and reservoir tillage effective techniques to reduce runoff under sprinkler irrigation in Mediterranean conditions," *Agricultural Water Management,* vol. 191, pp. 50-56, 2017.

[65] L. Silva, "The effect of spray head sprinklers with different deflector plates on irrigation uniformity, runoff and sediment yield in a Mediterranean soil," *Agric. Water Manage,* vol. 85, pp. 243-252, 2006.

[66] D. DeBoer, K. Stange and D. Beck, "Sprinkler and corn canopy effects on water application characteristics," *Irrig. Drain. Eng,* vol. 127, no. 5, pp. 272-276, 2001.

[67] J. Sitterson, C. Knightes, R. Parmar, K. Wolfe, M. Muche and B. Avant, "An Overview of Rainfall-Runoff," The U.S. Environmental Protection Agency (EPA) , 2017.

[68] W. Yang, D. Li, T. Sun and G. Nang Ni, "Saturation-excess and infiltration-excess runoff on green roofs," *Ecological Engineering,* vol. 74, pp. 327-336, 2015.

[69] R. S and S. S, "COMPARATIVE STUDY OF STREAM FLOW PREDICTION MODELS," *Internation joutnal of Life Science biology and Phama Research,* vol. 1, no. 2, 2012.

[70] A. Hayek, N. Tabaja, Z. Khraibani, S. Andalouss, J. Toufaily, M. Mrad, E. Garnie-Zarli and T. Hamieh, "Modeling of Runoff as a Function of Temperature and Precipitation: Application to the Litani River in Lebanon," *Journal of Earth Science & Climatic Change,* vol. 10, no. 9, 2019.

[71] L. Dar, "Rainfall-Runoff Modeling using Multiple Linear Regression Technique," *International Journal for Research in Applied Science & Engineering Technology (IJRASET) ,* 2017.

[72] A. Poul, Shourian.M and H. Ebrahimi, "A Comparative Study of MLR, KNN, ANN and ANFIS Models with Wavelet Transform in Monthly Stream Flow Prediction," *Water Resources Management* , vol. 33, pp. 2907-2923, 2019.

[73] S. Yirga, "Modelling reference evapotranspiration for Megecha catchment by multiple

linear regression," *Model. Earth Syst. Environ,* vol. 5, pp. 471-477, 2019.

[74] H. pGhobaria, A. Mattarab, M. El-Marazkyab, A. Dewidarab and M. Mattar, "Prediction of wind drift and evaporation losses from sprinkler irrigation using neural network and multiple regression techniques," *Agricultural Water Management,* vol. 195, no. 1, pp. 211-221, 2018.

[75] S. Hosseinia and N. Mahjourib, "Integrating Support Vector Regression and a geomorphologic Artificial Neural Network for daily rainfall-runoff modeling," *Applied Soft Computing,* vol. 38, pp. 329-345, 2016.

[76] M. Mohammed, R. Kolapalli, N. Golla and S. Maturi, "predicaton of rainfall using machine learning techniques," *International journal of Scienctific and Technology Research,* vol. 9, no. 1, 2020.

[77] F. Granata, R. Gargano and G. Marinis, "Support Vector Regression for Rainfall-Runoff Modeling in Urban Drainage: A Comparison with the EPA's Storm Water Management Model," *Water,* vol. 8, no. 69, 2016.

[78] W. Niu, Z. Feng, B. Feng, Y. Min, C. Cheng and J. Zhou, "Comparison of Multiple Linear Regression, Artificial Neural Network, Extreme Learning Machine, and Support Vector Machine in Deriving Operation Rule of Hydropower Reservoir," *Water,* vol. 11, no. 1, 2019.

[79] R. Sanchez, H. Navarro-Hellin, A. Guillamon-Frutos, R. San-Segundo and R.-A. 3. a. D.-M. , "A Decision Support System for Irrigation Management: Analysis and Implementation of Different Learning Techniques," *Water,* vol. 12, no. 2, 2020.

[80] P. Perea, C. Poyato, P. Montesinos and J. Díaz, "Prediction of irrigation event occurrence at farm level using optimal decision trees," *Computers and Electronics in Agriculture,* vol. 157, pp. 173-180, 2019.

[81] B. Bhattacharya and B. Solomatine, "Neural networks and M5 model trees in modeling water level-discharge relationship for an Indian river," in *ESANN'2003 proceedings - European Symposium on Artificial Neural Networks,* Bruges, 2003.

[82] D. SOLOMATINE and K. DULAL, "Model trees as an alternative to neural networks in rainfall—runoff modelling," *Hydrological Sciences Journal,* vol. XLVIII, no. 3, pp. 399-411, June 2003.

[83] A. Khan, Saadat and S. Linda, "Rainfall-Runoff Modelling using Data Driven and Statistical Methods," in *International Conference on Advances in Space Technologies ICAST,* Islamabad, Pakistan, 2006.

[84] H. ZIA, N. HARRIS and G. MERRETT, "Water quality monitoring , control and managment(WQMCM) framework using collaborative wireless sensor networks," in *11th*

International Conference on Hydroinformatics, New York, 2014.

[85] 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,* vol. 19, pp. 421-444, 2018.

[86] V. Kuzmanovski, "Integration of expert knowledge and predictive learning: Modelling water flows in agriculture," Jozef Stefan International Postgraduate School , Slovenia, 2012.

[87] A. Alipour, J. Yarahmadi and M. Mahdavi, "Comparative Study of M5 Model Tree and Artificial Neural Network in Estimating Reference Evapotranspiration Using MODIS Products," *Journal of climatology,* 2014.

[88] F. Chang, k. Hsu and L. Chang, Flood Forecasting Using Machine Learning Methods, 2019.

[89] S. Vyas, Y. Mathur, G. Sharma and V. Chandwan, "Rainfall-Runoff Modelling: Conventional regression and Artificial Neural Networks approach," in *2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE),* 2016.

[90] C. Juan, W. Genxu, M. Tianxu and S. Xiangyang, "ANN Model-Based Simulation of the Runoff Variation in Response to Climate Change on the Qinghai-Tibet Plateau," 2017.

[91] S. Vyas, Y. Mathur, G. Sharma and V. Chandwani, "Rainfall-Runoff Modelling: Conventional Regression and Artificial Neural Networks Approach," in *2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, Jaipur, India, 2016.

[92] A. Chakravarti, N. Joshi and H. Panjiar, "The comparison of observed and predicted runoff hydrograph reveals that the Artificial Neural Network (ANN) predicts the runoff data reasonably well in observed hydrograph.," *Indian Journal of Science and Technology,* vol. 8, no. 14, 2015.

[93] S. Prakash and S. Sahu, "Soil Moisture Prediction Using Shallow Neural Network," *International Journal of Advanced Research in Engineering and Technology,* vol. 11, no. 6, pp. 426-435, 220.

[94] Q. Noman, S. Samadianfard, S. Kheshtgar, S. Jarhan, S. Kisi and S. Shamshirband, "Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates," *Engineering Applications of Computational Fluid Mechanics,* vol. 13, no. 1, 2019.

[95] S. Ardabili, A. Mosavi, M. Dehghani and A. Varkonyi-Koczy2, "Deep learning and machine learning in hydrological processes climate change and earth".

[96] E. Apaydin, "Cross validation and resampling methods," in *Introduction to machine learning*, london, MIT press, 2010, pp. 486-488.

[97] M. Browne, "Cross-Validation Methods," *Journal of Mathematical Psychology,* vol. XLIV, no. 1, pp. 108-132, March 2000.

[98] stephanie, "statistics how to," November 2013. [Online]. Available: https://www.statisticshowto.datasciencecentral.com/mean-squared-error/. [Accessed 29 04 2020].

[99] "Wikipedia," january 2019. [Online]. Available: https://en.wikipedia.org/wiki/Mean_squared_error. [Accessed 29 04 2020].

[100] T. Cha and R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geoscientific model development,* vol. vii, no. 3, pp. 1247-1250, June 2014.

[101] M. Despotovic, V. Nedic, D. Despotovic and S. Cvetanovic, "Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation," *Renewable and Sustainable Energy Reviews,* vol. LVI, pp. 246-260, apirl 2016.

[102] stack exchange Inc, "Cross validated," Stack Exchange , january 2019. [Online]. Available: https://stats.stackexchange.com/questions/260615/what-is-the-difference-between-rrmse-and-rmsre. [Accessed 29 04 2020].

[103] J. Dufour, "Coefficients of determination," McGill university , canada, 2011.

[104] H. Zia, "Enabling Proactive Agricultural Drainage Reuse for Improved Water Quality through Collaborative Networks and Low-Complexity Data-Driven Modelling," university of southampton, SOUTHAMPTON, 2015.

[105] C. Willmott and K. Matsuura, "Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in Assessing Average Model Performance," *climate research,* vol. XXX, no. 1, pp. 79-82, December 2005.

[106] E. Kardakos, M. Alexiadis, S. Vagropoulos and A. Bakirtzis, "Application of time series and artificial neural network models in short-term forecasting of PV power generation," in *48th International Universities' Power Engineering Conference (UPEC)*, Dublin, Ireland, 2013.

[107] Dejan, "Arduino Wireless Communication – NRF24L01 Tutorial," Arduino, 1 March 2019. [Online]. Available: https://howtomechatronics.com/tutorials/arduino/arduino-wireless-communication-nrf24l01-tutorial/. [Accessed 1 Apirl 2020].

[108] Arduino, "NRF24L01 ANTENNA WIRELESS TRANSCEIVER MODULE," Arduino company, 1 March 2019. [Online]. Available: https://en-geniuses.com/product/nrf24l01-

antenna-wireless-transceiver-module/. [Accessed 1 Apirl 2020].

[109] "Arduino-Soil-Moisture-Sensor," [Online]. Available: https://www.instructables.com/id/Arduino-Soil-Moisture-Sensor/. [Accessed 1 March 2019].

[110] M. TAPAN, "ARDUINO TEMPERATURE SENSOR LM35," 18 August 2013. [Online]. Available: https://www.instructables.com/id/ARDUINO-TEMPERATURE-SENSOR-LM35/. [Accessed 1 March 2019].

Appendix

Arduino code

Arduino code 1

```
float temp = analogRead(A0);
  temp = temp * 0.48828125;
  Serial.print("Temperature: ");
  Serial.print(temp);
  Serial.println("C");
  delay(1000);
float moisture_percentage;
  int sensor_analog;
  sensor_analog = analogRead(sensor_pin);
  moisture_percentage = ( 100 - ( (sensor_analog/1023.00) * 100 )
  Serial.print("Moisture Percentage =");
  Serial.print(moisture_percentage);
  Serial.print("%\n\n");
  delay(1000);
char result[20];

    // connect to the server
   if (client.connect(server, 80))
{
  client.print("GET /SirProject/public/updateSensorStatus?sensor_i
  ///////////////Please Insert Arduino key bellow////////////
  client.print("1");
  client.print("&moisture_data");
  client.println(moisture_percentage);
  client.print("&temp_data=");
  client.println(temp);
Serial.println("\ncomplete\n");
```

Rest API code

```php
<?php

use Slim\Http\Request;
use Slim\Http\Response;

// Routes
// Routes
include "../include/DbConnection.php";

///////////////////////////UPDATE STATUS/////////////////////////////////////////////////////
$app->get( pattern: "/updateSensorStatus", function (Request $request, Response $response) {
    $dbConn = new DbConnection();
    $connection = $dbConn->connect();
    $sensor_id = $request->getParam( key: "sensor_id");
    $temp_sensor_data = $request->getParam( key: "temp_data");
    $moisture_sensor_data = $request->getParam( key: "moisture_data");
    $checkStatus=$connection->prepare( statement: "select sensor_status from iot_project_db.sensor
    $checkStatus->execute([$sensor_id]);
    $data=$checkStatus->fetch( fetch_style: PDO::FETCH_ASSOC);
    if($data["sensor_status"]=="true") {
        $stmt = $connection->prepare( statement: "UPDATE iot_project_db.sensor SET moisture_sensor_d

        if ($stmt->execute([$moisture_sensor_data,$temp_sensor_data,$sensor_id])) {

            return json_encode(array(
                "message" => "status update"
            ));

        } else {
            // echo "Data is not updated";
        }
    }
    else{
```

Algorithm codes and empirical equations

Multiple linear regression formula /code in Matlab

```
tic
mlr_coeff = regress(trn_tars, trn_data);
training_time_MLR= toc;
pred_sum= zeros(length(tst_data), 1);
for k=1:length(mlr_coeff)
    pred_sum = pred_sum + mlr_coeff(k)*tst_data(:, k);
end
pred_vals= pred_sum;
```

Artificial neural network hidden structure

```
x = trn_data';
t = trn_tars';

% Choose a Training Function
% For a list of all training functions type: help nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for challenging problems.
% 'trainscg' uses less memory. Suitable in low memory situations.
trainFcn = 'trainlm';  % Levenberg-Marquardt backpropagation.

% Create a Fitting Network
hiddenLayerSize = 10;
net = fitnet(hiddenLayerSize,trainFcn);

% Setup Division of Data for Training, Validation, Testing
net.divideParam.trainRatio = 70/100;
net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;
```

Regression tree formula /Hidden code Matlab

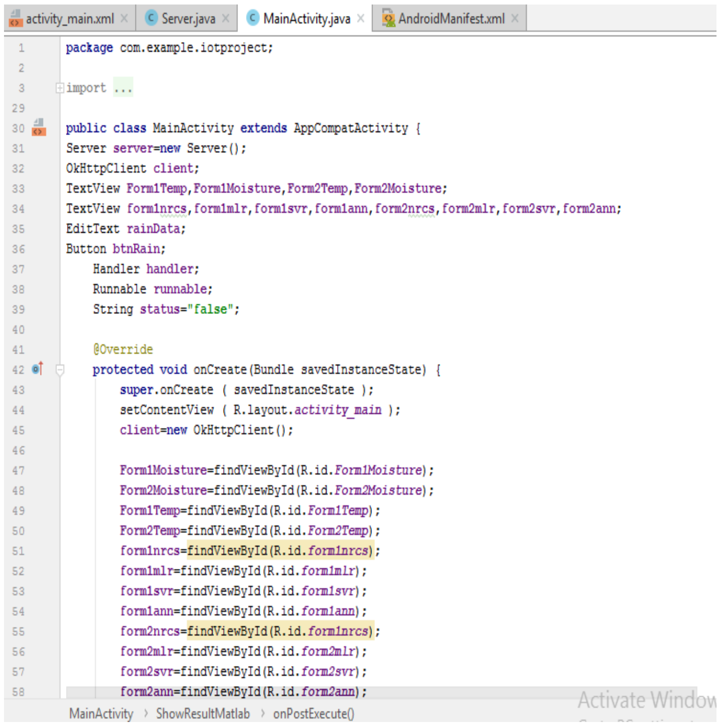
```
% Grow a regression tree using the entire data set.
tic
Mdl_decision_tree = fitrtree(trn_data, trn_tars);
training_time_DT= toc
pred_vals = predict(Mdl_decision_tree,tst_data);
```

Support vector regression formula /Hidden code Matlab

```
if size(trn_data, 2)==1
tbl_trn = table(trn_data(:, 1), trn_tars);
tbl_tst = table(tst_data(:, 1), tst_tars);
elseif size(trn_data, 2)==2
tbl_trn = table(trn_data(:, 1),trn_data(:, 2), trn_tars);
tbl_tst = table(tst_data(:, 1),tst_data(:, 2), tst_tars);
elseif size(trn_data, 2)==3
tbl_trn = table(trn_data(:, 1),trn_data(:, 2), trn_data(:, 3), trn_tars);
tbl_tst = table(tst_data(:, 1),tst_data(:, 2), tst_data(:, 3), tst_tars);
elseif size(trn_data, 2)==4
tbl_trn = table(trn_data(:, 1),trn_data(:, 2), trn_data(:, 3), trn_data(:, 4), trn_tars);
tbl_tst = table(tst_data(:, 1),tst_data(:, 2), tst_data(:, 3), tst_data(:, 4), tst_tars);
elseif size(trn_data, 2)==5
tbl_trn = table(trn_data(:, 1),trn_data(:, 2), trn_data(:, 3), trn_data(:, 4), trn_data(:, 5), trn_tars);
tbl_tst = table(tst_data(:, 1),tst_data(:, 2), tst_data(:, 3), tst_data(:, 4), tst_data(:, 5), tst_tars);
elseif size(trn_data,2)==8
tbl_trn = table(trn_data(:, 1),trn_data(:, 2), trn_data(:, 3), trn_data(:, 4), trn_data(:, 5), trn_data(:, 6

tbl_tst = table(tst_data(:, 1),tst_data(:, 2), tst_data(:, 3), tst_data(:, 4), tst_data(:, 5), tst_data(:, 6
end
tic
% Train a linear SVM regression model. Standardize the data.
Mdl_SVM = fitrsvm(tbl_trn,'trn_tars','Standardize',true); % Mdl is a RegressionSVM model.

training_time_SVR= toc;
% Predict responses for the test set
```

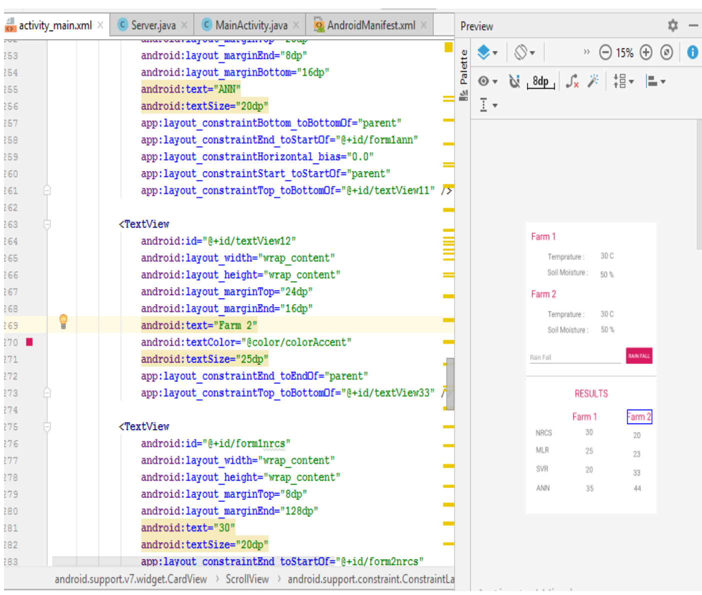
Android code

Android code 1



```java
package com.example.iotproject;

import ...

public class MainActivity extends AppCompatActivity {
Server server=new Server();
OkHttpClient client;
TextView Form1Temp,Form1Moisture,Form2Temp,Form2Moisture;
TextView form1nrcs,form1mlr,form1svr,form1ann,form2nrcs,form2mlr,form2svr,form2ann;
EditText rainData;
Button btnRain;
    Handler handler;
    Runnable runnable;
    String status="false";

    @Override
    protected void onCreate(Bundle savedInstanceState) {
        super.onCreate ( savedInstanceState );
        setContentView ( R.layout.activity_main );
        client=new OkHttpClient();

        Form1Moisture=findViewById(R.id.Form1Moisture);
        Form2Moisture=findViewById(R.id.Form2Moisture);
        Form1Temp=findViewById(R.id.Form1Temp);
        Form2Temp=findViewById(R.id.Form2Temp);
        form1nrcs=findViewById(R.id.form1nrcs);
        form1mlr=findViewById(R.id.form1mlr);
        form1svr=findViewById(R.id.form1svr);
        form1ann=findViewById(R.id.form1ann);
        form2nrcs=findViewById(R.id.form1nrcs);
        form2mlr=findViewById(R.id.form2mlr);
        form2svr=findViewById(R.id.form2svr);
        form2ann=findViewById(R.id.form2ann);
```

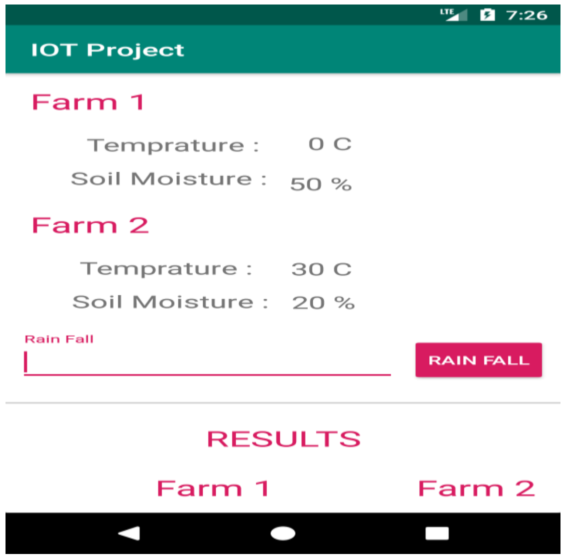
Andriod code 2



```xml
            android:layout_marginEnd="8dp"
            android:layout_marginBottom="16dp"
            android:text="ANN"
            android:textSize="20dp"
            app:layout_constraintBottom_toBottomOf="parent"
            app:layout_constraintEnd_toStartOf="@+id/form1ann"
            app:layout_constraintHorizontal_bias="0.0"
            app:layout_constraintStart_toStartOf="parent"
            app:layout_constraintTop_toBottomOf="@+id/textView11" />

        <TextView
            android:id="@+id/textView12"
            android:layout_width="wrap_content"
            android:layout_height="wrap_content"
            android:layout_marginTop="24dp"
            android:layout_marginEnd="16dp"
            android:text="Farm 2"
            android:textColor="@color/colorAccent"
            android:textSize="25dp"
            app:layout_constraintEnd_toEndOf="parent"
            app:layout_constraintTop_toBottomOf="@+id/textView33" />

        <TextView
            android:id="@+id/form1nrcs"
            android:layout_width="wrap_content"
            android:layout_height="wrap_content"
            android:layout_marginTop="8dp"
            android:layout_marginEnd="128dp"
            android:text="30"
            android:textSize="20dp"
            app:layout_constraintEnd_toStartOf="@+id/form2nrcs"
```

Emulator Snapshot

Emulator Snapshot 1



Emulator Snapshot 2