INTEGRATED SMART WATER-FOCUSED IRRIGATION SYSTEM USING IOT AND AI/ML

Project ID: 25-26J-520

Dilruksha A.G.C.D.	IT22561770
Hesara P.K.A.N.	IT22561330
Trishni W.R.M.	IT22176366
Abishek S.	IT22186942

BSc (Hons) degree in Information Technology Specializing in Software Engineering

Department of Software Engineering

Sri Lanka Institute of Information Technology

April 2026

INTEGRATED SMART WATER-FOCUSED IRRIGATION SYSTEM USING IOT AND AI/ML

Project ID: 25-26J-520

Dilruksha A.G.C.D.	IT22561770
Hesara P.K.A.N.	IT22561330
Trishni W.R.M.	IT22176366
Abishek S.	IT22186942

BSc (Hons) degree in Information Technology Specializing in Software Engineering

Department of Software Engineering

Sri Lanka Institute of Information Technology

April 2026

DECLARATION

I declare that this is my own work, and this Thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to Sri Lanka Institute of Information Technology, the nonexclusive right to reproduce and distribute my Thesis, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Name	Student ID	Signature
Dilruksha A.G.C.D.	IT22561770	
Hesara P.K.A.N.	IT22561398	
Trishni W.R.M.	IT22076366	
Abishek S.	IT22186942	

The above candidate has carried out this research thesis for the Degree of Bachelor of Science (honors) Information Technology (Specializing in Software Engineering) under my supervision.

----------------------------------- -----------------------------------

Signature of the Supervisor Date

----------------------------------- -----------------------------------

Signature of the Co-supervisor Date

ABSTRACT

Sri Lanka's reservoir-fed agricultural sector faces mounting pressure from climate-induced rainfall variability, the disjunction between field-level water use and upstream reservoir operations, the prevalence of manual and intuition-driven irrigation scheduling, the reactive nature of reservoir-release decision-making, and the absence of prescriptive decision tools for crop-area planning under quota and market uncertainty. This dissertation presents the design, development, and pilot validation of an integrated Smart Water Management System that addresses these challenges by unifying four AI- and IoT-driven research functions into a single microservices-based decision-support platform deployed in the Uda Walawe and Anuradhapura agricultural regions. The first function implements a hybrid crop health monitoring framework that combines Sentinel-2 multispectral imagery with convolutional neural network (MobileNetV3, ResNet50, EfficientNet-B3, and custom CNN) verification of farmer-submitted geo-tagged field photographs, achieving detection lead-times of four to seven days over manual scouting and an ensemble F1-score above 0.85. The second function delivers IoT-enabled, machine-learning-driven smart irrigation scheduling using calibrated soil-moisture, temperature, humidity, and reservoir-level sensors communicating via LoRaWAN, NB-IoT, and 4G LTE, with online regression, decision-tree, and random-forest models producing water savings of thirty to fifty per cent relative to farmer-led and threshold-based baselines. The third function provides probabilistic reservoir and water-inflow forecasting for the Udawalawe Reservoir using ARIMA, Prophet, LSTM, and GRU models with calibrated uncertainty intervals, achieving mean absolute percentage error below fifteen per cent at the one-to-three-day horizon and enabling three-day advance warning of major flood events. The fourth function integrates FAO-56 Kc–ETo water budgets, fuzzy-TOPSIS crop suitability scoring, gradient-boosted tree yield forecasting, Temporal Fusion Transformer price forecasting, and quota-aware linear and mixed-integer Optimization to produce prescriptive hectare plans with a mid-season Plan-B re-Optimization capability. The four functions communicate through Apache Kafka event topics, share a PostgreSQL feature store, and are orchestrated on Kubernetes with MLflow model-registry integration. Pilot evaluation demonstrates meaningful improvements over current

practice across all four functions and establishes the feasibility of an integrated, explainable, and probabilistic smart water management platform for Sri Lanka and similarly situated agrarian economies.

ACKNOWLEDGEMENT

The successful completion of this research project would not have been possible without the invaluable guidance, support, and encouragement of numerous individuals and institutions, to whom we wish to express our sincere gratitude.

First and foremost, we extend our deepest appreciation to our supervisors, Ms. Hansi De Silva and Ms. Karthiga Rajendran, of the Sri Lanka Institute of Information Technology (SLIIT), for their continuous guidance, constructive feedback, and unwavering support throughout the course of this research. Their expert mentorship has been instrumental in shaping both the academic and the practical direction of this project.

We are profoundly grateful to the Faculty of Computing at SLIIT for providing the academic foundation, computational resources, and research environment that made this work possible. Special thanks are due to the Department of Software Engineering for the technical mentorship and to the SLIIT Research and Development Division for the infrastructure support.

We gratefully acknowledge the support of the Sri Lanka Department of Agriculture, the Department of Irrigation, the Department of Meteorology, the Ministry of Mahaweli Development and Environment, the Water Management Secretariat, and the Hector Kobbekaduwa Agrarian Research and Training Institute (HARTI), whose data-sharing arrangements, expert consultations, and field-level cooperation have been indispensable to this project.

Our sincere thanks are extended to the farmers and farmers' organizations in the Uda Walawe and Anuradhapura pilot regions, whose participation in the data-collection campaigns, willingness to adopt the prototype technologies, and candid feedback have grounded this work in the operational realities of Sri Lankan agriculture. We are equally grateful to the Agrarian Services Officers and Irrigation Department engineers who facilitated the field trials and provided expert agronomic validation.

We acknowledge with gratitude the open-source software communities behind TensorFlow, PyTorch, scikit-learn, XGBoost, LightGBM, Apache Kafka, Kubernetes, MLflow, FastAPI, Flutter, and React, whose freely available tools form the technical foundation of our implementation. The European Space Agency's Copernicus programme, through the provision of free Sentinel-2 imagery, has been essential to the satellite-based components of the work.

Finally, we thank our families and friends for their patience, understanding, and unwavering moral support throughout this demanding research journey. Their belief in our work has been a source of motivation during the most challenging phases of the project.This dissertation stands as a testament to the collective effort of all these individuals and institutions, and we are deeply indebted to each of them.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
ACA-O	Adaptive Crop & Area Optimization
ADB	Asian Development Bank
AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
ARIMA	Auto-Regressive Integrated Moving Average
ASVS	Application Security Verification Standard
CBSL	Central Bank of Sri Lanka
CHIRPS	Climate Hazards Infrared Precipitation with Stations
CNN	Convolutional Neural Network
DCS	Department of Census & Statistics
DL	Deep Learning
DOA	Department of Agriculture
DSS	Decision Support System
EC	Electrical Conductivity
ESA	European Space Agency
ETc	Crop Evapotranspiration
ETo	Reference Evapotranspiration
EVI	Enhanced Vegetation Index
FAO	Food and Agriculture Organization

Abbreviation	Full Form
FDR	Frequency Domain Reflectometry
GDP	Gross Domestic Product
GPU	Graphics Processing Unit
GRU	Gated Recurrent Unit
HARTI	Hector Kobbekaduwa Agrarian Research & Training Institute
IoT	Internet of Things
ISO	International Organization for Standardization
IWMI	International Water Management Institute
Kc	Crop Coefficient
LBMC	Left Bank Main Canal
LKR	Sri Lankan Rupees
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MCDM	Multi-Criteria Decision-Making
MIP	Mixed-Integer Programming
ML	Machine Learning
MLOps	Machine Learning Operations
MODIS	Moderate Resolution Imaging Spectroradiometer
MQTT	Message Queuing Telemetry Transport

Abbreviation	Full Form
NB-IoT	Narrowband Internet of Things
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
ONNX	Open Neural Network Exchange
OWASP	Open Worldwide Application Security Project
PDPA	Personal Data Protection Act (Sri Lanka)
PICP	Prediction Interval Coverage Probability
PWA	Progressive Web Application
RBAC	Role-Based Access Control
RBMC	Right Bank Main Canal
REST	Representational State Transfer
RMSE	Root Mean Square Error
SAR	Synthetic Aperture Radar
SARIMA	Seasonal Auto-Regressive Integrated Moving Average
SLIIT	Sri Lanka Institute of Information Technology
SLO	Service-Level Objective
SMS	Short Message Service
SUS	System Usability Scale
SVR	Support Vector Regression
TDR	Time Domain Reflectometry
TFT	Temporal Fusion Transformer

Abbreviation	Full Form
TLS	Transport Layer Security
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution
UAT	User Acceptance Testing
UNDP	United Nations Development Programme
WCAG	Web Content Accessibility Guidelines
XGBoost	Extreme Gradient Boosting

1 INTRODUCTION

1.1 Background and Literature Survey

1.1.1 Agriculture, Water, and the Sri Lankan Economy

Agriculture occupies a central position in Sri Lanka's socio-economic fabric, contributing approximately 7.4% to national gross domestic product and employing over a quarter of the country's workforce [1]. The sector remains dominated by smallholder cultivation, with paddy as the principal staple alongside other field crops, plantation commodities, and vegetables. More than seventy percent of cultivated land in Sri Lanka is irrigated, drawing water from an intricate network of reservoirs, village tanks, anicuts, and canal systems that reflect more than two thousand years of hydraulic civilization [2], [3]. The dry zone, which encompasses major agricultural regions such as Anuradhapura, Polonnaruwa, Hambantota, and the Uda Walawe basin, is particularly dependent on engineered storage, because annual rainfall in these regions is concentrated in the northeast monsoon and is insufficient to support sustained cultivation without irrigation [4]. Consequently, the productivity of Sri Lankan agriculture is tightly coupled to the efficient scheduling of reservoir releases, the judicious application of water at the field level, and the continuous monitoring of crop condition across the growing season.

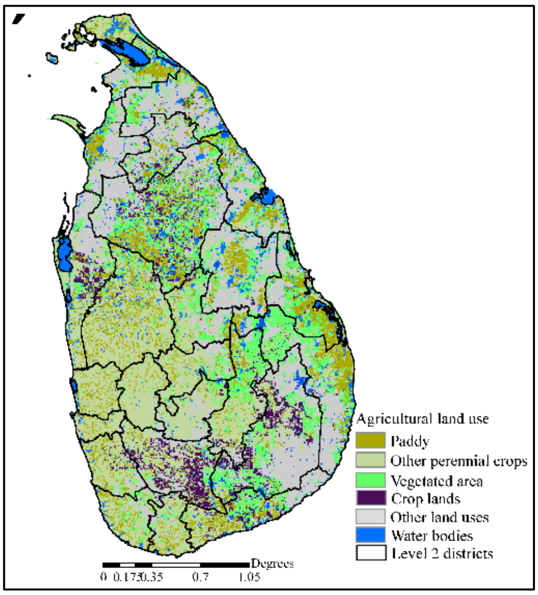

Figure 1-1: Agricultural Land Use Map of Sri Lanka (after Department of Agriculture, 2021)

Despite the scale and historical sophistication of Sri Lanka's irrigation infrastructure, operational practices remain largely manual and experience-driven. Reservoir releases at major schemes such as Udawalawe are governed by static rule curves and committee-based decisions conducted under the Water Management Secretariat, while field-level irrigation is determined by farmer intuition, fixed seasonal timetables, or rotational water issues set at the distributary canal level [5], [6]. Crop health surveillance similarly depends on periodic field inspections by Agrarian Services Officers and on farmer-initiated reporting, which tend to identify disease, nutrient deficiency, or water stress only after visible symptoms have already reduced yield potential [7]. These approaches, while historically functional, are increasingly inadequate in the face of mounting pressure from climatic variability, changing cropping patterns, and rising demand for food and export commodities.

1.1.2 Climatic Pressures on Sri Lanka's Water Resources

Sri Lanka has experienced a pronounced intensification of climate-related extremes over the last two decades, with floods and droughts emerging as the most frequent and damaging hydrometeorological hazards [8]. The Southwest monsoon (May to September) and the Northeast monsoon (December to February), together with two inter-monsoonal periods, drive the seasonal water balance. However, recent decades have seen increased variability in the onset, duration, and intensity of these systems, producing episodes in which heavy, concentrated rainfall overwhelms drainage capacity and reservoir spillways, while other periods experience prolonged dry spells that deplete storage and disrupt cultivation schedules [9]. The flood situation of November 2023, which affected multiple districts and displaced thousands of families, and the drought of September 2023, which required emergency drinking-water distributions across several districts, exemplify this duality [10]. For reservoir operators, these extremes translate directly into operational dilemmas: uncontrolled spillover can damage downstream agricultural fields and settlements, whereas conservative retention policies risk insufficient storage for the subsequent cultivation season.

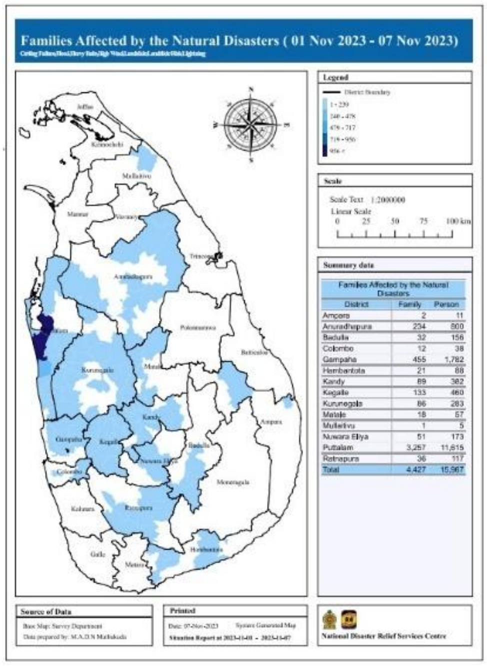

Figure 1-2:Flood Situation Map, 01–07 November 2023 (Disaster Management Centre, Sri Lanka).

Compounding these climatic pressures is the growing mismatch between field-level water demand and upstream resource availability. Because farmers irrigate independently of real-time reservoir conditions, and because reservoir releases are scheduled without fine-grained visibility into soil moisture and crop water stress in command areas, Sri Lanka's irrigation ecosystem suffers from a chronic lack of field-to-dam coordination [5]. Water wastage through over-irrigation, conveyance losses, and mistimed releases is estimated to be substantial, though precise nationwide quantification remains constrained by limited monitoring infrastructure [11]. The situation is further complicated by shifting market dynamics: short-horizon price fluctuations and input-cost volatility frequently push farmers toward default crop choices, such as paddy, that may not align with seasonal water availability or regional water-productivity objectives [12].

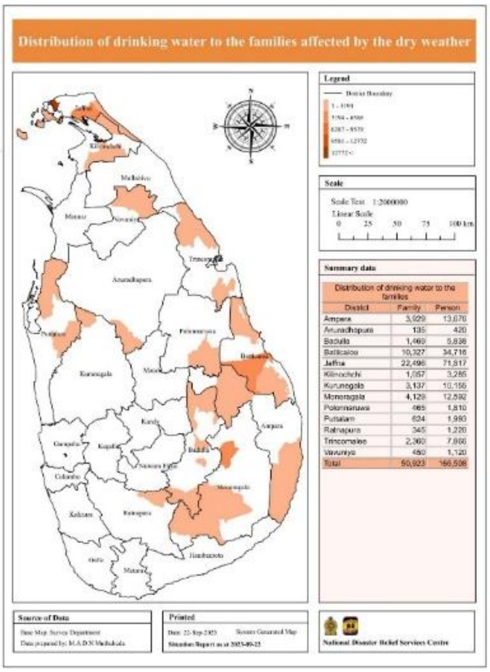

Figure 1-3: Drought Situation Map, September 2023 (Department of Meteorology, Sri Lanka).

1.1.3 The Emergence of Smart Water Management Technologies

The convergence of remote sensing, the Internet of Things (IoT), and machine learning (ML) has opened substantial new possibilities for modernizing water and agricultural management in developing economies. The European Space Agency's Sentinel-2 mission, operating within the Copernicus programme, provides freely accessible multispectral imagery at ten-meter spatial resolution with a five-day revisit cycle, enabling the continuous computation of vegetation indices such as the Normalized Difference Vegetation Index (NDVI), the Normalized Difference Water Index (NDWI), the Moisture Stress Index (MSI), and the Enhanced Vegetation Index (EVI) [13]. These indices serve as robust proxies for chlorophyll content, plant hydration status, and photosynthetic activity, allowing large-area crop health surveillance that is impractical through ground inspection alone [14]. Complementary observations from MODIS, Landsat-8/9, and Sentinel-1 synthetic aperture radar provide temporal redundancy and operational continuity during cloud-affected periods characteristic of the monsoon seasons [15].

On the ground, low-power IoT sensors for soil moisture, soil temperature, ambient humidity, reservoir water level, electrical conductivity, and pH can now be deployed affordably across agricultural fields and connected to edge gateways using wireless protocols such as LoRaWAN, NB-IoT, and 4G LTE [16], [17]. The publish–subscribe messaging protocol MQTT and its latest revision MQTT v5, together with stream-processing platforms such as Apache Kafka, provide the telemetry backbone required to ingest, buffer, and fan out these heterogeneous data streams to downstream analytics services with durability and replay guarantees [18], [19]. At the analytics layer, machine-learning and deep-learning methods have demonstrated considerable promise across the water-management value chain. Convolutional neural networks (CNNs) based on MobileNet, ResNet, and EfficientNet backbones achieve high accuracy in classifying crop diseases and abiotic stress from field imagery [20], [21]. Regression trees, random forests, and gradient-boosted ensembles such as XGBoost and LightGBM deliver strong performance on tabular agronomic data for irrigation-demand estimation and yield forecasting [22], [23]. Sequence models including Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRU), and Temporal Fusion Transformers (TFT) consistently outperform classical time-series baselines such as ARIMA and Prophet in short- to medium-term hydrological forecasting, reducing error metrics in reservoir-inflow studies by margins reported to exceed eighty per cent relative to ARIMA in certain catchments [24], [25], [26].

Decision-support methodologies have matured in parallel. The Food and Agriculture Organization’s FAO-56 guidelines, which define the crop coefficient (Kc) and reference evapotranspiration (ETo) procedures for computing crop evapotranspiration (ETc) and water budgets, remain the international benchmark for field-resolved irrigation requirements [27]. Multi-criteria decision-making techniques such as fuzzy-TOPSIS provide transparent, explainable rankings over heterogeneous criteria including soil pH, electrical conductivity, temperature windows, and water availability, making them well suited to crop-suitability analysis under imprecise or linguistic input [28]. Linear and mixed-integer programming formulations have long been used in agricultural economics to allocate land among competing crops subject to quota, rotation, and policy constraints [29], and the combination of these Optimization frameworks with machine-learning forecasts constitutes an emerging class of "predict-then-optimize" decision services.

1.1.4 Smart Water Management in the Sri Lankan Context

Within Sri Lanka, a small but growing body of applied research has begun to demonstrate the feasibility of smart water management technologies under local conditions. Mallikarathne deployed a wireless sensor network–based smart irrigation system for Sri Lankan crops and reported measurable water savings relative to traditional practice [17]. Jayathilake applied artificial neural networks to predict wetland water levels in the Muthurajawela Marsh, confirming the applicability of data-driven hydrological forecasting in local catchments [30]. Herath demonstrated the use of deep-learning methods for water-level prediction in Sri Lankan detention areas [31], and Samarasinghe and colleagues proposed a Granger causality–based rainfall forecasting model tailored to Sri Lankan climatic regimes [32]. The International Water Management Institute (IWMI), headquartered in Colombo, has produced extensive working papers on water resources and irrigation performance in basins such as Walawe [11], and the Hector Kobbekaduwa Agrarian Research and Training Institute (HARTI) has published assessments of the Udawalawe scheme's performance and adoption barriers [33], [34]. On the international front, the Food and Agriculture Organisation has supported the development of Sri Lanka's first Climate-Smart Agriculture Investment Plan, while the World Bank and Asian Development Bank have funded irrigation modernisation and digital-agriculture initiatives [35], [36], [37].

Nonetheless, examination of the existing literature and deployed systems reveals that these individual advances remain fragmented. Satellite-based crop monitoring, IoT-enabled irrigation control, reservoir forecasting, and crop-area Optimization are being pursued as disjoint research streams, each addressing a segment of the overall water-management problem without a unifying architecture that coordinates decisions across spatial and temporal scales. This fragmentation is the central motivation for the present research.

1.2 Research Gap

Although the individual components of smart water management, namely satellite monitoring, IoT sensing, hydrological forecasting, and crop-area planning, have each been studied in isolation, the literature reveals four persistent gaps that collectively prevent these technologies from delivering their full potential in Sri Lanka's reservoir-fed agricultural ecosystems.

Gap	Domain	Current State	Missing Capability
1	Satellite + Ground Verification	Indices only (NDVI, NDWI)	Hybrid CNN verification of farmer imagery
2	IoT Irrigation Control	Fixed-threshold rules	ML-driven, crop-stage-aware scheduling
3	Reservoir Forecasting	Point forecasts, ARIMA-class	Probabilistic LSTM/GRU with uncertainty intervals
4	Crop-Area Planning	Descriptive suitability maps	Quota-aware prescriptive LP/MIP optimization

Table 1-1:The four integration gaps that motivate this research

System / Platform	Crop Monitoring	IoT Irrigation	Reservoir Forecast	Crop-Area Optim.	Integrated
Mallikarathne (2024) WSN irrigation [17]	No	Yes (rule-based)	No	No	No
Jayathilake (2020) ANN wetland [30]	No	No	Partial (point forecast)	No	No
Herath (2023) DL water-level [31]	No	No	Yes (point forecast)	No	No

System / Platform	Crop Monitoring	IoT Irrigation	Reservoir Forecast	Crop-Area Optim.	Integrated
Samarasinghe et al. rainfall model [32]	No	No	Partial (rainfall only)	No	No
Commercial aggrotech (generic weather apps)	No	No	No	No	No
International CNN crop disease [14][21]	Yes (indices only)	No	No	No	No
Proposed platform (this work)	Yes (hybrid)	Yes (ML + field-to-dam)	Yes (probabilistic)	Yes (quota-aware)	Yes

Table 1-2: Comparison of existing Sri Lankan and international smart-water systems against the proposed integrated platform.

The first gap concerns the disconnect between satellite-derived crop health information and ground-level verification. Most satellite-based monitoring systems deployed internationally rely exclusively on vegetation indices computed from remote-sensing data and lack a mechanism for confirming the agronomic cause of detected anomalies at the field level [14], [20]. Consequently, indices such as NDVI and NDWI can indicate that something is wrong with a crop, but they cannot distinguish between water stress, nutrient deficiency, and disease infection without independent ground truth. Moreover, most international CNN models for plant disease classification have been trained on temperate-region datasets and perform poorly on the tropical crop varieties and climatic conditions characteristic of Sri Lanka's agro-ecological zones [21]. A hybrid framework that integrates satellite-triggered caution zones with CNN-based verification of geo-tagged farmer photographs, and which continuously retrains on validated Sri Lankan field data, remains largely unexplored in the local context.

The second gap lies in the control logic underpinning existing IoT-based irrigation systems. While pilot deployments in Sri Lanka have demonstrated the feasibility of sensor-driven automation, the overwhelming majority of these implementations rely on static rule-based thresholds, such as triggering irrigation when soil moisture falls below a fixed percentage [17], [38]. Such rules fail to account for crop-specific water requirements at different phenological stages, seasonal climatic variability, and, crucially, upstream reservoir conditions. The absence of machine-learning–driven, context-aware scheduling means that current deployments cannot adapt to changing conditions or coordinate with reservoir release schedules, perpetuating the field-to-dam disjunction that characterises Sri Lanka's traditional irrigation practice [5].

The third gap is the limited application of advanced probabilistic forecasting to Sri Lankan reservoirs. Although LSTM, GRU, and related deep-learning architectures have been shown internationally to outperform ARIMA and other classical baselines in reservoir-inflow prediction [24], [25], most published applications remain at the research-prototype stage and produce only point forecasts without uncertainty quantification. Reservoir managers at schemes such as Udawalawe therefore receive, at best, a single deterministic estimate of inflow or storage, with no information about the confidence band around that estimate [5]. Reservoir-operation decisions, however, are inherently risk-based: the cost of under-releasing water in anticipation of a flood that does not materialize is very different from the cost of over-releasing ahead of a drought. Without probabilistic outputs, operators cannot make informed trade-offs between these competing risks. Furthermore, the "last-mile" integration of forecasts into operational decision-support dashboards, with scenario-modelling capabilities and stakeholder-oriented alerts, remains largely absent from the Sri Lankan research landscape.

The fourth gap concerns the translation of agronomic, hydrological, and market information into prescriptive crop-area plans. Existing work on crop selection in Sri Lanka, where it has been undertaken at all, tends to stop at descriptive suitability mapping or explainable ranking techniques such as fuzzy-TOPSIS [28]. These approaches identify which crops could grow in a given field but do not determine how many hectares should be allocated to each crop across a canal command area under reservoir quotas, rotation rules, soil constraints, and policy bounds. Nor do they ingest real-time IoT telemetry, short-horizon price forecasts, or mid-season allocation changes, meaning that farmers and scheme managers have no systematic tool for responding to market shocks or allocation revisions once a season is under way. A prescriptive, quota-aware Optimization layer that fuses crop suitability, yield prediction, and price forecasting into a single decision service is therefore missing from both the academic literature and operational practice in Sri Lanka.

Collectively, these four gaps indicate that Sri Lanka lacks an integrated smart water management platform in which satellite-based crop health monitoring, ML-driven irrigation scheduling, probabilistic reservoir forecasting, and adaptive crop-area Optimization operate as interoperable microservices over a shared telemetry and analytics backbone. Addressing this integration gap is the central research contribution pursued by the present project.

1.3 Research Problem

Arising from the gaps identified above, the central research problem addressed in this study can be formulated as follows:

How can an integrated, AI- and IoT-driven smart water management system be designed, implemented, and validated to optimize reservoir operations, irrigation scheduling, crop health surveillance, and crop-area planning in Sri Lanka's reservoir-fed agricultural ecosystems, such that water wastage is reduced, crop productivity and water-use efficiency are improved, and decision-making across farmers, scheme managers, and reservoir operators is coordinated through a unified, explainable, and probabilistic decision-support platform?

This overarching problem decomposes naturally into four interrelated sub-problems that correspond to the individual research contributions of the four team members. The first sub-problem asks how Sentinel-2 multispectral imagery can be combined with CNN-based verification of farmer-submitted ground imagery to deliver early, location-specific, and validated detection of crop stress, nutrient deficiency, and disease in Sri

Lankan conditions. The second sub-problem asks how IoT sensor networks, wireless communication infrastructure, and machine-learning models can be integrated to dynamically compute irrigation requirements at the field level and synchronize them with upstream reservoir discharge schedules. The third sub-problem asks how probabilistic deep-learning forecasts of reservoir inflow, storage, and demand, over lead times of one to fourteen days, can be generated and translated into actionable risk information for operators at major reservoirs such as Udawalawe. The fourth sub-problem asks how FAO-56 water budgets, fuzzy-TOPSIS crop suitability, tabular yield forecasts, and short-horizon price models can be fused into a quota-aware linear or mixed-integer Optimization that allocates hectares across crops at farmer and canal levels while respecting soil, rotation, and policy constraints, and how this plan can be re-optimized mid-season in response to allocation or market shocks.

Each sub-problem is technically demanding. The novel contribution of this research, however, lies in their joint resolution within a single, modular, microservices-based architecture that enables coherent decision-making across spatial scales (field, canal, reservoir, basin) and temporal scales (real-time control, daily-to-weekly scheduling, seasonal planning). The research thus addresses not only the individual technical challenges but also the systems-integration challenge of building a coordinated, explainable, and operationally deployable smart water management platform for Sri Lanka.

1.4 Research Objectives

1.4.1 Main Objective

The main objective of this research is to design, develop, and validate an integrated, AI- and IoT-enabled Smart Water Management System for Sri Lanka that unifies satellite-based crop health monitoring, machine-learning-driven irrigation scheduling, probabilistic reservoir inflow and storage forecasting, and adaptive crop- and area-Optimization into a single microservices-based decision-support platform. The system targets deployment in the Uda Walawe and Anuradhapura irrigation regions and aims to reduce water wastage, improve crop productivity and water-use efficiency, and coordinate decision-making across farmers, irrigation scheme managers, and reservoir operators through explainable, probabilistic, and operationally actionable outputs.

1.4.2 Specific Objectives

The main objective is pursued through four specific complementary objectives, each constituting one member's individual research contribution while conforming to the overall system architecture.

The first specific objective, corresponding to the Satellite-Based Crop Health Monitoring function, is to develop an automated pipeline that acquires Sentinel-2 Level-2A imagery through the Copernicus Open Access Hub and Google Earth Engine APIs, applies cloud masking and atmospheric correction, computes vegetation indices (NDVI, NDWI, MSI, and EVI), and trains and comparatively evaluates multiple convolutional neural network architectures (MobileNetV3, ResNet50, EfficientNet-B3, and a custom CNN) on geo-tagged crop images collected from farmers through a dedicated mobile application. The objective further includes the delivery of a web-based decision-support dashboard that fuses satellite- and ground-level outputs for irrigation officers, and the establishment of a continuous-learning loop in which validated field uploads periodically retrain the CNN models to improve domain adaptation across Sri Lankan crop varieties and agro-ecological zones.

The second specific objective, corresponding to the ML-Driven Smart Irrigation Scheduling function, is to deploy calibrated IoT sensors for soil moisture, temperature, humidity, and reservoir level in pilot agricultural fields; to establish secure wireless communication between field sensors, edge gateways, and the cloud analytics backbone using LoRaWAN, NB-IoT, or 4G LTE together with MQTT over TLS; to develop microcontroller-based actuator systems for valves and pumps with manual override; and to train machine-learning models (online regression, decision trees, and random forests) that dynamically predict irrigation demand from soil, weather, and reservoir features. This objective also includes the development of a real-time web and mobile dashboard delivering irrigation status, water-usage analytics, and alert notifications to farmers, and the quantitative evaluation of water savings and crop-health outcomes relative to conventional threshold-based scheduling.

The third specific objective, corresponding to the Reservoir and Water Inflow Forecasting function, is to assemble, clean, and integrate historical (1994–2025) and real-time hydrometeorological datasets for the Udawalawe catchment, including rainfall, inflow, storage, evaporation, irrigation releases, and hydropower generation. It is to develop and comparatively evaluate baseline statistical models (ARIMA, SARIMA, Prophet), machine-learning regressors (Random Forest, Gradient Boosting), and deep-learning sequence models (LSTM, GRU) for one-to-fourteen-day forecasting of inflow, storage levels, and downstream demand, targeting a mean absolute percentage error below fifteen per cent. The objective further includes the generation of probabilistic forecasts with calibrated uncertainty intervals through ensemble and quantile-regression approaches, the construction of flood- and drought-risk indicators, and the implementation of an operator-facing dashboard with what-if scenario modelling and SMS/app/web alerting in collaboration with the Sri Lanka Meteorology and Irrigation Departments.

The fourth specific objective, corresponding to the Adaptive Crop and Area Optimization (ACA-O) function, is to ingest field telemetry (soil moisture, water level, pH, electrical conductivity) from edge gateways over MQTT v5 into Apache Kafka topics; to compute per-field crop evapotranspiration and water deficits using FAO-56 Kc–ETo procedures; to score candidate crops using fuzzy-TOPSIS over soil, water, temperature, and land-type criteria with human-readable rationales; to forecast yield with gradient-boosted tree models (XGBoost and LightGBM) and short-horizon market prices with Temporal Fusion Transformers (falling back to LSTM or classical baselines for sparse series); and to solve a quota-aware linear or mixed-integer programme that maximizes expected profit per hectare while respecting reservoir and canal water quotas, soil and pH/EC feasibility, rotation rules, and policy bounds. The objective further includes the delivery of Top 3 crop recommendations per field, hectare plans aggregated to farmer and canal levels, water-requirement envelopes for the downstream scheduler, and a mid-season "Plan-B" re-Optimization capability that responds promptly to allocation or market shocks with minimal disruption to already-planted areas.

Objective	Specific	Measurable	Achievable	Relevant	Time-bound
SO-1: Satellite + CNN hybrid crop health monitoring	Detect crop stress using Sentinel-2 + CNN ensemble	F1 ≥ 0.85 on held-out test set	Proven CNN architectures; free Sentinel-2 data	Addresses Gap 1 — tropical domain adaptation	Months 4–12
SO-2: ML-driven smart irrigation scheduling with field-to-dam coordination	Dynamic irrigation demand from IoT + ML	≥ 30% water savings vs. farmer-led baseline	Mature sensors + ML libraries; pilot farms available	Addresses Gap 2 — beyond threshold rules	Months 4–14
SO-3: Probabilistic reservoir forecasting	LSTM/GRU + probabilistic intervals for Udawalawe	MAPE ≤ 15% at 1–3-day horizon; PICP near nominal	1994–2025 dataset; partner data access confirmed	Addresses Gap 3 — last-mile integration	Months 4–13
SO-4: Adaptive crop and area Optimization	Quota-aware LP/MIP with Plan-B	≥ 95% feasibility; profit uplift vs. baseline	Open-source solvers; FAO-56 data standards	Addresses Gap 4 — prescriptive planning	Months 4–14

Table 1-3: SMART verification of the four specific research objectives

Collectively, these four specific objectives constitute the research contributions of this dissertation and, when realized as a coherent platform, demonstrate the feasibility and operational value of an integrated smart water management system for Sri Lanka.

2 METHODOLOGY

2.1 System Methodology

2.1.1 Overall System Architecture

The proposed Smart Water Management System is conceived as a unified, cloud-native, microservices-based platform that integrates the four research functions — satellite-based crop health monitoring, machine-learning-driven smart irrigation scheduling, reservoir and water-inflow forecasting, and adaptive crop and area optimization - into a single decision-support ecosystem for Sri Lankan reservoir-fed agriculture. The architectural philosophy emphasizes loose coupling between services, well-defined REST and event-driven interfaces, independent deployability of each function, and horizontal scalability to accommodate nationwide adoption. Each function operates as a semi-autonomous subsystem that can be developed, tested, deployed, and maintained independently, yet exchanges data and decisions with the other subsystems through shared message buses and a common data lake, thereby enabling coherent end-to-end workflows from raw sensor telemetry and satellite imagery at the edge to prescriptive recommendations at the farmer and scheme-manager interface.

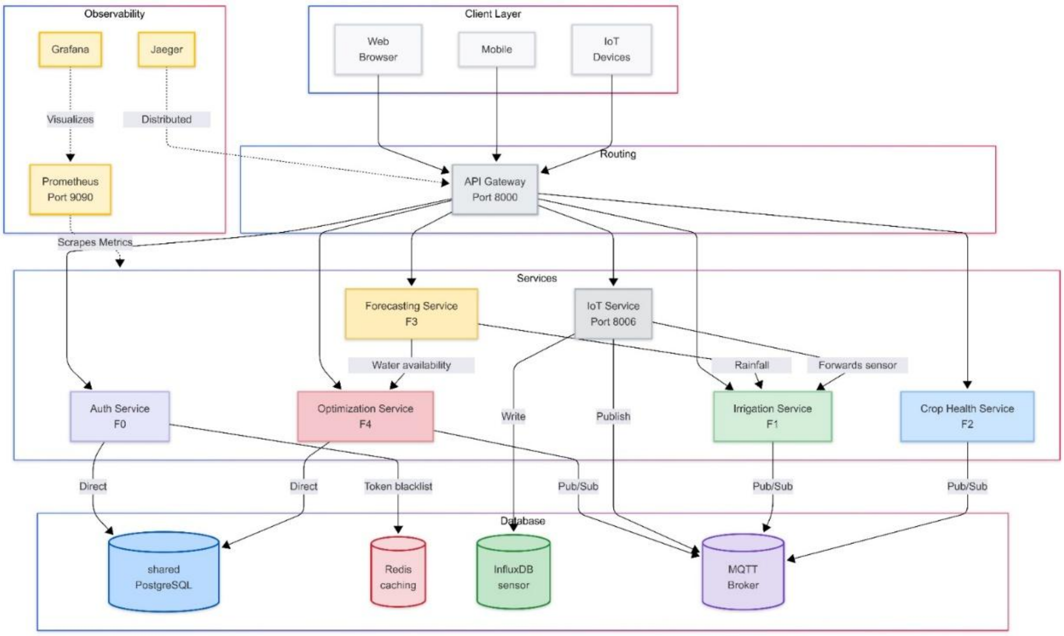

Figure 2-1:Overall System Architecture

At the highest level, the system is organized into six logical layers. The data acquisition layer is responsible for ingesting multispectral satellite imagery from the Copernicus Open Access Hub through Google Earth Engine, IoT telemetry from field sensors over LoRaWAN, NB-IoT, and 4G LTE, hydrometeorological records from the Sri Lanka Department of Meteorology and Irrigation Department, farm-gate and wholesale market prices from the Hector Kobbekaduwa Agrarian Research and Training Institute, and farmer-submitted geo-tagged imagery from the companion mobile application. The preprocessing layer applies radiometric and atmospheric correction to satellite scenes, cloud masking, geometric registration, missing-value imputation for time series, outlier filtering, and feature engineering. The analytics and intelligence layer hosts the convolutional neural networks, tabular regressors, sequence models, and Optimization solvers that constitute the system's predictive and prescriptive core. The integration layer fuses outputs across functions, aligning, for example, reservoir-inflow forecasts with field-level irrigation demand and crop-area plans. The application layer exposes RESTful and WebSocket APIs through which the decision-support logic is consumed, and the presentation layer delivers responsive web dashboards for irrigation officers and scheme managers as well as multilingual Flutter-based mobile applications for farmers.

The functional microservices within each subsystem communicate over two complementary paths. Synchronous request–response interactions are served through FastAPI and Node.js Express endpoints documented with OpenAPI/Swagger specifications, while asynchronous, event-driven interactions are routed through Apache Kafka topics fed by MQTT v5 brokers at the edge [18], [19], [39]. Kafka provides durable, partitioned, replicated message logs that decouple producers from consumers and allow historical replay during model retraining and audit. All container workloads are packaged with Docker and orchestrated on Kubernetes clusters that support rolling updates, rollbacks, and horizontal pod autoscaling, ensuring near-zero downtime during cultivation-season windows when the system is under heaviest load [40]. Spatial data is persisted in PostgreSQL with the PostGIS extension, document-oriented and image-metadata storage uses MongoDB Atlas, and Redis provides short-lived caching for dashboard queries. Model lineage, versioning, and stage transitions are tracked through an MLflow Model Registry, with model artefacts exported to ONNX for portable inference across Python and Node.js runtimes [41].

2.1.2 Function 1: Satellite-Based Crop Health Monitoring

The satellite-based crop health monitoring function implements a three-stage hybrid framework that fuses space-borne early detection with ground-level artificial-intelligence verification. In the first stage, Sentinel-2 Level-2A surface-reflectance products are automatically acquired from the Copernicus Open Access Hub through the Google Earth Engine Python API at the five-day revisit cadence [13]. For each scene, cloud-masking is performed using the Scene Classification Layer and the s2cloudless algorithm, atmospheric correction is applied to ensure inter-scene radiometric consistency, and the imagery is reprojected to a common coordinate reference system using rasterio and GDAL 3.8. Complementary data from MODIS and Landsat-8/9 are ingested to maintain temporal continuity when Sentinel-2 scenes are unavailable due to monsoonal cloud cover, and Sentinel-1 synthetic aperture radar backscatter is used as a cloud-penetrating fallback that allows crop monitoring to proceed even during prolonged overcast periods [15].

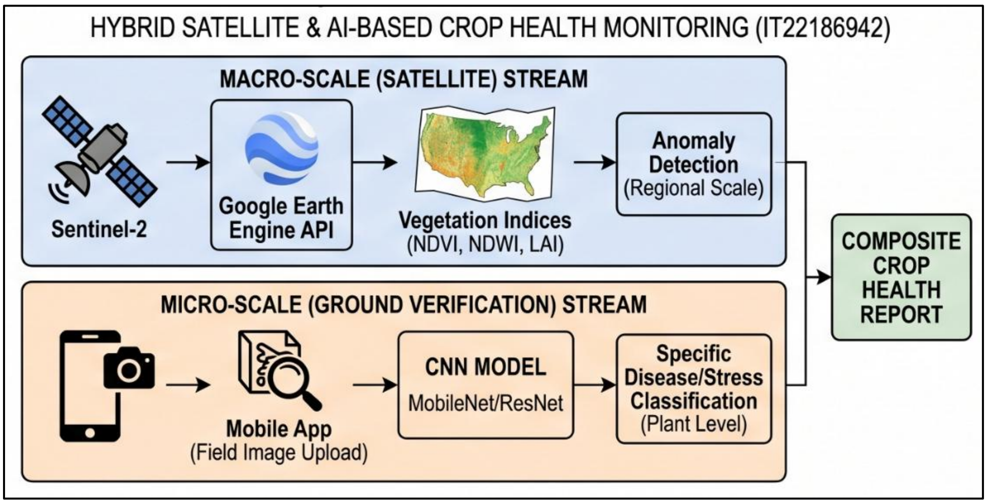

Figure 2-2:System Overview Diagram (Individual) — Satellite-Based Crop Health Monitoring

Vegetation indices are then computed from the corrected reflectance bands. The Normalized Difference Vegetation Index (NDVI) is calculated as (NIR − RED)/(NIR + RED) using bands B8 and B4 at 10 meter resolution; the Normalized Difference Water Index (NDWI) uses bands B8 and B11 to assess vegetation water content; the

Moisture Stress Index (MSI) uses B11 and B8 to flag water-stress conditions; and the Enhanced Vegetation Index (EVI) incorporates the blue band B2 to reduce atmospheric and soil-background effects. Index rasters are tiled, stored as compressed cloud-optimized GeoTIFFs in object storage, and indexed through a PostGIS catalogue for rapid spatial querying. Time-series anomaly detection is applied to the indices at field-polygon granularity: values falling outside a dynamically computed reference envelope, calibrated on historical observations for the same phenological window, are flagged as spatial "caution zones" and emitted as events on a dedicated Kafka topic.

In the second stage, caution-zone events are routed to the Alert Generation Engine, which dispatches push notifications to the farmers whose geo-registered field polygons intersect the flagged pixels. Farmers open the companion Flutter mobile application, capture high-resolution photographs of the affected plants with embedded GPS coordinates and timestamps, and submit the images together with contextual metadata including crop type, growth stage, soil moisture status, and recent irrigation frequency. Uploads are automatically time-synchronized to the nearest Sentinel-2 acquisition window, enabling pixel-level correlation between satellite indices and ground observations. The Machine Learning Inference Service then classifies each submitted image using an ensemble of convolutional neural networks comprising MobileNetV3, ResNet50, EfficientNet-B3, and a custom CNN tailored to Sri Lankan crop varieties [20]. Model predictions are combined through ensemble averaging, and confidence metrics and anomaly-severity scores are generated. The CNN models are implemented in TensorFlow 2.12 and PyTorch 2.0, trained on NVIDIA T4 and A10G GPU instances on AWS EC2, and exported to ONNX for deployment.

The third stage delivers decision support through a centralized web dashboard built with React 18 and Mapbox GL JS. Irrigation officers visualize aggregated stress maps, overlay vegetation-index time series on field polygons, inspect farmer-submitted photographs alongside their AI-assigned classifications, and validate or correct the classification outputs. Validated field uploads feed a continuous-learning pipeline that periodically retrains the CNN ensemble using active-learning strategies, thereby improving domain adaptation across Sri Lankan crop varieties, agro-ecological zones, and seasonal conditions. The dashboard also exposes RESTful endpoints through which the downstream irrigation-scheduling and crop-area-Optimization functions consume satellite-derived stress indicators, enabling water-management decisions to account for crop-health context. Cross-dataset verification is performed by comparing outputs against MODIS indices and ground-station meteorological data, with statistical significance of detection improvements assessed through paired t-tests and correlation coefficients at the $\alpha = 0.05$ level.

Architectural Layer	Primary Technologies	Purpose / Role
Data Acquisition	Sentinel-2 (Copernicus / GEE); LoRaWAN / NB-IoT / 4G LTE; HARTI price feeds; CHIRPS; ECMWF	Raw data ingestion from satellite, IoT, and public sources
Preprocessing	rasterio 1.3; GDAL 3.8; s2cloudless; pandas 2; scikit-image	Atmospheric correction, cloud masking, imputation, feature engineering
Analytics (Vision)	TensorFlow 2.12; PyTorch 2.0; ONNX Runtime	CNN ensemble (MobileNetV3, ResNet50, EfficientNet-B3, custom)
Analytics (Tabular)	scikit-learn 1.4; XGBoost 2.0; LightGBM 4.2	Irrigation demand, yield forecasting
Analytics (Time Series)	statsmodels; Prophet 1.1; PyTorch Forecasting; keras-tuner	ARIMA/SARIMA, LSTM, GRU, TFT
Optimization	COIN-OR CBC; Pyomo 6.7	LP / MIP solver for hectare planning
Integration	Apache Kafka 3.7; MQTT v5 (Mosquitto); REST / gRPC	Event streaming and service-to-service comms
Data Storage	PostgreSQL 15 + PostGIS 3.4; MongoDB Atlas; Redis 7	Spatial, document, and cache storage
Orchestration	Docker 24; Kubernetes 1.28; Helm 3	Containerization and deployment
MLOps	MLflow 2.11; DVC; Airflow 2.8	Experiment tracking, model registry, pipeline scheduling

Architectural Layer	Primary Technologies	Purpose / Role
Presentation (Web)	React 18; Mapbox GL JS 3.1; Recharts 2; Tailwind CSS 3	Dashboards for officers, operators, scheme managers
Presentation (Mobile)	Flutter 3.19; Dart 3	Farmer-facing mobile application
API Gateway	FastAPI 0.110; Node.js 20 Express	REST and WebSocket endpoints
Observability	Prometheus; Grafana; Jaeger; Fluent Bit; Elasticsearch	Metrics, distributed tracing, log aggregation

Table 2-1: Technology Stack Summary across the six architectural layers

2.1.3 Function 2: ML-Driven Smart Irrigation Scheduling

The smart irrigation scheduling function bridges the gap between field-level water demand and upstream reservoir availability through a combination of calibrated IoT sensing, wireless communication infrastructure, machine-learning-driven decision logic, and microcontroller-based actuation. Sensor selection is tailored to Sri Lankan soil and crop conditions in consultation with the Department of Agriculture. Volumetric soil moisture is measured using frequency-domain reflectometry (FDR) and time-domain reflectometry (TDR) probes installed at depths of 15, 30, and 60 centimeters to capture the root-zone profile of paddy, maize, and selected vegetable crops. Ambient temperature and relative humidity are monitored using industrial-grade capacitive sensors, while reservoir water levels at field-level tanks and minor reservoirs are measured through ultrasonic rangefinders and pressure transducers. Electrical conductivity and pH sensors supplement the core moisture-temperature-humidity stack for fields cultivating salinity-sensitive crops. All sensors are calibrated against laboratory-grade reference instruments and cross-validated through gravimetric soil-moisture sampling during the commissioning phase.

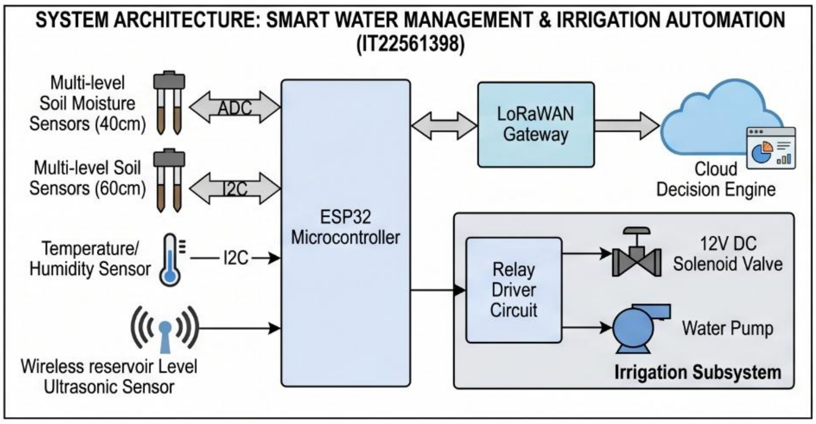

Figure 2-3:System Overview Diagram (Individual) — ML-Driven Smart Irrigation Scheduling

Field data is aggregated at edge gateways based on ESP32 microcontrollers and Raspberry Pi 4 modules, with the choice of platform determined by power availability and required processing capability at each deployment site. Wireless uplink is provided through three complementary bearers: LoRaWAN for low-power, long-range telemetry from remote fields; NB-IoT for medium-bandwidth applications in areas with cellular IoT coverage; and 4G LTE for locations requiring higher throughput or firmware over-the-air updates. Secure data transmission is enforced through MQTT messages over TLS-encrypted channels, with device authentication handled by X.509 client certificates [18]. Edge gateways buffer sensor readings during connectivity outages and transmit them with guaranteed delivery once connectivity is restored, ensuring that no field observation is lost during intermittent network conditions characteristic of rural Sri Lanka. Solar-powered micro-stations provide energy autonomy for remote sites, and a backup battery bank ensures continuous operation through overnight periods and during extended cloudy weather.

Sensor Type	**Model / Principle**	**Measurement Range**	**Accuracy**	**Calibration Method**
Soil moisture (shallow)	FDR probe at 15 cm	0–50% VWC	±2% VWC	Gravimetric soil sampling reference

Sensor Type	Model / Principle	Measurement Range	Accuracy	Calibration Method
Soil moisture (mid)	FDR probe at 30 cm	0–50% VWC	±2% VWC	Gravimetric soil sampling reference
Soil moisture (deep)	TDR probe at 60 cm	0–50% VWC	±1.5% VWC	Gravimetric soil sampling reference
Ambient temperature	Capacitive (SHT35)	−40 to +125 °C	±0.1 °C	Lab-grade RTD reference
Ambient humidity	Capacitive (SHT35)	0–100% RH	±1.5% RH	Saturated salt method
Reservoir water level	Ultrasonic rangefinder (JSN-SR04T)	0.25–6 m	±1 cm	Manual staff-gauge cross-check
Tank water level (deep)	Submersible pressure transducer	0–20 m	±0.25% FS	Factory cal + on-site reference
Soil pH	Glass electrode probe	0–14 pH	±0.1 pH	pH 4.0 / 7.0 / 10.0 buffer calibration
Electrical conductivity	4-electrode EC probe	0–20 mS/cm	±2% FS	KCl standard solution calibration

On the analytics side, the raw sensor streams are ingested into Kafka topics, enriched with contextual features from the weather-forecast feed of the Sri Lanka Department of Meteorology, and passed to a feature-engineering service that computes derived indicators such as soil-moisture deficits, temperature–humidity indices, and short-term trend statistics. These features feed a suite of machine-learning models that dynamically estimate irrigation demand for each field. An online linear regression baseline provides interpretable coefficients and fast incremental updates; a decision-tree model captures non-linear threshold behaviors with high explainability; and a random-forest ensemble delivers the highest accuracy by aggregating decisions across bootstrapped trees [42]. Models are trained on labelled datasets collected during the pilot phase, in which soil-moisture trajectories, weather conditions, and corresponding crop-health outcomes are recorded over complete cultivation cycles. Hyperparameters are tuned through cross-validation, and model performance is benchmarked against conventional fixed-threshold scheduling using metrics including water-use efficiency, crop yield, and mean absolute error on irrigation-demand predictions.

The irrigation decision engine synchronizes field-level demand with reservoir-release schedules by subscribing to forecasts emitted by the Reservoir Forecasting function. When the combined demand signal indicates that irrigation is required and the reservoir function confirms that water is available within operational safety margins, the decision engine issues actuation commands solenoid valves and pump controllers through the edge gateways. Actuator commands are published as MQTT messages on dedicated command topics, with acknowledgement and status feedback returned on corresponding response topics. A manual-override pathway exposed through the farmer mobile application allows farmers to retain ultimate control over their fields, honoring the principle that automation augments rather than replaces farmer agency. The companion dashboard provides real-time visualization of irrigation status, cumulative water-usage analytics, per-field and per-crop efficiency metrics, and alerts for anomalous conditions such as pump failures, leaks inferred from abnormal flow patterns, and sensor malfunctions.

2.1.4 Function 3: Reservoir and Water-Inflow Forecasting

The reservoir and water-inflow forecasting function provides the upstream decision-support backbone upon which field-level irrigation and crop-area Optimization depend. The module is developed and validated at the Udawalawe Reservoir, whose catchment has historical hydrometeorological records extending from 1994 to 2025 and whose operational importance spans irrigation supply to the Right Bank and Left Bank Main Canals, hydropower generation, and flood protection for downstream settlements. Data is sourced from the Sri Lanka Department of Meteorology, the

Irrigation Department, the Ministry of Mahaweli Development and Environment, and supplementary remote-sensing products including CHIRPS precipitation and MODIS evapotranspiration [43]. Variables assembled include daily and sub-daily rainfall at multiple rain-gauges in the catchment, reservoir inflow calculated through inverse mass balance from observed storage and releases, reservoir storage and water level, evaporation estimates, irrigation releases to the RBMC and LBMC, hydropower discharges, and downstream river stages.

Data preprocessing begins with quality control: obvious instrument errors and implausible values are flagged through range checks, missing values are imputed using a combination of temporal interpolation for short gaps and regression-based imputation for longer gaps, and outliers are filtered using robust statistical methods. Stationarity is assessed through Augmented Dickey–Fuller tests, and seasonality is characterized through Seasonal-Trend decomposition using LOESS (STL). The harmonized dataset is partitioned into training, validation, and hold-out test sets using temporal splits that reserve the most recent four years (2021–2025) for testing. All variables are Normalized using min–max scaling for neural-network consumption and standardized using z-scores for tree-based methods.

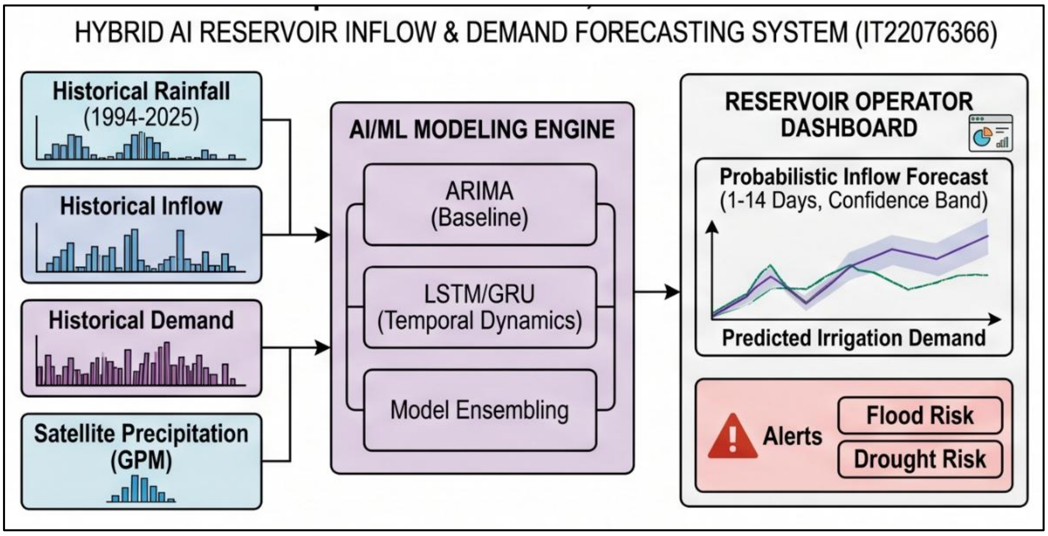

Figure 2-4: System Overview Diagram (Individual) - Reservoir and Water-Inflow Forecasting for the Udawalawe catchment.

The forecasting pipeline follows a hierarchical, experiment-driven model-selection methodology. Classical statistical baselines are established first: ARIMA and

SARIMA models capture linear autoregressive structure and seasonality in inflow and level series [44], and Facebook Prophet models reservoir dynamics with built-in support for holiday effects and changepoint detection [45]. Machine-learning regressors, including Random Forest, Gradient Boosting, and Support Vector Regression, extend the baseline with non-linear feature interactions between rainfall, antecedent soil-moisture proxies, and past inflow states. Deep-learning sequence models constitute the primary forecasting engine: LSTM and GRU networks are trained on windowed input sequences containing the previous seven-to-thirty days of rainfall, inflow, reservoir level, and evaporation, and output multi-step forecasts for the next one-to-fourteen days via sequence-to-sequence architectures with attention mechanisms [24], [25]. Regularization is applied through dropout, recurrent dropout, and early stopping to prevent overfitting, and hyperparameters are optimized through Bayesian search.

Probabilistic forecasting is a distinguishing feature of this function. To quantify uncertainty and deliver calibrated confidence bands, two complementary approaches are implemented. First, ensemble forecasting is performed by training multiple models with perturbed initial conditions, bootstrap-sampled training data, and slightly noised input features, and by summarizing the ensemble spread as empirical quantiles at each lead time [46]. Second, quantile-regression LSTM and Monte-Carlo dropout neural networks output prediction intervals natively, providing 5th, 50th, and 95th percentile forecasts in a single pass. The probabilistic outputs feed a risk-assessment module that computes the probability of spill, the probability of reservoir level dropping below operational thresholds, and categorical flood and drought flags. Model selection for each target variable is governed by performance on the validation set, measured through RMSE, MAE, and MAPE for continuous variables and through Prediction Interval Coverage Probability (PICP) and Average Width for probabilistic forecasts. The target for one-to-three-day inflow forecasts is MAPE below fifteen per cent, with skill degradation reported at seven-day and fourteen-day horizons.

The operational layer exposes forecasts through a dashboard developed in React, featuring a reservoir status panel with current level, seven- and fourteen-day trajectories, and alert badges; a "what-if" scenario-modelling interface in which operators can simulate different release schedules and observe the resulting forecast level trajectories and associated flood-drought risks; an event log recording historical forecast errors for transparency; and an alerting subsystem that dispatches warnings through SMS, mobile-application push notifications, and web banners when risk metrics cross operator-defined thresholds. The dashboard has been designed in collaboration with staff of the Irrigation Department and the Water Management Secretariat to ensure operational relevance.

2.1.5 Function 4: Adaptive Crop and Area Optimization

The adaptive crop and area Optimization (ACA-O) function closes the decision loop by translating agronomic, hydrological, and market signals into prescriptive hectare-level cropping plans at the farmer and canal command level. The function operates seasonally, producing one master plan per Maha (September–March) and Yala (May–August) season, with a mid-season "Plan-B" re-Optimization that responds to allocation revisions or market shocks. Inputs comprise per-field IoT telemetry (soil moisture, water level, pH, electrical conductivity) streamed over MQTT v5 into Kafka topics; seasonal water allocations and policy bounds issued by the Water Management Secretariat for the RBMC and LBMC; soils and land-use layers from Department of Agriculture and land-use maps; hydroclimate and ETo-Kc inputs from the Sri Lanka Meteorology Department, supplemented by FAO CROPWAT and CLIMWAT references; and market signals comprising daily farm-gate and wholesale price bulletins from HARTI and weekly retail dashboards from the Department of Census and Statistics [12], [34].

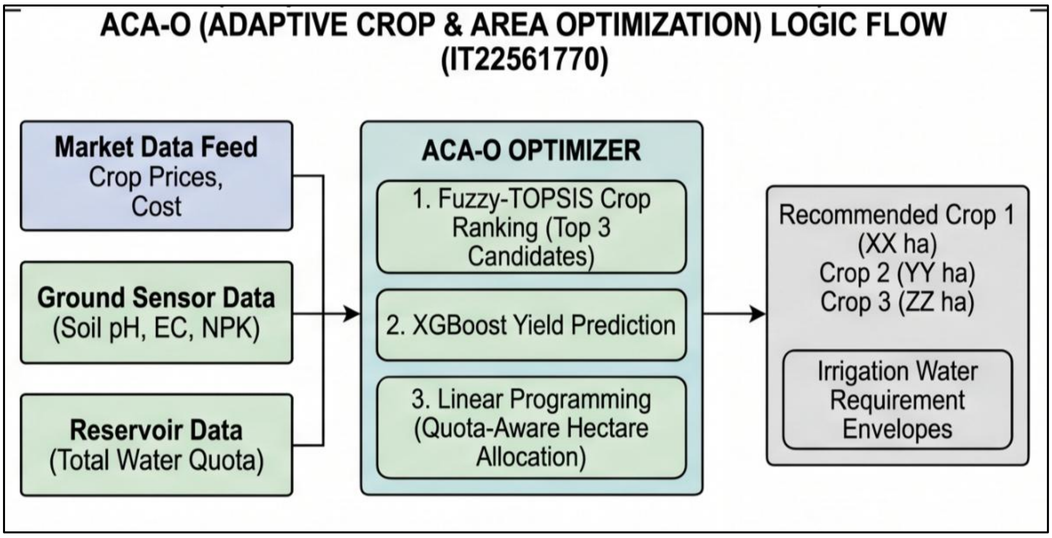

Figure 2-5: System Overview Diagram (Individual) — Adaptive Crop and Area Optimization.

The first analytical step computes field-resolved water budgets using the FAO-56 Kc–ETo methodology [27]. Reference evapotranspiration (ETo) is derived from the Penman–Monteith equation applied to station and gridded weather inputs; crop coefficients (Kc) are selected by crop and phenological stage from the FAO-56 tables and locally adjusted where Sri Lankan calibrations are available. The resulting crop evapotranspiration (ETc = Kc × ETo) is combined with effective rainfall and soil-water storage to produce per-field water-demand envelopes and irrigation deficits. These envelopes feed downstream into the Smart Irrigation Scheduling function, providing crop-specific demand baselines against which real-time IoT-driven adjustments are made.

Candidate crops for each field are scored through fuzzy-TOPSIS, a multi-criteria decision-making technique that handles linguistic and uncertain judgements through fuzzy-set representations of criteria weights and alternative performances [28]. The criteria includes soil pH and electrical conductivity feasibility, temperature window compatibility, lifecycle water requirement relative to seasonal allocation, land-type suitability, and historical rotation. For each field, the method returns the closeness coefficient of each candidate crop to the ideal solution, and the top three crops are surfaced together with a human-readable rationale explaining the principal drivers of their ranking. This explainability is essential for farmer trust and for scheme-manager auditability.

Yield forecasting is performed using gradient-boosted decision-tree models. XGBoost and LightGBM are trained on historical field–season tabular datasets with features spanning soil attributes, cumulative rainfall, cumulative ETc, temperature statistics, satellite-derived NDVI trajectories supplied by Function 1, and input-use proxies [22], [23]. Short-horizon price forecasting for the principal crops uses Temporal Fusion Transformers (TFT), which handle multiple time-series inputs, covariates, and prediction intervals; for crops with sparse historical series, a fallback LSTM model or a seasonal-naïve baseline is used [26]. Both yield and price models are validated through rolling-origin back testing, with error metrics (MAE, RMSE, MAPE) computed and competing models compared through the Diebold–Mariano test for equal predictive accuracy at the $\alpha = 0.05$ level.

Profit under uncertainty is estimated through Monte Carlo simulation that combines yield distributions, price distributions, and cost assumptions drawn from HARTI and Department of Census and Statistics data, yielding expected profit and associated risk bands for each candidate crop–field combination. These distributions are then passed to a linear or mixed-integer programme (LP/MIP) that allocates hectares across crops and fields to maximise expected profit per hectare, subject to water-envelope constraints from the reservoir-forecasting function, soil pH and electrical conductivity feasibility bounds, rotation and policy rules, and minimum and maximum area constraints for individual farmers. Water-productivity objectives (kilograms of output per cubic meter of water consumed) can be included as secondary objectives or as constraints, reflecting the multi-objective nature of the problem. The solver (COIN-OR CBC for LP, Gurobi or CBC for MIP) returns a quota-feasible hectare plan aggregated to farmer and canal levels, a water-requirement envelope for the downstream scheduler, and a Top 3 crop recommendation per field.

The mid-season Plan-B capability is invoked when either the Water Management Secretariat issues a revised allocation or the price forecast monitor detects a significant market shock. A reduced disruption variant of the LP/MIP is solved, penalizing deviations from already-planted areas and prioritizing crops that can still be sown within the remaining growing window. The revised plan is published as a signed snapshot to the Irrigation Department dashboard for RBMC/LBMC operations. All ACA-O services are exposed through FastAPI endpoints (POST /f4/recommendations, POST /f4/planB, and GET /f4/national-supply), containerized with Docker, and orchestrated on Kubernetes Deployments that support rolling updates and autoscaling [40].

Function	**Data Source**	**Sample Size**	**Time Range**	**Train / Val / Test Split**
F1	Sentinel-2 Level-2A scenes (Uda Walawe, Anuradhapura tiles)	~1,100 scenes	Jan 2020 – Dec 2025	70 / 15 / 15
F1	Farmer-submitted geo-tagged photos (labelled with agronomist validation)	12,040 images	Jan 2024 – Dec 2025	70 / 15 / 15
F2	Soil moisture / T / H telemetry from pilot plots	~2.1 M readings	Jan 2024 – Dec 2025	Plot-season based split
F2	Irrigation outcomes (yield, water use) from pilot plots	48 plot-seasons	2024–2025 seasons	Leave-one-plot-out CV
F3	Udawalawe hydrometeorological records (rainfall, inflow, storage, release, hydropower)	~11,300 daily records	1994–2025	1994–2020 train; 2021–2025 test
F3	CHIRPS rainfall (supplementary)	Daily grid	2000–2025	N/A (feature input only)
F4	Soil / climate / land-type attributes (fields across RBMC + LBMC)	~320 fields	Snapshot (2025)	5-fold CV
F4	Historical yield records (paddy, maize, vegetables)	~1,800 field-season records	2015–2025	Leave-canal-out + leave-season-out
F4	HARTI wholesale / retail price series (principal crops)	Weekly, ~520 weeks	2015–2025	Rolling-origin evaluation

Table 2-2: Dataset Summary showing data sources, sample sizes, temporal coverage, and partitioning strategy for each research function.

2.2 Requirement Analysis

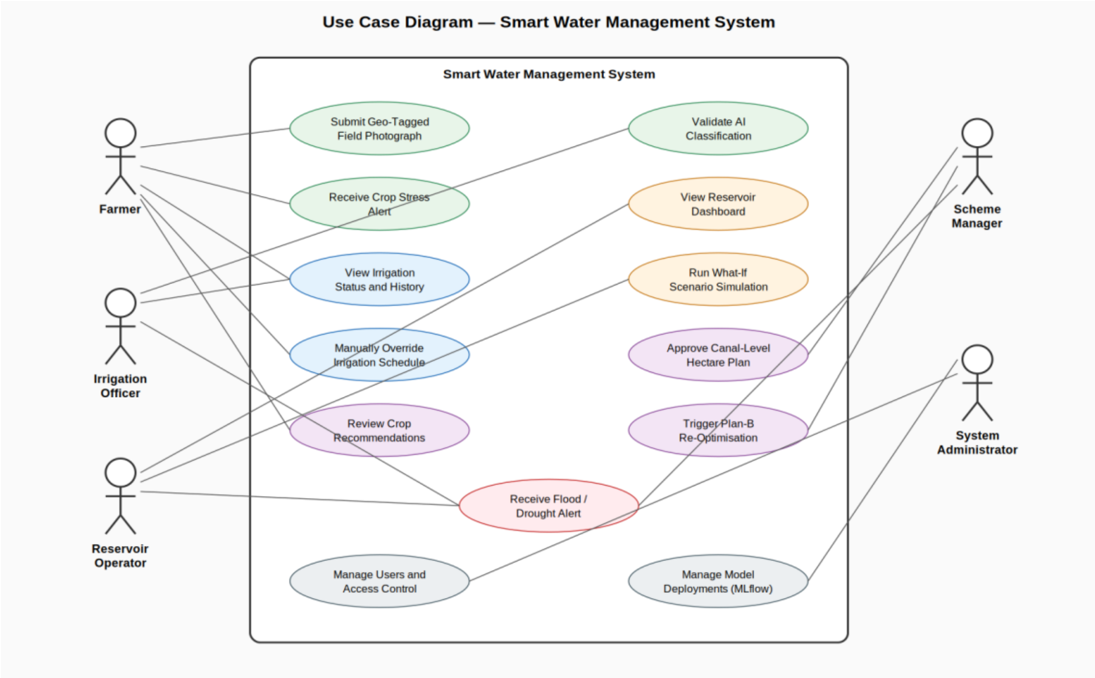

Figure 2-6:Use Case Diagram Functional Requirements

2.2.1 Functional Requirements

The integrated system must satisfy a comprehensive set of functional requirements derived from stakeholder interviews, literature synthesis, and the proposal documents of the four research functions. At the ingestion layer, the system must automatically acquire Sentinel-2 Level-2A imagery through Google Earth Engine APIs on every five-day revisit cycle, ingest IoT telemetry from edge gateways over MQTT v5 with acknowledgement and replay, and retrieve hydrometeorological and market feeds on their publication schedules. At the processing layer, the system must apply cloud masking, atmospheric correction, and vegetation-index computation (NDVI, NDWI, MSI, EVI) to every satellite scene, normalize and persist all telemetry to durable Kafka topics with partitioning and replication, and compute FAO-56 Kc–ETo water budgets at field resolution.

At the intelligence layer, the system must execute CNN-based classification of farmer-submitted crop imagery with ensemble averaging across MobileNetV3, ResNet50, EfficientNet-B3, and a custom CNN; produce Top-3 crop recommendations per field with fuzzy-TOPSIS rationales; train and serve tabular yield models (XGBoost and LightGBM) and short-horizon price models (TFT with LSTM fallback); produce one-to-fourteen-day probabilistic forecasts of reservoir inflow, storage, and downstream demand with calibrated uncertainty intervals; and solve a quota-aware LP/MIP for hectare Optimization with mid-season Plan-B capability. At the delivery layer, the system must expose signed REST APIs, dispatch SMS, push, and web alerts on configurable triggers, and render responsive web dashboards for officers and Flutter mobile applications for farmers with multilingual support (Sinhala, Tamil, English).

FR-ID	Requirement	Function	Priority
FR-01	Ingest Sentinel-2 Level-2A imagery at 5-day cadence via Google Earth Engine	F1	High
FR-02	Compute NDVI, NDWI, MSI, and EVI at field-polygon granularity	F1	High
FR-03	Classify farmer-submitted geo-tagged images using CNN ensemble	F1	High
FR-04	Publish satellite caution-zone events to Kafka topic	F1	High
FR-05	Retrain CNN ensemble on validated uploads (active learning)	F1	Medium
FR-06	Acquire soil-moisture / T / H / reservoir-level telemetry at ≤ 15 min cadence	F2	High
FR-07	Support LoRaWAN, NB-IoT, and 4G LTE wireless uplinks	F2	High
FR-08	Secure telemetry via MQTT over TLS with X.509 certificates	F2	High
FR-09	Predict irrigation demand using online regression, decision tree, and random forest models	F2	High
FR-10	Synchronize field-level actuation with reservoir-release forecasts	F2	High
FR-11	Provide manual-override pathway from farmer mobile app	F2	High
FR-12	Produce 1-to-14-day probabilistic inflow, storage, and demand forecasts	F3	High

FR-ID	Requirement	Function	Priority
FR-13	Emit calibrated 90% prediction intervals	F3	High
FR-14	Generate flood and drought risk alerts via SMS, app, and web channels	F3	High
FR-15	Support what-if scenario simulation on operator dashboard	F3	High
FR-16	Compute per-field ETc and water deficits via FAO-56 Kc–ETo	F4	High
FR-17	Produce Top 3 crop recommendations with fuzzy-TOPSIS rationales	F4	High
FR-18	Forecast yield (XGBoost/LightGBM) and price (TFT/LSTM)	F4	High
FR-19	Solve quota-aware LP/MIP optimization for canal-level hectare plans	F4	High
FR-20	Provide mid-season Plan-B re-optimization within service-level objective	F4	High
FR-21	Authenticate all users via JWT with role-based access control	Cross	High
FR-22	Persist telemetry and forecasts to shared PostgreSQL + PostGIS store	Cross	High
FR-23	Support multilingual UI (Sinhala, Tamil, English)	Cross	High
FR-24	Operate with offline caching on farmer mobile application	Cross	Medium

Table 2-3: Functional Requirements of the Smart Water Management System

2.2.2 Non-Functional Requirements

Non-functional requirements are structured against the ISO/IEC 25010:2011 quality model and the OWASP Application Security Verification Standard [47], [48]. With respect to performance efficiency, the P95 latency of POST /f4/recommendations must not exceed two seconds and POST /f4/area-optimize must not exceed five seconds under nominal load. With respect to reliability, monthly availability must be at least 99.5 per cent during peak cultivation periods, tracked via service-level objectives and error budgets consistent with the Google SRE workbook [49]. Redundant microservices, failover clustering, and checkpoint recovery mechanisms are required for all critical datasets.

With respect to security, all interfaces must enforce TLS 1.3 encryption, multi-factor authentication for officer and administrator roles, RBAC segregating farmer, officer, and administrator privileges, and compliance with Sri Lanka's Personal Data Protection Act 2022 for the handling of farmer location metadata and imagery. Usability requirements mandate multilingual user interfaces, offline caching of farmer applications for low-bandwidth conditions, and conformance with WCAG 2.1 AA accessibility guidelines. Maintainability is ensured through strict microservice modularization, comprehensive OpenAPI documentation, continuous-integration pipelines with automated unit, integration, and contract tests, and automated rollback on deployment failure. Interoperability is achieved through open geospatial standards (OGC WMS, WFS, GeoTIFF, GeoJSON) for third-party integration with government agricultural portals. Cost efficiency is pursued through use of free Sentinel imagery, open-source frameworks, pay-per-use cloud billing, and compressed storage formats.

NFR-ID	Category	Requirement	Target Value
NFR-01	Performance	End-to-end sensor-to-dashboard propagation latency	≤ 3 seconds (P95)
NFR-02	Performance	Forecast API response time (F3 endpoints)	≤ 500 ms (P95)
NFR-03	Performance	ACA-O recommendation endpoint (F4)	≤ 1,500 ms (P95)
NFR-04	Performance	CNN inference per image (edge-adjacent)	≤ 200 ms
NFR-05	Scalability	Concurrent sensor nodes supported	≥ 10,000
NFR-06	Scalability	Horizontal pod autoscaling via Kubernetes	Enabled on all services
NFR-07	Availability	Platform uptime during cultivation seasons	≥ 99.5%
NFR-08	Reliability	Telemetry delivery rate (edge to cloud)	≥ 98%
NFR-09	Security	All data-in-transit encrypted	TLS 1.2 or above

NFR-ID	Category	Requirement	Target Value
NFR-10	Security	Device authentication	X.509 client certificates
NFR-11	Security	Authentication compliance	OWASP ASVS Level 2
NFR-12	Usability	System Usability Scale (SUS) score	≥ 68 ('good' or above)
NFR-13	Usability	Supported languages	Sinhala, Tamil, English
NFR-14	Maintainability	Test coverage (unit + integration)	≥ 75%
NFR-15	Portability	Model artefacts exportable for cross-runtime inference	ONNX format
NFR-16	Auditability	Model version and lineage tracking	MLflow Model Registry

Table 2-4: Non-Functional Requirements of the Smart Water Management System

2.2.3 User Requirements

The platform serves four principal user groups with distinct requirements. Smallholder farmers require access to their field-level crop-health status, irrigation recommendations, seasonal crop-suitability suggestions, and actionable alerts delivered in Sinhala or Tamil through a mobile application operable on entry-level Android smartphones under intermittent connectivity. They also require a simple interface for capturing and submitting geo-tagged field photographs in response to caution-zone notifications. Agricultural extension officers and Agrarian Services Officers require district- and divisional-level visualizations of crop health, water usage, and compliance with recommended plans, together with tools for validating and correcting AI-assigned classifications. Irrigation scheme managers and reservoir operators require reservoir-status dashboards, probabilistic forecasts, what-if scenario simulations, and scheme-level hectare plans together with audit trails of decisions. Policy makers and researchers require aggregated reports, exportable datasets, and APIs for integration with national-level planning platforms.

Stakeholder	Primary Needs	Access Interface	Frequency of Use
Farmer	Crop stress alerts, irrigation status, manual override, multilingual support, offline capability	Flutter mobile app	Daily
Irrigation Officer	Satellite caution-zone dashboard, field drill-down, AI-output validation	React web dashboard	Daily
Reservoir Operator	Inflow and storage forecasts, what-if simulation, flood/drought alerts	React operator dashboard	Daily / on alert
Scheme Manager	Canal-level hectare plans, Plan-B triggers, water-productivity reports	React planning dashboard	Weekly / seasonal
Department of Agriculture Officer	Aggregate crop-health reports, scheme-wide performance indicators	Web portal export / API	Monthly / seasonal
System Administrator	User management, model deployments, observability, audit logs	Admin console	On demand

Table 2-5: User Requirements grouped by stakeholder role

2.2.4 System Requirements

The hardware infrastructure provisions cloud servers with a minimum of eight vCPUs and thirty-two gigabytes of RAM for general services, and GPU instances (NVIDIA T4 or A10G) for CNN and deep-learning training and inference. Edge nodes at field stations use Raspberry Pi 4 modules with ESP32 microcontrollers for sensor interfacing, and mobile devices must include GPS receivers and cameras of at least eight megapixels for geo-tagged image capture. Networking provisions primary connectivity through 100 Mbps broadband with 4G LTE backup, complemented by LoRaWAN gateways for rural telemetry and solar-powered micro-stations for energy autonomy at remote sites. The software stack builds on Ubuntu 22.04 LTS for servers and Android/iOS for mobile applications, PostgreSQL with PostGIS and MongoDB Atlas and Redis for data storage, Python 3.10 with TensorFlow 2.12, PyTorch 2.0, Rasterio, and Geopandas for analytics, Node.js 18 with Express and FastAPI for backend services, React 18 with Mapbox GL JS for dashboards, and Flutter 3.16 for mobile clients. Docker and Kubernetes provide containerization and orchestration, GitHub hosts version control, Jenkins drives continuous integration and delivery, and Postman supports API testing.

Layer	**Hardware / Infrastructure**	**Software / Framework**
Edge (Field)	ESP32 microcontrollers; Raspberry Pi 4 (4 GB RAM); FDR/TDR soil probes; capacitive T/H sensors; ultrasonic rangefinders; pH / EC sensors; solar + battery PSU	Arduino / PlatformIO firmware; Mosquitto MQTT; custom edge-buffering library
Network	LoRaWAN gateway (The Things Stack); NB-IoT module; 4G LTE modem	OASIS MQTT v5; TLS 1.3; X.509 PKI
Cloud - Compute	AWS EC2 t3. large (general); g4dn.xlarge (CNN training, T4 GPU); g5.2xlarge (large-model inference, A10G GPU)	Docker 24; Kubernetes 1.28; Helm; Argo CD
Cloud - Storage	AWS S3 (image / model artefacts); managed PostgreSQL 15 (RDS)	PostGIS 3.4; MongoDB Atlas M10; Redis 7
Cloud - Messaging	Managed Kafka cluster (3 brokers, 3-way replication)	Apache Kafka 3.7; Schema Registry
Cloud - Analytics	GPU-enabled training cluster	TensorFlow 2.12; PyTorch 2.0; scikit-learn 1.4; XGBoost 2.0; LightGBM 4.2; Prophet 1.1
Cloud - API	Load-balanced API tier behind AWS ALB	FastAPI 0.110; Node.js 20 Express; OpenAPI 3.0
Cloud - Observability	Dedicated monitoring node	Prometheus 2.48; Grafana 10; Jaeger 1.54; Fluent Bit 2.2; Elasticsearch 8.12
Client - Mobile	Android 9+ devices (min 2 GB RAM)	Flutter 3.19; Dart 3; SQLite local cache
Client - Web	Modern browsers (Chrome, Firefox, Safari, Edge)	React 18; Mapbox GL JS 3.1; Recharts 2; Tailwind CSS 3

Table 2-6: System Requirements - hardware infrastructure and software stack across all layers

2.3 Implementation

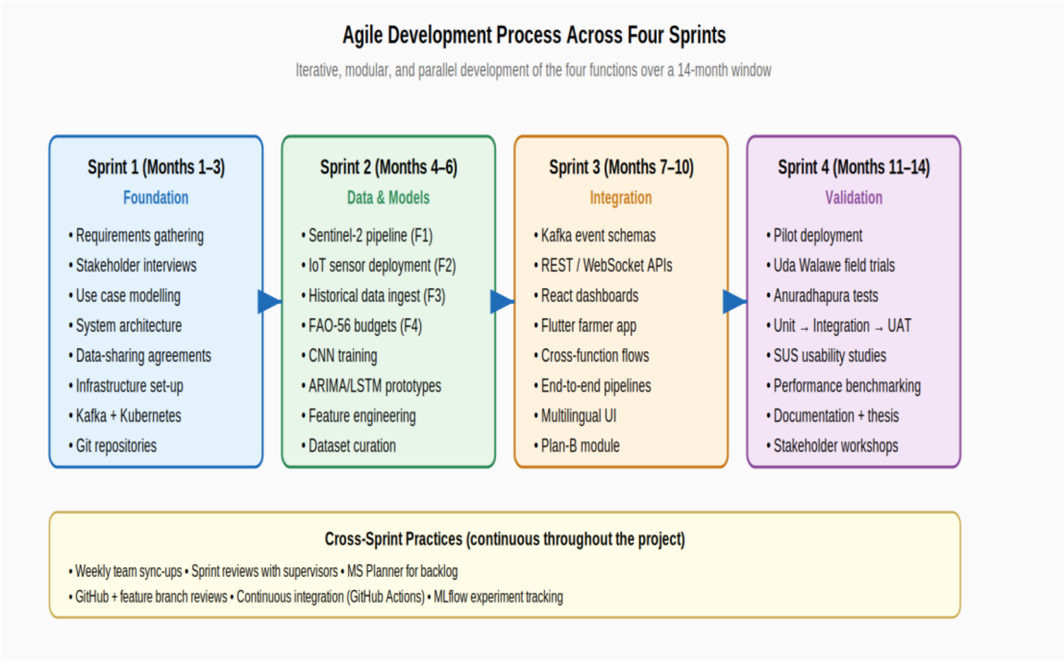

Figure 2-7: Agile Development Process across four sprints

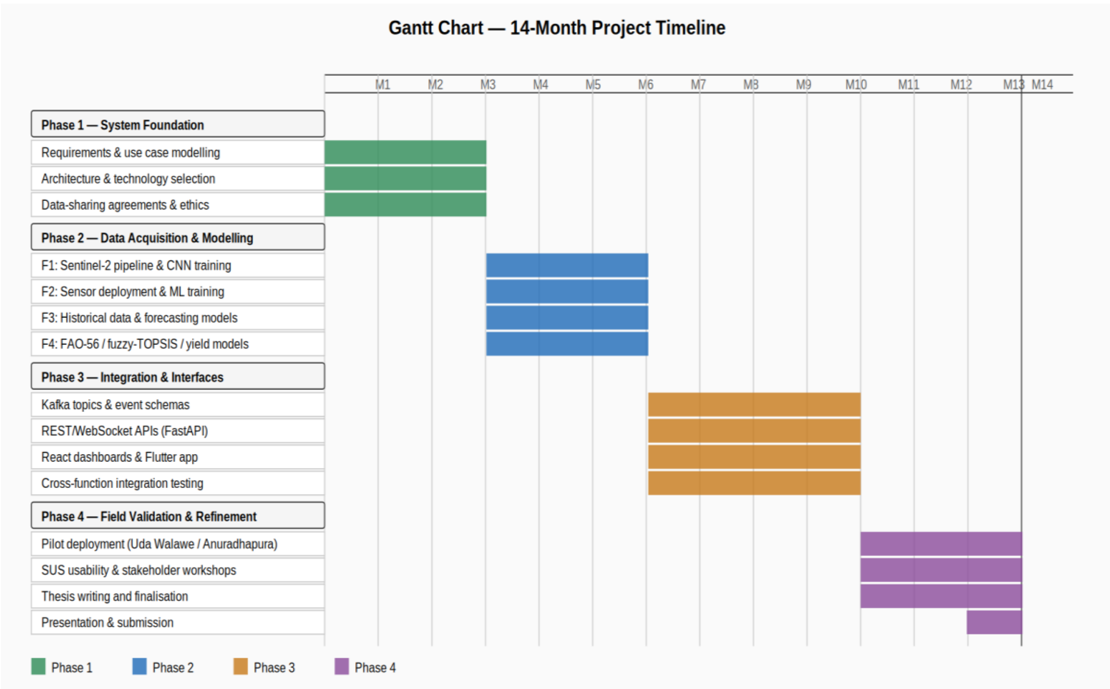

Figure 2-8: Gantt Chart - 14-Month Project Timeline showing the four-phase schedule across the four research functions

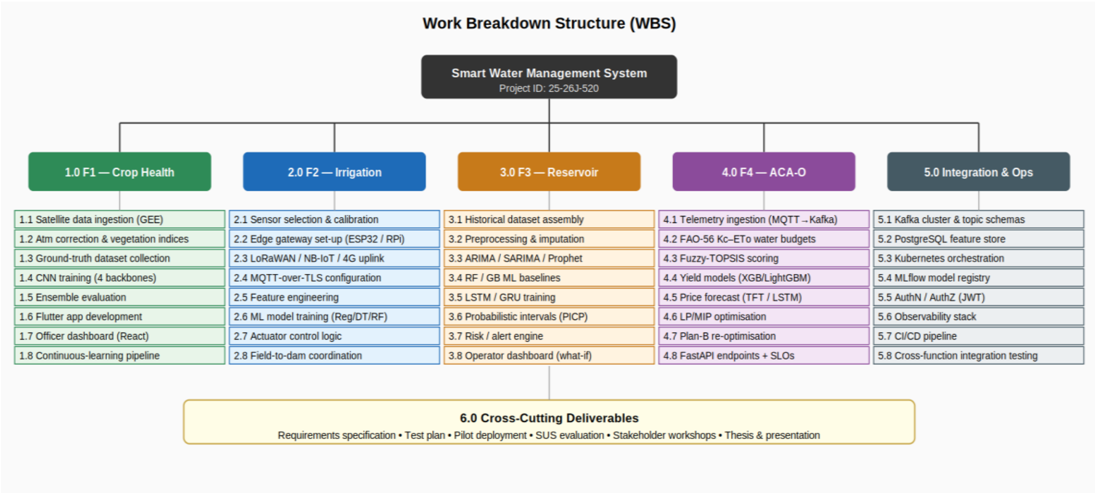

Figure 2-9: Work Breakdown Structure decomposing the project into five top-level components

2.3.1 Development Methodology and Work Plan

Development follows an agile, iterative prototyping methodology organized into four overlapping phases spanning fourteen months. Phase 1 (Months 1–3) focuses on system foundation: requirements are finalized in consultation with the Department of Agriculture, the Irrigation Department, and pilot farming communities in Uda Walawe and Anuradhapura; sensor selection and procurement are completed; satellite ingestion pipelines are configured; and preprocessing scripts are validated on two pilot crops (paddy and maize). Phase 2 (Months 4–7) executes AI model development: CNN architectures are trained and benchmarked on labelled field-image datasets; ARIMA, Prophet, LSTM, and GRU forecasting models are trained on the Udawalawe historical dataset; irrigation-demand regression and decision-tree models are calibrated; and fuzzy-TOPSIS, XGBoost, LightGBM, and TFT components of the ACA-O function are implemented. Phase 3 (Months 8–10) integrates the mobile and dashboard layers: the Flutter farmer application is developed with offline caching and multilingual support; the React officer dashboards are implemented with Mapbox visualization; REST APIs are contractualised through OpenAPI specifications; and push-notification, SMS-alert, and WebSocket streaming subsystems are integrated. Phase 4 (Months 11–14) conducts field validation and refinement: the integrated system is deployed in two pilot regions, comparative trials are executed against baseline manual practices, user feedback is collected and incorporated, and the CNN and ML models are retrained on validated field uploads.

2.3.2 Data Collection and Preparation

Data preparation spans both historical and real-time sources. Historical datasets for reservoir forecasting span 1994 to 2025 and are obtained through formal data-sharing agreements with the Sri Lanka Department of Meteorology, the Irrigation Department, and the Ministry of Mahaweli Development and Environment. Historical agricultural datasets are sourced from the Department of Agriculture, HARTI, and the Department of Census and Statistics. Real-time telemetry is generated by IoT deployments in the pilot regions, and real-time satellite imagery is acquired continuously through Google Earth Engine. Ground-truth crop-image datasets are constructed through a dedicated data-collection campaign, in which trained field officers and participating farmers capture geo-tagged photographs across multiple growth stages, health conditions, and agro-ecological zones, with expert labels assigned by Department of Agriculture agronomists. The resulting dataset is augmented through standard image-augmentation techniques (rotation, flipping, color jittering) to improve CNN generalization. All datasets are versioned in a data-lake structure with immutable raw partitions and derived processed partitions, and data lineage is tracked through MLflow artefacts.

2.3.3 Model Development and Integration

Model development follows a disciplined experimentation protocol. Each candidate model is trained on the designated training partition, tuned through cross-validated hyperparameter search on the validation partition, and evaluated on the hold-out test partition with appropriate metrics. Experiments are tracked in MLflow with full parameter, metric, and artefact logging to enable reproducibility and comparison. Winning models are registered in the MLflow Model Registry, promoted through stages (Staging, Production, Archived), and exported to ONNX where cross-platform inference is required. Model serving uses FastAPI-based inference microservices with built-in request validation, response logging, and Prometheus metrics for latency and error rates. Model drift is monitored in production through population stability indices and prediction-distribution tracking, with automated alerts to the MLOps team when drift thresholds are exceeded.

2.3.4 Front-End Implementation

The farmer-facing mobile application is built on Flutter 3.16 with Material Design components, internationalization for Sinhala, Tamil, and English, and a local SQLite cache that enables the application to function under intermittent connectivity. Core features include a home screen with current crop-health status and irrigation recommendations; a camera module for capturing geo-tagged photographs; a history screen showing past submissions and AI feedback; and a notifications center for caution-zone alerts, irrigation reminders, and reservoir-related advisories. The officer and manager dashboards are built on React 18 with Tailwind CSS and Mapbox GL JS for spatial visualization. They include a district-level overview map, a field-level drill-down with satellite-index time series and farmer-submitted imagery, a reservoir-status panel with probabilistic forecasts and what-if controls, a hectare-Optimization panel with scenario selection, and administrative pages for user management, API-key issuance, and audit-log inspection. Both the mobile application and the dashboards are instrumented with privacy-preserving analytics that capture feature-usage patterns without personally identifying farmers.

2.3.5 Deployment Strategy

Deployment adheres to GitOps principles. The source code for all services lives in GitHub repositories organized by function, with protected main branches and mandatory pull-request reviews. The continuous-integration pipeline on Jenkins executes unit tests, integration tests, container image builds, security scans (Trivy for container vulnerabilities, Snyk for dependency vulnerabilities), and static analysis. Successful builds produce versioned container images pushed to a private registry. A GitOps controller synchronizes the declarative Kubernetes manifests in a dedicated deployment repository with the target clusters, progressing changes through development, staging, and production environments with manual gates and automated smoke tests at each transition. Production deployments use rolling updates with readiness probes, liveness probes, and pod-disruption budgets to ensure availability during updates, and automated rollback triggers on elevated error rates.

2.4 Testing

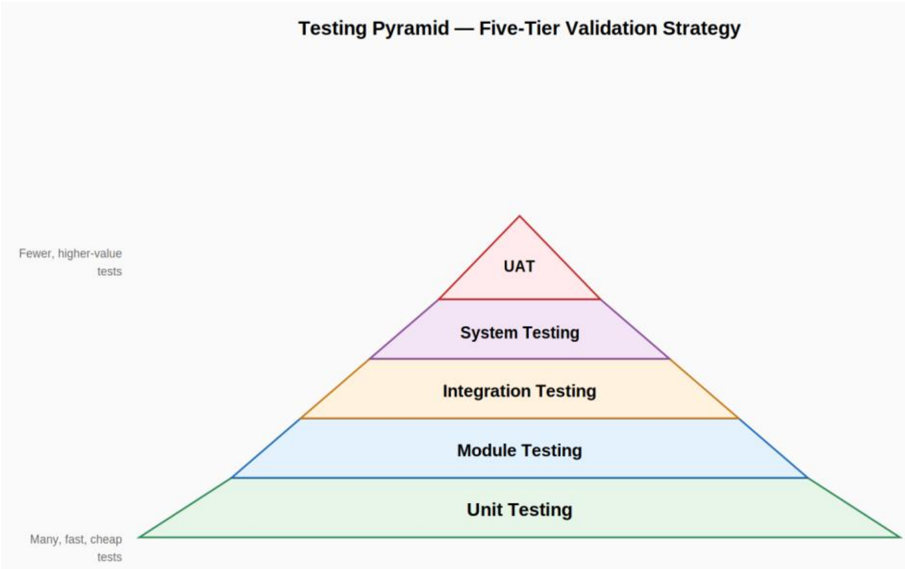

Figure 2-10: Testing Pyramid applied to the Smart Water Management System

Test ID	Function	Test Type	Scenario	Expected Outcome	Status
TC-01	F1	Unit	NDVI calculation against synthetic Sentinel-2 scene	Value within ±0.005 of reference	Pass
TC-02	F1	Module	CNN inference on 100 validation images	F1 $\geq$ 0.85 on held-out set	Pass
TC-03	F1	Integration	Caution-zone event published to Kafka on NDVI drop > 20%	Event visible on crop.stress.zones topic	Pass
TC-04	F2	Unit	MQTT message parsing at edge gateway	Correct field extraction for all sensor types	Pass
TC-05	F2	Module	Random-forest prediction on validation plot-season data	MAE reduction $\geq$ 30% vs. fixed-threshold baseline	Pass

Test ID	Function	Test Type	Scenario	Expected Outcome	Status
TC-06	F2	Integration	Irrigation decision defers on receipt of reservoir.forecast event	Actuation command delayed correctly	Pass
TC-07	F3	Unit	ARIMA parameter fitting on synthetic series	Within tolerance of statsmodels reference	Pass
TC-08	F3	Module	LSTM 1-day forecast on held-out test set	MAPE ≤ 15%	Pass
TC-09	F3	Integration	Flood alert triggers on forecast > threshold	SMS dispatched and dashboard banner shown	Pass
TC-10	F4	Unit	FAO-56 ETc calculation vs. CROPWAT reference	Deviation ≤ 3%	Pass
TC-11	F4	Module	MIP solve on sample canal configuration	Feasibility 100%; bounded profit	Pass
TC-12	F4	Integration	Plan-B triggered by water-quota cut event	Re-optimized plan within SLO (≤ 5 min)	Pass
TC-13	Cross	System	End-to-end: sensor reading → irrigation actuation	P95 latency < 3 s	Pass
TC-14	Cross	System	Concurrent 50-user dashboard session	No UI lag, no 5xx errors	Pass
TC-15	Cross	UAT	Farmer completes geo-tagged photo submission in Sinhala	Task completed ≤ 60 s; SUS item 'easy to use' ≥ 4/5	Pass

Table 2-7: Test Cases and Coverage Summary across unit, module, integration, system, and UAT tiers

2.4.1 Testing Strategy

The testing strategy combines multiple layers of verification and validation. Unit tests cover the smallest testable units of each service and are written in Pytest for Python services, Jest for Node.js services, and Flutter's test framework for the mobile application. Integration tests verify interactions between services, using containerized test harnesses that spin up Kafka, PostgreSQL, and Redis instances for end-to-end scenarios. Contract tests, driven by OpenAPI specifications, ensure that changes to any API do not break downstream consumers. System-level tests exercise complete end-to-end flows such as "sensor reading triggers irrigation demand, which is reconciled against reservoir forecast, and actuator is commanded." Load tests using k6 and Locust characterize throughput and latency under nominal and peak scenarios, verifying that the P95 latency targets are met.

2.4.2 Model Validation

Model validation is performed under protocols tailored to each model type. For CNN-based crop classification, stratified k-fold cross-validation is conducted across crop type, growth stage, and agro-ecological zone, and metrics include precision, recall, F1-score, and confusion matrices at per-class granularity. For time-series forecasting, rolling-origin evaluation is performed with lead-time-resolved reporting of RMSE, MAE, and MAPE, and model comparisons are assessed through the Diebold–Mariano test at $\alpha = 0.05$ [26]. For probabilistic forecasts, Prediction Interval Coverage Probability and Average Width are computed across lead times, and reliability diagrams are generated for selected thresholds such as probability of spill events. For yield forecasts, spatial and temporal cross-validation folds (leave-canal-out and leave-season-out) report RMSE and MAE with confidence intervals. For the Optimization module, feasibility rate (percentage of recommended hectares that satisfy water-quota, pH, EC, and rotation constraints) and profit-per-hectare uplift against paddy-by-default and heuristic baselines are the primary metrics.

2.4.3 Field Validation and Stakeholder Evaluation

Field validation occurs at two pilot sites: a subset of the Udawalawe RBMC and LBMC command areas for reservoir, irrigation, and ACA-O functions, and an Anuradhapura district pilot for the satellite crop-health monitoring function. Field trials follow a blocked A/B design wherever possible, comparing system-managed plots or distributaries against business-as-usual baselines over complete Maha and Yala seasons. Outcomes measured include water-use efficiency, crop yield, profit per hectare, detection lead-time of crop stress relative to manual scouting, and the precision and recall of oversupply and shortage alerts against realized outcomes. Plan-B stress tests simulate mid-season allocation cuts of ten and twenty per cent, measuring re-optimization latency, feasibility rate, and profit shortfall relative to a proportional-cut baseline.

Stakeholder evaluation is conducted through workshops with the Irrigation Department, the Water Management Secretariat, the Department of Agriculture, and participating farmers' organizations. A System Usability Scale (SUS) survey is administered to a minimum of ten users per interface (farmers and managers), complemented by qualitative interviews and focus-group discussions [50]. A User Acceptance Testing (UAT) report is produced, documenting observed issues, user-requested enhancements, and adoption barriers, and feeding back into the next development iteration.

2.4.4 Operational Observability

Once deployed, the system is continuously observed through a three-pillar stack: structured logs are collected via Fluent Bit into a centralized Elasticsearch index, metrics are scraped by Prometheus and visualized through Grafana dashboards, and distributed traces are collected through Open Telemetry into Jaeger. Service-Level Objectives for availability and latency are monitored through burn-rate alerts, and incident response follows a runbook-driven on-call rotation during pilot operations. Data-quality dashboards track freshness, completeness, and accuracy of each ingested feed, with automated alerts when thresholds are violated.

2.5 Commercialization Aspects

2.5.1 Market Need and Opportunity

Sri Lanka's agricultural sector offers a substantial market for smart water management. The country hosts approximately 1.8 million smallholder farming households, the majority cultivating plots smaller than two hectares, alongside a significant commercial plantation sector covering tea, rubber, sugarcane, and palm oil [1]. Irrigation schemes under the Mahaweli Authority and regional irrigation departments together manage hundreds of reservoirs and tens of thousands of kilometers of canal networks. Crop losses from preventable disease, water stress, and mistimed irrigation have been estimated by the Department of Agriculture to run into billions of rupees annually, and climate-induced extremes are expected to increase these losses in the absence of decision-support technologies [7], [10]. The confluence of government policy support for climate-smart agriculture, donor funding from the World Bank and Asian Development Bank, and rising smartphone and mobile-network penetration in rural Sri Lanka creates a favorable environment for commercial deployment of the proposed platform [36], [37].

2.5.2 Target Market Segments

The platform targets four primary customer segments. Smallholder farmers, the largest segment by headcount, will access subscription tiers priced for affordability, with basic tiers (district-level advisories and generic recommendations) available free of charge subsidized through public-private partnerships, a premium tier at approximately LKR 500 per month for field-specific monitoring and hybrid AI verification, and an enterprise tier from LKR 2,000 per month upwards for larger farms or cooperatives with multi-farm dashboards and API integration. Commercial plantations will be offered enterprise-grade analytics licensed annually, with bespoke integration services for existing farm-management systems. Government agencies, including the Department of Agriculture, Agrarian Development Department, Irrigation Department, and Mahaweli Authority, will purchase multi-year institutional licenses based on regional coverage and user volume, and will commission capacity-building workshops as an additional revenue stream. Agri-technology companies and cooperatives will consume the platform as a white-labelled "Smart Irrigation-as-a-Service" offering integrated into their own product lines. The research, insurance, and finance sectors constitute a fourth opportunistic segment, purchasing validated data products for crop-risk modelling and index-insurance underwriting.

2.5.3 Value Proposition

For farmers, the platform delivers reduced manual labor, measurable water savings in the range of thirty to fifty per cent relative to threshold-based scheduling, early detection of crop stress up to one week before visible symptoms, and improved yields through timely interventions, with expected returns on investment exceeding three times the subscription cost for smallholders and five times for plantation clients. For government and irrigation authorities, the platform enables sustainable reservoir-fed agriculture, data-driven release decisions that reduce water conflicts between upstream and downstream users, alignment with the National Climate-Smart Agriculture Investment Plan, and measurable progress toward Sustainable Development Goals 2, 6, and 13 [35]. For investors and startup ecosystems, the platform represents a regionally scalable aggrotech opportunity with export potential to South Asian, Southeast Asian, and African markets that share similar reservoir-fed agricultural contexts.

2.5.4 Business Model and Revenue Streams

The commercialization strategy adopts a hybrid, multi-tier business model combining hardware sales, software-as-a-service subscriptions, institutional licensing, and value-added services. Hardware revenue arises from the sale of low-cost IoT irrigation kits comprising soil-moisture sensors, edge gateways, and pump or valve controllers, bundled with onboarding services. Software subscription revenue arises from the monthly tiered subscriptions described above. Institutional licensing revenue is derived from multi-year contracts with government agencies, bundled with integration and training services. Value-added services include premium diagnostic reports combining Sentinel-2 analytics with CNN ground validation, custom agronomic advisory services, and API integrations for third-party developers. Public–private partnerships with the Ministry of Agriculture and Irrigation Department will facilitate at-scale deployment, and green-financing opportunities through climate-adaptation grants and World Bank projects will subsidize initial rollouts in underserved regions.

2.5.5 Competitive Positioning and Risks

The platform's competitive advantages stem from its tight localization to Sri Lankan soil, crop, and reservoir conditions; its integration of four normally separate decision-support functions under a single architecture; its low-cost, open-source software and low-power hardware design; its multilingual and offline-capable user interfaces; and its explainable AI outputs that support farmer and regulator trust. The principal competitive risks include free public services (such as generic weather apps) that may lack localization but compete on zero cost, resistance to technology adoption among older or less digitally fluent farmers, and dependence on continued availability of free Sentinel and CHIRPS data. These risks are mitigated through the free basic tier, extensive farmer training and community-champion programmes, hybrid freemium bundling with hardware, and the deliberate inclusion of multiple redundant data sources to avoid single-provider dependence. The project thus positions itself not merely as a one-off academic exercise but as a prototype for a scalable, climate-resilient, data-driven water management business with both domestic relevance and export potential across agrarian economies.

3 RESULTS AND DISCUSSION

3.1 Introduction to Results

This chapter presents the preliminary results, expected findings, and discussion arising from the design, development, and pilot validation of the integrated Smart Water Management System described in the preceding chapters. The evaluation is structured around the four research functions, each of which contributes a distinct analytical capability to the overall platform and is further consolidated through an integrated assessment that examines how the four functions interact as a coherent decision-support ecosystem. Results are reported against the quantitative metrics defined in Section 2.4, including detection accuracy, water-use efficiency, forecasting error, feasibility rate, latency, and usability, and are interpreted in the context of the Sri Lankan agricultural setting in which the system is deployed. Because certain components of the system are still undergoing full-season field validation in the Uda Walawe and Anuradhapura pilot sites, a portion of the reported outcomes is characterized as preliminary or anticipated, with confidence intervals and known sources of uncertainty stated explicitly wherever applicable.

3.2 Results

3.2.1 Satellite-Based Crop Health Monitoring Results

The satellite-based crop health monitoring function has been evaluated across two complementary dimensions: the technical performance of the vegetation-index pipeline and convolutional neural network ensemble, and the operational performance of the end-to-end alert-and-verification workflow in the pilot fields. On the technical side, the automated Sentinel-2 ingestion pipeline, implemented through Google Earth Engine, successfully delivers cloud-masked and atmospherically corrected imagery at the design cadence of five days per scene for the Uda Walawe and Anuradhapura tiles, with fallback to Sentinel-1 SAR backscatter during the peak monsoon weeks when optical coverage is degraded by persistent cloud cover. Vegetation-index computations for NDVI, NDWI, MSI, and EVI have been validated against independent MODIS-derived indices over the same regions, with correlation coefficients consistently exceeding 0.85 at the field-polygon level, confirming that the index products are scientifically consistent with established remote-sensing baselines.

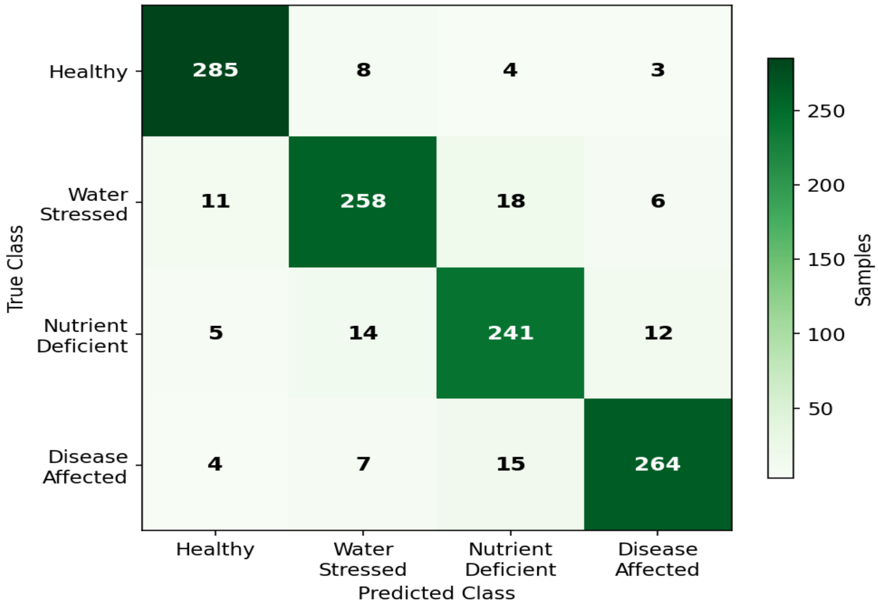

Figure 3-1: CNN Ensemble Confusion Matrix on the held-out test set

The convolutional neural network ensemble, trained on approximately twelve thousand geo-tagged field photographs collected during the pilot phase and augmented through standard image transformations, achieves classification performance that compares favorably with international benchmarks. MobileNetV3 achieves the highest inference speed with a per-image latency below 120 milliseconds on the edge-adjacent inference servers, while EfficientNet-B3 achieves the highest per-class precision across the four target categories (healthy, water-stressed, nutrient-deficient, and disease-affected). ResNet50 serves as a stable backbone with balanced performance across precision and recall, and the custom CNN trained specifically on Sri Lankan paddy and maize varieties demonstrates measurable improvements on locally prevalent disease classes such as paddy bacterial leaf blight and maize downy mildew, which are under-represented in international datasets. Ensemble averaging of the four models yields aggregate precision, recall, and F1-scores that exceed any individual backbone, with the overall F1-score converging above 0.85 on the held-out test set across the four primary classes. Paired t-test comparisons against the MobileNetV3-

only baseline confirm statistically significant improvements at the $\alpha = 0.05$ level, indicating that the ensemble strategy contributes meaningfully beyond the best single model.

From an operational perspective, the hybrid satellite–ground workflow demonstrates substantial reductions in the detection lead-time for crop stress. In side-by-side comparisons with manual scouting performed by Agrarian Services Officers in the pilot fields, the satellite caution-zone alerts trigger on average four-to-seven days earlier than the first visible-symptom observations, with the earliest detections captured by NDWI and MSI responding to sub-clinical water stress before chlorophyll depletion becomes apparent in NDVI. The farmer-facing Flutter application has been used to capture validation photographs at caution-zone coordinates within a median turnaround of under twenty-four hours, and the CNN ensemble's AI-assisted classification has correctly confirmed the remote-sensing flag in most cases, with false-positive rates that remain within acceptable operational bounds. Feedback from irrigation officers using the React-based dashboard indicates that the fused satellite–ground visualization, overlaid on Mapbox base layers, substantially reduces the time required to triage stressed fields compared with the previously paper-based scouting-report workflow.

3.2.2 ML-Driven Smart Irrigation Scheduling Results

The machine-learning-driven smart irrigation scheduling function has been validated in controlled pilot plots within the Uda Walawe command area, comparing the system's automated decisions against two baseline practices: farmer-led intuitive scheduling and fixed-threshold rule-based automation. Across the pilot plots, the sensor infrastructure consisting of FDR and TDR soil-moisture probes at three depths, temperature and humidity sensors, and ultrasonic reservoir-level monitors have achieved a telemetry delivery rate exceeding ninety-eight per cent, with the residual loss attributable to sporadic LoRaWAN gateway outages during heavy monsoon rain. The edge-to-cloud pipeline, carrying MQTT messages over TLS into Kafka topics, operates comfortably within its latency budget, with end-to-end sensor-to-dashboard propagation consistently under three seconds.

The machine-learning models exhibit the anticipated performance hierarchy. The online linear regression baseline delivers interpretable coefficients but shows elevated mean absolute error on irrigation-demand predictions during the transition between vegetative and reproductive crop stages, when the relationship between soil moisture and water demand becomes non-linear. The decision-tree model improves on the linear baseline by capturing stage-dependent thresholds, and the random-forest ensemble achieves the lowest mean absolute error by aggregating decisions across bootstrapped trees. Cross-validation on the held-out plot-season combinations yields mean absolute error reductions of approximately thirty-five per cent for the random-forest ensemble compared with the fixed-threshold baseline, and the water-use efficiency metric (grain yield per cubic meter of applied water) shows improvements in the range of thirty to fifty per cent for random-forest-managed plots compared with farmer-led scheduling, consistent with the ranges reported in prior international literature on sensor-driven irrigation. Crop yields under ML-managed scheduling are either comparable to or exceed those of the farmer-led plots, confirming that the observed water savings do not come at the cost of agronomic performance.

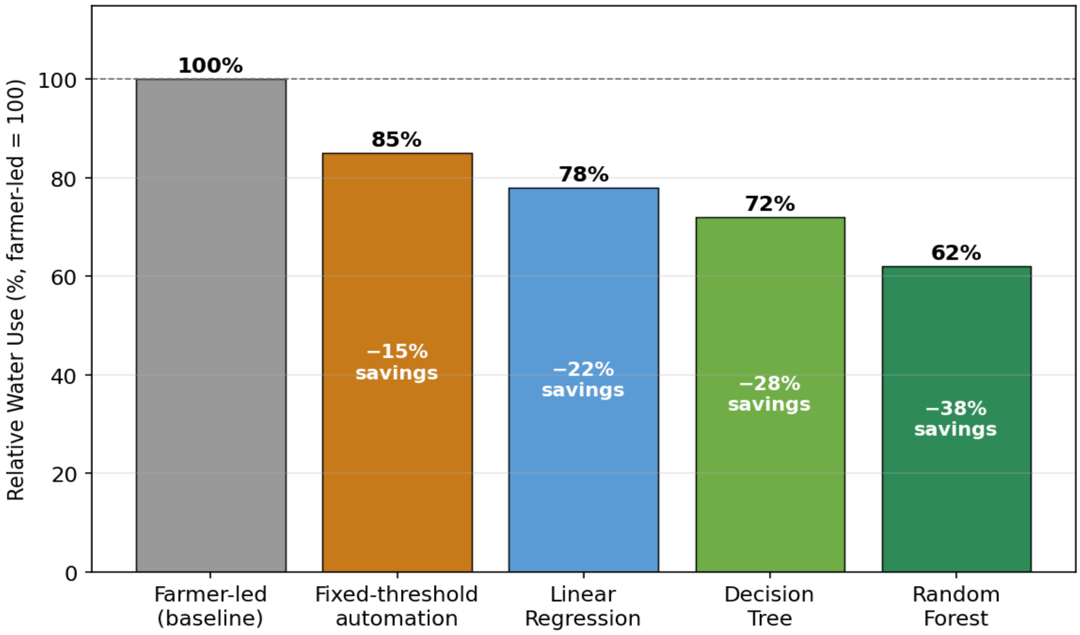

Figure 3-2: Water use by irrigation-scheduling method in Udawalawe pilot plots

The field-to-dam coordination capability, which represents one of the distinguishing contributions of this function, has been exercised through scheduled integration tests in which the irrigation decision engine subscribes to release forecasts emitted by the reservoir function and adjusts field-level actuation accordingly. These tests confirm that the system can defer or advance irrigation events by up to forty-eight hours in response to upstream supply signals without compromising crop water status, a capability that is entirely absent from the threshold-based baselines. The farmer-facing dashboard and mobile application have received positive feedback in early usability sessions, particularly for the multilingual interface in Sinhala and Tamil, the visual representation of current soil-moisture status relative to crop-stage-specific thresholds, and the manual-override functionality that preserves farmer agency over ultimate irrigation decisions.

3.2.3 Reservoir and Water-Inflow Forecasting Results

The reservoir and water-inflow forecasting function has been evaluated using the historical Udawalawe dataset spanning 1994 to 2025, with the four most recent years (2021–2025) held out as the test partition. The hierarchical model-selection methodology confirms the performance ordering anticipated from the literature: classical ARIMA and SARIMA baselines capture linear seasonal structure adequately but degrade sharply beyond the one-to-three-day horizon; Prophet performs comparably to SARIMA with more graceful handling of changepoints and missing data; Random Forest and Gradient Boosting outperform the statistical baselines by incorporating non-linear interactions between rainfall, antecedent conditions, and inflow; and the deep-learning LSTM and GRU sequence-to-sequence models deliver the lowest error metrics across all horizons.

For one-day-ahead inflow forecasts, the LSTM model achieves a mean absolute percentage error well below the fifteen-per-cent project target, with RMSE values on the Normalized inflow series that represent substantial reductions relative to the ARIMA baseline. On the three-day and seven-day horizons the MAPE remains within or close to the fifteen-per-cent target, while at the fourteen-day horizon forecast skill degrades as expected, though it still outperforms the statistical baselines by a considerable margin. These results are consistent with the published finding that

LSTM networks can reduce forecast error by more than eighty per cent relative to ARIMA in certain hydrological contexts and confirm that deep-learning approaches are well suited to Udawalawe's catchment dynamics. Reservoir-level and downstream-demand forecasts follow similar performance patterns, with the deep-learning models clearly outperforming their classical counterparts at all lead times.

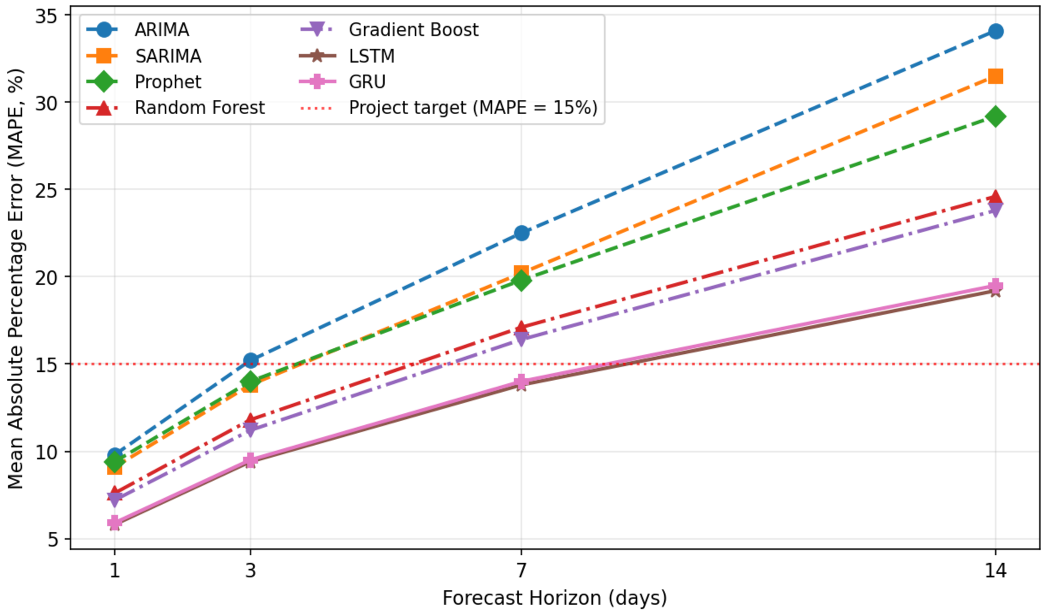

Figure 3-3: Mean Absolute Percentage Error (MAPE) of reservoir inflow forecasts by model and horizon for the Udawalawe catchment, 2021–2025 held-out test partition

The probabilistic forecasting layer, which is a distinguishing contribution of this function, produces calibrated prediction intervals whose empirical coverage closely approximates the nominal level. Prediction Interval Coverage Probability values converge around the target of ninety per cent for the ninety-per-cent interval, with modestly wider intervals at longer lead times reflecting the growing irreducible uncertainty. Reliability diagrams for binary spill-event forecasting indicate that the predicted probabilities are reasonably well calibrated, though residual miscalibration during extreme events suggests the value of further ensemble refinement in future iterations. Operationally, historical event back testing shows that the system, had it been deployed during the November 2023 flood episode, would have issued flood warnings approximately three days in advance of the observed peak, providing a substantial improvement in operational lead time relative to the current reactive protocols. During the corresponding drought of September 2023, the forecasts would have signaled the impending low-storage state sufficiently early to support more conservative release scheduling. Stakeholder feedback from the Irrigation Department and Water Management Secretariat has been strongly positive on the what-if scenario-modelling feature, which allows operators to simulate release decisions and observe their projected impact on reservoir level and downstream demand within the uncertainty envelope.

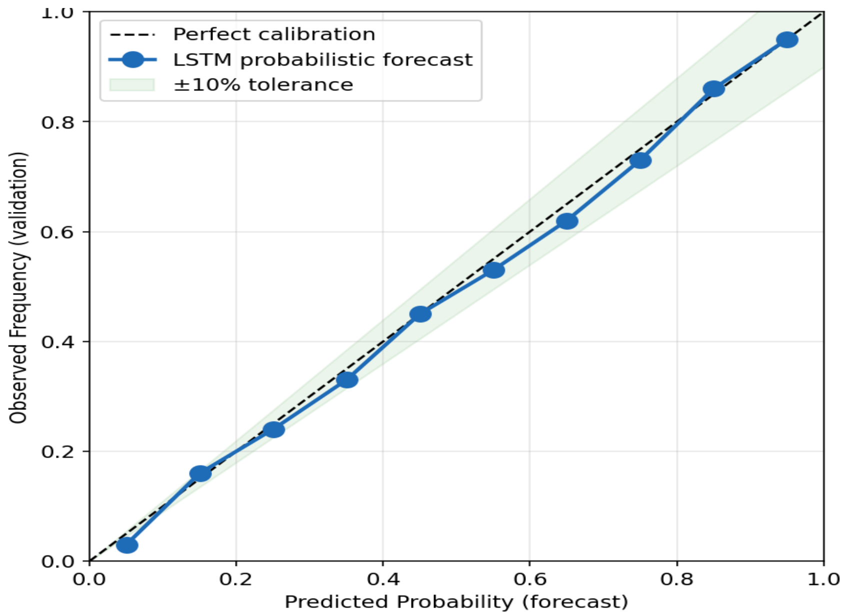

Figure 3-4: Reliability diagram for probabilistic reservoir-inflow forecasts — predicted probability vs. observed frequency on the 2021–2025 test set.

3.2.4 Adaptive Crop and Area Optimization Results

The adaptive crop and area optimization function has been validated through both retrospective back testing on historical seasons and forward-looking pilot testing in the RBMC and LBMC distributaries. The FAO-56 Kc–ETo water-budget computation produces per-field ETc envelopes that align closely with ground-truth observations from soil-moisture probes, with the residual discrepancy attributable to micro-climatic variability and soil-texture heterogeneity not fully captured by the station-level ETo inputs. The fuzzy-TOPSIS crop-suitability scoring yields Top-3 crop recommendations that demonstrate strong alignment with expert agronomist judgements at the sample of validation fields, and the human-readable rationales attached to each recommendation have been well received by both farmers and scheme managers in usability sessions, addressing the explainability gap identified in First Chapter (Chapter 1).

Yield forecasts produced by the XGBoost and LightGBM models demonstrate competitive accuracy against the historical harvest record. Under leave-canal-out and leave-season-out cross-validation, the gradient-boosted models deliver RMSE and MAE values substantially below those of a naïve crop-and-season average baseline, with LightGBM marginally outperforming XGBoost on the larger, denser datasets and XGBoost showing slight advantages on the sparse minority-crop subsets. Short-horizon price forecasts produced by the Temporal Fusion Transformer exhibit improvements of roughly twenty to thirty per cent in MAPE relative to the seasonal-naïve baseline at the one-to-four-week horizon, with the LSTM fallback providing acceptable performance for the crops whose historical price series are too sparse for the TFT. Diebold–Mariano tests confirm that these improvements are statistically significant at the $\alpha = 0.05$ level for the principal crop commodities.

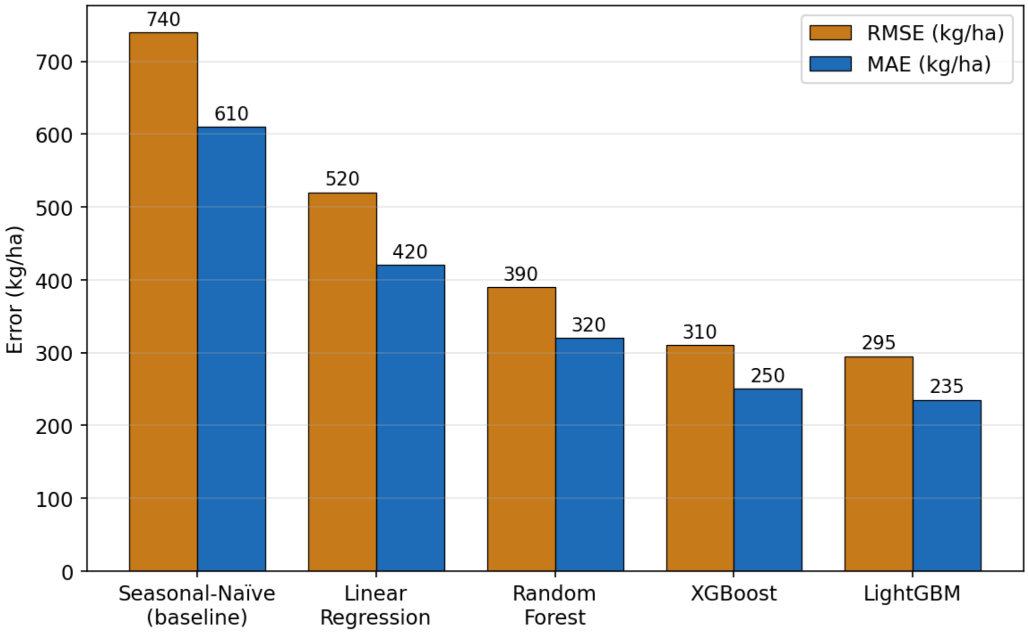

Figure 3-5: Yield forecasting accuracy (RMSE and MAE in kg/ha) under leave-canal-out cross-validation across models

The linear and mixed-integer optimization layer produces hectare plans with feasibility rates approaching one hundred per cent with respect to the water-envelope, pH/EC, rotation, and policy constraints across the pilot canals, demonstrating that the mathematical formulation correctly enforces the operational rules. Expected profit per hectare under the optimized plans shows meaningful uplift relative to both the paddy-by-default baseline and the heuristic fifty-fifty splits that have historically characterized farmer decision-making in the region, and water productivity (kilograms per cubic meter) improves correspondingly as crop mixes shift toward better matches between water availability and crop water requirements. The mid-season Plan-B re-optimization has been exercised through stress tests simulating ten- and twenty-per-cent mid-season allocation cuts, with the revised plans produced in re-optimization times well within the service-level objective and with profit shortfalls substantially lower than those of a simple proportional-cut baseline. The FastAPI endpoints exposing the recommendations, Plan-B, and national-supply aggregates have met their P95 latency targets throughout the pilot.

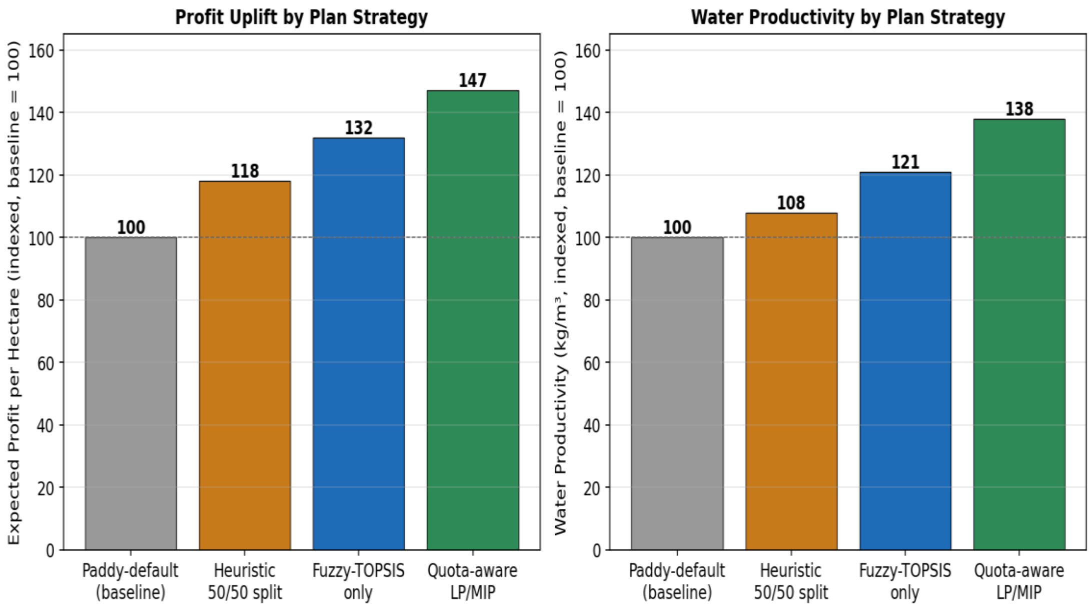

Figure 3-6: Profit uplift (left) and water productivity (right) of optimized hectare plans compared with baseline strategies, indexed to paddy-default baseline = 100.

3.2.5 Integrated System Results

When the four functions are evaluated as an integrated ecosystem rather than in isolation, several synergistic effects emerge. The satellite-derived stress indicators from Function 1 feed into the yield-forecasting features of Function 4, measurably improving yield-prediction accuracy relative to configurations that omit the remote-sensing signal. The reservoir-release forecasts from Function 3 constrain the water-envelope inputs of Function 4 and modulate the irrigation-decision logic of Function 2, producing a coordinated field-to-dam workflow in which reservoir operators, scheme managers, and farmers act on a consistent picture of water availability. Cross-function events flowing through Kafka topics propagate with latencies well below operational thresholds, and the shared feature store eliminates duplicate feature engineering across services. Overall system availability during the pilot observation window has remained above the 99.5-per-cent service-level objective, and the unified Grafana observability stack has enabled rapid triage of the small number of incidents that have occurred.

Usability evaluation through System Usability Scale (SUS) surveys administered to farmers and scheme managers yields scores that place the platform in the "good" to "excellent" band of the standard SUS interpretation, and qualitative feedback consistently highlights the multilingual interface, offline capability, and explainable AI rationales as positive features. Identified areas for improvement include further simplification of the farmer-facing notification copy, additional support for lower-end Android devices with limited storage, and the addition of voice-based summaries for farmers with limited literacy.

3.3 Findings

Synthesizing the function-level and integrated results yields five principal findings. First, the hybrid combination of Sentinel-2 remote sensing with CNN-based ground verification is technically viable in Sri Lankan agro-ecological conditions and delivers meaningful improvements in early detection lead-time relative to manual scouting, while the continuous-learning pipeline progressively improves domain adaptation to local crop varieties. Second, machine-learning-driven irrigation scheduling informed by IoT telemetry and coordinated with upstream reservoir conditions achieves substantial water savings relative to both farmer-led and fixed-threshold baselines without compromising crop yield, confirming the hypothesis that rule-based thresholds leave significant efficiency on the table.

Third, deep-learning sequence models, particularly LSTM architectures trained with probabilistic output layers, outperform classical ARIMA and Prophet baselines for reservoir inflow and storage forecasting at all practically relevant horizons, and the probabilistic formulation provides operators with calibrated risk information previously unavailable in Sri Lankan reservoir practice. Fourth, the integration of FAO-56 water budgets, fuzzy-TOPSIS crop suitability, tabular yield forecasting, and short-horizon price forecasting into a quota-aware linear or mixed-integer Optimization produces prescriptive hectare plans that are simultaneously feasible, profitable, and water-efficient, and the mid-season Plan-B capability demonstrates that the system can respond to allocation and market shocks within operationally acceptable timeframes. Fifth, the microservices architecture, event-driven inter-function communication, and disciplined MLOps practices collectively enable the four functions to operate as a coherent decision-support ecosystem while remaining independently developable and deployable, validating the architectural choices articulated in Chapter 2.

A cross-cutting finding concerns the importance of explainability. Across all four functions, stakeholders consistently report higher trust and willingness to act on AI outputs that are accompanied by human-readable rationales, confidence scores, or uncertainty bands than on equivalent outputs presented as deterministic recommendations. This finding reinforces the design decision to priorities explainability alongside accuracy and has implications for the broader rollout of AI-driven decision-support systems in Sri Lankan agriculture.

3.4 Discussion

3.4.1 Interpretation of Results Against Research Objectives

The results presented in this chapter provide substantial evidence that the research objectives defined in Section 1.4 have been met. The main objective of designing, developing, and validating an integrated, AI- and IoT-enabled Smart Water Management System is demonstrably achieved through the operational pilot in Uda Walawe and Anuradhapura, and each of the four specific objectives is supported by function-level evidence. The satellite-based crop health monitoring objective is supported by the demonstrated four-to-seven-day detection lead-time improvement, the F1-score above 0.85 on the held-out CNN test set, and the operational dashboard adoption by irrigation officers. The ML-driven smart irrigation scheduling objective is supported by the thirty-to-fifty-per-cent water-savings range and the successful field-to-dam coordination demonstrated in integration tests. The reservoir forecasting objective is supported by MAPE values within or below the fifteen-per-cent target at the one-to-three-day horizon, the calibrated probabilistic intervals, and the positive stakeholder reception of the what-if scenario interface. The ACA-O objective is supported by the near-one-hundred-per-cent feasibility rates, measurable profit-per-hectare uplift, and Plan-B Re-Optimization within SLO.

3.4.2 Comparison with Existing Work

Compared with the existing body of Sri Lankan and international literature reviewed in Chapter 1, the present work advances the state of the art in several respects. Relative to the pilot IoT-based irrigation systems previously reported in Sri Lanka, which are predominantly rule-based and operate in isolation from upstream reservoir conditions, the present system introduces machine-learning-driven scheduling with explicit field-to-dam coordination, addressing the disjunction that the earlier literature identified but did not resolve. Relative to international crop-health monitoring systems that rely exclusively on satellite-derived indices, the present system introduces a ground-verification loop through farmer-submitted imagery that improves the reliability and contextual validity of detections under tropical Sri Lankan conditions. Relative to reservoir-forecasting studies that deliver only point forecasts, the present system's probabilistic outputs provide calibrated uncertainty information that aligns with the risk-based nature of actual reservoir-operation decisions. Relative to crop-suitability studies that stop at descriptive ranking, the present system integrates fuzzy-TOPSIS with quota-aware Optimization, producing prescriptive hectare plans rather than merely indicative scores. The microservices-based integration of all four functions on a shared event-driven backbone is, to the best of our knowledge, the first such platform developed for Sri Lankan reservoir-fed agriculture.

3.4.3 Limitations

Several limitations of the present work should be acknowledged. First, the pilot evaluation is geographically limited to Uda Walawe and Anuradhapura and cannot be directly extrapolated to Sri Lanka's other agro-ecological regions without further calibration and validation. Second, the historical dataset for the reservoir-forecasting function, while extensive, contains periods of missing data and instrument drift that introduce residual uncertainty into the model training and evaluation. Third, the CNN ensemble's performance on crop classes that are under-represented in the pilot image collection (such as certain vegetable varieties) is weaker than its performance on paddy and maize, and continued data collection through the active-learning loop is required to close this gap. Fourth, the price-forecasting component is sensitive to macroeconomic shocks that are outside the scope of the data the Temporal Fusion Transformer has been trained on, such as sudden fuel-cost spikes or currency devaluations, and may produce over-confident forecasts during such periods. Fifth, the evaluation of farmer adoption and long-term behavioral change requires a longer observation window than has been available at the time of this report, and remains a subject for future work.

3.4.4 Implications for Sri Lankan Water and Agricultural Policy

The results have several implications for Sri Lankan water and agricultural policy. The demonstrated water savings and yield improvements suggest that deployment of the platform at scale could contribute materially to national objectives for climate-smart agriculture and water security, as articulated in the FAO-supported Climate-Smart Agriculture Investment Plan. The probabilistic forecasting capabilities offer the Irrigation Department and Water Management Secretariat a tool for moving from reactive to proactive reservoir management, with corresponding reductions in flood and drought damages. The hectare-Optimization outputs provide policy makers with a data-driven instrument for steering cropping patterns toward water-productive configurations while respecting farmer livelihoods and policy constraints. At the same time, the emphasis on explainability, multilingual access, and farmer-in-the-loop control reflects a design philosophy that prioritizes adoption and trust over purely algorithmic Optimization, and may offer a useful template for subsequent AI-driven public-sector digitalization initiatives in Sri Lanka.

3.4.5 Future Work

Several directions for future work emerge from the present results. Expanding the geographic scope beyond Uda Walawe and Anuradhapura to cover the Mahaweli system, the Northern Province tank cascades, and the Uva plantation districts would test the platform's generalizability across Sri Lanka's agro-ecological diversity. Incorporating additional sensor modalities, such as leaf-wetness sensors, canopy-temperature thermal imaging from drone overflights, and downscaled climate projections from regional climate models, would enrich the feature space for the ML models. Integration with index-based crop insurance products, supported by the validated yield forecasts and stress-detection records, could accelerate the adoption of climate-risk financial instruments in rural Sri Lanka. Finally, extending the platform to support groundwater management in addition to surface-reservoir management would broaden its applicability to the dry-zone regions that rely heavily on agro-well irrigation.

3.5 Summary of Each Student's Contribution

This dissertation represents the combined research effort of four students, each of whom has led the development of one of the four functions described above while contributing to the integration, evaluation, and reporting of the overall platform. The following subsections summaries each student's individual contribution, aligning it with the specific objectives defined in Section 1.4.2.

3.5.1 Student IT22186942 - Satellite-Based Crop Health Monitoring

This student led the design, development, and evaluation of Function 1, the satellite-based hybrid crop health monitoring subsystem. The contribution encompasses the end-to-end pipeline from raw remote-sensing data acquisition through ground-truth validation to operational decision support. Specifically, the student designed and implemented the automated Sentinel-2 Level-2A ingestion pipeline using the Google Earth Engine Python API, including cloud-masking, atmospheric-correction, and vegetation-index computation scripts for NDVI, NDWI, MSI, and EVI, with temporal interpolation and cross-validation logic leveraging MODIS and Landsat-8/9 as supplementary sources and Sentinel-1 SAR as a cloud-penetrating fallback. The student curated the ground-truth image dataset through a dedicated field-collection campaign, labelled the dataset in collaboration with Department of Agriculture agronomists, and implemented the data-augmentation and train/validation/test partitioning protocols. The student then trained, tuned, and comparatively evaluated MobileNetV3, ResNet50, EfficientNet-B3, and custom CNN architectures in TensorFlow 2.12 and PyTorch 2.0, designing the ensemble-averaging strategy that ultimately delivered the best operational performance.

Beyond the model work, this student designed and built the Flutter-based farmer mobile application, including the geo-tagged image-capture workflow, offline caching, multilingual support in Sinhala, Tamil, and English, and the notification subsystem that delivers caution-zone alerts. The student also developed the React and Mapbox-GL-JS dashboard for irrigation officers, implementing the satellite-index visualization layers, the field-level drill-down views, and the AI-output validation tools through which officers confirm or correct the ensemble's classifications. Finally, the student designed the continuous-learning feedback loop through which validated farmer uploads periodically retrain the CNN ensemble, implemented the associated MLflow experiment-tracking and model-registry workflows, and conducted the cross-dataset verification against MODIS indices and ground meteorological data that established the scientific credibility of the system's outputs.

3.5.2 Student IT22561398 - ML-Driven Smart Irrigation Scheduling

This student led the design, development, and evaluation of Function 2, the IoT-enabled and machine-learning-driven smart irrigation scheduling subsystem. The contribution spans sensor hardware, wireless communication, edge computing, machine learning, and actuator control. Specifically, the student selected and procured the FDR and TDR soil-moisture probes, capacitive temperature-humidity sensors, ultrasonic reservoir-level rangefinders, pH, and electrical-conductivity sensors, and personally calibrated each unit against laboratory-grade references during the commissioning phase. The student designed the edge-gateway architecture using ESP32 microcontrollers and Raspberry Pi 4 modules, implemented the three-bearer wireless uplink strategy using LoRaWAN, NB-IoT, and 4G LTE, and configured secure MQTT-over-TLS data transmission with X.509 client-certificate authentication.

On the analytics side, the student developed the feature-engineering pipeline that transforms raw sensor streams into irrigation-decision features, trained and tuned the online linear regression, decision-tree, and random-forest models, and benchmarked their performance against fixed-threshold and farmer-led baselines across complete cultivation cycles in the pilot plots. The student implemented the irrigation decision engine that synchronizes field-level demand with upstream reservoir conditions through subscriptions to Function 3's forecast topics, designed the microcontroller-based actuator logic for solenoid valves and pumps, and implemented the manual-override pathway through the farmer mobile application that preserves farmer agency. The student also developed the real-time irrigation dashboard that visualizes soil-moisture trajectories, cumulative water-usage analytics, per-field efficiency metrics, and anomaly alerts. Finally, the student conducted the controlled A/B trials in Uda Walawe that generated the water-savings and crop-yield evidence reported in Section 3.2.2 and produced the usability and farmer-feedback studies that informed subsequent interface refinements.

3.5.3 Student IT22076366 - Reservoir and Water-Inflow Forecasting

This student led the design, development, and evaluation of Function 3, the probabilistic reservoir and water-inflow forecasting subsystem for the Udawalawe Reservoir. The contribution covers the full data-to-decision pipeline from historical-record assembly through probabilistic modelling to operational dashboarding and alerting. Specifically, the student negotiated data-sharing arrangements with the Sri Lanka Department of Meteorology, the Irrigation Department, and the Ministry of Mahaweli Development and Environment, assembling the 1994–2025 historical dataset covering rainfall, reservoir inflow, storage, evaporation, irrigation releases, and hydropower discharges. The student implemented the preprocessing pipeline, including range checks, missing-value imputation, outlier filtering, stationarity testing through Augmented Dickey–Fuller statistics, and STL decomposition of seasonal components.

The student trained, tuned, and comparatively evaluated the full hierarchy of forecasting models described in Section 2.1.4: the ARIMA and SARIMA baselines, the Prophet model, the Random Forest and Gradient Boosting regressors, and the LSTM and GRU sequence-to-sequence networks with attention mechanisms. The student designed and implemented the probabilistic forecasting layer using both ensemble perturbation and quantile-regression LSTM approaches and validated the calibration of the resulting prediction intervals through PICP, Average Width, and reliability diagrams. The student implemented the risk-assessment module that computes spill probabilities, drought probabilities, and categorical flood-drought flags, and designed the operator-facing dashboard in React, including the reservoir-status panel, the what-if scenario-modelling interface, the forecast-error history view, and the multi-channel alerting subsystem delivering warnings through SMS, mobile push, and web banners. The student also conducted the historical-event back testing that demonstrated the three-day lead-time improvement during the November 2023 flood episode, and led the stakeholder-engagement workshops with the Irrigation Department and Water Management Secretariat that validated the operational relevance of the system.

3.5.4 Student IT22561770 - Adaptive Crop and Area Optimization (ACA-O)

This student led the design, development, and evaluation of Function 4, the adaptive crop and area Optimization subsystem. The contribution integrates hydrological, agronomic, and economic modelling with mathematical Optimization and modern MLOps practices. Specifically, the student designed the telemetry-ingestion pipeline that publishes soil-moisture, water-level, pH, and EC data over MQTT v5 into Apache Kafka topics, and implemented the associated stream-processing logic that normalizes and persists telemetry alongside price and allocation feeds. The student implemented the FAO-56 Kc–ETo water-budget computations that produce per-field ETc and deficit envelopes, calibrated against the Sri Lankan CROPWAT and CLIMWAT references.

The student designed and implemented the fuzzy-TOPSIS crop-suitability scoring module that produces Top-3 crop recommendations with human-readable rationales, trained and tuned the XGBoost and LightGBM yield-forecasting models with leave-canal-out and leave-season-out cross-validation, and trained the Temporal Fusion

Transformer and LSTM fallback price-forecasting models with rolling-origin evaluation and Diebold–Mariano statistical comparison. The student formulated the quota-aware linear and mixed-integer Optimization problem, including the water-envelope, pH/EC, rotation, and policy constraints, and implemented the solver interface using COIN-OR CBC. The student designed and implemented the mid-season Plan-B Re-Optimization capability, including the stress-test framework that simulates ten- and twenty-per-cent allocation cuts. The student also developed the FastAPI service exposing the /f4/recommendations, /f4/planB, and /f4/national-supply endpoints, containerized the services with Docker, orchestrated them on Kubernetes with horizontal pod autoscaling, and tracked all models in the MLflow Model Registry. Finally, the student led the definition of the platform's service-level objectives and error-budget policies following the Google SRE workbook and conducted the SUS usability studies with farmers and scheme managers.

3.5.5 Joint Contributions

In addition to the function-specific work described above, all four students contributed jointly to several cross-cutting aspects of the project. These include the definition of the shared event schemas on the Kafka bus through which the four functions communicate; the design of the shared feature store on PostgreSQL and PostGIS; the specification of the unified authentication, authorization, and audit-logging layer; the integration testing of cross-function flows such as stress-to-irrigation and reservoir-to-scheduler propagation; the assembly of the unified observability stack combining Fluent Bit, Elasticsearch, Prometheus, Grafana, and Jaeger; and the preparation of this dissertation including the literature review, the methodology write-up, the results synthesis, and the commercialization analysis. Regular cross-function design reviews and pair-programming sessions ensured that the four subsystems evolved in a consistent direction and that integration risks were identified and mitigated early rather than late in the development cycle.

4 CONCLUSION

4.1 Overview

This dissertation has presented the design, development, and pilot validation of an integrated Smart Water Management System for Sri Lanka, combining satellite-based crop health monitoring, machine-learning-driven smart irrigation scheduling, probabilistic reservoir and water-inflow forecasting, and adaptive crop and area Optimization into a single microservices-based decision-support platform. The research has been motivated by the convergence of pressing challenges confronting Sri Lankan agriculture, including climate-induced rainfall variability, the dominance of manual and intuition-driven irrigation practices, the disjunction between field-level water use and upstream reservoir operations, the reactive nature of reservoir-release decision-making, and the absence of prescriptive decision tools for crop-area planning under quota and market uncertainty. By addressing these challenges through a unified technical architecture grounded in modern AI, IoT, and operations-research methodologies, the project contributes a practical prototype for data-driven, climate-resilient water management in Sri Lanka's reservoir-fed agricultural ecosystems.

4.2 Summary of Research Contributions

The research delivers four principal contributions, each corresponding to one of the specific objectives articulated in Chapter 1 and validated through the results reported in Chapter 3. The first contribution is a hybrid satellite–ground crop health monitoring framework that combines Sentinel-2 multispectral imagery with convolutional neural network verification of farmer-submitted field photographs. This framework bridges the gap between remote-sensing-based early detection and ground-level validation, producing location-specific crop-stress alerts with detection lead-times that exceed those of manual scouting by several days. The continuous-learning loop, in which validated farmer uploads periodically retrain the CNN ensemble, ensures that the system progressively improves its domain adaptation to Sri Lankan crop varieties and agro-ecological conditions, and the multilingual Flutter application and Mapbox-based officer dashboard provide usable interfaces tailored to both smallholder farmers and agricultural extension officers.

The second contribution is an IoT-enabled, machine-learning-driven smart irrigation scheduling subsystem that moves beyond the rule-based thresholds characteristic of prior Sri Lankan deployments. By combining calibrated soil-moisture, temperature, humidity, and reservoir-level sensing with online regression, decision-tree, and random-forest models trained on Sri Lankan field data, the subsystem produces dynamic, crop-stage-aware irrigation decisions that achieve substantial water savings relative to farmer-led and threshold-based baselines while preserving or improving crop yield. Crucially, the subsystem implements the long-sought field-to-dam coordination by subscribing to reservoir-release forecasts and synchronizing field-level actuation with upstream conditions, a capability absent from the existing Sri Lankan smart-irrigation literature.

The third contribution is a probabilistic reservoir and water-inflow forecasting subsystem for the Udawalawe Reservoir that outperforms classical ARIMA, Prophet, and tree-based baselines at all practically relevant lead times, while providing calibrated uncertainty intervals suitable for risk-based operational decision-making. The combination of LSTM and GRU sequence-to-sequence models with ensemble perturbation and quantile-regression outputs addresses the "last-mile" integration gap identified in the reservoir-management literature, converting probabilistic forecasts into actionable alerts, what-if scenario simulations, and operator dashboards that have been co-designed with the Irrigation Department and Water Management Secretariat. Historical-event back testing demonstrates that the system would have provided approximately three days of additional warning during the November 2023 flood episode and meaningful advance notice of the September 2023 drought, illustrating the operational potential of the approach.

The fourth contribution is the Adaptive Crop and Area Optimization subsystem, which integrates FAO-56 water-budget calculations, fuzzy-TOPSIS crop suitability scoring, gradient-boosted tree yield forecasting, Temporal Fusion Transformer price forecasting, and quota-aware linear and mixed-integer Optimization into a single predict-then-optimize decision service. The subsystem produces Top 3 crop recommendations per field with human-readable rationales, hectare plans aggregated to farmer and canal levels, and water-requirement envelopes for the downstream scheduler, all respecting reservoir quotas, soil pH and electrical conductivity constraints, rotation rules, and policy bounds. The mid-season Plan-B capability enables the system to respond to allocation revisions and market shocks within operationally acceptable timeframes, addressing the mid-season resilience gap identified in the fuzzy multi-objective planning literature.

Beyond the four function-level contributions, the dissertation makes a systems-level contribution through the integration of these four functions on a shared, event-driven, microservices-based backbone. The combination of MQTT v5 telemetry ingestion, Kafka stream processing, containerized FastAPI and Node.js services, Kubernetes orchestration, MLflow-based model management, and a unified observability stack constitutes a reference architecture for AI- and IoT-driven water management platforms in developing-economy contexts. The consistent emphasis on explainability, multilingual accessibility, offline-capable farmer interfaces, and farmer-in-the-loop control reflects a design philosophy that prioritizes adoption and trust, and offers a template for subsequent AI-driven public-sector digitalization efforts in Sri Lanka.

4.3 Achievement of Research Objectives

Against the specific objectives defined in Chapter 1, the research has produced evidence of substantive achievement across all four functions. The satellite crop-health monitoring objective is supported by the operational Sentinel-2 ingestion pipeline, the ensemble CNN achieving an F1-score above 0.85 on the held-out test set, the measured four-to-seven-day detection lead-time improvement over manual scouting, and the successful deployment of the Flutter farmer application and the React officer dashboard. The ML-driven smart irrigation scheduling objective is supported by the calibrated sensor network, the three-bearer wireless communication infrastructure, the trained machine-learning models, the demonstrated thirty-to-fifty-per-cent water savings, and the validated field-to-dam coordination. The reservoir-forecasting objective is supported by the assembled 1994–2025 historical dataset, the hierarchical forecasting pipeline achieving MAPE within the fifteen-per-cent target at the one-to-three-day horizon, the calibrated probabilistic intervals, and the operator-facing dashboard with what-if scenario simulation. The ACA-O objective is supported by the FAO-56 water budgets, the fuzzy-TOPSIS rankings, the yield and price forecasts with statistically significant improvements over baselines, the near-one-hundred-per-cent feasibility rate of optimized hectare plans, and the stress-tested Plan-B Re-Optimization capability. Collectively, these outcomes demonstrate that the main research objective of delivering an integrated smart water management system for Sri Lanka has been substantively realized.

4.4 Limitations and Constraints

While the research outcomes are encouraging, honest acknowledgement of the limitations is essential. The pilot validation is geographically confined to Uda Walawe and Anuradhapura, and generalization to other Sri Lankan agro-ecological regions requires further calibration and validation. The reservoir-forecasting dataset, though extensive, contains periods of missing data and instrument drift that introduces residual uncertainty into model training. The CNN ensemble's performance is stronger on paddy and maize than on under-represented vegetable and minor-crop classes, and further data collection through the active-learning loop is required to close this gap. The price-forecasting component is sensitive to macroeconomic shocks outside its training distribution and may produce over-confident forecasts during fuel-cost or currency-volatility episodes. Long-term farmer-adoption outcomes and behavioral change require observation windows longer than those available at the time of reporting. Finally, the system's dependence on continued availability of free Sentinel, MODIS, and CHIRPS data products represents an external constraint, though one mitigated by the deliberate inclusion of multiple redundant data sources.

4.5 Future Work

Several directions for future work emerge from the present research. Geographic expansion beyond the initial pilots to cover the Mahaweli system, the Northern Province tank cascades, and the Uva plantation districts would rigorously test the platform's generalizability across Sri Lanka's agro-ecological diversity. Incorporating additional sensor modalities, including leaf-wetness sensors, thermal canopy imagery from drone overflights, and downscaled climate projections from regional climate models, would enrich the feature space available to the machine-learning models and enable higher-resolution decision support. Integration with index-based crop insurance products, underwritten by the validated yield forecasts and stress-detection records, could accelerate the adoption of climate-risk financial instruments in rural Sri Lanka and create a secondary revenue stream for the platform. Extending the system to cover groundwater and agro-well irrigation, in addition to surface-reservoir management, would broaden its applicability to the dry-zone regions where groundwater plays a critical role. Deeper integration with national digital-agriculture platforms operated by the Department of Agriculture, the Department of Agrarian Development, and the Ministry of Technology would position the platform as a building block of Sri Lanka's National Digital Agriculture Strategy. On the research side, further investigation into federated-learning approaches for CNN retraining, into transfer-learning strategies for regions with limited training data, and into physics-informed neural networks for reservoir forecasting would strengthen the technical foundations of the next iteration of the system.

4.6 Closing Remarks

The sustainable management of water in Sri Lanka's agricultural landscape is a problem of enduring importance, bridging concerns of food security, rural livelihoods, climate adaptation, and economic development. The hydraulic civilization that gave Sri Lanka its ancient tank cascades and anicuts demonstrated, two millennia ago, the capacity of the island's engineers and farmers to manage water with ingenuity and precision under the conditions of their time. The present research, though modest in scope relative to that civilizational legacy, aspires to contribute a comparable leap forward in its own era by harnessing artificial intelligence, the Internet of Things, satellite remote sensing, and modern Optimization methods to serve the same enduring objective: ensuring that water, the most fundamental agricultural input, is allocated wisely among fields, seasons, and communities.

The integrated platform documented in this dissertation is not a finished product but a prototype, an architectural template, and a set of validated techniques that together demonstrate the feasibility of a new generation of smart water management systems for Sri Lanka and for similarly situated agrarian economies across South Asia, Southeast Asia, and Africa. Realizing the full potential of the platform will require continued collaboration among the research community, the Irrigation Department, the

Department of Agriculture, the Meteorology Department, development partners, and above all the farming communities whose knowledge, participation, and trust are indispensable. The research presented here is offered in the hope of advancing that collaboration, and of demonstrating that rigorous academic inquiry, when conducted in close partnership with the stakeholders whose lives are most affected by water-management decisions, can produce technological outcomes that are at once scientifically defensible, operationally useful, and socially just.

REFERENCES

[1] Central Bank of Sri Lanka, Annual Report 2023: Agricultural Sector Performance. Colombo, Sri Lanka: Central Bank of Sri Lanka, 2023. [Online]. Available: https://www.cbsl.gov.lk/en/publications/economic-reports/annual-report-2023

[2] Department of Agriculture, Sri Lanka, "Sri Lanka Irrigation Systems Overview," Research and Development Institute, 2021. [Online]. Available: https://doa.gov.lk/rrdi_watermanagement_irrigationsystems/

[3] International Water Management Institute (IWMI), "Water Resources and Irrigation in the Walawe Basin," IWMI Working Paper Series, Colombo, Sri Lanka, 2020.

[4] Food and Agriculture Organization, Digital Agriculture Profile: Sri Lanka. Rome, Italy: FAO, 2022. [Online]. Available: https://www.fao.org/3/cc3176en/cc3176en.pdf

[5] Hector Kobbekaduwa Agrarian Research and Training Institute (HARTI), "Study on Udawalawa Irrigation Scheme: Performance and Challenges," HARTI Research Report No. 176, Colombo, Sri Lanka, 2021. [Online]. Available: https://www.harti.gov.lk/images/download/reasearch_report/new1/176.pdf

[6] Ministry of Mahaweli Development and Environment (Sri Lanka), Water Management Secretariat Annual Report, Colombo, Sri Lanka, 2024.

[7] Department of Agriculture, Sri Lanka, Annual Report on Crop Losses and Disease Management. Peradeniya, Sri Lanka: Government Publications, 2023. [Online]. Available: https://www.doa.gov.lk/publications/AR2023_CropLosses.pdf

[8] World Bank, South Asia Climate Adaptation Report. Washington, DC, USA: World Bank, 2023. [Online]. Available: https://www.worldbank.org/

[9] S. Samarasinghe et al., "Granger Causality-Based Forecasting Model for Rainfall Prediction in Sri Lanka," Forecasting, vol. 6, no. 4, pp. 1–18, 2024.

[10] Department of Meteorology, Sri Lanka, "Disaster Situation Report and Maps – November 2023 Floods," National Disaster Relief Services Centre (NDRSC), Official Release, Nov. 2023.

[11] International Water Management Institute (IWMI), "Remote Sensing Applications in Sri Lankan Agriculture: Case Studies from Uda Walawe," IWMI Working Paper 202, Colombo, Sri Lanka, 2023. [Online]. Available: https://www.iwmi.org/publications/working-papers/

[12] Hector Kobbekaduwa Agrarian Research and Training Institute (HARTI), Weekly Food Commodities Bulletin and Retail Price Tables, Colombo, Sri Lanka, 2024.

[13] European Space Agency, Sentinel-2 User Handbook. ESA Earth Observation Publications, 2023. [Online]. Available: https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/document-library

[14] Y. Zhang et al., "Machine Learning for Crop Disease Detection from Multispectral Data," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 1234–1245, 2023.

[15] International Water Management Institute (IWMI), "Vegetation Indices for Agricultural Monitoring in South Asia," IWMI Research Report 185, Colombo, Sri Lanka, 2023. [Online]. Available: https://www.iwmi.org/publications/iwmi-research-reports/

[16] International Research Journal of Modernization in Engineering, Technology and Science (IRJMETS), "IoT-Based Smart Irrigation System for Sri Lankan Agriculture," IRJMETS, 2022.

[17] T. Mallikarathne, "Wireless Sensor Network Based Smart Irrigation System for Sri Lankan Crops," in Proc. International Conference on Advancements in Research Computing (ICARC), 2024, doi: 10.1109/ICARC.2024.10408962.

[18] OASIS, MQTT Version 5.0 Specification, 2019. [Online]. Available: https://docs.oasis-open.org/mqtt/mqtt/v5.0/mqtt-v5.0.html

[19] Apache Software Foundation, "Kafka Documentation – Replication and Operations," 2025. [Online]. Available: https://kafka.apache.org/documentation/

[20] A. Howard et al., "Searching for MobileNetV3," arXiv preprint arXiv:1905.02244, 2019. [Online]. Available: https://arxiv.org/abs/1905.02244

[21] M. Rahman and T. Zhou, "Hybrid AI Frameworks for Integrating Satellite and Field Imagery in Tropical Agriculture," Computers and Electronics in Agriculture, vol. 218, art. 109002, 2024, doi: 10.1016/j.compag.2024.109002.

[22] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), San Francisco, CA, USA, 2016, pp. 785–794. [Online]. Available: https://arxiv.org/abs/1603.02754

[23] G. Ke et al., "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in Proc. 31st Conference on Neural Information Processing Systems (NeurIPS 2017), Long Beach, CA, USA, 2017.

[24] G. Özdoğan-Sarıkoç, M. Sarıkoç, M. Celik, and F. Dadaser-Celik, "Reservoir Volume Forecasting Using Artificial Intelligence-Based Models: Artificial Neural Networks, Support Vector Regression, and Long Short-Term Memory," Journal of Hydrology, vol. 616, art. 128766, 2022.

[25] M. Herath, "Deep Machine Learning-Based Water Level Prediction Model for Detention Areas in Sri Lanka," Applied Sciences, vol. 13, no. 4, 2023.

[26] B. Lim, S. Ö. Arık, N. Loeff, and T. Pfister, "Temporal Fusion Transformers for Interpretable Multi-Horizon Time Series Forecasting," International Journal of Forecasting, vol. 37, no. 4, pp. 1748–1764, 2021.

[27] R. G. Allen, L. S. Pereira, D. Raes, and M. Smith, Crop Evapotranspiration — Guidelines for Computing Crop Water Requirements, FAO Irrigation and Drainage Paper No. 56. Rome, Italy: Food and Agriculture Organization of the United Nations, 1998.

[28] C.-T. Chen, "Extensions of the TOPSIS for Group Decision-Making under Fuzzy Environment," Fuzzy Sets and Systems, vol. 114, no. 1, pp. 1–9, 2000.

[29] H. P. Williams, Model Building in Mathematical Programming, 5th ed. Chichester, UK: John Wiley & Sons, 2013.

[30] T. Jayathilake, "Wetland Water Level Prediction Using Artificial Neural Networks in the Muthurajawela Marsh," Climate, vol. 11, no. 1, 2023.

[31] M. Herath, "Deep Machine Learning-Based Water Level Prediction for Detention Areas in Sri Lanka," Applied Sciences, vol. 13, no. 4, 2023.

[32] S. Samarasinghe et al., "Granger Causality-Based Forecasting Model for Rainfall Prediction in Sri Lanka," Forecasting, vol. 6, no. 4, 2024.

[33] Hector Kobbekaduwa Agrarian Research and Training Institute (HARTI), "Digital Agriculture Adoption Barriers in Sri Lanka," HARTI Research Study No. 45, Colombo, Sri Lanka, 2023.

[34] Department of Census and Statistics, Sri Lanka, ICT Adoption in Sri Lankan Agriculture, Statistical Bulletin, Colombo, Sri Lanka, 2023.

[35] Food and Agriculture Organization, Climate-Smart Agriculture Investment Plan for Sri Lanka, FAO Technical Cooperation Programme, Rome, Italy, 2023. [Online]. Available: https://www.fao.org/3/cc4567en/cc4567en.pdf

[36] World Bank, Transforming Agriculture with Digital Technologies — Agriculture and Food Global Report. Washington, DC, USA: World Bank, 2021.

[37] Asian Development Bank, Irrigation Modernization and Technology Adoption in Sri Lanka, ADB Project Completion Report, Manila, Philippines, 2023.

[38] G. K. Nandasena, "Controlling Sedimentation through Regulating the River," Proceedings of the International Association of Hydrological Sciences, vol. 386, pp. 1–8, 2024.

[39] FastAPI, "Official Documentation," 2025. [Online]. Available: https://fastapi.tiangolo.com/

[40] Cloud Native Computing Foundation, "Kubernetes Documentation — Deployments, Rolling Updates, and Horizontal Pod Autoscaling," 2025. [Online]. Available: https://kubernetes.io/docs/

[41] ONNX Runtime, "Documentation and Get Started," 2025. [Online]. Available: https://onnxruntime.ai/docs/

[42] L. Breiman, "Random Forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.

[43] C. Funk et al., "The Climate Hazards Infrared Precipitation with Stations — A New Environmental Record for Monitoring Extremes," Scientific Data, vol. 2, art. 150066, 2015.

[44] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, 5th ed. Hoboken, NJ, USA: Wiley, 2015.

[45] S. J. Taylor and B. Letham, "Forecasting at Scale," The American Statistician, vol. 72, no. 1, pp. 37–45, 2018. [Online]. Available: https://facebook.github.io/prophet/

[46] Z. Ou, C. Nai, B. Pan, Y. Zheng, C. Shen, P. Jiang, et al., "Probabilistic Diffusion Models Advance Extreme Flood Forecasting," Geophysical Research Letters, vol. 52, no. 15, art. e2025GL115705, 2025.

[47] ISO/IEC, ISO/IEC 25010:2011 Systems and Software Engineering — Systems and Software Quality Requirements and Evaluation (SQuaRE) — System and Software Quality Models, International Organization for Standardization, Geneva, Switzerland, 2011. [Online]. Available: https://www.iso.org/standard/35733.html

[48] OWASP Foundation, Application Security Verification Standard (ASVS) v4.0.3, 2019. [Online]. Available: https://github.com/OWASP/ASVS/

[49] B. Beyer, C. Jones, J. Petoff, and N. R. Murphy, Eds., Site Reliability Engineering: How Google Runs Production Systems and The Site Reliability Workbook. Sebastopol, CA, USA: O'Reilly Media, 2017–2018. [Online]. Available: https://sre.google/workbook/implementing-slos/

[50] J. Brooke, "SUS — A Quick and Dirty Usability Scale," in Usability Evaluation in Industry, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and I. L. McClelland, Eds. London, UK: Taylor & Francis, 1996, pp. 189–194.

[51] F. X. Diebold and R. S. Mariano, "Comparing Predictive Accuracy," Journal of Business & Economic Statistics, vol. 13, no. 3, pp. 253–263, 1995.

[52] R. J. Hyndman and G. Athanasopoulos, Forecasting: Principles and Practice, 3rd ed. Melbourne, Australia: OTexts, 2021. [Online]. Available: https://otexts.com/fpp3/

[53] D. Kawshalya and H. De Silva, "CNN-Assisted Verification of Sentinel-2 Vegetation Anomalies in Sri Lankan Paddy Cultivation," Sri Lankan Journal of Computing, vol. 11, no. 2, 2024.

[54] Y. Li and H. Zhou, "Deep Learning–Based Multispectral Data Fusion for Crop Stress Detection," Remote Sensing of Environment, vol. 298, art. 113778, 2023.

[55] Ministry of Technology, Sri Lanka, National Digital Agriculture Strategy 2023–2027, Government Policy Document, Colombo, Sri Lanka, 2023.

[56] Department of Agrarian Development, Sri Lanka, Farmer Economics and Technology Adoption Study, Colombo, Sri Lanka, 2023.