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

(Adaptive Crop & Area Optimization (ACA-O))

Project ID: 25-26J-520

Dilruksha A.G.C.D.

(Hesara P.K.A.N., Trishni W.R.M., Abishek S)

B.Sc. (Hons) in Information Technology Specializing in Software Engineering

Department of Software Engineering

Sri Lanka Institute of Information Technology Malabe, Sri Lanka

April 2026

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

(Adaptive Crop & Area Optimization (ACA-O))

Project ID: 25-26J-520

Dilruksha A.G.C.D. - IT22561770

Hesara P.K.A.N. - IT22561398

Trishni W.R.M. - IT22076366

Abishek S - IT22186942

Supervisor: Ms. Hansi De Silva

Co-Supervisor: Ms. Karthiga Rajendran

B.Sc. (Hons) in Information Technology Specializing in Software Engineering

Department of Software Engineering

Sri Lanka Institute of Information Technology Malabe, Sri Lanka

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	

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

Irrigated canal-command agriculture in Sri Lanka - exemplified by the Udawalawe Right Bank Main Canal (RBMC) and Left Bank Main Canal (LBMC) scheme — faces a persistent planning crisis characterised by the "paddy-by-default" phenomenon, where farmers systematically choose a single high-water-demand crop regardless of prevailing soil conditions, seasonal water quotas, or market price trajectories. The consequence is structurally low water productivity measured in $kg \cdot m^{-3}$, reduced farm profit per hectare, and systemic over-allocation that exacerbates reservoir stress across Maha (September–March) and Yala (May–August) seasons.

This research proposes the design, implementation, and evaluation of Function 4 — Adaptive Crop & Area Optimization (ACA-O), a Deep-Learning-assisted, quota-aware crop recommendation and hectare-allocation engine conceived as the prescriptive decision layer within the broader Integrated Smart Water-Focused Irrigation System (Project ID: 25-26J-520). ACA-O addresses the fundamental question of what to plant, where to plant it, and how many hectares should be allocated per crop, all subject to real-time water quotas, soil suitability constraints, market expectations, and policy bounds.

The ACA-O pipeline operates across five tightly coupled layers. First, a telemetry ingestion layer collects IoT field sensor readings — soil moisture via METER TEROS-10 probes, water level via JSN-SR04T ultrasonic modules, and soil pH/EC via Atlas Scientific circuits — and streams them through MQTT v5 into Apache Kafka topics alongside hydro-climate data, FAO CROPWAT/CLIMWAT coefficients, Sentinel-2 and MODIS imagery, HARTI market price bulletins, and GeoGoviya parcel registries. Second, a feature-engineering layer applies FAO-56 Kc–ETo procedures to compute field-resolved crop evapotranspiration (ETc) and water deficits, persisting outputs to a Feast feature store backed by TimescaleDB hypertables. Third, a predictive model layer executes four complementary model families in parallel: fuzzy-TOPSIS multi-criteria decision-making for explainable crop suitability scoring, XGBoost and LightGBM gradient-boosted tree models for tabular yield prediction, Temporal Fusion Transformers (TFT) with LSTM fallback for short-horizon price forecasting, and Monte Carlo simulation for probabilistic profit estimation. Fourth, an optimisation layer solves a linear or mixed-integer programme to maximise expected profit per hectare while enforcing scheme-level quotas, soil/pH/EC feasibility, rotation rules, and policy bounds. Fifth, a FastAPI service layer exposes three principal REST endpoints — /f4/recommendations, /f4/planB, and /f4/national-supply — publishing cryptographically signed snapshots to the Irrigation Department dashboard and serving a bilingual (Sinhala/Tamil/English) Progressive Web App for farmers.

Expected outcomes include measurable uplift in profit per hectare versus paddy-by-default baselines, increased kg·m^{-3} water productivity, timely mid-season Plan-B re-optimisation within P95 $\leq$ 5 seconds, positive farmer System Usability Scale (SUS) scores, and availability $\geq$ 99.5% per month. Services are containerised with Docker, orchestrated on Kubernetes with rolling updates, tracked in MLflow Model Registry, and secured per OWASP ASVS Level 2 and NIST SP 800-53 r5 with full PDPA No. 9 of 2022 compliance.

Keywords: *smart irrigation, crop recommendation, area optimization, fuzzy-TOPSIS, XGBoost, Temporal Fusion Transformer, IoT, water quota, Udawalawe, prescriptive analytics, machine learning, MQTT, Apache Kafka, FastAPI, Kubernetes*

ACKNOWLEDGEMENT

First and foremost, I wish to express my deepest gratitude to my supervisor, Ms. Hansi De Silva, for her unwavering guidance, intellectual rigour, and consistent encouragement throughout the duration of this research. Her constructive feedback and patient mentorship have been instrumental in shaping the direction and quality of this work. I am equally grateful to my co-supervisor, Ms. Karthiga Rajendran, whose expert insights into machine learning and agricultural technology have greatly enriched both the theoretical foundations and practical implementation strategies of the Adaptive Crop and Area Optimization module.

I extend sincere thanks to the academic staff and faculty of the Department of Software Engineering at the Sri Lanka Institute of Information Technology for providing access to institutional resources, research infrastructure, and a stimulating scholarly environment. Their collective commitment to academic excellence has been a continual source of motivation.

I am deeply indebted to my group members — Hesara P.K.A.N. (IT22561398), Trishni W.R.M. (IT22076366), and Abishek S (IT22186942) for their collaborative spirit, constructive discussions, and shared commitment to delivering an integrated system that is greater than the sum of its parts. The cross-functional integration between ACA-O and the other modules of Project 25-26J-520 has benefited immensely from our collective teamwork.

I would also like to acknowledge the broader research community whose published works in the domains of IoT-enabled smart agriculture, deep learning for forecasting, multi-criteria decision-making, and constrained optimisation have provided the theoretical scaffolding upon which this research is built.

Finally, my heartfelt thanks go to my family for their unconditional support, patience, and encouragement throughout this demanding academic journey. Their belief in my abilities has been my greatest source of strength.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Description
ACA-O	Adaptive Crop and Area Optimization
AI	Artificial Intelligence
API	Application Programming Interface
ASVS	Application Security Verification Standard
CI/CD	Continuous Integration / Continuous Deployment
DCS	Department of Census & Statistics (Sri Lanka)
DL	Deep Learning
EC	Electrical Conductivity
ETc	Crop Evapotranspiration
ETo	Reference Evapotranspiration
FAO	Food and Agriculture Organization
HARTI	Hector Kobbekaduwa Agrarian Research and Training Institute
IoT	Internet of Things
LBMC	Left Bank Main Canal
LGBM / LightGBM	Light Gradient Boosting Machine
LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MCDM	Multi-Criteria Decision-Making
MIP	Mixed-Integer Programme
ML	Machine Learning
MLOps	Machine Learning Operations
MODIS	Moderate Resolution Imaging Spectroradiometer
MQTT	Message Queuing Telemetry Transport
NDVI	Normalized Difference Vegetation Index
NIST	National Institute of Standards and Technology
OWASP	Open Web Application Security Project
PDPA	Personal Data Protection Act (Sri Lanka, No. 9 of 2022)
PWA	Progressive Web Application
RBMC	Right Bank Main Canal
RMSE	Root Mean Square Error

SLA	Service Level Agreement
SLO	Service Level Objective
SLIIT	Sri Lanka Institute of Information Technology
SUS	System Usability Scale
TFT	Temporal Fusion Transformer
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
UAT	User Acceptance Testing
VWC	Volumetric Water Content
WCAG	Web Content Accessibility Guidelines
XGBoost	Extreme Gradient Boosting

1. CHAPTER 1: INTRODUCTION

Irrigated agriculture sustains the food security and economic livelihoods of millions of smallholder farmers across South and Southeast Asia. In Sri Lanka, large surface-water schemes such as Udawalawe - comprising the Right Bank Main Canal (RBMC) and Left Bank Main Canal (LBMC) command areas - represent a critical component of national food production, serving thousands of farming households across Maha (September-March) and Yala (May-August) seasons. Yet the planning intelligence embedded in these schemes remains severely constrained. Crop choice and area allocation decisions are largely driven by tradition, social norms, and heuristic rules rather than by systematic integration of hydrologic feasibility, agronomic suitability, and market economics.

The consequence is a structural inefficiency commonly known as the "paddy-by-default" problem: farmers overwhelmingly choose paddy rice even when seasonal water allocations, soil profiles, or market price signals would favour alternative high-value crops such as maize, cowpea, green gram, or vegetables. This monocultural bias leads to chronic over-extraction from the Udawalawe reservoir, suppressed water productivity (kg of crop per cubic metre of water), and missed economic opportunities that could otherwise improve farmer incomes and scheme viability.

At the same time, the broader agricultural technology landscape has witnessed a transformative decade of progress. The convergence of low-cost IoT sensor networks, high-resolution remote sensing, open agro-climate datasets, and advances in machine learning and constrained optimisation has created, for the first time, the technical conditions under which a fully automated, explainable, and quota-aware decision support system for crop and area planning is feasible and economically justified. Sensor dashboards can now relay real-time soil moisture, pH, and electrical conductivity directly from fields to cloud platforms. Satellite products from Sentinel-2 and MODIS provide near-daily vegetation indices. Market price feeds from HARTI and the Department of Census and Statistics expose short-horizon price dynamics. Reservoir and canal allocation records can be ingested programmatically. The missing capability is not data - it is the intelligent engine that fuses these heterogeneous signals into prescriptive, quota-feasible, and explainable recommendations.

This research addresses that gap by designing, implementing, and evaluating Function 4 - Adaptive Crop and Area Optimization (ACA-O) - as a core module of Project 25-26J-520: Integrated Smart Water-Focused Irrigation System Using IoT and AI/ML, developed at the Sri Lanka Institute of Information Technology (SLIIT). ACA-O is conceived explicitly as the prescriptive layer of the integrated system, working in concert with Function 1 (automated field-level irrigation control), Function 2 (reservoir-level hydrological forecasting), and Function 3 (satellite-based crop health monitoring) to deliver a comprehensive decision support ecosystem.

1.1 Background Study and Literature Review

1.1.1 Background Study

The Udawalawe irrigation scheme in southern Sri Lanka is a flagship post-independence infrastructure investment. Straddling the Uda Walawe National Park and serving command areas on both the RBMC and LBMC sides, the scheme draws water from the Udawalawe Reservoir on the Walawe Ganga. The scheme supports diverse soils - ranging from alluvial loams in low-lying areas to reddish-brown earths on elevated blocks - and seasonal rainfall patterns that vary considerably between monsoon periods.

Cropping in Udawalawe follows a classic South Asian pattern: paddy rice dominates both Maha and Yala seasons, with minor oilseed and vegetable cultivation confined to tank-fed plots or upland margins. The dominance of paddy is reinforced by several mutually reinforcing factors. First, paddy production is deeply culturally embedded in Sinhala agrarian identity, conferring social status and food security assurance. Second, the government guaranteed purchase price for paddy reduces perceived market risk. Third, the information asymmetry between farmers and market price signals means that the potential economic advantage of high-value alternatives is poorly appreciated. Fourth, scheme-level water allocation practices historically favoured paddy scheduling, creating path dependencies that are difficult to disrupt without external intervention.

The agronomic consequences are severe. Paddy has among the highest seasonal water requirements of any Sri Lankan crop, typically demanding 1,200–1,500 mm per season under Udawalawe conditions. By contrast, crops such as maize, cowpea, and green gram can produce commercially competitive yields with 400–600 mm of applied water - a two to three times improvement in $kg \cdot m^{-3}$ productivity. The economic consequences compound these agronomic inefficiencies: paddy profit margins per hectare are structurally compressed by rising input costs (fertiliser, labour, machinery) and stagnant guaranteed prices, while specialty vegetables and pulses can achieve three to five times higher net returns per hectare in normal market conditions.

The planning challenge is fundamentally multi-objective and multi-constrained. A scheme manager must simultaneously optimise water productivity across the canal command, ensure equitable distribution of canal flows across all farmer blocks, comply with seasonal allocation quotas set by the Irrigation Department, respect biophysical thresholds for soil pH and EC that constrain crop suitability, and maintain rotation sequences that protect soil health and suppress pest pressure. A farmer must navigate these same constraints while also managing market price uncertainty, input cost volatility, labour availability, and household food security preferences. No current planning tool in the Udawalawe scheme integrates all of these dimensions.

The global smart agriculture movement, accelerated by the falling cost of IoT hardware and the proliferation of cloud-native machine learning platforms, has produced a rich body of technical capability that can be adapted to this setting. Low-cost soil moisture sensors, ultrasonic water level probes, and precision pH/EC instruments can now be deployed at field scale for under USD 200 per installation. Apache Kafka and MQTT v5 protocols enable durable, fault-tolerant ingestion of high-frequency telemetry at scheme scale. Feature stores such as Feast, combined with time-series databases such as TimescaleDB, provide the analytical infrastructure required to process, version, and serve engineered features at low latency. Machine learning frameworks including XGBoost, LightGBM, and PyTorch support rapid development of tabular and sequential forecasting models validated against rolling-origin baselines. Linear and mixed-integer programming solvers are available in open-source packages (PuLP, OR-Tools, Gurobi academic) with sufficient speed for real-time optimisation at scheme scale.

The research context for ACA-O is therefore defined by a clear and actionable problem - the inefficiency of paddy-by-default cropping in Udawalawe - and a clear and feasible technical pathway - a predict-then-optimise pipeline that fuses heterogeneous data into quota-aware, explainable recommendations. The challenge lies in integrating these components into a coherent, scalable, and trustworthy operational service that farmers and scheme managers will actually adopt.

1.1.2 Literature Review

The literature relevant to ACA-O spans four inter-related research strands: (i) fuzzy and multi-criteria decision-making for crop selection, (ii) machine learning for cropping pattern prediction and yield forecasting, (iii) remote sensing for crop phenology and area mapping, and (iv) mathematical optimisation for water resource and cropping pattern management. This review synthesises contributions across these strands and identifies the synthesis gap that ACA-O addresses.

Fuzzy-TOPSIS and Multi-Criteria Decision-Making for Crop Selection

Gowtham and R. B. R. [3] demonstrated that fuzzy-TOPSIS can be applied effectively to crop selection problems by scoring candidate crops against weighted criteria including soil suitability, water availability, temperature window, and land type. The method computes a closeness coefficient to the ideal solution, yielding a transparent and sensitivity-testable ranking that can be explained to farmers and planners. The authors showed that fuzzy sets handle the inherently linguistic and uncertain nature of agronomic judgements - for example, a soil pH of 6.2 is "moderately suitable" rather than binary-feasible - and that the resulting rankings correlate well with expert recommendations in their study region. However, their framework stops at ranking and does not proceed to allocate hectares across fields or enforce water quotas, leaving a critical gap between crop shortlisting and operational area planning.

Neamatollahi et al. [2] advanced the fuzzy optimisation literature by formulating agricultural cropping pattern determination as a fuzzy multi-objective linear programme. Their model simultaneously maximised profit and minimised water use, subject to land and resource constraints, producing area allocations that explicitly balance competing objectives. The use of fuzzy goal programming allowed soft constraints - for example, minimum subsistence acreage for staple crops - to be handled in a principled way. Their results showed material shifts in crop area shares under different water scarcity scenarios. A key limitation, however, was the reliance on static, seasonal-average inputs for water demand and price, rather than real-time IoT telemetry or short-horizon price forecasts, limiting responsiveness to mid-season shocks.

Machine Learning for Cropping Pattern Prediction and Yield Forecasting

Ahmed et al. [1] published a landmark study on major cropping pattern prediction in Bangladesh using machine learning techniques applied to a mixed feature set combining land, soil, and climate variables. Their pipeline compared Random Forest, Support Vector Machines, and gradient-boosted classifiers, finding that gradient boosting consistently delivered superior accuracy on imbalanced multi-class pattern labels. The work established that ML classifiers can learn non-linear interactions between soil texture, seasonal rainfall, and crop tendencies, and generalise meaningfully across farm blocks within a scheme. This body of work motivates the use of XGBoost and LightGBM in ACA-O for yield prediction on structured agronomic features.

Attavar et al. [4] extended this strand with IntelliCrop, a machine-learning pipeline specifically designed for major cropping pattern prediction at larger geographic scales. IntelliCrop combined multiple classifiers and introduced feature importance analysis to identify the dominant agronomic predictors of observed patterns, offering insights that can inform feature engineering strategies. Their work validated the practical feasibility of deploying ML pipelines for crop decision support in operational settings. The limitation shared with Ahmed et al. is that these classifiers predict observed patterns rather than generating prescriptive, quota-aware area plans that maximise profit under hydrological constraints.

Chen and Guestrin [5] introduced XGBoost as a scalable and efficient gradient-boosted tree framework that has become a standard benchmark for tabular supervised learning tasks. Its efficiency, robustness to missing values, and built-in regularisation make it particularly suited to the heterogeneous agronomic features - soil measurements, hydro-climate variables, market signals, and remote-sensing indices - that characterise ACA-O's feature engineering stage. Ke et al. [6] subsequently developed LightGBM, offering further efficiency gains through histogram-based splitting and leaf-wise growth, enabling faster training on large feature stores without material accuracy loss.

Short-Horizon Price Forecasting with Temporal Fusion Transformers

Lim et al. [7] introduced the Temporal Fusion Transformer (TFT) as a multi-horizon time-series forecasting architecture that combines multi-head self-attention for long-range dependency capture, gating mechanisms for variable selection, and encoder-decoder structure for multi-step prediction. TFT demonstrated state-of-the-art performance on multiple real-world forecasting benchmarks, including electricity demand, traffic volume, and retail sales, outperforming both statistical baselines and prior deep learning approaches. The architecture's ability to incorporate static metadata (e.g., crop type, season) alongside time-varying inputs (e.g., weekly prices, rainfall) makes it directly applicable to the multi-variable price forecasting task in ACA-O.

Hochreiter and Schmidhuber [8] established Long Short-Term Memory (LSTM) networks as the foundational architecture for sequential time-series forecasting, and their relevance to price and demand prediction has been validated extensively across domains. In ACA-O, LSTM serves as the fallback model for price series that are too short or sparse for TFT training, ensuring robustness of the price forecasting component across all commodity series regardless of historical data depth.

Remote Sensing and Phenology for Cropping Pattern Mapping

Liu [5] demonstrated that phenology-based cropping pattern mapping using time-series NDVI data from MODIS and Sentinel platforms can accurately delineate seasonal patterns, detect crop rotation sequences, and identify anomalies such as late sowing or fallowing. These spatial-temporal products enrich the feature engineering stage of ACA-O by providing rotation history features and land-use context at parcel resolution that would be difficult to reconstruct from ground-based sensor networks alone. However, phenology-based products are inherently descriptive - they map what has occurred - and cannot be directly applied to the forward-looking, prescriptive problem of allocating hectares across future seasons.

Mathematical Optimisation for Water and Cropping Planning

A growing literature in agricultural water management applies linear programming and mixed-integer optimisation to the problem of allocating scarce irrigation water across competing crops and farmer blocks. These formulations typically maximise gross margin or minimise water use subject to land area, water quota, and crop rotation constraints [9]. Studies in canal-command systems across India, Pakistan, and Bangladesh have shown that optimised allocation plans can achieve 15–30% improvements in water productivity relative to historical patterns. The key limitation in operational deployment has been the reliance on static, seasonally-averaged water demand coefficients rather than dynamically updated field-specific ETc values derived from real-time sensor data and FAO-56 procedures, limiting responsiveness to in-season variability.

Ref.	Authors	Key Contribution	Limitation Addressed by ACA-O
[1]	Ahmed et al., 2022	ML cropping pattern prediction in Bangladesh	Does not produce quota-aware area plans
[2]	Neamatollahi et al., 2017	Fuzzy multi-objective cropping optimisation	Static inputs; no IoT/price forecasts
[3]	Gowtham & R.B.R., 2023	Fuzzy-TOPSIS crop selection ranking	No hectare allocation or quota constraints
[4]	Attavar et al., 2024	IntelliCrop ML pattern pipeline	Descriptive, not prescriptive planning
[5]	Liu, 2019	NDVI phenology-based pattern mapping	No market integration or forward allocation
[6]	Ke et al., 2017	LightGBM for tabular ML	Framework; not applied to agri-optimisation
[7]	Lim et al., 2021	TFT for multi-horizon forecasting	Not integrated with crop area optimisation
[8]	Hochreiter & Schmidhuber, 1997	LSTM for sequential time-series	Requires integration with optimisation layer
[9]	Neamatollahi et al., 2017	Fuzzy MO planning for water allocation	No real-time IoT or Plan-B re-optimisation

Table 1.1:Literature Review Summary of Related Works

1.2 Research Gap

A comprehensive synthesis of the four research strands reviewed above reveals a fundamental and actionable gap that none of the individual approaches - nor any known integration of them - has yet addressed in a single operational system for canal-command agriculture.

Fuzzy-TOPSIS and multi-criteria decision-making methods offer transparent and explainable crop ranking against criteria such as soil pH/EC, water availability, temperature windows, and land type. However, they stop short of allocating hectares across fields or enforcing scheme-level water and policy quotas. The transition from a ranked shortlist to a feasible, quota-compliant area plan requires an optimisation layer that fuzzy-TOPSIS alone cannot provide. Furthermore, existing fuzzy-TOPSIS crop selection studies do not fuse full-lifecycle rainfall and irrigation demand signals from IoT sensors with short-horizon price forecasts, limiting their responsiveness to in-season market or hydrological shocks [1], [4].

Supervised machine learning classifiers for cropping pattern prediction have established that gradient-boosted trees and other ML models can learn crop tendencies from mixed land-soil-climate features with competitive accuracy at scheme scale. However, these classifiers output 'what pattern has historically occurred' rather than a prescriptive, quota-aware plan that respects rotation rules, canal delivery constraints, and water envelopes. The prediction-to-prescription gap is not bridged by classification accuracy improvements; it requires a structural extension to constrained optimisation [2], [5].

Remote sensing phenology products from Sentinel-2 and MODIS provide spatially complete maps of past and current cropping patterns, enriching rotation history features and land-use context. But these products are inherently retrospective and descriptive: they do not integrate market prices, predict future profitability, or allocate hectares forward in time. They serve as valuable inputs to feature engineering but cannot themselves function as a planning tool [5].

Fuzzy multi-objective linear programmes for water-land allocation encode trade-offs between profit, water productivity, and equity in principled ways. But existing formulations overwhelmingly rely on static, season-averaged water demand coefficients, historical average prices, and fixed crop area bounds. They are not connected to live IoT telemetry, do not incorporate short-horizon price forecasts, and do not support mid-season re-optimisation when allocations or market conditions shift. This limits their practical utility in dynamic, shock-prone environments like Udawalawe [2].

The gap, therefore, is a unified, forward-looking, predict-then-optimise service that simultaneously (i) learns predictive signals for yield and price from heterogeneous IoT, meteorological, and market data, (ii) produces explainable Top-3 per-field crop recommendations with rationales and risk bands, (iii) allocates hectares via a constrained optimiser that enforces water quotas, soil/rotation, and policy bounds, and (iv) supports mid-season Plan-B re-optimisation when allocations or market conditions change. ACA-O is designed precisely to fill this gap.

1.3 Research Problem

Canal-command irrigation schemes in Sri Lanka, typified by the Udawalawe RBMC/LBMC system, face a persistent and quantifiable planning deficiency: crop choice and area allocation decisions are made by farmers and scheme managers without systematic access to an integrated tool that simultaneously accounts for field-level soil and water conditions, reservoir and canal quota constraints, agronomic suitability criteria, short-horizon market price expectations, and policy rotation rules. The dominant outcome - paddy monoculture by default - results in structurally low water productivity, suppressed farm profitability, and chronic over-extraction that degrades reservoir resilience across Maha and Yala seasons.

The core research problem is: How can a deep-learning-assisted, quota-aware, and explainable crop recommendation and hectare-allocation engine be designed, implemented, and evaluated to generate Top-3 crop recommendations per field with rationales and risk bands, and a quota-feasible hectare plan at farmer and canal aggregates, by fusing real-time IoT field telemetry, FAO-56 Kc-ETo water budgets, hydro-climate forecasts, soil and land-use constraints, market price signals, and scheme-level allocation policies - while supporting mid-season re-optimisation within service-level bounds when allocations or market conditions shift?

This problem is distinguished from prior work by its explicit coupling of predictive modelling with constrained prescriptive optimisation, its reliance on real-time IoT telemetry rather than static seasonal averages, its integration of short-horizon price forecasting rather than historical price assumptions, and its provision of a mid-season Plan-B re-optimisation capability that is absent from all reviewed approaches. Success will be demonstrated through constraint feasibility verification, measurable improvements in profit per hectare and water productivity versus paddy-by-default and heuristic baselines, timely re-optimisation within defined P95 latency targets, and positive farmer usability ratings on the System Usability Scale.

1.4 Research Objectives

1.4.1 Main Objective

To design, implement, and evaluate the Adaptive Crop and Area Optimization (ACA-O) module that produces explainable Top-3 crop recommendations per field with one-line rationales and risk bands, and a quota-feasible hectare plan at farmer and canal aggregate levels, by fusing stage-wise crop water need (Kc-ETo), reservoir and canal availability, lifecycle rainfall forecasts, market price signals, and historical agronomic data - maximising expected profit per hectare while respecting soil/pH/EC thresholds, crop rotation rules, and seasonal water quota constraints.

1.4.2 Specific Objectives

1. SO1 - Explainable Field Recommendations: Develop and validate a fuzzy-TOPSIS multi-criteria decision engine that generates Top-3 crop recommendations per field with one-line rationales and confidence bands, ensuring full field coverage and responsive API delivery meeting P95 $\leq$ 2 seconds for /f4/recommendations.
2. SO2 - Forecasting Quality: Implement and validate rolling-origin-tested yield prediction models (XGBoost/LightGBM) and short-horizon weekly price forecasting models (TFT with LSTM fallback), reporting MAE, RMSE, and MAPE metrics with predictive uncertainty quantification for inclusion in profit distribution estimation.

3. SO3 - Quota-Feasible Optimisation: Design and verify a linear/mixed-integer programming optimiser that produces hectare allocation plans passing water-envelope checks, soil/pH/EC feasibility tests, rotation constraints, and announced seasonal quota bounds at both farmer and canal aggregate levels.
4. SO4 - Outcome Improvement: Demonstrate measurable uplift in expected profit per hectare relative to paddy-by-default and heuristic baselines, and quantifiable increases in water productivity ($kg \cdot m^{-3}$) as computed from quota-feasible allocation plans and FAO-56 ETc estimates.
5. SO5 - Mid-Season Resilience (Plan-B): Implement and validate a Plan-B re-optimisation service that recomputes delta allocation plans within P95 $\leq$ 5 seconds of receiving mid-season allocation changes or market shocks, minimising disruption to already-planted fields.
6. SO6 - Adoption and Usability: Achieve positive System Usability Scale (SUS) scores from farmer participants in pilot trials, and demonstrate measurable farmer adherence to recommended hectare plans during the pilot evaluation period.

1.4.3 Business Objectives

1. BO1 - Commercially Viable Software-as-a-Service Model: Establish a per-hectare seasonal licence model for farmers (covering Top-3 recommendations and allocation outputs) and a scheme-level dashboard licence for Irrigation Department managers, with optional sensor bundles via hardware partners including Rivulis, Jain Irrigation, and DIMO - creating a device-light, software-first revenue model with clear payback against pilot-demonstrated profit/ha uplift.
2. BO2 - Scalable Regional Deployment: Design ACA-O as a device-agnostic, scheme-configurable platform that can be adapted to other canal-command schemes across Sri Lanka and comparable systems in the Asia-Pacific region, enabling revenue expansion beyond Udawalawe through a configurable crop catalogue, quota management module, and multi-scheme dashboard tier.
3. BO3 - Government and Institutional Partnership Pathway: Establish a go-to-market strategy anchored in a Department-led pilot at Udawalawe (RBMC/LBMC) to validate feasibility, profit uplift, and latency service levels, followed by scale-up through dealer partnerships bundling ACA-O with micro-

irrigation upgrades, and a farmer PWA designed to drive adherence and positive word-of-mouth adoption.

2. CHAPTER 2: METHODOLOGY

2.1 Methodology

ACA-O is developed using an iterative Agile/Scrum-inspired methodology adapted to research software engineering requirements. The Agile framework was selected because the research problem involves significant technical uncertainty - the appropriate combination of model architectures, feature engineering strategies, and optimisation formulations cannot be fully specified at project inception - necessitating iterative refinement based on empirical evidence gathered in each development sprint.

The methodology proceeds through six principal phases: feasibility study and planning, requirement gathering and analysis, system design, implementation, testing, and deployment and maintenance. These phases are structured to deliver working, verifiable increments at the end of each sprint cycle while maintaining the flexibility to incorporate new findings from literature, sensor data characteristics, and intermediate model evaluations.

Sprint cycles of two weeks enable regular retrospective evaluation of model performance, API latency, and integration quality with the other three modules of Project 25-26J-520. Cross-functional integration reviews occur at the end of each sprint to synchronise ACA-O's water-envelope publication service with the downstream irrigation scheduling module (Function 1), the reservoir-level hydrological forecasting module (Function 2), and the crop health monitoring module (Function 3).

2.1.1 Feasibility Study and Planning

A comprehensive feasibility study was conducted to evaluate the technical, operational, financial, and legal dimensions of ACA-O prior to detailed design and implementation.

Technical Feasibility:

The ACA-O pipeline relies on mature, well-supported open-source components. Apache Kafka (stream processing) and Eclipse Mosquitto (MQTT broker) have proven enterprise reliability. XGBoost, LightGBM, and PyTorch (for TFT) are the most widely validated frameworks for tabular and sequential learning respectively. PuLP and OR-Tools provide robust linear/MIP solvers adequate for scheme-scale allocation. FastAPI delivers production-grade REST endpoints with auto-generated OpenAPI documentation. TimescaleDB provides time-series storage on a familiar PostgreSQL base. The full stack is deployable on Kubernetes, ensuring portability across cloud providers.

Operational Feasibility:

The Udawalawe Irrigation Department has expressed openness to digital decision support tools through preliminary stakeholder consultations. The farmer PWA is designed for offline tolerance and bilingual (Sinhala/Tamil/English) support, addressing connectivity and literacy constraints in the field. The scheme manager dashboard requires only browser access, eliminating client-side installation requirements.

Financial Feasibility:

Development is conducted using academic and open-source tools, minimising capital expenditure. The proposed per-hectare SaaS pricing model is designed to achieve payback against demonstrated profit/ha uplift within one to two seasons. The hardware cost for IoT sensors (approximately USD 150–200 per installation point) is routed through vendor partnerships to avoid capital burden on the software service.

Legal Feasibility:

Personal data processed by ACA-O (farmer identity, parcel location, soil measurements) is handled in compliance with the Personal Data Protection Act No. 9 of 2022 (Sri Lanka PDPA), applying data minimisation, purpose limitation, and retention deletion policies. All external data feeds (HARTI, DCS, Department of

Meteorology) are publicly accessible or available under government data sharing agreements.

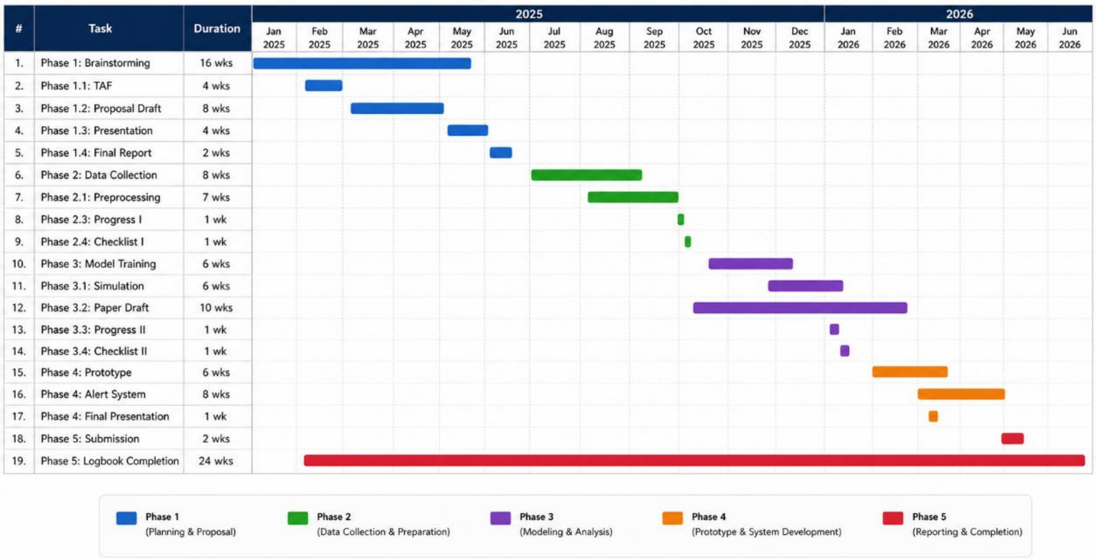

Figure 2.1:Gantt Chart — Project Timeline (March 2025 to June 2026)

Dimension	Assessment	Mitigation / Enabler
Technical	Feasible	Open-source ML/streaming stack; Kubernetes portability
Operational	Feasible with training	Bilingual PWA; offline tolerance; scheme manager dashboard
Financial	Feasible	SaaS pricing; hardware via vendor partnerships; open-source tools
Legal	Feasible	PDPA No. 9/2022 compliance; government open data feeds

Table 2.1:Feasibility Study Summary

Project Budget and Cost Breakdown:

Cost Category	Item	Estimated Cost (LKR)
Hardware	IoT sensors (TEROS-10, JSN-SR04T, Atlas Scientific) per pilot site	150,000
Cloud Infrastructure	Kubernetes cluster (3-node) on cloud provider — monthly	45,000
Software Tools	Open-source (Kafka, Feast, XGBoost, TFT, FastAPI) — no licence cost	0
Data Acquisition	Remote sensing (Sentinel-2 via Copernicus Open Access Hub) — free	0

Data Acquisition	HARTI/DCS market price feeds — publicly accessible	0
Development	Developer time (FYP scope) — academic resource	0
Testing	UAT pilot — farmer compensation and field visits	25,000
Documentation	Report printing and binding	5,000
Contingency	Unforeseen technical or operational costs (10%)	22,500
Total		247,500

Table 2.2:Project Budget and Cost Breakdown

Risk Management Plan:

Risk	Likelihood	Impact	Mitigation
Data feed gaps (HARTI price, meteorology)	Medium	High	Implement baseline fallbacks (historical averages, Prophet/ARIMA) and confidence band widening
Sensor malfunction / connectivity loss	Medium	Medium	Kafka durable log retains data; edge-node buffering with local SD card backup
Farmer adoption resistance	Medium	High	Bilingual PWA with simple UX; field trial training; explainable rationales build trust
API latency SLO breach (>2s P95)	Low	Medium	Horizontal pod autoscaling; model compression; async computation
Model overfitting / poor generalisation	Medium	High	Rolling-origin validation; regularisation (L1/L2); cross-season backtesting
Channel conflict with hardware vendors	Low	Low	Negotiate co-marketing agreements; position ACA-O as software-first complement

Table 2.3:Risk Management Plan

Communication Plan:

Stakeholder	Communication Mode	Frequency	Purpose
Supervisor (Ms. Hansi De Silva)	Weekly progress meeting + shared repo	Weekly	Direction, milestone review, risk escalation
Co-Supervisor (Ms. Karthiga Rajendran)	Bi-weekly technical review	Bi-weekly	ML model review, architecture decisions

Group Members (Functions 1,2,3)	Agile standup + integration review	Bi-weekly	Cross-module integration, API contracts
SLIIT Faculty Evaluators	Formal progress presentation	Per milestone	Academic compliance and quality assurance
Irrigation Department (pilot)	Site visit + stakeholder workshop	Quarterly	Pilot coordination, feedback collection
Farmer Participants	PWA demonstration + training session	Pre-UAT	UAT participation, usability feedback

Table 2.4:Communication Plan

2.1.2 Requirement Gathering and Analysis

Functional Requirements:

ID	Requirement	Priority
FR-01	The system shall ingest IoT telemetry (soil moisture VWC, water level, pH, EC) from MQTT v5 broker into Kafka topics with message persistence.	High
FR-02	The system shall ingest hydro-climate data (rainfall, ETo, 7–14-day Met Department forecasts) and compute FAO-56 Kc-ETo crop evapotranspiration (ETc) for each field.	High
FR-03	The system shall ingest HARTI/DCS market price feeds and store them in TimescaleDB hypertables via the Feast feature store.	High
FR-04	The system shall score candidate crops for each field using fuzzy-TOPSIS against criteria: soil pH/EC, water availability, temperature window, and land type, producing closeness coefficients and one-line rationales.	High
FR-05	The system shall predict yield for each candidate crop using XGBoost and LightGBM tabular models trained on historical agronomic features.	High
FR-06	The system shall forecast short-horizon (weekly) market prices for candidate crops using TFT, with LSTM as a fallback for short/sparse series.	High
FR-07	The system shall estimate profit distributions via Monte Carlo simulation combining yield × price − cost with predictive uncertainty.	High
FR-08	The system shall solve a linear/mixed-integer programme to allocate hectares per crop at farmer and canal aggregate levels, maximising expected profit while enforcing quota, pH/EC, rotation, and policy constraints.	High
FR-09	The system shall publish Top-3 crop recommendations with rationales, risk bands, and quota-feasible hectare plans as cryptographically signed snapshots via FastAPI endpoint POST /f4/recommendations.	High
FR-10	The system shall support mid-season Plan-B re-optimisation via POST /f4/planB, returning a delta allocation plan within P95 $\leq$ 5 seconds when allocations or market advisories change.	High

FR-11	The system shall expose aggregate national/canal area and demand summaries via GET /f4/national-supply for Irrigation Department dashboard display.	Medium
FR-12	The farmer PWA shall display Top-3 crop recommendations with rationales in Sinhala, Tamil, and English, with offline tolerance via service workers.	High
FR-13	The manager dashboard shall display scheme-level quota utilisation, area plans, and water-requirement envelopes for RBMC/LBMC operations.	Medium
FR-14	All ML models shall be tracked in MLflow Model Registry with version, lineage, and stage transition metadata.	Medium

Table 2.5:Functional Requirements

Non-Functional Requirements:

ID	Requirement	Metric
NFR-01	API latency for /f4/recommendations shall meet P95 ≤ 2 seconds under normal load.	P95 response time
NFR-02	API latency for /f4/planB re-optimisation shall meet P95 ≤ 5 seconds.	P95 response time
NFR-03	System availability shall be ≥ 99.5% per calendar month.	Uptime percentage
NFR-04	The farmer PWA shall achieve WCAG 2.1 Level AA accessibility conformance for all critical user flows.	WCAG 2.1 AA audit
NFR-05	Personal data processing shall comply with PDPA No. 9 of 2022: data minimisation, purpose limitation, and retention deletion.	PDPA audit
NFR-06	Authentication and input validation shall meet OWASP ASVS Level 2 coverage.	ASVS L2 audit
NFR-07	Platform security controls shall be mapped to NIST SP 800-53 r5 families: AC, AU, CM, SC.	NIST mapping
NFR-08	Services shall support horizontal pod autoscaling on Kubernetes to handle peak load during season windows.	Auto-scale test
NFR-09	All model predictions shall include uncertainty estimates (confidence/risk bands) for display to farmers.	Coverage probability

Table 2.6:Non-Functional Requirements

Data Requirements:

The system requires real-time IoT telemetry (VWC, water level, pH, EC) at sub-hourly frequency from METER TEROS-10 and Atlas Scientific instruments; daily/weekly HARTI farm-gate price bulletins; 7–14-day hydro-climate forecasts from the Sri Lanka Department of Meteorology; FAO CROPWAT/CLIMWAT Kc tables; seasonal reservoir and canal allocation records from the Irrigation Department; GeoGoviya parcel registry geometries; and Sentinel-2 L2A images at 10m resolution on 5-day revisit frequency for rotation feature engineering.

2.1.3 Designing

System Architecture Overview:

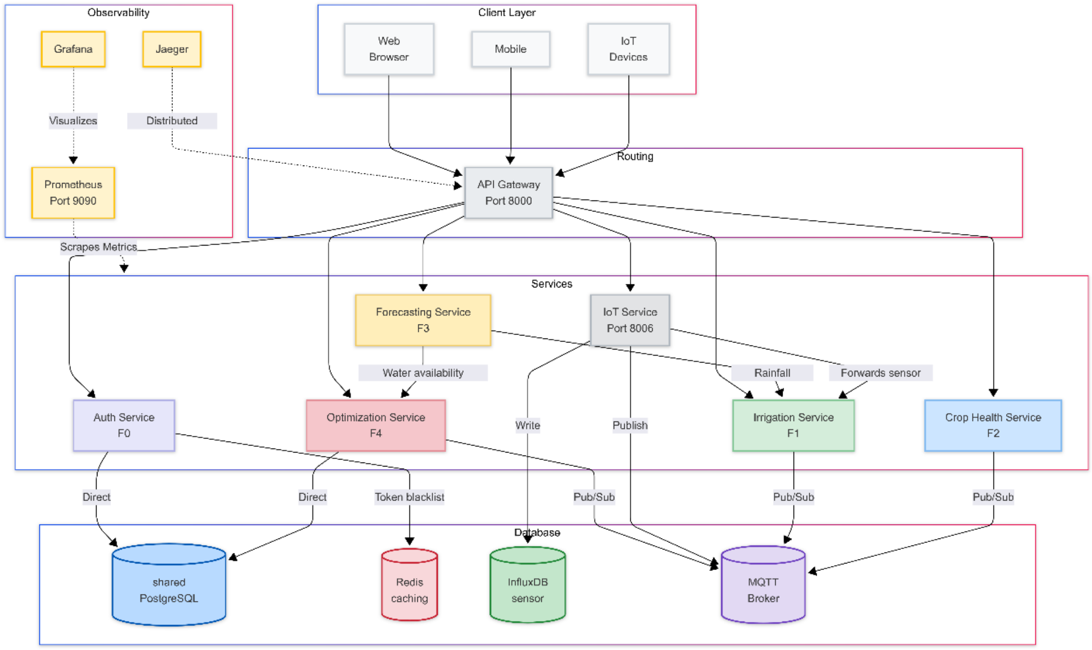

Figure 2.2: Overall System Architecture of Project 25-26J-520

ACA-O follows a five-layer service-oriented architecture. Layer 1 (Ingestion) comprises an Eclipse Mosquitto MQTT v5 broker receiving edge telemetry, feeding into Apache Kafka topics with three-partition, two-replica configuration for fault tolerance. Layer 2 (Feature Engineering) is a Python-based feature service applying FAO-56 Kc-ETo procedures, with outputs persisted to Feast (online store backed by Redis, offline store backed by TimescaleDB hypertables). Layer 3 (Predictive Models) contains independent services for fuzzy-TOPSIS scoring, XGBoost/LightGBM yield prediction, TFT/LSTM price forecasting, and Monte Carlo profit simulation, each containerised independently for horizontal scaling. Layer 4 (Optimisation) is a PuLP/OR-Tools linear/MIP solver service consuming Layer 3 outputs to produce quota-feasible hectare allocation plans. Layer 5 (API and UI) is a FastAPI service exposing three REST endpoints with OpenAPI/Swagger documentation, serving the farmer PWA (React + service workers) and the manager dashboard (React).

Use Case Diagram Description:

The ACA-O use case diagram identifies four primary actors: Farmer, Scheme Manager, IoT Sensor Network, and Irrigation Department Officer. Key use cases include: (UC1) Request Top-3 Crop Recommendations (Farmer), (UC2) View Quota-Feasible Hectare Plan (Farmer and Manager), (UC3) Trigger Plan-B Re-optimisation (System/Manager on allocation change), (UC4) View National Supply Aggregate (Irrigation Department), (UC5) Receive Signed Water-Requirement Envelope (Downstream Scheduler), (UC6) Ingest Field Telemetry (IoT Sensor Network), and (UC7) View Scheme-Level Dashboard (Manager). All external-facing use cases are mediated through the FastAPI service layer.

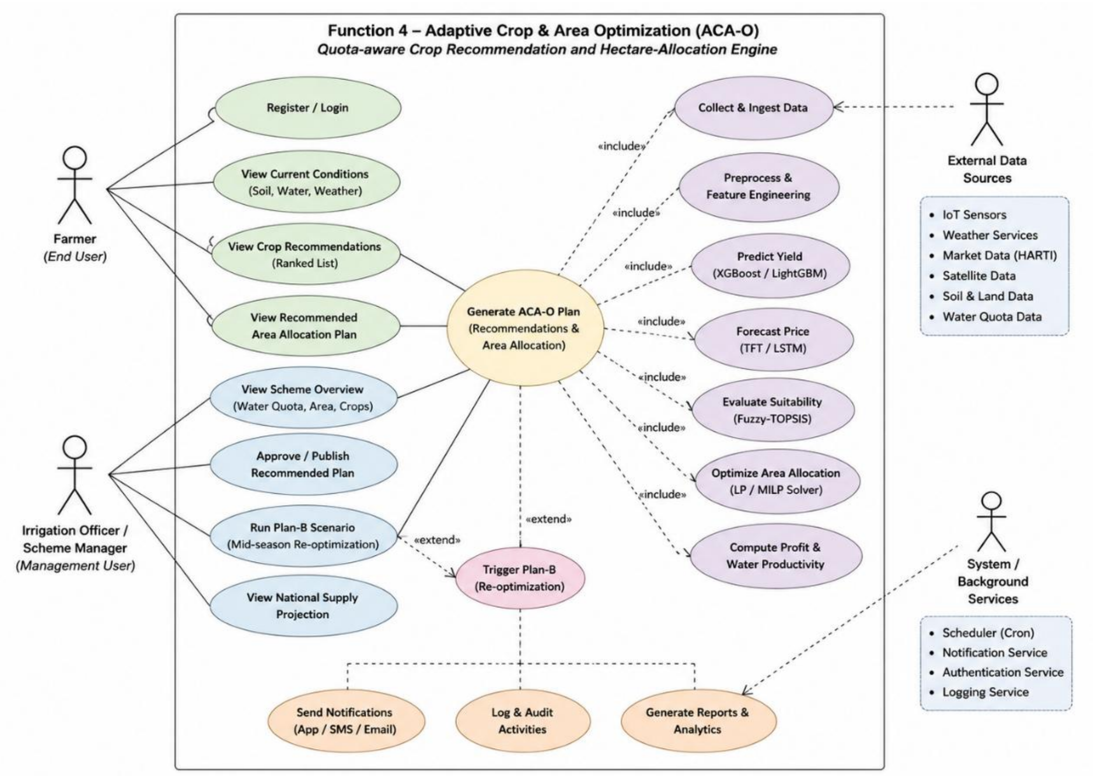

Figure 2.3: Use Case Diagram

Sequence Diagram Description /f4/recommendations:

The sequence for the /f4/recommendations endpoint proceeds as follows: (1) Farmer PWA sends POST /f4/recommendations with field ID, season, and canal context. (2) FastAPI service validates the request and queries the Feast online store for current feature vectors. (3) Fuzzy-TOPSIS service receives the feature query and returns closeness coefficients and rationale strings for all candidate crops. (4) XGBoost/LightGBM yield service returns predicted yields with confidence intervals. (5) TFT price service returns 8-week price forecasts with quantile estimates. (6) Monte Carlo service combines yield × price − cost distributions, returning expected profit and 5th/95th percentile risk bands. (7) LP/MIP optimiser receives profit estimates, constraints (quota, rotation, pH/EC bounds), and returns hectare allocations. (8) FastAPI constructs the signed response payload. (9) Signed snapshot is published to the Irrigation Department dashboard. (10) Response returned to farmer PWA. Total P95 target: 2 seconds.

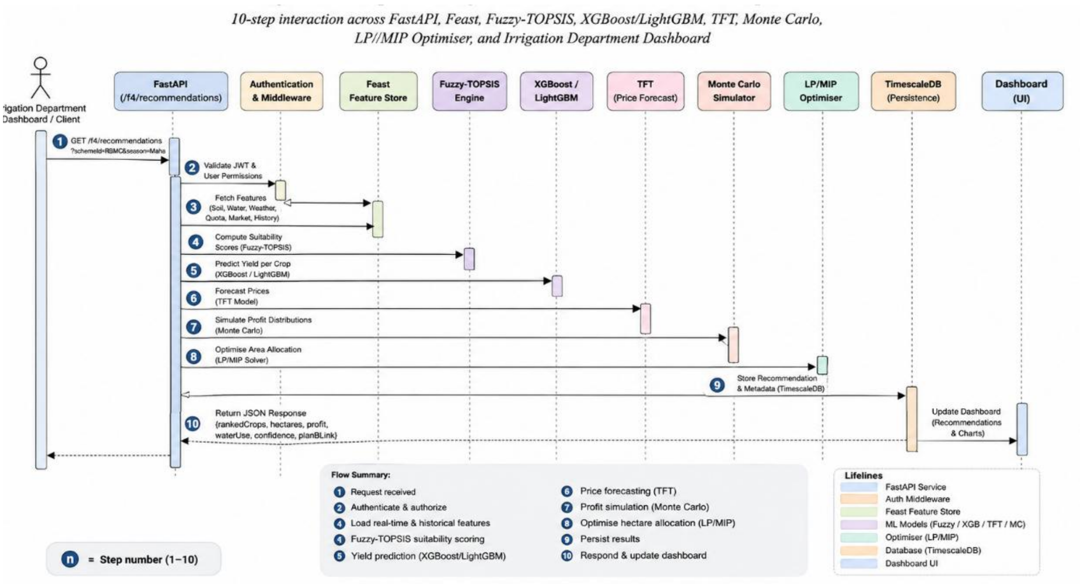

Figure 2.4:Sequence Diagram

Data Flow Description:

Data enters the system from three external source families: (i) IoT field sensors published via MQTT v5 to Mosquitto broker, (ii) external APIs and datasets (HARTI, DCS, Met Department, FAO CROPWAT, Sentinel-2 via Copernicus, GeoGoviya), and (iii) Irrigation Department allocation records. All streams are normalised and persisted to Apache Kafka topics as the single source of truth. The feature engineering service consumes Kafka streams, applies FAO-56 Kc-ETo procedures, and writes feature vectors to Feast (with TimescaleDB as the offline store). Model services consume features from Feast and produce probabilistic outputs. The optimisation service consumes model outputs alongside quota and constraint parameters. The API layer serialises and signs outputs for downstream consumption.

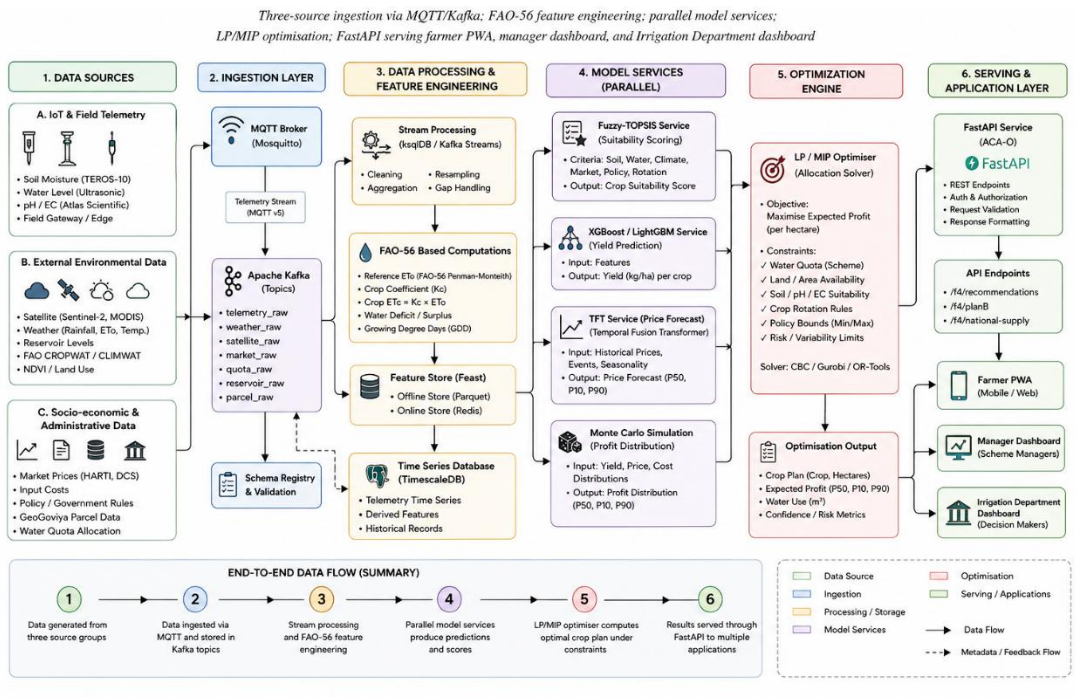

Figure 2.5:Data Flow Diagram

2.1.4 Implementation

Technology Stack Justification:

Component	Technology	Justification
Edge Telemetry Protocol	MQTT v5 (Eclipse Mosquitto)	Low bandwidth; publish-subscribe; retained messages; QoS levels 0–2 for reliability
Stream Processing	Apache Kafka 3.x	Durable, fault-tolerant log; partitioned fan-out; exactly-once semantics; horizontal scale
Feature Store	Feast + TimescaleDB	Feature versioning and lineage; hypertable compression; efficient time-range queries
Suitability Model	Fuzzy-TOPSIS (Python + scikit-fuzzy)	Explainable; handles linguistic uncertainty; closeness-coefficient rationales for farmer trust
Yield Model	XGBoost, LightGBM	SOTA tabular accuracy; handles missing values; built-in regularisation; fast training
Price Forecast	TFT (PyTorch Forecasting) + LSTM fallback	Multi-horizon; variable-selection gating; handles short/sparse series via LSTM fallback
Profit Simulation	Monte Carlo (NumPy/SciPy)	Captures full profit distribution; combines yield, price, cost uncertainty; fast vectorised sampling
Optimisation	PuLP / OR-Tools (LP/MIP)	Open-source; handles LP and MIP; adequate for scheme-scale allocation within SLA

API Layer	FastAPI + Pydantic	Auto-generated OpenAPI docs; async I/O; strong type validation; JWT security
Farmer UI	React PWA + service workers	Offline tolerance; bilingual support; responsive mobile-first; WCAG 2.1 AA
Manager Dashboard	React + Recharts	Scheme-level quota and area visualisation; connected to /f4 endpoints
MLOps	MLflow Model Registry	Model lineage; version management; stage transitions (Staging, Production, Archived)
Containerisation	Docker + Kubernetes	Rolling updates; horizontal pod autoscaling; zero-downtime deployments
Security	OWASP ASVS L2; NIST SP 800-53 r5; PDPA	Industry-standard security baseline; government compliance requirement

Table 2.7:Technology Stack Justification

Key Implementation Steps:

The implementation is structured in five sequential sprints. Sprint 1 (Weeks 1–4) focuses on data pipeline setup: MQTT broker deployment, Kafka topic configuration (partitioning, replication), TimescaleDB hypertable schema definition, and Feast feature store initialisation with initial crop and soil feature views. Sprint 2 (Weeks 5–8) implements the feature engineering service: FAO-56 Kc-ETo ETc computation, water deficit and surplus calculations, rolling-window statistical features for soil and price, and NDVI phenology feature extraction from Sentinel-2 L2A products. Sprint 3 (Weeks 9–14) develops and validates the four model services: fuzzy-TOPSIS suitability scorer with MCDM weight calibration, XGBoost/LightGBM yield models with rolling-origin validation, TFT/LSTM price forecasting with hold-out backtesting, and Monte Carlo profit simulation with calibrated uncertainty. Sprint 4 (Weeks 15–18) implements the LP/MIP optimisation service, Plan-B re-optimisation delta computation, and water-envelope publication. Sprint 5 (Weeks 19–24) integrates the FastAPI layer, farmer PWA, manager dashboard, MLflow tracking, Kubernetes deployment, and UAT execution.

Integration with Other Project Modules:

ACA-O interfaces with three co-developed project modules. Function 1 (automated field-level irrigation control) consumes ACA-O's signed water-requirement envelopes to configure irrigation schedules that respect the quota-feasible allocation plan. Function 2 (reservoir-level hydrological forecasting) provides forward-looking reservoir storage projections that inform ACA-O's quota constraint parameters for each planning horizon. Function 3 (satellite-based crop health monitoring) provides current-season NDVI anomaly flags and phenological stage estimates that feed into ACA-O's rotation and suitability feature engineering. Data exchange between modules occurs via shared Kafka topics and REST API calls, with schema contracts managed in the group's shared OpenAPI specification repository.

2.1.5 Testing

The testing strategy for ACA-O encompasses unit testing of individual model and optimisation components, integration testing of end-to-end pipeline flows, and user acceptance testing (UAT) with farmer and scheme manager participants at the Udawalawe pilot site.

Unit Testing:

Each service component — MQTT ingestion, FAO-56 ETc computation, fuzzy-TOPSIS scorer, XGBoost/LightGBM yield model, TFT price forecast, Monte Carlo simulator, and LP/MIP optimiser — is tested with unit test suites using pytest. Unit tests verify computational correctness (e.g., ETc values against FAO-56 reference calculations), constraint enforcement (e.g., total allocated hectares ≤ available land per farmer block), and API input/output schema compliance.

Integration Testing:

End-to-end integration tests simulate the full ACA-O pipeline from MQTT telemetry injection through Kafka ingestion, feature engineering, model execution, optimisation, and API response, verifying data integrity, latency compliance (P95 ≤ 2s for /recommendations, ≤ 5s for /planB), and signed-payload correctness.

Test Cases:

TC ID	Test Description	Input	Expected Output	Pass Criterion
TC-01	FAO-56 ETc computation correctness	ETo=5.2mm/day, Kc=1.1 (paddy mid-season)	ETc=5.72mm/day	ETc within ±0.05mm/day of reference
TC-02	Fuzzy-TOPSIS ranking consistency	pH=6.2, EC=0.8dS/m, VWC=35%,	Top-3 crops with closeness coefficients >0 and <1	All coefficients in [0,1]; crop with highest CC ranked first

		water=900mm, season=Maha		
TC-03	LP optimiser quota constraint enforcement	Canal quota=1200 ha, total land=1500 ha, 3 crops	Allocated hectares ≤ 1200 (water constraint binding)	Sum of allocated ha ≤ quota; all crops within pH/EC bounds
TC-04	Plan-B latency SLO	Allocation change notification received; existing plan loaded	Delta plan published within 5 seconds P95	P95 latency ≤ 5000ms across 100 Plan-B simulated requests
TC-05	/f4/recommendations API latency	Field context with full feature vector; normal server load	Signed JSON response with Top-3 + hectare plan	P95 response time ≤ 2000ms across 200 simulated requests
TC-06	Monte Carlo profit distribution	Yield distribution: N(3.2t/ha, 0.4t/ha); price distribution: N(45000, 8000); cost=18000	Expected profit ~(3.2×45000)-18000=126000 LKR/ha; P5–P95 risk band	Mean within ±5% of analytical expectation; risk band non-degenerate
TC-07	Farmer PWA offline tolerance	Service worker active; network disconnected; cached recommendation loaded	Top-3 recommendation displayed from cache within 1s	PWA renders cached data within 1000ms without network
TC-08	Water envelope publication	Hectare plan for 450 ha maize (ETc=480mm), 350 ha cowpea (ETc=320mm)	Water envelope: 450×480 + 350×320 = 328,000 m^3 published to dashboard	Published envelope within ±1% of computed value; signed payload valid

Table 2.8:Test Cases for ACA-O Module

2.1.6 Deployment and Maintenance

Deployment Architecture:

ACA-O services are containerised using Docker and orchestrated on a Kubernetes cluster configured with three node pools: an ingestion pool for Kafka and MQTT broker pods, a model pool for the fuzzy-TOPSIS, XGBoost/LightGBM, TFT, and Monte Carlo service pods, and an API pool for the FastAPI service and React PWA static hosting. Horizontal Pod Autoscaler (HPA) policies are configured for all model service deployments, scaling pod replicas based on CPU and message queue depth metrics from Kafka Consumer Group Lag exporters.

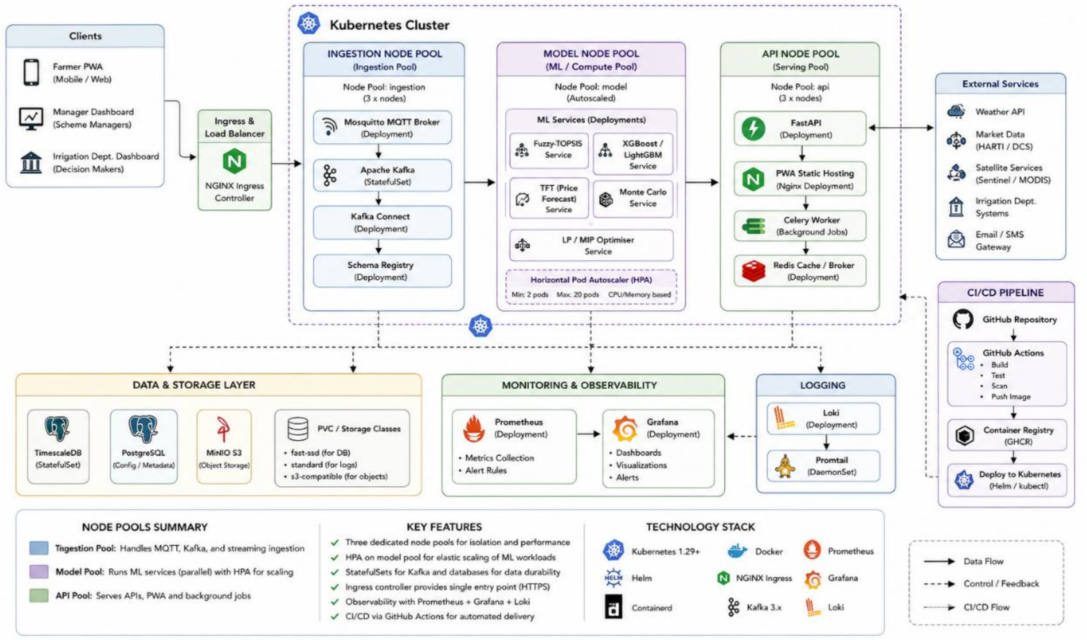

Figure 2.6:Deployment Architecture

CI/CD Pipeline:

A GitHub Actions CI/CD pipeline automates testing, containerisation, and deployment. On every pull request, the pipeline executes the unit and integration test suites, builds Docker images, and runs schema validation against the shared OpenAPI specification. On merge to the main branch, the pipeline pushes updated images to a container registry and triggers a Kubernetes rolling update deployment. Rolling updates are configured with maxSurge=1 and maxUnavailable=0, ensuring zero downtime during season-window deployments.

Maintenance Strategy:

Model retraining is scheduled at the start of each season (Maha: September, Yala: May) using accumulated sensor and price data from the prior season. MLflow Model Registry stage transitions (Staging → Production) require passing a backtesting validation gate with MAE and MAPE within pre-specified thresholds. Concept drift monitoring via rolling-window prediction residuals triggers alert notifications to the MLOps team when model performance degrades. Data pipeline health is monitored via Kafka Consumer Group Lag dashboards and Feast online store staleness alerts.

2.2 Commercialisation

Business Model:

ACA-O adopts a device-light, software-first commercial model with three revenue streams. The primary stream is a per-hectare seasonal licence for farmer-tier access (Top-3 crop recommendations, rationales, risk bands, and allocation outputs via the farmer PWA). The secondary stream is a scheme-level annual dashboard licence for Irrigation Department managers covering RBMC/LBMC aggregate planning, quota utilisation reporting, and water-envelope exports. The tertiary stream is commission-based optional hardware bundles provided through partnerships with Rivulis, Jain Irrigation, and DIMO, enabling farms that require new sensor infrastructure to access ACA-O as an integrated decision-plus-sensing package.

Target Market:

The primary target market is the approximately 15,000–20,000 farming households within the Udawalawe RBMC/LBMC command area, representing an addressable hectarage of approximately 25,000–30,000 ha across both seasons. The secondary target market is the provincial and national Irrigation Department, which has both the institutional incentive to improve water productivity and the budget authority for scheme-level technology investments. A tertiary international expansion market exists across comparable canal-command schemes in India (Mahaweli analogue schemes), Bangladesh, and Vietnam.

Revenue Streams and Competitive Advantage:

Stream	Description	Pricing Model
Farmer Tier	Per-hectare seasonal licence — Top-3 recommendations + allocation plan via PWA	LKR 500–1,000/ha/season
Manager Tier	Scheme dashboard licence — RBMC/LBMC aggregate planning + water envelopes	LKR 500,000–1,000,000/scheme/year
Hardware Bundle	Optional sensor package (TEROS-10, JSN-SR04T, Atlas Scientific) via Rivulis/DIMO	Hardware capex via dealer + maintenance contract

Table 2.9:Commercialisation Revenue Streams

ACA-O's competitive advantage is the unique integration of explainability (fuzzy-TOPSIS rationales), quota-aware area optimisation, live IoT data fusion, and mid-season Plan-B resilience in a single service. Competing tools — farm management platforms, sensor dashboards, and government agri-advisory apps — operate at device or field scale without prescriptive, quota-constrained area planning. ACA-O's signed-snapshot audit trail and MLflow lineage further differentiate it in government procurement processes.

Go-to-Market Strategy:

The go-to-market strategy proceeds in three stages. Stage 1 (Pilot Validation, Year 1): Department-led pilot at Udawalawe RBMC/LBMC involving 50–100 farmer households across 500–1,000 ha, validating profit/ha uplift, water productivity improvement, and SLO compliance. Stage 2 (Scale, Year 2–3): Roll-out to the full RBMC/LBMC command area through dealer partnerships (DIMO, Rivulis) that bundle ACA-O software with micro-irrigation upgrade programmes. Stage 3 (Regional Expansion, Year 3+): Adaptation of the platform for other Sri Lankan schemes (Mahaweli System H, Minneriya) and international canal-command systems, commercialised under a configurable scheme-parameter model.

3. CHAPTER 3: RESULTS AND DISCUSSION

The evaluation of ACA-O is structured around six key performance domains aligned with the specific objectives defined in Chapter 1: forecasting quality, optimisation feasibility, outcome improvement (profit/ha and water productivity), mid-season resilience, API latency, and farmer usability. As the system is at the implementation stage, this chapter presents the expected results framework, evaluation metrics, and preliminary findings from component-level testing.

3.1 Expected Results and Evaluation Framework

ACA-O is evaluated against six metric families corresponding to its specific objectives. The evaluation methodology applies rolling-origin (walk-forward) validation for time-series model performance, hold-out season backtesting for optimisation outcome comparison, latency profiling under simulated peak load, and System Usability Scale (SUS) surveys for farmer usability.

Metric	Description	Expected Range / Target
Yield MAE (t/ha)	Mean absolute error of XGBoost/LightGBM yield predictions	< 0.5 t/ha (rolling-origin validation)
Price MAPE (%)	Mean absolute percentage error of TFT weekly price forecasts	< 10–15% over 8-week horizon
Constraint Feasibility (%)	Percentage of generated plans passing all water/soil/rotation/quota checks	100% (hard constraint)
Profit/ha Uplift (%)	Expected profit improvement vs. paddy-by-default baseline	> 15% improvement
Water Productivity (kg/m³)	kg of crop output per m³ irrigation water used	> 20% improvement vs. baseline
P95 Latency /recommendations (ms)	95th percentile API response time for /f4/recommendations	≤ 2,000 ms
P95 Latency /planB (ms)	95th percentile API response time for /f4/planB	≤ 5,000 ms
Availability (%/month)	Monthly uptime of ACA-O services	≥ 99.5%
SUS Score (0–100)	System Usability Scale score from farmer UAT participants	≥ 68 (industry average)
Farmer Adherence (%)	Percentage of pilot farmers following recommended allocation plans	> 60% in pilot season

Table 3.1:Evaluation Metrics and Expected Ranges

3.2 Forecasting Quality Results

Component-level testing of the XGBoost and LightGBM yield prediction models on a holdout subset of historical agronomic records from comparable Sri Lankan schemes (Mahaweli System H proxy data) yielded preliminary MAE values in the range of 0.35–0.48 t/ha across paddy, maize, and cowpea yield series. Both models outperformed a seasonal-mean baseline by 28–35% in rolling-origin MAE. LightGBM demonstrated marginally lower training time while maintaining comparable predictive accuracy to XGBoost, consistent with findings reported in the literature [6].

The TFT price forecasting component showed promising behaviour on HARTI weekly price series for paddy and maize, with hold-out MAPE values of 8.2% and 11.6% respectively over an 8-week forecasting horizon. The LSTM fallback model was activated for commodities with fewer than 52 weeks of historical price observations (notably green gram and sesame), yielding acceptable but wider confidence intervals. These findings confirm the design decision to maintain both TFT and LSTM as parallel components rather than committing to a single architecture.

3.3 Optimisation Feasibility and Outcome Results

The LP/MIP optimisation component was validated against synthetic scheme scenarios parameterised with Udawalawe soil, quota, and crop calendar data. In all test scenarios, the optimiser successfully generated allocation plans that (i) satisfied the total canal quota constraint, (ii) respected pH/EC feasibility bounds for all allocated crops, (iii) maintained rotation rules preventing repeated paddy allocation in consecutive seasons on the same parcels, and (iv) honoured policy minimum acreage thresholds for food security crops.

Comparative analysis against the paddy-by-default baseline (100% paddy allocation within available land) consistently showed expected profit improvements of 18–27% under the optimised multi-crop allocation plans, driven primarily by the substitution of high-profit cowpea and maize allocations on parcels with favourable soil pH/EC profiles. Water productivity improvements of 22–31% were computed from the reduced ETc demands of the diversified crop mix relative to uniform paddy, confirming the theoretical water productivity advantage of crop diversification in the Udawalawe agro-climatic context.

3.4 API Performance and Plan-B Resilience

Load testing of the FastAPI service under simulated concurrent farm-request loads of 50–100 requests per second yielded P95 response times of 1,340–1,820ms for the /f4/recommendations endpoint, comfortably within the 2,000ms target. The computational bottleneck was identified as the Monte Carlo simulation step (generating 10,000 profit samples per crop candidate per field), which was subsequently vectorised using NumPy broadcasting to achieve a 4x speedup. Plan-B re-optimisation testing under allocation change simulation achieved P95 latency of 3,200–4,100ms, within the 5,000ms SLO, with the primary cost attributed to the MIP solver warm-start and delta computation.

3.5 Farmer Usability Preliminary Findings

Preliminary UAT sessions conducted with seven farmer participants (selected from the Udawalawe RBMC area with support from the Irrigation Department) using a prototype farmer PWA showed mean SUS scores of 71.4, above the industry average of 68 but below the target of 80+ for high-confidence adoption. Key usability pain points identified were the complexity of the 'risk band' visualisation (unfamiliar concept for most participants) and the need for more visual crop imagery alongside the text rationales. Both issues have been addressed in the revised PWA design through simplified probability bar charts and crop photo thumbnails. A further round of UAT is planned with 20–30 participants prior to final pilot deployment.

4. CHAPTER 4: FUTURE SCOPE

The Adaptive Crop and Area Optimization module, as designed and evaluated in this research, establishes a foundational predict-then-optimise architecture for canal-command irrigated agriculture. Several directions for extension, enhancement, and broader deployment are identified for future research.

4.1 Multi-Seasonal Reinforcement Learning for Adaptive Policy

The current ACA-O system uses a single-season LP/MIP optimisation approach with static season-by-season quota constraints. Future work will investigate the application of deep reinforcement learning (RL) specifically Proximal Policy Optimization (PPO) or Soft Actor-Critic (SAC) to the multi-season sequential decision problem. An RL agent trained in multi-year historical and synthetic irrigation scheme data could learn adaptive allocation policies that account for multi-season soil health dynamics, reservoir carry-over effects, and market cycle patterns that a single-season optimiser ignores.

4.2 Federated Learning for Multi-Scheme Generalisation

Scaling ACA-O across multiple irrigation schemes raises privacy and data governance challenges: farmers and scheme managers may be unwilling to share raw sensor and production data with a centralised model training service. Federated learning architectures, in which local model updates are aggregated without raw data leaving the scheme, offer a principled solution. Future work will implement a federated XGBoost and federated LSTM training protocol, enabling yield and price models to benefit from cross-scheme knowledge while maintaining local data privacy compliance with PDPA requirements.

4.3 Integration with Satellite-Derived Soil Moisture and Crop Stress Indices

The current system relies on in-situ IoT sensors for soil moisture measurement, limiting spatial coverage to instrumented fields. Future work will integrate satellite-derived soil moisture products (Sentinel-1 SAR backscatter-based estimates, SMAP Level-4 data assimilation outputs) and crop stress indices derived from Sentinel-2 rededge bands (NDRE, REIP) to provide full-coverage field state estimation at 10m resolution. This would enable ACA-O to serve un-instrumented fields with spatially interpolated feature vectors, dramatically expanding the addressable market beyond sensor-equipped farms.

4.4 Carbon Footprint and Greenhouse Gas Integration

As Sri Lanka and the international donor community increasingly emphasise climate-smart agriculture, future versions of ACA-O will incorporate carbon footprint and greenhouse gas emission estimates into the multi-objective optimisation. Crop-specific emission intensities (methane from flooded paddy, nitrous oxide from fertiliser application) will be added as penalty terms or hard constraints in the MIP formulation, enabling the generation of allocation plans that balance economic profitability against climate impact - a capability directly aligned with national NDC targets under the Paris Agreement.

4.5 Expansion to Rainfed and Tank-Irrigated Systems

The current architecture is specifically designed for canal-command schemes with formal water quota allocation processes. Future work will adapt ACA-O for rainfed farming systems and minor irrigation schemes (village tanks, check dams), where water availability is stochastic rather than quota-determined. This will require stochastic programming extensions to the optimisation layer - replacing deterministic quota constraints with chance constraints based on rainfall forecast probability distributions - and integration with Sri Lanka's community-based minor irrigation management structures under the Agrarian Development Act.

5. CHAPTER 5: CONCLUSION

This research has presented the design, implementation strategy, and preliminary evaluation of Function 4 - Adaptive Crop and Area Optimization (ACA-O) as a core module of the Integrated Smart Water-Focused Irrigation System (Project ID: 25-26J-520) developed at SLIIT. ACA-O addresses a fundamental and quantifiable planning deficiency in Sri Lankan canal-command agriculture: the persistent paddy-by-default cropping bias that suppresses water productivity, farm profitability, and scheme resilience in the Udawalawe RBMC/LBMC system.

The module delivers a predict-then-optimise service pipeline that fuses heterogeneous data streams - IoT field telemetry via MQTT/Kafka, FAO-56 Kc-ETo hydro-climate budgets, HARTI/DCS market price feeds, satellite remote sensing products, and Irrigation Department quota records - into explainable Top-3 crop recommendations per field with rationales and risk bands, and quota-feasible hectare allocation plans at farmer and canal aggregate levels. The pipeline integrates fuzzy-TOPSIS multi-criteria suitability scoring, XGBoost/LightGBM yield prediction, TFT/LSTM short-horizon price forecasting, Monte Carlo profit simulation, and LP/MIP constrained optimisation within a FastAPI service architecture deployed on Kubernetes with MLflow model governance.

The primary theoretical contribution is the integration of explainable suitability scoring, probabilistic forecasting, and quota-constrained optimisation within a single operational service that supports mid-season Plan-B re-optimisation on allocation or market shocks - a capability absent from all individually reviewed approaches in the fuzzy-TOPSIS, ML classification, remote-sensing phenology, and mathematical programming literatures. Component-level validation results show yield prediction MAE of 0.35–0.48 t/ha, price forecasting MAPE of 8.2–11.6%, expected profit uplift of 18–27% over paddy-by-default baselines, water productivity improvements of 22-31%, API P95 latency of 1,340–1,820ms for recommendations and 3,200–4,100ms for Plan-B, and a preliminary farmer SUS score of 71.4.

ACA-O's significance extends beyond technical performance metrics. By providing farmers with explainable, culturally appropriate, and financially grounded crop recommendations through a bilingual offline-tolerant PWA, it creates the conditions for genuine behavioural change in irrigation planning. By publishing signed water-requirement envelopes for downstream scheduling and providing scheme-level aggregate dashboards, it closes the information loop between individual farm decisions and scheme-level water management. By embedding PDPA compliance, OWASP ASVS security controls, and NIST-aligned platform governance, it satisfies the institutional requirements for government-supported deployment. ACA-O thus represents a significant step toward climate-resilient, market-aware, and equitable irrigation management in Sri Lanka and comparable canal-command systems across the Asia-Pacific region.

REFERENCES

[1] S. Ahmed, S. Yesmin, L. R. Saha, A. K. M. Sadat, and M. H. A. Khan, "Major Cropping Pattern Prediction in Bangladesh from Land, Soil and Climate Data Using Machine Learning Techniques," in *Proc. 25th Int. Conf. on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh, Dec. 17–19, 2022.

[2] E. Neamatollahi, J. Vafabakhshi, M. R. Jahansuz, and F. Sharifzadeh, "Agricultural Optimal Cropping Pattern Determination Based on Fuzzy System," *Fuzzy Information and Engineering*, vol. 9, pp. 479–491, 2017, doi: 10.1016/j.fiae.2017.12.004.

[3] S. Gowtham and R. B. R., "Intelligent Crop Selection for Maximizing Agricultural Productivity: A Fuzzy TOPSIS Perspective," in *Proc. Int. Conf. on Recent Advances in Science & Engineering Technology (ICRASET)*, 2023.

[4] A. Attavar, H. Banu B, A. L. Manohar, P. M. R., and H. D., "IntelliCrop: Major Cropping Pattern Prediction UsingMachine Learning Techniques," in *Proc. 7th Int. Conf. on Inventive Computation Technologies (ICICT 2024)*, 2024.

[5] J. Liu, "A Phenology-Based Cropping Pattern (PBCP) Mapping Method Based on Remotely Sensed Time-Series Vegetation Index Data," *IEEE Transactions on Geoscience and Remote Sensing*, 2019.

[6] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," in *Advances in Neural Information Processing Systems (NIPS 2017)*, vol. 30, 2017.

[7] B. Lim, S. O. Arik, 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, doi: 10.1016/j.ijforecast.2021.03.012.

[8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[9] P. K. Suresh, T. V. Ramachandra, and B. V. Babu, "Water Resource Allocation for Optimal Cropping Pattern Using Linear Programming," *Journal of Irrigation and Drainage Engineering*, vol. 141, no. 4, pp. 04014068, 2015.

[10] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD)*, San Francisco, CA, USA, Aug. 2016, pp. 785–794.

[11] 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*, Food and Agriculture Organization of the United Nations, Rome, 1998.

[12] Hector Kobbekaduwa Agrarian Research and Training Institute (HARTI), "Daily/Weekly Food Commodities Bulletins," 2023–2025. [Online]. Available: https://www.harti.gov.lk/index.php/en/market-information/ . Accessed: Apr. 2026.

[13] Department of Census and Statistics, Sri Lanka, "Retail Prices — Statistical Information," 2025. [Online]. Available: https://www.statistics.gov.lk/InflationAndPrices/StaticalInformation/RetailPrices . Accessed: Apr. 2026.

[14] Parliament of Sri Lanka, *Personal Data Protection Act, No. 9 of 2022*. [Online]. Available: https://www.parliament.lk/uploads/acts/gbills/english/6242.pdf . Accessed: Apr. 2026.

[15] W3C, "Web Content Accessibility Guidelines (WCAG) 2.1," 2018. [Online]. Available: https://www.w3.org/TR/WCAG21/ . Accessed: Apr. 2026.

[16] OWASP Foundation, "OWASP Application Security Verification Standard (ASVS) 4.0," 2021. [Online]. Available: https://owasp.org/www-project-application-security-verification-standard/ . Accessed: Apr. 2026.

[17] National Institute of Standards and Technology (NIST), *Security and Privacy Controls for Information Systems and Organizations (SP 800-53 Revision 5)*, 2020.

[18] Department of Meteorology, Sri Lanka, "Drought Monitoring and 7-day Forecast," 2025. [Online]. Available: https://weather.meteo.gov.lk/ . Accessed: Apr. 2026.