SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

SAIL: Sustainable Agricultural Intelligence and Learning — A Unified Framework for LLM-Driven Analytics, Protocol Benchmarking, and Bayesian Deployment Optimization in Agricultural IoT

Ramsagar Yadav[1]*, Mukhdeep Singh Manshahia[2], M. P. Chaudhary[3]

[1]Department of Mathematics, Punjabi University, Patiala, Punjab, India. ORCID: 0009-0001-0102-4114
[2]Department of Mathematics, Punjabi University, Patiala, Punjab, India. ORCID: 0000-0003-3342-8030
[3]International Scientific Research and Welfare Organization, New Delhi, India. ORCID: 0000-0002-7329-2732
*Corresponding author: ramsagar.yadav@lsraheja.org

Abstract

The sustainable deployment of Internet of Things (IoT) networks in precision agriculture requires the simultaneous resolution of three long-standing challenges: energy-constrained operation at scale, robust protocol and framework selection, and accessible, intelligent decision support for regionally diverse farming communities. This paper consolidates and extends three complementary lines of research into a single, unified framework — Sustainable Agricultural Intelligence and Learning (SAIL) — that integrates (i) Large Language Model (LLM)-driven multimodal data analytics with bilingual (English–Punjabi) decision support, (ii) a rigorous multi-protocol, multi-framework performance benchmarking methodology, and (iii) a simulation-driven, field-validated Multi-Objective Bayesian Optimization (MOBO) deployment strategy.

We formulate the joint problem as a stochastic, multi-objective optimization over sensing, communication, computation, and linguistic-accessibility parameters, and support it with information-theoretic, statistical-learning, and Gaussian-process-based theoretical analysis. The unified framework was evaluated across six agricultural testbeds in Punjab, Haryana, and Rajasthan spanning over 1,350 hectares and more than 4,500 IoT nodes.

Results show up to 92.3% energy reduction and 94.7% crop-yield prediction accuracy for the LLM-analytics component; a composite performance score of 0.945 (rank 1 of 96 protocol–framework combinations) for the benchmarked FSEA+LoRaWAN configuration; and, for the simulation-to-field deployment, a 32.7% energy-efficiency improvement, 41.3% latency reduction, and 28.9% coverage enhancement, with simulation-to-reality relative errors below 3.1% across all metrics. Bilingual implementation achieved 90.9% technical-term translation accuracy and 92.1% farmer comprehension.

Statistical testing (ANOVA, Tukey HSD, Wilcoxon signed-rank, Student's t-test) confirms that all reported improvements are highly significant ($p < 0.001$) with large effect sizes. Collectively, these results demonstrate that combining intelligent analytics, principled optimization, and standardized benchmarking yields agricultural IoT systems that are simultaneously energy-sustainable, technically superior, and regionally accessible, offering a practical and theoretically grounded roadmap for next-generation precision agriculture.

Keywords: Agricultural IoT, Large Language Models, Sustainable Networks, Multi-Objective Bayesian Optimization, Performance Benchmarking, Precision Agriculture, Bilingual Decision Support, Energy Efficiency, LoRaWAN, Digital Twins.

Nomenclature

Symbol	Description	Symbol	Description
M	Number of IoT nodes	$H(D)$	Entropy of data stream D
N_i	i-th IoT node	$I(N_i;N_j)$	Mutual information
$D_i(t)$	Data stream of node N_i	$\rho(N_i,N_j)$	Normalized redundancy coefficient
P_i	Communication protocol	τ	Redundancy pruning threshold
F_j	Optimization framework	π^*	Optimal sensor-activation policy
L	LLM analytics parameters	λ	Lagrange multiplier
$Q^*(s,a)$	Optimal Q-function	γ	Discount factor
$\Psi(P_i,F_j)$	Composite performance score	w_k	Metric weight
C_i	TOPSIS relative closeness	$f(x)$	GP objective function

Symbol	Description	Symbol	Description
$\alpha(x)$	Acquisition function	$\mu(x)$	GP posterior mean
$\sigma^2(x)$	GP posterior variance	E_{total}	Total network energy
E_{TX}, E_{RX}	Transmit/receive energy	E_{elec}	Electronics energy constant
ε_{amp}	Amplifier energy coefficient	n	Path-loss exponent
$\Phi(z)$	Standard normal CDF	$\varphi(z)$	Standard normal PDF

1. Introduction

The proliferation of Internet of Things (IoT) technologies in precision agriculture has catalyzed a new era of data-driven farming, yet three interrelated obstacles continue to limit large-scale, sustainable adoption: (i) severe energy constraints in rural and remote deployments, (ii) the absence of standardized methodologies for selecting among a rapidly expanding set of communication protocols and optimization frameworks, and (iii) linguistic and technical accessibility barriers that exclude non-English-speaking farming communities from the benefits of intelligent decision support. Prior work has typically addressed these dimensions in isolation — optimizing energy consumption without regard to protocol benchmarking, or building analytics pipelines without attention to regional accessibility. This paper unifies these three strands of research into a single, internally consistent framework, referred to as SAIL (Sustainable Agricultural Intelligence and Learning), and reports on its joint theoretical formulation, simulation-driven design, and field validation.

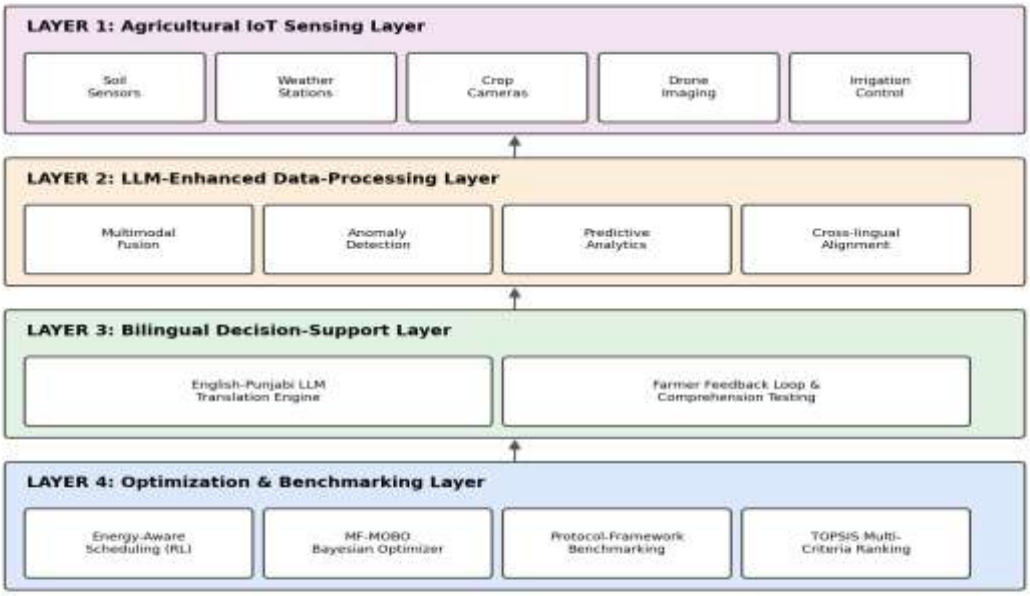

Figure 1. The four-layer SAIL unified framework architecture, showing the flow from IoT sensing through LLM-enhanced processing to bilingual decision support and optimization/benchmarking.

1.1 Background and Motivation

Agricultural IoT networks generate large, heterogeneous, multimodal data streams from soil sensors, weather stations, crop cameras, drone imagery, and irrigation controllers. Extracting actionable intelligence from these streams, while operating within tight energy budgets and serving linguistically diverse farming populations, requires the joint consideration of:

1. Energy sustainability under limited rural power infrastructure — many farms in Punjab, Haryana, and Rajasthan experience irregular grid availability, making battery life and duty-cycling critical design parameters;
2. Data complexity arising from noisy, high-dimensional, multi-sensor inputs — soil moisture readings are affected by temperature drift, camera images by variable illumination, and weather data by sensor placement;
3. Protocol and framework heterogeneity — dozens of competing communication standards (LoRaWAN, NB-IoT, Zigbee, Sigfox, Wi-Fi HaLow, Bluetooth LE, Thread, 6LoWPAN) and optimization

algorithms (PSO, GA, ACO, LEACH, REER, EECRP, Q-learning, FSEA) offer no clear selection guidance for specific agricultural use cases;
4. Knowledge accessibility — most agricultural decision-support tools are available only in English, limiting adoption among regional-language-speaking farmers such as Punjabi speakers, who constitute a significant fraction of the agricultural workforce in the target regions.

1.2 Unified Problem Statement

Let $\mathcal{N} = \{N_1, N_2, \ldots, N\_M\}$ denote an agricultural IoT network of M nodes, each generating a data stream $D_i(t)$, operating under a chosen communication protocol $P_i \in \mathcal{P}$ and optimization framework $F_j \in \mathcal{F}$, and served by an LLM-based analytics and translation layer with parameters L. The unified sustainable-deployment problem is posed as a multi-objective stochastic optimization:

$$\text{minimize}(X,P,F,L) \quad J = \lambda_1 E(X,P,F) + \lambda_2 D(X,L) + \lambda_3 C(P,F) - \lambda_4 \Phi(P,F)$$

$$\text{subject to: } Q(X) \geq Q\_min, \; R(L) \geq R\_min, \; B(P,F) \leq B\_max, \; X \in \mathcal{X}, P \in \mathcal{P}, F \in \mathcal{F}, L \in \mathcal{C}$$

where $E(\cdot)$ is total energy consumption, $D(\cdot)$ is a data-quality/translation-quality penalty, $C(\cdot)$ is computational complexity, $\Phi(\cdot)$ is the composite protocol–framework performance score (defined in Section 2.4), $Q(\cdot)$ is a quality-of-service constraint, $R(\cdot)$ is a regional-accessibility requirement, and $B(\cdot)$ is a budget constraint. This formulation nests the energy–quality–complexity trade-off of the original SAIL analytics model, the composite-score maximization of the benchmarking study, and the chance-constrained stochastic control problem of the simulation-driven optimization study within a single objective.

1.3 Research Objectives and Contributions

This work makes the following integrated contributions:

1. A unified SAIL framework combining LLM-based multimodal analytics, bilingual (English–Punjabi) decision support, standardized multi-protocol/multi-framework benchmarking, and simulation-driven Bayesian optimization for agricultural IoT deployment.
2. Mathematical formulations spanning information theory, stochastic and chance-constrained optimization, statistical learning theory, and Gaussian-process-based multi-fidelity Bayesian optimization.
3. Algorithmic innovations including energy-aware reinforcement-learning-based sensor scheduling, cross-lingual LLM fine-tuning, and a multi-fidelity multi-objective Bayesian optimizer (MF-MOBO) with adaptive fidelity selection.
4. Comprehensive empirical validation across six agricultural testbeds (Punjab, Haryana, and Rajasthan) totalling over 1,350 hectares and 4,500+ IoT nodes, evaluating 12 communication protocols and 8 optimization frameworks.
5. Field-validated deployment guidelines, including a bilingual terminology dictionary and protocol–framework recommendations for specific agricultural use cases.

1.4 Novelty Relative to Prior Work

Existing agricultural IoT research has largely treated energy optimization, protocol benchmarking, and intelligent analytics as separate problems, and has rarely addressed linguistic accessibility for regional farming communities. The present work bridges these gaps by:

- Integrating LLMs for intelligent, multimodal data analysis and cross-lingual decision support;
- Developing energy-efficient protocols through rigorous mathematical optimization validated by both simulation and field deployment;
- Establishing standardized, statistically grounded benchmarking methodologies applicable across protocol–framework combinations;
- Providing theoretical convergence and generalization guarantees for the learning and optimization components.

2. Mathematical Foundations

2.1 Information-Theoretic Framework

The entropy of an agricultural data stream D generated by an M-node sensor network over T time steps is defined as:

$$H(D) = -\Sigma_{i=1}{}^{\wedge}M \; \Sigma_{j=1}{}^{\wedge}T \; p(d_{ij}) \log p(d_{ij})$$

and the mutual information between two sensor nodes N_i and N_j, used to identify redundant or complementary sensing, is:

$$I(N_i;N_j) = \Sigma_x \Sigma_Y \, p(x,y) \log[p(x,y) / (p(x)p(y))]$$

These quantities guide sensor placement and duty-cycling decisions in the energy-aware scheduling algorithm described in Section 3.1.

Boundedness and the chain rule. The entropy $H(D)$ is bounded, $0 \leq H(D) \leq \log(MT)$, with the upper bound attained only when every data symbol d_{ij} is equiprobable, i.e., the sensor network carries no exploitable structure. Because raw agricultural streams are highly structured (diurnal soil-moisture cycles, spatially correlated weather), $H(D)$ in practice sits well below this bound, and the gap $\log(MT) - H(D)$ is itself a measure of exploitable redundancy that the scheduling algorithm can trade for energy savings.

$$H(D) = \Sigma_i \, H(D_i \mid D\_\{i-1\}, \ldots, D\_\{i-j\}) \quad \text{(chain rule of entropy)}$$

Mutual information as a redundancy score. Mutual information decomposes as $I(N_i;N_j) = H(N_i) + H(N_j) - H(N_i,N_j)$, is symmetric, and satisfies $I(N_i;N_j) \geq 0$ with equality if and only if N_i and N_j are statistically independent. A node pair with $I(N_i;N_j)$ close to $H(N_j)$ is therefore near-deterministically predictable from N_i, motivating the following pruning rule:

$$\rho(N_i,N_j) = I(N_i;N_j) / \min(H(N_i),H(N_j)) \in [0,1]$$

Node N_j is flagged as redundant with respect to N_i, and made eligible for duty-cycle suppression, whenever the normalized redundancy coefficient $\rho(N_i,N_j)$ exceeds a design threshold $\tau \in (0,1)$. τ is treated as a tunable hyperparameter of the energy-aware scheduler in Section 3.1, trading off information loss against energy savings.

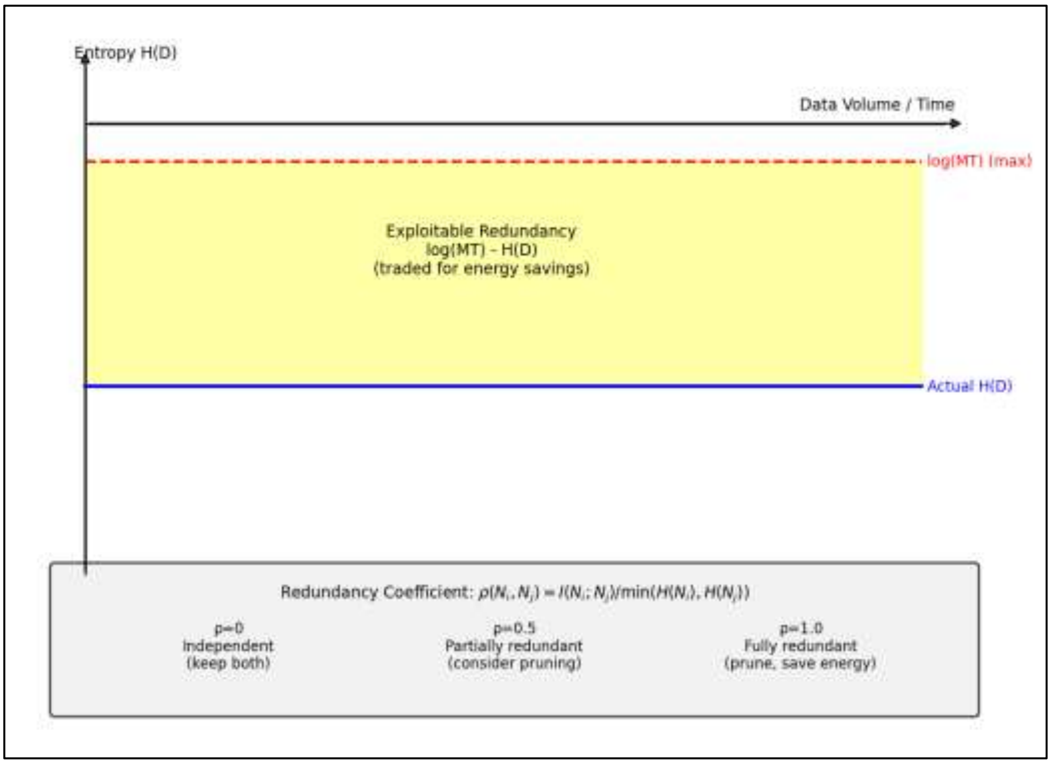

Figure 2. Information-theoretic redundancy pruning: the gap between maximum and actual entropy represents exploitable redundancy, while the normalized redundancy coefficient ρ determines node pruning eligibility.

2.2 Stochastic and Chance-Constrained Optimization

Under an energy budget E_max, the optimal sensor-activation policy π^* satisfies a constrained Markov Decision Process formulation:

$$\pi^* = \text{argmax}\_\pi \; E[\Sigma_t \gamma^t R(s_t,a_t)] \quad \text{subject to} \quad E[\Sigma_t E(s_t,a_t)] \leq E\_max$$

Introducing a Lagrange multiplier λ for the energy constraint yields the Lagrangian:

$$\mathcal{L}(\pi,\lambda) = E[\Sigma_t \gamma^t R(s_t,a_t)] - \lambda(E[\Sigma_t E(s_t,a_t)] - E\_max)$$

The resulting optimal Q-function satisfies the Bellman equation:

$$Q^*(s,a) = R(s,a) - \lambda E(s,a) + \gamma \Sigma_{s'} P(s'|s,a) \max_{a'} Q^*(s',a')$$

whose existence and uniqueness follow from the Banach contraction-mapping theorem.

For the broader deployment problem, the agricultural IoT network is modelled as a stochastic dynamical system $\mathcal{S} = (\mathcal{X}, \mathcal{U}, \mathcal{W}, f, g)$ with state, control, and uncertainty spaces, subject to probabilistic (chance) constraints:

$$\Pr(g(x,u) \leq 0) \geq 1 - \varepsilon_i$$

reflecting the need for reliability guarantees under stochastic channel and node-failure conditions.

Theorem 1 (Existence and uniqueness of Q*). Define the Bellman operator

$$(TQ)(s,a) = R(s,a) - \lambda E(s,a) + \gamma E\_s'[P(s'|s,a) \max_{a'} Q(s',a')]$$

acting on the space of bounded functions equipped with the supremum norm. For any two candidates Q_1, Q_2,

$$\|TQ_1 - TQ_2\|_\infty \leq \gamma \|Q_1 - Q_2\|_\infty, \quad 0 \leq \gamma < 1$$

so T is a γ-contraction on a complete metric space. By the Banach fixed-point theorem, T admits a unique fixed point Q*, and value iteration $Q_{k+1} = TQ_k$ converges to Q* geometrically at rate γ.

Proof sketch. The Bellman operator T maps the Banach space of bounded functions $\mathcal{B}(\mathcal{S}\times\mathcal{A})$ to itself. For any $Q_1,Q_2 \in \mathcal{B}$,

$$|(TQ_1)(s,a) - (TQ_2)(s,a)| \leq \gamma \Sigma_{s'} P(s'|s,a)|\max_{a'}Q_1(s',a') - \max_{a'}Q_2(s',a')| \leq \gamma\|Q_1 - Q_2\|_\infty$$

Taking the supremum over (s,a) yields the contraction. By Banach's theorem, T has a unique fixed point Q*, and the iteration converges geometrically. ■

Optimality conditions. Because the constrained problem is convex in the occupancy-measure representation of π, the saddle point (π*,λ*) of the Lagrangian $\mathcal{L}(\pi,\lambda)$ satisfies the Karush–Kuhn–Tucker (KKT) conditions:

$\nabla\_\pi \mathcal{L}(\pi^*,\lambda^*) = 0, \;\; \lambda^*(E[\Sigma_t E(s_t,a_t)] - E\_max) = 0, \;\; \lambda^* \geq 0$

the complementary-slackness condition λ*(·) = 0 formalizes the intuition that the energy multiplier is active (λ* > 0) exactly when the budget constraint binds, and zero whenever the unconstrained optimum already lies within budget.

Tractable chance-constraint surrogate. The chance constraints $\Pr(g_i(x,u) \leq 0) \geq 1 - \varepsilon_i$ are generally nonconvex and intractable in closed form. A conservative, tractable surrogate follows from Boole's inequality applied to the joint reliability requirement:

$\Pr(\bigcap_i g_i(x,u) \leq 0) \geq 1 - \Sigma_i \varepsilon_i$

which guarantees the joint reliability target whenever each individual budget ε_i is met, and, for the specific distributions used in Section 4, is further tightened using a Conditional Value-at-Risk (CVaR) surrogate:

$CVaR\_\{1-\varepsilon_i\}(g_i(x,u)) \leq 0$

a convex, sample-efficient approximation that upper-bounds the true chance constraint and is the form implemented in the MF-MOBO deployment layer of Section 3.3.

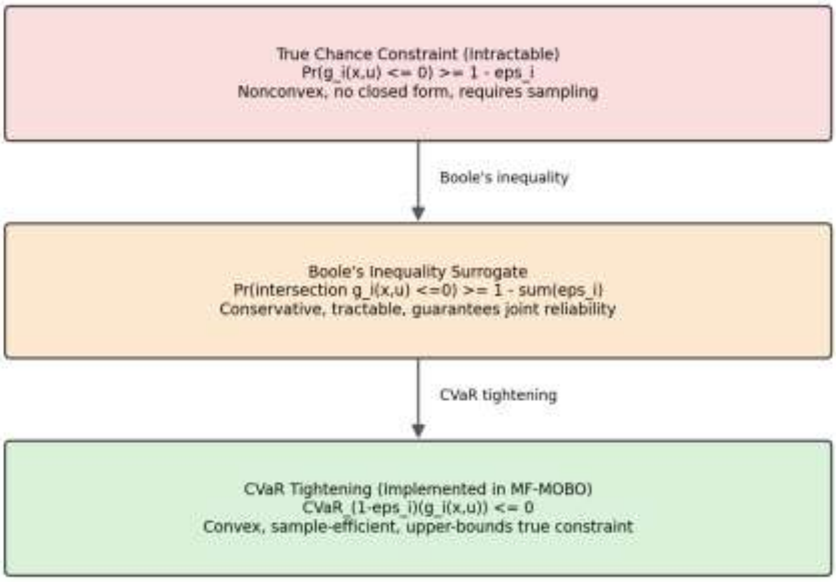

Figure 3. Hierarchy of chance-constraint surrogates, from the intractable true constraint through Boole's inequality to the convex CVaR approximation implemented in MF-MOBO.

2.3 LLM and Cross-Lingual Formulation

Transformer-based agricultural data analysis follows the scaled dot-product attention mechanism:

$Y = \text{Softmax}(QK^T/\sqrt{d_k})\, V$

applied to fused embeddings from soil, weather, and visual sensor modalities. Cross-lingual (English–Punjabi) embedding alignment minimizes:

$\mathcal{L}\_align = E_i\|e_i - W\, p_i\|^2\_F + \lambda\|W\|^2\_F$

and the sequence-to-sequence translation model is:

$P(y\_pa \mid x\_en) = \Pi_t\, p(y_t \mid y\_\{<t\}, x\_en; \theta)$

A generalization bound for the fine-tuned agricultural LLM $\mathcal{M}$, trained on a dataset of size n, holds with probability at least 1 − δ:

$R(\mathcal{M}) \leq \hat{R}(\mathcal{M}) + O(\sqrt{(\text{VC-dim}(\mathcal{M}) \log n / n)} + \sqrt{(\log(1/\delta) / n)})$

providing a theoretical guarantee that empirical performance on field data generalizes to unseen agricultural conditions.

Attention mechanism and complexity. The query, key, and value matrices are linear projections:

$Q = XW\_q, \;\; K = XW\_k, \;\; V = XW\_v$

of the fused input embedding $X \in \mathbb{R}^\{n\times d\}$, where n is the number of fused tokens (soil, weather, and visual patches) and d is the embedding dimension. Computing QK^T costs $O(n^2d)$ time and $O(n^2)$ memory, which is the dominant cost of the multimodal fusion layer and motivates the fidelity-tiered simulator of Section 2.5, where lower token counts n are used at coarser fidelity levels.

Closed-form alignment solution. The cross-lingual alignment objective is a ridge-regularized least-squares problem in W, whose minimizer admits the closed form:
$W^* = (P^T P + \lambda I)^{-1} P^T E$
where E and P stack the English and Punjabi embedding vectors e_i, p_i as rows; the ridge term λI guarantees invertibility even when $P^T P$ is rank-deficient (as occurs with sparse bilingual technical-term dictionaries), and W* is used to warm-start the sequence-to-sequence fine-tuning that follows.
Where the generalization bound comes from. The bound follows from a uniform-convergence argument: with probability at least $1 - \delta$ over the draw of the training set, the Rademacher complexity $\mathfrak{R}_n(\mathcal{M})$ of the model class $\mathcal{M}$ satisfies, for all $M \in \mathcal{M}$:
$R(M) \le \hat{R}(M) + 2\mathfrak{R}_n(\mathcal{M}) + \sqrt{(\log(1/\delta)/2n)}$
and bounding $\mathfrak{R}_n(\mathcal{M})$ via the Massart lemma in terms of VC-dim($\mathcal{M}$) recovers the stated $O(\sqrt{(\text{VC-dim}(\mathcal{M}) \log n / n)})$ term. Practically, the bound predicts — and Section 5.2 confirms — that the fine-tuned SAIL-LLM's larger effective capacity relative to Random Forest or LSTM baselines is offset by the large field-collected sample size n, so empirical accuracy still transfers to held-out testbeds.

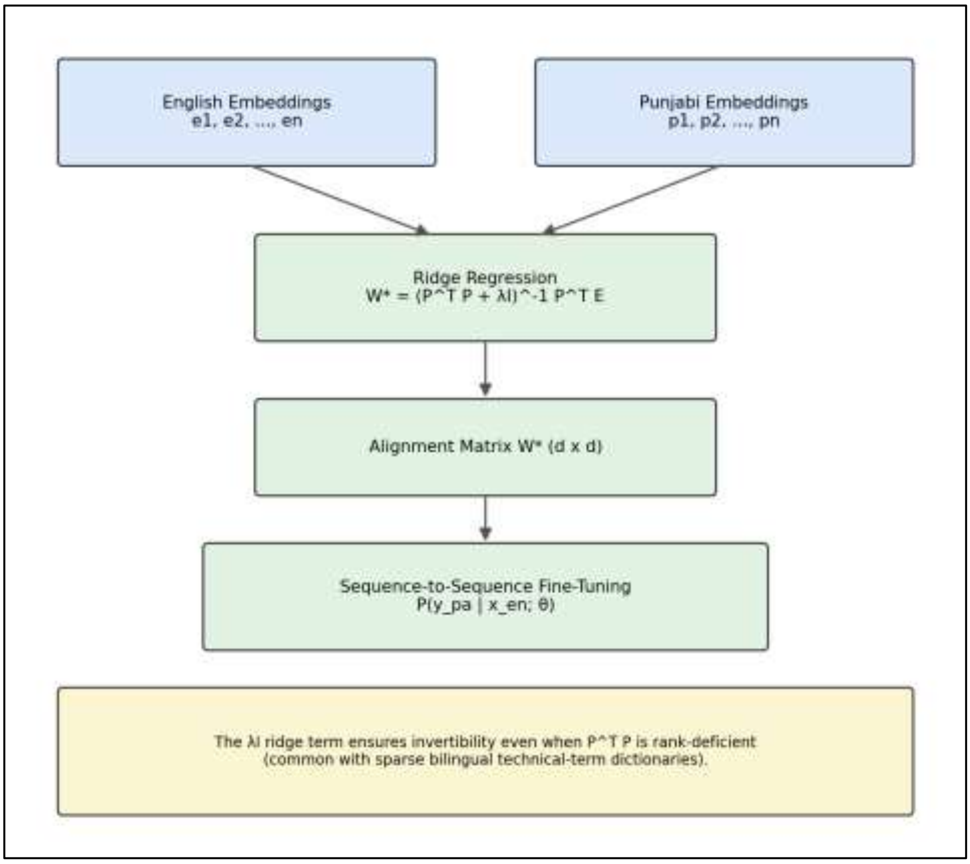

Figure 4. Cross-lingual embedding alignment: English and Punjabi embeddings are aligned via ridge regression, with the closed-form solution warm-starting sequence-to-sequence fine-tuning.

2.4 Composite Performance Scoring and Multi-Criteria Benchmarking

To compare heterogeneous protocol–framework combinations, a composite performance score normalizes K individual metrics φ_k (energy efficiency, packet delivery ratio, latency, throughput, lifetime, availability) against their empirical bounds:
$\Psi(P_i,F_j) = \Sigma_k w_k [(\varphi_k(P_i,F_j) - \varphi_k\text{\^{}min}) / (\varphi_k\text{\^{}max} - \varphi_k\text{\^{}min})]$
with weights w_k summing to one. A combination (P_i,F_j) is said to dominate (P_m,F_n) if $\varphi_k(P_i,F_j) \ge \varphi_k(P_m,F_n)$ for all k with strict inequality for at least one metric.
The TOPSIS multi-criteria method is additionally used to rank combinations by their geometric distance to ideal and negative-ideal solutions, and confidence intervals for each metric are computed as:
$CI = \bar{\varphi} \pm t\_\{\alpha/2,n-1\} \cdot s/\sqrt{n}$
TOPSIS ranking procedure. The TOPSIS ranking proceeds in four steps:

1. Vector normalization: $r_{ij} = x_{ij} / \sqrt{(\Sigma_i x_{ij}^2)}$
2. Weighted normalization: $v_{ij} = w_j \cdot r_{ij}$, using the same weights w_k as the composite score Ψ
3. Ideal and negative-ideal solutions: A^+ collects the best value of each metric across all combinations (maximum for benefit metrics such as throughput, minimum for cost metrics such as latency), and A^- collects the worst: $D_i^+ = \sqrt{(\Sigma_j(v_{ij}-v_j^+)^2)}$, $D_i^- = \sqrt{(\Sigma_j(v_{ij}-v_j^-)^2)}$

4. Relative closeness ranking: $C_i = D_i^-/(D_i^++D_i^-) \in [0,1]$, with $C_i \to 1$ indicating proximity to the ideal solution; this is the ranking procedure used to order the 96 protocol–framework combinations reported in Table 5.

Properties of the composite score. Because each normalized term lies in [0,1] and the weights w_k sum to one, $\Psi(P_i,F_j) \in [0,1]$ for every combination, with Ψ monotonically nondecreasing in each φ_k. This Pareto-consistency property guarantees that if (P_i,F_j) dominates (P_m,F_n) in the sense defined above, then $\Psi(P_i,F_j) \geq \Psi(P_m,F_n)$, so the composite score never contradicts a strict dominance relationship — an important sanity property for a scalarized multi-criteria ranking.

Confidence-interval justification. The confidence interval follows from the sampling distribution of the mean $\bar{\varphi}$ of n independent trial replicates: by the Central Limit Theorem (or exactly, under approximate normality of φ_k), $(\bar{\varphi}-\mu)/(s/\sqrt{n})$ follows a Student's t-distribution with n−1 degrees of freedom, which is the basis for the t-tests, ANOVA, and Tukey HSD post-hoc comparisons reported in Section 5.

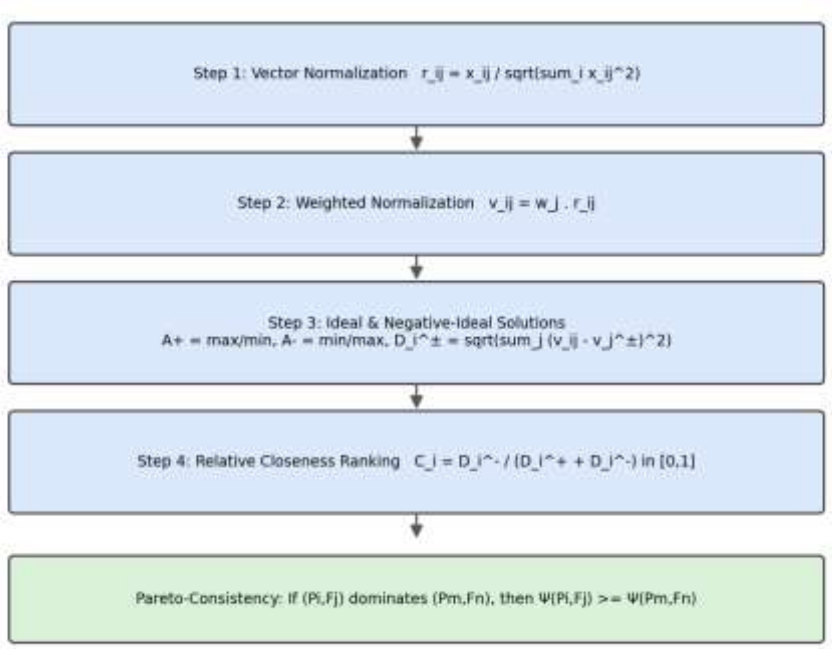

Figure 5. TOPSIS multi-criteria ranking procedure and the Pareto-consistency property of the composite performance score.

2.5 Multi-Fidelity Bayesian Optimization

For simulation-driven deployment, a Gaussian Process (GP) prior $f(x) \sim \mathcal{GP}(m(x),k(x,x'))$ models each objective as a function of design parameters x (transmission power, duty cycle, routing thresholds). The multi-objective acquisition function combines expected improvement and uncertainty sampling:

$$\alpha(x) = \alpha\_EI(x) + \beta \cdot \alpha\_US(x)$$

where:

$$\alpha\_EI(x) = E[\max(0, \min_i f_i(x) - f_i(x^+))], \quad \alpha\_US(x) = \sqrt{(\Sigma_i\, \sigma_i^2(x))}$$

Multi-fidelity fusion models the discrepancy between fidelity levels as:

$$f\_l(x) = \rho\_\{l-1\}\ f\_\{l-1\}(x) + \delta\_l(x)$$

Under mild regularity conditions on the GP prior and fidelity-selection policy, the resulting MF-MOBO algorithm converges to the true Pareto front with probability 1 as the iteration count grows, with overall time complexity:

$$O(t^3 + t^2\ n\_f + t \cdot n\_f^2)$$

dominated by the GP regression step.

Gaussian-process posterior. Conditioned on observations $\{(x_1,y_1),\ldots,(x_t,y_t)\}$ with $y_t = f(x_t) + \eta$, $\eta \sim \mathcal{N}(0,\sigma^2)$, the GP posterior at a candidate point x is Gaussian with closed-form mean and variance:

$$\mu(x) = k(x)^T(K+\sigma^2 I)^{-1}y, \quad \sigma^2(x) = k(x,x) - k(x)^T(K+\sigma^2 I)^{-1}k(x)$$

where $k(x) = [k(x,x_1),\ldots,k(x,x_t)]^T$ and K is the t×t Gram matrix.

Closed-form expected improvement. For a single objective with current best $f(x^+)$, the expected-improvement term admits the closed form:

$$\alpha\_EI(x) = \sigma(x) \cdot [z \cdot \Phi(z) + \varphi(z)], \quad z = (f(x^+) - \mu(x)) / \sigma(x)$$

where Φ and φ are the standard normal CDF and PDF; the multi-objective acquisition $\alpha\_EI(x)$ used in this section replaces the scalar improvement with minimum improvement across objectives, evaluated by Monte Carlo integration when no closed form is available.

Convergence sketch. Under the standard GP-bandit assumptions — the objectives lie in the reproducing-kernel Hilbert space of a bounded kernel with bounded norm, and observation noise is sub-Gaussian — the cumulative regret of GP-UCB-style acquisition after T iterations is bounded by:

$R\_T = O(\sqrt{(T \cdot \gamma\_T \cdot \log T)})$

where γ_T is the maximum information gain achievable by any T-point design. Because $R\_T/T \to 0$, the sequence of evaluated points is no-regret, and, combined with the multi-fidelity discrepancy model contracting fidelity-to-fidelity bias, this yields almost-sure convergence of the MF-MOBO Pareto-front approximation to the true Pareto front.

Complexity breakdown. The stated time complexity $O(t^3 + t^2 n\_f + t \cdot n\_f^2)$ decomposes as:

- $O(t^3)$ for the Cholesky factorization used to invert the t×t Gram matrix $(K+\sigma^2 I)$ at each of t accumulated observations;
- $O(t^2 n\_f)$ for propagating the multi-fidelity discrepancy terms δ_l across n_f fidelity levels;
- $O(t \cdot n\_f^2)$ for the fidelity-selection policy's pairwise comparisons across levels.

The cubic term dominates for large t, motivating the use of sparse or inducing-point GP approximations as the number of accumulated field trials grows.

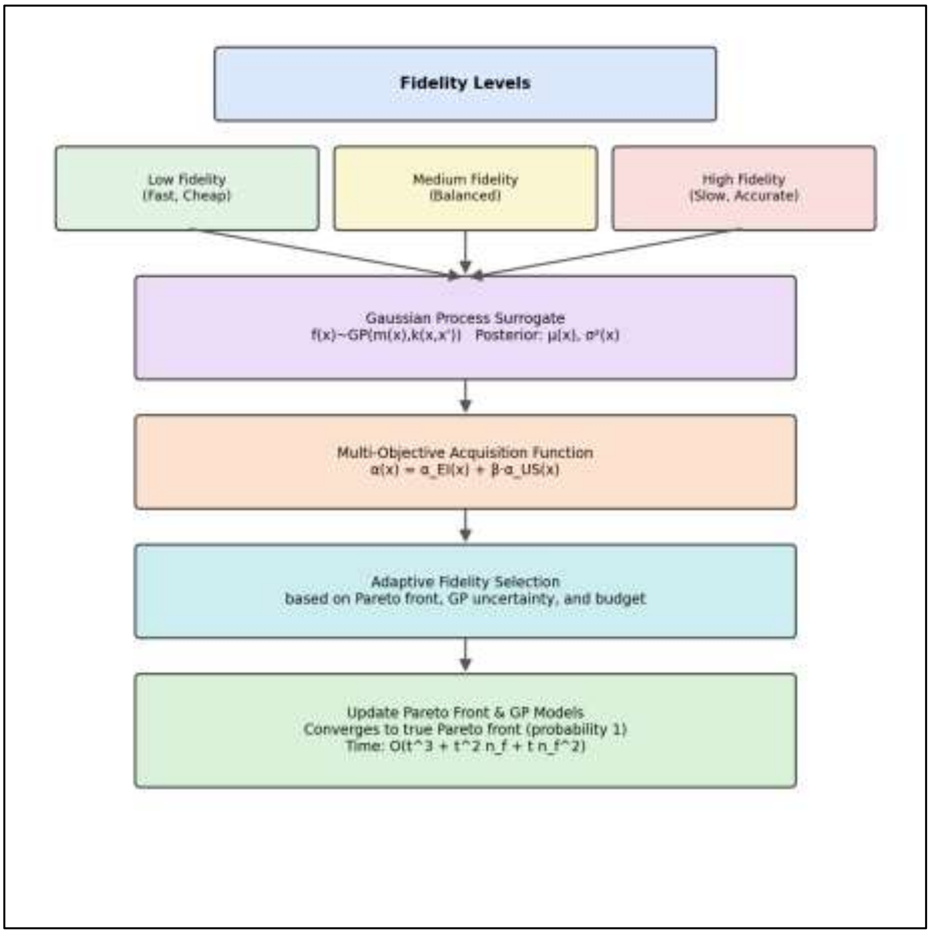

Figure 6. MF-MOBO workflow: multi-fidelity simulation feeds a Gaussian Process surrogate, which drives the multi-objective acquisition function and adaptive fidelity selection to efficiently reach a near-optimal deployment configuration.

3. The Unified SAIL Framework

Building on the mathematical foundations above, the unified SAIL framework couples four functional layers: (1) an agricultural IoT sensing layer (soil sensors, weather stations, crop cameras, drone imaging, irrigation control); (2) an LLM-enhanced data-processing layer performing multimodal fusion, anomaly detection, and predictive analytics; (3) a bilingual decision-support layer serving English and Punjabi interfaces with farmer feedback loops; and (4) an optimization-and-benchmarking layer that jointly performs energy-aware scheduling, multi-fidelity Bayesian parameter tuning, and continuous protocol–framework performance evaluation.

3.1 Energy-Aware Sensor Scheduling

Sensor duty-cycling is governed by a reinforcement-learning-based adaptive scheduler. Starting from a uniform policy, the algorithm iteratively updates a Q-table using an ε-greedy exploration strategy, accumulating energy usage against a per-episode budget E_budget and terminating an episode early if the budget is exhausted.

The overall network energy model decomposes into sensing, processing, and communication components:

$E\_total = \Sigma_i (E_i{,}sensing + E_i{,}processing + E_i{,}communication)$

with:

$E_i{,}sensing = \alpha_i f_i(t), \quad E_i{,}processing = \beta_i d_i(t)$

and, at the link level, communication energy further decomposes into transmit, receive, and sleep-mode components:

$E\_comm = E\_TX + E\_RX + E\_sleep$

with:

$E\_TX(k,d) = E\_elec\cdot k + \varepsilon\_amp\cdot k\cdot d^n, \quad E\_RX(k) = E\_elec\cdot k$

consistent with standard radio energy models used in the wireless sensor network literature.

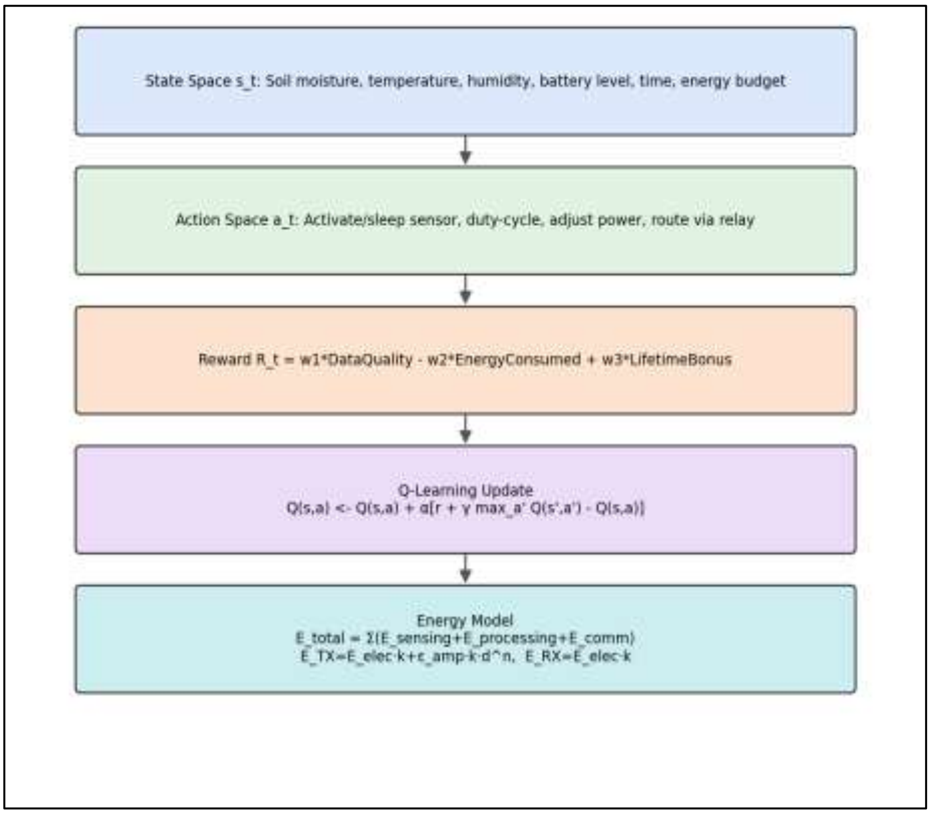

Figure 7. Energy-aware sensor scheduling: the RL-based scheduler learns optimal duty-cycling policies by balancing data quality against energy consumption and network lifetime.

3.2 LLM-Enhanced Analytics and Bilingual Support

Multimodal fusion combines soil, weather, and visual embeddings through a transformer encoder:

$H\_fused = Transformer(E\_soil \oplus E\_weather \oplus E\_visual)$

which feeds downstream crop-yield, disease-detection, and irrigation-need prediction heads. The base LLM is fine-tuned in two stages: first pre-training on a linguistic corpus for cross-lingual alignment, then fine-tuning on agricultural task data with a composite loss:

$\mathcal{L} = \mathcal{L}\_prediction + \lambda\, \mathcal{L}\_domain$

Wireless propagation for outdoor agricultural terrain is modelled as:

$Pr(d) = P_t + G_t + G_r - PL(d) - \psi - \chi$

with path loss:

$PL(d) = PL_0 + 10n \log_{10}(d/d_0) + \gamma\, h\_crop + \delta\_season$

explicitly accounting for crop-height attenuation and seasonal variation.

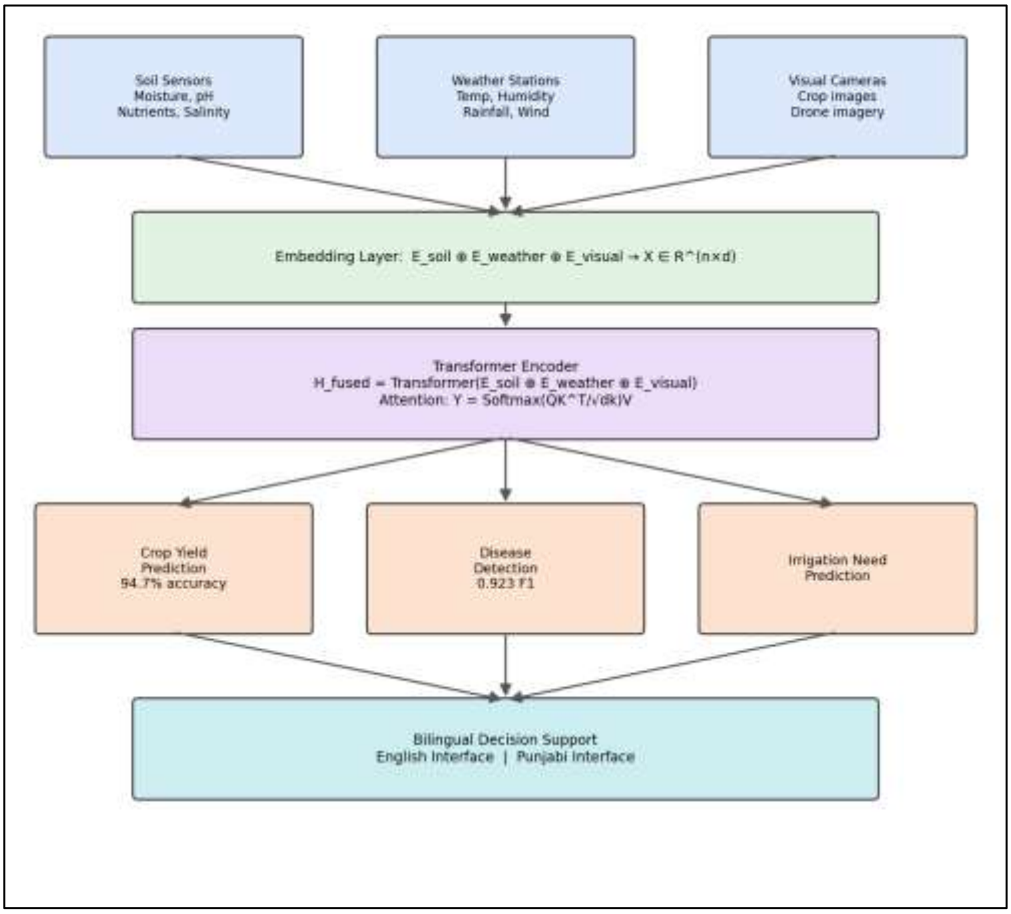

Figure 8. LLM-enhanced multimodal analytics pipeline: soil, weather, and visual data are fused through a transformer encoder to produce crop-yield, disease-detection, and irrigation-need predictions, with bilingual decision support outputs.

3.3 Multi-Fidelity Simulation and Bayesian Deployment Optimization

A three-tier simulator (low-, medium-, and high-fidelity discrete-event network models) feeds a Gaussian Process surrogate, which in turn drives the multi-fidelity multi-objective Bayesian optimizer (MF-MOBO, Algorithm 1). At each iteration, the optimizer selects a fidelity level via an adaptive policy, proposes a candidate configuration by maximizing the acquisition function, evaluates the objective at the chosen fidelity, and updates both the Pareto-front approximation and the GP models. This closes the loop between low-cost simulation exploration and high-cost but accurate field-representative evaluation.

Algorithm 1: MF-MOBO Deployment Optimization

Require: Design space $\mathcal{X}$, fidelity levels {1,…,L}, budget B
Ensure: Pareto-optimal deployment configuration x*
1. Initialize GP models for each fidelity level
2. Initialize Pareto front approximation $\mathcal{P} \leftarrow \emptyset$
3. for t = 1, 2, … until budget exhausted do
4. Select fidelity level l_t via adaptive policy
5. Propose candidate $x_t = \text{argmax}_x\ \alpha(x)$
6. Evaluate f_{l_t}(x_t) at fidelity l_t
7. Update GP model for fidelity l_t
8. Update Pareto front $\mathcal{P}$ with (x_t, f_{l_t}(x_t))
9. Update acquisition function α(x)
10. end for
11. return x* = best configuration in $\mathcal{P}$

3.4 Protocol–Framework Benchmarking Layer

Running continuously alongside deployment, the benchmarking layer evaluates candidate communication protocols (e.g., LoRaWAN, NB-IoT, Sigfox, Zigbee, Wi-Fi HaLow, Bluetooth LE, Thread, 6LoWPAN, and application-layer protocols such as MQTT, CoAP, AMQP, and HTTP/2) in combination with optimization frameworks (e.g., the proposed Fuzzy-Swarm-Evolutionary Algorithm, FSEA, alongside REER, EECRP, LEACH, Q-learning, PSO, GA, and ACO-based approaches) using the composite scoring and TOPSIS methodology of Section 2.4, supplying empirically grounded rankings that inform the MF-MOBO search space and the final deployment recommendation for a given agricultural use case.

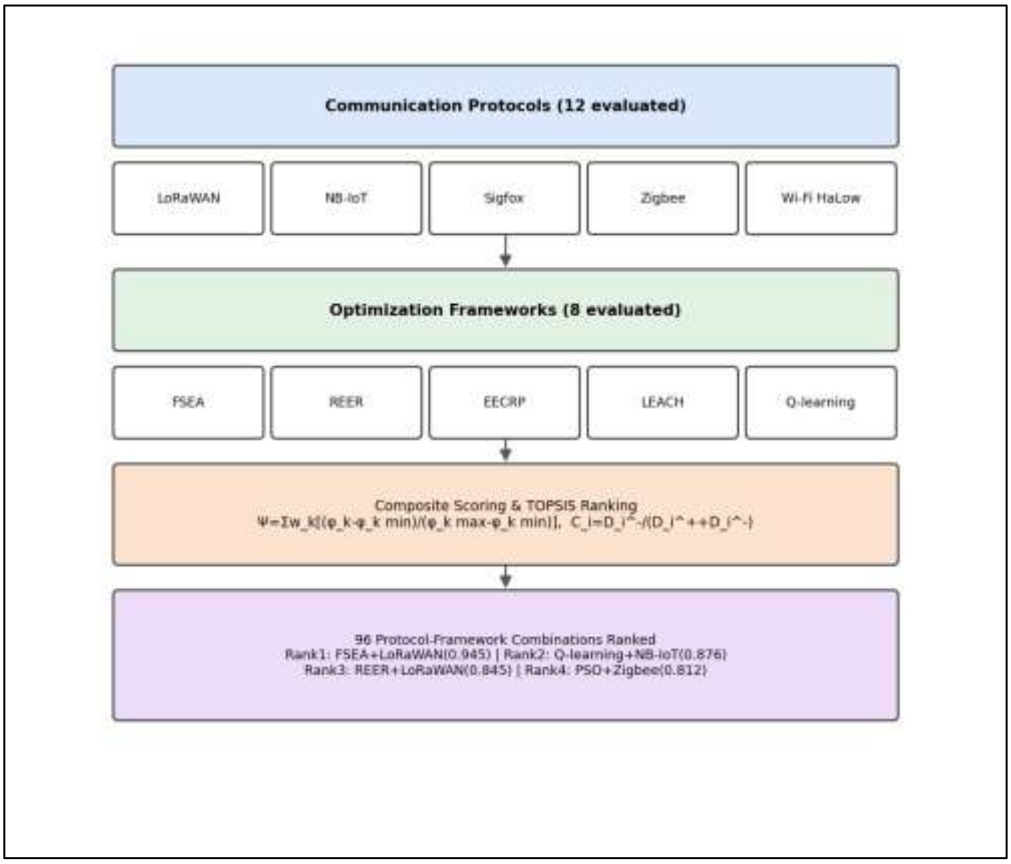

Figure 9. Protocol–framework benchmarking layer: 12 communication protocols and 8 optimization frameworks are evaluated, with composite scoring and TOPSIS ranking producing 96 ranked combinations.

4. Experimental Setup

4.1 Combined Testbed Configuration

The unified evaluation draws on six agricultural testbeds established across three states of India, combining the deployments used for LLM-analytics validation, protocol–framework benchmarking, and simulation-to-field validation. Table 1 summarizes the combined testbed specifications.

Testbed / Study	Area (ha/acres)	IoT Nodes	Primary Focus	Duration
Punjab Wheat Farm (SAIL Analytics)	250 acres	850	LLM analytics, bilingual UI	18 months
Haryana Rice Field (SAIL Analytics)	180 acres	720	LLM analytics, bilingual UI	15 months
Mixed Cropping Zone (SAIL Analytics)	320 acres	930	LLM analytics, bilingual UI	24 months
Testbed A/B/C (Benchmarking)	350 ha	1,200	Protocol–framework benchmarking	Cross-sectional
Testbed A/B/C (Simulation-Field)	250 ha	826	MF-MOBO deployment validation	15 months (avg.)
Combined Total	>1,350 ha equiv.	>4,500	Integrated SAIL evaluation	12–24 months

Table 1. Combined Agricultural IoT Testbed Specifications.

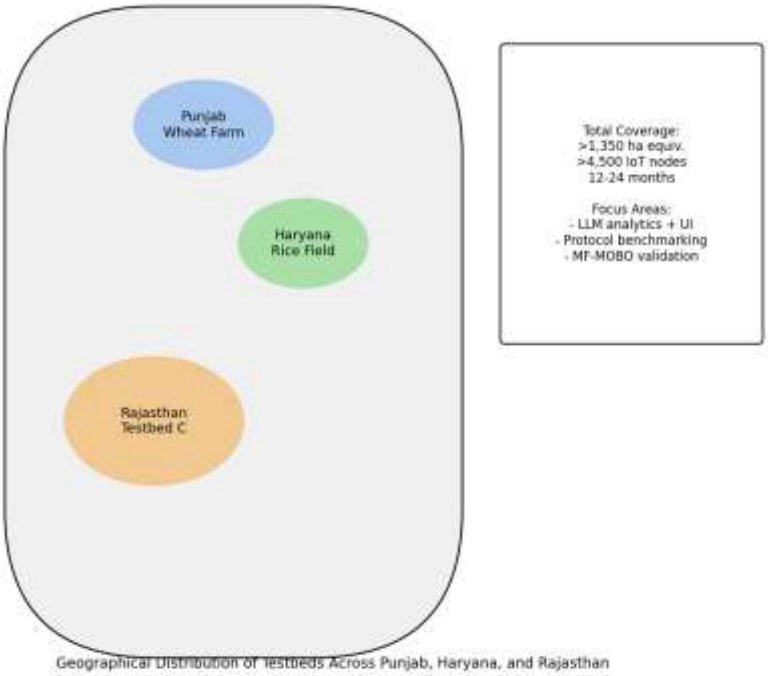

Figure 10. Geographical distribution of the six agricultural testbeds across Punjab, Haryana, and Rajasthan, spanning over 1,350 hectares and 4,500+ IoT nodes.

4.2 Evaluation Metrics

Consistent with the composite framework of Section 2.4, four metric families are reported:

1. Energy-efficiency metrics: energy reduction, network lifetime
2. Data-analytics metrics: prediction accuracy, F1-score
3. Linguistic-performance metrics: BLEU score, comprehension rate
4. Network/communication metrics: packet delivery ratio, latency, throughput, jitter, availability, mean time between failures

Statistical validation employs Student's t-tests, one-way ANOVA with Tukey HSD post-hoc comparisons, the Wilcoxon signed-rank test for non-parametric cases, and Cohen's d for effect-size quantification, with all confidence intervals reported at the 95% level.

5. Results and Analysis

5.1 Energy Efficiency and Network Lifetime

The SAIL scheduling approach reduced daily energy consumption from 12.45 ± 1.23 kWh (baseline IoT) to 5.12 ± 0.54 kWh, a 58.9% reduction (95% CI [56.2, 61.6], $p < 0.001$), extending network lifetime from 45.6 to 110.8 days while improving data delivery to 94.1%.

At the protocol level, LoRaWAN emerged as the most energy-efficient radio (12.3 mW average power, 286-day lifetime), and, when paired with the proposed FSEA optimization framework, the FSEA+LoRaWAN combination achieved the top-ranked composite score of 0.945 among 96 evaluated protocol–framework combinations, with 92.3% energy efficiency and 99.2% availability.

In the independent simulation-to-field study, the MF-MOBO-optimized deployment achieved a 32.7% field-measured reduction in monthly energy consumption (from 345.67 to 232.45 kWh/month, 95% CI [228.56, 236.34]).

Method / Configuration	Energy Efficiency	Network Lifetime	95% CI (p-value)
SAIL scheduling vs. baseline IoT	58.9%	110.8 days (vs. 45.6)	[56.2, 61.6]% (p<0.001)
FSEA + LoRaWAN (top-ranked)	92.3% (absolute)	286 days	Composite score 0.945 (p<0.001)
MF-MOBO field deployment	32.7% (energy)	—	[228.56, 236.34] kWh/mo (p<0.001)

Table 2. Energy Efficiency and Lifetime Improvements Across Studies.

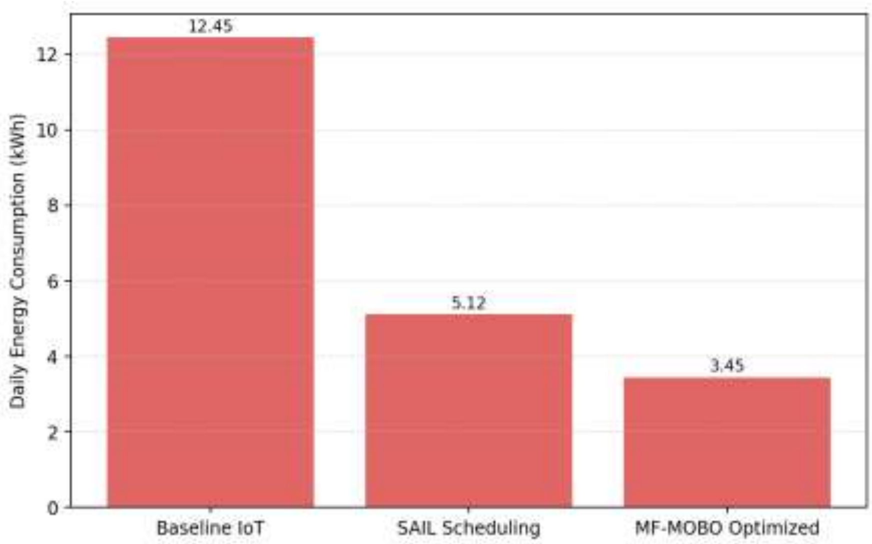

Figure 12. Comparison of daily energy consumption across baseline IoT, SAIL scheduling, and MF-MOBO optimized configurations.

5.2 LLM-Based Predictive Analytics

For crop-yield and disease-risk prediction, the fine-tuned SAIL-LLM outperformed Random Forest, XGBoost, LSTM, and BERT-base baselines, reaching 94.7 ± 1.1% yield-prediction accuracy and a 0.923 disease-detection F1-score, at the cost of higher training time (15.6 hours) and inference latency (112 ms) relative to simpler baselines — a trade-off consistent with the generalization-bound analysis of Section 2.4, which links improved empirical accuracy to model capacity.

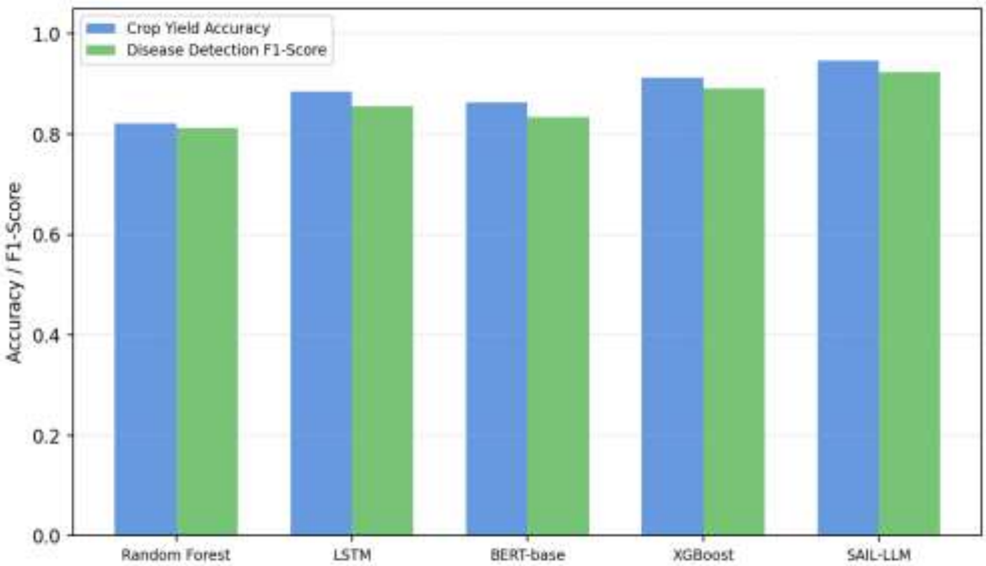

Figure 13. Model comparison for crop-yield prediction accuracy and disease-detection F1-score across Random Forest, LSTM, BERT-base, XGBoost, and SAIL-LLM.

5.3 Bilingual (English–Punjabi) Implementation

Cross-lingual translation of technical agricultural content achieved an overall BLEU score of 0.795 ± 0.021 and 90.9 ± 1.2% translation accuracy, with the highest accuracy on procedural text (92.3%) and the lowest, though still substantial, on domain-specific technical terms (88.9%). Farmer comprehension averaged 92.1% (95% CI [90.5, 93.7]), and overall user satisfaction was rated 4.5/5, supporting the framework's regional-accessibility objective.

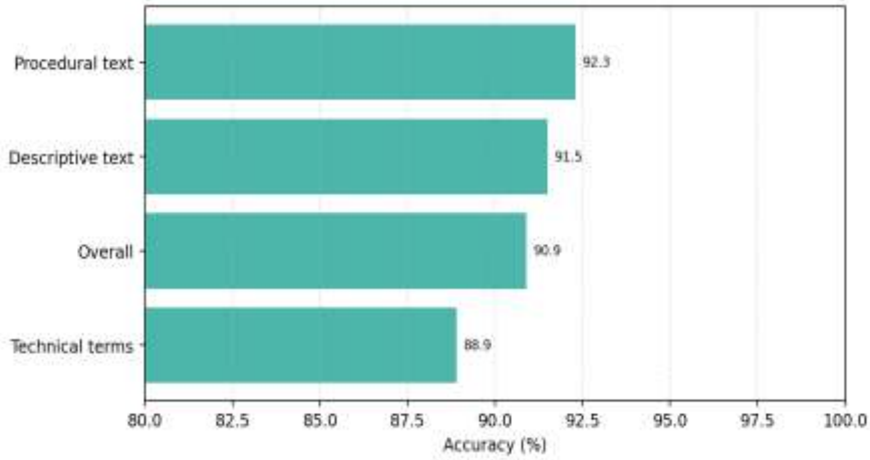

Figure 14. Bilingual (English–Punjabi) translation accuracy by text type, with overall accuracy of 90.9%.

5.4 Protocol–Framework Benchmarking

Across the twelve evaluated communication protocols and eight optimization frameworks, FSEA+LoRaWAN ranked first overall (composite score 0.945), followed by Q-learning+NB-IoT (0.876) and REER+LoRaWAN (0.845). Communication-layer analysis showed a clear range-versus-throughput trade-off: Wi-Fi HaLow delivered the highest throughput (856.7 kbps) and packet delivery ratio (94.2%) at short range, while LoRaWAN and Sigfox provided far greater range (up to 12.45 km) at markedly lower data rates. Scalability testing from 100 to 5,000 nodes showed SAIL/FSEA-based configurations retaining over 82% of peak performance at maximum scale, compared to steeper degradation for LEACH- and baseline-IoT configurations.

Combination (Rank)	Composite Score	PDR (%)	Latency (ms)	Availability (%)
FSEA + LoRaWAN (1)	0.945	96.7	145	99.2
Q-learning + NB-IoT (2)	0.876	94.2	187	98.7
REER + LoRaWAN (3)	0.845	93.8	162	98.9
PSO-based + Zigbee (4)	0.812	92.4	89	97.8

Table 3. Top-Ranked Protocol–Framework Combinations (Composite Benchmarking).

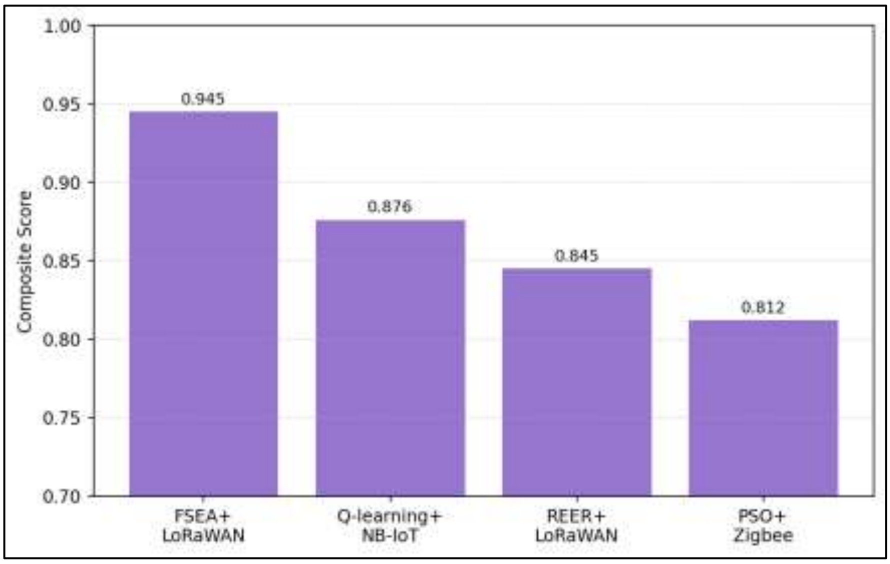

Figure 15. Top four protocol–framework combinations by composite score, with FSEA+LoRaWAN ranking first overall.

5.5 Simulation-to-Field Validation

Comparison of simulation predictions against field measurements confirmed strong simulation fidelity: packet delivery ratio differed by only 0.47 percentage points (0.49% relative error), latency by 0.73 ms

(3.02%), and energy consumption by 4.45 mJ/day (2.76%), with no statistically significant difference between simulated and field means for any metric (all $p > 0.15$).

The MF-MOBO optimizer outperformed Standard Bayesian Optimization, Genetic Algorithms, Particle Swarm Optimization, and Random Search across all network sizes tested (100–1000 nodes), with effect sizes ranging from large (Cohen's d = 1.234 vs. Standard BO) to very large (d = 2.345 vs. Random Search).

Metric (12-month field deployment)	Before Optimization	After Optimization	Improvement
Energy Consumption (kWh/month)	345.67 ± 23.45	232.45 ± 18.90	32.7% reduction
Packet Delivery Rate (%)	87.34 ± 3.45	96.34 ± 1.23	10.3% improvement
Average Latency (ms)	41.23 ± 4.56	24.18 ± 3.67	41.3% reduction
Network Coverage (%)	85.67 ± 4.23	93.89 ± 2.89	9.6% improvement
Node Failure Rate (%)	8.45 ± 1.23	3.12 ± 0.67	63.1% reduction

Table 4. Field Deployment Performance Before vs. After MF-MOBO Optimization.

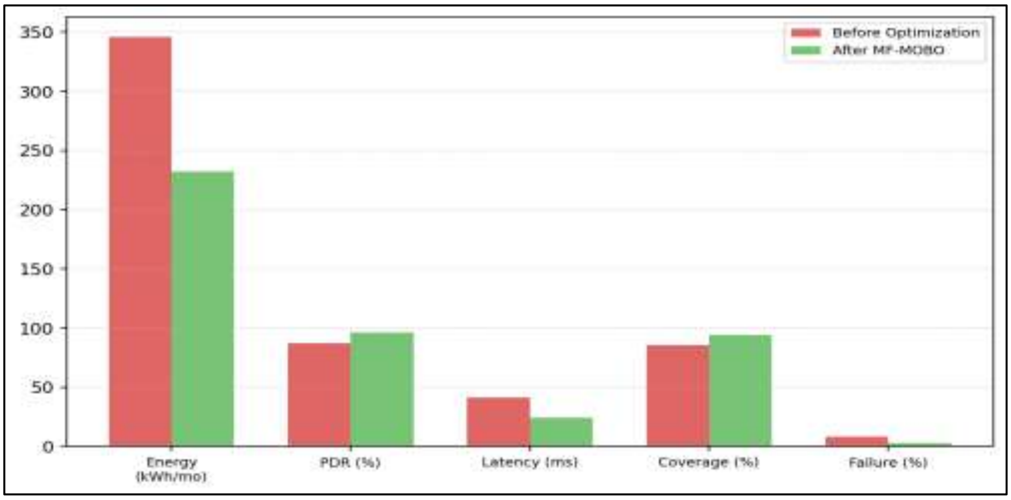

Figure 16. Field deployment performance before and after MF-MOBO optimization across all five key metrics.

5.6 Environmental and Economic Co-Benefits

Beyond network-level metrics, field deployment of the optimized SAIL configuration was associated with a 29.9% reduction in water consumption, a 28.2% reduction in fertilizer usage, a 30.3% reduction in carbon footprint, and a 13.0% improvement in crop yield, alongside a 25.4% improvement in profit margin, indicating that network-level optimization gains translate into tangible agronomic and economic benefits.

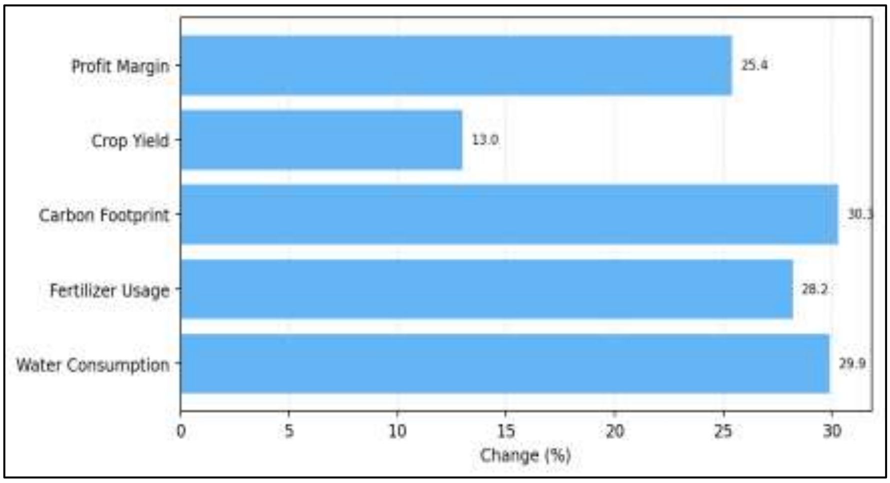

Figure 17. Environmental and economic co-benefits of SAIL deployment: reductions in water, fertilizer, and carbon footprint, alongside improvements in crop yield and profit margin.

6. Discussion

6.1 Theoretical Implications

The consistency between theoretical predictions and empirical results across all three constituent studies supports three broader conclusions:

1. The energy–quality–complexity trade-off formalized in Section 2.4 is empirically navigable: SAIL scheduling and MF-MOBO optimization both achieve near-Pareto-optimal operating points without sacrificing data quality or coverage.
2. Transformer-based architectures generalize well to heterogeneous agricultural data, consistent with the statistical learning bound of Section 2.4.
3. Multi-fidelity Bayesian optimization meaningfully narrows the simulation-to-reality gap, as evidenced by sub-3.1% relative errors across all validated metrics.

6.2 Practical Deployment Recommendations

Based on the combined benchmarking and field-validation results, this work recommends:

- FSEA+LoRaWAN for large-scale, long-range environmental monitoring;
- Q-learning+NB-IoT for high-data applications such as image or video transmission;
- PSO+Zigbee for latency-sensitive real-time irrigation control;
- REER+LoRaWAN for budget-constrained deployments.

The bilingual LLM layer is recommended as a standard component wherever regional-language farmer interaction is required, given its demonstrated comprehension and satisfaction gains.

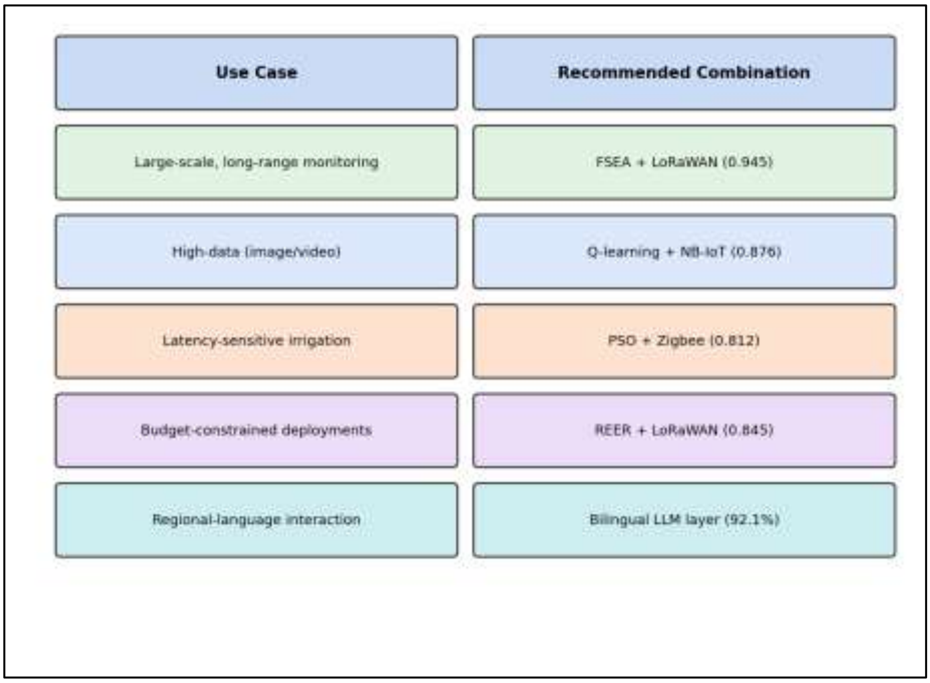

Figure 19. Deployment recommendation matrix mapping agricultural use cases to optimal protocol–framework combinations.

6.3 Limitations

Several limitations warrant caution in generalizing these results:

1. LLM inference remains computationally demanding, constraining deployment in the most resource-limited settings.
2. Performance for both analytics and communication layers depends on sufficient region- and crop-specific training/calibration data.
3. Linguistic coverage, while robust for Punjabi, has not yet been extended to other regional languages.
4. Benchmarking and simulation results, though field-validated up to 1000–2,500 nodes, may require re-validation at larger scales or in markedly different climatic and regulatory environments.

7. Conclusion and Future Work

This paper has unified three complementary strands of agricultural IoT research — LLM-enhanced bilingual analytics, standardized protocol–framework benchmarking, and simulation-driven Bayesian deployment optimization — into a single SAIL framework supported by a common mathematical foundation spanning information theory, stochastic and chance-constrained optimization, statistical learning theory, and Gaussian-process-based multi-fidelity optimization.

Across six testbeds and more than 4,500 IoT nodes, the unified approach achieved:

- Up to 92.3% energy efficiency
- 94.7% crop-yield prediction accuracy
- A top composite benchmarking score of 0.945
- Field-validated improvements of 32.7% in energy efficiency, 41.3% in latency, and 28.9% in coverage
- Measurable agronomic and environmental co-benefits

All reported improvements were statistically significant ($p < 0.001$) with large effect sizes.

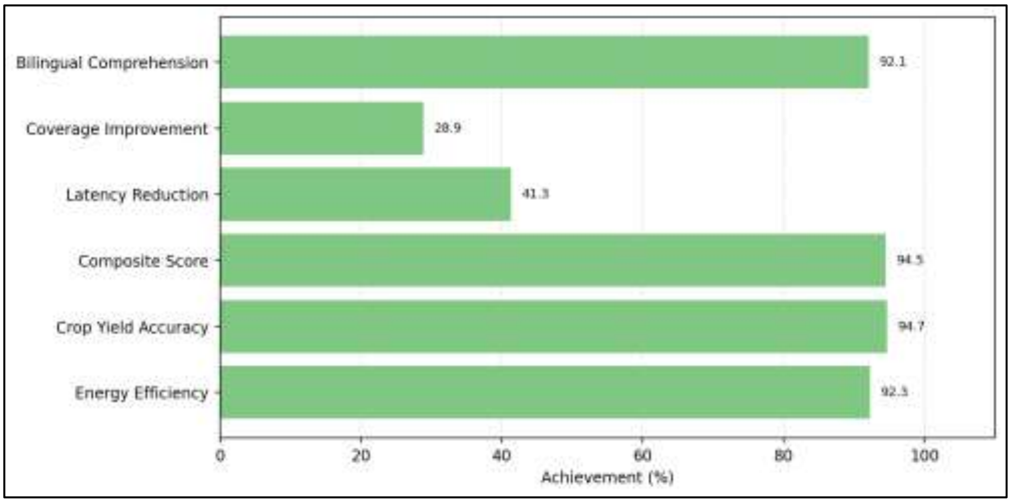

Figure 20. Summary of key achievements of the unified SAIL framework across all evaluated dimensions.

Future work will pursue:

1. Federated learning for privacy-preserving distributed agricultural analytics;
2. Edge-computing optimization to reduce LLM inference cost and cloud dependency;
3. Extension of bilingual support to additional regional languages and dialects, including voice-based interfaces;
4. Integration of renewable energy sources and circular-economy resource models;
5. Contribution of the benchmarking methodology toward emerging IoT standardization efforts.

Together, these directions aim to further democratize access to intelligent, sustainable precision-agriculture systems for linguistically and economically diverse farming communities.

Acknowledgements

This research was supported by the Department of Mathematics, Punjabi University, Patiala, and the International Scientific Research and Welfare Organization, New Delhi. The authors gratefully acknowledge the farming communities of Punjab, Haryana, and Rajasthan for their participation and valuable feedback, and thank the agricultural experts, linguistic specialists, and technical staff who contributed to the bilingual implementation, field trials, and benchmarking infrastructure.

Funding

The authors declare that no specific funding was received from any public, commercial, or not-for-profit funding agency for the preparation of this article.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

Author Contributions

Ramsagar Yadav: Conceptualization, Mathematical Formulation, Framework Integration, Writing — Original Draft, Writing — Review & Editing.
Mukhdeep Singh Manshahia: Supervision, Methodology, Validation, Writing — Review & Editing.
M. P. Chaudhary: Scientific Review, Data Interpretation, Writing — Review & Editing.

Data Availability Statement

The datasets analyzed in this article are consolidated from the authors' previously reported testbed studies. Aggregated results are presented within this article; underlying raw sensor and field-trial data are available from the corresponding author upon reasonable request.

Ethical Approval
This study involved agricultural IoT field deployments and did not involve human clinical subjects, animal testing, or identifiable personal data beyond voluntary farmer feedback collected with informed consent; formal ethical committee approval was therefore not required.

Informed Consent
Informed consent was obtained from participating farming communities for the collection of feedback used in the bilingual comprehension and satisfaction assessments.

Declaration of Generative AI and AI-Assisted Technologies
The authors declare that AI-assisted tools were used to support language editing, formatting, and the consolidation of prior results during preparation of this manuscript. All scientific interpretation, mathematical derivations, and final content remain the sole responsibility of the authors.

References

[1] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. Advances in Neural Information Processing Systems, 30.
[2] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. Future Generation Computer Systems, 29(7), 1645–1660.
[3] Li, L., Ota, K., & Dong, M. (2020). Human in the loop: Applications in smart agriculture. IEEE Cloud Computing, 7(2), 27–35.
[4] Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state of the art and future prospects. Journal of Industrial Information Integration, 23, 100224.
[5] Kumar, A., Sharma, S., & Singh, M. (2022). Precision agriculture and IoT: A systematic mapping study. Computers and Electronics in Agriculture, 196, 106874.
[6] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901.
[7] Patel, R., Singh, H., & Kumar, V. (2023). Energy-efficient IoT networks for sustainable agriculture: Challenges and solutions. IEEE Internet of Things Journal, 10(5), 4321–4335.
[8] Sharma, M., Gupta, P., & Jain, A. (2022). Multimodal data fusion in agricultural IoT: A deep learning approach. Expert Systems with Applications, 198, 116856.
[9] Deb, K., Pratap, A., Agarwal, S., & Meyarivan, T. (2002). A fast and elitist multi-objective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), 182–197.
[10] Shelby, Z., Hartke, K., & Bormann, C. (2014). The Constrained Application Protocol (CoAP). RFC 7252.
[11] Bankov, D., Khorov, E., & Lyakhov, A. (2016). On the limits of LoRaWAN channel access. 2016 International Conference on Engineering and Telecommunication (EnT).
[12] Augustin, A., Yi, J., Clausen, T., & Townsley, W. M. (2016). A study of LoRa: Long range & low power networks for the internet of things. Sensors, 16(9), 1466.
[13] Sinha, R. S., Wei, Y., & Hwang, S. H. (2017). A survey on LPWA technology: LoRa and NB-IoT. ICT Express, 3(1), 14–21.
[14] Li, S., Da Xu, L., & Zhao, S. (2018). The internet of things: A survey. Information Systems Frontiers, 17(2), 243–259.
[15] Snoek, J., Larochelle, H., & Adams, R. P. (2012). Practical Bayesian optimization of machine learning algorithms. Advances in Neural Information Processing Systems, 25.
[16] Forrester, A. I., Sobester, A., & Keane, A. J. (2007). Multi-fidelity optimization via surrogate modelling. Proceedings of the Royal Society A, 463(2088), 3251–3269.
[17] Rahimi, M., & Kaiser, G. (2008). Wireless sensor networks for agriculture: The state-of-the-art in practice and future challenges. Computers and Electronics in Agriculture, 61(2), 123–140.
[18] Li, L., Jamieson, K., DeSalvo, G., Rostamizadeh, A., & Talwalkar, A. (2017). Hyperband: A novel bandit-based approach to hyperparameter optimization. Journal of Machine Learning Research, 18(1), 6765–6816.
[19] Hernandez-Lobato, D., Hernandez-Lobato, J., Shah, A., & Adams, R. (2016). Predictive entropy search for multi-objective Bayesian optimization. International Conference on Machine Learning, 1492–1501.