Democratizing Prescriptive Agronomy: Quality-Preserving Edge AI for Sugar Beet

Ahmed Amine Tabbassi[1], Kristian Rother[1], Andrew Zakhary[1], Stefan Henkler[1]

[1] Hamm-Lippstadt University of Applied Sciences, Lippstadt, Germany

A paper from the Proceedings of the

17th International Conference on Precision Agriculture and 11th Brazilian Congress on Precision and Digital Agriculture

13-15 July 2026

Porto Alegre, Brazil

Abstract. *The global sugar beet sector faces a structural production paradox: agronomic interventions that maximize root yield frequently depress sucrose concentration and processing quality. While precision agriculture promises to navigate this trade-off, current methodologies have reached an impasse. Existing solutions are bifurcated between data-intensive machine learning (ML), which generalizes poorly across heterogeneous fields, and physiologically grounded Process-Based Models (PBMs), which are computationally rigid and, for sugar beet, mechanistically incomplete. Compounding this, the state of the art rarely accommodates the messy reality of farm data, where mismatched sensor formats and radiometric inconsistencies leave actionable insight inaccessible to non-experts. To close these gaps, this work proposes an end-to-end Decision Support System (DSS) built on the PVC-AWARE framework. The central objective is to stabilize the yield-quality trade-off by reframing it as a chance-constrained optimization problem, moving the system beyond prediction toward verifiable risk management. Strategically, we repair the DSSAT CERES-Sugarbeet module with lightweight ML augmentations, link applied nitrogen to productivity through a Mitscherlich saturation curve, estimate sucrose with a Gradient Boosting Regressor, and fit field-specific quadratic surrogates from Latin hypercube samples to enable rapid Quadratic Programming (QP). Tactically, an Advisory Mixture-of-Experts routes whatever field data is available to the most appropriate modelling expert, producing homogeneous decisions from heterogeneous inputs without costly hardware upgrades. Risk is enforced through a Plan-Validate-Cut (PVC) loop: when a proposed action, such as late-season fertilization, violates the probabilistic quality budget, the validator mints an "Assurance Cut" a persistent linear constraint that permanently excludes the unsafe decision region from future planning, converting transient runtime vetoes into durable, machine-checkable safety knowledge. A lightweight generative interface then translates the mathematical outputs into plain-language advice. On the repaired pipeline, fresh-yield prediction reached a mean absolute error of 6.4 t/ha and the sucrose model a pooled cross-validated R-squared of 0.61; constraint-aware QP delivered a mean theoretical sugar-yield gain of 2.71% over regional baselines, preserving a 2.45% gain even under a 170 kg N/ha regulatory cap, and a preliminary set of 30 paired field trials showed an observed uplift of +0.71 t/ha (+5.37%). Large-scale field deployment is ongoing. By enforcing safety through edge-resident constraints rather than exhaustive full-pipeline modelling, the architecture is designed to deliver prescriptive precision far more cost-effectively than high-compute infrastructures.*

Keywords. *Sugar beet, prescriptive agronomy, decision support system, chance-constrained optimization, crop modelling, DSSAT, surrogate optimization, edge AI, runtime safety.*

Introduction

Sugar beet (Beta vulgaris L.) is a foundational crop in the global agricultural economy, accounting for roughly 20% of the world's sugar production and forming the dominant source of sucrose in temperate climates (Draycott, 2006). The economic viability of sugar beet cultivation is dictated not by the gross tonnage of harvested roots, but by the total yield of extractable white sugar. This distinction introduces a profound management paradox: agronomic interventions that drive vegetative biomass accumulation most notably late-season nitrogen applications, and intensive irrigation often trigger a severe physiological dilution effect. The dilution reduces sucrose concentration and elevates impurities such as amino-nitrogen, potassium, and sodium, which impede the recovery of white sugar during factory processing (Hoffmann & Kenter, 2018).

The advent of precision agriculture (PA) and digital farming has promised to resolve this paradox by shifting from uniform, regional advisory guidelines toward spatially explicit, field-specific prescriptions. However, delivering prescriptive agronomy at scale requires predictive architectures capable of mathematically formulating and solving the yield-quality trade-off across a vast array of counterfactual scenarios. Current approaches fall short of this reliability because of a deep methodological bifurcation between unconstrained machine learning and rigid process-based models, neither of which simultaneously offers physiological grounding, computational tractability, and verifiable safety.

Generating a theoretically optimal management schedule is, moreover, only half of the prescriptive challenge. Executing that schedule autonomously in the real world introduces substantial epistemic uncertainty. Modern farms operate within chaotic digital ecosystems characterized by intermittent sensor connectivity, unmodeled biotic stresses, and abrupt weather shifts. If an optimized schedule dictates an irrigation event but a sudden localized storm occurs, the system must possess the autonomic intelligence to dynamically veto the action and to formulate a mathematical constraint that prevents the same error from recurring (Kephart & Chess, 2003).

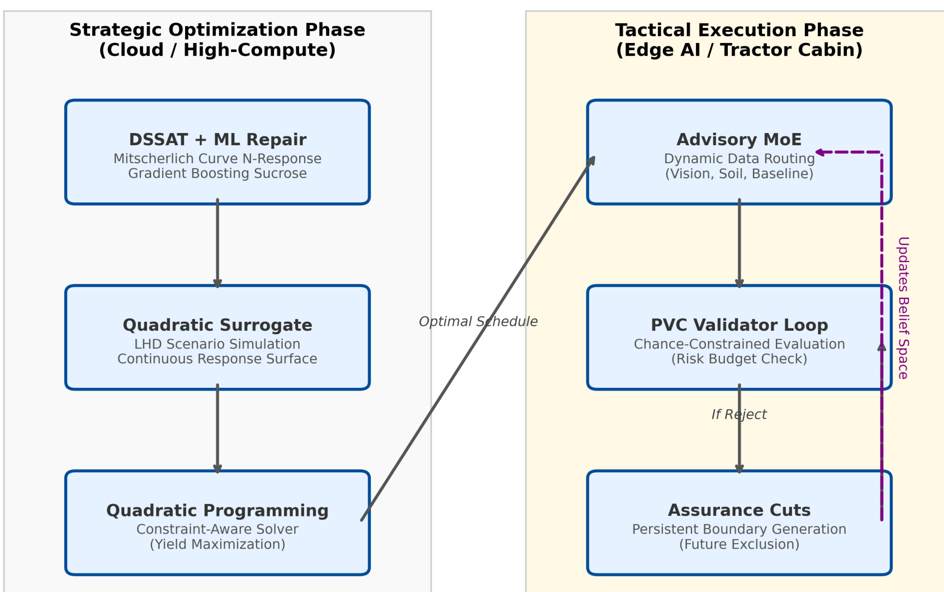

Figure 1. The unified end-to-end architecture, bridging strategic cloud-based optimization (DSSAT repair and quadratic surrogates) with tactical edge-AI execution (PVC-AWARE).

To resolve these strategic and tactical challenges jointly, this research introduces a unified, end-to-end framework (Figure 1). On the strategic side, we repair DSSAT with lightweight ML augmentations and recast the scheduling problem as a Quadratic Program (QP). On the tactical side, we deploy a Plan-Validate-Cut (PVC) belief-space planning architecture at the edge. The system treats the yield-quality balance as a strict chance-constrained risk budget, uses an Advisory Mixture-of-Experts (MoE) to route heterogeneous data, and mints persistent “Assurance Cuts” to guarantee safety during execution.Against this backdrop, the present study pursues four concrete objectives: (1) to repair the mechanistic deficiencies of the CERES-Sugarbeet nitrogen and sucrose routines so that DSSAT can serve as a credible optimization engine; (2) to convert the repaired simulator into fast, field-specific quadratic surrogates suitable for constraint-aware QP under regulatory limits; (3) to enforce processing-quality targets at execution time through a chance-constrained Plan-Validate-Cut loop that learns durable safety constraints; and (4) to package the resulting decisions in an accessible, low-compute interface that democratizes prescriptive agronomy for practitioners. Collectively, these objectives define the research goal of stabilizing the sugar beet yield-quality trade-off with verifiable, cost-effective precision.

Related Work

Democratizing prescriptive agronomy for an industrial crop such as sugar beet requires a multidisciplinary synthesis of plant physiology, predictive modelling, and autonomic systems engineering. This section reviews the foundational literature across these domains and identifies the structural deficiencies in current methodologies that motivate the proposed framework.

Agronomic Foundations: The Sugar Beet Yield-Quality Paradox

Sugar beet cultivation poses a distinctive optimization challenge relative to staple cereal crops. In global markets, the economic value of the harvest is determined not by gross biomass but by the net yield of extractable white sucrose (Draycott, 2006). This gives rise to a fundamental physiological conflict known as the yield-quality paradox.

Agronomic interventions intended to maximize taproot growth   principally nitrogen fertilization and intensive irrigation   stimulate cellular expansion and leaf area index. Excessive nitrogen, however, prolongs the vegetative phase and prevents the plant from transitioning fully into sucrose accumulation before harvest (Hoffmann & Kenter, 2018). As illustrated in Figure 2, root yield rises asymptotically with applied nitrogen while sucrose concentration falls through a pronounced dilution effect, producing a non-trivial optimization boundary.

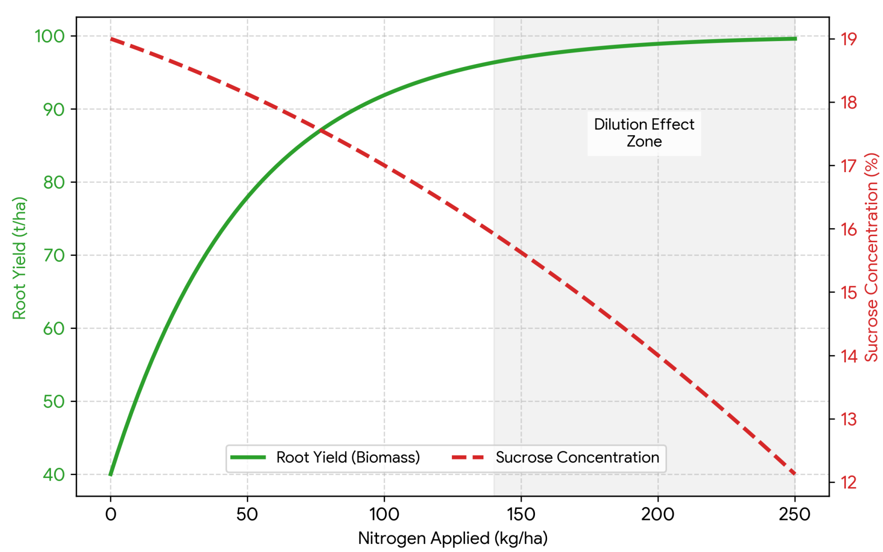

Figure 2. The sugar beet yield-quality paradox. As nitrogen application increases, vegetative root yield rises asymptotically while sucrose concentration declines through the physiological dilution effect, forming a complex optimization boundary.

Beyond dilution, excessive late-season nitrogen raises the concentration of impurities   specifically amino-nitrogen, potassium, and sodium   within the taproot. These melassigenic factors chemically bind sucrose during industrial processing, hindering crystallization at the refinery and depressing the extractability coefficient (Hoffmann & Kenter, 2018). Consequently, PA solutions cannot pursue yield maximization in isolation; they must model the multivariate trade-offs explicitly to identify the optimal intervention schedule for each field topology (Basso & Antle, 2020).

Process-Based Crop Models: Capabilities and Structural Deficiencies

To evaluate complex Genotype×Environment×Management (G×E×M) interactions without the prohibitive cost of physical trials, the agricultural community has historically relied on Process-Based Models (PBMs). Platforms such as the Decision Support System for Agrotechnology Transfer (DSSAT) and the Agricultural Production Systems sIMulator (APSIM) simulate crop growth dynamically from biophysical first principles, including daily solar radiation interception, carbon assimilation, and soil hydrology (Holzworth et al., 2014; Jones et al., 2003). The contemporary DSSAT ecosystem couples these crop modules with standardized soil, weather, and management files, enabling reproducible multi-scenario experimentation (Hoogenboom et al., 2019).

While PBMs are the gold standard for cereal crops, applying them to sugar beet exposes critical functional and architectural gaps. The CERES-Sugarbeet module (BSCER048) within DSSAT exhibits a fundamental flaw in its nitrogen routine. The module employs a demand-driven uptake mechanism: the crop computes its physiological nitrogen requirement from potential growth and then extracts that quantity from the simulated soil pools largely regardless of actual mineral availability (Jones et al., 2003). As a result, the model cannot natively register nitrogen stress; simulations parameterized at 0 kg N/ha and 250 kg N/ha frequently return indistinguishable dry-matter outputs. Because rigorous calibration of such modules depends on the model expressing the very responses being calibrated (Seidel et al., 2018; Wallach et al., 2019), this insensitivity undermines any attempt to tune the crop to observed nitrogen-response data.

Furthermore, CERES-Sugarbeet lacks a dedicated subroutine for sucrose accumulation; its terminal output is purely dry biomass. Because sugar yield is the ultimate economic objective, a PBM that outputs only dry matter is inherently incapable of supporting end-to-end optimization These structural deficiencies mandate that, before DSSAT can be used as an optimization engine for sugar beet, it must undergo significant mechanistic repair and augmentation.

Machine Learning and Surrogate-Assisted Optimization

To bridge the gaps left by rigid PBMs, prescriptive agronomy has increasingly turned to machine learning. ML architectures from Random Forests to deep Long Short-Term Memory (LSTM) networks excel at identifying non-linear patterns across massive, heterogeneous datasets (Liakos et al., 2018). Deploying pure ML models in agronomy nonetheless introduces real risks. Lacking intrinsic biological grounding, a model trained to optimize yield will frequently recommend agronomically disastrous interventions, such as massive over-fertilization, because it has no innate understanding of sucrose biochemistry or long-term soil-health degradation (Basso & Antle, 2020).

To combine the biological grounding of PBMs with the computational efficiency of ML, researchers have developed surrogate-assisted optimization frameworks (Zan Yang et al., 2023). Here, expensive crop models are replaced by lightweight, mathematically tractable proxy models, or surrogates. Building reliable surrogates requires intelligent sampling of the continuous management space. The Latin Hypercube Design (LHD) is widely used to guarantee near-orthogonal stratification across multiple dimensions for example nitrogen rate, planting date, and irrigation offering superior variance reduction relative to standard Monte Carlo sampling (McKay et al., 1979). Once sampled, the responses can be fitted with algorithms such as Gradient Boosting Regressors (Friedman, 2001; Chen & Guestrin, 2016) or second-order polynomial response surfaces. The latter is particularly advantageous because it translates a complex biological simulation into a standard Quadratic Programming (QP) problem, enabling robust convex optimization to locate global optima in milliseconds while strictly respecting linear regulatory constraints such as the EU Nitrates Directive (Boyd & Vandenberghe, 2004; Nocedal & Wright, 2006).

Autonomic Computing and Runtime Safety Assurance

While surrogate optimization solves the strategic planning problem, executing those schedules in the physical world introduces a new vector of complexity: runtime safety under epistemic uncertainty. Modern precision agriculture relies heavily on edge AI to interpret sensor streams in real time, operating within the paradigm of autonomic computing. Self-managing architectures are classically modelled via the MAPE-K (Monitor-Analyze-Plan-Execute over a shared Knowledge base) loop (Kephart & Chess, 2003).

Executing an optimized schedule safely calls for Chance-Constrained Optimization (CCO) (Charnes & Cooper, 1959) and Stochastic Model Predictive Control (SMPC), which bound the probability of system failure over a receding horizon (Mesbah, 2016). Running SMPC dynamically on resource-constrained edge hardware for instance within a tractor cabin is, however, often computationally infeasible. Many deployed systems therefore fall back on lightweight reactive safety monitors.

These reactive monitors suffer from a systemic flaw we term persistent ignorance (Figure 3). When a reactive monitor detects an unsafe condition for example, when a proposed irrigation delay threatens to breach a critical drought threshold it vetoes the action but discards the mathematical evidence that triggered the rejection. Because no persistent constraint is formulated, the planner is free to propose the same hazardous action on the next cycle, producing dangerous oscillation and rapidly depleting the available safety budget.

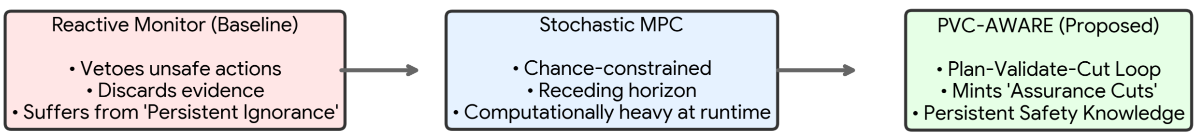

Figure 3. Architectural evolution in runtime safety assurance: the transition from baseline reactive monitors to belief-space frameworks that retain persistent safety knowledge.

Recent advances in conformal prediction and belief-space planning calibrate uncertainty bounds with finite-sample coverage guarantees (Angelopoulos & Bates, 2023) and formalize decision-making under partial observability (Kaelbling et al., 1998). Translating these uncertainty quantifications into durable, machine-checkable constraints that dynamically restrict the operational search space remains, however, a critical open challenge. Systems must move beyond merely identifying risk to converting runtime failures into persistent safety artifacts.

Taken together, the literature reveals a distinct methodological fragmentation. Strategic planning relies either on rigid crop models that are functionally incomplete for sugar beet or on unconstrained ML algorithms that violate physiological reality, while tactical execution relies on reactive edge-monitors that fail to learn from their operational boundaries. Addressing the yield-quality paradox therefore requires a unified mathematical architecture. The proposed framework answers this need by systematically repairing DSSAT to generate constraint-aware quadratic surrogates for strategic planning, while concurrently deploying the PVC-AWARE (Plan-Validate-Cut) architecture to mint persistent Assurance Cuts that guarantee tactical safety at the edge.

Materials and Methods

The proposed framework is organized as a two-tier pipeline (Figure 4). The strategic tier, executed in the cloud, repairs the DSSAT crop model, distils it into field-specific quadratic surrogates, and solves a constraint-aware Quadratic Program to produce an optimal management schedule. The tactical tier, executed at the edge, routes heterogeneous field data through an Advisory Mixture-of-Experts and enforces processing-quality targets at runtime through a chance-constrained Plan-Validate-Cut loop. The remainder of this section formalizes each component in turn, beginning with the data foundation and the mechanistic repair of DSSAT.

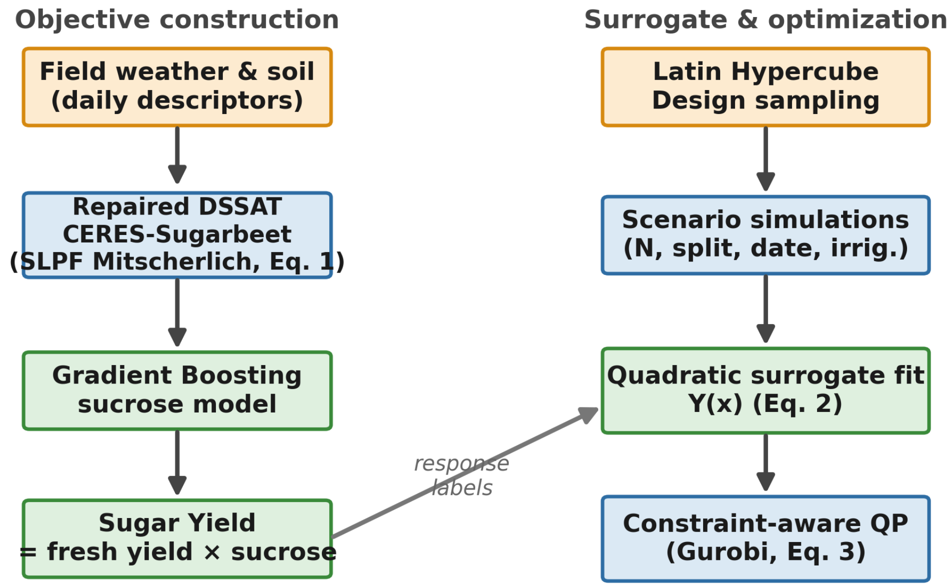

Figure 4. The strategic surrogate-generation and optimization pipeline. Field weather and soil descriptors drive the repaired DSSAT simulator and the Gradient Boosting sucrose model to construct the Sugar Yield objective, which labels Latin hypercube scenarios used to fit the quadratic surrogate and solve the constraint-aware QP.

Study Area and Data

The surrogate library underpinning the strategic tier was developed across a large set of commercial sugar beet production fields in North Rhine-Westphalia, Germany, a major European beet-growing region with temperate-maritime climate and predominantly loamy soils. For each field, georeferenced management records were paired with measured root yield and quality attributes, the latter cross-checked against sugar-factory delivery records to provide factory-verified labels for sucrose concentration and extractability. Daily weather series (solar radiation, maximum and minimum temperature, precipitation) and profile soil descriptors served as the environmental drivers supplied to DSSAT. Management variables total nitrogen, the early-season nitrogen split fraction, planting-date offset, and seasonal irrigation defined the decision space over which the framework optimizes. All records were harmonized to the DSSAT file conventions so that simulations and field observations could be compared on a common basis (Hoogenboom et al., 2019).

Repairing the DSSAT Nitrogen Response

The first methodological contribution is a targeted repair of the demand-driven nitrogen routine that renders CERES-Sugarbeet insensitive to fertilization. Rather than re-engineering the module's internal source code, we intervene through the Soil Productivity Factor (SLPF), a scalar parameter on [0, 1] that uniformly scales the crop's effective resource supply. By making the SLPF an explicit function of applied nitrogen, we induce the diminishing-returns growth response that the native routine fails to express. We model this relationship with a Mitscherlich saturation curve:

$$\mathrm{SLPF}(N) = s_{max} - (s_{max} - s_{min})\cdot\exp(-c\cdot N) \qquad (1)$$

where s_min (0.55) is baseline productivity under severe nitrogen deficiency, s_max (1.0) is the asymptotic maximum under non-limiting nitrogen, and c is the saturation-rate constant governing how quickly the response approaches its ceiling. Figure 5 plots the calibrated curve: productivity rises steeply at low nitrogen and flattens well before the EU Nitrates Directive limit of 170 kg N/ha, reproducing the physiologically realistic, diminishing-returns behaviour expected of the crop. This single, interpretable transformation restores the model's ability to distinguish between deficient and sufficient nitrogen regimes, which is a precondition for any downstream optimization (Seidel et al., 2018; Wallach et al., 2019).

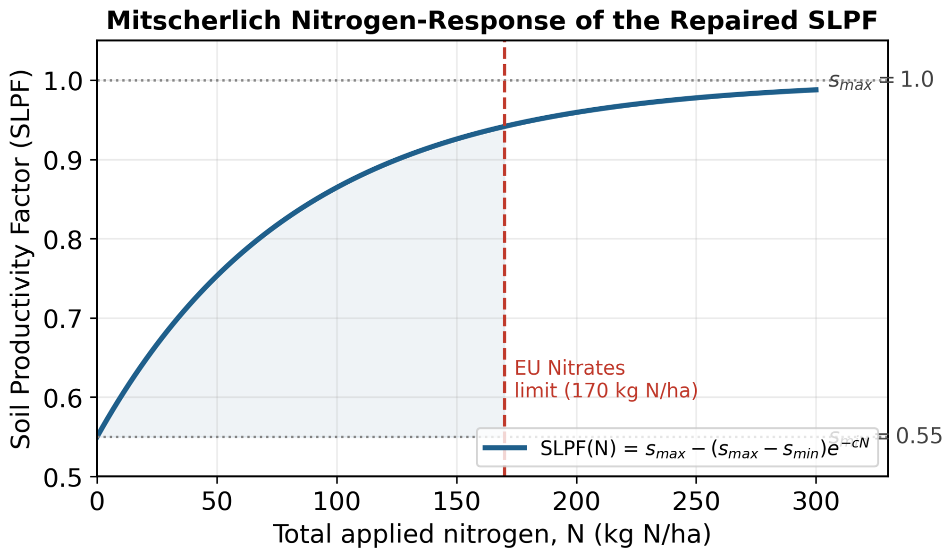

Figure 5. Calibrated Mitscherlich nitrogen-response of the repaired Soil Productivity Factor (Equation 1). Productivity saturates before the regulatory nitrogen ceiling, restoring sensitivity to fertilization that the native demand-driven routine lacks.

Sugar Content Estimation via Machine Learning

The second contribution addresses the absence of a sucrose subroutine. Because the repaired simulator still outputs only fresh and dry biomass, we attach a data-driven sucrose model on top of it. We train a Gradient Boosting Regressor (Friedman, 2001; Chen & Guestrin, 2016) to predict final sucrose concentration (%) from 17 meteorological and biometric features including cumulative thermal time, intercepted radiation, late-season temperature and precipitation aggregates, and simulated biomass-state variables exported by DSSAT. Gradient boosting was selected for its strong performance on tabular, moderately sized agronomic datasets and its robustness to correlated features and mild non-linearity (Hastie et al., 2009). Models were implemented in Python with established open-source libraries (Pedregosa et al., 2011), and generalization was assessed with grouped cross-validation that holds out whole fields, so that performance reflects transfer to unseen sites rather than memorization of within-field structure.

The sucrose prediction is then combined multiplicatively with the simulated fresh yield to form the economic objective that the optimizer maximizes:

Sugar Yield = fresh yield × predicted sucrose concentration. Treating sugar yield, rather than biomass, as the optimization target is what allows the framework to confront the yield-quality paradox directly instead of optimizing a proxy that the refinery does not reward.

Latin Hypercube Design and Quadratic Surrogates

Simulating the entire continuous management space for thousands of fields is computationally intractable. Instead, for each field we draw a Latin Hypercube Design (LHD) (McKay et al., 1979) over the n-dimensional decision vector x comprising total nitrogen, early-nitrogen split fraction, planting-date offset, and seasonal irrigation. The LHD stratifies each dimension into equiprobable intervals and permutes the samples so that the projection onto any single axis is near-uniform, yielding markedly lower estimator variance than naive Monte Carlo sampling for the same number of simulator calls. Each sampled scenario is evaluated by the repaired DSSAT-plus-sucrose pipeline to produce a labelled (x, Sugar Yield) pair.

The labelled scenarios are then used to fit a field-specific, second-order polynomial response surface (a quadratic surrogate) by ordinary least squares. The surrogate prediction function is:

$$Y(x) = \beta_0 + \Sigma_i \beta_i x_i + \Sigma\Sigma_{i\le j} \beta_{ij} x_i x_j \quad (2)$$

where the coefficients β capture the intercept, the main (linear) effects, and the pairwise interaction and curvature (squared) terms. Because the surrogate is a smooth, low-order polynomial, it can be differentiated analytically and evaluated in microseconds, and its coefficients map directly onto the matrices of a Quadratic Program.

Constraint-Aware Quadratic Programming

Substituting the fitted surrogate into the management problem yields a standard Quadratic Program. We maximize sugar yield subject to regulatory and agronomic constraints:

$$\max_x \ \tfrac{1}{2} x^T Q x + c^T x \quad \text{s.t.} \quad Ax \le b, \quad lb \le x \le ub \quad (3)$$

where Q holds the interaction and squared-term coefficients of Equation 2, c the linear terms, the bounds lb and ub bracket agronomically admissible settings of each decision variable, and $Ax \le b$ encodes linear policy constraints most importantly the EU Nitrates Directive cap (Council Directive 91/676/EEC). Because the constraints enter the program directly rather than as soft penalties, the optimizer cannot recommend agronomically or legally infeasible schedules. We solve the resulting programs with a commercial interior-point and active-set solver (Gurobi Optimization, 2024); where the fitted Q is indefinite, the problem is handled as a general (non-convex) QP with the solver's global routines, following standard practice in numerical optimization (Nocedal & Wright, 2006). Solving each field-specific program in milliseconds is what makes optimization across thousands of fields practical.

Advisory Mixture-of-Experts for Heterogeneous Data

The optimal schedules produced by the strategic tier presuppose clean, complete inputs. In the field, however, the available data are heterogeneous and frequently degraded: some fields offer high-resolution multispectral imagery, others only sparse in-situ soil probes, and others little more than regional agronomic priors. To produce consistent decisions from such uneven inputs without mandating expensive uniform sensing, the tactical tier employs an Advisory Mixture-of-Experts (MoE). A lightweight gating network inspects the quality and availability of each incoming data stream and routes the decision to the most reliable expert: a vision expert that interprets canopy reflectance and vegetation indices, a soil-sensor expert that assimilates in-situ probe measurements, and a baseline agronomic expert that falls back on calibrated regional priors when richer signals are absent (Figure 6).

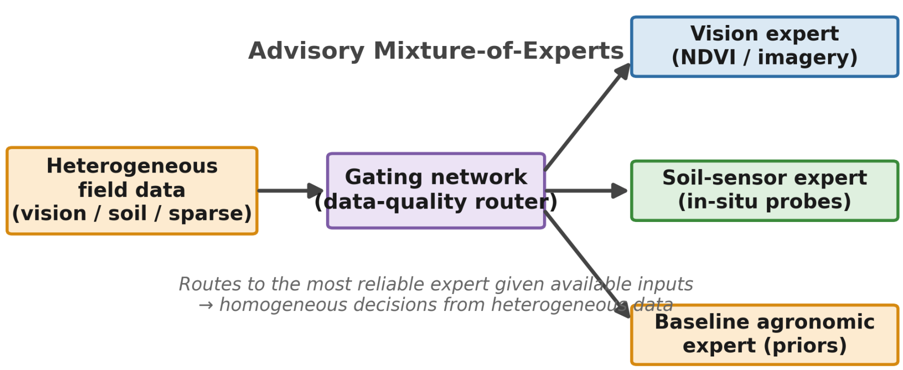

Figure 6. The Advisory Mixture-of-Experts. A gating network routes heterogeneous, possibly degraded field data to the most reliable expert, yielding homogeneous decisions without requiring uniform high-end sensing hardware.

Because routing is performed by a small gating model rather than by running every expert everywhere, the MoE delivers homogeneous, comparable recommendations across a fleet of fields at a fraction of the computational and hardware cost of a monolithic full-pipeline model. This design directly supports the democratization goal: a grower with only modest instrumentation still receives a coherent, quality-aware prescription.

Chance-Constrained Optimization Formulation

The optimal schedules generated by the surrogate models (Equation 3) act as strategic targets. To execute these targets safely under partial observability, we deploy the PVC-AWARE architecture at the edge. Rather than co-optimizing yield and quality through arbitrary scalar weights which conflates incommensurable objectives and hides the true risk being taken PVC-AWARE maximizes agronomic utility subject to a strict probabilistic risk budget α, following the chance-constrained programming paradigm (Charnes & Cooper, 1959). We define the chance constraint as:

$$P(\ Q(x, \omega) \geq Q_{min}\ ) \geq 1 - \alpha \tag{4}$$

where $Q(x, \omega)$ is the stochastic sucrose concentration resulting from action x under environmental uncertainty ω, Q_min is the critical quality threshold (e.g., 16.0%), and α is the maximum acceptable probability of failure. The constraint guarantees that the delivered crop meets its processing-quality target with confidence at least 1 − α. Crucially, because the threshold is expressed as a probability rather than a penalty weight, the grower sets a single, interpretable risk tolerance that the system then honours exactly.

The Plan-Validate-Cut Loop and Assurance Cuts

The operational core is the Plan-Validate-Cut (PVC) loop, which proceeds in three repeating phases. In the Plan phase, the system proposes an action x drawn from the optimal schedule. In the Validate phase, it assesses the chance constraint of Equation 4 by Monte Carlo rollouts over the current belief state, propagating environmental uncertainty ω through the surrogate to estimate the failure probability; this belief-space evaluation is the practical surrogate for exact reasoning under partial observability (Kaelbling et al., 1998). If the estimated failure probability respects the budget, the action executes. If it violates the budget, the system enters the Cut phase.

In the Cut phase, the validator mints an Assurance Cut (AC): a linear half-space constraint derived from the gradient of the failure probability at the rejected point x_0,

$$C_k(x) = \nabla P(x_0)^T (x - x_0) \leq 0 \tag{5}$$

where the gradient ∇P(x_0) is estimated locally (e.g., by finite differences over the Monte Carlo rollouts) and defines the orientation of the excluded half-space. Geometrically (Figure 7), the cut is a hyperplane through the rejected point whose normal points toward increasing failure probability; appending it to the constraint matrix A in Equation 3 permanently removes the unsafe region from the feasible decision space and steers subsequent planning toward a structurally safer optimum. This converts a transient runtime veto into durable, machine-checkable safety knowledge: because the cut persists across planning cycles, the planner can never re-propose a previously rejected hazardous action, eliminating the oscillation characteristic of memory-less reactive monitors.

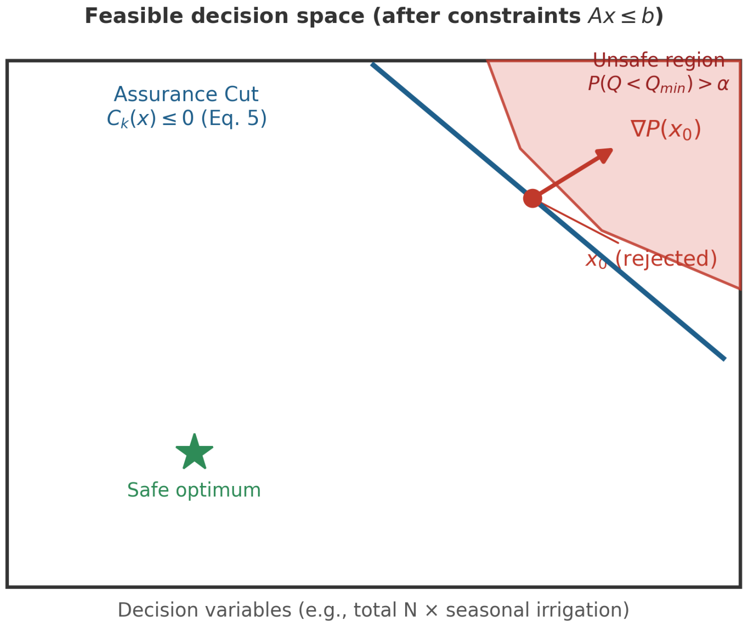

Figure 7. Geometry of an Assurance Cut. At a rejected point x_0, the gradient of the failure probability defines a half-space (Equation 5) that is appended to the feasible region, permanently excluding the unsafe region and steering the optimizer toward a safe optimum.

Generative Advisory Interface

Finally, to ensure practical adoption, the mathematical outputs of the optimizer and the PVC loop are passed to a lightweight generative interface that renders them as concise, plain-language advice for example, recommended nitrogen splits, irrigation timing, and an explicit statement of the residual quality risk. Translating QP solutions and probabilistic guarantees into human-readable guidance lowers the expertise barrier to acting on the prescription and closes the loop between formal optimization and on-farm decision-making.

Preliminary Results and Discussion

Surrogate Fidelity and Strategic Optimization

The repaired DSSAT pipeline demonstrated robust predictive accuracy. Baseline fresh-yield predictions achieved a Mean Absolute Error (MAE) of 6.4 t/ha, with a median error of just −1.2%, indicating low systematic bias. The Gradient Boosting sucrose model achieved a pooled cross-validation R-squared of 0.61, a reasonable fit given the inherent variability of field-measured sucrose concentration and the conservative grouped, field-held-out validation scheme.

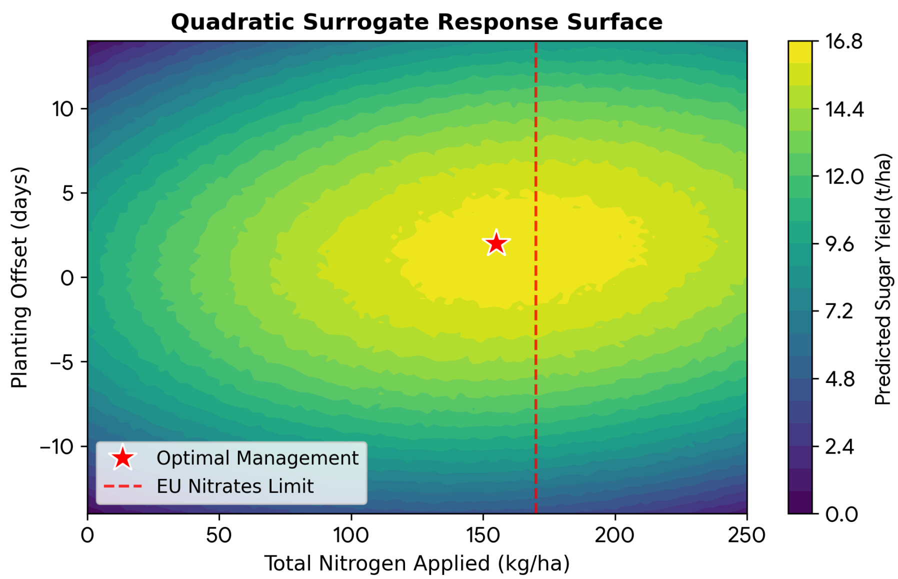

Figure 8. Quadratic surrogate response surface for a representative field, generated via Equation 2. The global optimum is identified while strictly respecting the EU Nitrates Directive linear constraint.

When the field-specific quadratic surrogates were passed to QP solvers, continuous joint optimization yielded a mean theoretical gain of 2.71% (+0.85 t/ha) over regional advisory baselines. Because the optimization is constraint-aware (Equation 3), the system navigated the EU Nitrates Directive (Council Directive 91/676/EEC), preserving a 2.45% yield gain even when fertilization was legally capped at 170 kg N/ha. In a prospective validation set of 30 paired field trials, the optimized schedules delivered a physically observed mean sugar-yield uplift of +0.71 t/ha (+5.37%). The gap between the theoretical and observed gains is consistent with unmodeled field heterogeneity and reinforces the need for the runtime safety layer described next.

Tactical Safety via PVC-AWARE

The tactical safety of the PVC-AWARE architecture was evaluated against a standard, memory-less reactive monitor in simulated quality-preservation scenarios in which a high-risk intervention repeatedly threatened the processing-quality threshold.

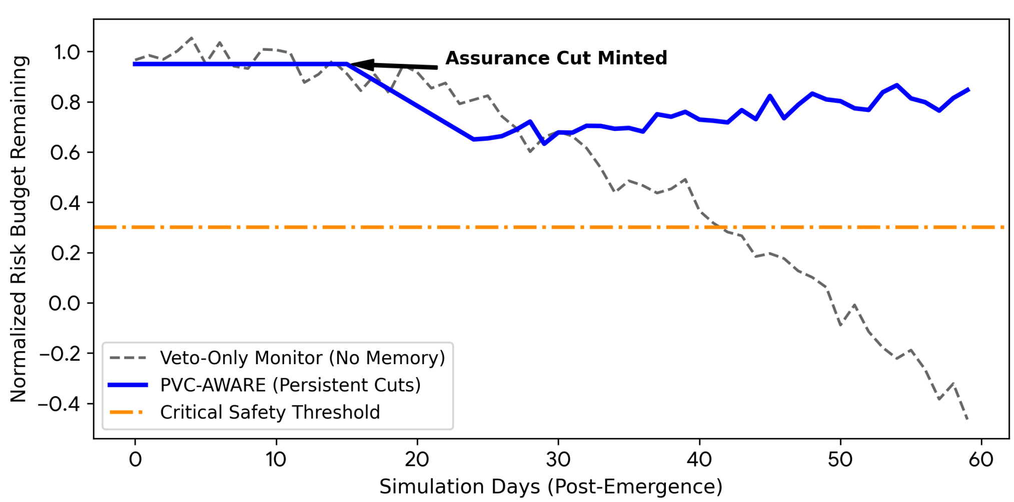

Figure 9. Simulation of safety-budget depletion. The PVC-AWARE system mints a persistent Assurance Cut (Equation 5) upon identifying high risk, stabilizing the safety budget over the receding horizon.

As illustrated in Figure 9, standard veto-only monitors suffered from persistent ignorance, producing an oscillating, steadily depleted risk budget. By contrast, the moment the PVC-AWARE validator identified a high-risk intervention it minted an Assurance Cut (Equation 5). This mathematical boundary immediately and permanently restricted the planner’s action space, forcing the system to adopt a structurally safer schedule and stabilizing the safety budget. The result confirms that converting runtime vetoes into durable constraints, rather than discarding them, is the mechanism that breaks the oscillation cycle and preserves quality guarantees over the horizon. Taken together with the strategic gains, the two tiers are complementary: the surrogate-QP stage supplies an aggressive, regulation-compliant target, while the PVC layer protects the processing quality that ultimately determines the value of the harvest.

Conclusion

The democratization of prescriptive agronomy requires systems that are computationally scalable, biologically grounded, and mathematically verifiable. This research presented a unified framework that addresses the sugar beet yield-quality paradox end to end. Strategically, repairing the demand-driven nitrogen routine of DSSAT through an interpretable Mitscherlich response (Equation 1) and attaching a gradient-boosted sucrose model restored the simulator’s sensitivity to management and re-centred it on the economically meaningful objective of sugar yield. Distilling the repaired simulator into field-specific quadratic surrogates (Equation 2) then enabled rapid, constraint-aware QP optimization (Equation 3) across many fields, yielding a mean theoretical sugar-yield gain of 2.71% and a +5.37% observed uplift on a preliminary set of paired trials while remaining compliant with the EU Nitrates Directive.

Tactically, deploying these schedules through the PVC-AWARE edge-AI architecture ensured their safe execution. The Advisory Mixture-of-Experts produced homogeneous decisions from heterogeneous, often degraded field data without mandating uniform high-end sensing, and the Plan-Validate-Cut loop enforced a single interpretable risk budget by combining chance-constrained evaluation (Equation 4) with persistent Assurance Cuts (Equation 5). In simulation, this mechanism converted transient runtime vetoes into durable, machine-checkable safety knowledge and stabilized the quality risk budget that memory-less monitors allowed to oscillate and deplete. The two tiers are complementary: aggressive, regulation-aware optimization upstream, and verifiable quality protection downstream.

Several limitations frame the contribution and motivate further work. The reported field evidence is preliminary, drawn from a modest set of paired trials, and the sucrose model’s moderate cross-validated fit leaves room for improvement; the tactical results, while encouraging, are so far established in simulation. Future work will therefore extend the present validation to large-scale, multi-season field deployment across the North Rhine-Westphalia production region, stress-test the Advisory Mixture-of-Experts routing under genuinely degraded and intermittent sensor conditions, and refine the sucrose estimator with additional biometric and remote-sensing features. We further intend to study the long-horizon behaviour of accumulated Assurance Cuts including strategies for retiring stale cuts as beliefs are updated and to evaluate the generative advisory interface with practitioners to confirm that the framework’s mathematical guarantees translate into trusted, actionable guidance. In bridging theoretical optimization and practical, risk-aware farming, the framework offers a cost-effective path toward prescriptive precision that does not depend on high-compute infrastructure, and its modular structure a repairable process model, tractable surrogates, and a verifiable runtime safety layer is readily transferable to other quality-sensitive industrial crops.

Acknowledgments

This work was supported by the Ministry of Economic Affairs, Industry, Climate Action and Energy of the State of North Rhine-Westphalia and co-funded by the European Union (Grant number EFRE-20800498).

References

Angelopoulos, A. N., & Bates, S. (2023). Conformal prediction: A gentle introduction. *Foundations and Trends in Machine Learning, 16*(4), 494–591.

Basso, B., & Antle, J. (2020). Digital agriculture to design sustainable agricultural systems. *Nature Sustainability, 3*(4), 254–256.

Boyd, S., & Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Charnes, A., & Cooper, W. W. (1959). Chance-constrained programming. *Management Science, 6*(1), 73–79.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM.

Council Directive 91/676/EEC. (1991). Concerning the protection of waters against pollution caused by nitrates from agricultural sources. *Official Journal of the European Communities, L 375*, 1–8.

Draycott, A. P. (2006). *Sugar beet*. Blackwell Publishing.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics, 29*(5), 1189–1232.

Gurobi Optimization, LLC. (2024). *Gurobi Optimizer reference manual*. Gurobi Optimization, LLC.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

Hoffmann, C. M., & Kenter, C. (2018). Yield potential of sugar beet – have we hit the ceiling? *Frontiers in Plant Science, 9*, 289.

Holzworth, D. P., Huth, N. I., de Voil, P. G., Zurcher, E. J., Herrmann, N. I., McLean, G., … Keating, B. A. (2014). APSIM – evolution towards a new generation of agricultural systems simulation. *Environmental Modelling & Software, 62*, 327–350.

Hoogenboom, G., Porter, C. H., Shelia, V., Boote, K. J., Singh, U., Pavan, W., … Jones, J. W. (2019). The DSSAT crop modeling ecosystem. In K. J. Boote (Ed.), *Advances in crop modelling for a sustainable agriculture* (pp. 173–216). Burleigh Dodds Science Publishing.

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., … Ritchie, J. T. (2003). The DSSAT cropping system model. *European Journal of Agronomy, 18*(3–4), 235–265.

Kaelbling, L. P., Littman, M. L., & Cassandra, A. R. (1998). Planning and acting in partially observable stochastic domains. *Artificial Intelligence, 101*(1–2), 99–134.

Kephart, J. O., & Chess, D. M. (2003). The vision of autonomic computing. *Computer, 36*(1), 41–50.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. (2018). Machine learning in agriculture: A review. *Sensors, 18*(8), 2674.

McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics, 21*(2), 239–245.

Mesbah, A. (2016). Stochastic model predictive control: An overview and perspectives for future research. *IEEE Control Systems Magazine, 36*(6), 30–44.

Nocedal, J., & Wright, S. J. (2006). *Numerical optimization* (2nd ed.). Springer.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

Seidel, S. J., Palosuo, T., Thorburn, P., & Wallach, D. (2018). Towards improved calibration of crop models – where are we now and where should we go? *European Journal of Agronomy, 92*, 130–149.

Wallach, D., Makowski, D., Jones, J. W., & Brun, F. (2019). *Working with dynamic crop models: Methods, tools and examples for agriculture and environment* (3rd ed.). Academic Press.

Zan Yang, Haobo Qiu, Liang Gao, Danyang Xu, Yuanhao Liu, (2023). A general framework of surrogate-assisted evolutionary algorithms for solving computationally expensive constrained optimization problems*, Information Sciences 619* ISSN 0020-0255