Machine Learning-Based Optimization of Site-Specific Fertilizer Recommendation

Oumnia Ennaji [1] **Abdelghani Belgaid** [1] **Achraf El Allali** [1]

Abstract

This paper presents a machine learning (ML) approach for generating site-specific fertilizer recommendations that maximize crop yield and nutrient efficiency while minimizing environmental impact. Using a rich agronomic dataset from the Al Moutmir program in Morocco, various ML models are trained (linear, tree-based, ensemble, and neural networks) to predict crop yield responses to nitrogen (N), phosphorus (P), and potassium (K) inputs under diverse soil and climate conditions. The best predictive model (an XGBoost regressor) achieved a Mean Absolute Percentage Error (MAPE) of 8.9%, substantially outperforming baseline approaches. The predictive model is then integrated with optimization algorithms (including Simulated Annealing and Particle Swarm Optimization) to identify the optimal N, P, K levels for each site. Simulated application of these recommendations indicates an average yield improvement of about 544 kg/ha over current practices, along with more efficient fertilizer use and low environmental impact. The importance of key features is analyzed in the recommendations, and an open analysis of the approach limitations is provided. All results are validated with statistical tests for significance. The proposed framework demonstrates how advanced ML and optimization techniques can enhance precision agriculture by tailoring fertilizer strategies to local needs.

1. Introduction

Global food security and environmental sustainability are driving a need for more efficient fertilizer use in agriculture. Conventional "one-size-fits-all" fertilization often leads to nutrient imbalances, water pollution, and greenhouse gas emissions. Precision agriculture (PA) has emerged as a promising solution to these issues by accounting for spatial variability in soil and crop conditions (Fanzo & Davis, 2019; Paudel et al., 2021). Site-specific fertilizer recommendations aim to apply optimal N, P, and K rates per field, improving yield and reducing waste. Recent studies highlight the potential of data-driven approaches (Barbosa et al., 2020; Ennaji et al., 2024). However, many existing approaches focus on yield prediction alone and do not directly recommend optimal fertilizer rates while assessing nutrient efficiency and environmental impact (Ennaji et al., 2023).

This study develops an end-to-end framework that predicts crop yield response and optimizes fertilizer recommendations for each site. This framework builds upon prior ML applications in agriculture and multi-objective optimization techniques for resource management. The key contributions are: (1) an integrated dataset of 7,180 on-farm trials covering varied regions and seasons, (2) a comparative evaluation of several ML algorithms for predicting yield from soil, and fertilizer inputs, (3) a formulation of site-specific fertilizer recommendation as a constrained optimization problem. By tailoring fertilizer strategies to local conditions, this approach aims to enhance farmers' productivity and sustainability.

2. Problem Formulation

The site-specific fertilizer task is formalized as a two-step problem: (a) yield prediction and (b) fertilizer optimization. First, an ML model $f(\mathbf{x}, N, P, K)$ is trained to predict crop yield given site features $\mathbf{x}$ (soil properties, weather, crop type, etc.) and fertilizer rates (N, P, K). This model approximates the complex agronomic response of yield to inputs. Second, for a new site with known features $\mathbf{x}$, the fertilizer combination (N, P, K) that maximizes predicted yield while considering nutrient use efficiency and environmental impact is identified (Basso et al., 2018). The objective function is defined as:

$$J(N, P, K \mid \mathbf{x}) = Y_{\text{pred}}(N, P, K \mid \mathbf{x}) - \lambda \cdot E(N, P, K), \tag{1}$$

[1]College of Computing, Mohammed VI Polytechnic University, Lot 660, Ben Guerir, 43150, Morocco. Correspondence to: Oumnia Ennaji <Oumnia.ENNAJI@um6p.ma>, Abdelghani Belgaid <Abdelghani.BELGAID@um6p.ma>, Achraf El Allali <Achraf.ELALLALI@um6p.ma>.

Proceedings of the 42nd International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025. Copyright 2025 by the author(s).

where Y_{pred} is the yield predicted by the ML model and $E(N, P, K)$ is an environmental penalty (e.g. for excessive N that could leach). The trade-off parameter λ adjusts the weight on environmental cost. In practice, the optimization maximizes Y_{pred} under agronomic constraints (e.g. upper/lower bounds for nutrients) instead of explicitly using λ (see below). The optimization problem is:

$$\max_{\substack{N \in [N_{\min}, N_{\max}] \\ P \in [P_{\min}, P_{\max}] \\ K \in [K_{\min}, K_{\max}]}} Y_{\text{pred}}(N, P, K) \tag{2}$$

with optional constraints on total fertilizer budget or environmental thresholds. This yields a recommended site-specific NPK rate. It is worth noting that the formulation inherently balances competing objectives: because the ML model is trained on real-world data, maximizing its output tends to favor high yields and avoid obviously wasteful over-fertilization (diminishing returns). To further safeguard against overuse, a custom penalty is included in the optimization for any nutrient application beyond agronomic needs (an approach inspired by adaptive penalty methods).

3. Methodology

3.1. Data and Features

Data from the Al Moutmir program was used (Ennaji et al., 2025), an open agriculture program in Morocco, which contains results from thousands of on-farm cereal trials. The dataset comprises 7,180 field observations collected over three growing seasons (2018–2019: 914 plots; 2019–2020: 1,736 plots; 2020–2021: 4,530 plots). Trials span 8 regions and 34 provinces, covering both rainfed and irrigated systems and mainly three cereal crops (soft wheat, durum wheat, barley). Each observation includes the site's soil properties (measured via soil tests), crop type and management information, fertilizer application rates, and the resulting yield.

Soil features: The key soil variables are pH (acidity/alkalinity), organic matter (%), total and active calcium carbonate ($CaCO_3$), texture (clay and sand %), Olsen phosphorus (P_2O_5 in ppm), exchangeable potassium (K_2O in ppm), and electrical conductivity (salinity). Significant variability was observed in these properties across sites, underscoring the need for site-specific management. Additional context variables include geographic coordinates (latitude, longitude) to capture spatial yield patterns (proxy for climate and other regional factors), and year/season indicators (Qin et al., 2018).

Data processing: Prior to modeling, rigorous data processing was applied. Missing values (some soil metrics) were imputed using median values stratified by region. Obvious outliers (e.g. unrealistically high yields or nutrient values) were winsorized. Numeric features were standardized to zero-mean, unit-variance for the ML models. Categorical variables (crop type, region) were one-hot encoded. A few redundant or future-known attributes (e.g. post-harvest measurements) were also dropped to avoid data leakage. These steps ensured a clean and consistent dataset for learning.

Train/test split: Models were evaluated under two scenarios: a random split (80% train, 20% test, stratified by crop yield quartiles) and a temporal split (earlier two seasons as train, the last season as test). The random split represents the case of independent and identically distributed samples, while the temporal split evaluates the model's ability to generalize to a new year (simulating deployment for future seasons). The random split was stratified to maintain yield distribution, with 1,421 training and 356 test instances per fold. The temporal split resulted in 1,356 train and 1,224 test points post data augmentation using SMOTE, with the test set containing a different distribution (notably, 70% of all data comes from 2020–2021, which was exclusively in the test set in this split).

3.2. Model Training

A range of ML algorithms was evaluated for yield prediction, including: (i) Linear models – Linear Regression, Ridge, Lasso; (ii) Tree-based ensemble models – Random Forests, Gradient Boosting Machines (GBM), XGBoost, LightGBM, CatBoost, AdaBoost; (iii) Neural networks – a deep fully-connected network (DNN), a feature-attention neural network, and a residual network; and (iv) Stacked ensembles – meta-learners combining multiple models. These models capture increasing levels of non-linearity and interactions. For fairness, all models were trained on the same input features and target (yield).

Hyperparameters: Default or literature-based hyperparameters were adopted, with limited tuning. Tree-based models were run with up to 2000 boosting iterations/trees and a learning rate of 0.05, except AdaBoost, which used 10 estimators (with a Random Forest base) due to its higher per-estimator cost. The forest models (Random Forest) used 100 trees. Neural networks were configured with 4 hidden layers (sizes 1024→512→256→128) with BatchNorm and Dropout regularization. Neural networks were trained for up to 2000 epochs with early stopping (patience 50) to prevent overfitting. All models were optimized using the Adam optimizer (for neural networks) or default boosting optimizers. Training used a Kaggle notebook with a Tesla P100 GPU.

Ensembling: To further improve robustness, a simple ensemble was implemented that averages the predictions of top-performing models. Specifically, the three models with the lowest validation error were selected and combined by weighted averaging (weights proportional to the reciprocal of their MAPE). This ensemble aimed to reduce variance

and model-specific bias.

Evaluation metrics: MAPE was primarily used and the coefficient of determination (R^2) to assess predictive accuracy. MAPE directly measures percentage error relative to actual yield, which is intuitive for agronomic relevance (e.g. 10% error on a 5 ton/ha yield = 0.5 ton/ha off). R^2 indicates the variance explained by the model. RMSE (kg/ha) was also tracked for interpretability in yield units. For each experiment, these metrics were computed on the test set.

3.3. Optimization for Fertilizer Recommendation

Once a predictive model $f(\mathbf{x}, N, P, K)$ was trained, it was used as a surrogate to optimize fertilizer rates. Because the response surface can be complex and non-convex, metaheuristic algorithms were employed instead of gradient-based optimization. A suite of optimization methods was evaluated: Simulated Annealing (SA), Particle Swarm Optimization (PSO), Genetic Algorithm (GA), Adaptive Differential Evolution, Bayesian Optimization (and a hybrid Bayesian-Evolutionary variant), Q-Learning-based Optimization, and others, totaling 10 methods. These algorithms explore the (N, P, K) space (continuous, and bounded within a reasonable range defined by the maximum and minimum values observed in the training dataset) to find the best predicted yield. The objective was the predicted yield itself, and a mild penalty was incorporated for total fertilizer amount to encourage parsimonious solutions (particularly in GA and SA via a weighted fitness function).

Each optimization run starts from the actual farmer's fertilizer rate and searches for improvements. The number of evaluations was limited to f to 1000 per method (to simulate a scenario of computational constraint in field use). The output is an optimal NPK recommendation for that site and the corresponding predicted yield. Each algorithm was applied to test set sites and recorded: the objective value reached (predicted yield minus any penalty), the predicted yield improvement over the farmer's practice, the reduction in unused nutrients, and the environmental impact.

Evaluation of recommendations: Algorithms were compared in terms of the yield gains, nutrient efficiency, and environmental impact. Nutrient Use Efficiency (NUE) is defined as the ratio of yield increase to additional nutrient applied (kg grain per kg nutrient) and Environmental Impact Metric as a proxy for potential N loss (e.g. surplus N beyond crop uptake). These were computed from the model's outputs. Additionally, Explained Variability was calculated for each recommendation: the percentage of the site's yield gap (difference between potential and current yield) that was closed by the recommended inputs. This metric indicates how well an algorithm's recommendation aligns with the site's actual needs. Higher values (close to 100%) mean the method found nearly optimal inputs, whereas lower values mean it fell short. Nutrient response surfaces were visualized for select crops to qualitatively assess the interactions between N, P, and K. For brevity, most detailed results are deferred to the Appendix.

The Supplementary Material Figure 1 shows the research workflow that involves data collection (soil and environmental features), data processing and feature engineering (cleaning, normalization, encoding), model training and evaluation (trying various ML models with performance metrics like R^2, RMSE, MAPE, MAE, and conducting a feature analysis), and finally an optimization step to derive site-specific NPK recommendations using various optimization algorithms. The process iterates with validation and ensures the recommendations are evaluated for yield improvement and environmental impact.

4. Results

4.1. Predictive Model Performance

On the random split (mixed seasons), ML models captured the yield variations with high accuracy. The best single model, XGBoost, attained $R^2 = 0.90$ and MAPE = 8.87% on the test set. This means the model's predictions were on average within 9% of actual yields, a substantial improvement over baseline linear regression (MAPE 24%, $R^2 \approx 0.70$) and a Null model (predicting mean yield, $R^2 = 0$ by definition). The stacking ensemble further reduced error to 8.5% MAPE (not a large gain, indicating XGBoost already captured most patterns). Most tree-based models and neural networks clustered in performance between 9–12% MAPE, whereas simpler models (linear, ridge, lasso) were worse (over 20% error). In contrast, on the temporal split, all models saw a performance drop. Stacking was top with MAPE $\approx$ 29% and a negative R^2 on the 2020–2021 season. Other models had MAPE ranging from 30% (random forest) to over 40% (for less flexible models), sometimes even yielding negative R^2 (indicating they performed worse than predicting the historical average). This highlights the challenge of capturing year-specific factors like weather or crop changes. Overall, the random split results demonstrate that given sufficient representative data, ML can accurately learn the yield response function. The temporal experiment underlines a limitation, indicating that additional temporal features or transfer learning may be needed to improve predictivity for future seasons.

To understand the models' decisions, SHAP values (Shapley Additive Explanations), feature importance, and permutation tests were examined. Soil properties and location emerged as critical factors. In particular, soil organic matter and pH were consistently top predictors of yield response across models, along with latitude (which correlates with climate zone). High organic matter and neutral pH tended to

increase yield potential and modify the optimal N, P requirements. Latitude/longitude features likely proxy regional rainfall patterns not explicitly in the data. Crop type also had a strong influence (different cereals respond differently to nutrients). Interestingly, initial soil phosphorus levels affected the yield response to P fertilizer additions fields with low Olsen-P saw greater yield increases from P applications, as expected agronomically. These findings are consistent with expert knowledge and lend credibility to the model. They show that it learns reasonable cause-effect relationships and not false patterns.

4.2. Optimization Outcomes

Using the best yield-prediction models, optimized NPK recommendations were generated for each test plot and evaluated. They substantially improved predicted yields compared to farmers' original practices. On average, the optimized rates gave a +543.8 kg/ha increase in yield (from a baseline of 3.5 t/ha, i.e. a 15% improvement) while reducing unused nutrients. Many sites showed that similar yields could be achieved with less fertilizer or higher yields with the same fertilizer, by reallocating N, P, K based on site conditions. For example, in some low-phosphorus soils, the optimizer suggested increasing P and slightly reducing N, improving nutrient balance and uptake.

Across the three major crops (barley, durum wheat, soft wheat), Simulated Annealing (SA) and Particle Swarm Optimization (PSO) were the top-performing optimization algorithms, consistently finding the highest-yield solutions. SA achieved the best objective values in most cases, likely due to effective search space exploration, while PSO was a close second with a faster convergence in some runs. For barley, SA improved yield by 768 kg/ha with a slight environmental penalty, whereas PSO achieved an even higher +770 kg/ha yield but with a bit more N usage (higher environmental cost). For durum wheat, SA was again best (yield +508 kg/ha), followed by a Bayesian-Evolutionary hybrid method and PSO (within 2%). Soft wheat showed a smaller relative gain (being the highest-yielding crop baseline), with SA and PSO still leading. GA and Adaptive Differential Evolution also found good solutions but were slightly outperformed by SA/PSO in the trials. Simpler local search methods (gradient-based, etc.) often got stuck in suboptimal points due to the complex response surface.

Qualitatively, the optimized nutrient rates reflected sensible agronomic trends. Supplementary Material Figure 2 visualizes the response surfaces for each crop: for barley, yield is most sensitive to N and K balance (insufficient K severely limits yield even if N is abundant) while P has a lesser effect. For durum wheat, P is more critical – the optimizer often increased P on P-deficient sites to improve root development, and managed K carefully to avoid luxury consumption. Soft wheat, being generally high-yielding, required a high N supply but the model showed diminishing returns beyond a point, so the optimizer sometimes reduced N from farmer levels if the model indicated no further gain, focusing on K instead. These patterns align with known crop-specific nutrient needs.

Finally, **Explained Variability (%)** is used to assess how well optimized nutrient rates align with actual crop needs. For each nutrient $i \in \{\text{N, P, K}\}$, it is computed as:

$$\text{Explained Variability (i\%)} = 100 \times \left(1 - \frac{|x_i - y_i|}{\max(x_i, 1)}\right),$$

comparing optimized (y_i) vs. observed (x_i) mean values. The average across nutrients gives a composite score. While SA and PSO had the highest yield gains, Gradient Descent (76%), Policy-Based (64.6%), and Q-Learning Optimization (61.2%) scored highest on Explained Variability, highlighting a trade-off between maximizing yield and aligning with real-world agronomic inputs. Moreover, the results suggest that no single optimizer consistently excels across all crops, emphasizing the importance of adopting crop-specific optimization strategies to fully capture the heterogeneity in nutrient responses.

5. Limitations

This study faces limitations in data, temporal generalization, optimization assumptions, and economic modeling. The dataset is imbalanced (70% from 2020–2021, 77% soft wheat) and relies on coarse external data (e.g., NASA POWER, SoilGrids), which may limit generalizability. Temporal performance drops significantly, likely due to unobserved covariate shifts (e.g., weather, farming practices); future work should consider temporal embeddings, domain adaptation, or transfer learning. The optimization assumes surrogate model correctness, errors in prediction may lead to suboptimal or harmful recommendations, though mitigated with bounds and penalties. Finally, the objective maximizes yield but does not account for input costs or crop prices. Incorporating profit-based objectives and farmer risk preferences is essential for real-world adoption. Synthetic data generation (Belgaid & Ennaji, 2025) and robust optimization are also promising future directions.

6. Conclusion

This study presents a machine learning and optimization framework for site-specific fertilizer recommendation, achieving strong yield gains and efficiency on Moroccan cereal plots. By addressing data limitations and incorporating economic, temporal, and user-centered elements, this approach can support scalable, responsible precision agriculture.

References

Barbosa, A., Trevisan, R., Hovakimyan, N., and Martin, N. F. Modeling yield response to crop management using convolutional neural networks. *Computers and Electronics in Agriculture*, 170(February):105197, 3 2020. ISSN 01681699. doi: 10.1016/j.compag.2019.105197.

Basso, B., Dumont, B., Maestrini, B., Shcherbak, I., Robertson, G. P., Porter, J. R., Smith, P., Paustian, K., Grace, P., Asseng, S., et al. Soil organic carbon and nitrogen feedbacks on crop yields under climate change. *Agricultural & Environmental Letters*, 3(1):180026, 2018.

Belgaid, A. and Ennaji, O. Sagda: Open-source synthetic agriculture data for africa, 2025. URL https://arxiv.org/abs/2506.13123.

Ennaji, O., Vergutz, L., and El Allali, A. Machine learning in nutrient management: A review. *Artificial Intelligence in Agriculture*, 2023.

Ennaji, O., Baha, S., Vergutz, L., and El Allali, A. Gradient boosting for yield prediction of elite maize hybrid zhengdan 958. *PLOS ONE*, 19(12):1–16, 12 2024.

Ennaji, O., Hamma, A., Vergütz, L., and El Allali, A. The assessment of soil variables relative importance for cereal yield prediction under rainfed cropping system in morocco. *Smart Agricultural Technology*, 100950, 2025. doi: 10.1016/j.atech.2025.100950.

Fanzo, J. and Davis, C. Can diets be healthy, sustainable, and equitable? *Current obesity reports*, 8:495–503, 2019.

Paudel, D., Boogaard, H., de Wit, A., Janssen, S., Osinga, S., Pylianidis, C., and Athanasiadis, I. N. Machine learning for large-scale crop yield forecasting. *Agricultural Systems*, 187(December 2020):103016, 2021. ISSN 0308521X. doi: 10.1016/j.agsy.2020.103016.

Qin, Z., Myers, D. B., Ransom, C. J., Kitchen, N. R., Liang, S., Camberato, J. J., Carter, P. R., Ferguson, R. B., Fernandez, F. G., Franzen, D. W., Laboski, C. A., Malone, B. D., Nafziger, E. D., Sawyer, J. E., and Shanahan, J. F. Application of Machine Learning Methodologies for Predicting Corn Economic Optimal Nitrogen Rate. *Agronomy Journal*, 110(6):2596–2607, 11 2018. ISSN 0002-1962. doi: 10.2134/agronj2018.03.0222.

Appendix: Machine Learning-Based Optimization of Site-Specific Fertilizer Recommendation

June 29, 2025

This appendix provides supplementary material for the paper "Machine Learning-Based Optimization of Site-Specific Fertilizer Recommendation". It includes:

- Overview of the data analysis and optimization workflow (Figure A1)
- Summary statistics and distributions of key agronomic variables (Table A1)
- Grain yield distribution across crops and growing seasons (Figures A2–A3)
- Comparative performance of optimization methods in terms of yield improvement, nutrient efficiency, and environmental impact (Figure A4)
- Three-dimensional NPK response surfaces for selected crops (Figure A5)

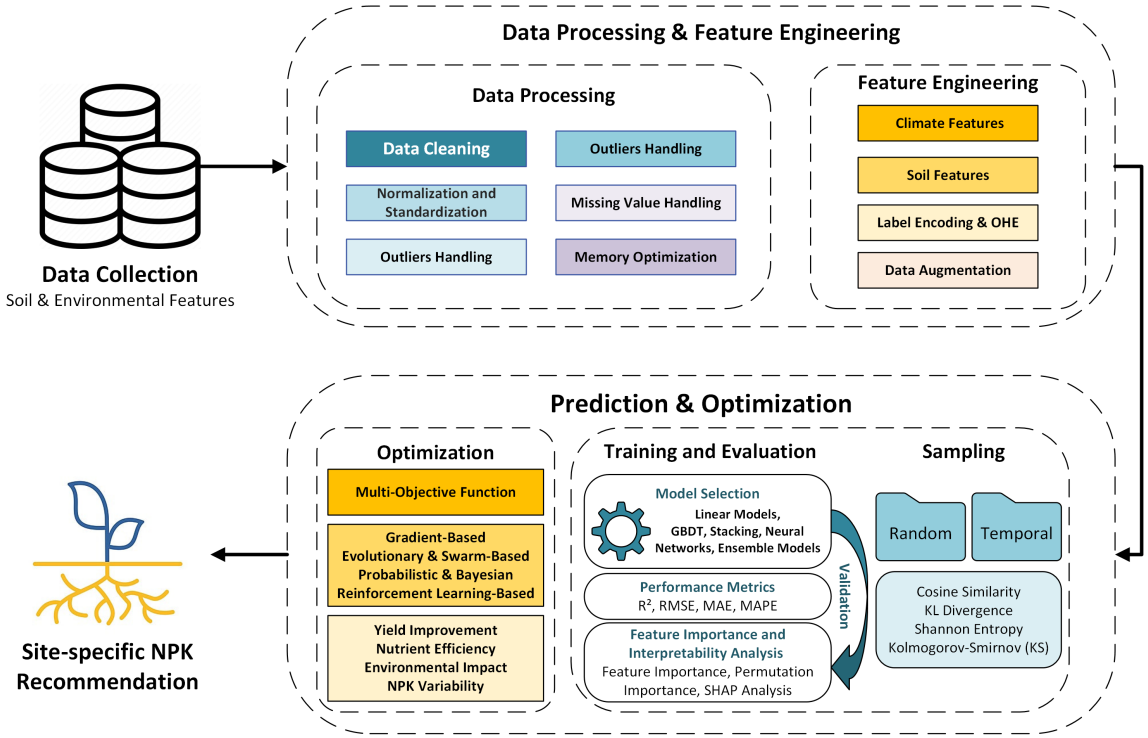

Figure 1: Research workflow.

Table 1: Descriptive Statistics and Histograms of Key Variables

Variable	Mean	SD	P0	P25	P50	P75	P100	Hist
soil_ph	7.97	0.37	7.40	7.68	8.00	8.30	8.50	
organic_matter_percent	2.40	0.62	1.60	1.83	2.30	2.90	3.40	
phosphorus_ppm	46.48	26.51	18.00	23.00	38.00	66.00	93.00	
potassium_ppm	304.48	102.11	169.50	210.00	293.60	388.00	463.60	
electrical_conductivity	0.09	0.48	0.00	0.00	0.00	0.00	8.00	
npk_nitrogen	28.76	5.81	20.00	24.00	30.00	34.00	36.75	
npk_phosphorus_p2o5	46.01	13.92	26.00	34.17	45.00	60.00	65.00	
npk_potassium_k2o	39.22	15.54	20.00	24.50	38.25	53.00	66.00	
npk_magnesium_mgo	1.14	0.69	0.36	0.53	0.97	1.64	2.37	
npk_calcium_cao	3.15	1.72	1.09	1.52	2.83	4.17	6.02	
electrical_conductivity_encoded	0.52	2.12	0.00	0.00	0.00	0.00	16.00	
grain_yield_kg	3524.55	1530.33	1365.00	2180.00	3500.00	5000.00	5700.00	

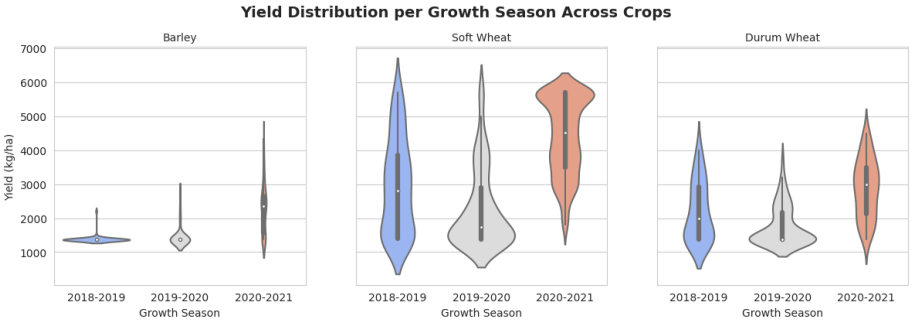

Figure 2: Distribution of grain yield across crops and growing seasons.

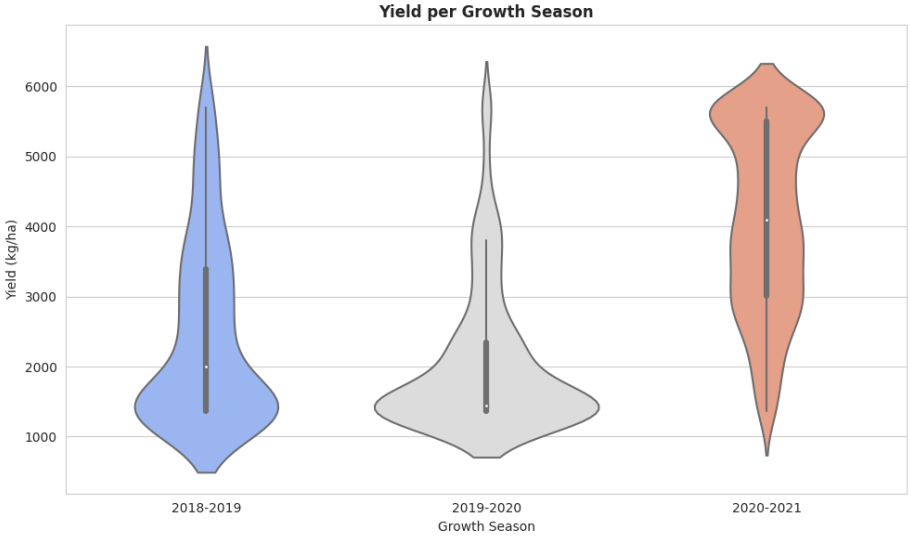

Figure 3: Average grain yield per growing season.

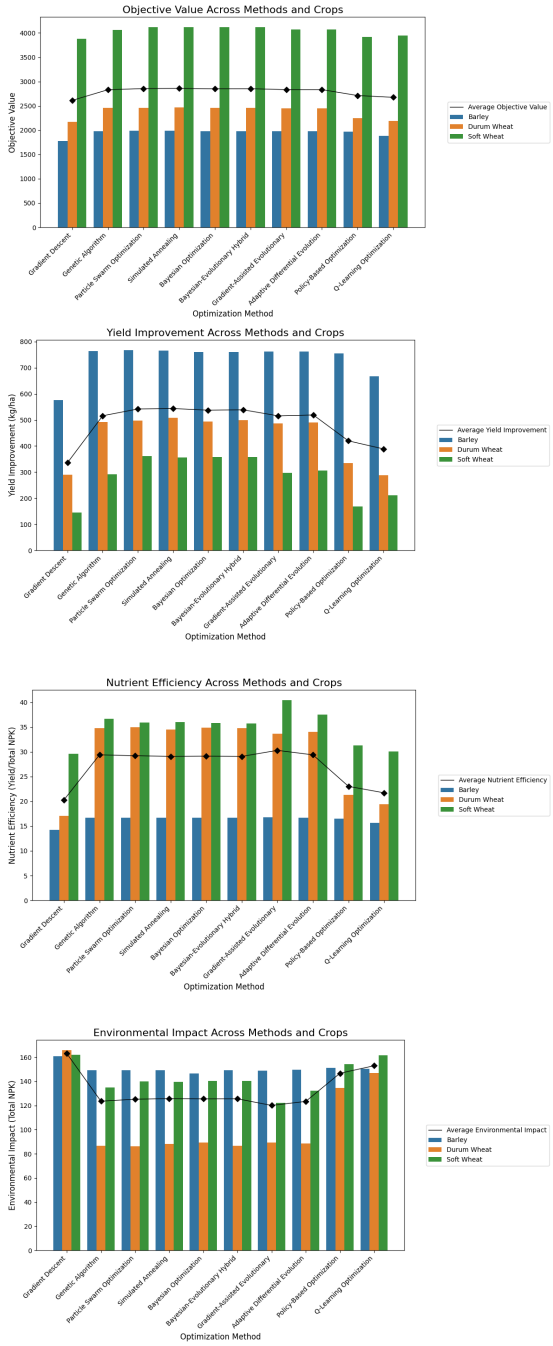

Figure 4: Comparative performance of optimization methods across crops: objective function value, yield improvement, nutrient use efficiency, and environmental impact.

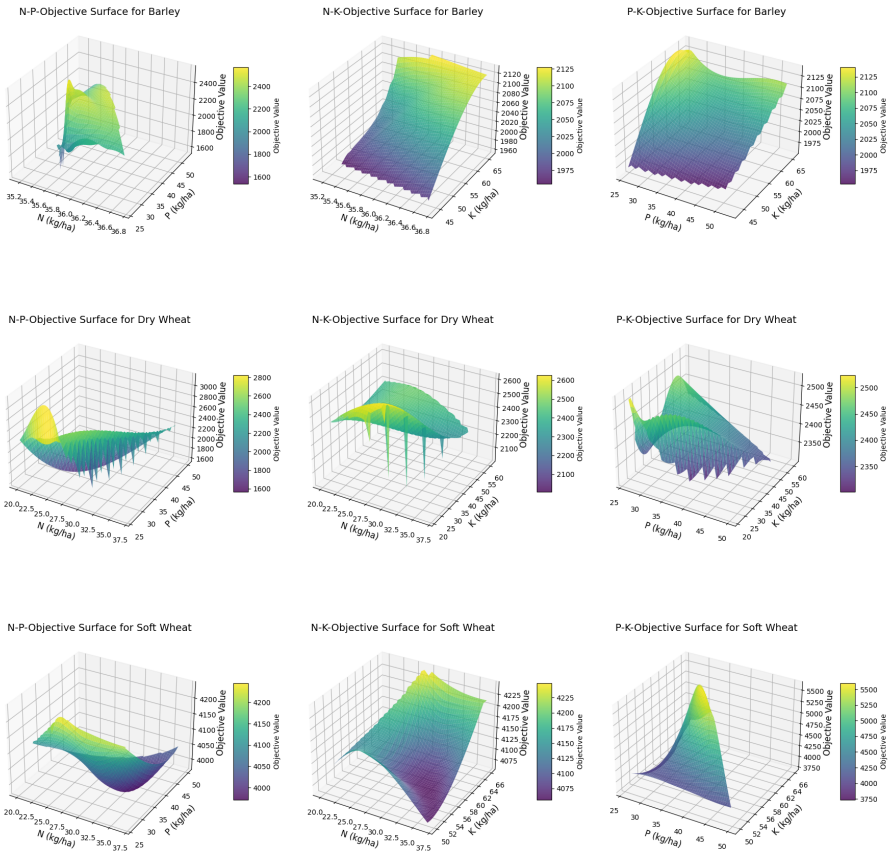

Figure 5: Three-dimensional NPK response surfaces for Barley, Durum Wheat, and Soft Wheat. The surfaces represent the objective function values combining yield prediction, nutrient use, and environmental cost.