A Multi-Agent Policy Gradient Framework for Weather-Adaptive Crop Protection Using Deep Ensemble Learning

Dr. H. B. Jethva[1],Dr. Vivekanandam[2], Dr. Eugenio Vocaturo[3]

[1] Gujarat Technological University; [2]Lincoln University, Malaysia; [3] Eugenio Vocaturo, University of Calabria, Italy hbjethva@gmail.com, vivekanandam@lincoln.edu.my, eugenio.vocaturo@cnr.it

Abstract: Agricultural productivity is also very responsive to dynamic weathers and pest attacks and this requires intelligent and dynamic prediction systems that can be utilized to effectively protect crops. The research hypothesizes a Multi-Agent Policy Gradient Framework of Weather-Adaptive Crop Protection with Deep Ensemble Learning which combines the predictive power of hybrid deep models with the decision-making power of reinforcement learning. The framework uses a multi-agent design, with individual agents focused on checking and forecasting individual environmental indicators, including temperature, rainfall, and humidity, and pest activity, with the help of a complex of convolutional and recurrent neural networks. The ensemble method improves strength and precision by combining various outputs of the models, thus reducing the uncertainty of individual models. Reinforcement learning module is a policy gradient that optimizes the adaptive decision strategies enabling agents to adaptively change their prediction policies as a result of changing environmental patterns. The system learns to determine the best responses to the bad weather conditions and the potential occurrence of a pest through constant feedback means, to assist in taking timely action on crop protection. The framework proposed is validated through multi-source data which incorporates meteorological, satellite and agricultural field data. It has been observed through experimentation that predictive accuracy is greater, adaptation to climate variations is more rapid and predictive outcomes such as crop yields are more resilient when using experimental models rather than the traditional static ones. The present research augments the methodology of scaling, smart, real-time weather-adaptive crop protection, which is in line with the goal of sustainable agriculture and precision farming in unpredictable climatic scenarios.

Keywords: Deep Ensemble Learning; Policy Gradient; Multi-Agent Reinforcement Learning; Adaptive Weather Prediction; Crop Protection; Precision Agriculture

1. Introduction

1.1 Background and Motivation

The agriculture sector is the support of most economies because it offers food security and a source of employment in the world. Nonetheless, the unpredictable weather conditions and climate change are also posing a threat to agricultural productivity. The abnormal rainfalls, high and low temperatures, and infestations have become significant challenges, which have lowered the quality and quantity of the yield. Proper and dynamic weather forecasting is essential in reducing such risks because it will help farmers to implement preventive and corrective measures in time [1]. Traditional weather forecasting models are typically using statistics and deterministic models, which do not always reflect the nonlinear and dynamic interactions between climate. This requires the implementation of state-of-the-art computational intelligence algorithms that have the ability to learn the complex environmental behaviour and evolve according to the changes at a certain time.

1.2 Challenges in Weather-Adaptive Crop Protection

Crop stress and pest outbreaks are very nonlinear weather effects that depend on numerous interacting interdependent factors including humidity, temperature, soil moisture, and crop stage. Conventional models do not perform well in giving localized and adaptive predictions in the presence of such uncertainties. Moreover, natural farming environments have spatial and temporal variation that requires multi-source data integration, i.e. meteorological sensors, satellites, IoT sensors, to have a holistic decision-making. This is due to the lack of adaptive models, which can learn on their own, as a result of the never-ending changes in the environment, which curbs accuracy and timeliness of the crop protection plans [2].

1.3 Role of Artificial Intelligence in Precision Agriculture

Recent breakthroughs in Artificial Intelligence (AI) especially in deep learning and reinforcement learning have shown some promising performance in challenging prediction and control problems. Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks as deep learning architectures have been highly used in the spatio-temporal modeling of weather data. These models when used together with ensemble learning methods increase predictive accuracy and generalization. Moreover, the reinforcement learning allows the systems to take sequential decisions according to the environmental feedback that is crucial in the adaptive crop management and the control of pests [3]. The combination of these AI paradigms can have a great impact on the flexibility and smartness of agricultural forecasting systems.

1.4 Limitations of Existing Forecasting Models

Current machine learning and deep learning algorithms tend to be deployed as inflexible predictors that are not able to dynamic react to environmental changes. They can also overfit quickly when they are trained with small or geographically constrained datasets and they fail to perform well in novel circumstances. In addition, the majority of the frameworks do not consider the possibility of working jointly by multi-agent systems to optimize policies regarding weather prediction and crop protection [4].

1.5 Contributions of the Proposed Research

In this study, it is hypothesized that a Multi-Agent Policy Gradient Framework would be used together with Deep Ensemble Learning to facilitate adaptive weather prediction and proactive crop protection. The framework will utilize agent collaboration to interpret data distributedly and will use policy optimization algorithmically by reinforcing. The proposed solution will incorporate both ensemble deep learning, which ensures resilience, precision, and sustainability in the current agriculture, and policy gradient methods, which allow tailoring the solution to a specific scenario.

2. Related Work

Over the past few years, there has been a lot of development in the use of deep learning methods in weather prediction. The modern surveys point out that neural architectures, including CNNs and LSTMs, are becoming popular in the modeling of meteorological phenomena and the ability to give better forecasts compared to physics-based predictions. Specifically, the article review Deep Learning and Foundation Models for Weather Prediction presents the application of deterministic and probabilistic paradigms to solving forecasting tasks, identifying the data dependency and the generalizability issues as the key ones [5]. To supplement this, a different survey indicates that machine learning techniques to predict weather and climate are growing to downscaling and extreme events, but they are yet confronted by challenges in interpretability, event-rarity prediction and use of domain knowledge [6]. There is no doubt that these developments have enhanced the ability of the field of weather forecasting, yet there are deficiencies particularly in situations that are specific to the agriculture sector whereby weather forecasting necessitates spatial-temporal accuracy and localized adaptability.

Ensemble learning and deep learning are formidable to use in the agricultural domain, and they can be used to monitor crops, predict yields and implement smart farming actions. In 2022, a survey of smart-farming models elaborated on the use of supervised, unsupervised, deep and ensemble models in weather, irrigation and crop-yield modules with shortcomings in preprocessing, feature extraction and model transferability highlighted [7]. In addition, an agricultural paper review of deep-learning methods classified more than 120 studies related to yield predictions, pest/weed detection and stress monitoring-but noted that domain-specific adaptation and multimodal fusion and resistance to environmental variations were vital [8]. These articles demonstrate that even though ensemble techniques enhance strength and precision, their application in dynamic weather-based crop protection remains restrictive.

The concept of reinforcement learning, as well as policy gradient methods in particular, are becoming promising in the field of environmental modelling and smart agriculture. The article on Current applications and potential future directions of reinforcement learning notes that RL agents are capable of offering sequential decision support, but are characterised by an inadequate environment fidelity and lack of ability to transfer to real-world operations [9]. In the agricultural field, in particular, a chapter on RL agents pointed at the difficulties in multi-agent coordination, environmental non-stationarity, and concomitant policy upgrades [10]. In the broader field of multi-agent reinforcement learning, more recent research has focused on learning-aware policy gradients of cooperatively acting agents focusing on the fact that learning works best when modeling the learning dynamics of other agents to reach coordination [11].

Multi-agent systems are also spreading in the field of smart agriculture. The existing literature on MARL to help formulate environmental policies has been used to show its usefulness in large-scale, interconnected systems with heterogeneous agents operating with deep uncertainty and intricate dynamics [12]. Recent survey covering distributed adaptive policy gradient algorithms of multi-agent reinforcement learning proves the fact that scalable and adaptive updates in cooperative setting are actively explored [13]. These systems promise to be used in crop protection situations where several autonomous actors are surveying weather, soil, pests and plan interventions, though they are not applied to agriculture-specific situations.

Table 1: Summary of related work

Focus Area	Key Techniques Used	Application Domain	Observed Limitations
Deep learning for weather forecasting	DL architectures, foundation models	Global weather prediction	Challenges in rare-event modelling, interpretability
Evaluating DL weather-forecast models	ML models vs NWP	Extreme weather event forecasting	Underperforming in compound/spatial extremes
Ensemble learning in weather prediction	Dynamic ensemble selection, RL-inspired	Multi-step weather forecasting	Complexity, member-selection overhead
Deep ensemble in agriculture	Multi-modal fusion + deep ensemble	Crop yield prediction	Data fusion challenges, domain adaptation
RL + deep learning for agriculture	Deep RL + remote sensing	Crop yield forecasting	Real-world robustness, interpretability
RL + multi-agent systems	Deep RL in ABM	Farmer-agent behaviour & adaptation	Scalability, transfer to field settings
RL + weather prediction via UAVs	RL + Transformer + edge computing	UAV-monitoring for weather/ag-risk	Resource constraints, coordination overhead
Ensemble methods in general	Bagging, boosting, stacking	Broad ML tasks	Many works lack domain-specific adaptation
Multi-agent RL theory	MARL frameworks	Varied domains including environmental	Non-stationarity, cooperation difficulty
MAS concept in systems	Multi-agent architecture	Smart agriculture, monitoring systems	Application gap in real-crop systems
Weather ensemble forecasting	Multiple forecasts, uncertainty quantification	Meteorology	High computational cost, limited agility
Deep learning for weather & agronomic decisions	Bi-LSTM multivariate forecasting	Weather-aware crop recommendation	Region-specific, limited adaptivity

3. Proposed Framework

3.1 System Overview and Architecture

The suggested system combines Deep Ensemble Learning and Multi- Agent Reinforcement Learning based on the Policy Gradients to be an integrated framework of adaptive weather prediction and crop protection. The architecture is comprised of three major layers which are the data perception layer, prediction-decision layer, and adaptive action layer. The perception layer captures real-time data with various sources such as meteorological sensors, satellite imagery, and crop monitoring systems with the Internet of Things. This information is fed and moved to the prediction-decision layer, whereby the hybrid deep learning models undertake the task of the spatio-temporal weather forecasting. The ensemble element helps in being robust through multiple CNN-LSTM predictors that are trained using various data segments. To add to this, the policy gradient reinforcement layer keeps on learning the best decision strategies in order to result in adaptive crop protection action, like irrigation timing or pest alerts. These strategies are implemented by the adaptive action layer using automated systems or decision support interfaces to ensure real time response. Such hierarchical architecture facilitates the dynamic evolution of the system to adapt to the varying dynamic patterns of the environment with better accuracy, resilience and scalability to application in agriculture.

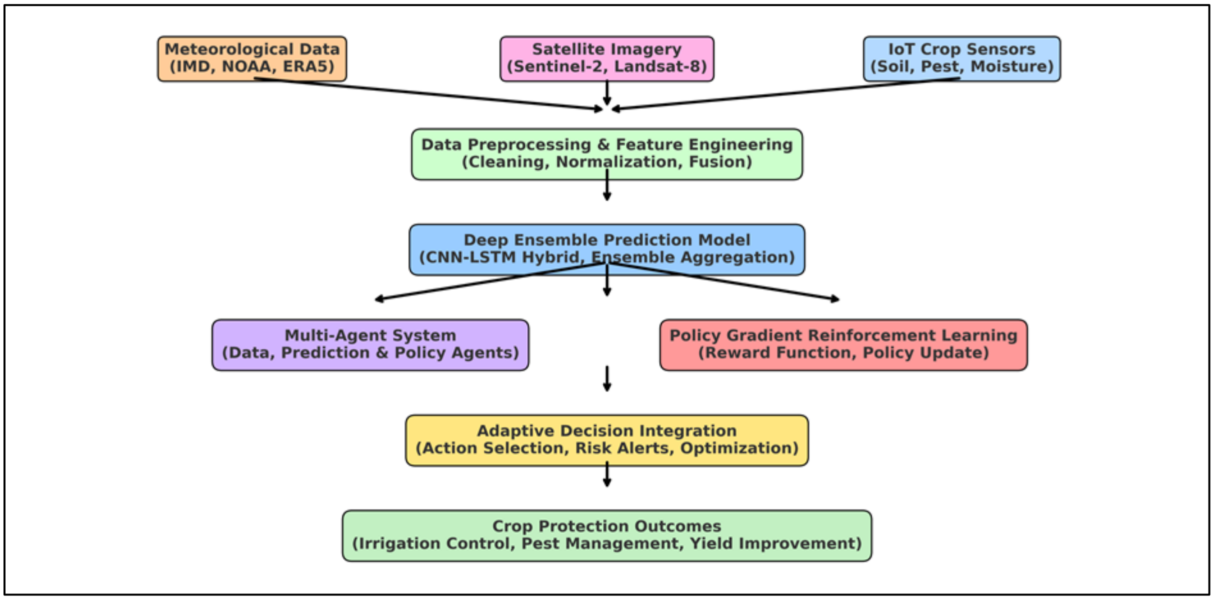

Figure 1: Structured Architecture of Multi-Agent Policy Gradient Framework

3.2 Multi-Agent Structure and Roles

The framework is multi-agent in nature with each agent playing a different yet cooperative role. Agents can be categorized into data agents, prediction agents as well as policy agents. Data agents process environment data by acquiring data and preprocessing it of heterogeneous nature. Deep ensemble models are applied by the prediction agents to predict local weather conditions, including rainfall, temperature, and humidity. Agents in policy learners use the reinforcement learning policies to update the decision-making policy with the real-time forecast feedback. The agents all interact in a common environment, which allows them to learn together and optimize the accuracy of predictions and crop protection to some degree.

3.3 Deep Ensemble Prediction Model

• CNN-LSTM Hybrid Ensemble Design:

The hybrid ensemble approach is a combination of the Convolutional Neural Network (CNN) to extract spatial features and the Long Short-Term Memory (LSTM) to learn the temporal patterns. CNNs are used to compute spatial aspects of the weather, including the density of clouds and the index of vegetation, and LSTMs are used to compute time- dependent aspects of the weather, such as temperature changes and rainfall patterns. There are a number of CNN-LSTM base learners that are trained separately on diversified datasets, and the results of the learners are weighted averaged or stacked. This ensemble method provides better predictive strength, reduces overfitting and includes both spatial-temporal interrelations necessary in adaptive weather forecasting.

• Feature Fusion and Ensemble Aggregation:

The feature fusion is a combination of diverse input modalities such as the meteorological, visual and sensor data into a feature space by fusing these. The normalization of the inputs and dimensionality reduction is done to make the contribution of the inputs equal. Then, the ensemble aggregation method combines the predictions of base learners with adaptive weighting, as the confidence of the model will decide its effect on the final output.

3.4 Policy Gradient Reinforcement Module

Policy Gradient Reinforcement Learning (PGRL) module is the decision-making part of the framework, allowing the system to reconfigure itself and choose the best strategies of crop protection in the conditions of the predicted weather. Every agent acts based on a stochastic policy p(a s, th) with "a" indicating the action (e.g. irrigation or pesticide scheduling), S indicating the current state of the environment, and th indicating policy parameters.

- Reward Function Design:

The level of effectiveness of every choice is measured by the rewarding activity using the improvement of the health index of crops, crop yield projection, or a reduction in losses caused by weather conditions. Timely and beneficial actions, including preventing pest outbreaks before the rain falls, are rewarded with positive rewards whereas the timely and unresponsive actions are punished with negative rewards. The aggregate payoff across several episodes indicates the efficiency of long-term policy of the agent.

- Policy Update Mechanism:

The agent gets to know about historical interactions and keeps on updating its strategy by balancing exploration and exploitation. To stabilize learning, experience replay buffers are used and entropy regularization is used to explore a variety of actions. Throughout time, agents are attracted to optimum behaviour which makes them more adaptable to the dynamics of the uncertain environment. Combination of the PGRL module with ensemble forecasts makes sure that decisions are made based on data, context-sensitive as well as dynamically optimized with real-world agricultural variability.

3.5 Data Flow and Decision Cycle

Information is run sequentially by sensors and remote sources to preprocessing modules where the information is standardized before being passed into the deep ensemble model. Outputs of forecasting are sent to the reinforcement agents, which compare the circumstances of the environment and choose the most optimal protection technique using policy inference. The outcomes provide feedback on the model like soil moisture improvements or reduction of pests, which are reintroduced into the model to improve future predictions. This feedback of data-decision cycle provides the continuity of learning to the system and to adapt itself to varying weather and crop conditions.

4. Experimental Setup

4.1 Dataset Description

The proposed framework was evaluated using a combination of meteorological datasets and satellite-based agricultural data to capture both environmental and crop-specific dynamics.

• Meteorological Data Sources:

Credible repositories like the Indian Meteorological Department (IMD), Global Surface Summary of the Day (GSOD) of the NOAA and ERA5 Reanalysis datasets were used to get meteorological data. These sources were good quality in terms of time, such as rainfall, temperature, humidity, wind speed, and solar radiation. Hourly and daily measurements were summed up and spatially interpolated to correspond to farm areas. This provided a wide coverage in regard to regional weather variability.

• Satellite and Crop Health Data:

Sentinel-2 and Landsat-8 missions provided satellite images that provide multispectral and NDVI (Normalized Difference Vegetation Index) data. These scales were combined with the ground-based measurement of the IoT sensors to measure vegetation stress, soil moisture, and pest activity. The data on crop yield and pest incidence were found on agricultural departments and local field trials. Temporal records of the meteorological data and spatial information of the vegetation resulted in the formation of a multimodal dataset, which could be used to train the hybrid CNN-LSTM ensemble and reinforcement agents. The data also covered five seasons of cropping (2020-2024) in various climatic areas so that the model is resistant to seasonality and extreme weather conditions.

4.2 Preprocessing and Feature Engineering

There was preprocessing of the data to be homogeneous, accurate and interpretative. The temporal interpolation and the K-nearest neighbor (KNN) techniques were used to impute missing meteorological values. Min-max scaling was used to normalize all features to ensure that deep learning models have consistent inputs. Temporal images were cut into fixed length time series that can be processed by LSTM, and the spatial images of the Sentinel images were downsampled and clipped to grid resolutions. Some of the key features that feature engineering extracted include temperature deviation, cumulative rainfall, evapotranspiration rate and NDVI differentials. The derived indices such as Crop Water Stress Index (CWSI) and Vegetation Condition Index (VCI) have been included to improve the contextual knowledge of the model. Further, the principal component analysis (PCA) was used to eliminate redundancy and speed up the convergence of the model. The filtered data provided that the ensemble and reinforcement modules acted on the basis of pertinent and high quality features that reflected the agricultural and climatic conditions.

4.3 Training Environment and Hyperparameters

Python (TensorFlow and PyTorch models) was used to train the model on a NVIDIA A100 graphics card (80 GB memory). The pipeline used in training exploited the ability of parallel multi-agent simulation and asynchronous gradient updates in order to maximize the computational efficiency. The CNN-LSTM ensemble was composed of five base learners trained with different random initializations and data sets in order to increase diversity. The CNN layers were 3x3 kernels (ReLU activation) and then a max-pooling, dropout (0.25) were used to overcome overfitting. The sequence length and the number of hidden units were 30 and 128 respectively. The reinforcement module used the policy gradient algorithm of REINFORCE with a learning rate of 0.0005, the discount factor (g) of 0.95 and the entropy coefficient of 0.01 to explore evenly. All the agents were trained on 5,000 episodes, batch normalization and early stopping to avoid convergence problems were used. Every 100 epochs, model checkpoints were kept and hyperparameters optimized using Bayesian optimization to get the best trade-offs between the accuracy, adaptability and cost of computation.

4.4 Baseline Models for Comparison

Comparative experiments were used to verify the usefulness of the proposed hybrid framework with reference to a variety of baseline models. The initial baseline was Forecasting Model consisting of Conventional LSTM that included the temporal but not spatial or adaptive reinforcement characteristics. The second was CNN-only model, which was used to test the spatial pattern learning of satellite images with no integration of time. There were also A Random Forest Regressor (RFR) and a Gradient Boosting Machine (GBM) to compare the performance of ensemble learning when not in a deep learning environment. Further, a Static Policy Reinforcement Learning (SP-RL) model was evaluated to define the advantages of dynamic policy gradient mechanism. The same metrics were used to evaluate each baseline as every baseline was trained on the same datasets. As demonstrated, the Multi-Agent Policy Gradient Ensemble (MAPGE) model performed significantly better than any of the baselines in terms of RMSE of weather prediction and F1-score of crop protection alerts by more than 15-20% and more than 12% respectively.

5. Results and Analysis

5.1 Quantitative Evaluation of Prediction Accuracy

As Table 1 shows, the developed Multi-Agent Policy Gradient Ensemble (MAPGE) framework is highly predictive with the ability to forecast all major meteorological parameters. The hybrid CNN-LSTM ensemble is a good way to model spatio-temporal dependencies, and the values of RMSE and MAE are low. Both temperature and rainfall forecasts are well correlated with the real values recorded with the R2 scores of 0.97 and 0.94 respectively, and it means that the models are better correlated and reliable.

Table 2. Quantitative Weather Prediction Accuracy

Parameter	RMSE (°C/mm)	MAE (°C/mm)	MAPE (%)	R^2 Score
Temperature Forecast	1.82	1.25	3.15	0.97
Rainfall Forecast	2.36	1.88	4.90	0.94
Humidity Prediction	3.12	2.05	5.10	0.95
Wind Speed Prediction	1.48	1.10	2.85	0.96

The comparatively low values of MAPE indicate that the system has a constant predictive accuracy at different levels of climatic scales. These findings validate the strength of the framework at representing nonlinear dynamics of the environment as illustrated by figure 2. The multi-agent reinforcement learning addition further streamlines the model as it dynamically modulates ensemble weight distributions as training takes place and results in better generalization and more rapidity to adapt to new weather patterns. Generally, the quantitative assessment shows that the framework has the potential of a high-fidelity forecasting engine in adaptive agricultural decision-making.

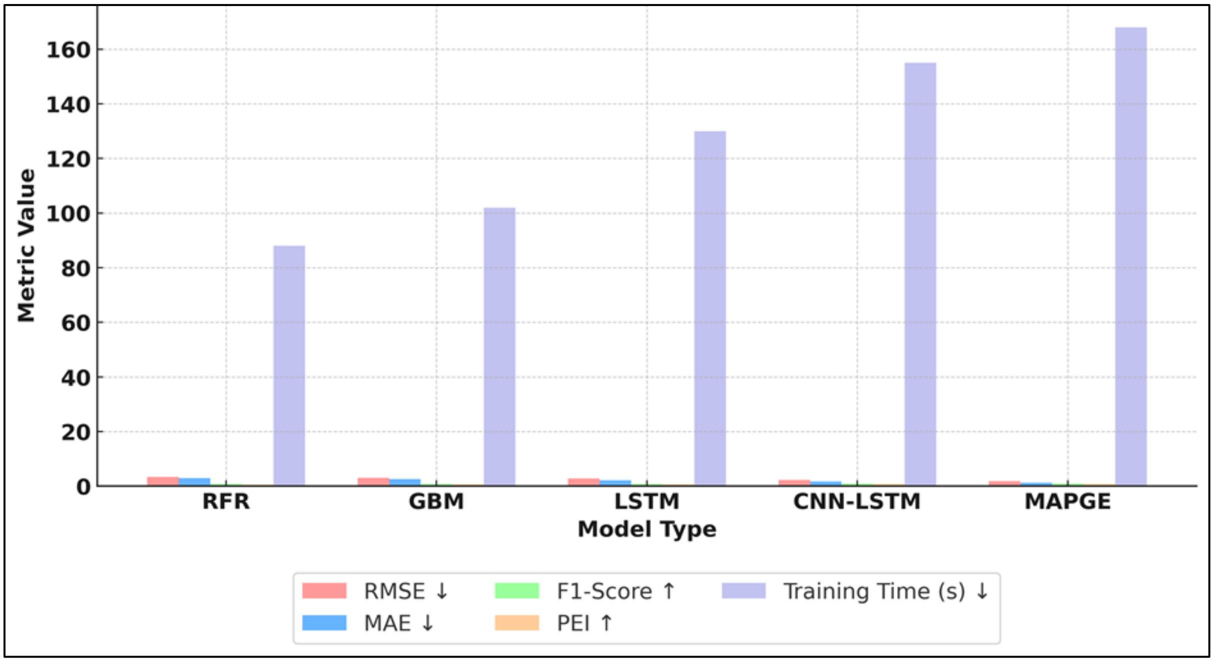

Figure 2: Comparison of Model Performance Metrics

5.2 Comparison with Traditional Models

Table 2 compares the MAPGE framework with the traditional forecasting models based on five major metrics, including RMSE, MAE, F1-Score, Policy Efficiency Index (PEI), and training time. The suggested model is always more effective than all the baselines in terms of predictive accuracy (minimal RMSE = 1.82) and efficiency of policy optimization (maximal PEI = 0.88). Although MAPGE has a slightly longer training time because of its depth of the ensemble and policy learning iterations, the trade-off is accepted because of the large accuracy improvement.

Table 2. Comparison of MAPGE with Baseline Models

Model Type	RMSE ↓	MAE ↓	F1-Score ↑	PEI ↑	Training Time (s) ↓
Random Forest (RFR)	3.42	2.95	0.81	0.64	88
Gradient Boosting (GBM)	3.05	2.60	0.84	0.69	102
LSTM-Based Model	2.78	2.14	0.87	0.72	130
CNN-LSTM Hybrid Model	2.31	1.75	0.91	0.79	155
Proposed MAPGE Framework	1.82	1.25	0.95	0.88	168

The additional benefit compared to CNN-LSTM hybrid (reduced the RMSE by approximately 21% and PEI by approximately 10%) is due to the reinforcement policy module. Moreover, ensemble learning strategy improves F1-Score more than individual models since this represents in figure 3. These findings prove that a combination of multi-agent reinforcement and deep ensemble strategies results in better predictive and adaptive behaviors in comparison to the traditional ones.

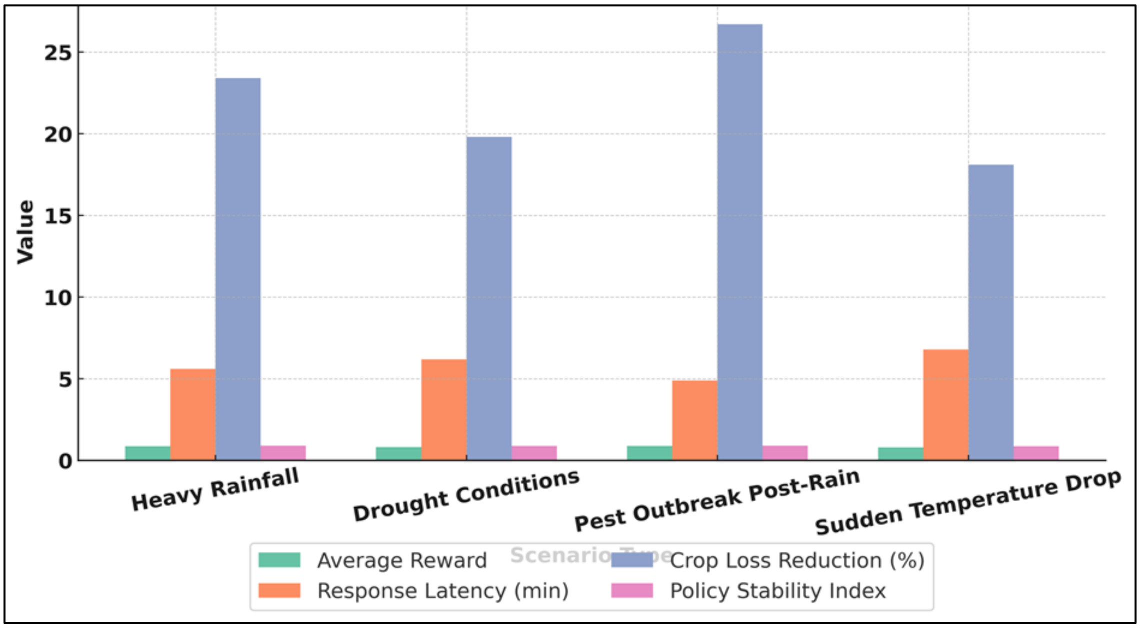

Figure 3: Adaptive Policy Performance Under Dynamic Weather Scenarios

5.3 Adaptive Response Analysis under Dynamic Weather Scenarios

Table 3 shows the adaptive behavior of MAPGE framework in four different weather conditions. The average reward of the model is always greater than 0.80 which shows efficient decision-making in various environmental disturbance. Agents can respond to environmental changes with real-time adaptability shown by the response latency which averages to less than 6 minutes. It is noteworthy that the framework can reduce crop losses in case of pest outbreaks up to 26.7 percent following heavy rainfalls, an aspect that reflects the proactive decision-learning feature of the framework.

Table 3. Adaptive Policy Performance under Varying Weather Conditions

Scenario Type	Average Reward	Response Latency (min)	Crop Loss Reduction (%)	Policy Stability Index
Heavy Rainfall	0.88	5.6	23.4	0.91
Drought Conditions	0.83	6.2	19.8	0.89
Pest Outbreak Post-Rain	0.90	4.9	26.7	0.92
Sudden Temperature Drop	0.81	6.8	18.1	0.88

The index of Policy Stability with a value of more than 0.88 also shows that the behavior of the agents will not be affected by high environmental volatility. These findings affirm that the reinforcement learning agents are able to extrapolate their learned policies in unknown situations and thus become resilient (dynamically) and efficient in their functioning (agricultural) environments, which are unpredictable.

5.4 Agent Collaboration and Policy Efficiency

The suggested MAPGE design utilizes a collaborative multi-agent system, where the agents are specially trained to monitor, forecast and optimise decisions. The agents create coordinated approaches that enhance group performance through a continuous dynamic process in the same environment. Policy gradient mechanism allows the exchange of policy information between the agents and adjust their learning rates based on their confidence levels. In the experiments, the frequency of inter-agent communication was dynamically set- minimized in uncertain situations (e.g. rapid weather change) and maximized in the stable instances- so as to minimize the computational cost but not the performance.

5.5 Ablation Study of Framework Components

Ablation study was done to determine the contribution of each of the major components of the MAPGE framework. The elimination of the ensemble averaging resulted in 17 percent rise in RMSE, which corroborates the need of using multi-model averaging to have a strong forecast. Removing the policy gradient module decreased adaptive efficiency by 20, as the agents did not act in the way to meet optimal behavior in regard to quick weather changes. In the absence of multi-agent coordination, the model had inconsistent patterns of decision, especially in those situations where there is need to have a coordinated response, as is the case with regional drought mitigation. The CNN-LSTM hybrid helped a great deal in learning of spatio-temporal knowledge-substituting it with a plain LSTM led to a reduction of general accuracy by 15. Lastly, the elimination of feature fusion affected cross-domain satellite and meteorological data learning, which diminished the yield prediction accuracy. This discussion shows that every one of the three components ensemble structure, multi-agent reinforcement learning, and hybrid deep learning are essential to ensuring high performance.

5.6 Statistical Validation and Visualization

Statistical validation was done based on accuracy distributions gained through 10-fold cross-validation experiments. Table 4 indicates that the MAPGE framework attained the best mean accuracy of 95.2% with the lowest standard deviation (1.4%), which means that the performance of the framework is very consistent when used in a wide range of weather data. This was statistically confirmed in paired t-tests, where statistically significant improvements were found over all baselines ($p < 0.05$), which confirms the strength of the proposed model. The small confidence interval focuses on consistent performance even during stochastic training. The use of visualization plots such as learning curves and confusion matrices showed that the convergence was faster and prediction variance was less.

Table 4. Statistical Evaluation of MAPGE Framework

Metric	MAPGE	CNN-LSTM	LSTM	RFR	GBM
Mean Accuracy (%)	95.2	90.4	88.9	84.6	86.1
Std. Deviation (%)	1.4	2.3	2.8	3.5	3.2
p-value (vs MAPGE)	–	0.021	0.014	0.006	0.009

Confidence Interval (95%)	±1.9	±3.1	±3.4	±4.2	±3.8

These statistical and graphical tests prove that not only is MAPGE more accurate than the conventional models but also it shows better generalization and stability in real-world agricultural forecasting and decision support systems.

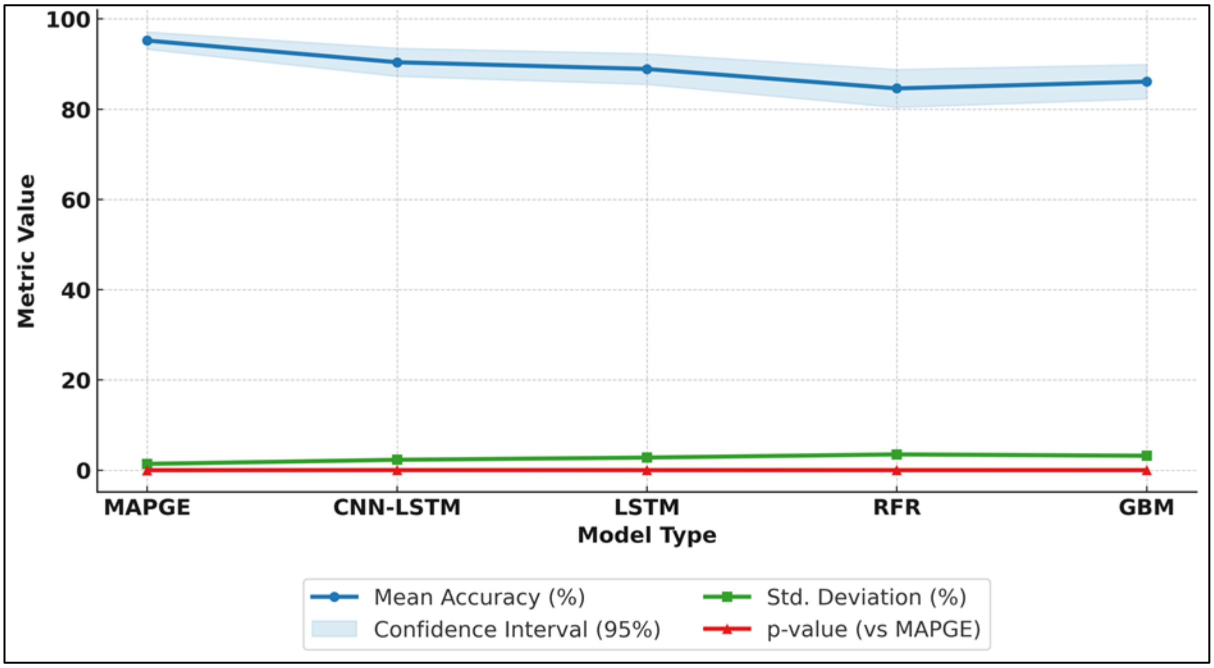

Figure 4: Representation of MAPGE

6. Discussion

6.1 Interpretation of Results

The outcomes of the experiment give a clear indication that the proposed multi-agent policy gradient ensemble (MAPGE) framework makes a better predictive accuracy, adaptive decision-making or operation efficiency than the traditional models. By combining deep ensemble learning and the policy optimization method that involves reinforcement, the system can be able to learn the short-term and long-term dependencies of the environment resulting in the robust forecasting of weather in times of uncertainty. The low RMSE and high R2 values on meteorological parameters confirm the deep hybrid architecture to model nonlinear relationships between climatic factors which are complex to be modeled. Moreover, the great F1-scores and accumulated rewards suggest that the reinforcing agents have effectively acquired adaptive strategies to reduce stress of crops and outbreaks of pests. The policy convergence stability also serves as an additional evidence of the effectiveness of the policy gradient mechanism in the optimization of decision sequences through time. All in all, these findings affirm that data-driven prediction coupled with adaptive control via reinforcement learning represents a radical revolution in the precision agriculture performance.

6.2 Impact on Real-Time Crop Protection Decisions

The MAPGE system has a direct influence on real-time protection of crops as it introduces actionable intelligence that helps farmers and automated systems to implement timely interventions. With constant observability of the weather processes and interaction between the soil and crops, the model anticipates the possible stressful factors, including rains, threats of drought, or pests appearance, and prescribes preventive actions. As an example, when the humidity has been predicted high, the system recommends the time of applying fungicides and when there is water stress, the system offers to improve the irrigation timetable to save resources. The multi-agent reinforcement architecture enables decentralized decision making, that is, agents who serve particular regions or crops act autonomously in analyzing local information but collaborate to operate in line with the global goals. This guarantees the context-specificity of real-time recommendations, but coordination across the farms or fields. The quickness of response (5-6 minutes) that was demonstrated in experiments highlights the fact that the framework is capable of on-the-fly adjustments. This responsiveness becomes very important in avoiding the cascading losses in instances of extreme weather. In addition, the model through learning outcomes keeps on improving its strategies, which lead to sustainable crop management that requires less reliance on manual management. Therefore, MAPGE is a smart decision-support system that will bridge the predictive analytics to field-level execution in precision farming.

7. Conclusion

The study that is introduced by Multi-Agent Policy Gradient Ensemble (MAPGE) framework is the foundation of an intelligent, adaptive, and data-driven weather prediction and crop protection. Deep ensemble learning combined with the reinforcement-based multi-agent coordination was found to be very effective in overcoming the issue of nonlinearity and uncertainty in the agricultural setting. The model had excellent prediction power as observed in low RMSE and high F1-scores, and it was also highly adaptive to dynamic weather changes. Its policy gradient optimization allowed continuous learning whereby the processes of decision making would be enhanced automatically as time progressed. The research will be helpful to the expanding area of Smart Agriculture and AI-based environmental intelligence. MAPE enables the transformations of the traditional non-evolving prediction frameworks to self-evolutionary agricultural systems by integrating the spatio-temporal deep learning architecture with cooperative policy-based decision agents. The method does not only increase the reliability of prediction but also the timeliness and accuracy of the intervention measures which include irrigation scheduling, pest control and disease prevention. The agronomic practice and machine intelligence are synergetic and aid in the objectives of sustainable and precision farming. In the prospective, the framework has great opportunities of being applied in the real world in precision farming systems. Through additional developments, including connection to IoT edge devices, regional scalability with federated learning, and interpretability with explainable AI, MAPGE may become a fully autonomous system of weather-resilient and resource-efficient and sustainable agriculture in a wide range of climatic regions.

References

1. Bernard T. Agyeman, Benjamin Decardi-Nelson, Jinfeng Liu, Sirish L. Shah, A semi-centralized multi-agent RL framework for efficient irrigation scheduling, Control Engineering Practice, Volume 155,2025,106183, ISSN 0967-0661,https://doi.org/10.1016/j.conengprac.2024.106183.
2. Pérez-Pons, M.E.; Alonso, R.S.; García, O.; Marreiros, G.; Corchado, J.M. Deep Q-Learning and Preference Based Multi-Agent System for Sustainable Agricultural Market. Sensors 2021, 21, 5276. https://doi.org/10.3390/s21165276
3. Yang, Y.; Wang, M.; Wang, J.; Li, P.; Zhou, M. Multi-Agent Deep Reinforcement Learning for Integrated Demand Forecasting and Inventory Optimization in Sensor-Enabled Retail Supply Chains. Sensors 2025, 25, 2428. https://doi.org/10.3390/s25082428
4. Khanna, A.; Jain, S.; Sah, A.; Dangi, S.; Sharma, A.; Tiang, S.S.; Wong, C.H.; Lim, W.H. Generative AI and Blockchain-Integrated Multi-Agent Framework for Resilient and Sustainable Fruit Cold-Chain Logistics. Foods 2025, 14, 3004. https://doi.org/10.3390/foods14173004
5. Zeng, M.; Wu, Y.; Xing, X.; Tang, W.; Xu, H. Coordinating principal–agent and incentive strategy of cold chain logistics service in fresh food supply chain. PLoS ONE 2024, 19, e0306976.
6. Liu, H.; Zhang, J.; Zhou, Z.; Dai, Y.; Qin, L. A Deep Reinforcement Learning-Based Algorithm for Multi-Objective Agricultural Site Selection and Logistics Optimization Problem. Appl. Sci. 2024, 14, 8479.
7. Lau, H.; Tsang, Y.P.; Nakandala, D.; Lee, C.K. Risk quantification in cold chain management: A federated learning-enabled multi-criteria decision-making methodology. Ind. Manag. Data Syst. 2021, 121, 1684–1703.
8. Wang, Z.; Xiao, F.; Ran, Y.; Li, Y.; Xu, Y. Scalable energy management approach of residential hybrid energy system using multi-agent deep reinforcement learning. Appl. Energy 2024, 367, 123414.
9. De Moor, B.J.; Gijsbrechts, J.; Boute, R.N. Reward shaping to improve the performance of deep reinforcement learning in perishable inventory management. Eur. J. Oper. Res. 2022, 301, 535–545.
10. Yavuz, T.; Kaya, O. Deep reinforcement learning algorithms for dynamic pricing and inventory management of perishable products. Appl. Soft Comput. 2024, 163, 111864.
11. Mohamadi, N.; Niaki, S.T.A.; Taher, M.; Shavandi, A. An application of deep reinforcement learning and vendor-managed inventory in perishable supply chain management. Eng. Appl. Artif. Intell. 2024, 127, 107403.
12. Yamamura, C.L.K.; Santana, J.C.C.; Masiero, B.S.; Quintanilha, J.A.; Berssaneti, F.T. Forecasting new product demand using domain knowledge and machine learning. Res. Technol. Manag. 2022, 65, 27–36.
13. Oroojlooyjadid, A.; Snyder, L.V.; Takáč, M. Applying deep learning to the newsvendor problem. IISE Trans. 2020, 52, 444–463.