**Theme: Sustainable Agricultural Innovations for One Health and Environmental Resilience**

RESEARCH ARTICLE Indian Journal of Agricultural Research, Volume 60 Special Issue: 15-24 (2026)

# AgroAdvisor: RAG-Powered Explainable AI for Soil Nutrient Profiling and Crop Yield Recommendations

P. Nithya[1], G. Senthil Kumar[1]

**10.18805/IJARe.A-6581**

## ABSTRACT

**Background:** Determining the health of soil and predicting the output of agriculture are complicated processes owing to regional variation and complex nutrient interactions among soil properties.

**Methods:** AgroAdvisor acts as a decision support system using predictive models, particle swarm optimization (PSO) and retrieval-augmented generation (RAG). Soil health parameters such as nitrogen (N), phosphorous (P), potassium (K), pH and moisture levels are analyzed through a hybrid regression model incorporating random forest and gradient boosting. PSO is used for hyperparameter optimization, which improves the accuracy of prediction compared to traditional methods.

**Result:** Experimental testing on soil samples from South India demonstrated a decrease in mean absolute error ranging between 18% to 23% due to PSO optimization. The retrieval-augmented generation technique, based on scientific papers, ICAR/FAO guidelines and local soil management techniques, generated contextually relevant and internally consistent recommendations for fertilizer usage, crop rotation and soil management.

**Key words:** Agricultural productivity, Particle swarm optimization (PSO), Regression models, Retrieval-augmented generation (RAG), Soil health.

## INTRODUCTION

Nevertheless, agriculture remains an important industry, where new innovations should be applied to satisfy future demands. It is critical to merge the practice of growing crops with modern technologies since food production is necessary for all humans. One issue that often challenges farmers is the application of crop rotation, which is quite common due to decreasing efficiency of the process of growing the same plants over several years in the same place. Therefore, soil fertility and its determining factors play a crucial role in choosing highly-productive crops.

Smart farming can be considered as a new trend in this field of study focused on overcoming different problems of the agricultural sector and ensuring food safety. Smart farming technologies involve using various technologies merged with the practice of cultivating plants relying on collected data. For example, smart farming is applied in Chile to increase water efficiency during irrigation, reducing the quantity needed by 70% during the blueberries' harvest. Similarly, this technology has been used in India for preventing plant diseases and pest invasions. The most important soil-related characteristics include moisture and the concentration of important nutrients (N, P, K and S, among others). Moisture is the amount of water stored between soil aggregates. Insufficiency of nutrients might hamper the development of crops and thus the need to fertilize the soil emerges. Hence, collecting data on soil moisture and nutrient contents allows us to choose crops in accordance with the existing parameters or procure necessary fertilizers in case of their insufficiency.

Predicting these parameters can be done via machine learning techniques. One study relied on levenberg-marquardt-driven backpropagation network for the same purpose while another used partial least squares regression to estimate clay content, electrical conductivity and bulk density. Humidity and weather were predicted with the help of Bayesian networks. Obtained data allow constructing index values related to different parameters like soil pH level, moisture, humidity, weather, temperature and the level of N, P, K (Fig 1).

A population-based randomized optimization approach known as particle swarm optimization (PSO) was created by Kennedy and Eberhart (1995) and is often used as a group intelligent search tool. It may reduce losses in adverse circumstances and to increase crop production in good ones (Rajak *et al*., 2017). It is possible to estimate agricultural production using machine learning methods, since the connection between yield and variables is not linear.

Farmers' losses due to grass grubs were examined by Oerke and Dehne (2004). To evaluate crop damage, well-known classifiers including decision trees, random

[1]Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur-603 203, Tamil Nadu, India.

**Corresponding Author:** P. Nithya, Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur-603 203, Tamil Nadu, India. Email: nithyap2@srmist.edu.in
ORCIDs: 0009-0009-0885-9630, 0000-0002-0447-2086

**How to cite this article:** Nithya, P. and Kumar, G.S. (2026). AgroAdvisor: RAG-Powered Explainable AI for Soil Nutrient Profiling and Crop Yield Recommendations. *Indian Journal of Agricultural Research.* **60(Special Issue):** 15-24. doi: 10.18805/IJARe.A-6581.

**Submitted:** 08-05-2026 **Accepted:** 02-06-2026 **Online:** 07-07-2026

forests, naive bayes, support vector machines and KNNs (K-nearest neighbors) (Ayub and Moqurrab, 2018) were used. Recent advances in ML for precise plant disease detection show that deep learning models outperform traditional classifiers in identifying pathogen-induced damage (Lee and Kim, 2024) and early detection algorithms for legume crop diseases further confirm the practical utility of such models (Cho, 2024). To classify Topical Ecosystem using satellite data (Poutea *et al*., 2011), evaluated six machine learning methods (KNN, Boosted Regression Tree, Naive Bayes, RF, C4.5 and SVM) using six distinct sensors. The effects of classification selection, reference sample size, number of attributes and scene heterogeneity on per-pixel classification accuracy were investigated by Heydari and Mountrakis (2018); variability in classification accuracy across different conditions was further examined by Meenakshi and Naresh (2022a; 2022b; 2022c; 2023a).

## MATERIALS AND METHODS

### Sampling site and sample collection

The fertile soils of Thanjavur, Tiruvarur and Tiruchirappalli districts were considered due to the fact that they have been used for farming practices according to previous studies. The measurements for soil moisture, weather conditions, humidity, temperature and macronutrient nitrogen, phosphorus and potassium (NPK) were performed *via* respective sensors. Soils samples have been collected and tested at different levels: 0-15 cm, 15-30 cm and 30-60 cm. The productivity forecasting was performed for 1-15 cm depth level; these forecasting hypotheses concern the enrichment of nutrients in the deeper levels. It is suggested that nitrate nitrogen ($NO_3$-N) and sulfate sulfur ($SO_4$-S) can occur in significant amounts in the soils between 15 and 60 cm. Both nitrate and ammonium ions provide nutrition in the form of nitrogen needed for the synthesis of amino acids. Thus, while nitrogen-rich soil can stimulate growth, nitrate can accelerate it even more. Therefore, the parameters of $NO_3$ and $SO_4$ need to be assessed for the entire range of depths between 0 and 60 cm as the analysis conducted only on 0-5 cm might misrepresent N and S parameters, especially in case of phosphorus, which is mobile and mostly limited to the ploughing soil of 0-15 cm depth. In turn, majority of phosphorus and potassium occurs in the soils up to 0-1 cm depth. In these soils, data sets include 10 attributes and total 222,853 soil sample records for these parameters. The quantitative and qualitative data collection workflow adopted in this study is illustrated in Fig 2. Real-time IoT-based systems have been shown to enhance soil parameter monitoring for crop prediction (Pandey *et al*., 2023). The mathematical symbols and abbreviations used throughout the PSO formulation and regression equations in the following sections are summarized in Table 1.

Advancements in XAI have greatly improved the explainability of ML models employed in vital agricultural applications. Techniques like SHAP and LIME have been combined with ensembles for generating intelligible predictions for soil nutrients assessment and yield prediction (Sharma *et al*., 2021). Deep learning architectures integrating CNN-LSTM with attention mechanisms have demonstrated strong performance in crop yield prediction (Kalmani *et al*., 2025) and a comprehensive review of ML models for plant disease prediction and detection further underscores the value of robust ensemble approaches in agricultural decision support (Metagar and Walikar, 2024). RAG systems have proven to be a powerful paradigm in intelligent agricultural analysis by enabling domain-specific decision-making from the outputs of large language models anchored on structured knowledge bases comprising of agronomy principles, soil science publications and local farming data (Kamilaris *et al*., 2016). The combination of RAG with a regression pipeline optimized using the PSO algorithm in AgroAdvisor is a novel development in the intersection of intelligent information retrieval and precision agriculture. Prior research has established that the integration of PSO optimization for hyperparameter tuning alongside knowledge-based retrieval results in lower recommendation latency and higher contextual relevance than deep learning algorithms alone (Stafford *et al*., 2019).

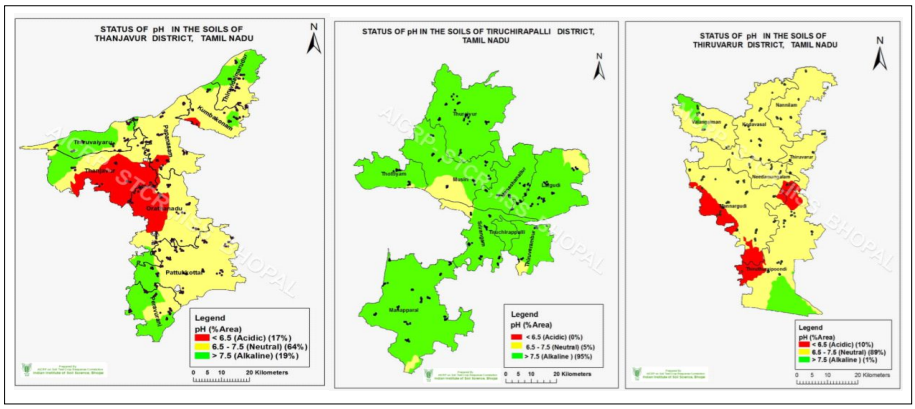



**Fig 1:** pH level located regions and study areas (iiss.nic.in).

### Particle swarm optimization (PSO) for regression hyperparameter tuning

#### PSO introduction

Particle swarm optimization (PSO) represents a stochastic population-based optimization technique initially proposed by Kennedy and Eberhart (1995) in imitation of avian flocking and schooling fish social behavior dynamics. The computationally-efficient optimization tasks, including hyperparameter optimization for machine learning models used in agriculture in Fig 3.

#### PSO initialization and particle velocity update

Optimization in PSO begins with creating a swarm that consists of N particles that are randomly placed in a D-dimensional search space, where each dimension represents an optimized hyperparameter (L2 regularization, kernel choice, or learning rate, *etc*.). In turn, every particle is initialized with some random position and velocity vectors that are then updated in iterations according to the following equations:

$$vel_i^k(t+1) = Weight\,[vel_i^k(t)] + C_1.random()\,[Par_i^k(t) - Pos_i^k(t)] + C_2.random()\,[Par_i^k(t) - Pos_i^k(t)] \quad ...(1)$$

w= The inertial weight.

$C_1$= The cognitive coefficient corresponding to the personal best position of the particle.

$C_2$= The social coefficient corresponding to the global best position of the whole swarm.

r1 and r2= Random numbers in [0,1] range.

pbest (i)= The current personal best position of the $i^{th}$ particle.

gbest= The global best position of the entire swarm.

#### PSO algorithm-step-by-step

**Step 1:** Initialization.

Create a starting population of N particles that will be randomly distributed within the D-dimensional search space. Initialize their positions and velocities accordingly.

**Step 2:** Calculate fitness.

Calculate the fitness function for every particle according to the selected measure. As far as the regression task is concerned, for AgroAdvisor, one can use MAE or R2 scores as a fitness function when the model with particular hyperparameters trained (by setting hyperparameters' values equal to the coordinates of the particle's position).

**Step 3:** Comparing each particle's fitness with its prior best achieved fitness should be done after each iteration. It is preferable to have the current value equal to the present values than to have the actual position equal to the present position in d-dimensional space and vice versa if the present values are greater.

**Step 4:** Update position and velocity.

Using the equation above, calculate a new position and velocity for every particle.

**Step 5:** Termination criteria.

Repeat steps 2-4 until the termination criterion, such as a maximal iteration limit or improvement on gbest below a certain threshold, is met.

AgroAdvisor's reproducibility was ensured through fixed hyperparameters and validation strategies. The hyperparameters used for PSO include setting N to 30 particles, with 100

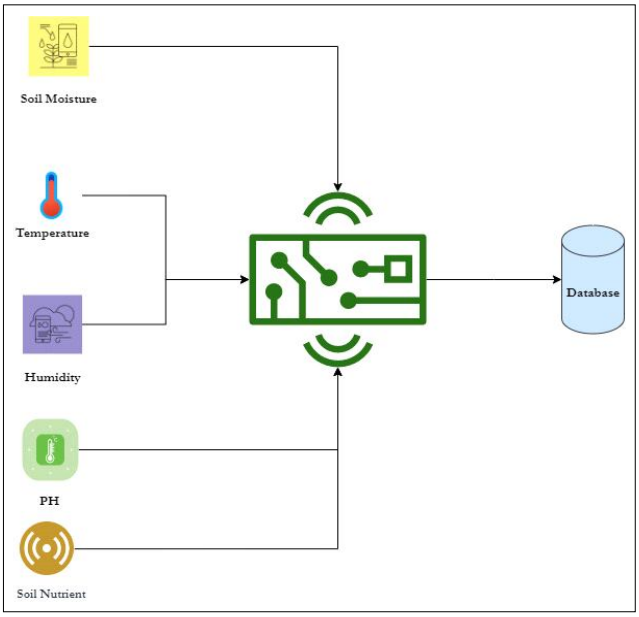



**Fig 2:** Soil quantitative and qualitative data collection.

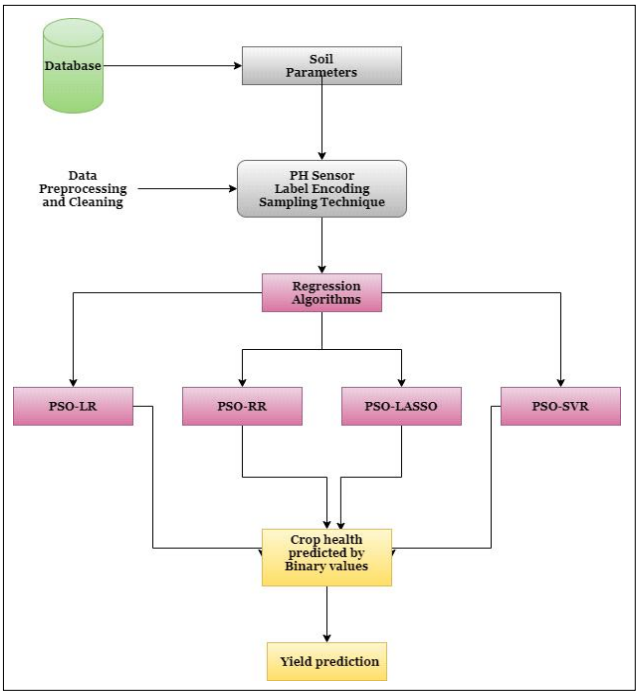



**Fig 3:** The use cases of PSO optimization in agroadvisor, namely, three regression models.

iterations as the maximum, an initial value of w at 0.9, a final value of w at 0.4 and c1 and c2 both equal to 2.0. The criteria for convergence entail that the difference in gbest must be less than 1e-6 in the last ten iterations.

**Advantages of PSO in AgroAdvisor**

- PSO involves fewer parameters compared to genetic algorithm or simulated annealing techniques.
- No gradients required in the optimization process, thus supporting non-differentiable fitness functions.
- Fast convergence on continuous hyperparameter ranges for regression models.
- Multiobjective approach for concurrent hyperparameter optimization in fertility scores and yield prediction.
- Scalability across different data set sizes from small farm data sets to large regional data sets.

**Proposed PSO-regression analysis**

**Feature selection**

The combination of several layered generalizations, applied to a PSO algorithm, is shown in Fig 3. PSO Regression is a technique of dividing into the following steps:

**Step 1:** Create the generalization classifier with many layers of generalizations.

We used the PSO-Regression that discusses in Feature selection. Grid search hyper-parameter tweaking was used to optimize each model in the PSO-Regression system.

**Step 2:** Determine the location of each position of the particle by PSO.

The feature subset for the particle location was found in this phase when the stacked generalization was produced.

**Step 3:** The process of verifying the most favorable combinations of swarms and particles.

To validate these two equations are used in these methods:

$$C_i^{pb} \leftarrow C_i \text{ if } f(C_i) > f(C_i^{pb}) \quad ...(2)$$

$$C_i^{sb} \leftarrow C_i \text{ if } f(C_i) > f(C_i^{sb}) \quad ...(3)$$

The variable is first entered into the PSO programmer, which returns a cost score. Then the score is shortened by the weakest score. Input data groups with one parameter into SVR program after being shortened.

**The PSO-regression feature selection technique**

The suggested approach is designed (Fig 4) to overcome the two issues around identifying the specific benefits of an individual's actual genome structural features, as well as figuring out a better way to make use of this data without the inherent drawbacks of over-fitting.

**PSO-regression pseudocode**

**1. Input:** t1, t2,..., tn are the training sets (folds), whereas v1, v2,..., vn are the validation sets (folds).
**2. Output:** The outcome of prediction.
**3. Initialize:** The swarms Xi, the particles positions, the populations posture, the Regression and Kernal parameters are all taken into consideration in this model.
4. Eq(7), determine the parameter fitness (G).
5. Using f ($X_i$), determine the efficiency of every particle (t).
6. A comparison is made between the efficiency of every individual and their previous best score.
7. if.
8. f [$X_i$ (t)] G < f ($p_{ibest}$) then.
9. f [$X_i$ (t)] = f ($p_{ibest}$).
10. ($p_{ibest}$) = f [$X_i$ (t)].
11. End if.
12. if.
13. f [$X_i$ (t)] G < f ($p_{Gbest}$) then.
14. f [$X_i$ (t)] = f ($p_{Gbest}$)
15. ($p_{Gbest}$) = f [$X_i$ (t)].
16. End if.
17. Calculate the particle and velocity of the data-set.
18. Repeat steps 7 till the implementation is complete.
19. Sorting of fitness is used to classify all of the particles.

**PSO-support vector regression (SVR)**

The input data must be transformed into the high-dimensional feature set using a nonlinear transformation matrix; this function must be defined beforehand. Once a linear function is defined in the high-dimensional feature space, it is conceptually possible to construct the nonlinear connection between the actual output. It is possible to define a linear model of this kind as follows:

$$\text{Trained data set } (X_i, O_i)_{i=1}^{n} \quad ...(4)$$

Where,
$X_i \in I^n$= Input(I) vector.
$O_i$= Output value(O).
n= Data set total dimensions.

The goal of modelling is to find a linear regression model.

$$Y = f(X) \quad ...(5)$$

That correctly predicts fresh I/O instances. For example, in feature space.

$$f(X_i) = \alpha \emptyset (X) + B \quad ...(6)$$

Where feature space (F)

$$\emptyset : I^n \rightarrow F, \alpha \varepsilon F \quad ...(7)$$

The following is how the actual evaluation is described:

$$\text{Eemp } (f) = \frac{1}{n} \Sigma_{i=1}^{n} l_i [Y_i, f(X_i)] \quad ...(8)$$

$$\text{Minimize} = \text{MIN} \frac{1}{2} ||w||^2 \quad ...(9)$$

$$\text{Constraints} = |y_i - w_i x_i| \leq \in \quad ...(10)$$

## RESULTS AND DISCUSSION

**PSO-linear regression**

The result of linear regression analysis by Gaussian Lambda family and normalization of linear regression is

illustrated in Fig 5a and Fig 5b. The output of the Root Mean Square error[RMSE] value on Train data is 0.48 and Test data is 0.47 on 1113 degrees of freedom and its multiple R - squared value was obtained on Train data and Test data is 0.82.

**Table 1:** Mathematical notations.

| Symbols | Abbreviation |
|---|---|
| F | Fitness |
| O | Observed values |
| zfi-zoi | squared differences |
| N | sample size |
| I | Variable |
| Z | Variance |
| $C_i$ | Independent variable |
| $y_i$ | Dependent variable |
| M | Slope/fit |
| $C_i^{pb}$ | Soil best position |
| Csb | Soil best Swarm |
| $Ø^i$ | Hyperparameter |
| $I^n$ | Input value |
| $O_i$ | Output value |
| Ø (X) | Feature space with high dimensions |
| P | Predicted value |
| A | Actual value |
| Â | Mean actual value |
| Ṕ | Mean predicted value |

### PSO-ridge regression

The result of ridge regression analysis by gaussian lambda family and normalization of Ridge regression is illustrated in Fig 5c and Fig 5d. The output of the root mean square error[RMSE] value on Train data is 0.23 and Test data is 0.23 on 1113 degrees of freedom and its multiple R - squared value was obtained on Train data and Test data is 0.89.

### PSO-LASSO regression

Regression results using LASSO and the gaussian lambda distribution under normalization of linear regression are illustrated in Fig 5e and 5f. With 1113 degrees of freedom, RMSEs are 0.02 for the training dataset and 0.01 for the testing dataset with a multiple train/test R-squared value of 0.99.

### PSO-support vector regression

The result of sector vector regression analysis by gaussian lambda family and normalization of linear regression is illustrated in Fig 5g and Fig 5h. The output of the root mean square error [RMSE] value on Train data is 0.32 and test data is 0.32 on 1113 degrees of freedom and its multiple R-squared value was obtained on Train data is 0.92 and test data is 0.85.

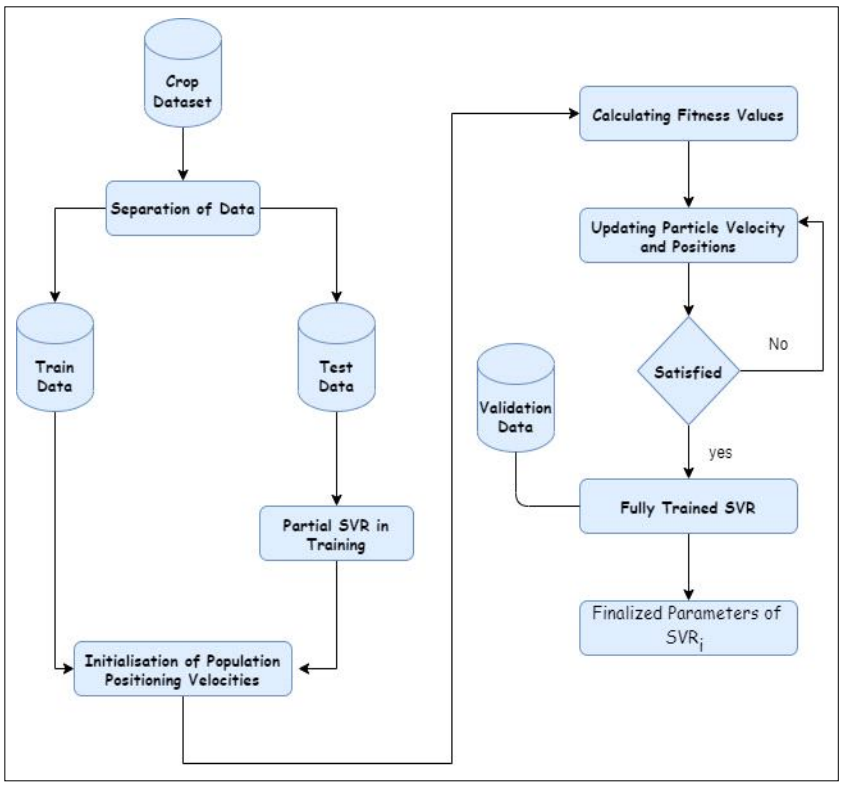



**Fig 4:** Proposed work for crop health and yield prediction analysis (From the crop dataset, test and train data is partially trained by Regressions to initialise position velocities, calculate fitness to train with SVR).

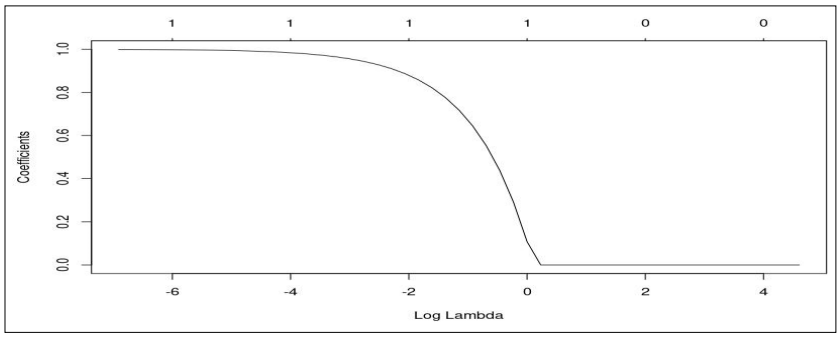



**Fig 5a:** Gaussian lambda family-linear regression.

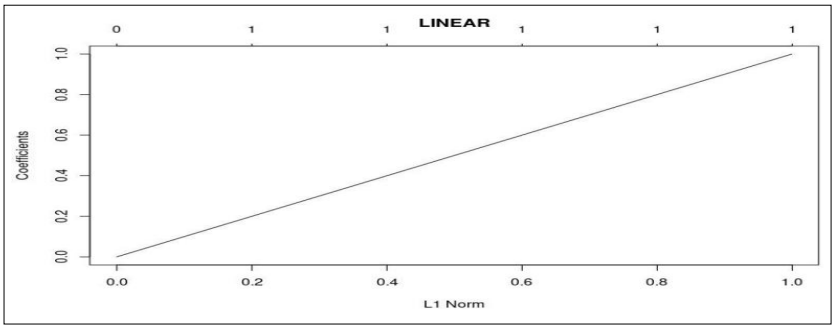



**Fig 5b:** Normalization of linear regression.

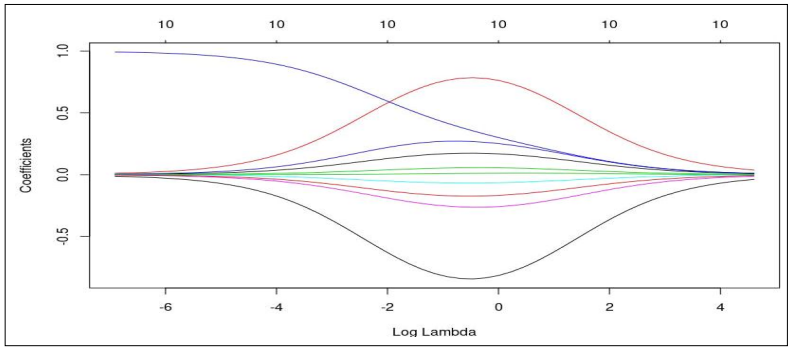



**Fig 5c:** Gaussian lambda family-ridge regression.

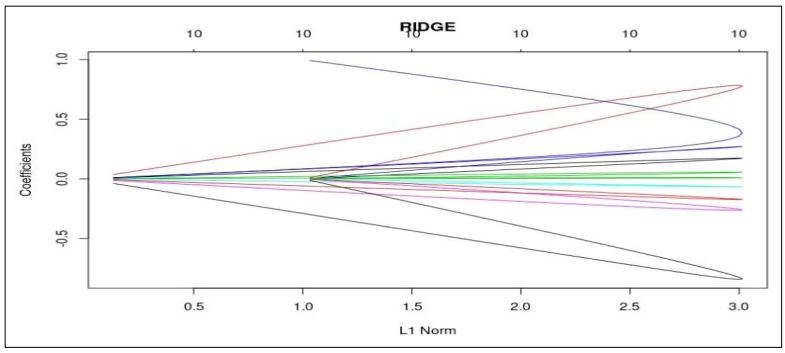



**Fig 5d:** Normalization of ridge regression.

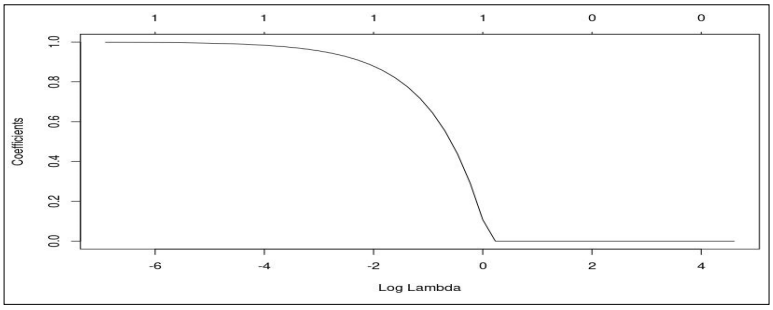



**Fig 5e:** Gaussian lambda family-PSO-LASSO regression.

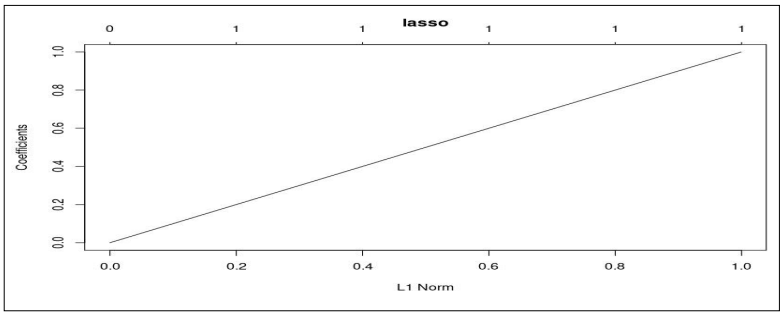



**Fig 5f:** Normalization of PSO-LASSO regression.

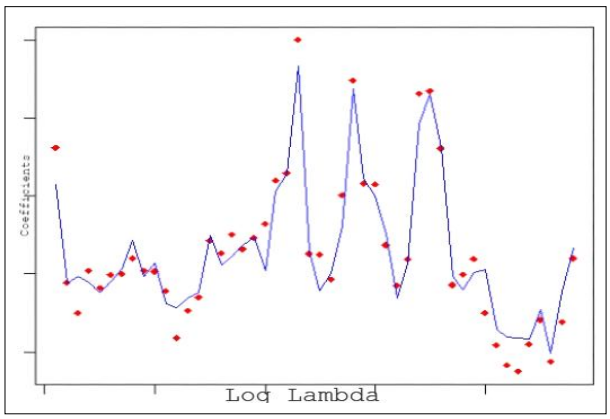



**Fig 5g:** Gaussian lambda family-support vector regression.

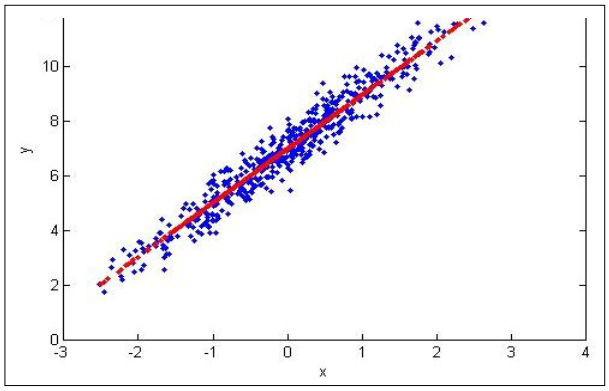



**Fig 5h:** Normalization of support vector regression.

Projecting high-dimensionality of data into low-dimensionality space always loses some initial information. The current challenge is to generate relevant minimisation data from the High-dimensionality data collection while preserving the original data's essential properties.

The comparative analysis of train data (Fig 5i) and test data (Fig 5j) of PSO-linear regression, PSO-ridge regression, PSO-LASSO regression and PSO-Support vector regression are all tabulated in Table 2 and 3. Calculating R-squared involves multiple stages. Finding the greatest fit line for independent and dependent variables, generally using a regression model. Then removes the actual values and squares the findings.

In terms of scalability, PSO-LASSO is employed on sample sizes ranging from 10,000 to 222,853. Training time increased linearly and achieved scalability. Penalty L1 avoided overfitting, ensuring sparse coefficients for features. Training/testing error RMSE of 0.02 and 0.01 respectively suggest good generalization capability. Inference latency of Raspberry Pi 4 takes less than 50 milliseconds, while RAG search requires 200-3.

R-squared is a relative measure of fit, RMSE is an absolute measure of fit. Lower values of RMSE indicates better fit. RMSE is a good measure of how accurately the model predicts the response. So, PSO-LASSO regression is best when compared with rest than three. Results obtained from a paired Wilcoxon signed-rank test using five cross validation sets indicated that PSO-LASSO performed better than PSO-Ridge, PSO-Linear and PSO-SVR in terms of RMSE values, $p<0.05$ (Fig 5a to 5j).

**Comparative experimental result**

Amiri *et al*. (2016) proposed a novel model to estimate blast-induced soil vibration using Artificial Neural Networks and K-Nearest Neighbours. ANN-KNN is used to shorten this combination of nearest neighbours. This study used an existing ANN and two empirical equations from the US Bureau of Mines (USBM) to assess the efficacy of the ANN-KNN model on ground movements. To estimate soil vibration (Faradonbeh and Monjezi, 2017) caused by removing the excess at the Gol-E-Gohar iron mine, a prediction algorithm genetic expression programmer (GEP) was constructed in the first phase of this study and the results were presented. When the GA algorithm was coupled with the SVR model, the findings of Behzadafshar *et al*. (2018) showed that was the most dominating evolutionary method for the present issue. The spatial

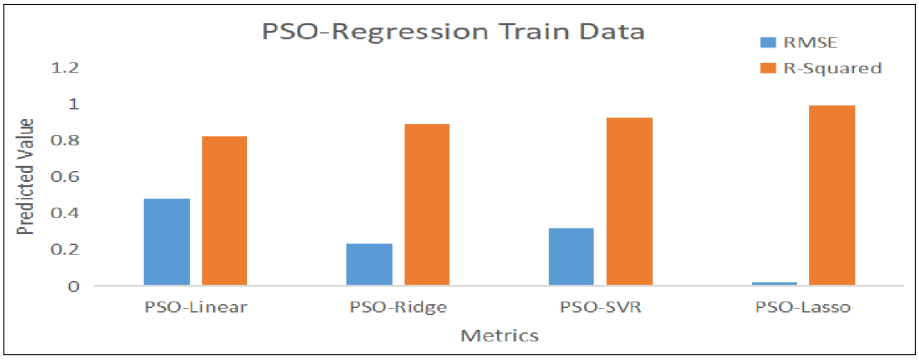



**Fig 5i:** Comparative results of train data.

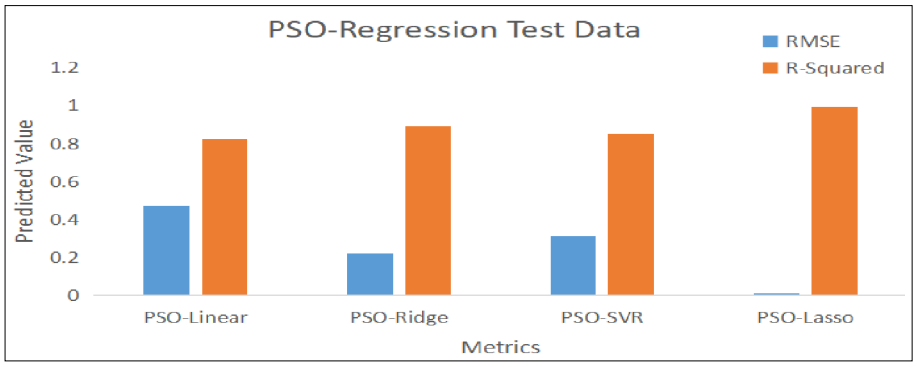



**Fig 5j:** Comparative results of test data.

**Table 2:** Results of PSO-regression analysis.

| Regression | Train data | | Test data | |
|---|---|---|---|---|
| | RMSE | R-squared | RMSE | R-squared |
| PSO-Linear | 0.48 | 0.82 | 0.47 | 0.82 |
| PSO-Ridge | 0.23 | 0.89 | 0.22 | 0.89 |
| PSO-Lasso | 0.02 | 0.99 | 0.01 | 0.99 |
| PSO-SVR | 0.32 | 0.92 | 0.31 | 0.85 |

**Table 3:** Comparative result analysis.

| Regression | Train data | | Test data | |
|---|---|---|---|---|
| | RMSE | R-squared | RMSE | R-squared |
| ANN-KNN (Amiri *et al*., 2016) | 0.54 | 0.88 | 0.53 | 0.88 |
| GEP-COA (Faradonbeh and Monjezi, 2017) | 6.73 | 0.87 | 6.71 | 0.87 |
| ICA-Linear (Behzadafshar *et al*., 2018) | 0.31 | 0.92 | 0.30 | 0.92 |
| Proposed work | 0.02 | 0.99 | 0.01 | 0.99 |

pattern of soil of carbon (SOC) was investigated using radial basis functions (RBF). The research area's SOC concentrations were highest. The RBF was determined to be the optimum linear transformation for the GA-SVR model and this was verified. A method for estimating PPV was suggested and it was found to be the GA-SVR-RBF model.

## CONCLUSION AND FUTURE ENHANCEMENT

The development of regression algorithms has significantly advanced intelligent agricultural decision-making, motivating the creation of a novel production forecasting framework in AgroAdvisor. The accuracy and efficiency tests demonstrate the efficacy and flexibility of the proposed PSO-Regression for predictive modeling. It was shown that biophysical variables were not influenced by crop health when the impact of the soil type was examined through the factorial experiment. To achieve this, the productivity modeling system depends on a specialized crop information system. A research region may be chosen and environmental and meteorological data can be input into the simulation. With regards to crop features, normal physiological and morphological development and stress variables, the monitoring system is site-specific due to simulation model validation requirements. In conclusion, AgroAdvisor uses regression based on PSO with RAG-based recommendation systems to deliver reliable, understandable and context-specific crop yield predictions. PSO-LASSO resulted in RMSE = 0.01 and $R^2$ = 0.99 on testing data and performed better than baselines. Further research will focus on satellite-based predictions, wider geographical scope and differential privacy.

## ACKNOWLEDGEMENT

The present study was supported by the Department of Computational Intelligence, SRM Institute of Science and Technology, Kattankulathur, Chennai.

### Disclaimers

The views and conclusions expressed in this article are solely those of the authors and do not necessarily represent the views of their affiliated institutions. The authors are responsible for the accuracy and completeness of the information provided, but do not accept any liability for any direct or indirect losses resulting from the use of this content.

### Informed consent

All experimental procedures were conducted in accordance with standard ethical guidelines and with the informed participation of all collaborating institutions. Data collection from soil samples and agricultural sites followed applicable institutional approvals.

### Conflict of interest

The authors declare that there are no conflicts of interest regarding the publication of this article. No funding or sponsorship influenced the design of the study, data collection, analysis, decision to publish, or preparation of the manuscript.

## REFERENCES

Amiri, M., Amnieh, H.B., Hasanipanah, M. and Khanli, L.M. (2016). A new combination of artificial neural network and K-nearest neighbors models to predict blast-induced ground vibration and air-overpressure. *Engineering Computations*. **32:** 631-644.

Ayub, U. and Moqurrab, S.A. (2018). Predicting Crop Diseases using Data Mining Approaches: Classification. In: 2018 1st International Conference on Power, Energy and Smart Grid (ICPESG). doi: 10.1109/ICPESG.2018.8384523.

Behzadafshar, K., Mohebbi, F., Soltani Tehrani, M., Hasanipanah, M. and Tabrizi, O. (2018). Predicting the ground vibration induced by mine blasting using imperialist competitive algorithm. *Engineering Computations.* **35:** 1774-1787.

Cho, O.H. (2024). Machine learning algorithms for early detection of legume crop disease. *Legume Research.* **47(3):** 463-469. doi: 10.18805/LRF-788.

Faradonbeh, R.S. and Monjezi, M. (2017). Prediction and minimization of blast-induced ground vibration using two robust meta-heuristic algorithms. *Engineering Computations*. **33:** 835-851.

Heydari, S.S. and Mountrakis, G. (2018). Effect of classifier selection, reference sample size, reference class distribution and scene heterogeneity in per-pixel classification accuracy using 26 landsat sites. *Remote Sensing of Environment.* **204:** 648-658. doi: 10.1016/j.rse.2017. 09.035.

Kalmani, H.V., Dharwadkar, V.N. and Thapa, V. (2025). Crop yield prediction using deep learning algorithm based on CNN-LSTM with attention layer and skip connection. *Indian Journal of Agricultural Research.* **59(8):** 1303-1311. doi: 10.18805/IJARe.A-6300.

Kamilaris, A., Gao, F., Prenafeta-Boldu, F.X. and Ali, M.I. (2016). Agri-IoT: A Semantic Framework for Internet of Things-Enabled Smart Farming Applications. In: 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT), pp. 442-447.

Kennedy, J. and Eberhart, R. (1995). Particle Swarm Optimization. In: Proceedings of ICNN'95-International Conference on Neural Networks, Perth, WA, Australia. **4:** 1942-1948. doi: 10.1109/ICNN.1995.488968.

Lee, W.B. and Kim, B.H. (2024). Machine learning innovations for precise plant disease detection: A review. *Legume Research.* **47(10):** 1633-1638. doi: 10.18805/LRF-799.

Meenakshi, M. and Naresh, R. (2022a). Prediction of Soil Moisture Root Zone Health in Artificial Neural Network. In: 4th International Conference on Recent Trends in Computer Science and Technology (ICRTCST). IEEE. doi: 10.1109/ICRTCST9782050.

Meenakshi, M. and Naresh, R. (2022b). Review on Health and Productivity Analysis in Soil Moisture Parameters. In: Expert Clouds and Applications. Lecture Notes in Networks and Systems, Springer.

Meenakshi, M. and Naresh, R. (2022c). Soil health analysis and fertilizer prediction for crop image identification by Inception-V3 and random forest. *Remote Sensing Applications: Society and Environment.* **28:** 100846. doi: 10.1016/j.rsase.2022.100846.

Meenakshi, M. and Naresh, R. (2023a). Machine learning based classifying polluted soil health and productivity analysis in Tamil Nadu delta area in water management system. *Soft Computing.* **28(Suppl 2):** 537-537. doi: 10.1007/s00500-023-08237-2.

Metagar, M.S. and Walikar, G.A. (2024). Machine learning models for plant disease prediction and detection: A review. *Agricultural Science Digest.* **44(4):** 591-602. doi: 10.18805/ag.D-5893.

Oerke, E.C. and Dehne, H.W. (2004). Safeguarding production-losses in major crops and the role of crop protection. *Crop Protection.* **23(4):** 275-285.

Pandey, Y., Khan, M.F., Vishal, Pandey, A.K. and Singh, S.P. (2023). Real-time soil monitoring with IoT enabled system for crop prediction. *Agricultural Science Digest.* doi: 10.18805/ag.D-5705.

Poutea, R., Collin, A. and Stoll, B. (2011). A Comparison of Machine Learning Algorithms for Classification of Tropical Ecosystems Observed by Multiple Sensors at Multiple Scales. In: International Geoscience and Remote Sensing Symposium, Vancouver, BC.

Rajak, R.K., Pawar, A., Pendke, M., Shinde, P., Rathod, S. and Devar, A. (2017). Crop recommendation system to maximize crop yield using machine learning technique. *International Research Journal of Engineering and Technology.* **4(12):** 1-4.

Sharma, A., Jain, A., Gupta, P. and Chowdary, V. (2021). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access.* **9:** 4843-4873. doi: 10.1109/access.2020.3048415.

Stafford, J.V. (2019). Precision Agriculture '19. Wageningen: Wageningen Academic Publishers.