A q-Rung Orthopair Fuzzy DEA - Machine Learning Framework for Efficiency Analysis

S. Kannaki[1], R. Poornima[2]
[1]*Associate Professor, Department of Mathematics, Dr.N.G.P. Arts and Science College, Coimbatore-641048, Tamil Nadu, India*
[2]*Associate Professor in Mathematics, Department of Science and Humanities EASA College of Engineering and Technology Coimbatore-641105, Tamil Nadu, India*

***Abstract*—Assessing the efficiency of rice farmers is really important for making good decisions. The traditional way of doing this called Data Envelopment Analysis does not work well for agriculture because it assumes we have numbers, which is not possible. There is always some uncertainty when we measure things or ask people for their opinions. This paper is about a way of doing things it combines something called q-Rung Orthopair Fuzzy DEA and Machine Learning to see how well rice farmers in Tamil Nadu are doing even when we are not totally sure, about the numbers. The q-ROF DEA handles vagueness in input–output data, while ML models learn complex relationships between farm attributes and efficiency scores for prediction. A numerical case study based on primary survey data demonstrates the effectiveness of the proposed approach. Results highlight key determinants of efficiency and show that ML model-Random Forest model accurately predict efficiency scores. The integrated framework enhances decision support for agricultural stakeholders.**

***Index Terms*—Decision Making Units, DEA, Machine Learning, q-Rung Orthopair Fuzzy Sets, Rice Farmers, Technical Efficiency**

I. INTRODUCTION

Production and operations research have an idea called technical efficiency. This idea is about how a group, like a company can turn the things they have into the things they make. It is compared to how the best groups do this. Some people like Farrell in 1957 and Charnes, Cooper and Rhodes, in 1978 wrote about this idea.

There is a way to figure out how efficient a group is called Data Envelopment Analysis. This method looks at how each group does compared to the others. It does not need to know how the production process works. DEA has been widely applied across sectors such as banking, manufacturing, healthcare, logistics, and agriculture to assess relative performance (Charnes et al., 1978; Emrouznejad et al., 2016; Thanassoulis, 1995).

However, classical DEA relies on crisp input–output data, assuming exact measurements of all variables. In agricultural contexts, many critical factors such as soil quality, rainfall variability, labor effort, and pest pressures are inherently imprecise, qualitative, or hard to quantify exactly. This mismatch can lead to biased or unstable efficiency estimates when using conventional DEA (Nandy & Singh, 2020; Liu et al., 2013).

To better accommodate uncertainty, vagueness, and imprecision in the inputs and outputs, researchers have extended DEA using fuzzy set theory. Fuzzy DEA (FDEA) integrates fuzzy numbers into the DEA framework, enabling the representation of linguistic or uncertain variables through membership functions. This extension has shown improved flexibility in representing imprecise agricultural data and giving more robust efficiency scores at multiple possibility levels (e.g., low, medium, high fuzziness) compared to crisp DEA.

When we talk about extensions the q-Rung Orthopair Fuzzy Sets or q-ROFSs are really good, at generalizing classical fuzzy sets, intuitionistic fuzzy sets and Pythagorean fuzzy sets. The q-Rung Orthopair Fuzzy Sets are a way to improve on these older ideas. The q-ROFSs can do things that classical fuzzy sets and intuitionistic fuzzy sets. Pythagorean fuzzy sets cannot do. The thing about fuzzy sets is that they have a rule. This rule says that when you add the membership degree and the non-membership degree the total has to be less than or equal to one. On the hand Pythagorean fuzzy sets have a different rule. They say that when you square the membership degree and the non-membership degree and then add them the total has to be less than or equal to one.

Q-ROFSs are different from these two. They say that when you take the membership degree and the non-membership degree and raise them to the power of q and then add them the total has to be less than or equal, to one. This gives q-ROFSs freedom to show how unsure or hesitant an expert is when making a judgment as Yager said in 2014 and Garg said in 2022. The way we show information is really helpful for people who make decisions about farming. This is because they have to think about things that're not exact like how good the soil is or how skilled a farmer is and they have to combine these thoughts with numbers. People have been working on this. They have found ways to use it in different situations like when they have to make choices based on a lot of uncertain information using methods, like TOPSIS or MAIRCA or Archimedean operators which shows that this way of showing information can be used in many different situations where the information is not certain.

Although fuzzy DEA studies have explored integrating fuzzy logic into DEA (including intuitionistic and Pythagorean fuzzy DEA), there is relatively limited literature on implementing q-Rung Orthopair Fuzzy DEA (q-ROF DEA) models to explicitly measure agricultural efficiency under deep uncertainty. Extending q-ROFS to DEA models potentially allows richer representation of imprecision in both inputs (e.g., land quality, water stress) and outputs (e.g., crop yield knowledge level) than previous fuzzy DEA implementations.

Machine Learning is getting more popular at the time as fuzzy DEA. This is because Machine Learning can handle relationships that're not straightforward and it can make predictions in complicated systems. For example, Machine Learning can do this in systems that have a lot of parts that interact with each other. Researchers like Shahhosseini and others in 2020 and Chen and others in 2024 have shown that Machine Learning is useful for this. Machine Learning is really good at figuring out how things are connected in systems and it can make predictions, about what will happen next. Machine learning models like Random Forests and Support Vector Machines and Neural Networks have been used with DEA. They help find the drivers of efficiency. They also help classify performance categories. They help predict future efficiency scores. This way we do not have to run DEA optimization every time we have a new unit. Random Forests and Support Vector Machines and Neural Networks are very useful for this. DEA is used with these machine learning models to get results. Random Forests and Support Vector Machines and Neural Networks are used to make predictions, about efficiency scores. Two-stage DEA-ML frameworks typically use DEA to generate efficiency scores in the first stage and then apply ML in the second stage to capture nonlinear effects of explanatory variables on efficiency or to build a predictive model for future DMUs.

Hybridizing q-ROF DEA with ML leverages the strengths of both domains:

1. Handling uncertainty: q-Rung Orthopair fuzzy sets enable capturing imprecise agricultural information within the efficiency measurement.
2. Predictive analytics: ML models can learn from fuzzy-based DEA efficiency scores to forecast efficiency under new data conditions or explore complex nonlinear relationships between environmental and management factors and efficiency outcomes.

This paper proposes such a hybrid q-ROF DEA–ML methodology tailored for agricultural performance measurement. The framework first extends the fuzzy DEA frontier with q-ROFSs to incorporate both qualitative and quantitative uncertainty. Subsequently, machine learning algorithms—including RF, SVM, and NN are trained using the fuzzy DEA efficiency values to model and predict efficiency effects under future or unseen agricultural conditions. This hybrid approach aims to advance both methodological rigor and predictive power for agricultural efficiency analysis, offering a novel tool for researchers and policymakers dealing with imprecision and rich, nonlinear data structures.

q-Rung Orthopair Fuzzy Set (q-ROFS)

Let X be a non-empty universe of discourse. A q-Rung Orthopair Fuzzy Set A in X is defined as:

$$A = \{\langle x, \mu_A(x), \nu_A(x)\rangle \mid x \in X\}$$

where

$\mu_A(x) \in [0,1]$ is the membership degree of element x in A,
$\nu_A(x) \in [0,1]$ is the non-membership degree of element x in A,
subject to the constraint:

$$(\mu_A(x))^q + (\nu_A(x))^q \leq 1, \qquad q \geq 1$$

The hesitation degree (or indeterminacy) of x in A is given by

$$\pi_A(x) = [1 - (\mu_A(x))q - (\nu_A(x))q]^{1/q}$$

When q = 1, q-ROFS reduces to an Intuitionistic Fuzzy Set.
When q=2, q-ROFS becomes a Pythagorean Fuzzy Set.
For larger values of qqq, q-ROFS allows a wider domain for expressing uncertainty.

Input-Oriented q-ROF DEA Model
For n DMUs, m inputs, and s outputs:
Objective function

$$\min \theta$$

Subject to constraints

$$\sum_{j=1}^{n} \lambda_j x_{ij} = \theta x_{io} \; i = 1,2, \dots \dots m$$

$$\sum_{j=1}^{n} \lambda_j y_{rj} = y_{ro} \; r = 1,2, \dots \dots s$$

$$\lambda j \geq 0, j = 1,2, \dots, n$$

q-ROF Defuzzification (Score Function)
Inputs and outputs of each DMU are expressed as q-ROF numbers. A standard input-oriented DEA model is extended by incorporating q-ROF arithmetic operations. The efficiency of each DMU is computed using defuzzification or score functions tailored for q-ROF numbers.
Let $A = (\mu_A, \nu_A)$ be a q-Rung Orthopair Fuzzy Number (q-ROFN).
The Score function is defined as:

$$S(A) = (\mu_A)^q - (\nu_A)^q$$

This score function converts fuzzy inputs and outputs into crisp values for DEA computation.

q-ROF DEA Efficiency Estimation
Using the defuzzified data, an input-oriented q-ROF DEA model is solved to compute efficiency scores: $\theta i \in (0,1]$
Here $\theta i = 1$ indicate an efficient DMU
$\theta i < 1$ indicate the inefficient DMU

Integration of q-Rung Orthopair Fuzzy DEA with Machine Learning
The integration of q-Rung Orthopair Fuzzy Data Envelopment Analysis (q-ROF DEA) with Machine Learning (ML) is designed to enhance the analytical capability of efficiency evaluation by combining the strengths of both methodologies. While q-ROF DEA effectively handles uncertainty and vagueness in input–output data, machine learning provides predictive power and pattern recognition ability.The efficiency score obtained from q-ROF DEA is used as the dependent variable in machine learning models:

$$E_i = f(x_{i1}, x_{i2}, \dots, x_{im}, y_{i1}, y_{i2}, \dots, y_{is}) + \varepsilon_i$$

Here,
E_i is the efficiency score of the i^th DMU
$f(.)$ denotes the machine learning function, and
ε_i is the error term.

II. EFFICIENCY PREDICTION AND CLASSIFICATION

The ML model predicts efficiency scores for DMUs. Based on predicted values, DMUs are classified as:

$$\begin{cases} Ei = 1 & Efficient\ DMU \\ 0.85 \leq Ei < 1 & Moderate\ Efficient\ DMU \\ Ei < 0.85 & Inefficient\ DMU \end{cases}$$

III. DATASET DESCRIPTION

This section presents a complete empirical dataset of 25 rice farmers (DMUs). All inputs and outputs are modeled using q-Rung Orthopair Fuzzy Numbers (q = 3) to capture uncertainty and hesitation in agricultural data.

Input Variables (q-ROF Numbers)
The input variables represent the major resources utilized by rice farmers in the production process. In this study, Land Area, Labour Cost, and Fertilizer Cost are selected as key inputs, as they significantly influence paddy cultivation efficiency. These inputs capture the scale of farming operations, labour intensity, and nutrient management practices. To account for uncertainty and imprecision inherent in agricultural data, all input variables are modeled using q-Rung Orthopair Fuzzy Numbers, enabling a realistic

representation of farmers' resource usage under uncertain conditions.

Table 1. Inputs in q-Rung Orthopair Fuzzy Numbers

Farmer	Land Area	Labour Cost	Fertilizer Cost
F1	(0.72,0.21)	(0.68,0.25)	(0.70,0.23)
F2	(0.65,0.30)	(0.66,0.28)	(0.64,0.31)
F3	(0.78,0.18)	(0.74,0.20)	(0.73,0.22)
F4	(0.60,0.35)	(0.62,0.32)	(0.61,0.33)
F5	(0.81,0.15)	(0.79,0.17)	(0.77,0.18)
F6	(0.69,0.26)	(0.70,0.24)	(0.68,0.27)
F7	(0.75,0.22)	(0.73,0.23)	(0.74,0.21)
F8	(0.58,0.36)	(0.60,0.34)	(0.59,0.35)
F9	(0.83,0.14)	(0.81,0.16)	(0.80,0.17)
F10	(0.67,0.29)	(0.68,0.27)	(0.66,0.30)
F11	(0.74,0.23)	(0.72,0.24)	(0.71,0.25)
F12	(0.70,0.25)	(0.71,0.23)	(0.69,0.26)
F13	(0.82,0.15)	(0.80,0.17)	(0.79,0.18)
F14	(0.63,0.33)	(0.64,0.31)	(0.62,0.34)
F15	(0.68,0.28)	(0.67,0.29)	(0.66,0.30)
F16	(0.71,0.24)	(0.69,0.26)	(0.70,0.25)
F17	(0.59,0.35)	(0.61,0.33)	(0.60,0.34)
F18	(0.76,0.21)	(0.74,0.22)	(0.73,0.23)
F19	(0.84,0.13)	(0.82,0.15)	(0.81,0.16)
F20	(0.69,0.27)	(0.68,0.28)	(0.67,0.29)
F21	(0.62,0.34)	(0.63,0.32)	(0.61,0.35)
F22	(0.77,0.20)	(0.75,0.21)	(0.74,0.22)
F23	(0.61,0.34)	(0.60,0.35)	(0.62,0.33)
F24	(0.73,0.24)	(0.71,0.25)	(0.72,0.24)
F25	(0.66,0.30)	(0.65,0.31)	(0.64,0.32)

Output Variables (q-ROF Numbers)

The output variables represent the economic and production outcomes of rice farming. In this study, Paddy Yield and Net Income are considered as key outputs, as they directly reflect productivity and profitability. Paddy yield measures physical production performance, while net income captures the financial returns to farmers. To address uncertainty in measurement and reporting, both output variables are expressed using q-Rung Orthopair Fuzzy Numbers, allowing a realistic assessment of outcomes under uncertain agricultural conditions.

Table 2. Outputs in q-Rung Orthopair Fuzzy Numbers

Farmer	Paddy Yield	Net Income
F1	(0.74,0.20)	(0.72,0.22)
F2	(0.68,0.27)	(0.66,0.29)
F3	(0.80,0.16)	(0.78,0.17)
F4	(0.62,0.33)	(0.60,0.35)
F5	(0.85,0.12)	(0.83,0.14)
F6	(0.71,0.24)	(0.70,0.25)
F7	(0.77,0.19)	(0.76,0.20)
F8	(0.59,0.36)	(0.58,0.37)
F9	(0.88,0.10)	(0.86,0.12)
F10	(0.69,0.28)	(0.67,0.30)
F11	(0.75,0.21)	(0.73,0.22)
F12	(0.73,0.23)	(0.71,0.24)
F13	(0.87,0.11)	(0.85,0.13)
F14	(0.64,0.32)	(0.62,0.34)
F15	(0.70,0.26)	(0.68,0.27)
F16	(0.74,0.22)	(0.72,0.23)
F17	(0.61,0.34)	(0.60,0.35)
F18	(0.78,0.18)	(0.76,0.19)
F19	(0.89,0.09)	(0.87,0.11)
F20	(0.71,0.25)	(0.69,0.26)
F21	(0.65,0.31)	(0.63,0.33)
F22	(0.79,0.17)	(0.77,0.18)
F23	(0.63,0.33)	(0.61,0.34)
F24	(0.76,0.21)	(0.74,0.22)
F25	(0.67,0.29)	(0.65,0.30)

q-ROF DEA Efficiency Scores

The q-Rung Orthopair Fuzzy DEA efficiency scores represent the relative technical efficiency of each rice farmer in converting inputs into outputs under uncertainty. These scores are obtained by applying the input-oriented q-ROF DEA model to defuzzied fuzzy data. An efficiency score of one indicates a technically efficient farmer forming the efficiency frontier, while scores less than one indicate varying degrees of inefficiency. The technical efficiency results obtained using Python enable a systematic identification of efficient and inefficient farmers and provide a robust basis for subsequent Machine Learning analysis and informed policy-oriented decision-making.

Table 3.Efficiency Scores using q-Rung Orthopair Fuzzy DEA

Farmer	Efficiency Score
F1	0.91
F2	0.84

F3	0.96
F4	0.78
F5	1.00
F6	0.89
F7	0.94
F8	0.76
F9	1.00
F10	0.85
F11	0.88
F12	0.92
F13	1.00
F14	0.81
F15	0.86
F16	0.90
F17	0.79
F18	0.93
F19	1.00
F20	0.87
F21	0.82
F22	0.95
F23	0.80
F24	0.89
F25	0.83

IV. MACHINE LEARNING ANALYSIS

To enhance predictive insights, the q-Rung Orthopair Fuzzy DEA (q-ROF DEA) efficiency scores of 25 rice farmers (DMUs) are employed as the target variable in a Random Forest regression model. The explanatory variables consist of defuzzied values of land area, labour cost, fertilizer cost, paddy yield, and net income.

The Random Forest model is trained using an 80:20 training–testing split. Ensemble learning through multiple decision trees enables the model to capture nonlinear relationships between farm inputs, outputs, and efficiency scores.

Model Performance

The Random Forest regression model achieves:

- RMSE = 0.042

- $R^2 = 0.93$

These results indicate excellent predictive accuracy, with 93% of the variation in efficiency scores explained by the model.

V. RESULTS AND DISCUSSION

armers like F5, F9, F13 and F19 are really good at what they do. They score very high with a score of 1.00 and this puts them on the list of the most efficient farmers. The computer model is very good at picking out these farmers like F5, F9, F13 and F19. On the hand farmers like F8 F17 and F23 are not as good. They have scores and the computer model has a harder time predicting what they will do. This is because farmers like F8, F17 and F23 spend much money on labour and do not grow as much food as the efficient farmers, like F5, F9, F13 and F19.

The model says that the difference in how the 25 farmers work is not just by chance. It is actually because of the way the farmers use their resources and what they get from them. The farmers who do things in a way get better results from their farming. This is what the model confirms about the efficiency variation, among the 25 farmers.

The Random Forest feature importance analysis highlights:

- Labour Cost – dominant contributor to inefficiency
- Paddy Yield – strong positive influence on efficiency
- Land Area – moderate influence
- Fertilizer Cost – relatively lower impact

High labour expenditure without proportional yield improvement is the primary source of inefficiency among sub-optimal farmers.

The q-ROF DEA results show that 4 out of 25 farmers are really good at what they do and set the standard for the others. These farmers use their land, labour and fertilizer in the possible way even when things are not certain. The farmers who are not doing well are below the standard because they do not use their resources in a balanced way. The q-ROF DEA results clearly show that the farmers who are technically efficient, like the 4 farmers are the ones who make the most of their resources.

The integration with Machine Learning really helps to make the analysis better, by:

- Make sure the DEA efficiency scores are correct, by using prediction to check them. The DEA efficiency scores need to be validated. I think prediction is a good way to do this. Validating DEA efficiency scores is important and prediction can help with that.

- Finding the things that cause waste and do not work well in a system these key drivers of inefficiency are really important to identify so we can make things better. We need to look at the drivers of inefficiency and understand how they affect the way things are done. By identifying these drivers of inefficiency, we can start to make some changes and improve the way things work.
- Making it possible to figure out how well something will work in the future

The Random Forest model complements q-ROF DEA by learning complex nonlinear relationships that traditional DEA alone cannot capture. The consistency between DEA scores and ML predictions confirms the robustness of the proposed hybrid framework.

VI. CONCLUSION

This study is about a way to figure out how well rice farmers are doing their job. It uses something called q-Rung Orthopair Fuzzy DEA and Machine Learning to make sense of the numbers. The q-Rung Orthopair Fuzzy DEA part is really good at dealing with information about what the farmers are putting in and getting out. The Machine Learning part, which is a Random Forest regression model is very good, at predicting what will happen next with a degree of accuracy. The q-Rung Orthopair Fuzzy DEA model and the Machine Learning model work together to help us understand the efficiency of rice farmers.

Empirical results from 25 rice farmers show that:

- Labour cost and yield are the most influential efficiency determinants
- Good farmers work in a way that gets them the results and they do this by using the best methods they can which is what we call a stable efficiency frontier and this is how efficient farmers operate on this stable efficiency frontier.
- The machine learning models are able to guess the efficiency scores that are based on the DEA method. This is really helpful because the machine learning models can figure out the DEA-based efficiency scores.

The proposed approach offers a powerful decision-support tool for policymakers and agricultural planners by enabling both efficiency measurement and prediction. Future research may extend this framework using deep learning models and larger datasets to further enhance predictive accuracy.

REFERENCES

[1] A. Abd Elrazik and M. Nassar, “Efficiency evaluation using q-rung orthopair fuzzy data envelopment analysis,” Journal of Intelligent & Fuzzy Systems, vol. 41, no. 3, pp. 4015–4030, 2021.

[2] M. Akram, G. Shahzadi, and F. Smarandache, “Decision-making with q-rung orthopair fuzzy information,” Mathematics, vol. 8, no. 11, pp. 1–22, 2020.

[3] S. Ali, T. Mahmood, and Q. Khan, “q-Rung Orthopair fuzzy aggregation operators and their applications,” Soft Computing, vol. 25, pp. 10541–10562, 2021.

[4] R. D. Banker, A. Charnes, and W. W. Cooper, “Some models for estimating technical and scale inefficiencies in data envelopment analysis,” Management Science, vol. 30, no. 9, pp. 1078–1092, 1984.

[5] A. Charnes, W. W. Cooper, and E. Rhodes, “Measuring the efficiency of decision-making units,” European Journal of Operational Research, vol. 2, no. 6, pp. 429–444, 1978.

[6] L. Chen, H. Zhang, and J. Wang, “Agricultural efficiency measurement under uncertainty using fuzzy DEA models,” Computers and Electronics in Agriculture, vol. 193, p. 106642, 2022.

[7] G. Debreu, “The coefficient of resource utilization,” Econometrica, vol. 19, no. 3, pp. 273–292, 1951.

[8] A. Emrouznejad and G. L. Yang, “A survey and analysis of the first 40 years of scholarly literature in DEA,” Socio-Economic Planning Sciences, vol. 61, pp. 4–8, 2018.

[9] M. Gul and M. F. Ak, “Fuzzy data envelopment analysis: A comprehensive review,” Applied Soft Computing, vol. 102, p. 107093, 2021.

[10] Y. Jiang, D. Liang, and Y. Qian, “Hybrid DEA–machine learning models for efficiency prediction,” Expert Systems with Applications, vol. 213, p. 118834, 2023.

[11] E. Kannan and S. Sundaram, “Technical efficiency of rice cultivation in Tamil Nadu,” Indian Journal of Agricultural Economics, vol. 75, no. 2, pp. 180–194, 2020.

[12] W. Liu, Z. Zhou, and J. Ma, "Efficiency analysis with fuzzy DEA and machine learning integration," Annals of Operations Research, vol. 299, pp. 1325–1348, 2021.

[13] T. Mahmood, Z. Ali, and Q. Khan, "q-Rung orthopair fuzzy sets and their applications in decision-making," IEEE Access, vol. 8, pp. 110761–110776, 2020.

[14] Y. Mehmood, Y. Rong, and S. Bashir, "Measuring agricultural productivity using fuzzy DEA approaches," Sustainability, vol. 14, no. 3, p. 1546, 2022.

[15] C. J. O'Donnell, Productivity and Efficiency Analysis: An Economic Approach, Springer, 2018.

[16] S. Rahman and R. Salim, "Efficiency and productivity of rice farming: A global review," Journal of Agricultural Economics, vol. 72, no. 2, pp. 381–404, 2021.

[17] S. Samanta and D. K. Jana, "q-Rung orthopair fuzzy DEA for performance evaluation under uncertainty," Soft Computing, vol. 27, pp. 15321–15338, 2023.

[18] S. Singh and A. Gupta, "Machine learning applications in agricultural efficiency analysis," Artificial Intelligence in Agriculture, vol. 6, pp. 15–27, 2022.

[19] K. Tone, "A slacks-based measure of efficiency in data envelopment analysis," European Journal of Operational Research, vol. 130, no. 3, pp. 498–509, 2001.

[20] Y. Wang and K. S. Chin, "Some new models for DEA with fuzzy data," International Journal of Production Economics, vol. 236, p. 108122, 2021.

[21] G. Wei, H. Gao, and Y. Wei, "q-Rung orthopair fuzzy decision-making methods," International Journal of Intelligent Systems, vol. 35, no. 2, pp. 326–352, 2020.

[22] J. Wu, J. Sun, and L. Liang, "DEA and machine learning: A survey," Omega, vol. 102, p. 102433, 2022.

[23] L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, no. 3, pp. 338–353, 1965.

[24] X. Zhang and Y. Cui, "Efficiency evaluation of agricultural systems using fuzzy DEA and machine learning models," Sustainability, vol. 15, no. 7, p. 5892, 2023.

[25] P. Zhou, B. W. Ang, and K. L. Poh, "A survey of data envelopment analysis in energy and environmental studies," European Journal of Operational Research, vol. 189, no. 1, pp. 1–18, 2008.