The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 2023

Volume 23, Pages 332-337

ICRETS 2023: International Conference on Research in Engineering, Technology and Science

A Comparative Analysis of Uncertainty Assessment for Annual Yield Prediction of Citrus Growth Using FIS and ANFIS Models

Filiz Alshanableh
Near East University

Abstract: Accurate prediction of citrus fruit yield is essential for effective agricultural planning, resource allocation, and decision-making. This study aims to compare the uncertainty analysis of developed Fuzzy Inference System (FIS) and Adaptive Neuro-Fuzzy Inference System (ANFIS) models in the context of predicting the annual yield of citrus growth. To achieve this, a comprehensive dataset comprising relevant features such as climate variables, soil conditions, and historical yield records is collected. FIS and ANFIS models were constructed using average temperature, average rainfall, and average relative humidity as input parameters and annual citrus yield as output parameters for the 1980–2019 harvesting season. Out of 40 historical data sets, 35 of them were used to train the models. The last five years were utilized for testing the proposed models. The proposed FIS and ANFIS models were found to be in close agreement with their actual counterparts, i.e. R2 values were 0.913 and 0.935 for FIS and ANFIS, respectively. To evaluate the uncertainty associated with the predictions of both models, a Monte Carlo simulation technique is employed. Preliminary results indicate that the FIS and ANFIS models exhibit promising performance in predicting the annual yield of citrus growth. However, a detailed comparison of uncertainty metrics suggests that the ANFIS model tends to provide more precise and reliable predictions, with narrower confidence intervals, than the FIS model. This could be attributed to the adaptive learning capabilities of ANFIS, allowing it to effectively capture complex nonlinear relationships between input variables and citrus yield.

Keywords: Uncertainty analysis, FIS, ANFIS, Citrus fruits yield

Introduction

The prediction of citrus fruit yield plays a crucial role in agricultural planning and decision-making processes in the Northern Cyprus economy. The citrus industry, including oranges, mandarins, clementine, grapefruit, and lemons, accounts for 40% of Northern Cyprus's total agricultural production of which Guzelyurt and Lefke districts in West Mesaria region produce almost 98 % of all citrus cultivation (TRNC, 2023). Agriculture is a significant sector of many economies, and accurate yield predictions help in economic planning. Governments and businesses can anticipate fluctuations in crop production and adjust their policies, investments, and trading strategies accordingly. Crop yield prediction allows for better management of agricultural resources. Farmers can make informed decisions about the amount of seeds, fertilizers, water, and other inputs to use, optimizing resource allocation and minimizing waste.

In recent years, fuzzy inference systems (FIS) and adaptive neuro-fuzzy inference systems (ANFIS) have emerged as powerful soft computing tools for yield prediction in various agricultural domains. In the work of Khoshnevisan et al. (2014), the ANFIS model yielded better than the artificial neural networks (ANN) model, because it implements fuzzy rules, to predict potato yield. ANFIS was also used earlier by Pankaj (2011) to predict wheat yields. Forecasting crop yields also frequently makes use of fuzzy time series. Narendra et al. (2012) forecasted wheat yield using the same methods as Sakin did in 2010, who generated fuzzy time series for the prediction of rice production. Al-Shanableh in 2022, as a part of the current study, used FIS to predict annual citrus yield grown in Northern Cyprus.

- Selection and peer-review under responsibility of the Organizing Committee of the Conference

This study aims to compare the uncertainty analysis associated with the annual yield prediction of citrus growth using FIS and ANFIS models. To achieve this, a comprehensive dataset comprising relevant features such as climate variables in terms of temperature, humidity, and precipitation, and historical yield records for the 1980–2019 harvesting season were collected. The FIS model is developed by defining linguistic rules based on expert knowledge and using fuzzy logic to approximate the relationships between input variables and citrus yield. On the other hand, the ANFIS model incorporates a learning algorithm to optimize fuzzy rules and adaptively adjust model parameters. By analyzing the variations in predictions across multiple Monte Carlo simulation runs, the uncertainties associated with each model were quantified.

Method

Data Collection

Historical data sets for the 1980–2019 harvesting season were gathered in order to anticipate citrus yield. The West Mesaria region's Guzelyurt area provided the majority of the data, which also included information on the overall amount of oranges collected and the weather during this time. Irrigation, weather, and soil types and conditions all have an impact on citrus tree productivity. In this study, weather conditions such as temperature, precipitation, and relative humidity were chosen as the input parameters. Table 1 lists the value range for input/output data sets within a 40-year time window.

Table 1. Input/output data sets for FIS and ANFIS modelling

	Data Description	**Range**
	Average temperature (^{o}C)	17.0 - 25.0
Inputs	Average rainfall (mm)	140.0 - 600.0
	Average relative humidity (%)	55.0 – 80.0
Output	Citrus yield (kg/donum)	6,000 – 17,000

Fuzzy Logic Modelling

FIS is a soft computing technique that mimics human reasoning, allowing for the representation of uncertainty and ambiguity in the data. FIS has been used successfully to solve different nonlinear engineering problems creating a relationship between input and output parameters (Al-Shanableh et al., 2017; Al-Shanableh & Evcil, 2022). FIS modeling consists of three basic stages; fuzzification stage, fuzzy inference system (FIS), and defuzzification stage (Al-Shanableh et al., 2020).

A total of 40 historical data sets which were covering the 1980-2019 period were provided to train and test Mamdani type FIS model using MATLAB R2015a (8.5.0.197613., Mathworks Inc., Natick, USA) Fuzzy Logic Designer. Data sets consisted of average temperature, average rainfall, the average relative humidity for inputs, and actual citrus fruit yield for output. The numerical values for both input and output variables were transformed into various linguistic variables such as low, mid, and high through fuzzification. Several membership functions (MFs) such as triangular, trapezoidal, and bell-shaped were then identified along with their parameters, including the base width for each parameter. The second stage, FIS, is the core part of fuzzy logic that comprises fuzzy IF-THEN rules that link fuzzy input variables to fuzzy output variables (Mamdani, 1976). In this section 35 data sets were used to create if-then rules between the fuzzified inputs and output. The last stage called defuzzification converts fuzzy output sets back into crisp values to obtain the final prediction. Citrus yields were predicted by providing input data sets to the rule reviewer section of MATLAB Fuzzy Logic Designer (Al-Shanableh, 2022). The predicted values were compared with the actual yield values and relevant evaluation metrics, such as mean squared error (MSE) and R-squared were calculated.

ANFIS Modelling

The Adaptive Neuro-Fuzzy Inference System (ANFIS) is a powerful hybrid model that combines the strengths of fuzzy logic and neural networks to tackle complex prediction tasks (Al-Shanableh et.al., 2019). ANFIS comprises five interconnected layers as shown in Figure 1, each serving a specific purpose in the prediction process: fuzzification layer, fuzzy inference-rule layer, product layer, defuzzification layer, and overall output summation layer.

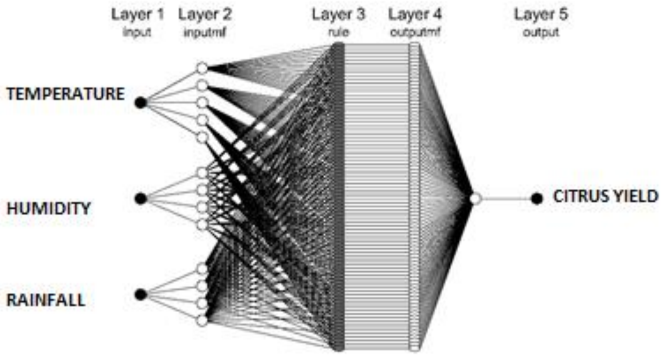

Figure 1. Structure of the ANFIS model used to predict the citrus yield

FIS model was extended by training an ANFIS model using the backpropagation algorithm with the same dataset. Prediction models for the ANFIS were generated using MATLAB R2015a (8.5.0.197613., Mathworks Inc., Natick, USA) Neuro-Fuzzy Designer. A significant number of simulations were conducted during the development of prediction models. In Neuro-Fuzzy Designer, the training and verification data sets are loaded into the workspace. Those crisp data sets were fuzzified into fuzzy linguistic variables using membership functions. Based on a predetermined number of MFs, a Sugeno-type FIS was generated using the grid partition method. The number of MF could be specified in association with each input. After FIS generated, the firing strength of each rule was divided by the sum of all rule firing strengths to obtain normalized values. The consequent parameters of the fuzzy rules were adjusted using a linear combination of the normalized firing strengths and multiplied with the function of Sugeno fuzzy rules. Next, defuzzification occurred by calculating the centroid area under MFs. A single node is used to compute the sum of all outputs in the last layer. In the end, the weighted average method results in a crisp result from the fuzzy result (Jang, 1993). The predicted values were compared with the actual yield values and relevant evaluation metrics, such as mean squared error (MSE) and R-squared were calculated.

Uncertainty Analysis of Predicted Models

The Monte Carlo simulation technique is a statistical method used to understand the uncertainty associated with a model's predictions. It involves creating a large number of random samples for each input variable, drawing from appropriate probability distributions. These random samples are then fed into the prediction model to generate a corresponding set of output values. By analyzing the variations in predictions Monte Carlo simulation were run and the uncertainties associated with each model were quantified by *d*-factor (Noori et.al., 2010). *d*-factor calculated as follow;

$$d - factor = \frac{\overline{d_x}}{\sigma_x}$$

where σ_x is the standard deviation of the measured variable X and $\bar{d}_x$ is degree of uncertainty or the average distance between the upper and lower 95 percent prediction uncertainties. $\bar{d}_x$ can be express as

$$\overline{d_x} = \frac{1}{k}\sum_{i=1}^{k}(X_U - X_L)$$

where *k* is the number of observed data points. Zero is the ideal d-factor value. The d-factor should have a value lower than 1 (Noori et.al., 2010). Larger uncertainty corresponds to a larger d-factor.

Results and Discussion

Proposed FIS Model

A fuzzy model was generated to predict the annual citrus fruit harvest in Northern Cyprus's Guzelyurt area. 35 historical data sets from the 1980–2014 time period were used to train the model, and 5 data sets from the 2015–

2019 time period were used to test it. The proposed FIS model was constructed by [6 4 5] trapezoidal MFs (trapmf) for input parameters which mean average temperature made up of 6 trapmf average rainfall had 4 and relative humidity had 5 trapmf. While the output layer that was citrus yield had 5 trapmf. The base width of each input/output parameter could be obtained from previous research by Al-Shanableh (2022). In FIS section 32 rules were created to train the fuzzy model and by using those rules lowest RMSE was obtained. Trained FIS was used to test for the data sets which were covered last 5 years. Citrus yields were predicted for the last 5 years and listed in Table 2. Because only training data sets were used to generate rules, the test phase results were impressive with a tighter prediction of real values.

Table 2. Predicted citrus yield for testing data sets

Year	Temp. (°C)	Rainfall (mm)	Relative Humidity (%)	Actual Citrus Yield (kg/donum)	Predicted Citrus Yield by FIS (kg/donum)	Predicted Citrus Yield by ANFIS (kg/donum)
2015	19.3	357.2	61.4	9,368	10,100	8,541
2016	19.9	282.9	58.5	7,819	8,060	6,160
2017	19.4	146.6	60.7	9,192	8,950	9,754
2018	20.1	405.8	57.2)	8,441	8,950	8,805
2019	19.7	465.4	62.6	10,736	11,000	11,876

Proposed ANFIS Model

Various input and output MF topologies with different types and numbers were used to simulate alternative models for annual citrus yield predictions. The goal was to achieve low RMSE values both in the training and verification phases. For each alternative model, the output was taken into account as constant MF, since linear MF produced high verification errors. The best conditions for both the training and verification phases were obtained for input MF using a generalized bell with grid partition FIS. Table 3 shows some alternative models generated during the trial stage. ANFIS-7 with the least error tolerance both in the training and verification phase was chosen as the predictive model. Citrus yields were predicted using generated ANFIS model for the last 5 years and listed in Table 2.

Table 3. Some simulation results for citrus yield prediction models

Model name	Type of input MF	Number of Input MF	Number of Rule	Epochs	RMSE Training	Verification
ANFIS-1	gbellmf	[2 2 2]	28	610	1.9854	2.1047
ANFIS-2	gaussmf	[3 3 3]	32	410	2.9873	2.8074
ANFIS-3	gauss2mf	[3 3 3]	32	780	2.3489	2.4573
ANFIS-4	gbellmf	[3 3 3]	32	660	1.7641	2.0814
ANFIS-5	gbellmf	[4 4 4]	44	780	1.5524	1.8932
ANFIS-6	gbellmf	[5 4 4]	50	900	1.3381	2.1205
ANFIS-7 (*)	gbellmf	[5 5 4]	62	980	0.7524	1.3932
ANFIS-8	gbellmf	[5 5 5]	68	1100	0.2381	2.1205

(*) Ideal model for prediction

The prediction capabilities of the proposed FIS and ANFIS models

Figure 2 shows the prediction capabilities of the proposed FIS model together with the proposed ANFIS model. Prediction error was found to be highest for 1987 with 2273 kg/donum. Both 1993 and 2002 have relatively high prediction errors of 1216 and 2130, respectively. The reason largely depended on data sets with different input parameters but the same output value, which led to high prediction errors. On the other hand, for the ANFIS Model, the largest prediction error was obtained for 1984 with 1470 kg/donum.

R^2 and RMSE of the proposed FIS model for citrus yield prediction are tabulated at Table 4 along with the ANFIS model comparison. Overall R^2 for the ANFIS model with 0.935 showed better capability compared to the FIS model that had 0.913. When predicting unseen last five years, it was found that developed FIS predicted citrus yield with better accuracy. This can be explained, by expert knowledge that can be translated into linguistic rules effectively in FIS Modelling. If domain experts can articulate rules and membership functions to represent uncertainties and relationships between variables clearly, FIS can capture this knowledge and provide interpretable results. ANFIS might not be able to provide similar transparency, as its structure involves neural network components, making it harder to extract explicit rules.

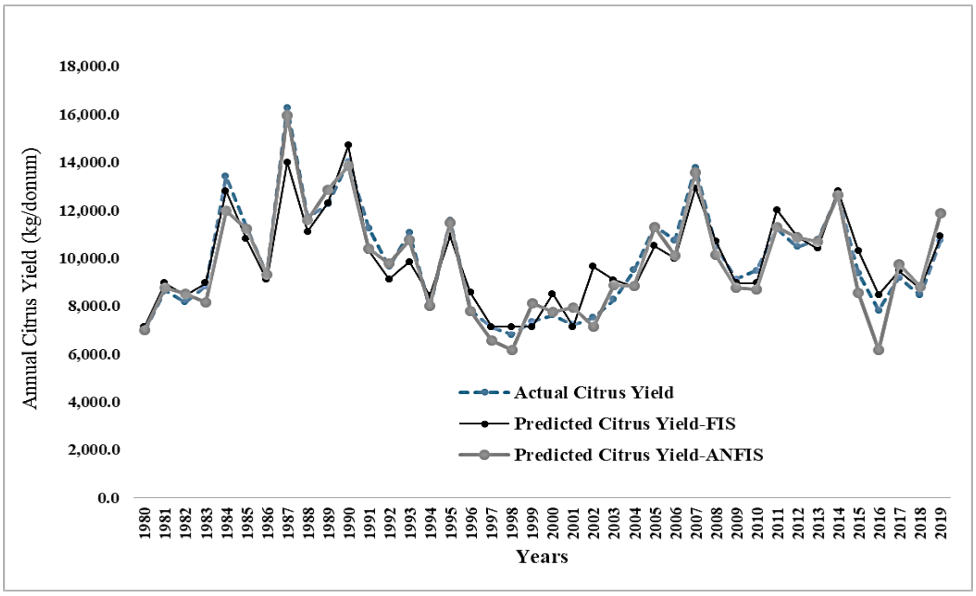

Figure 2. Comparison of citrus yield values of the proposed FIS and ANFIS models with actual counterparts

Table 4. Evaluation of statistical metrics for developed FIS and ANFIS models

Data Sets	FIS RMSE	R^2	ANFIS RMSE	R^2
Training	0.952	0.938	0.943	0.966
Verification	-	-	1.393	0.897
Testing	1.109	0.899	1.669	0.855
Overall	1.066	0.913	1.160	0.935

Uncertainty Analysis of Proposed Models

By measuring the confidence intervals of the simulation findings, uncertainty analysis of the expected annual citrus yield during the training and testing processes had been quantified. While the *d*-factor for the developed FIS Model was calculated as 0.72, for the ANFIS Model, it was found as 0.88. Both FIS and ANFIS can forecast the annual citrus yield with a d-factor lower than 1, which is acceptable. Uncertainty analysis revealed that the developed ANFIS Model forecasts yearly citrus yield more accurately than FIS Model.

Conclusion

This study investigates the prediction of citrus fruit yield using fuzzy inference system and adaptive neuro-fuzzy inference system. Input parameters used in FIS and ANFIS model construction were average temperature, average rainfall and average relative humidity for the 1980–2019 harvesting season. Out of 40 historical data set, 35 of them used to train FIS model. For ANFIS model construction, approximately one third of the training sets were used for verification of the model developed. The last five years were utilized for testing the proposed models and estimated citrus yield values by the proposed FIS and ANFIS models were found to be in close agreement with their actual counterparts, i.e. R^2 values were 0.913 and 0.935 for FIS and ANFIS, respectively. By employing the Monte Carlo simulation technique, the uncertainties associated with these models were compared. The findings shed light on the reliability, robustness, and predictive capabilities of each model in capturing the complexities and uncertainties present in citrus fruit production. This research contributes to the agricultural domain by providing insights into effective modeling techniques for yield prediction and aiding in decision-making processes for citrus growers and agricultural planners.

Further research may focus on incorporating additional environmental variables, exploring different membership functions and fuzzy rules, and applying advanced optimization techniques to improve the accuracy and reliability of yield predictions.

Scientific Ethics Declaration

The author declares that the scientific ethical and legal responsibility of this article published in EPSTEM journal belongs to the author.

Acknowledgements

* This article was presented as an oral presentation at the International Conference on Research in Engineering, Technology and Science (www.icrets.net) held in Budapest/Hungary on July 06-09, 2023.

References

Al-Shanableh, F. (2022). Prediction of the annual yield of citrus growth in the Guzelyurt district using fuzzy inference systems. In: C. Kahraman, A.C. Tolga, S. Cevik Onar, S. Cebi, B. Oztaysı, & I.U. Sarı (Eds.). *Intelligent and Fuzzy Systems*. Cham: Springer.

Al-Shanableh, F., Bilin, M., Evcil, A., & Savas, M.A. (2023). Estimation of cold flow properties of biodiesel using ANFIS-based models, *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 45*(2), 5440-5457.

Al-Shanableh, F., Bilin, M., Evcil, A., Savas, M.A. (2020). a study of jojoba oil extraction based on a fuzzy logic Model. *4th International Symposium on Multidisciplinary Studies and Innovative Technologies. ISMSIT 2020 - IEEE Conferences Proceedings*.

Al-Shanableh, F., Evcil, A., & Savas, M.A. (2017). Fuzzy logic model for prediction of cold filter plugging point of biodiesel from various feedstock. *Procedia Computer Sciences, 120*, 245–252.

Al-Shanableh, F., & Evcil, A., (2022). Prediction of energy consumption of residential buildings in Northern Cyprus using fuzzy interference system. *Energy and Buildings, 256,* 111555.

Jang, J.R. (1993). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on Systems, Man, and Cybernetics, 23*, 665-685.

Khoshnevisan, B., Rafiee, S., Omid, M., & Mousazadeh, H. (2014). Prediction of potato yield based on energy inputs using multi-layer adaptive neuro-fuzzy inference system. *Measurement, 47,* 521–530.

Mamdani, E.H. (1976). Advances in the linguistic synthesis of fuzzy controllers. *International Journal of Man-Machine Studies, 8*(6), 669–678.

Narendra, K., Ahuja, S., Kumar, V., & Kumar, A. (2010). Fuzzy time series forecasting of wheat production. *International Journal of Computer Science and Engineering. 2*, 635–640.

Noori, R., Hoshyaripour, G., Ashrafi, K., & Araabi, B.N. (2010). Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration. *Atmospheric Environment, 44,* 476-482.

Pankaj, K. (2011). Crop yield forecasting by adaptive neuro fuzzy inference system. *Mathematical Theory and Modeling, 11*(1),1–7.

Sakin, K., & Kumar, N. (2012). A novel method for rice production forecasting using fuzzy time series. *International Journal of Computer Science Issues, 9,* 455–459.

TRNC. (n.d.). *Agriculture and natural resources ministry homepage*. http://tarim.gov.ct.tr/tr-tr/istatistik.aspx

Author Information

Filiz Al-Shanableh
Near East University
Nicosia, Cyprus
Contact e-mail: *filiz.shanableh@neu.edu.tr*

To cite this article:

Alshanableh, F. (2023). A comparative analysis of uncertainty assessment for annual yield prediction of citrus growth using FIS and ANFIS models. *The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 23*, 332-337