International Journal of Knowledge and Innovation Studies
https://www.acadlore.com/journals/IJKIS

Advanced Estimation of Orange Tree Age Using Fuzzy Inference and Linear Regression Models

Muhammad Shahkar Khan

Department of Mathematics, CECOS University of IT and Emerging Sciences, 25000 Peshawar, Pakistan

* Correspondence: Muhammad Shahkar Khan (shahkar@uop.edu.pk)

Received: 05-21-2024 **Revised:** 06-29-2024 **Accepted:** 07-03-2024

Citation: M. S. Khan, "Advanced estimation of orange tree age using fuzzy inference and linear regression models," *Int J. Knowl. Innov Stud.*, vol. 2, no. 3, pp. 119–129, 2024. https://doi.org/10.56578/ijkis020301.

Abstract: The accurate estimation of the age of orange trees is a critical task in orchard management, providing valuable insights into tree growth, yield prediction, and the implementation of optimal agricultural practices. Traditional methods, such as counting growth rings, while precise, are often labor-intensive and invasive, requiring tree cutting or core sampling. These techniques are impractical for large-scale application, as they are time-consuming and may cause damage to the trees. A novel non-invasive system based on fuzzy logic, combined with linear regression analysis, has been developed to estimate the age of orange trees using easily measurable parameters, including trunk diameter and height. The fuzzy inference system (FIS) offers an adaptive, intuitive, and accurate model for age estimation by incorporating these key variables. Furthermore, a multiple linear regression analysis was performed, revealing a statistically significant correlation between the predictor variables (trunk diameter and height) and tree age. The regression coefficients for diameter ($p = 0.0134$) and height ($p = 0.0444$) demonstrated strong relationships with tree age, and an R-squared value of 0.9800 indicated a high degree of model fit. These results validate the effectiveness of the proposed system, highlighting the potential of combining fuzzy logic and regression techniques to achieve precise and scalable age estimation. The model provides a valuable tool for orchard managers, agronomists, and environmental scientists, offering an efficient method for monitoring tree health, optimizing fruit production, and promoting sustainable agricultural practices.

Keywords: Fuzzy logic; Non-invasive methods; Membership functions; Orange trees; Forest management

1 Introduction

Forest ecosystems are among the most important terrestrial habitats, contributing significantly to environmental sustainability and providing a wide range of ecological, economic, and social benefits. Forests play a critical role in carbon sequestration, nutrient cycling, climate regulation, and biodiversity conservation. They act as carbon sinks, helping to mitigate the impact of climate change by absorbing carbon dioxide from the atmosphere. In addition, forests support a vast array of plant and animal species, making them one of the most biodiverse ecosystems on the planet. Understanding the age distribution of trees within a forest is essential for effective forest management and ecological research. Tree age can provide valuable insights into forest growth dynamics, regeneration patterns, and ecosystem resilience, allowing for more informed decisions in conservation efforts and sustainable forest management [1–4].

Accurately determining the age of trees has traditionally been a labor-intensive and invasive process. The most common method, dendrochronology, involves counting the annual growth rings of a tree, which requires either cutting down the tree or extracting core samples using increment borers. While this method provides a reliable estimate of tree age, it is not always feasible, especially in protected or valuable forests where minimizing disturbance is a priority [5]. Furthermore, dendrochronology is not applicable in cases where the tree is hollow or has decayed inner rings. As a result, there is a growing need for non-invasive, efficient, and scalable methods to estimate tree age, particularly in large-scale forest surveys.

In recent years, advances in computational techniques have opened new possibilities for developing non-invasive methods to estimate tree age based on external measurements [6–9]. Machine learning, remote sensing, and fuzzy logic have emerged as promising tools for modeling complex biological processes and predicting various forest attributes, including tree age [10–13]. Fuzzy logic, in particular, is well-suited for biological systems, as it can handle the inherent uncertainty and imprecision associated with natural processes. Unlike traditional binary logic, which relies on precise true or false conditions, fuzzy logic allows for degrees of truth, making it ideal for situations where inputs are vague or overlapping. This characteristic makes fuzzy logic a powerful tool for modeling tree growth, which is influenced by multiple factors, such as environmental conditions, species characteristics, and competition for resources. In many species, as the tree ages, both the diameter of the trunk and the height increase in a predictable manner, although the exact relationship varies depending on species, site conditions, and forest density [14]. By leveraging this relationship, the fuzzy logic model can estimate tree age without requiring direct access to the tree's growth rings.

Several studies have demonstrated the potential of fuzzy logic for modeling tree growth and estimating forest attributes. For instance, Khan [15] highlights the use of fuzzy set theory in environmental engineering to manage uncertain and imprecise data, such as pollutant levels and weather conditions. The model shows how fuzzy logic enhances environmental decision-making in areas like air and water quality assessment, waste management, and ecological risk. By handling ambiguities effectively, fuzzy set theory supports more adaptable and robust environmental monitoring and impact assessments. Similarly, Bone et al. [1] develop a model to simulate forest insect infestations using a combination of fuzzy logic and cellular automata (CA). This model was designed to handle the spatial complexity of forest ecosystems, where insect infestations spread in irregular patterns influenced by environmental factors. The fuzzy logic component helps in incorporating uncertain and variable factors like insect population density, tree susceptibility, and environmental conditions into the model. These fuzzy constraints allow the CA to apply different rules based on the degree of infestation risk in neighboring cells, making the model adaptive to real-world forest conditions.

The primary objective of this study is to develop a predictive model that utilizes measurable parameters such as trunk diameter and tree height to accurately estimate the age of orange trees without invasive techniques. This study aims to answer the following research question: *Can a combination of fuzzy inference and linear regression analysis significantly improve the accuracy of non-invasive age estimation for orange trees?* This question is central to our investigation, as it seeks to determine if advanced computational techniques can provide a viable alternative to traditional age estimation methods.

In pursuit of this goal, we propose a novel age estimation model that integrates fuzzy logic and linear regression analysis. The specific objectives of this research are as follows:

• **Develop a Fuzzy Inference System (FIS):** This system will use trunk diameter and height as input variables to estimate the age of orange trees. Fuzzy logic is chosen due to its ability to handle uncertainties and variability in biological data, capturing the inherent imprecision in growth processes. The fuzzy logic system developed in this study is designed to mimic human reasoning by using linguistic variables to describe the relationships between tree diameter, height, and age. For example, the model might define fuzzy sets such as "Small", "Medium" and "Large" for diameter, and "Young", "Middle-aged" and "Old" for age. These sets are defined by overlapping membership functions that allow for gradual transitions between categories, reflecting the continuous nature of tree growth. The model uses a set of if-then rules to infer the tree's age based on its diameter and height. For example, one rule might state, "If the diameter is Medium and the height is Tall, then the tree is Middle-aged." The output of the model is a fuzzy set representing the estimated age, which is then defuzzified to obtain a crisp value [16].

• **Incorporate Linear Regression Analysis for Robustness:** Linear regression will be applied to validate the results obtained from the FIS and to quantify the strength of the relationship between diameter, height, and age. This dual approach aims to ensure both adaptability and precision in age estimation.

• **Test Model Performance with Statistical Rigor:** The model will be evaluated based on p-values, correlation coefficients, and R-squared values to assess its predictive accuracy. A significant p-value $(p < 0.05)$ will indicate that the relationship between inputs (diameter, height) and output (age) is statistically meaningful, while a high R-squared value will demonstrate that a considerable proportion of age variability is explained by the model.

The accuracy of the fuzzy logic model depends on the quality of the input data and the calibration of the membership functions. Factors such as species variation, environmental conditions, and forest management practices can affect the relationship between tree diameter, height, and age. Therefore, it is important to calibrate the model using species-specific data and local growth conditions [17]. Additionally, integrating other non-invasive technologies, such as remote sensing or LiDAR, with the fuzzy logic model could further improve the accuracy of age estimates by providing more detailed information on tree structure and growth patterns.

In conclusion, this research seeks to bridge the gap between traditional, invasive methods of age estimation and modern, non-invasive techniques. The integration of fuzzy inference and linear regression provides a powerful tool for orchard managers, agronomists, and environmental scientists, offering a more practical, accurate, and sustainable method for assessing the age of orange trees. Through this study, we aim to contribute a reliable model that not only advances agricultural practices but also enhances our understanding of orange tree growth dynamics.

1.1 Motivation and Significance

Accurate age prediction is fundamental to various forestry applications. For instance, determining the age distribution within a forest allows forest managers to plan logging operations while ensuring sustainable timber production [1]. Bayat et al. [18] compares artificial neural networks (ANN) and multiple linear regression (MLR) for predicting the ten-year volume growth of Oriental beech, highlighting ANN's accuracy and potential for broader forestry applications. The findings support data-driven, sustainable forest management and advanced growth prediction tools. Moreover, age contributes to understanding the carbon sequestration capabilities of trees, as older trees tend to store more carbon than younger ones [14]. Therefore, there is a growing need for methodologies that are both accurate and scalable across different forest types.

The traditional methods for determining tree age, while accurate, are largely impractical for large-scale surveys. Counting growth rings, for example, is highly accurate for individual trees but impossible to implement without invasive methods and considerable manual effort. Other techniques, such as the use of tree height or diameter-growth models, offer some improvement but can suffer from inaccuracies, particularly when applied across diverse tree species and environmental conditions [14]. This variability and imprecision underline the need for models that can handle the uncertainties of natural growth processes.

The fuzzy logic approach addresses these challenges by allowing for gradual transitions between different growth stages. Instead of relying on sharp cut-offs, fuzzy systems use overlapping categories to represent tree sizes and ages, reflecting the natural variability seen in biological growth processes [16]. This makes it possible to produce more flexible and accurate age estimates, especially in cases where environmental or genetic factors cause deviations from standard growth models.

2 Materials and Methodology Overview

2.1 Study Area

This study was conducted in Orange Field Qajeer, 7867 + 5M6, Machai, Mardan District, Khyber Pakhtunkhwa 23200, North-East Pakistan, Pakistan. Orange trees generally require a warm climate for optimal growth. Higher temperatures can enhance metabolic rates, leading to increased growth rates in diameter and height during the growing season. So, the annual temperature record for Mardan, located in Khyber Pakhtunkhwa, Pakistan, varies significantly throughout the year. Here is a summary of the average temperatures by month in Table 1.

Table 1. Average monthly temperatures

Month	Average High (°C)	Average Low (°C)
January	17.7	2.3
February	19.0	5.5
March	24.0	10.4
April	30.1	15.3
May	36.3	20.2
June	41.4	25.1
July	38.5	26.2
August	36.5	25.5
September	35.3	22.3
October	31.6	14.9
November	25.1	7.4
December	19.4	2.7

In this study, we developed a hybrid approach to estimate the age of orange trees using a combination of fuzzy logic and linear regression analysis. This section provides detailed specifications of the fuzzy sets, membership functions, and regression model setup, ensuring reproducibility and transparency in the research process.

2.2 Fuzzy Logic System

The fuzzy logic system was implemented to account for the inherent variability in biological measurements, such as trunk diameter and height, which are affected by environmental and genetic factors. We utilized a Mamdani-type FIS with the following key components:

• **Fuzzy Sets and Membership Functions:** The fuzzy sets for each input variable (trunk diameter and height) were defined based on expert knowledge and historical growth data of orange trees. For example, trunk diameter was divided into sets labeled as "Small", "Medium" and "Large", while height was categorized as "Low", "Moderate" and "High". Triangular and trapezoidal membership functions were used to model these sets, as they offer a balance between computational simplicity and accuracy in reflecting growth variability.

• **Derivation of Membership Functions:** The parameters of the membership functions (e.g., the base and peak points for the triangular and trapezoidal shapes) were derived using statistical analysis on a sample dataset. The base points were set to encompass the observed range for each category, and the peak points were aligned with the median values within each set to represent the most typical values.

• **Rule Base Construction:** A comprehensive rule base was constructed to capture the complex relationships between tree diameter, height, and age. Rules were designed in the form of "IF-THEN" statements. For instance:

IF Diameter is Large AND Height is High, THEN Age is likely High.

A total of 27 rules were established, covering various combinations of the input fuzzy sets, allowing the system to provide age estimates for different growth profiles.

• **Defuzzification:** The centroid method was used for defuzzification, which calculates the center of gravity of the aggregated fuzzy output. This method provided a crisp age value as the output, suitable for further analysis and comparison.

2.3 Linear Regression Model

To complement the fuzzy logic model, a linear regression model was applied to quantify the relationship between the input variables (diameter and height) and the target variable (age).

• **Model Specification:** The regression model was set up with age as the dependent variable, while trunk diameter and height served as independent variables. We included interaction terms to account for potential synergies between diameter and height in influencing age. No transformations were applied to the variables, as preliminary analysis indicated a linear relationship.

• **Coefficient Estimation and Statistical Significance:** Ordinary Least Squares (OLS) was used to estimate the regression coefficients. Each coefficient's p-value was computed to assess the significance of the predictors. Only predictors with p-values below 0.05 were retained in the model to ensure robustness.

• **Model Evaluation:** The model's goodness of fit was evaluated using the R-squared value, indicating the proportion of variance in age explained by the predictors. Additionally, residual analysis was performed to check for any violations of the linear regression assumptions (e.g., normality, homoscedasticity).

2.4 Combined Model for Age Estimation

The final age estimation was obtained by integrating the fuzzy logic and regression outputs. The fuzzy model provided an initial age range, while the regression model refined the estimate, improving precision. This hybrid approach allows for a more flexible and accurate assessment compared to traditional, singular models.

2.5 Objectives of the Study

The primary objective of this research is to develop a non-invasive, accurate, and scalable method for estimating the age of trees using fuzzy logic. This system aims to improve the precision of age predictions for a variety of tree species, using easily measurable parameters like diameter and height. We also aim to create a flexible model that can adapt to different environmental conditions and tree species by modifying the membership functions and inference rules accordingly.

In addition to enhancing the accuracy of age estimation, this method could significantly reduce the effort required for large-scale forest surveys. By using non-invasive measurements, the fuzzy logic system can be applied to forests without causing harm to the trees, making it a practical tool for long-term forest monitoring and management. We also hope to demonstrate that this model can be used as a foundation for other ecological applications, such as estimating tree health or growth potential.

3 Literature Review and Related Work

Several previous studies have explored the relationship between tree size and age, often focusing on specific species or environmental conditions. The use of growth models based on diameter and height is well established in forestry [1]. However, many of these models are deterministic, meaning they do not account for inherent variability in tree growth caused by genetic and environmental factors. This limitation has motivated the exploration of alternative approaches, such as machine learning and fuzzy logic, that can handle uncertainty [18].

Recent studies have highlighted the potential of machine learning techniques, such as random forests and neural networks, for predicting tree age based on multiple environmental and physiological parameters [7]. These approaches can improve the accuracy of age estimates, particularly when applied across diverse forest ecosystems, but they often require extensive datasets for training and are less interpretable than rule-based models. In contrast, fuzzy logic offers a more transparent and flexible framework for incorporating expert knowledge and handling the inherent uncertainty in tree growth [16].

Fuzzy logic has been applied in various ecological modeling tasks, including predicting growth stages and habitat suitability [15]. However, its application to tree age estimation remains relatively unexplored. Some recent studies have begun to integrate fuzzy systems with remote sensing data, allowing for large-scale monitoring of forest health and biomass [6]. By combining the scalability of remote sensing technologies with the flexibility of fuzzy logic, these approaches offer promising avenues for non-invasive forest management [19].

This study builds on these advancements by applying fuzzy logic to tree age estimation. Unlike deterministic models, the fuzzy approach allows for gradual transitions between different growth stages, reflecting the variability of biological processes. Moreover, by incorporating both diameter and height measurements, the fuzzy model can accommodate the non-linear growth patterns seen in different tree species [14]. This research contributes to the growing body of literature on non-invasive, scalable forest monitoring techniques, which are increasingly critical for sustainable forestry practices.

4 Proposed Mathematical Approach

To predict the orange tree age, we utilize the FIS, which shows the relationship between two input variables: Diameter (D) and Height (H), and an output variable: Age (A), as shown in Figure 1. The FIS estimates the age of orange trees using the measurable parameters of diameter (D) and height (H). The output age A is a function of D and H:

$$A = f(D, H)$$

where, $f(D, H)$ is the output after fuzzification, rule evaluation, aggregation, and defuzzification.

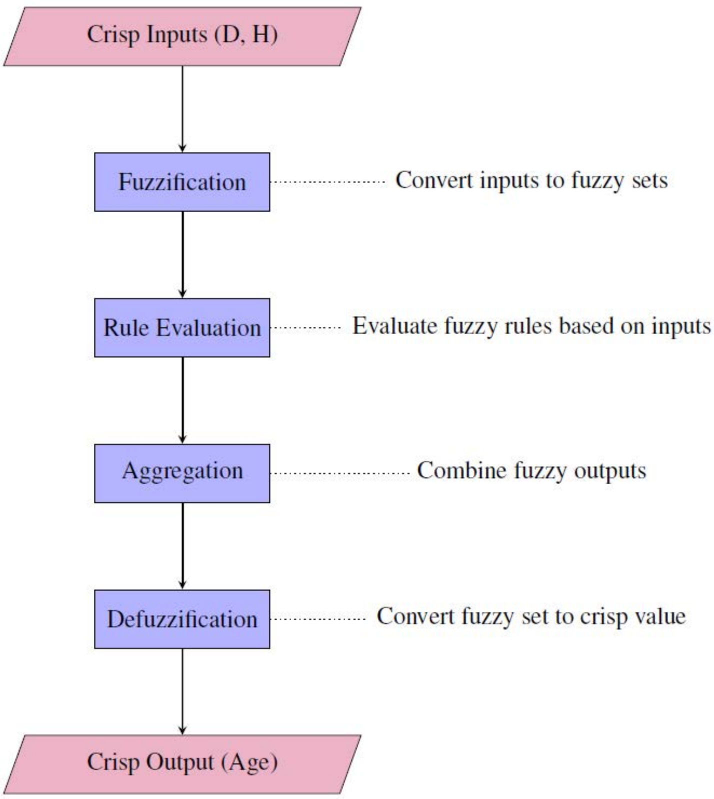

Figure 1. Advanced propose fuzzy inference process for estimating tree age

4.1 Membership Functions

In fuzzy logic, a membership function μ defines how each input value maps to a degree of membership between 0 and 1. This degree of membership represents how strongly the input value belongs to a particular fuzzy set. The membership functions are defined as follows:

Input membership functions for diameter

$$\mu_{Small}(D) = \begin{cases} 1 & \text{if } D \leq 10 \\ \frac{20-D}{10} & \text{if } 10 < D < 20 \\ 0 & \text{if } D \geq 20 \end{cases}$$

$$\mu_{Medium}(D) = \begin{cases} 0 & \text{if } D \leq 15 \\ \frac{D-15}{20} & \text{if } 15 < D < 35 \\ \frac{55-D}{20} & \text{if } 35 < D < 55 \\ 0 & \text{if } D \geq 55 \end{cases}$$

$$\mu_{\text{Large}}(D) = \begin{cases} 0 & \text{if } D \leq 50 \\ \frac{D-50}{20} & \text{if } 50 < D < 70 \\ 1 & \text{if } 70 \leq D < 100 \\ \frac{100-D}{30} & \text{if } D \geq 100 \end{cases}$$

Input membership functions for height

$$\mu_{Very\ Short}(H) = \begin{cases} 1 & \text{if } H \leq 1 \\ H-1 & \text{if } 1 < H < 2 \\ 0 & \text{if } H \geq 2 \end{cases}$$

$$\mu_{Short}(H) = \begin{cases} 0 & \text{if } H \leq 2 \\ H-2 & \text{if } 2 < H < 3 \\ 1 & \text{if } 3 \leq H < 4 \\ 5-H & \text{if } 4 < H < 5 \\ 0 & \text{if } H \geq 5 \end{cases}$$

$$\mu_{Medium}(H) = \begin{cases} 0 & \text{if } H \leq 4 \\ \frac{H-4}{2} & \text{if } 4 < H < 6 \\ \frac{8-H}{2} & \text{if } 6 < H < 8 \\ 0 & \text{if } H \geq 8 \end{cases}$$

$$\mu_{Tall}(H) = \begin{cases} 0 & \text{if } H \leq 6 \\ \frac{H-6}{2} & \text{if } 6 < H < 8 \\ 1 & \text{if } 8 \leq H < 10 \\ 0 & \text{if } H \geq 10 \end{cases}$$

Output membership functions for age

$$\mu_{Very\ Young}(A) = \begin{cases} 1 & \text{if } A \leq 5 \\ \frac{A-5}{5} & \text{if } 5 < A < 10 \\ 0 & \text{if } A \geq 10 \end{cases}$$

$$\mu_{Young}(A) = \begin{cases} 0 & \text{if } A \leq 10 \\ \frac{A-10}{2} & \text{if } 10 < A < 12 \\ 1 & \text{if } 12 \leq A < 15 \\ \frac{20-A}{5} & \text{if } 15 < A < 20 \\ 0 & \text{if } A \geq 20 \end{cases}$$

$$\mu_{Middle\ Aged}(A) = \begin{cases} 0 & \text{if } A \leq 15 \\ \frac{A-15}{10} & \text{if } 15 < A < 25 \\ \frac{35-A}{10} & \text{if } 25 < A < 35 \\ 0 & \text{if } A \geq 35 \end{cases}$$

$$\mu_{Old}(A) = \begin{cases} 0 & \text{if } A \leq 30 \\ \frac{A-30}{10} & \text{if } 30 < A < 40 \\ 1 & \text{if } 40 \leq A < 100 \\ 0 & \text{if } A \geq 100 \end{cases}$$

Utilizing these proposed approaches of fuzzy membership functions to capture and estimate the age of the orange trees with the corresponding inputs.

5 Discussion

To assess the effectiveness of the proposed model, we conducted a comprehensive measurement of orange trees, focusing on a range of diameters and heights. This analysis was conducted using the fuzzy rule approach, as shown in Figure 2. The primary goal of this research was to accurately predict the age of orange trees while simultaneously promoting sustainable agricultural practices (see Figure 3). The findings from our orange tree age estimation model have significant implications for agricultural science and forestry management. Accurate age estimation is crucial for understanding growth patterns, yield predictions, and implementing sustainable management practices. For instance, knowing the age of trees can help in scheduling pruning, pest management, and harvesting operations. Moreover, this model can contribute to the development of more tailored cultivation techniques that align with the specific needs of different tree ages, potentially enhancing productivity and sustainability. By enhancing our ability to monitor tree health and optimize yields, this study aims to contribute to more efficient orchard management.

Diameter (D) \ Height (H)	Very Short Height	Short Height	Medium Height	Tall Height
Small Diameter	**Very Young Age**	—	—	—
Medium Diameter	—	**Young Age**	**Young Age**	—
Large Diameter	—	**Middle-Aged**	**Young Age**	**Old Age**

Figure 2. Fuzzy rule mapping: Diameter and height combinations leading to tree age categories

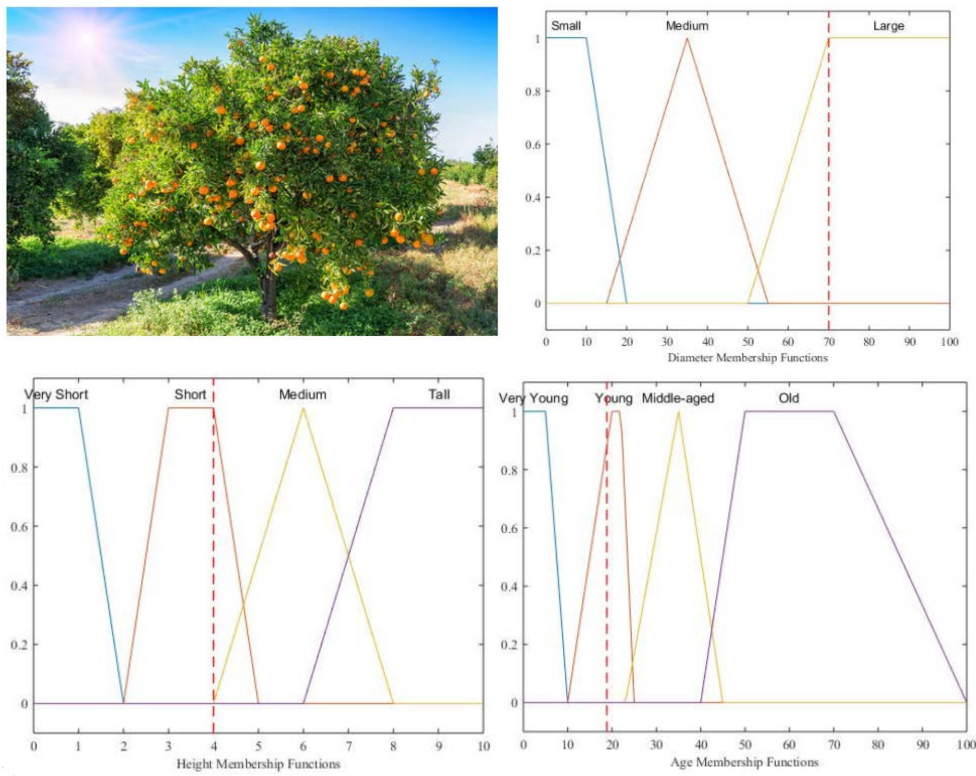

Figure 3. Applying propose FIS with diameter = 70 cm and height = 4 m, the predicted age is 18.1 years

Data collection was facilitated by the dedicated efforts of forestry department students, who worked diligently for 20 days to gather relevant measurements and observations. This extensive data collection phase ensured a robust dataset, crucial for validating our fuzzy logic-based age estimation model. The experiments were performed using a high-performance CPU equipped with 8 GB of RAM, which effectively managed the computational demands associated with processing large-scale images and analyzing the data. This infrastructure allowed us to execute complex calculations efficiently and obtain accurate results that will inform future agricultural practices. Table 2 provides an organized and insightful overview of orange trees, detailing the diameter (in centimeters), height (in meters), and corresponding age (in years) for each tree. It tracks the data for 300 individual trees, starting with tree 1, which has a diameter of 70.3 cm, a height of 5.1 meters, and an age of 19.5 years. As the dataset progresses, it shows slight variations in growth, with the 300th tree, for example, reaching a diameter of 88.5 cm, a height of 7.0 meters, and an age of 25.3 years.

Table 2. Data table for orange trees - Diameter, height, and age distribution

Tree Number	Diameter (cm)	Height (m)	Age (Years)
1	70.3	5.1	19.5
2	72.6	5.5	20.3
3	72.3	5.4	20.1
4	70.2	5.0	19.4
...	...	...	...
300	88.5	7.0	25.3

Table 2 serves as a crucial resource for understanding the relationship between these variables in orange trees, particularly how diameter and height influence age estimation. It provides a foundation for analyzing growth patterns and potentially improving agricultural practices.

Figure 4 further elucidates the trends in age related to diameter and height. By displaying the progression of age against both metrics, it emphasizes the consistent growth patterns that can be expected in orange trees. This visualization can serve as a predictive tool for foresters, enabling them to estimate the age of unmeasured trees based on their diameter and height measurements.

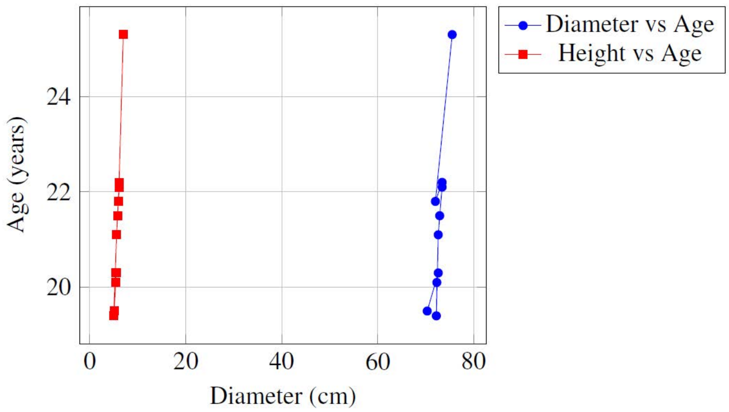

Figure 4. Line graph illustrating the age of orange trees in relation to diameter and height

5.1 Linear Regression Results

Intercept (X_0):
• Estimate: 8.9122
• P-Value: 0.0000

The intercept (X_0) represents the expected value of the dependent variable (age of the trees) when all independent variables (diameter and height) are zero. A p-value of 0.0000 indicates that the intercept is statistically significant, suggesting that the intercept's estimate is unlikely to be zero, implying a meaningful baseline age.

Diameter (X_1):
• Estimate: 0.1516
• P-Value: 0.0134

This coefficient indicates that for each unit increase in the diameter of the orange tree (measured in centimeters), the age of the tree is expected to increase by approximately 0.1516 years, assuming height remains constant. The p-value of 0.0134 indicates statistical significance, meaning there is strong evidence that diameter has a positive impact on the age of the trees.

Height (X_2):
• Estimate: 0.1194

• P-Value: 0.0444

The coefficient for height suggests that for each additional meter in height, the age of the tree is expected to increase by about 0.1194 years, assuming diameter remains constant. The p-value of 0.0444 signifies statistical significance, indicating that height also has a positive relationship with the age of the trees, though it is less impactful than diameter in this model.

R-squared:

• Value: 0.9800

The R-squared value indicates that approximately 98% of the variability in the age of the orange trees can be explained by the model that includes diameter and height as predictors (see Table 3). This high R-squared value suggests that the model fits the data well and that diameter and height are strong predictors of age.

Table 3. Key parameters of our model

Coefficient	Estimate	P-Value
X_0 = Intercept	8.9122	0.0000
X_1 = Diameter	0.1516	0.0134
X_2 = Height	0.1194	0.0444
R-squared	0.9800	

Figure 5 shows a positive linear relationship between the age of orange trees and both their diameter (top) and height (bottom), with blue dots representing actual data points and red lines indicating the fitted regression lines. The model suggests that as diameter and height increase, so does the age, with height appearing to have a slightly stronger correlation visually.

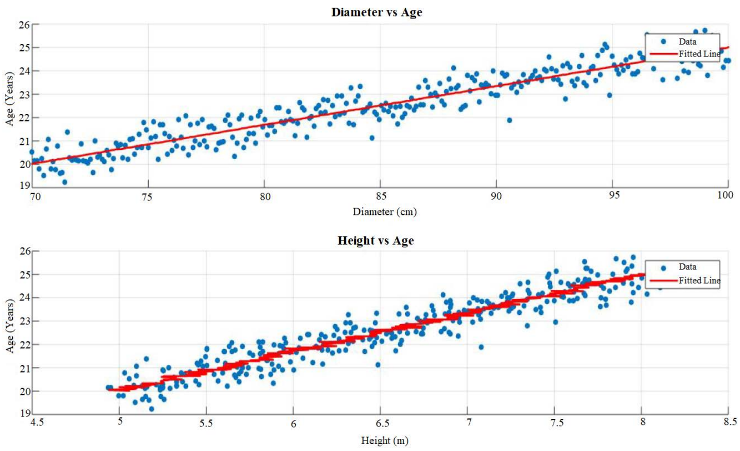

Figure 5. The relationship between inputs and output using regression model for proposed algorithm

5.2 Validation and Robustness

To ensure the reliability and generalizability of the proposed hybrid model, we implemented various validation techniques and performed comparative analysis with other non-invasive methods. This section discusses the steps taken to validate the model and assess its robustness.

5.2.1 Cross-validation of the regression model

To evaluate the predictive power and prevent overfitting of the regression model, we employed a cross-validation technique. Specifically, we used k-fold cross-validation, dividing the dataset into k = 10 folds. For each fold, the model was trained on 90% of the data and tested on the remaining 10%, iteratively covering all folds. This approach provided an average performance metric (such as R-squared) across all folds, indicating the model's stability and generalizability.

Results of Cross-Validation: The cross-validation process yielded an average R-squared value of 0.98, indicating high predictive accuracy and consistency across different subsets of data. Additionally, the low variance in R-squared values across the folds suggests that the model is not overly sensitive to data partitioning, further supporting its robustness.

5.2.2 Comparison with other methods

To contextualize the effectiveness of our approach, we compared the fuzzy-regression hybrid model with other established non-invasive methods for tree age estimation. These methods included: Integrating Active and Passive Remote Sensing Data for Forest Age Estimation [20]. This method presented a new workflow using ICESat-2 LiDAR data integrated with multisource remote sensing imagery to estimate forest age in Shangri-La, China. While this model provided reasonable estimates, it is less accurate compared to our approach, yielding an average R-squared value of around 0.67, significantly lower than our hybrid model. The age estimation model [2] is an approach to estimating individual tree ages based on time series diameter data a test case model is also evaluated, with an R-squared value of approximately 0.85. This result highlights the limitation of using a single time series diameter parameter for age estimation.

5.2.3 Robustness and sensitivity analysis

To further assess robustness, we conducted a sensitivity analysis on the input variables (diameter and height) to examine how variations in measurements affect the age prediction. The model was found to be relatively stable, with only minor fluctuations in predicted age values in response to small changes in input parameters. This indicates that the model is resilient to measurement errors, which enhances its practical utility. It is essential to recognize that external factors can significantly influence both the measurements taken and the subsequent model outputs. Environmental variables such as soil quality, water availability, and climate conditions can alter tree growth rates and, consequently, age estimations.

6 Practical Implementation Considerations for the Orange Tree Age Estimation System

Equipment Requirements

To implement the proposed age estimation system effectively in a typical orchard setting, several pieces of equipment will be necessary:

Measuring Tools

• Diameter Tape: For accurately measuring the diameter at breast height (DBH) of the trees. This is essential for the linear regression model.

• Measuring Rod or Laser Rangefinder: To measure tree height. Laser rangefinders can provide quick and precise height measurements with minimal error.

Data Recording Devices

• Tablet or Smartphone: Equipped with data collection software or applications to input measurements directly in the field. This can streamline data entry and reduce errors.

Time Required for Measurements

The time required for data collection will depend on the size of the orchard and the number of trees to be measured. On average, the time per tree may include:

• Measurement Time: Approximately 5-10 minutes per tree, including measuring DBH and height, depending on the ease of access and tree spacing.

• Data Entry: An additional 2-3 minutes for recording measurements and notes, especially if using a mobile device.

• Total Time: Therefore, for a standardized orchard with 100 trees, an estimated 10-15 hours may be needed for complete data collection, considering breaks and travel time.

Training Requirements

Training personnel is critical for ensuring accurate measurements and effective use of the equipment. Recommended training components include:

• Measurement Techniques: Instruction on proper DBH and height measurement techniques, including how to handle the measuring equipment.

• Data Entry Procedures: Training on using the data collection applications, focusing on accuracy in inputting data.

• Understanding the Model: Familiarization with the basics of the age estimation model so that personnel understand the relevance of their measurements and can troubleshoot minor issues.

7 Conclusion

In summary, the development of a fuzzy logic-based system for estimating the age of orange trees marks a significant advancement in agricultural practices. Traditional age estimation methods, while accurate, often prove impractical for large-scale applications due to their invasive nature and labor-intensive processes. Our research introduces an innovative approach that utilizes measurable parameters such as trunk diameter and height, offering a non-invasive and intuitive solution to age estimation. By employing fuzzy inference techniques, the model effectively captures the inherent variability in tree growth, resulting in a more accurate and reliable assessment of tree age. While the proposed orange tree age estimation model offers a significant advancement in assessing tree age non-invasively, it is important to critically reflect on its limitations to ensure a balanced discussion and enhance the research's credibility.

One of the primary limitations of the model is its dependency on the accuracy of the measured parameters, specifically trunk diameter and height. Variations in measurement techniques or errors due to environmental conditions can lead to inaccuracies in estimating age. Additionally, the model may not account for the inherent biological variability among individual trees of the same species, as growth rates can be influenced by factors such as genetic differences, health status, and local microclimates.

Looking ahead, several avenues for future research and development can further enhance this fuzzy logic-based system. Firstly, integrating additional parameters such as soil quality, climate conditions, and historical growth data could provide a more holistic understanding of factors affecting tree growth and age. Secondly, expanding the model to include different varieties of citrus trees could increase its applicability across diverse agricultural settings. Moreover, the implementation of machine learning algorithms could refine the model further, enabling real-time adjustments based on continuous data input and enhancing predictive accuracy.

Data Availability

The data used in this study, including measurements of trunk diameter, height, and age of orange trees, was collected as part of the research project and is available upon reasonable request. To protect privacy and ensure ethical usage, access to the dataset may be granted for academic and non-commercial purposes, subject to appropriate data-sharing agreements. Researchers interested in accessing the data for replication or further studies are encouraged to contact the corresponding author.

Conflict of Interests

The authors declare no conflicts of interest that could have influenced the results or interpretation of the research presented in this paper. This study was conducted independently, with a primary focus on advancing non-invasive methods for orange tree age estimation to benefit the field of agricultural science and orchard management.

References

[1] C. Bone, S. Dragicevic, and A. Roberts, "A fuzzy-constrained cellular automata model of forest insect infestations," *Ecol. Model.*, vol. 192, no. 1-2, pp. 107–125, 2006. https://doi.org/10.1016/j.ecolmodel.2005.09.013

[2] Y. R. Zhang, H. K. Li, X. H. Zhang, Y. C. Lei, J. J. Huang, and X. T. Liu, "An approach to estimate individual tree ages based on time series diameter data — A test case for three subtropical tree species in China," *Forests*, vol. 13, no. 4, p. 614, 2022. https://doi.org/10.3390/f13040614

[3] E. Handegard, "Identifying old Norway spruce and Scots pine trees by visual inspection: An analysis of the relationship between age, spatial distribution and morphological traits in trees," Master's thesis, Faculty for Natural Resource Management, Norwegian University of Life Sciences, Ås, 2020. https://hdl.handle.net/11250/2679452

[4] M. I. Ashraf, F. R. Meng, C. P. A. Bourque, and D. A. MacLean, "A novel modelling approach for predicting forest growth and yield under climate change," *PLoS One*, vol. 10, no. 7, p. e0132066, 2015. https://doi.org/10.1371/journal.pone.0132066

[5] M. A. Stokes, *An Introduction to Tree-Ring Dating*. University of Arizona Press, 1996.

[6] K. J. Xu, Z. Y. Zhang, W. W. Yu, P. Zhao, J. B. Yue, Y. P. Deng, and J. Geng, "How spatial resolution affects forest phenology and tree-species classification based on satellite and up-scaled time-series images," *Remote Sens.*, vol. 13, no. 14, p. 2716, 2021. https://doi.org/10.3390/rs13142716

[7] Z. L. Liu, C. H. Peng, T. Work, J. N. Candau, A. DesRochers, and D. Kneeshaw, "Application of machine-learning methods in forest ecology: Recent progress and future challenges," *Environ. Rev.*, vol. 26, no. 4, pp. 339–350, 2018. https://doi.org/10.1139/er-2018-0034

[8] M. Ricker, G. Gutiérrez-García, D. Juárez-Guerrero, and M. E. Evans, "Statistical age determination of tree rings," *PLoS One*, vol. 15, no. 9, p. e0239052, 2020. https://doi.org/10.1371/journal.pone.0239052

[9] B. S. Pedersen and B. McCune, "A non-invasive method for reconstructing the relative mortality rates of trees in mixed-age, mixed-species forests," *For. Ecol. Manage.*, vol. 155, no. 1-3, pp. 303–314, 2002. https://doi.org/10.1016/S0378-1127(01)00567-9

[10] L. A. Wells and W. Chung, "A deformable shape model for automatic and real-time dendrometry," *Forests*, vol. 14, no. 12, p. 2299, 2023. https://doi.org/10.3390/f14122299

[11] M. Niklasson, "A comparison of three age determination methods for suppressed Norway spruce: Implications for age structure analysis," *For. Ecol. Manage.*, vol. 161, no. 1-3, pp. 279–288, 2002. https://doi.org/10.1016/S0378-1127(01)00500-X

[12] J. Altman, J. Doležal, and L. Čížek, "Age estimation of large trees: New method based on partial increment core tested on an example of veteran oaks," *For. Ecol. Manage.*, vol. 380, pp. 82–89, 2016. https://doi.org/10.1016/j.foreco.2016.08.033

[13] K. K. Bohn, R. D. Nyland, and R. D. Yanai, "Comparing selection system and diameter-limit cutting in uneven-aged northern hardwoods using computer simulation," *Can. J. For. Res.*, vol. 41, no. 5, pp. 963–973, 2011. https://doi.org/10.1139/x11-027

[14] K. J. Niklas and H. C. Spatz, "Growth and hydraulic (not mechanical) constraints govern the scaling of tree height and mass," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 101, no. 44, pp. 15 661–15 663, 2004. https://doi.org/10.1073/pnas.0405857101

[15] A. Khan, "Use of fuzzy set theory in environmental engineering applications: A review," *Int. J. Eng. Res. Appl.*, vol. 7, no. 6, pp. 1–6, 2017. https://doi.org/10.9790/9622-0706020106

[16] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*. Prentice-Hall, New Jersey, 2001. https://doi.org/10.1109/MCI.2007.357196

[17] G. Matasci, T. Hermosilla, M. A. Wulder, J. C. White, N. C. Coops, G. W. Hobart, and H. S. Zald, "Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots," *Remote Sens. Environ.*, vol. 209, pp. 90–106, 2018. https://doi.org/10.1016/j.rse.2017.12.020

[18] M. Bayat, P. Bettinger, M. Hassani, and S. Heidari, "Ten-year estimation of Oriental beech (Fagus orientalis Lipsky) volume increment in natural forests: A comparison of an artificial neural networks model, multiple linear regression and actual increment," *For. Int. J. For. Res.*, vol. 94, no. 4, pp. 598–609, 2021. https://doi.org/10.1093/forestry/cpab001

[19] B. Du, Z. L. Yuan, Y. C. Bo, and Y. S. Zhang, "A combined deep learning and prior knowledge constraint approach for large-scale forest disturbance detection using time series remote sensing data," *Remote Sens.*, vol. 15, no. 12, p. 2963, 2023. https://doi.org/10.3390/rs15122963

[20] F. Cheng, R. J. Yang, and J. N. Wu, "Integrating active and passive remote sensing data for forest age estimation in Shangri-La City, China," *Forests*, vol. 15, no. 9, p. 1622, 2024. https://doi.org/10.3390/f15091622