Fuzzy Logic Modelling for Disease Severity Prediction in Cotton Crop using Weather Parameters

Kiran Devi[1], Vikash Siwach[2], Manju Singh Tonk[3], Anu[4]

[1,2,3,4] *Department of Mathematics & Statistics, COBS&H, CCS HAU, Hisar*

---***---

Abstract - *The relationship between weather parameters and disease severity has long been established. Disease severity measurement in crops is a major challenge, as it plays a crucial role in yield estimation and determining the control factors necessary to enhance crop productivity. Mathematical models assist in disease severity assessment and management. Fuzzy logic models provide a flexible framework for dealing with imprecise and uncertain input, making them suited for modelling complicated relationships in real-world systems. Cotton leaf curl disease (CLCuD) is a viral disease that affects cotton plants and it is a significant threat to cotton production in many regions in India. In this study, a fuzzy logic system (FLS) based model has been proposed for prediction of disease severity based on potential weather variables such as temperature, relative humidity, sunshine and rainfall etc. which are indicators of disease severity. This model has been using comprehensive weather dataset and corresponding disease severity levels for model formulation, training and validation. Different performance indicators, including accuracy, precision, recall, and F1-score, were used to evaluate the model's predictive ability. Furthermore, sensitivity analysis was performed to identify the most influential weather parameters in disease severity prediction. The experimental results demonstrated that the developed FLS achieved significant accuracy in predicting disease severity based on weather parameters. The model exhibited the ability to capture complex non-linear relationships between weather conditions and disease severity, providing valuable insights for disease management.*

Key words: Fuzzy logic system, Cotton leaf curl disease, Disease severity, Weather parameters, Sensitivity analysis etc.

1. INTRODUCTION

A mathematical paradigm for thinking in the presence of uncertainty and ambiguity is fuzzy logic [1]. It's particularly helpful when dealing with complex and imprecise data that can't be expressed easily using standard binary logic. Fuzzy logic, rather than using firm true or false values, provides a mechanism to handle and manipulate this uncertain data. Control systems, pattern recognition, decision-making systems, expert systems, and automotive systems are all applications of fuzzy logic models. They're employed in a variety of applications, including industrial process control, robotics, image processing, voice recognition, and automobile control systems [2]. Fuzzy logic is a versatile and robust method for dealing with uncertainty and imprecision in a variety of fields.

Agriculture has a crucial part in the Indian economy because it directly or indirectly supports more than half of the country's population [3]. The agricultural sector in India is noted for its size and production, which can be linked to the country's unique agro-climatic conditions. Despite its economic importance, farmers encounter numerous hurdles in increasing crop yields. Numerous efforts have been made to identify the primary causes causing poor crop productivity. Pest and disease modelling using diverse methodologies is one key approach to reducing the influence on agricultural yields [4]. Using cutting-edge techniques, agricultural forecasting approaches, including pest and disease forecasting, have been developed recently [5], [6]. In addition, internet-based forecasting tools and decision support systems have been introduced to successfully predict and manage crop diseases [7], [8]. Furthermore, advances in dispersal modelling and projected meteorology have aided in disease warnings and allowed farmers to adopt preventive actions [9], [10], [11]. These collaborative activities aim to assist Indian farmers in overcoming barriers and increasing crop yield in a sustainable manner.

India is the world's second-largest producer of cotton, after China. Cotton has been cultivated in India for centuries, and it plays a significant role in the country's agricultural sector and economy [12]. The country accounts for about 25% of the global cotton production. The production of cotton in India has been steadily increasing over the years. In the 2020-2021 season, India produced approximately 30 million bales (each weighing 170 kilograms) of cotton. India cultivates various cotton varieties, including both hybrid and genetically modified (GM) cotton. Cotton cultivation in India faces several challenges, including pests and diseases, water scarcity, fluctuations in market prices, and the availability of quality seeds [13]. Pest management, irrigation facilities, and access to credit for farmers are areas of continuous focus for improving cotton production.

A severe viral hazard to plants in the cotton crop, cotton leaf curl disease (CLCuD) offers a considerable challenge to cotton production in numerous Indian states. The adverse effects of this disease result in reduced crop yield and significant economic losses for farmers. Understanding the contributing factors to disease severity is essential for implementing effective disease management and crop protection strategies [15]. Management of CLCuD involves a combination of cultural, chemical, and genetic strategies [16]. Cultural practices such as crop rotation, removal of infected plants, and timely sowing can help reduce the disease incidence. Chemical control through the use of insecticides to manage whitefly populations is also employed, although it must be done judiciously to avoid the development of insecticide resistance. Additionally, the development and cultivation of genetically modified (GM) cotton varieties, such as Bt cotton, which carry genes for resistance against CLCuD, have proven effective in mitigating the disease's impact. Ongoing research and monitoring are essential to understand the epidemiology of CLCuD, identify new virus strains, and develop improved management strategies. By addressing the challenges posed by CLCuD, researchers and farmers aim to sustain cotton production and safeguard the livelihoods of those involved in the cotton industry.

The assessment of disease severity in crops has long been intertwined with the influence of weather parameters, as they play a vital role in quantitatively predicting yield and determining the necessary control factors to enhance crop productivity [17]. Mathematical models have emerged as valuable tools for disease severity assessment and management, providing a means to capture complex relationships within real-world systems [18]. Weather parameters are most important factor for disease severity in crops, predicting crop yield and implementing effective control measures [19]. Predicting disease severity in cotton crops is also crucial for cotton farmers and the industry [20]. It enables timely interventions to mitigate yield losses and optimize resource allocation based on disease severity. Understanding disease severity patterns helps improve pest management strategies and reduce costs, contributing to the economic sustainability of cotton farming [21]. Disease severity predictions aid in crop planning and decision-making, ensuring optimal yields and minimizing the impact of diseases. Furthermore, accurate severity predictions facilitate breeding programs for developing disease-resistant cotton varieties, enhancing the long-term viability of cotton production. Disease severity prediction is important in effective disease management, maximizing profitability, and ensuring the success of the cotton industry. Accurately measuring disease severity in crops presents a significant challenge for agricultural researchers and practitioners.

Fuzzy logic plays a significant role in the prediction of disease severity [22] particularly when dealing with complex and uncertain data. Fuzzy logic is a mathematical framework that allows you to represent and deal with unclear and ill-defined data. When exact mathematical models aren't appropriate or practical, it offers a flexible approach to modelling and decision-making. When applied to the prediction of disease severity, fuzzy logic models utilize fuzzy sets and linguistic variables to capture the inherent uncertainty and imprecision in the relationship between disease-related factors, such as weather parameters, and disease severity [23].

2. MATERIALS AND METHODS

2.1 DATA

In the developed fuzzy model data were taken from the website Indian Council of Agricultural Research [Ministry of Agriculture and Farmers Welfare] of cotton crop in the region HAU, hisar. We used 4 years' data from 2019 to 2022 in this paper for the developed model the data include weekly maximum and minimum temperature(°c), weekly morning and evening relative humidity (%), weekly rainfall(mm), weekly sunshine(hours), CLCuD incidence (%), Percent Disease Intensity (PDI) on the weekly basis. Further, we divide the data into two parts for training and testing of the developed fuzzy logic model.

2.2 Fuzzy Logic System

The fuzzy logic model, which was introduced by Zadeh in 1965, is a mathematical procedure that aims to replicate the human thought process using an 'IF-THEN' rule system. This approach permits for the translation of human reasoning into a formal mathematical framework. According to Zadeh [24], successful fuzzy modelling relies on four fundamental components which are respectively:

(i) Fuzzification: Fuzzification is the process of transforming traditional inputs into fuzzy inputs. It entails using fuzzy sets with membership functions to express data as fuzzy values. Typically, these fuzzy sets map data to a region of interest within the range of [0,1]. The fuzzy set F ⊆ U is given by:

$$F = \{(a, u_F(a));\ a \in U\}, 0 \le u_F(a) \le 1 \qquad \dots \qquad (1)$$

Where a is any element of U and $u_F(a)$ is the membership function of a in F.

(ii) Fuzzy rule: The fuzzy rule base's set of rules is meant to address a wide variety of possible fuzzy interactions between input and output variables. The IF-THEN expressions in Equation 2 represent these rules [25]. The quantity of input parameters and their corresponding membership functions determines the number of rules in the base.

$$if\ a_1\ is\ A_{i,1}\ and\ a_2\ is\ A_{i,2}\ and\ \ldots a_p\ is\ A_{i,p}\ then\ b\ is\ B_{i,1} \qquad (2)$$

where $a_1, a_2, \ldots, a_p$ are the input variables and b is the output variable. The linguistic input variables are assumed the terms $A_{i,j}$, and the rule k use the index i. The linguistic output variable is assumed the term B_i.

(iii) Fuzzy inference engine: As the tactical element in charge of processing all accessible fuzzy rules from the rule base, the fuzzy inference engine plays a critical role. Its primary function is to determine how to transform a set of inputs into the corresponding outputs. It accomplishes this by utilizing IF-THEN rules in conjunction with the connectors "OR" or "AND" to construct critical decision rules [24]. Among the frequently used strategies for the fuzzy inference system, Mamdani [26] and Takagi and Sugeno [27] both proposed two noteworthy methods. Due to their efficiency in handling fuzzy rule-based systems and producing useful output from ambiguous or inaccurate input data, these techniques are widely used.

(iv) Defuzzification: Defuzzification is the process of converting the inference engine's fuzzy outputs into numerical outputs [28].

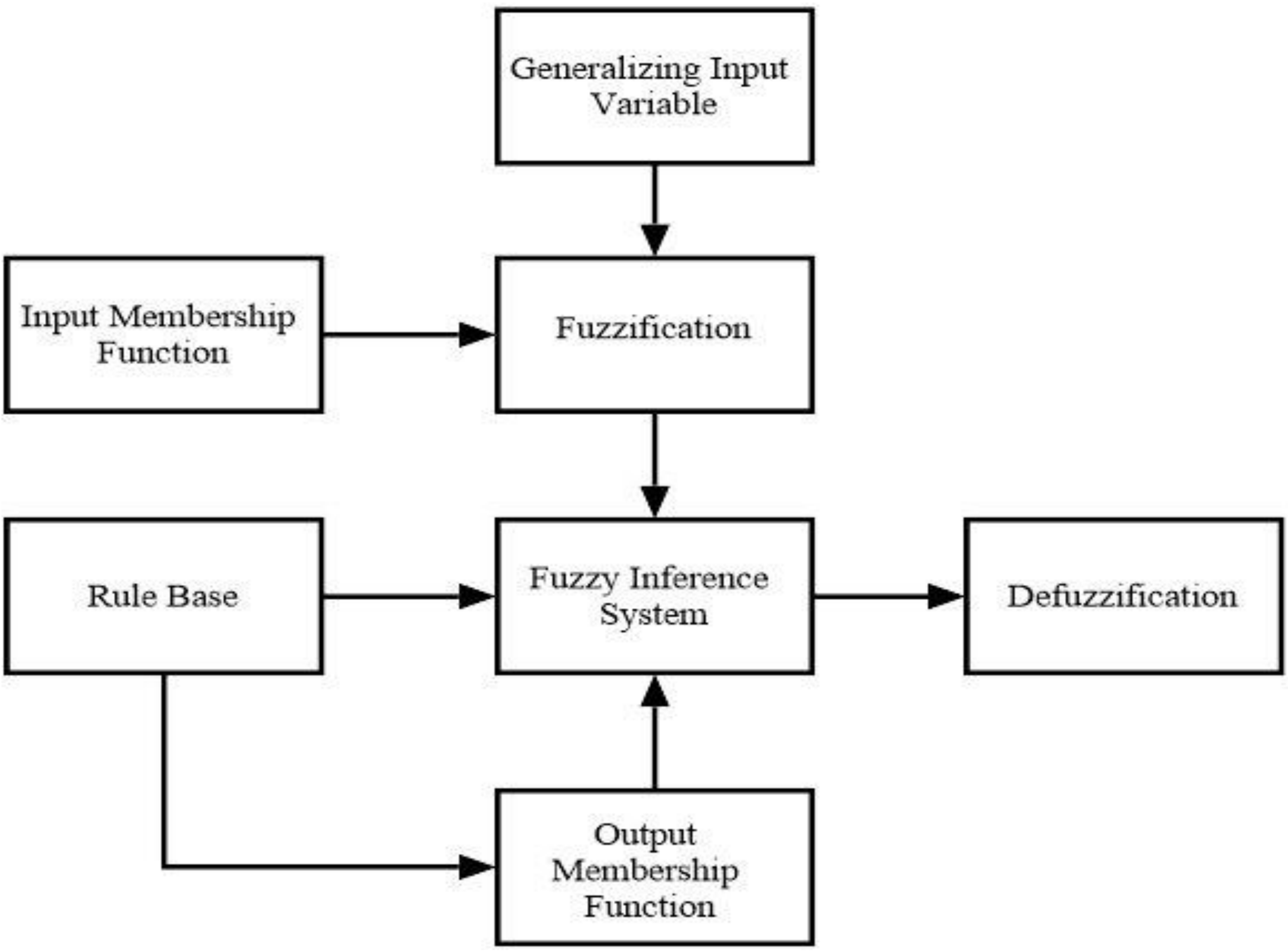
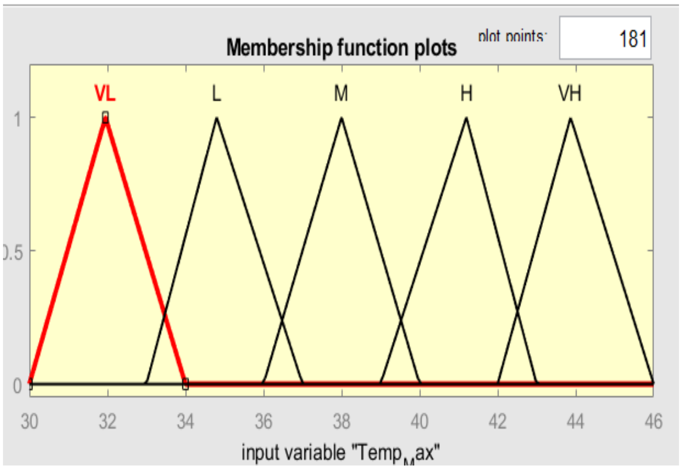
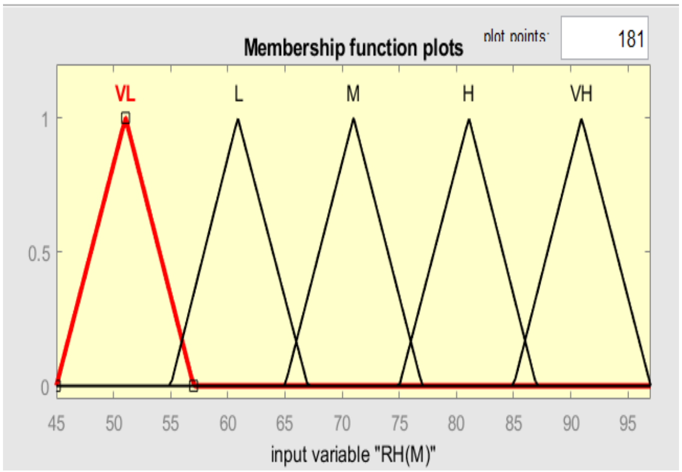


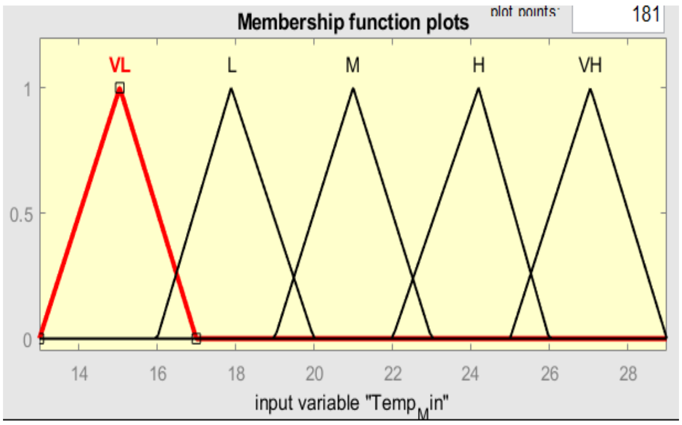
Fig – 1: Flowchart of fuzzy logic model

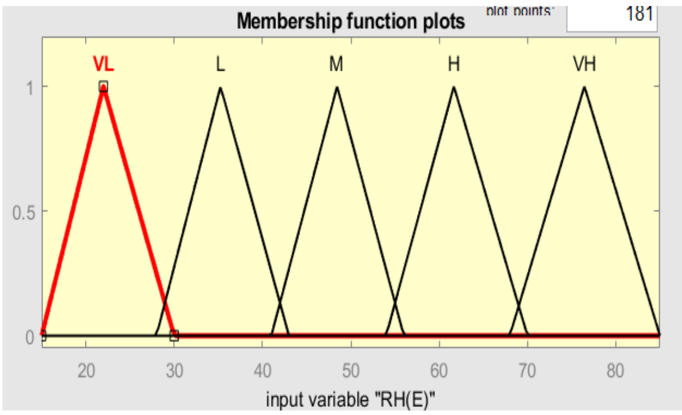
2.3 Developing the fuzzy model

The fuzzy framework was built with the Mamdani minimum-maximum inference engine and the MATLAB fuzzy logic toolbox (Mathworks, 2018). The following input and output variables were taken into account in this model: maximum temperature (Temp Max), minimum temperature (Temp Min), morning relative humidity (RH(M)), evening relative humidity (RH(E)), rainfall RF, sunshine (SH), CLCuD incidence, and PDI. Linguistic phrases were created to explain the input variables' fuzzy inputs in order to represent the input variables. These designations, which covered all conceivable input variations, included "very low" (VL), "low" (L), "medium" (M), "high" (H), and "very high" (VH). Similar to CLCuD incidence, and the output variable for PDI was divided into seven classes to account for all possible fuzzy outputs: immune (I), highly resistant (HR), resistant (R), moderately resistant (MR), moderately susceptible (MS), susceptible (S), and highly susceptible (HS). The input and output parameters of the fuzzy model were both described by triangular membership functions. The expert's knowledge and the experimental conditions were taken into account while determining the number and range of membership functions. A rule editor was used to develop the fuzzy model, which produced 89 rules that represented all of the interactions between the input and output variables. Finally, the Center of Gravity (COG) approach, a popular defuzzification technique, was employed to provide a real-valued (numerical) output.









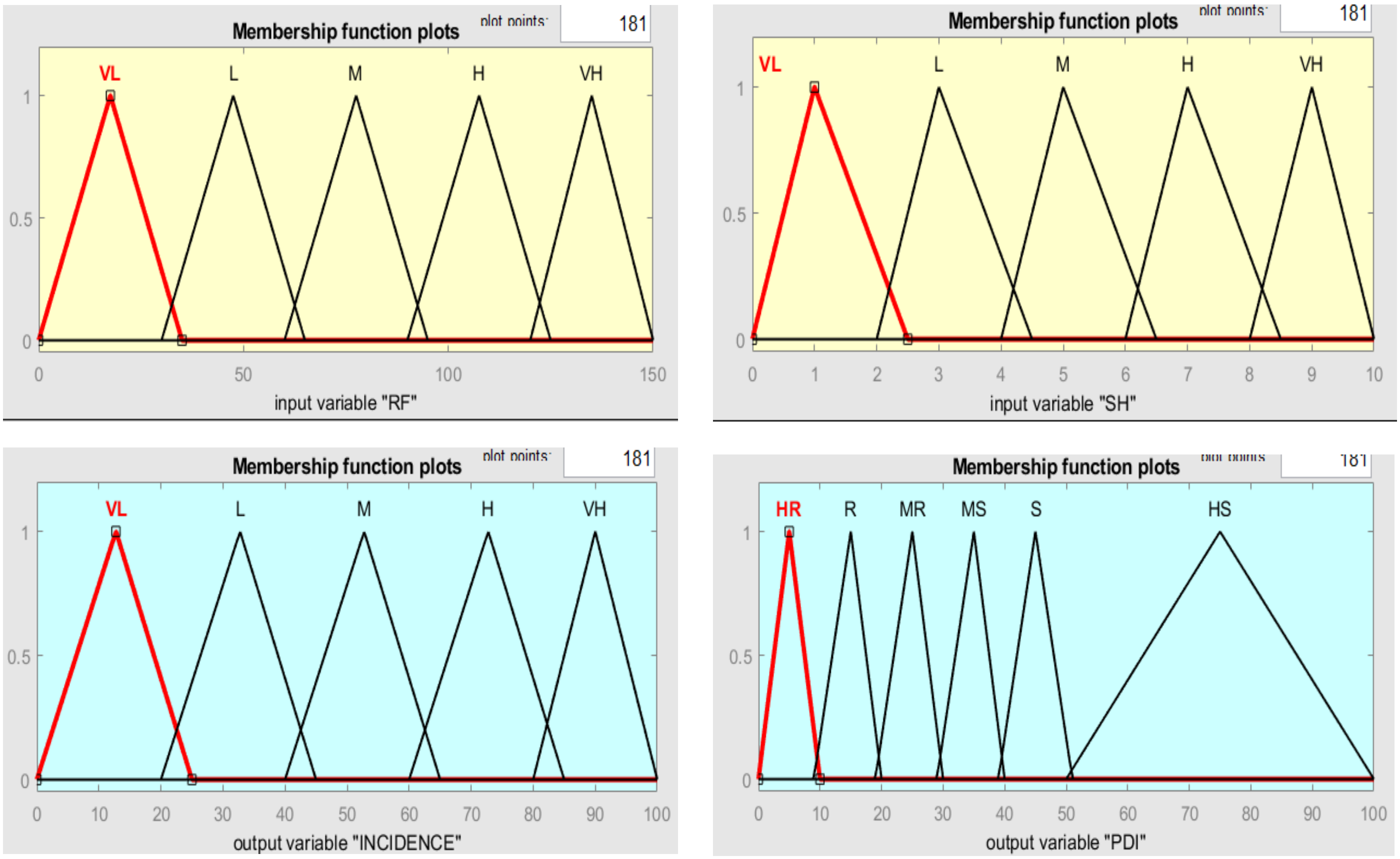


Fig – 2: Membership function defined for fuzzy input and output variables

Table 1: Fuzzy rules to estimate plant disease

Rule no.	rule
1.	If (Temp max is VH) and (Temp min is H) and (RH(M) is VL) and (RH(E) is VL) and (RF is VL) and (SH is VH) then (Disease incidence is VL) (PDI is I)
2.	If (Temp max is VH) and (Temp min is VH) and (RH(M) is VL) and (RH(E) is VL) and (RF is VL) and (SH is VH) then (Disease incidence is VL) (PDI is I)
3.	If (Temp max is H) and (Temp min is VH) and (RH(M) is M) and (RH(E) is L) and (RF is L) and (SH is H) then (Disease incidence is VL) (PDI is HR)
.	
.	
.	
50.	If (Temp max is L) and (Temp min is H) and (RH(M) is VH) and (RH(E) is M) and (RF is VL) and (SH is H) then (Disease incidence is VH) (PDI is HS)
.	
.	
.	
89.	If (Temp max is VL) and (Temp min is L) and (RH(M) is VH) and (RH(E) is M) and (RF is VL) and (SH is L) then (Disease incidence is VH) (PDI is S)

2.4 Criteria for evaluation

The predictive performance of the model was assessed using various performance metrics such as accuracy, precision, recall, and F1-score. Sensitivity Analysis (SA) is a valuable tool for evaluating and comprehending the influence and relative significance of parameters inside the modelling process. It is beneficial to determine how changes in input values affect the model's output values. The cosine amplitude method (CAM) was employed in this study to identify the variables influencing CLCuD that are most responsive to modification [25]. The level of sensitivity for each input factor was determined by determining the strength of the magnitude M_{ij} between the output and the specific input parameters under evaluation. A larger CAM value show the greater impact on output parameters. Consider the 'n' data samples that were gathered from the datasets that made up the common data array A, which is designated as

$$W = \{w_1, w_2, w_3, \ldots, w_n\} \qquad \ldots (3)$$

In the data array A, each member a_i is a vector with length m, so

$$w_i = \{w_{i1}, w_{i2}, w_{i3}, \ldots, w_{im}\} \quad \ldots \quad (4)$$

In this case, it is important to think of each data pair w_i and w_j as a point in an m-dimensional space, where each point needs m coordinates for a complete representation. The following equation is used to calculate and depict how strongly certain data pairs are related.

$$M_{ij} = \frac{\sum_{k=1}^{m} w_{ik} w_{jk}}{\sqrt{\sum_{k=1}^{m} w_{ik}^2 \sum_{k=1}^{m} w_{jk}^2}}, \quad 0 \leq M_{ij} \leq 1 \qquad \ldots (5)$$

Where i, j = 1, 2, 3, …, n

3. RESULTS AND DISCUSSION

3.1 Fuzzy Framework

In this study, we developed a fuzzy logic model to predict the occurrence of the Cotton Leaf Curl Disease (CLCuD) based on various weather parameters. The model achieved an impressive accuracy of 90%, indicating its ability to correctly classify instances as either diseased or non-diseased with a high level of precision. Furthermore, the model exhibited a recall of 100%, suggesting that it effectively identified all true positive cases of CLCuD, thereby minimizing the chances of false negatives. The high precision of 80% implies that when the model predicted an instance to be CLCuD positive, it meant that it was correct 80% of the time. This is crucial in the context of agricultural disease prediction, as false positive predictions could lead to unnecessary and potentially costly treatments for healthy crops. The model's ability to achieve a recall of 100% is particularly noteworthy since it means that it did not miss any cases of CLCuD, avoiding the risk of false negatives, which could have devastating consequences for crop health and yield. The F1 score provides a fair evaluation of the model's overall performance by accounting for both false positives and false negatives. The high F1 score indicates that the fuzzy logic model strikes a good balance between precision and recall, effectively capturing the trade-off between these two metrics. The F1 score for this model was found to be 88%.

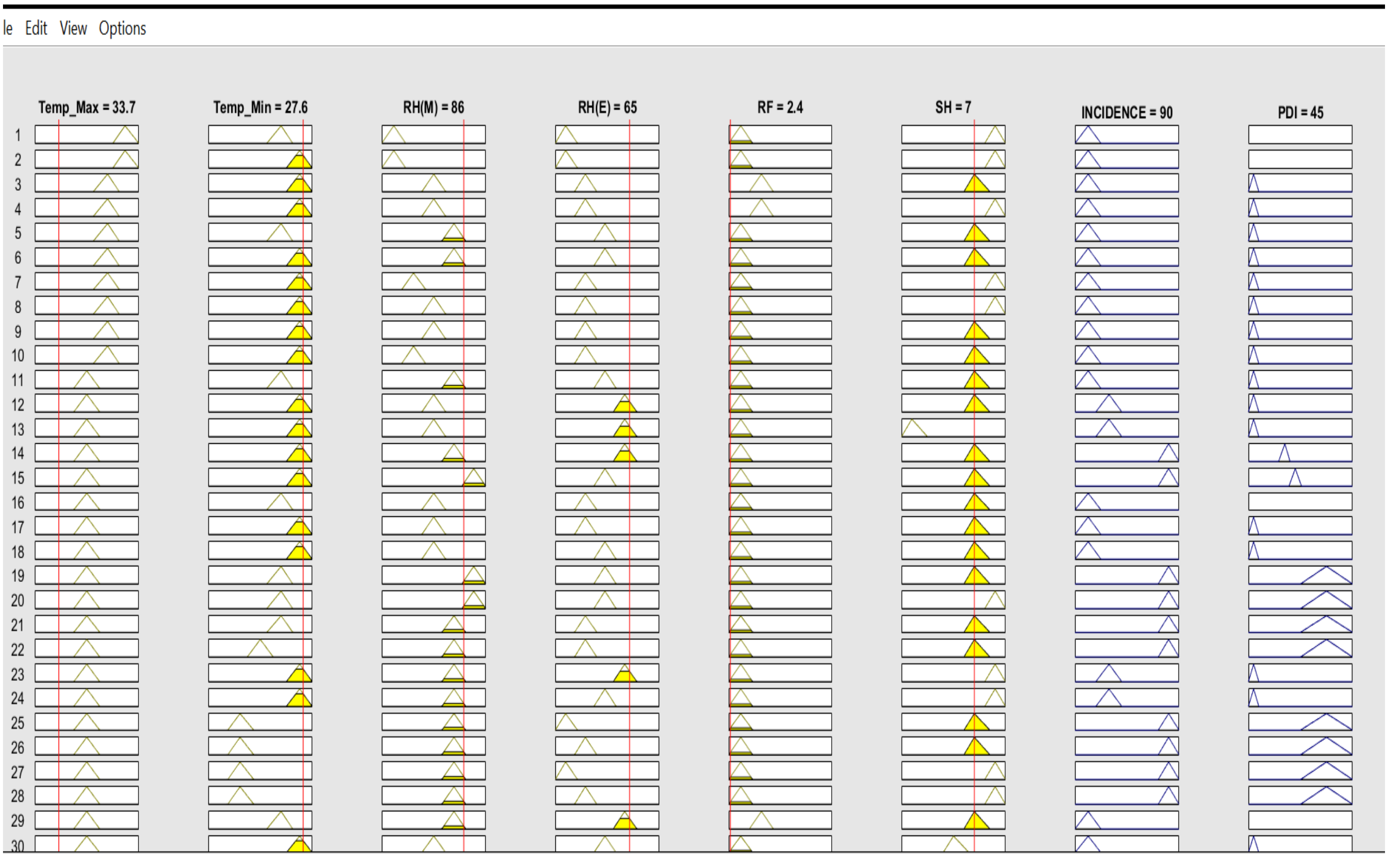
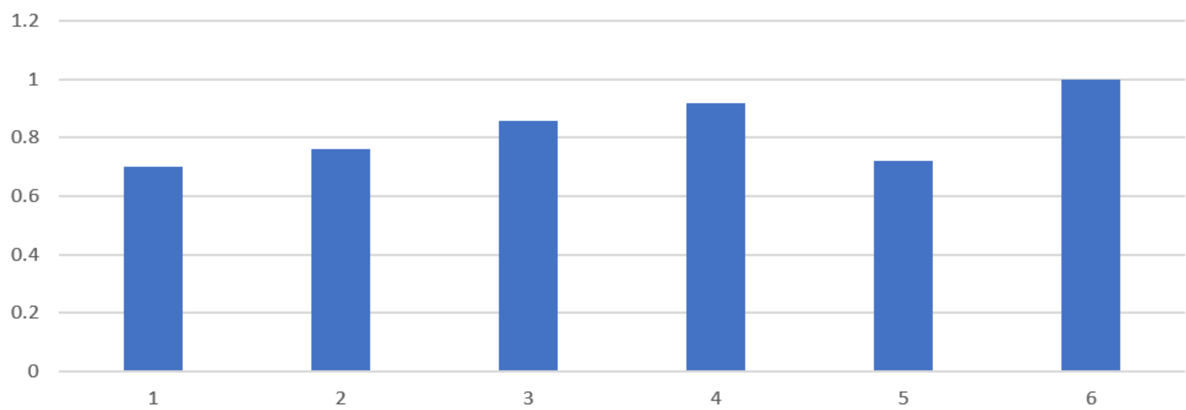


Fig – 3: rule viewer window shows some rules of the developed fuzzy model

3.2 Sensitivity Analysis

By utilizing Eq. 5, a comprehensive set of sensitivity analyses (SAs) was performed on both input and output parameters. The CAM method was employed to determine the values of M_{ij} representing the relationships between the predicted values of disease by the fuzzy model and the associated input factors. The results, depicted in Fig. 4, illustrate that the obtained M_{ij} values for all inputs were consistently close to 1. This indicates that all input factors play an important role in contributing to the development of the fuzzy models. Furthermore, these results emphasize that all the inputs were crucial in predicting the disease, and no parameter should be overlooked or disregarded.



where 1,2,3, 4, 5 and 6 represent Temp (max), Temp(min), RH(M), RH(E), RF and SH respectively.

Fig – 4: Sensitivity analysis of disease and each input parameter for the fuzzy logic model

4. FUTURISTIC APPROACH

It is essential to acknowledge the limitations of the study. While the model demonstrated promising accuracy, precision, recall, and F1 score, its performance could still be affected by various factors. For instance, the availability and quality of historical data, the choice of fuzzy logic membership functions, and the selection of weather parameters could all influence the model's outcomes. To further enhance the model's predictive capabilities, future research could focus on incorporating additional data sources, such as soil moisture levels, crop management practices, and genetic factors of the cotton plants. Furthermore, combining ensemble approaches or incorporating machine learning algorithms with the fuzzy logic model may improve accuracy and robustness.

5. CONCLUSION

The fuzzy logic model successfully predicted the occurrence of CLCuD based on weather parameters with an accuracy of 90%, a precision of 80%, a recall of 100%, and an F1 score of 88%. These results underline the potential of fuzzy logic as a valuable tool in agricultural disease prediction, providing farmers and policymakers with valuable insights to make informed decisions and mitigate the impact of diseases on cotton crops. Nevertheless, continuous refinement and validation of the model are necessary to ensure its reliability and applicability in real-world scenarios. Additionally, the sensitivity analysis carried out in this work offered insightful insights into the robustness of the fuzzy logic model to changes in meteorological factors. In the context of disease prediction in cotton crops, the results of the sensitivity analysis together with the model's performance indicators lead to a thorough review and a deeper understanding of its capabilities and limitations.

6. REFERENCES

[1] L. A. Zadeh, "Fuzzy sets," Information and Control, vol. 8, no. 3, 1965, pp. 338-53.

[2] Belohlavek, Radim, Joseph W. Dauben, and George J. Klir, "Fuzzy Logic and Mathematics: A Historical Perspective," New York, 2017; online edn, Oxford Academic, 22 June 2017, doi:10.1093/oso/9780190200015.001.0001.

[3] L. Praburaj, "Role of Agriculture in the Economic Development of a Country," Shanlax International Journal of Commerce, vol. 6, no. 3, 2018, pp. 1-5.

[4] M. Donatelli, R.D. Magarey, S. Bregaglio, L. Willocquet, J.P.M. Whish, and S. Savary, "Modelling the impacts of pests and diseases on agricultural systems," Agric. Syst., vol. 155, 2017, pp. 213-224.

[5] K. Leksakul, P. Holimchayachotikul, and A. Sopadang, "Forest of off-season longan supply using fuzzy support vector regression and fuzzy artificial neural network," Comput. Electron. Agric., vol. 118, 2015, pp. 259- 269.

[6] M.W. Shi, "Based on time series and RBF network plant disease forecasting," Procedia Eng., vol. 15, 2011, pp. 2384-2387.

[7] W. Pavan, C.W. Fraisse, and N.A. Peres, "Development of a web – based disease forecasting system for strawberries," Comput. Electron. Agric., vol. 75, no. 1, 2011, pp. 169-175.

[8] I.M. Small, L. Joseph, and W.E. Fry, "Development and implementation of the BlightPro decision support system for potato and tomato late blight management," Comput. Electron. Agric., vol. 115, 2015, pp. 57-65.

[9] K.A. Garrett, S.P. Dendy, E.E. Frank, M.N. Rouse, and S.E. Travers, "Climate change effects on plant disease: genomes to ecosystem," Annu. Rev. Phytopathol., vol. 44, 2006, pp. 489-509.

[10] P. Skelsey, G.J.T. kessel, A.A.M. Holtslag, A.F. Moene, and W. Van Der Werf, "Regional spore dispersal as a factor in disease risk warnings for potato late blight: a proof of concept," Agric. For. Meteorol., vol. 149, 2009, pp. 419-430.

[11] W. Luo, S. Pietravalle, S. Parnell, F. van den Bosch, T.R. Gottwald, M.S. Irey, and S.R. Parker, "An improved regulatory sampling method for mapping and representing plant disease from a limited number of samples," Epidemics, vol. 4, no. 2, 2012, pp. 68-77.

[12] D. S. Jadhav, S. S. More, S. D. Choudhari, and S. A. Chaudhari, "Growth and Instability in Area, Production and Yield of Cotton in India," International Journal of Current Microbiology and Applied Sciences, vol. 11, 2020, pp. 1411-1417.

[13] M. Mollaee, A. Mobli, N. K. Mutti, S. Manalil, and B. S. Chauhan, “Challenges and opportunities in cotton production,” cotton production, 2019, pp. 371-390.

[14] A. Farooq, J. Farooq, A. Mahmood, A. Shakeel, K. A. Rehman, A. Batool, and S. Mehboob, “An overview of cotton leaf curl virus disease (CLCuD) a serious threat to cotton productivity,” Australian Journal of Crop Science, vol. 5, no. 13, 2011, pp. 1823-1831.

[15] M. U. Rahman, A. Q. Khan, Z. Rahmat, M. A. Iqbal, and Y. Zafar, “Genetics and Genomics of Cotton Leaf Curl Disease, Its Viral Causal Agents and Whitefly Vector: A Way Forward to Sustain Cotton Fiber Security,” Front Plant Sci., vol. 4, no. 8, 2017 Jul, doi: 10.3389/fpls.2017.01157. PMID: 28725230; PMCID: PMC5495822.

[16] J. A. Singh, S. Sohi, D. S. Brar, I. Denhdm, D. Russel, and R. Briddon, “Management of cotton leaf curl virus disease in India,” Proc. ICAC regional consultation insecticide resistance management in cotton. CCRI, Multan, 1999, pp. 277-284.

[17] R. E. Gaunt, “The relationship between plant disease severity and yield,” Annual review of phytopathology, vol. 33, no. 1, 1995, pp. 119-144.

[18] L. M. Contreras-Medina, I. Torres-Pacheco, R. G. Guevara-Gonzalez, R. J. Romero-Troncoso, I. R. Terol-Villalobos, and R. A. Osornio-Rios, “Mathematical modeling tendencies in plant pathology,” African Journal of Biotechnology, vol. 8, no. 25, 2009.

[19] P. Sharma, B. K. Singh, and R. P. Singh, “Prediction of potato late blight disease based upon weather parameters using artificial neural network approach,” In 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), IEEE, 2018 July, pp. 1-13.

[20] M. A. Khan, and H. A. Khan, “Cotton leaf curl virus disease severity in relation to environmental conditions,” Pak. J. Biol. Sci, vol. 3, no. 10, 2000, pp. 1688-1690.

[21] H. Saeed, M. Ehetisham-ul-Haq, M. Atiq, M. Kamran, M. Idrees, S. Ali, and S. Yasin, “Prediction of cotton leaf curl virus disease and its management through resistant germplasm and bio-products,” Archives of Phytopathology and Plant Protection, vol. 51, 2018, pp. 170-186.

[22] V. Tilva, J. Patel, and C. Bhatt, “Weather Based Plant Diseases Forecasting Using Fuzzy Logic,” Nirma University International Conference on Engineering (NUiCONE), 2013, pp. 1-5.

[23] I. Hernandez, S. Gutierrez, S. Ceballos, F. Palacios, S. L. Toffolatti, G. Maddalena, and J. Tardaguila, “Assessment of downy mildew in grapevine using computer vision and fuzzy logic,” Development and validation of a new method, 2022.

[24] L.A. Zadeh, “The concept of a linguistic variable and its application to approximate reasoning—I,” Inf. Sci., vol. 8, no. 3, 1975, pp. 199–249.

[25] T.J. Ross, “Fuzzy Logic with Engineering Applications (2nd edn),” Wiley, Hoboken, 2005.

[26] E.H. Mamdani, “Application of fuzzy algorithms for control of simple dynamic plant,” In: Proc. Inst. Elec. Eng., vol. 121, no. 12, 1974, pp. 1585–1588.

[27] T. Takagi, and M. Sugeno, “Fuzzy identification of systems and its applications to modeling and control,” IEEE Trans. Syst. Man Cybernetics SMC, vol. 15, no. 1, 1985, pp. 116–132.

[28] W. Van-leekwijck, and E.E. Kerre, “Defuzzification: criteria and classification,” Fuzzy Sets Syst., vol. 108, no. 2, 1999, pp. 159–178.