JAISCR, 2023, Vol. 13, No. 4, pp. 247 – 272
10.2478/jaiscr-2023-0018

AN EXPLAINABLE AI APPROACH TO AGROTECHNICAL MONITORING AND CROP DISEASES PREDICTION IN DNIPRO REGION OF UKRAINE

Ivan Laktionov[1,*], Grygorii Diachenko[1], Danuta Rutkowska[2], Marek Kisiel-Dorohinicki[3]

[1]*Dnipro University of Technology,*
av. Dmytra Yavornytskoho, 19, Dnipro, UA49005, Ukraine

[2]*University of Social Sciences,*
90-113, Łódź, Poland

[3]*AGH University in Kraków, 30-059,*
Krakow, Poland

[*]*E-mail: Laktionov.I.S@nmu.one*

Submitted: 31st May 2023; Accepted: 11th September 2023

Abstract

The proliferation of computer-oriented and information digitalisation technologies has become a hallmark across various sectors in today's rapidly evolving environment. Among these, agriculture emerges as a pivotal sector in need of seamless incorporation of high-performance information technologies to address the pressing needs of national economies worldwide. The aim of the present article is to substantiate scientific and applied approaches to improving the efficiency of computer-oriented agrotechnical monitoring systems by developing an intelligent software component for predicting the probability of occurrence of corn diseases during the full cycle of its cultivation. The object of research is non-stationary processes of intelligent transformation and predictive analytics of soil and climatic data, which are factors of the occurrence and development of diseases in corn. The subject of the research is methods and explainable AI models of intelligent predictive analysis of measurement data on the soil and climatic condition of agricultural enterprises specialised in growing corn. The main scientific and practical effect of the research results is the development of IoT technologies for agrotechnical monitoring through the development of a computer-oriented model based on the ANFIS technique and the synthesis of structural and algorithmic provision for identifying and predicting the probability of occurrence of corn diseases during the full cycle of its cultivation.

Keywords: IoT, ANFIS, explainable AI, agrotechnical monitoring, disease prediction, crop.

1 Introduction

1.1 Relevance of the topic

The current trends in the agro-industrial sector imply the need to continuously search for scientifically based methods to enhance the efficiency of growing agricultural crops in the open-field. Such need results primarily from the global negative dynamics of environmental factors and the limitations of soil and climatic resources [1]. Crop preservation agrotechnical challenges throughout the crop cultivation cycle are critical in determining and maintaining the long-term export potential of many countries worldwide, not only in years of unexpected yield deviation due to adverse weather conditions but also in favourable years.

It is worth noting that a wide range of computer and information technologies for various purposes is currently used in the cultivation of various types of agricultural crops. To date, most of the technologies used are intellectualised by integrating functional components based on artificial intelligence methods and tools. Taking into account the significant theoretical and applied achievements in the field of intelligent systems for agrotechnical purposes, it should be noted that the theory of development and use of information intelligent technologies to support decision-making on preventive measures to preserve crop yields is currently being actively developed and formed [2]. Thus, the topic of the article is relevant and the significance of the expected scientific and applied result is the development of information technologies for distributed aggregation and intelligent analysis of soil and climatic data. This will allow for precise diagnosis and prediction of the probability of occurrence of corn diseases during the full cycle of its cultivation.

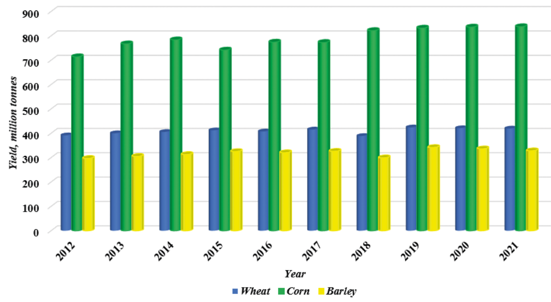
The main socio-economic effect of the research conducted in this article is the substantiation of ways of software and technical modernisation of domestic agricultural production facilities to increase the indicators of investment attractiveness, food security and export potential of Ukraine while decreasing the risks of food crisis in countries reliant on the volume of cereal crops production in Ukraine.

1.2 Review, critical analysis and systematisation of modern literature sources

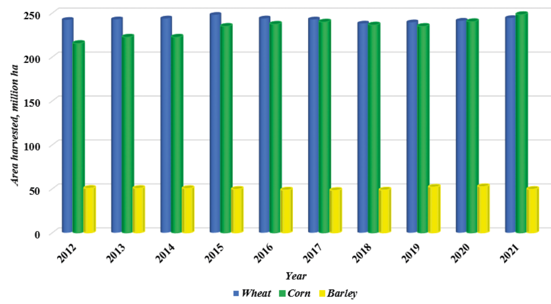
Based on the statistical analysis of data accumulated by the globally recognised FAO [3, 4], the most popular crops grown in the world in open-field conditions were identified. Cereals are the most commonly cultivated crops in the global agricultural practice. Their production has risen by a third over the last 20 years.

As of 2021 (the most recent update of available data) the total yield has exceeded 3 billion tonnes [4]. In turn, the commonly produced cereals in the countries of Eastern, Southeastern and Central Europe are wheat, corn and barley. The trend in the dynamics of yields, areas used for cultivation and specific yields worldwide according to FAO [3] is shown in Figure 1.

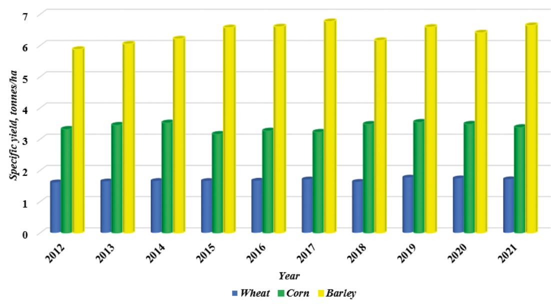
The data shown in Figure 1 confirms the global trend of steady dynamics in the popularity of cereal crop cultivation across diverse agroclimatic conditions and countries worldwide, emphasizing the need to synthesise scientific and applied approaches to crop preservation and rational use of resources and fertilisers during the full cycle of cultivation.



a) Yield



b) Area harvested



c) Specific yield

Figure 1. Statistical data on the cultivation of the most popular cereals worldwide

Table 1. Up-to-date technologies for intelligent analysis of time series of distributed measurement monitoring results

Title of the technique	**Purpose of use**	**References**
Artificial neural networks	This technique is commonly utilised for problem-solving in prediction, decision-making, image recognition, optimisation, and data mining. Depending on the complexity of the mathematical provision, it can be deployed in systems based on cloud, fog and edge architectures.	[5–13]
Fuzzy logic	It is used to solve decision-making support and data mining problems in expert monitoring, control and management systems of various architectures and levels of complexity, including those based on microprocessor devices.	[14–17]
Adaptive neuro-fuzzy inference system	It is an integral solution of artificial neural networks and fuzzy logic. One can be used to solve problems of prediction, optimisation and data mining in monitoring, control and management systems where the input data is implicitly specified. Depending on the complexity of the mathematical provision, it can be deployed in systems based on cloud, fog and edge architectures, including those built on microprocessor devices.	[18–22]
Genetic algorithms	They are used to solve a wide range of tasks related to artificial intelligence and machine learning, such as optimisation, regression and approximation, pattern recognition, decision-making, etc. GAs are mainly used at the modelling and optimisation stages in the synthesis of architectural solutions for intelligent information and computer-oriented monitoring and control systems.	[23, 24]
TinyML	This technique is utilised in solving machine learning tasks related to classification and prediction while considering the constraints of lightweight hardware and software suitable for deployment on low-cost microcontroller devices for use in automatic monitoring and control systems.	[25–27]

Table 2. Up-to-date scientific and applied advances in the development and research of explainable AI

Subject of development and research	Results obtained	References
A unified taxonomic approach to the methods of XAI based on the analysis of concepts and approaches in modern research.	More than 50 most cited review articles on XAI methods, metrics and method characteristics were identified and summarised into a single structured taxonomy of XAI methods.	[28]
A systematic meta-review of challenges and directions for future research in the field of XAI.	An approach to presenting the difference between the terms ”explainability” and ”interoperability” was proposed. Significant challenges and future directions of XAI research arising from the selected 73 articles were identified.	[29]
Analysis of the growing field of XAI with the aim of systematically summarising concepts and approaches to explanability and evaluating XAI methods.	A significant number of scientific studies were clustered using a hierarchical system that classifies theories and concepts related to the concept of explainability and approaches to evaluating XAI methods.	[30]
Methods and solutions for ensuring the explainability and interpretability of artificial intelligence, in particular for machine learning models.	A number of solutions were provided for tasks that require explainability and interpretability, and various methods and approaches to ensuring transparency of decision-making by artificial intelligence and machine learning models were covered.	[31]
Methods from the field of XAI to understand and explain the predictions obtained from artificial intelligence and machine learning models in agricultural sciences, in particular, to study the impact of various factors on crop yields.	New patterns between conditions and corn yields were identified. It was also found that the use of XAI can improve trust and clarity in the application of artificial intelligence in agriculture.	[32]
Factors influencing the adoption of artificial intelligence systems by agricultural farmers using the technology adoption model and the theory of planned behaviour.	It was shown that the perception of behavioural control has the greatest impact on the adoption of artificial intelligence systems in agriculture, followed by the personal position of farmers on the use of such systems. The modelled relationships explain 59 % of the total variation in adoption. The results of the study point to several possible directions and implications for increasing the adoption of AI systems in agriculture.	[33]
Vital system, which uses low-cost sensors and combines them with a highly efficient decision-making system based on artificial intelligence.	A fuzzy rule-based system has been developed that efficiently makes irrigation decisions for fields. Compared to other open-field agriculture systems, the researched and developed system uses the XAI approach, providing maximum efficiency and increased interpretability.	[34]
AI and XAI methods used in the context of Industry 4.0.	An overview of the main AI and XAI methods and their application in Industry 4.0 are provided and prospects and challenges for future research are pointed out.	[35]
Artificial intelligence and deep learning methods for detecting and classifying plant diseases in agriculture.	The use of Inception V3 and ResNet-9 models on the Plant Village and New Plant Disease datasets for disease detection and classification was proposed. XAI methods, such as LIME and Grad-CAM, were also used to understand the work of deep learning models.	[36]

One possible approach to adjusting cultivation regimes by practitioners in the agricultural sector is to use modern intellectualised information technologies for aggregating and precision analysis of measurement data with real-time decision-making support.

As of today, a significant amount of qualitative research findings are available on the development of methods and software and hardware implementation of tools and models for processing time- and space-distributed measurement data, as shown in Table 1.

When substantiating the selection of intelligent technology for aggregation and transformation of agrotechnical monitoring data to identify the probability of occurrence and development of crop diseases, the following characteristics of the monitoring object were taken into account:

– the input data is not clearly defined, but is the result of a generalisation of expert experience in the field of open-field agriculture;
– the main tasks of transformation: precision detection, prediction and decision-making support;
– the ability to deploy data aggregation and transformation software on low-cost microprocessor and sensor devices.

Given the applied specifics of the research problem being solved, an analysis of current scientific and applied advances in the development and application of explainable AI (XAI) in creating intelligent monitoring and control systems is necessary, as summarised in Table 2.

Thus, based on a detailed analysis of the results of the logical generalisation of known intelligent data transformation technologies (see Tables 1 and 2), taking into account the characteristic features of the monitored object, as mentioned above, it was found that it is advisable to conduct the research of the article using the ANFIS technique.

It is worth noting that the ANFIS technique is currently widely used in the construction of intelligent technical systems with decision-making support, including for agrotechnical purposes, as shown in Table 3.

It is also important to stress the fact that as of today, the scientific literature contains a significant number of research findings on solving the problems of plant disease identification using artificial intelligence methods. However, most of these studies are based on the technology of using artificial neural networks for identifying and recognising patterns in graphic images [42–46]. Given the significant qualitative scientific and practical achievements presented in the above articles, it is also noteworthy that such a technology based on graphic image recognition has several limitations in terms of the possibility of its practical integration into currently used low-cost agrometeorological stations as software and hardware nodes. That is, the software and hardware solution under development should be implemented as a functional component of the intellectual transformation of data on the soil and climatic condition of agricultural production facilities based on microprocessor devices as part of the currently used IoT networks for agrometeorological monitoring without any fundamental modification of their architecture.

Another fundamental requirement for the development is the aggregation and processing of data in near real-time with the ability to predict the probability of a particular disease, and most image recognition systems can detect and classify crop diseases that have already occurred, rather than predict the probability of their occurrence.

Therefore, the anticipated scientific and practical advancement of the known methods and technologies of intelligent data transformation for identifying and predicting the probability of disease occurrence in crops is as follows: taking into account typical agroclimatic conditions and the dynamics of soil and climatic factors when developing a computer model; consideration of many years of expert experience in the field of open-field crop production (including the specific influence of soil and climate factors depending on the types of crops cultivated and the differentiation of common plant diseases depending on the agroclimatic zones of cultivation) during the formal description of the stages of data aggregation and processing; the potential to integrate the developed computer model as a software and functional module based on low-cost microprocessor devices into currently used agrometeorological stations, which increase the efficiency of their use through the practical transition to fog- and edge-computing technologies.

Table 3. Examples of the use of the ANFIS technique in the construction of intelligent technical systems for various purposes

Subject of development and research	Application area	References
A computer model of an adaptive automatic temperature and humidity control system for the greenhouse growing area.	Protected ground vegetable growing	[37]
IoT technology for adaptive regulation of acidity and nutrient content in hydroponic irrigation solution.	Protected ground vegetable growing	[38]
An optimised computer model of a decision-making support system for the selection of construction machines.	Technical machines in the construction industry	[39]
Computer technology for heart disease prediction with decision-making support.	Medicine	[40]
Embedded microcomputer device for intelligent control of the drip irrigation system using alternative energy sources.	Crop production	[41]

1.3 Aim, object, subject and structure of the research

The aim of the present article is to substantiate scientific and applied approaches to improving the efficiency of computer-oriented agrotechnical monitoring systems by developing an intelligent software component for predicting the probability of occurrence of corn diseases during the full cycle of its cultivation.

The object of research is non-stationary processes of intelligent transformation and predictive analytics of soil and climatic data, which are factors of the occurrence and development of diseases in corn.

The subject of the research is methods and XAI models of intelligent predictive analysis of measurement data on the soil and climatic condition of agricultural enterprises specialised in growing corn.

The article is structured as follows: Section 1 provides details on the relevance of the topic, outlines the direction of research, and the current state of the subject area along with explaining the scientific and practical novelty of the results obtained; Section 2 contains information on the materials, methods and approaches used in the research; Section 3 presents the main quantitative and qualitative results of the research; Section 4 contains information on promising areas for future research; the conclusions containing the quantitive and qualitative results are drawn in Section 5.

1.4 Scientific and practical significance of the obtained results

The scientific and applied significance of the results of the article lies in the development of IoT technologies for agrotechnical monitoring through the development of the explainable AI model and the synthesis of structural and algorithmic provisions for identifying and predicting the probability of occurrence of typical corn diseases during the full cycle of its cultivation. The developed computer model is based on the ANFIS technique for transforming real-time soil and climatic data. In contrast to known models, it is compatible with most low-cost agrometeorological stations currently in operation. The hardware basis for the deployment of the developed model is low-cost microprocessors and sensor devices, and the information basis is measuring signals from sensors of soil and climatic parameters of agricultural crop production enterprises. Therefore, the implementation of the studied software and hardware solution does not necessitate a fundamental change in the architecture of currently used agrometeorological monitoring systems and networks, which corresponds to current trends in the development of fog and edge computing technologies [47].

2 General approaches, materials and research methods

2.1 Approaches to the research

The methodological basis for solving the scientific and applied problem is an integral approach using the following methods and techniques: logical generalisation and critical analysis of known scientific research and expert data; structural and algorithmic synthesis of information technologies; computer modelling; theory of adaptive neuro-fuzzy inference systems; theory of identification of nonlinear dynamic systems.

The primary research results were acquired through a computer experiment using the Matlab & Simulink environment in the specialised laboratories of Dnipro University of Technology (Dnipro, Ukraine).

The experimental data utilised in the research of this article were acquired with the help of:

– own implemented computerised weather station [48];

– Metos by Pessl Instruments weather station using the FieldClimate IoT platform, access to which was provided by Metos Ukraine LLC.

2.2 Model limitations

The developed computer-oriented model based on XAI takes into account the following factors and parameters:

– crop type: corn;

– diagnosed diseases: Fusarium Head Blight, Leaf Blight Helminthosporium turcicum (Southern Corn Leaf Blight, Northern Corn Leaf Blight);

– agroclimatic zone of experimental data acquisition: the northern steppe of Ukraine (arid and warm zone; hydrothermal coefficient ranges from 0.7 to 1.0; typical annual temperature sum ranges from 2900 °C to 3300 °C);

– informative soil and climatic parameters: air temperature, air relative humidity, precipitation and time of leaf wetness of agricultural crops.

2.3 The structure of the computer-oriented model

The generalised structure of the computer-based model developed for the predictive analysis of the probability of disease occurrence in corn is presented in this Subsection (see Figure 2). The ANFIS technology was used as the basis for the construction of the computer-oriented model under study. The decision was made due to the specific details of the problem being studied. Namely, there is a significant number of input parameters such as air temperature, air humidity, precipitation, and time of leaf wetness. The regularity of the influence of these parameters on the output function (i.e., the probability of the occurrence of the disease) is based on empirical evidence drawn from many years of expert experience in the field of cereal crops. Additionally, the choice of ANFIS technology is influenced by the specifics of the research problem in terms of mathematical substantiation. It involves multi-parameter dynamic nonlinear regression analysis with data extrapolation.

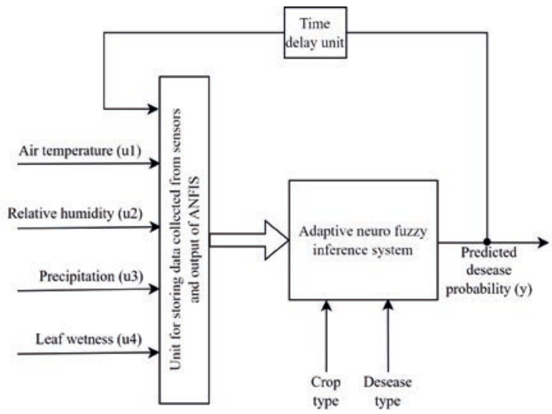


Figure 2. Block diagram of a computer-oriented model of predictive analysis

The symbols in Figure 2:

1. Air temperature (u_1): this parameter, expressed in degrees Celsius, reflects the prevailing thermal conditions in the corn-growing environment. It is a key factor that determines the growth and reproduction of fungi, thereby having a significant impact on the likelihood of disease.

2. Relative humidity (u_2): measured as a percentage, the relative humidity parameter determines the moisture content of the atmosphere. It is generally recognised that high humidity levels promote the development and spread of fungal infections, making this parameter informative in ANFIS technology studies.

3. Precipitation (u_3): measured in millimetres, it reflects the amount of moisture. Excessive precipitation can lead to prolonged wetting of agricultural crops, which creates a favourable environment for the development of pathogens.

4. Leaf wetness time (u_4): expressed in minutes, leaf wetness is a direct indicator of the length of time that leaves remain wet. Longer periods of leaf moisture increase the likelihood of disease, which determines the importance of taking this parameter into account during ANFIS technology studies.

In addition, the model includes a feedback mechanism through the output signal of ANFIS (y), which means the probability of occurrence of disease in percentage terms. This mechanism is included in the model based on the fact that the probability of occurrence of disease in crops depends not only on the momentary values of the input physicochemical parameters but also on the preceding values of the input and output model parameters within a specific timeframe. Incorporating feedback allows the model to adapt and improve predictions by utilising previous data, thereby enhancing its overall predictive capabilities.

By taking into account the main environmental factors (air temperature and relative humidity, precipitation and leaf wetness time), as well as using the adaptive capabilities of the ANFIS architecture, this model is potentially suitable for accounting for the intricate relationship between these variables and the probability of occurrence of specific types of diseases in specific types of crops. Hence, this makes it possible to develop and research hardware and software solutions for detecting the probability of occurrence of diseases, taking into account the diversity of cultivated crops, their growing cycles and the intricacy of the interrelationships between soil and climatic parameters.

The implementation of the ANFIS technology under study will allow for an increase in the efficiency and promptness of decision-making for practitioners in the field of open-field crop production through the use of software and hardware decision-making support tools for planning agrotechnical measures for growing agricultural crops.

2.4 Algorithm of neuro-fuzzy transformation technique

Taking into account the theoretical and applied features of the research problem and prior experience in developing applied information systems based on fuzzy logic [16, 17, 49], the Takagi-Sugeno algorithm (type 1) was chosen as the baseline algorithm.

The structural-algorithmic framework of ANFIS meets the criteria of the research problem of predicting the probability of occurrence of crop diseases, which is described in this article, due to its hybrid nature, combining the strengths of fuzzy logic and neural networks. This combination enables the model to take into account the intricate relationships between environmental factors and the probability of occurrence of the disease underlying the model. The architecture optimises its parameters through iterative learning, ensuring that the output of the system agrees with accurate estimates of the probability of disease occurrence.

The following physical variables were selected as input variables in the respective ranges (see Section 2.4 for details): air temperature (u_1 – from -10.9 °C to 35.6 °C), relative humidity (u_2 – from 27 % to 99 %), precipitation (u_3 – from 0 mm to 5.7 mm), leaf wetness time (u_4 – from 0 min to 60 min), probability of disease occurrence (y – from 0 % to 100 %).

The next step is the process of fuzzification, the process of converting crisp (numerical) input values into fuzzy sets. Each input characteristic x_i is associated with a set of membership functions (MFs) that determine the degree of membership of the input data in each fuzzy set. In the present study, Gaussian membership functions are used.

A fuzzy rule base defines the relationship between fuzzy input sets and output sets. Each fuzzy rule corresponds to a combination of fuzzy input feature sets. For a Sugeno-type ANFIS, the rule has the form: IF x_1 is A_1 AND x_2 is A_2 AND ... AND x_n is A_n THEN $y = f(x_1, x_2, ..., x_n)$, where A_1, A_2, ... A_n – are the selected membership functions for each input feature, and $y = f(x_1, x_2, ..., x_n)$ – represents a linear function of the inputs.

For each input combination, the membership values of the fuzzy sets in the antecedent (IF) part of the rules are determined based on the fuzzified inputs. This indicates how much each rule is 'activated' by the input data.

The activated rules contribute to the final output through a weighted average. The weights are calculated based on the rule's activation strength and the linear function defined in the consequent (THEN) part of the rule.

The weighted values obtained from the previous step are then normalised to ensure that the sum of the weights equals 1. This step ensures that the output of the system remains interpretable and consistent.

The normalised weighted outputs from all the rules are summed up to produce the final output of the ANFIS system. The fuzzy result is converted back to a crisp value to facilitate interpretation. In this study, the 'wtaver' was used for defuzzification, which is the weighted average of all rule outputs.

The model training process involves the use of previous data on the probability of occurrence of a specific disease and relevant environmental conditions. Through in-depth learning, the ANFIS system adjusts its membership functions and rules to reflect the intricate interactions between the input data and the probability of occurrence of the disease in the output. In this article, the following learning algorithm was used: Hybrid, Clustering method – Subtractive Clustering.

2.5 Experimental data

In this Subsection, a detailed overview of the experimental data collected and utilised to train and validate the proposed computer-oriented model based on XAI for predicting the probability of diseases occurrence during corn cultivation is presented. The data covers the period from 12 September 2022 to 31 July 2023 and includes the main input parameters such as air temperature, relative humidity, precipitation and leaf wetness time. The data sampling time is 60 minutes. In addition, the probabilities of infection with the main types of corn diseases: Fusarium Head Blight, Helminthosporium turcicum Leaf Blight (Southern Corn Leaf Blight, Northern Corn Leaf Blight) were carefully recorded and analysed. These probabilities were determined by practitioners with expertise in open-field agriculture through a combination of field observations and laboratory analyses, which created a broad data set for model validation. The data presented in Figure 3 covers the change of seasons and environmental dynamics that play a key role in shaping the probability of diseases occurrence. The selected timeframe provides a thorough understanding of the correlation of climatic factors and their relationship to disease occurrence at different stages of corn growth.

Correlation matrices (see Figures 4, 6, 8) and scatter plots (see Figures 5, 7, and 9) were created to identify intricate relationships between input parameters and disease probabilities. These visual representations serve as tools to identify potential correlations and trends in the data.

By analysing the correlation matrices for Fusarium Head Blight and Helminthosporium turcicum Leaf Blight, the environmental factors that predominantly contribute to the disease were identified. Strong correlations can identify leading indicators of disease probability, which can help determine agricultural management strategies. Thus, from Figure 4, it can be seen that the correlation coefficient between air temperature and the probability of occurrence of the disease is negative. While for other parameters, the coefficients are positive. This indicates that with increasing temperature, the probability of occurrence of the disease decreases. From Figures 6 and 8 it can be seen that all of the aforementioned parameters have a positive correlation coefficient with Southern Corn Leaf Blight and Northern Corn Leaf Blight.

Scatter plots provided the authors with a visual representation of the relationship between all possible variables, where each data point represents a particular dimension in the data set. These results were used to solve the model identification problem (see Subsection 3.1).

The occurrence and development of Fusarium Head Blight is highly dependent on climatic conditions during flowering. If the surrounding environment is warm and humid, the risk of spore occur-

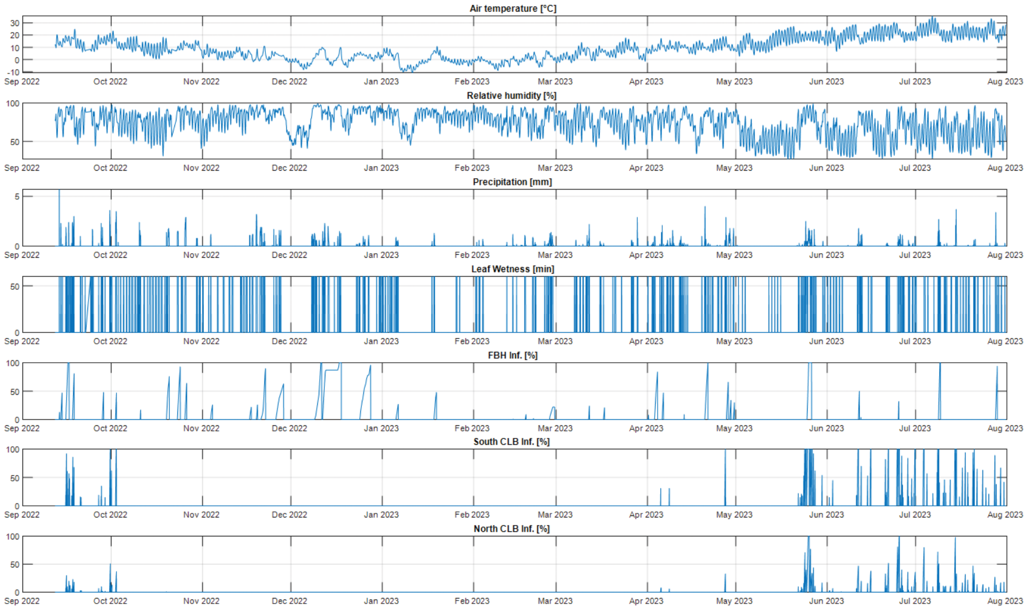


Figure 3. Measurement data for Fusarium Head Blight and Leaf Blight Helminthosporium turcicum (Southern Corn Leaf Blight, Northern Corn Leaf Blight)

rence and further development of the corresponding corn disease is significantly increased [50]. Southern Corn Leaf Blight is a widespread disease of corn worldwide. Precipitation, air temperature and relative humidity are crucial for the occurrence and development of the disease. An environment with high temperatures (from 20 °C to 32 °C) and high levels of relative humidity are factors that increase the risk of the disease occurrence. High risks of Northern Corn Leaf Blight are observed during moderate air temperatures (from 18 °C to 29 °C), high relative humidity and significant dew during the corn vegetation period. Dry and sunny climatic conditions reduce the risk of Southern Corn Leaf Blight and Northern Corn Leaf Blight [51].

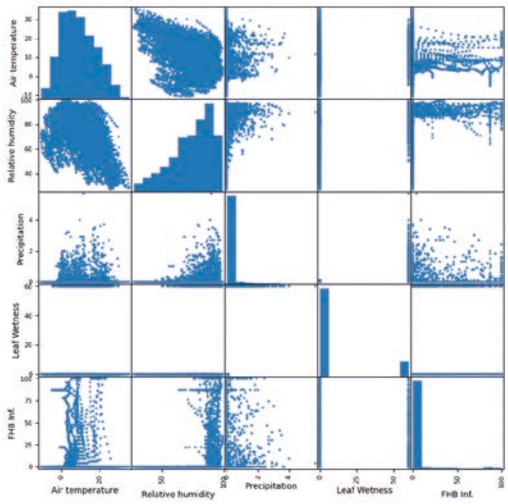


Figure 5. Scatter plot for FHB

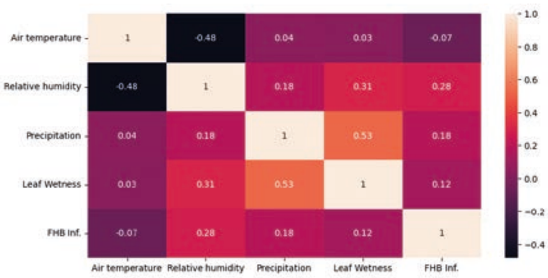


Figure 4. Correlation matrix for FHB

The experimental data form the foundation for the creation of training and validation datasets for training and testing the predictive capabilities of the proposed and developed computer-oriented model based on XAI. For Fusarium Head Blight, the training set was selected from 12 September 2022 to 31 May 2023. The validation set is from 01 June 2023 to 31 July 2023. For the two types of Leaf Blight Helminthosporium turcicum, the training set was selected from 12 September 2022 to 30 June 2023. The validation set is from 01 July 2023 to 31

July 2023. These data groups were selected based on a preliminary statistical analysis to maximise the consideration of the seasonality factor in the model development.

The adaptability and accuracy of the model can be assessed by correlating the input parameters with the probability of disease occurrence and examining the associated trends. Such data-driven validation ensures the relevance of the model and its applicability in real-world scenarios, thus increasing its reliability as a tool for predicting the probability of occurrence of diseases in corn cultivation.

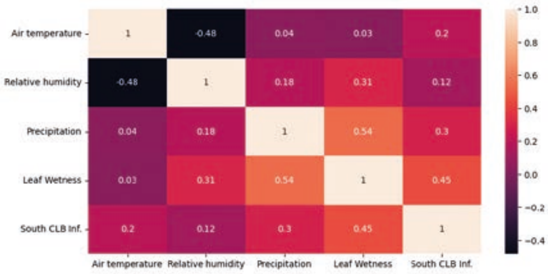


Figure 6. Correlation matrix for South CLB

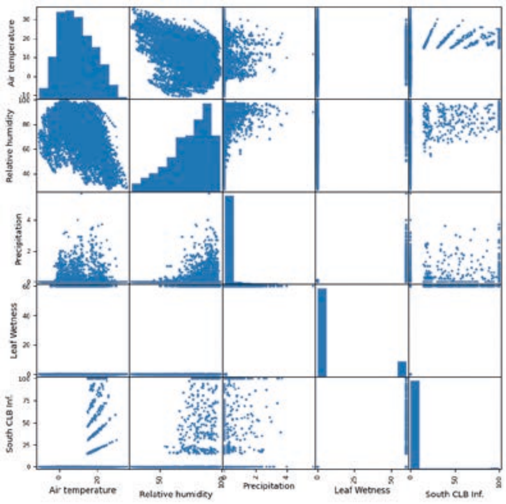


Figure 7. Scatter plot for South CLB

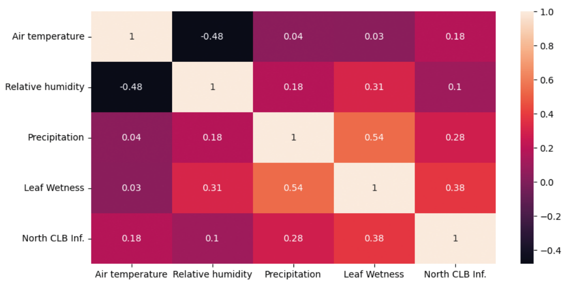


Figure 8. Correlation matrix for North CLB

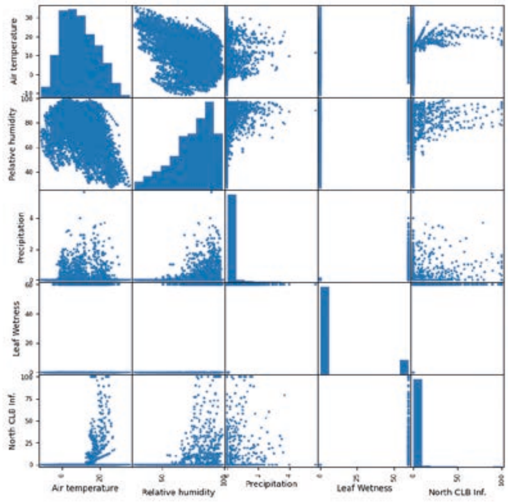


Figure 9. Scatter plot for North CLB

The resulting visualisations, including scatter plots and correlation matrices, illustrate how the probability of disease occurrence is affected by changes in each input variable. In this way, the influence of each factor can be intuitively understood by the authors. It should also be noted that enriching the dataset with additional raw data, as mentioned in the discussion Section, would further enhance the overall understanding of these relationships.

Therefore, research on the development of ANFIS technology is imperative and should be based on the principle of decomposition of the research task.

The process involves identifying informative parameters and the corresponding regression model, creating a simulation model in a computer modelling environment, training the identified model, validating the results of the computer experiment, and performing qualitative and quantitative analysis of the results. Finally, it involves identifying promising areas for future research.

3 Research results

3.1 System identification

The main results of the nonlinear model identification of the dynamic system based on experimental data (Subsection 2.5) using an adaptive neuro-fuzzy inference system (ANFIS) are presented in this Subsection.

In order to use ANFIS for identification, it was first determined which variables should be used as input arguments for each disease type separately. In essence, a 'black box' modelling is performed: a case where it is not possible to obtain an exact mathematical representation of the system for physical reasons, and hence the shape of the model and its numerical characteristics are extracted from the data. In this case, the mathematical basis of the developed identification model is the approach to the identification of nonlinear systems [52], taking into account the results of a priori correlation analysis of the data (see Subsection 2.5). For Fusarium Head Blight, 23 input factors were used in this article based on preliminary analysis: *y(k-1), y(k-2), y(k-3), y(k-6), y(k-7), y(k-8), y(k-9),* $u_1(k-1)$, $u_1(k-2)$, $u_1(k-3)$, $u_1(k-4)$, $u_2(k-1)$, $u_2(k-2)$, $u_2(k-3)$, $u_3(k-1)$, $u_3(k-2)$, $u_3(k-3)$, $u_4(k-1)$, $u_4(k-2)$, $u_4(k-3)$, $u_4(k-7)$, $u_4(k-8)$, $u_4(k-9)$. The parameter *y(k)* was used as the output of the system. These parameters include: the output function with a recency in the range from 1 to 9 hours; air temperature with a recency in the range from 1 to 4 hours; relative humidity with a recency in the range from 1 to 3 hours; precipitation with a recency in the range from 1 to 3 hours; and leaf wetness time with a recency in the range from 1 to 9 hours. Such attributes of the identification model were established through a priori qualitative analysis of experimental data (see Figure 3) and test computer experiments of the system identification model.

Based on the computer identification experiment, the number of model inputs was chosen to be seven (two inputs for y, two inputs for u_4, and one for u_1, u_2 and u_3). Afterwards, a sequential forward search was performed on the inputs. In the process, combinations of input variables are successively selected to minimise the root mean square error (RMSE). A graph for all combinations of inputs for the training and validation samples sorted by decreasing RMSE is shown in Figure 10.

According to the search results *y(k-1), y(k-6),* $u_1(k-4)$, $u_2(k-1)$, $u_3(k-3)$, $u_4(k-1)$ and $u_4(k-7)$ were selected as inputs since the model with these inputs has the lowest training RMSE of ±4.33 % and validation RMSE of ±5.32 %.

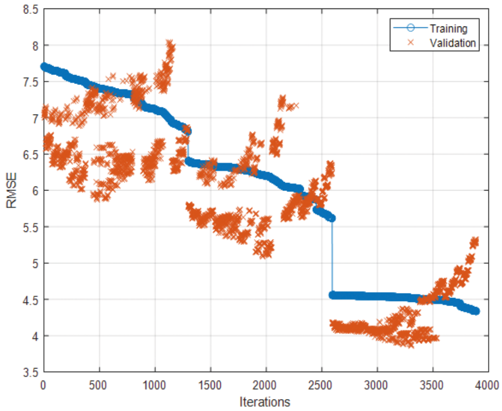


Figure 10. Error for corresponding combinations of inputs for FHB

The next step in model identification was to determine the training parameters. Subtractive clustering is a fast, one-pass algorithm for estimating the number of clusters and cluster centres in a dataset. The cluster influence range was set to 0.5. This value indicates the range of influence of the cluster when the data space is considered as a unit hypercube. Specifying a small cluster radius typically results in the creation of many small clusters in the data. This, in turn, leads to the generation of a FIS with a large number of rules. These numerous rules might make it easier to explain the behaviour of the model but could also make it more complex and harder in terms of parametrisation and analisys. On the other hand, using a large cluster radius creates fewer, larger clusters in the data, resulting in fewer rules in the FIS. This can make the model more computationally efficient but might reduce its sensitivity and interpretability of the output function because the resulting rules are more generalised. Thus, a balance between interpretability and model performance should be sought.

An important advantage of using the clustering method for rule discovery is that the resulting rules are better suited to the input data than rules generated without clustering. This adjustment reduces the total number of rules when the input data is high-dimensional. This approach was used on the basis that the implemented model should be further integrated into low-cost microcontroller devices in the form of embedded software, which corresponds to the applied principles of fog and edge computing.

To improve the FIS performance, the system was optimised using the anfis function. A training period of 100 epochs was used (see Figure 11). Also, during this experiment, training with validation sampling was immediately set up to detect overfitting when the validation error starts to increase while the training error continues to decrease (see Figure 12).

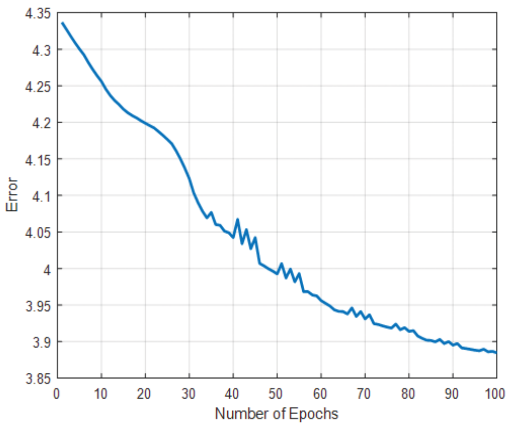


Figure 11. Training error for FHB

In Figure 11 it is shown that the learning error continues to decrease even at epoch 100. However, this tendency is insignificant from epoch 90 onwards.

In Figure 12 it is shown that the lowest value of the data validation error occurs at epoch 36. After this point, it increases slightly, even though anfis continues to minimise the error compared to the training data. This pattern is a sign of overfitting. Depending on the specified accuracy to error, the validation error plot can also indicate the ability of the model to generalise the test data. Based on the training results, the system whose settings correspond to the 36th training epoch was selected, with a training RMSE of ± 4.06 % and a validation RMSE of ± 4.50 %. The performance of the model shows a slight improvement on the training data, and slightly better on the validation data.

For Southern Corn Leaf Blight, a limit of 21 input factors was set: $y(k\text{-}1)$, $y(k\text{-}2)$, $y(k\text{-}3)$, $y(k\text{-}6)$, $y(k\text{-}7)$, $y(k\text{-}9)$, $u_1(k\text{-}1)$, $u_1(k\text{-}3)$, $u_1(k\text{-}4)$, $u_2(k\text{-}1)$, $u_2(k\text{-}2)$, $u_2(k\text{-}3)$, $u_3(k\text{-}1)$, $u_3(k\text{-}2)$, $u_3(k\text{-}3)$, $u_4(k\text{-}1)$, $u_4(k\text{-}2)$, $u_4(k\text{-}3)$, $u_4(k\text{-}7)$, $u_4(k\text{-}8)$, $u_4(k\text{-}9)$. The output of the system is $y(k)$. A plot for all combinations of inputs for the training and validation samples sorted by decreasing training RMSE is shown in Figure 13.

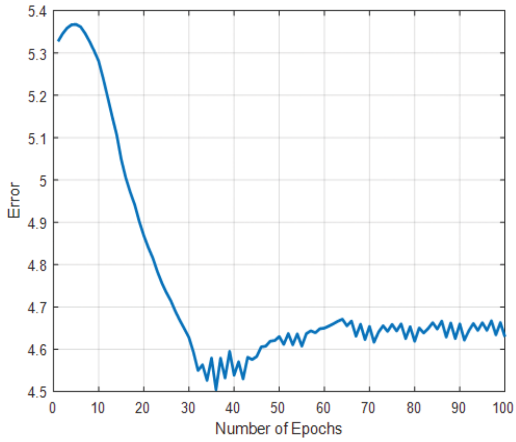


Figure 12. Validation error for FHB

Based on the search results, $y(k\text{-}1)$, $y(k\text{-}9)$, $u_1(k\text{-}4)$, $u_2(k\text{-}3)$, $u_3(k\text{-}1)$, $u_4(k\text{-}1)$ and $u_4(k\text{-}9)$ were selected as inputs, since the model with these inputs has the lowest training RMSE of ± 6.02 % and validation RMSE of ± 13.07 %.

The result of FIS performance improvement over the training period of 100 epochs is shown in Figures 14 and 15.

Based on the results of model training for Southern Corn Leaf Blight, the system corresponding to the 97th training epoch was selected, with a training RMSE of ± 5.92 % and a validation RMSE of ± 13.62 %.

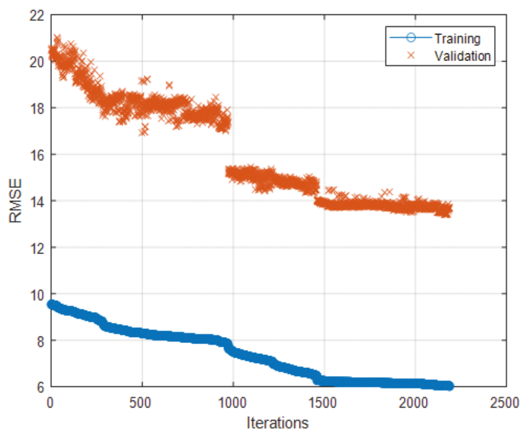


Figure 13. Error for corresponding combinations of inputs for South CLB

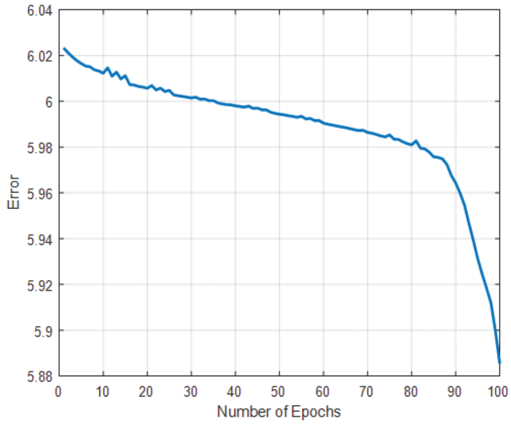


Figure 14. Training error for South CLB

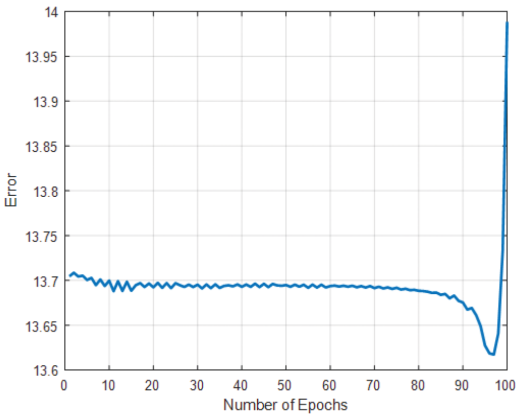


Figure 15. Validation error for South CLB

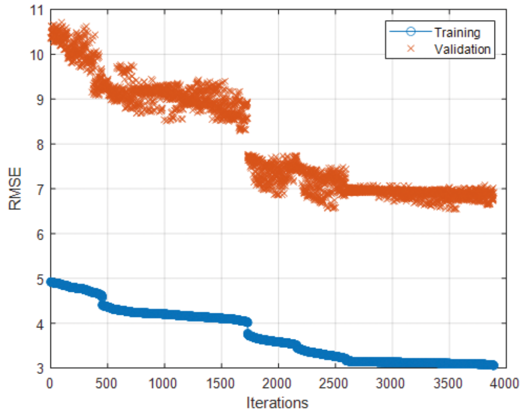


Figure 16. Error for corresponding combinations of inputs for North CLB

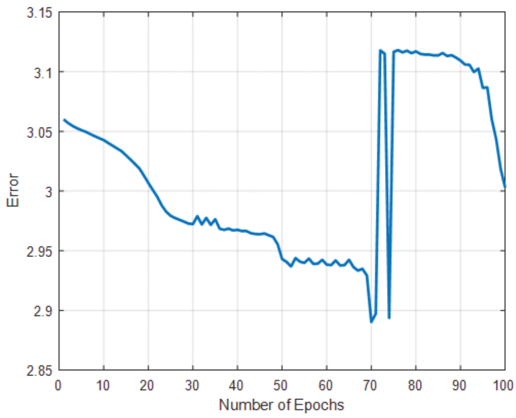


Figure 17. Training error for North CLB

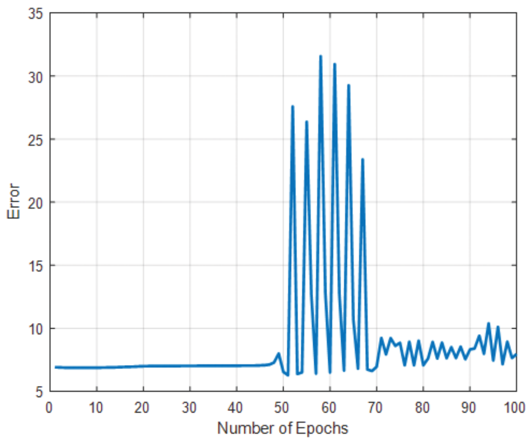


Figure 18. Validation error for North CLB

For Northern Corn Leaf Blight, the input candidates and training parameters are identical to the Fusarium Head Blight case.

In Figure 16 it is shown a plot for all combinations of inputs for the training and validation samples sorted by decreasing training RMSE.

According to the search results, $y(k\text{-}1)$, $y(k\text{-}9)$, $u_1(k\text{-}2)$, $u_2(k\text{-}3)$, $u_3(k\text{-}3)$, $u_4(k\text{-}3)$ and $u_4(k\text{-}8)$ were selected as inputs since the model with these inputs has the lowest training RMSE of ± 3.06 % and validation RMSE of ± 6.92 %.

The result of FIS performance improvement over the training period of 100 epochs is shown in Figures 17 and 18.

Based on the results of training the model for Northern Corn Leaf Blight, the system whose settings correspond to the 51st training epoch was selected, with a training RMSE of ± 2.94 % and a validation RMSE of ± 6.28 %.

Thus, the results of the identification of the model under study made it possible to refine (expand) the list of input parameters to improve the accuracy of predictive analytics of the probability of occurrence of corn diseases. In the identified model, in addition to the primary soil and climatic parameters (temperature and relative humidity, precipitation and leaf wetness), the values of the output function (probability of disease occurrence) and input soil and climatic parameters at previous points in time are used as input values, which include:

– for FHB: air temperature with a 4-hour delay, relative humidity with a 1-hour delay, precipitation with a 3-hour delay, leaf wetness time with a 1-hour and 7-hour delay, and the output function (probability of FHB disease occurrence) with a 1-hour and 6-hour delay;

– for South CLB: air temperature with a delay of 4 hours, relative humidity with a delay of 3 hours, precipitation with a delay of 1 hour, leaf wetness time with a delay of 1 hour and 9 hours, and the output function (probability of occurrence of South CLB disease occurrence) with a delay of 1 hour and 9 hours;

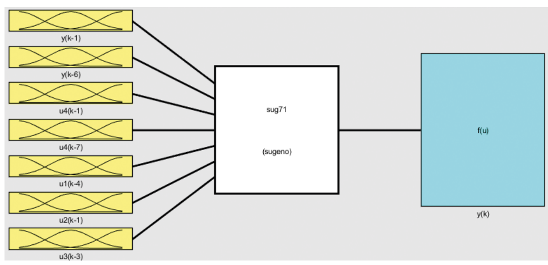
– for North CLB: air temperature with a delay of 2 hours, relative humidity with a delay of 3 hours, precipitation with a delay of 3 hours, leaf wetness time with a delay of 3 hours and 8 hours, and the output function (probability of occurrence of North CLB occurrence) with delays of 1 hour and 9 hours.

The performed feature importance analysis confirmed that all physical and chemical parameters (air temperature, relative humidity, precipitation and leaf wetness) are influential for disease prediction. During the analysis, the permutation importance technique was partially applied to identify appropriate combinations of input variables for the ANFIS model. In addition, the determination of appropriate time delays for the selected variables was emphasised. In particular, a sequential search was carried out, limiting the investigation to delays between *(k-1)* and *(k-10)*. This limitation was imposed due to the significant computational load on a typical personal computer. Consequently, further investigation in this regard remains a potential avenue of research. Thus, the specific list of informative attributes of the predictive computer-oriented model is determined by the type of diagnosed disease. This approach was implemented in the form of computer models in Matlab & Simulink (see Subsection 3.2).

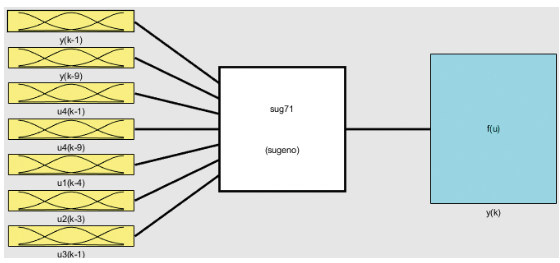
3.2 ANFIS computer models

Based on the results of the identification of informative parameters (see Subsection 3.1), the corresponding Type-1 Sugeno systems were created in the Fuzzy Logic Designer app for each type of disease (see Figure 19).

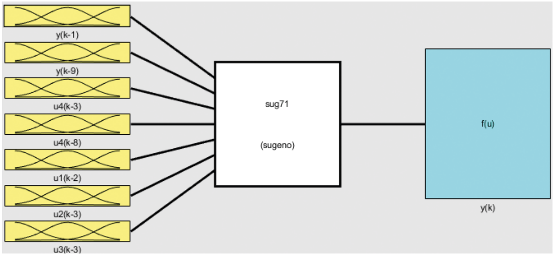
The ANFIS models (see Figure 19) assimilate four key input parameters, namely air temperature (u_1), air relative humidity (u_2), precipitation (u_3), and leaf wetness time (u_4), as well as the values of identified soil and climatic parameters and the output function (y) at previous time points. The combination of these input data with a delay of a fixed number of sampling periods (the sampling step is 60 minutes), taking into account the previous data on the probability of occurrence of a particular disease (see Section 3.1), serves as the information basis for the functioning of the developed model for detecting the probability of occurrence of a particular type of corn disease. The output of the models is the probability of occurrence of the disease (y), expressed in percentage terms. This result encapsulates the cumulative effect of the input parameters on the probability of occurrence of the disease in corn crops.



a) ANFIS model for FHB



b) ANFIS model for South CLB



c) ANFIS model for North CLB

Figure 19. Developed ANFIS models

Thus, the developed ANFIS models establish the relationship between input parameters and the probability of occurrence of diseases, providing practical information for making management decisions in agriculture.

In addition, through an iterative learning process, the ANFIS model optimises its membership functions and rules, allowing it to accurately reflect intricate relationships between inputs and outputs. This mechanism, influenced by experimental data and system identification, allows the model to refine its predictions over time, ensuring its adaptability to dynamic environmental conditions. The simulation models presented in Figure 19 serve as a functional basis for the computer experiment to validate the ANFIS model for detecting the probability of occurrence of a specific disease in corn, as described in detail in Section 3.4.

3.3 Fuzzification of input and output variables

The procedure for fuzzification of input and output variables is implemented in the Fuzzy Logic Designer environment of the Matlab & Simulink application package according to the procedure described in Subsection 2.4 'Algorithm of neuro-fuzzy transformation technique'.

A Sugeno fuzzy system was created using membership functions derived directly from the data clusters found by Subtractive clustering of the input and output data. Each input and output variable contains one membership function for each cluster. The input variables use Gaussian membership functions. The number of phasing terms in the input variables is 3 for each Fusarium Head Blight parameter, 4 for each Southern Corn Leaf Blight parameter, and 4 for each Northern Corn Leaf Blight parameter. The output variables use linear membership functions. The coefficients for the original linear function are given in Tables 4–6.

The coefficients in Tables 4–6 are associated with each cluster and are responsible for converting fuzzy input data into crisp output. The physical meaning of these coefficients is the partial sensitivity coefficients of the output function to the input parameters. In combination with the Gaussian form of the fuzzy terms, they allow for precise predictions that are tuned to the specific characteristics of the relationship between the inputs and outputs of each cluster. The inclusion of these coefficients increases prediction accuracy and improves the capacity of the model to account for intricate relationships between input parameters and the probability of occurrence of a disease.

The obtained results of fuzzification are the information basis for further research on the validation of computer models with further qualitative and quantitative analysis of the results.

3.4 Results of the computer experiment

As a result of the systematisation of the results obtained in Subsections 3.1–3.3 of the study of the intelligent software component for predicting the probability of occurrence of diseases of corn during the full cycle of its cultivation based on the ANFIS technique, a series of computer simulation tests were conducted in the Fuzzy Logic Designer environment of the Matlab & Simulink application package, in which the predicted values of the probability of disease occurrence are compared with the actual data. The results of these tests are time series plots, as shown in Figures 20, 21, 24, 25, 28, 29; comparison of actual and predicted data, as shown in Figures 22, 23, 26, 27, 30, 31, as well as a comprehensive comparison of performance indicators based on the following statistical distribution estimates: mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination (R^2), as shown in Table 7.

To evaluate the accuracy of the proposed model the following metrics were used:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|;$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2};$$

Table 4. Coefficients of the ANFIS model output function for Fusarium Head Blight

Input	Cluster1	Cluster2	Cluster3
extra coefficient	1.047	0.3066	1.03
y(k-1)	-0.01584	0.04864	-0.05155
y(k-6)	1627	3298	-1.5
u4(k-1)	-0.00605	0.005124	-0.01502
u4(k-7)	0.01749	0.008965	0.0956
u1(k-4)	0.005681	-0.0016	0.135
u2(k-1)	1.379	-4.52	0.04577
u3(k-3)	-0.4342	-0.05988	78.27

Table 5. Coefficients of the ANFIS model output function for Southern Corn Leaf Blight

Input	Cluster1	Cluster2	Cluster3	Cluster4
extra coefficient	-1728	-9175	-0.5165	0.812
y(k-1)	-510.2	-227.7	0.01905	0.1412
y(k-9)	-1114	2172	-0.6571	-0.1249
u4(k-1)	42.47	-42.56	-0.06345	0.09056
u4(k-9)	0.009674	0.05373	0.1225	0.2887
u1(k-4)	-0.0008521	0.03742	0.02445	-0.007766
u2(k-3)	-2.149	55.35	-19.39	0.4443
u3(k-1)	0.06405	-2.612	1.212	6.292

Table 6. Coefficients of the ANFIS model output function for Northern Corn Leaf Blight

Input	Cluster1	Cluster2	Cluster3	Cluster4
extra coefficient	1.253	1.408	0.2345	0.8651
y(k-1)	0.02065	0.02942	0.0503	0.1478
y(k-9)	-2.406	93.65	-0.05027	-0.04665
u4(k-3)	-553.7	-14.47	-0.02527	0.008054
u4(k-8)	0.001263	0.02038	0.07957	0.06213
u1(k-2)	0.0001716	0.01448	0.006851	0.01
u2(k-3)	0.3012	4.254	10.77	-0.5421
u3(k-3)	-0.01968	-1.009	0.4468	1.799

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2},$$

where y_i is the observed data, $\hat{y}_i$ is the predicted data and $\bar{y}$ is the mean value of the observed data.

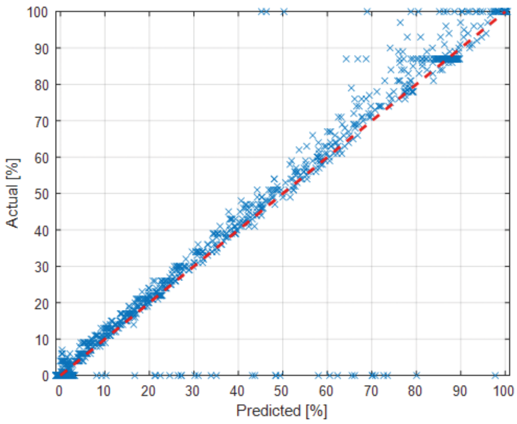


Figure 22. FHB: Actual disease % vs predicted disease % for training data

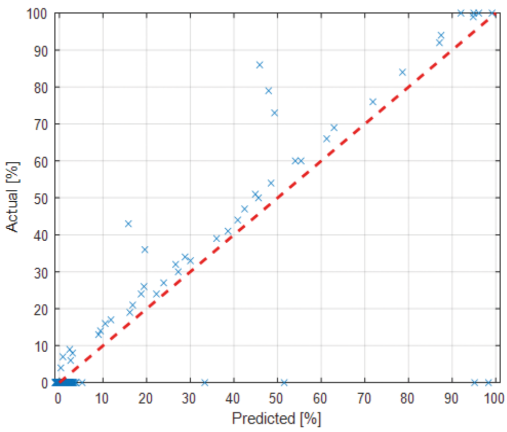


Figure 23. FHB: Actual disease % vs predicted disease % for validation data

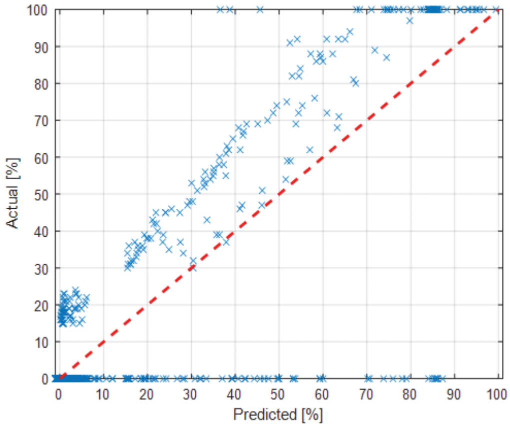


Figure 26. South CLB: Actual disease % vs predicted disease % for training data

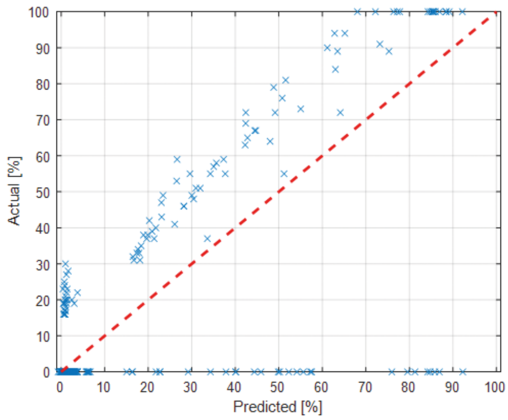


Figure 27. South CLB: Actual disease % vs predicted disease % for validation data

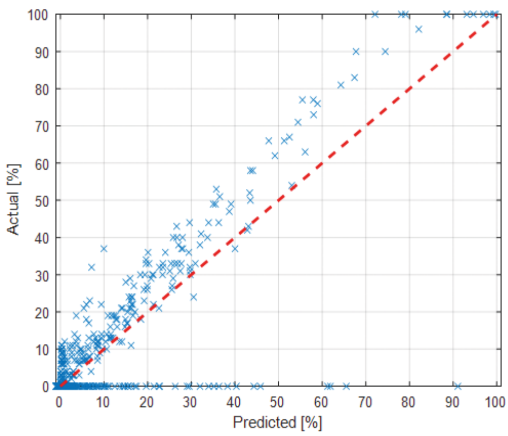


Figure 30. North CLB: Actual disease % vs predicted disease % for training data

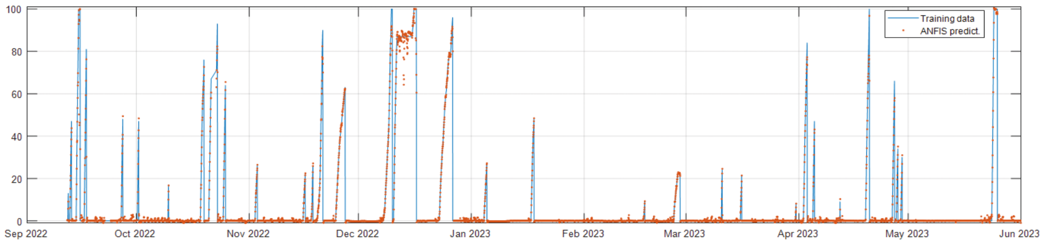


Figure 20. FHB: Training data (solid), ANFIS Prediction (dots)

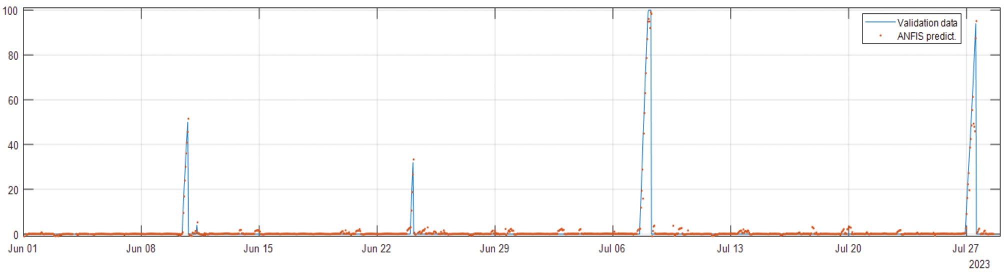


Figure 21. FHB: Validation data (solid), ANFIS Prediction (dots)

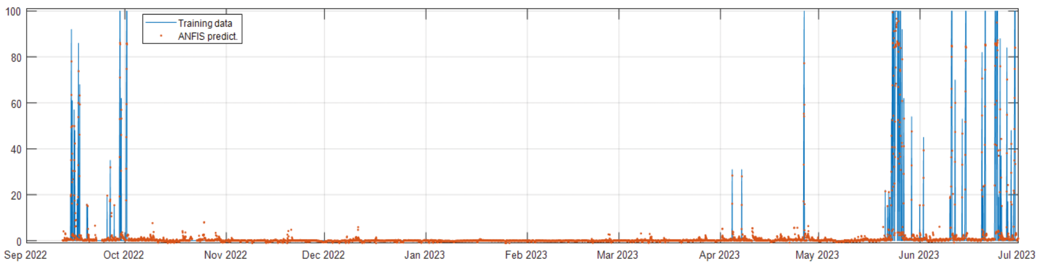


Figure 24. South CLB: Training data (solid), ANFIS Prediction (dots)

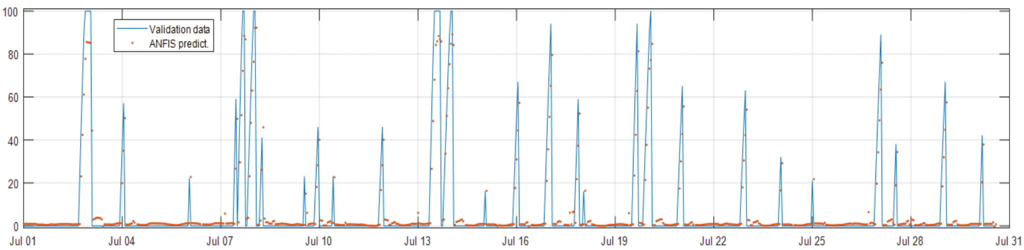


Figure 25. South CLB: Validation data (solid), ANFIS Prediction (dots)

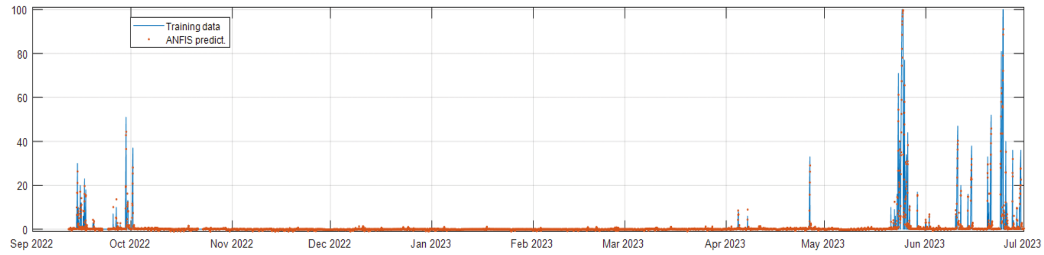


Figure 28. North CLB: Training data (solid), ANFIS Prediction (dots)

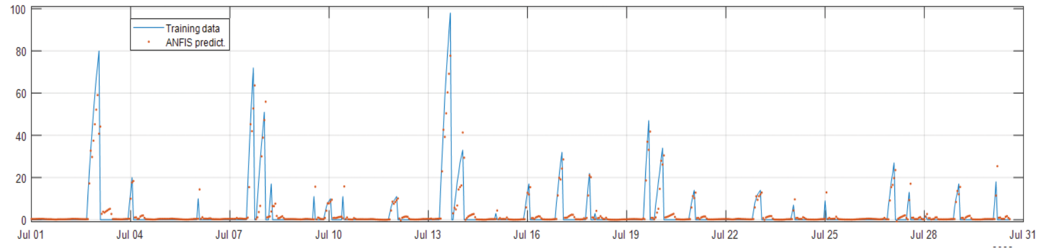


Figure 29. North CLB: Validation data (solid), ANFIS Prediction (dots)

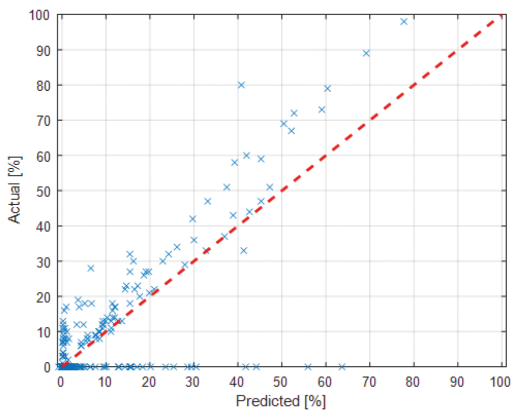


Figure 31. North CLB: Actual disease % vs predicted disease % for validation data

The results of a computer experiment demonstrate the potential of the model to predict the probability of diseases occurrence in corn cultivation. By comparing predicted and actual data, analysing performance indicators, and identifying time dynamics, this Section demonstrates the ability of the ANFIS-based model to predict disease probability and proactively manage agriculture.

Thus, through a series of computer experiments, the effectiveness of building software and hardware for predicting corn diseases based on precision monitoring of soil and climatic parameters with their subsequent transformation using the ANFIS technique was proved. It was also found that in order to accurately and efficiently detect the probability of occurrence of corn diseases, it is necessary to use a specific set of informative input attributes for each type of disease, including consideration of their previous dynamics, as described in detail in Subsection 3.1.

Table 7. Performance Metrics Comparison for Various Datasets

Metric	Type of dataset	FHB	South CLB	North CLB
MAE, %	training	0.76	1.36	0.56
	validation	0.72	5.57	2.3
RMSE, %	training	±4.06	±5.92	±2.94
	validation	±4.5	±13.62	±6.28
R^2	training	0.96	0.75	0.8
	validation	0.8	0.61	0.7

The proposed scientific and technical invention allows software implementation of a system of precision analysis of measurement data with real-time decision-making support, taking into account a significant number of functional interrelationships of soil and climatic factors of a wide range of crops grown in different periods of their vegetation.

This, in turn, makes it possible to integrate the proposed algorithm into low-cost microcontroller devices in the form of embedded software that meets the applied principles of fog and edge computing.

4 Discussion and suggestions for future investigations

The main scientific and practical effect of the research results is the development of IoT technologies for agrotechnical monitoring through the development of the computer-oriented model based on the ANFIS technique and the synthesis of structural and algorithmic software for identifying and predicting the probability of occurrence of corn diseases (Fusarium Head Blight, Southern Corn Leaf Blight, Northern Corn Leaf Blight) during the full cycle of its cultivation. This model can be used as an embedded IoT monitoring software, which allows for real-time decision-making support for practitioners in the field of open-field crop production. The physical principle of the developed model is based on the intelligent analysis of soil and climatic data (air temperature in the previous time intervals with a delay from 1 to 4 hours, relative humidity and precipitation in the previous time intervals with a delay from 1 to 3 hours, leaf wetness time in the previous time intervals with a delay from 1 to 9 hours, according to the type of diagnosed disease) coming from the corresponding distributed sensors of physical and chemical quantities. It is also worth emphasising that the developed model assumes the use of an output function (the probability of a particular type of disease occurrence) in previous time intervals with a delay from 1 to 9 hours, depending on the type of diagnosed disease, as an input argument. Discussing the impact of missing inputs, it should be noted that the model performance may degrade if one or two critical inputs are missing. This emphasises the importance of data completeness and the potential consequences of data gaps in agrotechnical monitoring systems.

This method differs from the well-known approaches based on image recognition, as it is designed to predict the probability of diseases occurrence and prevent them. Image recognition can only diagnose and classify diseases that have already emerged in agriculture. Thus, the integration of the developed model in the form of embedded software into the currently used computerised agroclimatic weather stations allows them to significantly expand their functionality by implementing decision-making support for planning agrotechnical procedures during the full cycle of growing various types of crops. Typical agrotechnical procedures that can be planned and implemented on the basis of the developed model are: optimal planning of sowing time, selection of the optimal types of fertilisers, selection of the optimal time of fertilisation, soil irrigation, etc.

Promising priority areas for further research on this information technology are: additional research to expand the dataset and include a wider range of crops and types of diagnosed diseases; consideration of fertilisation practices, irrigation strategies and other agronomic measures to obtain a holistic view of the dynamics of disease probability; long-term experimental tests in real-world conditions in different climatic zones; comprehensive technical and economic assessment of the investment attractiveness of implementing the solutions. The most informative attributes that require priority research to improve the developed model by introducing additional weighting factors when assessing the final probability of occurrence of crop diseases are:

- taking into account the currently implemented agrotechnical measures: spatial isolation of crops, liming of soils and crop rotation;
- taking into account the physical and mechanical measures implemented, such as soil warming before sowing;
- taking into account the chemical pesticides used in the cultivation of cereal crops.

In addition, the priority areas include those related to the study of the influence of input physicochemical values on the reliability of plant disease detection at the quantitative level:

- partial dependency plots can be used to show the effect of individual input signals on the prediction while keeping other inputs constant. This can provide a clear view of how changes in specific environmental factors affect the probability of disease occurrence;

– sensitivity analysis can be used to assess the robustness of the model to variations in input signals. This can involve perturbing input values within a certain range and observing the response of the model to analyse how sensitive or insensitive the model is to variations in each input signal. Moreover, it can be used to determine which input variables have the most significant impact on the prediction result;

– in addition to quantitative analysis, qualitative insights into domain knowledge can be gained by explaining why certain input signals may have a greater influence on disease prediction based on agronomic or environmental principles. This can add depth to the understanding of the behaviour of the model;

– case studies or real-world examples that demonstrate the practical implications of the predictions of the model can be conducted to show how changes in specific input signals have resulted in different disease outcomes in actual agricultural scenarios.

Research in the above areas will improve the adequacy, adaptability and scale of crop disease diagnosis, which will have a positive impact on the investment attractiveness and long-term sustainability of the agricultural sector by increasing the innovation component of the agricultural sector.

5 Conclusions

An important problem of developing scientific and applied provisions for improving the efficiency of computer-oriented agrotechnical monitoring systems by developing the intelligent software component for predicting the probability of disease occurrence in corn during the full cycle of its cultivation based on the ANFIS as the explainable AI technique has been solved in this article.

The main results of the article are: a review, critical analysis and systematisation of modern scientific and applied achievements in the field of technology for intelligent analysis of time series of measurement monitoring results; a computer-oriented data processing model based on the ANFIS technique, which, unlike the known ones, takes into account the simultaneous influence of a set of soil and climatic parameters of crop production enterprises, their previous dynamics, as well as the prehistory of the output parameter (probability of disease occurrence), which allowed to detect the probability of occurrence of such diseases in corn with the value of the coefficient of determination on the validation data: Fusarium Head Blight – 0.8, Southern Corn Leaf Blight – 0.61 and Northern Corn Leaf Blight – 0.7; developed and studied a model for identifying computer model parameters that takes into account possible combinations of input and output data with a delay of a fixed number of sampling periods, which allowed to establish a set of the most influential parameters with the lowest error values of 4.33 % on the training sample for Fusarium Head Blight, 6.02 % for Southern Corn Leaf Blight and 5.92 % for Northern Corn Leaf Blight; computer models were developed in Matlab and Simulink, which allowed a series of computer experiments where the predicted values of the probability of disease occurrence are compared with actual data to be carried out. The models are dynamic and can adapt to changing environmental conditions. Furthermore, the models could be scaled to the input data set.

Finally, a set of promising areas of research to improve the adequacy, adaptability and scalability of the developed computer-oriented model using XAI for diagnosing crop diseases has been substantiated in the article.

Acknowledgements

The authors would like to thank the company Metos Ukraine LLC for the information support provided. We also gratefully acknowledge the funding support by the project ARTIQ - AI Centers of Excellence ARTIQ/0004/2021.

6 List of abbreviations

ANFIS	– Adaptive neuro-fuzzy inference system
FAO	– Food and Agricultural Organization
FHB	– Fusarium Head Blight
FIS	– Fuzzy inference system
FL	– Fuzzy logic
GA	– Genetic algorithm
IoT	– Internet of Things

MAE	– Mean absolute error
North CLB	– Northern Corn Leaf Blight
RMSE	– Root mean square error
South CLB	– Southern Corn Leaf Blight
XAI	– Explainable artificial intelligence

References

[1] Food and Agriculture Organization of the United Nations: Climate change and food security: risks and responses. Available at: fao.org/3/i5188e/I5188E.pdf [Accessed 14 June 2023].

[2] R. Lopez-Lozano, B. Baruth, An evaluation framework to build a cost-efficient crop monitoring system. Experiences from the extension of the European crop monitoring system, Agric. Systems, 168, 2019, pp. 231–246. doi.org/10.1016/j.agsy.2018.04.002.

[3] FAOSTAT: Crop and livestock products. Available at: fao.org/faostat/en/#data/QCL [Accessed 18 June 2023].

[4] FAO: Agricultural production statistics 2000 – 2021. Available at: fao.org/3/cc3751en/cc3751en.pdf [Accessed 20 June 2023].

[5] M.F. Mushtaq, U. Akram, M. Aamir, H. Ali, M. Zulqarnain, Neural Network Techniques for Time Series Prediction: A Review, JOIV: International Journal on Informatics Visualization, 3 (3), 2019, pp. 314–320. dx.doi.org/10.30630/joiv.3.3.281.

[6] E. Funes, Y. Allouche, G. Beltrán, A. Jiménez, A Review: Artificial Neural Networks as Tool for Control Food Industry Process, Journal of Sensor Tech., 5, 2015, pp. 28–43. doi.org/10.4236/jst.2015.51004.

[7] K. Suzuki (Ed.), Artificial Neural Networks – Industrial and Control Engineering Applications, InTech, 2011. doi.org/10.5772/2041.

[8] S. Kujawa, G. Niedbała (Ed.), Artificial Neural Networks in Agriculture, MDPI, 2021. doi.org/10.3390/books978-3-0365-1579-3.

[9] S. Francik, Z. Slipek, J. Fraczek, A. Knapczyk, Present Trends in Research on Application of Artificial Neural Networks in Agricultural Engineering, Agricultural Engineering, 20 (4), 2017, pp. 15–25. doi.org/10.1515/agriceng-2016-0060.

[10] L. Rutkowski, Generalized Regression Neural Networks in Time-Varying Environment, IEEE Transactions on Neural Networks, 15 (3), 2004, pp. 576–596. doi.org/10.1109/TNN.2004.826127.

[11] L. Rutkowski, Adaptive Probabilistic Neural Networks for Pattern Classification in Time-Varying Environment, IEEE Transactions on Neural Networks, 15 (4), 2004, pp. 811–827. doi.org/10.1109/TNN.2004.828757.

[12] M. Pérez-Pons, J. Parra-Dominguez, S. Omatu, E. Herrera-Viedma, J. Corchado, Machine Learning and Traditional Econometric Models: A Systematic Mapping Study, Journal of Artificial Intelligence and Soft Computing Research, 12 (2), 2022, pp. 79–100. doi.org/10.2478/jaiscr-2022-0006.

[13] J. Lv, M. Pawlak, Bandwidth Selection for Kernel Generalized Regression Neural Networks in Identification of Hammerstein Systems, Journal of Artificial Intelligence and Soft Computing Research, 11 (3), 2021, pp. 181–194. https://doi.org/10.2478/jaiscr-2021-0011.

[14] T.J. Ross, J.M. Booker, W.J. Parkinson (Eds.), Fuzzy Logic and Probability Applications: Bridging the Cap, ASA-SIAM Series on Statistics and Applied Probability, SIAM, Philadelphia, ASA, Alexandria, VA, 2002. URL: pzs.dstu.dp.ua/logic/bibl/prob.pdf.

[15] X. Li, M.J. Er, B.S. Lim, J.H. Zhou, O.P. Gan, L. Rutkowski, Fuzzy Regression Modeling for Tool Performance Prediction and Degradation Detection, Int. J. Neural. Syst., 20 (5), 2010, pp. 405–419. doi.org/10.1142/S0129065710002498.

[16] I. Laktionov, O. Vovna, M. Kabanets, Computer-Oriented Method of Adaptive Monitoring and Control of Temperature and Humidity Mode of Greenhouse Production, Baltic J. Modern Computing, 11 (1), 2023, pp. 202–225. doi.org/10.22364/bjmc.2023.11.1.12.

[17] I. Laktionov, O. Vovna, M. Kabanets, Information Technology for Comprehensive Monitoring and Control of the Microclimate in Industrial Greenhouses Based on Fuzzy Logic, Journal of Artificial Intelligence and Soft Computing Research, 13 (1), 2023, pp. 19–35. doi.org/10.2478/jaiscr-2023-0002.

[18] L. Rutkowski, Flexible Neuro-Fuzzy Systems, Springer New York, NY, 2004. doi.org/10.1007/b115533

[19] D. Rutkowska, Neuro-Fuzzy Architectures and Hybrid Learning, Physica Heidelberg, 2002. doi.org/10.1007/978-3-7908-1802-4.

[20] D. Karaboga, E. Kaya, Adaptive network based fuzzy inference system (ANFIS) training approaches: a comprehensive survey, Artif. Intell. Rev., 52, 2019, pp. 2263–2293. doi.org/10.1007/s10462-017-9610-2.

[21] P. Dziwiński, A. Przybył, P. Trippner, J. Paszkowski, Y. Hayashi, Hardware Implementation of a Takagi-Sugeno Neuro-Fuzzy System Optimized by a Population Algorithm, Journal of Artificial Intelligence and Soft Computing Research, 11 (3), 2021, pp.243–266. doi.org/10.2478/jaiscr-2021-0015.

[22] A. Niewiadomski, M. Kacprowicz, Type-2 Fuzzy Logic Systems in Applications: Managing Data in Selective Catalytic Reduction for Air Pollution Prevention, Journal of Artificial Intelligence and Soft Computing Research, 11 (2), 2021, pp. 85–97. doi.org/10.2478/jaiscr-2021-0006.

[23] Y.K. Al-Douri, H. Hamodi, J. Lundberg, Time Series Forecasting Using a Two-Level Multi-Objective Genetic Algorithm: A Case Study of Maintenance Cost Data for Tunnel Fans, Algorithms, 11 (123), 2018, pp. 1–19. doi.org/10.3390/a11080123.

[24] I. Aouadni, A. Rebai, Decision support system based on genetic algorithm and multi-criteria satisfaction analysis (MUSA) method for measuring job satisfaction, Ann. Oper. Res., 256, 2017, pp. 3–20. doi.org/10.1007/s10479-016-2154-z.

[25] H. Han, J. Siebert, TinyML: A Systematic Review and Synthesis of Existing Research, In: 2022 Int. Conf. on Artificial Intelligence in Information and Communication, Jeju Island, Republic of Korea, 2022, pp. 269–274. doi.org/10.1109/ICAIIC54071.2022.9722636.

[26] P.P. Ray, A review on TinyML: State-of-the-art and prospects, Journal of King Saud University – Computer and Information Sciences, 34 (4), 2022, pp. 1595–1623. doi.org/10.1016/j.jksuci.2021.11.019.

[27] N. Schizas, A. Karras, C. Karras, S. Sioutas, TinyML for Ultra-Low Power AI and Large Scale IoT Deployments: A Systematic Review, Future Internet, 14 (363), 2022, pp. 1–45. doi.org/10.3390/fi14120363.

[28] G. Schwalbe, B. Finzel, A comprehensive taxonomy for explainable artificial intelligence: a systematic survey of surveys on methods and concepts, Data Mining and Knowledge Discovery, 2023, pp. 1–59. doi.org/10.1007/s10618-022-00867-8.

[29] W. Saeed, C. Omlin, Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities, Knowledge-Based Syst., 263, 2023, pp. 1–24. doi.org/10.1016/j.knosys.2023.110273.

[30] G. Vilone, L. Longo, Notions of explainability and evaluation approaches for explainable artificial intelligence, Information Fusion, 76, 2021, pp. 89–106. doi.org/10.1016/j.inffus.2021.05.009.

[31] P. Mishra, Explainable AI Recipes: Implement Solutions to Model Explainability and Interpretability with Python, Springer, 2023. doi.org/10.1007/978-1-4842-9029-3.

[32] M. Ryo, Explainable artificial intelligence and interpretable machine learning for agricultural data analysis, Artif. Intelligence in Agric., 6, 2022, pp. 257–265. doi.org/10.1016/j.aiia.2022.11.003.

[33] S. Mohr, R. Kühl, Acceptance of artificial intelligence in German agriculture: an application of the technology acceptance model and the theory of planned behavior, Precision Agric., 22, 2021, pp. 1816–1844. doi.org/10.1007/s11119-021-09814-x.

[34] N.L. Tsakiridis, T. Diamantopoulos, A.L. Symeonidis, J.B. Theocharis, A. Iossifides, P. Chatzimisios, G. Pratos, D. Kouvas, Versatile Internet of Things for Agriculture: An eXplainable AI Approach, Artificial Intelligence Applications and Innovations, 584, 2020, pp. 180–191. doi.org/10.1007/978-3-030-49186-4_16.

[35] I. Ahmed, G. Jeon, F. Piccialli, From Artificial Intelligence to Explainable Artificial Intelligence in Industry 4.0: A Survey on What, How, and Where, IEEE Transactions on Industrial Informatics, 18 (8), 2022, pp. 5031–5042. doi.org/10.1109/TII.2022.3146552.

[36] S. Rakesh, M. Indiramma, Explainable AI for Crop disease detection, In: 2022 4th Int. Conf. on Advances in Computing, Communication Control and Networking, Greater Noida, India, 2022, pp. 1601–1608. doi.org/10.1109/ICAC3N56670.2022.10074303.

[37] H. Oubehar, A. Selmani, A. Ed-Dahhak, A. Lachhab, M. El Hassane Archidi, B. Bouchikhi, ANFIS-Based Climate Controller for Computerized Greenhouse System, Adv. Sci. Technol. Eng. Syst. J., 5 (1), 2020, pp. 8–12. doi.org/10.25046/aj050102.

[38] V. Vincentdo, N. Surantha, Nutrient Film Technique-Based Hydroponic Monitoring and Controlling System Using ANFIS, Electronics, 12 (1446), 2023, pp. 1–26. doi.org/10.3390/electronics12061446.

[39] D. Bozanic, D. Tesic, D. Marinkovic, A. Milic, Modeling of neuro-fuzzy system as a support

in decision-making processes, Reports in Mechanical Engineering, 2 (1), 2021, pp. 222–234. doi.org/10.31181/rme2001021222b.

[40] N. Ziasabounchi, I.N. Askerzade, ANFIS Based Classification Model for Heart Disease Prediction, Int. J. of Engineering & Computer Science IJECS-IJENS, 14 (02), 2014, pp. 7–12, URL: ijens.org/IJECSVol14Issue02.html.

[41] S. Bellahirich, D. Mezghani, A. Mami, Design and Implementation of an Intelligent ANFIS Controller on a Raspberry Pi Nano-Computer for Photovoltaic Pumping Intended for Drip Irrigation, Energies, 14 (5217), 2021, pp. 1–19. doi.org/10.3390/en14175217.

[42] A.S. Keceli, A. Kaya, C. Catal, B. Tekinerdogan, Deep learning-based multi-task prediction system for plant disease and species detection, Ecological Informatics, 69, 2022, pp. 1–14. doi.org/10.1016/j.ecoinf.2022.101679.

[43] S.H. Lee, H. Goëau, P. Bonnet, A. Joly, Conditional Multi-Task learning for Plant Disease Identification, In: 2020 25th Int. Conf. on Pattern Recognition, Milan, Italy, 2021, pp. 3320–3327. doi.org/10.1109/ICPR48806.2021.9412643.

[44] M.S. Hema, N. Sharma, Y. Sowjanya, Ch. Santoshini, R. Sri Durga, V. Akhila, Plant disease prediction using convolutional neural network, EMITTER Int. J. of Engineering Technology, 9 (2), 2021, pp. 283–293. doi.org/10.24003/emitter.v9i2.640.

[45] T.V. Reddy, K.S. Rekha, Plant Disease Detection Using Advanced Convolutional Neural Networks with Region of Interest Awareness, Stem Cell Res Int, 6 (2), 2022, pp. 121–131. doi.org/10.47363/JIRR/2022(2)117.

[46] T. Domingues, T. Brandão, J.C. Ferreira, Machine Learning for Detection and Prediction of Crop Diseases and Pests: A Comprehensive Survey, Agriculture, 12 (1350), 2022, pp. 1–23. doi.org/10.3390/agriculture12091350.

[47] H. Bangui, S. Rakrak, S. Raghay, B. Buhnova, Moving to the Edge-Cloud-of-Things: Recent Advances and Future Research Directions, Electronics, 7 (11:309), 2018, pp. 1–31. doi.org/10.3390/electronics7110309.

[48] I.S. Laktionov, O.V. Vovna, M.M. Kabanets, H.O. Sheina, I.A. Getman, Information model of the computer-integrated technology for wireless monitoring of the state of microclimate of industrial agricultural greenhouses, Instrumentation Mesure Metrologie, 20 (6), 2021, pp. 289–300. doi.org/10.18280/i2m.200601.

[49] I. Laktionov, L. Rutkowski, O. Vovna, A. Byrski, M. Kabanets, A Novel Approach to Intelligent Monitoring of Gas Composition and Light Mode of Greenhouse Crop Growing Zone on the Basis of Fuzzy Modelling and Human-in-the-Loop Techniques, Engineering Applications of AI, 126 (B), 2023, pp. 1–21, doi.org/10.1016/j.engappai.2023.106938.

[50] Crop Science Australia: Fusarium head blight. Available at: crop.bayer.com.au/pests/diseases/fusarium-head-blight [Accessed 25 July 2023].

[51] University of Illinois: Report on plant diseases. Common leaf blights and spots of corn. Available at: ipm.illinois.edu/diseases/rpds/202.pdf [Accessed 25 July 2023].

[52] MathWorks: Nonlinear System Identification. Available at: mathworks.com/help/fuzzy/nonlinear-system-identification.html [Accessed 30 July 2023].

Ivan LAKTIONOV is a Professor in the Department of Computer Systems Software at Dnipro University of Technology (Dnipro, Ukraine). He received his Doctor of Science degree in 2021 majoring in computer systems and components. He specialises in advanced computer, sensor, infocommunication and microprocessor technologies of physical and chemical parameters monitoring and control systems. He is the author of more than 80 scientific works, including more than 25 articles in the world's authoritative publications Scopus and WoS.
https://orcid.org/0000-0001-7857-6382

Grygorii DIACHENKO is an Associate Professor in the Department of Electric Drive at Dnipro University of Technology (Dnipro, Ukraine). He received his PhD degree in 2021 majoring in electrotechnical complexes and systems. His research interests include machine learning, mechatronics, control theory and IoT. He is the author of more than 20 scientific works, including 6 articles in the world's authoritative publications Scopus and WoS.
https://orcid.org/0000-0001-9105-1951

Danuta Rutkowska is a professor of computer science. She graduated from Wrocław University of Science and Technology, Wrocław, Poland, from where she also received her Ph.D. in automation and then D.Sc. in computer science. In 2002 she was given the title of professor conferred by the President of Poland. She is an author or co-author of numerous publications, mostly in computational intelligence, including several books, book chapters, and many scientific papers. Her research interests are in the area of artificial intelligence, especially computational intelligence, focusing on artificial neural networks, fuzzy systems, genetic/evolutionary algorithms, as well as hybrid intelligent systems, i.e. neuro-fuzzy or genetic-neuro-fuzzy systems, and their applications. After working for many years at the Częstochowa University of Technology, Częstochowa, Poland, she is currently a professor in the Institute of Information Technology at the University of Social Sciences in Łódź, Poland. Since 2022 she has been serving as vice-rector for science at this university.
https://orcid.org/0000-0003-0217-2589

Marek Kisiel - Dorohinicki obtained Ph.D. in 2001 and D.Sc. in 2013 at the Department of Computer Science of the AGH University of Science and Technology in Krakow, Poland. His main research interests are metaheuristics, agent-based systems and criminal analysis. He works as a Full Professor in the Institute of Computer Science at the AGH University of Science and Technology.
https://orcid.org/0000-0002-8459-1877