

Artificial Neural Networks and Adaptive Neuro Fuzzy Inference System for wheat yield analysis and prediction

Shailesh Rao¹, Manoj Vishal¹, Anjana Krishnan²

¹Nitte Meenakshi Institute of Technology, Bangalore, India. E-mail: shailesh.rao@nmit.ac.in, vishalmanojvs@gmail.com

²CSIR. National Aerospace Laboratories, Bangalore, India. E-mail: anjana2k3@rediffmail.com

Received: 10/05/2023; Accepted: 25/08/2023.

ABSTRACT

The current study evaluated the prediction of the yield of wheat crops in the Bagalkot district of Karnataka State, India. The study aimed to provide crop yield predictions to help farmers optimize their cultivation and marketing strategies. The model used various independent variables, such as temperature, humidity of air, and water resources, to predict growth in the yield of wheat crops. The correlation analysis helps determine the strength and direction of the relationship between the variables based on the results. The statistical analysis identifies the variables that have a significant impact on crop yield growth. The work developed and tested two different models (the Artificial Neural Network (ANN) model and the Adaptive Neuro-fuzzy Interference System (ANFIS) to predict crop yield growth based on the selected independent variables. The ANFIS model was particularly interesting as it can predict a mapping between the input and output parameters, which can be useful for understanding the relationships between different variables. ANFIS was considered a better predictor than ANN as the error percentage ranged from 0-3%. Overall, the work highlighted the importance of crop yield predictions and the potential benefits that simulations can generate for farmers and the agriculture sector in general.

Keywords: Crop Yield, Correlation, ANN, ANFIS.

Redes Neurais artificiais e Sistema Adaptativo de Inferência Neuro Fuzzy para análise e previsão da produtividade do trigo

RESUMO

O presente estudo avaliou a previsão do rendimento das culturas de trigo no distrito de Bagalkot, do Estado de Karnataka, Índia. O estudo teve como objetivo fornecer previsões de rendimento das colheitas para ajudar os agricultores a otimizar suas estratégias de cultivo e comercialização. O modelo usou várias variáveis independentes tais como temperatura, humidade do ar e recursos hídricos para prever o crescimento no rendimento das culturas de trigo. O trabalho se desenvolveu e testou dois modelos diferentes: Modelo de Rede Neural Artificial (Artificial Neural Network – ANN) e Sistema de Interferência Neuro-fuzzy Adaptativo (Adaptive Neuro-fuzzy Interference System - ANFIS) a fim prever o crescimento do rendimento das culturas com base nas variáveis independentes selecionadas. O modelo ANFIS foi particularmente interessante, pois pôde prever um mapeamento entre os parâmetros de entrada e saída, os quais podem ser úteis para compreender a relação entre diferentes variáveis. ANFIS foi considerado um modelo de predição melhor que o modelo ANN, com uma percentagem de erro variando de 0-3%. De maneira geral, o trabalho destacou a importância das previsões do rendimento das culturas e os potenciais benefícios que as simulações podem gerar para os agricultores e para o setor agrícola em geral.

Palavras-chave: Rendimento da colheita, Correlação, ANN, ANFIS.



1. Introduction

Crop yield prediction, a crucial task in agriculture, plays a vital role in making informed decisions about crop management, harvest planning, and market forecasting. It helps farmers determine the best time for planting, selecting suitable crops, optimizing resource allocation, and maximizing yield. The prediction of crop yield can be done using various approaches, including statistical modeling, machine learning, and remote sensing. The yield of crops depends on several parameters, such as rainfall, temperature, humidity, soil attributes, and agricultural factors. Precision agriculture is one of the challenging problems, wherein yield prediction depends on several datasets such as climate, soil conditions, and seed quality (Beloti et al., 2017; Xu et al. 2019). Statistical modeling is used for crop yield prediction and is based on historical data on weather, soil, and crop characteristics. The statistical models can identify the relationships between these parameters and crop yield, allowing for the prediction of future yields. Machine learning models, on the other hand, can handle complex nonlinear relationships between input and output variables and have been widely used for crop yield prediction. However, these models require training from the past huge set of variable data and depend on the appropriate algorithm for developing the models, which becomes a real challenge in determining a sophisticated predictive model. Similarly, some researchers (Rahman et al., 2022) employed statistical models to forecast the production of winter wheat depending on environmental conditions. They compared the performance of three statistical models, including multiple linear regression, support vector regression, and random forest models, in predicting crop yield. The study showed that all three models achieved high accuracy in predicting crop yield, with the random forest model outperforming the other two models. The study also revealed that the most important environmental factors for crop yield prediction are precipitation, temperature, solar radiation, and soil organic matter content. Su et al. (2012) employed linear regression models to predict maize yield in China based on weather and soil data and reported that temperature, precipitation, and soil nutrients are important factors for predicting crop yield. Another study used time series analysis to predict rice yield in West Bengal, India, based on historical yield data and weather information (Bhavani, et al., 2018) and demonstrated that the autoregressive integrated moving average (ARIMA) model was effective in predicting rice yield.

Machine learning (ML) models have been widely used for crop yield prediction, as they can handle complex nonlinear relationships between input and output variables. The ML approaches are used in a wide range of applications, such as manufacturing processes,

supply chain management, agriculture sectors, and many more (Ayyappa et al., 2021; Manoj et al., 2020). The linear models make it easy to understand the relationship between the dependent and independent variables, but they fail to achieve high prediction accuracy. This is mainly due to their inability to relate the nonlinear interactions among the variables (Ansarifar et al., 2021). Over the years, researchers have employed various ML methods to address crop yield prediction. For instance, Drummond et al. (2003) utilized stepwise linear regression models to examine the relationship between yield and topographic characteristics; they minimized overfitting and predicted an optimal yield model using a neural network. Similarly, another study employed the crop yield with a random forest model, compared it to the yield achieved in the MLR method, and reported that the random forest method is effective in crop yield prediction, providing acceptable predictions in smart agriculture with reasonable data (Jeong et al., 2016).

Despite the effectiveness of the random forest method, its usage in predicting crop yield and related output is limited (Fukuda et al., 2013) as each dataset is unique, and some ML models may not be suitable for a particular system (Gromping et al., 2009). Moreover, crop yield prediction changes over time, as it behaves nonlinearly depending on several factors, such as soil characteristics, climate, the genetic makeup of the crop, and pest infestations. To overcome these challenges, several ML techniques have been proposed for crop yield prediction, including support vector machines, deep neural networks, and ensemble models, among others. Each of these models has its strengths and weaknesses, and selecting the appropriate model depends on the data available and the research objectives. A deep neural network (DNN) can capture complex nonlinear relationships between crop yield and its predictors. Ensemble models such as random forests and gradient boosting can handle high-dimensional data and noisy datasets. ML methods have shown immense potential in predicting crop yield and related outputs. However, selecting an appropriate model for a particular dataset is critical for making accurate predictions. Additionally, researchers should continually evaluate the performance of their models over time, as crop yield prediction changes with time and depends on several factors.

Of late, several models for yield production with suitable methods have been proposed by the research community. They relate to corn yield predictions with detailed correlation and regression algorithms (Sarturi et al., 2022). Predicting crop yield and enhancing crop management are two associated and crucial concepts (Russo et al., 2015). Furthermore, support vector machines (SVM) have proven to be a better method for

performing regression analysis and predicting crop yield. The correlation of a sediment yield from monsoon data in determining the accuracy of prediction was established (Misra et al., 2009). Another study examined the details of crop yield prediction using different models and their importance in the ML processes (Elavarasan et al., 2018). Another research aimed at predicting the soil organic matter content (SOM) in crop fields using artificial neural networks (ANNs) and regression models (Achieng et al., 2019) and demonstrated that the ANN model outperformed the regression models, with an R-squared value of 0.804 and a root mean squared error (RMSE) of 0.261% for the ANN model compared to an R-squared value of 0.694 and an RMSE of 0.323% for the best regression model. The ANN and adaptive neuro-fuzzy inference system (ANFIS) for the prediction of crop yield using two different models were evaluated and the use of the ANFIS model was suggested as a reliable tool for crop yield prediction in agriculture (Kalpana et al., 2020). A study on the modeling of wheat yield using ANN and ANFIS models employed 16-year-long data about wheat yield observations and weather data and evaluated the models using the coefficient of determination (R^2) and RMSE (Kumar et al., 2023). The ANFIS model showed better performance with R^2 values of 0.85 and 0.83 for training and testing, respectively. Razavi et al. (2011) used ANFIS to predict pear yield based on weather parameters and soil properties and demonstrated that the ANFIS model outperformed traditional regression models in predicting crop yield. The prediction of soil nutrients based on soil properties and weather data was carried out with the ANFIS method (Ramsauer et al., 2022). The study revealed that ANFIS could accurately predict soil nutrients, which can help farmers in fertilizer management. The use of ANFIS to predict dairy cow milk yield based on environmental factors and management practices was also studied (Gholami et al., 2018). They showed that ANFIS can accurately predict milk yield, thereby helping farmers optimize feeding and management practices. Based on photographs of the crop leaves, the ANFIS was used to identify tomato illnesses and ANFIS was shown to achieve high accuracy in diagnosing diseases, which can help farmers in timely disease management (Kamath et al., 2021).

A literature survey on crop yield prediction reveals that various modeling techniques have been applied for predicting crop yields, including statistical modeling, ML, and deep learning-based approaches. Some of the commonly used statistical techniques are regression, multivariate regression, and time series analyses. ML-based approaches include artificial neural networks, support vector regression, decision trees, and random forests. Deep learning-based approaches such as

convolutional neural networks and recurrent neural networks have also been used for crop yield prediction using remote sensing and weather data.

To elaborate on the importance of using tailored modeling techniques for specific crops and regions, understanding the varying requirements of different crops for growth and development is essential. For instance, wheat has specific temperature and water requirements for optimal growth and yield production. In addition, environmental factors, such as soil quality, sunlight exposure, and pests and diseases, vary significantly across different regions, even within the same country. Thus, modeling techniques that work well for one crop or region may not necessarily perform well for another.

Bagalkot district in Karnataka, India, predominantly cultivates wheat as a major crop. However, the region experiences irregular rainfall patterns throughout the year, which poses a significant challenge to wheat cultivation. Accurately predicting the atmospheric conditions is therefore crucial to aid farmers in planning and managing their crops. In this regard, novel modeling techniques that consider the specific requirements of wheat in terms of temperature and water sources can provide valuable insights into the potential yield production of the crop under varying environmental conditions. In conclusion, employing tailored modeling techniques for specific crops and regions can significantly enhance the accuracy and reliability of yield predictions. Timely and optimum use of temperature and water sources as predictors for wheat yield prediction in Bagalkot can aid farmers in planning and managing their crops to mitigate the adverse effects of irregular rainfall patterns. Therefore, it is essential to assess and compare the efficiency of different modeling approaches tailored to specific crops and regions to optimize crop yield and ensure food security.

2. Material and Methods

Bagalkot district, located in the state of Karnataka, India, is well-known for cultivating wheat crops on a large scale. The farmers in this region begin sowing seeds in September and complete the harvest by March in the ensuing year. To collect data regarding crop yield and water sources in Bagalkot district, the government website (<https://eands.dacnet.nic.in>) was referred to. Additionally, temperature and humidity data spanning from the year 2000 to 2020 were obtained from <https://www.timeanddate.com/weather/>. Table 1 summarizes the dataset obtained from various sources, including the four independent variables: temperature, humidity, rainfall, and irrigation sources (canal, well, or other means) for an average month from September to March.

Table 1. Agricultural Dataset

Year	Temperature (in °C)	Humidity (in %)	Average rain fall (mm)	Area (mm ²)	Production	Yield	Irrigation Provision for Crop Cultivation				
							Demand	Canal	Well	Other	Total
999	27.31	75.33	1007	21584	29875	1.38	66.22	5.93E4	5.21E4	7.63E4	1.88E5
2000	27.83	75.5	871	23969	37181	1.55	85.61	6.34E4	5.79E4	7.73E4	1.99E5
2001	25.75	74.83	568	25855	34969	1.35	72.77	5.75E4	6.52E4	5.91E4	1.82E5
2002	27.66	74.91	464	23327	32386	1.39	65.76	2.58E4	8.12E4	6.89E4	1.76E5
2003	26.83	74.25	439	15300	18327	1.20	72.16	1.56E4	7.27E4	6.93E4	1.58E5
2004	26.16	79.66	571	21202	32903	1.55	68.64	5.06E4	8.39E4	7.77E4	2.12E5
2005	26.11	75.66	1119	21840	34948	1.60	69.35	5.97E4	8.56E4	7.70E4	2.22E5
2006	26.83	76.41	994	20992	27344	1.30	75.81	4.65E4	8.15E4	8.62E4	2.14E5
2007	27.25	75.16	899	23208	38300	1.65	78.57	4.75E4	9.19E4	8.85E4	2.28E5
2008	26.41	74.83	1022	25668	37486	1.46	80.68	5.22E4	9.78E4	1.01E5	2.51E5
2009	27.33	74.83	1054	26356	36213	1.37	80.8	5.48E4	9.98E4	1.06E5	2.61E5
2010	27.33	75.16	990	24685	40235	1.63	86.87	5.48E4	1.04E5	1.17E5	2.75E5
2011	27.33	74.75	960	25676	41483	1.62	94.88	5.78E4	1.08E5	1.18E5	2.83E5
2012	27.51	74.54	980	20920	18662	0.89	93.51	2.35E4	1.09E5	1.20E5	2.53E5
2013	26.83	75.32	967	17901	20713	1.16	95.86	5.15E4	1.08E5	1.14E5	2.74E5
2014	26.33	73.75	947	23237	36071	1.55	86.53	5.27E4	1.06E5	1.16E5	2.75E5
2015	28.00	77.33	1013	20118	27808	1.38	92.29	2.03E3	1.22E5	1.08E5	2.33E5
2016	27.66	77.41	851	22101	28303	1.28	98.51	2.41E4	1.25E5	9.66E4	2.46E5
2017	27.52	75.83	967	22675	33798	1.49	99.87	3.58E4	1.29E5	1.03E5	2.67E5
2018	27.41	76.6	1197	24665	39647	1.61	103.63	5.90E4	1.36E5	1.21E5	3.15E5
2019	27.58	75.80	990	21038	33757	1.60	107.86	5.31E4	1.30E5	1.10E5	2.93E5
2020	27.75	76.16	1109	24685	40235	1.63	109.59	4.40E4	1.80E5	1.38E5	3.61E5

The variables considered for a given year of wheat crop cultivation were the area, production, and yield. Initial examination of the data revealed that wheat crop cultivation had steadily increased over the years, which could be attributed to the variation in the availability of water resources and changing levels of rainfall in the region. Correlation techniques are statistical methods that allow us to explore and quantify the relationship between two or more variables. Table 2 presents the coefficient of correlation a measure of the strength and direction of the relationship between the dependent and independent variables.

Scatter diagrams are visual representations of the relationship between two variables, with each point representing an observation of the two variables. In the present study, Figure 1 shows the relationship between different variables and provides insight into the nature of the relationship between them. The results indicate a positive relationship between rainfall and crop production. This suggests that when there is more rainfall, there is a higher crop yield. Additionally, moderate correlations ranging from 0.552 to 0.850 were observed between different sources of water used for crop cultivation, including wells, canals, and other sources. This implies that the type of water source used for crop cultivation affects crop yield.

The scatter diagram between water sources and crop yield also established a positive correlation, further

highlighting the importance of water as the main resource for crop cultivation. The use of different water sources, such as canals, wells, or tankers, showed a strong correlation with the demand for crops. This suggests that the availability of water resources influences the demand for crops. The study also showed that the main water resource for wheat crop cultivation was from wells or bore water resources. This indicates that the type of water source used for crop cultivation may vary depending on the type of crop grown with other variable parameters.

Furthermore, the study revealed that the surrounding temperature, humidity, and water resources are key factors affecting wheat crop cultivation. Farmers choose to cultivate crops based on these environmental conditions, which can vary throughout the year. As shown in Figure 1B, the yield increases with an increase in the area of land used for cultivation, which also depends on the available environmental conditions.

Table 2 Coefficient of correlations

	Temperature	Humidity	Rainfall	Well Water	Canal water	Other water resources
Yield	0.35	0.45	0.56	0.6	0.52	0.54
Demand	0.52	0.42	0.53	0.85	0.82	0.84

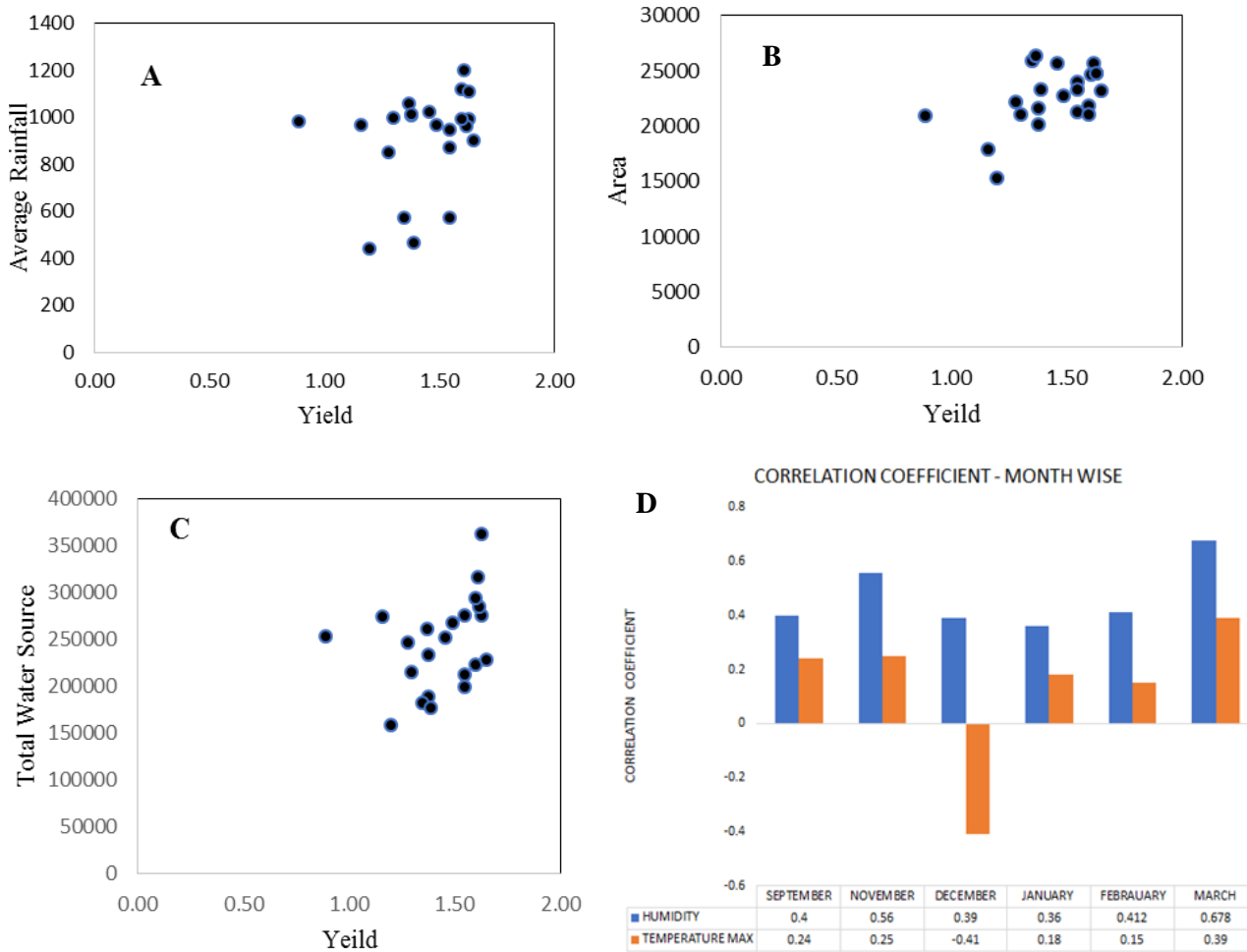


Figure 1 (A) Variation of yield for average rainfall, (B) Variation of yield for area, (C) Variation of yield for Total water source, and (D) Monthly coefficient of correlations.

Water resources are critical for crop cultivation as they have a strong positive correlation with crop yield over time. Figure 1C illustrates this relationship. Adequate availability of water plays a crucial role in crop growth and development. Without adequate water resources, crops cannot grow properly and their yield decreases. Therefore, farmers must ensure to have enough water for their crops. They can do this by using irrigation techniques or relying on rainfall. Overall, the study highlights the importance of water resources in crop cultivation, with a strong positive correlation between water resources and crop yield over time.

Temperature and humidity also play a predominant role in farming. Temperature affects plant growth and development by influencing the rate of photosynthesis, respiration, and transpiration. The optimal temperature range for most crops is between 20°C and 30°C. Above or below this range, the yield decreases. Similarly, humidity affects crop growth by regulating transpiration, which also involves water loss from plants. The optimal humidity range for most crops is between 60% and 80%. In dry weather conditions, the duration of sunshine has a good correlation with crop

yield (Chmielewski and Pots, 1995). Longer periods of sunshine can compensate for the lack of humidity to some extent.

In Bagalkot, crop cultivation begins in September and harvesting happens in March (seven months). During this time, humidity and temperature strongly affect crop yield. Figure 1D shows the correlation between the maximum temperature and minimum humidity value for each month and how it affects crop yield for different years. The maximum humidity was always found between 12 am to 5 am, and the minimum humidity was found in mid-afternoon. There was a positive correlation between temperature and humidity towards crop yield. This explains that crop yield mainly depends on the duration of sunshine and humidity during the afternoon. In the months between September to November, the rainfall had improved humidity, resulting in improved crop cultivation. The higher correlation was found during the harvesting period, i.e. in March when the higher humidity and temperature during the daytime had improved crop cultivation. Therefore, farmers must pay attention to weather patterns and adjust their farming practices accordingly to ensure optimal crop yield.

3. Results and Discussion

In the current study, various input parameters were considered, such as year, temperature, humidity, average rainfall, total water source, area, production, and yield, out of which temperature, humidity, average rainfall, and total water source were found to be significant predictors of Yield. The ANN and ANFIS models were used to predict the yield, and a MATLAB toolbox was employed for the same. For the ANN model, 22 iterations were performed, and the model was trained with 70% of the data, validated with 15%, and tested with 15% using the Levenberg-Marquardt training method and feedforward backpropagation.

The optimal architecture was obtained through a trial-and-error method and achieved at 4, 10, 1, and 1, as shown in Figure 2A. The regression plot drawn in Figure 3 shows the training, testing, and validation stages, based on which the optimal level of the model was decided. The membership functions used were log sigmoid and pure line functions, and the mean square error was chosen as the performance metric for evaluation. On the other hand, the Adaptive Neuro-fuzzy Inference System (ANFIS) model was trained with 81 rules, 3 input layers, and 1 output layer, using trim membership function type with constant variation and 4 hidden layers (Figure 2B).

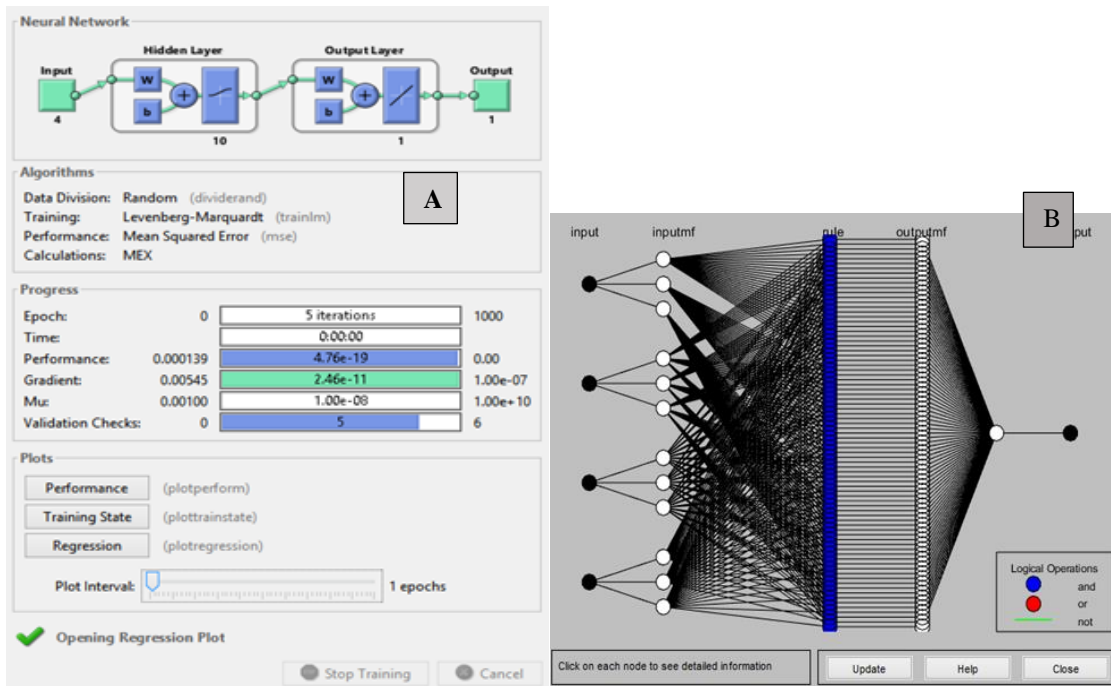


Figure 2 Architecture of Optimal (A) ANN model and (B) ANFIS model

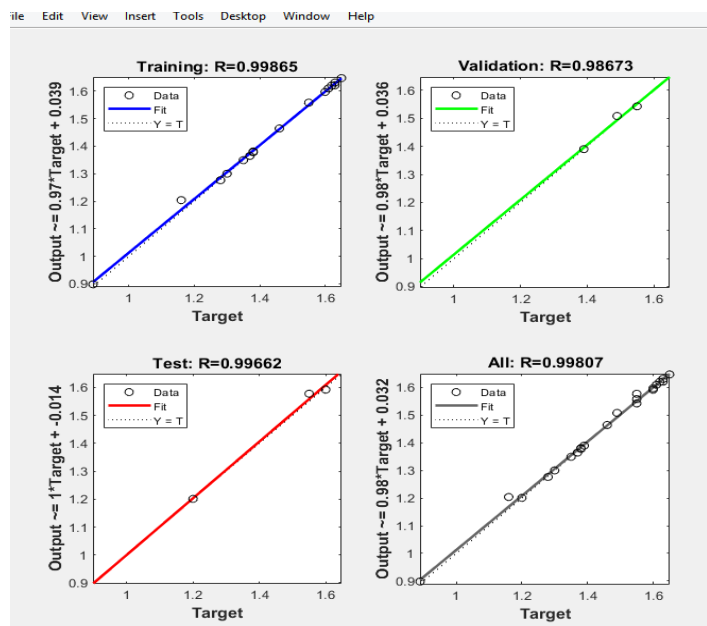


Figure 3 Regression plot for ANN optimal model.

The membership function was chosen based on the average training and testing errors by trial-and-error methods. For the optimal model, the average training and testing errors were 3.05E-08 and 0.29934, respectively. The model was trained with an epoch of 1000. It employed 17 training and 5 testing data for optimal performance. Table 3 shows that the Adaptive Neuro-fuzzy Interference System (ANFIS) model has 2 times less error (0-3%) than the ANN model (0-6%), indicating that the ANFIS model is a better predictor compared to the ANN model. The performance of the models was evaluated using Mean square error (MSE), Mean Absolute error (MAE), and Root mean square error (RMSE), as depicted in Table 4. The optimal ANFIS model also provides a 3-D mapping between input and output parameters, as shown in Figure 4. Similar analogies were used to design optimal models (Manoj et al., 2021; 2022).

This study showed a positive relationship between rainfall and crop production, indicating that higher rainfall results in a higher crop yield. Additionally, moderate correlations were observed between different sources of water used for crop cultivation, including wells, canals, and other sources. This suggests that the type of water source used for crop cultivation may affect crop yield. The present study also revealed that the type of water source used for crop cultivation may vary depending on the type of crop being grown. Wells or bore water resource was found to be the main water

resource for wheat crop cultivation. Overall, these findings suggest that both rainfall and water sources influence crop yield. While rainfall can be unpredictable and subject to seasonal variations, the availability and management of water resources can have a significant impact on crop cultivation. Therefore, farmers need to consider both rainfall and water sources in making informed decisions regarding crop cultivation and water resource management (Figure 4A).

Rainfall and temperature are crucial factors that can significantly affect crop yield. Adequate rainfall is necessary for crop growth, but excessive or inadequate rainfall can result in adverse effects on crop yield, such as waterlogging or drought stress. Similarly, temperature plays a critical role in crop growth and development, and each crop has its own optimal temperature range. Extreme temperatures can result in decreased agricultural output, impaired crop quality, and increased vulnerability to pests and diseases, even if they are unexpectedly high or low.

The model employed in the present study examined the correlation between average rainfall, temperature, and wheat crop yield, considering the local environmental conditions. Notably, this model is specific to wheat crops and may not be applicable to other crops or different environmental conditions, as each crop has distinct water and temperature requirements that vary based on different factors such as soil type, topography, and altitude.

Table 3: Agricultural data set used for ANN and ANFIS prediction of Yield with percentage error

zz	Humidity (in %)	Average rain fall (mm)	Total Water source	Yield	ANN Predicted	Percentage Error	ANFIS Predicted	Percentage Error
27.30	75.33	1007.9	1.8E5	1.38	1.38	0.060	1.381	0.060
27.80	75.51	871.3	1.98E5	1.55	1.57	1.788	1.578	1.788
25.80	74.83	568	1.81E5	1.35	1.34	0.045	1.349	0.045
27.73	74.91	464	1.75E5	1.39	1.39	0.007	1.390	0.007
26.83	74.25	439	1.57E5	1.20	1.20	0.070	1.211	0.903
26.16	79.66	571	2.12E5	1.55	1.55	0.536	1.558	0.536
26.11	75.66	1119	2.22E5	1.60	1.59	0.154	1.598	0.154
26.83	76.41	994	2.14E5	1.30	1.30	0.022	1.310	0.791
27.25	75.16	899	2.27E5	1.65	1.64	0.152	1.687	2.272
26.41	74.83	1022	2.51E5	1.46	1.46	0.305	1.464	0.305
27.33	74.83	1054	2.60E5	1.37	1.36	0.363	1.365	0.363
27.33	75.16	990	2.75E5	1.63	1.62	0.510	1.622	0.510
27.33	74.75	960	2.83E5	1.62	1.62	0.042	1.621	0.042
27.50	74.51	980	2.52E5	0.89	0.89	0.935	0.898	0.935
26.83	75.12	967	2.73E5	1.16	1.20	3.805	1.174	1.219
26.33	73.75	947	2.75E5	1.55	1.54	0.462	1.543	0.462
28.00	77.33	1013	2.32E5	1.38	1.37	0.119	1.378	0.119
27.66	77.41	851	2.42E5	1.28	1.27	0.269	1.277	0.269
27.51	75.83	967	2.67E5	1.49	1.50	1.230	1.508	1.230
27.41	76.62	1197	3.15E5	1.61	1.61	0.000	1.610	0.000
27.58	75.81	990	2.93E5	1.60	1.50	6.111	1.592	0.486
27.75	76.01	1109	3.61E5	1.63	1.63	0.091	1.631	0.091

Table 4: Evaluation matrices for ANN and ANFIS models

Sl. No.	Measurement Index	Yield	
		ANN	ANFIS
1.	RMSE	0.15767	0.128002
2.	MSE	0.02486	0.016384
3.	MAE	0.01260	0.008192

Nonetheless, the developed model can be useful for predicting crop yield with respect to temperature and rainfall (Figure 4B). The model reveals that the correlation between humidity, temperature, and crop yield is necessary for crop growth and development. Temperature and humidity have a direct relationship and impact the atmospheric water vapor content, which is essential for crop growth (psychrometry). Maintaining moderate humidity and temperature values

is crucial for improving crop yield. High humidity can cause moisture stress and increase the probability of pest and disease attacks, whereas low humidity can lead to moisture loss and lower crop yield. High-temperature levels can increase the rate of evapotranspiration and result in water stress, whereas low temperatures can slow down crop growth.

Thus, while there may not be a straightforward correlation between humidity, temperature, and crop yield in the study, these environmental factors have been shown to play a crucial role in crop cultivation and can affect crop yield in various ways (Figure 4C). Furthermore, the results suggested that the type of water source used for crop cultivation may affect crop yield, and there is a strong correlation between water sources and the demand for crops.

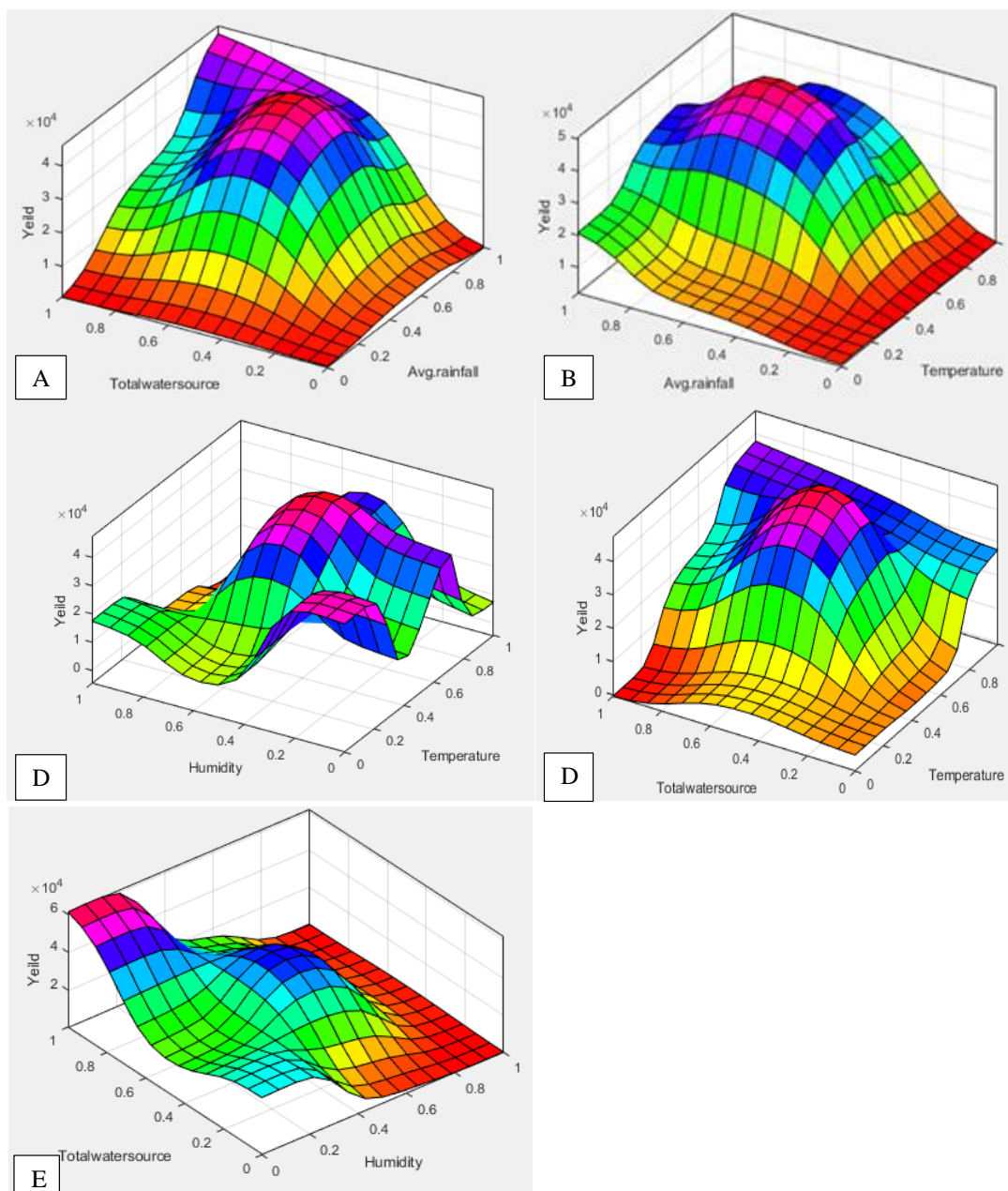


Figure 4. ANFIS input and output parametric variation obtained from the optimal model.

The main water resource for wheat crop cultivation is wells or bore water resources. This suggests that the type of water source used for crop cultivation may vary depending on the type of crop being grown. Additionally, surrounding temperature and water resources are found to be the key factors affecting wheat crop cultivation, and farmers choose to cultivate crops based on these environmental conditions, which can vary throughout the year. The crop yield was observed to increase with the rise in the area of land used for cultivation, which also depends on the available environmental conditions. Therefore, a correlation between water sources, temperature, and crop yield seems probable, with the optimal combination of these factors varying depending on the type of crop being grown and the local environmental conditions (Figure 4D).

This study examined the relationship between humidity, water source, and crop yield. High humidity can increase the atmospheric water vapor content, while water sources provide the necessary water for crop growth. However, too much humidity can lead to moisture stress and increased susceptibility to pests and diseases. On the contrary, inadequate water supply can lead to drought stress and reduced crop yield. Therefore, both humidity and water sources should be balanced to ensure optimal crop growth and yield. The study also found a moderate correlation between different sources of water used for crop cultivation, including wells, canals, and other sources, and crop yield. This indicates that the type of water source used for crop cultivation can have an impact on crop yield.

The scatter diagram between water sources and crop yield (Figure 1A) also revealed a positive correlation, further supporting the importance of water as the main resource for crop cultivation. The use of different water sources, such as canals, wells, or tankers, showed a strong correlation with the demand for crops, indicating that the availability of water resources may influence the demand for crops. In terms of humidity, while there may not be a direct correlation between humidity and crop yield in the study, high humidity can increase the atmospheric water vapor content, leading to moisture stress and increased susceptibility to pests and diseases. Therefore, humidity levels should be monitored and balanced to ensure optimal crop growth and yield (Figure 4E).

4. Conclusions

This study has implications for policy formulation and farmer decision-making regarding crop cultivation and water resource management, enabling policymakers to shape sustainable agricultural policies and enabling farmers to increase yield, enhance profitability, and meet market demands. Temperature, humidity, average rainfall, and total water source are

significant environmental determinants driving crop growth and development. Rainfall and temperature play a significant role in crop growth, with extreme weather conditions hampering yield and deviations from optimal temperature range affecting crop yield and quality. The interplay of humidity, temperature, and crop yield is crucial for crop growth and development, with wells and borewater dominating as primary water sources for wheat cultivation. The results also showed that the type of water source used, such as wells, canals, or tankers, may influence crop productivity and demand. The correlation ranges from 0.55 to 0.8, are in accordance with the investigation. Temperature, humidity, average rainfall, and total water source were found to be important predictors of crop yield in the study.

Authors' Contribution

Dr. Shailesh Rao spearheaded the data collection process from various sources and conducted correlation and regression analyses. Dr. Manoj contributed by performing ANFIS analysis for the study. Dr. Anjana Krishnan played a crucial role in detecting issues throughout the correlation, regression, and ANFIS analysis phases.

Bibliographic References

- Achieng, Kevin, O. 2019. Modeling of Soil Moisture Retention Curve Using Machine Learning Techniques: Artificial and Deep Neural Networks Vs Support Vector Regression Models. *Computers & Geosciences*, 133, 104320. DOI: <https://doi.org/10.1016/j.cageo.2019.104320>.
- Ansarifar, J., Wang, L., Archontoulis, S.V. 2021. An interaction regression model for crop yield prediction. *Scientific Reports*, 11(1), 1-10. DOI: <https://doi.org/10.1038/s41598-021-97221-7>.
- Ayyappa, T., Manoj, I.V., Kurse, S., Kumar, J., Bhattacharyya, T. 2021. Evaluation of prediction performance for K-Nearest Neighbour, Random Forest algorithm and Extreme Gradient Boosting algorithms in the forecasting of fault diagnosis for roller bearings. *Recent Advances in Material, Manufacturing, and Machine Learning*, 491-500, CRC Press Books., DOI: <https://doi.org/10.1201/9781003358596-41>
- Beloti, I.F., Nunes Rodrigues de Azevedo, B., Maciel, G.M., Peres, H.G., Pivetta Momesso, M., Alves, I.M. 2019. Agronomic Performance and Physiological Quality of Pepper Seeds Under Different Planting System. *Revista de Agricultura Neotropical*, 6(4), 57-62. DOI: <https://doi.org/10.32404/rean.v6i4.3432>.
- Bhavani, A.K., Singh, G, Bhavani, A.K., Singh. 2018. ARIMA Forecast of Area, Production and Productivity of Rice in Amalgamated Raipur Using R. *International Journal of Agricultural Science and Research*, 8(1), 29-36. DOI: <https://doi.org/10.24247/ijasrfeb201818>.

- Chmielewski, F.M., Potts, J.M. 1995. The relationship between crop yields from an experiment in southern England and long-term climate variations. *Agricultural and Forest Meteorology*, 73(1-2), 3–66. DOI: [https://doi.org/10.1016/0168-1923\(94\)02174-I](https://doi.org/10.1016/0168-1923(94)02174-I)
- Drummond, S.T., Sudduth, K.A., Joshi, A., Birrell, S.J., Kitchen, N.R., Palm, H.L. 2003. Statistical and neural methods for site-specific yield prediction. *Transactions of the ASAE*, 46(1), 5-14. DOI: <https://doi.org/10.13031/2013.12541>.
- Elavarasan, D., Vincent, D. R., Sharma, V., Zomaya, A. Y., Srinivasan, K. 2018. Forecasting yield by integrating agrarian factors and machine learning models: A survey. *Computers and Electronics in Agriculture*, 155, 257–282. DOI: <https://doi.org/10.1016/j.compag.2018.10.024>
- Fukuda, S., Spreer, W., Yasunaga, E., Yuge, K., Sardud, V. 2013. Random forests modelling for the estimation of mango (*Mangifera indica* L. cv. Chok Anan) fruit yields under different irrigation regimes. *Agricultural Water Management*, 116, 142-150. DOI: <https://doi.org/10.1016/j.agwat.2012.07.003>.
- Gholami, V., Khaleghi, M.R. 2018. A comparative study of the performance of artificial neural network and multivariate regression in simulating springs discharge in the Caspian Southern Watersheds, Iran. *Applied Water Science*, 9(1), 1-10. DOI: <https://doi.org/10.1007/s13201-018-0886-4>
- Gromping, U. 2009. Variable importance assessment in regression: linear regression versus random forest. *The American Statistician*, 63(4), 308-319. DOI: <https://doi.org/10.1198/tast.2009.08199>
- Huang, J., Chen, X., Zhou, L., Xue, Y., Lin, J. 2017. Statistical analysis of the relationship between climate-induced maize yield and rainy-season precipitation across Inner Mongolia, North China. *Theoretical and Applied Climatology*, 129(3-4), 1145-1156. DOI: <https://doi.org/10.1007/s00704-016-1839-1>
- Jeong, J.H., Resop, J.P., Mueller, N.D., Fleisher, D.H., Yun, K., Butler, E.E., Timlin, D.J. 2016. Random forests for global and regional crop yield predictions. *PLoS One*, 11(6), e0156571. DOI: <https://doi.org/10.1371/journal.pone.0156571>
- Kamath, P., Patil, P., Shrilatha, S., Sushma, Sowmya, S. 2021. Crop yield forecasting using data mining. *Global Transitions Proceedings*, 2(2), 402-407. DOI: <https://doi.org/10.1016/j.gltp.2021.08.008>
- Kumar, V., Ramesh, K. 2023. Development of Artificial Neural Network Model for Indian Steel Consumption Forecast. *Journal of the Institution of Engineers (India): Series D*, 2023. DOI: <https://doi.org/10.1007/s40033-023-00482-x>.
- Kumar, V., Ramesh, K. 2023. Development of Artificial Neural Network Model for Indian Steel Consumption Forecast. *Journal of the Institution of Engineers (India): Series D*, 2023. DOI: <https://doi.org/10.1007/s40033-023-00482-x>.
- Manoj, I.V. 2021. Slant Type Taper Profiling and Prediction of Profiling Speed for a Circular Profile during Wire Electric Discharge Machining using Hastelloy-X. *Proceedings of the Institution of Mechanical Engineers, Part B. Journal of Engineering Manufacture*. 235(21), 5511-5524. DOI: <https://doi.org/10.1177/0954406221992398>
- Manoj, I.V., Narendranath, S. 2020. Variation and artificial neural network prediction of profile areas during slant type taper profiling of triangle at different machining parameters on Hastelloy X by wire electric discharge machining. *Proceedings of the Institution of Mechanical Engineers, Part E. Journal of Process Mechanical Engineering*, 234(6), 982-996. DOI: <https://doi.org/10.1177/0954408920938614>
- Manoj, I.V., Narendranath, S. 2023. Evaluation of WEDM Performance Characteristics and Prediction of Machining Speed during Taper Square Profiling on Hastelloy-X. *Australian Journal of Mechanical Engineering*, 21(4), 1152-1167. DOI: <https://doi.org/10.1080/14484846.2021.1960670>
- Manoj, I.V., Soni, H., Narendranath, S., Mashinini, P.M., Kara, F. 2022. Examination of Machining Parameters and Prediction of Cutting Velocity and Surface Roughness Using RSM and ANN Using WEDM of Altemp HX. *Advances in Materials Science and Engineering*, (special issue), 1-9. DOI: <https://doi.org/10.1155/2022/5192981>
- Misra, D., Oommen, T., Agarwal, A., Mishra, S.K., Thompson, A.M. 2009. Application and analysis of support vector machine-based simulation for runoff and sediment yield. *Biosystems Engineering*, 103(4), 527-535. DOI: <https://doi.org/10.1016/j.biosystemseng.2009.04.017>
- Rahman, A.B.M., Salman. 2022. A Comparative Study Between Linear Regression and Support Vector Regression Model Based on Environmental Factors of a Smart Bee Farm, *Korean Institute of Smart Media*, 11(5), 38–47. DOI: <https://doi.org/10.30693/smj.2022.11.5.38>
- Ramsauer, T., Wei, T., Low, A., Marzahn, P. 2021. An Hourly Soil Moisture Data Set Based on Weather Radar, Soil Properties and Reanalysis Temperature Data. *Remote Sensing*, 13(9), 1712. DOI: <https://doi.org/10.3390/rs13091712>
- Razavi, M.S., Golmohammadi, A., Sedghi, R., Asghari, A. 2019. Prediction of bruise volume propagation of pear during the storage using soft computing methods. *Food Science & Nutrition*, 8(2), 884–893. DOI: <https://doi.org/10.1002/fsn3.1365>
- Russo, T.A., Devineni, N., Lall, U. 2015. Assessment of agricultural water management in Punjab, India, Using Bayesian Methods. *Sustainability of Integrated Water Resources Management*, 147-162. DOI: https://doi.org/10.1007/978-3-319-12194-9_9
- Sarturi, M.V.R., Menegassi Bianchi Teixeira, C. A., Carvalho, I. R., Demari, G. H., Vieira Loro, M., Pradebon, L. C., Donadel Port, E. 2022. Prediction Of Corn Grain Productivity as A Function Of Altitude And Plant Population. *Revista de Agricultura Neotropical*, 9(4), e7070. DOI: <https://doi.org/10.32404/rean.v9i4.7070>.
- Su, Tao. 2012. Regional Yield Estimation for Spring Maize With Multi-Temporal Remotely Sensed Data in Junchuan, China, *Advanced Materials Research*, vol. 610–613, Trans Tech Publications. 3601–05. DOI: <https://doi.org/10.4028/www.scientific.net/amr.610-613.3601>.
- Xu, X., Gao, P., Zhu, X., Guo, W., Ding, J., Li, C., Wu, X. 2019. Design of an integrated climatic assessment indicator (ICAI) for wheat production: a case study in Jiangsu Province, China. *Ecological Indicators*, 101, 943-953. DOI: <https://doi.org/10.1016/j.ecolind.2019.01.059>