



OPEN

## Evaluating crop yield prediction models in illinois using aquacrop, semi-physical model and artificial neural networks

Vishal Gautam<sup>1</sup>, Abdul Gani<sup>2</sup>, Shray Pathak<sup>1</sup> & Anoop Kumar Shukla<sup>3</sup>✉

Crop yield is important for agricultural productivity and the country's economy. While crop yield estimation is an essential aspect of modern agriculture, it continues to be one of the most challenging tasks to manage effectively. Corn and soybean are the important crops in Illinois, USA, considerably enhancing the region's agricultural output and economy. The present study integrates semi-physical model, AquaCrop and Artificial Neural Network (ANN) Models for estimating corn and soybean yields. Data of different meteorological parameters including precipitation, maximum and minimum temperature, relative humidity, wind speed, solar radiation, photosynthetically active radiation and fraction of photosynthetically active radiation, land surface water index were collected for a period of 25 years from 2000 to 2024 from NASA POWER, USDA and NASS. The observed yield of soybean and corn was ranges from 2.49 to 4.37 ton/ha and 7.06 to 14.66 ton/ha. The predicted corn yield using the AquaCrop, semi-physical, and ANN models ranged from 7.60 to 14.42 ton/ha, 9.01 to 13.42 ton/ha, and 6.81 to 15.63 ton/ha, respectively. For soybean, the predicted yield ranged from 2.80 to 4.34 ton/ha, 2.92 to 3.84 ton/ha, and 2.45 to 4.43 ton/ha, respectively. The ANN model achieves the highest coefficient of determination ( $R^2 = 0.96$ ) in predicting soybean yield, while the semi-physical model records the lowest  $R^2$  value of 0.42, indicating the superior predictive capability of the ANN model. For both corn and soybean yields, the ANN model showed the highest prediction accuracy among the other models. Thus, the study underscores the significance of employing the ANN model for crop yield estimation, particularly in the regions that share similar physiographic and meteorological conditions with Illinois.

**Keywords** *AquaCrop, Crop modelling, Illinois, Remote sensing, Yield*

According to the United Nations Food and Agriculture Organization<sup>1</sup>, there were around 1.23 billion people employed in the world's agri-food systems in the year 2019. Crop production prediction is one of the major components of agricultural planning, food security strategy development, and economic forecasting. Agriculture is one of the pillars of economy of the United States of America (USA) and worldwide food supply. Yield forecasting is significantly important in guiding trade strategies, policy decisions, and resource distribution<sup>2</sup>. The varying climatic conditions across states, diverse soil types and agricultural practices present challenges and opportunities for crop yield modelling. Climate change, with its associated increase in weather variability has raised higher demands for solid yield prediction models. Extreme events are bound to occur more frequently, and models must modify and incorporate these nonlinear effects on crop productivity<sup>3</sup>. The major factors of agricultural crop production include temperature, rainfall, pH, and nutrient content present in soil, as well as irrigation water in the form of surface and groundwater. Additionally, agronomic factors like crop management practices, technology use, and plant protection measures significantly influence productivity<sup>4-6</sup>.

Crop development and growth monitoring, and yield prediction are the key components of agricultural planning and management<sup>7,8</sup>. Numerous techniques are employed in this technological age to estimate crop yields. The most frequently utilized of these are statistical techniques<sup>9,10</sup>, crop simulation models, and remote sensing<sup>11</sup>. These methods are also employed to determine how the climate affects the production of various crop

<sup>1</sup>Department of Civil Engineering, Indian Institute of Technology Ropar, Rupnagar 140001, Punjab, India.

<sup>2</sup>Department of Civil Engineering, Netaji Subhas University of Technology, New Delhi 110073, India. <sup>3</sup>Manipal School of Architecture and Planning, Manipal Academy of Higher Education, Manipal 576104, Karnataka, India.

✉email: anoop.shukla@manipal.edu

varieties<sup>12,13</sup>. Various annual crops, including wheat, soybean, corn, sugarcane, and others, have predicted their yields using machine learning techniques.

This has led to a growing interest in crop modeling, driven by the pressing need to address global food security amid challenges such as climate change, population growth, and resource constraints<sup>14</sup>. Advanced crop production prediction models have been developed as a result of recent advancements in data gathering instruments, increased processing capacity, and creative modelling approaches. Crop models have been shown to be essential tools for decision-making, evaluating how management practices and climate change/variability affect the environmental performance and productivity of alternative cropping systems, and advancing more sustainable and effective agriculture<sup>15</sup>. They are a less costly and quicker way to investigate the impacts of agricultural land management strategies on crop yields and the environment and determine the best management to achieve economically efficient yields<sup>16</sup>. Crop simulation models may be used as decision support tools to evaluate the financial and risk implications of agricultural management techniques. The modeling technique may produce results that are quite accurate for developing agricultural land management plans if the models are calibrated and validated using trustworthy observable field data<sup>17</sup>.

AquaCrop model is used to assess crop productivity, evapotranspiration, and water use efficiency under varying degrees of water stress, as well as to analyze crop responses to diverse irrigation strategies and water levels<sup>18,19</sup>. The model further evaluates yield responses to the soil moisture content, fertilization practices, soil organic amendments, and their interactions<sup>20</sup>. It also offers insights into canopy cover, panicle biomass, and grain yield under different irrigation and nitrogen management regimes<sup>21</sup>, and assesses the crop yield to climate change scenarios, accounting for changes in rainfall and temperature<sup>22</sup>. Additionally, AquaCrop has been employed to study crop performance under different irrigation and manure management practices at high altitudes<sup>23</sup>, as well as responses to water stress<sup>24,25</sup>. The AquaCrop model has been extensively applied in studies involving maize, rice, and wheat water balance, particularly in Asia.

According to Basso & Liu<sup>26</sup>, these models employ a range of methodologies, such as process-based models, advanced machine learning techniques, and traditional statistical methods. Since each model has unique benefits and drawbacks, comparing them is crucial to comprehending how applicable they are in a variety of geophysical and agricultural settings. Recent developments have created chances to integrate precision agricultural data, long-term climate forecasts, and high-resolution satellite images to increase model accuracy and geographical resolution<sup>27</sup>. Recent developments in crop modeling provides a validation for enhanced complexity, integration with diversified data sources, and application of computational techniques<sup>28</sup>. Semi-physical models simulate crop growth from a physiological processes and interaction with the environment, continually improving the dynamics of the plant soil atmosphere system. The crop models are characterized by better representations of root architecture, nutrient cycling, and crop-specific behaviours, enables more accurate simulations of crop responses to abiotic environmental stressors<sup>29</sup>. Meanwhile, the emergence of big data and machine learning has further transformed agricultural research<sup>30</sup>.

In USA, 97 million hectares area used for crop production yearly. USA is one of the biggest producers and exporters of major crops, including maize, soybean, wheat, and rice, making up 87% of the total agricultural area. USA contributed 27%, 32% and 6% of the world's yearly production of maize, soybeans, and wheat, during the last ten years (2013–2022). In turn, the US is one of the main exporters of rice, even if its area and production are relatively minor when compared to Asian nations<sup>31</sup>. Table 1 depicts the average production, harvested area and average yields of various crops.

According to the United State of Department of Agricultural<sup>32</sup>, the average US corn production in the US between 2021 and 2023 is mainly concentrated in the Midwest, with Iowa accounting for 18% of the national output, followed by Illinois at 16%. Nebraska and Minnesota each contribute 10%, while Indiana produces 7%. South Dakota, Wisconsin, Kansas, and Ohio each account for 4–5% of the total. Similarly, soybean production is heavily centered in the Midwest. Illinois leads with 16% of national output, followed by Iowa at 14%, Minnesota at 8%, and Indiana and Ohio each contributing 8%. Nebraska and Missouri add 6% each, while Arkansas and Mississippi, located in the southern region, each contribute a smaller share of 3%<sup>29</sup>.

The present study is design with the objectives (i) to estimate the corn and soybean yield in Illinois, USA by using AquaCrop, semi-physical and ANN models; (ii) comparison of the performance of these models for predicting corn and soybeans yields, and therefore identifying which model is most effective in predicting corn and soybean yield.

## Materials and methods

### Study area

Illinois state located in the middle of the US with significant historical, natural, and economic significance and lies in between 37° to 42°N and 87° to 91°W<sup>33</sup>. The state shares borders with Indiana to the East, Kentucky to the

S. No.	Crops	Total Production (Million ton)	Harvested Area (Million Ha)	Average Yield (ton/Ha)
1.	Maize (Corn)	364.1	33.6	10.8
2.	Soybean	111.2	33.6	3.3
3.	Wheat	52.3	16.5	3.1
4.	Rice	9.1	1.1	8.4

**Table 1.** Average production, harvested area and yield of maize, soybean, wheat, and rice in the USA<sup>31</sup>.

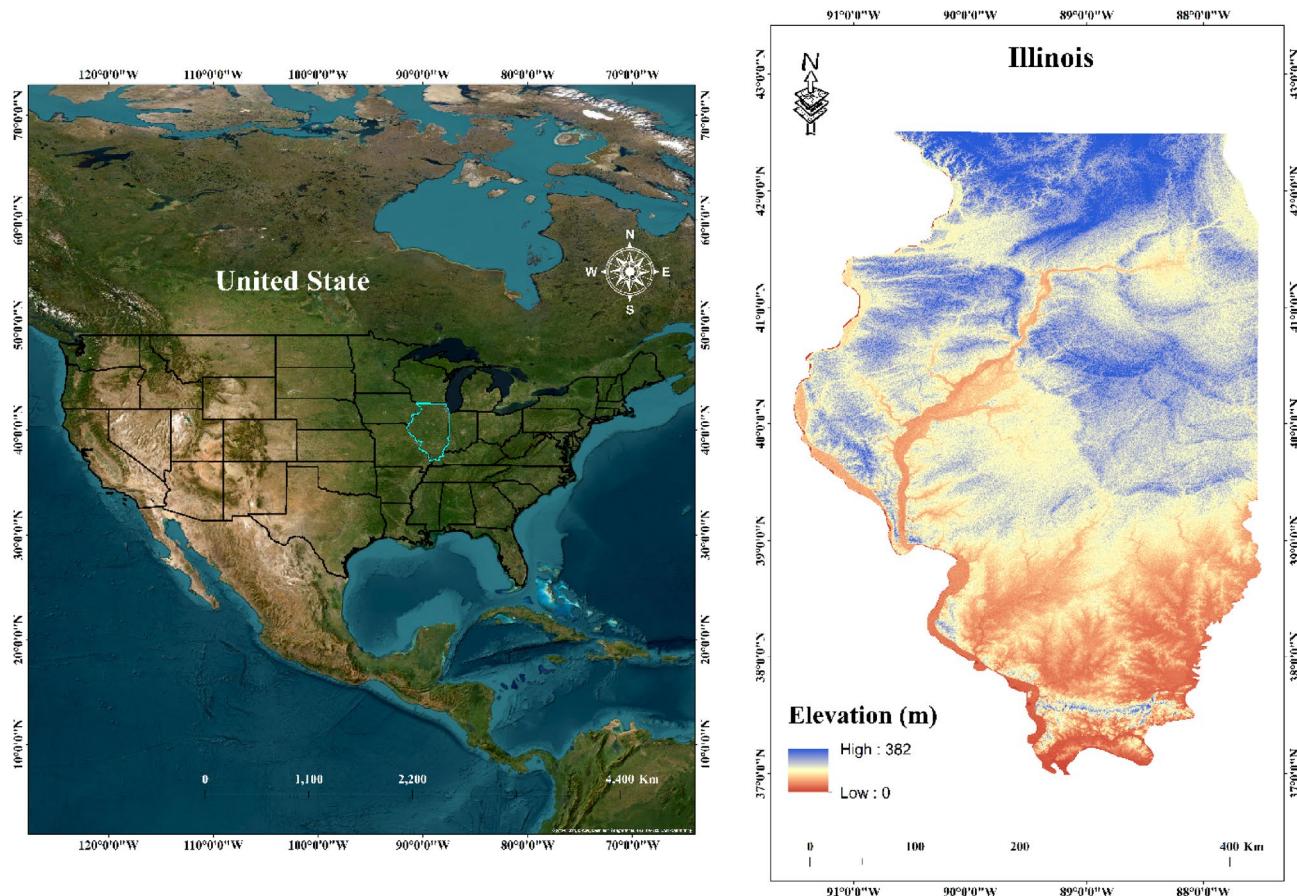
South, Iowa and Missouri to the West, and Wisconsin to the North. According to the USDA, Illinois accounts for 16% of the country's corn production and 16% of its soybean production. These statistics highlight Illinois' vital role in the national agricultural sector. Illinois is known for its diverse geography, which ranges from hills and forests in the South to flat grasslands in the Centre and North. In Illinois, average yearly temperatures range from 9 °C in the North to 14 °C in the South. Seasonal highs vary from 14 °C to 19 °C during spring and fall, while summer highs average 27 °C. Annual precipitation exceeds 48 inches in Southern Illinois, dropping to less than 32 inches in the North. Snowfall patterns are reversed, with Northern regions receiving an average of 36 inches annually, compared to less than 10 inches in the South. The Chicago area sees the most significant winter snowfall, heavily influenced by lake-effect snow from Lake Michigan. Figure 1 shows the study area of Illinois and its county boundaries.

### Data collection

The data of various parameters used in the AquaCrop, semi-physical and ANN models were collected for a period of 25 years, from 2000 to 2024, in Illinois. Daily Data of temperature, precipitation, solar radiation, wind speed, relative humidity, surface soil-wetness, soil moisture profile and root zone soil wetness were collected from NASA POWER, while photosynthetically active radiation (PAR), fraction of absorbed photosynthetically active radiation (fAPAR) and LSWI data were extracted from Moderate Resolution Imaging Spectroradiometer (MODIS) imagery and crop data layer data was collected from USDA NASS imagery. Data of PAR and fAPAR was collected from MODIS MCD18C2 and MODIS MCD15A3H imagery respectively from the month of May to September.

### Semi physical model

Semi-physical models refer to the modeling technique which uses both physical and biological understanding of crop growth with empirical methods. Physical and biological processes include photosynthesis, evapotranspiration and soil-water balance whereas empirical methods include regression models or machine learning techniques to account for local factors. The semi-physical approach for assessing the corn and soybean yields combines crop growth physiology knowledge with remote sensing data to create a thorough framework for evaluating yield dynamics under various environmental circumstances. For the assessment of crop yield, PAR, fAPAR, light use efficiency (LUE), water stress, temperature stress and vegetation growth 8-day composite data were used. The model uses high resolution satellite data to estimate yield and accounts for the varied spatial



**Fig. 1.** Study Area Map (created using QGIS software, version 3.40.7; available at <https://qgis.org/download/>).

distribution of environmental factors across numerous agroecological regions. Figure 2 depicts the methodology adopted for the semi-physical model.

### Light use efficiency ( $\epsilon$ )

The light use efficiency ( $\epsilon$ ) is calculated by using Eq. 1<sup>34</sup>:

$$\epsilon = \epsilon_{max} \times T_s \times W_s \times P_s \quad (1)$$

Where,

$\epsilon_{max}$  is the Maximum light use efficiency (gC/MJ) and its value is 0.34<sup>35</sup>.

$W_s$  is the water stress;

$T_s$  is the temperature stress;

$P_s$  is the vegetation growth.

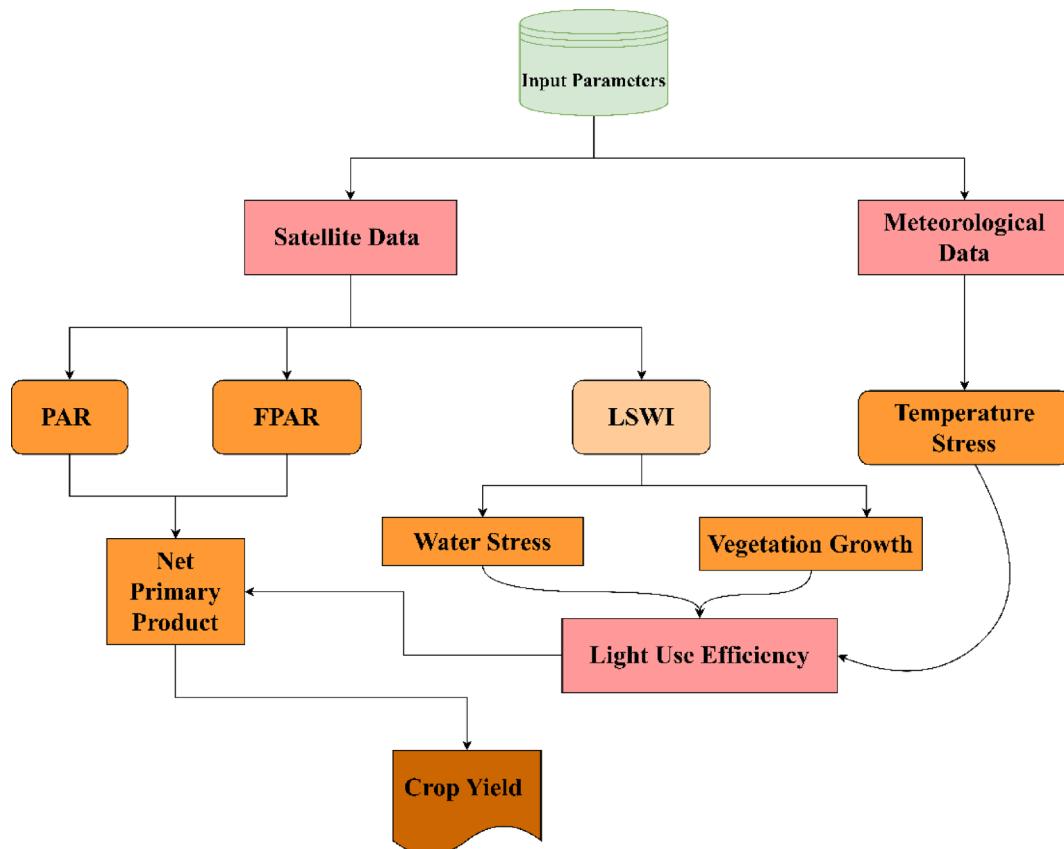
### Water stress ( $W_s$ )

Remote sensing data from MOD09A1 MODIS product is used to assess water stress. Land surface water indicator (LSWI) measures moisture content in both plant and soil and calculated by using reflectance data from the shortwave infrared (SWIR) and near-infrared (NIR) bands. Using both LSWI and  $W_s$ , water-stressed areas can be effectively identified and mapped, supporting enhanced water management strategies and improving yield predictions across various agricultural landscapes. LSWI is calculated by using the Eq. 2<sup>36</sup>:

$$LSWI = \frac{NIR - SWIR}{NIR + SWIR} \quad (2)$$

LSWI ranges from  $-1$  to  $+1$ , where higher values correspond to increased moisture availability. In order to calculate  $W_s$  throughout the growing season, the maximum LSWI value ( $LSWI_{max}$ ) is determined, and water stress is calculated using the Eq. 3:

$$W_s = \frac{(1 - LSWI)}{(1 + LSWI_{max})} \quad (3)$$



**Fig. 2.** Flowchart of Semi-Physical Model.

$W_s$  lies from 0 to 1, where values closer to 1 indicates higher levels of water stress.

### Temperature stress ( $T_s$ )

Temperature stress during the cropping period is calculated by using Eq. 4<sup>36</sup>:

$$T_s = \frac{(T - T_{\min})(T - T_{\max})}{(T - T_{\min})(T - T_{\max}) - (T - T_{opt})/2} \quad (4)$$

Where,

$T$  is the Mean temperature (°C),

$T_{\min}$  is the Minimum temperature (°C),

$T_{\max}$  is the Maximum temperature (°C),

$T_{opt}$  is the Optimum temperature for photosynthesis (°C) and its value is 30 °C for corn and soybeans.

### Vegetation growth

The two phases of the vegetation growth ( $P_s$ ) influence index are the leaf growth phase ( $P_{s1}$ ) and grain ripening phase ( $P_{s2}$ ) and is calculated as given by Xiao et al.<sup>37</sup>. The leaf growth phase is a significant phase in the crop development because it comprises of beginning and expansion of leaves, which play an important role in canopy formation and photosynthesis. This stage is crucial in determining the total amount of biomass produced, the health of the plants, and eventually the crop yield. The grain ripening phase marks the final stage of crop development and is critical for determining the harvestable yield of crops. This phase follows grain filling and is characterized by physiological maturation, moisture loss, and stabilization of seed dry matter.

$$P_{s1} = 1 \quad (5)$$

$$P_{s2} = \frac{1 + LSWI}{2} \quad (6)$$

### Estimation of crop yield

Crop yield is assessed by determining net primary productivity (NPP) and NPP is calculated by using the Eq. 7<sup>34</sup>:

$$NPP = \sum PAR \times fAPAR \times \epsilon \times \Delta t \quad (7)$$

After determining NPP, crop yield was further calculated by using a yield model, represented by the Eq. 8<sup>34</sup>:

$$Yield = a \times \sum NPP + b \quad (8)$$

Where, a and b are the regression constants for the relationship between NPP and crop yield and value of a is 0.12 and 0.073 for corn and soybean respectively, while the value of b is 0.739 and 1.298 for corn and soybean respectively as given by Na et al.<sup>34</sup>.

### AquaCrop model

AquaCrop is the water-driven simulation tool developed by the Food and Agriculture Organization (FAO) of the United Nation Land and Water Division for planting and scenario analysis<sup>38,39</sup>. The model connects the soil, crop, and atmospheric components through a comprehensive soil-water balance framework<sup>40</sup>. The model particularly focuses on the crop canopy, which is the upper layer of leaves and stems, to better understand how different factors influence crop development and yield in different farming environments. The model also evaluates crop yields and the associated carbon, energy, and water footprints, offering valuable insights into food production challenges and solutions to water scarcity<sup>41</sup>. Moreover, it supports assessments related to food security, crop productivity, sustainability, and efficiency, with a focus on water availability and consumption in agricultural production<sup>42</sup>.

AquaCrop model incorporated daily data of various parameters including maximum and minimum temperature, precipitation, reference evapotranspiration, initial soil water content, and soil characteristics. In the absence of ground truth data, initial parameters were employed to estimate crop yields. The growing season for corn and soybeans extended from May to September across all years. Due to variations in sowing dates across whole region, we used the average planting date of 20th May and the average harvesting date of 30th of September.

The reference evapotranspiration ( $ET_o$ ) is calculated by the modified Hargreaves method and given by Eq. 9<sup>43</sup>:

$$ET_o = 0.0013 \times 0.408 \times RA \times (T_{avg} + 17.0) \times (TD - 0.0123 \times P)^{0.76} \quad (9)$$

Where,

$ET_o$  is the reference evapotranspiration,

$RA$  is the extra-terrestrial radiation expressed in  $MJm^{-2}d^{-1}$ ,

$T_{avg}$  is the average daily temperature in °C, and,

$TD$  is the temperature range in °C.

The RA is calculated as given in Eq. 10<sup>43</sup>:

$$RA = \frac{24(60)}{\pi} \times G_{sc} \times d_r \times [w_s \times \sin(\phi) \sin \delta + \cos(\phi) \cos(\delta) \sin(w_s)] \quad (10)$$

Where,

$G_{sc}$  is the solar constant equal to  $0.0820 \text{ MJ m}^{-2} \text{ min}^{-1}$ ,  
 $d_r$  is the inverse relative distance Earth-Sun,  
 $\delta$  is the solar declination in radians,  
 $w_s$  is the sunset hour angle in radians, and.  
 $\phi$  is latitude in radians.

### Artificial neural network (ANN) model

Artificial neural networks (ANNs) are computational models used in machine learning that consist of interconnected processing units. By processing inputs through interconnected neurons, they handle difficult issues such as image recognition and computer vision, outperforming traditional rule-based programming<sup>44</sup>. ANN consists of interconnected artificial neurons that form specialized topologies such as feedforward, feedback, and lateral networks. ANN consists of three layers: input, hidden, and output. The flow diagram of the ANN process is shown in Fig. 3.

### Optimization of network architecture

The optimization of network design for crop yield prediction is an important step in developing precise and reliable agricultural prediction models. ANN provides a non-linear regression model to capture complex, non-linear correlations between input factors and crop production<sup>45</sup>. An input layer, one or more hidden layers, and an output layer are the usual layers of a well-structured ANN model. For developing optimal neural network, architecture for the crop yield prediction involves three important factors: (i) selection of input neurons that accurately represent relevant predictor variables, (ii) developing hidden layer configurations that demonstrate intricate nonlinear interactions, and (iii) developing output layer to effectively depict projected yield<sup>46</sup>. Data preprocessing is part of the optimization process for assuring that all input variables are on comparable scales and no variables dominate the learning process<sup>47</sup>. To assure the model's prediction accuracy and resilience, models' performance was evaluated by using root mean squared error (RMSE), and coefficient of determination ( $R^2$ ).

The ANN model was used to predict crop yield by utilizing composite data for the growing season from May to September including maximum temperature, minimum temperature, precipitation, solar radiation, wind speed, relative humidity, surface soil wetness, profile soil moisture, and root zone wetness<sup>48</sup>. These parameters were integral in enhancing the model's prediction accuracy. The ANN architecture consisted of 10 hidden layers, designed to capture the complex, non-linear relationships between the input variables and crop yield. The model was trained on the 70% dataset, and its performance was assessed by using 15% datasets for each testing and validation, ensuring a robust evaluation of the model's predictive capabilities for both corn and soybean yields.

### Performance evaluation and model selection

The performance evaluation and model selection process for ANN in crop yield prediction follows a comprehensive approach to ensure optimal results. The model's performance is assessed using RMSE and  $R^2$  during the initial phase. RMSE improves interpretability by expressing errors in the same units as the crop yields, facilitating a clearer understanding of the model's accuracy<sup>49</sup>.  $R^2$  measures how well a model predicts crop yields, with values closer to 1 indicating greater performance.

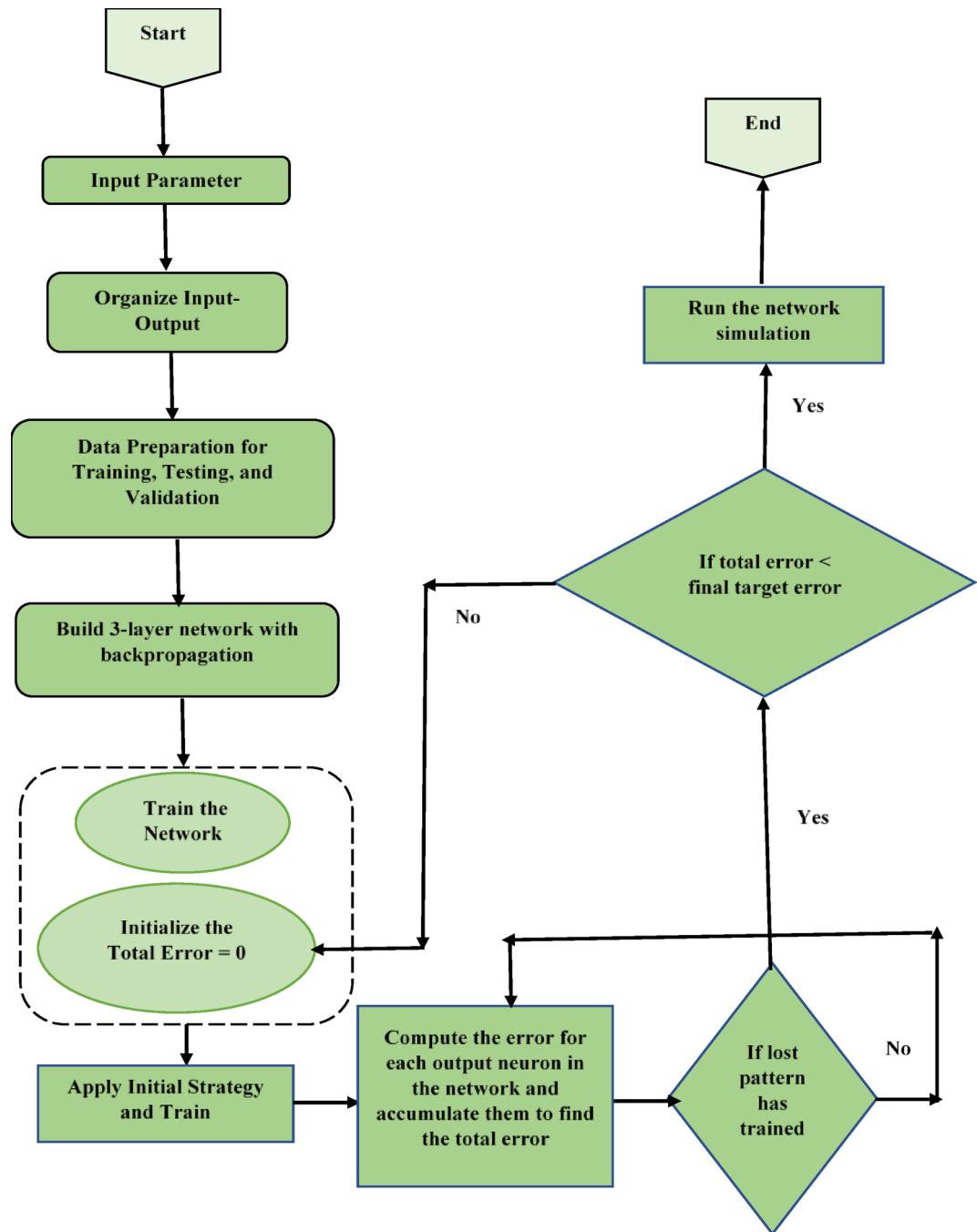
## Results and discussion

Figure 4 depicts the variation of various climatic parameters including maximum temperature, minimum temperature, precipitation and evapotranspiration from the year 2000 to 2024. The maximum temperature varies from  $-19.23$  to  $41.31$  °C in which the highest and lowest value of maximum temperature was observed on the 4th July 2012 and 2nd January 2014. The minimum temperature varies from  $-28.37$  to  $25.89$  °C in which the highest and lowest value of maximum temperature was observed on the 3rd July 2012 and 26th January 2019. The daily precipitation in the study area varies from 0 to 55.15 mm, in which the maximum precipitation was observed on 16th February 2018. The reference evapotranspiration was varied in the range of 0 to 10.2 mm, in which the highest value was found on 25th June 2012.

The observed yield data was collected from USDA NASS from 2000 to 2024. The yield of corn and soybeans were simulated for the growing season from May to September in Illinois, USA. The observed yield of corn and soybean varies from 7.06 ton/ha to 14.66 ton/ha and 2.49 ton/ha to 4.37 ton/ha respectively. Figure 5 depicts the land use land cover (LULC) map of Illinois, USA for the year 2023. From LULC map it has been found that corn and soybean crop area were 30.86% and 26.87% respectively. The minimum and maximum observed yield of corn was observed in 2012 and 2024 respectively, while the minimum observed yield of soybean was found in 2003 and maximum yield of soybean was found in 2021 and 2024. The yield of corn and soybean was estimated by using semi-physical model, AquaCrop model and ANN model for 25-year duration from 2000 to 2024 for Illinois, USA.

### Semi-physical model

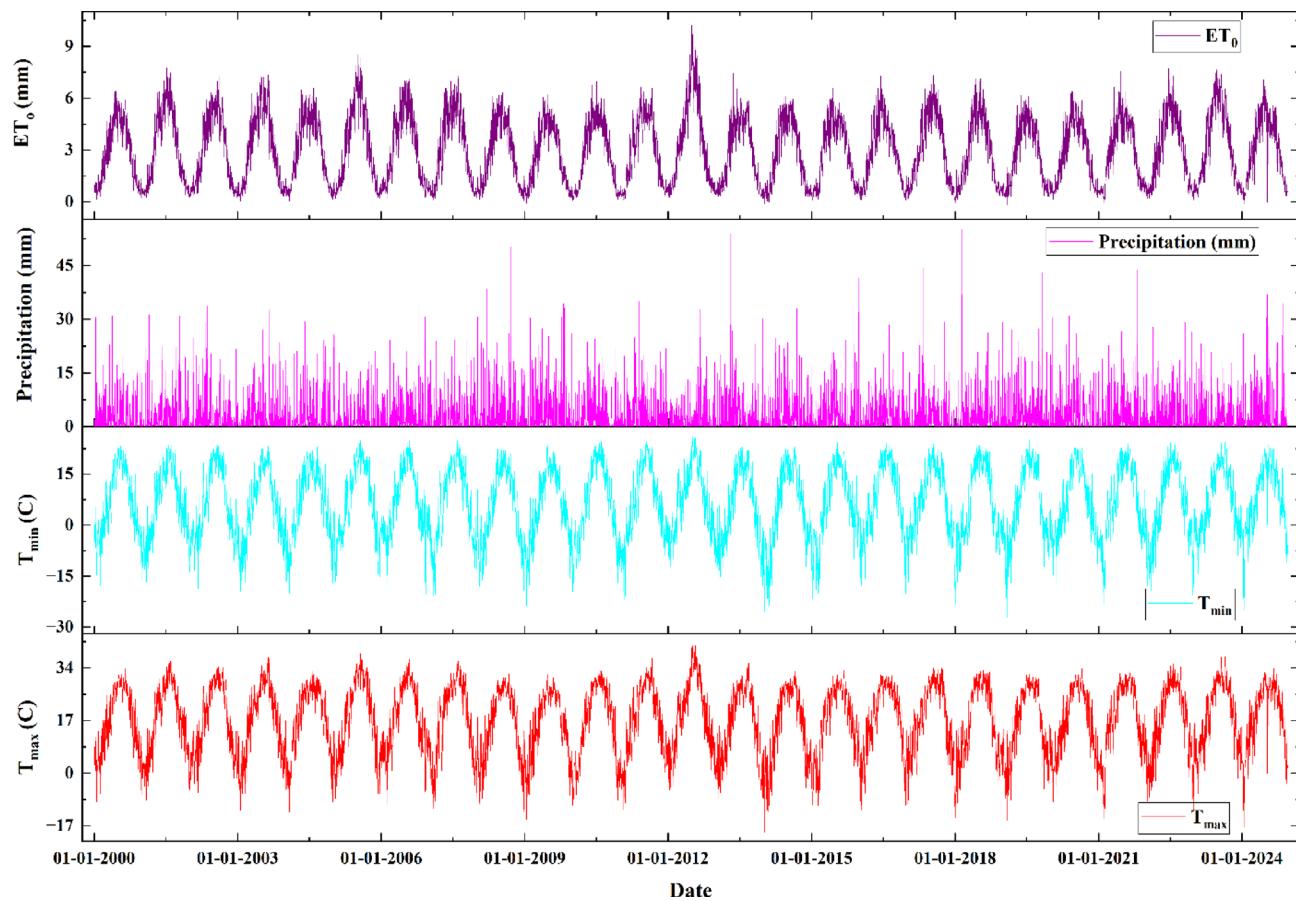
The yield of corn and soybeans was estimated for the 25-year duration from 2000 to 2024 by using the semi-physical model for the state of Illinois, USA. The assessed yield of corn varied from 9.01 to 13.42 ton/ha. The minimum yield of corn was found in 2012, whereas the maximum yield of corn was found in 2024.  $R^2$  value of 0.60 and RMSE value of 1.4377 ton/ha were obtained for the estimation of corn yield by using semi-physical



**Fig. 3.** Flowchart of ANN.

model which indicates the model's moderate performance of model for the corn yield estimation. The estimated yield of soybeans ranged from 2.92 ton/ha to 3.84 ton/ha. The maximum and minimum yield of soybeans were found in 2014 and 2012 respectively.  $R^2$  value of 0.42 and RMSE value of 0.4888 ton/ha were obtained for the estimation of soybean yield by using semi-physical model which indicates weaker performance of model for the estimation of soybean yield. Figure 6 depicts the regression plots of observed and estimated yield of corn and soybean.

PAR and LUE are the main factors influencing the yield estimations. PAR is a significant factor in promoting photosynthesis and LUE measures how well plants convert PAR to NPP. Consequently, changes in PAR and LUE influences spatial variation in yields. A mechanistic knowledge of the variables affecting agricultural production is provided by this semi-physical method, which also offers insightful information about the complex interactions between radiation, crop physiology, and yield results. Ji et al.<sup>50</sup> assessed the yield of wheat in China with  $R^2$  value of 0.64 and RMSE value of 0.948 ton/ha. Figure 7 shows the spatial variation of estimated yields of corn and soybean for the year 2024. The corn yield in North-East of Illinois was ranges from 11.81 ton/ha to 14.05 ton/



**Fig. 4.** Variation of Maximum Temperature, Minimum Temperature, Precipitation and Evapo-transpiration.

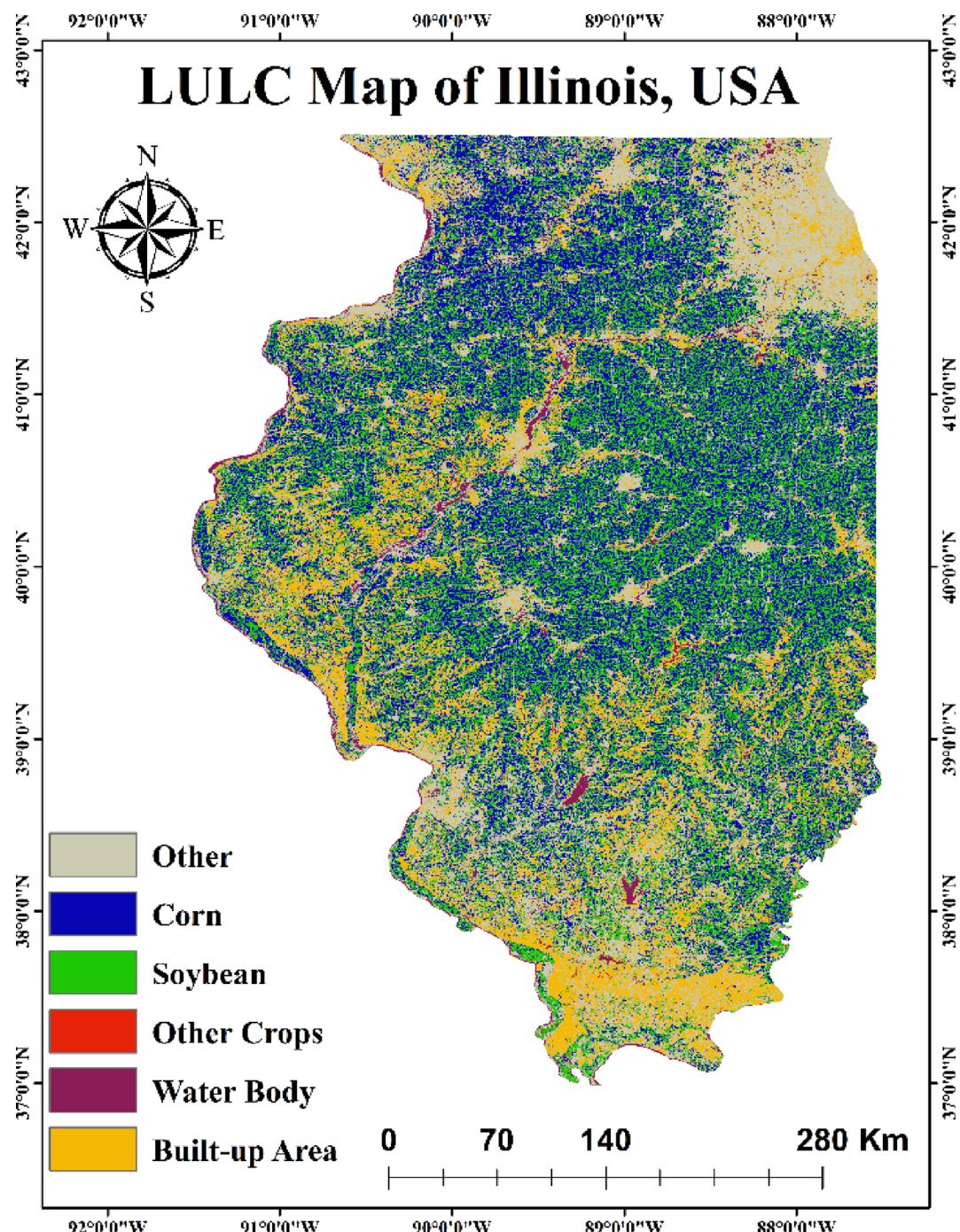
ha, while in North-West Illinois it varies from 13.33 ton/ha to 19.09 ton/ha. Wheras in South-East Illinois corn yield ranges from 3.6 ton/ha to 13.32 ton/ha and in South-West Illinois, it ranges from 3.6 ton/ha to 19.09 ton/ha. The soybean yield in North-East and North-West of Illinois was ranges from 1.19 ton/ha to 5.91 ton/ha, while in South-East and South-West Illinois it varies from 1.19 ton/ha to 4.23 ton/ha.

#### Estimation of yield using aquacrop model

The yield of corn and soybean was also estimated by using AquaCrop model which includes maximum and minimum temperature, precipitation, reference evapotranspiration, initial soil water content and soil characteristics. This study employs AquaCrop model for predicting the corn and soybean yield in the Illinois. The model uses 70% and 30% of the dataset was utilized for training and validation respectively. The model was calibrated and validated by using 18 and 7 year dataset respectively.

The estimated yield of corn by using AquaCrop model ranges from 7.60 ton/ha to 14.42 ton/ha. The minimum estimated corn yield assessed in 2005, which reflects a significant deviation from the observed value, which was 9.62 ton/ha. The maximum estimated corn yield was assessed in 2024 which closely aligned with the observed yield of 14.66 ton/ha, showing the model's accuracy in predicting higher yields in the more recent years. Figure 8 shows the regression curve of the AquaCrop model that was used in the present study to predict the corn yields during the training, and validation phases. The  $R^2$  value of 0.72 and RMSE value of 1.37 ton/ha were assessed for the estimation of corn yield by using AquaCrop model which indicates moderate model performance. The model achieved  $R^2$  values of 0.71, and 0.58 for calibration, and validation phases, respectively, indicating consistent and reliable performance across different stages. The model successfully predicted the general trends of corn yield improvement over time. Jin et al.<sup>51</sup> applied the AquaCrop model to simulate maize yield and found the  $R^2$  value for grain yield was 0.78 and RMSE was 1.44 ton/ha in Beijing, China.

The estimated soybean yields range from 2.80 ton/ha to 4.34 ton/ha. The minimum estimated soybean yield assessed in 2005, aligns closely with the observed yield of 3.13 ton/Ha, while the maximum estimated soybean yield in 2020 was very similar to the observed yield of 4.33 ton/Ha. This indicates that the AquaCrop model has been more effective in predicting soybean yields than corn yield in recent years. While earlier years showed some variability in estimates, AquaCrop model demonstrated improved accuracy in capturing soybean yield trends over time. Figure 9 shows the regression curve of the AquaCrop model to predict the soybean yields during the training, and validation phases. The regression plot for soybean indicates a strong correlation between predicted and observed yields, with an overall  $R^2$  value of 0.77 and RMSE value of 0.29 tons per hectare, demonstrating the model's precision. The model achieved  $R^2$  values of 0.68, and 0.58 for calibration, and validation phases,

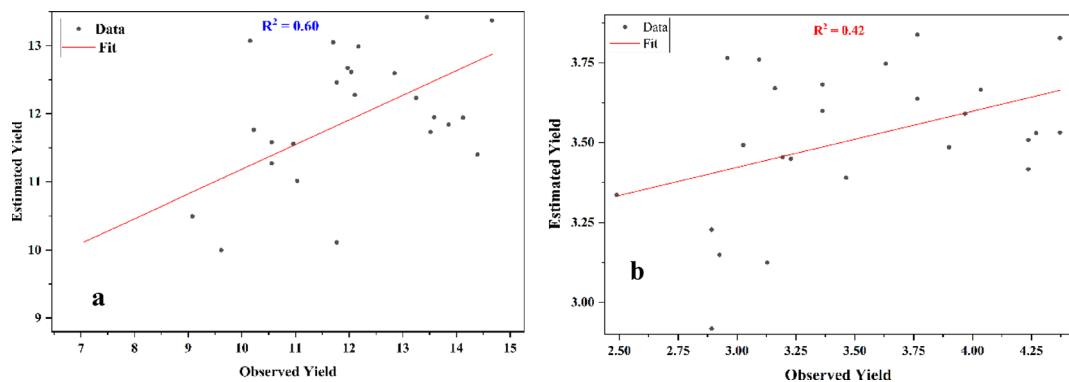


**Fig. 5.** Land Use Land Cover map of Illinois USA (created using QGIS software, version 3.40.7; available at <https://qgis.org/download/>).

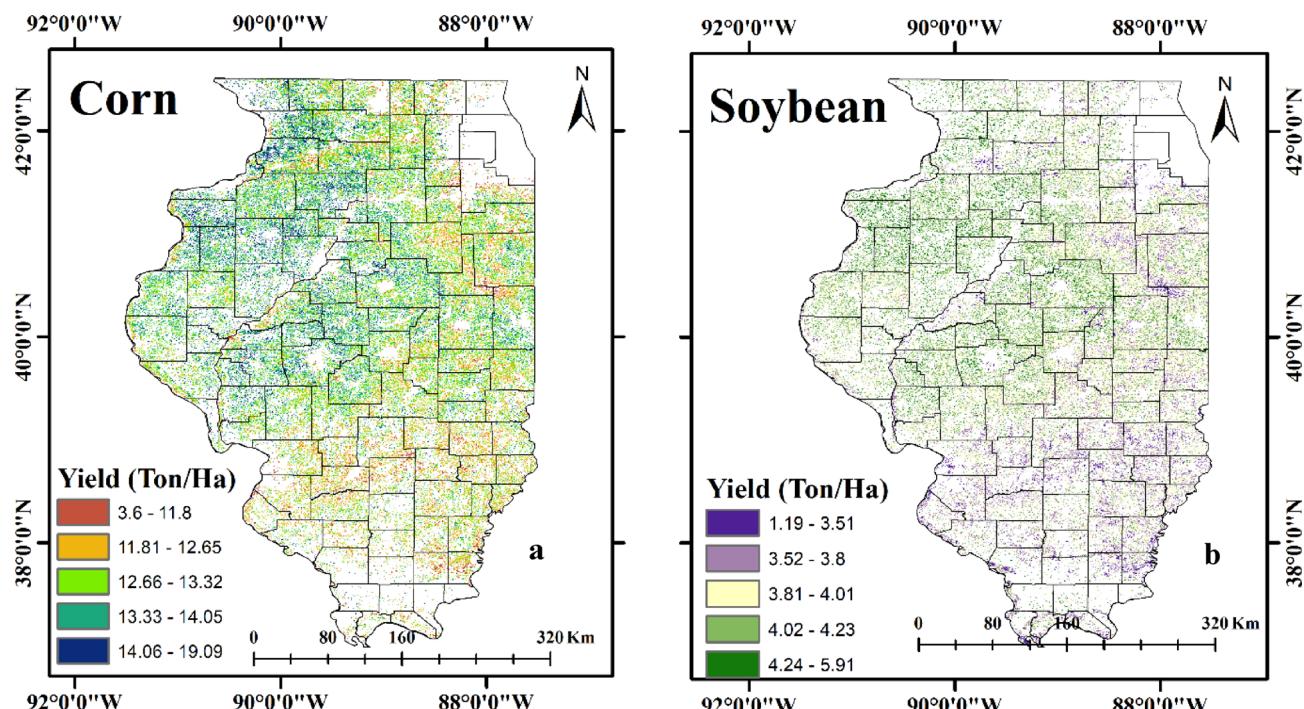
respectively, indicating consistent and reliable performance across different stages. These results demonstrate the model's strong predictive capabilities, though its performance could be further improved by including additional ground truth data and relevant parameters, thereby enhancing the overall accuracy and reliability of the crop yield predictions. Adeboye et al.<sup>52</sup> estimate the soybean yield in Nigeria and finds the  $R^2$  value of 0.99 and RMSE value of 0.03 ton/ha.

#### Artificial neural network model

ANN model including one input layer, ten hidden layers, and one output layer has been employed. The z-score technique was used to standardize the input data and provide values between 0 and 1. Nine and one neurons were selected for the input and output layers, respectively. After that, the best network design and data partitioning plan were determined using the neural network fitting tools in the MATLAB 2024b program.



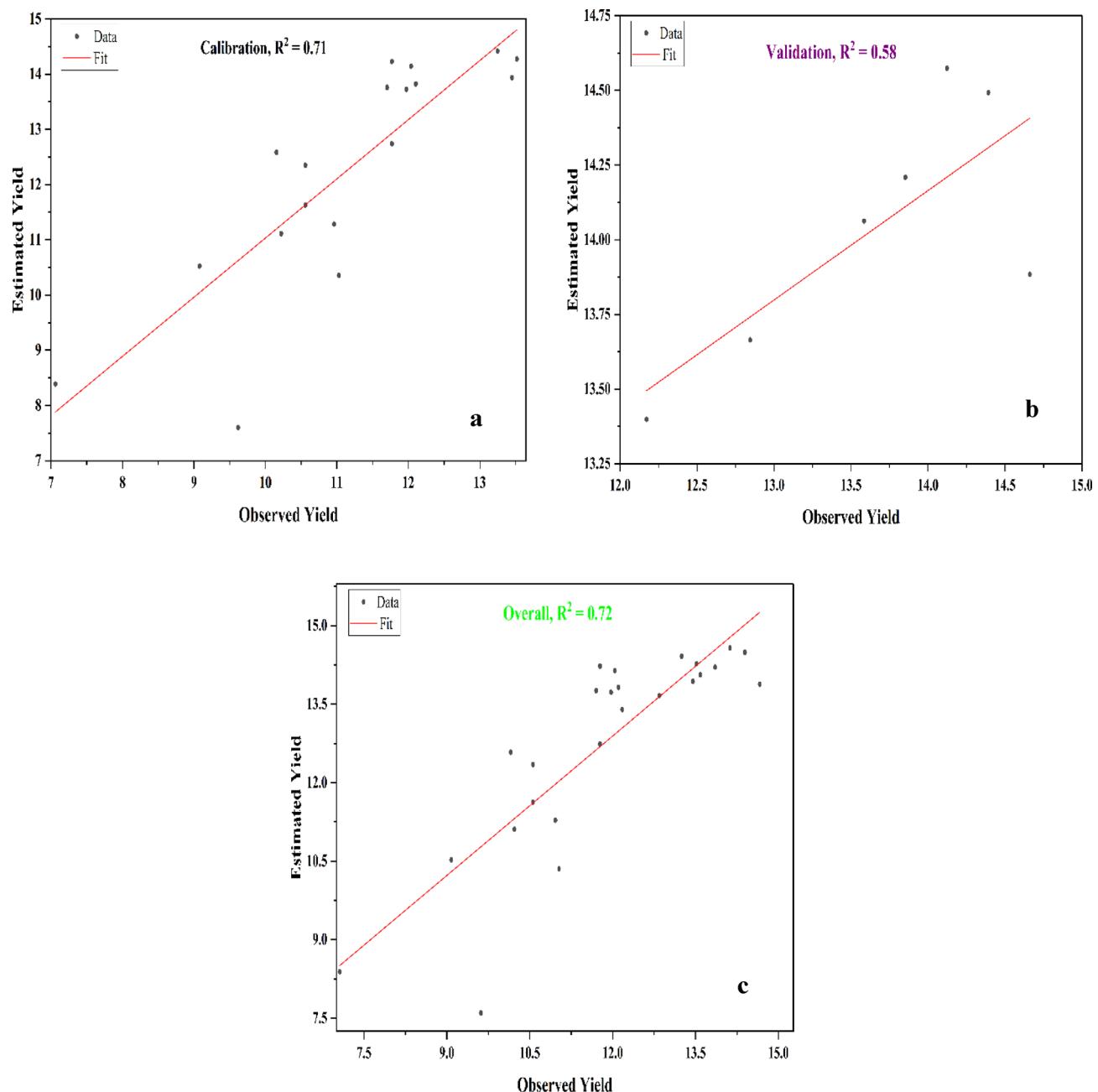
**Fig. 6.** (a-b): Regression Plots of Observed and Estimated Yields by using Semi-physical Model; a: Corn and b: Soybean.



**Fig. 7.** Estimated Yields in the year 2024 by using Semi-physical Model; a: Corn and b: Soybean (created using QGIS software, version 3.40.7; available at <https://qgis.org/download/>).

This serves as the main tool for identifying optimal network architecture at a later, more thorough investigation. According to the model used in this study, 70% of the dataset was utilized for training, and 15% was used for validation. However, testing has been conducted using the remaining 15% of the dataset. The yield was then calculated using the ANN model because it works well with this partitioning approach. The training outcomes validate that the ANN model with architecture 9–10–1 is the best combination for forecasting the research area's yield. We split the data samples for system validation testing, validation, and training in order to identify the best partitioning technique. Additionally, data was divided using a random technique that chose specific fundamental points from the whole dataset, which included data that had not been identified before. Figure 10 shows the performance plot for the estimation of corn and soybean yields. The performance plot highlights a convergence with the MSE reducing to 1.3194 at 6 epoch for the estimation of corn yield which signifies efficient learning whereas for the estimation of soybean yield, MSE decreases steadily to 0.039 at 3 epoch, which confirms the model's capability to adapt and capture soybean yield unpredictability.

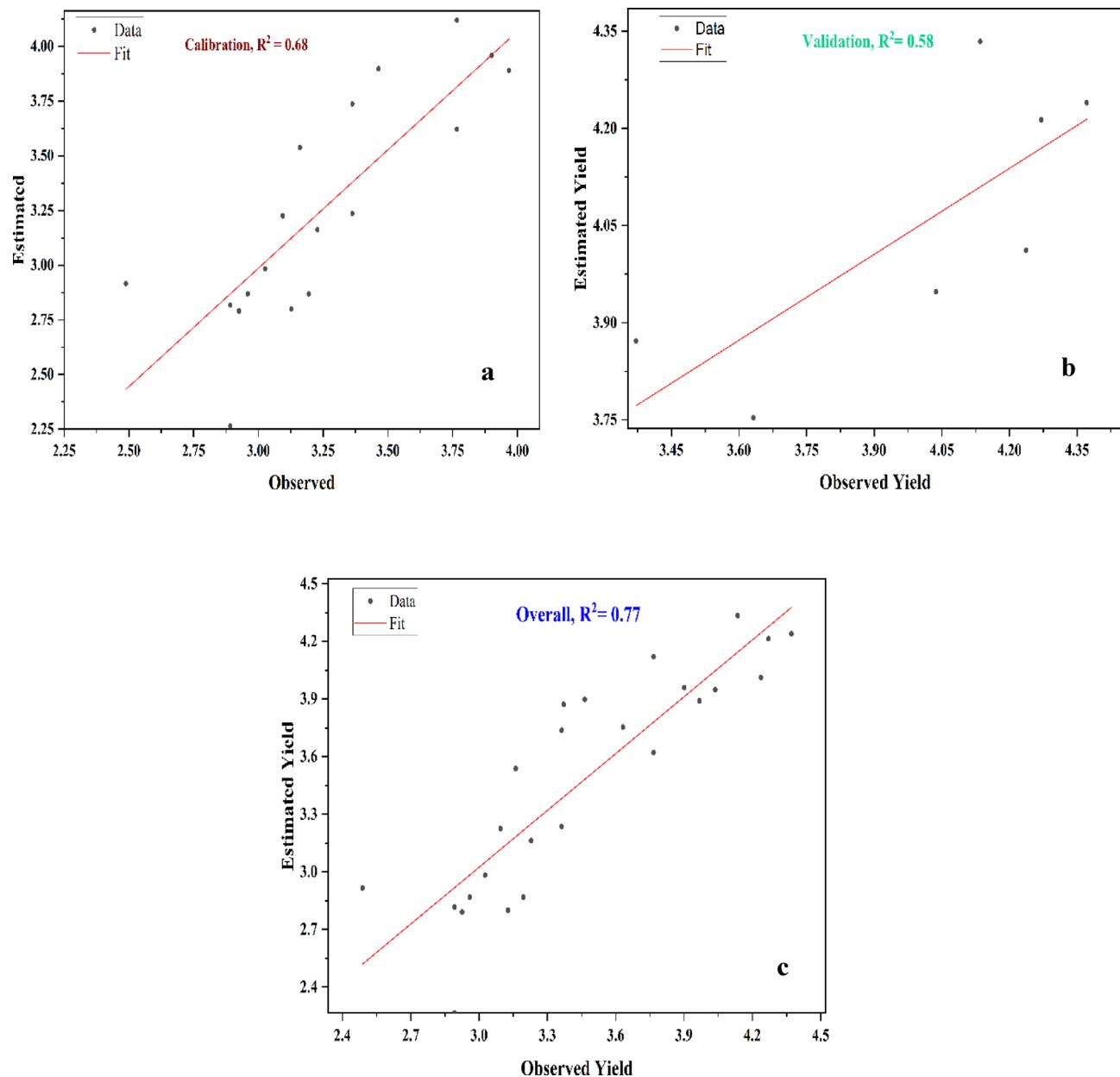
In contrast, we employ the linear regression approach to objectively and statistically assess validation results using both observed and estimated data. An additional criterion to assess the effectiveness of the final model is ordinary least square multiple linear regression. The rapid propagation approach was used to train yield predictions using a 9–10–1 network combination. The ANN network showed a significant positive and exceptionally good correlation ( $r=0.98$  at  $p=0.01$ ) with observed corn yield values in the training phase.



**Fig. 8.** (a-c): Regression Plots of Observed and Estimated Yields of Corn by using AquaCrop Model; a: Calibration, b: Validation and c: overall.

Figure 11 shows the regression curve of the ANN model that was used in the present study to forecast the corn yields during the training, validation, and testing phases. The regression plot for corn indicates a strong correlation between predicted and observed yields, with an overall  $R^2$  value of 0.89, demonstrating the model's precision. The model achieved  $R^2$  values of 0.98, 0.88, and 0.56 for training, validation, and testing phases, respectively, indicating consistent and reliable performance across different stages. Ma et al.<sup>53</sup> predicts the corn yield for the USA with  $R^2$  value of 0.77 and RMSE value of 1.03 ton/ha by using Bayesian neural network.

The ANN network showed a significant positive and exceptionally good correlation ( $r=0.97$  at  $p=0.01$ ) with observed soybean yield values in the training phase. Figure 12 shows the regression plot of the ANN model that was used in the present study to forecast the soybean yields during the training, validation, and testing phases. The regression plot for soybean indicates a strong correlation between predicted and observed yields, with an overall  $R^2$  value of 0.96, reflecting effective generality by the model. The model achieved  $R^2$  values of 0.97, 0.99 and 0.83 for training, validation and testing phases respectively, underscoring the model's capability to predict yields accurately. Najafabadi et al.<sup>54</sup> predicts the soybean yield in Canada with  $R^2$  value of 0.76 with RMSE value of 0.225 ton/ha. A comparative analysis indicates that the ANN model demonstrates slightly better performance



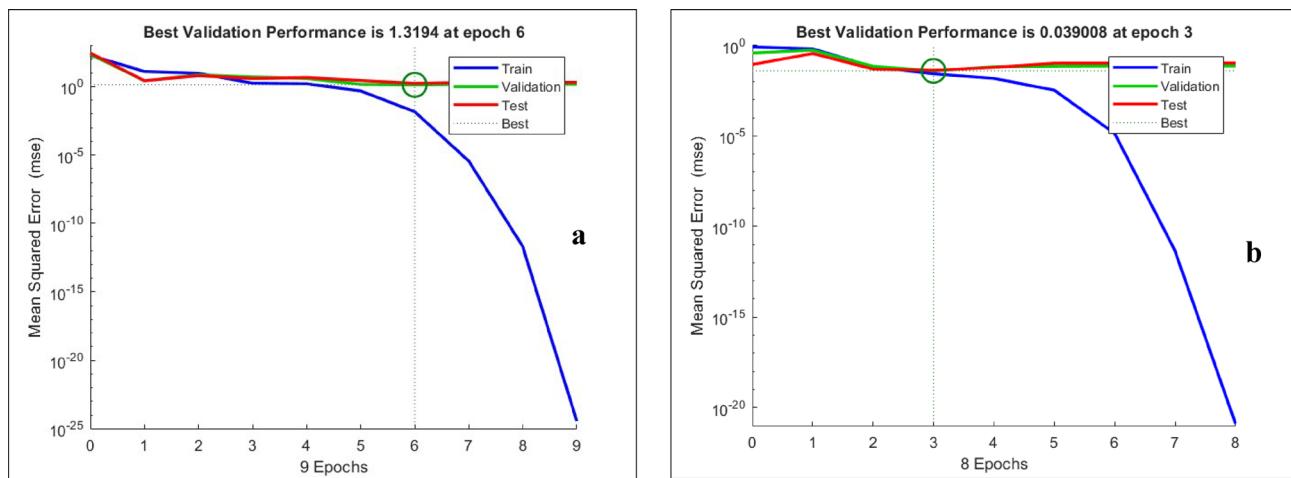
**Fig. 9.** Regression Plots of Observed and Estimated Yields of Soybean by using AquaCrop Model; a: Calibration, b: Validation and c: overall.

for soybean than corn, although ANN model achieves satisfactory accuracy for the prediction of both corn and soybean.

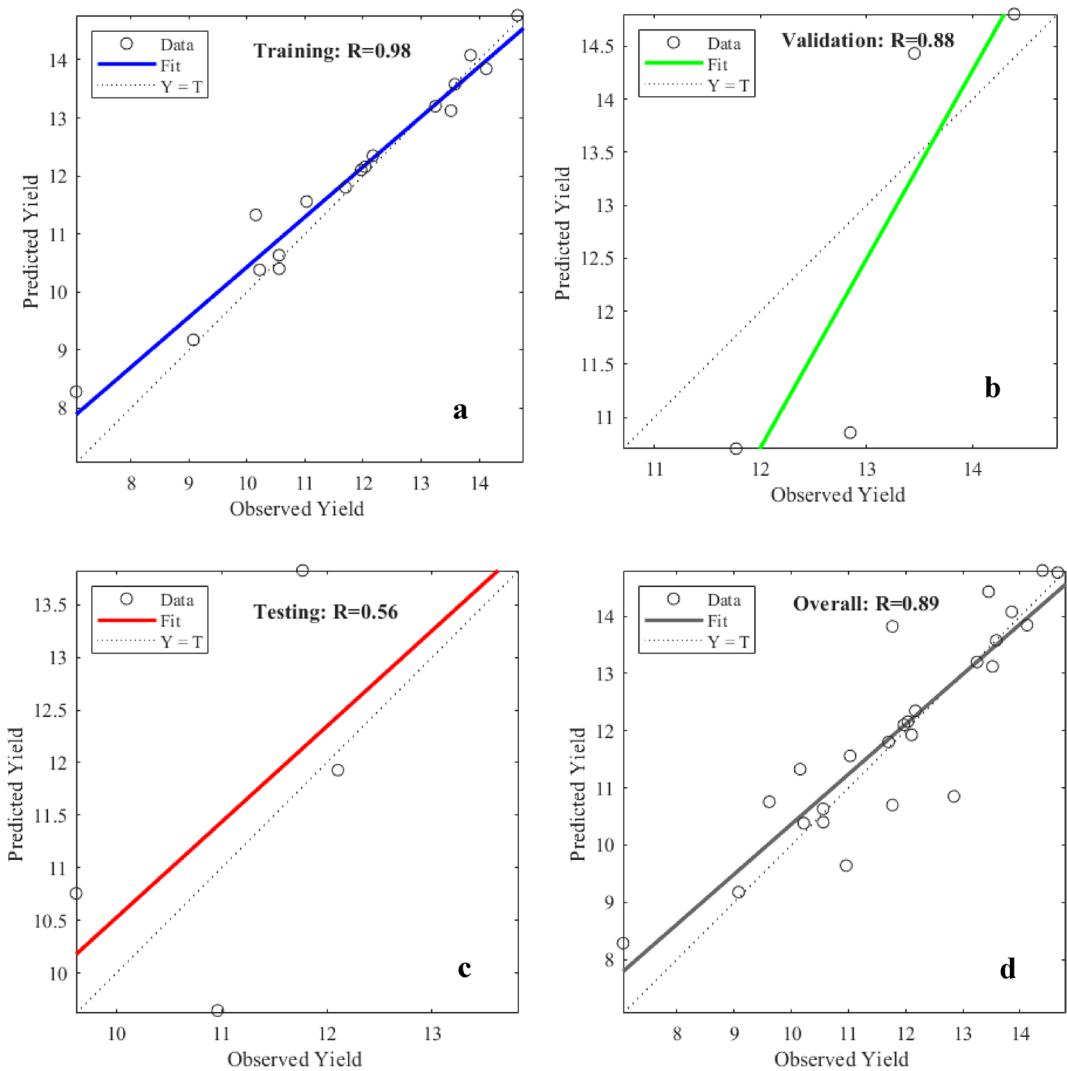
The ANN model predicted the corn yield that ranges from 6.81 ton/ha to 15.63 ton/ha, whereas ANN model predicted the soybean yield which ranges from 2.45 ton/ha to 4.43 ton/ha. The maximum predicted corn yield was found in 2014 and the minimum predicted corn yield in 2012. The maximum predicted soybean yield was observed in 2024 while minimum predicted yield of soybean was observed in 2003. The overall robust performance of the ANN models for both crops highlight their potential as reliable tools for yield prediction. These findings can support policymakers and farmers in making informed decisions regarding resource allocation and agricultural planning, ultimately contributing to sustainable agricultural practices.

#### Comparison between different models

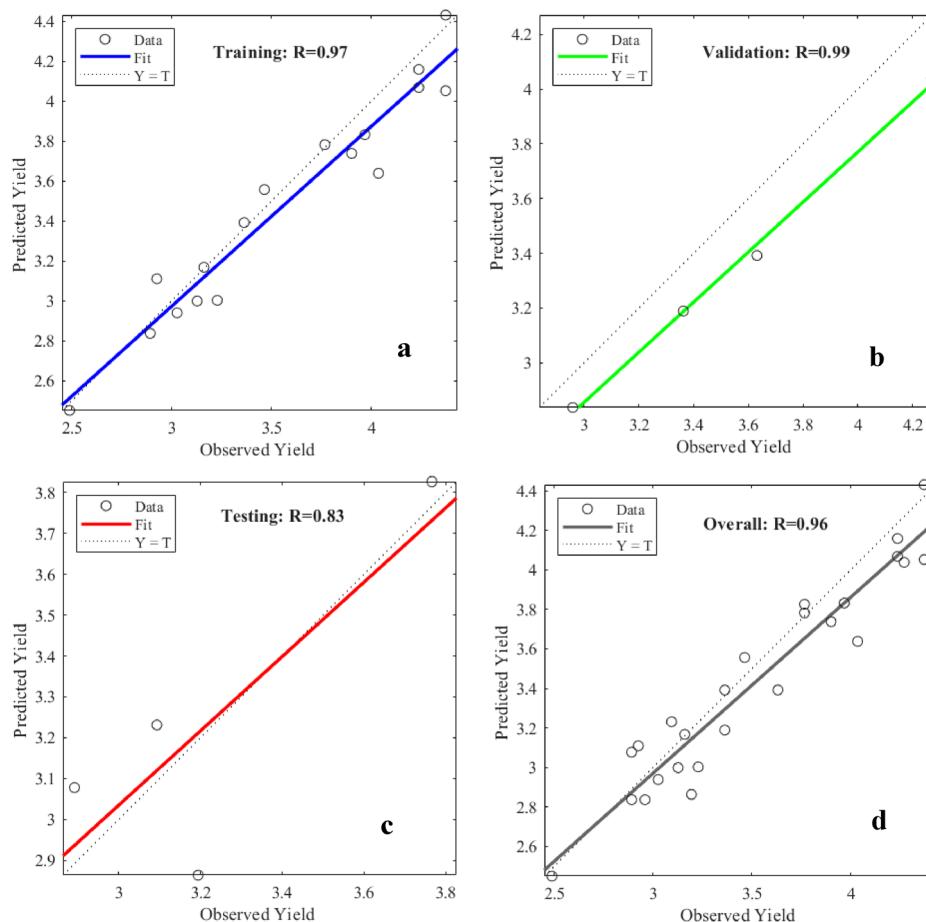
The yield of corn and soybean was estimated by using semi-physical model, AquaCrop model and ANN model for 25-year duration from 2000 to 2024 for Illinois, USA. The ANN model predicted the corn yield to range from 6.81 ton/ha to 15.63 ton/ha and the maximum predicted corn yield was found in 2014 and the minimum predicted corn yield in 2012. The assessed yield of corn by using semi-physical method varied from 9.01 ton/ha to 13.42 ton/ha in which the minimum yield of corn was found in 2012, whereas the maximum yield of corn



**Fig. 10.** (a-b): Performance Plot for the estimation of Yield; a: Corn and b: Soybean.



**Fig. 11.** (a-d): Regression Plots of Observed and Estimated Yields of Corn for the ANN Model; a: Training, b: Validation, c: Testing and d: Overall.

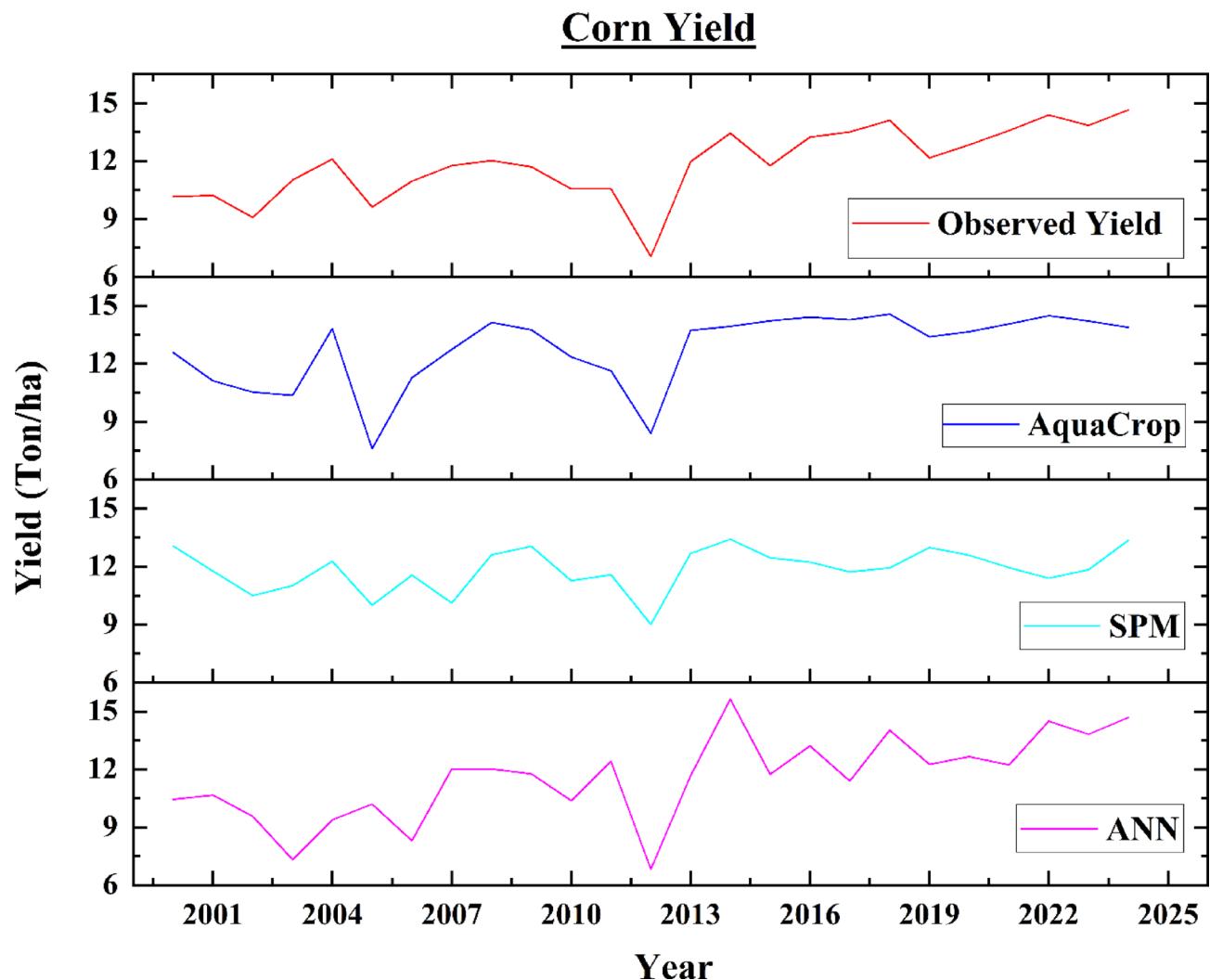


**Fig. 12.** (a-d): Regression Plots of Observed and Estimated Yields of Soybean for the ANN Model; a: Training, b: Validation, c: Testing and d: Overall.

was found in 2024. Estimated yield of corn by using AquaCrop model ranges from 7.60 ton/ha to 14.42 ton/ha in which the minimum estimated corn yield was found in 2005 and maximum estimated corn yield was found in 2024. The observed yield of corn ranges between 7.06 ton/ha to 14.66 ton/ha in which the minimum and maximum observed yield of corn was observed in 2012 and 2024 respectively. Figure 13 shows the variation of corn observed yield, and estimated yield from different models.

The ANN model predicted the soybean yield which ranges from 2.45 ton/ha to 4.43 ton/ha in which the maximum predicted soybean yield was observed in 2024 while minimum predicted yield of soybean was observed in 2003. The estimated yield of soybeans using semi-physical method ranges from 2.92 ton/ha to 3.84 ton/ha in which the maximum and minimum yield of soybeans were found in 2014 and 2012 respectively. The estimated soybean yields using AquaCrop model range from 2.80 ton/ha to 4.34 ton/ha in which the minimum estimated soybean yield was assessed in 2005, and the maximum estimated soybean yield in 2020, which was very similar to the observed yield of 4.33 ton/Ha. The observed yield of soybean varies from 2.49 ton/ha to 4.37 ton/ha in which the minimum observed yield of soybean was found in 2003 and maximum yield of soybean was found in 2021 and 2024. Figure 14 shows the variation of soybean observed yield, and estimated yields from different models.

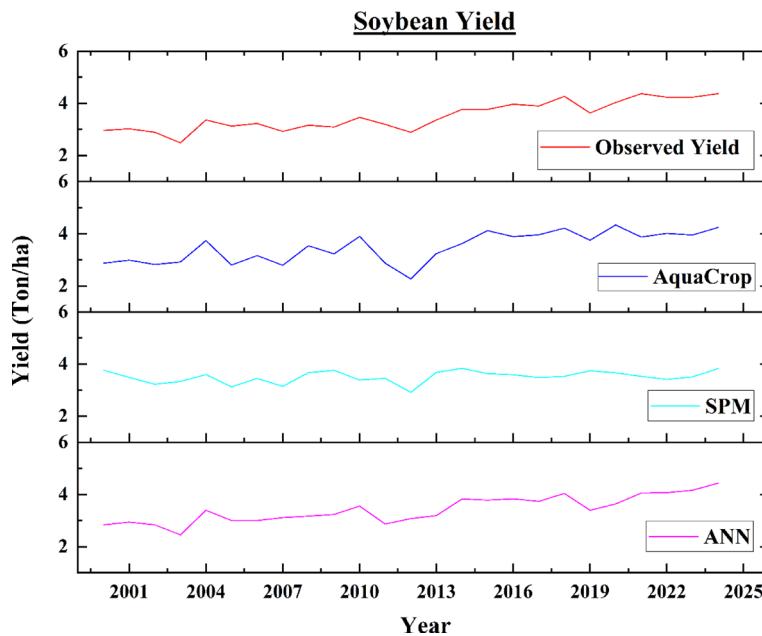
The comparison of the performance of AquaCrop model, semi-physical model, and ANN model for corn and soybean yield prediction highlights the higher accuracy of the ANN model. Table 2 depicts the  $R^2$  and RMSE values for different models. For the estimation of corn yields, ANN achieves the highest  $R^2$  value of 0.89 and semi-physical model achieves the lowest  $R^2$  value of 0.60, while ANN model gives the lowest RMSE value of 1.3194 ton/ha and semi-physical model gives the highest RMSE value of 1.4377 ton/ha. But for the estimation of soybean yields, ANN model gives the highest  $R^2$  value of 0.95 and semi-physical model gives the lowest  $R^2$  value of 0.42 while ANN model gives the lowest RMSE value of 0.039 ton/ha and semi-physical model gives the highest RMSE value of 0.4888 ton/ha. The ANN model achieved the highest  $R^2$  and lowest RMSE values for corn and soybeans. This indicated that ANN effectively captures complex, nonlinear relationships in crop yield predictions. High value of  $R^2$  for the ANN model confirms the accurate prediction of soybean yield. In this study we utilized the ANN models which gives the highest  $R^2$  values and minimum RMSE for both corn and soybean that indicate the relation between statistical and crop prediction is high interconnected. For better accuracy of AquaCrop we go for ground truth data and experimental data. Accurate forecasting of crop yields before harvest is of paramount importance, particularly in regions prone to the climatic uncertainties.



**Fig. 13.** Variation of Observed and estimated yields from different models for corn.

### Conclusion

The present study aimed to evaluate the effectiveness of three distinct crop yield prediction models: AquaCrop, semi-physical, and ANN models for predicting corn and soybean yield. The findings demonstrate the advantages and disadvantages of each strategy, showing that semi-physical models offer a compromise between computational efficiency and complexity, whereas AquaCrop offers a process-based understanding of crop reactions to environmental conditions. Conversely, when trained on large, high-quality datasets, ANNs demonstrated a high degree of prediction accuracy. The findings of the present study indicates that no single model is universally superior; rather than model selection should be driven by the availability of input data, computational resources, and the specific goals of the research. For crop climate interactions, AquaCrop is still a useful tool, while ANNs offer plenty of possibility for data-driven forecasting, particularly when there are big and varied datasets available. As an alternative, the semi-physical model is practical and captures important agronomic processes without requiring a lot of calibration. Hybrid modeling techniques that combine the predictive capabilities of ANNs with the mechanistic insights of AquaCrop should be investigated in future studies. Model accuracy may also be increased by integrating data from remote sensing and refining data collection techniques. Further improvements can be achieved by conducting field experiments on agricultural land to validate and refine the model's predictive abilities. Additionally, early yield prediction facilitates the development of contingency plans by governments to ensure reasonable redistribution of food resources during periods of shortage or scarcity. Farmers, agronomists, and legislators can all make better decisions to maximize agricultural productivity in Illinois and elsewhere by improving these prediction frameworks.



**Fig. 14.** Variation of Observed Yield and Estimated Yields from different models for soybean.

Model	Crop	R <sup>2</sup>	RMSE (ton/ha)
AquaCrop	Corn	0.72	1.37
	Soybean	0.77	0.28
SPM	Corn	0.60	1.4377
	Soybean	0.42	0.4888
ANN	Corn	0.89	1.3194
	Soybean	0.96	0.039

**Table 2.** Comparison between different models.

### Data availability

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Received: 15 February 2025; Accepted: 24 July 2025

Published online: 28 July 2025

### References

1. Davis, B. et al. *Estimating Global and country-level Employment in Agrifood Systems* (Food & Agriculture Org, 2023).
2. Rosenzweig, C. et al. Assessing agricultural risks of climate change in the 21st century in a global gridded crop model intercomparison. *Proc. Natl. Acad. Sci. USA.* **111**, 3268–3273 (2014).
3. Lobell, D. B. & Asseng, S. Comparing estimates of climate change impacts from process-based and statistical crop models. *Environ. Res. Lett.* **12**, 015001 (2017).
4. Hlatshwayo, S. I. et al. The determinants of crop productivity and its effect on food and nutrition security in rural communities of South Africa. *Front. Sustain. Food Syst.* **7**, 109133 (2023).
5. Nuno, D. B. & Baker, M. M. The determinants of agricultural crop productivity among smallholder households in Haramaya distinct, Eastern Ethiopia. *Grassr J. Nat. Resour.* **4**, 146–153 (2021).
6. Aggarwal, P. K. Determinants of crop growth and yield in a changing climate. In *Rainfed Agriculture: Unlocking the Potential* 73–80CABI, (2009).
7. Challinor, A. J. et al. Improving the use of crop models for risk assessment and climate change adaptation. *Agric. Syst.* **159**, 296–306 (2018).
8. Ray, D. K. et al. Climate change has likely already affected global food production. *PLoS One.* **14**, e0217148 (2019).
9. Kern, A. et al. Statistical modelling of crop yield in central Europe using climate data and remote sensing vegetation indices. *Agric. Meteorol.* **260**, 300–320 (2018).
10. Lobell, D. B. & Burke, M. B. On the use of statistical models to predict crop yield responses to climate change. *Agric. Meteorol.* **150**, 1443–1452 (2010).
11. Kuwata, K. & Shibasaki, R. Estimating crop yields with deep learning and remotely sensed data. In *IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)* 858–861 (IEEE, 2015). (2015).
12. Crane-Droesch, A. Machine learning methods for crop yield prediction and climate change impact assessment in agriculture. *Environ. Res. Lett.* **13**, 114003 (2018).

13. Beringer, T., Kulak, M., Müller, C., Schaphoff, S. & Jans, Y. First process-based simulations of climate change impacts on global tea production indicate large effects in the world's major producer countries. *Environ. Res. Lett.* **15**, 034023 (2020).
14. Lobell, D. B. & Gourdji, S. M. The influence of climate change on global crop productivity. *Plant. Physiol.* **160**, 1686–1697 (2012).
15. Singh, P. Crop models for assessing impact and adaptation options under climate change. *J. Agrometeorology* **25** (1), 18–33 (2023).
16. Timlin, D., Paff, K. & Han, E. The role of crop simulation modeling in assessing potential climate change impacts. *Agrosystems Geosci. Environ.*, **7**(1), 1–23 (2024).
17. Zhao, Y., Chen, S. & Shen, S. Assimilating remote sensing information with crop model using ensemble Kalman filter for improving LAI monitoring and yield Estimation. *Ecol. Model.* **270**, 30–42 (2013).
18. Katerji, N., Campi, P. & Mastorilli, M. Productivity, evapotranspiration, and water use efficiency of corn and tomato crops simulated by aquacrop under contrasting water stress conditions in the mediterranean region. *Agric. Water Manag.* **130**, 14–26 (2013).
19. Maniruzzaman, M., Talukder, M. S. U., Khan, M. H., Biswas, J. C. & Nemes, A. Validation of the aquacrop model for irrigated rice production under varied water regimes in Bangladesh. *Agric. Water Manag.* **159**, 331–340 (2015).
20. Qin, W., Wang, D., Guo, X., Yang, T. & Oenema, O. Productivity and sustainability of rainfed wheat-soybean system in the North China plain: results from a long-term experiment and crop modelling. *Sci. Rep.* **5**, 17514 (2015).
21. Amiri, E. Calibration and testing of the aquacrop model for rice under water and nitrogen management. *Commun. Soil Sci Plant. Anal.* **47**, 387–403 (2016).
22. Stevens, T. & Madani, K. Future climate impacts on maize farming and food security in Malawi. *Sci. Rep.* **6**, 36241 (2016).
23. Babel, M. S., Deb, P. & Soni, P. Performance evaluation of aquacrop and DSSAT-CERES for maize under different irrigation and manure application rates in the Himalayan region of India. *Agric. Res.* **8**, 207–217 (2019).
24. Jalil, A., Akhtar, F. & Awan, U. K. Evaluation of the aquacrop model for winter wheat under different irrigation optimization strategies at the downstream Kabul river basin of Afghanistan. *Agric. Water Manag.* **240**, 106321 (2020).
25. Zhu, X., Xu, K., Liu, Y., Guo, R. & Chen, L. Assessing the vulnerability and risk of maize to drought in China based on the aquacrop model. *Agric. Syst.* **189**, 103040 (2021).
26. Basso, B. & Liu, L. Seasonal crop yield forecast: methods, applications, and accuracies. *Adv. Agron.* **154**, 201–255 (2019).
27. Bolton, D. K. & Friedl, M. A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. Meteorol.* **173**, 74–84 (2013).
28. Jones, J. W. et al. Toward a new generation of agricultural system data, models, and knowledge products: state of agricultural systems science. *Agric. Syst.* **155**, 269–288 (2017).
29. Poorter, H., Böhler, J., van Dusschoten, D., Climent, J. & Postma, J. A. Pot size matters: a meta-analysis of the effects of rooting volume on plant growth. *Funct. Plant. Biol.* **39**, 839–850 (2012).
30. Kamilaris, A. & Prenafeta-Boldú, F. X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **147**, 70–90 (2018).
31. Food and Agriculture Organization of the United Nations (FAO). FAO Statistical Databases. FAOSTAT Jan. (2025). <https://faostat.fao.org/> (Accessed).
32. United States Department of Agriculture – National Agriculture Statistics Service (USDA-NASS). Quick Stats. USDA-NASS (2024). [https://www.nass.usda.gov/Quick\\_Stats/index.php](https://www.nass.usda.gov/Quick_Stats/index.php). Accessed January (2025).
33. Li, H., Di, L., Zhang, C., Lin, L. & Guo, L. Improvement Of In-season Crop Mapping For Illinois Cropland Using Multiple Machine Learning Classifiers. In 10th International Conference on Agro-geoinformatics (Agro-Geoinformatics) (pp. 1–6). IEEE (2022). (2022).
34. Na, S., Hong, S., Kim, Y. & Lee, K. Estimation of corn and soybean yields based on MODIS data and CASA model in Iowa and Illinois, USA. *Korean J. Soil. Sci. Fertilizer.* **47** (2), 92–99 (2014).
35. Potter, C., Klooster, S., Myneni, R. B. & Genovese, V. Terrestrial carbon sinks predicted from MODIS satellite data and ecosystem modeling. *Earth Observer.* **16** (2), 15–20 (2004).
36. Jha, M. N., Kumar, A., Dubey, S. & Pandey, A. Yield Estimation of rice crop using semi-physical approach and remotely sensed data. In: Jha, C. S., Pandey, A., Chowdary, V., Singh, V. (eds) *Geospatial Technologies for Resources Planning and Management* 331–349 (Springer International Publishing, 2022).
37. Xiao, X. et al. Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote Sens. Environ.* **89**, 519–534 (2004).
38. Steduto, P., Hsiao, T. C., Raes, D. & Fereres, E. AquaCrop—The FAO crop model to simulate yield response to water: I. Concepts and underlying principles. *Agron. J.* **101**, 426–437 (2009).
39. Hsiao, T. C. et al. AquaCrop—the FAO crop model to simulate yield response to water: III. Parameterization and testing for maize. *Agron. J.* **101**, 448–459 (2009).
40. Araya, A., Habtu, S., Hadgu, K. M., Kebede, A. & Dejene, T. Test of aquacrop model in simulating biomass and yield of water deficient and irrigated barley (*Hordeum vulgare*). *Agric. Water Manag.* **97**, 1838–1846 (2010).
41. Xu, Z. et al. Impacts of irrigated agriculture on food–energy–water–CO<sub>2</sub> nexus across metacoupled systems. *Nat. Commun.* **11**, 5837 (2020).
42. Karandish, F., Nouri, H. & Brugnach, M. Agro-economic and socio-environmental assessments of food and virtual water trades of Iran. *Sci. Rep.* **11**, 15022 (2021).
43. Pathak, S. et al. Social, economic and environmental assessment of urban sub-catchment flood risks using a multi-criteria approach: A case study in Mumbai city, India. *J. Hydrol.* **591**, 125216 (2020).
44. Dahikar, S. S. & Rode, S. V. Agricultural crop yield prediction using artificial neural network approach. *Int. J. Innov. Res. Electr. Electron. Instrum. Control Eng.* **2**, 683–686 (2014).
45. Al-Bashiti, M. K. et al. Predicting wildfire ember Hot-Spots on gable roofs via deep learning. *Fire* **7**, 153 (2024).
46. Gani, A., Singh, M., Pathak, S. & Hussain, A. Groundwater quality index development using the ANN model of Delhi metropolitan city, India. *Environ. Sci. Pollut. Res.* **30**, 1–16 (2023).
47. Khan, N. et al. Prediction of oil palm yield using machine learning in the perspective of fluctuating weather and soil moisture conditions: evaluation of a generic workflow. *Plants* **11**, 1697 (2022).
48. Joshi, B., Sudhakaran, P. & Varma, M. Comprehensive study of sequester-based carbon concrete in an acidic environment using artificial neural networks. *Asian J. Civ. Eng.* **26**, 207–220 (2025).
49. Aravind, K. S., Vashisth, A., Krishnan, P. & Das, B. Wheat yield prediction based on weather parameters using multiple linear, neural network and penalised regression models. *J. Agrometeorol.* **24**, 18–25 (2022).
50. Ji, F., Meng, J., Cheng, Z., Fang, H. & Wang, Y. Crop yield Estimation at field scales by assimilating time series of Sentinel-2 data into a modified CASA-WOFOST coupled model. *IEEE Trans. Geosci. Remote Sens.* **60**, 1–14 (2021).
51. Jin, X., Li, Z., Feng, H., Ren, Z. & Li, S. Estimation of maize yield by assimilating biomass and canopy cover derived from hyperspectral data into the aquacrop model. *Agric. Water Manag.* **227**, 105846 (2020).
52. Adeboye, O. B., Schultz, B., Adekalu, K. O. & Prasad, K. C. Performance evaluation of aquacrop in simulating soil water storage, yield, and water productivity of rainfed soybeans (*Glycine max L. merr*) in Ile-Ife, Nigeria. *Agric. Water Manag.* **213**, 1130–1146 (2019).
53. Ma, Y., Zhang, Z., Kang, Y. & Özdogan, M. Corn yield prediction and uncertainty analysis based on remotely sensed variables using a bayesian neural network approach. *Remote Sens. Environ.* **259**, 112408 (2021).
54. Yoosefzadeh-Najafabadi, M., Tulpan, D. & Eskandari, M. Using hybrid artificial intelligence and evolutionary optimization algorithms for estimating soybean yield and fresh biomass using hyperspectral vegetation indices. *Remote Sens.* **13** (13), 2555 (2021).

## Acknowledgements

We are grateful to the Civil Engineering Department, Indian Institute of Technology Ropar, Punjab for providing us with the resources and support throughout the research process. We are also thankful to the anonymous reviewers for their insightful and constructive suggestions, which have helped in improving the quality of the manuscript.

## Author contributions

Vishal Gautam and Abdul Gani has designed the research, including data collections, data analysis, manuscript writing and Shrey Pathak and Anoop Kumar Shukla helped with the supervision, conceptual correction and revision of manuscript.

## Funding

Open access funding provided by Manipal Academy of Higher Education, Manipal. This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit organizations.

## Declarations

### Competing interests

The authors declare no competing interests.

### Declaration of competing interest

The authors declare no conflict of interest.

### Additional information

**Correspondence** and requests for materials should be addressed to A.K.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025