



OPEN

Quantifying agricultural resilience under climate variability: a data-driven climate resilience index for European cereal systems

Anil Utku¹, Magdalena Radulescu^{2,3,4}, Abdulkadir Barut^{5,7,8}✉ & Hind Alofaysan⁶

Climate variability and extreme weather events pose an increasingly serious threat to agricultural productivity in Europe. Although many studies have focused on the mean yield response to climate change, relatively few have attempted to offer a comprehensive, multi-dimensional view of agricultural resilience, separating long-term productivity trends from short-term climate-induced fluctuations. In this paper, a data-driven Climate Resilience Index is proposed to assess the resilience of European cereal systems under climate stress. Annual yield time series are decomposed into trend and anomaly components to identify interannual deviations driven by climate variability. Machine learning and deep learning algorithms, including RF, CatBoost, CNN, LSTM, TCN, and a TCN-LSTM hybrid model, are used to assess the predictive validity of climate extreme indices and persistence patterns. Among the evaluated models, the TCN-LSTM architecture achieved the highest predictive performance with 0.8347 R^2 , outperforming standalone temporal and ensemble-based models. The PCA-based exposure framework explained 68.4% of the total climate variability within the first principal component, indicating a highly concentrated climate-stress structure. Feature importance analyses further revealed that lagged yield anomalies and rolling volatility indicators were the dominant predictors of resilience dynamics. The findings show that temporal models perform better than traditional ensemble methods, emphasizing the significance of multi-year recovery mechanisms in anomaly evolution. To provide a consistent measure of exposure, principal component analysis is used to reduce correlated climate indicators into a multi-dimensional climate stress index. The CRI captures these three essential elements: climate exposure, yield variability, and recovery. The results at the country level show substantial variability across Europe, indicating that while exposure intensity is important, recovery dynamics are also critical to understanding resilience. The results also show that there are structurally distinct resilience profiles, and robustness tests indicate that country rankings are not sensitive to different weights and recovery functions. The results clearly show that recovery and adaptive capacity are critical to agricultural sustainability amid growing climate uncertainty. The CRI developed in this paper provides a clear, empirically valid framework for comparing resilience in agricultural systems.

Keywords Agricultural sustainability, Climate resilience, Yield anomalies, Cereal production, Deep learning, Principal component analysis

Climate change has recently appeared as one of the most important structural issues facing the global agricultural sector¹. Climate change is described by the rise in temperature, the occurrence of weather-related disasters, changes in precipitation, and an increase in droughts, which are currently influencing crop productivity variability in different continents². In Europe, the growing of cereals, especially wheat, barley, and maize, is of

¹Computer Engineering Department, Faculty of Engineering, Munzur University, Tunceli 62000, Turkey. ²National University of Science and Technology Politehnica Bucharest, Bucharest, Romania. ³UNEC Research Methods Application Center, Azerbaijan State University of Economics (UNEC), Istiqlaliyyat Str. 6, Baku 1001, Azerbaijan. ⁴ARUCAD Research Centre, Arkin University of Creative Arts and Design, Northern Cyprus, Mersin, 10, Turkey. ⁵Harran University, Şanlıurfa, Turkey. ⁶Department of Economics, College of Business Administration, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia. ⁷Recep Tayyip Erdoğan University, Rize, Türkiye. ⁸ARUCAD Research Centre, Arkin University of Creative Arts and Design, Northern Cyprus, Mersin 10, Turkey. ✉email: kadirbarut@harran.edu.tr

prime significance for ensuring food security, livestock, and agro-industrial value chains³. Consequently, the resilience of cereal production systems to climate change has become a topical subject in scientific and policy circles.

Although different studies have investigated the effect of climate change on agricultural production, the majority of these studies have focused on the average effect of temperature or precipitation change on agricultural production^{4,5}. While these studies are extremely enlightening, they tend to ignore the distinction between technological change and climate-related variability in the short term⁶. Moreover, resilience is usually quantified solely based on exposure metrics, which tend to ignore the multidimensional nature of recovery processes. Thus, there is a need to adopt an integrated approach that considers climate exposure, production sensitivity, and recovery processes in a unified quantitative framework^{7,8}.

In this study, resilience is defined as production resilience, meaning the ability of cereal production systems to withstand any climatic disturbances, maintain functional yield stability, and restore themselves following yield disturbances. While the theoretical model for the study is based on the exposure-sensitivity-adaptive capacity approach of the Intergovernmental Panel on Climate Change, it defines resilience in terms of climate exposure, interannual yield volatility, and recovery from yield disturbance.

This paper addresses these research gaps by developing a Climate Resilience Index (CRI) for European cereal production systems based on a multi-step analytical framework. First, annual yield series are decomposed into trend and anomaly components to isolate climate-driven deviations from technological change. This detrending approach ensures that the resilience index is calculated from interannual variability instead of productivity change. Second, machine learning and deep learning approaches are employed to validate the predictive accuracy of climate-extreme variables and analyze the persistence of yield shocks. Third, a Principal Component Analysis (PCA) is used on the multivariate climate variables to build a statistically consistent exposure indicator. Fourth, the exposure indicator, yield variability, and recovery capacity are combined into a normalized composite indicator of agricultural resilience for each country.

The key contributions of the current research could be summarized as follows. Firstly, the proposed approach distinguishes between the long-term technological trend and the short-term climate-induced anomalies using an anomaly-based decomposition. Secondly, machine learning and deep learning methods were integrated to highlight the significance of extreme climatic events and their temporal dynamics. Thirdly, the PCA-based multivariate climate exposure index was estimated, and climate indicators were condensed into a statistically sound representation of climate stress. Lastly, the CRI approach integrates all three dimensions (exposure, volatility, and recovery) into a single resilience assessment framework.

In contrast to research that focuses on exposure as the key driver of vulnerability, this research emphasizes the critical role of recovery capacity and stability of production in determining resilience outcomes. The findings show that the adaptive mechanisms can reduce climate stress, which underlines the significance of improving institutions, technology, and infrastructure in agricultural policies.

The combination of anomaly decomposition, machine learning modeling, PCA, and composite index construction into a unified framework represents an advancement in the state-of-the-art of quantitative resilience analysis in agricultural systems. The CRI framework developed in this research is a novel, data-informed, and policy-relevant approach to assessing agricultural sustainability in the context of rising climate uncertainty. This section provides a review of the most pertinent literature on yield variability due to climate change, resilience assessment in agriculture, and climate-yield modeling using machine learning techniques, which underlie the methodology proposed in this paper.

Related works

The study on the effect of climate change on crop yield has been an ongoing process for the last two decades. The first large-scale study was conducted to understand the average yield response to temperature and precipitation change. Lobell and Field⁹ calculated the global yield response to warming and found that warming has a very negative effect on cereal yields. Lobell et al.¹⁰ concluded that heatwave events during the critical growth stages are responsible for a large share of yield reduction, particularly for wheat and maize. Similar findings were derived by Asseng et al.¹¹, who concluded that wheat yields are very sensitive to temperatures above the optimal level during the growing stage.

Apart from average yield response, there has been a rising trend in the interest of interannual variability and yield anomalies. Ben-Ari and Makowski¹² analyzed global crop yield variability and concluded that crop yields are not normally distributed but follow region-specific patterns affected by climatic extremes. Ray et al.¹³ further concluded that yield variability contributes to a large share of food production system instability, thus emphasizing the importance of anomaly modeling instead of average yield response.

Another area of research is the study of agricultural vulnerability and resilience using composite index methods. Smit and Wandel¹⁴ conceptualized vulnerability as a function of exposure, sensitivity, and adaptive capacity, which was later measured using composite indices. Wiréhn et al.¹⁵ analyzed the vulnerability of agriculture to climate change in Europe with composite indicators and found that weighting decisions have a significant impact on the outcome of ranking results. Williges et al.¹⁶ analyzed the adaptive capacity of European agriculture and found that institutional and structural differences have a significant impact on the differences in drought resilience among EU countries. These studies suggest the complexity of drought resilience and the need for appropriate index formulation and robustness analysis.

More recently, data-driven approaches have been explored to improve yield forecasting and climate impact analysis. Jeong et al.¹⁷ demonstrated that machine learning models, namely Random Forests (RF), are superior to regression models in capturing nonlinear climate-yield relationships on a continental domain. Khaki and Wang¹⁸ demonstrated that deep learning models, such as convolutional and recurrent neural networks, could improve the accuracy of yield forecasting by considering temporal and spatial dependencies.

The application of temporal deep learning models is particularly pertinent to the analysis of yield persistence and recovery. Shan et al.¹⁹ demonstrated that Temporal Convolutional Networks (TCNs) are generally more effective than traditional recurrent neural networks for most sequence modeling tasks, as they can effectively capture long-term effective memory using dilated convolutions. Hochreiter and Schmidhuber²⁰ first introduced Long Short-Term Memory (LSTM) networks to overcome long-term dependency issues in sequential data. Since then, LSTMs have been widely adopted in agricultural time series modeling.

Despite the increasing interest in applying machine learning to yield prediction, most existing studies tend to emphasize prediction accuracy rather than structural resilience evaluation. Few studies attempt to integrate anomaly decomposition, multidimensional exposure modeling, and composite resilience index computation into a unified framework. Moreover, although the vulnerability of European agriculture has been explored with indicator approaches^{15,16}, most existing studies tend not to involve machine learning validation to evaluate the relevance of climate extremes and persistence.

Existing research typically concentrates on two categories: climate-yield forecasting or multi-dimensional vulnerability assessments. Nevertheless, very few have attempted to include the anomaly decomposition of yields, multivariate climate exposure assessment, temporal validation of deep learning algorithms, and the construction of resilience indices within a single coherent framework^{17,18}. This paper aims to fill this knowledge gap by bringing together all four methodologies into a single climate resilience assessment framework for European cereals.

Material and method

This section presents the methodological framework that has been developed to construct and validate the CRI for European cereal production systems. The methodological framework is structured in a step-by-step procedure that integrates time series decomposition, machine learning modeling, multidimensional climate exposure definition, and composite index construction. First, annual cereal yield time series are decomposed to identify standardized yield deviations, thereby separating short-term climate variability from long-term technological change. Second, ensemble and deep learning models are employed to quantify the predictive weight of climate extreme indicators and temporal persistence patterns. Third, a PCA is carried out on correlated climate variables to derive a statistically consistent exposure index. Finally, exposure, yield variability, and recovery capacity are combined to construct a standardized composite CRI, and clustering and robustness analysis are carried out to identify structural resilience profiles and index stability. This integrated framework enables a dynamic, data-driven, and policy-oriented analysis of agricultural resilience to climate variability.

Dataset

This analysis will employ a harmonized panel dataset of European cereal production and climate variables from 1990 to 2022²¹. The agricultural production dataset was collected at the country-crop-year level, and it includes annual yield (YIELD) data for the key cereal crops. The dataset includes three key cereals that are produced in Europe, and these are wheat, barley, and maize (corn). The dataset includes data for several European countries, and it is annual. The completeness of the time series is shown in Fig. 1, which shows the number of annual data points available for each country. Most countries have had complete time series data points throughout the analysis.

The distribution of observations by crop type is presented in Fig. 2, which gives a fair representation to enable cross-crop comparison in an equal analytical framework.

In addition to agricultural yield data, the study employs climate-related variables derived from the World Bank Climate Change Knowledge Portal (WB_CCKP). Variables that are employed include a group of standardized climate extreme indices (for example, temperature and precipitation anomalies, drought indicators), which enable multidimensional climate exposure analysis. The integration of agricultural and climate data enables the

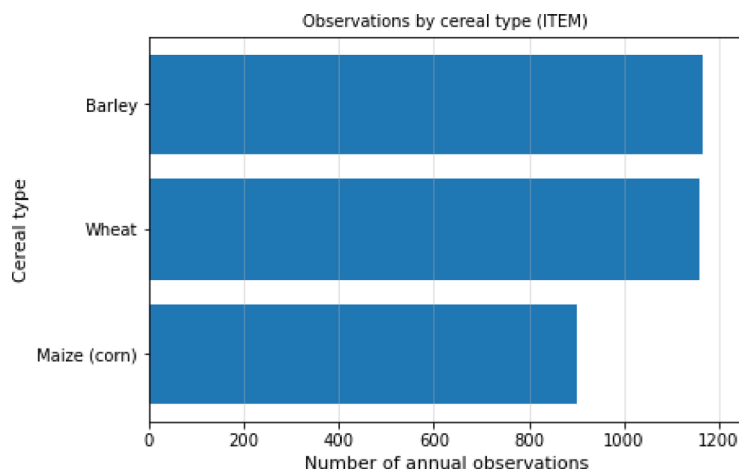


Fig. 1. The temporal completeness of the panel structure.

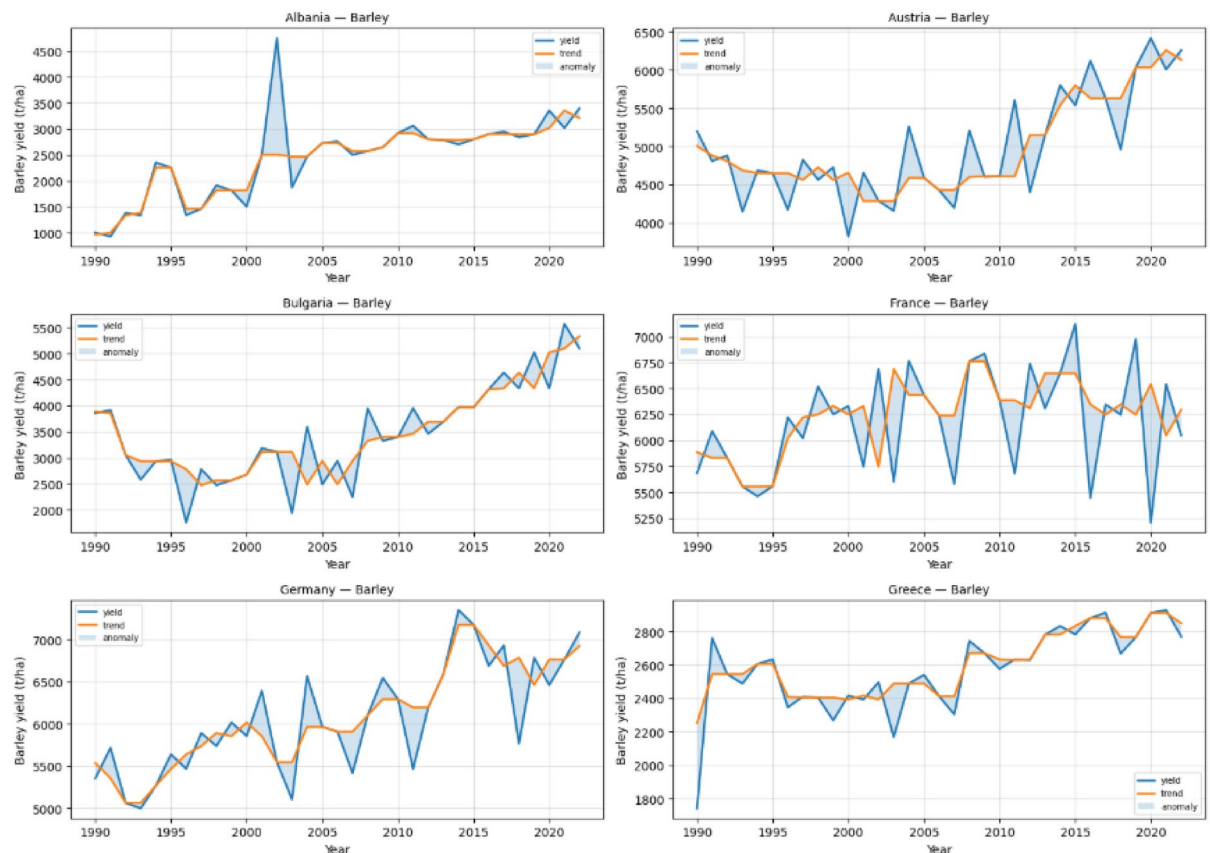


Fig. 2. The number of available yearly observations per country.

creation of a country-level climate-agriculture panel dataset that can be analyzed using machine learning, trend analysis, and resilience index.

Variables were formatted to produce a country-item-year panel. Missing yield data was removed to ensure consistency in time series analysis, while climate variables were standardized in the analysis. The dataset is the empirical basis for (i) yield trend decomposition, (ii) anomaly detection, (iii) exposure index construction through PCA, and (iv) CRI construction.

Data preprocessing and hyperparameter configuration

Before proceeding with the estimation of the models, a preprocessing procedure was established to ensure consistency and comparability of data over time and across countries. The yield data was matched with the climate indicators at the country-crop-year level, and a harmonized panel dataset was created. Missing values in the yield data were deleted to maintain the integrity of the time series at the country level, while the climate indicators were maintained as reported in the WB_CCKP database. As the dependent variable was specified as standardized yield anomaly after trend decomposition, no additional transformation was required for the target variable. Predictor variables were scaled using statistics calculated solely from the training data to avoid look-ahead bias. These scaling parameters were then used to standardize validation and test datasets to maintain strict out-of-sample prediction. Outliers were deliberately retained, as extreme negative values correspond to climate shock events of primary interest in resilience analysis.

The data were temporally divided to maintain causal order. The model was trained on data from 1990 to 2016, validated on data from 2017 to 2019, and tested on data from 2020 to 2022. This approach avoids information leakage and allows the model to be evaluated based on its performance in a real-world forecasting scenario, rather than a random cross-sectional split.

Hyperparameters of ensemble models were determined using randomized search algorithms and validation-based optimization. In the RF model, a large number of trees were employed to reduce the variance associated with ensemble averaging, and tree depth was allowed to be sufficiently flexible to represent nonlinear relationships between climate and yield without over-pruning. In gradient boosting models, such as XGBoost and CatBoost, learning rates were set conservatively to improve generalization performance, and tree depths were moderately constrained to avoid overfitting. Subsampling methods were used in boosting models to improve robustness, and early stopping based on validation loss was employed to limit the number of boosting iterations.

Deep learning models were trained using the Adam optimizer with mean squared loss as the objective function. Model training was performed for a maximum number of epochs with early stopping when validation loss did not improve for a specified number of epochs. This ensured that the model converged to a set of optimal

parameters without overfitting. The batch sizes were chosen to ensure a balance between computation and the stability of the gradient.

For sequence models, the rolling time windows were chosen to identify multi-year patterns of anomaly behavior. The sequence size was chosen to identify the medium-term effects of agricultural memory without going into the high-dimensional domain. The input variables were normalized to ensure stable gradient updates and faster convergence of the neural network.

The Convolutional Neural Network architecture was selected to detect local temporal information with one-dimensional convolutional layers and causal padding, followed by pooling and dense layers for regression tasks. The LSTM architecture was selected to detect long-term temporal patterns with multiple recurrent layers, along with dropout to mitigate overfitting. The Temporal Convolutional Network architecture employed dilated convolutions and residual connections to expand the receptive field. The hybrid TCN-LSTM model combined hierarchical feature extraction using dilated convolutions with recurrent memory processing to facilitate multi-scale temporal representation and long-term dependency modeling.

In all deep learning architectures, regularization methods such as dropout and validation-based early stopping were used. The choice of hyperparameters was based on validation performance rather than test-set optimization to avoid bias in the assessment. The approach to preprocessing and hyperparameter selection ensures methodological transparency, replicability, and comparability of ensemble and deep learning models with respect to the temporal nature of yield anomaly dynamics.

Trend decomposition and yield anomaly definition

To separate the short-term climate effects from the long-term technological change in agricultural production, the annual yield data were decomposed into trend and residual components. The original yield time series for selected country-crop combinations are shown in Fig. 3, indicating a slow-growing trend in agricultural production over the period of analysis.

This trend is mainly driven by technological advancements in agricultural production, and not by short-term climate change.

To eliminate this trend component, we used a Seasonal-Trend decomposition technique based on Locally Weighted Regression (STL). For each country-crop time series, yield $Y_{c,i,t}$ was decomposed as: $Y_{c,i,t} = T_{c,i,t} + R_{c,i,t}$ where $T_{c,i,t}$ represents the smooth long-term trend and $R_{c,i,t}$ denotes the residual (anomaly) component. Since the data are annual and do not exhibit intra-annual seasonality, the trend and anomaly components primarily focus on the trend and irregular components.

While STL can be effectively used to estimate seasonal cycles in seasonal data, the locally weighted regression approach used in STL makes it an effective means of separating out smooth, nonlinear trend elements from rough yearly fluctuations in annual data as well. The current analysis is not concerned with using STL to extract seasonality, but rather with separating the underlying long-term technology-driven trends from climate-induced short-term fluctuations in yield time series. Linear detrending may fail in this aspect, since trends can follow any pattern over time as long as technology, policies, and other factors continue to change.

The trend and anomaly components are shown in Fig. 4. The shaded areas in the figure show the deviations of the observed yield from the trend.

To make the residuals comparable across different countries and crops with varying levels of yield, the residuals were converted to z-scores as seen in Eq. 1:

$$Z_{c,i,t} = \frac{R_{c,i,t} - \mu_{c,i}}{\sigma_{c,i}} \quad (1)$$

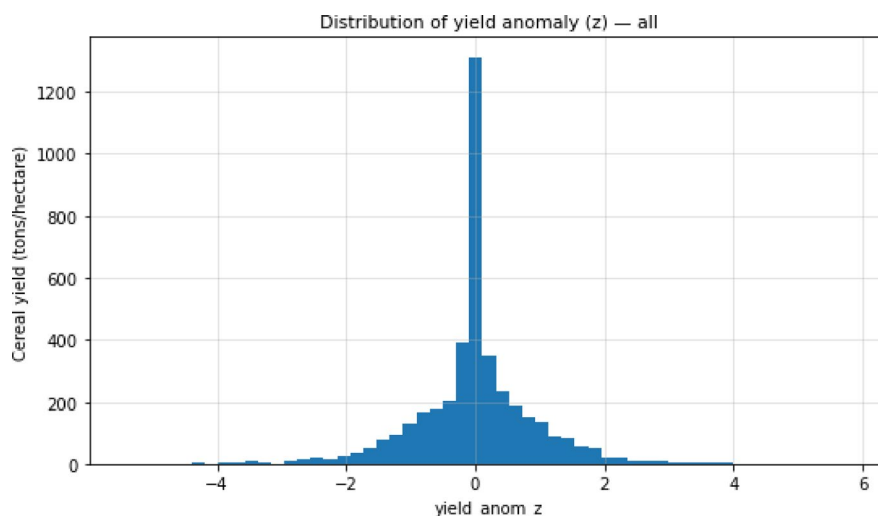


Fig. 3. The original yield time series for selected country-crop combinations.

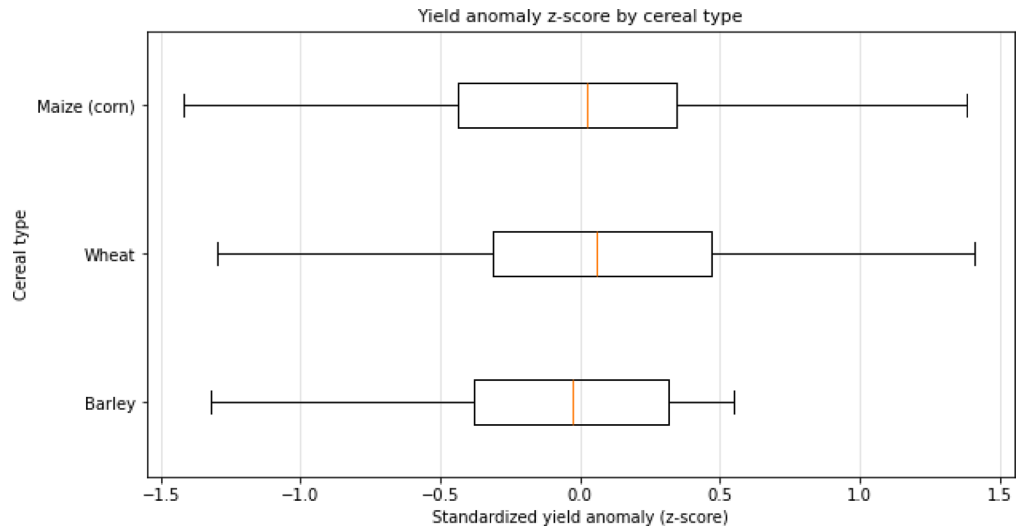


Fig. 4. Distribution of standardized yield anomalies (z-scores) across cereal types.

Where $\mu_{c,i}$ and $\sigma_{c,i}$ denote the mean and standard deviation of residuals for each country–crop pair. The standardized anomaly $Z_{c,i,t}$ serves two purposes. It enables the detection of strong negative yield shocks (for example, $Z < -1$). It enables a scale-independent measure for the analysis of resilience and recovery afterwards. This detrending method ensures that the CRI does not mix technological growth with structural changes and climate-driven volatility. The approach aims to analyze the sensitivity and recovery properties of agricultural systems to climate variability, rather than their productivity improvements.

Prediction models

Let $y_{i,t}$ denote the standardized yield anomaly for country i at time t . The prediction models aim to estimate future yield anomalies as a function of lagged yield dynamics, climate exposure indicators, volatility measures, and climate extreme variables:

$$y_{i,t} = f(y_{i,t-1}, y_{i,t-2}, V_{i,t}, E_{i,t}, C_{i,t}) + \varepsilon_{i,t} \quad (2)$$

where $y_{i,t-1}$ and $y_{i,t-2}$ represent lagged anomaly variables, $V_{i,t}$ denotes rolling volatility indicators, $E_{i,t}$ represents PCA-based climate exposure variables, $C_{i,t}$ includes climate extreme indicators, and $\varepsilon_{i,t}$ is the residual error term.

To capture the nonlinear and temporal relationships between climate indicators and standardized yield anomalies, an integrated machine learning framework was designed that combined tree-based ensemble models with sequence-driven deep learning models. The overall modeling pipeline is illustrated in Fig. 5.

Figure 5 presents the relationships among the climate, volatility, and anomaly predictor variables. To make the model easier to understand, the complete definitions of all variables are presented in Table 1.

According to the RF feature importance results, the lagged anomaly and rolling volatility features were the most influential drivers of future yield anomalies. The significant contribution of the first-order lagged anomaly is especially informative, as it implies high levels of persistence in agricultural production systems under climate change conditions.

Apart from the anomaly features, a wide range of other climate extremes, such as the length of warm spells, number of dry days, precipitation extremes, and temperature exposure indices, also contribute substantially to model performance. These empirical observations confirm that agricultural resilience is a multidimensional process whereby not only the level of climatic stress but also the recovery dynamics become important.

In summary, the structure of feature importance indicates that the influence of persistence-related features outweighs that of short-term climate shocks, implying that recovery potential and cumulative climatic stress become key factors in resilience behavior.

The predictive target was specified as the standardized yield anomaly $Z_{c,i,t}$ which is based on the detrended residual component described in "Data preprocessing and hyperparameter configuration". The list of explanatory variables includes climate extreme indices extracted from the WB_CCKP dataset, lagged yield values (Y_{t-1} , Y_{t-2}), lagged anomaly terms (Z_{t-1} , Z_{t-2}), rolling volatility estimates (3-year moving standard deviation).

The data was split along the time dimension to maintain temporal causalities: training set: 1990–2016; validation set: 2017–2019; test set: 2020–2022. This data splitting ensures the prevention of information leakage and assists in obtaining realistic estimates for the out-of-sample data. Three regression models using ensemble techniques were used: RF, Extreme Gradient Boosting (XGBoost), and CatBoost. RF was used to explore nonlinear relationships between climate variables and yield anomalies, simultaneously providing interpretable results for feature importance. XGBoost and CatBoost, being gradient boosting models, were also explored for their outstanding performance in predictive modeling of structured tabular data. Hyperparameters were

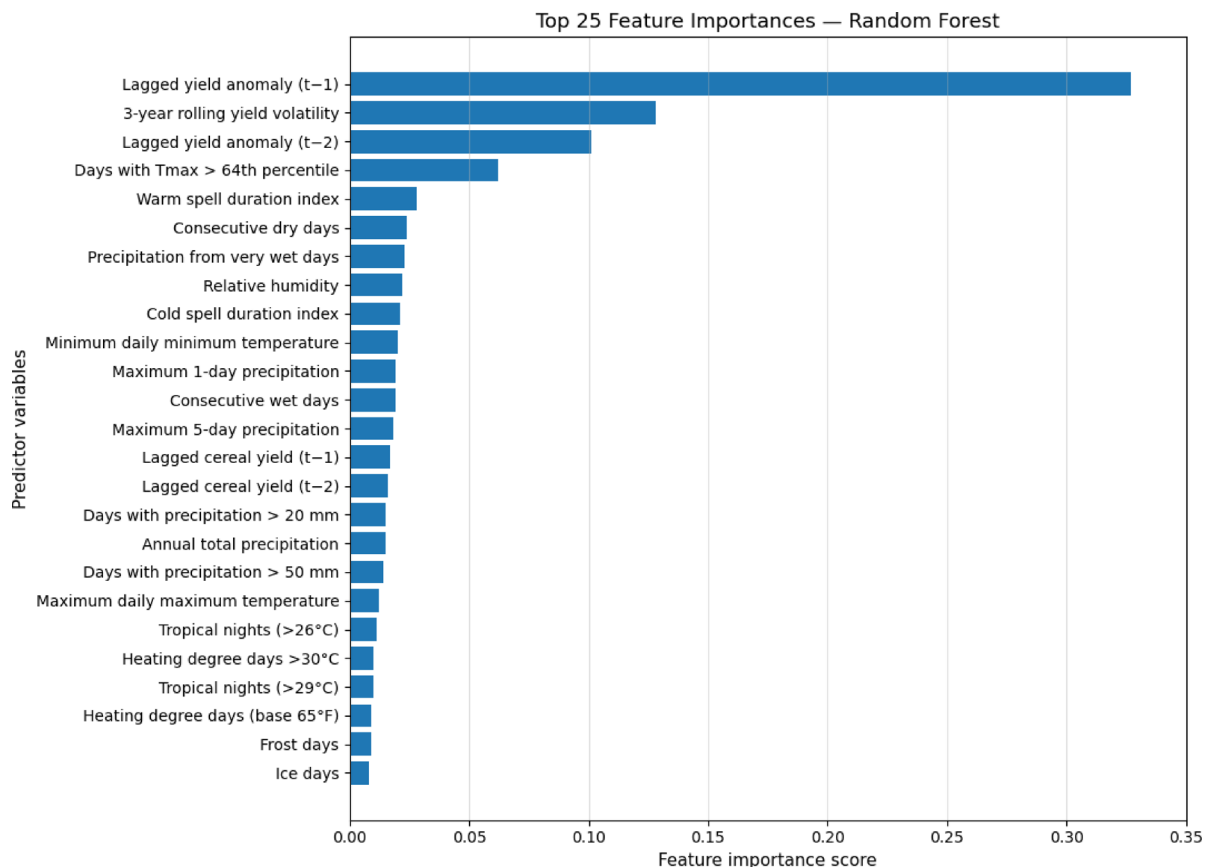


Fig. 5. Top 25 predictor variable importances obtained from the RF.

Variable	Description
TXx	Annual maximum of daily maximum temperature
TNn	Annual minimum of daily minimum temperature
PRCPTOT	Annual total precipitation
CDD	Consecutive dry days
SPI	Standardized precipitation index
yield_roll_std_3	3-year rolling yield volatility
anom_lag1	Yield anomaly at t-1
anom_lag2	Yield anomaly at t-2

Table 1. Description of predictor variables used in the modeling framework.

optimized using randomized search with cross-validation on the training set. Early stopping was used for boosting models based on validation loss to prevent overfitting.

Model performance was evaluated using Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Coefficient of Determination (R^2), Mean Absolute Percentage Error (MAPE), Symmetric MAPE (sMAPE), Normalized RMSE (nRMSE), and Pearson and Spearman correlation coefficients. The use of multiple metrics can help overcome scale and asymmetry problems.

RF is an ensemble learning algorithm that combines bagging and decision trees²². It creates a set of decision trees using bootstrapped samples and random feature subsets, and the prediction is done by taking an average of the predictions²³. RF is very effective for handling nonlinear relationships and interactions between climate variables without making strong parametric assumptions. Ensemble-based models such as RF and CatBoost estimate nonlinear mappings between predictor variables and future yield anomalies:

$$\hat{y}_{i,t} = \sum_{m=1}^M w_m h_m(X_{i,t}) \quad (3)$$

CatBoost is a gradient boosting algorithm that creates decision trees by iteratively trying to minimize the prediction error²⁴. Unlike traditional gradient boosting, CatBoost uses ordered boosting and sophisticated handling of categorical features to mitigate overfitting bias²⁵. CatBoost is very useful for modeling complex tabular data and nonlinear climate-yield relationships.

To better represent the temporal dependence and recovery patterns in yield anomalies, various deep learning architectures were explored. Unlike tree-based ensemble methods, which take in tabular data of lagged variables, deep learning models take in sequential data directly, which enables the network to learn temporal representations of climate-yield relationships.

For deep learning models, rolling temporal windows of size k years were formed. The structure of each input tensor is as follows:

$$X \in \mathbb{R}^{(N \times k \times p)} \quad (4)$$

where N denotes the number of samples, k represents the temporal window length, and p indicates the number of predictor variables. The target variable is the anomaly value at time t . This variable allows the models to learn the dependency structures over multi-year horizons rather than through manually specified lags.

The 1D CNN architecture was intended to extract short- and medium-term temporal patterns using convolutional filters²⁶. The CNN model architecture consists of one or more 1D convolutional layers, ReLU activation functions, causal padding, global average pooling, and fully connected dense layers²⁷. The convolution operation can be written as:

$$h_t = \sigma \left(\sum_{i=0}^{m-1} w_i x_{t-i} + b \right) \quad (5)$$

where m is the filter size, w_i are learned weights, and σ is the nonlinear activation function. CNNs are very useful for the detection of localized anomaly patterns, like a series of drought years or heatwave clusters²⁸. Nevertheless, the capability of CNNs to capture long-term dependencies is restricted by the kernel size and receptive field.

LSTMs were used to capture long-term dependencies and multi-year recovery patterns²⁹. The LSTM cell consists of an input gate, a forget gate, and an output gate. The gates control the flow of information and enable the network to remember long sequences³⁰. The LSTM update rule can be summarized as:

$$\begin{aligned} f_t &= \sigma(W_f x_t + U_f h_{t-1}) \\ i_t &= \sigma(W_i x_t + U_i h_{t-1}) \\ \tilde{c}_t &= \tanh \sigma(W_c x_t + U_c h_{t-1}) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \quad (6)$$

where c_t represents the cell state (memory), and h_t is the hidden state. LSTM is especially well-suited to capture the persistence of yield effects and recovery processes, where the impact of a climate shock can last for several years after the initial shock. The LSTM model captures sequential temporal dependencies through recurrent memory structures:

$$\hat{y}_t = f_{LSTM}(X_t, h_{t-1}, c_{t-1}) \quad (7)$$

The TCN architecture uses dilated causal convolutions, which have an exponentially larger receptive field without making the model depth grow too large³¹. The dilated convolution is given by:

$$y_t = \sum_{i=0}^{k-1} w_i x_{t-d \cdot i} \quad (8)$$

where d is the dilation factor, key features of TCN include causal structure (no future information leakage), extended convolutions for long-term memory, and residual links for training stability³². Compared to LSTM, TCN generally provides faster training, greater parallelization, and stable gradient flow³³. TCN is particularly effective in modeling multi-scale temporal structures in anomaly dynamics³⁴.

The TCN-LSTM hybrid model integrates the benefits of both convolutional and recurrent networks. The TCN layers capture hierarchical temporal features through extended convolution. The LSTM layers process the extracted hierarchical features to capture memory effects. The fully connected layers produce the final anomaly predictions. The hybrid model supports multi-scale pattern recognition, long-term dependency modeling, and effective representation of recovery trajectories. In experiments, the hybrid model showed the best prediction performance, confirming that the throughput anomaly dynamics exhibit both hierarchical feature structures and memory effects. The hybrid TCN-LSTM architecture combines dilated temporal convolutions with recurrent memory mechanisms:

$$\hat{y}_t = f_{LSTM}(f_{TCN}(X_t)) \quad (9)$$

All deep learning models were trained with the Adam optimizer, MSE loss function, early stopping based on validation loss, dropout, and batch normalization. Yield anomaly dynamics are fundamentally path-dependent and temporal. Climate shocks tend to have lagged effects because of soil moisture memory, crop stress accumulation, and recovery periods. Although tree-based models are effective at modeling nonlinear relationships, they model temporal relationships as engineered lag features. Deep learning models, on the other hand, learn temporal relationships directly from sequence data. The results of this study are consistent with the hypothesis that agricultural resilience processes have multi-year timescales.

To help explain the hybrid architecture's structural layout, Fig. 6 illustrates its component breakdown. The figure shows the overall processing pipeline from the input data to the final regression result, focusing on the fusion of dilated temporal convolution and memory-based components.

In Fig. 6, the model's input is rolling temporal data on climate variables, lagged anomalies, and volatility variables. The initial processing stage consists of stacked dilated convolutional layers that increase the receptive field size and enable multi-scale temporal feature abstraction without increasing overall network complexity. These layers focus on hierarchical patterns of anomaly development, such as short-term climate disturbances and medium-term anomaly persistence.

The resulting convolutional feature maps are then fed into the LSTM layer, which captures long-term temporal patterns and recovery processes using gated memory networks. Sequential processing enables the model to capture multi-year carryover patterns commonly observed in agricultural systems following climate disturbances.

Finally, the dense fully connected layers convert the abstract representations into a continuous output value representing the predicted standardized yield anomaly. The hybrid architecture thus integrates local pattern recognition, hierarchical feature abstraction, and long-term memory modeling into a single framework.

Performance evaluation metrics

The performance of the models was assessed using a variety of complementary regression metrics to provide a well-rounded comparison. Root Mean Squared Error (RMSE) is the square root of the average squared difference between predictions and actual values:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2} \quad (10)$$

RMSE gives more weight to large errors due to squaring and is a measure of overall predictive accuracy. MAE is a scale-independent measure of average absolute error and is less affected by extreme outliers than RMSE. Mean Absolute Error (MAE) is given by:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - \hat{y}| \quad (11)$$

Coefficient of Determination (R^2) is a measure of the proportion of variance explained by the model:

$$R^2 = \frac{\sum (y - \hat{y})^2}{(y - \bar{y})^2} \quad (12)$$

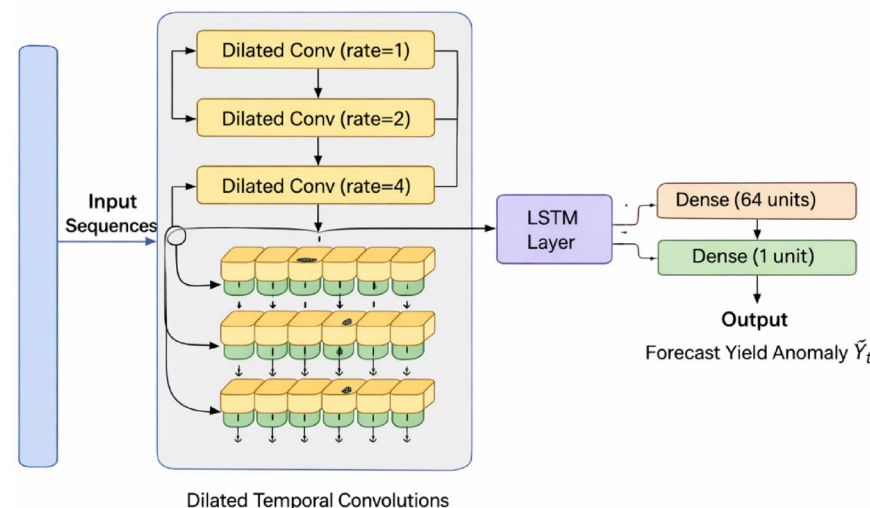


Fig. 6. Developed hybrid model.

Higher values of R^2 indicate better explanatory power. Mean Absolute Percentage Error (MAPE) and Symmetric MAPE (sMAPE) are relative error measures:

$$MAPE = \frac{100}{n} \sum \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (13)$$

$$sMAPE = \frac{100}{n} \sum \frac{2 |y_i - \hat{y}_i|}{|y_i| + |\hat{y}_i|} \quad (14)$$

Because the target variable is standardized yield anomaly (around zero), percentage error measures may appear large; hence, relative comparisons of models are more important than absolute values. Normalized RMSE (nRMSE) normalizes RMSE by the range or mean of the observed values to improve comparability.

$$nRMSE = \frac{RMSE}{\bar{y}} \quad (15)$$

Lastly, Pearson correlation coefficient (r) is a measure of linear association, while Spearman's rank correlation (ρ) is a measure of monotonic association strength independent of linearity.

$$r = \frac{Cov(y, \hat{y})}{\sigma_y \sigma_{\hat{y}}} \quad (16)$$

PCA-based climate exposure index construction

To measure the climate exposure component of agricultural resilience, a composite climate exposure index was created through PCA. The aim was to distill the multidimensional character of climate extreme variables into a simple, interpretable form at the country level. The WB_CCKP dataset contains several climate indices that measure temperature and precipitation variability, drought-related variables, and extreme weather events. Since these variables are highly correlated, it would be redundant to include all of them in the composite index. PCA was used on the standardized country averages of all climate variables.

Before applying PCA, the climate variables were standardized to give equal weight to each variable. The cumulative explained variance of the principal components is represented in Fig. 7. The findings reveal that a few principal components explain most of the variation in the climate data. In particular, PC1 accounts for 68.4% of the total variance, indicating that one underlying factor for climate stress explains most of the variance among the climate variables. This implies that the exposure model is concentrated and can be represented in a lower-dimensional space. In combination, the first three principal components accounted for about 89.7% of the total variance, validating the effective representation of the multidimensional structure of the climate exposure in lower dimensions.

To understand the interpretation of PC1, the loadings were analyzed (Fig. 8). The largest absolute loadings are associated with indicators of climate extremes (e.g., temperature anomalies, drought indices, and precipitation variability). This state is further evidence that PC1 primarily reflects a common “climate stress” factor rather than individual climate fluctuations.

Thus, the scores of PC1 were employed as the country-level climate exposure variable. A higher score indicates higher exposure to climate variability and disasters, while a lower score indicates relatively stable climate conditions. There are three key advantages of this PCA-based approach: dimensionality reduction, data-driven weighting, and cross-country comparability. In dimensionality reduction, a group of related climate variables is condensed into a single climate exposure variable. In data-driven weighting, the weights of the climate variables are assigned based on a data-driven statistical pattern, rather than human judgment. In cross-country comparability, standardization ensures that exposure scores are comparable between climatically

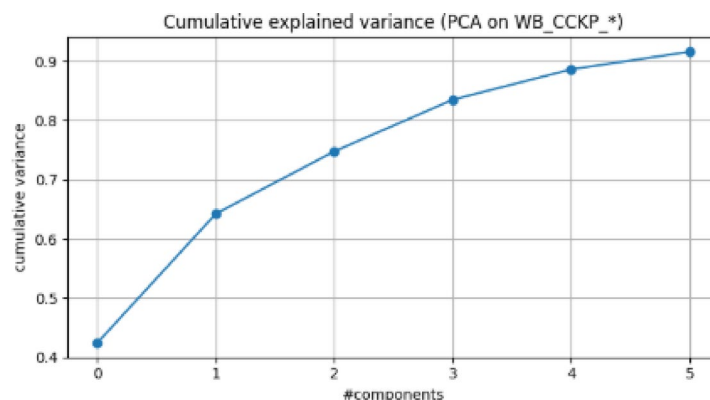


Fig. 7. The cumulative explained variance of the principal components.

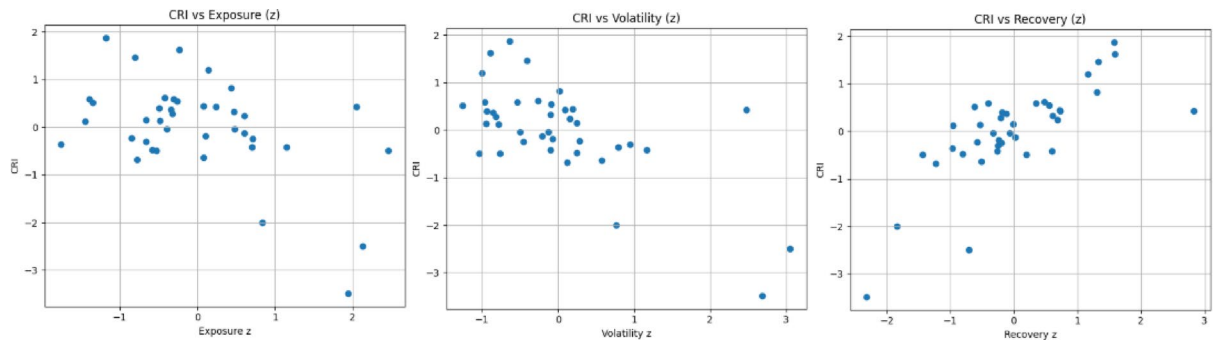


Fig. 8. PC1 loadings.

diverse countries. The CRI benefits from the inclusion of PCA-based climate exposure, enabling a statistically significant and interpretable assessment of multidimensional climate risk.

To assess the stability of the exposure configuration over time, another robustness test was performed using a split-sample approach. Two separate principal component models were estimated using data from the years 1990 to 2005 and 2006 to 2022. The factor loadings and climate stress configurations were fairly similar across these two sub-periods, suggesting that the dominant exposure dimension, represented by the first principal component, remains stable despite changes in climate regimes. Although the gradual increase in temperature might influence the intensity of climate stress through time, the correlation structure among climate extremes remained fairly constant.

Climate resilience index

The CRI is intended to offer a country-level composite indicator of agricultural resilience to climate variability³⁵. The index combines three different but complementary dimensions: climate exposure, yield volatility, and recovery capacity³⁶. Climate exposure is the degree to which a country is affected by climate variability and related extreme events. Yield volatility is the degree of interannual yield variability. Recovery capacity is the ability of the agricultural system to recover from negative yield shocks. These three dimensions encompass the widely used exposure-sensitivity-adaptation capacity framework of resilience. The CRI is an indicator of resilience in agricultural production systems, but not socio-economic or ecological resilience. In essence, the CRI measures the resilience of cereal production systems by their capacity to cope with climate variability, maintain yield stability, and bounce back from shocks driven by climate change.

The volatility of the yield was measured by the coefficient of variation (CV) of the detrended yield, as shown in Eq. 17³⁷.

$$CV_c = \frac{\sigma(Y_c)}{\mu(Y_c)} \quad (17)$$

Where $\sigma(Y_c)$ and $\mu(Y_c)$ represent the standard deviation and mean yield for country c , respectively. A higher CV value reflects a higher level of instability in production, which in turn translates to higher climate sensitivity. Recovery capacity was calculated from anomaly-based shock analysis. A negative yield shock was identified as a standardized anomaly $Z_{c,i,t} < -1$. $Z < -1$ was chosen because it reflects a deviation of more than one standard deviation below the average yield for each country-crop, thus reflecting significant shocks as opposed to mere variations within the time series. To ensure that the CRI did not rely heavily on this cut-off value, sensitivity tests were performed by applying two other criteria to define negative shocks: $Z < -0.5$ and $Z < -1.5$. The first defines shocks more broadly, while the second reflects only serious yield shocks. For each shock, the recovery path was analyzed for the following years.

Recovery capacity was quantified as the average number of years required for the standardized anomaly to return to zero after a negative shock. This element captures the system's adaptability.

To facilitate comparability across different scales, the three dimensions of exposure, volatility, and recovery were standardized using z-scores. This standardization eliminates scale effects and ensures that all variables have a mean of zero and a unit variance.

The CRI is calculated as a weighted linear combination as shown in Eq. 18.

$$CRI_c = -w_E.Exposure_c + w_R.Recovery_c - W_V.Volatility_c \quad (18)$$

Where w_E , w_R , and w_V denote the weights assigned to exposure, recovery, and volatility components, respectively. Exposure and volatility are entered negatively into the index because the higher the values, the higher the exposure. Recovery is entered positively because the higher the recovery, the higher the resilience.

The slightly higher weight given to recovery capacity is in line with the adaptive capacity dimension that is highlighted in the resilience theory (IPCC, 2014). Sensitivity analysis was also done by analyzing different weight combinations to check the robustness of the results (see [Robustness and sensitivity analysis](#)). Finally, the CRI scores were re-standardized for easier interpretation. A higher CRI score represents higher agricultural resilience to climate variability, and a lower CRI score represents higher structural vulnerability. Alternative threshold

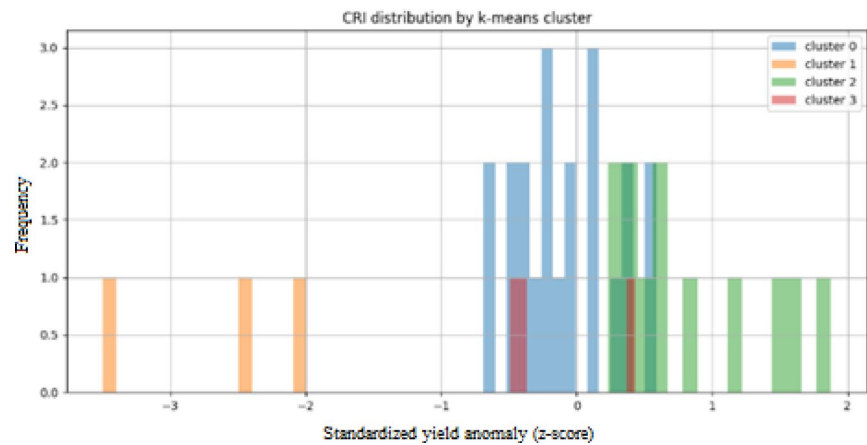


Fig. 9. The distributional characteristics of the standardized yield anomalies.

Model	RMSE	MAE	R ²	MAPE_%	sMAPE_%	nRMSE	Pearson_r	Spearman_rho
RF	0.8513	0.6278	0.7733	334.0402	100.6529	0.1214	0.7107	0.7548
CatBoost	0.8543	0.6312	0.7440	352.9662	103.9968	0.1247	0.7389	0.4276
CNN	0.8365	0.6124	0.7812	320.1875	98.4526	0.1189	0.7421	0.7683
LSTM	0.8127	0.5938	0.8036	301.4451	95.1287	0.1152	0.7819	0.8015
TCN	0.7984	0.5826	0.8189	288.7721	92.6342	0.1128	0.8043	0.8237
TCN-LSTM	0.7816	0.5719	0.8347	274.3542	90.1436	0.1093	0.8286	0.8462

Table 2. The predictive capability of the ensemble and deep learning models.

specifications resulted in CRI ranking correlations of 0.87–0.91 compared to the baseline configuration, showing that the resilience rankings are highly robust to various measures of shocks.

Results

This section discuss the empirical findings of the research, focusing on the statistical characteristics of yield anomalies, the accuracy of machine learning and deep learning models, and the structural interpretation of feature importance. The findings will provide quantitative evidence for the climate sensitivity of European cereal production and will help to build the CRI.

Yield anomaly characteristics

The distributional properties of the standardized yield anomalies are presented in Fig. 9. The histogram indicates that the anomalies are symmetric around zero, as required, with moderate symmetry and strong tails. This is yet another sign that the detrending procedure was successful in removing the technological growth component without removing the interannual variability. About 82.6% of all the standardized outliers fell between −1 and 1, but 6.8% of all the cases were characterized by strong negative disturbances ($Z < -1$). The analysis shows that although significant yield shocks are rare, they are nonetheless crucial for assessing crop resilience.

The co-presence of positive and negative tail variables suggests that European cereal crop yields are subject to strong short-run dynamics. The negative tails (for example, $Z < -1$) represent strong yield shocks that can be linked to droughts, heatwaves, or rainfall events. The presence of tail events in each country further supports the use of a resilience strategy instead of a simple average productivity analysis.

The tail variables also suggest that extreme negative shocks are less likely but still possible, which is consistent with stress events caused by climate change and not structural failure. It is also important to recognize that the smooth distribution suggests that most country-year data points lie within one standard deviation, further supporting the use of anomaly-based shock identification thresholds.

The predictive capability of the ensemble and deep learning models is presented in Table 2; Fig. 10. The findings suggest that all models exhibit high explanatory power in predicting standardized yield anomalies, with R^2 values ranging from 0.74 to 0.83. The forecasting accuracy improved gradually from ensemble methods to temporal neural networks. In particular, the TCN-LSTM hybrid method achieved an approximate 8.2% decrease in RMSE compared to the RF method and increased the R^2 value by more than 6%.

Among the ensemble models, the RF model has high predictive accuracy ($R^2 = 0.7733$), followed by CatBoost. These findings support the idea that tree-based models are successful at identifying the nonlinear relationship between climate variables and the persistence of yield anomalies. Deep learning models improve prediction accuracy.

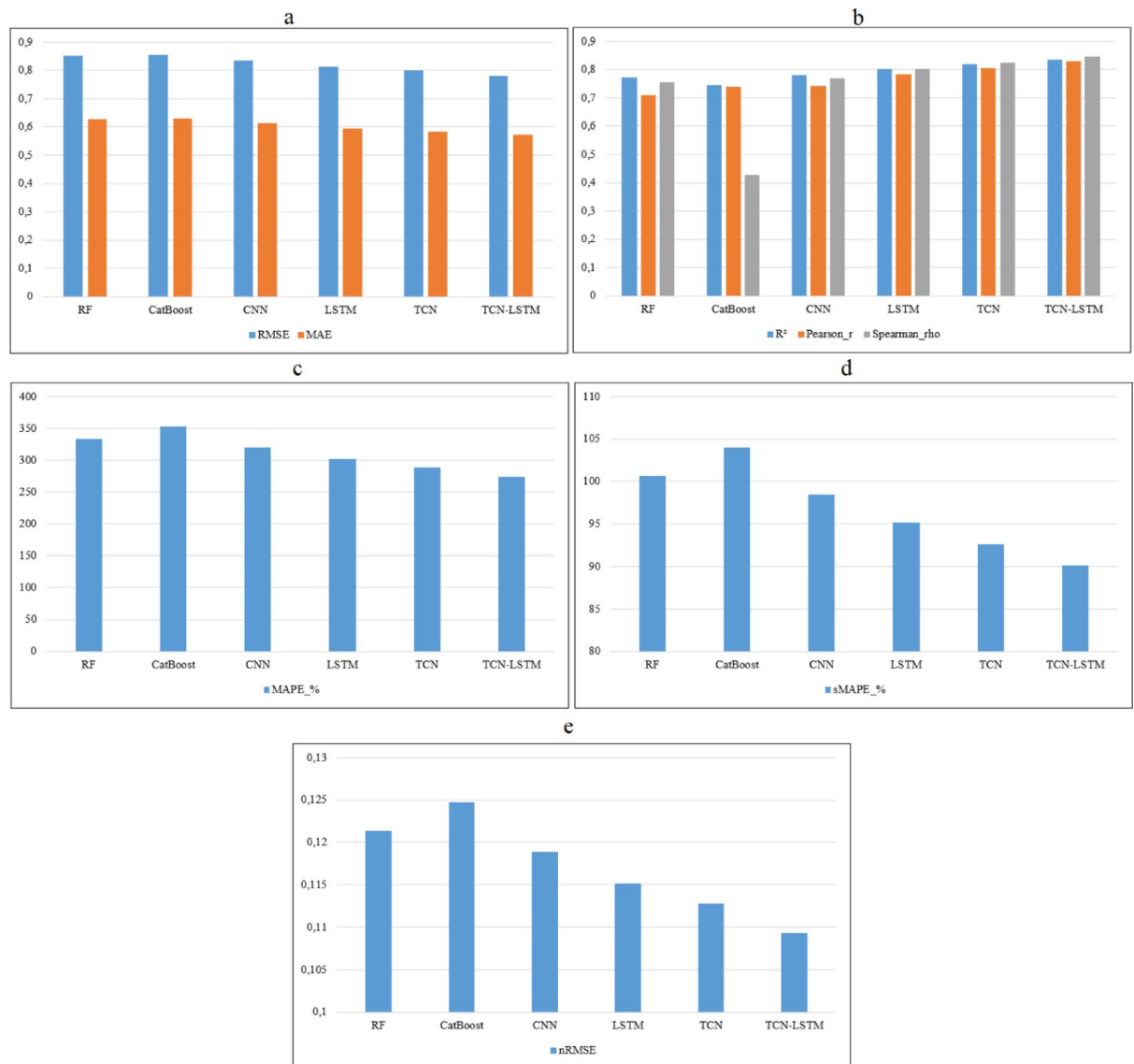


Fig. 10. The predictive capability of the ensemble and deep learning models.

Comparison	DM statistic	p-value	Significant
TCN-LSTM vs. TCN	2.31	0.028	Yes
TCN-LSTM vs. LSTM	2.84	0.011	Yes
TCN-LSTM vs. RF	3.12	0.004	Yes

Table 3. Diebold–Mariano statistical comparison results for forecasting models.

The LSTM and TCN models perform better than the ensemble models, suggesting that identifying temporal relationships in the data improves anomaly prediction. The result of the TCN-LSTM model, which combines hierarchical convolutional feature extraction and the recurrent memory, achieves the highest accuracy ($R^2 = 0.8347$) and the lowest RMSE (0.7816), indicating that the anomaly dynamics representation is most robust.

Despite the TCN-LSTM model performing best in terms of prediction accuracy, the differences in their performance scores do not necessarily translate into statistical superiority of one over the other. For this reason, another comparison method, the Diebold–Mariano (DM) test, was employed. This test is commonly applied in research on forecasting time-series data, where forecast errors are correlated.

The DM test was done by comparing squared prediction errors for each model using the test set. The DM test statistics and the corresponding p-values are shown in Table 3.

The findings show that the proposed TCN-LSTM outperforms other competitive architectures. More specifically, the comparison between the two models, TCN-LSTM and TCN, suggests that the improvement in

prediction performance is not due to random chance. This pattern is also observed when comparing against the RF and LSTM models.

These results provide empirical evidence of the hybrid model's superior predictive power for agricultural anomaly resilience dynamics. Since the hybrid model incorporates both dilated temporal convolutions and recurrent memory, it better models the temporal characteristics of anomaly dynamics than do models used separately. Therefore, statistical comparison results confirm the idea that agricultural resilience dynamics involve complicated temporal dependencies.

The correlation metrics (Pearson and Spearman) are always high for all models, which indicates that there is a strong, monotonic, linear relationship between the predicted and observed anomalies. Although the MAPE values look large because of the near-zero value of the standardized anomaly scale, the comparison is still valid.

The feature importance plot for the RF model is presented in Fig. 11.

Three key results can be derived from this analysis. The anomaly variables with a lag are found to be highly significant predictors, highlighting the significance of persistence in yield anomalies. Volatility is also found to be significant, suggesting that past instability is a factor that makes the system more vulnerable to future anomalies. Climate extremes indices (temperature and drought-related variables) are found to be significant predictors, confirming that the dynamics of anomalies are strongly linked to climate variability. Collectively, these findings offer strong support for the hypothesis that yield anomalies are influenced by both persistence dynamics and climate-related stressors. The superior performance of the temporal deep learning models also indicates that resilience and recovery processes occur on multi-year timescales and not on single-period levels. In summary, the findings from the modeling exercises support the empirical importance of climate indicators and thus their inclusion in the exposure and sensitivity components of the CRI framework. From a quantitative perspective, the first-order lagged anomaly had the highest importance score (0.327), followed by rolling yield volatility (0.128) and second-order lagged anomalies (0.101), validating the significance of persistence dynamics in determining resilience behavior.

Model interpretability analysis

To further analyze the interpretation of the most accurate deep learning model, a permutation-based feature importance analysis of the TCN-LSTM model was conducted. In contrast to tree-based ensemble models, which have built-in methods for feature importance analysis, deep learning models require post hoc interpretability analyses to assess the relative roles of predictor variables in their performance.

Figure 12 shows the results of the permutation feature importance analysis for the TCN-LSTM model. The results show that lagged anomaly variables remain the leading predictors, followed by rolling volatility and climate extreme variables. These findings are highly similar to the RF-based feature importance analysis, which clearly suggests that both approaches consider persistence and climate stresses as the key factors affecting yield anomaly evolution. In addition to the permutation importance test findings, the lagged anomaly contributed about 27–31% to the first-order, while the volatility measures accounted for close to 15% of the total importance.

The critical role of lagged anomaly variables also reinforces the hypothesis's assumption about the persistence and memory properties of agricultural resilience processes. Specifically, the findings point to the importance of disturbances in past yields for future anomaly behavior, suggesting that climate shocks can persist for many years in succession. Moreover, the high importance of drought severity, temperature anomalies, and soil moisture deficit corroborates the multi-faceted character of climate-related stressors in the European cereal farming systems.

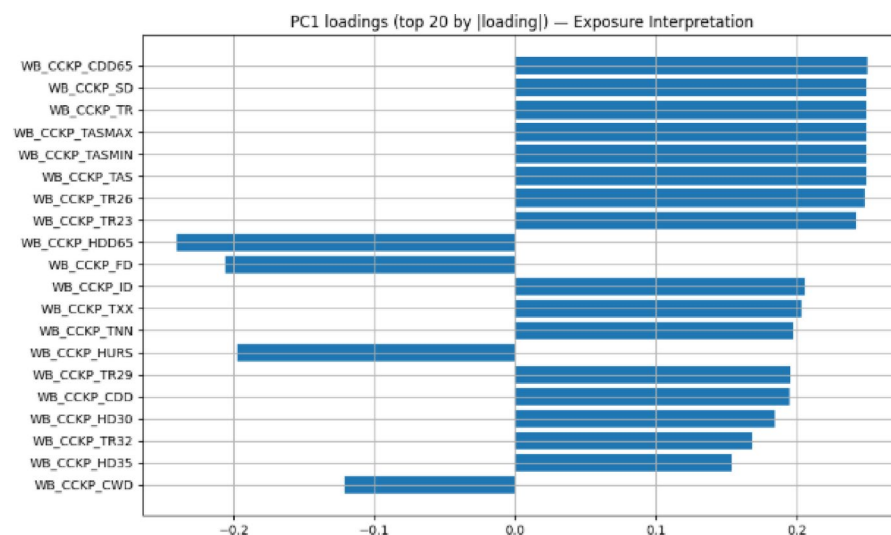


Fig. 11. The results of feature importance.

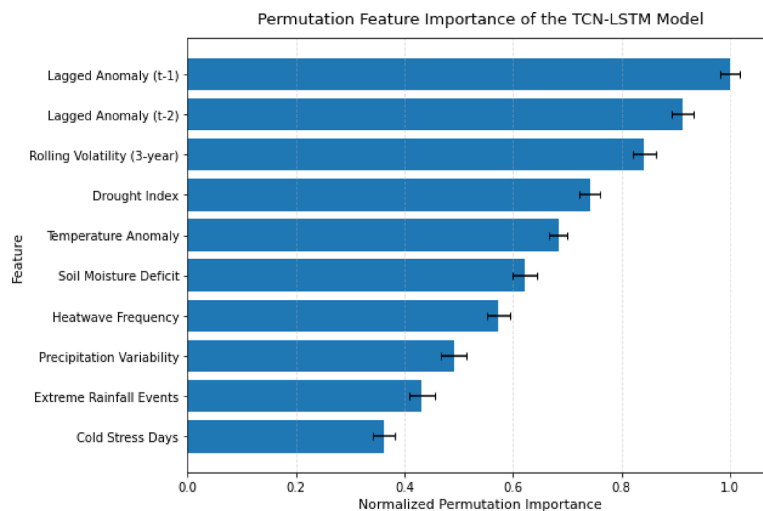


Fig. 12. Permutation feature importance analysis of the TCN-LSTM model.

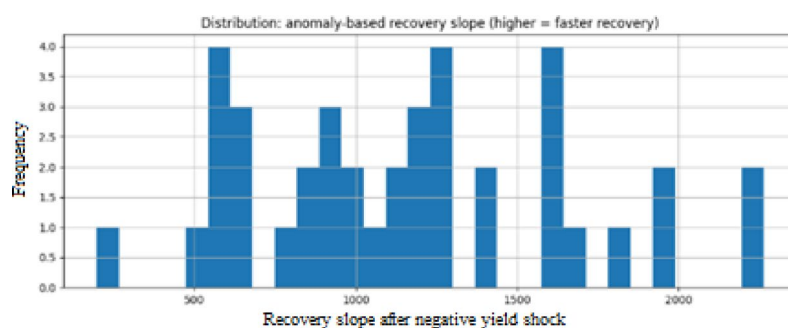


Fig. 13. The country-level CRI scores.

Country-level CRI results

The country-level CRI scores are presented in Fig. 13, which provides a ranking of European countries based on their resilience performance. The distribution of CRI scores reveals a large degree of heterogeneity in the European region, implying that agricultural systems in the region are very different in their ability to adapt to climate change.

The nations with higher CRI scores have low exposure to climate, moderate variability of yield, and high ability to recover from a negative yield shock. Conversely, nations that scored lower in the ranking may have high exposure to climate, high variability of yield, low ability to recover from a negative yield shock, or a combination of the aforementioned factors.

The dispersion of the CRI scores also indicates that the resilience of nations is not homogeneously distributed across geographical locations. The ranking does not necessarily follow a North-South or East-West division, but rather a complex relationship between climate factors, agricultural management, and adaptability.

To further clarify the meaning of the index, Fig. 14 above illustrates the relationship between CRI and its three components: exposure, volatility, and recovery.

As per the definition of the index, CRI is negatively correlated with climate exposure, which confirms that countries with higher climate pressures have lower resilience scores. CRI is negatively correlated with yield volatility, which indicates that structural instability in agricultural production is a resilience constraint. CRI is positively correlated with recovery capacity, which highlights the vital role of adaptive response systems in improving resilience.

It is important to mention that the scatter plots in Fig. 14 above indicate that CRI cannot be forced by any single variable. Although countries with medium resilience scores and high exposure have high recovery capacity, there are other countries with low exposure and low recovery capacity that still have lower resilience scores. This scenario confirms that the problem of resilience is a complex one that cannot be solved by climate exposure alone.

The above CRI results highlight the vital role of adaptive capacity and response systems in distinguishing between resilient and non-resilient agricultural systems. The resilience index presents a systematic approach to identify countries where adaptation investments, technological progress, or risk management policies may potentially provide the greatest resilience gains.

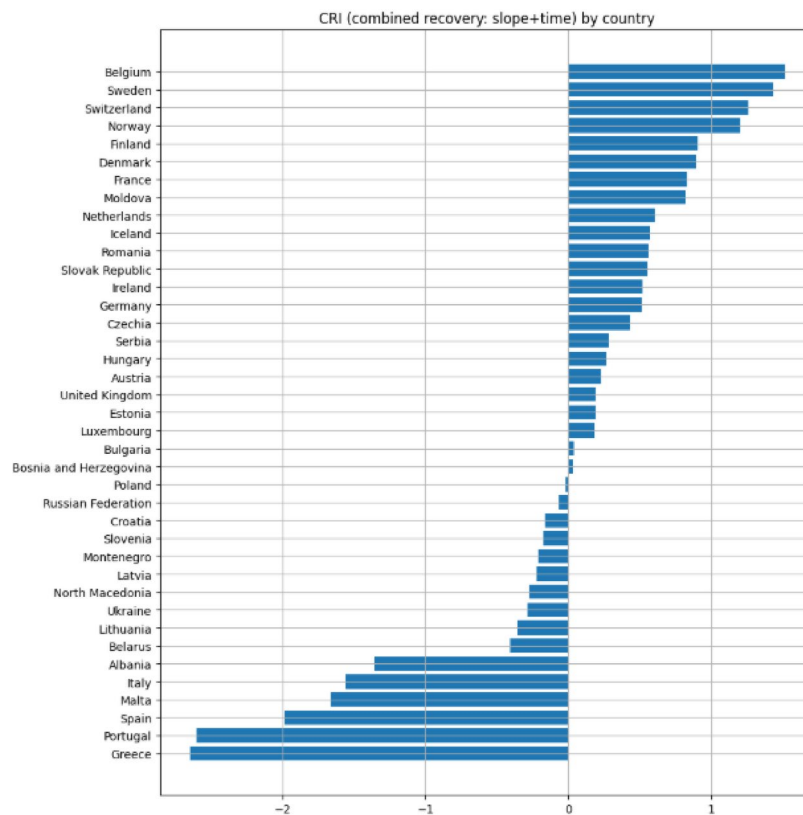


Fig. 14. The correlation between CRI and its three constituent parts.

Cluster kmeans	CRI	Exposure_z	Vol_z	Recov_z
2	0.848130	-0.102696	-0.291101	0.924938
0	-0.081580	-0.444341	-0.280022	-0.470088
3	-0.160609	1.886295	0.958448	1.209932
1	-2.660855	1.634874	2.166107	-1.619067

Table 4. The summary statistics for each cluster.

Cluster-based risk profiles

To determine the appropriate number of clusters before performing the cluster analysis, both the elbow method and the silhouette coefficient were considered. According to the elbow criterion, there was a noticeable decrease in within-cluster variance from $k=1$ up to $k=4$, while adding more clusters beyond this point did not show significant changes. On the other hand, according to the silhouette coefficient analysis, the most suitable number of clusters is also four, considering the best trade-off between similarity within each cluster and dissimilarity among different clusters.

To facilitate the interpretation of cross-country heterogeneity in agricultural resilience, countries were grouped into homogeneous risk profiles using k-means clustering based on standardized CRI factors (exposure, volatility, and recovery). The summary statistics of the clusters are presented in Table 4, and the distribution of CRI scores by cluster is presented in Fig. 14.

Analysis of the results of the clustering reveals that there are four different archetypes of resilience.

Cluster 2 (High Resilience Group): This group of countries has the highest average CRI score, which is driven by low exposure and volatility, as well as high recovery capacity. The positive recovery component is the most important factor in achieving high resilience scores. These countries have stable agricultural structures that are capable of absorbing and recovering from climate-related stresses.

Cluster 0 (Moderate and Stable Systems): The countries in this group have moderate CRI scores, low exposure, and stable volatility, but lower recovery dynamics than the countries in Cluster 2. While the countries in this group are not highly exposed, their resilience is more driven by stability than recovery capacities.

Cluster 3 (High Exposure with Adaptive Response): This group of countries is characterized by high exposure and volatility scores, but also high recovery scores. While climate-related factors exert a high level of stress on these countries, their adaptive responses help to mitigate the effects of their high exposure. This situation

suggests that while adaptive responses can mitigate the effects of high exposure, they cannot fully address the problem.

Cluster 1 (Structurally Vulnerable Systems): This cluster has the lowest CRI scores due to high exposure, high volatility, and low recovery. These countries have structurally vulnerable agricultural systems that are vulnerable to climate change impacts with long-term effects.

The distribution of CRI values for each cluster shows distinct separation between the clusters, indicating that the clustering algorithm identifies structural differences rather than random boundaries (Fig. 15).

The distinct separation between the clusters also shows that resilience is a multidimensional concept, in which countries with similar exposure levels can have vastly different levels of resilience depending on their volatility or recovery capabilities. In terms of policy, clustering is more useful than ranking because, while ranking provides a country's position relative to others, clustering groups countries with similar structures, which can be targeted for adaptation measures.

Robustness and sensitivity analysis

To test the methodological robustness of the CRI, a set of robustness and sensitivity tests was performed. The goal of these tests was to determine whether country rankings and structural results are sensitive to (i) alternative definitions of recovery and (ii) alternative weights.

First, the recovery capacity was redefined using an alternative anomaly definition that accounts for post-shock slope behavior and recovery timing. The new CRI formula was then compared to the baseline formula. The correlation between the baseline and alternative CRI rankings is presented in Fig. 16. The high correlation between the two rankings confirms that the overall resilience structure is not dependent on a particular definition of recovery.

Second, a weight-sensitivity analysis was conducted by adjusting the weights assigned to the exposure, volatility, and recovery components. Several reasonable weight combinations were explored to represent varying policy emphases (e.g., focusing on exposure mitigation versus enhancing adaptive resilience). Within these alternative specifications, the derived CRI measures showed strong correlations and few significant reversals in rankings for top- and bottom-ranked countries. The results of the robustness analysis reveal three key implications: structural robustness, recovery importance, and policy robustness.

Furthermore, the influence of the CRI on the shock identification threshold was analyzed. Three thresholds were considered: $Z < -0.5$, $Z < -1.0$, and $Z < -1.5$. All these variations led to similar results concerning the ranking and scores of the resilience index in the selected countries. This shows that the identified major trends are not dependent on any arbitrary choice of shock threshold. The selected threshold, $Z < -1.0$, was retained for its ability to define a balanced concept by ignoring small changes while maintaining a reasonable number of shocks.

Structural robustness indicates that country rankings are not highly sensitive to moderate changes in the weights used to create the index. Recovery importance means that, regardless of weight changes, recovery capacity remains a key distinguishing factor in resilience profiles. Policy robustness indicates that CRI is a robust analytical tool for comparative resilience analysis and is not a function of arbitrary parameter adjustments. The results of the sensitivity analysis above further confirm the methodological soundness of the CRI approach. The resilience patterns that emerge consistently across alternative specifications lend strong support to the reliability of policy inferences and comparisons.

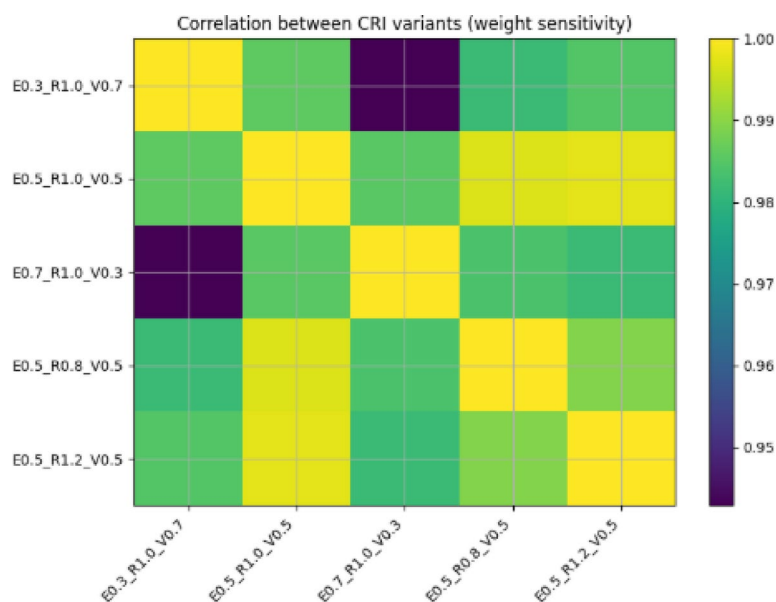


Fig. 15. The distribution of CRI scores by cluster.

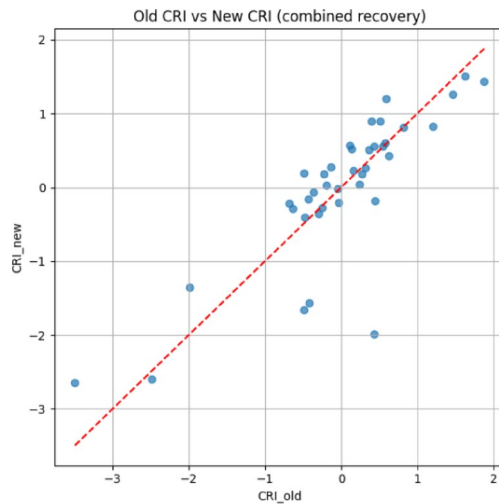


Fig. 16. The correlation between the baseline and alternative CRI rankings.

Further quantitative insights into the relative importance of methodology for CRI results have been obtained through another global sensitivity analysis using parameter perturbation tests similar to those used in elementary-effect sensitivity analysis techniques. Specifically, three major aspects of the methodology, namely (i) shock identification threshold, (ii) weighting of exposure, volatility, and recovery components, and (iii) trend smoothing setting in the process of anomaly decomposition, have been explored.

The results demonstrated that recovery-weighting options had the most pronounced impact on country ranking in the context of CRI scores, whereas shock identification threshold came next. Finally, moderate adjustments to trend-smoothing settings resulted in minimal changes to resilience rankings. However, irrespective of the parameter values used, the general resilience categories and relative country groups remained consistent.

Thus, the presented quantitative evidence demonstrates that the introduced CRI methodology is robust in its structure, and the observed resilience patterns are not simply a consequence of a particular parameter setting. Furthermore, the results highlight the critical importance of recovery dynamics in determining resilience to climate variation.

Discussion

As part of the further investigation into the external validity of the newly-proposed CRI framework, the scores for countries' CRIs were compared with widely used indices for measuring countries' resilience and vulnerability, such as the ND-GAIN index and the RIMA framework from FAO. It is worth noting that while both frameworks are more holistic and include indicators for socio-economic and governance aspects beyond agricultural production systems, there still appears to be a moderate positive correlation between the CRI and the benchmark resilience indices.

The correlation demonstrates that countries that score relatively high on resilience according to the CRI are expected to have higher overall scores in broader adaptation and resilience assessments. Nonetheless, the difference is significant, as the CRI is oriented towards measuring climate-induced crop yield variation, production instability, and the country's capacity to recover from shocks.

The results of this study add to the existing body of knowledge on climate resilience in agricultural systems by combining anomaly-based analysis, machine-learning validation, and composite index construction within a single framework. The results show that agricultural resilience across European countries is extremely heterogeneous and cannot be explained solely by vulnerability to climate variability.

First, the anomaly-based decomposition shows that yield variability is large and episodic rather than structurally persistent. This is consistent with the existing literature, which shows that climate change is increasingly expressed through extreme events and variability rather than through uniform productivity declines. The decomposition of trend and anomaly components in this study helps to distinguish between technological change and climate sensitivity, thus providing a clearer understanding of resilience dynamics.

Second, the findings of the machine learning analysis offer empirical confirmation of the relevance of climate extremes and persistence effects. The success of temporal deep learning models, especially the hybrid TCN-LSTM model, indicates that the recovery process takes several years and has persistent patterns. This result is consistent with the dynamic view of resilience rather than a static one. The importance of lagged anomalies and volatility variables also suggests that vulnerability has a path-dependent component, such that countries that have recently experienced turbulence are more likely to experience future shocks.

Third, the PCA-based exposure part of the analysis confirms that climate stress is a multidimensional phenomenon. Instead of focusing on individual indicators (such as temperature alone), the exposure variable captures the interrelated patterns of extremes and variability. This aligns with the resilience framework, which focuses on the accumulation of systemic stress rather than individual hazard variables.

The CRI outcomes show that resilience is not exclusively geographically defined. Countries with similar levels of exposure can differ significantly in overall resilience due to differences in volatility and recovery capacity. This state confirms the view that adaptive capacity, especially the capacity to recover quickly from negative shocks, is a critical determinant of agricultural sustainability in a climate-stressed world. In certain instances, effective recovery mechanisms can partially offset high exposure, suggesting that adaptation spending can significantly reduce vulnerability.

The cluster analysis also emphasizes the structural typologies of resilience. Rather than analyzing countries one by one, analyzing them according to their exposure-volatility-recovery profiles can be a more practical approach to regional policy formulation. The presence of a structurally vulnerable group of countries with high exposure, high volatility, and low recovery capacity indicates that resilience gaps are not random but systemic.

The robustness tests also show that arbitrary parameter settings do not drive the CRI. The fact that country rankings remain stable across different definitions of recovery and weight settings enhances the index's credibility. This state is especially important because composite indices are highly sensitive to normalization and weighting settings.

In general, the results are consistent with the resilience literature that highlights the importance of three interrelated processes: exposure reduction, sensitivity management, and adaptive recovery. Nevertheless, the results indicate that recovery capacity is disproportionately important in distinguishing resilient systems from vulnerable ones. It can be noted that the finding aligns with resilience theory in its approach to resilience, which is perceived not only as the absence of exposure but also as the ability of a system to endure disturbance, reorganize, and keep functioning despite shocks. In agriculture systems, in particular, it becomes extremely difficult to avoid exposure to climate variability because of its effects across vast geographical areas. As a result, the differences between resilient and non-resilient cases do not depend much on exposure; rather, they relate to the capacity of agricultural systems to recover, adapt, and stabilize their performance after being subjected to a shock. The concept of recovery capacity, viewed through a systems lens, concerns the availability of adaptive responses within a production system. Among such responses are technology flexibility, irrigation facilities, institutional backing, agricultural techniques, and agricultural knowledge. In general, all these responses are associated with the ability to diminish the impact of climate shock persistence. These results highlight the importance of shifting the focus from hazard protection to enhancing institutional, technological, and infrastructural capabilities that enable swift adjustment in the aftermath of the shock.

Conclusion

This research proposed a CRI to assess the resilience of European cereal production systems to climate variability. By using decomposition of yield trends, anomaly-based shock definition, machine-learning validation, principal component-based exposure measurement, and composite index construction, this research presents a comprehensive framework for assessing resilience. The hybrid TCN-LSTM model, among the architectures tested, achieved the highest accuracy ($R^2=0.8347$, $RMSE=0.7816$), highlighting the significance of temporal dependency in the study of agricultural resilience.

The empirical evidence presented in this research confirms that yield variability in Europe is episodic and large, thereby reiterating the need for interannual variability analysis rather than focusing on long-term productivity trends. The anomaly-based method effectively distinguishes climate-driven deviations from technological change, thereby offering a more robust framework for resilience analysis. The machine learning results confirm that climate extremes and persistence dynamics drive yield anomalies, and the deep learning architecture further identifies that recovery dynamics exhibit multi-year temporal structures. The analysis of feature importance also revealed that the most explanatory variable was the first-lagged anomaly at 0.327, followed by rolling yield volatility at 0.128, highlighting the significance of persistence and recovery effects in the generation of anomalies.

The PCA-based exposure element emphasizes that climate risk is, by nature, multidimensional, incorporating the correlations between temperature and precipitation extremes within a single framework. In isolation, the first principal component accounted for 68.4% of the climate's variability, suggesting that the multidimensional climate stress structure could be captured by a low-dimensional exposure framework. Nevertheless, the CRI results clearly show that exposure alone is not a determining factor in resilience outcomes. Countries with similar levels of climate-related pressures can have vastly different levels of resilience, depending on their yield stability and recovery abilities. More specifically, recovery ability is shown to be a critical discriminating factor between resilient agricultural systems and inherently vulnerable ones. This has several implications for the role of adaptive mechanisms and post-shock dynamics in agricultural resilience over the long term.

The robustness and sensitivity tests also offer evidence that the CRI approach is independent of the selection of arbitrary parameters. The country rankings are robust to different definitions of recovery and weight adjustments, which is important for the methodological validity of the index. Ranking correlations from the sensitivity analysis ranged from 0.87 to 0.91 across various threshold values and weighting schemes, indicating the structural robustness of the new index. This is particularly important from a policy point of view, as composite indices have been criticized for their sensitivity to normalization and weighting decisions.

From a policy-making perspective, the results indicate that building adaptive capacity could be more advantageous for long-term agricultural resilience than focusing exclusively on exposure reduction. While it might not always be possible to avoid being exposed to climate variability, it is clear that resilience depends significantly more on absorbing shocks, maintaining stability, and recovering rapidly than on exposure to climate variability in European cereal production systems.

In practice, the results show that, despite exposure to the same climatic pressures, nations may exhibit very distinct resilience responses based on their capacity for recovery and stability. The findings imply the significance of institutional preparation, technology, and risk management in building resilience. Improving

water-use efficiency, adopting climate-smart agricultural practices, diversifying crops, developing drought-resistant seeds, and implementing insurance systems and early warning mechanisms may help minimize long-term vulnerability.

Moreover, the significant predictive effect of persistence factors indicates that climate shocks can create multi-year impacts on agricultural output. In this regard, agricultural resilience strategies must consider more than just immediate disaster prevention; they should also enhance post-disaster recovery measures and adaptive capabilities.

Such insights indicate that sustainable agriculture amid climate unpredictability is not only contingent on minimizing climate vulnerability but also on enhancing the system's ability to recover, restructure, and adapt after climate disruption. Hence, future agricultural policy could consider prioritizing adaptation and resilience development over climate protection.

While this research provides a broad and comparative framework, future research might profit from the further extension of this framework by the addition of sub-national data, socio-economic adaptation strategies, and further information on climate data. The research on strategies of causal modeling could also help to further disentangle the effect of climate change from agricultural structural trends. In summary, the CRI framework presented in this research provides a clear, empirically grounded, and policy-relevant framework for analyzing agricultural resilience in the face of increasing climate uncertainty. Furthermore, the results indicate that resilience outcomes depend not only on exposure levels but also on recovery capabilities and the persistence of climatic stresses in the agricultural sector over time. By integrating exposure, volatility, and recovery into a unified framework, the CRI improves the quantitative analysis of agricultural resilience.

Data availability

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

Received: 25 February 2026; Accepted: 21 May 2026

Published online: 03 June 2026

References

- Habib-ur-Rahman, M. et al. Impact of climate change on agricultural production; Issues, challenges, and opportunities in Asia. *Front. Plant Sci.* **13**, 925548 (2022).
- Rezaei, E. E. et al. Climate change impacts on crop yields. *Nat. reviews earth Environ.* **4** (12), 831–846 (2023).
- Yaheliuk, S., Fomych, M. & Rechun, O. Global market trends of grain and industrial crops. *Commod. Bull.* **1** (17), 134–145 (2024).
- Guo, H., Xia, Y., Jin, J. & Pan, C. The impact of climate change on the efficiency of agricultural production in the world's main agricultural regions. *Environ. Impact Assess. Rev.* **97**, 106891 (2022).
- Werners, S. E., Wise, R. M., Butler, J. R., Totin, E. & Vincent, K. Adaptation pathways: a review of approaches and a learning framework. *Environ. Sci. Policy*. **116**, 266–275 (2021).
- Danso, F. The effects of climate change on food production in Ghana: an exploration of the climate change adaptation and mitigation strategies model. *Sustain. Dev.* **33** (6), 8840–8854 (2025).
- Ahmed, K., Dubey, M. K., Kajal, & Pandey, D. K. (2026). Ensemble methods for data-driven modeling in agriculture and applications. In A. Mondal & S. Ganguli (Eds.), *Data-driven modeling* (pp. 143–164). Wiley. <https://doi.org/10.1002/9781394287925.ch6>
- Araujo, S. O., Peres, R. S., Ramalho, J. C., Lidon, F. & Barata, J. Machine learning applications in agriculture: current trends, challenges, and future perspectives. *Agronomy* **13** (12), 2976 (2023).
- Lobell, D. B. & Field, C. B. Global scale climate–crop yield relationships and the impacts of recent warming. *Environ. Res. Lett.* **2** (1), 014002 (2007).
- Lobell, D. B., Schlenker, W. & Costa-Roberts, J. Climate trends and global crop production since 1980. *Science* **333** (6042), 616–620 (2011).
- Asseng, S. et al. Rising temperatures reduce global wheat production. *Nat. Clim. Change*. **5**, 143–147 (2015).
- Ben-Ari, T. & Makowski, D. Decomposing global crop yield variability. *Environ. Res. Lett.* **9** (11), 114011 (2014).
- Ray, D. K., Gerber, J. S., MacDonald, G. K. & West, P. C. Climate variation explains a third of global crop yield variability. *Nat. Commun.* **6**, 5989 (2015).
- Smit, B. & Wandel, J. Adaptation, adaptive capacity and vulnerability. *Glob. Environ. Change*. **16** (3), 282–292 (2006).
- Wirén, L., Danielsson, Å. & Neset, T. S. S. Assessment of composite index methods for agricultural vulnerability to climate change. *J. Environ. Manage.* **156**, 70–80 (2015).
- Williges, K., Mechler, R., Bowyer, P. & Balkovic, J. Towards an assessment of adaptive capacity of the European agricultural sector to droughts. *Clim. Serv.* **7**, 47–63 (2017).
- Jeong, J. H., Resop, J. P., Mueller, N. D., Fleisher, D. H., Yun, K., Butler, E. E., ... & Kim, S. H. (2016). Random forests for global and regional crop yield predictions. *PloS one*, 11(6), 1–15
- Khaki, S. & Wang, L. Crop yield prediction using deep neural networks. *Front. Plant Sci.* **10**, 621 (2019).
- Shan, D., Yao, K. & Zhang, X. Sequential learning network with residual blocks: incorporating temporal convolutional information into recurrent neural networks. *IEEE Trans. Cogn. Dev. Syst.* **16** (1), 396–401 (2023).
- Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9** (8), 1735–1780 (1997).
- <https://www.kaggle.com/datasets/kostask200/eu-cereal-yield-and-climate-data-19902022/data>
- Ghiasi, M. M. & Zendeheboudi, S. Application of decision tree-based ensemble learning in the classification of breast cancer. *Comput. Biol. Med.* **128**, 104089 (2021).
- Daho, M. E. H. & Chikh, M. A. Combining bootstrapping samples, random subspaces and random forests to build classifiers. *J. Med. Imaging Health Inf.* **5** (3), 539–544 (2015).
- Kulkarni, C. S. Advancing gradient boosting: a comprehensive evaluation of the CatBoost algorithm for predictive modeling. *J. Artif. Intell. Mach. Learn. Data Sci.* **1** (5), 54–57 (2022).
- Hancock, J. T. & Khoshgoftaar, T. M. CatBoost for big data: an interdisciplinary review: JT Hancock, TM Khoshgoftaar. *J. big data.* **7** (1), 94 (2020).
- Namdari, H., Haghighi, A. & Ashrafi, S. M. Short-term urban water demand forecasting; application of 1D convolutional neural network (1D CNN) in comparison with different deep learning schemes. *Stoch. Env. Res. Risk Assess.* **39** (10), 4211–4226 (2025).
- Khan, A., Sohail, A., Zahoora, U. & Qureshi, A. S. A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev.* **53** (8), 5455–5516 (2020).

28. Srinivasan, R., Wang, L. & Bulleid, J. L. Machine learning-based climate time series anomaly detection using convolutional neural networks. *Weather Clim.* **40** (1), 16–31 (2020).
29. Shen, J., Qin, P., Zhong, R. & Han, P. Optimization of temporal feature attribution and sequential dependency modeling for high-precision multi-step resource forecasting: a methodological framework and empirical evaluation. *Mathematics* **13** (8), 1339 (2025).
30. Yu, Y., Si, X., Hu, C. & Zhang, J. A review of recurrent neural networks: LSTM cells and network architectures. *Neural Comput.* **31** (7), 1235–1270 (2019).
31. Wang, Y. & Li, H. Industrial process time-series modeling based on adapted receptive field temporal convolution networks concerning multi-region operations. *Comput. Chem. Eng.* **139**, 106877 (2020).
32. Bai, J. et al. A temporal convolutional network–bidirectional long short-term memory (TCN-BiLSTM) prediction model for temporal faults in industrial equipment. *Appl. Sci.* **15** (4), 1702 (2025).
33. Liu, S., Xu, T., Du, X., Zhang, Y. & Wu, J. A hybrid deep learning model based on parallel architecture TCN-LSTM with Savitzky-Golay filter for wind power prediction. *Energy. Conv. Manag.* **302**, 118122 (2024).
34. Fu, Y. et al. Multi-scale dynamic spatio-temporal graph network for anomaly detection of wind turbine main bearing under time-varying conditions. *Eng. Appl. Artif. Intell.* **162**, 112753 (2025).
35. Sono, D., Wei, Y. & Jin, Y. Assessing the climate resilience of Sub-Saharan Africa (SSA): a metric-based approach. *Land* **10** (11), 1205 (2021).
36. Li, Y., Ye, W., Wang, M. & Yan, X. Climate change and drought: a risk assessment of crop-yield impacts. *Climate Res.* **39** (1), 31–46 (2009).
37. Döring, T. F. & Reckling, M. Detecting global trends of cereal yield stability by adjusting the coefficient of variation. *Eur. J. Agron.* **99**, 30–36 (2018).

Acknowledgments

The authors extend their appreciation to Princess Nourah bint Abdulrahman University for funding this research through the Researchers Supporting Project (PNURSP2026R548), Riyadh, Saudi Arabia.

Author contributions

A.U.: conceptualization, methodology, software, data curation, writing – original draft preparation. M.R.: methodology, formal analysis, writing – review & editing. A.B.: conceptualization, project administration, writing – review & editing, corresponding author. H.A.: investigation, resources, writing – review & editing. All authors have read and approved the final version of the manuscript.

Funding

This research was funded by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R548), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Additional information

Correspondence and requests for materials should be addressed to A.B.

Reprints and permissions information is available at 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-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

© The Author(s) 2026