https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

SMN-AgroCLA: Comparison of Different Normalization Methods for Improving Rice Yield Prediction Accuracy Using Remote Sensing Data in Eastern China from 2008 to 2017

Li Liu[a, b], Kehan Zhang[a, b], Yang Zhang[a, b], Tao Lin[c], Weiwei Sun[d], Hang Sun[a, b], Peilin Song[e], Weiwei Liu[d], Biao Xiong[a, b], Dong Ren[a, b, *], Jingfeng Huang[f, g, h, *]

[a] Hubei Key Laboratory of Intelligent Vision Based Monitoring for Hydroelectric Engineering, China Three Gorges University, Yichang 443002, China

[b] College of Computer and Information Technology, China Three Gorges University, Yichang 443002, China

[c] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China

[d] Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo 315211, China

[e] Key Laboratory of Physical Electronics and Devices, Ministry of Education, Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China

[f] Institute of Applied Remote Sensing and Information Technology, Zhejiang University, Hangzhou 310058, China

[g] Key Laboratory of Agricultural Remote Sensing and Information Systems, Hangzhou 310058, China

[h] Key Laboratory of Environment Remediation and Ecological Health, Ministry of Education, College of Natural Resources and Environmental Science, Zhejiang University, Hangzhou 310058, China

* Corresponding author.

E-mail address: dren@ctgu.edu.cn (D. Ren).

E-mail address: hjf@zju.edu.cn (J. Huang).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

Abstract

Yield prediction is crucial for national food security and the formulation of trade policies. Most deep learning (DL) models rely on normalization methods to process input data, aiming to enhance the stability of model training and accelerate convergence speed. However, the importance of data preprocessing (i.e., input data normalization) in DL-based yield prediction is underemphasized. Furthermore, conventional methods fail to address distortions in feature scaling caused by extreme values, such as abnormally high precipitation, leading to increased prediction errors. In this study, we proposed a Sequential Midrange Normalization (SMN) method and integrated it with the newly designed Agricultural-CNN-LSTM-Attention (AgroCLA) model, collectively termed the SMN-AgroCLA framework, to improve rice yield prediction accuracy under extreme weather conditions. To validate the efficacy of the SMN, we compared it with four other commonly used normalization methods and conducted yield prediction experiments across six different DL models, by using Moderate Resolution Imaging Spectroradiometer, Global Precipitation Measurement and other multi-source remote sensing data of the Eastern China from 2008-2017. The results shown that SMN method consistently outperformed superior yield prediction performance even in years affected by extreme meteorological disasters (e.g., 2015), achieving an R^2 of 0.815, which was 17.3% higher than the next best method, ZSN (Z-Score Normalization). Based on SMN, the accuracy and generalization of all models were optimized, with the AgroCLA achieved the highest accuracy (with R^2=0.841). Additionally, the model's performance peaked around the flowering stage (around mid-August, R^2 =0.859), which is two months ahead of the harvest season. This study demonstrates the critical role of data normalization in deep learning-based yield prediction and offers a practical solution to mitigate the threat of increasing extreme meteorological disasters to food security.

Keywords: Paddy rice yield prediction, Data normalization, Deep learning, Remote sensing, Time series

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

1. Introduction

Using remote sensing (RS) technology to reliably predict crop yields is crucial for assisting policymakers in managing local food supply effectively, regulating timely imports and exports of grains, and achieving food security. Additionally, climate change and extreme weather events poses a significant threat to the reliable prediction of crop yields (Ben-Ari et al., 2018; Rattis et al., 2021). China, one of the most disaster-prone countries in the world, with a wide variety of disasters, extensive geographical distribution, and high frequency of occurrence, has suffered severe crop yields losses (Wakatsuki et al., 2023). Seasonal assessment of crop yields provides more insightful evaluations of their responses to environmental stressors (Guan et al., 2017).

Currently, crop yield prediction methods are primarily based on two types: crop growth models and statistical regression (Wang et al., 2023). Due to the numerous parameters and high accuracy requirements, crop growth models often lack sufficient parameters for regional-scale or global crop prediction. Statistical regression-based methods (linear and nonlinear) do not require a large number of input parameters, making them more suitable for large-scale crop yield prediction. Given the complexity of agricultural ecosystems, researchers often employ nonlinear models to monitor crop growth and predict single yields, such as Support Vector Regression (SVR) and Random Forest (RF) (Li et al., 2009; Fortin et al., 2011). Deep learning (DL) models excel in extracting and integrating features across multiple scales and levels of abstraction, and effectively distill complex features into higher-level representations, enhancing their ability to capture intricate patterns within the data and showing significant potential in using RS data for crop yield prediction (Zhang et al., 2016a; Reichstein et al., 2019). CNN models are well-suited for handling spatial autocorrelation in RS images, while RNN models have an advantage in analyzing temporal changes in RS data (You et al., 2017; Cai et al., 2019). The Long Short Term Memory (LSTM) neural networks, which use memory units with gating capabilities

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

compared to ordinary RNNs, enhance the network's long-term memory capability and are therefore widely used in handling time series data (Jiang et al., 2020; Schwalbert et al., 2020).

In recent years, studies have progressively shifted from singular neural network architectures to more intricate models derived from improved versions of earlier architectures (Xiong et al. 2024; Guo et al., 2024). Tian et al. (2021) proposed an attention mechanism-based LSTM (LSTM neural network with an attention mechanism, ALSTM) model, which had six layers, namely one input layer, one LSTM layer, one attention mechanism layer, two dropout layers, and one output layer, for winter wheat yield prediction in the Guanzhong Plain of Shaanxi. Jeong et al. (2022) predicted paddy rice (*Oryza sativa* L.) yield at the county and pixel levels in the Korean Peninsula using RS techniques. They investigated the influence of different model structural combinations (1D-CNN and LSTM) on yield prediction.

The aforementioned study enhanced yield prediction accuracy by comparing DL models with traditional machine learning models or different DL architectures. However, the accuracy does not solely depend on the model structure; it is also typically influenced by input data (such as source, quality, scale, selection of key variables) (Zhang et al., 2016a). Data normalization is a crucial step in DL tasks as it contributes to enhancing the efficiency of model training and testing and assists in mitigating the risk of model overfitting (Ioffe et al., 2015; Ba et al., 2016). Chu et al. (2020) developed an end-to-end model for predicting rice yield, where data normalization techniques played a crucial role during the model's preprocessing stage. They discussed various normalization methods like mean normalization and Z-score normalization, and elaborately explained why the Min-Max normalization was ultimately chosen, primarily based on its simplicity and efficiency in ensuring all numerical variables are on the same scale.

However, the traditional fixed normalization method might not be suitable for all scenarios, especially when

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

the data distribution might change (Passalis et al., 2019). In general, the accuracy improvement brought by the input data normalization method in DL yield prediction tasks is seldom discussed in depth, and its errors mainly originate from the following three aspects: (1) Information loss error. Normalization methods typically map the original data to a specific numerical range, which may lead to the loss of some information; (2) Outlier handling. Some normalization methods are sensitive to outliers in the data used to calculate the data's mean and variance, affecting the performance of the model; (3) Hyperparameter selection. Some methods, such as Min-Max normalization, require manually selected hyperparameters (e.g., minimum and maximum values), and improper selection can lead to inaccurate scaling and reduced prediction accuracy (Yang et al., 2020; Bischl et al., 2023).

To address the feature scaling distortion caused by extreme values (such as abnormally high precipitation) in common normalization methods and to improve the accuracy and robustness of rice yield prediction. This study designed the SMN method during the data preprocessing phase to monitor the relative change characteristics of actual values of various factors with high quality, combined with newly designed Agricultural-CNN-LSTM-Attention (AgroCLA) model, collectively termed SMN-AgroCLA framework for county-level rice yield prediction in eastern China. Our specific objectives are:

(1) To focus on the importance of data pre-processing in the DL applications for yield prediction and to meticulously design comparative experiment of various data normalization and input methods;

(2) Developed a novel rice prediction model (AgroCLA) and compared its performance with five widely used models in the field of rice yield prediction.

(3) Exploring the interpretability of the SMN and AgroCLA model from typical actual disaster events and feature importance analysis, respectively.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

2.Materials

2.1 Study area and yield records

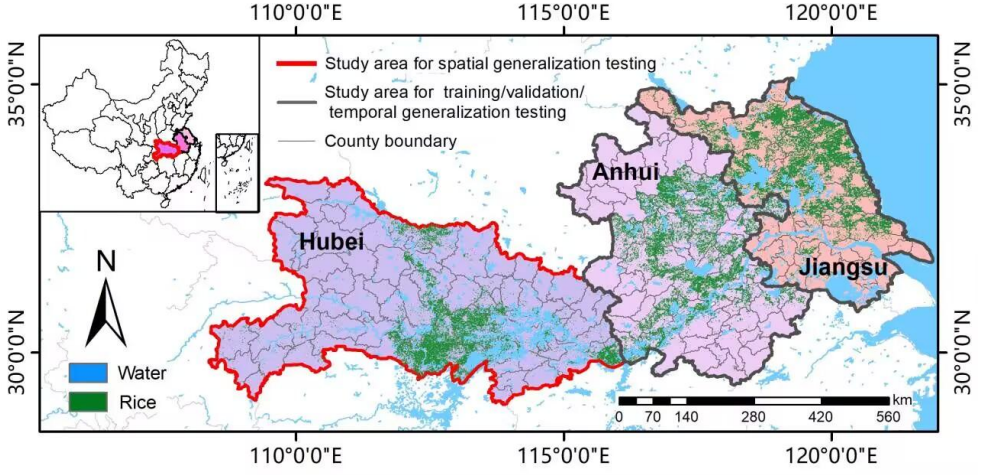


Fig. 1. Paddy rice distribution map in study area.

Anhui, Jiangsu, and Hubei Provinces are located in eastern China (29° - 36° N, 108° - 122° E), with a combined area of approximately 405,000 square kilometers, spanning the Yangtze River Delta and the Huaihe River Basin (Xiao et al., 2002; Fig. 1). The topography is predominantly characterized by plains, with the majority of the area having an elevation of 50 meters or lower. This region falls within the subtropical monsoon climate zone, receiving abundant solar radiation and ample rainfall, which makes it highly suitable for crop cultivation. Major crops grown in three provinces include paddy rice, winter wheat (*Triticum aestivum* L.), and winter rapeseed (*Brassica napus* L.). In the rice cultivation process, rice seeds are sown in dry soil and allowed to grow until they develop 2-4 leaves. Subsequently, rice seedlings are transplanted into irrigated fields and receive continuous watering. Typically, rice transplanting occurs in June, heading takes place from mid-August to mid-September, and grain harvesting occurs from early to mid-October to early November (Liu et al., 2019).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

The census yields (unit: kg/ha) at the county level from 2008 to 2017 were obtained from the Agricultural Yearbook of the provinces and county level statistics bureaus (http://www.stats.gov.cn). Yield records were available for 73 districts in Jiangsu Province, 78 in Anhui Province, and 79 in Hubei Province. The growth stage data of field observations was downloaded from the China Meteorological Data Service Centre (CMDSC, http://data.cma.cn/). To enhance the training efficiency of the model, we utilized the Min-Max method to process the yield data, DEM, Lat and Lon. It is worth noting that from August to October, when crops mature, the East Asian Summer Monsoon retreats southward. Droughts and floods happen easily in this season and have caused serious economic losses and environmental damage (Wang et al., 2019). The alternating occurrence of extreme climatic events such as droughts and waterlogging in the study area is highly representative (Guo et al., 2016). Moreover, this study covers a long research period, providing diversified remote sensing input data, which effectively tests the model's applicability and robustness under various disaster scenarios.

2.2 Data description and preprocessing

For the model development and validation, ten variables were extracted from seven data sources. A summary of the input features is shown in Table 1, and the details of each extracted variable are given below. Three products of the Moderate Resolution Imaging Spectroradiometer (MODIS, Collection 6) were used in this study, including the daily Land Surface Temperature (LST) products (MO[Y]D11A1), the 8-day composite surface reflectance products (MO[Y]D09A1), and 8-day composite gross primary production products (MOD17A2HGF).

All environmental and remote sensing variables used in this study were collected from 2003 to 2019, except for solar radiation, which was available from 2003 to 2018. This extended time span ensures a comprehensive capture of interannual climatic and vegetation variability. However, the selection of specific study years was

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

determined based on the availability and completeness of county-level rice yield data. Therefore, while the predictor variables span a longer period, the actual modeling and evaluation were conducted only for the years with reliable and continuous yield records.

Table 1

Primary data sources and relevant information.

Variable	Abbreviation	Data name	Temporal resolution	Spatial resolution	Source
Precipitation	P	GPM 3IMERGDF	1-day	0.1°	NASA-GES DISC DAAC [1]
Soil Moisture	SM	Daily all-weather surface soil moisture data set	1-day	1000 m	TPDC [2]
Air Temperature	AT	MO[Y]D11A1 LST	1-day	1000 m	NASA-LAADS DAAC [3]
Solar Radiation	SR	China Meteorological Forcing Dataset	1-day	0.1°	TPDC
Vegetation Index	EVI2	MO[Y]D09A1 reflectance	8-day	500 m	NASA-LAADS DAAC
Gross Primary Productivity	GPP	MOD17A2HGF gross primary productivity	8-day	500 m	NASA-LAADS DAAC
DEM/ Longitude/ Latitude	DEM/Lon/ Lat	Digital Elevation Model	--	90 m	SRTM [4]
Yield	Y	Statistical data	year	County level	Agricultural Yearbook
Rice Mask	--	Rice distribution map	year	500 m	Liu et al. 2020

NASA-GES DISC DAAC [1]: NASA Goddard Earth Sciences Data and Information Services Center Distributed Active Archive Center

TPDC [2]: National Tibetan Plateau Third Pole Environment Data Center

NASA-LAADS DAAC [3]: NASA Land Atmosphere Near real-time Capability for EOS Data Active Archive Center

SRTM [4]: Shuttle Radar Topography Mission

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

2.2.1. Precipitation

The Global Precipitation Measurement (GPM) program is a collaborative effort initiated by the National Aeronautics and Space Administration (NASA) of the United States and the Japan Aerospace Exploration Agency (JAXA). Building upon the successful experience and achievements of the Tropical Rainfall Measuring Mission (TRMM), GPM not only improves traditional precipitation retrieval algorithms but also provides more accurate calibration references for multi-satellite joint precipitation retrieval using the GPM Microwave Imager (GMI) and the Dual-frequency Precipitation Radar (DPR) sensors. The Version 06 IMERG Final Run daily datasets with 0.1° spatial resolution from 2003 to 2017 were used in this study, which were obtained from the NASA archive (https://gpm.nasa.gov/data-access/downloads/gpm).

2.2.2. Soil Moisture

The "Daily all-weather surface soil moisture dataset with 1 km resolution in China (2003-2019)" was utilized to obtain soil moisture information (Song et al., 2022; Song et al., 2021). This dataset, originating from previous research study conducted by the authors' laboratory, can be freely accessed and downloaded from the National Tibetan Plateau Third Pole Environment Data Center (TPDC). The dataset incorporates data from the AMSR-E and AMSR-2 passive microwave radiometers, covering all on-orbit times between 2003 and 2019. It also combines MODIS optical reflectance data (MCD43A4) and daily thermal infrared land surface temperature data (MYD21A1 LST). Surface Soil Moisture (SSM) at a 36 km resolution is retrieved and downscaled based on the 36 km brightness temperature data from the two radiometers.

2.2.3. Solar Radiation

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

The research utilized solar radiation data derived from the Downward Solar or Shortwave Radiation (SR) data product, which originates from the "China Meteorological Forcing Dataset (CMFD)" (He et al., 2020; Yang et al., 2010) and were obtained from the National Tibetan Plateau Data Center (Download link: http://data.tpdc.ac.cn/zh-hans/data/8028b944-daaa-4511-8769-965612652c49/). This dataset is created by blending conventional meteorological observations from the China Meteorological Administration with background fields from existing international sources, including the Princeton University's Global Land Surface Model Data, GLDAS data, GEWEX-SRB (Global Energy and Water Exchanges-Surface Radiation Budget) radiation data and TRMM precipitation data. The dataset accuracy falls between the meteorological bureau observational data and satellite RS data, and it has been demonstrated to be superior to the accuracy of pre-existing international reanalysis data. The formula for calculating solar radiation (SR) is as follows:

$$SR = \frac{sard \times 24 \times 3600}{1000} \quad (1)$$

where SR represents the daily solar radiation in units of MJ/(m2•day), and sard denotes the obtained downward shortwave radiation value from the product.

2.2.4. Air Temperature

The spatiotemporally comprehensive air temperature dataset for the study area was derived from the author's previous research (Dou et al., 2020) and calculated using MODIS Land Surface Temperature (LST) products. After mosaicking and clip, we extracted 4 sub-datasets (LST_Day_1km, QC_Day, LST_Night_1km, QC_Night) and selected good quality pixels whose QC value equals zero. Then we used LST data, digital elevation model (DEM), the two-band Enhanced Vegetation Index (EVI2), latitude and longitude data to predict daily mean air temperatures based on RF model. For reconstructing predicted air temperature, we adopted the

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

cloud gap-filling method referred to as a typical spatiotemporal data fusion (STDF) (Song et al., 2019). It was built using clear-sky predicted air temperature of spatially neighboring pixels observed at proximal dates, with concurrent EVI2 and DEM also employed as additional data inputs.

2.2.5. The two-band Enhanced Vegetation Index

The spatiotemporally comprehensive vegetation index dataset for the study area originated from the author's previous research (Liu et al., 2020). The 500m 8-day composite surface reflectance products (MOD09A1 and MYD09A1) of Terra and Aqua satellite from 2003 to 2019 were downloaded from NASA's Level 1 and Atmosphere Archive and Distribution System (LAADS) (https://ladsweb.nascom.nasa.gov/index.html). With the data processing method combination of EVI2_BLUE_MYO, the procedures mainly included image mosaicking, sub-setting, two-band Enhanced Vegetation Index (EVI2) calculation, data quality labelling, cloudy pixel removal, interpolation of vegetation index images, image stacking, and Savitzky-Golay smoothing. The two-band Enhanced Vegetation Index (EVI2) (Jiang et al., 2008) was computed as follows (Eq. (2)):

$$EVI2 = 2.5 \times \frac{(\rho_{858} - \rho_{654})}{(\rho_{858} + 2.4 \times \rho_{654} + 1)} \quad (2)$$

where ρ_{654} and ρ_{858} represent the reflectance of b1 and b2 respectively, of the MOD09A1 and MYD09A1 products.

2.2.6. Gross Primary Productivity

The MOD17A2HGF images from 2001–2017 were downloaded from the NASA Land Processes Distributed Active Archive Center (https://lpdaac.usgs.gov/dataset_discovery/modis/). It is a standard satellite product with a spatial resolution of 500 m and a temporal resolution of 8 days and calculated based on the biome-biogeochemical cycles (BIOME-BGC) model which uses light use efficiency (LUE) (Coops et al.,

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

2009). The LUE-based GPP model uses the relationships among solar radiation, vegetation types, and LUE. And the year-end gap-filling method is used to solve the problem of missing data caused by cloud and rain pollution (Zhao et al., 2005).

2.2.7. DEM, Longitude and Latitude

A Digital Elevation Model (DEM) is a digital representation of the morphology of the Earth's surface. The DEM data utilized in this study were sourced from the Shuttle Radar Topography Mission (SRTM), a collaborative effort between NASA and the National Imagery and Mapping Agency (NIMA) of the U.S. Department of Defense. Specifically, version 4.1 of the SRTM DEM data with a spatial resolution of 90 meters was employed (Reuter et al., 2007). This version has been improved by the International Center for Tropical Agriculture (Centre International Pour Agriculture Tropical, CIAT) through the application of new interpolation algorithms to enhance the previously available data. The data can be downloaded from the SRTM 90m DEM Digital Elevation Database website (http://srtm.csi.cgiar.org/). Additionally, latitude (Lat) and longitude (Lon) information for the study area was extracted from this dataset.

2.2.8. Rice map

The distribution of rice from 2003 to 2017 was obtained by the author's previous research results with a resolution of 500m (Liu et al., 2020). Both maps are based on decision tree classification named PhenoRice algorithm which combined with the phenology information of crops and with accuracies greater than 90%. The algorithm, considering the cropping systems of the study area, is capable of extracting extensive rice area distribution, sowing dates, and flowering dates across various environmental conditions by analyzing the

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

temporal profile of individual pixels. Further details regarding the PhenoRice algorithm can be found in previous studies (Boschetti et al., 2017; Busetto et al., 2019).

We resampled the different spatial resolutions of the input variables (except Yield) at a resolution of 1km for our reference using the nearest-neighbor interpolation method. All the sequential variables were collected or aggregated to an 8-day interval with a total of 46 periods for every year. And the rice map of each year was used to delete the non-rice pixels of all input variable images. All of data preprocessing were implemented using the Python v.3.7 programming language.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

3.Experimental setups and methodology

After data preprocessing was completed, the 8-day 1-km datasets of precipitation (P), soil moisture (SM), solar radiation (SR), air temperature (AT), the two-band Enhanced Vegetation Index (EVI2), and gross primary productivity (GPP) covering the entire study area from 2008 to 2017 were obtained.

3.1 Experimental setup

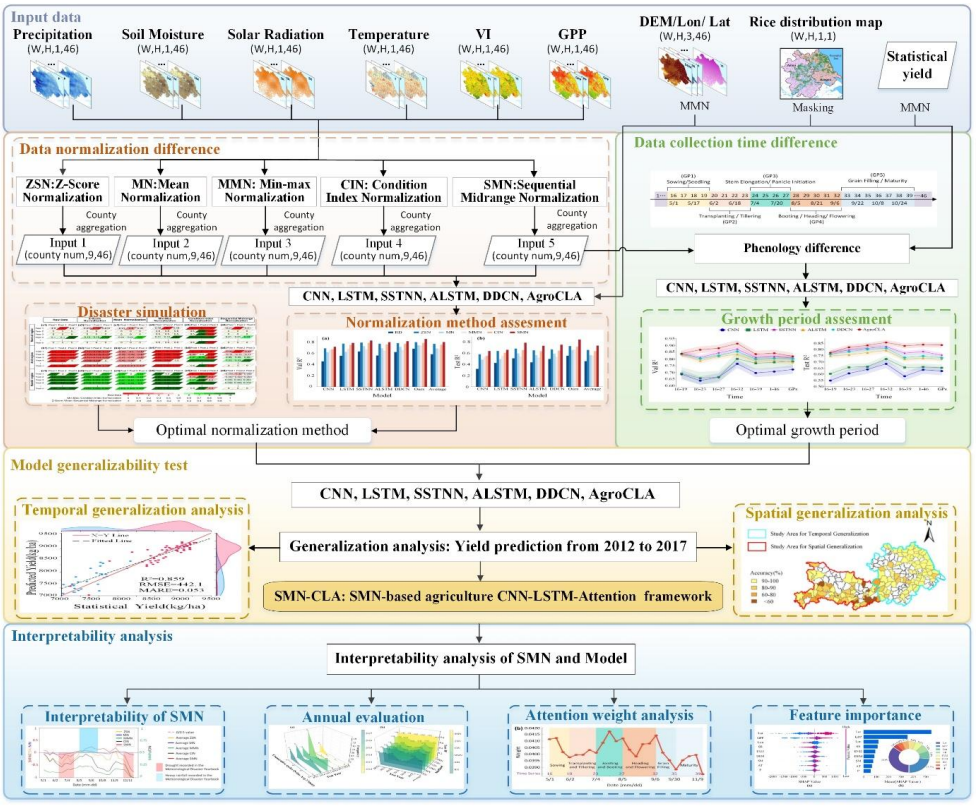


Fig. 2. Flowchart of SMN based AgroCLA rice yield prediction framework (SMN-AgroCLA).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

Differences in normalization methods lead to variations in data quality, while different data collection times leads to variations in data scale. To evaluate the influence of normalization methods on rice yield prediction accuracy, our specific experimental designs are shown in Fig. 2: (1) Five different data normalization methods: Z-Score Normalization (ZSN), Mean Normalization (MN), Min-Max Normalization (MMN), Condition Indices Normalization (CIN), and the newly proposed Sequential Midrange Normalization (SMN), were used to process input data from Jiangsu and Anhui Provinces. Five types of data were individually fed into six different DL models including CNN (Nevavuori et al., 2019), LSTM (Jiang et al., 2020), SSTNN (Qiao et al., 2021), ALSTM (Tian et al., 2021), DDCN (Xiong et al., 2024) and the proposed AgroCLA, to evaluate the differences in accuracy arising from different normalization methods in yield prediction study in Section 4.1. (2) Seven distinct phenological stage data combinations were designed (Fig. 3), including 16-19 (sowing/seedling), 16-23 (from sowing to the end of tillering), 16-27 (from sowing to the end of booting), 16-32 (from sowing to the end of flowering), 16-39 (from sowing to the end of maturity), 1-46 (full-year data), and GPs (average of five growth periods), and separately input into six different DL models to evaluate the accuracy difference due to different data collection times in Section 4.2. (3) To validate the robustness of SMN-AgroCLA which is integrated by selected SMN method and AgroCLA model, we adhered to the "training-validation-testing" strategy and conducted generalization test in Section 4.3 and feature importance analysis in Section 4.4.

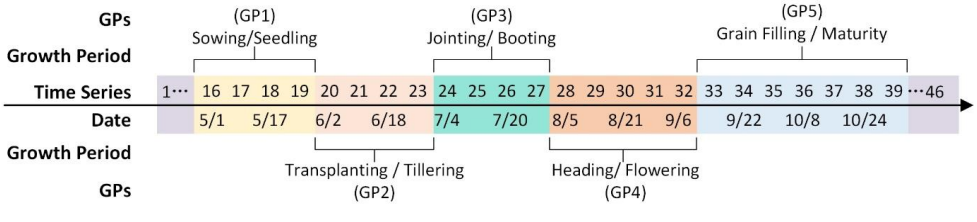


Fig. 3. Introduction to the rice growth period in the study area.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

To evaluate the performances of different models, four metrics including coefficient of determination (R^2), absolute relative error (ARE), root mean squared error (RMSE), and mean absolute relative error (MARE) were selected and were calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3)$$

$$ARE = \left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \quad (5)$$

$$MARE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \quad (6)$$

where y_i and $\hat{y}_i$ respectively represent the actual reference value and predicted value of pixel i, $\bar{y}$ denotes the average value of the measured samples, and n indicates the sample scale.

3.2 Data normalization method

To reduce errors caused by different units and improve the training efficiency of the model (Ma et al., 2021; Patro and Sahu, 2015), each input value was normalized. We chose ZSN, MN, MMN, CIN, and SMN to compare differences in accuracy caused by variations in data quality in the yield prediction study.

3.2.1. Z-Score Normalization

Z-Score Normalization (Curtis et al., 2012), commonly employed in DL models for feature scaling, standardizes the input features to have a mean of zero and a standard deviation of one. This method formulated as follows (Eq. (7)):

$$ZSN_{i,t,s} = \frac{F_{i,t,s} - mean(F_i)}{SD}, i \in (P, SM, SR, AT, EVI2, GPP) \quad (7)$$

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

where $F_{i,t,s}$ is the pixel current value of the feature i, $mean(F_i)$ and SD represent the mean and standard deviation of all pixels over time and space of the feature i. $ZSN_{i,t,s}$ change from -1 to 1, and the value of 0 is set as the threshold to monitor the anomalous change. When $ZSN_{i,t,s}$ equals 0, $F_{i,t,s}$ equals $mean(F_i)$, which is defined as $Standard_{ZSN}$. Take precipitation as an example, during a drought event, the $ZSN_{i,t,s}$ is less than 0, it means that the current state of the pixel ($F_{i,t,s}$) is less than $Standard_{ZSN}$ ($mean(F_i)$), and vice versa.

3.2.2. Mean Normalization

Mean Normalization (D'haene et al., 2012), commonly used in DL tasks to improve model training, transforms input features by subtracting the mean and dividing by the data range, thereby centering the values around zero. This normalization typically maps the data to the range [−1, 1], facilitating faster and more stable convergence in neural networks. The mathematical formulation is given in Eq. (8):

$$MN_{i,t,s} = \frac{F_{i,t,s} - mean(F_i)}{max(F_i) - min(F_i)}, i \in (P, SM, SR, AT, EVI2, GPP) \quad (8)$$

where $F_{i,t,s}$ is the pixel current value of the feature i, $mean(F_i)$, $min(F_i)$, and $max(F_i)$ are the mean, minimum, and maximum value of all pixels over time and space of the feature i. When $MN_{i,t,s}$ equals 0, $F_{i,t,s}$ equals $mean(F_i)$, which is defined as $Standard_{MN}$.

3.2.3. Min-Max Normalization

Min-Max Normalization (Chu and Yu, 2020), widely used in DL-based yield prediction, linearly scales input features to the range [0, 1]. The transformation is defined as follows (Eq. (9)):

$$MMN_{i,t,s} = \frac{F_{i,t,s} - min(F_i)}{max(F_i) - min(F_i)}, i \in (P, SM, SR, AT, EVI2, GPP, DEM, Lon, Lat) \quad (9)$$

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

where $F_{i,t,s}$ is the pixel current value of the feature i, $min(F_i)$ and $max(F_i)$ are the minimum and maximum value of all pixels over time and space of the feature i. The value of 0.5 is set as the threshold to monitor the anomalous change. When $MMN_{i,t,s}$ equals 0.5, $Standard_{MMN}$ equals $[max(F_i) + min(F_i)]/2$.

3.2.4. Condition Indices Normalization

The Temperature Condition Index (TCI), Soil Moisture Condition Index (SMCI), and Vegetation Condition Index (VCI) are widely used in meteorology and remote sensing applications (Kogan, 1995a; Kogan, 1995b; Zhang and Jia, 2013; Wei et al., 2021). These indices are commonly derived using the Condition Indices Normalization (CIN) method, as formulated in Eq. (10):

$$CIN_{i,t,s} = \frac{F_{i,t,s} - min(F_{i,s})}{max(F_{i,s}) - min(F_{i,s})}, i \in (P, SM, SR, AT, EVI2, GPP) \quad (10)$$

where $F_{i,t,s}$ is the pixel current value, $max(F_{i,s})$ and $min(F_{i,s})$ represent the maximum and minimum value recorded for this pixel over time of the feature i. $CIN_{i,t,s}$ change from 0 to 1, and the value of 0.5 is set as the threshold to monitor the anomalous change. When $CIN_{i,t,s}$ equals 0.5, the $Standard_{CIN}$ equals $[max(F_{i,s}) + min(F_{i,s})]/2$ (Ghaleb et al., 2015; Zhang et al., 2019; Wei et al., 2020).

3.2.5. Sequential Midrange Normalization

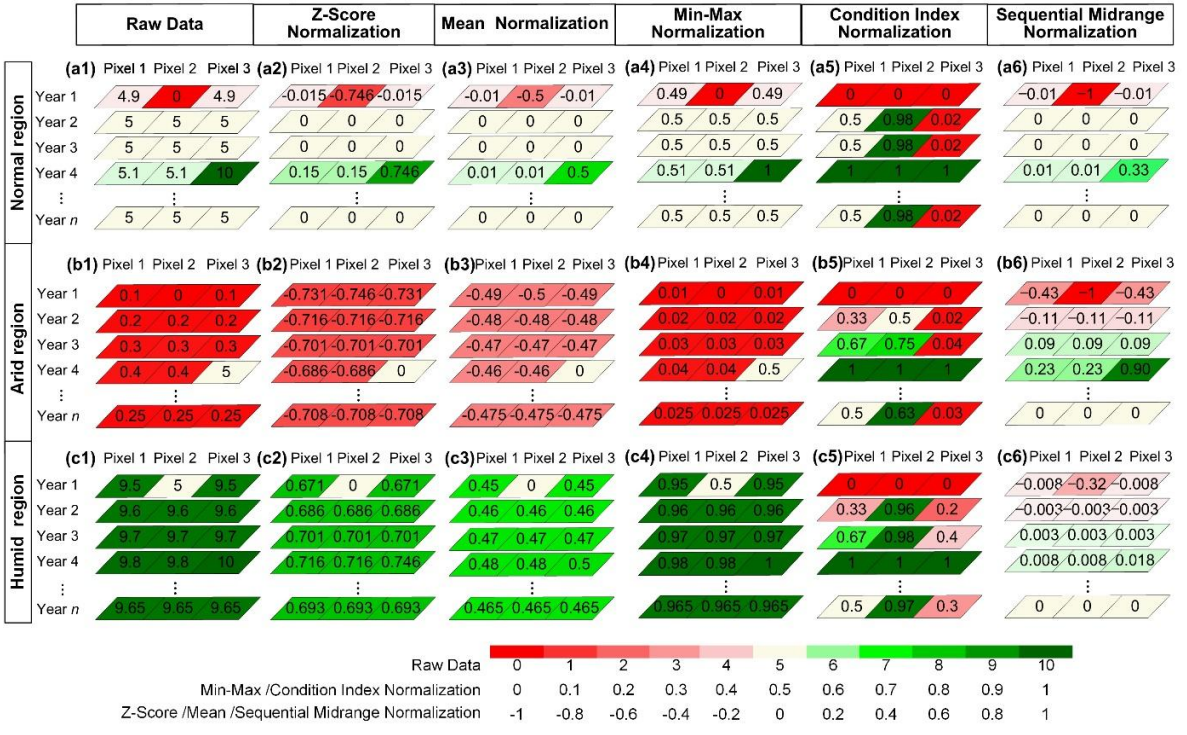
The Sequential Midrange Normalization (SMN) method is derived by optimizing Normalized Indices which are proposed in our previous study (Liu et al., 2021) and used to monitor changes in precipitation (Normalized Precipitation Index, NPI), soil moisture (Normalized Soil Moisture Index, NSMI), and crop growth status (Normalized Vegetation Index, NVI). etc. They are defined as (Eq. (11)):

$$SMN_{i,t,s} = \frac{F_{i,t,s} - \bar{F}_{i,s}}{F_{i,t,s} + \bar{F}_{i,s}}, i \in (\mathrm{P}, \mathrm{SM}, \mathrm{SR}, \mathrm{AT}, \mathrm{EVI2}, \mathrm{GPP}) \quad (11)$$

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

where $F_{i,t,s}$ is the pixel current value, the $\overline{F}'_{i,s}$ equals the mean value from 40% percentile to 70% percentile of the target pixel (as shown in Fig. 4) over study time based on their precipitation (or soil moisture or vegetation index, etc.) values. $SMN_{i,t,s}$ change from -1 to 1, and the value of 0 is set as the threshold to monitor the anomalous change. When $SMN_{i,t,s}$ equals 0, $F_{i,t,s}$ equals $\overline{F}'_{i,s}$, which is defined as $Standard_{SMN}$.

For DL model training, the ZSN, MN, MMN, or CIN methods are usually used to reduce errors caused by different units and improve the training efficiency. However, due to uncertainties in data collection quality or the data normalization method, the data used as inputs to the model often cannot accurately reflect the true state of the paddy rice growth environment, resulting in the model's low yield prediction accuracy. For illustration, we consider precipitation as an example (Fig. 4). Assume that the precipitation range (True value) is between 0 and 10, where 0 means no precipitation, 10 means the maximum precipitation recorded in the history of all regions, and 5 is the normal precipitation in an ordinary semiarid and semi-humid region. Pixel 1 represents normal pixels that show no extreme drought or extraordinary rainstorm event occurring, or places where both have occurred with similar severity in all monitoring years; Pixel 2 represents only severe drought events that occurred in a certain year; Pixel 3 represents only severe humid events that occurred (such as a sudden increase in precipitation, sudden irrigation, dry land becoming paddy field, etc.).



329

330 **Fig. 4.** The simulation of (2) Z-Score Normalization, (3) Mean Normalization, (4) Min-Max Normalization, (5) Condition Indices Normalization, and (6) Sequential Midrange

331 Normalization methods used in (a) normal region, (b) arid region and (c) humid region, using the (1) Raw Data of various pixels for many years. Pixel 1 represents normal

332 pixels; Pixel 2 represents pixels where only severe drought events occurred; Pixel 3 represents pixels where only severe flood events occurred.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

When using multiyear RS data to predict crop yield, the purpose of the input data normalization is to compare current state with a standard value ($Standard_{ZSN}$/$Standard_{MN}$/$Standard_{MMN}$/$Standard_{CIN}$/ $Standard_{SMN}$) assessing the degree of drought (deficit) or moisture (excess). Therefore, this standard value needs to be typical and can represent the normal level of the pixel over a long period of time. Because the extreme values are added to the calculation of standard value of MN, MMN, or CIN, they cannot represent well the real normal level of the pixels (Fig. 4(b3, b4, b5, c3, c4, c5)). These extreme values are often caused by the occurrence of unusually extreme flood (Fig. 4(b1)) or drought (Fig. 4(c1)) events. In addition, during ZSN, MN, and MMN calculation, the detailed information and outliers of the original data may be clipped or scaled, thereby damaging the integrity of the data. Especially when incorporating all pixel values over time and space in our study, the details of the original data are severely scaled. SMN was devoid of the aforementioned limitations inherent to these indices, enabling it to effectively monitor the relative changes in the true values of pixels in both arid and humid regions (as illustrated in the fourth column of Fig. 4). This was attributed to the following two reasons: (1) Accurate outliers handling. SMN was not sensitive to outlier in the data which often affect the mean and variance of the data, thus impairing the accuracy performance of the model; (2) Appropriate hyperparameter selection. The computation method of $Standard_{SMN}$ can represent pixel's consistent normative state over prolonged durations.

Based on the 8-day 1-km resolution datasets of P, SM, SR, T, EVI2, GPP, we calculated the six features for all rice-growing counties in the study area from 2008 to 2017 using ZSN, MN, MMN, CIN, and SMN. For the independent variables DEM, Lon, and Lat, which each have only one set of data, we replicated them 10 times across 46 instances to match the spatiotemporal resolution of the other variables and applied only the MMN method. From an initial set of approximately 120 rice-growing counties, we selected 83 that had consistently

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

cultivated rice for 10 consecutive years and had no missing data (Among them, there are 52 counties in Hubei Province and Jiangsu Province, and 31 counties in Anhui Province). Before inputting the data of each year into the model, we implemented a duplication procedure to augment the dataset, serving the purpose of data augmentation. We divided the data into training, validation, and testing sets by year, as detailed in Section 4.1.1.

3.3 The structure of the developed new model

The model AgroCLA (Agro-CNN-LSTM-A) consists of three main components (Fig. 5): (a) Convolutional Neural Network, (b) Long Short-Term Memory, and (c) Attention Mechanism. This model has seven hierarchical layers, which include: an input layer, a 1D CNN layer, two LSTM layers, an attention mechanism layer, a fully connected layer, and an output layer. The input layer receives multiple feature sequences $\{x_1, x_2 \cdots x_t\}$. These features include: P, SM, SR, AT, EVI2, GPP, DEM, Lon, and Lat, for each of the 46 instances annually. The output layer is responsible for predicting rice yield.

Similar to local feature extraction techniques in signal processing, the CNN can capture local temporal dependencies and pattern features in time series data. This characteristic is particularly effective in capturing short-term trends and seasonal changes (Tiao et al., 1981; Turner et al., 1999; Heaton et al., 2016). Although traditional time series analysis often relies on intricate feature engineering, such as selecting appropriate lagged variables and decomposing trends, CNN can automatically extract key features from raw data, significantly reducing the complexity of manual feature engineering. Additionally, CNN can be combined with other DL network structures, such as RNN, LSTM, and Gated Recurrent Unit (GRU). Compared to ALSTM, our model enhances its capability in feature extraction and sequence modeling. 1D CNN excels at extracting local

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

features from sequential data, such as short-term patterns and local trends, while LSTM is adept at capturing long-term dependencies within the sequence. By integrating both architectures, the model can simultaneously leverage local feature representations and long-range dependencies, enabling a more comprehensive understanding of the data's characteristics. (Jeong et al., 2022). The attention mechanism layer can generate different attention values based on the output of the LSTM to reflect the importance of feature vectors during the rice growth phase for yield (Tian et al., 2021; Liu et al., 2022a).

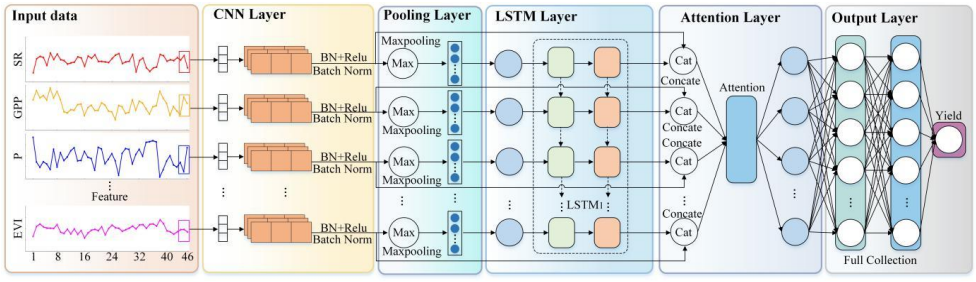


Fig. 5. Overall structure of the AgroCLA model for county-level rice yield prediction.

To ensure the reproducibility of our experiments, our development and experimental environments are as follows: The operating system is Windows 11; the programming language environment is Python 3.9; the DL framework is PyTorch 2.0.1. Other libraries and tools include Numpy 1.25.1, matplotlib 3.3.4, scikit-learn 1.1.1, pandas 2.0.3, and so on. We employed a Grid Search strategy (Breiman et al., 1985) to systematically explore the parameter space ensuring the identification of the optimal model configuration, and identified the optimal parameters: the number of units in the hidden layer is 128; there are two layers in the LSTM; the batch size is 512; the learning rate is 0.00001; the momentum is 0.9; the L2 regularization coefficient is 0.001; and the chosen optimizer is Adam (Table S1).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

4.Experiment result and discussion

4.1. Performance comparison of normalization method

4.1.1 The result of different normalization method

To evaluate the influence of normalization methods on rice yield prediction accuracy, we compared the performance of ZSN, MN, MMN, CIN, and SMN, and used data from 2008 to 2014 as the training set, 2015 and 2016 as the validation set. and 2017 as the test set. The five different datasets were fed into six models: CNN, LSTM, SSTNN, ALSTM, DDCN, and AgroCLA to calculate the yield prediction accuracy. Fig. 6(a) and (b) display the R^2 results computed for the validation (rice yield prediction accuracy) and test datasets (temporal generalization test), respectively. As shown in Fig. 6, different normalization methods have significantly different effects on the accuracy improvement of the model. Among them, the SMN method performed the best when applied to all six models, with an average R^2 of 0.733 on the test set. The ZSN was suboptimal, with an average test R^2 of 0.655. The MMN input obtained the lowest test R^2 of 0.557 (Table S2 and S3). Compared to traditional MMN and MN, the average test R^2 of SMN increased by 0.176 and 0.170, respectively, with the designed AgroCLA model achieving the highest accuracy (with R^2=0.841).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

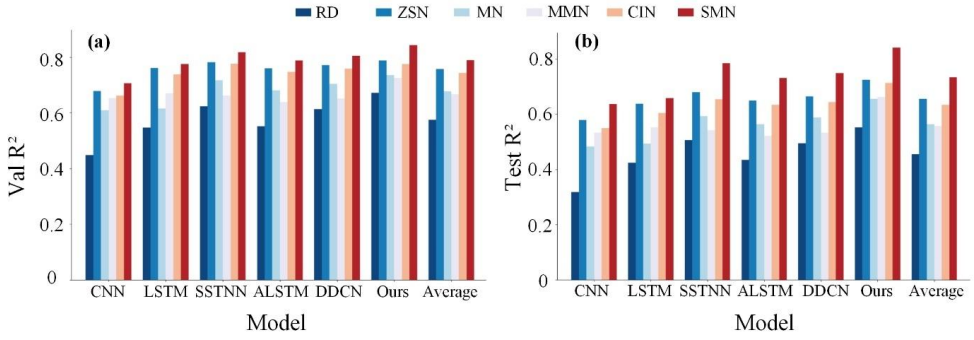


Fig. 6. The Raw Data (RD), Z-Score Normalization (ZSN), Mean Normalization (MN), Min-Max Normalization (MMN), Condition Indices Normalization (CIN), and Sequential Midrange Normalization (SMN) prediction performance of six models in which (a) represented the R^2 of validation dataset and (b) represented the R^2 of test dataset. The A in model name represented the Attention layers in deep learning.

This discrepancy arose from the variations in input data quality caused by different data normalization methodologies (Singh et al., 2020). The criteria for normalization method selection necessitated its ability to reflect the relative changes in the true values of pixels across varied regions, whether those values represented precipitation, soil moisture, or vegetation conditions, etc. ZSN, MN, and MMN methods, during the computation, normalized the input data across both time and space. Such a computational approach tended to overlook crops cultivated across diverse regions, which may have had distinct requirements due to differences in topography, crop variety, and other factors.

4.1.2 The interpretability of SMN

Using a real case study, in mid-June to mid-July 2014, precipitation in December was also much lower than in previous years. However, from early August to mid-September 2014, Jiangsu and Anhui Province experienced continuous rainy days (Yin et al., 2020), with the average precipitation being 60% higher than usual for the

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

same period in normal years. This period included several heavy rainfalls, sometimes accompanied by typhoons, resulting in field waterlogging and crop damage (Hu et al., 2021). Five normalization methods, ZSN, MN, MMN, CIN, and SMN, were applied to the original precipitation data (Fig. 7). Compared to the actual situation, the precipitation represented by the ZSN, MMN and CIN methods was significantly lower than the actual precipitation, they failed to detect the excessive precipitation events in August to September. MN could relatively reflect the occurrence of drought and precipitation events. However, it maintained a relatively stable value throughout all time periods leading to an inability to accurately reflect the severity of disasters, thereby reducing the accuracy of yield prediction.

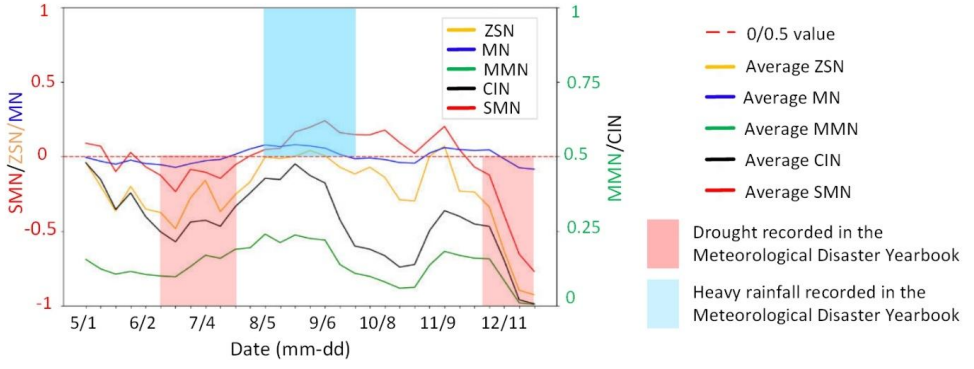


Fig. 7. Comparison of ability to reflect reality with different normalization methods. The average rice pixel values calculated by the ZSN, MN, MMN, CIN, and SMN methods based on the original precipitation data of Jiangsu Province from May 2014 to December 2014, which were updated every eight days and color-coded differently. 0.5 is the threshold of MMN/CIN, whereas the other indices use 0.

Moreover, due to the unreasonable calculation methods of MMN and CIN, the occurrence of this extreme precipitation event resulted in a sustained severe overestimation of drought. In regions where both droughts and floods incidents manifest with similar severity, CIN offered reliable accuracy in monitoring. However, regardless

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

of the actual occurrence of extreme events, every pixel invariably exhibited values of 0 and 1 (Fig. 4(c5)) symbolizing extreme drought and flood scenarios, a representation that often diverges from reality. The MMN and CIN calculation method fall short in reflecting the genuine relative variations in meteorological conditions faced by crops in different regions. This misrepresentation of normal conditions ultimately reduces the accuracy of crop yield predictions (Liu et al., 2021).

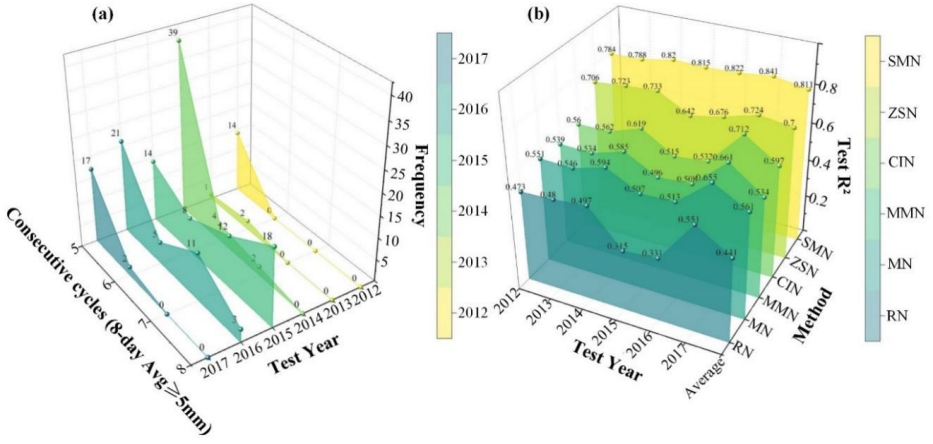


Fig. 8. Performance of different normalization methods during extreme precipitation events (a) Statistical frequency of events with an average precipitation greater than 5mm for 5-8 consecutive cycles (each cycle is 8 days) from 2012 to 2017. (b) Comparison of model performance using different normalization methods in normal years/disaster-affected years.

In contrast, the SMN was devoid of the aforementioned limitations inherent to these indices. This was attributed to the computation method of SMN, which can effectively monitor the relative changes in the true values of pixels in both arid and humid regions (as illustrated in Section 3.2.5). The inclusion of SMN, due to its superior ability to portray the authentic variations of meteorological factors across diverse regions, not only bolstered the integrity of the model's input data but also enhanced the precision of yield predictions and

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

reinforced the model's generalization. These are why SMN was still able to achieve an accurate rice yield prediction during the super El Niño events of 2015-2016 (Fig. 8(b) with yellow color).

The super El Niño events occurred in 2015-2016, which enhanced monsoon activity, thereby triggering persistent precipitation in southern China (Fig. 8(a)) and South America (Guo et al., 2016; Wang et al., 2019). When using different normalization methods combined with the AgroCLA model to predict rice yields during extreme precipitation years, except the SMN method, the prediction accuracy of all normalization methods showed a sharp decline (Fig. 8(b)). The persistent precipitation during 2015-2016 caused deviations in the characterization of precipitation data by other normalization methods, whereas the SMN method did not encounter this issue. This is because the SMN method can monitor the relative changes of actual values of various factors with high quality (Fig. 4) and accurately characterize whether flooding/drought events occur (Fig. 7), enabling the SMN method to maintain high prediction accuracy even under extreme climate scenarios. The R^2 values of the SMN method in 2015 and 2016 were 0.815 and 0.822, respectively, which were 17.3% and 14.6% higher than those of the second-best ZSN method ($R^2 = 0.642$ and $R^2 = 0.676$).

4.2. Validation of model robustness at different data collection times

4.2.1 The result of different data collection times

During the actual growth process of rice, the data volume of various variables accumulated over time. Starting from early May when rice sowing commenced in our study area (Fig. 3), we evaluated the model's predictive performance as it varied over time. We designed five growth stage data combinations: 16-19, 16-23, 16-27, 16-32, 16-39, combined with 1-46, and GPs, yielding a total of seven different data input scales. Based on

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

the SMN selected in Section 4.1, we used data from 2008 to 2014 as the training set, data from 2015 and 2016 as the validation set, and data from 2017 as the test set.

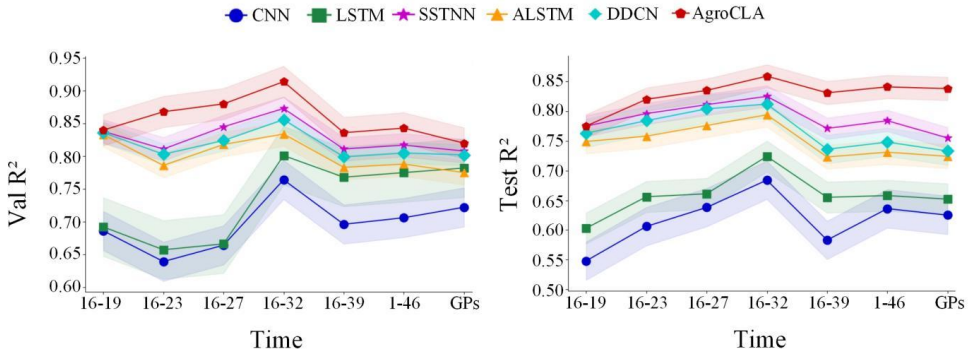


Fig. 9. The average R^2 and standard deviation changes in model (a) accuracy validation and (b) temporal generalization testing under different combinations of input data across various growth periods. The 16-19 is from sowing to seedling, 16-23 is from sowing to the end of tillering, 16-27 is from sowing to the end of booting, 16-32 is from sowing to the end of flowering, 16-39 is from sowing to the end of maturity, 1-46 represents the data for a whole year, and GPs is the average of five growth periods.

The study findings indicated that the AgroCLA model (with test R^2 ranging from 0.775 to 0.859) substantially outperformed other models (with test R^2 ranging from 0.548 to 0.825) (as shown in Table S4-S5). Moreover, the performance of all models was at its lowest during the early growth stages of rice, as depicted in Fig. 9(b). As rice progressed through its active growth phase and the amount of information increased, the predictive accuracy of the model gradually improved (Sheng et al., 2022). The model exhibited optimal performance around mid-August during the heading and flowering stage, achieving a test R^2 value of 0.859, which is two months ahead of the harvest season. This could be attributed to the fact that, in the early growing

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

season, the correlation between RS and weather features with rice yield was relatively weak, resulting in subpar performance across all methods (Johnson, 2014).

Notably, as rice entered the later stages of growth, there was a decrease in yield prediction accuracy, which was because changes in vegetation indices were highly correlated with crop chlorophyll content. During the reproductive growth stage of rice, as the leaves turn from green to yellow, the vegetation indices cannot accurately reflect changes in rice yield (Liu et al., 2022b). Additionally, research by Gu et al., (2022) has confirmed that the vegetation index during the heading/flowering stage has higher accuracy in predicting rice yield in the later stages. After the flowering stages, the model's performance became stable. Compared to the other five methods, the developed AgroCLA model exhibited the best performance, achieving a validation R^2 of 0.914 and a testing R^2 of 0.859.

4.2.2 The interpretability of AgroCLA

Upon further analysis of all the study findings, the AgroCLA model exhibited the most outstanding overall performance. This was attributable to CNN's capability to effectively extract local features from time series data, while LSTM captured long-range temporal dependencies. The fusion of these two facilitated the model's aptitude to discern intricate patterns within time sequences. The attention mechanism permitted the model to place heightened emphasis on more pivotal time steps when making predictions. For instance, in certain forecasting tasks, recent observations are often more important than earlier ones (Shen et al., 2023). By employing the attention mechanism, the model can autonomously learn these weights, thereby focusing more on inputs that are paramount for prediction. To demonstrate the reliability of the attention mechanism, we visualized the attention weights of the AgroCLA model when predicting rice yield (Fig. 10).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

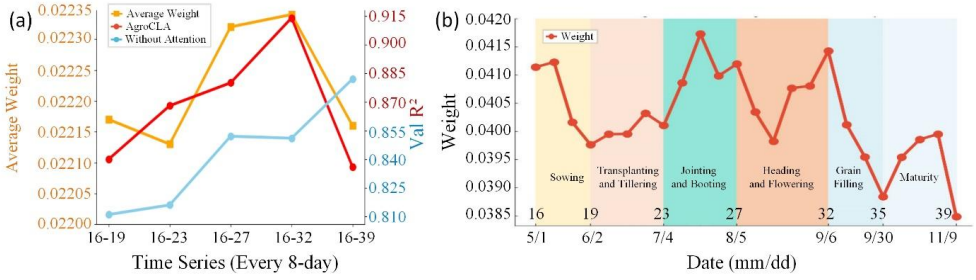


Fig. 10. The attention changes of AgroCLA model (a) in five different growth period combinations and (b) in each 8-day time step of which each growth period is color-coded differently. In (a), the red line represents the R^2 of the AgroCLA model, while the blue line represents the R^2 for the AgroCLA model without the add the attention mechanism.

Compared to AgroCLA without attention (with validation average R^2 = 0.842), the model with the attention mechanism (with validation average R^2 = 0.867) Moreover, the accuracy trend closely mirrored the trend of average attention weight, with both peaking at the end of the flowering stage (Fig. 10(a)). Further analysis of the 8-day interval attention weights revealed two significant peaks at the end of the jointing stage and the end of the flowering stage. This indicated that the attention mechanism in AgroCLA assigned greater importance to data from the jointing and booting stage, as well as the flowering and grain filling stage during yield prediction (Fig. 10(b)). The results in Fig. 10 have shown that the incorporation of the attention mechanism further bolsters the model's predictive prowess and interpretability.

4.3. Model generalization test

4.3.1 Temporal generalization

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

Based on the highest-quality data normalization methods and optimal data collection times identified in Sections 4.1 and 4.2, all models were evaluated over six test years from 2012 to 2017 (Table 2). For each test year, data from all preceding years since 2008 were used for model training and validation. Specifically, data from the two years prior to the test year were designated as the validation dataset, while the remaining served as the training dataset. For instance, when predicting yields for the year 2016, we used data from 2008 to 2013 as the training set, data from 2014 and 2015 as the validation set, and data from 2016 as the test set. AgroCLA model consistently achieved the highest predictive accuracy and temporal generalization in all years.

Table 2

The R^2 for temporal generalization testing of models.

Test Year	CNN	LSTM	SSTNN	ALSTM	DDCN	AgroCLA
2012	0.621	0.640	0.733	0.704	0.712	**0.783**
2013	0.663	0.699	0.789	0.753	0.761	**0.799**
2014	0.697	0.709	0.794	0.787	0.791	**0.822**
2015	0.674	0.722	0.796	0.779	0.798	**0.816**
2016	0.679	0.724	0.803	0.782	0.809	**0.818**
2017	0.684	0.731	0.825	0.794	0.812	**0.859**
Average	0.669	0.704	0.790	0.767	0.781	**0.816**

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

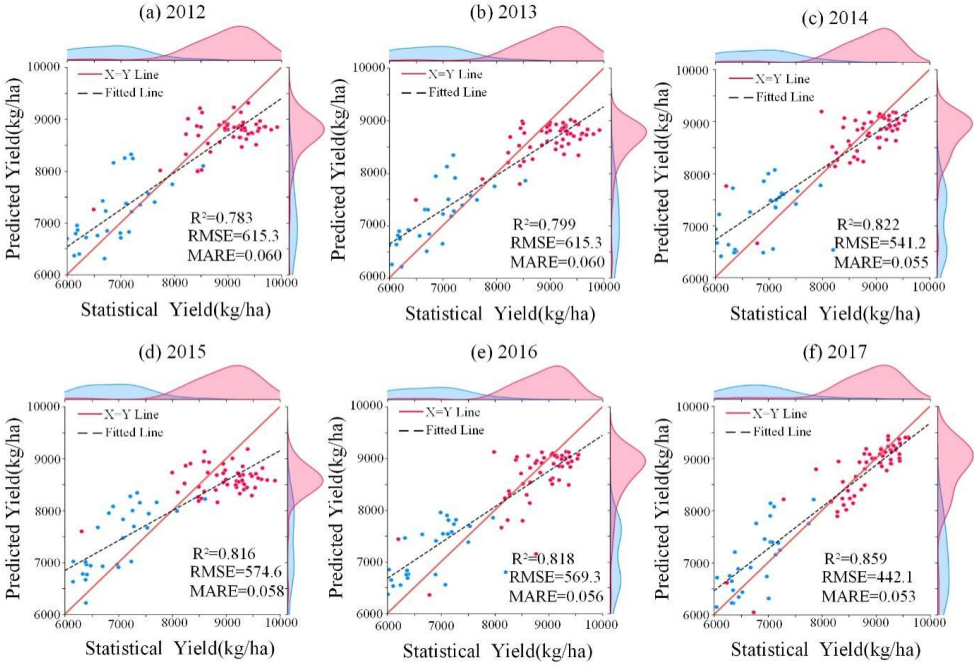


Fig. 11. Scatter plots of statistical yields vs. predicted yields of AgroCLA in six testing years: (a) 2012; (b) 2013; (c) 2014; (d) 2015; (e) 2016; (f) 2017. The counties of Jiangsu Province and Anhui Province are represented by red and blue, respectively.

The CNN model excelled in processing image data but exhibits reduced performance in multi-feature time-series tasks (Hao et al., 2024). While it can effectively extract local spatiotemporal patterns in such tasks, its performance diminished with longer and more complex sequences (Wibawa et al., 2022). Furthermore, manually defined convolutional window sizes may lead to loss of crucial information, especially when temporal dependencies exceed the convolution window size (Zhang et al., 2016). In contrast, the LSTM model can naturally handle sequences of arbitrary lengths through its recurrent structure, comprehensively evaluating each event in the time series and retaining memory of past events, thereby capturing long-term and complex temporal dependencies. The AgroCLA model can effectively combine the strengths of both models, capturing both local

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

and global information simultaneously. Incorporating attention mechanisms can assign higher weights to important information in the time series, thereby enhancing predictive accuracy. The multi-year average R^2 of AgroCLA model was 0.816, ranking first, and its temporal generalization ability was also the best (Fig. 11.). As shown in Table 3, it can be observed that, the AgroCLA model consistently exhibited commendable performance in the RMSE tests. The multi-year average RMSE is 559.6, the lowest among all tested models.

Table 3

The RMSE (kg/ha) for temporal generalization testing of models.

Test Year	CNN	LSTM	SSTNN	ALSTM	DDCN	AgroCLA
2012	802.3	753.2	683.3	664.7	652.2	615.3
2013	778.6	722.5	631.8	659.8	655.6	615.3
2014	687.1	625.2	573.1	588.3	581.4	541.2
2015	676.4	588.6	565.9	593.3	579.1	574.6
2016	664.7	582.4	524.8	585.7	571.4	569.3
2017	588.4	574.9	470.3	545.7	508.5	442.1
Average	699.6	641.1	574.9	606.3	591.4	559.6

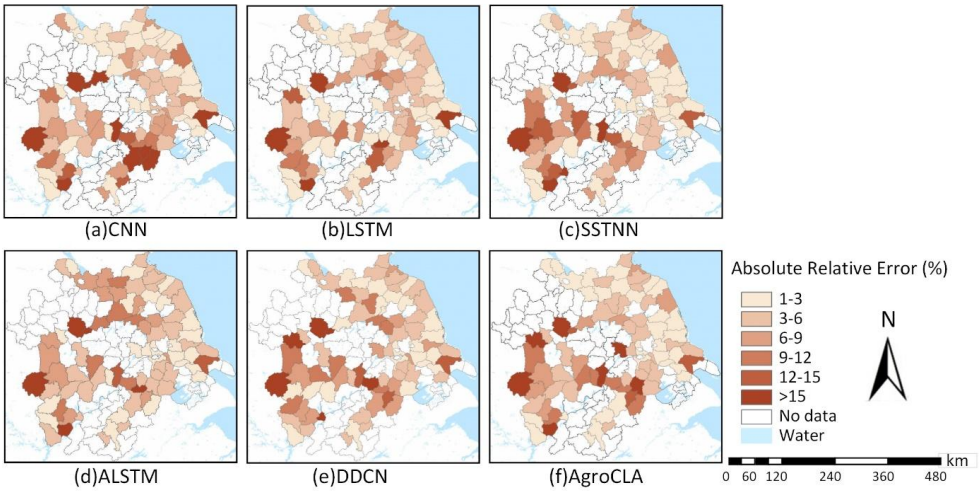


Fig. 12. The average absolute relative error maps of (a) CNN, (b) LSTM, (c)SSTNN, (d)ALSTM, (e) DDCN, and (f)AgroCLA from 2012 to 2017.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

To further elucidate the temporal generalization of different models, we averaged the results from 2012 to 2017 and presented the absolute relative error diagrams for each model (Fig. 12). The results revealed that AgroCLA had lower error distributions than the other models. Compared to Jiangsu Province, higher prediction errors were observed in Anhui Province across all methods. This can be attributed to the superior accuracy of rice distribution extraction results in Jiangsu Province compared to Anhui Province (Liu et al., 2020). In this study, we used previously developed rice distribution products as crop masks, aggregating multi-source data at the county level for all years. However, inaccuracies in crop area identification may have introduced noise into the training data. Additionally, the economic level of Jiangsu Province surpassed that of Anhui Province. A higher economic level suggests well-educated farmers, advanced agricultural techniques, and comprehensive irrigation systems. These factors contributed to Jiangsu Province achieving consistently high rice yields, leading to higher model accuracy and reduced prediction errors (Zhuo et al., 2022).

To further demonstrate the consistency between statistical and predicted yields in 2017, we plotted the scatter diagrams for all methods in Fig. 13. The results showed that among the six models, AgroCLA once again achieved the best consistency (as shown in Fig. 13(f)), with the highest testing R^2 of 0.859 and the lowest RMSE and MARE values of 442.1 and 0.053, respectively. Compared with the ALSTM model (R^2 = 0.749), AgroCLA improved the yield prediction accuracy by approximately 0.11. The second-best model was SSTNN, with an R^2 of 0.825, followed by DDCN with an R^2 of 0.812. As demonstrated in previous studies, AgroCLA consistently exhibits superior temporal generalization and predictive stability.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

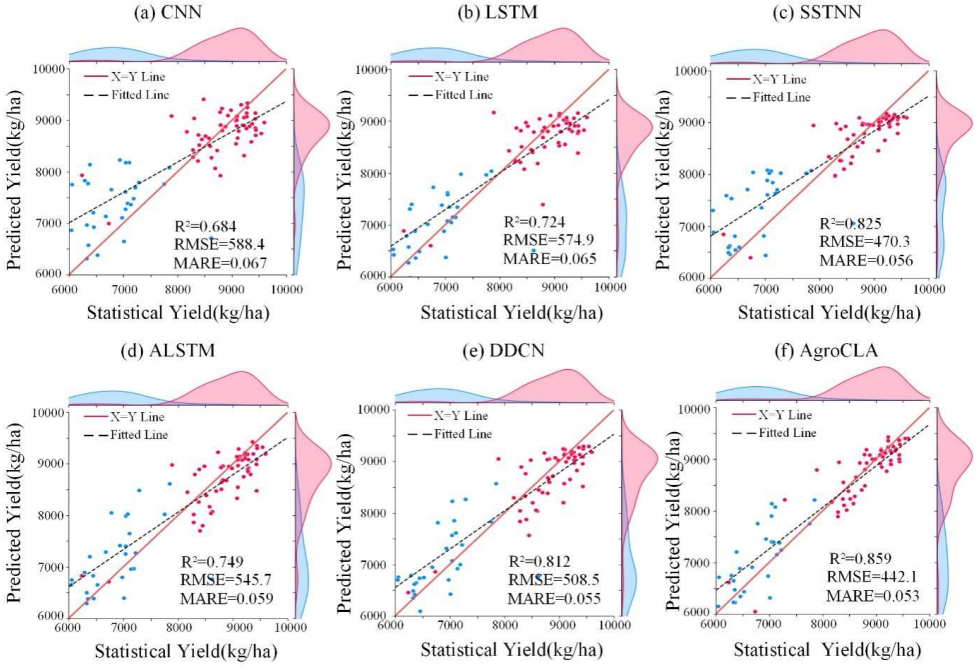
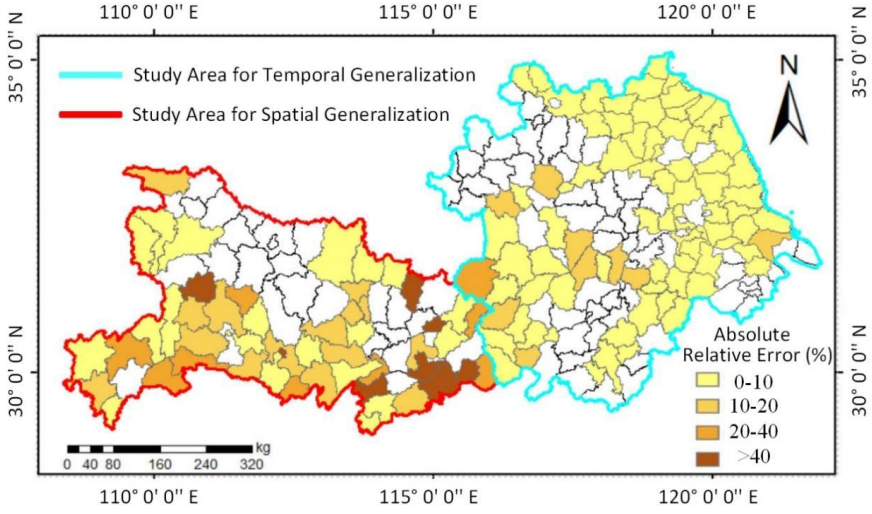


Fig. 13. Scatter plots of statistical yields vs. predicted yields of (a) CNN, (b) LSTM, (c) SSTNN, (d) ALSTM, (e) DDCN, and (f) AgroCLA in 2017. The counties of Jiangsu Province and Anhui Province are represented by red and blue, respectively.

4.3.2 Spatial generalization

To validate the spatial generalization capability of the AgroCLA model, we conducted experiments in county-level regions of Hubei Province in 2017 (where the area of rice paddies ranges from 50 to 70,000 hectares).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025



577

578 **Fig. 14.** The absolute relative error maps of temporal and spatial generalization testing in 2017.

579 As shown in Fig.14, we visualized the prediction accuracy for each county. Although 68.6% of the counties
580 showed ARE below 20%, the ARE in some counties in Huangshi and Huanggang was greater than 40%. The
581 larger errors observed in Huangshi and Huanggang were attributed to the relatively small rice cultivation areas in
582 these cities. DL models rely on a large number of input features to learn spatial and temporal patterns. When the
583 rice cultivation area in a county is relatively small, the proportion of effective agricultural features (such as EVI
584 and GPP) in the remote sensing signals is low within each pixel. As a result, these signals are more susceptible to
585 interference from non-agricultural land types (such as urban areas, forests, and water bodies), thereby
586 exacerbating the mixed pixel effect. (Joshi et al., 2023).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

Table 4

The impact of the rice cultivation area on the model's spatial generalization in Hubei province.

Area (ha)	County Number	Average ARE (%)
>50000	1	4.5
20000-50000	8	9.2
5000-20000	18	21.1
0-5000	24	23.3

Therefore, we compiled data on the rice cultivation area in each county of Hubei Province along with the corresponding spatial generalization testing ARE (Table 4). There was a positive correlation between the rice cultivation area and the average ARE of the models. Specifically, counties with the largest rice cultivation area (over 50,000 hectares) exhibited the lowest ARE, reaching 4.5%, while counties with smaller areas (0-5,000 hectares) showed an average ARE of 23.3%. This trend indicated that as the area increased, the predictive accuracy of the model improved (Cao et al., 2025). These results indicated that the AgroCLA model showed strong spatial generalization.

Table 5

Comparison of different models' efficiency in spatial generalization testing.

Models	**R^2 of Spatial Generalization testing**	**Time**	**Para**
CNN	0.654	**21.38s**	**106.8K**
LSTM	0.469	38.29s	203.6K
SSTNN	0.802	45.52s	258.1K
ALSTM	0.672	46.73s	267.7K
DDCN	0.691	76.42s	548.8K
AgroCLA	**0.833**	48.19s	268.4K

As shown in Fig. 13. and Table 5, the AgroCLA model achieved the best temporal generalization (R^2=0.859) and spatial generalization (R^2=0.833). At the same time, it also performed well in terms of efficiency, with a runtime of 48.19 seconds and 268.4K parameters. In contrast, although other models had advantages in terms of runtime and parameter count, such as the CNN model which had the lowest parameter count (106.8K)

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

and the shortest runtime (21.38 seconds), it had lower accuracy with a temporal generalization R^2 of 0.684 and spatial generalization R^2 of 0.654. Despite its more complex structure and higher computational cost, the AgroCLA model's superior temporal and spatial generalization should not be overlooked.

4.4. Feature importance analysis

The Shapley Additive Explanations (SHAP) based global explanation suggests that the latitude, Gross Primary Productivity (GPP), longitude and the Two-Band Enhanced Vegetation Index (EVI2) (Huang et al., 2013) are the main features used to predict the rice yield by our AgroCLA model (Fig. 14(b)). Broadly, these nine factors can be categorized into three groups, ranked in decreasing order of feature importance as: geolocation-related, crop growth-related, and climate-related. Our study area was located on both sides of the Qinling-Huaihe line, which serves as the climatic boundary between northern and southern China. South of the Qinling-Huaihe line, the region was predominantly paddy fields, while north of the line, it was mainly dryland (Ge et al., 2024). Factors like latitude, longitude, and soil properties collectively depict the long-term environmental characteristics and information of an area. High-yielding regions are typically characterized by fertile soils, abundant water resources, well-educated farmers, advanced agricultural techniques, comprehensive irrigation systems, and favorable climatic conditions. These features can be holistically described through their spatial attributes, whereas climatic factors lack this property (Fowler et al., 2024).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

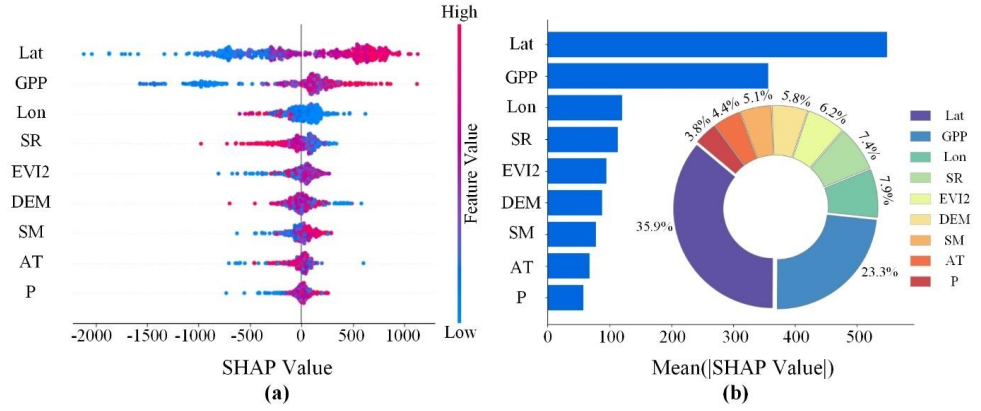


Fig. 15. SHAP analysis for drivers of AgroCLA prediction. (a) The bee swarm plot of SHAP values for the local interpretation of each sample at the specific feature dimension. (b) The bar plot and pie chart of the averaged absolute SHAP values for the global interpretation of input feature importance.

This aligns with existing research. For instance, Wang et al. (2020) demonstrated through spatial analysis that spatial factors not only influence the distribution of air temperature, solar radiation, and soil but also significantly influence crop growth stages. Additionally, the positive correlations of GPP and EVI2 with rice yield (Fig. 15(a)) are consistent with the findings of Zhang et al. (2016b) and Huang et al. (2021). Among meteorological factors, solar radiation is the most critical, followed by soil moisture, reaffirming findings from Liu et al. (2021). This is particularly relevant in the mid-lower Yangtze River plain, where prolonged rainy seasons make flooding more frequent than droughts. While short-term precipitation appears less critical due to modern irrigation systems, long-term precipitation patterns, particularly those influenced by El Niño events, remain pivotal for predicting rice yields (Stuecker et al., 2018). Therefore, in rice yield estimation, both short-term feature importance and long-term precipitation dynamics should be considered to capture their cumulative impact on crop growth and productivity (Fan et al., 2024).

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

5. Conclusion and future work

A proper normalization method can mitigate distortions in feature scaling caused by extreme values, which is critical for improving the accuracy of rice yield prediction. In this study, we implemented five data normalization methods: ZSN, MN, MMN, CIN, and SMN, which were fed into six models (CNN, LSTM, SSTNN, ALSTM, DDCN, and AgroCLA) based on data from 2008 to 2017 in eastern China. The predicted rice yield accuracy of all models based on the SMN index (R^2=0.733) surpassed that of traditional MMN (R^2=0.557) and MN (R^2=0.563) by average R^2 increases of 0.176 and 0.170, respectively. The SMN method significantly enhanced the quality of the input data, which in turn indirectly improved the model's predictive accuracy and generalization performance. It effectively captures the relative changes in the actual values of multiple variables, ensuring stable feature representation under varying conditions. Among them, the AgroCLA model exhibited the highest accuracy, with an average R^2 of 0.841. Using seven data combinations corresponding to different growth stages, we found that model performance peaked during the heading and flowering stage, approximately two months before harvest. Additionally, all models were evaluated over the six test years from 2012 to 2017. Results showed the AgroCLA model consistently achieved superior accuracy and generalization in all years, with the best test set R^2 reaching 0.859, and the lowest RMSE and MARE values being 442.1 and 0.053, respectively. The inclusion of the attention mechanism further enhanced the model's predictive capability and interpretability. Within our study area, geographical location exerted the most significant impact on yield prediction, followed by crop growth-related factors. Variations in crop mask accuracy and regional economic conditions amplified prediction uncertainties.

Improvements in subsequent phases can be achieved by further exploring aspects such as data quality discrepancies, data resolution and model lightweighting. This study leveraged the outcomes of prior research to

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

obtain high-precision rice distribution data and high-quality input from various factors. This approach minimized the potential negative impacts on yield prediction accuracy caused by noise in remote sensing data, cloud contamination, and sensor errors, all of which can increase data uncertainty. Furthermore, the discrepancies arising from the singular use of RS data or meteorological data inputs, essentially the differences in data resolution, introduce another layer of uncertainty that warrants in-depth discussion.

Funding

This work was supported by the National Key Research and Development Program of China (2023YFD2300300). It was also funded by the National Natural Science Foundation of China (No. 42401404, 42171314, 42001304) and Hubei Provincial Natural Science Foundation (No. 2024AFB130, 2024AFB217).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We are very grateful for the support by the Advanced Computing Center of the China Three Gorges University.

Author contributions

L.L.: Conceptualization, Data curation, Funding acquisition, Writing – original draft, Writing

K.Z.: Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing

Y.Z. and H.S.: Methodology, Project administration, Result Analysis

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

T.L., W.S. and W.L.: Validation, Supervision

P.S. and B.X.: Funding acquisition, Data curation

D.R. and H.J.: Writing-Reviewing and Editing.

All authors read and approved the final manuscript.

Data availability statement

The Rice map and Rice yield for the study sites in this paper is accessible for download at no cost from the following link: https://www.scidb.cn/en/s/bU3e2m

All the datasets used in this study are open access from various sources:

Precipitation: https://gpm.nasa.gov/data-access/downloads/gpm

Soil Moisture: https://doi.org/10.11888/Hydro.tpdc.271762

Solar Radiation: http://data.tpdc.ac.cn/zh-hans/data/8028b944-daaa-4511-8769-965612652c49

Air Temperature: https://www.earthdata.nasa.gov/centers/laads-daac

The two-band Enhanced Vegetation Index: https://ladsweb.nascom.nasa.gov/index.html

Gross Primary Productivity: https://lpdaac.usgs.gov/dataset_discovery/modis

DEM, Longitude and Latitude: http://srtm.csi.cgiar.org

All authors have read and agreed to the submitted version of the manuscript.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

References

Ba, J. L., Kiros, J. R., Hinton, G. E., 2016. Layer Normalization. Stat. 1050: 21. https://doi.org/10.48550/arXiv.1607.06450

Ben-Ari, T., Bo´e, J., Ciais, P., Lecerf, R., Van der Velde, M., Makowski, D., 2018. Causes and implications of the unforeseen 2016 extreme yield loss in the breadbasket of France. Nat. Commun. 9, 1627 https://doi.org/10.1038/s41467-018-04087-x

Bischl, B., Binder, M., Lang, M., Pielok, T., Richter, J., Coors, S., Thomas, J., Ullmann, T., Becker, M., Boulesteix, A. L., Deng, D., Lindauer, M., 2023. Hyperparameter optimization: Foundations, algorithms, best practices, and open challenges. Wires. Data. Min. Knowl. 13(2): e1484. https://doi.org/10.1002/widm.1484

Boschetti, M., Busetto, L., Manfron, G., Laborte, A., Asilo, S., Pazhanivelan, S., Nelson, A., 2017. PhenoRice: A method for automatic extraction of spatio-temporal information on rice crops using satellite data time series. Remote Sens. Environ. 194, 347-365. https://doi.org/10.1016/j.rse.2017.03.029

Breiman, L., Friedman, J. H., 1985. Estimating optimal transformations for multiple regression and correlation. J. Am. Stat. Assoc. 80(391): 580-598. https://doi.org/10.1080/01621459.1985.10478157

Busetto, L., Zwart, S. J., Boschetti, M., 2019. Analysing spatial–temporal changes in rice cultivation practices in the Senegal River Valley using MODIS time-series and the PhenoRice algorithm. Int. J. Appl. Earth Obs. Geoinf. 75, 15-28. https://doi.org/10.1016/j.jag.2018.09.016

Cai, Y. P., Guan, K. Y., Lobell, D., Potgieter, A. B., Wang, S. W., Peng, J., Xu, T. F., Asseng, S., Zhang, Y. G., You, L. Z., Peng, B., 2019. Integrating satellite and climate data to predict wheat yield in Australia using

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

709 machine learning approaches. Agric. For. Meteorol. 274, 144-159.
710 https://doi.org/10.1016/j.agrformet.2019.03.010

711 Cao, J., Zhang, Z., Luo, X., Z. Xu, J. L., Xie, J., Han, J. C., Tao, F. L., 2025. Mapping global yields of four major
712 crops at 5-minute resolution from 1982 to 2015 using multi-source data and machine learning. Sci Data. 12,
713 357. https://doi.org/10.1038/s41597-025-04650-4

714 Chu, Z., Yu, J., 2020. An end-to-end model for rice yield prediction using deep learning fusion. Comput.
715 Electron. Agr. 174, 105471. https://doi.org/10.1016/j.compag.2020.105471

716 Coops, N. C., Ferster, C. J., Waring, R. H., Nightingale, J., 2009. Comparison of three models for predicting
717 gross primary production across and within forested ecoregions in the contiguous United States. Remote
718 Sens. Environ. 113, 680-690. https://doi.org/10.1016/j.rse.2008.11.013

719 Curtis, A. E., Smith, T. A., 2016. Ziganshin, B. A., Elefteriades, J. A., The mystery of the Z-score. Aorta. 4(04):
720 124-130. https://doi.org/10.12945/j.aorta.2016.16.014

721 D'haene, B., Mestdagh, P., Hellemans, J., Vandesompele, J., 2012. miRNA expression profiling: from reference
722 genes to global mean normalization. Next-Generation MicroRNA Expression Profiling Technology:
723 Methods and Protocols. 822, 261-272. https://doi.org/10.1007/978-1-61779-427-8_18

724 Dou, Y. J., Huang, R., Mansaray, L. R., Huang, J. F., 2020. Mapping high temperature damaged area of paddy
725 rice along the Yangtze River using Moderate Resolution Imaging Spectroradiometer data. Int. J. Remote
726 Sens. 41, 471-486. https://doi.org/10.1080/01431161.2019.1643936

727 Fan, L., Fang, S. B., Fan, J. L., Wang, Y., Zhan, L. Q., He, Y. K., 2024. Rice Yield Estimation Using Machine
728 Learning and Feature Selection in Hilly and Mountainous Chongqing, China. Agriculture. 14(9): 1615.
729 https://doi.org/10.3390/agriculture14091615

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

730 Fortin, J. G., Anctil, F., Parent, L., Bolinder, M. A., 2011. Site-specific early season potato yield forecast by
731 neural network in Eastern Canada. Precis Agric. 12, 905-923. https://doi.org/10.1007/s11119-011-9233-6
732 Fowler, A., Basso, B., Maureira, F., Millar, N., Ulbrich, R., Brinton, W. F., 2024. Spatial patterns of historical
733 crop yields reveal soil health attributes in US Midwest fields. Sci. Rep. 14(1): 465.
734 https://doi.org/10.1038/s41598-024-51155-y
735 Ge, S. Q., He, Q. Q., Sarangi, S., 2024. Reaping what you sow: Historical rice farming and contemporary
736 cooperative behavior in China. J ECON BEHAV ORGAN. 217: 581-613.
737 https://doi.org/10.1016/j.jebo.2023.11.021
738 Ghaleb, F., Mario, M., Sandra, A. N., 2015. Regional Landsat-based drought monitoring from 1982 to 2014.
739 Climate. 3(3), 563-577. https://doi.org/10.3390/cli3030563
740 Gu, C., Ji, S., Xi, X., Zhang, Z. H., Hong, Q. Q., Huo, Z. Y., Li, W. X., Mao, W., Zhao, H. T., Zhang, R. H., Li,
741 B., Tan, C. W., 2022. Rice yield estimation based on continuous wavelet transform with multiple growth
742 periods. Front. Plant. Sci. 13: 931789. https://doi.org/10.3389/fpls.2022.931789
743 Guan, K. Y., Wu, J., Kimball, J. S., Anderson, M. C., Frolking, S., Li, B., Hain, C. R., Lobell, D. B., 2017. The
744 shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating
745 large-scale crop yields. Remote Sens. Environ. 199, 333-349. https://doi.org/10.1016/j.rse.2017.06.043.
746 Guo D., Wang, L. W., Li, Z. K., Su, Y. C., Tan, H., Huang, Y., 2016. Comparison between anomalies of summer
747 rainfall in China in decaying years during super El Niño events of 2015/2016 and 1997/1998. Trans Atmos
748 Sci. 39(6): 835-844. https://doi.org/10.13878/j.cnki.dqkxxb.20160828010.
749 Hao, J., Liu, F., 2024. Improving long-term multivariate time series forecasting with a seasonal-trend
750 decomposition-based 2-dimensional temporal convolution dense network. Sci. Rep. 14(1): 1689.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

751 https://doi.org/10.1038/s41598-024-52240-y

752 He, J., Yang, K., Tang, W. J., Lu, H., Qin, J., Chen, Y. Y., Li, X., 2020. The first high-resolution meteorological
753 forcing dataset for land process studies over China. Sci. Data. 7, 25.
754 https://doi.org/10.6084/m9.figshare.11558439

755 Heaton. J., 2016. An empirical analysis of feature engineering for predictive modeling. SoutheastCon 2016.
756 IEEE. 2016: 1-6. https://doi.org/10.1109/SECON.2016.7506650

757 Hu, R. X., Wang, L. J., 2021. Variation of high and low level circulation of Meiyu in Jiangsu Province in recent
758 30 years. Atmosphere. 12(10): 1258. https://doi.org/10.3390/atmos12101258

759 Huang, D., Chi, H., Xin F. F., Miyata, A., Kang, M., Liu, K. W., L, R. D., Dang, H. S., Qin, Y. W., Xiao, X. M.,
760 2021. Improved estimation of gross primary production of paddy rice cropland with changing model
761 parameters over phenological transitions. Ecological Modelling. 445: 109492.
762 https://doi.org/10.1016/j.ecolmodel.2021.109492

763 Huang, J. F., Wang, X. Z., Li, X. X., Tian, H. Q., Pan, Z. K., 2013. Remotely Sensed Rice Yield Prediction Using
764 Multi-Temporal NDVI Data Derived from NOAA's-AVHRR. PLoSONE. 8(8): e70816.

765 Ioffe, S., Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal
766 covariate shift[C]//International conference on machine learning. Pmlr. 2015: 448-456.
767 https://doi.org/10.48550/arXiv.1502.03167

768 Jeong, S., Ko, J., Yeom, J. M., 2022. Predicting rice yield at pixel scale through synthetic use of crop and deep
769 learning models with satellite data in South and North Korea. Sci. Total. Environ. 802,149726.
770 https://doi.org/10.1016/j.scitotenv.2021.149726

771 Jiang, H., Hu, H., Zhong, R., Xu, J. F., Xu, J. L., Huang, J. F., Wang, S. W., Ying, Y. B., Lin T., 2020. A deep

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

learning approach to conflating heterogeneous geospatial data for corn yield prediction: A case study of the US Corn Belt at the county level. Glob. Chang.Biol. 26, 1754-1766. https://doi.org/10.1111/gcb.14885

Jiang, Z. Y., Huete, A. R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. Remote Sens. Environ. 112, 3833-3845. https://doi.org/10.1016/j.rse.2008.06.006

Johnson, D. M., 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. Remote Sens. Environ. 141, 116-128. https://doi.org/10.1016/j.rse.2013.10.027

Joshi, A., Pradhan, B., Gite, S., Chakraborty, S., 2023. Remote-sensing data and deep-learning techniques in crop mapping and yield prediction: A systematic review. Remote Sens. 15(8): 2014. https://doi.org/10.3390/rs15082014

Kogan, F. N. 1995a. Application of vegetation index and brightness temperature for drought detection. Adv. Space Res. 15, 91-100. https://doi.org/10.1016/0273-1177(95)00079-T

Kogan, F. N., 1995b. Droughts of the Late 1980s in the United States as Derived from NOAA Polar-Orbiting Satellite Data. B. Am. Meteorol. Soc. 76, 655-668. https://doi.org/10.1175/1520-0477(1995)076<0655:DOTLIT>2.0.CO;2

Li, R., Li, C., Xu, X., Wang, J., Pan, Y., 2009. Winter wheat yield estimation based on support vector machine regression and multi-temporal remote sensing data. Trans. Chin. Soc. Agric. Eng. 25, 114-117. https://doi.org/10.3969/j.issn.1002-6819.2009.07.021

Liu, J. H., Li, L., Huang, X., Liu, Y. M., Li, T. S., 2019. Mapping paddy rice in Jiangsu Province, China, based on phenological parameters and a decision tree model. Front Earth. Sci-Prc. 13, 111-123. https://doi.org/10.1007/s11707-018-0723-y

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

793 Liu, L. W., Lu, C. T., Wang, Y. M., Lin, K. H., Ma, X. M., Lin, W. S., 2022b. Rice (Oryza sativa L.) growth
794 modeling based on growth degree day (GDD) and artificial intelligence algorithms. Agriculture. 12(1): 59.
795 https://doi.org/10.3390/agriculture12010059
796 Liu, L., Huang, J. F., Xiong, Q. X., Zhang, H. J., Song, P. L., Huang, Y. H., Dou, Y. J., Wang, X. Z., 2020.
797 Optimal MODIS data processing for accurate multi-year paddy rice area mapping in China. Gisci. Remote
798 Sens. 57, 687-703. https://doi.org/10.1080/15481603.2020.1773012
799 Liu, L., Huang, R., Cheng, J. F., Liu, W. W., Chen, Y., Shao, Q., Duan, D. D., Wei, P. L., Chen, Y. Y., Huang, J. F.,
800 2021. Monitoring meteorological drought in southern China using remote sensing data. Remote Sens.
801 13(19), 3858 https://doi.org/10.3390/rs13193858
802 Liu, Y. Y., Wang, S. Q., Chen J. H., Chen, B., Wang, X. B., Hao, D. Z., Sun, L. G., 2022a. Rice yield prediction
803 and model interpretation based on satellite and climatic indicators using a transformer method. Remote Sens.
804 14(19): 5045. https://doi.org/10.3390/rs14195045
805 Ma, Y. C., Zhang, Z., Kang, Y. H., Özdoğan, M., 2021. Corn yield prediction and uncertainty analysis based on
806 remotely sensed variables using a Bayesian neural network approach. Remote Sens. Environ. 259, 112408.
807 https://doi.org/10.1016/j.rse.2021.112408
808 Nevavuori, P., Narra, N., Lipping, T., 2019. Crop yield prediction with deep convolutional neural networks.
809 Comput. Electron. Agr. 163: 104859.
810 Passalis, N., Tefas, A., Kanniainen, J., Gabbouj, M., Iosifidis, A., 2019. Deep Adaptive Input Normalization for
811 Time Series Forecasting. IEEE. T. Neur. Net. Lear. 31, 3760-3765.
812 https://doi.org/10.1109/TNNLS.2019.2944933
813 Patro, S., Sahu, K. K., 2015. Normalization: A preprocessing stage. Int. Adv. Res. J. Sci. Eng. Technol. 1503,

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

06462. https://doi.org/10.17148/IARJSET.2015.2305

Qiao, M., He, X., Cheng, X., Li, P., Luo, H., Zhang, L., Tian, Z., 2021. Crop yield prediction from multi-spectral, multi-temporal remotely sensed imagery using recurrent 3D convolutional neural networks. Int. J. Appl. Earth Obs. Geoinf. 102, 102436. https://doi.org/10.1016/j.jag.2021.102436.

Rattis, L., Brando, P. M., Macedo, M. N., Spera, S. A., Castanho, A. D., Marques, E. Q., Costa, N. Q., Silverio, D. V., Coe, M. T., 2021. Climatic limit for agriculture in Brazil. Nat. Clim. Change. 11, 1098-1104. https://doi.org/10.1038/s41558-021-01214-3

Reichstein, M., Camps, V. G., Stevens, B., Jung M., Denzler, J., Carvalhais, N., Prabhat., 2019. Deep learning and process understanding for data-driven Earth system science. Nature. 566, 195-204. https://doi.org/10.1038/s41586-019-0912-1

Reuter, H. I., Nelson, A., 2007. Jarvis, A., An evaluation of void-filling interpolation methods for SRTM data. Int. J. Geogr. Inf. Sci. 21, 983-1008. https://doi.org/10.1080/13658810601169899

Running, S. W., Zhao, M., 2019. MOD17 User's Guide. User's Guide. https://lpdaac.usgs.gov/documents/103/MOD17_User_Guide_V6.pdf

Schwalbert, R. A., Amado, T., Corassa, G., Pott, L. P., Prasad, P. V. V., Ciampitti, I. A., 2020. Satellite-based soybean yield forecast: Integrating machine learning and weather data for improving crop yield prediction in southern Brazil. Argic. For. Meteorol. 284, 107886. https://doi.org/10.1016/j.agrformet.2019.107886

Shen, L. F., Kwok J. 2023. Non-autoregressive conditional diffusion models for time series prediction. International Conference on Machine Learning. PMLR. 202:31016-31029. https://doi.org/10.48550/arXiv.2306.05043

Sheng, R. T. C., Huang, Y. H., Chan, P. C., Bhat, S. A., Wu, Y. C., H, N. F., 2022. Rice growth stage

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

classification via RF-based machine learning and image processing. Agriculture. 12(12): 2137. https://doi.org/10.3390/agriculture12122137

Singh, D., Singh, B., 2020. Investigating the impact of data normalization on classification performance. Appl Soft Comput. 97: 105524. https://doi.org/10.1016/j.asoc.2019.105524

Song, P. L., Huang, J. F., Mansaray, L. R., 2019. An improved surface soil moisture downscaling approach over cloudy areas based on geographically weighted regression. Argic. For. Meteorol. 275, 146-158. https://doi.org/10.1016/j.agrformet.2019.05.022

Song, P., Zhang, Y. Q., Guo, J. P., Shi, J. C., Zhao, T. J., Tong, B., 2022. A 1-km daily surface soil moisture dataset of enhanced coverage under all-weather conditions over China in 2003-2019. Earth. Syst. Sci. Data. Discuss. 14, 2613-2637. https://doi.org/10.5194/essd-2021-428

Song, P., Zhang, Y. Q., Tian, J., 2021. Improving surface soil moisture estimates in humid regions by an enhanced remote sensing technique. Geophys. Res. Lett. 48, 91459. https://doi.org/10.1029/2020GL091459

Stuecker, M. F., Tigchelaar, M., Kantar, M. B., 2018. Climate variability impacts on rice production in the Philippines. PLoSONE. 13(8): e0201426. https://doi.org/10.1371/journal.pone.0201426

Tian, H. R., Wang, P. X., Tansey, K., Han, D., Zhang, J. Q., Zhang, S. Y., Li, H. M., 2021. A deep learning framework under attention mechanism for wheat yield estimation using remotely sensed indices in the Guanzhong Plain, PR China. In: 2021 Int. J. Appl. Earth Obs. Geoinf. 102, 102375. https://doi.org/10.1016/j.jag.2021.102375

Tiao, G. C., Box, G. E. P., 1981. Modeling multiple time series with applications. J. Am. Stat. Assoc. 76(376): 802-816. https://doi.org/10.1080/01621459.1981.10477728

Turner, C. R., Fuggetta, A., Lavazza, L., Wolf, A. L., 1999. A conceptual basis for feature engineering. J. Syst.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

Software. 49(1): 3-15. https://doi.org/10.1016/S0164-1212(99)00062-X

Wakatsuk, H., Ju, H., Nelson, G. C., Farrell, A. D., Deryng, D., Meza, F., Hasegawa, T., 2023. Research trends and gaps in climate change impacts and adaptation potentials in major crops. Curr. Opin. Env. Sust. 60, 101249. https://doi.org/10.1016/j.cosust.2022.101249

Wang, J. W., Chen, J. S., Zhang, J. H., Yang, S. S., Zhang, S., Bai, Y., Xu, R. Z., 2023. Consistency and uncertainty of remote sensing-based approaches for regional yield gap estimation: A comprehensive assessment of process-based and data-driven models. Field. Crop. Res. 302, 109088. https://doi.org/10.1016/j.fcr.2023.109088

Wang, P., Tam, C. Y., Xu, K., 2019. El Niño–East Asian monsoon teleconnection and its diversity in CMIP5 models. Climate Dynamics. 53: 6417-6435. https://doi.org/10.1007/s00382-019-04938-3

Wang, X. L., Huang, J. X., Feng, Q. L., Yin, D. Q., 2020. Winter wheat yield prediction at county level and uncertainty analysis in main wheat-producing regions of China with deep learning approaches. Remote Sens. 12(11), 1744 https://doi.org/10.3390/rs12111744

Wei, W., Pang, S. F., Wang, X. F., Zhou, L., Xie, B. B., Zhou, J. J., Li, C. H., 2020. Temperature Vegetation Precipitation Dryness Index (TVPDI)-based dryness-wetness monitoring in China. Remote Sens. Environ. 248, 111957. https://doi.org/10.1016/j.rse.2020.111957

Wei, W., Zhang, H. Y., Zhou, J. J., Zhou, L., Xie, B. B., Li, C. H., 2021. Drought monitoring in arid and semi-arid region based on multi-satellite datasets in northwest, China. Environ. Sci. Pollut. R. 28: 51556-51574. https://doi.org/10.1007/s11356-021-14122-y

Wibawa, A. P., Utama, A. B. P., Elmunsyah, H., Pujianto, U., Dwiyanto, F. A., Hernandez, L., 2022. Time-series analysis with smoothed Convolutional Neural Network. Journal of big Data. 9(1): 44.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

https://doi.org/10.1186/s40537-022-00599-y

Xiao, X., S. Boles., S. Frolking., W. Salas., B. Moore., C. Li., L. He., R. Zhao., 2002. Observation of flooding and rice transplanting of paddy rice fields at the site to landscape scales in China using VEGETATION sensor data. Int. J. Remote. Sens. 23, 3009-3022. https://doi.org/10.1080/01431160110107734

Xiong, X. G., Yang, J., Zhong, R. H., Dong, J. W., Huang, J. H., Ting, K. C., Ying, Y. B., Lin, T., 2024. Integration of harvester trajectory and satellite imagery for large-scale winter wheat mapping using deep positive and unlabeled learning. Comput Electron Agric. 216: 108487. https://doi.org/10.1016/j.compag.2023.108487

Yang, K., He, J., Tang, W. J., Qin, J., Cheng, C. C. K., 2010. On downward shortwave and longwave radiations over high altitude regions: Observation and modeling in the Tibetan Plateau. Agr. Forest. Meteorol. 150, 38-46. https://doi.org/10.1016/j.agrformet.2009.08.004

Yang, L., Shami, A., 2020. On hyperparameter optimization of machine learning algorithms: Theory and practice. Neurocomputing. 415: 295-316. https://doi.org/10.1016/j.neucom.2020.07.061

Yin, Y. X., Han, C., Yang G. Y., Huang, Y. H., Liu, M. Y., Wang, X. J., 2020. Changes in the summer extreme precipitation in the Jianghuai plum rain area and their relationship with the intensity anomalies of the south Asian high. ATMOS RES. 236: 104793. https://doi.org/10.1016/j.atmosres.2019.104793

You, J. X., Li, X. C., Low, M., Lobell, D., Ermon, S., 2017. Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data. Proc. AAAI Conf. Artif. Intell. 4559-4566. https://doi.org/10.1609/aaai.v31i1.11172

Zhang, A. Z., Jia, G. S., 2013. Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data. Remote Sens. Environ. 134, 12-23.

https://doi.org/10.5194/egusphere-2025-3901
Preprint. Discussion started: 15 August 2025

https://doi.org/10.1016/j.rse.2013.02.023

Zhang, L. P., Zhang, L. F., Du, B., 2016a. Deep learning for remote sensing data: a technical tutorial on the state of the art. IEEE. Geosc. Rem. Sen. M. 4, 22-40. https://doi.org/10.1109/MGRS.2016.2540798

Zhang, Q., Yu, H. Q., Sun, P., Singh, V. P., Shi, P. J., 2019. Multisource data based agricultural drought monitoring and agricultural loss in China. Glob. Planet. Change. 172, 298-306. https://doi.org/10.1016/j.gloplacha.2018.10.017

Zhang, R., Lee, H., Radev, D., 2016. Dependency sensitive convolutional neural networks for modeling sentences and documents. arxiv preprint arxiv:1611.02361.

Zhang, X., Zhang, Q., 2016b. Monitoring interannual variation in global crop yield using long-term AVHRR and MODIS observations. ISPRS J Photogramm Remote Sens. 114: 191-205. https://doi.org/10.1016/j.isprsjprs.2016.02.010.

Zhao, M., Heinsch, F. A., Nemani, R. R., Running, S. W., 2005. Improvements of the MODIS terrestrial gross and net primary production global data set. Remote Sens. Environ. 95, 164-176. https://doi.org/10.1016/j.rse.2004.12.011

Zhuo, W., Fang, S. B., Gao, X. R., Lei, W., Wu, D., Fu, S. L., Wu, Q. L., Huang, J. X., 2022. Crop yield prediction using MODIS LAI, TIGGE weather forecasts and WOFOST model: A case study for winter wheat in Hebei, China during 2009-2013. Int. J. Appl. Earth. Obs. Geoinf. 106, 102668. https://doi.org/10.1016/j.jag.2021.102668