

## Article

# Winter Wheat Yield Estimation by Fusing CNN–MALSTM Deep Learning with Remote Sensing Indices

Changchun Li <sup>1</sup>, Lei Zhang <sup>1</sup>, Xifang Wu <sup>1,\*</sup>, Huabin Chai <sup>1</sup>, Hengmao Xiang <sup>2,\*</sup> and Yinghua Jiao <sup>2</sup>

<sup>1</sup> School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China; lcc@hpu.edu.cn (C.L.); 212304020070@hpu.edu.cn (L.Z.); chaih@hpu.edu.cn (H.C.)

<sup>2</sup> Shandong Provincial Institute of Land Surveying and Mapping, Jinan 250102, China; jiaoyinghua@shandong.cn

\* Correspondence: wxf2020@hpu.edu.cn (X.W.); xianghengmao@shandong.cn (H.X.)

**Abstract:** A rapid and accurate determination of large-scale winter wheat yield is significant for food security and policy formulation. In this study, meteorological data and enhanced vegetation index (EVI) were used to estimate the winter wheat yield in Henan Province, China, by constructing a deep learning model. The deep learning model combines CNN feature extraction and makes full use of the sequence data processing capability of the LSTM and a multi-head attention mechanism to develop a novel CNN–MALSTM estimation model, which can capture the information of input sequences in different feature subspaces to enhance the expressiveness of the model. A CNN–LSTM baseline model was also constructed for comparison. Compared with the baseline model ( $R^2 = 0.75$ , RMSE = 646.53 kg/ha, and MAPE = 8.82%), the proposed CNN–MALSTM model ( $R^2 = 0.79$ , RMSE = 576.01 kg/ha, MAPE = 7.29%) could more accurately estimate the yield. Based on the cross-validation with one year of left-out data and the input of the fertility period by fertility period to explore the sensitivity of the model to data from different fertility periods to the final yield, an annual yield distribution map of Henan Province was constructed. Through cross-validation, the stability of the model in different years was assessed. The results showed that the model could obtain the best prediction of the yield approximately 20 days in advance. In terms of the spatial distribution of the yield in Henan Province on a yearly basis, the estimated yield showed an overall uptrend from west to east, consistent with the trend in the statistical yearbook of the yield for Henan Province. Thus, it can be concluded that the proposed CNN–MALSTM model can provide stable yield estimation results.

**Keywords:** enhanced vegetation index (EVI); meteorological data; long short-term memory networks with multiple attention mechanism (MALSTM); yield estimation; winter wheat



**Citation:** Li, C.; Zhang, L.; Wu, X.; Chai, H.; Xiang, H.; Jiao, Y. Winter Wheat Yield Estimation by Fusing CNN–MALSTM Deep Learning with Remote Sensing Indices. *Agriculture* **2024**, *14*, 1961. <https://doi.org/10.3390/agriculture14111961>

Academic Editor: Francesco Marinello

Received: 27 September 2024

Revised: 24 October 2024

Accepted: 30 October 2024

Published: 1 November 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

As one of the important cereal crops, wheat is a vital food source for human beings. With global climate change, the frequent occurrence of various extreme climate events, and the growing global population, food security has become a major concern [1]. In this context, an accurate estimation of wheat production plays a significant role in the stable development of agricultural production, ensuring the effective supply of important agricultural products, improving the agro-ecological environment, and promoting food security [2].

Conventional crop yield estimation mainly relies on human labor, and it is difficult to perform on a large scale. In recent years, multi-source data have played an increasingly important role in crop yield estimation [3–5]. Crop yield estimation methods can be mainly classified into crop growth model estimation based on the crop growth process and empirical model estimation based on conventional statistics. Crop growth models (e.g., DSSAT, WOFOST, ASPIM) can simulate the growth of crops at different time steps,

as well as the process of yield formation under extreme climatic conditions, with good physical interpretation [3,6–9]. However, due to the excessive number of input parameters (e.g., varieties, field management data, soil properties), the difficulty in obtaining these parameters, and the strong influence of spatial heterogeneity, the model applicability is low, the simulation is time-consuming, and it is difficult to realize an accurate yield estimation over a large-scale region [10,11].

Empirical statistical models (conventional statistical models and machine learning models) have shown advantages in large-scale crop yield estimation compared with crop growth models based on crop growth processes. In recent years, conventional empirical statistical models have produced better yield estimation results [12]. For example, Ren [13] et al. constructed a linear relationship between the Normalized Difference Vegetation Index (NDVI) and yield; this relationship effectively helped estimate the winter wheat yields in the Yellow–Huaihua Plain of China, with an estimation accuracy of  $R^2 \geq 0.86$  for different prefecture-level cities. However, since conventional empirical statistical models rely only on the linear relationship between vegetation indices (VIs) and target yields, it is difficult to dig deeper into the relationship between data and yields when the linear relationship is not satisfied. Machine learning algorithms [14,15], such as random forest (RF), support vector machine (SVM), gaussian process regression (GPR), partial least squares (PLS), etc., can be used to construct nonlinear relationships between environmental factors and yield. Various optimization algorithms can be used to optimize the model and identify features from data that play a key role in yield estimation [16,17]. For example, Son [18] et al. used the Moderate Resolution Imaging Spectroradiometer (MODIS) NDVI series data to develop a machine learning model based on the RF to estimate rice yields in both seasons in Taiwan, China; the coefficient of determination  $R^2$  was greater than 0.84.

With the continuous development of satellite technology, large amounts of remote sensing data have been applied in various fields, and multi-source remote sensing data are widely used in crop yield estimation [19,20]. Currently, most studies have employed the MODIS image data and Landsat to obtain relevant VIs, such as the NDVI, EVI, leaf area index (LAI), etc., to construct a relationship between the index and yield [21–23]. Huang [10] et al. predicted the yield in Hebei Province, China, by assimilating the LAI obtained from Landsat TM and MODIS data into a crop growth model with an RMSE = 151.29 kg/ha. Douglas K. Bolton [24] developed empirical models for predicting corn and soybean yields in central U.S. The use of the MODIS two-band enhanced vegetation index (EVI2) for corn yield estimation, compared with the widely used NDVI, resulted in better estimates. Due to the limitations of MODIS and the spatiotemporal resolution of Landsat, although better yield information can be obtained, there are some drawbacks in interpreting the complex relationship between crop yields over time. Therefore, an increasing number of studies have adopted Sentinel-2 image data and calculated the correlation indices for yield estimation, achieving better results for both conventional linear and nonlinear models.

The booming development of computer technology has provided powerful arithmetic support. Several algorithmic models have emerged, and the deep learning technique has been widely used. Deep learning, which is an advanced machine learning algorithm, is more advantageous than conventional linear models in dealing with nonlinear relationships by continuously learning layer by layer from low-level features to high-level features, and it has been widely used in various fields [25–28]. As a result, the application of deep learning models has gradually gained attention, especially combining the feature extraction capability of convolutional neural networks (CNNs) [29,30] with the advantages of long short-term memory networks (LSTMs) in processing time series data [31]. This algorithmic fusion not only improves the prediction performance of the model but also identifies the key factors affecting yield during crop growth more effectively. By using CNN to extract spatial features, the model is able to analyze important information in satellite images in greater depth, while LSTM is able to deal with dependencies in time-series data, thus capturing dynamic changes at different growth stages. In addition, the introduction of the multi-attention mechanism enables the model to focus on multiple relevant features,

thus enhancing the identification and weighting of important variables. This integrated approach effectively enhances the accuracy and reliability of yield estimation and can provide farmers and policymakers with more accurate forecasting information.

In the yield estimation process for crops, accumulating and learning from historical information is crucial. However, since 1D-CNNs lack a recurrent mechanism, integrating historical data into the learning process can be challenging. Xiao [32] et al. constructed a novel deep learning framework, ACNN, with Sentinel-2 raw bands as input data and effectively identified the key spectral bands by placing a weighted calculation of the extracted features after a number of convolutional layers. By placing attention channels and weighting the extracted features after several convolutional layers, the key spectral bands were effectively identified, overcoming the problems of high-yield underestimation and low-yield overestimation; the  $R^2$  and RMSE values of the model were 0.44 and 1544 kg/ha, respectively. Clearly, the CNN has strong feature extraction capability and the advantage of the attention mechanism in determining the contribution of features. Tian [23] constructed a deep learning framework based on the attention mechanism (ALSTM). The LSTM network has the advantage of processing temporal data, and the incorporation of the attention mechanism improves the internal interpretability of the model. It can also identify the importance of input variables in determining the yield at different fertility periods. Wang [33] et al. combined a CNN with a GRU network to estimate the wheat yield in the Guanzhong Plain of Shaanxi; in terms of prediction accuracy, the  $R^2$  and RMSE values of the model were 0.64 and 462.56 kg/ha, respectively. This demonstrates the advantages of the CNN combined with RNN for yield estimation. Sun [34] introduced an innovative multi-level CNN-LSTM model designed to extract spatiotemporal features for predicting maize yield, showcasing the benefits of integrating CNN with RNN. Nonetheless, research focusing on the use of CNN in combination with LSTM and attention mechanisms for crop yield estimation remains limited.

Hence, in this study, a CNN-ALSTM deep learning model was developed, and its performance and effectiveness in estimating wheat yield across different years and fertility periods were assessed. A CNN-MALSTM yield estimation framework was constructed by combining remote sensing and meteorological data to estimate winter wheat yield in the counties of Henan Province. The advantage of the LSTM in processing time-series data was utilized, combined with the powerful spatial feature extraction capability of the CNN and the identification capability of different input features contributing to the yield by the multi-attention mechanism, in order to improve the interpretability of the model between the input data and the target yield. The EVI, temperature, and precipitation were used as input data for the yield estimation, and the reason for choosing these parameters as the model input features is that the EVI often responds to crop growth, while temperature and precipitation as meteorological data are crucial in the crop growth process. The main objectives of this study were as follows:

- (1) Investigating the differences between the CNN-MALSTM model and the baseline model for county crop yield estimation in terms of their accuracy and performance.
- (2) Investigating the stability of the CNN-MALSTM model in yield estimation by fertility period for each year and the spatial and temporal distribution of production estimates in counties in different years.
- (3) Determining the optimal fertility period for yield estimation.

## 2. Material and Methods

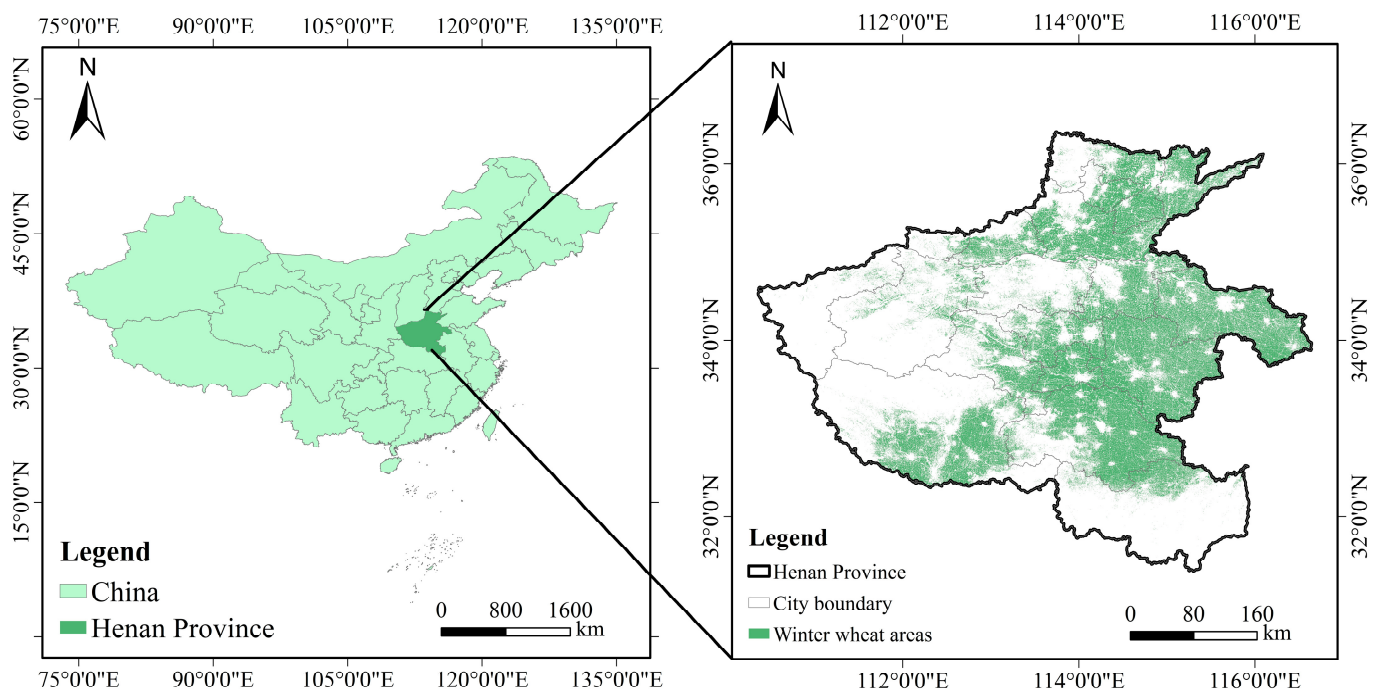
### 2.1. Study Area

The study area selected for this work was the main wheat-growing area in Henan Province, China. The area ranges from 110°21' to 116°39' E and 31°23' to 36°22' N, with Shandong and Anhui to the east, Shaanxi to the west, Hebei and Shanxi to the north, and Hubei to the south. The province covers a total area of 167,000 km<sup>2</sup>. The terrain is high in the west and low in the east, with most of it located in the warm temperate zone, which experiences four distinct seasons and simultaneous rain and heat. The average

annual temperature of the province from south to north ranges from 10.5 to 16.7 °C, and the average annual precipitation ranges from 407.7 to 1295.8 mm. Rainfall peaks during the June–August period, and the province receives an average annual sunshine of 1285.7–2292.9 h. October through June is the main fertility period of winter wheat (Table 1). This study was based on eight fertility periods for modeling and estimating yields (Figure 1).

**Table 1.** Winter wheat fertility period.

Month	Fertility Period
mid-October to mid-November	seedling stage
late November to mid-December	tillering stage
late December to late February	overwintering stage
early March to late March	greening stage
early April to mid-April	jointing stage
late April to early May	heading stage
mid-May to late May	anthesis stage
early June to mid-June	maturity stage



**Figure 1.** Study area.

## 2.2. Data Acquisition and Pre-Processing

### 2.2.1. Satellite Data

Studies have shown that the EVI (Enhanced Vegetation Index) provides more information about crop growth and plays an important role in yield estimation [35]. The Sentinel-2 satellite, operated by the European Space Agency (ESA), is an Earth observation project that provides high spatial resolution and multispectral observation data globally. The high spatial and temporal resolution of the Sentinel-2 satellite facilitates the calculation of EVI. We collected the Sentinel-2 (5 d, 10 m) Level 2A image dataset of winter wheat fertility in Henan Province to calculate EVI and used the B2, B4, and B8 bands to calculate EVI as shown in Equation (1).

$$EVI = 2.5 \frac{NIR - R}{NIR + 6R - 7.5B + 1} \quad (1)$$



Here, the *NIR* denotes the near-infrared band (B8), *R* denotes the infrared band (B4), and *B* denotes the blue band (B2).

Due to the effects of clouds, rain, etc., which can cause discontinuity in EVI calculation, Level 2A images (5 d, 10 m) with a cloud cover of less than 40% were screened. The median-value synthesis method was used to generate image data within each fertility period according to the division of the fertility period (Table 1). According to Equation (1), the EVI values were calculated, and the EVI time-series values were reconstructed using the Savitzky–Golay (SG) filtering [36], along with the winter wheat distribution maps for masking. Finally, the mean EVI values for each county were calculated on a fertility-by-fertility basis. All the above calculations were performed on the Google Earth Engine (GEE, Google, Mountain View, CA, USA) platform.

### 2.2.2. Meteorological Data

Winter wheat is often affected by meteorological factors in different fertility stages, and the formation of the final yield is inextricably linked to changes in climate. Studies have shown that accurate access to meteorological data in all fertility stages of the winter wheat can help improve the accuracy of yield estimation.

In this study, the ERA5 dataset (ERA5, European Centre for Medium-Range Weather Forecasts, Reading, UK), which is the fifth generation of atmospheric reanalysis of the global climate provided by the European Centre for Medium-Range Weather Forecasts (ECMWF), was used. Temperature and precipitation data were selected based on meteorological driver data commonly used in crop growth models to simulate winter wheat growth. The meteorological data used in this study were obtained from the GEE platform. Hour-by-hour data with a spatial resolution of  $0.1^\circ \times 0.1^\circ$  were selected according to the needs of this study, using 24-h averages as the precipitation and temperature for the day. These averages were then computed to each fertility stage using the mean method based on the fertility stages of winter wheat (Table 1). The same approach was used for the mask aggregation to the county level using the distribution map of winter wheat.

### 2.2.3. Winter Wheat Mask Data and Yield Data

To accurately obtain the distribution data of winter wheat cultivation in Henan Province, a winter wheat extraction model based on spectral features, vegetation features, texture features, and terrain features was constructed in past research. This model can provide high-precision extraction of winter wheat planting areas (OA = 92.7%, Kappa = 0.902) [37]. Using GEE as the data processing platform, the NDVI values of the features in the study area were calculated, the climatic periods that were easy to distinguish between winter wheat and other features were selected, a synthesis window for the remote sensing images was determined, and the Sentinel-2 images were synthesized to finally obtain the pre-wintering and post-wintering images. The RF algorithm was used to identify the preferred feature images and produce the distribution map of the winter wheat for Henan Province.

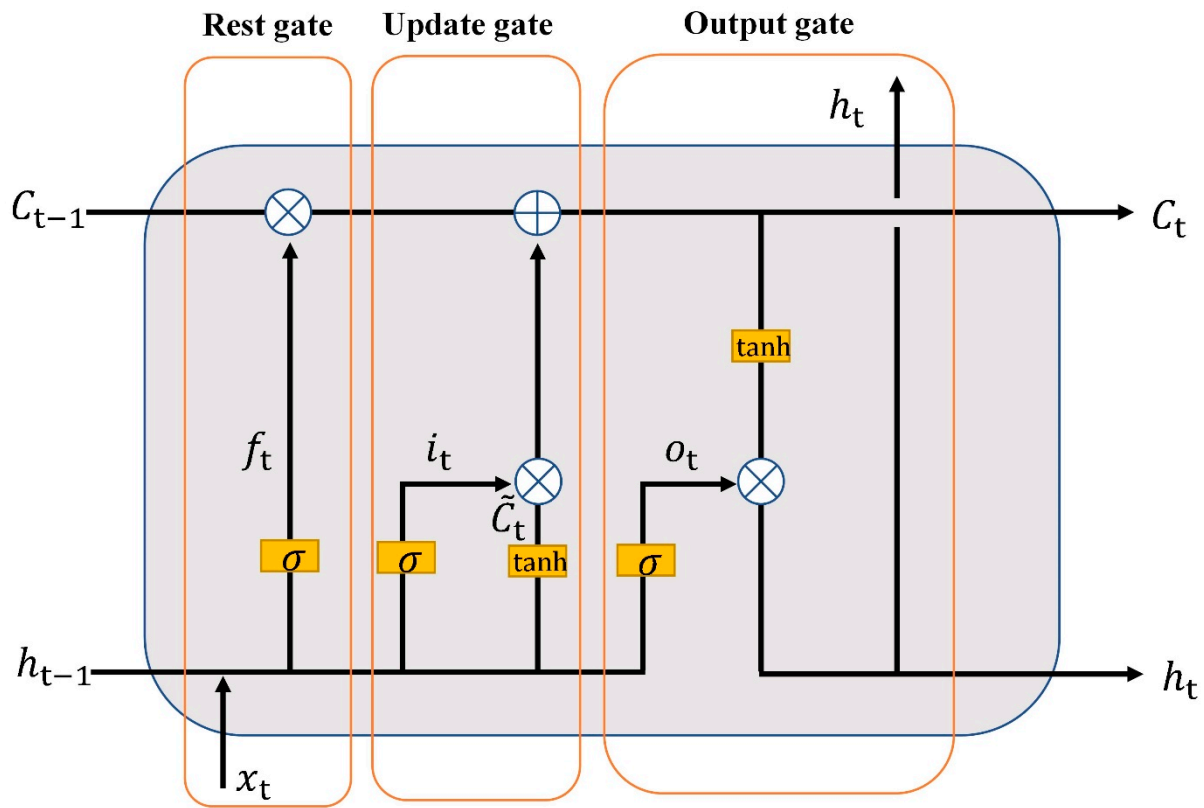
According to the Henan Provincial Statistical Yearbook (<https://www.stats.gov.cn/>, accessed on 31 May 2024), the change in the winter wheat planting area in Henan Province over the past five years was less than 1%. Therefore, the winter wheat distribution in 2019 was used instead of the winter wheat distribution in the 2019–2023 period. The sown area and yield data for the winter wheat in the 2019–2023 period were obtained from the Henan Provincial Statistical Yearbook. Based on the yield data, 101 major wheat-growing counties with yields in the range of 3000–9000 kg/ha were selected for modeling and analysis.

## 2.3. Methods

### 2.3.1. LSTM Neural Network Models and Multi-Head Attention Mechanisms

The LSTM is a special type of RNN structure that typically processes and learns from time-series data through three key gating units: an input gate, a forget gate, and an output gate. These gating units control the flow of information during learning, which facilitates

the LSTM network to better deal with long-term dependencies. Figure 2 shows the internal structure of the LSTM.



$\sigma$  Sigmoid function

$\tanh$  Hyperbolic tangent function

$\otimes$  Matrix elements multiplication

$\oplus$  Addition operation

**Figure 2.** Diagram of the internal parameters of the LSTM model.

For a given time step  $t$  and input  $X_t$ , each LSTM unit is defined as follows (Figure 2):

- (1) Forget gate:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (2)$$

- (2) Input gate:

$$i_t = \sigma(W_i [h_{t-1}, x_t] + b_i) \quad (3)$$

- (3) Updating unit status:

$$\tilde{C}_t = \tanh(W_C [h_{t-1}, x_t] + b_C) \quad (4)$$

- (4) Unit state:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

- (5) Output gate:

$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o) \quad (6)$$

- (6) Hide status updates:

$$h_t = o_t * \tanh(C_t) \quad (7)$$

In this context,  $\sigma$  represents the sigmoid function.

The multi-head attention (MHA) mechanism splits the input query, key, and value matrices into multiple heads and computes the attention independently in each head. It then combines the outputs of these heads into a linear transformation to achieve the simultaneous capture and integration of multiple interactions in different representation subspaces, enhancing the expressive power of the model. The output of the LSTM,  $h_i$ , is used as the input to the MHA mechanism, which is calculated as follows:

For each input  $h_i$ , a linear transformation is performed using different weight matrices  $W_q$ ,  $W_k$ , and  $W_v$  to generate the query,  $Q$ , key  $K$ , and value  $V$ .

$$Q_i = h_i W_q, \quad V_i = h_i W_v, \quad K_i = h_i W_k \quad (8)$$

For each query  $Q_i$  and all keys  $K_i$ , the attention score  $e_{ij}$  is computed as follows:

$$e_{ij} = \frac{Q_i K_j}{\sqrt{d_k}} \quad (9)$$

where  $d_k$  is the dimension of the key vector, which is scaled to avoid excessive inner product results as the dimension increases.

The attention score  $e_{ij}$ , was subjected to softmax operation to obtain the attention weight  $a_{ij}$ .

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_k \exp(e_{ik})} \quad (10)$$

For each value  $V_j$ , a weighted summation is performed using the attention weights  $a_{ij}$  to obtain the final attention output.

$$z_i = \sum_j a_{ij} V_j \quad (11)$$

The MHA mechanism parallelizes the above computation process  $h$  times (i.e., the number of multi-heads), using a different weight matrix each time. The output of each head is as follows:

$$z_i^{(1)} = \sum_j a_{ij}^{(1)} V_j^{(1)} \quad (12)$$

$$z_i^{(2)} = \sum_j a_{ij}^{(2)} V_j^{(2)} \quad (13)$$

The final output is obtained by a linear layer transformation:

$$Output_i = Z_i W_o \quad (14)$$

where  $W_o$  is the output projection matrix of the MHA mechanism.

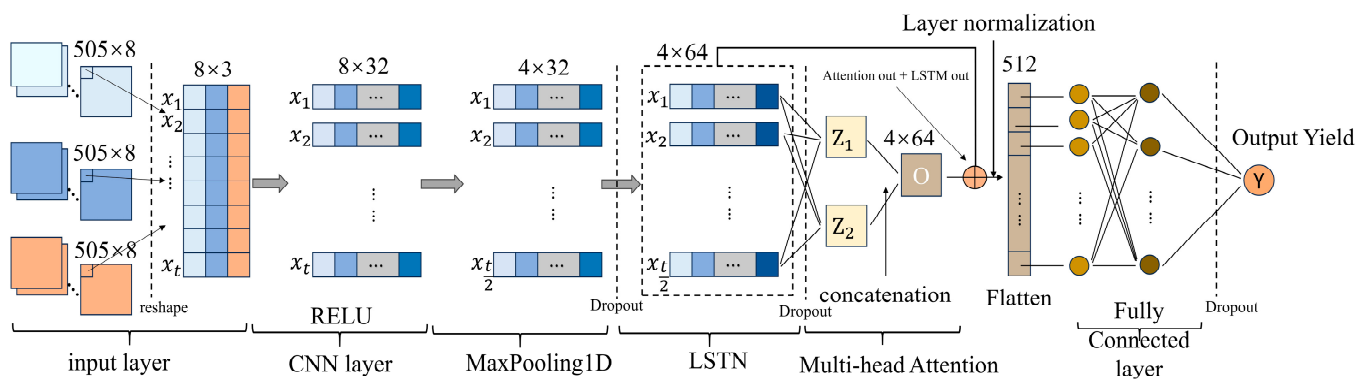
The output processed by the multi-attention mechanism passes through the residual connectivity and normalization layers to ensure stability and information flow through the network. It is subsequently flattened to fit subsequent fully connected layers for further prediction of the yield.

### 2.3.2. CNN–MALSTM Model

The CNN has made significant achievements in the field of object detection and image recognition with its powerful feature extraction capability. Traditionally, CNNs have made significant strides in target detection and image recognition due to their robust feature extraction capabilities. Similarly, 1D-CNNs are specifically designed for one-dimensional data, such as time series or text sequences. By applying convolution operations to this data, 1D-CNNs effectively capture local features and enhance feature learning. The MHA mechanism captures richer features and contextual information on the basis of self-attention by adding multiple attention heads to allocate attention to different dimensions of the input

information in parallel. This enables the model to capture key features in the input information from different angles and dimensions, which significantly enhances the expressive ability and learning efficiency of the model [38].

The model proposed in this study combines a convolutional neural network (1D-CNN), a long short-term memory network (LSTM), and a multi-head attentional mechanism, which is the new fusion method we propose to enhance the prediction of winter wheat yield. The model first extracts features through a 1D convolutional layer and reduces the dimensionality using a maximum pooling layer, followed by the introduction of a Dropout layer to prevent overfitting. Next, an LSTM layer processes the sequence data, and another Dropout layer is added to enhance the robustness of the model. On this basis, the multi-head attention mechanism is used to further optimize the feature extraction capability, and the attention output is summed with the LSTM output to stabilize the training process through layer normalization, thus enhancing the effectiveness of the information flow. Finally, the output is passed through the fully connected layer, which is set to univariate prediction, aiming to improve the accuracy and stability of the prediction. The specific model framework is shown in Figure 3.



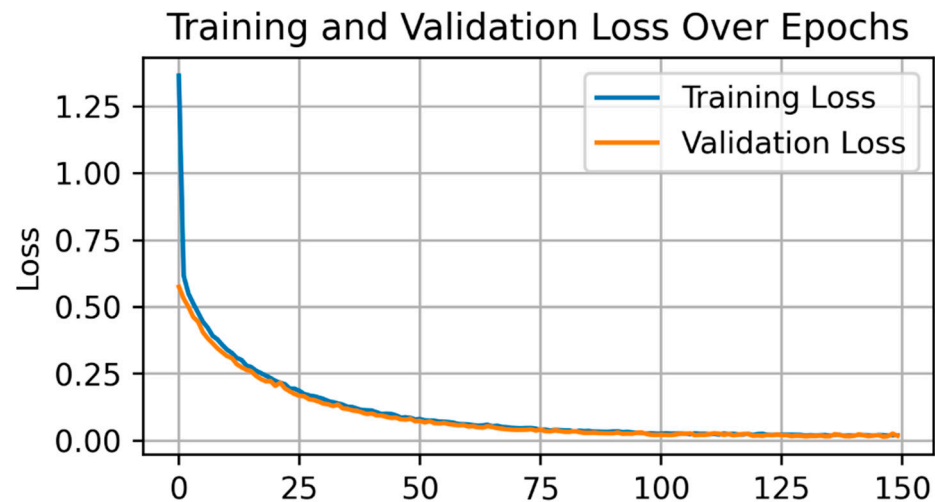
**Figure 3.** Diagram of the CNN–MALSTM fusion framework.

### 2.3.3. Hyperparametric Sensitivity Analyses

The choice of hyperparameters for the model has a significant impact on the prediction accuracy. Specifically, the following (Table 2) several parameters affect the model's performance to varying degrees. The model was trained with different combinations of parameters through grid search and the performance (mean squared error or  $R^2$  value) of the validation set was compared. This identifies which parameters have a greater impact on the accuracy of the model and adjusts the parameters to optimize model performance. After extensive experiments, hyperparameter results were obtained as shown in Table 2. The obtained hyperparameters are inputted into the model to obtain the training error plot (Figure 4). The training error plot shows the trend of the model's loss (error) during the training process, which consists of two curves, the training error and the validation error, with the increase in the number of training rounds (Epochs), the training error gradually decreases with the training and tends to stabilize, while the validation error decreases with the increase of training rounds and then also tends to stabilize. This indicates that the model is gradually learning and improving, and can be adjusted to optimize the model's performance without any obvious overfitting phenomenon.

**Table 2.** Hyperparameter ranges.

Hyperparameter	Range of Analyses	Optimal Results
kernel_size	(2, 5)	2
lstm_units	(16, 256)	64
num_heads	(1, 16)	2
key_dim	(16, 128)	64
Dropout	(0.1, 0.5)	0.1
learning_rate	(0.0001, 0.01)	0.001



**Figure 4.** CNN–MALSTM training error plot (Epochs = 150, batch size = 15).

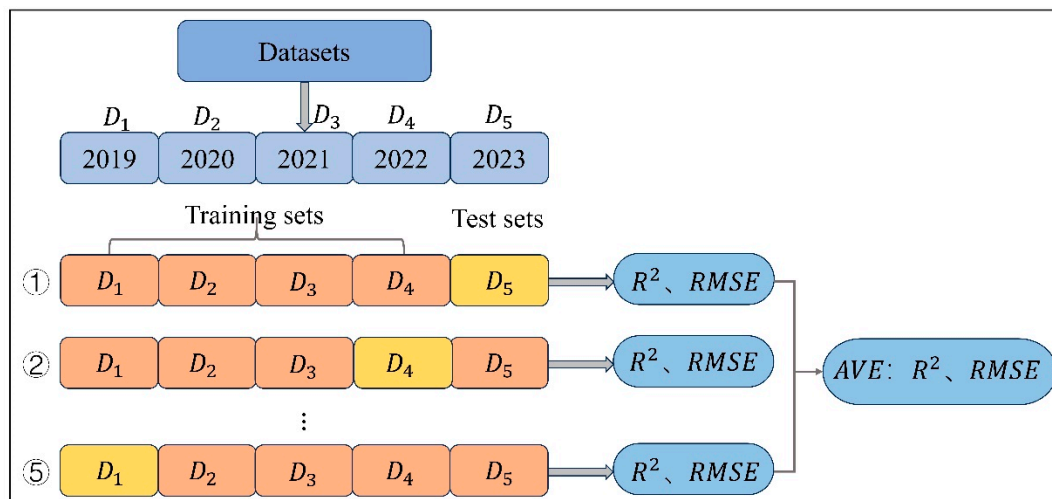
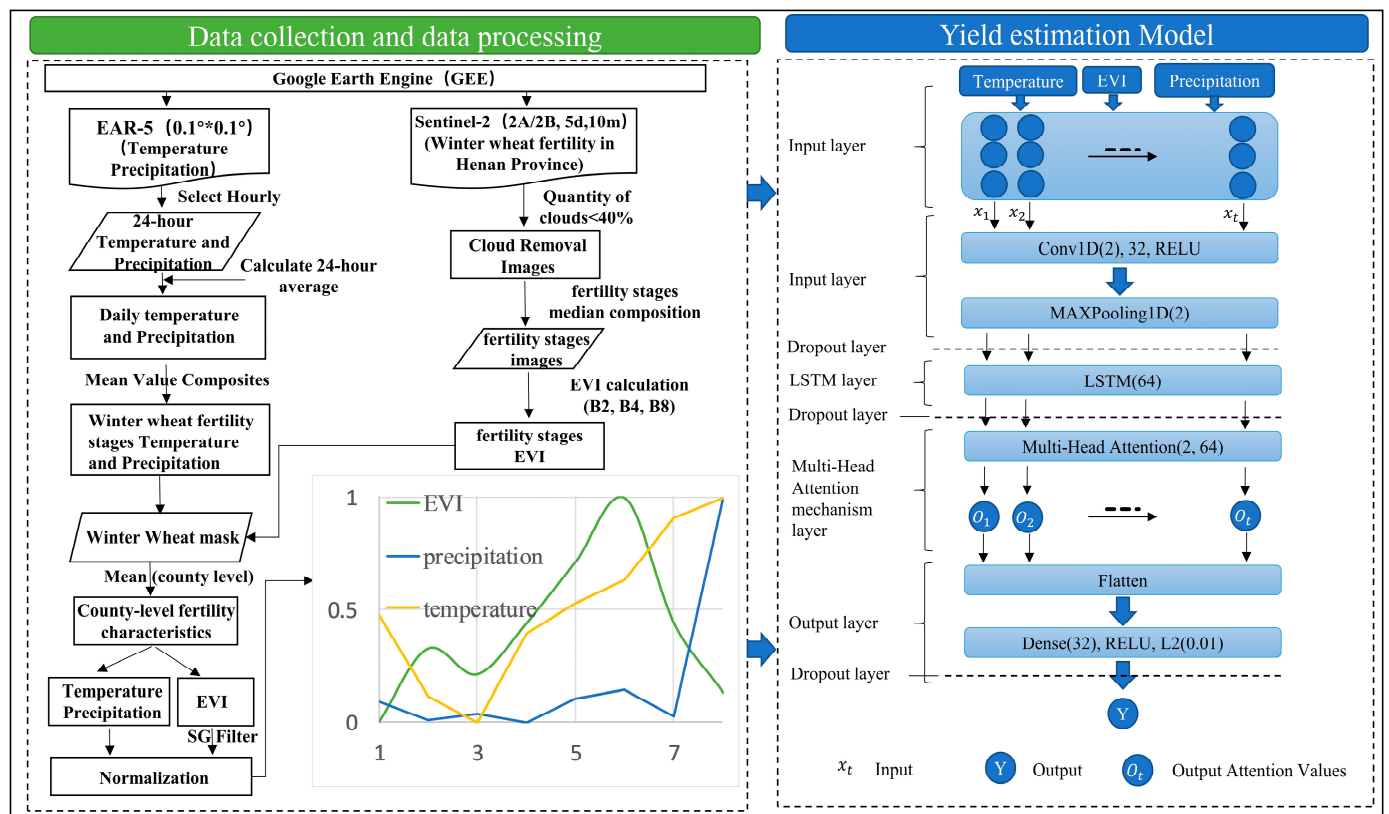
#### 2.3.4. CNN–MALSTM Estimation Technology Route

The data collection and pre-processing steps are shown in Figure 5, where the processed data are normalized to the range [0, 1] for each feature before being input into the model. This step ensures that all input variables contribute equally to the learning process of the model, thus avoiding bias towards features with a large range of values. The input data consisted of three variables (EVI, mean temperature, and mean precipitation) for each of the 101 counties for the period 2019 to 2023, reshaped into a  $505 \times 8 \times 3$  format. Here, the data were reshaped into a three-dimensional tensor, with 505 indicating a total of 505 data points, where 8 represents the time step (eight reproductive periods), and 3 corresponds to the three features (EVI, temperature, and precipitation). This allows the model to exploit temporal dependencies in the data. The input model was preceded by a random dataset, i.e., 80% (404) training and 20% (101) validation. The model was built on TensorFlow version 2.0. After conducting a grid search and sensitivity analysis, the following model parameters were determined: a single convolutional layer with a 32 filter, a kernel size of 1, and a step size of 2, an activation function ReLu, a maximal pooling layer with a pooling window of 2, a 64-unit LSTM, and lastly a 32-unit fully connected layer. The optimizer is a gradient descent method with a learning rate of 0.001. The loss function is the mean squared error, and L2 regularization is applied. The model was trained for 150 batches with a batch size of 15. Figure 5 shows the flowchart of data processing and the construction of the CNN–MALSTM model to estimate the yield.

#### 2.3.5. Assessment of Models

K-fold cross-validation is a technique commonly used in model evaluation, aiming to effectively assess the performance and generalization ability of models. In this article, five-fold cross-validation was chosen to evaluate deep learning models, and its specific steps are as follows (Figure 6): First, the original dataset is divided into five disjoint subsets according to the year (2019–2023), with each year's data in this study comprising 101 entries. Then, iterations are performed; in the five iterations, one of the subsets is sequentially selected as the test set, while the remaining four subsets (404 entries) are grouped into the training set. The model is trained on the training set, and the model's performance is evaluated on the test set. In each iteration, the model is given a performance metric, i.e.,  $R^2$  and RMSE on the test set. Averaging these metrics gives the final cross-validation performance. The performance metrics at different vintages show the stability of the model at different vintages and its generalization ability in the face of variations, helping to identify whether the model has overfitting or underfitting problems.





To compare the model performance and the advantages of the MHA mechanism, a separate CNN-LSTM was selected as a comparison model, with the input data and basic parameter settings the same as those of the CNN-MALSTM. This comparison was used to assess the advantages of the MHA mechanism in nonlinear information extraction. Data from different years and fertility-by-fertility data were input to evaluate the model's ability to explore the input data at the county scale and across different fertility stages. Specifically, model performance at the county scale and the model's ability to recognize the importance of different fertility data as they accrue were explored. To quantify the importance of different fertility data for the final yield and the model's ability for early predictions, i.e., how far in advance and how accurate the model can predict wheat yield, three metrics

were chosen: the correlation coefficient  $R^2$ , root mean squared error (RMSE), and mean absolute percentage error (MAPE). The formulas for  $R^2$  and RMSE are given below:

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{y}_i)^2} \quad (15)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (16)$$

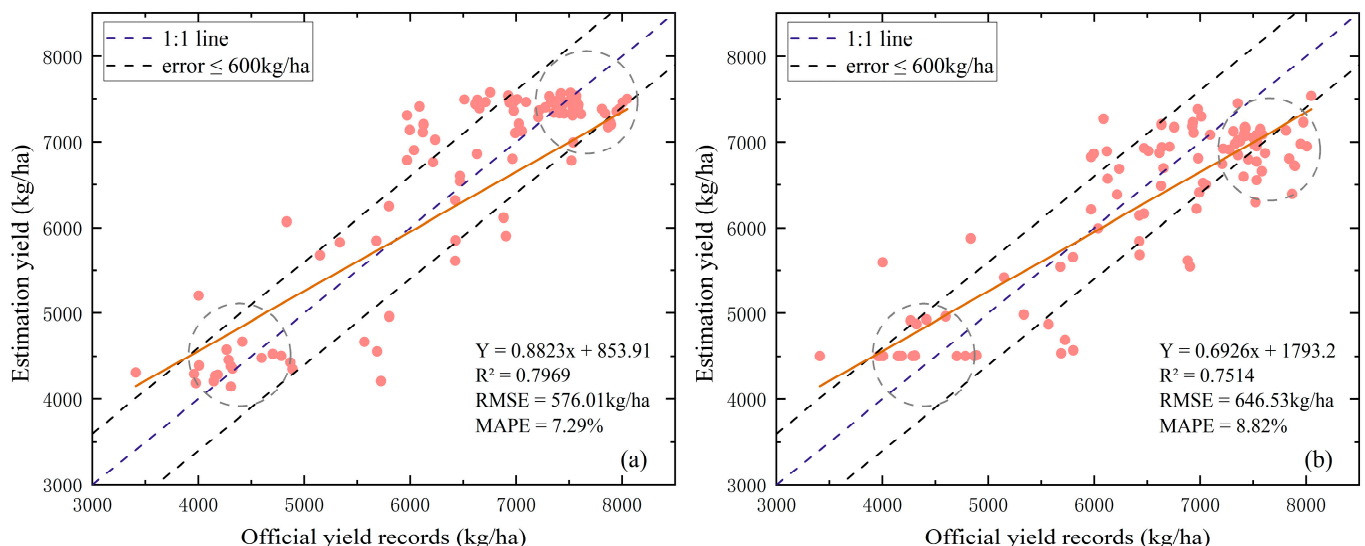
$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - x_i}{x_i} \right| 100\% \quad (17)$$

where  $x_i$  is the statistical yield of the wheat,  $y_i$  is the estimated yield of the winter wheat,  $\bar{y}_i$  is the mean statistical yield of wheat, and  $n$  is the number of samples in the test set.

### 3. Results and Analysis

#### 3.1. Yield Estimation Using CNN–MALSTM and Baseline Model

Figure 7 shows the scatter plot of the yield estimation for the test set across 101 counties of Henan Province. The results indicated that the  $R^2$ , RMSE, and MAPE of the CNN–MALSTM model and the baseline model, CNN–LSTM, were 0.79 and 0.75, 576.01 kg/ha and 646.53 kg/ha, and 7.29% and 8.82%, respectively. The overall performance of the CNN–MALSTM model was better; in comparison with the CNN–LSTM model, the  $R^2$  increased by 0.04, and the RMSE and MAPE decreased by 70.52 kg/ha and 1.53%, respectively. From the yield scatter plot, the data are distributed on both sides of the 1:1 line, exhibiting a linear relationship. The MAPE of the CNN–MALSTM model was 7.29%, which means that the model could explain more than 90% of the yield variations. This demonstrates the effectiveness of the combination of the CNN and multi-attention mechanism in extracting the relevant features from the original data and the fact that the model can pay attention to the differences in the contribution of different features.



**Figure 7.** Comparison of the performance of the two models in terms of the county-level yield statistics and estimated data, where (a) is the CNN–MALSTM model and (b) is the CNN–LSTM baseline model.

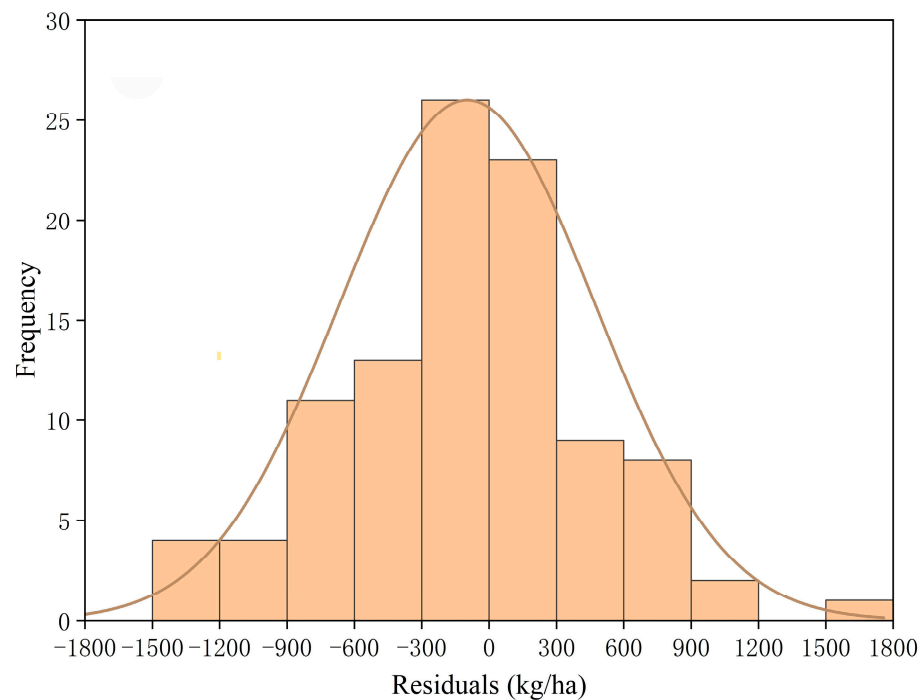
Compared with the CNN–LSTM model, the slope of the trend line for the CNN–MALSTM model was closer to 1. In the yield intervals of 3000–5000 kg/ha and 7000–8000 kg/ha, considered low-yield and high-yield, respectively, the scatter plots of the two models showed that the CNN–MALSTM model can alleviate the problems of high-yield underestimation and low-yield overestimation in the yield estimation process. This also verifies the effectiveness

of the multi-attention mechanism to solve these problems. Compared with the CNN-LSTM model, the proposed model performed well for both high- and low-yield samples; however, the overall slope was still less than 1, and there was still an underestimation of high yields and an overestimation of low yields in some areas. Based on the density of the scatterplot in different yield intervals, the yield data were mainly concentrated in the range of 5000–7500 kg/ha, compared with the low-yield samples and high-yield samples, which also explains the lower accuracy of the two models for the high and low yields from one aspect. Deep learning typically requires a large sample size, and the sample size in this study resulted in the model's inability to learn high-yield features. From the 600 kg/ha error line on the three scatter plots, combined with the evaluation metrics  $R^2$  and RMSE, the CNN-MALSTM model fell the least outside of the error line and was closer to the trend line, whereas the CNN-LSTM model suffered from significant underestimation of high yields. This further suggests that a combination of CNN and the multi-attention mechanism allows for better extraction of important features.

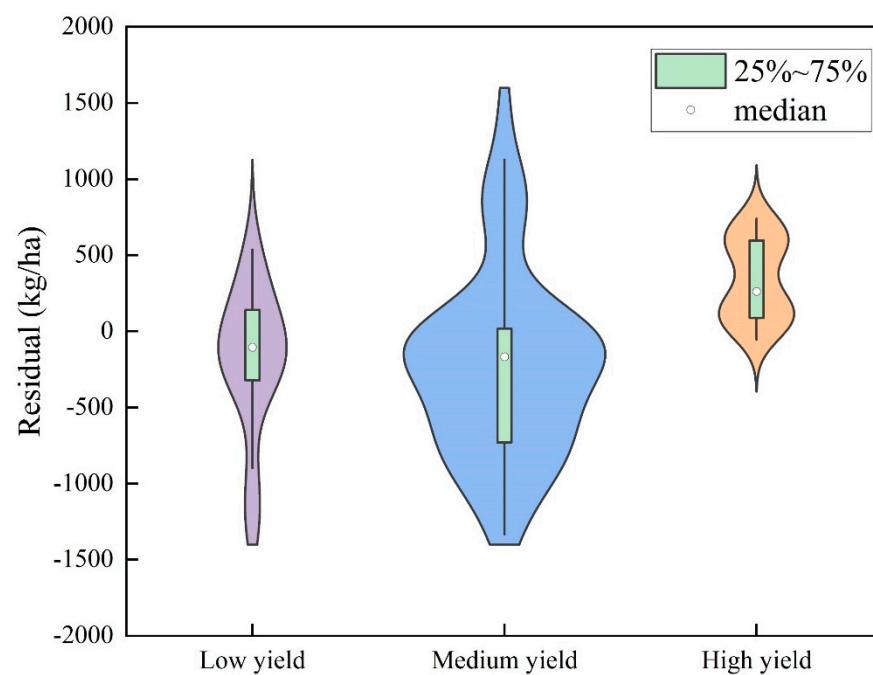
### 3.2. Comparison of CNN-MALSTM Model in Terms of County Residuals

In this study, the difference between the official statistical yields and the predicted yields was used to calculate the residuals. We took into account the specific yield characteristics of the study area and referred to the threshold-setting method for wheat yield estimation from similar areas [39]. We used the absolute value threshold method to set a fixed absolute value threshold of 600 kg/ha. Specifically, residual values greater than 600 kg/ha were considered underestimates, while those less than 600 kg/ha were considered overestimates. Figure 8 shows the histograms of the residuals in counties with the CNN-MALSTM model. Overall, the yield errors in most areas were within 600 kg/hm<sup>2</sup>, and the overall residuals followed a normal distribution and were uniformly distributed on both sides of zero. Based on the residual box plots, high yields showed different degrees of underestimation, while the medium- and low-yield residuals were both positive and negative. The average residuals for the high yield were 337.50 kg/ha, which was much higher than the residuals for the medium and low yields of −230.29 kg/ha and −183.05 kg/ha. These results indicated that the model performed well in medium- and low-yielding areas. From the size of the residual violin plot shown in Figure 9, it can be seen that the number of medium-yield samples is high, consistent with the distribution of the collected yield data. During model training, the use of fewer high-yield data resulted in the model performing poorly in estimating high yields and better in estimating low and medium yields, with the CNN-MALSTM model performing the best. Overall, the model can learn the relationship between different features to better relate to the final yield.

Box plots are an effective statistical chart used to demonstrate the variability and distributional characteristics of data. They help researchers quickly understand the distribution of data by displaying key statistical indicators of the dataset, such as minimum, first quartile (Q1), median (Q2), third quartile (Q3), and maximum values. Thus, box plots can help in analyzing the variations of different ranges of yields in winter wheat yield prediction results. Further analysis of the cause of the residuals through box plots reveals that the unevenness of the data distribution may be a key factor contributing to this phenomenon. The scarcity of high-yield samples makes the model's learning ability in these areas insufficient to capture the key factors affecting high yields, such as climate change and corresponding remote sensing data. In addition, feature selection and model complexity also affect the prediction results. In order to improve the performance of the model in high-yield areas, future research could consider introducing more high-yield data or applying data enhancement techniques to improve the accuracy of the model in high-yield estimation.



**Figure 8.** Residual distribution.



**Figure 9.** Residual violin plots for different yield ranges.

### 3.3. Model Robustness Assessment

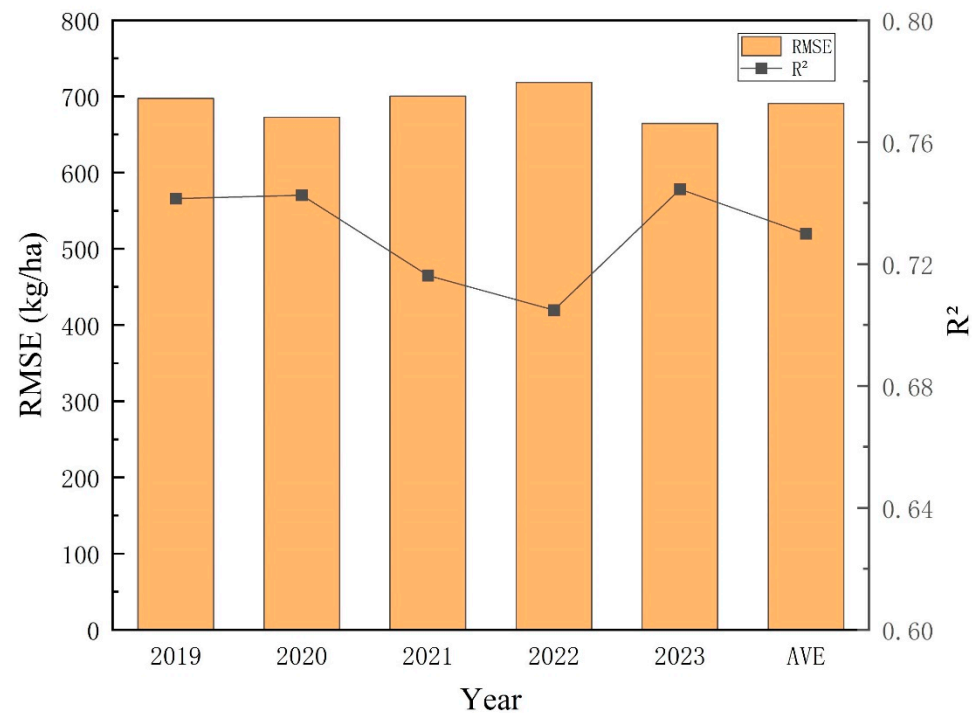
To further assess the robustness and generalizability of the model, cross-validation was used for model training, i.e., one year's data were used for model validation, while the data from other years were used for model training. The model performance was then assessed for each year until all the year's data had been used as the test set in turn.

Figure 10 shows the results of the cross-validation evaluation of the model for different years, with  $R^2$  ranging from a minimum of 0.70 to a maximum of 0.74 and the corresponding RMSE ranging from 718.40 kg/ha to 697.58 kg/ha. Overall, the estimation accuracy of the model was more stable across all years, with an average  $R^2$  of 0.73 and an RMSE of 690.82 kg/ha. The interannual differences in the estimation results of the different years

were relatively small, indicative of the robustness of the CNN–ALSTM model (Table 3 and Figure 10).

**Table 3.** “One-year-out” cross-validation.

Maturity	2019	2020	2021	2022	2023	Average
RMSE (kg/ha)	697.58	672.85	700.44	718.40	664.82	690.82
R <sup>2</sup>	0.74	0.74	0.72	0.70	0.74	0.73



**Figure 10.** RMSE and R<sup>2</sup> trends for “one-year sabbatical” cross-validation for the 2019–2023 period.

### 3.4. Fertility-by-Fertility CNN–MALSTM Model Performance Analysis

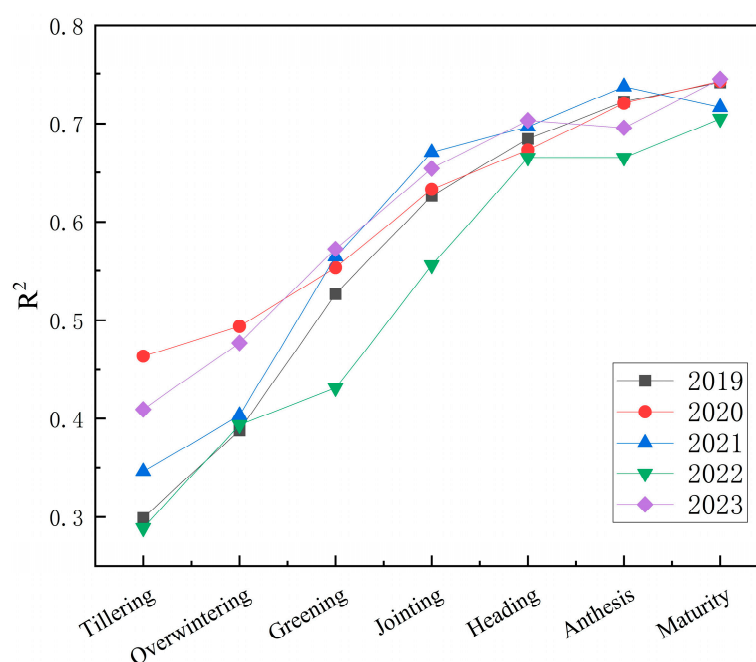
To explore the cumulative effect of input feature time on the final yield estimation, a stepwise sensitivity analysis was used to quantify the changes in the model’s yield estimation performance with the increase in the feature data. The evaluation metrics used R<sup>2</sup> and RMSE to compare and analyze the accuracy of the yield estimation for 2019–2023 in each county by the fertility period. Figures 11 and 12 show the results.

Before the overwintering period, winter wheat grows slowly, soil background information has a large impact, and it is difficult to obtain yield-related information from satellite images. On the other hand, meteorological data inputs are small, and it is difficult to distinguish the impact of environmental factors on winter wheat, resulting in the model being inaccurate in the early stage.

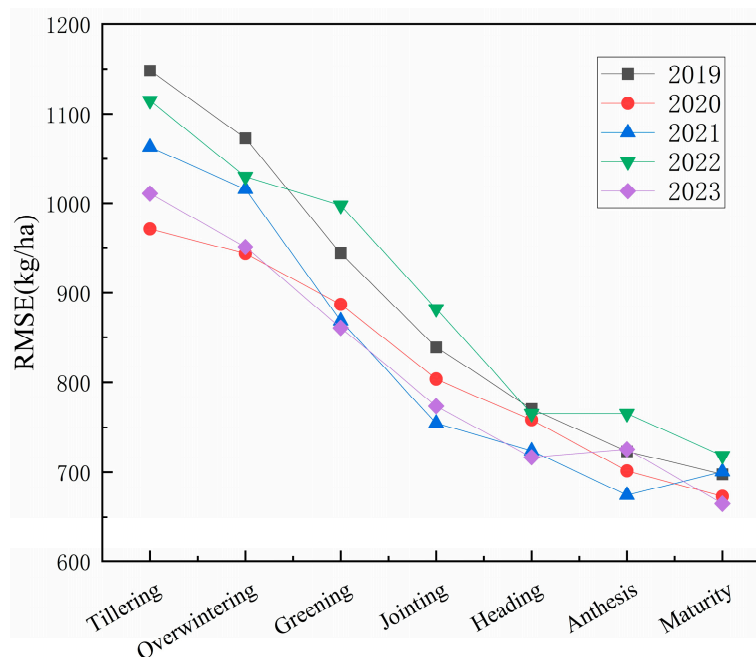
As shown in Table 4, in five years, the mean values of the yield estimation for the same fertility period were 0.43 and 999.84 kg/ha for R<sup>2</sup> and RMSE in the overwintering period, respectively. After winter, wheat enters the greening period; the temperature gradually rises, and under the influence of photosynthesis, winter wheat grows rapidly, and the chlorophyll content of the plant body rises substantially until it reaches the maximum value before the tassel filling period. Accordingly, this is reflected in remote sensing imagery. After overwintering, with the increase in the input model data over time, the model R<sup>2</sup> increased significantly, and the RMSE decreased significantly, indicating that the complex nonlinear relationship between the input characteristics and yield could be significantly captured; the model accuracy leveled off from the late stage of filling to maturity. During the gradual input of data from different fertility stages, the effect of the data on final



yield varies across fertility periods, and there was an evident rising interval in the model accuracy. The mean  $R^2$  increased by 0.25, and the RMSE decreased by 261.52 kg/ha from the overwintering to the heading stage. In terms of wheat growth stage, the dramatic increase in chlorophyll content, enhanced photosynthesis, and increased demand for water and nutrients in winter wheat from jointing to heading are important factors in the formation of wheat yields, and the model could successfully identify the contribution of the characteristics of this stage. The performance of the model was consistent from year to year, as seen in different years.



**Figure 11.** Trends in the multi-feature  $R^2$  by the fertility period, 2019–2023.



**Figure 12.** Trends in the multi-characteristic RMSE by fertility period for the 2019–2023 period.

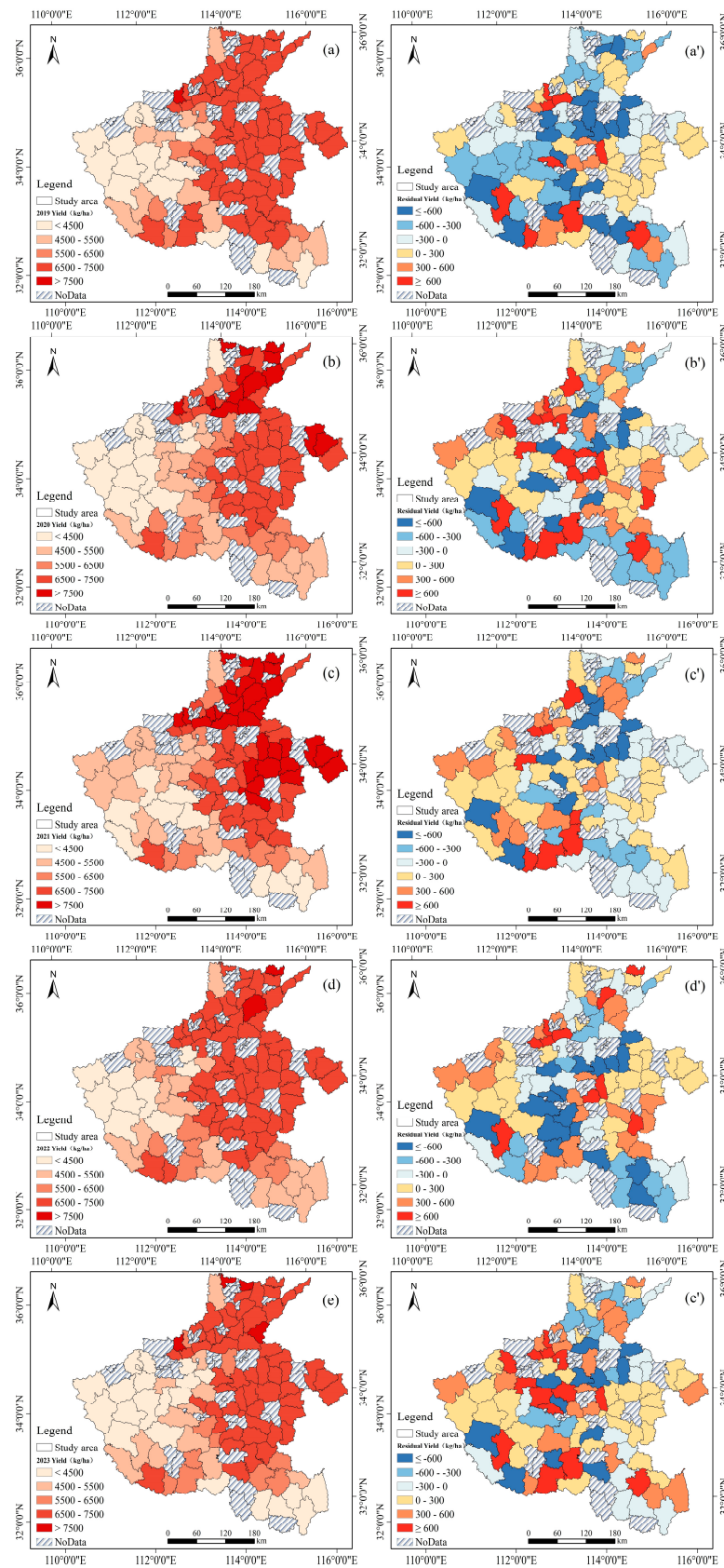
**Table 4.** Mean values of  $R^2$  and RMSE by the fertility period, 2019–2023.

Fertility Period	$R^2$	Change		RMSE (kg/ha)	Change (kg/ha)	
Tillering stage	0.36	0.07	0.10	1062.78	62.94	90.37
Overwintering stage	0.43			999.84		
Greening stage	0.53	0.10	0.05	909.47	104.19	66.96
Jointing stage	0.63			805.28		
Heading stage	0.68	0.03	0.02	738.32	17.49	27.21
Anthesis stage	0.71			720.83		
Maturity stage	0.73			693.62		

An analysis of Table 4 shows that with the gradual increase in the input data, the model can better capture the critical period of wheat growth and gradually improve the ability of yield estimation. In the later stage of growth, as the growth of winter wheat reaches its peak and gradually undergoes senescence, the chlorophyll and water contents in the plant body decline rapidly. The input data reflecting the yield information also tends to be saturated, and the model accuracy is slow to improve and tends to be stable, achieving a better effect in the anthesis period. The five-year average  $R^2$  and RMSE were 0.71 and 720.83 kg/ha, respectively. In summary, the model can obtain good yield estimation around the anthesis period.

### 3.5. Analysis of the Spatiotemporal Distribution of County-Level Estimates of the Winter Wheat Yield for the CNN–MALSTM Model

Figure 13 shows the predicted and residual distributions of the model in each year, where Figure 13a–e are the county-level estimated yield distribution maps of the model for the 2019–2023 period, and Figure 13a'–e' are the differences between the statistical and predicted values of the model in each year. The estimated yield distribution map shows that the forecast results are overall consistent with the trend in the yield distribution in Henan Province, with lower values in the west and higher values in the center and east. This is related to the undulating terrain. The western part of the country is at a higher altitude, and the central and northern plains are suitable for wheat cultivation, while the southern Xinyang area has limited wheat cultivation due to topographical and climatic factors, resulting in lower yields. From the residual plot, the high-yielding areas in the east-central part, such as Boai County, Wuzhi County, Wen County, Yanling County, Xiangcheng County, Joon County, Changge City, and Shangshui County, were underestimated to varying degrees. Low-yielding areas in the west, such as Luanchuan County, Luoning County, Ruzhou City, Xixia County, Xichuan County, and others, were overestimated. Among them, Xixia County had an error greater than 600 kg/ha in all five years of yield estimation, and the area has a high degree of topographic relief, which is affected by topographic shading, leading to a decline in image quality and variations that may be significant across years and locations, further exacerbating the fluctuations in yield. Therefore, these factors may have contributed to the bias in yield estimates. Overall, the yield production in Henan Province increased steadily; the model's estimation results for each year were consistent with the official statistical distribution, the absolute value of the model's overall residuals remained within 600 kg/ha, and the performance of the CNN–MALSTM model in estimating the yields in each county in Henan Province was overall better.



**Figure 13.** Spatiotemporal distribution of the winter wheat yield estimates and residuals for the 2019–2023 period, (a–e) spatial and temporal distributions of the yield estimates for the 2019–2023 period, and (a’–e’) spatiotemporal distributions of the residuals of the yield estimates for the 2019–2023 yield.

## 4. Discussion

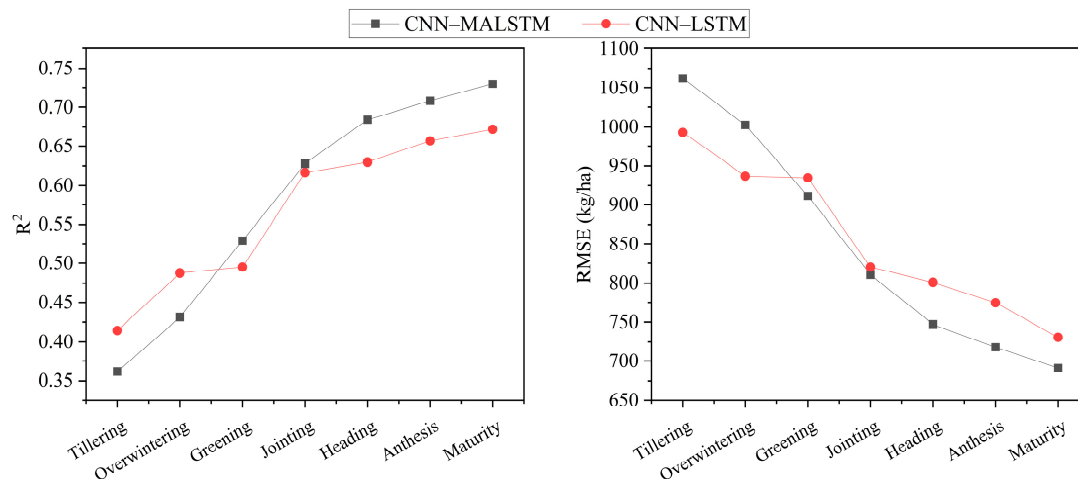
### 4.1. Analysis of the Advantages of the CNN–MALSTM Model

The CNN–MALSTM model proposed in this paper effectively estimates the yield of winter wheat at the county level in Henan Province by integrating multi-source remote sensing data. The model not only realistically reflects the distribution of the data but also captures the key variables that affect the yield of winter wheat throughout the growth and development process. Compared with the baseline model CNN–LSTM, the CNN–MALSTM model incorporates a multi-attention mechanism, which significantly improves the accuracy of yield estimation, especially in overcoming the problem of underestimation of high yield. The yield estimation of the model highly matches the actual yield distribution in Henan Province.

In the cross-validation of the “one year on leave” approach, the model’s yield predictions show stability, although the prediction accuracy varies slightly across years. In addition, the model performed well in the fertility-based sensitivity analysis in different years. Using the sensitivity analyses in Table 4, we found that gradually adding fertility data to the model was effective in quantifying the accuracy of predicted yields and that the model accurately predicted yields as early as 20 days before the harvest of winter wheat, which is in line with the results of a previous study [40]. This 20-day lead time is important for early decision-making and planning by farmers and policymakers to help them adjust their strategies in time to maximize yield potential and resource allocation.

In order to deeply explore the changes in the contribution of CNN–MALSTM to winter wheat yield estimation compared to CNN–LSTM in different years and different fertility periods, we conducted fertility-by-fertility sensitivity analyses of CNN–LSTM and calculated the  $R^2$  and RMSE in each year. By averaging the results of the two models over five years in each fertility period (shown in Figure 14), we are able to intuitively observe the trend of changes in the estimation accuracy of the models at different fertility periods. Before the greening period of winter wheat, the factors affecting the yield are relatively simple due to weak crop growth, such as temperature, precipitation, and other basic meteorological data. These factors can be effectively captured by the CNN–LSTM model, so the model performs well at this stage even without the help of the attention mechanism. However, as the growth of winter wheat advances, the meteorological data (e.g., temperature and precipitation) and remote sensing data are gradually fused, at which point the advantage of the attention mechanism begins to emerge. It can help the model focus more precisely on key yield-related features, which in turn improves the prediction ability. Both models were able to identify important fertility stages of winter wheat (e.g., the jointing to anthesis stage), where the link between key crop characteristics and yield potential became clearer, especially when the EVI reached saturation. The introduction of the multiple attention mechanism allows the model to capture these key characteristics at an early stage, resulting in earlier and more accurate yield predictions, providing relatively accurate predictions about 20 days in advance.

In summary, by combining CNN–LSTM models with multiple attention mechanisms, we revealed the cumulative effects and nonlinear relationships between the influencing factors and yield, and explored the timeliness of crop yield prediction as well as the contribution of multivariate variables to yield estimation at different growth stages. Simple feature relationships in the early stage may allow the model without the attention mechanism to perform better, while in the later stage, as the feature relationships become more complex, the inclusion of the attention mechanism can effectively improve the model’s performance. This phenomenon reflects the different needs for model structure and feature extraction capability at different stages. The model not only improves the reliability and timeliness of yield prediction but also provides a solid scientific basis for relevant management decisions.



**Figure 14.** Mean  $R^2$  and RMSE for fertility-by-fertility projections, 2019–2023.

#### 4.2. Generalisation Performance of the Model

Based on the data processing and preparation in Chapter 2, we applied the CNN–MALSTM model to Pingyuan County, Shandong Province, for a heterogeneous generalization experiment to estimate winter wheat yield between 2019 and 2023. Under the premise of keeping the model structure and parameters unchanged, the prediction accuracy performance of the model in these five years is shown in Table 5.

**Table 5.** Winter wheat acreage and production in Pingyuan County, 2019–2023.

	2019	2020	2021	2022	2023
Official yield (kg/ha)	6612.75	6872.70	7022.55	7085.55	7100.85
Estimation yield (ka/ha)	6373.43	6375.14	6588.78	6647.14	6559.97
Accuracy (%)	96.38%	92.76%	93.82%	93.81%	92.38%

During the study period from 2019 to 2023, the model achieved an  $R^2$  value of 0.73 and an RMSE of 422.21 kg/ha, showing good performance of the model in different years. The average accuracy was as high as 93.80%, and the maximum inter-annual gap was only 4.0%. This result indicates that despite the increased fluctuation in yields in the county after 2019 (Table 5) and the decrease in the accuracy of the model compared to 2019, the overall high predictive ability was still maintained. There is a subtle underestimation of the overall yield, potentially because the distribution of training data features in Henan may be significantly different from those in Shandong, and this distributional difference can affect the predictive accuracy of the model, especially when regional extrapolation is prone to bias. Specifically, the model is able to adapt to climate change and agricultural management practices in different years, ensuring that a high level of prediction accuracy is maintained despite data changes. In addition, the ability to effectively integrate remotely sensed data with meteorological information enhances its applicability to different regions and time periods.

In summary, the performance of the CNN–MALSTM model in heterogeneous generalization confirms its reliability and adaptability in winter wheat yield estimation, and lays the foundation for future applications in other regions.

#### 4.3. Analysis of Model Limitations

Although the proposed CNN–MALSTM model performs well at the county level, there are still some limitations. One of the main challenges is related to data collection. Due to the relatively short runtime of Sentinel-2’s binary star network, the model may have a more limited sample size, which affects its robustness and generalization ability. Faced with the discrepancy between the high-frequency temperature and precipitation data and the



low-frequency satellite imaging data, we unified all the data to the same time step, i.e., the fertility interval, according to the research needs. The other high-frequency meteorological data can be fulfilled within the current time interval only if the satellite images satisfy the prerequisites. To cope with this problem, data augmentation methods such as generative adversarial networks (GANs) and variational autoencoders (VAEs) can be considered in the future to extend the diversity of the dataset. In addition, when Sentinel-2 data are not available at a certain growth stage of winter wheat due to factors such as clouds and rain, a potential solution is to integrate Landsat and Sentinel-2 data through multi-source image fusion techniques to ensure the completeness of the images during the reproductive period. Second, this study explored the ability of the model to be applied in winter wheat yield estimation, mainly from the perspective of the fertility period. Thanks to the high temporal resolution of Sentinel-2, future studies can further explore yield estimation at finer time scales (e.g., 5-, 10-, and 15-day intervals). It is also important to perform a detailed data quality assessment, including the detection of missing values, outliers, and erroneous data, before using meteorological datasets. Consider using interpolation techniques (e.g., mean interpolation, linear interpolation, or model-based interpolation) to fill in missing values and ensure the integrity of the dataset. In cases where there are fewer missing values, there is an option to remove missing data points to avoid any impact on model training.

The diversity of cultivars and field management practices within the vast winter wheat growing region of Henan Province, coupled with the fact that it spans a significant north-south distance, poses additional challenges for yield estimation. For example, there are differences in the phenological periods of winter wheat in northern and southern Henan Province. Although this study used a uniform standard to delineate fertility periods across the province, this may not fully explain the regional differences. Future studies could adjust fertility intervals based on recognized phenological differences to improve the accuracy of input data and ultimately enhance the precision of yield estimates. Additionally, certain factors that significantly influence yield, such as elevation, soil properties, human field management, and irrigation practices, were not included in this study due to limitations in feature quantification and data acquisition. Jiang et al. [41] showed that the introduction of these factors can effectively help the model identify spatial heterogeneity in yield. Although the model performed well on the existing dataset, the inclusion of these additional factors is expected to further enhance its predictive ability. For example, solar-induced fluorescence (SIF), which reflects crop photosynthesis, and the vegetation temperature condition index (VTCI), which is sensitive to moisture stress during the growing season [42], can be integrated into the yield estimation model to further improve its accuracy.

Although deep learning models like CNN-MALSTM are good at revealing complex relationships between data, they are often criticized as “black boxes” due to the black-box nature of their internal computational processes. This lack of transparency can lead to uncertainty, as the intrinsic relationship between input features and target yield remains unclear. In contrast, traditional crop growth models provide a more explanatory characterization of crop growth processes and mechanisms. Future research could attempt to combine crop growth models with deep learning methods to improve the interpretability of the models and apply them to large-scale yield estimation. This hybrid approach could combine both the predictive power of deep learning and the explanatory advantages of crop growth models.

In addition, the model in this study has only been tested on winter wheat and has not been validated for other crops (e.g., corn and soybean). This limits the applicability of the model to some extent, as its accuracy and reliability for other crops cannot be guaranteed. Different crops have different growth cycles and phenological characteristics and therefore require independent modeling and validation to ensure generalization and accuracy of the predictive model. In addition, the model has only been tested in wheat-producing areas in Henan Province and Pingyuan County, and further validation in other areas in the future will help to improve the applicability of the model. Therefore, the generalization and wide applicability of the method still need further research and validation.

## 5. Conclusions

In this study, we developed a CNN–MALSTM deep learning framework that combines the spatial feature extraction capability of convolutional neural networks (1D-CNN) with the advantages of long short-term memory (LSTM) networks in processing time-series data and integrates a multi-attention mechanism to enable the model to efficiently capture the contribution of various features to the final yield. All data processing steps were performed on the Google Earth Engine (GEE) platform, including the pre-processing of Sentinel-2 imagery (resampling and masking) and computation of the EVI. In addition, meteorological data (precipitation and temperature) were derived from the ERA5 dataset, extracted hour by hour through GEE, and aggregated to the fertility scale. This processing method makes the data processing more efficient and guarantees the spatial and temporal consistency of the data.

The results show that the CNN–MALSTM model exhibits excellent accuracy in estimating winter wheat yield at the county level in Henan Province, outperforming the baseline model CNN–LSTM. In particular, the model is able to accurately predict the yield about 20 days before harvest, showing good robustness and reliability. In addition, the spatio-temporal distribution map of county-level winter wheat yields in Henan Province from 2019 to 2023 demonstrated a trend of lower yields in the western region and higher yields in the central-eastern region, along with a reduction in inter-annual variability. The results are highly correlated with the official statistics, indicating that the proposed CNN–MALSTM framework has significant potential for yield estimation.

Future research could explore the application of the CNN–MALSTM framework to other crops or regions to assess its generality and adaptability. In addition, incorporating other variables that affect yield, such as soil, geographic, or pest and disease information, may further improve the predictive power and robustness of the model. These research avenues will help expand the range of model applications and provide broader technical support for global agricultural yield prediction.

**Author Contributions:** Conceptualization, C.L. and L.Z.; Methodology, C.L. and L.Z.; Resources, C.L., X.W. and H.C.; Data curation, C.L., L.Z., X.W., H.X. and Y.J.; Writing—original draft, L.Z.; Writing—review & editing, C.L.; Supervision, C.L.; Funding acquisition, C.L. and H.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by the National Major Scientific Research Achievement Cultivation Fund (NSFRF240101), the Fundamental Research Funds for the Universities of Henan Province (242300420221), and the Dual-Class Program in Surveying and Mapping Science and Technology (BZCG202301). Additional support was provided by the Henan Provincial Postdoctoral Research Launch Project (202103072), the Henan Polytechnic University Doctoral Fund Project (B2021-19), Henan Polytechnic University’s High-Level Talent Development Program for the Establishment of “Double First-Class” Discipline in Surveying and Mapping Science and Technology (GCCYJ202427), and the Key Research Project of Higher Education Institutions in Henan Province (25A420002).

**Data Availability Statement:** Dataset available on request from the authors.

**Conflicts of 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.

## References

1. Sundström, J.F.; Albiñ, A.; Boqvist, S.; Ljungvall, K.; Marstorp, H.; Martiin, C.; Nyberg, K.; Vågsholm, I.; Yuen, J.; Magnusson, U. Future threats to agricultural food production posed by environmental degradation, climate change, and animal and plant diseases—A risk analysis in three economic and climate settings. *Food Secur.* **2014**, *6*, 201–215. [\[CrossRef\]](#)
2. Liu, B.; Asseng, S.; Müller, C.; Ewert, F.; Elliott, J.; Lobell David, B.; Martre, P.; Ruane Alex, C.; Wallach, D.; Jones James, W.; et al. Similar estimates of temperature impacts on global wheat yield by three independent methods. *Nat. Clim. Chang.* **2016**, *6*, 1130–1136. [\[CrossRef\]](#)
3. Alvarez, R. Predicting average regional yield and production of wheat in the Argentine Pampas by an artificial neural network approach. *Eur. J. Agron.* **2009**, *30*, 70–77. [\[CrossRef\]](#)

4. Cai, Y.; Guan, K.; Lobell, D.; Potgieter, A.B.; Wang, S.; Peng, J.; Xu, T.; Asseng, S.; Zhang, Y.; You, L.; et al. Integrating satellite and climate data to predict wheat yield in Australia using machine learning approaches. *Agric. For. Meteorol.* **2019**, *274*, 144–159. [\[CrossRef\]](#)
5. Ali, A.M.; Abouelghar, M.; Belal, A.A.; Saleh, N.; Yones, M.; Selim, A.I.; Amin, M.E.S.; Elwesemy, A.; Kucher, D.E.; Maginan, S.; et al. Crop Yield Prediction Using Multi Sensors Remote Sensing (Review Article). *Egypt. J. Remote Sens. Space Sci.* **2022**, *25*, 711–716. [\[CrossRef\]](#)
6. Brown, J.N.; Hochman, Z.; Holzworth, D.; Horan, H. Seasonal climate forecasts provide more definitive and accurate crop yield predictions. *Agric. For. Meteorol.* **2018**, *260–261*, 247–254. [\[CrossRef\]](#)
7. Feng, P.; Wang, B.; Liu, D.L.; Waters, C.; Yu, Q. Incorporating machine learning with biophysical model can improve the evaluation of climate extremes impacts on wheat yield in south-eastern Australia. *Agric. For. Meteorol.* **2019**, *275*, 100–113. [\[CrossRef\]](#)
8. Xie, Y.; Huang, J. Integration of a Crop Growth Model and Deep Learning Methods to Improve Satellite-Based Yield Estimation of Winter Wheat in Henan Province, China. *Remote Sens.* **2021**, *13*, 4372. [\[CrossRef\]](#)
9. Zhuang, H.; Zhang, Z.; Cheng, F.; Han, J.; Luo, Y.; Zhang, L.; Cao, J.; Zhang, J.; He, B.; Xu, J.; et al. Integrating data assimilation, crop model, and machine learning for winter wheat yield forecasting in the North China Plain. *Agric. For. Meteorol.* **2024**, *347*, 109909. [\[CrossRef\]](#)
10. Huang, J.; Tian, L.; Liang, S.; Ma, H.; Becker-Reshef, I.; Huang, Y.; Su, W.; Zhang, X.; Zhu, D.; Wu, W. Improving winter wheat yield estimation by assimilation of the leaf area index from Landsat TM and MODIS data into the WOFOST model. *Agric. For. Meteorol.* **2015**, *204*, 106–121. [\[CrossRef\]](#)
11. Zhuo, W.; Huang, J.; Xiao, X.; Huang, H.; Bajgain, R.; Wu, X.; Gao, X.; Wang, J.; Li, X.; Wagle, P. Assimilating remote sensing-based VPM GPP into the WOFOST model for improving regional winter wheat yield estimation. *Eur. J. Agron.* **2022**, *139*, 126556. [\[CrossRef\]](#)
12. Peng, B.; Guan, K.; Zhou, W.; Jiang, C.; Frankenberg, C.; Sun, Y.; He, L.; Köhler, P. Assessing the benefit of satellite-based Solar-Induced Chlorophyll Fluorescence in crop yield prediction. *Int. J. Appl. Earth Obs. Geoinf.* **2020**, *90*, 102126. [\[CrossRef\]](#)
13. Ren, S.; Guo, B.; Wu, X.; Zhang, L.; Ji, M.; Wang, J. Winter wheat planted area monitoring and yield modeling using MODIS data in the Huang-Huai-Hai Plain, China. *Comput. Electron. Agric.* **2021**, *182*, 106049. [\[CrossRef\]](#)
14. Besalatpour, A.A.; Ayoubi, S.; Hajabbasi, M.A.; Jazi, A.Y.; Gharipour, A. Feature Selection Using Parallel Genetic Algorithm for the Prediction of Geometric Mean Diameter of Soil Aggregates by Machine Learning Methods. *Arid Land Res. Manag.* **2014**, *28*, 383–394. [\[CrossRef\]](#)
15. Zeraatpisheh, M.; Ayoubi, S.; Jafari, A.; Tajik, S.; Finke, P. Digital mapping of soil properties using multiple machine learning in a semi-arid region, central Iran. *Geoderma* **2019**, *338*, 445–452. [\[CrossRef\]](#)
16. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Comput. Electron. Agric.* **2018**, *151*, 61–69. [\[CrossRef\]](#)
17. Naimi, S.; Ayoubi, S.; Demattê, J.A.M.; Zeraatpisheh, M.; Amorim, M.T.A.; Mello, F.A.d.O. Spatial prediction of soil surface properties in an arid region using synthetic soil image and machine learning. *Geocarto Int.* **2022**, *37*, 8230–8253. [\[CrossRef\]](#)
18. Son, N.-T.; Chen, C.-F.; Chen, C.-R.; Guo, H.-Y.; Cheng, Y.-S.; Chen, S.-L.; Lin, H.-S.; Chen, S.-H. Machine learning approaches for rice crop yield predictions using time-series satellite data in Taiwan. *Int. J. Remote Sens.* **2020**, *41*, 7868–7888. [\[CrossRef\]](#)
19. Sellers, P.J.; Berry, J.A.; Collatz, G.J.; Field, C.B.; Hall, F.G. Canopy reflectance, photosynthesis, and transpiration. III. A reanalysis using improved leaf models and a new canopy integration scheme. *Remote Sens. Environ.* **1992**, *42*, 187–216. [\[CrossRef\]](#)
20. Guan, K.; Wu, J.; Kimball, J.S.; Anderson, M.C.; Frolking, S.; Li, B.; Hain, C.R.; Lobell, D.B. The shared and unique values of optical, fluorescence, thermal and microwave satellite data for estimating large-scale crop yields. *Remote Sens. Environ.* **2017**, *199*, 333–349. [\[CrossRef\]](#)
21. Johnson, D.M. A comprehensive assessment of the correlations between field crop yields and commonly used MODIS products. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *52*, 65–81. [\[CrossRef\]](#)
22. Zhou, X.; Wang, P.; Tansey, K.; Zhang, S.; Li, H.; Wang, L. Developing a fused vegetation temperature condition index for drought monitoring at field scales using Sentinel-2 and MODIS imagery. *Comput. Electron. Agric.* **2020**, *168*, 105144. [\[CrossRef\]](#)
23. Tian, H.; Wang, P.; Tansey, K.; Han, D.; Zhang, J.; Zhang, S.; Li, H. A deep learning framework under attention mechanism for wheat yield estimation using remotely sensed indices in the Guanzhong Plain, PR China. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102375. [\[CrossRef\]](#)
24. Bolton, D.K.; Friedl, M.A. Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics. *Agric. For. Meteorol.* **2013**, *173*, 74–84. [\[CrossRef\]](#)
25. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Noda, K.; Yamaguchi, Y.; Nakadai, K.; Okuno, H.G.; Ogata, T. Audio-visual speech recognition using deep learning. *Appl. Intell.* **2015**, *42*, 722–737. [\[CrossRef\]](#)
27. Kamilaris, A.; Prenafeta-Boldú, F.X. Deep learning in agriculture: A survey. *Comput. Electron. Agric.* **2018**, *147*, 70–90. [\[CrossRef\]](#)
28. Khoshroo, A.; Emrouznejad, A.; Ghaffarizadeh, A.; Kasraei, M.; Omid, M. Sensitivity analysis of energy inputs in crop production using artificial neural networks. *J. Clean. Prod.* **2018**, *197*, 992–998. [\[CrossRef\]](#)
29. Chen, Y.; Jiang, H.; Li, C.; Jia, X.; Ghamisi, P. Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6232–6251. [\[CrossRef\]](#)

30. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)]
31. Chung, J.; Gülçehre, Ç.; Cho, K.; Bengio, Y. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *arXiv* **2014**, arXiv:abs/1412.3555.
32. Xiao, G.; Zhang, X.; Niu, Q.; Li, X.; Li, X.; Zhong, L.; Huang, J. Winter wheat yield estimation at the field scale using sentinel-2 data and deep learning. *Comput. Electron. Agric.* **2024**, *216*, 108555. [[CrossRef](#)]
33. Wang, J.; Wang, P.; Tian, H.; Tansey, K.; Liu, J.; Quan, W. A deep learning framework combining CNN and GRU for improving wheat yield estimates using time series remotely sensed multi-variables. *Comput. Electron. Agric.* **2023**, *206*, 107705. [[CrossRef](#)]
34. Sun, J.; Lai, Z.; Di, L.; Sun, Z.; Tao, J.; Shen, Y. Multilevel Deep Learning Network for County-Level Corn Yield Estimation in the U.S. Corn Belt. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 5048–5060. [[CrossRef](#)]
35. Zhou, M.; Ma, X.; Wang, K.; Cheng, T.; Tian, Y.; Wang, J.; Zhu, Y.; Hu, Y.; Niu, Q.; Gui, L.; et al. Detection of phenology using an improved shape model on time-series vegetation index in wheat. *Comput. Electron. Agric.* **2020**, *173*, 105398. [[CrossRef](#)]
36. Chen, J.; Jönsson, P.; Tamura, M.; Gu, Z.; Matsushita, B.; Eklundh, L. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sens. Environ.* **2004**, *91*, 332–344. [[CrossRef](#)]
37. Li, C.; Chen, W.; Wang, Y.; Wang, Y.; Ma, C.; Li, Y.; Li, J.; Zhai, W. Mapping Winter Wheat with Optical and SAR Images Based on Google Earth Engine in Henan Province, China. *Remote Sens.* **2022**, *14*, 284. [[CrossRef](#)]
38. Wang, D.; Zhang, Z.; Jiang, Y.; Mao, Z.; Wang, D.; Lin, H.; Xu, D. DM3Loc: Multi-label mRNA subcellular localization prediction and analysis based on multi-head self-attention mechanism. *Nucleic Acids Res.* **2021**, *49*, e46. [[CrossRef](#)]
39. Guo, F.; Wang, P.; Tansey, K.; Zhang, Y.; Li, M.; Liu, J.; Zhang, S. A novel transformer-based neural network under model interpretability for improving wheat yield estimation using remotely sensed multi-variables. *Comput. Electron. Agric.* **2024**, *223*, 109111. [[CrossRef](#)]
40. Wang, Y.; Zhang, Z.; Feng, L.; Ma, Y.; Du, Q. A new attention-based CNN approach for crop mapping using time series Sentinel-2 images. *Comput. Electron. Agric.* **2021**, *184*, 106090. [[CrossRef](#)]
41. Jiang, H.; Hu, H.; Zhong, R.; Xu, J.; Xu, J.; Huang, J.; Wang, S.; Ying, Y.; Lin, T. A deep learning approach to conflating heterogeneous geospatial data for corn yield estimation: A case study of the US Corn Belt at the county level. *Glob. Chang. Biol.* **2020**, *26*, 1754–1766. [[CrossRef](#)] [[PubMed](#)]
42. Xie, Y.; Wang, P.; Bai, X.; Khan, J.; Zhang, S.; Li, L.; Wang, L. Assimilation of the leaf area index and vegetation temperature condition index for winter wheat yield estimation using Landsat imagery and the CERES-Wheat model. *Agric. For. Meteorol.* **2017**, *246*, 194–206. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.