



Phenological piecewise modelling is more conducive than whole-season modelling to winter wheat yield estimation based on remote sensing data

Xin Huang, Wenquan Zhu, Cenliang Zhao, Zhiying Xie & Hui Zhang

To cite this article: Xin Huang, Wenquan Zhu, Cenliang Zhao, Zhiying Xie & Hui Zhang (2022) Phenological piecewise modelling is more conducive than whole-season modelling to winter wheat yield estimation based on remote sensing data, European Journal of Remote Sensing, 55:1, 338-352, DOI: [10.1080/22797254.2022.2073916](https://doi.org/10.1080/22797254.2022.2073916)

To link to this article: <https://doi.org/10.1080/22797254.2022.2073916>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 24 May 2022.



[Submit your article to this journal](#)



Article views: 76



[View related articles](#)



[View Crossmark data](#)

Phenological piecewise modelling is more conducive than whole-season modelling to winter wheat yield estimation based on remote sensing data

Xin Huang^{a,b}, Wenquan Zhu^{a,b}, Cenliang Zhao^{a,b}, Zhiying Xie^{a,b} and Hui Zhang^{a,b}

^aState Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Aerospace Information Research Institute of Chinese Academy of Sciences, Faculty of Geographical Science, Beijing Normal University, Beijing, China; ^bBeijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Beijing Normal University, Beijing, China

ABSTRACT

Most of the existing remote sensing-based yield estimation methods adopt the mean or cumulative value of meteorological factors within the whole growing season, which may ignore the impact of adverse meteorological conditions on the growth of winter wheat in a certain phenological period. In this study, we distinguished the developmental progression of winter wheat as three phenological periods. In each phenological period, the vegetation indices and meteorological factors were optimized. Then the accuracy and spatiotemporal transferability of the phenological piecewise modelling was compared with that of the whole-season modelling based on four regression methods (i.e. multiple linear regression, artificial neural network, support vector regression and random forest). The results showed that the optimal combinations of variables for the whole-season modelling and the phenological piecewise modelling were different. Compared with the whole-season models, the R^2 for the phenological piecewise models improved by 1.4% to 7.6%, the root mean square error (RMSE) decreased by 1.1% to 8.2% among four regression methods. In addition, compared with the whole-season models, the spatiotemporal transferability for the phenological piecewise models was generally better. The accuracies after spatiotemporal transfer for the phenological piecewise models were still higher than that for the whole-season models.

ARTICLE HISTORY

Received 28 June 2021
Revised 24 March 2022
Accepted 2 May 2022

KEYWORDS

Meteorological factors; phenology; remote sensing; winter wheat; yield estimation; machine learning

Introduction

The capacity of global food supply is facing increasing pressure due to the population growth, climate change, and ecological environment deterioration (Godfray et al., 2010; Rounsevell et al., 2005; Searchinger et al., 2008). Winter wheat is one of the most widely cultivated food crops in the world. Its growth and yield are closely related to global food security (Thenkabail et al., 2012). Therefore, timely and accurate estimation of the winter wheat yield in a large region is of great significance to guide agricultural production and improve the agricultural disaster response ability.

Remote sensing statistical models, crop growth models, and data assimilation models are the three commonly used methods to estimate crop yield (Huang et al., 2019). Crop growth models predict the crop yield by simulating the growth processes of crop from sowing to harvest (Challinor et al., 2004; Jones et al., 2003). It has a clear physiological mechanism and can be successfully applied at a field scale (Jego et al., 2012). However, the various model parameters of crop growth models are hard to obtain at a large scale and crops growth models are usually based on simplified growth processes, which may lead to some uncertainties (Huang et al., 2019). The data

assimilation method can solve this problem to some extent by assimilating remote sensing data and crop growth models, but its calculation processes are complicated and the computational efficiency is not satisfactory at a large scale (Lee et al., 2010). Compared with the two methods above, the remote sensing statistical model is a type of empirical models. Although the established model may only perform well on specific crop cultivars or certain geographical regions (Doraiswamy et al., 2003; Fang et al., 2011), it is still commonly adopted to estimate the crop yield due to its advantages of few data requirements, fast computational efficiency, and applicable to a large scale (Kowalik et al., 2014; Liao et al., 2019; Nolasco et al., 2021). Since the vegetation indices calculated by remote sensing reflectance data can reflect the growth status of vegetation and has a high correlation with yield, the early developed statistical methods were based on the linear statistical regression models between the vegetation index and the actual yield (Mkhabela et al., 2011; Rasmussen, 1997; Siyal et al., 2015). After that, more and more studies notice the improvement effect of combining the remote sensing vegetation index with meteorological factors (such as temperature, precipitation and sunlight; Kamir et al.,

2020; Prasad et al., 2006; Sharifi, 2020; Stepanov et al., 2020). In addition, the phenology variables such as start of season, heading date and mature date can reflect the duration of crop key growth periods, which control the accumulation of effective dry matter (Saseendran et al., 2000). In recent years, some studies improved the accuracy of crop yield statistical estimation models by introducing crop phenology variables as additional independent variables (Guo et al., 2021; Ji et al., 2021; Sun et al., 2020). Except for the abundance of model variables, the spatial resolution of remote sensing data has been also improved with the development of satellite sensors. The main sensor products used in crop yield estimation include: SPOT-VEGETATION composite products (1000 meters, 10 days; Kowalik et al., 2014), MODIS composite products (250/500 meters, 8/16 days; Funk & Budde, 2009; Leroux et al., 2019), Landsat products (30 meters, 16 days; Liao et al., 2019; Siyal et al., 2015), and Sentinel-2 products (10 meters, 3–5 days; He & Mostovoy, 2019; Hunt et al., 2019). Different satellite sensors may have different characteristics (e.g. with a higher spatial resolution but a lower temporal resolution). It is necessary to select appropriate remote sensing products according to the study area and research objectives.

Although the current remote sensing statistical yield estimation methods are relatively mature, there is still room for improvement. Firstly, in terms of the vegetation index, some studies have compared the performance of commonly used vegetation indices on crop yield estimation, such as the normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI; Johnson, 2016; Nolasco et al., 2021; Tuvdendorj et al., 2019). However, some new vegetation indices with great potential in yield estimation are rarely evaluated, such as the near-infrared index (NIRv), which is closely related to solar-induced chlorophyll fluorescence (Badgley et al., 2017), and the normalized difference phenology index (NDPI), which is extremely insensitive to soil and contains vegetation water information (Chen et al., 2019; Wang et al., 2017). In addition, different vegetation indices may show different applicability in different phenophases due to their different emphasis. Secondly, when introducing meteorological factors to improve the yield estimation model, most studies considered the meteorological factors within the whole growing season. Even in the study that integrating meteorology and phenology, the phenology was usually taken as new independent variables into the model (Guo et al., 2021; Ji et al., 2021). Since crops have different requirements for hydrothermal conditions in different growth phases, using the mean/accumulated value of meteorological factors within

the whole growing season may weaken/ignore the impact of adverse meteorological conditions on crop growth in a certain phase. Therefore, the phenological piecewise modelling, i.e. separating the growing season into different phenological periods and selecting appropriate input variables in each phenological period to construct the yield estimation model, may be more conducive to the improvement of yield estimation.

The main objectives of this study are: (1) to compare the optimized variables between the whole-season yield modelling and the phenological piecewise yield modelling; (2) to compare the estimation accuracy between the whole-season yield models and the phenological piecewise yield models; and (3) to compare the spatiotemporal transferability (i.e. the ability of models to predict the winter wheat yield in other years or regions whose samples did not participate in the model training) between the whole-season models and the phenological piecewise models.

Materials

Study area

The Henan Province is located in central China (110.4°E – 116.6°E, 31.4°N – 36.4°N), covering approximately 167,000 square kilometres (Figure 1a). Under the temperate monsoon climate, the annual averaged temperature in Henan Province in 2018 was 15.5°C, and the annual precipitation was 748.9 mm (Henan Statistics Bureau, 2019). Henan Province is one of the main grain-producing provinces in China, and the main crops are winter wheat and summer maize (Figure 1b). According to the China Rural Statistical Yearbook (National Bureau of Statistics of China, 2019), in 2018, the sown area of winter wheat in the province was 5.74 million ha, and the total yield of winter wheat in the province was 36.03 million tons, ranking the first among all provinces in China. Henan Province mainly consists of plains, and also has some mountainous regions (Figure 1c). In the mountainous regions in the west of Henan Province, the fields of winter wheat are relatively small and scattered (Figure 1d). In contrast, the fields of winter wheat cultivated in the plains in the east of Henan Province are large and concentrated (Figure 1e and 1f). Based on the FROM-GLC10 land cover product with 10-m spatial resolution (Gong et al., 2019), the statistical results of the area of each cropland patch in Henan Province showed that more than 75% of the cropland patches are over 25 hectares (i.e. 500 m × 500 m) (Figure S1).

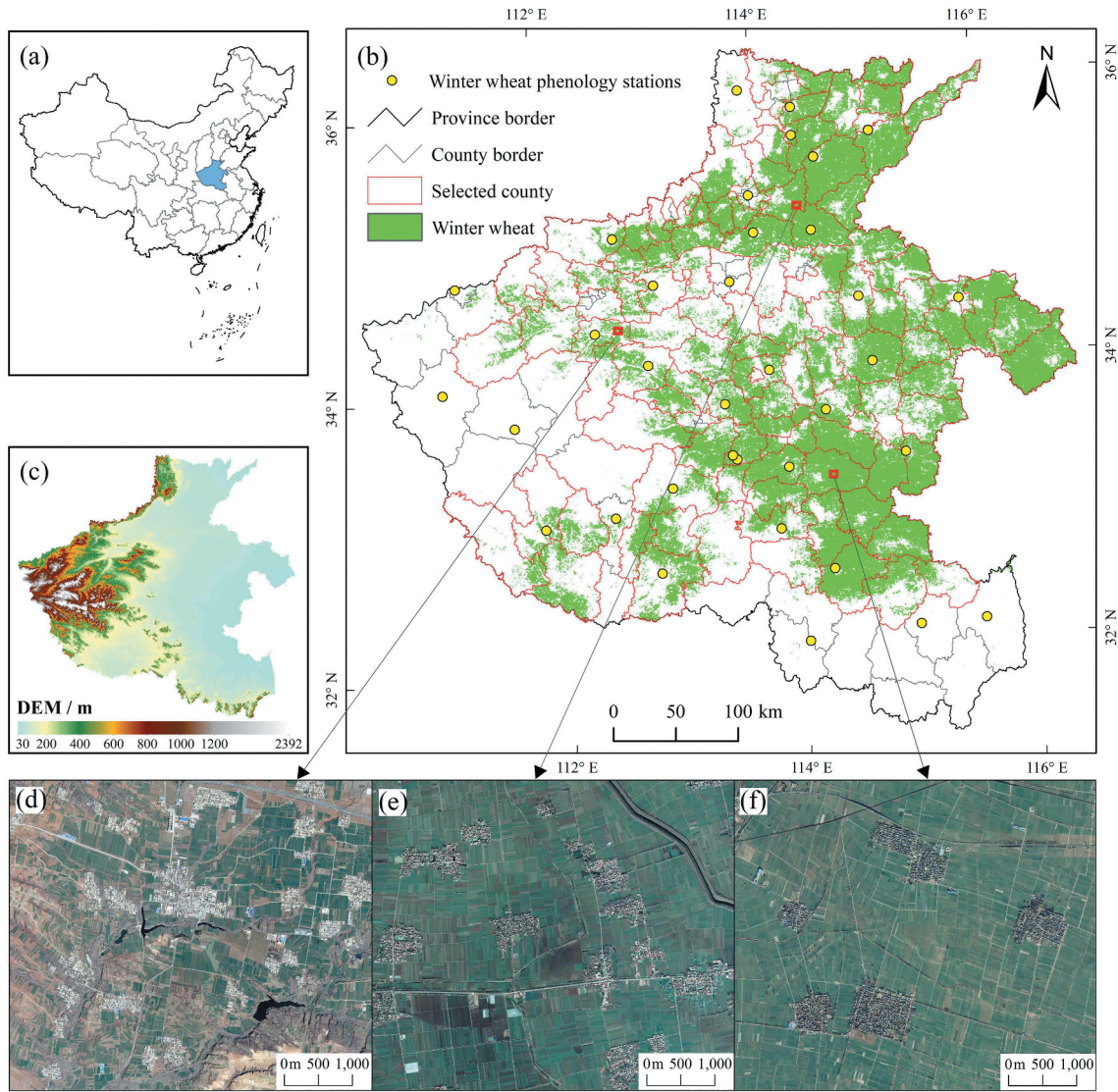


Figure 1. Study area. (a) location map, (b) spatial distributions of winter wheat and phenology observation stations, (c) topographic map, (d) google earth image of the mountainous regions in the west of Henan province, (e) google earth image of the plains in the northeast of Henan province, (f) google earth image of the plains in the southeast of Henan province. The data sources of winter wheat distribution and winter wheat phenology stations refer to section 2.2.4 and Appendix A, respectively.

Data

Remote sensing data

The 8-day composited product at 500 m resolution from 2017 to 2018, MODIS surface reflectance data set (MOD09A1 Version 6) was used to calculate the vegetation indices (Vermote, 2015). There are seven reflectance bands including blue band (459–479 nm), green band (545–565 nm), red band (620–670 nm), near-infrared bands (841–876 nm and 1230–1250 nm) and short-wave infrared bands (1628–1652 nm and 2105–2155 nm). The vegetation indices, NDVI (Rouse et al., 1974), EVI (Huete et al., 2002), NIRv (Badgley et al., 2017) and NDPI (Wang et al., 2017), used in this study were calculated based on Equation (1)–(4), respectively.

$$\text{NDVI} = \frac{\rho_{\text{NIR}} - \rho_{\text{R}}}{\rho_{\text{NIR}} + \rho_{\text{R}}} \quad (1)$$

$$\text{EVI} = 2.5 \times \frac{\rho_{\text{NIR}} - \rho_{\text{R}}}{\rho_{\text{NIR}} + 6\rho_{\text{R}} - 7.5\rho_{\text{B}} + 1} \quad (2)$$

$$\text{NIRv} = \text{NDVI} \times \rho_{\text{NIR}} \quad (3)$$

$$\text{NDPI} = \frac{\rho_{\text{NIR}} - (0.74 \times \rho_{\text{R}} + 0.26 \times \rho_{\text{SWIR}})}{\rho_{\text{NIR}} + (0.74 \times \rho_{\text{R}} + 0.26 \times \rho_{\text{SWIR}})} \quad (4)$$

where ρ_{NIR} , ρ_{R} , ρ_{B} , ρ_{SWIR} represent the surface reflectance of the near-infrared band (841–876 nm), red band (620–670 nm), blue band (459–479 nm) and short-wave infrared band (1628–1652 nm), respectively.

To reduce the effects of clouds and atmospheric conditions on vegetation index time-series data, the change-weight filtering method (Zhu et al., 2012), which can effectively reduce noise and preserve crop phenology characteristics, was used to reconstruct the vegetation index time-series data with good quality.

Meteorological data

The meteorological data set CFSv2 (Climate Forecast System Version 2) was used in this study, which was released by the National Centers for Environmental Prediction (NCEP; Saha et al., 2014). The CFSv2 reanalysis product is produced by the Climate Forecast System, which is a fully coupled model representing the interaction between the Earth's atmosphere, oceans, land, and sea ice. The temporal resolution is 6 h, and the spatial resolution is 0.2°. The data set contains a variety of meteorological data worldwide since 1979 and was commonly used in related researches (Carbin et al., 2016; Pegion & Alexander, 2013). Considering the growth and development of crops are mainly affected by temperature, moisture and sunlight conditions, the daily mean temperature 2 m above the ground, total precipitation, mean specific humidity, mean soil moisture content 5 cm depth below the surface layer and total short-wave solar radiation were used and computed on the Google Earth Engine platform in this study.

Phenology data

The phenology data used in this study were the spatial distributions of the start dates of four winter wheat phenological events in Henan province: the start date of regreening (BBCH 21), the start date of jointing (BBCH 31), the start date of heading (BBCH 51) and the start date of milking maturity (BBCH 73). The regreening date is the date when the leaves of winter wheat begin to turn green after overwintering. It is the date that marks the beginning of winter wheat growth, that is, the start of the growing season. The jointing date is the date when the stem of winter wheat begins to elongate fast, which indicates that winter wheat has entered a period of rapid growth. The heading date is the date when the fully developed spike of winter wheat extends out of the top leaf with the elongation of the stem, which marks the transition from vegetative growth to reproductive growth. The milking maturity date is the date when the grain filling of winter wheat is complete, which indicates that winter wheat has ended its growth and entered the mature stage. According to the phenology calendar released by the Ministry of Agriculture and Rural Affairs of the People's Republic of China (<http://www.moa.gov.cn/>), in Henan Province, these four dates usually occur in early March, early April, late April, and late May. These four dates are important time nodes during

the growing season of winter wheat, and the demands of winter wheat for hydrothermal conditions vary greatly during these periods (Table S1).

The start date of regreening was extracted by the dynamic threshold method (White et al., 1997) based on NDVI time-series data, and the threshold was set as 20%, which was verified to be suitable for extracting the regreening date of winter wheat (RMSE is about 10 days; Gan et al., 2020). The start date of heading was provided by Huang et al. (2020), which was extracted by the accumulated temperature method with a high accuracy (RMSE is about 5 days). The start dates of jointing and milking maturity were generated by adding a fixed duration (number of days) to the start date of regreening and the start date of heading, respectively. The fixed duration was obtained from local phenological observation data (see details in Appendix A).

Winter wheat map and yield statistics data

Since this paper focuses on the comparison of whole-season yield modelling and phenological piecewise yield modelling, we pursue a high user accuracy to ensure the classified winter wheat pixels are as much as possible the true winter wheat pixels. The winter wheat map used in this study (Figure 1) was the intersection of the winter wheat maps in 2017 and 2018, which were both produced by a double threshold method (Huang et al., 2020; see details in Supplementary Materials). The spatial resolution is 250 m. The overall accuracies of winter wheat maps in 2017 and 2018 are 90.0% and 89.2%, respectively. The producer accuracies are 87.5% and 85.4%, respectively. The user accuracies are 95.4% and 96.1%, respectively. The kappa coefficients are 0.80 and 0.79, respectively.

The county-level statistical data of winter wheat yield from 2017 to 2018 in Henan Province, China were downloaded from the EPS data platform (<http://olap.epsnet.com.cn/>) as the reference data to train the yield estimation models and assess accuracy. To reduce the uncertainty from some counties with small winter wheat planting area, the top 100 counties with winter wheat planting area in Henan Province (the number of winter wheat pixels for each county is between 1721 and 34,005, and the winter wheat cultivated area for each county is between 10.8×10^3 ha and 212.5×10^3 ha) were finally selected in this study (Figure 1).

Methods

This study was conducted on the county level and the spatial aggregation of the data was realized by calculating the average value of pixels in each county. Therefore, the dependent variable is the winter wheat yield (kg/ha) in each county from 2017 to 2018, and

the independent variables are the spatial mean values of optimized vegetation indices or meteorological factors corresponding to winter wheat pixels in each county from 2017 to 2018. The variable optimization, model building and accuracy assessment were conducted based on the samples from both years. In addition, to ensure the robustness of the results, these steps were also conducted based on the samples from each year respectively.

Variable optimization

Variable optimization was divided into two parts and based on the multiple linear regression model. The first part was the optimization of vegetation indices, and then the meteorological factors were optimized. In the optimization of vegetation indices, four vegetation indices, NDVI, EVI, NIRv and NDPI, as well as two calculation methods, the maximum value in a growth phase (VI_{max}) and the mean value in a growth phase (VI_{mean}), were considered. The combination of vegetation index and calculation method which showed the highest R^2 with the county-level yield statistics was selected. Then the optimal vegetation index was combined with different meteorological factors to build multiple linear regression models with different independent variables. Finally, the multiple linear regression model was selected based on two criteria: (1) it showed the highest R^2 with the county-level yield statistics; and (2) each independent variable of the model passed the $P < 0.05$ significance test. The independent variables of the selected model were considered as the optimal input variables. The meteorological factors included daily mean air temperature (T), total precipitation (P), mean soil moisture content (SM), mean specific humidity (SH) and total solar radiation (SR). In addition, since the precipitation, soil moisture content and specific humidity all characterize the water condition, only the variable with the highest R^2 among them was selected when combined with other meteorological factors.

The growth and development process of winter wheat was divided into three phenophases: from greening to joining, from jointing to heading and from heading to milking maturity. In each phenophase, the vegetation indices and meteorological factors were optimized according to the above method. The selected variables in each phenophase were all used as independent variables to participate in the modelling.

Yield estimation models

To test whether the phenological piecewise modelling always performed better than the whole-season modelling based on different regression methods, in addition to the traditional multiple linear regression

(MLR), three commonly used and state-of-the-art machine learning regression models, artificial neural network (ANN), support vector regression (SVR) and random forest regression (RF), were selected in this study. ANN, SVR, and RF were implemented based on the “nnet”, “e1071”, and “randomForest” packages within the R environment software, respectively. The zero-mean normalization was conducted and the parameters were optimized by comparing the errors of the models with different combinations of parameters (Table 1).

Multiple linear regression

Multiple linear regression is a linear regression model applied to the case that multiple independent variables jointly affect the dependent variable. Compared with the simple linear regression model, it is more effective and closer to reality (Sousa et al., 2007).

Artificial neural network

Artificial neural network imitates the natural neural network, and it can effectively solve the complex regression problem with a large number of correlated variables. The structure of an artificial neural network includes an input layer, multiple hidden layers and an output layer. Each hidden layer is formed by several artificial neural nodes, and the estimation error is transmitted to the feedforward neural network through the backpropagation algorithm. The weight of each node is adjusted through continuous iteration until the estimation error reaches an acceptable level (Tsai & Lee, 1999; Zhou, 1999).

Support vector regression

Support vector regression is a kind of regression models based on the kernel function. When solving non-linear problems, low-dimensional variables are mapped to high-dimensional space by the kernel function. By finding the optimal hyperplane (the sum of the distances from all points to the hyperplane is the

Table 1. Parameter settings of the three machine learning regression models.

Models	Parameters
Artificial Neural Network (ANN)	The number of hidden layers = 3 The number of nodes in each hidden layer = 5 Learning rate = 0.01 Momentum = 0.4 Error criterium = “LMS” Hidden layer = “Tansig” Output layer = “Tansig” Method = “BATCHgdwm”
Support Vector Regression (SVR)	Kernel = “radial” Gamma = 0.25 Cost = 1 Episode = 0.2
Random Forest (RF)	The number of trees = 500 Mtry = 3

shortest) in the high-dimensional space, the nonlinear problem is transformed into a linear problem (Bennett & Demiriz, 1998; Suykens & Vandewalle, 1999).

Random forest regression

Random forest regression is an ensemble learning algorithm, which uses the bootstrap resampling method to generate different sample subsets to build different decision trees, and calculates the average of the prediction results of all decision trees as the final prediction result (Briem et al., 2001). It has become one of the most widely used machine learning regression models because it has few parameters, high prediction accuracy and can effectively solve the problem of interaction among independent variables (Leroux et al., 2019; Peng et al., 2020).

Accuracy assessment

This study adopted the 10-fold cross-validation method, and it was repeated 20 times to take the average of each accuracy indicator. The coefficient of determination (R^2), BIAS, mean absolute error (MAE), root mean square error (RMSE), and relative root mean square error (RRMSE) were used as the accuracy indicators. Among them, R^2 is used to evaluate the goodness-of-fit of a model, and a larger R^2 indicates a better goodness-of-fit. BIAS is used to evaluate the deviation between the estimation and the reference. If the value is greater than 0, it indicates that the estimated result is overall greater than the reference result, and vice versa. MAE and RMSE are used to evaluate the overall absolute error between the estimation and the reference, and RRMSE is used to evaluate the overall relative error between the estimation and the reference. The smaller the absolute values of these four indicators indicate the smaller error and the higher accuracy. The calculations of the five indicators were shown in Equations (5)-(9), respectively.

$$R^2 = \frac{\text{cov}(\hat{Y}, Y)^2}{\text{var}(\hat{Y})\text{var}(Y)} \quad (5)$$

$$\text{BIAS} = \frac{\sum_{i=1}^N (\hat{Y}_i - Y_i)}{N} \quad (6)$$

$$\text{MAE} = \frac{\sum_{i=1}^N |\hat{Y}_i - Y_i|}{N} \quad (7)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (\hat{Y}_i - Y_i)^2}{N}} \quad (8)$$

$$\text{RRMSE} = \frac{\text{RMSE}}{\bar{Y}} \quad (9)$$

where \hat{Y}_i is the estimated result of the i th sample, Y_i is the corresponding reference value, \bar{Y} is the average of samples, N is the number of samples, $\text{cov}(\hat{Y}, Y)$ represents the covariance between the estimated results and the references, and $\text{var}(\hat{Y})$ and $\text{var}(Y)$ represent the variance of the estimated results and the references, respectively.

In addition, the independent samples t -test was used to test whether there were statistically significant differences ($P < 0.05$) in the accuracies between the phenological piecewise models and the whole-season models. The experimental results being tested were the accuracy indicators (e.g. RMSEs) from 10-fold cross-validation (conducted 20 times). The Kolmogorov-Smirnov (K-S) test was performed in advance to verify that all samples were normally distributed.

Evaluation of the spatiotemporal transferability

In addition to the estimation accuracy, the spatiotemporal transferability of the whole-season models and the phenological piecewise models were also compared. In the evaluation of temporal transferability, the data in 2017 was first selected as training data to build the models, and the data in 2018 was selected as test data to assess the accuracy. Then the data in 2018 was selected as training data, and the data in 2017 was selected as test data. It is to evaluate the ability of models to predict the winter wheat yield in other years whose samples did not participate in modelling (i.e. the temporal transferability). In the evaluation of spatial transferability, the 100 counties were randomly divided into two equal groups, and the two groups also took turns as the training samples to build the models or test samples to assess the accuracy. Correspondingly, it is to evaluate the ability of models to predict the winter wheat yield in other regions whose samples did not participated in modelling (i.e. the spatial transferability). The assessment indicators were R^2 , MAE, RMSE and RRMSE mentioned above.

Results

The optimal combination of variables for different models

For the whole-season model, the R^2 between mean value of EVI (EVI_{mean}) within the growing season and yield was the highest (Figure S2a). After combining the vegetation index with meteorological factors, the R^2 was greatly improved (Figure 2). Among those combinations, the combination of air temperature, soil moisture content and solar radiation performed the best. The R^2 was the highest and all independent

variables passed the significance test ($P < 0.05$). Therefore, the combination of EVI_{mean} , air temperature, soil moisture content and solar radiation was the optimal combination for the whole-season yield model. Similar results were found based on samples from 2017 and 2018 respectively (Figure S3 & Table S2).

For the phenological piecewise model, EVI also performed the best among the four vegetation indices in each phenophase. Specifically, EVI_{max}

was the best choice in the period of regreening-jointing, and EVI_{mean} was the best choice in the periods of jointing-heading and heading-milking maturity (Figure S2b, c and d). However, the importance of meteorological factors was different in different phenophases (Figure 3 & Table 2). For the period of regreening-jointing, EVI_{max} combined with soil moisture content performed the best. The independent variables all passed the $P < 0.001$ significance test and the R^2 (0.42)

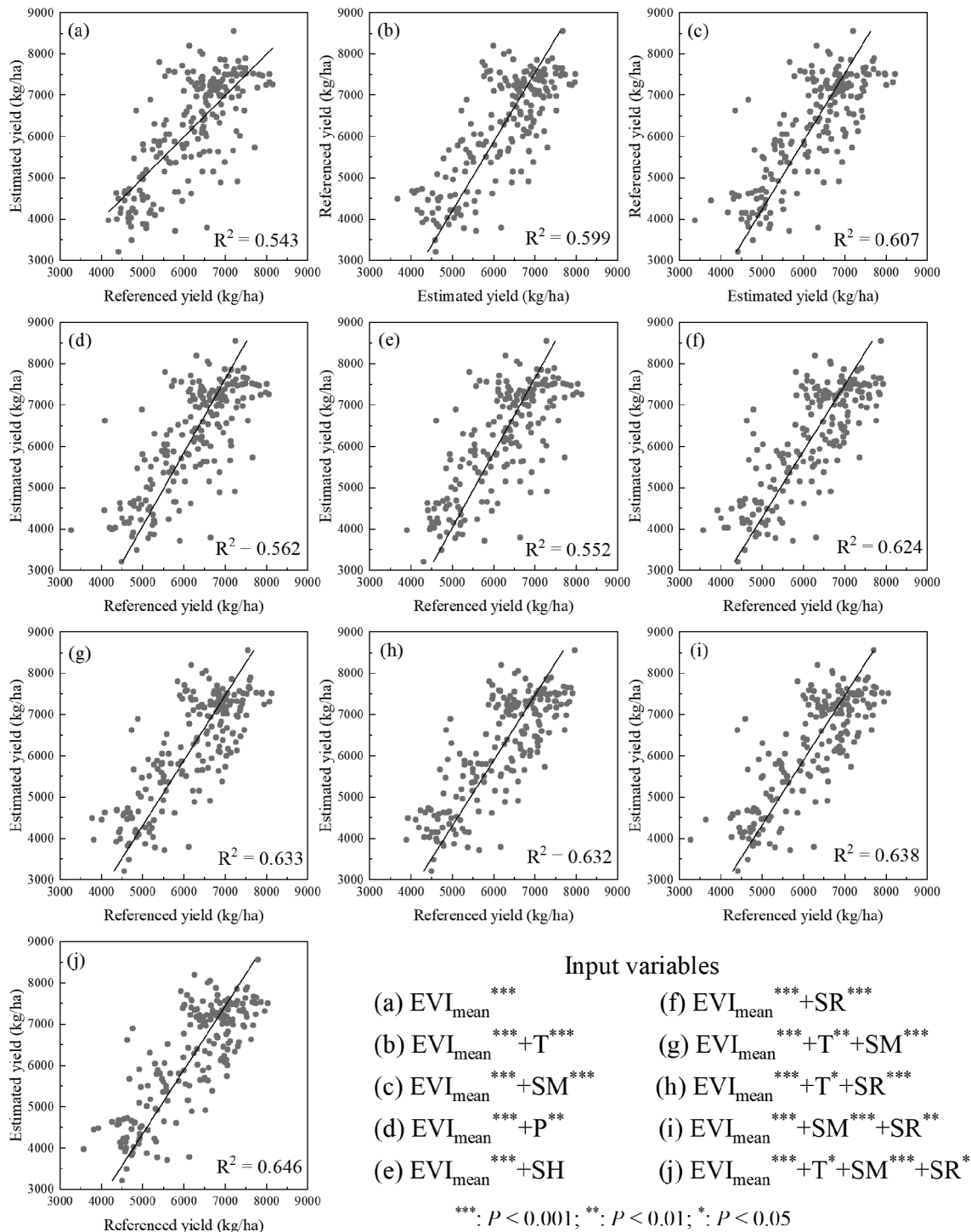


Figure 2. The scatterplots between estimated yield and referenced yield for the multiple linear regression models based on different combinations of vegetation index and meteorological factors within whole-season.

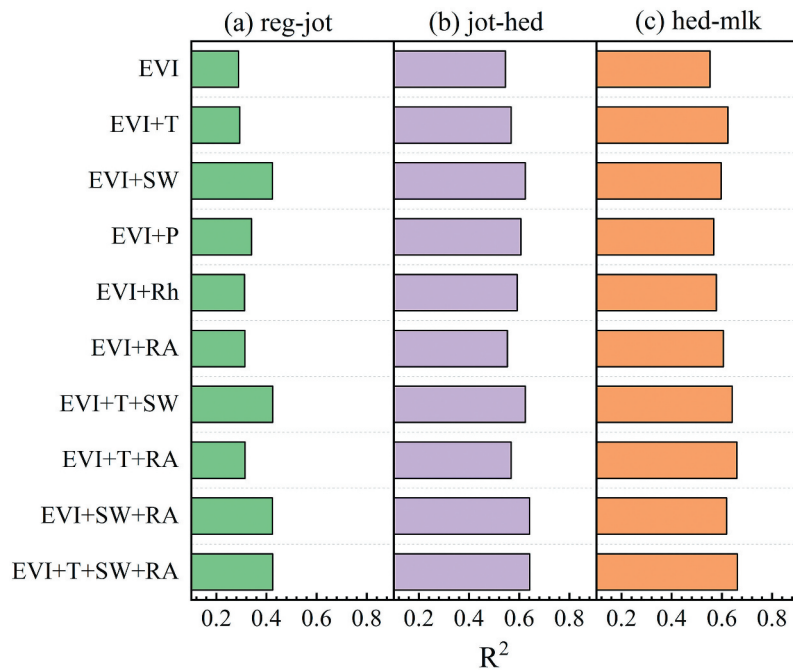


Figure 3. The R^2 of the multiple linear regression models established by different combinations of vegetation index and meteorological factors in different phenophases. (a) regreening-jointing (reg-jot), (b) jointing-heading (jot-hed), and (c) heading-milking maturity (hed-mlk).

Table 2. The significance between independent variables and yield in different phenophases.

Regreening-jointing	Jointing-heading	Heading-milking maturity
EVI_{max}^{***}	EVI_{mean}^{***}	EVI_{mean}^{***}
$EVI_{max}^{***}+T$	$EVI_{mean}^{***}+T^{***}$	$EVI_{mean}^{***}+T^{***}$
$EVI_{max}^{***}+SM^{***}$	$EVI_{mean}^{***}+SM^{***}$	$EVI_{mean}^{***}+SM^{***}$
$EVI_{max}^{***}+P^{***}$	$EVI_{mean}^{***}+P^{***}$	$EVI_{mean}^{***}+P^{**}$
$EVI_{max}^{***}+SH^{**}$	$EVI_{mean}^{***}+SH^{***}$	$EVI_{mean}^{***}+SH^{**}$
$EVI_{max}^{***}+SR^{**}$	$EVI_{mean}^{***}+SR$	$EVI_{mean}^{***}+SR^{***}$
$EVI_{max}^{***}+T+SM^{***}$	$EVI_{mean}^{***}+T+SM^{***}$	$EVI_{mean}^{***}+T^{***}+SM^{**}$
$EVI_{max}^{***}+T+SR$	$EVI_{mean}^{***}+T^{**}+SR$	$EVI_{mean}^{***}+T^{***}+SR^{***}$
$EVI_{max}^{***}+SM^{***}+SR$	$EVI_{mean}^{***}+SM^{***}+SR^{**}$	$EVI_{mean}^{***}+SM^{**}+SR^{***}$
$EVI_{max}^{***}+T+SM^{***}+SR$	$EVI_{mean}^{***}+T+SM^{***}+SR^{**}$	$EVI_{mean}^{***}+T^{***}+SM+SR^{***}$

***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$. Note that the words in bold represent the optimal combination of variables.

was higher than that of other combinations. For the period of jointing-heading, the combination of EVI_{mean} , soil moisture content and solar radiation was the best. For the period of heading-milking maturity, air temperature and solar radiation both significantly contributed to the yield model. Therefore, the combination of EVI_{max} and SM in the period of regreening-jointing, EVI_{mean} , SM and SR in the period of jointing-heading, and EVI_{mean} , T and SR in the period of heading-milking maturity was the optimal combination for the phenological piecewise model. Similar results were found based on samples from 2017 and 2018 respectively (Figure S4, Table S3 & Table S4).

The estimation accuracy for different models

In general, the accuracies of the phenological piecewise models were higher than that of the whole-season models (Figure 4). Except for SVR only with a little improvement (failed to pass $P < 0.05$ significance test), the accuracies of MLR, ANN and RF were all improved significantly (passed $P < 0.05$ significance test). The R^2 increased by 4.3% to 7.6% among the three models, the absolute value of BIAS decreased by 19.9% to 33.3%, the MAE decreased by 7.1% to 8.4%, and the RMSE and the RRMSE decreased by 3.0% to 8.2%. Similar improvements were found based on samples from 2017 and 2018 respectively (Figure S5).

The spatiotemporal transferability for different models

After spatiotemporal transfer, the accuracies of the whole-season models and the phenological piecewise models both decreased to some extent (Figure 5). After temporal (spatial) transfer, the R^2 for the whole-season models and the phenological piecewise models decreased by 0.01 to 0.03 (0.04 to 0.05) and 0.01 to 0.05 (0.03 to 0.05) among four regression models, respectively. The MAE increased by 23 kg/ha to 89 kg/ha (28 kg/ha to 87 kg/ha) and 9 kg/ha to 44 kg/ha (22 kg/ha to 67 kg/ha), respectively. The RMSE increased by 13 kg/ha to 79 kg/ha (23 kg/ha to

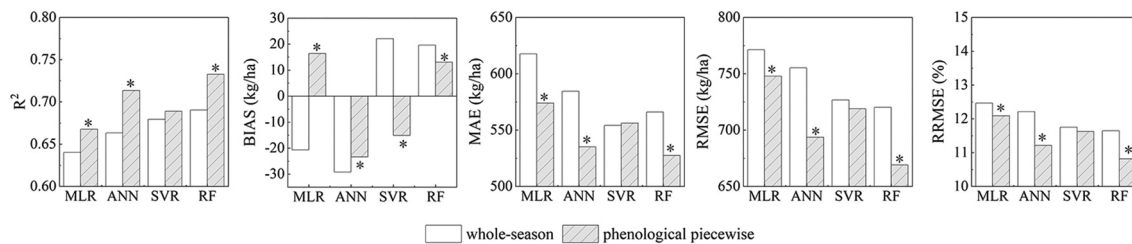


Figure 4. The estimation accuracy for the whole-season models and the phenological piecewise models. * above the column indicates that there is a significant difference ($P < 0.05$) in the accuracy between the whole-season model and the phenological piecewise model based on this regression method.

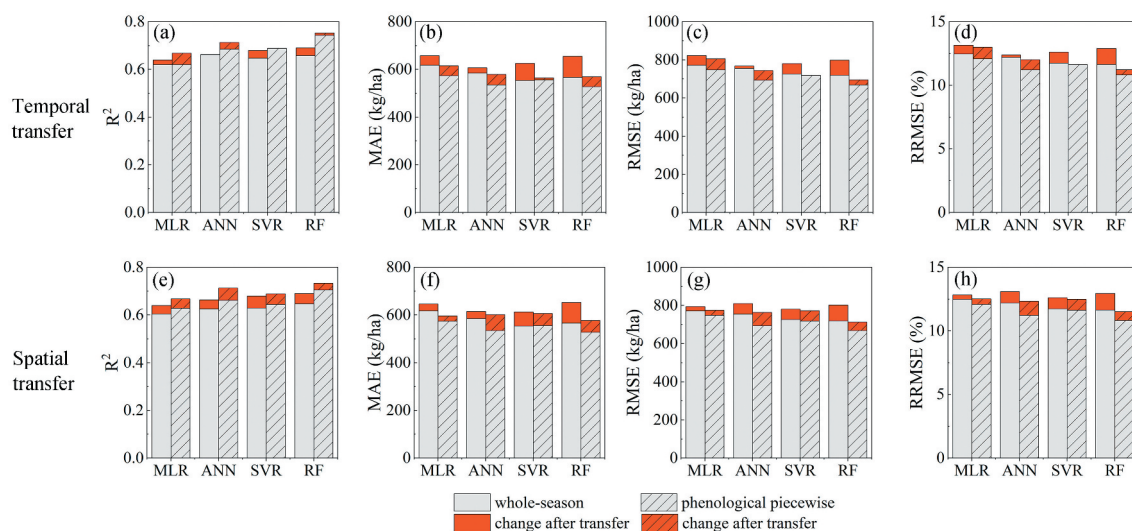


Figure 5. Assessment of the spatiotemporal transferability for the whole-season models and the phenological piecewise models. Note that the change after transfer for R^2 is decreased, while the change after transfer for the other indicators is increased.

81 kg/ha) and 1 kg/ha to 58 kg/ha (27 kg/ha to 70 kg/ha), respectively. The RRMSE increased by 0.2% to 1.3% (0.4% to 1.3%) and 0.0% to 1.0% (0.4% to 1.1%), respectively.

Although the difference for decreased accuracies between the whole-season models and the phenological piecewise models was unobvious and varied depending on the specific regression model, the accuracies after spatiotemporal transfer for all four regression models based on phenological piecewise modelling were still higher than that based on whole-season modelling.

Discussion

The difference for the optimal combination of variables between the whole-season modelling and the phenological piecewise modelling

For winter wheat, temperature, water and sunlight are all important to its growth and development (Xiao et al., 2016). Therefore, the air temperature, soil moisture content, and solar radiation were all selected to build the whole-season models. However, the degrees of importance of these meteorological factors are different in different

phenophases in our study, which may be related to the different demands of crops for hydrothermal conditions in different phenophases and leads to the difference for the optimal combination of variables between the whole-season modelling and the phenological piecewise modelling.

During the period of regreening-jointing, winter wheat needs much water to differentiate its spike primordium, and the soil moisture content should reach 75%-85% of the field capacity (Song et al., 2006). Therefore, the soil moisture content may be the most important factor during this period. The jointing-heading is the period when winter wheat grows the fastest. The study area in this study can generally meet the needs of sunlight and temperature for the growth of winter wheat during this period, while the lack of water is an important factor restricting the growth of winter wheat in this area (Liu et al., 2018). Therefore, the meteorological factors that characterize the water condition show the highest importance during this period in our study. Similarly, Lei et al. (2010) found that the water deficit had a significant effect on the yield of winter wheat especially during the periods of

regreening-jointing and jointing-heading. Heading-milking maturity is a critical period for the maturity of winter wheat. During this period, the requirements for temperature, water and sunlight conditions are very strict. Too high or too low will both have a great negative impact on the growth of winter wheat. In our study, the temperature contributes the most among the meteorological factors during this period. This may be because the period of heading-milking maturity is the most sensitive temperature stage for wheat (Sánchez et al., 2014), and the high temperature stress is one of the most important influence factors (Duncan et al., 2015; Teixeira et al., 2013).

The difference for the accuracy between the whole-season models and the phenological piecewise models

Due to the different optimal combinations of variables, the accuracies for the phenological piecewise models and the whole-season models are different. Compared with the accuracies of whole-season models, the accuracies of phenological piecewise models are higher. This may be explained by the stronger biophysical mechanism of phenological piecewise models. In addition to the different demands of crops for hydrothermal conditions in different phenophases mentioned above, the other reason is that the meteorological factors within the whole growing season cannot reflect the influence of adverse meteorological conditions in a certain phenophase. For example, the range of suitable temperature for the growth of winter wheat is 4–6°C during the period of regreening-jointing, 12–16°C during the period of jointing-heading, and 18–22°C during the period of heading-milking maturity (Gong, 1988; Song et al., 2006). The suitable temperature for the growth of winter wheat varies greatly in different phenophases, so the mean temperature during the whole growing season cannot effectively reflect whether the growth of winter wheat is always in a suitable condition. Therefore, the phenological piecewise modelling is closer to the actual growth process of crops and can achieve a higher accuracy.

The difference for the spatiotemporal transferability between the whole-season models and the phenological piecewise models

In general, there is little difference for spatiotemporal transferability between the whole-season models and the phenological piecewise models. However, the difference for the spatiotemporal transferability between the specific regression models (i.e. MLR, ANN, SVR, RF) is relatively obvious. The spatiotemporal transferability of SVR and RF based on phenological piecewise modelling is stronger than that based on whole-season

modelling, but for MLR and ANN, the results are opposite. Compared with the whole-season modelling, the phenological piecewise modelling needs more independent variables and contains more crop growth information. In this case, the spatiotemporal transferability or generalization ability of MLR and ANN is limited due to their disadvantage of easy overfitting (Karystinos & Pados, 2000; Lawrence & Giles, 2000). In contrast, the spatiotemporal transferability of SVR and RF is stronger by their advantage of applicable to handle high-dimensional data (Belgiu & Dragut, 2016; Chang & Lin, 2011; Maulik & Chakraborty, 2013).

The performance of different vegetation indices and regression methods

Among the four vegetation indices, EVI and NIRv performed the best in winter wheat yield estimation, followed by NDPI and NDVI. This may be related to the characteristics of the vegetation indices. NDVI can effectively reflect the greenness of vegetation, but it is easily saturated and susceptible to atmospheric conditions and soil background. Compared with NDVI, EVI solves these problems to a certain extent and can better reflect the growth status of vegetation. It shows a higher accuracy in yield estimation, which is consistent with the results of Johnson (2016) and Sharifi (2020). NDPI is mainly designed to eliminate the interference of soil background changes on vegetation signals in the early growth phase of vegetation (Wang et al., 2017), and its ability to resist the interference of soil background is stronger than that of NDVI (Chen et al., 2019). However, in the early growth phase of crops, the contribution of vegetation index to yield estimation is limited according to our results, which may explain there is no obvious advantage in yield estimation when comparing NDPI with other vegetation indices. NIRv can reflect the photosynthesis of vegetation because of its high correlation with the sun-induced chlorophyll fluorescence. Therefore, it has great advantages in estimating the gross primary productivity and yield of crops (Peng et al., 2020; Wu et al., 2020).

Among the four regression methods, RF performed the best, with the highest accuracy, significant improvement, and the best spatiotemporal transferability. This is similar to the results of Peng et al. (2020) and Guo et al. (2021). The response function between crop growth rate and meteorological factors generally presents a nonlinear relationship (Soltani & Sinclair, 2012). In this case, the machine learning regression methods, which are more suitable for building nonlinear relationships, perform better in crop yield estimation (Ludwig & Asseng, 2006). Moreover, there are

many factors affecting crop growth, and the machine learning regression methods also perform better in dealing with the problem of the strong interaction between independent variables. In addition, compared with the other two machine learning regression methods (i.e. ANN and SVR), RF can better solve the curse of dimensionality. It requires fewer parameters and is less dependent on training data (Belgiu and Dragut, 2016), which may explain the superiority of RF.

Limitations and prospects

In this study, we aim to compare the difference between the whole-season modelling and the phenological piecewise modelling for winter wheat yield estimation. Except for input variables, other aspects of the two modelling schemes are generally identical. However, if the objective is to estimate crop yield, the uncertainties from data sources (Zhang et al., 2019), parameter settings of models (Rodriguez-Galiano et al., 2015), scale effects (Tarnavsky et al., 2008), or some other factors should be more carefully considered. Moreover, a main potential limitation of this study is the resolution of MODIS data we used. The fields of winter wheat cultivated in the plains in the east of our study area are large and concentrated (Figure 1e & 1f), thus the resolution of MODIS can be well matched. However, in the mountainous regions in the west of our study area, there are a few small and scattered winter wheat fields cultivated (Figure 1d). These fields may be influenced by sub-pixels due to the sparse spatial resolution of MODIS. In this case, the remote sensing data with finer resolution (e.g. Sentinel-2) can be used to solve the problem. In addition, the phenological piecewise modelling requires additional phenology data. Considering the availability of phenology data and the relatively fixed time frame of crop growth, the phenophases distinguished by a fixed time frame may be also effective. Finally, it was found in this study that the phenological piecewise modelling is better than the whole-season modelling for winter wheat yield estimation, while for other crops and at other scales (e.g. the field scale), the performance of phenological piecewise modelling needs to be further evaluated.

Conclusions

In this study, the optimized variables, accuracy and spatiotemporal transferability of the phenological piecewise models were compared with that of the whole-season models. The accuracies of the phenological piecewise models were higher than that of the whole-season models. The difference for spatiotemporal transferability between the phenological piecewise

models and the whole-season models was little and the results varied depending on the specific regression model. However, the accuracies after spatiotemporal transfer for the phenological piecewise models were still higher than that for the whole-season models.

About the optimized variables, the enhanced vegetation index, air temperature, soil moisture content, and solar radiation were selected for the whole-season modelling, while for the phenological piecewise modelling, the optimal combinations of variables in different phenophases were different. It was mainly related to the different demands of crops for hydrothermal conditions in different phenophases. The phenological piecewise model has stronger biophysical mechanism and we highlight the empirical yield estimation models in the future should considered it.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the National Natural Science Foundation of China (NO. 41771047).

References

- Badgley, G., Field, C. B., & Berry, J. A. (2017). Canopy near-infrared reflectance and terrestrial photosynthesis. *Science Advances*, 3(3), e1602244. <https://doi.org/10.1126/sciadv.1602244>
- Belgiu, M., & Dragut, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Bennett, K. P., & Demiriz, A. (1998). Semi-supervised support vector machines. In M.S. Kearns, S.A. Solla, & D. A. Cohn (Eds.), *Advances in neural information processing systems 11* (pp.368–374). MIT Press.
- Briem, G. J., Benediktsson, J. A., & Sveinsson, J. R. (2001). Boosting, bagging, and consensus based classification of multisource remote sensing data. In J. Kittler & F. Roli (Eds.), *Multiple classifier systems* (pp. 279–288). Springer Berlin Heidelberg.
- Carbin, G. W., Tippet, M. K., Lillo, S. P., & Brooks, H. E. (2016). Visualizing long-range severe thunderstorm environment guidance from CFSv2. *Bulletin of the American Meteorological Society*, 97(6), 1021–1031. <https://doi.org/10.1175/BAMS-D-14-00136.1>
- Challinor, A. J., Wheeler, T. R., Craufurd, P. Q., Slingo, J. M., & Grimes, D. (2004). Design and optimisation of a large-area process-based model for annual crops. *Agricultural and Forest Meteorology*, 124(1–2), 99–120. <https://doi.org/10.1016/j.agrformet.2004.01.002>
- Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 1–27. <https://doi.org/10.1145/1961189.1961199>

- Chen, X. H., Guo, Z. F., Chen, J., Yang, W., Yao, Y. M., Zhang, C. S., Cui, X. H., & Cao, X. (2019). Replacing the red band with the red-swir band (0.74(red)+0.26(swir)) can reduce the sensitivity of vegetation indices to soil background. *Remote Sensing*, 11(7), 851. <https://doi.org/10.3390/rs11070851>
- Doraiswamy, P. C., Moulin, S., Cook, P. W., & Stern, A. (2003). Crop yield assessment from remote sensing. *Photogrammetric Engineering and Remote Sensing*, 69(6), 665–674. <https://doi.org/10.14358/PERS.69.6.665>
- Duncan, J. M. A., Dash, J., & Atkinson, P. M. (2015). Elucidating the impact of temperature variability and extremes on cereal croplands through remote sensing. *Global Change Biology*, 21(4), 1541–1551. <https://doi.org/10.1111/gcb.12660>
- Fang, H. L., Liang, S. L., & Hoogenboom, G. (2011). Integration of MODIS LAI and vegetation index products with the CSM-CERES-Maize model for corn yield estimation. *International Journal of Remote Sensing*, 32(4), 1039–1065. <https://doi.org/10.1080/01431160903505310>
- Funk, C., & Budde, M. E. (2009). Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe. *Remote Sensing of Environment*, 113(1), 115–125. <https://doi.org/10.1016/j.rse.2008.08.015>
- Gan, L., Cao, X., Chen, X., Dong, Q., Cui, X., & Chen, J. (2020). Comparison of MODIS-based vegetation indices and methods for winter wheat green-up date detection in Huanghuai region of China. *Agricultural and Forest Meteorology*, 288, 108019. <https://doi.org/10.1016/j.agrfor.2020.108019>
- Godfray, H. C. J., Beddington, J. R., Crute, I. R., Haddad, L., Lawrence, D., Muir, J. F., Pretty, J., Robinson, S., Thomas, S. M., & Toulmin, C. (2010). Food security: The challenge of feeding 9 billion people. *Science*, 327(5967), 812–818. <https://doi.org/10.1126/science.1185383>
- Gong, S. X. (1988). *Grain crops and meteorology*. Beijing Agricultural University Press.
- Gong, P., Liu, H., Zhang, M., Li, C., Wang, J., Huang, H., Clinton, N., Ji, L., Li, W., Bai, Y., Chen, B., Xu, B., Zhu, Z., Yuan, C., Suen, H. P., Guo, J., Xu, N., Li, W., Zhao, Y., Song, L. (2019). Stable classification with limited sample: Transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. *Science Bulletin*, 64(6), 370–373. <https://doi.org/10.1016/j.scib.2019.03.002>
- Guo, Y. H., Fu, Y. S., Hao, F. H., Zhang, X., Wu, W. X., Jin, X. L., Bryant, C. R., & Senthilnath, J. (2021). Integrated phenology and climate in rice yields prediction using machine learning methods. *Ecological Indicators*, 120, 106935. <https://doi.org/10.1016/j.ecolind.2020.106935>
- He, L., & Mostovoy, G. (2019). Cotton yield estimate using sentinel-2 data and an ecosystem model over the southern US. *Remote Sensing*, 11(17), 2000. <https://doi.org/10.3390/rs11172000>
- Henan Statistics Bureau. (2019). *Henan statistical yearbook*. China Statistics Press.
- Huang, J. X., Gomez-Dans, J. L., Huang, H., Ma, H. Y., Wu, Q. L., Lewis, P. E., Liang, S. L., Chen, Z. X., Xue, J. H., Wu, Y. T., Zhao, F., Wang, J., & Xie, X. H. (2019). Assimilation of remote sensing into crop growth models: Current status and perspectives. *Agricultural and Forest Meteorology*, 276(277), 107609. <https://doi.org/10.1016/j.agrformet.2019.06.008>
- Huang, X., Zhu, W., Wang, X., Zhan, P., Liu, Q., Li, X., & Sun, L. (2020). A method for monitoring and forecasting the heading and flowering dates of winter wheat combining satellite-derived green-up dates and accumulated temperature. *Remote Sensing*, 12(21), 3536. <https://doi.org/10.3390/rs12213536>
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*, 83(1–2), 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- Hunt, M. L., Blackburn, G. A., Carrasco, L., Redhead, J. W., & Rowland, C. S. (2019). High resolution wheat yield mapping using Sentinel-2. *Remote Sensing of Environment*, 233, 111410. <https://doi.org/10.1016/j.rse.2019.111410>
- Jego, G., Pattey, E., & Liu, J. G. (2012). Using leaf area index, retrieved from optical imagery, in the STICS crop model for predicting yield and biomass of field crops. *Field Crops Research*, 131, 63–74. <https://doi.org/10.1016/j.fcr.2012.02.012>
- Ji, Z., Pan, Y., Zhu, X., Wang, J., & Li, Q. (2021). Prediction of crop yield using phenological information extracted from remote sensing vegetation index. *Sensors*, 21(4), 1406. <https://doi.org/10.3390/s21041406>
- Johnson, D. M. (2016). A comprehensive assessment of the correlations between field crop yields and commonly used MODIS products. *International Journal of Applied Earth Observation and Geoinformation*, 52, 65–81. <https://doi.org/10.1016/j.jag.2016.05.010>
- Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., Wilkens, P. W., Singh, U., Gijsman, A. J., & Ritchie, J. T. (2003). The DSSAT cropping system model. *European Journal of Agronomy*, 18(3–4), 235–265. [https://doi.org/10.1016/S1161-0301\(02\)00107-7](https://doi.org/10.1016/S1161-0301(02)00107-7)
- Kamir, E., Waldner, F., & Hochman, Z. (2020). Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods. *ISPRS Journal of Photogrammetry and Remote Sensing*, 160, 124–135. <https://doi.org/10.1016/j.isprsjprs.2019.11.008>
- Karystinos, G. N., & Pados, D. A. (2000). On overfitting, generalization, and randomly expanded training sets. *IEEE Transactions on Neural Networks*, 11(5), 1050–1057. <https://doi.org/10.1109/72.870038>
- Kowalik, W., Dabrowska-Zielinska, K., Meroni, M., Raczka, T. U., & de Wit, A. (2014). Yield estimation using SPOT-VEGETATION products: A case study of wheat in European countries. *International Journal of Applied Earth Observation and Geoinformation*, 32, 228–239. <https://doi.org/10.1016/j.jag.2014.03.011>
- Lawrence, S., & Giles, C. L. (2000). Overfitting and neural networks: Conjugate gradient and backpropagation. *IJCNN 2000: Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, 1, 114–119.
- Lee, V. W., Kim, C., Chhugani, J., Deisher, M., Kim, D., Nguyen, A. D., Satish, N., Smelyanskiy, M., Chennupaty, S., Hammarlund, P., Singhal, R., & Dubey, P. (2010). Debunking the 100X GPU vs. CPU myth: An evaluation of throughput computing on CPU and GPU. *Proceedings of the 37th annual international symposium on Computer architecture*, 37, 451–460.
- Lei, Y., Zhang, F., Kou, W., & Feng, L. (2010). Effect of water deficit at different growth stages and nitrogen fertilizer on yield and water use efficiency of winter wheat. *Journal of Northwest A&F University (Natural Science Edition)*, 38(5), 167–174. <https://doi.org/10.13207/j.cnki.jnwafu.2010.05.017>
- Leroux, L., Castets, M., Baron, C., Escorihuela, M. J., Begue, A., & Lo Seen, D. (2019). Maize yield estimation in West Africa from crop process-induced combinations

- of multi-domain remote sensing indices. *European Journal of Agronomy*, 108, 11–26. <https://doi.org/10.1016/j.eja.2019.04.007>
- Liao, C. H., Wang, J. F., Dong, T. F., Shang, J. L., Liu, J. G., & Song, Y. (2019). Using spatio-temporal fusion of Landsat-8 and MODIS data to derive phenology, biomass and yield estimates for corn and soybean. *Science of the Total Environment*, 650, 1707–1721. <https://doi.org/10.1016/j.scitotenv.2018.09.308>
- Liu, X., He, X., & You, S. (2018). Characteristics of agricultural meteorological disasters in North China from 1986 to 2015. *Agricultural Research in the Arid Areas*, 36(5), 257–263. <https://doi.org/10.7606/j.issn.1000-7601.2018.05.36>
- Ludwig, F., & Asseng, S. (2006). Climate change impacts on wheat production in a Mediterranean environment in Western Australia. *Agricultural Systems*, 90(1–3), 159–179. <https://doi.org/10.1016/j.agsy.2005.12.002>
- Maulik, U., & Chakraborty, D. (2013). Learning with transductive SVM for semisupervised pixel classification of remote sensing imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 77, 66–78. <https://doi.org/10.1016/j.isprsjprs.2012.12.003>
- Mkhabela, M. S., Bullock, P., Raj, S., Wang, S., & Yang, Y. (2011). Crop yield forecasting on the Canadian Prairies using MODIS NDVI data. *Agricultural and Forest Meteorology*, 151(3), 385–393. <https://doi.org/10.1016/j.agrformet.2010.11.012>
- National Bureau of Statistics of China. (2019). *China rural statistical yearbook*. China Statistics Press.
- Nolasco, M., Ovando, G., Sayago, S., Magario, I., & Bocco, M. (2021). Estimating soybean yield using time series of anomalies in vegetation indices from MODIS. *International Journal of Remote Sensing*, 42(2), 405–421. <https://doi.org/10.1080/01431161.2020.1809736>
- Pegion, K., & Alexander, M. (2013). The seasonal footprinting mechanism in CFSv2: Simulation and impact on ENSO prediction. *Climate Dynamics*, 41(5–6), 1671–1683. <https://doi.org/10.1007/s00382-013-1887-5>
- Peng, B., Guan, K. Y., Zhou, W., Jiang, C. Y., Frankenberg, C., Sun, Y., He, L. Y., & Kohler, P. (2020). Assessing the benefit of satellite-based solar-induced chlorophyll fluorescence in crop yield prediction. *International Journal of Applied Earth Observation and Geoinformation*, 90, 102126. <https://doi.org/10.1016/j.jag.2020.102126>
- Prasad, A. K., Chai, L., Singh, R. P., & Kafatos, M. (2006). Crop yield estimation model for Iowa using remote sensing and surface parameters. *International Journal of Applied Earth Observation and Geoinformation*, 8(1), 26–33. <https://doi.org/10.1016/j.jag.2005.06.002>
- Rasmussen, M. S. (1997). Operational yield forecast using AVHRR NDVI data: Reduction of environmental and inter-annual variability. *International Journal of Remote Sensing*, 18(5), 1059–1077. <https://doi.org/10.1080/014311697218575>
- Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., & Chica-Rivas, M. (2015). Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geology Reviews*, 71, 804–818. <https://doi.org/10.1016/j.oregeorev.2015.01.001>
- Rounsevell, M. D. A., Ewert, F., Reginster, I., Leemans, R., & Carter, T. R. (2005). Future scenarios of European agricultural land use II. Projecting changes in cropland and grassland. *Agriculture Ecosystems & Environment*, 107(2–3), 117–135. <https://doi.org/10.1016/j.agee.2004.12.002>
- Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., & Harlan, J. C. (1974). Texas A & M University, Remote Sensing Center. *Monitoring the vernal advancement and retrogradation (greenwave effect) of natural vegetation. NASA/GSFC Type III Final Report 371*.
- Saha, S., Moorthi, S., Wu, X. R., Wang, J., Nadiga, S., Tripp, P., Behringer, D., Hou, Y. T., Chuang, H. Y., Iredell, M., Ek, M., Meng, J., Yang, R. Q., Mendez, M. P., Van Den Dool, H., Zhang, Q., Wang, W. Q., Chen, M. Y., & Becker, E. (2014). The NCEP climate forecast system version 2. *Journal of Climate*, 27(6), 2185–2208. <https://doi.org/10.1175/JCLI-D-12-00823.1>
- Sánchez, B., Rasmussen, A., & Porter, J. R. (2014). Temperatures and the growth and development of maize and rice: A review. *Global Change Biology*, 20(2), 408–417. <https://doi.org/10.1111/gcb.12389>
- Saseendran, S. A., Singh, K. K., Rathore, L. S., Singh, S. V., & Sinha, S. K. (2000). Effects of climate change on rice production in the tropical humid climate of Kerala, India. *Climatic Change*, 44(4), 495–514. <https://doi.org/10.1023/A:1005542414134>
- Searchinger, T., Heimlich, R., Houghton, R. A., Dong, F. X., Elobeid, A., Fabiosa, J., Tokgoz, S., Hayes, D., & Yu, T. H. (2008). Use of US croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science*, 319(5867), 1238–1240. <https://doi.org/10.1126/science.1151861>
- Sharifi, A. (2020). Yield prediction with machine learning algorithms and satellite images. *Journal of the Science of Food and Agriculture*, 101(3), 891–896. <https://doi.org/10.1002/jsfa.10696>
- Siyal, A. A., Dempewolf, J., & Becker-Reshef, I. (2015). Rice yield estimation using Landsat ETM plus data. *Journal of Applied Remote Sensing*, 9(1), 095986. <https://doi.org/10.1117/1.JRS.9.095986>
- Soltani, A., & Sinclair, T. R. (2012). *Modeling physiology of crop development, growth and yield*. CAB International.
- Song, Y. B., Wang, J. L., & Yang, F. Y. (2006). *Meteorological services for food security*. China Meteorological Press.
- Sousa, S. I. V., Martins, F. G., Alvim-Ferraz, M. C. M., & Pereira, M. C. (2007). Multiple linear regression and artificial neural networks based on principal components to predict ozone concentrations. *Environmental Modelling & Software*, 22(1), 97–103. <https://doi.org/10.1016/j.envsoft.2005.12.002>
- Stepanov, A., Dubrovin, K., Sorokin, A., & Aseeva, T. (2020). Predicting soybean yield at the regional scale using remote sensing and climatic data. *Remote Sensing*, 12(12), 1936. <https://doi.org/10.3390/rs12121936>
- Sun, J., Lai, Z. L., Di, L. P., Sun, Z. H., Tao, J. B., & Shen, Y. L. (2020). Multilevel deep learning network for county-level corn yield estimation in the US corn belt. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13, 5048–5060. <https://doi.org/10.1109/JSTARS.2020.3019046>
- Suykens, J. A. K., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters*, 9(3), 293–300. <https://doi.org/10.1023/A:1018628609742>
- Tarnavsky, E., Garrigues, S., & Brown, M. E. (2008). Multiscale geostatistical analysis of AVHRR, SPOT-VGT, and MODIS global NDVI products. *Remote Sensing of Environment*, 112(2), 535–549. <https://doi.org/10.1016/j.rse.2007.05.008>
- Teixeira, E. I., Fischer, G., van Velthuizen, H., Walter, C., & Ewert, F. (2013). Global hot-spots of heat stress on agricultural crops due to climate change. *Agricultural and Forest Meteorology*, 170, 206–215. <https://doi.org/10.1016/j.agrformet.2011.09.002>

- Thenkabail, P. S., Knox, J. W., Ozdogan, M., Gumma, M. K., Congalton, R. G., Wu, Z. T., Milesi, C., Finkral, A., Marshall, M., Mariotto, I., You, S. C., Giri, C., & Nagler, P. (2012). Assessing future risks to agricultural productivity, water resources and food security: How can remote sensing help? *Photogrammetric Engineering and Remote Sensing*, 78(8), 773–782.
- Tsai, C. P., & Lee, T. L. (1999). Back-propagation neural network in tidal-level forecasting. *Journal of Waterway Port Coastal and Ocean Engineering-Asce*, 125(4), 195–202. [https://doi.org/10.1061/\(ASCE\)0733-950X\(1999\)125:4\(195\)](https://doi.org/10.1061/(ASCE)0733-950X(1999)125:4(195))
- Tuvdendorj, B., Wu, B. F., Zeng, H. W., Batdelger, G., & Nanzad, L. (2019). Determination of appropriate remote sensing indices for spring wheat yield estimation in Mongolia. *Remote Sensing*, 11(21), 2568. <https://doi.org/10.3390/rs11212568>
- Vermote, E. (2015). MOD09A1 MODIS/Terra surface reflectance 8-day L3 global 500m sin grid v006 [data set]. NASA EOSDIS Land Processes DAAC. <https://doi.org/10.5067/MODIS/MOD09A1.006>
- Wang, C., Chen, J., Wu, J., Tang, Y. H., Shi, P. J., Black, T. A., & Zhu, K. (2017). A snow-free vegetation index for improved monitoring of vegetation spring green-up date in deciduous ecosystems. *Remote Sensing of Environment*, 196, 1–12. <https://doi.org/10.1016/j.rse.2017.04.031>
- White, M. A., Thornton, P. E., & Running, S. W. (1997). A continental phenology model for monitoring vegetation responses to interannual climatic variability. *Global Biogeochemical Cycles*, 11(2), 217–234. <https://doi.org/10.1029/97GB00330>
- Wu, G. H., Guan, K. Y., Jiang, C. Y., Peng, B., Kimm, H., Chen, M., Yang, X., Wang, S., Suyker, A. E., Bernacchi, C. J., Moore, C. E., Zeng, Y. L., Berry, J. A., & Cendrero-Mateo, M. P. (2020). Radiance-based NIRv as a proxy for GPP of corn and soybean. *Environmental Research Letters*, 15(3), 034009. <https://doi.org/10.1088/1748-9326/ab65cc>
- Xiao, D. P., Shen, Y. J., Zhang, H., Moiwo, J. P., Qi, Y. Q., Wang, R. D., Pei, H. W., Zhang, Y. C., & Shen, H. T. (2016). Comparison of winter wheat yield sensitivity to climate variables under irrigated and rain-fed conditions. *Frontiers of Earth Science*, 10(3), 444–454. <https://doi.org/10.1007/s11707-015-0534-3>
- Zhang, Y. S., Chipanshi, A., Daneshfar, B., Koiter, L., Champagne, C., Davidson, A., Reichert, G., & Bedard, F. (2019). Effect of using crop specific masks on earth observation based crop yield forecasting across Canada. *Remote Sensing Applications-Society and Environment*, 13, 121–137. <https://doi.org/10.1016/j.rsase.2018.10.002>
- Zhou, W. Y. (1999). Verification of the nonparametric characteristics of backpropagation neural networks for image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2), 771–779. <https://doi.org/10.1109/36.752193>
- Zhu, W., Pan, Y., He, H., Wang, L., Mou, M., & Liu, J. (2012). A changing-weight filter method for reconstructing a high-quality NDVI time series to preserve the integrity of vegetation phenology. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4), 1085–1094. <https://doi.org/10.1109/TGRS.2011.2166965>

Appendix A Determination of the duration of phenophases

The fixed durations from regreening to jointing and heading to milking maturity were obtained from the observation data of 35 winter wheat phenology stations in Henan Province (Figure 1) from 2017 to 2018. The phenology stations were established by the China Meteorological Administration (CMA). The stations monitor the growth and development status of field crops through direct human observation. The specific records include the crop variety, the name of the crop development stage and its date, the anomaly of the development stage, etc.

A total of 61 (recording both regreening and jointing dates) and 63 (recording both heading and milking maturity dates) station-year samples were used. Considering the days from regreening to jointing and heading to milking maturity recorded by most stations did not vary greatly (the differences between the upper quartiles and the lower quartiles in Figure A1 were within 10 days), the median value of samples was taken as the corresponding fixed duration (i.e. 26 days for the duration from regreening to jointing and 34 days for the duration from heading to milking maturity).

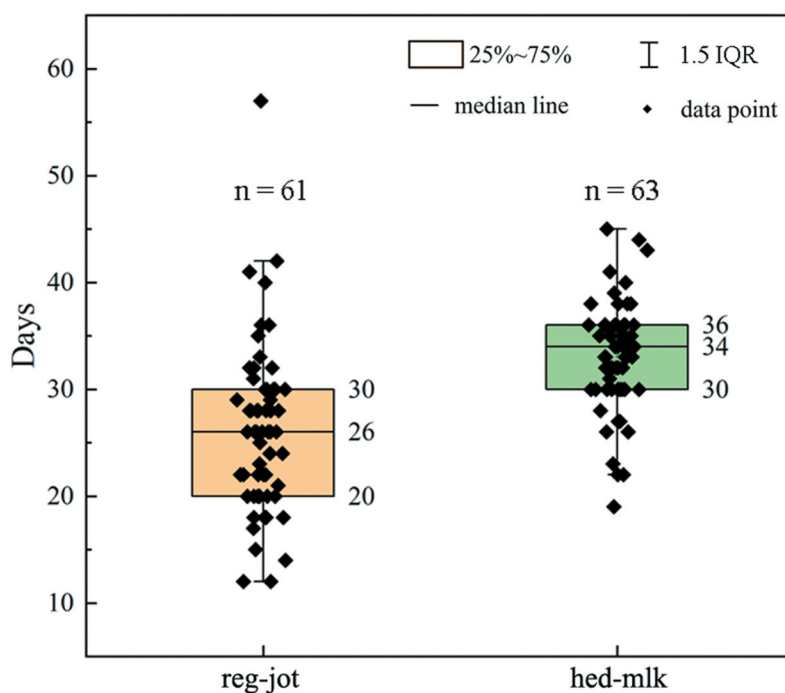


Figure A1. The days from regreening to jointing (reg-jot) and heading to milking maturity (hed-mlk) recorded by phenology stations from 2017 to 2018. n: the number of samples.