

RESEARCH

Open Access



# Nondestructive estimation of potato yield using relative variables derived from multi-period LAI and hyperspectral data based on weighted growth stage

Shanjun Luo<sup>1,2</sup>, Yingbin He<sup>1\*</sup> , Qian Li<sup>3</sup>, Weihua Jiao<sup>4</sup>, Yaqiu Zhu<sup>3</sup> and Xihai Zhao<sup>1</sup>

## Abstract

**Background:** The accurate estimation of potato yield at regional scales is crucial for food security, precision agriculture, and agricultural sustainable development.

**Methods:** In this study, we developed a new method using multi-period relative vegetation indices (rVIs) and relative leaf area index (rLAI) data to improve the accuracy of potato yield estimation based on the weighted growth stage. Two experiments of field and greenhouse (water and nitrogen fertilizer experiments) in 2018 were performed to obtain the spectra and LAI data of the whole growth stage of potato. Then the weighted growth stage was determined by three weighting methods (improved analytic hierarchy process method, IAHP; entropy weight method, EW; and optimal combination weighting method, OCW) and the Slogistic model. A comparison of the estimation performance of rVI-based and rLAI-based models with a single and weighted stage was completed.

**Results:** The results showed that among the six test rVIs, the relative red edge chlorophyll index ( $rCI_{red\ edge}$ ) was the optimal index of the single-stage estimation models with the correlation with potato yield. The most suitable single stage for potato yield estimation was the tuber expansion stage. For weighted growth stage models, the OCW-LAI model was determined as the best one to accurately predict the potato yield with an adjusted  $R^2$  value of 0.8333, and the estimation error about 8%.

**Conclusion:** This study emphasizes the importance of inconsistent contributions of multi-period or different types of data to the results when they are used together, and the weights need to be considered.

**Keywords:** Yield estimation, Remote sensing, Potato, Relative variables, Slogistic model, Weighted growth stage

## Background

Potato (*Solanum tuberosum* L.), a mixed grain, forage, and vegetable crop [1], is the fourth most important crop in the world [2, 3]. Since the launch of the potato staple food strategy in 2015 in China, potato has become another major staple food crop after rice, wheat, and

corn [4]. Timely forecasting potato yield data is a vital reference index for variety breeding determined by the combination of genes and growth environment [5]. The accurate prediction of potato yield, especially at the regional level, is of great significance for ensuring food security and promoting the sustainable development of agriculture, which is related to the formulation of major policies and guidelines of the national economy and people's livelihood.

The method of crop growth models (CGM), costly, time-consuming, and not always accurate, is often used

\*Correspondence: heyingbin@caas.cn

<sup>1</sup> Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China  
Full list of author information is available at the end of the article



© The Author(s) 2020. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

in conventional yield estimation, which relies on a large amount of data collection [6, 7]. It is reported that there are approximately 32 types of CGM combining multiple data sources and methods to monitor the potato yield under conditions of water, nitrogen fertilizer, and CO<sub>2</sub> atmospheric levels [8]. However, the difficulty of obtaining large amounts of input data is one of the major limitations of the widespread employment of models due to their complexity [9–11]. Furthermore, field investigation, another traditional method, is a destructive estimation way. Although the accuracy of the final results can be guaranteed by comprehensive surveys, it is undoubtedly a laborious and time-intensive work [12, 13].

Remote estimation of yield is an approach to establish the relationship between crop spectra and yield data [14]. Remote sensing (RS), an emerging technique, can be used to effectively obtain spectral data of vegetation canopy from space in a non-destructive manner, which carries much valuable information indicating the interaction between canopy and solar radiation such as vegetation absorption and scattering [15]. Vegetation canopy spectrum is closely related to crop growth, especially the visible range affected by pigment and the near-infrared (NIR) bands affected by cell tissue and canopy structure [16, 17]. Therefore, the vegetation index (VI) calculated by these bands has been widely used for the monitoring and estimation of vegetation characteristic parameters, such as leaf area index (LAI) [18], biomass [19], chlorophyll content [20], nitrogen content and carbon content [21], and achieved high accuracy. In addition, various VIs showed great differences when applied in diverse scenarios. For example, when the fractional vegetation cover (FVC) is over 50%, the ratio vegetation index (RVI) has a high sensitivity for vegetation [22]. The normalized difference vegetation index (NDVI) is commonly used to research the vegetation growth and distinguish vegetation from non-vegetation with eliminating most of the radiation errors, but it is prone to saturation [23]. Not only that, VIs also have many applications in the yield estimation of different crops on account of the sensitivity to plant photosynthesis. Gong et al. [24] found that NDVI is of great help to the prediction of rapeseed yield using unmanned aerial vehicle (UAV) imagery. Moreover, VI also contributes significantly to yield estimation for crops such as rice [25, 26], maize [27, 28], and wheat [29, 30]. The simulation results of crop characteristic parameters can be obtained by constructing the linear or nonlinear empirical relationship [31] or by machine learning methods [32] like support vector machine (SVM), random forest (RF), partial least squares (PLS) and artificial neural network (ANN) between VIs and these parameters. So far, the VI-based parameter statistics is the simplest and most widely studied estimation method, which has been

extensively applied in crop growth monitoring [33]. And the crop growth status monitored by RS directly determines the final crop yield. Hence, remote estimation of crop yield based on VI exhibits good potential, especially in a large-scale domain of estimation scenarios [34].

LAI is one of the vital parameters of crop canopy structure that related to photosynthesis, respiration, and transpiration [35]. Peng et al. [36] proved that LAI can be applied to estimate yield in oilseed rape using UAV data with the estimation error below 15%. Liu et al. [37] calculated the canopy density (Chl) using LAI and then constructed the simple linear prediction model of rice yield with an R<sup>2</sup> value of 0.81. Therefore, LAI can be determined for yield estimation.

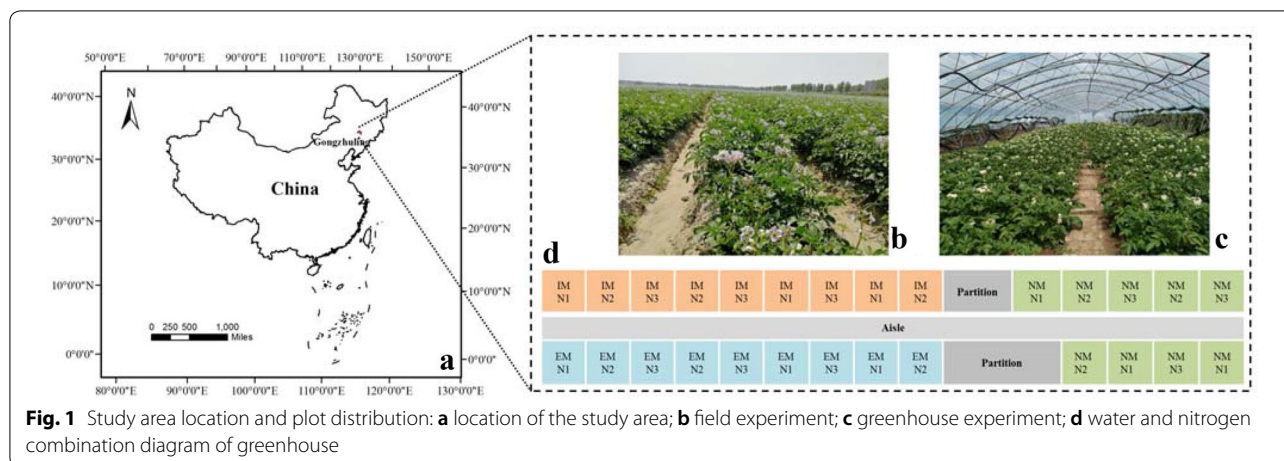
Employing RS technique including UAV, satellite, and ground measurement, massive data of multiple time series can be obtained. However, there are still some issues worthy of our attention. The atmospheric environment, soil background, and solar radiation conditions all will change during the process of obtaining data for many times [38]. Actually, eliminating the interference caused by illumination, aerosol, and background environment among multi-period data is a prerequisite for accurate yield estimation. For example, some reference whiteboards can be used for radiometric correction of remotely sensed images, but it is still difficult to obtain absolutely accurate data [39]. Therefore, we try to use the method of relative variables by subtraction to reduce the differences of data caused by the external environment.

In our experiment, the whole-stage canopy spectra of potato field were remotely measured from ground platforms, which had the advantage of reflecting the field variations well. Meanwhile, the LAI data in the same period were obtained. With potato grown under different water and nitrogen fertilizer treatments, our objectives are (1) to determine the optimal VI for single-stage yield estimation of potato; (2) to determine the optimal single stage for potato yield prediction, and (3) to compare the performance of rVI-based and rLAI-based models using single-stage and weighted stage data, and determine the final potato yield prediction model.

## Methods

### Study area

The experiment area (Fig. 1a) was located at the experimental base of Economic Plants Research Institute (43.45°N, 124.99°E), Jilin Academy of Agricultural Sciences, Gongzhuling City, Jilin Province, China. The greenhouse experiment (Fig. 1c) was conducted in May–September 2018. Shepody [40], a widely planted potato variety in Jilin Province, was selected as the experimental object. Potatoes were sown on May 2nd and harvested on September 10th including the whole growth stages.



Through the combination of nitrogen fertilizer and water, 27 plots (Fig. 1d) including three nitrogen levels (N1: half of the normal nitrogen fertilizer, N2: normal nitrogen fertilizer, and N3: two times normal nitrogen fertilizer) and three water levels (EM: excessive moisture, NM: normal moisture, and IM: insufficient moisture) were set up. The water–nitrogen combination experiment was divided into 9 treatments, and each treatment was repeated 3 times randomly. To ensure that there is no water interference between the treatments, two partitions between IM and NM, and three partitions between EM and NM were set. The same potato variety was planted in the field experiment (Fig. 1b) to avoid the influence of sampling on the greenhouse experiment. The field experiment was used to study the change of dry weight with time to simulate the growth of potato, while the greenhouse experiment was applied to estimate potato yield by measuring hyperspectral and LAI.

The experimental area was located in the middle of the Songliao Plain, with a temperate continental monsoon climate, an average temperature from May to August of 18–20 °C, and abundant natural resources. It is a key commodity grain base in China and a demonstration area for potato cultivation.

**Data collection and multi-period data processing**

The collection of data covered five key stages of potato growth: seeding stage (SS), tuber formation stage (TFS), tuber expansion stage (TES), starch accumulation stage (SAS), and harvest stage (HS). The field data, including LAI and hyperspectral, were collected for five times from SS (14 June), TFS (28 June), TES (23 July), SAS (9 August), to HS (27 August).

The SUNSCAN Canopy Analysis System (Delta-T Devices, Ltd., Burwell, Cambridge, UK) [41] was used to acquire the potato LAI data under conditions of

windless and stable light. Since potatoes were planted following the ridges, our measurements were made 5 times parallel to the ridges and perpendicular to the ridges, respectively. Five different places were selected for measurement in each plot, and the mean values of 25 measurements in total were taken as canopy LAI values of the plot.

The USB 2000 spectrometer (Ocean Optics, Inc., Dunedin, Florida, United States) [42] was adopted to collect potato canopy hyperspectral under cloudless and windless conditions, with a spectral sampling interval of 0.46 nm. The spectral measurement was performed daily from 10:00 to 14:00 with the field-of-view angle of 25°, the probe vertically downward and about 1 m away from the top of the potato canopy. The observation was repeated five times for each plot, and the average value was regarded as the canopy spectral reflection. The reference whiteboard (chemical composition is BaSO<sub>4</sub>) was used for relative radiometric correction prior to measurement.

Dry weight measurements of potato plants (including stems, leaves, roots, flowers, etc.) were conducted by destructive sampling. In each growth stage, the sampling interval is 3–6 days. Ten points were randomly selected for each measurement. The collected plants were dried in the laboratory after drying in the field until their weights remained unchanged when weighing again. The average value was taken as the dry weight data of this measurement. In total, 17 times of sampling were taken on June 14 (SS), June 22, June 25, June 28, July 1 (TFS), July 9, July 13, July 16, July 19, July 23 (TES), July 31, August 3, August 6, August 9, August 16, August 21 (SAS), and August 27 (HS), respectively.

At HS, the potatoes in all plots were harvested manually. Then plot-level potatoes were weighed immediately.

For LAI and VI of multiple periods, the utilization of relative VI (rVI) and relative LAI (rLAI) is expected to reduce the limitation of uncertain information about background, light and atmospheric conditions at different growth stages. Firstly, plot-level rVI and rLAI were proposed under the premise of the hypothesis that solar radiation, atmospheric conditions, and field background were similar at each data acquisition. A standard plot can then be selected as a reference to help diminish the difference caused by time. In this study, rVI, rLAI, and relative yield were calculated based on a reference of an appropriate plot. The calculation of rVI, rLAI, and relative yield was carried out through the differences of VI, LAI, and yield between the study plot and reference plot (Eqs. 1–3). The method of eliminating the influence of external factors by subtraction can keep the correlation between original data unchanged.

$$rLAI = LAI_{(mea)} - LAI_{(Ref)} \tag{1}$$

where *rLAI* is the plot-level relative LAI, *LAI<sub>(mea)</sub>* is the measured LAI of a study plot, *LAI<sub>(Ref)</sub>* is the measured LAI of reference plot.

$$rVI = VI_{(mea)} - VI_{(Ref)} \tag{2}$$

where *rVI* is the plot-level relative VI, *VI<sub>(mea)</sub>* is the plot-level VI calculated by measured spectra, *VI<sub>(Ref)</sub>* is the VI calculated by measured spectra of reference plot.

$$relative\ yield = yield_{(mea)} - yield_{(Ref)} \tag{3}$$

where *yield<sub>(mea)</sub>* is the measured yield of a study plot, *yield<sub>(Ref)</sub>* is the measured yield of the reference plot.

### Vegetation index selection

Many scholars have determined that the optimal bands for studying the relationship between vegetation spectra and biophysical parameters lie in the visible and near-infrared ranges [43, 44]. According to this, VIs of NDVI, *CI<sub>red edge</sub>*, *CI<sub>green</sub>*, EVI2, NDRE, and MTCI (Table 1) calculated by the green (550 nm), red (670 nm), red edge (720 nm),

and near-infrared (800 nm) bands were built. The reason why these six VIs were selected is that many scholars have achieved good results in relevant studies.

### Algorithms for determining the weights of growth stages

#### Slogistic model

The curve expression of the Slogistic model is shown as Eq. (4). With the increase of independent variable, the value of the dependent variable increases slowly at first, but rapidly in a certain range later. When the independent variable reaches a certain limit, the growth of the dependent variable tends to be slow, and the whole curve shows a shape of flat "S". This equation is extensively used in epidemiology and agrometeorology [50].

$$y = \frac{a}{1 + be^{-kx}} \tag{4}$$

where *a* refers to the maximum value of the dependent variable, *b* and *k* are the characteristic parameters of the Slogistic curve equation.

The first-order and second-order partial derivatives of the independent variable of Eq. (4) were calculated to obtain Eqs. (5) and (6). According to the trend of curve change, the Slogistic model can be divided into three parts: the range of  $[0 \sim [\ln b - \ln(2 + \sqrt{3})]/k]$  is the gradually increasing stage,  $[[\ln b - \ln(2 + \sqrt{3})]/k \sim [\ln b + \ln(2 + \sqrt{3})]/k]$  is the rapidly increasing stage, and  $[[\ln b + \ln(2 + \sqrt{3})]/k \sim \infty]$  is the slowly increasing stage. When the independent variable is  $\ln b/k$ , the increasing speed of the dependent variable reaches the maximum value. The establishment of the model is helpful to judge the potato growth stages and determine their weights.

$$\frac{dy}{dx} = \frac{abk^{-kt}}{(1 + bk^{-kt})^2} \tag{5}$$

$$\frac{d^2y}{dx^2} = \frac{abk^{-kt}(abk^{-kt} - k)}{(1 + be^{-kt})^3} \tag{6}$$

**Table 1** Vegetation indices used in this study

Vegetation indices	Formula	References
Normalized Difference Vegetation Index (NDVI)	$(R_{800} - R_{670}) / (R_{800} + R_{670})$	Rouse et al. [45]
Red edge Chlorophyll Index ( <i>CI<sub>red edge</sub></i> )	$R_{800} / R_{720} - 1$	Gitelson et al. [46]
Green edge Chlorophyll Index ( <i>CI<sub>green</sub></i> )	$R_{800} / R_{550} - 1$	Gitelson et al. [46]
Two-band Enhanced Vegetation Index (EVI2)	$2.5(R_{800} - R_{670}) / (1 + R_{800} + 2.4R_{670})$	Jiang et al. [47]
Normalized Difference Red edge (NDRE)	$(R_{800} - R_{720}) / (R_{800} + R_{720})$	Gitelson et al. [48]
MERIS Terrestrial Chlorophyll Index (MTCI)	$(R_{800} - R_{720}) / (R_{720} - R_{670})$	Dash et al. [49]



**Improved analytic hierarchy process**

Analytic hierarchy process (AHP) is a system analysis method that combines qualitative and quantitative analysis, which was put forward by T.L. Saaty, a famous American operational research scientist in the early 1970s [51]. The judgment matrix of the traditional AHP adopted a nine-scale method (1–9). The subjective factors of experts play a leading role, which will lead to the deviation of the evaluation results. In addition, if the judgment matrix is not consistent in the consistency test, it will destroy the main function of the AHP’s scheme optimization and sorting, with a large amount of calculation and low accuracy. The improved analytic hierarchy process (IAHP) developed a new three-scale method (0–2), which made it easy for experts to make a comparison of the relative importance of the two factors, without the need for a consistency test. Moreover, IAHP can greatly reduce the number of iterations, improve the convergence speed, and meet the requirements of calculation accuracy [52]. The specific calculation steps are as follows:

1. Construction of comparison matrix  $A(a_{ij})$ .

As shown in Eq. (7), according to the relative importance of potato growth stages, a comparison matrix  $A(a_{ij})_{5 \times 5}$  was constructed.

$$A(a_{ij}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 2 \\ 2 & 2 & 1 & 2 & 2 \\ 2 & 2 & 0 & 1 & 2 \\ 2 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{7}$$

where 0 indicates that the stage  $i$  is not as important as stage  $j$ ; 1 indicates that the stage  $i$  is as important as stage  $j$ ; 2 indicates that the stage  $i$  is more important than stage  $j$ .

2. Construction of judgment matrix  $B(b_{ij})$ .

Firstly, the importance coefficients ( $r_j = \sum_{i=1}^5 b_{ij}$ ) of five potato growth stages were calculated, and then the judgment matrix  $B(b_{ij})$  was constructed as shown in Eq. (8):

$$B(b_{ij}) = \begin{cases} \frac{r_i - r_j}{r_{max} - r_{min}} \times (k - 1) + 1 & r_i \geq r_j \\ \left[ \frac{|r_i - r_j|}{r_{max} - r_{min}} \times (k - 1) + 1 \right]^{-1} & r_i < r_j \end{cases} \tag{8}$$

where  $r_{max} = \max\{r_j\}$ ,  $r_{min} = \min\{r_j\}$ ,  $k = r_{max}/r_{min}$ .

3. Calculation of transfer matrix  $C(c_{ij})$  and quasi-optimal uniform matrix  $C^*(c_{ij}^*)$ .

The elements in transfer matrix  $C(c_{ij})$  and quasi-optimal uniform matrix  $C^*(c_{ij}^*)$  need to meet Eqs. (9) and (10).

$$C(c_{ij}) = \frac{1}{5} \sum_{t=1}^5 \left( \lg \frac{b_{it}}{b_{jt}} \right) \tag{9}$$

$$C^*(c_{ij}^*) = 10^{c_{ij}} \tag{10}$$

4. Weight determination.

The maximum eigenvalue and the maximum eigenvector of the quasi-optimal matrix  $C^*$  were calculated, and the weight of each growth stage can be obtained after normalization.

**Entropy weight method**

The entropy weight method (EW) determines the index weight according to the variation degree of each index value. It is an objective weighting method, which has been widely used in the fields of economy, engineering, and finance [53]. The advantage of this method is that it can avoid the influence of human factors, but it ignores the importance of the index itself. Sometimes the weight of the index determined is far from the expected result, and the dimension of the evaluation index cannot be reduced [54]. The data matrix of  $G(g_{ij})_{5 \times 5}$  was constructed based on the potato characteristic parameters of different plot-level in different stages, then the entropy value ( $e_j$ ) and the difference coefficient ( $d_j$ ) of each growth stage were calculated as shown in Eqs. (11) and (12).

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n \left( \frac{g_{ij}}{\sum_{i=1}^n g_{ij}} \ln \frac{g_{ij}}{\sum_{i=1}^n g_{ij}} \right) \tag{11}$$

$$d_j = \frac{1 - e_j}{5 - \sum_{j=1}^5 e_j} \tag{12}$$

The weight  $w_j$  of the growth stage  $j$  can be obtained by normalizing the difference coefficient  $d_j$  as shown in Eq. (13).

$$w_j = \frac{d_j}{\sum_{j=1}^5 d_j} \quad (0 < w_j < 1, \sum_{j=1}^5 w_j = 1) \tag{13}$$

**Optimal combination weighting method**

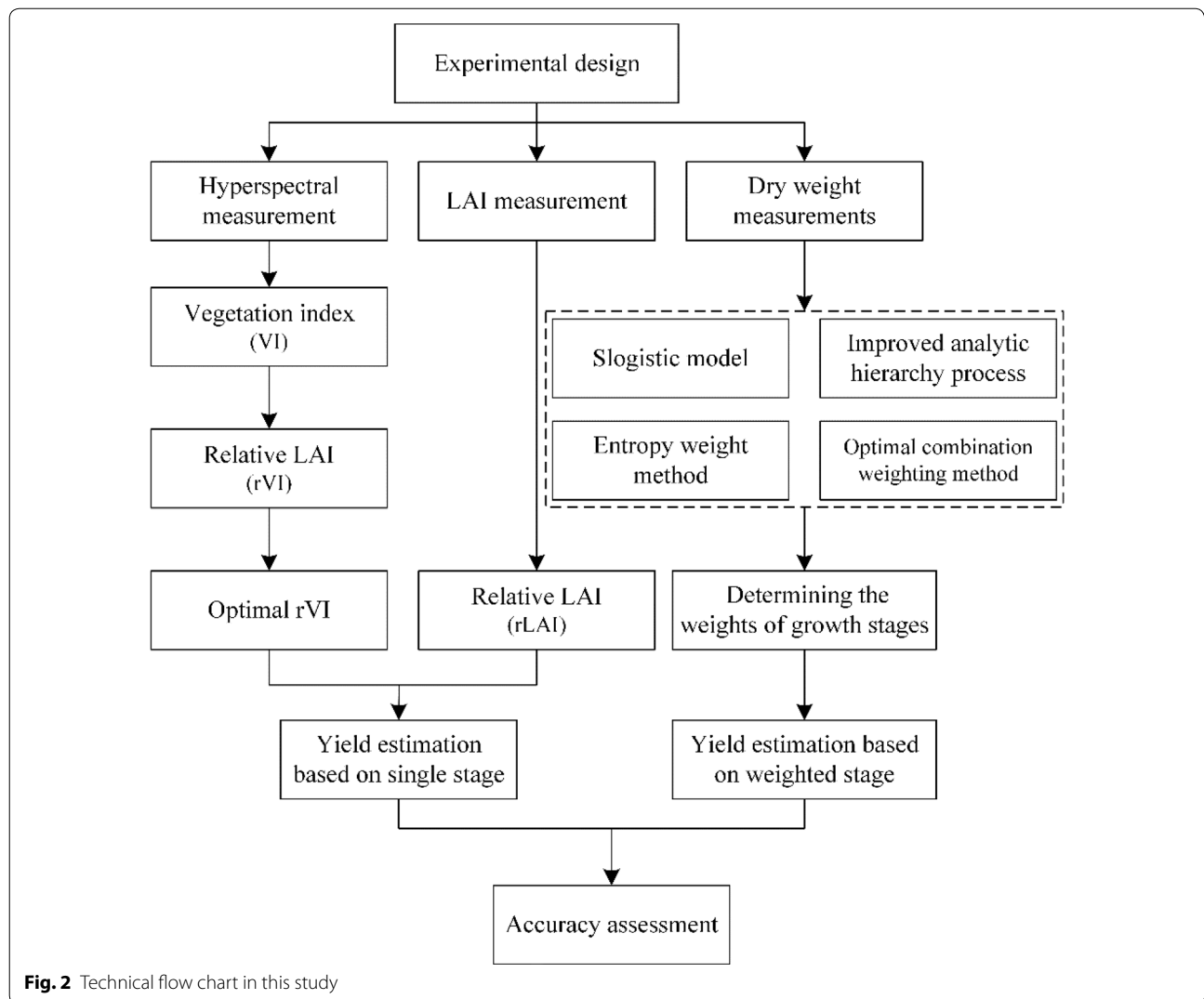
An optimal combination weighting method (OCW) was employed to solve the proportion of weights in the combined decision-making based on obtaining subjective and

objective weights, then the decision weights considering both subjective will and objective existence were obtained [55]. To select a set of weights with the largest total distance (R) between the subjective weights and objective weights, the weight determined by the subjective weighting method was written as  $W_1 = (w_1^1, w_2^2, w_3^3, w_4^4)$ , the weight determined by the objective weighting method was written as  $W_2 = (w_1^2, w_2^2, w_3^2, w_4^2)$ , and the combined weight determined by OCW was written as  $W = (w_1, w_2, w_3, w_4)$ . The optimal combination weight can be obtained by constructing the optimization model of Eq. (14) below.

$$\begin{cases} \max R = \sum_{m=1}^2 (1 - \sqrt{\frac{1}{5} \sum_{j=1}^5 (w_j - w_j^m)^2}) \\ \sum_{j=1}^5 w_j = 1 \end{cases} \quad (14)$$

**Leave-one-out cross-validation**

The technical flow chart (Fig. 2) demonstrates the experimental methodology in this study, including experimental design, data collection, data processing, methods, and writing logic. The estimation and validation models of potato yield were established using leave-one-out cross-validation (LOOCV). This method is widely employed in model construction and validation to reduce the dependence on a single random part of the calibration and validation datasets [56]. Firstly, the original population samples were divided into K mutually exclusive sets (K=26 in this study), of which K - 1 sets were used iteratively as training data for calibrating the coefficients ( $Coef_i$ ) of the algorithm, and then the remaining single sample was retained as the validation to obtain  $R^2_i$  and the estimation error ( $E(y_i) - y_i$ ). The whole training and validation process should be repeated K times until each sample



**Fig. 2** Technical flow chart in this study

participates in the validation process. After  $K$  iterations, the coefficients and precision of the final algorithm can be expressed as follows:

$$Coef = \frac{\sum_{i=1}^K Coef_i}{K} \tag{15}$$

$$R^2 = \frac{\sum_{i=1}^K R_i^2}{K} \tag{16}$$

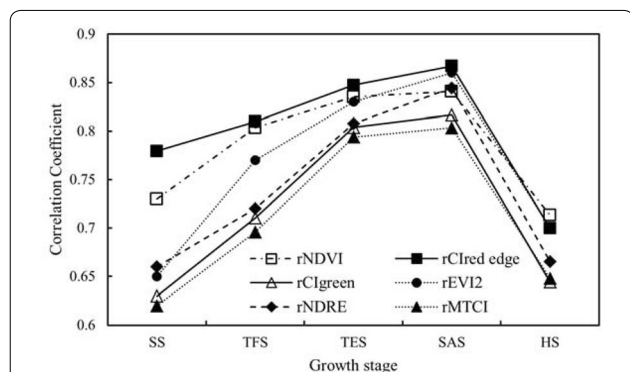
$$RMSE = \sqrt{\frac{\sum_{i=1}^K (E(y_i) - y_i)^2}{K}} \tag{17}$$

where  $E(y)$  is the actual observed value, and  $y$  is the predicted value simulated by the model.

### Results

#### Determination of the optimal rVI

Each VI in this study was converted to rVI through the transformation of Eq. (2). The correlation coefficients between rVIs of different growth stages and relative yield are shown in Fig. 3. It can be seen that the correlation coefficients between each rVI and its corresponding relative yield showed an overall trend of increasing first (SS to SAS) and then decreasing (SAS to HS) during the whole growth stage. Correlated with relative yield, correlation coefficients of all selected rVIs in different stages exhibited consistent changes: reaching maximum values at SAS (with a correlation coefficient of 0.867 for  $rCI_{red\ edge}$ , 0.860 for  $rEVI2$ , 0.845 for  $rNDRE$ , 0.841 for  $rNDVI$ , 0.817 for  $rCI_{green}$ , and 0.803 for  $rMTCI$ .) and showing smaller values at SS and HS. At each potato growth stage, there was the strongest correlation between  $rCI_{red\ edge}$  and relative yield. Therefore, it can be concluded that SAS is the most effective



**Fig. 3** Comparison of correlation coefficients between different relative vegetation indices and yield

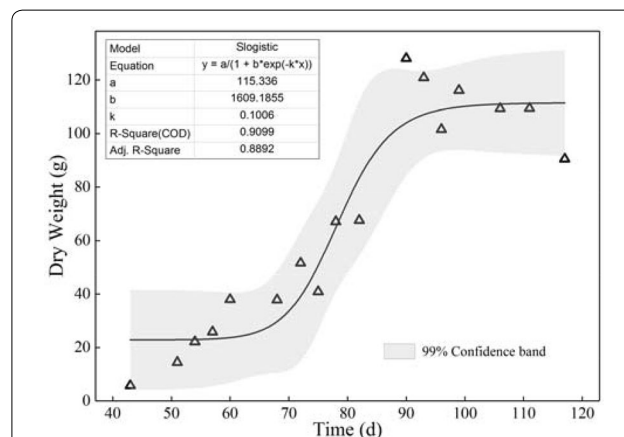
stage for potato yield estimating using VI, and  $rCI_{red\ edge}$  has the best performance. When using rVI to construct yield models, only  $rCI_{red\ edge}$  will be considered.

#### Simulation of potato growth based on Slogistic model

As shown in Fig. 4, the dry weight data of the whole growth stage were used to construct the Slogistic model to characterize the growth process of potatoes. It can be found that the simulation accuracy is high, with the adjusted  $R^2$  close to 0.9. Generally speaking, the growth speed of potato is relatively slow in the early and late stages, and faster in the middle stage. Equations (5) and (6) were utilized to calculate the length of the growth stage, and the time nodes of the gradually increasing stage, the rapidly increasing stage, and the slowly increasing stage were 60th and 86th days, respectively. Based on these three stages, the five growth stages of this study can be obtained by increasing the seeding stage and harvest stage. The importance degree of each growth stage relative to yield can be sorted according to the growth rate of different stages. Combined with the actual planting experience, the final importance ranking was determined as  $TES > SAS > TFS > HS > SS$ . This result can provide a reference for the determination of the weights of different growth stages.

#### Estimation of potato yield based on a single developmental stage

The new VI (rVI) and LAI(rLAI) datasets were compared with the relative yield data at five different developmental stages respectively. Adjust coefficient of determination ( $R^2$ ) and root mean square error (RMSE) of all estimation models of single-stage rLAI and  $rCI_{red\ edge}$  at each growth stage are shown in Table 2. At the same time, F-test was conducted on the



**Fig. 4** Variation of potato dry weight with time during the whole growth stage

**Table 2** Potato yield estimation models for single stage

Growth stage	Regression equation	Adjusted R <sup>2</sup>	RMSE	Significance test
SS	$y = 1.9752a_1 - 0.1321$	0.5912	0.3361	$P < 0.001$
	$y = 1.6494b_1 + 0.0935$	0.4064	0.4043	$P < 0.001$
TFS	$y = 0.4626a_2 - 0.1644$	0.6423	0.3145	$P < 0.001$
	$y = 0.4223b_2 - 0.0487$	0.5661	0.3463	$P < 0.001$
TES	$y = 0.4421a_3 - 0.1823$	0.7415	0.2671	$P < 0.001$
	$y = 0.4305b_3 - 0.1747$	0.7034	0.2864	$P < 0.001$
SAS	$y = 0.3436a_4 + 0.0572$	0.7063	0.2844	$P < 0.001$
	$y = 0.3494b_4 + 0.0593$	0.6835	0.2956	$P < 0.001$
HS	$y = 0.3269a_5 + 0.0961$	0.4692	0.3835	$P < 0.001$
	$y = 0.3984b_5 - 0.0706$	0.5174	0.3654	$P < 0.001$

a1, a2, a3, a4, a5 are plot-level rVI ( $CI_{red\ edge}$  in this study) at SS, TFS, TES, SAS, and HS, respectively

b1, b2, b3, b4, b5 are plot-level rLAI at SS, TFS, TES, SAS, and HS, respectively

whole regression models at 0.01 probability level, and the results were measured by P-value. Potato rLAI and  $rCI_{red\ edge}$  at TES closely related to the relative yield having the adjusted R<sup>2</sup> above 0.7, much lower correlations were found at SS and HS. From the perspective of different stages, TES is the optimal stage when using rVI and rLAI to estimate potato yield, and the models' expressions are shown in Eqs. (18) and (19). In this stage, the prediction performance of VI is better than that of LAI (Adjusted R<sup>2</sup> of 0.7415 vs. 0.7034, RMSE of 0.2671 vs. 0.2864).

$y$  is the relative yield

$$yield_{(VI)} = 0.4421 \times rCI_{red\ edge[TES]} - 0.1823 + yield_{(Ref)} \tag{18}$$

where  $yield_{(VI)}$  is the estimated potato yield using single-stage rVI ( $rCI_{red\ edge}$  at TES in this study).

$$yield_{(LAI)} = 0.4305 \times rLAI_{[TES]} - 0.1747 + yield_{(Ref)} \tag{19}$$

where  $yield_{(LAI)}$  is the estimated potato yield using single-stage rLAI.

**Estimation of potato yield based on weighted growth stage**

Three weight calculation methods (subjectivity, objectivity, and their combined form) were used to determine the weights of potato growth stages (Table 3). The results showed that the weights of each growth stage determined by EW were very close. The weights determined by IAHP and OCW were the largest at TES, followed at SAS and TFS, and the smallest at HS and SS. Based on the weighting results of the three methods (IAHP, EW, and OCW), the rVI and rLAI data of the potato's critical growth stages in the study area were calculated (ie, the weighted  $rCI_{red\ edge}$  and rLAI), and then the linear regression models between the weighted relative variables and the relative potato yield were obtained. It can be found that the correlation between the potato yield and the weighted variables ( $rCI_{red\ edge}$  and rLAI) obtained by the three weighting methods was very significant ( $P < 0.001$ ). For the three different weighting method models, the EW-based and OCW-based methods had the lowest and the highest model accuracy, respectively. But the results obtained by these three methods were significantly improved compared to the single-stage models. By comparing the fitting models of the two relative variables,

**Table 3** Potato yield estimation models based on weighted growth stage

Weighting method	Vegetation parameter	SS	TFS	TES	SAS	HS	Regression model	Adjusted R <sup>2</sup>	RMSE	Significance test
IAHP	$CI_{red\ edge}$	0.0329	0.1296	0.5101	0.2638	0.0636	$y = 0.4484a - 0.1318$	0.8092	0.2340	$P < 0.001$
	LAI						$y = 0.4717b - 0.1778$	0.8127	0.2318	$P < 0.001$
EW	$CI_{red\ edge}$	0.1441	0.1556	0.2906	0.1633	0.2464	$y = 0.5504a - 0.2357$	0.7949	0.2426	$P < 0.001$
	LAI	0.1808	0.1701	0.2958	0.1679	0.1854	$y = 0.6056b - 0.3255$	0.7960	0.2419	$P < 0.001$
OCW	$CI_{red\ edge}$	0.0663	0.1374	0.4442	0.2337	0.1185	$y = 0.4775a - 0.1788$	0.8225	0.2257	$P < 0.001$
	LAI	0.0773	0.1417	0.4458	0.2350	0.1002	$y = 0.5169b - 0.2330$	0.8333	0.2187	$P < 0.001$

a and b represent  $rCI_{red\ edge}$  and rLAI of weighted growth stage respectively.  $y$  is the relative yield of potato



the results obtained by the three weighting methods all showed that the weighted rLAI-based models had higher accuracy than the weighted rCI<sub>red edge</sub>-based models. The optimal estimation models of potato yield can be determined as Eqs. (20) and (21). As the final estimation models of potato yield were based on the relative yield model by adding the yield of the reference spot, their prediction ability remains unchanged (Table 3).

$$\begin{aligned} \text{yield}_{(VI)} = & 0.4775 \times rCI_{red\ edge}[\text{weighted stage}] \\ & - 0.1788 + \text{yield}_{(Ref)} \end{aligned} \quad (20)$$

where  $\text{yield}_{(VI)}$  is the estimated potato yield using rVI based on the weighted growth stage.

$$\text{yield}_{(LAI)} = 0.5169 \times rLAI[\text{weighted stage}] - 0.2330 + \text{yield}_{(Ref)} \quad (21)$$

where  $\text{yield}_{(LAI)}$  is the estimated yield using rLAI based on the weighted growth stage.

#### Accuracy assessment using leave-one-out cross-validation

The leave-one-out cross-validation (LOOCV) method was utilized to obtain the potato yield validation models (Fig. 5).  $R^2$ , RMSE, and mean relative error (MRE) were taken as evaluation indices. The results indicated that the accuracy of all models was acceptable ( $R^2 > 0.75$  and  $RMSE < 0.26$ ). In general, models with high simulation accuracy also have high verification accuracy, with the minimum error less than 9%. Based on the combination models of three weighting methods and two different variables, the EW-based LAI model has the lowest accuracy, while the OCW-based LAI model has the highest accuracy ( $R^2 = 0.8234$ ,  $RMSE = 0.2267$ ,  $MRE = 0.0833$ ), explaining 82% of the variability. Therefore, combining the estimation and the verification models, the LAI model based on the OCW method to determine the weights of different growth stages is the optimal model for potato yield estimation.

#### Discussion

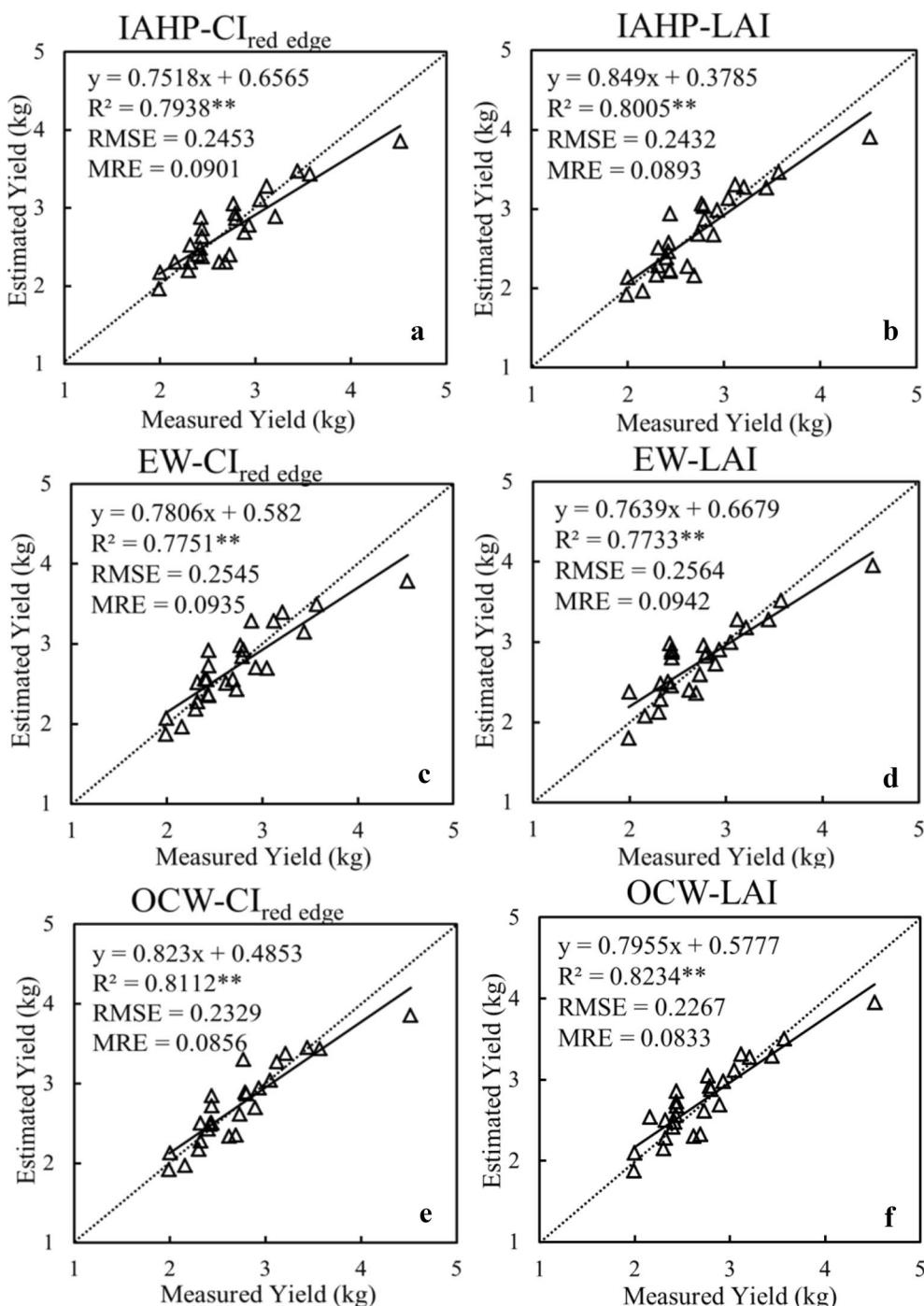
For potato yield estimation, most scholars used to employ some crop growth models derived from general crop growth models or from gramineous (rice, wheat, corn, etc.) crop growth models [57]. Based on the principle and structure of the original model, the corresponding parameters were modified to conform to the growth characteristics of potato, and the growth process of potato was simulated, so as to output the physiological characteristic parameters and yield data and realize the model simulation function. Quiroz et al. [3] proposed that the incorporation of remotely sensed data in crop growth models with different temporal resolutions and

levels of complexity could help to improve the yield estimation in potato. Moreover, it was identified that LAI at the initiation of stem elongation stage was closely related to yield, thus the remote estimation of LAI at this stage could be used to indicate the yield in oilseed rape [36]. Sharma et al. [58] tested Trimble GreenSeeker® (TGS) and Holland Scientific Crop Circle™ ACS-430 (HCCACS-430) wavebands to predict potato yield using LAI and NDVI with  $R^2$  reaching 0.7. These studies indicate that both remote sensing and LAI data have potential for yield prediction. Therefore, spectra and LAI data were selected in this paper to estimate potato yield.

Six VIs of NDVI, CI<sub>red edge</sub>, CI<sub>green</sub>, EVI2, NDRE, and MTCI were utilized to non-destructively estimate potato yield in this study (Table 1) and CI<sub>red edge</sub> showed the most excellent performance in the correlation with potato yield (Fig. 3). Gong et al. [24] also proved that CI<sub>red edge</sub> had a good effect on the estimation of rapeseed yield using UAV data. Ma et al. [43] pointed out that at the seeding and bolting stage, the CI<sub>red edge</sub> exhibited good performance compared to the other VIs. These conclusions are consistent with the results of this study, proving the credibility of this study.

In this study, the concepts of rVI and rLAI were proposed to solve the problem that the data acquired in different stages would be affected by solar radiation, aerosol, and soil background. Under the assumption of constant external conditions, subtraction can effectively remove these interferences, so that multi-period data can be used in combination. Furthermore, this method has the advantage of not changing the degree of aggregation and preserving the deviation of the original data. Wang et al. [59] used division to construct several relative vegetation indices ( $\Delta VI$ ) to estimate rice yield with hyperspectral imagery. Although the influence of external conditions such as background can be eliminated to some extent, the problem of changing the aggregation degree of data is ignored, resulting in lower RMSE and larger  $R^2$ .

The dry weight data of the whole growth period were used to fit the Slogistic model (Fig. 4) by analyzing the growth process of potato (the growth rate is slow in the early and late stages, and fast in the middle stage). According to the model, we can not only divide the different growth stages of potato but also provide the basis for determining the weights of each growth stage. There are few systematic and specific divisions of potato growth stages in the existing literature. The main reason is that potato tubers are buried in the soil, and the changes can not be observed directly by the eyes. Therefore, the joint utilization of potato multi-period data is subject to certain restrictions [60]. With a clear division of growth stages, more refined research can be carried out like crops such as rice [61] and wheat [62].



**Fig. 5** Validation models of potato yield estimation based on weighted growth stages and different variables: **a** IAHP- $CI_{red\ edge}$ ; **b** IAHP-LAI; **c** EW- $CI_{red\ edge}$ ; **d** EW-LAI; **e** OCW- $CI_{red\ edge}$ ; **f** OCW-LAI

At present, there are many problems about the joint use of multiple-period or various kinds of data. To improve the accuracy of the research results, many scholars blindly used the data of multiple-period or diversified data directly. For example, Zhou et al. [63] predicted rice

grain yield using multiple linear regression (MLR) with multi-temporal VIs derived from the multi-spectral and digital images to improve the estimation accuracy. Obviously, the contributions of different developmental stages to yield estimation are not consistent, so it can not be

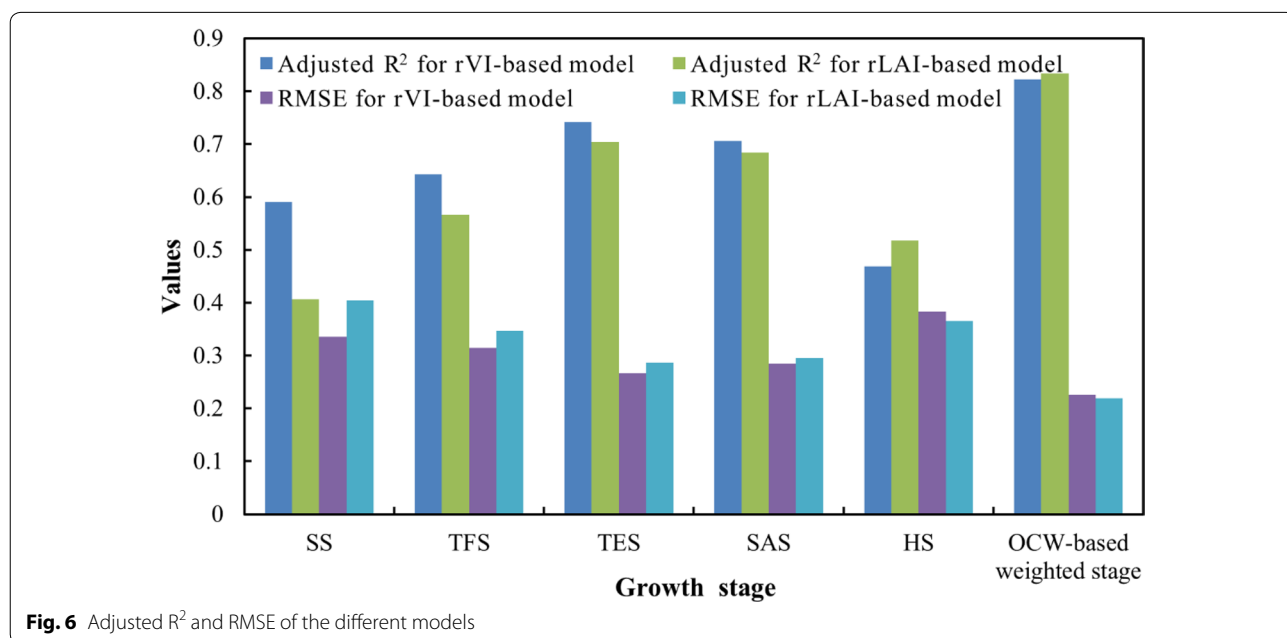
directly used for MLR, ignoring the weights of growth stages. Wang et al. [64] estimated LAI of paddy rice using MLR, partial least squares (PLS) regression, and least squares support vector machines (LS-SVM) regression with 15 optimal hyperspectral bands to product more accuracy. No research has shown that the contributions of these 15 bands to LAI estimation are the same, so these data can not be directly used together. Of course, if multiple data are obtained in the same period, it can be used directly in combination. For example, Duan et al. [35] predicted rice LAI using SVM regression with spectral features and the texture features to determine the texture feature effective. In this study, IAHP, EW, and OCW methods were employed to confirm the weights of different stages of potato. From the perspective of subjectivity, objectivity, and the combination of them, the most suitable method (OCW) was selected, which solved the problem of joint use of multi-period data. The weighting results (Table 3) of different potato growth stages determined by EW are relatively close, thus they cannot reflect the degree of impact of different growth stages on yield. The calculation results of IAHP and OCW are in accordance with the actual situation.

When the spectra and LAI data of a single stage were used to predict the potato yield, the estimation accuracy of each stage basically met (1) TES > SAS > TFS > HS > SS for the same variable, which is consistent with the ranking of weights determined by IAHP, EW, and OCW; (2) VI > LAI for the same stage (Table 2). At HS, the simulation accuracy of VI is lower than that of LAI (adjusted R<sup>2</sup> of 0.4692 vs. 0.5174), and when using the variable of

VI, the accuracy at HS is lower than at SS (adjusted R<sup>2</sup> of 0.4692 vs. 0.5912). The reason for this result is probably that the withering of potato leaves at HS resulted in the change of canopy spectra and the decrease of yield prediction ability.

To compare the accuracy of yield estimation, the linear regression models were constructed based on plot-level weighted variables (rVI and rLAI) and relative yield (Table 3). The accuracy of different models was shown in Fig. 6. It can be found that OCW-based models have the highest accuracy. Unlike the single-stage results, in the OCW-based models, the accuracy of the rLAI model is higher than that of the rVI model. Because the appearance of saturation phenomenon in yield estimation using spectral index will limit the accuracy of models to some extent [25]. The LAI data is the three-dimensional (3D) information of the crop, and the limitations will be reduced.

To improve the suitability of the model, this experiment was set as water and fertilizer conditions, which can meet the current situation of water stress in potato planting areas in China and even the world [65]. Our future work will contain more data from different platforms for analysis, especially the UAV and satellite data because they can well express data at the spot-level. In addition, we will conduct experiments in more regions to verify the robustness of the models. And a new instrument of the LI-3100C table leaf area meter, (LI-COR Inc., Lincoln, Nebraska, USA) [66] will be used to avoid the impacts of the stems and flowers on the output of LAI, and more realistic LAI data will be obtained to improve



and validate the accuracy of potato yield estimation by ground-measurement data.

## Conclusions

In this study, we developed a technique to improve the estimation of potato yield using weighted relative variables at plot-level derived from multi-period LAI and hyperspectral data. Plot-level relative vegetative index and LAI (rVI and rLAI) were proposed to eliminate the influence of external factors (solar radiation, aerosol, and soil background). The weights of different growth stages of potato were determined based on the Slogistic model and three weight calculation methods (IAHP, EW, and OCW). The linear regression was performed to estimate potato yield using single-stage and weighted multiple-stage variables respectively. The results indicated that  $rCI_{red\ edge}$  was the optimal index for the potato yield estimation among all the test rVIs. TES is most suitable for potato yield estimation using a single growth stage. When multi-period data were applied to estimate the potato yield, the accuracy was greatly improved. The estimation model of LAI using the OCW-based method combining subjectivity and objectivity (OCW-LAI) showed the best performance with the estimation error about 8%.

Although the idea of weighted developmental stage based on the Slogistic model and weighting calculation methods proposed in this study were tested in potato yield estimation, this work may offer a theoretical reference for other key parameters retrieving in crops that have an apparent division of growth stages. In future work, we will attempt to apply this technique to predict other growth parameters in potato and other crops.

## Acknowledgements

We thank the Potato Science Institute of Jilin Academy of Vegetables and Flower Sciences for preparing the seed and planting for the experiments, and Professor Shengli Zhang, Dr. Fei Xu, and Mr. Zhongcai Han for designing the trail.

## Authors' contributions

SL and YH conceived and designed the experiments. QL and YZ performed the experiments. SL analyzed the data and wrote this manuscript. WJ and XZ checked the language. YH, QL, WJ, and YZ contributed to the discussion. All authors read and approved the final manuscript.

## Funding

This study was supported by the National Natural Science Foundation of China "Study on temporally and spatially precise assessment on potato cultivation suitability based on dynamic process-oriented mode" (41771562) and "Innovation Project" of the Chinese Academy of Agricultural Sciences (2016–2020, IARRP).

## Availability of data and materials

The remotely sensed and yield data used in this study is available upon the approval of Dr. Yingbin He from the Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, China.

## Consent for publication

All authors agreed to publish this manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup> Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China. <sup>2</sup> School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China. <sup>3</sup> School of Economics and Management, Tiangong University, Tianjin 300387, China. <sup>4</sup> Center for Agricultural and Rural Economic Research, Shandong University of Finance and Economics, Jinan 250014, China.

Received: 14 June 2020 Accepted: 2 November 2020

Published online: 10 November 2020

## References

- Luo SJ, He YB, Duan DD, Wang ZZ, Zhang JK, Zhang YT, et al. Analysis of hyperspectral variation of different potato cultivars based on continuum removed spectra. *Spectrosc Spec Anal*. 2018;38:3231–7.
- Sulli M, Mandolino G, Sturaro M, Onofri C, Diretto G, Parisi B, et al. Molecular and biochemical characterization of a potato collection with contrasting tuber carotenoid content. *PLoS ONE*. 2017;12:e0184143.
- Quiroz R, Loayza H, Barreda C, Gavilan C, Posadas A, Ramirez DA Linking process-based potato models with light reflectance data: does model complexity enhance yield prediction accuracy? *Eur J Agron*. 2017;82:104–12.
- Duan DD, He YB, Luo SJ, Wang ZZ. Analysis on the ability of distinguishing potato varieties with different hyperspectral parameters. *Spectrosc Spec Anal*. 2018;38:3215–20.
- Al-Gaadi KA, Hassaballa AA, Tola E, Kayad AG, Madugundu R, Alblewi B, et al. Prediction of potato crop yield using precision agriculture techniques. *PLoS ONE*. 2016;11:e0162219.
- Reynolds CA, Yitayew M, Slack DC, Hutchinson CF, Huete A, Petersen MS. Estimating crop yields and production by integrating the FAO crop specific water balance model with real-time satellite data and ground-based ancillary data. *Int J Remote Sens*. 2000;21:3487–508.
- Campos I, Neale CMU, Arkebauer TJ, Suyker AE, Goncalves IZ. Water productivity and crop yield: a simplified remote sensing driven operational approach. *Agric Forest Meteorol*. 2018;249:501–11.
- Raymundo R, Asseng S, Cammarano D, Quiroz R. Potato, sweet potato, and yam models for climate change: a review. *Field Crops Res*. 2014;166:173–85.
- Setiyono TD, Quicho ED, Holecz FH, Khan NI, Romuga G, Maunahan A, et al. Rice yield estimation using synthetic aperture radar (SAR) and the ORYZA crop growth model: development and application of the system in South and South-east Asian countries. *Int J Remote Sens*. 2019;40:8093–124.
- Novelli F, Vuolo F. Assimilation of sentinel-2 leaf area index data into a physically-based crop growth model for yield estimation. *Agronomy*. 2019;9:255.
- Luo SJ, He YB, Wang ZZ, Duan DD, Zhang JK, Zhang YT, et al. Comparison of the retrieving precision of potato leaf area index derived from several vegetation indices and spectral parameters of the continuum removal method. *Eur J Remote Sens*. 2019;52:155–68.
- Li SY, Ding XZ, Kuang QL, Ata-UI-Karim ST, Cheng T, Liu XJ, et al. Potential of UAV-based active sensing for monitoring rice leaf nitrogen status. *Front Plant Sci*. 2018;9:1934.
- Yao YJ, Liu QH, Liu Q, Li XW. LAI retrieval and uncertainty evaluations for typical row-planted crops at different growth stages. *Remote Sens*. 2008;12:94–106.
- Battude M, Al Bitar A, Morin D, Cros J, Huc M, Sicre CM, et al. Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sens Environ*. 2016;184:668–81.
- Liu NF, Budkewitsch P, Treitz P. Examining spectral reflectance features related to Arctic percent vegetation cover: implications for hyperspectral remote sensing of Arctic tundra. *Remote Sens Environ*. 2017;192:58–72.
- Woolley JT. Reflectance and transmittance of light by leaves. *Plant physiol*. 1971;47:656–62.

17. Gausman HW, Allen WA, Cardenas R. Reflectance of cotton leaves and their structure. *Remote Sens Environ.* 1969;1:19–22.
18. Towers PC, Strever A, Poblete-Echeverria C. Comparison of vegetation indices for leaf area index estimation in vertical shoot positioned Vine canopies with and without grenbiule hail-protection netting. *Remote Sens.* 2019;11:1073.
19. Wu JD. Developing general equations for urban tree biomass estimation with high-resolution satellite imagery. *Sustain.* 2019;11:4347.
20. Zhang XH, He Y, Wang C, Xu F, Li XH, Tan CW, et al. Estimation of corn canopy chlorophyll content using derivative spectra in the O-2-A absorption band. *Front Plant Sci.* 2019;10:1047.
21. Chen JX, Li F, Wang R, Fan YF, Raza MA, Liu QL, et al. Estimation of nitrogen and carbon content from soybean leaf reflectance spectra using wavelet analysis under shade stress. *Comput Electron Agric.* 2019;156:482–9.
22. Anderson GL, Hanson JD, Haas RH. Evaluating landsat thematic mapper derived vegetation indexes for estimating aboveground biomass on semiarid rangelands. *Remote Sens Environ.* 1993;45:165–75.
23. Miller JR, Hare EW, Wu J. Quantitative characterization of the vegetation red edge reflectance. 1. An inverted-gaussian reflectance model. *Int J Remote Sens.* 1990;11:1755–73.
24. Gong Y, Duan B, Fang SH, Zhu RS, Wu XT, Ma Y, et al. Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. *Plant Methods.* 2018;14:70.
25. Duan B, Fang SH, Zhu RS, Wu XT, Wang SQ, Gong Y, et al. Remote estimation of rice yield with unmanned aerial vehicle (UAV) data and spectral mixture analysis. *Front Plant Sci.* 2019;10:204.
26. Shiu YS, Chuang YC. Yield estimation of paddy rice based on satellite imagery: comparison of global and local regression models. *Remote Sens.* 2019;11:111.
27. Joshi VR, Thorp KR, Coulter JA, Johnson GA, Porter PM, Strock JS, et al. Improving site-specific maize yield estimation by integrating satellite multispectral data into a crop model. *Agronomy.* 2019;9:719.
28. Sakamoto T, Gitelson AA, Arkebauer TJ. Near real-time prediction of US corn yields based on time-series MODIS data. *Remote Sens Environ.* 2014;147:219–31.
29. Mateo-Sanchis A, Piles M, Munoz-Mari J, Adsua JE, Perez-Suay A, Camps-Valls G. Synergistic integration of optical and microwave satellite data for crop yield estimation. *Remote Sens Environ.* 2019;234:12.
30. Becker-Reshef I, Justice C, Sullivan M, Vermote E, Tucker C, Anyamba A, et al. Monitoring global croplands with coarse resolution earth observations: the global agriculture monitoring (GLAM) project. *Remote Sens.* 2010;2:1589–609.
31. Dong TF, Liu JG, Shang JL, Qian BD, Ma BL, Kovacs JM, et al. Assessment of red-edge vegetation indices for crop leaf area index estimation. *Remote Sens Environ.* 2019;222:133–43.
32. Li SY, Yuan F, Ata-Ul-Karim ST, Zheng HB, Cheng T, Liu XJ, et al. Combining color indices and textures of UAV-based digital imagery for rice LAI estimation. *Remote Sens.* 2019;11:1763.
33. Verrelst J, Camps-Valls G, Munoz-Mari J, Rivera JP, Veroustraete F, Clevers JGPW, et al. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties—a review. *ISPRS J Photogramm Remote Sens.* 2015;108:273–90.
34. Sun L, Gao F, Anderson MC, Kustas WP, Alsina MM, Sanchez L, et al. Daily mapping of 30 m LAI and NDVI for grape yield prediction in California Vineyards. *Remote Sens.* 2017;9:317.
35. Duan B, Liu YT, Gong Y, Peng Y, Wu XT, Zhu RS, et al. Remote estimation of rice LAI based on Fourier spectrum texture from UAV image. *Plant Methods.* 2019;15:124.
36. Peng Y, Zhu TE, Li YC, Dai C, Fang SH, Gong Y, et al. Remote prediction of yield based on LAI estimation in oilseed rape under different planting methods and nitrogen fertilizer applications. *Agric For Meteorol.* 2019;271:116–25.
37. Liu XJ, Zhang K, Zhang ZY, Cao Q, Lv ZF, Yuan ZF, et al. Canopy chlorophyll density based index for estimating nitrogen status and predicting grain yield in rice. *Front Plant Sci.* 2017;8:1829.
38. Wang ZX, Liu C, Huete A. From AVHRR-NDVI to MODIS-EVI: advances in vegetation index research. *Acta Ecol Sin.* 2003;23:979–87.
39. Du Y, Teillet PM, Cihlar J. Radiometric normalization of multitemporal high-resolution satellite images with quality control for land cover change detection. *Remote Sens Environ.* 2002;82:123–34.
40. Xu F, Liu W, Huang YJ, Liu QN, Zhang CJ, Hu HH, et al. Screening of potato flour varieties suitable for noodle processing. *J Food Process Preserv.* 2020;44:e14344.
41. Zhao P, Fan WJ, Liu Y, Mu XH, Xu XR, Peng JJ. Study of the remote sensing model of FAPAR over rugged terrains. *Remote Sens.* 2016;8:309.
42. Delgado AJ, Castellanos EM, Sinhorette MAC, Oliveira DC, Abdulhameed N, Geraldini S, et al. The use of different photoinitiator systems in photopolymerizing resin cements through ceramic veneers. *Oper Dent.* 2019;44:396–404.
43. Ma Y, Fang SH, Peng Y, Gong Y, Wang D. Remote estimation of biomass in winter oilseed rape (*Brassica napus* L.) using canopy hyperspectral data at different growth stages. *Appl Sci.* 2019;9:545.
44. le Maire G, Francois C, Soudani K, Berveiller D, Pontailleur JY, Breda N, et al. Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sens Environ.* 2008;112:3846–64.
45. Rouse JW, Haas RH, Schell JA, Deering DW. Monitoring vegetation systems in the great plains with ERTS. *NASA Spec Publ.* 1974;309–317.
46. Gitelson AA, Vina A, Ciganda V, Rundquist DC, Arkebauer TJ. Remote estimation of canopy chlorophyll content in crops. *Geophys Res Lett.* 2005;32:L08403.
47. Jiang ZY, Huete AR, Didan K, Miura T. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens Environ.* 2008;112:3833–45.
48. Gitelson AA, Merzlyak MN. Remote estimation of chlorophyll content in higher plant leaves. *Int J Remote Sens.* 1997;18:2691–7.
49. Dash J, Curran PJ. The MERIS terrestrial chlorophyll index. *Int J Remote Sens.* 2004;25:5403–13.
50. van Smeden M, Moons KGM, de Groot JAH, Collins GS, Altman DG, Eijkemans MJC, et al. Sample size for binary logistic prediction models: beyond events per variable criteria. *Stat Methods Med Res.* 2019;28:2455–74.
51. Sun HY, Wang SF, Hao XM. An improved analytic hierarchy process method for the evaluation of agricultural water management in irrigation districts of north China. *Agric Water Manag.* 2017;179:324–37.
52. Geng ZQ, Yang X, Han YM, Zhu QX. Energy optimization and analysis modeling based on extreme learning machine integrated index decomposition analysis: application to complex chemical processes. *Energy.* 2017;120:67–78.
53. Zou ZH, Yun Y, Sun JN. Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. *J Environ Sci.* 2006;18:1020–3.
54. Zhang ML, Li BZ. How to improve regional innovation quality From the perspective of green development? Findings from entropy weight method and Fuzzy-Set qualitative comparative analysis. *IEEE ACCESS.* 2020;8:32575–86.
55. Wang JJ, Jing YY, Zhang CF. Fuzzy multi-criteria evaluation model of HVAC schemes in optimal combination weighting method. *Build Serv Eng Res Technol.* 2009;30:287–304.
56. Fielding AH, Bell JF. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environ Conserv.* 1997;24:38–49.
57. Oliveira JS, Brown HE, Gash A, Moot DJ. An explanation of yield differences in three potato cultivars. *Agron J.* 2016;108:1434–46.
58. Sharma LK, Bali SK, Dwyer JD, Plant AB, Bhowmik A. A case study of improving yield prediction and sulfur deficiency detection using optical sensors and relationship of historical potato yield with weather data in maine. *Sensors.* 2017;17:1095.
59. Wang FL, Wang FM, Zhang Y, Hu JH, Huang JF, Xie JK. Rice yield estimation using parcel-level relative spectral variables from UAV-based hyperspectral imagery. *Front Plant Sci.* 2019;10:453.
60. Li B, Xu XM, Zhang L, Han JW, Bian CS, Li GC, et al. Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging. *ISPRS J Photogramm Remote Sens.* 2020;162:161–72.
61. Nemoto M, Hamasaki T, Matsuba S, Hayashi S, Yanagihara S. Estimation of rice yield components with meteorological elements divided according to developmental stages. *J Agric Meteorol.* 2016;72:128–41.
62. Fu ZP, Jiang J, Gao Y, Krienke B, Wang M, Zhong KT, et al. Wheat growth monitoring and yield estimation based on multi-rotor unmanned aerial vehicle. *Remote Sens.* 2020;12:508.



63. Zhou X, Zheng HB, Xu XQ, He JY, Ge XK, Yao X, et al. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS J Photogramm Remote Sens.* 2017;130:246–55.
64. Wang FM, Huang JF, Lou ZH. A comparison of three methods for estimating leaf area index of paddy rice from optimal hyperspectral bands. *Precision Agric.* 2011;12:439–47.
65. Rodríguez PL, Sanjuanelo CD, Núñez LCE, Moreno-Fonseca LP. Growth and phenology of three Andean potato varieties (*Solanum tuberosum* L.) under water stress. *Agron Colomb.* 2016;34:141–54.
66. Brandao ZN, Zonta JH. Hemispherical photography to estimate biophysical variables of cotton. *Rev Bra Eng Agric Ambient.* 2016;20:789–94.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

