International Journal of Applied Earth Observations and Geoinformation 101 (2021) 102362

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Improving the spatiotemporal fusion accuracy of fractional vegetation cover in agricultural regions by combining vegetation growth models

Guofeng Tao [a,b], Kun Jia [a,b,*], Xiangqin Wei [c], Mu Xia [a,b], Bing Wang [a,b], Xianhong Xie [a,b], Bo Jiang [a,b], Yunjun Yao [a,b], Xiaotong Zhang [a,b]

[a] State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[b] Beijing Engineering Research Center for Global Land Remote Sensing Products, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[c] Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, 100101, China

A R T I C L E I N F O

Keywords:
Spatiotemporal fusion
Vegetation growth model
Fractional vegetation cover
GF-1 WFV

A B S T R A C T

Spatiotemporal fusion has provided a feasible way to generate fractional vegetation cover (FVC) data with high spatial and temporal resolution. However, when the currently available spatiotemporal fusion methods are applied over agricultural regions, they usually underestimate high FVC values at the peak vegetation growth stage with medium FVC values as base data. This mainly results from inconsistencies in the temporal variations between fine- and coarse-resolution data if substantial temporal changes occur in vegetation. Therefore, a Spatial and Temporal Fusion method combining with Vegetation Growth Models (STF-VGM) was proposed to address this problem in this study, which incorporates vegetation growth models into the fusion process. Unlike other spatiotemporal fusion methods that mainly rely on changes in coarse-resolution data for prediction, STF-VGM fully utilizes available coarse- and fine-resolution time series data, including uncontaminated information in cloud/cloud shadow contaminated images. By establishing vegetation growth models with time series data, a conversion relationship between coarse- and fine-resolution FVC that changes along with the nonlinear vegetation change process can be extracted. STF-VGM makes prediction based on this variable relationship. A typical agricultural region located in the North China Plain was selected as the study area. The validation results indicated that the prediction accuracy for high FVC values was significantly improved using STF-VGM compared to the commonly used Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model (ESTARFM) and Flexible Spatiotemporal DAta Fusion (FSDAF) methods (STF-VGM: coefficient of determination (R^2) = 0.9491, root mean square error (RMSE) = 0.0650, average difference (AD) = -0.0092; ESTARFM: R^2 = 0.9341, RMSE = 0.1127, AD = -0.0631; FSDAF: R^2 = 0.9224, RMSE = 0.1110, AD = -0.0599). The satisfactory performance of STF-VGM was also achieved in predicting FVC values at other vegetation growth stages (early growth stage: R^2 = 0.8277, RMSE = 0.0440, AD = 0.0027; rapid growth stage: R^2 = 0.9183, RMSE = 0.0844, AD = 0.0500). In addition, STF-VGM also has the potential to improve the spatiotemporal fusion accuracy of other vegetation parameters and vegetation indices, which will be evaluated in the future.

1. Introduction

Fractional vegetation cover (FVC), referring to the fraction of green vegetation seen from nadir, is an important parameter to characterize vegetation conditions in the horizontal direction (Camacho et al., 2013; Jia et al., 2016). Agriculture monitoring is one of the main application fields of FVC (Jia et al., 2018). In the agricultural region, FVC follows the strong seasonal change patterns of crops and can be highly variable within a short period (Atzberger, 2013). Since the growth of crops and agricultural activities mostly occur at small scales, the fine-resolution monitoring of crop conditions is necessary. These characteristics of agricultural monitoring lead to the requirements for reliable FVC data with both high spatial and temporal resolutions.

Remote sensing provides an effective way to generate FVC data due

* Corresponding author at: State Key Laboratory of Remote Sensing Science, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China.
E-mail addresses: 201921051079@mail.bnu.edu.cn (G. Tao), jiakun@bnu.edu.cn (K. Jia), weixq@radi.ac.cn (X. Wei), xiamu@mail.bnu.edu.cn (M. Xia), icebergwb1995@mail.bnu.edu.cn (B. Wang), xianhong@bnu.edu.cn (X. Xie), bojiang@bnu.edu.cn (B. Jiang), yaoyunjun@bnu.edu.cn (Y. Yao), xtngzhang@bnu.edu.cn (X. Zhang).

https://doi.org/10.1016/j.jag.2021.102362
Received 15 November 2020; Received in revised form 23 April 2021; Accepted 10 May 2021

to its inherent ability to repeatedly observe the Earth's surface with wide coverage. There is a variety of remote sensing data for FVC generation, from low spatial resolutions (e.g., Advanced Very High Resolution Radiometer (AVHRR), Moderate Resolution Imaging Spectroradiometer (MODIS), and SPOT VEGETATION) (Bullock, 1992; Gutman and Ignatov, 1998; Jia et al., 2015; Propastin and Erasmi, 2010) to medium and high spatial resolutions (e.g., Landsat, Sentinel-2, and GaoFen-1 (GF-1)) (Hu et al., 2020; Jia et al., 2016; Tao et al., 2019; Wang et al., 2017) and from multispectral to hyperspectral (e.g., AVRIS, Hyperion, and HJ-1) (Zhang et al., 2013). Each sensor has its advantages and limitations. Commonly used large-scale FVC products are mainly derived from low spatial resolution data, such as the POLarization and Directionality of the Earth's Reflectances (POLDER) (Roujean and Lacaze, 2002), Medium Resolution Imaging Spectrometer (MERIS) (Baret et al., 2006), Carbon Cycle and Change in Land Observational Products from an Ensemble of Satellites (CYCOLPES) (Baret et al., 2007), GEOV1 (an improved product based on CYCOLPES) (Baret et al., 2013), and Global LAnd Surface Satellite (GLASS) (Jia et al., 2015; Liang et al., 2013) FVC products. Although these products can provide frequent monitoring, the spatial resolutions ranging from 300 m to 6 km prevent them from capturing detailed surface features that are necessary for regional applications at fine resolution. FVC generated from medium and high spatial resolution data is well suited to describe the spatial details, but its ability to collect frequent observations is usually not comparable to that of low spatial resolution data. The difficulty in obtaining high resolution data over both temporal and spatial dimensions from single satellite sensor has promoted the development of spatiotemporal fusion methods, which generate synthesized data with high temporal frequency from coarse-resolution (high temporal but low spatial resolution) data and rich spatial details from fine-resolution (low temporal but high spatial resolution) data (Zhu et al., 2018). With the recent emergence of new satellites, such as Sentinel-2 and GaoFen-1 Wide Field View (GF-1 WFV), the trade-off between spatial and temporal resolution has no longer been the main obstacle to obtaining high spatiotemporal resolution data. However, due to the late launch of these satellites (e.g., GF-1 WFV was launched in 2013 and Sentinel-2 was launched in 2015), there are still limitations in acquiring historical dense time series data from sub 100 m sensors for long-term studies. In addition, contaminations from clouds, cloud shadows and other unfavorable weather conditions can lengthen the time interval between two valid observations, which is generally longer than a month instead of the designed short revisit cycles. Therefore, spatiotemporal fusion is still of great significance for FVC monitoring in agricultural regions.

A number of spatiotemporal fusion methods have been proposed over the past several years. Based on at least one pair of coarse- and fine-resolution data acquired on the same date or temporally close dates, typical fusion methods can predict the high spatial resolution data on desirable dates as long as the corresponding coarse-resolution data are available, so as to improve the monitoring frequency at high spatial resolution. Spatiotemporal fusion was originally designed for land surface reflectance data, such as the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM), which was proposed by Gao et al. (Gao et al., 2006). STARFM conducts prediction by adding a weight combination of temporal changes derived from neighboring coarse-resolution pixels to the fine-resolution pixels on base date (the date with a known pair of coarse- and fine-resolution data). Coarse-resolution pixels that are mixed by lesser land cover type can get higher weights and contribute more to the prediction. Therefore, STARFM is more suitable for homogeneous landscapes. To improve the application ability in heterogeneous areas, STARFM was modified to various methods, such as the Spatial Temporal Adaptive Algorithm for mapping Reflectance CHange (STAARCH (Hilker et al., 2009)), the Enhanced STARFM (ESTARFM (Zhu et al., 2010)), Robust Adaptive Spatial and Temporal Fusion Model (RASTFM (Zhao et al., 2018)), Inpainting-based Steering Kernel Regression Fusion Model (ISKRFM (Wu et al., 2017)). Because of the involvement of the weight function, these STARFM-like methods are commonly categorized as weight function-based methods (Zhu et al., 2018). Based on the increasing demand for monitoring vegetation dynamics using time series data at high temporal resolutions, these STARFM-like approaches have been extended to directly fuse the vegetation indices and vegetation parameters, such as Normalized Difference Vegetation Index (NDVI) (Tian et al., 2013), Leaf Area Index (LAI) (Houborg et al., 2016), and FVC (Tao et al., 2019). Unmixing-based methods, another type of widely used spatiotemporal fusion method, estimate the values of fine-resolution pixels or the changes in fine-resolution pixels by unmixing coarse-resolution pixels using linear spectral unmixing theory. Because there may be large errors in the spectral unmixing process (Zhu et al., 2018), unmixing-based methods are usually combined with constraint rules or other methods, such as weight function-based methods. The Flexible Spatiotemporal DAta Fusion (FSDAF) algorithm is a representative of this type (Zhu et al., 2016). FSDAF combines the ideas of spectral unmixing analysis and weight assignment, having advantages in fusion tasks when changes in land cover type occur. Recently, FSDAF has been modified and improved to several methods, such as the improved FSDAF method for time series NDVI production (IFSDAF) (Liu et al., 2019b), the modified FSDAF method that uses downscaled MODIS data as input (Zhai et al., 2020), and the enhanced FSDAF method that considers subpixel class fraction change information (SFSDAF) (Li et al., 2020). In addition, various recent studies have focused on the temporal evolution in coarse-resolution data and utilized the entire coarse-resolution time series dataset instead of only using the coarse-resolution data on few pair dates and prediction date. For example, Liu et el. (Liu et al., 2018) proposed a modified ESTARFM method that extracts phenological information from coarse-resolution time series data to guide the similar pixel selection. STGDFM (Kim et al., 2020), a recently proposed method, uses the land-cover-specific temporal profiles from coarse-resolution time series to consider the temporal change information.

All these methods can increase the density of fine-resolution data in the temporal dimension and assist in vegetation monitoring. However, these methods face challenges if strong temporal changes in fine-resolution pixels cannot be fully captured by the corresponding coarse-resolution pixels. This problem is particularly evident when these spatiotemporal fusion methods are used for agricultural regions. For instance, in an agricultural area where many residential patches of various shapes and sizes are randomly distributed, the information in a mixed coarse-resolution pixel is contributed by residential patches with slight changes and crops with substantial changes during the growing season (e.g., from medium FVC to high FVC values). Therefore, the large changes in crops are difficult to observe fully with coarse-resolution data. For most spatiotemporal fusion methods, the fusion process is to reproduce the variations that occur in fine-resolution pixels. However, the acquisition of such variations with these methods relies entirely on the variations in coarse-resolution pixels during the prediction period. STARFM considers the linear weighted combination of coarse-resolution variations as the predicted variations in fine-resolution pixels. Unmixing-based methods, such as FSDAF, obtain the variations that occur in fine-resolution pixels by unmixing the variations in their corresponding coarse-resolution pixels. Even though some residual compensation terms are added to modify the results, the errors caused by spectral unmixing still have a great impact on the final prediction. The fusion methods that use coarse-resolution time series also face such limitations. To overcome this problem, auxiliary fine-resolution information needs to be considered. ESTARFM includes the information on the changes in fine-resolution data in the prediction instead of using only the changes in coarse-resolution data. However, the assumptions of linear reflectance changes and the constant conversion coefficient in ESTARFM are not appropriate for the significant nonlinear changes that exist during the prediction period. Therefore, due to the assumptions and inherent designs of the abovementioned methods, their performances are unsatisfactory when temporal changes in coarse- and fine-resolution pixels occur to different degrees. In particular, because of the inadequacy of capturing the strong fine-resolution variations, there may be significantly underestimated predictions on the peak vegetation growth stage if the medium values are used as baseline. Furthermore, cloud contamination is frequently present at the peak vegetation growth stage in agricultural regions; thus, the fusion accuracy must be significantly improved in such situations.

To cope with the abovementioned problems and improve the spatiotemporal fusion accuracy of FVC, a Spatial and Temporal Fusion method combining with Vegetation Growth Model (STF-VGM) is proposed in this study, which incorporates vegetation growth models into the fusion process. In addition to coarse-resolution time series data, the available fine-resolution time series data, including partly uncontaminated data, are also employed in STF-VGM. Based on time series FVC data, vegetation growth models at coarse- and fine-resolutions are built, and the nonlinear vegetation changes are considered. Then, a conversion relationship between coarse- and fine-resolution FVC data that changes along with the vegetation growth process can be obtained through the vegetation growth models, and the prediction process is based on this variable relationship. To assess the effectiveness of STF-VGM, ESTARFM method and the unmixing-based method FSDAF are compared in this study. The GaoFen-1 Wide Field View (GF-1 WFV) sensors launched by China provide multispectral data with 16 m spatial resolution, which are valuable data sources for describing small objects. Therefore, FVC derived from GF-1 WFV reflectance data are selected as the fine-resolution data. For the coarse-resolution FVC data, the available GLASS FVC product, which was generated from MODIS and AVHRR reflectance data, is a suitable choice.

2. Method

2.1. Estimating FVC from GF-1 WFV data

There are three main types of FVC estimation methods from remote sensing data, including empirical methods, pixel unmixing methods, and physical model-based methods (Jia et al., 2016). Considering its good generalization ability, a physical model-based method is adopted to estimate FVC from GF-1 WFV reflectance data, which is described in detail in a previous study (Tao et al., 2019). The development of this method consists primarily of three parts. First, based on the radiative transfer model PROSAIL (Jacquemoud and Baret, 1990; Jiménez-Muñoz et al., 2009; Verhoef, 1984), a simulated dataset containing FVC and the corresponding vegetation canopy reflectance under various conditions is generated. Then, the simulated dataset is used to train a random forest regression (RFR) model to learn the relationships between FVC and canopy reflectance. Finally, the trained RFR model is used to estimate FVC from real GF-1 WFV reflectance data. Because the effectiveness and accuracy of this FVC estimation method have been verified in a previous study (Tao et al., 2019), its performance is not discussed here.

2.2. STF-VGM

Fig. 1 shows a flowchart of the proposed STF-VGM method. The input data include all the available coarse- and fine-resolution FVC data during a specific period (e.g., one crop growing season). The dates when fine-resolution FVC data are available are denoted as $(\cdots, t_{base1}, t_{base2}, t_{base3},$

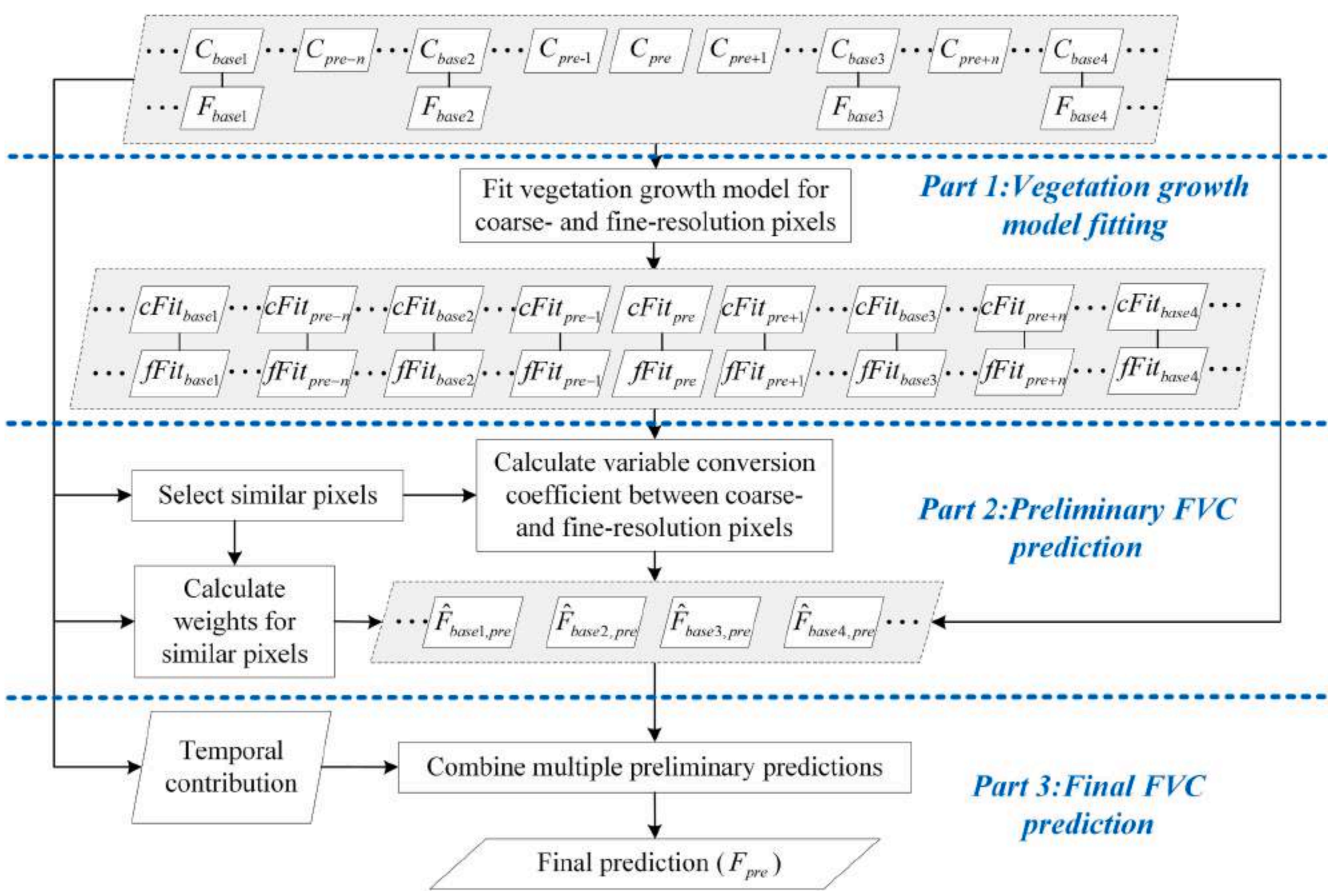

Fig. 1. Flowchart of STF-VGM.

$t_{base4}, \cdots)$, and the pairs of coarse- and fine-resolution FVC data on these dates are denoted as $(\cdots, C_{base1}, C_{base2}, C_{base3}, C_{base4}, \cdots)$ and $(\cdots, F_{base1}, F_{base2}, F_{base3}, F_{base4}, \cdots)$, respectively. The other dates in the coarse-resolution data time series when the corresponding fine-resolution FVC data are not available are denoted as $(\cdots, t_{pre-n}, \cdots, t_{pre-1}, t_{pre}, t_{pre+1}, \cdots, t_{pre+n}, \cdots)$, and the coarse-resolution FVC data on these dates are $(\cdots, C_{pre-n}, \cdots, C_{pre-1}, C_{pre}, C_{pre+1}, \cdots, C_{pre+n}, \cdots)$. STF-VGM can be divided into three parts: vegetation growth model fitting, preliminary FVC prediction, and final FVC prediction. In the first part, the vegetation growth model for each coarse- and fine-resolution pixel is fitted using the time series data. The input fine-resolution FVC image does not need to be entirely free of clouds, but there should be at least four valid values at different times for each fine-resolution pixel to fit the vegetation growth model. For the preliminary FVC prediction, each input coarse–fine pair is a "base pair" and can generate its corresponding result at the prediction date, which is considered as the "preliminary prediction". Specifically, there are three major steps in this part. The first step is the selection of similar pixels for each fine-resolution pixel. The second step is the weight calculation for similar pixels. The third step is the estimation of preliminary predictions using the conversion coefficient that changes with the nonlinear vegetation growth process. These three steps are conducted for each base pair separately, and multiple preliminary predictions$(\cdots, \widehat{F}_{base1,pre}, \widehat{F}_{base2,pre}, \widehat{F}_{base3,pre}, \widehat{F}_{base4,pre}, \cdots)$ can be obtained. In the calculation of the final FVC prediction F_{pre}, all preliminary predictions are combined with their temporal weights. The details of these implementations are introduced in the following sections.

2.2.1. *Vegetation growth model*

The vegetation growth model describes the characteristics of the temporal changes in vegetation. The construction of a reasonable vegetation growth model is a key step to ensure the prediction accuracy of the proposed method. To date, two types of vegetation growth models have been developed: the first type includes mechanical models with clear physical mechanism, reflecting the nature of vegetation growth (Jones et al., 2003; Ritchie, 1985; Boogaard et al., 1998); the other includes empirical models based mainly on statistical analyses of time series data (Bindraban, 1999; Wu et al., 2003). However, the complexity of the natural principles and the large number of input parameters make mechanical models inconvenient to use. Empirical models are more flexible and easier to build than mechanical models due to the accessibility of time series data (field measurements or remote sensing products). Therefore, an empirical model is adopted in this study. As an intrinsic parameter describing vegetation conditions (Wang et al., 2016), FVC is suitable for constructing a statistical vegetation growth model because it changes regularly with the vegetation growth process. The modified Verhulst logistic equation (Lin et al., 2003) is selected to construct the vegetation growth model:

$$FVC = \frac{d}{1 + exp(a \times t^2 + b \times t + c)} \tag{1}$$

where a, b, c and d are coefficients that can be fitted using FVC time series data and t is the day of year (DOY).

For each coarse- and fine-resolution vegetation pixel, a unique model is established. The widely used Levenberg-Marquardt (LM) algorithm, which combines the steepest descent and the Gauss–Newton methods, is employed to fit the model parameters (Lourakis, 2005). It is meaningless to fit vegetation growth models for non-vegetation pixels or parse-vegetation pixels with low FVC values and inapparent temporal changes. Therefore, a threshold is set to determine whether a fit is required. Considering the observed valid FVC values for each pixel are obtained during the whole growing season, if the average of all valid FVC values of a fine-resolution pixel is less than 0.15, its vegetation growth model will not be fitted. The FVC changing curves of such pixels are directly obtained by the temporal interpolation function.

2.2.2. *Estimation of preliminary prediction*

For simplicity, suppose there is one pair of coarse- and fine-resolution data C_{base} and F_{base} acquired at t_{base} and one coarse-resolution data C_{pre} acquired at t_{pre}. The task of estimating a preliminary prediction is to generate a synthesized fine-resolution image $\widehat{F}_{base,pre}$ at t_{pre} from the data at t_{base}.The main idea can be represented as the following formula:

$$\widehat{F}_{base,pre} = F_{base} + \Delta F \tag{2}$$

The preliminary FVC prediction $\widehat{F}_{base,pre}$ is the sum of the FVC value at t_{base} and the fine-resolution variation (ΔF) between t_{base} and t_{pre}. Accordingly, to estimate $\widehat{F}_{base,pre}$, ΔF should be first calculated.

Due to the heterogeneity of the land surface, most of the coarse-resolution pixels are mixed by multiple signals from various land cover types. The diversities in inter-class and intra-class make the fine-resolution pixels within the same coarse-resolution pixel show different variations between t_{base} and t_{pre}. Therefore, the conversion coefficient, indicating the ratio of the FVC change in a fine-resolution pixel to the FVC change in the corresponding mixed coarse-resolution pixel (Zhu et al., 2010), is used to estimate the changes in different fine-resolution pixels. The conversion coefficient for each fine-resolution pixel is unique. Due to the consideration of the nonlinear growth pattern of crops, the conversion coefficient in STF-VGM is variable following the process of crop growth, which is the key of STF-VGM.

STF-VGM follows the idea of weight function-based methods that predict a pixel value using the auxiliary information from its neighboring fine-resolution similar pixels (referred to as similar pixels hereafter) and their corresponding weights.

In a given local window, ΔF at the central target pixel is predicted by:

$$\Delta F(x,y) = \sum_{i=1}^{N} W_i \times V_i \times \left(C_{pre}(x_i,y_i) - C_{base}(x_i,y_i)\right) \tag{3}$$

where (x,y) represents the central target pixel, w is the size of the local window centered at (x,y), (x_i,y_i) represents the ith similar pixel and N is the number of similar pixels, respectively. C means the resampled coarse-resolution data whose spatial resolution is the same as that of the fine-resolution data, and C_{pre} and C_{base} represent the data acquired on prediction date and base date. W_i is the weight for the ith similar pixel. V_i is the conversion coefficient at the ith similar pixel.

To get the key parameters (W_i and V_i) in Eq. (3), three steps are implemented as the following sequence: selection of similar pixels, weight calculation of similar pixels and estimation of preliminary prediction using the variable conversion coefficient. These processes are described in detail in the following subsections.

2.2.2.1. Selection of similar pixels. Integrating the information from neighboring similar fine-resolution pixels helps to improve the fusion accuracy and remove the cell boundaries from coarse-resolution data in the fine-resolution prediction (Gao et al., 2015). Similar pixels play an important role in two aspects of STF-VGM. One aspect is to obtain the conversion coefficients between fine- and coarse-resolution data for their central pixels and themselves, and the other is to reduce the uncertainty in the prediction. It is worth noting that the growth stage of vegetation is constantly changing. Although FVC between two pixels are similar on one date, such similarity is not promised on other dates. Therefore, the criterion for similar pixel selection on a base image F_{base} in this study is: (a) the pixel values are similar to the central pixel on F_{base}, and (b) the temporal changing trajectories of pixels should be close to that of the central pixel.

Fig. 2 shows a schematic diagram of the similar pixel selection for a pixel. The selection principle is based on the following equations:

$$S_{base}(x_l,y_l) = |F_{base}(x_l,y_l) - F_{base}(x,y)| \tag{4}$$

$$S_{all}(x_l, y_l) = \sum_{j=1}^{m} |F_j(x_l, y_l) - F_j(x, y)| \tag{5}$$

$$Rank_{S_{base}} = \{Rank_{S_{base}}(x_1, y_1), Rank_{S_{base}}(x_2, y_2), \cdots, Rank_{S_{base}}(x_l, y_l), \cdots, Rank_{S_{base}}(x_{w^2}, y_{w^2})\} \tag{6}$$

$$Rank_{S_{all}} = \{Rank_{S_{all}}(x_1, y_1), Rank_{S_{all}}(x_2, y_2), \cdots, Rank_{S_{all}}(x_l, y_l), \cdots, Rank_{S_{all}}(x_{w^2}, y_{w^2})\} \tag{7}$$

$$Rank(x_l, y_l) = Rank_{S_{base}}(x_l, y_l) \times Rank_{S_{all}}(x_l, y_l) \tag{8}$$

$$Rank = \{Rank(x_1, y_1), Rank(x_2, y_2), \cdots, Rank(x_l, y_l), \cdots, Rank(x_{w^2}, y_{w^2})\} \tag{9}$$

where $F_j (j = 1, 2, \ldots, m)$ denotes the jth fine-resolution FVC image in the time series, m is the number of fine-resolution FVC images. The subscript of x and y represents the identification of a pixel within the search window. For example, (x_1, y_1) is the first neighboring pixel of the central target pixel (x, y) within the search window, and (x_l, y_l) represents the lth one (Fig. 2). $S_{base}(x_l, y_l)$ is the absolute FVC difference between (x_l, y_l) and (x, y) on the base date, and $S_{all}(x_l, y_l)$ is the sum of the absolute FVC difference between (x_l, y_l) and (x, y) on all dates during the time series of fine-resolution FVC. These two indicators reflect the similarity between (x_l, y_l) and (x, y). Each fine-resolution pixel in the search window has its own S_{base} and S_{all}. All the pixels in the search window are sorted from small to large based on S_{base} and S_{all}, respectively, and the two vectors $Rank_{S_{base}}$ and $Rank_{S_{all}}$ contain the corresponding sorting orders of (x_1, y_1), $(x_2, y_2), \cdots, (x_l, y_l), \cdots, (x_{w^2}, y_{w^2})$. As the example shown in Eq. (8), $Rank_{S_{base}}$ and $Rank_{S_{all}}$ are combined through multiplication, and a vector $Rank$ containing the combination result can be obtained. The first N (the number of similar pixels) pixels with the smallest $Rank$ values are selected as the final similar pixels for $F_{base}(x, y)$. The introduction of sorting order vectors and their multiplication ensures that F_{base} still plays an important role in the selection of similar pixels when considering the information from all available fine-resolution FVC data.

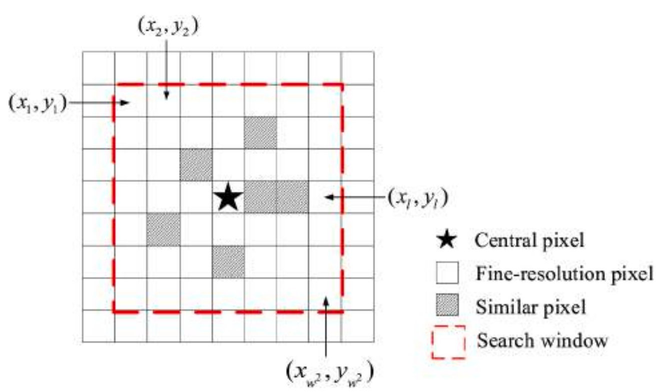

Fig. 2. Schematic diagram of the similar pixel selection for a pixel.

Fig. 3 shows a comparison between three principles: selecting the first 20 pixels with the smallest $Rank_{S_{base}}$ (*Principle 1*), $Rank_{S_{all}}$ (*Principle 2*) and $Rank$ (*Principle 3*) as similar pixels, respectively. The asterisks (denoted as '*') are true FVC of the selected pixels on the five fine-resolution data acquisition dates, and one color represents one pixel. The curves are the fitted vegetation growth curves of each pixel and can reflect the temporal change. For the base date marked as red in Fig. 3, the FVC values of pixels selected based on $Rank_{S_{base}}$ differ little on this date, whereas they are quite different on other dates, indicating that these selected pixels show inconsistent phenological changes and fail to meet the second criterion for similar pixels. Although pixels selected based on $Rank_{S_{all}}$ are similar on each date, their differences on base date are slightly larger than those of pixels selected based on $Rank_{S_{base}}$ (Fig. 3 (b)). While pixels selected based on $Rank$ not only have little difference on base date like *Principle 1*, but also show similar temporal changing trajectories, indicating the advantages of using $Rank$ for similar pixel selection.

2.2.2.2. Weight calculation of similar pixels. Different similar pixels have different contributions to the prediction of the central pixel value. Therefore, a weight coefficient for each similar pixel is necessary. Because the FVC values of the selected pixels are generally very similar, the FVC similarity between the central pixel and similar pixels cannot provide effective guidance for the weight assignment. Therefore, the weight for a similar pixel is determined according to the spatial distance. Pixels spatially closer to the central pixel have higher weights. The weight W_i for the ith similar pixel can be described as follows:

$$W_i = \frac{1/D_i}{\sum_{i=1}^{N} (1/D_i)} \tag{10}$$

$$D_i = 1 + \sqrt{(x_i - x)^2 + (y_i - y)^2}/(w/2) \tag{11}$$

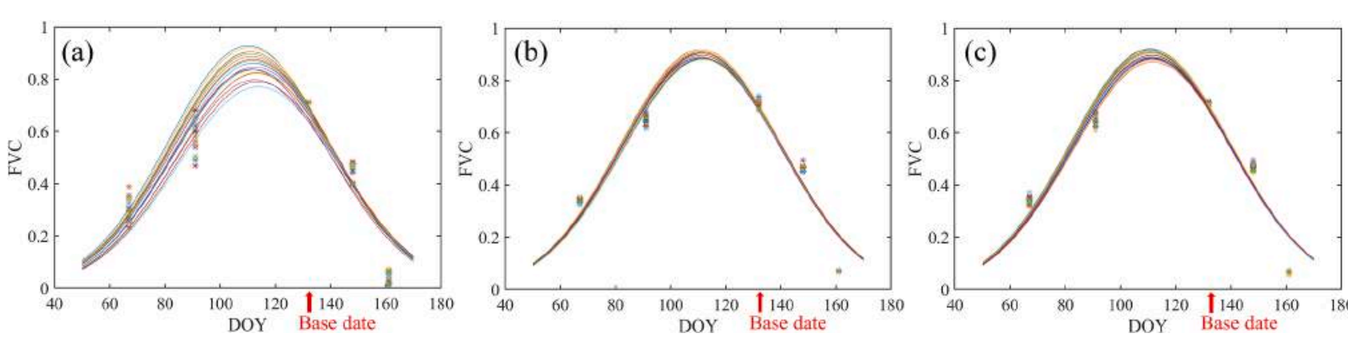

Fig. 3. The fitted vegetation growth curves of the selected 20 similar pixels for the same target pixel based on the three selection principles: $Rank_{S_{base}}$ (a), $Rank_{S_{all}}$ (b) and $Rank$ (c). The asterisks (denoted as '*') in same color represent the FVC values of the same selected similar pixel on its different acquired dates.

where N is the number of similar pixels, D_i is the spatial distance between the central pixel (x, y) and the ith similar pixel (x_i, y_i), which can be calculated from Eq. (11). The value 1 and $(w/2)$ are introduced to normalize the spatial distance such that D_i always ranges from 1 to $1+\sqrt{2}$ in different search window sizes.

2.2.2.3. Estimation of preliminary prediction using the variable conversion coefficient. The conversion coefficient V_i is defined as:

$$V_i = \frac{\Delta F_i}{\Delta C_i} \tag{12}$$

where ΔF_i and ΔC_i are the fine- and coarse-resolution variations occurring at the ith similar pixel during the same period. As evidenced in ESTARFM (Zhu et al., 2010), V_i is constant with the assumption that the proportion of each endmember within the coarse-resolution pixel and the change rate of each endmember are stable during a relatively short period. The vegetation growth model can fit FVC at any date, making it possible to establish the relationship between coarse- and fine-resolution FVC over any short-term. Therefore, ΔF_i and ΔC_i can be obtained from the growth model fitting results, and V_i can be concretely described as follows:

$$V_i = \frac{fFit_m(x_i, y_i) - fFit_n(x_i, y_i)}{cFit_m(x_i, y_i) - cFit_n(x_i, y_i)} \tag{13}$$

where $fFit_m(x_i, y_i)$ and $fFit_n(x_i, y_i)$ are the fitted fine-resolution FVC values of the ith similar pixel at t_m and t_n, respectively. $cFit_m(x_i, y_i)$ and $cFit_n(x_i, y_i)$ are the fitted coarse-resolution FVC values of the ith similar pixel at t_m and t_n, respectively.

For the calculation of the conversion coefficient, similar pixels are utilized to reduce uncertainty. Linear regression analysis between the fitted coarse- and fine-resolution FVC data at t_m and t_n is conducted on similar pixels within the same coarse-resolution pixel. The slope of the linear regression is considered as the conversion coefficient for similar pixels. Similar pixels within the same coarse-resolution pixel share the same conversion coefficient. Because the search window size for selecting similar pixels is usually set to be larger than the coarse-to-fine-resolution ratio, similar pixels may be distributed in different coarse-resolution pixels. In this case, these similar pixels are divided according to their corresponding coarse-resolution pixels. Linear regressions for similar pixels located in different coarse-resolution pixels are conducted separately.

The definition of conversion coefficient in Eq. (12) and Eq. (13) is suitable for short-term in which the change of vegetation can satisfy the linear approximation. In practice, for the given base date t_{base} and prediction date t_{pre}, it is usually to divide the period t_{base}-t_{pre} into several short-terms that are equivalent to the temporal resolution of coarse data. In this case, the conversion coefficients for these short-terms can form a variable V_i that changes along with the nonlinear vegetation growth. Specifically, assuming that t_{base}, t_k , and t_{pre} are three successive dates of coarse-resolution observations, the preliminary prediction $\widehat{F}_{base,pre}$ can be obtained by generating the interim prediction at t_k. This process is described as follows:

$$\widehat{F}_{base,pre} = F_{base} + \Delta F_{base,k} + \Delta F_{k,pre} \tag{14}$$

where $\Delta F_{base,k}$ and $\Delta F_{k,pre}$ are the fine-resolution variations during t_{base}-t_k and t_k-t_{pre}, respectively, which can be calculated from their corresponding conversion coefficients and coarse-resolution variations through Eq. (3). In this way, the changes in vegetation growth could be described as a nonlinear pattern.

2.2.3. Combination of multiple preliminary predictions

Based on different base pairs, multiple preliminary predictions, $\cdots$, $\widehat{F}_{base1,pre}, \widehat{F}_{base2,pre}, \widehat{F}_{base3,pre}, \widehat{F}_{base4,pre}, \cdots$, can be generated. A final prediction with less uncertainty may be obtained by a weighted combination of these multiple preliminary predictions. The weight for each preliminary prediction is mainly determined by two aspects: the absolute FVC difference and the temporal distance between the coarse-resolution data at its base date and the prediction date. The calculation process is still based on the search window in Section 2.2.2. Accordingly, the weight for a preliminary prediction $\widehat{F}_{base,pre}$ can be calculated by the following equations:

$$T1_{base,pre}(x, y) = \frac{1/\left|\sum_{l=1}^{w^2} C_{base}(x_l, y_l) - \sum_{l=1}^{w^2} C_{pre}(x_l, y_l)\right|}{\sum_{j=1}^{m}\left(1/\left|\sum_{l=1}^{w^2} C_j(x_l, y_l) - \sum_{l=1}^{w^2} C_{pre}(x_l, y_l)\right|\right)} \tag{15}$$

$$T2_{base,pre}(x, y) = \frac{1/|t_{base} - t_{pre}|}{\sum_{j=1}^{m}\left(|t_j - t_{pre}|\right)} \tag{16}$$

$$T_{base,pre}(x, y) = \frac{T1_{base,pre}(x, y) \times T2_{base,pre}(x, y)}{\sum_{j=1}^{m}\left(T1_{j,pre}(x, y) \times T2_{j,pre}(x, y)\right)} \tag{17}$$

where (x_l, y_l) is the location of the lth pixel in the search window, $C_{base}(x_l, y_l)$ and $C_{pre}(x_l, y_l)$ are the coarse-resolution FVC values at base date t_{base} and prediction date t_{pre}, m is the number of preliminary predictions as well as the number of base pairs, and $C_j(x_l, y_l)$ represents the lth resampled coarse-resolution FVC value of the jth base pair. $T1_{base,pre}(x, y)$ and $T2_{base,pre}(x, y)$ are the weights calculated by the temporal FVC difference in coarse-resolution data and the temporal distance of dates, respectively. $T_{base,pre}(x, y)$, the combination of $T1$ and $T2$, is the final temporal weight for $\widehat{F}_{base,pre}$. Based on this weight, the combination of multiple preliminary predictions can be described as follows:

$$F_{pre} = \sum_{j=1}^{m}\left(T_{j,pre} \times \widehat{F}_{j,pre}\right) \tag{18}$$

Considering that the FVC acquired at a date too far from the prediction date is weakly correlated to the predicted FVC, base pairs with a temporal distance of more than two months to the prediction date are not introduced to the combination. In addition, different base pairs involved in Eq. (18) could lead to different accuracies of the final prediction. The determination of the base pairs will be discussed in the following sections.

3. Case study to validate the STF-VGM method

3.1. Study area and field-measured FVC

Hengshui city, located in Hebei Province, northern China (115°10′E ~ 116°34′E, 37°03′N ~ 38°23′N) (Fig. 4), is a typical agricultural region where the vegetation condition changes dramatically. The patches of cropland and human settlement area are irregularly staggered in this area. Thus, spatial heterogeneity also exists in this area, making it a suitable choice to evaluate the performance of the proposed method.

The field-measured FVC data for winter wheat were collected at the sample sites distributed in 11 counties of Hengshui (Fig. 4(c)). There were two ground survey periods. The first period lasted from March 29, 2017 to April 1, 2017, during which the FVC values of winter wheat were generally medium. The second period was from May 4, 2017 to May 6, 2017, and the winter wheat was at the peak growth stage with high FVC. Two sample sites were selected in each county. For each sample site, five sample points with sizes of 30 m × 30 m covering relatively homogeneous areas were selected.

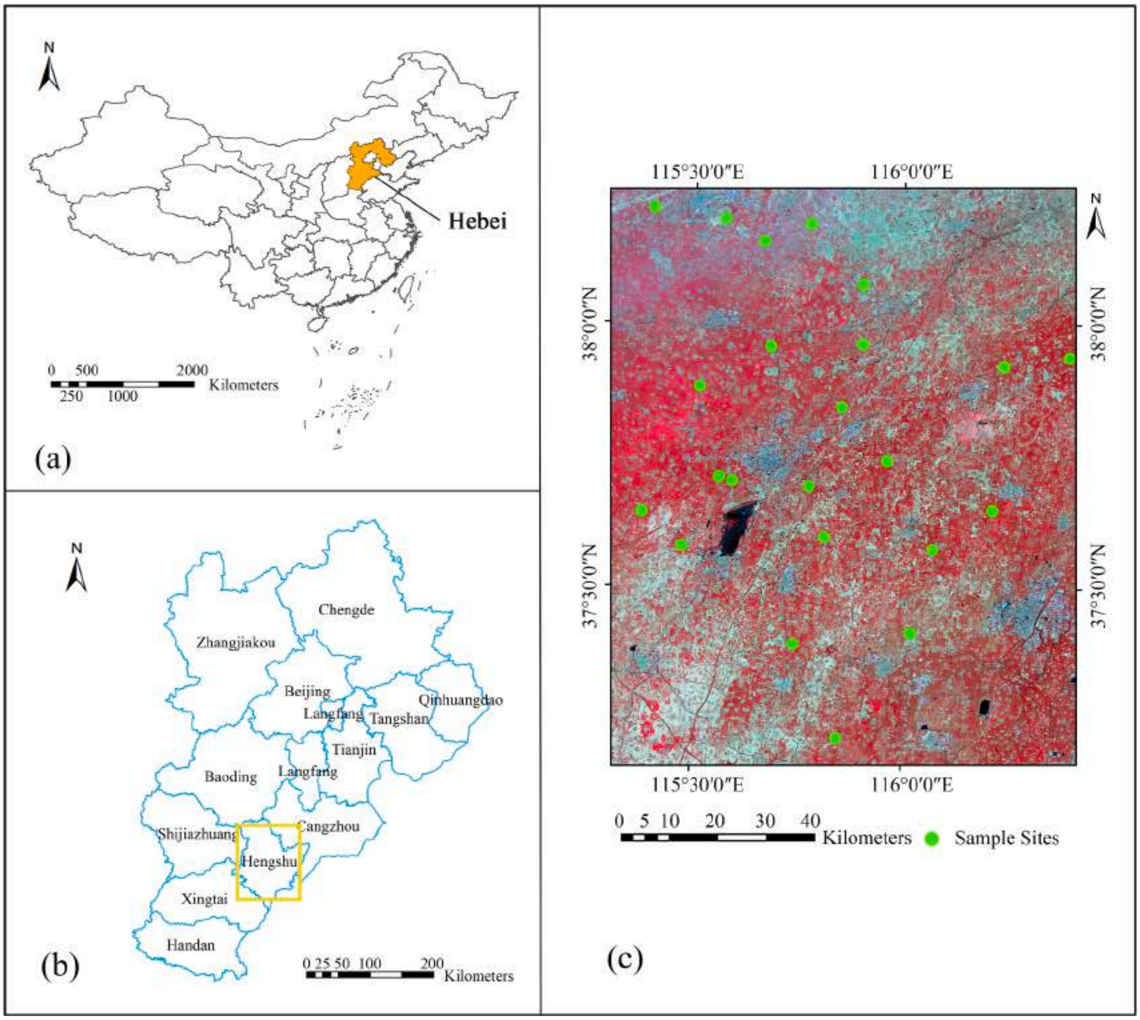

Fig. 4. Geographic location of the study area is marked by the yellow rectangle in (b), and the distribution of the sample sites of field-measured FVC are represented by the solid green circles on the standard false color composited GF-1 WFV data in (c). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2. Data and preprocessing

The fine-resolution FVC data were generated from the Chinese GF-1 WFV multispectral data, which were designed with a revisit cycle of 4 days and a spatial resolution of 16 m. However, due to the unfavorable weather conditions and occasionally poor quality of the acquired data, the time interval between two cloud-free GF-1 WFV data for specific regions is generally longer than one month instead of 4 days. Therefore, it is necessary to conduct fusion to the GF-1 WFV-derived FVC data. Seven GF-1 WFV data were collected during the winter wheat growing season in Hengshui, and their acquisition dates are shown in the timeline in Fig. 5. The preprocessing of GF-1 WFV data involved radiance calibration, atmospheric correction and geometric correction. The method proposed in Section 2.2 was used to estimate the FVC from the preprocessed GF-1 WFV reflectance data. GF-1 WFV FVC data on April 26, 2017 (DOY116), April 1, 2017 (DOY91) and March 8, 2017 (DOY67) were selected as the reference images for accuracy validation, which represented high FVC values at peak vegetation growth stage, medium FVC values at the fast-growing stage and low FVC values at early growth stage, respectively.

The GLASS FVC product was selected as the coarse-resolution data due to its satisfactory accuracy and spatiotemporal continuities (Jia et al. 2018; Jia et al. 2019; Liu et al. 2019a,b). GLASS FVC data were generated from the MODIS surface reflectance product MOD09A1, with

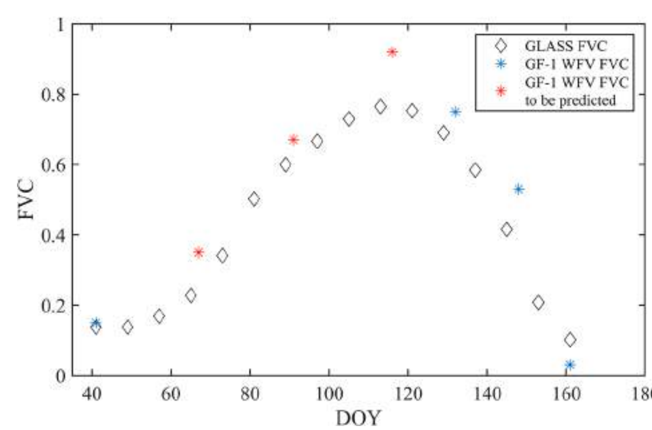

Fig. 5. Acquisition dates of GF-1 WFV and GLASS FVC and the corresponding FVC values at a randomly selected position in the study area. The red asterisks denote the values to be predicted in the following experiments. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

a temporal resolution of 8 days and a spatial resolution of 500 m. Prior to the production of GLASS FVC, cloud-cleared MODIS reflectance data were obtained by preprocessing MOD09A1 using the temporal-spatial filtering method developed by Tang et al. (Tang et al., 2013), which ensured that the GLASS FVC data were spatially and temporally complete. Thirteen GLASS FVC data were used (Fig. 5), and all of them were resampled to the same spatial resolution (16 m) as GF-1 WFV FVC data. Additionally, Fig. 5 shows that the acquisition dates of GLASS FVC could not always perfectly match those of GF-1 WFV FVC. Hence, the cubic spline interpolation method was used to match the acquisition dates of GLASS FVC and GF-1 WFV FVC. In addition, it is obvious that the FVC change in the coarse-resolution pixel is slight between DOY 90 and DOY 130, while the change in the fine-resolution pixel is quite strong. This is a typical phenomenon of the previously mentioned situation where the changes in coarse- and fine-resolution pixels vary greatly.

Three indices are selected for quantitative accuracy assessment: the coefficient of determination R^2, root mean square error (RMSE), and average difference (AD). The ideal values of R^2, RMSE, and AD are 1, 0, and 0, respectively.

3.3. Performance comparison of different base pairs for the final prediction

To evaluate the performance of different base pairs employed for the final prediction, experiments with one, two (before and after the prediction), and multiple base pairs were conducted for the three prediction dates at different vegetation growth stages. For simplicity and convenience, notations of the predictions were created to represent the experimental results (Table 1). Data on June 10, 2017 were not selected as a base pair because the vegetation condition was largely different from the predictions. A region with 400 × 400 GF-1 WFV pixels was extracted from the study area for performance comparison.

Fig. 6 provides a visual comparison between FVC predictions on April 26, 2017, based on different numbers of input pairs. All the predictions display generally high similarities to the actual FVC image, indicating the effectiveness of the proposed STF-VGM method. There is little difference between the four strategies to restore spatial details, whereas they show different performances in predicting temporal changes. *Pre0426-1–0512* and *Pre0426-2* more closely resemble the actual FVC. Scatter plots and quantitative assessments (Fig. 7) also illustrate that these two predictions achieve higher accuracies than *Pre0426-1–0401* and *Pre0426-4*. *Pre0426-1–0401* and *Pre0426-1–0512* are the two preliminary predictions for *Pre0426-2*. As evidenced by both the visual performance (Fig. 6) and the scatter plots (Fig. 7), *Pre0426-2* is closer to *Pre0426-1–0512* than Pre0426-1–0401. This is because the vegetation condition on April 26 was more similar to that on May 12, and the difference in GLASS FVC between April 26 and May 12 was less, resulting in a larger temporal weight to *Pre0426-1–0512*. Therefore, *Pre0426-1–0512* plays the dominant role in the combination of the two preliminary predictions. Although in terms of assessment metrics, *Pre0426-1–0512* slightly outperforms *Pre0426-2* and *Pre0426-4*, its uncertainty is larger in some places than *Pre0426-2* and *Pre0426-4*. The scatter plots (Fig. 7) show that there are some gathering points with obviously large deviations from the reference line near the coordinates of (0.3, 0.6) and (0.8,0.6) in Fig. 7 (b), whereas such points in Fig. 7 (c) are much less due to the integrated accurate information from Fig. 6 (a) near this coordinate.

Table 1
Dates of the different base pairs for fusion.

Prediction date	Number of base pairs	Dates of base pairs	Notations for the predictions
2017–04-26	1	2017–04-01	*Pre0426-1–0401*
	1	2017–05-12	*Pre0426-1–0512*
	2	2017–04-01 and 2017–05-12	*Pre0426-2*
	4	2017–03-08, 2017–04-01, 2017–05-12 and 2017–05-28	*Pre0426-4*
2017–04-01	1	2017–03-08	*Pre0401-1–0308*
	1	2017–04-26	*Pre0401-1–0426*
	2	2017–03-08 and 2017–04-26	*Pre0401-2*
	4	2017–03-08, 2017–04-26, 2017–05-12 and 2017–05-28	*Pre0401-4*
2017–03-08	1	2017–02-10	*Pre0308-1–0210*
	1	2017–04-01	*Pre0308-1–0401*
	2	2017–02-10 and 2017–04-01	*Pre0308-2*
	5	2017–02-10, 2017–04-01, 2017–04-26, 2017–05-12 and 2017–05-28	*Pre0308-5*

A visual comparison between *Pre0401-1–0308*, *Pre0401-1–0426*, *Pre0401-2*, and *Pre0401-4* is shown in Fig. 8. All predictions capture the general temporal change in cropland and are visually consistent with the actual FVC. The overall performances of the four strategies are similar, although some differences also exist. Compared with the actual FVC, *Pre0401-1–0308* (Fig. 8 (a)) overestimates some cropland around the human settlement area, while this phenomenon is not obvious in the other three predictions. Unlike other croplands in the study area, vegetation in the croplands with overestimation in *Pre0401-1–0308* actually did not grow from March 8 to April 1, resulting in constant low FVC values during this period. However, the corresponding coarse-resolution FVC values significantly increased from March 8 to April 1. Therefore, the deviation over these croplands in *Pre0401-1–0308* is larger than those in other predictions. Although the scatter points of *Pre0401-1–0308* are concentrated (Fig. 9 (a)), there are large overestimations in areas with low FVC values. *Pre0401-2* performs best among the four predictions, with the lowest RMSE (0.0844) and the relatively high R^2 (0.9183).

Vegetations in the study area were going through the early growing stage on March 8, 2017, and the FVC values were much lower than those on April 1 and April 26. Fig. 10 shows the prediction results with different base pairs on March 8, 2017. Visually, compared with *pre0308-1–0401* (Fig. 10 (b)), *pre0308-1–0210* (Fig. 10 (a)), *pre0308-2* (Fig. 10 (c)) and *pre0308-5* (Fig. 10 (d)) are more similar with the real FVC (Fig. 10 (e)), although they slightly overestimate in some areas. *pre0308-1–0401* shows an overall underestimation, and the predicted spatial details are different from the real FVC, which is mainly manifested as the more blurred boundaries between croplands and residential areas. The poor performance of *pre0308-1–0401* is closely related to the base image on April 1, 2017. Scatter plots and quantitative assessments (Fig. 11) more intuitively show the performances of different base pairs. The accuracies of *pre0308-2* and *pre0308-5* are nearly equal, both higher than that of *pre0308-1–0210* and *pre0308-1–0401*. Accordingly, accurate predictions are more accessible when using more than one base pair.

3.4. Performance comparison with the ESTARFM and FSDAF methods

Two commonly used spatiotemporal fusion methods, ESTARFM (Zhu et al., 2010) and FSDAF (Zhu et al., 2016), were selected as benchmark methods for comparison. Considering that ESTARFM requires at least two coarse–fine pairs as inputs for the prediction, two base pairs were also used in both STF-VGM and FSDAF for equitable comparison. For STF-VGM, the predictions with two base pairs in Section 3.3 were directly used in this comparison. For FSDAF, the predictions obtained separately from the two base pairs were combined using the method in Section 2.2.3.

Fig. 12 shows the FVC predictions on April 26, 2017, using the three methods. FVC predictions by STF-VGM are much more consistent with the reference FVC than those by the other two methods. In contrast, ESTARFM (Fig. 12 (a)) and FSDAF (Fig. 12 (b)) produce large underestimations. The FVC values on April 26 were generally approximately 0.9, but the FVC values on the base date of April 1 or May 12 were approximately 0.6–0.7. For either ESTARFM or FSDAF, the fine

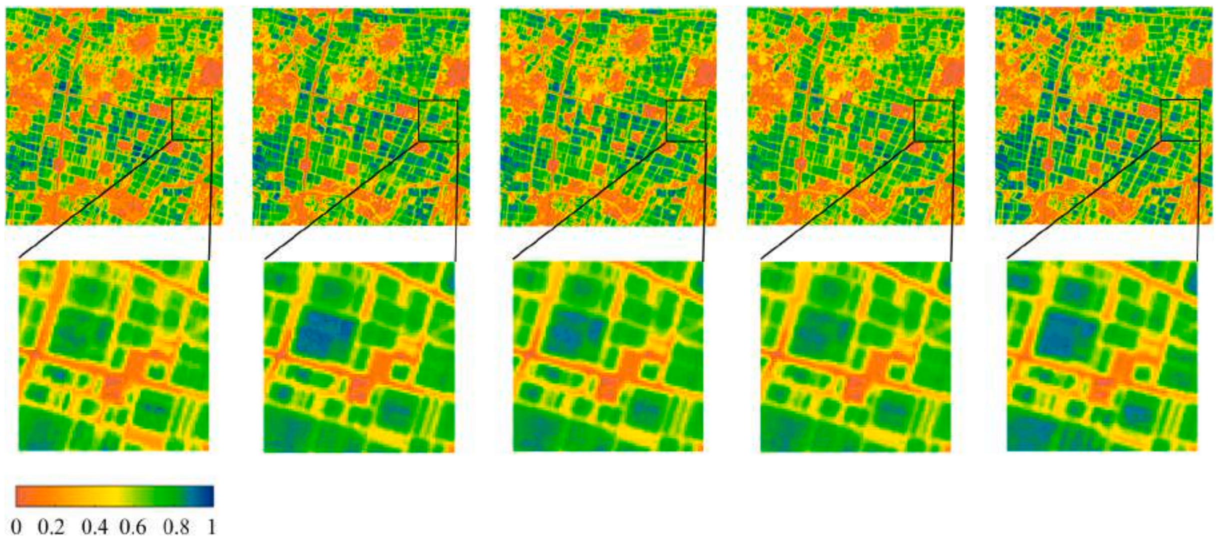

Fig. 6. FVC predictions on April 26, 2017, based on one base pair on April 1, 2017 (a), one base pair on May 12, 2017 (b), two base pairs (c), and four base pairs (d), as well as the actual FVC (e).

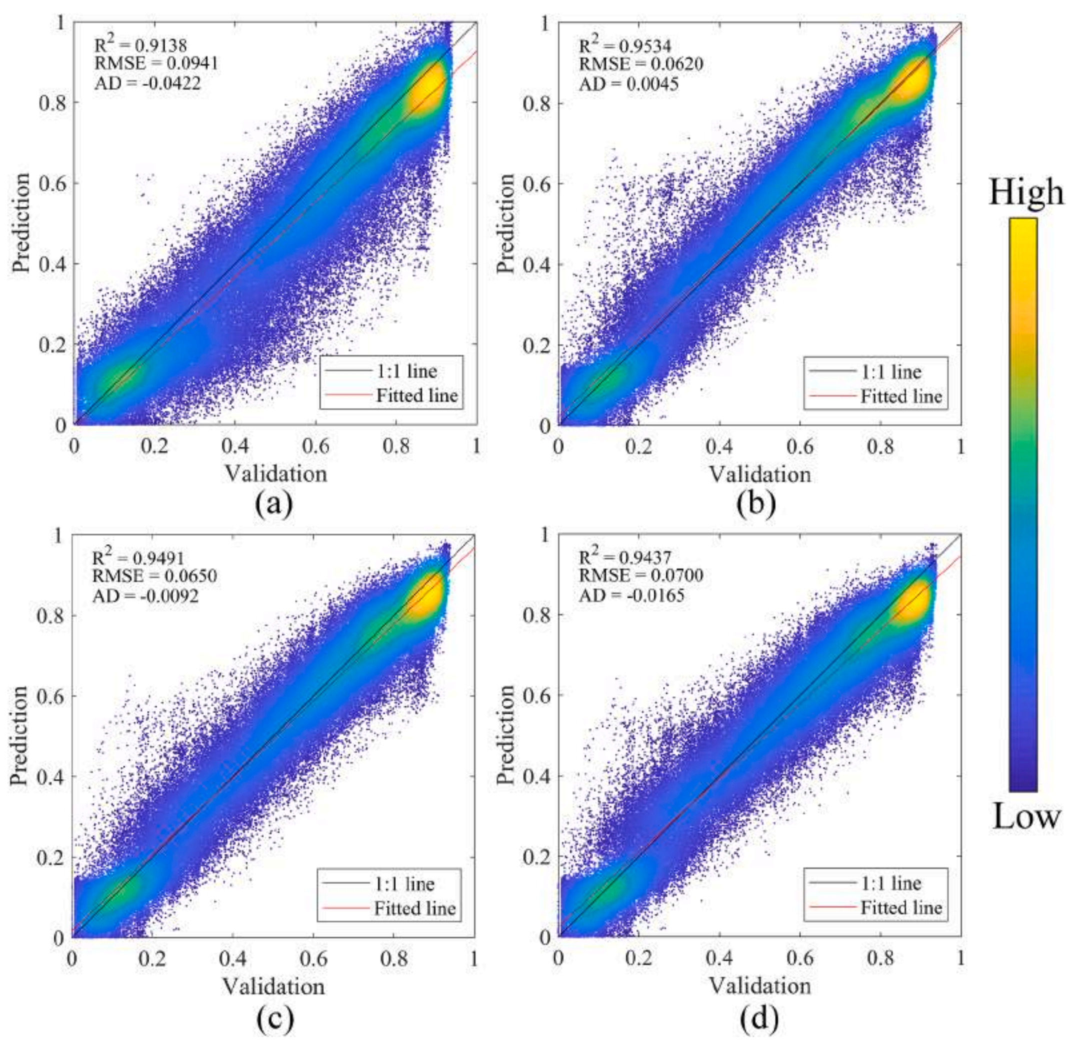

Fig. 7. Scatter plots of the FVC predictions on April 26, 2017, based on one base pair on April 1, 2017 (a), one base pair on May 12, 2017 (b), two base pairs (c), and four base pairs (d).

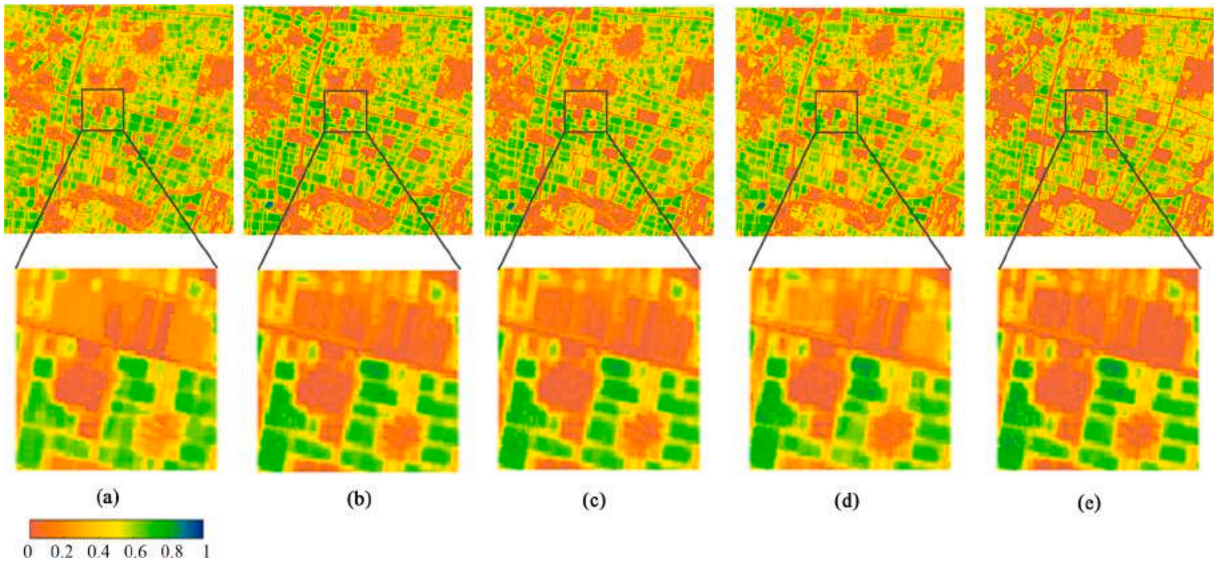

Fig. 8. FVC predictions on April 1, 2017, based on one base pair on March 8, 2017 (a), one base pair on April 26, 2017 (b), two base pairs (c), and four base pairs (d), as well as the actual FVC (e).

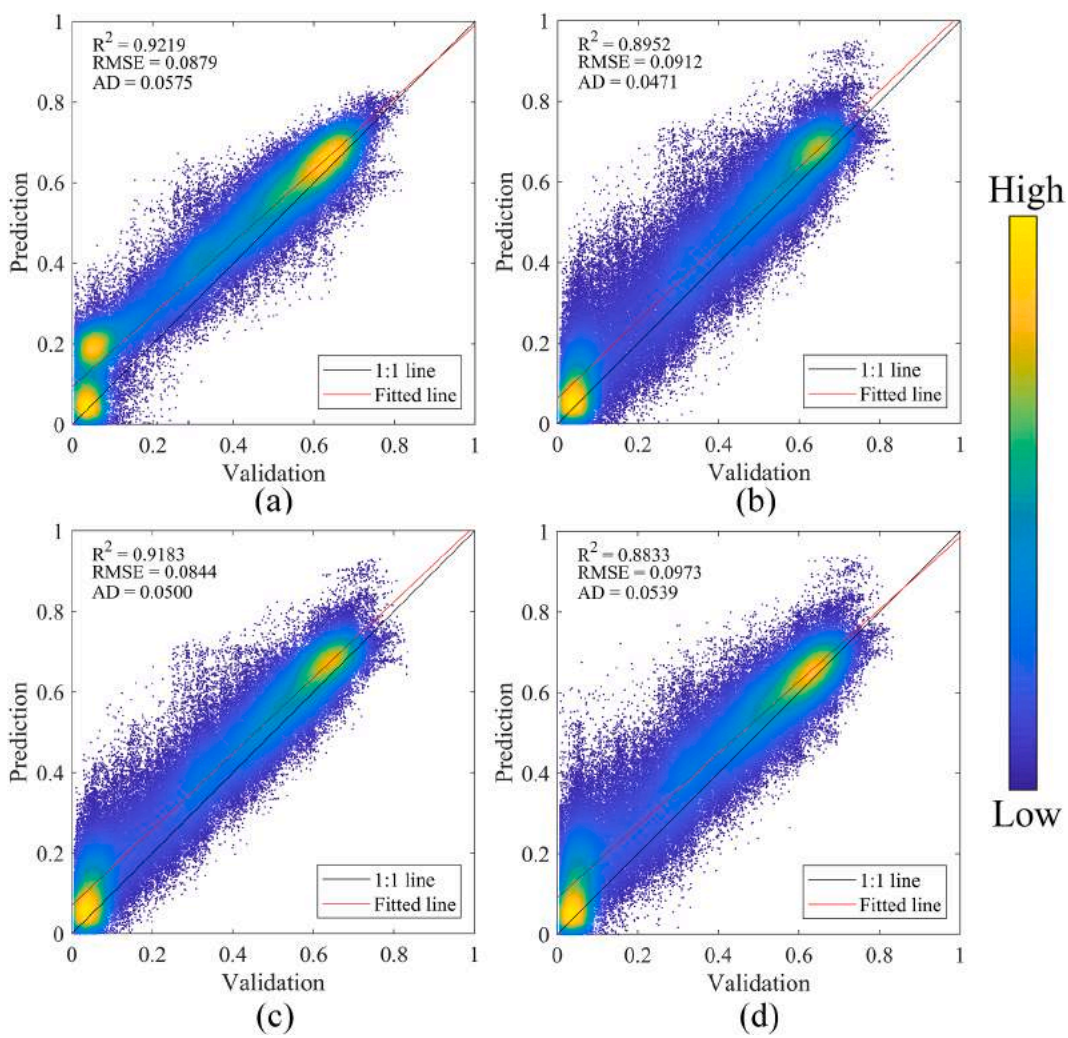

Fig. 9. Scatter plots of the FVC predictions on April 1, 2017, based on one base pair on March 8, 2017 (a), one base pair on April 26, 2017 (b), two base pairs (c), and four base pairs (d).

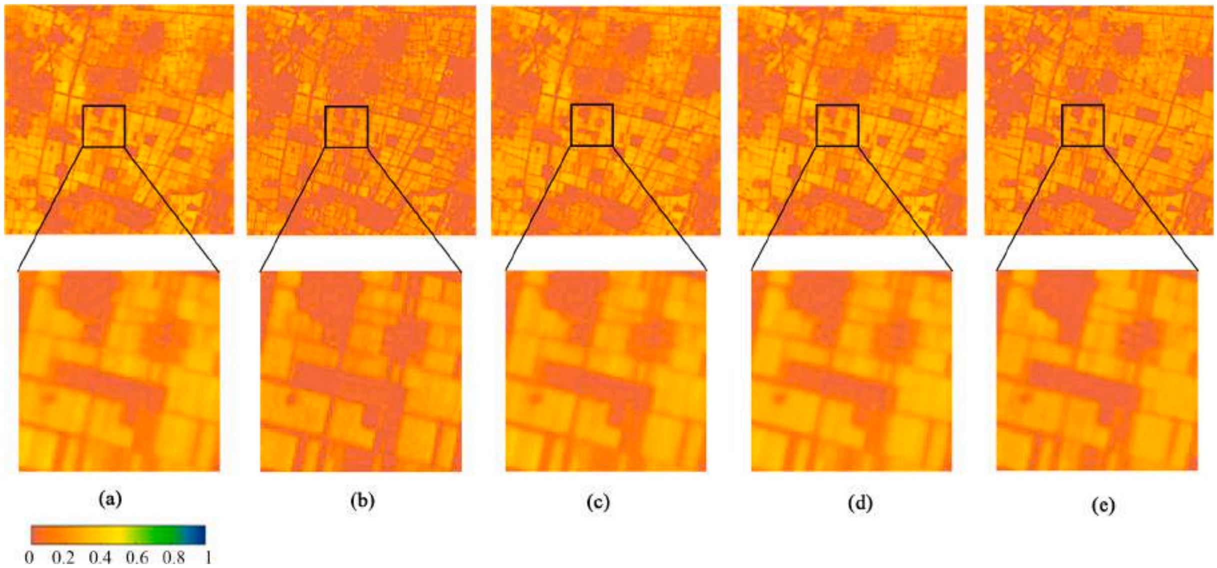

Fig. 10. FVC predictions on March 8, 2017, based on one base pair on February 10, 2017 (a), one base pair on April 1, 2017 (b), two base pairs (c), and five base pairs (d), as well as the actual FVC (e).

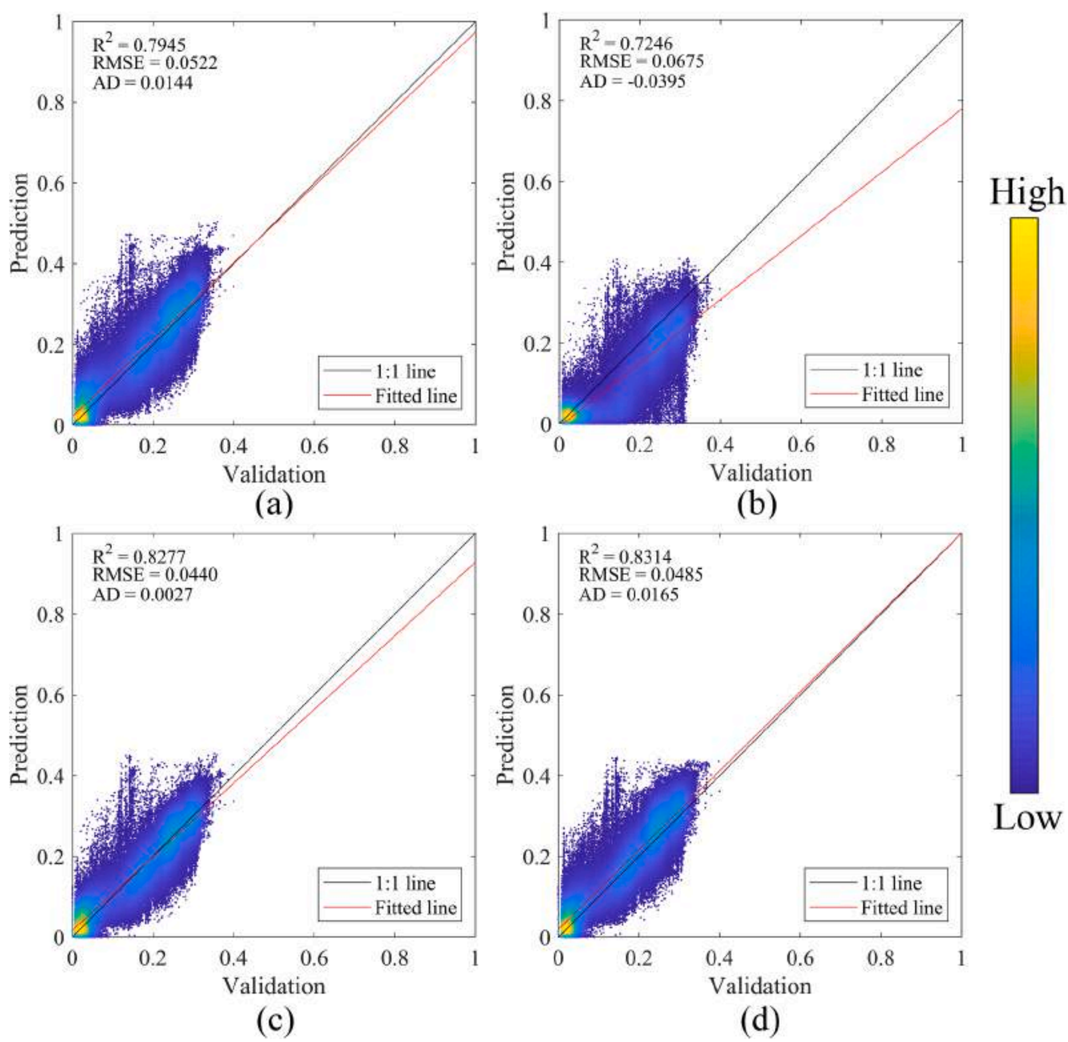

Fig. 11. Scatter plots of the FVC predictions on March 8, 2017, based on one base pair on February 10, 2017 (a), one base pair on April 1, 2017 (b), two base pairs (c), and five base pairs (d).

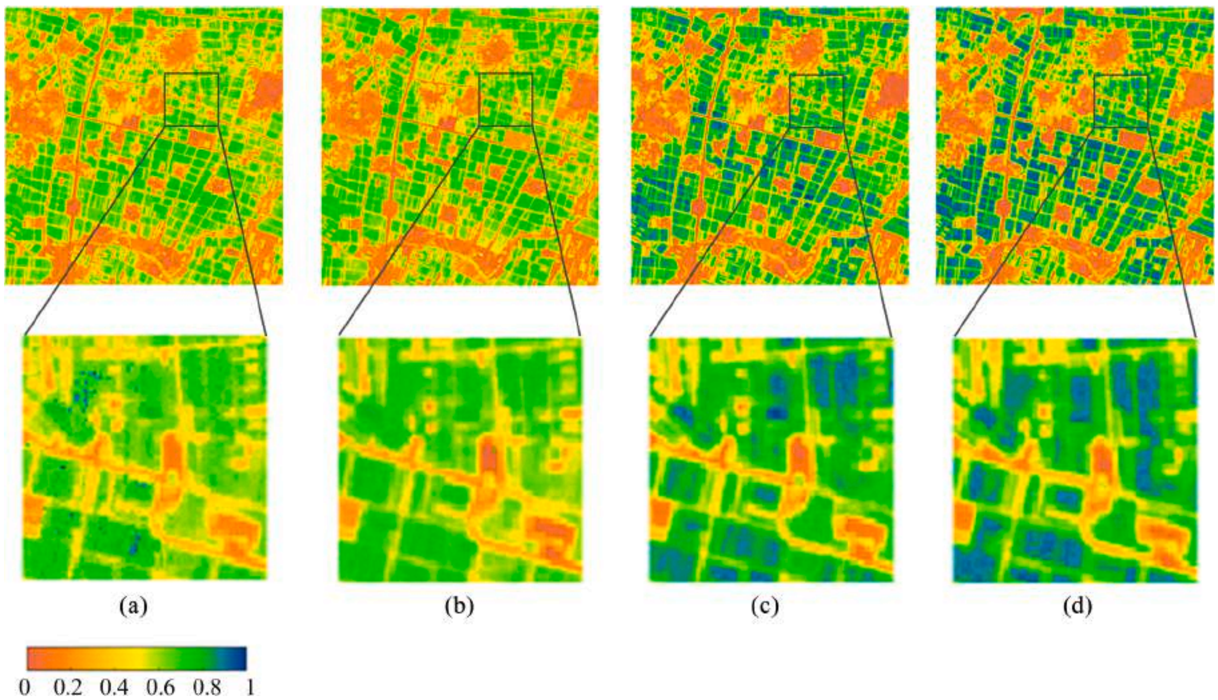

Fig. 12. FVC predictions on April 26, 2017, generated by ESTARFM (a), FSDAF (b), and STF-VGM (c), as well as the actual FVC (d).

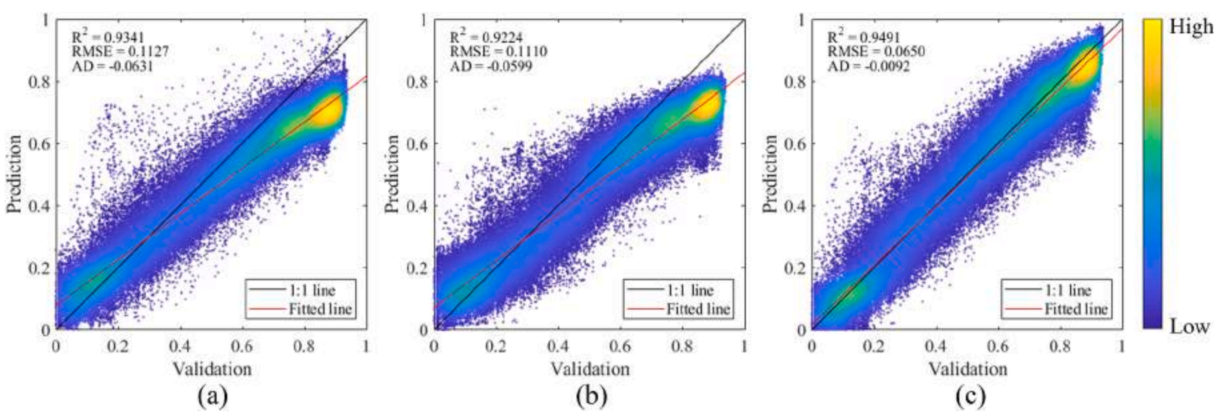

Fig. 13. Scatter plots of the FVC predictions on April 26, 2017, generated by ESTARFM (a), FSDAF (b), and STF-VGM (c).

FVC changes between the base dates (April 1, 2017, or May 12, 2017) and prediction date can only be predicted at approximately 0.1–0.15, while the actual variations might reach more than 0.2. The failure of the two methods in capturing the full variations in fine-resolution pixels leads to underestimations in their prediction results. In addition, due to the low predictions in the cropland where the actual FVC values are high, the spatial contrast between different cropland patches is not prominent in the predictions using ESTARFM and FSDAF. Thus, the restoration of spatial details by ESTARFM and FSDAF is not as good as that of the proposed STF-VGM method.

Scatter plots (Fig. 13) and quantitative evaluation also confirm that the proposed STF-VGM method achieves the highest accuracy. The yellow points with high density in Fig. 13 correspond to the pixels with high FVC values in the actual FVC image. In the scatter plots of ESTARFM (Fig. 13 (a)) and FSDAF (Fig. 13 (b)), the yellow area is noticeably located below the 1:1 reference line, indicating that the predictions of these two methods are much lower than the reference. The yellow points in the scatter plot of STF-VGM (Fig. 13 (c)) are generally distributed along the reference line with the highest R^2 (0.9491) and lowest RMSE (0.0650). Furthermore, the AD (-0.0092) of STF-VGM is closer to zero than that of the other methods, which indicates that STF-VGM produces the least biased prediction. All the accuracy metrics illustrate that the proposed STF-VGM method can effectively improve the predicted FVC value at the peak vegetation growth stage when using medium FVC values as base data.

For other periods (e.g., the stage when vegetation is growing rapidly), STF-VGM also demonstrates its superior predictive performance. Fig. 14 presents a visual comparison of the FVC predictions on April 1, 2017, generated by ESTARFM, FSDAF, and STF-VGM. The prediction fused by STF-VGM is the most consistent result with the actual reference FVC. The color of the cropland in the predictions using ESTARFM and FSDAF is greener than that using STF-VGM and the actual FVC, indicating that there is an overall overestimation by the two methods. Such overestimations can be explained by the unsatisfactory ability of the two methods to capture strong fine-resolution variations. There are two prediction periods for April 1 (from March 8 to April 1 and from April 1 to April 26). As shown in Fig. 5, the change of GLASS FVC and GF-1 WFV FVC are inconsistent during the second period (DOY91-DOY113). Since the reconstructions of fine-resolution variations by ESTARFM and FSDAF heavily depend on the coarse-resolution

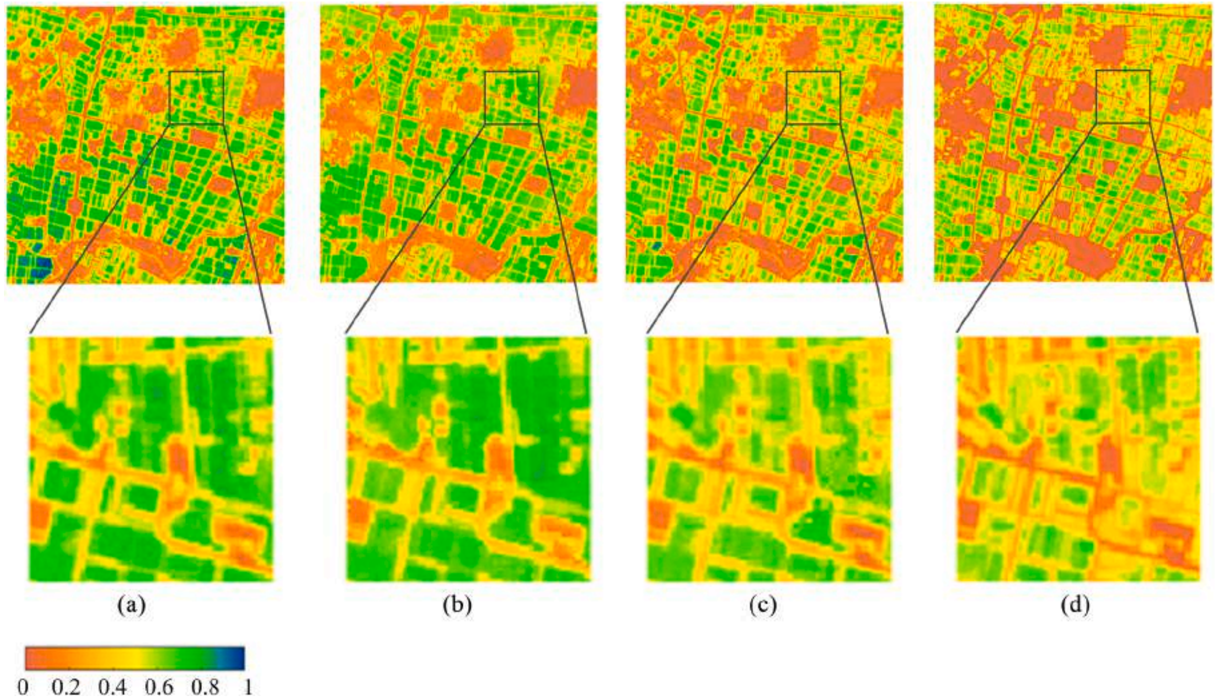

Fig. 14. FVC predictions on April 1, 2017, generated by ESTARFM (a), FSDAF (b), and STF-VGM (c), as well as the actual FVC (d).

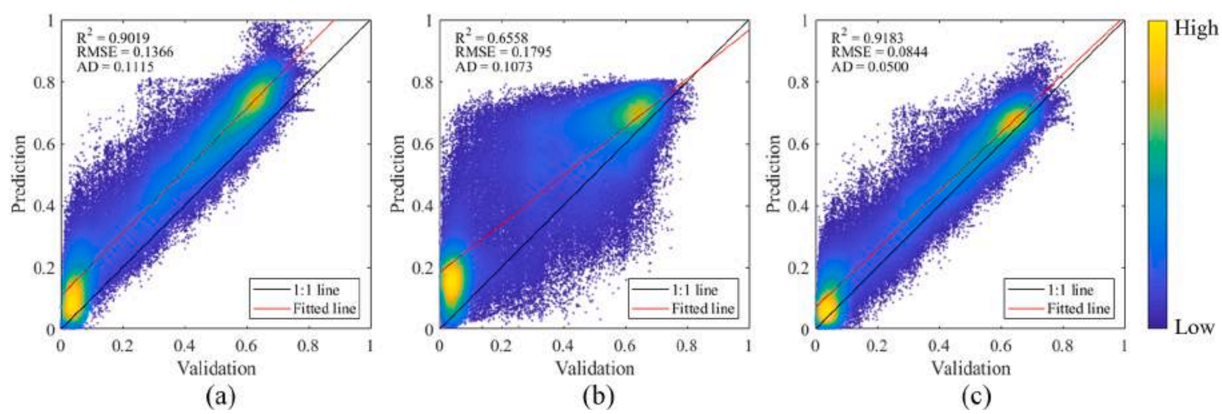

Fig. 15. Scatter plots of the FVC predictions on April 1, 2017, generated by ESTARFM (a), FSDAF (b), and STF-VGM (c).

variations, the low variations in GLASS FVC lead to the low predicted fine-resolution variations. Therefore, the FVC results, derived by subtracting the predicted low variations from the high base data on April 26, are generally higher than the real FVC. Because the vegetation condition on April 1 was more similar to that on April 26, and the difference in GLASS FVC between April 1 and April 26 was smaller, prediction with April 26 as base date has a higher weight to the final prediction. Consequently, the overestimations of ESTARFM and FSDAF are generated. Since STF-VGM has advantages in capturing the strong vegetation change, the predictions by STF-VGM are not highly overestimated like those by ESTARFM and FSDAF. Moreover, STF-VGM performs well in terms of maintaining the spatial details at the boundaries of the small cropland patches or human settlement area, whereas such boundaries are more blurred in the predictions using ESTARFM and FSDAF.

The scatter plots and the evaluation metrics are shown in Fig. 15. The scatter points of FSDAF are unsatisfactorily scattered, which indicates that FSDAF is not effective for this experiment, while the distribution patterns of the scatter points from ESTARFM and STF-VGM display higher similarity and fit the 1:1 line better. However, STF-VGM shows higher accuracy than ESTARFM, with a higher R^2 (0.9183), lower RMSE (0.0844), and an AD (0.05) closer to zero. These qualitative and quantitative comparisons indicate that the proposed STF-VGM method is also capable of achieving satisfactory predictions for medium FVC values in the period of rapid vegetation growth and outperforms the other two methods.

The above two comparisons focus on the phonological phases with high and medium vegetation cover. A comparison of predictions on March 8, 2017 can show the performances of the three methods for the periods with low FVC (e.g., the early stage of crop growth). From the visual comparison (Fig. 16), it can be seen that the results derived from ESTARFM and STF-VGM have high similarity and are generally consistent with the real FVC. While the result of FSDAF (Fig. 16(b)) is noteworthy because of the obvious anomalies with high overestimations in some residential areas. The evaluation metrics also correspondingly show the unsatisfactory accuracy of FSDAF (Fig. 17(b)). As for the quantitative comparison between ESTARFM and STF-VGM, although the former achieves slightly higher R^2 and lower RMSE, the AD of the latter is closer to 0. In general, there is almost no difference between ESTARFM and STF-VGM, and the accuracy between the two methods

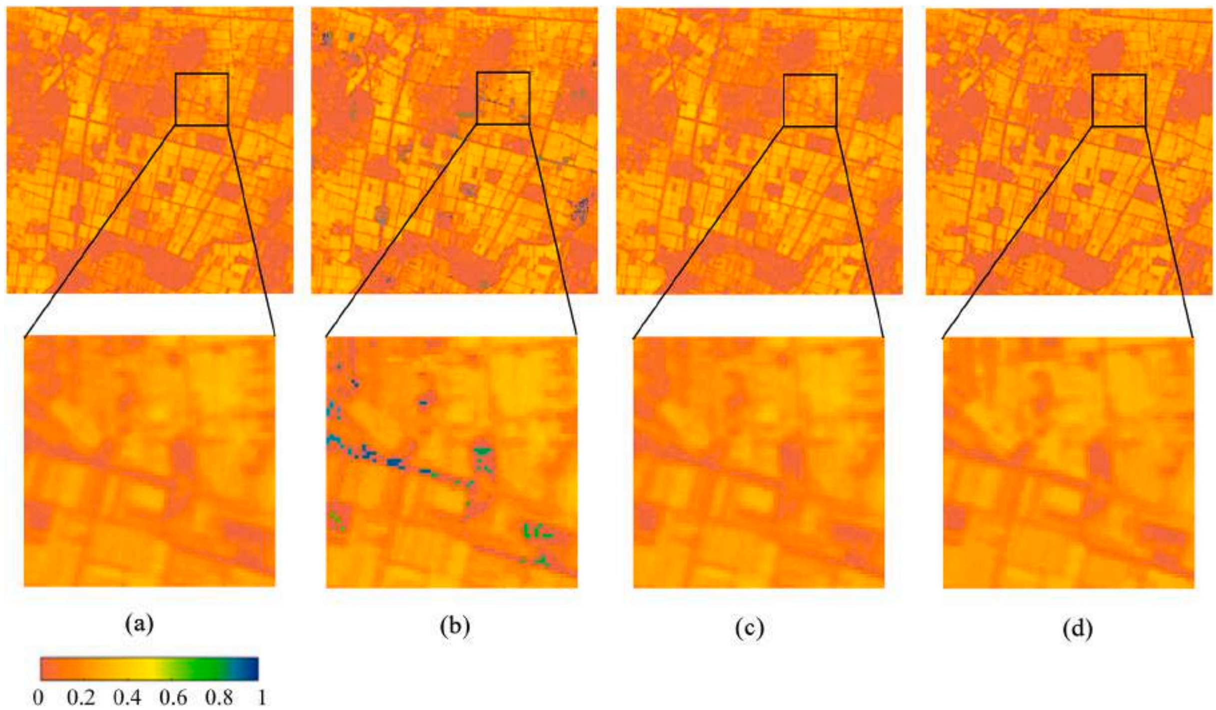

Fig. 16. FVC predictions on March 8, 2017, generated by ESTARFM (a), FSDAF (b), and STF-VGM (c), as well as the actual FVC (d).

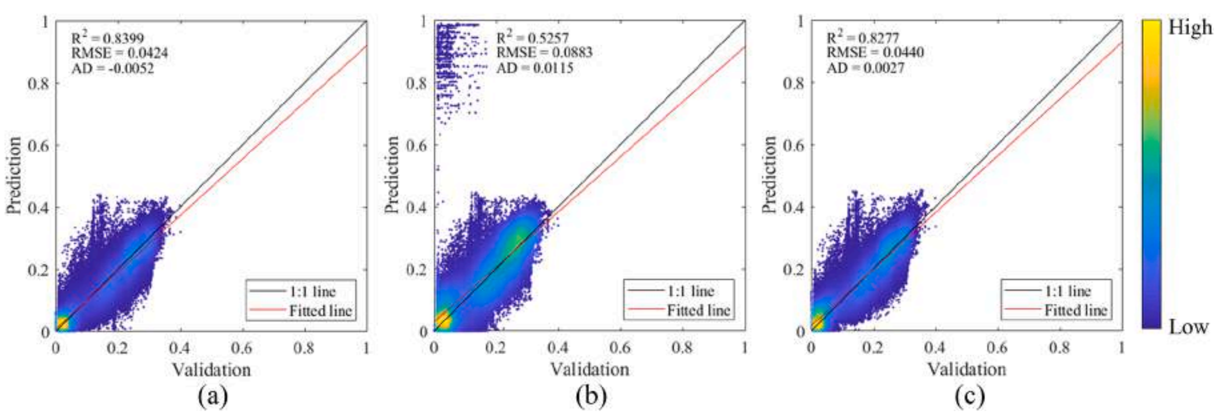

Fig. 17. Scatter plots of the FVC predictions on March 8, 2017, generated by ESTARFM (a), FSDAF (b), and STF-VGM (c).

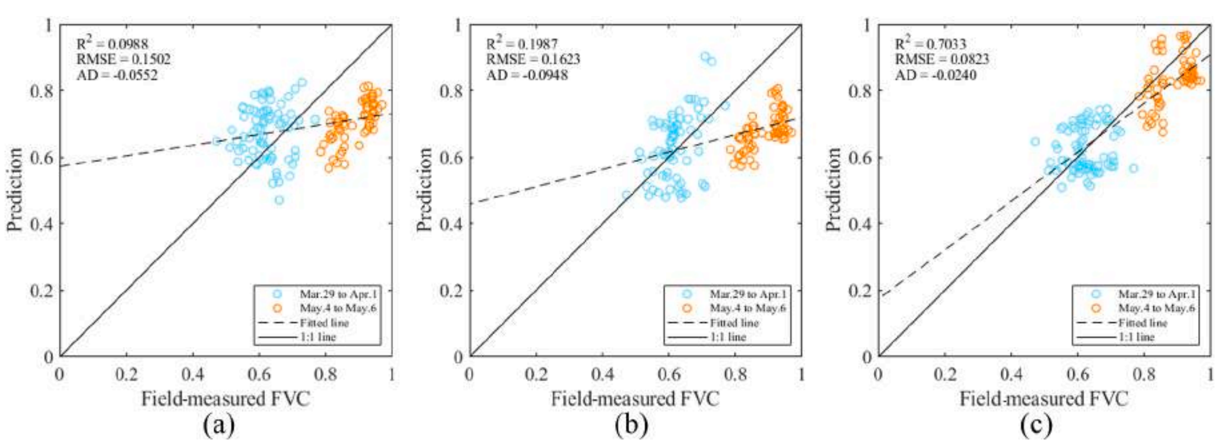

Fig. 18. Validation of the FVC predictions generated by ESTARFM (a), FSDAF (b), and STF-VGM (c) based on the field-measured FVC.

can be considered equivalent. Accordingly, STF-VGM also performs well for the early vegetation stage with low FVC and rapid growth.

3.5. Performance comparison based on the field-measured FVC

To further validate and compare the fusion methods, field-measured FVC data were also used to evaluate the accuracy. To ensure the validity of the comparisons, the number of base pairs was also set to two for each method. For the first ground survey period from March 29, 2017, to April 1, 2017, one pair on March 8, 2017, and another pair on April 26, 2017, were fused to obtain the prediction on April 1, 2017. For the second ground survey period from May 5, 2017, to May 8, 2017, the base dates were set to May 12, 2017, and April 1, 2017.

The proposed STF-VGM method achieves the best overall accuracy (R^2 = 0.7033; RMSE = 0.0823; AD = -0.0240) (Fig. 18). As shown in Fig. 18 (c), most of the points from STF-VGM fall close to the 1:1 line, and the high FVC values of the second ground survey period are well predicted. In contrast, the predictions generated by ESTARFM and FSDAF for the second period are biased low, as evidenced by all the points falling below the 1:1 line, which is in agreement with the results of the pixel-to-pixel validations in Section 3.4. This comparison further indicates the reliability of the proposed method in predicting high FVC values when using medium values as base data.

4. Discussion

This study aims to improve the spatiotemporal fusion accuracy in the agricultural region where strong temporal changes exist, especially focusing on the underestimation of high FVC values at the peak vegetation growth period when using medium FVC values as the baseline. In most previous studies, the prediction of changes in fine-resolution pixels relied entirely on the information extracted from coarse-resolution variations. However, in periods of rapid vegetation growth, the changes in fine-resolution pixels representing vegetation are significantly stronger than those in mixed coarse-resolution pixels in most situations. It is difficult to infer accurate variations at fine resolution from the slight variations at coarse resolution. The proposed STF-VGM method introduces the vegetation growth model into spatiotemporal fusion and takes full advantage of the available fine-resolution time series information, including partly uncontaminated data. STF-VGM shows significant advantages in reproducing temporal variations in fine-resolution pixels, as evidenced by the validation results.

For STF-VGM, addressing the inconsistency in the changes in fine- and coarse-resolution data is the key to improve the fusion accuracy. The inconsistency existing in certain areas is mainly attributed to two aspects. One aspect is that the variation in the fine-resolution vegetation endmember is reduced in its corresponding coarse-resolution pixel because the coarse-resolution pixel is often mixed with the non-vegetation part, which was briefly discussed in the Introduction. The other aspect may be the inconsistency of the FVC data at the two resolutions. Spatiotemporal fusion methods usually assume that the data obtained from different sensors are consistent, e.g., fine-resolution data can be aggregated and linearly related to the coarse-resolution data acquired on the same date (Gao et al., 2015). However, when using these spatiotemporal fusion methods to generate vegetation parameters, such as FVC, the consistency between the input fine- and coarse-resolution data may not be as perfect as the assumption due to differences in the sensors and the inversion algorithms. Therefore, the combination of inconsistent data leads to a contrast between the variations in fine- (strong) and coarse- (slight) resolution when rapid changes exist in vegetation. The proposed STF-VGM method can reduce the influence of the abovementioned inconsistency on the predictions through the introduction of fine-resolution data with good observation quality. As discussed above, due to the existence of inconsistency between fine- and coarse-resolution data, there may be certain errors in the fine-resolution variations derived only from the temporal information at coarse-resolution scale compared with the actual situation. The effective information of the fine-resolution pixels in the temporal dimension can be added as an additional constraint to the implementation of spatiotemporal fusion, so that the derivation of the fine resolution variation does not only dependent on the single temporal signal from coarse-resolution time series data. The inclusion of fine-resolution information means that part of the actual fine-resolution time series information is considered, which can provide supplementary guidance for the construction of the complete fine-resolution time series. Specifically, STF-VGM involves the free-cloud fine-resolution data through the vegetation growth model. STF-VGM simulates the temporal change characteristics of vegetation growth at coarse and fine resolutions by establishing the vegetation growth model separately. The fitted curves generated from the actual FVC at the two resolutions are relatively independent of each other. Therefore, the conversion coefficient calculated from the fitted FVC values can reflect a relatively reliable relationship between coarse and fine resolutions. Even though inconsistency exists, STF-VGM can still capture the fine variations close to reality using the reliable conversion coefficients.

In the part of estimating preliminary prediction in STF-VGM, there are two key steps worth being noted. One is the selection of similar pixels. In this subpart, the indicators $Rank_{S_{all}}$ and $Rank_{S_{base}}$ are used to select the most similar neighboring pixels for the target pixel. $Rank_{S_{base}}$ reflect the FVC similarity of pixels on F_{base}. If there is only F_{base} being used to recognize similar pixels, pixels that are at different phenological stages and belong to different crop species may be selected (e.g. Fig. 3 (a)). All available fine-resolution data during the growing season are considered in the indicator $Rank_{S_{all}}$. Accordingly, $Rank_{S_{all}}$ can determine which pixels have similar FVC change curves to the target pixel. In general, there are large probabilities for pixels with small $Rank_{S_{all}}$ of belonging to the same crop type as the target pixel. However, due to the differences in intra-class, it is not guaranteed that pixels selected by $Rank_{S_{all}}$ have the perfect FVC similarities with the target pixel at t_{base} (e.g., Fig. 3 (b)), which will also influence the prediction accuracy. $Rank$, the combination of $Rank_{S_{all}}$ and $Rank_{S_{base}}$, can summarize the functions of $Rank_{S_{all}}$ and $Rank_{S_{base}}$, and can find pixels that not only have same FVC change characteristics as the target pixel but also have high similarities to the target pixel on F_{base}. The pixels selected by $Rank$ are more consistent with the definition of "similar pixels" in spatiotemporal fusion and may provide positive contributions to the improvement of prediction accuracy. Another key point is the variable conversion coefficient that can change along with the nonlinear process of crop growth. Based on the vegetation growth model constructed at the coarse- and fine-resolutions, continuous curves reflecting the crop growth status can be obtained. With the fitting result as the support, it is easy to extract the conversion relationship between the coarse- and fine-resolution data at any time through Eq. (13). Due to the nonlinear change pattern of crops and the inconsistancy between coarse- and fine-resolution data, such conversion relationship is not stable during the vegetation growth process. Therefore, compared with the constant conversion coefficient used in ESTARFM, the conversion relationship derived in STF-VGM can change with time because of its calculability for any short period during different crop growth stages.

STF-VGM provides a framework to combine multiple preliminary predictions calculated from different base dates. The case study indicates that improved predictions can be achieved using only the two available nearest base pairs before and after the prediction date rather than more base pairs. Increasing the number of base pairs does not necessarily lead to reduced errors. The base pairs distant from the prediction date are less correlated with the target FVC. The surface status at a certain time is related to the previous moment and closely affects the following moment. Therefore, data on the prediction date are most closely related to the nearest base pairs before and after the prediction date. The addition of more base pairs does not reduce the uncertainty in the final prediction, and the possibility of introducing error information can also be increased. Therefore, instead of integrating all preliminary predictions, two preliminary predictions obtained based on the two nearest base pairs before and after the prediction date are recommended.

The experiments in Section 3.4 and 3.5 compare the three methods of ESTARFM, FSDAF and STF-VGM. In Section 3.4, the performances of the three methods at peak vegetation growth stage (high FVC), rapid vegetation growth stage (medium FVC) and early growth stage (low FVC) were evaluated, respectively. For the peak vegetation growth stage, FVC predictions of crops on April 26, 2017 generated by ESTARFM and FSDAF were mainly underestimated, while STF-VGM showed satisfactory performance because of the relatively accurate predictions on the FVC between 0.7 and 1. The fusion process of ESTARFM, FSDAF and STF-VGM is to reproduce the variations that occur in fine-resolution pixels during a prediction period (ΔF). The underestimation of ESTARFM and FSDAF are both caused by their underestimated ΔF. ESTARFM converts the coarse-resolution variations into fine-resolution variations through a conversion coefficient ($V = \Delta F/\Delta C$). The vegetation nonlinear change patterns are not considered in the construction of V. For the period from April 1 to May 12, there was a change inconsistency between coarse- (slight) and fine- (strong) resolution data, and V showed a trend of increasing first and then decreasing (Fig. 5). However, V calculated by ESTARFM was constant and smaller than the real changing V, resulting in the underestimated ΔF. FSDAF obtains ΔF by unmixing ΔC that occur in the corresponding coarse-resolution pixels. Due to the slight change in coarse-resolution pixels from April 1 to May 12 and the error caused by unmixing method, ΔF calculated by FSDAF was also smaller than the real values. In summary, due to the lack of consideration of vegetation nonlinear change and the strong dependency on coarse-resolution variations during fusion process, when vegetation changes rapidly and temporal variations in coarse- and fine-resolution pixels occur to different degrees, ESTARFM and FSDAF may fail to capture the full strong change in fine resolution and generate underestimated FVC predictions with only medium or low FVC as base data. STF-VGM has two main advantages over ESTARFM and FSDAF, which lead to its ability to obtain a more satisfactory predicted ΔF. The first one is that STF-VGM utilizes the effective information of fine-resolution data over temporal dimension, making the derivation of ΔF not entirely relied on ΔC during the same prediction period. The second one lies in the changing conversion coefficient which can adapt to the vegetation nonlinear growth pattern. As for the prediction on rapid vegetation growth stage, the temporal change of coarse- and fine-resolution data were generally consistent from March 8 to April 1 and appeared gradually different from each other from April 1 to April 26 (Fig. 5). Due to the relatively accurate predicted ΔF during the second period (from April 1 to April 26), STF-VGM also achieved higher accuracy than ESTARFM and FSDAF on this stage. During the early vegetation growth stage, the temporal changes occurring at coarse and fine resolutions were both strong and showed consistent trend. The advantage of STF-VGM in addressing the inconsistency is not obvious for the prediction on this period, and STF-VGM thus achieved the equivalent accuracy with ESTARFM. In summary, STF-VGM can significantly improve the fusion accuracy for the period when the inconsistency between coarse- and fine-resolution data exists, especially for the peak vegetation growth stage. And when the temporal changes at the two resolutions are generally consistent, STF-VGM can still maintain good performance compared with ESTARFM and FSDAF.

There are also some limitations of the proposed method. One limitation lies in the ability to address abrupt changes, which is a long-standing problem in spatiotemporal fusion. For example, in the case study in Section 3.3, there is a small region with abrupt changes in the center of the test area between the prediction date and May 12, 2017, which is possibly caused by the land cover type changing from bare soil with low FVC to crops with much higher FVC. In this situation, the vegetation growth model could not provide helpful information for such unpredictable land cover change. The use of two input pairs before and after the prediction date, as recommended in the previous paragraph, may potentially reduce the large errors caused by abrupt land cover changes. Another limitation is that STF-VGM is currently more suitable for the fusion of one crop growing season. This is because that the used vegetation growth model describes the unimodal vegetation growth pattern. For the areas with 2–3 growing seasons throughout the year, STF-VGM needs to conduct fusion for each growing season separately. For further applications, other suitable models can also be considered for bimodal and multimodal crops, which will be explored in future work.

The proposed STF-VGM method was investigated for improving the spatiotemporal fusion accuracy of FVC data in this study. STF-VGM is also suitable for improving the fusion accuracy of other variables related to vegetation, such as NDVI, LAI, and FAPAR, as long as their temporal changes are consistent with the growing process of vegetation. Although GF-1 WFV data were selected as the fine-resolution data for test, other satellite data, such as Sentinel-2 data, could also be fused with the GLASS FVC product. Moreover, jointly using GF-1 WFV and Sentinel-2 data in one fusion framework can increase the amount of temporal observations, which could largely improve the data availability for vegetation growth model fitting. In addition, more satellite observations can also provide more auxiliary information temporally close to the prediction date, which is helpful for handling abrupt changes.

5. Conclusion

In this study, the STF-VGM method was proposed to improve the spatiotemporal fusion accuracy of FVC in the agricultural region by incorporating vegetation growth models, which can help describe nonlinear vegetation changes. Based on the available coarse- and fine-resolution time series data, STF-VGM can address the problem of the inconsistency between fine- and coarse-resolution variations and achieve satisfactory performance. The experimental results fused from GF-1 WFV FVC and GLASS FVC data by STF-VGM, ESTARFM, and FSDAF indicated that STF-VGM achieved the most precise temporal variations in fine-resolution pixels. Therefore, STF-VGM significantly improved the spatiotemporal fusion accuracy of high FVC values at the peak vegetation growth stage compared to the commonly used ESTARFM and FSDAF methods. Additionally, STF-VGM also achieved satisfactory performance for other vegetation growth stages. Moreover, STF-VGM is applicable to other satellite data, e.g., Sentinel-2 data, or other variables related to vegetation, such as NDVI, LAI, and FAPAR. Future work will focus on further applications of STF-VGM in this regard.

CRediT authorship contribution statement

Guofeng Tao: Methodology, Writing - original draft, Investigation, Visualization. **Kun Jia:** Conceptualization, Writing - review & editing. **Xiangqin Wei:** Writing - review & editing. **Mu Xia:** Writing - review & editing. **Bing Wang:** Resources. **Xianhong Xie:** Writing - review & editing. **Bo Jiang:** Writing - review & editing. **Yunjun Yao:** Writing - review & editing. **Xiaotong Zhang:** Writing - review & editing.

Declaration of Competing Interest

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

Acknowledgement

This study was supported by the National Key Research and Development Program of China (2016YFB0501404 and 2016YFA0600103), the National Natural Science Foundation of China (No. 41671332) and the Tang Scholar Program (K. Jia is a Tang Scholar of Beijing Normal University).

References

Atzberger, C., 2013. Advances in remote sensing of agriculture: context description, existing operational monitoring systems and major information needs. Remote Sens. 5, 949–981.

Baret, F., Bacour, C., Béal, D., Weiss, M., Berthelot, B., Regner, P., 2006. Algorithm theoretical basis document for MERIS top of canopy land products (TOC_VEG). Contract 1–25.

Baret, F., Hagolle, O., Geiger, B., Bicheron, P., Miras, B., Huc, M., Berthelot, B., Niño, F., Weiss, M., Samain, O., 2007. LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION: Part 1: Principles of the algorithm. Remote Sens. Environ. 110, 275–286.

Baret, F., Weiss, M., Lacaze, R., Camacho, F., Makhmara, H., Pacholcyzk, P., Smets, B., 2013. GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part1: principles of development and production. Remote Sens. Environ. 137, 299–309.

Bindraban, P.S., 1999. Impact of canopy nitrogen profile in wheat on growth. F. Crop. Res. 63, 63–77.

Boogaard, H.L., Diepen, C.A. van, Rötter, R.P., Cabrera, J.C.M.A., van Laar, H.H., 1998. WOFOST 7.1. User guide for the WOFOST 7.1 crop growth simulation model and WOFOST Control Center 5.1., Winand Staring Centre. SC-DLO.

Bullock, P.R., 1992. Operational estimates of western Canadian grain production using NOAA AVHRR LAC data. Can. J. Remote Sens. 18, 23–28.

Camacho, F., Cernicharo, J., Lacaze, R., Baret, F., Weiss, M., 2013. GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: validation and intercomparison with reference products. Remote Sens. Environ. 137, 310–329.

Gao, F., Hilker, T., Zhu, X., Anderson, M., Masek, J., Wang, P., Yang, Y., 2015. Fusing Landsat and MODIS data for vegetation monitoring. IEEE Geosci. Remote Sens. Mag. 3, 47–60.

Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. IEEE Trans. Geosci. Remote Sens. 44, 2207–2218.

Gutman, G., Ignatov, A., 1998. The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. Int. J. Remote Sens. 19, 1533–1543.

Hilker, T., Wulder, M.A., Coops, N.C., Linke, J., McDermid, G., Masek, J.G., Gao, F., White, J.C., 2009. A new data fusion model for high spatial-and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. Remote Sens. Environ. 113, 1613–1627.

Houborg, R., McCabe, M.F., Gao, F., 2016. A spatio-temporal enhancement method for medium resolution LAI (STEM-LAI). Int. J. Appl. Earth Obs. Geoinf. 47, 15–29.

Hu, Q., Yang, J., Xu, B., Huang, J., Memon, M.S., Yin, G., Zeng, Y., Zhao, J., Liu, K., 2020. Evaluation of global decametric-resolution LAI, FAPAR and FVC estimates derived from sentinel-2 imagery. Remote Sens. 12, 912.

Jacquemoud, S., Baret, F., 1990. PROSPECT: A model of leaf optical properties spectra. Remote Sens. Environ. 34, 75–91.

Jia, K., Liang, S., Gu, X., Baret, F., Wei, X., Wang, X., Yao, Y., Yang, L., Li, Y., 2016. Fractional vegetation cover estimation algorithm for Chinese GF-1 wide field view data. Remote Sens. Environ. 177, 184–191.

Jia, K., Liang, S., Liu, S., Li, Y., Xiao, Z., Yao, Y., Jiang, B., Zhao, X., Wang, X., Xu, S., Cui, J., 2015. Global land surface fractional vegetation cover estimation using general regression neural networks from MODIS Surface reflectance. IEEE Trans. Geosci. Remote Sens. 53, 4787–4796.

Jia, K., Liang, S., Wei, X., Yao, Y., Yang, L., Zhang, X., Liu, D., 2018. Validation of global land surface satellite (GLASS) fractional vegetation cover product from MODIS data in an agricultural region. Remote Sens. Lett. 9, 847–856.

Jia, K., Yang, L., Liang, S., Xiao, Z., Zhao, X., Yao, Y., Zhang, X., Jiang, B., Liu, D., 2019. Long-Term Global Land Surface Satellite (GLASS) fractional vegetation cover product derived from MODIS and AVHRR Data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12, 508–518.

Jiménez-Muñoz, J.C., Sobrino, J.A., Plaza, A., Guanter, L., Moreno, J., Martínez, P., 2009. Comparison between fractional vegetation cover retrievals from vegetation indices and spectral mixture analysis: case study of PROBA/CHRIS data over an agricultural area. Sensors 9, 768–793.

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. Eur. J. Agron. 18, 235–265.

Kim, Y., Kyriakidis, P.C., Park, N.-W., 2020. A cross-resolution, spatiotemporal geostatistical fusion model for combining satellite image time-series of different spatial and temporal resolutions. Remote Sens. 12, 1553.

Li, X., Foody, G.M., Boyd, D.S., Ge, Y., Zhang, Y., Du, Y., Ling, F., 2020. SFSDAF: An enhanced FSDAF that incorporates sub-pixel class fraction change information for spatio-temporal image fusion. Remote Sens. Environ. 237, 111537.

Liang, S., Zhao, X., Liu, S., Yuan, W., Cheng, X., Xiao, Z., Zhang, X., Liu, Q., Cheng, J., Tang, H., Qu, Yonghua, Bo, Y., Qu, Ying, Ren, H., Yu, K., Townshend, J., 2013. A long-term Global LAnd Surface Satellite (GLASS) data-set for environmental studies. Int. J. Digit. Earth 6, 5–33.

Lin, Z., Xiang, Y., Mo, X., Li, J., Wang, L., 2003. Normalized leaf area index model for summer maize. Chinese J. Eco-Agriculture 11, 69–72.

Liu, D., Jia, K., Wei, X., Xia, M., Zhang, Xiwang, Yao, Y., Zhang, Xiaotong, Wang, B., 2019a. Spatiotemporal comparison and validation of three global-scale fractional vegetation cover products. Remote Sens. 11, 2524.

Liu, M., Liu, X., Wu, L., Zou, X., Jiang, T., Zhao, B., 2018. A modified spatiotemporal fusion algorithm using phenological information for predicting reflectance of paddy rice in southern China. Remote Sens. 10, 772.

Liu, M., Yang, W., Zhu, X., Chen, J., Chen, X., Yang, L., Helmer, E.H., 2019b. An Improved Flexible Spatiotemporal DAta Fusion (IFSDAF) method for producing high spatiotemporal resolution normalized difference vegetation index time series. Remote Sens. Environ. 227, 74–89.

Lourakis, M.I.A., 2005. A brief description of the Levenberg-Marquardt algorithm implemented by levmar. Found. Res. Technol. 4, 1–6.

Propastin, P., Erasmi, S., 2010. A physically based approach to model LAI from MODIS 250 m data in a tropical region. Int. J. Appl. Earth Obs. Geoinf. 12, 47–59.

Ritchie, J.T., 1985. Description and performance of CERES wheat: a user-oriented wheat yield model. ARS wheat yield Proj. 159–175.

Roujean, J.L., Lacaze, R., 2002. Global mapping of vegetation parameters from POLDER multiangular measurements for studies of surface-atmosphere interactions: A pragmatic method and its validation. J. Geophys. Res. D Atmos. 107.

Tang, H., Yu, K., Hagolle, O., Jiang, K., Geng, X., Zhao, Y., 2013. A cloud detection method based on a time series of MODIS surface reflectance images. Int. J. Digit. Earth 6, 157–171.

Tao, G., Jia, K., Zhao, X., Wei, X., Xie, X., Zhang, Xiwang, Wang, B., Yao, Y., Zhang, Xiaotong, 2019. Generating High Spatio-Temporal Resolution Fractional Vegetation Cover by Fusing GF-1 WFV and MODIS Data. Remote Sens. 11, 2324.

Tian, F., Wang, Y., Fensholt, R., Wang, K., Zhang, L., Huang, Y., 2013. Mapping and evaluation of NDVI trends from synthetic time series obtained by blending Landsat and MODIS data around a coalfield on the Loess Plateau. Remote Sens. 5, 4255–4279.

Verhoef, W., 1984. Light scattering by leaf layers with application to canopy reflectance modeling: the SAIL model. Remote Sens. Environ. 16, 125–141.

Wang, X., Jia, K., Liang, S., Li, Q., Wei, X., Yao, Y., Zhang, X., Tu, Y., 2017. Estimating fractional vegetation cover from landsat-7 ETM+ Reflectance data based on a coupled radiative transfer and crop growth model. IEEE Trans. Geosci. Remote Sens. 55, 5539–5546.

Wang, X., Jia, K., Liang, S., Zhang, Y., 2016. Fractional vegetation cover estimation method through dynamic Bayesian network combining radiative transfer model and crop growth model. IEEE Trans. Geosci. Remote Sens. 54, 7442–7450.

Wu, B., Huang, B., Cao, K., Zhuo, G., 2017. Improving spatiotemporal reflectance fusion using image inpainting and steering kernel regression techniques. Int. J. Remote Sens. 38, 706–727.

Wu, R., Ma, C.-X., Zhao, W., Casella, G., 2003. Functional mapping for quantitative trait loci governing growth rates: a parametric model. Physiol. Genomics 14, 241–249.

Zhai, H., Huang, F., Qi, H., 2020. Generating high resolution lai based on a modified FSDAF model. Remote Sens. 12, 150.

Zhang, X., Liao, C., Li, J., Sun, Q., 2013. Fractional vegetation cover estimation in arid and semi-arid environments using HJ-1 satellite hyperspectral data. Int. J. Appl. Earth Obs. Geoinf. 21, 506–512.

Zhao, Y., Huang, B., Song, H., 2018. A robust adaptive spatial and temporal image fusion model for complex land surface changes. Remote Sens. Environ. 208, 42–62.

Zhu, X., Cai, F., Tian, J., Williams, T.K.A., 2018. Spatiotemporal fusion of multisource remote sensing data: literature survey, taxonomy, principles, applications, and future directions. Remote Sens. 10, 527.

Zhu, X., Chen, J., Gao, F., Chen, X., Masek, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. Remote Sens. Environ. 114, 2610–2623.

Zhu, X., Helmer, E.H., Gao, F., Liu, D., Chen, J., Lefsky, M.A., 2016. A flexible spatiotemporal method for fusing satellite images with different resolutions. Remote Sens. Environ. 172, 165–177.