Comparison of three NDVI time-series fitting methods in crop phenology detection in Northeast China

Meng Wang\textsuperscript{1,2}, Fulu Tao\textsuperscript{1}

\textsuperscript{1}Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, 11A, Datun Road, Chaoyang District, Beijing, 100101, China
\textsuperscript{2}University of Chinese Academy of Sciences, Beijing, 10049, China

Email: wangm.10b@igsnrr.ac.cn

Abstract. Phenological changes of cropland are the pivotal basis for farm management, agricultural production, and climate change research. Over the past decades, a range of methods have been used to extract phenological events based on satellite derived continuous vegetation index time series, however, large uncertainties still exist. In this study, three smoothing methods were compared to reduce the potential uncertainty and to quantify crop green-up dates over Northeast China. The results indicated that the crop spring onset dates estimated by three methods show variance in the dates, but with similar spatial pattern. In 60\% of the study area, the standard deviation (SD) of the estimated starting date from different method is less than 10 days, while 39.5\% of total pixels have SDs between 10 days and 30 days. Through comparative analysis against the observation phenological data, we concluded that Asymmetric Gaussians produced the most approximative results of all, followed by Double Logistic algorithm, and Savizky-Glolay algorithm performed worst. The starting dates of crops occur mostly between May and June in this region. The Savitzky-Golay has the earliest estimates, while the Asymmetric Gaussians and Double logistic fitting method show similar and later estimates, which are more consistent with the observed data.

1. Instruction
The cropland phenophase is a critical parameter in global climate change research [1-3] and is closely related to farming management and planning [4]; in addition, it also plays an important role in crop models that are used to monitor crop production. The aim of cropland phenophases detection is to identify the dates on which the phenological events in crops take place. There are at least three methods to derive vegetation phenology: (1) networks observing (2) phenology modeling (3) remote sensing. The first method is visual observation in field, which is difficult to be used in large areas. Extracting cropland phenophases through remote sensing data can make up for the deficiency of visual observation and is essential for the capturing of cropland phenophases on a large scale [5].

Northeast China is an important production base of grain in China and plays a critical role in the national food supply and agricultural production of the country [6]. To date, many studies have been made to detect vegetation phenophases with time series data of remote sensing, however, the studies mainly focused on the phenophases of grassland, shrub land and forest but not cropland [6-8]. The human managements make the estimation of phenology for cropland much more complicated. Accurate understanding of the spatial pattern of cropland phenology appears especially important. Several studies have been conducted to capture cropland phenology in Northeast China based on...
remote sensing data. Most of the study used one single method, which failed in comparing the accuracy of different method, and large differences exist among the results of the estimations. White et al. (2009) showed that the central and largely uncertainty is due to the algorithm selected to extract the phenological parameters [9]. It is crucial to conduct an inter-comparison of different algorithm.

The study presented here performed a three method investigation of the start of season of cropland in Northeast China using NDVI data produced from SPOT Vegetation satellites. We choose three provinces include Heilongjiang, Jilin and Liaoning as our study area (figure 1). We focus on the comparison of three different methods which are used to simulate the spatial pattern of the start of spring phenology.

2. Data and methods

2.1. Overview of the study area
Northeast China covers an area of 791,800km$^2$, in which of 264,400km$^2$ is cropland, accounting for 16.5% of the total arable land in China [6]. The major crops of this region include rice, corn, wheat and soybean. The average summer temperature is 20-25°C, and the total precipitation is 500-800mm. In most parts of this region, the accumulated temperature (AT) $\geq$0°C is 2000-4200°C·days, and the active AT $\geq$10°C is 1600-3600°C·days.

![Figure 1. A geographical location map of the study area (bottom right inset) and the location of the observation stations used in the study (main map).](image)

2.2. Data set

The NDVI data set for the period 1999-2009 derived from the vegetation instrument of SPOT 4 and 5 satellites was used, which was downloaded from the Vegetation Distribution Site (http://free.vgt.vito.be) of the Flemish Institute for Technological Research (VITO). The spatial resolution is 1km and the temporal resolution is about 10 days, so there are 36 composites each year. The SPOT NDVI dataset has been corrected to remove the effects of sensor degradation and satellite shift. The atmospheric contamination has also been corrected through simplified method for atmospheric corrections (SMAC) [10]. The non-vegetation effects were also minimized through the method of maximum value composite for each 10-day interval [11]. Despite of the efforts in improving the data quality mentioned above, there are still spurious variations in the vegetation indexes caused by the remaining atmospheric contamination [12] [13].

The observed phenology data for 1999-2009 for the crops in Northeast China are from experiments conducted at the national agro-meteorological stations, which are maintained by the Chinese Meteorological Administration (CMA) [14]. The dates of greenup for each crop are recorded field
observations. In this study, we took rice as an example to test the consistency between the extracted and observed phenology data.

The land use data of Northeast China was derived from a combination of remotely sensed data from AVHRR and geographical data sets [15]. The arable land in Northeast China was extracted by the ArcGIS software.

2.3. Analyses
The smoothing reconstruction is a crucial basis for using the multi-temporal NDVI dataset to extract crop phenological features. In the process, vegetation indices which are acquired biweekly or less frequently are expanded into time series suitable for phenological study, usually on daily base [16]. In the same time, noisy points are also filtered. In this study we used three principle filter methods in reconstructing NDVI, which are Adaptive Savitzky-Golay fitting method, Asymmetric Gaussians function fitting method, and Double logistic function fitting method.

Base on the reconstructed satellite signal time series, the thresholds for phenological events are determined. The rule selected for the extraction of phenological information is consistent for the three fitting method. In this study, the starting stage of crop growth is defined as the date when the NDVI time-series curve grew to 20% of the overall increase [17].

Lastly, correlation analysis was performed between the starting stage of crop growth extracted through the two steps aforementioned and the observed data from crop phenological observation stations to verify the accuracy. The spatial distribution of the crop phenological observation stations is shown in figure 1.

3. Results and discussion

3.1. Spatial patterns of green-up dates
The spatial patterns of the starting date of crop growth in Northeast China calculated with three different methods using the average SPOT NDVI data from 1999 to 2009 are shown in figure 3. In eastern Sanjiang Plain and northern Songnei Plain, the dates of onset-of-growth are earlier. The spatial pattern of the starting date of crop growth did not exhibit apparent south-north differences. In most of the region, the starting dates of crops occur between May and June. The spatial pattern of the starting dates of crops revealed by the Asymmetric Gaussians fitting method and Double Logistic fitting method are similar, while the variances between those two methods and the result of Adaptive Savitzky-Golay fitting method are significant. The starting dates of crops are usually earlier from the Adaptive Savitzky-Golay fitting method.

The standard deviation (SD) of the estimated starting dates of crops from the three different methods was calculated to test the methodological variance. Nearly 60% of total pixels have SDs less than ten days, while only 39.5% of total pixels have SDs between 10 days and 30 days. Moreover, about 0.5% of total pixels show SDs more than one month. The regions with the largest discrepancy among different methods mainly distributed in the midwest of the study area (figure 2D).

3.2. Model comparison and evaluation
Results in the above show the differences in the spatial pattern of starting dates derived from different methods. Here we further compare the extracted starting dates with the observed seedling stage of rice. First, the starting dates value for each station were derived by averaging the starting dates values over a window of 3 pixels by 3 pixels centered on each rice phenological station. Then the consistency between the average starting date and asynchronous observed seedling stage of rice was analyzed. The results revealed that an apparent linear relationship exist between them. The correlation coefficient between extracted starting date and observed starting date from 1999 to 2009 is 0.82 (P<0.01), 0.79 (P<0.01) and 0.66 (P<0.01) for Asymmetric Gaussians, Double logistic, and Adaptive Savitzky-Golay fitting method respectively (table 1). The analyses indicated that the extracted starting date can accurately represent the seedling date of rice.
4. Conclusion
In this study, we investigated the starting date of crop growth in Northeast China with three different fitting methods. We demonstrated that the starting dates of crops occur mostly between May and June in this region. Discrepancy of the estimates exists among the three methods. The Savitzky-Golay has the earliest estimates, while the Asymmetric Gaussians and Double logistic fitting method show similar results. The largest discrepancy among different methods mainly distributed in the midwest of the study area. The starting dates generated in this study had a significant correlation with the observed starting date of rice, which indicated that the extracted starting date can accurately represent the observed seedling stage of crop.
The NDVI time-series were used to extract the spatial-temporal patterns of the phenological features of cropland, compared with the traditional phenological research, the gridded results could provide more detailed information on a regional scale, and may be applied to the model to simulate biogeochemical cycles and cropland productivity. This comparative study is useful to understand the performance of each fitting method, and choose the most appropriate method to improve the accuracy in the application of the phenological parameters.

However, further studies are required in the future. For example, more principal algorithms should be included in the comparison, i.e., Harmonic Analyses of NDVI Time Series (HANTS), Ployfit, wavelet-based filter, and so on. In addition, the optimization of parameters of the methods should be conducted according to the observed phenological records.

References
[1] Brown M E and de Beurs K M 2008 Evaluation of multi-sensor semi-arid crop season parameters based on NDVI and rainfall Remote Sens. Environ 112 5 2261–71.
[2] Parmesan C 2006 Ecological and evolutionary responses to recent climate change Annu Rev Ecol Evol S 37 637–669.
[3] Piao S L, Wang X H, Ciais P, Zhu B, Wang T and Liu J 2011 Changes in satellite-derived vegetation growth trend in temperate and boreal Eurasia from 1982 to 2006 Global Change Biol 17 3228–39.
[4] Boschetti M, Stroppiana D, Brivio P A and Bocchi S 2009 Multi-year monitoring of rice crop phenology through time series analysis of MODIS images Int. J. Remote Sens 30 18 4643–62.
[5] Cong N, Wang T, Nan H, Ma Y, Wang X, Myneni R B and Piao S L 2013 Changes in satellite-derived spring vegetation green-up date and its linkage to climate in China from 1982 to 2010: a multimethod analysis Global Change Biol 19 881–891.
[6] Li Z G, Tang H J, Yang P, Wu W B, Chen Z X, Zhou Q B, Zhang L and Zou J Q 2012 Spatio-temporal responses of cropland phenophases to climate change in Northeast China. J. Geogr. Sci. 22 1 29–45.
[7] Piao S, Fang J Y, Zhou L M and Ciais P and Zhu B 2006 Variations in satellite-derived phenology in China’s temperate vegetation Global Change Biol 12 672–685.
[8] Wang X, Piao S L, Ciais P, Li J, Friedlingstein P, Koven C and Chen A P 2011 Spring temperature change and its implication in the change of vegetation growth in North America from 1982 to 2006. P Natl Acad Sci USA 108 1240–45.
[9] White M A, De Beurs K M, Didan K, Inouye D W, Richardson A D, Jensen O P, O’Keefe J, Zhang G, Nemani R R, Van Leeuwen W J D, Brown J F, De Wit A, Schaepman M, Lin X, Dettinger M, Bailey A S, Kimball J, Schwartz M D, Baldocchi D D, Lee J T and Lauenroth W K 2009 Intercomparison, interpretation, and assessment of spring phenology in North America estimated from remote sensing for 1982–2006 Global Change Bio 15 10 2335–59.
[10] Rahman H and Dedieu G 1994 SMAC: a simplified method for atmospheric correction of satellite measurements in the solar spectrum. Int. J. Remote Sens 15 123–143.
[11] Maisongrande P, Duchemin B and Dedieu G 2004 VEGETATOPM/SPOT – an operational mission for the earth monitoring: presentation of new standard products. Int. J. Remote Sens 25 9–14.
[12] Samanta A, Costa M H, Nunes E L, Vieira S A, Xu L and Myneni R B 2011 Comment on drought-induced reduction in global terrestrial net primary production from 2000 through 2009 Science 333 1093.
[13] Chen J, Jössson P, Tamura M, Gu Z H, Matsushita B and Eklundh L 2004 A simple method for reconstructing a high-quality NDVI time-series data set based on the Savizky–Golay filter. Remote Sens. Environ 91 332–344.
[14] Xiao D P, Tao F L, Liu Y J, Shi W J, Wang M, Liu F S, Zhang S and Zhu Z 2013 Observed changes in winter wheat phenology in the North China Plain for 1981-2009 Int J
Liu J Y, Zhuang D F, Luo D and Xiao X M 2003 Land cover classification of China: integrated analysis of AVHRR imagery and geophysical data. *Int. J. Remote Sens.* **24** 12 2485–2500.

Cong N, Piao S L, Chen A P, Wang X H, Lin X, Chen S P, Han S J, Zhou G S and Zhang X P 2012 Spring vegetation green-up date in China inferred from SPOT NDVI data: A multiple model analysis. *Agr Forest Meteorol* **165** 104–113.

Heumann B W, Seaquist J W, Eklundh L and Jonsson P 2007 AVHRR derived phenological change in the Sahel and Soudan, Africa, 1982-2005 *Remote Sens. Environ.* **108** 4 385–392.