

# Corn Growth Stage Estimation Using Time Series Vegetation Index

Genong (Eugene) Yu, Liping Di  
Center for Spatial Information  
Science & Systems  
George Mason University  
Fairfax, VA, USA  
{gyu,ldi}@gmu.edu

Zhengwei Yang  
National Agricultural Statistics  
Service  
USDA  
Fairfax, VA, USA  
Zhengwei.Yang@nass.usda.gov

Yonglin Shen, Zeqiang Chen, Bei  
Zhang  
Center for Spatial Information  
Science and Systems  
George Mason University  
Fairfax, VA, USA

**Abstract**— Crop growth stage is important information for decision making in many related agricultural sectors. In-time accurate estimation of crop growth stage is desired. Kernel-fitting of time series vegetation indices have shown potential in estimating crop growth stage while tolerant to noisy data and missing data. The challenge to apply such models is dealing with current year when incomplete data are available. This study proposed a progressive double sigmoid model that leverages the existing best model to compensate the incompleteness of data. The progressive double sigmoid modeling algorithm has three stages of estimation: pre-peak, early post-peak, and late post-peak. Simulation results and experiments showed that the progressive version of double sigmoid algorithm solved the problem of fitting model with insufficient data at early stages. Double sigmoid models have been compared with other alternative approaches in different treatment of data analysis. The results showed that double sigmoid models performed better than moving median window smoothing and Savitzky-Golay alone. Further studies may consider optimizing season partitions and thresholds.

**Keywords**—component; corn growth stage, MODIS, NDVI, time series, phenology

## I. INTRODUCTION

Crop growth stage is an important piece of information for decision making activities related to crop management and food security[1]. In-time accurate estimation of crop growth stage is required during growing season. Remote sensing approaches have been proved to be efficient in determining several featured stages[2–4]. Phenological stages are often estimated from the detected feature stages[5–10]. The detection of feature stages relies on the underlying profile of vegetation index. Different crops have different profiles and hence different threshold or feature points should be used in determining the stages[11]. Another challenge is the noises in remotely sensed data due to clouds and atmospheric condition[12], [13]. To enable the operational detection of crop growth stage, we developed an algorithm for automatic detection of crop growth stage from a multi-stage kernel-fitted time series of Normalized Differentiate Vegetation Index (NDVI) during crop growing seasons. As shown in another paper[14], different smoothing approaches have different effect on crop condition assessment results. Hence, adopting different smoothed data in modeling of crop growth will generate

different results. In this paper, multiple experiments have been evaluated.

## II. METHODOLOGY

### A. Remote Sensing Growth Stage Estimation

Use of time series NDVI and other Earth Observation based indices starts with detection of special feature stages and then interpolates other non-detectable stages. The most studied feature stages are onset of green and end of green[15]. They are somewhat loosely related to physiological phenology of vegetations[15], [16]. Depending on the smoothing algorithms and extraction strategies of feature stages, methods can be grouped into thresholds[17], empirical modeling,[18] change ratio thresholds[19], [20], maximum change rate[21], and logistic function fitting[22], [23]. It was reported that double sigmoid function fitting performed well in crop growth stage estimation[7], [22]. Therefore, double sigmoid function was chosen as the primary model to be fitted.

Our study on crop condition assessment has experimented with other alternative models[14]. These models are also considered for comparative study on crop stage estimation.

### B. Double Sigmoid Model Fitting

Double sigmoid is used as the base kernel. The crop growth stage estimation algorithm uses different fitting strategies at different stages. Base kernel is determined from modeling annual NDVI profiles of previous years at pixel level. Double sigmoid (DSig) fitting combines two sigmoid functions as shown in (1).

$$Y = \frac{1}{1+e^{-\frac{t-p_1}{w_1}}} - \frac{1}{1+e^{-\frac{t-p_2}{w_2}}} \quad (1)$$

Where, Y is NDVI value, t is time in days or weeks mostly, and  $p_1$ ,  $p_2$ ,  $w_1$  and  $w_2$  are parameters to be estimated.

### C. Progressive Double Sigmoid Model Fitting Algorithm

To allow the operational detection of crop growth stage in current year when incomplete data are available in early growing season, we differentiate the stages for parameters to be fitted. Three conditioned models are used depending on what stage it is on the NDVI profile position: (1) pre-peak, (2) early post-peak, and (3) later post-peak. At early two stages,

previous year or closely related existing model are used as the base to allow partial parameter fitting with limited data. The transactions between stages of estimation methods are determined by evaluating the root mean square error of modeled results against actual values up to the present day. An empirically determined, small RMSE threshold value is used. In the sequence of three stage models, two models are tried and evaluated. If the next stage model produces lower RMSE than the previous model and lower RMSE than the threshold, the switching of stage model happens from that day on. The following describe the major steps of the algorithm of the adapted crop growth model for estimating the crop growth stages.

- 1) *Selection of good crop DSig model*: The program computes a good model for every pixel (250 m spatial resolution). Considering the locality of models induced by many uncontrolled factors, we select the model as close as the one under study in terms of temporal and spatial closeness. Crop type should be the same. The good model is described using the four parameters,  $p_1$ ,  $p_2$ ,  $w_1$  and  $w_2$ . The “closeness” is defined as follows: a) same crop type, b) temporally close, i.e. within the past one to three years, and c) spatially close, i.e. close pixel’s models if no previous pixel’s model meets the requirements along the temporal dimension.
- 2) *Pre-peak DSig model and crop stage estimate*: Before the occurrence of peak NDVI, all four parameters cannot be estimated using the data of current year. The best estimate of model can be assumed to be similar to the good crop DSig model on the same crop but a shift of position. Therefore, the parameter to be estimated is the shift,  $s$ , as shown in the following function.  $p_1$ ,  $p_2$ ,  $w_1$  and  $w_2$  will be those of the previous good model as shown in (2).

$$Y = \frac{1}{1+e^{-\frac{t-p_1-s}{w_1}}} - \frac{1}{1+e^{-\frac{t-p_2-s}{w_2}}} \quad (2)$$

Where,  $s$  is the only parameter to be estimated.

- 3) *Early post-peak DSig model and crop stage estimate*: After the peak value occurred, it is possible to accurately model the first half sigmoid function and positions, but the span of second sigmoid function is not easy to accurately model. In this case, the width of previous good DSig model will be used. Therefore, the parameters,  $p'_1$ ,  $p'_2$ , and  $w'_1$ , are to be estimated. The width  $w_2$  is directly borrowed from the previous good DSig model as shown in (3).

$$Y = \frac{1}{1+e^{-\frac{t-p'_1}{w'_1}}} - \frac{1}{1+e^{-\frac{t-p'_2}{w_2}}} \quad (3)$$

Where,  $p'_1$ ,  $p'_2$ , and  $w'_1$  are parameters to be estimated.

- 4) *Later post-peak DSig model and crop stage estimate*: Close to the end of growing season, the available data

of the current year allow the model construction with a reasonable accuracy. All four parameters can be estimated using the data of current year as shown in (4).

$$Y = \frac{1}{1+e^{-\frac{t-p'_1}{w'_1}}} - \frac{1}{1+e^{-\frac{t-p'_2}{w'_2}}} \quad (4)$$

Where,  $p'_1$ ,  $p'_2$ ,  $w'_1$ , and  $w'_2$  are parameters to be estimated.

#### D. Corn Phenological Stages

Once the model is built up for the whole year, two searching strategies are used with a given threshold – local threshold and global threshold [24], [25]. Global threshold uses the value given as the actual threshold value to conduct the search while local threshold uses an adjusted threshold against the range of actual data model. Local threshold guarantees that a match threshold day to be found. Global threshold does not guarantee a match threshold day if the fitted model is off.

Five phenological stages of corn are monitored during the growing season in the US [26], [27]. They are 1) *emerged*, 2) *silking*, 3) *dough*, 4) *dent*, and 5) *mature*. Threshold approach is adopted in the study, similar to those in [17], [24], [25]. For the convenience of comparison across different models, we fixedly using thresholds 0.55, 0.75, 0.99, 0.75, and 0.55 [24] to respectively correspond to the five phenological stages of corn - emerged, silking, dough, dent, and mature. These threshold values are tried out with both global and local threshold search. The first half of the threshold search starts from minimum and the second half does from maximum. In other words, threshold search for emerged (0.55), silking (0.75), and dough (0.99) search from minimum, and dent (0.75) and mature (0.55) from maximum.

#### E. Alternatives for Pre-processing, Smoothing, and Stage Estimator

Noises and different ways of handling scaling (mixed pixels) can affect the model performances as shown [14]. The factors taken into consideration are: 1) smoothing algorithms or underline model, 2) threshold strategy, 3) pre-processing approach, and 4) masking strategies. As an extension of the comparative study of crop condition assessment to crop growth stage, we compared 32 experiments as shown in Table 1. For pre-processing of data, we implemented two typical approaches, Best Index Slope Extraction (BISE) [25] and linear interpolation (LI) [24], to deal with missing data and noisy data. In Table 1, we used the same name convention for smoothing algorithm and masking as those in [14], to facilitate the easy cross reference to the behaviors of experiments between those on crop condition assessment and on crop stage estimation.

TABLE I. CROP STAGE ESTIMATION EXPERIMENTS

Experiment	Description			
	Smooth <sup>a</sup>	T <sup>b</sup>	Pre <sup>c</sup>	Mask <sup>d</sup>
dsigglobndvi	DSIG	G	B	M0
dsigglobndvi.mask100	DSIG	G	B	M100

Experiment	Description			
	Smooth <sup>a</sup>	T <sup>b</sup>	Pre <sup>c</sup>	Mask <sup>d</sup>
dsigglobndvi.mask90	DSIG	G	B	M90
dsigglobsm4253Htwice3ndvi.mask100	4253	G	B	M100
dsigglobsmbsplinendvi.mask100	BS	G	B	M100
dsigglobsmdsigndvi.mask100	DSIG2	G	B	M100
dsigglobsmpoly5ndvi.mask100	POL5	G	B	M100
dsigglobsmavgol3ndvi.mask100	SG	G	B	M100
dsigndvi	DSIG	L	N	M0
dsigndvi.mask100	DSIG	L	N	M100
dsigndvi.mask90	DSIG	L	N	M90
dsigsm4253Htwice3ndvi.mask100	4253	L	B	M100
dsigsmbbsplinendvi.mask100	BS	L	B	M100
dsigsmdsigndvi.mask100	DSIG	L	B	M100
dsigsmpoly5ndvi.mask100	POL5	L	B	M100
dsigsmsavgol3ndvi.mask100	SG	L	B	M100
globndvi	S0	G	LI	M0
globndvi.mask100	S0	G	LI	M100
globndvi.mask90	S0	G	LI	M90
globsm4253Htwice3ndvi.mask100	4253	G	LI	M100
globsmbsplinendvi.mask100	BS	G	N	M100
globsmdsigndvi.mask100	DSIG	G	N	M100
globsmpoly5ndvi.mask100	POL5	G	N	M100
globsmavgol3ndvi.mask100	SG	G	N	M100
locndvi	S0	L	LI	M0
locndvi.mask100	S0	L	LI	M100
locndvi.mask90	S0	L	LI	M90
locsm4253Htwice3ndvi.mask100	4253	L	LI	M100
locsmbsplinendvi.mask100	BS	L	N	M100
locsmdsigndvi.mask100	DSIG	L	N	M100
locsmpoly5ndvi.mask100	POL5	L	N	M100
locsmavgol3ndvi.mask100	SG	L	N	M100

- a. Smoothing algorithms include 1) S0 – no smoothing algorithm applied; 2) 4253 – smoothing algorithm 4253H, Twice; 3) BS-cubic B-Spline; 4) DSIG – double sigmoid fitting; 5) POL5 – degree 5 polynomial fitting; 6) SG – Savitzky-Golay filtering; 7)DSIG2 – twice double sigmoid fitting.
- b. Threshold approach: 1) G – global threshold; 2) L – local threshold;
- c. Pre-processing: 1) B - Best Index Slope Extraction; 2) LI – linear interpolation; 3) N – no special pre-processing.

## F. Evaluation Framework

Fig. 1 shows the overall workflow to evaluate different approaches. USDA NASS QuickStats[28] was used to retrieve the survey data online. There are steps to estimate crop growth stages using remotely sensed data, i.e. 1) pre-processing, 2) model fitting/smoothing, and 3) growth stage extraction.

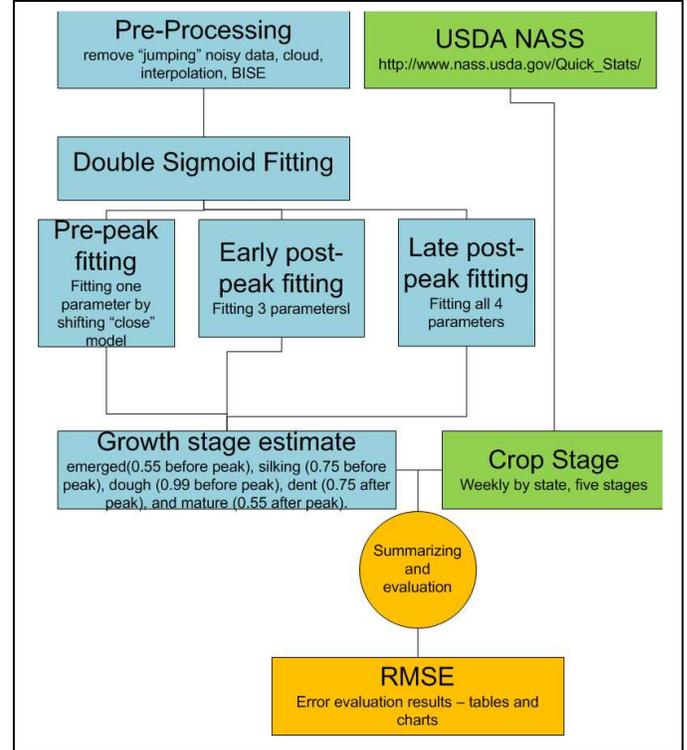


Figure 1. A workflow for evaluating crop stage estimate

The root mean square error (RMSE) between estimated and surveyed percentage is done at state level which is the level where crop growth reports publicly available during growing season. The estimated stage percentage is summarized from corn pixels of a state.

## G. Evaluation Dataset

Daily NDVI dataset for the whole year of 2006 were retrieved from VegScape[29]. Quality layer of original MODIS Surface Reflectance Dataset was used to flag cloud pixels[30]. For testing the progressive model and evaluate its behavior in details, we used the average daily NDVI series of 2009 and 2010 in Black Hawk, Iowa, U.S.A.

## III. RESULTS AND DISCUSSIONS

### A. Double sigmoid crop stage estimation

Fig. 2 shows the double sigmoid model fitted to the average daily NDVI series of year 2009 in Black Hawk County, Iowa, USA. The fitted model represents well all the important features of interest. Peak has good match and two tails end at the same days. With BISE, out envelope (green circles) was captured well.

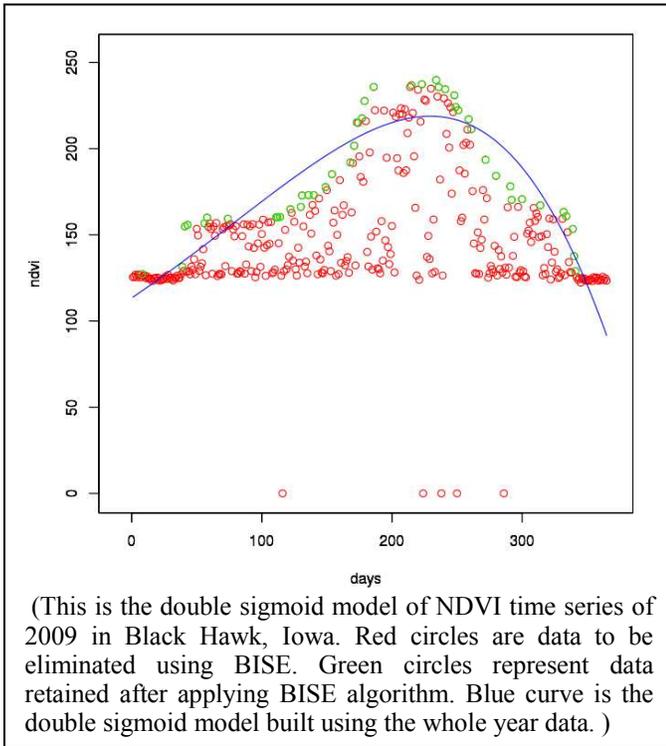


Figure 2. Double sigmoid fitting for year 2006, Black Hawk, Iowa, US

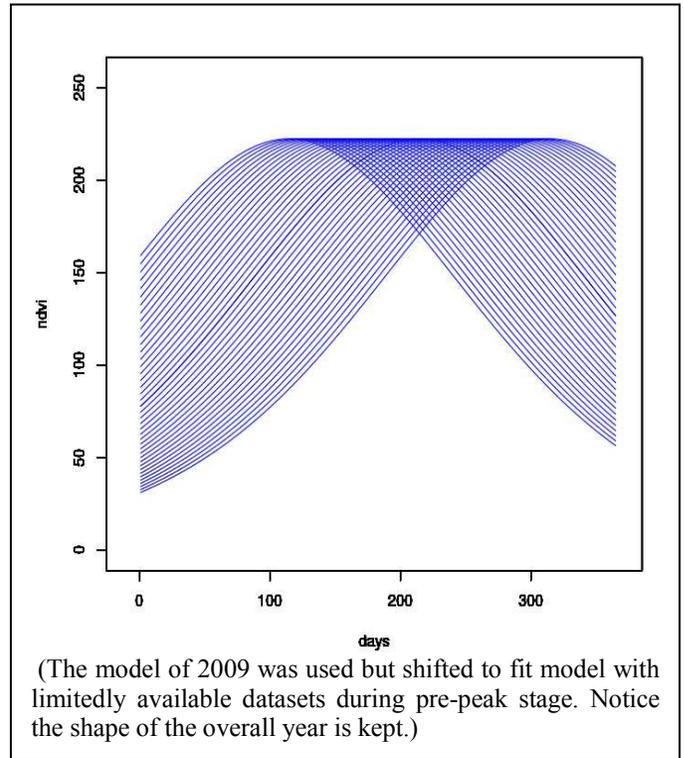


Figure 3. Effect of shifting existing model

Fig. 3 shows the animated effect of shifting an existing model. This demonstrates what the progressive DSig algorithm does during pre-peak stage. The resulted model reserves the curve form. Fig. 4 shows the case when the program forcefully constructs the model without using conditioned models with partial parameters from an existing good model. The severe distortion of constructed models is obvious. Fig. 5 shows the results when the progressive double sigmoid model is applied. With the special handling of three stages, the distortions of models were significantly reduced.

Overall, the results of constructing and executing the models against time series NDVI in Black Hawk, Iowa showed the feasibility of the algorithm in detecting corn growth stages. Further development of the algorithm is to seek improvements in model selection and comparing with varieties of Markov models.

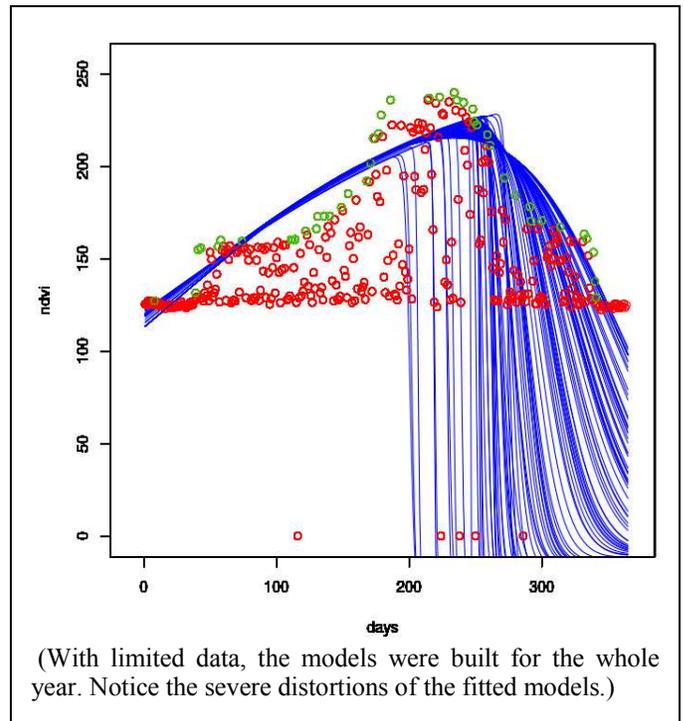


Figure 4. Yearly double sigmoid models

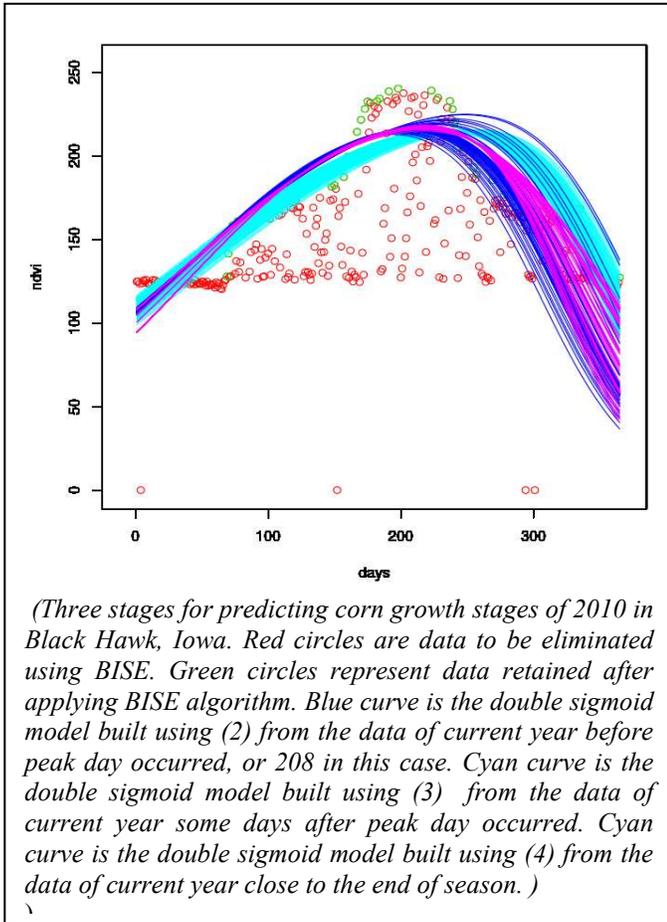


Figure 5. Progressive double sigmoid models

### B. Results of Experiments

Fig. 7 and 8 give the RMSE results of 32 experiments for the year of 2006, Iowa, US. The following can be observed from the results.

1) *Double sigmoid models have relatively low RMSE:* All experiments with double sigmoid models applied have relatively lower RMSE than the rest. All experiments with DSig are less than 50% while others without DSig, such as 4253H, Twice, Savitzky-Golay, and non-smoothed model, have RMSE higher than 50%.

2) *The lowest RMSE is still higher than 20%:* The experiments with lowest RMSE are locmpoly5ndvi.mask100 (RMSE: 24.56%), globmsbsplindvi.mask100 (RMSE: 24.6%) and dsigglobmsavgov3ndvi.mask100 (RMSE: 35.4%). Their values of RMSE are still quite high. The causes for such high RMSE may be related to the selection of thresholds. The thresholds are fixed without going through any adjustment to the area under study. Lower RMSE should be achieved if improved thresholds are chosen based on analyzing historical dataset.

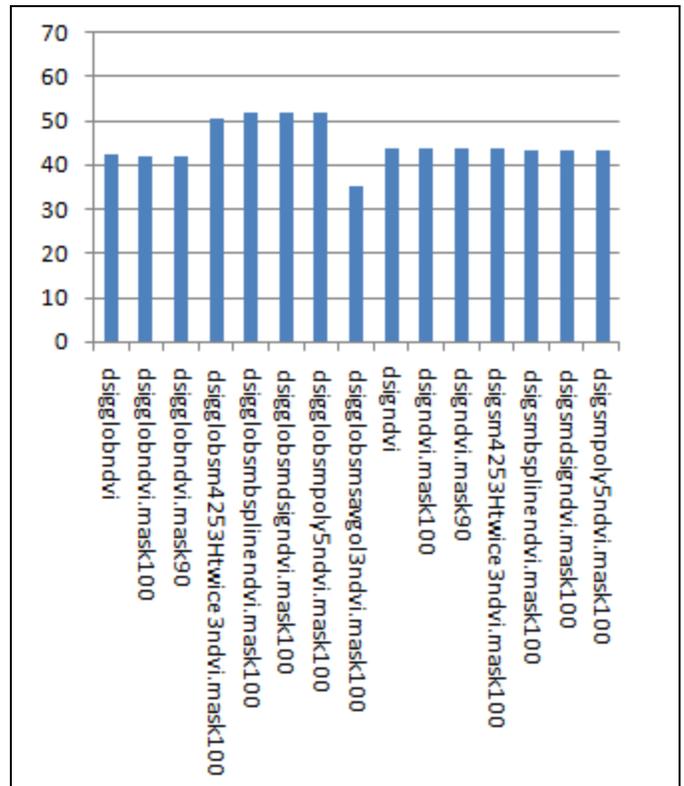


Figure 6. RMSE Results of Corn Growth Stage with Different Experiments in 2006, Iowa, USA (Part 1/2)

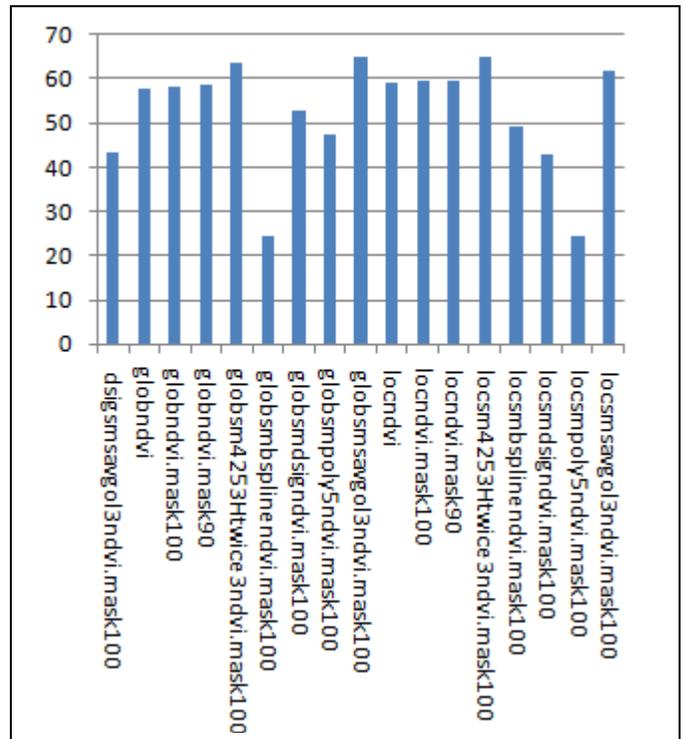


Figure 7. RMSE Results of Corn Growth Stage with Different Experiments in 2006, Iowa, USA (Part 2/2)

#### IV. CONCLUSIONS

A progressive double sigmoid modeling algorithm was proposed to improve the model fitting of current year when limited data are available. Illustrations from modeling time series of NDVI showed that the progressive model maintained the curve shape and reduced the effect of distorted models in early half of the growing season. Comparing to other alternative models with different smoothing algorithms, pre-processing measures, threshold search strategies, and scaling, DSig models had a relative low RMSE.

Further improvements on the models should consider the optimization of thresholds by using historical data to train and select thresholds. Another potential to improve the accuracy of DSig model fitting is partition the growing season and limit the DSig model to fit mono-modal part over a year of daily NDVI series.

#### ACKNOWLEDGMENT

This study is currently supported by NASA project NNX09AO14G (PI: Dr. Liping Di, Co-I: Dr. Zhengwei Yang and Dr. Genong Yu).

#### REFERENCES

- [1] P. Garcia, S. H. Irwin, R. M. Leuthold, and L. Yang, "The value of public information in commodity futures markets," *Journal of Economic Behavior & Organization*, vol. 32, no. 4, pp. 559–570, Apr. 1997.
- [2] J. Meng and B. Wu, "Study on the Crop Condition Monitoring Methods with Remote Sensing," in *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.*, Beijing, China, 2008, vol. Vol. XXXVII, Part. B8, pp. 945–956.
- [3] P. C. Doraiswamy, J. L. Hatfield, T. J. Jackson, B. Akhmedov, J. Prueger, and A. Stern, "Crop condition and yield simulations using Landsat and MODIS," *Remote Sensing of Environment*, vol. 92, no. 4, pp. 548–559, Sep. 2004.
- [4] M. S. Moran, D. C. Hymer, J. Qi, and Y. Kerr, "Comparison of ERS-2 SAR and Landsat TM imagery for monitoring agricultural crop and soil conditions," *Remote Sensing of Environment*, vol. 79, no. 2–3, pp. 243–252, Feb. 2002.
- [5] H. Zhao, Z. Yang, L. Di, and Z. Pei, "Evaluation of Temporal Resolution Effect in Remote Sensing Based Crop Phenology Detection Studies," in *Computer and Computing Technologies in Agriculture V*, vol. 369, D. Li and Y. Chen, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 135–150.
- [6] T. Sakamoto, M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka, and H. Ohno, "A crop phenology detection method using time-series MODIS data," *Remote Sensing of Environment*, vol. 96, no. 3–4, pp. 366–374, Jun. 2005.
- [7] H. Zhao, Z. Yang, L. Di, L. Li, and H. Zhu, "Crop phenology date estimation based on NDVI derived from the reconstructed MODIS daily surface reflectance data," in *2009 17th International Conference on Geoinformatics*, Fairfax, VA, USA, 2009, pp. 1–6.
- [8] P. Jonsson and L. Eklundh, "Seasonality extraction by function fitting to time-series of satellite sensor data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 8, pp. 1824 – 1832, Aug. 2002.
- [9] K. P. Gallo and T. K. Flesch, "Large-area crop monitoring with the NOAA AVHRR: Estimating the silking stage of corn development," *Remote Sensing of Environment*, vol. 27, no. 1, pp. 73–80, Jan. 1989.
- [10] A. Pimstein, J. U. H. Eitel, D. S. Long, I. Mufradi, A. Karnieli, and D. J. Bonfil, "A spectral index to monitor the head-emergence of wheat in semi-arid conditions," *Field Crops Research*, vol. 111, no. 3, pp. 218–225, Apr. 2009.
- [11] I. Masialetti, S. Egbert, and B. D. Wardlow, "A Comparative Analysis of Phenological Curves for Major Crops in Kansas," *GIScience & Remote Sensing*, vol. 47, no. 2, pp. 241–259, Apr. 2010.
- [12] J. N. Hird and G. J. McDermid, "Noise reduction of NDVI time series: An empirical comparison of selected techniques," *Remote Sensing of Environment*, vol. 113, no. 1, pp. 248–258, Jan. 2009.
- [13] NASS, "NASS - Surveys - Remotely Sensed Data for Crop Condition and Crop Yield," 27-May-2009. [Online]. Available: [http://www.nass.usda.gov/Surveys/Remotely\\_Sensed\\_Data\\_Crop\\_Yield/index.asp](http://www.nass.usda.gov/Surveys/Remotely_Sensed_Data_Crop_Yield/index.asp). [Accessed: 04-May-2011].
- [14] G. Yu, L. Di, Z. Yang, Z. Chen, and B. Zhang, "Crop condition assessment using high temporal resolution satellite images," in *The First International Conference on Agro-Geoinformatics 2012*, Shanghai, China, 2012.
- [15] M. D. Schwartz, B. C. Reed, and M. A. White, "Assessing satellite-derived start-of-season measures in the conterminous USA," *International Journal of Climatology*, vol. 22, no. 14, pp. 1793–1805, 2002.
- [16] B. C. Reed, M. D. Schwartz, and X. Xiao, "Remote Sensing Phenology," in *Phenology of Ecosystem Processes*, A. Noormets, Ed. Springer New York, 2009, pp. 231–246.
- [17] D. LLOYD, "A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery," *International Journal of Remote Sensing*, vol. 11, no. 12, pp. 2269–2279, 1990.
- [18] S. Moulin, L. Kergoat, N. Viovy, and G. Dedieu, "Global-Scale Assessment of Vegetation Phenology Using NOAA/AVHRR Satellite Measurements," *Journal of Climate*, vol. 10, no. 6, pp. 1154–1170, Jun. 1997.
- [19] M. A. White, P. E. Thornton, S. W. Running, and R. R. Nemani, "Parameterization and Sensitivity Analysis of the BIOME–BGC Terrestrial Ecosystem Model: Net Primary Production Controls," *Earth Interactions*, vol. 4, no. 3, pp. 1–85, Jan. 2000.
- [20] M. A. White, P. E. Thornton, and S. W. Running, "A continental phenology model for monitoring vegetation responses to interannual climatic variability," *Global Biogeochem. Cycles*, vol. 11, no. 2, pp. 217–234, 1997.
- [21] F. Yu, K. P. Price, J. Ellis, and P. Shi, "Response of seasonal vegetation development to climatic variations in eastern central Asia," *Remote Sensing of Environment*, vol. 87, no. 1, pp. 42–54, Sep. 2003.
- [22] P. S. A. Beck, C. Atzberger, K. A. Hogda, B. Johansen, and A. K. Skidmore, "Improved monitoring of vegetation dynamics at very high latitudes: A new method using MODIS NDVI," *Remote Sensing of Environment*, vol. 100, no. 3, pp. 321–334, Feb. 2006.
- [23] X. Zhang, M. A. Friedl, C. B. Schaaf, A. H. Strahler, J. C. F. Hodges, F. Gao, B. C. Reed, and A. Huete, "Monitoring vegetation phenology using MODIS," *Remote Sensing of Environment*, vol. 84, no. 3, pp. 471–475, Mar. 2003.
- [24] M. Lange, "biseVec: Auxiliary functions for phenological data analysis." CRAN, 28-Jun-2011.
- [25] N. Viovy, O. Arino, and A. S. Belward, "The Best Index Slope Extraction (BISE): A method for reducing noise in NDVI time-series," *International Journal of Remote Sensing*, vol. 13, no. 8, pp. 1585–1590, 1992.
- [26] NASS, "NASS - Publications - Current State Crop Progress and Condition Reports." [Online]. Available: [http://www.nass.usda.gov/Publications/State\\_Crop\\_Progress\\_and\\_Condition/index.asp](http://www.nass.usda.gov/Publications/State_Crop_Progress_and_Condition/index.asp). [Accessed: 22-Aug-2012].
- [27] NASS, "National Crop Progress - Terms and Definitions." [Online]. Available: [http://www.nass.usda.gov/Publications/National\\_Crop\\_Progress/Terms\\_and\\_Definitions/index.asp](http://www.nass.usda.gov/Publications/National_Crop_Progress/Terms_and_Definitions/index.asp). [Accessed: 22-Aug-2012].
- [28] NASS, "NASS - Quick Stats." [Online]. Available: [http://www.nass.usda.gov/Quick\\_Stats/](http://www.nass.usda.gov/Quick_Stats/). [Accessed: 22-Aug-2012].
- [29] CSISS, "VegScape - Vegetation Condition Explorer." [Online]. Available: <http://dss.csiss.gmu.edu/VegScape/>. [Accessed: 22-Aug-2012].
- [30] NASA, "MOD09GQ: Surface Reflectance Daily L2G Global 250m." [Online]. Available: [https://lpdaac.usgs.gov/products/modis\\_products\\_table/mod09gq](https://lpdaac.usgs.gov/products/modis_products_table/mod09gq). [Accessed: 22-Aug-2012].