Model of soybean NDVI change based on time series

Zhang Zhitao¹,²,³, Yubin Lan ²*, Wu Pute³, Han Wenting³

¹. College of Water Resources and Architectural Engineering, Northwest A&F University, Yangling 712100, China; ². Aerial Application Technology, USDA-ARS-SPARC-AATRU, College Station, TX 77845, USA; ³. Institute of Water Saving Agriculture in Arid Area of China, Northwest A&F University, Yangling 712100, China

Abstract: Normalized Difference Vegetation Index (NDVI) has been found to have good correlations with many physical properties of soybean surfaces. Due to the factors of air temperature, humidity, solar radiation, soil moisture, etc., NDVI of soybean varies dynamically in a day. The establishment of the soybean NDVI prediction model at different times in a day can effectively modify this variation. The soybean NDVI values are continuously monitored in hours during soybean seeding, flowering & podding and maturating stages by way of Green Seeker. Results show that the trend of NDVI change every day in the three stages is taken on as a reverse parabola. The NDVI value reaches to the maximum at 8 am or 9 am and decreases to its minimum at 2 pm before a moderate rise. A model for intraday and long-term NDVI change for soybean is built. The test of the model with independent data indicates that the precision meets the demands, with the root mean square error (RMSE) of each day being 3.95, 5.45 and 2.86 for the seeding stage, the bean podding stage and the maturation period, respectively. The prediction RMSEs of the soybean NDVI model for soybeans of the three stages for the fifth day are 5.75, 2.65 and 5.51, respectively and the prediction RMSEs for the sixth day are 9.74, 2.82 and 14.04, respectively according to the data from the first four days.

Keywords: model, NDVI, monitoring time, time series, atmospheric radiation, soybean

DOI: 10.3965/j.ijabe.20140705.007


1 Introduction

Normalized difference vegetation index (NDVI) is defined as a ratio of the difference between the infrared (NIR) and the red bands versus the sum of the two bands \(^1,2\). NDVI has been used in agriculture to characterize the length of the growing seasons and to estimate vegetation water content\(^3-6\). These indexes can be used in conjunction with agronomic references to quantify basic nutrient response, crop condition, yield potential, stress, and pest impact\(^7,9\).

NDVI has become a popular tool for assessing different aspects of soybean, but it may be affected by several factors. For example, NDVI can saturate at relatively low leaf area index (LAI) values\(^10,11\). NDVI is based on the red (RED) and near infrared (NIR) bands, so they are located in the strong chlorophyll absorption region and high reflectance plateau of vegetation canopies, respectively. Therefore, NDVI represents chlorophyll rather than water content\(^12,13\). Some researches demonstrated that the reflection in the NIR and red bands can be related to the fraction of absorbed photosynthetic-active radiation\(^14-17\). However, such relationships are not completely linear with the deviations from the linearity caused by atmospheric effects,
sun/sensor geometry, soil background influence, and the presence of non-photosynthetic vegetation within the canopies\cite{18,19}. Various works reported a decline in the photosynthetic exchange rates during the afternoon when light intensities are similar to those in the morning\cite{20,21}. The combination of an afternoon decline and a midday depression leads to a two-peak and dissymmetric diurnal curve of photosynthesis. These curves are referred to as a transpiration plateu by Lynn and Carlson\cite{22} and Carlson et al\cite{23}.

These factors change every day with temporal fluctuations leading to false variations in the NDVI of soybean, resulting in fictitious variations of the surface even without actual change. The NDVI of soybean thus allows monitoring vegetation photosynthesis through time and enables easy temporal and spatial comparisons\cite{24,25}. However, the uncertainties relevant to the false change are not quantified generally in the studies. It is therefore unclear to what extent the NDVI-based monitoring would be reliable.

In this study, the soybean NDVI values are continuously monitored by hour during soybean’s seeding, flowering & podding, and maturating stages by way of Green Seeker. The soybean NDVI intraday variations in the three stages are analyzed. The changing process is fitted in mathematic formulae to establish a prediction model for soybean NDVI daily variations.

2 Material and methods

2.1 Overview of study sites

The study was conducted at the USDA-ARS experiment station in College Station, Texas (30°31’17 N, 96°23’56 W). The subtropical location has an average annual temperature at 20.7°C and an average annual precipitation at 1008 mm. The average altitude here is 94 m and the soil is sandy loam.

2.2 Materials and study design

Soybean was planted on April 23, 2012 with a row spacing of 1 m. The study was conducted in three growth stages: the seeding stage from May 21, 2012 to May 26, the flowering & podding stage from June 17 to June 22, and the maturating stage from July 16 to July 21. During the test, three sample points were randomly selected in the field, each with a length of about 3 m along the ridge. From 8:00 to 18:00, these sample points were hourly measured. In each stage, measurement was carried out continuously for 6 days when it was sunny.

2.3 NDVI determination

In this study, gauge Trimble Green Seeker was applied to collect NDVI data in three stages. The NDVI measurement range varied from 0 to 0.99\cite{26,27}. When measuring, the head of spectral searchlight was made parallel to the crop canopy, leaving a vertical height of about 0.6 m from the crop canopy. Every sample point was used to collect around 100 NDVI data at a time. The measurement results were directly stored in Green Seeker Hand-held, and then were transmitted to the computer in Excel data format through Hand-held. Their average values were calculated as the NDVI value for each sample point and the average of these three NDVI values was calculated as the final value\cite{28}.

2.4 Data analyses

Microsoft Excel 2003 and SPSS13.0 were employed to analyze data. Sigma Plot 10.0 was used to draw figures and to fit mathematical models.

3 Results and analyses

3.1 Soybean NDVI variation curves

Figures 1a – c are soybean NDVI variation curves for 6 days in a row. Figure 1a is in the seeding stage, 1b in the flowering & podding stage and 1c in the maturating stage. Figure 1d stands for the average daily NDVI variation curve during these three stages.

It can be seen from Figures 1a – c that soybean NDVI values varied dynamically at different times in a day. NDVI decreased gradually from 8:00 am, reached to its minimum at 14:00 or 15:00, and then rose again by degrees. The whole dynamic variation process was approximately like a reverse parabola shape. This kind of change is inversely proportional to the change of temperature in a day. In other words, the crop NDVI value reaches to its minimum when the air temperature in a day is at its highest. This result reveals a fact that temperature is the main factor that affects the crop evapotranspiration amount. As the temperature
increases, so does the crop evapotranspiration amount. But in this process, in order to avoid excessive water loss, the blade’s stoma gradually closes, thus weakening the crop evapotranspiration and degrading the photosynthesis. At the point, the crop decreases in absorbing infrared light but increases in reflection. It also gradually decreases in reflecting near-infrared light. Therefore, the crop NDVI value accordingly decreases. Meanwhile, the solar radiation amount exercises the same influence on the crop NDVI as the temperature does.

3.2 Models and analyses

3.2.1 Model theoretical bases

After analyzing the soybean NDVI variations in three stages, it can be concluded that the soybean NDVI daily variations were in a reverse parabola correlation over time, as the formula (1) shows.

\[ NDVI = aT^2 + bT + c \]  \hspace{1cm} (1)

where, \( T \) stands for time in 24 hours; \( a, b, c \) are the parabolic coefficients.

In the study, NDVI measurement in each growth period was carried out continuously for 6 days. Data collected in the other 5 days were used to construct models and data collected in the fifth day were used to examine models and evaluate their accuracy.

3.2.2 Model in the seeding stage

The NDVI data in soybean seeding stage are used to fit models via formula (1). The fitting parameters concerning the relation between soybean NDVI variations and time changes are shown in Table 1.

As can be seen from Table 1, the determination fitting parameters vary from 0.932 to 0.737 with a good correlation. Hence, the relation between soybean NDVI variations and time changes in seeding stage can be described as:
\[ NDVI = m(0.004T^2 - 0.112T + 0.582) + 0.685 \]  
(2)

where, \( m \) stands for the calculation coefficient.

**Table 1  The fitting parameters between soybean NDVI and time in seeding stage**

<table>
<thead>
<tr>
<th>Days</th>
<th>Seeding</th>
<th>( a )</th>
<th>( b )</th>
<th>( c )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td></td>
<td>0.0044</td>
<td>-0.1231</td>
<td>1.4476</td>
<td>0.843</td>
</tr>
<tr>
<td>Second</td>
<td></td>
<td>0.0045</td>
<td>-0.1218</td>
<td>1.4682</td>
<td>0.743</td>
</tr>
<tr>
<td>Third</td>
<td></td>
<td>0.0052</td>
<td>-0.1417</td>
<td>1.5934</td>
<td>0.737</td>
</tr>
<tr>
<td>Fourth</td>
<td></td>
<td>0.0063</td>
<td>-0.1744</td>
<td>1.8122</td>
<td>0.932</td>
</tr>
<tr>
<td>Sixth</td>
<td></td>
<td>0.0047</td>
<td>-0.1394</td>
<td>1.6888</td>
<td>0.905</td>
</tr>
</tbody>
</table>

It can be inferred from Table 1 that parameters \( a, b \) and \( c \) first increased and then decreased as the measurement went on day by day. They were in a parabolic correlation with the number of days. Hence, the relation between the variation of coefficient \( m \) and the number of days in the seeding stage can be described as:

\[ m = eD^2 - fD + g \]  
(3)

where, \( D \) stands for the number of days for continuous measurement; \( e, f \) and \( g \) are the parabolic coefficients.

After the data in Table 1 being used to fit models via the formula (3), the relation between the variation of coefficient \( m \) and the number of days in the seeding stage can be described as in Formula (4). Its correlation coefficient \( R^2 = 0.898 \).

\[ m = 0.013D^2 - 0.1763D + 1.5199 \]  
(4)

**3.2.3 Model in the flowering & podding stage**

The NDVI data in soybean’s flowering & podding stage were used to fit models via Formula (1). The fitting parameters concerning the relation between soybean NDVI variations and time changes are shown in Table 2.

**Table 2  The fitting parameters between soybean NDVI and time in flowering & podding stage**

<table>
<thead>
<tr>
<th>Days</th>
<th>Flowering &amp; Podding</th>
<th>( a )</th>
<th>( b )</th>
<th>( c )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td></td>
<td>0.0025</td>
<td>-0.0698</td>
<td>1.3467</td>
<td>0.804</td>
</tr>
<tr>
<td>Second</td>
<td></td>
<td>0.002</td>
<td>-0.0573</td>
<td>1.2543</td>
<td>0.765</td>
</tr>
<tr>
<td>Third</td>
<td></td>
<td>0.0033</td>
<td>-0.0909</td>
<td>1.4602</td>
<td>0.807</td>
</tr>
<tr>
<td>Fourth</td>
<td></td>
<td>0.0028</td>
<td>-0.0806</td>
<td>1.3958</td>
<td>0.804</td>
</tr>
<tr>
<td>Sixth</td>
<td></td>
<td>0.0028</td>
<td>-0.0788</td>
<td>1.3779</td>
<td>0.783</td>
</tr>
</tbody>
</table>

As can be seen from Table 2, the determination fitting parameters vary from 0.807 to 0.765 with a good correlation. Hence, the relation between soybean NDVI variations and time changes in flowering & podding stage can be described as:

\[ NDVI = m(0.002T^2 - 0.056T + 0.281) + 0.965 \]  
(5)

After the data in Table 2 being used to fit models via Formula (3), the relation between the variation of coefficient \( m \) and the number of days in the flowering & podding stage can be described as in Formula (6). Its correlation coefficient \( R^2 = 0.951 \).

\[ m = -0.0499D^2 + 0.3663D - 0.5097 \]  
(6)

**3.2.4 Model in the maturating stage**

The NDVI data in the soybean maturating stage were used to fit models via Formula (1). The fitting parameters concerning the relation between soybean NDVI variations and time changes are shown in Table 3.

**Table 3  The fitting parameters between soybean NDVI and time in maturating stage**

<table>
<thead>
<tr>
<th>Days</th>
<th>Maturating</th>
<th>( a )</th>
<th>( b )</th>
<th>( c )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>First</td>
<td></td>
<td>0.0038</td>
<td>-0.1032</td>
<td>1.3852</td>
<td>0.796</td>
</tr>
<tr>
<td>Second</td>
<td></td>
<td>0.0044</td>
<td>-0.1229</td>
<td>1.5274</td>
<td>0.924</td>
</tr>
<tr>
<td>Third</td>
<td></td>
<td>0.0049</td>
<td>-0.1338</td>
<td>1.5767</td>
<td>0.894</td>
</tr>
<tr>
<td>Fourth</td>
<td></td>
<td>0.0057</td>
<td>-0.1548</td>
<td>1.7029</td>
<td>0.892</td>
</tr>
<tr>
<td>Sixth</td>
<td></td>
<td>0.0048</td>
<td>-0.1344</td>
<td>1.5452</td>
<td>0.942</td>
</tr>
</tbody>
</table>

As can be seen from Table 3, the determination fitting parameters vary from 0.942 to 0.796 with a good correlation. Hence, the relation between soybean NDVI variations and time changes in the maturating stage can be described as:

\[ NDVI = m(0.0039T^2 - 0.111T + 0.573) + 0.687 \]  
(7)

After the data in Table 3 being used to fit models via Formula (3), the relation between the variation of coefficient \( m \) and the number of days in maturating stage can be described as in Formula (8). Its correlation coefficient \( R^2 = 0.975 \).

\[ m = 0.0154D^2 - 0.0401D + 0.9285 \]  
(8)

**3.2.5 Model for all stages**

As can be seen from Formulas (2), (5) and (7), the soybean NDVI values share the same trend in seeding, flowering & podding and maturating stages. Therefore, these three formulae can be integrated into the general Formula (9).

\[ NDVI = (N+1)(0.002T^2 - 0.056T + 0.281) + 0.965 \]  
(9)

where, \( N \) stands for the coefficient in different soybean growth stages.
And the relation between coefficient $N$ and the number of days for continuous measurement can be shown as:

$$N = iD^2 + jD + k$$

(10)

where, $D$ stands for the number of days for continuous measurement; $i$, $j$ and $k$ are coefficients.

As long as a soybean NDVI value was measured at any time during a day, the value of $N$ in that day can be calculated through Formula (9). After the value of $N$ being put in Formula (9), soybean NDVI value at any point in that day can be calculated.

If the values of $N$ in three successive days were obtained, coefficients $i$, $j$ and $k$ in Formula (10) can be calculated. Then Formula (10) can be employed to predict the values of $N$ in following days. Moreover, these future values of $N$ can be put in Formula (9) to predict the soybean NDVI values at different times in following days.

### 3.3 Model accuracy validation

The measured data in fifth day during soybean seeding, flowering & podding and maturating stages are used to validate the model accuracy of Formula (9). In the first place, the measured soybean NDVI value at 10:00 in fifth day is put in Formula (9), thus getting the value of $N$ in fifth day. Then, this value of $N$ is put in Formula (9), thus getting NDVI values at different times in fifth day. Lastly, by virtue of the measured data in fifth day, root-mean-square errors (RMSEs) at different times and the whole root-mean-square error can be calculated. The results are presented in Table 4.

<table>
<thead>
<tr>
<th>Stages</th>
<th>8:00</th>
<th>9:00</th>
<th>10:00</th>
<th>11:00</th>
<th>12:00</th>
<th>13:00</th>
<th>14:00</th>
<th>15:00</th>
<th>16:00</th>
<th>17:00</th>
<th>18:00</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeding</td>
<td>-3.15</td>
<td>3.47</td>
<td>2.73</td>
<td>0.01</td>
<td>0.16</td>
<td>-1.83</td>
<td>-7.8</td>
<td>-1.96</td>
<td>-1.9</td>
<td>-6.85</td>
<td>-4.84</td>
<td>3.95</td>
</tr>
<tr>
<td>Flowering &amp; podding</td>
<td>-5.84</td>
<td>-1.75</td>
<td>-0.73</td>
<td>-2.06</td>
<td>-4.77</td>
<td>-5.84</td>
<td>-8.96</td>
<td>-7.95</td>
<td>-5.58</td>
<td>-5.19</td>
<td>-5.06</td>
<td>5.45</td>
</tr>
<tr>
<td>Maturating</td>
<td>-7.53</td>
<td>0.28</td>
<td>-0.81</td>
<td>0.96</td>
<td>2.22</td>
<td>-0.08</td>
<td>0.35</td>
<td>3.6</td>
<td>2.58</td>
<td>-0.02</td>
<td>-2.61</td>
<td>2.86</td>
</tr>
</tbody>
</table>

As can be concluded from Table 4, the prediction RMSEs of the soybean NDVI models were 3.95, 5.45 and 2.86, respectively in soybean’s seeding, flowering & podding and maturating stages. In the seeding stage, the maximum of prediction RMSE was -7.80 and the minimum was 0.01. In the flowering & podding stage, the maximum of prediction RMSE was -8.96 and the minimum was -0.73. In the maturating stage, the maximum of prediction RMSE was -7.53 and the minimum was -0.02. It can be also concluded from Table 4 that the maximum prediction errors usually occur at 8:00 am and any time between 14:00 to 15:00. To analyze with a combination of results from Figure 1, the crop NDVI was largely affected by air humidity at 8:00, so its prediction accuracy is low; the NDVI value was greatly influenced by air temperature, evapotranspiration amount and photosynthesis between 14:00 to 15:00, so its prediction error was large.

The NDVI data at 10:00 in the first four days were measured in soybean’s seeding, flowering & podding and maturating stages. After that, the value of NDVI in each day was calculated through Formula (9). Then, the coefficient in Formula (10) was gained by virtue of the values of $N$ in the first four days. After the calculated coefficient being put in Formula (10), the values of $N$ in fifth and sixth days were prediction through Formula (10). Finally, these prediction values of $N$ were put in Formula (9) to calculate NDVI values at different times in the fifth and sixth days. Compared with the measured data in the fifth and sixth days, prediction RMSEs and the whole RMSE were calculated separately. The results are presented in Table 5.

<table>
<thead>
<tr>
<th>Days</th>
<th>Stages</th>
<th>8:00</th>
<th>9:00</th>
<th>10:00</th>
<th>11:00</th>
<th>12:00</th>
<th>13:00</th>
<th>14:00</th>
<th>15:00</th>
<th>16:00</th>
<th>17:00</th>
<th>18:00</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fifth</td>
<td>Seeding</td>
<td>-0.01</td>
<td>6.77</td>
<td>8.53</td>
<td>12.13</td>
<td>3.48</td>
<td>8.11</td>
<td>0.54</td>
<td>4.19</td>
<td>0.67</td>
<td>-0.76</td>
<td>-1.09</td>
<td>5.75</td>
</tr>
<tr>
<td></td>
<td>Flowering &amp; podding</td>
<td>-4.45</td>
<td>0.36</td>
<td>2.01</td>
<td>1.25</td>
<td>-0.98</td>
<td>-1.77</td>
<td>-4.69</td>
<td>-3.8</td>
<td>-1.77</td>
<td>-1.79</td>
<td>-2.21</td>
<td>2.65</td>
</tr>
<tr>
<td></td>
<td>Maturating</td>
<td>-9.46</td>
<td>-2.74</td>
<td>-5.05</td>
<td>-4.23</td>
<td>-3.71</td>
<td>-6.68</td>
<td>-6.4</td>
<td>-2.76</td>
<td>-3.33</td>
<td>-5.26</td>
<td>-6.92</td>
<td>5.51</td>
</tr>
<tr>
<td>Sixth</td>
<td>Seeding</td>
<td>1.75</td>
<td>9.54</td>
<td>12.31</td>
<td>16.67</td>
<td>9.27</td>
<td>14.12</td>
<td>7.23</td>
<td>10.46</td>
<td>6.64</td>
<td>4.45</td>
<td>3.08</td>
<td>9.74</td>
</tr>
<tr>
<td></td>
<td>Flowering &amp; podding</td>
<td>-1.87</td>
<td>0.08</td>
<td>2.21</td>
<td>1.65</td>
<td>0.5</td>
<td>-4.63</td>
<td>-5.59</td>
<td>-3.48</td>
<td>-2.85</td>
<td>-0.81</td>
<td>-1.56</td>
<td>2.82</td>
</tr>
</tbody>
</table>
As can be concluded from Table 5, the prediction RMSEs of the soybean NDVI models on the fifth day were 5.75, 2.65 and 5.51, respectively and the prediction RMSEs in the sixth day were 9.74, 2.82 and 14.04, respectively. On the fifth day, the maximum of prediction RMSEs was 12.13 and the minimum was -0.01. On the sixth day, the maximum of prediction RMSEs was 16.67 and the minimum was 0.08. It can be also concluded from Table 5 that the prediction accuracy on the sixth day is lower than that on the fifth day. The measurement error of the soybean NDVI data may account for this prediction error. Therefore, in order to improve the model prediction accuracy, NDVI values at different times in a day can be chosen to calculate the value of N, and then the average is obtained to be work as the value of N.

4 Conclusions

Owing to the influence of air humidity, temperature, solar radiation, soil moisture content and crop photosynthesis, soybean NDVI values vary dynamically at different times of a day. The establishment of the soybean NDVI prediction model can effectively modify this variation.

The model is a simple empirical model. When the soybean NDVI value is prediction, the involved calculations and observations are limited. So it is user-friendly.

Since the study is only concerned with measured data in soybean seeding, flowering & podding and maturing stages, and the studied soil type is but sandy soil. Further study is still called for concerning the soybean NDVI variation trend and prediction models for other climates and with other soil types.

Acknowledgements

This paper was supported by Governmental Public Industry Research Special Funds for Projects (201301016), and National Key Technology R&D Program of the Ministry of Science and Technology (2012BAH29B04-02), and Chinese Universities Scientific Fund, Northwest A&F University (QN2011132), and the Introduction of Intelligence from Abroad Project for Innovation in Academic Schools, China (B12007).

[References]


