Reframing the Red Band with the Red-SWIR Band
(0.74\(\rho_{\text{red}}+0.26\rho_{\text{swir}}\)) Can Reduce the Sensitivity of Vegetation Indices to Soil Background

Xuehong Chen\(^1,2\), Zhengfei Guo\(^1,2\), Jin Chen\(^1,2,\*\), Wei Yang\(^3\), Yanming Yao\(^4\),
Chishan Zhang\(^1,2\), Xihong Cui\(^1,2\) and Xin Cao\(^1,2\)

1 State Key Laboratory of Earth Surface Processes and Resource Ecology, Faculty of Geographical Science,
Beijing Normal University, Beijing 100875, China; chenxuehong@bnu.edu.cn (X.C.);
201621480077@mail.bnu.edu.cn (Z.G.); zhangchishan@mail.bnu.edu.cn (C.Z.);
cuixihong@bnu.edu.cn (X.C.); caoxin@bnu.edu.cn (X.C.)

2 Beijing Engineering Research Center for Global Land Remote Sensing Products,
Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science,
Beijing Normal University, Beijing 100875, China

3 Center for Environmental Remote Sensing, Chiba University, Chiba 263-8522, Japan; yangwei@chiba-u.jp

4 Key Laboratory of Agricultural Remote Sensing (AGRIRS), Ministry of Agriculture/Institute of Agricultural
Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China;

* Correspondence: chenjin@bnu.edu.cn; Tel.: +86-135-2288-9711

Received: 19 February 2019; Accepted: 4 April 2019; Published: 9 April 2019

Abstract: Most vegetation indices (VIs) of remote sensing were designed based on the concept
of soil-line, which represents a linear correlation between bare soil reflectance at the red and
near-infrared (NIR) bands. Unfortunately, the soil-line can only suppress brightness variation,
not color differences of bare soil. Consequently, soil variation has a considerable impact on vegetation
indices, although significant efforts have been devoted to this issue. In this study, a new soil-line is
established in a new feature space of the NIR band and a virtual band that combines the red and
shortwave-infrared (SWIR) bands (0.74\(\rho_{\text{red}}+0.26\rho_{\text{swir}}\)). Then, plus versions of vegetation indices
(VI\(^+\)), i.e., normalized difference vegetation index plus (NDVI\(^+\)), enhanced vegetation index plus
(EVI\(^+\)), soil-adjusted vegetation index plus (SAVI\(^+\)), and modified soil-adjusted vegetation index plus
(MSAVI\(^+\)), are proposed based on the new soil-line, which replaces the red band with the red-SWIR
band in the vegetation indices. Soil spectral data from several spectral libraries confirm that bare soil
has much less variation for VI\(^+\) than the original VI. Simulation experiments show that VI\(^+\) correlates
better with fractional vegetation coverage (FVC) and leaf area index (LAI) than original VI. Ground
measured LAI data collected from BigFoot, VALERI, and other previous references also confirm
that VI\(^+\) derived from Moderate Resolution Imaging Spectroradiometer (MODIS) data correlates
better with ground measured LAI than original VI. These data analyses suggest that replacing the red
band with the red-SWIR band can reduce the sensitivity of VIs to soil background. We recommend
employing the proposed NDVI\(^+\), EVI\(^+\), SAVI\(^+\), and MSAVI\(^+\) in applications of large area, sparse
vegetation, or when soil color variation cannot be neglected, although sensitivity to soil moisture and
clay content might cause slight side effects for the proposed VI\(^+\)s.

Keywords: vegetation index; soil-line; soil sensitivity; red-SWIR band

1. Introduction

Vegetation indices (VIs) provide proxies for measuring canopy greenness and vigor by enhancing
absorptive and reflective features of vegetation in mathematical combinations of different spectral
bands [1,2]. As simple and effective tools, VIs have been widely applied in various vegetation research studies, including vegetation parameter retrieval, phenology monitoring, carbon cycle modeling, and soil erosion estimation [2–7]. An ideal vegetation index is expected to maintain good correlation with vegetation parameters (e.g., leaf area index (LAI), fractional vegetation cover (FVC), etc.) while minimizing sensitivity to soil backgrounds, atmospheric conditions, and sun-target-sensor geometries [8]. Numerous VIs have been developed in the past few decades using various mathematical formulas and spectral band combinations to meet these objectives [9].

Soil effect is found to have considerable impact on VIs [10,11], especially in sparsely-vegetated areas [12]. Soil effect can be divided into two parts: primary variations associated with the varying brightness of bare soils, and secondary variations attributed to “color” differences among bare soils [13]. As shown in Figure 1, brightness differences produce a principal axis of soil spectral variance (a soil-line) in the feature space of the red and near-infrared (NIR) bands, while secondary curve-forming variations add width to the soil-line [13]. In order to reduce the soil effect, a group of soil-adjusted vegetation indices were proposed using different algebraic formulations of the red and NIR bands, including soil-adjusted vegetation index (SAVI) [14], modified soil-adjusted vegetation index (MSAVI) [15], transformed soil-adjusted vegetation index (TSAVI) [16], and optimized soil-adjusted vegetation index OSAVI [17]. The soil and atmosphere resistant vegetation index (SARVI) and enhanced vegetation index (EVI) were further proposed to reduce soil and atmospheric effects [18]. These VIs were designed based on the soil-line concept and they measure the distances or angles between the vegetation-related isolines and the soil-line. Such efforts minimize brightness-related soil effects by considering first-order soil vegetation interaction with soil-adjustment parameters [19]. However, the soil spectrum is affected by various factors (e.g., minerals, organic compounds, moisture, particle size, and soil structure), producing large color variation of soil backgrounds and deviation of soil-lines [20,21]. Consequently, there is no unique global soil-line [22]. Previous efforts in soil adjustment have not been able to address the uncertainty induced by soil color variation, which increases the soil-line width, especially with large-area satellite imagery [19]. In summary, soil heterogeneity, especially color variation, cannot be reduced by existing VIs, and it affects their consistence in large scale analyses [13,23].

![Illustration of soil-line and leaf area index (LAI) isolines (adapted from [24]).](image-url)
Unlike previous efforts to adjust the distances or angles between the vegetation-related isolines and the soil-line, this study attempts to address the uncertainty induced by soil-line width. Aside from the commonly used red and NIR bands, the SWIR band at 1.6 or 2.1 µm has been proven to be helpful for improving the estimation accuracy of LAI and FVC in previous references, as soil background or aerosol contamination can be suppressed if the SWIR band is introduced [25–28]. Unfortunately, the optimal formula incorporating the SWIR band to reduce soil background is not widely accepted. Recently, Wang et al. (2017) [29] used a weighted red-SWIR (shortwave infrared) combination to replace the red band in NDVI, and designed a normalized difference phenology index (NDPI) with a very similar formula as NDVI to reduce the impacts of snow effect. Inspired by these studies, we established a thinner soil-line combining the NIR, red, and SWIR bands, on which the commonly used vegetation indices could be easily modified. Several analyses were employed to investigate whether soil effect can be reduced by the modified VIs. The performance of the proposed VIs in retrieving vegetation parameters (e.g., FVC and LAI) based on simulation experiments and in situ measurement data was evaluated and compared with that of original VIs. The sensitivity to soil moisture and the parameters used to establish a new soil line were also explored to confirm the robustness of the proposed VIs.

2. Improved Vegetation Indices by Replacing Red Band with Red-SWIR Band

2.1. A New “Soil-Line” in the Feature Space of NIR and Red-SWIR Band

In this study, the red band in the feature space of the red-NIR bands is replaced with a weighted combination of the red and SWIR bands used in NDPI [29].

\[
\rho_{\text{swir}}^{\text{red}} = \alpha \rho_{\text{red}} + (1 - \alpha) \rho_{\text{swir}}, \quad \alpha = 0.74, \tag{1}
\]

where \(\rho_{\text{red}}, \rho_{\text{swir}},\) and \(\rho_{\text{swir}}^{\text{red}}\) denote reflectance in the red band (620–670 nm), SWIR band (1628–1652 nm), and the proposed red-SWIR band. The red and SWIR bands correspond to bands 1 and 6 of the Moderate Resolution Imaging Spectroradiometer (MODIS), respectively; \(\alpha\) is the weight of the combination of the red and SWIR bands. To investigate characteristics of traditional and new soil-lines in two feature spaces, 4890 soil spectra were selected from several spectral libraries, including LARS, ICRAF, and ASTER [30,31], which include most worldwide soil types that have large ranges of soil color and water conditions. These spectra were then resampled by the MODIS spectral response function.

As shown in Figure 2a, the traditional soil-line can be observed in the feature space of the red and NIR bands, which is approximately parallel to the 1:1 line and has a considerable width. Traditional VIs are often designed to measure the departure of observed data from this line, using Euclidean distance or angular difference. The new soil-line in the feature space of the NIR and red-SWIR bands exhibits similar linear characteristics but has less width (Figure 2b). The square of correlation coefficient (R squared) the new soil-line is 0.95, which is higher than that of the traditional one (0.91). The root mean square error (RMSE) is also reduced from 0.036 to 0.027. This suggests that the new soil-line not only maintains the utility of a traditional soil-line, but also reduces its width, which may induce uncertainty in VI performance. We further investigated whether 0.74 is the optimal \(\alpha\) value for the new soil-line. As shown in Figure 3, correlation of the new soil-line achieves a high \(\alpha\) value of 0.74, which is consistent with that of NDPI [23].
Figure 2. Comparison of the original (a) and the new (b) soil-lines. The colors, from blue to red, correspond with low to high density.

Figure 3. RMSE (Root mean square error) and R squared (square of correlation coefficient) results of the new soil-line with different $\alpha$ values.

2.2. Improving Vegetation Index-Based New “Soil-Lines”

By replacing the red band with the new red-SWIR band, commonly used VIs can be improved as modified versions (VI$^+$). Four VIs were modified as plus versions in this study:

\[
\text{NDVI}^+ = \frac{\rho_{\text{nir}} - \rho_{\text{swir}}}{\rho_{\text{nir}} + \rho_{\text{swir}}},
\]

\[
\text{SAVI}^+ = \frac{(\rho_{\text{nir}} - \rho_{\text{swir}})(1 + L)}{\rho_{\text{nir}} + \rho_{\text{swir}} + L},
\]

\[
\text{EVI}^+ = \frac{2.5(\rho_{\text{nir}} - \rho_{\text{swir}})}{1 + \rho_{\text{nir}} + 6\rho_{\text{swir}} - 7.5\rho_{\text{blue}}},
\]

\[
\text{MSAVI}^+ = \frac{2\rho_{\text{nir}} + 1 - \sqrt{(2\rho_{\text{nir}} + 1)^2 - 8(\rho_{\text{nir}} - \rho_{\text{swir}})}}{2},
\]
where $\rho_{\text{nir}}$ and $\rho_{\text{blue}}$ are reflectance in the NIR and blue bands corresponding to bands 2 and 3 of MODIS, respectively. It is noticed that NDVI* is the same as NDPI [29]. NDPI is renamed here to maintain consistent naming with the other VI* s. The adjustment factor ($L$) in SAVI is dependent on vegetation coverage. Considering that prior knowledge of vegetation coverage is often unknown, $L$ is usually set to 0.5 [12]. In this study, SAVI and SAVI* s were also calculated with $L = 0.5$.

VI* s and original VIs were calculated for the 4890 soil spectra from spectral libraries. As shown in Figure 4, the variations of the four VI* s (0.0009–0.0016) for bare soils are much lower than those of the original VIs (0.0023–0.0063), indicating that VI* s are much less sensitive to soil spectral variation than original VIs.

**Figure 4.** Histograms of plus versions of vegetation indices (VI* s) and original VIs for soil spectra. (a) Normalized difference vegetation index (NDVI) and NDVI*; (b) enhanced vegetation index (EVI) and EVI*; (c) soil-adjusted vegetation index (SAVI) and SAVI*; (d) modified soil-adjusted vegetation index (MSAVI) and MSAVI*.

3. Experiments

Two simulation experiments and a global in situ measured LAI dataset were employed to compare the performance of VI* s and original VIs in vegetation parameter retrieval under different soil backgrounds. The first simulation experiment based on a linear spectral mixing model, which considers only single-scatter, was employed to evaluate FVC retrieval [32]. The second simulation experiment, based on a scattering by arbitrary inclined leaves (SAIL) model, which considers multiple scattering...
between vegetation and soil, was employed to evaluate LAI retrieval [33]. Finally, in situ LAI data measured around the world were collected to validate the performance of VIs under real scenarios.

3.1. A Simulation Experiment Based on Linear Mixing Model

In this experiment, the 4890 soil spectra mentioned in Section 2 and 35 vegetation spectra from the United States Geological Survey (USGS) spectral library [34] were selected as endmembers. Mixed spectra were, thus, simulated with a linear mixing model:

\[ \rho_{\text{mix}} = f \rho_{\text{veg}} + (1 - f) \rho_{\text{soil}} \]  

where \( \rho_{\text{mix}} \) is the generated mixed spectrum, \( \rho_{\text{veg}} \) and \( \rho_{\text{soil}} \) are the spectra of vegetation and soil, respectively, and \( f \) is the FVC. FVC changed from 0 to 1 with steps of 0.01. Mixed spectra were generated for each step in FVC with \( 4890 \times 35 \) combinations of vegetation and soil endmembers. In total, 17,115,000 mixed spectra were simulated and corresponding VIs were calculated. Figure 5 shows the relationship between FVC and different VIs. VI+ correlate better with FVC than the four original VIs. R squared is improved with 0.05–0.10 and RMSE decreases with 0.015–0.03. The uncertainties of VI+ are much lower than original VIs, especially when FVC is low, indicating that soil effect is well suppressed in VI+.

3.2. A Simulation Experiment Based on PROSAIL (Leaf Optical Properties Spectra Model and Scattering from Arbitrarily Inclined Leaves Canopy Model)

The PROSAIL model [33] was used to generate canopy reflectance in this experiment. A typical leaf spectrum was simulated by PROSPECT5 and the 4890 soil spectra were used as underlying surfaces in the SAIL model. LAI changed from 0.1 to 6 with steps of 0.025. Parameters other than soil spectrum and LAI were fixed (Table 1) to focus on the impact of soil background. Thus, 1,173,600 canopy spectra were simulated in total, and the corresponding VIs were calculated. As the relationships between LAI and VIs are nonlinear (Figure 6), an exponential model was employed [8]. As shown in Figure 6, the VI+ have smaller fitting errors for LAI estimation than original VIs. The improvement of VI+ in this experiment is less obvious because soil effect becomes weak with high LAI.

<table>
<thead>
<tr>
<th>Table 1. Parameter settings in PROSAIL (leaf optical properties spectra model and scattering from arbitrarily inclined leaves canopy model).</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Leaf Parameters in PROSPECT</strong></td>
</tr>
<tr>
<td>Leaf Structural: 1.5</td>
</tr>
<tr>
<td>Chlorophyll: 40 µg/cm²</td>
</tr>
<tr>
<td>Carotenoids: 0.77 µg/cm²</td>
</tr>
<tr>
<td>Water thickness: 0.79 cm</td>
</tr>
<tr>
<td>Dry Matter: 0.80 g/cm²</td>
</tr>
<tr>
<td>Brown pigments: 0</td>
</tr>
</tbody>
</table>
Figure 5. The relationship between fractional vegetation coverage (FVC) and VIs derived from the simulated mixed spectra. (a–d) FVC vs. original VIs, and (e–h) FVC vs. VI+s. The colors from blue to red correspond to the density from low to high.
Figure 6. The relationship between LAI and VIs derived from the simulated canopy spectra. (a–d) LAI vs. original VIs, and (e–h) LAI vs. VI+s. The colors from blue to red correspond with low to high density.
3.3. Global Validation Using In-Situ-Measured LAI

To further confirm the effectiveness of VI+s in real scenarios, we collected several LAI datasets, including BigFoot [35], VALERI [36], and other data in previous references [37–54] for global validation. The ground-measured LAI, location (LAT/LONG), and measured date were collected. Figure 7 shows the spatial distribution of ground-measured LAIs. Meanwhile, the corresponding spectra of a MODIS 8-day surface reflectance product (MOD09A1) were extracted to calculate VIs. In total, 89 pairs of ground-measured LAIs and MODIS-derived VIs were collected and their relationships were investigated. Both linear and exponential regression models were regarded, considering the possible nonlinear relationship between VIs and LAI. Compared with original VIs, VI+s are better explained by LAI for both linear and exponential models (Figure 8). The improvement in R squared ranged from 0.03 to 0.04, indicating that the soil effect in real scenarios can also be effectively reduced by the proposed VI+s.

Considering that the performance of SAVI could be affected by \( L \) value, exponential regression models were used to investigate the relationship between LAI and SAVI/SAVI+ for different \( L \) values. As shown in Figure 9, the R squared values of the LAI-\( \text{SAVI}^+ \) model are always higher than those of the LAI-\( \text{SAVI} \) model. Moreover, the improvement in R squared by SAVI+ ranges from 0.03 to 0.05 compared to SAVI, whereas the improvement in R squared is up to 0.01 with \( L \) adjustment compared to the case with \( L = 0 \). There is no consistent value of \( L \) adjustment for global analysis, because it largely relies on vegetation coverage. In contrast, \( \alpha \) value in VI+s is globally consistent, as it is derived from several spectral libraries, including nearly all the global soil samples.

![Figure 7. Locations of the ground measured LAI.](image-url)
Figure 8. The relationship between ground measured LAI and VIs derived from moderate resolution imaging spectroradiometer (MODIS) data ($n = 89$). (a–d) LAI vs. original VIs, and (e–h) LAI vs. VI+s.
Weight value ($\alpha$) is critical for building the new soil-line and VI+s. It probably varies for different onboard sensors because band ranges and corresponding spectral response functions are inconsistent. Therefore, we resampled the soil spectra with the spectral response functions of several commonly used onboard sensors, including Landsat 5, Landsat 8, Sentinel-2, SPOT 5, and Worldview 3. Then, optimal $\alpha$ values corresponding to the best soil-line (with the highest correlation between the NIR and red-SWIR bands) were determined for different onboard sensors. As shown in Table 2, the optimal $\alpha$ varies from 0.74 to 0.80 for different sensors. Therefore, different $\alpha$ values are recommended when calculating VI+s from different remotely sensed data.

Table 2. Optimal $\alpha$ values for different onboard sensors.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Red Band</th>
<th>Near-Infrared (NIR) Band</th>
<th>Shortwave-Infrared (SWIR) Band</th>
<th>$\alpha$ Value</th>
<th>R Squared of Soil-Line</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat 8</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>0.74</td>
<td>0.951</td>
</tr>
<tr>
<td>Sentinel-2</td>
<td>4</td>
<td>8</td>
<td>11</td>
<td>0.78</td>
<td>0.949</td>
</tr>
<tr>
<td>SPOT 5</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>0.77</td>
<td>0.952</td>
</tr>
<tr>
<td>Landsat 5</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>0.79</td>
<td>0.950</td>
</tr>
<tr>
<td>Worldview 3</td>
<td>6</td>
<td>8</td>
<td>11</td>
<td>0.80</td>
<td>0.946</td>
</tr>
<tr>
<td>MODIS</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>0.74</td>
<td>0.953</td>
</tr>
</tbody>
</table>

3.5. Sensitivity to Soil Moisture

Soil moisture sensitivity for VI+s is a concern because the introduced SWIR band corresponds to water absorption. Soil spectra with different moisture contents were collected for investigating soil moisture sensitivity for different VIs. The spectra of 112 black soil samples with moisture content ranging from 5% to 45% were measured in a dark room using an ASD FieldSpec Pro FR [55]. As shown in Figure 10, the proposed VI+s are more sensitive to soil moisture than original VIs (except NDVI and NDVI+). Fortunately, the sensitivity to soil moisture is very low, although significant. VI+s variance induced by varying soil moisture is approximately 5.0e–5, which is much smaller than variance from different soil types (0.0009–0.0016). This indicates that soil moisture is not a critical issue for the proposed VI+s.
variance from different soil types (0.0009–0.0016). This indicates that soil moisture is not a critical issue for the proposed VI+s.

Figure 10. The relationship between soil moisture and VIs (n = 112). (a–d) Soil moisture vs. original VIs, and (e–h) soil moisture vs. VI+s.

4. Discussion and Conclusions

In this study, a thinner soil-line was established based on the NIR and red-SWIR bands and modified versions of vegetation indices were proposed by replacing the original red band with the
red-SWIR band. All of the experiments based on simulated and in situ measurement show that VI’s are more effective for FVC estimation and LAI retrieval than the original VIs. Aside from soil color variation suppression, introducing the SWIR band might also alleviate atmospheric effect because of the longer wavelength [56]. Moreover, the proposed VI’s have similar mathematical formulas as commonly used VIs, making them familiar to use. In previous efforts, different VIs were designed to enhance performance in some respects, whereas they usually lose performance in others. For example, EVI has better aerosol effect control than NDVI, but it suffers from much higher sensitivity to topographic variation [57]. These strengths and weaknesses also manifest in the proposed VI’s, thus users will find VI’s familiar to use.

Soil moisture and clay content could cause side effects when using the SWIR band because of their absorptions. In this study, we explored an optimal weighting of the SWIR and red bands using nearly all the world soil spectra, including wet soils and clay-bearing soils. Consequently, sensitivity to soil moisture and clay content could be constrained to a reasonable extent.

In conclusion, we recommend employing the proposed NDVI+, EVI+, SAVI+, and MSAVI+ in applications when soil color variation cannot be neglected, such as large areas, and sparsely vegetated areas or seasons.


Funding: This study was supported by National Key Research and Development Program of China (No. 2017YFD0300201).

Conflicts of Interest: The authors declare no conflict of interest.

References
25. Jurgens, C. The modified normalized difference vegetation index (mNDVI) a new index to determine frost damages in agriculture based on Landsat TM data. *Int. J. Remote Sens.* 1997, 18, 3583–3594. [CrossRef]
57. Matsushita, B.; Yang, W.; Chen, J.; Onda, Y.; Qiu, G. Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: A case study in high-density cypress forest. *Sensors* **2007**, *7*, 2636–2651. [CrossRef] [PubMed]