Spectral Considerations for Modeling Yield of Canola

John J. Sulik and Dan S. Long

USDA-ARS Columbia Plateau Conservation Research Center, PO Box 370, Pendleton, OR 97801, USA,

e-mail: sulikjj@gmail.com (Sulik); dan.long@ars.usda.gov (Long)

Corresponding author (Tel.: +1-541-278-4370, e-mail: sulikjj@gmail.com)

This research was supported by Research Grant Award No. 2012-10008-19727 from the USDA National

Institute of Food and Agriculture.

Abstract

Prominent yellow flowers that are present in a *Brassica* oilseed crop such as canola require careful consideration when selecting a spectral index for yield estimation. This study evaluated spectral indices for multispectral sensors that correlate with the seed yield of *Brassica* oilseed crops. A small-plot experiment was conducted near Pendleton, Oregon in which spring canola was grown under varying water regimes and nitrogen treatments to create a wide range in oilseed yield. Plot measurements consisted of canopy reflectance at flowering using a hand-held spectroradiometer and seed yield at physiological maturity. Spectroradiometric measurements were converted to MODIS band equivalent reflectance. Selected indices were computed from spectra obtained with the radiometer and correlated with seed yield. A normalized difference yellowness index (NDYI), computed from the green and blue wavebands, overcame limitations of the normalized difference vegetation index (NDVI) during flowering and best modeled variability in relative yield potential. NDYI was more linear and correlated with county-wide oilseed yield data and MODIS satellite data from North Dakota ($r^2 \leq 0.72$) than NDVI ($r^2 \leq 0.66$). NDYI only requires wavebands in the visible region of the spectrum and can be applied to any satellite or aerial sensor that has blue and green channels. These findings highlight the benefit of using a spectral index that is sensitive to reproductive growth of vegetation instead of vegetative growth for crops with spectrally prominent reproductive canopy elements. Our results indicate that NDYI is a better indicator of yield potential than NDVI during mid-season development stages, especially peak flowering.

Keywords

Canola; *Brassica* oilseed; Yield estimation; Optical remote sensing

1. Introduction

Remote sensing has provided a means for regional assessment of the yield of economically important crops (Doraiswamy & Cook, 1995; Groten, 1993; Maas, 1988; Tucker & Sellers, 1986) and within-field characterization of spatial differences in crop growth and development (Dalla Marta et al., 2015; Pinter et al., 2003; Yang et al., 2006). The normalized difference vegetation index (NDVI) is the most commonly used spectral index for estimating yield from either a mid-season image or time-series of images, and has long been recognized for its value in monitoring crop conditions and forecasting crop yields (Doraiswamy, 2004; Johnson, 2014; Rouse et al., 1974; Tucker, 1979).

The NDVI is theoretically an effective yield estimator because it correlates well with photosynthetic capacity (Becker-Reshef et al., 2010). NDVI's link to photosynthetic capacity is largely due to the use of a NIR waveband in its formulation. Within a species, NIR reflectance may positively correlate with photosynthetic potential since higher numbers of intercellular air spaces in leaf mesophyll cause more refractions and therefore increased multiple scattering. Specifically, the ratio of mesophyll cell surface area exposed to intercellular air spaces per unit leaf area corresponds to greater surface area for CO₂ diffusion into Rubisco, where carboxylation takes place (carbon fixation) (Slaton et al., 2001). In addition, low reflectance in the red waveband used in NDVI is strongly related to chlorophyll absorption, with greater absorption being proportional to higher chlorophyll content and therefore greater potential for carbon fixation since chlorophyll molecules receive internally diffused CO₂ along with water to synthesize carbohydrates. These carbohydrates are used for respiration, energy storage, and soluble photosynthate that is translocated from the leaf to other organs as plant development progresses.

NDVI is very useful for remote sensing of such crops as wheat and corn that have inconspicuous flowers and simply "green-up" and then "green-down" after entering the reproductive growth stages. However, many *Brassica* oilseeds "green-up" then "yellow-up" with the appearance of conspicuous yellow flowers and have an overlap of "yellow-down and green-down" during maturation. This spectral-temporal variability requires careful consideration when selecting a spectral index to correlate with a physiologically meaningful quantity such as seed yield. Spectral-temporal variability is a function of the

morphological development of canola (*Brassica napus* L.)- an edible oilseed crop that is grown for both food and biofuel in many parts of the world. Canola undergoes three distinct canopy morphologies during development. The first is dominated by leaves, the second is characterized by yellow (less commonly white or beige) flower petals, and the last is mainly pods and branches. Each canopy morphology strongly influences how solar radiation is intercepted (Leach et al., 1989).

During the flowering stage there is reduced interception of photosynthetically active radiation (Yates & Steven, 1987). Yellow flowers contribute red light (yellow = green+red) to a canopy-level signal and this added radiation reduces NDVI values (Behrens et al., 2006; Piekarczyk, 2011; Shen et al., 2009), which adversely affects the performance of NDVI to map variability in biomass (Shen et al., 2010; Sulik & Long, 2015) or yield potential. For instance, if there is a field of flowering radish, turnip, mustard, or other *Brassica* crop- any flower where there is a substantial red light component (i.e., white, red, yellow), then the red channel's value will be a combination of foliar (what NDVI is supposed to index) and floral reflectance.

The floral contribution to the reflectance is manifest most strongly in the green waveband (Sulik & Long, 2015). Yellow canola petal coloration is due to carotenoid absorption (Zhang et al., 2015) at ~450 nm. Canola petals are yellow because carotenoids in the petal absorb blue light and therefore reflect a mixture of green and red light that we perceive as yellow. Flowering canopies reflect more and absorb less between 500 nm and 700 nm (Yates & Steven, 1987), reflect less and absorb slightly more than vegetative canopies between 400 nm and 500 nm (Yates & Steven, 1987), and have little impact on the red edge or NIR (Migdall et al., 2010). Therefore, NDVI does not index solely vegetative growth during flowering.

The fact that NDVI is confounded by spectral variation of the yellow flowers in a canola canopy motivated us to implement a spectral index that is directly sensitive to variation in yellowness. It turns out that the reflection of radiation from flowers can provide us with clues about reproductive potential. The authors found that a ratio of green to blue wavebands positively correlates with the number of flowers per unit area (Sulik & Long, 2015). By directly capturing the spectral variation in yellow flower density, we

are able to acquire information about the relative yield potential further into the growing season than an index that is more suitable for plant development stages that occur before reproductive growth commences. Moreover, final seed yield for canola does not solely hinge on information about photosynthetic capacity or vegetative growth potential because canola yield components are composed of more than merely the result of photosynthesis. Yield per unit area depends on stand density, the number of pods per plant, the number of seeds per pod, and seed weight (Diepenbrock, 2000). Though canola produces roughly twice as many flowers than pods, remote sensing can give us an indication of variability in future pod development by gleaned information about flower density. Flower density reflects accumulated dry matter (ADM) and leaf area index (LAI) at anthesis (Faraji, 2010), and indicates potential for pod development. For example, the number of pods and seeds generated after flowering depends on the supply of assimilates. Stocks of carbon assimilate at the onset of flower initiation, as well as shading factors, can have strong influence on flower, pod and seed number (Tayo & Morgan, 1979). Photosynthetic capacity alludes to the accumulation of carbon and nitrogen that may contribute to reproductive growth; however, those nutrients must also flow to reproductive organs. For canola, available stocks of carbon and nitrogen determine the growth of flowers and pods (Brunel-Muguet et al., 2013).

During flowering, nitrogen is mainly sourced from leaves and sunk into flowers; however, the factors that determine the efficiency of amino acid remobilization are not fully understood and may depend on canopy structural characteristics that affect leaf shading (Brunel-Muguet et al., 2013), among other things. Fortunately, the distribution and intensity of flowers reflect the ability of a canopy to translocate nitrogen and carbon to reproductive growth, which builds upon the vegetative growth accumulated as a function of photosynthetic capacity. As the number of flowers per plant increases, the number of pods increases (Faraji et al., 2008). As the number of pods increases there is more opportunity for seeds to set, barring any high stress impacts at this stage such as heat stress (Morrison, 1993) or a combination of heat and water stress (Angadi et al., 2000) and hence greater seed yield. Therefore, more flowers per unit area should correspond to more pods and therefore more seed yield.

1.1. Canola yield modeling using remote sensing

Prior remote sensing research on canola has demonstrated the importance of timing for image acquisition if estimating relationships with grain yield. It has already been established that flowers in a canola canopy are problematic for relating NDVI to yield (Basnyat et al., 2004; Piekarczyk, 2011). An important temporal finding is that the relationship between NDVI and yield declines as flowering increases (Piekarczyk, 2011); however, no spectral solution to this problem has been offered to-date. Holzapfel et al. (2009) found that NDVI data acquired between the six-leaf stage and the beginning of flowering were correlated to canola seed yield ($R^2 = 0.35$; $p < 0.001$). Optical sensing is conducted at such an early development stage in order to estimate nitrogen requirements for variable rate fertilizer recommendations before flowering and is useful for back-calculating fertilizer requirements in order to help meet yield expectations rather than for actually determining final seed yield.(Holzapfel, 2007; Osborne, 2007).

A forecasting exercise was carried out in order to predict final seed yield employed MODIS 10-day composite NDVI data from late July to early August in Canada (Mkhabela, 2011), which is typically during the flowering and grain filling period. Satellite imagery was combined with canola yield data for Census Agricultural Regions in the Canadian Prairies and resulted in very low correlations for canola, especially in semi-arid environments. However, a crop mask was not applied to the satellite sensor imagery analyzed and so the results cannot strictly be interpreted with respect to the spectral-temporal properties of canola.

Indeed, canola crop yield estimates have been reported to be too low, which may be attributed to the fact that remote sensing indices such as NDVI are intended to sense vegetative growth (Pratt, 2013). For crops like canola there is interest in the seed, which is reproductive growth. Therefore, the most proximate thing to the seed that can be remotely sensed is the flower that precedes the pods that eventually bear seed. The objective of this study was to evaluate the performance of remote sensing indices that are designed to sense reproductive growth in predicting canola yield. Results from a small

plot experiment with spring canola in Oregon are presented to support the theoretical arguments outlined above. To cross-verify our biophysical theory and reinforce the findings from the small-plot study, the practical application of reproductive indices is illustrated with results from remotely sensed prediction of county-wide canola yields in North Dakota. Study sites and sensor systems are disparate with respect to geographic scale as needed to assess robustness of correlation between index values and yield at mid-season.

2. Methods

2.1. Small-plot field experiment

A two-year (2013-2014) experiment was conducted on Walla Walla silt loam soils (coarse-silty, mixed, superactive, mesic Typic Haploxerolls) at the USDA-ARS Columbia Plateau Conservation Research Center near Adams, Oregon, USA (45°43.1' N. Lat.; -118°37.7' W. Lon.). The study was moved to a different site before the beginning of each season. An irrigation system was used to create three distinct water regimes according to the following scheduling scheme. The high water regime applied water during the growing season to maintain soil water content at 20% on a dry-mass basis. The intermediate regime received approximately one half the amount of water that was applied in the high regime. In the low regime, no water was applied and the crop was grown under rain fed conditions.

Within each water regime, a spoke wheel applicator was used to apply liquid fertilizer (urea ammonium nitrate) and create small plots of three nitrogen (N) levels in a randomized complete block design. Rates of applied N were adjusted for each water regime and were equivalent to 0, 45, and 90 kg ha⁻¹ in the low regime, 0, 90, and 179 kg ha⁻¹ in the intermediate regime, and 0, 134, and 269 kg ha⁻¹ in the high regime.

Planting of spring canola (cv. HyCLASS 940 Roundup Ready) occurred on 4 April and 8 April in 2013 and 2014. In 2013, planting was accomplished with a 3-m wide no-till plot drill (Seed Hawk, Inc., Langbank, SK, Canada) equipped with single sideband, seed knife openers. Row spacing and seeding

168 rates were 25 cm and 118 pure live seeds m^{-2} . Individual plots were approximately 3 m wide and 23 m
 169 long (Fig. 1a). In 2014, a 1.83 m wide no-till plot drill (Great Plains Ag, model 606NT, Salina, KS,
 170 USA) was used to plant canola into 3.7×11 m plots at 124 pure live seeds m^{-1} on 18 cm row spacing (Fig.
 171 1b).

(a)

(b)

172

Figure 1. Plot layout of 2013 field experiment (a) and 2014 field experiment (b) where spring canola was grown under varying water and nitrogen levels. Green dots in 2013 denote subplots where measurements of canopy spectral reflectance and crop yield were obtained.

In both years, irrigation was applied with 16 mm dia. T-Tape (John Deere and Company, Moline, IL, USA) drip tape with 20 cm hole spacing that effectively delivers a flow rate of 12.7 mm hr⁻¹ at 55-60 kPa. Soil cores were sampled with a hand probe to the 30 cm depth from within each water regime. Gravimetric water content was determined by weighing soil core before and after oven drying at 40°C for 24 h. Broadleaf weeds were controlled by application of appropriate herbicides at the rosette stage of crop growth. Insecticides were applied during flowering to control infestations of aphid.

2.1.1. Sampling 2013 plots

An ASD FieldSpecProTM Radiometer (PANalytical, Boulder, CO, USA) was used to obtain canopy spectral reflectance measurements on 7 June (Fig. 2a), 18 June (Fig. 2b), 27 June (Fig. 2c), and 2 July (Fig. 2d) under clear or mostly clear sky conditions. Spectral sampling dates were chosen to capture varying levels of flowering intensity. Spectral measurements were obtained within three 0.5 m² quadrats within each of the 30 plots (green dots in Fig. 1a), which brought to 90 the total number of subplots. Spectral reflectance was calculated as the mean of 40 scans per sample reading. The fiber optic probe has a 25° field of view (FOV) and was held approximately 1.25 m at nadir so the instantaneous FOV was about 0.5 m². Spectral measurements at each plot location were preceded by a dark current measurement and a white reference measurement of a SpectralonTM panel (Labsphere Inc., North Sutton, New Hampshire, USA) to convert radiance values into reflectance factors. Spectroradiometric measurements were converted to Band Equivalent Reflectance (BER) values (Trigg & Flasse, 2000) of the spectral response functions of the MODIS satellite sensor. At physiological maturity, each of the 30 plots was hand sampled for seed yield within the three sub-plots. Final yield samples for six plots were missing and

one other plot had a bad radiometer reading so these sample points were omitted, leaving a total of 83 data points for analysis.

Figure 2. Appearance of canola with respect to flowering on four spectral sampling dates in 2013.

2.1.2. Sampling 2014 plots

In 2014, canopy spectral reflectance measurements were obtained on different dates (Fig. 3) by means of digital aerial imaging. A Tetracam Mini-MCA6 multispectral imager mounted to a multi-rotor, unmanned aerial vehicle was used to record imagery on 30 May (Fig. 3a), 5 June Fig. (3b), 9 June (3c), 12 June (Fig. 3d), 19 June (Fig. 3e), 28 June, 1 July (Fig. 3f), 9 July, and 17 July. Flights were within three hours of solar noon under clear or mostly clear sky conditions. The camera wavebands are 10 nm full width half max (FWHM) and are centered at 450 nm, 551 nm, 673 nm, 703 nm, 752 nm, and 812 nm. Each image was captured at approximately 120 m above ground level, which resulted in a ground sampling distance of approximately 6.5 cm. Each Tetracam Mini-MCA imager is independent and has a focal length of 8.5 mm and a 1.3 megapixel ($1,280 \times 1,024$ pixels) CMOS sensor arranged in a 2×3

array. An alignment file was used with Tetracam's Pixel Wrench software to align the six bands. No mosaicing was necessary since all 36 plots were within the camera's field of view at 120 m.

Figure 3. Appearance of canola with respect to flowering on six spectral sampling dates in 2014.

Unfortunately, the 812 nm waveband channel saturated at intermediate values and the 450 nm waveband saturated at very high values on every image capture date. Therefore, empirical line calibration was not feasible for radiometric calibration of the image channels. Fortunately, Tracor Aerospace canvas calibration tarps were in the field of view of the Tetracam during each flight. A black calibration tarp had consistently low enough pixel values in each image so it was used to flat field correct (FFC) the imagery (Carrere & Abrams, 1988). Since the tarp is spectrally flat, dividing an average of the tarp's pixel values into each waveband effectively normalized the imagery and adjusted the shape of the spectral reflectance curve of in-scene features by removing the solar curve from digital numbers band-by-band. Although the imagery was not calibrated to reflectance factors and are not suited for comparison to reference spectra,

the normalized digital numbers accurately map within-field spatial variability in the relative intensity of reflected light.

The 450 nm waveband only saturated over bright features so all of the pixel values over vegetation were fine. However, the 812 nm waveband saturated over dense areas of vegetation so the 752 nm waveband was used for the NIR term in the NDVI calculations. Spectroradiometric measurements of the canola canopy indicate that 752 nm is on the cusp of the NIR plateau and the rate of change of the red edge slows down considerably around 750 nm, after which the NIR plateau stabilizes.

Each of the 36 plots was sampled for seed yield using a small plot combine after the crop reached physiological maturity in August. Two short edges of the ends of each plot were trimmed away. The combine harvested the remaining core of each plot and excluded two rows along the outside length to avoid edge effects. The dimensions of each plot were measured after trimming to normalize the seed weight by harvested area. For every image capture date, a spectral index was calculated and then all pixels in the core area of each plot, within the harvested areas, were averaged for use in regression analysis (described in section 3.).

2.2. North Dakota 2012 study area

Canola is an economically important crop in North Dakota where it is grown in rotation with more predominant crops such as wheat, soybeans, and corn. Canola is grown in all 53 counties of the state; however, it is predominantly grown in the northern half of the state. Canola cultivated in North Dakota accounts for the majority of this crop in the U.S. It is sown in the spring and typically reaches anthesis by mid-July. Average annual precipitation varies longitudinally across the state from 350 mm in the west to 550 mm in the east. Consequently, canola yields are nearly 100% greater in the sub-humid, eastern portion than in the semi-arid, western portion (Fig. 4). In 2012, total canola acreage in North Dakota was approximately 1.43 M acres versus 7.58 M acres planted to wheat (USDA Farm Service,

2013). A typical canola field in North Dakota is about 0.65 square km (160 acres) although smaller fields are common.

Figure 4. Canola yields aggregated to the county-level in North Dakota.

2.3. Sources of Data for Satellite Remote Sensing

2.3.1. Yield data

North Dakota county-level statistics on canola yield, area harvested, and production were obtained from the USDA National Agricultural Statistics Service (NASS) Quick Stats database (NASS, 2013b). NASS is the agency responsible for administering the USDA's program for collecting and publishing agricultural statistics at the national, state and county levels. Yield data are based on county totals. The NASS crop statistics are based on data obtained from multiple frame-based sample surveys of farm operators, objective yield surveys, agribusinesses, shippers, processors and commercial storage firms.

The NASS statistics data are inconsistent across years, and some years, including 2012, aggregate several counties and are also missing some data. For 2012 there are 25 counties with information at the county level and 6 counties that fall under the combined counties category (total number of ND counties is 53). In 2011, there is yield data for 31 individual counties and 3 in the combined category. Yield data is significantly reduced for 2013 and 2014 so these years were excluded. There are other years of data available; however, the focus of this study is on experimental data so we only analyzed yield data for 2012, 2011, and 2010.

2.3.2. MODIS time-series data

There are several different MODIS products to choose from. Daily-level data would be ideal but North Dakota typically has extensive cloud cover in June and July. Therefore, we used the MOD09A1 Version-5 MODIS product, downloaded from the United States Geological Survey Global Visualization Viewer, because it is a temporal composite that uses the best observations (i.e. clouds, view angle, etc.) over an 8-day period (Holben, 1986). Most fields in North Dakota are laid out on the US Land Grid system of 640 acre square sections and each section is divided into four quarter sections. A typical canola field is cultivated as a quarter section, which is about 650 m on each side. In comparison, MOD09A1 pixels are 500 m wide whereas a typical field in North Dakota is about 650 m wide so a 250 m MODIS product would be more suitable; however, no MOD09A1 products are available below 500 m spatial resolution.

2.4. Spectral indices

We chose to evaluate three spectral indices for canola yield estimation. Indices were selected so as to include wavebands that should be available on many multispectral imaging systems including those that are typically mounted on unmanned aircraft and employ silicon-based detectors that have an upper limit of sensitivity in the NIR.

NDVI is fundamentally a contrast between blue light and green light (Eq. 1) whereas NDVI is a contrast between NIR and red light (Eq. 2), and Blue NDVI (BNDVI) (Eq. 3) is similar to NDVI except its denominator incorporates a blue waveband instead of red. The rationale for analyzing BNDVI is that chlorophyll absorbs blue light and is not affected by reflected light from yellow flowers so it may be a good alternative to NDVI. BNDVI is sometimes referred to as “pseudo-NDVI” but we disagree with this terminology since there is nothing false about it. BNDVI is simply less commonly employed than NDVI, largely due to atmospheric considerations, but this in no way implies that its relationship to any measured quantity is more spurious than other indices.

$$\text{NDVI} = (\text{Green} - \text{Blue}) / (\text{Green} + \text{Blue}) \quad \text{Eq. 1}$$

$$\text{NDVI} = (\text{NIR} - \text{Red}) / (\text{NIR} + \text{Red}) \quad \text{Eq. 2}$$

$$\text{BNDVI} = (\text{NIR} - \text{Blue}) / (\text{NIR} + \text{Blue}) \quad \text{Eq. 3}$$

2.5 Crop mask

Spectral index values were masked (Genovese et al., 2001) as needed to process only the appropriate pixels. A North Dakota canola map was generated from a NASS product called the Cropland Data Layer (CDL). The CDL is a land cover map of specific crop types that is developed from ground surveys, farmer reports, and remotely sensed data from Landsat products (NASS, 2013a). The 2012 CDL raster layer for North Dakota is distributed at 30 m spatial resolution. Therefore, we used majority resampling to convert this raster layer into a binary map of canola at 500 m resolution in order to match the spatial resolution of MODIS data. Thus, spectral data are based on patches within counties after the masking process so that only pixels associated with canola production pixels are related to the yield values for each county.

Operational canola yield forecasts at regional scales will require a canola map to mask out other surface features and crop types. The manner in which a crop layer is generated often varies by crop and may be subject to multiple iterations of refinement (Sakamoto et al., 2014). Towards this end, an

automated method for *Brassica* oilseed identification using medium-resolution satellite imagery is under development.

3. Modeling

3.1. Simple linear regression modeling

Robust yield models typically include agro-meteorological variables in addition to a spectral index such as NDVI (Johnson, 2014). However, a spectral index is the most arbitrary variable in such a model because indices lack physical dimension. Accordingly, the authors chose a bivariate analysis in order to assess the best dimensionless proxy for yield.

With respect to a process model, we believe that a measurement of yield potential at mid-season should be linearly related to final yield because an index will ideally have a linear relationship to what it is measuring. The model slopes tell us how much yield changes as a linear function of spectral index values. Often, the curvature found between an index and a biophysical variable is not due to a non-linear biological or physical process but is due to a mathematical procedure such as the normalized differencing of two wavebands (Gitelson, 2004). Therefore, a spectral index that has a linear relationship with a biophysical variable such as yield may employ a statistical model that is congruent with the process model. Our intent is to analyze and theoretically develop an alternative to NDVI, not to propose an operational yield model. NDVI is the status quo and is treated as a benchmark for performance. We hope this will provide a basis for future research to combine NDVI with agro-meteorological variables.

3.1.1. Regression models for 2013 small-plot experiment

For the 2013 field experiment we fit a separate regression model for each spectral sampling date and spectral index. An additional regression model using the maximum spectral index value of each sampling plot across all sampling dates was included to compensate for geographic variation in development stage and allow NDVI to be evaluated before peak flowering (which deflates NDVI) and

NDVI to be evaluated during peak flowering, regardless of the development stage of a sub-plot during a spectral sampling date.

3.1.2. Regression models for 2014 small-plot experiment

For the 2014 field experiment we fit a separate regression model for each image capture and each spectral index. An additional regression model using the maximum spectral index value of each plot across all sampling dates was included to compensate for between-plot differences in development stage. This allows NDVI to be evaluated before peak flowering and NDVI to be evaluated during peak flowering, regardless of between-plot differences in development stage.

For comparison, all 2014 regression analyses were run using multispectral imagery with pixels processed by FFC as well as digital numbers that represent at-sensor scaled radiance and the results did not differ at a meaningful level of precision. This is likely because the imagery was not subject to univariate analyses across time; rather, the spectral data for each image capture date were related to yield in a bivariate context and so only the covariance between each image date's spectra and yield are important for this study. Calibration to reflectance would have shifted the origin of each image capture's pixel values and enabled univariate and/or multivariate temporal analysis.

For each image, a threshold value for NDVI was interactively determined by viewing the NDVI histogram and used to create a canopy segmentation mask. The NDVI-derived canopy mask was applied to all spectral index rasters in order to restrict the analysis to vegetation pixels. Finally, all vegetation pixels were averaged and these values were related to the total seed yield for each of the 36 plots.

3.1.3. Regression models for 2012 North Dakota counties

For the 2012 MODIS data we fit a separate regression model for each 8-day period and each spectral index. An additional regression model using the maximum spectral index value of each pixel across all sampling dates was included to compensate for geographic variation in development stage. This allows NDVI to be evaluated before peak flowering and NDVI to be evaluated during peak flowering,

regardless of the average development stage within a county during the dates that comprise each 8-day composite image. However, if there were several fields within a county that reached anthesis before 25 June, which corresponds to Julian day 177 for 2012, then this will adversely affect the performance of NDVI.

NDVI, (BNDVI), and NDYI were calculated from the MODIS surface reflectance product after band-level processing based on the bit encoding for MOD09A1 QA/QC state flags and quality assessment data (Roy et al., 2002). Specifically, a custom R-script was written to generate logical masks based on band quality to only include land surface cover, without cloud cover or high aerosol content. All R code used for the North Dakota MODIS analysis is available from the corresponding author upon request.

3.1.4. Regression models for 2010 and 2011 North Dakota counties

For the 2010 and 2011 MODIS data we fit a separate regression model for each 8-day period and each spectral index. An additional regression model using the maximum spectral index value of each pixel across all sampling dates was included to compensate for geographic variation in development stage. The data was processed in the same manner as the 2012 MODIS data.

4. Results

For our experimental data, dividing the study site into three distinct water environments and five nitrogen levels was effective in creating a wide range in total seed yield. Regression modeling utilized data collected during mid-season crop development stages, immediately before anthesis to the beginning of pod fill, to demonstrate the spectral-temporal dynamics of relationships between spectral indices and total canola yield.

The North Dakota data also had a wide range in spectral index values and total seed yield but the results are not as consistent. This may be due to inconsistencies in yield reporting across years, varying accuracy of the CDL crop mask, and different environmental conditions across tabulated counties for each year.

4.1. Regression analysis of 2013 data

Scatter plots of spectral index values versus total yield show more scattering in the high yielding plots across all dates (Figs. 5 and 6). This may be due to racemes branching out very far and eventually extending beyond the predetermined hand-harvest area. In addition, manual handling of these racemes may have resulted in varying degrees of pod shatter. BNDVI maintains a consistent relationship with yield regardless of sampling date (Fig. 5). The r^2 value is slightly reduced during peak flowering ($r^2=0.54$). Like NDVI, the scatter plot has a curvilinear shape and most of the high watering sample points are above the regression line, especially by pod fill, and so the models will under-predict yield values. As expected, NDVI is most strongly affected by peak flowering ($r^2=0.22$). However, when flowers are not present its performance is almost as good as BNDVI. During peak flower the BNDVI and NDVI regression lines effectively partition the sample points of the high water regime from those of the low and intermediate water regimes. Yields in the low water regime had consistently low yields but a wide range of BNDVI and NDVI values. The number of exceptionally low index values in the low water regime affects the r^2 of BNDVI and NDVI more than it does for the NDYI, which may be partly attributable to differences in dynamic range.

NDYI was not adversely affected by flowers. In addition, NDYI had the strongest relationship with yield during peak flowering ($r^2=0.66$). The regression lines did not partition the sample points of the high water regime from points of the other regimes nor did the scatter plots exhibited the most linear relationships with yield regardless of the normalized differencing. As an aside, a simple ratio vegetation index (NIR/R) regression was run and there was no curvature in the scatter plot, but these models displayed the same pattern in r^2 values. This reinforced our notion that the curvature in the NDVI and BNDVI scatter plots is an artifact of the normalized difference, which is a non-linear transformation of data. NDYI may not demonstrate the same pronounced curvature because it has a narrower dynamic range.

Figure 5. Regression models for each Julian sampling date. Refer to Figure 2 in order to visualize the impact of flowering magnitude

Using the maximum index value across sampling dates (Fig. 6) resulted in similar patterns to the models that varied by sampling date (Fig. 5). In terms of absolute value, BNDVI had the largest slope, but the strength of the relationship is intermediate between NDYI and NDVI. BNDVI and NDVI both exhibit very large slope coefficients and this may be due to the large curvature in their relationships with yield. Again, NDYI has the most linear and strongest relationship with yield ($r^2=0.67$).

Figure 6. Regression models for each index using the maximum index value recorded for each subplot during the sampling period.

4.2. Regression analysis of 2014 data

A separate linear model was fit for each image capture date. The same processing was performed on the FFC imagery as well as the raw digital numbers and the differences were negligible (data not shown). Though image pixel values were not calibrated to reflectance, the FFC normalization recovered the expected shape of the spectral reflectance curves for the vegetation and soil within the imagery. For both the 2013 field experiment and North Dakota county data, NDYI has a linear relationship with canola yield whereas NDVI has a pronounced curvilinear relationship and overall lower r^2 values (Fig. 7). Scatter plots of the spectral index values versus total yield along with regression lines and coefficients of variation are presented in Figures 5-11. All regression coefficients are significant at $p < 0.01$.

Figure 7. Regression models for each Julian sampling date. Refer to Figure 3 in order to visualize the impact of flowering magnitude

Julian day 143 corresponds to 22 May and the canola was still in rosette at this point. Figure 3a indicates that stand establishment was set, but the canopy still had not closed and no spectral index correlated well with yield. NDYI performed the worst ($r^2 < 0.1$) because no flowers were present yet. Stem elongation initiated by Julian day 151 (Fig. 3b) but upper leaves were still undeveloped and no index correlated well with yield, indicating that a later-season image is required. There was enough canopy development by Julian day 157 (Fig. 3c) so that each index finally explained at least half of the variability in yield (Fig. 7). The leaves were more fully developed at 157 than 151 and so NDVI's performance improved; however, there were enough flowers present so that it did not perform as well as NDYI, especially in the low water regime where bloom was most intense at this point.

Bloom mostly intensified in the low and intermediate water regimes by Julian 161, at which point NDVI's r^2 value dropped by 23 percentage points and BNDVI's r^2 increased by 20 points. Since flowering was more uniform throughout all water regimes, NDYI's r^2 rose 15 percentage points to 0.78, which is the highest r^2 of any index but not by much. NDVI's r^2 value dropped to 0.18 on 12 June, the lowest at any time point. During this same period from Julian days 161-164, BNDVI and NDYI both maintained very strong relationships with yield. Based on the scatterplots, BNDVI appeared to do better explaining high yield values whereas NDYI did better with low yield values.

Bloom in the high water regime lagged behind until Julian day 171 (Fig. 3d), at which point NDYI better explained variability in the high water regime (Fig. 7). In addition, BNDVI performed worse on this date, especially for the low water regime. Canopy development shifted from a predominance of bloom to a greater presence of pod-filling (canola development stages can have considerable overlap) by Julian day 183 (Fig. 3f), after which NDYI performs worse than the other indices. From this point on NDVI performs better than the other indices because most flowers have shed and the red waveband is progressively less responsive to carotenoid reflectance and more responsive to chlorophyll absorption.

The increase in r^2 values for all indices on Julian days 183 and 191 (Figs. 7 and 8) can be explained by weather events. The timing of the rain on Julian days 178-179 (Fig. 8) stimulated more foliar and floral production. Heavy rainfall (> 10 mm) on 26-27 June likely initiated new branches and

flower clusters, effectively prolonging flowering to over four weeks and reinvigorating the remaining leaves (Fig. 3).

Figure 8. Temporal pattern of the strength of relationship between each index and total seed yield (top) with total rainfall for each day during the study period. Note the spike in rainfall from Julian days 177-179 and subsequent increase in strength of relationship for each index.

Based on the coefficient of variation, NDVI gave the poorest performance in line with theoretical expectations (Fig. 9). However, the scatterplots indicate that NDVI had the tightest fit for the intermediate water regime (triangles). BNDVI performed better than NDVI; however, NDYI clearly relates better to the yield values, especially for low and high yielding plots. There are plots with yields that are greater than two standard deviations from the mean, which are both high water and high nitrogen treated plots, and no index appears to explain their variability well at any time. These values are also well above the county-level yield rates for 2012 which suggests the yields for these plots were subject to measurement error.

Figure 9. Regression models for each index using the maximum index value recorded for each plot during the sampling period.

4.3. Regression analysis of 2012 North Dakota counties

Simple linear regression was run for each 8-day interval to determine the influence of timing on mid-season yield estimates. For each 8-day composite, NDVI and BNDVI had comparable performance (Fig. 10). Each index appeared to underestimate very high yield values on Julian Days 185 and 201. Based on the r^2 values, 4 August (Julian Day 217) corresponds the weakest relationship for each index and NDVI has the most linear scatter and strongest relationship ($r^2=0.41$) for this 8-day period. The model slopes appear to be more stable across time than for our small plot experiment and there is not as much curvature in the scatter for BNDVI or NDVI. NDYI consistently performed the best except after Julian Day 209 (end of July/early August), which is late in flowering and early in pod fill. NDVI appears to perform well after pod fill commences and most petals are shed, probably because light is able to reflect from basal leaves since higher-positioned canopy components have thinned out.

Figure 10. Regression models for each MODIS 8-day composite image product.

To account for temporal heterogeneity in crop development due to geography and management factors, separate models were based on the maximum value of each spectral index (Fig. 11), which resulted in a more linear point scatter for BNDVI and NDVI than exhibited by the models in Figure 10. There still appears to be more curvature than for NDVI, especially at extremely low values. In addition, NDVI has the most linear and strongest relationship ($r^2=0.76$).

Figure 11. Regression models for each index using the maximum index value recorded for each county across all 8-day periods.

4.4. Regression analysis of 2010 and 2011 North Dakota counties

For both years, individual simple linear regression models were run for each 8-day interval in order to determine the influence of timing on mid-season yield estimates. However, the 8-day results do not suggest that any spectral index is useful for modeling canola yield. Therefore, we only report the results for maximum index value regressions, which are provided Table 1.

For 2011, all spectral indices had comparable performance (see Table 1). Each spectral index was weakly related to total seed yield, especially for high yield values. For each index, the variance of residuals increased with higher values of yield. For 2010, the coefficient of variation was low for all indices. BNDVI and NDVI had comparable performance whereas maximum NDVI values resulted in the most homogeneous scatter around the regression line.

Table 1. Coefficient of determination values for the maximum value of each spectral index for 2011 and 2010.

Year	BNDVI	NDVI	NDYI
2011	0.52	0.5	0.47
2010	0.33	0.31	0.47

5. Discussion

The same general patterns are found in the North Dakota data and our small plot experiment thus suggesting that NDYI is relatively better than NDVI for modeling canola seed yield. NDVI has been shown to work best for estimating wheat yield either at maximum green leaf biomass (Tucker, et al., 1980) or maximum green canopy cover (Mahey et al., 1993). Similar to how NDVI performs during maximum greenness, NDYI performs best during maximum yellow flower cover. With respect to timing, late flowering branches typically produce fewer flowers and pods. Canola will flower from the bottom to the top of the plant and the pods on the bottom are typically bigger than those on top. In addition, ability to compensate for flower and pod losses diminishes as flowering progresses. Therefore, peak flower may be the most informative time to acquire mid-season imagery for yield modeling.

For the North Dakota data, BNDVI likely outperformed NDYI for Julian Day 209 because NDYI exploits variability in flowering and flowering was much less pronounced, or finished, by this time. Pre-anthesis leaf contribution to dry matter assimilation diminishes as pods develop, but pod and branch photosynthesis provides assimilate for seed development (Bilbrough, 1983) so BNDVI and NDVI may be useful tools once flowers are shed.

Even though NDVI does not perform well during flowering, it and BNDVI both do well before and after flowering. This may be due to the fact that NDVI is an indicator of photosynthetic activity, which is decreased during flowering. LAI decreases at the start of flowering and pod area index (PAI) does not increase appreciably until after flowering ends (Diepenbrock, 2000). Therefore, NDVI is able to index photosynthetic capacity again once pods, which are photosynthetically active, develop and basal leaves become more visible after flower abscission.

For both study sites, our results demonstrate that the NDVI-yield relationship weakens as bloom intensifies in agreement with Piekarczyk (2011). The Oregon and North Dakota study sites show the same patterns/trends regardless of growing environment and sensor resolution. Although the study sites are not directly related, our results are cohesive because NDYI performs better during peak flower for both data sets and exploits the same variability that confounds NDVI. However, the blue waveband is sensitive to atmospheric contamination and adequate processing is required for its use.

NDYI may also be used to monitor the duration and intensity of flowering, which would allow yield estimates to incorporate critical heat thresholds and other events that can reduce yield potential during flowering (Morrison, 1993; Morrison, 2002; Malcolm et al., 2002). After all, seed yield depends on the survival of flowers and pods, in addition to seed weight and the number of seeds per pod (Diepenbrock, 2000), and heat stress can prematurely end flowering and reduce seed set (Johnston et al., 2002). Further, low water and high temperature stress conditions during flowering and seed filling can result in poor pod formation and seed yield because number of pods is determined during the flowering period (Abolfazl Faraji et al., 2008).

Although our main emphasis has been on the flowering period, the other development stages are still extremely important to monitor with remote sensing. For example, NDVI time series can be used along with ancillary data from a weather station to determine if high temperatures during vegetative development cause a reduction in flower number. Reduction in flower number due to high heat during vegetative development was demonstrated for all *Brassica* species (*B. napus*, *B. rapa* L., and *B. juncea*) (Malcolm et al., 2002). NDYI mapping and analysis could also be used in applications of precision agriculture such as delineating management zones within fields, informing marketing decisions, and locating weed and disease infestations. Further research will involve specification of a rule-based classifier to develop an in-season canola map instead of relying on government distributed data to mask out canola pixels.

Remote sensing data alone were insufficient to explain variability in county-level yield within North Dakota. An identical analysis was attempted for 2011 and 2010; however, the relationships were

not significant and the r^2 values were low for all indices. Since the counties with published yield values vary by year, it is possible that the 2011 and 2010 counties had greater soil and weather differences than in 2012. For example, North Dakota had a cold, wet spring that delayed seeding in 2011. More robust models that incorporate soils and weather information would be necessary to explain the variability in yield across space and time. Therefore, the 2012 analysis is only meant to reinforce our experimental findings, which are spectrally explanatory and would need ancillary data to be geographically transferable across different soils and climates.

6. Conclusions

This study used empirical remotely sensed yield models to evaluate the performance of three different spectral indices during mid-season development stages. The same patterns were evident at the small-plot and county levels. NDVI is designed to be sensitive to a signal coming from vegetative growth (leaves) and is reduced by the spectral signal that comes from the reproductive growth (flowers) in the canopy. The regression-based models in this study tested the hypothesis that an index sensitive to the spectral characteristics of a canola canopy's reproductive vigor, as opposed to vegetation vigor, is a more linear mid-season indicator of total seed yield than NDVI. We used remotely sensed data collected at different spatial extents, different sensors, and different geographies to demonstrate that NDVI does not perform as well as BNDVI or NDYI.

Our aim was not to create an operational yield model; rather, to demonstrate consistent patterns in how the chosen spectral indices perform at different development stages, regardless of spatial resolution. In addition, our findings suggest that it is not necessary to rely on a remotely sensed proxy for vegetative photosynthetic capacity, such as NDVI, and we provide both theoretical reasoning and empirical results that demonstrate how the spectrally prominent reproductive growth (yellow flowers) of canola compromise the ability of NDVI as an effective estimator of total seed yield.

Furthermore, the goal of this research is not simply to propose another vegetation index. It is important to think critically about how indices are used as covariates of yield, especially when canopy

spectral properties are as dynamic as we illustrated. Many researchers have used ratios of green and blue light for different purposes and our study has furthered understanding of how this band combination, as well as more frequently used band combinations, behave as a function of canola development and yield. Future research should focus on understanding how these indices change as a function of time as well as how to incorporate the implications of our research into more sophisticated process-based and/or agrometeorological models.

This research was completed as part of an effort to provide regionalized strategies to integrate sustainable oil seed production into existing land uses in ways that increase farm profitability and rural economic opportunities, while providing biofuel refiners dependable supplies of high quality feedstock. This requires strategies for using space- and ground-based sensors to predict oilseed quality and yield, and apply information to harvest and crushing operations to improve consistency of feedstock quality and extraction efficiency.

Acknowledgements

This research was supported by Research Grant Award No. 2012-10008-19727 from the USDA National Institute of Food and Agriculture. Grateful appreciation to John McCallum, Steve Umbarger, Aron Boettcher, and Mandy Wuest for assistance in the field and laboratory. Use of trade names does not constitute official endorsement by the USDA.

References

- Angadi, S. V., Cutforth, H. W., Miller, P. R., McConkey, B. G., Entz, M. H., Brandt, S. A., & Volkmar, K. M. (2000). Response of three Brassica species to high temperature stress during reproductive growth. *Canadian Journal of Plant Science*, 80(4), 693–701. <http://doi.org/10.4141/P99-152>
- Basnyat, P., McConkey, B., Lafond, G. P., Moulin, A., & Pelcat, Y. (2004). Optimal time for remote sensing to relate to crop grain yield on the Canadian prairies. *Canadian Journal of Plant Science*, 84(1), 97–103. <http://doi.org/10.4141/P03-070>
- Becker-Reshef, I., Vermote, E., Lindeman, M., & Justice, C. (2010). A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sensing of Environment*, 114(6), 1312–1323. <http://doi.org/10.1016/j.rse.2010.01.010>
- Behrens, T., Müller, J., & Diepenbrock, W. (2006). Utilization of canopy reflectance to predict properties of oilseed rape (*Brassica napus* L.) and barley (*Hordeum vulgare* L.) during ontogenesis. *European Journal of Agronomy*, 25(4), 345–355. <http://doi.org/10.1016/j.eja.2006.06.010>
- Bilsborrow. (1983). A Consideration of Factors Affecting the Yield of Oilseed Rape. *Aspects of Applied Biology*, 6, 91–99.
- Brunel-Muguet, S., Beauclair, P., Bataillé, M.-P., Avice, J.-C., Trouverie, J., Etienne, P., & Ourry, A. (2013). Light Restriction Delays Leaf Senescence in Winter Oilseed Rape (*Brassica napus* L.). *Journal of Plant Growth Regulation*, 32(3), 506–518. <http://doi.org/10.1007/s00344-013-9317-9>
- Carrere, V. & Abrams, M.J. (1988), An assessment of AVIRIS data for hydrothermal alteration mapping in the Goldfield mining district, Nevada. In Proceedings of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Performance Evaluation Workshop (G. Vane, Ed.), JPL Publ. 88-38, Jet Propulsion Laboratory, Pasadena, CA, 134-154.
- Dalla Marta, A., D. Grifoni, M. Mancini, F. Orlando, F. Guasconi, S. O. (2015). Durum wheat in-field monitoring and early-yield prediction: assessment of potential use of high resolution satellite imagery in a hilly area of Tuscany, Central Italy. *The Journal of Agricultural Science*, 153(01), 68–77.
- Diepenbrock, W. (2000). Yield analysis of winter oilseed rape (*Brassica napus* L .): a review, 67(March).
- Doraiswamy, P. (2004). Crop condition and yield simulations using Landsat and MODIS. *Remote Sensing of Environment*, 92(4), 548–559. <http://doi.org/10.1016/j.rse.2004.05.017>
- Doraiswamy, P. C., & Cook, P. W. (1995). Spring Wheat Yield Assessment Using NOAA AVHRR Data. *Canadian Journal of Remote Sensing*, 21(1), 43–51. <http://doi.org/10.1080/07038992.1995.10874595>
- Faraji, A. (2010). Flower formation and pod/flower ratio in canola (*brassica napus* l.) affected by assimilates supply around flowering. *International Journal of Plant Production*, 4(4), 271–280. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-77957072268&partnerID=tZOtx3y1>

659 Faraji, A., Latifi, N., Soltani, A., Hossain, A., Rad, S., Resources, N., ... Kavous, G. (2008). Effect of
660 high temperature stress and supplemental irrigation on flower and pod formation in two canola
661 (*Brassica napus* L.) cultivars at Mediterranean climate. *Asian Journal of Plant Sciences*, 7(4), 343–
662 351.

663 Genovese, G., Vignolles, C., Negre, T., & Passera, G. (2001). A methodology for a combined use of
664 normalised difference vegetation index and CORINE land cover data for crop yield monitoring and
665 forecasting. A case study on Spain. *Agronomie*, 21(1), 91–111. <http://doi.org/10.1051/agro:2001111>

666 Gitelson, A. (2004). Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical
667 Characteristics of Vegetation. *Journal of Plant Physiology*, 161(2), 2004, 165-173.
668 <http://dx.doi.org/10.1078/0176-1617-01176>.

669 Groten, S. M. E. (1993). NDVI—crop monitoring and early yield assessment of Burkina Faso.
670 *International Journal of Remote Sensing*, 14(8), 1495–1515.
671 <http://doi.org/10.1080/01431169308953983>

672 Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data.
673 *International Journal of Remote Sensing*, 7(11), 1417–1434.
674 <http://doi.org/10.1080/01431168608948945>

675 Holzapfel, C. B., Lafond, G. P., Brandt, S. A., May, W. E., & Johnston, A. M. (2007). In-soil banded
676 versus post-seeding liquid nitrogen applications in no-till spring wheat and canola, (2002).

677 Johnson, D. M. (2014). An assessment of pre- and within-season remotely sensed variables for
678 forecasting corn and soybean yields in the United States. *Remote Sensing of Environment*, 141, 116–
679 128. <http://doi.org/10.1016/j.rse.2013.10.027>

680 Johnston, A. M., Tanaka, D. L., Miller, P. R., Brandt, S. A., Nielsen, D. C., Lafond, G. P., & Riveland, N.
681 R. (2002). Oilseed crops for semiarid cropping systems in the Northern Great Plains. *Agronomy*
682 *Journal*, 94(2), 231–240. <http://doi.org/10.2134/agronj2002.2310>

683 Leach, J. E., Milford, G. F. J., Mullen, L. A., Scott, T., & Stevenson, H. J. (1989). Accumulation of Dry
684 Matter in Oilseed Rape. *Aspects of Applied Biology*, 23, 117–123.

685 Maas, S. J. (1988). Using Satellite Data to Improve Model Estimates of Crop Yield. *Agronomy Journal*,
686 80, 655–662. <http://doi.org/10.2134/agronj1988.00021962008000040021x>

687 Mahey, R. K., Singh, R., Sidhu, S. S., Narang, R. S., Dadhwal, V. K., Parihar, J. S., & Sharma, A. K.
688 (1993). Pre-harvest state level wheat acreage estimation using IRS-IA LISS-I data in Punjab (India).
689 *International Journal of Remote Sensing*, 14(6), 1099–1106.
690 <http://doi.org/10.1080/01431169308904398>

691 Migdall, S., Ohl, N., & Bach, H. (2010). Parameterisation of the land surface reflectance model SLC for
692 winter rape using spaceborne hyperspectral CHRIS data. In *Hyperspectral 2010 Workshop* (Vol.
693 ESA SP-683). Frascati, Italy.

694 Mkhabela, M. S., Bullock, P., Raj, S., Wang, S., & Yang, Y. (2011). Crop yield forecasting on the
695 Canadian Prairies using MODIS NDVI data. *Agricultural and Forest Meteorology*, 151(3), 385–
696 393. <http://doi.org/10.1016/j.agrformet.2010.11.012>

697 Morrison, M. J. (1993). Heat stress during reproduction in summer rape. *Canadian Journal of Botany*,
698 71(2), 303–308. <http://doi.org/10.1139/b93-031>

699 Morrison, M. J., & Stewart, D. W. (2002). Heat Stress during Flowering in Summer Brassica. *Crop*
700 *Science*, 42, 797. <http://doi.org/10.2135/cropsci2002.0797>

701 NASS. (2013a). *Cropland data layer*. USDA NASS, Research and Development Division.

702 NASS. (2013b). *North Dakota County Estimates for Canola*. USDA National Agricultural Statistics
703 Service. Retrieved from http://www.nass.usda.gov/Quick_Stats/

704 Osborne, S. L. (2007). Determining N and Yield of Canola. *Agricultural Journal*, 2(2), 180–184.

705 Piekarczyk, J. (2011). Winter Oilseed-Rape Yield Estimates From Hyperspectral Radiometer
706 Measurements. *Quaestiones Geographicae*, 30(1), 77–84. [http://doi.org/10.2478/v10117-011-0007-](http://doi.org/10.2478/v10117-011-0007-z)
707 [z.A](http://doi.org/10.2478/v10117-011-0007-z)

708 Pinter Jr., P. J., Hatfield, J. L., Schepers, J. S., Barnes, E. M., Moran, M. S., Daughtry, C. S. T., &
709 Upchurch, D. R. (2003). Remote sensing for crop management. *Photogrammetric Engineering and*
710 *Remote Sensing*, 69(6), 647–664. Retrieved from [http://www.scopus.com/inward/record.url?eid=2-](http://www.scopus.com/inward/record.url?eid=2-s2.0-0037903234&partnerID=40&md5=6c5ce957efeb0b8158b7bdcacfa14ab6)
711 [s2.0-0037903234&partnerID=40&md5=6c5ce957efeb0b8158b7bdcacfa14ab6](http://www.scopus.com/inward/record.url?eid=2-s2.0-0037903234&partnerID=40&md5=6c5ce957efeb0b8158b7bdcacfa14ab6)

712 Pratt, S. (2013, October). Satellite crop estimate too low: analysts. *The Western Producer*. Retrieved from
713 <http://www.producer.com/2013/10/satellite-crop-estimate-too-low-analysts/>

714 Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the
715 Great Plains with ERTS. In *NASA Goddard Space Flight Center 3d ERTS-1 Symposium* (pp. 309–
716 317). United States. Retrieved from citeulike-article-id:12009708

717 Roy, D. P., Borak, J. S., Devadiga, S., Wolfe, R. E., Zheng, M., & Descloitres, J. (2002). The MODIS
718 Land product quality assessment approach. *Remote Sensing of Environment*, 83(1-2), 62–76.
719 [http://doi.org/10.1016/S0034-4257\(02\)00087-1](http://doi.org/10.1016/S0034-4257(02)00087-1)

720 Sakamoto, T., Gitelson, A. A., & Arkebauer, T. J. (2014). Near real-time prediction of U.S. corn yields
721 based on time-series MODIS data. *Remote Sensing of Environment*, 147, 219–231.
722 <http://doi.org/10.1016/j.rse.2014.03.008>

723 Shen, M., Chen, J., Zhu, X., Tang, Y., & Chen, X. (2009). Yellow flowers can decrease NDVI and EVI
724 values: Evidence from a field experiment in an alpine meadow. *Canadian Journal of Remote*
725 *Sensing*, 35(2), 99–106.

726 Shen, M., Chen, J., Zhu, X., Tang, Y., & Chen, X. (2010). Do flowers affect biomass estimate accuracy
727 from NDVI and EVI? *International Journal of Remote Sensing*, 31(8), 2139–2149.
728 <http://doi.org/10.1080/01431160903578812>

- 729 Slaton, M. R., Raymond Hunt, E., & Smith, W. K. (2001). Estimating near-infrared leaf reflectance from
730 leaf structural characteristics. *American Journal of Botany*, 88 (2), 278–284. Retrieved from
731 <http://www.amjbot.org/content/88/2/278.abstract>
- 732 Sulik, J., & Long, D. S. (2015). Spectral Indices for Yellow Canola Flowers. *International Journal of*
733 *Remote Sensing*, 36(10). <http://doi.org/10.1080/01431161.2015.1047994>
- 734 Tayo, T. O., & Morgan, D. G. (1979). Factors influencing flower and pod development in oil-seed rape (*Brassica napus* L.). *The Journal of Agricultural Science*, 92(02), 363–373.
735 <http://doi.org/10.1017/S0021859600062894>
- 737 Trigg, S., & Flasse, S. (2000). Characterizing the spectral-temporal response of burned savannah using in
738 situ spectroradiometry and infrared thermometry. *International Journal of Remote Sensing*, 21(16),
739 3161–3168. <http://doi.org/10.1080/01431160050145045>
- 740 Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation.
741 *Remote Sensing of Environment*, 8(2), 127–150. [http://doi.org/10.1016/0034-4257\(79\)90013-0](http://doi.org/10.1016/0034-4257(79)90013-0)
- 742 Tucker, C. J., Holben, B. N., Elgin, J. H., & McMurtrey, J. E. (1980). Relationship of spectral data to
743 grain yield variation. *Photogrammetric Engineering and Remote Sensing*, 45, 600–608.
- 744 Tucker, C. J., & Sellers, P. J. (1986). Satellite remote sensing of primary production. *International*
745 *Journal of Remote Sensing*, 7(11), 1395–1416. <http://doi.org/10.1080/01431168608948944>
- 746 USDA Farm Service. (2013). *2012 North Dakota Acreage Report Summary*. Retrieved from
747 http://www.fsa.usda.gov/Internet/FSA_File/2012acreagereportingsummary.pdf
- 748 Yang, C., Everitt, J., & Bradford, J. (2006). Comparison of QuickBird Satellite Imagery and Airborne
749 Imagery for Mapping Grain Sorghum Yield Patterns. *Precision Agriculture*, 7(1), 33–44.
750 <http://doi.org/10.1007/s11119-005-6788-0>
- 751 Yates, D. J., & Steven, M. D. (1987). Reflexion and absorption of solar radiation by flowering canopies
752 of oil-seed rape (*Brassica napus* L.). *The Journal of Agricultural Science*, 109(3), 495 – 502.
- 753 Zhang, B., Liu, C., Wang, Y., Yao, X., Wang, F., Wu, J., ... Liu, K. (2015). Disruption of a
754 CAROTENOID CLEAVAGE DIOXYGENASE 4 gene converts flower colour from white to
755 yellow in Brassica species. *New Phytologist*, n/a–n/a. <http://doi.org/10.1111/nph.13335>

756

757 List of Figure Captions

- 758 Figure 1. Plot layout of 2013 field experiment (a) and 2014 field experiment (b) where spring canola was grown under varying
759 water and nitrogen levels. Green dots in 2013 denote subplots where measurements of canopy spectral reflectance and
760 crop yield were obtained.
- 761
- 762 Figure 2. Appearance of canola with respect to flowering on four spectral sampling dates in 2013.
- 763
- 764 Figure 3. Appearance of canola with respect to flowering on six spectral sampling dates in 2014.
- 765
- 766 Figure 4. Canola yields aggregated to the county-level in North Dakota.

767
768 Figure 5. Regression models for each Julian sampling date. Refer to Figure 2 in order to visualize the impact of flowering
769 magnitude
770
771 Figure 6. Regression models for each index using the maximum index value recorded for each subplot during the sampling
772 period.
773
774 Figure 7. Regression models for each Julian sampling date. Refer to Figure 3 in order to visualize the impact of flowering
775 magnitude
776
777 Figure 8. Temporal pattern of the strength of relationship between each index and total seed yield (top) with total rainfall for
778 each day during the study period. Note the spike in rainfall from Julian days 177-179 and subsequent increase in strength
779 of relationship for each index.
780
781 Figure 9. Regression models for each index using the maximum index value recorded for each plot during the sampling period.
782
783 Figure 10. Regression models for each MODIS 8-day composite image product.
784
785 Figure 11. Regression models for each index using the maximum index value recorded for each county across all 8-day periods.

786 Table 1. Coefficient of determination values for the maximum value of each spectral index for 2011 and 2010.
787