2012

Vegetation Outlook (VegOut): Predicting Remote Sensing–Based Seasonal Greenness

Tsegaye Tadesse
University of Nebraska-Lincoln, ttadesse2@unl.edu

Brian D. Wardlow
University of Nebraska - Lincoln, bwardlow2@unl.edu

Mark D. Svoboda
University of Nebraska - Lincoln, msvoboda2@unl.edu

Michael J. Hayes
University of Nebraska - Lincoln, mhayes2@unl.edu

Published in *Remote Sensing of Drought: Innovative Monitoring Approaches*, edited by Brian D. Wardlow, Martha C. Anderson, & James P. Verdin (CRC Press/Taylor & Francis, 2012), pp. 75–94

Vegetation Outlook (VegOut): *Predicting Remote Sensing–Based Seasonal Greenness*

Tsegaye Tadesse, Brian D. Wardlow,
Mark D. Svoboda, and Michael J. Hayes

University of Nebraska–Lincoln

Contents

1 Introduction

Accurate and timely prediction of vegetation conditions enhances knowledge-based decision making for drought planning, mitigation, and response. This is very important in countries that are highly dependent on rainfed agriculture. For example, studies show that remote sensing–based observations and vegetation condition prediction have great potential for estimating crop yields (Verdin and Klaver, 2002; Ji and Peters, 2003; Seaquist et al., 2005; Tadesse et al., 2005a, 2008; Funk and Brown, 2006), which in turn may help to address agricultural development and food security issues, as well as improve early warning systems.

Many studies have demonstrated the value of Vegetation Indices (VIs), such as the Normalized Difference Vegetation Index (NDVI), calculated from satellite observations for assessing vegetation cover and conditions (Tucker et al., 1985; Roerink et al., 2003; Anyamba and Tucker, 2005; Seaquist et al., 2005), and such data have become a common source of information for vegetation monitoring. The term *vegetation condition* in this chapter refers to vegetation greenness or vegetation health, as inferred from canopy reflectance values measured by satellite observations (Mennis, 2001; Anyamba and Tucker, 2005). The vegetation greenness metric is commonly calculated from time-series NDVI (Reed et al., 1994) and represents the seasonal, time-integrated NDVI at a specific date, which has been shown to be representative of indicators of general vegetation health including net primary production (NPP) and green biomass (Tucker et al., 1985; Reed et al., 1996; Yang et al., 1998; Eklundh and Olsson, 2003; Hill and Donald, 2003). As a result, VIs and VI derivatives such as time-integrated VI can be used to characterize the temporal and spatial relationships between climate and vegetation and improve our understanding of the lagged relationship between climate (e.g., precipitation and temperature) and vegetation response (Roerink et al., 2003; Anyamba and Tucker, 2005; Seaquist et al., 2005; Camberlin et al., 2007; Groeneveld and Baugh, 2007). Quantitative descriptions of climate-vegetation response lags can then be used to identify and predict vegetation stress during drought.

Predicting vegetation conditions over large geographic areas is imperative for a wide range of applications such as crop and rangeland condition assessments, drought monitoring, fire risk potential, and ecological studies. However, predicting vegetation conditions and understanding the impact of drought on vegetation are challenging because vegetation health is dependent not only on climatic patterns but also on complex relationships involving soil characteristics, land use/land cover (LULC), topography, and other ecological characteristics. Improvements in our predictive capabilities in this area are becoming possible with the increasing availability of many high-quality environmental data sets (e.g., climate, ocean, and remote sensing observations), longer historical records of observations, improved computing capabilities, and the emergence of advanced data analysis techniques useful for data mining.

Several studies have shown significant associations between indices of large-scale oceanic/atmospheric variables and climate over North America (e.g., Panu and Sharma, 2002; Tadesse et al., 2005b; Baigorria et al., 2008; Martinez et al., 2009). For example, Tadesse et al. (2005b) indicated a connection between the occurrences of drought over Nebraska and the Southern Oscillation Index (SOI), Multivariate El Niño Southern Oscillation Index (MEI), Pacific North American Oscillation (PNA), Pacific Decadal Oscillation (PDO), and North Atlantic Oscillation (NAO). The importance of the Atlantic and tropical Pacific sea surface temperature (SST) on past and present drought occurrences across North America is emphasized by Feng et al. (2008) in understanding North American drought variability and predictability. Martinez et al., (2009) showed a strong correlation between the climate indices and oceanic indices that are derived from the Pacific–North American pattern and tropical North Atlantic and eastern tropical Pacific SSTs to predict corn yields in the southeastern United States. Since ocean–atmosphere interactions can drive precipitation patterns affecting vegetation health, a suite of variables should be incorporated into predictive models of vegetation conditions to capture the teleconnected climate–vegetation response linkages.

In addition, the integration of satellite data with climate and oceanic data holds considerable potential for improving our capabilities to predict future vegetation conditions, as demonstrated in the work of Ji and Peters (2003) and Funk and Brown (2006). Using climate (monthly precipitation and relative humidity) and satellite (Advanced Very High Resolution Radiometer (AVHRR) NDVI) data, these studies indicated that NDVI is an effective indicator of vegetation-moisture conditions, but seasonal timing should be taken into consideration when monitoring drought with NDVI (Ji and Peters, 2003). At present, various high-quality climate, ocean, and remote sensing data sets with increasing length of records of more than 20 years are available to provide the historical basis to develop predictive techniques. In addition, the availability of advanced statistical data mining techniques such as regression tree analysis allows these diverse data sets to be effectively integrated into new vegetation-related models such as the Vegetation Drought Response Index (VegDRI) (Brown et al., 2008), upon which similar predictive models could be developed.

The National Drought Mitigation Center (NDMC), in partnership with the USDA Risk Management Agency (RMA), has developed a new drought monitoring tool called the Vegetation Outlook (VegOut) (Tadesse et al., 2005a, 2010). VegOut provides outlooks of general vegetation conditions based on prior climate and ocean index measurements, satellite-based observations of current vegetation conditions, and other environmental information including ecological setting, elevation, soil characteristics, and LULC type. Regression tree modeling was used to analyze historical time-lag relationships between satellite-observed vegetation conditions and oceanic and climatic observations and to develop empirically-based models, which are applied to a suite of "current" observations to predict future vegetation conditions at multiple time steps such as 2, 4, and 6 week outlooks.

In this chapter, the VegOut methodology will be presented in terms of the specific data inputs and predictive modeling approach. Results from two contrasting growing seasons (the 2008 drought and the 2009 nondrought years) over the central United States will be presented to demonstrate the utility and potential of VegOut for vegetation and drought monitoring. Future work to improve the current VegOut method and other possible alternative approaches that have potential use for operational drought monitoring will also be discussed.

2 DATA AND METHODS

2.1 Study Area

The 15-state region of the central United States (Figure 1) provides a geographically diverse study area in terms of land cover types, land use practices, and climate across which to test the capability of VegOut. Land cover in this study area varies from alpine forests along the Rocky Mountains in the west and the forested regions of northern Minnesota to the west–east transition of shortgrass to tallgrass prairie across the Great Plains states (e.g., Kansas, Nebraska, North Dakota, and South Dakota) and the sparsely vegetated shrubland of southern Texas and New Mexico. In addition, many parts of the study area are intensively cultivated, including the

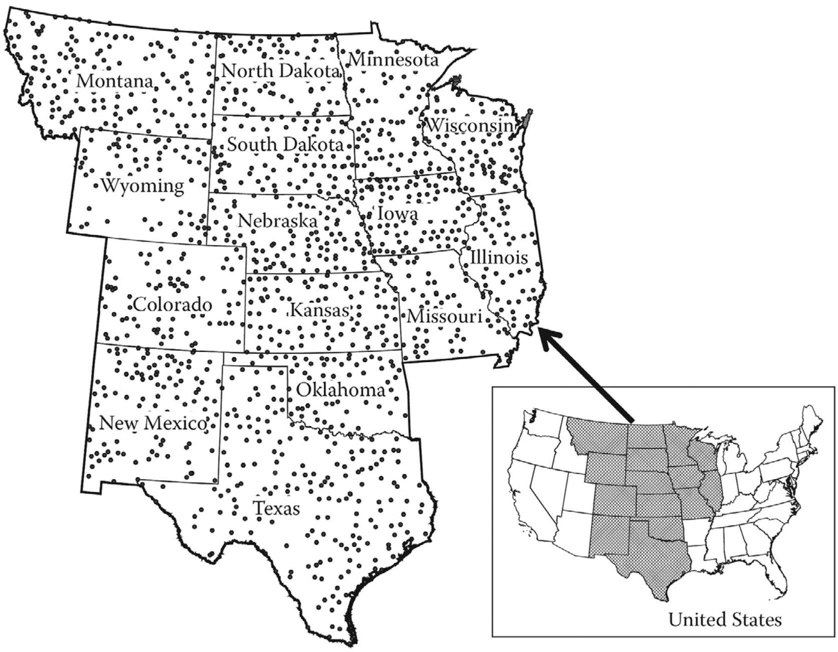

Figure 1. Central U.S. study area, showing 1420 weather stations providing training and testing data used to develop the VegOut models.

corn–soybean–dominated Corn Belt (central Nebraska eastward through Illinois and northward into Minnesota), the Winter Wheat Belt (northern Texas, central Oklahoma, and south-central Kansas), and extensive tracts of irrigated and rainfed cropland stretching the length of the Great Plains from North Dakota to Texas. The study area also has a marked precipitation gradient ranging from 255 to 510 mm in the semiarid western locations to more than 1020 mm in the east. Growing season length is also highly variable, ranging from ~125 days (mid-May to late September) in the extreme northern part of the study area to more than 250 days (late February to late November) in southern Texas.

2.2 General Overview of VegOut

The fundamental basis for developing a predictive vegetation condition tool such as VegOut is building a comprehensive and integrated database of long-term historical records of key observed variables (e.g., climate-based indices, ocean teleconnections, remote sensing–based VI, and other environmental characteristics) that contribute information regarding the complex nature of vegetation growth. These data sets must be readily available over large areas to accurately represent the range of conditions that might be encountered over the spatial modeling domain. In addition, access to these data sets in near real time is essential to the application of VegOut as an operational tool that is capable of generating informational products in a timely manner to support a variety of decision-making

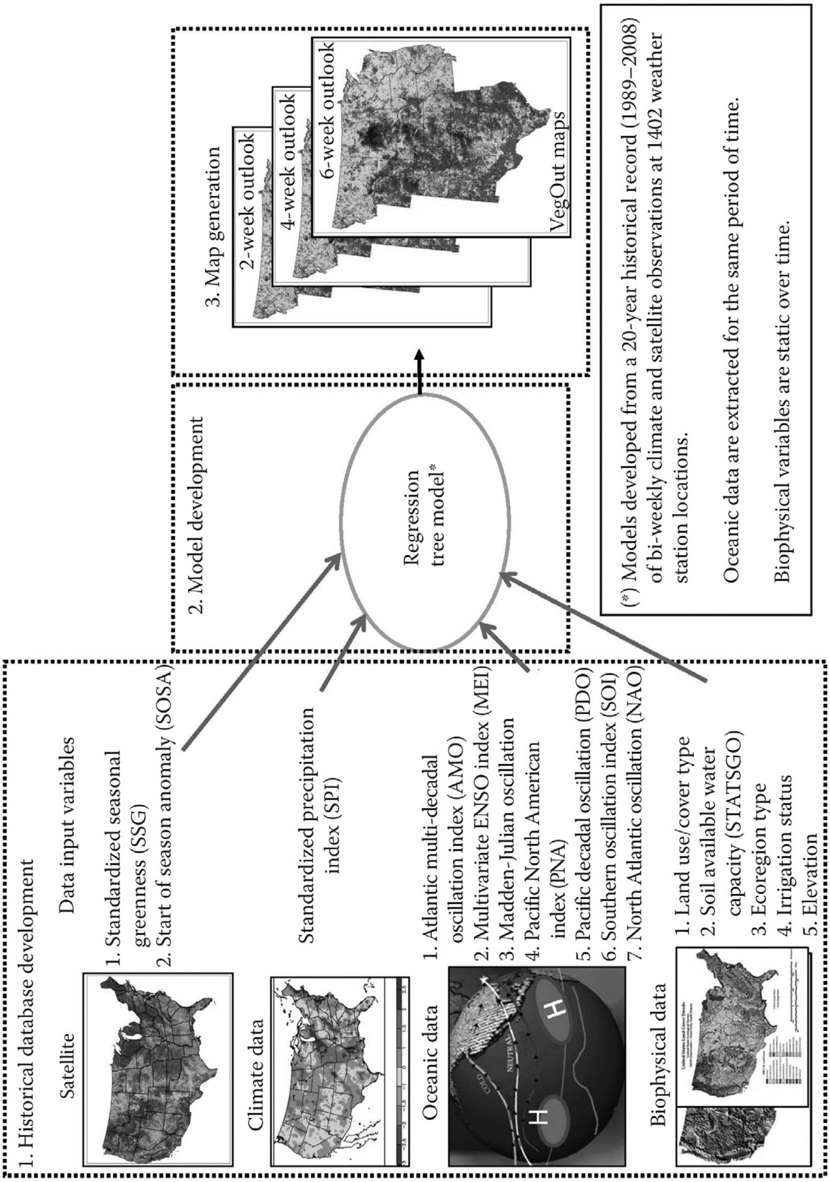

Figure 2. VegOut database development and outlook map generation process.

activities. Figure 2 shows a graphical overview of the specific variables incorporated into VegOut as well as the general methodology, which includes historical database development, rule generation for the model, and model application in the gridded image domain to produce the predicted seasonal greenness (*SG*) maps. VegOut uses rule-based regression tree models that make predictions of future vegetation conditions based on historical temporal and spatial relationships and patterns identified among satellite-derived VIs, climatic drought indices, oceanic indices, and other environmental variables.

The capability exists to predict *SG* several months into the future using VegOut. However, the predictive accuracy of VegOut decreases linearly as the prediction time interval increases (see Figure 3). The predictive accuracy of VegOut has been found to be greater than $R^2 = 0.8$ for model outlook periods of 6 weeks or less. Although testing across a variety of potential predictive outlook intervals is important for understanding model capabilities and limitations, only the results from VegOut models for 2, 4, and 6 week vegetation outlook periods are presented in this chapter to illustrate the potential of this tool for predicting the vegetation *SG*.

SG is defined as the NDVI accumulated through time from the start of the growing season (*SOS*) to a specific date in the growing season, with the accumulation continuing until the end of the growing season (*EOS*). Both the *SOS* and *EOS* are determined at the pixel level over the image domain from satellite-based time-series NDVI data. Reed et al. (1994) and Brown et al. (2008) provide more detailed descriptions of the *SG*, *SOS*, and *EOS* calculations from time-series NDVI data. For each 1 km grid cell and for a given biweekly period, the predicted *SG* patterns produced by the VegOut model are based on the analysis of patterns observed in the historical records of satellite, climate, and oceanic observations over a 20 year period (1989–2008). Also considered in the model are a set of general environmental characteristics that remain static over time but provide a baseline geographic framework to facilitate spatial differentiation of the dynamic patterns. These historical records of patterns provide the basis to forecast future *SG* in the VegOut model. For example, if *SG* patterns are being predicted on July 11 for July 25 (i.e., a 2 week vegetation outlook), then records in the historical database that exhibited similar relationships between the climatic, oceanic, satellite, and environmental variables would be used to predict the *SG* values. In short, a forward/backward time step approach (i.e., forward or backward time-lag relationship) involving the historical record among these variables is used to build the model to predict the *SG*. The specific input variables in the VegOut model (shown in Figure 2) are described in the following section.

2.3 Historic Database Development

2.3.1 Satellite Data

A time series of standardized *SG* (SSG) observations produced from biweekly AVHRR 1 km NDVI data was calculated for each year in a 20 year historical record (1989–2008) to provide vegetation condition information for the VegOut models. First, *SG* is calculated, representing the accumulated NDVI through time from the *SOS* to the last day of each biweek using the following formula:

$$SG = \sum_{p = SOS\ (=P_1)}^{EOS\ (=P_n)} (NDVI_p - NDVI_b) \tag{1}$$

where

SG is the seasonal greenness

$P_1, P_2, \ldots, P_n$ refer to individual biweekly periods

$NDVI_p$ is the observed value in the AVHRR composite data

$NDVI_b$ is the latent (or baseline) NDVI value (representative of the nonvegetated background signal) defined at the *SOS* for each pixel (Reed et al., 1994)

The SSG metric is then calculated for each biweekly time step across each year in the historical record using the following standardization formula:

$$SSG_i = \frac{SG_i - \overline{SG_i}}{\sigma_i} \tag{2}$$

where

SG_i is the current SG

$\overline{SG_i}$ is the average SG observed in the historical record up to time period i

σ is the standard deviation of these historical SG_i values (Tadesse et al., 2010)

The result is a 20 year historical time series of *SSG* images, which have zero-centered values in deviation units (ranging from −4.0 to +4.0) reflecting general vegetation conditions that are spatiotemporally comparable over both space and time because of the standardization process.

The other satellite-derived variable used in VegOut is the *SOS* anomaly (SOSA), which is the difference between the current year *SOS* date and the median 20 year historical *SOS* date. For each year in the historical record, a single SOSA value is calculated at the pixel level across the image domain using a delayed moving average approach developed by Reed et al. (1994). SOSA is used to distinguish areas with low SSG attributable to a shift to a substantially later *SOS* date, which results in a shorter interval of accumulated NDVI and thus lower SSG. Such shifts often result from human-induced LULC change, and they also frequently occur in areas of low SSG that have an *SOS* date similar to the historical average but much lower NDVI values over the same period due to some type of environmental stress that can include drought and late-spring freeze.

2.3.2 Climate Data

The Standardized Precipitation Index (SPI) (McKee et al., 1995) is used to identify climatic patterns of meteorological dryness at 2 week intervals corresponding to the biweekly periods of the satellite data across the 20 year record. The SPI is based on precipitation data and has the flexibility to detect both short- and long-term precipitation deficits. Because the SPI has the inherent flexibility to be calculated over various time spans, an optimal SPI interval had to be selected for the central United States. Exhaustive testing of all SPI intervals ranging from 1 to 51 weeks was conducted for each biweekly period across the growing season to determine the specific interval that provided the best predictive accuracy within VegOut. For each SPI interval, a 20 year analysis of the 2, 4, and 6 week VegOut model results incorporating that specific SPI was conducted for each growing season biweekly period across 1420 weather station locations within the 15-state study region (Figure 1). The test results showed the 36 week SPI consistently provided the highest VegOut model accuracy across most of the growing season, and it was selected as the SPI input for the VegOut models.

Point-based, tabular SPI data for each weather station location shown in Figure 1 were used to develop the empirically-based VegOut models. For model implementation and VegOut map production, the point-based SPI data were spatially interpolated using an inverse distance weighting technique to create 1 km gridded SPI images.

2.3.3 Oceanic/Atmospheric Index Data

As stated earlier in this chapter, several studies have shown a teleconnective link between the oceanic and climate indices (Asner et al., 2000; Los et al., 2001; Barnston et al., 2005; Tadesse et al., 2005b; Baigorria et al., 2008; Martinez et al., 2009). Understanding these relationships and using oceanic/atmospheric data help improve vegetation monitoring and prediction by incorporating various complex parameters that influence vegetation health. Schubert et al. (2007) stated that modeling work for drought prediction has largely attributed the major North American droughts of the last 150 years to global circulation anomalies that were forced by tropical SST. Based on the correlation coefficient values, however, it was observed that not all oceanic indices had a strong relationship with climate and vegetation response over the central United States (Tadesse et al., 2009).

Seven of the most commonly used oceanic/atmospheric indices were selected for integration into the VegOut predictions of SSG to account for the temporal and spatial relationships between ocean–atmosphere dynamics and climate–vegetation interactions (i.e., teleconnection patterns) that have been observed over the Central United States. These indices include the Atlantic Multidecadal Oscillation (AMO), MEI, Madden–Julian Oscillation (MJO), PNA, PDO, SOI, and NAO. Data for each of these oceanic indices are freely available online from different sources (Tadesse et al., 2009). For each oceanic index, a single value is reported in a tabular format for each time interval across the historical record, which can vary from bimonthly to monthly updates. The historical, tabular oceanic index data adapted to the biweekly time step were used to develop VegOut models. For the mapping portion of VegOut, the single oceanic index value for a specific biweek was gridded as a constant value over the 15-state study area to produce the series of 1 km oceanic index raster images.

2.3.4 Environmental Data

A set of five general environmental variables that describe aspects of the environment that influence climate–vegetation interactions were incorporated into the VegOut model. These variables include LULC type, soil available water holding capacity (AWC), ecosystem type (Eco), percent of irrigated land (Percent_Irrig), and elevation (Elev). The LULC input was derived from the 2001 National Land Cover Dataset (NLCD, Vogelmann et al., 2000) and is essential because the climate–vegetation *SG* of different land cover types such as crops, forest, and grassland (DeBeurs and Henebry, 2004) may also have different response in vegetation condition. Soil AWC, which was derived from the USDA STATSGO data set (USDA, 1994), is also a critical parameter because it defines available moisture for plant growth, and variations in soil AWC can result in different responses from the same land cover type under similar climatic conditions. The percent irrigated agriculture variable derived from the Moderate Resolution Imaging Spectroradiometer (MODIS) Irrigated Agriculture Dataset (MIrAD) (Brown et al., 2009) was incorporated to stratify the landscape into areas of rainfed vegetation that are more sensitive to climate variations such as drought versus those irrigated areas that are less affected because of targeted water applications. Elevation derived from the U.S. Geological Survey (USGS) 30 m digital elevation model (DEM) was also included to account for the different altitudinal climate regions across which a land cover type may be found (e.g., alpine versus coastal evergreen forest). An ecoregion input (Omernik, 1987) was included to account for regional differences in the collective environmental setting (e.g., climate, topography, soils, and vegetation types) that a specific land cover type (e.g., grassland) might be located (e.g., mountains versus plains). Each of these environmental variables was held constant over the 20 year study period because consistent, seamless data sets reflecting changes in these variables over time are not available for the study area.

Zonal calculations within a 3-by-3 km square window (snapped to the 1 km AVHRR pixel grid) surrounding each weather station were used to extract data from the gridded climate, environmental, and satellite variables for inclusion in the historical training database used to develop the VegOut models. For continuous variables such as percent irrigated agriculture, the mean value within each station window was calculated, and for categorical variables such as land cover type, the majority class within the window was used. Once the data were extracted for all station locations, they were merged with the tabular SPI and oceanic index data in a database used to train the VegOut models.

2.4 VegOut Model Development and Implementation

For VegOut model development, the dynamic climate, oceanic, and satellite variables in the training database were organized into a continuous time series of biweekly observations across the 20 year historical record, while the environmental variables were assumed static over this period. Biweekly VegOut models were developed from historical data extracted for 1420 weather station locations across the study area (Figure 1) using commercial classification and regression tree (CART) modeling software called Cubist (Quinlan, 1993; Rulequest, 2010). These models serve to identify historical relationships over the 20 year record at each training site between observed vegetation *SG*, climate, and oceanic conditions for a specific biweekly period and the corresponding vegetation *SG* that occurred after that date at some future biweekly time period (e.g., 2 weeks). Individual models for the 2, 4, and 6 week vegetation outlooks were generated for each biweekly period across the May–October growing season.

2.5 Model Development

The algorithm underlying the VegOut model to predict SSG is based on a series of multiple linear regression equations defined by the CART-based Cubist software through the analysis of the historical data discussed in the previous section. The model calculates the SSG value for future biweekly period $t = i$ (e.g., $t = 2$ weeks into the future) by applying a set of linear regression equations associated with historical periods in the database that exhibited similar patterns (or behavior) among the set of independent variables. To calculate the predicted value, the regression equation(s) is applied using the conditions of the current week ($t = 0$). The following is the general form of the linear regression equation defined by Cubist that is applied to calculate the SSG for a future biweekly time period $t = i$:

$$\begin{aligned}\mathrm{VegOut}(i) = {} & f_{1,i}(\mathrm{SSG, SOSA})_{t=0} + f_{2,i}(\mathrm{SPI})_{t=0} \\ & + f_{3,i}(\mathrm{LULC, Eco, Percent\_Irrig, AWC})_{t=0} \\ & + f_{4,i}(\mathrm{MEI, MJO, NAO, PDO, SOI, AMO, PNA})_{t=0} \end{aligned} \tag{3}$$

where VegOut(i) is the predicted SSG at future biweekly time period i as a function of the current ($t = 0$) values of the input variables. The equation shows that the VegOut is defined as four functions (f_1, f_2, f_3, and f_4) of the current (i.e., the date on which the *SG* prediction is made) climate, environmental, and satellite variables and the values of the oceanic indices, respectively.

2.6 VegOut Map Generation

For VegOut map production, the regression tree rules in the VegOut model for a specific biweekly period in the growing season are applied to the gridded image input data (as shown in Figure 2) for the corresponding biweekly period in a given year (e.g., June 10, 2010) using MapCubist software developed at the USGS Center for Earth Resources Observation and Science (EROS). The capability exists to apply the model in near real time to current observational inputs to produce an up-to-date VegOut map or to apply it retrospectively to generate a map for that biweekly period for any year in the historical record. During model implementation, the values of all input variables for that specific period at each pixel are considered to identify which rule(s) in the VegOut rule set should be used, which in turn determines the linear regression equation(s) that will be applied to input data values to calculate a VegOut SSG value for each pixel across the study area. In many instances, multiple rules may apply to each pixel, resulting in multiple linear regression equations being applied, and the average value across all regression calculations is used as the predicted SSG. Operationally, the period-specific VegOut models can be sequentially applied for each biweekly period across the year to generate a complete time series of 2, 4, and 6 week VegOut maps for the growing season.

3 Results and Discussion

3.1 VegOut Predictive Accuracy across the Growing Season

Figure 3 shows the average correlation between predicted and observed SSG across all periods of the growing season for 20 years (1989–2008). Examination of the predictive accuracy of the VegOut model across the growing season shows that the model's accuracy decreases linearly as the forecast interval increases. Based on these analyses, VegOut predictions presented in this chapter are limited to forecast intervals associated with historical R^2 values of 0.8 or higher (which were observed for 2, 4, and 6 week forecasts) to illustrate the potential of this new predictive approach.

Individual 2, 4, and 6 week VegOut forecasts for each biweekly period across the growing season (Figure 4) were constructed to assess their accuracy across the year as vegetation progresses through its various phenological stages. The results of this evaluation showed that the lowest predictive accuracy (R^2 = 0.7–0.8) occurred in the early spring (April and early May) for all three outlooks. By late May, the accuracy of the outlooks exceeded an R^2 value of 0.8 and was relatively stable for the remainder of the growing season. The lower R^2 values during the spring phase may be due to low green biomass associated with early stages of vegetation green-up, resulting in greater fluctuation of the SSG values during this part of the year that is magnified by early season interannual temperature variations (e.g., late spring freeze) and land management decisions (e.g., crop planting times). The relatively high and stable predictive accuracy of VegOut throughout the late spring and summer is encouraging

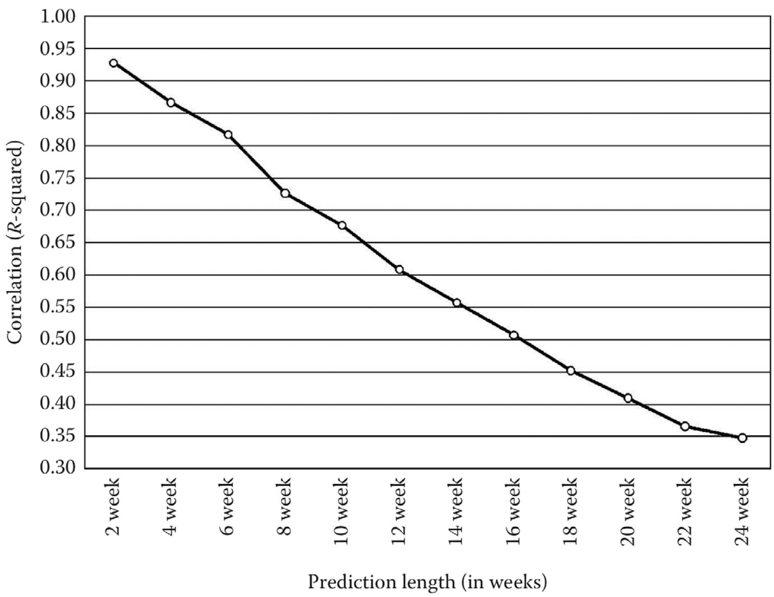

Figure 3. Twenty-year (1989–2008) average R^2 between the observed and predicted SSG values across the May–October growing season for outlook periods ranging from 2 to 24 weeks.

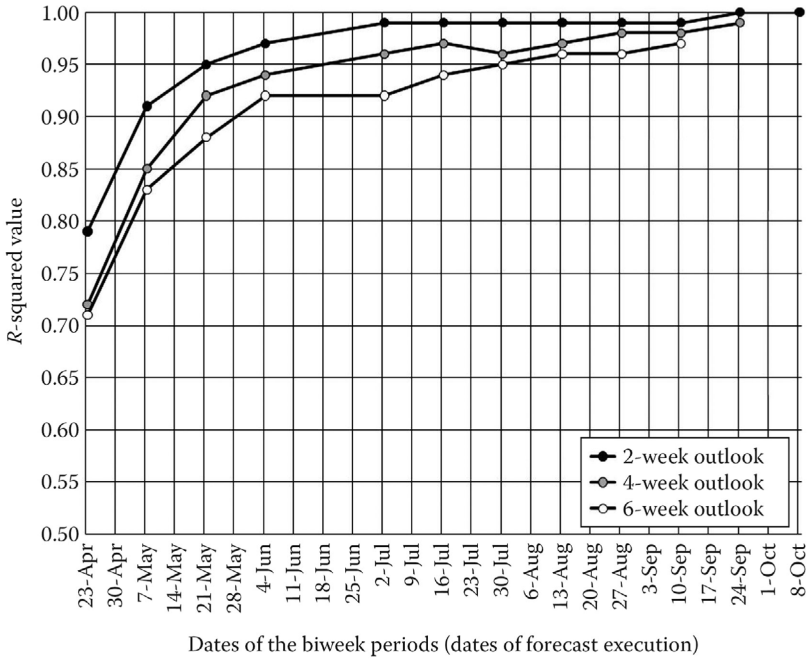

Figure 4. Twenty-year (1989–2008) average R^2 between the observed and predicted SSG values for the 2-, 4-, and 6-week outlooks for each biweekly period during the growing season.

for drought monitoring because this is an important period that determines crop yields and grassland production. The ability of VegOut to provide outlooks of vegetation SSG with reasonable accuracy over these critical months could provide new insights into the early-stage identification of emerging agricultural drought conditions.

The predictive accuracy across the growing season was consistently highest for the 2-week outlook, with the R^2 values slightly declining as the outlook interval increased. This is expected because uncertainty in future SSG values will generally increase with longer prediction intervals because of the increasing uncertainty of future states of the complex land–atmosphere system being modeled (Cushman-Roisin and Beckers, 2008).

3.2 Spatial Pattern Assessment for Drought and Nondrought Years

VegOut maps showing 2, 4, and 6 week outlooks generated for a midsummer biweekly period during a drought year (2008) and nondrought year (2009) over the study area are examined to demonstrate the capabilities of VegOut to predict SSG patterns under contrasting climatic conditions. During these 2 years, with the exception of southern Texas, a large portion of the 15-state study area experienced drought conditions in 2008 and nondrought conditions in 2009, as shown by the U.S. Drought Monitor (USDM) maps in Figure 5a and b. In 2008, for example, large areas of extreme drought (D4 classification in the USDM) over western North Dakota and

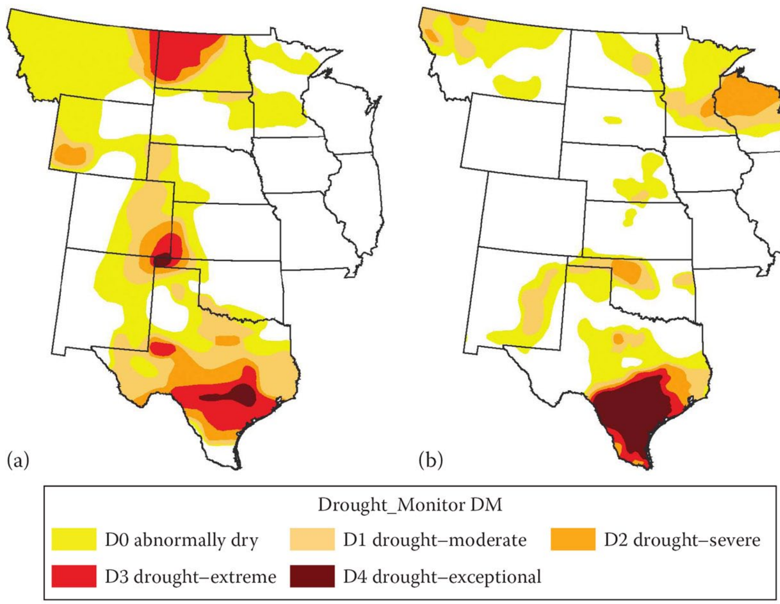

Figure 5. USDM maps over the study area for (a) July 29, 2008 and (b) July 28, 2009.

moderate to extreme drought (D2–D4 designation in the USDM) from western Nebraska southward through eastern Colorado, western Kansas, northwest Oklahoma, and parts of northern Texas were observed in the USDM (Figure 5a). However, in 2009, the conditions had improved to a nondrought or abnormally dry (D0) classification over most these areas (Figure 5b). In contrast, eastern Minnesota, northern Wisconsin, northwest Montana, and parts of eastern New Mexico and central Oklahoma were drier in 2009 than 2008.

Maps of the predicted SSG values for the 2, 4, and 6 week outlooks as forecast on June 30, 2008, and June 29, 2009, are shown in Figures 6b through d and 7b through d, respectively. The initial SSG conditions observed from AVHRR NDVI image data on forecast submission dates in 2008 and 2009 are presented in Figures 6a and 7a, respectively. Figures 6e through g and 7e through g show the observed SSG patterns from AVHRR NDVI on the targeted dates of the three vegetation outlooks in 2008 and 2009. The broad-scale spatial patterns of SSG depicted in the 2, 4, and 6 week VegOut forecasts produced across both summer seasons were in general agreement with observed SSG patterns across the 15-state area in each corresponding period. Generally, the most substantial differences between the predicted and observed SSG were limited to small, localized areas in both years (Figures 6h through j and 7h through j). In an effort to highlight major differences between the predicted and observed SSG patterns in the difference maps, ±1 standard deviation thresholds were used to indicate pixels with excessive error. In 2008, there was good spatial agreement between the SSG patterns predicted in the three outlook maps and those observed

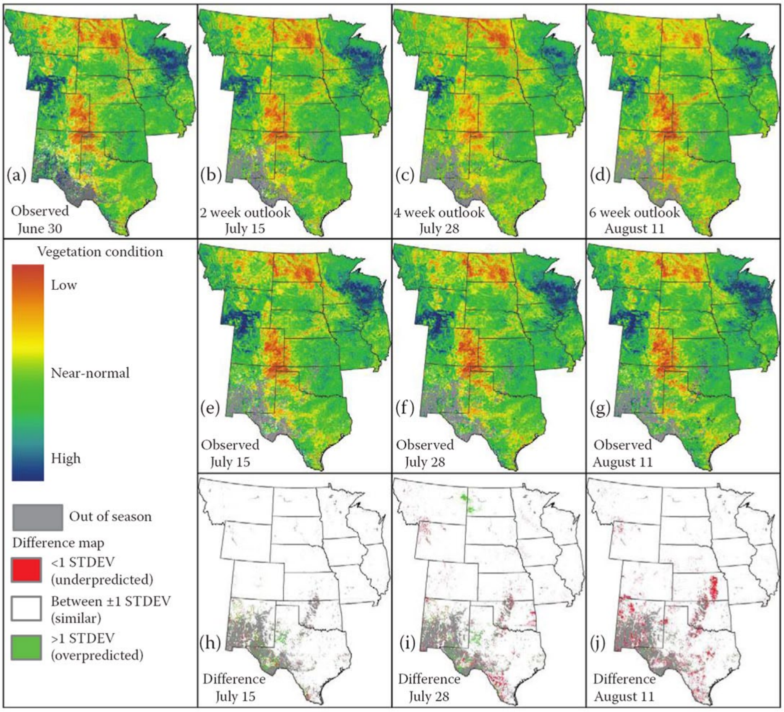

Figure 6. (a) Observed SSG for June 30, 2008; (b), (c), and (d) are 2, 4, and 6 week outlooks; (e), (f), and (g) are observed SSG for July 15, July 28, and August 11 that correspond to the 2, 4, 6 week outlooks, respectively; and (h), (i), and (j) show the difference between the predicted and observed greenness for the corresponding 2, 4, and 6 week outlooks, respectively.

from satellite over the drought-impacted areas. For example, the spatial extent and evolution of the lower SSG values observed by satellite over North Dakota and the High Plains (i.e., eastern Colorado and western Kansas) (Figure 6e through g) were consistent with those predicted by VegOut for the July and August dates (Figure 6b through d). In 2009, the spatial extent and magnitude of the SSG values were comparable between the VegOut results for the three predictive periods and the satellite observations over the drought-impacted areas in northern Wisconsin and eastern New Mexico. These results suggest that the information presented in the series of vegetation outlooks could be used as an early indicator of the impact of drought conditions on vegetation in the near future.

In addition, the major high SSG landscape features observed for these 2008 dates in Wisconsin and southwest Wyoming were also depicted in the series of VegOut maps. The most notable difference during the 2008 drought year was the slight underestimation of SSG values over some locations with either extremely high (Wisconsin) or low (south-central North Dakota) SSG values, particularly in the longer 6

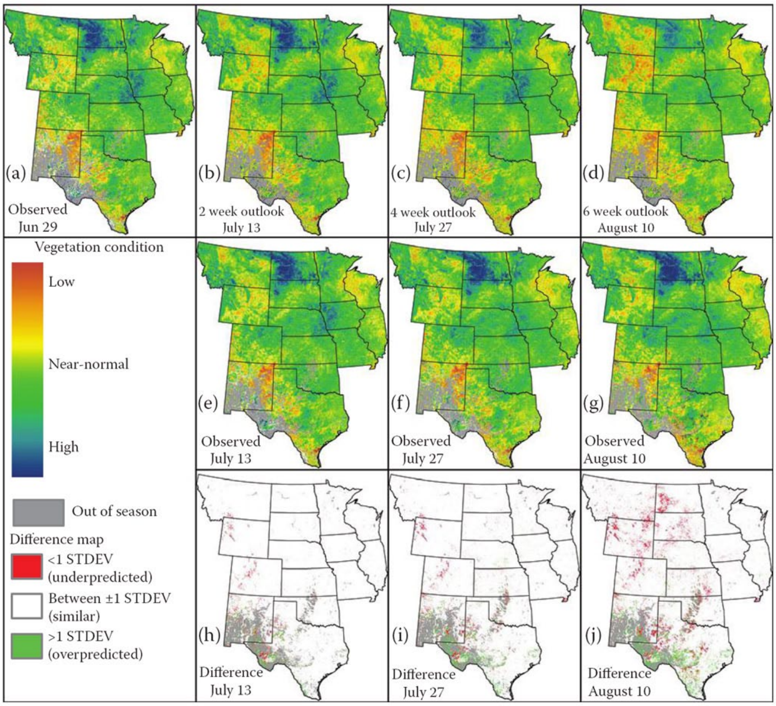

Figure 7. (a) Observed SSG for June 29, 2009; (b), (c), and (d) are 2, 4, and 6 week outlooks; (e), (f), and (g) are observed SSG for July 13, July 27, and August 10 that correspond to the 2, 4, 6 week outlooks, respectively; and (h), (i), and (j) show the difference between the predicted and observed greenness for the corresponding 2, 4, and 6 week outlooks, respectively.

week outlook maps (Figure 6d and g). In 2009, strong spatial agreement was found between the predicted and observed SSG values across the 15 states for all three outlook periods. The main exceptions were the underestimation of some high SSG values observed over western North Dakota and intermediate SSG values over Wyoming and south-central Montana in the 4 and 6 week VegOut maps.

In general, the performance of VegOut was fairly robust over most of the central United States under both drought and nondrought conditions. The exception was over sparsely vegetated areas of New Mexico and southwest Texas, where persistent differences were found in the form of both under- and over-predicted SSG values for the longer 4 and 6 week outlooks in 2008 and 2009. This discrepancy could be due to the dynamic range of SSG values (i.e., minimum to maximum SSG value range over the year), which is quite low in these sparsely vegetated landscapes; thus, a minimal difference between the predicted and observed SSG values will often exceed the one standard deviation threshold used to detect differences. However, in reality, those values may be very similar in comparison with the complete SSG value range for the entire study area.

A visual comparison of the series of difference maps showed that the majority of the study area did not have a notable difference between the observed and predicted SSG across all three outlook intervals for both study years. The majority of differences identified in the series of difference maps were distributed as small pockets for varying locations across the study area in both years. However, a closer inspection of these differences revealed several notable patterns and trends. First, the majority of marked differences flagged were associated with the underprediction of SSG values by VegOut. Second, the areal extent of these differences slightly increased as the outlook period lengthened. This would be expected given the decline in predictive accuracy of the VegOut models observed from the 2 to 6 week interval shown in Figure 3. Third, there was tendency by VegOut in some locations with extreme SSG values to underpredict the highest values and overpredict the lowest values. For example, the extremely low SSG values observed for western North Dakota and eastern Montana on July 28, 2008, (Figure 6f) were substantially underpredicted in the 4 week outlook (see the green areas in Figure 6i). Another example for high SSG values occurred over western North Dakota on August 10, 2009, where the high observed SSG values (dark blue area in Figure 7g) were substantially underestimated in the 6 week outlook (red area in Figure 7j). In general, the prediction of extreme values is challenging for any type of empirically based forecasting because there may not be representative events in the historical record used to develop the predictive models. However, these results show that the VegOut predictions do not contain a consistent bias to underpredict or overpredict extreme SSG values. Overall, the VegOut was found to predict comparable SSG values to those observed from satellite over the majority of the central United States, with the most substantial differences isolated to small geographic areas that had little impact on the overall SSG patterns depicted in the VegOut maps.

4 Future Directions

In an effort to enhance and extend VegOut as a predictive tool for mapping future vegetation conditions, several research activities are currently underway or planned in the near future. These include (1) expanded testing of VegOut over both the western and eastern United States and other regions of the world; (2) continued testing of longer outlook periods ranging from 3 to 6 months; (3) incorporating new variables such as remote sensing–based evapotranspiration, soil moisture, and land surface temperature, as well as refined sets of climate and oceanic indicators; and (4) testing and transitioning to VI data collected from new satellite sensors such as the MODIS and the Visible/Infrared Imager/Radiometer Suite (VIIRS).

A "scenario-based" VegOut modeling approach called Scenario-VegOut is also being developed to complement the "diagnostic-based" VegOut approach presented in this chapter. Scenario-VegOut is designed to predict *SG* for different climatic episodes (i.e., dry, normal, and wet conditions) using the same regression tree–based modeling and input variables as the diagnostic VegOut model. This approach provides users the flexibility to project future vegetation *SG* under these different precipitation scenarios during defined outlook periods (e.g., 2 week interval of the 2 week outlook). In this approach, Scenario-VegOut predictions are calculated for three possible scenarios over 2, 4, and 6 week outlook periods that represent below-normal, near-normal, and above-normal precipitation conditions. Scenario-VegOut will predict *SG* based on scenarios over each outlook period that represent dry conditions (e.g., 0%–50% of average precipitation), near normal conditions (e.g., 75%–125% of normal precipitation), and wet conditions (e.g., more than 150% of average precipitation). For each scenario, the SPI grids will be generated from historical, station-based precipitation values for each of these targeted precipitation percentages. This approach has the flexibility to base outlooks on different percentages if desired.

5 Summary and Conclusion

Because of the varied and potentially costly losses caused by drought events, better tools for monitoring and predicting general vegetation conditions are needed to more effectively deal with this natural hazard. VegOut attempts to fill this need by predicting vegetation *SG* patterns based on analysis of satellite, climate, and oceanic data sets and other general environmental variables using an advanced data mining technique. VegOut capitalizes on historical climate–vegetation interactions and teleconnections between the ocean and climate (such as El Niño and Southern Oscillation [ENSO]) to generate these outlooks, while considering several static environmental characteristics such as LULC type, irrigation status, soil characteristics, and ecological setting, which can influence vegetation's response to weather conditions. The goal of VegOut is to provide timely information about future vegetation conditions across large geographic areas, which can be used by drought experts to identify the early stages of vegetation drought stress and gain insight into the possible near-term trends in vegetation conditions. In addition to drought monitoring, VegOut information could also be used by agricultural producers, natural resource managers, and policy makers to make more informed decisions at local to regional scales.

The evaluation of the spatiotemporal performance of VegOut presented in this chapter across the 2008 and 2009 growing seasons found the models to have high predictive accuracy ($R^2 > 0.8$) for the central United States and predicted SSG patterns in 2, 4, and 6 week outlook maps to have strong spatial agreement with observed SSG patterns. The comparisons of the predicted and observed SSG patterns of the VegOut maps of the 2008 and 2009 summer seasons in this study showed that major differences between the predicted and observed SSG values (both underprediction and overprediction) occurred primarily at a local scale over sparsely vegetated areas. This discrepancy occurs because the narrower, possibly more tail-heavy dynamic ranges of *SG* values that frequently characterize sparsely vegetated areas reflect the more rapid changes in actual *SG* values for these regions (i.e., these regions are more sensitive to short-term climate fluctuations), which leads to increased predictive uncertainty in the models and thus more frequent, larger model "misses." Some disagreement was also found for some locations exhibiting both extreme high and low SSG values, but was restricted to relatively isolated locations within the longer outlook periods. Although the examples shown in this chapter illustrate the potential value of VegOut for predicting large-area vegetation conditions, additional validation work is needed to fully understand this new predictive tool's performance. This study was restricted to a limited number of midsummer dates, and similar evaluations should be performed across the entire growing season and across the entire 20+ year historical record of observations that are available to generate the VegOut maps. In addition, further assessment of VegOut results under varying levels of drought severity should be carried out to determine its ability to characterize both rapid and slow onset drought stress events.

Because VegOut maps and products integrate climate, satellite, and oceanic data as well as incorporate the environmental characteristics of the local areas to predict the vegetation condition with a reasonable accuracy, they can be used by agricultural producers, extension agents, early warning institutes, policy makers, and other stakeholders to make more informed decisions at the local levels.

References

Anyamba, A. and C.J. Tucker. 2005. Analysis of Sahelian vegetation dynamics using NOAA-AVHRR NDVI data from 1981–2003. *Journal of Arid Environments* 3(3):596–614.

Asner, G.P., A.R. Townsend, and B.H. Braswell. 2000. Satellite observation of El Niño effects on Amazon forest phenology and productivity. *Geophysical Research Letters* 27(7):981–984.

Baigorria, G.A., J.W. Hansen, N. Ward, J.W. Jones, and J.J. O'Brien. 2008. Assessing predictability of cotton yields in the southeastern USA based on atmospheric circulation and surface temperatures. *Journal of Applied Meteorology and Climatology* 47:76–91.

Barnston, A.G., A. Kumar, L. Goddard, and M.P. Hoerling. 2005. Improving seasonal prediction practices through attribution of climate variability. *Bulletin of the American Meteorological Society* 86:59–72.

Brown, J.F., S. Maxwell, and M.S. Pervez. 2009. Mapping irrigated lands across the United States using MODIS satellite imagery. In *Remote Sensing of Global Croplands for Food Security*, eds. P.S. Thenkabail, J.G. Lyon, H. Turral, and C.M. Biradar, pp. 177–198. Boca Raton, FL: Taylor & Francis.

Brown, J.F., B.D. Wardlow, T. Tadesse, M.J. Hayes, and B.C. Reed. 2008. The vegetation drought response index (VegDRI): A new integrated approach for monitoring drought stress in vegetation. *GIScience and Remote Sensing* 45(1):16–46.

Camberlin, P., N. Martiny, N. Philippon, and Y. Richard. 2007. Determinants of the interannual relationships between remote sensed photosynthetic activity and rainfall in tropical Africa. *Remote Sensing of Environment* 106(2):199–216.

Cushman-Roisin, B. and J-M. Beckers. 2008. *Introduction to Geophysical Fluid Dynamics: Physical and Numerical Aspects* 2nd edn. Amsterdam, the Netherlands: Academic Press, Elsevier.

DeBeurs, K.M. and G.M. Henebry, 2004. Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan. *Remote Sensing of Environment* 89(4):497–509.

Eklundh, L. and L. Olsson, 2003. Vegetation index trends for the African Sahel 1982–1999. *Geophysical Research Letters* 30:131–134.

Feng, S., R.J. Oglesby, C.M. Rowe, D.B. Loope, and Q. Hu. 2008. Atlantic and Pacific SST influences on Medieval drought in North America simulated by the

Community Atmospheric Model. *Journal of Geophysical Research* 113:D11101, doi:10.1029/2007JD009347.

Funk, C.C. and M.E. Brown. 2006. Intra-seasonal NDVI change projections in semi-arid Africa. *Remote Sensing of Environment* 101:249–256.

Groeneveld, D.P. and W.M. Baugh. 2007. Correcting satellite data to detect vegetation signal for eco-hydrologic analyses. *Journal of Hydrology* 344:135–145.

Hill, M.J. and G.E. Donald. 2003. Estimating spatio-temporal patterns of agricultural productivity in fragmented landscapes using AVHRR NDVI time series. *Remote Sensing of Environment* 84:367–384.

Ji, L. and A.J. Peters. 2003. Assessing vegetation response to drought in the northern Great Plains using vegetation and drought indices. *Remote Sensing of Environment* 87:85–98.

Los, S.O., G.J. Collatz, L. Bounoua, P.J. Sellers, and C.J. Tucker. 2001. Global interannual variations in sea surface temperature and land surface vegetation, air temperature, and precipitation. *Journal of Climate* 14(7):1535–1549.

Martinez, C.J, G.A. Baigorria, and J.W. Jones. 2009. Use of indices of climate variability to predict corn yields in the Southeast USA. *International Journal of Climatology* 29:1680–1691.

McKee, T.B., N.J. Doesken, and J. Kleist. 1995. Drought monitoring with multiple time scales. *Preprints, 9th Conference on Applied Climatology*, Dallas, TX, pp. 233–236. Boston, MA: American Meteorological Society.

Mennis, J. 2001. Exploring relationships between ENSO and vegetation vigour in the south-east USA using AVHRR data. *International Journal of Remote Sensing* 22(16):3077–3092.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. *Annals of the Association of American Geographers* 77(1):118–125.

Panu, U.S. and T.C. Sharma. 2002. Challenges in drought research: Some perspectives and future directions. *Hydrological Sciences Journal* 47(S):19–29.

Quinlan, J.R. 1993. *C4.5 Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers.

Reed, B.C., J.F. Brown, D. VanderZee, T.R. Loveland, J.W. Merchant, and D.O. Ohlen. 1994. Measuring phenological variability from satellite imagery. *Journal of Vegetation Science* 5:703–714.

Reed, B.C., T.R. Loveland, and L.L. Tieszen. 1996. An approach for using AVHRR data to monitor U.S. Great Plains grasslands. *Geocarto International* 11(3):13–22.

Roerink, G.J., M. Menenti, W. Soepboer, and Z. Su. 2003. Assessment of climate impact on vegetation dynamics by using remote sensing. *Physics and Chemistry of the Earth* 28:103–109.

Rulequest. 2010. An overview of Cubist. http://www.rulequest.com/cubist-win.html (accessed on December 12, 2011).

Schubert, S., R. Koster, M. Hoerling, R. Seager, D. Lettenmaier, A. Kumar, and D. Gutzler. 2007. Predicting drought on seasonal-to-decadal time scales. *Bulletin of the American Meteorological Society* 88:1625–1630.

Seaquist, J.W., L.L. Olsson, and J. Ardö. 2005. A remote sensing-based primary production model for grassland biomes. *Ecological Modeling* 169(1):131–155.

Tadesse, T., J.F. Brown, and M.J. Hayes. 2005a. A new approach for predicting drought-related vegetation stress: Integrating satellite, climate, and biophysical data over the U.S. central plains. *ISPRS Journal of Photogrammetry and Remote Sensing* 59(4):244–253.

Tadesse, T., M. Haile, G. Senay, C. Knutson, and B.D. Wardlow. 2008. Building integrated drought monitoring and food security systems in sub-Saharan Africa. *Natural Resources Forum* 32(4):245–279.

Tadesse, T., B.D. Wardlow, M.J. Hayes, M.D. Svoboda, and J.F. Brown. 2010. The vegetation condition outlook (VegOut): A new method for predicting vegetation seasonal greenness. *GIScience and Remote Sensing* 47(1):25–52.

Tadesse, T., B.D. Wardlow, and J.H. Ryu. 2009. Discovering the spatial and temporal relationships between vegetation condition and climate in monitoring drought. *Proceedings of the 2009 Climate Prediction Applications Science Workshop*, Norman, OK, March 23–25.

Tadesse, T., D.A. Wilhite, M.J. Hayes, S.K. Harms, and S. Goddard. 2005b. Discovering associations between climatic and oceanic parameters to monitor drought in Nebraska using data-mining techniques. *Journal of Climate* 18(10):1541–1550.

Tucker, C.J., J.R.G. Townshend, and T.E. Goff. 1985. African land cover classification using satellite data. *Science* 9227(4685):369–375.

USDA (U.S. Department of Agriculture), Natural Resource Conservation Service. 1994. State Soil Geographic (STATSGO) Data Base: Data Use Information, Washington, DC: U.S. Soil Conservation Service. *Miscellaneous Publication Number* 1492.

Verdin, J. and R. Klaver. 2002. Grid cell based crop water accounting for the famine early warning system. *Hydrological Processes* 16:1617–1630.

Vogelmann, J.E., S.M. Howard, L. Yang, C.R. Larson, B.K. Wylie, and N. Van Driel. 2000. Completion of the 1990s National Land Cover Dataset for the coterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering & Remote Sensing* 67(6):650–661.

Yang, L., B.K. Wylie, L.L. Tieszen, and B.C. Reed. 1998. An analysis of relationships among climate forcing and time-integrated NDVI of grasslands over the U.S. northern and central Great Plains. *Remote Sensing of the Environment* 65:25–37.