

2015

## Developing a 30-m grassland productivity estimation map for central Nebraska using 250-m MODIS and 30-m Landsat-8 observations

Yingxin Gu

US Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, [yingxin.gu.ctr@usgs.gov](mailto:yingxin.gu.ctr@usgs.gov)

Bruce K. Wylie

USGS EROS, [wylie@usgs.gov](mailto:wylie@usgs.gov)

Follow this and additional works at: <http://digitalcommons.unl.edu/usgsstaffpub>

 Part of the [Geology Commons](#), [Oceanography and Atmospheric Sciences and Meteorology Commons](#), [Other Earth Sciences Commons](#), and the [Other Environmental Sciences Commons](#)

---

Gu, Yingxin and Wylie, Bruce K., "Developing a 30-m grassland productivity estimation map for central Nebraska using 250-m MODIS and 30-m Landsat-8 observations" (2015). *USGS Staff-- Published Research*. 917.

<http://digitalcommons.unl.edu/usgsstaffpub/917>

This Article is brought to you for free and open access by the US Geological Survey at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in USGS Staff -- Published Research by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



# Developing a 30-m grassland productivity estimation map for central Nebraska using 250-m MODIS and 30-m Landsat-8 observations



Yingxin Gu<sup>a,\*</sup>, Bruce K. Wylie<sup>b</sup>

<sup>a</sup> ASRC InuTeq, Contractor to U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, 47914 252nd Street, Sioux Falls, SD 57198, USA

<sup>b</sup> USGS EROS, 47914 252nd Street, Sioux Falls, SD 57198, USA

## ARTICLE INFO

### Article history:

Received 17 April 2015

Received in revised form 25 August 2015

Accepted 22 October 2015

Available online 5 November 2015

### Keywords:

Satellite vegetation index (NDVI)

Grassland productivity

Downscaling MODIS GSN

Landsat

Regression tree model

Central Nebraska

## ABSTRACT

Accurately estimating aboveground vegetation biomass productivity is essential for local ecosystem assessment and best land management practice. Satellite-derived growing season time-integrated Normalized Difference Vegetation Index (GSN) has been used as a proxy for vegetation biomass productivity. A 250-m grassland biomass productivity map for the Greater Platte River Basin had been developed based on the relationship between Moderate Resolution Imaging Spectroradiometer (MODIS) GSN and Soil Survey Geographic (SSURGO) annual grassland productivity. However, the 250-m MODIS grassland biomass productivity map does not capture detailed ecological features (or patterns) and may result in only generalized estimation of the regional total productivity. Developing a high or moderate spatial resolution (e.g., 30-m) productivity map to better understand the regional detailed vegetation condition and ecosystem services is preferred. The 30-m Landsat data provide spatial detail for characterizing human-scale processes and have been successfully used for land cover and land change studies. The main goal of this study is to develop a 30-m grassland biomass productivity estimation map for central Nebraska, leveraging 250-m MODIS GSN and 30-m Landsat data. A rule-based piecewise regression GSN model based on MODIS and Landsat ( $r = 0.91$ ) was developed, and a 30-m MODIS equivalent GSN map was generated. Finally, a 30-m grassland biomass productivity estimation map, which provides spatially detailed ecological features and conditions for central Nebraska, was produced. The resulting 30-m grassland productivity map was generally supported by the SSURGO biomass production map and will be useful for regional ecosystem study and local land management practices.

© 2015 Elsevier Inc. All rights reserved.

## 1. Introduction

Accurately describing and assessing aboveground vegetation biomass productivity is very important for best land management practices. Previous aboveground vegetation biomass production and yield information have been derived from ground observations (e.g., National Agricultural Statistics Service (NASS) crop yield data, Soil Survey Geographic (SSURGO) rangeland productivity estimates, and flux tower observations). The limitations of these approaches include (1) county level statistics with very low spatial resolutions; (2) sparse field observations that lack the continuous spatial coverage of vegetation canopy; and (3) spatial discontinuities, such as differences across state and county lines due to the slightly different criteria for their soil surveys (Curran & Williamson, 1985, 1986; Gu, Wylie, & Bliss, 2013; Han, Yang, Di, & Mueller, 2012; Tieszen, Reed, Bliss, Wylie, & DeJong, 1997). Satellite remote sensing data, which have wide coverages and high spatial and temporal resolutions, have been widely used for monitoring and characterizing landscape scale vegetation dynamics and ecosystem services (Anderson, Hardy, Roach, & Witmer, 1976; Gu,

Brown, Verdin, & Wardlow, 2007; Gu & Wylie, 2010; Gu, Wylie, & Howard, 2015; Potter et al., 1993; Reed et al., 1994; Tucker, Vanpraet, Sharman, & Van Ittersum, 1985; Wylie et al., 2008). Satellite-derived growing season integrated (or averaged) Normalized Difference Vegetation Index (NDVI) has been used as a proxy for vegetation biomass productivity because the growing season integrated NDVI (GSN) captures the seasonal dynamics throughout the growing season (Becker-Reshef, Vermote, Lindeman, & Justice, 2010; Funk & Budde, 2009; Gu et al., 2013; Gu, Wylie, Howard, Phuyal, & Ji, 2013b; Hobbs, 1995; Ji et al., 2012; Tieszen et al., 1997; Wang, Rich, Price, & Kettle, 2004; Wylie et al., 1995).

In our previous study, we assessed the relationship between Moderate Resolution Imaging Spectroradiometer (MODIS) GSN and SSURGO annual grassland productivity for the Greater Platte River Basin (GPRB) and developed an empirical equation to estimate grassland biomass productivity based on the MODIS GSN (Gu et al., 2013). The advantage of MODIS is that it has high temporal resolution (1 to 2 days revisit time) and a wide range of wavelengths (36 spectral bands, wavelengths range from 0.4  $\mu\text{m}$  to 14.4  $\mu\text{m}$ ), which made the MODIS land surface products reliable and robust (e.g., maximum value composites of NDVI data, cloud removal, and atmospherically corrected for cloud, cloud shadows, and aerosols, [https://lpdaac.usgs.gov/products/modis\\_](https://lpdaac.usgs.gov/products/modis_)

\* Corresponding author at: 47914 252nd Street, Sioux Falls, SD 57198, USA.  
E-mail address: [ygu@usgs.gov](mailto:ygu@usgs.gov) (Y. Gu).

[products\\_table/modis\\_overview](#)). A 250-m MODIS GSN-based grassland productivity estimation map was generated over the GPRB. The derived grassland productivity estimation map improved the regional consistency of the SSURGO grassland productivity map and provided useful information for land management (Gu et al., 2013). However, this 250-m grassland productivity map can only provide a generalized ecological condition and coarser scale pattern information of a region and could not capture more detailed site-specific information for specific areas. Therefore, a high spatial or moderate resolution (e.g., 30-m) productivity map is needed to better understand the regional detailed ecosystem condition and services.

An approach was recently developed that combined 250-m MODIS GSN and 30-m Landsat multiple-date observations to down-scale MODIS GSN to 30 m and generated a quality improved, atmospherically corrected, high spatial resolution (30 m) GSN map for a single Landsat path/row located in northeastern Colorado (Gu & Wylie, 2015). This multiple sensor method retains the detailed seasonal dynamic information captured by MODIS while leveraging the high-resolution information from Landsat. The derived 30-m GSN map provides detailed biophysical information for high and moderate scale ecological features (Gu & Wylie, 2015). The objectives of this study are to (1) apply this previous method (Gu & Wylie, 2015) to a large area (e.g., two adjacent Landsat path/rows in central Nebraska with different scene acquisition dates), (2) develop a single rule-based piecewise regression GSN model that spans the two Landsat path/row data sets (facilitates automation and regional analysis), (3) predict the 30-m MODIS-Landsat GSN and investigate the GSN spatial dynamic for the study area, and (4) develop a 30-m grassland biomass productivity estimation map for central Nebraska, which will provide spatially detailed ecological features and conditions of central Nebraska and will be useful for regional ecosystem study and local land management.

## 2. Materials and method

### 2.1. Study area

The study area is located in central Nebraska (within the red box in Fig. 1) and crosses two Landsat path/rows (light blue polygon for path/row 30/31 and green polygon for path/row 29/31 in Fig. 1). The main vegetation cover types in the study area are grasslands (~49%) and cultivated crops (~41%). Other vegetation cover types include forest, pasture and hay, and shrubland (Jin et al., 2013). The broad range of vegetation productivities in the study area (i.e., GSN ranges from 0.3 to 0.8) can help to develop a robust, unbiased, and reliable GSN prediction model.

### 2.2. Data

#### 2.2.1. Landsat 8 data

Our previous study indicated that GSN can be successfully predicted based on the three-date (across the growing season) Landsat data and that the cloud mask data were not useful in the GSN model (Gu et al., 2015). In this study, we selected three similar dates for each Landsat path/row (Julian dates 184 and 193, 232 and 241, and 264 and 273 for each path/row) with low cloud cover during the 2013 growing season. The 30-m Landsat 8 Level 1 T (Standard Terrain Corrected data which were the best available Landsat 8 data when this study started, [http://landsat.usgs.gov/Landsat\\_Processing\\_Details.php](http://landsat.usgs.gov/Landsat_Processing_Details.php)) data for the seven Landsat 8 bands (bands 2–7 for vegetation mapping and band 9 for cloud detection) for the three selected dates of each Landsat path/row were obtained through the Global Visualization Viewer (<http://glovis.usgs.gov/>). All Landsat data layers were clipped within the boundary of the study area (Fig. 1). Pixels outside each Landsat scene but within the study area were set to “0” (lower panel in Fig. 1) in order to allow

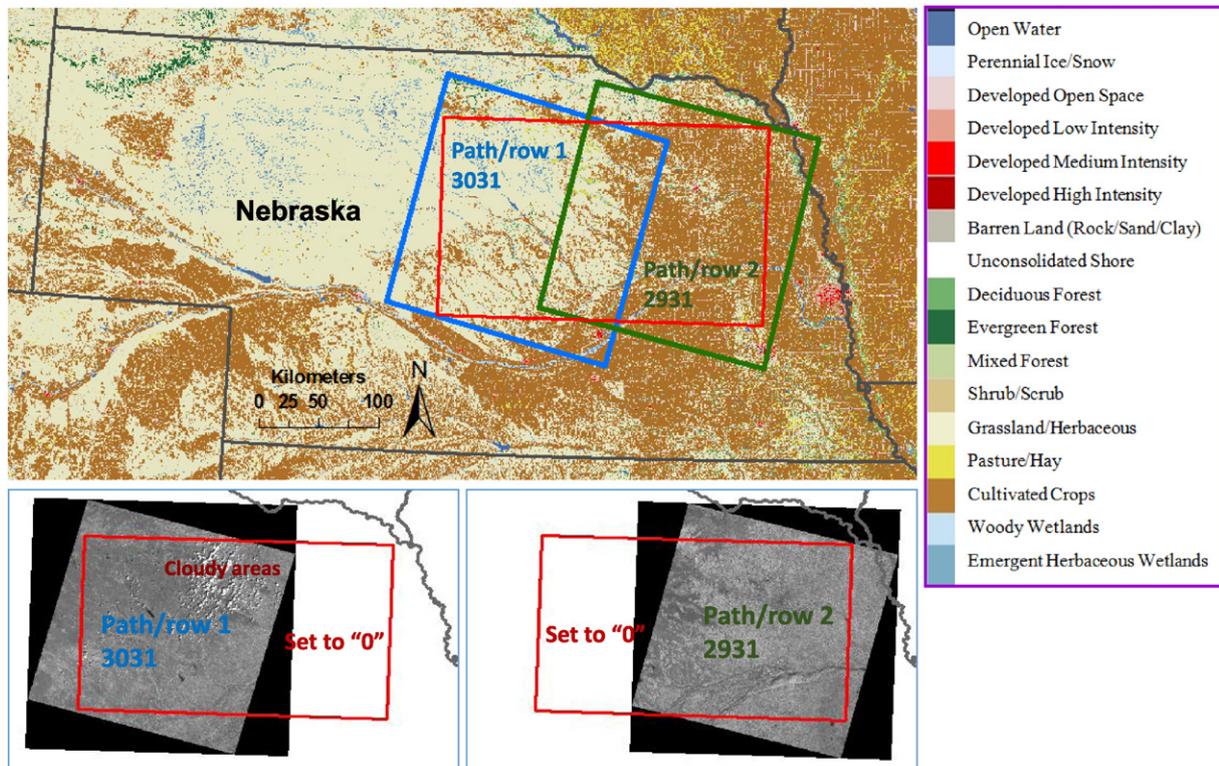


Fig. 1. Location of the study area (inside the red outline), the land cover types as identified by the National Land Cover Database (NLCD) 2011, and the locations of the two Landsat path/rows (light blue and green polygons). All Landsat 8 variable values were set to “0” when the pixels were outside the Landsat scene but within the study area (lower panel).

the single overall GSN model to successfully recognize the data gaps between the two selected Landsat path/rows (i.e., “0” value regions in Fig. 1) and force the GSN predictions to use only the available Landsat path/row scene data (i.e., based on the none “0” data). The 30-m Landsat data were then upscaled to 250-m using the “spatial averaging” method.

2.2.2. MODIS GSN and grassland mask

The 7-day maximum value composites of 250-m MODIS NDVI data for 2013 were obtained from the USGS expedited MODIS (eMODIS) data archive (<https://lta.cr.usgs.gov/emodis>). An enhanced maximum value composite algorithm, which filters out input surface reflectance data with bad quality, negative values, clouds, snow cover, and low view angles based on the MODIS quality assurance (QA) data, was used in the eMODIS 7-day composite NDVI generation (Jenkerson, Maersperger, & Schmidt, 2010). The 2013 time series of NDVI data were smoothed using a weighted least-squares approach to reduce additional atmospheric noise (Swets, Reed, Rowland, & Marko, 1999). The 2013 growing season averaged NDVI was then calculated using the start of season time as early April (~Julian date 100) and the end of season time as late October (~Julian date 300). The derived GSN map was approximately classified into three classes of productivity (low, medium, or high) based on the GSN values. The USGS 30-m Multi-Resolution Land Characteristics (MRLC, 2011) data (<http://www.mrlc.gov/nlcd2011.php>) were used to identify grassland pixels. A total of 24,295,466 grassland pixels (about 22,000 km<sup>2</sup>) were located within the study area.

2.3. Developing a 30-m MODIS-Landsat GSN map for the two Landsat path/rows

In this study, we used a data mining method to generate a 30-m MODIS-Landsat GSN map for the two adjacent Landsat path/rows located in central Nebraska. Cubist software (<http://www.rulequest.com>), which develops a generalized set of rules with associated multiple regression models (a series of piecewise regressions) constrained by the data ranges of the input variables, was used to build the GSN models. The main procedures for developing the MODIS-Landsat GSN model and generating the 30-m GSN map for the 2-Landsat path/row study area included the following steps:

1. Selected ~9000 randomly stratified samples from the study area based on the three MODIS GSN (productivity) classes, with each productivity category having ~3000 random samples.
2. Extracted the 250-m MODIS GSN (dependent variable) and the 3-date 7-band 250-m averaged Landsat 8 data for the two Landsat path/rows (independent variables) for the selected samples (pixels).

3. Developed a data-driven rule-based piecewise regression MODIS-Landsat GSN model at 250-m resolution using Cubist.
4. Estimated the 250-m MODIS-Landsat predicted GSN based on the 250-m averaged Landsat 8 data and the derived 250-m GSN model.
5. Generated the GSN absolute difference map at 250-m based on the actual MODIS GSN and the predicted GSN, evaluated the high error (i.e., absolute GSN difference ≥ 0.2) areas, and confirmed the accuracy and consistency of the derived GSN model prediction. No high error regions were observed in this study, but if the high error regions were large enough, additional training samples would need to be selected in these regions and added into the entire training data set to insure a robust GSN mapping model.
6. Mapped the 30-m MODIS-Landsat predicted GSN by applying the 250-m GSN mapping model to the original 30-m Landsat 8 data.

Fig. 2 is a flowchart summarizing the “2-Landsat path/row down-scaling GSN” approach presented in this study. The derived 30-m MODIS-Landsat GSN map was evaluated by comparing it with Google Earth images, the original 30-m Landsat RGB (Red, Green, Blue) images, and the 250-m MODIS GSN map.

2.4. Generating a 30-m grassland productivity map for central Nebraska

A 30-m grassland biomass productivity estimation map for central Nebraska was generated based on the 30-m predicted GSN derived from this study and an empirical equation (Eq. (1)) in Gu et al. (2013):

$$\text{Grassland biomass productivity (kg ha}^{-1} \text{ year}^{-1}) = 9936.5 \times \text{GSN} - 1554 \quad (1)$$

Eq. (1) was developed based on the eMODIS GSN, and the 30-m predicted GSN is downscaled eMODIS GSN; therefore, data consistencies for Eq. (1) are preserved. The spatial patterns and the locations of the productive and the unproductive grasslands in central Nebraska were evaluated using the derived 30-m grassland productivity estimation map. The resulting 30-m grassland productivity estimation map was assessed and validated by the 30-m SSURGO total biomass productivity map.

3. Results and discussion

3.1. MODIS-Landsat GSN model for the two Landsat path/rows

A data-driven rule-based piecewise regression MODIS-Landsat GSN mapping model was developed for the two Landsat path/rows in the study area at a 250-m spatial resolution. The correlation coefficient

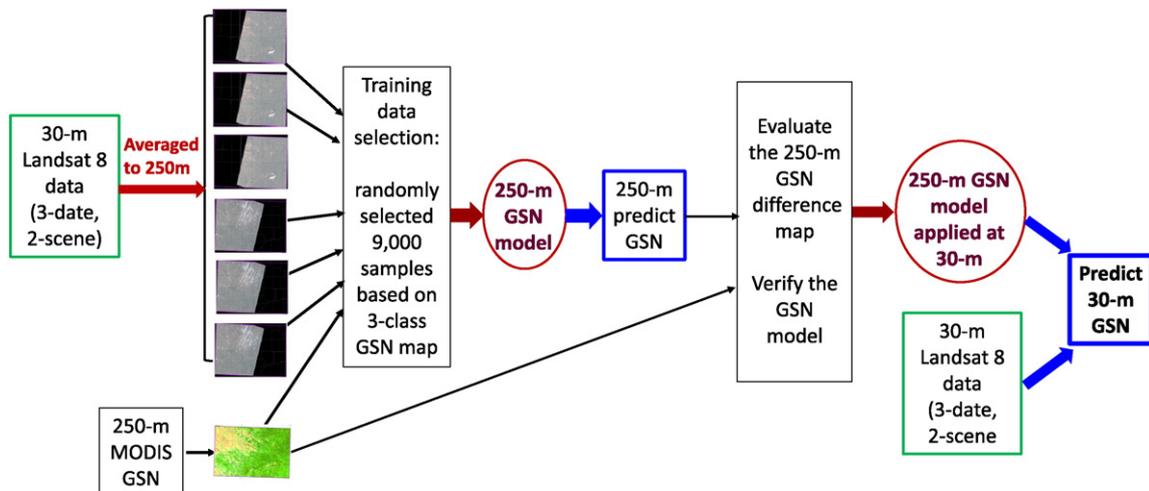


Fig. 2. Flowchart on developing the 2-Landsat path/row 30-m MODIS-Landsat GSN map.

(r) between the predicted GSN and the actual GSN at 250-m resolution is 0.91 with a mean absolute error of 0.026 for the 1491 test samples (Fig. 3a), indicating that the derived 250-m resolution GSN model can successfully predict GSN across the study area based on the two Landsat path/row data. The important Landsat variables that were used in the GSN model are data from Landsat bands 2 to 7. Landsat band 9 (Cirrus 1.36–1.38 μm band, used for detection of cirrus cloud contamination; [http://landsat.usgs.gov/best\\_spectral\\_bands\\_to\\_use.php](http://landsat.usgs.gov/best_spectral_bands_to_use.php)) is the least important variable used in the GSN model (Table 1).

The absolute difference map between the actual MODIS GSN and the 250-m predicted GSN indicated that only a very small portion (~0.2%) of the study area had high absolute GSN difference (≥0.2). These small areas primarily consisted of open water, cultivated crops, pasture and hay, and some grasslands. The good spatial agreement at the 250-m resolution indicates that the GSN mapping model can successfully recognize useful data availability across the two different Landsat path/rows and that the GSN mapping model was robust and reliable. This also implies that the 2-Landsat path/row downscaled GSN method can be applied to the multiple adjacent path/rows (e.g., four or more adjacent Landsat path/rows).

3.2. 30-m MODIS-Landsat predicted GSN map for central Nebraska

Figs. 3b and 3c are the 30-m MODIS-Landsat predicted GSN and the 250-m MODIS actual GSN maps, respectively, for central Nebraska. The 30-m MODIS-Landsat predicted GSN map (Fig. 3b) provides a spatially explicit and seamless GSN image despite being estimated from two different Landsat path/rows with different scene dates (Fig. 1). The general spatial patterns of the GSN are similar for the 30-m and 250-m GSN maps (Fig. 3b and c). For example, the GSN values (productivity) increase from west to east through the study area because of the different climate conditions (dry to wet climate conditions). Both 30-m and 250-m GSN maps illustrate very low GSN values (i.e., low productivity) in the Sand Hills ecoregion (within the purple boundary) because of the sandy soil condition. These results demonstrated the robustness and reliability of the 2-Landsat path/row MODIS-Landsat GSN model at 250-m resolution.

Fig. 4 provides a detailed view of a region in Holt County, Nebraska (Fig. 3b, Box 1): (a) 30-m MODIS-Landsat predicted GSN, (b) Google Earth map, and (c) 250-m MODIS actual GSN. The center pivot irrigation systems are clearly shown in the 30-m GSN and the Google Earth maps (Fig. 4a and b) but could not be identified from the 250-m MODIS GSN

Table 1

Attribute usage in the 2-Landsat path/row rule-based piecewise GSN regression model (at 250-m resolution). Only those variables with ≥40% average usage in the GSN model are shown in the table. Name explanation: (1) the first three characters of the name represent the Julian dates and (2) B1-9 represent Landsat 8 bands 1–9.

Name	Usage in rule stratification (%)	Usage in regression model (%)	Average usage (%)
264B2	61	91	76
184B2	52	90	71
273B4	33	96	65
241B2	26	97	62
273B5	29	93	61
264B5	44	66	55
184B5	22	77	50
241B5	16	82	49
241B3	7	90	49
193B2	6	90	48
241B4	8	87	48
193B6	0	95	48
273B2	5	88	47
264B7	2	90	46
193B4	7	85	46
184B4	11	80	46
264B4	0	90	45
241B6	0	90	45
232B4	13	75	44
273B9	2	86	44
193B5	0	88	44
264B3	0	87	44
241B7	7	79	43
273B6	3	82	43
193B7	8	76	42
193B3	2	81	42
184B7	1	82	42
184B6	5	76	41
273B3	1	79	40

map (Fig. 4c) because of the coarse spatial resolution. The 30-m MODIS-Landsat GSN estimation map can even provide very small environment features of the region. For example, the “Sand blow out” areas and the “Fence lines” detected from the Google Earth map (Fig. 4 zoom maps 1 and 2) were characterized as very low GSN values in the 30-m GSN estimation map (red areas indicated by the blue arrow in Fig. 4a). Furthermore, the 30-m GSN estimation map can reflect the seasonal dynamics of a region rather than a single time observation map derived from Google Earth, since the GSN represents vegetation conditions

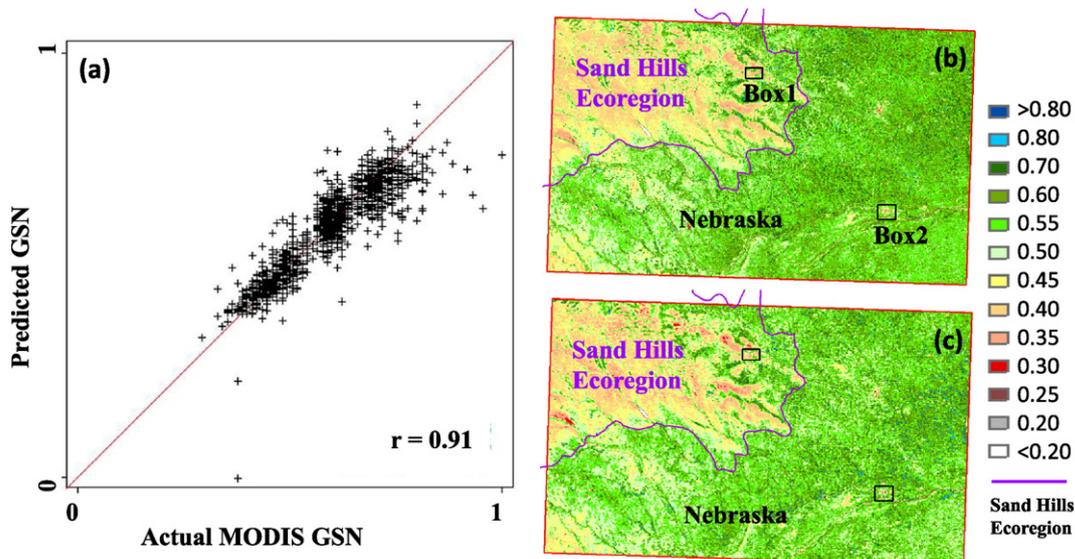
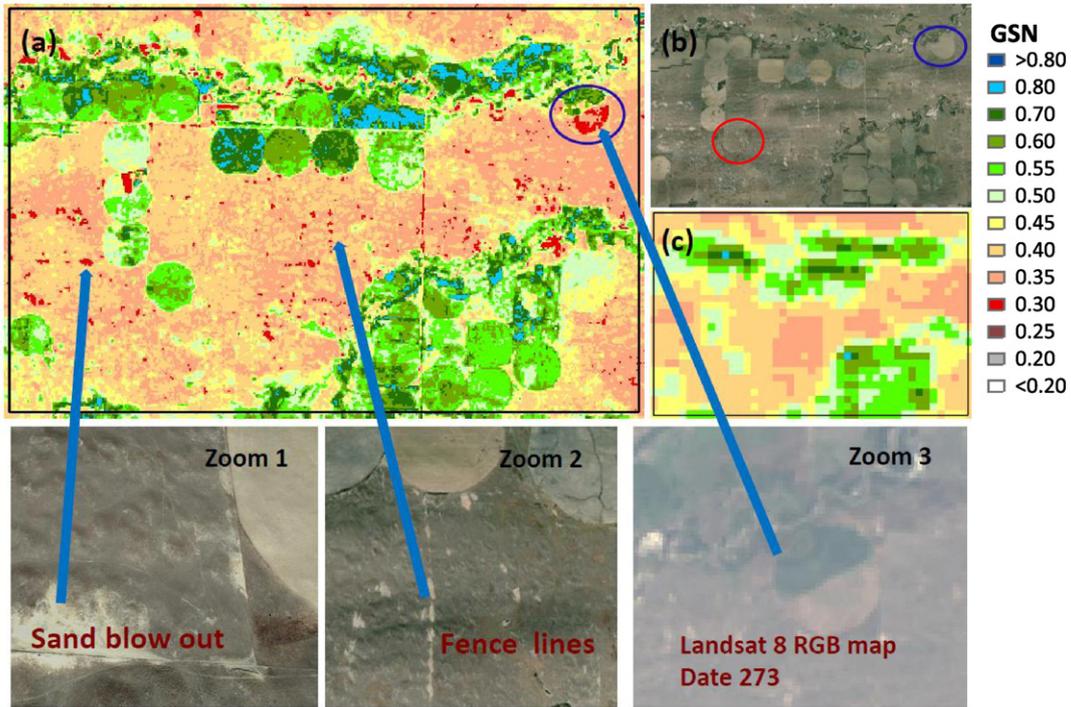


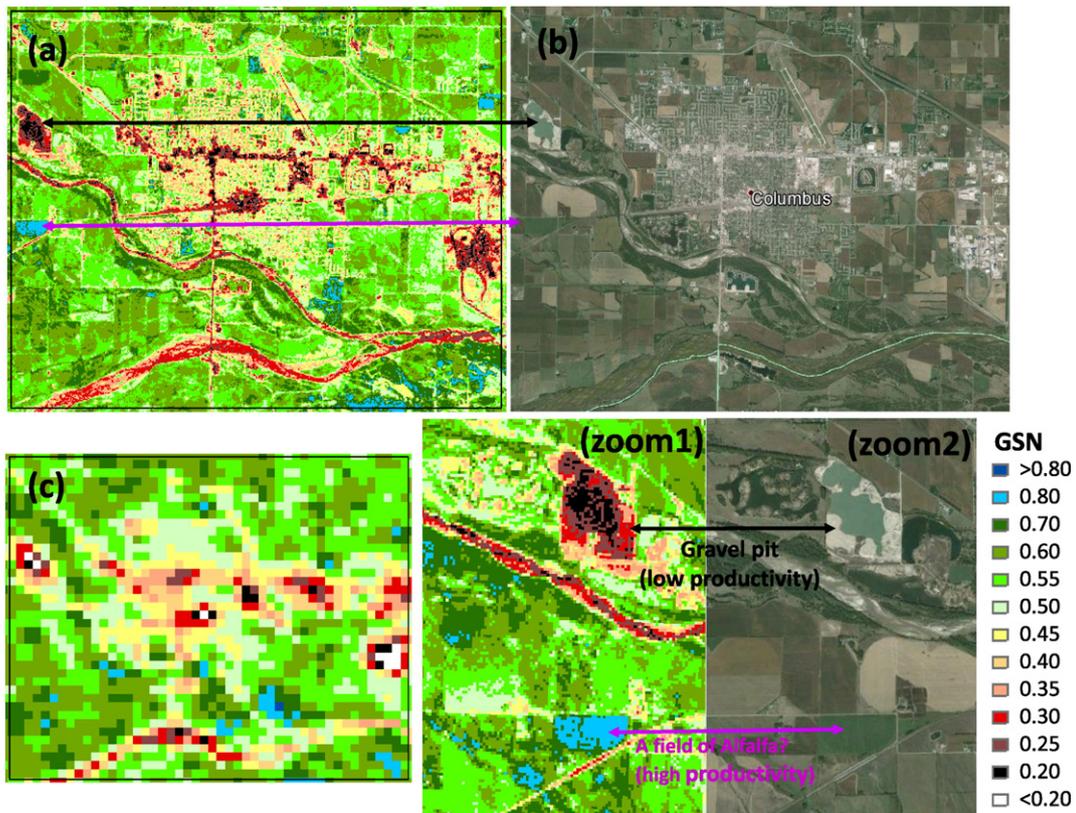
Fig. 3. (a) Scatterplots for the actual MODIS GSN and the predicted GSN for the 1491 test samples, (b) 30-m MODIS-Landsat based predicted GSN, and (c) 250-m MODIS actual GSN maps for the study area (includes all land covers).



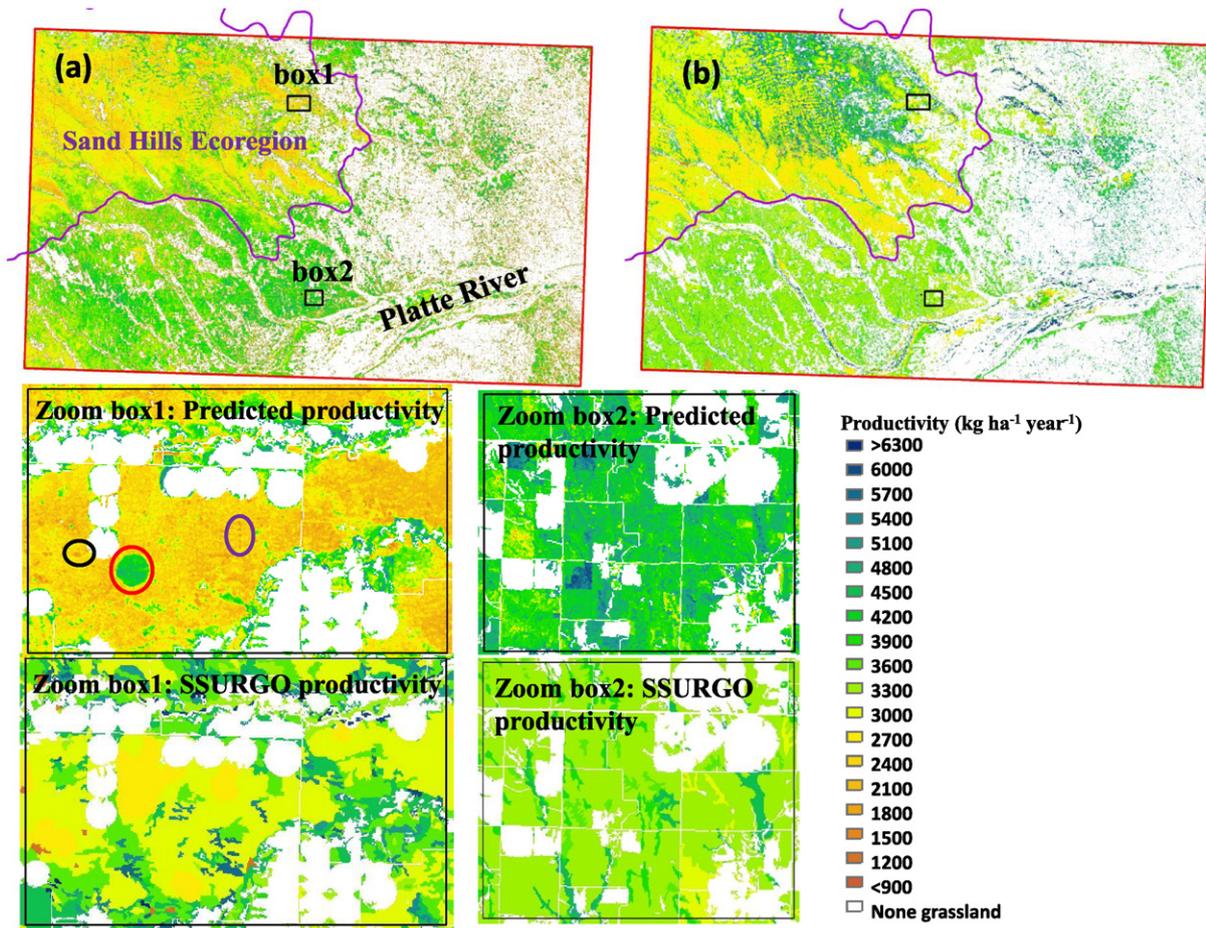
**Fig. 4.** A detailed view of a region in Holt County, Nebraska (Box 1 in Fig. 3b). (a) 30-m MODIS-Landsat predicted GSN for 2013, (b) Google Earth map with imagery dated March 11, 2012, and (c) 250-m MODIS actual GSN map for 2013. Zoom 1 and Zoom 2 are the maps from Google Earth. Zoom 3 is the Landsat 8 RGB map obtained from September 30, 2013 (Julian date 273), for the “dark blue oval” region.

over the growing season. For example, there is a very low productivity region located in the southern part of a center pivot system in the 30-

m GSN map (i.e., within the “dark blue oval” in Fig. 4a), implying that this region might be abandoned land or half harvested land (chopped



**Fig. 5.** A detailed view of Columbus, Nebraska (Box 2 in Fig. 3b). (a) 30-m MODIS-Landsat predicted GSN for 2013, (b) Google Earth map dated October 06, 2013, and (c) 250-m MODIS actual GSN map for 2013. Zoom 1 is the 30-m MODIS-Landsat predicted GSN map. Zoom 2 is the map from Google Earth.



**Fig. 6.** (a) 30-m predicted grassland biomass productivity map for 2013, and (b) 30-m SSURGO total productivity map for a normal year for central Nebraska. Zoom box 1 and box 2 in the lower panel are the 30-m predicted grassland productivity and the 30-m SSURGO biomass productivity close-up maps for the two selected areas.

for silage or hayed) (Fig. 4 Zoom 3, Landsat RGB image), but this phenomenon could not be captured by the single-time observation of the Google Earth map (Fig. 4b “dark blue oval”).

Fig. 5 is an example for a developed region around Columbus, Nebraska (Box 2 in Fig. 3b). The detailed ecological and environmental conditions for the Columbus region are clearly shown in the 30-m GSN estimation map and the Google Earth map (e.g., the downtown area, the roads, and the river) (Fig. 5a and b), but did not show up in the 250-m GSN map (blurry image in Fig. 5c). An extremely low GSN region located in the 30-m GSN map (Fig. 5a and Zoom 1, black arrow region) was identified as a gravel pit from the Google Earth map (Fig. 5b and Zoom 2, black arrow region). On the other hand, a very high GSN region located in the 30-m GSN map (Fig. 5a and Zoom 1, pink arrow region) was also identified from the Google Earth map (Fig. 5b and Zoom 2, pink arrow region), which might be a field of alfalfa.

In summary, the derived 2-Landsat path/row based MODIS-Landsat GSN model can successfully predict the 30-m GSN for central Nebraska. The 30-m MODIS-Landsat predicted GSN map is seamless for the two Landsat path/rows and provides detailed ecological and vegetation dynamic information of the study area, which will be useful for regional ecosystem study.

### 3.3. 30-m grassland biomass productivity estimation map for central Nebraska

Fig. 6a is the final estimated grassland biomass productivity map for central Nebraska. Large seamless spatial variations of productivity in the study area can be found in the figure. For example, productivity is low within the Sand Hills ecoregion because of sandy soils. On the other

hand, productivity is high along the Platte River because of the favorable vegetation growth conditions and possible intensive management (irrigation and fertilization) in that area (Fig. 6a). To illustrate the spatial variations of productivity in the study area more clearly, two small boxes representing low productivity (the same Box 1 in Figs. 3b and 6a) and high productivity (Box 2 in Fig. 6a) were selected. Box 1 is located within the Sand Hills ecoregion with low grassland productivity; productivities in most of this box were less than 3300 kg ha<sup>-1</sup> year<sup>-1</sup> (Fig. 6 Zoom box 1). The sand blow out area and the fence line area that were previously identified by the Google Earth map (Fig. 4 Zoom 1 and Zoom 2) can also be recognized in the 30-m productivity estimation map (i.e., very low productivity areas within the black circle and the purple oval regions, Fig. 6 Zoom box 1). The highly productive grassland area within the red circle in Fig. 6 Zoom box 1 is assumed to be an irrigated cropland region, which was not identified by the 2011 NLCD map but apparent in Google Earth (Fig. 4b red circle). This center pivot was either misclassified or was a new center pivot irrigation system. Box 2 is located along the Loup River with very high grassland productivity; productivities in most of this region were greater than 3300 kg ha<sup>-1</sup> year<sup>-1</sup> (Fig. 6 Zoom box 2). This 30-m grassland biomass productivity estimation map provides spatially detailed biomass productivity information for central Nebraska, which can be used as a reference for local land management practices.

### 3.4. Validation of the 30-m predicted grassland productivity map using SSURGO biomass production data

To confirm and validate the derived 30-m grassland productivity map (Fig. 6a), the 30-m SSURGO total biomass production map for a

normal year for central Nebraska (Fig. 6b) was used to evaluate the resulting map. The SSURGO total biomass production map is considered as the best available high spatial resolution ground observation based production map. SSURGO data were created county by county through collaboration with the states and the Natural Resources Conservation Service (NRCS), each county uses slightly different criteria for their soil surveys. One disadvantage of SSURGO data is that they have spatial discontinuities (e.g., differences across state and county lines) (Gu et al., 2013).

Fig. 6a–b indicates that the general spatial pattern of the 30-m predicted grassland productivity map is in good agreement with the SSURGO total biomass productivity map. Both productivity maps show relatively low productivities in the Sand Hills ecoregion because of the sandy soil condition. Productivity increases from west to east through the study area in the two production maps (Fig. 6a and b) because of the dry to wet climate conditions. The two zoom-in boxes also illustrate the good agreement in spatial variations of the two productivity maps and the detailed ecological features that the 30-m predicted productivity map provided (e.g., “Sand blow out” areas and the “Fence lines”).

Some differences, however, existed between the two productivity maps because of the different sampling time with the different weather and climate conditions. For example, the 30-m predicted grassland productivities in the Zoom box 1 area are lower than the 30-m SSURGO biomass productivities. The reason for this difference is because the SSURGO data were derived from a normal climate year with a normal productivity condition, whereas the 30-m predicted grassland productivity map was derived from a severe drought year (2013) for the box 1 region (<http://droughtmonitor.unl.edu/MapsAndData/MapArchive.aspx>) and thus lead to a low productivity in that region. One can also notice that the 30-m predicted grassland productivities are relatively higher than the 30-m SSURGO biomass productivities in some regions within the Zoom box 2, and may be due to the intensive agricultural management (e.g., irrigation) in those regions.

To further assess the agreement of the two productivity maps, the percentage differences between the 30-m predicted productivity and the SSURGO productivity was estimated (Fig. 7). The result shows that there is only a small portion (<15%) of the study area that has large productivity difference (>30% difference). The large negative difference regions are mainly located within the Sand Hills ecoregion (Fig. 7), and were mainly caused by the severe drought in 2013. Overall, the derived 30-m grassland productivity map generally agreed with (<30%

difference) the SSURGO biomass productivity map in most regions (>85%) of the study area, implying that the 30-m grassland productivity map was generally supported by the ground observations.

### 3.5. Discussion

In this study, Landsat 8 data were used to develop the 250-m MODIS-Landsat GSN model and predicted the GSN at a 30-m resolution. Landsat 8 data were used instead of Landsat 7 data in this study (1) to demonstrate the high quality of 12-bit Landsat 8 data, (2) to avoid the data gaps that were caused by the failure of the Scan Line Corrector (U.S. Geological Survey, 2012) in the Landsat 7 data, and (3) to ensure the high quality of the model training data. Landsat 8 provides an improved signal-to-noise performance that enables more accurate characterization of land cover state and condition (<http://landsat.usgs.gov/landsat8.php>) and thus helps to improve the MODIS-Landsat GSN model robustness and reliability. However, there was only one year (2013) of Landsat 8 data available for this investigation. In a future study, we plan to use multiple year, high-level Landsat 8 data to develop a long-term 30-m grassland biomass productivity data set for central Nebraska when more data are available and to advance automation of this downscaling technique.

Because all the training data used for developing the MODIS-Landsat GSN model were done at the 250-m resolution to match the MODIS data, some extreme values (e.g., extreme high or low values) in the 30-m Landsat data may be smoothed by 250-m model development. Therefore, the 250-m MODIS-Landsat GSN model may not capture the extreme cases when estimating the 30-m GSN. For example, some extreme low GSN areas ( $GSN < 0.3$ , red color in Fig. 3c) in the Sand Hills ecoregion in the 250-m MODIS GSN map did not show up in the 30-m predicted GSN map (Fig. 3b). On the other hand, some extreme high GSN areas ( $GSN > 0.8$ , light blue color in Fig. 3c) around the Platte river region did not show up in the 30-m predicted GSN map (Fig. 3b). Therefore, adding more model training data points for these extreme regions or applying an “extrapolation” approach for the GSN model to estimate the 30-m predicted GSN (Gu, Howard, Wylie, & Zhang, 2012) is proposed in future studies. An “extrapolation” approach is to set a relatively high extrapolation allowance in the GSN model, which will allow a wide range of the predicted GSN values (i.e., beyond the model development data range).

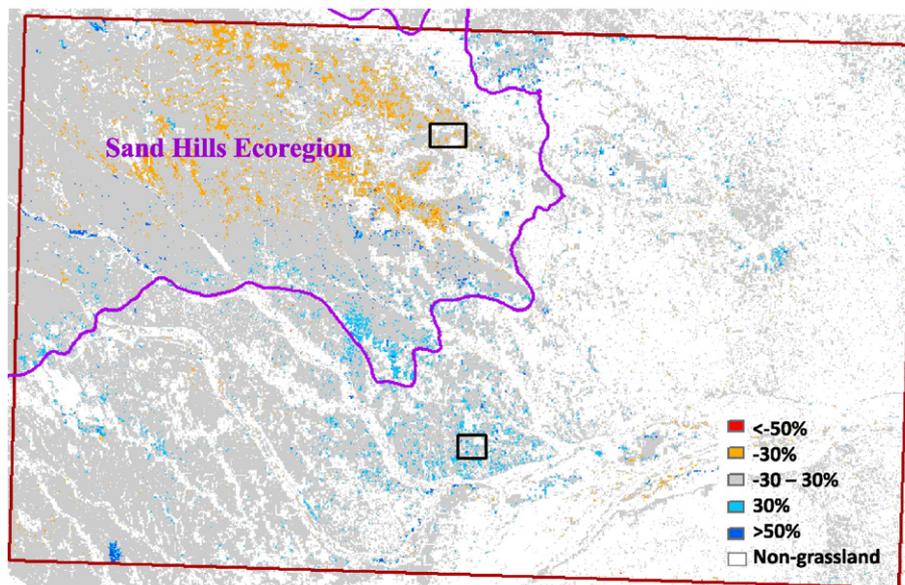


Fig. 7. Percentage difference (PD) map between the 30-m predicted grassland productivity (P) and the SSURGO biomass productivity (S).  $PD = 100 \times ((P - S) / S)$ .

Using the original 30-m Landsat NDVI data to calculate the 30-m GSN might be simple and easy; however, the derived GSN values may be much lower than the actual GSN values for the cloud-impacted pixels (Gu & Wylie, 2015). Based on the results from Gu and Wylie (2015), using the original Landsat NDVI data to calculate the GSN can cause large errors, our proposed GSN model can improve the correlation from the original Landsat scene. The advantage of the approach presented in this study is that it uses both original Landsat data and atmospherically corrected MODIS GSN data to develop the MODIS-Landsat GSN model, the derived GSN model captures the growing season long vegetation dynamics and handles the infrequent cloudy conditions well (Gu et al., 2015). In summary, this approach retains the advantages of the MODIS data (e.g., atmospherically corrected for cloud and aerosols, high temporal resolution that can capture the detailed seasonal dynamic information) but leverages the high-resolution spatial context from Landsat, which provides an effective method for downscaling the MODIS GSN data.

Moreover, collecting ground observation data to further validate the derived 30-m MODIS-Landsat productivity map is needed for future studies. Examples of ground observations include identifying the specific land cover types based on ground observations (e.g., NASS Cropland Data Layer data or the local land use records) to explain the high productive areas (e.g., pink arrow area in Fig. 5) and land cover change (i.e., red circle area in Fig. 6 Zoom box 1 map); investigating the cause of the half harvested region in Fig. 4a and Zoom 3 map (e.g., is it an abandoned land?) based on ground observations (e.g., interview the land owner about their crop histories); and further verifying the grassland productivity based on NASS or other ground record data.

#### 4. Conclusion

This study developed a 30-m MODIS-Landsat GSN map and estimated grassland biomass productivity for the two Landsat path/rows located in central Nebraska. The derived 2-Landsat path/row rule-based piecewise regression GSN model can successfully predict the GSN ( $r = 0.91$  and the mean absolute error = 0.026). The resulting seamless 30-m grassland biomass productivity estimation map, which provided spatially detailed ecological features and conditions of central Nebraska, was generally supported by the SSURGO biomass production map and will be useful for regional ecosystem study and local land management practices.

#### Acknowledgments

This work was performed under USGS contract G13PC00028 and funded by the USGS Land Change Science Program in support of Renewable Energy-Biofuels. The authors thank Lei Ji, Thomas Adamson, Sandra C. Cooper, and three anonymous reviewers for their valuable suggestions and comments. The authors thank Norman Bliss for processing SSURGO data. Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

#### References

Anderson, J.R., Hardy, E.E., Roach, J.T., & Witmer, R.E. (1976). *A land use and land cover classification system for use with remote sensor data*. (p. 28) Reston: U. S. Geological Survey.

- 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, 1312–1323.
- Curran, P.J., & Williamson, H.D. (1985). The accuracy of ground data used in remote-sensing investigations. *International Journal of Remote Sensing*, 6, 1637–1651.
- Curran, P.J., & Williamson, H.D. (1986). Sample size for ground and remotely sensed data. *Remote Sensing of Environment*, 20, 31–41.
- Funk, C., & Budde, M.E. (2009). Phenologically-tuned MODIS NDVI-based production anomaly estimates for Zimbabwe. *Remote Sensing of Environment*, 113, 115–125.
- Gu, Y., Brown, J.F., Verdin, J.P., & Wardlow, B. (2007). A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophysical Research Letters*, 34.
- Gu, Y., Howard, D., Wylie, B., & Zhang, L. (2012). Mapping carbon flux uncertainty and selecting optimal locations for future flux towers in the Great Plains. *Landscape Ecology*, 27, 319–326.
- Gu, Y., & Wylie, B.K. (2015). Downscaling 250-m MODIS growing season NDVI based on multiple-date Landsat images and data mining approaches. *Remote Sensing*, 7, 3489–3506.
- Gu, Y., & Wylie, B.K. (2010). Detecting ecosystem performance anomalies for land management in the upper Colorado river basin using satellite observations, climate data, and ecosystem models. *Remote Sensing*, 2(8), 1880–1891.
- Gu, Y., Wylie, B.K., & Bliss, N.B. (2013a). Mapping grassland productivity with 250-m eMODIS NDVI and SSURGO database over the Greater Platte River Basin, USA. *Ecological Indicators*, 24, 31–36.
- Gu, Y., Wylie, B.K., & Howard, D.M. (2015). Estimating switchgrass productivity in the Great Plains using satellite vegetation index and site environmental variables. *Ecological Indicators*, 48, 472–476.
- Gu, Y., Wylie, B.K., Howard, D.M., Phuyal, K.P., & Ji, L. (2013b). NDVI saturation adjustment: A new approach for improving cropland performance estimates in the Greater Platte River Basin, USA. *Ecological Indicators*, 30, 1–6.
- Han, W., Yang, Z., Di, L., & Mueller, R. (2012). CropScape: A web service based application for exploring and disseminating US conterminous geospatial cropland data products for decision support. *Computers and Electronics in Agriculture*, 84, 111–123.
- Hobbs, T.J. (1995). The use of NOAA-AVHRR NDVI data to assess herbage production in the arid rangelands of central Australia. *International Journal of Remote Sensing*, 16, 1289–1302.
- Jenkinson, C.B., Maiersperger, T.K., & Schmidt, G.L. (2010). eMODIS — A user-friendly data source. *U.S. Geological Survey Open-File Report*. 1055. (pp. 10).
- Ji, L., Wylie, B.K., Nossov, D.R., Peterson, B.E., Waldrop, M.P., McFarland, J.W., ... Hollingsworth, T.N. (2012). Estimating aboveground biomass in interior Alaska with Landsat data and field measurements. *International Journal of Applied Earth Observation and Geoinformation*, 18.
- Jin, S., Yang, L., Danielson, P., Homer, C., Fry, J., & Xian, G. (2013). A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment*, 132, 159–175.
- Potter, C.S., Randerson, J.T., Field, C.B., Matson, P.A., Vitousek, P.M., Mooney, H.A., & Klooster, S.A. (1993). Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles*, 7, 811–841.
- Reed, B.C., Brown, J.F., Vanderzee, D., Loveland, T.R., Merchant, J.W., & Ohlen, D.O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5, 703–714.
- Swets, D.L., Reed, B.C., Rowland, J.R., & Marko, S.E. (1999). A weighted least-squares approach to temporal smoothing of NDVI. *ASPRS Annual Conference, From Image to Information*. Portland, Ore.: Am. Soc. for Photogram. and Remote Sens.
- Tieszen, L.L., Reed, B.C., Bliss, N.B., Wylie, B.K., & DeJong, D.D. (1997). NDVI, C<sub>3</sub> AND C<sub>4</sub> production, and distributions in Great Plains grassland land cover classes. *Ecological Applications*, 7, 59–78.
- Tucker, C.J., Vanpraet, C.L., Sharman, M.J., & Van Ittersum, G. (1985). Satellite remote sensing of total herbaceous biomass production in the Senegalese Sahel: 1980–1984. *Remote Sensing of Environment*, 17, 233–249.
- Wang, J., Rich, P.M., Price, K.P., & Kettle, W.D. (2004). Relations between NDVI and tree productivity in the central Great Plains. *International Journal of Remote Sensing*, 25, 3127–3138.
- Wylie, B.K., Denda, I., Pieper, R.D., Harrington, J.A., Reed, B.C., & Southward, G.M. (1995). Satellite-based herbaceous biomass estimates in the pastoral zone of Niger. *Journal of Range Management*, 48, 159–164.
- Wylie, B.K., Zhang, L., Bliss, N.B., Ji, L., Tieszen, L.L., & Jolly, W.M. (2008). Integrating modeling and remote sensing to identify ecosystem performance anomalies in the boreal forest, Yukon River Basin, Alaska. *International Journal of Digital Earth*, 1, 196–220.