Monitoring Drought Impact on Annual Forage Production in Semi-arid Grasslands: A Case Study of Nebraska Sandhills

Markéta Podebradská 1,2,*, Bruce K. Wylie 3, Michael J. Hayes 1, Brian D. Wardlow 1,4, Deborah J. Bathke 1,2, Norman B. Bliss 3 and Devendra Dahal 5

1 School of Natural Resources, University of Nebraska-Lincoln, 3310 Holdrege Street, Lincoln, NE 68583-0988, USA
2 National Drought Mitigation Center, School of Natural Resources, University of Nebraska-Lincoln, 3310 Holdrege Street, P.O. Box 830988, Lincoln, NE 68583-0988, USA
3 U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, 47914 252nd Street, Sioux Falls, SD 57198, USA
4 Center for Advanced Land Management Information Technologies, School of Natural Resources, University of Nebraska-Lincoln, 3310 Holdrege Street, Lincoln, NE 68583-0988, USA
5 KBR, contractor to the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center, 47914 252nd Street, Sioux Falls, SD 57198, USA

* Correspondence: podebradska.marketa@huskers.unl.edu

Received: 1 August 2019; Accepted: 4 September 2019; Published: 9 September 2019

Abstract: Land management practices and disturbances (e.g., overgrazing, fire) have substantial effects on grassland forage production. When using satellite remote sensing to monitor climate impacts, such as drought stress on annual forage production, minimizing land management practices and disturbance effects sends a clear climate signal to the productivity data. This study investigates the effect of this climate signal by: (1) providing spatial estimates of expected biomass under specific climate conditions, (2) determining which drought indices explain the majority of interannual variability in this biomass, and (3) developing a predictive model that estimates the annual biomass early in the growing season. To address objective 1, this study uses an established methodology to determine Expected Ecosystem Performance (EEP) in the Nebraska Sandhills, US, representing annual forage levels after accounting for non-climatic influences. Moderate Resolution Imaging Spectroradiometer (MODIS)-based Normalized Difference Vegetation Index (NDVI) data were used to approximate actual ecosystem performance. Seventeen years (2000–2016) of annual EEP was calculated using piecewise regression tree models of site potential and climate data. Expected biomass (EB), EEP converted to biomass in kg ha⁻¹ yr⁻¹, was then used to examine the predictive capacity of several drought indices and the onset date of the growing season. Subsets of these indices were used to monitor and predict annual expected grassland biomass. Independent field-based biomass production data available from two Sandhills locations were used for validation of the EEP model. The EB was related to field-based biomass production ($R^2 = 0.66$ and $0.57$) and regional rangeland productivity statistics of the Soil Survey Geographic Database (SSURGO) dataset. The Evaporative Stress Index (ESI), the 3- and 6-month Standardized Precipitation Index (SPI), and the U.S. Drought Monitor (USDM), which represented moisture conditions during May, June and July, explained the majority of the interannual biomass variability in this grassland system (three-month ESI explained roughly 72% of the interannual biomass variability). A new model was developed to use drought indices from early in the growing season to predict the total EB for the whole growing season. This unique approach considers only climate-related drought signal on productivity. The capability to estimate annual EB by the end of May will potentially enable land managers to make informed decisions about stocking rates, hay purchase needs, and other management issues early in the season, minimizing their potential drought losses.
Keywords: rangeland productivity; remote sensing; eMODIS NDVI; rule-based piecewise regression tree; drought; forage production; biomass prediction; Sandhills

1. Introduction

According to the 2019 Global Risks Perception Survey [1], the top three “global risks” of highest concern in terms of likelihood are: extreme weather events, failure of climate-change mitigation and adaptation, and natural disasters. These three concerns, together with water crises, also appear in the top five risks in terms of impact on multiple countries and sectors [1]. Drought, characterized as a natural departure from expected water availability [2], is a major factor in each of these risk categories of concern. Over the past 20 years, the National Centers for Environmental Information (NCEI) identified 17 droughts in the United States that have each resulted in more than one billion-dollar loss per event. These losses often originate in the agricultural sector [3]. This paper presents a strategy to use remote sensing to quantify the impacts of drought on forage biomass and to predict total growing season biomass from drought indices measured early in the growing season.

With 92% of the total land area used for farming and ranching, Nebraska’s agricultural production annually contributes approximately $25 billion to the state’s economy, accounting for more than one-fifth of the state’s gross domestic product. Cattle are the state’s leading commodity in terms of value. Beef production is the largest agricultural sector and generates approximately $7.2 billion in annual cash receipts [4]. This makes Nebraska one of the largest beef producers in the United States and the world. Roughly 45% of Nebraska’s total area is classified as rangeland and pastures, half of which are located in the Sandhills region (approximately 12 million acres) that covers much of the north-central part of the state [5]. Beef production in this region primarily relies upon grassland forage, which makes grassland productivity of the Sandhills important to Nebraska’s economy. Harsh environmental conditions, such as drought, can have direct consequences for ranching operations, including reduced forage yield, livestock losses, and water quality and quantity issues. To replace forage lost to drought, producers often purchase commercial feed that increases costs for ranching operations. These impacts potentially lead to other indirect issues like physical and emotional stress [6] or financial and social instability [7], often characterized as secondary drought impacts.

Large portions of Nebraska have experienced an extreme or exceptional drought four times over the past 15 years, with nearly 95% of Nebraska’s pastures and rangeland in “poor” or “very poor” conditions during the severe 2012 drought [8,9]. Drought is an event that every land manager must consider and try to prepare for, especially in regions with lower water-holding capacity like the Sandhills, where interannual changes in precipitation greatly influence forage production [10–12]. Drought planning can help manage these climate shocks. According to the concept of the three pillars of drought risk management [13], the key areas that should be addressed are: (1) monitoring and early warning, which identifies drought status in a timely fashion; (2) vulnerability and impact assessment, determining the location and what is at risk of drought and why; and (3) mitigation and response, describing actions and measures needed to mitigate drought impacts and respond to drought emergencies. Our research contributes to the establishment of connections between these three pillars by describing links between drought indices and drought impacts, with the latter quantified as expected changes in biomass and forage.

Monitoring and early warning are important components of drought planning strategies. Over the past 20 years, the accuracy and precision of drought monitoring has improved due to technological advancements in meteorological instrumentation and the ability to archive, analyze, and disseminate the available data [14]. However, in-situ observations are spatially and temporally limited, and are unable to meet the increasing monitoring demand, especially on large spatial scales. Compared to ground-based observations, satellites provide a solution with global, near-real-time observations over large areas, consistent data records, and improved spatial resolution [15]. Over the course of the last few decades, a suite of indices and tools measuring the severity and extent of a drought has been developed. Some of these are based on in-situ measurements of various parts of
the hydrologic cycle, while others are a blend of input variables originating from different data sources, including satellite remote sensing.

Time-consuming, ground-based monitoring of range conditions using visual estimates or destructive sampling methods is often performed on a small scale. Applications that utilize modeling approaches to monitor and estimate annual range conditions on a large scale are being developed [16,17]. So far, however, the spatial resolution or the meaning (e.g., quantity of forage) of the provided range condition information is not appropriate for use by individual land managers. Information on location-specific productivity is important, especially in hydrologically complex and land cover diverse areas such as the Sandhills, where relatively high and low productivity can be found on a small spatial extent (see more information about upland areas and wet meadows in the Study Area section). Therefore, enhancing the monitoring of range conditions on a landscape scale with sufficient spatial resolution, while providing meaningful and simple information for range management decisions, is needed to more effectively mitigate and respond to drought.

Satellite remote sensing techniques and applications have been widely used for landscape assessment of vegetation health, as well as estimation of vegetation biomass and grass forage [18–22]. However, changes in vegetation health and the amount of produced biomass can be difficult to interpret because the remotely sensed signal does not capture information about the specific causes of these changes [23]. When studying the impact of drought on biomass production, it is important to separate the influence of climate from other factors. While weather refers to short-term variations (i.e., minutes to days) in variables such as temperature, precipitation, and wind; climate describes conditions over larger regions and for extended periods of time. Climate extremes, such as drought, are characterized by the deviation of climate statistics from what is expected over a given period of time (e.g., a month, season, or year). Indices are frequently calculated to describe, monitor, and project the state of the climate.

The goal of this study is to establish a relationship between annual biomass production and climate conditions, particularly those related to drought that occurs before and during the growing season to identify the climate variables and indices that explain the majority of interannual biomass variability. Biomass production depends on various site characteristics, such as soil properties, topography, long-term climate, and historic land use and management [24]. The interannual variation in biomass production is mainly influenced by climate conditions that occur before or during the growing season and by disturbances such as fire, grazing, or management [25]. Therefore, when studying the impact of drought on biomass production, it is important to use a methodology that separates the influence of climate from other factors.

The Normalized Difference Vegetation Index (NDVI) has been widely used for the assessment of ecosystem performance. NDVI values averaged over the growing season (GSN) have proven to have a strong relationship with ground-based observations of biomass productivity [26–28]. We separate the influence of seasonal climate on the amount of total seasonal biomass from other factors with a previously established methodology [29] that uses historic remotely sensed NDVI, climate data, and regression tree modeling techniques. The product of the model is the Expected Ecosystem Performance (EEP) that represents the annual biomass expected to be produced at a specific site under certain climate conditions, without the influence of other factors.

The first (1) objective of this study is to develop an annual EEP model and corresponding maps for the years 2000 to 2016 at 250 meter (m) resolution for the upland grasslands of the Sandhills Major Land Resource Area (MLRA), which represents roughly half of the rangeland and pastures in Nebraska. These maps serve the purpose of isolating the effect of climate on growing season total forage production. This is achieved by using model inputs that are related only to climate and site characteristics. The output of the model also provides information about the overall importance of the different explanatory input variables. The second (2) objective is to examine the relationship between various drought indices, the timing of drought, and biomass production derived from the model. This objective uniquely combines ecosystem performance methodology with the use of various drought indices and tools that are based on diverse sources of information. We identify the indices and time frames that explain the majority of interannual seasonal biomass variability. The
third (3) objective is to create a piecewise regression tree model that uses drought indices, summarizing the moisture conditions before and during the early stages of the growing season to predict annual forage production. The result of this study will improve the knowledge available to land managers for more informed decision making during the early stages of the growing season.

2. Background, Materials and Methods

2.1. Study Area

The Sandhills MLRA is located almost entirely in Nebraska (98%) with a small area in South Dakota. Only the area located in Nebraska was considered for this study. The region is moderately vulnerable to drought because of its sandy soil type, land type (ranching), and seasonal crop moisture deficiency probabilities [30]. The entire region covers over 53,000 km². It is one of the largest, grass-stabilized dune regions in the world, and the largest in the Western Hemisphere. The average annual temperature ranges from 8 °C to 10 °C, with an average of 155 freeze-free days. The average low temperature in January is −9.6 °C and the average high temperature in July is 29.9 °C (measured in Whitman, central part of the Sandhills). The average annual precipitation ranges from 660 mm in the east to 380 mm in the west. The majority of the precipitation occurs from April to October and originates either from frontal storms or convective thunderstorms. The elevation gradually decreases from west (1200 meters above sea level [masl]) to east (600 masl). The dominant soil orders are Entisols and Mollisols, which are generally sandy, very deep, and excessively to somewhat poorly drained, depending on the location [31].

Sandhills grasslands are characterized by a matrix of steep, irregular sand dunes stabilized by grass vegetation, and narrow, elongated valleys between the dunes. The dunes and valleys are oriented in a northwest-southeast direction and commonly extend for several miles. Sharp contrasts in topography influence the hydrological variability and subsequently the vegetation communities [32,33]. The main vegetation communities are mid and tall grasses that differ in composition, based on a geographic and topographic location. The major grass species of sand dune uplands, mostly warm-season grasses, are: Little bluestem (*Schizachyrium scoparium*), sand bluestem (*Andropogon hallii*), prairie sandreed (*Calamovilfa longifolia*), switchgrass (*Panicum virgatum*), Indiangrass (*Sorghastrum nutans*), sand lovegrass (*Eragrostis trichodes*), and needle-and-thread grass (*Hesperostipa comata*) [30]. Upland areas and dry valleys, which are locations of groundwater recharge, are more affected by variations in climate, and their productivity is dependent upon the amount and timing of precipitation. Wet meadows are located in topographic depressions, where the depth to groundwater is shallow (approximately 1 m depending on the groundwater levels), where groundwater discharges, and are characterized by a different grass species composition—mostly cool season grasses, sedges, and rushes. The area is hydrologically complex. Understanding the hydrologic regimes remains challenging due to the local and regional geomorphologic differences and variability in location, duration, and timing of precipitation episodes [34]. Due to sharp contrasts in topography and major biophysical differences between uplands and meadows, data with adequate spatial resolution are required for the analysis.

2.2. Drought Indices

Drought indices are used to identify and quantify drought in terms of intensity, duration, and spatial extent [35]. They can be based on a single variable like precipitation, temperature, evaporation or soil moisture or on a composite of several variables (e.g. the U.S. Drought Monitor—USDM, Vegetation Drought Response Index—VegDRI). These variables can be measured either in-situ at meteorological stations or remotely estimated from satellite observations and models.

Drought indices such as the Standardized Precipitation Index (SPI) [36], Evaporative Stress Index (ESI) [37], and Evaporative Demand Drought Index (EDDI) [38] can be computed for a variety of timescales. Knowledge of the timescale is important for the interpretation of the index value and the impacts that might be connected with it. These timescales range from one week to more than a year. The shorter timescales can be applied in relation with short-term soil moisture or crop stress
during the growing season. The 3- to 6-month index represents seasonal moisture conditions compared to historical climate normals. This timescale is useful for monitoring impacts on agriculture and vegetation, capturing moisture deficits in critical phenological stages of plant development. Longer timescales (12 months or more) are tied to hydrological impacts (i.e., changes in stream flows, reservoir, and groundwater levels [39]) and ecological impacts on a landscape level [40]. Indices and tools used in this analysis are SPI, ESI, EDDI, VegDRI, and USDM. Table 1 provides a brief description of each of these datasets as well as the spatial resolution, time step(s), and justification for their use in this study.

**Table 1.** Drought indices and tools used in the correlation analysis and in the predictive model. This table provides a simple description of each index, its spatial resolution and timescales used in our analysis, and a justification of why this index was selected for the analysis.

<table>
<thead>
<tr>
<th>Index</th>
<th>Description</th>
<th>Spatial resolution</th>
<th>Timescales</th>
<th>Justification</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPI</td>
<td>Measures the rarity of a dry or wet event based on probability</td>
<td>Gridded at 0.125°</td>
<td>1, 3, 6 months</td>
<td>Recommended by World Meteorological Organization as a global measure of meteorological drought.</td>
<td>[36,41–43]</td>
</tr>
<tr>
<td>ESI</td>
<td>Quantifies anomalies of ET to PET*</td>
<td>5 km</td>
<td>4, 12 weeks</td>
<td>Provides an indication of water stress on vegetation. Not limited by cloud or snow cover; expected to detect drought earlier than other indices.</td>
<td>[37,44,45]</td>
</tr>
<tr>
<td>EDDI</td>
<td>Measures anomalies of E0* to reference evapotranspiration</td>
<td>Gridded at 0.125°</td>
<td>1, 3 months</td>
<td>Indicates plant stress specifically due to drought.</td>
<td>[38,46,47]</td>
</tr>
<tr>
<td>VegDRI</td>
<td>Integrates climate, vegetation stress, and biophysical data</td>
<td>1 km</td>
<td>Weekly</td>
<td>Comprehensive indicator of drought; informs USDA drought relief payments.</td>
<td>[14]</td>
</tr>
<tr>
<td>USDM</td>
<td>Synthesizes quantitative drought data with qualitative local expertise</td>
<td>Sub-county vector data converted to raster at 500 m</td>
<td>Weekly</td>
<td></td>
<td>[48,49]</td>
</tr>
</tbody>
</table>

*Actual evaporation (ET), potential evaporation (PET), atmospheric evaporative demand (E0)*

The USDM map of drought conditions differs from traditional datasets and consequently warrants additional explanation. The map is produced weekly, combining objective drought indicators with subjective expertise. The map classifies drought into severity categories using a percentile approach [48]. Our analysis required seasonally integrated continuous representation of the USDM. To accomplish this, we created an integrated percentile index for a variety of time intervals (winter, December-February; spring, March-May; summer, June-August; fall, September-November; growing season, April-October; and annual). Each weekly USDM value was assigned an approximate percentile ranking (“None” = 50, D0 = 25, D1 = 15, D2 = 8, D3 = 4, D4 = 1) and integrated over the course of the specific season.

2.3. PRISM Weather Dataset
Weather variables (precipitation; minimum, maximum, and mean temperature) for the years 1999 to 2016 were obtained from PRISM Climate Group [50]. The data were acquired with 4 km resolution and bilinearly interpolated to achieve a 250-m resolution. Monthly data for each variable were compiled into seasons (winter, December-February; spring, March-May; summer, June-August; fall, September-November). Precipitation was integrated over the course of each season, while temperature variables were represented by the mean of each season. We ensured that climate variables selected for the model were biologically meaningful by excluding conditions that occurred after the end of the growing season (winter) or by excluding those that do not have a large influence on annual biomass production (fall—end of the growing season).

2.4. Site Potential

Site potential represents a long-term GSN value at a certain location under “good” conditions for vegetation growth, a measure of the land’s inherent productivity. It accounts for long-term spatial variation in growing conditions. These conditions can be primarily affected by soils, vegetation type, long-term climate, or topography. To approximate the site potential, we used a long historical time series of GSN and eliminated years when the growth was reduced to represent the site’s potential to produce biomass under good conditions [51]. We define site potential as the mean above the long-term (2000-2016) GSN median. In other words, it is an average of all GSN values from years 2000 to 2016 that are higher than a GSN median for each pixel in the same time period. Therefore, site potential estimates moderately high to high potential productivity for each pixel. We employed this method to dramatically reduce the potential spatial artifacts that are often present when other methods of approximation are used.

2.5. Suitable Areas for Training Pixels

Due to the different hydrological characteristics and responses to climate variability, wet meadows and upland areas need to be treated separately in an ecosystem performance model. Past research emphasizes the influence of climate on semi-arid grasslands [25,52,53]. Therefore, the EEP model was developed only for the upland areas that represent this type of grassland and a majority of the land cover in the Sandhills. To distinguish between the two types of grassland ecosystems growing either in the upland or wet meadows (both included in the same National Land Cover Database [54] category), the Nebraska GAP land cover [55] product with 30-m resolution was used (upland = “Sandhills upland prairie”, meadows = “Lowland tallgrass prairie”). Even though the 250-m resolution of the NDVI data that we used for the analysis is adequate to retain the upland areas of interest [56], there might still be an influence of mixed pixel signal. We wanted to train the EEP model only on pure upland areas. Therefore, we calculated the percent of 250-m grid areas that were covered by the “Sandhills upland prairie”. Only areas covered with 100% of the desired land cover were identified as suitable for placement of model training points. However, less-restrictive criteria were used for final mapping. Areas used for mapping had to be at least 75% upland prairie and less than 10% water.

2.6. Data Analysis

A modelling technique for ecosystem performance using a rule-based piecewise regression approach was originally developed for a boreal forest ecosystem [29]. Since then, the approach has been used in multiple other studies focusing on grassland ecosystems [51,57,58]. The technique provides precise modelling of complex systems and produces an outcome that enhances the understanding of relationships between dependent and independent variables [29]. Cubist® software [59] was used to perform the rule-based piecewise regression model. This modeling software divides the input data space into many small multi-dimensional zones (the stratification) and fits a linear regression model to predict the dependent variable from some (or all) of the independent variables within that zone. Each zone is represented by a “rule” that includes both the definition of the zone and the regression equation. It is useful to evaluate which variables are
“important” by counting how many times a given variable is used in a rule for stratification, and how many times it is used in the regression equations for prediction. We use the average of these counts as a measure of importance for each variable.

We used two different models to address our objectives. The EEP model was designed to create the EEP and expected biomass (EB) time series of annual maps. The predictive model assessed which combination of existing drought products best predict variations in grass annual biomass as quantified by the EEP time series early in the growing season (Figure 1). The EB data were validated using long-term ground observations at two experimental sites located in the eastern and central Sandhills and also using the Soil Survey Geographic Database (SSURGO) range productivity dataset [60]. Additionally, we performed a regression analysis between annual biomass deviation and drought indices at various temporal scales for each week of the growing season.

Figure 1. Flowchart of the modeling process and data used to generate the Expected Ecosystem Performance (EEP), interannual biomass deviation (EEP model), and predictions of expected biomass at different time steps (predictive model).

The modelling process of the EEP model consists of three major steps: (1) Determining an actual ecosystem performance that is represented by GSN, (2) determining the EEP in terms of annual EB based on annual site conditions and climate inputs [26], and (3) calculating interannual biomass deviations that represent a deviation of annual EB from the long-term EB median.

2.6.1. Actual Ecosystem Performance (AEP)

The AEP represents growing season vegetation dynamics that can be approximated using the NDVI [18,56]. For the purpose of this research, we used NDVI with a 250-m resolution derived from the Expedited Moderate Resolution Imaging Spectroradiometer (eMODIS) [61]. The 7-day NDVI composites from eMODIS Terra collection 5 NDVI (2000-2002) and eMODIS Aqua collection 6 NDVI...
(2003-2016) have been quality masked, temporally smoothed [62] to remove temporary dips in NDVI associated with cloud contamination, and scaled to 0–200. We define these modified NDVI values as mNDVI. All mNDVI data have been obtained from the US Geological Survey Earth Resources Observation and Science Center (USGS EROS) eMODIS data archive. The GSN was calculated for each year using time series of 7-day mNDVI composites over the course of the growing season (week 14 to 44), which remained the same for the 17 years of the analysis. The growing season span was aligned with growing season in Gu et al. [26], which allowed the conversion of GSN to grass biomass production in kg ha⁻¹ yr⁻¹. The AEP serves as the dependent variable to train and evaluate the EEP model. EB values, described more completely below, approximate climate-driven variations in AEP.

2.6.2. EEP, EB, and Interannual Biomass Deviation

The AEP serves as the predicted (dependent) variable, and the site potential and seasonally averaged climate variables are included as the explanatory (independent) variables in the EEP model. No variables that would explain variability caused by management or disturbances are included in the model. Thus, the EEP is solely based on climate variations. More information about the EEP model can be found in Wylie et al. [29]. We trained the model on a set of randomly placed points (n = 7650) in areas that were identified as suitable for training and stratified based on annual GSN (low, medium, high). Points were placed randomly across the entire geographic area of the Sandhills region and the 17-year time span to represent spatial and temporal variability of the productivity in the model. To minimize the over- and under-fitting tendencies of the model, the number of points, rules, and other parameters of the model was based on the recommendations of Gu et al. [63]. The output of the model was then used for developing maps of the EEP using MapCubist [59], an internal USGS EROS code for applying Cubist models with spatial inputs to output predicted maps. In our case, these spatial inputs were the seasonal climate for each year and the site potential. The EEP maps were converted into total biomass productivity using the empirical equation developed by Gu et al. [26] for the Greater Platte River Basin:

$$\text{Biomass productivity} (\text{kg ha}^{-1}\text{yr}^{-1}) = 9936.5 \times \frac{\text{EEP} - 100}{100} - 1554$$

We defined interannual biomass deviation as the EB (EEP converted to biomass) in a given year minus a long-term (17 years) biomass productivity median for any given pixel. We calculated the mean absolute error (MAE) for these data and created maps of significant anomalies in each year when significant anomaly is defined by the overall value of MAE (median +/- MAE). We used these statistical metrics because they do not assume normality of the data and the median is less affected by outliers compared to the mean.

2.6.3. Validation of the EEP Model

We obtained long-term ground clipping data from two locations—Barta Brothers Ranch, located in the eastern part of the Nebraska Sandhills; and Gudmundsen Sandhills Laboratory, located in the central Sandhills. These data were collected by clipping 0.25 m² quadrats from pastures that had been grazed the previous year and exclusion cages that were used to prevent grazing of the clipped quadrat areas. Clipping was done in mid-August, which was considered peak production for warm-season grasses for that year, as plants were generally mature at that time. For more information on the data collection methodology, see Stephenson et al. [11]. We spatially matched the plots where the clipping data were obtained with our EB maps. We compared the annual biomass production retrieved from the model and the ground clipping data.

We also compared the characteristics of the modelled biomass distribution (minimum, maximum, mean, median, 1st and 3rd quantiles, and 5% and 95% percentile values) to the distribution of values from the SSURGO productivity data derived from the same area. In other words, the regional statistics for modelled biomass and SSURGO data were compared. SSURGO productivity data were obtained from the U.S. Department of Agriculture (USDA) Natural Resources Conservation Service (NRCS) soil survey database and gaps in those data were filled with the State Soil Geographic (STATSGO) data (compiled at 1:250,000 map scale) [64]. We used the representative
value of SSURGO (“rsprod_r” attribute), which estimates the annual potential production for rangeland.

2.6.4. Correlation Analysis of Drought Indices

A correlation analysis was performed between weekly drought indices (SPI, ESI, EDDI, and VegDRI) at several time scales (1, 3, and 6 months) and interannual biomass deviation was derived from Model 1 for years 2000 to 2016 for SPI and EDDI, 2001 to 2016 for ESI, and 2009 to 2016 for VegDRI. The number of years in the analysis depends on the availability of the data. Values of drought indices and biomass anomaly were extracted from 100 randomly placed points within the Sandhills upland grasslands for each year of the analysis. Only months of the growing season were used for the analysis (the beginning of April through the end of October). Correlation coefficients were calculated for each week and each drought index separately by correlating the EEP anomalies with the corresponding drought index values.

Seasonal USDM percentile index values were used in this analysis. A correlation analysis was performed between values of USDM index for each time step and productivity anomaly. The USDM captures both short- and long-term drought conditions. However, lingering long-term drought conditions identified by the USDM might not be representative of grassland productivity conditions. Therefore, we excluded two years (2004 and 2013) from the USDM correlation analysis and compared it with correlations where all years were evaluated. Years 2004 and 2013 were classified as D4 in a long-term drought perspective; however, both of the years were characterized by normal or above normal precipitation.

2.6.5. The Predictive Model

The prediction modelling process uses the EB that is derived from the EEP model in a piecewise regression with various drought indices from before the peak of the growing season to predict the end of growing season forage amounts.

We extracted 100 points from grassland areas within the Sandhills MLRA for each year of the analysis (2001-2016) and for each variable represented in the model. These included EB; site potential; seasonal integrated percentile index of USDM for winter and spring; annual integrated percentile index of USDM from the previous year; and SPI, ESI and EDDI from the last recorded date of winter and spring season representing conditions during the past 3 months (Figure 1).

3. Results

3.1. Site Potential

The value of site potential, estimated based on long-term GSN and representing the inherent productivity of a certain location generally decreased from west to east (Figure 2a), following the decreasing gradient of precipitation. We observed abrupt changes in productivity around wet meadows (Figure 2b) and land cover changes, especially for irrigated agricultural areas (Figure 2c,d). When developing a productivity model for a certain land cover class, it is optimal to exclude other land cover classes that might have different long-term productivities and responses to climate variability with the highest possible spatial resolution. An adequate spatial resolution is necessary for areas like the Sandhills, where there is a high spatial variability of topography and land cover.
Figure 2. The Nebraska Sandhills study area with site potential that represents the inherit productivity of any location (a). We observed abrupt changes in productivity around areas of wet meadows (b) that are represented by emergent wetlands (cross-hatched areas) and by lowland prairie (unhatched areas) land cover (from Nebraska GAP land cover analysis) in contrast to upland prairie that is less productive (hatched areas). Other abrupt changes were observed for other land cover classes, especially irrigated agricultural areas (center pivots) or open water (c and d).

3.2. EEP Model

The EEP model, a function of site potential and seasonal climate, was trained on 80% of the random points (n = 6120) and the other 20% of data points were used for independent testing. The model was limited to six rules, which is in compliance with recommendations of Gu et al. [63] to maximize accuracy while minimizing overfitting tendencies. Detailed model results and a list of all rules are provided in Supplementary Materials (S1). Table 2 summarizes the predictive variables that were used in the model and their overall importance (the percentage of the rules in which the variable was used in training the model calculated as the mean of stratification and percentages). The most important variables in the model were summer precipitation and site potential. Both the training and testing data $R^2$ values were strong at 0.86, which indicates a robust model with minimal overfitting. The average and relative error for these two datasets were similar and, importantly, the relative error was substantially less than one, which can be interpreted as a useful model. The relative error magnitude is the ratio of the average error magnitude to the error magnitude that would result from always predicting the mean value. Relative error indicates a useful mode when this value is less than 1 [65].

<table>
<thead>
<tr>
<th>Driving variable</th>
<th>Overall importance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Driving variables and their overall importance in the EEP model. Overall importance was calculated as a mean of frequency of usage in model stratification and prediction (%).
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Summer Precipitation</td>
<td>100</td>
</tr>
<tr>
<td>Site Potential</td>
<td>87</td>
</tr>
<tr>
<td>Winter Precipitation</td>
<td>63.5</td>
</tr>
<tr>
<td>Minimum Spring Temperature</td>
<td>50</td>
</tr>
<tr>
<td>Maximum Spring Temperature</td>
<td>44.5</td>
</tr>
<tr>
<td>Mean Summer Temperature</td>
<td>43</td>
</tr>
<tr>
<td>Spring Precipitation</td>
<td>36.5</td>
</tr>
<tr>
<td>Minimum Summer temperature</td>
<td>34</td>
</tr>
<tr>
<td>Mean Spring Temperature</td>
<td>33</td>
</tr>
<tr>
<td>Mean Winter Temperature</td>
<td>26</td>
</tr>
<tr>
<td>Previous Fall Precipitation</td>
<td>20.5</td>
</tr>
<tr>
<td>Maximum Summer Temperature</td>
<td>10</td>
</tr>
</tbody>
</table>

**Model Structure** 6 rules, 80% training points

**Training Dataset**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(R^2)</td>
<td>0.86</td>
</tr>
<tr>
<td>Average Error</td>
<td>1.5</td>
</tr>
<tr>
<td>Relative Error</td>
<td>0.36</td>
</tr>
</tbody>
</table>

**Testing Dataset**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(R^2)</td>
<td>0.86</td>
</tr>
<tr>
<td>Average Error</td>
<td>1.6</td>
</tr>
<tr>
<td>Relative Error</td>
<td>0.36</td>
</tr>
</tbody>
</table>

Figure 3 shows EEP model estimates against AEP observations where the distance of each point from the 1:1 line represents the residual of the model. Residuals that fell outside of the 90% confidence interval represent significant over- or underperforming anomalies. Circles located around the regression line are within the 90% confidence interval, triangles above the regression line and outside of the confidence interval represent over-performing pixels, and squares below the regression line and outside of the confidence interval are under-performing pixels. Over- and under-performing residuals illustrate the influence of land management and disturbances (vertical variation of pixels), while the EEP varies with site potential and interannual climate variability (horizontal variation of pixels).
Figure 3. AEP regressed on EEP predicted from Model 1 with the regression and 1:1 line and the 90% confidence limits displayed. Over- and underperformance points outside of the 90% confidence interval capture the influence of land management practices and disturbances (overgrazing, fire, irrigation, etc.) while the normal points represent the inherit noise in the model.

3.2.1. EB and Interannual Biomass Deviation

EEP, the output of the EEP model, was converted to biomass and is displayed for each year of the study period (2000–2016) on maps of the Sandhills MLRA in Figure 4. These maps reflect the interannual changes in biomass that are caused by climate conditions. There are apparent low biomass dry years (2002, 2006 and 2012), high biomass wet years (2009, 2015 and 2016), and normal years (2005, 2008 and 2013) in the maps. The maps also capture the gradient from lower productivity in the west to higher productivity in the east, which is consistent with the site potential map and is reflective of the high overall importance of site potential in the EEP model.
Figure 4. EB maps for the study period (2000–2016) in the Sandhills MLRA that were masked to include only desired grassland areas.

To be able to isolate the influence of climate from the site potential, we developed interannual biomass deviation maps that capture the deviation of biomass in a certain year from a long-term median (Figure 5). These maps effectively show the spatial distribution of areas that experienced significantly lower or higher amounts of biomass compared to the long-term median. The maps put the biomass of each year in a perspective of long-term values and reveal the impact of annual climate conditions on the amount of produced biomass on a regional scale. Years that were identified as dry (2002, 2006, and 2012) show that the entire Sandhills region experienced significantly lower than
normal biomass amounts. Similarly, we observed higher than normal biomass amounts in the majority of the study area during the wet years. In years that were identified as “normal” (2005, 2008, 2013), the majority of the study area was covered with near-normal biomass, while small areas of higher and lower biomass were also present.

Figure 5. Annual biomass deviation maps developed for the Sandhills MLRA from 2000 to 2016. Sandhills MLRA was masked to include only the desired grassland areas. Green colors represent significant positive deviation from long-term median, while red colors represent significant negative deviation.
3.2.2. Validation of the EEP Model

The time series of modelled and observed biomass suggests that the model captures correctly the interannual variability of biomass production (Figure 6). We observed general overprediction in this plot, especially during the early years of the analysis. The comparisons between the model-derived biomass and the field clipping data showed moderately strong relationships (Gudmundsen Sandhills Laboratory \( R^2 = 0.57 \), Barta Brothers Ranch \( R^2 = 0.66 \)). In general, our model slightly overpredicted the field data (Figure 7). The overprediction was more conspicuous for lower values than for higher values, and this phenomenon was more pronounced for the eastern Barta Brothers Ranch location than for the Gudmundsen Sandhills Laboratory located in the central region of the Sandhills. The more pronounced overprediction may be due to the fact that the Barta Brothers Ranch is located in the most eastern part of the region and might not be as representative of the entire area as the Gudmundsen Sandhills Laboratory, which is more centrally located. Given the difference between spatial resolution of eMODIS and the fields plots (62,500 m² and 0.25 m², respectively), the general relationship between modelled and field-observed biomass gives a reasonable verification of our model, but may indicate possible overestimation of biomass in severe drought conditions.

![Figure 6](image1)

**Figure 6.** Time series of biomass derived from the EEP model (dashed lines) and field observation (solid lines) in two locations in Sandhills—the Barta Brothers Ranch (orange) located in the eastern part of the region and the Gudmundsen Sandhills Laboratory (blue) located in the central part.

![Figure 7](image2)

**Figure 7.** Scatter plots for two locations of ground-based validation collected at the Gudmundsen Sandhills Laboratory (a) and the Barta Brothers Ranch (b). Clipping data (actual) are represented on the x-axis and modelled data (predicted) are on the y-axis. Each year of data collection (2000–2016) is represented by one data point (Gudmundsen data starting with the year 2004). The blue line
represents regression and the black line is the 1:1 relationship between actual and predicted biomass. Error bars are based on variation between sampling locations for each year.

We also analyzed the differences between the modelled biomass data and the SSURGO range productivity derived from the Sandhills region. The broad polygons of the same productivity in the SSURGO database are a result of the mapping methodology that is used for estimating the amount of biomass. Multiple soil associations within a polygon are weighted by area using percentage of the polygon attributes to determine a mean value for each polygon. This methodology could contribute significant noise when compared to individual 250 m pixels. Therefore, we compared the regional statistics of the long-term mean modelled productivity from the years (2000–2016) to the SSURGO dataset (Figure 8) rather than comparing spatial variability of biomass in these two datasets. We used major statistical measures (minimum, 5th percentile, 1st quantile, mean, median, 3rd quantile, 95th percentile, and maximum) describing the distribution of values in the two datasets. We found that the distribution of modelled values sufficiently captures the range of the expected values derived from the SSURGO dataset.
3.3. Correlation of Drought Indices

The results of the correlation analysis of the various drought indices reveal differences between indices, their accumulation periods, and the time of the growing season that each index represents (Table 3). The maximum strength of correlation for all drought indices combinations was 0.85 for the 3-month ESI at the end of July—representing conditions from the beginning of May, explaining roughly 72% of the biomass productivity. Generally, the longer the accumulation period of an index, the higher the correlation. The 1-month ESI yielded above 0.65 correlation from 3 May to the middle of August. The strength of the 3-month ESI correlation above 0.75 was observed from the mid of June to the middle of September, spanning across the growing season. For the 1-month SPI, the correlation is notably lower with the highest value of 0.65, which represents the time period from 10 May to 10 June. For the 3-month SPI, the highest correlation value was 0.81, representing moisture conditions from the middle of April through the middle of July, and for the 6-month SPI, the highest correlation (0.84) was achieved in a period that roughly encompasses the length of the whole growing season (24 March to 14 September). The correlation of 1- and 3-month EDDI did not exceed -0.75 (negative correlation due to a different scale of EDDI when low values correspond to wet conditions and vice versa). The VegDRI correlation was above 0.65 from 8 July and reached the highest value (0.73) at the end of the growing season.

Table 3. Summary of the R values for each drought index used in the regression analysis and the time period that each index represents. Different numbers of points and years were evaluated by various indices: SPI and EDDI (n = 1700, 2000–2016), ESI (n = 1696, 2001–2016), VegDRI (n = 800, 2009–2016). Points that did not have an assigned value were excluded from the analysis. Five weeks (first four weeks and the last week) of the growing season for VegDRI analysis were not evaluated due to the lack of data (“out of season” according to the VegDRI methodology). Correlation coefficient values that are equal to or higher than 0.65 are indicated in green, higher than 0.75 in yellow, and the highest correlation (0.85) in orange.
The correlation of the USDM percentile index revealed the strongest relationship with the productivity anomaly for fall \((R = 0.76)\) and summer \((R = 0.69, \text{Figure 9})\). Excluding years that represent long-term drought (2004 and 2013) notably improved the correlation of all analyzed seasons \((0.84 \text{ and } 0.82 \text{ for fall and spring, respectively})\). The relationship with winter and spring USDM percentile index appeared weak (not shown).

<table>
<thead>
<tr>
<th></th>
<th>3-Jun</th>
<th>3-Mar</th>
<th>3-Jun</th>
<th>5-Mar</th>
<th>3-Jun</th>
<th>3-Dec</th>
<th>3-Jun</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3-May</td>
<td>0.61</td>
<td>−0.66</td>
<td>0.68</td>
<td>0.67</td>
<td>−0.44</td>
<td>0.67</td>
<td>0.70</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>10-May</td>
<td>0.65</td>
<td>−0.60</td>
<td>0.71</td>
<td>0.74</td>
<td>−0.54</td>
<td>0.73</td>
<td>0.73</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>17-May</td>
<td>0.61</td>
<td>−0.60</td>
<td>0.71</td>
<td>0.76</td>
<td>−0.62</td>
<td>0.73</td>
<td>0.75</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>24-May</td>
<td>0.58</td>
<td>−0.56</td>
<td>0.73</td>
<td>0.77</td>
<td>−0.68</td>
<td>0.77</td>
<td>0.76</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>1-Jun</td>
<td>0.55</td>
<td>−0.63</td>
<td>0.74</td>
<td>0.78</td>
<td>−0.74</td>
<td>0.79</td>
<td>0.77</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>8-Jun</td>
<td>0.54</td>
<td>−0.56</td>
<td>0.76</td>
<td>0.78</td>
<td>−0.73</td>
<td>0.81</td>
<td>0.79</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>15-Jun</td>
<td>0.62</td>
<td>−0.64</td>
<td>0.79</td>
<td>0.81</td>
<td>−0.70</td>
<td>0.83</td>
<td>14-Jan</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>22-Jun</td>
<td>0.61</td>
<td>−0.63</td>
<td>0.81</td>
<td>0.80</td>
<td>−0.70</td>
<td>0.84</td>
<td>21-Jan</td>
<td>0.71</td>
<td></td>
</tr>
<tr>
<td>29-Jun</td>
<td>0.60</td>
<td>−0.65</td>
<td>0.83</td>
<td>0.80</td>
<td>−0.72</td>
<td>0.85</td>
<td>28-Jan</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>5-Jul</td>
<td>0.62</td>
<td>−0.67</td>
<td>0.80</td>
<td>0.78</td>
<td>−0.74</td>
<td>0.84</td>
<td>4-Feb</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>12-Jul</td>
<td>0.61</td>
<td>−0.68</td>
<td>0.77</td>
<td>0.78</td>
<td>−0.72</td>
<td>0.85</td>
<td>11-Feb</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>19-Jul</td>
<td>0.43</td>
<td>−0.66</td>
<td>0.72</td>
<td>0.73</td>
<td>−0.69</td>
<td>0.83</td>
<td>18-Feb</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>26-Jul</td>
<td>0.27</td>
<td>−0.50</td>
<td>0.62</td>
<td>0.69</td>
<td>−0.68</td>
<td>0.82</td>
<td>25-Feb</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>2-Aug</td>
<td>0.20</td>
<td>−0.34</td>
<td>0.58</td>
<td>0.69</td>
<td>−0.67</td>
<td>0.81</td>
<td>3-Mar</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>9-Aug</td>
<td>−0.07</td>
<td>−0.18</td>
<td>0.47</td>
<td>0.60</td>
<td>−0.62</td>
<td>0.80</td>
<td>10-Mar</td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td>16-Aug</td>
<td>0.05</td>
<td>−0.08</td>
<td>0.45</td>
<td>0.56</td>
<td>−0.62</td>
<td>0.78</td>
<td>17-Mar</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>23-Aug</td>
<td>0.21</td>
<td>−0.11</td>
<td>0.30</td>
<td>0.53</td>
<td>−0.60</td>
<td>0.75</td>
<td>24-Mar</td>
<td>0.84</td>
<td></td>
</tr>
<tr>
<td>30-Aug</td>
<td>0.20</td>
<td>−0.02</td>
<td>0.20</td>
<td>0.48</td>
<td>−0.54</td>
<td>0.70</td>
<td>31-Mar</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>7-Sep</td>
<td>0.50</td>
<td>0.03</td>
<td>0.06</td>
<td>0.48</td>
<td>−0.48</td>
<td>0.63</td>
<td>7-Apr</td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td>14-Sep</td>
<td>0.50</td>
<td>0.00</td>
<td>0.06</td>
<td>0.42</td>
<td>−0.40</td>
<td>0.55</td>
<td>14-Apr</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td>21-Sep</td>
<td>0.41</td>
<td>0.07</td>
<td>0.03</td>
<td>0.39</td>
<td>−0.30</td>
<td>0.41</td>
<td>21-Apr</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>28-Sep</td>
<td>0.12</td>
<td>0.26</td>
<td>−0.16</td>
<td>0.29</td>
<td>−0.12</td>
<td>0.27</td>
<td>28-Apr</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>
Figure 9. Scatter plots of the interannual biomass deviation and the USDM percentile index. Summer (a, c) and fall (b, d) were selected as two seasons where USDM percentile index explained the majority of the biomass variation. Upper plots (a, b) display all years of the analysis (2000–2016). Bottom plots (c, d) display values without two years that were identified as long-term drought (2004 and 2013). Specific years are identified by unique colors. The blue dashed line represents a significant deviation from normal and the black solid line is the regression line.

3.4. The Predictive Model

The predictive model was trained on 80% of the training points and evaluated on the remaining 20% of points. The EB in this model is a function of site potential and drought indices that represented the moisture conditions before and during the early growing season until the end of May. The independent variables used in the model and their overall importance is provided in Table 4. Detailed model results and a list of rules created by the model are provided in Supplementary Materials (S2). The site potential importance was strong, similar to the EEP model, and was the most important variable, followed by the three-month spring SPI and the previous year USDM percentile index. The $R^2$ of both training and testing datasets was 0.9, and their relative errors were substantially below 1. In Figure 10, we plotted predicted EB against observed EB derived from the testing dataset. The linear relationship appeared to be strong along the entire distribution with minimal bias. The training dataset consisted of a relatively small number of training points ($n = 1280$), which can make the model less robust to various environmental and climatic conditions.
Figure 10. Evaluation of the predictive model performance on a testing dataset. The observed biomass is seen to have regressed on the predicted biomass; linear regression is displayed as a dashed line and the 1:1 line is a solid one.

Table 4. Driving variables and their overall importance in the late spring biomass model (Model 2). Overall importance was calculated as a mean of frequency of usage in model stratification and prediction (%).

<table>
<thead>
<tr>
<th>Driving Variable</th>
<th>Overall importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site Potential</td>
<td>82</td>
</tr>
<tr>
<td>SPI 3m Spring</td>
<td>66.5</td>
</tr>
<tr>
<td>USDM Previous Year</td>
<td>50.5</td>
</tr>
<tr>
<td>SPI 3m winter</td>
<td>47</td>
</tr>
<tr>
<td>EDDI Spring</td>
<td>46.5</td>
</tr>
<tr>
<td>USDM Winter</td>
<td>46.5</td>
</tr>
<tr>
<td>USDM Spring</td>
<td>37</td>
</tr>
<tr>
<td>ESI Spring</td>
<td>33.5</td>
</tr>
<tr>
<td>EDDI winter</td>
<td>33</td>
</tr>
<tr>
<td>ESI winter</td>
<td>6.5</td>
</tr>
</tbody>
</table>

Model Structure: 24 total rules, 3 committee models, 80% training points

Training Dataset
- $R^2 = 0.9$
- Average Error: 129.83
- Relative Error: 0.31

Testing Dataset
- $R^2 = 0.9$
- Average Error: 143.67
- Relative Error: 0.32
4. Discussion

The current study provides a landscape-scale understanding of seasonal and interannual climate effects on the dynamics of grassland forage production in semi-arid upland locations of the Nebraska Sandhills. The EEP has been previously found to capture the variation of grassland productivity due to spatial changes in site potential and temporal changes in climate, while minimizing the influence of disturbances and land management [29]. Other studies, however, have focused on locating the impacts of land management practices and disturbances creating EEP anomalies, subtracting EEP from AEP [e.g., 29,43]. The strong use of site potential in the EEP model suggests that biomass production is dependent on a specific location [29,53,66] and corresponds well to the west-east productivity gradient [26,58]. Summer precipitation was the most important variable in the EEP model, appearing in all stratification and prediction rules. Total summer precipitation was previously found to explain majority of the time-integrated NDVI variation for warm-season grasses, the dominant grasses of the Sandhills uplands. The inclusion of summer season total precipitation in net primary productivity models from the Central Great Plains was supported [67].

The EB maps were found to successfully capture the wet and dry years, as well as the productivity gradient. The maps of interannual biomass deviation provide information about specific spatial distributions of below or above normal biomass for each year caused by interannual climate variability. A differential response of vegetation in various areas to climate conditions can influence the amount of produced biomass. However, we tried to eliminate this factor by choosing only one MLRA that should have similar biophysical characteristics (soils, climate, vegetation, etc.) [31], and by eliminating areas of wet meadows and other land cover classes. Maps of EB and the interannual biomass deviation, if provided as growing season biomass estimates during the early phases of growing seasons, might help land managers make decisions about appropriate stocking rates and other management considerations. For example, in 2007, a land manager from the drought-affected northwestern part of the Sandhills could have sought additional hay supply from the southern portion of the region, where there was above-average biomass production (Figures 4 and 5). A moderately high spatial resolution of our product could also provide information about areas with higher forage availability that would be suitable for potential relocation of animals during severe drought conditions, and be especially beneficial for Bureau of Land Management (BLM) lands or private landowners with large-area ownership.

To validate our findings from the EEP model, we compared modelled biomass data with long-term grass clipping data from two locations in the Sandhills. The relationship between these two datasets was relatively modest. The R² values in our analysis were in the range of values observed in other studies, despite the fact that the ratio of field-observed versus remotely sensed area was much larger in those studies, compared to the area of our validation [22,68,69]. Considering the relatively small sample size (compared to the potential variability in the area) and the difference in plot sizes of grass clipping (0.25-m²) and the EB data (62,500-m²), the model captured well the interannual variability in biomass production, perhaps with some underestimation of biomass during severe drought conditions. It is important to note that ground clipping observations can also introduce certain errors. Additional errors could have been introduced when converting the EEP to biomass using an empirical equation developed for a much larger area than the Sandhills [26]. Developing a biomass equation for a more specific area could lead to improvement of biomass estimates. The regional statistics of the SSURGO range productivity dataset further validated the overall model performance. The regression line’s close alignment with the 1:1 line in Figure 8 indicated minimal bias, although we observed a slight overprediction of lower productivity values. Models tend to underpredict high values and overpredict low values. We mitigated this tendency using various classes of productivity and a long period of observed NDVI [70]. Selecting more training points from wet and dry years that are characterized by higher and lower biomass, respectively, or using more classes of productivity stratification could lead to an improvement in the model.

We examined a suite of drought indices driven by different climate variables (SPI—precipitation, EDDI and ESI—evaporation, VegDRI—strong NDVI component, USDM—a blend of various indices and expert observations) and their relationship with biomass production. The
correlation analysis revealed strong relationships between certain drought indices and EB, derived from the model during various stages of the growing season. The highest correlation was observed for the three-month ESI, capturing the period of May, June and July. This period of time has been found to be sensitive to precipitation variability in Sandhills grasslands [12]. Longer drought index timescales had stronger explanatory power with regards to the interannual variation of EB. Strong relationships were achieved when the index covered the period of June and July, which is the peak growth period for warm-season dominant grasslands in Nebraska Sandhills [71]. Weaker relationships were observed for periods that captured the moisture amounts during the greenup and senescence of grasslands [12].

The USDM percentile index proved to be a successful method to compare the original categorical data derived from the USDM with interannual biomass anomalies. The strongest relationship was found for the summer and fall integrated index. Two years (2004 and 2013) that were classified as severe drought in the USDM, however, showed near-normal productivity values. Both of these years followed severe drought conditions in the previous years. One of the properties of the USDM is that it captures both short- and long-term drought, which is delineated in the weekly USDM online maps, but is not captured in the provided geographic information system (GIS) data. Long-term drought might not be entirely alleviated by shorter-term precipitation events during the growing season; however, this moisture supply can be crucial for forage production. Therefore, there is a need to distinguish between short- and long-term drought conditions when explaining grassland biomass variability using the USDM. We excluded the years 2004 and 2013 from our analysis, which notably improved the correlation coefficients for all seasons. Non-linear relationships are also to be noted, with stronger relationships in the lower-left quadrants of the scatterplots, indicating good relationships during the dry years, which is most likely caused by the inability of the USDM to capture wet conditions. This approach should be considered when using the USDM with respect to biomass production in semi-arid grasslands, for example in the Livestock Forage Program (LFP). The LFP program’s payment amounts for Nebraska statewide were comparable in years 2012 and 2013, despite the differences in total annual precipitation and the amount of biomass derived from our model. We recognize that ecosystem performance is not a perfect depiction of ecosystem health; other processes not captured by remote sensing can influence ecosystem health. Grassland species’ composition might be affected by severe drought. For example, in 2013, following the severe drought in 2012, the Sandhills grasslands showed more annual grasses and forbs than in other years [72], which might have an effect on the nutritional value or palatability of forage. The relationships described in this study might be different in other regions that are characterized by different physical and biological properties [73].

Our predictive model correctly captured the interannual variability in EB. The overall importance of variables in the model showed the importance of spring moisture for the exponential growth phase of warm-season grasses [74]. We examined the training and testing mean absolute error for overfitting (testing MAE–training MAE)/training MAE), which showed a slightly higher, but still acceptable value (10.6%, while overfitting is considered negligible when <10%). We also developed an additional predictive model that used expected biomass anomalies as the dependent variable without the use of site potential as an explanatory variable. This model yielded weaker regression, but was still significant ($R^2 = 0.88$, $MAE = 73.8$, $p < 0.01$). We did not include the VegDRI dataset in the predictive model because of its short period of record and the high use of NDVI in the VegDRI modeling process, which our predicted variable is derived from. We will continue to improve the predictive model using larger sample size, biomass stratification, and longer time series. A measure of soil moisture in the model is currently not represented because the sandy soils do not have high water-holding capacity. Therefore, it might not be important for sites like the Sandhills. We plan to further include soil moisture in the model, especially for areas characterized by higher water-holding capacity. We also plan to investigate other equivalent NDVI products that are being developed in the event of a possible MODIS decommission. Further research is needed to test the applicability of this methodology in other semi-arid grasslands.
When compared to other forage prediction tools, this model ties information about biomass production to specific locations, which allows the exclusion of areas that are not representative of major land cover. In the case of Sandhills, the excluded areas were wet meadows with different species composition, different growth curves, and a connection to near-surface groundwater levels that can be decoupled from climate variations for one or more years.

5. Conclusions

We developed models of Expected Ecosystem Performance (EEP) [75] and expected biomass (EB) based on site conditions, climate, and drought indices. The EEP model separated the influence of land management and disturbances from the influence of climate variations, which is an important step when focusing on the impacts of drought on biomass production. The novelty of this approach is that we did not focus on the impacts of land management and disturbances themselves, but rather on isolating the impacts of climate variations on variability in the EEP. The correlation analysis of drought indices showed important time periods and indices that can be observed with regards to the interannual biomass to make predictions of growing season biomass production. The strongest relationships were found either during or at the end of the growing season (end of July and September for ESI and SPI and fall for the USDM). These findings might be useful for a landscape post-season biomass production evaluation (e.g. for the LFP program). However, this information would not be timely for land managers and the decisions that they need to make as early as possible in the growing season. Therefore, we developed a model that utilized drought indices characterizing moisture conditions before and in the early stages of the growing season, which can be used for early predictions of annual biomass production.

Our study can enhance the understanding of interannual productivity variability with respect to drought. This information could be used to develop an operational tool that would inform land managers about the anticipated growing season conditions and the end of season biomass predictions. Such a tool has the potential to reduce the uncertainty that is connected to decision making, and can enhance the development of more informed decisions about stocking rates, purchasing hay, and other management actions. This research contributes to the three-pillar drought risk management framework by providing: 1) enhanced monitoring, 2) identifying vulnerable grassland areas, and 3) enhancing understanding of the interannual variability of productivity with respect to climate conditions.


Funding: This research was funded by the U.S. Geological Survey Land Change Science Program though US Geological Survey Cooperative Agreement no. G19AC00074 and was part of the “Drought Risk Management Research Center”, and is supported by the National Oceanic and Atmospheric Administration Competitive Grant No. NA15OAR4310110 in support of the National Integrated Drought Information System.

Acknowledgments: We thank Jerry Volesky for his professional guidance and useful discussions, and researchers working at the Gudmundsen Sandhills Laboratory and Barta Brothers Ranch for providing long-term ground clipping data for model validation. We also thank John Swigart and Tonya Haigh from the National Drought Mitigation Center for providing drought indices data and professional assistance. We would also like to thank Chris Funk from the U.S. Geological Survey for his review of the manuscript and his valuable comments and revisions. The authors would like to thank all the anonymous reviewers for their valuable comments and helpful suggestions, which lead to substantial improvements in the manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

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

References
4. Nebraska Department of Agriculture. *Nebraska Agriculture Facts Brochure*; Nebraska Department of Agriculture: Lincoln, NE, USA, 2016.
5. Nebraska Department of Agriculture. *Nebraska Agricultural Fact Card*; Nebraska Department of Agriculture: Lincoln, NE, USA, 2018.


72. Volesky, J.D. (University of Nebraska – Lincoln, North Platte, NE, USA). Personal communication, July 2019.


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).