January 2021

Scaling Near-Surface Remote Sensing To Calibrate And Validate Satellite Monitoring Of Grassland Phenology

Morgen W.v. Burke

SCALING NEAR-SURFACE REMOTE SENSING TO CALIBRATE AND VALIDATE SATELLITE MONITORING OF GRASSLAND PHENOLOGY

by

Morgen W. V. Burke
Bachelor of Environmental Science, Brandon University, 2014
Master of Science, University of North Dakota, 2016

A Dissertation

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Grand Forks, North Dakota

December
2021

This dissertation, submitted by Morgen W. V. Burke in partial fulfilment of the requirements for the Degree of Doctor of Philosophy from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

DocuSigned by:
Bradley Rundquist
Brad Rundquist

DocuSigned by:
Jeff VanLooy
Jeff VanLooy

DocuSigned by:
Haochi Zheng
Haochi Zheng

DocuSigned by:
Sean Hammond STH
Sean Hammond

DocuSigned by:
Gregory Vandeberg
Gregory Vandeberg

DocuSigned by:
Jesslyn Brown
Jesslyn Brown

This document is being submitted by the appointed advisory committee as having met all the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

DocuSigned by:
Chris Nelson
Chris Nelson
Dean of the School of Graduate Studies

11/23/2021
Date

PERMISSION

Title	Scaling Near-Surface Remote Sensing to Calibrate and Validate Satellite Monitoring of Grassland Phenology
Department	Earth System Science and Policy
Degree	Doctor of Philosophy

In presenting this dissertation in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my dissertation work or, in his absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my dissertation.

Morgen W. V. Burke
December 1, 2021

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I wish to express my greatest appreciation to Dr. Rundquist who was willing to take on my dissertation, and help me immensely through out the process, without your efforts this would not have been possible. Dr. VanLooy I am very appreciative of all you did to get me going. When I started my research in ESSP I had a lot of ideas, but without you I would not have realised what I was truly interested in, and how to get the ball rolling. Dr. Zheng, Dr. Hammond, Dr. Vandeberg, and Ms. Brown all helped me in so many ways during my Ph.D., and I appreciate any parts they played in the completion of it.

Dedicated to Justine, Diane, and Cam without whom the inspiration and dedication required to complete this research would not have been possible.

ABSTRACT

Phenology across the U.S. Great Plains has been modeled at a variety of field sites and spatial scales. However, combining these spatial scales has never been accomplished before, and has never been done across multiple field locations. We modeled phenocam Vegetation Indices (VIs) across the Great Plains Region. We used coupled satellite imagery that has been aligned spectrally, for each imagery band to align with one another across the phenocam locations. With this we predicted the phenocam VIs for each year over the six locations.

Using our method of coupling the phenocam VIs and the meteorological data we predicted 38 years of phenocam VIs. This resulted in a coupled dataset for each phenocam site across the four VIs. Using the coupled datasets, we were able to predict the phenocam VIs, and examine how they would change over the 38 years of data. While imagery was not available for modeling the 38 years of weather data, we found weather data could act as an acceptable proxy. This means we were able to predict 38 years of VIs using weather data. A main assumption with this method, it that no major changes in the vegetation community took place in the 33 years before the imagery. If a large change did take place, it would be missed because of the data lacking to represent it.

Using the phenocam and satellite imagery we were able to predict phenocam GCC, VCI, NDVI, and EVI2 and model them over a five-year period. This modeled six years of phenocam imagery across the Great Plains region and attempted to predict the phenocam VIs for each pixel of the satellite imagery. The primary challenge of this method is aggregating

grassland predicted VIs with cropland. This region is dominated by cropland and managed grasslands. In many cases the phenology signal is likely driven by land management decisions, and not purely by vegetation growth characteristics. Future models that take this into account may provide a more accurate model for the region.

CHAPTER I

I. INTRODUCTION

1.1 Background

Phenology is the study of reoccurring biological events or life stages, with particular interest in their timing and relationship with weather and climate (M.D. Schwartz, 2013). The reoccurrence of biological events can be influenced by both abiotic and biotic factors, and the interactions between them (A. D. Richardson, Keenan, et al., 2013). An understanding of phenology is thought to be as old as the existence of early civilizations, particularly those that understood the seasonality of crop growth (M.D. Schwartz, 2013). In Japan, starting in 705 AD there is a written record marking the beginning of cherry flowering in the spring (Menzel, 2013). In North America, the Smithsonian Institution started a phenology network in 1851, observing species of plants, birds and insects across 33 states (Mark D. Schwartz et al., 2013).

Phenology of vegetation is primarily concerned with the life cycle of plants throughout the growing season. This includes growth through photosynthesis in the spring time and the onset of senescence in the fall (Morisette et al., 2009). The link between climate and plant phenology has been well established with environmental factors such as temperature, photoperiod and precipitation acting as primary forcing to the onset of plant growth and senescence (Elsa E. Cleland et al., 2007; Kathuroju et al., 2007; Menzel et al., 2005; Morisette et al., 2009; A. D. Richardson, Keenan, et al., 2013; Mark D. Schwartz et al.,

2006). The relationship between plant phenology and climate provides a strong foundation for long-term monitoring of vegetation phenology as a way to track plant and ecosystem responses to global climate change (Elsa E. Cleland et al., 2007). For example, since vegetation is a primary producer, monitoring of vegetation can provide information on organisms at higher trophic levels, such as how early onset flowering requires pollinator insects to adjust their life cycles (Mark D. Schwartz et al., 2006). Monitoring of vegetation phenology can also detect unexpected feedback mechanisms, for example while global climate change in the northern hemisphere might be triggering an earlier start of season, and is expected to lengthen the growing season (Mark D. Schwartz et al., 2006). Earlier spring vegetation growth can reduce soil moisture causing drought and early onset senescence, which in turn shortens the total growing season (Toomey et al., 2015). Feedback mechanisms like this are still poorly understood, and an area where further research is needed (A. D. Richardson, Keenan, et al., 2013).

Monitoring of vegetation phenology takes place on many spatial scales. While phenology monitoring has been done at a local single-species scale historically, for example the cherry flowering in Japan (Menzel, 2013), more recently an area of research within the field of remote sensing called land surface phenology (LSP) has permitted much larger geographic areas to be monitored using satellite sensors (G. Henebry & Su, 1995). Broad-scale LSP has been occurring since the 1980s with the deployment of the Advanced Very High-Resolution Radiometer (AVHRR) with a 1km spatial resolution. This has been followed by the deployments of the Moderate Resolution Imaging Spectroradiometer

(MODIS) in 1999 and the Visible Infrared Imager/Radiometer Suite (VIIRS) in 2017 (Reed et al., 2009; X. Zhang, Liu, et al., 2017). Also the U.S. Geological Survey (USGS) Landsat satellite series with a spatial resolution of 30 m to 60 m provides imagery dating back to the 1970s and has been used in several phenology studies (Baumann et al., 2017; Jeremy I. Fisher & Mustard, 2007; Jeremy Isaac Fisher et al., 2006; Y. Liu et al., 2017; X. Zhang, Wang, et al., 2017). The more recent launch of the Sentinel-2 satellites in 2015 - 2017 with a 10 m to 60 m spatial resolution has already proved useful for LSP research (Vrieling et al., 2018). Since satellite-based LSP does not necessarily contain the spectral information for a single species, it is assumed that each pixel or grid cell contains the aggregate spectral information for a variety of plants with an ecosystem (G. Henebry & Su, 1995; Vrieling et al., 2018). This means that LSP requires a conceptual shift away from historical phenology research focused on a single species to a focus on the seasonality and characteristics of landscape surface reflectance (G. Henebry & Su, 1995).

Near-surface remote sensing (NSRS) is an even more recent approach to phenological research that is helping to bridge the gap between historical observer-based phenological data and LSP (A. Richardson & Braswell, 2009). NSRS includes any remote sensing instrument that can record electromagnetic radiation at an ecosystem level, and is typically mounted to a permanent structure such as a tower, mast or building and can include radiometric or other imaging sensors (A. D. Richardson, Klosterman, et al., 2013). While radiometers have been used to study the reflective properties of natural surfaces for many years (Coulson et al., 1965), a more recent technological shift is the use of digital web-enabled cameras to act as

imaging sensors for phenological research (A. D. Richardson et al., 2007). Web-enabled cameras that have been configured to record near-continuous imagery of vegetation are called “phenocams” (T. B. Brown et al., 2016). However, this term originated from the digital cameras specifically used for the PhenoCam Project (https://phenocam.sr.unh.edu). Phenocams provide two main benefits that are often difficult with other phenology observations. First, they can provide information about a small footprint or area of an ecosystem but still provide the aggregated pixel information used in LSP. Second, they provide continuous monitoring of an ecosystem and are less effected by environmental factors such as cloud cover that will obstruct satellite imagery, or weather conditions that may stop an individual from collecting observational data (A. D. Richardson, Klosterman, et al., 2013). Using phenocam data we can examine the characteristics and relationships of different environmental drivers on the vegetation phenology (T. B. Brown et al., 2016). This is achieved using the phenocams’ digital repeat photography and either visual interpretation by an individual (Kosmala et al., 2016), or quantitative analysis with the use of a vegetation index (VI) (Petach et al., 2014).

Phenocams have proved to be an invaluable resource for improving our understanding of ecosystem scale LSP (T. B. Brown et al., 2016). However, further improvements in the way phenocam data are calibrated and standardized is needed (A. D. Richardson, Klosterman, et al., 2013). Changing illumination conditions can cause variation in derived Vegetation Indices (VIs), and developing a better way to calibrate and remove this noise would be beneficial (Petach et al., 2014). The green chromatic coordinate (GCC) VI

has been used by several studies derived from phenocam imagery (L. A. Brown et al., 2017; Browning et al., 2017; Keenan et al., 2014; Morisette et al., 2009; Sonnentag et al., 2012), while only a few have used the Normalized Difference Vegetation Index (NDVI) (Filippa et al., 2018; Klosterman et al., 2014; Petach et al., 2014). This is likely because NDVI requires the phenocam to record infrared radiation, which is usually secondary to imagery with blue, green, and red radiation (A. D. Richardson, Klosterman, et al., 2013). This is likely because not all phenocams have the ability to record infrared radiation, and those that do require addition data processing to extract and calibrate the infrared imagery to make it compatible with the color imagery (Petach et al., 2014). Improving the use of infrared from phenocam imagery may improve our ability to detect vegetation health and stress (Petach et al., 2014), as well as improve our ability to compare the imagery to satellite-derived NDVI (Filippa et al., 2018). While NDVI and GCC indices are both measures of vegetation greenness, they do not always posses strong correlation for different phenology transitions and across ecosystem types. Better understanding of the physiological processes associated with variation between these two indices is needed (Filippa et al., 2018; Morisette et al., 2009). With this study, we hope to improve upon the calibration and standardization issues currently facing phenocam imagery. Improving the ways phenocam imagery is calibrated and compared with satellite derived VIs will bring great value to the use of phenocams for monitoring phenology, and environmental change (Filippa et al., 2018; Petach et al., 2014). Also developing methods to identify primary environmental forcing that are driving change in phenology over the

growing season will help improve the usefulness of co-located weather sensors for ecosystem modeling (T. B. Brown et al., 2016).

1.2 Literature Review

Vegetation phenology started with individuals monitoring and recording biological events for various flora, for example the cherry tree flowering recorded in Japan starting in 705 AD (Menzel, 2013). In the UK the Marsham family recorded the flowering date of wood anemone starting in 1736 until 1958, representing one of the longest records kept by amateur naturalists (Sparks & Menzel, 2002). More recently, researchers in North America have recorded timing of bud burst for cloned lilac and honeysuckle plants starting in the 1950s (Mark D. Schwartz, 2003). The National Oceanic and Atmospheric Administration (NOAA) created a network with the two plants in 1967, recording several phenophases of leaf opening and blooming for the lilac, and nine different phenophases for the honeysuckle (Lieth, 1974). Using these observations an advancing of spring green up has been detected in the northern hemisphere (Mark D. Schwartz et al., 2006). Schwartz (1994) also suggested calibrating observations of lilac and honeysuckle phenophases with spring green up detected from satellite derived NDVI. Even today observations of these two plant species are still used by the U.S. National Phenology Network to track blooming and leaf-out on an annual basis, to detect variations in seasonal phenology, such as the timing of spring green up across the U.S. (USA National phenology Network, 2018).

Phenology is often carried out as a hobby, or pastime of naturalists and citizen scientists (A. D. Richardson, Keenan, et al., 2013). The recognition that phenological data are a valuable resource for understanding the response of ecosystems to climate change (Fig. 1) have made them desirable for use in research and modeling (Sparks & Menzel, 2002). Phenology may even be the simplest way to monitor ecosystem and species change in response to climate change (The Intergovernmental Panel on Climate Change, 2007).

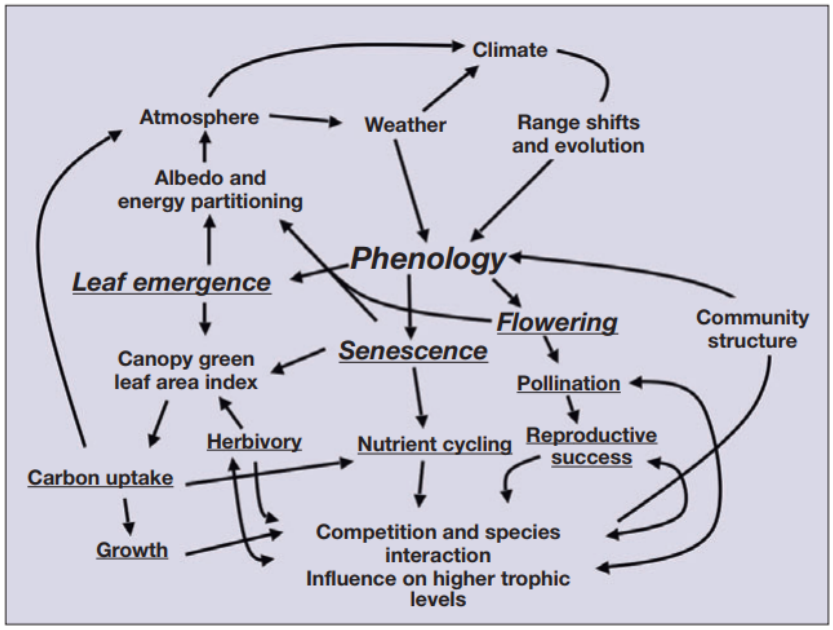


Fig. 1: Web of connections between phenology and the environment. Weather and climate influence vegetation phenology, but feedback mechanisms exist through nutrient cycling and atmospheric interactions. The underlined terms represent an ecosystem service with management or economic benefits. Figure originally published in Morisette et al. (2009).

1.2.1 Weather and Climate

Recent understanding of the feedback mechanisms between vegetation phenology and climate has brought recognition to the important value of the phenological records that have

been kept for many centuries (Sparks & Menzel, 2002). Seen in Fig. 1, phenology is not only influenced by changes in the weather and climate (Rundquist & Harrington, 2000), but has recognized feedback mechanisms that can influence atmospheric conditions (Morisette et al., 2009). Vegetation is able to influence climate through mechanisms such as albedo, surface roughness length, canopy conductance, water and energy fluxes, and photosynthesis and CO_2 fluxes (A. D. Richardson, Keenan, et al., 2013).

The albedo of an ecosystem is the proportion of incident solar radiation that is reflected. Any vegetation on the surface will have a direct effect on the albedo, and this will in turn affect the energy budget and the climate (Pitman, 2003). A good example of this is the mathematical model known as Daisyworld, in which a biosphere exists with only two species, black and white daisies (A. J. Watson & Lovelock, 1983). The albedo of the white daisies is greater than that of the black daisies, and the daisies can self-regulate their local temperature so that they can survive under a greater range of incoming solar radiation then if they were passive actors unable to modify the surrounding climate. While the model is simplified compared with the complexities found in the natural world, it is still able to show us how organisms can feedback to their surrounding environment, and have an effect on it (Lenton & Lovelock, 2001).

Photosynthesis and the ability of vegetation to uptake CO_2 and sequester carbon into organic matter is a very important feedback mechanism for affecting global temperature (A. D. Richardson, Keenan, et al., 2013). Measurements of CO_2 flux from the biosphere became technologically feasible starting in the 1980s with some of the first measurements taken

using the eddy covariance method. Since this time, we have been able to measure the uptake and respiration of CO_2 from the surrounding ecosystem and establish networks of measurements (Baldocchi, 2008). Using flux networks it has been estimated that the global terrestrial vegetation is able to remove 2 - 4 Pg. (2 - 4 trillion kg, or 4.4 - 8.8 trillion lbs.) of carbon from the atmosphere annually (A. D. Richardson, Keenan, et al., 2013). The expected increase in growing season brought on by increased atmospheric CO_2 (Mark D. Schwartz et al., 2006) means that vegetation will have more time to use photosynthesis and remove carbon from the atmosphere, creating a negative feedback loop (Baldocchi, 2008). However, this might not be the case in all regions, such as locations where drought caused by less spring soil moisture results in a shortening of the growing season (Toomey et al., 2015). Exceptions such as this demonstrate how important it is for a diverse array of ecosystems to by monitored to provide a better understanding of how vegetation phenology will change under future climate conditions.

1.2.2 Grassland Phenology

Grasslands cover approximately 59 million km^2 of the Earth surface (Hufkens et al., 2016) making up between 10 and 30 percent of the global carbon stock (Scurlock & Hall, 1998), this makes grasslands the second largest carbon sink after forests (Anderson, 1991). In North America, the Great Plains cover approximately 2.9 million km^2 within a gradient of tall and short grass prairie. However, the conversion of grassland to cropland has drastically reduced the remaining native prairie ecosystems. In 2018, it is estimated that only half of these grassland ecosystems remain, with 87 percent of them located on poor and marginal

quality soils (World Wildlife Fund, 2018). The variation within the Great Plains creates a variety of community types typically dominated by C_3 grasses in the north and east having more precipitation and cooler temperatures, and C_4 grasses in the south and west having lower precipitation and higher temperatures (Petrie et al., 2016). Along with a large amount of spatial variability, grasslands are also characterized with high amounts of temporal variability (Flanagan & Adkinson, 2011). This means that climate change induced shifts in grassland phenology will likely only be detectable using long-term monitoring over several years to decades (G. M. Henebry, 2013).

Across the northern hemisphere the onset of spring is predicted to occur earlier under the warmer conditions predicted from climate change (Mark D. Schwartz et al., 2006). However, a controlled test of grassland phenology using plants grown with a warmer temperature, elevated CO_2 increased nitrogen, and increased precipitation has shown an array of responses that were not all anticipated. Additions of CO_2 had a delay on spring greenness while increased nitrogen slowed down the growth acceleration. Precipitation had no effect suggesting it was not a limiting factor for the controlled plants, while increased temperature was the only factor to have the expected outcome, causing plants to flower earlier by 2-5 days (E. E. Cleland et al., 2006). Field observations of arid grasslands using both phenocams as well as satellite imagery are also in agreement that warmer temperatures bring an earlier start of season to the grasslands. However, in an arid environment precipitation has been found to affect the recorded vegetation indices, even causing a second peak greenness in the growing season after a large precipitation event (Browning et al., 2017). Even though

precipitation may have a small influence on grassland phenology, modeled scenarios under future climate conditions still suggest that North America will see an increase in both the length of the growing season as well as the productivity of grasslands. This is because the model grasslands are expected to become more efficient in retaining moisture under higher CO_2 levels allowing for a more efficient use of water and a reduction in the amount of water lost in transpiration (Hufkens et al., 2016). This suggests that precipitation may need to fall below a certain threshold before it will have a noticeable effect on growing season length (Browning et al., 2017).

1.2.3 Satellite-Scale Land Surface Phenology

The use of satellites to track LSP on a global scale required a conceptual shift in the way phenology was understood (G. Henebry & Su, 1995). Satellite sensors such as AVHRR, MODIS, and VIIRS have been used to track global LSP with a near-daily temporal resolution starting in the 1980s. The spatial resolution of these sensors ranges from 1 km for AVHRR down to 250 m for some MODIS and VIIRS bands (Reed et al., 2003), however, at this resolution the spectral properties of several land cover types are mixed, making measured changes in the spectral reflectance a property of entire vegetation communities, rather than a single species (G. Henebry & Su, 1995). These satellite sensors cover large areas, and they can provide almost daily coverage of the globe, giving more opportunities to collect imagery that is cloud free. This allowed studies of global LSP starting in the 1980s with organizations such as NOAA producing global or regional multiday composites that were more manageable for researchers (Reed et al., 1994). Unfortunately, the use of imagery covering

such a large area has proven difficult to validate since on the ground measurements cover only a small fraction of the area covered by these sensors. Also, many of the validation methods that have been used, such as field measurements of a few species, or various climate and hydrosphere models, are not directly comparable since they do not measure the same biophysical properties as satellite LSP (X. Zhang et al., 2018).

The Landsat satellite series began in 1972 and provides a long-term record of LSP that can be studied at a local or regional level (Robinson et al., 2017). The Landsat satellites have traded the higher temporal resolution of AVHRR, MODIS, and VIIRS for a significant increase in spatial resolution. The 16-day period between successive Landsat imagery makes it difficult to detect quickly changing phenological stages (Reed et al., 2009). The introduction of the Thematic Mapper (TM) sensor on Landsats 4-5 saw an improved spatial resolution of 30 m from the Multispectral Scanner's 68 m by 83 m bands used on Landsats 1-5. Landsat 7 introduced the Enhanced Thematic Mapper Plus (ETM+), which added a 15 m resolution panchromatic band, and Landsat 8 carries the Operational Land Imager (OLI), which added three more bands but had no additional improvements in spatial resolution (U.S. Geological Survey, 2018). Using the Landsat imagery it is possible to generate VIs, and in particular the NDVI, which has been used in hundreds of thousands of journal articles starting from its first use in 1973 (Robinson et al., 2017). To improve the temporal resolution of the Landsat series data, data fusion with MODIS and VIIRS imagery has been used in studies, using models such as the Spatial and Temporal Adaptive Reflectance Fusion Model

(STARFM) to provide near daily imagery (Gao et al., 2006; Hilker et al., 2009; Walker et al., 2014).

Calibration of global land surface products is still a challenge and has only been partly addressed, for example by compositing multiple daily images to smooth out image variation. This unfortunately means that many LSP products are unavailable at temporal resolutions higher than several days to weeks based on their composite periods (G. M. Henebry & de Beurs, 2013). Curve-fitting and function fitting are other methods used to remove noise from LSP, unfortunately these methods also have downfalls such as removing actual phenological variations, or even introducing new errors to the time series (Jeremy I. Fisher & Mustard, 2007). Because of the difficulties in validating data, LSP products often vary from ground measurements by more than ten days. This makes LSP difficult for use in climate monitoring when changes detected over decades often vary by only by a few days (X. Zhang, Wang, et al., 2017).

Recent deployments of different satellite constellations may help remove some of the issues faced by satellites with either poor spatial or poor temporal resolution. The European Space Agency (ESA) recently launched the Sentinel 2A and Sentinel 2B satellites each with a Multi Spectral Instrument with a spatial resolution between 10 and 60 m. With the two sensors combined the repeat imagery from these sensors is approximately five days. This higher image frequency could remove the need to fuse satellite imagery like has been done with Landsat and MODIS (Vrieling et al., 2018). The PlanetScope constellation of 120 CubeSat 3U satellites has achieved an almost daily repeat image of the Earth with a spectral

resolution of 3.125 m (Planet Team, 2016). While these satellites are not recording bands in the middle and thermal-infrared end of the spectrum like Landsat and Sentinel 2, they are recording the visible and near-infrared data needed for VIs such as NDVI and GCC. Currently no phenology product exists with PlanetScope data, likely because the data record began in 2017, however future use of satellite constellations such as this could drastically improve the temporal difficulties faced by current long-term satellite records.

1.2.4 Near-Surface Remote Sensing Phenology

Near surface remote sensing (NSRS) provides an approach that may bridge traditional and satellite scale remote sensing. This is done by providing field level imaging that can target specific species or vegetation communities using spectral information that is then comparable with other forms of remote sensing (A. Richardson & Braswell, 2009). Since NSRS instruments use spectral information in the same form as satellite scale remote sensing, and often use similar or identical VIs, the physiological changes detected in vegetation are comparable. This is an improvement over satellite produced VIs being compared against records of spring budding or senescence. Not only does NSRS often combine spectral properties of entire vegetation communities, it also does not depend on these single observer records (A. D. Richardson, Klosterman, et al., 2013).

NSRS can include any remote sensing device that records spectral information where the sensor is mounted to a post, structure, tower or held. This includes hyperspectral devices such as spectrometers (Mohler & Goodin, 2013), and goniometers (Coburn & Peddle, 2006), or it can include multispectral imaging devices such as a Tetracam (Higgins et al., 2011), or a

phenocam (T. B. Brown et al., 2016). The hyperspectral devices can record hundreds of spectral bands, but often lack the ability to detect spatial variation, while imaging devices often only record spectral information for three to six bands (A. D. Richardson, Klosterman, et al., 2013).

Phenocam started as a single web-enabled camera, recording NSRS imagery in the Bartlett Experimental Forest in New Hampshire, U.S. (A. D. Richardson et al., 2007). Ever since, the term “phenocam” has become more generalized to mean any web-enabled camera used to study phenology (T. B. Brown et al., 2016). The Phenocam Network (https://phenocam.sr.unh.edu/webcam/) now has more than 500 active cameras in various field sites across the globe, however approximately 480 are within the contiguous U.S., while there is little coverage in South America, Africa and Asia (T. B. Brown et al., 2016). Many different vegetation community types are represented by phenocam imagery, including cropland, deciduous forest, coniferous forest, mixed forest, grassland, shrubs, tundra, wetlands (A. D. Richardson et al., 2018) and even a tidal salt marsh (O’Connell & Alber, 2016). Phenocams started monitoring a forest, and their use for forest phenology has continued to grow (Hufkens et al., 2012; Keenan et al., 2014; Klosterman et al., 2014; Melaas, Sulla-Menashe, et al., 2016; Sonnentag et al., 2012; Toda & Richardson, 2018; Yingying et al., 2018). Phenocam research has also been used in various grassland studies (Browning et al., 2017; Cremonese et al., 2017; Inoue et al., 2015; Julitta et al., 2014; Y. Liu et al., 2017; Migliavacca et al., 2011; Q. Zhou et al., 2019). However, work is still needed to improve the way phenocam imagery is compared and scaled with satellite sensors (Tang et

al., 2016), as well as which VIs are used to detect phenophases in different ecosystem types (Helman, 2018). Calibration and protocol development for the ways in which phenocam imagery can be used and accessed is still a challenge that needs to be addressed (A. D. Richardson, Klosterman, et al., 2013).

Calibration of phenocam imagery is a work in progress, and a few studies have made improvements in the way phenocam imagery are used (Filippa et al., 2016, 2018; Petach et al., 2014). The NetCam SC IR (StarDot Technologies, Buena Park, CA) camera that is standard in many phenocam studies was chosen because of its ability to record both a visible blue, green, and red color image, and an infrared (IR) image. However, several specifications about this camera were unknown, and this made it difficult to integrate the infrared feature into the camera. Since the IR image is not taken at the same time as the color image, and the IR image does not filter out visible light, but instead includes both, the raw digital number (DN) values between the color and IR image are not directly comparable. Petach et al. (2014) were able to fix this issue by using a lab controlled environment to determine ratios for the three color bands to subtract from the IR image, so that the DN values represent only the IR radiation. They were also able to adjust exposure values using a square root function to make sure the color and IR imagery taken at slightly offset times are comparable and can be used to derive VIs that rely on color and IR information such as NDVI (Filippa et al., 2018). Finally, for the four spectral bands they were able to determine what portion of the electromagnetic spectrum (EMS) the camera sensor is sensitive to (Petach et al., 2014). This provides valuable information when comparing phenocam imagery with other sensors that

may have differences in the wavelengths of the EMS that they are sensitive to, which could result in differences in their measured phenology signal. With this information, comparing a phenocam and satellite NDVI signal has been proven possible, with a linear scaling equation (Filippa et al., 2018). While Petach *et al*. (2014), made some crucial findings in improving phenocam calibration, they recognized that further calibration studies are needed. Changing illumination conditions can have a large effect on a phenology signal, and this has only been managed by smoothing the signal, for example by using a three-day composite (A. D. Richardson, Klosterman, et al., 2013). Unfortunately, this lowers the temporal resolution of the imagery, and relies on composites, rather than being able to extract an index value from any given image. However, calibration panels or other standardised object in a phonecam's' field of view may be able to remove this noise, and is an area of calibration that should be addressed (Petach et al., 2014; A. D. Richardson, Klosterman, et al., 2013).

The development of a software package that can be utilized through R (R Foundation for Statistical Computing, Vienna, Austria) has provided a foundation for protocol development using phenocam imagery. The package is called Phenopix, and with it users can analyze phenocam imagery more readily without needing a strong background in computer science, or digital data analytics (Filippa et al., 2016). Within Phenopix users can define an area of interest within a phenocam's field of view, then choose a VI to use. From here, the user can filter the data using one of five methods, fit a curve to the VI using one of five methods, and extract various phenophases from the imagery using the four available methods. Lastly, the user can estimate the statistical uncertainty in the imagery to measure

how much variation exists for the phenophases (Filippa et al., 2016). The R package can also implement the IR correction developed by Petach et al. (2014), which is important for the use of NDVI derived from phenocam imagery. Continued development of frameworks such as this will allow for widespread use of phenocam imagery in research. The establishment of open-source development environments, such as those using R or Python, allow collaborative work to add future improvements to phenocam calibration and protocol, and make these improvements more accessible to a broader community of researchers.

1.2.5 Vegetation Indices

VIs are dimensionless measures derived from spectral data acquired from a remote sensing instrument, they often measure the amount of green vegetation present (Jones & Vaughan, 2010). Vegetation will interact with electromagnetic radiation by either absorbing, transmitting or reflecting the radiation. The way in which plants interact with electromagnetic radiation is dependent on the photosynthetic and other pigmented tissues in the plants leaves (Helman, 2018). VIs take advantage of these relationships and allow us to remotely monitor plant growth and development throughout the growing season (Xue & Su, 2017).

There are many different VIs that have been developed using various portions of the electromagnetic spectrum (Bannari et al., 1995). For example, Xue and Su (2017) have documented more than 125 different VIs that have been developed to take advantage of the ways in which plants interact with electromagnetic radiation, and to reduce noise created by inaccuracies in measurement. Inaccuracies in VIs are often caused by electromagnetic

radiation interacting with atmospheric aerosols, clouds, soil, and water (Jones & Vaughan, 2010). While there have been many VIs produced, this has also led to some criticism: because many VIs are based upon similar spectral information, they are often significantly correlated and may not provide more information than the VI they are attempting to improve on (Glenn et al., 2008). Several VIs takes advantage of plants' interaction with the red (~620 nm) and the near infrared (NIR) (~800 nm) electromagnetic radiation. This is because red light is absorbed by healthy vegetation through photosynthetic tissues, while NIR light can damage plant cells and is typically highly reflected by or transmitted through plant tissues (Helman, 2018). The interaction plants have with red and NIR light has led to the creation of VIs such as the Difference Vegetation Index (DVI) (Eq. 1), the Ratio Vegetation Index (RVI) (Eq. 2) and the NDVI (Eq. 3) (Bannari et al., 1995; Jones & Vaughan, 2010; Xue & Su, 2017). NDVI is the most widely used index. NDVI is sensitive to vegetation, and is normalized so that values typically fall between 0 and 1 (Xue & Su, 2017). However, values as low as -1 can occur when imaging water, clouds or snow (Jones & Vaughan, 2010). Having normalized values makes it easy to compare values across multiple locations (Xue & Su, 2017). While NDVI has been widely adopted many modified versions exist that try and manage problems found with atmospheric and soil background effects, as well as saturation in areas of dense vegetation (Helman, 2018). A modified NDVI has even been produced to be used in environments that are often snow covered prior to the onset of spring green-up, since snow cover can have a large effect on NDVI values (Cong Wang et al., 2017).

$$DVI = NIR - Red \quad (1)$$

$$RVI = \frac{NIR}{Red} \quad (2)$$

$$\mathrm{N}DVI = \frac{\mathrm{N}IR - \mathrm{R}ed}{\mathrm{N}IR + Red} \quad (3)$$

In addition to NDVI, another VI that has become prominent in phenology is the use of chromatic coordinates (Gillespie et al., 1987), more specifically the Green Chromatic Coordinate Index (GCC) which can be found in several publications (L. A. Brown et al., 2017; Browning et al., 2017; Filippa et al., 2018; Julitta et al., 2014; Keenan et al., 2014; A. D. Richardson et al., 2007; Sonnentag et al., 2012; Toda & Richardson, 2018). GCC does not depend on having NIR and instead only needs imagery containing red, green, and blue (RGB) spectral information. Using RGB, total brightness can be calculated be simply summing together the digital number (DN) of each color. We can then calculate normalized brightness for any color by dividing its DN by the total brightness (A. D. Richardson et al., 2007). Equation 4 shows this using green as the numerator to produce the GCC index. Using GCC it is possible to compare the relative percent greenness between sequential measurements taken at the same location. GCC, like NDVI, can be used for the detection of phenophases over the growing season (Fig. 2).

$$\mathrm{GCC} = \frac{\mathrm{Green}_{\mathrm{DN}}}{\mathrm{Blue}_{\mathrm{DN}} + \mathrm{Green}_{\mathrm{DN}} + \mathrm{Red}_{\mathrm{DN}}} \quad (4)$$

NDVI and GCC are both being used to derive estimates of vegetation greenness over the growing season using phenocam imagery (Filippa et al., 2018; Petach et al., 2014). However, the two indices do not always correlate with to various phenophases, or across ecosystem types (Filippa et al., 2018). Research is still needed to identify how life stages in different vegetation species affect VI measurements, and how these affect scaling to various remote sensing platforms (Morisette et al., 2009). In grassland ecosystems one of the largest differences between GCC and NDVI occurs in their prediction of the end-of-season DOY, in which the GCC date can be up to 50 days earlier then NDVI (Filippa et al., 2018). In addition, NDVI has been used to model changes in the leaf area index (LAI) (Steltzer & Welker, 2006), while GCC has been used in correlation with gross primary productivity (GPP) (Toomey et al., 2015).

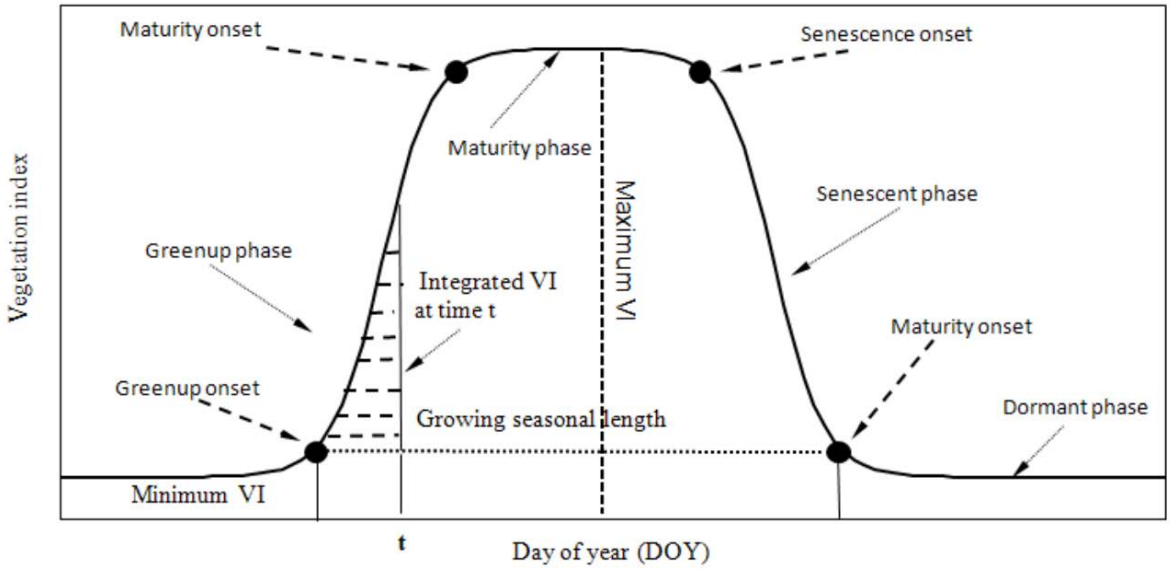


Fig. 1: Various phenophases that can be captured using a vegetation index, originally published in Zhang (2012).

Determining phenophase transition dates using VIs relies on developing methods that can identify specific transition periods in vegetation growth and senescence (Fig. 2) (X. Zhang, 2012). While the determination of different phenophase events seems straight forward, many different approaches and terms have been used. For example, the identification of spring green-up, or the start-of-season, or the green-wave (Mark D. Schwartz, 1994) that occurs every growing season, may all represent vegetation growth in the springtime but could be referring to different biological processes (White et al., 2009). Three main methods exist that have been used to determine the day of year (DOY) in which a phenophase occurred, and within these methods, there are a few variations. Global thresholds are perhaps the simplest method to measure phenophase transitions. A phenophase is defined by the DOY on which the VI reaches a set threshold value. For example, the DOY recorded as onset green-up could be set based on when NDVI reaches a value of 0.2 or 0.3 (White et al., 2009). These thresholds may need to be adjusted for any given location, and do not allow for easy spatial comparison. The global threshold approach can be modified into a second category called local thresholds. Within local thresholds the VI is adjusted so that the minimum and maximum values over the growing season or a historic record are scaled between zero and one. Then a threshold value is set, such as the midpoint, to determine a given phenophase DOY (White et al., 2009; White & Nemani, 2006). The percent-above-threshold model (PAT) modifies the local threshold model further by requiring a majority of imagery pixels to reach the threshold value. This reduces the ability of data noise to reach a threshold early, but requires VI data to have multiple pixels within the study region (White &

Nemani, 2006). Mathematical models make up the third method used to determine a given phenophase DOY value. Instead of relying on thresholds, a function is fit to the VI values, and changes in slope over the growing season are used to identify shifts in phenophase (X. Zhang, 2012).

Using both the mathematical models and the local thresholds to determine phenophase transitions have several advantages. Since VIs are a unit-less measures of vegetation growth, no constant value is necessarily going to describe the same growth stage in vegetation for any given location. This means that using a constant global threshold is not highly effective, and does a poor job of describing similar vegetation growth stages at different locations (White & Nemani, 2006). Instead using historic minimum and maximum values to adjust local thresholds allows values to be comparable between locations, and they are more likely going to describe similar changes in vegetation growth (White et al., 2009). In addition, unlike mathematical models, local thresholds do not smooth out variations in data that could be caused by true disturbance events, and because they use historic records to establish minimum and maximum values they can be used for predictive modeling (White & Nemani, 2006). In contrast, mathematical models are less sensitive to variation in VI values caused by noise from non-vegetation components such as snow or cloud cover, and are less likely to inaccurately detect true phenophase transitions (X. Zhang, 2012).

1.3 Project Objectives

This study will answer the following groups of questions:

1) How can calibrated imagery within a phenocam's field of view be used to normalize phenocam imagery under illumination conditions that change daily and seasonally? Assuming satellite calibrated measurements are correct, can we develop an algorithm to correct phenocam GCC, EVI2 and NDVI?

2) What are the statistical relationships between phenocam GCC, VCI, NDVI, EVI2 and the co-located weather station measurements, including solar radiation, soil moisture, precipitation, soil temperature and air temperature? How can we identify the primary forcing between start of season, peak greenness, senescence, and end of season? What are the differences between the four VIs in their ability to detect changing environmental conditions?

3) How can we couple together VI information from multiple phenocam sites to allow for both temporal and radiometric calibration of phenocam imagery. Can we use this information to model VIs across the study region, and will aligning the imagery allow us to predict VI values at any given location? How do the predicted models compare with datasets that have been left out of the modeling process?

This dissertation is organised so that each major chapter seeks to address each of the three objectives. Each chapter is organized and formatted in the style of a scientific journal article to facilitate future publication.

II. Scaling Phenocam GCC, NDVI, and EVI2 with Harmonized Landsat-Sentinel using Gaussian Processes

Burke, Morgen W.V., and Bradley C Rundquist. 2021. “Scaling Phenocam GCC, NDVI, and EVI2 with Harmonized Landsat-Sentinel Using Gaussian Processes.” *Agricultural and Forest Meteorology* 300 (January). Elsevier B.V.: 108316. doi:10.1016/j.agrformet.2020.108316.

2.1. Introduction

Phenology is the study of biological events that have reoccurring cycles. These cycles can be influenced by both abiotic and biotic factors, and the interactions between these factors (A. D. Richardson, Keenan, et al., 2013). The study of vegetation phenology is often done at two spatial scales (Zeng et al., 2020). Traditional monitoring of vegetation phenology consisted of ground-based observations, often with a small sample size across a small geographic extent (A. Richardson & Braswell, 2009). For example, since 801 AD there is a written record marking the beginning of cherry flowering in Japan (Aono & Kazui, 2008; Aono & Saito, 2010). Phenology networks have improved upon this by increasing the observational extent, and standardising protocols (Mark D. Schwartz et al., 2013).

Satellite-scale land surface phenology (LSP) has been made possible with the increased temporal resolution of satellite sensors, often repeating imagery on a bi-weekly to weekly schedule (Helman, 2018). LSP required a conceptual shift away from traditional single species monitoring to satellites covering a broad range of species across an ecosystem (G. Henebry & Su, 1995). However, satellites often have trade offs between spatial and temporal resolution that can make it difficult to detect changes in phenology, particularly

when changes over a decadal period often vary by only a few days (X. Zhang, Liu, et al., 2017). Harmonization of different satellite sensors, such as the Harmonized Landsat-8 and Sentinel-2 (HLS) surface dataset (Claverie et al., 2018), can help address some of these challenges (Q. Zhou et al., 2019), but spatial and temporal gaps remain.

Near-surface remote sensing (NSRS) can help fill the gap between traditional and satellite-scale phenology (Browning et al., 2017). NSRS instruments can have a high temporal resolution with repeat measurements made multiple times a day, and capture information over a given area instead of focusing on individual species (A. Richardson & Braswell, 2009). The Phenocam Network (https://phenocam.sr.unh.edu) is a system of web-enabled digital cameras used as NSRS devices to capture time-lapse photography of various ecosystems across the U.S., with a few cameras in other countries across the globe. The cameras are referred to as phenocams, and are used to capture changes in the radiometric properties of vegetation within their field of view (FOV) (T. B. Brown et al., 2016). There are several other phenology networks that are also gathering data across the globe including the U.S. National Phenology Network (Denny et al., 2014), the European Phenology Network (Van Vliet et al., 2003), and the Phenological Eyes Network (Nagai et al., 2018).

Phenocams allow the phenological changes of vegetation, called phenophases, such as spring green-up and fall senescence, to be tracked throughout the growing season (A. Richardson & Braswell, 2009). The color information provided by the camera can be used to calculate vegetation indices (VIs) (Filippa et al., 2018) and these have been used within mathematical models to examine changes in the timing of vegetation phenology (Elmore et

al., 2012; Ren et al., 2018). The most commonly used VI from phenocam data is the Green Chromatic Coordinate (GCC), a proportional measure of relative channel brightness that has been shown to reduce noise in the phenology signal (L. A. Brown et al., 2017; Browning et al., 2017; Cremonese et al., 2017; Cui et al., 2019; Julitta et al., 2014; A. Richardson & Braswell, 2009; Sonnentag et al., 2012; Toda & Richardson, 2018; Vrieling et al., 2018). A modification of the GCC has also been proposed with the Vegetation Contrast Index (VCI), which has an increased dynamic range (X. Zhang et al., 2018). The VIs calculated using phenocam data can also be used for validation with data gathered from other sources, such as satellite imagery (Q. Zhou et al., 2019) or individual observation (Kosmala et al., 2016).

While NSRS using phenocams has grown over the past decade, there are still some challenges when using data from the cameras. Phenocams provide an image based on digital numbers (DNs) that do not represent true measures of reflectance or radiance. This means that imagery needs to be calibrated or filtered to get a useful signal (Filippa et al., 2018; Piao et al., 2019). To use VIs that exploit the near-infrared (NIR) portion of the electromagnetic spectrum, many phenocams have a mechanical cut-filter that allows the camera to capture a color and a color-infrared image. Then, with post-processing, the NIR information can be extracted for use in VIs such as the Normalized Difference Vegetation Index (NDVI) (Filippa et al., 2018; Petach et al., 2014). NDVI relies on the NIR band that is highly reflected by the mesophyll leaf structure, and the red band that is absorbed by chlorophyll (Jones & Vaughan, 2010; Pettorelli et al., 2005). NDVI has a long history in earth

observation (Rouse et al., 1973), and has been increasingly used with satellite imagery for ecological studies (Pettorelli et al., 2005).

Changes in solar illumination between consecutive phenocam images can have a large effect on the recorded DN values (Goodin et al., 2004). To manage the noise introduced by changing solar illumination, images are often filtered, either by time of day or using a classification system (O'Connell & Alber, 2016), or they are aggregated into daily or multiday composites to smooth out the signal (Filippa et al., 2018; Petach et al., 2014). While this has generally been found to reduce noise throughout the growing season, the DN values used for calculation still do not match true measures of surface reflectance (SR) and cannot be directly compared against other remote sensing products such as satellite imagery, or between phenocam sites (Sonnentag et al., 2012). This has been somewhat addressed by scaling phenocam metrics to that of satellite imagery or *in-situ* field spectrometers (Filippa et al., 2018; Petach et al., 2014). However, having an *in-situ* spectrometer to scale phenocam imagery is redundant since the spectrometer is likely to record the radiometric signal more accurately. Also, phenocams are considerably lower cost and do not require regular sensor cleaning and calibration making them more desirable for widespread use over spectrometers (A. D. Richardson, Klosterman, et al., 2013). Scaling phenocam data with satellite imagery relies on having imagery at high enough temporal resolution to provide a reliable measure between various phenophases. However, even with high temporal resolution imagery, the shape of the phenology signal may differ between satellite and phenocam sensors caused by the "scale effect" leading to different phenophase dates being derived from each sensor

(Licong Liu et al., 2019). Reference panels have also been suggested as a means to remove noise between consecutive images (Browning et al., 2017; Ide & Oguma, 2010; Petach et al., 2014; A. Richardson & Braswell, 2009), however, panels can fade and change in their spectral properties over time making them unreliable for long-term measurements (A. D. Richardson, Klosterman, et al., 2013).

Gaussian process (GP) regression is a machine learning method that is characterized as a nonparametric Bayesian approach (Aghighi et al., 2018). A GP regression fits a defined covariance function to a dataset by recreating the underlying signal and removing the noise (Rasmussen & Williams, 2006). The GP does require optimization of observational noise and covariance function hyperparameters, however this can be achieved using stochastic gradient-based optimization (Kingma & Lei Ba, 2015). In LSP, GPs have typically been used for smoothing and gap filling time series VIs such as the leaf area index (LAI) (Belda et al., 2020; Verrelst et al., 2012) and NDVI (Aghighi et al., 2018; Jönsson & Eklundh, 2002; Rodrigues et al., 2012).

We propose applying a machine learning approach using a GP to scale phenocam imagery to HLS, and to derive scaled VIs from the phenocam DN values. We will use the GP primarily to smooth the HLS VIs and scale the phenocam VIs to the GP HLS models. This will allow us to calculate VIs that utilize the red, green, blue (RGB) and NIR portions of the electromagnetic spectrum and align them with VIs calculated from standardised reflectance measures. This could greatly improve the ability to make comparisons between phenocam measurements at different sites or make comparisons with VIs measured using satellite

sensors. Also, VIs calculated using SR are better able to maintain relationships with measures of canopy cover and LAI (Jones & Vaughan, 2010).

2.2. Methods and Data

2.2.1 Study Area

We selected six phenocam field locations to carry out our proposed methods (Fig. 3). The three northern stations located within the temperate prairies are the Oakville Prairie station (Oakville), a part of the University of North Dakota located in Grand Forks County, North Dakota (47.8993°N, 97.3161°W); the USGSEROS station at the Earth Resources Observation and Science (EROS) Data Center near Sioux Falls, South Dakota (43.7343°N, 96.6234°W); and the Nine Mile Prairie station (Nine-Mile), a part of the University of Nebraska – Lincoln (40.8680°N, 96.8221°W), located in Lancaster County, Nebraska. The three southern stations, located within the south-central semiarid prairies, are a part of the National Ecological Observatory Network (NEON). These sites include the NEON.D06.KONZ.DP1.00033 station (Konza) (39.1008°N, 96.5631°W), located at the Konza Prairie Biological Station near Manhattan, Kansas, the NEON.D10.ARIK.DP1.20002 station (ARIK) (39.7582°N, 102.4471°W) located near the Arikaree River in Colorado, and the NEON.D11.OAES.DP1.00033 station (OAES) (35.4106°N, 99.0588°W) located at the Klemme Range Research Station in Oklahoma. These stations represent a latitudinal gradient from North Dakota (47.8993°N) to Oklahoma (35.4106°N) through the North American Great Plains Region. Grassland or prairie phenocam sites will be the focus of this research. Grasslands can readily be monitored from both near-surface and satellite-scale remote

sensing having a strong seasonal signal (Petach et al., 2014) and a relatively homogenous canopy when compared with forests or other more complex land cover types (Ali et al., 2016; Y. Liu et al., 2017).

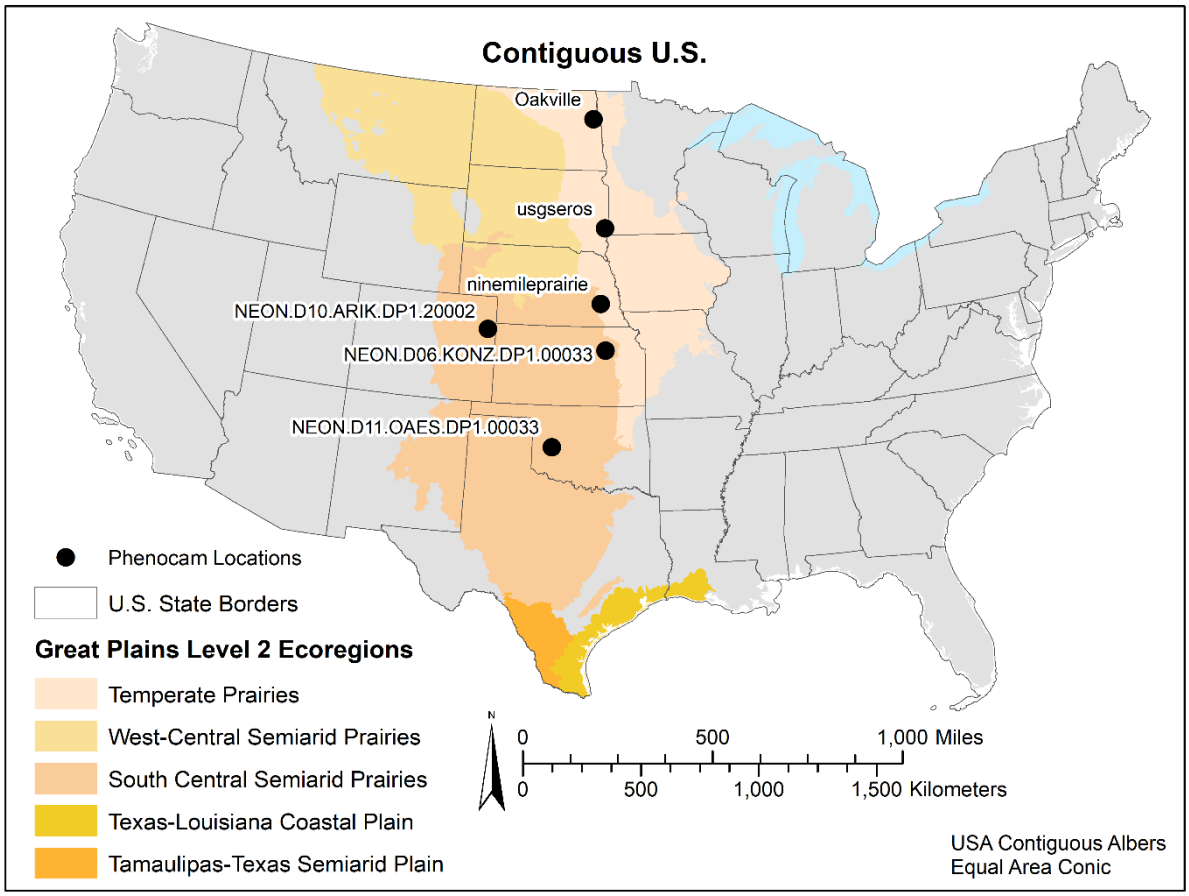


Figure 3: Study area showing the six phenocam locations situated within the Great Plains of the Contiguous U.S.

2.2.2 Determining Phenocam Field of View

Mapping each phenocam's FOV was necessary to determine the spatial extent that the phenocam is capable of imaging. For the Oakville station, we had access to the field site, and were able to carry out an *in-situ* method to determine the phenocam's FOV. To do this

several datasets are needed. First, a series of points need to be collected within the phenocam's FOV. This was done using a Trimble Geoexplorer (Trimble Inc., Sunnyvale, CA) GPS receiver in which 20 georeferenced points were recorded within the FOV, with the help of a guide to ensure georeferenced points were captured within the FOV while standing as close as possible to the FOV boundaries (Fig. 4a).

In addition, ancillary measurements were needed from the phenocam including the azimuth (352°), height above ground (1.57 m), the angle towards the ground, where parallel with ground is 0° with a downward angle being negative (-18°), and an accurate georeferenced point location for the Phenocam (Fig. 4b). The last data source needed was a digital elevation model (DEM) for the phenocam region. For the Oakville station a $1/9^{th}$ arc second DEM is available from the U.S. Geological Survey (USGS) National Map (https://viewer.nationalmap.gov/basic/). Using the gathered datasets, we used Arcpy within ArcGIS Desktop 10.6 (Environmental Systems Research Institute, Redlands, CA) to run the visibility tool within the 3D Analyst toolset (Krienert, 2015). This produced a polygon showing the FOV for the phenocam, allowing us to determine the region of vegetation that is visible within the phenocam imagery (Fig. 4b).

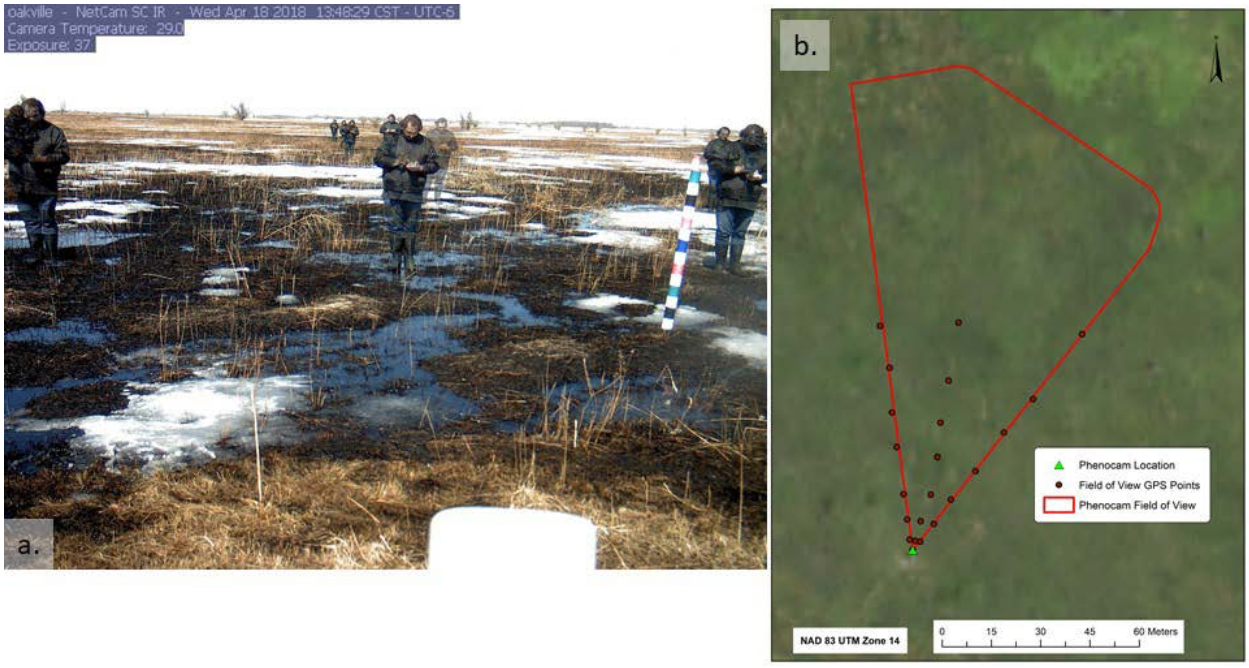


Figure 4: (a) Using a handheld GPS to map the FOV extent (b) The FOV for the Oakville phenocam. Twenty-point locations were recorded within the FOV and used to determine the spatial extent visible by the phenocam.

For the five other field locations we choose to use an *ex-situ* method to determine the FOV. Using the satellite-based basemap imagery available within ArcGIS Desktop 10.6, and imagery from U.S. Department of Agriculture's National Agricultural Imagery Program (NAIP), we visually located each of the phenocam stations and estimated the spatial extent, cross-referencing with the phenocam imagery to identify landmarks, such as trees and roads, within the imagery (Fig. 5). These spatial extents were used to extract pixel information for the HLS imagery, which has a spatial resolution of 30m x 30m, and we do not expect errors in over or underestimating the FOV to have significant effects on our results since grasslands are relatively homogenous.

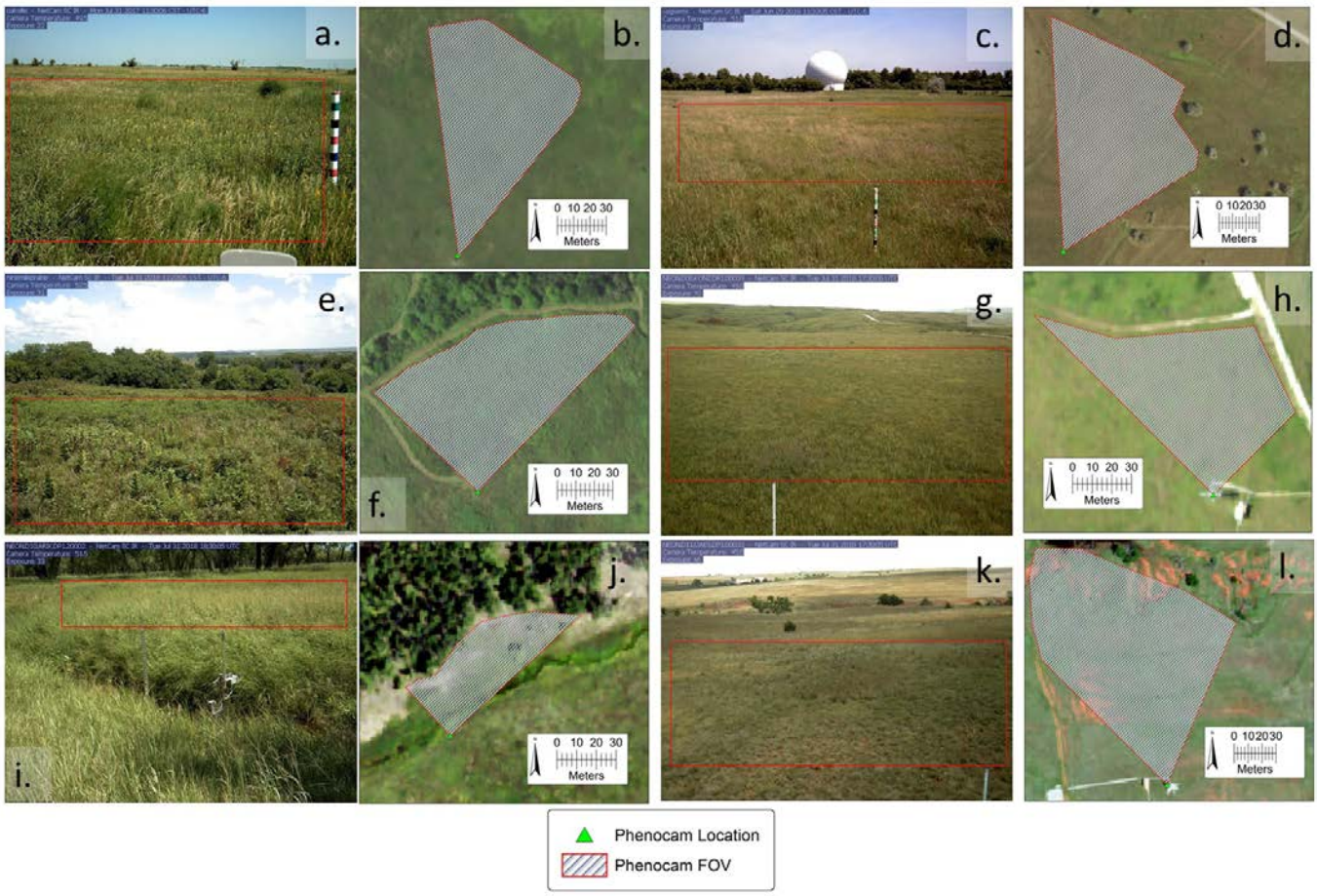


Figure 5: Field of View (FOV) for five of the phenocam stations. FOV boxes were drawn in each of the phenocam images (a, c, e, g, i, and k) along with maps showing the estimated spatial extent (b, d, f, h, j, and l) on NAIP imagery. (a) Oakville phenocam and (b) spatial extent; (c) USGSEROS phenocam and (d) spatial extent; (e) Nine-Mile phenocam and (f) spatial extent; (g) Konza phenocam and (h) spatial extent; (i) ARIK phenocam and (j) spatial extent; and (k) OAES phenocam and (l) spatial extent.

2.2.3 Processing HLS Imagery and Extraction of Reflectance

We acquired HLS from the National Aeronautics and Space Administration's (NASA) data portal (https://hls.gsfc.nasa.gov/data/v1.4/). By combining imagery taken from both the Landsat-8 Operational Land Imager (OLI), and the Sentinel-2 Multispectral Instrument (MSI) sensors, we can acquire many measurements taken over the growing

season. We used the Sentinel-2 S30, and the Landsat-8 L30 products, which provided 30m x 30m spatial resolution datasets. We identified imagery tiles for each of the field sites, and retrieved all imagery recorded between January 1, 2013, and December 31, 2019, giving us seven years of data at each site.

Typically, for low-resolution imagery such as the Moderate-Resolution Imaging Spectroradiometer (MODIS), a single pixel or the average value of a 3x3 or 5x5 pixel window centered over the phenocam site is used to represent satellite reflectance (Hufkens et al., 2012). However, with the HLS's higher resolution, we chose to incorporate all pixels in the FOV into our reflectance calculation. For each field site we iterated through each of the HLS images and extracted the pixels for the RGB and NIR bands that fell within the sites' FOV (Fig. 5). Image processing was done using the Python 3.7 programming language (Python Software Foundation, Beaverton, Oregon). To filter out cloud and snow cover images we used the median value from the blue band for each image and identified any images that had a median reflectance greater than 8.5 percent. We visually inspected these images and found that this threshold was able to identify images with clouds or snow covering our FOV. Therefore, we removed any imagery from our dataset that had a blue median value above 8.5 percent. Next, we used Inverse Distance Weighting (IDW) to derive a weighted average reflectance value from each image that would be associated with the phenocam FOV. IDW was calculated with Eq. 1, by measuring the distance between the phenocam and the center of each pixel (d_n), as well as the pixel's value (x_n). This provided an

interpolated measure in which pixels closer to the phenocam have a greater influence on the resulting value (Shepard, 1968).

$$IDW = \frac{\left((x_1 / d_1) + (x_2 / \mathrm{d}_2)\ldots + (x_n / \mathrm{d}_n)\right)}{\left(\left({}^{1}/_{d_1}\right) + \left({}^{1}/_{d_2}\right)_{\ldots} + \left({}^{1}/_{d_n}\right)\right)} \quad (1)$$

2.2.4 Calculating VIs from HLS imagery

We selected three different VIs to calculate using the HLS IDW reflectance values and the phenocam DN values. GCC (Eq. 2), was originally developed for use with phenocams, and provides a measure of 'greenness' that is relatively stable under changing illumination conditions (A. Richardson & Braswell, 2009). NDVI (Eq. 3) has a long history of use in Earth Observation (Rouse et al., 1973), and provides a normalized value with a theoretical range from zero to one, except for clouds, snow and water that often produce negative values (Jones & Vaughan, 2010; H. Q. Liu & Huete, 1995). The Enhanced Vegetation Index (EVI) was produced as a modification of NDVI that could better minimize effects caused by soil background and atmospheric noise (H. Q. Liu & Huete, 1995). A variation of EVI was later produced (EVI2) (Eq. 4), that could be calculated using only red and NIR reflectance, allowing it to be used on sensors that lack blue reflectance, such as the Advanced Very High Resolution Radiometer (AVHRR) (Jiang et al., 2008). While several other VIs exist that could be used for our analysis, we found these three were the most prominent in the literature and should offer a good gauge of how well our methods performed.

$$GCC = \frac{Green}{Blue + Green + Red} \tag{2}$$

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{3}$$

$$EVI2 = 2.5\frac{NIR - Red}{NIR + 2.4 * Red + 1} \tag{4}$$

2.2.5 Curve fitting the HLS Imagery with a GP

For each of the field locations we used a GP to fit a locally periodic covariance function (Eq. 5) to the IDW interpolated HLS VIs calculated in section 2.2.4 against the Julian century for each image (Rasmussen & Williams, 2006). This allowed us to calculate the value for the VIs at any point in time across the seven years of HLS imagery. We used TensorFlow within Python 3.7 to carry out the GP regression. TensorFlow provides an interface, and framework to execute machine learning algorithms (Abadi et al., 2015).

$$\mathrm{k}_{LocallyPeriodic}(\mathrm{x},\mathrm{x}') = \sigma^2 \exp\left(-\frac{2\sin^2\left(\frac{\pi|\mathrm{x}-\mathrm{x}'|}{\mathrm{p}}\right)}{\ell^2}\right)\left(1+\frac{(\mathrm{x}-\mathrm{x}')^2}{2\alpha\ell^2}\right)^{-\alpha} \tag{5}$$

The covariance function is made up of both a periodic and a rational quadratic covariance function, also known as a kernel. The multiplication of these two kernels allows the model to align with the yearly periodicity of the data, while still being flexible to changes between years (Camps-Valls et al., 2016; Duvenaud, 2014; Rasmussen & Williams, 2006). The periodic kernel required a period hyperparameter (p), while the rational quadratic kernel

required a mixing hyperparameter (α), and both required an amplitude(σ) and length-scale hyperparameter (ℓ). All hyperparameters were given an initial value of 1, except for the period hyperparameter, which was set at 0.009, the approximate length of a year in Julian centuries. An observational noise hyperparameter was also set at 0.01 to manage the balance between bias and variance. After these initial hyperparameter settings, each one was further fit to the HLS data using AdaMax optimization (Kingma & Lei Ba, 2015) run over 2,000 permutations with a learning rate of 0.01. This produced a HLS GP model for each of the three VIs at each field site.

2.2.6 Phenocam Image Exposure Correction

We used the phenocam imagery for each of the six field sites from the Phenocam Network website (phenocam.sr.unh.edu). Imagery is available from the sites across multiple years, with Oakville and Nine-Mile covering 2016-2019; USGSEROS covering 2015-2017, 2019; and the three NEON sites covering 2017-2019. Throughout the year an image was recorded every half hour during the day, typically starting around 4:00 am and ending around 10:30 pm local standard time. Each of our chosen sites has the NetCam SC IR with automatic exposure, and we selected only images with an exposure from 1/1,000 s to 1 s. Imagery taken in the early morning or late at night often had exposure values above 1 s where the prairie vegetation was not visible because of the low light levels. For each image recorded four sets of information were produced: an RGB image, a mixed RGB-NIR image, and two metadata files containing the imagery parameters such as time, date, and exposure (Fig. 6).

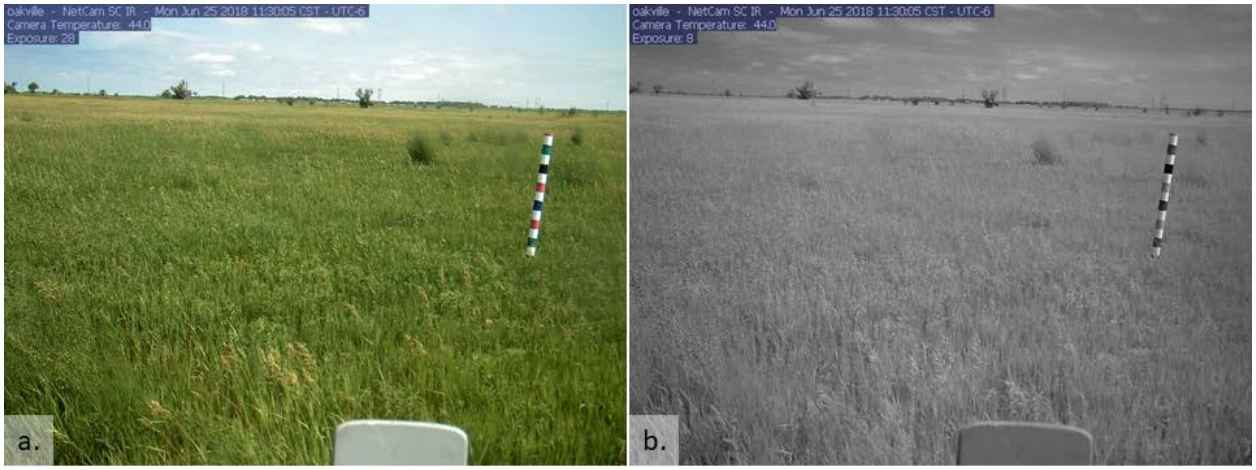


Figure 6: The FOV from the phenocam showing an example of (a) an RGB image and (b) the mixed RGB-infrared image. The resolution of the imagery is 1296 x 960 pixels (5 megapixels).

The first step in processing the phenocam imagery required combining RGB imagery with NIR information present in the mixed RGB-NIR images. To accomplish this task the methods laid out by Petach *et al.* (2014) were automated using the Python 3.7. Using these methods, the RGB-NIR mixed image (Z_{DN}) was corrected (Z'_{DN}) using its exposure (E_Z) (Eq. 6), then using the RGB image the visible component (Y_{DN}) was calculated (Eq. 7) and was corrected (Y'_{DN}) using the RGB image's exposure (E_Y) (Eq. 8), and finally the NIR image (X'_{DN}) was extracted (Eq. 9). The same exposure correction was applied to each band of the RGB imagery (Eq. 8). This resulted in four-band phenocam images ($RGBIR_{DN}$) with RGB and NIR information in which exposure was standardized between bands.

$$Z'_{DN} = \frac{Z_{DN}}{\sqrt{E_z}} \tag{6}$$

$$Y_{DN} = 0.30 \cdot Red_{DN} + 0.59 \cdot Green_{DN} + 0.11 \cdot Blue_{DN} \tag{7}$$

$$Y'_{DN} = \frac{Y_{DN}}{\sqrt{E_Y}} \quad (8)$$

$$X'_{DN} = Z'_{DN} - Y'_{DN} \quad (9)$$

2.2.7 Scaling Phenocam Imagery to HLS Reflectance

For each of the six sites we separately applied a method to scale the phenocam DN values to the HLS GP models from section 2.2.5. We used Python 3.7 to iterate over each $RGBIR_{DN}$ image and extract the pixel values within the FOV (Fig. 6). We then grouped the DN values by day and calculated the 90th percentile for each band (DN90), as well as the standard deviation of the blue band, and the number of images recorded each day. Using the 90th percentile across multiple images has been shown to reduce variability, and is often done across a three-day window (Hufkens et al., 2012; Sonnentag et al., 2012).

We found snow presence in the phenocam imagery introduced variability into the calculated DN values. Research detecting snow cover within phenocam imagery has been conducted (Kosmala et al., 2018), however the research required the use of neural networks and data from 133 field sites, and is not easily implemented at an individual station, or small subset of stations. Since this was not available for our field sites, we instead relied on a simple measure of standard deviation, which proved effective for detecting daily snow presence. The highly reflective surface of the snow results in blue band DN values approximately three times higher than is typical from vegetation, and this also produces greater variation of blue band DN values throughout the day from changing solar

illumination. To exclude snow-covered images, we removed days that had a standard deviation in the blue band of greater than 4.5 or had less than 20 images recorded throughout the day. We found this simple measure did an accurate job of identifying imagery with snow present and was able to help reduce noise within the DN values.

Using the DN90 data we calculated GCC (Eq. 2), NDVI (Eq. 3), and EVI2 (Eq. 4). These VIs were then scaled to the HLS GP VIs. We used linear regressions (Eq. 10) to scale the phenocam VIs to the HLS GP modeled VIs. This produced a set of the three scaled VIs (GCC_s, $NDVI_s$, $EVI2_s$) for each phenocam site. The entire methods process used to derive these three VIs is summarised in Figure 7.

$$VI_{scaled} = VI_{phenocam} * m + b \quad (10)$$

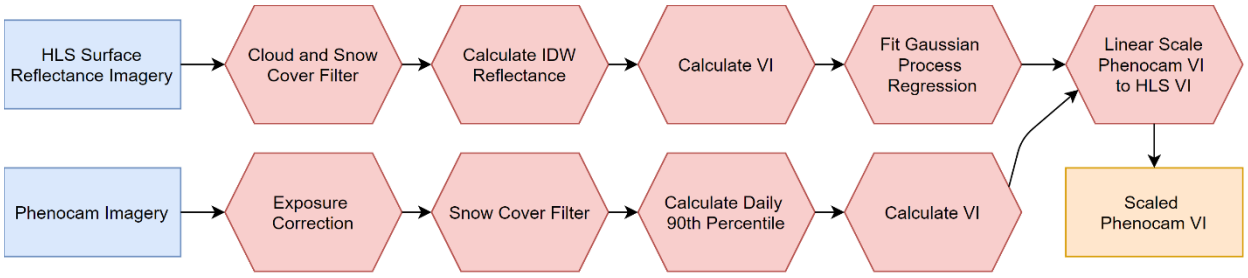


Figure 7: Flow chart of the methods used to scale the phenocam VIs using VIs derived from HLS imagery.

2.3. Results

2.3.1 HLS Data Analysis

2.3.1.1 Extracting HLS FOV reflectance

We used IDW to calculate HLS reflectance at the site level. We also calculated the mean and standard deviation for each image's FOV and compared the relationship between the IDW and mean values. Across all bands the IDW and mean values were highly related with R^2 ranging from 0.96 to 0.99, and root-mean-square error (RMSE) ranging from 0.0033 to 0.0075 (Fig. 8). Since both values have such a high correlation, and RMSEs of less than 1 percent, it is likely acceptable to use either approach. However, the standard deviation of FOV pixels, particularly in the blue band was quite noisy, often spanning half the data range. Because of this, we decided to use the IDW reflectance, which provides a better control on spatial variability.

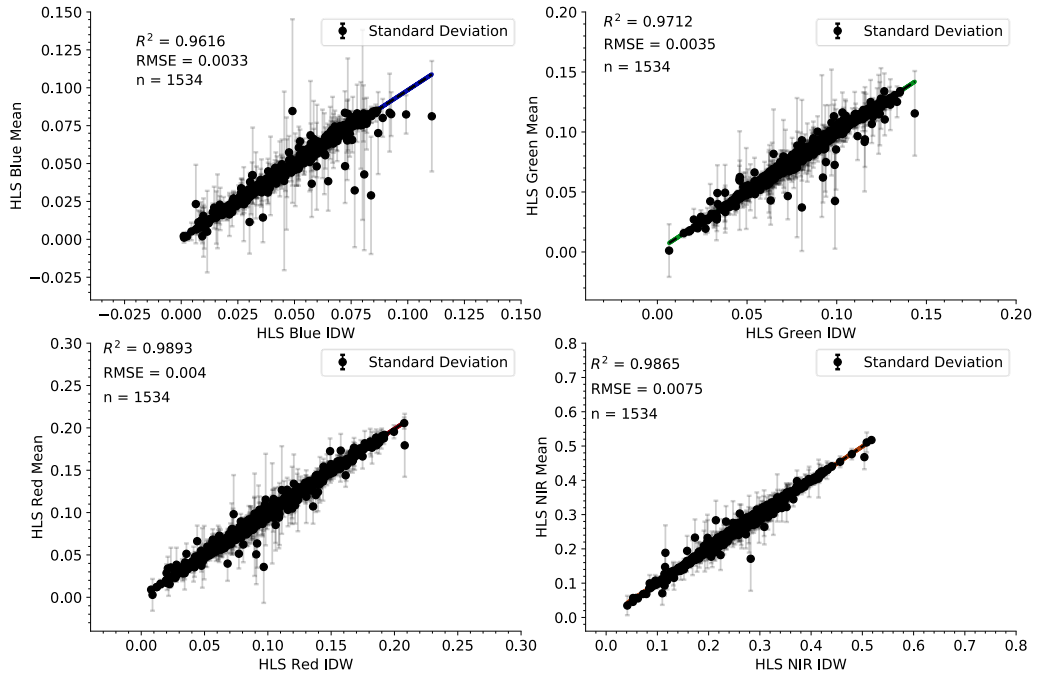


Figure 8: Scatter plot comparing the mean and IDW values for each images red, green, blue, and NIR bands across all six field sites.

2.3.1.2 HLS GP Models

For the six field locations a GP model was produced for each of the HLS VIs (Fig. 9, 10 and Appendix A). The total number of HLS images available from 2013 to 2019 was 1,536 and ranged from 153 at the USGSEROS site to 371 at the Nine-Mile site. Comparing the predicted VIs values from the HLS GP regression models and the actual HLS data calculated using IDW, NDVI had the highest average R2 at 0.82 with a RMSE of 0.073 and ranged from 0.71 at OAES to 0.90 at ARIK. EVI2 was the poorest performing band with an average R2 of 0.73 and a RMSE of 0.044 and ranged from 0.58 at OAES to 0.84 at ARIK.

GCC bands had an average R2 of 0.78, with RMSE of 0.024. EVI2 at the OAES site was the poorest performing model with an R2 of 0.58 and a RMSE of 0.037.

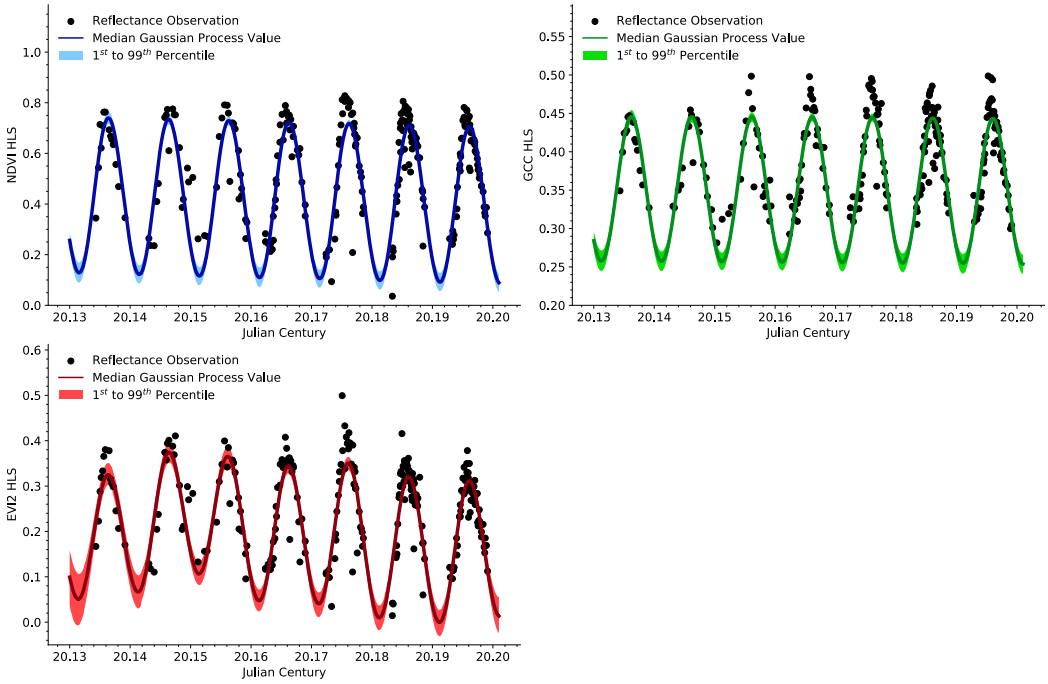


Figure 9: HLS GP model for the Oakville station, showing the median predicted value fit to the seven years of IDW HLS VI data. The 1st to 99th percentile is also shown to depict the regions of uncertainty in the model, these regions tend to widen in areas with little or no HLS data.

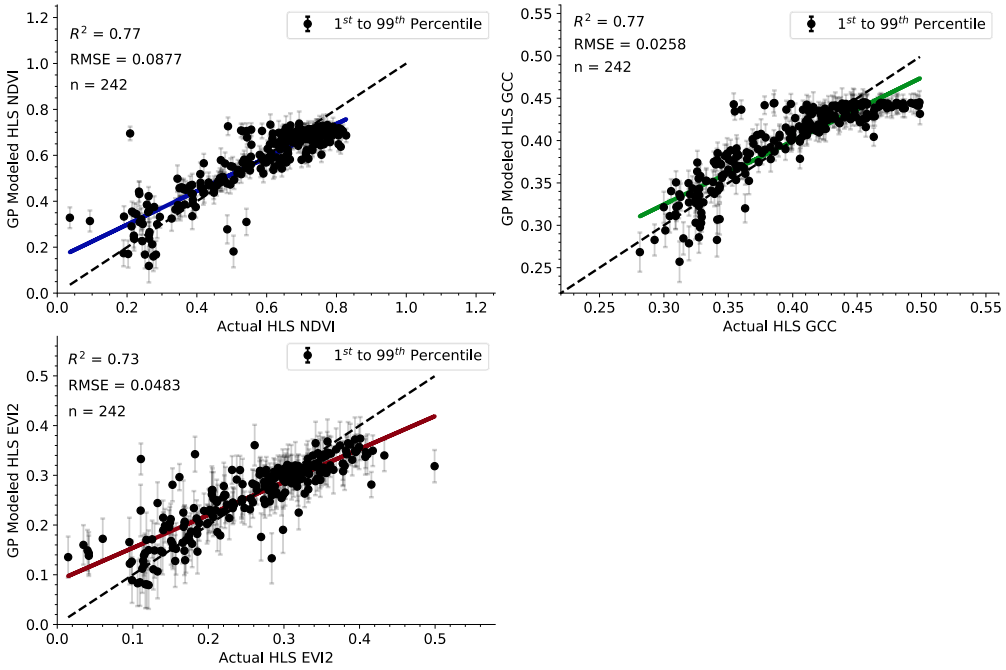


Figure 10: Scatter Plot of Actual IDW HLS VI values against the GP modeled HLS VI values for the Oakville station. The dashed line shows an ideal one-to-one linear fit where y=x (1:1 line).

2.3.2 Phenocam Calculated VIs

Across the six field locations we used the phenocam calculated GCC, NDVI and EVI2 and scaled it linearly with the HLS VIs (Fig. 11). We also compared the GP modeled HLS VI measurements with the phenocam DN 90th percentile data and the scaled phenocam data using a linear regression (Fig. 12, 13 and Appendix B). This allowed us to examine the correlation between the HLS GP models and the phenocam VIs. The correlation measured using R^2 was the same for both scaled and unscaled phenocam VIs since we used a linear transformation. On average GCC had an R^2 of 0.816, NDVI had an R^2 of 0.800, and EVI2 had an R^2 of 0.818. Examining each site ARIK had the highest average R^2 at 0.904, while the

USGSEROS site had the lowest average R^2 of 0.740. The RMSE of these graphs (Fig. 12, 13 and Appendix B) provides a measure to determine how much better the scaled phenocam data aligns with the GP modeled HLS data over the unscaled phenocam data. Across all six phenocam sites, for each of the three VIs the scaled phenocam VIs had a lower RMSE than the unscaled data. GCC had the smallest difference between the unscaled GCC with an average RMSE of 0.044 and the scaled GCC with an average RMSE of 0.023. NDVI had the largest difference in RMSE values with the unscaled NDVI having an average RMSE of 0.648 while the scaled NDVI had an average RMSE of 0.085. For EVI2 the average RMSE for the unscaled data was 0.516, while the average RMSE for the scaled data was 0.037.

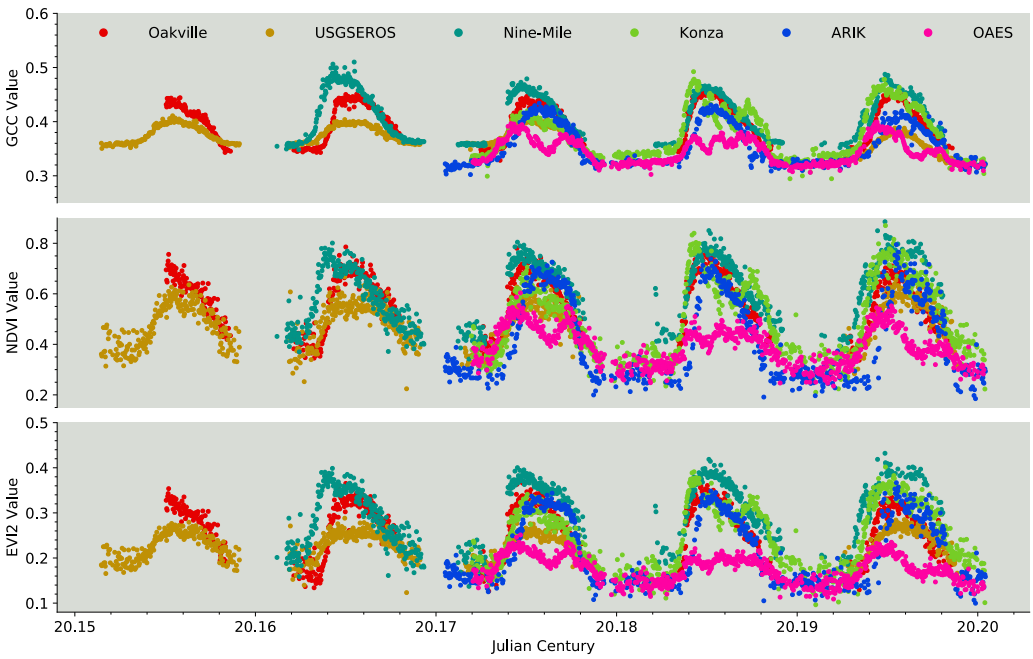


Figure: 11: The three VIs calculated for all six field locations using the scaled phenocam reflectance models. Absolute differences between field sites can be compared since VIs were scaled to the HLS GP VIs.

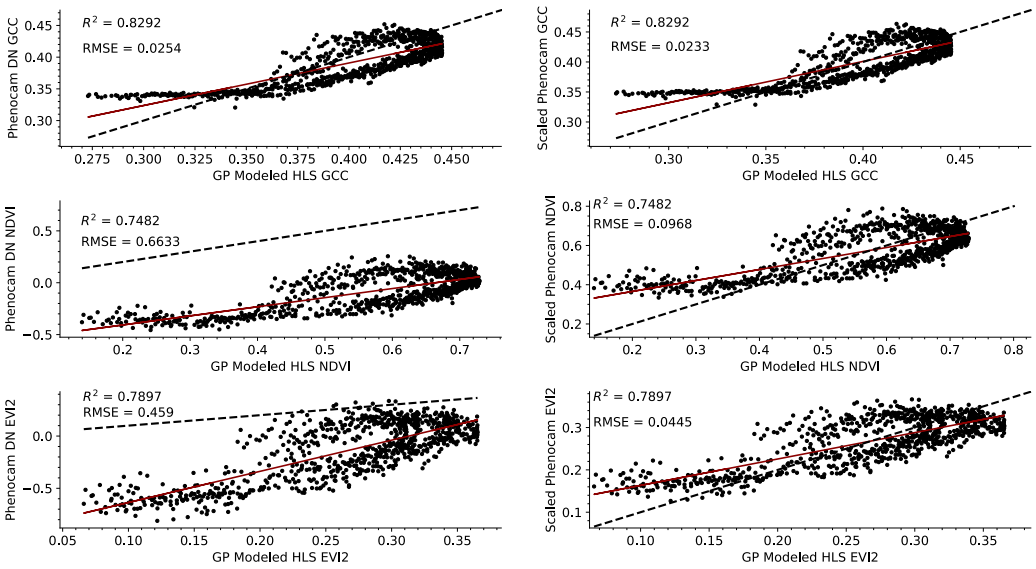


Figure 12: VIs calculated for the Oakville station, showing linear regressions between both the phenocam DN 90th percentile VIs and the scaled phenocam VIs against the scaled GP modeled HLS VIs. The dashed line shows an ideal one-to-one linear fit where y=x (1:1 line).

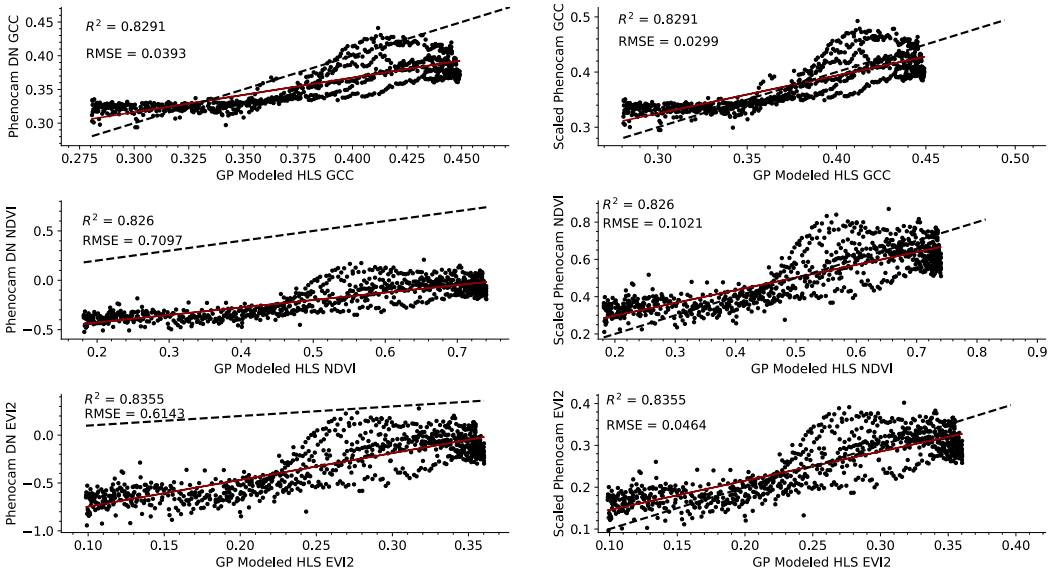


Figure 13: VIs calculated for the Konza station, showing linear regressions between both the phenocam DN 90th percentile VIs and the scaled phenocam VIs against the scaled GP modeled HLS VIs. The dashed line shows an ideal one-to-one linear fit where y=x (1:1 line).

2.4 Discussion

Using GPs, we have developed a novel method to scale phenocam imagery to HLS reflectance values and provided a method to calculate VIs from phenocam imagery, including VIs such as NDVI and EVI2 that incorporate NIR reflectance. With the GP we were able to fit a covariance function (Eq. 5) to GCC, NDVI, and EVI2 for HLS imagery. We conducted this across six grassland field locations, to examine how well this method could be applied at different phenocam sites. Overall, we found our resulting VIs (Fig. 12, 13 and Appendix B) to have a high level of agreement with the GP modeled HLS VIs having an R^2 ranging from 0.67 to 0.91. Also, the scaled phenocam VIs had a smaller RMSE with the GP modeled HLS VIs averaging 0.049 when compared with the unscaled phenocam VIs which averaged 0.403. Across the three VIs the average RMSE was highest for NDVI with an average scaled RMSE of 0.085 and an average unscaled RMSE of 0.648, while GCC had the lowest RMSE with an average scaled RMSE of 0.023 and an average unscaled RMSE of 0.044. The linear models for the unscaled NDVI and EVI2 always fell below the 1:1 line (Fig. 12, 13 and Appendix B), indicating that phenocam NDVI and EVI2 values were always lower than the GP modeled HLS dataset. The linear models for the GCC unscaled phenocam data fell on both sides of the 1:1 line, displaying the robustness of GCC to match satellite-derived indices even without scaling, though on average the RMSE was 0.021 greater than the scaled GCC. The unscaled GCC at the OAES site was an exception since it fell above the 1:1 line. This in not a surprise since the RMSE for the unscaled OAES GCC against the GP modeled HLS was the largest for GCC with a value of 0.108. Scaling the OAES GCC

brought the RMSE down to 0.015 and confirms that scaling the phenocam GCC data is still important to ensure the best alignment with satellite imagery, and to standardise datasets between phenocam sites.

Using the HLS imagery, we used IDW to extract reflectance values that represented the phenocams' FOV. Phenology studies that have made similar comparisons between phenocam and satellite imagery have typically used the average value of a 3x3 or 5x5 pixel window centered over the phenocam site (Filippa et al., 2018; Hufkens et al., 2012). Examining the relationship between the two methods we found a strong agreement with and average R^2 of 0.98 and a maximum RMSE of 0.75 percent across the four bands (Fig. 8). It is likely acceptable to use either method because of the high correlation between them. Using the mean value is easier to implement since it does not require spatial information about the phenocam location and the distance of each pixel to the phenocam. However, this spatial information is readily available for each phenocam site and using the IDW may provide better stability across a less homogenous landcover. While our study focused on grassland field sites that are relatively homogenous, further work should be done to examine if using IDW allows HLS data to be used at more heterogenous sites where pixels may include mixed landcover adjacent to the phenocams' FOV.

The HLS GP regression models provided a way of recovering the underlying VI signal across the HLS imagery while removing the noise (Rasmussen & Williams, 2006). The covariance function fit to the HLS VIs was made up of two kernels, the periodic kernel aligns with the annual growth signal of the vegetation, while the rational quadratic kernel is

able to adapt to changes in the signal from year to year (Camps-Valls et al., 2016). The HLS GP regression models relies on satellite data to extract the surface reflectance values and calculate the VIs used to scale the phenocam VIs. Unfortunately, this means that our developed methods do not likely provide a validation method for satellite-derived VIs since this would result in circularity. However, the linear scaling (Eq. 10) only relies on the magnitude of the HLS data, and the shape of phenocam VIs is not altered by the shape of the HLS data. This means phenophase transition dates can still be derived from our model and compared with satellite sensors, however the magnitude of the VIs cannot be used to validate the magnitude of HLS derived VIs since this would result in circularity. To detect phenophase transition dates methods such as the double logistic function (Elmore et al., 2012; Ren et al., 2018), or the pruned exact linear time method, that has been used with phenocam data (Killick et al., 2012; A. D. Richardson et al., 2019) can be used with our scaled dataset.

With the HLS GP models (Fig. 9, and Appendix A.) we produced three VIs and compared them against VIs calculated using the phenocam imagery (Fig. 12, 13 and Appendix B). Across all three VIs distributions, the HLS GP models had a high level of agreement with the observational data, producing an average R^2 of 0.811. Other studies such as Filippa et al. (2018) used a linear scaling between satellite-derived NDVI and a phenocam NDVI. From the grassland phenocam sites examined they found linear scaling to produce R^2 values ranging from 0.52 to 0.79. Their R^2 values were a little lower then our NDVI R^2 range from 0.735 to 0.905, however they relied on MODIS imagery with a 250m x 250m

resolution, which is a much lower spatial resolution then the HLS imagery, and did not use a GP on the satellite imagery first to remove noise and gap fill the time series. Petach et al. (2014) found an R^2 of 0.89 between their NDVI measurements with a radiometer and a phenocam. The fact that this R^2 is 0.08 higher than our HLS measurement is to be expected since a radiometer measurement taken at the field should have a high degree of agreement with the phenocam. Having our HLS scaling come close to that of a field-based radiometer is a positive sign that our GP modeled HLS was able to align well with the phenocam VIs.

The six grassland sites we selected had relatively unimodal seasonal characteristics across their growing seasons. The Konza and OAES sites (Fig. 11) within the south-central semiarid prairies did however have some unique growth patterns during the 2018 growing season. From early June until October of 2018 the two sites experienced severe to extreme drought conditions that likely influenced the VIs (NDMC et al., 2020). These drought conditions likely produced the rapid decrease in the VIs that can be seen in the phenocam VIs after their peak values in the spring. The presence of the drought characteristics in the VIs is evidence of the phenocams' VIs abilities to fit the data, without over generalizing these growth characteristics. We believe this gives strong support for future work using this method to focus on many of the remaining phenology knowledge gaps such as scaling phenocam metrics with other sensors, and in working with climate data to further understand the driving forces of phenological observations (Piao et al., 2019).

2.5 Conclusions

The aim of our study was to align phenocam VI values with HLS imagery. The model we developed likely can be applied at any phenocam site that has HLS imagery available and may be further applicable with other satellite SR products for scaling the phenocam bands. Our HLS GP models were able to fit well with the phenocam observations. Using our model additional VIs that rely on the RBG and NIR portions of the electromagnetic spectrum can also be calculated. By scaling the VIs with VIs calculated from measures of SR we are further able to use phenocam imagery to make comparisons between phenocam measurements at different sites or make comparisons with VIs measured using satellite sensors. This allows for not only VI shape to be compared, but the magnitude of the VIs can also be used when examining phenophase values or for using phenocams to validate satellite-derived VIs other than the HLS data we used as a model input (Hufkens et al., 2012; Robinson et al., 2017).

III. Modelling Vegetation Phenology at Six Field Stations within the U.S. Great Plains: Constructing a 38-year Timeseries of GCC, VCI, NDVI, and EVI2 Using Phenocam Imagery and DAYMET Meteorological Records

3.1 Introduction

Grasslands cover approximately 59 million km^2 of the Earth's surface (Hufkens et al., 2016) making up between 10 and 30 percent of the global carbon stock (Scurlock & Hall, 1998); this makes grasslands the second largest carbon sink after forests (Anderson, 1991). In North America, the Great Plains cover approximately 2.9 million km^2 within an east-to-west gradient of tall to short-grass prairie. However, the conversion of grassland to cropland has drastically reduced the remaining native prairie ecosystems. In 2018, it was estimated that only half of these grassland ecosystems remain, with 87 percent of them located on poor and marginal quality soils (World Wildlife Fund, 2018). The variation within the Great Plains creates a variety of community types typically dominated by C_3 grasses in the north and east (more precipitation and cooler temperatures), and C_4 grasses in the south and west (lower precipitation and higher temperatures) (Petrie et al., 2016). The C_3-pathway for photosynthesis is common in temperate regions in grasses such as wheatgrass (*Agropyron*), bentgrass (*Agrostis*), and foxtail (*Alopecurus*), while the C_4-pathway is common in arid regions where the weather is typically hotter and drier with grasses such as bluestem (*Bothriochloa*), threeawn (*Aristida*), and grama (*Bouteloua*) (Jones & Vaughan, 2010; Stubbendieck et al., 2017). Along with a large amount of spatial variability, grasslands are also characterized with high amounts of temporal variability (Flanagan & Adkinson, 2011).

This means that climate change induced shifts in grassland phenology will likely only be detectable using long-term monitoring over several years to decades (G. M. Henebry, 2013).

Across the northern hemisphere the onset of spring is predicted to occur earlier under the warmer conditions predicted from climate change (Mark D. Schwartz et al., 2006). However, a controlled test of grassland phenology using plants grown within a warmer temperature, elevated CO_2, increased nitrogen, and increased precipitation has shown an array of responses that were not all anticipated. Additions of CO_2 delayed spring greenness while increased nitrogen slowed down the growth acceleration. Precipitation had no effect, suggesting it was not a limiting factor for the controlled plants, while increased temperature was the only factor to have the expected outcome, causing plants to flower earlier by 2-5 days (E. E. Cleland et al., 2006). Field observations of arid grasslands using both phenocams (A. D. Richardson et al., 2018) as well as satellite imagery are also in agreement that warmer temperatures bring an earlier start of season to the grasslands. However, in an arid environment precipitation has been found to influence the recorded vegetation indices (VIs), even causing a second peak of greenness in the growing season after a large precipitation event (Browning et al., 2017). Even though precipitation may have a small influence on grassland phenology, modeled scenarios under future climate conditions still suggest that North America will see an increase in both the length of the growing season and the productivity of grasslands. This is because the modeled grasslands are expected to become more efficient in retaining moisture under higher CO_2 levels, allowing for a more efficient use of water and a reduction in the amount of water lost in transpiration (Hufkens et al.,

2016). This suggests that precipitation needs to fall below a threshold before it has a noticeable effect on growing season length (Browning et al., 2017).

Identifying the limiting factor for growth of grassland phenology is a challenging task, with factors such as temperature and precipitation fluctuating throughout the growing season to limit plant growth (J. Wang et al., 2003). Many phenology models still rely on temperature as the primary limiting factor to growth, and because of this they under-perform by not recognizing the importance of photoperiod and water availability (Piao et al., 2019). Temperature-driven models may fail to help predict future phenology patterns from climate change since plants can have a reduced sensitivity to temperature (Fu et al., 2015). Instead, new models should be developed to account for the interactions between the many environmental factors that drive plant growth.

Machine learning has gained traction in Earth sciences and ecology, with many machine learning models outperforming traditional statistical models (Dai et al., 2019). Machine learning algorithms apply non-linear techniques that can often identify complex underlying relationships in the data (H. Zhang et al., 2019). Regardless of these advantages, there are few phenology models that take advantage of the benefits provided by machine learning (Dai et al., 2019). One recently developed machine learning algorithm, known as XGBoost (XGB), is a gradient boosted decision tree model capable of both regression and classification tasks (Chen & Guestrin, 2016). Improvements made in XGB make it more robust at handling noise, as well as dealing with unbalanced and skewed datasets (H. Zhang et al., 2019). This makes it an excellent choice when working with empirical data that often

fails to meet the requirements of parametric statistical analysis. However, using machine learning for phenology requires long timeseries datasets with few data gaps, although, even then, there analysis can be challenging when noise is present in the data (Belda et al., 2020).

Phenocams are digital web-enables cameras that are capable of imaging ecosystems with high temporal resolution (A. D. Richardson, 2019). Phenocams record changes in vegetation throughout the growing season by capturing multiple images per day using the visible, and sometime the near-infrared portions of the electromagnetic spectrum. Changes in vegetation phenology are known as phenophases and include green-up in the spring, and senescence in the fall (A. Richardson & Braswell, 2009). The imagery captured by phenocams is used to calculate VIs that record changes in vegetation growth, and have been used to calculate other growth indices such as leaf area index (Keenan et al., 2014). The VIs calculated from phenocam imagery can also be used to record changes in the timing of phenophase transitions to detect how vegetation is responding to changes in their local environment, such as changes brought on by climate change (Elmore et al., 2012; Killick et al., 2012; Ren et al., 2018). Four VIs that are prominent in phenology research include the green chromatic coordinate (GCC) (A. Richardson & Braswell, 2009), the vegetation contrast index (VCI) (X. Zhang et al., 2018), the normalized difference vegetation index (NDVI) (Rouse et al., 1973) and the two-band enhanced vegetation index (EVI2) (Jiang et al., 2008).

The high temporal availability of phenocam imagery makes it a suitable data source for machine learning analysis. Also, the need for phenology models capable of detecting the underlying relationships between many environmental factors makes machine learning an

important method to consider for the development of new models. The North American Great Plains provide an interesting study area to examine the interactions of different meteorological variables because of the spatial gradients that exist in temperature and precipitation. Because of this we decided to: 1) develop a regression model using XGB that can predict GCC, VCI, NDVI and EVI2 values using meteorological data at multiple grassland phenocam locations, 2) determine the primary meteorological variables within the model, and how these differ between VIs, and 3) predict the four VIs and measure their phenophases to establish trends in phenophase transitions using 38 years of historic meteorological data.

3.2 Methods and Data

3.2.1 Study Area

The Great Plains of North America occupy 281 million ha with 224 million ha located within the contiguous U.S. (U.S. Environmental Protection Agency, 2020). The Great Plains Ecoregion is subdivided into Level-2 regions that represent the diversity within the Great Plains (Fig. 14). The temperate prairies in the east are wetter and contain more croplands than the drier west-central and south-central semiarid prairies, while the west-central semiarid prairies are on average cooler than south-central semiarid prairies (Omernik & Griffith, 2014).

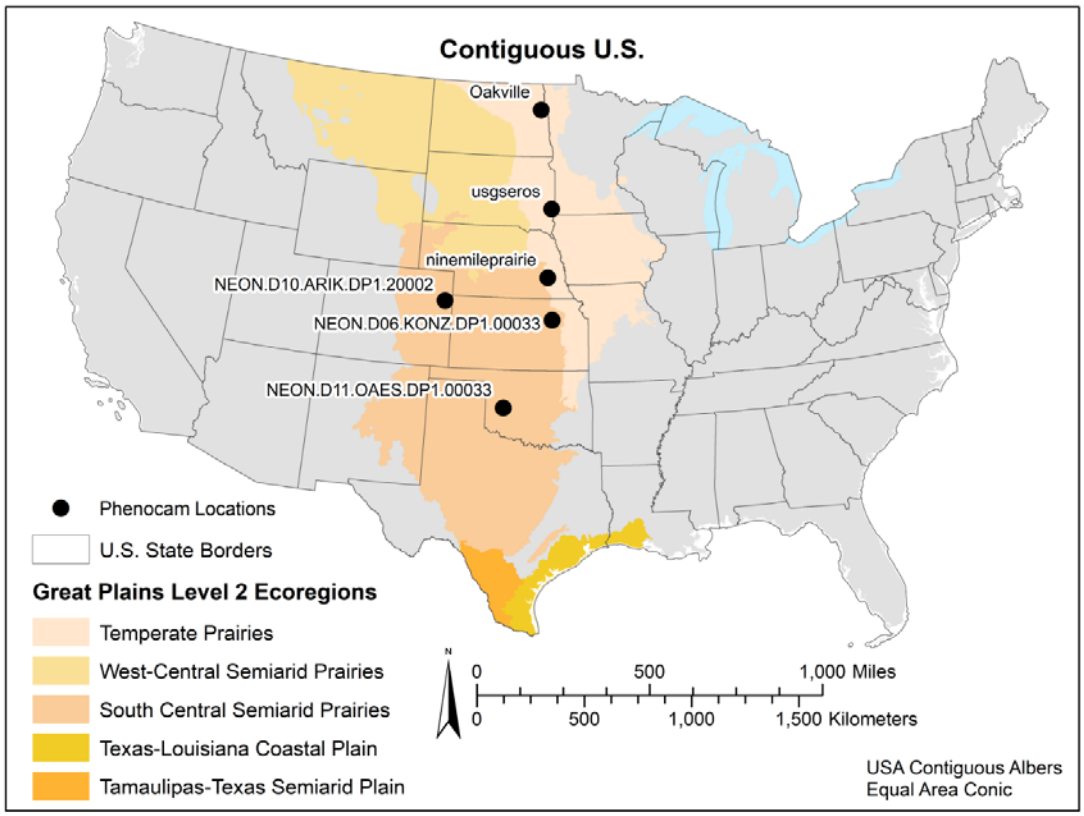


Figure 14: The phenocam locations within the study area. Showing the six phenocam locations situated within the Great Plains of the contiguous U.S. Figure taken from Burke and Rundquist (2021).

We selected six grassland locations within the Great Plains of the contiguous U.S. (Fig. 14) each of which has a phenocam with at least three years of data (Table 1). Three of the sites are located within the temperate prairie ecoregion; the Oakville Prairie (Oakville), a part of the University of North Dakota, located in Grand Forks County, North Dakota (47.8993°N, 97.3161°W); the USGSEROS station at the Earth Resources Observation and Science (EROS) Data Center in South Dakota (43.7343°N, 96.6234°W); and the Nine Mile Prairie station (Nine-Mile), a part of the University of Nebraska – Lincoln (40.8680°N, 96.8221°W), located in Lancaster County, Nebraska. The other three phenocam sites are

within the south-central semiarid prairie and are a part of the National Ecological Observatory Network (NEON). These sites include the NEON.D06.KONZ.DP1.00033 station (Konza) (39.1008°N, 96.5631°W) located at the Konza Prairie Biological Station in Kansas; the NEON.D10.ARIK.DP1.20002 station (ARIK) (39.7582°N, 102.4471°W) located near the Arikaree River in Yuma County, Colorado; and the NEON.D11.OAES.DP1.00033 station (OAES) (35.4106°N, 99.0588°W) located at the Klemme Range Research Station in Washita County, Oklahoma. The six sites represent a 1,470 km latitudinal transect through the Great Plains region ranging from 35.4°N to 47.9°N.

Table 1: Years of data available for each of the phenocam site locations.

Station Name	Years with available data
Oakville	2016 - 2019
USGSEROS	2015 - 2017, 2019
Nine Mile Prairie	2016 - 2019
NEON.D06.KONZ.DP1.00033	2017 - 2019
NEON.D10.ARIK.DP1.20002	2017 - 2019
NEON.D11.OAES.DP1.00033	2017 - 2019

3.2.2 Phenocam Data Source and Calculating the VIs

We choose to derive four VIs from the phenocam imagery at the six field stations. GCC (Eq. 1) is a proportional measure of relative ‘greenness’ that was originally developed for use with phenocams because of its relative stability under changing illumination conditions (A. Richardson & Braswell, 2009). GCC has be used in a diverse array of ecosystem types, and can be measured using any digital camera capable of capturing a color

(red, green, and blue) image (A. D. Richardson, 2019). VCI (Eq. 2) was created as a nonlinear transformation of GCC that has a higher dynamic range relative to GCC by contrasting the green band to the sum of red and blue (X. Zhang et al., 2018). NDVI (Eq. 3) has a long history in Earth Observation (Rouse et al., 1973), and has been derived from phenocams that are sensitive to near-infrared wavelengths (Burke & Rundquist, 2021; Filippa et al., 2018; Petach et al., 2014; A. D. Richardson, 2019). EVI2 (Eq. 4) was developed as an adjustment to NDVI, with an enhanced ability to remove soil background noise, and atmospheric effects (Jiang et al., 2008).

$$GCC = \frac{Green}{Blue + Green + Red} \quad (1)$$

$$VCI = \frac{Green}{Blue + Red} \quad (2)$$

$$NDVI = \frac{NIR - Red}{NIR + Red} \quad (3)$$

$$EVI2 = 2.5\frac{NIR - Red}{NIR + 2.4 * Red + 1} \quad (4)$$

To calculate each of the chosen VIs from the phenocam imagery, we first downloaded all available imagery from the six phenocam locations (phenocam.sr.unh.edu). We then applied the exposure correction to both the color and mixed color-infrared imagery to extract the near-infrared and three color bands (Petach et al., 2014). Using the image digital numbers (DNs) for the red, green blue (RGB) and near-infrared (NIR) bands the three VIs were calculated using Eq. 1, 2, 3 and 4 for each day of the year in which phenocam imagery was available (Table 1). Finally, the phenocam VIs were linearly scaled to Gaussian Process

Regression modeled VIs calculated with Harmonized Landsat-Sentinel surface reflectance imagery (described in detail in Burke and Rundquist, 2021). This standardised the VI values between all phenocam sites, allowing them to be used together within a single XGB model.

3.2.3 Meteorological Data

We used Daily Surface Weather and Climatological Summaries (DAYMET) data made available by the Oak Ridge National Laboratory (ORNL) within the Distributed Active Archive Center (DAAC) (Thornton et al., 2018). DAYMET provides 1km x 1km gridded data for North America starting in 1980, with several different weather variables available (Table 2). We retrieved the data for each of the six phenocam locations (Fig. 14), for the phenocam imagery time periods (Table 1).

We also used the DAYMET data to derive a few accumulative variables for precipitation, snow water equivalent (SWE) and temperature. Previous research has shown that precipitation often has a lag period before its has a measured effect on a VI's signal (Potter & Brooks, 1998; J. Wang et al., 2003). Based on this research we decided to accumulate precipitation over both 15 and 30 days to see if this would have a stronger relationship with the VI signals compared with the daily total precipitation. We did the same with the SWE, except changed the lag periods to 60 and 90 days to reflect the longer lag periods for snowfall. To calculate these values, we summed together the precipitation or SWE for the set number of days prior to each day of the year. To estimate the accumulated heat for vegetation growth we used growing degree days (GDD) calculated for each day of the year (Eq. 4) (Burke et al., 2018). GDD have historically been used for predicting

agricultural crop growth and development, with Tbase set at 0°C for winter wheat a C_3 plant and 10°C for corn a C_4 plant (McMaster & Wilhelm, 1997). We choose to calculate GDD for three Tbase values set at 0, 5 and 10°C and examine the relationship these three datasets have with our grassland VIs. This resulted in a total of 13 variables being included in our model.

$$GDD = \frac{Tmax + Tmin}{2} - Tbase, if \frac{Tmax + Tmin}{2} > Tbase \quad (4)$$
$$GDD = 0, otherwise$$

Table 2: DAYMET daily surface weather data variables used to model the phenocam VIs, including both DAYMET provided data and the variables derived from the DAYMET data, such as SWE and GDD.

Data Field	Description	Units	Source
Dayl	Duration of the daylight period for the day	seconds	DAYMET
Prcp	Daily total precipitation	mm	DAYMET
Srad	Incident shortwave radiation flux density	w/m^2	DAYMET
Swe	Snow water equivalent	kg/m^2	DAYMET
tmax	Daily maximum 2-meter air temperature	degrees Celsius	DAYMET
Tmin	Daily minimum 2-meter air temperature	degrees Celsius	DAYMET
acc prcp 15	Accumulated precipitation over 15 days	mm	Derived
acc prcp 30	Accumulated precipitation over 30 days	mm	Derived
acc swe 60	Accumulated SWE over 60 days	kg/m^2	Derived
acc swe 90	Accumulated SWE over 90 days	kg/m^2	Derived
gdd 0	GDD with Tbase = 0 degrees Celsius	degrees Celsius	Derived
gdd 5	GDD with Tbase = 5 degrees Celsius	degrees Celsius	Derived
gdd 10	GDD with Tbase = 10 degrees Celsius	degrees Celsius	Derived

3.2.4 Statistical Analysis of Daily VIs

To produce a regression model for the four VIs we used XGB, a gradient boosted decision tree model (Chen & Guestrin, 2016). We trained our XGB models using a randomly selected 80 percent (n = 2,815) of the available data, leaving 20 percent (n = 704) for model validation. To help prevent overfitting of the model, and to prune any branches with a negative gain, we set lambda to 1 and both alpha and gamma to 0. We also set the learning rate to 0.1, max depth to 10 and number of estimators to 50,000. We choose parameters that would help prevent overfitting of the model, and were recommended to produce a more conservative algorithm (Chen & Guestrin, 2016). Subsampling, also know as bootstrap aggregating, was used so that a random selection of half (subsample = 0.5) the training samples were used to grow each tree with gradient-based selection (Chen & Guestrin, 2016; H. Zhang et al., 2019).

Using the XBG model we fit each of the VIs against all the meteorological data variables including the accumulated precipitation, accumulated SWE and GDD. We combined the data sets across all six phenocam sites and created a model that could predict the four phenocam-based VIs at any one of the grassland sites given the daily meteorological data. By examining the total gain, a relative measure of a variable's contribution to the model, we refined each of the VIs models further by removing the variables with the lowest total gain in a stepwise fashion until the R^2 declined by more than 3 percent from the first model containing all variables, then selecting the model directly before the 3 percent decline. We used 3 percent as a threshold to minimize loss of model performance, while allowing

enough of a reduction to the model to remove the variables that added little prediction power. Using the refined models for each of the four VIs we used the meteorological data to predict the VI values for each day of the year starting in 1981 and ending in 2019, producing a dataset for each VI ranging 38 years for each of the six phenocam locations.

3.2.5 Determining Phenophase Transitions Dates

Using the 38 years of data for the four modeled VIs at the six phenocam locations we identified phenophase transitions dates using the same methods applied to the Collection 6 Moderate Resolution Imaging Spectrometer (MODIS) Land Cover Dynamics Product (CMCD12Q2) (Gray et al., 2019). The CMCD12Q2 product identifies seven phenophase stages throughout a growth cycle (Fig. 15), starting with greenup in the spring and ending with dormancy in the fall. This procedure was completed 24 times to account for the four VIs at 6 different sites. A natural cubic spline (Drury, 2020) was fit to the full 38-year time series. To find the optimal number of knots to fit the spline we used Akaike's Information Criterion (AIC) to balance under-overfitting of the model (Hurvich et al., 1998). To do this we randomly set aside one third of the dataset and fit the spline starting at 38 knots (1 knot per year of data) and ending at 570 knots (15 knots per year of data). Using the AIC we measured the models fit against the randomly removed data and selected the number of knots that produced the lowest AIC value. The spline was then re-fit to the entire dataset using the determine optimal number of knots.

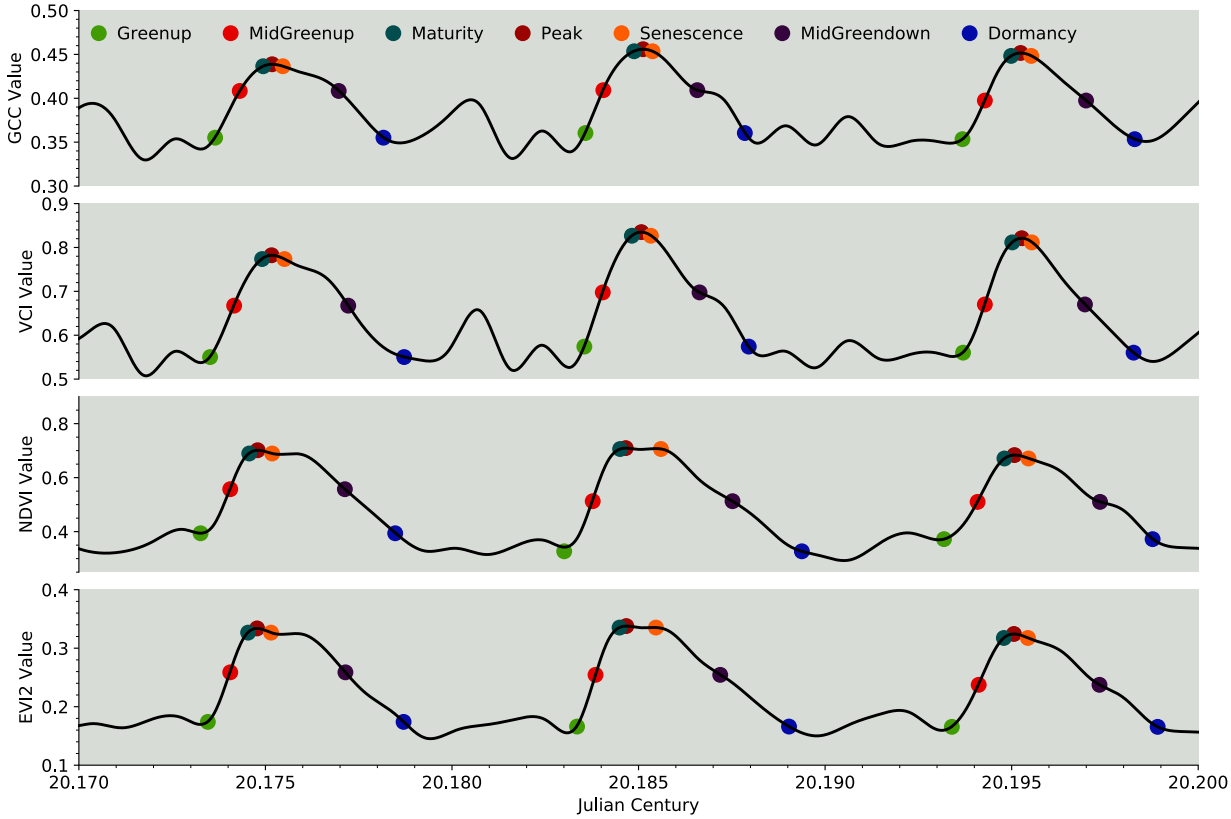


Figure 15: Phenophase transitions dates for the four VIs at the Oakville station determined using the same methods applied to the CMCD12Q2 product. The colored circles denote the beginning of their corresponding phenophase. The graph shows three years of data (2017-2019) taken from the modeled 38-year dataset.

Valid vegetation cycles were identified from the 24 spline models using methods similar to the CMCD12Q2 product (Gray et al., 2019). Local minima and maxima were identified for each year with a half year overlap at the beginning and end of the year. The maxima were examined for validity as a peak in vegetation growth while the minima were examined to be either the start or end of a vegetation cycle. However, the methods used for the CMCD12Q2 product was produced for EVI2 specifically and has a set value of 0.1 required in the amplitude of any greenup or greendown period for it to be considered a valid

cycle. The three other VIs have a varying range of values that do not necessarily align with EVI2. Instead of using a constant value of 0.1, we modified this step by requiring greenup and greendown periods to have an amplitude that is at least 70% that of the current year's amplitude. Once the valid growth periods were identified we then extracted the seven phenophase periods using the same methods as the CMCD12Q2 product. The peak is reached at the maximum value for the VI. The greenup, mid-greenup, and maturity occur at a 15, 50, and 90 percent increase in amplitude, while senescence, mid-greendown, and dormancy occur after the peak as amplitude decreases past 90, 50, and then 15 percent. Using these values, we also measured the length of greenup, the number of days between greenup and maturity, the length of maturity, the number of days between maturity and senescence, and the length of greendown, the number of days between senescence and dormancy, and the length of season, the number of days between greenup and dormancy.

3.3 Results

3.3.1 XGB Regression Models

Using the GCC, VCI, NDVI, and EVI2 datasets we produced four XGB regression models capable of predicting the VIs value based on all variables within the meteorological DAYMET data (Fig. 16). For each of the VIs a total of 2,815 data points were used in model training, while 704 data points were set aside for model validation (Fig. 16). Examining the validation results GCC was the best fitting model with an R^2 of 0.946 and a root mean square error (RMSE) of 0.01, while EVI2 was the lowest with an R^2 of 0.8954, and an RMSE of 0.02. Examining the total gain for each of the variables in the four models provides a relative

measure of importance. Across all four models the photoperiod as day length, and temperature as GDD with a base of 0°C were the two most important variables. While the minimum temperature and 30-days of accumulated precipitation were the third and fourth most important variables (Fig. 17). These four variables had the highest total gain across all four VIs, however they did not all occur in the same order. For example, day length had the highest total gain for GCC and VCI while GDD with a base of 0°C was the highest for NDVI, and EVI2.

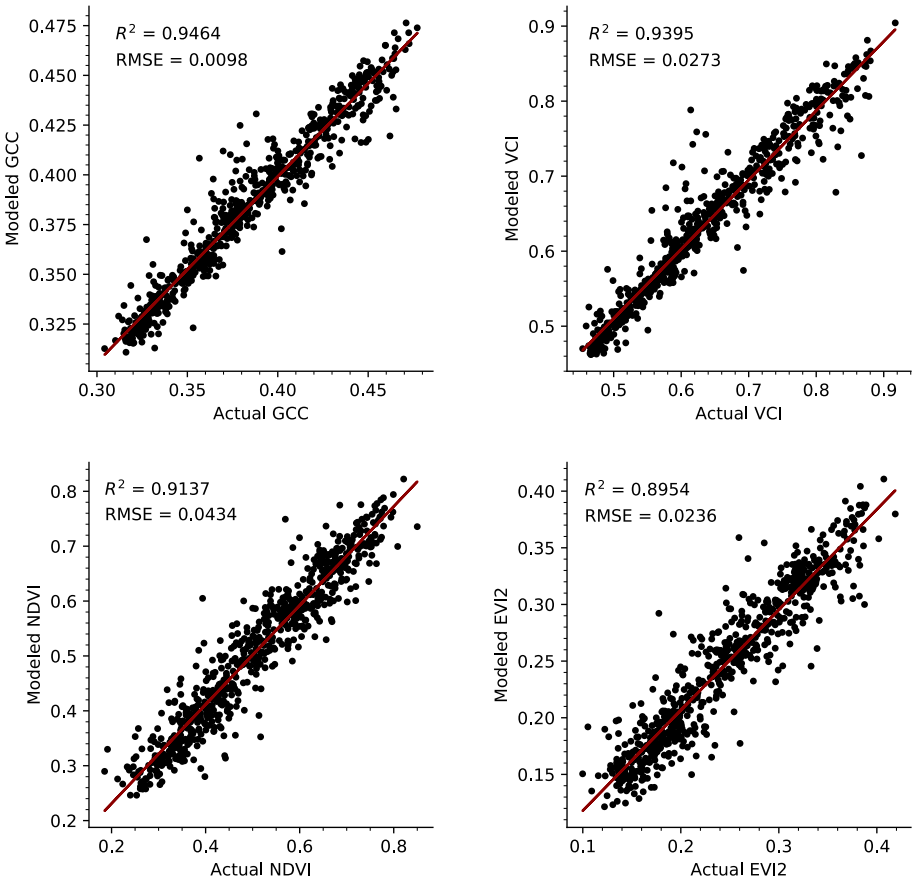


Figure 16: The four XGB modeled VIs against the validation datasets, showing the models ability to predict the VIs values given all 13 meteorological variables.

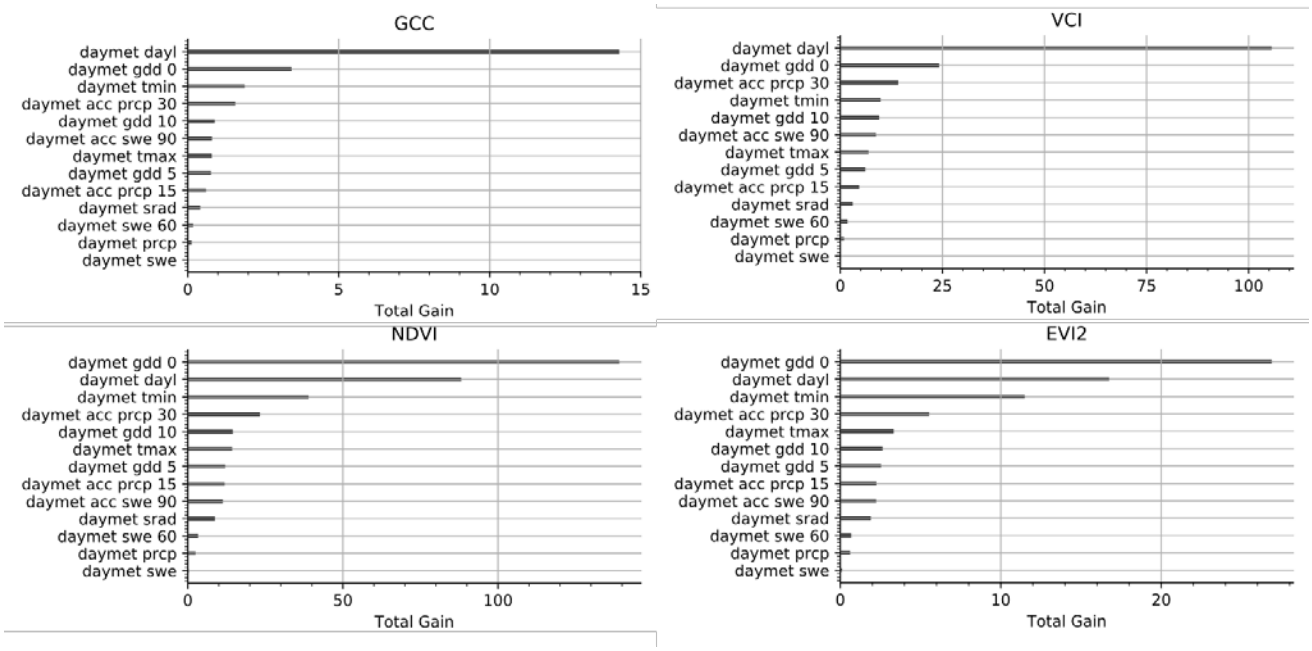


Figure 17: The total gain for each of the 13 meteorological variables used in the four XBG models.

3.3.2 Reducing the XGB Regression Models

With each of the four XGB regression models we removed variables one at a time for each VI independently, starting with the variable with the lowest total gain. We then refit the XGB models and assessed them with the validation dataset. We continued to remove variables until the R^2 value of the validation dataset decreased by greater than 3 percent from the XGB models that contained all 13 meteorological variables, then selected the previous model. For the GCC and VCI XGB models this resulted in a final model using only four variables: day length, GDD with a base of 0°C, 30-days of accumulated precipitation, and GDD with a base of 10°C (Fig. 18). For the NDVI and EVI2 XGB models the final model required five variables: GDD with a base of 0°C, day length, daily minimum temperature, 30-day accumulated precipitation, and GDD with a base of 5°C (Fig. 18). These four XGB

models were able to account for between 89.6 and 93.1 percent of the variation in the VIs datasets given 6 of the 13 meteorological variables (Fig. 19).

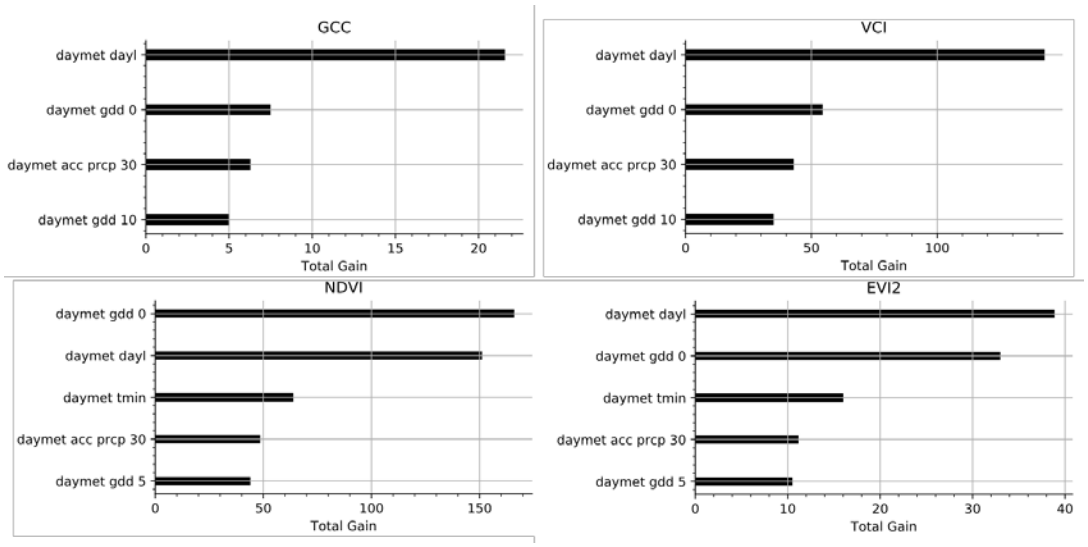


Figure 18: The total gain for the remaining variables used in the reduced XGB models for each of the four VIs.

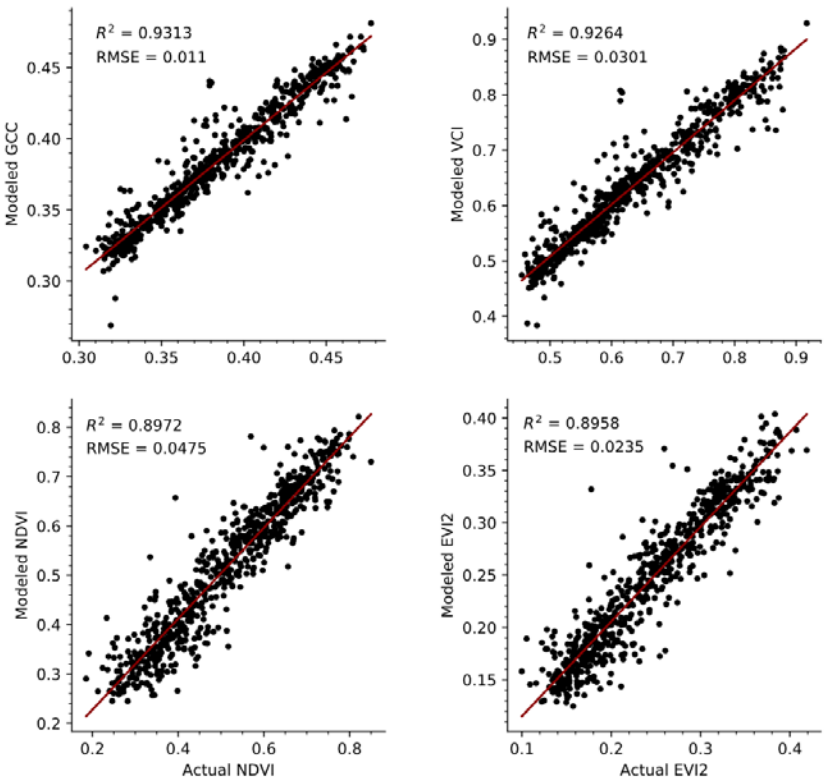


Figure 19: The reduced XGB modeled VIs against the validation datasets. For GCC and VCI four meteorological variables were used, while for NDVI and EVI2 five of the variables were used.

Using the four reduced VIs XGB regression models we conducted a sensitivity analysis to determine how a change in any of the variables effects the resulting VI value (Fig. 20). To do this we calculated the minimum, maximum and mean values for each of our variables, and then predicted the VI value at 100 evenly spaced sample points between each variable's minimum and maximum while holding all other variables at their mean value. This analysis shows many of the nonlinearities between the meteorological variables and the VIs. For example, across all four VIs an increase in the lower values (< ~1,000) of GDD 0°C tends to cause an increase in the VI value. However, as GDD 0°C increases (> ~1,000),

eventually the VI value either reaches a plateau or the VI starts decreasing as GDD 0°C increases.

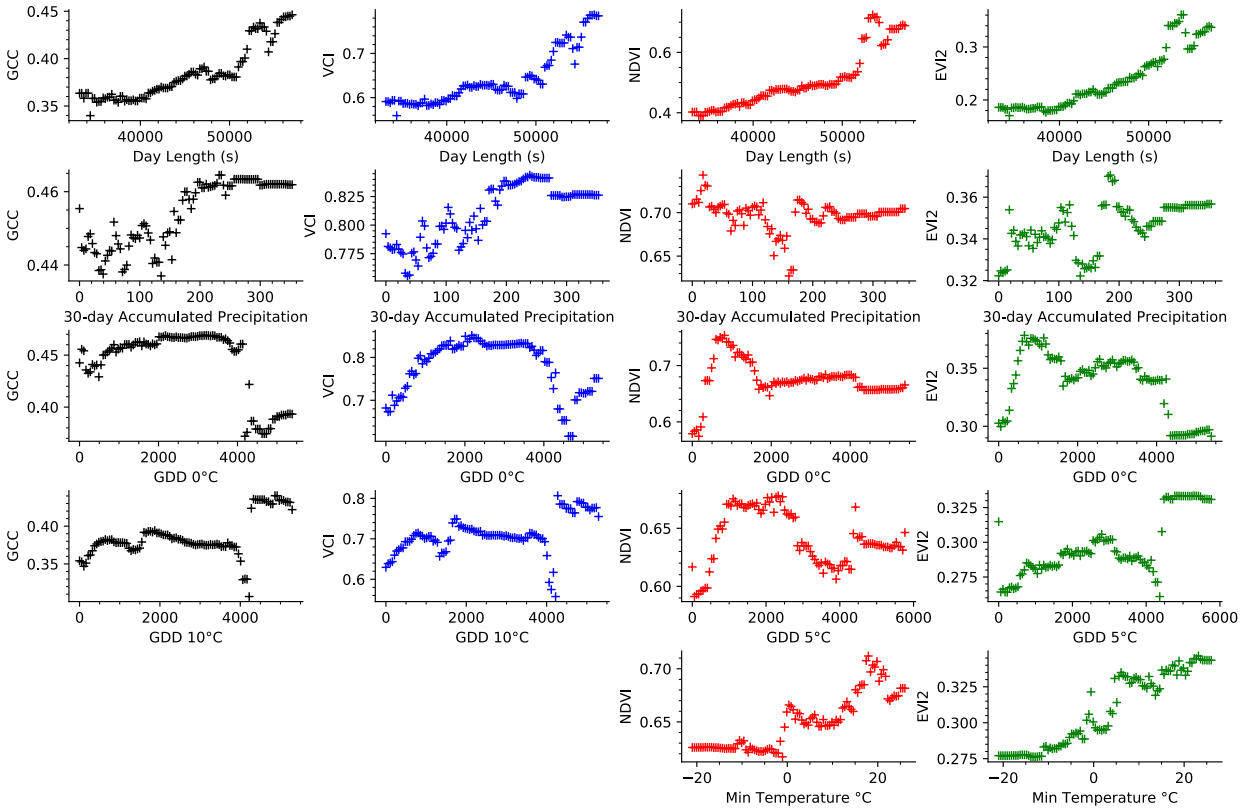


Figure 20: Sensitivity analysis showing how the variables in the four reduced XGB models effect the VIs values as their value is increases from it minimum to maximum value while all other variables are held at their mean value.

3.3.3 Trends in Phenophase Transitions

Using the XGB models with the 38 years of meteorological data we predicted the four VIs values for each day of the year. Then using these predictions splines were fit for the four VIs across the six phenocam locations. For example, at the Oakville station a spline model was fit to the predicted NDVI values (Fig. 21). Comparing the XGB predicted values with

the spline models, we found that the splines were able to align well with an R^2 and a RMSE ranging from 0.83 and 0.017 for GCC to 0.92 and 0.039 for NDVI (Fig. 22). Noticeably the spline did reduce extreme values within the predicted VI values, for example in GCC where XGB predicted values below 0.2 were closer to 0.3 in the spline models. We examined the quantile range for both the XGB models and spline models and found little difference between the 1st, 2nd, and 3rd quantile for the two models, while the minimum and maximum values for the spline models were always closer to the median than the XGB models (Table 3).

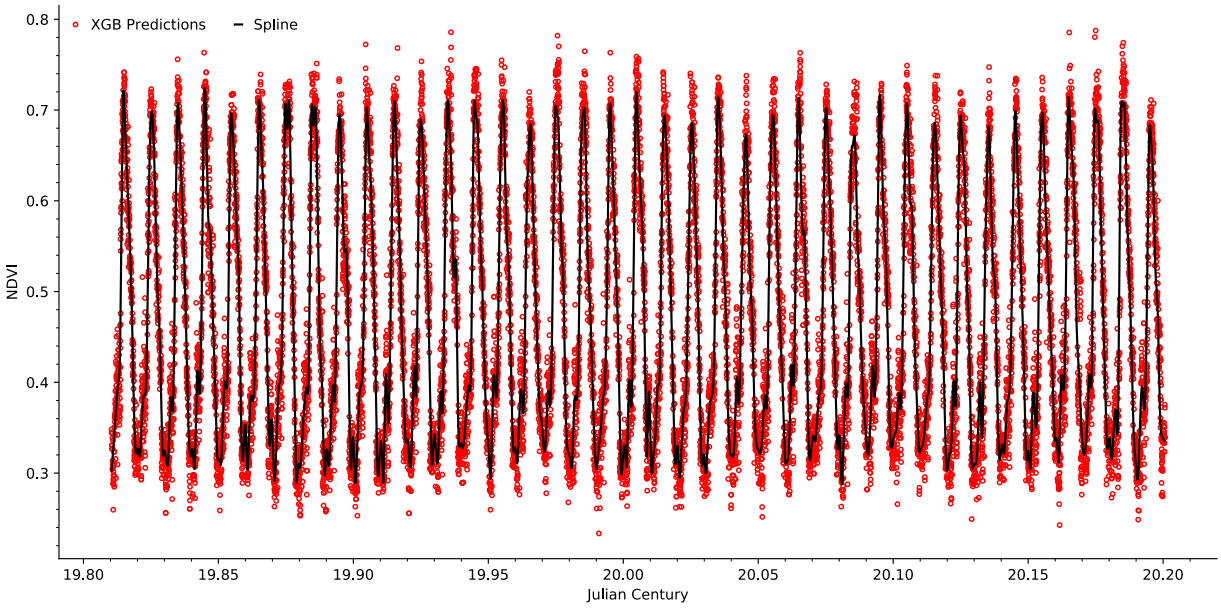


Figure 21: The XGB predicted NDVI values for the Oakville phenocam, using the meteorological data starting in 1981 to 2019, covering 38 years. The solid line depicts the spline fit to the model predictions showing the yearly vegetation cycles.

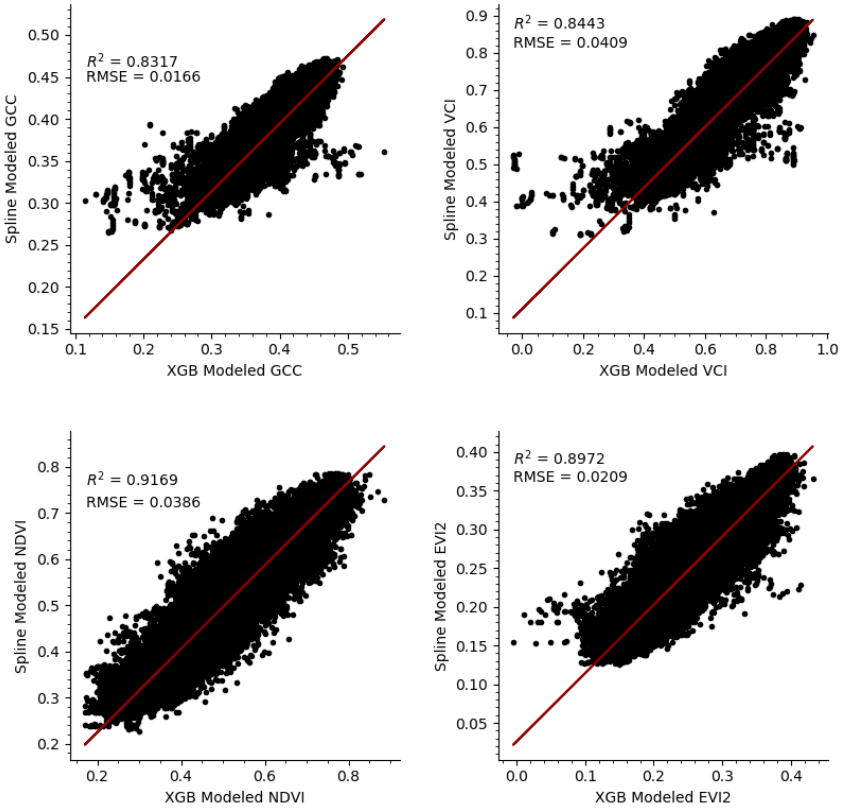


Figure 22: Scatter plot showing the relationship between the XGB modeled VIs and the splines fit to the vegetation cycles. This includes all six of the spline models for each phenocam location across the four VIs.

Table 3: The quantile range of the XGB regression models and the spline models for the four VIs. The 1st, 2nd, and 3rd quantiles of the two model types have very little difference, while the minimum and maximum values of the spline are always closer to the median value then the XGB model.

Percentile	XGB Model (GCC, VCI, NDVI, EVI2)	Spline Model (GCC, VCI, NDVI, EVI2)
Maximum	0.554, 0.954, 0.886, 0.432	0.472, 0.890, 0.785, 0.397
3rd Quantile	0.396, 0.655, 0.561, 0.270	0.394, 0.655, 0.559, 0.269
Median (2nd Quantile)	0.361, 0.576, 0.420, 0.203	0.361, 0.571, 0.417, 0.201
1st Quantile	0.339, 0.525, 0.344, 0.168	0.341, 0.530, 0.347, 0.170
Minimum	0.114, -0.028, 0.170, -0.005	0.265, 0.310, 0.228, 0.125

For each of the spline models we predicted seven day of year (DOY) values as phenophases occurring within the vegetation growth cycles. We also calculated the length of greenup, the length of maturity, the length of greendown, and the total length of season, as the number of days between the greenup, maturity, senecence, and dormancy DOY values, respectively. This allowed us to examine trends in the seven phenophases to determine if over the 38-year data period they are occurring earlier of later in the growth cycle, and to determine if the lengths of time between them is increasing or decreasing. We calculated 66 linear regressions (Appendix C), one for each phenophase and length between them at the 6 phenocam locations. Of these linear regressions we found 14 to have a significant trend within a 90 percent confidence interval (Table 4). The slope of these linear models provides us the change per year in each of the phenophases. For example, at the Oakville phenocam the dormancy phenophase produced a slope of 0.27, suggesting that dormancy is occurring 0.27 days later every year, which across our 38 years of data results in dormancy occurring 10 days later in 2019 compared to 1981.

Table 4: The linear regressions for the phenophases that had a significant trend within a 90 percent confidence interval across the 38-year data period. A 95 percent confidence interval is denoted by a *, while a 99 percent confidence interval is denoted with **.

Site	Phenophase	Slope	Intersect	R^2	p	Standard error
Oakville	Dormancy	0.2716	-245.50	0.029	0.0341*	0.127
	Length of Greendown	0.2431	-374.94	0.022	0.0626	0.130
USGSEROS	Peak	-0.3452	886.31	0.020	0.0789	0.195
	Senescence	-0.3069	823.81	0.049	0.0057**	0.109
	Length of Maturity	-0.1948	419.02	0.019	0.0889	0.114
	Length of Greendown	0.5063	-923.46	0.077	0.0005**	0.142
Konza	Peak	-0.3034	770.91	0.043	0.0095**	0.116
	Greenup	-0.2495	596.77	0.032	0.0250*	0.110
	Maturity	-0.2735	698.67	0.044	0.0085**	0.103
ARIK	Mid-greendown	0.3600	-466.60	0.031	0.0277*	0.162
	Senescence	0.1963	-210.26	0.024	0.0535	0.101
	Length of Season	0.6131	-1021.37	0.033	0.0236*	0.268
	Length of Greenup	0.4260	-787.17	0.028	0.0376*	0.203
OAES	Length of Maturity	0.2866	-544.54	0.021	0.0694	0.157

3.4 Discussion

Using the XGB regression we developed a model capable of explaining 90 to 93 percent of the variability in four VIs (Fig. 19) across six grassland phenocam sites over the growing season. Our models demonstrate the importance of including photoperiod, temperature, and precipitation information when modeling vegetation phenology. Piao *et al.* (2019) reviewed the importance of including these different meteorological driving factors

for modeling vegetation phenology and remarked that many current phenology models underperform because of their dependence on temperature without considering the interactions of other weather variables. A study by Wang, Rich and Price (J. Wang et al., 2003) examined the Konza prairie, one of our six phenocam sites, and found that temperature was highly correlated with NDVI at the beginning and end of the growing season. Of the three GDD Tbase values explored, 0°C remained the most import variable within our model, having the highest total gain and remaining in all four reduced models. A Tbase of 0°C typically represents vegetation that uses the C_3-pathway for photosynthesis such as grasslands in the temperate prairie region, while the C_4-pathway is represented by a Tbase of 10°C and would be more common in the hotter and drier south-central semiarid prairie (Jones & Vaughan, 2010; McMaster & Wilhelm, 1997). Because of this we anticipated that either the 0 °C and the 10 °C GDD variables would both be included in the reduced model or the 5°C variable would better represent both regions and would have the highest total gain within the XGB regression. Instead, we found a mix of the three GDD Tbase values were used depending on the VI (Fig. 18). Both reduced GCC and VCI models contained Tbase values 0°C and 10°C, while the NDVI and EVI2 contained Tbase values 0°C, and 5°C.

The stepwise backwards elimination in XGB regression model variables we used to refine our final model was a simple approach to limiting regression variables, while allowing the model to identify the most important variables to include. XGB models developed with 50 to hundreds of independent variables can use more advanced feature selection models eliminating multiple features at a time with optimization algorithms that speed up processing

time (Pan et al., 2009; H. Zhang et al., 2019). With our approach, we were able to reduce our model from 13 variables down to four or five, depending on the VI, with a negligible change in model performance reflected in the average model R^2 decreasing by 0.011 and RMSE increasing by 0.002. This reduction in model variables allowed us to examine the importance of the variables as well as the calculated lag times for precipitation and SWE, and the relationship between different Tbase values for GDD. Wang, Rich and Price (2003) found a two-week lag in NDVI's response to precipitation events, however they also note that the response varied based on environmental conditions. For example, during a drier period the response to precipitation would often happen quicker. Our reduced models all selected precipitation with an accumulation of 30 days to best predict the phenology signals, suggesting that precipitation events occurring up to 30 days prior can control vegetation growth. This may be particularly true for the three phenocam sites in the south-central semiarid prairies since they are more susceptible to drought.

The four VIs we used across our analysis, GCC, VCI, NDVI, and EVI2, are all measures of vegetation phenology across the growing season. Of the three VIs, NDVI has the longest history in remote sensing (Rouse et al., 1973), while GCC has been well recognized within the phenocam literature because of its stability with uncalibrated imaging sensors (A. Richardson & Braswell, 2009). VCI provides a nonlinear transformation of GCC, providing a higher range of values by contrasting green with the sum of red and blue (X. Zhang et al., 2018). EVI2 has also increased in use recently (Bolton et al., 2020; Peng et al., 2021),

particularly with remotely sensed data from the Visible Infrared Imaging Radiometer Suite (VIIRS) system that lacks the blue band (X. Zhang et al., 2018).

Using the four VIs we were able to construct a 38-year phenology record at each phenocam location using the meteorological data and the reduced XGB models. Being able to use a combination of near-surface remote sensing and meteorological data to derive these VIs provides a valuable dataset for validation of satellite-based phenology products. It should be noted that these models reflect the vegetation from the period in which they were trained, 2015 to 2019. Any change in vegetation composition that may have occurred between 1981 and 2015 can not be accounted for since this period of the models is based entirely on meteorological data, and not on imagery from the phenocam stations. While this is a limitation of our models, it also acts as a control on our results since the trends in phenophase transition identified by the models are not affected by a change in species composition and are instead driven entirely by changes in climate. Changes in species composition can have a large effect on a phenology signal and presents a challenge in identifying climate change driven modification of phenophase transition periods (Prevéy & Seastedt, 2014; Wilsey et al., 2018). Because our models are not based on imagery of the vegetation across the 38 years, and instead depend on meteorological data, we are able to model the timeseries under the assumption that the species composition did not change.

The spline models used for detecting the phenophase transitions were on average able to account for 87 percent of the variation in the models with RMSE ranging from 0.017 for GCC to 0.041 for VCI. One feature of the spline models we did note, was their tendency to

be less influenced by extreme VI values (Table 3). Using the four splines for each VI at the six phenocam locations we measured seven phenophases and four phenophase periods. This resulted in 66 linear regression models (Appendix C) to determine if any trends appeared in phenophase transitions over the 38-year timeseries. Examining the significant trends within a 90 percent confidence interval (Table 4) we found 14 phenophases that have shifted across the phenocam sites except for the Nine-Mile station which had no significant trends. For the two northern phenocams in the temperate prairies the length of greendown has increased by 9.2 days (0.24 days/year) at the Oakville station, and 19.2 days (0.51 days/year) at the USGSEROS station over the 38 years. The 10-day difference between the two stations is likely attributed to the fact that the USGSEROS station has seen an earlier onset of peak greenness by 13.1 days (-0.35 days/year), and an earlier onset of senescence by 11.7 days (-0.31 days/year), which has also shortened the length of maturity by 7.4 days (-0.19 days/year). This suggests that the growing season at the USGSEROS station is trending towards a quicker occurrence of peak greenness followed by a shorter period of greenness between maturity and senescence, with an extension in the greendown period. In a study using imagery from the Advanced Very High Resolution Radiometer (AVHRR) from 1982 to 2002, Reed (2006) found grasslands to have a later dormancy period by 6.52 days (0.33 days/year), while greenup also started later by 8.01 days (0.40 days/year). A similar study used AVHRR from 1982 to 2006, Zhu *et al.* (2012) found grasslands in North America to have a later onset of greenness by 7.6 days (0.32 days/year), and a later dormancy by 2.1 days (0.09 days/year) causing a shortening of the growing season by 5.6 days (-0.23

days/year). The offset of dormancy occurring later into the season agrees with our study with dormancy at the Oakville station occurring 10 days later (0.27 days/year). This falls within the range found by Liu *et al.* (2016) with dormancy in the Northern Hemisphere occurring between 0.19 and 0.45 days later each year. For five of the six phenocam sites greenup did not have a significant trend, with no sites finding greenup occurring later. The one site with a greenup trend was the Konza station in which greenup occurs 9.5 days (-0.25 days/year) earlier in 2019 then in 1981. This value is close to the 2.8 days per decade (-0.28 days/year) in which spring phenology is predicted to have advanced for both plants and animals in the northern hemisphere (Hoegh-Guldberg, Jacob, et al., 2018). At the Konza station maturity and peak greenness is also occurring earlier in the year by 10.4 days (-0.27 days/year) and 11.5 days (-0.30 days/year), respectively. For this station, the earlier onset of greenness seems to be to be followed by an earlier onset of maturity and peak greenness for the vegetation. Of the six stations ARIK was the only station to find a significant trend in the overall length of the growing season with it increasing by 23.3 days (0.61 days/year). This station also had its length of greenup increase by 16.2 days (0.43 days/year) while its senescence and mid-greendown dates are occurring 7.5 days (0.20 days/year) and 13.7 days (0.36 days/year) later, respectively. The ARIK increase in length of season agrees with Zhou *et al.* (2001) who used AVHRR from 1981 to 1999 finding length of season in North America to increase on average by 12 days (0.65 days/year) and finding dormancy to occur 4 days (0.22 days/year) later. Overall across the five phenocam locations the significant trends we found align with studies of vegetation phenology over North American grasslands. Jeong

et al. (2011) used AVHRR to assess phenology from 1982 to 2008 and found both temporal and spatial variations in different phenology trends. They identified a reduction in the trend of an earlier onset greenness starting in 2000, while at the same time found an increase rate in later onset of dormancy, with both contributing to a lengthening of the growing season. While we did have variability across our six field stations, this is to be expected with increasing trends in spring temperature variability for North American grasslands that can have an influence on both spring and fall phenology (Lingling Liu & Zhang, 2020). Across our study area the results indicate that changing temperature and precipitation patterns are driving a significant change in phenology of the grasslands.

3.5 Conclusion

We used the machine learning based XGB regression model to predict changes in GCC, VCI, NDVI, and EVI2 across the growing season at six phenocam sites. With this model we were able to accurately predict 90 to 93 percent of variability in the VI values. This allowed us to reconstruct the VIs signals to derive a 38-year timeseries. With these modeled timeseries we were able to examine the trending changes in the phenophases at each of the grassland field sites. The significant trends we identified agreed with the many AVHRR and other satellite-based analysis that have been done for North American grasslands. We believe the methods used to develop our model provide a valuable framework for future work modeling vegetation phenology. Using near-surface remote sensing and meteorological data provides a valuable validation dataset for satellite-based phenology. Our model can be applied to additional phenocam sites, including ecosystem types other than grasslands, to

examine the interactions between photoperiod, temperature, and precipitation in these regions. Also, additional environmental factors could be considered such as soil moisture or nutrient availability. Future work that would help improve our understanding of grassland phenology should focus on identifying the spatial and temporal variability that exists in the phenology of the North American Great Plains, and how this will affect phenology across the region under future climate scenarios.

IV. Modeling Near-Surface Phenocam GCC, VCI, NDVI, and EVI2 Across the US Great Plains: Using Harmonized Landsat-Sentinel and MODIS Imagery

4.1 Introduction

Spatiotemporal changes in vegetation phenology identified in recent decades have been linked with changes in the global climate (Hoegh-Guldberg, O., et al., 2018). Vegetation phenology provides a measure of ecosystem response to climate change, and the feedbacks between vegetation and the climate (Buitenwerf et al., 2015). As vegetation phenology changes it has the potential to disrupt various functions within the ecosystem, such as the timing of flowering plants becoming mismatched with the life cycle of pollinator species (Tierney et al., 2013). Historically vegetation phenology has been monitored at individual field sites with species specific observations coordinated by networks such as the U.S. National Phenology Network and the Pan European Phenology project, under which citizen scientists collect the phenological records of various plant species (Crimmins et al., 2017; Rodriguez-Galiano et al., 2015).

Land surface phenology (LSP) involves the use of remotely sensed imagery often acquired from a satellite-based platform. Instead of focusing on a single species, or individual plants at a field site, LSP integrates entire regions of an ecosystem into a single measurement aggregated by the spatial resolution of pixels within remotely sensed imagery (Helman, 2018). The spatial resolution of satellite imagery used for LSP can vary in scale from meters to kilometers with each pixel representing a patchwork of vegetation. Depending on the heterogeneity of a given ecosystem, a single pixel may represent many different

species as well as varying life stages within a single species (Snyder et al., 2019). Vegetation Indices (VIs) are commonly derived from satellite imagery to provide a unitless measure that tracks plant growth and CO_2 uptake (Helman, 2018). This includes VIs such as the normalized difference vegetation index (NDVI), which exploits the interactions between photosynthetic tissues and electromagnetic radiation (Helman, 2018; Jones & Vaughan, 2010) and has a long history in earth observation (Rouse et al., 1973). Using these VIs it is possible to derive various phenophases across the growing season such as the start of season in the spring, followed by maturity through the summer, the onset of senescence in the fall, and finally dormancy of the vegetation until the beginning of the next growing season (Bolton et al., 2020; Gray et al., 2019).

While acquiring vegetation phenology from satellite-based imagery has been conducted using many sensors, difficulties arise when comparing phenophase dates measured using imagery with various spatial scales (X. Zhang, Wang, et al., 2017). Differences in phenophase dates have been identified, particularly between imagery that ranges in spatial resolution from meters to kilometers (Peng, Zhang, et al., 2017; F. Wang et al., 2018; X. Zhang, Wang, et al., 2017). The difference in phenophases between sensors of various spatial resolutions has been called the “scale effect” and has been identified even when comparing the same satellite imagery rescaled to a courser resolution (Licong Liu et al., 2019; Peng, Zhang, et al., 2017). Validation efforts have been made between fine and course resolution satellite imagery, as well as with ground-based observations and phenological models, and

often resulted in a difference between observations of greater than 10 days (X. Zhang, Wang, et al., 2017).

Studies have used ground-based phenology cameras known as phenocams to attempt to overcome the challenges with relying on satellite imagery (A. Richardson & Braswell, 2009). Phenocams can provide valuable information about the vegetation at single locations across the globe (Filippa et al., 2018; Hufkens et al., 2012; Petach et al., 2014; Sonnentag et al., 2012; Toomey et al., 2015; X. Zhang, Wang, et al., 2017). However, none have attempted to couple the temporal and radiometric information in the phenocam imagery. To overcome the difficulties in using single site ground-based phenocam observations as a method of calibration, we propose a multisite validation method. In this manner, phenocam observations from multiple sites are coupled together and radiometric calibration is completed. This aligns the multiple phenocam measurements and provides a method to couple observations from multiple sites, in which the observations from various sites can be compared against one another in a standardised way.

4.2 Methods and Data

4.2.1 Study Area

The Great Plains of North America occupy 281 million ha with 224 million ha located within the contiguous U.S. (U.S. Environmental Protection Agency, 2020). The Great Plains Ecoregion is subdivided into Level-2 regions that represent the diversity within the Great Plains (Fig. 23). The temperate prairies in the east are wetter and contain more croplands than the drier west-central and south-central semiarid prairies, while the west-

central semiarid prairies are on average cooler than south-central semiarid prairies (Omernik & Griffith, 2014).

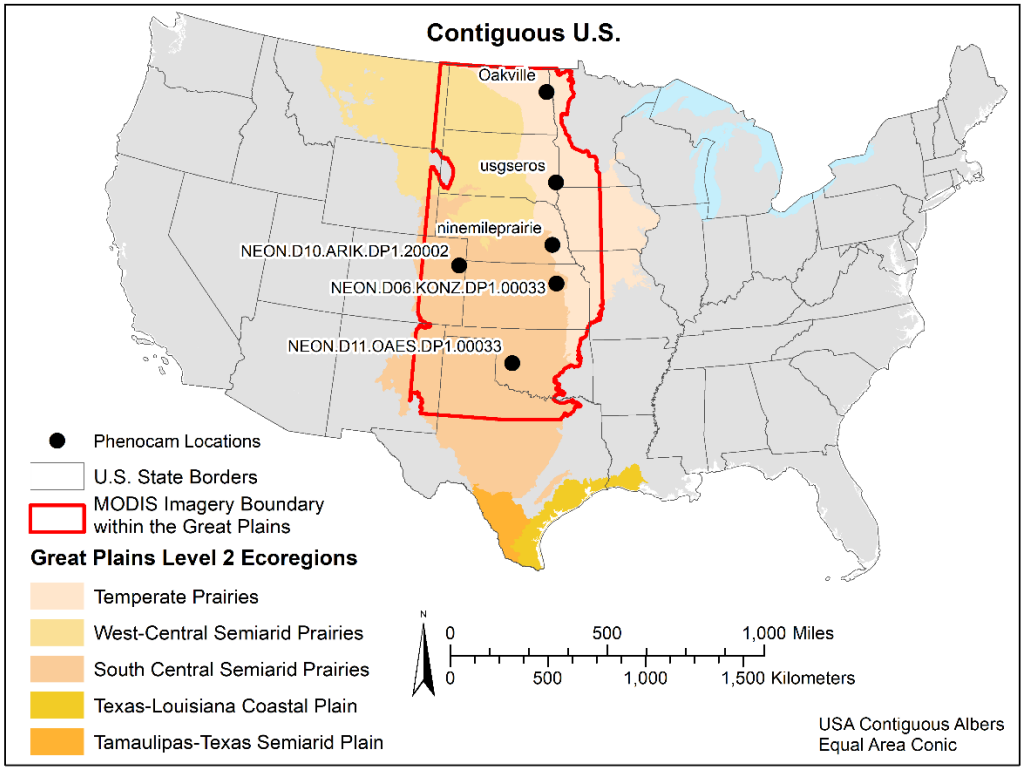


Figure 23: The phenocam locations within the study area. Showing the six phenocam locations situated within the Great Plains of the contiguous U.S. The MODIS imagery boundary is based on the MODIS acquired for this study and the boundary of the Great Plains region. Modified from Burke and Rundquist (2021).

We selected six grassland locations within the Great Plains of the contiguous U.S. (Fig. 23) each of which has a phenocam with at least three years of data (Table 5). Three of the sites are located within the temperate prairie ecoregion; the Oakville Prairie (Oakville), a part of the University of North Dakota, located in Grand Forks County, North Dakota (47.8993°N, 97.3161°W); the USGSEROS station at the Earth Resources Observation and

Science (EROS) Data Center in South Dakota (43.7343°N, 96.6234°W); and the Nine Mile Prairie station (Nine-Mile), a part of the University of Nebraska – Lincoln (40.8680°N, 96.8221°W), located in Lancaster County, Nebraska. The other three phenocam sites are within the south-central semiarid prairie and are a part of the National Ecological Observatory Network (NEON). These sites include the NEON.D06.KONZ.DP1.00033 station (Konza) (39.1008°N, 96.5631°W) located at the Konza Prairie Biological Station near Manhattan, Kansas; the NEON.D10.ARIK.DP1.20002 station (ARIK) (39.7582°N, 102.4471°W) located near the Arikaree River in Yuma County, Colorado; and the NEON.D11.OAES.DP1.00033 station (OAES) (35.4106°N, 99.0588°W) located at the Klemme Range Research Station in Washita County, Oklahoma. The six sites represent a 1,470 km latitudinal transect through the Great Plains region ranging from 35.4°N to 47.9°N.

Table 5. Years of data available for each of the phenocam site locations.

Station Name	Years with available data
Oakville	2016 - 2019
USGSEROS	2015 - 2017, 2019
Nine Mile Prairie	2016 - 2019
NEON.D06.KONZ.DP1.00033	2017 - 2019
NEON.D10.ARIK.DP1.20002	2017 - 2019
NEON.D11.OAES.DP1.00033	2017 - 2019

4.2.2 Phenocam Data

We choose to derive four VIs from the phenocam imagery at the six field stations. GCC (Eq. 1) is a proportional measure of relative 'greenness' that was originally developed for use with phenocams because of its relative stability under changing illumination conditions (A. Richardson & Braswell, 2009). GCC has be used in a diverse array of ecosystem types, and can be measured using any digital camera capable of capturing a color (red, green, and blue) image (A. D. Richardson, 2019). VCI (Eq. 2) was created as a nonlinear transformation of GCC that has a higher dynamic range relative to GCC by contrasting the green band to the sum of red and blue (X. Zhang et al., 2018). NDVI (Eq. 3) has a long history in Earth Observation (Rouse et al., 1973), and has been derived from phenocams that are sensitive to near-infrared wavelengths (Burke & Rundquist, 2021; Filippa et al., 2018; Petach et al., 2014; A. D. Richardson, 2019). EVI2 (Eq. 4) was developed as an adjustment to NDVI, with an enhanced ability to remove soil background noise, and atmospheric effects (Jiang et al., 2008).

$$GCC = \frac{Green}{Blue + Green + Red} \tag{1}$$

$$VCI = \frac{Green}{Blue + Red} \tag{2}$$

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{3}$$

$$EVI2 = 2.5\frac{NIR - Red}{NIR + 2.4 * Red + 1} \tag{4}$$

To calculate each of the chosen VIs from the phenocam imagery, we first downloaded all available imagery from the six phenocam locations (phenocam.sr.unh.edu). We then applied the exposure correction to both the color and mixed color-infrared imagery to extract the near-infrared and three color bands (Petach et al., 2014). Using the image digital numbers (DNs) for the red, green blue (RGB) and near-infrared (NIR) bands the four VIs were calculated using Eq. 1, 2, 3 and 4 for each day of the year in which phenocam imagery was available (Table 5). Finally, the phenocam VIs were linearly scaled to Gaussian Process Regression modeled VIs calculated with Harmonized Landsat-Sentinel surface reflectance imagery (described in detail in Burke and Rundquist 2021). This standardised the VI values between each phenocam site, allowing them to be used together within a single regression model.

4.2.3 MODIS and HLS Imagery

MODIS imagery was acquired from the U.S. Geological Survey through the Application for Extracting and Exploring Analysis Ready Samples (AρρEEARS). We acquired all available imagery for the MOD09GA v006 product from January 1, 2015, to December 31, 2019, using the boundary shown in Fig. 23. From the MOD09GA product we used the provided blue, green, red, and near-infrared (NIR) color bands, bands 3,4,1 and 2, with radiometric, and atmospheric correction providing measures of surface reflectance with a 500m x 500m spatial resolution. These images cover 146 million ha within the Great Plains region and in total 1,816 images were recorded over the 5-year period.

We acquired Harmonized Landsat Sentinel-2 (HLS) from the National Aeronautics and Space Administration's (NASA) data portal (https://hls.gsfc.nasa.gov/data/v1.4/). We selected the imagery tiles covering each of the six phenocam sites and downloaded all imagery available for each site starting January 1, 2013 ending on December 31, 2019. This included imagery taken from both the Sentinel-2 S30, and the Landsat-8 L30 products, providing a 30m x 30m spatial resolution. Each of the HLS images covers approximately 1,211,000 ha of land, and in total we acquired 2,027 Landsat-8 L30 images and 2,385 Sentinel-2 S30 images. We used the blue, green, red and NIR color bands for both sets of imagery, with Landsat-8 bands 2,3,4, and 5, and Sentinel-2 bands 2,3,4 and 8.

4.2.4 Cloud Removal

To reduce cloud contamination in both the MODIS and HLS imagery we used the quality bands to identify pixels in the imagery that contained cloud contamination and either removed these images from the dataset or modeled the pixel values. Images were removed from the dataset if 500 or fewer pixels were determined to be cloud free, or if after attempting to model cloud-contaminated pixels the images were visually inspected and found to still contain cloud contamination. To model pixels that contained cloud cover, images were aligned in sequential order and processed in that order. For each image, the 20 images prior and 20 images post were used to build a regression model. Using an XGB regression model (Chen & Guestrin, 2016), cloud-free pixels from the image being modeled were used in conjunction with cloud-free pixels from the 40 other images as a training dataset. We trained our XGB models using a randomly selected 80 percent of the available cloud free

pixels, leaving 20 percent for model validation. To help prevent overfitting of the model, and to prune any branches with a negative gain, we set lambda to 1 and both alpha and gamma to 0. We also set the learning rate to 0.1, max depth to 10 and number of estimators to 50,000. We choose parameters that would help prevent overfitting of the model, and were recommended to produce a more conservative algorithm (Chen & Guestrin, 2016). Subsampling, also known as bootstrap aggregating, was used so that a random selection of half (subsample = 0.5) of the training samples were used to grow each tree with gradient-based selection (Chen & Guestrin, 2016; H. Zhang et al., 2019). Using the XGB models, we iterated though all 1,816 MODIS images, and then through each phenocam site's HLS images totalling 4,412 images. We followed up with a visual inspection of all imagery and removed images if they still contained cloud cover pixels.

4.2.5 Imagery Fusion

To increase the temporal resolution of the HLS imagery, we used the MODIS imagery with the Flexible Spatiotemporal Data Fusion (FSDAF) model to predict HLS imagery for any day in which MODIS imagery were available but not HLS imagery (X. Zhu et al., 2016). This process was done at all six phenocam locations independent of one another from January 1, 2015, to December 31, 2019. We paired together the HLS and MODIS imagery and identified all days in which the higher resolution HLS imagery was already available. We also clipped the MODIS imagery to align with the spatial extent of the HLS image. Then, we identified days with a MODIS image but no HLS image and used this list to iterate through and execute the FSDAF model. To run the model, a single pair of

HLS/MODIS images was needed along with the MODIS image for the day during which the fused image was to be generated. For each day without an HLS image, the closest paired HLS and MODIS images chronologically were used. This produced a fused HLS-MODIS image for any day of year in which no HLS imagery was available and increased the temporal resolution of the HLS imagery timeseries. Across the six phenocam locations over the five-year period 9,532 FSDAF fused images were produced.

4.2.6 Modeling Phenocam VIs Using XGB Decision Trees

Using the HLS, MODIS and the FSDAF fused images we calculated the four VIs for each image. To do this we used Eq. 1,2,3 and 4 to calculate each of the VIs for each pixel within the satellite images. This produced a timeseries for each of the VIs for both the HLS and FSDAF imagery at each phenocam site, as well as the MODIS imagery. From here we used the HLS and FSDAF imagery to model the phenocam VIs using an XGB regression model. To do this we first needed to calculate a single value for each image across all six phenocam sites. We used Inverse Distance Weighting (IDW) (Shepard, 1968) with the phenocam imagery footprints to calculate a single VI value for each image, following the same procedure as used by Burke and Rundquist (2021). This provided a timeseries of values for the four VIS across the six phenocam sites.

For each of the four VIs we produced an XGB regression model (Chen & Guestrin, 2016) capable of predicting the difference between the phenocam VIs and the HLS VIs values. For each of the four VIs, the model was trained using the VIs values, and the accumulated VIs values, as well as the day of year (DOY) for each image. The accumulated

VIs are calculated by adding all previous VI values within a given year up to and including the date for a given image. On January 1 of any given year this value was reset to zero and then the VI was accumulated throughout the year. Using the accumulated VIs values, and the DOY we then trained the XGB model to predict the phenocam VI value. We trained our XGB models using a randomly selected 80 percent of the available phenocam pixels, leaving 20 percent for model validation. To help prevent overfitting of the model, and to prune any branches with a negative gain, we set lambda to 1 and both alpha and gamma to 0. We also set the learning rate to 0.1, max depth to 20 and number of estimators to 50,000. Subsampling, also known as bootstrap aggregating, was used so that a random selection of half (subsample = 0.5) the training samples were used to grow each tree with gradient-based selection (Chen & Guestrin, 2016; H. Zhang et al., 2019).

We used the XGB models to then predict the phenocam VIs values across the MODIS imagery given the MODIS VIs values and the DOY. To test the ability of the XGB models to predict pixel values at locations the model was not trained on, we retrained the XGB models for each VI withholding one of the phenocam sites, and then used this site as the validation dataset. We did this six times for each of the VIs, excluding each of the six phenocam locations. This gave us a measure of how well the XGB models were able to account for locations they were not trained with, which for the MODIS imagery is any pixel other than those at the six phenocam locations.

4.2.7 Phenophase Detection

To measure phenophase changes across the MODIS imagery using both the standard and modeled VIs, we used the protocols provided for phenophase detection with MODIS imagery (Gray et al., 2019). We identified seven phenophase stages throughout a growth cycle, starting with greenup in the spring, followed by midgreenup, maturity, peak greenness, senescence, midgreendown, and ending with dormancy in the fall. This process was carried out for each pixel of the MODIS imagery and done for each of the four VIs. A natural cubic spline (Drury, 2020) was fit to the five-year time series. To find the optimal number of knots to fit the spline we used Akaike's Information Criterion (AIC) to balance under-overfitting of the model (Hurvich et al., 1998). To do this we randomly set aside one third of the dataset and fit the spline starting at 5 knots (1 knot per year of data). Using the AIC we measured the models fit against the randomly removed data and selected the number of knots that produced the lowest AIC value. The spline was then re-fit to the entire dataset using the determine optimal number of knots.

Valid vegetation cycles were identified from the spline models using methods similar to the CMCD12Q2 product (Gray et al., 2019). Local minima and maxima were identified for each year with a 6-month overlap at the beginning and end of the year. The maxima were examined for validity as a peak in vegetation growth while the minima were examined to be either the start or end of a vegetation cycle. However, the methods used for the CMCD12Q2 product was produced for EVI2 specifically and has a set value of 0.1 required in the amplitude of any greenup or greendown period for it to be considered a valid cycle. The three

other VIs have a varying range of values that do not necessarily align with EVI2. Instead of using a constant value of 0.1, we modified this step by requiring greenup and greendown periods to have an amplitude that is at least 70 percent that of the current year's amplitude. Once the valid growth periods were identified we then extracted the seven phenophase periods using the same methods as the CMCD12Q2 product. The peak is reached at the maximum value for the VI. The greenup, mid-greenup, and maturity occur at a 15, 50, and 90 percent increase in amplitude, while senescence, mid-greendown, and dormancy occur after the peak as amplitude decreases past 90, 50, and then 15 percent. Using these values, we were able to produce an image for each phenophase across the five years of data for the four VIs. This resulted in a total of 140 images with 7 phenophases across 5 years of data and 4 different VIs.

4.3 Results

4.3.1 XGB Models

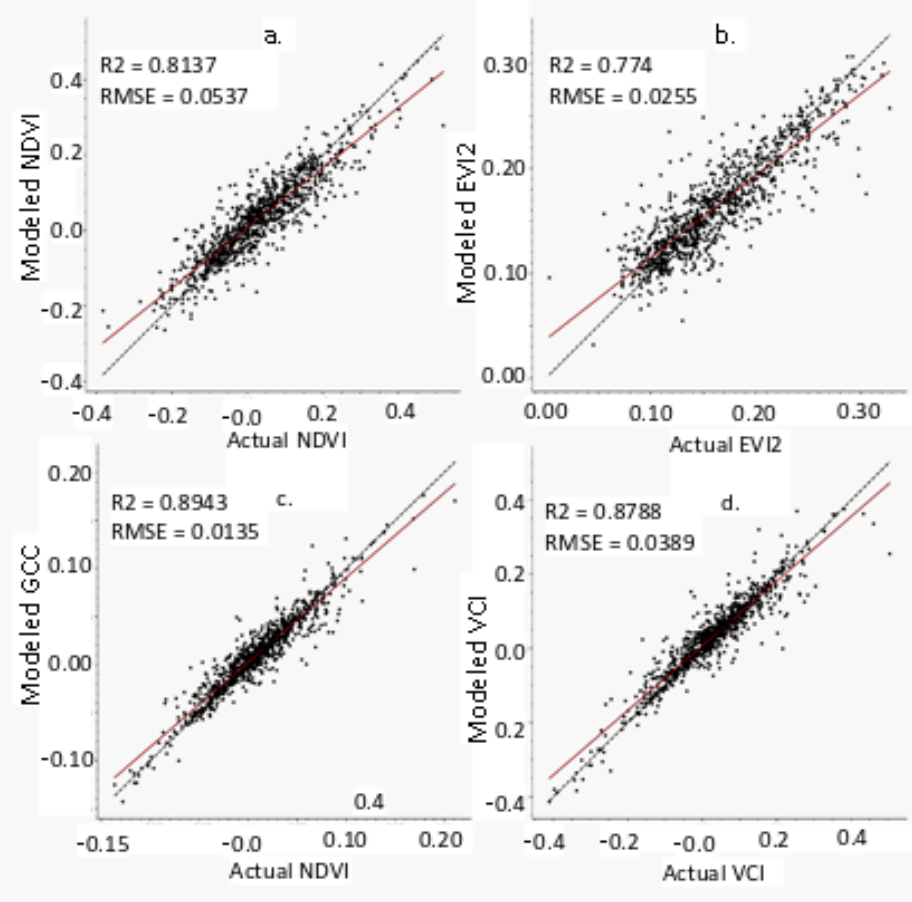


Figure 24: Final XGB Models for predicting the four VIs, showing the relationship between the predicted VI values and the actual values taken from the phenocam imagery. The dashed black line depicts a 1:1 line showing a perfect fit.

Using the four VIs we modeled the six phenocam sites to predict the VIs values at any given location across the study region. This produced four models, one for each of the VIs (Fig. 24). The GCC model produced had the highest R^2 at 0.894 with RMSE of 0.0135

while EVI2 had the lowest R^2 at 0.774 with a RMSE of 0.0255. The NDVI had the largest RMSE at 0.0537 with R^2 at 0.814 and VCI fell in between with R^2 of 0.879 and RMSE of 0.0389. By separating each of the phenocam sites we were able to evaluate how well each sites data fit to a site that was not included in the training data. While this does not provide a direct assessment of the models created using all available phenocam data, it provides a good assessment of how well these models performed at sites not included in training the model (Figs. 25, 26, 27, and 28). Across the four VIs, and the six phenocam sites on average the R^2 value was 0.5953 with an average RMSE of 0.0617. The best performing site was GCC at Oakville with an R^2 0.808 and an RMSE of 0.0222, while the worst performing site was EVI2 at Nine-Mile with an R^2 0.207 and a RMSE of 0.0678.

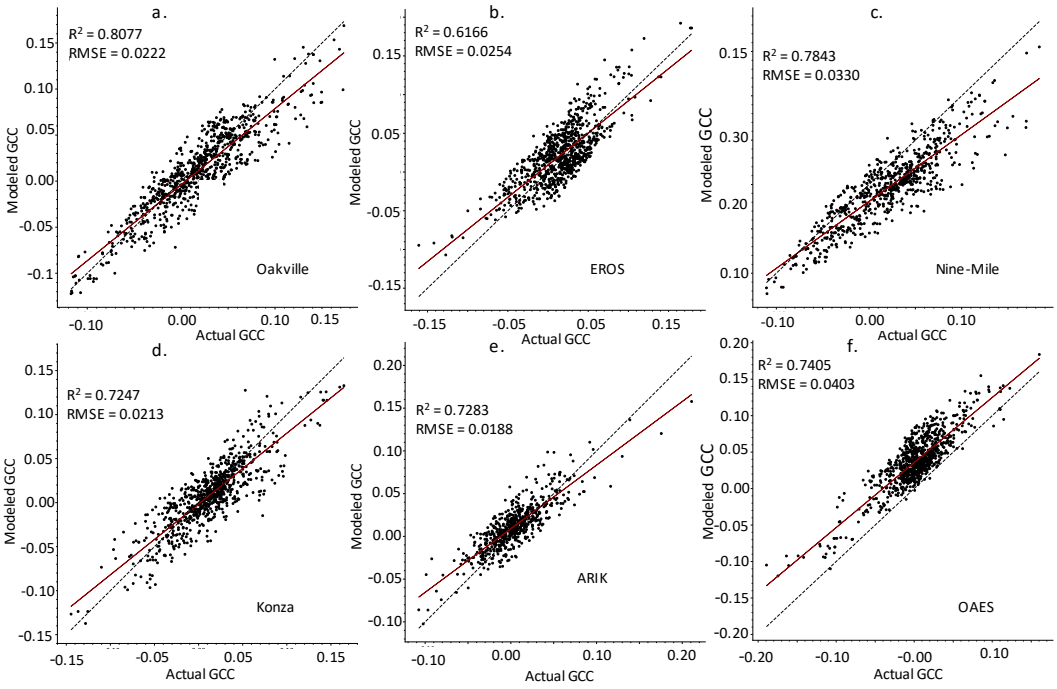


Figure 25: How each of the XGB models preformed after being trained using all data except for one of the six phenocam sites. The site listed in each scatterplot is the site that the model

was not trained with and was instead fit against. This shows how well the model was able to predict GCC values at a location in which the model was not previously trained.

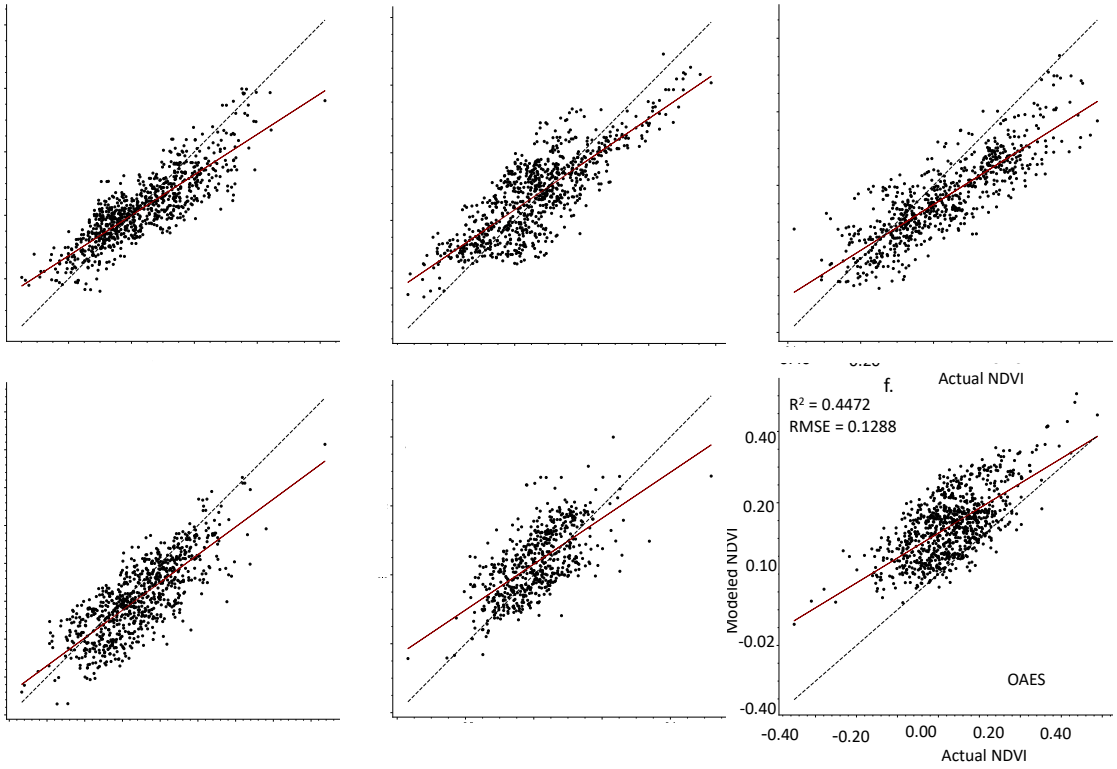


Figure 26: How each of the XGB models preformed after being trained using all data except for one of the six phenocam sites. The site listed in each scatterplot is the site that the model was not trained with and was instead fit against. This shows how well the model was able to predict NDVI values at a location in which the model was not previously trained.

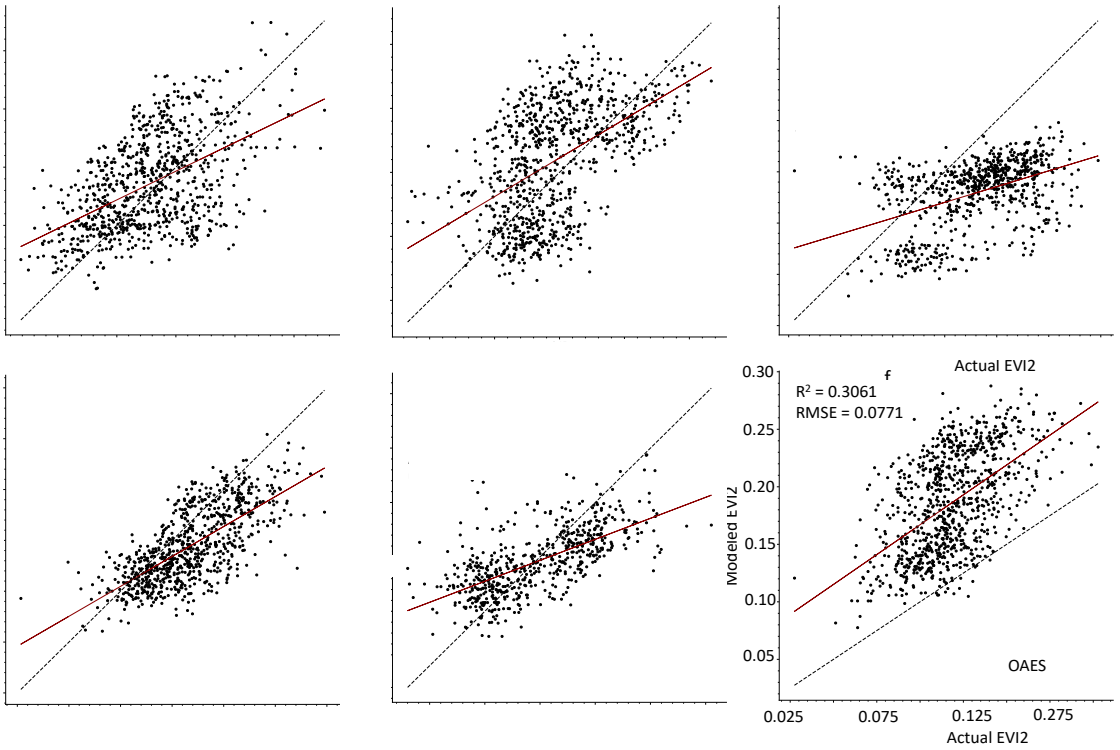


Figure 27: How each of the XGB models preformed after being trained using all data except for one of the six phenocam sites. The site listed in each scatterplot is the site that the model was not trained with and was instead fit against. This shows how well the model was able to predict EVI2 values at a location in which the model was not previously trained.

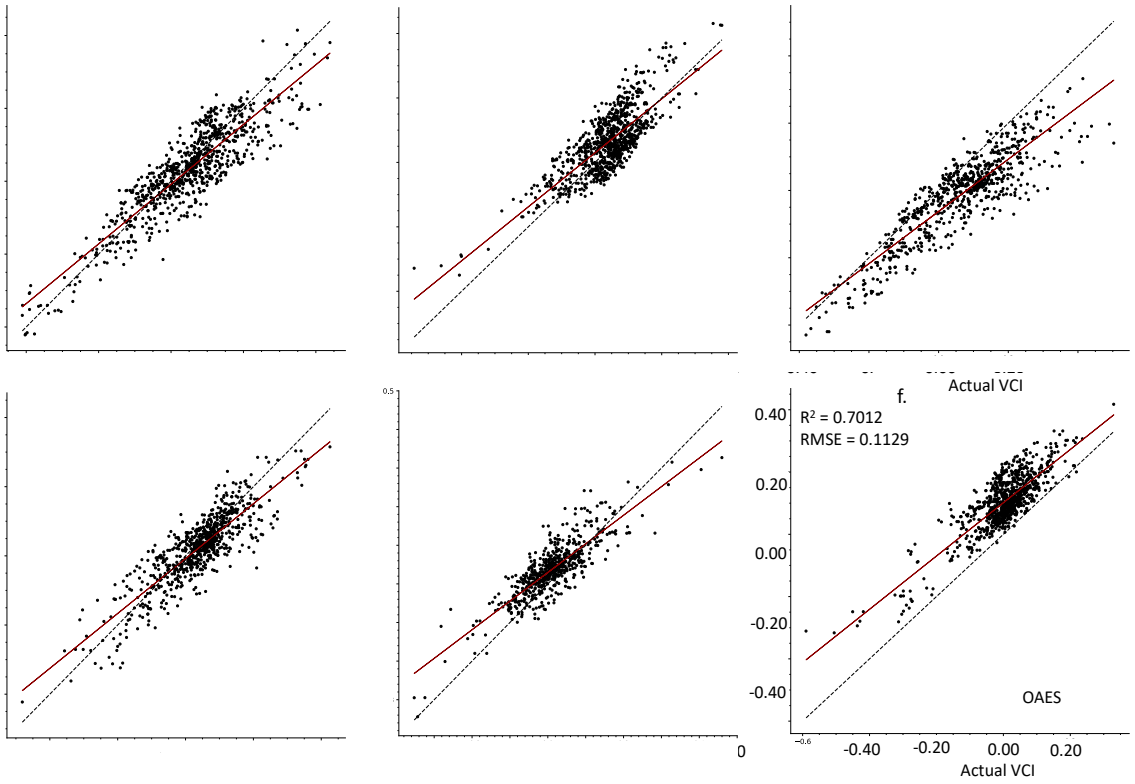


Figure 28: How each of the XGB models preformed after being trained using all data except for one of the six phenocam sites. The site listed in each scatterplot is the site that the model was not trained with and was instead fit against. This shows how well the model was able to predict VCI values at a location in which the model was not previously trained.

4.3.2 Histograms of Original and Modeled Phenophase Detection

Each of the phenophases for the four VIs were graphed so that each phenophase was presented with a separate graph for each year as well as separate graphs between the original and the modeled VI values. The four sets of graphs, one set for each VI are presented in Figs. 29, 30, 31, 32 and Appendix D. Across the graphs the modeled VIs often have peaks occurring closer together than the original VIs with peaks more spread across the growing season. For example, in Fig. 25 the NDVI values in the modeled graphs

cluster togethers with peaks between DOY 100 and 300, while the original graphs spread out with peaks covering a larger range of values across the year between DOY 50 and 350. Values of zero occur for all phenophases and across all the graphs, these are pixels that contain no data occurring on the edge of the images in the empty pixels. Values less than zero are an error in the modeling procedure in which a given pixel is predicted to have phenophases occurring before the first day of the year.

Figure 29: Histogram of NDVI-based phenophase values, showing the predicted date for the seven phenophases starting with greenup in the spring and dormancy in the fall. The left column of graphs shows NDVI values from the original MODIS imagery, while the right column shows the predicted NDVI values. The two columns start with 2015 at the top and end with 2019 at the bottom.

Figure 30: Histogram of GCC-based phenophase values, showing the predicted date for the seven phenophases starting with greenup in the spring and dormancy in the fall. The left column of graphs shows GCC values from the original MODIS imagery, while the right column shows the predicted GCC values. The two columns start with 2015 at the top and end with 2019 at the bottom.

Figure 31: Histogram of EVI2-based phenophase values, showing the predicted date for the seven phenophases starting with greenup in the spring and dormancy in the fall. The left column of graphs shows EVI2 values from the original MODIS imagery, while the right column shows the predicted EVI2 values. The two columns start with 2015 at the top and end with 2019 at the bottom.

Figure 32: Histogram of VCI-based phenophase values, showing the predicted date for the seven phenophases starting with greenup in the spring and dormancy in the fall. The left column of graphs shows VCI values from the original MODIS imagery, while the right column shows the predicted VCI values. The two columns start with 2015 at the top and end with 2019 at the bottom.

4.3.3 Phenophase Maps

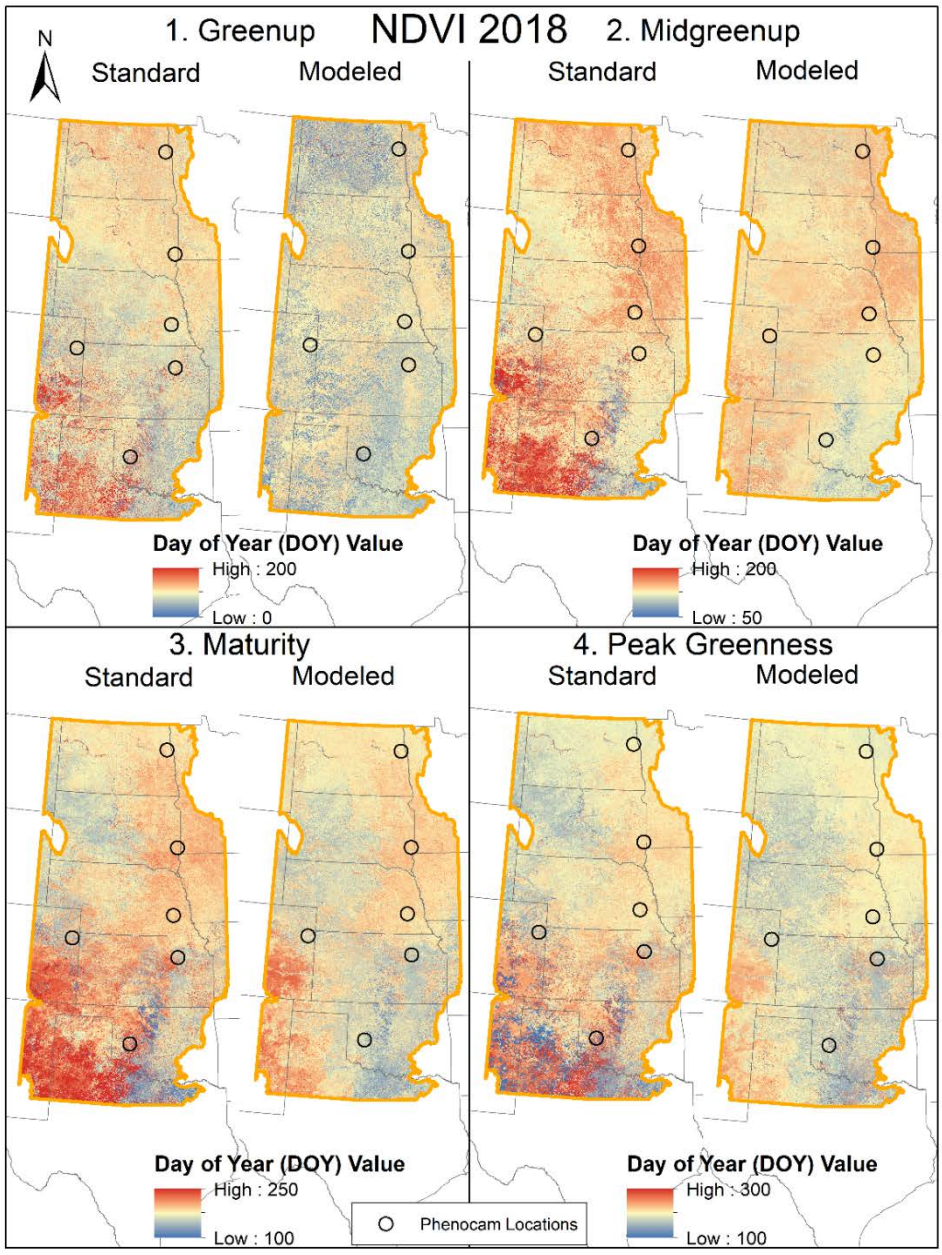


Figure 33: Maps showing the date of the first four NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2018.

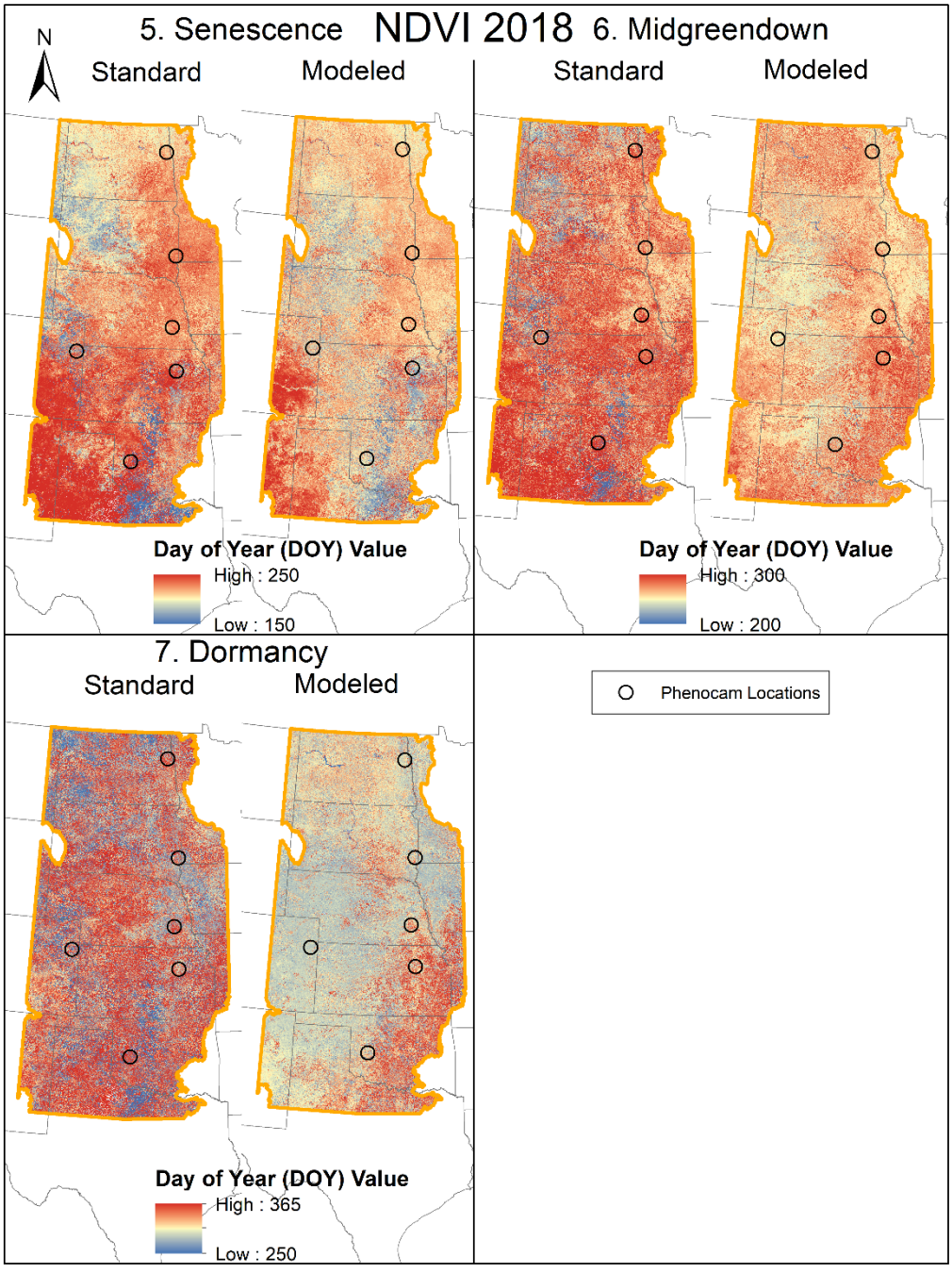


Figure 34: Maps showing the date of the last three NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2018.

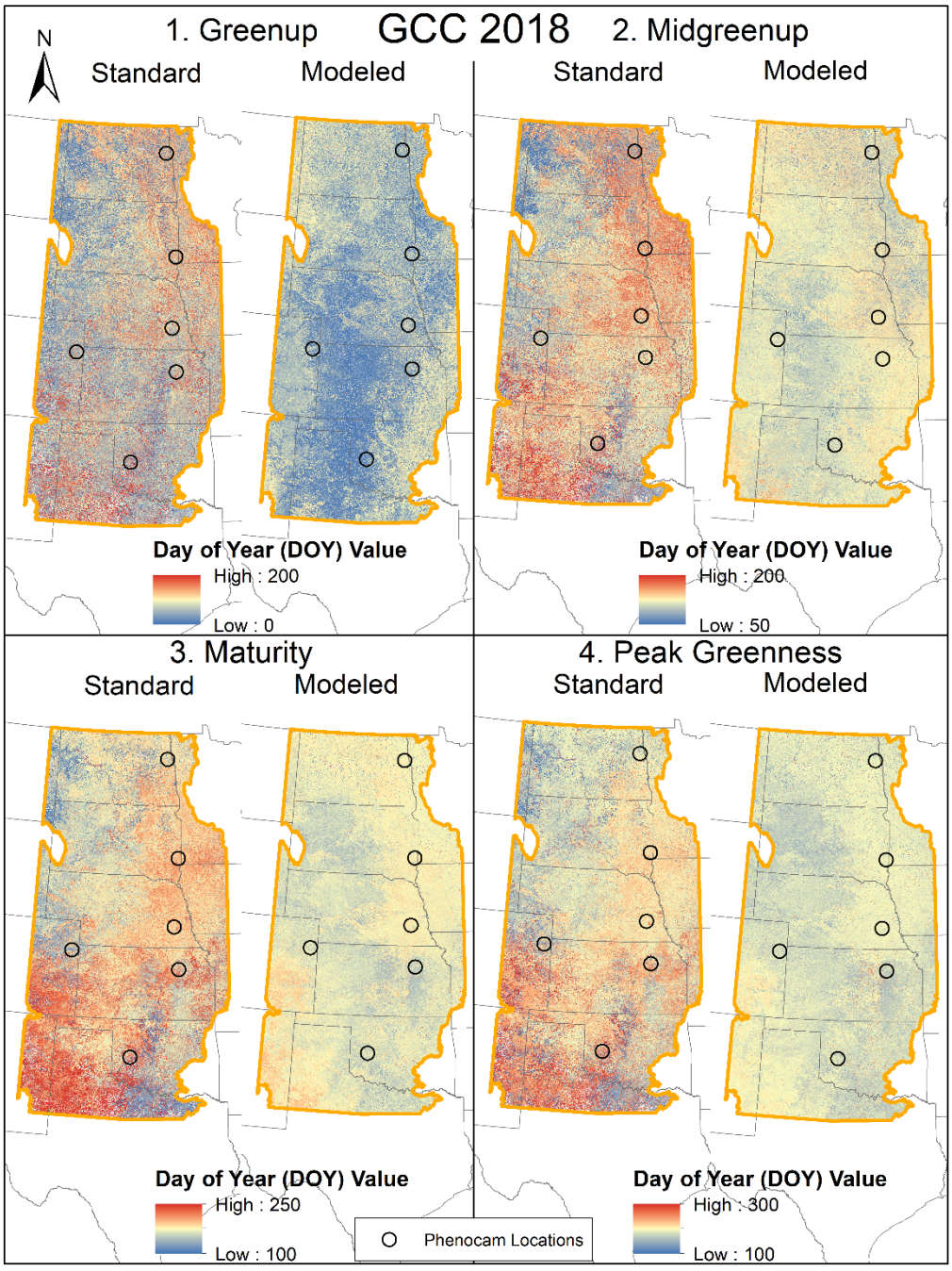


Figure 35: Maps showing the date of the first four GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2018.

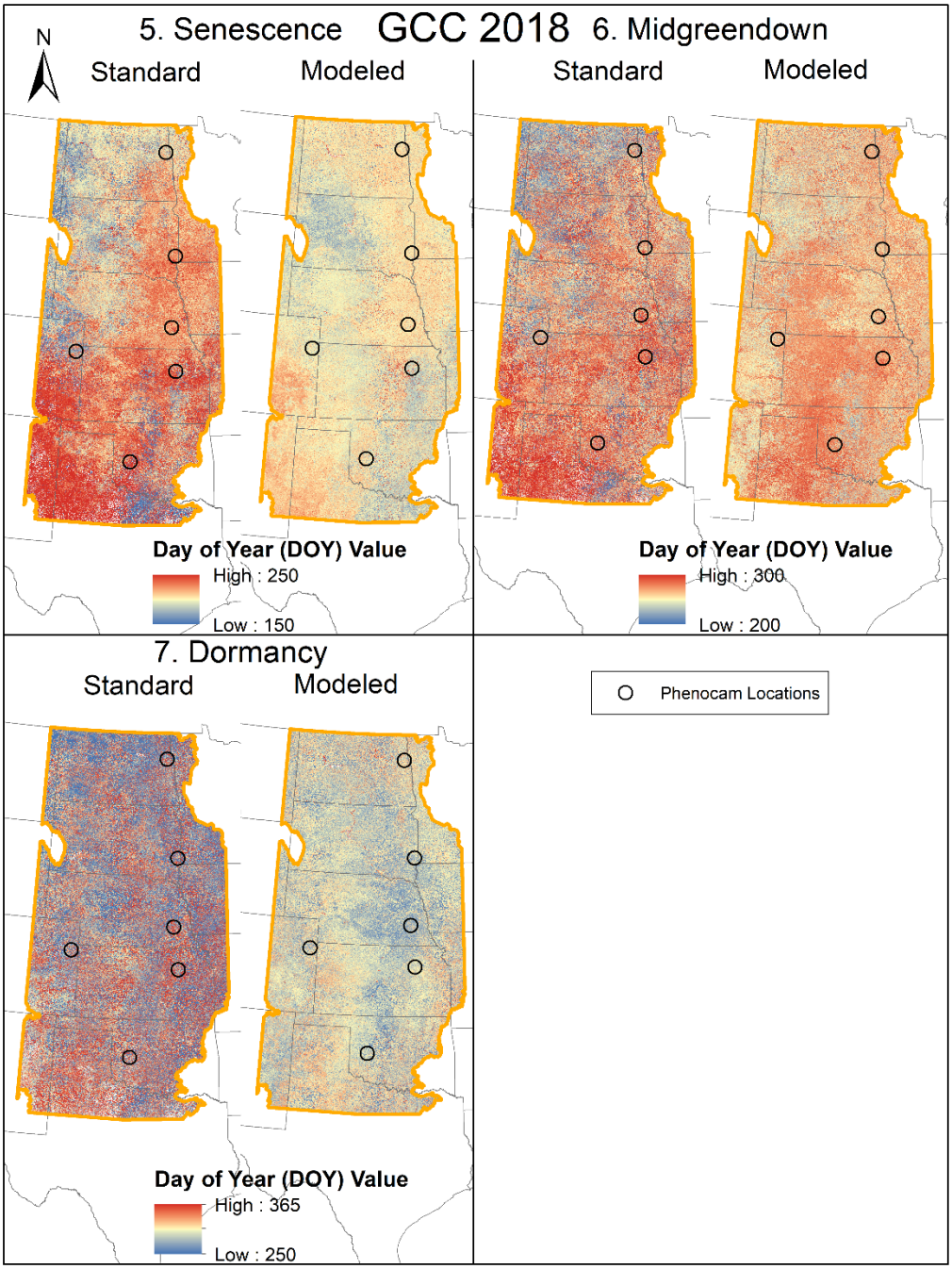


Figure 36: Maps showing the date of the last three GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2018.

Using the five years of data across the four different VIs produces forty sets of maps (Figs. 33, 34, 35, 36 and Appendix D). Each of these map sets show how the seven phenophases vary between the five years and across the four VIs. The map sets also show the variation between the standard imagery and the modeled imagery. In the 2018 NDVI Imagery (Figs. 29 and 30) we found the modeled datasets often had fewer extreme values across the pixel values. This was particularly true in the southwest corner of the maps that tended to have higher pixel values across the phenophases. Greenup in the modeled NDVI tended to occur earlier in the year, while dormancy in the fall happened earlier for most of the modeled pixels. The 2018 GCC followed similar trends as NDVI with the modeled pixels having fewer extreme values, particularly in the southwest where the standard pixels had much higher values.

4.4 Discussion

Using a combination of HLS and MODIS satellite imagery we were able to produce near daily measurements of the six phenocam locations and the surrounding study region. With blue, green, red, and NIR color information available we calculated four VIs. This produced a total of five years of data for each of the VIs from 2015 to 2019, or twenty years of data total. Using the timeseries for each of the four VIs we then predicted the dates for greenup, midgreenup, maturity, peak greenness, senescence, midgreendown and dormancy.

In recent years the number of available phenocam's within the North American prairies has increased to cover several ecozones within the region (Seyednasrollah et al.,

2019). This is an improvement over previous studies that found this region lacking in phenocam data (Cui et al., 2020). Using the improved spatial coverage of phenocam data in the region in conjunction with the temporal resolution of our modeled phenocams we calculated VIs over five growing seasons. This provides a great improvement both spatially and temporally over methods that often rely on data from a limited spatial region or temporal timeline (Cui et al., 2020; Walker et al., 2014).

While satellite sensors such as AVHRR, MODIS and VIIRS can cover a much larger region, including being used for global coverage (X. Zhang, Liu, et al., 2017), it is not able to account for regional differences present in fine-resolution satellite imagery. Instead, we developed a method that has allowed us to capture the fine spatial details present in Landsat and Sentinel-2 data, while at the same time we were able to couple together multiple phenocam datasets from across the Great Plains. This can then be used to improve the spatial and temporal resolution of models depending entirely on MODIS imagery (Ganguly et al., 2010). This may provide some insights in global phenology dynamics that have similar aspects (Cuizhen Wang et al., 2015), or in interpreting phenology metrics within North America (Cui et al., 2019).

Using imagery from several phenocams has been done to incorporate imagery from multiple phenocam sites (C. J. Watson et al., 2019). However, this is the first study that has attempted to fuse together the spatial and spectral information across multiple phenocam sites. Having fused imagery across multiple phenocam locations provides important correction information, allowing the imagery to be compared between the phenocam sites,

and providing coupled ground control locations which can be used to predict VIs using a calibrated method across the study region. With the fused imagery we were then able to predict VI values at any location in our study region for each DOY.

Examining Figures 25, 26, 27, and 28 shows the quality in which each of the VIs were predicted by our model. EVI2 appears to have the most variability across its values, with GCC, NDVI, and VCI maintaining average R^2 values of 0.767, 0.597, 0.726 across the models when comparing with EVI2 having an average R^2 of 0.325. This suggest that EVI2 is more sensitive to variations in the VI values when comparing with the other three VIs. However, EVI2 has been used to monitor phenology across North America, and provides a valuable metric on the growth of vegetation in the region (Jiang et al., 2008; X. Zhang et al., 2020). Of the four VIs, GCC had the best R^2 values ranging from 0.82 to 0.73, providing a higher degree of fit between the XGB models and the predicted phenocam GCC values. GCC has traditionally been used as a VI taken from phenocam imagery, and because it is uncalibrated, however relatively stable under changing illumination conditions (A. D. Richardson et al., 2018), it not typically used for satellite imagery (Cui et al., 2019). However, by coupling together spectral information from both satellite and multiple phenocam data sources we were able to calibrate the imagery and derive GCC data across our study area.

We predicted the seven phenophases across the five years of imagery for each of the VIs (Figs. 29, 30, 31, and 32). Using this information, we were then able to compare each VI in the original MODIS and HLS imagery with the modeled imagery. Across the five years

and four VIs the modeled VIs typically had a smaller range of values and was better able to reduce the variation in the predicted VIs. However, some variations still existed within the histograms, and did not correct for all errors. One notable error in our phenophase detection was in both the original and modeled phenophases for 2019 across the four VIs, a large portion of the fourth VI always occurred at the end of the year instead of in the middle where it is expected. This did not occur for any other VIs and is an error in the VI detection process. Otherwise, both the original and modeled 2019 VIs appear to have properly predicted the six other phenophases across 2019.

The four VIs were used to map the values in the standard MODIS imagery as well as the modeled VI values across the five years (Figs. 33, 34, 35, 36 and Appendix D). This provided the expected DOY value that the phenophase occurred across the seven phenophases and using the four different VIs for 2015, 2016, 2017, 2018, and 2019. This produced 140 maps representing the modeled phenophase values for each year, across the four VIs, for the seven different phenophases (Figs. 29, 30, 31, 32 and Appendix D). Among the four VIs we found many similar trends over the study region with the modeled phenophase VIs often reducing or smoothing many of the extreme values found in the original VIs. Peng, Wu, et al. (2017) used NDVI and EVI to predict spring phenophase dates across the U.S. and found the two often predicted very similar transition dates. This was similar for our study in which the VIs often predicted similar transition dates for the modeled VIs.

Using HLS data to predict phenology in grasslands has been attempted in the past with some success (Pastick et al., 2020). However, often the limited temporal resolution of the data has made it challenging to predict VI values across the growing season, or limits the spatial resolution to only specific vegetation types (Pastick et al., 2020). Limited spatial resolution has also been a challenge in the Great Plains region with the largest difficulty presented by a small spatial and temporal data collection period (Q. Zhou et al., 2019). Using our method, we attempted to overcome these challenges by coupling together data from multiple spatial and temporal resolutions. This provides a much stronger control on the phenology signal across the study area and normalizes differences between phenocam locations.

4.5 Conclusion

Using our phenology modeling approach, we were able to couple together data from both multiple spatial and temporal resolutions using MODIS, Landsat, Sentinel-2 and phenocam datasets. This allowed us to model phenology across our study region over a five-year period. With this we used XGB to model the phenology signal and were able to predict changes in grasslands across the Great Plains Region.

A challenge of our approach is using the method in a spatially complex location where multiple canopies are present and would interfere with one another. Y. Liu et al. (2017) found this when a grass understory mixed with a tree canopy that increased the complexity of the phenology signal. For our study region the vegetation is largely a single grassland or cropland canopy. However, croplands do introduce a complexity in which

cropland fields do not typically follow the same phenology signal as grasslands. Since we only based our method on grassland phenology, we expect this presented many of the errors in non-grassland locations. In the future this may be addressable by the addition of a phenocam dataset taken from agricultural locations, such as has been done for agricultural areas in Kazakhstan (De Beurs & Henebry, 2004).

Grasslands provide a relatively simple landscape when compared with forested areas that can have multiple canopy layers (Peng et al., 2021; A. D. Richardson et al., 2009). This provided us a more simplistic location to develop our phenocam modeling method. Transitioning these methods to a forest or location with multiple canopies would require further research into methods that can separate the phenology signal across complex vegetation regimes. While studies have been conducted to use digital repeat photography in these regions (Liang et al., 2011; Melaas, Friedl, et al., 2016; Sonnentag et al., 2012). Future research is needed in mixed canopy locations to see if these phenology signals can be separated, and if our modeling approach can be applied. Perhaps by developing a method that can produce multiple canopy datasets across complex environments such as these. In addition, future research can focus on refining and improving the quality of the modeled phenology. This includes using multiple vegetation types, such as grasslands and croplands, to determine if the accuracy of the modeled phenology can be improved when aggregating signals from various landcovers.

V. Discussion

In Ch. 2 we (Burke & Rundquist, 2021) scaled phenocam imagery to align each of the field locations with one another. To do this we fit a Gaussian Process (GP) model to each of six phenocam locations. Then we used a linear regression to align our GCC, NDVI, and EVI2 models with Harmonized Landsat Sentinel-2 satellite imagery (HLS). Later for our Ch. 3 and Ch. 4, we also carried out this same process using VCI to provide a similar metric for GCC, as we had EVI2 that uses the same spectral bands as NDVI. This provided a comparable band that could be used in discussion of the four VIs. Having the GP models, we were then able to adjust the VIs at each of the phenocam sites to align them with the HLS imagery and from here produce scaled VIs at the six locations that are comparable and can be used within a single model having scaled magnitudes.

The methods carried out in Ch. 1 were necessary for the two following chapters because the dataset produced was used in both Ch. 3 and Ch. 4. It provided the initial phenocam dataset taken across the six field locations and allowed them to be used within a single model since they were scaled to match one another. This chapter was published in (Burke & Rundquist, 2021), which was an important step in having the methods in the paper peer reviewed, before carrying on with the next stages of the research. The GP models produced did range in quality of fit to the raw phenocam imagery. Phenocam imagery can be quite noisy with R^2 of 0.73 at the Oakville phenocam for EVI2. Future research into methods to control model errors or improve modeling the phenocam site would be useful in producing

more accurate models and provide a stronger connection to the true phenology signal at the field site.

In Ch. 3 we used the scaled GCC, VCI, NDVI and EVI2 from Ch. 1 to model the vegetation growth over the six years. Using DAYMET weather data we either used already available weather variables or derived accumulated snow water equivalent, temperature, and precipitation. We used the available weather data from 1981 to 2019 to predict the weather variables across 38 years. Using this dataset, we were able to model the weather variables at the six phenocam locations using XGBoost. Using the five years of data we identified the transition dates between each of the phenophases. Across all four VI models, we found day length, GDD with a base of 0°C to be the two most important variables, while other variables that ended up in the models includes growing degree days with a base temperature of 5°C, minimum temperature, 30-day of accumulated precipitation, and minimum temperature. Using our developed 38-year model we were able to predict the phenology signal across the six phenocam sites and identify and significant trends in the models. We found 14 of the VIs across the five of the six phenocam sites had significant trends in phenophase transitions, excluding the Nine-Mile station. This chapter identified phenology trends at the six phenocam stations within the North American Great Plains. It provides a valuable method to identify phenophase transition and to do this with coupled together datasets, and to extend the phenology datasets into the past using historical weather data as a substitute for phenocam imagery.

Ch. 3 provides the review of the temporal weather data, examining how the VIs at the six phenocam sites changed across 38 years. It allows us to examine the vegetation even before the phenocams were available at the six field locations. This means that the data before 2015 was entirely synthetic, based on historical weather data, and is not taken from imagery. Having the imagery may provide insight into a changing vegetation dynamic that occurred over the 38 years, but without the ability to get this historic imagery, we felt our method was the best option. The XGB model we used predicted 90 to 93 percent of the variability in the four VIs. Further refinement in the modeling procedure, or additional weather data, could improve the model accuracy. Other weather data may be available over this time that could help improve the accuracy of the model as well.

In Ch. 4 we used MODIS, Harmonized Landsat-Sentinel-2, and phenocam imagery, to model phenology across a portion of the Great Plains region. To do this we relied on 1,816 MODIS, and 2,027 HLS images taken from the six phenocam locations from 2015 to 2019. We removed cloud contaminated imagery from the modeling process and the carried-out imagery fusion to produce synthetic HLS imagery from the MODIS imagery. XGB was used to predict the VIs values across the five years. Using both standard and modeled VIs we predicted the phenophase dates across the 5 years of data, and across the four VIs. We found GCC produced the best R^2 values at 0.894, while EVI2 had the lowest R^2 values at 0.774. In testing each of the VIs by removing one of the six datasets before training we identified a similar trend where GCC was the best performing VI with an average R^2 of 0.734 while NDVI had the lowest average R^2 of 0.587.

We were able to model the four VIs across the Great Plains region with R^2 values ranging from 0.894 to 0.774. This paper provided the spatial method for modeling phenology across the Great Plains region. With this method we were able to predict the VIs values at any given location using the Harmonized Landsat-Sentinel-2 imagery. Our standard and modeled VIs were similar performing with the modeled VIs often having a lower spread in values suggesting a better control on VI values. Both the standard and modeled VIs did however have model errors in the 2019 VIs, where phenophase number 4, representing peak greenness often occurred at the end of the year, instead of the middle of the season where it was expected. Future work for this method may require incorporating data from additional locations or satellite sensors. Vegetation growth data can have a high degree of variability and incorporating more data, including data from outside the grasslands, such as the croplands may provide a more accurate description of the vegetation dynamics.

APPENDICES

Appendix A: HLS GP Results

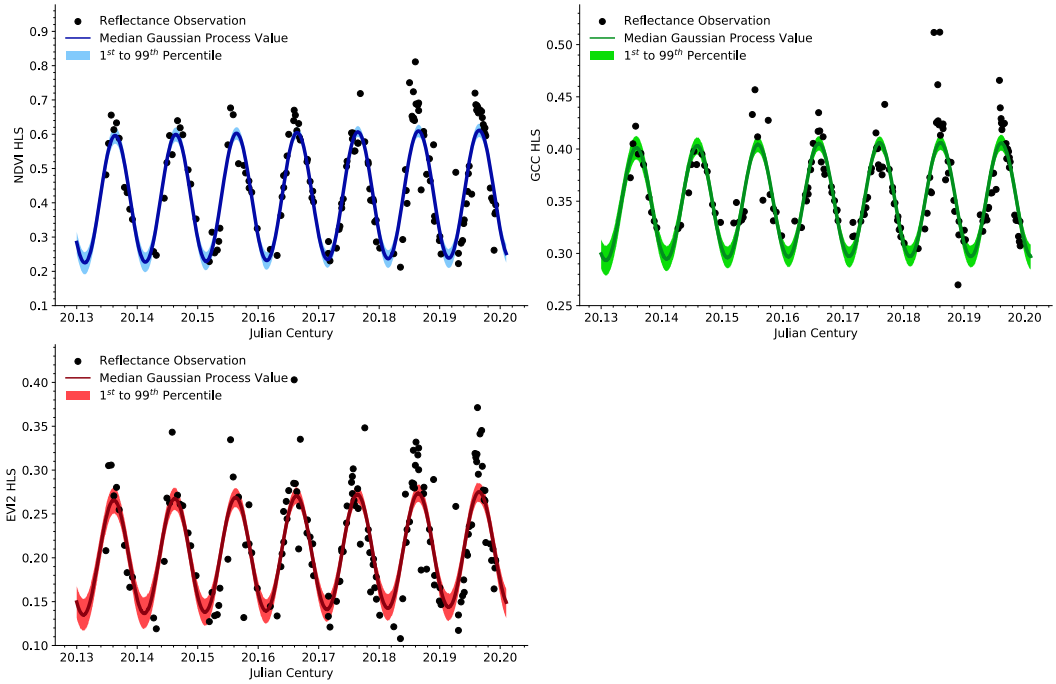


Figure A.1: HLS GP model for the USGSEROS station, showing the median predicted value fit to the seven years of IDW HLS data.

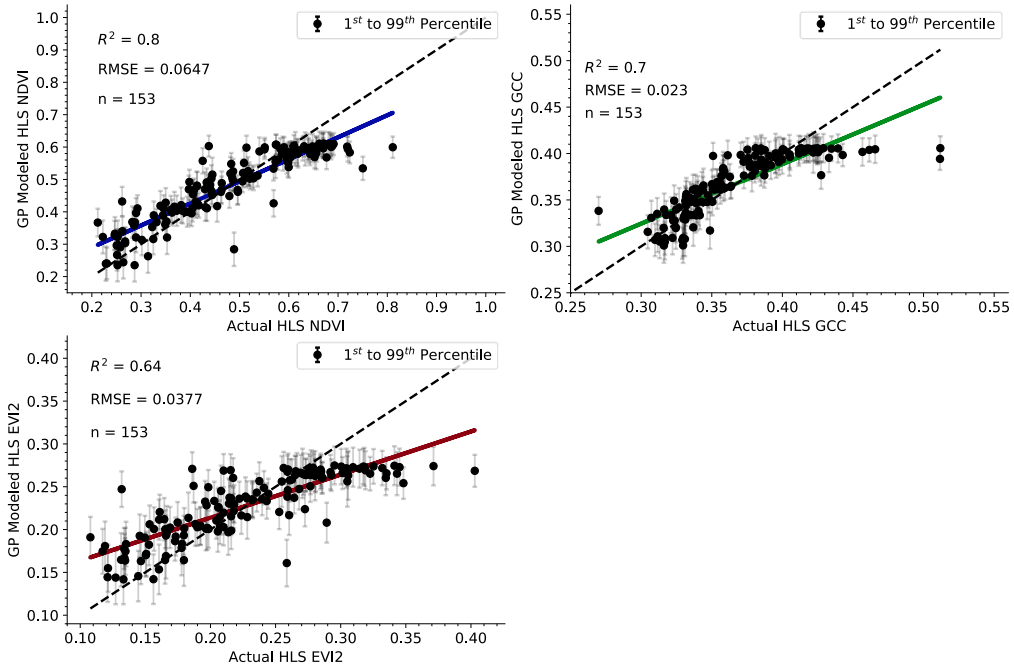


Figure A.2: Scatter Plot of Actual IDW HLS VI values against the GP modeled HLS VI values for the USGSEROS station. The dashed line represents an ideal one-to-one linear fit where y=x (1:1 line).

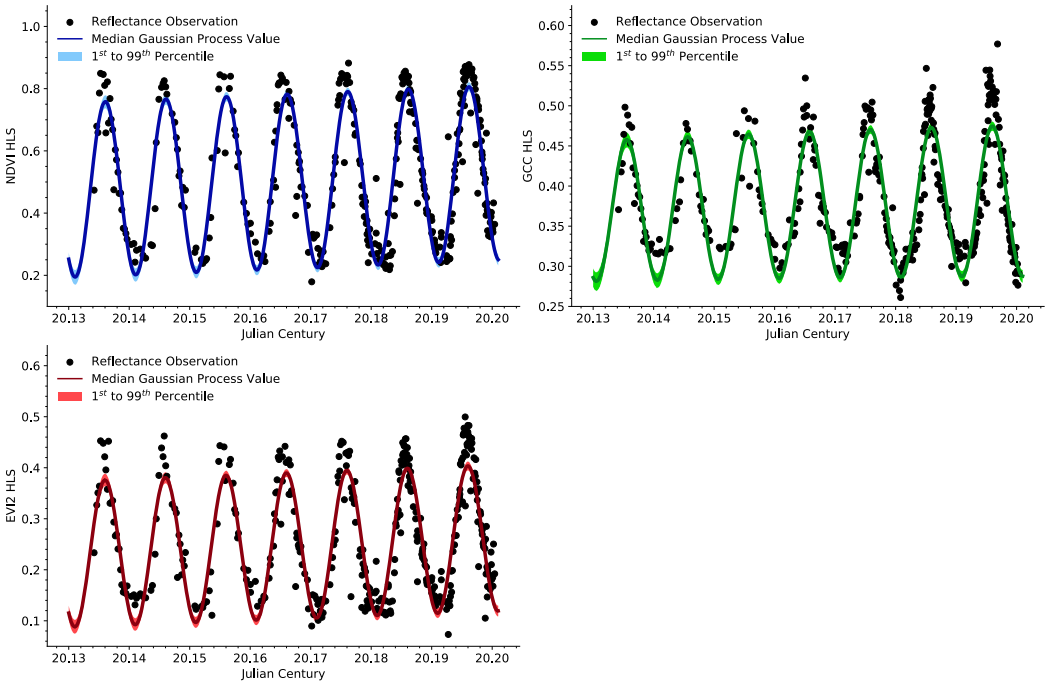


Figure A.3: HLS GP model for the Nine-Mile station, showing the median predicted value fit to the seven years of IDW HLS data.

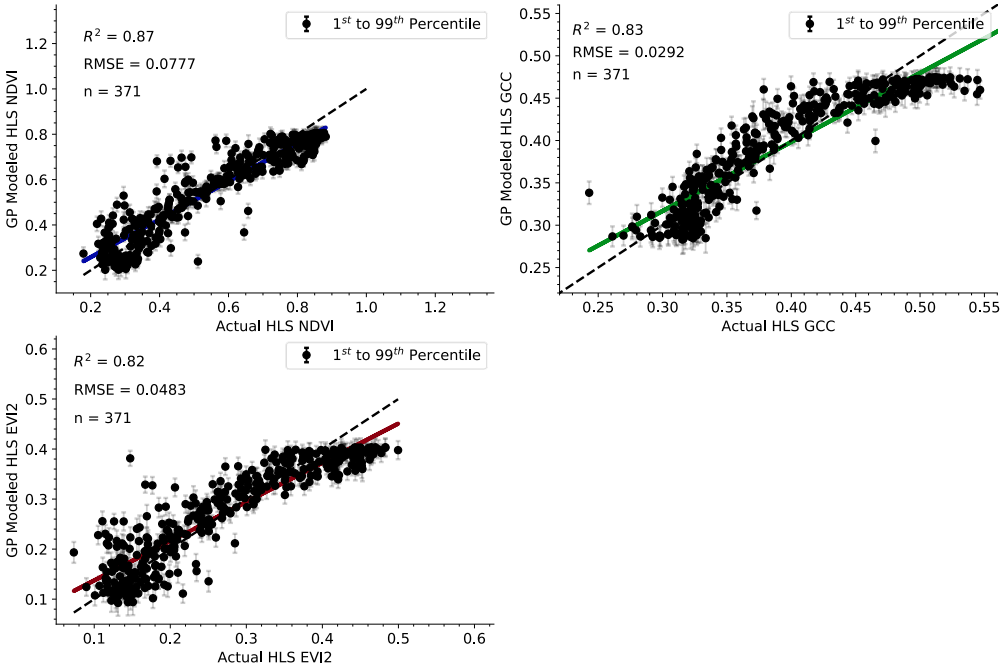


Figure A.4: Scatter Plot of Actual IDW HLS VI values against the GP modeled HLS VI values for the Nine-Mile station. The dashed line represents an ideal one-to-one linear fit where y=x (1:1 line).

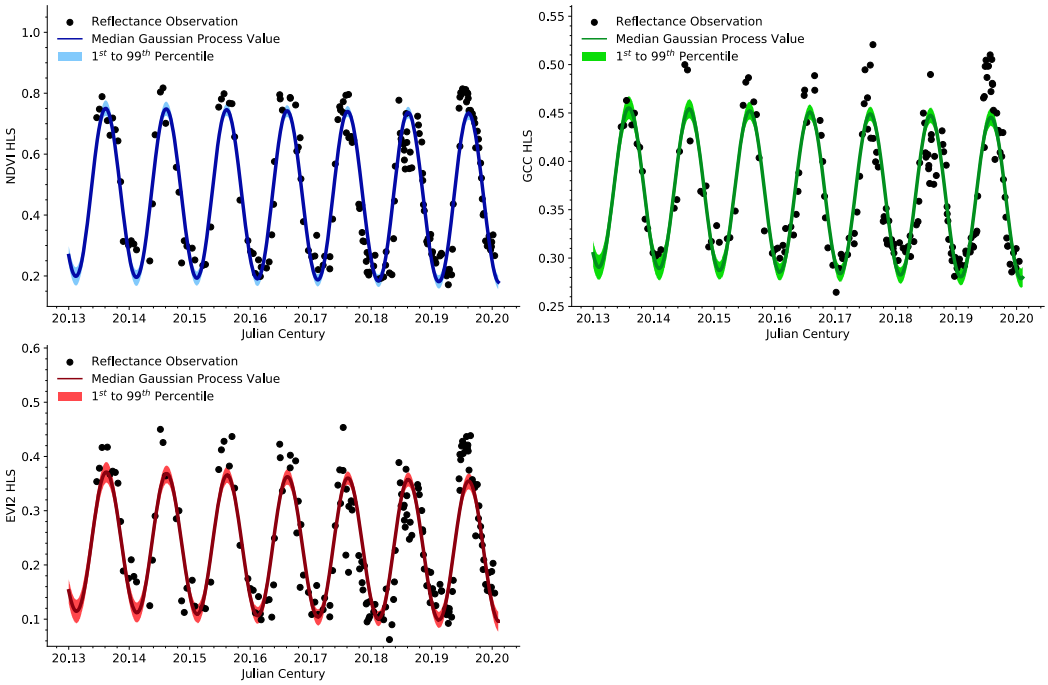


Figure A.5: HLS GP model for the Konza station, showing the median predicted value fit to the seven years of IDW HLS data.

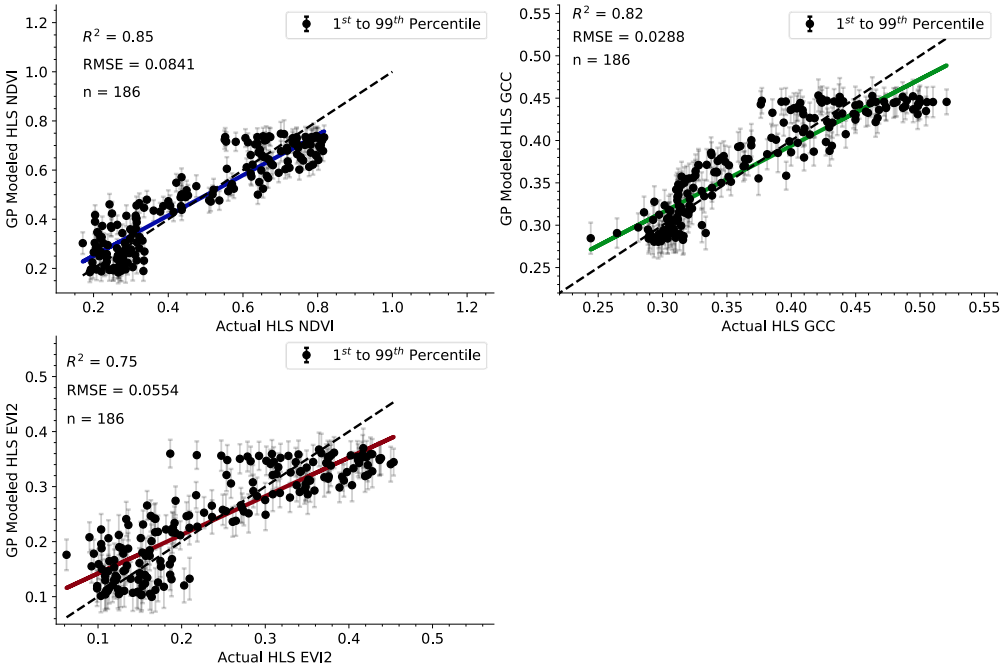


Figure A.6: Scatter Plot of Actual IDW HLS VI values against the GP modeled HLS VI values for the Konza station. The dashed line represents an ideal one-to-one linear fit where y=x (1:1 line).

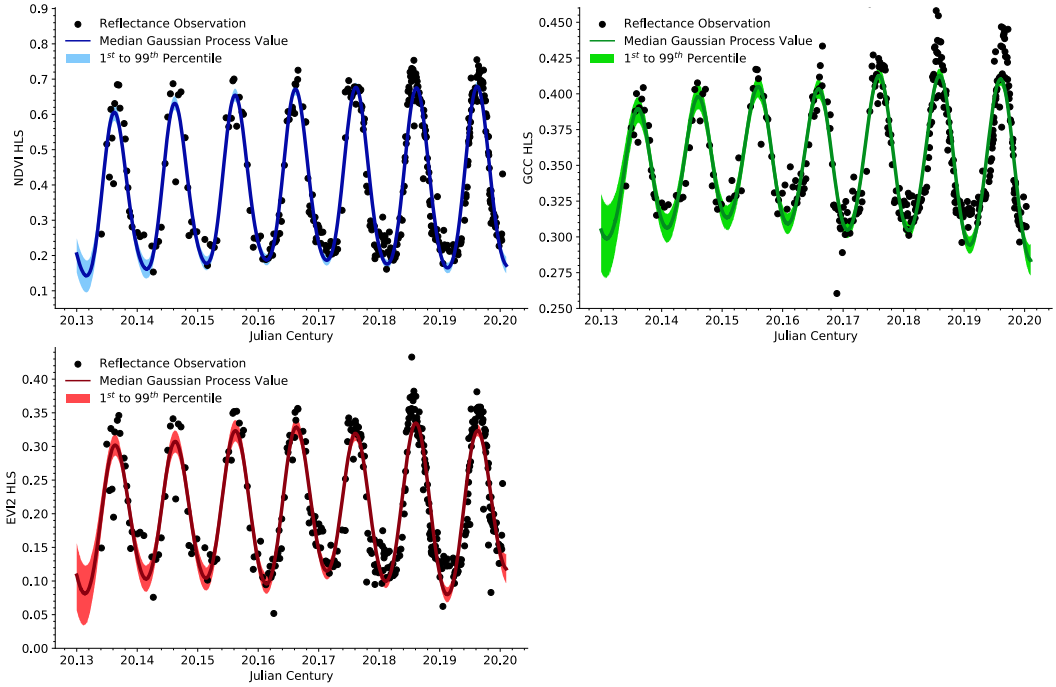


Figure A.7: HLS GP model for the ARIK station, showing the median predicted value fit to the seven years of IDW HLS data.

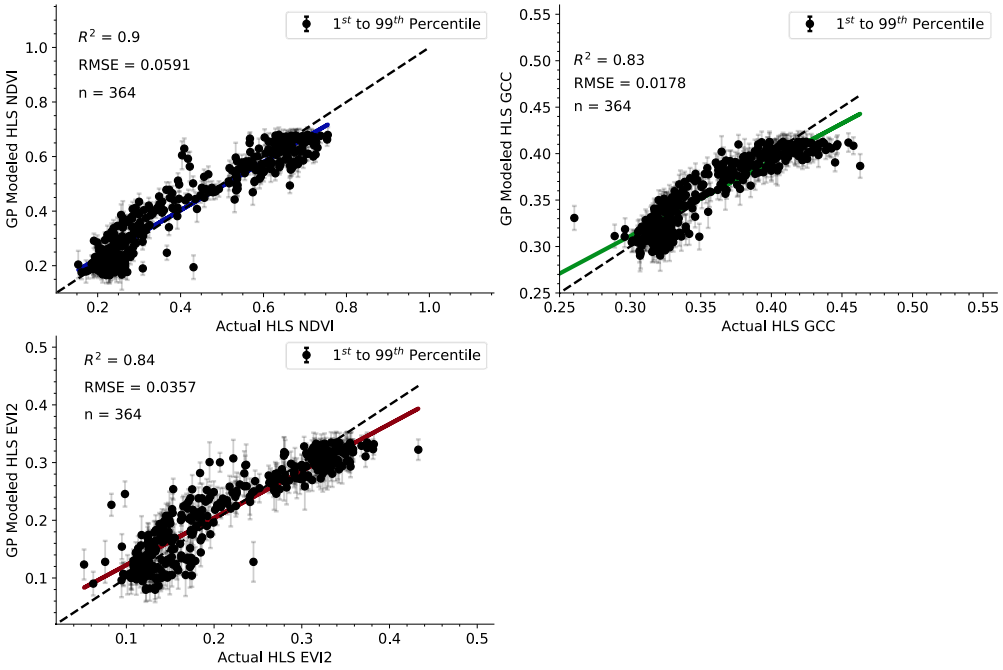


Figure A.8: Scatter Plot of Actual IDW HLS VI values against the GP modeled HLS VI values for the ARIK station. The dashed line represents an ideal one-to-one linear fit where y=x (1:1 line).

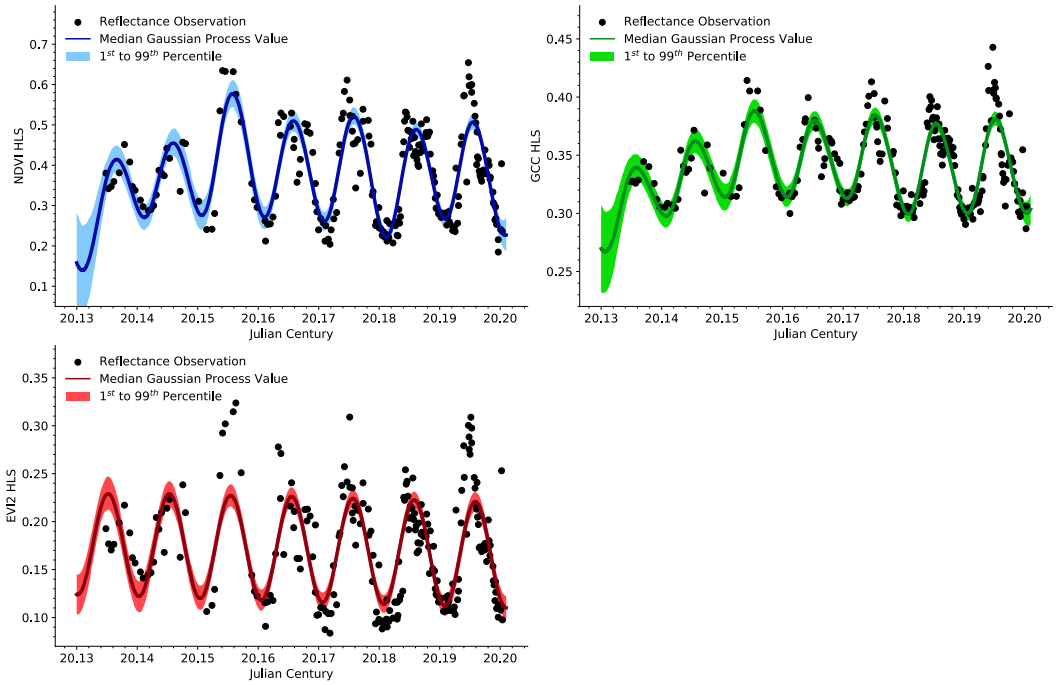


Figure A.9: HLS GP model for the OAES station, showing the median predicted value fit to the seven years of IDW HLS data.

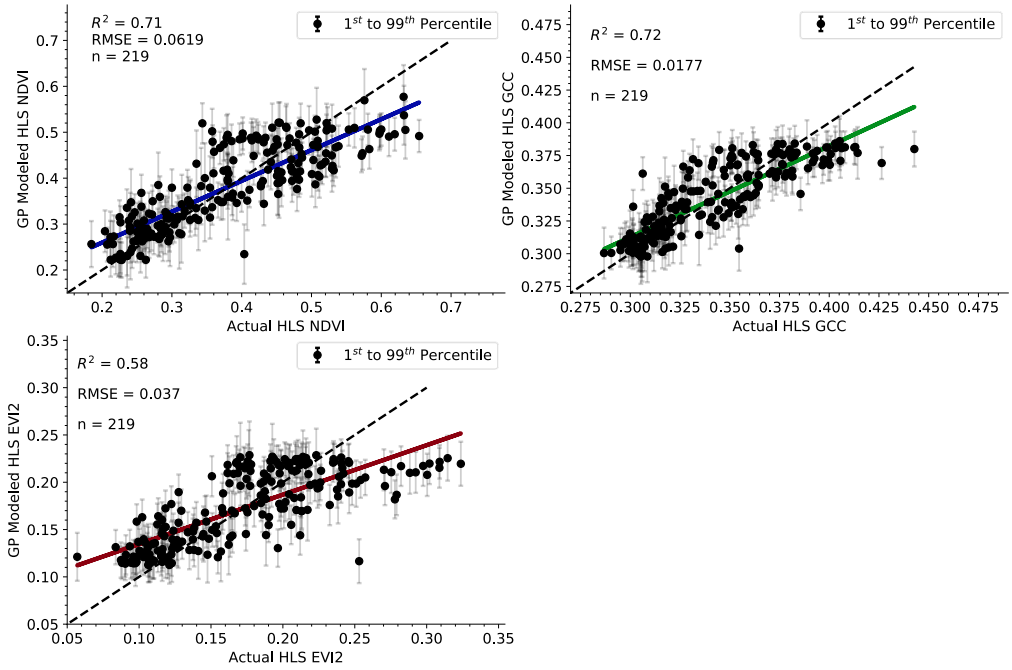


Figure A.10: Scatter Plot of Actual IDW HLS VI values against the GP modeled HLS VI values for the OAES station. The dashed line represents an ideal one-to-one linear fit where y=x (1:1 line).

Appendix B: Phenocam VIs Comparing Observations with GP Models

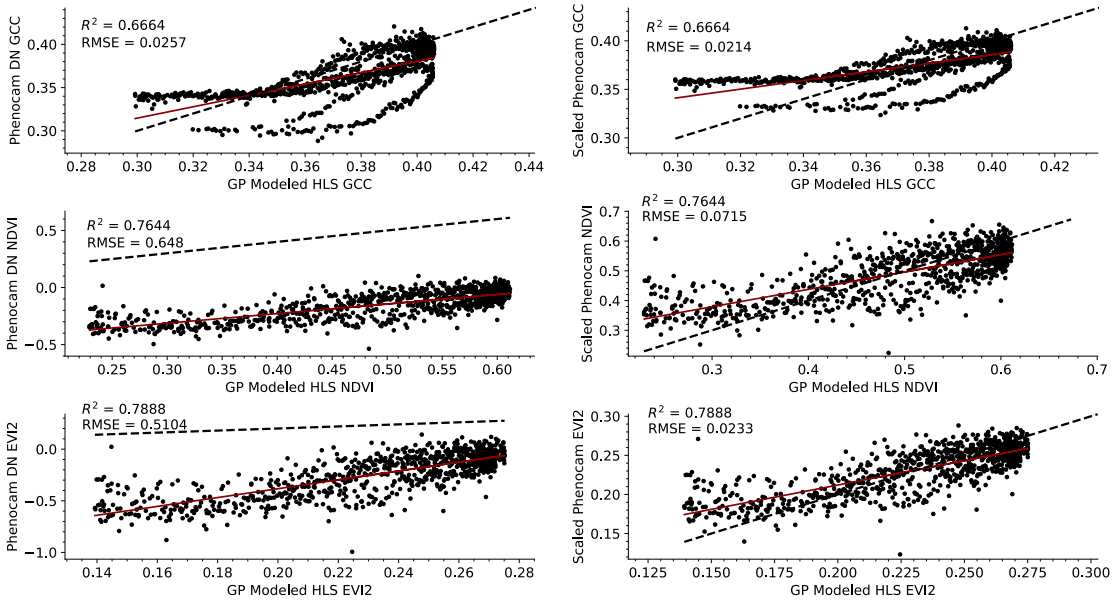


Figure B.1: VIs calculated for the USGSEROS station, showing linear regressions between both the phenocam DN 90th percentile VIs and the scaled phenocam VIs against the scaled GP modeled HLS VIs. The dashed line shows an ideal one-to-one linear fit where y=x (1:1 line).

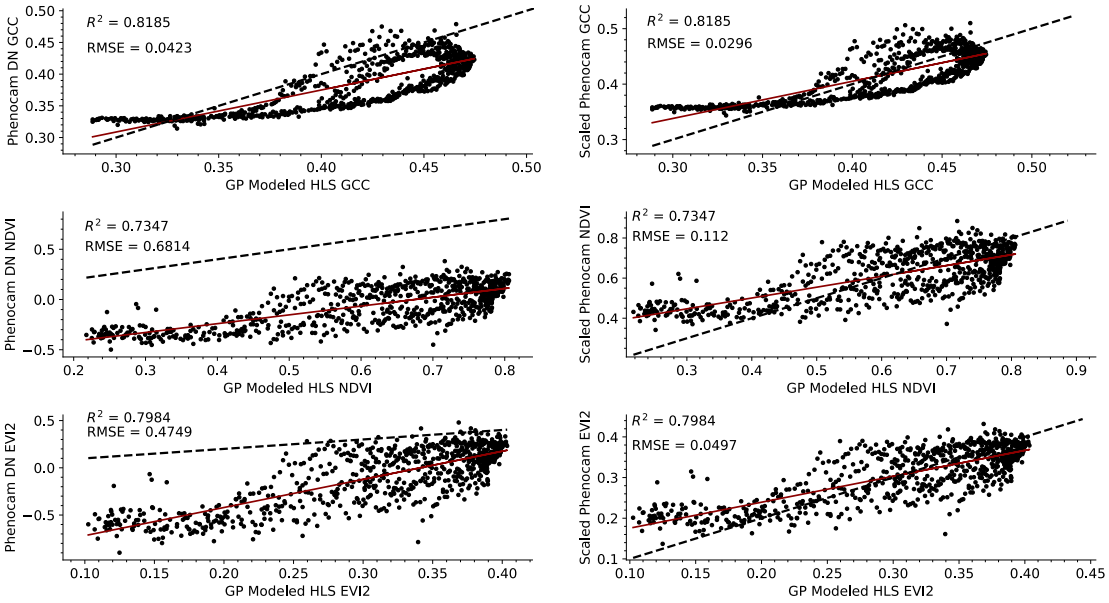


Figure B.2: VIs calculated for the Nine-Mile station, showing linear regressions between both the phenocam DN 90th percentile VIs and the scaled phenocam VIs against the scaled GP modeled HLS VIs. The dashed line shows an ideal one-to-one linear fit where y=x (1:1 line).

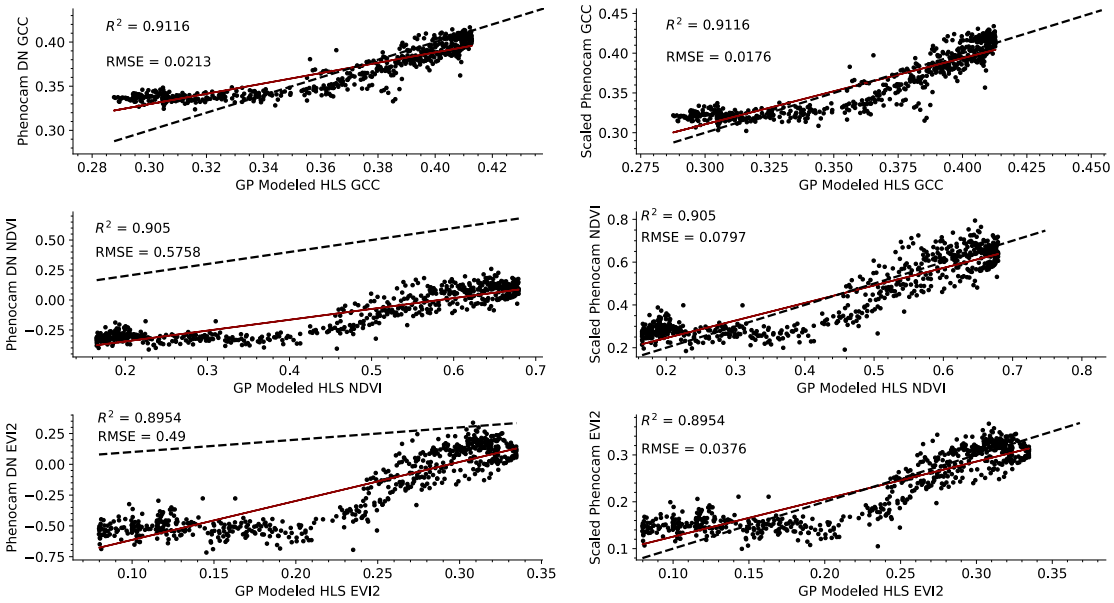


Figure B.3: VIs calculated for the ARIK station, showing linear regressions between both the phenocam DN 90th percentile VIs and the scaled phenocam VIs against the scaled GP modeled HLS VIs. The dashed line shows an ideal one-to-one linear fit where y=x (1:1 line).

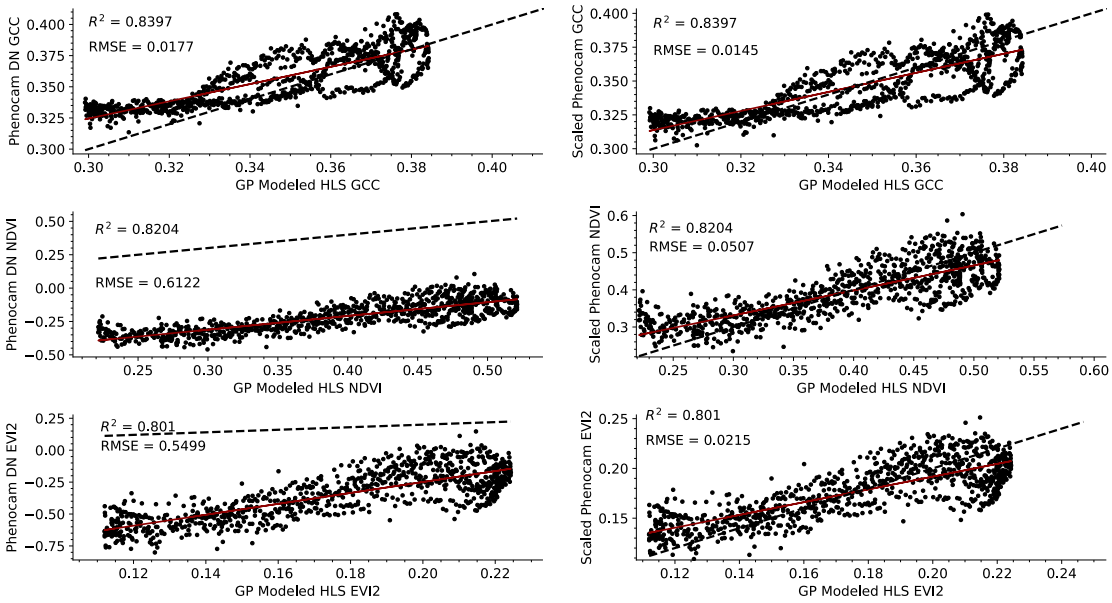
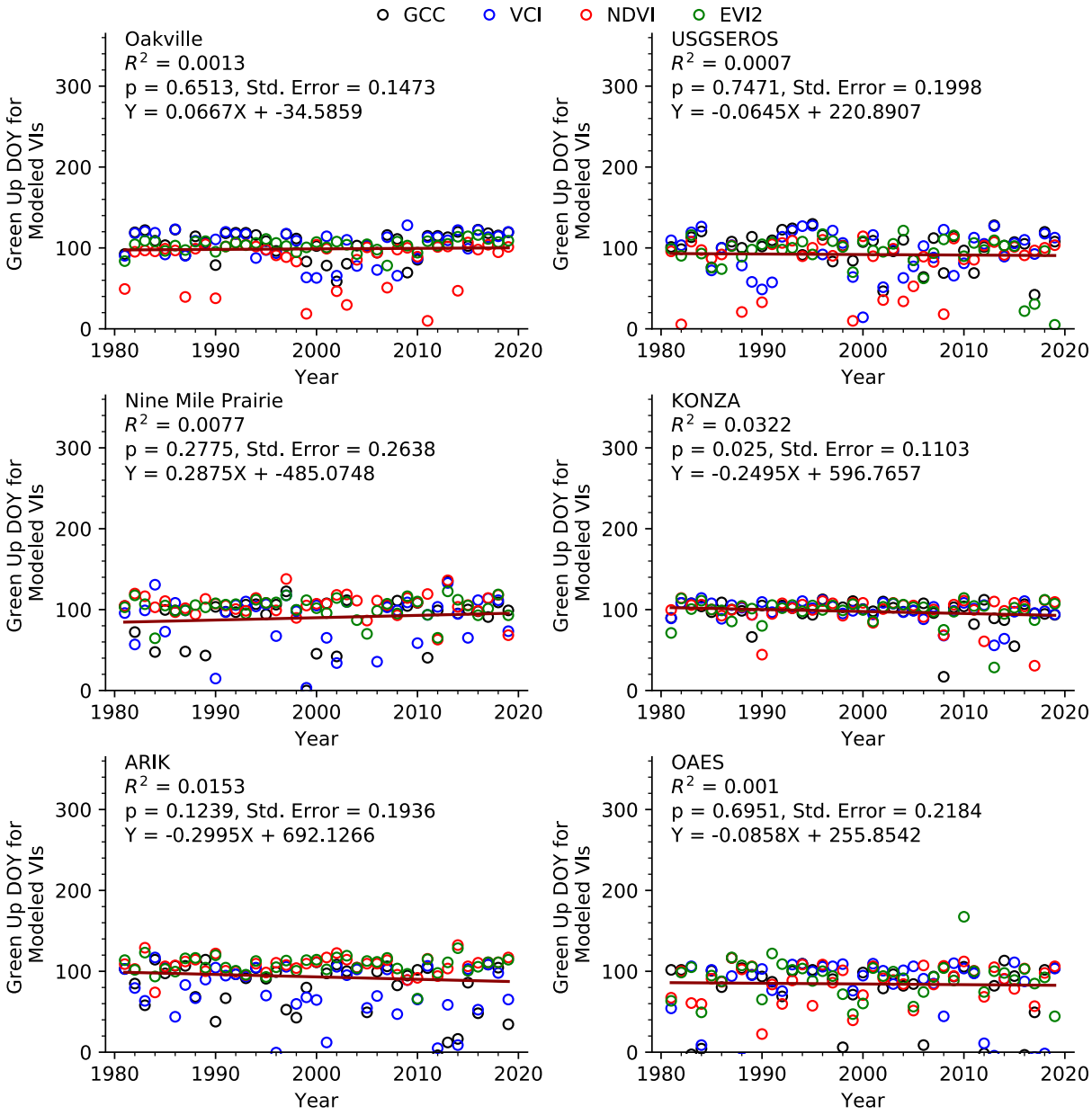
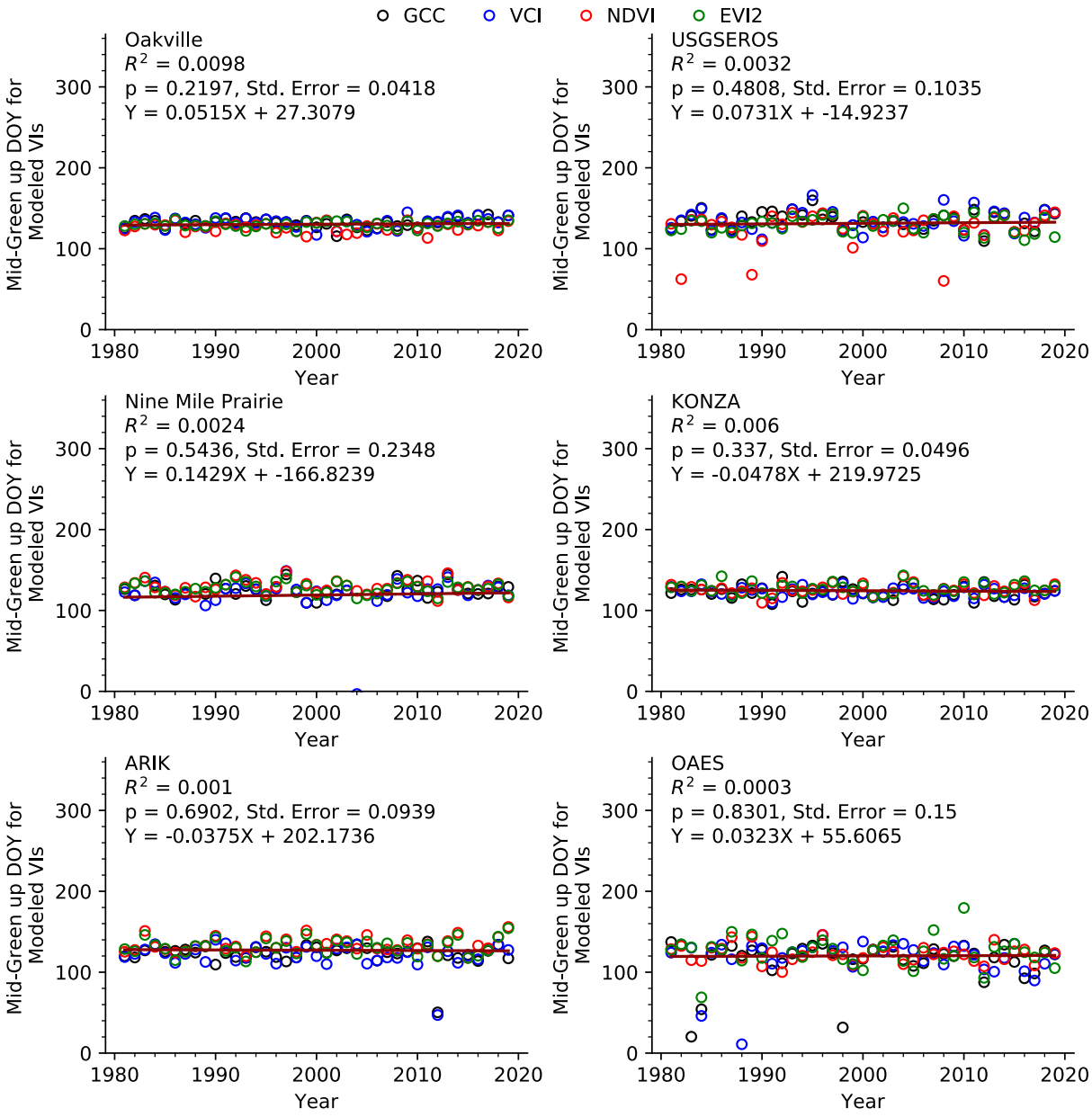


Figure B.4: VIs calculated for the OAES station, showing linear regressions between both the phenocam DN 90th percentile VIs and the scaled phenocam VIs against the scaled GP modeled HLS VIs. The dashed line shows an ideal one-to-one linear fit where y=x (1:1 line).

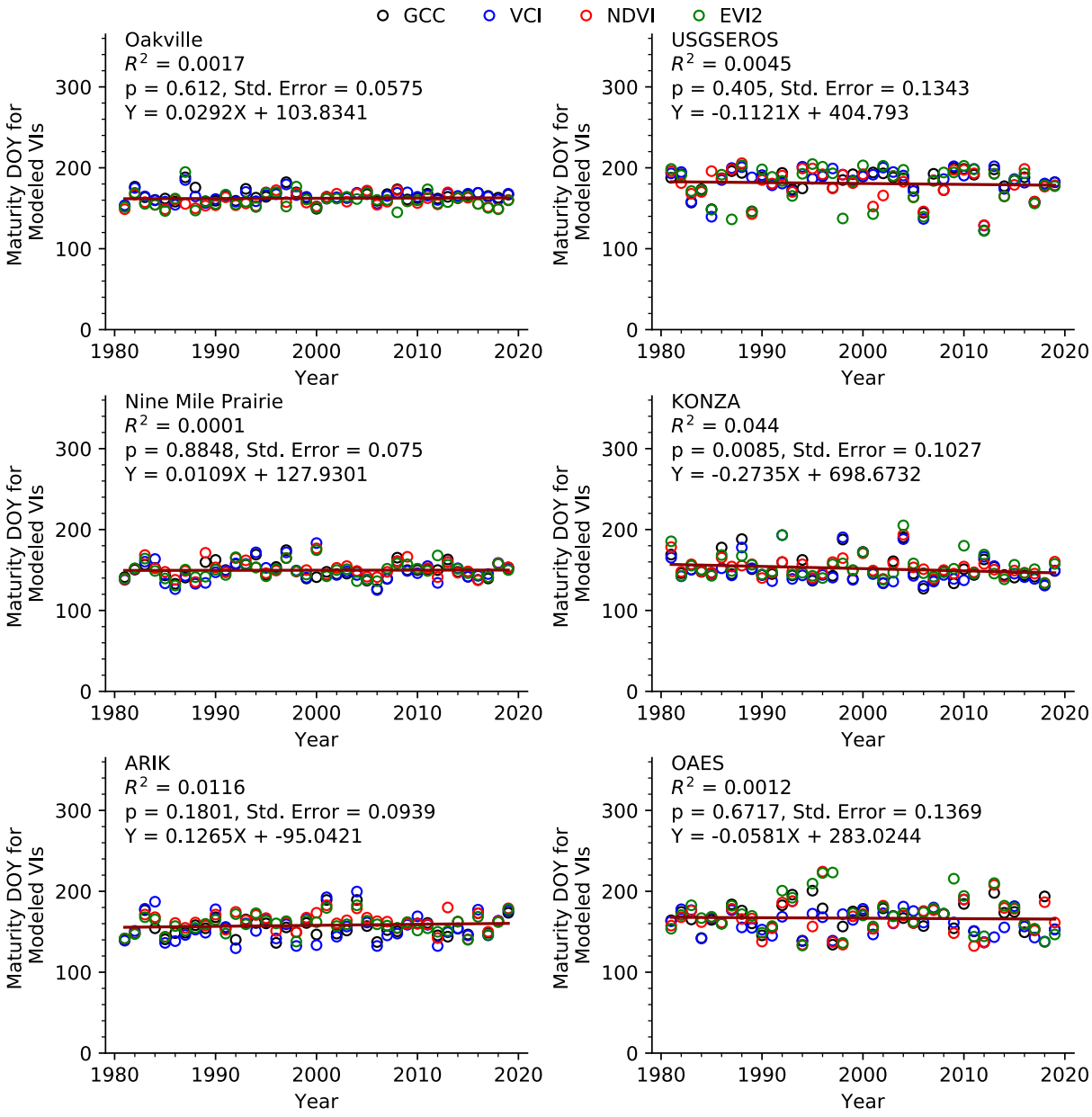
Appendix C: Linear Regression Modeling the Trends for the XGB Modeled Phenophases Across the Six Phenocam Sites



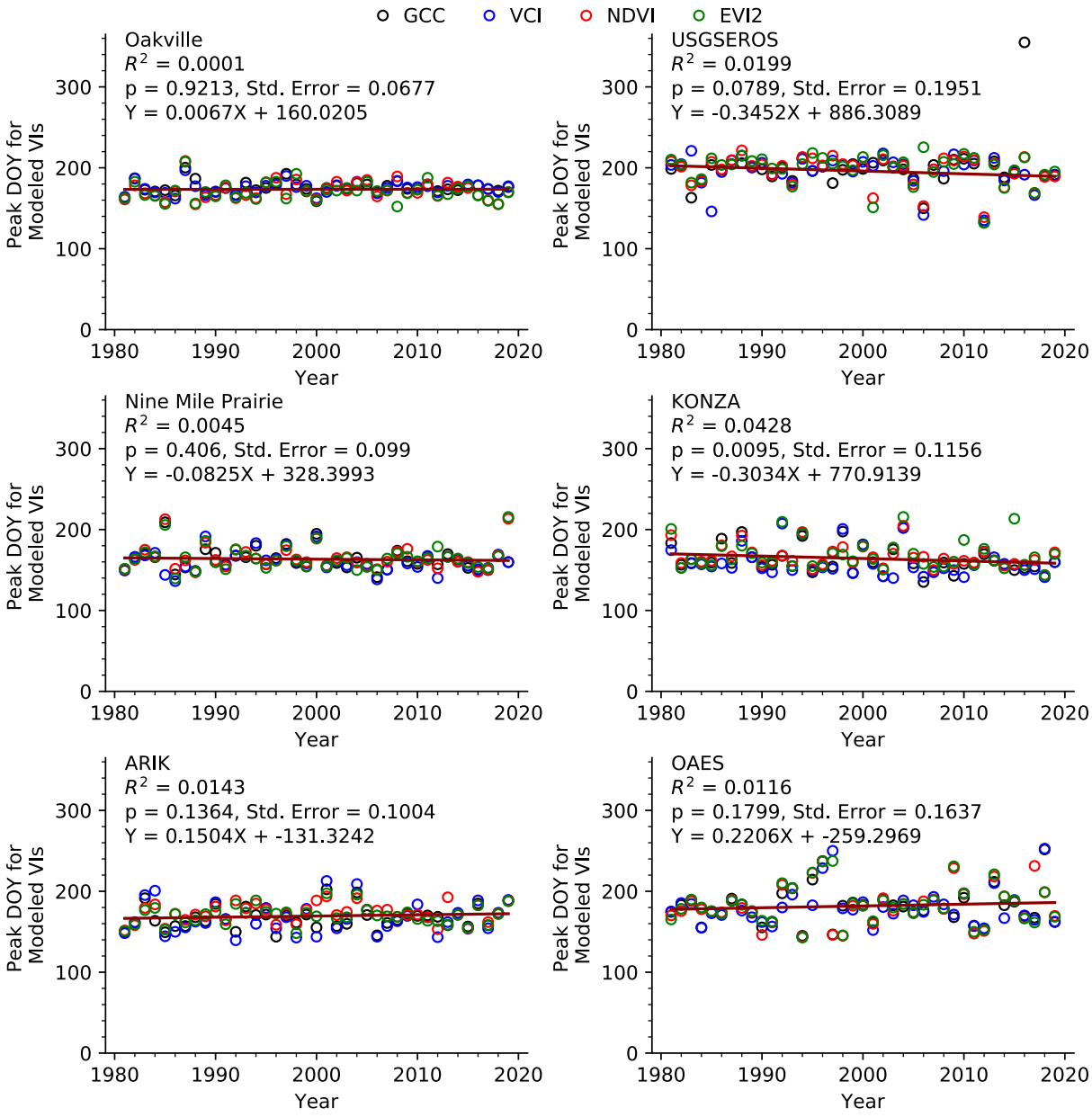
Appendix C.1: The day of year (DOY) in which the greenup phenophase was detected for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



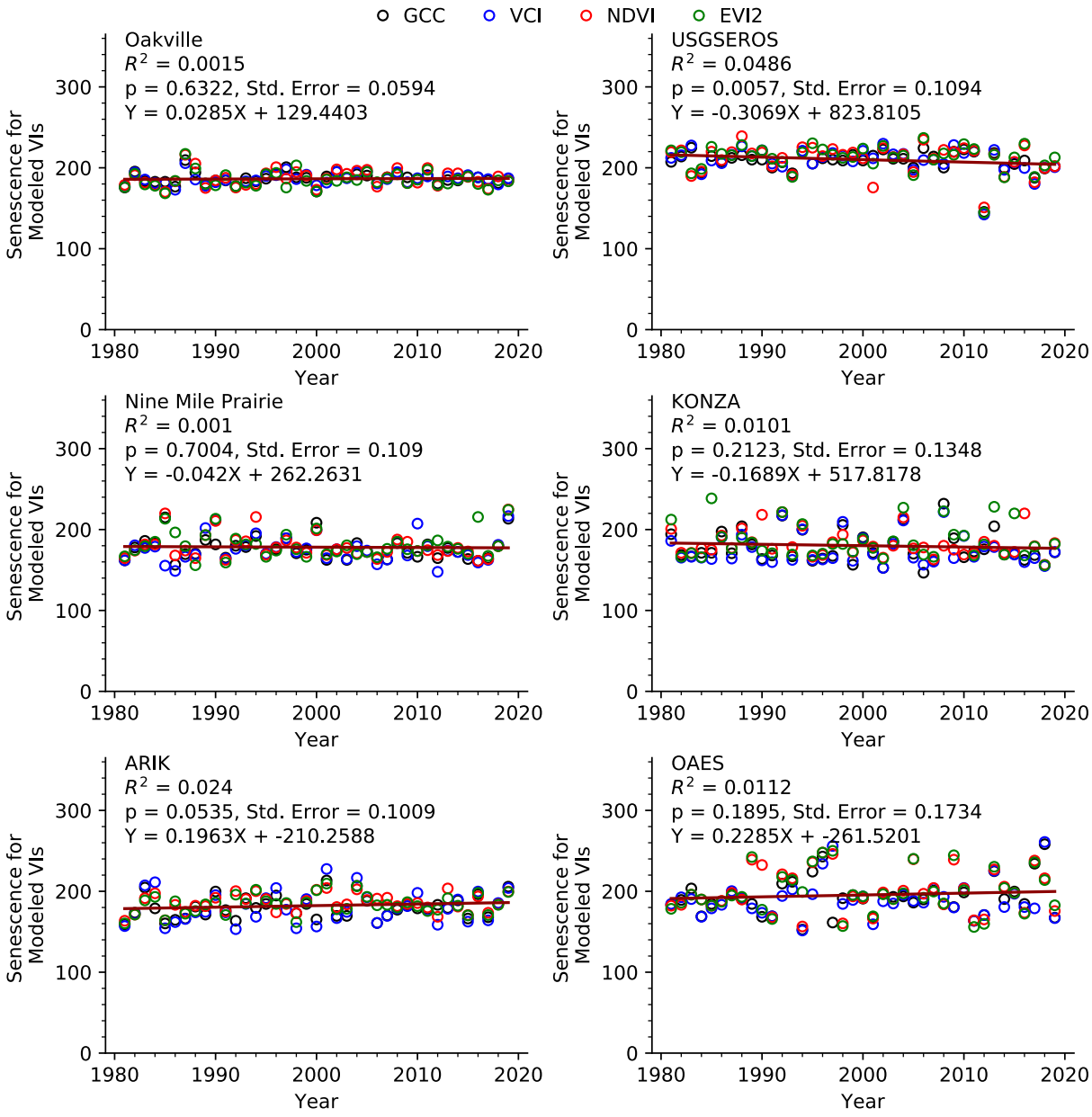
Appendix C.2: The day of year (DOY) in which the mid-greenup phenophase was detected for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



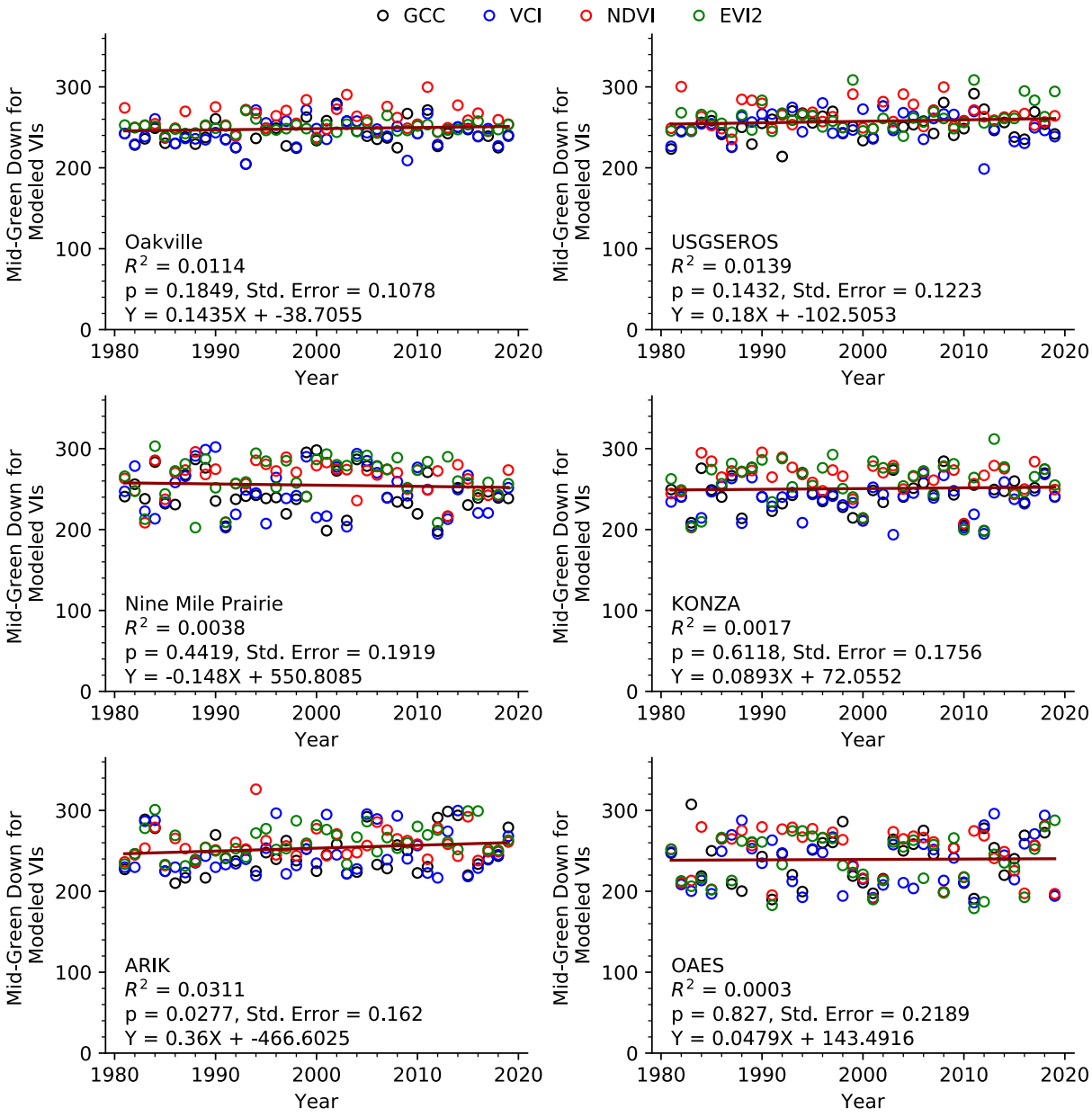
Appendix C.3: The day of year (DOY) in which the maturity phenophase was detected for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



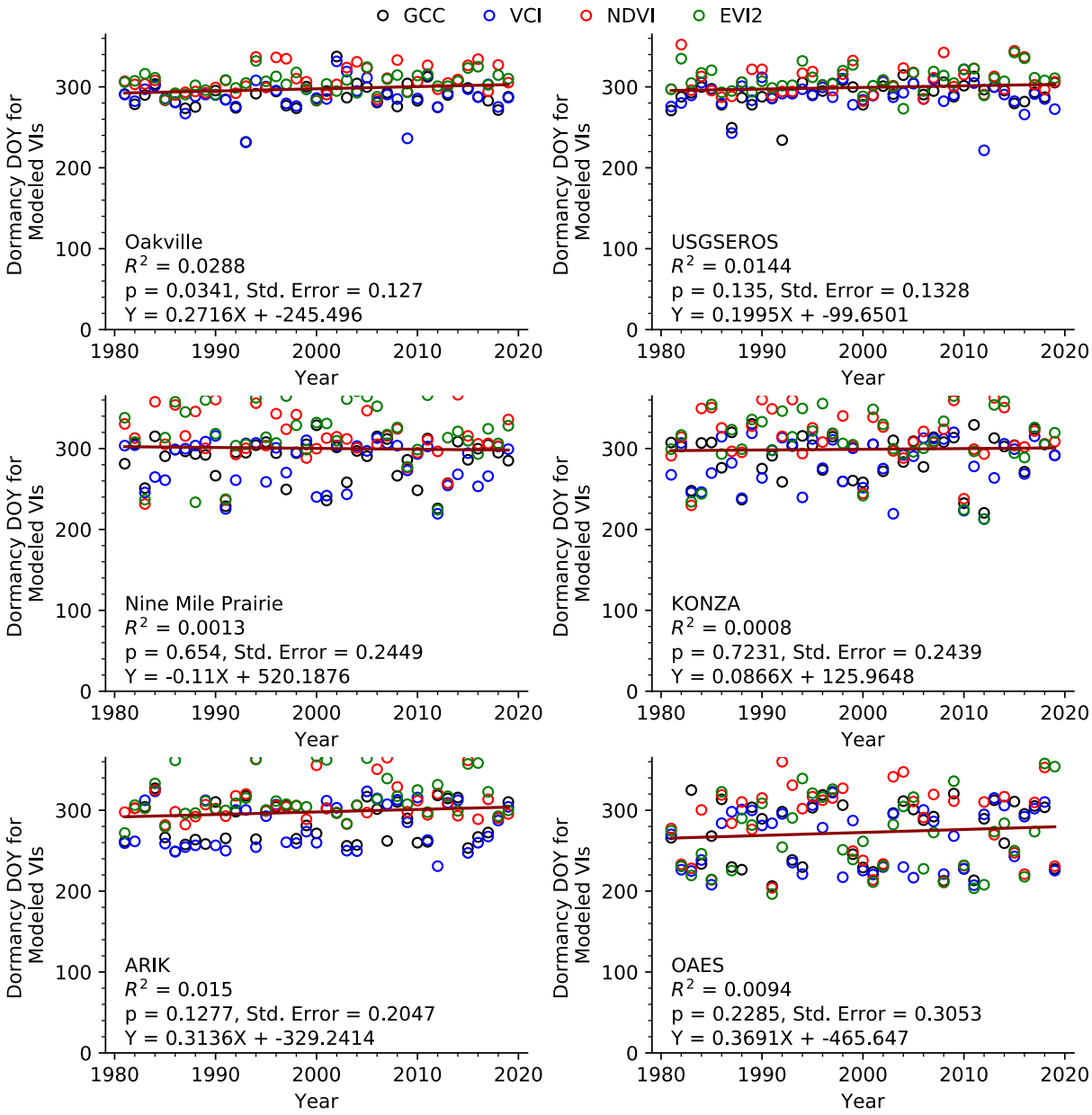
Appendix C.4: The day of year (DOY) in which the peak phenophase was detected for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



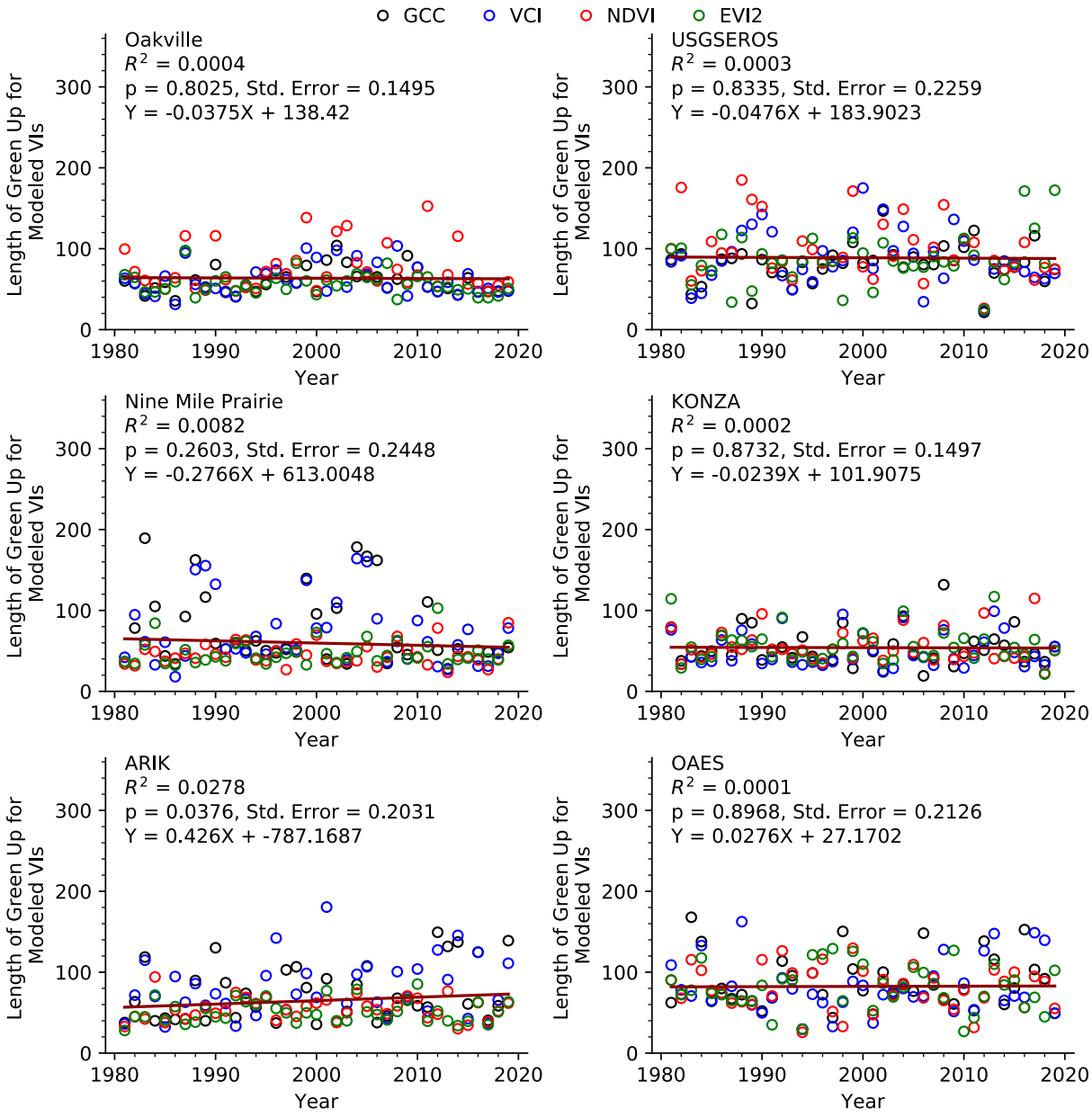
Appendix C.5: The day of year (DOY) in which the senescence phenophase was detected for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



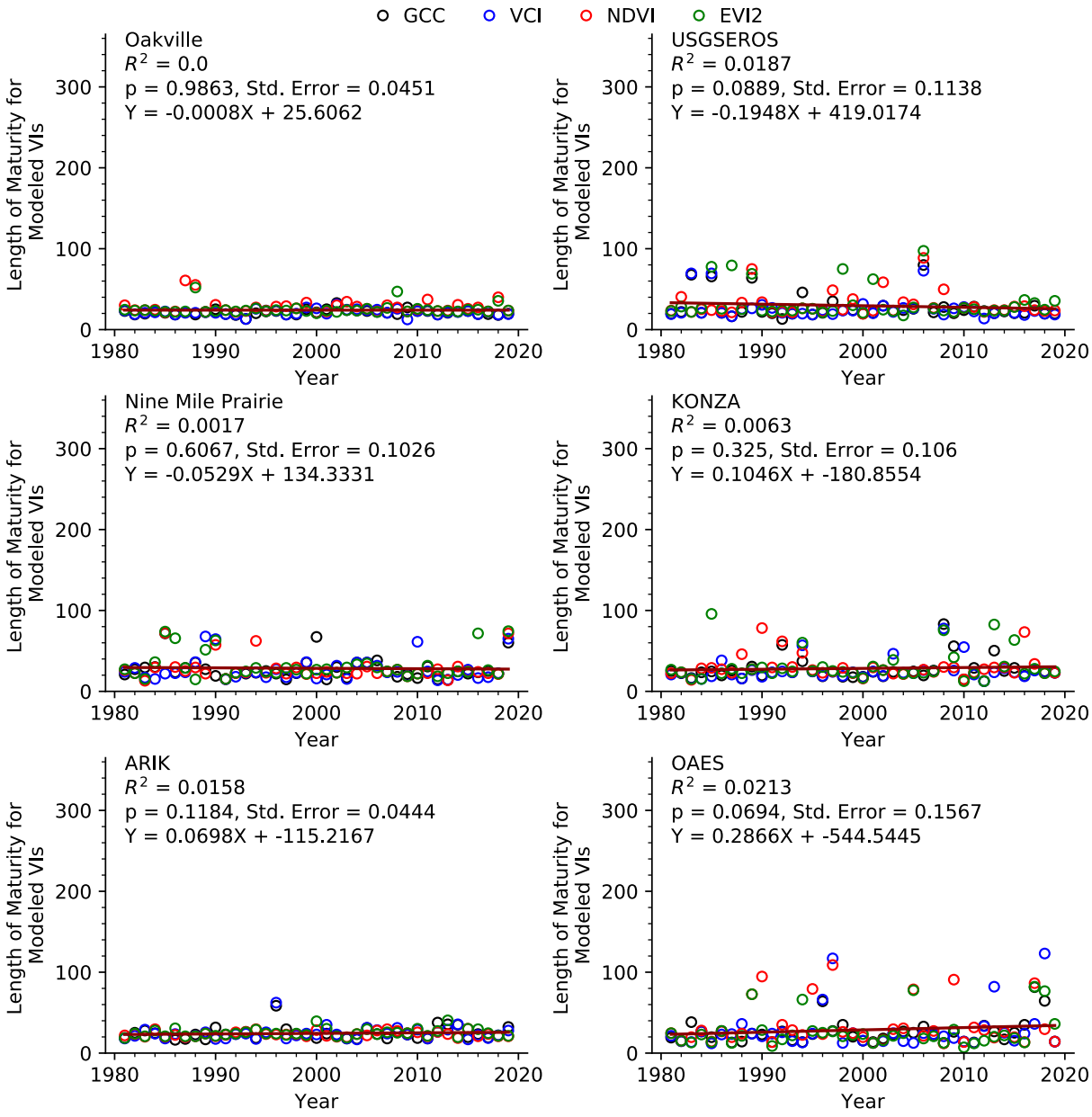
Appendix C.6: The day of year (DOY) in which the mid-greendown phenophase was detected for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



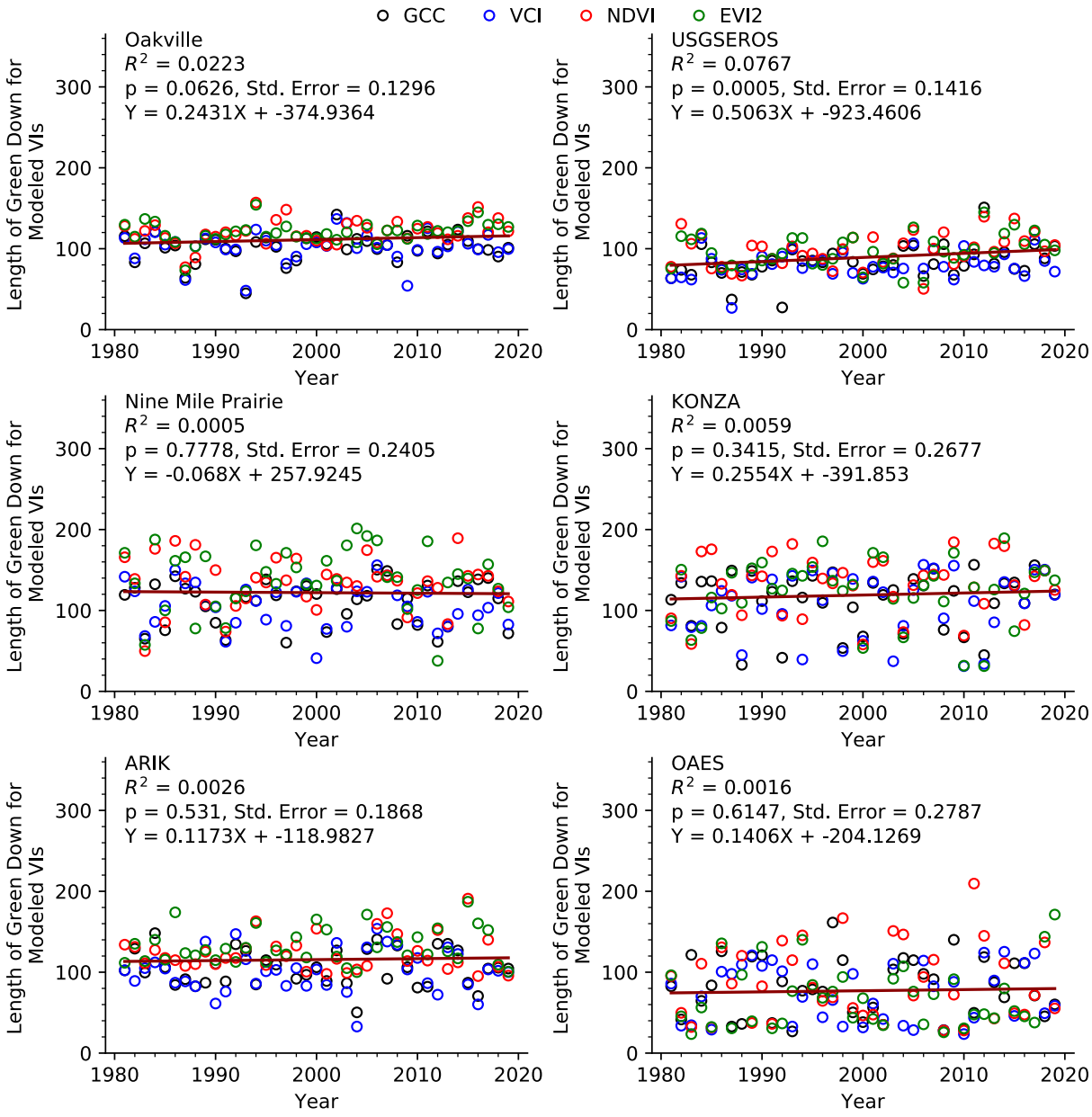
Appendix C.7: The day of year (DOY) in which the dormancy phenophase was detected for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



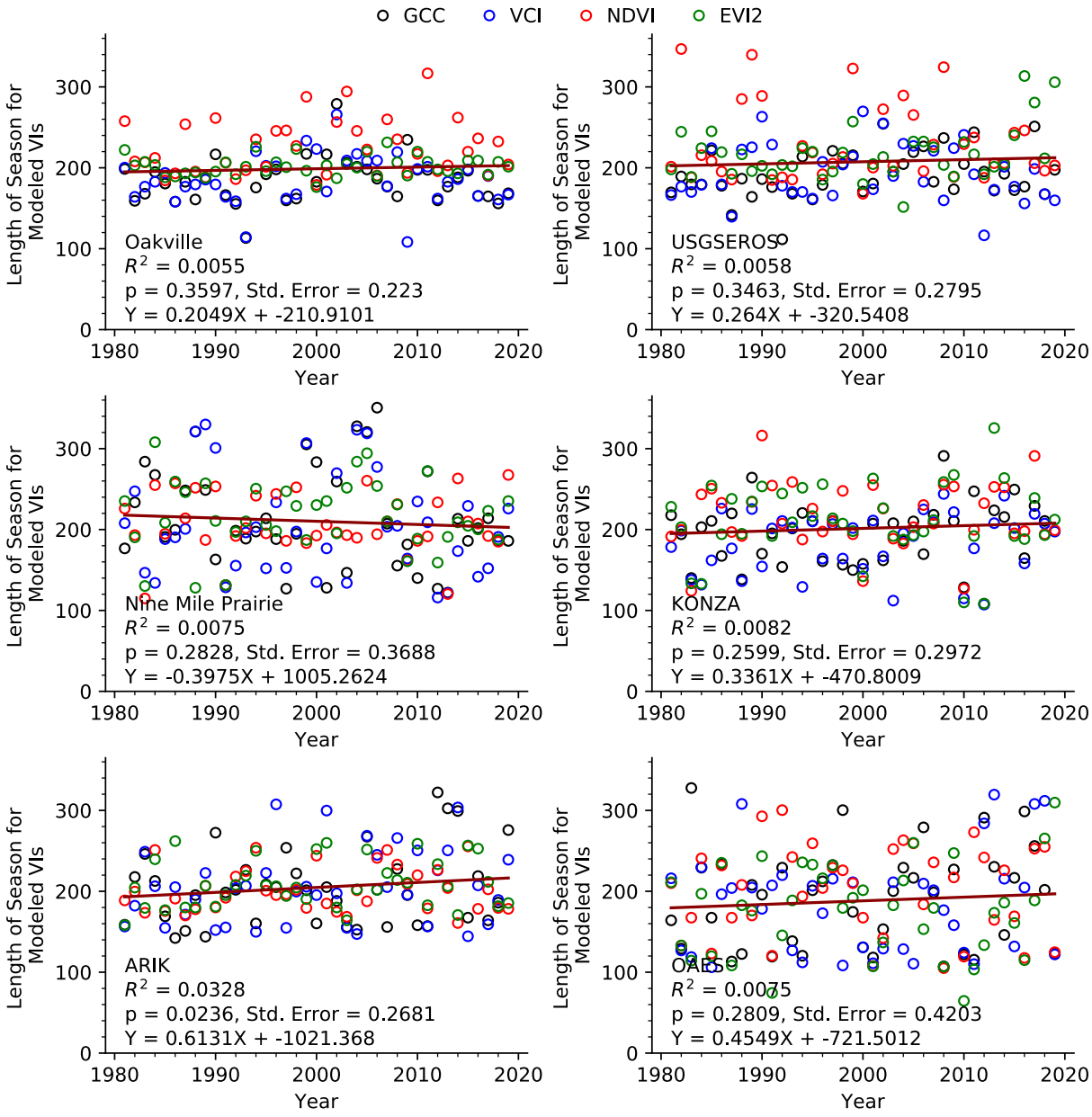
Appendix C.8: The length of time in days between the greenup and maturity phenophases for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



Appendix C.9: The length of time in days between the maturity and senescence phenophases for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



Appendix C.10: The length of time in days between the senescence and dormancy phenophases for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.



Appendix C.11: The length of time in days between the greenup and dormancy phenophases for the four VIs across the six phenocam sites. Showing a linear regression for each of the phenocam sites with all four VIs used in the regression model.

Appendix D: Phenophase Dates Across the VIs and Five Years

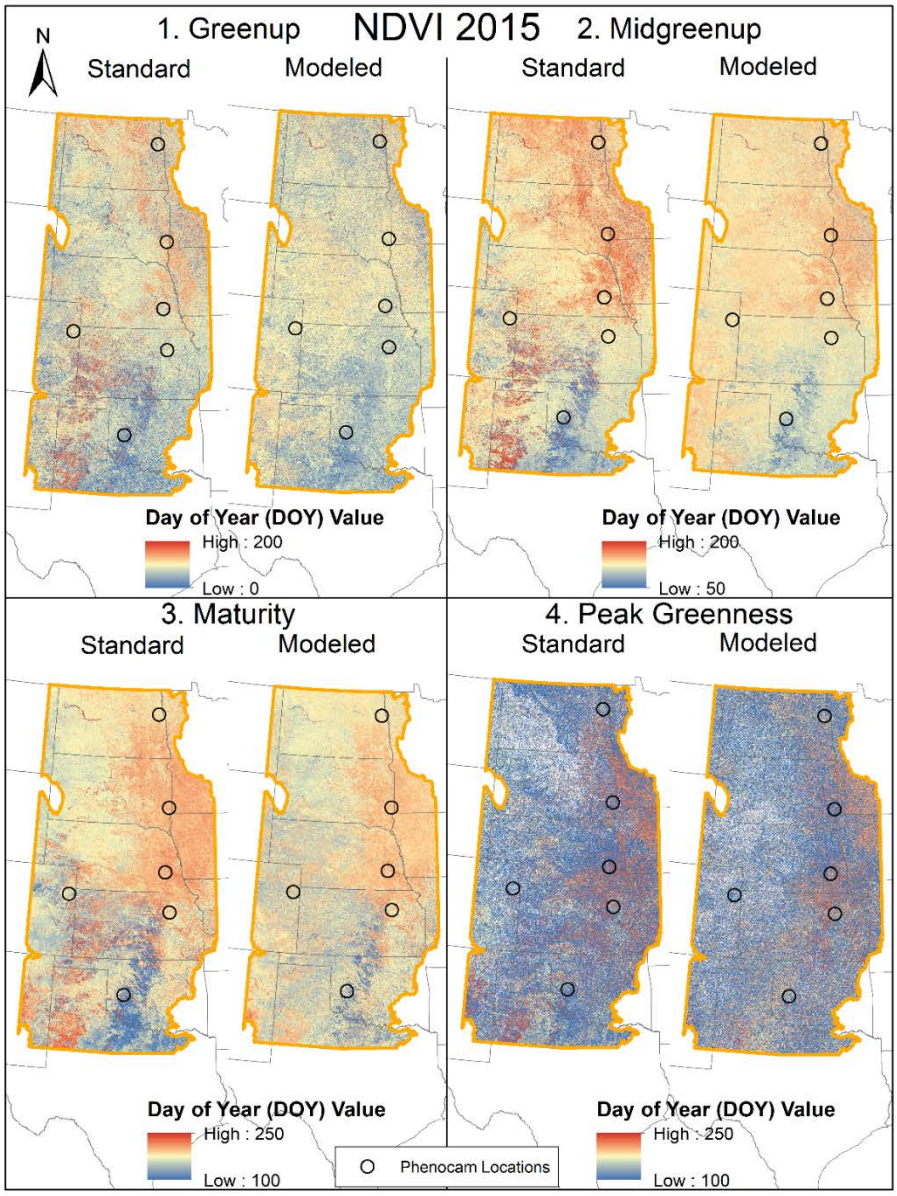


Figure A1: Maps showing the date of the first four NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2015.

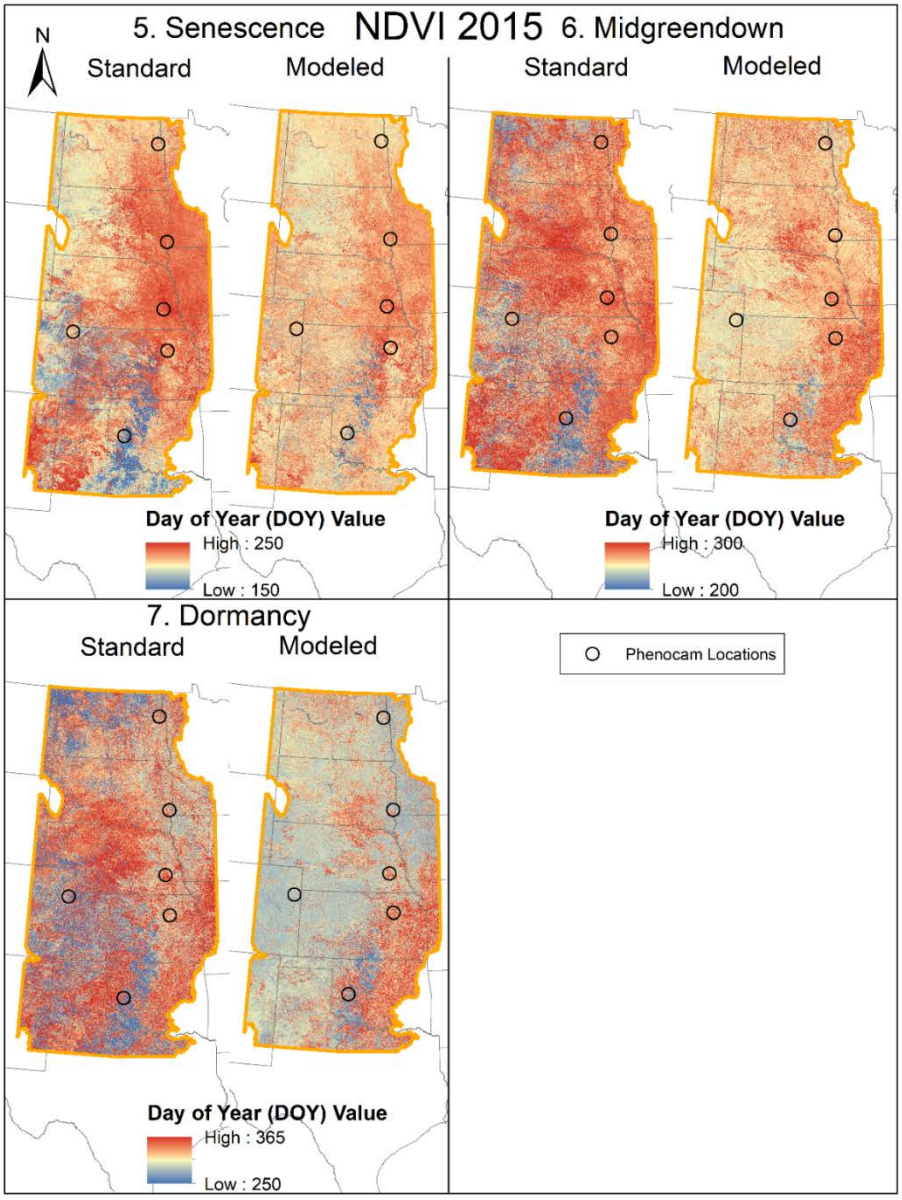


Figure A2: Maps showing the date of the last three NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2015.

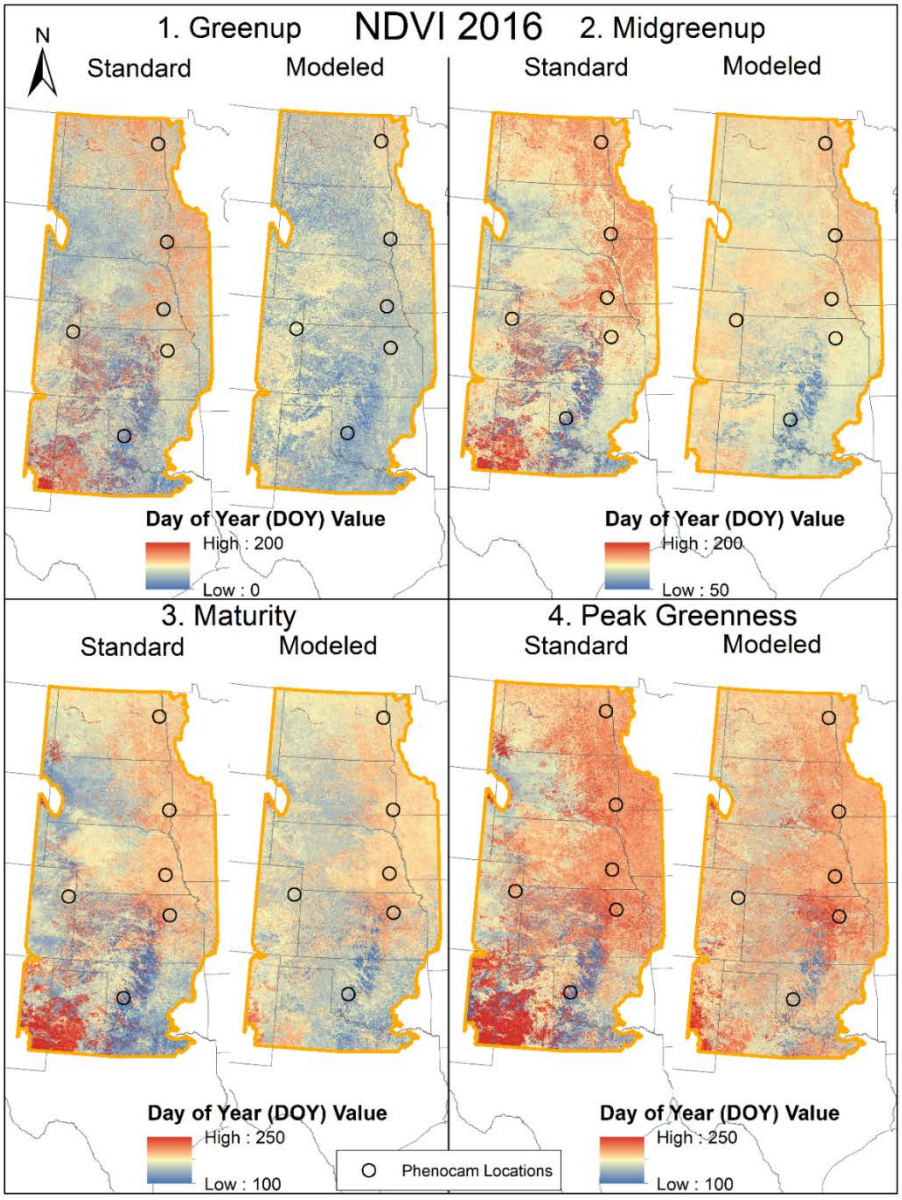


Figure A3: Maps showing the date of the first four NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2016.

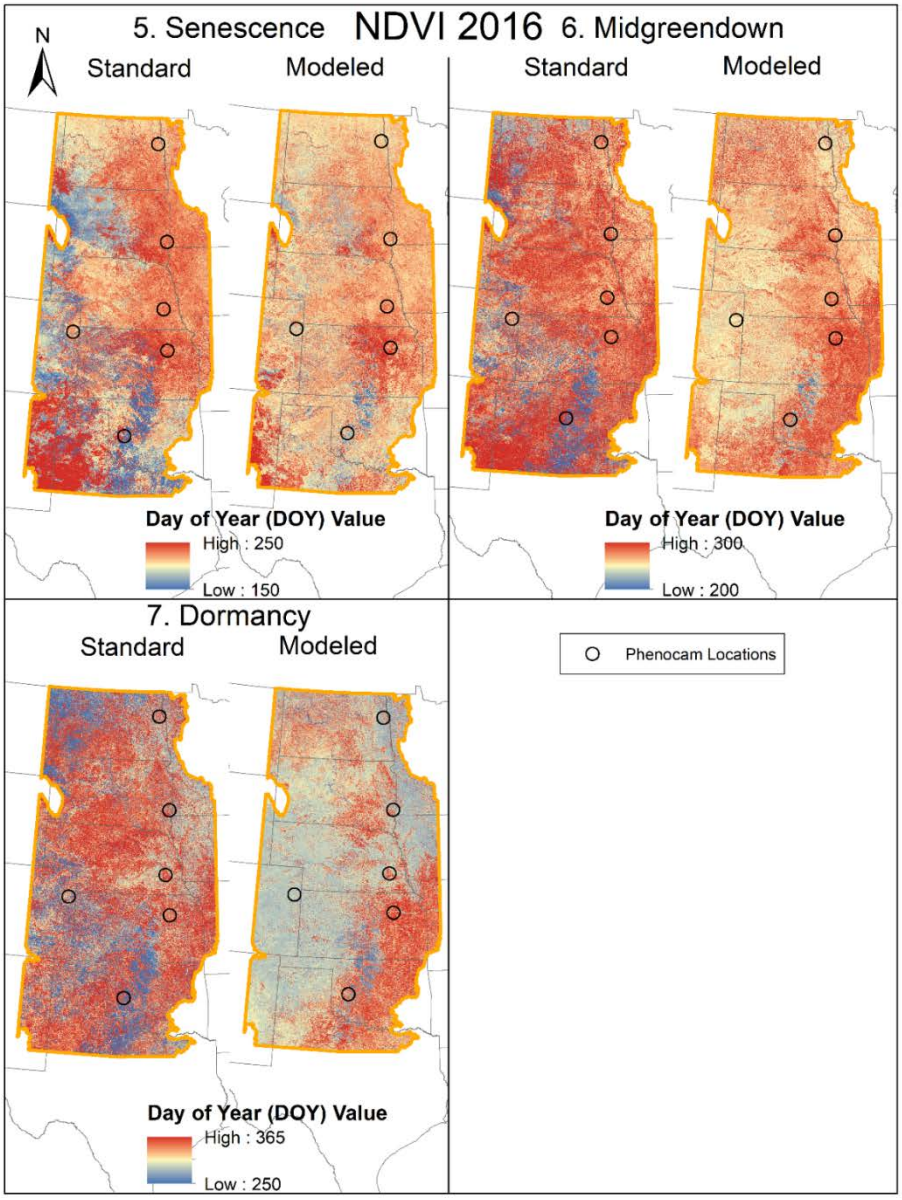


Figure A4: Maps showing the date of the last three NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2016.

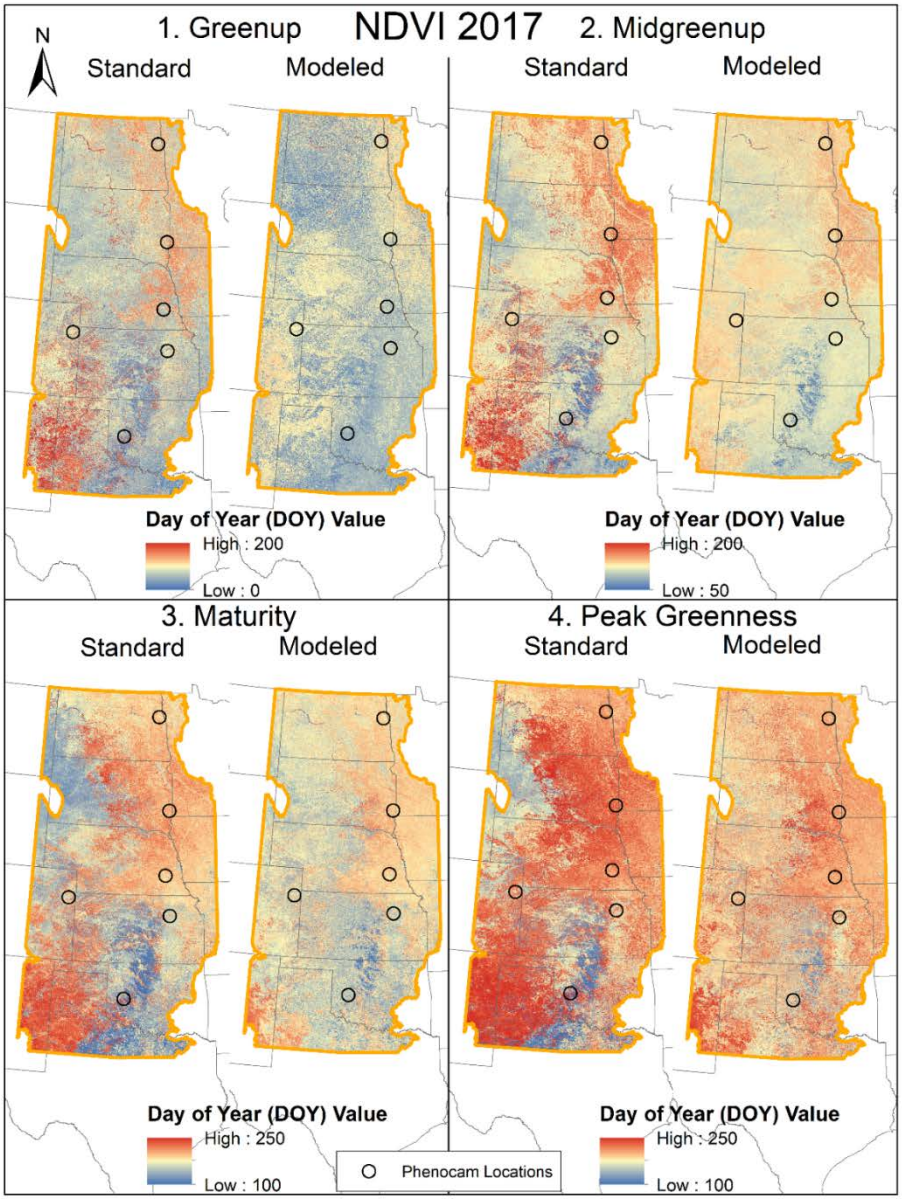


Figure A5: Maps showing the date of the first four NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2017.

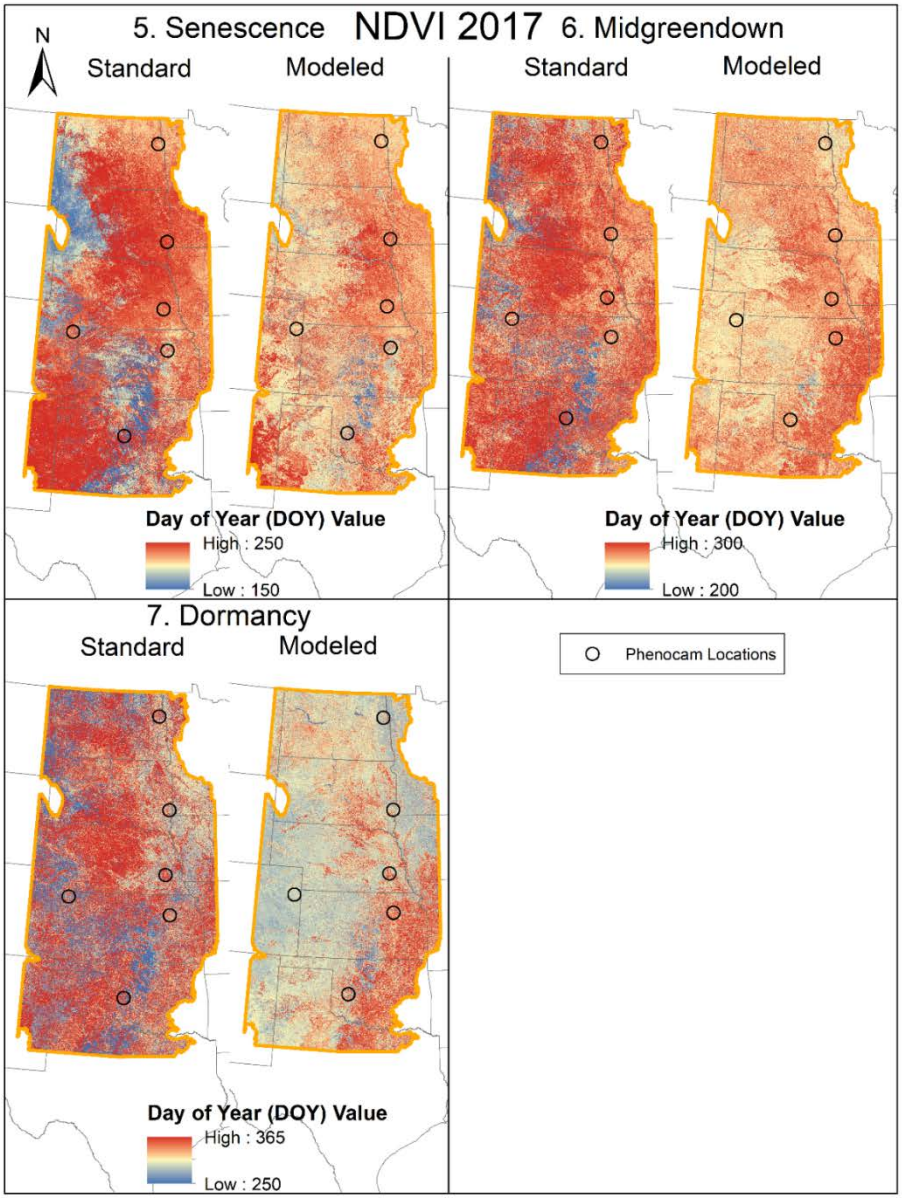


Figure A6: Maps showing the date of the last three NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2017.

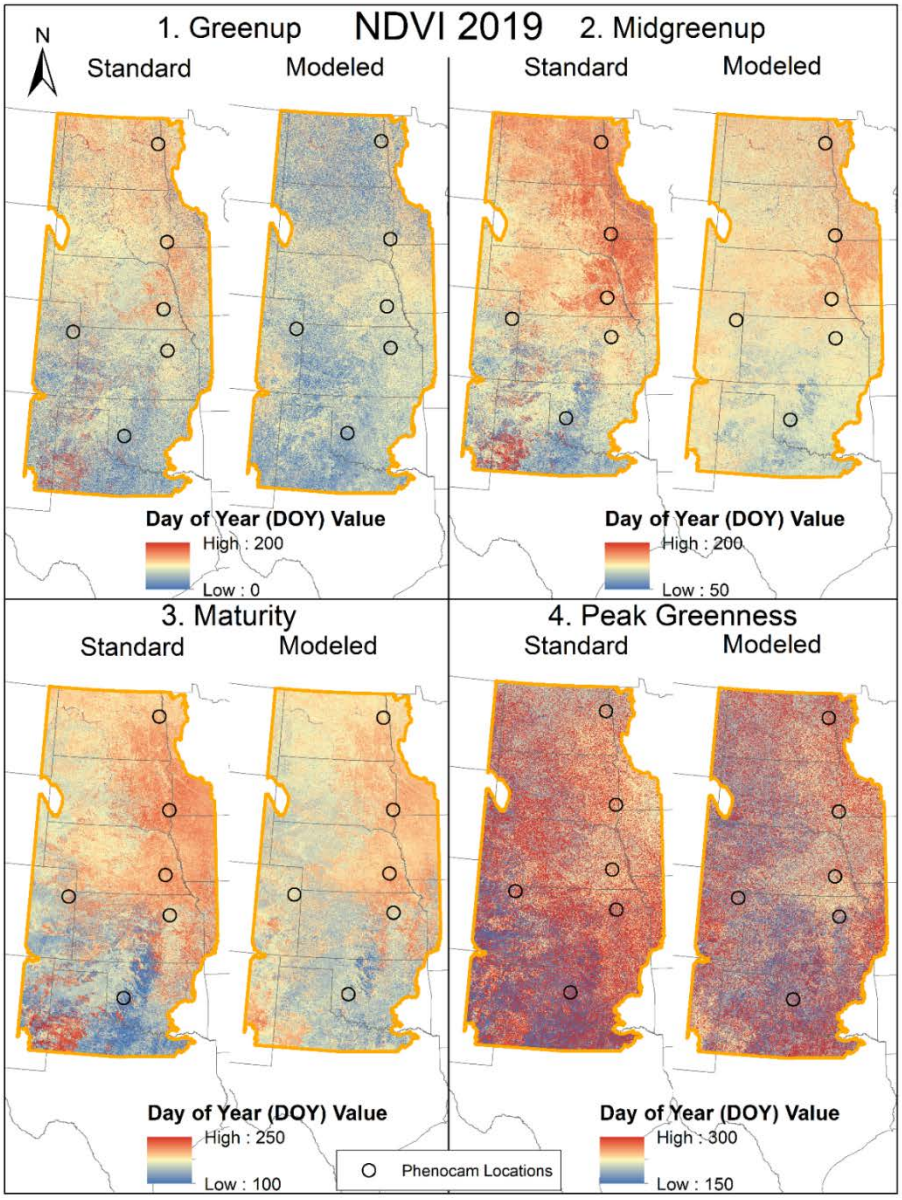


Figure A7: Maps showing the date of the first four NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2019.

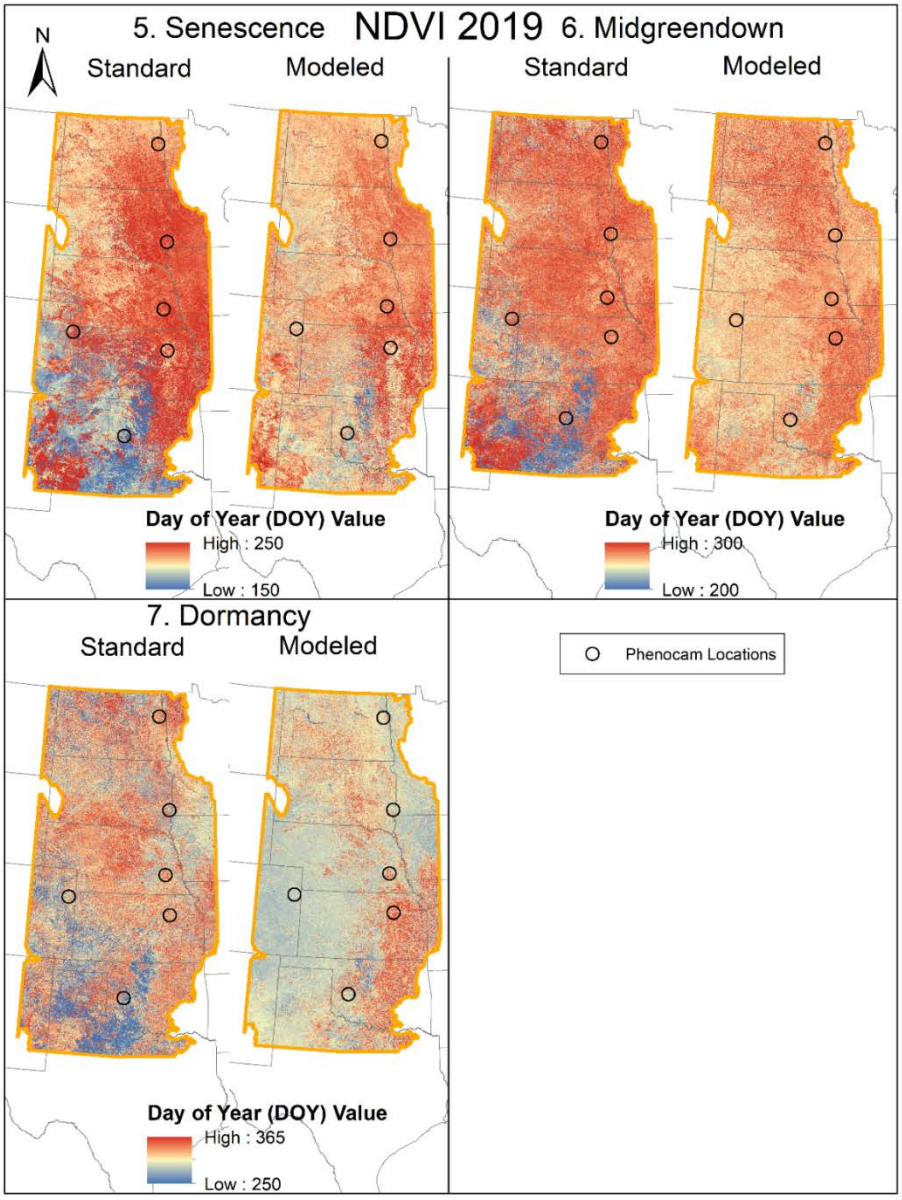


Figure A8: Maps showing the date of the last three NDVI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2019.

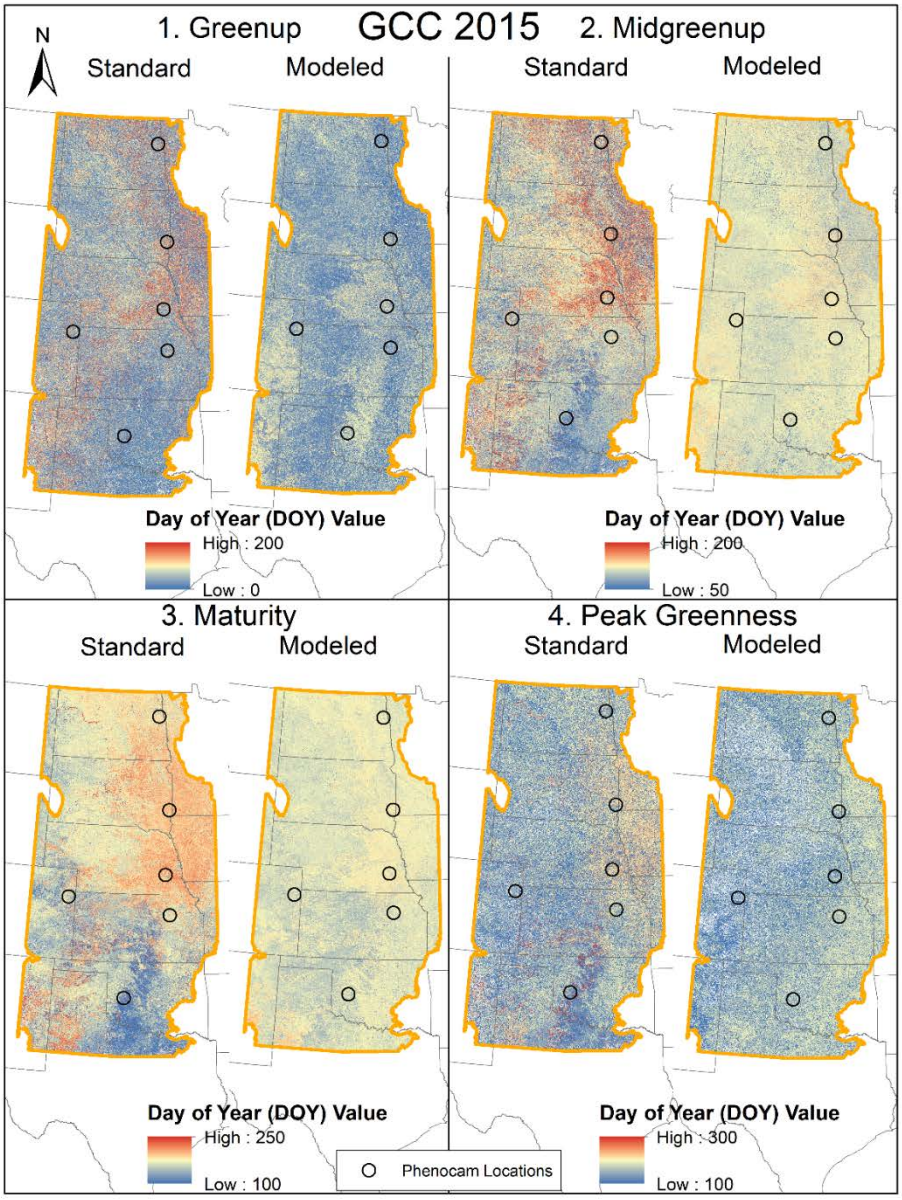


Figure A9: Maps showing the date of the first four GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2015.

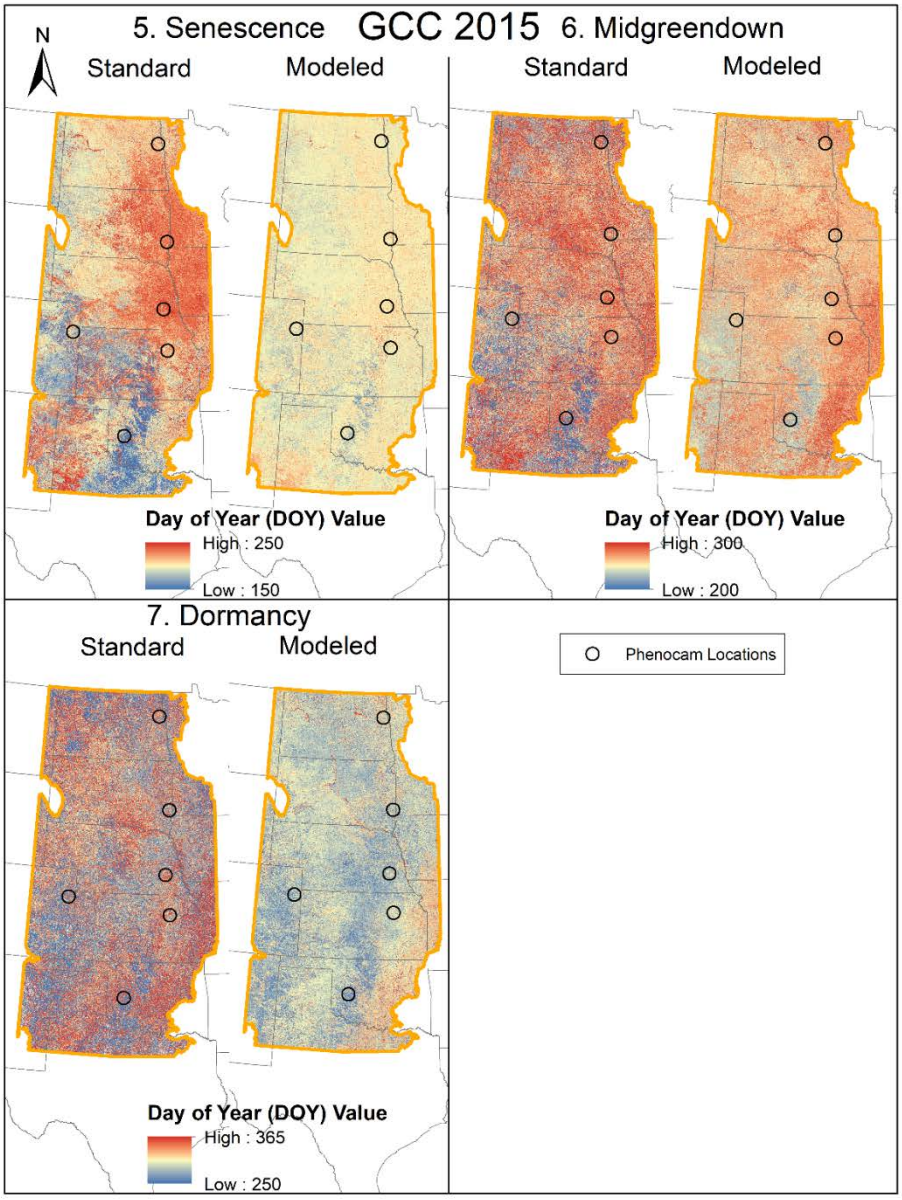


Figure A10: Maps showing the date of the last three GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2015.

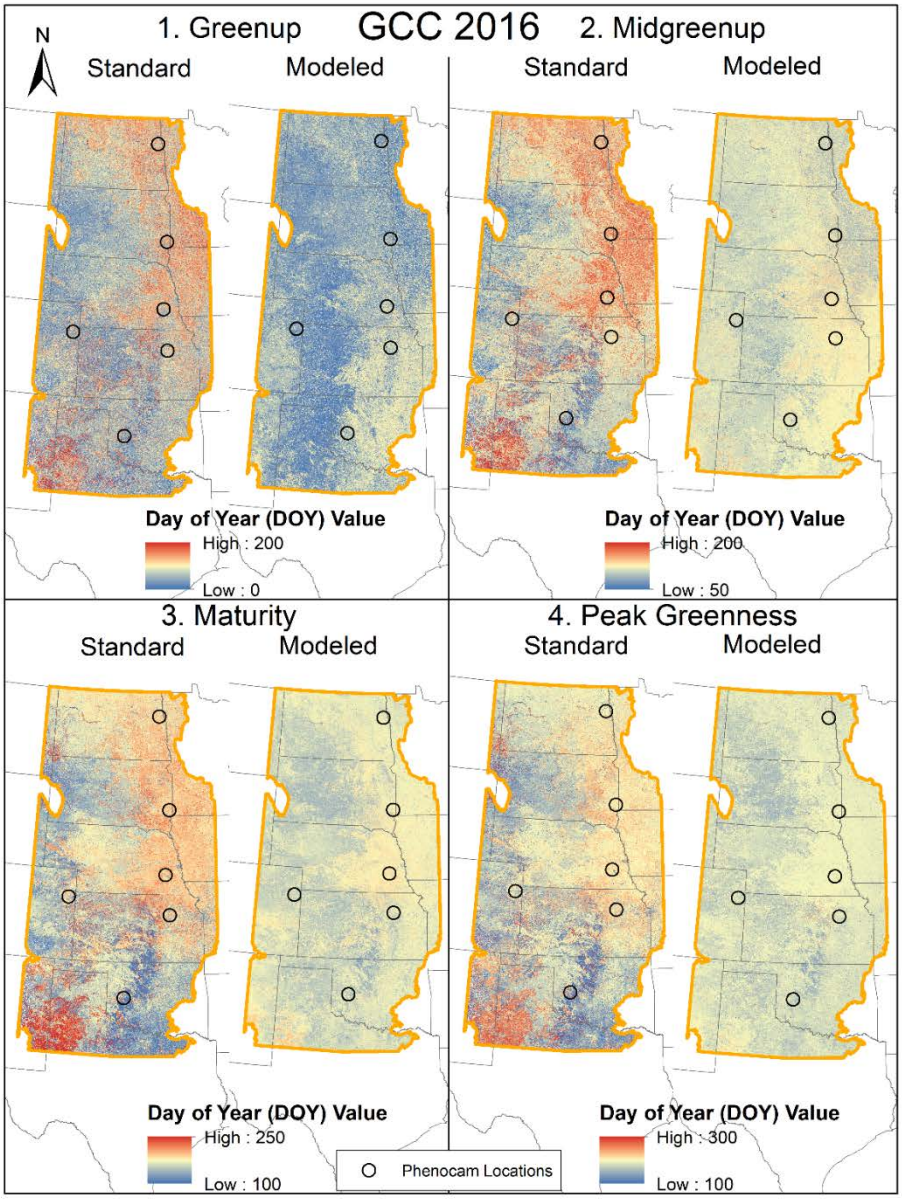


Figure A11: Maps showing the date of the first four GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2016.

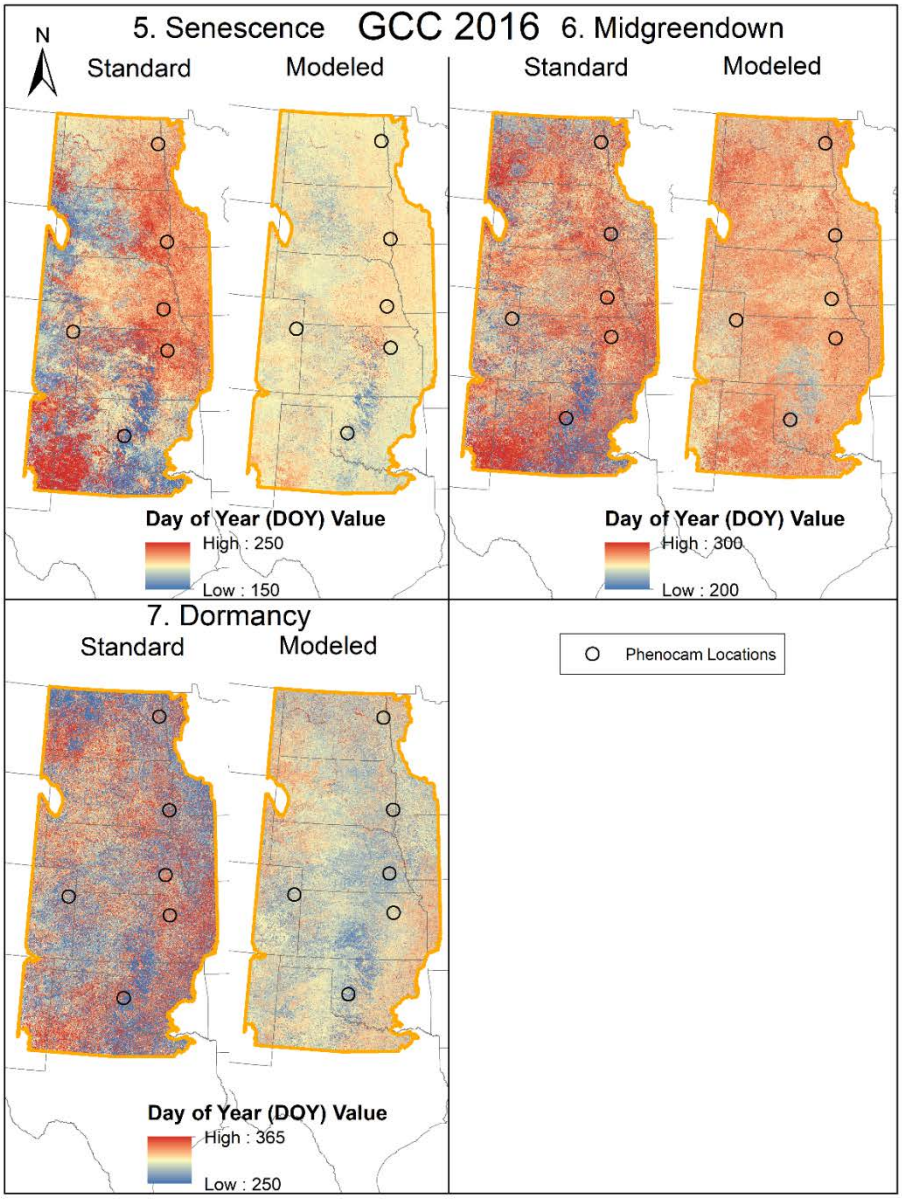


Figure A12: Maps showing the date of the last three GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2016.

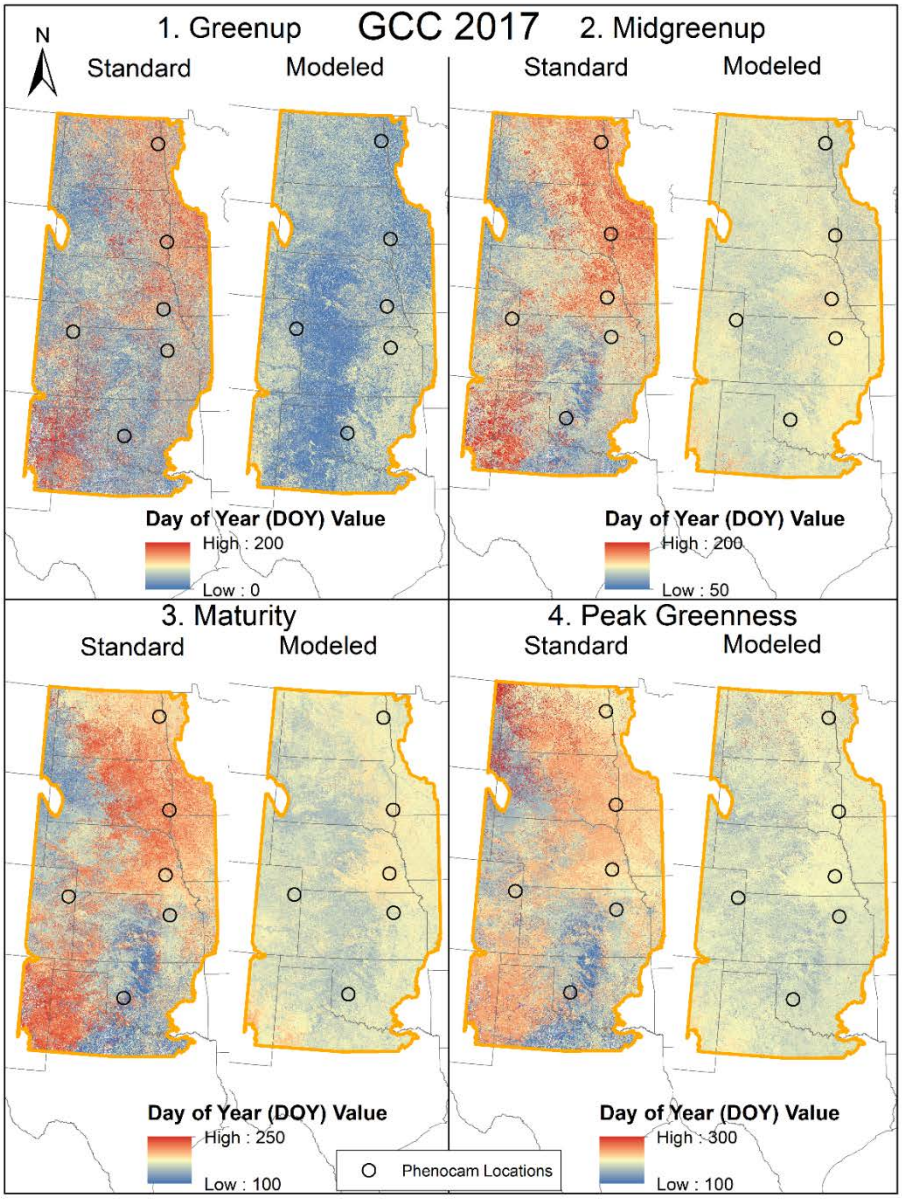


Figure A13: Maps showing the date of the first four GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2017.

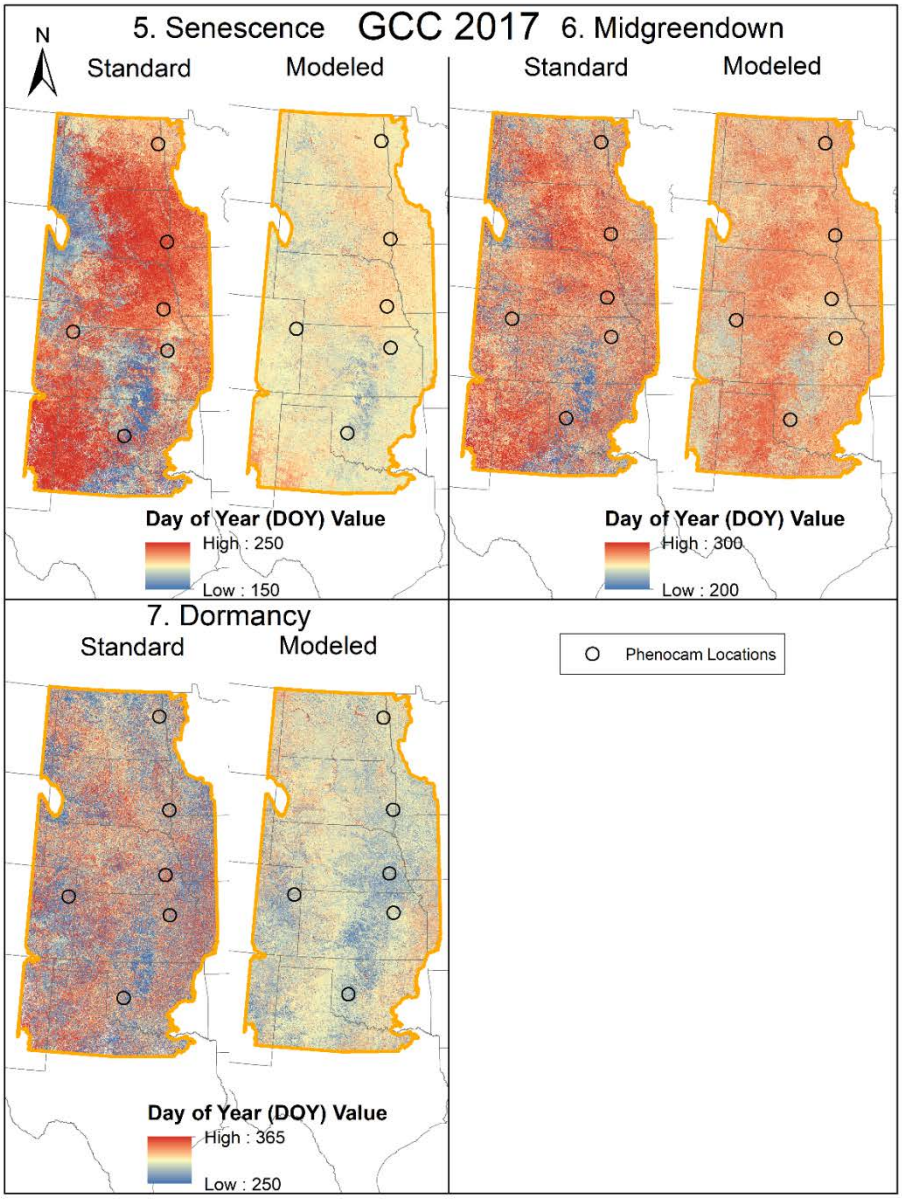


Figure A14: Maps showing the date of the last three GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2017.

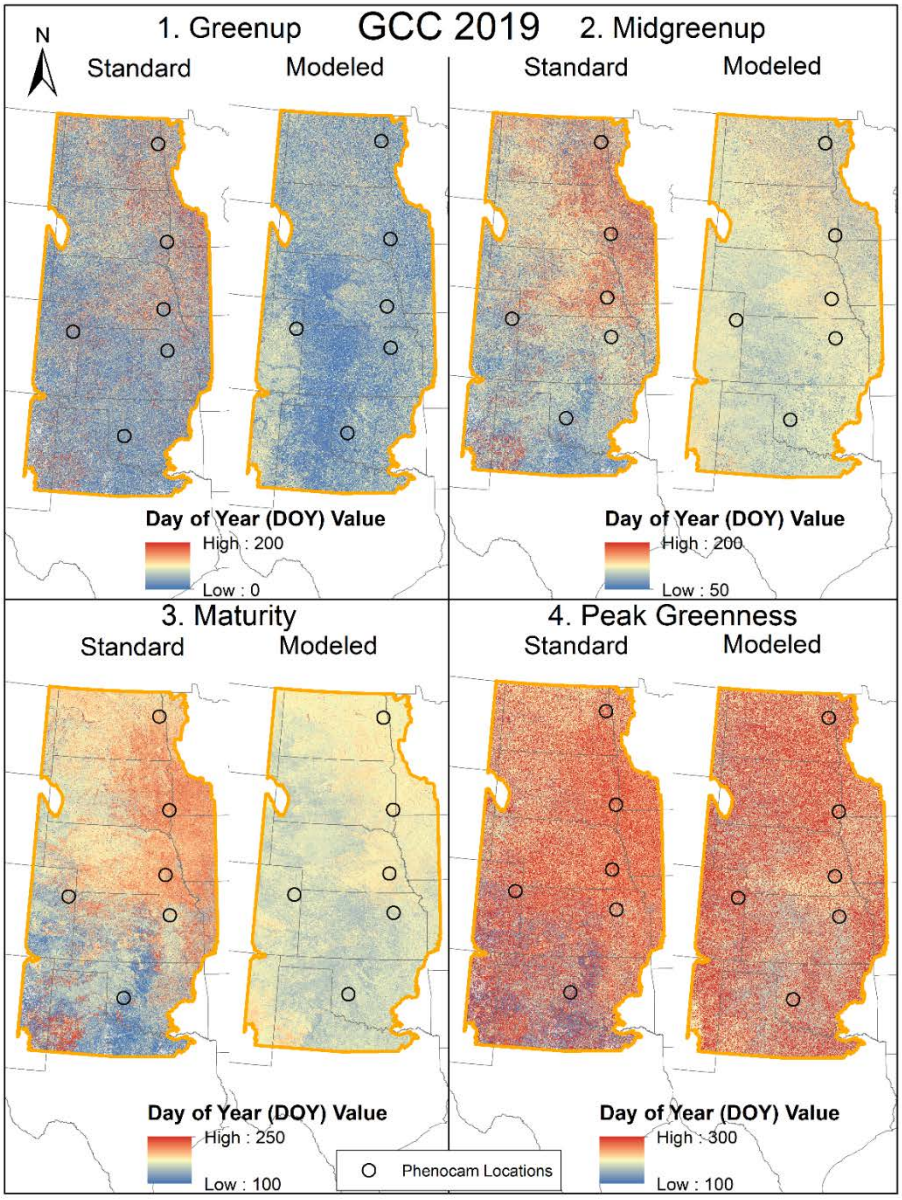


Figure A15: Maps showing the date of the first four GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2019.

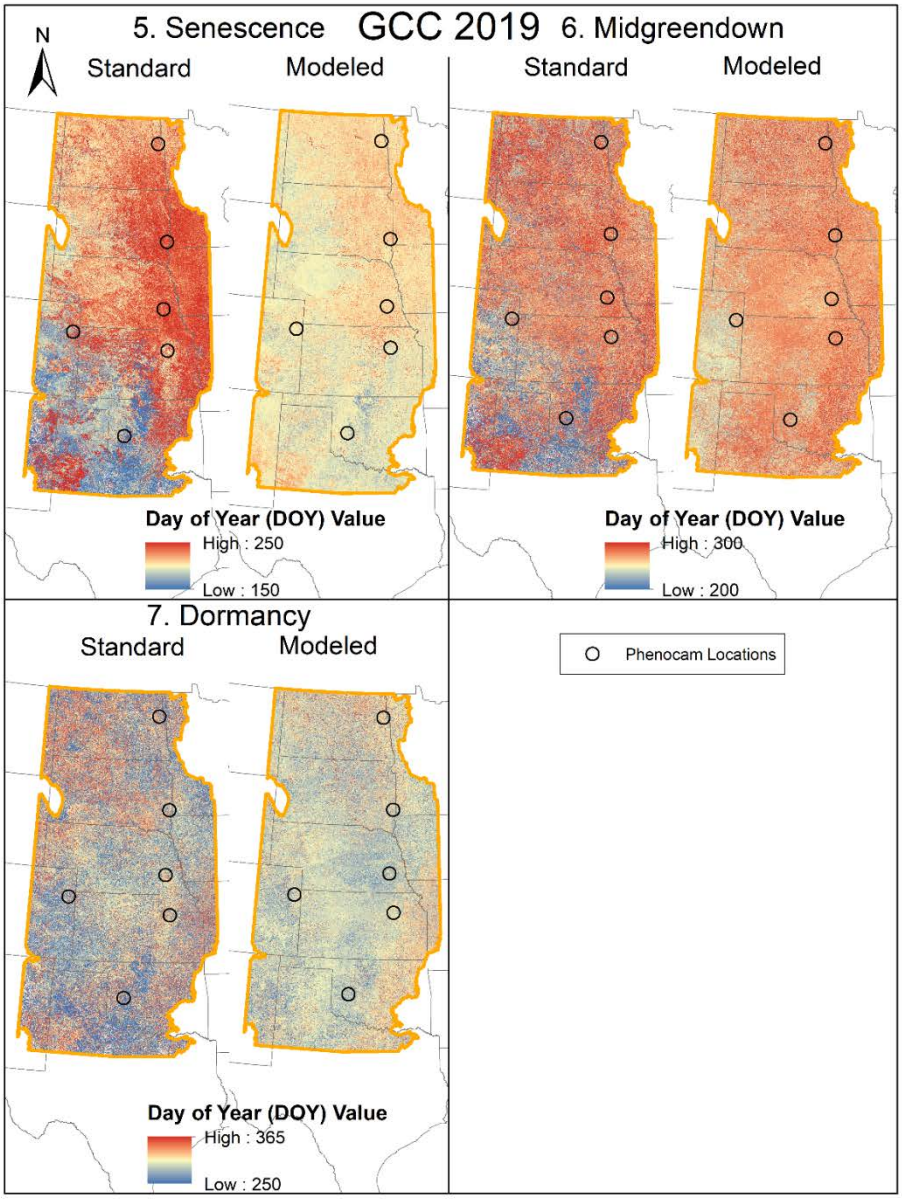


Figure A16: Maps showing the date of the last three GCC phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2019.

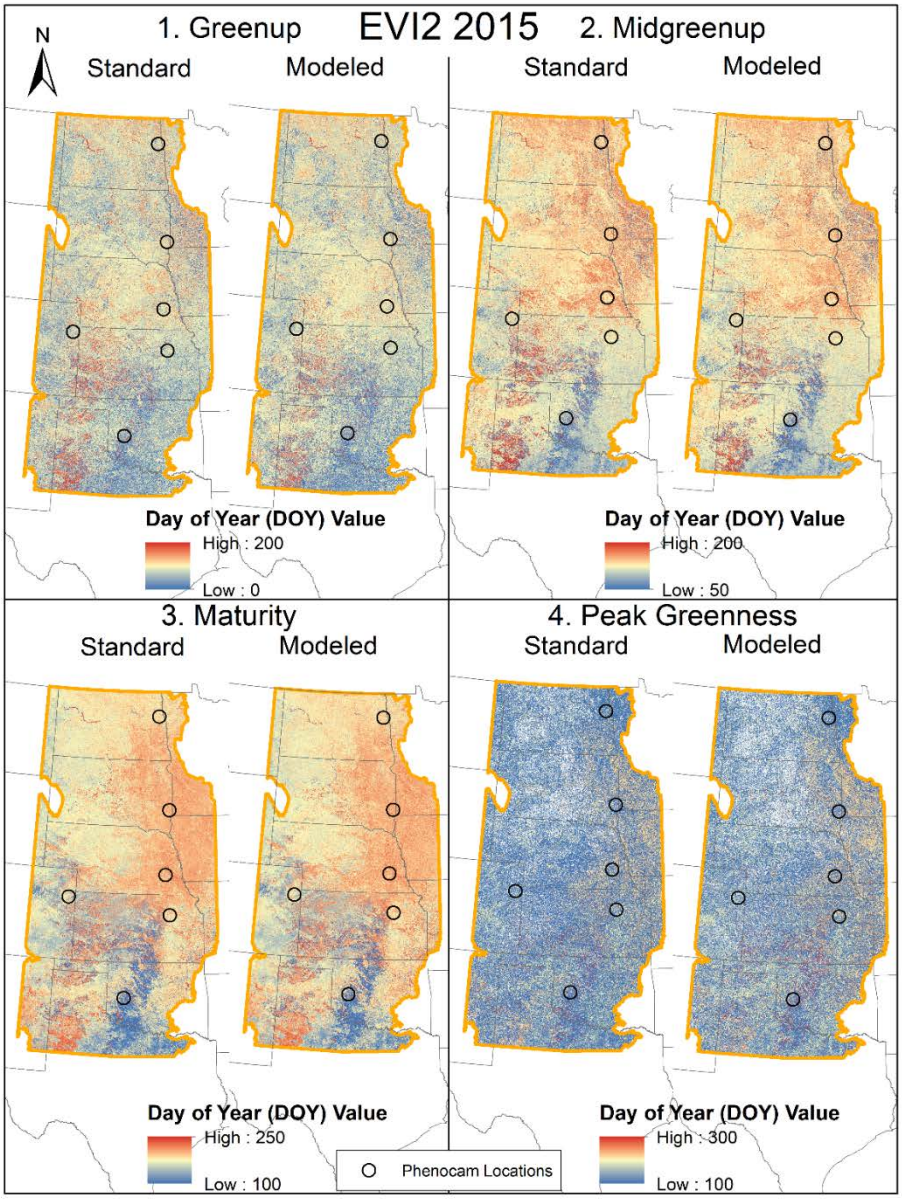


Figure A17: Maps showing the date of the first four EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2015.

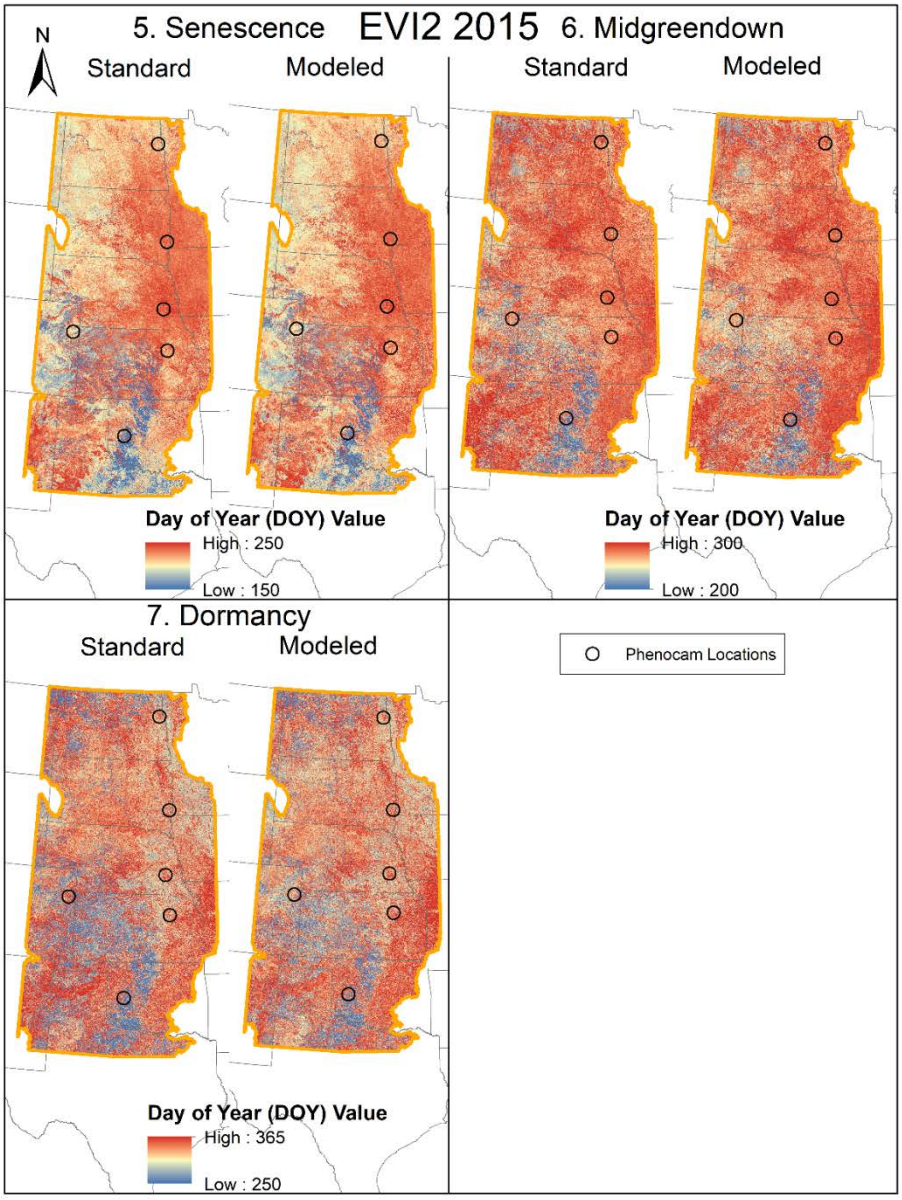


Figure A18: Maps showing the date of the last three EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2015.

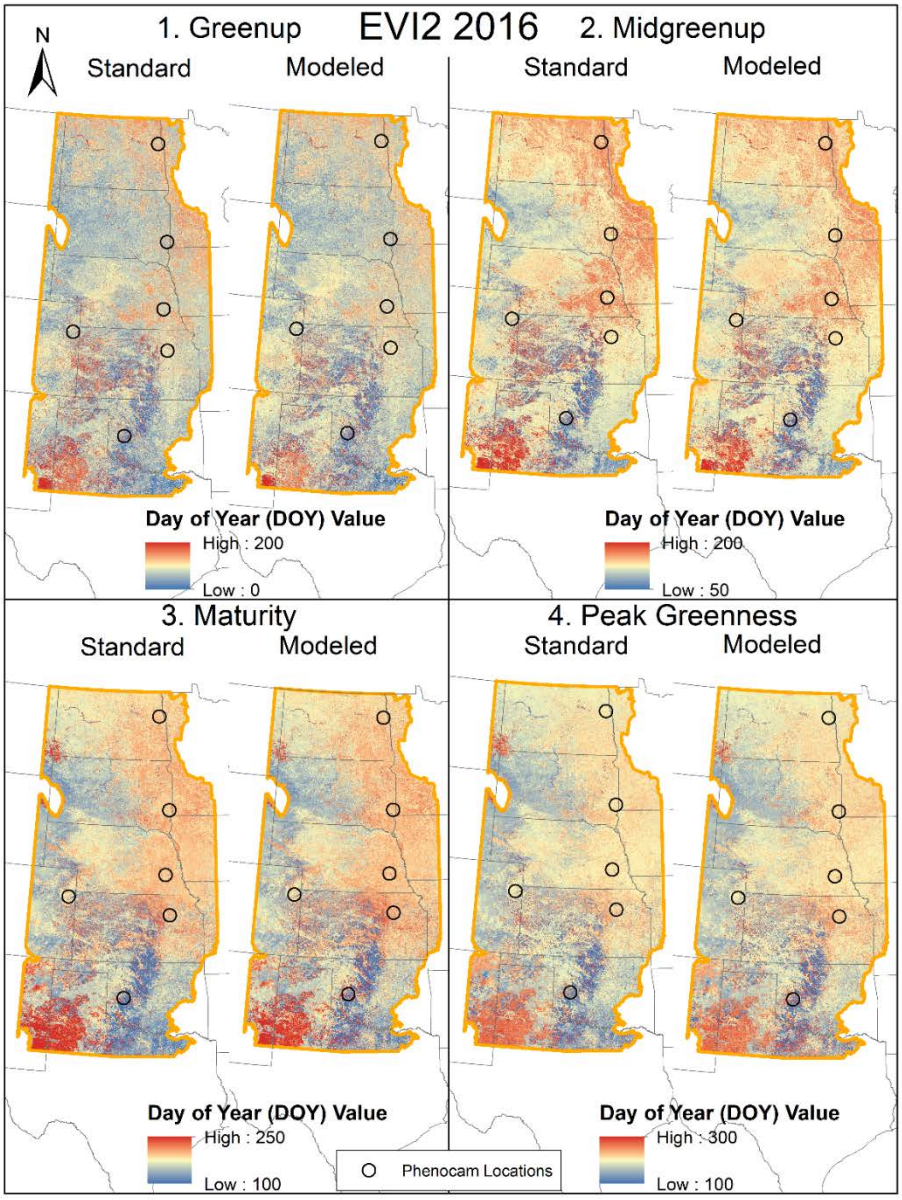


Figure A19: Maps showing the date of the first four EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2016.

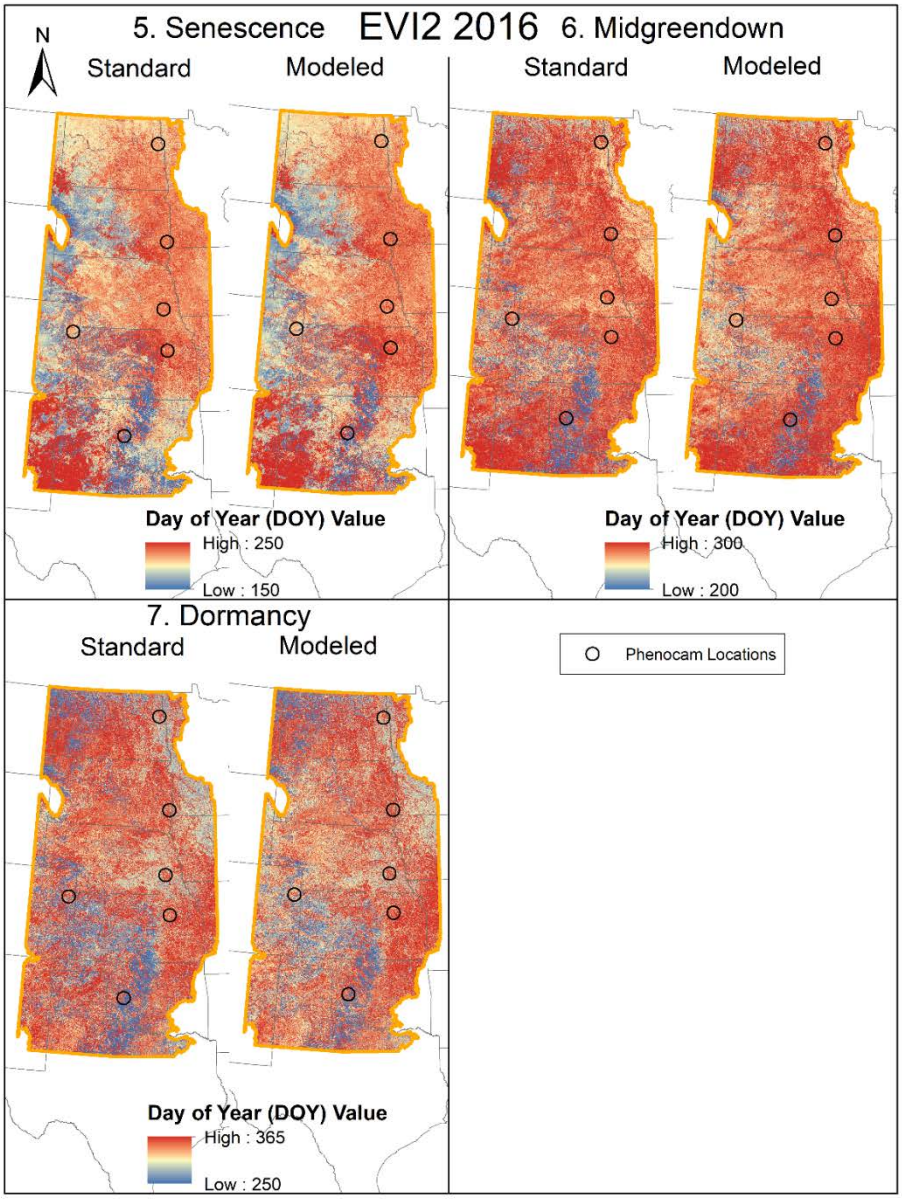


Figure A20: Maps showing the date of the last three EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2016.

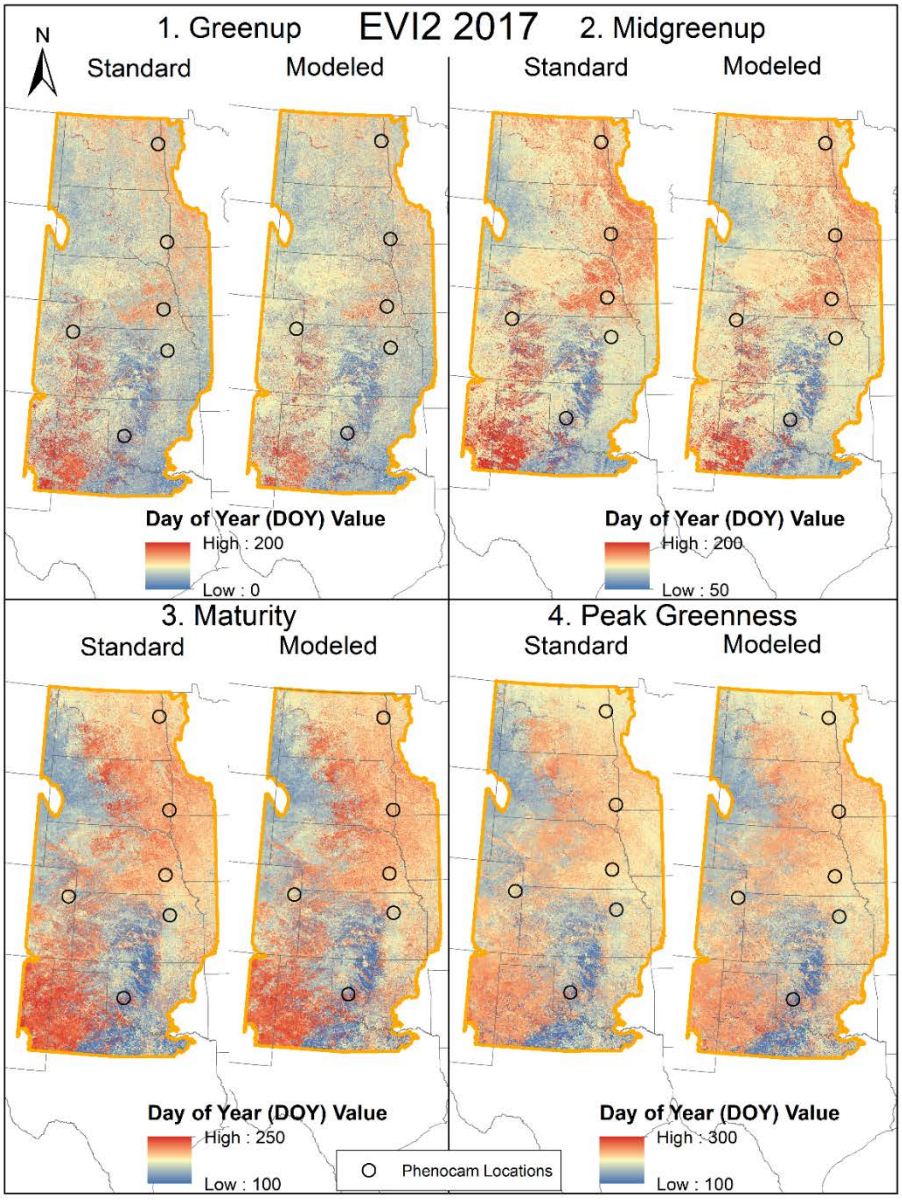


Figure A21: Maps showing the date of the first four EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2017.

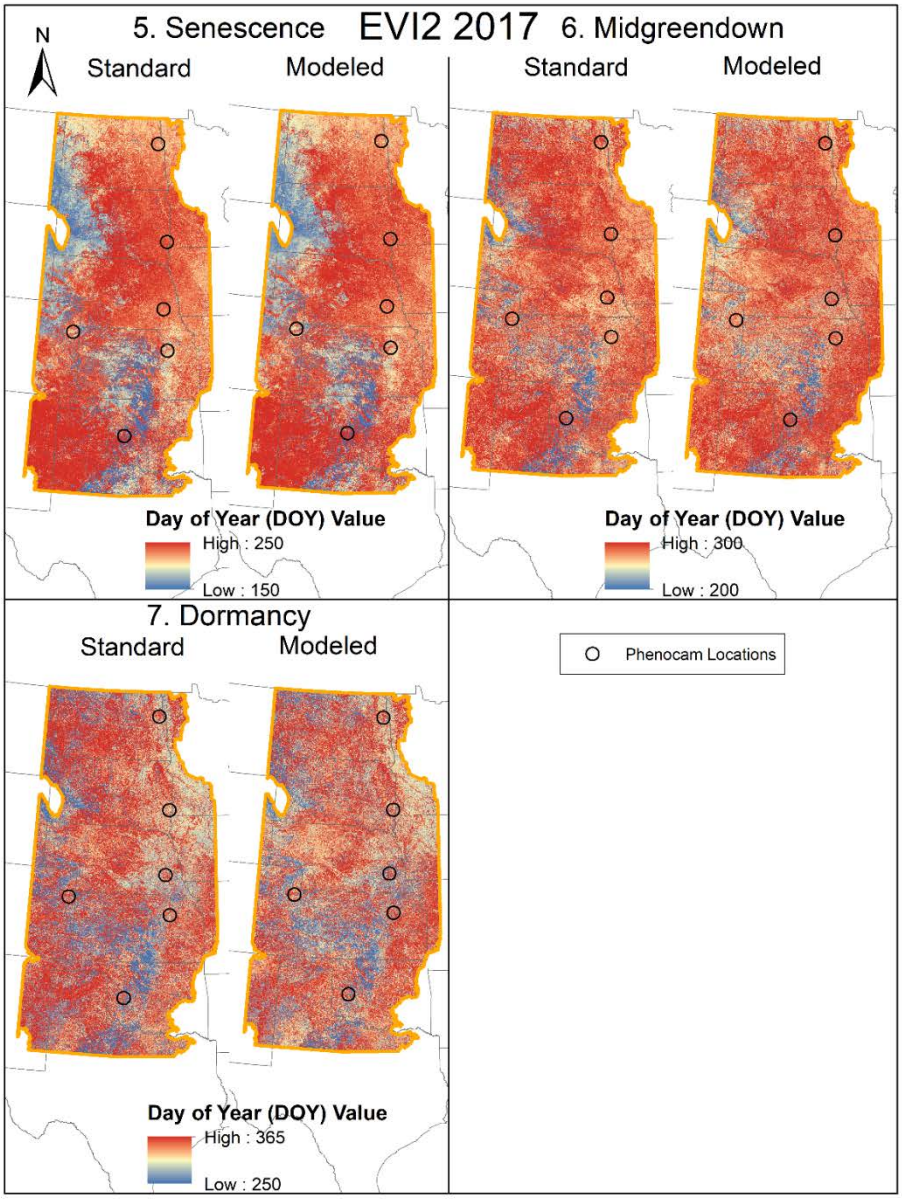


Figure A22: Maps showing the date of the last three EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2017.

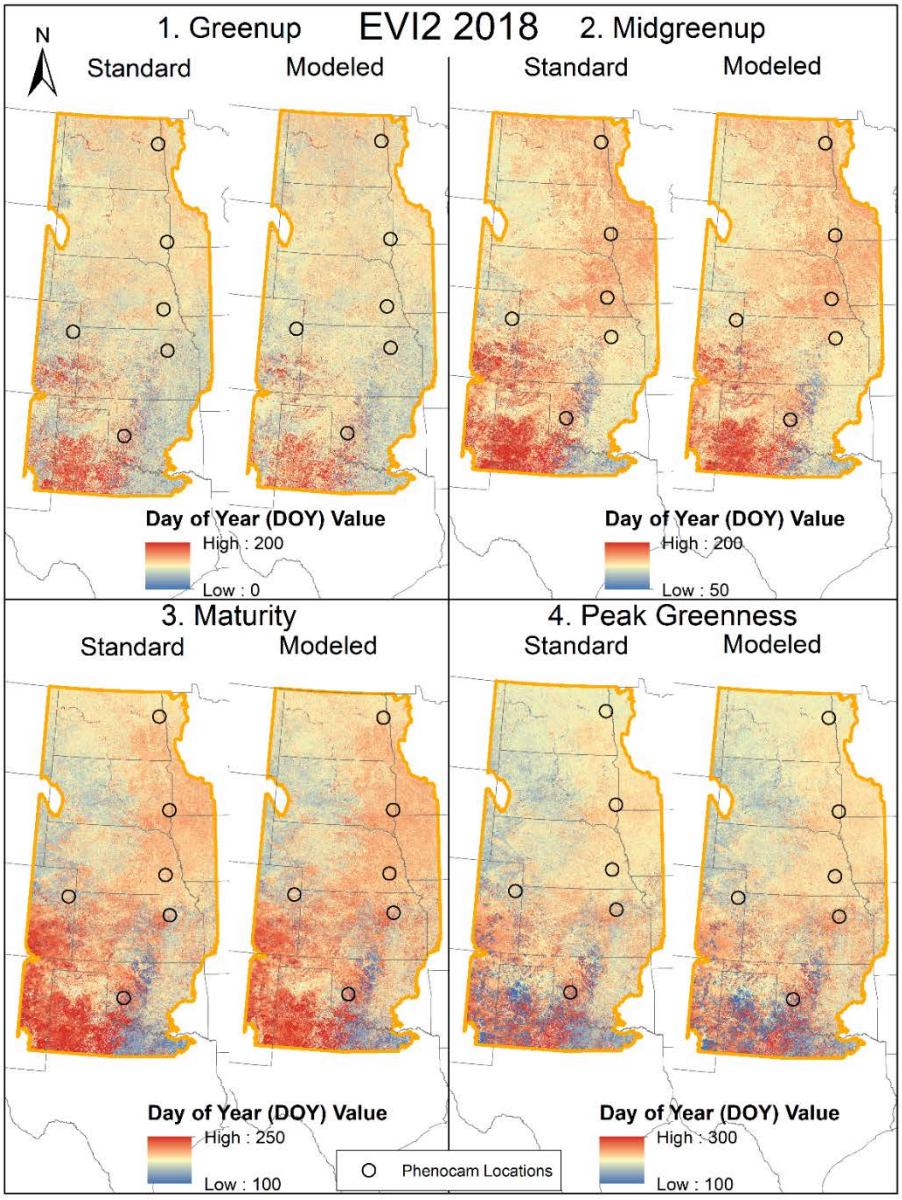


Figure A23: Maps showing the date of the first four EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2018.

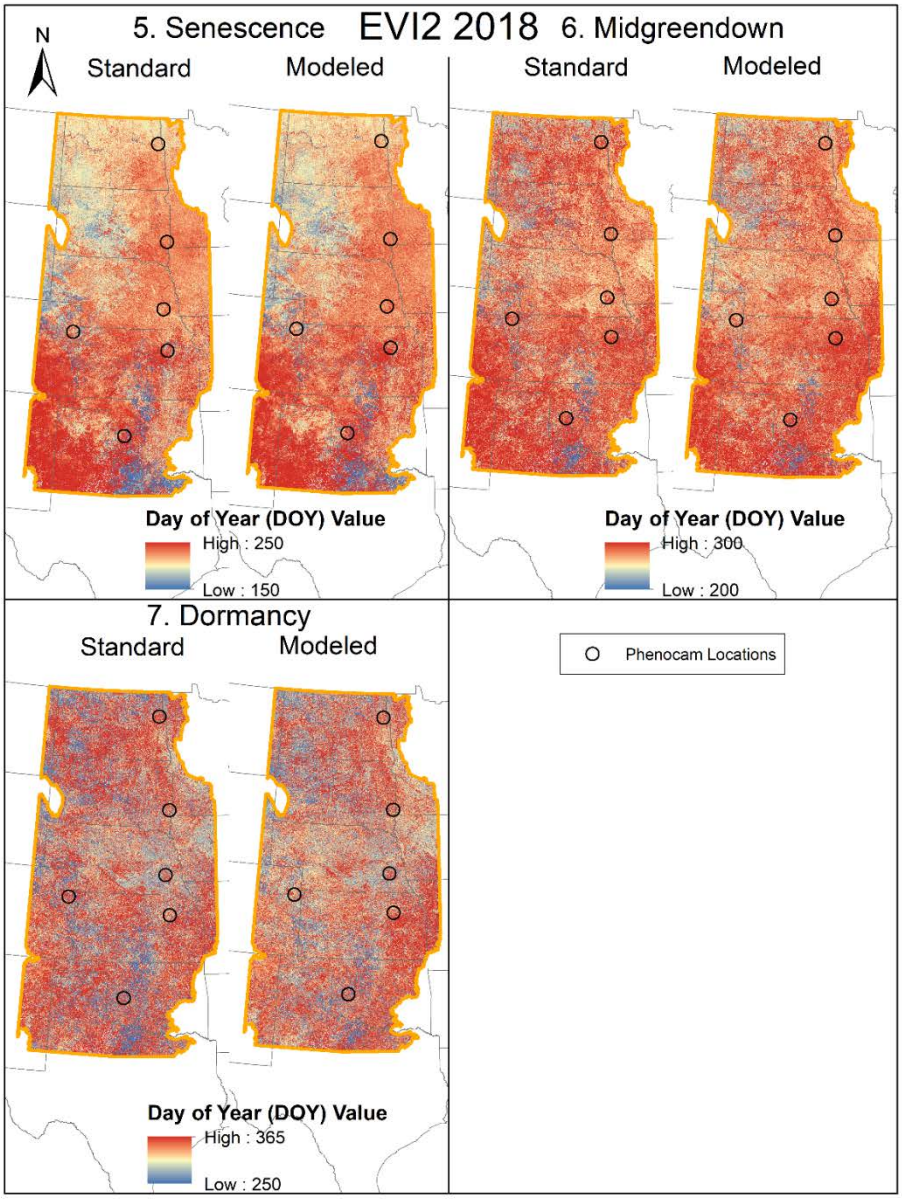


Figure A24: Maps showing the date of the last three EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2018.

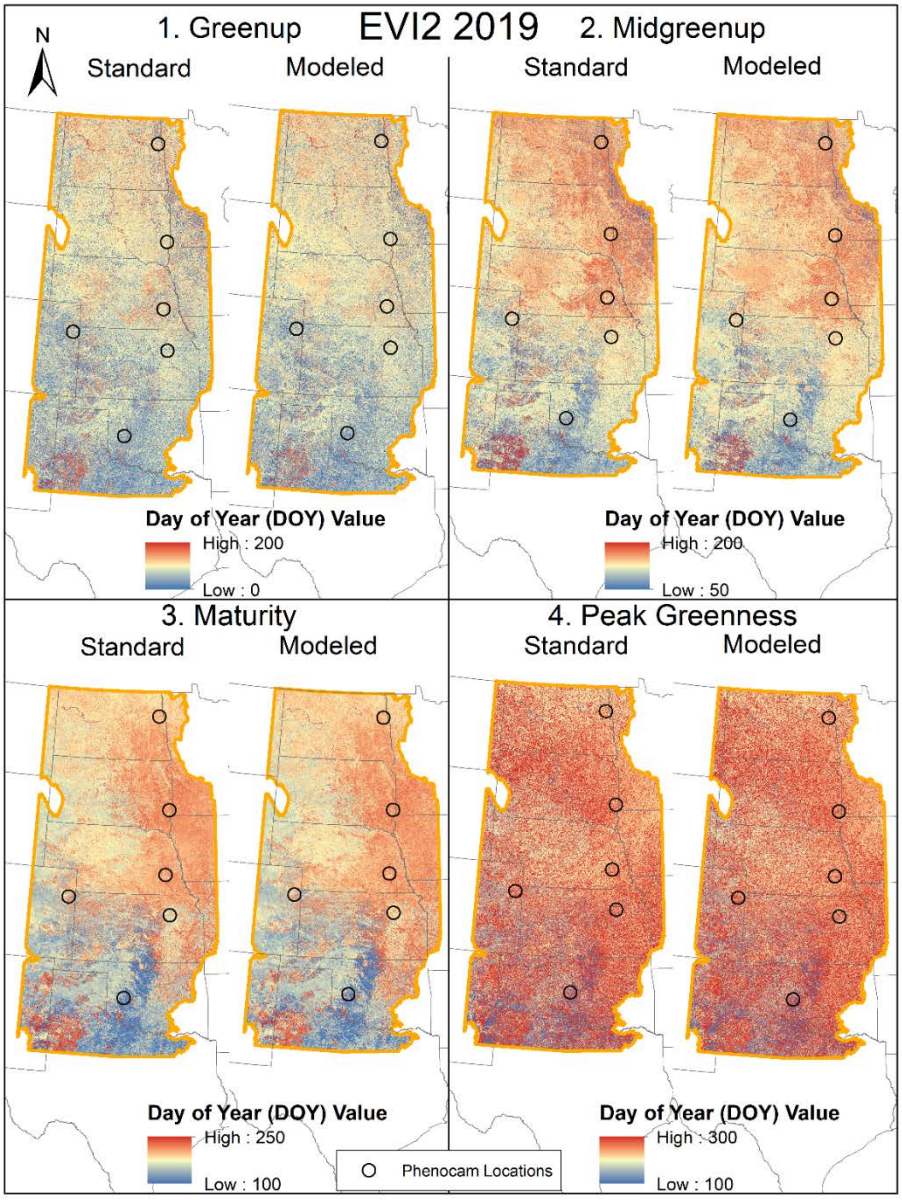


Figure A25: Maps showing the date of the first four EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2019.

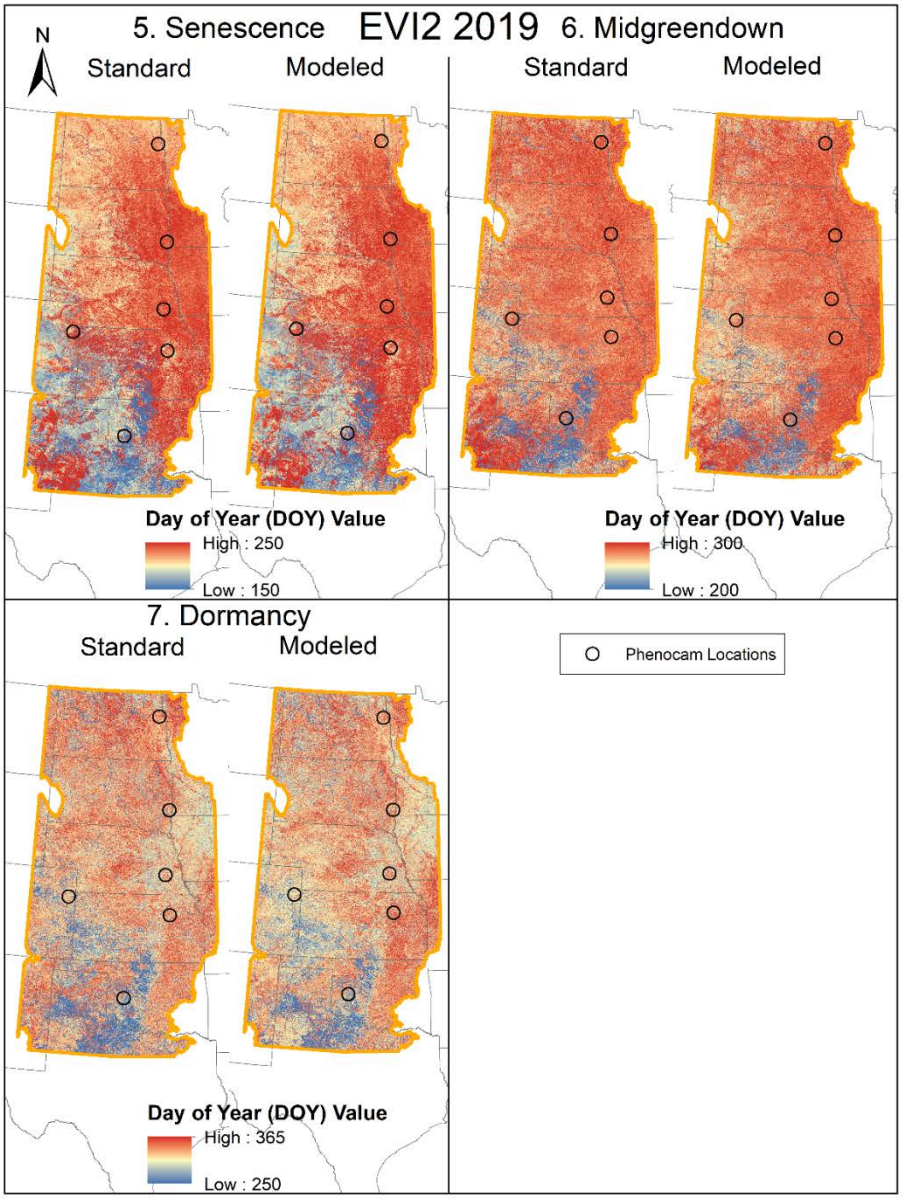


Figure A26: Maps showing the date of the last three EVI2 phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2019.

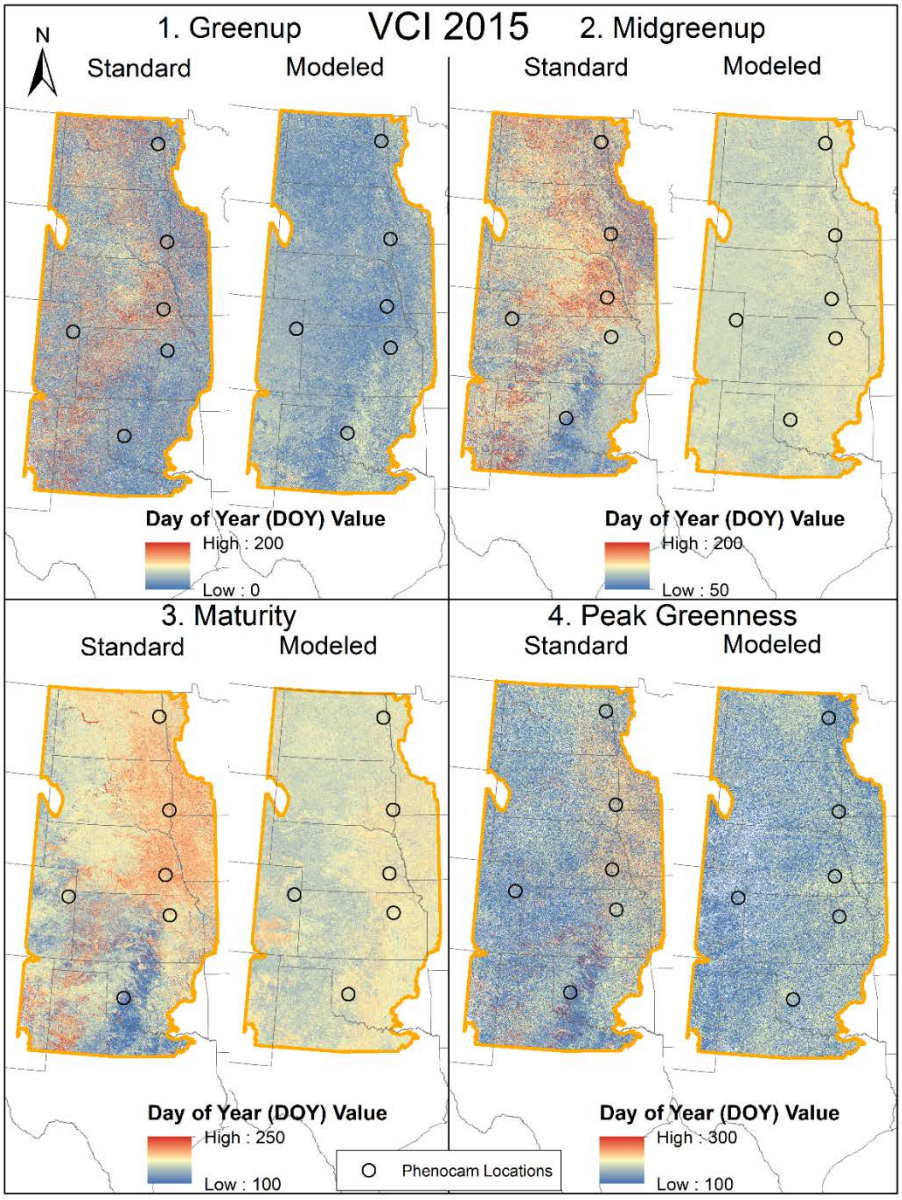


Figure A27: Maps showing the date of the first four VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2015.

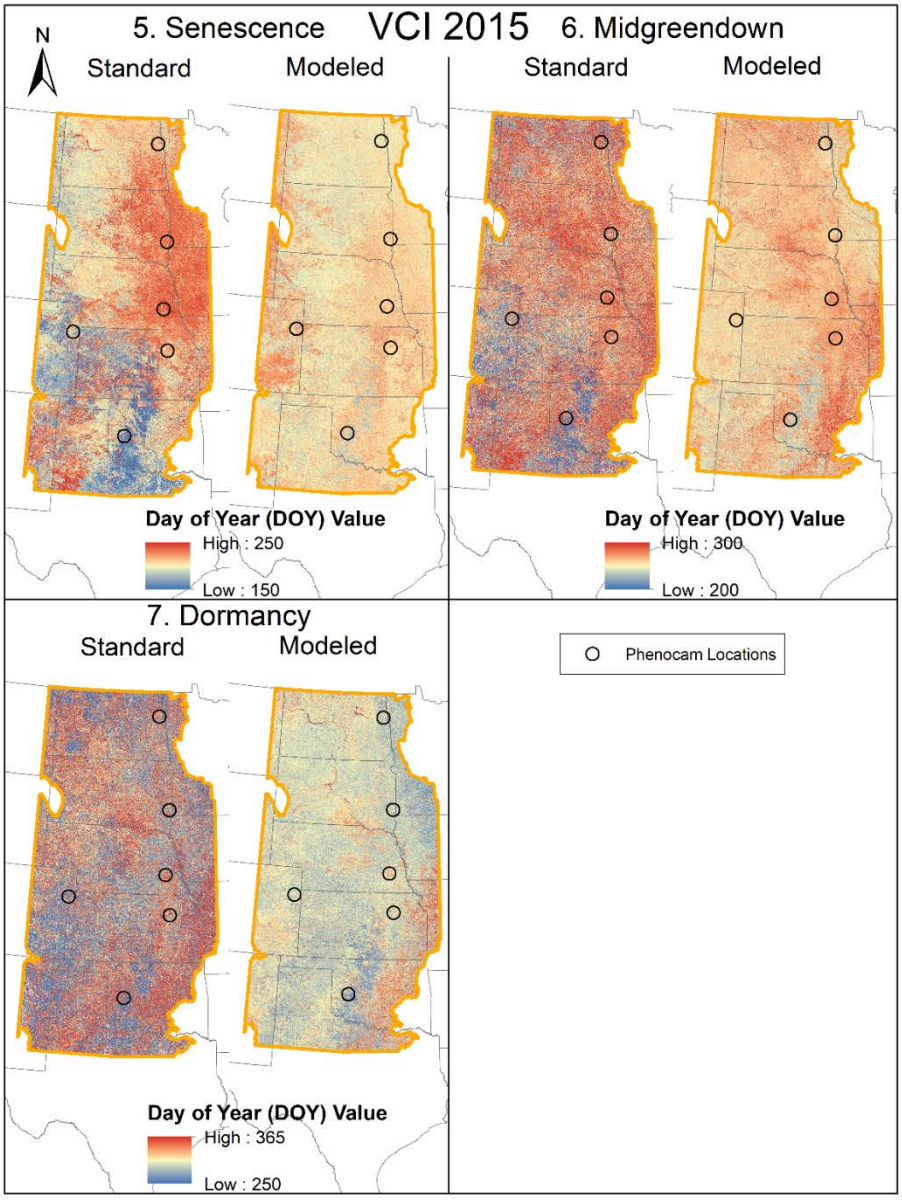


Figure A28: Maps showing the date of the last three VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2015.

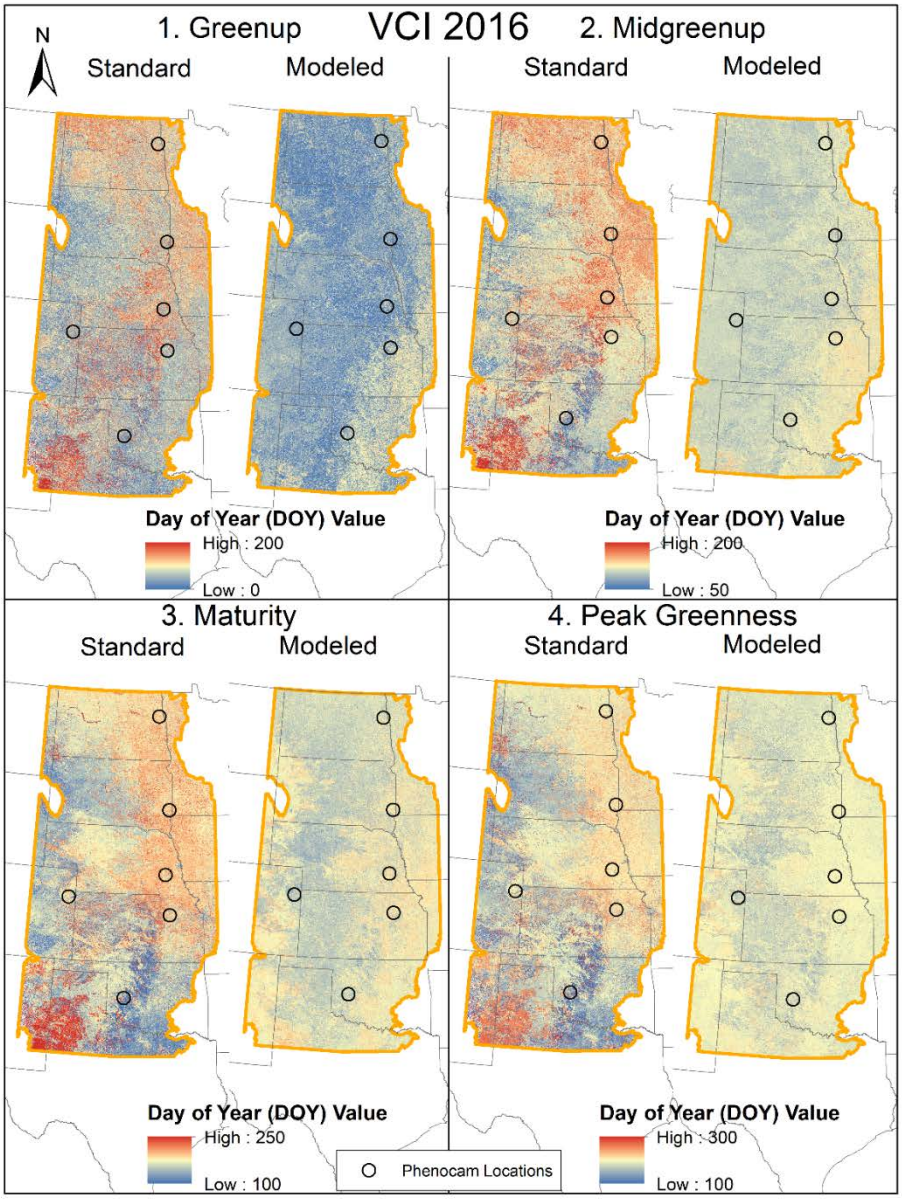


Figure A29: Maps showing the date of the first four VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2016.

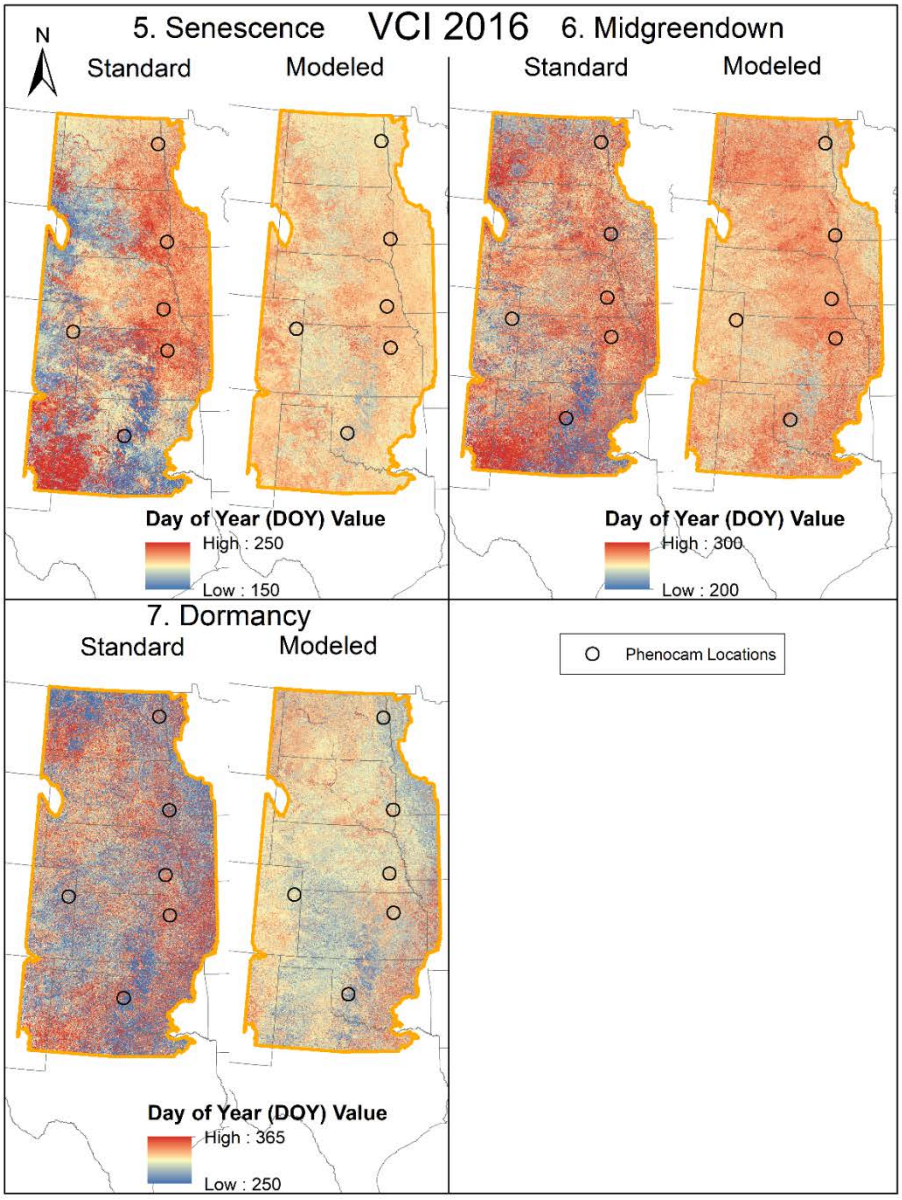


Figure A30: Maps showing the date of the last three VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2016.

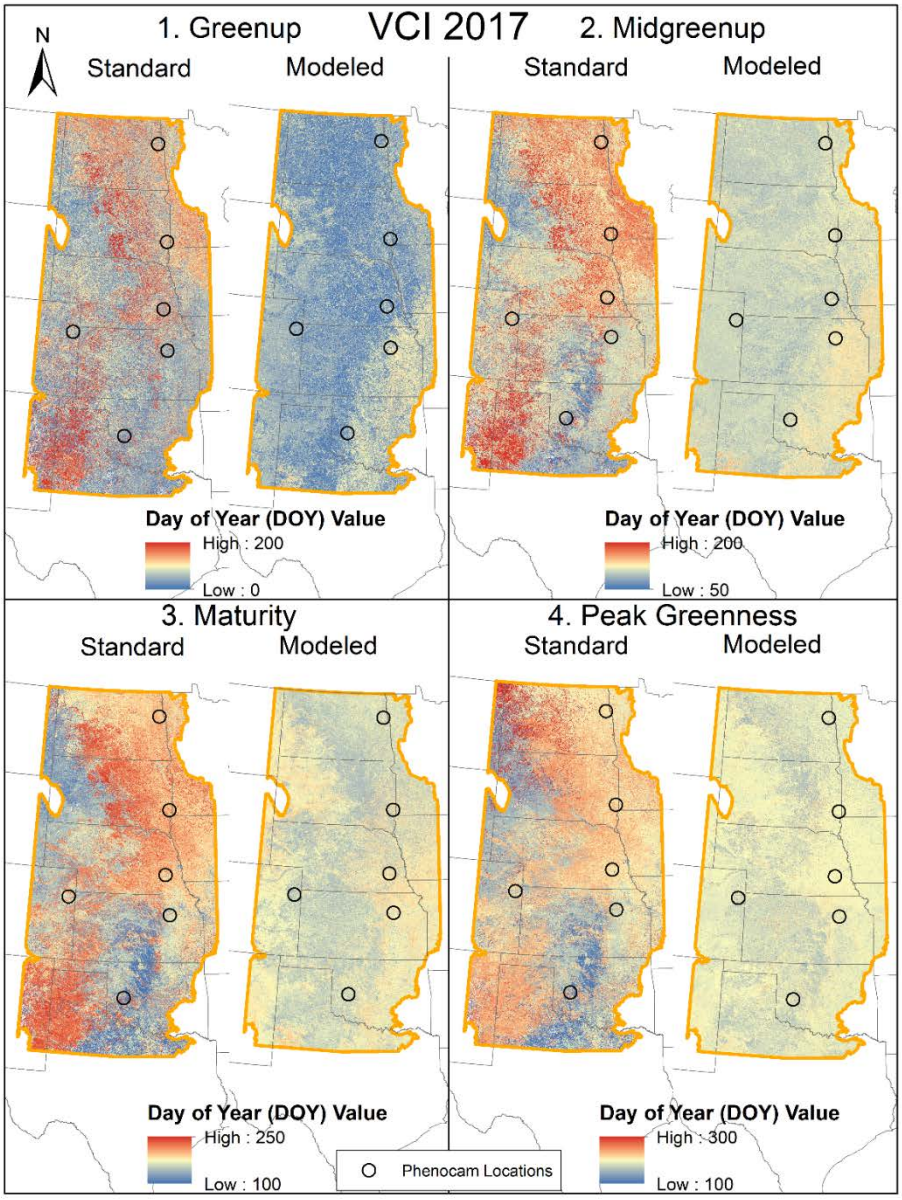


Figure A31: Maps showing the date of the first four VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2017.

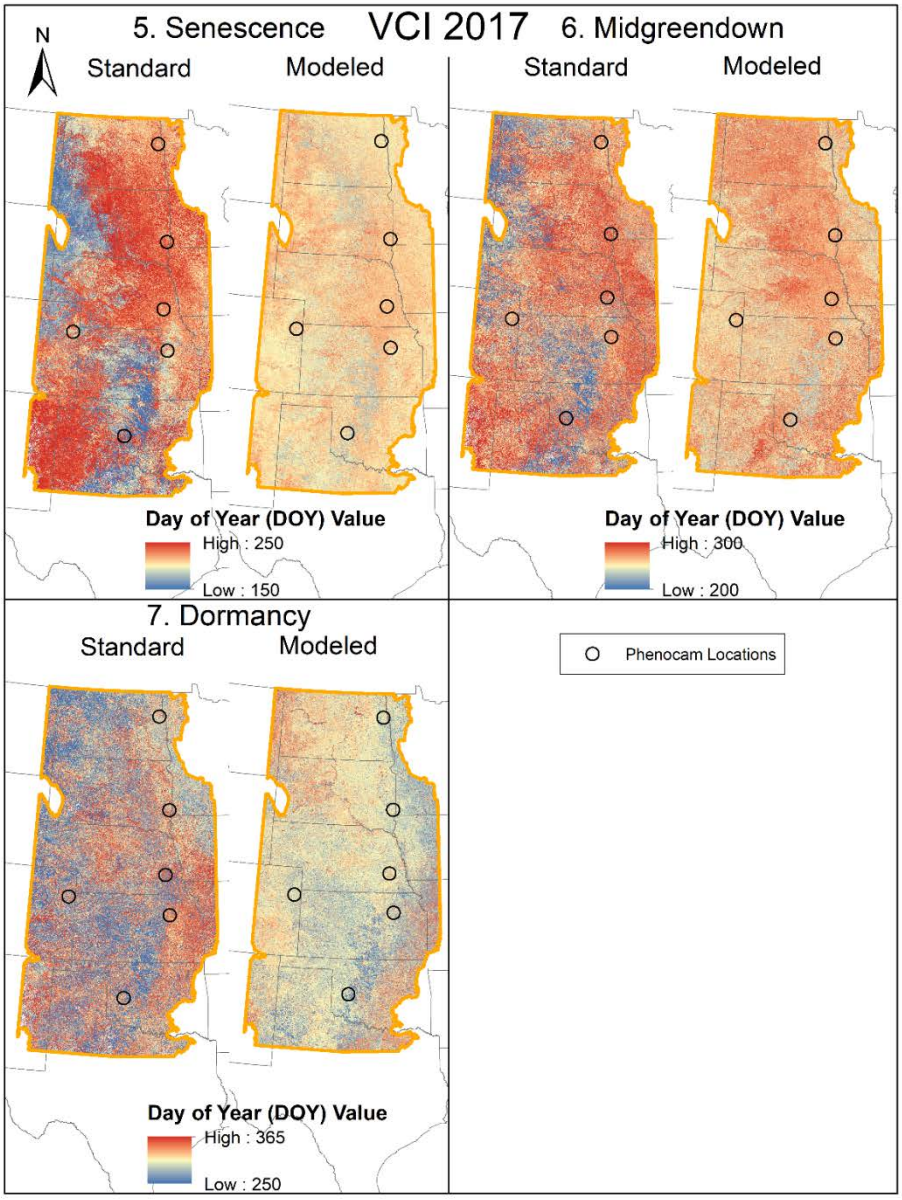


Figure A32: Maps showing the date of the last three VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2017.

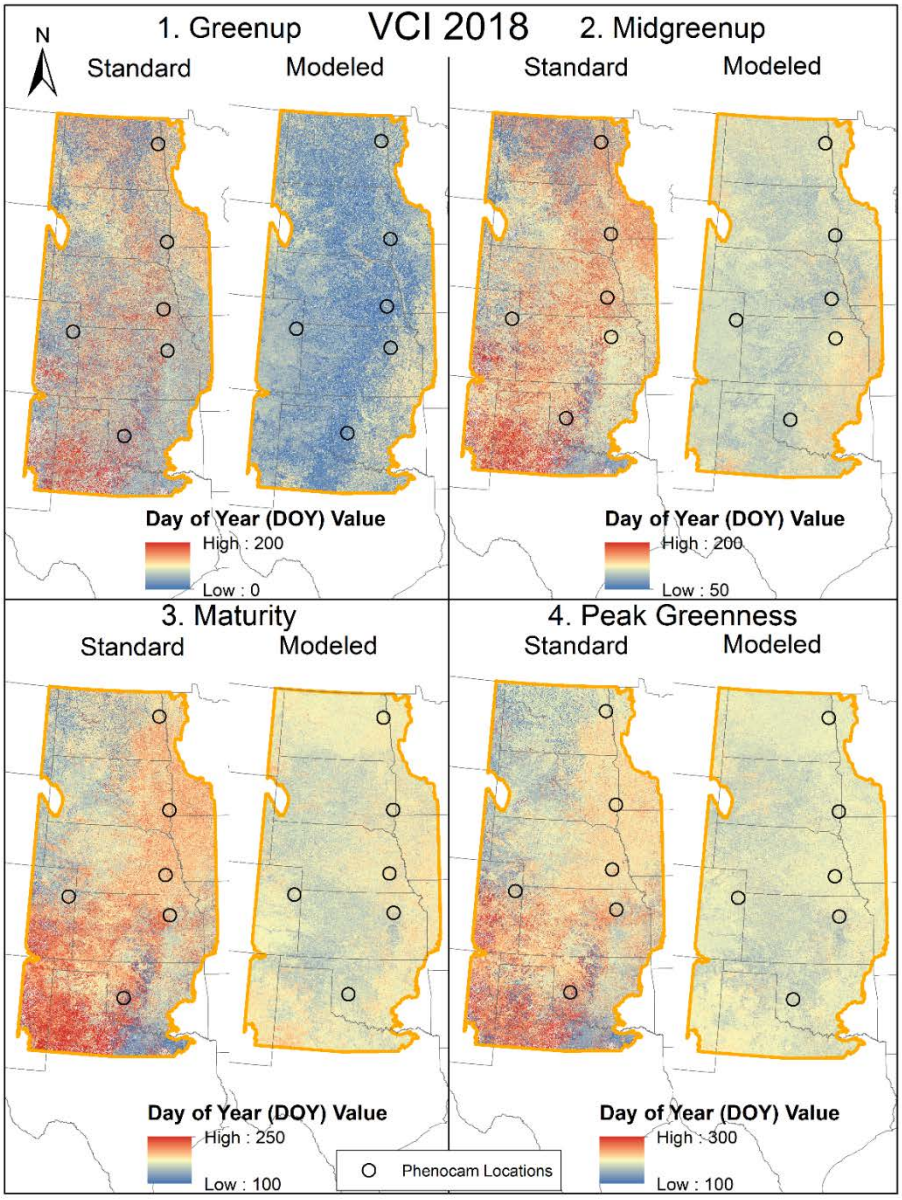


Figure A33: Maps showing the date of the first four VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2018.

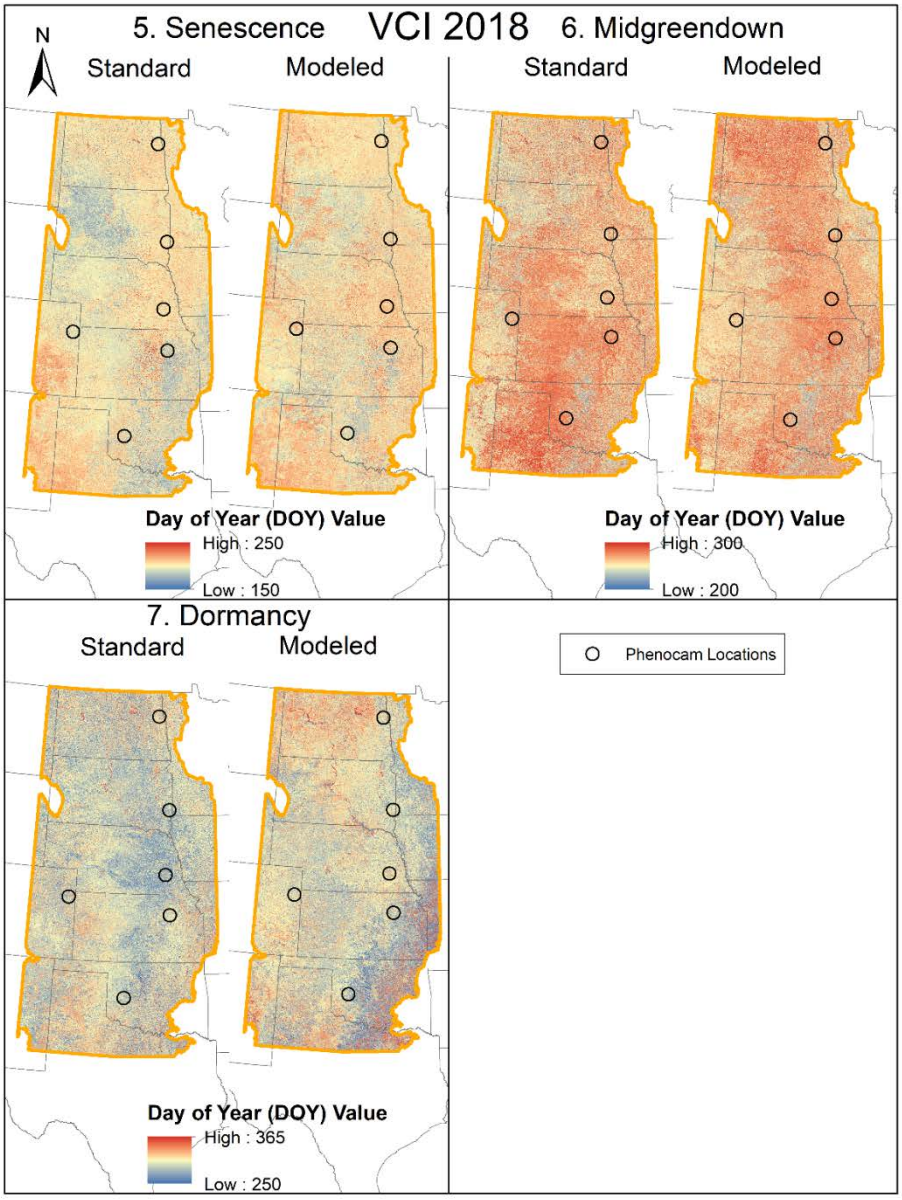


Figure A34: Maps showing the date of the last three VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2018.

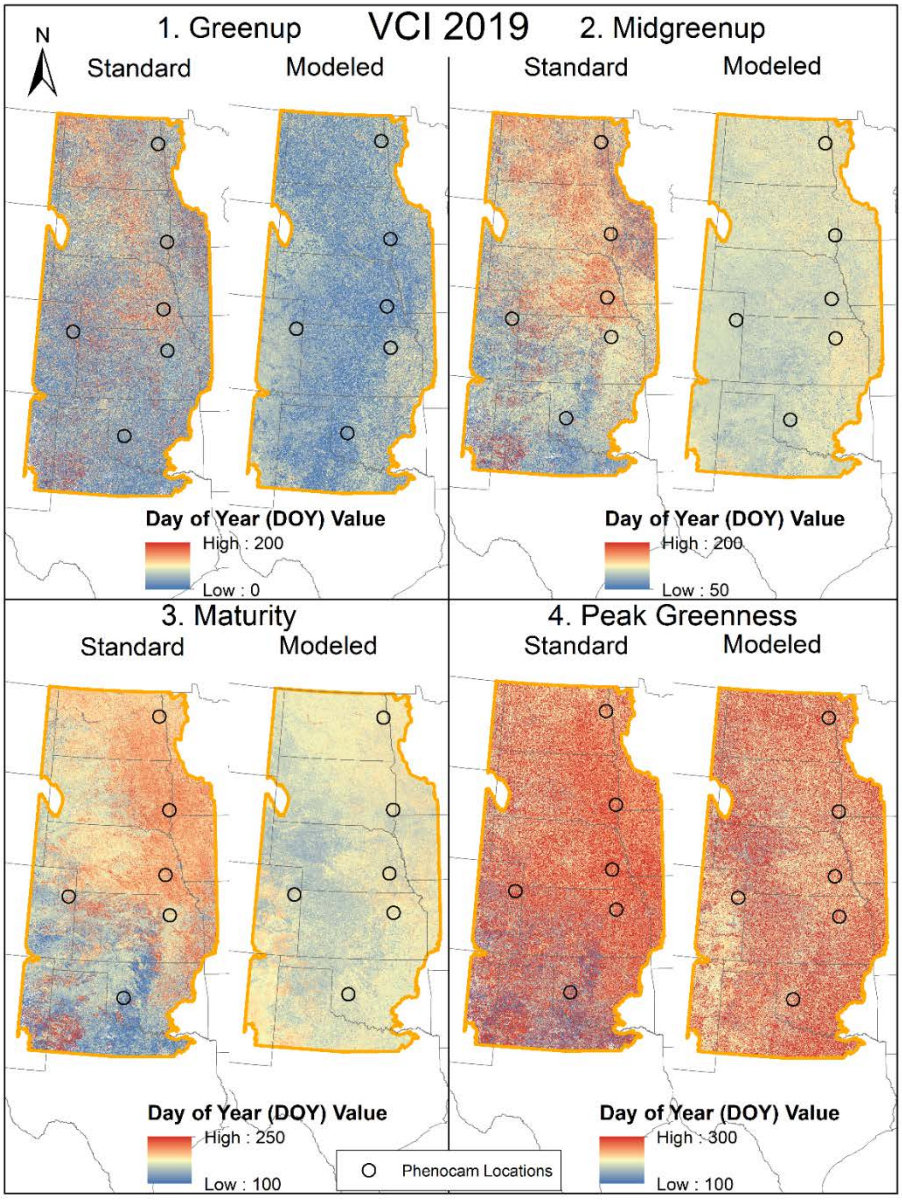


Figure A35: Maps showing the date of the first four VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2019.

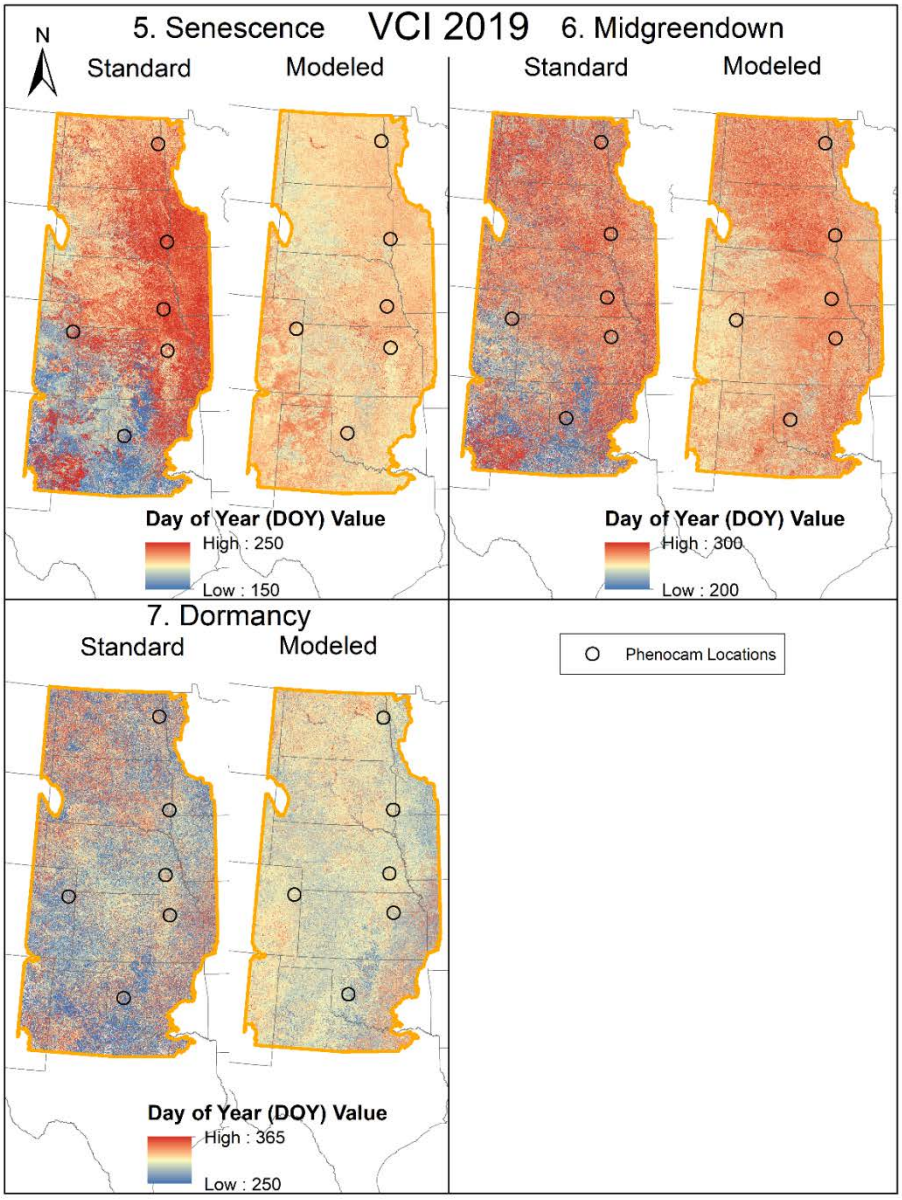


Figure A36: Maps showing the date of the last three VCI phenophases for both the standard MODIS imagery and the modeled MODIS imagery in 2019.

REFERENCES

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jozefowicz, R., Jia, Y., Kaiser, L., Kudlur, M., … Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems*. tensorflow.org

Aghighi, H., Azadbakht, M., Ashourloo, D., Shahrabi, H. S., & Radiom, S. (2018). Machine Learning Regression Techniques for the Silage Maize Yield Prediction Using Time-Series Images of Landsat 8 OLI. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(12), 4563–4577. https://doi.org/10.1109/JSTARS.2018.2823361

Ali, I., Cawkwell, F., Dwyer, E., Barrett, B., & Green, S. (2016). Satellite remote sensing of grasslands : from observation to management. *Journal of Plant Ecology*, *9*(6), 649–671. https://doi.org/10.1093/jpe/rtw005

Anderson, J. M. (1991). The Effects of Climate Change on Decomposition Processes in Grassland and Coniferous Forests. *Ecological Applications*, *1*(3), 326–347.

Aono, Y., & Kazui, K. (2008). Phenological data series of cherry tree flowering in Kyoto, Japan, and its application to reconstruction of springtime temperatures since the 9th century. *International Journal of Climatology*, *28*(7), 905–914. https://doi.org/10.1002/joc.1594

Aono, Y., & Saito, S. (2010). Clarifying springtime temperature reconstructions of the medieval period by gap-filling the cherry blossom phenological data series at Kyoto,

Japan. *International Journal of Biometeorology*, *54*(2), 211–219. https://doi.org/10.1007/s00484-009-0272-x

Baldocchi, D. (2008). “Breathing” of the terrestrial biosphere: lessons learned from a global network of carbon dioxide flux measurement systems. *Australian Journal of Botany, 56*, 1–26.

Bannari, A., Morin, D., Bonn, F., & Huete, A. R. (1995). A review of vegetation indices. *Remote Sensing Reviews, 13*(1–2), 95–120. https://doi.org/10.1080/02757259509532298

Baumann, M., Ozdogan, M., Richardson, A. D., & Radeloff, V. C. (2017). Phenology from Landsat when data is scarce: Using MODIS and Dynamic Time-Warping to combine multi-year Landsat imagery to derive annual phenology curves. *International Journal of Applied Earth Observation and Geoinformation*, *54*, 72–83. https://doi.org/10.1016/j.jag.2016.09.005

Belda, S., Pipia, L., Morcillo-Pallarés, P., Rivera-Caicedo, J. P., Amin, E., De Grave, C., & Verrelst, J. (2020). DATimeS: A machine learning time series GUI toolbox for gap-filling and vegetation phenology trends detection. *Environmental Modelling and Software*, *127*(November 2019). https://doi.org/10.1016/j.envsoft.2020.104666

Bolton, D. K., Gray, J. M., Melaas, E. K., Moon, M., Eklundh, L., & Friedl, M. A. (2020). Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sensing of Environment*, *240*(January), 111685. https://doi.org/10.1016/j.rse.2020.111685

Brown, L. A., Dash, J., Ogutu, B. O., & Richardson, A. D. (2017). On the relationship

between continuous measures of canopy greenness derived using near-surface remote sensing and satellite-derived vegetation products. *Agricultural and Forest Meteorology*, *247*(August), 280–292. https://doi.org/10.1016/j.agrformet.2017.08.012

Brown, T. B., Hultine, K. R., Steltzer, H., Denny, E. G., Denslow, M. W., Granados, J., Henderson, S., Moore, D., Nagai, S., Sanclements, M., Sánchez-Azofeifa, A., Sonnentag, O., Tazik, D., & Richardson, A. D. (2016). Using phenocams to monitor our changing earth: Toward a global phenocam network. *Frontiers in Ecology and the Environment*, *14*(2), 84–93. https://doi.org/10.1002/fee.1222

Browning, D. M., Karl, J. W., Morin, D., Richardson, A. D., & Tweedie, C. E. (2017). Phenocams bridge the gap between field and satellite observations in an arid grassland ecosystem. *Remote Sensing*, *9*(10). https://doi.org/10.3390/rs9101071

Buitenwerf, R., Rose, L., & Higgins, S. I. (2015). Three decades of multi-dimensional change in global leaf phenology. *Nature Climate Change*, *5*(April), 364–368. https://doi.org/10.1038/NCLIMATE2533

Burke, M. W. V., & Rundquist, B. C. (2021). Scaling Phenocam GCC, NDVI, and EVI2 with Harmonized Landsat-Sentinel using Gaussian Processes. *Agricultural and Forest Meteorology*, *300*(January), 108316. https://doi.org/10.1016/j.agrformet.2020.108316

Burke, M. W. V., Shahabi, M., Xu, Y., Zheng, H., Zhang, X., & Vanlooy, J. (2018). Identifying the driving factors of water quality in a sub-watershed of the republican river basin, Kansas USA. *International Journal of Environmental Research and Public Health*, *15*(5). https://doi.org/10.3390/ijerph15051041

Camps-Valls, G., Verrelst, J., Munoz-Mari, J., Laparra, V., Mateo-Jimenez, F., & Gomez-Dans, J. (2016). A survey on Gaussian processes for earth-observation data analysis: A comprehensive investigation. *IEEE Geoscience and Remote Sensing Magazine*, *4*(2), 58–78. https://doi.org/10.1109/MGRS.2015.2510084

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, *13-17-Augu*, 785–794. https://doi.org/10.1145/2939672.2939785

Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J. C., Skakun, S. V., & Justice, C. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment*, *219*(August 2017), 145–161. https://doi.org/10.1016/j.rse.2018.09.002

Cleland, E. E., Chiariello, N. R., Loarie, S. R., Mooney, H. A., & Field, C. B. (2006). Diverse responses of phenology to global changes in a grassland ecosystem. *Proceedings of the National Academy of Sciences*, *103*(37), 13740–13744. https://doi.org/10.1073/pnas.0600815103

Cleland, Elsa E., Chuine, I., Menzel, A., Mooney, H. A., & Schwartz, M. D. (2007). Shifting plant phenology in response to global change. *Trends in Ecology and Evolution*, *22*(7), 357–365. https://doi.org/10.1016/j.tree.2007.04.003

Coburn, C. A., & Peddle, D. R. (2006). A low-cost field and laboratory goniometer system for estimating hyperspectral bidirectional reflectance. *Canadian Journal of Remote Sensing*, *32*(3), 244–253. https://doi.org/10.5589/m06-021

Coulson, K. L., Bouricius, G. M., & Gray, E. L. (1965). Optical Reflection Properties of Natural Surfaces. *Journal of Geophysical Research*, *70*(18), 4601–46111. https://doi.org/10.1029/JZ070i018p04601

Cremonese, E., Filippa, G., Galvagno, M., Siniscalco, C., Oddi, L., Morra di Cella, U., & Migliavacca, M. (2017). Heat wave hinders green wave: The impact of climate extreme on the phenology of a mountain grassland. *Agricultural and Forest Meteorology*, *247*(August), 320–330. https://doi.org/10.1016/j.agrformet.2017.08.016

Crimmins, T. M., Crimmins, M. A., Gerst, K. L., Rosemartin, A. H., & Weltzin, J. F. (2017). USA National Phenology Network's volunteer-contributed observations yield predictive models of phenological transitions. *PLoS ONE*, *12*(8), 1–17. https://doi.org/10.1371/journal.pone.0182919

Cui, T., Martz, L., Lamb, E. G., Zhao, L., & Guo, X. (2019). Comparison of Grassland Phenology Derived from MODIS Satellite and PhenoCam Near-Surface Remote Sensing in North America. *Canadian Journal of Remote Sensing*, *45*(5), 707–722. https://doi.org/10.1080/07038992.2019.1674643

Cui, T., Martz, L., Zhao, L., & Guo, X. (2020). Investigating the impact of the temporal resolution of MODIS data on measured phenology in the prairie grasslands. *GIScience and Remote Sensing*, *57*(3), 395–410. https://doi.org/10.1080/15481603.2020.1723279

Dai, W., Jin, H., Zhang, Y., Liu, T., & Zhou, Z. (2019). Detecting temporal changes in the temperature sensitivity of spring phenology with global warming: Application of machine learning in phenological model. *Agricultural and Forest Meteorology*,

279(August), 107702. https://doi.org/10.1016/j.agrformet.2019.107702

De Beurs, K. M., & Henebry, G. M. (2004). Land surface phenology, climatic variation, and institutional change: Analyzing agricultural land cover change in Kazakhstan. *Remote Sensing of Environment*, *89*(4), 497–509. https://doi.org/10.1016/j.rse.2003.11.006

Denny, E. G., Gerst, K. L., Miller-Rushing, A. J., Tierney, G. L., Crimmins, T. M., Enquist, C. A. F., Guertin, P., Rosemartin, A. H., Schwartz, M. D., Thomas, K. A., & Weltzin, J. F. (2014). Standardized phenology monitoring methods to track plant and animal activity for science and resource management applications. *International Journal of Biometeorology*, *58*(4), 591–601. https://doi.org/10.1007/s00484-014-0789-5

Drury, M. (2020). *Basis Expansions*. https://github.com/madrury/basis-expansions

Duvenaud, D. K. (2014). *Automatic Model Construction with Gaussian Processes* (Issue June) [University of Cambridge]. https://www.cs.toronto.edu/~duvenaud/thesis.pdf

Elmore, A. J., Guinn, S. M., Minsley, B. J., & Richardson, A. D. (2012). Landscape controls on the timing of spring, autumn, and growing season length in mid-Atlantic forests. *Global Change Biology*, *18*(2), 656–674. https://doi.org/10.1111/j.1365-2486.2011.02521.x

Filippa, G., Cremonese, E., Migliavacca, M., Galvagno, M., Forkel, M., Wingate, L., Tomelleri, E., Morra di Cella, U., & Richardson, A. D. (2016). Phenopix: A R package for image-based vegetation phenology. *Agricultural and Forest Meteorology*, *220*, 141–150. https://doi.org/10.1016/j.agrformet.2016.01.006

Filippa, G., Cremonese, E., Migliavacca, M., Galvagno, M., Sonnentag, O., Humphreys, E.,

Hufkens, K., Ryu, Y., Verfaillie, J., Morra di Cella, U., & Richardson, A. D. (2018). NDVI derived from near-infrared-enabled digital cameras: Applicability across different plant functional types. *Agricultural and Forest Meteorology*, *249*(August 2017), 275–285. https://doi.org/10.1016/j.agrformet.2017.11.003

Fisher, Jeremy I., & Mustard, J. F. (2007). Cross-scalar satellite phenology from ground, Landsat, and MODIS data. *Remote Sensing of Environment*, *109*(3), 261–273. https://doi.org/10.1016/j.rse.2007.01.004

Fisher, Jeremy Isaac, Mustard, J. F., & Vadeboncoeur, M. A. (2006). Green leaf phenology at Landsat resolution: Scaling from the field to the satellite. *Remote Sensing of Environment*, *100*(2), 265–279. https://doi.org/10.1016/j.rse.2005.10.022

Flanagan, L. B., & Adkinson, A. C. (2011). Interacting controls on productivity in a northern Great Plains grassland and implications for response to ENSO events. *Global Change Biology*, *17*(11), 3293–3311. https://doi.org/10.1111/j.1365-2486.2011.02461.x

Fu, Y. H., Zhao, H., Piao, S., Peaucelle, M., Peng, S., Zhou, G., Ciais, P., Huang, M., Menzel, A., Peñuelas, J., Song, Y., Vitasse, Y., Zeng, Z., & Janssens, I. A. (2015). Declining global warming effects on the phenology of spring leaf unfolding. *Nature*, *526*(7571), 104–107. https://doi.org/10.1038/nature15402

Ganguly, S., Friedl, M. A., Tan, B., Zhang, X., & Verma, M. (2010). Land surface phenology from MODIS: Characterization of the Collection 5 global land cover dynamics product. *Remote Sensing of Environment*, *114*(8), 1805–1816. https://doi.org/10.1016/j.rse.2010.04.005

Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006). On the Blending of the MODIS and Landsat ETM + Surface Reflectance：Predicting Daily Landsat Surface Reflectanc. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(8), 2207–2218.

Gillespie, A. R., Kahle, A. B., & Walker, R. E. (1987). Color enhancement of highly correlated images. II. Channel ratio and "chromaticity" transformation techniques. *Remote Sensing of Environment*, *22*(3), 343–365. https://doi.org/10.1016/0034-4257(87)90088-5

Glenn, E., Huete, A., Nagler, P., & Nelson, S. (2008). Relationship Between Remotely-sensed Vegetation Indices, Canopy Attributes and Plant Physiological Processes: What Vegetation Indices Can and Cannot Tell Us About the Landscape. *Sensors*, *8*(4), 2136–2160. https://doi.org/10.3390/s8042136

Goodin, D. G., Gao, J., & Henebry, G. M. (2004). The effect of solar illumination angle and sensor view angle on observed patterns of spatial structure in tallgrass prairie. *IEEE Transactions on Geoscience and Remote Sensing*, *42*(1), 154–165. https://doi.org/10.1109/TGRS.2003.815674

Gray, J., Sulla-Menashe, D., & Friedl, M. A. (2019). *User Guide to Collection 6 MODIS Land Cover Dynamics (MCD12Q2) Product*. *6*, 1–8.

Helman, D. (2018). Land surface phenology: What do we really 'see' from space? *Science of the Total Environment*, *618*, 665–673. https://doi.org/10.1016/j.scitotenv.2017.07.237

Henebry, G. M. (2013). Phenologies of North American Grasslands and Grasses. In Mark D. Schwartz (Ed.), *Phenology: An Integrative Environmental Science* (2nd ed.). Springer.

https://doi.org/10.1007/978-94-007-6925-0_11

Henebry, G. M., & de Beurs, K. M. (2013). Remote Sensing of Land Surface Phenology: A Prospectus. In M.D. Schwartz (Ed.), *Phenology: An Integrative Environmental Science* (2nd ed., pp. 385–411). Springer. https://doi.org/10.1007/978-94-007-6925-0_21

Henebry, G., & Su, H. (1995). Observing Spatial Structure in the Flint Hills using AVHRR Biweekly Composites of Maximum NDVI. In *Proceedings of the... North American Prairie Conference* (Vol. 14, pp. 143–151).

Higgins, S. I., Delgado-Cartay, M. D., February, E. C., & Combrink, H. J. (2011). Is there a temporal niche separation in the leaf phenology of savanna trees and grasses? *Journal of Biogeography*, *38*(11), 2165–2175. https://doi.org/10.1111/j.1365-2699.2011.02549.x

Hilker, T., Wulder, M. A., Coops, N. C., Seitz, N., White, J. C., Gao, F., Masek, J. G., & Stenhouse, G. (2009). Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sensing of Environment*, *113*(9), 1988–1999. https://doi.org/10.1016/j.rse.2009.05.011

Hoegh-Guldberg, O., D. J., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K. L., Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S. I., Thomas, A., Warren, R., & Zhou, G. (2018). Impacts of 1.5ºC Global Warming on Natural and Human Systems. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of. *[Masson-*

Delmotte, V., P. Zhai, H.-O. Pörtner, D. Roberts, J. Skea, P.R. Shukla, A. Pirani, W. Moufouma-Okia, C. Péan, R. Pidcock, S. Connors, J.B.R. Matthews, Y. Chen, X. Zhou, M.I. Gomis, E. Lonnoy, T. Maycock, M. Tignor, and T. Waterfield (Eds.)]. In Pr, 175–311. https://www.ipcc.ch/sr15

Hoegh-Guldberg, O., Jacob, D., Taylor, M., Bindi, M., Brown, S., Camilloni, I., Diedhiou, A., Djalante, R., Ebi, K. L., Engelbrecht, F., Guiot, J., Hijioka, Y., Mehrotra, S., Payne, A., Seneviratne, S. I., Thomas, A., Warren, R., & Zhou, G. (2018). Impacts of 1.5ºC Global Warming on Natural and Human Systems. In: Global Warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of . In *Special Report, Intergovernmental Panel on Climate Change* (Issue ISBN 978-92-9169-151-7).

Hufkens, K., Friedl, M., Sonnentag, O., Braswell, B. H., Milliman, T., & Richardson, A. D. (2012). Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology. *Remote Sensing of Environment*, *117*, 307–321. https://doi.org/10.1016/j.rse.2011.10.006

Hufkens, K., Keenan, T. F., Flanagan, L. B., Scott, R. L., Bernacchi, C. J., Joo, E., Brunsell, N. A., Verfaillie, J., & Richardson, A. D. (2016). Productivity of North American grasslands is increased under future climate scenarios despite rising aridity. *Nature Climate Change*, *6*(7), 710–714. https://doi.org/10.1038/nclimate2942

Hurvich, C. M., Simonoff, J. S., & Tsai, C.-L. (1998). Smoothing parameter selection in

nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *60*(2), 271–293. https://doi.org/10.1111/1467-9868.00125

Ide, R., & Oguma, H. (2010). Use of digital cameras for phenological observations. *Ecological Informatics*, *5*(5), 339–347. https://doi.org/10.1016/j.ecoinf.2010.07.002

Inoue, T., Nagai, S., Kobayashi, H., & Koizumi, H. (2015). Utilization of ground-based digital photography for the evaluation of seasonal changes in the aboveground green biomass and foliage phenology in a grassland ecosystem. *Ecological Informatics*, *25*, 1–9. https://doi.org/10.1016/j.ecoinf.2014.09.013

Jeong, S. J., Ho, C. H., Gim, H. J., & Brown, M. E. (2011). Phenology shifts at start vs. end of growing season in temperate vegetation over the Northern Hemisphere for the period 1982-2008. *Global Change Biology*, *17*(7), 2385–2399. https://doi.org/10.1111/j.1365-2486.2011.02397.x

Jiang, Z., Huete, A. R., Didan, K., & Miura, T. (2008). Development of a two-band enhanced vegetation index without a blue band. *Remote Sensing of Environment*, *112*(10), 3833–3845. https://doi.org/10.1016/j.rse.2008.06.006

Jones, H. G., & Vaughan, R. A. (2010). *Remote sensing of vegetation: principles, techniques, and applications* (1st ed.). Oxford university press.

Jönsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(8), 1824–1832. https://doi.org/10.1109/TGRS.2002.802519

Julitta, T., Cremonese, E., Migliavacca, M., Colombo, R., Galvagno, M., Siniscalco, C., Rossini, M., Fava, F., Cogliati, S., Morra di Cella, U., & Menzel, A. (2014). Using digital camera images to analyse snowmelt and phenology of a subalpine grassland. *Agricultural and Forest Meteorology*, *198–199*, 116–125. https://doi.org/10.1016/j.agrformet.2014.08.007

Kathuroju, N., White, M. A., Symanzik, J., Schwartz, M. D., Powell, J. A., & Nemani, R. R. (2007). On the use of the advanced very high resolution radiometer for development of prognostic land surface phenology models. *Ecological Modelling*, *201*(2), 144–156. https://doi.org/10.1016/j.ecolmodel.2006.09.011

Keenan, T. F., Darby, B., Felts, E., Sonnentag, O., Friedl, M. A., Hufkens, K., O'Keefe, J., Klosterman, S., Munger, J. W., Toomey, M., & Richardson, A. D. (2014). Tracking forest phenology and seasonal physiology using digital repeat photography: A critical assessment. *Ecological Applications*, *24*(6), 1478–1489. https://doi.org/10.1890/13-0652.1

Killick, R., Fearnhead, P., & Eckley, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, *107*(500), 1590–1598. https://doi.org/10.1080/01621459.2012.737745

Kingma, D. P., & Lei Ba, J. (2015). Adam: A method for stochastic optimization. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 1–15.

Klosterman, S. T., Hufkens, K., Gray, J. M., Melaas, E., Sonnentag, O., Lavine, I., Mitchell,

L., Norman, R., Friedl, M. A., & Richardson, A. D. (2014). Evaluating remote sensing of deciduous forest phenology at multiple spatial scales using PhenoCam imagery. *Biogeosciences*, *11*(16), 4305–4320. https://doi.org/10.5194/bg-11-4305-2014

Kosmala, M., Crall, A., Cheng, R., Hufkens, K., Henderson, S., & Richardson, A. D. (2016). Season spotter: Using citizen science to validate and scale plant phenology from near-surface remote sensing. *Remote Sensing*, *8*(9), 1–22. https://doi.org/10.3390/rs8090726

Kosmala, M., Hufkens, K., & Richardson, A. D. (2018). Integrating camera imagery, crowdsourcing, and deep learning to improve high-frequency automated monitoring of snow at continental-to-global scales. *PLoS ONE*, *13*(12), 1–19. https://doi.org/10.1371/journal.pone.0209649

Krienert, J. M. (2015). *NE9mile_AOI-Analysis*. Github. https://github.com/treystaff/PhenoAnalysis/blob/master/Python/FOV Analysis/Scripts/NE9mile_AOI-Analysis.py

Lenton, T. M., & Lovelock, J. E. (2001). Daisyworld revisited: Quantifying biological effects on planetary self-regulation. *Tellus, Series B: Chemical and Physical Meteorology*, *53*(3), 288–305. https://doi.org/10.1034/j.1600-0889.2001.01191.x

Liang, L., Schwartz, M. D., & Fei, S. (2011). Validating satellite phenology through intensive ground observation and landscape scaling in a mixed seasonal forest. *Remote Sensing of Environment*, *115*(1), 143–157. https://doi.org/10.1016/j.rse.2010.08.013

Lieth, H. (1974). *Phenology and Seasonality Modeling* (Vol. 8). https://doi.org/10.1007/978-3-642-51863-8

Liu, H. Q., & Huete, A. (1995). Feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Transactions on Geoscience and Remote Sensing*, *33*(2), 457–465. https://doi.org/10.1109/36.377946

Liu, Licong, Cao, R., Shen, M., Chen, J., Wang, J., & Zhang, X. (2019). How does scale effect influence spring vegetation phenology estimated from satellite-derived vegetation indexes? *Remote Sensing*, *11*(18). https://doi.org/10.3390/rs11182137

Liu, Lingling, & Zhang, X. (2020). Effects of temperature variability and extremes on spring phenology across the contiguous United States from 1982 to 2016. *Scientific Reports*, *10*(1), 1–14. https://doi.org/10.1038/s41598-020-74804-4

Liu, Q., Fu, Y. H., Zhu, Z., Liu, Y., Liu, Z., Huang, M., Janssens, I. A., & Piao, S. (2016). Delayed autumn phenology in the Northern Hemisphere is related to change in both climate and spring phenology. *Global Change Biology*, *22*(11), 3702–3711. https://doi.org/10.1111/gcb.13311

Liu, Y., Hill, M. J., Zhang, X., Wang, Z., Richardson, A. D., Hufkens, K., Filippa, G., Baldocchi, D. D., Ma, S., Verfaillie, J., & Schaaf, C. B. (2017). Using data from Landsat, MODIS, VIIRS and PhenoCams to monitor the phenology of California oak/grass savanna and open grassland across spatial scales. *Agricultural and Forest Meteorology*, *237–238*, 311–325. https://doi.org/10.1016/j.agrformet.2017.02.026

McMaster, G. S., & Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, *87*(1), 291–300. https://doi.org/10.1016/S0168-1923(97)00027-0

Melaas, E. K., Friedl, M. A., & Richardson, A. D. (2016). Multiscale modeling of spring phenology across Deciduous Forests in the Eastern United States. *Global Change Biology*, *22*(2), 792–805. https://doi.org/10.1111/gcb.13122

Melaas, E. K., Sulla-Menashe, D., Gray, J. M., Black, T. A., Morin, T. H., Richardson, A. D., & Friedl, M. A. (2016). Multisite analysis of land surface phenology in North American temperate and boreal deciduous forests from Landsat. *Remote Sensing of Environment*, *186*, 452–464. https://doi.org/10.1016/j.rse.2016.09.014

Menzel, A. (2013). Europe. In M.D. Schwartz (Ed.), *Phenology: An Integrative Environmental Science* (2nd ed., pp. 53–65). Springer. https://doi.org/10.1007/978-94-007-6925-0_4

Menzel, A., Estrella, N., & Testka, A. (2005). Temperature response rates from long-term phenological records. *Climate Research*, *30*(1), 21–28. https://doi.org/10.3354/cr030021

Migliavacca, M., Galvagno, M., Cremonese, E., Rossini, M., Meroni, M., Sonnentag, O., Cogliati, S., Manca, G., Diotri, F., Busetto, L., Cescatti, A., Colombo, R., Fava, F., Morra di Cella, U., Pari, E., Siniscalco, C., & Richardson, A. D. (2011). Using digital repeat photography and eddy covariance data to model grassland phenology and photosynthetic CO2 uptake. *Agricultural and Forest Meteorology*, *151*(10), 1325–1337. https://doi.org/10.1016/j.agrformet.2011.05.012

Mohler, R. L., & Goodin, D. G. (2013). Temporal burn scar evolution in tallgrass prairie based on field spectroscopy. *International Journal of Remote Sensing*, *34*(20), 7199–7217. https://doi.org/10.1080/01431161.2013.817713

Morisette, J. T., Richardson, A. D., Knapp, A. K., Fisher, J. I., Graham, E. A., Abatzoglou, J., Wilson, B. E., Breshears, D. D., Henebry, G. M., Hanes, J. M., & Liang, L. (2009). Tracking the rhythm of the seasons in the face of global change: Phenological research in the 21 st century. *Frontiers in Ecology and the Environment*, *7*(5), 253–260. https://doi.org/10.1890/070217

Nagai, S., Akitsu, T., Saitoh, T. M., Busey, R. C., Fukuzawa, K., Honda, Y., Ichie, T., Ide, R., Ikawa, H., Iwasaki, A., Iwao, K., Kajiwara, K., Kang, S., Kim, Y., Khoon, K. L., Kononov, A. V., Kosugi, Y., Maeda, T., Mamiya, W., … Nasahara, K. N. (2018). 8 million phenological and sky images from 29 ecosystems from the Arctic to the tropics: the Phenological Eyes Network. *Ecological Research*, *33*(6), 1091–1092. https://doi.org/10.1007/s11284-018-1633-x

NDMC, USDA, & NOAA. (2020). *U.S. Drought Monitor*. National Drought Mitigation Center, U.S. Department of Agriculture, National Oceanic and Atmospheric Association. https://droughtmonitor.unl.edu/

O'Connell, J. L., & Alber, M. (2016). A smart classifier for extracting environmental data from digital image time-series: Applications for PhenoCam data in a tidal salt marsh. *Environmental Modelling and Software*, *84*, 134–139. https://doi.org/10.1016/j.envsoft.2016.06.025

Omernik, J. M., & Griffith, G. E. (2014). Ecoregions of the Conterminous United States: Evolution of a Hierarchical Spatial Framework. *Environmental Management*, *54*(6), 1249–1266. https://doi.org/10.1007/s00267-014-0364-1

Pan, F., Converse, T., Ahn, D., Salvetti, F., & Donato, G. (2009). Feature selection for ranking using boosted trees. *International Conference on Information and Knowledge Management, Proceedings*, 2025–2028. https://doi.org/10.1145/1645953.1646292

Pastick, N. J., Dahal, D., Wylie, B. K., Parajuli, S., & Boyte, S. P. (2020). Characterizing Land Surface Phenology and Exotic Annual Grasses in Dryland Ecosystems Using Landsat and Sentinel-2 Data in Harmony. *Remote Sensing*, *12*(725), 17. https://doi.org/10.3390/rs12040725

Peng, D., Wang, Y., Xian, G., Huete, A. R., Huang, W., Shen, M., Wang, F., Yu, L., Liu, L., Xie, Q., Liu, L., & Zhang, X. (2021). Investigation of land surface phenology detections in shrublands using multiple scale satellite data. *Remote Sensing of Environment*, *252*(October 2020). https://doi.org/10.1016/j.rse.2020.112133

Peng, D., Wu, C., Li, C., Zhang, X., Liu, Z., Ye, H., Luo, S., Liu, X., Hu, Y., & Fang, B. (2017). Spring green-up phenology products derived from MODIS NDVI and EVI: Intercomparison, interpretation and validation using National Phenology Network and AmeriFlux observations. *Ecological Indicators*, *77*, 323–336. https://doi.org/10.1016/j.ecolind.2017.02.024

Peng, D., Zhang, X., Zhang, B., Liu, L., Liu, X., Huete, A. R., Huang, W., Wang, S., Luo, S., Zhang, X., & Zhang, H. (2017). Scaling effects on spring phenology detections from MODIS data at multiple spatial resolutions over the contiguous United States. *ISPRS Journal of Photogrammetry and Remote Sensing*, *132*, 185–198. https://doi.org/10.1016/j.isprsjprs.2017.09.002

Petach, A. R., Toomey, M., Aubrecht, D. M., & Richardson, A. D. (2014). Monitoring vegetation phenology using an infrared-enabled security camera. *Agricultural and Forest Meteorology*, *195–196*, 143–151. https://doi.org/10.1016/j.agrformet.2014.05.008

Petrie, M. D., Brunsell, N. A., Vargas, R., Collins, S. L., Flanagan, L. B., Hanan, N. P., Litvak, M. E., & Suyker, A. E. (2016). The sensitivity of carbon exchanges in Great Plains grasslands to precipitation variability. *Journal of Geophysical Research: Biogeosciences*, *121*(2), 280–294. https://doi.org/10.1002/2015JG003205

Pettorelli, N., Vik, J. O., Mysterud, A., Gaillard, J. M., Tucker, C. J., & Stenseth, N. C. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology and Evolution*, *20*(9), 503–510. https://doi.org/10.1016/j.tree.2005.05.011

Piao, S., Liu, Q., Chen, A., Janssens, I. A., Fu, Y., Dai, J., Liu, L., Lian, X., Shen, M., & Zhu, X. (2019). Plant phenology and global climate change: current progresses and challenges. *Global Change Biology*, *February*, gcb.14619. https://doi.org/10.1111/gcb.14619

Pitman, A. J. (2003). The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, *23*(5), 479–510. https://doi.org/10.1002/joc.893

Planet Team. (2016). *Planet Imagery Product Specification: Planetscope & Rapideye* (Issue November). https://www.planet.com/products/satellite-

imagery/files/1611.09_Spec_Sheet_Combined_Imagery_Product_Letter_DraftV3.pdf

Potter, C. S., & Brooks, V. (1998). Global analysis of empirical relations between annual climate and seasonality of NDVI. *International Journal of Remote Sensing*, *19*(15), 2921–2948. https://doi.org/10.1080/014311698214352

Prevéy, J. S., & Seastedt, T. R. (2014). Seasonality of precipitation interacts with exotic species to alter composition and phenology of a semi-arid grassland. *Journal of Ecology*, *102*(6), 1549–1561. https://doi.org/10.1111/1365-2745.12320

Rasmussen, C. E., & Williams, C. K. I. (2006). Gaussian processes for machine learning. In *The MIT Press, Cambridge, MA, USA* (Vol. 38, Issue 2).

Reed, B. C. (2006). Trend analysis of time-series phenology of North America derived from satellite data. *GIScience and Remote Sensing*, *43*(1), 24–38. https://doi.org/10.2747/1548-1603.43.1.24

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*(5), 703–714. https://doi.org/10.2307/3235884

Reed, B. C., Schwartz, M. D., & Xiao, X. (2009). Remote sensing phenology: Status and the way forward. In A. Noormets (Ed.), *Phenology of Ecosystem Processes - Application in global change research* (1st ed., pp. 231–246). Springer. https://doi.org/10.1007/978-1-4419-0026-5

Reed, B. C., White, M., & Brown, J. F. (2003). Remote Sensing Phenology. In Mark D. Schwartz (Ed.), *Phenology: An Integrative Environmental Science* (pp. 365–381).

Springer.

Ren, S., Chen, X., Lang, W., & Schwartz, M. D. (2018). Climatic Controls of the Spatial Patterns of Vegetation Phenology in Midlatitude Grasslands of the Northern Hemisphere. *Journal of Geophysical Research: Biogeosciences*, *123*(8), 2323–2336. https://doi.org/10.1029/2018JG004616

Richardson, A., & Braswell, B. (2009). Near-surface remote sensing of spatial and temporal variation in canopy phenology. *Ecological Applications*, *19*(6), 1417–1428. http://www.esajournals.org/doi/abs/10.1890/08-2022.1

Richardson, A. D. (2019). Tracking seasonal rhythms of plants in diverse ecosystems with digital camera imagery. *New Phytologist*, *222*(4), 1742–1750. https://doi.org/10.1111/nph.15591

Richardson, A. D., Hollinger, D. Y., Dail, D. B., Lee, J. T., Munger, J. W., & O'Keefe, J. (2009). Influence of spring phenology on seasonal and annual carbon balance in two contrasting New England forests. *Tree Physiology*, *29*(3), 321–331. https://doi.org/10.1093/treephys/tpn040

Richardson, A. D., Hufkens, K., Li, X., & Ault, T. R. (2019). Testing Hopkins' Bioclimatic Law with PhenoCam data. *Applications in Plant Sciences*, *7*(3). https://doi.org/10.1002/aps3.1228

Richardson, A. D., Hufkens, K., Milliman, T., Aubrecht, D. M., Chen, M., Gray, J. M., Johnston, M. R., Keenan, T. F., Klosterman, S. T., Kosmala, M., Melaas, E. K., Friedl, M. A., & Frolking, S. (2018). Tracking vegetation phenology across diverse North

American biomes using PhenoCam imagery. *Scientific Data*, *5*, 180028. https://doi.org/10.1038/sdata.2018.28

Richardson, A. D., Jenkins, J. P., Braswell, B. H., Hollinger, D. Y., Ollinger, S. V., & Smith, M. L. (2007). Use of digital webcam images to track spring green-up in a deciduous broadleaf forest. *Oecologia*, *152*(2), 323–334. https://doi.org/10.1007/s00442-006-0657-z

Richardson, A. D., Keenan, T. F., Migliavacca, M., Ryu, Y., Sonnentag, O., & Toomey, M. (2013). Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agricultural and Forest Meteorology*, *169*, 156–173. https://doi.org/10.1016/j.agrformet.2012.09.012

Richardson, A. D., Klosterman, S., & Toomey, M. (2013). Near-Surface Sensor-Derived Phenology. In M.D. Schwartz (Ed.), *Phenology: An Integrative Environmental Science* (2nd ed., pp. 413–430). Springer. https://doi.org/10.1007/978-94-007-6925-0_22

Robinson, N. P., Allred, B. W., Jones, M. O., Moreno, A., Kimball, J. S., Naugle, D. E., Erickson, T. A., & Richardson, A. D. (2017). A dynamic landsat derived normalized difference vegetation index (NDVI) product for the conterminous United States. *Remote Sensing*, *9*(8), 1–14. https://doi.org/10.3390/rs9080863

Rodrigues, A., Marcal, A. R. S., & Cunha, M. (2012). Phenology parameter extraction from time-series of satellite vegetation index data using phenosat. *International Geoscience and Remote Sensing Symposium (IGARSS)*, 4926–4929. https://doi.org/10.1109/IGARSS.2012.6352507

Rodriguez-Galiano, V. F., Dash, J., & Atkinson, P. M. (2015). Intercomparison of satellite sensor land surface phenology and ground phenology in Europe. *Geophysical Research Letters*, *42*(7), 2253–2260. https://doi.org/10.1002/2015GL063586

Rouse, J. W. J., Haas, R. H., Schell, J. A., & Deering, D. W. (1973). *Monitoring the Vernal Advancement and Retrogradation (Green Wave Effect) of Natural Vegetation*.

Rundquist, B. C., & Harrington, J. A. (2000). The effects of climatic factors on vegetation dynamics of tallgrass and shortgrass cover. *Geocarto International*, *15*(3), 33–38. https://doi.org/10.1080/10106040008542161

Schwartz, M.D. (2013). Introduction. In M.D. Schwartz (Ed.), *Phenology: An Integrative Environmental Science* (2nd ed., pp. 1–5). Springer. https://doi.org/10.1007/978-94-007-6925-0_1

Schwartz, Mark D. (1994). Monitoring global change with phenology: The case of the spring green wave. *International Journal of Biometeorology*, *38*(1), 18–22. https://doi.org/10.1007/BF01241799

Schwartz, Mark D. (2003). *Phenology: An integrative environmental science* (M. D. SCHWARTZ (ed.)).

Schwartz, Mark D., Ahas, R., & Aasa, A. (2006). Onset of spring starting earlier across the Northern Hemisphere. *Global Change Biology*, *12*(2), 343–351. https://doi.org/10.1111/j.1365-2486.2005.01097.x

Schwartz, Mark D., Beaubien, E. G., Crimmins, T. M., & Weltzin, J. F. (2013). North America. In M.D. Schwartz (Ed.), *Phenology: An Integrative Environmental Science*

(2nd ed., pp. 67–89). Springer. https://doi.org/10.1007/978-94-007-6925-0_5

Scurlock, J. M. O., & Hall, D. O. (1998). The global carbon sink: a grassland perspective. *Global Change Biology*, *4*, 229–233.

Seyednasrollah, B., Young, A. M., Hufkens, K., Milliman, T., Friedl, M. A., Frolking, S., & Richardson, A. D. (2019). Tracking vegetation phenology across diverse biomes using Version 2.0 of the PhenoCam Dataset. *Scientific Data*, *6*(1), 222. https://doi.org/10.1038/s41597-019-0229-9

Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM National Conference On* -, 517–524. https://doi.org/10.1145/800186.810616

Snyder, K. A., Huntington, J. L., Wehan, B. L., Morton, C. G., & Stringham, T. K. (2019). Comparison of landsat and land-based phenology camera normalized difference vegetation index (NDVI) for dominant plant communities in the great basin. *Sensors (Switzerland)*, *19*(5). https://doi.org/10.3390/s19051139

Sonnentag, O., Hufkens, K., Teshera-Sterne, C., Young, A. M., Friedl, M., Braswell, B. H., Milliman, T., O'Keefe, J., & Richardson, A. D. (2012). Digital repeat photography for phenological research in forest ecosystems. *Agricultural and Forest Meteorology*, *152*(1), 159–177. https://doi.org/10.1016/j.agrformet.2011.09.009

Sparks, T. H., & Menzel, A. (2002). Observed changes in seasons: An overview. *International Journal of Climatology*, *22*(14), 1715–1725. https://doi.org/10.1002/joc.821

Steltzer, H., & Welker, J. M. (2006). Modeling the effect of photosynthetic vegetation properties on the NDVI–LAI relationship. *Ecology*, *87*(11), 2765–2772. https://doi.org/10.1890/0012-9658(2006)87

Stubbendieck, J., Hatch, S. L., & Dunn, C. D. (2017). *Grasses of the Great Plains* (First). Texas A&M University Press.

Tang, J., Körner, C., Muraoka, H., Piao, S., Shen, M., Thackeray, S. J., & Yang, X. (2016). Emerging opportunities and challenges in phenology: A review. *Ecosphere*, *7*(8), 1–17. https://doi.org/10.1002/ecs2.1436

The Intergovernmental Panel on Climate Change. (2007). *Climate Change 2007: Impacts, Adaptation and Vulnerability* (M. Parry, O. Canziani, J. Palutikof, P. van der Linden, & C. Hanson (eds.)). Cambridge University Press.

Thornton, P. E., Thornton, M. M., Mayer, B. W., Wei, Y., Devarakonda, R., Vose, R. S., & Cook., R. B. (2018). Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 3. In *ORNL DAAC*. https://doi.org/10.3334/ORNLDAAC/1328

Tierney, G., Mitchell, B., Miller-Rushing, A., Katz, J., Denny, E., Brauer, C., Donovan, T., Richardson, A. D., Toomey, M., Kozlowski, A., Weltzin, J., Gerst, K., Sharron, E., Sonnentag, O., & Dieffenbach, F. (2013). Phenology monitoring protocol: Northeast Temperate Network. Natural Resource Report NPS/NETN//NRR—2013/681. *National Park Service Natural Resource Report*, 272. http://pubs.er.usgs.gov/publication/70047249

Toda, M., & Richardson, A. D. (2018). Estimation of plant area index and phenological

transition dates from digital repeat photography and radiometric approaches in a hardwood forest in the Northeastern United States. *Agricultural and Forest Meteorology*, *249*(August 2017), 457–466. https://doi.org/10.1016/j.agrformet.2017.09.004

Toomey, M., Friedl, M. A., Frolking, S., Hufkens, K., Klosterman, S., Sonnentag, O., Baldocchi, D. D., Bernacchi, C. J., Biraud, S. C., Bohrer, G., & Brzostek, E. (2015). *Greenness indices from digital cameras predict the timing and seasonal dynamics of canopy-scale photosynthesis*.

U.S. Environmental Protection Agency. (2020). *Ecoregions of North America*. https://www.epa.gov/eco-research/ecoregions-north-america

U.S. Geological Survey. (2018). *What are the band designations for the Landsat satellites?* https://landsat.usgs.gov/what-are-band-designations-landsat-satellites

USA National phenology Network. (2018). *Status of Spring*. https://www.usanpn.org/news/spring

Van Vliet, A. J. H., De Groot, R. S., Bellens, Y., Braun, P., Bruegger, R., Bruns, E., Clevers, J., Estreguil, C., Flechsig, M., Jeanneret, F., Maggi, M., Martens, P., Menne, B., Menzel, A., & Sparks, T. (2003). The European Phenology Network. *International Journal of Biometeorology*, *47*(4), 202–212. https://doi.org/10.1007/s00484-003-0174-2

Verrelst, J., Alonso, L., Camps-Valls, G., Delegido, J., & Moreno, J. (2012). Retrieval of vegetation biophysical parameters using Gaussian process techniques. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(5 PART 2), 1832–1843.

https://doi.org/10.1109/TGRS.2011.2168962

Vrieling, A., Meroni, M., Darvishzadeh, R., Skidmore, A. K., Wang, T., Zurita-Milla, R., Oosterbeek, K., O'Connor, B., & Paganini, M. (2018). Vegetation phenology from Sentinel-2 and field cameras for a Dutch barrier island. *Remote Sensing of Environment*, *November 2017*, 0–1. https://doi.org/10.1016/j.rse.2018.03.014

Walker, J. J., de Beurs, K. M., & Wynne, R. H. (2014). Dryland vegetation phenology across an elevation gradient in Arizona, USA, investigated with fused MODIS and landsat data. *Remote Sensing of Environment*, *144*, 85–97. https://doi.org/10.1016/j.rse.2014.01.007

Wang, Cong, Chen, J., Wu, J., Tang, Y., Shi, P., Black, T. A., & Zhu, K. (2017). A snow-free vegetation index for improved monitoring of vegetation spring green-up date in deciduous ecosystems. *Remote Sensing of Environment*, *196*, 1–12. https://doi.org/10.1016/j.rse.2017.04.031

Wang, Cuizhen, Guo, H., Zhang, L., Liu, S., Qiu, Y., & Sun, Z. (2015). Assessing phenological change and climatic control of alpine grasslands in the Tibetan Plateau with MODIS time series. *International Journal of Biometeorology*, *59*(1), 11–23. https://doi.org/10.1007/s00484-014-0817-5

Wang, F., Luo, S., Wu, C., Zhang, X., Liu, X., Peng, D., Yu, L., Huete, A. R., & Zhang, H. (2018). Scaling up spring phenology derived from remote sensing images. *Agricultural and Forest Meteorology*, *256–257*(April 2017), 207–219. https://doi.org/10.1016/j.agrformet.2018.03.010

Wang, J., Rich, P. M., & Price, K. P. (2003). Temporal responses of NDVI to precipitation and temperature in the central Great Plains, USA. *International Journal of Remote Sensing*, *24*(11), 2345–2364. https://doi.org/10.1080/01431160210154812

Watson, A. J., & Lovelock, J. E. (1983). Biological homeostasis of the global environment: the parable of Daisyworld. *Tellus B*, *35 B*(4), 284–289. https://doi.org/10.1111/j.1600-0889.1983.tb00031.x

Watson, C. J., Restrepo-Coupe, N., & Huete, A. R. (2019). Multi-scale phenology of temperate grasslands: Improving monitoring and management with near-surface phenocams. *Frontiers in Environmental Science*, *7*(FEB), 1–18. https://doi.org/10.3389/fenvs.2019.00014

White, M. A., de Beurs, K. M., Didan, K., Inouye, D. W., Richardson, A. D., Jensen, O. P., O'Keefe, J., Zhang, G., Nemani, R. R., van Leeuwen, W. J. D., Brown, J. F., de Wit, A., Schaepman, M., Lin, X., Dettinger, M., Bailey, A. S., Kimball, J., Schwartz, M. D., Baldocchi, D. D., … Lauenroth, W. K. (2009). Intercomparison, interpretation, and assessment of spring phenologyReal-time monitoring and short-term forecasting of land surface phenology in North America estimated from remote sensing for 1982-2006. *Global Change Biology*, *15*(10), 2335–2359. https://doi.org/10.1111/j.1365-2486.2009.01910.x

White, M. A., & Nemani, R. R. (2006). Real-time monitoring and short-term forecasting of land surface phenology. *Remote Sensing of Environment*, *104*(1), 43–49. https://doi.org/10.1016/j.rse.2006.04.014

Wilsey, B. J., Martin, L. M., & Kaul, A. D. (2018). Phenology differences between native and novel exotic-dominated grasslands rival the effects of climate change. *Journal of Applied Ecology*, *55*(2), 863–873. https://doi.org/10.1111/1365-2664.12971

World Wildlife Fund. (2018). *The Plowprint Report: 2018*. worldwildlife.org/ngp

Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, *2017*. https://doi.org/10.1155/2017/1353691

Yingying, X. I. E., Civco, D. L., & Silander, J. A. (2018). Species-specific spring and autumn leaf phenology captured by time-lapse digital cameras. *Ecosphere*, *9*(1). https://doi.org/10.1002/ecs2.2089

Zeng, L., Wardlow, B. D., Xiang, D., Hu, S., & Li, D. (2020). A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment*, *237*(August 2019), 111511. https://doi.org/10.1016/j.rse.2019.111511

Zhang, H., Eziz, A., Xiao, J., Tao, S., Wang, S., Tang, Z., Zhu, J., & Fang, J. (2019). High-resolution vegetation mapping using eXtreme gradient boosting based on extensive features. *Remote Sensing*, *11*(12). https://doi.org/10.3390/rs11121505

Zhang, X. (2012). *Phenology and climte change* (X. Zhang (ed.)). InTech. www.intechopen.com

Zhang, X., Jayavelu, S., Liu, L., Friedl, M. A., Henebry, G. M., Liu, Y., Schaaf, C. B., Richardson, A. D., & Gray, J. (2018). Evaluation of land surface phenology from VIIRS

data using time series of PhenoCam imagery. *Agricultural and Forest Meteorology*, *256–257*(February), 137–149. https://doi.org/10.1016/j.agrformet.2018.03.003

Zhang, X., Liu, L., & Yan, D. (2017). Comparisons of global land surface seasonality and phenology derived from AVHRR, MODIS, and VIIRS data. *Journal of Geophysical Research: Biogeosciences*, *122*(6), 1506–1525. https://doi.org/10.1002/2017JG003811

Zhang, X., Wang, J., Gao, F., Liu, Y., Schaaf, C., Friedl, M., Yu, Y., Jayavelu, S., Gray, J., Liu, L., Yan, D., & Henebry, G. M. (2017). Exploration of scaling effects on coarse resolution land surface phenology. *Remote Sensing of Environment*, *190*, 318–330. https://doi.org/10.1016/j.rse.2017.01.001

Zhang, X., Wang, J., Henebry, G. M., & Gao, F. (2020). Development and evaluation of a new algorithm for detecting 30 m land surface phenology from VIIRS and HLS time series. *ISPRS Journal of Photogrammetry and Remote Sensing*, *161*(August 2019), 37–51. https://doi.org/10.1016/j.isprsjprs.2020.01.012

Zhou, L., Tucker, C. J., Kaufmann, R. K., Slayback, D., Shabanov, N. V., & Myneni, R. B. (2001). Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. *Journal of Geophysical Research Atmospheres*, *106*(D17), 20069–20083. https://doi.org/10.1029/2000JD000115

Zhou, Q., Rover, J., Brown, J., Worstell, B., Howard, D., Wu, Z., Gallant, A., Rundquist, B., & Burke, M. (2019). Monitoring Landscape Dynamics in Central U.S. Grasslands with Harmonized Landsat-8 and Sentinel-2 Time Series Data. *Remote Sensing*, *11*(3), 328. https://doi.org/10.3390/rs11030328

Zhu, W., Tian, H., Xu, X., Pan, Y., Chen, G., & Lin, W. (2012). Extension of the growing season due to delayed autumn over mid and high latitudes in North America during 1982-2006. *Global Ecology and Biogeography*, *21*(2), 260–271. https://doi.org/10.1111/j.1466-8238.2011.00675.x

Zhu, X., Helmer, E. H., Gao, F., Liu, D., Chen, J., & Lefsky, M. A. (2016). A flexible spatiotemporal method for fusing satellite images with different resolutions. *Remote Sensing of Environment*, *172*, 165–177. https://doi.org/10.1016/j.rse.2015.11.016