Research Article

Analysis of ET and NDVI Correlation in Different Land Cover Types of Colorado

Zhibin Sun[1,2], Runqing Liu[2], Kate Laidlaw[1,3], Maosi Chen[1] and Wei Gao[1,3,*]

[1] United States Department of Agriculture UV-B Monitoring and Research Program, Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, Colorado 80521 USA; zhibin.sun@colostate.edu (Z.S.), kate.laidlaw@colostate.edu (K.L.), maosi.chen@colostate.edu (M.C.), wei.gao@colostate.edu (W.G.)

[2] School of Mathematical Sciences, Ministry of Education Key Laboratory of NSLSCS, Nanjing Normal University, Nanjing, Jiangsu 210023, China; zhibin.sun@njnu.edu.cn (Z.S.), 220902056@njnu.edu.cn (R.L.)

[3] Department of Ecosystem Science and Sustainability, Colorado State University, Fort Collins, Colorado 80523 USA; kate.laidlaw@colostate.edu (K.L.), wei.gao@colostate.edu (W.G.)

[*] Corresponding Author: wei.gao@colostate.edu; +1-970-491-3609.

Academic Editor: Carter Wang
Received: 21 May 2025
Revised: 28 July 2025; 30 July 2025
Accepted: 31 July 2025
Published: 24 October 2025

Abstract: Climate change and water scarcity are pressing public concerns, severely impacting agricultural production and ecological balance. Leveraging advanced remote sensing, this study analyzed 30m-resolution Landsat normalized difference vegetation index (NDVI) and evapotranspiration (ET) data during Colorado's growing seasons (April–August) from 2000 to 2018. Examining correlations across various land cover types, it detailed data processing, relation analysis at different lead times, and correlation classification. Results show a mostly positive correlation in Colorado, aligning with coarser-resolution studies, with some negative exceptions. Uniquely, ET's influence on NDVI weakens as lead time extends, regardless of correlation. Combining with land cover analysis, forests with high water storage maintain stable ET. Additionally, human intervention significantly affects these correlations.

Keywords: Landsat, NDVI, ET, land cover, Colorado

1. Introduction

Plant production is a function of many variables. Besides obvious factors like sunlight and precipitation, plant growth is also affected by other variables like soil moisture, surface temperature, and evapotranspiration. Plants of all species can exhibit metabolic responses to extremes in moisture and even temperature, not just sunlight. Briggs *et al*. (1995) highlighted that the majority of interannual variations in net primary productivity (NPP) cannot be explained solely by a single factor, such as precipitation. This finding aligns with the concept that the NPP patterns in high grassland ecosystems are a result of both spatial and temporal fluctuations in light, water, and nutrients.

The normalized difference vegetation index (NDVI) is a widely used metric for assessing Earth's vegetation (Crow *et al*., 2012). It is robust, versatile, and easy to interpret. Regardless of vegetation type, NDVI shows a strong correlation with ground observations. It is versatile across various fields. NDVI values range from −1 to 1, with 0 representing a vegetation-free region. As vegetation abundance and vitality increase, NDVI increases over time or space. A maximum NDVI of 1.0 indicates dense, healthy vegetation. Values below zero indicate the absence of terrestrial vegetation, and a lake or ocean exhibits an NDVI of −1.

The NDVI exhibits a correlation with vegetation biomass, making it a useful proxy for assessing habitat quality across a wide range of species where ground-based observational data is lacking. However, the accuracy of NDVI may be limited in sparsely vegetated arid and semi-arid environments due to signal contamination from substrate reflectance (Formica *et al*., 2017). Numerous studies have indicated a linear relationship between NDVI and NPP or aboveground net primary productivity (ANPP) under specific conditions. Multiple investigations have demonstrated a positive correlation between NDVI derived from AVHRR/NOAA satellite data and biomass or annual ANPP in different geographical regions and ecosystems

Citation: Sun, Z., Liu, R., Laidlaw, K., Chen, M., & Gao, W. (2025). Analysis of ET and NDVI correlation in different land cover types of Colorado. *Journal of Earth Observation and Geospatial Applications, 1*(1), 1–20. DOI: https://doi.org/10.65372/4z6zq224

(Paruelo *et al*., 1997). Strong correlations were observed between NDVI and annual ANPP at five study sites in cumulative growing season (i.e., from May to September), suggesting that NDVI can serve as an alternative indicator for ANPP. In the Great Plains region of the United States, the correlation between county-level accumulated actual evapotranspiration (AET) from April to July, and ANPP was higher than that with precipitation, especially in dry areas. After considering the ratio of AET to potential evapotranspiration (PET), the correlation was improved with integrated NDVI (iNDVI) and county-level NRCS ANPP (Chen *et al*., 2019). In addition, Zeng *et al*. (2013) has found that precipitation and temperature are powerful controlling factors for NDVI variability, exhibiting distinct spatial patterns.

Evapotranspiration (ET), a variable influencing plant growth and crop production, is poorly understood and difficult to calculate directly. It is the combination of two key parts of water cycle: evaporation and transpiration. ET is the conversion of liquid water to water vapor from land surfaces, water body surfaces, and plant stomata, so it represents a loss of usable water from plants. Therefore, high ET rates are a limit on plant growth. Variables contributing to ET include latent heat flux, sensible heat flux, ground heat flux, relative humidity, wind, temperature, and solar radiation (Rocha *et al*., 2020). Areas with high ET are mainly arid, sunny, and windy, such as the Great Plains of the United States and entire Colorado. ET is a key variable of the hydrologic cycle and agricultural water demand, especially in arid and semi-arid places like Colorado.

In the Great Plains region of the United States, the cumulative AET and transpiration (Tr) from April to July are closely associated with annual ANPP, particularly in arid areas, where AET and Tr show a stronger correlation with ANPP compared with precipitation. In more humid regions, the accumulated precipitation from April to July exhibits the highest interannual variation in relation to ANPP (Chen *et al*., 2019). It has been found that late July is the optimal period for predicting ANPP using NDVI, with a Pearson correlation coefficient of 0.86. A linear model called TAM (tallgrass ANPP model) was developed, confirming the linear association between ANPP and NDVI (An *et al*., 2013). In addition, by studying the relationship between various variables related to water changes and plant growth, one can gain a deeper insight into the consequences arising from water scarcity and climate change (Mukheibir *et al*., 2010), thereby formulating more efficient agricultural development strategies (Nikolaou *et al*., 2020).

Grassland ANPP is strongly influenced by the amount and distribution of annual precipitation (PPT) (Sala *et al*., 1988). The magnitude and direction of precipitation changes associated with climate change are highly uncertain, leading to uncertainties in ecosystem responses. Periods with above or below-average rainfall over multiple years may indicate consequences of changing precipitation conditions. Long-term data on ANPP and PPT were compiled for eight North American grasslands, quantifying the relationships between ANPP and PPT during the periods of above and below-average rainfall, applicable to temperate, semi-arid cool, and semi-arid warm grassland types (Petrie *et al*., 2018). On the other hand, understanding the dynamics of the global carbon cycle relies on the functional relationship between PPT and NPP (Wilcox *et al*., 2016).

As a categorical variable, land cover has a significant impact on ET and NDVI. Accurate information regarding land cover is crucial for the study of global change. In the past decade, data sources and methods for creating global land cover maps through remote sensing have rapidly advanced. The National Land Cover Database (NLCD) is an extensive resource that provides comprehensive nationwide data on land cover and land cover change in the United States. It utilizes a 30-meter resolution and follows a 16-class legend based on the modified Anderson Level II classification system. The version of NLCD used in this study (i.e., NLCD 2019) focuses on delivering innovative information on land cover and land cover change for the entire country (Dewitz *et al*., 2021; Wickham *et al*., 2021; Homer *et al*., 2020; Jin *et al*., 2019; Yang *et al*., 2018). It includes eight integrated epochs of land cover data, spanning the years of 2001, 2004, 2006, 2008, 2011, 2013, 2016, and 2019. The developed classes for these years are derived directly from the percentage of developed impervious surface and are accompanied by descriptor labels that specify the type of impervious surface found in each image pixel. Appendix A provides specific information on land cover in Colorado.

Moreover, research has revealed that global water resources are significantly impacted by human interventions (Haddeland *et al*., 2014). For example, human activities led to an 11%–16% increase in evapotranspiration over the Yellow River between 2003 and 2010 (Zhang *et al*., 2020).

The Great Plains region of the United States serves as a vital agricultural production hub in the global market and constitutes a significant source of greenhouse gas emissions (Parton *et al*., 2015). In spite of its arid and highly variable climate, Colorado is a key state for agricultural production in the United States. Livestock, especially cattle along with alfalfa, sugar beets, corn, melons, peaches, apples, and wheat are the most important parts of Colorado's agriculture industry (USDA/NASS, 2021). However, drought is one of the most devastating afflictions on agricultural success, sustainability, and production in Colorado (Bauman *et al*., 2013). Drought conditions persist almost continually throughout the state, with scarce periods where

some regions are not affected to some extent (McKee *et al*., 2000). Droughts are usually accompanied by extremely high temperatures (Hood *et al*., 2020), adding stresses to plants. The changing climate is likely to decrease water availability and agricultural yields in Colorado (What Climate Change Means for Colorado, 2016). As a result, this study was undertaken in order to investigate what and if any effects ET have on agricultural plant growth in Colorado.

To investigate the impact of ET or cumulative ET (cu ET) on NDVI and its variation across different land types, thereby shedding light on the close relationship between water availability and agriculture, this study incorporates three variables: ET, NDVI, and land cover.

Long-term datasets, which are essential for advancing research in ecology and guiding agricultural production, offer vital information for ecosystem management and evidence-based policymaking at appropriate scales (Lindenmayer *et al*., 2012). For example, Joiner *et al*. (2018) utilized long-term ET and NDVI data from 2003 to 2016 to conclude that ET generally exhibits a quicker response to water deficits and enhancements compared with NDVI. Using a similar length of the above time span, this study explores the relationship between ET and NDVI across diverse land cover types in the region within the growing seasons from April to August between 2000 and 2018.

ET and NDVI are hard to measure since it depends on so many variables that can extremely change over spatiotemporal scales. Additionally, monitoring any hydro-meteorological in-situ variable requires expensive sensors and frequent maintenance. Consequently, satellite remote sensing has emerged as an effective way to obtain these kinds of data worldwide, fill in data coverage gaps since sensors cannot be installed everywhere, and calibrate in-situ physical sensors. National Aeronautics and Space Administration (NASA) and United States Geologic Survey's (USGS) Landsat mission is the longest-lasting continuous remote sensing survey of the Earth's surface. It provides trustworthy geospatial raster data in the form of tagged image file format (TIFF or TIF) at fine resolutions with extensive, informative metadata at the levels of understanding ranging from professional to amateur. These factors make Landsat data trustworthy and frequently used in geographic information systems (GIS). In addition, the relationship between NDVI and ANPP is linear for the data obtained from NOAA/AVHRR and Landsat TM (Paruelo *et al*., 2000).

2. Data and Methodology

In this study, ET and NDVI data are sourced from NASA's USGS Landsat mission and are treated as temporally continuous variables, renowned for its impressive 30m resolution. There are some limitations for both data. For example, (1) there are missing values in some areas; (2) some pixels are contamination by clouds/rain/snow, and (3) there are overlapping pixels at scene boundaries for the same acquisition date. We selected the data sets spanning from the year 2000 to 2018, specifically focusing on the growing seasons of each year, from April to August. The land cover data, obtained from the National Land Cover Database (NLCD), encompasses the entire state of Colorado and comprises only 16 land cover types. Consequently, the land cover data is treated as a categorical variable with 16 categories, as presented in Table A1 in Appendix A. The specific information for all variables is detailed in Table 1. The download process for all variables started on EarthExplorer with setting Colorado as the predefined area.

Table 1. Data sources and resolution characteristics of satellite products.

Variable	Dataset	Data Source/ Landsat Mission/ Sensor	Spatial Resolution	Temporal Resolution
ET	Landsat C2 Level-2 (L2)	Landsat 7 and 8 (ETM+, OLI/TIRS) C2 L2	30m	about half monthly
NDVI	Landsat C2 Level-2 (L2)	Landsat 7 and 8 (ETM+, OLI/TIRS) C2 L2	30m	about half monthly
Land Cover	Landsat Collection 2 (C2) U.S. Analysis Ready Data (ARD)	Landsat 7 and 8 (ETM+, OLI/TIRS) C2 L2	30m	2001, 2004, 2006, 2008, 2011, 2013, 2016, 2019

Each variable exhibits distinct spatial and temporal distributions, reflecting changes over space and time. To uncover the relationships among these variables, it is necessary to firstly process the downloaded raw remote sensing data. The data processing techniques, as shown in Figure 1, take into account the temporal and spatial characteristics of the raw data, enabling the establishment of the linear relationship between ET

and NDVI. This processing facilitates the analysis of correlations and potential influences across different land covers. Initially, the linear relationships are examined following spatial and temporal alignment and interpolation. Moreover, specific land types are selected to study the correlation between ET and NDVI, allowing for a focused analysis of their relationship within a specific area. All programs were executed using MATLAB R2022a on two Intel Xeon Silver 4215 CPUs. The computer has 128GB of memory.

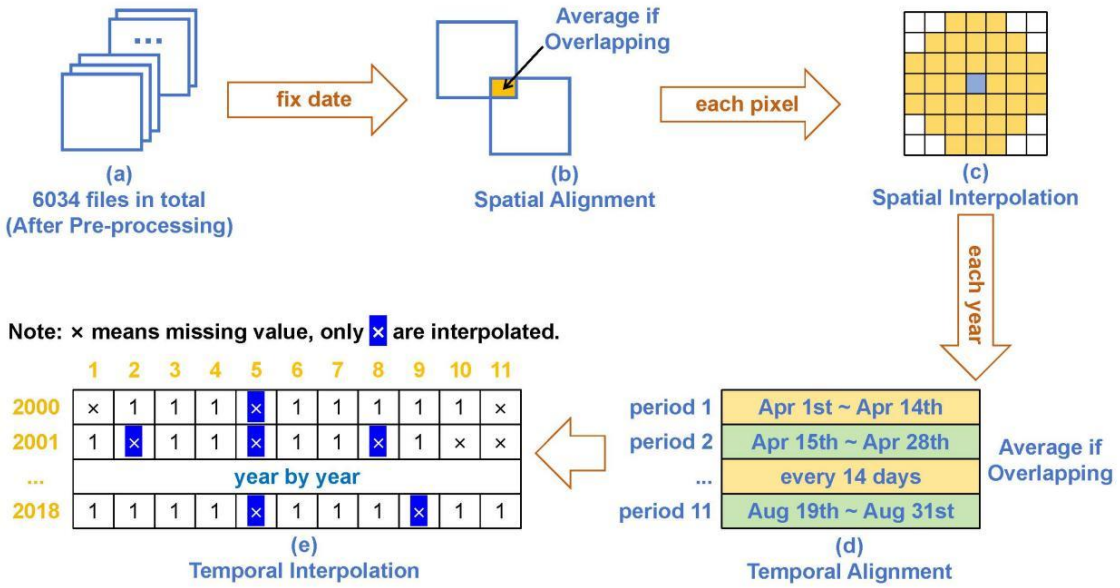

Figure 1. Data Processing Workflow. (a) Preprocess the downloaded data and convert a ".tif" file into ". mat" format compatible with MATLAB. (b) When a fixed date is considered, there may be multiple valid values in a pixel within the yellow area. The average value is taken as the value at that pixel. (c) Illustration of Spatial Interpolation. The blue square represents a pixel requiring interpolation, indicating its original missing values. The yellow squares represent the area involved in interpolating the pixel at the blue square. Each pixel has a resolution of 30 meters, and these yellow squares encompass a radius of 100 meters around the pixel requiring interpolation. (d) From April 1st to August 31st, there are a total of 153 days, divided into 11 periods of 14 days each, starting from April 1st. This may result in a pixel having multiple values within the same period. The average of these values is taken as the value of the pixel for that period. (e) Linear interpolation is performed on an annual basis, excluding the missing values only during the periods between years, as the last period of one year and the first period of the following year have a substantial time gap.

The detailed processing method represented by Figure 1 is outlined as follows:

1) Pre-processing (Figure 1a)
   - 1.1) Remove pixel values affected by cloud, cloud shadow, and snow from the QA_pixel dataset (in Landsat product) and set them to NaN.
   - 1.2) Based on the vertex coordinates in the source data, delineate the areas captured by the satellite images, primarily to distinguish NaN values originating from blank edges or remote sensing algorithms.
2) Spatial Alignment (Figure 1b)
   - 2.1) Extract dates and UTM projection zones from 6034 files, categorize them, and generate all possible combinations of dates and zones.
   - 2.2) Read all combinations, iterate through each spatial coordinate in the zones, and examine the uniqueness of variable values within all files in that combination.
   - 2.3) If all values at the same coordinate across files are NaN, assign NaN as the value for that coordinate. Otherwise, calculate the average of all non-NaN values at the coordinate as the pixel value.
   - 2.4) Based on the combinations of dates and zones, derive the final outcome associated with dates and zones.

3) Spatial Interpolation (Figure 1c)
   3.1) Identify the date and zone, read the results from the previous step, and perform spatial interpolation.
   3.2) As depicted in Figure 1c, for each pixel (blue pixel), compute the average of all valid values (i.e., yellow pixels) within a circle of 90-meter radius.
   3.3) If all values at the same coordinate across files are NaN, assign NaN as the value for that coordinate. Otherwise, calculate the average of all non-NaN values as the pixel value at that location.
   3.4) The interpolated results are linked to the date and zone.
4) Temporal Alignment (Figure 1d)
   4.1) Divide each year from April to August into periods of two weeks (14 days), resulting in 11 periods per year.
   4.2) Determine the year and period, and retrieve all files within the specified time range.
   4.3) If all values at the same coordinate across files are NaN, assign NaN as the value for that coordinate. Otherwise, calculate the average of all non-NaN values as the pixel value at that location.
   4.4) Iterate through different years and periods to obtain final results.
5) Temporal Interpolation (Figure 1e)
   5.1) Fix the year and spatial coordinates, examine the values of the corresponding pixels across 11 periods, sorted chronologically.
   5.2) For each period, perform linear interpolation if non-NaN values are present both before and after that period.
   5.3) Ultimately, results are obtained for different years and periods.

2.1. Preprocessing of Raw Data

The datasets preprocessed in this study consist of two variables: ET and NDVI. Both variables exhibit the same spatial and temporal distribution. Therefore, in order to explain the data processing methodology, we will focus on elucidating the approach using the ET variable. Our data processing approach follows a two-step principle: first, focusing on maximizing the utilization of raw data (i.e., fine-grained processing), and then employing a coarser approach that involves merging overlapping data and performing interpolation (i.e., rough-grained processing). The data processing is carried out in a sequential order, prioritizing spatial considerations before temporal aspects.

Taking the ET variable as an example, we obtained a total of 6034 products from the EarthExplorer website (https://earthexplorer.usgs.gov/) based on the metadata. These products covered the period from 2000 to 2018, specifically from April to August of each year, and were located in different spatial and temporal contexts. The downloaded products had a resolution of 30m and provided information about variable value, pixel quality, and coordinate information for each pixel. To prepare the raw data for further analysis, we performed preprocessing, which involved two steps: removing pixels with bad quality and annotating the valid measurement areas.

In the raw data, certain pixels were affected by factors such as clouds, cloud shadows, and snow, resulting in poor quality. To address this, we utilized the QA_pixel product, where "QA" stands for Quality Assessment. By referring to this product, we identified and removed pixels that were affected by cloud, cloud shadow, and snow. This process involved replacing the values of these pixels with missing values, effectively excluding them from subsequent analyses.

Additionally, we observed that the actual measurement areas of the products differed from the entire represented regions. Upon visual examination of products, we identified missing values within the actual measurement area, which corresponds to a rectangular-like region in the image, as well as the region outside the actual measurement area. We refer to the actual measurement area as the valid region, while the rest is considered invalid. During subsequent spatial interpolation, it is crucial to exclude the invalid regions from the interpolation process. Therefore, our first step was to annotate the valid and invalid regions of the products. The metadata files provided the latitude and longitude coordinates of the vertices that defined the actual measurement area for each product. By utilizing these coordinates, we constructed a quadrilateral in the latitude and longitude coordinate system and labeled the boundaries and interior of the quadrilateral as the valid region, while the exterior was labeled as the invalid region. This labeling process is illustrated in Figure 2, where the left image represents the annotated result, with the yellow color indicating the valid region and

the blue color representing the invalid region. The consistency between the annotated valid region and the right image confirms the accuracy of our labeling approach.

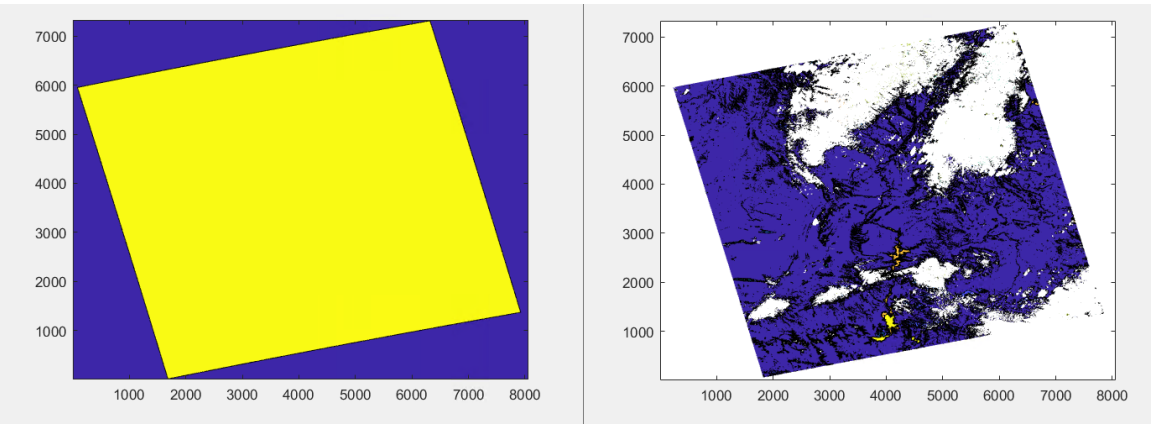

Figure 2. Illustration of the annotated valid region for a specific .tif product. An image from a randomly selected tile on a particular day was chosen as an example for annotating the valid region. In the left image, the yellow area represents the valid region, while the blue area represents the invalid region. The right image shows the white area indicating missing values, and it is noticeable that there is a substantial block of missing values outside the edges of the valid region.

2.2. Spatial Alignment and Interpolation

Within a given day, a product may consist of multiple tiles, and upon comparison, overlapping areas between adjacent tiles can be identified. In these overlapping regions, pixels representing the same geographic coordinates may possess values derived from at least two distinct tiles. On a given date, variations in the time of image acquisition (including hour, minute, and second) and disparities resulting from the utilization of different algorithms to calculate satellite variables using diverse raw data can potentially lead to multiple distinct values for the same variable at a specific pixel within the same day. As shown in Figure 1b, disregarding invalid values, we compute the average of these divergent data values to determine the pixel value, thereby ensuring a more representative and reliable estimate of the variable's value for further analysis and interpretation.

Upon processing the overlapping areas, it can be observed from Figure 2 that numerous missing values exist within valid regions. To ensure the data reflects the overall characteristics of the entire area, spatial interpolation is conducted on the regions with missing values. Let $X_{date} = \left(x_{date}^{(location)}\right)$ represents all variable values on the same date, with location fixed. Let $N^{(location)} = \{neighbour \mid 0 < distance(location, neighbour) < 100\ \text{m}\}$ denotes all locations within 100 meters of the given location. For each location, compute

$$y_{date}^{(location)} = mean\,\{x_{date}^{(neighbour)} \mid neighbour \in N^{(location)}\}, \tag{1}$$

resulting in $Y_{date} = \left(y_{date}^{(location)}\right)$. Let $M_{date} = \left(m_{date}^{(location)}\right)$ be the mask for X_{date}: if there is a value at $x_{date}^{(location)}$, then $m_{date}^{(location)} = 1$; otherwise, $m_{date}^{(location)} = 0$.

After interpolation, the interpolated X^* is given by Equation (2) where $\odot$ represents the Kronecker product.

$$X^* = X + Y \odot (1 - M) \tag{2}$$

As depicted in Figure 1c, we perform interpolation at the center blue pixel by taking the average of all valid values (i.e., yellow pixels) within a radius of 100 meters. This interpolation expands the area with available values while preserving the similarity of values in neighboring regions. During the interpolation process, only the data after handling the overlapping areas is utilized, and no additional spatial interpolation

data is employed. Consequently, some missing values still exist within the valid regions after the completion of spatial interpolation. Due to inadequate satellite data and models for computing these missing values, we consider them as genuinely missing and therefore do not subject them to further spatial interpolation.

2.3. Temporal Alignment and Interpolation

As depicted in Figure 1d, considering an average of approximately two data points per month for each pixel, we conducted temporal interpolation by partitioning the time into discrete 14-day periods and consolidating data from different dates within the same period. The analysis focused on utilizing data from April to August, encompassing the years from 2000 to 2018. Specifically, each year from April 1st to August 31st, totaling 153 days, was divided into 11 periods of 14 days. After considering valid values, the pixel value was assigned as the mean value across different dates within the same period.

Then, we performed linear interpolation for each of the 11 periods, progressing year by year. Let $X_{year}^{(location)} = \left(x_{year,period}^{(location)}\right)$ denotes all variables for a certain location within a given year across all periods. For each location and year, linear interpolation was applied from the first non-missing period to the last valid period. If $x_{year,period}^{(location)}$ was missing, interpolation was performed via

$$x_{year,period}^{(location)} = (1-\alpha)x_{year,period_0}^{(location)} + \alpha x_{year,period_1}^{(location)}, \quad (3)$$

where $\alpha = \frac{period - period_0}{period_1 - period_0}$. Here, $period_0$ and $period_1$ denote the periods to which the nearest valid values preceding and following the missing value belong, respectively. As illustrated in Figure 1e, linear interpolation was applied for areas where the first and last valid values of each year were present. However, no interpolation was conducted for the regions where there were no valid values at the beginning or end of the time period.

2.4. ET Control on NDVI

Previous studies have identified the controlling effect of precipitation on NDVI (Zeng *et al*., 2013), indicating a potential lagged relationship between precipitation and NDVI (Joiner *et al*., 2018). To further investigate this mechanism, this study introduces a variable called Lead, varying from 0 to 9. This involves comparing ET data and NDVI data within same periods, with ET data leading by Lead periods respectively. Pearson correlation coefficients are calculated, and the proportion of significant correlations ($p<0.05$) is determined.

After performing spatial and temporal alignment and interpolation on the data, for each pair of longitude and latitude, we obtain a column of time-related data $\{x_{year,period}\}$, where x can represent ET or NDVI, as depicted in Figure 3. It is important to note that when calculating the correlation coefficients and their corresponding p-values, the specific dates corresponding to the periods in this study are not continuous. For

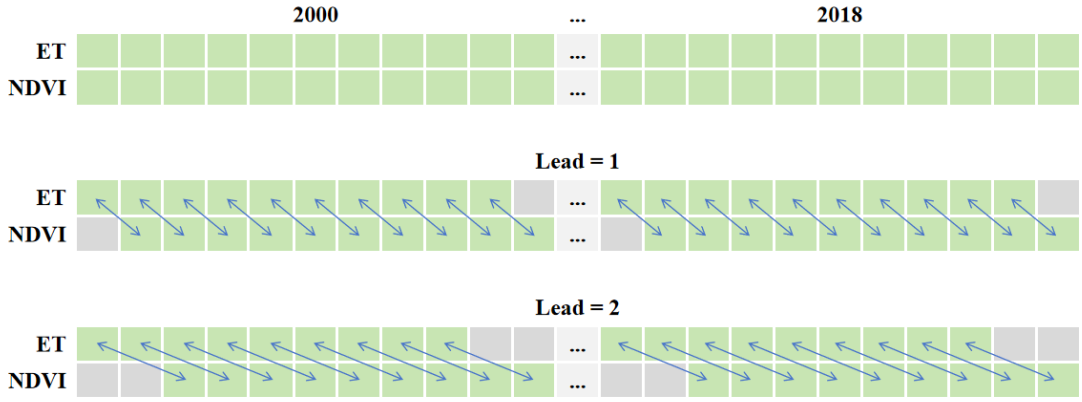

Figure 3. Illustration of calculating correlation coefficients at different Lead levels. The green squares represent the data utilized in the analysis, while the gray squares indicate the data that were not included. The bidirectional arrows demonstrate how ET and NDVI data from different time points correspond to each other.

instance, the last period of 2001 ends on August 31st, while the first period of 2002 begins on April 1st. Given this discontinuity, we do not consider the preceding period of the first period in 2002 as the last period of 2001, but rather replace it with missing values. Additionally, when calculating the correlation coefficients, the calculation is performed only when both the leading ET and NDVI values are non-missing. Introducing the variable Lead, we define the following two sequences:

$$E_{Lead} = \{ET_{year,period} \mid period = 1, 2, \ldots, 11 - Lead\}, \quad (4)$$

$$N_{Lead} = \{NDVI_{year,period} \mid period = Lead + 1, Lead + 2, \ldots, 11\}. \quad (5)$$

The correlation between E_{Lead} and N_{Lead} at different Lead levels is evaluated using Pearson correlation coefficients and their corresponding p-values.

2.5. Cumulative ET Control on NDVI

Previous studies have demonstrated the controlling effect of cumulative precipitation on NDVI (Zeng *et al*., 2013; Koster *et al*., 2014). To investigate this mechanism, this study also introduces a variable called CuLead, varying from 0 to 9. In this context, Lead refers to the cumulative ET data obtained by summing the ET values from the preceding CuLead periods up to the current period. This cumulative ET data is then compared with the NDVI data of the current period, as illustrated in Figure 4. To examine whether the NDVI during the growing season (i.e., from late July to August) is influenced by cumulative precipitation, this study selects the last two periods of each year for NDVI data. These NDVI data points are paired with their corresponding cumulative ET data, and Pearson correlation coefficients and p-values are calculated for analysis. Introducing the variable CuLead, we define the following two sequences:

$$E_{CuLead} = \{\sum_{period=n-CuLead}^{n} ET_{year,period} \mid n = 10, 11\}, \quad (6)$$

$$N_{CuLead} = \{NDVI_{year,period} \mid period = 10, 11\}. \quad (7)$$

The correlation between E_{CuLead} and N_{CuLead} at different CuLead levels is evaluated using Pearson correlation coefficients and their corresponding p-values.

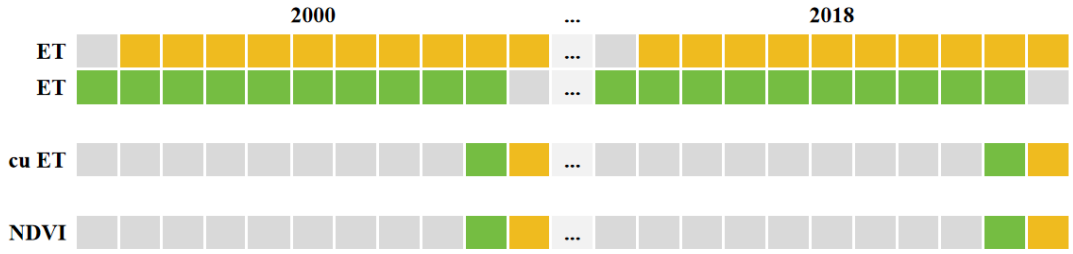

Figure 4. Illustration of calculating correlation coefficients at CuLead = 9 level. The yellow and green squares representing cu ET for each year are obtained by summing the yellow and green squares of ET, respectively. These cu ET squares are then matched with the corresponding yellow and green squares of NDVI to calculate the correlation coefficients.

3. Results

This presenting study consists of three main parts: the influence of individual monthly ET on NDVI, the influence of cumulative ET over multiple months on NDVI, and the impact of land cover on the correlation coefficient between ET and NDVI.

3.1. Influence of ET on NDVI

Individual monthly ET may exert control on NDVI. The control of ET on NDVI could exhibit a lag effect, differing across different regions. To investigate this mechanism, we separately examined the control of ET on NDVI, with ET leading by different periods.

Through observing Figure 5a, it can be noticed that the majority of regions in the image exhibit significant positive correlation, while a few regions show significant negative correlation. Building upon this observation, approximately 92% of the regions display positive correlation, with approximately 86% of the regions exhibiting statistical significance. Notably, the correlation appears to be less stable in the western regions of Colorado, whereas the eastern regions demonstrate more pronounced and consistent positive correlation. In Section 3.3., we will introduce the variable land cover to explore the reasons behind the negative correlation and lack of correlation in certain regions.

Upon examining the comparison between the left and right pictures in Figure 5, a clear pattern emerges: the regions with stronger positive or negative correlation exhibit a more significant correlation. Figure 6 provides a detailed depiction of the distribution of correlation coefficients (Lead=0) between ET and NDVI in Colorado. From the graph, it is evident that when the correlation coefficient is near zero, it is predominantly non-significant. Conversely, when the correlation coefficient deviates from zero, it tends to be statistically significant. This result suggests that the relationship between the two variables in these regions is non-random and likely exhibits a systematic connection. On the other hand, for the regions with lower correlation, various factors may influence the relationship, resulting in a larger p-value and non-significant correlation between the variables. To investigate this phenomenon, it is necessary to introduce additional variables and conduct further research to uncover the relationships among these variables and identify generalized patterns.

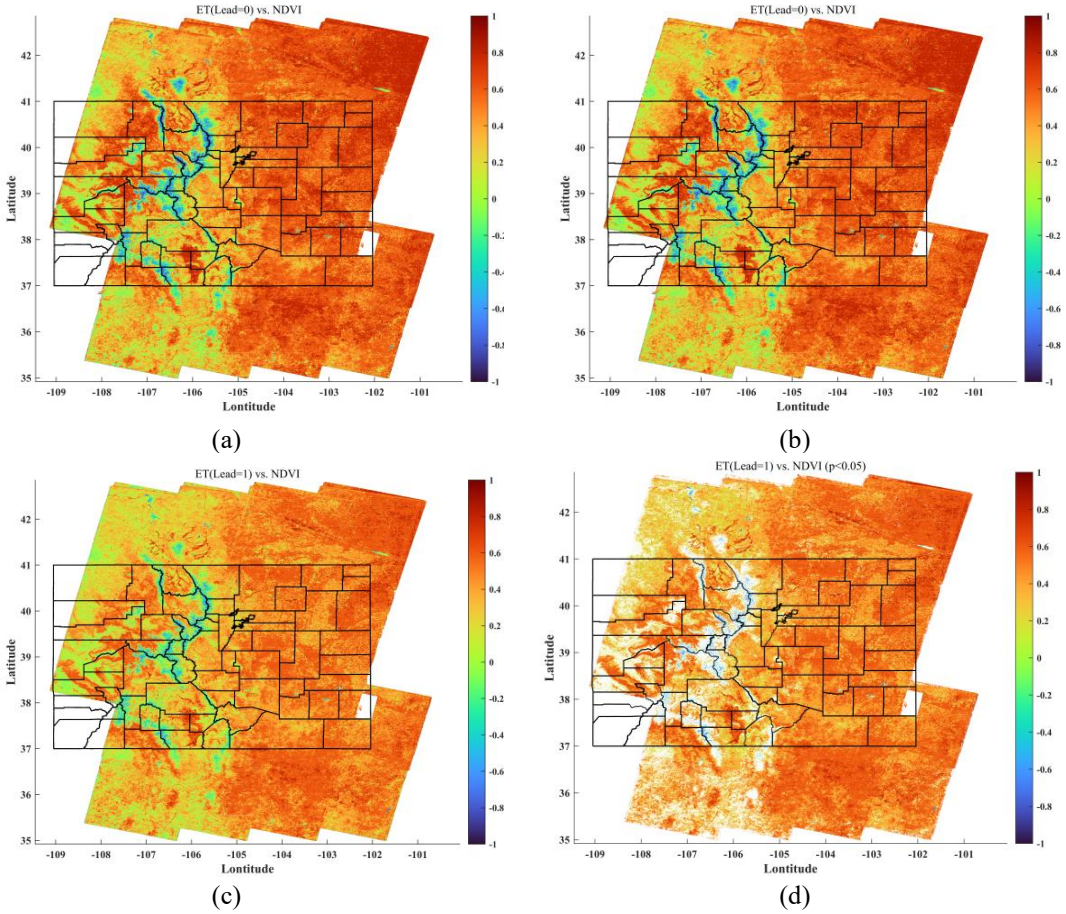

Figure 5. Pearson correlation coefficients between ET and NDVI in Colorado. The black box represents the state of Colorado and its administrative regions. Panels (a) and (c) represent the Pearson correlation coefficients for ET and NDVI at lead times of 0 and 1, respectively. Panels (b) and (d) display the Pearson correlation coefficients that are statistically significant ($p<0.05$) between ET and NDVI at lead times of 0 and 1, respectively.

We also discover that for the majority of regions in Colorado, the lagged control of ET on NDVI gradually weakens as the Lead increases, regardless of whether it is positive or negative correlation, through the comparison of results in Figure 5a & 5c. This phenomenon can be attributed to two main factors. Firstly, as

the Lead increases, the replacement of missing values between two-year intervals leads to a reduction in data volume, increasing the uncertainty of the results. Secondly, it could be related to the vegetation types in Colorado, which will be discussed in detail in Section 3.3.

Figure 7 illustrates the proportion of positive and negative correlation coefficients between ET and NDVI, as well as their significance, at different Lead levels within Colorado. Observation of the image reveals that as Lead increases, the proportion of statistically significant data decreases for both positive and negative correlations. Furthermore, by calculating the distance between the probability distribution curves in Figure 6 and other Lead levels, which we refer to as significant distance, Figure 8 is obtained. The significant distance increases with the growth of Lead, indicating a gradual weakening of the lagged control of ET on NDVI. Additionally, Figure 9 shows a decreasing trend in the expected correlation coefficients as Lead increases, further emphasizing the diminishing lagged control of ET on NDVI with increasing Lead.

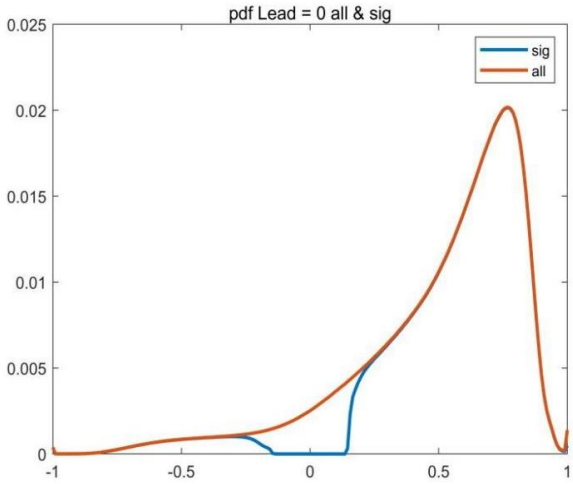

Figure 6. The probability density depicts the correlation between ET and NDVI when Lead=0. The horizontal axis represents the values of the correlation coefficient between ET and NDVI, while the vertical axis represents the proportion of the overall distribution. The upper curve represents the probability distribution of the correlation coefficient, while the lower curve represents the distribution that is statistically significant ($p<0.05$).

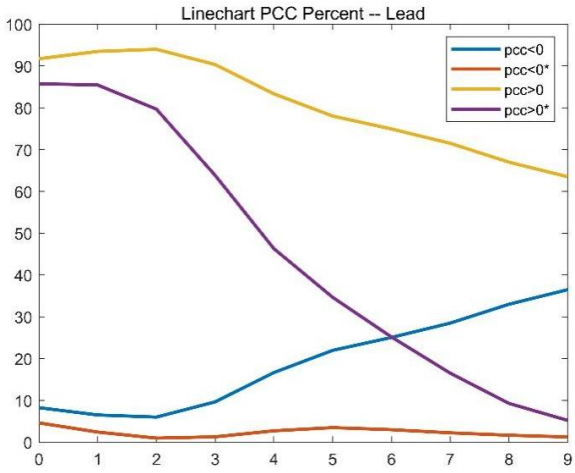

Figure 7. The line graph illustrates the proportion of positive and negative correlation coefficients as Lead increases. The horizontal axis represents Lead, while the vertical axis represents the proportion. The yellow and blue curves represent positive and negative correlation coefficients, respectively, while the purple and orange curves represent their statistically significant occurrences.

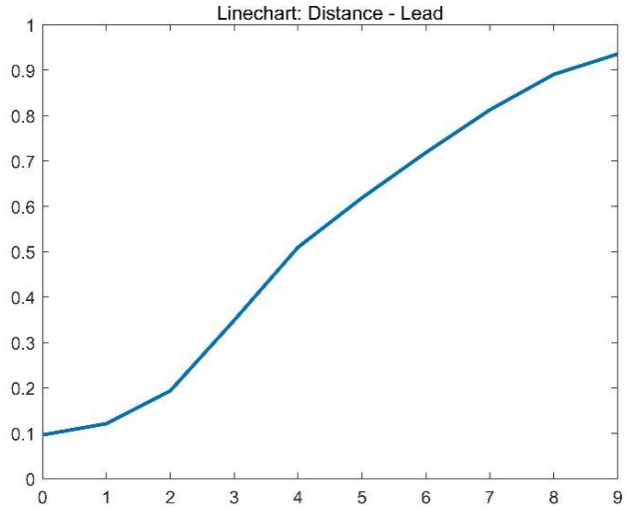

Figure 8. The graph presents the significant distance at different Lead levels. The horizontal axis represents Lead, while the vertical axis represents the significant distance, given by $\text{Distance} = \int_{-1}^{1} \text{p(x)} - \text{p}^{*}\text{(x)dx}$.

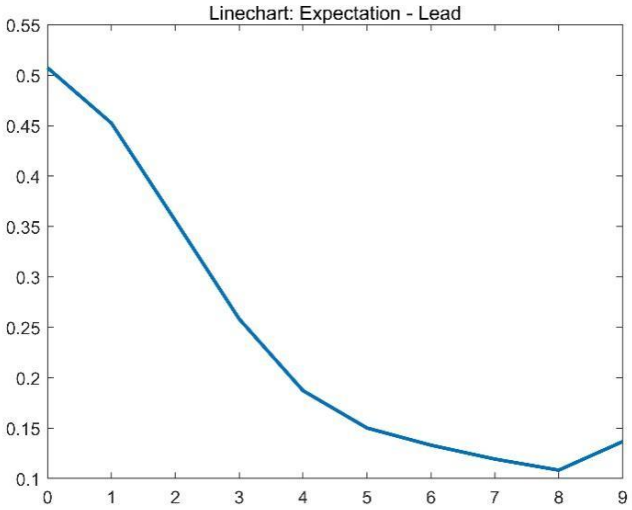

Figure 9. The graph illustrates the expected correlation coefficients at different Lead levels. The horizontal axis represents Lead, while the vertical axis represents the expected correlation, given by $\text{Expectation} = \int_{-1}^{1} \text{x p(x)dx}$.

3.2. Control of Cumulative ET on NDVI

Previous studies have demonstrated the controlling effect of cumulative precipitation on plant growth. In this study, we investigated the relationship between NDVI and cumulative ET during late growing seasons, specifically from late July to August. The NDVI were selected for the last two periods of each year, while the cumulative ET were calculated for the accumulation of ET over from 0 to 9 months. We computed the Pearson correlation coefficient and p-value between NDVI and cumulative ET.

As shown in Figure 10, almost all regions exhibit a positive correlation between cumulative ET and NDVI, indicated by positive correlation coefficients. However, upon comparing the two panels, it can be observed that this positive correlation is not very significant, with only a few regions (36%) showing statistical significance ($p < 0.05$). There are hardly any regions with significant negative correlation (<1%) or lack of correlation. Furthermore, Figure 10 vividly illustrates that the proportion of statistically significant positive correlation coefficients notably decreases compared with Figure 5.

By comparing the upper and lower panels in Figure 10, a slight enhancement in both positive and negative

correlation can be observed as the cumulative ET periods increase from 8 to 9. This phenomenon is further illustrated in Figure 11, which demonstrates the varying influence of cumulative ET on NDVI as the ET accumulation period lengthens. Specifically, it indicates a diminishing effect followed by a strengthening effect. This observation suggests that the NDVI during the late growth season is significantly influenced by both the end of the growth period and the cumulative ET throughout the entire growth season. In other words, the final growth state of the plants is influenced by the entire growth process and recent conditions, while the impact of intermediate processes on plant growth outcomes is relatively minor.

Upon observing Figure 10, it becomes apparent that the regions displaying significant positive correlation are predominantly concentrated in the eastern part of Colorado. Moreover, certain regions exhibit variations in correlation as the cumulative periods increase. Further investigation is required to delve into the mechanisms underlying these fluctuations. These phenomena are closely related to land cover, which will be explored in Section 3.3.

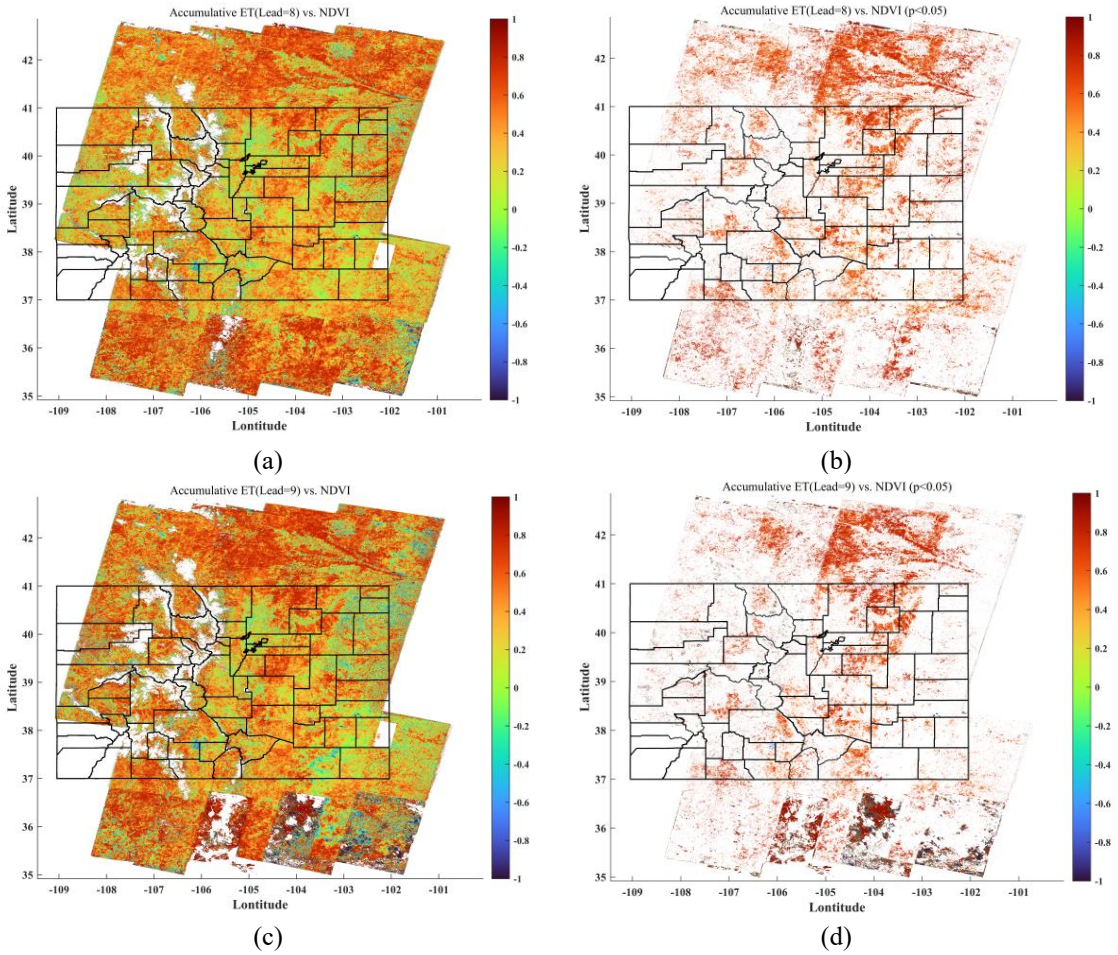

Figure 10. Pearson correlation coefficients between the cumulative ET and NDVI in Colorado. Panels (a) and (c) represent the Pearson correlation coefficients for ET and NDVI at lead times of 8 and 9, respectively. Panels (b) and (d) display the Pearson correlation coefficients that are statistically significant ($p < 0.05$) between cumulative ET and NDVI at lead times of 8 and 9, respectively.

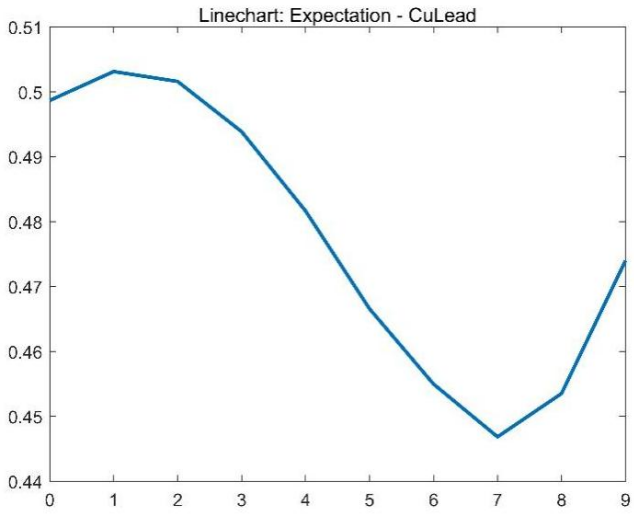

Figure 11. The graph illustrates the expectation of correlation coefficients at different levels of CuLead. The horizontal axis represents CuLead, while the vertical axis represents the expected correlation.

3.3. Impact of Land Cover on the PCC between ET and NDVI

Figure 12 presents a schematic map depicting the land cover distribution in Colorado, where different colors represent distinct land types. The proportions of each land cover category can be observed in Figure 13. By combining these two figures, it becomes evident that the western half of Colorado is predominantly comprised of Evergreen Forest and Deciduous Forest, whereas the eastern half is mainly characterized by Grassland/Herbaceous and Cultivated Crops. Additionally, Shrub/Scrub and Developed Lands also contribute to a portion of the land cover. From Figure 5 and Figure 10, it is apparent that in many areas, ET (or cumulative ET) displays a negative correlation with NDVI. By statistically analyzing the correlation coefficients between ET (or cumulative ET) and NDVI for each land cover category, we can generate Figure 14, Figure 15, and Figure 16.

Based on Figure 14 and Figure 16, it is evident that Perennial Ice/Snow, Barren Land (Rock/Sand/Clay), and Open Water, regardless of ET or cumulative ET, exhibit a predominant negative correlation with NDVI. These three land cover types are characterized by a scarcity of vegetation, resulting in ET primarily driven by water evaporation and minimal or no involvement of plant transpiration. Consequently, the dominant relationship between NDVI and ET is negative. In Figure 15, when Lead=0, the proportion of pixels with significant correlation reveals that Open Water and Barren Land exhibit a similar number of pixels with both positive and negative correlations. This indicates that ET is unrelated to NDVI in these land cover types, suggesting that evaporation is not influenced by vegetation and primarily driven by water evaporation processes.

Both trees and the land they inhabit a higher water storage capacity, allowing for stable and continuous water evaporation and plant transpiration throughout the entire growing season. This indicates that trees possess strong control over water dynamics within their surroundings, including water acquisition, storage, and consumption. As shown in Figure 14, among the remaining land cover categories, Forests, which includes Evergreen Forest, Mixed Forest, and Deciduous Forest, exhibit a higher proportion of negatively correlated regions compared with other land cover types. When considering the cumulative effect of ET on NDVI throughout the growing season, as depicted in Figure 16, the number of pixels with both positive and negative correlations is nearly equal in all Forest categories. This suggests that trees have a stable capacity for water regulation, enabling them to adapt well to environmental changes and provide a relatively consistent water cycle within ecosystems.

On the other hand, human intervention has a significant impact on correlation coefficients. Specifically, land cover categories influenced by human activities include Developed Space with different intensities, Pasture/Hay, and Cultivated Crops. As depicted in Figure 14, due to human intervention, plants can effectively utilize water and grow healthily, with most of the ET originating from plant transpiration. Therefore, there is a high positive correlation at various stages. As shown in Figure 16, due to human intervention, the cumulative ET exhibits a decreased positive correlation with end-of-growing-season NDVI compared with the influence of ET on NDVI. This phenomenon can be attributed to the "marginal effect",

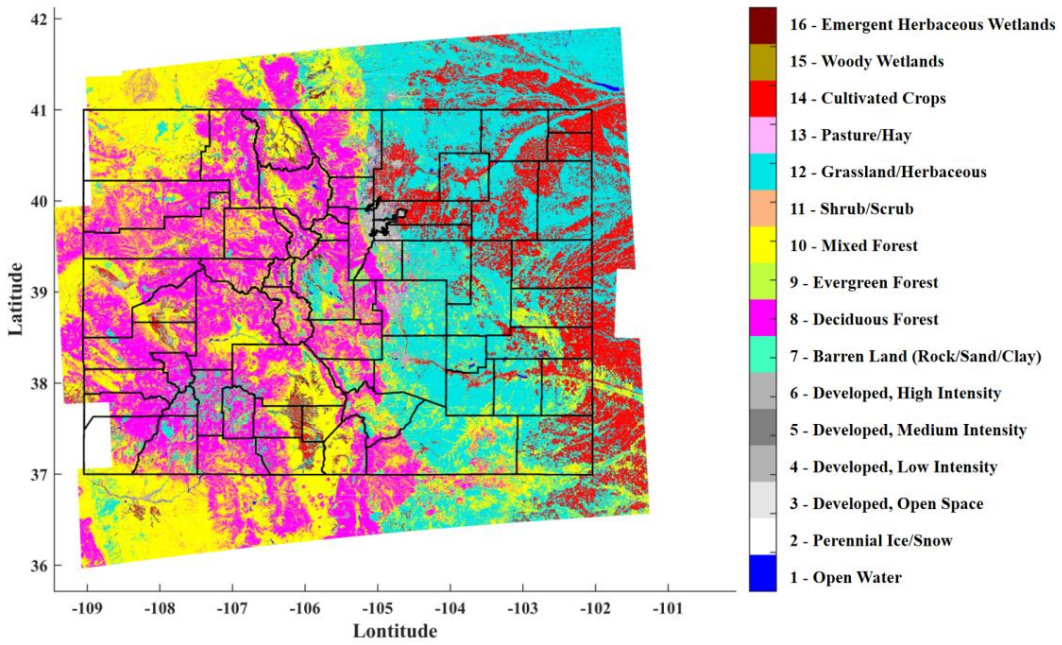

Figure 12. The schematic representation of land cover in Colorado. Land cover is a categorical variable, where integers ranging from 1 to 16 are used to represent the 16 land types in Colorado. The specific meaning of each number can be found in Table A1 of Appendix A, under the Analysis ID column.

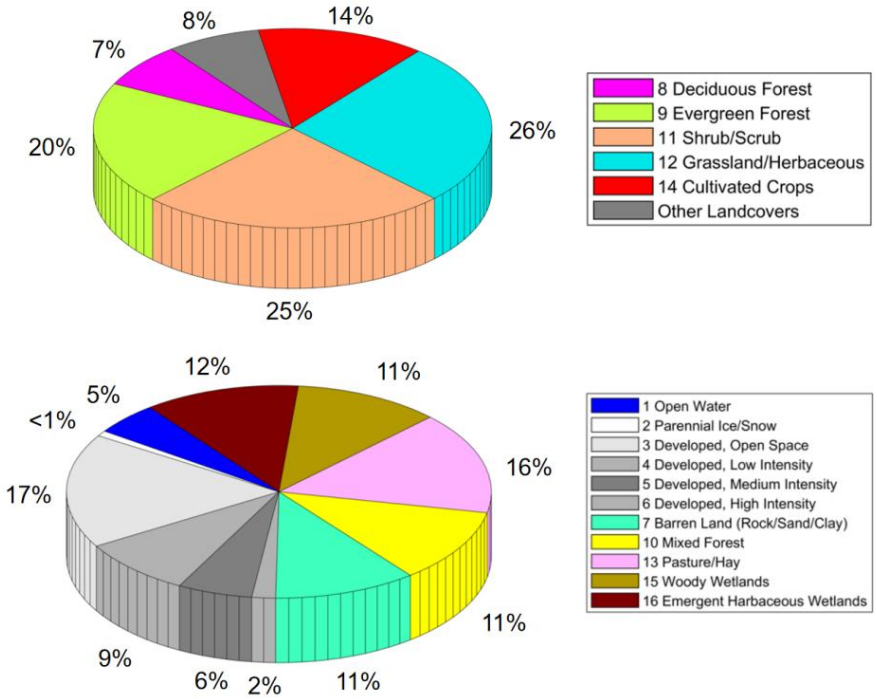

Figure 13. The pie chart illustrates the land cover composition in Colorado state. (Upper panel): It presents the proportion of the top 5 land cover categories with the largest share in Colorado, while the remaining land cover categories, each accounting for less than 1%, are combined into a single category. (Lower panel): It displays the proportion of the remaining land cover categories in Colorado, excluding the top 5 categories mentioned in upper panel.

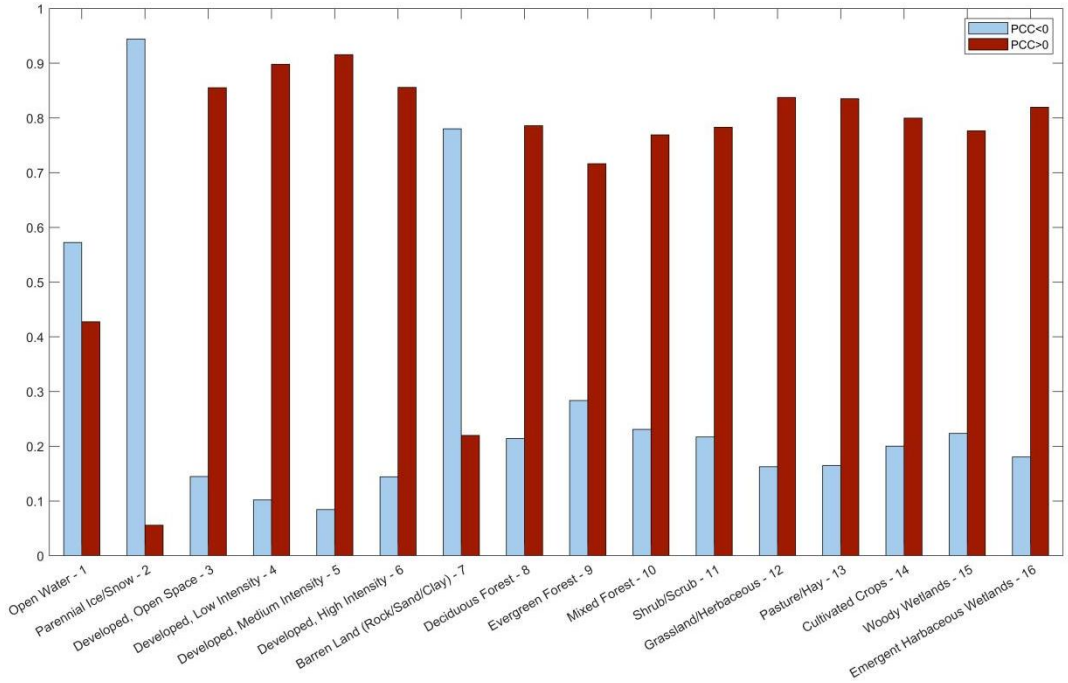

Figure 14. The bar chart illustrates the proportion of positive and negative correlation coefficients for different land cover types when Lead=0. In the chart, the blue bars represent the proportion of pixels with negative correlation coefficients, while the red bars represent the proportion of pixels with positive correlation coefficients. The combined values of the blue and red bars sum up to one.

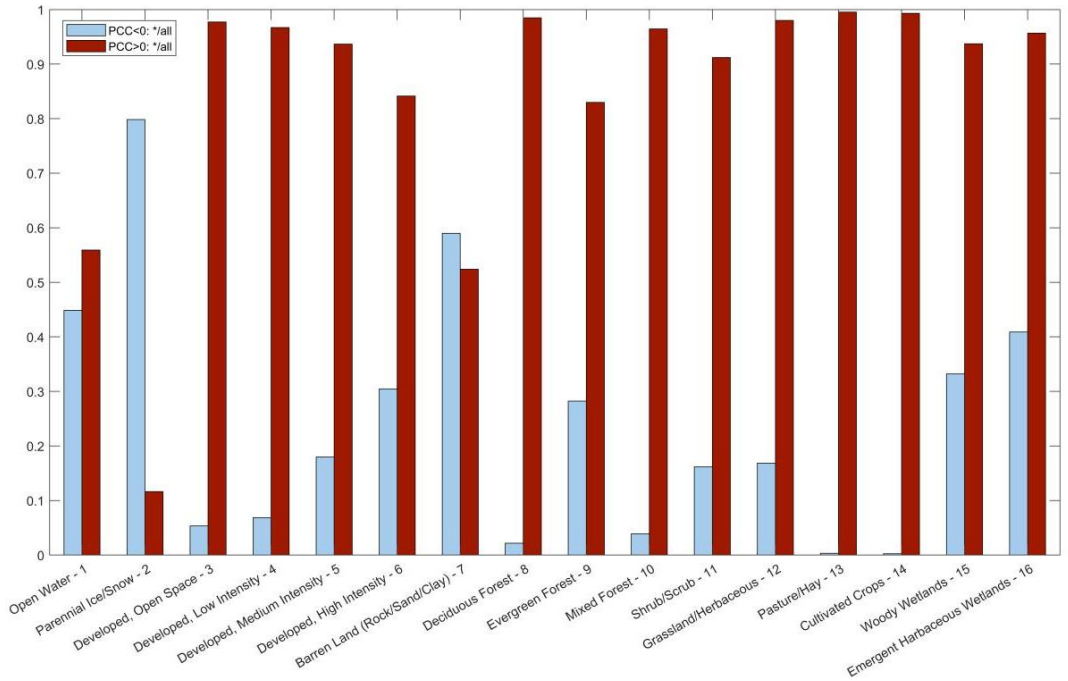

Figure 15. The bar chart illustrates the proportion of significant positive and negative correlation coefficients for different land cover types when Lead=0. Within each land cover category, the blue bars represent the proportion of pixels with significantly negative correlations among the significant pixels, while the red bars represent the proportion of pixels with significantly positive correlations.

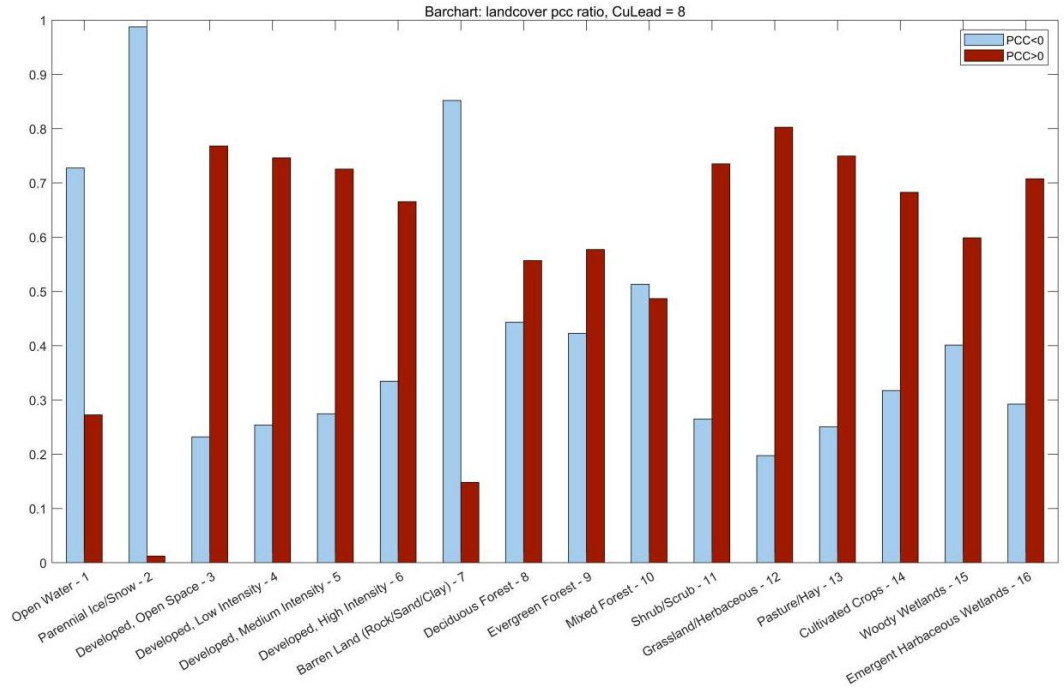

Figure 16. The bar chart illustrates the proportion of positive and negative correlation coefficients for different land cover types when CuLead=8. In the chart, the blue bars represent the proportion of pixels with negative correlation coefficients, while the red bars represent the proportion of pixels with positive correlation coefficients. The combined values of the blue and red bars sum up to one.

where the impact of water availability on plant growth diminishes as sufficient water is provided. Proper and scientifically managed agricultural practices can provide adequate water at different stages of plant growth, resulting in a cost-effective relationship between plant growth and water input. As illustrated in Figure 15, considering significance, both Pasture/Hay and Cultivated Crops land cover categories exhibit significant positive correlations.

Analyzing the results from Section 3.1. and Section 3.2., it becomes evident that the eastern regions generally exhibit stronger and more significant positive correlation compared with the western regions. Considering Figure 12, we can infer that the eastern and western sides of Colorado have distinct predominant land types, which contribute to the varying effects of ET (or cumulative ET) on NDVI. The red areas represent Cultivated Crops, which often display stronger and more significant positive correlation. Therefore, these regions are well-suited for crop cultivation, and this knowledge can inform agricultural planning and development strategies based on these findings.

4. Discussion

Chen *et al*. (2017) conducted a study that revealed a strong relationship between ET and NDVI with vegetation productivity (ANPP). Specifically, AET demonstrated a high correlation with ANPP, and there was also a significant correlation between NDVI and ANPP. These findings indicate the close relationship between ET and NDVI in influencing grassland productivity, providing valuable insights into the climate–vegetation interactions within grassland ecosystems.

By utilizing NDVI as an indicator of vegetation abundance and vitality, we observed that both ET and land cover had significant influences on plant production. Moreover, different land cover types demonstrated varying sensitivities to ET. These findings have significant implications for agriculture and environmental management, providing insights into the factors that affect plant growth and crop production. For example, planners can use NDVI as a proxy for water demand, optimizing irrigation schedules to reduce water waste. A comprehensive understanding of the relationship between ET, land cover, and NDVI can inform the development of more effective strategies for agricultural resource management. However, it is important to acknowledge certain limitations of this study, such as its focus on a specific region and time period, its

reliance on remote sensing data rather than direct measurements of plant growth, and possibly degrading usable scene availability due to Landsat's cloud-contaminated data. Future research should aim to expand the study's scope and incorporate more detailed data sources to address these limitations.

In other previous studies, DayCent model has been employed to simulate the impacts of climate and land use changes on various crops, grassland, and forest systems worldwide (Del Grosso *et al*., 2011). Regardless of rainy or non-rainy conditions, eMODIS NDVI explains approximately 50–52% of the variability in AET (Del Grosso *et al*., 2018). Del Grosso *et al*. conclude that remote sensing NDVI, when combined with the knowledge of Volumetric Soil Water Content (VSWC) and rainfall events, serves as a powerful tool for predicting Net Ecosystem Exchange (NEE) and AET across multiple temporal scales (from daily to seasonal) in semiarid grasslands.

In the future, additional research is necessary to precisely uncover the causal link between ET and NDVI. Supplementary datasets, such as soil moisture data, will be incorporated into the study. This exploration might uncover the fundamental mechanisms that dictate their relationship, thereby enhancing our understanding of the positive correlation in different land types.

Previous studies have shown a stronger correlation between ET and biomass. For instance, Ritchie *et al*. (1983) proposed an alternative approach for calculating biomass production independently of ET and T. Purcell *et al*. (2007) assessed the relationship between yield and estimated cumulative transpiration (T) or ET in well-watered soybean crops. Benjamin *et al*. (2015) found that soil water dynamics and soil water storage are critical for determining irrigation scheduling and impact various factors such as soil water use, water content, biomass production, grain yield, and water use efficiency in a continuous corn system in Colorado. Based on those existing research findings, as well as the results of this study, the future study will progress to predicting end-of-season biomass by utilizing ET, NDVI, land cover, and other relevant variables.

Through these potential studies, we expect to introduce usable methods using cutting-edge remote sensing technologies and recommend practical mathematical models for policymaking in agriculture area. Simultaneously, we seek to address issues such as water scarcity and climate change, providing support for the future sustainable development of agriculture.

5. Conclusions

This study aimed to investigate the correlation between ET, NDVI and land cover. By employing remote sensing data and statistical methodologies, our investigation revealed that approximately 92% of regions demonstrated a notable positive correlation between ET and NDVI. However, incorporating land cover into the analysis unveiled consistent patterns in Colorado, where Perennial Ice/Snow, Barren Land (Rock/Sand/Clay), and Open Water consistently showed a predominant negative correlation, ranking as the top three land types with NDVI, regardless of ET or cumulative ET. The findings presented in this study are comparable to those studies from Joiner *et al*. (2018) and Zeng *et al*. (2013). However, this study distinguishes itself by employing data with an unprecedentedly high spatial resolution of 30 meters, a novelty in this field.

Specific land types characterized by extensive root systems, enabling stable and continuous water evaporation and plant transpiration throughout the growing season, such as various forest types, exhibited a significant negative correlation with ET. Additionally, human intervention significantly influences correlation coefficients. Pasture/Hay and Cultivated Crops, two primary land types where plants are directly affected by human activity, exhibit correlation coefficients close to 1 between ET and NDVI, which is statistically significant.

In investigating the relationship between ET and NDVI through the introduction of Lead and CuLead variables, we noted that during the growing season, the lagged influence of ET on NDVI is less prominent compared with its concurrent month. Initially, the cumulative impact of ET on NDVI diminishes gradually, but it starts to intensify when CuLead≥7. This suggests that NDVI exerts a stronger influence on the cumulative ET throughout the entire growing season than during the recent period within the season.

In summary, this study pioneers the use of two-decade (2000–2018), high-resolution (30m) Landsat data to resolve fine-scale ET-NDVI dynamics in semi-arid Colorado, uncovering land cover-dependent mechanisms overlooked in coarse-scale studies. It demonstrates that agricultural systems (e.g., Cultivated Crops) exhibit strong positive correlations, indicating efficient conversion of water inputs to productivity, whereas natural ecosystems (e.g., Forests) show buffered responses due to soil water storage and canopy interception. Perennial Ice/Snow and Barren Land consistently display negative correlations. Temporally,

cumulative ET effects intensify at CuLead≥7 months, revealing legacy impacts on seasonal biomass. These findings deliver actionable pathways: optimizing irrigation to reduce agricultural water waste, guiding conservation planning for climate resilience for arid regions.

Data Availability Statement: Data are available for public at the sources in Table 1.

Acknowledgment: This work is partly supported by the US Department of Agriculture (USDA) UV-B Monitoring and Research Program, Colorado State University, under USDA National Institute of Food and Agriculture Grant 2022-34263-38472. This material is based upon work supported by the U.S. Geological Survey under Grant/Cooperative Agreement No. G18AP00077 (for GY18-GY22) or G23AP00683 (GY23-GY27).

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

An, N., Price, K. P., & Blair, J. M. (2013). Estimating above-ground net primary productivity of the tallgrass prairie ecosystem of the central Great Plains using AVHRR NDVI. *International Journal of Remote Sensing, 34*(11), 3717–3735.

Bauman, A., Goemans, C., Pritchett, J., & McFadden, D. T. (2013). Estimating the economic and social impacts from the drought in Southern Colorado. *Journal of Contemporary Water Research & Education, 151*(1), 61–69.

Benjamin, J. G., Nielsen, D. C., Vigil, M. F., Mikha, M. M., & Calderon, F. (2015). Cumulative deficit irrigation effects on corn biomass and grain yield under two tillage systems. *Agricultural Water Management, 159*, 107–114.

Briggs, J. M., & Knapp, A. K. (1995). Interannual variability in primary production in tallgrass prairie: Climate, soil moisture, topographic position, and fire as determinants of aboveground biomass. *American Journal of Botany, 82*(8), 1024–1030.

Chen, M., Parton, W. J., Del Grosso, S. J., Hartman, M. D., Day, K. A., Tucker, C. J., … Gao, W. (2017). The signature of sea surface temperature anomalies on the dynamics of semiarid grassland productivity. *Ecosphere, 8*(12), e02069.

Chen, M., Parton, W. J., Hartman, M. D., Del Grosso, S. J., Smith, W. K., Knapp, A. K., … Gao, W. (2019). Assessing precipitation, evapotranspiration, and NDVI as controls of US Great Plains plant production. *Ecosphere, 10*(10), e02889.

Crow, W. T., Kumar, S. V., & Bolten, J. D. (2012). On the utility of land surface models for agricultural drought monitoring. *Hydrology and Earth System Sciences, 16*(9), 3451–3460. https://www.hydrol-earth-syst-sci.net/16/3451/2012/

Del Grosso, S. J., Parton, W. J., Derner, J. D., Chen, M., & Tucker, C. J. (2018). Simple models to predict grassland ecosystem C exchange and actual evapotranspiration using NDVI and environmental variables. *Agricultural and Forest Meteorology, 249*, 1–10.

Del Grosso, S. J., Parton, W. J., Keough, C. A., & Reyes-Fox, M. (2011). Special features of the DayCent modeling package and additional procedures for parameterization, calibration, validation, and applications. In *Methods of introducing system models into agricultural research* (Vol. 2, pp. 155–176).

Dewitz, J., & U.S. Geological Survey. (2021). National Land Cover Database (NLCD) 2019 Products (ver. 2.0, June 2021) [Data set]. U.S. Geological Survey. https://doi.org/10.5066/P9KZCM54

ESPA - LSRD. (n.d.). https://espa.cr.usgs.gov/. Last access: 15 May 2023.

Formica, A. F., Burnside, R. J., & Dolman, P. M. (2017). Rainfall validates MODIS-derived NDVI as an index of spatio-temporal variation in green biomass across non-montane semi-arid and arid Central Asia. *Journal of Arid Environments, 142*, 11–21.

Haddeland, I., Heinke, J., Biemans, H., Eisner, S., Flörke, M., Hanasaki, N., … Wisser, D. (2014). Global water resources affected by human interventions and climate change. *Proceedings of the National Academy of Sciences, 111*(9), 3251–3256.

Homer, C., Dewitz, J., Jin, S., Xian, G., Costello, C., Danielson, P., … Riitters, K. (2020). Conterminous United States land cover change patterns 2001–2016 from the 2016 national land cover database. *ISPRS Journal of Photogrammetry and Remote Sensing, 162*, 184–199. https://doi.org/10.1016/j.isprsjprs.2020.02.019

Jin, S., Homer, C., Yang, L., Danielson, P., Dewitz, J., Li, C., … Howard, D. (2019). Overall methodology design for the United States national land cover database 2016 products. *Remote Sensing, 11*(24), 2971. https://doi.org/10.3390/rs11242971

Joiner, J., Yoshida, Y., Anderson, M., Holmes, T., Hain, C., Reichle, R., … Zeng, F. W. (2018). Global relationships among traditional reflectance vegetation indices (NDVI and NDII), evapotranspiration (ET), and soil moisture variability on weekly timescales. *Remote Sensing of Environment, 219*, 339–352.

Koster, R. D., Walker, G. K., Collatz, G. J., & Thornton, P. E. (2014). Hydroclimatic controls on the means and variability of vegetation phenology and carbon uptake. *Journal of Climate, 27*(14), 5632–5652.

Lindenmayer, D. B., Likens, G. E., Andersen, A., Bowman, D., Bull, C. M., Burns, E., … Wardle, G. M. (2012). Value of long-term ecological studies. *Austral Ecology, 37*(7), 745–757.

McKee, T. B., Doesken, N. J., Kleist, J., Shrier, C. J., & Stanton, W. P. (2000). A history of drought in Colorado: Lessons learned and what lies ahead.

Mukheibir, P. (2010). Water access, water scarcity, and climate change. *Environmental Management, 45*, 1027–1039.

Nikolaou, G., Neocleous, D., Christou, A., Kitta, E., & Katsoulas, N. (2020). Implementing sustainable irrigation in water-scarce regions under the impact of climate change. *Agronomy, 10*(8), 1120.

Parton, W. J., Gutmann, M. P., Merchant, E. R., Hartman, M. D., Adler, P. R., McNeal, F. M., & Lutz, S. M. (2015). Measuring and mitigating agricultural greenhouse gas production in the US Great Plains, 1870–2000. *Proceedings of the National Academy of Sciences, 112*(34), E4681–E4688.

Paruelo, J. M., Epstein, H. E., Lauenroth, W. K., & Burke, I. C. (1997). ANPP estimates from NDVI for the central grassland region of the United States. *Ecology, 78*(3), 953–958.

Paruelo, J. M., Oesterheld, M., Di Bella, C. M., Arzadum, M., Lafontaine, J., Cahuepé, M., & Rebella, C. M. (2000). Estimation of primary production of subhumid rangelands from remote sensing data. *Applied Vegetation Science, 3*(2), 189–195.

Petrie, M. D., Peters, D. P., Yao, J., Blair, J. M., Burruss, N. D., Collins, S. L., … Steiner, J. L. (2018). Regional grassland productivity responses to precipitation during multiyear above- and below-average rainfall periods. *Global Change Biology, 24*(5), 1935–1951.

Purcell, L. C., Edwards, J. T., & Brye, K. R. (2007). Soybean yield and biomass responses to cumulative transpiration: Questioning widely held beliefs. *Field Crops Research, 101*(1), 10–18.

Ritchie, J. T. (1983). Efficient water use in crop production: Discussion on the generality of relations between biomass production and evapotranspiration. In *Limitations to efficient water use in crop production* (pp. 29–44).

Rocha, N. S. D., Käfer, P. S., Skokovic, D., Veeck, G., Diaz, L. R., Kaiser, E. A., … Rolim, S. B. A. (2020). The influence of land surface temperature in evapotranspiration estimated by the S-SEBI model. *Atmosphere, 11*(10), 1059. https://doi.org/10.3390/atmos11101059

Sala, O. E., Parton, W. J., Joyce, L. A., & Lauenroth, W. K. (1988). Primary production of the central grassland region of the United States. *Ecology, 69*(1), 40–45.

U.S. Geological Survey. (n.d.-a). Collection 2 Landsat 8–9 OLI (Operational Land Imager) and TIRS (Thermal Infrared Sensor) Level-2 Science Product. https://doi.org/10.5066/P9OGBGM6

U.S. Geological Survey. (n.d.-b). *Collection 2 Landsat 7 Enhanced Thematic Mapper Plus (ETM+) Level-2 Science Product*. https://doi.org/10.5066/P9C7I13B

U.S. Geological Survey. (n.d.-c). *Landsat Collection 2 U.S. Landsat Analysis Ready Data (ARD) Level-2*. https://doi.org/10.5066/P960F8OC

United States Environmental Protection Agency. (2016). *What climate change means for Colorado* [PDF]. Last access: 1 July 2023.

USDA/NASS. (2021). *State agriculture overview for Colorado*. https://www.nass.usda.gov/Quick_Stats/Ag_Overview/stateOverview.php?state=COLORADO. Last access: 1 July 2023.

Wickham, J., Stehman, S. V., Sorenson, D. G., Gass, L., & Dewitz, J. A. (2021). Thematic accuracy assessment of the NLCD 2016 land cover for the conterminous United States. *Remote Sensing of Environment, 257*, 112357. https://doi.org/10.1016/j.rse.2021.112357

Wilcox, K. R., Blair, J. M., Smith, M. D., & Knapp, A. K. (2016). Does ecosystem sensitivity to precipitation at the site-level conform to regional-scale predictions? *Ecology, 97*(3), 561–568.

Yang, L., Jin, S., Danielson, P., Homer, C., Gass, L., Bender, S. M., … Xian, G. (2018). A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS Journal of Photogrammetry and Remote Sensing, 146*, 108–123.

Zeng, F. W., Collatz, G. J., Pinzon, J. E., & Ivanoff, A. (2013). Evaluating and quantifying the climate-driven interannual variability in Global Inventory Modeling and Mapping Studies (GIMMS) Normalized Difference Vegetation Index (NDVI3g) at global scales. *Remote Sensing, 5*(8), 3918–3950.

Zhang, M., & Yuan, X. (2020). Crucial role of natural processes in detecting human influence on evapotranspiration by multisource data analysis. *Journal of Hydrology, 580*, 124350.

Appendix A

Table A1. Land types and an introduction to land types in Colorado.

Analysis ID	NLCD ID	Land types
1	11	Open Water - All areas of open water, generally with less than 25% cover or vegetation or soil
2	12	Perennial Ice/Snow - All areas characterized by a perennial cover of ice and/or snow, generally greater than 25% of total cover.
3	21	Developed, Open Space - Includes areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses.
4	22	Developed, Low Intensity -Includes areas with a mixture of constructed materials and vegetation.
5	23	Developed, Medium Intensity - Includes areas with a mixture of constructed materials and vegetation.
6	24	Developed, High Intensity - Includes highly developed areas where people reside or work in high numbers.
7	31	Barren Land (Rock/Sand/Clay) - Barren areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits and other accumulations of earthen material.
8	41	Deciduous Forest - Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover.
9	42	Evergreen Forest - Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover.
10	43	Mixed Forest - Areas dominated by trees generally greater than 5 meters tall, and greater than 20% of total vegetation cover.
11	52	Shrub/Scrub - Areas dominated by shrubs; less than 5 meters tall with shrub canopy typically greater than 20% of total vegetation.
12	71	Grassland/Herbaceous - Areas dominated by graminoid or herbaceous vegetation, generally greater than 80% of total vegetation.
13	81	Pasture/Hay - Areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle.
14	82	Cultivated Crops - Areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards.
15	90	Woody Wetlands - Areas where forest or shrub land vegetation accounts for greater than 20 percent of vegetative cover and the soil or substrate is periodically saturated with or covered with water.
16	95	Emergent Herbaceous Wetlands - Areas where perennial herbaceous vegetation accounts for greater than 80 percent of vegetative cover and the soil or substrate is periodically saturated with or covered