Analysis of the intra-annual variations in soil moisture throughout the Missouri and Arkansas-White-Red River basins

by

Emily Nottingham

B.S., Kansas State University, 2019

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Carl and Melinda Helwig Department of Biological and Agricultural Engineering
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2021

Approved by:

Major Professor
Dr. Stacy Hutchinson

Copyright

Abstract

Soil moisture is listed as an Essential Climate Variable by the Global Climate Observing System Program. This is because soil moisture is a key factor in controlling the exchange of water and energy fluxes between the hydrosphere, biosphere, and the atmosphere through its impact on the partitioning of moisture for evapotranspiration and surface-sensible and latent heat fluxes. This characteristic of soil moisture also plays an important role within the hydrologic cycle due to its ability to control the rainfall-runoff response. Therefore, soil moisture is important in determining available water content and can have implications on water resources management for food and energy production. Thus, the evaluation of surface soil moisture at basin-scale is needed to understand spatiotemporal soil moisture trends and their implications on water resources management. Soil moisture To evaluate basin-level soil moisture trends, surface soil moisture estimates from SPoRT-LIS (0-10 cm layer) were used. Managed by NASA's Short-term Prediction Research and Transition (SPoRT) Center, the SPoRT-LIS is an observation-driven, real-time simulation of the Noah land surface model at a 3-km resolution over the full continental United States. This soil moisture product is at a higher spatial and temporal resolution than is currently available with remotely sensed satellite estimates or in situ measurements of the same product. Seasonal trend analysis was done using TIMESAT to determine soil moisture hydrometrics. Hydrometrics characterize the important seasonal components of soil moisture such as the start and end of the season and the corresponding levels of soil moisture throughout the season. To determine the TIMESAT parameter settings, a sensitivity analysis was done using soil hydrologic groups. Results from the TIMESAT analysis captured intra-annual soil moisture variability and highlighted the impact of soil texture and climate on the availability of soil moisture.

Next, the hydrometrics were compared to climate and soil variables to determine the impact that they have on the seasonality trends. This was done using a regression model with a space site effect. The results showed that all three variables, precipitation, temperature, and hydrologic soil groups significantly impacted hydrometrics. Precipitation had the largest impact on the available water content, field capacity, and wilting point of the soil where temperature had the largest impact on the start, middle, and end of season dates. This shows that precipitation drives soil water storage capacity where temperature is the driver of the seasonal timing of soil water storage. However, season length was the only hydrometric that was impacted the most by hydrologic soil groups. Ecoregions were also compared to the hydrometrics. This showed that there are additional drivers that impact hydrometrics which could include land cover, land use, and topography. From this, there is a better understanding of the spatiotemproal soil moisture variations throughout the Great Plains region which can help scientists, land managers, and policy makers to make decisions concerning reservoir management, irrigation applications, and farming practices.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

4DDA	–	Four- dimensional data assimilation
AMP	–	Seasonal Amplitude
AMSR-E	–	Advanced Microwave Sounding Radiometer
ASCAT	–	Advanced scatterometer
AVHRR	–	Advanced Very High-Resolution Radiometer
AWC	–	Available Water Content
BAS	–	Base Value
EOS	–	End of Season
ET	–	Evapotranspiration
ERS	–	European Remote Sensing
FC	–	Field Capacity
GSL	–	Growing Season Length
HSG	–	Hydrologic Soil Group
LSM	–	Land surface model
MOS	–	Middle of Season
MXF	–	Maximum Value
NDVI	–	Normalized Difference Vegetation Index
SMAP	–	Soil Moisture Active Passive
SMOS	–	Soil Moisture and Ocean Salinity
SMP	–	Soil Moisture Percentiles
SOS	–	Start of Season
SPI	–	Standardized Precipitation Index

SPoRT-LIS – Short-term Prediction Research and Transition Center Land Information System

STATSGO2 – Digital General Soil Map of the United States

SSURGO – Soil Survey Geographic Database

PDSI – Palmer Drought Severity Index

WP – Wilting Point

θ – Volumetric Water Content (m^3/m^3)

θ_{AWC} – Available Water Volumetric Water Content (m^3/m^3)

θ_{FC} – Field Capacity Volumetric Water Content (m^3/m^3)

θ_{WP} – Wilting Point Volumetric Water Content (m^3/m^3)

Acknowledgments

First and foremost, I would like to express my sincere gratitude to my major professor, Dr. Stacy Hutchinson, for her support and guidance. Her mentorship, enthusiasm, and immense knowledge have not only helped me complete this thesis but have inspired and shaped me into the person I am today. Also, I would like to recognize Dr. Shawn Hutchinson and Dr. Vaishali Sharda for their assistance on the research project and for being members of my supervisory committee.

Thank you KSU NRT faculty for providing a supportive community allowing me to grow as a young scientist and researcher. With this, I am grateful to Dr. Melanie Derby and Dr. Mathew Sanderson for being my faculty mentors. I would also like to acknowledge the help of Kelsey McDonough for her guidance on the research project as well as her mentorship. Also, I would like to thank all of the graduate students that I have had the opportunity to get to know during my time at Kansas State University for their friendship and support.

Finally, I would like to acknowledge my family and friends for their constant love and encouragement. I owe a huge thank you to Kyle for always being there for me. You listened to all my presentations and read over all my writings, even though you have no idea what I am talking about, you still show interest in what I am doing. I also owe a special thank you to my parents. You have supported me in all of my dreams and have taught me how to be a compassionate, curious, and adventurous woman, and for that, I am forever grateful. This research was supported by the National Science Foundation Grant 1828571.

Chapter 1 - Introduction

1.1 Problem Statement

Soil moisture is listed as an Essential Climate Variable by the Global Climate Observing System Program (GCOS, 2010). It is a key factor in controlling the exchange of water and energy fluxes between the hydrosphere, biosphere, and the atmosphere through its impact on the partitioning of moisture for evapotranspiration and surface-sensible and latent heat fluxes (Wang et al., 2016; Blankenship et al., 2016; Griesfeller et al., 2016). Soil moisture also plays an important role in the hydrologic cycle due to its control of the rainfall-runoff response (Dobriyal et al., 2012; Al-Shrafany et al., 2014; Alvarez-Garreton et al., 2014; Grillakis et al., 2016; Li et al., 2018; Meng et al., 2017). Soil moisture is important in determining available water content and can influence the productivity of natural and agricultural ecosystems (Dobriyal et al., 2012).

By understanding soil moisture variations in space and time, scientists can better improve flood and drought forecasting, weather and climate predictions, and crop growth modeling and monitoring (Wang et al., 2015, An et al., 2016; Blankenship et al., 2016; Dobriyal et al., 2012; Fascetti et al., 2016; Griesfeller et al., 2016). One promising way to examine the spatiotemporal variability of soil moisture on a large scale is to use soil moisture values derived from remote-sensing platforms. This powerful tool has enhanced our ability to understand land-atmosphere processes (Griesfeller et al., 2016) by providing global coverage at regular time intervals (Alvarez-Garreton et al., 2014). It has also improved streamflow prediction (Alvarez-Garreton et al., 2014; Brocca et al., 2012; Li et al., 2018) and has demonstrated the potential to advance short-term flood forecasting (Meng et al., 2017).

In addition to satellite-derived soil moisture, land surface models can simulate soil moisture based upon forcing variables, such as precipitation and wind speed, and physical properties such as soil texture and land cover (Koster et al., 2009; Blankenship et al., 2016). One major advantage of model-based soil moisture estimates is the ability to use real-time weather and climate information (Xia et al., 2015) and provide estimates of soil moisture to several meters of depth at hourly, daily, and monthly time steps (Moran et al., 2004; Cammalleri et al., 2015). Satellite observations have been combined with land surface models to improve surface soil estimates through data assimilation (Pinnington et al., 2018). These improved land surface models, with fine to moderate temporal scale, are useful for understanding the seasonality trends of soil moisture throughout large aerial extents.

1.2 Objectives

The overall goal of this research is to characterize the intra- annual variations of near-surface soil moisture throughout the Missouri and Arkansas-White-Red river basins using model-derived soil moisture estimates from the SPoRT-LIS software. SPoRT-LIS is managed by NASA's Short-term Prediction Research and Transition (SPoRT) Center. It produces surface soil moisture estimates (0-10 cm layer) from an observation-driven, real-time simulation of the Noah land surface model at a 3-km resolution over the full continental United States (https://weather.msfc.nasa.gov/sport/modeling/lis.html).

TIMESAT, a software package designed to analyze time-series satellite sensor data, is used for seasonal trend analysis, (Jonsson and Eklundh, 2002; http://web.nateko.lu.se/timesat/timesat.asp). This produces hydrometrics which characterize the important seasonal components of soil moisture, this includes when the start, middle, and end of the season occurs which defines the time when the soil is drying down. The hydrometrics also provide the available water content (AWC), wilting point (WP), and field capacity of the soil (FC; Figure 1.1). Field Capacity defines the maximum amount of water the soil can hold after excess drainage and WP defines the minimum amount of water the soil can hold before the plant wilts (NRCS, 1998). Available water content is defined as the amount of water that is available for plant uptake and is determined by the difference between FC and WP (NRCS, 1998). This method of characterizing the soil water storage is in contrast with the traditional method of using the rainfall-runoff approach because instead of estimating the amount of water within a system by measuring the inputs (precipitation and runon) and outputs (runoff and evapotranspiration) of the system, there is now the ability to measure and analyze soil moisture over large aerial extents.

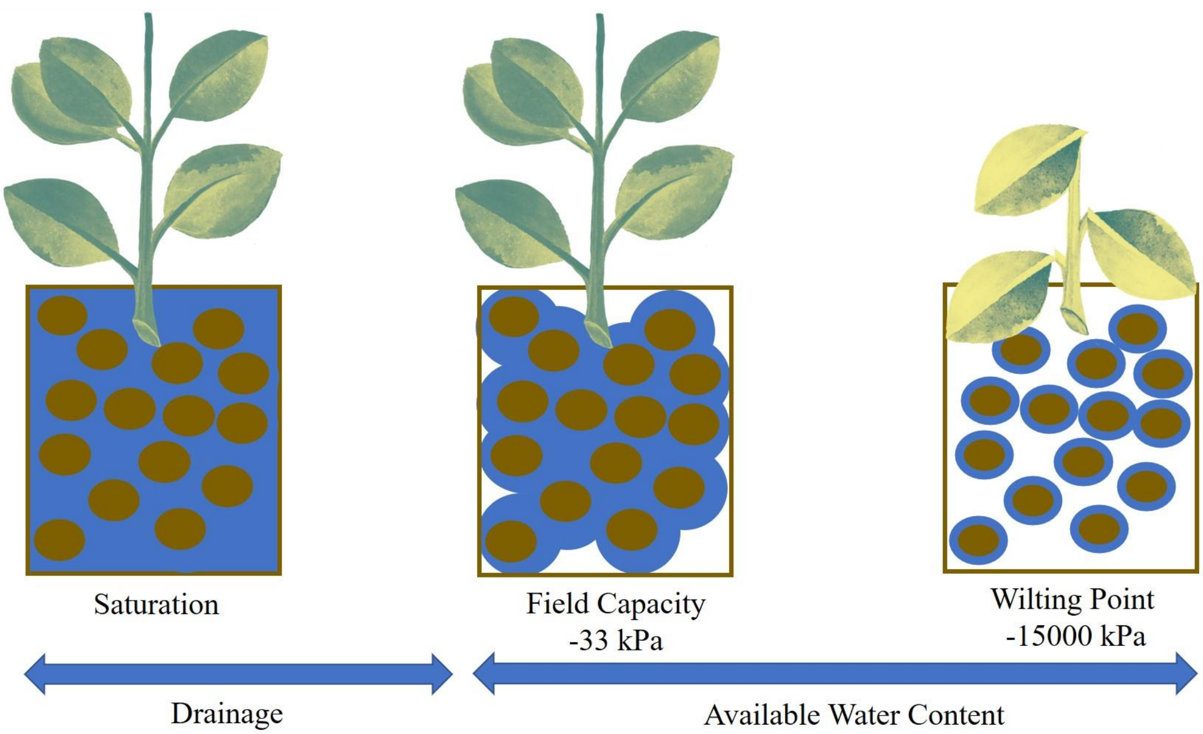

Figure 1.1. A graphical depiction of soil water holding capacity showing saturation, field capacity (FC), wilting point (WP), and available water content (AWC).

From the soil moisture seasonality trend analysis, it is hypothesized that the soil dry season will follow the growing season. Three different environmental variables are then used to characterize the hydrometrics to determine the individual influence each has on the region's soil moisture. These include:

- Hydrologic soil groups (HSG) from USDA-NRCS STATSGO2 summarizes soils into 4 different classifications, A, B, C, and D based upon the ability for water to enter the soil and the soil's retention capacity and ranges from sandy, loose soils (HGS A) to tighter clay soils (HSG D).
- 30-year normal (1981-2010) mean annual precipitation and temperature from PRISM Climate Group (2015). Precipitation and temperature can impact the soil moisture by impacting the amount of water that infiltrates into the soil as well as controlling evapotranspiration which extracts water from the soil (Akuraju et al., 2016).
- Level III ecoregions are classified based upon the biotic and abiotic phenomena that define ecosystems and their function (EPA, 2013). This includes topography, land cover, soil type, and climate which all impact soil moisture (Famiglietti et al., 1998).

HSGs are used to determine the influence the physical component of the soil has on soil moisture seasonality trends. The physical component of the soil is summarized into HSG based upon the soil type, particle size, and water retention capacity. It is hypothesized that the HSG will determine the AWC of the soil, the WP, and the FC values throughout the region. Following this, climate variables and soil characteristics are used to find out what impact they have on defining hydrometrics and if temperature, precipitation, or HSG play a larger role in the spatial patterns of each hydrometric. With this, it is hypothesized that precipitation and temperature will be the driving factor for the soil moisture characteristics of the region. Finally, level III ecoregions are used to determine if the seasonal soil moisture trends can be distinguished between each ecoregion. It is hypothesized that there will be variability between each ecoregion to support the different biotic and abiotic phenomena for the region.

Chapter 2 - Literature Review

2.1 Soil Moisture and its Role within the Hydrologic Cycle

Earth is known as the "Blue Planet" since approximately 70% of its surface is covered in water. Even though water mass only makes up 0.02% of Earth's mass, it is essential for the survival of all organisms (Oki and Kanae, 2006). Unlike most natural resources, water circulates naturally creating a continuous movement of water throughout the Earth's system. This cycling of water in and out of the atmosphere plays a significant role in regional and global weather patterns.

The hydrologic cycle follows a series of processes to circulate water in the form of a liquid, solid, or gas and ties together the major parts of the Earth's climate system, air, clouds, oceans, lakes, vegetation, snowpack, and glaciers (Oki and Kim, 2016). To begin the water is evaporated from the surface of the ocean turning it from a liquid state to a gas phase (Figure 1). From there the moist air condenses due to colder temperatures in the atmosphere where it forms clouds. The moisture is transported throughout the atmosphere until it returns to the surface as precipitation in the form of rain, snow, or ice. Once the water reaches the ground it could penetrate the surface and become groundwater or could be moved on top of the surface via runoff. For both groundwater and runoff, the water is then transported back to the oceans, rivers, and streams completing the cycle. However, water can also be returned to the atmosphere from the surface by evapotranspiration (ET). ET is a combination of moisture evaporated from the soil and moisture transpired from plants. ET approximately accounts for 66% of the precipitation while also making up 15% of the atmosphere's water vapor (Oki and Kim, 2016).

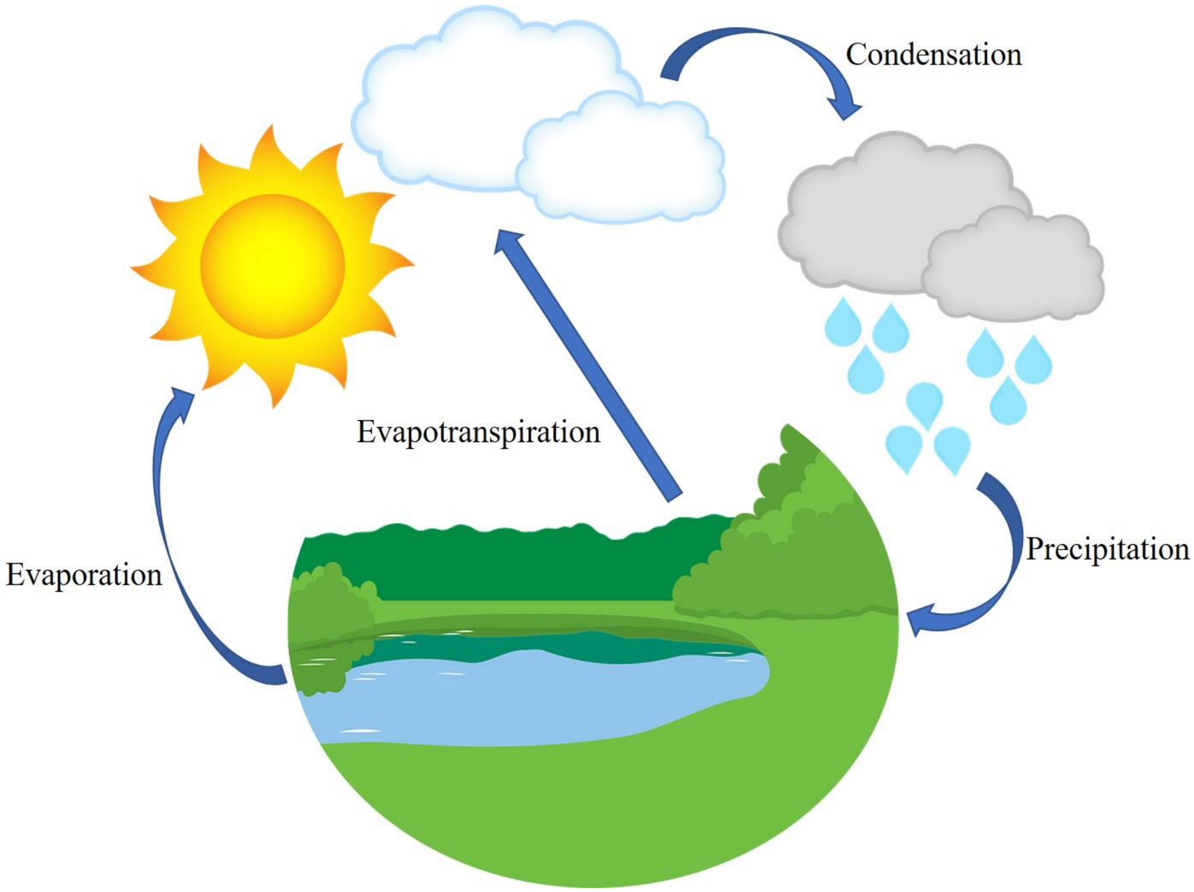

Figure 2.1. The global water cycle and its major components.

Earth's water can be stored in various reservoirs and various states throughout the hydrologic cycle. Oceans make 96.5% of the Earth's water, where 2.5% is considered freshwater (Shiklomanov, 1993; Figure 2.2). Of the small fraction of water on Earth that is considered freshwater, 68.7% is stored in glaciers and ice caps, 30.1% is groundwater, and 1.2% is surface water. The surface water can then be further divided into ice and permafrost, lakes, soil moisture, swamps and marshes, rivers, atmosphere, and living things.

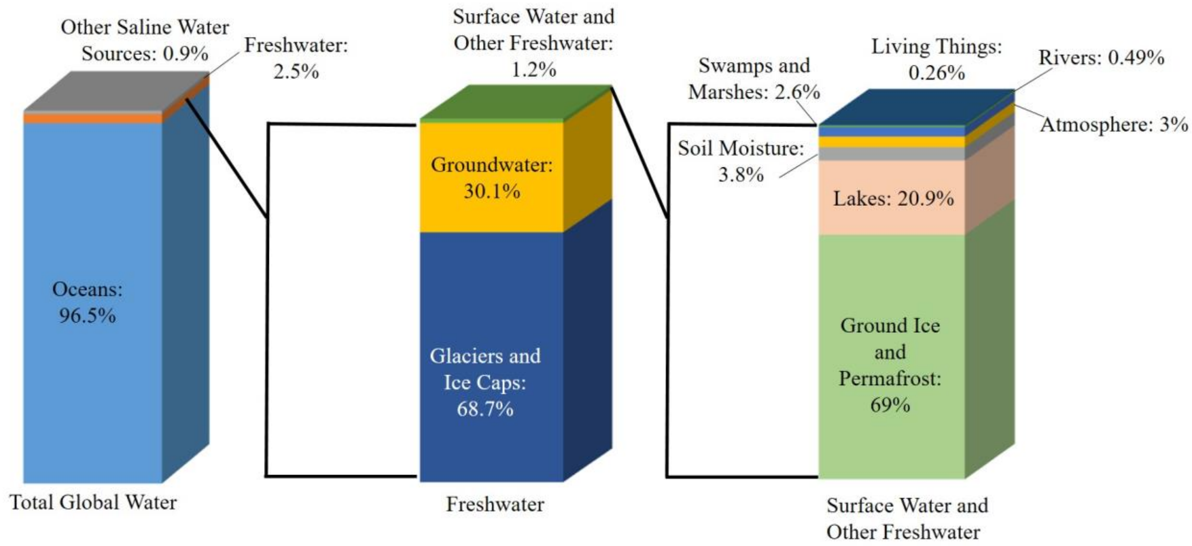

Figure 2.2. The distribution of Earth's water in the percentage of total water (adapted from Shiklomanov, 1993).

Since the 1980s and the start of four-dimensional data assimilation (4DDA) of the global atmosphere, the water cycle has been characterized by the water budget to determine the amount of water within the surface soil system (Oki and Kim, 2016). The water budget consists of inputs into the system such as precipitation and infiltration subtracted by the outputs of the system, including runoff, evapotranspiration, and deep seepage to determine the change in water storage (Figure 2.3). The equation is most often defined as,

$$\frac{dS}{dt} = P - R - ET, \tag{2.1}$$

where $\frac{dS}{dt}$ is the change in water storage over time, P is precipitation, R is runoff, and ET is evapotranspiration. Throughout time, there has been the ability to measure the inputs and outputs of the system but what water remains in the system has been estimated.

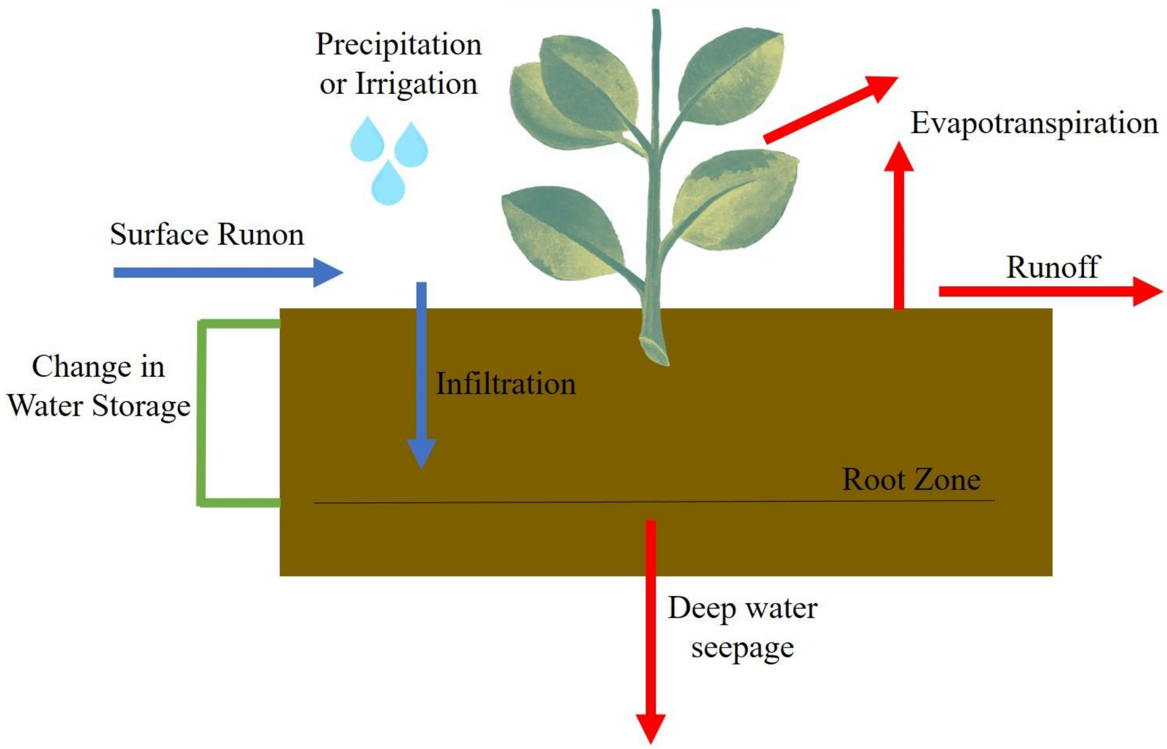

Figure 2.3. The water balance showing the inputs into the soil water system in blue (surface runon and infiltration), the outputs of the system in red (runoff, evapotranspiration, and deep water seepage), and the change in water storage in green.

It is important to accurately understand the amount of water in the soil. This is because it is a controlling factor in the exchange of water and energy fluxes between the hydrosphere, biosphere, and the atmosphere through its impact on the partitioning of moisture for evapotranspiration and surface-sensible and latent heat fluxes (Famiglietti et al., 1998; Wang et al., 2015; Chen et al., 2016; McDonough et al., 2018). Soil moisture also provides thermal inertia within the climate system, sorts and later releases heat, and dampens out diurnal and seasonal variation in surface temperature (Famiglietti et al., 1998). This is why soil moisture is listed as an Essential Climate Variable by the Global Climate Observing System Program (GCOS, 2010).

Soil moisture variability is influenced by many factors including topography, soil properties, vegetation type and density, mean moisture content, depth to the water table, precipitation depth, solar radiation, and other meteorological factors (Famiglietti et al., 1998). Topography influences soil moisture by affecting the infiltration, drainage, and runoff through the slope and angle of the landscape. The curvature of the landscape also influences the convergence of lateral flow and the upslope surface area influences the distribution of soil moisture by controlling the potential volume of subsurface moisture flowing past a particular point on the landscape. Soil properties affect soil moisture through variations in color, organic matter content, structure, and the existence of macroporosity. Soil texture also impacts soil moisture through the rate of evaporative drying. This is supported by Fernandez-Illesca et al., (2001) which found that sandy soils tend to have the lowest mean levels of ET while silty loam soils have the highest ET.

Surface soil moisture also shows a significant response to rainfall and temperature. A study by Akuraju et al., (2016), found that ET is strongly constrained when soil moisture is between the wilting point and field capacity but if soil moisture is below the wilting point, ET is not constrained by soil moisture. It was also shown that under energy-limited conditions ET is not strongly related to soil moisture. However, under water-limited conditions, the variability in ET is constrained by soil moisture and crop growth stage. Energy-limited conditions impact the rate of ET through limited atmospheric energy controlled by incoming radiation (Eagleson, 1978). For water-limited conditions, the rate of ET is limited by water availability and not energy.

Land cover also plays an important role in controlling spatial patterns of soil moisture by influencing infiltration, runoff, and ET, particularly during the growing season (Chen et al., 2016). Vegetation influences the pattern of rainfall that is intercepted by the canopy. It also shades the land surface affecting the rate of evaporative drying along with impacting soil hydraulic conductivity through root activity and the addition of an organic matter layer. Vegetation then extracts moisture from the soil profile through transpiration (Famiglietti et al., 1998). This is especially true in water-limited regions where vegetation and soil moisture play a coupled role in the ecosystem dynamics (Fernandez-Illesca et al., 2001). Therefore, soil moisture is important in determining available water and can dictate the productivity of natural and agricultural ecosystems; especially since 75% of the global freshwater is used for agriculture annually with a majority of the water being returned to the atmosphere via ET (Akuraju et al., 2016). Feng (2016) found that vegetation expansion has the potential to mitigate soil wetting trends where vegetation degradation in these wetting regions has the potential to lengthen the dry season and increase streamflow, thus decreasing rainfall infiltration. When looking at drying regions, the opposite was found where vegetation degradation can help mitigate the drying trend. It is thus important to assess the long-term and large-scale historical patterns and trends in regional soil moisture, which provides useful information to understand the individual effects from land cover and climate variability.

2.2 Soil Moisture and Climate Change

Global climate change is having a direct impact on the global water and energy cycles (Rasul and Sharma, 2015; Ehsani et al., 2017). Climate change is increasing global average surface temperatures, increasing the frequency and intensity of heat waves and droughts, impacting precipitation frequency and intensity, reducing snow cover, and causing widespread melting of ice. All of which impact the soil moisture content and make it increasingly important to understand the amount of soil moisture within the system and the variability across the landscape. This is particularly critical since the global water and energy cycles are driven by certain soil moisture-controlled land-atmosphere interactions, including latent and sensible heat fluxes (Dorigo et al., 2012; Qiu et al., 2016; Lai et al., 2016).

Soil moisture also has an impact on floods and droughts (Ault et al., 2016; Albergel et al., 2013) because soil moisture integrates the effects of moisture supply, storage, and atmospheric demand (Ault et al., 2016). Under future climate projections, droughts are expected to become more prevalent as it has been found that higher temperatures shift the moisture balance toward conditions that are drier on average (Famiglietti and Rodell, 2013a; Cook et al., 2014; Cook et al., 2015; Ault et al., 2016; Ehsani et al., 2017). This is caused by the projected increases in atmospheric demand for moisture from the land's surface, thus increasing the ET rates and shifting the soil moisture baseline (Cook et al., 2014; Ault et al., 2016).

Over the central plains, drying is driven primarily by the increased evaporative demand during the spring and summer seasons (Hoerling et al., 2012; Cook et al., 2015). This increased evaporative demand is likely to be sufficient to overcome precipitation increases and result in a decrease of 3 to 12 percent annual water availability throughout the United States (Ehsani et al., 2017). With this decline, existing dams and reservoirs are incapable of storing the added water in the wet season to supplement lower flows in the dry months, affecting the ability of the system to meet human and environmental water demands as well as hydropower production targets and thermoelectric generation. All of this can have a long-term impact on critical water resources, agricultural production, and economic activity, presenting major adaption challenges for managing ecological and anthropogenic water needs in the region. (Cook et al., 2014; Cook et al., 2015). Also, the human population in the region, and their associated water resource demands, have been increasing rapidly in recent decades, and these trends are expected to continue for years to come (Cook et al., 2015).

On the other hand, in some regions, extreme flooding is becoming more frequent and with greater intensity. There is an expected increase in global precipitation by 1 to 3 percent and an expected increase in extreme precipitation by 5 to 10 percent (Famiglietti and Rodell, 2013; Berghuijs, et al., 2016; Swain et al, 2020) Models project that areas that are already dry will become drier and wet areas will get wetter (Cook et al., 2014; Greve et al., 2014; Roderick et al., 2014; Feng and Zhang, 2015) This has implications on portions of the Gulf Coast, the southeastern United States, and the Carolinas which have all seen catastrophic floods between 2015 and 2020 (Swain et al., 2020). Throughout most of the United States, snowmelt and soil moisture are thought to be the controlling factor of flood response (Berghuijs, et al., 2016).

All of these results brought on by climate change can have long-term impacts on critical water resources, agricultural production, and economic activity, presenting major adaption challenges for managing ecological and anthropogenic water needs in the region. (Cook et al., 2014; Cook et al., 2015; Berghuijs et al., 2016). Thus, it is important to be able to measure the water storage within a region. Feng and Zhang (2015) found that using derived indices such as Palmer Drought Severity Index (PDSI), Standardized Precipitation Index (SPI), or Soil Moisture Percentiles (SMP) resulted in significant errors and uncertainties and overestimated future drought conditions. Using a hydrologic variable such as soil moisture that connects precipitation and ET can result in better predictions of drying and wetting trends (Greve et al., 2014; Feng and Zhang, 2015). Thus, by understanding soil moisture variation in space and time, scientists can better improve flood and drought forecasting, weather and climate prediction, and crop growth modeling and monitoring (Albergel et al., 2013; Wang et al., 2015; Qui et al., 2016).

2.3 Soil Moisture Assessment Methods

Soil moisture data is primarily collected via three methods: in situ observations, remotely sensed satellite observations, and climate and land surface models (Xia et al., 2015). Due to the varying assessment methods, the spatial and temporal scale of soil moisture data varies widely.

2.3.1 In-Situ Observations

In situ data provides point-scale estimates of soil moisture by measuring the soil moisture content at a specific location in space and time. One method for doing this is gravimetric sampling which takes a fresh soil sample or soil core. The soil sample is weighed, then oven-dried, and weighed again. The difference between the fresh soil weight and the dried soil weight provides the mass of water per mass of dry soil (Erback, 1987). This can also be done on a volumetric basis where the soil water content is measured as a percentage of soil water volume. This process is relatively easy to achieve but can require extensive sampling time. For larger study areas, such as field sites, automated probes are a more viable option (Srivastava et al., 2016). Automated sampling probes are placed in the field and measure the soil water content at the location and depth of the sensor (Sample et al., 2016). The ease of this method comes with its inaccuracies and concerns such as calibration difficulties due to salinity, temperature, and soil texture (Escoriheula et al., 2006; Vaz et al., 2013).

In situ data has been useful for the evolution of remotely sensed and modeled soil moisture data (Albergel et al., 2013; Xia et al., 2015). It is also valuable in validating other soil moisture estimates (Jacobs et al., 2004; Albergel et al., 2013; Xia et al, 2015; McDonough et al., 2018). However, this method of data collection lacks large-scale applications because of very limited in situ hydrologic observation networks that have global coverage and representativeness across space and time (Dorigo et al., 2012). This type of representativeness requires dense ground-based networks that would be necessary to truly capture the spatial and temporal variations of soil moisture due to soil characteristics, precipitation patterns, temperature patterns, landcover, and other physical and climatological factors (Dobriyal et al., 2012; Dorigo et al., 2012; Al-Shrafany et al., 2014). Thus, it is a challenge to infer spatiotemporal patterns of total water stored throughout a region with in situ observations alone (Famiglietti and Rodell, 2014). This lack of robust data has caused fine-scaled soil moisture variations and the impacts of regional patterns of precipitation, temperature, and ET to not be well understood (Cook et al., 2014).

2.3.2 Remotely Sensed Soil Moisture Observations

Soil moisture estimates derived from remotely sensed platforms provide a path for analyzing spatial-temporal variability of soil moisture over regional or global scales, unlike in situ observations alone (Wang et al., 2015; McDonough, 2020). The satellite-derived soil moisture data provides global coverage at regular time intervals from 1979 onwards (Dorigo, et al., 2012; Srivastava et al., 2013; Alvarez-Garreton et al., 2014). This has allowed for the continued understanding of land-atmosphere interactions and processes along with an insight into climate change (Griesfeller et al., 2016; Srivastava et al., 2015).

The first set of remotely sensed soil moisture estimates were obtained through short wave measurements and relied on the fact that soils get darker when wet (Srivastava et al., 2016). This method resulted in measurement error due to atmospheric effects, cloud cover, and vegetation cover (Kerr 2007). Another method for soil moisture measurement included latent heat effects where wet soils have higher thermal inertia compared to dry soils (Srivastava et al., 2016). This too had inaccuracies due to atmospheric effects, cloud masking, vegetation cover opacity, and atmospheric interactions with the soil's surface layer. An improvement to this method, microwave systems, can be used to measure the dielectric constant of soils that relates directly to the water content. Microwave systems also offer the ability to be used in all weather conditions, can penetrate vegetation, and operate at nighttime (Srivastava et al., 2015).

Several satellites have been launched to perform global soil moisture observations. One of these satellites, the European Remote Sensing satellite (ERS), offers a fine spatial resolution of ten meters but a low temporal resolution of 35 days (Wagner et al., 2007). Another includes the advanced scatterometer (ASCAT) which gives a freely available global soil moisture data set derived from backscatter measurement with an active microwave remote sensor (Wagner et al., 1999). Also, the Soil Moisture and Ocean Salinity (SMOS) mission were launched in 2009 (Kerr et al., 2001). Similarly, NASA launched the Soil Moisture Active Passive (SMAP) in 2015 (Entekhabi et al., 2010). Both SMOS and SMAP are passive remote sensors that have a spatial resolution of 40 km and a 3-day revisit time. Another passive microwave instrument is the Advanced Microwave Sounding Radiometer (AMSR-E; Reichle et al., 2007). This satellite was not specifically designed for soil moisture retrieval but has shown good performance (Al-Yaari et al., 2014).

These satellites allow for the global observation of soil moisture, but the accuracy of the data is variable through both space and time (Brocca et al., 2012). There are also concerns over the coarse spatial and temporal resolution along with the capability to only measure the first few centimeters of the soil layer with these products (Al-Shrafany et al., 2014; Alvarez-Garreton et al., 2014; Brocca et al., 2012).

2.3.3 Modeled Soil Moisture Data

Land surface models (LSMs) are another method for soil moisture observation. LSMs work to solve the coupled fluxes of water, energy, and carbon between the land surface and the atmosphere (Fisher and Koven, 2020). This is done through forcing variables including precipitation, wind speed, and physical properties such as soil texture and land cover (Koster et al., 2009; Blankenship et al., 2016). As LSMs progress, the proposal of additional axes of variation has been introduced to represent particular land surface processes, including the representation of soil moisture dynamics beginning in the 1980s (Fisher and Koven, 2020). This has allowed for LSMs to provide global coverage of soil moisture data, similar to satellite-derived data, but it can provide estimates of soil moisture to several meters of depth at hourly, daily, and monthly time steps (Moran et al., 2004; Cammalleri et al., 2015).

A major advantage to LSMs is that they can use real-time climate and weather data (Xia et al., 2015). However, it has been shown that modeling soil moisture can be very complex along with large sensitives to meteorological forcing data and LSM parameterization (Pitman et al., 1999; Koster et al., 2009; Camalleri et al., 2015; Xia et al., 2015; Yang et al., 2016). Satellite observations and LSMs have been combined to improve surface soil moisture estimates through data assimilation (Pinnington et al., 2018). This has allowed for LSMs to be a valuable tool in measuring soil moisture at a high spatial and temporal resolution (McDonough et al., 2018).

2.4 Soil Moisture Trend Analysis

Soil moisture trend analysis is important for understanding climate change effects along with helping manage water resources to combat global water security challenges (Dobriyal et al., 2012; Albergel et al., 2013; Qui et al., 2016). McDonough et al., (2020) noted that a majority of the past research over soil moisture trend analysis has been conducted at the global and continental scales. This is seen in a study by Dorigo et al., (2012) which looked at the merged microwave-based surface soil moisture dataset, SM-MW (Liu et al., 2011), from 1988 to 2010. Intra-annual trend analysis was conducted using the non-parametric Mann-Kendall test along with defined seasons from December-February, March-May, June-August, and September-November. It was found that a majority of the significant trends were drying trends were not explained by precipitation alone. Instead, other drivers of soil moisture variations such as evaporation, soil type, irradiation, vegetation, and topography had an impact on the trend. Albergel et al., (2013) found similar results by using soil moisture values from ERA-Land and MERRA-Land and averaged them across the defined seasons stated above. These trends were investigated with the Koppen-Geiger climate classification to understand precipitation and temperature impacts.

Also, Feng and Zhang (2015) along with Feng (2016) defined global climate regions of wetting and drying using the Climate Change Initiative data set of the European Space Agency that contains fusion data of active and passive microwave satellite observations at a resolution of 25 km from 1978 to 2013. Feng (2016) analyzed these trends with climate change and vegetation data. It was found that climate change dominates the soil moisture trends, but vegetation can have a negative effect on soil moisture trends in the dry and sparsely vegetated regions but has an opposite impact in wet and densely vegetated regions.

This research has shown the long-term soil moisture trends at the global scale ranging in spatial resolution from 25 km to 80 km, which helps in understanding regions that are becoming wetter or dryer (Dorigo et al., 2012; Albergel et al., 2013; Feng and Zhang 2015; Feng 2016). This has implications on global climate change impacts and contributes to improved land components of climate models and global circulation models (Grayson et al., 1997). However, these global soil moisture trend analysis studies, along with coarse spatial resolutions, are not capable of capturing details of local and regional water availability (Hall, 2014; Mekonnen and Hoekstra, 2016).

To better understand regional soil moisture trends, McDonough et al., (2020) characterized the long-term trend in soil moisture across the Missouri and Arkansas-White-Red River basins in the United States. This was done using SPoRT-LIS near-surface soil moisture estimates with a 3 km spatial resolution from 1987 to 2016. It was found that there is a drying trend in soil moisture throughout most of the region. This long-term trend analysis of basin-scale soil moisture has implications on important hydrological processes and water resources management, but there is evidence that average soil water content changes seasonally and that seasonal changes in precipitation and ET tend to lead to periods where soils are persistently wetter or drier than average (Grayson et al., 1997; Deliberty and Legates, 2003). Soil moisture can be extremely dynamic and heterogeneous with large day-to-day variability over relatively small areas. The seasonal variability of soil moisture plays a role critical role in irrigation and reservoir management, flood and drought predictions, impacts fire frequencies, and has an effect on the development of perennial vegetation (Deliberty and Legates, 2003).

Soil moisture trend analysis had been made possible through recent advancements in satellite and modeling technology that has shifted our ability from measuring soil moisture at a single point scale with in situ data to now having coarse soil moisture data over large aerial extents (Moran et al., 2004; Alvarez-Garreton et al., 2014; Cammalleri et al., 2015; Griesfeller et al., 2016; Pinnington et al., 2018). This new technology has allowed soil moisture to be characterized at different spatial and temporal scales (Table 2.1) allowing for a better understanding of the soil water storage within a system. This is in contrast to the traditional rainfall-runoff approach that considers the water budget and measures the inputs and outputs of the system to estimate the soil water storage.

Table 2.1. A summary of the reviewed soil moisture trend analysis studies and the soil moisture datasets.

Soil moisture dataset	Spatial Resolution	Spatial Extent	Temporal Resolution	Temporal Extent	Source
ERA-Land	80 km	Global Scale	6-hourly	1980-2010	Albergel et al., 2013
MERRA-Land	1/2 degree and 2/3 degree in latitude and longitude	Global Scale	2-3 days	1980-2010	Albergel et al., 2013
SM-MW	0.25 degrees	Global Scale	Daily	1988-2010	Dorigo et al., 2012
Climat Change Initiative (CCI) data set of the European Space Agency (ESA)	25 km	Global Scale	Daily	1987-2013	Feng and Zhang, 2015; Feng, 2016
SPoRT-LIS	3 km	Basin Scale	Daily	1987-2016	McDonough et al., 2020

Chapter 3 - Methods and Materials

3.1 Study Area

The Missouri and Arkansas-White-Red River basins are located in the central part of the United States and cover most of the Great Plains region, and portions of Montana, South Dakota, Nebraska, Kansas, Oklahoma, and portions of Wyoming, North Dakota, Colorado, Iowa, Missouri, Arkansas, Texas, New Mexico, and Louisiana (Figure 3.1). The basins are characterized by varying climate, topology, geology, and land cover.

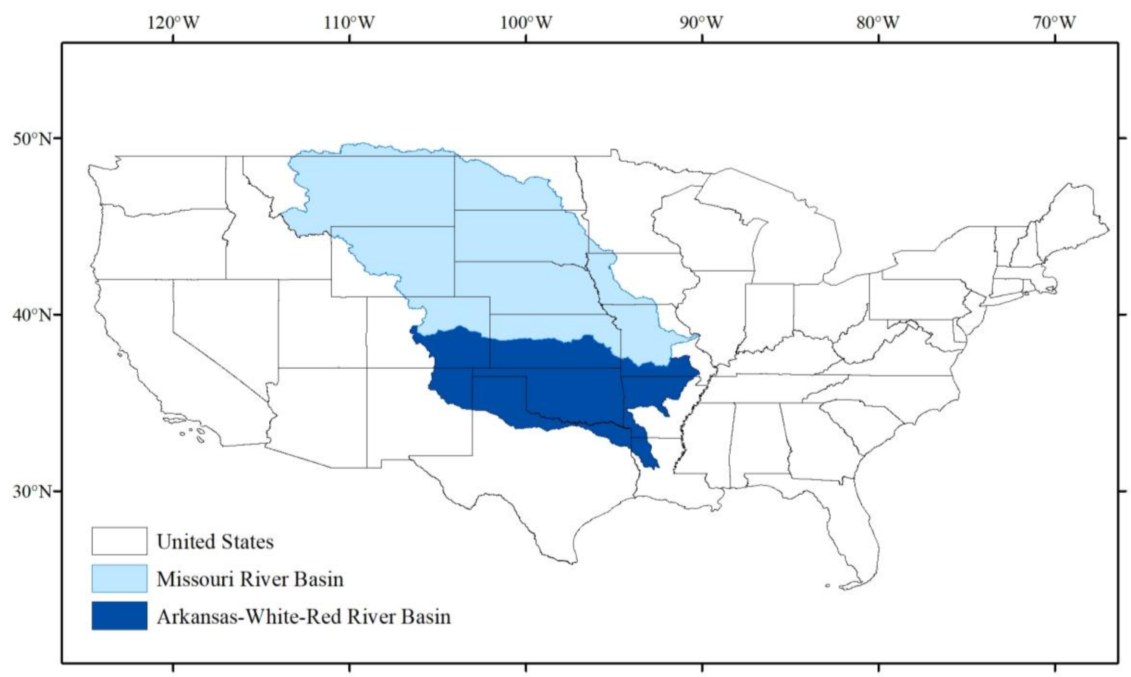

Figure 3.1. The location of the Missouri and Arkansas-White-Red river basins within the United States.

Throughout the study area, the climate is transitional between humid in the east and arid conditions in the west (Cook et al., 2007). The region experiences multiple climate and weather extremes including floods, droughts, severe storms, tornadoes, and winter storms (Antle, et al., 2014). The mean annual precipitation ranges from 142 mm in the northwest to 2077 mm in the southeast portion of the region (Figure 3.1; PRISM, 2015); throughout most of the Great Plains, this is not enough to replace the water demands of humans, plants, and animals (Antle et al., 2014). The Great Plains features relatively flat plains that increase in elevation from 20 m to more than 3,900 m at the base of the mountain ranges along the Continental Divide (Figure 3.3). This is mirrored by mean annual temperature where the mountainous regions to the north and west have an average annual temperature less than 0 °C and the southern portion of the plains experience mean annual temperatures greater than 19 °C (Figure 3.4). This varying climate results in challenges for climate change adaption and sustainable water resources management that add to already stressed communities (Antle et al., 2014).

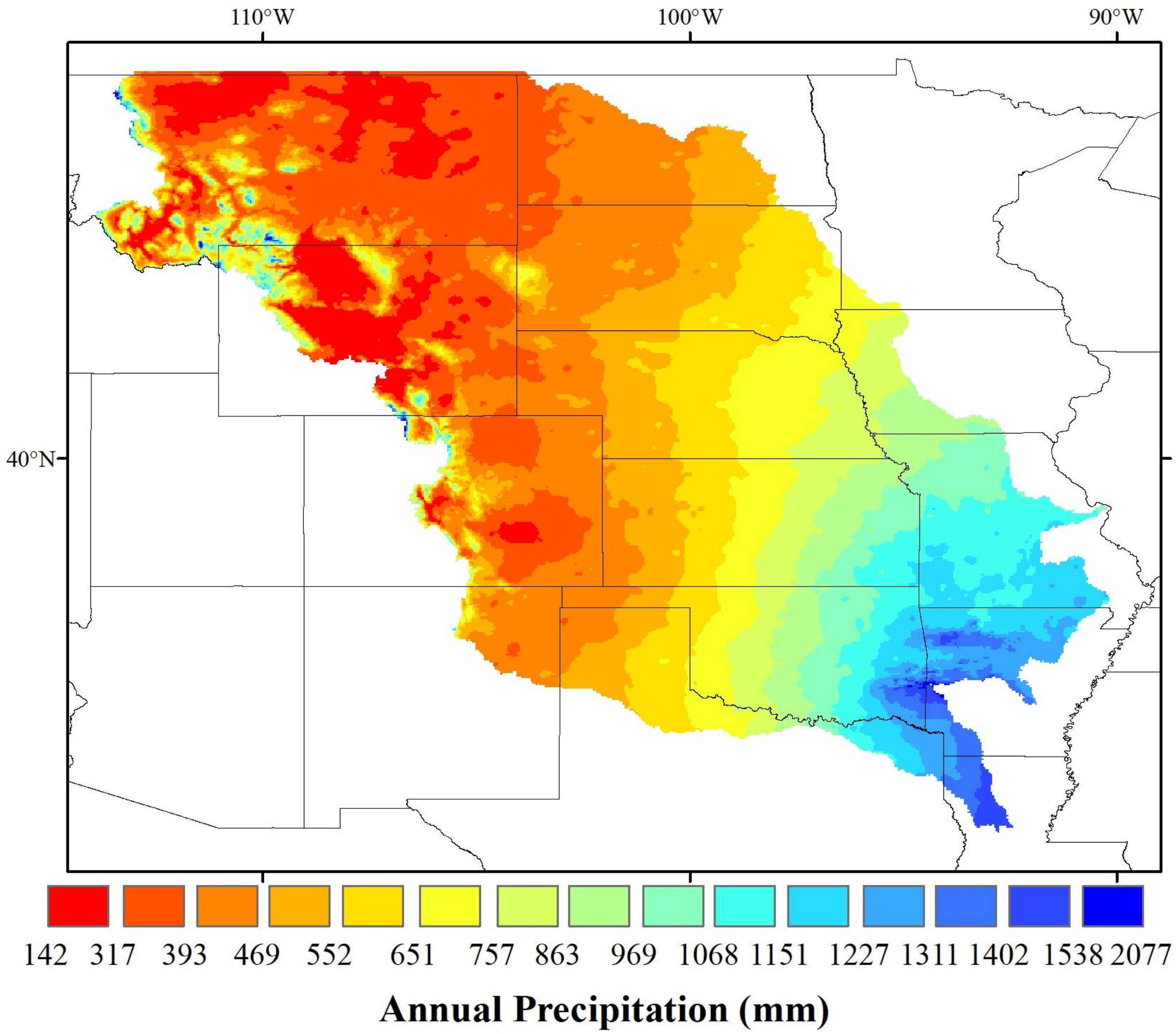

Figure 3.2. The 30-year-normal (1981-2010) mean annual precipitation (mm) across the Missouri and Arkansas-White-Red river basins (PRISM, 2015).

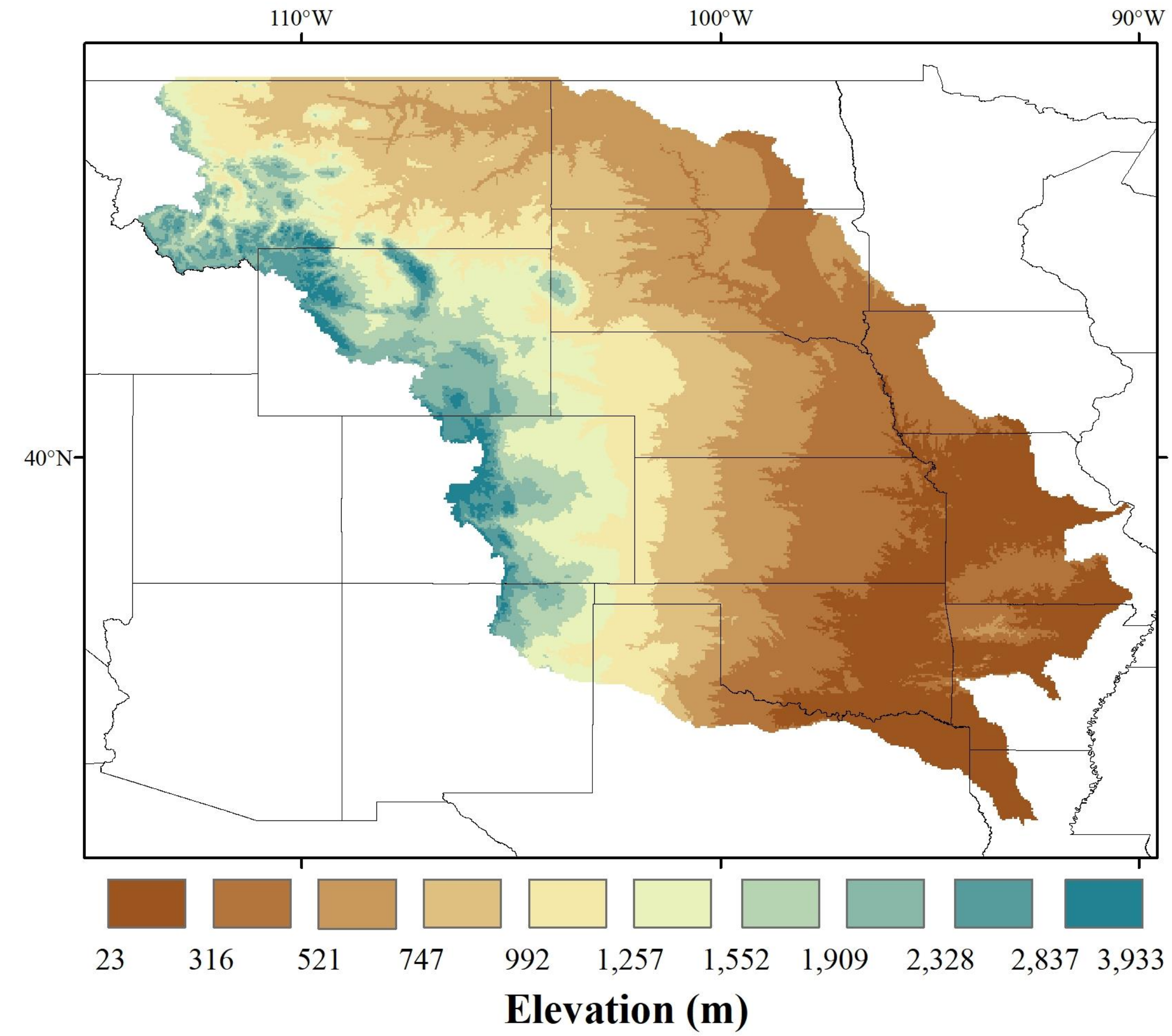

Figure 3.3. Elevation (m) throughout the Missouri and Arkansas-White-Red river basins (PRISM, 2014).

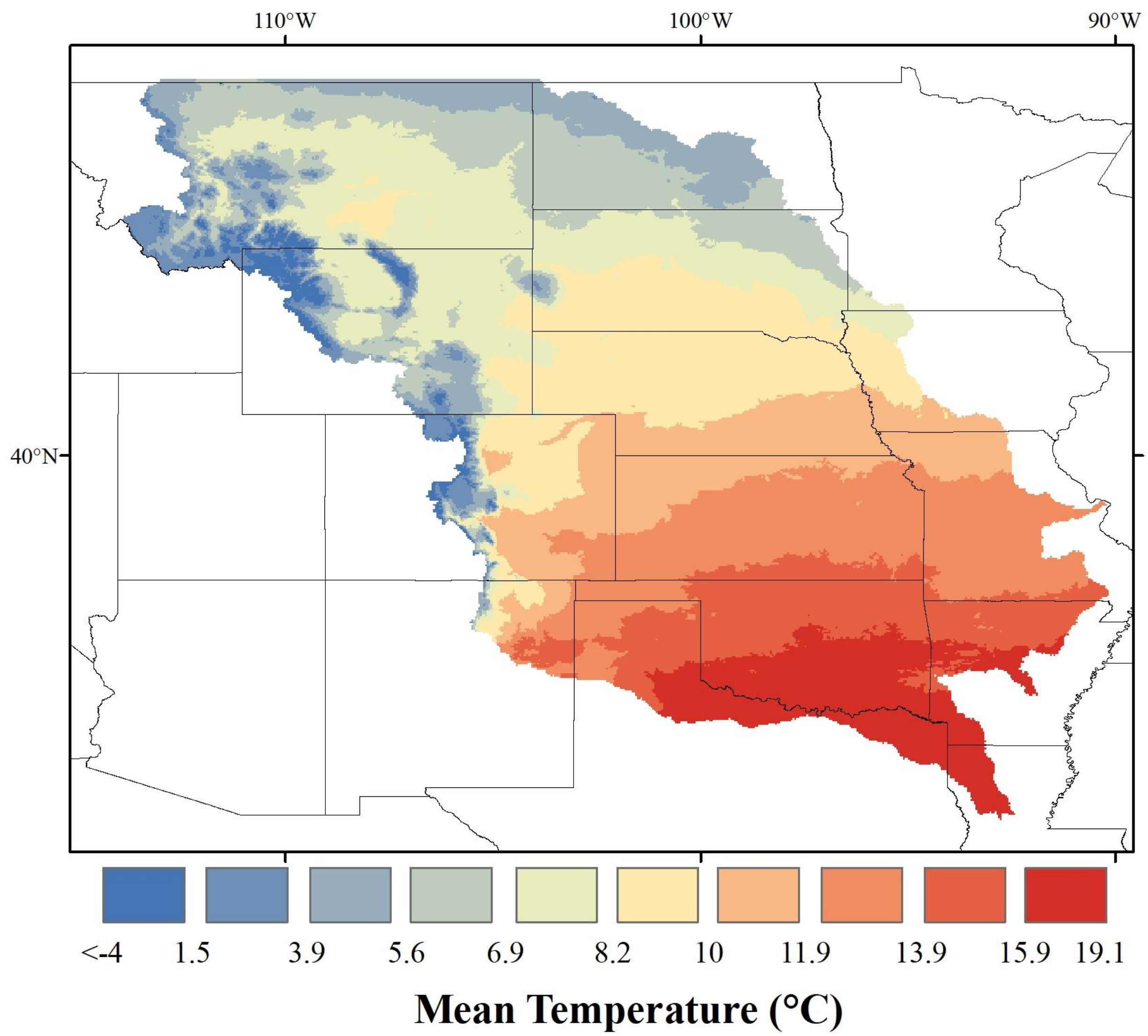

Figure 3.4. The 30-year-normal mean annual temperature (°C) over the Missouri and Arkansas-White-Red river basins (PRISM, 2015).

The land cover for the Great Plains region is predominantly herbaceous grassland which makes up approximately 35% of the study area (Figure 3.5; Dewitz, 2019). Cultivated cropland along with hay and pasture make up 25% and 7% of the area respectively with only 3% of the area classified as urban. The agricultural, municipal, and industrial water needs for the region are dependent on the accessibility as well as the availability of water resources (Moore et al., 2015). Less than 10% of the farmland area in the region is irrigated (Figure 3.6; USDA-NASS, 2012).

The area with the most irrigation is in Kansas and Nebraska, overlaying the Ogallala Aquifer. The Ogallala Aquifer is considered the largest contiguous area of local water stress in the United States due to declining groundwater and surface water availability (Moore et al., 2015). As for surface water storage, there are more than 1,500 reservoirs throughout the Missouri and Arkansas-White-Red River basins (Figure 3.7; USGS, 2006). These reservoirs provide flood protection, hydropower, electricity, and recreation. However, reservoirs in the area are threatened due to declining storage capacity from siltation, degraded streambanks, and increases in annual flood magnitudes (Melillo et al., 2014; Contant et al., 2018). With declining groundwater and surface water availability, research is needed to understand the patterns of regional intra-annual hydrologic variability to maintain the water resources that provide the livelihood of the people within the area.

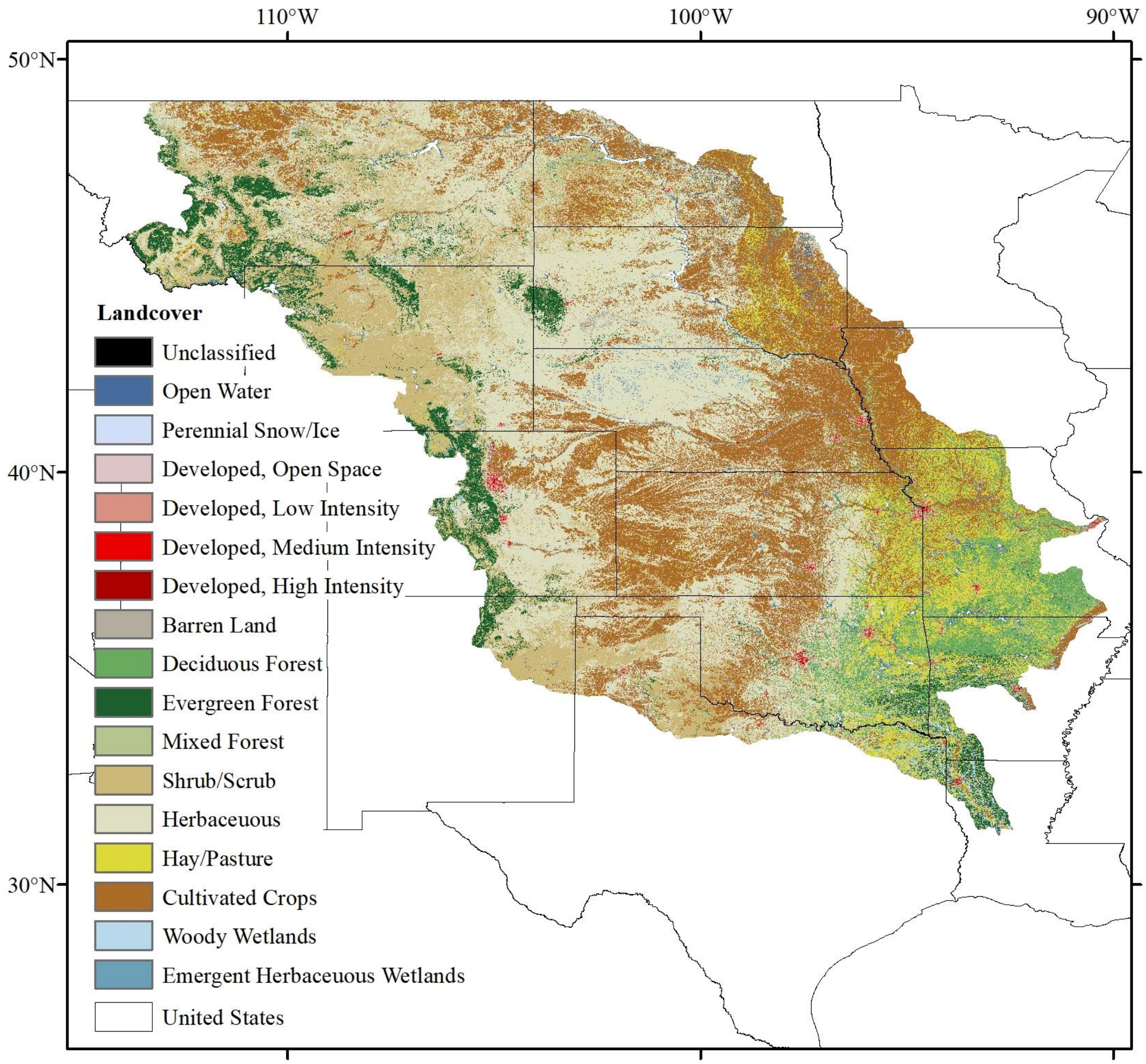

Figure 3.5. The land cover classification for the Missouri and Arkansas-White-Red river basins (Dewitz, 2019).

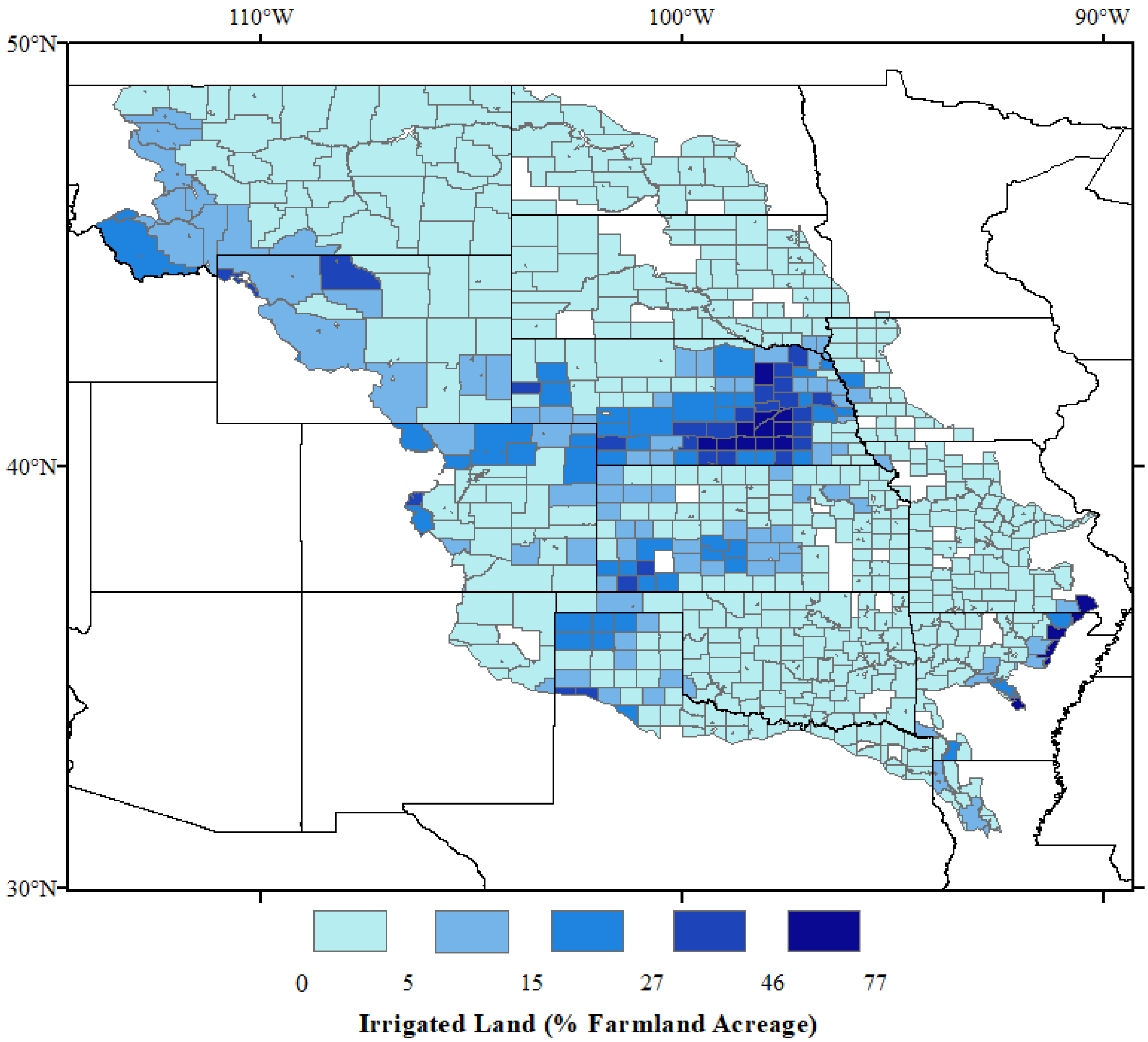

Figure 3.6. Irrigated land per county, as a percentage of farmland, for the Missouri and Arkansas-White-Red river basins (USDA-NASS, 2012).

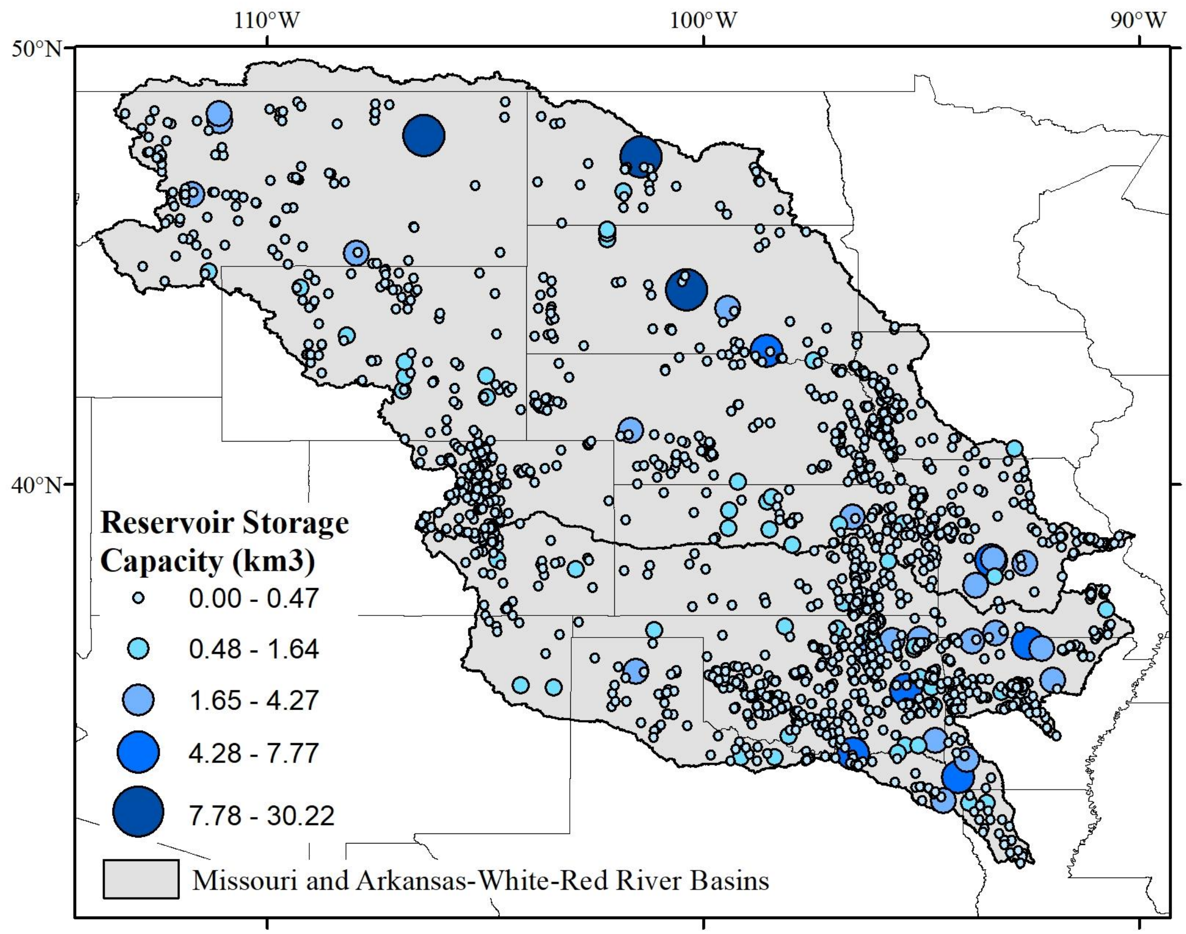

Figure 3.7. Maximum storage capacity (km3) of major dams in the study area. This figure shows dams that are 15 m or more in height, or with a normal storage capacity of 0.006 km3 or more, or with a maximum storage capacity of 0.03 km3 or more from the U.S. Army Corps of Engineers National Inventory of Dams (USGS, 2006).

3.2 SPoRT-LIS

NASA's Marshall Space Flight Center's Short-term Prediction Research and Transition (SPoRT) Center (Jedlovec, 2013) has developed a real-time application of the NASA Land Information System (LIS) for use in experimental operations by both domestic and international operational weather forecasters (Chen and Dudhia, 2001; Ek et al., 2003Case, 2016; Case et al., 2016; Case and Zavodsky, 2018; Zavodsky et al., 2013). LIS is a high-performance land surface modeling and data assimilation system that can be used to run a variety LSMs. This is done by integrating satellite-derived datasets, ground-based observations, and model re-analyses (Kumar et al., 2006; Peters-Lidard et al., 2007). LIS features an Ensemble Kalman Filter algorithm (Evensen, 2003) to conduct land surface data assimilation (Kumar et al., 2008; Kumar et al., 2009) for a variety of datasets and variables such as soil moisture, land surface temperature, and snow (e.g. Liu et al., 2013).

The SPoRT-LIS provides soil moisture estimates with a 3-km grid over a 2-meter deep soil column for the full contiguous United States (McDonough et al., 2018). The SPoRT-LIS top-layer volumetric soil moisture (0-10 cm) estimate was selected for analysis because surface soil moisture provides insight into the structural moisture changes for the entire soil column (Dorigo et al., 2012). The SPoRT-LIS surface soil moisture values were validated and shown to provide accurate volumetric soil moisture estimates at a higher spatial and temporal resolution than is currently provided by other datasets (McDonough et al., 2018). The soil moisture estimates were obtained on a daily scale from 1987 to 2018.

3.3 TIMESAT

TIMESAT is a software package designed to analyze time-series satellite sensor data (Jonsson and Eklundh, 2002). In particular, TIMESAT was created to model the seasonality of dynamic properties of vegetation from NOAA Advanced Very High-Resolution Radiometer (AVHRR) Normalized Difference Vegetation Index (NDVI) data. To do this, TIMESAT implements three processing methods based upon the least-squares fits to the upper envelope of the data (Jonsson and Eklundh, 2004). The three methods include the Savitzky-Golay filter that uses a local polynomial function in the fitting, asymmetric Gaussian, and double logistic which both use ordinary least squares fitting.

TIMESAT was originally designed to analyze satellite-derived vegetation data to estimate seasonality parameters where it is assumed that there would be a peak of greenness during the growing season (Jonsson and Eklundh, 2002). However, the research objective was to analyze the seasonality trends of soil moisture data. The soil moisture seasonality curve mirrors the vegetation curve because increasing leaf area index or greening of vegetation decreases soil water content (Zeng et al., 2018). To modify the SPoRT-LIS soil moisture data to best fit the TIMESAT assumptions made for NDVI data analysis which looks for a concave curve to define the season, the original soil moisture data were inverted using a conversion of one minus the original value (Figure 3.9). This conversion was made so that TIMESAT analyzes the "drying season" of soil moisture that corresponds to the growing season of vegetation. This is important since TIMESAT is designed to look for a concave curve to define the seasonality parameters, where the non-inverted soil moisture data results in a convex curve during the soil drying season. From there, the seasonality parameters, or hydrometrics, were computed for soil moisture. There are 13 hydrometrics computed from TIMESAT that give a range of important seasonal characteristics such as the start and end of the soil drying season and the AWC of the soil during that season (Table 3.1).

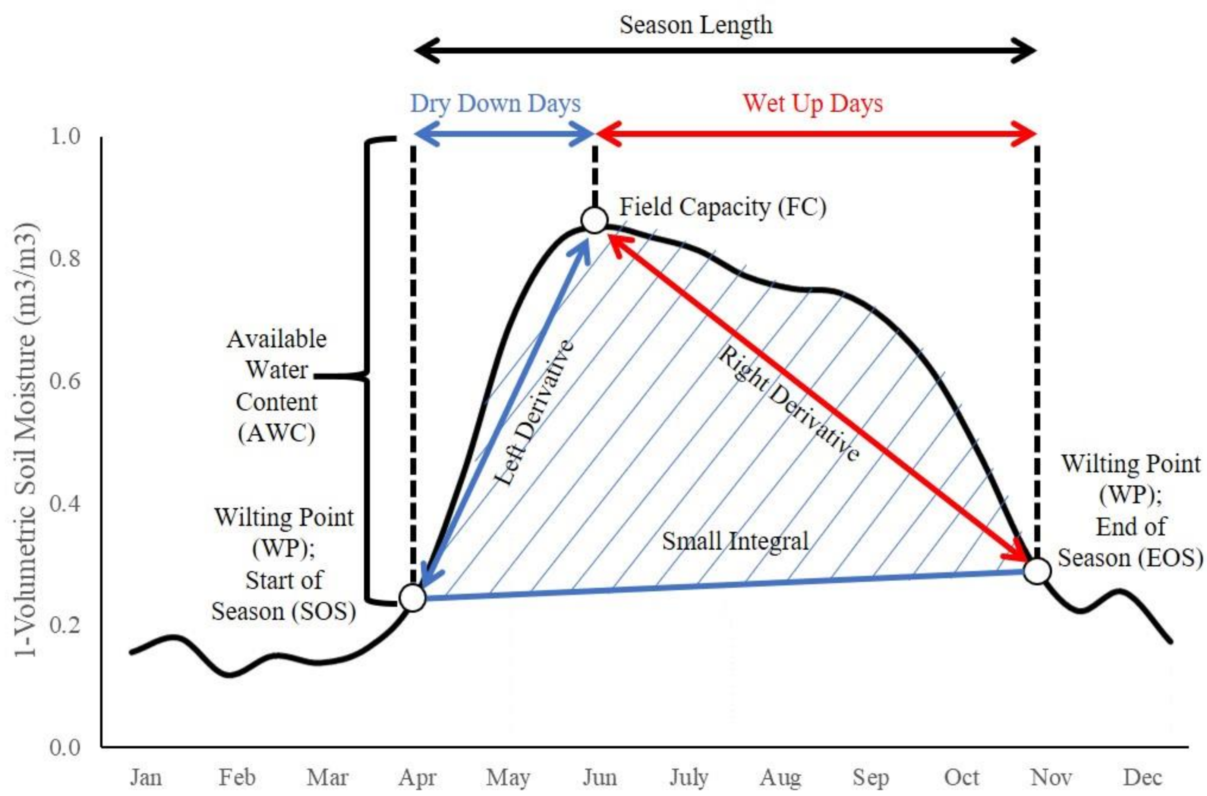

Figure 3.8. The inverted soil moisture data showing the peak of soil moisture dryness along with the corresponding hydrometrics computed in TIMESAT (Shawn Hutchinson, personal communication, 24 Septemenber 2020).

Table 3.1. List, description, and biological significance of TIMESAT soil moisture seasonality parameters (Eklundh and Jonsson, 2015).

Hydrometric	TIMESAT Definition	Soil Moisture Interpretation
Seasonal Amplitude (AWC)	The difference between the maximum value and base value.	The difference in moisture content between Maximum and Base Value can also be described as the available water content.
Base Level (FC)	Average of the left and right minimum values.	The highest percent of soil wetness where 1- soil moisture content is the field capacity point.
End of Season (EOS)	Time at which the right edge has decreased to a user-defined level measured from the right minimum value	End of soil drying season where the evapotranspiration demand is less than the precipitation inputs.
Season Length (SL)	Time from start to end of season.	Length of soil drying season. Defined between Start of Season and End of Season.
Rate of Increase at the Beginning of Season	The ratio of the difference between the left 20% and 80% levels and the corresponding time difference.	Rate of soil drying.
Large Seasonal Integral	Integral of the function describing the season from season start to season end.	Proxy for the 1-soil moisture content without regarding minimum values.
Middle of Season (MOS)	The mean value of the times at which the left edge has increased to the 80% level and the right edge has decreased to the 80% level.	The day when the middle of the season occurs.
Maximum Value (WP)	The largest data value for the fitted function during the season.	The highest percent of soil dryness where 1-soil moisture content corresponds to the wilting point.
Rate of Decreasing at the End of Season	The absolute value of the ratio of the difference between the right 20% and 80% levels and the corresponding time difference.	.Rate of soil wetting
Small Season Integral	Integral of the difference between the function describing the season and the base level from season start to season end.	.Proxy for the 1-soil moisture content while regarding minimum values
Start of Season (SOS)	Time at which the left edge has increased to a user-defined level measured from the left minimum value.	Beginning of soil drying season where the precipitation inputs into the soil become less than evapotranspiration.
Value End of Season	The value at the season end point.	The value of 1-soil moisture content at the End of Season day.
Value Start of Season	The value at the season start point.	The value of 1-soil moisture content at the Start of Season day.

[a]The Timesat parameters selected for analysis in this study are highlighted in grey

3.3.1 Sensitivity Analysis

To determine the model settings, a sensitivity analysis was conducted based upon the soil hydrologic groups. Hydrologic soil properties determine the ability of water to enter the soil and the soil's retention capacity (USDA, 2015). HSGs summarize the soil physical properties that can also impact soil moisture (Famiglietti et al., 1998; Fernandez-Illesca et al., 2001). Soil properties affect soil moisture through the distribution of water by variations in texture, organic matter content, structure, and the existence of macroporosity, thus impact the soil water holding capacity and infiltration rates. Hydrologic soil properties are categorized into four groups (HSG-A, B, C, D) ranging from sandy, loose soils (HSG A) to tighter clay soils (HSG D). The main HSG within the Missouri and Arkansas-White-Red River basins, covering 49% of the area, is HSG B which corresponds to soils with moderate infiltration rates that have moderately well-drained fine to moderately coarse texture (Figure 3.10). Type A soil covers 6% of the study area and is characterized by having a high infiltration rate and well-drained sandy soil. Type C soil covers 17% of the area and has a slow infiltration rate with a slow rate of water transmission. Type D soil covers 22% of the study area and has very slow infiltration rates with a high-water holding capacity.

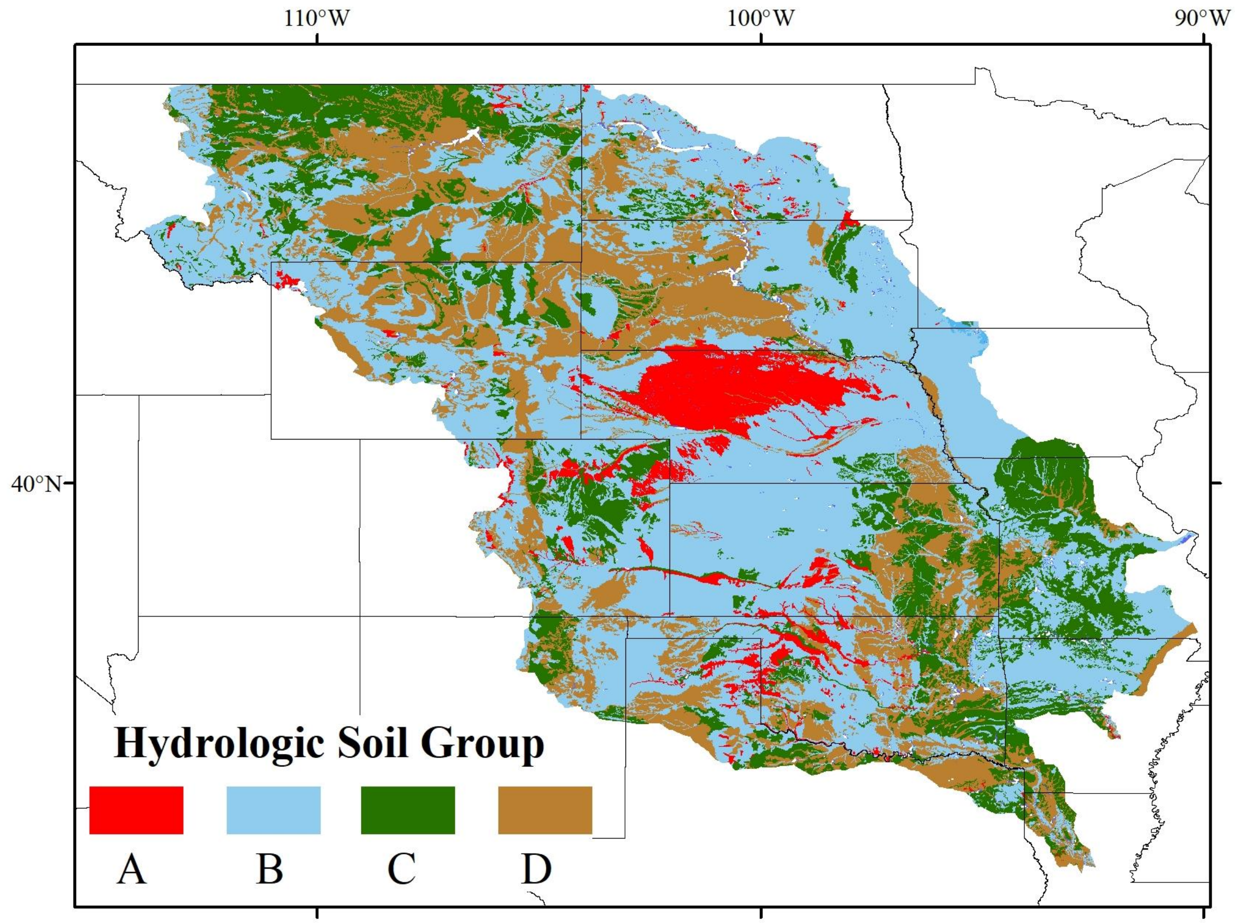

Figure 3.9. The hydrologic soil groups (HSG) throughout the Missouri and Arkansas-White-Red river basins (USDA-NRCS, 2019).

The HSG data was retrieved from the Digital General Soil Map of the United States, STATSGO2, which is broad-based inventory of soils and non-soil areas that occur in a repeatable pattern on the landscape (USDA-NRCS, 2019). STATSGO2 was chosen over SSURGO data because it provides coarse vector data at a spatial resolution of 1:250,000 (approximately 125 m raster resolutaion) where the Soil Survey Geographic Database (SSURGO) provides vector data at a spatial resolution of 1:12,000 (approximately 6 m raster resolution). The HSG data was then transformed into a raster form with a 3-km spatial resolution by defining a threshold where the pixel was classed by a specific HSF if 80% of the HSG polygon layer was within the 3-km pixel. If one of the four HSG did not cover at least 80% of the 3 km area, then the pixel was labled as no data and was not considered in future analysis.

A power analysis was performed to determine the number of cells of each soil type that was necessary for the sensitivity analysis. The sample size was calculated with a 95% confidence interval with a 5% margin of error (Cohen, 1992). The power analysis showed that 371 pixels were required for A HSG, 383 pixels for B HSG, 380 for C HSG, and 381 for D HSG. Pixels for each soil type were randomly selected throughout the study area and used to extract the SPoRT-LIS surface soil moisture data.

The sensitivity analysis was performed using the selected pixels for each HSG as the input for TIMESAT. Different smoothing parameters were tested (Table 3.2). These include selecting between three smoothing algorithms, Gaussian, Logistic, or Savitsky-Golay, common settings, and class-specific settings. The parameters that showed the largest impact on the data were the Savitzky-Golay window size and the season start, and season stop values (Pockrandt, 2012). The window size determines the number of days that the curve is smoothed over and the season start and stop values determine the point along the curve that defines when the soil drying season begins and ends. The parameters that were kept constant throughout the sensitivity analysis were amplitude cutoff of 0, median spike method with a spike parameter of 4, seasonality parameter of 1, number of envelope iterations of 2, adaption strength of 3, and a Savitzky-Golay fitting method (Eklundh and Jonsson, 2010; Pockrandt, 2012).

Table 3.2. TIMESAT setting parameters and their descriptions (Eklundh and Jonsson, 2015).

Setting	Description
Fitting Method	
Asymmetric Gaussian	Fitting with least-squares where data is fitted to a Gaussian function
Double Logistic	Fitting with least-squares where data is fitted to a double logistic function
Savitsky-Golay	Fitting with local polynomial functions
Common Settings	
Data range	The lower and upper data values for the valid range. Data outside the range will be assigned weight 0
Amplitude value	Cutoff for low amplitude. Series with amplitude smaller than this value will not be processed. 0 processes all data
Spike method	3 = weights from STL multiplied with original weights, 2 = weights from STL, 1 = method based on median filtering, 0 = no spike filtering
Spike value	If spike method = 1 the spike value determines the degree of spike removal. A low value will remove more spikes
Class-Specific Settings	
Seasonal parameter	A value of 1 will attempt to fit one season per year, a value of close to zero will attempt to fit two seasons
No. of envelope iterations	No. of iterations for the upper envelope adaptation 1, 2, 3
Adaption strength	Strength of the envelope adaptation. 10 is the maximum strength
Window size	Half window for the Savitzky-Golay filtering. A large value of the window will give a high degree of smoothing.
Start of season method	Method for determining start/end of season. 2 = start and end where the fitted curve crosses a threshold value. 1 = start and end where the fitted curve reaches a proportion of the seasonal amplitude measured from the left/right minimum value
Season start	Value for determining the season start. If start method = 1 the values must be between 0 and 1
Season stop	Value for determining the season stop. If start method = 1 the values must be between 0 and 1

Fourteen different iterations of smoothing parameters were tested for each HSG, and the top five iterations that resulted in the least number of no data seasons within the TIMESAT

Graphical User Interface were selected for further analysis (Figure 3.11). The iterations ranged in window size from 2-30 days and the season start and stop values ranged from 0.7-0.85 volumetric soil moisture content (Figure 3.12). The selected iterations were then compared based upon 5 hydrometrics including, season start day, season stop day, amplitude value, base value, and maximum value. The final TIMESAT smoothing parameters were determined by which smoothed curve fit the data well along if the hydrometrics resulted in reasonable values based upon the physical characteristics of the soil.

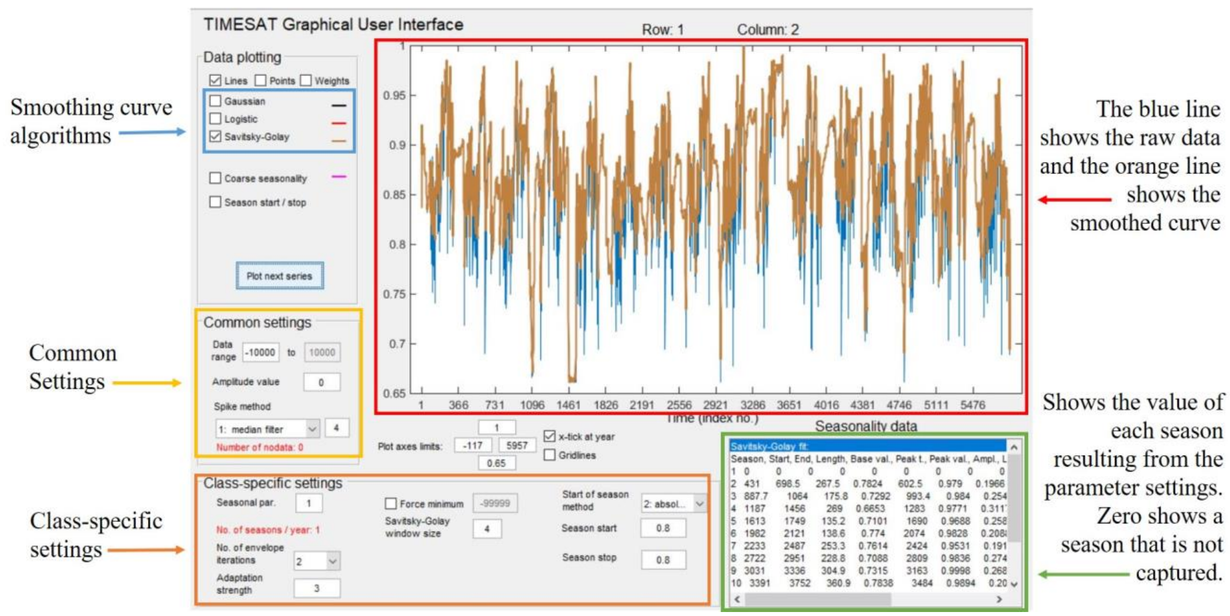

Figure 3.10. The TIMESAT Graphical User Interface showing the three types of settings, smoothing curve algorithms, common settings, and class-specific settings, that users select from. The red box shows the smoothing curve window where the blue curve is the raw data and the orange curve is the smoothed curve. The green box also shows the resulting hydrometric values for each season where a zero shows that the parameters are not capturing that season.

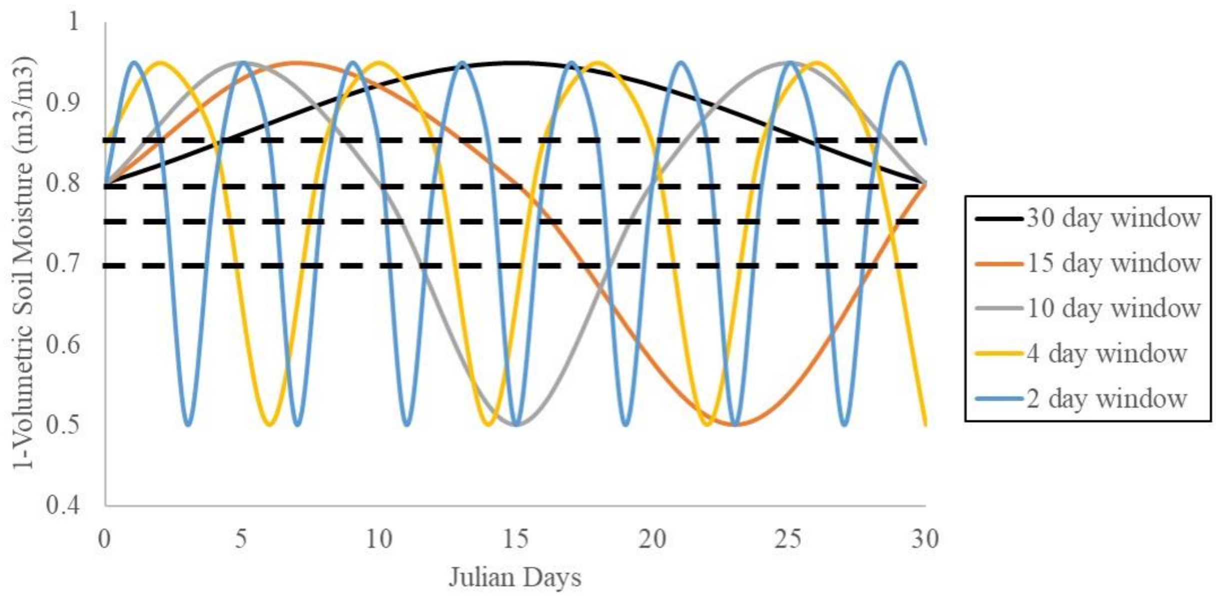

Figure 3.11. The variation in the TIMESAT window size parameter and the season start and top values for the fourteen iterations in the sensitivity analysis. The curved lines show different window size smoothing curves and the dotted black lines show the thresholds selected for the season start and stop values.

From the sensitivity analysis, the smoothing parameters used for HSG A were a window size of 2, a season start value of 0.8, and a season stop value of 0.85 volumetric soil moisture content. For HSGs B, C, and D the parameters used were a window size of 2 and a season start and stop of 0.75 volumetric soil moisture content. From there, TIMESAT ran with the inverted surface soil moisture values along with a mask layer defining each pixel's set smoothing parameters based upon the sensitivity analysis. TIMESAT then defined 13 hydrometrics for surface soil moisture throughout the Missouri and Arkansas-Red-White River basins for each year from 1987-2018, however, the focus of this research is only on 7 of those hydrometrics (Table 3.1). From there, the median value from 1987-2018 for each pixel was calculated since the temporal distribution for each hydrometric was non-normal. The summary hydrometrics were then used for statistical analysis.

3.4 Statistical Analysis

To better understand what factors are impacting the hydrometric results, statistical analysis was completed to determine the impact that climate and HSG have on hydrometrics. Climate variables such as precipitation and temperature can have a major impact on the water balance and soil moisture by impacting ET rates, runoff, and soil water storage capacity (Famiglietti et al., 1998; Akuraju et al., 2016). Soil physical properties impact soil moisture through the distribution of water by variations in texture, organic matter content, structure, and the existence of macroporosity, thus impact the soil water holding capacity and infiltration rates (Famiglietti et al., 1998; Fernandez-Illesca et al., 2001).

Therefore, mean annual temperature and mean annual precipitation from the PRISM Climate Group (2015) were chosen as independent variables along with HSG. Mean annual temperature and precipitation were chosen over seasonal temperature and precipitation because the soil water storage is being characterized over an annual water budget by assuming an annual hydrologic cycle. From there, seven hydrometrics were chosen for the analysis based upon their importance to plant growth where amplitude defines the AWC of the soil, base value shows the FC of the soil, and maximum value defines the WP of the soil (Table 3.1). Also, the season start, middle, and end provide information on when and how long the soil is drying. For AWC, FC, and WP, the statistical model was given by a beta distribution

$$[y_l|\mu_l,\phi] \equiv Beta(\mu_l,\phi), \tag{3.1}$$

with a logistic link function, and the SOS, MOS, EOS, and SL hydrometrics are represented by a gamma distribution

$$[y_l|\mu_l,\phi] \equiv Gamma(\mu_l,\phi)\ , \tag{3.2}$$

with a log link function where y is the true median value of the hydrometric over the years1987-2018 at the l^{th} location. The median value of the hydrometric was used instead of the mean because the temporal distribution was not normally distributed. The location, l, is defined as the central point in each three-kilometer grid cell given a latitude and longitude. The dispersion parameter is given by ϕ. The expected value of the hydrometric is represented by μ_l and is defined as

$$\mu_l = \beta_0 + \beta_1 x_{1,l} + \beta_2 x_{2,l} + \beta_3 x_{3,l} + \eta_l, \tag{3.3}$$

where $x_{1,l}$ is the mean annual precipitation value in mm for the three-kilometer grid cell at location l, $x_{2,l}$ is the mean annual temperature in °C for the three-kilometer grid cell at location l, $x_{3,l}$ is the HSG for the three-kilometer grid cell at location l, and $\eta_{l,t}$ is the site level effect at the l^{th} location.

The beta distribution was chosen as the distribution for the AWC, FC, and WP hydrometrics because the support is continuous from 0 to 1 and these hydrometrics are measured in volumetric soil moisture content. The gamma distribution was then chosen as the distribution for the SOS, MOS, EOS, and SL because these hydrometrics are represented by the number of days, and the gamma distribution has positive continuous support.

3.4.1 Ecoregions and Hydrometrics

In addition to looking at the impact of climate and soil variables, ecoregions were compared to the hydrometric results to determine if hydrometrics can be distinguished between each ecoregion. Ecoregions are defined by the biotic and abiotic phenomena that define ecosystems and their function (Omernik 1987). This includes geology, topography, soils, vegetation, climate, and land use. There are 25 level III ecoregions within the study area (Figure 3.13).

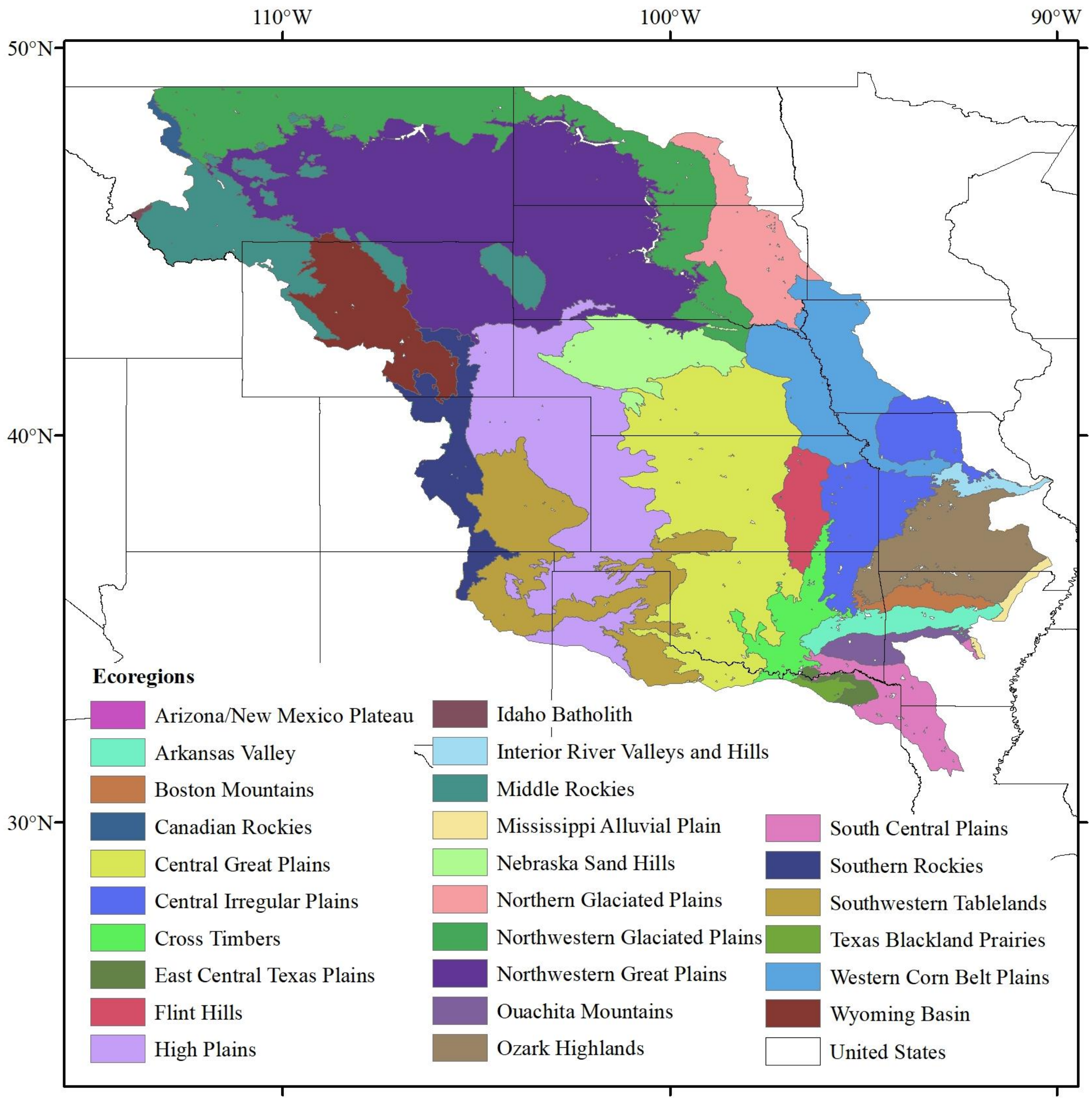

Figure 3.12. Level III ecoregions within the Missouri and Arkansas-White-Red river basins (EPA, 2013).

To be able to determine if the hydrometrics can be distinguished between each ecoregion, density plots were created to represent the distribution of the median SOS, MOS, and EOS for each ecoregion within the study area. Then the hydrometrics were analyzed with the Kruskal Wallis test to determine if there is a significant difference between each ecoregion. The Kruskal

Wallis test was used because of the non-uniform distribution of the hydrometrics temporal component. From there, the pairwise Wilcox test was performed to determine what pairs are significantly different from each other.

Chapter 4 - Results and Discussion

Results from the TIMESAT analysis showed that it captured intra-annual soil moisture variability and that the seasonal soil moisture characteristics can be characterized at a coarse spatial resolution of 3 km. Furthermore, the statistical analysis highlighted the impact of soil texture and climate on the availability of soil moisture where it was shown that precipitation drives soil water storage capacity and temperature is the driver of the seasonal timing of soil water storage. Additionally, it was found that for a majority of the ecoregions, the seasonal soil moisture characteristics can be distinguished between each region, highlight the impact of soil moisture availability on the biotic and abiotic phenomena.

4.1 Field Capacity, Wilting Point, and Available Water Content TIMESAT Analysis

The results from the TIMESAT analysis showed that A soils had the lowest WP ($\theta_{WP\ median}$ = 0.03 m^3/m^3, range 0-0.07 m^3/m^3) and lowest FC ($\theta_{FC\ median}$ = 0.25 m^3/m^3, range 0.20-0.31 m^3/m^3; Figure 4.1). HSGs B and D have similar WP and FC values with C soils having the highest WP ($\theta_{WP\ median}$ = 0.13 m^3/m^3, range 0.04-0.22 m^3/m^3) and FC ($\theta_{FC\ median}$ = 0.36 m^3/m^3, range 0.23-0.48 m^3/m^3).

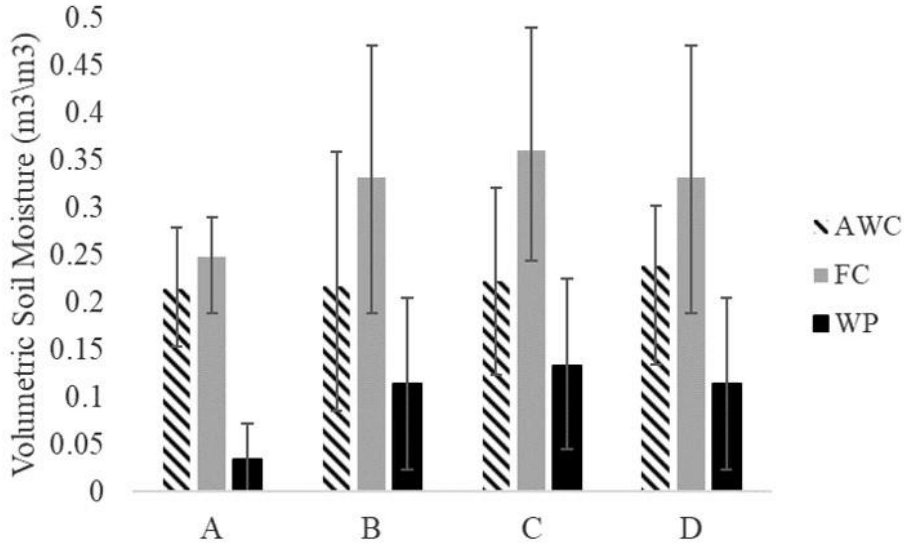

Figure 4.1. The median available water content (AWC), field capacity (FC), and wilting point (WP) from 1987-2018 showing the spatial median values for each hydrologic soil group (HSG) with error bars showing the upper and lower extreme values.

The permanent WP happens at approximately -1500 kPa which is the stage at which the soil contains some water, but it is difficult for the roots to extract from the soil. The median values for each HSG fell within expected values, except for HSG D where the WP is less than the assumed range (Table 4.1). This could be because D soil groups have small particle sizes that are held tightly together resulting in them being less responsive to precipitation impacts (USDA, 2015). Also, clay soils are typically found at higher elevations and on hillslopes compared to sandier soils (Collins and Foster, 2008). This variation in HSG D location and structure impact the model's ability to accurately fit a smoothing curve over the variations of soil moisture.

Table 4.1. Field capacity (FC), wilting point (WP), and available water content (AWC) for different hydrologic soil groups (HSG) comparing the expected values (NRCS, 1998; Saxton and Rowls, 2006) and the observed upper and lower extreme values from the TIMESAT analysis.

	θ_{FC} (m^3/m^3)		θ_{WP} (m^3/m^3)		θ_{AWC} (m^3/m^3)	
HSG	Expected	Observed	Expected	Observed	Expected	Observed
A	0.10-0.15	0.20-0.31	0.0-0.05	0.0-0.07	0.0-0.10	0.15-0.28
B	0.15-0.20	0.19-0.48	0.05-0.10	0.02-0.20	0.10-0.15	0.086-0.36
C	0.25-0.40	0.23-0.48	0.05-0.20	0.04-0.22	0.10-0.20	0.12-0.32
D	0.35-0.45	0.19-0.48	0.25-0.30	0.02-0.20	0.15-0.25	0.13-0.30

Field capacity is assumed to be approximately -33 kPa which defines the point where the soil contains the maximum amount of water after excess drainage has occurred (Figure 1.1; NRCS, 1998). The median values of FC for HSGs C and D fell within the expected range, but the median FC for HSGs A and B fell above the expected range (Table 4.1). The results also showed that for all HSGs, the median AWC was approximately the same ($\theta_{AWC\ median}$ = 0.22 m^3/m^3) where HSG A had the shortest range between the upper and lower extreme values from the box-plot (Figure 4.1; Appendix A). It was unexpected that all soil types would have roughly the same AWC value. This could be because of the varying scales of soil moisture observations and the scale at which soil water movement occurs. The expected values of soil moisture for FC, WP, and AWC is observed with in situ soil samples and looks at the soil pore scale (Saxton and Rowls, 2006) where the SPoRT-LIS data considers a spatial scale of 3 km at the near-surface soil layer (0-10 cm). This demonstrates the limitation of characterizing the soil storage capacity using coarse soil moisture observations.

4.2 Soil Moisture Start, Middle, and End of Season TIMESAT Analysis

The soil dry season was hypothesized to follow the growing season. The growing season is defined by the period between the last frost of spring and the first frost of fall, where the air temperature drops below the freezing point (Kukal and Irmak, 2018). This shows that for the Missouri and Arkansas-White-Red River basins, the last spring frost occurs between March 2nd to May 30th, and the first fall frost occurring September 8th to November 20th. The results show that the median soil moisture start of season for each HSG lags behind the growing season start, but the median soil moisture end of season occurs before the end of the growing season (Figure 4.2; Appendix A). The results also showed that the lower extreme MOS values are later in the year than the EOS values, however when looking at the full set of data, only 3 MOS pixels (0.0005% of the study area) are greater than the EOS. Also, the median MOS and EOS values are closer than expected for all HSGs. This is showing that the soil is wetting up quicker than it dries down.

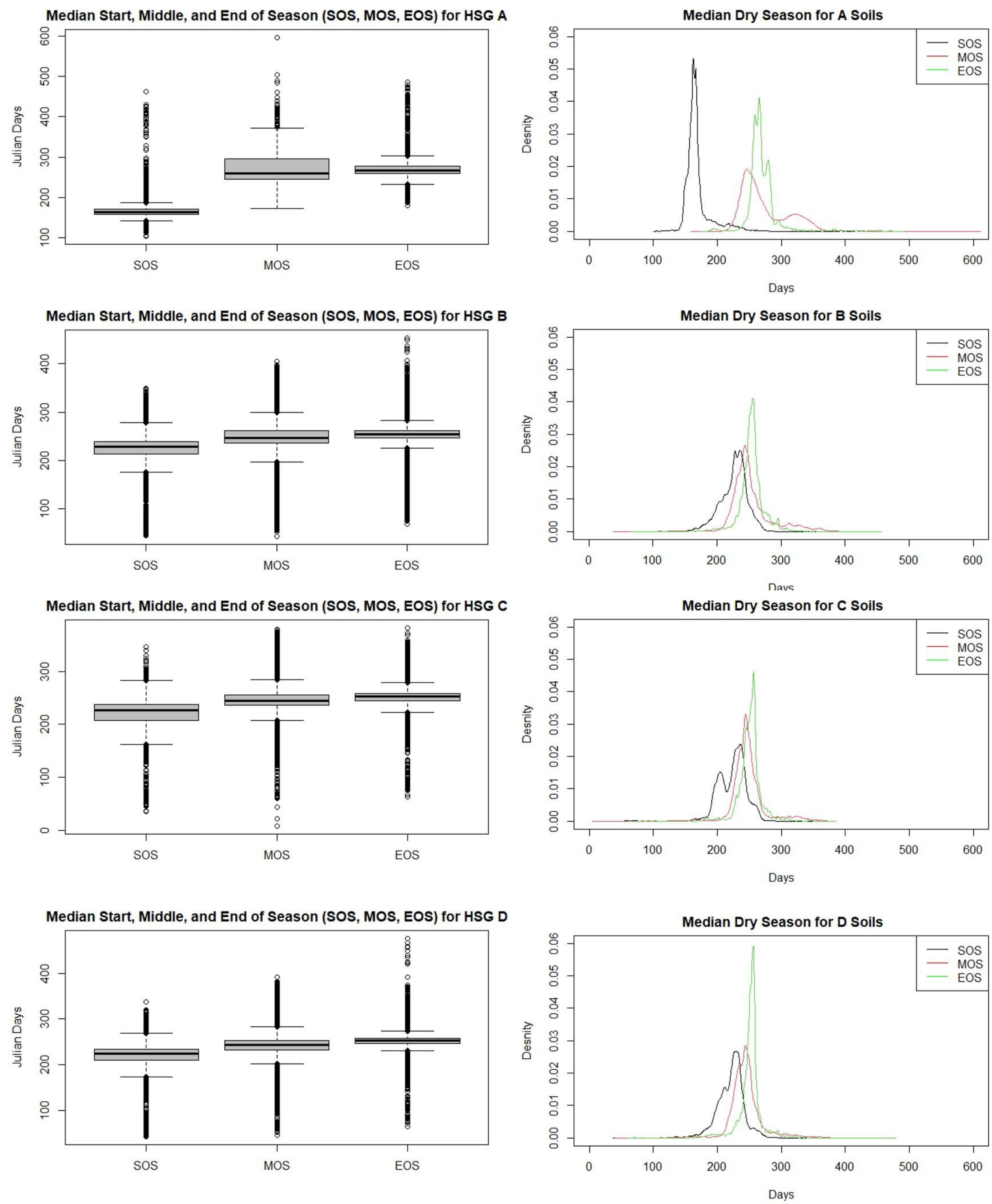

Figure 4.2. Box-plots and density plots showing the distribution of the median soil season's start (SOS), middle (MOS), and end (EOS) for each hydrologic soil group (HSG).

Density plots and box-plots of the SOS, MOS, and EOS demonstrate the range and distribution of each hydrometric (Figure 4.2). For HSG A, all three are right-skewed with a tri-modal EOS and a long season length. B soil SOS, MOS, and EOS distributions are normally distributed with each occurring close in time to each other showing a short season length. C soil distributions are slightly left-skewed with a bimodal SOS. Then for HSG D, the SOS, MOS, and EOS are normally distributed and show a shorter season length. For B, C, and D groups, EOS has a narrow distribution where the A soil group SOS has a narrower distribution. The SOS, MOS, and EOS hydrometrics were then analyzed with the Kruskal Wallis test to determine if there is a significant difference between the HSGs. This showed that all HSGs were significantly different from each other, demonstrating the impact that the physical characteristics of the soil have on soil moisture. The box-plots also gave insight into what values are considered outliers for each hydrometric within each HSG (Appendix A).

4.3 Statistical Analysis

To further understand the impact of soil and climate characteristics and address the variations in results from what was hypothesized, the regression model with a space site effect was used. The results show that for season length (SL), there was a non-significant negative impact for precipitation where every unit increase of precipitation would result in a log decrease of season length day by -1.211 x 10-5 (Table 4.3). All other impacts were significant.

Table 4.2. A summary of each hydrometrics estimated regression coefficients and their corresponding p-values.

Hydrometric	Independent variable	Regression Coefficient	Estimate	p-value
SOS	Precipitation	β_1	6.17E-05	<2e-16
	Temperature	β_2	6.48E-03	<2e-16
	HSG	β_3	2.63E-02	<2e-16
MOS	Precipitation	β_1	-1.47E-04	<2e-16
	Temperature	β_2	-1.14E-02	<2e-16
	HSG	β_3	-7.11E-03	<2e-16
EOS	Precipitation	β_1	2.81E-05	<2e-16
	Temperature	β_2	4.18E-03	<2e-16
	HSG	β_3	-7.61E-03	<2e-16
SL	Precipitation	β_1	-1.21E-05	0.467
	Temperature	β_2	-1.55E-02	<2e-16
	HSG	β_3	-1.96E-01	<2e-16
AWC	Precipitation	β_1	3.54E-04	<2e-16
	Temperature	β_2	-2.58E-02	<2e-16
	HSG	β_3	4.53E-03	<2e-16
FC	Precipitation	β_1	5.51E-04	<2e-16
	Temperature	β_2	-2.30E-02	<2e-16
	HSG	β_3	6.96E-02	<2e-16
WP	Precipitation	β_1	4.63E-04	<2e-16
	Temperature	β_2	-1.01E-02	<2e-16
	HSG	β_3	1.68E-01	<2e-16

Three model checking techniques were performed to determine if any assumptions were violated. The first model checking technique performed was the test the assumption of normally distributed residuals with a mean of zero. This assumption was violated by MOS, SL, AWC, and WP hydrometrics which all had a residual mean of zero but were not normally distributed (Appendix A). However, it is found that violating the normaility of residuals assumption is rarely problematic for hypothesis testing and parameter esitamtes (Warton et al., 2016). The second assumption tested was spatially auto-correlated residuals by looking at the semivariograms for each model. This showed that all seven models violated this assumption. This assumptions is often violotated for environmental and ecological data because the data itself is inheratnly spatially autocorrelated as observations closer in space are more similar than those farther apart and (Miralha and Kim, 2018; Gaspard et al., 2019). The third assumption tested was that the space site effect is independent of the predictor variables, $x_{i,l}$. The concurvity test showed that all seven models violated this assumption, however, all models converged successfully. Given that the soil moisture data is representing the natural system, it was expected that some model assumptions would be violated, thus it was determined that even with the violation of assumptions, the results were valid and can be used for further analysis.

4.3.1 Field Capacity, Wilting Point, and Available Water Content Statistical Analysis

The results showed that precipitation and HSG had positive impacts on WP and FC hydrometrics where the temperature had a negative impact, however, the precipitation has the largest impact (Table 4.2). These results align with what is hypothesized since precipitation would increase soil moisture (Appendix B). Similarly, the opposite is true for temperature where an increase in temperature results in a decrease in soil moisture content due to increased ET rates (Famiglietti et al., 1998; Cook et al., 2014; Ault et al., 2016). As for HSG, the finer the soil is the more water holding capacity it has, increasing the soil moisture content (NRCS, 1998; Saxton and Rowls, 2006; USDA, 2015).

These trends can be seen in Figures 4.3 and 4.4 where HSG A stands out for having the lowest FC and WP and HSG D stands out for having high WP and FC values. In the western portion of the study area, there are high FC values and low temperatures (Figure 3.4), aligning with the results from the statistical analysis. This region also experiences low precipitation (Figure 3.2), which does not align with the results. From this, the temperature is shown to be the main driver of FC in the western portion of the study area compared to precipitation. As for WP (Figure 4.4), a majority of the region follows HSG patterns (Figure 3.9). However, for each hydrometric, some areas cannot be explicitly explained by increased precipitation, decreased temperature, or HSGs alone. This is seen in the eastern portion of the study area for FC where high soil moisture values occur that cannot be seen in the climate and soil trends. Additionally, developed land cover were deemed outliers and were masked out of the WP map (Figure 3.5). This shows that additional variables such as land cover, land management, elevation, and slope should be considered in future analysis.

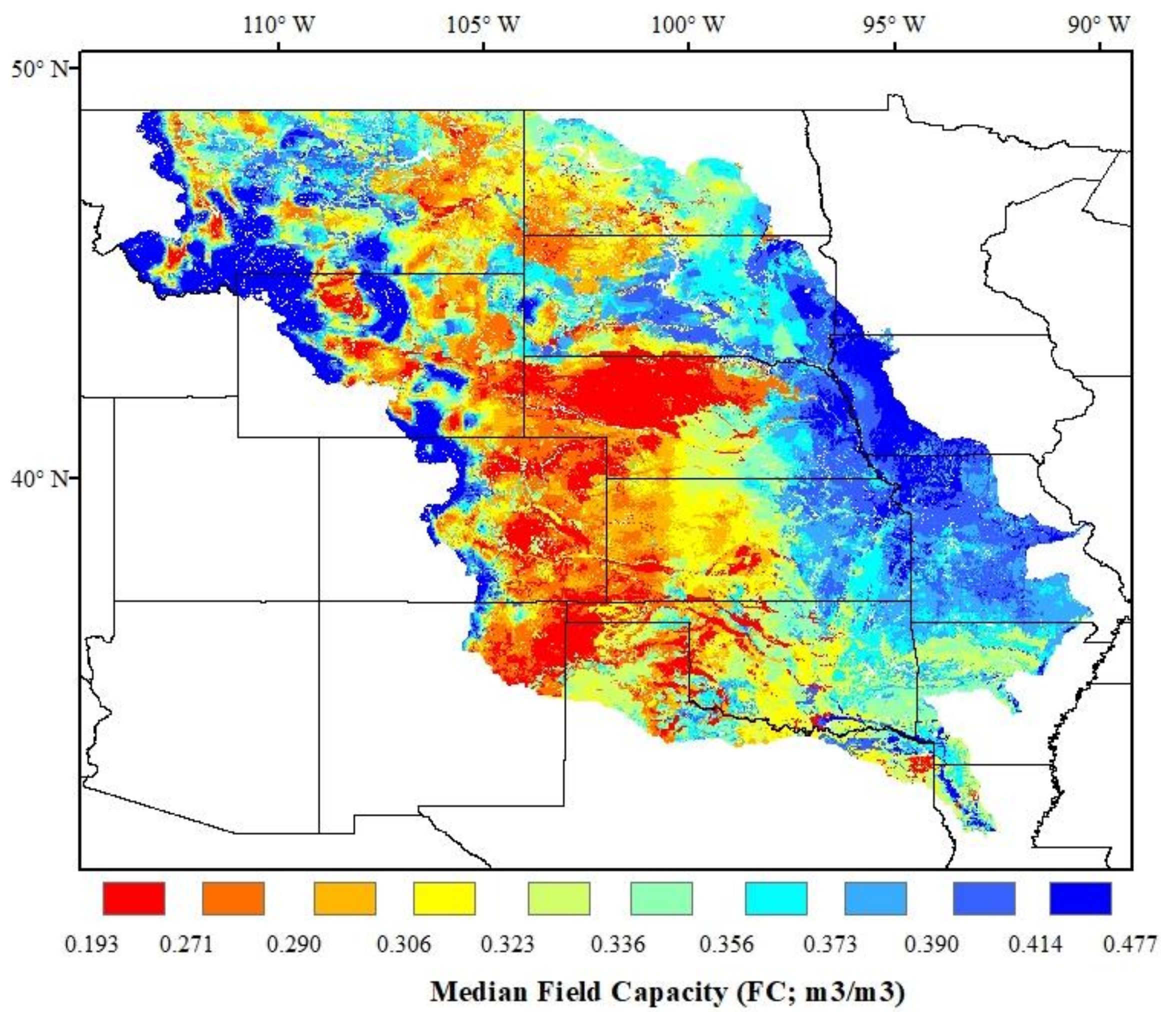

Figure 4.3. A map of the spatial variations of the median field capacity (1-BAS) in volumetric soil moisture content (m3/m3) for the Missouri and Arkansas-White-Red river basins with outliers exluded.

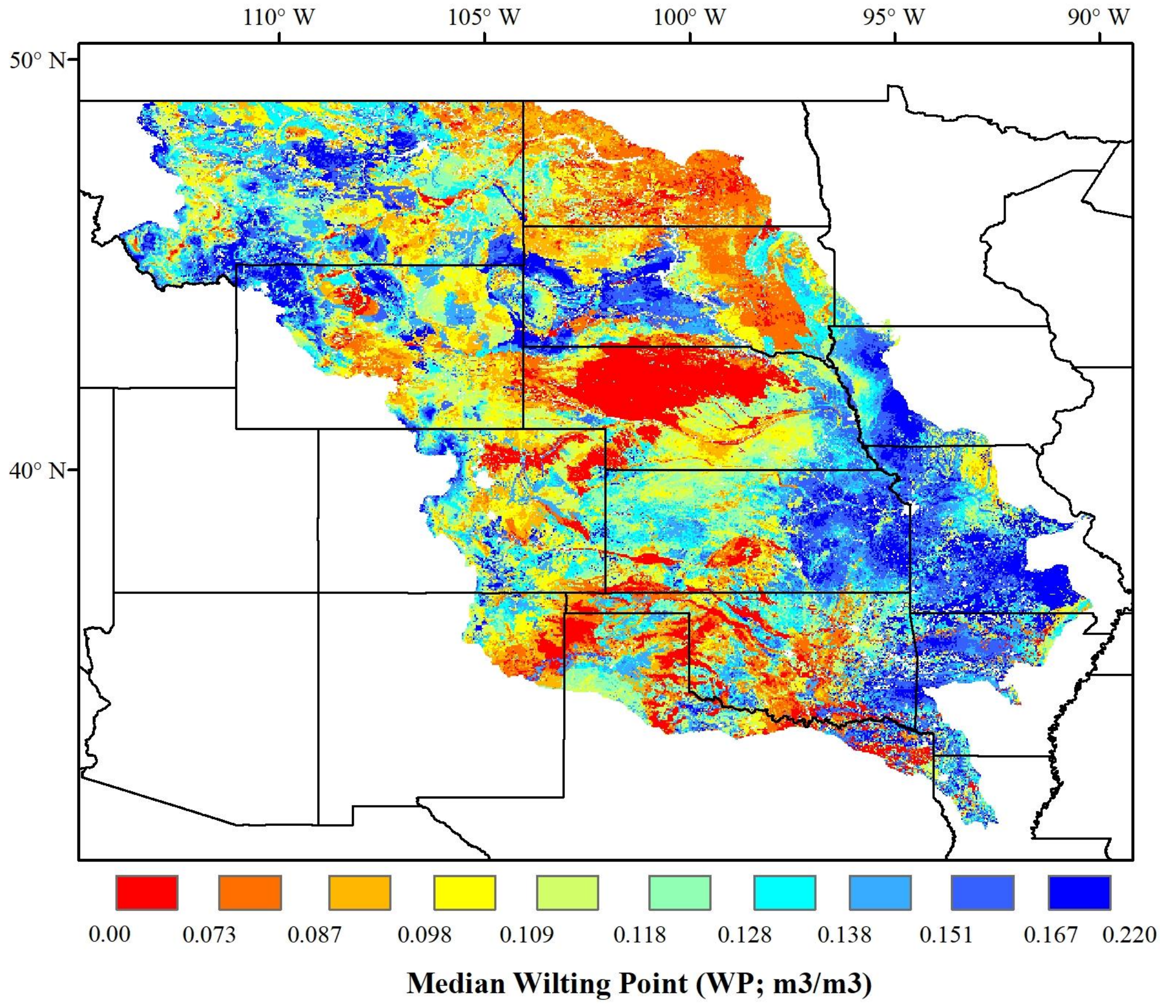

Figure 4.4. A map of the spatial variations of the median wilting point (1-MXF) in volumetric soil moisture content (m3/m3) for the Missouri and Arkansas-White-Red river basins with outliers exluded.

The AWC of the soil, shown with the AMP hydrometric, resulted in positive impacts from precipitation and HSG but a negative impact from temperature (Table 4.2). These results align with what was hypothesized since precipitation would increase the amount of water in the soil where the temperature would result in the opposite (Appendix B). Also, as you go from A, B, C, to D soil groups, the water holding capacity of the soil increases. However, the largest impact was seen with precipitation.

These trends can be seen in the southwestern portion of the region (Figure 4.5). The area shows that a combination of low mean annual precipitation (142-469 mm; Figure 3.2) and high mean annual temperatures (10.1-13.9 °C; Figure 3.4) play a role in the low AWC (0.12-0.17 m3/m3). However, in other areas, these trends do not explicitly stand out by looking at the spatial variations of the median AWC. For example, the northwestern and eastern portions of the study show the greatest AWC. The increased AWC trend in the west aligns with the decreased temperature trend for that region but does not demonstrate the positive impact from precipitation similarly to the FC. As for the increased AWC in the eastern portion of the river basins, neither, temperature, precipitation, or HSG patterns seem to play a dominant role. Instead, the land cover could the driving factor since the area is defined as cropland (Figure 3.5). From this, land management might be causing irrigation and other farming practices to out weight the natural climate drivers of soil moisture (Jasa, 2013; Lawston, et al., 2017). Also supporting the impact of land cover on the AWC of the soil are the areas defined as being developed, which show the least amount of AWC for the region. This could be because urban areas have decreased infiltration and evapotranspiration rates due to surface sealing and increased impervious areas (Wessolek and Facklam 1997; Easton et al., 2007).

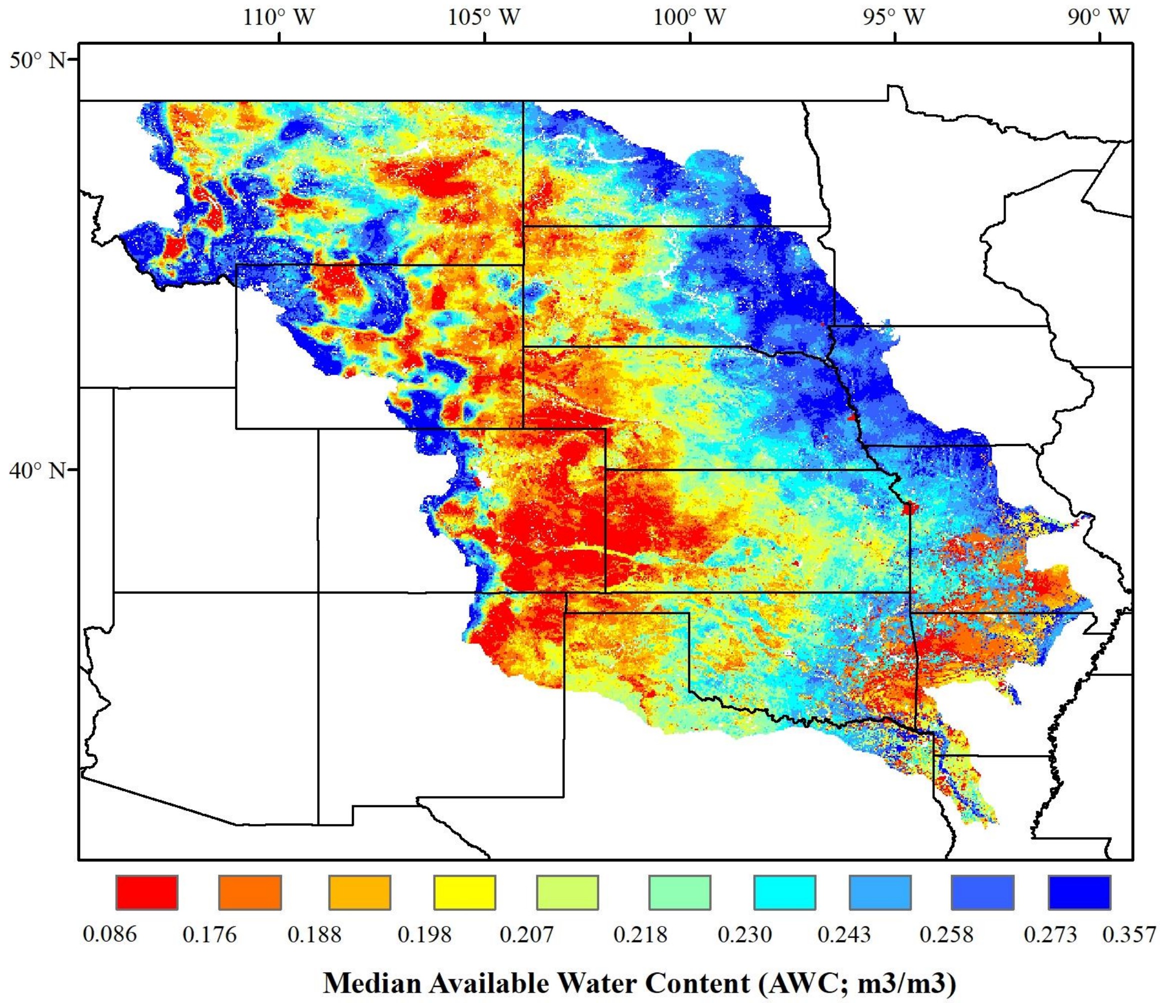

Figure 4.5. A map of the spatial variations of the median available water content (AMP) in volumetric soil moisture content (m3/m3) for the Missouri and Arkansas-White-Red river basins with outliers excluded.

4.3.2 Start, Middle, and End of Season Statistical Analysis

The start of season hydrometric begins when the soil moisture wets up to 0.2 volumetric soil moisture content and ends when it dries back down to 0.15 volumetric soil moisture content for A soils, and for B, C, and D soils, the season begins at 0.25 volumetric soil moisture content ends when it dries back down to this level. A positive impact on the start of the season would mean that the soil season started later where a negative impact would mean that the independent variable is resulting in the soil season beginning earlier in the year.

The results showed that precipitation, temperature, and HSG all had positive effects on the SOS (Table 4.2). The greater the mean annual precipitation an area receives, the earlier in the year the SOS occurs supporting what was hypothesized since the soil season is defined by the soil drying down (Appendix B). As for temperature, the hotter an area is on average throughout the year, the later in the year the soil dry season begins. It was expected that the SOS would occur earlier as the hotter temperatures would dry the soil out sooner. This variation in what was expected and what occurred could be because the temperature variable represents average annual temperature, and it is says nothing about the seasonal variations in temperature that would drive soil moisture drying (Famiglietti et al., 1998; Cook et al., 2014; Ault et al., 2016). As For HSG, the finer the soil particles, the later in the year that the soil dry season occurred. This is because a more clayey soil has increased water holding capacity showing that the soil stays wetter, thus taking longer for the dry season to begin.

The middle of season is defined as the average of the days between the last 20% of the start of season and the first 20% of the end of season days (Figure 3.8; Table 3.1). A positive impact on the MOS would mean that it is occurring later in the year, but a negative impact would show that the climate or soil variables are causing the MOS to happen earlier in the year. The results showed that all of the variables had a negative impact on the MOS (Table 4.2). It was hypothesized that precipitation and HSG would have a positive impact resulting in the MOS occurring later in the year, where the temperature would have a negative impact resulting in the MOS occurring sooner in the year. The hypothesis was only met for temperature (Appendix B). One reason for the MOS to occur earlier in the year with increased precipitation is that the increased precipitation would shorten the season length, thus causing the MOS to happen earlier.

As for HSG, the negative impact was significant but the magnitude of the impact between an A soil and a D soil was approximately 8 days.

For the end of the season, a positive impact from an independent variable would mean that the soil is staying dry later in the year, but a negative impact would mean that it starts to wet up earlier in the year. For the EOS hydrometric, precipitation and temperature both had positive effects where HSG showed a negative effect (Table 4.2). The positive impact from precipitation shows that the soil is wetting back up later in the year (Appendix B). This is the opposite of what is hypothesized since increased precipitation would cause the soil to wet up not stay dry for longer. But, the precipitation is an annual average and says nothing in regards to the seasonal variations of precipitation. As for the positive impact from temperature, it shows that increased temperatures drive the soil to stay dry longer. This goes along with what is to be expected since increased temperatures can result in increased ET rates, thus a decrease in soil moisture. The negative impact from HSG on EOS aligns with the positive impact on the SOS showing that a sandier soil is more responsive to precipitation due to its larger particle size and increased pore space in the soil, however, sandy soils do not hold water well, thus resulting in a long dry season (USDA, 2015).

For the SOS, MOS, and EOS hydrometrics the largest impact was seen with the temperature where precipitation showed a slightly smaller impact on the SOS. This aligns with the climatologic growing season that is defined by temperature gradients (Kukal and Irmak, 2018). For these three hydrometrics, the earlier days occurred on the western portion of the study area (Figures 4.6; Figure 4.7; Figure 4.8) where the cooler temperatures occur along with higher elevations (Figure 3.4; Figure 3.3). An exception to this was seen with SOS where HSG A stands out for having the earliest start of season day of the study area. Another distinct feature that shows up in all three hydrometrics is the southwestern portion of the study area that shows the hydrometrics occurring later in the year. It is hypothesized that a combination of hotter temperatures and decreased precipitation could be a cause for this.

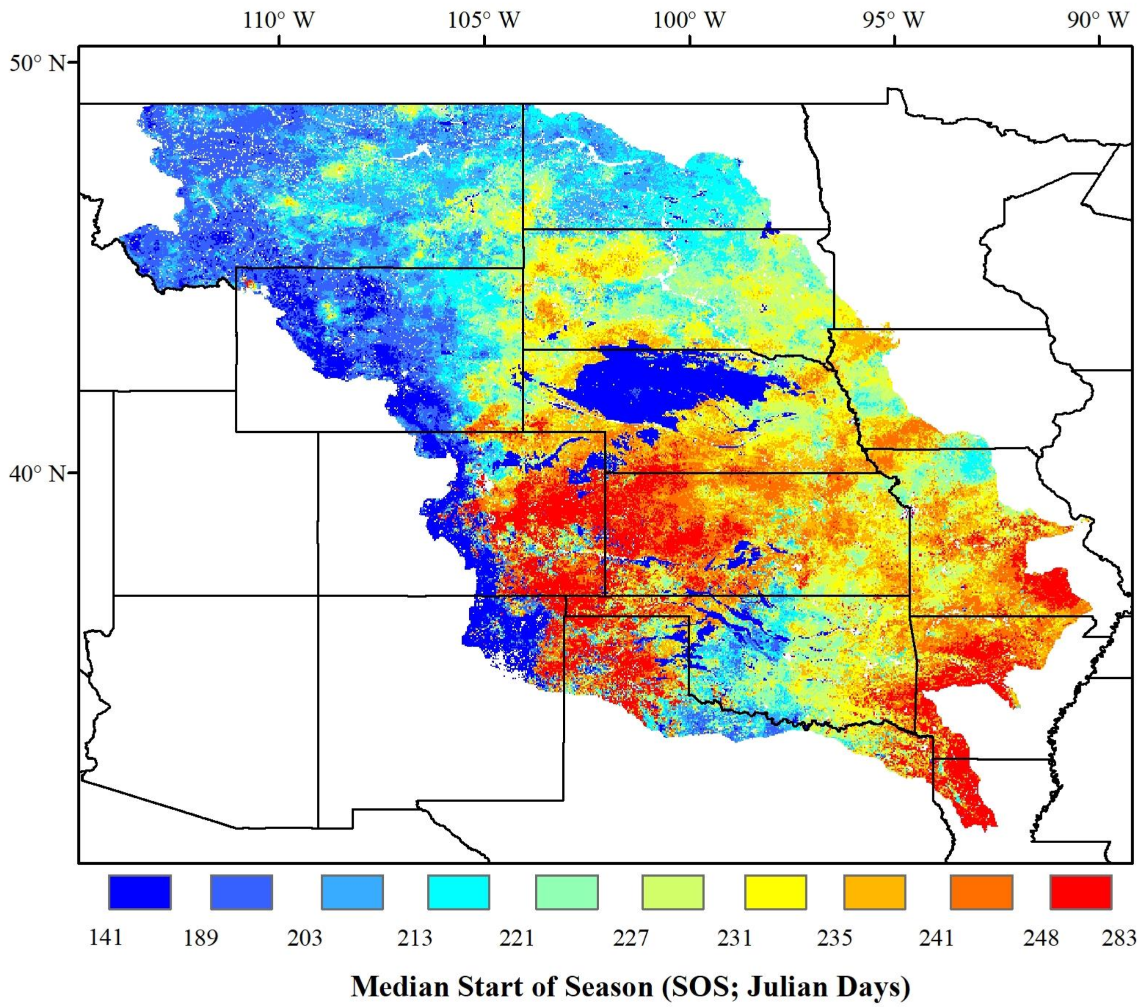

Figure 4.6. A map of the spatial variations of the median start of season (SOS) in Julian days for the Missouri and Arkansas-White-Red river basins with outliers excluded.

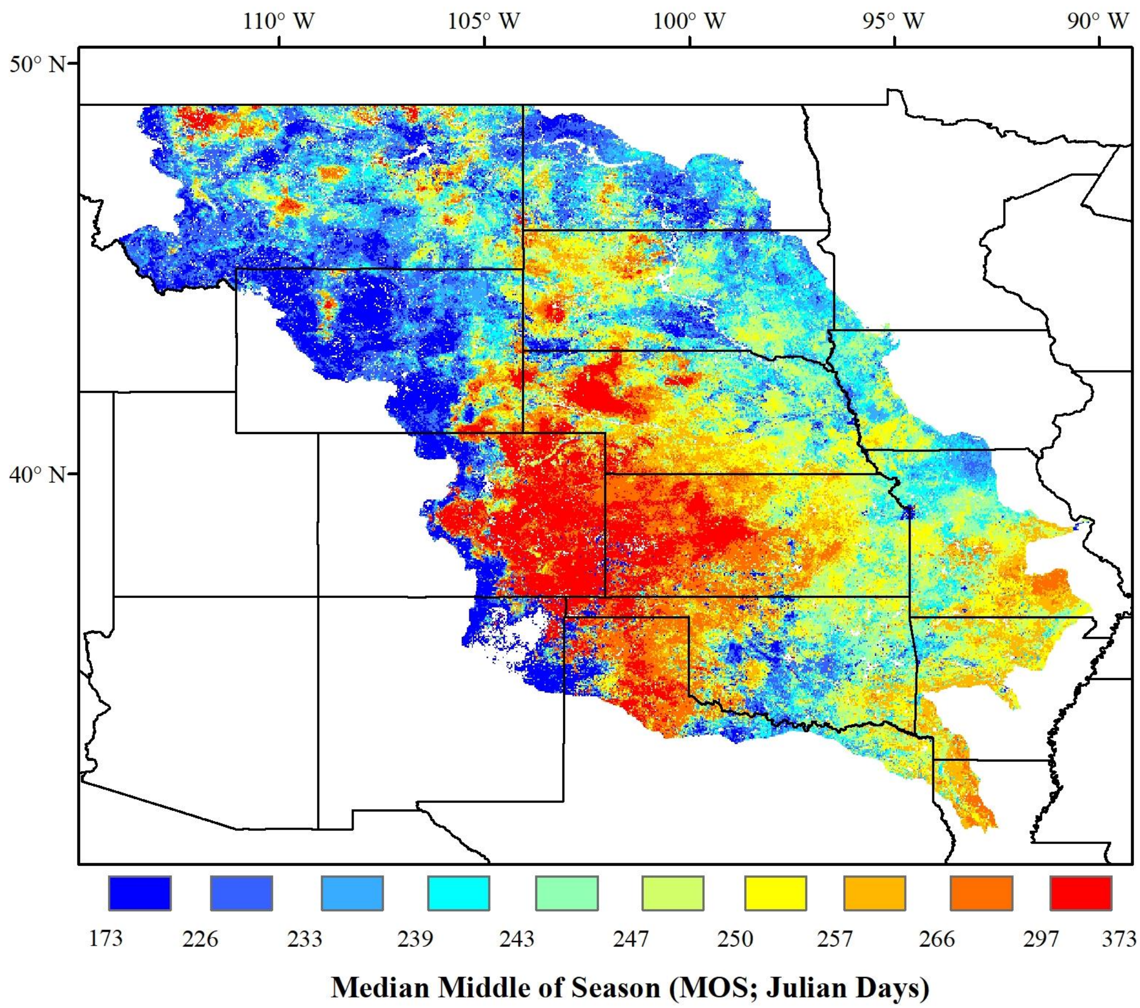

Figure 4.7. A map of the spatial variations of the median middle of season (MOS) in Julian days for the Missouri and Arkansas-White-Red river basins with outliers excluded.

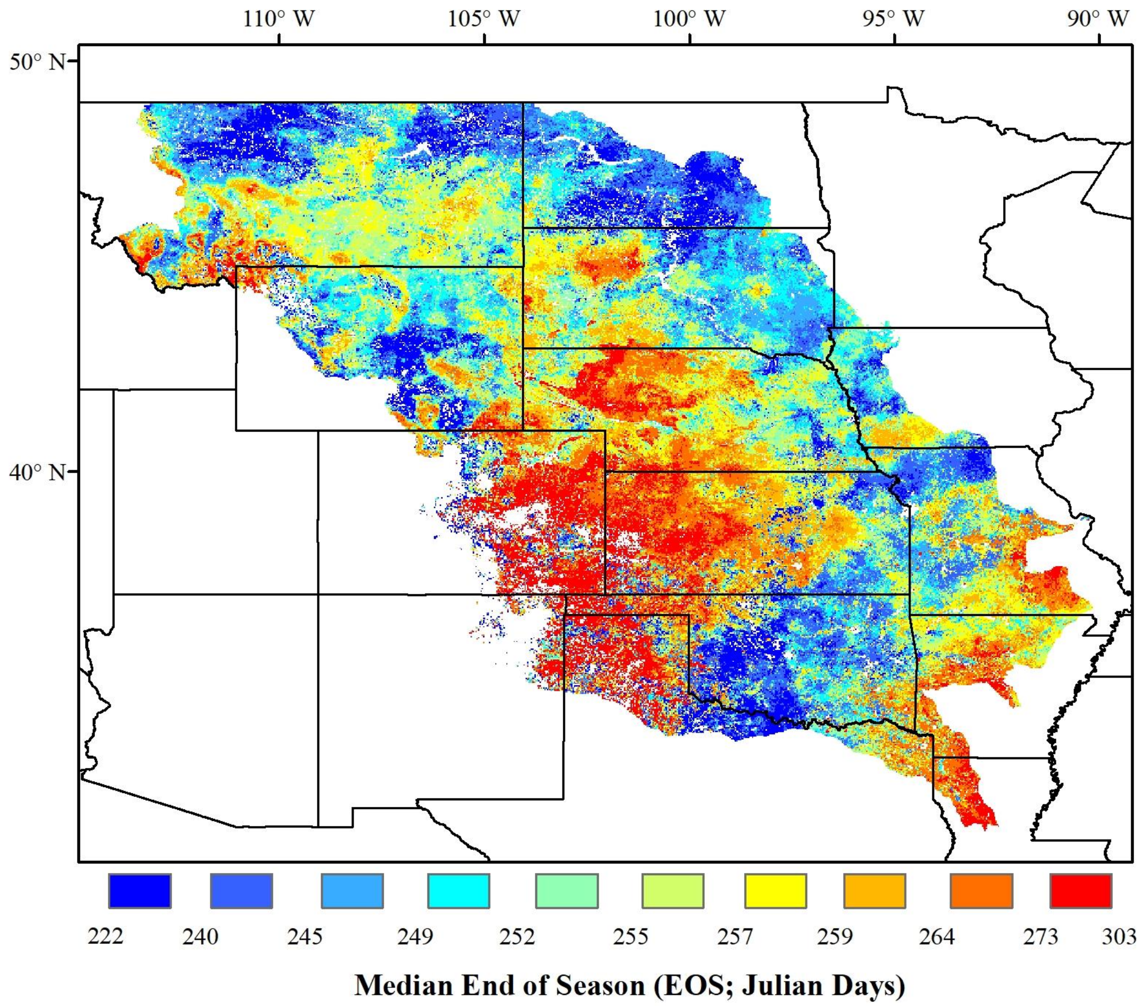

Figure 4.8. A map of the spatial variations of the median end of season (EOS) in Julian days for the Missouri and Arkansas-White-Red river basins with outliers excluded.

Another interesting feature of the median SOS, MOS, and EOS maps is that urban areas stand out, in the SOS and MOS maps where the hydrometrics occur earlier but in the EOS map the urban areas were deemed outliers and masked out of the map. The reason that urban areas are considered outliers is because the amount of impervious areas in developed regions is impacting TIMESAT's ability to capture the hydrometrics accurately for those areas. An improvement to the TIMESAT model would be to exclude all developed areas from the analysis.

To be able to determine if the hydrometrics can be distinguished between each ecoregion, density plots were created to represent the distribution of the median SOS, MOS, and EOS for each ecoregion within the study area (Appendix D). From this, it can be seen that some ecoregions start, middle, and end of season hydrometrics are normally distributed, where others are right or left-skewed (Table 4.4). Also, there are differences in the number of peaks for each distribution where the Nebraska Sand Hills SOS shows the greatest number of peaks with 4. Additionally, for the SOS, seven pairs were not significantly different from each other (Figure 4.9). For the MOS hydrometric, there are six non-significant pairs and for the EOS hydrometric, there are five pairs. This shows that for a majority of the ecoregions, the seasonal hydrometrics can be distinguished between each region, supporting the different biotic and abiotic phenomena for the region.

Table 4.3. A summary of the start (SOS), middle (MOS), and end of season (EOS) median distribution for each level III ecoregion within the Missouri and Arkansas-White-Red river basins.

Ecoregion	Distribution	Number of peaks		
		SOS	MOS	EOS
Idaho Bathalith	Normal	1	2	2
Middle Rockies	Normal	1	1	1
Wyoming Basin	Normal	3	2	2
Southern Rockies	Right-skewed	1	2	2
High Plains	Normal	2	3	3
South West Tablelands	Left-skewed	2	3	3
Central Great Plains	Right-skewed	3	2	3
Flint Hills	Normal	2	1	2
Cross Timbers	Normal	1	2	2
Texas Blackland Prairies	Right-skewed	1	1	3
East Central Texas Plains	Right-skewed; Normal SOS	1	1	1
South Central Plains	Left-skewed	2	2	2
Ouachita Mountains	Right-skewed	2	1	1
Arkansas Valley	Right-skewed	1	1	1
Boston Mountains	Normal	1	2	3
Ozark Highlands	Normal	2	2	3
Central Irregular Plains	Normal	2	1	2
Canadian Rockies	Normal	1	1	1
North West Glaciated Plains	Normal	1	1	1
North West Great Pains	Left-skewed	2	2	1
Nebraska Sand Hills	Right-skewed	4	1	3
North Glaciated Plains	Normal	1	1	1
Interior River Valleys & Hills	Right-skewed	2	2	2
Mississippi Alluvial Plains	Left-skewed; Right-skewed EOS	1	2	1
Western Corn Belt Plains	Normal	1	1	1

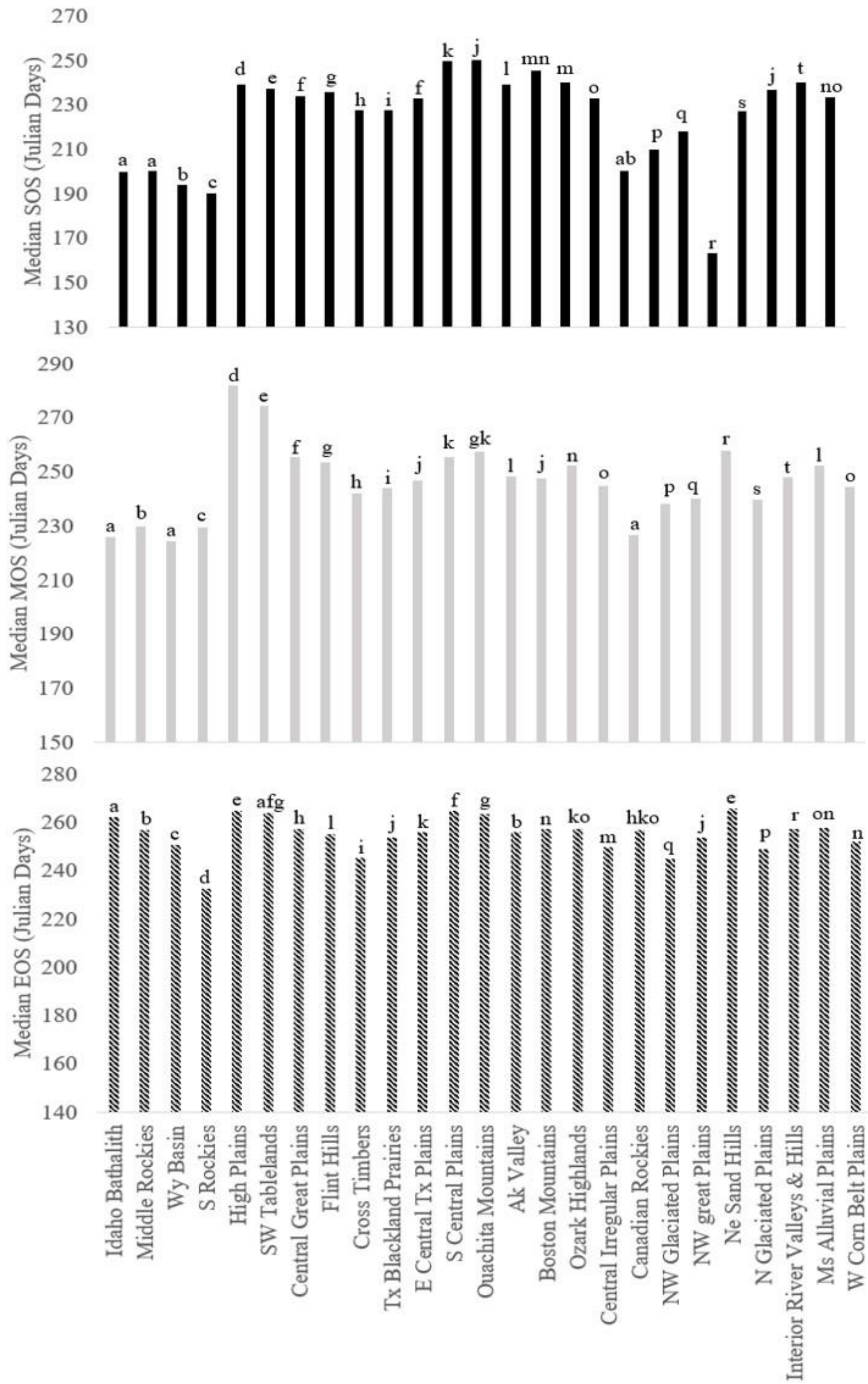

Figure 4.9. The mean start (SOS), middle (MOS), and end (EOS) values in Julian days for each level III ecoregion within the Missouri and Arkansas-White-Red river basins. The letters above plots indicate similarities among dates (same letters stand for no statistical difference).

4.3.3 Season Length Statistical Analysis

The SL is defined as the number of days between the SOS and EOS (Figure 3.8; Table 3.1). A positive impact would show a longer season and a negative impact would result in a shorter season. The results of the statistical analysis showed that all three variables had a negative effect on the SL (Table 4.2). The precipitation and HSG negative effects align with what was hypothesized. This shows that the more precipitation an area receives and the finer the soil for the area, the shorter the SL (Appendix B). The negative impact on the SL from temperature does not align with what was hypothesized since increased temperature results in increased ET rates, but like what has been mentioned, the seasonal temperature variations are not represented in the statistical model. The SL hydrometric was the only hydrometric that showed the largest impact from the HSG variable, β_3. This is shown in Figure 4.10 where HSG A stands out to have a significantly longer SL than the rest of the study area.

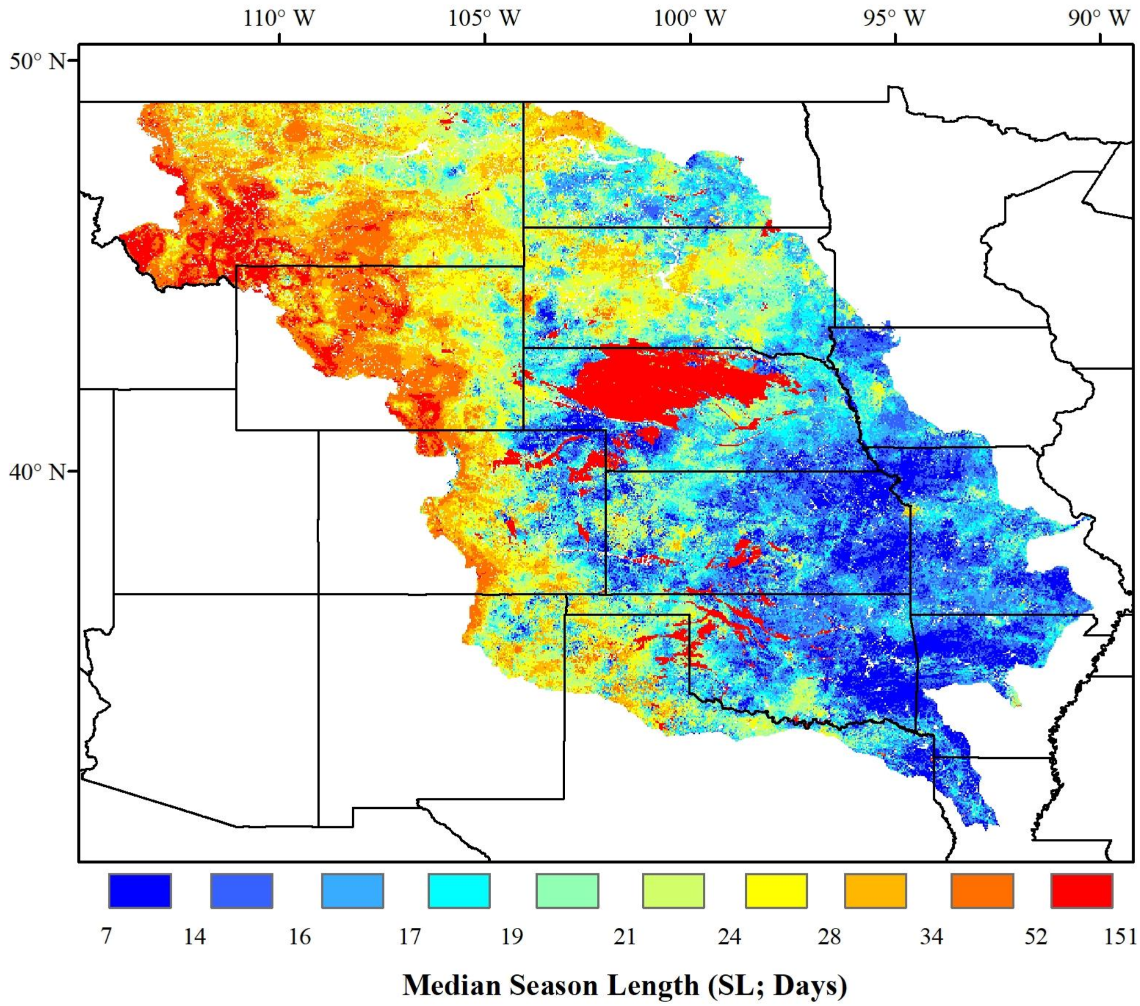

Figure 4.10.A map of the spatial variations of the median soil season length (SL) in Julian days for the Missouri and Arkansas-White-Red river basins with outliers excluded.

There is also a trend of shorter SL in the southeastern portion of the study area with increased SL in the northwestern portion where there is higher elevation (Figure 3.3), decreased temperatures (Figure 3.4), and decreased precipitation (Figure 3.2). However, the negative impact on SL from temperatures shows that precipitation could be the cause for this trend. Topography could also play a role in this trend. In general, higher elevations and steeper slopes, like what is found in mountain regions in the western portion of the study area, resulting in decreased soil moisture where recharge occurs from snowmelt during the spring (Osenga, et al., 2019). To further understand this trend, elevation and slope would need to be added to the statistical model as independent variables.

The SL hydrometric was then compared between each ecoregion. The results showed that the ecoregions with the longest SL were located in the northwestern portion of the study area and those with the shortest SL are in the southeast (Figure 4.11). This supports the results above where decreased precipitation and temperature resulted in a long dry season. Also, the ecoregion with the longest SL was the Nebraska Sand Hills which is characterized by HSG A. This also aligns with the results showing a negative trend for SL for finer soil particles.

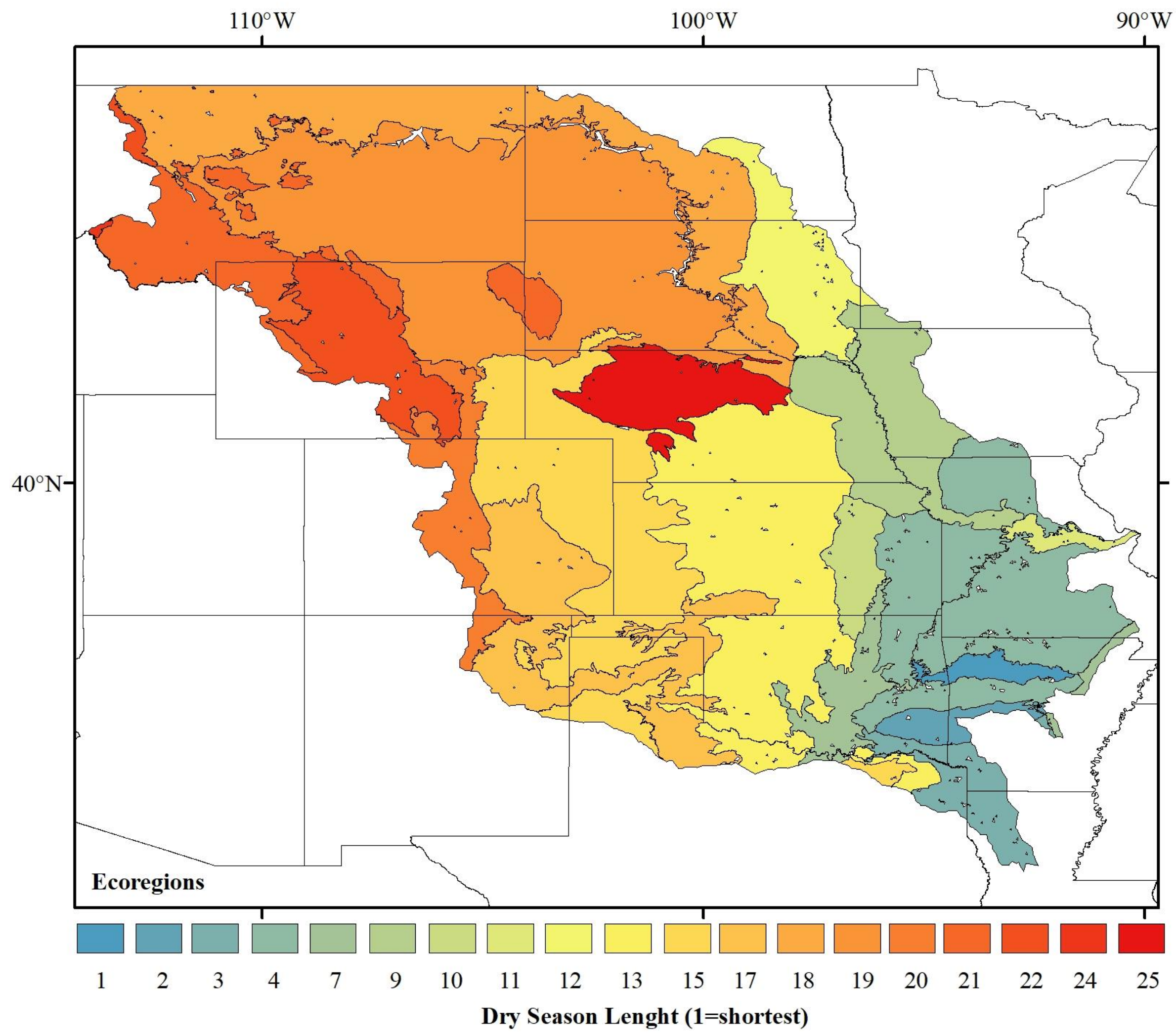

Figure 4.11. The season length (SL) rank for each level III ecoregion within the Missouri and Arkansas-White-Red river basins where a rank of 1 shows the shortest season and a rank of 25 shows the ecoregion with the longest season.

However, the ecoregions that do not follow these trends are Texas Blackland Prairie and East Central Texas Plains (Figure 3.12). The Texas Blackland Prairie is distinguished by its clay soils where the land cover for the area is cropland and rangeland with large portions of the area being converted to industrial and urban use (EPA, 2013). The East Central Texas Plains are mostly clay soils as well but the area is mainly used for rangeland and pasture and is distinguished by the post oak savanna vegetation that surrounding areas do not have. The Interior River Valley and Hills do not appear to follow the precipitation and temperature distributions. This could be because the Interior River Valley and Hills is mainly covered in cropland, about 30 percent of the area is pasture, and the rest is forested. The topology of the area could also play a role in the increased season length compared to the surrounding ecoregions since this region is made up of many wide, flat-bottomed terraced valleys, forested valley slopes, and dissected glacial till plains.

Chapter 5 - Conclusion

The goal of this research was to characterize the intra- annual variations of near-surface soil moisture throughout the Missouri and Arkansas-White-Red river basins. This was done with model-derived soil moisture estimates from the SPoRT-LIS software and TIMESAT to define soil moisture hydrometrics. Results from the TIMESAT analysis showed that it captured intra-annual soil moisture variability and that the seasonal soil moisture characteristics can be characterized at a spatial resolution of 3 km. However, the soil water storage capacity is not captured well at this spatial resolution. Additionally, the mean soil start and end of season occurred earlier than was hypothesized. This shows that refinement of the model smoothing parameters is necessary to accurately capture the soil moisture seasonality trends. This refinement can be done by adjusting the parameters based on different environmental variables such as land cover, topography, and climate.

The statistical analysis highlighted the impact of soil texture and climate on the availability of soil moisture. The results also showed that precipitation, temperature, and HSG all significantly impact hydrometrics except a non-significant impact on SL from precipitation. However, precipitation showed to have the largest impact on AMP, BAS, and MXF hydrometrics, where the temperature had the largest impact on SOS, MOS, and EOS hydrometrics. This shows that precipitation drives soil water storage capacity where temperature is the driver of the seasonal timing of soil water storage. The SL hydrometric was the only hydrometric where HSG had the largest impact. When comparing these results to the spatial variation in the median hydrometric values across the study area, some areas do follow these trends showing that additional biological or physical characteristics are playing a role in the soil moisture for that area. Additionally, it was found that for a majority of the ecoregions, the seasonal soil moisture characteristics can be distinguished between each region, highlight the impact of the biotic and abiotic phenomena on soil moisture availability. To further understand the magnitude of this impact, elevation, slope, land cover, and land management are needed to be analyzed.

Overall, this research used a novel approach to define soil water storage by characterizing the seasonality of water storage capacity using model-derived estimates compared to the traditional rainfall-runoff approach. This showed that soil moisture varies temporally throughout the year and that it also varies spatially given climate and HSG variations. From this research, there is a better understanding of the soil moisture variations throughout a majority of the Great Plains region which can help scientists, land managers, and policy makers to make decisions concerning reservoir management, irrigation applications, and farming practices. This is especially important as climate change impacts are expected to cause increased demands for water and energy within the Great Plains region (Melillo et al., 2014). This can result in shifts from irrigated cropland to dryland farmland, causing a predicted reduction in crop yields by a factor of 2 (Melillo et al., 2014; Deines et al., 2020). This can further cause serious impacts on the global economy and global food security systems since more than 30 percent of the U.S. ag land area along with more than 30 percent of the beef production for the United States is produced within the Great Plains which generates exports with an estimated total market value of $92 billion (NASS, 2012; Milello et al., 2014).

Chapter 6 - Future Work

Future work will include additional sensitivity analyses where land cover, land use, and topography are considered in defining the TIMESAT smoothing parameters. This is necessary because land cover impacts the rate of infiltration by the number of impervious materials there are in an area (Wessolek and Facklam 1997; Easton et al., 2007). Additionally, vegetation and ecosystems soil moisture have a coupled relationship that is basic to ecosystem dynamics where vegetation impacts the rate of ET, causes shading underneath the tree canopy, and intercepts rainfall from reaching the ground surface (Fernandez-Illesca et al., 2001; Lozano-Parra et al., 2018). Feng (2016) also found that vegetation degradation in wetting regions, can lengthen the dry season and increase the streamflow, which subsequently decreases the rainfall infiltration. As for land-use practices, the use of irrigation and farming practices such as no-till and crop residue can all impact soil moisture (Jasa, 2013; Lawston, et al., 2017). The topography is shown to influence soil moisture by affecting the infiltration, drainage, and runoff through the slope and angle of the landscape where high elevations and steeper slopes result in decreased soil moisture (Famiglietti et al., 1998; Osenga, et al., 2019). The curvature of the landscape also influences the convergence of lateral flow and the upslope surface area influences the distribution of soil moisture by controlling the potential volume of subsurface moisture flowing past a particular point on the landscape (Famiglietti et al., 1998). Furthermore, extending the study area to the contiguous United States will provide soil moisture trends over more diverse climates, topography, and land use. This in turn will allow for a better understanding of the soil moisture seasonality drivers.

References

Akuraju, V. R., Ryu, D., George, B., Ryu, Y., Dassanayake, K. (2016). Seasonal and inter-annual variability of soil moisture stress function in dryland wheat field, Australia. *Ag.* and *Forest Meteor.*, 232, 489-499. http://doi.org/10.1016/j.agrformet. 2016.10.007

Al-Shrafany, D., Rico-Ramirez, M.A., Han, D., Bray, M. (2014). Comparative assessment of soil moisture estimation from land surface model and satellite remote sensing based on catchment water balance. *Meteorol. App.*, 21, 521-534. http://10.1002/met.1357

Al-Yaari, A., Wigneron, J.P., Ducharne, A., Kerr, Y., de Rosnay, P., de Jeu, R., Govind, A., Al Bitar, A., Albergel, C., Munoz-Sabater, J., Richaume, P., Mialon, A. (2014). Global-scale evaluation of two satellite-based passive microwave soil moisture datasets (SMOS and AMSR-E) with respect to Land Data Assimilation System estimates. *Remote Sens. Environ*., 149, 181 – 195.

Albergel, C., Dorigo, W., Reichle, R. H., Balsamo, G., De Resonay, P., Munoz-Sabater, J., Isakesen, L., De Jeu, R., Wagner, W. (2013). Skill and global trend analysis of soil moisture from reanalyses and microwave remote sensing. *American Meteor*. *Soc.,* 14, 1259-1277. http://doi.org/10.1175/JHM-D-12-0161.1

Alvarez-Garreton, C., Ryu, D., Western, A.W., Crow, W.T., Robertson, D.E. (2014). The impacts of assimilating satellite soil moisture into a rainfall-runoff model in a semi-arid catchment. *J. Hydrol.*, 519, 2763-2774. http://doi.org/10.1016/j.jhydrol.2014.07.041

Amin, M., Khan, A. A., Perveen, A., Rauf, Z., Hassan, S. S., Goheer, M. A., Ijaz, M. (2019). Drought risk assessment: a case study in Punjab, Pakistan. *Sarhad J. Agr.*, 35(1), 234-243. http://doi.org/10.17582/journal.sja/2019/35.1.234.243.

An, R., Zhang, L., Wang, Z., Quaye-Ballard, J.A., You, J., Shen, X.,… Ke, Z. (2016). Validation of the ESA CCI soil moisture product in China. *Int. J. Appl. Earth Obs. Geoinfo.*, 48, 28-36. http://doi.org/10.1016/j.jag.2015.09.009

Antle, J. M., Kluck, D., McPherson, R. A., Petersen, S., Scanlon, B., Sherman, K. (2014). Great Plains. In *National climate assessment*. U.S. Global Change Research Program. Retrieved from https://nca2014.globalchange.gov/report

Ault, T. R., Mankin, J. S., Cook, B. I., Smerdon, J. E. (2016). Relative impacts of mitigation, temperature, and precipitation on 21st-century megadrought risk in the American Southwest. *Sci. Adv.*, 2(10), 1-8. http://doi.org/10.1126/sciadv.1600873

Berghuijs, W. R., Woods, R. A., Hutton, C. J., Sivapalan, M. (2016). Dominant flood generating mechanisms across the United States. *Geophys. Res. Lett.*, 43, 4382-4390. http://doi.org/10.1002/2016GL068070

Blankenship, C.B., Case, J.L., Zavodsky, B.T., Crosson, W.L. (2016). Assimilation of SMOS retrievals in the land information system. *IEEE Transac. Geosci. Rem. Sens.*, 51(11), 6320-6332. http://doi.org/10.1109/TGRS.2016.2579604

Brocca, L., Tullo, T., Melone, F., Moramarco, T., Morbidelli, R. (2012). Catchment scale soil moisture spatial-temporal variability. *J. Hydrol*., 422–423, 63–75. http://doi. org/10.1016/j.jhydrol.2011.12.039.

Cammalleri, C., Micale, F., Vogt, J. (2015). On the value of combining different modeled soil moisture products for European drought monitoring. *J. Hydrol.*, 525, 547-558. DOI: 10.1016/j.jhydrol.2015.04.021

Case, J.L. (2016). From drought to flooding in less than a week over South Carolina. *Res. Phys.*, 6, 1183-1184.

Case, J.L., White, K.D., Guyer, B., Meyer, J., Srikishen, J., Blankenship, C.B., Zavodsky, B.T. (2016). Real-time land information system over the continental U.S. for situational awareness and local numerical weather prediction applications. *30th Conference on Hydrology*, New Orleans, LA.

Case, J.L., Zavodsky, B.T. (2018). Evolution of 2016 drought in the southeastern United States from a land surface modeling perspective. *Res. Phys.*, 8, 654-656.

Chen, F., Dudhia, J. (2001). Coupling an advanced land-surface/hydrology model with the Penn State/NCAR MM5 modeling system. Part I: Model description and implementation. *Monthly Weath. Rev.*, 129, 569-585.

Chen, X., Liu, X., Liu, Z., Zhou, P., Zhou, G., Liao, J., Liu, L. (2016). Spatial clusters and temporal trends of seasonal surface soil moisture across China in responses to regional climate and land cover changes. *Ecohydrol.*, 10, 1-12. https://doi.org/10.1002/eco.1800

Cohen, J. (1992). Statistical power analysis. *Physcho. Sci.*, 1(3), 98-101

Collins, C. D., Foster, B. L. (2008). The Role of topography and soil characteristics in the relationship between species richness and primary productivity in a Kansas grassland. *Trans. Kansas Acad. Sci.*, 111, 105-117.

Conant, R. T., Kluck, D., Anderson, M., Badger, A., Boustead, B. M., Derner, J., ... Small, V. (2018). Northern Great Plains. In *Impacts, risks, and adaptation in the United States: Fourth National Climate Assessment* (Vol. II, pp. 941-986). Washington, DC: U.S. Global Change Research Program. https://doi.org/10.7930/NCA4.2018.CH22

Cook, B. I., Smerdon, J. E., Seager, R., Coats, S. (2014). Global warming and 21st century drying. *Clim. Dyn.*, 43, 2607-2627. http://doi.org/10.1007/s00382-014-2075-y

Cook, B. I., Ault, T. R., Smerdon, J. E. (2015). Unprecedented 21st century drought risk in the American Southwest and Central Plains. *Sci. Adv.*, 1 (1). http://doi.org/10.1126/sciadv.1400082

Cook, E. R., Seager, R., Cane, M. A., Stahle, D. W. (2007). North American drought: Reconstructions, causes and consequences. *Earth Sci. Rev.*, 81, 93-134. http://doi.org/10.1016/j.earscirev.2006.12.002

Deines, J. M., Schipanski, M. E., Golden, B., Zipper, S. C., Nozari, S., Rottler, C., Guerrero, B., Sharda, V. (2020). Transitions from irrigated to dryland agriculture in the Ogallala Aquifer: Land use suitability and regional economic impacts. *Agric. Water Manag.*, 223, http://doi.org/10.1016/j.agwat.2020.106061

Deliberty, T. L., Legates, D. R. (2003). Interannual and seasonal variability of modelled soil moisture in Oklahoma. *Int. J. Climato.*, 23(9), 1057-1086. http://doi.org/10.1002/joc.904

Dewitz, J., 2019, National Land Cover Database (NLCD) 2016 Products (ver. 2.0, July 2020): U.S. Geological Survey data release, http://doi.org/10.5066/P96HHBIE

Dobriyal, P., Qureshi, A., Badola, R., Hussain, S.A. (2012). A review of the methods available for estimating soil moisture and its implications for water resource management. *J. Hydrol.*, 458-459, 110-117. http://doi.org/10.1016/j.jhydrol.2012.06.021

Dorigo, W.A., Wagner, W., Hohensinn, R., Hahn, S., Paulik, C., Xaver, A., Gruber, A., Drusch, M. Mecklenburg, S., van Oevelen, P., Robock, A., Jackson, T. (2011). The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrol. Earth Sys. Sci.*, 15, 1675-1698. http://doi.org/10.5194/hess-15-1675-2011

Dorigo, W., Jeu, R., Chung, D., Parinussa, R., Liu, Y., Wagner, W., Fernandez-Prieto, D. (2012). Evaluating global trends (1988-2010) in harmonized multi-satellite surface soil moisture. *Geophys. Res. Lett.,* 39, L18495. http://doi.org/10.1029/2012GL052988

Eagleson, P. S. (1978). Climate, soil, and vegetation: The expected value of annual evapotranspiration. *Wat. Resour. Res.*, 14(5), 731–739. https://doi.org/10.1029/WR014i005p00731

Easton, Z. M., Gerard-Marchant, P., Walter, M. T., Petrovic, A. M., Steenhuis, T. S. (2007). Hydrologic assessment of an urban variable source watershed in the northeast United States, *Water Resour. Res.,* 43. http://doi.org/10.1029/2006WR005076

Ehsani, N., Vorosmarty, C. J., Fekete, B. M., Stakhiv, E. Z. (2017). Reservoir operations under climate change: storage capacity options and mitigate risk. *J. Hydrol.*, 555, 435-446. http://doi.org/10.1016/j.jhdrol.2017.09.008

Ek, M.B., Mitchell, K.E., Lin, Y., Rogers, E., Grunmann, P., Koren, V.,…Tarpley, J.D. (2003). Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale eta model. *J. Geophys. Resour.,* 108(D22), 8851. http://doi.org/10.1029/2002JD003296

Eklundh, L., Jonsson, P. (2015). TIMESAT 3.2 with parallel processing software manual. Retrieved from http://www.nateko.lu.se/TIMESAT/

Entekhabi, D., Njoku, E.G., O'Neill, P.E., Kellogg, K.H., Crow, W.T., Edelstein, W.N., Entin, J.K., Goodman, S.D., Jackson, T.J., Johnson, J., Kimball, J., Piepmeier, J.R., Koster, R.D., Martin, N., McDonald, K.C., Moghaddam, M., Moran, S., Reichle, R., Shi, J.C., Spencer, M.W., Thurman, S.W., Tsang, L., Van Zyl, J. (2010). The Soil Moisture Active

Passive (SMAP) mission. *Proc*. *IEEE*, 98, 704 – 716. http://doi.org/10.1109/JPROC.2010.2043918

EPA. (2013). Level III ecoregions of the conterminous United States. Environmental Protection Agency. Retrieved from https://www.epa.gov/eco-research/level-iii-and-iv-ecoregions-continental-united-states

Erback, D. C. (1987). Measurement of soil bulk density and moisture. *Trans. ASAE,* 30(4), 0922-0931. http://doi.org/10.13031/2013.30500

Escorihuela, M.J., de Rosnay, P., Kerr, Y.H. (2006). Temperature Dependency of Bound Water Spectral Parameters and Its Influence in Soil Moisture Measurements. EGU, Wien. EGU General assembly.

Famiglietti, J. S., Rudnicki, J. W., Rodell, M. (1998). Variability in surface moisture content along a hillslope transect: Rattlesnake Hill, Texas. *J. Hydrol.*, 210, 259-281.

Famiglietti, J. S. and Rodell, M. (2013). Water in balance. *Sci.,* 340, 1300-1301. http://doi.org/10.1126/science.1240843

Fascetti, F., Pierdicca, N., Pulvirenti, L., Crapolicchio, R., Muñoz-Sabater, J. (2016). A comparison of ASCAT and SMOS soil moisture retrievals over Europe and northern Africa from 2010 to 2013. *Int. J. Appl. Earth Obs. Geoinfo.*, 45, 135-142. http://doi.org/10.1016/j.jag.2015.09.008

Feng, H. and Zhang, M. (2015). Global land moisture trends: drier in dry and wetter in wet over land. *Sci. Rep.*, 5: 18018. http://doi.org/10.1038/srep18018

Feng, H. (2016). Individual contributions of climate and vegetation change to soil moisture trends across multiple spatial scales. *Sci. Rep.,* 6: 32782. http://doi.org/10.1038/srep32782

Fernandez-Illescas, C. P., Porporato, A., Laioi, F., Rodriguez-Iturbe, I. (2001). The ecohydrological role of soil texture in a water-limited ecosystem. *Wat. Resour. Res.,* 37(12), 2863-2872.

Fisher, R. A., Koven, C. D. (2020). Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. *J. Adv. Mod. Earth Sys.* 12(4). http://doi.org/10.1029/2018MS001453

Gaspard, G., Kim, D., Chun, Y.(2019). Residual spatial autocorrelation in macroecological and biogeographical modeling: a review. *J. Ecology. Environ.* 43, 19. https://doi.org/10.1186/s41610-019-0118-3

GCOS. (2010). Implementation plan for the Global Observing System for climate in support of the UNFCCC—2010 update, report. Global Climate and Observation System. *World Meteorol. Organ.,* Geneva Switzerland.

Gleick, P. H. (2016). Water strategies for the next administration. *Sci.,* 354(6312), 555-556. http://doi.org/10.1126/science.aaj2221

Gracia-Estringana, P., Latron, J., Llorens, P., Gallart, F. (2013). Spatial and temporal dynamics of soil moisture in a Mediterranean mountain area (Vallcebre, NE Spain). *Ecohydrol.,* 6, 741-753. http://doi.org/10.1002/eco.1295

Grayson et al., (1997). Preferred states in spatial soil moisture patterns: Local and nonlocal controls. *Wat. Resour. Res.,* 33(12), 2897-2908. http://doi.org/10.1029/97WR02174

Griesfeller, A., Lahoz, W.A., de Jeu, R.A.M., Dorigo, W., Haugen, L.E., Svendby, T.M., Wagner, W. (2016). Evaluation of satellite soil moisture products over Norway using ground-based observations. *Int. J. Appl. Earth Obs. Geoinfo.,* 45, 155-164. http://doi.org/10.1016/j.jag.2015.04.016

Grillakis, M.G., Koutroulis, A.G., Komma, J., Tsanis, I.K., Wagner, W., Blöschl, G. (2016). Initial soil moisture effects on flash flood generation – a comparison between basins of contrasting hydro-climatic conditions. *J. Hydrol.,* 541, 206-217. http://doi.org/10.1016/j.jhydrol.2016.03.007

Greve, P., Orlowsky, B., Mueller, B., Sheffield, J., Reichstein, M., Seneviratne, S. I. (2014). Global assessment of trends in wetting and drying over land. *Nat. Geoscie.,* 7, 716-721. http://doi.org/10.1038/NGEO2247

Hall, A. (2014). Projecting regional change. *Sci.,* 345(6216), 1461-1462. http://doi.org/10.1126/science.aaa0629

Hoerling, M. P., Eischeid, J. K., Diaz, H. F., Webb, R. S., Dole, R. M., Easterling, D. R. (2012). Is a transition to semipermanent drought conditions imminent in the U.S. Great Plains? *J. Clim.,* 25(24), 8380-8386. http://doi.org/10.1175/JCLI-D-12-00449.1

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G.,…Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States – representing a decade of land cover change information. *Photogram. Eng. Rem. Sens.,* 81(5), 345-354.

Jabobs, J. M., Mohanty, B. P, Hsu, E. C., Miller, D. (2004). SMEX02: Field scale variability, time stability and similarity of soil moisture. *Rem. Sens. Environ.,* 92(4), 436-446. http://doi.org/10.1016/j.rse.2004.02.017

Jasa, P. (2013). Conserving soil and water with no-till and crop residue. Nebraska Institute of Agricultural and Natural Resources. Retrieved from https://cropwatch.unl.edu

Jedlovec, G. (2013). Transitioning research satellite data to the operational weather community: the SPoRT paradigm. *IEEE Geosci. Rem. Sens. Mag.*, 1(1), 62-66. http://doi.org/10.1109/MGRS.2013.2244704

Jonsson, P., Eklundh, L. 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transac. Geosci. Rem. Sens.*, 40 (8), 1824-1832

Jonsson, P., Eklundh, L. 2004. TIMESAT- a program for analyzing time-series of satellite sensor data. *Comp. Geosci.*, 30(8), 833-845. http://doi.org/10.1016/j.cageo.2004.05.006

Kerr, Y., Waldteufel, P., Wigneron, J.P., Martinuzzi, M., Font, J., Berger, M. (2001). Soil moisture retrieval from space: the soil moisture and ocean salinity (SMOS) mission. *IEEE Trans. Geosci. Remote Sens.*, 39, 1729–1736. http://doi.org/10.1109/36.942551.

Kerr, Y.H. (2007). Soil moisture from Space: Where are we? *Hydrogeol. J.*, 15, 117 – 120.

Koster, R.D., Guo, Z., Yang, R., Dirmeyer, P.A., Mitchell, K., Puma, M.J. (2009). On the nature of soil moisture in land surface models. *J. Clim.*, 22, 4322-4335. http://doi.org/10.1175/2009JCLI2832.1

Koster, R. D., Milly, P. C. D. (1997). The interplay between transpiration and runoff formulations in land surface schemes used with atmospheric models. *J. Clim.*, 10(7), 1578-1591

Kukal, M. S., Irmak, S. (2018). U.S. agro-climate in 20th century: Growing degree days, first and last frost, growing season length, and impacts on crop yields. *Sci. Rep.*, 8(6977), 1-14. http://doi.org/10.1038/s41598-018-25212-2

Kumar, S.V., Peters-Lidard, C.D., Tian, Y., Houser, P.R., Geiger, J., Olden, S.,…Sheffield, J. (2006). Land information system: An interoperable framework for high resolution land surface modeling. *Environ. Model. Softw.*, 21(10), 1402-1415. http://doi.org/10.1016/j.envsoft.2005.07.004

Kumar, S.V., Reichle, R.H., Peters-Lidard, C.D., Koster, R.D., Zhan, X., Crow, W.T.,…Houser, P.R. (2008). A land surface data assimilation framework using the land information system: description and applications. *Adv. Water Resour.*, 31, 1419-1432. http://doi.org/10.1016/j.advwatres.2008.01.013

Kumar, S.V., Reichle, R.H., Koster, R.D., Crow, W.T., Peters-Lidard, C.D. (2009). Role of subsurface physics in the assimilation of surface soil moisture observations. *J. Hydrometeorol.*, 10, 1534-1547. http://doi.org/10.1175/2009JHM1134.1

Lai, X., Wen, J., Cen, S., Huang, X., Tian, H., Shi, X. (2016). Spatial and temporal soil moisture variations over China from simulations and observations. *Adv. Meteor.*, 1-14. http://doi.org/10.1155/2016/4587687

Lawston, P. M., Stanello, J. A., Kumar, S. V. (2017). Irrigation signals detected from SMAP soil moisture retrievals. *Geophys. Res. Let.*, 44(23), 11,860-11,867. http://doi.org/10.1002/2017GL075733

Lehmann A., Stahr K. (2007). Nature and significance of anthropogenic urban soils. *J. Soils Sed.*, 7, 247–260

Li, Y., Grimaldi, S., Pauwels, V.R.N., Walker, J.P. (2018). Hydrologic model calibration using remotely sensed soil moisture and discharge measurements: the impact on predictions at gauged and ungauged locations. *J. Hydrol.* 557, 897–909. http://doi.org/10.1016/j.jhydrol.2018.01.013.

Liu, Y. Y., Parinussa, R. M., Dorigo, W. A., De Jeu, R. A. M., Wagner, W., Van Dijk, A. I. J. M., McCabe, M. F., Evans, J. P. (2011). Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals, *Hydrol. Earth Syst. Sci.,* 15, 425–436, http://doi.org/10.5194/hess-15-425-2011.

Liu, Y., Peters-Lidard, C.D., Kumar, S., Foster, J.L., Shaw, M., Tian, Y., Fall, G.M. (2013). Assimilating satellite-based snow depth and snow cover products for improving snow predictions in Alaska. *Adv. Water Resour.*, 54, 208-227.

Lozano-Parra, J., Pulido, M., Lozano-Fondon, C., Schnabel, S. (2018). How do soil moisture and vegetation covers influence soil temperature in drylands of Mediterranean regions? *Water*, 10, 1747. http://doi.org/10.3390/w10121747

McDonough, K. R., Hutchinson, S. L., Hutchinson, J. M. S., Case, J. L., Rahmani, V. (2018). Validation and assessment of SPoRT-LIS surface soil moisture estimates for water resources management applications. *J. Hydrol*., 566, 43-54

McDonough, K. R., Hutchinson, S. L., Hutchinson, J. M. S. (2020). Declining soil moisture threatens water availability in the Great Plains U.S. . *ASABE*, 63(5), 1147-1156. http://doi.org/10.13031/trans.13773

Mekonnen, M. M., Hoekstra, A. Y. (2016). Four billion people facing severe water scarcity. *Sci. Adv.*, 2(2). http://doi.org/10.1126/sciadv.1500323

Meng, S., Xie, X., Liang, S. (2017). Assimilation of soil moisture and streamflow observations to improve flood forecasting with considering runoff routing lags. *J. Hydro.*, 550, 568-579. http://doi.org/10.1016/j.jhydrol.2017.05.024

Melillo, J M, Richmond, T C, Yohe, G W (2014) Climate Change Impacts in the United States: The Third National Climate Assessment (U.S. Global Change Research Program), pp 841. http://doi.org/10.7930.J0Z31WJ2

Miralha, L., Kim, D. (2018). Accounting for and predicting the influence of spatial autocorrelation in water quality modeling. *Int. J. Geo-Info.*, 7, 64. http://doi.org/10.3390/ijgi7020064

Moore, B. C., Coleman, A. M., Wigmosta, M. S., Skaggs, R. L., & Venteris, E. R. (2015). A high spatiotemporal assessment of consumptive water use and water scarcity in the conterminous United States. *Water Resour. Mgmt*., 29(14), 5185-5200. https://doi.org/10.1007/s11269-015-1112-x

Moran, M.S., Peters-Lidard, C.D., Watts, J.M., McElroy, S. (2004). Estimating soil moisture at the watershed scale with satellite-based radar and land surface models. *Can. J. Rem. Sens.,* 30(5), 805-826.

National Agricultural Statistics Service (2012) Census of Agriculture (United States Department of Agriculture). Retrieved from https://www.nass.usda.gov/Publications/AgCensus/2012/#highlights

NRCS. (1998). Soil quality resource concerns: Available water capacity. Retrieved from https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs142p2_051279.pdf

NRCS. (2006). U.S. General Soil Map (STATSGO2). Retrieved from https://sdmdataaccess.sc.egov.usda.gov/

Oki, T., Kanae, S. (2006). Global hydrologic cycles and world water resources. *Sci.,* 313, 1068-1072

Oki, T., Kim, H. (2016). Macroscale hydrological modeling and global water balance. *In Terrestrial water cycle and climate change: Natural and Human-induced impacts* (221, pp. 3-17). American Geophysical Union.

Omernik, J.M. 1987. Ecoregions of the conterminous United States. Map (scale 1:7,500,000). *Ann. Assoc. Amer. Geogr*. 77(1), 118-125.

Osenga, E. C., Arnott, J. C., Endsley, K. A., Katzenberger, J. W. (2019). Bioclimatic and soil moisture monitoring across elevation in a mountain watershed: Opportunities for research and resource management. *Wat. Resour. Res.*, 55, 2493-2503. http://doi.org/10.1029/2018WR023653

Peters-Lidard, C.D., Houser, P.R., Tian, Y., Kumar, S.V., Geiger, J., Olden, S.,…Sheffield, J. (2007). High-performance earth system modeling with NASA/GSFC's land information system. *Innov. Sys. Softw. Eng.*, 3(3), 157-165.

Pinnington, E., Quaife, T., Black, E. (2018). Impact of remotely sensed soil moisture and precipitation on soil moisture predications in a data assimilation system with the JULES land surface model. *Hydrol. Earth Syst. Sci.*, 22, 2575-2588. http://doi.org/10.519/hess-22-2575-2018

Pitman A. J., Slater A. G., Desborough C. E., Zhao M. (1999). Uncertainty in the simulation of runoff due to the parameterization of frozen soil moisture using the GSWP methodology. *J. Geophys. Resear.*, 104, 16879–16888.

Pockrandt, B. R. 2012. A multi-year comparison of vegetation phenology between military training lands and native tallgrass prairie using TIMESAT and moderate-resolution satellite imagery. MS thesis. Manhattan, KS: Kansas State University, Department of Geography.

PRISM Climate Group. (2014). Terrain Elevation. Retrieved from http://www.prism.oregonstate.edu/normals/

PRISM Climate Group. (2015). 30-yr normal precipitation and temperature: Annual. Retrieved from http://www.prism.oregonstate.edu/normals/

Qiu, J., Gao, Q., Wang, S., Su, Z. (2016). Comparison of temporal trends from multiple soil moisture data sets and precipitation: The implication of irrigation on regional soil moisture trend. *Int. J. App. Earth Obs. Geoinfo.,* 48, 17-27. http://doi.org/10.1016/j.jag.2015.11.012

Rasul, G., Sharma, B. (2015). Climate policy the nexus approach to water-energy-food security: An option for adaption to climate change. *Clim. Policy,* http:doi.org/10.1080/14693062.2015.1029865

Reichle, R.H., Koster, R.D., Liu, P., Mahanama, S.P.P., Nkoju, E.G., Owe, M. (2007). Comparison and assimilation of global soil moisture retrievals from the advanced microwave scanning radiometer for the earth observing system (AMSR-E) and the scanning multichannel microwave radiometer (SMMR). *J. Geophys. Res.,* 112, D09108. https://doi.org/10.1029/2006JD008033.

Roderick, M. L., Sun, F., Lim, W. H., Farquhar, G. D. (2014). A general framework for understanding the response of the water cycle to global warming over land and ocean. *Hydrol. Earth. Syst. Sci.,* 18, 1575-1589

Rubel, F., Kottek, M. (2010). Observed and projected climate shifts 1901-2100 depicted by world maps of the Köppen-Geiger climate classification. *Meteorologische Zeitschrif.,* 19, 135-141. http://doi.org/10.1127/0941-2948-2010-0430

Sample, D. J., Owen, J. S., Fields, J. S., Barlow, S. (2016). Understanding soil moisture sensors: A fact sheet for irrigation professionals in Virginia. *Virginia Cooperative Extension,* BSE-198P

Saxton, K. E., Rowels, W. J. (2006). Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci. Soc. Am. J.*¸70 (5), 1569-1578. http://doi.org/10.2136/sssaj2005.0117

Shaifullah, K, Mezbahuddin, S., Sujauddin, M., Haque, S. S. (2008). Effects of coastal afforestation on some soil properties in Lakshmipur coast of Bangladesh. *J. For. Res.*, 19, 32-36. http://doi.org/10.1007/s11676-008-0005-8.

Shiklamonov, I. (1993). Water in crisis: A guide to the world's freshwater resources.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. U.S. General Soil Map (STATSGO2). Available online. Accessed [8/15/2019].

Srivastava, P.K., Han, D., Ramirez, M.A., Islam, T., (2013). Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application. *Water Resour. Manag.*, 27, 3127 – 3144.

Srivastava, P.K., O'Neill, P., Cosh, M., Kurum, M., Lang, R., Joseph, A. (2015). Evaluation of dielectric mixing models for passive microwave soil moisture retrieval using data from ComRAD ground-based SMAP simulator. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 8 (9), 4345 – 4354.

Srivastava, P. K., Petropoulos, G. P., Kerr, Y. H. (2016). *Satellite soil moisture retrieval: Techniques and applications*. Elsevier. (pp. 2-9).

Swain, D. L., Wing, O. E. J., Bates, P. D., Done, J. M., Johnson, K. A. Cameron, D. R. (2020). Increased flood exposure due to climate change and population growth in the United States. *Earth's Fut.,* 8. http://doi.org/10.1029/2020EF001778

USDA-NASS. (2012). Census of agriculture. Washington, DC: USDA National Agricultural Statistics Service. Retrieved from https://www.nass.usda.gov/Publications/AgCensus/2012/#highlights

USDA-NASS. (2015). Updated hydrologic soils group (HSG): Question and answers. Retrieved from https://www.nrcs.usda.gov/wps/PA_NRCSConsumption/download?cid=stelprd b1262857&ext=pdf

USGS. (2006). Major dams of the United States. In *National atlas of the United States.* Reston, VA: U.S. Geological Survey. Retrieved from https://www.sciencebase.gov/catalog/item/4f4e4b32e4b07f02db6b42c6

Wagner, W., Lemoine, G., Rott, H. (1999). A method for estimating soil moisture from ERS scatterometer and soil data. *Remote Sens. Environ.,* 70, 191–207. https//doi.org/10. 1016/S0034-4257(99)00036-X.

Wagner, W., Bloschl, G., Pampaloni, P., Calvet, J.C., Bizzarri, B., Wigneron, J.P., Kerr, Y. (2007). Operational readiness of microwave remote sensing of soil moisture for hydrologic applications. *Nord. Hydrol.*, 38, 1 – 20.

Wang, C., Meng, Q., Zhan, Y., Peng, J., Wei, X., Yang, J., Li, J. (2015). Ground sampling methods for surface soil moisture in heterogeneous pixels. *Environ. Earth Sci.,* 73, 6427-6436. http://doi.org/10.1007/s12665-014-3866-5

Wang, S., Mo, X., Liu, S., Lin, Z., Hu, S. (2016). Validation and trend analysis of ECV soil moisture data on cropland in north China plain during 1981-2010. *Int. J. App. Earth Obs. Geoinfo.,* 48, 110-121. http://doi.org/10.1016/j.jag.2015.10.010

Warton, D.I., Lyons, M., Stoklosa, J., Ives, A.R. (2016) Three points to consider when choosing a LM or GLM test for count data. *Meth. Ecol. Evo.*, 7, 882–890. http://doi.org/ : 10.1111/2041-210X.12552

Wessolek G., Facklam M. (1997). Site characteristics and hydrology of sealed areas. *Z Pflanzenernähr Bodenkd* , 160, 41–46

Vaz, C. M. P., Manieri, J. M., De Maria, I. C., Genuchten, M. T. V. (2013). Scaling the dependency of soil penetration resistance on water content and bulk density of different soils. *Soil Sci. Soc. Amer. J.*, 77(5). http://doi.org/10.2136/sssaj2013.01.0016

Xia, Y., Ek, M.B., Wu, Y., Ford, T., Quiring, S. (2015). Comparison of NLDAS-2 simulated and NASMD observed daily soil moisture. Part I: comparison and analysis. *J. Hydrometeorol.*, 16, 1962-1980. http://doi.org/10.1175/JHM-D-14-0096.1

Yang, K., Zhu, L., Chen, Y., Zhao, L., Qin, J., Lu, H., …, Fang, N. (2016). Land surface model calibration through microwave data assimilation for improving soil moisture simulations. *J. Hydrol.*, 533, 266–276. https://doi.org/10.1016/j.jhydrol.2015.12.018.

Zavodsky, B.T., Case, J.L., Blankenship, C.B., Crosson, W.L., White, K.D. (2013). Application of next-generation satellite data to a high-resolution, real-time land surface model. *EarthZine.*

Zeng, Z., Piao, S., Li, L. Z. X., Wang, T., Ciais, P., Lian, X., Yang, Y., Mao, J., Shi, X., Myneni, R. B. (2018). Impact of earth greening on the terrestrial water cycle. *J. Clim.*, 31(7), 2633-2650. http://doi.org/10.1175/JCLI-D-17-0236.1

Appendix A - TIMESAT Analysis

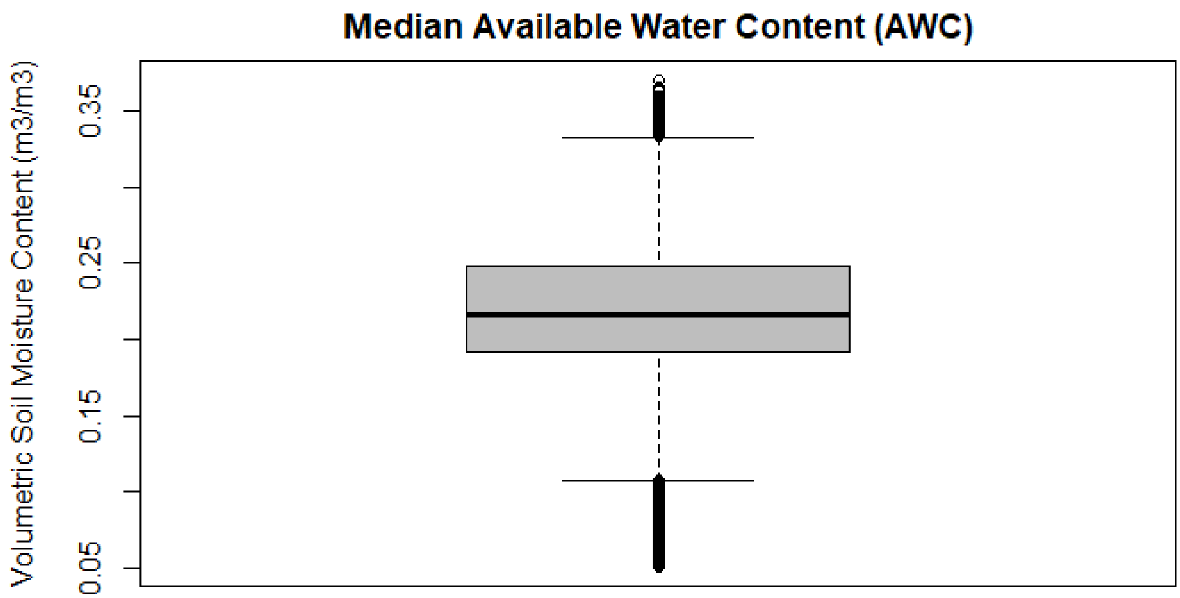

Figure A.6.1. Box-plot of the median available water content (AWC) for the Missouri and Arkansas-White-Red River basins.

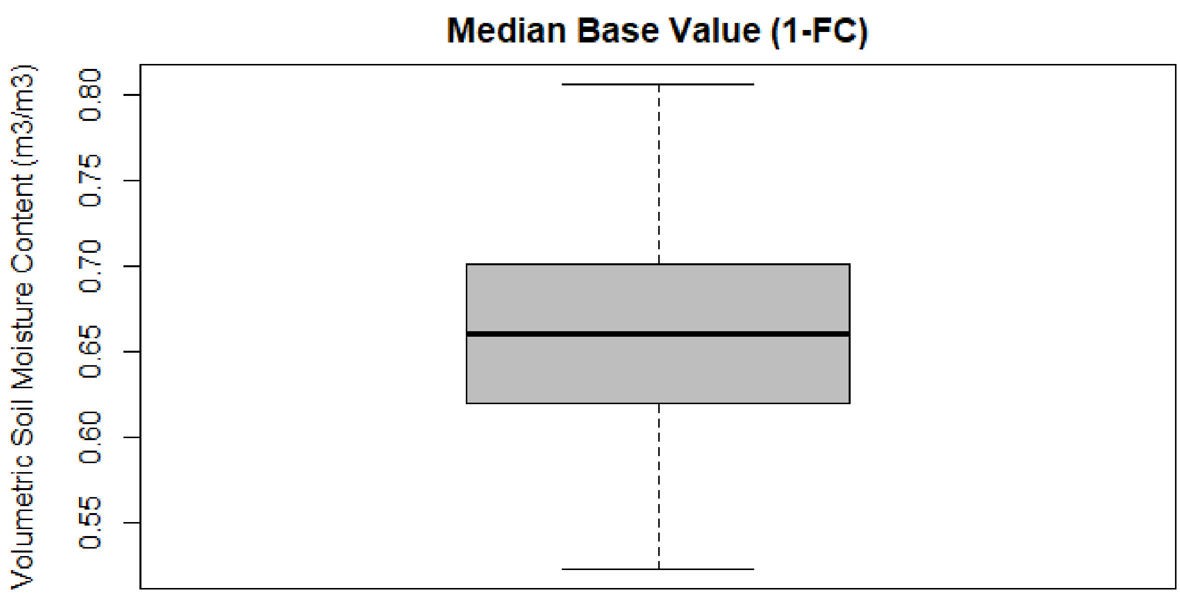

Figure A.6.2. Box-plot of the median base value (1-Field Capacity) for the Missouri and Arkansas-White-Red River basins.

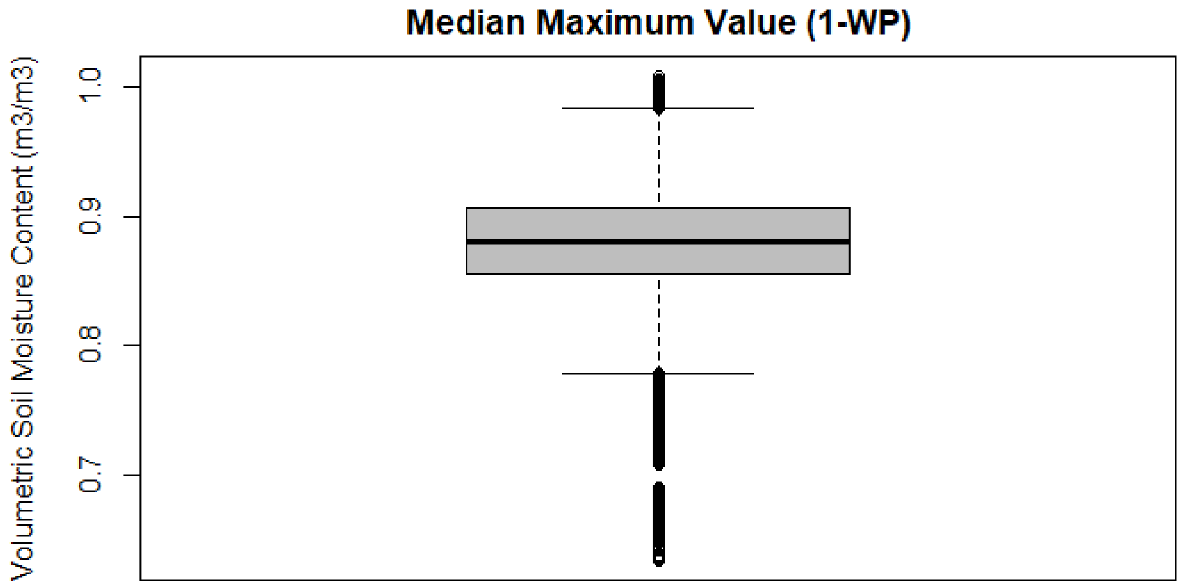

Figure A.6.3. Box-plot of the median maximum value (1-Wilting Point) for the Missouri and Arkansas-White-Red River basins.

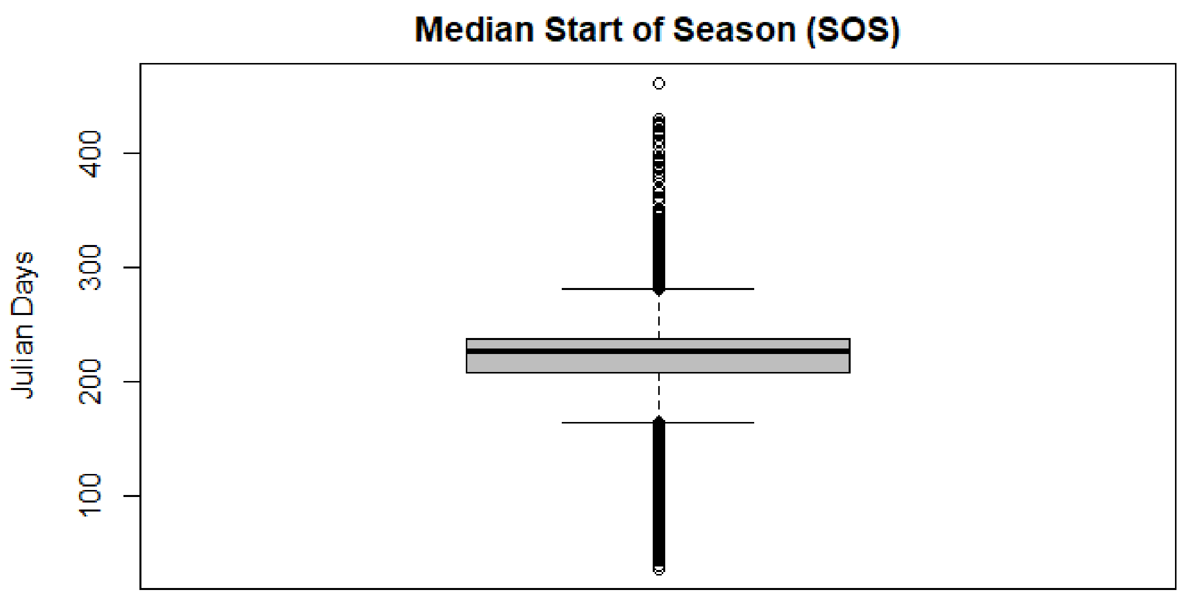

Figure A.6.4. Box-plot of the median start of season (SOS) for the Missouri and Arkansas-White-Red River basins.

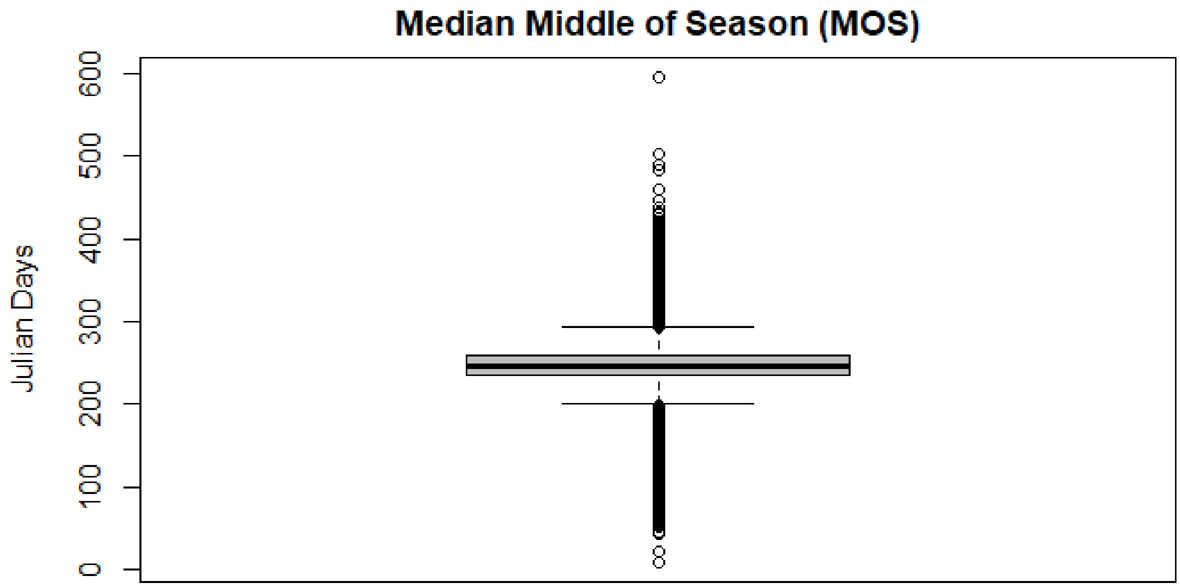

Figure A.6.5. Box-plot of the median middle of season (MOS) for the Missouri and Arkansas-White-Red River basins.

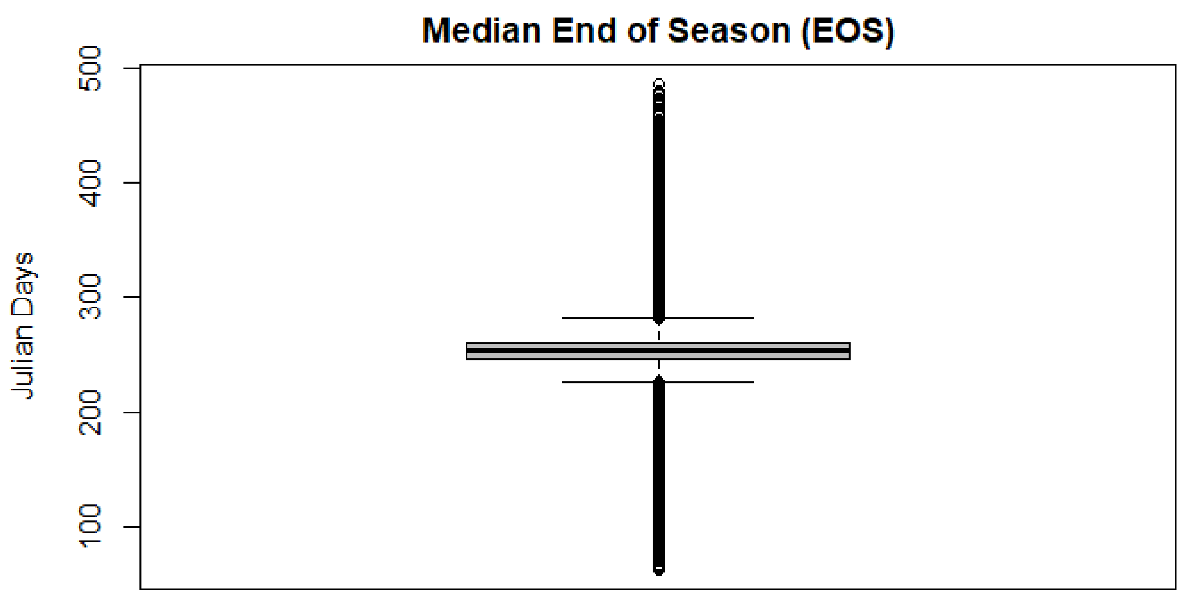

Figure A.6.6. Box-plot of the median end of season (EOS) for the Missouri and Arkansas-White-Red River basins.

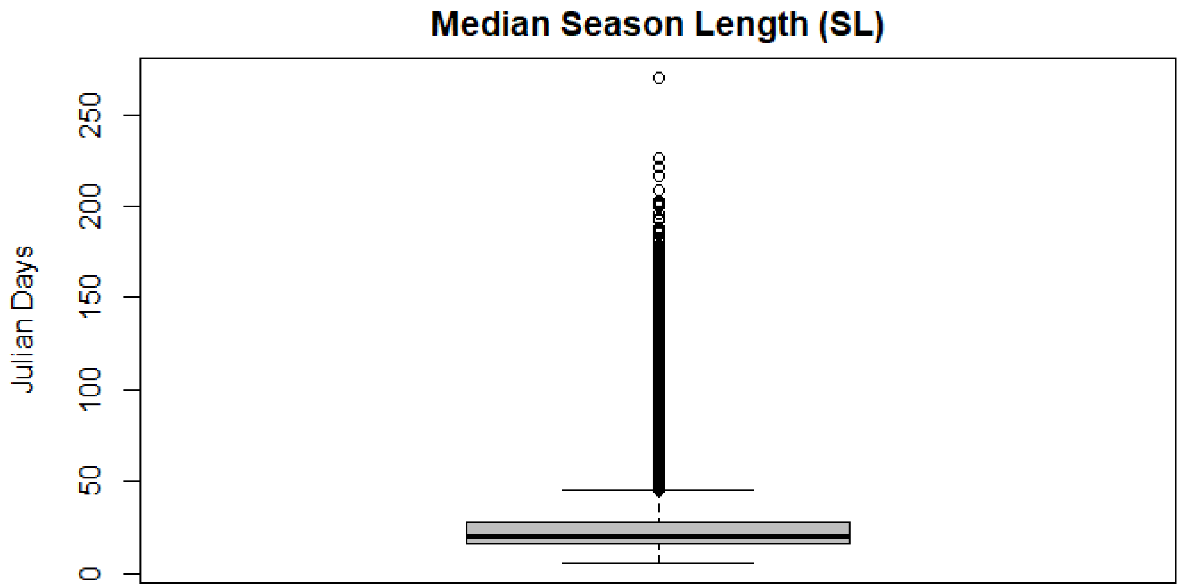

Figure A.6.7. Box-plot of the median season length (SL) for the Missouri and Arkansas-White-Red River basins.

Table A.6.1. The median soil moisture season start (SOS), middle (MOS), and end (EOS) by hydrologic soil group (HSG) over 30 years showing the spatial upper extreme, median, and lower extreme values

		A	B	C	D
UPPER EXTREME	SOS	20-May	23-Jun	10-Jun	21-Jun
	MOS	20-Jun	14-Jul	25-Jul	19-Jul
	EOS	19-Aug	11-Aug	9-Aug	17-Aug
MEDIAN	SOS	12-Jun	16-Aug	14-Aug	11-Aug
	MOS	15-Sep	1-Sep	1-Sep	29-Aug
	EOS	21-Sep	10-Sep	8-Sep	9-Sep
LOWER EXTREME	SOS	5-Jul	4-Oct	9-Oct	25-Sep
	MOS	6-Jan	25-Oct	9-Oct	9-Oct
	EOS	29-Oct	9-Oct	5-Oct	29-Sep

Appendix B - Results of the Statistical Analysis

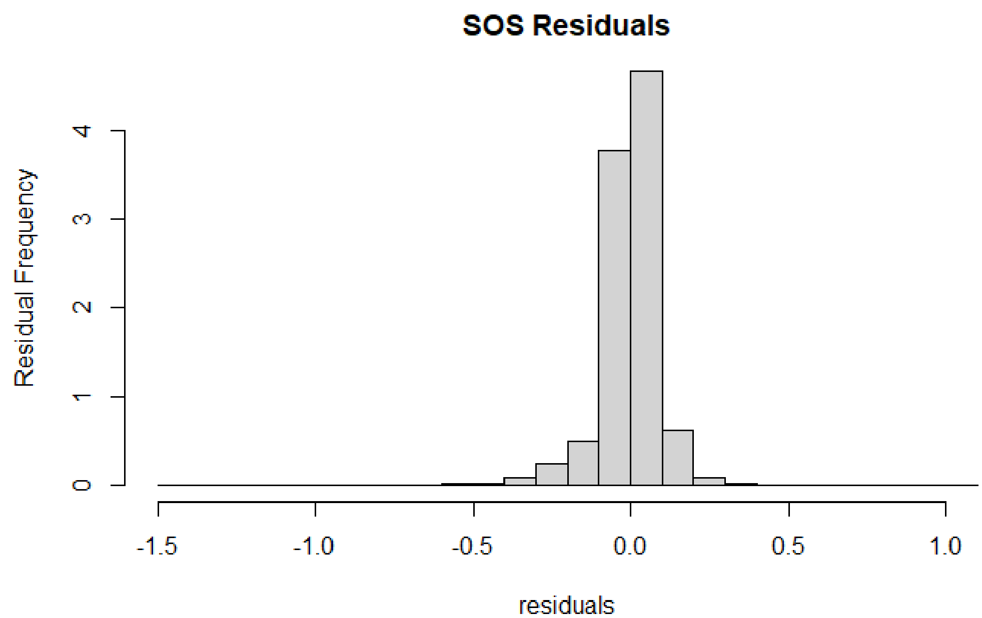

Figure B.6.8. A histogram plot showing the start of season (SOS) residuals.

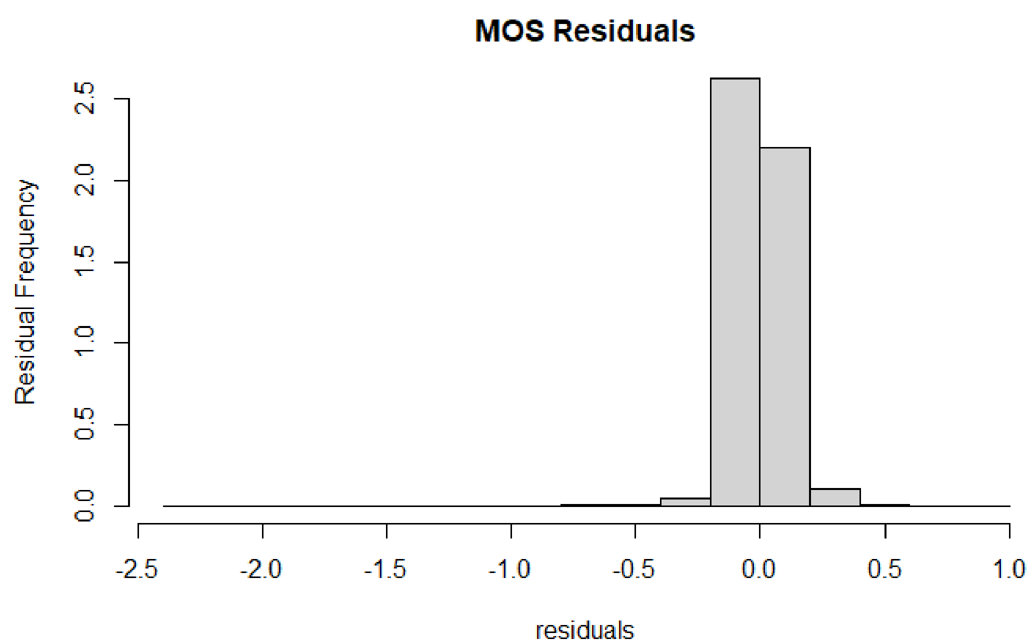

Figure B.6.9. A histogram plot showing the middle of season (MOS) residuals.

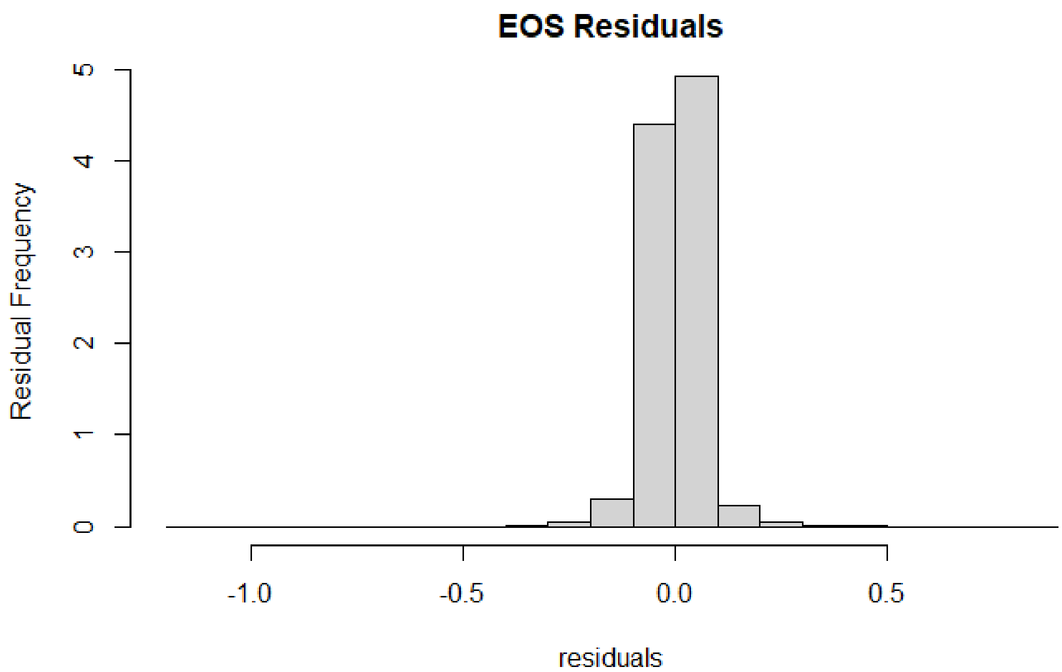

Figure B.6.10. A histogram plot showing the end of season (EOS) residuals.

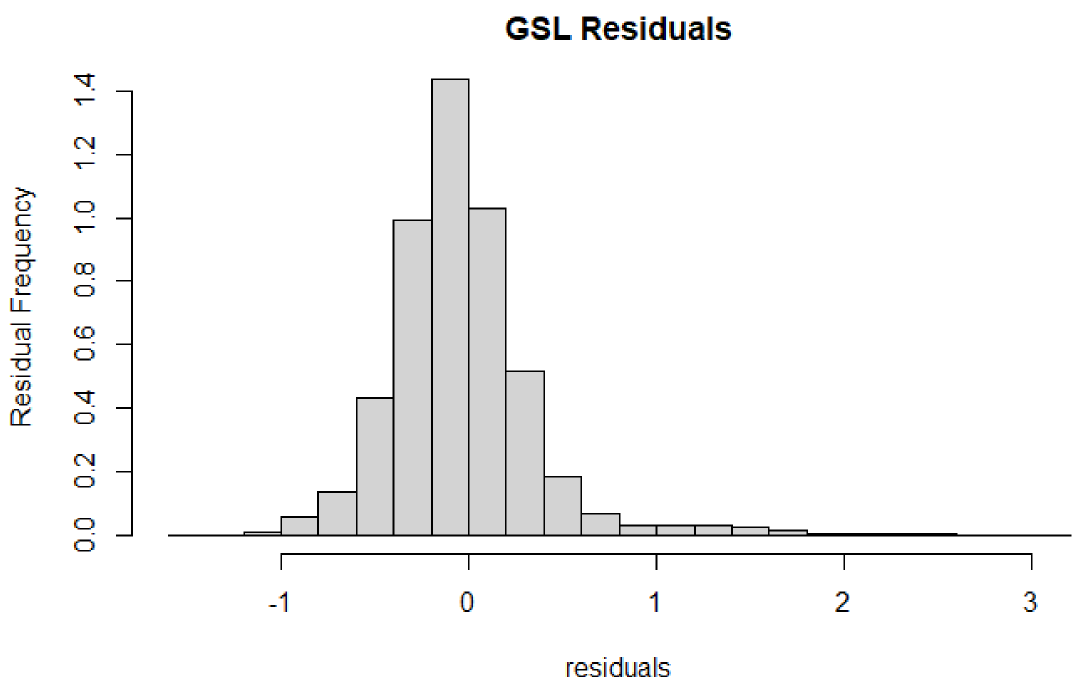

Figure B.6.11. A histogram plot showing the season length (GSL) residuals.

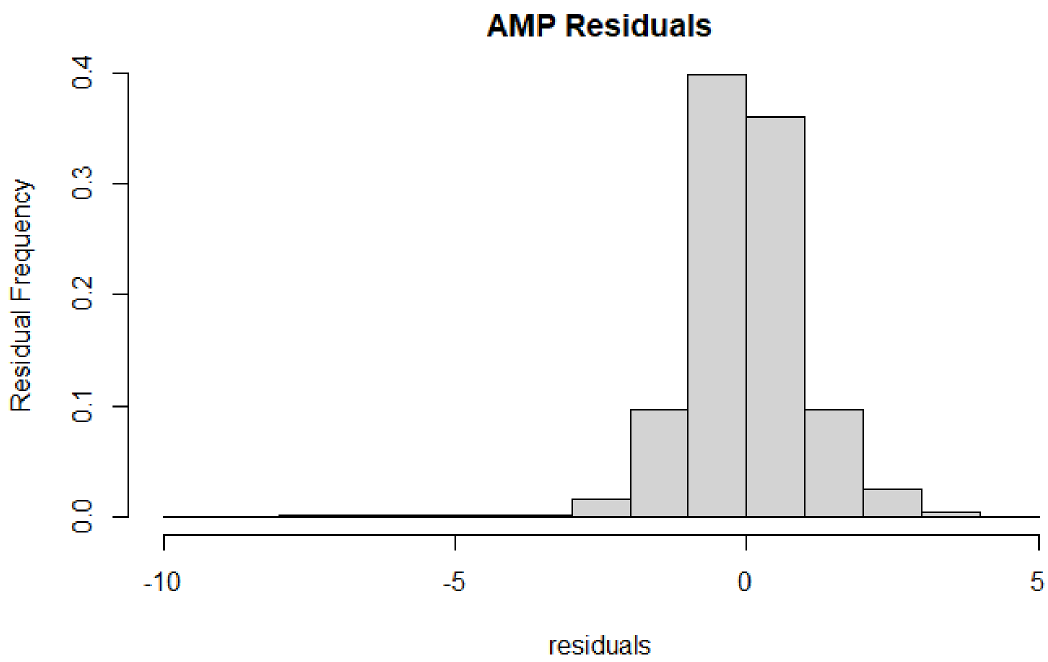

Figure B.6.12. A histogram plot showing the available water content (AMP) residuals.

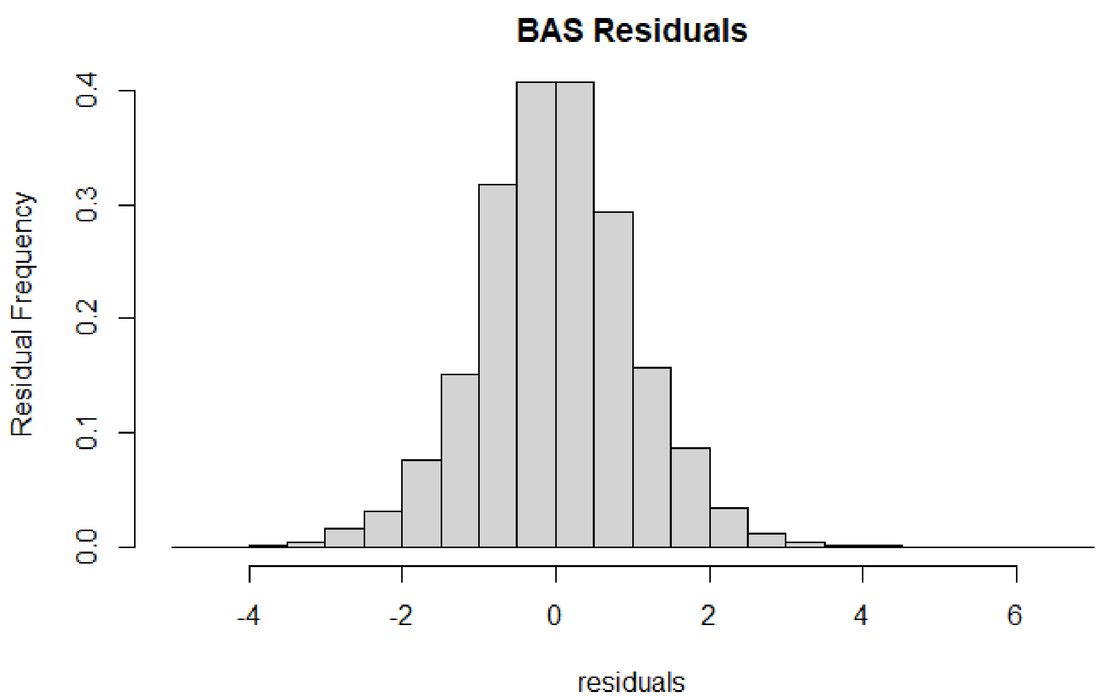

Figure B.6.13. A histogram plot showing the field capacity (BAS) residuals.

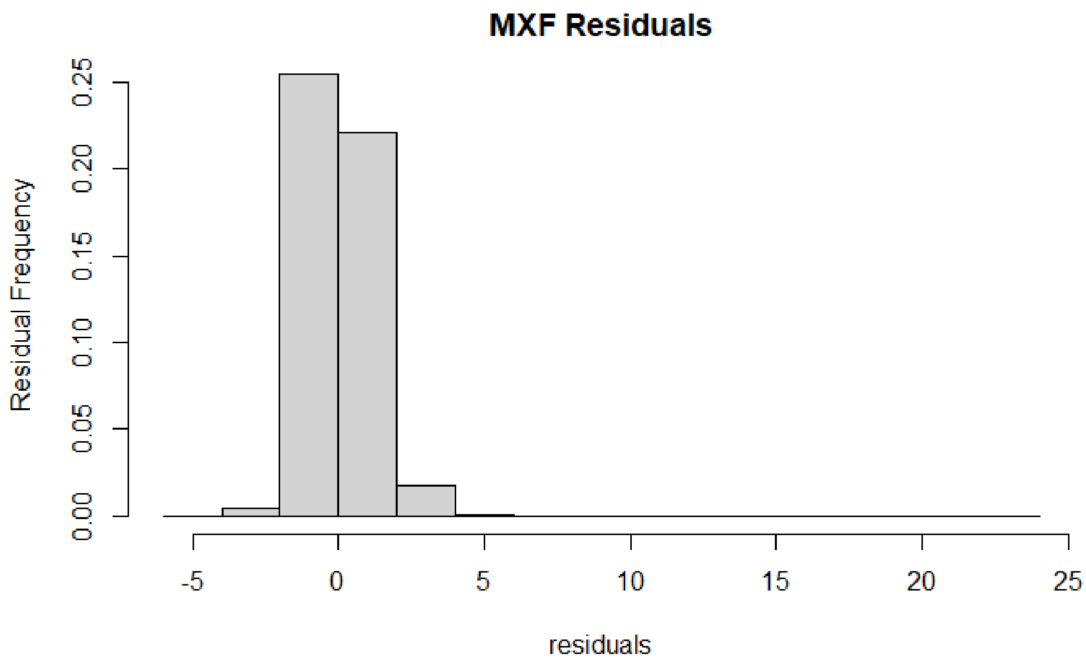

Figure B.6.14. A histogram plot showing the wilting point (MXF) residuals.

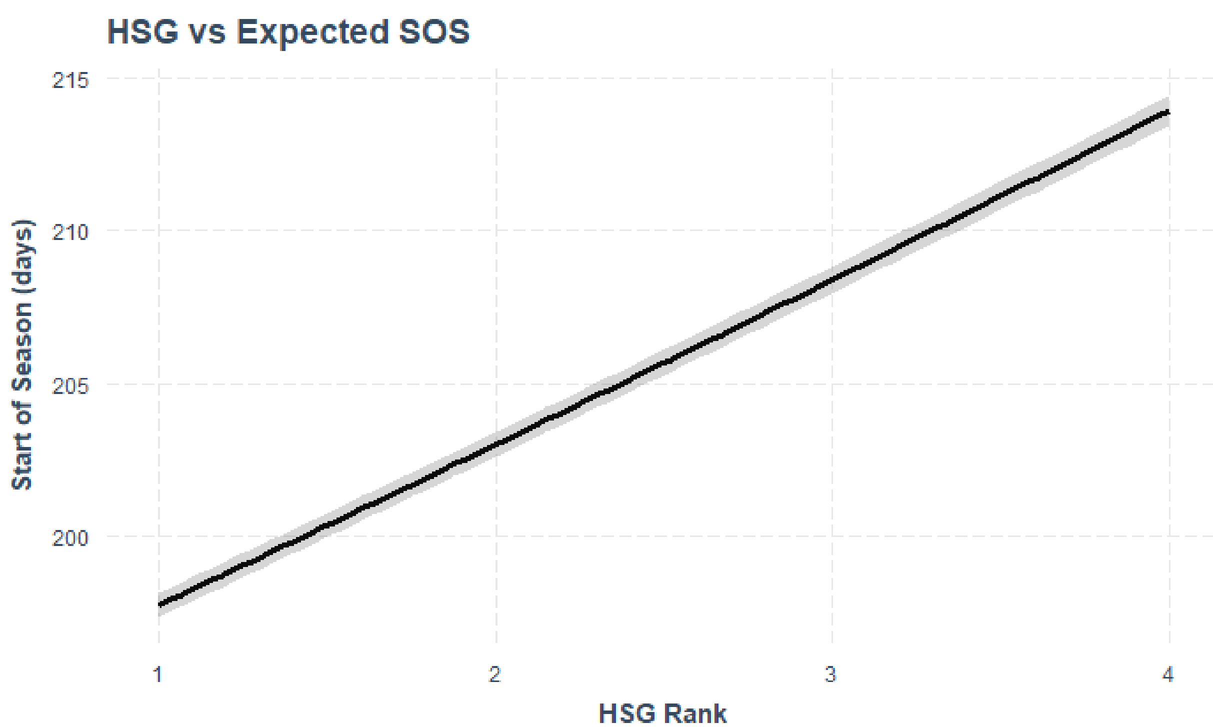

Figure B.6.15. A plot showing the impact of HSG (1=A, 2=B, 3=C, 4=D) on the expected value of the start of season (SOS) in Julian days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

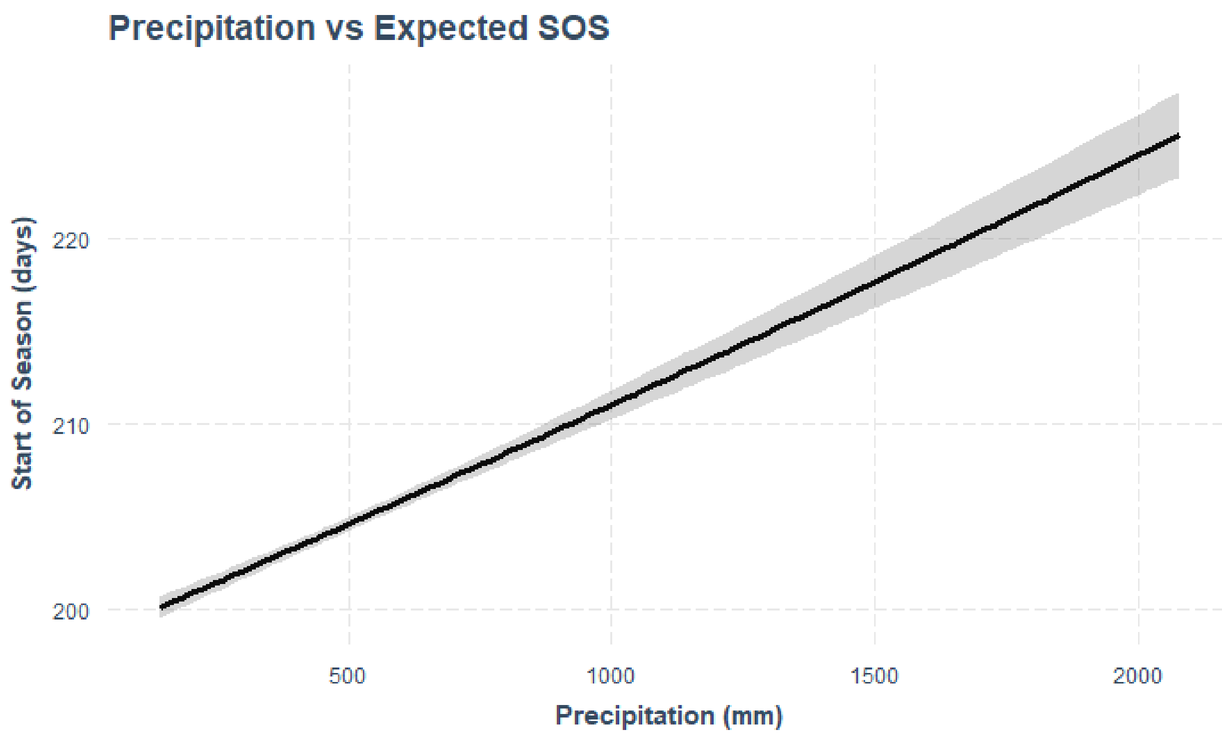

Figure B.6.16. A plot showing the impact of mean annual precipitation (mm) on the expected value of the start of season (SOS) in Julian days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

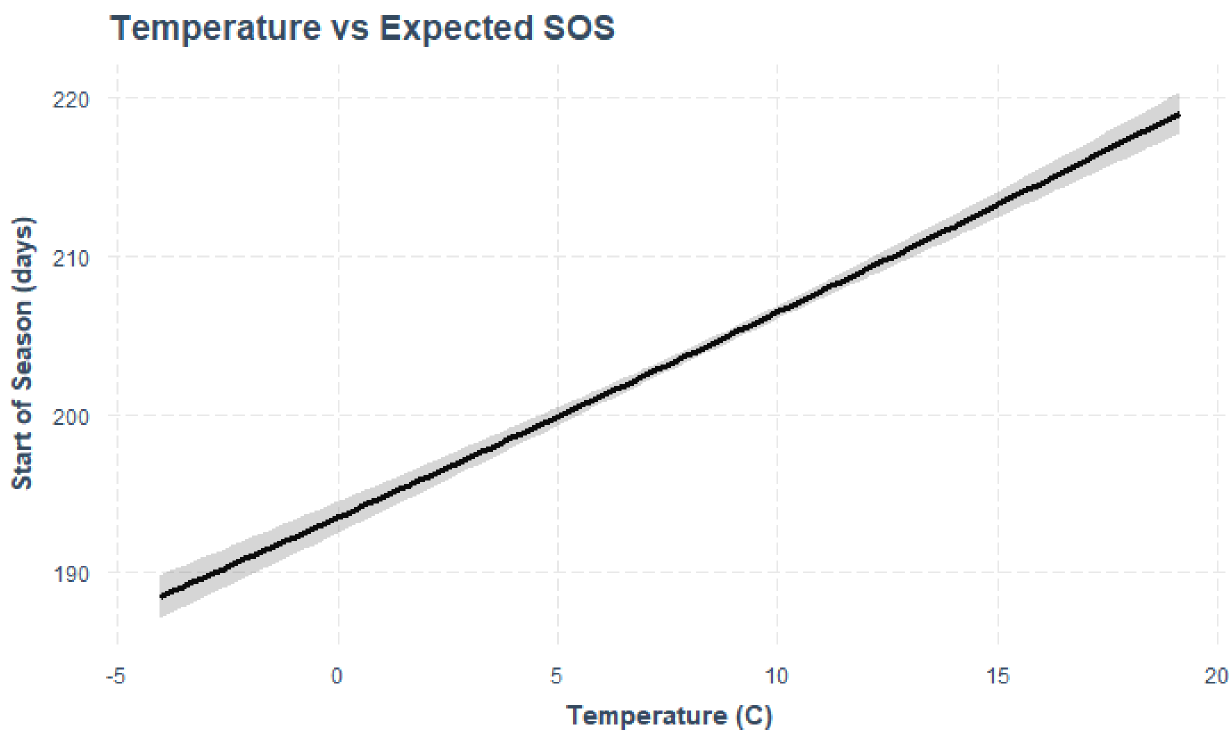

Figure B.6.17. A plot showing the impact of mean annual temperature (°C) on the expected value of the start of season (SOS) in Julian days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

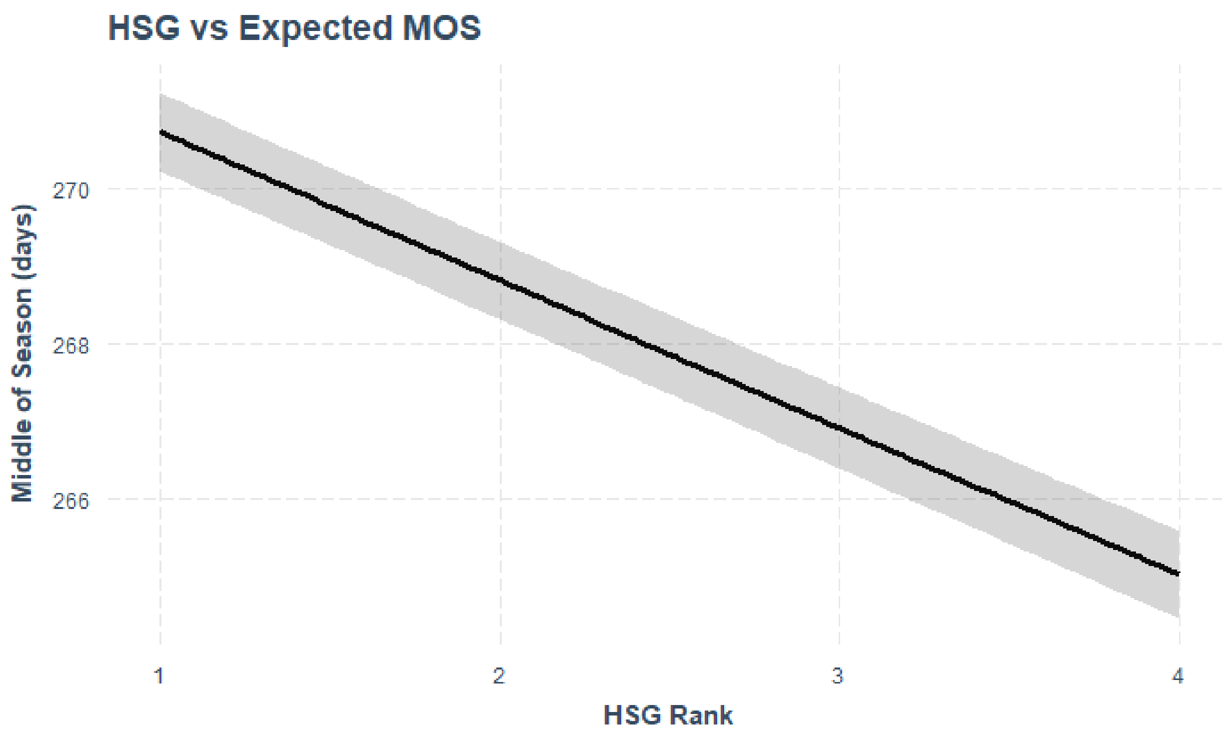

Figure B.6.18. A plot showing the impact of HSG (1=A, 2=B, 3=C, 4=D) on the expected value of the middle of season (MOS) in Julian days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

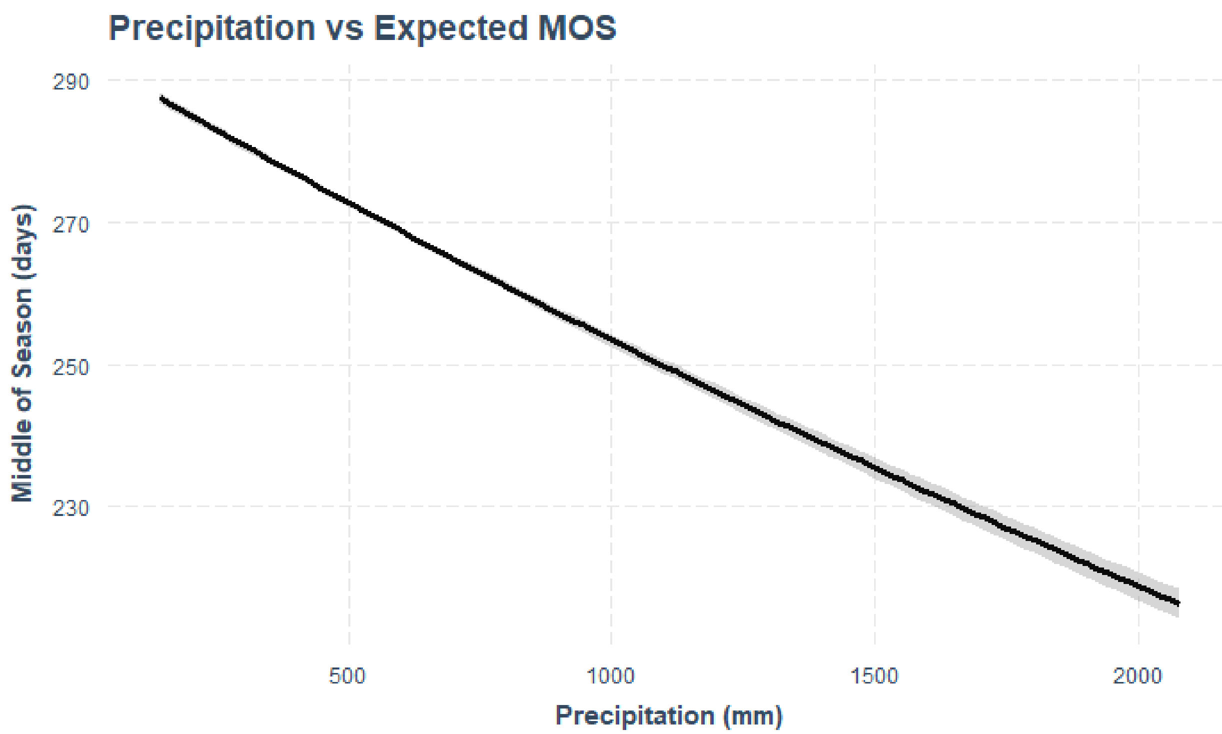

Figure B.6.19. A plot showing the impact of mean annual precipitation (mm) on the expected value of the middle of season (MOS) in Julian days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

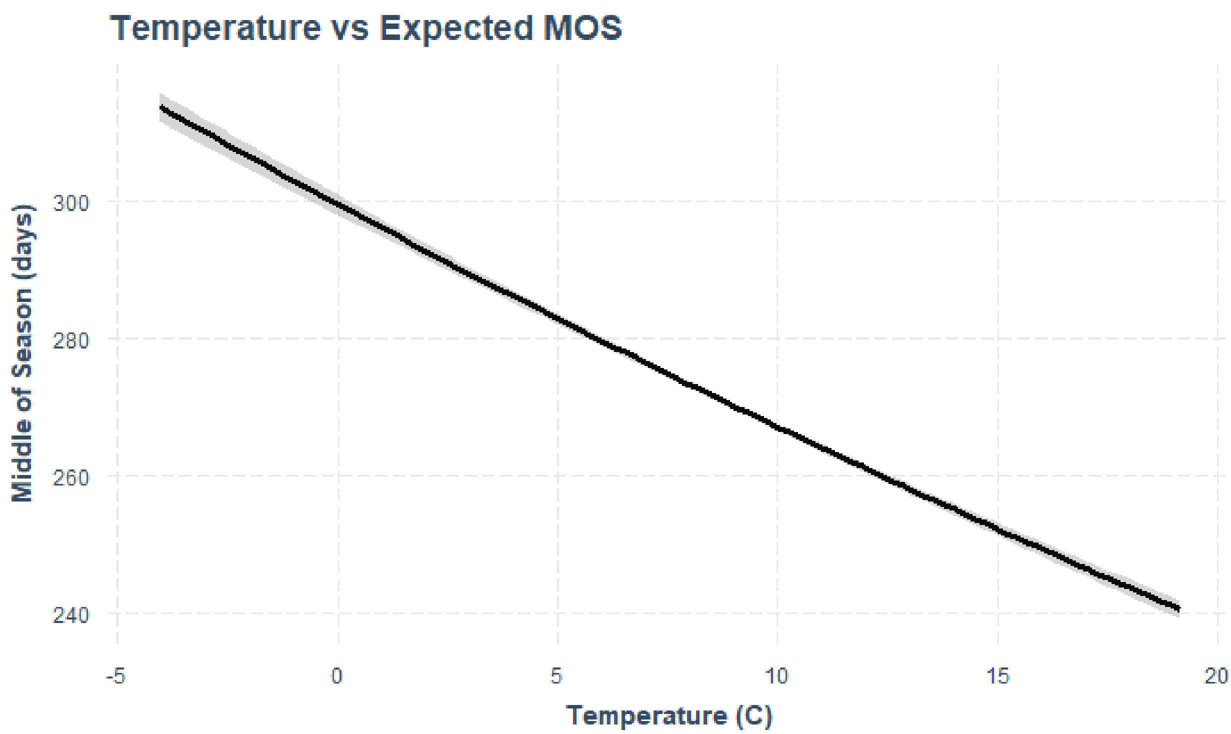

Figure B.6.20. A plot showing the impact of mean annual temperature (°C) on the expected value of the middle of season (MOS) in Julian days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

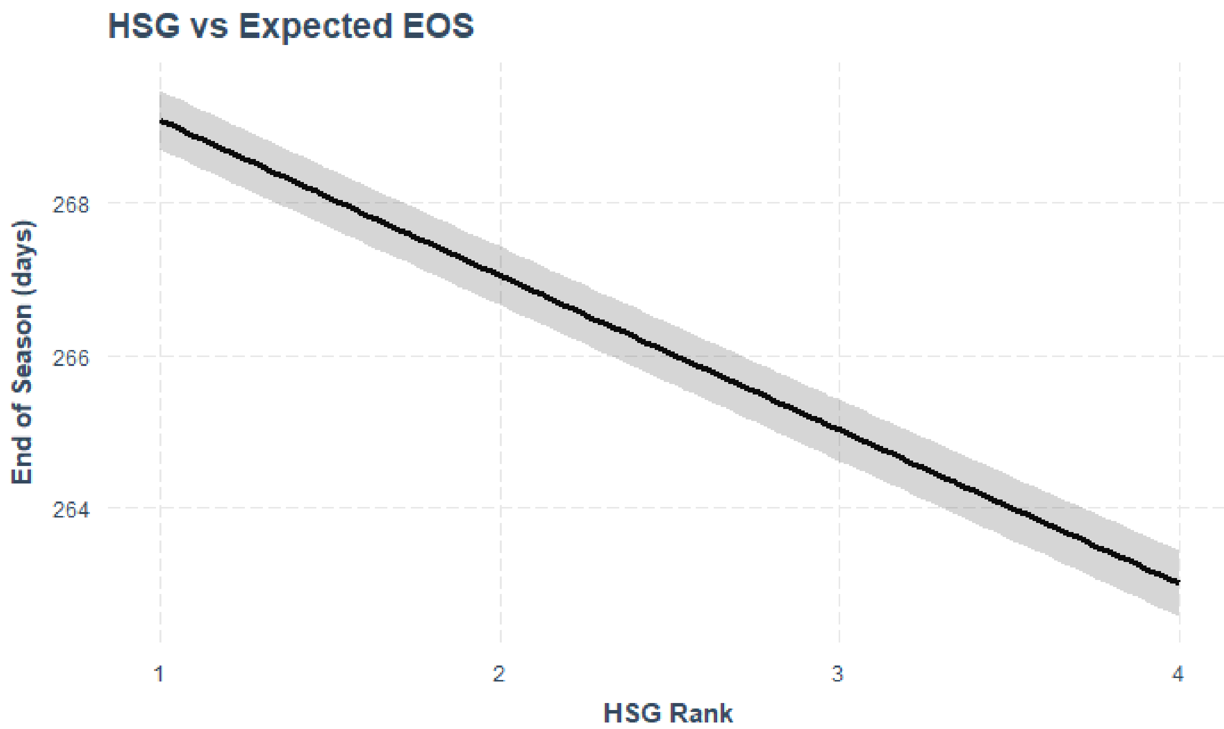

Figure B.6.21. **A plot showing the impact of HSG (1=A, 2=B, 3=C, 4=D) on the expected value of the end of season (EOS) in Julian days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.**

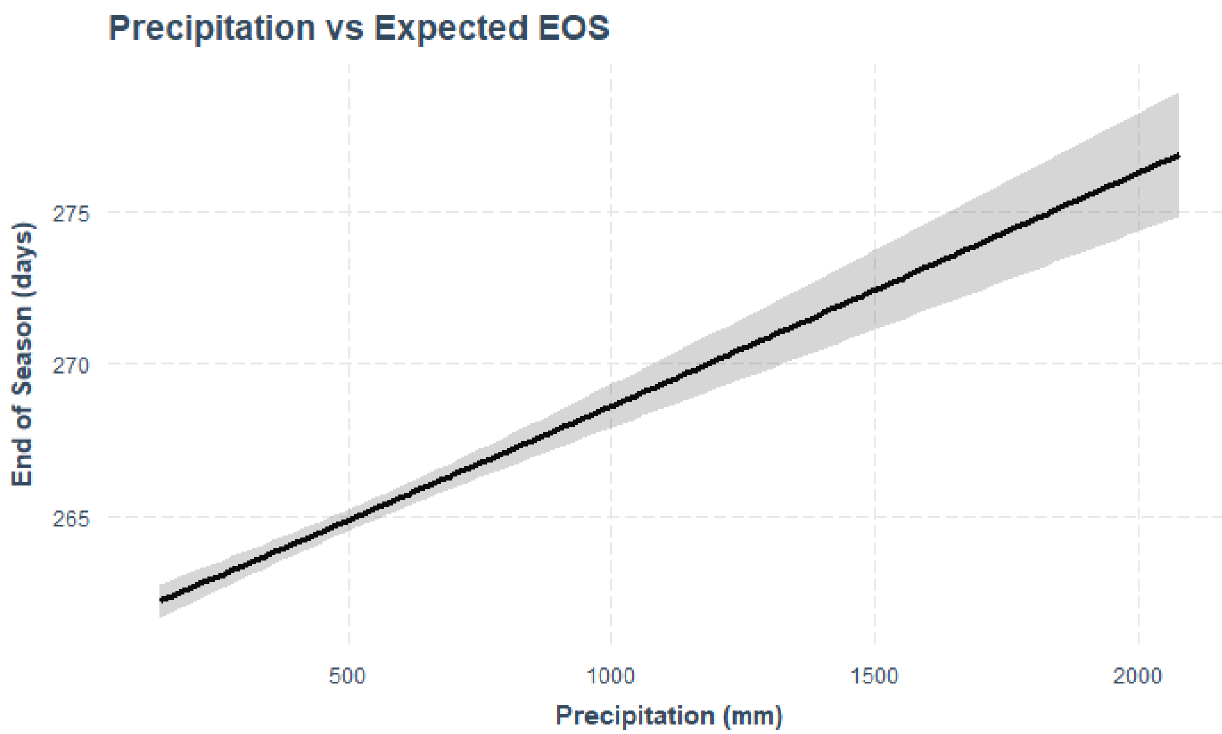

Figure B.6.22. A plot showing the impact of mean annual precipitation (mm) on the expected value of the end of season (EOS) in Julian days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

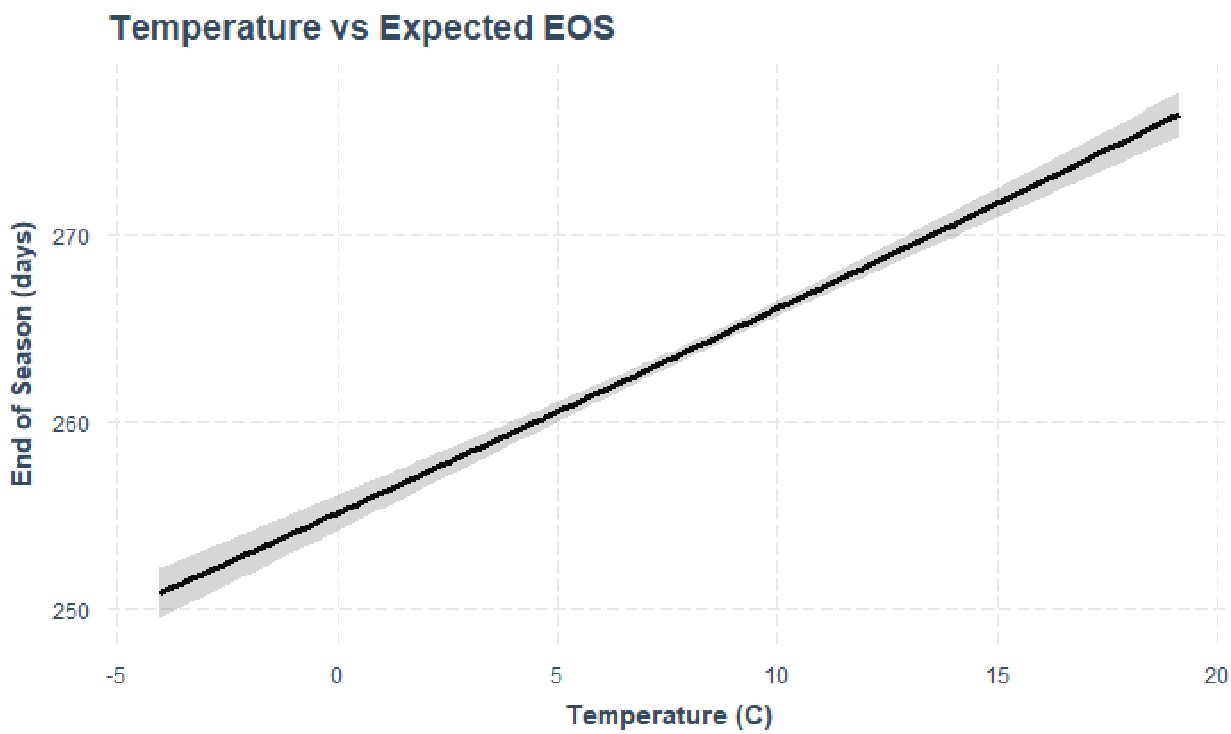

Figure B.6.23. A plot showing the impact of mean annual temperature (°C) on the expected value of the end of season (EOS) in Julian days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

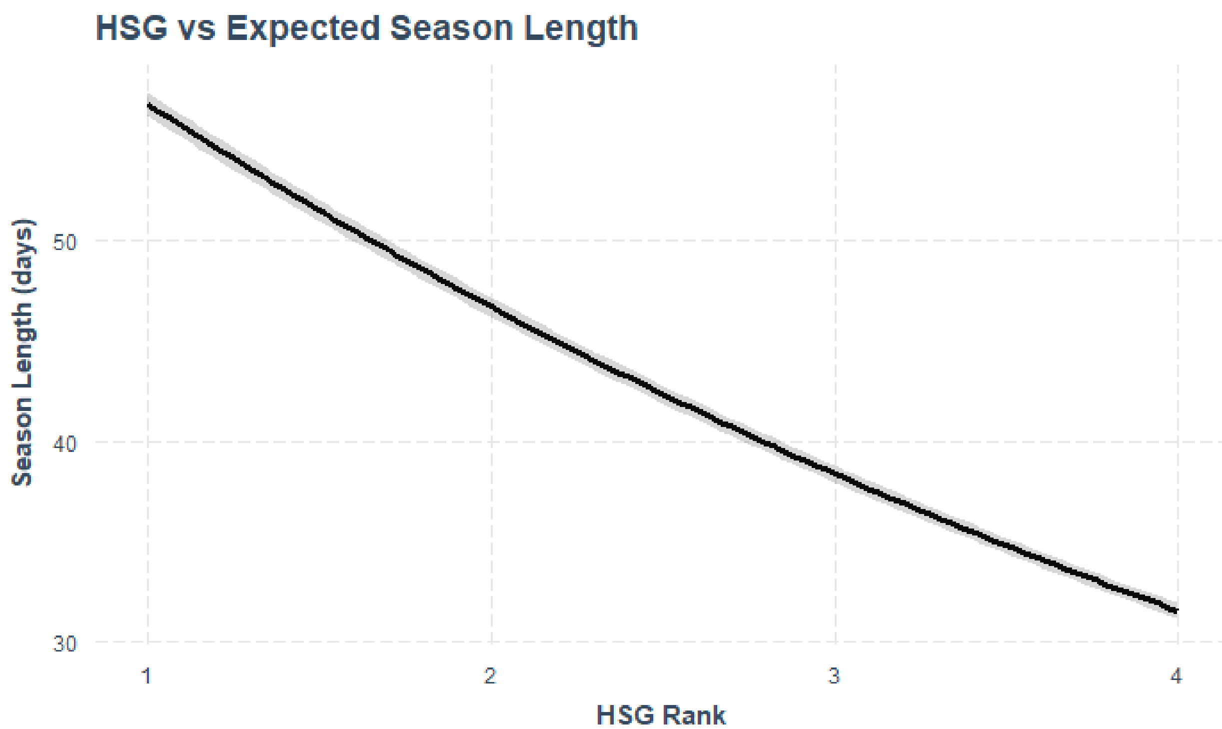

Figure B.6.24. A plot showing the impact of HSG (1=A, 2=B, 3=C, 4=D) on the expected season length (GSL) in days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

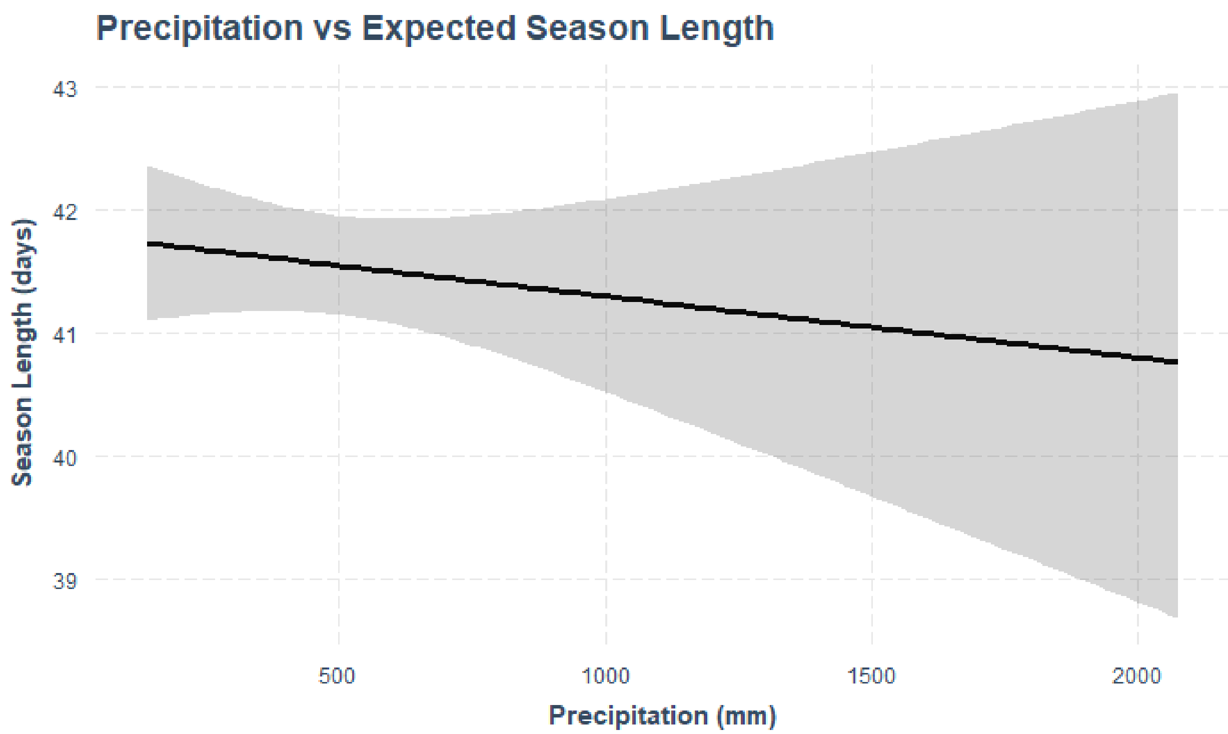

Figure B.6.25. A plot showing the impact of mean annual precipitation (mm) on the expected season length (GSL) in days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

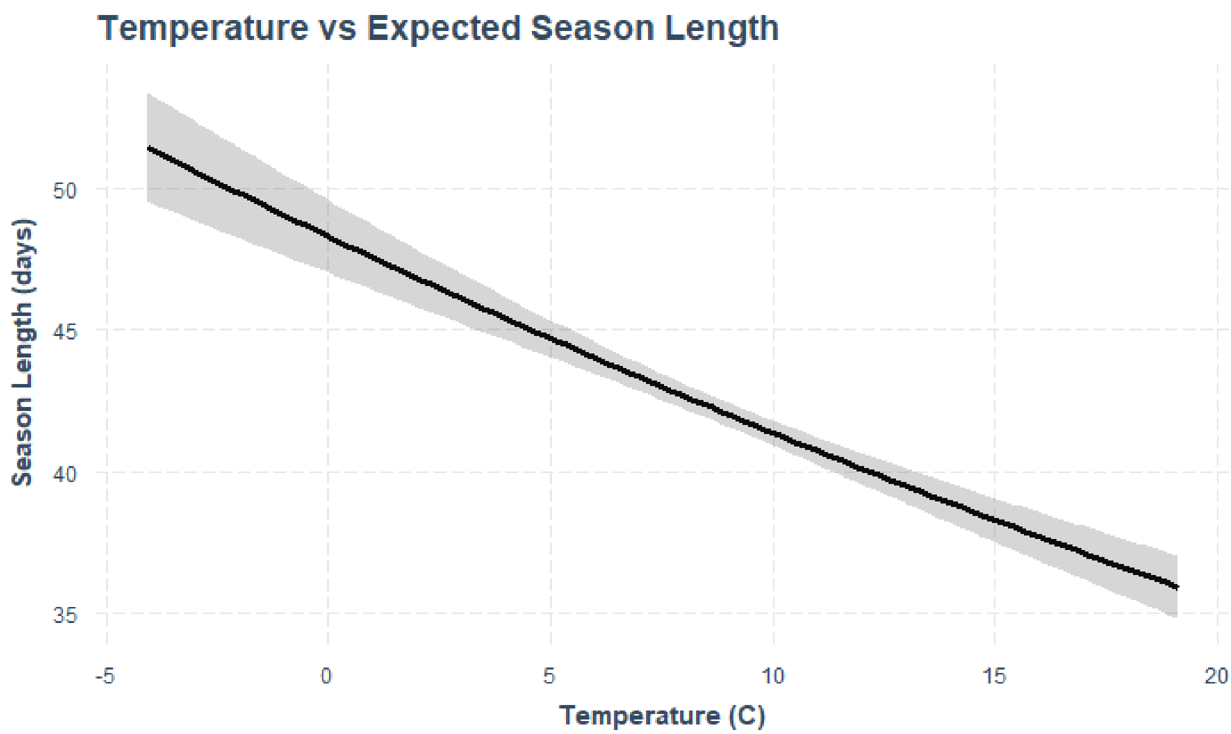

Figure B.6.26. A plot showing the impact of mean annual temperature (°C) on the expected season length (GSL) in days for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

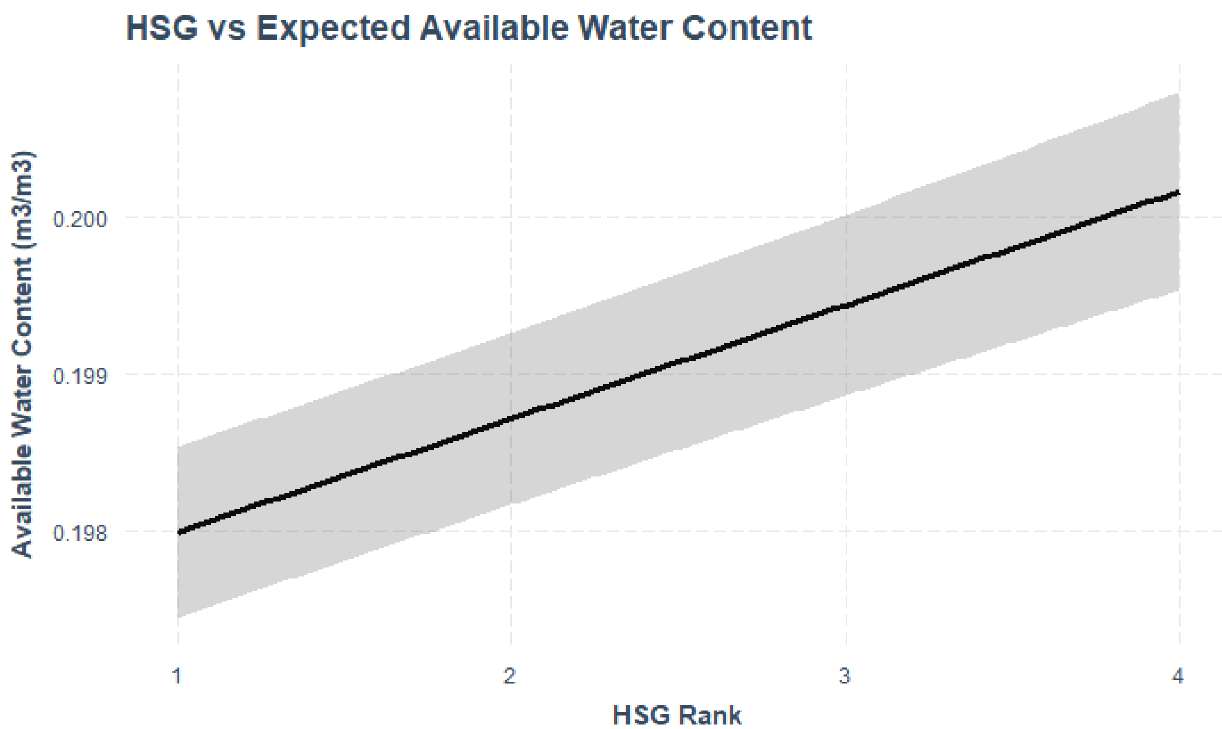

Figure B.6.27. A plot showing the impact of HSG (1=A, 2=B, 3=C, 4=D) on the expected value of the available water content (AMP) in m^3/m^3 for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

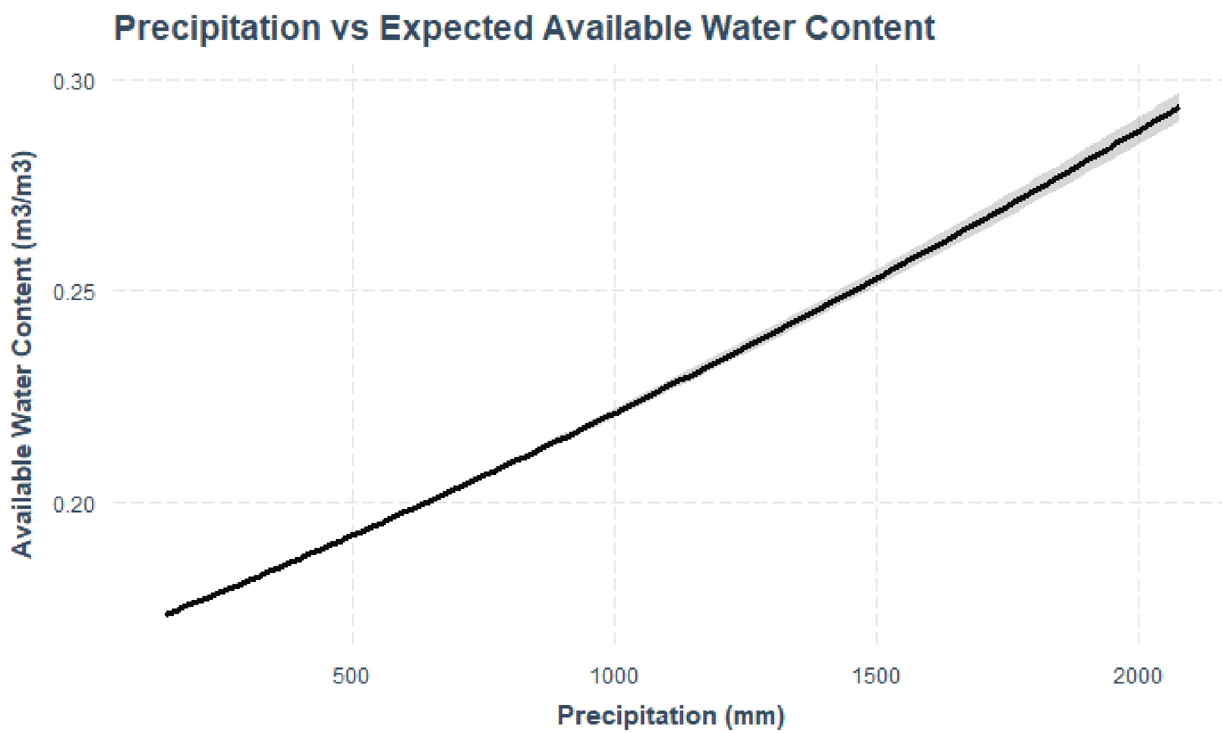

Figure B.6.28. A plot showing the impact of mean annual precipitation (mm) on the expected value of the available water content (AMP) in m^3/m^3 for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

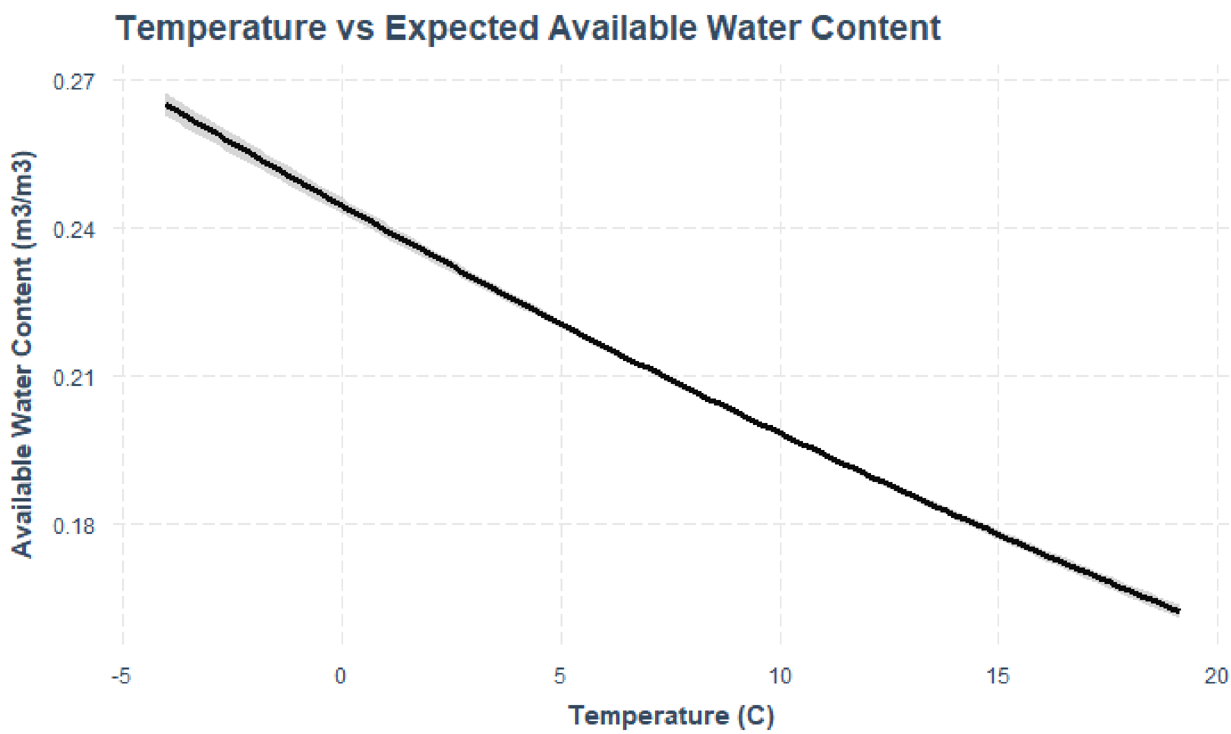

Figure B.6.29. A plot showing the impact of mean annual temperature (°C) on the expected value of the available water content (AMP) in m^3/m^3 for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

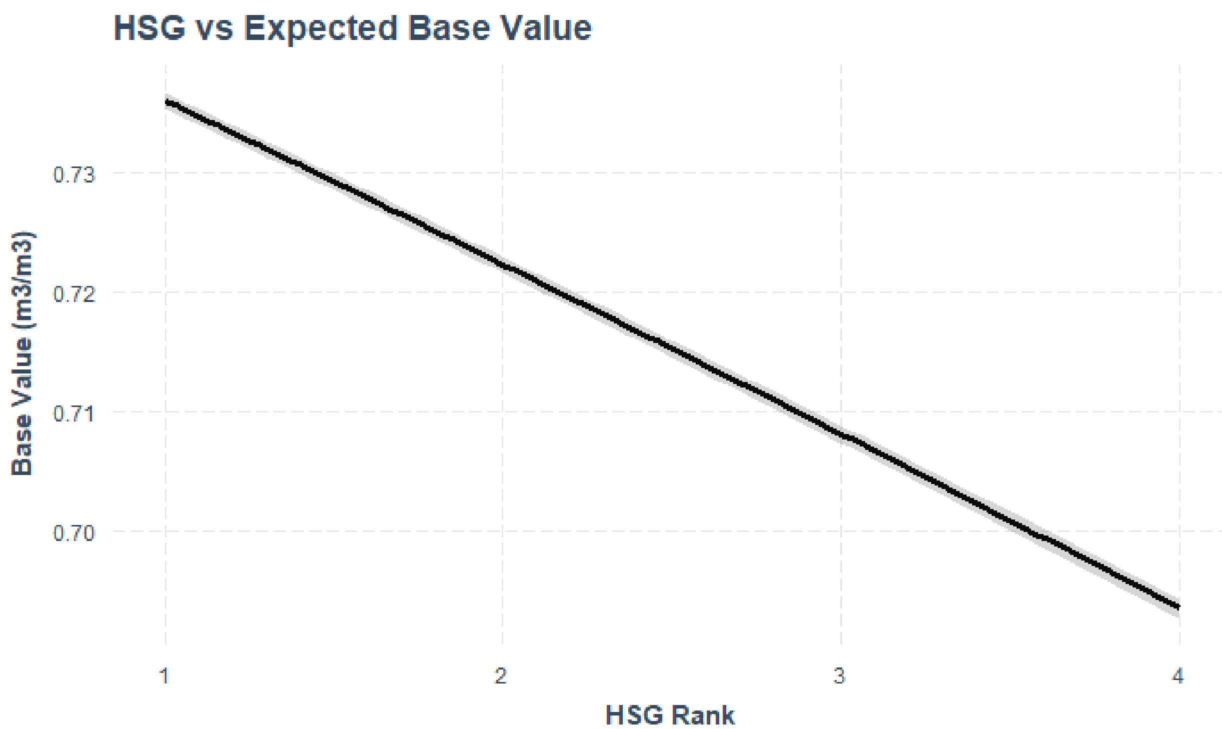

Figure B.6.30. A plot showing the impact of HSG (1=A, 2=B, 3=C, 4=D) on the expected value of the base value (BAS) in m^3/m^3 for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

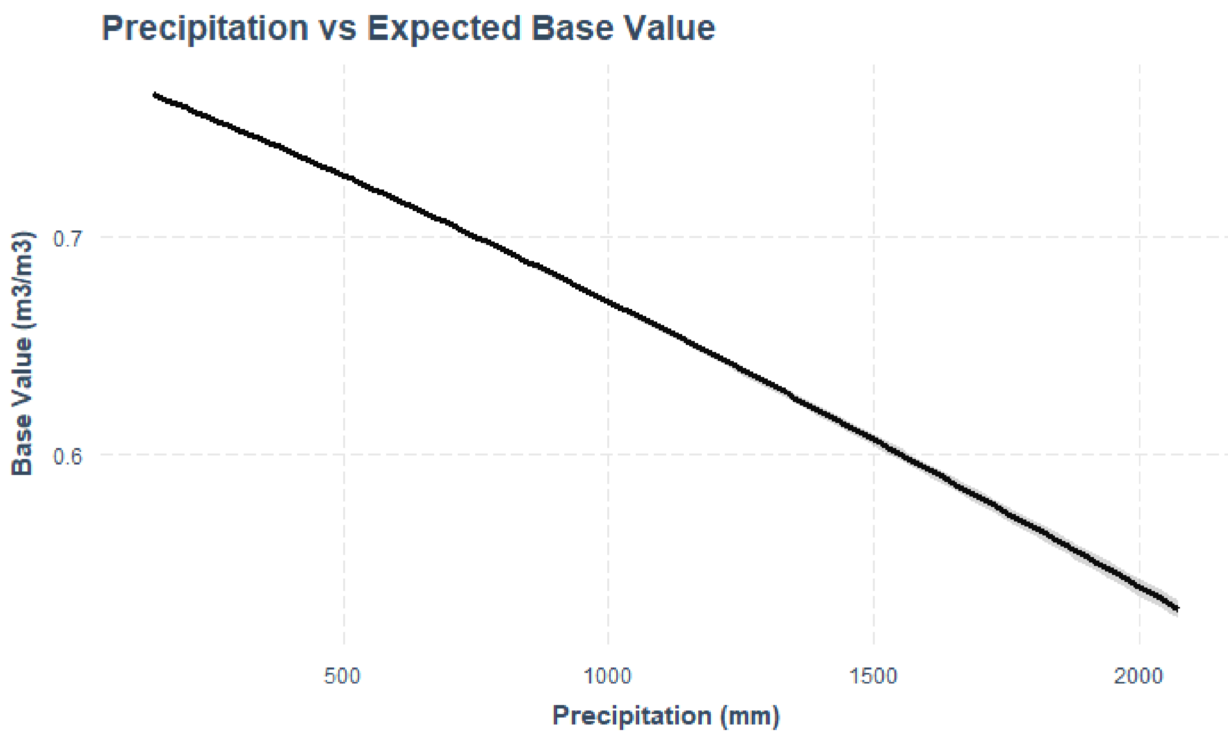

Figure B.6.31. A plot showing the impact of mean annual precipitation (mm) on the expected value of the base value (BAS) in m^3/m^3 for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

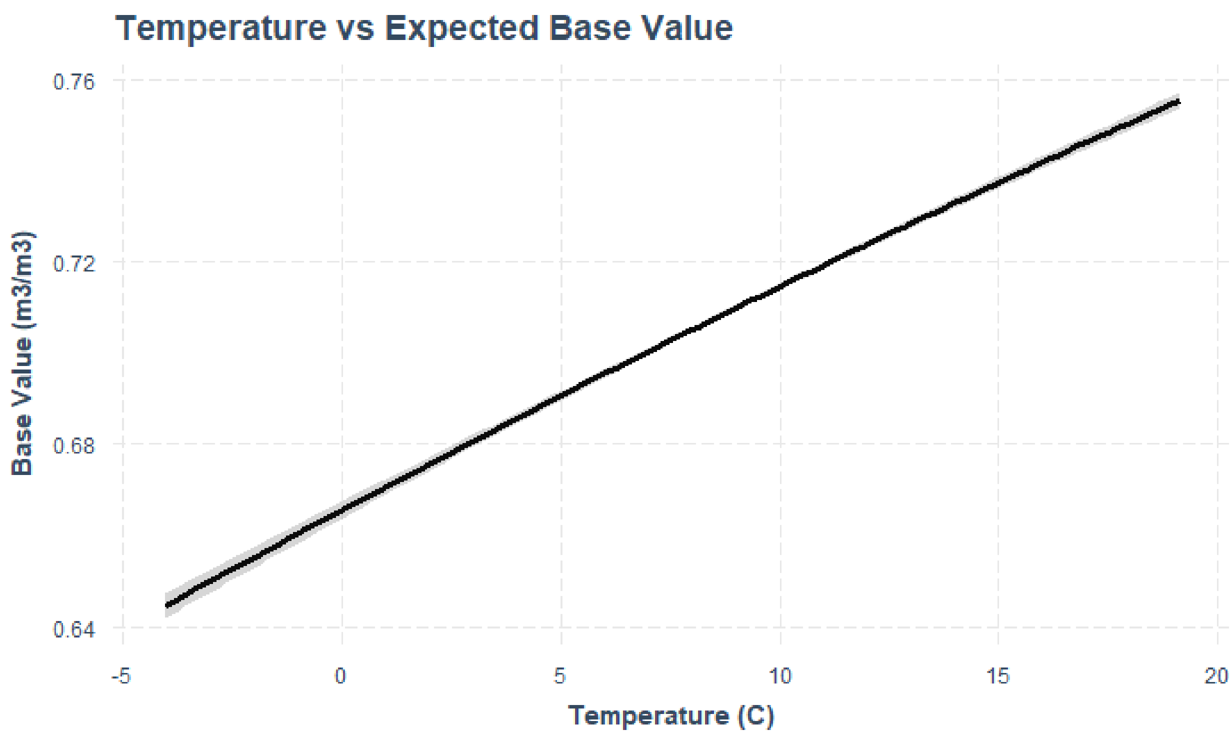

Figure B.6.32. A plot showing the impact of mean annual temperature (°C) on the expected value of the base value (BAS) in m^3/m^3 for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

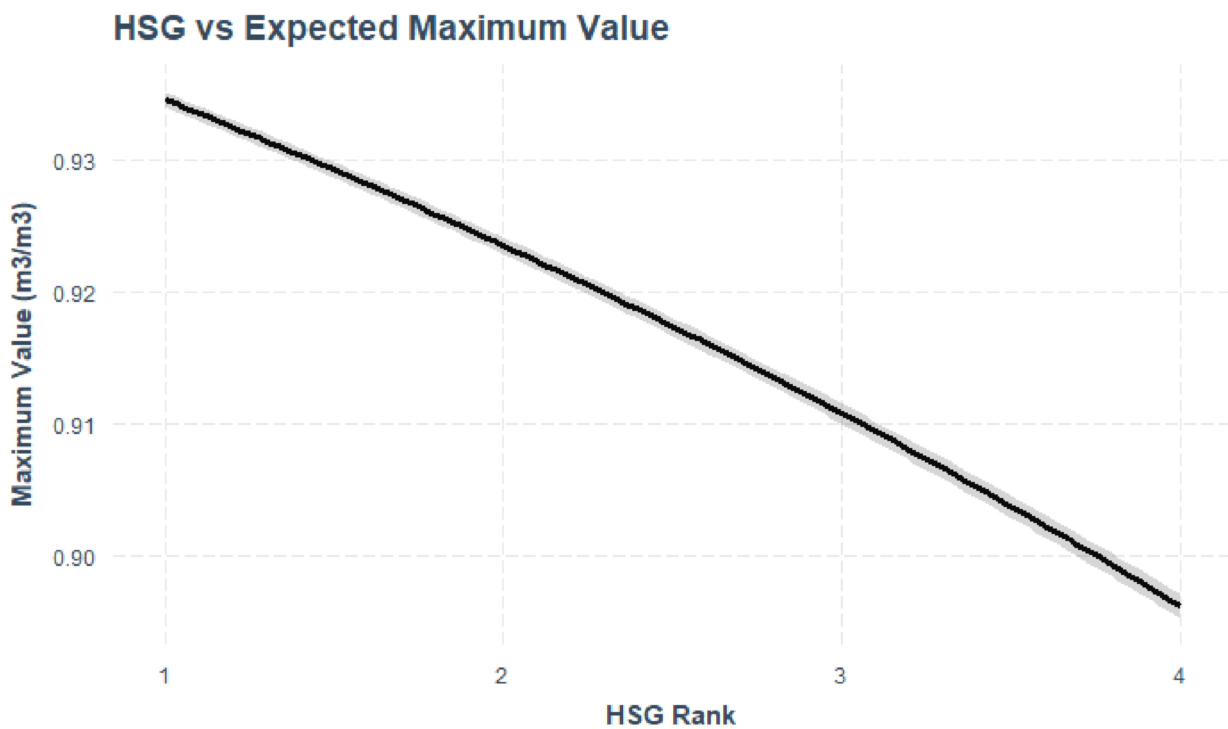

Figure B.6.33. A plot showing the impact of HSG (1=A, 2=B, 3=C, 4=D) on the expected value of the maximum value (MXF) in m^3/m^3 for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

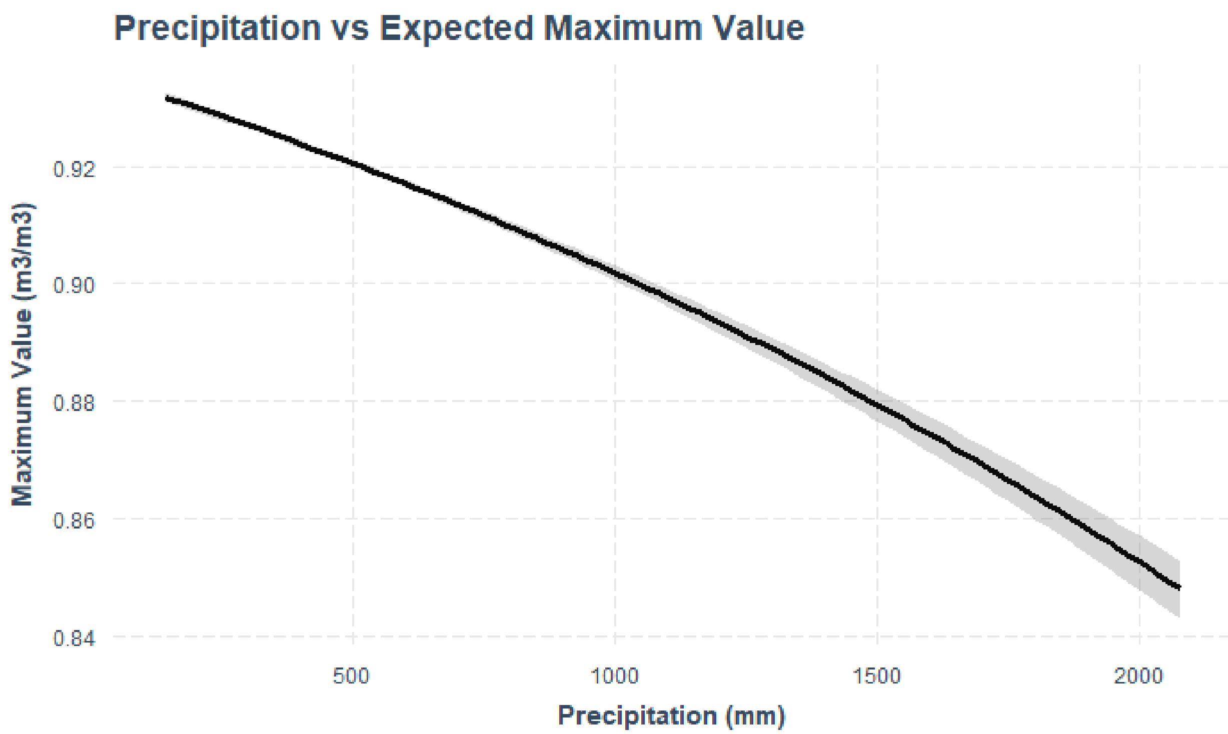

Figure B.6.34. A plot showing the impact of mean annual precipitation (mm) on the expected value of the maximum value (MXF) in m^3/m^3 for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

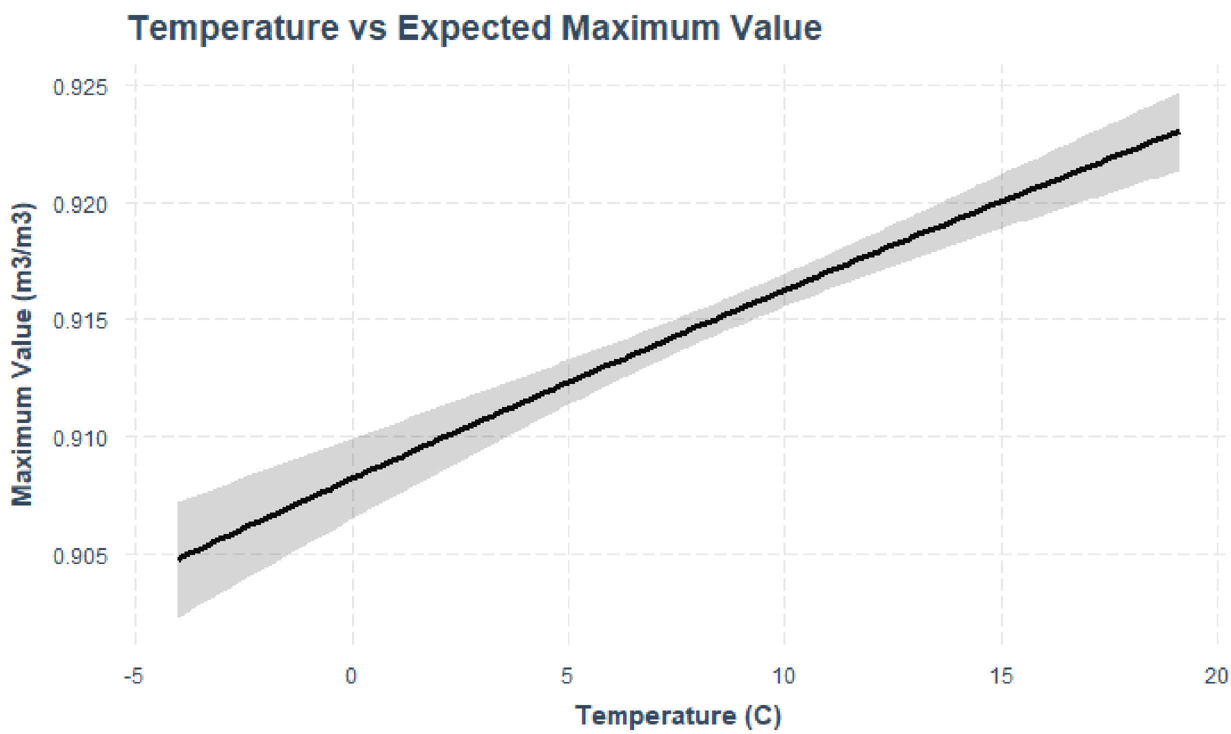

Figure B.6.35. A plot showing the impact of mean annual temperature (°C) on the expected value of the maximum value (MXF) in m^3/m^3 for the Missouri and Arkansas-White-Red river basins. The grey region shows 95% confidence intervals.

Appendix C - Statistical Analysis R Code

Emily Nottingham

1/29/2021

```
library(sf)
library(dplyr)
library(rgdal)
library(mgcv)
library(maps)
library(maptools)
library(plotrix)
library(gstat)
library(ggplot2)
library(jtools)
library(raster)

#load directories and subset dataframes

sos_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Media
n/SOS_0130_MEDIAN.tif")
mos_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Media
n/MOS_0130_MEDIAN.tif")
eos_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Media
n/EOS_0130_MEDIAN.tif")
gsl_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Median
/GSL_0130_MEDIAN.tif")
amp_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Medi
an/AMP_0130_MEDIAN.tif")
bas_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Media
n/BAS_0130_MEDIAN.tif")
mxf_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Media
n/MXF_0130_MEDIAN.tif")
Precip <-raster(x="C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/temp_precip/Precip_30yr_PRI
SM_AOI1.tif")
Temp <-raster(x="C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/temp_precip/Tmean_prism_AO
I1.tif")
HSG <- raster (x="C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/test/TIMESAT
_CODES1.tif")

stack_All <- stack(sos_median, mos_median, eos_median, gsl_median, amp_median,
                   bas_median, mxf_median, Precip, Temp, HSG)
median_All <- as.data.frame(stack_All)
median_All <- na.omit(median_All)
```

```r
#get lat and long
sos_pts <- rasterToPoints(sos_median)
mos_pts <- rasterToPoints(mos_median)
eos_pts <- rasterToPoints(eos_median)
gsl_pts <- rasterToPoints(gsl_median)
amp_pts <- rasterToPoints(amp_median)
bas_pts <- rasterToPoints(bas_median)
mxf_pts <- rasterToPoints(mxf_median)
precip_pts <- rasterToPoints(Precip)
temp_pts <- rasterToPoints(Temp)
HSG_pts  <- rasterToPoints(HSG)
head(sos_pts)
all_merge <- merge(data.frame(sos_pts, row.names=NULL),
                data.frame(mos_pts, row.names=NULL), all = TRUE)
all_merge <- merge (data.frame(all_merge, row.names=NULL),
                 data.frame(eos_pts, row.names=NULL), all = TRUE)
all_merge <- merge (data.frame(all_merge, row.names=NULL),
                 data.frame(gsl_pts, row.names=NULL), all = TRUE)
all_merge <- merge (data.frame(all_merge, row.names=NULL),
                 data.frame(amp_pts, row.names=NULL), all = TRUE)
all_merge <- merge (data.frame(all_merge, row.names=NULL),
                 data.frame(bas_pts, row.names=NULL), all = TRUE)
all_merge <- merge (data.frame(all_merge, row.names=NULL),
                 data.frame(mxf_pts, row.names=NULL), all = TRUE)
all_merge <- merge (data.frame(all_merge, row.names=NULL),
                 data.frame(precip_pts, row.names=NULL), all = TRUE)
all_merge <- merge (data.frame(all_merge, row.names=NULL),
                 data.frame(temp_pts, row.names=NULL), all = TRUE)
all_merge <- merge (data.frame(all_merge, row.names=NULL),
                 data.frame(HSG_pts, row.names=NULL), all = TRUE)

# Fit model to data for sos

#Assign model variables
y <- all_merge$SOS_0130_MEDIAN
x1 <- all_merge$Precip_30yr_PRISM_AOI1
x2 <- all_merge$Tmean_prism_AOI1
x3 <- all_merge$TIMESAT_CODES1
x4 <- all_merge$x
x5 <- all_merge$y


mSos <- bam(y~x1+x2+x3+
         s(x4,x5,bs="gp"),family = Gamma(link="log"))
summary(mSos)
```

```
# Fit model to data for mos

#Assign model variables
y <- all_merge$MOS_0130_MEDIAN
x1 <- all_merge$Precip_30yr_PRISM_AOI1
x2 <- all_merge$Tmean_prism_AOI1
x3 <- all_merge$TIMESAT_CODES1
x4 <- all_merge$x
x5 <- all_merge$y


mMos <- bam(y~x1+x2+x3+
        s(x4,x5,bs="gp"),family = Gamma(link="log"))
summary(mMos)

# Fit model to data for eos

#Assign model variables
y <- all_merge$EOS_0130_MEDIAN
x1 <- all_merge$Precip_30yr_PRISM_AOI1
x2 <- all_merge$Tmean_prism_AOI1
x3 <- all_merge$TIMESAT_CODES1
x4 <- all_merge$x
x5 <- all_merge$y


mEos <- bam(y~x1+x2+x3+
        s(x4,x5,bs="gp"),family = Gamma(link="log"))
summary(mEos)

# Fit model to data for gsl

#Assign model variables
y <- all_merge$GSL_0130_MEDIAN
x1 <- all_merge$Precip_30yr_PRISM_AOI1
x2 <- all_merge$Tmean_prism_AOI1
x3 <- all_merge$TIMESAT_CODES1
x4 <- all_merge$x
x5 <- all_merge$y


mGsl <- bam(y~x1+x2+x3+
        s(x4,x5,bs="gp"),family = Gamma(link="log"))
summary(mGsl)

# Fit model to data for amp

#Assign model variables
```

```r
y <- all_merge$AMP_0130_MEDIAN
x1 <- all_merge$Precip_30yr_PRISM_AOI1
x2 <- all_merge$Tmean_prism_AOI1
x3 <- all_merge$TIMESAT_CODES1
x4 <- all_merge$x
x5 <- all_merge$y


mAmp <- bam(y~x1+x2+x3+
        s(x4,x5,bs="gp"),family = betar(link="logit"))
summary(mAmp)

# Fit model to data for bas

#Assign model variables
y <- all_merge$BAS_0130_MEDIAN
x1 <- all_merge$Precip_30yr_PRISM_AOI1
x2 <- all_merge$Tmean_prism_AOI1
x3 <- all_merge$TIMESAT_CODES1
x4 <- all_merge$x
x5 <- all_merge$y


mBas <- bam(y~x1+x2+x3+
        s(x4,x5,bs="gp"),family = betar(link="logit"))
summary(mBas)

# Fit model to data for MXF

#Assign model variables
y <- all_merge$MXF_0130_MEDIAN
x1 <- all_merge$Precip_30yr_PRISM_AOI1
x2 <- all_merge$Tmean_prism_AOI1
x3 <- all_merge$TIMESAT_CODES1
x4 <- all_merge$x
x5 <- all_merge$y


mMxf <- bam(y~x1+x2+x3+
        s(x4,x5,bs="gp"),family = betar(link="logit"))
summary(mMxf)

# Examine regression coefficient estimates and 95% CI
beta.1.hat <- c(coef(mSos)[2], coef(mMos)[2], coef(mEos)[2], coef(mGsl)[2],
            coef(mAmp)[2], coef(mBas)[2], coef(mMxf)[2])
beta.1.hat # order is sos, mos, eos, gsl, amp, bas, mxf

beta.2.hat <- c(coef(mSos)[3], coef(mMos)[3], coef(mEos)[3], coef(mGsl)[3],
```

```
              coef(mAmp)[3], coef(mBas)[3], coef(mMxf)[3])
beta.2.hat

beta.3.hat <- c(coef(mSos)[4], coef(mMos)[4], coef(mEos)[4], coef(mGsl)[4],
              coef(mAmp)[4], coef(mBas)[4], coef(mMxf)[4])
beta.3.hat

#uncertainty for Precipitation
ucl2 <- c(confint.default(mSos,parm="x1")[2],confint.default(mMos,parm="x1")[2],
        confint.default(mEos,parm="x1")[2],confint.default(mGsl,parm="x1")[2],
        confint.default(mAmp,parm="x1")[2],confint.default(mBas,parm="x1")[2],
        confint.default(mMxf,parm="x1")[2])

lcl2 <- c(confint.default(mSos,parm="x1")[1],confint.default(mMos,parm="x1")[1],
        confint.default(mEos,parm="x1")[1],confint.default(mGsl,parm="x1")[1],
        confint.default(mAmp,parm="x1")[1],confint.default(mBas,parm="x1")[1],
        confint.default(mMxf,parm="x1")[1])

par(mar=c(4,7,1,1))
plotCI(c(1:7), beta.1.hat, ui=ucl2, li=lcl2, pch=20, xaxt="n",xlab="",
      ylab="Estimated regression coefficient \n (Precipitation)")
lines(c(0,7),c(0,0),col="gold",lwd=3)
axis(at=c(1:7),lab=c("SOS", "MOS", "EOS", "GSL", "AMP", "BAS", "MXF"),side=1)

#uncertainty for Temperature
ucl2 <- c(confint.default(mSos,parm="x2")[2],confint.default(mMos,parm="x2")[2],
        confint.default(mEos,parm="x2")[2],confint.default(mGsl,parm="x2")[2],
        confint.default(mAmp,parm="x2")[2],confint.default(mBas,parm="x2")[2],
        confint.default(mMxf,parm="x2")[2])

lcl2 <- c(confint.default(mSos,parm="x2")[1],confint.default(mMos,parm="x2")[1],
        confint.default(mEos,parm="x2")[1],confint.default(mGsl,parm="x2")[1],
        confint.default(mAmp,parm="x2")[1],confint.default(mBas,parm="x2")[1],
        confint.default(mMxf,parm="x2")[1])

par(mar=c(4,7,1,1))
plotCI(c(1:7), beta.2.hat, ui=ucl2, li=lcl2, pch=20, xaxt="n",xlab="",
      ylab="Estimated regression coefficient \n (Temperature)")
lines(c(0,7),c(0,0),col="gold",lwd=3)
axis(at=c(1:7),lab=c("SOS", "MOS", "EOS", "GSL", "AMP", "BAS", "MXF"),side=1)

#uncertainty for HSG
ucl2 <- c(confint.default(mSos,parm="x3")[2],confint.default(mMos,parm="x3")[2],
        confint.default(mEos,parm="x3")[2],confint.default(mGsl,parm="x3")[2],
        confint.default(mAmp,parm="x3")[2],confint.default(mBas,parm="x3")[2],
        confint.default(mMxf,parm="x3")[2])

lcl2 <- c(confint.default(mSos,parm="x3")[1],confint.default(mMos,parm="x3")[1],
```

```
        confint.default(mEos,parm="x3")[1],confint.default(mGsl,parm="x3")[1],
        confint.default(mAmp,parm="x3")[1],confint.default(mBas,parm="x3")[1],
        confint.default(mMxf,parm="x3")[1])

par(mar=c(4,7,1,1))
plotCI(c(1:7), beta.3.hat, ui=ucl2, li=lcl2, pch=20, xaxt="n",xlab="",
     ylab="Estimated regression coefficient \n (HSG)")
lines(c(0,7),c(0,0),col="gold",lwd=3)
axis(at=c(1:7),lab=c("SOS", "MOS", "EOS", "GSL", "AMP", "BAS", "MXF"),side=1)

#plot expected value of y vs individual predictor variables
effect_plot(mSos, pred = x1, interval = TRUE,
         main.title="Precipitation vs Expected SOS",
         x.label="Precipitation (mm)",
         y.label="Start of Season (days)")

effect_plot(mSos, pred = x2, interval = TRUE,
         main.title ="Temperature vs Expected SOS",x.label="Temperature (C)",
         y.label="Start of Season (days)")

effect_plot(mSos, pred = x3, interval = TRUE,
         main.title = "HSG vs Expected SOS", x.label ="HSG Rank",
         y.label="Start of Season (days)")

#MOS
effect_plot(mMos, pred = x1, interval = TRUE,
         main.title = "Precipitation vs Expected MOS",
         x.label ="Precipitation (mm)",
         y.label="Middle of Season (days)")

effect_plot(mMos, pred = x2, interval = TRUE,
         main.title ="Temperature vs Expected MOS",x.label="Temperature (C)",
         y.label="Middle of Season (days)")

effect_plot(mMos, pred = x3, interval = TRUE,
         main.title = "HSG vs Expected MOS", x.label ="HSG Rank",
         y.label="Middle of Season (days)")

#EOS
effect_plot(mEos, pred = x1, interval = TRUE,
         main.title = "Precipitation vs Expected EOS",
         x.label ="Precipitation (mm)",
         y.label="End of Season (days)")

effect_plot(mEos, pred = x2, interval = TRUE,
         main.title ="Temperature vs Expected EOS",x.label="Temperature (C)",
         y.label="End of Season (days)")
```

```r
effect_plot(mEos, pred = x3, interval = TRUE,
        main.title = "HSG vs Expected EOS", x.label ="HSG Rank",
        y.label="End of Season (days)")

#GSL
effect_plot(mGsl, pred = x1, interval = TRUE,
        main.title = "Precipitation vs Expected Season Length",
        x.label ="Precipitation (mm)",
        y.label="Season Length (days)")

effect_plot(mGsl, pred = x2, interval = TRUE,
        main.title ="Temperature vs Expected Season Length",
        x.label="Temperature (C)",
        y.label="Season Length (days)")

effect_plot(mGsl, pred = x3, interval = TRUE,
        main.title = "HSG vs Expected Season Length", x.label ="HSG Rank",
        y.label="Season Length (days)")

#AMP
effect_plot(mAmp, pred = x1, interval = TRUE,
        main.title = "Precipitation vs Expected Available Water Content",
        x.label ="Precipitation (mm)",
        y.label="Available Water Content (m3/m3)")

effect_plot(mAmp, pred = x2, interval = TRUE,
        main.title = "Temperature vs Expected Available Water Content",
        x.label ="Temperature (C)",
        y.label="Available Water Content (m3/m3)")

effect_plot(mAmp, pred = x3, interval = TRUE,
        main.title = "HSG vs Expected Available Water Content",
        x.label ="HSG Rank",
        y.label="Available Water Content (m3/m3)")

#BAS
effect_plot(mBas, pred = x1, interval = TRUE,
        main.title = "Precipitation vs Expected Base Value",
        x.label ="Precipitation (mm)",
        y.label="Base Value (m3/m3)")

effect_plot(mBas, pred = x2, interval = TRUE,
        main.title = "Temperature vs Expected Base Value",
        x.label ="Temperature (C)",
        y.label="Base Value (m3/m3)")

effect_plot(mBas, pred = x3, interval = TRUE,
        main.title = "HSG vs Expected Base Value", x.label ="HSG Rank",
        y.label="Base Value (m3/m3)")
```

```
#MXF
effect_plot(mMxf, pred = x1, interval = TRUE,
          main.title = "Precipitation vs Expected Maximum Value",
          x.label ="Precipitation (mm)",
          y.label="Maximum Value (m3/m3)")

effect_plot(mMxf, pred = x2, interval = TRUE,
          main.title = "Temperature vs Expected Maximum Value",
          x.label ="Temperature (C)",
          y.label="Maximum Value (m3/m3)")

effect_plot(mMxf, pred = x3, interval = TRUE,
          main.title = "HSG vs Expected Maximum Value", x.label ="HSG Rank",
          y.label="Maximum Value (m3/m3)")

#Checking model assumptions

#Semivariogram to check for spatial autocorrelation among residuals
sos_pts_df <- data.frame(sos_pts)
vg1 <- variogram(residuals.gam(mSos, type = "response") ~ 1, loc = ~x +
              y, data = sos_pts_df)
plot(vg1, main = "SOS Semivariogram")

mos_pts_df <- data.frame(mos_pts)
vg2 <- variogram(residuals.gam(mMos, type = "response") ~ 1, loc = ~x +
              y, data = mos_pts_df)
plot(vg2, main = "MOS Semivariogram")

eos_pts_df <- data.frame(eos_pts)
vg3 <- variogram(residuals.gam(mEos, type = "response") ~ 1, loc = ~x +
              y, data = eos_pts_df)
plot(vg3, main = "EOS Semivariogram")

gsl_pts_df <- data.frame(gsl_pts)
vg4 <- variogram(residuals.gam(mGsl, type = "response") ~ 1, loc = ~x +
              y, data = gsl_pts_df)
plot(vg4, main = "GSL Semivariogram")

amp_pts_df <- data.frame(amp_pts)
vg5 <- variogram(residuals.gam(mAmp, type = "response") ~ 1, loc = ~x +
              y, data = amp_pts_df)
plot(vg5, main = "AMP Semivariogram")

bas_pts_df <- data.frame(bas_pts)
vg6 <- variogram(residuals.gam(mBas, type = "response") ~ 1, loc = ~x +
              y, data = bas_pts_df)
plot(vg6, main = "BAS Semivariogram")
```

```
mxf_pts_df <- data.frame(mxf_pts)
vg7 <- variogram(residuals.gam(mMxf, type = "response") ~ 1, loc = ~x +
              y, data = mxf_pts_df)
plot(vg7, main = "MXF Semivariogram")

#Concurvity
concurvity(mSos)
concurvity(mMos)
concurvity(mEos)
concurvity(mGsl)
concurvity(mAmp)
concurvity(mBas)
concurvity(mMxf)

#check if residuals mean is zero
sos_pts_df$residuals <- residuals(mSos)
mean(sos_pts_df$residuals)
hist(sos_pts_df$residuals, freq = FALSE, xlab = "residuals", ylab =
     "Residual Frequency", main = "SOS Residuals")

mos_pts_df$residuals <- residuals(mMos)
mean(mos_pts_df$residuals)
hist(mos_pts_df$residuals, freq = FALSE, xlab = "residuals", ylab =
     "Residual Frequency", main = "MOS Residuals")

eos_pts_df$residuals <- residuals(mEos)
mean(eos_pts_df$residuals)
hist(eos_pts_df$residuals, freq = FALSE, xlab = "residuals", ylab =
     "Residual Frequency", main = "EOS Residuals")

gsl_pts_df$residuals <- residuals(mGsl)
mean(gsl_pts_df$residuals)
hist(gsl_pts_df$residuals, freq = FALSE, xlab = "residuals", ylab =
     "Residual Frequency", main = "GSL Residuals")

amp_pts_df$residuals <- residuals(mAmp)
mean(amp_pts_df$residuals)
hist(amp_pts_df$residuals, freq = FALSE, xlab = "residuals", ylab =
     "Residual Frequency", main = "AMP Residuals")

bas_pts_df$residuals <- residuals(mBas)
mean(bas_pts_df$residuals)
hist(bas_pts_df$residuals, freq = FALSE, xlab = "residuals", ylab =
     "Residual Frequency", main = "BAS Residuals")

mxf_pts_df$residuals <- residuals(mMxf)
mean(mxf_pts_df$residuals)
hist(mxf_pts_df$residuals, freq = FALSE, xlab = "residuals", ylab =
     "Residual Frequency", main = "MXF Residuals")
```

Appendix D - Results of the Ecoregion Statistical Analysis

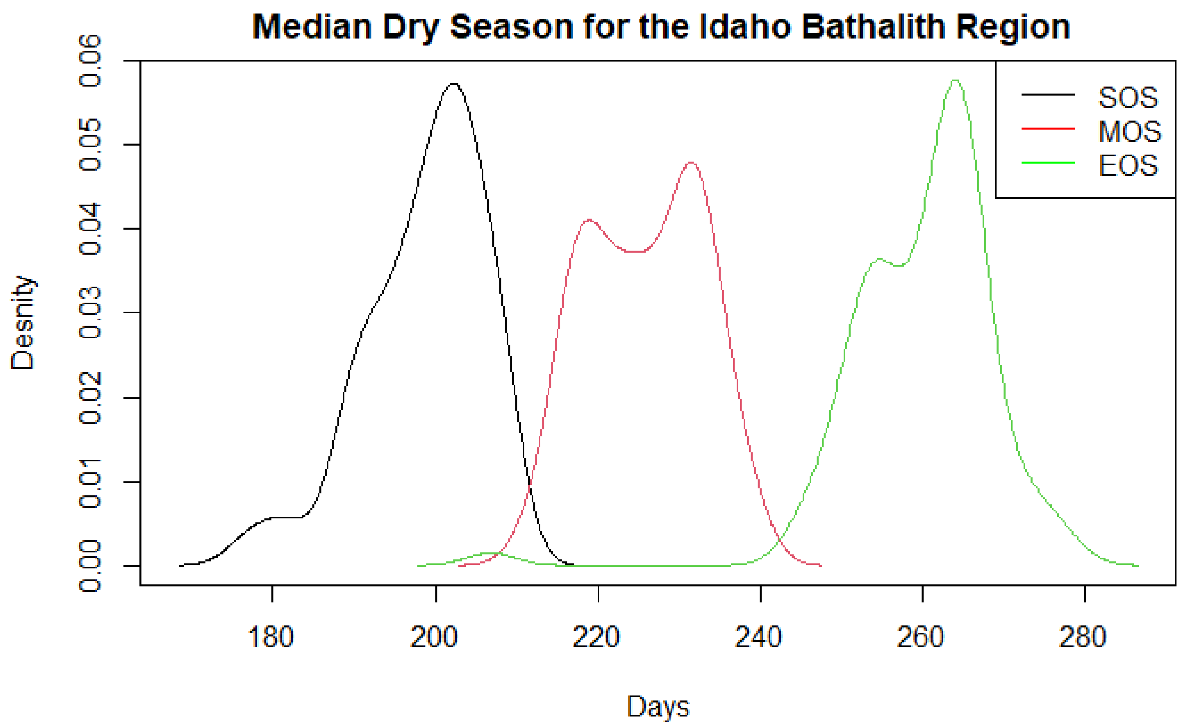

Figure D.6.36. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Idaho Bathalith ecoregion.

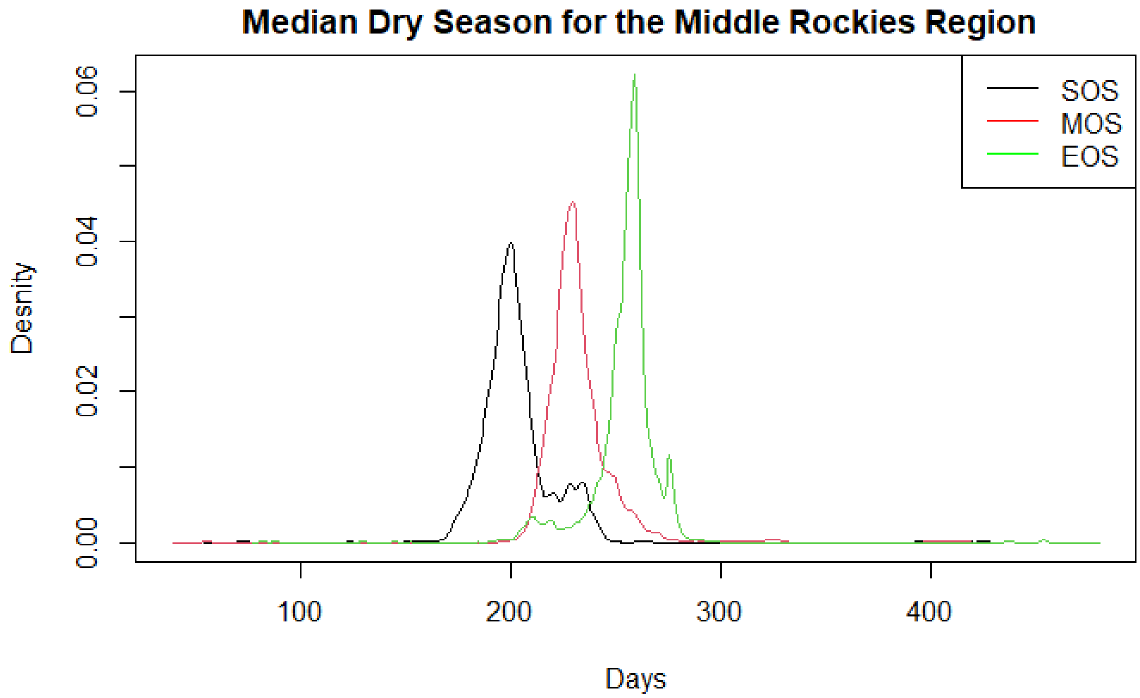

Figure D.6.37. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Middle Rockies ecoregion.

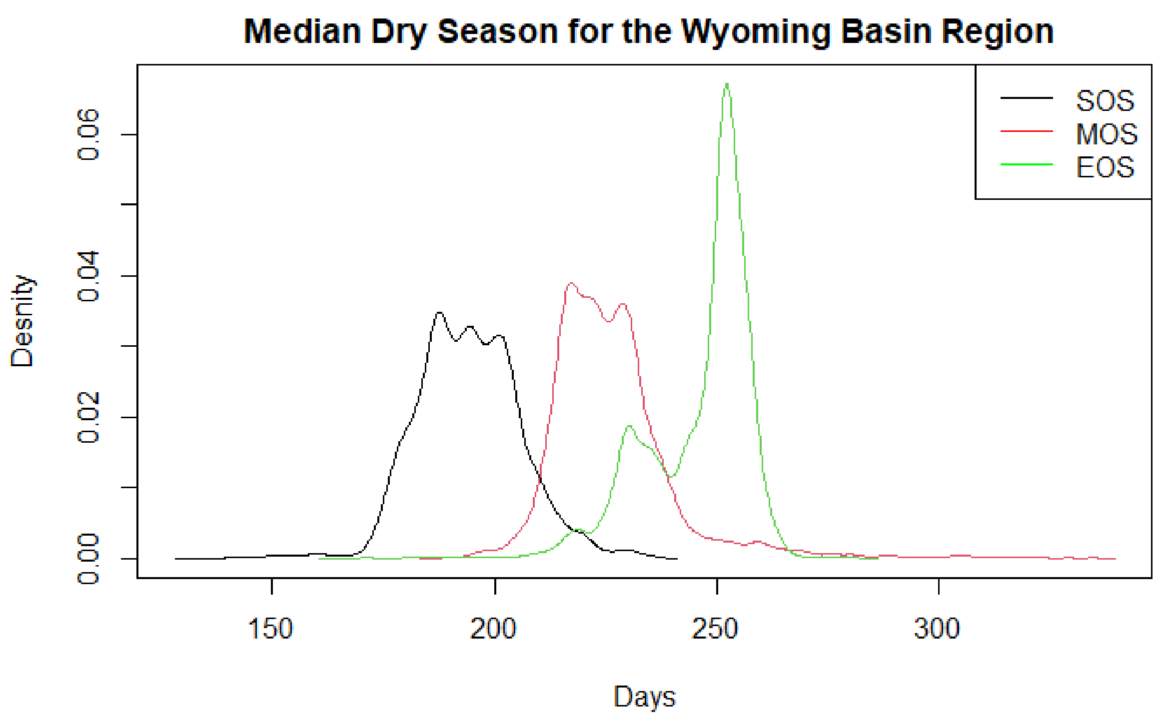

Figure D.6.38. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Wyoming Basin ecoregion.

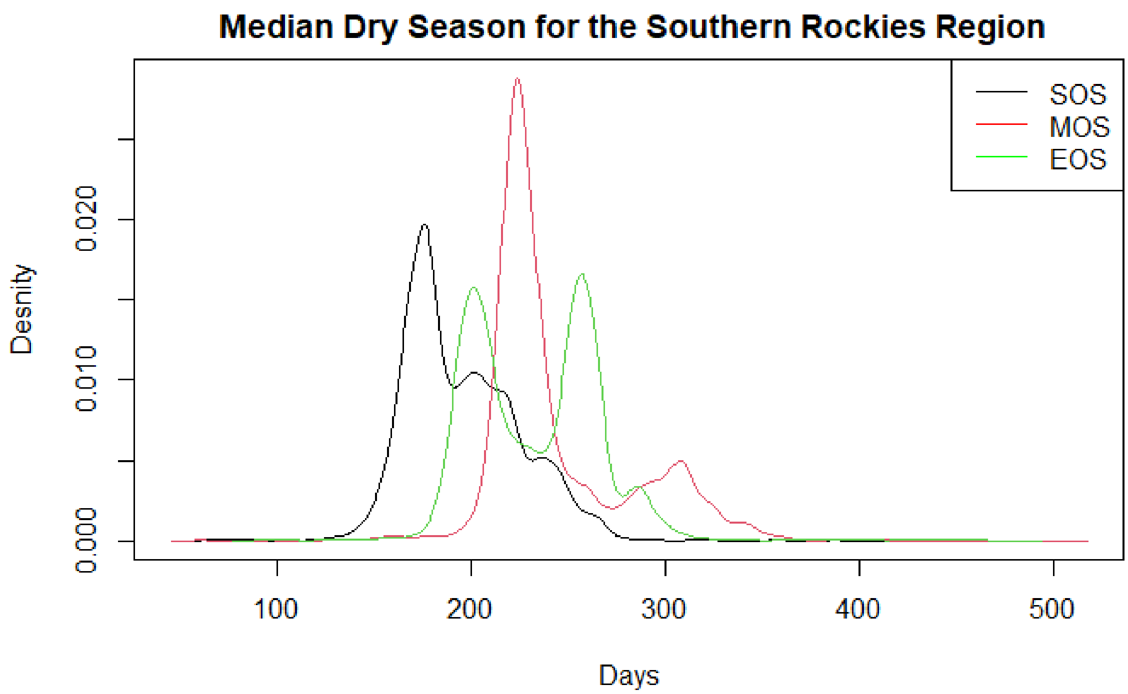

Figure D.6.39. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Southern Rockies ecoregion.

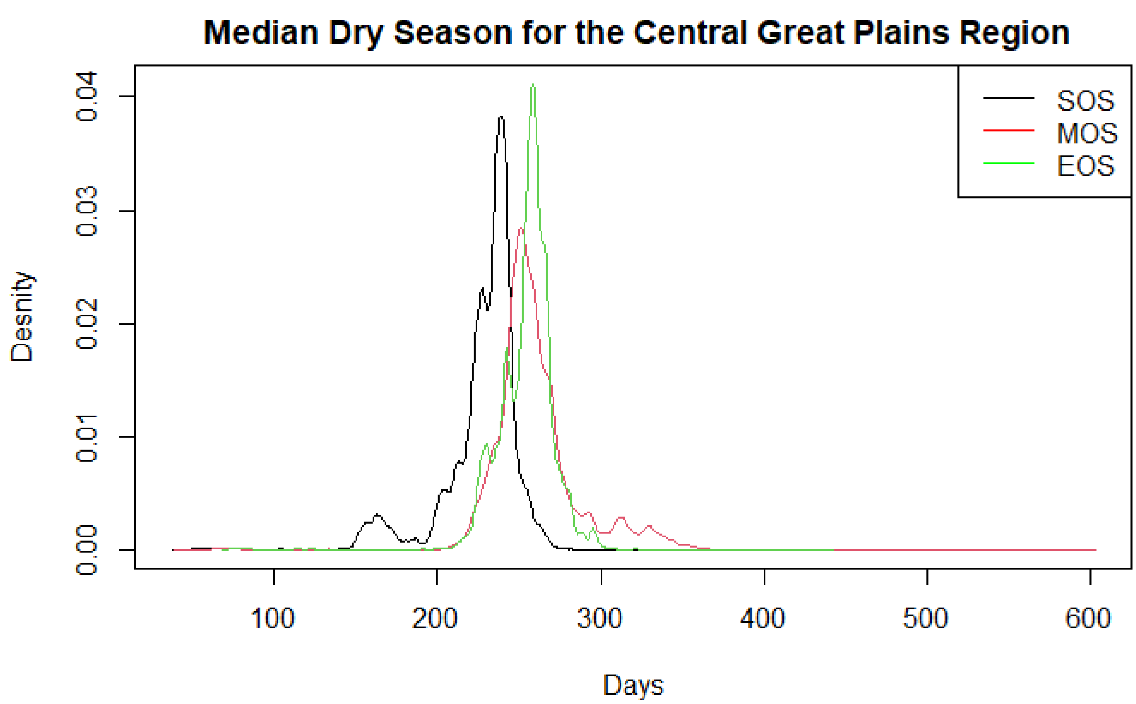

Figure D.6.40. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Central Great Plains ecoregion.

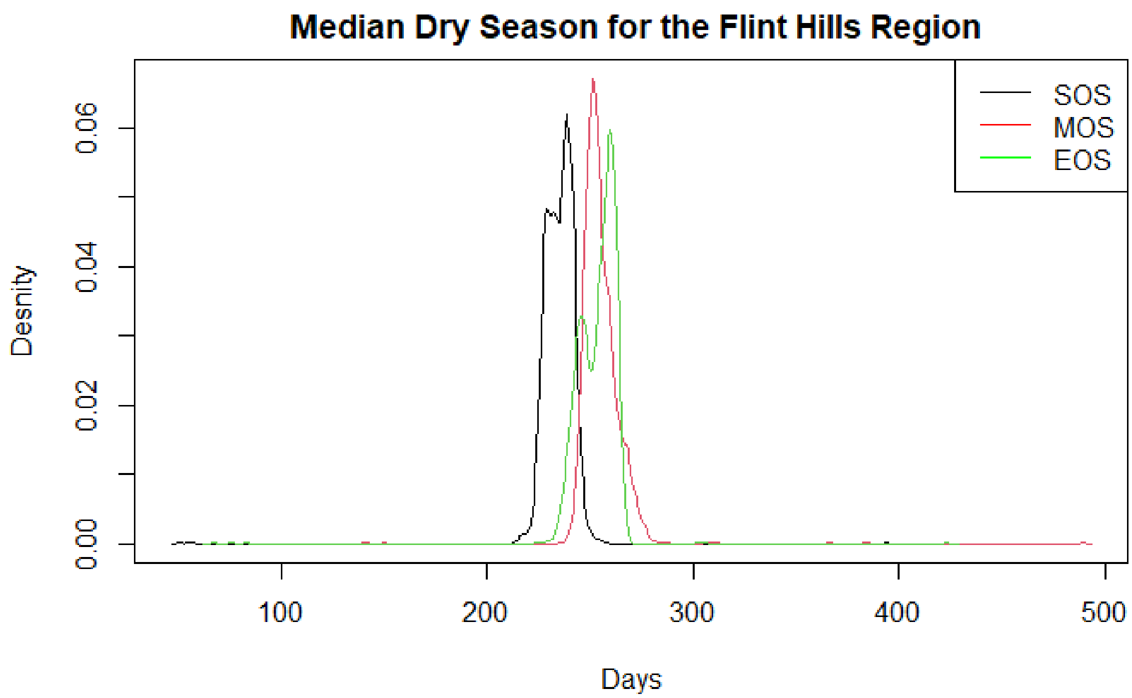

Figure D.6.41. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Flint Hills ecoregion.

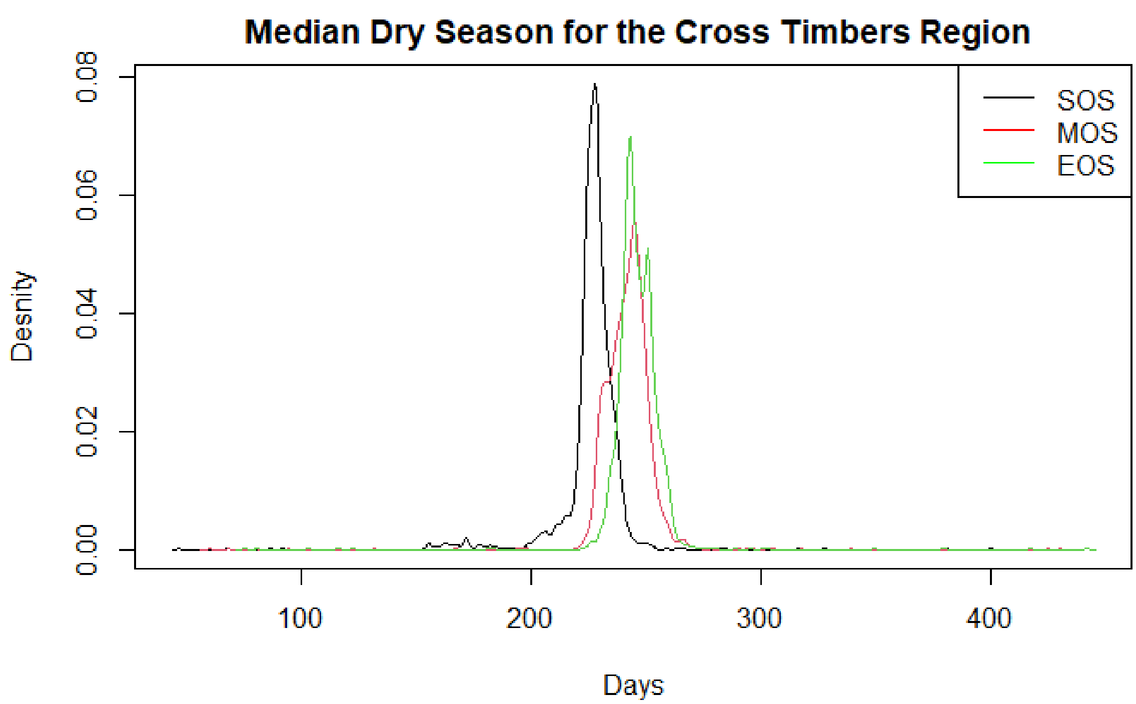

Figure D.6.42. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Cross Timbers ecoregion.

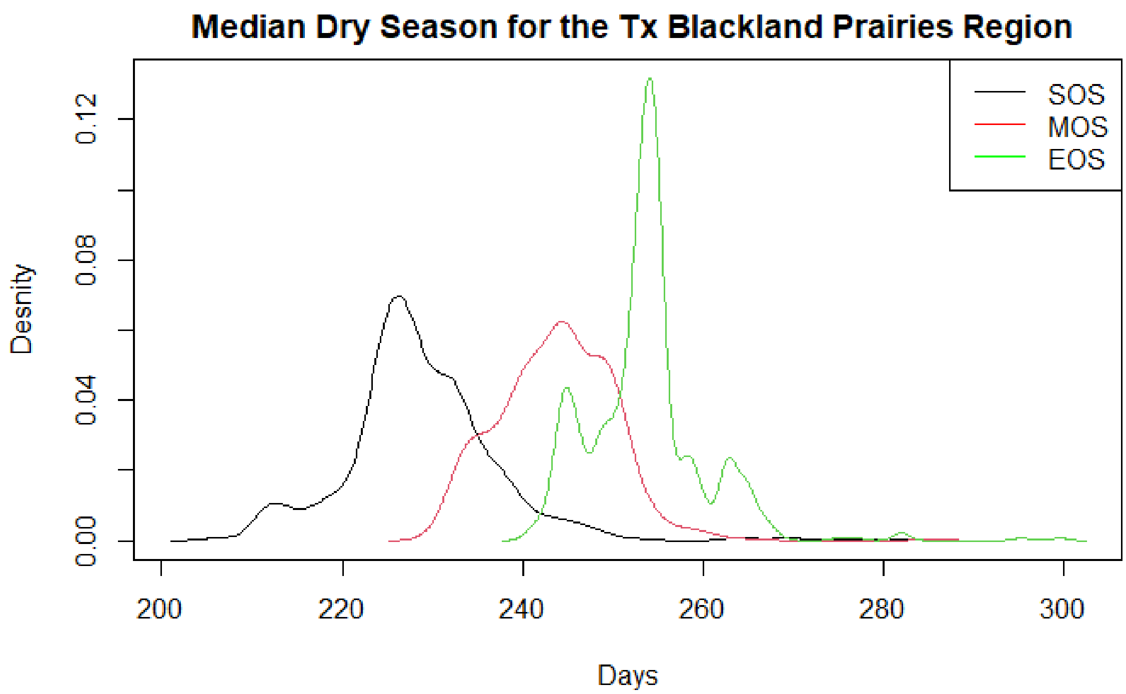

Figure D.6.43. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Texas Blackland Prairies ecoregion.

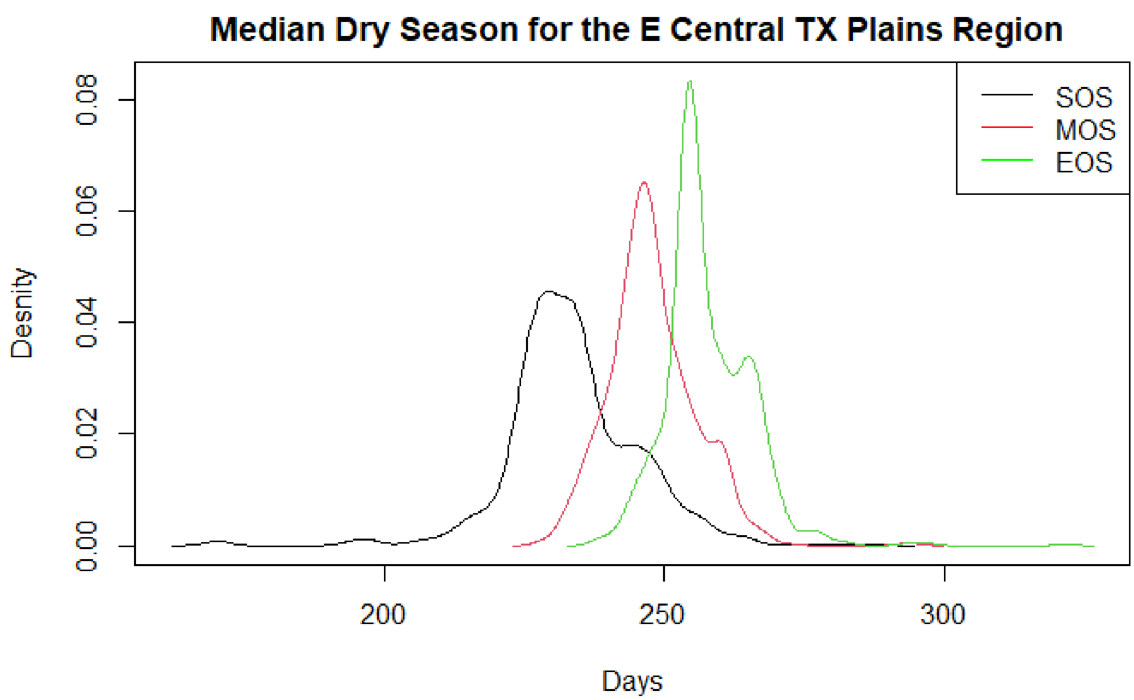

Figure D.6.44. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the East Central Texas Plains ecoregion.

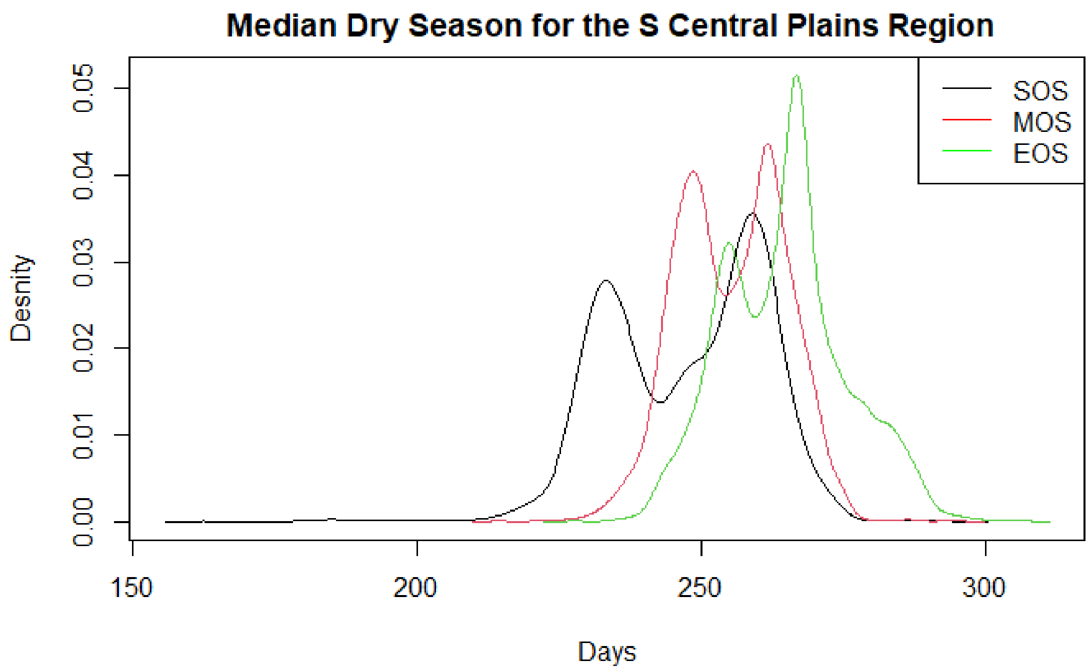

Figure D.6.45. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the South Central Plains ecoregion.

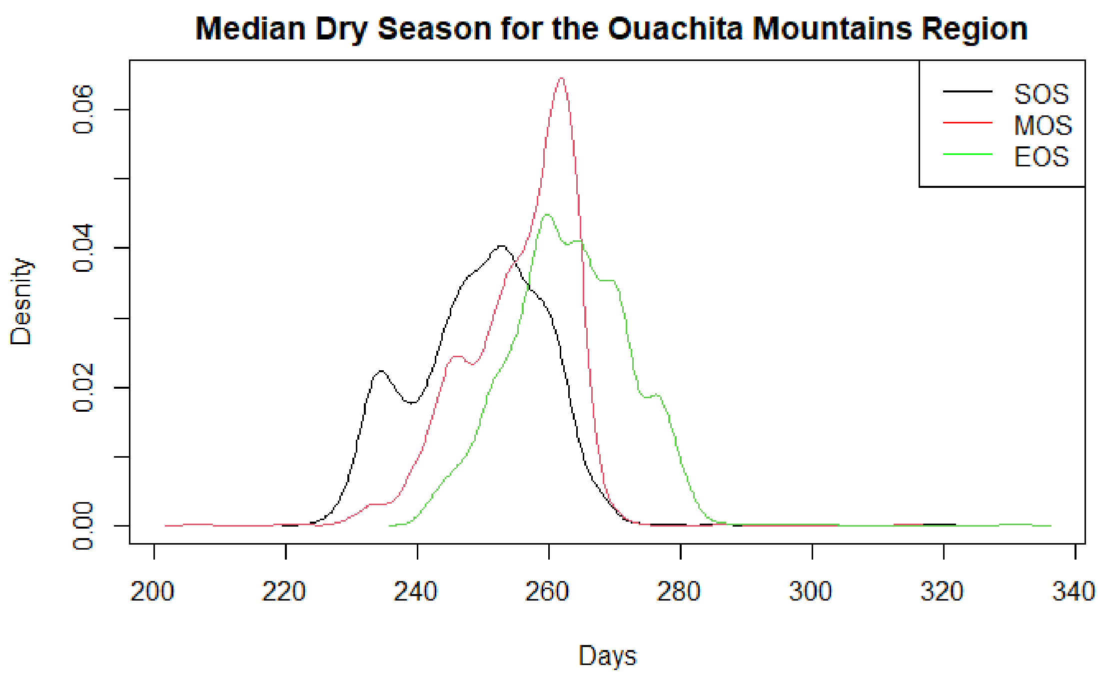

Figure D.6.46. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Ouachita Mountains ecoregion.

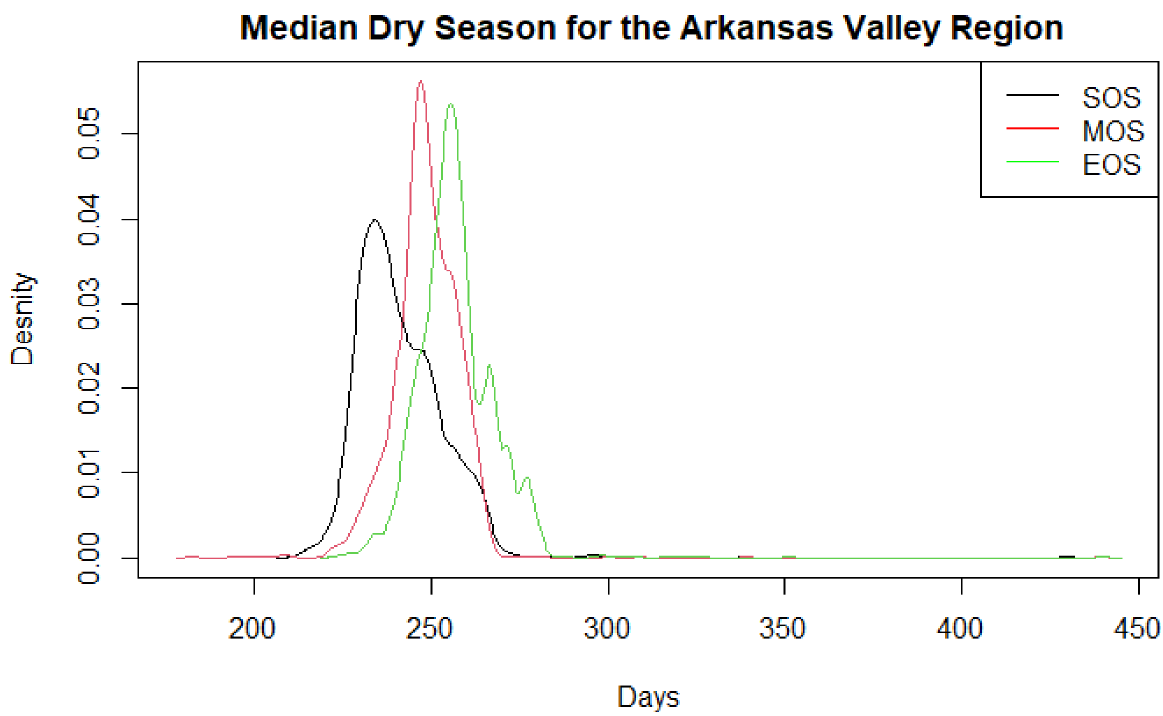

Figure D.6.47. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Arkansas Valley ecoregion.

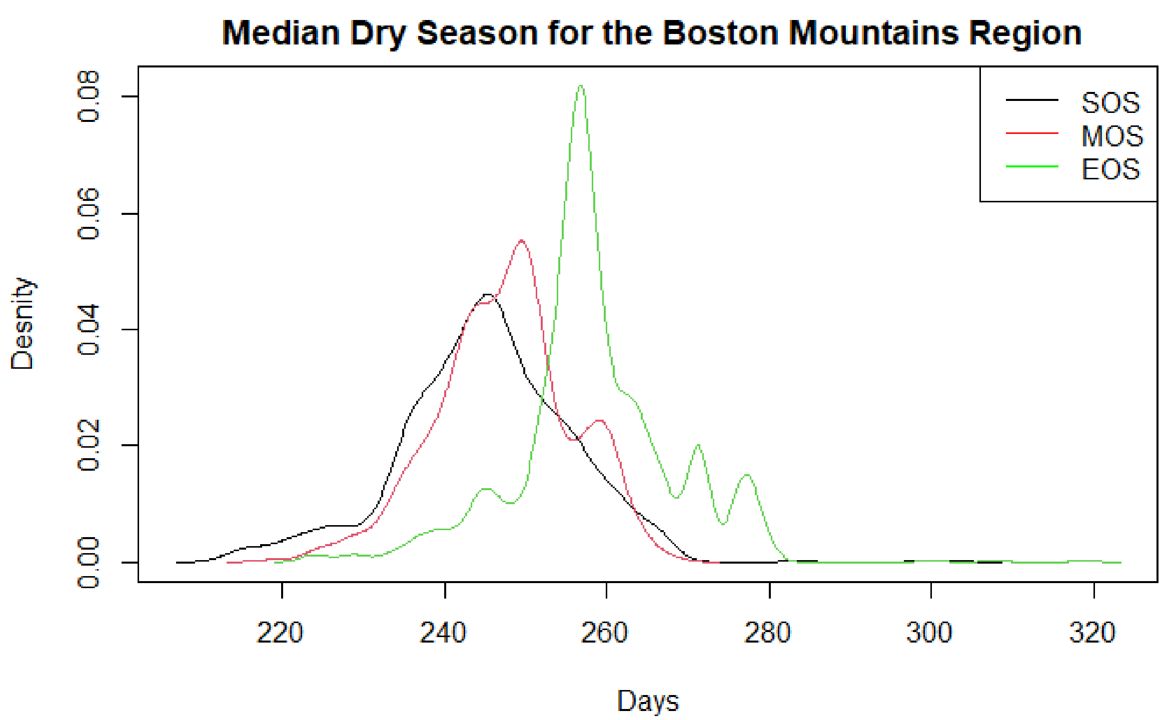

Figure D.6.48. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Boston Mountains ecoregion.

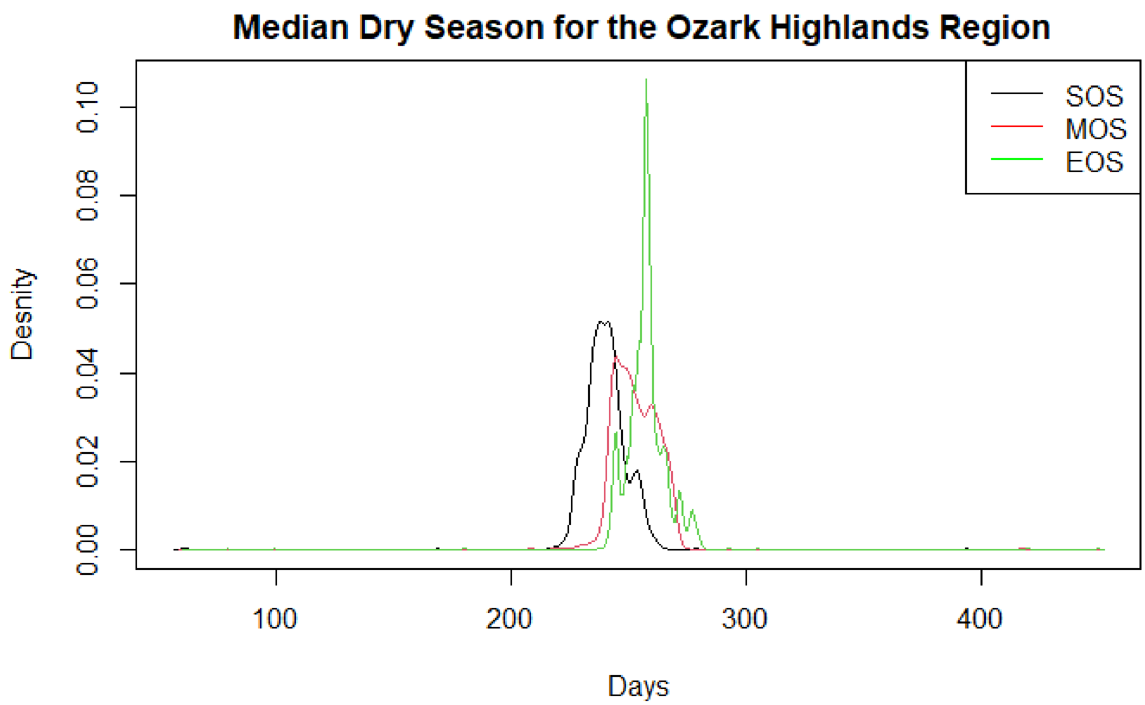

Figure d.6.49. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Ozark Highlands ecoregion.

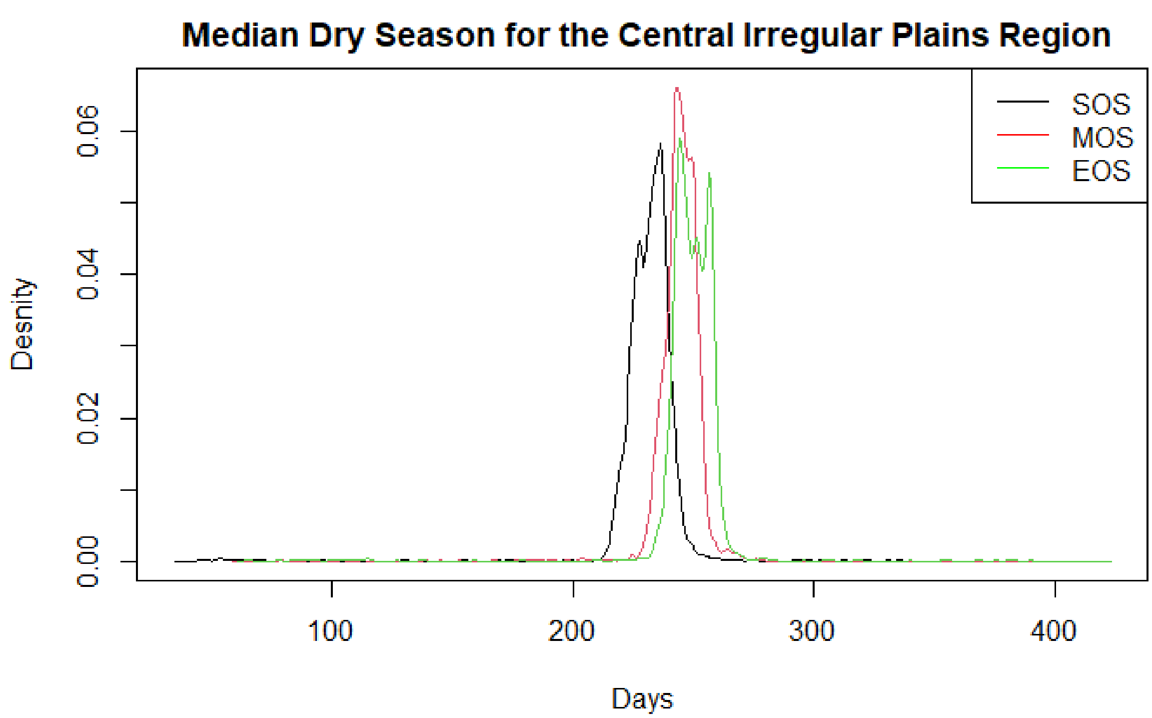

Figure D.6.50. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Central Irregular Plains ecoregion.

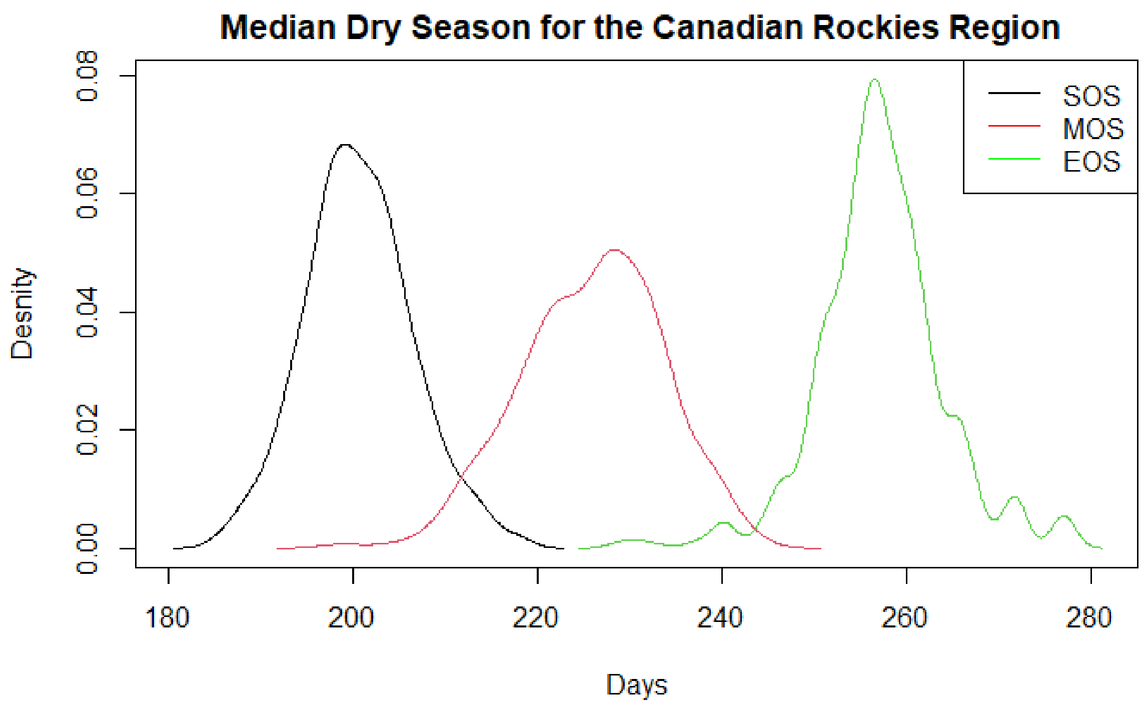

Figure D.6.51. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Canadian Rockies ecoregion.

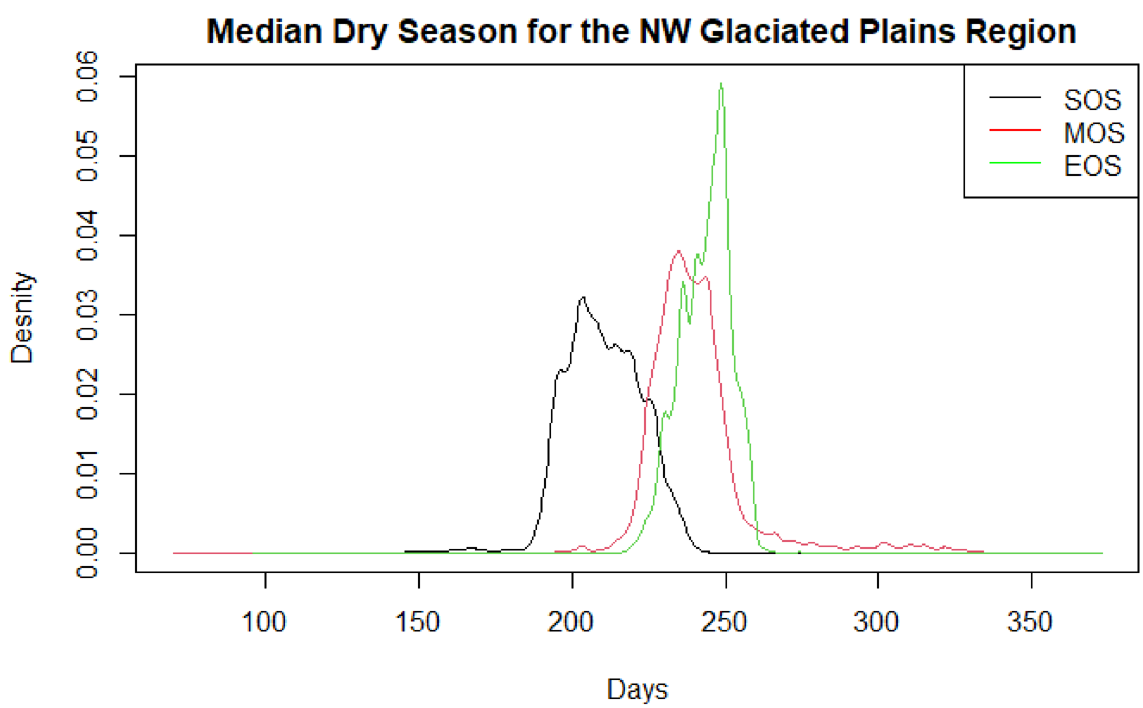

Figure D.6.52. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Northwest Glaciated Plains ecoregion.

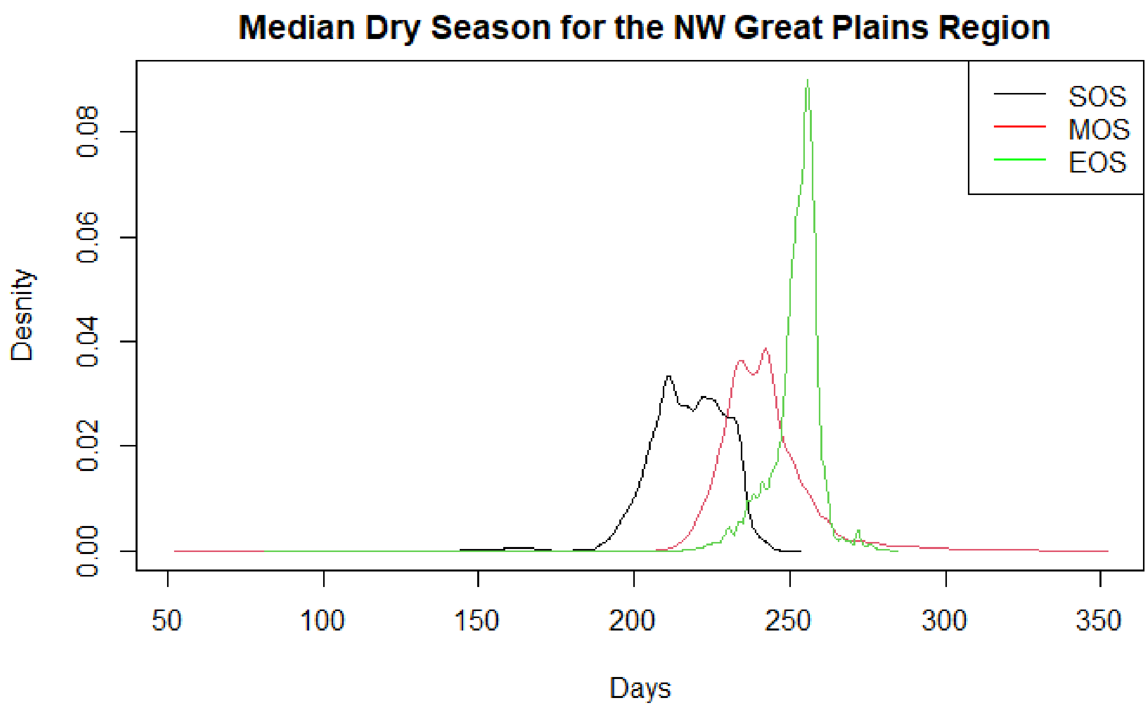

Figure D.6.53. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Northwest Great Plains ecoregion.

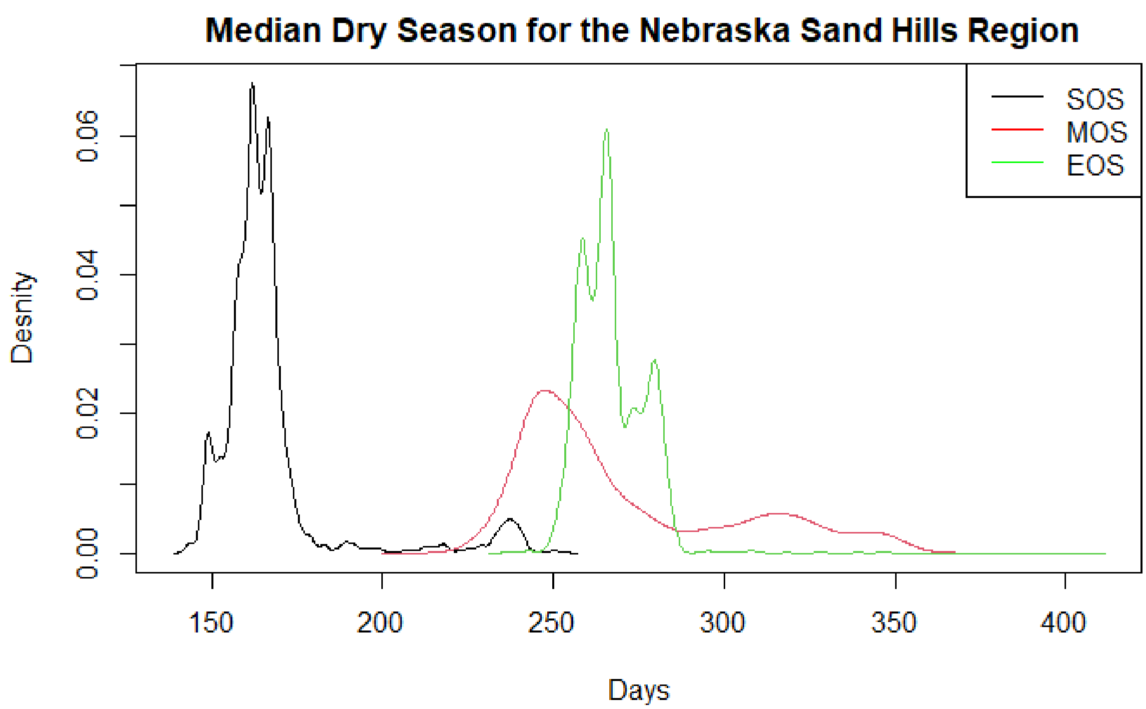

Figure D.6.54. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Nebraska Sand Hills ecoregion.

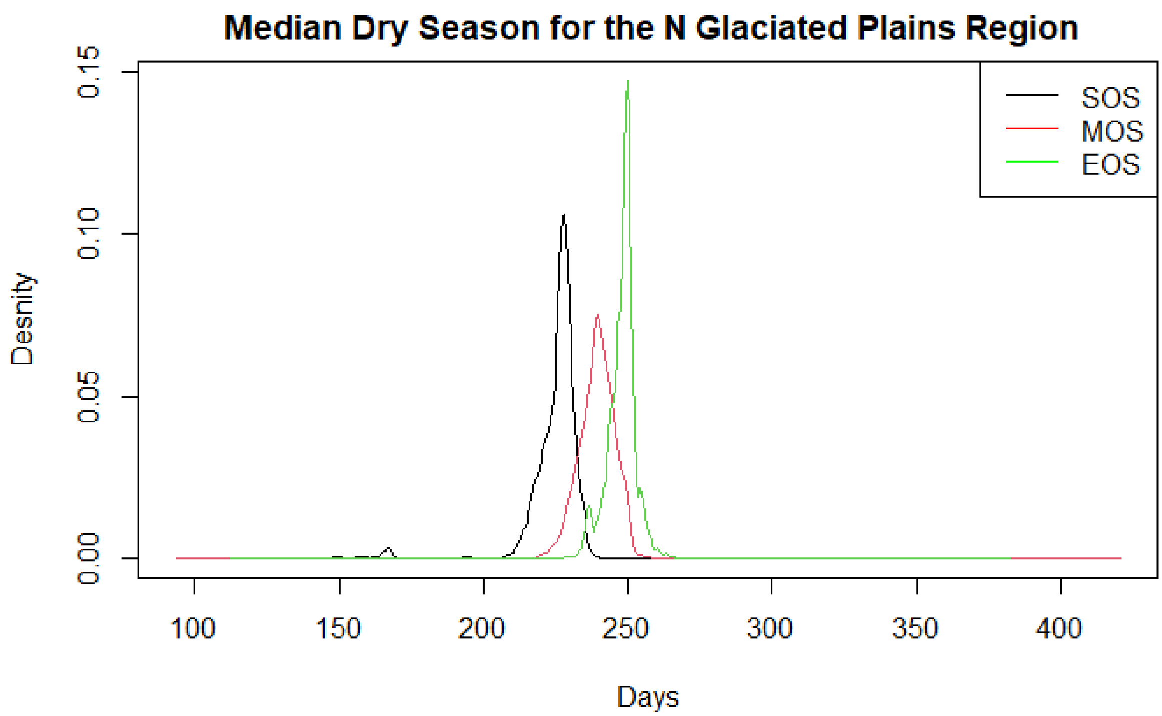

Figure D.6.55. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the North Glaciated Plains ecoregion.

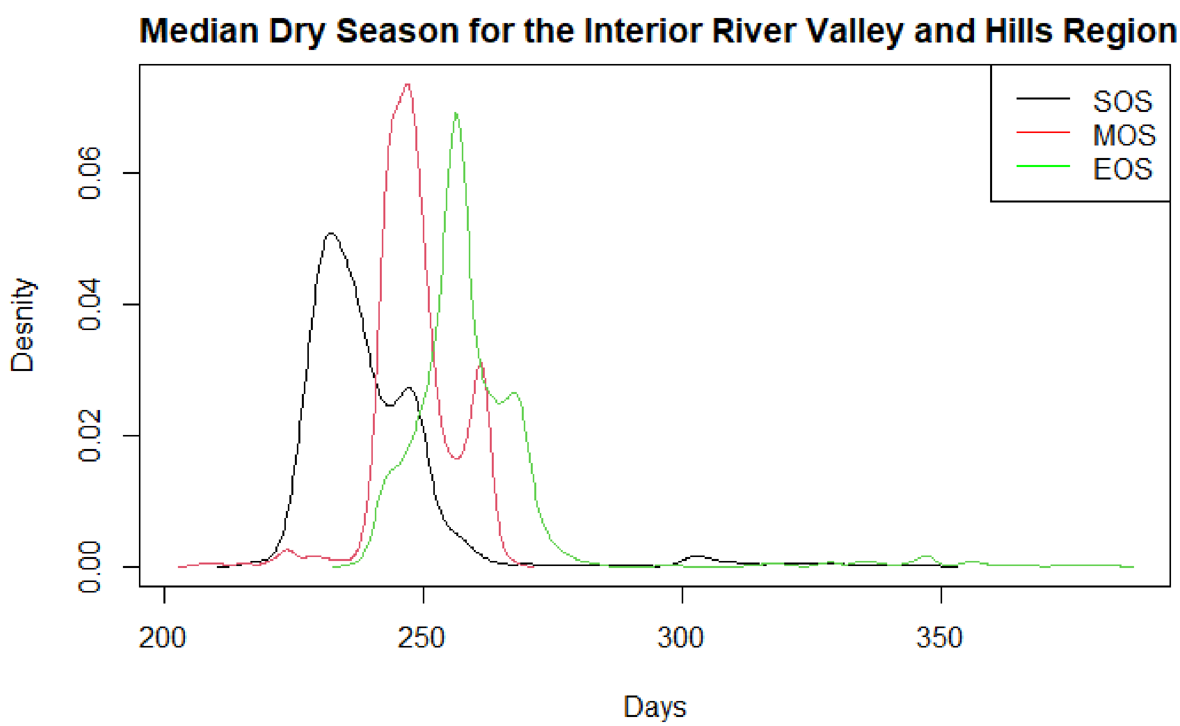

Figure D.6.56. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Interior River Valley and Hills ecoregion.

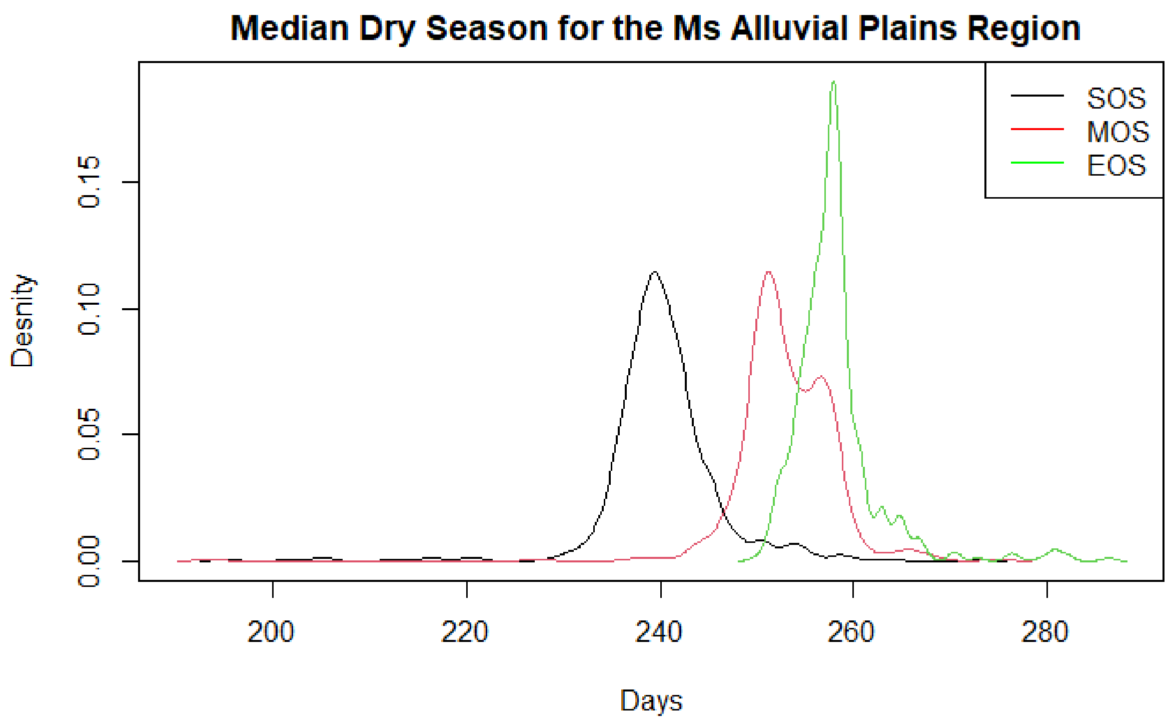

Figure D.6.57. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Mississippi Alluvial Plains ecoregion.

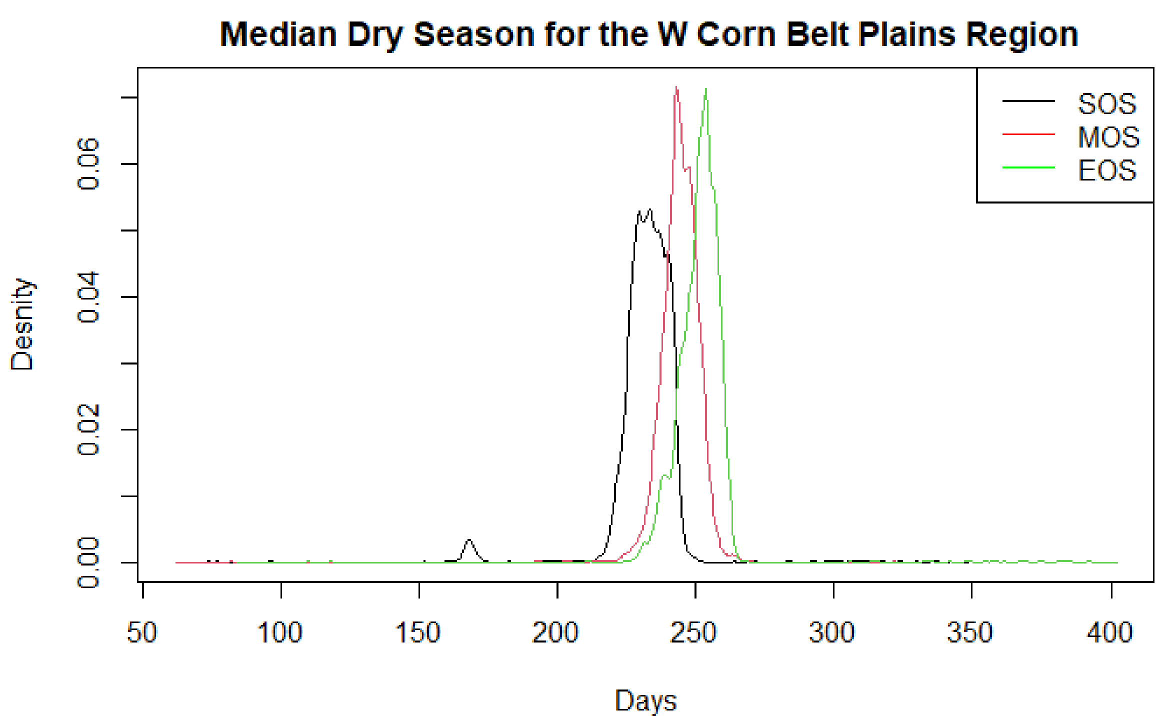

Figure D.6.58. A histogram plot showing the median soil season's start (SOS), middle (MOS), and end of season (EOS) for the Western Corn Belt Plains ecoregion.

Appendix E - Ecoregion Statistical Analysis R Code

Emily Nottingham

1/15/2021

```
#load libraries
library(sf)
library(dplyr)
library(rgdal)
library(mgcv)
library(maps)
library(maptools)
library(plotrix)
library(gstat)
library(ggplot2)
library(raster)
library(sm)
library(multcompView)

#load directories and subset dataframes

sos_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Median/
SOS_0130_MEDIAN.tif")
mos_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Median/
MOS_0130_MEDIAN.tif")
eos_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/Median/
EOS_0130_MEDIAN.tif")
Ecoregions <-raster(x="C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HUC_IN/test/eco_ras
ter_1.tif")

stack_eco <- stack(sos_median, mos_median, eos_median, Ecoregions)
median_eco <- as.data.frame(stack_eco)
median_eco <- na.omit(median_eco)

flint_hills <- subset(median_eco, median_eco$eco_raster_1 == 21)
flint_hills <- subset(flint_hills, select = -c(eco_raster_1))

Idaho_Bathalith <- subset(median_eco, median_eco$eco_raster_1 == 8)
Idaho_Bathalith <- subset(Idaho_Bathalith, select = -c(eco_raster_1))

Middle_Rockies <- subset(median_eco, median_eco$eco_raster_1 == 9)
Middle_Rockies <- subset(Middle_Rockies, select = -c(eco_raster_1))

Wy_basin <- subset(median_eco, median_eco$eco_raster_1 == 10)
Wy_basin <- subset(Wy_basin, select = -c(eco_raster_1))

S_Rockies <- subset(median_eco, median_eco$eco_raster_1 == 14)
S_Rockies <- subset(S_Rockies, select = -c(eco_raster_1))

High_Plains <- subset(median_eco, median_eco$eco_raster_1 == 18)
```

```r
High_Plains <- subset(High_Plains, select = -c(eco_raster_1))

SW_Tablelands <- subset(median_eco, median_eco$eco_raster_1 == 19)
SW_Tablelands <- subset(SW_Tablelands, select = -c(eco_raster_1))

Central_Great_Plains <- subset(median_eco, median_eco$eco_raster_1 == 20)
Central_Great_Plains <- subset(Central_Great_Plains, select = -c(eco_raster_1))

Cross_Timbers <- subset(median_eco, median_eco$eco_raster_1 == 22)
Cross_Timbers <- subset(Cross_Timbers, select = -c(eco_raster_1))

Tx_Blacklands_Prairies <- subset(median_eco, median_eco$eco_raster_1 == 26)
Tx_Blacklands_Prairies <- subset(Tx_Blacklands_Prairies,
                                 select = -c(eco_raster_1))

EC_Tx_Plains <- subset(median_eco, median_eco$eco_raster_1 == 27)
EC_Tx_Plains <- subset(EC_Tx_Plains, select = -c(eco_raster_1))

SC_Plains <- subset(median_eco, median_eco$eco_raster_1 == 29)
SC_Plains <- subset(SC_Plains, select = -c(eco_raster_1))

Ouachita_Mountains <- subset(median_eco, median_eco$eco_raster_1 == 30)
Ouachita_Mountains <- subset(Ouachita_Mountains, select = -c(eco_raster_1))

Ak_Valley <- subset(median_eco, median_eco$eco_raster_1 == 31)
Ak_Valley <- subset(Ak_Valley, select = -c(eco_raster_1))

Boston_Mountains <- subset(median_eco, median_eco$eco_raster_1 == 32)
Boston_Mountains <- subset(Boston_Mountains, select = -c(eco_raster_1))

Ozark_highlands <- subset(median_eco, median_eco$eco_raster_1 == 33)
Ozark_highlands <- subset(Ozark_highlands, select = -c(eco_raster_1))

C_Irregular_Plains <- subset(median_eco, median_eco$eco_raster_1 == 35)
C_Irregular_Plains <- subset(C_Irregular_Plains, select = -c(eco_raster_1))

Can_Rockies <- subset(median_eco, median_eco$eco_raster_1 == 36)
Can_Rockies <- subset(Can_Rockies, select = -c(eco_raster_1))

NW_Glaciated_Plains <- subset(median_eco, median_eco$eco_raster_1 == 37)
NW_Glaciated_Plains <- subset(NW_Glaciated_Plains, select = -c(eco_raster_1))

NW_Great_Plains <- subset(median_eco, median_eco$eco_raster_1 == 38)
NW_Great_Plains <- subset(NW_Great_Plains, select = -c(eco_raster_1))

Ne_Sand_Hills <- subset(median_eco, median_eco$eco_raster_1 == 39)
Ne_Sand_Hills <- subset(Ne_Sand_Hills, select = -c(eco_raster_1))

N_Glaciated_Plains <- subset(median_eco, median_eco$eco_raster_1 == 41)
N_Glaciated_Plains <- subset(N_Glaciated_Plains, select = -c(eco_raster_1))
```

```r
W_Corn_Belt_Plains <- subset(median_eco, median_eco$eco_raster_1 == 42)
W_Corn_Belt_Plains <- subset(W_Corn_Belt_Plains, select = -c(eco_raster_1))

Int_River_Valleys_Hills <- subset(median_eco, median_eco$eco_raster_1 == 70)
Int_River_Valleys_Hills <- subset(Int_River_Valleys_Hills,
                                  select = -c(eco_raster_1))

Ms_Alluvial_Plains <- subset(median_eco, median_eco$eco_raster_1 == 71)
Ms_Alluvial_Plains <- subset(Ms_Alluvial_Plains, select = -c(eco_raster_1))

#Retrieve median values
sapply(Idaho_Bathalith, median, na.rm=TRUE)
sapply(Middle_Rockies, median, na.rm=TRUE)
sapply(Wy_basin, median, na.rm=TRUE)
sapply(S_Rockies, median, na.rm=TRUE)
sapply(High_Plains, median, na.rm=TRUE)
sapply(SW_Tablelands, median, na.rm=TRUE)
sapply(Central_Great_Plains, median, na.rm=TRUE)
sapply(flint_hills, median, na.rm=TRUE)
sapply(Cross_Timbers, median, na.rm=TRUE)
sapply(Tx_Blacklands_Prairies, median, na.rm=TRUE)
sapply(EC_Tx_Plains, median, na.rm=TRUE)
sapply(SC_Plains, median, na.rm=TRUE)
sapply(Ouachita_Mountains, median, na.rm=TRUE)
sapply(Ak_Valley, median, na.rm=TRUE)
sapply(Boston_Mountains, median, na.rm=TRUE)
sapply(Ozark_highlands, median, na.rm=TRUE)
sapply(C_Irregular_Plains, median, na.rm=TRUE)
sapply(Can_Rockies, median, na.rm=TRUE)
sapply(NW_Glaciated_Plains, median, na.rm=TRUE)
sapply(NW_Great_Plains, median, na.rm=TRUE)
sapply(Ne_Sand_Hills, median, na.rm=TRUE)
sapply(N_Glaciated_Plains, median, na.rm=TRUE)
sapply(Int_River_Valleys_Hills, median, na.rm=TRUE)
sapply(Ms_Alluvial_Plains, median, na.rm=TRUE)
sapply(W_Corn_Belt_Plains, median, na.rm=TRUE)

#create Density Plots

den <- apply(Idaho_Bathalith,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Idaho
    Bathalith Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))

legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Middle_Rockies,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Middle
    Rockies Region",xlim=range(sapply(den, "[", "x")),
```

```
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Wy_basin,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
    Wyoming Basin Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(S_Rockies,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
    Southern Rockies Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(High_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the High
    Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(SW_Tablelands,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the SW
    Tablelands Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Central_Great_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
    Central Great Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(flint_hills,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Flint
    Hills Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))
```

```
den <- apply(Cross_Timbers,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Cross
    Timbers Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Tx_Blacklands_Prairies,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Tx
    Blackland Prairies Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(EC_Tx_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the E
    Central TX Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(SC_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the S
    Central Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Ouachita_Mountains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
    Ouachita Mountains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Ak_Valley,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
    Arkansas Valley Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Boston_Mountains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Boston
    Mountains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
```

```
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Ozark_highlands,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Ozark
    Highlands Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(C_Irregular_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
    Central Irregular Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Can_Rockies,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
    Canadian Rockies Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(NW_Glaciated_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the NW
    Glaciated Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(NW_Great_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the NW
    Great Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Ne_Sand_Hills,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
    Nebraska Sand Hills Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))
```

```
den <- apply(N_Glaciated_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the N
    Glaciated Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Int_River_Valleys_Hills,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
    Interior River Valley and Hills Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(Ms_Alluvial_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Ms
    Alluvial Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

den <- apply(W_Corn_Belt_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the W
    Corn Belt Plains Region",xlim=range(sapply(den, "[", "x")),
    ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
      col = c("black","red", "green"))

#Kruskal Wallis test for ecoregions

S_eco <- pairwise.wilcox.test(median_eco$SOS_0130_MEDIAN,
                        median_eco$eco_raster_1, p.adjust.method = "BH")
M_eco <- pairwise.wilcox.test(median_eco$MOS_0130_MEDIAN,
                        median_eco$eco_raster_1, p.adjust.method = "BH")
E_eco <- pairwise.wilcox.test(median_eco$EOS_0130_MEDIAN,
                        median_eco$eco_raster_1, p.adjust.method = "BH")

S_eco
M_eco
E_eco

S_eco_let <-multcompLetters(S_eco$p.value)
M_eco_let <-multcompLetters(M_eco$p.value)
E_eco_let <-multcompLetters(E_eco$p.value)

S_eco_let
M_eco_let
E_eco_let
```

Appendix F - Hydrologic Soil Group Statistical Analysis R Code

Emily Nottingham

1/15/2021

```
#load libraries
library(sf)
library(dplyr)
library(rgdal)
library(mgcv)
library(maps)
library(maptools)
library(plotrix)
library(gstat)
library(ggplot2)
library(raster)
library(sm)
library(multcompView)

#load directories and subset dataframes

sos_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG
_HUC_IN/Median/SOS_0130_MEDIAN.tif")
mos_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG
_HUC_IN/Median/MOS_0130_MEDIAN.tif")
eos_median <- raster(x = "C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG
_HUC_IN/Median/EOS_0130_MEDIAN.tif")
Ecoregions <-raster(x="C:/Users/emiro/Stat764 Dropbox/Emily Nottingham/HSG_HU
C_IN/test/eco_raster_1.tif")

stack_eco <- stack(sos_median, mos_median, eos_median, Ecoregions)
median_eco <- as.data.frame(stack_eco)
median_eco <- na.omit(median_eco)

flint_hills <- subset(median_eco, median_eco$eco_raster_1 == 21)
flint_hills <- subset(flint_hills, select = -c(eco_raster_1))

Idaho_Bathalith <- subset(median_eco, median_eco$eco_raster_1 == 8)
Idaho_Bathalith <- subset(Idaho_Bathalith, select = -c(eco_raster_1))

Middle_Rockies <- subset(median_eco, median_eco$eco_raster_1 == 9)
Middle_Rockies <- subset(Middle_Rockies, select = -c(eco_raster_1))

Wy_basin <- subset(median_eco, median_eco$eco_raster_1 == 10)
Wy_basin <- subset(Wy_basin, select = -c(eco_raster_1))

S_Rockies <- subset(median_eco, median_eco$eco_raster_1 == 14)
S_Rockies <- subset(S_Rockies, select = -c(eco_raster_1))
```

```r
High_Plains <- subset(median_eco, median_eco$eco_raster_1 == 18)
High_Plains <- subset(High_Plains, select = -c(eco_raster_1))

SW_Tablelands <- subset(median_eco, median_eco$eco_raster_1 == 19)
SW_Tablelands <- subset(SW_Tablelands, select = -c(eco_raster_1))

Central_Great_Plains <- subset(median_eco, median_eco$eco_raster_1 == 20)
Central_Great_Plains <- subset(Central_Great_Plains, select = -c(eco_raster_1
))

Cross_Timbers <- subset(median_eco, median_eco$eco_raster_1 == 22)
Cross_Timbers <- subset(Cross_Timbers, select = -c(eco_raster_1))

Tx_Blacklands_Prairies <- subset(median_eco, median_eco$eco_raster_1 == 26)
Tx_Blacklands_Prairies <- subset(Tx_Blacklands_Prairies,
                                 select = -c(eco_raster_1))

EC_Tx_Plains <- subset(median_eco, median_eco$eco_raster_1 == 27)
EC_Tx_Plains <- subset(EC_Tx_Plains, select = -c(eco_raster_1))

SC_Plains <- subset(median_eco, median_eco$eco_raster_1 == 29)
SC_Plains <- subset(SC_Plains, select = -c(eco_raster_1))

Ouachita_Mountains <- subset(median_eco, median_eco$eco_raster_1 == 30)
Ouachita_Mountains <- subset(Ouachita_Mountains, select = -c(eco_raster_1))

Ak_Valley <- subset(median_eco, median_eco$eco_raster_1 == 31)
Ak_Valley <- subset(Ak_Valley, select = -c(eco_raster_1))

Boston_Mountains <- subset(median_eco, median_eco$eco_raster_1 == 32)
Boston_Mountains <- subset(Boston_Mountains, select = -c(eco_raster_1))

Ozark_highlands <- subset(median_eco, median_eco$eco_raster_1 == 33)
Ozark_highlands <- subset(Ozark_highlands, select = -c(eco_raster_1))

C_Irregular_Plains <- subset(median_eco, median_eco$eco_raster_1 == 35)
C_Irregular_Plains <- subset(C_Irregular_Plains, select = -c(eco_raster_1))

Can_Rockies <- subset(median_eco, median_eco$eco_raster_1 == 36)
Can_Rockies <- subset(Can_Rockies, select = -c(eco_raster_1))

NW_Glaciated_Plains <- subset(median_eco, median_eco$eco_raster_1 == 37)
NW_Glaciated_Plains <- subset(NW_Glaciated_Plains, select = -c(eco_raster_1))

NW_Great_Plains <- subset(median_eco, median_eco$eco_raster_1 == 38)
NW_Great_Plains <- subset(NW_Great_Plains, select = -c(eco_raster_1))

Ne_Sand_Hills <- subset(median_eco, median_eco$eco_raster_1 == 39)
Ne_Sand_Hills <- subset(Ne_Sand_Hills, select = -c(eco_raster_1))
```

```r
N_Glaciated_Plains <- subset(median_eco, median_eco$eco_raster_1 == 41)
N_Glaciated_Plains <- subset(N_Glaciated_Plains, select = -c(eco_raster_1))

W_Corn_Belt_Plains <- subset(median_eco, median_eco$eco_raster_1 == 42)
W_Corn_Belt_Plains <- subset(W_Corn_Belt_Plains, select = -c(eco_raster_1))

Int_River_Valleys_Hills <- subset(median_eco, median_eco$eco_raster_1 == 70)
Int_River_Valleys_Hills <- subset(Int_River_Valleys_Hills,
                                  select = -c(eco_raster_1))

Ms_Alluvial_Plains <- subset(median_eco, median_eco$eco_raster_1 == 71)
Ms_Alluvial_Plains <- subset(Ms_Alluvial_Plains, select = -c(eco_raster_1))

#Retrieve median values
sapply(Idaho_Bathalith, median, na.rm=TRUE)
sapply(Middle_Rockies, median, na.rm=TRUE)
sapply(Wy_basin, median, na.rm=TRUE)
sapply(S_Rockies, median, na.rm=TRUE)
sapply(High_Plains, median, na.rm=TRUE)
sapply(SW_Tablelands, median, na.rm=TRUE)
sapply(Central_Great_Plains, median, na.rm=TRUE)
sapply(flint_hills, median, na.rm=TRUE)
sapply(Cross_Timbers, median, na.rm=TRUE)
sapply(Tx_Blacklands_Prairies, median, na.rm=TRUE)
sapply(EC_Tx_Plains, median, na.rm=TRUE)
sapply(SC_Plains, median, na.rm=TRUE)
sapply(Ouachita_Mountains, median, na.rm=TRUE)
sapply(Ak_Valley, median, na.rm=TRUE)
sapply(Boston_Mountains, median, na.rm=TRUE)
sapply(Ozark_highlands, median, na.rm=TRUE)
sapply(C_Irregular_Plains, median, na.rm=TRUE)
sapply(Can_Rockies, median, na.rm=TRUE)
sapply(NW_Glaciated_Plains, median, na.rm=TRUE)
sapply(NW_Great_Plains, median, na.rm=TRUE)
sapply(Ne_Sand_Hills, median, na.rm=TRUE)
sapply(N_Glaciated_Plains, median, na.rm=TRUE)
sapply(Int_River_Valleys_Hills, median, na.rm=TRUE)
sapply(Ms_Alluvial_Plains, median, na.rm=TRUE)
sapply(W_Corn_Belt_Plains, median, na.rm=TRUE)

#create Density Plots

den <- apply(Idaho_Bathalith,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Ida
ho
     Bathalith Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
```

```
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Middle_Rockies,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Mid
dle
     Rockies Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Wy_basin,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
     Wyoming Basin Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(S_Rockies,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
     Southern Rockies Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(High_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Hig
h
     Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(SW_Tablelands,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the SW
     Tablelands Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Central_Great_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
     Central Great Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
```

```r
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(flint_hills,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Fli
nt
     Hills Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Cross_Timbers,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Cro
ss
     Timbers Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Tx_Blacklands_Prairies,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Tx
     Blackland Prairies Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(EC_Tx_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the E
     Central TX Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(SC_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the S
     Central Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Ouachita_Mountains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
     Ouachita Mountains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
```

```
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Ak_Valley,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
     Arkansas Valley Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Boston_Mountains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Bos
ton
     Mountains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Ozark_highlands,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Oza
rk
     Highlands Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(C_Irregular_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
     Central Irregular Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Can_Rockies,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
     Canadian Rockies Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(NW_Glaciated_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the NW
     Glaciated Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
```

```r
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(NW_Great_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the NW
     Great Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Ne_Sand_Hills,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
     Nebraska Sand Hills Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(N_Glaciated_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the N
     Glaciated Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Int_River_Valleys_Hills,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the
     Interior River Valley and Hills Region",xlim=range(sapply(den, "[", "x")
),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(Ms_Alluvial_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the Ms
     Alluvial Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))

den <- apply(W_Corn_Belt_Plains,2,density)
plot(NA, ylab="Desnity", xlab = "Days",main = " Median Dry Season for the W
     Corn Belt Plains Region",xlim=range(sapply(den, "[", "x")),
     ylim=range(sapply(den, "[", "y")))
mapply(lines, den, col=1:length(den))
legend("topright", c("SOS","MOS","EOS"), lty = c(1,1),
       col = c("black","red", "green"))
```

```r
#Kruskal Wallis test for ecoregions

S_eco <- pairwise.wilcox.test(median_eco$SOS_0130_MEDIAN,
                              median_eco$eco_raster_1, p.adjust.method = "BH"
)
M_eco <- pairwise.wilcox.test(median_eco$MOS_0130_MEDIAN,
                              median_eco$eco_raster_1, p.adjust.method = "BH"
)
E_eco <- pairwise.wilcox.test(median_eco$EOS_0130_MEDIAN,
                              median_eco$eco_raster_1, p.adjust.method = "BH"
)

S_eco
M_eco
E_eco

S_eco_let <-multcompLetters(S_eco$p.value)
M_eco_let <-multcompLetters(M_eco$p.value)
E_eco_let <-multcompLetters(E_eco$p.value)

S_eco_let
M_eco_let
E_eco_let
```