FUEL FOR WILDFIRE:

CONTROLS ON THE DISTRIBUTUION OF WILDFIRE IN THE SOUTHEASTERN

UNITED STATES

A Thesis

by

MAX ALEXANDER MORRIS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Charles Lafon
Co-Chair of Committee,	David Cairns
Committee Member,	Michelle Lawing
Head of Department,	David Cairns

May 2019

Major Subject: Geography

ABSTRACT

Ecologists and wildlife managers alike have explored the role of fire as an ecosystem disturbance for decades and, yet, the role of scale remains poorly understood in pyrogeography. Understanding how wildfire occurs on the landscape and, furthermore, how these trends will change in the future provides an enhanced understanding of vegetative patterns, successional changes and biome distributions. As scientific research begins to account for the effects of climate change, predictive modeling will remain one of the foremost tools in understanding how present-day trends will begin to change. This study employs a series of spatial modeling techniques to examine which factors are most influential on the presence of wildfire hotspots on the landscape and which factors may be influential on areas devoid of wildfire occurrence entirely. Clustering algorithms were used to identify wildfire hotspots across the study area and targeted pseudo-absence points were created outside the bounds of these clusters. The resulting presence/absence points were analyzed within physiographic regions and a predictive model was fit to the data. Analysis of common covariates, such as climatic variables, land use, and topography allowed this study to not just fit a model to wildfire distribution, but inform comparable studies conducted anywhere similar data are available. As different aspects of climate change begin to exert influence on ecosystems globally, this study sheds light on how fire regimes may change with it.

DEDICATION

> “Every one thing exists for the sake of all things and all for the sake of one; for the one is of course the all as well. Nature, despite her seeming diversity, is always a unity, a whole… a relationship to the rest of the system.”

- Alexander von Humboldt, *Personal Narrative of a Journey to the Equinoctial Regions of the New Continent*

This thesis is dedicated to the Geographers that have come before and those that have yet to leave their mark. However small or grand the contributions of this study, it was driven by the idea that all research builds to a greater understanding of our world and those that inhabit it. It is my sincerest hope this study instills on its readers the same desire for knowledge and passion for adventure that inspired its creation.

ACKNOWLEDGEMENTS

"Much like a turtle on a fence post, you know I didn't get here by myself" - Unknown

It's said that when President Kennedy visited a NASA facility, he asked a custodian he met what they did there. The custodian simply replied that he was helping to put a man on the moon. In much the same way, my successes at Texas A&M University would not have been possible without the support of the staff and faculty of the Department of Geography, College of Geosciences, and Department of Ecosystem Science and Management.

I would like to especially thank my committee chair, Dr. Lafon, and my committee members, Dr. Cairns, and Dr. Lawing for their guidance and support throughout the course of this research. This study would not have been possible if not for the immense time Dr. Lafon spent with me from start to finish, the verification and conceptual support of Dr. Cairns, and the sage advice and kind encouragement of Dr. Lawing. I could not have asked for a better committee.

I owe my sanity in this trying period of my life to my friends, family, and colleagues at Texas A&M University and around the country. Your guidance and friendship always kept me on track through this difficult endeavor.

Finally, though I know you cannot read this, Izzy, my best friend and dog, I couldn't have done this without you, little buddy. You are a constant ray of sunshine in my life and I couldn't be more thankful for you.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a thesis committee consisting of Dr. Charles Lafon, Dr. David Cairns of the Department of Geography and Dr. Lawing of the Department of Ecosystem Science and Management.

Funding Sources

The work of this study would not have been possible if not for the generous funding provided by the Department of Geography at Texas A&M University through Graduate Teaching Assistantships and tuition waivers. Further funding was made available through the College of Geosciences from the M.T. Halbouty '30 AAPG Endowed Fellowship and Lechner Scholarship and through independent research assignments through the School of Innovation.

NOMENCLATURE

ESDA	Exploratory Spatial Data Analysis
SAC	Spatial Autocorrelation
EBF	Eastern Broadleaf Forest Physiographic Sub-region
APP	Appalachian Physiographic Sub-region
CSTL	Coastal Plains Physiographic Sub-region
EPD	Expanded Piedmont Physiographic Sub-region
EVR	Everglades Physiographic Sub-region
SE	Southeastern Regional Study Area
MFI	Predicted Mean Fire Interval
USCB	United States Census Bureau
USGS	United States Geological Survey
GAP	USGS Gap Analysis Program
MRLC	Multi-Resolution Land Cover Characteristics
DEM	Digital Elevation Model
KDE	Kernel Density Estimation
OR	Odds Ratio

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

1. INTRODUCTION

Wildfire is a foundational element of global ecology, of ecosystem destruction, and revitalization, but most of all: change. Often, wildfire is captured in the public perception as a force of destruction and this is not without cause. In 2017 alone, roughly 71,500 wildfires were responsible for burning over 10 million acres of land (Hoover 2018). When the paths of these fires intersect with areas of human-development and land use, the results can be devastating. Wildfire is not, however, solely a destructive force. Fire, whether from natural or anthropogenic forces, often works as the spark of change; it shifts ecosystem dynamics, encourages successional cycles, and sets the conditions for the growth of new life. Pyrophillic vegetation are highly adaptive species the utilize fire to grow; species such as the longleaf pine (Pinus palustris) require fire, using their fire-resistance as a competitive advantage to grow and spread in ecosystems (Latham 2013). Cyclical disturbance by fire furthermore serves to open habitat for wildlife that cannot exist in thick woody vegetation otherwise. Fire serves an essential role in how ecological systems experience change and how fire-adapted species grow and develop. Fire-adapted species use wildfire to expand to new territories and garner resources from otherwise highly-competitive vegetation. The magnitude and frequency of fire serves not just to aid these species, but in many cases works to define the vegetation characteristics of an entire region (Lafon et al. 2017).

Despite the known ecological benefits of fire, suppression remains a common reaction to the presence of wildfire. The answer to lasting fire control may lie in the age-old idiom ‘fighting fire with fire’. Evidence has shown that frequent low-understory fires are effective in curbing the creation and spread of large, destructive wildfires (Brose et al. 2001, Latham 2013).

Fire is a universal, global phenomena but wildfire distribution is far from consistent across different scales and regions. Different ecosystem types experience a greater or lesser frequency of fire, with these variations exerting strong influence on their vegetation and successional cycles. This research aims to analyze and interpret where wildfire is most and least prevalent across the Southeastern United States, and which factors are the most influential on these conditions. This study will advance the scientific understanding of what controls wildfire distribution. With that knowledge, ecosystem managers and fire-suppression agencies alike will be better prepared to control fire before it threatens human populations, but allow it to flourish where it is a vital part of ecosystem dynamics. The fire regimes of the United States will likely to continue to change, but understanding how they are developed is the first step to asserting any kind of control upon them.

To advance this regional understanding of wildlife, this research will focus on two primary objectives. The first objective is to characterize the known distribution of wildfire in the Southeastern United States. Wildfire distribution will be evaluated to identify the presence of clustering, hyper-dispersion, or random distributions of wildfire across spatial scales and physiographic sub-regions. Non-random distributions of wildfire indicate a control on the distribution of wildfire. Therefore, the second objective of this study is to identify potential factors most influential on the distribution of both wildfire presence and absence. Variables found to be consistent across all spatial scales and sub-regions tested can be identified as having the strongest association on the presence of wildfire.

2. LITERATURE REVIEW

Research attempting to identify patterns in fire regimes is of the utmost importance to ecologists and ecosystem managers. Fire exerts a heavy influence on the disturbance and subsequent successional ecology of all continents, excluding Antarctica (Pausas & Ribeiro 2013). Increasingly, research in pyrogeography has been focused on defining large-scale patterns and interactions of fire as a disturbance regime. Bond & Keeley (2005) proposed that the ubiquity of fire as a disturbance around the world was so profound that it could be compared to the effects of grazing or browsing commonly done by large ungulates. This idea substantiated the role of fire in global ecosystems but only addresses the results of interconnected fire regimes. The interconnectedness and importance of fire regimes on global ecology is not a novel topic, however, several years later, Archibald et al. (2013) popularized the term ‘Pyromes’ as a way of classifying how fire regimes differ in association with vegetative and climatic patterns often used in defining biomes. Like Bond et al. (2005), Archibald et al. (2013) contribute to the global understanding of fire regimes, even providing a potential classification scheme based on the physical characteristics, frequency, and location of fires. It does not, however, offer evidence as the environmental, anthropogenic, or topographic factors that set the conditions for the trends observed in these regimes.

These advances in the global fire regime research have demonstrated the importance of fire research at large spatial extents, but offered little evidence on what factors influenced these trends or how relevant factors themselves may change regionally. Previous studies have identified criteria for determining the presence or absence of fires based on the ecological or anthropogenic preconditions of the area (Bowman et al. 2014, Bradstock et al. 2010, Pausas & Ribeiro 2013), but none have sought to identify the statistical influence of both specific factors on fire distribution. Bowman et al. (2014) further express the value in understanding the underlying causes of cyclical disturbance regimes, such as fire, on ecosystems. When dealing with dynamic systems such as large-scale ecological succession, it is imperative that testing variables be as independent as possible. Their study suggests that determining global fire regimes or pyromes using only macroecological factors is too broad; these factors are interdependent and reliant on the same cycles that wildfire is an integrated part of. Therefore, a study that identifies a set of predictive criteria for large-scale fire regimes based on the core factors that make up ecological designations (such as biomes) is required. This study seeks to determine how large-scale trends in pyrogeography are influenced or controlled by simple environmental, anthropogenic, or spatial factors.

To define a clear set of criteria for the prediction of wildfires will require more than just a point-based analysis. Studies of a similar nature have, thus, used the clustering or aggregation of fire locations to define areas of high or low fire susceptibility. While these data can further contribute to determining the influence of controlling factors on fire regimes, it also has the potential to demonstrate change in fire regimes over time. Serra et al. (2013) provides a framework to model these factors and evaluate wildfire clusters but was limited in both scope and depth. Their study evaluated covariates to fire clusters such as land use and ignition yet did not include any environmental controls on fire distribution. Additionally, it was limited to a small study area of Catalonia, Spain over a brief time scale of four years. Serra et al. (2013) was a useful study in identifying controls on the distribution of fire clusters and Chou (1992) and Parente et al. (2016) help complete the picture. These studies successfully tested relevant controls on fire clusters. The factors that these studies tested covered topics such as climate, fire prevention policy, spatial influence of past fires, and anthropogenic land use. Neither study uses, nor generates, a comprehensive list of covariates to test wildfire clusters on the landscape. These three studies together support a methodology for testing the influence of covariates on the spatial and temporal distributions of fires, but leave a gap to examine a study area large enough to draw conclusions on changes in large-regional or global fire regimes. Research building upon the groundwork laid out by these studies will be able to identify the covariates most influential on the spatial distribution of wildfires on a regional scale.

Bradstock (2010) was one of the few studies that sought to define criteria for the prediction of wildfires at a continental extent. This study uses variations in primary productivity and vegetation cover to suggest changes in Australian fire regimes. While this study was able to identify major controls on fire regimes, it nonetheless came under criticism in its validity. Bowman et al. 2014 questions whether such a model could be valid considering the inherent feedback cycles between climate, fire, and primary productivity.

Krawchuck et al. (2009) provides one of the most comprehensive analyses of the distribution of wildfire, with particular interest given to how these distributions change in relation to climate change. This study does face its limitations, however. While a fine example of how the environmental factors influential in the distribution of wildfire will change global fire regimes, this study limits potentially relevant predictive criteria to environmental or ignition criteria. Furthermore, this does not consider how variations in significant predictive factors occur within their study region. Co-authors of the study expanded on this work in Parisen & Moritz (2009) by analyzing the large-regional contiguous United States amongst regional California and smaller ecoregions. By analyzing several spatial scales and incorporating more potential covariates, this study begins to build on the comprehensive understanding of which factors are most influential in the creation of fire regimes at different scales. My research attempts to build on this foundation by expanding the non-environmental testing covariates included in this analysis. In this way, this research will be able to provide support to, or evidence to refute the importance of non-environmental variables on regional and sub-regional wildfire distribution.

Contrary to public perception, the bulk of wildfire is not caused by natural or environmental factors alone. Balch et al. (2017) analyzed wildfire records from 1992-2012 to determine the influence of human impact on wildfire in the United States. They discovered that of fires analyzed, 84% were human-ignited in some way. Furthermore, Balch et al. (2017) identified that 44% of the total area burned in the time period was the result of these human-ignited fires. The study found that not only is human-impact causing more fires and burning more area than purely natural causes, but human-impact is expanding the annual temporal and spatial niche of wildfire. Outside of the United States, similar studies have brought forth similar results. Calviño-Cancela et al. (2016)

analyzed the wildfire risk associated with urban-wildland interfaces in Northwestern Spain. Like Balch et al. (2017), the study found that areas where urban and wildland areas interfaced increased ignition risk in most land cover types. The importance in human-fire interaction is, therefore, monumental, further illustrating the value in identifying how human influence effects fire *distribution*.

Ricotta et al. (2018) found that ignition points were generally close to roads in all land cover types, with the ignition risk varying amongst these types. Roads are not only a method for human transportation but are a sign of built infrastructure and human development. Roads are expensive to build and maintain; generally, the higher the road density of an area, the higher the volume of persons that they have been built to cater to. Additionally, with increases in road density, it can be expected in most cases that other increases in infrastructure will follow (e.g., powerlines, buildings, etc.). These shifts in infrastructure represent an increased risk of fire ignition both through direct or indirect human influence. Their research has thus set the framework for the inclusion of road density as a relevant factor in the prediction of wildfire.

Another measure of human-involvement on the landscape, population density, has also been found to correlate with wildfire. Increases in population density have resulted in both increases and decreases in the occurrence wildfire largely depending on land cover (Blistinas et al. 2013). While it would be expected that increases in population density would lead to a wildfire rise in fire-prone areas, it would be equally expected that wildfire would be found less frequently as areas begin to heavily urbanize and suppression becomes more prevalent (Calviño-Cancela et al. 2016). As with roads, land cover/land-use is essential to measuring the impact of human influence on the presence and distribution of wildfire. These studies substantiate that the role of human-impact on the distribution of wildfire is interconnected with the natural environment. The relationship between these groups of variables is, thus, essential to the prediction of wildfire.

Just as environmental predictors alone are not wholly satisfactory in explaining the distribution of wildfire, land cover is not a catch-all. To effectively build a more comprehensive understanding of wildfire, it is imperative that different aspects contributing to wildfire be analyzed together. Conducted in a similar study region, Carmo et al. (2011) explored the relationship between land cover and slope on the presence of fire. They found, along with studies of a similar nature (Nunes et al. 2005, Moreira et al. 2009) shrub land cover to be the most fire-prone land cover with cultivated lands being the least fire-prone. It is not necessarily valid to assume that results generated from studies in different regions the world will hold true for research in the United States, but Carmo et al. (2011) helps to identify trends expected to hold true. Carmo et al. (2011) attributes falling fire prevalence in cultivated land to the controlled moisture content of irrigated crops and ranchland. The study also theorizes that rapidly changing elevation is likely a powerful indicator in the link between land cover and wildfire prevalence. Their research found that strong slopes increased the fire-risk of all land cover types. This relationship illustrates the importance of both elevation and a metric of surface roughness in evaluating the distribution of wildfire.

Carmo et al. (2011) focuses on the relationship between land cover and elevation in explaining wildfire distributions, whereas within the Great Lakes Region of the U.S., Cardille & Ventura (2001) explores the relationship between land cover and land-ownership on the occurrence of wildfire. Cardille & Ventura (2001) found that wildfire was more likely to occur in areas that were not forested than areas that were within forested patches, but the presence or absence of wildfire otherwise varied amongst land cover types. Furthermore, they found that areas within the bounds of national or state-protected forests were less likely to have wildfire than comparably sized areas outside of them. These results were found to be consistent across their testing scales demonstrating a strong relationship between wildfire and these variables. Their study asserts that correlation between the absence of wildfire and the federally or state protected land may be the result of fire-suppression practices or the absence of high-density human-involvement. Just as land cover is a powerful covariate in conjunction with human-involvement, their study supports the idea that ownership of the land may also play an important role.

Factors that contribute to the distribution of wildfire are only valuable to the degree that they can be tested. This research aims to not just explain, but to quantify the contributions of these significant factors on the distribution of wildfire. Before these analyses can be tested, however, the distribution must be understood. The methods used to test distribution in this study are rooted in a collection of techniques called exploratory spatial data analysis (ESDA). ESDA was largely developed by Anselin (1998) as a series of techniques that focused on accounting for spatial autocorrelation

(SAC) in spatial data. It was developed as a broad mechanism for the exploration of spatial phenomena. ESDA sets the framework for research on the distribution of points in space; it is inherently geographic, with applications from spatial epidemiology to biogeography. ESDA can be broken down into 3 core segments. First, the dataset must be visualized, or presented. In many cases, data must be appropriately ascertained and presented in a way that could be tested using the techniques involved in ESDA. Secondly, the distribution of the data must be reviewed for evidence of SAC. To understand what influences SAC of a dataset, and how it can be predicted, a spatial relationship, the SAC of a dataset must be verifiable. Finally, the third core ESDA segment is the prediction of spatial patterns.

This combined body of literature substantiates the scientific dedication to research in pyrogeography. Each study, paper and piece of journal correspondence contributes to the greater understanding of the intermingled factors that contribute to the distribution and spread of wildfire: a global ecological phenomenon. From global to local scales and nearly all scales in-between, literature has identified the value of studying wildfire, its importance in global ecology, its relationship with human-interaction and land cover. By understanding these relationships and by utilizing techniques such as those outlined in ESDA, this research will build on the foundation that came before it. This research will not just quantify the importance of different factors on the distribution of wildfire, but synthesize how the relationships defined in previous literature are represented across this study area.

3. DATA

3.1. Study Sites

The influence of wildfire is a common topic for ecologists, ecosystem managers and even homeowners throughout the world. This study focuses on the influencing factors of wildfire within the southeastern United States (see *Figure 1).* The southeastern U.S. was chosen not just to help fill a gap in scientific literature, but because the distribution of major wildfires in this region fit the right criteria to explain which factors most influence the presence or absence of wildfire. This study focused on large wildfires (>500 acres burned) in the Southeastern U.S. (see *figure 1).* From a cursory observation it appears that to some degree, there are areas where wildfire is clustered, and some areas that appear almost completely devoid of wildfire. The disparity in the distribution of wildfire across the region makes it optimal in investigating how these trends occur.

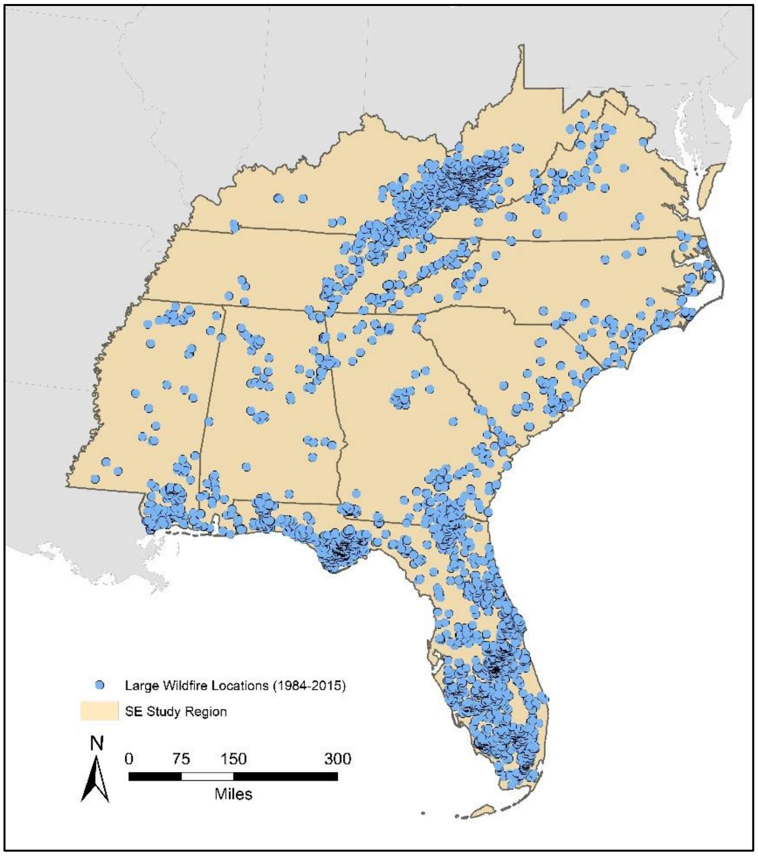

Figure 1: Presence of Wildfire in the S.E. U.S.

Just as there are areas of tightly clustered points and areas sparsely populated, the land, too, changes drastically between sub-regions within this selected region. The vast changes in physiography between the Appalachian Mountains and the Everglades identify the need for analyses at differing spatial scales. To help explain the role of scale in distribution, the southeastern study region was further divided into sub-regions based on their physiography. *Figure 2* displays the five physiographic regions that this study analyzes. These regions: The Eastern Broadleaf Forest (EBF), Appalachian (APP), Expanded Piedmont (EPD), Coastal Plain (CSTL), and Everglades (EVR) were loosely derived from Bailey's Ecoregions (Bailey 1998). Sub-regions delineated in Bailey 1998 were aggregated based on physiography to define the sub-regions for this study. Each of the 5 defined sub-regions as well as the full regional study area (Southeast – SE), were defined, and prepared for further analyses. By analyzing each sub-region individually and comparing the results between them, this study seeks to further the scientific understanding of what causes variations in the distribution of wildfire.

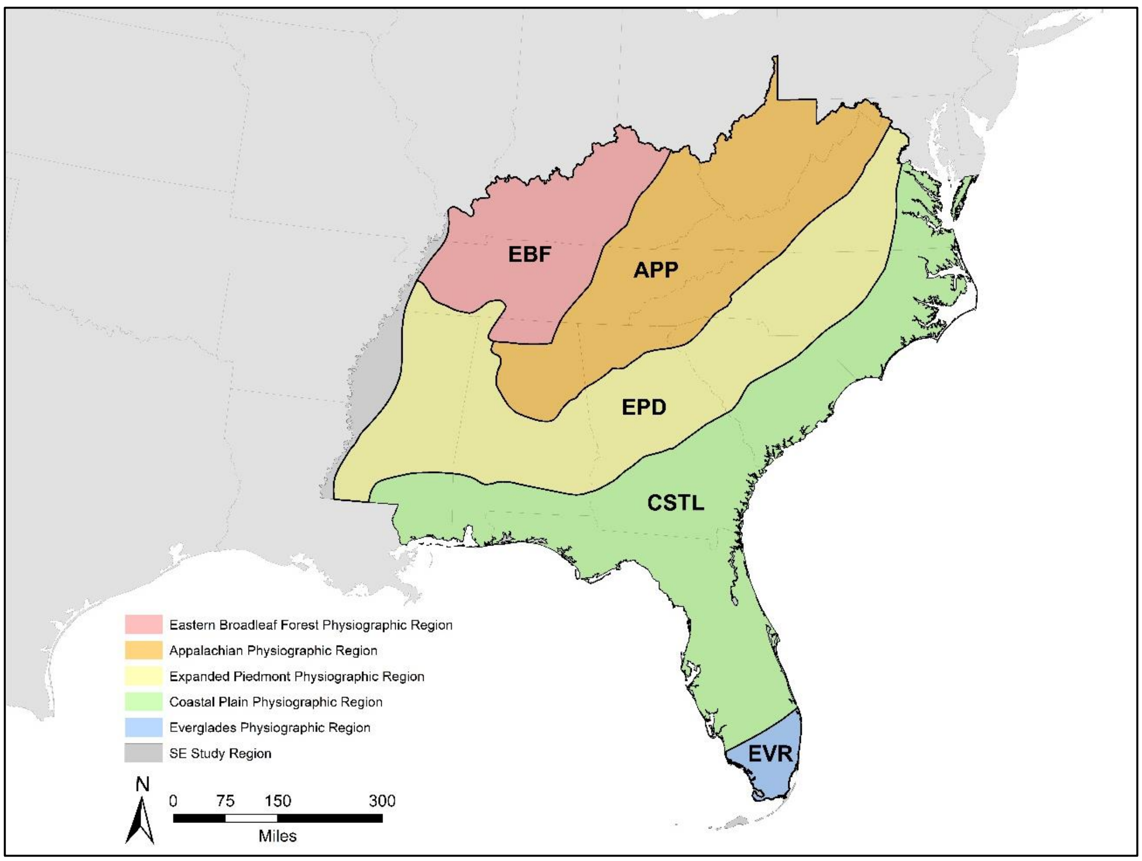

Figure 2: Physiographic Sub-regions of the S.E. U.S.

3.2. Vector Point Data - Wildfire Data Points

The variations in the distribution of wildfire would not be possible if not for data that accurately captures this information. This study derived wildfire presence data from

Monitoring Trends in Burn Severity (Finco et al. 2012, www.MTBS.gov). These data provide the position of all fire occurrence from 1984 – 2015. The data was constricted to only locations of major wildfires (acres burned >= 500). Data was then divided amongst the six testing groups (SE, EBF, APP, EPD, CSTL, and EVR) for pseudo-absence creation and analysis. Table 1 displays the distribution of point data among regions.

LOCATION	PRESENCE	PSEUDO-ABSENCE
SE	3000	4501
EVR	296	444
CSTL	1773	2660
EPD	90	135
APP	828	1242
EBF	13	20

Table 1: Wildfire Presence and Pseudo-absence by Study Region

3.3. Raster Data - Covariate Testing Data

This study is exploratory in nature and thus compared a total of 36 covariate raster datasets to the dependent wildfire dataset established. Wildfire occurrence has been attributed to a widespread series of variables ranging from anthropogenic/societal to environmental/topographic (Parente et al. 2016, Serra et al. 2013, Guyette et al. 2012, Lafon et al. 2017). As a result, this study sought to evaluate the influence of a series of covariates on the distribution of wildfire. Table 2 displays the covariates, their class-type, a short description and the source of the data.

NAME	CLASS	DESCRIPTION	SOURCE
BIO1	Bioclimatic	Annual mean temperature	Worldclim.org
BIO2	Bioclimatic	Mean monthly diurnal range	Worldclim.org
BIO3	Bioclimatic	Isothermality	Worldclim.org
BIO4	Bioclimatic	Temperature seasonality	Worldclim.org
BIO5	Bioclimatic	Max. temp. of warmest month	Worldclim.org
BIO6	Bioclimatic	Min. temp. of coldest month	Worldclim.org
BIO7	Bioclimatic	Temp. annual range	Worldclim.org
BIO8	Bioclimatic	Mean temp. of wettest quarter	Worldclim.org
BIO9	Bioclimatic	Mean temp. of driest quarter	Worldclim.org
BIO10	Bioclimatic	Mean temp. of warmest quarter	Worldclim.org
BIO11	Bioclimatic	Mean temp. of coldest quarter	Worldclim.org
BIO12	Bioclimatic	Annual precipitation	Worldclim.org
BIO13	Bioclimatic	Precip. of wettest month	Worldclim.org
BIO14	Bioclimatic	Precip. of driest month	Worldclim.org
BIO15	Bioclimatic	Precip. seasonality	Worldclim.org
BIO16	Bioclimatic	Precip. of wettest quarter	Worldclim.org
BIO17	Bioclimatic	Precip. of driest quarter	Worldclim.org
BIO18	Bioclimatic	Precip. of warmest quarter	Worldclim.org
BIO19	Bioclimatic	Precip. of coldest corner	Worldclim.org
WIND	Bioclimatic	Mean annual wind speed (m/s)	Worldclim.org
NAITO	Processed	Predicted mean fire interval (MFI)	Guyette et al. 2012, Lafon et al. 2017
POP	Anthropogenic	Population density	USCB

Table 2: Covariate Types, Descriptions, and Sources.

All datasets resampled to 1km resolution

NAME	CLASS	DESCRIPTION	SOURCE
ROADS	Anthropogenic	Road density	Author/USGS
PTCD	Anthropogenic	Binary protected land designation	USGS (GAP)
DECID	Land Cover	Binary deciduous land designation	MRLC
BARREN	Land Cover	Binary barren land designation	MRLC
SHRUB	Land Cover	Binary shrubland designation	MRLC
MXDFRST	Land Cover	Binary mixed forest land designation	MRLC
EVERGRN	Land Cover	Binary evergreen forest land designation	MRLC
WETLAND	Land Cover	Binary wetland designation	MRLC
DEVELOP	Land Cover	Binary developed land designation	MRLC
CULTIV	Land Cover	Binary cultivated land designation	MRLC
GRASSLND	Land Cover	Binary grassland designation	MRLC
RELIEF	Topographic	Topographic relief/roughness	Author/USGS
DEM	Topographic	Digital elevation model	USGS
ASPECT	Topographic	Categorical, directional slope	Author/USGS

Table 2 (Continued): Covariate Types, Descriptions, and Sources.

All datasets resampled to 1km resolution

This study also seeks to identify the strongest predictive criteria for wildfire using resources readily or easily available in many areas. The use of bioclim variables (a free, global, online dataset) is one of the ways that this objective is made possible. Although less certain to be available, many of the other datasets are derived from basic geospatial or population data that may be collected and processed by government entities around the world.

Many of the tested datasets were freely available through private or government entities but some had to be developed for this study. These datasets: roads, relief, and aspect were processed in ArcGIS Desktop Version 10.5 using data derived from the USGS (USGS 2015). The roads dataset was developed by attributing a count of 1 to each road in the USGS dataset then creating a summation spatial join to a fishnet polygon of the appropriate resolution. This polygon was then rasterized for further processing. The relief dataset was created as a measure of the standard deviation of elevation in a roaming 3x3 grid cell zone. This measure of topographic roughness or relief, attributed to Ascione et al. 2008 and Klinkenberg 1992, was derived from the USGS DEM Dataset (USGS 2015) and calculated in ArcGIS Desktop Version 10.5. Also based off the USGS DEM dataset, the aspect dataset was created using the 'Aspect' tool in the Spatial Analyst toolbox (ArcGIS Desktop Version 10.5) then converted to directional categories (cardinal directions and NE, SE, NW, SW) using RStudio Version 1.1456.

4. METHODS

4.1. Pseudo-Absence Creation

This research seeks to advance the scientific understanding of which factors contribute most strongly to the presence of wildfire on a landscape and furthermore, if these factors change across regions or sub-regions. To attain this goal, it is not satisfactory to explain the conditions of where wildfires are located; it is equally imperative to explain the areas where wildfire is not occurring. For this reason, this research developed a metric to be used to represent not the absence of wildfire, necessarily, but the pseudo-absence: a statistically random distribution of geographic points representing the absence of wildfire. Point locations of absence do not, by definition, exist. In composing an ESDA of the absence of wildfire, this inhibits the first segment (displaying the data). To circumvent this issue, a metric must be derived to represent absence. A statistically random distribution of points restricted to areas least likely to be afflicted by wildfire would accomplish the first two tasks of ESDA, priming these data to be processed alongside wildfire presence data.

Figure 3A represents the points of known wildfire locations on the landscape within the study period. There are some areas in the southeast where these locations seem concentrated and others where they appear to almost non-existent. To have pseudo-absence points represent areas where wildfire was not noted to occur, a metric had to be utilized to determine where they were most likely to occur. To represent the areas of wildfire concentration, a kernel density estimation (KDE) was performed. KDE estimates the probability density of a variable over a smoothed area (Rosenblatt 1956). The KDE is calculated using the presence of wildfire within a search distance or 'bandwidth'. This study made use of the h_{opt} formula presented by Fotheringham et al. 2000 to determine the optimal bandwidth where:

$$h_{opt} = [2/3n]^{1/4}\boldsymbol{\sigma}$$

Where h_{opt} (the optimal bandwidth) is a function of both the standard distance of wildfire presence data, $\boldsymbol{\sigma}$, and its sample size, n. Table 3 represents the calculation of the optimal bandwidth used for the KDE analysis in this research.

DATASET	n	$\boldsymbol{\sigma}$	h_{opt}
SE MTBS WILDFIRE POINTS	3001	513387.6882m	62676.702m

Table 3: H_{opt} Calculation for S.E. Wildfire Points

Once the KDE was performed, concentrations of wildfire could be reliably determined. *Figure 3B* represents the concentration of wildfire on the landscape by 20th percentiles. The 80th percentile of wildfire concentration on the landscape was used to create 'Pseudo-Absence Fire Exclusion Zones' (See *Figure 3C*) that represented areas where wildfire occurrence was most concentrated across the landscape. With these exclusion zones delineated, pseudo-absence points could then be populated in areas outside of these exclusion zones to represent areas where wildfire was absent.

ArcGIS Desktop version 10.5 was used to create points locations outside of the Fire Exclusion zones using the 'Create Random Points' tool in the Data Management toolbox. These points were populated at a 40/60 presence to pseudo-absence ratio within each delineated physiographic sub-region. This ratio was selected to provide a nearly even number of values for each category while providing more user-generated points to ensure a large sample size for analysis. Table 1 represents the total points produced for each study area and *Figure 3D* represents the distribution of these points and their respective coverage of the study regions as a whole.

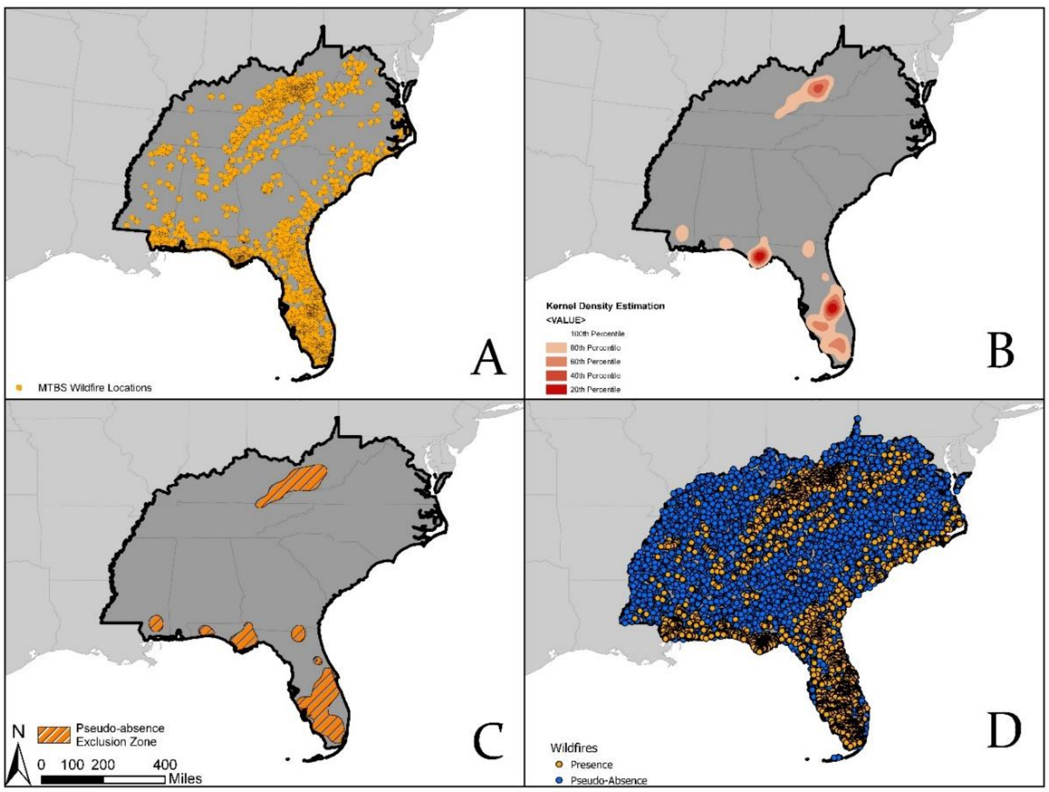

Figure 3: Illustrated Method for the Creation of Pseudo-Absence of Wildfire in the S.E. U.S.

4.2. Covariate Data Creation and Formatting

To properly relate wildfire presence and newly created pseudo-absence data it was essential that 36 raster testing covariates be standardized. Additionally, in further predictive modeling of the study regions, the covariates selected as predictive variables must be aligned in order to be used as a predictive stack (group of aligned rasters used for modiling). As a result, this research not only acquired, but cleaned, standardized and aligned the 36 raster datasets. Raster processing was completed in ArcGIS Desktop Version 10.5 using the Data Management toolbox. All datasets that deviated from the set 1km resolution were standardized using the 'Resample' tool and the 'Majority' algorithm. This method ensures that rasters at a finer resolution were aggregated to a value encompassing the majority of the cells within it, as well as coarser resolution rasters defining the same value evenly across newly delineated cells. Once all raster datasets were standardized to the appropriate resolution, they were fit to a common extent and aligned. These steps, essential to forming a raster stack to be used in predictive analysis, were also completed using the Data Management toolbox in ArcGIS Desktop version 10.5. Each of the covariate rasters was separately, but in turn, mosaiced onto the blank raster using the 'Mosaic to New Raster' tool. The resulting rasters, each with a standardized extent, resolution and alignment were then all masked to encompass only the study region.

Once the raster datasets were aligned and standardized across extent and resolution, the covariate data at each of the wildfire presence and pseudo-absence points were extracted and joined. For each of the study regions, the 'Extract Multi Values to

Points' tool in the Spatial Analyst toolbox was used to extract the raster value of each of the covariate rasters to each presence and pseudo-absence point. The resulting attribute files were then exported from ArcGIS for processing.

4.3. Measuring Spatial Autocorrelation

Kernel density estimation provides a visual representation of the probability density of a variable across a study area. While useful in identifying concentrations, it fails to adequately quantify spatial autocorrelation, particularly as it changes amongst different spatial scales. This research is centered on explaining distributions of wildfire, therefore a metric to quantify this distribution was needed. To accomplish this task, this study made use of the Ripley's K Statistic (Ripley 1976). The Ripley's K Statistic is a point pattern analysis that measures spatial autocorrelation at multiple spatial scales. Measuring spatial autocorrelation at different scales quantifies the spatial dependence of the feature data. This study focuses on both regional and sub-regional data and, as a result, understanding the spatial dependence of wildfire distribution furthers the scientific understanding what factors are most significant in controlling the distribution of wildfire in the Southeastern United States. *Figure 4* represents the graphic output of the Ripley's K function through ArcGIS. The output graphic visualizes the relationship between the autocorrelation of feature and space as it differs from the pattern expected of a sample. Within ArcGIS Pro Version 1.4 the 'Multi-Distance Spatial Cluster Analysis (Ripley's K Function)' tool from the Spatial Statistics Toolbox was used to accomplish this task. Each of the study regions were tested separately using 10 distance bands, the Ripley Edge correction formula and repeated for 999 permutations to create a confidence envelope. Distances used for each region were scaled according to relative region size. The output of these analyses was generated graphically for further evaluation.

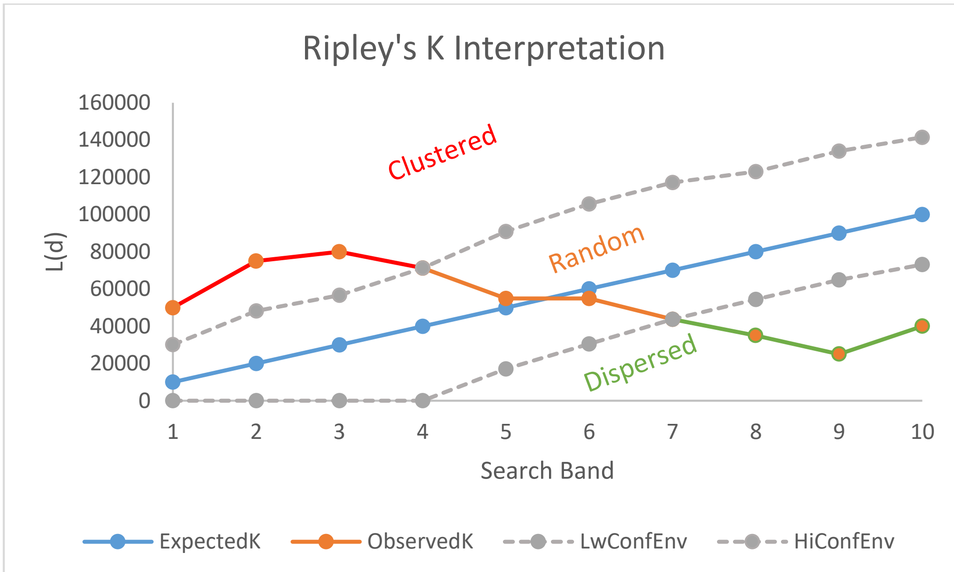

Figure 4: Ripley's K Interpretation

4.4. Regression Analyses

This study makes use of variations of logistic regression analyses to evaluate the relationship between wildfire presence/pseudo-absence, a binary dependent variable, and 36 Bioclimatic, anthropogenic, and topographic independent variables. Logistic regression models calculate not just the relationship between dependent and independent variables, but the probability that dependent feature data will be within a testing category (Menard 2010). The result of a logistic regression is, thus, a measurement of the expected odds that a dependent feature falls within an independent category at some magnitude.

Before any analyses could be performed, the testing vector data had to be cleaned. At a regional to sub-regional scale, errors in raster data are not uncommon. The analyses performed for this study however cannot be run without complete datasets. As a result, any data points either in the presence or pseudo-absence categories that represented an incomplete recording of attributed covariate data were not processed. *Table 4* displays the data reductions attributed to the cleaning of this data. While the loss in data is not ideal, the inflation of data by the creation of pseudo-absence ensures that there remain enough data points to be processed.

LOCATION	PROCESSED PRES	PROCESSED ABS	PROCESSED TOTAL	PRE-PROCESSED TOTAL
SE	1622	3092	4714	7501
EVR	150	196	346	740
CSTL	377	792	1169	4433
EPD	27	47	74	225
APP	1620	2910	756	2070
EBF	9	16	25	33

Table 4: Data Reductions by Study Region

With a complete dataset created and covariate data attributed, regression analyses could be performed. The preliminary logistic regression model run for each study area was a generalized linear model processed in RStudio Version 1.1456 using the ‘glm’ function within the ‘stats’ package. Model fit for these analyses was evaluated using the ‘extractAIC’ function in the ‘stats’ package. AIC, or Akaike Information Criterion, is a method for measuring the relative quality of statistical models (Aho et al. 2014). This value is relative to the model being tested and can be used to tweak models to achieve the lowest AIC, indicating the highest quality model. The ‘glm’ method accomplishes the basic task of performing a logistic regression, but it doesn’t adjust covariate use to account for model fit. Therefore, any changes to the variables used in a model using a ‘glm’ must be removed or added manually before the model reevaluated. To ensure the best fit model, this study made use of a stepwise regression model.

Stepwise regression models optimize the fit of the model by analyzing the significance of covariates in different ways. The first way that this can be accomplished is in a ‘forward’ method, where each covariate is added to the model in forward inclusion and the model fit is measured at each step (Menard 2010). Conversely, the ‘backward’ method works through a backward elimination of variables, starting with full covariate inclusion in the model and removing them at each step. Both methods can be evaluated using the ‘stepAIC’ function in the ‘MASS’ package and the better fit model of the two selected.

Stepwise regression analyses have been criticized as representative of a researcher’s poor understanding of a subject (Studenmund & Cassidy 1987). This criticism is derived from the method’s using an automated stepwise selection as opposed to the previously presented manual method. Critics of this method argue that automated stepwise regression is more like taking a ‘shot in the dark’ at valuable covariates rather than demonstrating true understanding. Conversely, stepwise regression has been defended as a valuable resource in directed exploratory research (Agresti & Finlay 1997). This study seeks to narrow down the most statistically significant factors in the distribution of wildfire. This study is, at its core, an exploratory analysis. Exploratory research has and always will be an essential part of the scientific process and in cases of directed exploration, stepwise regression is a powerful tool to aid in that process. Not only do the included variables matter in stepwise regression, but so do does the order that they are processed in. To ensure model validity, this study will first see if relevant covariates change based on the method of stepwise regression (forward or backward). If the relevant covariates from these match, than the model remains valid. Otherwise, the organization of covariates must be randomized to ensure each covariate is not exerting greater influence on the model than any other. Once a method has been selected, processed and verified, it can be interpreted.

Once a final regression model has been selected for each study region, their interpretation will be done using a measure of the regression coefficient: their associated Odds Ratio (OR). In a logistic regression model, the regression coefficient is equal to the log(OR). To simplify this interpretation the OR is calculated from the coefficient. OR is defined as the ratio of the odds of presence to the odds of absence, given each particular covariate. OR values, therefore, range from zero to positive infinity with values below one representing a negative correlation between the presence (positive correlation with absence) of the dependent variable and an increase in the covariate, one representing no correlation, and values greater than one showing a positive correlation between the presence of the dependent variable and an increase in the covariate. This study will use the OR generated by each region's model to quantify the significance of each variable. An important consideration when interpreting the OR of any variable in a model is that the OR is a measure of the probability or absence of a dependent variable, given an increase in the independent variable. For binary variables, this simply means with probability of dependent variable presence, given the presence of the independent variable. For continuous variables, the probability is given an increase in the independent variable. In this way, interpreting the OR of each model run will allow this study to quantify which variables are exerting the most influence on the presence or absence of wildfire.

5. RESULTS

5.1. Objective 1: Spatial Autocorrelation

5.1.1. Southeastern Regional Study Area

The multi-distance spatial cluster analysis for the southeastern region was an important step in characterizing the nature of wildfire distribution across this large regional area. *Figure 5* displays how the spatial autocorrelation of wildfire points changes across spatial scales. From search bands 1 through 6, or measuring distances 100 km to 600 km out from any given point, results indicated that these data are strongly, positively autocorrelated. This clustered distribution becomes increasingly more random at higher search distances, but never crosses the threshold into a random distribution within these confidence envelopes for a random distribution.

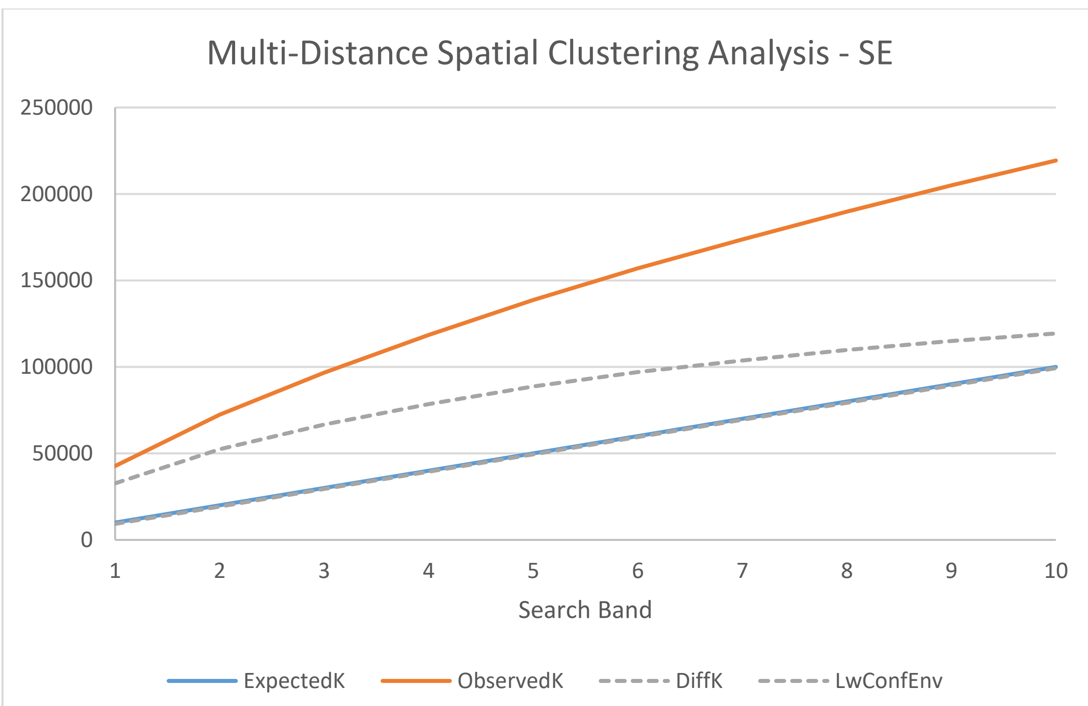

Figure 5: Multi-Distance Spatial Clustering Analysis - SE

5.1.2. Eastern Broadleaf Forest Physiographic Sub-region (EBF)

The EBF region contained the fewest number of points of any of the study regions and offered the most unique distribution of wildfire. Each of the regions studied found a measure (of varying degrees) of positive spatial autocorrelation amongst them indicating a clustered distribution. Conversely, the EBF region fell within the confidence envelope for a complete spatially random distribution at all search bands. *Figure 6* displays a graphic representation of the random distribution of wildfires across all the spatial scales in this region.

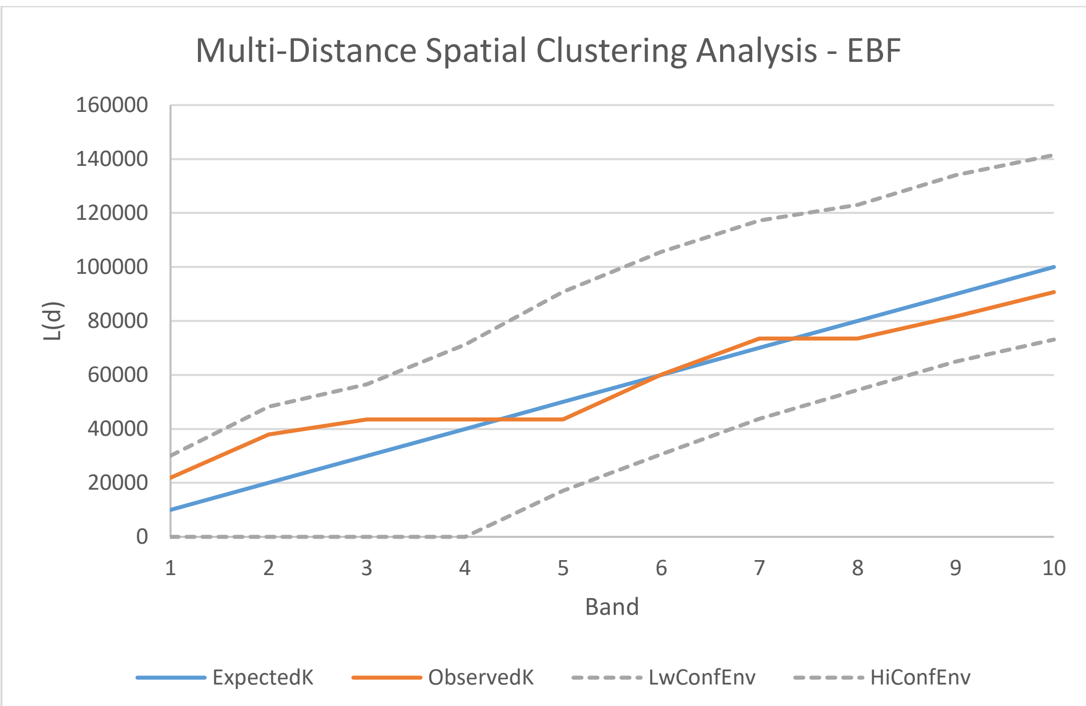

Figure 6: Multi-Distance Spatial Clustering Analysis – EBF

5.1.3. Appalachian Physiographic Region (APP)

The APP region demonstrated strong spatial autocorrelation across all spatial scales (see *Figure 7*). The predicted Ripley's K result, L(d), far surpassed the expected distribution indicating strong, positive spatial autocorrelation, or clustering. This result was consistently positive, but lower degrees of clustering were found at both smaller and larger distance bands, with the highest disparity between expected and observed K values occurring in the middle search bands.

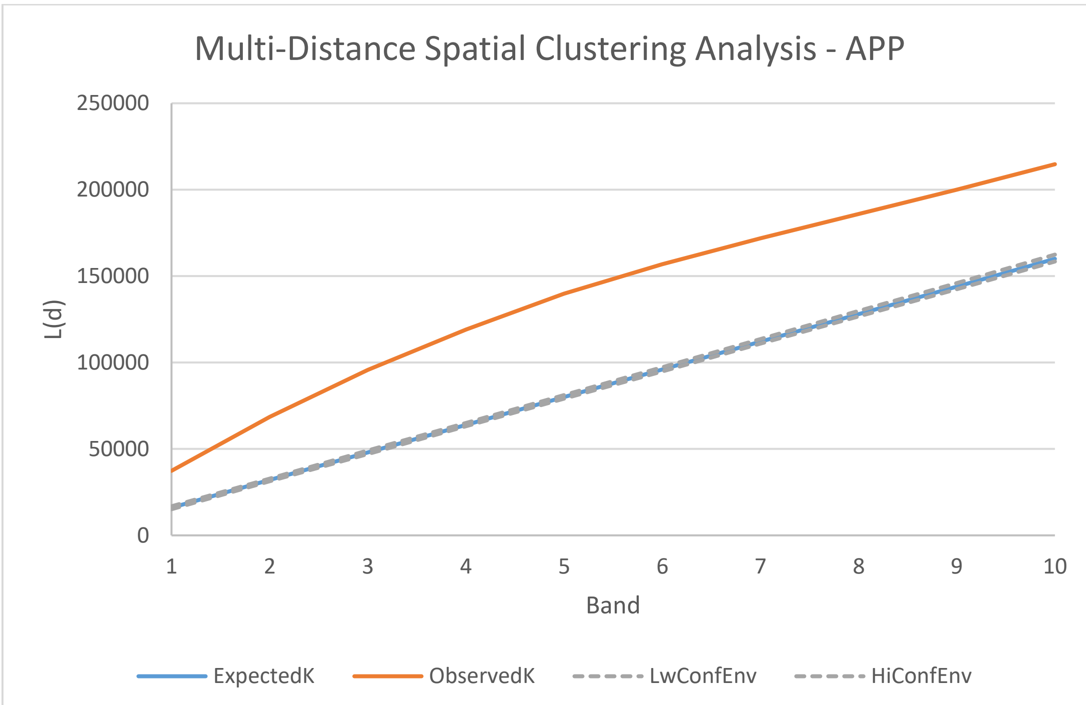

Figure 7: Multi-Distance Spatial Clustering Analysis – APP

5.1.4. Expanded Piedmont Physiographic Region (EPD)

Like the EBF region, the EPD region had relatively few points when compared to many of the other study regions. Despite this, this region exhibited a vastly different distribution. Following a trend observed in the results of many of the other study regions, the EPD region exhibited strong spatial autocorrelation across all search bands (see *Figure 8*). Unlike these other regions, however, a greater variance was observed in the confidence envelope of this region. Furthermore, the greatest deviance from the expected K values was exhibited in the smaller bands (bands 3 & 4) before leveling out to a more consistent representation at larger search bands.

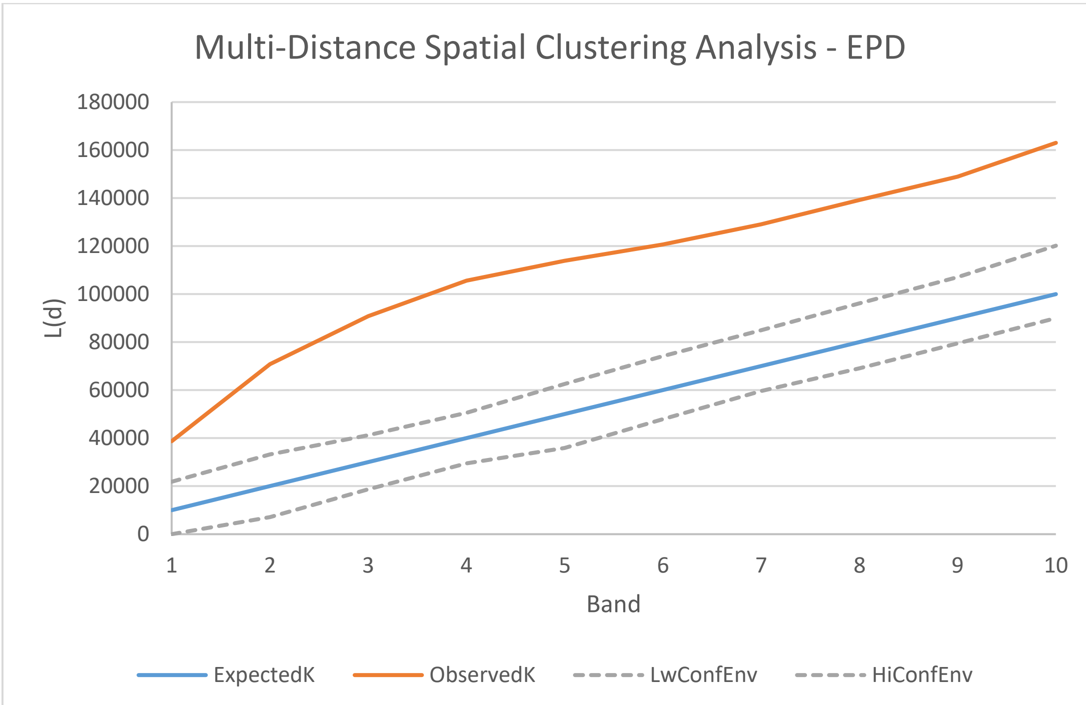

Figure 8: Multi-Distance Spatial Clustering Analysis – EPD

5.1.5. Coastal Plain Physiographic Region (CSTL)

The CSTL region generated similar results to those found in the APP region. While the search bands were greater to accommodate the larger size of this region, strong, positive spatial autocorrelation was found at all scales (see *Figure 9*). The disparity between the observed and expected K statistic was consistently large, though the degree to this disparity fluctuated across spatial scales. This region proved to be most clustered at middle-scales (bands 5 & 6) and less clustered at smaller scales (band 1). Despite this variation, the region remained overwhelmingly clustered.

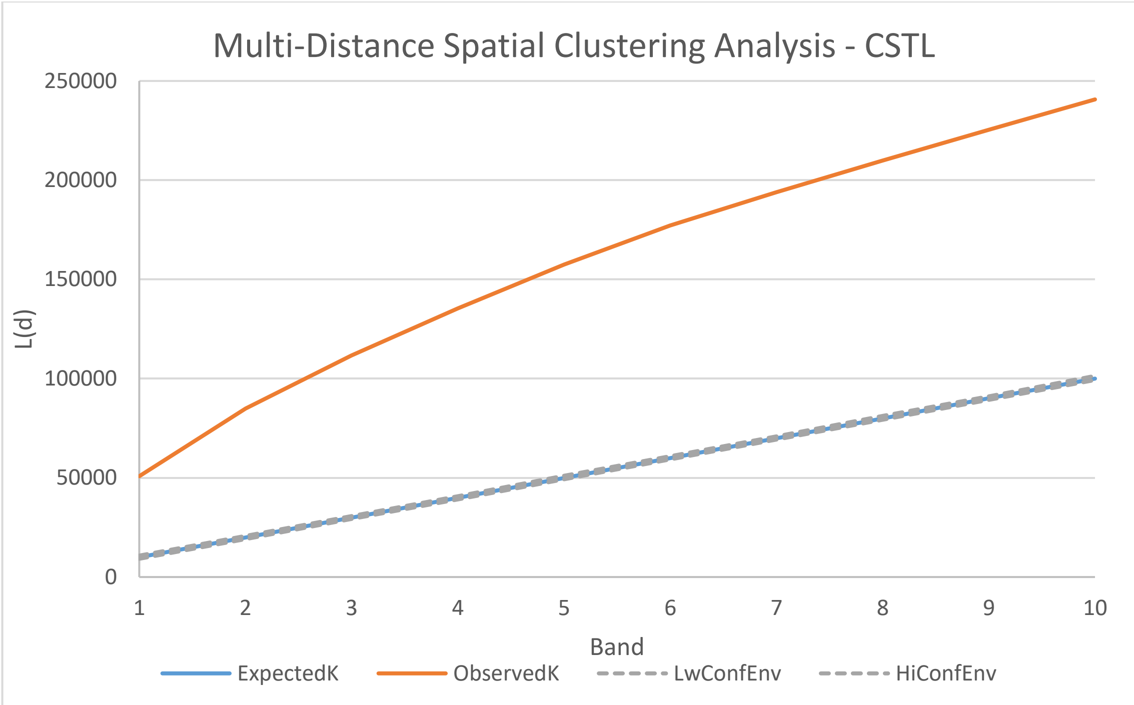

Figure 9: Multi-Distance Spatial Clustering Analysis – CSTL

5.1.6. Everglades Physiographic Region (EVR)

The EVR region exhibited the most consistent results of any of the study regions. The EVR region, like all other testing regions other than the EBF, exhibited positive spatial autocorrelation across all spatial scales (see *Figure 10*). This result, however did not change drastically across spatial scales, and was furthermore less strong than the relationship between expected and observed K values in other regions. This relationship nonetheless defines the wildfire distribution of this region clustered at all tested spatial scales.

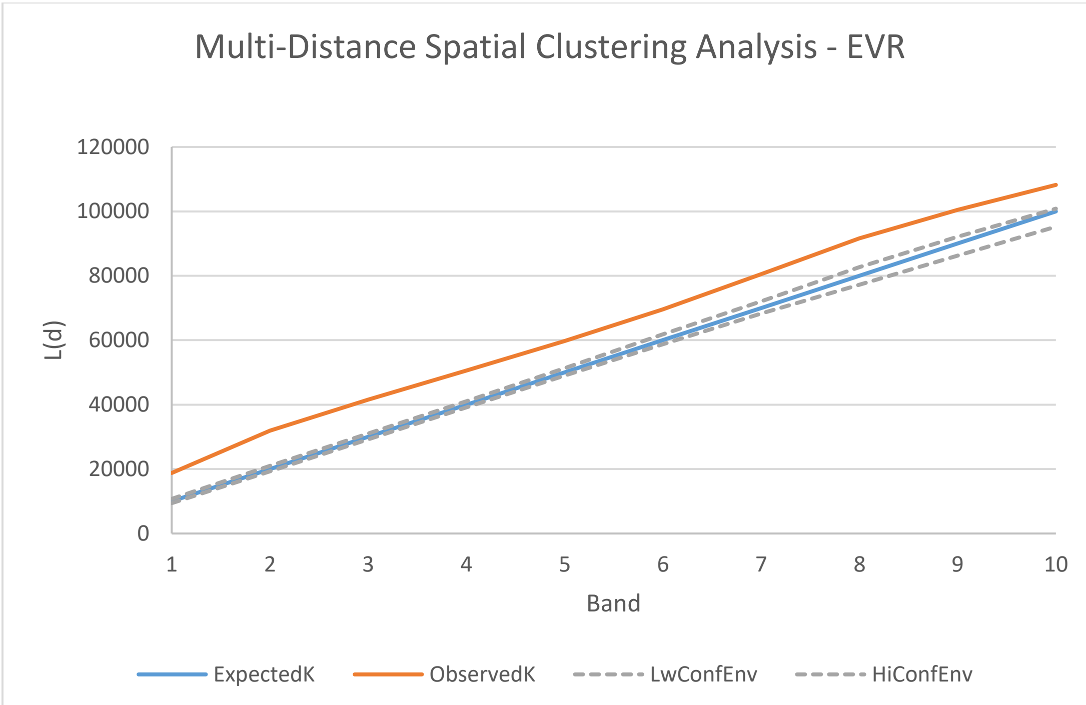

Figure 10: Multi-Distance Spatial Clustering Analysis - EVR

5.2. Objective 2: Logistic Regression & Stepwise Regression

5.2.1. Southeastern Regional Study Area

The Southeastern United States region was evaluated as a whole to best represent how factors influential to wildfire might change with spatial scale. An analysis of the AICs for the different regression models proved the 'backward' stepwise regression of the 'glm' to provide the lowest AIC and, thus, the best fit model. Furthermore, both the 'forward' and 'backward' methods resulted in the same selected covariates. The resulting selected covariates, their significance, and their OR could then be evaluated (*Table 5*). The optimal model built through this analysis utilized 25 of the 36 testing covariates, 24 of which were found to be significant at a p-value of <= .01. Overall, bioclimatic variables offered less predictive power across the southeastern region than land cover, and topographic variables had the lowest ORs of any group. Barren land cover, protected lands, grassland land cover, and shrubland cover had the strongest positive relationships with wildfire presence. Wetland cover, deciduous forest cover and evergreen forest cover also has positive relationships to a lesser degree. Conversely, cultivated land had the strongest relationship with the absence of wildfire in the southeast, with average annual wind and maximum temperature of the warmest month (bio5) also demonstrated strong correlations with the absence of wildfire.

COVARIATE	Z-VALUE	PR(>|Z|)	SIGNIFICANCE	OR
WETLAND	4.44	9.04E-06	***	2.1460358
SHRUB	4.85	1.25E-06	***	2.665786
ROADS	-9.327	<2e-16	***	0.9433885
NAITO	-4.56	5.06E-06	***	0.8956563
GRASSLND	4.36	1.31E-05	***	3.011816
DECID	3.616	0.0003	***	1.8162177
CULTIV	-5.23	1.69E-07	***	0.3328653
EVERGRN	5.16	2.51E-07	***	2.3529292
BARREN	2.622	0.008742	**	5.1816792
RELIEF	10.128	<2e-16	***	1.0132553
PTCD	11.255	<2e-16	***	3.4142484
WIND	-7.04	1.96E-12	***	0.5223172
DEM	-3.305	0.000949	***	0.9977846
ASPECT	-1.455	0.145791		0.9992389
BIO03	5.19	2.16E-07	***	1.2070583

Table 5: SE Covariate Significance and OR
Significance Codes: '*' 0.001, '**' 0.01, '*' 0.05, '.' 0.1**
Average Calculated Pseudo-R^2 = 0.484

COVARIATE	Z-VALUE	PR(>|Z|)	SIGNIFICANCE	OR
BIO04	3.561	0.000369	***	1.0042172
BIO05	-7.13	1.03E-12	***	0.6117248
BIO08	6.79	1.16E-11	***	1.1017459
BIO09	3.815	0.000136	***	1.0503353
BIO13	2.815	0.004876	**	1.0257956
BIO14	-4.77	1.86E-06	***	0.9507472
BIO16	2.964	0.003033	**	1.0124701
BIO17	10.386	<2e-16	***	1.0460073
BIO18	-4.10	4.20E-05	***	0.989777
BIO19	-10.131	<2e-16	***	0.9786264

Table 5 (Continued): SE Covariate Significance and OR
Significance Codes: '*' 0.001, '**' 0.01, '*' 0.05, '.' 0.1**
Average Calculated Pseudo-R^2 = 0.484

The SE region had the greatest variance within covariates of any of the regions studied in this research due to the large spatial scale it incorporates. Despite this, the model for this region still generated significant results for many of the covariates. The result with the highest OR for the SE region was barren land cover. *Figure 11* displays the distribution of barren land cover relative to the 80th percentile concentrations of wildfire in the SE region. There is prevalent barren land cover across the coastal plains, but the highest concentration falls within the APP region, particularly in or near the wildfire cluster it contains. The model for this analysis does not base its results simply on how the presence of wildfire relates to the presence of a covariate, but likewise how the absence relates. Whereas barren land is over represented in areas with wildfire, it is vastly underrepresented in areas with a dearth of wildfire.

Figure 11: Barren Land Cover Presence of the SE Region

Similarly, the distribution of wildfire in the SE region closely matches the distribution of protected lands. *Figure 12* displays the presence of protected lands relative to the 80th percentile of wildfire in this region. The correlation in the presence of wildfire and barren land was strongest in the APP region, but the correlation between protected land and wildfire is strongest in the EVR and CSTL regions, particularly in Florida. These areas represent the highest concentration of wildfire in the region incorporating the upper 20th percentile of concentration. The correlation of wildfire to protected lands also correlates strongly with the absence of roads in these same areas (*Figure 13*). The percentage of APP roads in the fire cluster zone was comparable to proportion in protected lands (7.2% in fire cluster, 6.9% in PTCD land).

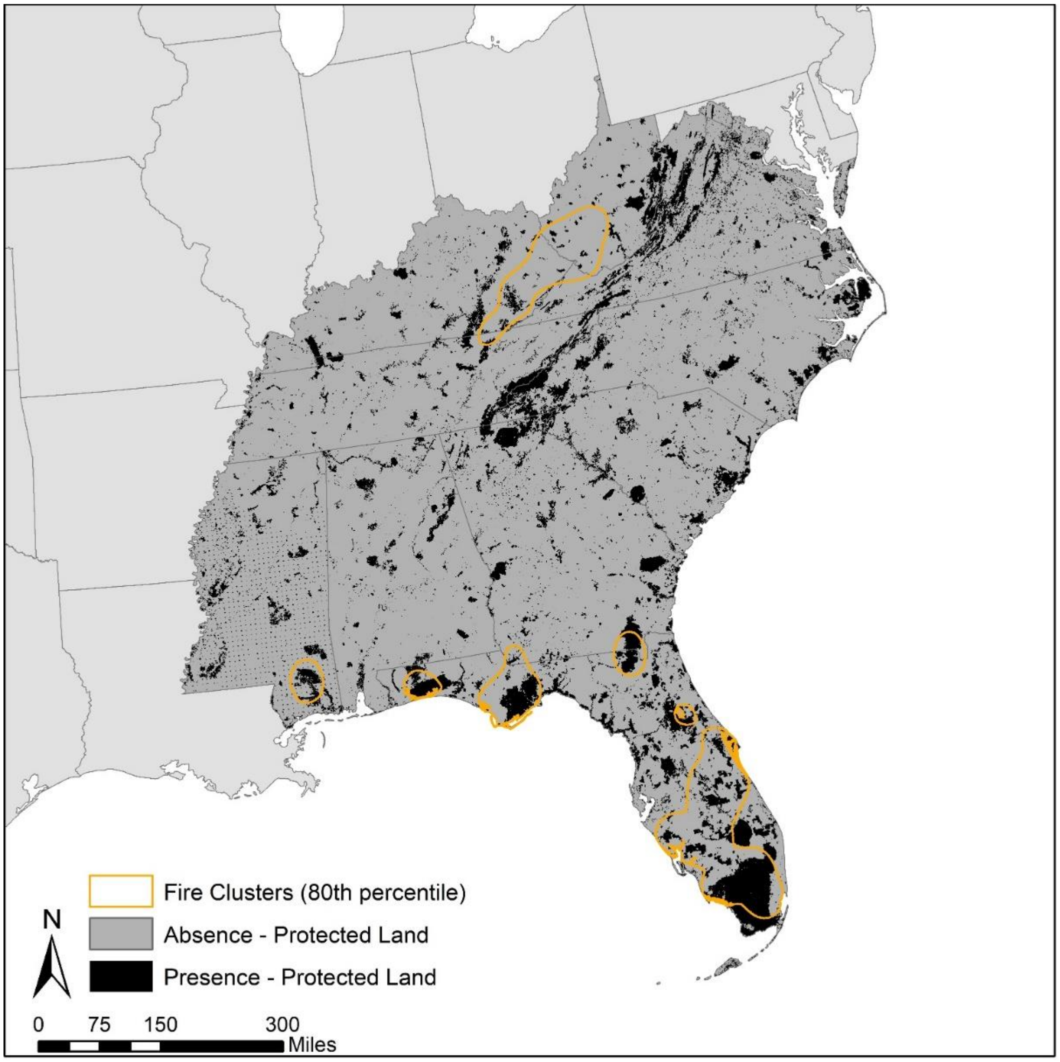

Figure 12: Protected Lands in the SE Region

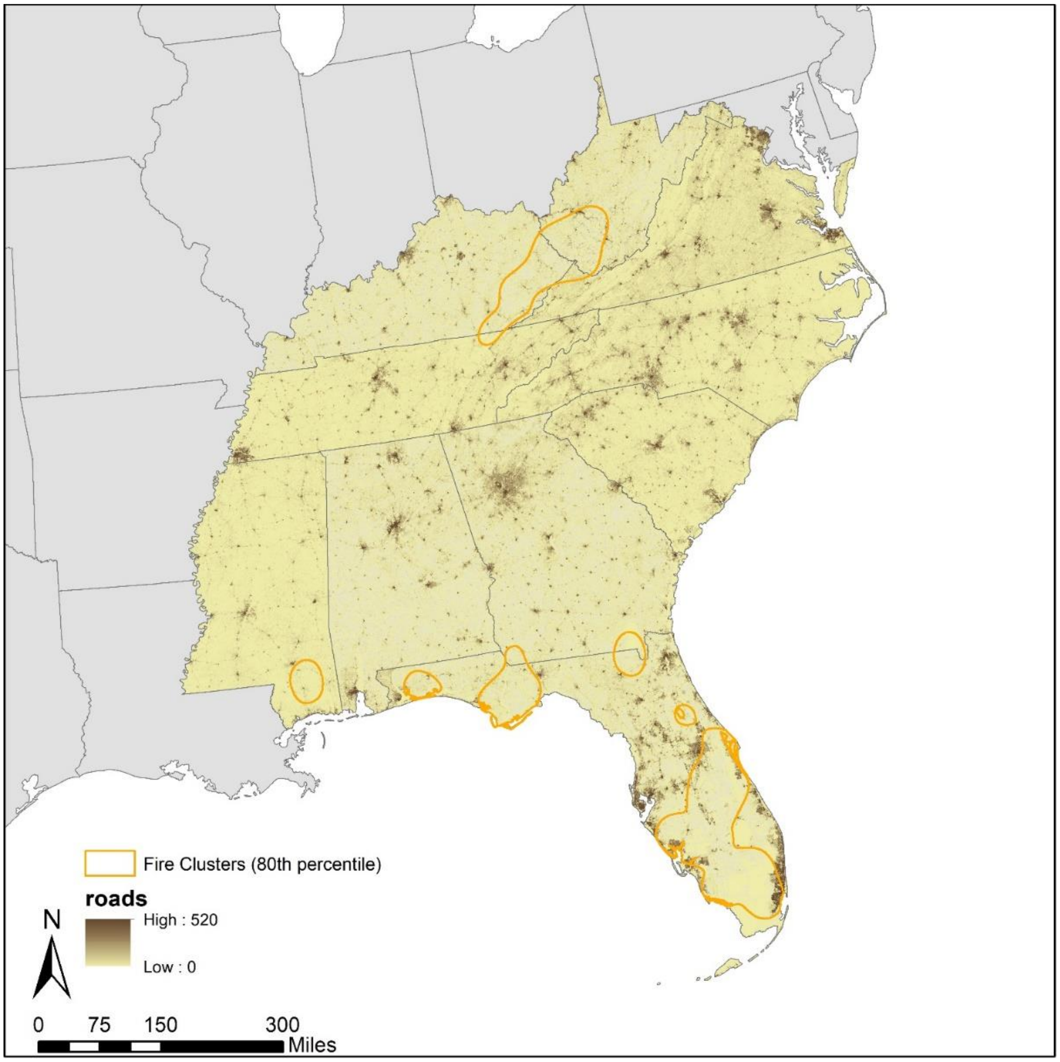

Figure 13: Road Density of the SE Region

This model has not only exhibited strong examples of variables that are influential on the presence of wildfire in the SE region, but has to provided support for valuable predictors in understanding the absence of wildfire. Cultivated land, including farmlands and pasturelands had the strongest relationship with the absence of wildfire in the SE. Wind, another strong indicator of absence was strongest on the coastal edges of the study region and in the highest parts of the Appalachian Mountains.

The final relevant indicator to the absence of wildfire in the SE region was maximum temperature in the warmest month (bio5). Max. temp. in the warmest month is lowest in the Appalachian Mountains and EBF, but exhibited lower values on the coasts of Florida where wildfire was prevalent. *Figure 14* displays the distribution of max. temp. in the warmest month values across the study region in comparison to the prevalence of wildfire. More important for the results of this model, the highest values of max. temp. in the warmest month (>=30 degrees Celsius) were located throughout the EPD region, where the highest number of pseudo-absence points were located.

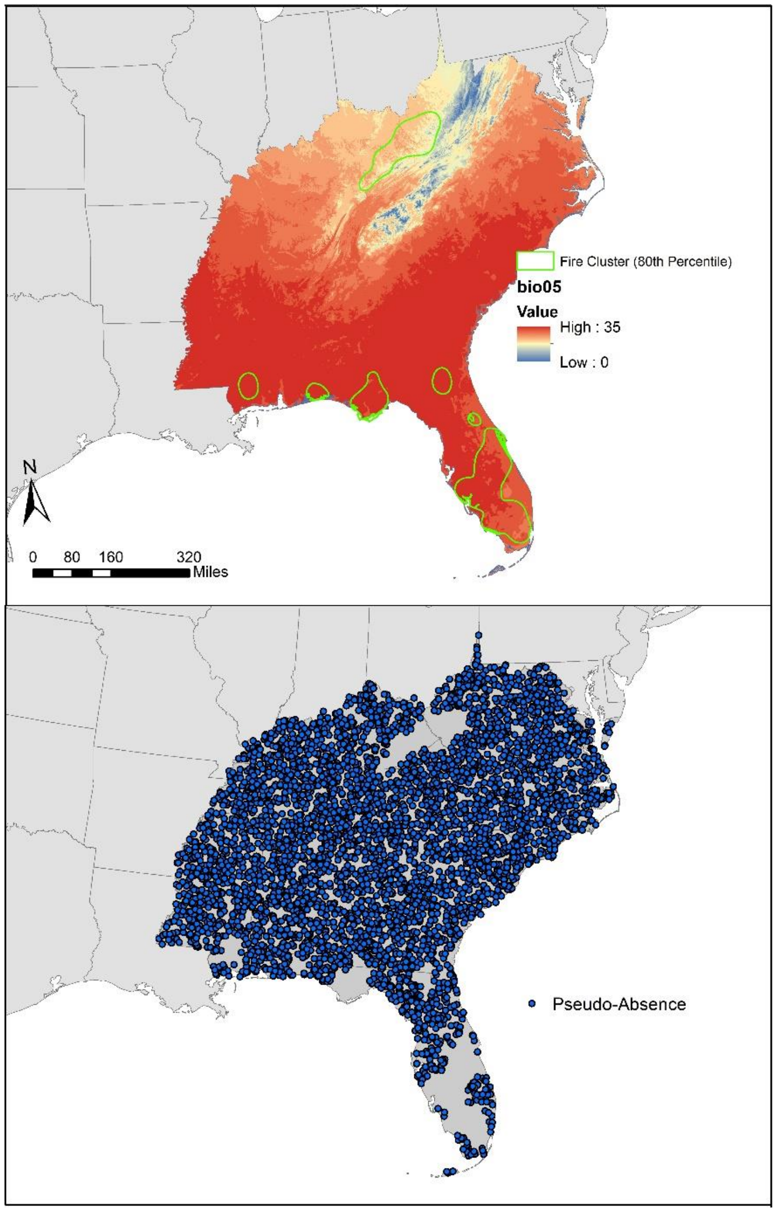

Figure 14: Mean Temperature of Warmest Month and Wildfire Distribution, S.E. U.S.

5.2.2. Eastern Broadleaf Forest Physiographic Sub-region (EBF)

The EBF region had the fewest number of wildfire points of any physiographic region studied. While these numbers were bolstered by the creation of pseudo-absence points, there were still not enough points for any statistical model attempting to explain distribution to maintain significance. As a result, regression results from the EBF region hold no validity. Due to the absence of wildfire and wildfire clusters, however, this region is heavily represented in the pseudo-absence points that were created for the SE regional analysis (see *Figure 14*).

5.2.3. Appalachian Physiographic Region (APP)

Similar to the analysis of the entire southeastern U.S. region, the APP regression model found that the 'backward' stepwise regression generated the model with the lowest AIC and best fit (See table 6). The APP region also utilized a large number of variables in its model formation, 23 of the 36 total, though far fewer were deemed significant. Of the 23 used, only 12 were deemed significant at a p-value <= 0.05 standard. Unlike the analysis with the entire southeast, this sub-regional analysis of the APP region found bioclimatic variables to play a larger role in the presence or absence of fire found across the study area. The strongest positive correlations of the significant variables were protected lands, mean temperature of coldest quarter (bio11), and mean temperature of the driest quarter (bio9). Mean temperature of the warmest quarter (bio10), minimum temperature of the coldest month (bio6), and Isothermality (bio3) offered the strongest relationship with the absence of wildfire in this region. Two of the metrics of human-related activity, road density and population density, were also flagged as significant and had a small correlation with the absence of wildfire in this region.

COVARIATE	Z-VALUE	PR(>|Z|)	SIGNIFICANCE	OR
SHRUB	0.014	0.98911		1.37E+08
ROADS	-1.711	0.08715	.	0.89
POP	-3.093	0.00198	**	0.89
MXDFRST	0.013	0.98982		4.07E+07
GRASSLND	0.014	0.98897		1.76E+08
DECID	0.013	0.98939		8.49E+07
EVERGRN	0.013	0.98993		3.34E+07
BARREN	0.018	0.98585		3.78E+10
PTCD	4.549	5.39E-06	***	18.27
DEM	1.724	0.08473	.	1.01
BIO01	1.323	0.18578		2.52
BIO03	-1.833	0.06681	.	0.47
BIO05	1.186	0.23567		2.53
BIO06	-1.641	0.10084	**	0.38
BIO09	3.289	0.00101	*	1.93
BIO10	-2.029	0.04249	.	0.18
BIO11	1.654	0.0982	*	3.62
BIO12	2.143	0.03208	**	1.06
BIO13	-3.288	0.00101	**	0.74
BIO16	3.06	0.00221		1.10
BIO17	-1.565	0.11749	.	0.93
BIO18	-1.705	0.08817		0.95
BIO19	-1.611	0.10719		0.91

Table 6: APP Covariate Significance and OR
Significance Codes: '*' 0.001, '**' 0.01, '*' 0.05, '.' 0.1**
Average Calculated Pseudo-R^2 = 0.558

Contrary to the results observed in the regional SE study area, presence or absence of wildfire was much more strongly correlated to bioclimatic variables in the APP region. The significant variable with the highest OR, however, was protected lands. Within the SE regional study area, this covariate was easily recognizable within the 80th percentile of wildfire concentration. The APP sub-region offers a less direct correlation (see *Figure 15)*. Wildfire presence in protected lands occurs most commonly outside of the delineated 80th percentile wildfire cluster but in nearly all locations outside of the clustered area wildfire presence falls within the bounds of protected land.

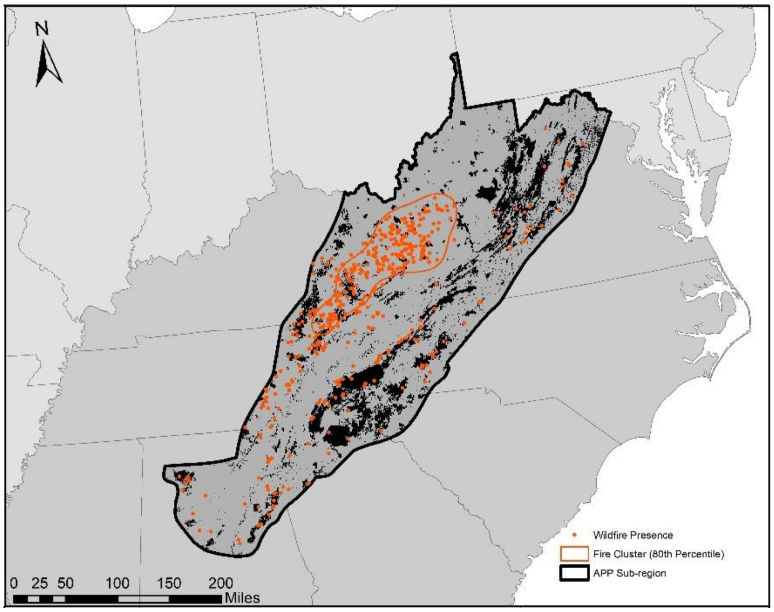

Figure 15: Wildfire Presence and Protected Lands

Both the average temperature of the coldest quarter and the average temperature of the driest quarter also exhibited a strong correlation with the presence of wildfire within the APP sub-region. These two variables are similar in nature and, likewise, exhibit a similar distribution across the study area. Wildfire is almost entirely excluded from areas where these bioclimatic variables are low, and is heavily present in areas with medium to high values (see *Figure 16)*. Testing of the pseudo-absence of this region yielded similar results. Bio10, or Average temperature of the warmest quarter, and Bio6, or minimum temperature of the coldest month, exerted the greatest influence on the absence of wildfire in this sub-region. The strong relationship between wildfire presence and average temperature in the coldest quarter and the strong relationship between wildfire absence and average temperature in the warmest quarter substantiate that the role of uncharacteristically out-of-season temperatures are a strong predictor of wildfire.

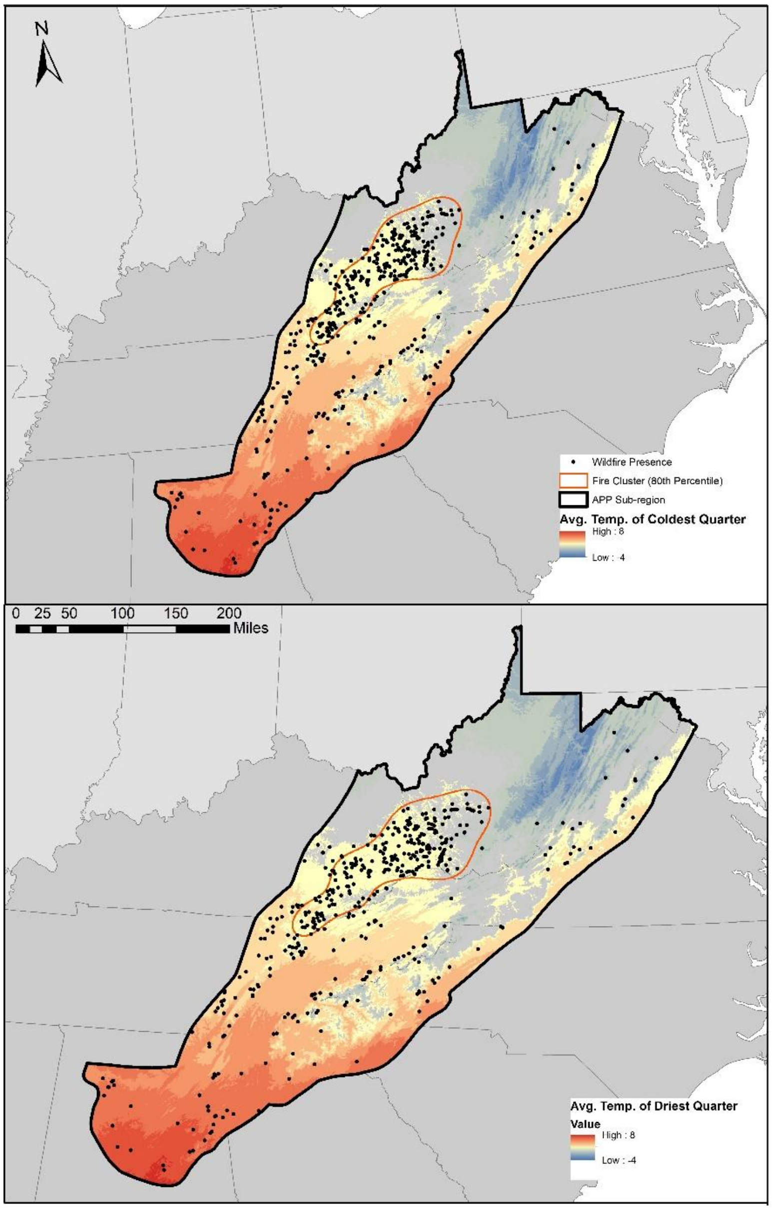

Figure 16: Avg. Temp. of Coldest and Driest Quarters, APP Region

5.2.4. Expanded Piedmont Physiographic Region (EPD)

Like the EBF region, the EPD contained only sparse amounts of wildfire points. The difference, however proved to assist in running regression models. Whereas the EBF region could not be run with any confidence, the EPD region was able to be run, but lacked no statistical relevance in any of the covariates. Few points across large regions do not allow a model to confidently evaluate the variance (or lack of) between points in the region.

5.2.5. Coastal Plain Physiographic Region (CSTL)

The CSTL region offered some of the strongest results of the study using a model with numerous significant covariates and strong ORs (see table 7). This region's model was also constructed using the 'backward' method of stepwise regression resulting in a model that incorporated 22 of the 36 covariate datasets. Of these data, 21 were found to be significant at a p-value <= 0.05 cutoff. Barren land cover in this region offered the highest OR of the study with evergreen land cover, wetland cover, shrub land cover, grassland cover and developed land cover all indicating strong positive relationships with the presence of wildfire in this region. Also indicating positive relationships to wildfire presence at a lesser degree were cultivated land cover, protected land, average annual wind, average annual temperature (bio1), and precipitation seasonality (bio15). Conversely mean temperature of the driest quarter (bio9) and mean temperature of the warmest quarter (bio10) demonstrated strong relationships with the absence of wildfire in this region.

COVARIATE	Z-VALUE	PR(>|Z|)	SIGNIFICANCE	OR
WETLAND	5.012	5.39E-07	***	36.64
SHRUB	4.751	2.02E-06	***	39.07
ROADS	-4.207	2.59E-05	***	0.93
POP	-1.455	0.145599		1.00
GRASSLND	2.573	0.010071	*	12.15
DEVELOP	3.873	0.000107	***	33.52
CULTIV	2.018	0.04359	*	4.73
EVERGRN	5.157	2.51E-07	***	53.89
BARREN	3.864	0.000111	***	1731.41
PTCD	3.243	0.001183	**	2.37
WIND	2.5	0.012419	*	2.13
DEM	-3.118	0.001819	**	0.97
BIO01	2.108	0.035063	*	2.00
BIO02	3.055	0.002254	**	1.93
BIO04	-2.307	0.021035	*	0.99
BIO09	-4.241	2.22E-05	***	0.57
BIO10	-2.487	0.012869	*	0.46
BIO15	6.713	1.91E-11	***	1.91
BIO16	-2.751	0.005948	**	0.96
BIO17	3.509	0.00045	***	1.06
BIO18	-2.79	0.005273	**	0.97
BIO19	2.997	0.002726	**	1.04

Table 7: CSTL Covariate Significance and OR
Significance Codes: '*' 0.001, '**' 0.01, '*' 0.05, '.' 0.1**
Average Calculated Pseudo-R^2 = 0.640

The CSTL region had the largest amount of significant factors of any of the sub-regional study areas. Overall, land cover exerted the most influence on the presence of wildfire in CSTL sub-region and bioclimatic variables exerted the most influence on the absence of wildfire. The CSTL sub-region occupies the largest spatial extent of any of the sub-regions and similarly shares many traits with the SE regional study area. Successional land cover exhibited a similar distribution to that observed in the SE regional analysis. One of the strongest correlations in the CSTL exists with wetland cover. Wetland cover is heavily present in the 80th percentile wildfire clusters and coincides strongly with protected lands in the region (see *Figure 17).* Wetlands present at or near the Okefenokee Swamp, the Apalachicola National Forest, and the feed areas to the Everglades/Big Cypress (Okeechobee region) exhibited the strongest concentrations of wildfire within the entire SE and likewise correspond strongly to wetlands and protected lands in the CSTL sub-region.

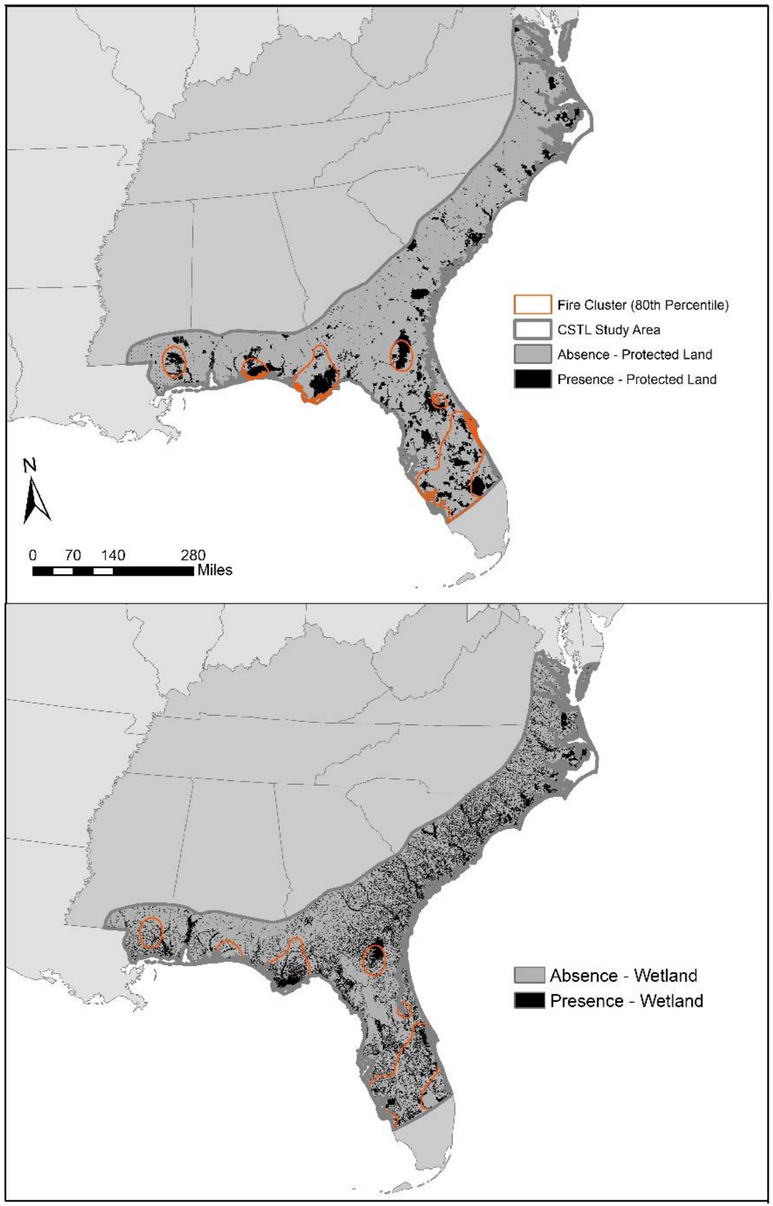

Figure 17: Protected Lands and Wetlands of CSTL Sub-region

Adding to the fire-susceptibility of this area are two additional bioclimatic factors flagged as significantly influential on wildfire: bio1 and bio15. Both bio1 (average annual temperature) and bio15 (precipitation seasonality) follow a standard latitudinal gradient, increasing in values as they move southward (see *figure 18)*. The highest values of these factors, thus, occurs in the southernmost part of the sub-region. Likewise, wildfire is most strongly concentrated in the southern region, more particularly in the panhandle and greater peninsula of Florida.

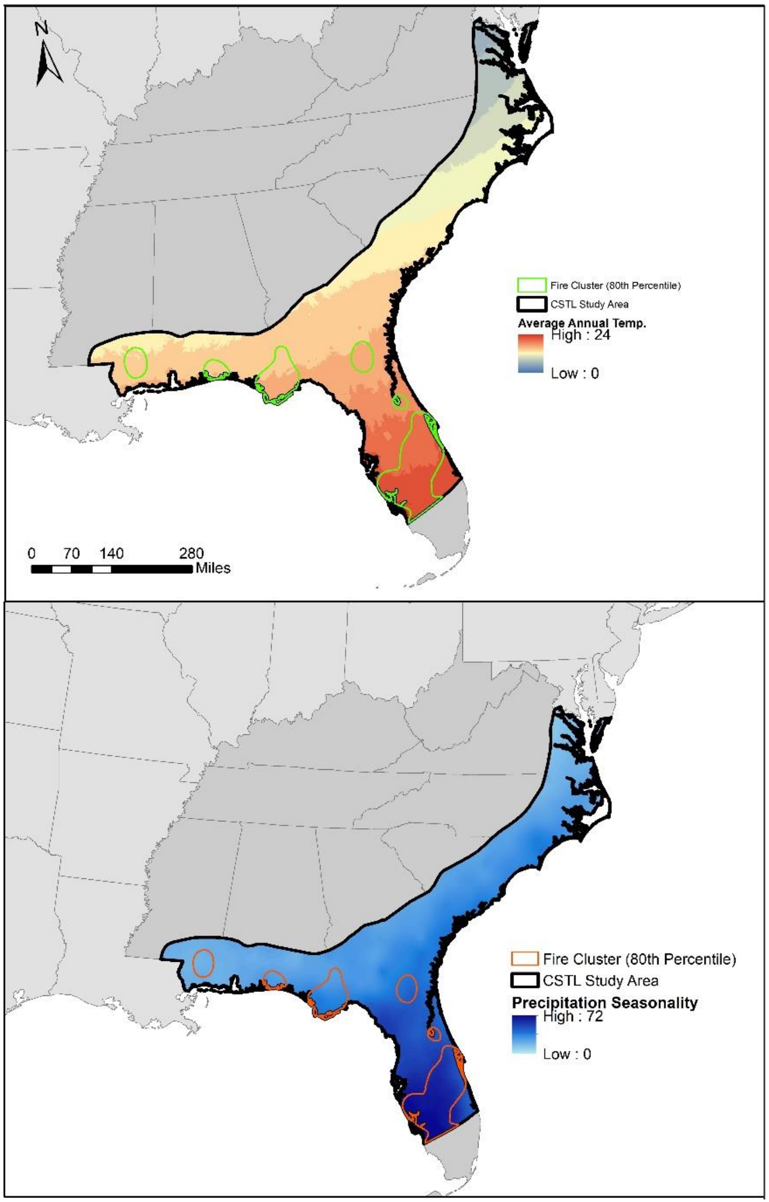

Figure 18: Avg. Temp. and Precip. Seasonality of CSTL Sub-region

Bioclimatic variables were also found to be correlated to the absence of wildfire within the study area. The calculated average temperature of the driest quarter (bio9) and average temperature of the warmest quarter (bio10) offered the strongest correlation with the absence of wildfire (see *Figure 19)*. High values of bio09 fell consistently out of calculated wildfire concentration. While a strong indicator of drought conditions, bio9 was typically lower in areas with wetlands: an area associated with the presence of wildfire. With little variation amongst the entire sub-region, bio10 was likely flagged due to larger ratio of pseudo-absence to presence points in the area. The lack of value in this predictor is further substantiated by the low significance of this factor in the model.

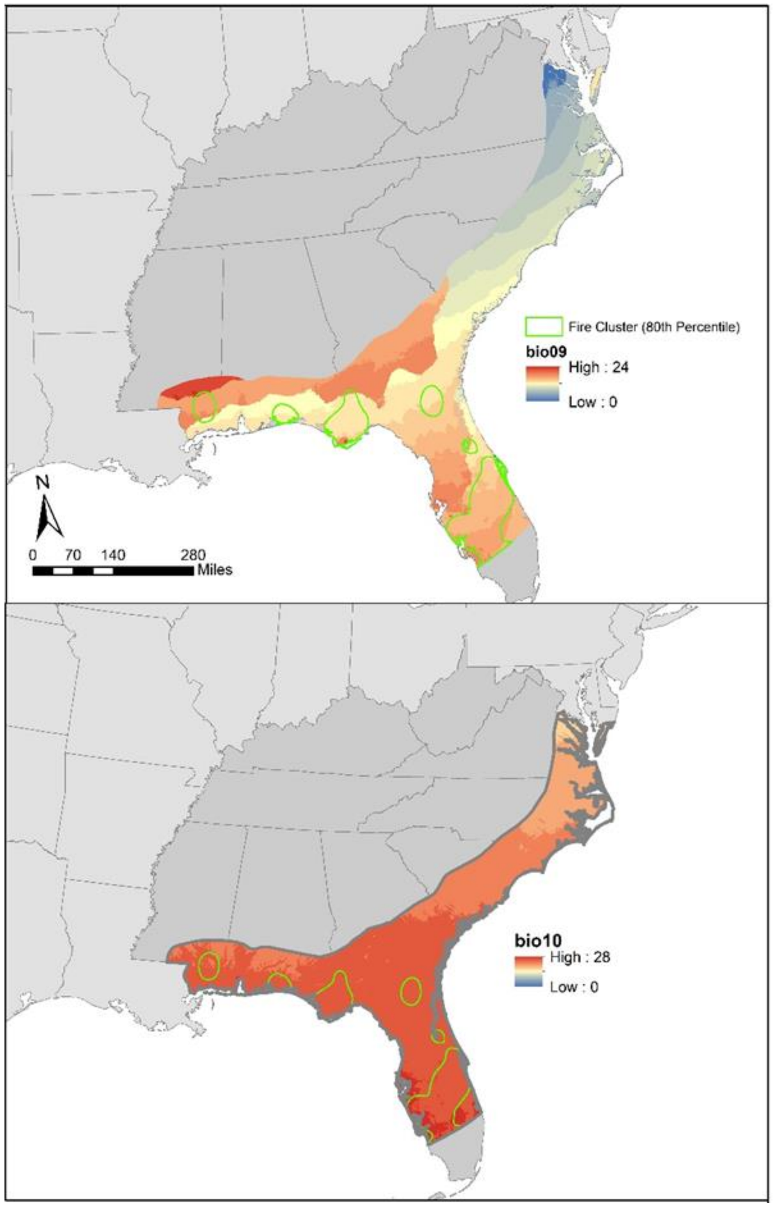

Figure 19: Avg. temp of driest (bio09) and wettest (bio10) quarters, CSTL Sub-region

5.2.6. Everglades Physiographic Region (EVR)

The optimized model for this region was created through the 'backward' stepwise regression but only utilized 10 of the 36 covariate datasets (See table 8). Of these, only 8 datasets were deemed statistically significant at a p-value <= 0.05 cutoff. Whereas the CSTL model contained the highest OR values of any of the regions tested, the EVR model demonstrated the lowest OR values of any of the study regions. Predicted mean fire interval (naito), cultivated land cover, and mean annual temperature (bio1) had the strongest relationships with the absence of wildfire in this region each with an OR lower than .01. Conversely, elevation (dem), and precipitation of the wettest quarter (bio16) had the strongest relationship with the presence of wildfire. Temperature seasonality (bio4) was also correlated to the presence of wildfire, to a lesser degree.

COVARIATE	Z-VALUE	PR(>|Z|)	SIGNIFICANCE	OR
NAITO	-3.159	0.001581	**	0.003
CULTIV	-2.189	0.028599	*	0.00002
WIND	-1.616	0.106102		0.085
DEM	1.962	0.049764	*	5.56
BIO01	-2.98	0.002879	**	0.0009
BIO04	3.357	0.000787	***	1.41
BIO06	1.463	0.143468		0.001
BIO10	-2.935	0.003332	**	1.03
BIO12	2.238	0.025229	*	1.39
BIO16	4.146	3.38E-05	***	7.81

Table 8: EVR Covariate Significance and OR
Significance Codes: '*' 0.001, '**' 0.01, '*' 0.05, '.' 0.1**
Average Calculated Pseudo-R^2 = 0.860

The EVR region had both the smallest spatial extent and the least number of presence and pseudo-absence points of any region a model could be made for. As a result, the EVR exhibited the fewest significant factors of any of the sub-regions with results. Despite this, it had the highest pseudo-R^2 of any model. Elevation, as represented by a digital elevation model (DEM) was denoted as one of the most significantly correlated factors with the presence of wildfire in the EVR sub-region. Elevation followed a roughly latitudinal gradient, decreasing as it moved south. Wildfire presence was almost exclusively found in areas of higher elevation whereas generated pseudo-absence was found across all gradients of elevation in the region (see *Figure 20)*.

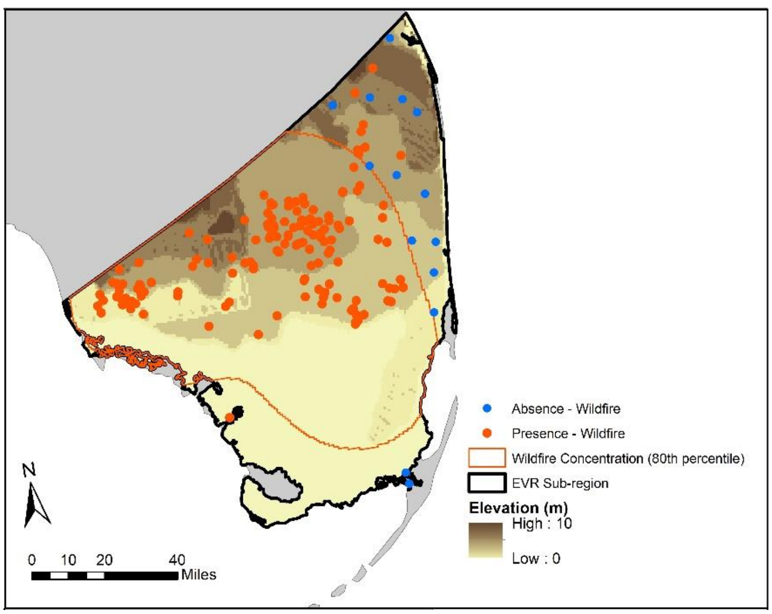

Figure 20: Elevation and Wildfire in the EVR sub-region

Other than elevation, precipitation of the wettest quarter (bio16) and temperature seasonality (bio4) were found to be significantly correlated with the presence of wildfire (see *Figure 21)*. These values were stronger in the internal section of the study area, particularly in the Everglades and the Big Cypress National preserver. This relationship closely overlaps with the distribution of both wetlands and protected lands in the sub-region.

Cultivated lands was had the strongest relationship with the absence of wildfire on the landscapes, largely occupying the northern edge of the study region but occurring exclusively in areas without the presence of wildfire. To a lesser degree, mean annual temperature (bio1) and mean fire interval (naito) were significant correlated to the absence of wildfire, but exhibited little variation across the region. Though wildfire was found almost exclusively within the lowest value groups of these variables, there is not enough variability across the sub-region to conclude that these variables exert any real influence on either the presence or the absence of wildfire.

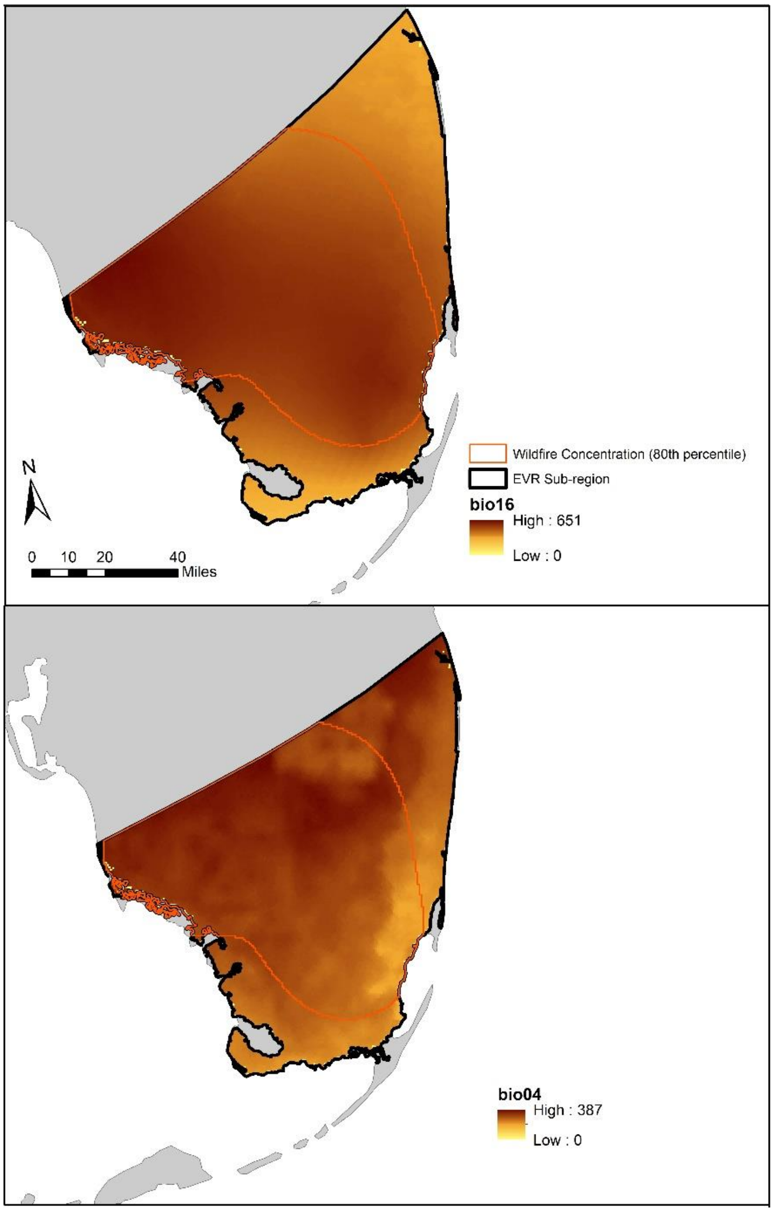

Figure 21: Precipitation of the Wettest Quarter (bio16) and Temperature Seasonality (bio4) of the EVR Sub-region

5.3. Regional/Sub-regional Trends

Analysis of sub-regional variations provides insight on how relevant covariates change amongst different physiographic regions, but in order to generate a list of predictive criteria that begins to encompass the greater regional area, trends among sub-regions must be analyzed. Trends were measured by the number of sub-regions the variable was significant (out of the three results were gained from), if the trend of the OR was toward presence, absence or split amongst sub-regions and if the sub-regional analysis was consistent with the regional analysis of the SE (see table 9).

VARIABLE	SUB-REGIONS SIGNIFICANT	OR TREND	CONSISTENT WITH SE ANALYSIS
DEM	3	SPLIT	
BIO10	3	ABSENCE	NO
ROADS	2	ABSENCE	YES
CULTIV	2	SPLIT	
PTCD	2	PRESENCE	YES
BIO1	2	SPLIT	
BIO4	2	SPLIT	
BIO9	2	SPLIT	
BIO12	2	PRSENCE	NO
BIO16	2	SPLIT	
BIO19	2	SPLIT	

Table 9: Regional/Sub-Regional Trends

6. DISCUSSION

The difference in scales across the study regions for this research is essential to building an understanding of the relationship between contributing wildfire factors and scale. Equally important, however, was conducting analyses across sub-regions of similar scales. Both the similarities and differences in predictive factors across these regions provide insight on the role of physiography on wildfire. The base covariates tested in this study will help to predict and understand the distribution of wildfire in the southeastern U.S., but the insight gained from these analyses will better scientific understanding of wildfire across the world. This study revealed significant factors related to the concentration of wildfire, building a foundation for other similar, regional analyses.

6.1. Southeastern Regional Study Area

Barren land was found to be one of the strongest predictors of wildfire across the SE region, but the relationship with wildfire may not be so direct. Literature such as Kalabokidis et al. 2002 dismiss the use of barren land as a relevant covariate in predicting the prevalence of wildfire as it does not support the necessary fuel load to sustain large fires. Furthermore, barren land does not have the fuel load needed to support the type of wildfires analyzed in this study (>500 acres burned). Therefore, wildfire is not the result of barren land on the landscape, but the same cannot be said for the reverse. Wildfire instigates land cover change and, for some portion of time, barren land is a part of the successional cycle. It is subtle, though relevant that the presence of barren land is not explicitly a predictor of wildfire, but the continued and concentrated presence of barren land is a strong implicit indicator of land cover change, of which wildfire is a common perpetrator. As land cover for this study was derived from satellite imagery, this research did not account for temporal trends in barren land. These areas could be strongly related to successional cycles, or in many cases permanent areas of barren land, such as mountain tops.

Protected lands offer large, unfragmented sections of natural vegetation that is highly susceptible to the ignition and spread of wildfire. Human-interaction with protected land is one of the most prevalent causes of wildfire (Balch et al. 2017) and protected lands are some of the most vulnerable areas. The relationship between wildfire and protected lands is, therefore, heavily explained by their very nature. Furthermore, protected lands are not only ripe for the spread of wildlife, but wildfire management policies may also play a significant role in the presence of wildfire in these areas. Wildfire practices such as fire breaks and 'let it burn' policies often control the spread of wildfire without direct suppression. These practices make it easier for wildfires to reach the size required for consideration in this study. Despite this, there are areas with large amounts of protected lands that do not fall within the bounds of heavy fire concentration. This variation may be further explained by fire suppression tactics. Whereas protected lands in remote areas may be allowed to burn, proximity to infrastructure human development may serve to reduce the amount of wildfire, regardless if nearby lands are protected or not. Lafon et al. 2017 details how in the Appalachian region, suppression tactics have worked to reduce ignitions and keep wildfires small, but these same tactics may be less effective in protected land in areas with less resources or less desire to control wildfire. Important considerations in examining the ability to adequately suppress wildfire are derived from both the accessibility and vulnerability of the areas. Wildfires that spread in the northeast APP sub-region are less accessible and the surrounding areas less vulnerable to the loss of human development. Road density may, therefore, be a useful tool in the evaluation of protected lands as a fire-prediction metric. Prior research exploring the cluster of wildfires in the APP region corroborates the relationship between roads and the absence of wildfire (Maingi & Henry 2007). Despite this utility, the absence of roads in rural, but not fire-prone areas inhibits the use of road density as a powerful predictor for wildfire in the SE region. As a powerful predictor of wildfire presence in the SE region, protected land is still limited by how human-interaction responds to it.

Furthermore, land cover related to stages of succession, such as grassland cover and shrub land cover proved to be relevant covariates. Each of these represent further stages in successional cycles prevalent in this region and of which fire is a classic disturbance (Clements 1916). Each of these variables were flagged as significant predictor of wildfire which aligns well with the prevalence of wildfire as a disturbance in these same ecological systems. Whether wildfire clears the way for the expansion of grassland and later shrubland, or in the development of pyrophilic evergreens, wildfire is both essential and common in natural S.E. U.S. ecosystems.

Variables related to the absence of wildfire offer differing interpretations. Cultivated land cover was a variable strongly correlated to the absence of wildfire. This closely corroborates the results of other studies of a similar nature (Nunes et al. 2005, Moreira et al. 2009). Fire suppression and the high moisture content of cultivated areas remain important factors in controlling the spread of wildfire. Though these areas generally contain an unfragmented fuel load, their value makes them strong targets for fire suppression. The absence of wildfire in cultivated land cover is most likely the result of human-intervention.

Another indicator of absence, high winds, may be expected to be an indicator of large fires. High winds can increase the size and spread rate of wildfires (Beer 1991), but this is not the correlation observed in the SE region. This disparity is likely more due to the geography of the locations with high wind as opposed to any inherent relationship between wind and wildfire absence. Coastal areas have higher precipitation rates and populations than areas inland. Both of these factors have a negative influence on the occurrence of wildfire. Furthermore, the contributing dataset only separated wind speed into 6 classes, limiting the specific wind trends that can be observed.

Temperature exerts influence on the distribution of wildfire by controlling drought conditions and vegetation growth (as well as successional regrowth). As a result, the converse of the relationship observed between mean temperature of the warmest month (bio5) and wildfire absence might be expected. Like the defined relationship between wildfire absence and wind, proximity to the coast may be an important controlling factor. Without coastal winds cooling the peninsula of Florida and the CSTL region, max. temp. in the warmest month values would likely follow a more linear latitudinal gradient. Under this distribution, there would be little to no relationship between wildfire and the max. temp. in the warmest month in this region.

In a much broader sense, the SE regional model offers interesting results as to which covariate types are the most relevant. In the SE region, land cover was typically flagged as variable of increased importance in the prediction of both the presence and absence of wildfire. Bioclimatic variables, on the other hand, were largely insignificant in this same regional prediction. Bioclimatic variables typically associated with wildfire, such as high temperatures or low precipitation, have shown little influence on the regional presence of wildfire. The highest clusters of wildfire in the SE region are found in the peninsula of Florida and in the APP region. The latitudinal differences in these locations accounts for vastly different bioclimatic conditions and proximity to the coast furthers the disparity. At this scale, bioclimatic conditions cannot fully account for the presence and absence of fire as the variability between fires in the mountains and fires in the everglades is too great. The strong correlation of wildfire to land cover is a strong indication that, at least in this region, the variability amongst bioclimatic variables is not strong enough to overcome the superior influence of other covariates. These results, therefore, substantiate the idea that land cover may be a more important predictor for the presence of wildfire than bioclimatic or topographic variables at large scales. As vulnerable land cover interacts with ignition points, the presence of wildfire will continue to be observed.

The SE model does not simply help to explain the significant influencers at this regional scale, but also works to help explain discrepancies at sub-regional scales. The

EBF and EPD regions did not generate useful results from the attempted logistic regression models, but these areas were nonetheless accounted for in the SE regional model. The EBF and EPD regions had very few wildfires and none of the fire-exclusion zones used to limit the creation of pseudo-absence. As a result, the pseudo-absence strongly reflects the lack of wildfire in these regions. While this limits the insight that can be gained from the presence of wildfire in these regions, the results most influential on the absence of wildfire in the SE region are most reflected in the EBF and EPD regions.

6.2. Sub-regional Analyses

In regions where results could be calculated, a dichotomy of relevant results must be interpreted. The standard interpretation of sub-regional analyses focuses on the relevant covariates in each individual sub-region's presence or absence of wildfire. Equally important is which variables are commonly significant in either the presence or absence of wildfire across numerous sub-regions. The trends observed across study regions and scales illustrate the strength or weakness of each variable. This two-pronged interpretation of the results will not only provide insight on which variables factor most heavily into specific physiographic regions, like the EVR, but also what commonality exists between vastly different regions such as the EVR and APP.

6.2.1. APP Physiographic Sub-region

The APP region offered significant results as to which covariates were the most influential in the most mountainous sub-region of this research. It, too, had the largest number of insignificant variables included in a model. This result illustrated one of the potential issues with the use of stepwise regression. Stepwise regression utilizes the variables necessary to generate a best fit model, not necessarily with significant points. This sub-region serves as a strong example of the line that must be walked between model fit and variable significance. The results from this sub-region were based largely on the substantial concentration of points found in the Appalachian Plateau, but were likewise influenced by the presence of wildfire observed across the rest of the region, generally in protected lands. The strong relationship found between protected lands and wildfire indicate that protected lands are the most susceptible to wildfire in the APP sub-region. A discrepancy exists in the large amount of wildfires in the Appalachian Plateau that are not in protected lands. This area correlates strongly to absence of roads which could stand to be a potential explanation for this concentrations (Maingi & Henry 2007). As this study only analyzed wildfires above 500 acres in size, the link between rural land access and fire suppression in national and state parks could play a major role in preventing the clustering of wildfire distribution in parts of the sub-region with large expanses of protected land. Suppression in this region is nothing new (Lafon et al. 2017) and without suppression and control tactics used in many of the protected lands of this sub-region, the concentrated region of wildfire presence may expanded to include a large amount of protected land. Though it is difficult to conclude any one explanation for the presence of wildfire in this region, the link between fire-suppression, rural-land access, and land designation appears to be a significant influencer on the distribution of wildfire in this sub-region.

Despite the relationship between wildfire and road density, the significant concentration of wildfire in the northwestern section of the APP sub-region remains largely unexplained. This area, the Appalachian/Kentucky coalfields, contains the strongest concentration of wildfires in the region, without matching many of the significant characteristics of the APP, such as protected lands. Speculation has been made that this relationship may exist due to human land-use in the region (Maingi & Henry 2007), but could be related to the geology and vegetation of the region as well.

Average temperature of the coldest quarter is another variable that was correlated with the presence of wildfire. This variable, a strong indicator of warm winter seasons, coincided with average temperature of the driest quarter, a strong indicator of the drought conditions of an area. As these variables increase, the odds of wildfire presence increase, thereby demonstrating a correlation between the two. The distribution of wildfire relative to the values of these factors across the study area indicate that while high values of these two factors is an indicator of the presence of wildfire, the converse is equally true. Areas experiencing low values of these factors exhibited little to no wildfire. These factors were highest outside of the mountainous areas of the sub-region supporting the idea that any role that topographic factors have in the presence and spread of wildfire may be offset by the changes in the bioclimate associated with topographic differences.

Warm winters exert a higher influence on the presence of wildfire than warm summers or cold winters (as exhibited in the significance of bio6). The role of mild bioclimatic conditions on the distribution of wildfire is further substantiated by the relationship between isothermality (bio3) and the *absence* of wildfire. Isothermality is a ratio of monthly temperature variation to annual temperature variation, making it an indication of the strength of seasonality. High isothermality is an indication of mild seasons, as monthly temperature variation is comparable to annual temperature variation. The APP model correlates high isothermality to the absence of wildfire providing further evidence that the absence of wildfire is correlated to mild seasonal variation. The final variable flagged in the absence of wildfire was bio13, a measure of the precipitation in the wettest month. Unsurprisingly, wetter areas were more indicative of the absence of wildfire.

Overall, this study of the APP sub-region provides insight on the role of bioclimatic variables on wildfire distribution. Potential drought conditions exerted a heavy influence on wildfire occurrence, as substantiated by the positive correlation between presence and avg. temp. of the driest quarter (bio9) and absence with precip of the wettest month (bio13). Furthermore, uncharacteristically warm winters were highly influential on wildfire presence, while warm summers, cold winters and high isothermality correlated strongly with absence.

6.2.2. CSTL Physiographic Sub-region

Given the large spatial extent of this sub-region, barren, evergreen and grassland land cover are likely significantly correlated to the presence of wildfire for similar reasons as their relevance at the SE regional scale. These different land cover types paint a picture as to the complex and often essential role of wildfire in successional cycles and unsurprisingly, this role is reflected on the correlation between them and the presence of wildfire. The interdependent role of wetlands, wind, and protected land, however, reflect a different phenomenon. Wind, as previously identified, is strongest along the coasts and peninsula of Florida and weakest on the internal regions of the SE. The distribution of strong winds across the CSTL region correlates well to the presence of wildfire. Hotspots of wildfire in the CSTL region were found in the panhandle and peninsula of Florida where wind was greatest. Strong winds can aid the spread of fire, but this correlation was likely aided by the presence of overlapped wetland and protected lands. Previous studies have observed that increases in wildfire corresponded to the draining of traditionally wet soils (Bacchus 2000, Harden et al. 2003). The drying of large fuel loads leaves wetlands at high risk of ignition. Florida and southern Georgia are home to the greatest concentration of wetlands in the SE region. Under the wildfire observation period of this study (1984 -2015), the aquifers that underlie the wetlands in the CSTL region experiences the heaviest toll of water-use (Marella 2014). Adding to the fire susceptibility of these lands, most major wetlands in the region are under either federal or state protection. As previously noted, protected lands often face challenges to fire-suppression agencies. The combination of high winds, protected status, and wetlands facing both anthropogenic and natural drought conditions help to define the southern CSTL sub-region as a wildfire hotspot. Though not often thought of as being ideal land cover for wildfire, the strong correlation and possible interpretation of this relationship indicate that wetlands, particularly those facing draining, represent a potentially important predictor of wildfire.

Further working to exacerbate this relationship, this model noted a correlation between precipitation seasonality and the presence of wildfire. High precipitation seasonality is an indication of high disparities between dry and wet seasons. While heavy precipitation in wet seasons helps to build the fuel load, strong dry seasons contribute to the draining of these lands and the building of potential wildfire fuel loads (Harden et al. 2003). Longer dry seasons, combined with high average annual temperatures work to further drought conditions that aid in the presence of wildfire. The correlation exhibited between these factors and wildlife help to define how the distribution of wildfire became so clustered in the southern wetlands of the CSTL sub-region.

Measures of absence for wildfire in the CSTL sub-region were significantly less insightful than measures of absence. Average temperature of the driest quarter (bio9) was one variable that correlated to the absence of wildfire, where the reverse might be expected. Drought, as previously identified, plays an integral role in the prediction of the presence of wildfire, but this measure of high average temperatures in the dry seasons offers contradictory results. As was the case in the SE regional analysis, this distribution could be more the result of coastal/internal land differences. The interdependence of many of bioclimatic variables with physical climatic processes such as wind may be the reason for contradictory results. Likewise, mean temperature of warmest quarter (bio10) offers further contradictory results. This discrepancy has potentially a more obvious explanation. There exists little variation within this dataset for the CSTL sub-region, so even subtle differences between the values of presence and absence points are magnified.

Similar to the analysis of previous models, the CSTL region also indicates the weight that land cover exerts on the presence of wildfire is superior to that of bioclimatic or topographic variables. While human-influence was not directly correlated to the either the presence or absence of wildfire, human land-use and human-impact on land cover may play a significant role in the relationships observed by this model.

6.2.3. EVR Physiographic Sub-region

The strongest predictor of wildfire in this region was precipitation of the wettest quarter (bio16). Strong precipitation does not inherently have a relationship with the presence of wildfire, but is essential in maintaining the primary productivity of wetlands. Precipitation in the wettest quarter (bio16) alone may seem counterintuitive to the presence of wildfire on the landscape, but heavy precipitation plays greatly into the creation and sustainability of the wetlands that make up much of the concentrated wildfire area in this sub-region. The picture further clears when temperature seasonality is considered. High temperature seasonality is an indication of high variability in temperature, further indicating these wetlands experience high summer temperatures that can contributed to the draining and ignition of this area.

This relationship is further explained with the addition of another significant variable: elevation. As previously discussed, wetlands that experience drainage or drought are at the highest risk of wildfire presence. The lowest elevation areas of the everglades are coincidentally the most saturated areas, encompassing the Everglades and the Big Cypress Swamp (Davis et al. 1994). These major wetlands also define the east/west border of wildfire presence. Presence of wildfire was found universally across the northern bounds of the sub-region, excluding the eastern coast. Wetlands, and more appropriately, wetland soils are consistent across the calculated wildfire concentration, with sandy or limestone-based soils more prevalent on the eastern coast (Davis et al. 1994, Lockwood et al. 2003). Both wetlands and protected lands are likely still strong indicators of wildfire in the region, but these factors encompass nearly the entire sub-region. As a result, they account for both the presence and pseudo-absence of wildfire in the region, making them statistically insignificant. A finer-scale analysis of the soils in this sub-region, particularly discriminating wetland soils may offer strong predictors to the presence or absence of wildfire in the EVR and similar regions. Despite this, the relationship between strong precipitation and elevation help to solidify that wetlands that experience draining, as would be more prevalent in higher elevation areas of wetlands, are at increased risk of wildfire.

Like other sub-regions, cultivated land was one of the strongest predictors of the absence of wildfire. The EVR region is home to large expanses of agriculture, producing sugarcane, vegetables, sod, and rice (Davis et al. 1994). Farmers in these areas have a vested interest in fire suppression and likely contribute to the strong absence of large wildfire in cultivated areas of the EVR sub-region. Other strongly correlated variables to absence of wildfire in the sub-region offered little interpretability.

All in all, the fire regime of the EVR region is largely defined by the prevalence of the protected wetlands and the factors that contribute to its creation and stability. The interrelationship between wetlands, elevation and the climatic conditions that support them define the presence of wildfire. Large, unfragmented swaths of wetlands under state or federal protection offer the ideal wildfire 'habitat' under the right conditions. Though this study area was likely too small to evaluate factors with little variance, a smaller-scale analysis of both wildlife and contributing variables for this small sub-region would benefit the claims asserted by this research.

6.3. Regional/Sub-regional Trends

The APP, CSTL and EVR sub-regions offer vastly different ecologies, climates, and topography making fitting one set of predictive criteria to all three regions extremely difficult. The differences between different sub-regions and between the sub-regions and the SE regional study area highlight the importance of scale and study area consideration in research utilizing predictive modeling. Despite these differences, existing commonalities are a strong indication of robust predictive variables. Of the 11 variables consistent among sub-regions, only 4 were found to have a consistent interpretation. The extremely weak correlation between annual precipitation (bio12) and the presence of wildfire (average OR = 1.04) offers little interpretation. Protected lands, however, reigns as the most prominent covariate related to the presence of wildfire. Its strong OR (average = 8.02) and consistency with SE results indicated that this correlation is universally strongly correlated with wildfire. This relationship is enhanced by the interpretation of the EVR region; though not statistically significant, protected lands encompassed nearly all wildfire presence in the sub-region. Protected lands are not physiography-dependent. As a human construct, their presence or absence is the result of land use and ownership making protected lands universal across the entire study region. While protected land's relationship with wildfire is heavily influenced by human-activity (both through ignitions and suppressions), it exists as the strongest predictor of wildfire in this study area.

Just as important as universal predictors of wildfire presence are universal predictors of wildfire absence on the landscape. Average temperature of the warmest quarter (bio10) was a predictor of wildfire across all 3 sub-regions that generated results. Despite this, the variable was not found significant among the SE regional analysis. Avg. temp of the warmest quarter (bio10) offered little variation (average range = 8.67) amongst values in the study region making it expansive at the sub-regional scale. At the regional-scale and full range (range = 29) of its values avg. temp of the warmest quarter (bio10) loses significance. Though not found to be significant in the EVR region, roads were found to be consistently correlated with the absence of wildfire across the APP, CSTL and SE study areas. The roads variable closely mirrors the protected lands variable. Both were not found significant in the EVR region, but could be significant if not for the large proportion of the land occupied by protected, uninhabited land. These variables are also both human constructs that exert influence on the land use of their locations. Just as federal or state protection creates conditions possible for lands to host large wildfires, roads offer the development and infrastructure that make lands more susceptible to fire suppression. While topographic, bioclimatic, and land cover variables were each significant amongst different study areas, measures of human-influence on the landscape proved to be the most universal predictors of both wildfire presence and absence on the landscape.

7. CONCLUSIONS

An inherent, intermingled relationship exists between humans and the natural environment: a testament the humanity's struggle to control nature. Humanity has, and always will grapple with natural disturbance and wildfire is no exception. There is, therefore, no 'one size fits all' for modeling wildfire distribution in the Southeastern United States, or any region for that matter. Despite this, commonalities can be found and trends can be recognized. This research sought to help define the variables most influential on the distribution of wildfire in the Southeastern United States. By understanding this, it will be possible to expand scientific understanding of both current wildfire regimes and the ever-changing fire regimes across the globe. Pyrogeography's relevance lies not in just its impact on today's ecosystems but on ecosystems subject to climate change, land use change, and continued human development.

This study has found that regional and sub-regional influencers of wildfire vary both within and amongst spatial scales. Bioclimatic and land cover variables hold significant influence on the prevalence of both the presence and absence of wildfire across each physiographic region analyzed. Human-influence, however, reigns as the most consistent effect on both the presence and absence of wildfire in the Southeastern United States. Predicting the presence or absence of wildfire depends not just on the bioclimatic or land use characteristics of a study area but how these lands are impacted by human activity. Just as human-impact will continue to affect the climate, topography and land cover across the world, knowingly or not, it will continue to change the world's fire regimes.

7.1. Further Directions

There exists opportunity for the expansion of both the breadth and impact of this study given greater time and resources. As noted in the interpretation of sub-regions analyzed in this study, covariates could be scaled more appropriately to the size of the study regions. While the resolution analyzed was more than appropriate for the SE regional study area or large sub-regions like the APP or CSTL, this resolution was less appropriate for smaller regions like the EVR or EBF. A more proportional resolution to spatial extent would allow research to identify smaller-scale variation across the landscape. Similarly, a decrease in the size of wildfires analyzed for this study would increase available data and would help to strengthen results. Likewise, this study used the standard distance derived from the SE study region to define pseudo-absence points for the whole region. By defining this metric on a regional basis, the estimation of absence points could be more fine-tuned, and better represent areas with absence in small region. Future studies incorporating these scaled changes would better reflect the nature and trends observable in smaller regions. The results of this study could also be strengthened by testing models on similar eco-regions in different study areas. More replications of these models and relevant predictors would serve to enhance these results by supporting the claims of this study and testing its hypotheses on new areas. Future studies incorporating these changes would enhance the strength of their findings.

This research provides a strong foundation for future research exploring the relationship between wildfire and land-change. The coupling of this model with climate change and land-use change models would expand the results of this study by accounting for the temporal aspect. While primary effects of climate change are increasingly studied in scientific literature, the secondary effects to both disturbance regimes and the ecosystems they inhabit remains largely unrecognized. Expansions on this work would continue to fill these gaps and explore the impacts of climate change on global pyrogeography.

REFERENCES

Agresti, A., and B. Finlay. 1997. *Statistical Methods for the Social Sciences*. 3rd ed: Prentice Hall.

Aho, K., D. Derryberry, and T. Peterson. 2014. Model selection for ecologists: the worldviews of AIC and BIC. *Ecology* 95 (3):631-636.

Anselin, L. 1998. *Exploratory spatial data analysis in geocomputational environment*.

Archibald, S., C. E. R. Lehmann, J. L. Gómez-Dans, and R. A. Bradstock. 2013. Defining pyromes and global syndromes of fire regimes. *Proceedings of the National Academy of Sciences of the United States of America* (16):6442.

Ascione, A., A. Cinque, E. Miccadei, F. Villani, and C. Berti. 2008. The Plio-Quaternary uplift of the Apennine chain: new data from the analysis of topography and river valleys in Central Italy. *Geomorphology* 102 (1):105-118.

Bacchus, S. T. 2000. UNCALCULATED IMPACTS OF UNSUSTAINABLE AQUIFER YIELD INCLUDING EVIDENCE OF SUBSURFACE INTERBASIN FLOW1. *JAWRA Journal of the American Water Resources Association* 36 (3):457-481.

Bailey, R. G., and L. Ropes. 1998. *Ecoregions : the ecosystem geography of the oceans and continents : with 106 illustrations, with 55 in color. Robert G. Bailey ; illustrations by Lev Ropes*: New York : Springer, 1998.

Balch, J. K., B. A. Bradley, J. T. Abatzoglou, R. C. Nagy, E. J. Fusco, and A. L. Mahood. 2017. Human-started wildfires expand the fire niche across the United States. *Proceedings of the National Academy of Sciences*:201617394.

Beer, T. 1991. The interaction of wind and fire. *Boundary-Layer Meteorology* 54 (3):287-308.

Bistinas, I., D. Oom, A. C. L. Sá, S. P. Harrison, I. C. Prentice, and J. M. C. Pereira. 2013. Relationships between human population density and burned area at continental and global scales. *PLoS ONE* 8 (12):e81188-e81188.

Bond, W. J., and J. E. Keeley. 2005. Fire as a global ‘herbivore’: the ecology and evolution of flammable ecosystems. *Trends in Ecology & Evolution* 20:387-394.

Bowman, D. M. J. S., B. P. Murphy, G. J. Williamson, and M. A. Cochrane. 2014. Pyrogeographic models, feedbacks and the future of global fire regimes. *Global Ecology and Biogeography* 23 (7):821-824.

Bradstock, R. A. 2010. A biogeographic model of fire regimes in Australia: current and future implications. *Global Ecology and Biogeography* 19 (2):145-158.

Brose, P., T. Schuler, D. van Lear, and J. Berst. 2001. Bringing Fire Back: The Changing Regimes of the Appalachian Mixed-Oak Forests. *Journal of Forestry* 99 (11):30-35.

Calviño-Cancela, M., M. L. Chas-Amil, E. D. García-Martínez, and J. Touza. 2016. Wildfire risk associated with different vegetation types within and outside wildland-urban interfaces. *Forest Ecology and Management* 372:1-9.

Cardille, J. A., and S. J. Ventura. 2001. Occurrence of wildfire in the northern Great Lakes Region: Effects of land cover and land ownership assessed at multiple scales. *International Journal of Wildland Fire* 10 (2):145-154.

Carmo, M., F. Moreira, P. Casimiro, and P. Vaz. 2011. Land use and topography influences on wildfire occurrence in northern Portugal. *Landscape and Urban Planning* 100 (1):169-176.

Chou, Y. H. 1992. Spatial autocorrelation and weighting functions in the distribution of wildland fires. *International Journal of Wildland Fire* 2 (4):169-176.

Clements, F. E. 1916. *Plant succession ; an analysis of the development of vegetation. by Frederic E. Clements*: Washington : Carnegie institution of Washington, 1916.

Davis, S. M., J. C. Ogden, and W. A. Park. 1994. *Everglades : the ecosystem and its restoration. editors, Steven M. Davis, John C. Ogden ; associate editor, Winifred A. Park*: Delray Beach, Fla. : St. Lucie Press, [1994].

Fick, S. E., and R. J. Hijmans. 2017. Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*.

Finco, M., B. Quayle, Y. Zhang, J. Lecker, K. A. Megown, and C. K. Brewer. 2012. Monitoring trends and burn severity (MTBS): monitoring wildfire activity for the past quarter century using Landsat data. Paper read at In: Morin, Randall S.; Liknes, Greg C., comps. Moving from status to trends: Forest Inventory and Analysis (FIA) symposium 2012; 2012 December 4-6; Baltimore, MD. Gen. Tech. Rep. NRS-P-105. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station.[CD-ROM]: 222-228.

Fotheringham, A. S., C. Brunsdon, and M. Charlton. 2000. *Quantitative geography : perspectives on spatial data analysis. A. Stewart Fotheringham, Chris Brunsdon, Martin Charlton*: London : SAGE, 2000.

Guyette, R. P., M. C. Stambaugh, D. C. Dey, and R.-M. Muzika. 2012. Predicting Fire Frequency with Chemistry and Climate. *Ecosystems* (2):322.

Harden, J., R. Meier, C. Silapaswan, D. Swanson, and A. D. McGuire. 2003. *Soil Drainage and Its Potential for Influencing Wildfires in Alaska*.

Homer, C. G., J. Dewitz, L. Yang, S. Jin, P. Danielson, Xian, J. Coulston, N. Herold, J. Wickham, and K. Megown. 2015. Completion of the 2011 National Land Cover Database for the conterminous United States – representing a decade of land cover change information, Photogrammetric Engineering and Remote Sensing. 81:345-353.

Hoover, K. 2018. Wildfire Statistics - IF10244, ed. CRS: Congressional Research Service.

Kalabokidis, K. D., P. Konstantinidis, and C. Vasilakos. 2002. GIS analysis of physical and human impact on wildfire patterns. Paper read at Proceedings of the IV international conference on forest fire research & wildland fire safety, Coimbra. Millpress, Rotterdam.

Klinkenberg, B. 1992. Fractals and morphometric measures: is there a relationship? *Geomorphology* 5 (1):5-20.

Krawchuk, M. A., M. A. Moritz, M. A. Parisien, J. Van Dorn, and K. Hayhoe. 2009. Global pyrogeography: The current and future distribution of wildfire. *PLoS ONE* 4 (4).

Lafon, C. W., A. T. Naito, H. D. Grissino-Mayer, S. P. Horn, and T. A. Waldrop. 2017. *Fire history of the Appalachian Region : a review and synthesis. Charles W. Lafon [and four others]*: [Ashville, N.C.] : United States Department of Agriculture, Forest Service, Southern Research Station, 2017.

Latham, D. 2013. *Painting the landscape with fire : longleaf pines and fire ecology. Den Latham, foreword by Shibu Jose*: Columbia, S.C. : University of South Carolina Press, 2013.

Lockwood, J. L., M. S. Ross, and J. P. Sah. 2003. Smoke on the water: the interplay of fire and water flow on Everglades restoration. *Frontiers in Ecology and the Environment* 1 (9):462-468.

Maingi, J. K., and M. C. Henry. 2007. Factors influencing wildfire occurrence and distribution in eastern Kentucky, USA %J International Journal of Wildland Fire. 16 (1):23-33.

Marella, R. L. 2014. Water withdrawals, use, and trends in Florida, 2010. In *Scientific Investigations Report*, 72. Reston, VA.

Menard, S. W. 2010. *Logistic Regression: From Introductory to Advanced Concepts and Applications*. Thousand Oaks, Calif.: SAGE.

Mitchell, A. 1999. *The ESRI guide to GIS analysis. 1st ed. Andy Mitchell*: Redlands, Calif. : ESRI, ©1999-©2012.1st ed.

Moreira, F., P. Vaz, F. Catry, and J. S. Silva. 2009. Regional variations in wildfire susceptibility of land-cover types in Portugal: implications for landscape management to minimize fire hazard. *International Journal of Wildland Fire* 18 (5):563-574.

Nunes, M. C. S., M. J. Vasconcelos, J. M. C. Pereira, N. Dasgupta, R. J. Alldredge, and F. C. Rego. 2005. Land Cover Type and Fire in Portugal: Do Fires Burn Land Cover Selectively? *Landscape Ecology* 20 (6):661-673.

One Meter Digital Elevation Models (DEMs) - USGS National Map 3DEP Downloadable Data Collection. 2015. ed. U. S. G. Survey. The National Map: Data.gov.

Parente, J., M. Ł. G. Pereira, and M. Tonini. 2016. Space-time clustering analysis of wildfires: The influence of dataset characteristics, fire prevention policy decisions, weather and climate [electronic resource]. *Science of the total environment* 559:151-165.

Parisien, M.-A., and M. A. Moritz. 2009. Environmental controls on the distribution of wildfire at multiple spatial scales. *Ecological Monographs* 79 (1):127-154.

Pausas, J. G., and E. Ribeiro. 2013. The global fire–productivity relationship. *Global Ecology and Biogeography* 22 (6):728-736.

Protected Areas Database of the United States (PAD-US), version 1.4 Combined Feature Class. 2016. U.S. Geological Survey, Gap Analysis Program (GAP).

Ricotta, C., S. Bajocco, D. Guglietta, and M. Conedera. 2018. Assessing the Influence of Roads on Fire Ignition: Does Land Cover Matter? *Fire* 1 (2).

Ripley, B. D. 1976. The second-order analysis of stationary point processes. *Journal of Applied Probability* 13 (2):255-266.

Rosenblatt, M. 1956. Remarks on Some Nonparametric Estimates of a Density Function. *The Annals of Mathematical Statistics* 27 (3):832.

Serra, L., P. Juan, D. Varga, J. Mateu, and M. Saez. 2013. Spatial pattern modelling of wildfires in Catalonia, Spain 2004–2008. *Environmental Modelling & Software* 40:235-244.

Studenmund, A., and H. Cassidy. *Using Econometrics: A Practical Guide (5 th)*.