NOVA IMS Information Management School

MEGI

Mestrado em Estatística e Gestão de Informação
Master Program in Statistics and Information Management

The Contributing Factors of Large Wildfires

Exploring the main structural factors driving large wildfire ignition and spread in central Portugal (2005-2015)

David Rodrigues Alves de Sousa

Dissertation presented as partial requirement for obtaining the Master's degree in Statistics and Information Management, with a specialization in Information Analysis and Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

THE CONTRIBUTING FACTORS OF LARGE WILDFIRES: EXPLORING THE MAIN STRUCTURAL FACTORS DRIVING LARGE WILDFIRE IGNITION AND SPREAD IN CENTRAL PORTUGAL (2005-2015)

David Rodrigues Alves de Sousa

Supervisor: Marco Painho, Ph.D.

Co-supervisor: Frederico Cruz Jesus, Ph.D.

Dissertation presented as partial requirement for obtaining the Master's degree in Statistics and Information Management, with a specialization in Information Analysis and Management

September 2018

AGRADECIMENTOS

Muitas foram as pessoas que me ajudaram e tornaram possível a elaboração desta dissertação.

Em primeiro lugar, devo um agradecimento especial ao orientador, Professor Doutor Marco Painho, e ao co-orientador desta tese, Professor Doutor Frederico Jesus, pela sua disponibilidade, por todo o trabalho desenvolvido em conjunto e pelos conselhos e orientações imprescindíveis à elaboração deste trabalho.

À Professora Doutora Ana Cristina Ferreira, pela ajuda inestimável durante desenvolvimento da metodologia adotada e por tudo o resto.

Em igual medida, aos meus colegas Bartholomeus Schoenmakers, André Sousa e Sandra Lagarto no Instituto Nacional de Estatística, que disponibilizaram parte do seu tempo para me ajudar com inúmeras questões metodológicas ao longo do percurso e sem os quais o resultado deste trabalho não seria o mesmo.

À minha família, amigos e colegas pelo apoio que me proporcionaram todos os dias e pela motivação que me transmitiram sempre que necessário.

Um sincero Obrigado!

ABSTRACT

Large wildfires have devastating human, environmental and economic consequences and are responsible for the majority of total burned area in Mediterranean Europe, even though they account for only a marginal portion of all fire occurrences. Most predictions suggest a global intensification of fire danger, and among all European Mediterranean countries Portugal displays the highest fire incidence. The purpose of this work is to examine the main factors driving large wildfire ignition and spread in central Portugal between 2005 and 2015, contributing with empiric knowledge on their importance and variability throughout the study area.

This research was successful at listing a comprehensive set of elements contributing to fire occurrence and at gathering data on these phenomena. Spatial cluster analysis was used to find homogeneous regions within the study area concerning the main factors influencing both fire ignition and burned area. Probit and two-part regression techniques were used to model the contribution of the different elements driving large fire occurrence and propagation.

The main findings of this analysis confirm the presence of spatial variability in the contribution exerted by most structural factors driving large wildfire ignition and spread in central Portugal. Additionally, while vegetation characteristics appear much more relevant for fire propagation, socioeconomic elements seem to be connected to fire incidence.

All in all, this research provides relevant input with implementation in different fields, from large fire awareness and prevention to the development of wildfire policies, as well as appropriate contributions to methodological concerns in fire danger and fire risk analyses.

KEYWORDS

Large wildfires; fire danger; large fire ignition; large fire spread; driving factors; spatial clustering; probit regression; two-part models

RESUMO

Os grandes incêndios rurais têm como consequência impactos socioeconómicos e ambientais devastadores e são responsáveis pela maior parte do total de área ardida na Europa mediterrânica, ainda que representem apenas uma fração pouco expressiva do total de ocorrências. A maioria dos estudos prevê uma intensificação do perigo de incêndio, sendo que, entre todos os países europeus da bacia mediterrânica, é Portugal quem apresenta a mais alta incidência deste fenómeno. O objetivo deste trabalho é estudar os fatores que mais contribuíram para a ignição e propagação de grandes incêndios rurais no centro de Portugal entre 2005 e 2015, concorrendo assim com conhecimento empírico relativamente à sua importância e variabilidade na área de estudo.

Esta investigação conseguiu listar um conjunto abrangente de elementos que contribuem para a ocorrência de incêndios rurais, assim como reunir os dados necessários. Uma análise de *clusters* espacial foi aplicada para identificar regiões homogéneas dentro da área de estudo no que respeita aos principais fatores influenciando a ignição e o alastrar dos grandes incêndios. Modelos *probit* e em duas partes foram utilizados para analisar a contribuição dos diferentes elementos para a ocorrência e propagação dos fogos.

Os resultados deste estudo confirmam a presença de variação espacial no impacto exercido pela maioria dos fatores estruturais que contribuem para a ocorrência e propagação dos grandes incêndios rurais. Por outro lado, enquanto as características da vegetação se revelam mais relevantes na perspetiva do alastrar dos incêndios, os fatores socioeconómicos parecem estar relacionados com a ignição destes fenómenos.

Em suma, este estudo contribui com informação relevante, a implementar em diferentes âmbitos, desde a consciencialização das populações à prevenção e ao desenvolvimento de políticas na área dos fogos rurais. Este apresenta ainda contributos apropriados na área de metodologias de análise do perigo e risco de incêndio.

PALAVRAS-CHAVE

Grandes incêndios rurais; perigo de incêndio; ignição; propagação; fatores determinantes; análise de *clusters* espacial; regressão *probit*; modelos em duas partes

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF MAPS

LIST OF ABBREVIATIONS AND ACRONYMS

CAOP	Carta Administrativa Oficial de Portugal
COS	Carta de Ocupação e Uso do Solo
DEM	Digital Elevation Model
DGT	Direção-Geral do Território
FWI	Fire Weather Index
GIS	Geographic Information Systems
ICNF	Instituto da Conservação da Natureza e das Florestas
INE	Instituto Nacional de Estatística
NOAA	National Oceanic and Atmospheric Administration (USA)
PNDFCI	Plano Nacional de Defesa da Floresta contra Incêndios
SAU	Used Agricultural Surface
SNDFCI	Sistema Nacional de Defesa da Floresta contra Incêndios
WUI	Wildland-Urban Interface

1. INTRODUCTION

This research seeks to investigate the main factors driving large wildfire occurrences in central Portugal between 2005 and 2015, focusing on their spatial patterns, dynamics and impact on the likelihood of fire ignition and burned area. This document is divided into 5 chapters: (1) the introduction; (2) a review of literary sources; (3) an account of the chosen methodology; (4) the presentation of the main findings of this work, complemented by a discussion; and (5) the concluding remarks.

1.1. BACKGROUND AND RESEARCH DEVELOPMENT

Fire is perceived as an important agent for both ecological development and deterioration of forest ecosystems around the world (Ferreira-Leite, Lourenço, & Bento-Gonçalves, 2013; Verde & Zêzere, 2010). In fact, even though fire has always be present as a landscape transformation and renewal factor (Goudie, 2006), fire events can have disastrous human, environmental and economic consequences, especially when these develop into large wildfires (Tedim et al., 2013).

Mediterranean Europe is particularly vulnerable to these destructive phenomena. This region is home to the second most diverse community of species worldwide after the tropics, and even though Mediterranean ecosystems are historically fire resilient, intense human pressure on the environment has favoured fire recurrence, has increased fire size and has made fire events overall more widespread (Moreno, Vallejo, & Chuvieco, 2013). Large fires play a significant role in this context because they are responsible for the majority of total burned area, while only expressing a small share of all fire events (San-Miguel-Ayanz, Moreno, & Camia, 2013).

These trends in fire dynamics have been observable across European Mediterranean countries, although at different rates. In Portugal, there is evidence of higher fire incidence than in other Mediterranean countries (Rego & Silva, 2014), a steady increase in the frequency of large wildfires and a rise in the extent of burned area during the second half of the 20th century (Ferreira-Leite et al., 2013). Climatic characteristics, land abandonment and other socioeconomic transformations have left the Portuguese inland territory susceptible to the occurrence of large wildfires (Oliveira, Pereira, & Carreiras, 2012), and climate change is expected to reinforce fire danger significantly (Rego & Silva, 2014).

One aspect that is important in shaping wildfire frequency and size is the role of humans and their activities, which is known to both stimulate fire ignition and prevent large burned areas. Human-caused fires represent 90% of all occurrences in the European Mediterranean (Moreira, Catry, Rego, & Bação, 2010), and in Portugal, the main known causes of forest fires are the use of fire, arson and rekindling (ICNF, 2014).

Wildfire risk is partially determined by fire danger, in which fire ignition and propagation likelihood play a central role (Chuvieco, Aguado, et al., 2014). In turn, the driving forces of fire occurrence and spread can be divided into vegetation and soil factors (the available fuel and its spatial continuity), climatic factors (favourable weather conditions), topographic factors, and human factors (usually, a source of ignition) (Ganteaume et al., 2013; Martínez-Fernández, Chuvieco, & Koutsias, 2013; Mhawej, Faour, & Adjizian-Gerard, 2015). These structural elements display various relationships among themselves, explaining wildfire incidence and determining the behaviour of fire events.

1.2. RESEARCH GOALS

The overall purpose of this research can be summarised into one main objective and four specific research goals, as presented below:

Table 1 – Main objective and research goals

Main Objective: To study the main determinants of large wildfire occurrences in central Portugal between 2005-2015 as to their spatial patterns, dynamics and contribution.
Research Goal #1: To recognise and to explain the main general factors driving wildfire ignition and spread, specifically in the Mediterranean context. **Research Goal #2:** To investigate the spatial patterns of large wildfire incidence and the distribution of the main drivers of fire occurrence and propagation. **Research Goal #3:** To identify homogenous regions within the study area with respect to the distribution of the main drivers of fire ignition and burned area. **Research Goal #4:** To model the occurrence and spread of large wildfires from the contribution of its main factors and to explore the variability of the factors importance among regions.

In order to accomplish these objectives, a comprehensive literature review on the factors driving wildfire ignition and spread will be carried out, along with multivariate data analysis techniques. Cluster analysis will be used to identify distinctive regions within the study area. Probit and two-part modelling procedures will be used to assess the influence of the main elements driving large wildfire occurrence and burned area, as well as the lack of stationarity across regions.

1.3. STUDY RELEVANCE

The relative influence of a great array of factors driving large wildfire occurrence in Portugal, in which we can count biophysical, land use, ownership and socioeconomic aspects, needs to be further studied for a richer understanding of fire dynamics. As Tedim et al. (2013) and Calviño-Cancela et al. (2017) put it, an insight into the various interactions among the main factors driving wildfire risk and fire behaviour – studying their arrangements, patterns and variations – should translate into an improvement in risk management strategies, such as more efficient prevention measures, the development of existing regulation and the creation of campaigns aimed at increasing the awareness of different stakeholders in connection with specific activities and settings.

Additionally, a better understanding of wildfire occurrences and their spatial behaviour is essential for prevention efforts and firefighting planning, also in the legislative framework (Moreira et al., 2010). In fact, the spatial patterns of fire events are counted as one of the essential elements in wildfire danger assessment, and the study of these varying distributions helps understanding the reasons behind different fire regimes in otherwise similar regions (Rego & Silva, 2014).

The results of empirical studies in this area are still considered valuable input for model parameterisation in the perspective of risk analysis (Miller & Ager, 2013). There is a high degree of uncertainty involved in the study of wildfires, especially of those anthropogenic in nature, connected to deliberate, accidental and negligent causes (Rodrigues, Jiménez, & de la Riva, 2016). Furthermore, research studying large wildfires has not been very extensive in respect to their main factors, even though they are responsible for the majority of wildfire damages in the Mediterranean (Dimitrakopoulos, Gogi, Stamatelos, & Mitsopoulos, 2011; Ganteaume & Jappiot, 2013). Deeper insights into the mechanisms triggering large wildfire occurrence will hopefully provide the necessary knowledge to strengthen prevention efforts (San-Miguel-Ayanz et al., 2013).

The focus of this study will be on large events considering burned area, making use of the findings of previous works. This knowledge should result in beneficial developments in awareness, prevention and wildfire policies, and in advances in the effect mitigation of large wildfires (Ganteaume & Jappiot, 2013; Grala et al., 2017), specifically in the Portuguese context.

2. LITERATURE REVIEW

This chapter provides an indicative review of the main literary sources on this topic. To start, general considerations about wildfires are presented, as well as the most striking and important characteristics of large wildfire dynamics, specifically in the Portuguese context. Later, wildfire risk assessment frameworks are discussed and the main contributing factors of wildfires are described and explained. To conclude a short account of the methods used for analysing wildfire danger is given, with attention to the main goals of this research.

2.1. WILDFIRES – SOME CONSIDERATIONS

A quick look through the reference literature provides a list of different words or expressions to designate the phenomenon under study, from the most frequent *wildfire* and *forest fire* to a variety of other technical expressions. Mhawej, Faour and Adjizian-Gerard (2015) agree that a "*wildfire* is generally the unplanned or unwanted natural or human-caused fire that spreads over a minimum area of one hectare, where one or more types of vegetation are concerned" while asserting that the term *forest fire* can be used interchangeably.

Although the latter is the preferred expression in Portuguese-language studies, the term *wildfire* appears to extend beyond the specific context of forest environments and is the most widely used expression in English-language academic settings and elsewhere, making it the choice of this study.

Fire has a necessary and natural role in the environment as an agent for landscape transformation and wildlife development (Verde & Zêzere, 2007). Goudie (2006) discusses at length the different ways in which fire contributes to species diversity, pest control and vegetative reproduction. However, wildfires also represent the source of great destruction to forested and populated areas alike, with recent examples of extensive damage in Greece and other regions with Mediterranean characteristics (Mitsopoulos, Mallinis, & Arianoutsou, 2014), including Portugal. All in all, Ferreira-Leite, Lourenço and Bento-Gonçalves (2013) consider it to be a fundamental aspect of ecological development or deterioration of global forest environments.

In the literature, the expression large fire is used to identify all wildfires that either have the possibility of developing into very large fires or are liable to burn a large area, while the size varies depending on the author or study (Ganteaume & Jappiot, 2013).

In line with the Resolution from the Portuguese Parliament n. 35/2013, of March 19th (Resolução da Assembleia da República n.º 35/2013, de 19 de março in Diário da República n.º 55/2013, Série I), large wildfires refer to all events with a burned area over 500 ha, replacing the previous institutional classification which considered all occurrences with a burned area over 100 ha (Ferreira-Leite et al., 2013). It is important to mention, however, that this second threshold is still widely used in the reference literature to identify large wildfires, and represents the classification used in this research work.

On the other hand, there is an important difference between large fires and mega-fires, which represent a specific subgroup of the former. These are considered complex, multidimensional and unusual events, specifically in what concerns burned area extension, damage severity and associated firefighting efforts (Dimitrakopoulos et al., 2011; Tedim et al., 2013). According to San-Miguel-Ayanz et al. (2013), megafires can be charcaterised on the basis of their damaging impacts, specifically economic losses and human casualties. Tedim et al. (2013) also mention the main attribute of these phenomena to be their short and long-term impacts.

2.1.1. The consequences of wildfire events

The negative consequences of wildfires are several, wide-ranging and are well documented throughout the literature. Tedim et al. (2013) separate direct damages and indirect losses, stressing the environmental and economic impacts of fire events inside and outside the burned area, even though many times, in Portugal, damages are measured as a sole function of burned area size. These same authors consider that wildfire impacts may be classified as to their type and their spatial and temporal scales, for a comprehensive assessment and monitoring of this phenomenon.

According to Oliveira, Pereira and Carreiras (2012) the three decades of intense fire activity starting in 1980 and ending in 2010, in Portugal, have resulted in a replacement of large areas of maritime pine forests for eucalypt plantations and shrublands, with all the negative conditions it entails for future fire regimes. Ricotta and Di Vito (2014) agree that, in Mediterranean ecosystems, recurrent wildfires are found to thoroughly disturb landscapes, reducing biodiversity and favouring the growth of shrubland areas, since many species require a sufficiently longer time interval to regenerate. In turn, this degradation promotes wildfire occurrence, given the apparent lower value of such areas. Countless other environmental dangers have been verified, such as the aggradation of rivers, which is a real danger in the Mondego basin in central Portugal (Ferreira-Leite, Lourenço, et al., 2013).

On the other hand, wildfires pose numerous threats to human lives, their resources and their activities (Chuvieco, Aguado, et al., 2014), namely a serious impact on the forestry sector, reducing earnings and decreasing investment (Álvarez-Díaz, González-Gómez, & Otero-Giraldez, 2015) and negative economic effects on the farming activity, damaging machines, burning profitable crops and killing animals (Tedim et al., 2013). It is expected that demographic changes increase wildfire impacts on human safety, as in some areas of the world, such as California, the urban sprawl has provided a source for rising ignition risk in particularly vulnerable settings (Keeley & Syphard, 2016).

In the words of Moreno, Vallejo and Chuvieco (2013), the Mediterranean Europe is currently classified as degraded, as a consequence of long-term human presence and exploitation. This anthropogenic impact is known to influence both wildfire occurrence, as well as the ecological response to such extreme events. Additionally, destructive effects are more likely to ensue as a result of a high number of infrastructures and a prevalence of densely populated areas, particularly during the summer months.

Other important aspects include carbon emissions (Amiro et al., 2001; van der Werf et al., 2003), and high economic impacts, especially connected to fire suppression (Westerling & Bryant, 2007), other aid expenditures, like evacuation procedures, and the recovery process (Tedim et al., 2013). In fact, Moreno, Vallejo and Chuvieco (2013) acknowledge that more than 2.5 billion euros are spent every year in prevention and suppression across the European Mediterranean region.

When focusing specifically on the consequences of large forest fires, great damages are an expected outcome. The work of Tedim et al. (2013) covering the mega-fires occurred in Algarve, south of Portugal, in 2003, gives an account of the effects of these extreme events on dwellings, forests and agricultural lands, the shortage of stored water and the destruction of power and telephone lines. An association has been made by the same study between these high damages and the location of the wildfires, which spread to areas on the Wildland-Urban Interface, where human presence is dispersed and elements are particularly vulnerable.

Conversely, Künzli et al. (2006) have demonstrated that health effects on children, namely respiratory conditions, are one of the serious concerns of smoke from large wildfires, while Dimitrakopoulos et al. (2011) have focused their attention on the high costs of firefighting, particularly due to extended suppression times and the use of airborne means.

It is worth considering the perceptions of different stakeholders when discussing the main effects of wildfires. A study conducted in the Portuguese municipality of Mação (NUTS III Médio Tejo) has shown that, while forestry technicians mentioned soil erosion and the loss of biodiversity as their main concerns, the local community pointed out the loss of timber and the destructive impacts to the local economy and private property (Ribeiro et al., 2015).

2.1.2. An insight into fire dynamics

The term fire regime defines the regular and constant features of wildfire behaviour and dynamics occurring in a given area over a period of time (Jiménez-Ruano, Rodrigues, & de la Riva, 2017). According to Seol, Lee and Chung (2012), fire regimes differ in size, regularity, pattern, intensity, type, magnitude and seasonality of fire occurrences.

There is evidence of an increasing frequency and severity of wildfires over the 2nd half of the 20th century (Ferreira-Leite et al., 2013). Holsinger, Parks and Miller (2016) argue that this trend is related to forest fire deficits created by a decrease in the prevalence of managed wildfires. In fact, there is evidence of a rise in wildfire occurrences in the Mediterranean during the 1990s. However, since 2000 a noticeable but inconsistent decrease has been witnessed by some researchers (San-Miguel-Ayanz et al., 2013).

In Portugal, where the territory is already prone to this phenomenon as a result of climatic characteristics, population-related and socioeconomic transformations, as well as climate change, have reinforced susceptibility (Oliveira et al., 2012). This tendency is only expected to become more prominent for Portugal and other Mediterranean countries given future predictions of warmer and drier weather, in connection to structural changes occurring to the social and economic systems (Moreira et al., 2010; Rego & Silva, 2014).

While increasing frequency is an alarming scenario, given that short wildfire recurrence produces the most hazardous impacts to land degradation (Ricotta & Di Vito, 2014), an increase in severity has given way to wider-ranging damages, specifically to the environment, property and human life (Álvarez-Díaz et al., 2015). A clear example of more severe fire regimes is a general intensification of burned area.

Although there is evidence of large wildfires occurring in Portuguese territory in the 19th century, their frequency has been steadily rising during the last three decades (1980-2010), and a very substantial increase in the extent of burned area has been witnessed, with the last decade accounting for more than 1 million burned hectares (Ferreira-Leite, Lourenço, et al., 2013). For the southern Europe and elsewhere, there is ample agreement in the literature that large fires (>100 ha), which represent a marginal portion of all occurrences, are responsible for the great majority of total burned area (Dimitrakopoulos et al., 2011; Ganteaume & Jappiot, 2013; Moreno et al., 2013; Ricotta & Di Vito, 2014; San-Miguel-Ayanz et al., 2013).

Nevertheless, the analysis conducted by San-Miguel-Ayanz et al. (2013) for a series of countries in the Mediterranean basin, including Portugal, has failed to prove a growing trend in the number of large fires in the last decades. In fact, there is evidence of negligible variability in the number of large occurrences throughout the years and even a slight decreasing trend.

In Portugal, the same human-induced changes linked to an increase in wildfire risk, specifically what concerns population trends and land cover transformations, are believed to be responsible for more extensive burned areas (Moreira et al., 2010). Nevertheless, some authors suggest that their occurrence follows a cyclic behaviour, either connected to post-fire vegetation dynamics (Moreira et al., 2010) or to prolonged drought periods (Ganteaume & Jappiot, 2013).

When compared with other countries from the European Mediterranean region, there is evidence of higher fire incidence in Portugal (Moreira et al., 2010; Oliveira et al., 2012). Nevertheless, Oliveira et al. (2012) go on to describe the substantial regional variability of this phenomenon: the north is dominated by short fire recurrence intervals, frequent and overall small events considering burned area, and is mainly composed of shrubland; the central region, composed predominantly of forested areas, is also characterised by short intervals between wildfires, where occurrences are not many but usually originated large burned areas; apart from Algarve, in southern Portugal fire intervals are longer and burned area is small, considering its composition of mostly agricultural land.

The results from the work of Moreira et al. (2010) agree with these statements, since the places where most ignitions occur do not correspond to the locations subject to large burned areas. Whereas ignitions seem to be mostly related to the presence of human activities, depopulated areas seem to be more susceptible to fire spread. Therefore, wildfire events igniting along the oceanic shore or in the south of the country are more likely to remain small, while wildfire occurrences located near the eastern border with Spain, in the central and north-eastern regions or in the mountains in Algarve (south) have a greater probability of spreading to large areas, which is a similar pattern to that identified by Tedim et al. (2013).

Other authors have reached the same conclusions, as is the case with Rego and Silva (2014). They confirm there are clear geographical differences, in Portugal, between the distribution of ignitions developing into large wildfires and that of smaller sized occurrences. They also sustain that the role of wildfire effects differs among regions, which is a firm demonstration of the variability of fire dynamics.

Future perspectives on fire regimes cannot be detached from global climate projections, specifically in the Mediterranean. Several studies agree on the most distinctive foreseen weather patterns for this region of the world. Global warming will translate into an irreversible increase in annual temperatures and a rise in the intensity and frequency of heat waves. Moreover, it will represent a decrease in precipitation and wet days and an intensification of drought periods (Keeley & Syphard, 2016; Moreno et al., 2013; Rego & Silva, 2014). All in all, these conditions represent a rise in fire danger and imply an increase in the duration of the fire season (San-Miguel-Ayanz et al., 2013)

Adding to these aspects, the process of land-use changes is believed to keep affecting fire dynamics in southern Europe, either related to land abandonment or to a decrease in agricultural areas promoted by the Common Agricultural Policy (CAP) (Moreno et al., 2013).

All these elements bring about significant impacts to wildland and urban landscapes and suggest an increase in wildfire risk, especially regarding the occurrence of large fires (Moreno et al., 2013). However, climate change might also trigger a decline in fire severity, given the effects of drier and warmer weather on fuel availability (Keeley & Syphard, 2016).

2.1.3. The role of human activities

The role of humans and their activities in shaping wildfire size and frequency is multidimensional, representing both causative and deterrent aspects. Man has been an important player in the transformation of ecosystems since he has been able to control fire (Chuvieco, Aguado, et al., 2014), and this long-lasting impact is known to have altered fire regimes in most regions of the world (Rego & Silva, 2014), reducing fire-return intervals considerably (Moreno et al., 2013).

Long and concentrated human presence, together with weather features and the presence of flammable vegetation, has shaped fire regimes in the Mediterranean, making it susceptible to wildfire occurrence (Dimitrakopoulos et al., 2011; Ferreira-Leite et al., 2013). Although in some parts of the world natural ignitions are the origin of up to half of all fires (Goudie, 2006), and propagation is favoured from natural elements such as fuel availability and low humidity, the European Mediterranean faces an overwhelming preponderance of human-caused wildfires, which represent more than 90% of all wildfires in that region (Chuvieco, Aguado, et al., 2014; Moreira et al., 2010; Rego & Silva, 2014; Rodrigues et al., 2016; San-Miguel-Ayanz et al., 2013; Vilar et al., 2016).

The public report containing the analysis of wildfire causes in Portugal, for the period 2003-2013, goes as far as attributing 98% of all investigated fire occurrences to humans, either intentionally or negligently (ICNF, 2014). In fact, anthropogenic factors are the main structural elements influencing geographic and temporal ignition patterns in Portugal, as several studies have demonstrated (Rego & Silva, 2014).

Even though the proximity to human settlements has an overwhelming effect on wildfire dynamics, this effect is not necessarily harmful. A positive aspect connected to the importance of human factors for fire incidence and behaviour stems from our ability to better predict these events (Moreno et al., 2013), and therefore plan ahead to prevent them. Moreover, humans are a key element of fire suppression, reducing fire severity (Rego & Silva, 2014).

2.1.4. The causes of wildfire ignition

The understanding of the ignition sources and generic causes of wildfires is an integral part of the study of this phenomenon and of its main determinant factors. Some authors agree that the ignition typology has impact on fire behaviour and dynamics, particularly the extension of burned area, as well as wildfire frequency, duration and patterns (Ganteaume & Jappiot, 2013; Grala et al., 2017; Moreno et al., 2013). However, caution is necessary when analysing this information, since there is an associated component of high subjectivity related to the work of technicians (Álvarez-Díaz, González-Gómez, & Otero-Giraldez, 2015; Grala et al., 2017).

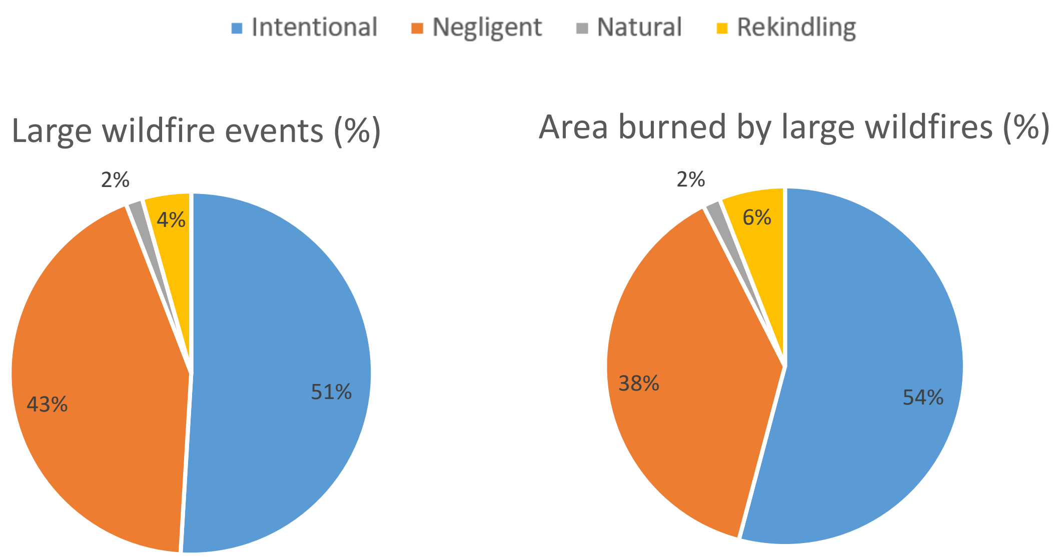
According to ICNF's report on the causes of forest fires in Portugal during the period 2003-2013, 76% of all events registered in the mainland have been subject to police investigation (ICNF, 2014). This has been standard procedure since 2006, when forest rangers were assimilated into the military police force (GNR). During this decade a considerable amount of inquiries were inconclusive; this occurred to 28% of all investigated events in 2013.

The methodology selected for the investigation of wildfire causes relies on physical evidences, since 1991, such as fire behaviour patterns and other indicators. The categories linked to wildfire causes comprise an extensive list of 71 subclasses, which distinguish specific activities and conducts and can be grouped into seven main types: use of fire, accidental, structural, arson (intentional), natural, undetermined and rekindling.

Between 2003 and 2013, the majority of wildfires whose cause has been determined was attributed to the use of fire (e.g. fires, debris burning, smoking; at around 30%), to arson (approximately 25%) and to rekindling (around 15%). Wildfires originated by the use of fire, although very frequent during this period, resulted in small burned areas, whereas natural or machinery (accidental) caused fires represented the vastest burned areas. The districts of Santarém, Braga and Viana do Castelo are identified as the regions most associated with intentional causes, specifically arson.

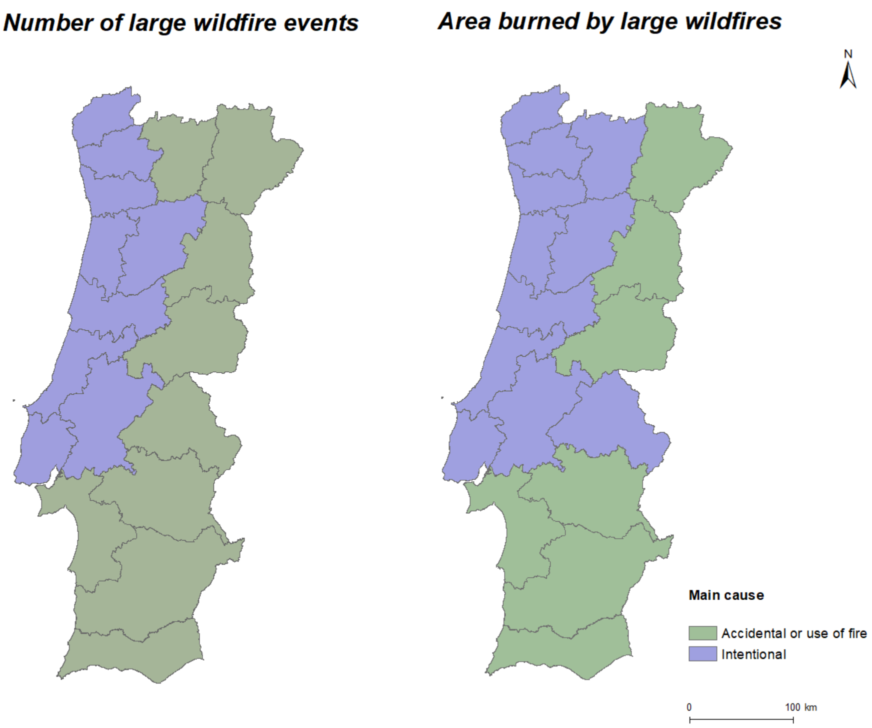
When focusing specifically on large wildfires (>100ha), data for the period of 2005-2015 and the two graphs below (Figure 1) show that the single most common cause typology refers to arson without an apparent reason (vandalism), standing at about 35% of all occurrences whose inquiries were successful, followed by the use of fire for grazing purposes (roughly 23%). The largest burned areas can also be attributed to these two causes (38% and 16%, respectively), as well as to other intentional sources (8%) and fire rekindling (6%). Overall, the grounds for a total of 570 occurrences, which represent more than 1/3 of all large fire events and around 1/3 of all burned area during this decade, were left undetermined.

Figure 1 – Large wildfire events and area burned by large wildfires, by cause of ignition, in Portugal (2005-2015)



An analysis of the spatial distribution of the main causes of large wildfires, in mainland Portugal, shows a balance among districts in respect to the most common category connected to fire events, between intentional motivations and accidental or use of fire sources. As shown on the first map below (Map 1), there is a clear contrast between north coastline and south and inland areas in this case. When considering only burned area extent, however, there is a higher incidence of deliberate causes throughout the territory (addition of Vila Real and Portalegre). It is interesting to realise that large wildfires originating from unknown causes were common in Aveiro (85%), Castelo Branco (76%) and Porto (68%), during this period.

Map 1 – The main causes of large wildfires, by district (2005-2015)



Arson is a particular aspect among human motivations. Álvarez-Díaz et al. (2015) underline the importance of this phenomenon in northern Portugal and Galicia, stemming from social conflicts, while Ganteaume and Jappiot (2013) describe it as the most frequent cause of wildfires in southern France, mostly related to the interests of hunters, shepherds and real estate mediators. This last study has been successful in associating arson and burned area, giving emphasis to its role in the occurrence of large wildfires.

It is interesting to observe that the perception of the main causes driving fire spread is not exactly coincident among stakeholder groups. Returning to the results of the study conducted in the Portuguese municipality of Mação regarding the perspectives of different stakeholders on the subject of forest wildfires, local forestry technicians mentioned the density of fuel accumulation, specifically shrub, and the type of vegetation as the most important aspects favouring the propagation of fires. Local community respondents have referred shrubland areas as well, while favouring the role of weather conditions (Ribeiro et al., 2015).

2.1.5. The Portuguese wildfire protection framework

Forest protection against wildfires is regulated in Portugal in the framework of the National Plan of Forest Protection against Wildfires (PNDFCI) and the National System of Forest Protection against Wildfires (SNDFCI). These are established in the scope of comprehensive legislation which arose from the need for a modern framework tackling the issue of forest fires in Portugal, in the wake of tragic wildfire events which affected the country in 2003 and 2005 and set new burned area records (C. Ribeiro et al., 2015). In fact, some authors mention wildfire catastrophes as powerful drivers of change in forest administration in different Mediterranean regions (San-Miguel-Ayanz et al., 2013).

The National Plan of Forest Protection against Wildfires stems from the Resolution of the Council of Ministers n. 65/2006, of May 26th (Resolução do Conselho de Ministros n.º 65/2006, de 26 de maio in Diário da República n.º 102/2006, Série I-B), establishing forest protection measures, lines of action and political reforms aiming at increasing the efficiency of forest management and reducing the number of wildfire occurrences (Viana, 2010). The SNDFCI was created shortly after the creation of the PNDFCI in order to regulate and specify the general measures presented on the national plan.

The Decree n. 124/2006, of June 28th (Decreto-lei n.º 124/2006, de 28 de junho in Diário da República n.º 123/2006, Série I-A) established the measures to be developed under the SNDFCI. This regulation has been subject to five revisions since its first publication, the last of which took place in 2017 and strengthened the role of the military in firefighting and prevention actions. However, the general aspects of the document, including the responsibilities of the different stakeholders, have been kept constant throughout the years.

According to this legislation, three action axes define the National System of Forest Protection against Wildfires, which together convey the protection of people, property and the forest: (1) structural prevention; (2) vigilance, detection and inspection; and (3) firefighting and post-fire vigilance (C. Ribeiro et al., 2015). The responsibility for each of these pillars belongs to a specific entity, with the Nature Preservation and Forests Institute (ICNF) coordinating awareness and planning actions, the police (GNR) taking charge of operational aspects such as alertness and assessment, and the National Authority for Civil Protection (ANPC) managing the combat and post-fire vigilance stages (Viana, 2010).

Apart from these overarching concerns, the Decree contemplates a reference to the different levels of planning. In reality, districts and municipalities have in their structures specific branches in charge of producing local forest protection plans and managing the participation of different stakeholders into this process (C. Ribeiro et al., 2015). Additionally, it legislates on the circumstances where the use of fire is not allowed, specifically in relation with fire-weather information, and the foreseen practices regarding wildfire detection and alarm (Viana, 2010).

Although one of the three main pillars of the Portuguese framework of protection against wildfires stands for structural prevention, some experts criticise the present paradigm as predominantly reactive, focusing efforts on firefighting (Rego & Silva, 2014). This fact can be demonstrated by the disproportionate annual resource allocation in the context of the SNDFCI: 1/3 for prevention and 2/3 for firefighting (according to the Portuguese State Secretary for Forests, Miguel Freitas, in *"Não me repugna que os bombeiros intervenham na prevenção", Público, 01/09/2017*).

It is important to note, however, that forest fire protection corresponds to one branch of the current Portuguese forest planning framework, which encompasses the broad technical and policy areas of forest management, forest health and desertification. Together, these comprise the four main levels of the Portuguese Forest Planning System (Pinho, 2014).

2.2. WILDFIRE RISK ASSESSMENT

2.2.1. Conceptual approaches to wildfire risk

The importance of gathering knowledge on the destructive characteristics of this phenomenon has led to the development of many studies on wildfire risk assessment. This is seen as a critical stage in resource planning, fire prevention and fire management activities generally speaking, by approximating the times and places most susceptible to increased wildfire occurrence and damages (Chuvieco, Aguado, et al., 2014).

According to several authors, the main guiding principles of risk analysis transpose soundly to the field of wildfires, for its focus on stochastic events of extreme consequences (Miller & Ager, 2013). Fire is a process determined by probability (Oliveira et al., 2012; Seol et al., 2012). As such, wildfire risk is the product of aspects related to ignition and burning likelihood, as well as a component linked to fire intensity and effects, in which propagation, damage and the difficulty of fire control are contemplated (Cao et al., 2013; Mhawej et al., 2015; Moreira et al., 2010).

Miller and Ager's (2013) wildfire risk assessment framework follows this general structure. As they see it, wildfire likelihood can be predicted from ignition probability, or simply the probability of burning, accounting for subsequent fire spread. On the other hand, fire intensity and effects, which are closely related in most models and come together to denote hazard, refer respectively to fire behaviour and the positive or negative modifications in environmental, social and economic values affected by this phenomenon.

On a similar yet distinct note, Mitsopoulos et al. (2014) refer to fire behaviour individually as a specific component of wildfire risk, underlining the view of some authors who identify it as the "likely behaviour given that a fire does occur". As we have stated before, in many contexts, the occurrence of specific fire types is natural and represents a necessity. This component of wildfire risk is also stressed by Tedim et al. (2013), which consider fire intensity, connected to the release of energy, and rate of propagation separate from the effects and possible damages of wildfires.

These views are enriched by those of Chuvieco, Aguado, et al. (2014), which recognise wildfire risk as a result of both fire hazard (used interchangeably with danger) and ecosystem vulnerability. Nevertheless, while Miller and Ager (2013) understand hazard as a combination of fire intensity and effects, separate from the probability of fire occurrence, for the first authors it encompasses ignition and propagation danger and is explained by the specific driving factors affecting these two fire components.

In fact, terminology has been an unlimited topic of discussion among researchers in this area of knowledge. Bachmann and Allgöwer's (2001) efforts in consistent language standardisation have originated a set of concepts borrowed from the field of technical risk engineering. These authors agree with Chuvieco, Aguado, et al. (2014) in what concerns the general notion of wildfire risk, integrating both the probability of fire occurrence and the chance of damage, i.e. vulnerability.

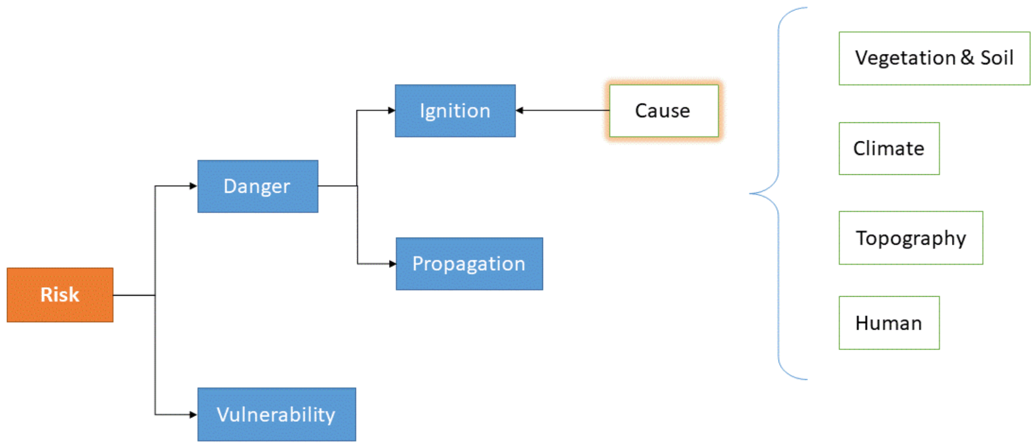
However, fire hazard is perceived by these authors as the phenomenon itself, distancing itself from the preconditions that drive fire incidence, as mentioned by Chuvieco, Aguado, et al. (2014), and from the potential damage given a fire event stressed by Miller and Ager (2013). Conversely, the term "danger" is deemed by these authors too subjective and abstract to be of any relevance to the field of study.

As mentioned before, the concept of fire risk is not complete without vulnerability. It refers to the susceptibility of communities to the effects of wildfires, or simply the magnitude of the damages they entail, whether environmental or economic (Chuvieco, Martínez, Román, Hantson, & Pettinari, 2014). It should be noted that it considers the time of recovery after the event and the adaptive capacity of landscapes (Chuvieco, Aguado, et al., 2014; Miller & Ager, 2013)

Some authors believe that vulnerability also affects hazard itself, by identifying the main human, economic and ecological values that drive firefighting operations (Chuvieco, Aguado, et al., 2014), thus influencing fire behaviour and spread. Additionally, Tedim et al. (2013) agree that the different categories of effects play a bigger role in wildfire risk, interacting with fire behaviour, the landscape and firefighting to outline fire severity.

The wildfire risk assessment framework defined in this study (Figure 2) draws extensive influence from the works presented above. This characterisation is important in the context of the identification of the main contributing factors of large wildfires in central Portugal. Even though the vulnerability component of fire risk stands beyond the scope of this study, its impact on fire danger is mentioned in connection to specific human and physical variables influencing fire spread.

Figure 2 – Adopted wildfire risk framework, adapted from Chuvieco, Aguado, et al. (2014)



2.2.2. Factors influencing wildfire ignition and burned area

The main factors related to fire occurrence and spread, and therefore decisively contributing to wildfire risk, have been determined through extensive scientific research in different settings and under different assumptions. According to Álvarez-Díaz et al. (2015), the four decisive conditions for wildfire occurrence and spread are favourable meteorological conditions, the presence of fuel, its spatial continuity and a source of ignition. From these stem all significant driving factors of wildfires under discussion in this study.

With a particular interest to this work lies a considerable list of literary references on the topic of driving factors, either specifically concerned with burned area extent or focused solely on large fire events (Annex A: Table A1). Similarly, most references refer to works studying Mediterranean regions around the world and Europe in particular.

In order to present a structured review of the main elements influencing wildfire incidence, we will follow the categorization adapted from Mhawej et al. (2015) and Ganteaume et al. (2013) in which driving forces of wildfires are divided into: (i) vegetation and soil factors; (ii) climatic factors; (iii) topographic factors; and (iv) human factors. Nevertheless, it is worth stressing the large variety of relationships displayed among factors, which are particularly relevant to consider for methodological purposes.

An association can be established between the different categories of factors driving wildfire occurrence and specific indices representing wildfire danger. In their 2009 work, Vasilakos, Kalabokidis, Hatzopoulos and Matsinos make use of climate, vegetation and terrain, and human variables, respectively, to model the Fire Weather Index (FWI), the Fire Hazard Index (FHI) and the Fire Risk Index (FRI), with a broader aim to assess the significance of the original variables for the Fire Ignition Index (FII). Their study was successful at identifying important factors from all areas in connection to fire ignition.

The categorisation of factors must also take into account the relevant effects produced by each fire danger element. Spatial wildfire patterns diverge among ignition and propagation focused analyses, meaning that higher fire incidence locations may not necessarily translate into areas affected by large wildfires or specifically vulnerable to fire spread. This statement is sustained by Miller and Ager's (2013) views on this subject, who believe that fire likelihood can be described by ignition characteristics in the case of small occurrences but not in the case of large wildfires. As such, the following review will take effort in distinguishing structural aspects driving fire ignition from those explaining fire spread.

2.2.2.1. Vegetation and soil factors

Vegetation conditions represent the decisive component of any fire. In fact, fuel is the main requirement for fire ignition and spread (Cao et al., 2013; Holsinger et al., 2016) and different vegetation patterns promote the fire susceptibility of landscapes. Even though many fire are human caused, specifically in southern Europe, the main features of local vegetation remain a determinant factor driving wildfire risk, as they determine the success of the ignition event – irrespective of cause – and, most importantly, fire behaviour (Calviño-Cancela et al., 2016).

Vegetation factors are believed to be poorly stationary over space, as is the influence they exert (Fernandes, Monteiro-Henriques, et al., 2016; Martínez-Fernández, Chuvieco, & Koutsias, 2013). As active management of vegetation is possible, knowledge of its effects provides great meaning to wildfire prevention strategies (Calviño-Cancela et al., 2017)

2.2.2.1.1. Type

Fuel type is known to influence both burned area and wildfire occurrence (Kalabokidis, Koutsias, Konstantinidis, & Vasilakos, 2007; Mitsopoulos et al., 2014; Moreira et al., 2010). Although the link is not straightforward, many references prefer to label this specific factor as land cover type. In fact, while fuel type presents a more detailed view of vegetation characteristics and is expressly connected to fuel flammability, land cover involves a more general and human presence related outlook.

As Calviño-Cancela et al. (2017) explain, the type of vegetation determines the flammability, density and continuity of fuels. The susceptibility of specific types of vegetation is well documented within the literature; shrublands and other light vegetation are particularly mentioned in connection to higher ignition risk and increased burned area (Fernandes, Monteiro-Henriques, et al., 2016; Ganteaume & Jappiot, 2013; Martínez-Fernández et al., 2013; Moreira et al., 2010; Nunes, Lourenço, & Meira, 2016; Salis et al., 2015). Nunes et al. (2016) and Martínez-Fernández et al. (2013) attribute this fact to a higher flammability and rate of fire propagation, as well as to decreased firefighting efforts given the low value of these areas, while Calviño-Cancela et al. (2016) attest a higher flammability of these areas in connection to great loads of ground fuel.

Additionally, forest land covers are mentioned in relation to fast-spreading wildfires and increased burned area (Calviño-Cancela et al., 2017; Fernandes, Monteiro-Henriques, et al., 2016; Ganteaume & Jappiot, 2013; Nunes, Lourenço, Bento-Gonçalves, & Vieira, 2013; Nunes et al., 2016), except for the case of native forest cover types. Calviño-Cancela et al. (2017) and Nunes et al. (2016) relate this phenomenon to the density of maritime pine tree (*pinus pinaster*) and eucalyptus (*eucalyptus globulus*) forests, for their specific characteristics favouring moisture loss, heat absorption and fire propagation. Among the most common vegetation species of the Portuguese forest, these two are considered the most prone to fire, followed by unspecified broadleaf and coniferous forests (Rego & Silva, 2014).

The simplification of the vegetation structure gives way to an increase in fuel loads and promotes the extent of burned area (Ganteaume et al., 2013). A mixed-fuel vegetation, on the other hand, encourages preferential burning, reducing the size of wildfires (Fernandes et al., 2016). However, the same authors who agree with this last statement suggest that the importance of fuel type or land cover composition might be mitigated by the overwhelming effect of favourable fire weather.

2.2.2.1.2. Density

Ganteaume et al. (2013) underline the role vegetation plays in fire hazard. While fuel availability represents the crucial component of any wildfire event (Srivastava, Saran, de By, & Dadhwal, 2014), high vegetation densities promote vertical and horizontal fuel continuity which in turn facilitates fire propagation and increases fire size (Fernandes, Monteiro-Henriques, et al., 2016; Nunes et al., 2013; Ricotta & Di Vito, 2014). This is particularly valid for Mediterranean ecosystems.

According to Ferreira-Leite et al. (2013), extreme wildfires which are particularly difficult to extinguish result from accumulated biomass and more fuel materials in forests, which are a direct consequence of socioeconomic changes in rural landscapes.

On the other hand, (Vasilakos et al., 2009) correlate fuel density to wildfire ignition. In fact, Nunes et al. (2016) propose that, in order to limit the development of shrub vegetation, prepare the land for agriculture and encourage forage renewal, farmers make deliberate use of fire.

The works of Dondo Bühler, de Torres Curth and Garibaldi (2013) and Martínez-Fernández et al. (2013) present contrasting views on this topic. The first study takes evidence from Bariloche, Argentina, where more vegetation availability is not a driver of wildfire occurrence, attributing a greater weight to anthropogenic factors. The second study relates fire density in Spain to fragmented landscapes, even though this result might also be connected to human presence.

2.2.2.1.3. Moisture

Vasilakos et al. (2009) describe fuel moisture as "the amount of water in a fuel particle", meaning that the higher the content the smaller the amount of fuel that is subject to combustion. Salis et al. (2015) attest the variability of this variable and its dependence on climatic conditions.

Therefore, vegetation moisture is connected to a decreased ignition risk (González-Olabarria, Mola-Yudego, & Coll, 2015), with the spatial distribution of lightning-caused wildfires depending decisively on fuel moisture (Ganteaume et al., 2013).

Similarly, there is ample consensus among experts that dry fuels help propagate fire and increase intense fire behaviour (Holsinger et al., 2016). Low dead fuel moisture content is specifically mentioned by Riley et al. (2013) as the main driver of surface fire, with extreme crown fires often resulting from large accumulations of dead fuels such as woody debris and litter.

2.2.2.1.4. Past wildfire activity

The effect of previous burns is expected to reduce the probability of wildfire occurrence and fire growth. Lack of burnable fuel after the wildfire event is the most straightforward explanation for this phenomenon (Holsinger et al., 2016; Nunes et al., 2016). However, as Fernandes, Monteiro-Henriques, et al. (2016) describe, small-scale pyrodiversity (i.e. the heterogeneity of past fire activity) is assumed to control fire size. Recent wildfire activity is likely to restrict fire spread for at least eight years (Fernandes, Pacheco, Almeida, & Claro, 2016), irrespective of climate, and is usually considered a proxy for vegetation accumulation (Fernandes, Monteiro-Henriques, et al., 2016).

Contrary to this belief, however, the work of Ricotta and Di Vito (2014) gives evidence of areas previously affected by wildfires displaying an increased recurrence probability. This phenomenon is related to the rapid post-fire development of susceptible vegetation, an increase of its continuity and spatial homogeneity and a proliferation of dead fuels.

Additionally, fuel age, which is also connected to vegetation build-up, influences fire spread, because fire is known to burn increasingly older fuels as it develops (Fernandes, Monteiro-Henriques, et al., 2016; Fernandes, Pacheco, et al., 2016)

2.2.2.1.5. Soil

In the same way vegetation conditions affect wildfire occurrence and propagation, soil characteristics have an impact on fire events. Sarris et al. (2014) consider limited soil moisture, in connection to climatic aspects, to have important consequences to wildfire behaviour. On the other hand, duff layer continuity (i.e. organic soil matter) promotes fire spread (Dimitrakopoulos et al., 2011). Besides moisture and organic matter, texture is also pointed out by Mhawej et al. (2015) as a factor driving increased wildfire probability.

2.2.2.1.6. Large wildfires

The occurrence of large wildfires is specifically connected to vegetation and soil factors by different authors. Dimitrakopoulos et al. (2011) relate an increased rate of large wildfire propagation to vegetation density and duff cover continuity, while asserting that this is true for both surface and crown fire types.

The studies by Fernandes, Monteiro-Henriques, et al. (2016) and Fernandes, Pacheco, et al. (2016), on the other hand, discuss the reputation of forests as the areas most affected by large wildfires. Pine and eucalyptus trees, along with evergreen oaks (*quercus suber*), vegetation types that are well present in the Portuguese landscape, are highlighted as particularly vulnerable species in connection with extreme fire events.

Additionally, large extensions of fuel connexion promote fire spread for long distances and there is specific evidence that wildfires with a burned area greater than 500 ha are largely reliant on fuel-related elements (Fernandes, Monteiro-Henriques, et al., 2016).

2.2.2.2. Climatic factors

Evidence from several studies, particularly in Mediterranean regions, establishes the importance of climatic factors in the analysis of wildfire patterns. Firstly, climate is well noted for its impact in shaping fire regimes in those areas (Ganteaume et al., 2013). Moreover, the impact weather exerts during fire events is deemed very relevant (Hernandez, Drobinski, & Turquety, 2015).

Nevertheless, some authors question the key role of weather conditions for wildfire incidence. Rodrigues et al. (2016) maintain the less influential effect of climatic factors in connection to fire frequency. In addition, Ricotta & Di Vito (2014) argue that the weather-fire relationship is particularly weakened in human-dominated landscapes, such as the Mediterranean. Weather aspects are subject to significant space-time variability (Keeley & Syphard, 2016), which might support such interpretations.

2.2.2.2.1. Temperature

Temperature represents one of the most immediate climatic factors controlling fire potential. For once, there is a long-established strong relationship between high temperatures and the probability of fire ignition (Ganteaume et al., 2013; Martínez-Fernández, Chuvieco, & Koutsias, 2013; Vasilakos et al., 2009). On the other hand, Hernandez et al. (2015) explain that extreme heat may also promote wildfire propagation by enabling fires to spread in many different directions, posing difficulties to suppression forces.

Interestingly, Keeley and Syphard's (2016) research has shown that while not strongly associated with the average yearly temperatures, burned area appeared to be connected to spring and summer temperatures in a significant manner.

A noteworthy feature connected to temperature and encouraging fire occurrence is the number sunshine hours. Either together with high temperature or acting alone, this factor promotes evaporation and vegetation dryness (Guo et al., 2016).

2.2.2.2.2. Precipitation

Research has demonstrated the importance of precipitation for the occurrence and propagation of wildfires, although Martínez-Fernández et al. (2013) restrict its high influence to dry areas. The effects of this element are twofold.

Initially, precipitation during the fire season is noted for its role limiting fire ignition and spread. The findings of many studies support this sentence by describing significant negative relationships between recent rainfall episodes and wildfire occurrence or propagation (eg. Balsa-Barreiro & Hermosilla, 2013; Martínez-Fernández et al., 2013; Vasilakos et al., 2009).

In contrast, off-season precipitation, months before the fire season, boosts the occurrence of wildfires, (Ganteaume & Jappiot, 2013; Martínez-Fernández et al., 2013). Ferreira-Leite et al. (2013) owe this phenomenon to the direct effect of rainfall on vegetation development. Nunes et al. (2013) share this view, by connecting the highest number ignitions to the rainiest areas in Portugal.

Relative humidity is closely related to precipitation, expressing fuel moisture, and shares most of its characteristics in regard to wildfire occurrence. Nevertheless, Vasilakos et al. (2009) has failed to find significance in relative humidity's influence in fire ignition, mainly due to the limited range of values this variable takes during summer in Greece.

Another phenomenon associated with precipitation activity is drought. An important characteristic shared by this variable, as well as other wildfire driving forces, is its spatial and temporal variability (Ganteaume et al., 2013).

There is a collection of studies establishing the significant effects of both long and short-term drought on wildfire occurrence and spread. As Riley et al. (2013) explain, the duration of the drought period determines its relationship to fire behaviour. According to their study, short-term precipitation deficits upset the moisture levels of dead fuels, while live fuels are instead affected by longer drought episodes. The death of vegetation, a consequence of extended periods of low to no precipitation, is a direct cause of fire intensity and propagation and is connected to an increase in crown fire probability. A positive relationship between drought episodes and burned area has also been proven by Keeley and Syphard (2016).

In their study, Riley et al. (2013) discuss the suitability of two drought indices for modelling wildfire incidence and spread: the Palmer Drought Severity Index (PDSI) and the Standardized Precipitation Index (SPI). Although the former is widely applied within the fire literature to account for drought periods driving wildfire events and burned area, the authors in this study suggest that large wildfires are rather connected to the SPI based on the 3 previous months.

2.2.2.2.3. Wind

Wind, on the other hand, is another determinant aspect included in the literature, with many authors indicating the influence of wind direction and speed particularly in connection to fire spread (Hernandez et al., 2015; Holsinger et al., 2016; Mhawej et al., 2015).

In fact, Dimitrakopoulos et al. (2011) believe it to be the leading climatic factor driving fire spread and relate strong winds to surface and crown propagation modes. However, the high spatial variability of wind speed and direction, inducing the formation of microclimates, is pointed out by Vasilakos et al. (2009) as a possible reason why these variables showed a negligible association with wildfire occurrence in the Greek island of Lesbos.

2.2.2.2.4. Large wildfires

Other authors claim that there is a strong connection between weather conditions and the duration of the fire event, relating the occurrence of large wildfires to climatic factors such as temperature and wind speed (Hernandez et al., 2015).

In fact, large wildfires are specifically mentioned within the literature in connection to a collection of climatic factors, which often display a significant influence in respect to area burned or large wildfire occurrence.

Dimitrakopoulos et al. (2011) analysed the weather conditions in which large wildfires propagate in Greece, underlining the effect produced by low relative humidity, very high temperatures, and strong northern winds, which are also mentioned by Hernandez et al. (2015) for the entire Mediterranean Europe. These specific meteorological conditions help produce wildfires with extreme fire behaviour, which override firefighting efforts, causing larger burned areas and increasing related losses (Fernandes et al., 2016; Salis et al., 2015).

On the other hand, the importance of high temperature and low precipitation extremes has been noted by Sarris et al. (2014), also in a Greek setting. Some of their conclusions relate to the role played by summer drought and extreme maximum summer temperatures in burned area extent. Ganteaume and Jappiot (2013) are also successful at determining higher likelihood of large burned areas during drought periods in southern France.

2.2.2.2.5. Climate-Vegetation link

A determinant aspect of climate shaping wildfire behaviour is its influence on vegetation. The moisture levels of dead and live fuels and the vegetation growth are equally controlled by short and long-term weather conditions (Hernandez et al., 2015).

Weather is related to vegetation conditions for its influence in humidity levels (Holsinger et al., 2016). In fact, as mentioned before, sporadic episodes of lack of precipitation, as well as pronounced dry periods, help to reduce the moisture content of live vegetation and increase dead surface fuel loadings, which strongly encourage fire ignition and spread immediately afterwards (Riley et al., 2013; Tedim et al., 2013). According to Nunes et al. (2013) and Hernandez et al. (2015), high temperatures produce a similar scenario, increasing fuel flammability and promoting the propagation of wildfires.

Additionally, the growth of vegetation translates into an increase in fuel availability. This phenomenon is particularly dependent on higher precipitation levels and lower temperature, with some authors relating this to the development of shrub cover and specific forest habitats (Nunes et al., 2013; Sarris et al., 2014). This apparent dichotomy is well stressed by Keeley and Syphard (2016), who verify the role of climate as affecting both fuel moisture and fuel volume.

The Canadian Forest Fire Weather Index System (FWI) (Van Wagner, 1987) is a good example of the role played by the underlying relationships between vegetation and climate in shaping fire danger. From a complete set of records of daily weather observations (temperature, rainfall, relative humidity and wind speed), three codes accounting for fuel moisture are calculated. Fire behaviour is then derived from two indices combining fuel moisture and weather information, indicating a measure of the risk of fire spread. An important output of this system is the daily severity rating, which relates to the necessary fire suppression capacity.

Among the indices comprised by the Fire Weather Index, Fernandes et al. (2016) emphasise the importance of the Initial Spread Index (ISI) and the Buildup Index (BUI), which account separately for the atmospheric and drought elements of fire occurrence and describe a measure of fire intensity. While the ISI infers fire propagation rate from weather conditions, the BUI reflects fuel consumption as a consequence of past precipitation values.

2.2.2.3. Topographic factors

Topography was found to be associated with wildfire risk by many authors, with Ganteaume et al. (2013) considering it one of the most determinant environmental factors driving wildfire occurrence in Mediterranean Europe. Although the significance and direction of this association varies among studies, several authors have mentioned the effect of different topographic features on burned area and ignition density, as revealed by Nunes et al. (2016).

2.2.2.3.1. Slope

Many references discuss the role of terrain slope in contributing to fire behaviour. First of all, Salis et al. (2015) explain that most of the importance of this factor lies on its effect on fire intensity. Additionally, landscape ruggedness, which is connected to steep slopes, plays a dual role in its connection to fire occurrence and burned area. This variable operates as a significant negative factor and its reverse, influencing accessibility and affecting human caused fires and firefighting efforts (Ganteaume & Jappiot, 2013).

2.2.2.3.2. Aspect

The impact of aspect on wildfire ignition is associated with climatic factors such as temperature, solar radiation and wind. In fact, north oriented slopes in the northern hemisphere are characterised by a diminished flammability because they receive less sunlight which translates into cooler and more humid environments (Calviño-Cancela et al., 2017). Vasilakos et al. (2009) go as far as to sustain a rise in fire risk during the day, depending on the different aspects. These authors also distinguish between types of fires according to main orientation, with south facing slopes enduring the largest number of ignitions and north facing slopes being prone to increased fire intensity, mainly due to favourable fuel conditions.

2.2.2.3.3. Elevation

Elevation is evident among topographic factors as that more clearly linked to wildfire ignition. A higher density of fire occurrences at low elevations might be connected to different aspects of human activity (Guo et al., 2016). Nevertheless, this same reason might explain frequent wildfires at higher altitudes, largely affected by the use of fire for grazing purposes (Moreira et al., 2010).

Vasilakos et al. (2009) confirm a higher fire hazard at lower areas and a particular link to climatic conditions at higher altitudes, affecting fire behaviour through time. An increase in the number of ignitions at low elevations is also pointed out by Calviño-Cancela et al. (2017) and by Ricotta and Di Vito (2014), this last study specifically focusing fire recurrence. All in all, elevation is known to display opposing relationships with wildfire behaviour, with studies showing either a positive impact motivated by lightning frequency at high altitudes or a significant positive relationship connected to vegetation dryness at low elevations (Martínez-Fernández et al., 2013).

2.2.2.3.4. Other factors related to topography

The importance of topographic elements, and most of all elevation, for the study of wildfires is due to their prominent connection to an array of different environmental and human factors. For instance, the elevation-climate link is supported by the study of Calviño-Cancela et al. (2017), which associates higher elevations to a drier and windier environment.

Additionally, vegetation conditions are a determinant aspect in connection with the landscape topographic features. As Holsinger et al. (2016) put it, topography is related to vegetation moisture levels and fuel concentration. Areas of steep slopes and uneven surfaces, which are less suitable for agricultural activity, promote forest and shrub vegetation density (Calviño-Cancela et al., 2017; Nunes et al., 2013). The highest mountain elevations, on the other hand, present high fuel humidity values, which limit fire occurrence (González-Olabarria et al., 2015).

Finally, the occurrence and intensity of wildfires is dependent on topography for its marked effect on human activities (Calviño-Cancela et al., 2017). Among all human aspects disturbed by elevation and slope, accessibility is one of the most important for the study of wildfires. First of all, as Calviño-Cancela et al. (2017) describe, lower elevations are associated with an increased number of ignitions as these areas are generally more accessible, a fact that is deemed particularly relevant for the occurrence of deliberate fires. However, higher elevations may produce larger and more intense fires due to a greater propagation risk, connected to the complexity of firefighting operations in these areas.

Accordingly, lower altitudes are also the place of greater human and infrastructural densities and intensive land use, either for agriculture or other activities, with an increase in human fire impact (Ganteaume et al., 2013; Martínez-Fernández et al., 2013; Ricotta & Di Vito, 2014).

2.2.2.4. Human factors

According to the works by Balsa-Barreiro and Hermosilla (2013) and Nunes et al. (2013), human factors have been identified as some of the main drivers of wildfires, particularly in the last decades, influencing wildfire intensity, recurrence and the unbalanced density and distribution of ignition spots.

These factors act in combination with the different sets of other identified elements to shape fire behaviour. Rodrigues et al. (2016) maintain the importance of this link in the light of specific human-related variables losing explanatory power in the models, although there is conflicting evidence that these same variables are slowly prevailing over the biophysical factors in determining fire incidence (Ganteaume et al., 2013).

Human-related factors have a distinct influence on fire occurrence and spread, depending directly on wildfire cause (Ganteaume & Jappiot, 2013). Naturally, human factors are directly associated with human-caused wildfires and their great significance lies in the overwhelming preponderance of these type of fire events in the Iberian Peninsula (Romero-Calcerrada, Barrio-Parra, Millington, & Novillo, 2010).

Above all, it is important to note that human factors, unlike other identified aspects, are predominantly non-stationary in time and space (Rodrigues et al., 2016).

2.2.2.4.1. Population

As revealed by Oliveira, Pereira, San-Miguel-Ayanz and Lourenço (2014) and Rodrigues, de la Riva and Fotheringham (2014), human-caused fire events are linked to the importance of human presence and density as predictors of fire occurrence. This view is supported by Nunes et al. (2013) who agree that an increase in population translates into greater fire risk.

Population density is deemed by many authors an important force driving fire frequency, specifically in southern Europe (Balsa-Barreiro & Hermosilla, 2013; Dondo Bühler et al., 2013; Ganteaume & Jappiot, 2013; Mhawej et al., 2015; Oliveira et al., 2014; Romero-Calcerrada et al., 2010). This variable is usually considered a measure of human pressure (Álvarez-Díaz et al., 2015) and is related to the presence of urban areas (González-Olabarria et al., 2015). Additionally, Rodrigues et al. (2016) attest the increase in influence this driving factor is displaying nowadays in the case of Spain.

It is relevant to consider the contrasting implications of human presence and density on wildfire behaviour. There is a broad consensus in the literature about the double role of population as a wildfire enabler and detractor. While human pressure is known to increase ignition risk, it is also believed to promote reduced burned areas as a result of less available fuel, more fire prevention efforts, easier accessibility, earlier detection of fire events and a more effective firefighting (Dondo Bühler et al., 2013; Ganteaume et al., 2013; Moreira et al., 2010; Oliveira et al., 2014).

Conversely, the issues of ageing and depopulation are also mentioned in connection with general wildfire patterns (Dondo Bühler et al., 2013; Tedim et al., 2013). Martínez-Fernández et al. (2013) discuss at length the significant effect of these factors on wildfire behaviour, mainly due to fuel accumulation and hazardous agricultural practices, confining them to rural areas. These areas are affected by a demographic decline, intensified by land abandonment, while urban areas signal positive population growth rates, including the balancing effect of migration inflows (Balsa-Barreiro & Hermosilla, 2013).

2.2.2.4.2. Human activities and infrastructures

Proximity to human activities and infrastructures is considered by Mhawej et al. (2015) an important force driving fire ignition. Firstly, as explained by Balsa-Barreiro & Hermosilla (2013) for the Spanish region of Galicia, population density is greater in proximity to urban areas and infrastructures. Furthermore, negligent and accidental wildfires, which make up a large portion of all fire occurrences in the densely populated region of the Mediterranean, are related to a diverse range of human activities, such as agriculture, forestry, camping and animal grazing (González-Olabarria et al., 2015). Contrary to this belief, however, there is evidence suggesting that fire occurrence is becoming less reliant on human activities in recent years (Rodrigues et al., 2016).

Oliveira et al. (2014) highlight the importance of non-wildland areas, such as human-managed agricultural and forest lands, as locations susceptible to wildfire occurrences. Although other classes of vegetative cover, where fuel conditions are more favourable, present higher fire risk, a considerable amount of ignitions arise on non-wildland areas in connection with human activities.

Among human activities, agriculture has been deemed particularly significant in regard to both ignition probability and burned area (Álvarez-Díaz et al., 2015), even if describing opposing trends (Nunes et al., 2016; Rodrigues et al., 2016). Actually, Rodrigues et al. (2016) argue that agricultural activities and fire frequency are still positively linked, specifically in the European Mediterranean region, although the significance of this aspect displays a non-stationary behaviour across territorial units.

There are several well-recognised reasons accounting for these positive patterns. One of these is the use of fire connected to farm waste clean-up activities (Grala et al., 2017; Rodrigues et al., 2014). The age of farmers and rural populations in general is often considered a reliable proxy variable, mirroring the employment of traditional agricultural management practices (Ganteaume et al., 2013). This aspect is still considered to be a determinant cause of wildfires in southern European regions (Álvarez-Díaz et al., 2015).

Accidents related to the operation of agricultural machinery, particularly near forested areas, are equally mentioned by some authors (Martínez-Fernández et al., 2013; Rodrigues et al., 2014). Indeed, the issues connected to agricultural activity encouraging fire occurrence are aggravated by the proximity to forest stands. As Rodrigues et al. (2016) explain, the wildland-agricultural interface (WAI) encompasses the border between land used for agricultural purposes and wildland areas. In these locations, there is an additional risk of human-caused fire events spreading to nearby areas with dense fuel loadings.

The burning activities conducted in rural areas are also mentioned in connection to animal grazing (Álvarez-Díaz et al., 2015; Rodrigues et al., 2014). Firstly, this phenomenon is linked to the process of gaining or preserving cattle grazing through fire (Ferreira-Leite, Lourenço, et al., 2013; Vilar et al., 2016). Secondly, as Balsa-Barreiro and Hermosilla (2013) discuss, fire is equally used as a tool to reduce predatory wildlife and promote herd management and control.

Nevertheless, livestock presents mixed results throughout the literature as to its influence on fire ignition and behaviour. It has been found to have a negative association with ignition points (Romero-Calcerrada et al., 2010; Romero-Calcerrada, Novillo, Millington, & Gomez-Jimenez, 2008), a perspective supported by Oliveira et al. (2014), who understand animal grazing as a powerful fuel management tool, particularly effective in less accessible areas. These authors also discuss the importance of discriminating among different animal types when studying fire occurrence, for their specific characteristics.

Human activities related to forest stands represent a great economic value, as can be attested by the importance of the eucalyptus timber business (Álvarez-Díaz et al., 2015). These areas are under constant economic pressure and generally face an absence of steady and effective management, both of which incite wildfire occurrence (Balsa-Barreiro & Hermosilla, 2013). In fact, the type of land ownership is significantly associated with fire behaviour, with declining conservation accounting for an increase in human-caused fires. Martínez-Fernández et al. (2013) explain that private forest land in

Spain is mostly subject to less control, as opposed to community managed areas, and therefore liable to a greater wildfire risk.

The environmental value of protected sites, on the other hand, plays a similar role in reducing fire risk. Protected forested areas have been found to contribute as a deterrent factor, preventing fire ignition and propagation (Rodrigues et al., 2014). This link has become more pronounced in recent years, showing a negative relationship between the degree of protection and the amount of wildfire events (Rodrigues et al., 2016).

These authors are not alone in describing the limiting effect of human activities for wildfire incidence. Nunes et al. (2016) associate agriculture, forestry and grazing with a reduction of fire ignitions and burned area, whereas Oliveira et al. (2014) discuss a negative association between the amount of cultivated land and fire occurrence. Moreover, Martínez-Fernández et al. (2013) refer to the high fragmentation and uneven distribution of farms and other land properties as a significant factor decreasing human-caused fire hazard throughout Spain.

The abandonment of rural and agricultural areas is related to Nunes et al. (2016) experience of traditional wildland activities such as agriculture and livestock being an agent of a decrease in fire frequency and burned area. In fact, the depopulation of large country areas and the subsequent disappearance of these activities leads to an accumulation of fuel, determinant in the occurrence and spread of human and natural-caused wildfires alike (Ganteaume et al., 2013; Nunes et al., 2016; Rodrigues et al., 2014; Vilar et al., 2016). Other factors such as the change in energy sources and specific forest and agriculture-related European and national policies have further intensified this trend (Nunes et al., 2013).

In many cases, this land use-land cover transformations have been mirrored in the development of areas of extremely flammable vegetation types and a general loss of species diversity, with serious consequences for wildfire patterns (Tedim et al., 2013). However, Martínez-Fernández et al. (2013) found that, despite intense depopulation of rural areas, regions with a higher number of farmers are increasingly affected by wildfire events.

This rural flight phenomenon has been coined as rural exodus in many regions. In Portugal and Spain rural depopulation has been in place since 1960, mostly in inland and mountainous areas, and has been characterised by an abandonment of agricultural and forested areas of both high and low productivity, implying a socioeconomic regression (Almeida, Nunes, & Figueiredo, 2013; Balsa-Barreiro & Hermosilla, 2013; Ferreira-Leite, Lourenço, et al., 2013).

A decrease of population in rural areas during the last decades meant instead a significant increase of human concentration in urbanised regions, mostly by the coast in the case of mainland Portugal (Moreira et al., 2010). This phenomenon of increased urbanisation is directly linked to urban growth and the ensuing increase of the wildland-urban interface area (WUI), in the words of Rodrigues et al. (2016) and Romero-Calcerrada et al. (2008).

In general terms and as previously mentioned, the pressure of humans on the environment, especially in forested and other wildland areas, is seen as the central hazard posed by population to fire regimes in the Mediterranean (Nunes et al., 2016). The WUI denotes a dispersed system of human settlements bordering wildland areas (Rodrigues et al., 2014) and is considered a central aspect in most studies focusing on the human impacts on wildfire occurrence, with some authors arguing that its contribution to fire incidence has increased over time (Vilar et al., 2016) and others discussing the loss of predictive influence (Rodrigues et al., 2016).

Indeed, housing density is usually mentioned on the subject of the main human factors driving wildfire occurrence (Álvarez-Díaz et al., 2015). Moreover, the density of secondary housing has been found to have a significant relationship with fire ignition spots (Romero-Calcerrada et al., 2010, 2008).

Another factor representing increased human pressure on natural areas is the intensification of recreational activities, including a growth in the number of tourists, which is particularly valid in Mediterranean regions (Ganteaume & Jappiot, 2013). The works by Romero-Calcerrada et al. (2010, 2008) are successful at linking wildfire ignition and a proximity to recreational areas and camping sites in Spain, while underlining the significant role of these aspects during the height of the fire season, given an increase in human pressure during summer months. These driving factors are related to the occurrence of negligent fires (Moreira et al., 2010).

Vilar et al. (2016) also suggest that the services workforce is a significant determinant of wildfire occurrence, although it is not clear if this is a confounding predictor, related instead to the presence of leisure activities and population density. The proximity to industrial sites is also mentioned by some authors (Romero-Calcerrada et al., 2010), although further studies are necessary to evaluate the concrete influence it exerts.

As Ganteaume et al. (2013) describe, spatially explicit variables are of utmost importance for an all-inclusive assessment of human-caused wildfire events. The distance to infrastructures is considered a meaningful factor in regard to fire incidence (Álvarez-Díaz et al., 2015). Among these, communication infrastructures such as roads and railways are fully covered by the literature on this topic.

There is ample agreement that proximity to roads has a favourable impact on the probability of fire ignitions, mainly in connection with high accessibility and strengthened human presence (Martínez-Fernández et al., 2013; Moreira et al., 2010; Oliveira et al., 2014; Ricotta & Di Vito, 2014; Rodrigues et al., 2016; Romero-Calcerrada et al., 2010, 2008). Although mentioning that distance to communication networks are usually considered within the scope of fire risk modelling, Ganteaume et al. (2013) suggest that small road density increases the likelihood of wildfires in places where human interference in this phenomenon is limited.

Additionally, proximity to railroads is connected to a similar positive impact on fire incidence (Rodrigues et al., 2014; Romero-Calcerrada et al., 2010). This effect is mostly explained by accidents producing fire ignitions (González-Olabarria et al., 2015). In fact, negligent and accidental causes are numerous among factors contributing to wildfire ignition. Power lines are a concrete example of human infrastructures producing fire ignitions from electric discharges (Oliveira et al., 2014; Rodrigues et al., 2014; Rodrigues et al., 2016).

2.2.2.4.3. Socioeconomic aspects

The study conducted by Grala et al. (2017) proved an association between the population's socioeconomic characteristics and the occurrence of specific types of human-caused fires. Economically deprived regions were found to be more prone to wildfires caused by campfires, children, debris burning, smoking, equipment use and railroads than to incendiary fires. The case of higher income areas being linked by Grala et al. (2017) to arson ignitions needs further scientific evidence. Rodrigues et al. (2016) sustain that the study of incendiarism related variables might provide new insights into human-caused wildfire patterns.

Similarly, Dondo Bühler et al. (2013) refer to socioeconomic vulnerability as an important force driving wildfire ignition in their Argentinean case study. Conflicts may arise from economic difficulties promoting an increase in arson events. Furthermore, high unemployment rates might encourage incendiarism either as an indication of dissatisfaction or as a means to provide a job for firefighters and other seasonal workers.

The same study also describes the influence of two other socioeconomic aspects on wildfire behaviour. First of all, the community's educational level associates negatively with fire ignition because more educated individuals are expected to present a greater understanding of the value of natural resources, to be more aware of the effects of wildfires and to have access to more information on preventive measures. Secondly, house ownership is linked to increased fire density, because of its underlying connection with unemployment in the study area. Nevertheless, the authors agree that this finding goes against the results from previous studies, which associate house ownership to building conservation and the maintenance of surrounding areas, and therefore to a decreased probability of wildfire occurrence

Other socioeconomic aspects related to political activity (e.g. raising anxiety in politically intensive contexts) and criminality (e.g. police arrests) have also been associated with wildfire occurrence (Álvarez-Díaz et al., 2015).

2.2.2.4.4. Firefighting

Firefighting represents a different human aspect influencing wildfire behaviour and spread (Cao et al., 2013). Suppression efforts, which entail both ground and airborne firefighting forces, have got a definite effect on wildfire development, as mentioned by several authors (Dimitrakopoulos et al., 2011; Moreira et al., 2010; Tedim et al., 2013). Additionally, Moreira et al. (2010) and Dimitrakopoulos et al. (2011) underline the importance of response times to the success of these efforts.

The role of firefighting as an aspect increasing wildfire severity is discussed by Tedim et al. (2013). The rapid suppression of small wildfires brings about an intensification of vegetation growth, promoting landscape flammability and increasing the risk of extreme fire events.

2.2.2.4.5. Large wildfires

Moreira et al. (2010) are successful at associating human factors and burned area extent in the Portuguese mainland, focusing on the role played by population density, distance to roads and land use-land cover.

The importance of population density as an anti-causative factor of large burned areas is well stressed by these authors, who assert that an increase in human presence results in reduced likelihood of fire propagation. Moreover, ignitions occurring far from roads are related to an increased risk of developing into medium/large fire events, given a combined effect of low accessibility and retarded fire detection.

Other human induced trends, such as the decline of agricultural activity and the rise of shrublands and other unmanaged vegetation covers, discussed over the previous chapters, also display a clear impact on wildfire behaviour, promoting fire spread, as ignitions resulting in larger burned areas occurred preferentially in these land cover types. Landscapes exhibiting mixed land use types, particularly urban-rural areas, on the other hand, are at less risk of large wildfire occurrence.

2.2.3. Modelling wildfire danger

Modelling the probability of fire occurrence and spread and mapping areas of increased hazard has been a recurrent topic for researchers in this discipline.

The temporal scales of wildfire risk assessment are important to consider in a comprehensive review of methodologies, specifically what concerns the difference between short-term and structural estimations. In fact, the main objectives of these perspectives are contrasting in nature. Whereas short-term fire risk assessment is mostly connected to the operational planning of suppression activities and the issuing of fire alerts and rely mainly on weather data, long-term exercises intend to inform strategic decisions, such as resource management and regional planning, making use of a diverse array of structural contributing factors (Yakubu, Mireku-Gyimah, & Duker, 2015).

The previously mentioned Fire Weather Index (FWI) is a particular case of fire risk assessment, dependent on a system of equations integrating information from a different set of weather related codes (Van Wagner, 1987). The Portuguese system of fire risk warnings combines the FWI and the structural risk, attributed to permanent topographic and vegetation conditions, into a wildfire risk index (RCM), as described by IPMA and ICNF.

On the contrary, the assessment of the fire hazard component of structural fire risk entails the combination of many different factors and responds successfully to a great range of statistical and data analysis procedures, from each we can count regression models, data mining techniques and statistical tests. Other non-statistical methodologies, such as integrated multi-criteria decision-making analysis, are also pointed out as viable solutions by some authors, although they stand beyond the scope of this work and therefore are not considered in this review.

The suitability of generalized regression techniques, namely logistic regression, has been long-established in the field of environmental studies for its ability to model binary response variables (Kalabokidis et al., 2007). For wildfire risk assessment, where several different factors play into the modelling process, this statistical technique is widespread (Cao et al., 2013).

Logistic regression is considered a flexible tool, allowing for the integration of numerous continuous and categorical variables (Yakubu et al., 2015). As an example, the work developed by Grala et al. (2017) makes use of multinomial logistic regression to analyse the influence of different socioeconomic and other human-related elements in human-caused wildfire occurrence in Mississippi, USA. In fact, human-caused wildfires are commonly modelled with the help of logistic regression (Rodrigues, 2015).

Nevertheless, some authors maintain that the assumptions of logistic regression regarding spatial stationarity are not appropriate for the study of large geographical areas since models are unable to adequately reveal the different underlying regional dynamics (Nunes et al., 2016). This statement is especially true for the study of wildfire occurrence, but also its human driving factors, as Rodrigues et al. (2016) and Oliveira et al. (2012) reveal, for the case of Spain and Portugal respectively.

These works integrate the findings of Brunsdon, Fotheringham and Charlton (1998) by applying a geographically weighted regression (GWR). This statistical technique incorporates spatial non-stationarity into model parameter estimation so that the diverse contribution of predictors can be measured and model errors can be assessed over space, making use of geographic weights attributed to neighbours on a basis of distance to location being estimated.

Therefore, GWR is different from a simple moving window regression, whose result is simply a set of local regression statistics, because it assigns different weights to all observations inside the selected bandwidth. Choosing this neighbourhood is a key step for the development of any GWR model and an optimal solution can be reached through the minimization of the errors (Rodrigues, 2015).

The regional variation of wildfire drivers, specifically the case of human-related explanatory factors, has been deemed of extreme importance and taken into account by Chuvieco, Aguado, et al. (2014) in their estimation of human influence on wildfire occurrence in Spain. They developed a geographically weighted logistic regression (GWLR) model as part of a broader strategy to map integrated wildfire risk in this region.

However, Vilar et al. (2016) suggest that generalised linear models might not be applicable at all in the particular case of studies making use of human factors, as these tend not to follow a known statistical distribution, proposing nonparametric techniques like Maxent instead. This machine learning technique works by maximizing entropy in approximating a uniform distribution to empirical data (Phillips, Anderson, & Schapire, 2006).

According to Rodrigues (2015), machine learning methods have a great predictive accuracy in the field of data mining, meaning they perform well with a great quantity of input data. Rodrigues and de la Riva (2014) employed random forests, boosted regression trees and support vector machines for modelling human-caused wildfire occurrence and compared their performance to that of logistic regression.

Methodological concerns in the development of predictive models are discussed by some authors. First of all, it is important to realise that the assessment of all relevant factors connected to wildfire occurrence and burned area is unrealistic (Kalabokidis et al., 2007). Holsinger et al. (2016) mention the robustness of logistic regression dealing with confounding variables, which can prove to be an obstacle in the identification of major driving forces, while also warning of spatial autocorrelation in the data. The issues of collinearity are also pointed out, with Nunes et al. (2016) underlining the importance of this analysis at the local level. Although not influencing the coefficient estimates, a correlation among variables generates an increase in their standard error (Rodrigues, 2015).

Many of the models developed for wildfire risk assessment have taken into consideration solely biophysical and other environmental factors. However, many authors advocate the inclusion of anthropogenic predictor variables alongside biophysical factors as a way to improve model performance (Rodrigues et al., 2016; Vilar et al., 2016). Additionally, as we have previously noted, since the majority of Mediterranean wildfires is human-caused, explanatory variables connected to population and socioeconomic activities display a key impact on this phenomenon by increasing fire frequency and changing its spatial distribution (Grala et al., 2017; Nunes et al., 2016).

3. METHODOLOGY

This chapter presents and describes in detail the data and methods employed in the development of this research work. To begin with, the territorial boundaries of the study area are defined, and the main biophysical and socioeconomic characteristics of this region are briefly portrayed. Secondly, the process of data handling and integration is summarised, and retained variables for analysis are presented, along with corresponding data sources. Lastly, all methodological aspects concerning the proposed study are discussed, namely the methods applied in exploratory and statistical analyses.

3.1. STUDY AREA

The study area of this research is limited to the Portuguese mainland territory contained between latitude parallels 39°30'0''N and 40°40'0''N. This strip of land covers most of the area corresponding to NUTS II "Centro", which is an approximate match to the central region of Portugal, as well as a small portion of northern Alentejo (NUTS II "Alentejo"). Altogether, this region encompasses the entire area of NUTS III "Coimbra" and "Beira Baixa", and parts of "Aveiro", "Leiria", "Viseu Dão-Lafões", "Médio Tejo", "Beiras e Serra da Estrela", "Alto Alentejo" and "Oeste", which make up 1 004 parishes (before the 2013 administrative reform) and represented around 1,6 million inhabitants, in 2011.

Although Portugal cannot be considered a particularly mountainous country, with less than 1/8 of its territory rising above 700m, the central region is home to some important mountain ranges. The study area makes up an altitude average of 364 m, from the low-lying western coastal extents to the Serra da Estrela ridge (2000m, to the northeast), which marks the highest point in mainland Portugal. This area is also characterised by a Mediterranean climate (Köppen-Geiger classification *Csa* and *Csb*) (Kottek et al., 2006), like the rest of the Portuguese territory, although there is a noticeable contrast in annual precipitation totals between northern coastal and southern inland areas.

Wildland areas are particularly relevant to consider in the context of wildfire research. In Portugal, forests are the predominant land-use type, corresponding to approximately 35% of the total area, and are mainly composed of eucalyptus plantations, cork and holm oak forests and maritime pine trees (ICNF, 2013; Rego & Silva, 2014). Around 75% of the entire forest area is owned by non-industrial private holders, specifically in northern and central Portugal, which creates an obstacle to efficient forest management (Ribeiro et al., 2015). Agricultural land makes up about 24% of the total area of the country, shrublands and grasslands about 32%, and only 5% is taken up by urban spaces (ICNF, 2013).

According to the national land cover and land use cartography (DGT, 2010), in 2010, approximately 65% of the study region accounted for forest areas. The main species composing these forests were pine trees (38.5%), followed by eucalyptus (24.5%), evidencing a divergence with national data. In fact, whereas pine tree forests were relatively abundant throughout the study area, and particularly in the centre, the north-eastern part of the study region (Guarda) displayed virtually no eucalyptus presence.

On the other hand, shrublands were very common in this north-eastern area (14% overall), and forests of other species (predominantly oak trees and sclerophyllous vegetation) were extensive mainly along the border (23% overall). Protected areas were also frequent all over the inland extent of this region, with three important natural parks, one natural reserve, one protected landscape and one natural monument falling on this half of the study area (in total, there are 12 protected areas intersecting the territory under analysis). Grazing grounds, which made up roughly 20% of the entire area, were mostly restricted to the south-eastern portion of the study region (Castelo Branco and Portalegre districts).

Two different reasons explain the selection of this area as the focus of the research work. First of all, the study area has been affected by extreme wildfire events, including the recent and infamous megafire of Pedrógão Grande, which occurred in June of 2017 and caused unprecedented damage, including 65 fatalities and a burned area of 45,328 ha (Tedim et al., 2018). This makes it an interesting location for studying the specific dynamics of large wildfire occurrence.

Additionally, there is a great variation in landscape throughout the territory, which is apparent in forest and vegetation characteristics, diverse human activities and type of settlements, as well as other biophysical conditions such as elevation and climate. In fact, this region includes both coastline and inland features and is considered a borderline area between northern and southern Portugal, displaying many of the contrasting Mediterranean and Atlantic features that characterise this country (Ribeiro, 2011). This spatial variability provides a rich insight into the factors driving large wildfire ignition and spread, depending on location.

On the other hand, the decision concerning the study period (2005-2015) stems from the need to provide the analysis with recent and coherent data. Relevant data sources in the context of this research work include the last population and housing census (INE, 2011), the last agricultural census (INE, 2009) and the 2010 exercise of national land cover and land use classification (COS 2010) (DGT, 2010), which represent the years falling in the middle of the reference period.

3.2. DATA COLLECTION AND PROCESSING

The proposed objectives of this research require the gathering of extensive data on the structural factors driving wildfire ignition and spread. For the purpose of this study, some of the variables that have been presented in the context of the literature review have been left out, either because of a lack of agreement among authors, a higher effort in collecting data or a general impracticality in measuring a given phenomenon. A complete list of all identified factors and corresponding scientific references can be found in the Annex (Annex A: Table A1), as well as an indication on whether the underlying phenomenon has been considered for the subsequent analysis or not. In any case, 58 variables have been retained, including four describing wildfire events and burned area.

The spatial framework of the proposed investigation lies on the European ETRS89-LAEA 1x1km Inspire grid, taking evidence from previous analysis in this field of study and benefiting from standardisation opportunities. In fact, the adoption of a 1x1km grid in wildfire research is not new and has been deemed appropriate in other contexts (Rodrigues, 2015; Vilar, Nieto, & Martín, 2010).

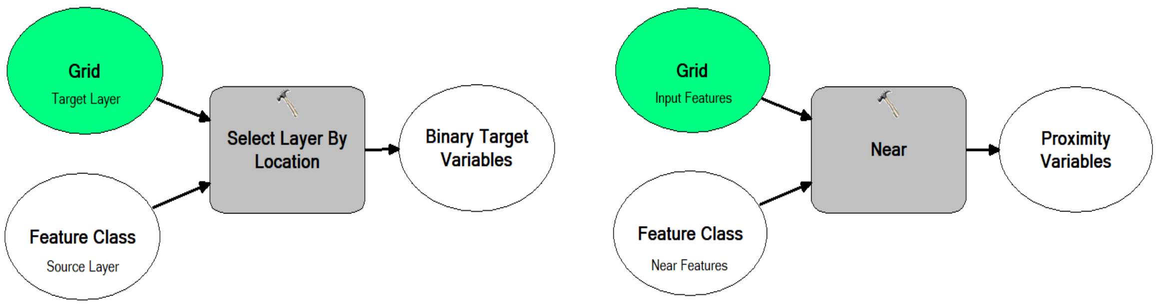
This spatial data infrastructure is a feature class based on the Lambert Azimuthal Equal Area and ETRS89 projections, with the centre of the projection at the point 52°N, 10°E, and false easting: x0 = 4321000 m, false northing; y0 = 3210000 m. This grid has been developed in the context of the Inspire Directive for spatial analysis and statistical reporting purposes and is composed by 1x1km cells. It has been used by the Portuguese National Statistical Office (INE) for the dissemination of specific census variables, namely: total population, total buildings and total housing (data that have been retained for the analysis).

The study area (central Portugal) corresponds to 21 570 unique grid cells. Variables have been transposed to the 1x1km grid framework following different methodologies. The workflow of data processing depends mainly on file format, but also on the type of variable. Data processing and integration was performed with the help of ArcMap 10.5.1 GIS software (ESRI), MS Excel and MS Access (Microsoft). It is possible to identify six different integration procedures:

i) **Shapefiles (binary target variable):** performing a select by location, on ArcGIS, of all grid cells that intersect the source feature classes (fire ignitions).

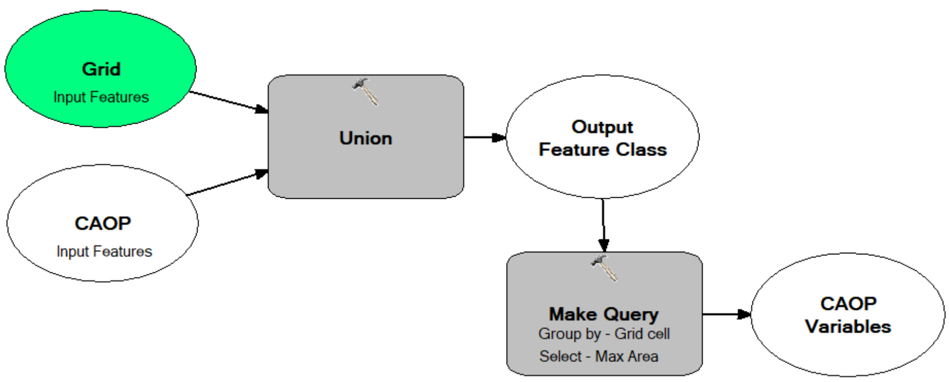
ii) **Shapefiles (proximity variables):** distance between the limits of the grid cells and the nearest shapefile feature are calculated using the analysis tool *Near*, on ArcGIS, which creates new attribute columns on the feature class data frame indicating the distance in meters to the proximity feature and its ID.

Figure 3 – Data integration for binary target (i) and proximity variables (ii) (respectively)



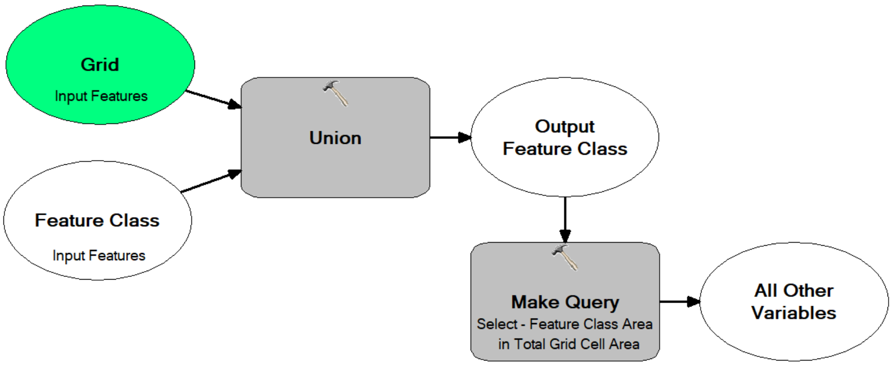
iii) **Shapefiles (administrative units):** correspondence between grid cells and the smallest Portuguese administrative unit (parishes) is retrieved by computing the intersection between feature classes (parishes and grid) by using the overlay tool *Union*, on ArcGIS, and selecting the corresponding parish to the largest area of each grid cell, on MS Access.

Figure 4 – Data integration for variables related to administrative units (iii)



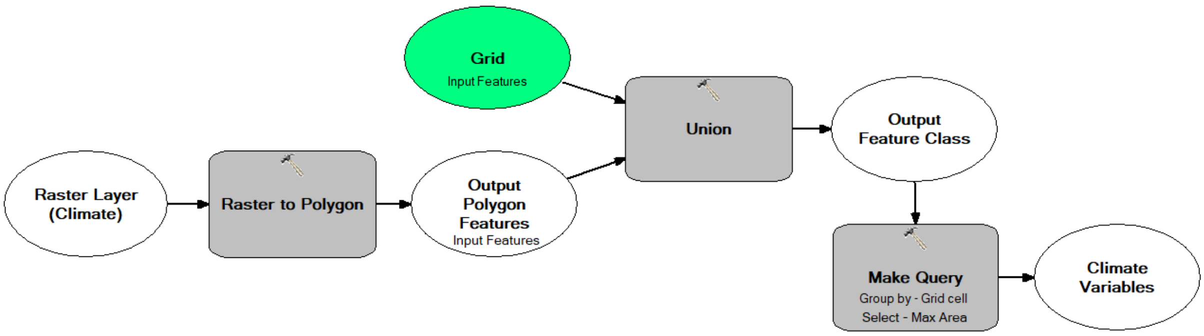
iv) **Shapefiles (other variables, including one continuous target variable):** calculation of the percentage of each feature class on each grid cell, on MS Access, by performing a *Union* of input variable and grid, on ArcGIS**.**

Figure 5 - Data integration for other variables in shapefile format (iv)



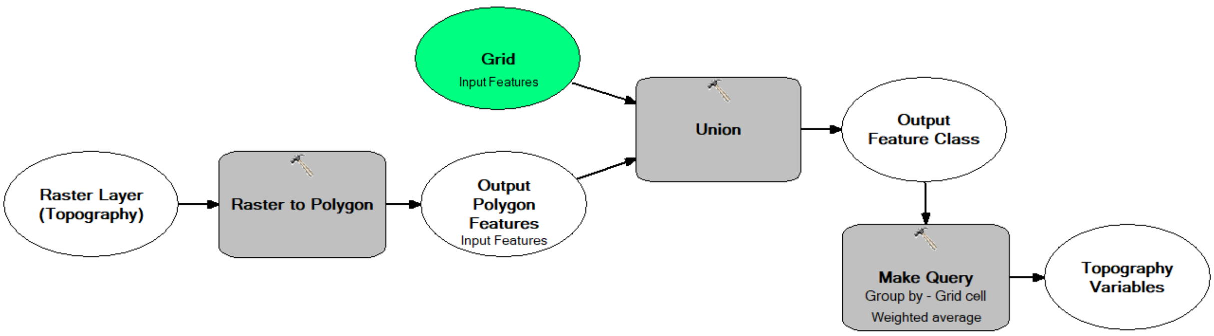
v) **Raster files (interpolated climate variables):** on ArcGIS, conversion of interpolated raster files to polygon feature shapefiles (integer values), then using the tool *Union* for calculating the intersection between polygons and the grid feature class, and finally selecting for each grid cell, on MS Access, the value corresponding to most of the cell's area.

Figure 6 – Data integration for interpolated climate raster variables (v)



vi) **Raster files (topographic variables):** following the first two steps of the previous procedure (raster to polygon conversion and *Union*) and computing an average of the unique values of each portion of the grid cell, weighted by the percentage of the corresponding area (MS Access). This procedure includes the climate variable PRECTOT (mean annual precipitation), in addition to all topographic factors.

Figure 7 – Data integration for topographic raster variables and mean annual precipitation (vi)



3.2.1. Wildfire events and burned area

The Decree n. 124/2006, of June 28th (Decreto-lei n.º 124/2006, de 28 de junho in Diário da República n.º 123/2006, Série I-A), which establishes the main measures to be implemented under the National System of Forest Protection against Wildfires (SNDFCI) and has been since revised by four Decrees and the Law n. 76/2017, of August 17th (Lei n.º 76/2017, de 17 de Agosto in Diário da República n.º 158/2017, Série I), imparts on the Portuguese Institute for Nature Conservation and Forests (ICNF) the responsibility of maintaining a national scale wildfire database connected to a geographic information system. This database assembles information from several sources, from wildfire to budget management systems, providing the necessary data for devising pre and post-fire vigilance, detection, firefighting and supervision strategies and processes.

Part of these data are available to the public on ICNF's website, namely a complete list of all annual wildfire events with information on ignition location, dates and times for alarm, first intervention and extinction, total burned area and ignition cause, as well as shapefiles with burned area extensions for each year, extracted from satellite imagery and aerial photographs. For the purpose of this study, only fires larger than or equal to 100 ha were retained.

The large fire ignitions binary target variable was taken from the intersection of these data with the grid, with 1 indicating at least an ignition, and 0 representing no ignition. The generalisation of this variable has been computed from these data: adjacent cells to ignition locations, which correspond to a 3x3 grid cell window, have been retained given the uncertainty of ignition locations and the reduced number of target observations, which might compromise the proposed analysis.

Additionally, proximity to ignition locations has been computed in the ArcGIS environment using the analysis tool *Near*, which determines the distance between the boundaries of the grid cells and the nearest ignition point. Burned area cells, on the other hand, were translated into a percentage of the affected grid cell, for target variable purposes, following the methodology explained above (Figure 3).

Incomplete data make it difficult to connect wildfire occurrence records with burned areas in some cases. However, this link is not considered relevant for the objectives of this study and the data are considered trustworthy given their source and the fact that they are subject to ongoing validation procedures. During the study period, 378 ignitions that developed into large wildfires occurred in the study area, which translate into 2,810 grid cells (considering adjacent cells). Similarly, 383 different fire perimeters have been considered in the analysis, corresponding to 5,343 burned grid cells.

3.2.2. Vegetation and soil factors

Data on vegetation and soil moisture can be estimated from climatic variables such as temperature, relative humidity and precipitation (Sharples & Matthews, 2011). Additionally, there is evidence from Holsinger et al. (2016) that vegetation moisture is easily approximated through satellite imagery, making use of the normalised difference vegetation index (NDVI). However, since these variables are intrinsically connected to short-term weather conditions, they lose their meaning if their values were to be averaged for a decade. Therefore, these two factors have not been be included in the analysis. Moreover, all other information on soil has also been left out because it was judged unobtainable.

Information on past fire activity is accessible through the national cartography of annual burned areas, which is made available in shapefile and is maintained by ICNF, as required by specific legislation. Despite it being considered of great importance to future wildfire occurrence, it is impossible to incorporate this variable into the proposed analysis because it is ever-changing and depends on the fire events.

Data on land cover or type of vegetation is available in the Portuguese land-use cartography (COS 2010) from Direção-Geral do Território (DGT). The fifth level of breakdown in the land use cartography provides a 118 category typology of forests and agro-forestry systems. However, as previously mentioned, it is well known from several studies that eucalyptus and pine tree forests are particularly fire prone (Rego & Silva, 2014), as are shrublands (Martínez-Fernández et al., 2013; Nunes et al., 2016; Salis et al., 2015). Thus, four variables have been created, indicating the percentage of the following types of vegetation on a given grid cell: (1) eucalyptus forests, (2) pine tree forests, (3) shrubland and (4) all other species.

The vegetation density features are also reflected on the fifth level of breakdown in COS 2010. The work developed by Dimitrakopoulos et al. (2011) separates dense from sparse vegetation on a given area by the threshold of 40% of vegetation cover. The different categories of the land use cartography, on the other hand, establish 50% as the splitting point for shrubland and sclerophyllous vegetation, and 30% for forests (as opposed to open forests). Therefore, the variable describing dense vegetation takes in consideration the forests with over 30% cover, as well as the dense shrubs and dense sclerophyllous vegetation areas.

What concerns the vegetation's flammable characteristics, these are specifically represented in the National Fuel Model (ICNF), which organises the different types of vegetation into fine fuel classes. This description of Portuguese forest spaces, comprising 18 representative fuel categories, has been compiled by Fernandes et al. (2009), and one of the outputs of this study links specific fire behaviour characteristics (namely speed of propagation and released energy) to these fuel models. As a result, a variable describing highly flammable fuel varieties has been created from the categories which denote a higher fuel load per hectare (7 out of 18 fuel varieties of forest, shrub or mixed vegetation).

3.2.3. Cimatic factors

Information for all identified climatic factors is made available on demand by Instituto Português do Mar e da Atmosfera (IPMA), for 20 weather stations in Portugal. This can be either monthly or seasonal data (for drought indices) or daily records (for all other variables). Nevertheless, these data are not free of charge. Additionally, some of these climate indicators are highly reliant on time windows, such as drought indices, or only make sense at short time intervals (days), which is the case of wind direction.

Daily land-based station data is available online, for the Portuguese territory, for some of the climate variables, curtesy of the North American governmental agency National Oceanic and Atmospheric Administration (NOAA). Daily data on mean temperature, mean dew point, mean windspeed and total precipitation, for 29 weather stations in Portugal and border areas in Spain, were collected, for the study period, for the purpose of this study.

At first, these raw data were subject to minor processing, namely unit conversion (°F to °C, knots to km/h and inches to mm). Secondly, and contrary to previous beliefs, since large wildfire occurrences were not limited to summer months, annual totals (for precipitation) and averages (for all other variables) were computed for the years that had at least 330 records. Given the high number of daily records, the fact that some series were not complete was considered of little relevance. Nevertheless, this analysis was conducted, and some stations were left out when computing precipitation-related variables (total annual precipitation and number of dry months).

Afterwards, these yearly values were averaged for each station, providing real information on average daily climate conditions during the decade. These annual weather variables were also used by Guo et al. (2016), Martínez-Fernández et al. (2013) and Nunes et al. (2013) in the identification of wildfire long-term occurrence factors and their spatial distribution.

To complete, these data points were interpolated for the entire study area using the geostatistical method inverse distance weighting (IDW). With IDW, values at unsampled locations are estimated as the weighted average of their neighbours, and the power function provides weighting proportional to the inverse of the distance between data point and predicted point (Lu & Wong, 2008). The calculations were conducted in ArcGIS with the help of the geostatistical analyst tool *IDW*, with default parameters, including an output raster cell of 1,5x1,5km.

Regarding the inclusion of a drought-related variable, this study takes inspiration from Ganteaume & Guerra (2018) and Ganteaume & Jappiot (2013), which make use of the Gaussen index for identifying the number of dry months. According to this methodology, a given month is considered dry if total rainfall (mm) is less than two times the average mean temperature (°C) (Bagnouls & Gaussen, 1957). This variable (number of dry months) was calculated from the processed daily climate data presented above and follows the same geostatistical interpolation procedure.

All things considered, data on wind direction, atmospheric pressure, evapotranspiration and sun radiation were either unobtainable or impossible to account for given the envisioned analysis.

3.2.4. Topographic factors

Topographic data in raster files is made available by ESRI for the Portuguese territory. The digital elevation models (DEM), as they are called, are available with a resolution of 30m, providing information on elevation. The DEM for the Portuguese continental territory was used, generalising its resolution to 450m as a way of simplifying data processing.

From this information and with the help of GIS technology, it is possible to compute other variables, such as slope gradient and orientation. ArcGIS was chosen for this task, making use of the spatial analyst tools *Slope* and *Aspect*. The first one identifies the steepness from each cell of a raster file and the output ranges from 0° to 90°. The second tool derives the aspect from the compass direction that the downhill slope faces for each location, returning a value between 0° and 360° (measured clockwise from north), or -1 if there is no slope. Both these techniques rely on a 3x3 moving window.

3.2.5. Human factors

There are three data sources used as the foundation for the majority of the identified human factors. These are official statistical records from Instituto Nacional de Estatística (INE), land use cartography from Direção-Geral do Território (COS 2010) and open source data from the Open Street Map platform (OSM).

Most statistical data on human activities such as agriculture and livestock are available from the agricultural census of 2009. Variables such as the percentage of farms with agricultural machines, the average used agricultural surface (SAU) per farm (ha), the farm density (no./km²) and the average number of livestock per farm, are available for one year for the smallest Portuguese administrative division (parish). The only exception is made to animal density, which is represented by the number of normal heads per SAU (no./ha) and is only available at municipal level (aggregation of parishes).

Additionally, the agricultural workforce is taken from the last population and housing census (INE, 2011) and represents the percentage of population employed in agricultural, animal production, fishing, forestry and hunting, corresponding to the entire sector A of CAE Rev. 3 (INE, 2007).

Moreover, COS 2010 is a rich source of evidence on human activities. First of all, it provides information on grazing, whether permanent or mixed with other land uses, according to the following breakdown levels: 2.3, 2.4.1.03 and 2.4.4.03. Furthermore, it allows for a separation between land-use types, such as farming and woodland areas. Agricultural regions are represented by the entire level 2. Forest spaces, with less human intervention, are represented by all categories in level 3, apart from entries 3.3.1 (beaches, dunes and sands) and 3.3.4.01 (non-forested burned areas).

It was impossible to find data from these or other sources in the subject of forest management, hunting and the size and density of agricultural plots. It is worth mentioning that the proportion of single agricultural holders over 65 years of age, available from the agricultural census (INE, 2009) for all parishes, is used as a proxy variable for the use of fire for agriculture and grazing activities, as it is mostly considered a traditional method (Fernandes et al., 2013).

Variables related to population, employment, educational level and housing also refer to census data and are available for all parishes for 2011. The unemployment rate is the sole variable related to employment. The percentage of secondary residence housing also denotes seasonal use or empty homes, while house ownership refers to the percentage of total housing owned by residents. The variables connected to education comprise the proportion of resident population with secondary education (high school diploma), the proportion of resident population with post-secondary education (university degree) and the illiteracy rate.

Population variables, taken from the census and considered in the context of this study, include the rate of population change since the 2001 census and the ageing index, which corresponds to the ratio between children (<14) and the elderly (>65). Nevertheless, for population, building and housing density, which are also taken from the census results, the values have been spatialized to the 1x1km INSPIRE grid cells, providing an accurate representation of population dynamics at that level.

Information on other human factors is regularly made available by INE for the municipalities, either with an infra-annual, yearly or biennial frequency. This is the case with tourism statistics (number of nights at hotel establishments per 100 inhabitants), the per capita purchasing power, the criminality rate and the potentiality index, which is related to female fertility and provides a measure of demographic potential. The annual records of these variables have been averaged for a representative value of the 2005-2015 decade.

Still, data on economic difficulties, the rural exodus and the ageing of rural population are unobtainable. The per capita purchasing power, the population change (2001-2011) and the age of farm holders can be considered as proxy variables.

An important group of human-related factors, mainly in the perspective of wildfire ignitions but concerning fire propagation in some cases, are proximity features. All considered proximity variables have been computed with the help of the ArcGIS analysis tool *Near*, as previously explained.

In this context, it is significant to take note of the role of the road, track, rail and power line networks. These data are available at the open source platform Open Street Map. Roads have been separated into two categories: primary roads, which comprise primary roads, trunks and motorways, and secondary roads, which also cover tertiary and unclassified roads. Tracks are not included into any of these two categories, representing mostly rural makeshift roads. The railway, on the other hand, indicates all rail types except for light rail and the subway.

The cartography of the National Network of Protected Areas (RNAP) is made available by ICNF in shapefile format, including sites of national and local significance such as national parks, natural parks, natural reserves, protected landscapes, natural monuments and privately owned protected areas. Even though there have been additions to RNAP during the study period and after that, these have taken place outside of the study area, meaning that the current geography is valid.

This entity also provides the cartography of the primary network of fuel management fire lines (RPFGC), for the year of 2014. This is considered a structural component of the landscape, aiming at protecting people, property and forests against wildfires, and has been created in the context of the Decree n. 124/2006, of June 28th (Decreto-Lei n.º 124/2006, de 28 de Junho in Diário da República n.º 123/2006, Série I-A). The value of this variable is specifically connected to fire spread.

Distance measures to land-use classes can be determined from COS 2010. Industrial areas, campsites and landfills correspond to the breakdown levels 1.2.1.01.1, 1.4.2.02.1 and 1.3.2.01.1 plus 1.3.2.02.1 respectively. In addition, recreational and touristic areas include all sublevels of category 1.4, apart from 1.4.1.02.1 (cemeteries) and 1.4.2.02.1 (campsites), and the proximity to urban areas and infrastructures encompasses all subclasses of levels 1.1 and 1.2.1, provided that the associated polygons have a minimum dimension of 1km^2.

The COS 2010 is also one of the sources of two other variables. In the case of changes in land use, a selection has been made of all forest areas in COS 2010 (all sublevels of category 3) and all agricultural areas in COS 1995 (all sublevels of category 2). The areas resulting from the intersection of these selections were retained as indication of a decrease in agricultural area and an increase in area covered by forest or semi-natural vegetation (Martínez-Fernández et al., 2013; Nunes et al., 2013; Rodrigues et al., 2014).

The other variable is an indication of whether the grid cell belongs or not to the Wildland-Urban Interface, approximated from land-use (COS 2010) and census data. The computation of this variable follows the guidelines presented in Stewart, Radeloff, Hammer and Hawbaker (2007) and Platt (2010), adapted to the Portuguese reality, which split the WUI into the intermix and interface areas. While the intermix WUI describes a wildland area of scattered structures, the interface WUI represents a borderline region between an urban community and wildland fuels. This is a short description of the steps followed on ArcGIS for the calculation of each type of WUI:

i) **Intermix:** (1) select grid cells with more than 6 buildings, from census data (Output 1); (2) select forest and natural areas (category 3 of COS 2010), except for subclasses 3.3.1 and 3.3.4.01 (Output 2); (3) intersect Output 1 with Output 2 and select areas with more than 500m² (Output 3); (4) select grid cells that intersect Output 3, with a search distance of -1 (Final Output 1).

ii) **Interface:** (1) select areas with more than 5km² from Output 2 (Output 4); (2) compute a 400m buffer from Output 4, as suggested by Calviño-Cancela et al. (2017) (Output 5); intersect Output 1 with Output 5 and select areas with more than 500m² (Output 6); select grid cells that intersect Output 6, with a search distance of -1 (Final Output 2).

The WUI corresponds to the union of Final Output 1 and Final Output 2 and makes up about half of the whole study area. It is interesting to observe that all WUI grid cells are linked to the intermix type, with roughly half of these belonging to the interface group as well.

To conclude, apart from what is presented on the wildfires database made available by ICNF, which enables the calculation of the time lag between alarm and first intervention and the duration of firefighting, there is no other information on firefighting efforts (namely, suppression resources). Nevertheless, the database provides too many null values, which renders unfeasible the addition of these variables to the analysis.

3.3. DATA ANALYSIS

Two different datasets, referring to fire ignition and propagation factors respectively, have been kept as the groundwork matter of all ensuing analyses. These datasets combine and synthesise the contributing factors described along the literature review, meeting the target set by the first research goal (**Research Goal #1**): “to recognise and to explain the main general factors driving wildfire ignition and spread, specifically in the Mediterranean context”.

The approach chosen for reaching the other goals proposed by this research work took evidence from a comprehensive review of methods for modelling wildfire danger, as previously described, as well as other appropriate methodologies employed in different scientific fields. It laid specific emphasis on a series of methodological considerations discussed earlier and related to the development of predictive models. The methods proposed as an attempt to reach research goals 2 to 4 are the following:

- **Research Goal #2** - Exploratory univariate data analysis;
- **Research Goal #3** - Multivariate data analysis: Cluster analysis;
- **Research Goal #4** - Multivariate data analysis: Probit and two-part regression models.

3.3.1. Exploratory univariate data analysis

All variables contemplated on the fire ignition and propagation datasets were subject to an exploratory analysis, in order to investigate the spatial distribution of the main drivers of fire occurrence and propagation. Particular emphasis was put in describing the spatial patterns of large wildfire incidence.

Histograms and box plots, along with basic descriptive statistics (range, average and standard deviation) were computed and presented in order to summarise the main characteristics of the identified driving factors. Maps were created on ArcMap 10.5.1 GIS software (ESRI) and histograms with box plots were plotted on Stata 14.0 statistical software (StataCorp), using the *histbox* command (Ender, 2002).

3.3.2. Cluster analysis

Cluster analysis is a widely used technique in multivariate data analysis, with its main purpose being the combination of observations into homogeneous groups regarding certain characteristics (Sharma, 1996). These groups should also be the most different among themselves with respect to the same characteristics.

Cluster analysis has been used by other authors to analyse the spatial patterns and behaviours of wildfires, in different contexts (see, e.g., Dimitrakopoulos et al., 2011; Parente, Pereira, & Tonini, 2016; Pereira et al., 2015; Seol et al., 2012). Additionally, it has also been used to group territorial units according to land cover dynamics and other sets of variables, in the same field of study (see, e.g., Ganteaume & Guerra, 2018; Oliveira et al., 2017). Therefore, this method has been deemed appropriate for identifying homogeneous regions within the study area with respect to the distribution of the main drivers of fire ignition and burned area extent.

Two different clustering exercises were conducted, to account for both fire ignition and propagation factors. The results were two different classifications of the grid cells composing the study area, in relation to two specific sets of elements contributing to wildfire occurrence.

First of all, given the fairly high amount of retained variables (58), Pearson's correlation coefficient was calculated as a measure of the degree of association (or closeness) between all pairs of continuous variables (*X,Y*). All correlation coefficients of $\rho_{X,Y} > 0.5$ or $\rho_{X,Y} < -0.5$ were considered high, meaning either *x* or *y* could be discarded because part of the information provided by one of these variables is present on the other. The chosen threshold was considered adequate in this situation because of (1) the great number of observations ($N = 21.750$) and (2) the fact that in cluster analysis lower correlation coefficients can be perceived as relevant (Asuero, Sayago, & González, 2006).

Data reduction was completed with a measure of variable worth in predicting the phenomenon under study (target variable), which had a determinant role in the choice of the variables to remove from the analysis in case of high correlation coefficients. This ordering of variables by their apparent value was achieved by calculating the Gini split worth statistic, which makes use of a reduction of the Gini index. This method corresponds to building a decision tree of depth one (SAS Institute Inc., 2009), which can be described as a Boolean function of one feature.

The variables retained as a result of the combination of these procedures were standardised with z-scores, to enable value comparison, and the transformed data were used in the cluster analysis. Two geographical standardised variables denoting the cell's x and y coordinates were added to the analysis, as it provided better results.

According to Sharma (1996), hierarchical clustering methods rely on a similarity matrix, which represents the distances between observations and clusters, building clusters at each step from *n* to 1. However, these methods have been considered unsuitable in the presence of a great number of observations, due to the chain effect present in these methods, which leaves us with non-hierarchical clustering approaches, which divide the data into *k* groups, depending on the location of *k* initialisation seeds and a process of reassigning observations to the closest cluster in each iteration. These methods require *k* to be defined at the beginning of the analysis, and have proven to be sensitive to the location of the initialisation seeds. Hence, both approaches, hierarchical and non-hierarchical, have its own advantages and disadvantages.

Given the restrictions of both approaches, the same author proposes that hierarchical and non-hierarchical methods be used as complementary techniques, in order to boost clustering performance. The following methodology has been adopted in the context of this research:

1. Performing a **hierarchical clustering analysis**.
   a. Five different methods were employed and the one performing the best was selected. These approaches differ in regard to the preferred methodology for computing distances between clusters, i.e. **Centroid** relies on the clusters' centroids, **Single** calculates the distance to the nearest neighbour, **Complete** the distance to the farthest neighbour and **Average** computes the average distance. **Ward's method** works differently, creating clusters which maximize within-clusters homogeneity.
   b. The chosen measure of similarity was the Euclidean squared distance, which is the default parameter for the different methods used within SAS PROC CLUSTER.
2. Defining the **number of ideal clusters to retain**, from the R^2 plot and the dendrogram.
   a. The R^2 denotes the ratio of the between-clusters sum of squares to the sum of the between-clusters sum of squares and within-clusters sum of squares[1], evaluating the extent to which clusters are different among themselves. The R^2 values were analysed through a scree plot, with the choice of *k* falling on the number of clusters immediately before a considerable drop in R^2 (explained information). The rationale behind this criterion is to prevent that two very different clusters are joined, which would lead to a meaningful drop in the R^2.
   b. The dendrogram is a tree diagram depicting the clustering process, which presents an immediate visualisation of imbalances in cluster dimension. The vertical lines, depending upon the orientation of the dendrogram, are proportional to the distance between the clusters that are joined. Hence, looking at it, is possible to choose the number of clusters prior to the point where very far apart clusters are joined, in a similar criterion to that of the R^2.
3. Conducting a **non-hierarchical clustering analysis**, following the K-means method and placing the initial seeds on the hierarchical clusters' centroids. This way, the non-hierarchical algorithm will work as an optimization technique to the hierarchical methods. In other words,

[1] $R^2 = \frac{SS_b}{SS_t}$ where $SS_t = SS_b + SS_w$

the limitation of the randomness in the initial placement of the seeds is surpassed. If there is any change in the clusters' composition it will only be those that improve the initial result.

a. The **K-means method** follows the principles of the non-hierarchical clustering methods explained above, reassigning observations based on the calculation of their distance to clusters' centroids (Tan, Steinbach, Karpatne, & Kumar, 2018).

The entire cluster analysis has been conducted with the help of SAS Enterprise Guide 7.1 statistical software (SAS Institute Inc.). The main commands used were PROC CLUSTER, PROC TREE, PROC MEANS and PROC FASTCLUS with the seed= option.

3.3.3. Probit and two-part regression models

The short review of the wildfire danger assessment methods carried out previously was evident at stressing the suitability of generalised regression techniques, such as logistic regression, in this field of study. These regression methods have been chosen for exploring the contribution of the main factors of large wildfire occurrence and spread, as well as for recognising the variability of the factors importance among regions. Two different methodologies were employed in this analysis, depending on the target variable: large fire ignition and propagation.

The entire analysis, including model development and the assessment of results, was conducted with the help of Stata 14.0 statistical software (StataCorp).

3.3.3.1. Ignition

According to Wooldridge (2012), binary response models, as they are known, are sophisticated regression methods that restrict fitted probabilities to the values between 0 and 1, such that:

$$P(y = 1|X) = G(\beta_0 + X\beta)$$

where *G* is a function presenting values between 0 and 1 and $X\beta = \beta_1 x_1 + \cdots + \beta_j x_j$.

The probit model is a specific case of binary response models where *G* is the standard normal cumulative distribution function (cdf). One of the ways it differs from logistic regression, another binary response modelling technique, is due to the model's errors, which follow the standard normal distribution and not the standard logistic distribution. The probit model is usually favoured for this reason, as well as for the fact that their results are interpreted as probabilities instead of odds. In both cases, however, errors are assumed to be independent of *X*.

Given the fact that the probit model is primarily derived from a latent variable model, the magnitudes of the model's estimates ($\hat{\beta}_j$) do not represent the effect of x_j on $P(y = 1|X)$. This fact is made more difficult because *G* does not follow a linear distribution.

For a continuous variable x_j, the calculation of its partial effect on the response probability $p(X) = P(y = 1|X)$ lies on the partial derivative:

$$\frac{\partial p(X)}{\partial x_j} = g(\beta_0 + X\beta)\beta_j$$

In the case of a binary explanatory variable x_k, the partial effect of changing x_k from 0 to 1 with the other variables remaining constant is given by:

$$G(\beta_0 + \beta_k + \beta_j x_j) - G(\beta_0 + \beta_j x_j)$$

These results are usually interpreted as average partial effects, meaning the average of the partial effect across the population, or simply as partial effects at the average, i.e. the average value of the explanatory variable (Wooldridge, 2012). As it is explained by this author, these are two distinct measures: the first corresponds to using the average of the nonlinear function, the second refers to the calculation of the nonlinear function of the average.

For the purpose of this study, partial effects and associated probabilities were estimated and plotted for the entire range of population values. These were the results interpreted as the magnitude of the factors contribution to large fire ignition events in the study area during this specific period. One model was fitted to the entire study area, which was mostly regarded as a measure of comparison. Because the main focus of this work is identifying underlying regional patterns, this process was then repeated for each of the previously defined clusters in order to explore the variability of the factors importance throughout the region.

All models were fitted through the forward-stepwise selection procedure. It works by beginning model specification with no covariates, which are added to the model step by step, based on the coefficients statistical significance. Variables can be included and excluded from the model at each step, as coefficients are revaluated along the process. Significance thresholds are defined beforehand both for adding and removing variables.

As mentioned by Hosmer and Lemeshow (2000), this methodology is very useful for an effective screening of a large number of variables. The chosen thresholds for adding and excluding variables were set at $p < 0.1$ and $p \geq 0.2$ respectively.

For model assessment, the employed goodness-of-fit statistic was the Pearson χ^2, together with an analysis of classification tables, particularly sensitivity values and the percentage of correctly classified observations, and the area under the receiver-operator (ROC) curve. This methodology follows the advice of some authors who believe that model performance evaluation must consider both specification and discrimination power (Hosmer & Lemeshow, 2000). This double assessment was particularly meaningful in the context of ignition models as the number of target observations is particularly low (roughly 13% of all cells).

The Pearson χ^2 goodness-of-fit statistic measures the difference between observed and fitted values. It is a summary statistic based on the Pearson residuals and given by:

$$\chi^2 = \sum_{j=1}^{J} r(y_j, \hat{y}_j)^2$$

In theory, the lower the Pearson χ^2 the better the model. P-values are calculated using the χ^2 distribution with degrees of freedom $J - (p + 1)$. We reject the null hypothesis that the model is poorly specified if $P(\chi^2) \leq 0.1$.

Classification tables result from model estimations, with observations being divided into four groups representing correct classification on one hand and target variable behaviour (0,1) on the other. Sensitivity values refer to correctly classified positive target outcomes and therefore presents a good measure of the model's power. The percentage of accurate classifications is also a valuable assessment measure, although a great unbalance in target variable distribution may increase the number of correct observations with no connection to model performance, as this method favours classification into the larger group (Hosmer & Lemeshow, 2000).

The ROC curve is a similar method, plotting the relationship between sensitivity and the false positive rate, i.e. 1-specificity (Phillips et al., 2006). The area under the curve is calculated as a measure of accuracy, with 50% meaning no accuracy and 100% meaning maximum accuracy.

3.3.3.2. Propagation

Large wildfire propagation can be derived from the burned area extent of large fire occurrences. The percentage of burned area in each cell follows a highly skewed distribution, with only 24.8% of the study area displaying any sign of burning, which prevents the use of the classic multivariate Ordinary Least Squares (OLS) linear regression model. We can assume that $P(y=0)>0$, which means that using the OLS estimator would bias our estimation.

The applicability of non-linear two-stage estimation procedures, such as two-part models (2PM), has been successfully demonstrated in a variety of fields where observed data (either count or continuous data) are characterised by a heavy presence of 0 in the response variable (Farewell et al., 2017). Two-part models are composed of two distinct stages, the first predicting the probability of occurrence (0,1), through a binary response model, and the second predicting the target variable conditional on nonzero outcomes, using the linear regression model (Buntin & Zaslavsky, 2004).

The unconditional expectation on $E(y)$ is therefore given by the multiplication of the probabilities of occurrence from the first part by the expected levels from the second part:

$$E(y|X) = P(y > 0|X)E(y|X, y > 0)$$

For the purpose of this study, burned area percentage was first generalised to a binary target variable and the methodology defined for modelling large wildfire ignition occurrences (probit model) was replicated, including coefficient interpretation and model assessment. Subsequently, an OLS regression was fitted to the subpopulation of cells displaying burning activity during the decade, also with the forward-stepwise variable selection procedure.

Different phenomena are being estimated on each of the two stages. In the first part it is the probability of a cell burning being modelled, whereas in the second part the extent of burned area is the one being predicted, restricted to the occurrence of burning activity.

Model assessment was determined with the help of the F statistic, the adjusted R^2 and by plotting the residuals against the fitted values. The models were also checked for normality of the residuals.

Both the F statistic and the adjusted R^2 rely on the residual and total sum of squares. The F statistic is useful for model assessment when the overall goal of the analysis is explanation (Sweet & Martin, 2012). For the F-test, we are able to reject the null hypothesis that the predictors' coefficients are statistically equal to 0 if $P(F) \leq 0.1$, meaning there is at least one coefficient different from 0.

Adjusted R^2 indicates the model's goodness-of-fit for providing a value for the amount of variance explained by the model, with the benefit of only increasing with the addition of explanatory variables if the increase in model fit pays off the loss of degrees of freedom. A high adjusted R^2, although desirable, is not essential when the final objective is assessing the relationship between the different factors and not prediction (Sweet & Martin, 2012).

The plot of the model's residuals against the fitted values is particularly useful to identify violations of the OLS assumptions. Points should be randomly scattered around $y = 0$ and no pattern should be visible from the data plot, particularly linear relations and increased or decreased variability.

4. RESULTS AND DISCUSSION

4.1. LARGE WILDFIRE INCIDENCE

Large wildfire incidence in central Portugal can be characterised by the amount of large wildfire ignitions, the burned area extent and their locations. Summary tables for the four target variables of this research are displayed in the Annex (Annex B: Tables B1 to B4), describing the occurrence of large wildfires in the study area.

The spatial distribution of these variables shows that, even though large wildfires occur in all locations throughout the study area, both ignitions and burned area are concentrated in northern and southern central regions. Large ignitions appear to occur far from the Spanish border, while larger burned areas are very rare near the coastline. The overwhelming majority of non-ignition and non-burned cells accounts for the unbalanced statistical distributions.

4.2. MAIN DRIVING FACTORS OF LARGE WILDFIRE IGNITION AND SPREAD

The main driving factors of large wildfire ignition and spread have been identified according to the literature review and the elements considered in this research have been specified along with the details of the data collection methodology. The summary tables included in the Annex (Annex B: Tables B5 to B58) provide a description of the selected driving factors, listing their main features, displaying their spatial arrangements and characterising the statistical distributions.

There is a strong pattern common to the majority of the human factors driving large wildfire ignition and spread: the opposition dynamic urban-coastline and depopulated rural-inland. This trend is particularly noticeable in the case of agricultural and livestock related variables, variables connected to the proximity to infrastructures and urban areas, population and housing density measures, the demographic variables POPCHANG_RT and AGE_INDEX, as well as education factors.

Many socioeconomic factors influencing large fire events display extremely asymmetric distributions, resulting from a preponderance of high or low values. Noteworthy exceptions are agricultural connected variables AGRMAQ_PERC and AGR65_PERC, housing related variables SSEHOUS_PERC and HOBR_PERC, the potentiality index (POTENT_INDEX), the unemployment and crime rates (UNEMP_PERC and CRIME_RT) and the rate of population change from 2001 to 2011 (POPCHANG_RT).

It is possible to observe different vegetation cover patterns across the study area. Eucalyptus forests are scattered everywhere except in the highland regions of the northeast. In these regions shrublands predominate. Pine cover is very common across the study area, except in the Castelo Branco region, and concentrates along the coastline and in the heart of Portugal. Other forest species are mostly restricted to the areas along the border with Spain. Additionally, a striking match is visible between highly flammable and dense vegetation covers, as expected.

Apart from annual precipitation values (PRECTOT), climate variables show a small range of possible values and are bound by north-south dynamics and elevation patterns. Highest peaks and most rugged landscapes are located inland, with low altitudes spreading from the coastline to central areas.

After the previously described process of data reduction, the following variables were kept for the statistical analyses:

Table 2 – List of variables kept for the statistical analyses

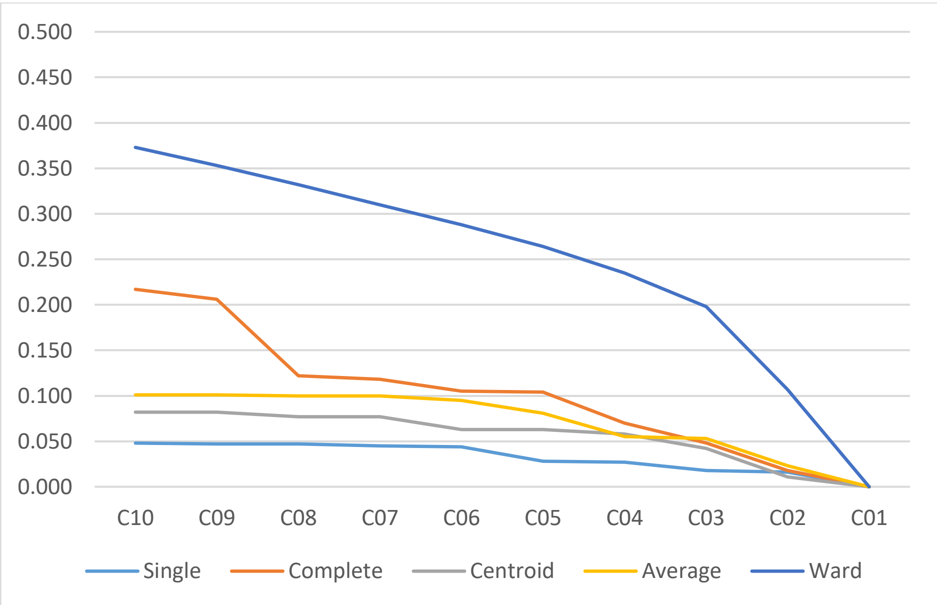
EFFECT	TARGET	VARIABLES		
Ignition	IGN_PLUS	PRECTOT CRIME_RT NHEST_PERC HEADS_NSAU* PCPP POTENT_INDEX* SLOPE* URB_DIST AGE_INDEX*	PRIM_PERC* UNEMP_PERC RPSUP_PERC AGRMAQ_PERC LVSTK_NFARM* POWER_DIST GRZ_COS* POP_GRID*	SHRUB_COS* AGR_COS* SROAD_DIST WUI PIN_COS ASPECT* AGR65_PERC PROAD_DIST*
Spread	BA_PERC	IGN_DIST SLOPE* SHRUB_COS* HEADS_NSAU* ELEVATION AGR_COS* POTENT_INDEX*	LVSTK_NFARM* AP2015_DIST SSEHOUS_PERC AGE_INDEX* SAUFARM_HA EUC_COS PRIM_PERC*	POP_GRID* GRZ_COS* FARMDEN_KM DRYMONTH PROAD_DIST* ASPECT* OUTR_COS

* Variables used on both effects.

4.3. Large wildfire ignition

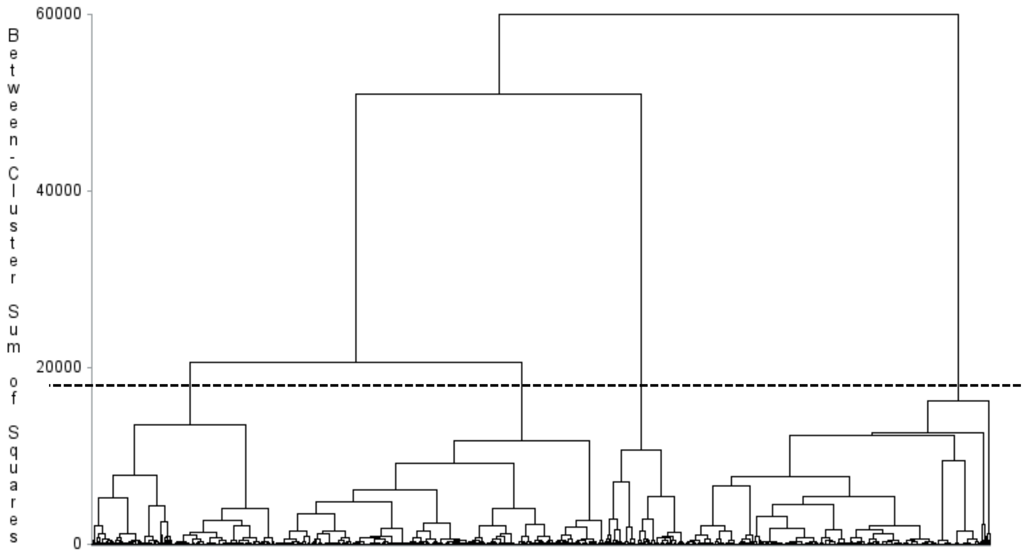
The best performing hierarchical clustering method was Ward's method. This evaluation is clearly presented on the graph below (Figure 8), which shows the plot of the R^2 values associated to 10 different clustering partitions (from 10 groups to one) for each of the previously described hierarchical methods:

Figure 8 – R^2 values associated to different clustering partitions and hierarchical methods



The Ward's method curve also suggests that three clusters might be a desirable result, since there is an abrupt decrease in explained variance immediately thereafter. However, when interpreting the dendrogram (Figure 9) it is evident that the four clusters solution is the most balanced partition. It corresponds to an R^2 of 0.235 which is not a high value.

Figure 9 – Ward's method dendrogram with the cut-off line at four clusters

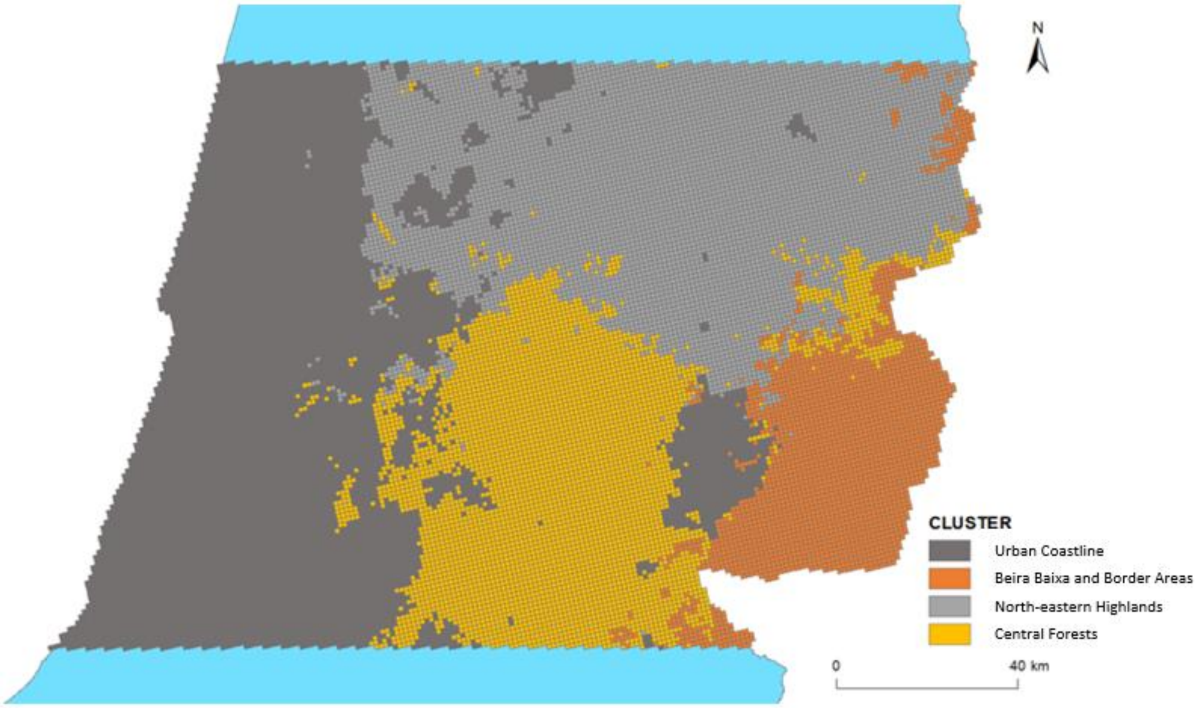


Different sets of variables were used in exploratory exercises, including an experience containing only socioeconomic factors. Nevertheless, none of these attempts displayed an increase in explained variance that could make up for the apparently random exclusion of variables from the study. The Wildland-Urban Interface was not used, as it is a binary variable. Lastly, two geographical variables representing the longitude and latitude (x and y) values of grid cells were added to the analysis, which improved the boundaries of the different clusters. This explicit geospatial procedure has been used extensively for clustering purposes in different fields, such as epidemiology (Li, 2018), and accounts for the data's spatial autocorrelation (Ruß & Kruse, 2011).

K-means non-hierarchical clustering was then performed for four clusters. This process needed a total of seven iterations for completion and improved the value of the R^2 to 0.258. This small increase in explained variance is considered enough given the main objective of this research, the overall size of the dataset and the interpretability of the clustering solution.

The map presenting the clustering solution can be found below (Map 2). It is important to understand that not all variables displayed the same proportion of explained variance (R^2). The longitude (X_COORD), the total annual precipitation (PRECTOT), the distance to power lines (POWER_DIST) and the distance to urban centres and infrastructures (URB_DIST) exhibited the highest values (above 45%). On the other hand, the average animal headcount in each farm per parish (LVSTK_NFARM), the population density (POP_GRID) and the orientation (ASPECT) showed the poorest performance among all variables, standing well below 10%.

Map 2 – Clustering solution (4 clusters) based on large wildfire ignition driving factors



The following table (Table 3) summarises the clustering solution in regard to group dimension, cluster location and specific ignition-related characteristics:

Table 3 – Description of the ignition clustering solution based on the selected variables

FEATURES	URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
Dimension	7,933	2,358	6,611	4,668
Population Density	High	Low	Medium-Low	Low
Distance to Infrastructures	High density of urban centres and roads	Very low density of urban centres and roads	Average density of urban centres and medium to high road density	Low density of urban centres and medium to low road density
Agriculture and Livestock	Few people employed in primary sector activities although agricultural activity is still present, younger agricultural population, medium to high livestock activity and high density of animal headcount per utilised agricultural surface, high mechanisation of agriculture	Comparatively high number of people employed in primary sector activities, large agriculture and grazing areas, old rural population, average to low livestock activity and low density of animal headcount, very low mechanisation of agriculture	Average number of people employed in primary sector activities and medium to low extents of agriculture and grazing areas, reasonably old rural population, medium livestock activity and low density of animal headcount, low mechanisation of agriculture	Average number of people employed in primary sector activities and very few agriculture and grazing areas, very old rural population, average to low livestock activity and average to low density of animal headcount, low mechanisation of agriculture

FEATURES	URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
Demographic and Socioeconomic Aspects	Young population and medium-high potentiality index, high purchasing power, medium to low unemployment rate, medium to high criminality rate, large number of residents with higher education, intense touristic activity	Very old population although an average to high potentiality index, medium to low purchasing power, reasonably high unemployment rate, high criminality rate, small number of residents with a higher education degree, relatively intense touristic activity	Moderately old population with a very low potentiality index, average to low purchasing power, medium to high unemployment rate, low criminality rate, comparatively low proportion of residents with higher education, moderate to little touristic activity	Old population with a medium-high potentiality index, low purchasing power, average to low unemployment rate, low criminality rate, very low proportion of residents with a higher education degree, little touristic activity
Biophysical Aspects	Flat landscape, medium incidence of pine trees and few shrub, low precipitation levels	Flat landscape, very few pine trees and minor shrub areas, low precipitation levels	Hilly terrain, little presence of pine forests and large shrubland extensions, high precipitation levels	Very rugged landscape, large pine tree forests and medium to high shrub incidence, low to medium precipitation levels

There is an evident contrast between groups, with “Urban Coastline” showcasing a clear urban profile. “Beira Baixa and Border Areas”, on the other hand, is characterised by a predominance of traditional extensive agricultural activities, mostly connected to grazing, olive groves and cork oak forests. “North-eastern Highlands” is mainly defined in respect to biophysical aspects, with a large extension of shrub vegetation and high precipitation values. Lastly, “Central Forests” refers particularly to a rugged landscape dominated primarily by pine tree forests.

This cluster partition is consistent with administrative limits and with widespread knowledge on the characteristics of these regions. In fact, the opposition between rural abandonment in inland areas and increasing demographic pressure along the coast is a notorious trend in Portugal, well-noted in the literature for its connection to wildfire events (Almeida et al., 2013; Moreira et al., 2010). Land use trends follow this pattern, with most agricultural areas being located along the central coastal plain (“Urban Coastline”) and shrublands overtaking considerable land stretches in eastern Portugal, in mountainous and sparsely populated regions (“North-eastern Highlands”) (Oliveira et al., 2012). Forest monocultures are also common in central Portugal (“Central Forests”) (Almeida et al., 2013).

These results can also be compared with findings from the reference literature, connected to the difference in large wildfire ignition patterns among regions, with interesting conclusions. One cannot dissociate large ignition events from burned area, as they are intrinsically connected. Oliveira et al. (2012) discuss how forest-dominated regions are characterised by not many ignitions but where wildfire events originate large burned areas (“Central Forests”).

Inland areas of central Portugal fall into this category, displaying large burned areas (Oliveira, Zêzere, Queirós, & Pereira, 2017). The area corresponding approximately to “Central Forests” (Pinhal Interior) is historically connected to the occurrence of large wildfires, it is characterised by an ageing population, by great extents of uninterrupted forest and shows a strong burning probability dependence on fuel age (Oliveira et al., 2012; Pereira, Carreiras, Silva, & Vasconcelos, 2006). Overall, the coastline displays the opposite features.

Moreira et al. (2010), on the other hand, relate ignitions producing large burned areas to low population density and place them preferably along the border with Spain ("North-eastern Highlands" and "Beira Baixa and Border Areas"), which, in the case of "Beira Baixa and Border Areas", is not entirely consistent with the data. For "North-eastern Highlands", however, this is an accurate assessment and is linked in some measure to the increase in fuel availability.

With these regions in mind, probit models were developed for the entire study area and for each cluster according to the methodology presented above. The results of this analysis are presented below, including the estimated models' quality assessment measures (Table 4), as well as the average partial effects (APE) and statistical significance of the estimates (Table 5).

Table 4 – Estimated models quality assessment measures (ignition)

MODEL ASSESSMENT	GLOBAL	URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
Pearson χ^2	22027.86 **	7787.44	1773.06	6481.36	4530.27
Sensitivity	0.07%	0%	0%	1.42%	2.2%
Correct classifications	86.96%	88.26%	97.33%	80.77%	88.41%
Area under the ROC curve	69.12%	67.23%	80.42%	66.76%	72.35%

Significance of the Pearson χ^2: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$)

The ROC curve plots can be found in Annex C, along with the classification tables. The goodness-of-fit measures do not look too cheerful for any of the regional models, essentially due to the disparity between ignition and non-ignition cells (13% ignition cells overall). Nevertheless, of all methodologies, the forward stepwise selection criterion, with the chosen thresholds for adding and excluding variables set at $p < 0.1$ and $p \geq 0.2$ respectively, was the one yielding the best results sensitivity-wise.

Table 5 – Average partial effects (APE) and estimates significance (ignition)

VARIABLES	GLOBAL	URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
PRIM_PERC		0.0036834 ***	0.0020299 ***	-0.0020661 **	
AGRMAQ_PERC	0.0004325 ***		0.0012495 ***	0.0011656 ***	0.0008776 **
LVSTK_NFARM	0.00001 ***	0.00000498 **	-0.0003175 ***	0.0000415 ***	-0.0002832 **
HEADS_NSAU		-0.0026729 **		-0.013785 ***	-0.0230373 ***
AGR65_PERC	0.0009325 ***	0.0014434 ***		0.0015176 ***	
NHEST_PERC	-0.000047 ***			-0.0002209 ***	-0.0003535 ***
POTENT_INDEX	-0.0010581 **			0.0047426 ***	
AGE_INDEX	-0.0000113 **			0.000057 ***	-0.0000297 ***
PCPP					0.0020026 ***

VARIABLES	GLOBAL	URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
UNEMP_PERC	0.0025004 ***	0.0025886 *	0.0013594 *		0.0051223 ***
CRIME_RT	-0.0009958 **	-0.0032611 ***			
RPSUP_PERC	-0.0065255 ***	-0.00625 ***	-0.0038194 **	-0.0056018 ***	-0.004297 *
POP_GRID		0.0000271 ***			0.0002133 *
AGR_COS					
GRZ_COS	-0.0021489 ***	-0.0013045 *	-0.0003569 ***	-0.0074459 ***	
PIN_COS	-0.0004579 ***		-0.0013038	-0.0007647 ***	-0.0008468 ***
SHRUB_COS	0.0006049 ***			0.0012196 ***	-0.0011757 ***
WUI	0.0387819 ***	0.0403855 ***		0.0377415 ***	0.0425118 ***
PROAD_DIST			0.0032228 ***		-0.005423 ***
SROAD_DIST	-0.0510909 ***		-0.0102155 ***	-0.0882342 ***	-0.004297 ***
URB_DIST	-0.0041204 ***	0.0056221 ***	-0.002652 ***	-0.0051581 ***	-0.0045913 ***
POWER_DIST					
PRECTOT	0.0001594 ***		-0.000148 *	0.0003796 ***	0.0005267 ***
SLOPE	0.0043096 ***	0.0096275 ***		0.0033764 ***	0.0068886 ***
ASPECT	0.0001844 ***	0.0002757 ***		0.000322 ***	

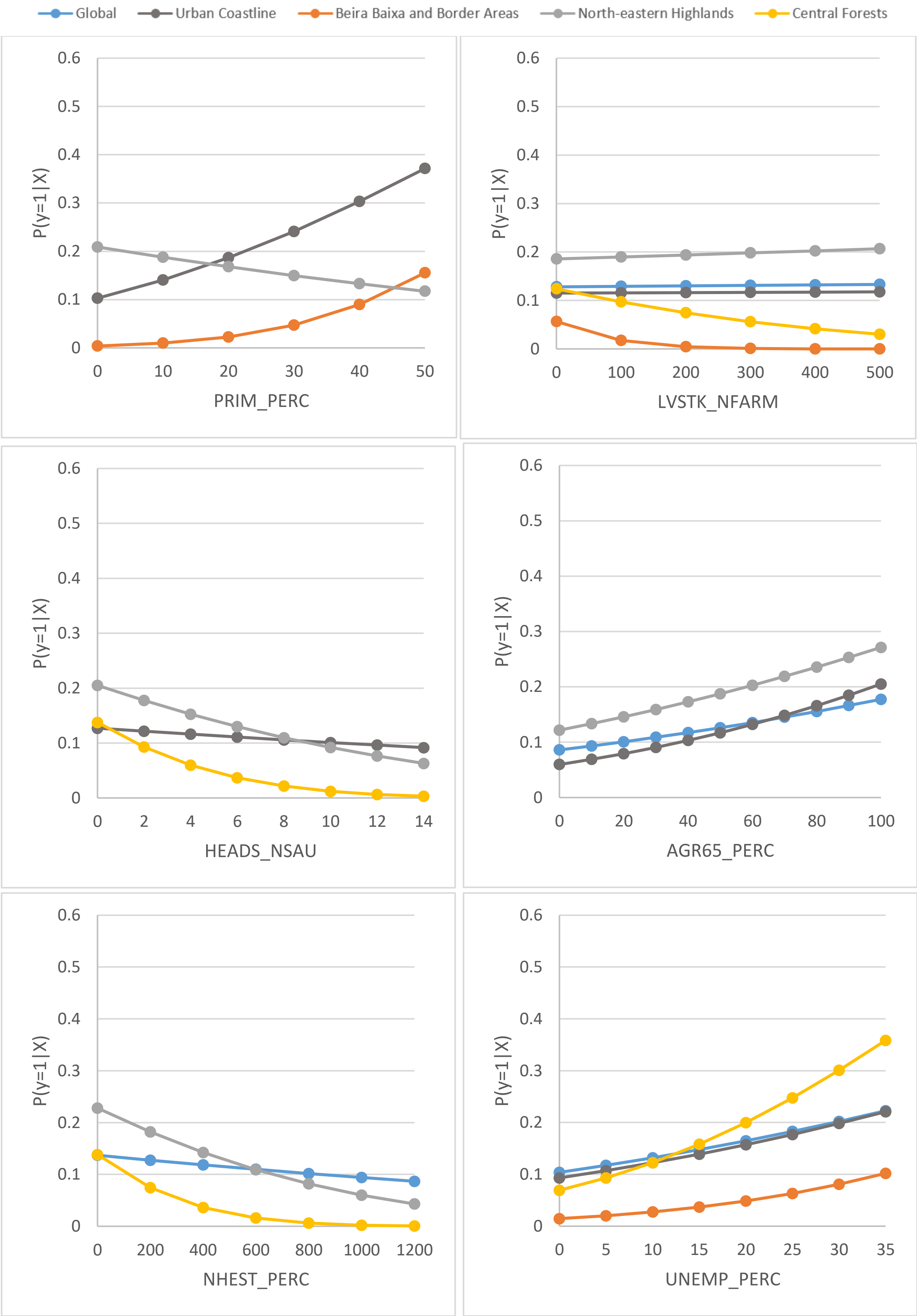
Significance of average partial effects (APE): *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$)

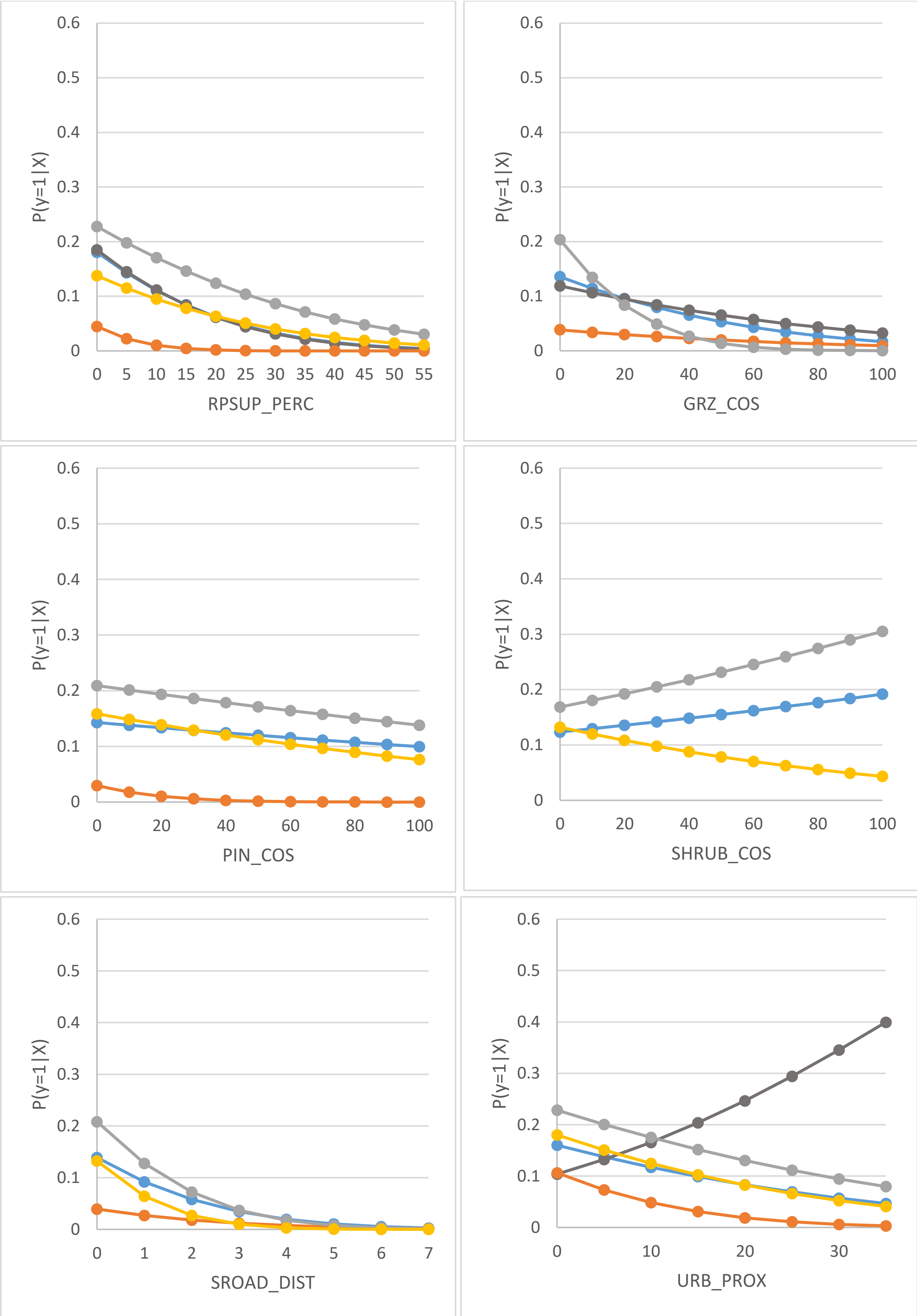
The table above (Table 5) presents the average partial effects (APE), following the methodology illustrated before, with emphasis to the estimates significance. All models have a significant (***) negative constant ($\hat{\beta}_0$), except for “Beira Baixa and Border Areas” whose intercept is not statistically significant.

It is interesting to observe that the only variables left out of the models were the percentage of agricultural areas (AGR_COS) and the distance to power lines (POWER_DIST), even though these same factors were moderate to well represented in the cluster analysis. On the other hand, three variables were included into all five model specifications, these being the average number of livestock per farm (LVSTK_NFARM), the percentage of resident population with a university degree (RPSUP_PERC) and the proximity to urban areas and infrastructures (URB_DIST). However, for LVSTK_NFARM and URB_DIST the signs shifted among clusters, which suggests great spatial variability.

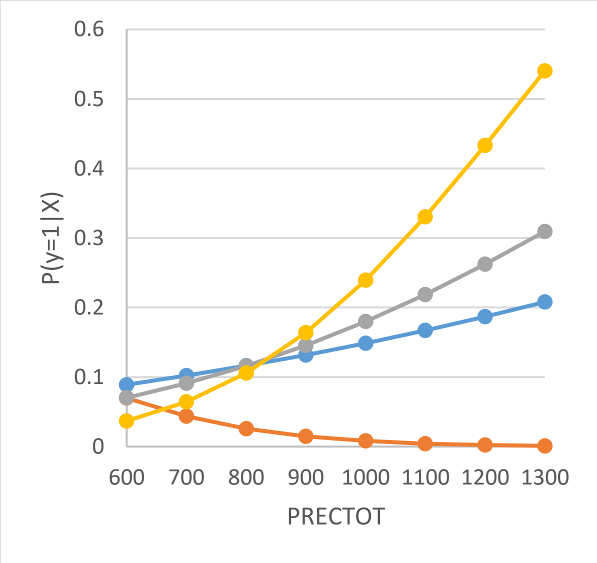
Specific driving factors have been highlighted, given the perceived higher importance or regional variability of their estimated partial effects and associated probabilities. The following figure (Figure 10) summarises the main results of the models by calculating the ignition probabilities associated to each real value of the variables included in the modelling exercise. For a correct interpretation of these results, it is also important to consider the individual variable plots per cluster, which present the estimates’ corresponding 95% confidence intervals, in Annex D.

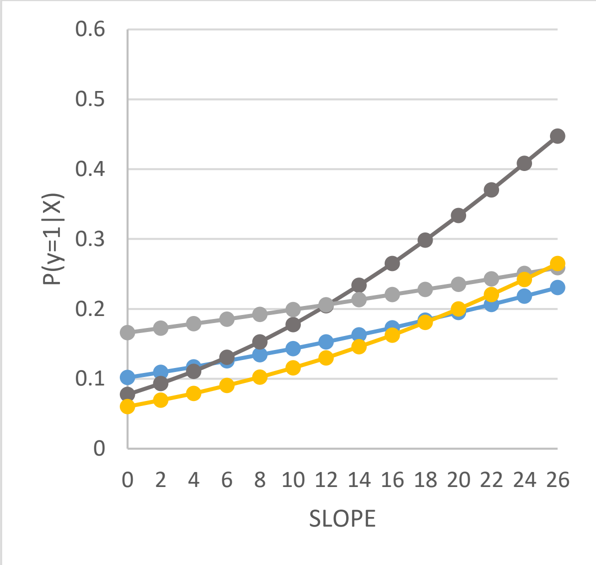
Figure 10 – Main results of the ignition probit model





P(y=1|X)
RPSUP_PERC
GRZ_COS
PIN_COS
SHRUB_COS
SROAD_DIST
URB_PROX





The proportion of residents employed in primary sector jobs (PRIM_PERC) seems to display a positive relationship with the occurrence of ignitions developing into large burned areas, in the case of Clusters “Urban Coastline” and “Beira Baixa and Border Areas”, and a negative relationship, although not significant (see Annex D), in the case of “North-eastern Highlands”. For “Urban Coastline”, 50% of primary sector employment corresponded to a probability of almost 40% for a large ignition, while for “Beira Baixa and Border Areas” the associated probability was roughly 16%. This might suggest that in locations with a stronger incidence of land use activities, such as agriculture, this factor plays a determinant role in boosting the number of large wildfire ignitions.

Oliveira et al. (2014) have shown a positive association between agricultural activities and ignition locations in Portugal, while Rodrigues et al. (2016) have proven this link to be non-stationary in Spain, with several regions displaying no relationship between both phenomena. Both accounts support the findings of this analysis.

In turn, an increase in the age of farm holders (AGR65_PERC) provides a rise in ignition probability for all contemplated models (Global, “Urban Coastline” and “North-eastern Highlands”). This trend is particularly valid for “North-eastern Highlands”, where 100% of farm holders over 65 corresponds to an ignition probability of approximately 30%. It is important to recall that this variable was used as a proxy for the use of traditional methods in agriculture, such as burning pastures.

These results reinforce the positive association between agricultural activities and large fire ignitions. As mentioned by several authors, the use of fire for agricultural purposes, which is common practice in many traditional methods, is a well-known contributing factor of fire ignitions (Álvarez-Díaz et al., 2015; Rodrigues et al., 2014). The modelling outcomes suggest that the same conditions are valid for large fires in central Portugal.

The extent of grazing areas (GRZ_COS) displays a weak negative relationship with the probability of large ignition events in the case of all clusters. For “Urban Coastline” this relationship does not seem sufficiently strong for interpretation (see Annex D). For “North-eastern Highlands”, however, the decrease in ignition probability with an increase in grazing area is very pronounced, starting at 20% for no grazing spaces to almost 0% likelihood for 40% pasture land.

What concerns animal creation, the results are mixed. Animal density (HEADS_NSAU) displays a negative relationship with large wildfire ignition probability for “Urban Coastline”, “North-eastern Highlands” and “Central Forests”, not expressive in the case of “Urban Coastline” (see Annex D). For “North-eastern Highlands”, an absence of animals per agricultural area utilised for farming represents an ignition probability of 20%, while the maximum animal density provides a probability of roughly 6%.

For “Beira Baixa and Border Areas” and “Central Forests”, the increase in the average number of animals per farm (LVSTK_NFARM) also decreases the chances of fire ignition, although these trends are somewhat weak. The global model and the other clusters, however, seem to display the opposing relationship, with “North-eastern Highlands” showing a very pronounced increasing trend. As seen in Annex D, the corresponding probability of no animal activity was less than 20%, whereas a marked presence (5 thousand animals per farm on average) resulted in a probability of approximately 45%.

Livestock activities and animal density are believed to affect ignition patterns in different ways. The use of fire is assumed to be widespread, in this context, for gaining or preserving cattle grazing, specifically in less accessible areas (Oliveira et al., 2014; Vilar et al., 2016). Apart from “North-eastern Highlands”, where a marked livestock activity resulted in a higher large fire ignition probability, the results of this modelling exercise seem to go against previous knowledge. It might be the case that, in the study area, increased animal creation and larger grazing grounds are associated to a higher density of human activities and better territorial management practices, which prevent the development of fire ignitions into large burned areas.

Some socioeconomic aspects exhibit a strong positive relationship with ignition probability. This is the case of the unemployment rate (UNEMP_PERC) for some regions. Among all models, “Central Forests” shows the most distinct link, with the maximum unemployment rate (35%) corresponding to almost 40% of large fire ignition probability.

Previous studies have found a positive connection between unemployment and wildfire ignitions. Dondo Bühler et al. (2013) uncovered a link between arson and high unemployment, in Argentina, as a result of social unrest. This explanation is not deemed sufficiently relevant in the case of central Portugal, particularly “Central Forests”, and further studies focused solely on intentional wildfire events are needed to better interpret these results.

Touristic activity, measured by the number of nights at hotel establishments per 100 inhabitants (NHEST_PERC), shows a negative association with ignition likelihood. This relationship can be considered somewhat strong for “North-eastern Highlands” where the maximum values are linked to a probability of only 5%.

These findings seem to go against previous knowledge on this subject. In the European Mediterranean, tourism is known to provide a source of growing pressure on the environment, which translates specifically into an increase in wildfire incidence (Ganteaume & Jappiot, 2013). These results suggest that either this trend is not present in central Portugal, particularly in inland areas, or that this variable is not appropriate for measuring touristic activity.

The proportion of resident population with a university degree (RPSUP_PERC) follows a distinct pattern. It seems to be linked to urban dynamics, where most highly educated people are located, and therefore results show higher large wildfire ignition probabilities in connection to small percentages of university educated residents. “North-eastern Highlands” displays a stronger association than the rest of the models, spanning from around 22% probability at 0% university educated inhabitants to less than 5% probability at maximum values.

Input from a study set in Argentina shows that education is negatively associated with wildfire ignitions, mainly because educated people have more access to information on fire prevention, are aware of the damaging effects of wildfires and understand the importance of preserving the environment (Dondo Bühler et al., 2013). Nevertheless, as mentioned before, the inhabitants of urban centres in the study area are, comparatively speaking, the people displaying the highest education level. Knowing that fire ignitions occur outside of these locations, this driving factor can be assumed to mask the effect of the urban dynamics and fuel availability.

In fact, the distance to urban areas (URB_DIST) seems to be greatly related to large wildfire ignition likelihood. Most models show expressive negative relationships (see Annex D), with the highest probabilities of large ignitions occurring within 0 to 10 km from urban centres and infrastructures. “Urban Coastline”, which corresponds to the most urban region, displays the opposite link, with maximum distance from urban centres (35 km) showing an associated ignition likelihood of 40%.

As these correspond to ignitions originating large burned areas, increasingly urban locations tend to enable earlier wildfire containment, due to the presence of population and better accessibility, which is the case of “Urban Coastline”. This fact might explain the regional variability found with this driving factor.

The distance to secondary roads also seems to display pronounced negative relationships throughout the study area. Higher ignition probabilities are linked to the proximity to these infrastructures (maximum for “North-eastern Highlands” at 20%). After 4 km all models show an approximately 0% likelihood of large wildfire ignition. As was the case with urban proximity, higher accessibility and human presence have also been the aspects mentioned in connection to the roads’ effect on ignition probability (Martínez-Fernández et al., 2013; Moreira et al., 2010). Thus, these results suggest that large fire ignitions have a similar behaviour to that of wildfires of all sizes, in respect to road proximity.

Land cover types seem to be associated to the variability in the likelihood of large ignitions. The percentage of pine forest cover (PIN_COS) shows a weak negative connection, with higher probabilities linked to small areas or to an absence of pine trees (maximum for “North-eastern Highlands” at 20%).

The shrub extents (SHRUB_COS), on the other hand, show different trends. The global model and “North-eastern Highlands” display a positive relationship with ignition probability, prominently in the latter case where 100% shrub cover corresponds to a likelihood of over 30%. This area shows a considerable presence of this type of vegetation. For “Central Forests” the relationship between shrub and ignition is negative, with higher probabilities occurring in other cover types (0% shrub, at roughly 15% incidence likelihood).

In the literature, vegetation types are primarily related to fire propagation and not ignition likelihood. However, apart from associating shrublands to faster fire spread, Martínez-Fernández et al. (2013) and Nunes et al. (2016) link this specific cover type to a decreased value, corresponding to one of the least priorities during firefighting. Both these reasons explain that ignitions occurring in shrublands develop into large fire events before being extinguished.

Two other biophysical aspects show an association with the probability of large wildfire ignition occurrences, although to different extents: precipitation and slope. The total annual precipitation displays a strong positive relationship in the case of "North-eastern Highlands" and "Central Forests" and the global model. In fact, for "Central Forests" this positive link is very pronounced ranging from 5% likelihood at the minimum precipitation values (600 mm) to almost 55% at the maximum (1300 mm). In the case of "Beira Baixa and Border Areas", the association is negative, even though not strong (see Annex D).

These findings, which show an important connection between precipitation totals and large wildfire ignition, are consistent with common knowledge in this field of study. In fact, Nunes et al. (2013) show that the rainiest areas in Portugal correspond to the highest ignition incidence zones. It is then possible to assume that this trend also applies to large wildfires.

Strong positive relationships are also found with slope for all models, except "Beira Baixa and Border Areas". This statement is particularly apparent in the case of "Urban Coastline", which displays a very fast rise in ignition probability with an increase in terrain gradient. For maximum slope (26°) the associated probability is approximately five times the likelihood of large ignition in the case of no slope.

These results seem to contradict previous knowledge that ignitions tend to occur in flatter, more accessible and densely populated areas (Ganteaume & Jappiot, 2013), given the overwhelming effect of human interference. However, the fact that only ignitions originating large burned areas are being studied, events taking place in rugged surfaces are more likely to spread faster and present a higher fire intensity, because of inclination, land cover and low accessibility (Salis et al., 2015).

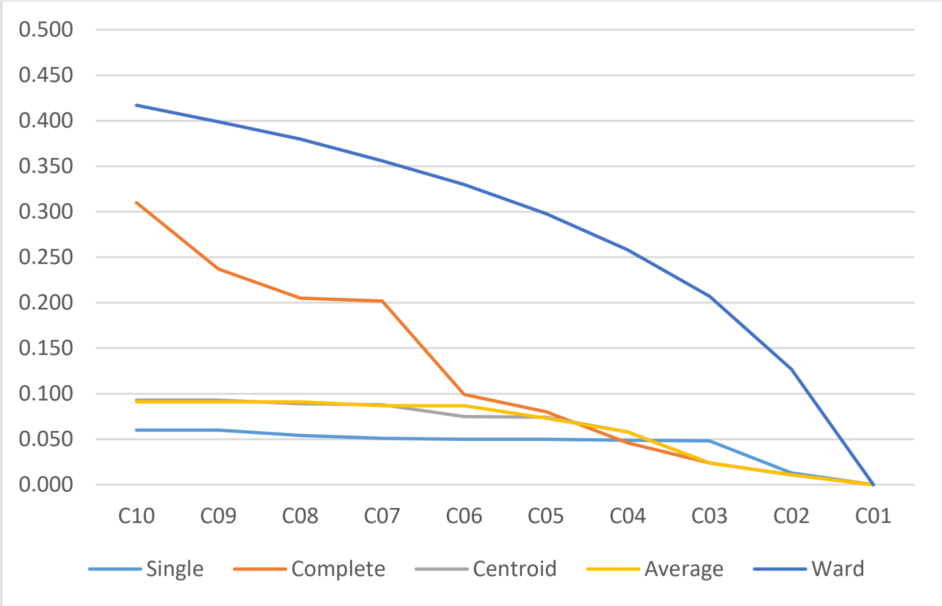
The rest of the driving factors included in the models failed to provide enough evidence to support further discussion (see Annex D). The Wildland-Urban Interface (WUI) has not shown considerable differences in ignition probability, depending on category. Previous studies had already shown that the relationship between WUI areas and large wildfires in Portugal is not significant (Modugno, Balzter, Cole, & Borrelli, 2016).

4.4. LARGE WILDFIRE SPREAD

The same methodological steps employed before were repeated for the large wildfire spread analysis. All methodological decisions regarding common procedures of the clustering and modelling techniques were kept consistent throughout both analyses.

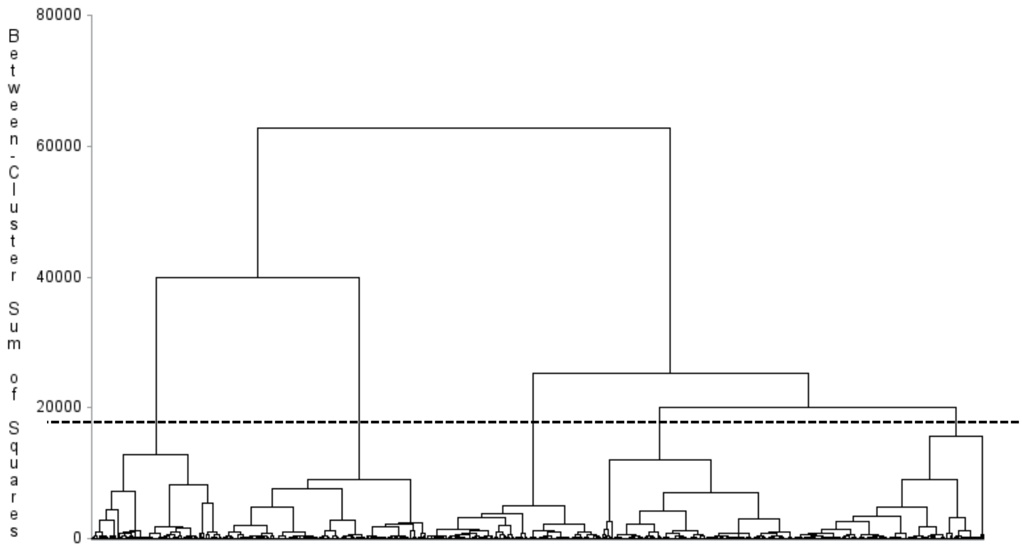
The best performing hierarchical clustering method was also Ward's method. This evaluation is made evident by the graph below (Figure 11), which shows the plot of the R^2 values for each clustering partition and hierarchical method:

Figure 11 – R^2 values associated to different clustering partitions and hierarchical methods



The Ward's method curve suggests that either four or three clusters might be a desirable result, although there is no clear sharp drop in explained variance visible. When interpreting the dendrogram (Figure 12), however, it is evident that the five clusters solution is the most balanced partition. It corresponds to an R^2 of 0.298, which cannot be considered a high value. It is higher than the previous ignition clustering solution, partly because it contains one more group.

Figure 12 – Ward's method dendrogram with the cut-off line at five clusters



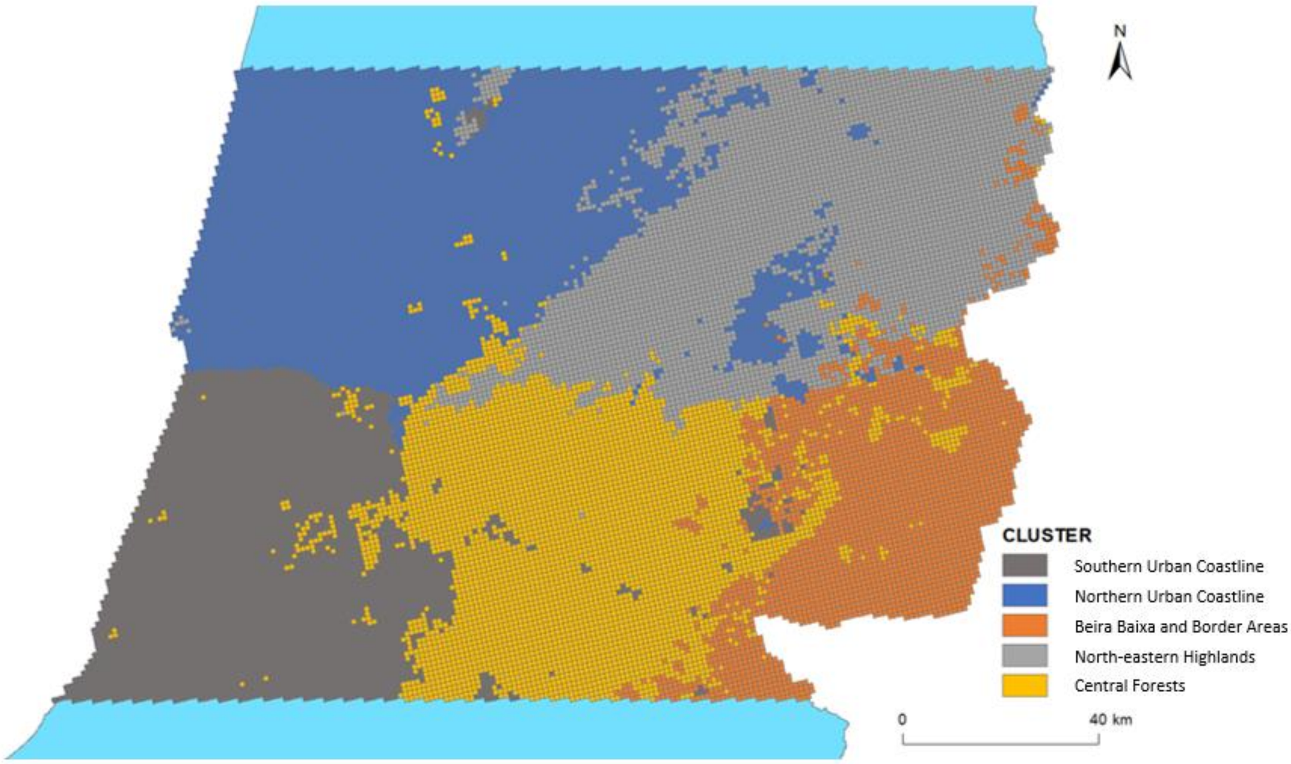
Similarly to what was done previously for the ignition clustering partition, two geographical variables representing the latitude and longitude values of grid cells were added to the analysis. This decision also improved the geographical boundaries of the different clusters.

K-means non-hierarchical clustering was then performed for five clusters. This process needed a total of 5 iterations for completion and improved the value of the R^2 to 0.324. Again, the small increase in explained variance is considered enough for the purpose of this study and given the characteristics of the data and the interpretability of the clustering solution.

The map presenting the clustering solution can be found below (Map 3). It is important to understand that not all variables displayed the same proportion of explained variance (R^2). The coordinates (X_COORD, Y_COORD), the average number of dry months (DRYMONTH), the elevation (ELEVATION) and the proportion of seasonal use, secondary residence or empty housing (SSEHOUS_PERC) exhibited the highest values (50% or above).

On the other hand, the average animal headcount in each farm per parish (LVSTK_NFARM), the population density (POP_GRID) and the orientation (ASPECT) showed the poorest performance among all variables, standing well below 10%. These variables were also the least well represented on the ignition clustering solution.

Map 3 – Clustering solution (5 clusters) based on large wildfire propagation driving factors



The following table (Table 6) summarises the clustering solution in regard to group dimension, cluster location and specific fire spread-related characteristics:

Table 6 – Description of the propagation clustering solution based on the selected variables

FEATURES	SOUTHERN URBAN COASTLINE	NORTHERN URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
Dimension	3,899	5,546	2,726	4,853	4,546
Population	Average to high population density,	Medium to high population density,	Average to low population density,	Medium to low population density,	Average to low population density,

FEATURES	SOUTHERN URBAN COASTLINE	NORTHERN URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
	young population with a medium potentiality index	young population with an average to high potentiality index	very old population with a medium potentiality index	relatively old population with a very low potentiality index	moderately old population with medium to high potentiality index
Roads and Housing	Average to high density of primary roads and high ongoing housing occupation	Medium to high road density and high ongoing housing occupation	Very low density of primary roads and high number of not permanently occupied housing	Average road density and high number of empty housing	Average density of primary road and considerable to high number of not permanently occupied housing
Agriculture and Livestock	Very few people employed in primary sector activities, agricultural activity is still present and grazing areas are less dominant, average to high density of small to medium farms, high livestock activity and very high density of animal headcount per utilised agricultural surface	Average to low proportion of population employed in the primary sector, average extent of agricultural areas but smaller grazing areas, very high density of small to medium farms, moderate livestock activity and animal density	High number of people employed in primary sector activities, large agriculture areas, especially for grazing, low density of high dimension farms, average to low livestock activity and low density of animal headcount	Considerable number of people employed in primary sector activities and medium to low extents of agriculture and grazing areas, moderate to low density of regular sized farms, medium to low livestock activity and low density of animal headcount	Average number of people employed in primary sector activities and very few agriculture and grazing areas, medium to low density of moderate to small farms, average to low livestock activity and animal density
Biophysical Aspects	Flat and low altitude landscape, moderate presence of eucalyptus and forests of other types, average to small shrub extents, extended drought season	Terrain is reasonably flat and at low elevation, medium to high presence of eucalyptus trees, average to low extents of shrublands and forests of other species, short drought season	Flat landscape at medium-low altitude, very few eucalyptus trees and shrubs, large forests of other types, extended drought season	Rugged surface at high elevation, very few eucalyptus trees, average to high presence of other types of forest and very large shrublands, short drought season	Rugged surface at medium elevation, high eucalyptus presence, medium to low extents of other species, moderate shrub presence, extended drought season
Ignition Density	Medium-high	Medium-high	Very low	High	Medium

It is interesting to observe the similarities between both clustering solutions (ignition and spread). The results are identical for “Beira Baixa and Border Areas” in both exercises, displaying the same features: extensive agriculture and grazing areas, low density of infrastructures and an ageing population. Additionally, it presents a very low ignition density. Although some of the selected variables were the same for both analysis, many were different, which suggests that the overall regional partition is consistent across several sets of variables.

Again, a great contrast between groups is visible, although “Southern Urban Coastline” and “Northern Urban Coastline” seem very close in respect to their main characteristics, exhibiting a distinct urban profile. In fact, they seem to be mostly distinguished based on the number of dry months (extended drought season in “Southern Urban Coastline” and smaller in “Northern Urban Coastline”), the livestock activity (high in “Southern Urban Coastline” and moderate in “Northern Urban Coastline”) and the animal density (high in “Southern Urban Coastline” and moderate in “Northern Urban Coastline”).

“North-eastern Highlands” and “Central Forests” also share a number of resemblances, including an ageing population and a high proportion of empty or secondary-use housing, although ignition patterns differ significantly. “North-eastern Highlands” is mainly defined in respect to biophysical aspects, including high elevations and large shrub extensions. “Central Forests” has very few agriculture and grazing areas and is instead dominated by eucalyptus forests.

As seen before, the overall clustering partition is coherent with previous knowledge on the characteristics of these regions, namely the urban coastline-rural inland dichotomy. When comparing these results with findings from other studies, connected to the different patterns in burned area among regions, some resemblances can be emphasised.

Shrublands are mentioned as one of the most significant wildland cover varieties in mountainous areas of northern and central Portugal (“North-eastern Highlands” and “Central Forests”), specifically in connection to their high flammability and preferential burning (Fernandes, Luz, & Loureiro, 2010). These vegetation features are coupled with a rugged landscape, a higher fuel connectivity and a lower population density, favouring fire spread (Mateus & Fernandes, 2014). In fact, larger burned areas in Portugal concentrate in these regions (Benali et al., 2016; Tedim et al., 2013).

Nunes et al. (2013) also point out the northern-central districts of Guarda, Viseu and Castelo Branco and southern-central Santarém (districts composing part of “North-eastern Highlands” and “Central Forests”) as regions affected by large wildfires. They discuss the essential role of socioeconomic transformations, such as the rural exodus, and favourable climatic conditions (higher precipitation values) for the proliferation of available fuel in uncultivated areas.

The characteristics of the coastal areas contrast with these patterns, particularly in the north-western pocket of the study area (“Northern Urban Coastline”). Even though ignitions are concentrated along the coast and around urban areas (medium to high ignition density), higher population density, fragmented settlements, intense agricultural activities, higher accessibility and fire suppression efforts prevent the occurrence of large wildfire events (Barros & Pereira, 2014).

The first step of the two-part modelling exercise for burned area follows the same methodology as before. Probit models were developed for the entire study area and for each region according to the clustering solution. The results of this analysis are presented below, including the estimated models quality assessment measures (Table 7), as well as the average partial effects (APE) and statistical significance of the estimates (Table 8).

Table 7 – Estimated models quality assessment measures (fire spread)

MODEL ASSESSMENT	GLOBAL	SOUTHERN URBAN COASTLINE	NORTHERN URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
Pearson χ^2	272234.06 ***	4525.97 ***	4862.66	2594.80	18346.06 ***	3733.25
Sensitivity	42.58%	47.76%	39.33%	2.08%	64.14%	50.86%
Correct classifications	80.06%	82.92%	83.01%	92.81%	72.88%	81.74%
Area under the ROC curve	84.44%	85.39%	86.33%	84.07%	81.51%	86.40%

Significance of the Pearson χ^2: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$)

The ROC curve plots can be found in Annex C, along with the classification tables. From the goodness-of-fit measures it is possible to recognise that the fire spread models performed considerably better than the attempts at predicting ignition locations, especially when considering sensitivity values and the area under the ROC curve. This might be easily explained by the increase in positive response cells (corresponding to area burned by large wildfires), roughly double the generalised ignition cells.

Nevertheless, three models displayed non-significant Pearson χ^2 values and the percentage of accurate positive classifications for Cluster 3 (both the smallest and the least affected by fire) is still very small.

Table 8 – Average partial effects (APE) and estimates significance (fire spread)

VARIABLES	GLOBAL	SOUTHERN URBAN COASTLINE	NORTHERN URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
PRIM_PERC	-0.00387 ***	-0.00599 ***	-0.00442 ***	-0.00332 ***	-0.00267 ***	-0.00383 ***
FARMDEN_KM						
SAUFARM_HA	0.000896 ***	-0.01097 ***			-0.00384 ***	
LVSTK_NFARM	0.00000765 ***		-0.0000239 ***	0.000499 ***	0.00026 ***	
HEADS_NSAU		-0.00602 ***			-0.05414 ***	-0.0110341 **
SSEHOUS_PERC	0.002147 ***	0.001721 ***		0.001509 ***	0.004694 ***	
POTENT_INDEX	-0.00765 ***		-0.00215 **	-0.00621 ***	-0.01081 ***	-0.00431 ***
AGE_INDEX	-0.0000381 ***		0.00016 **		-0.0000336 ***	-0.0000232 **
POP_GRID	-0.0000644 ***		-0.0000243 *			
AGR_COS	-0.00074 ***		-0.00092 ***	-0.00034 *	-0.00319 ***	
EUC_COS	0.001402 ***	0.004372 ***	0.000937 ***	0.000547 **	-0.0031 ***	0.002259 ***
GRZ_COS						

VARIABLES	GLOBAL	SOUTHERN URBAN COASTLINE	NORTHERN URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
OUTR_COS		0.000614 *			-0.00303 ***	
SHRUB_COS	0.0027 ***	0.004314 ***	0.006178 ***		0.001626 ***	0.003562 ***
AP2015_DIST	0.001526 ***	0.001709 **		0.001672 *	0.007463 ***	0.00385 ***
PROAD_DIST		0.010724 ***	0.018557 ***	-0.00194		0.008157 ***
DRYMONTH			0.131673 ***	-0.07283 ***	-0.28573 ***	-0.10114 ***
SLOPE	0.005608 ***	0.014619 ***	0.013525 ***		0.003301 **	
ASPECT	0.000117 *					
ELEVATION		0.000339 ***		-0.00015 ***		-0.00034 ***
IGN_DIST	-0.0654134 ***	-0.06913 ***	-0.07194 ***	-0.01363 ***	-0.07668 ***	-0.07202 ***

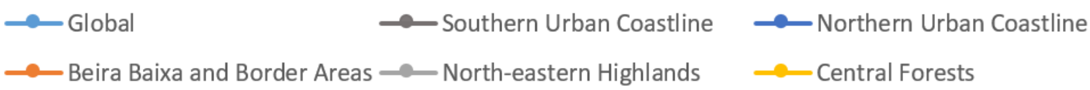
Significance of average partial effects (APE): *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$)

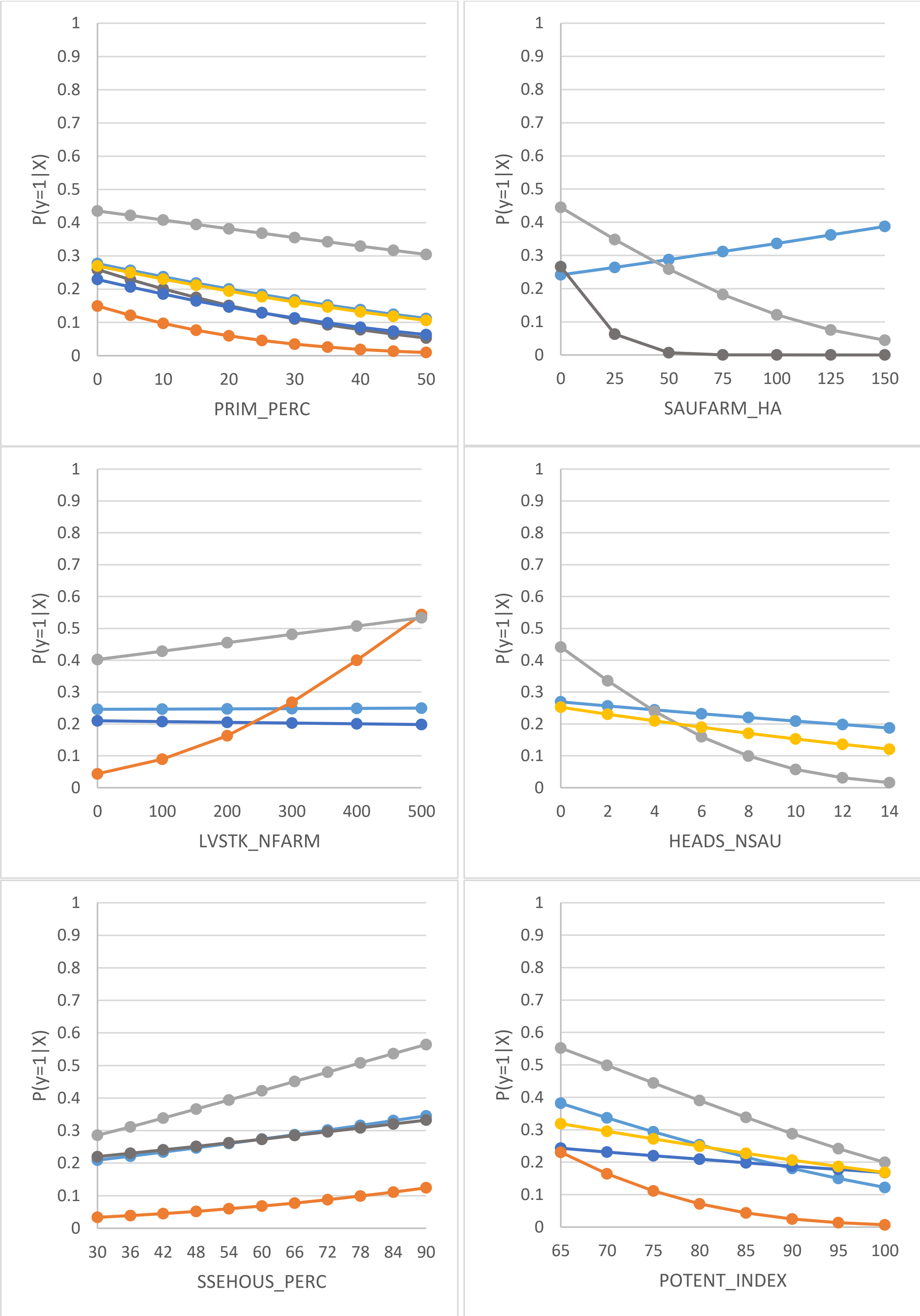
The table above (Table 8) presents the estimated average partial effects (APE) and associated statistical significance. The models for Clusters "Beira Baixa and Border Areas", "North-eastern Highlands" and "Central Forests" and the global model have exhibited a significant (***) positive constant ($\hat{\beta}_0$). For "Southern Urban Coastline" and "Northern Urban Coastline" the intercept was negative, with different significance levels ($p < 0.01$ and $p < 0.1$ respectively).

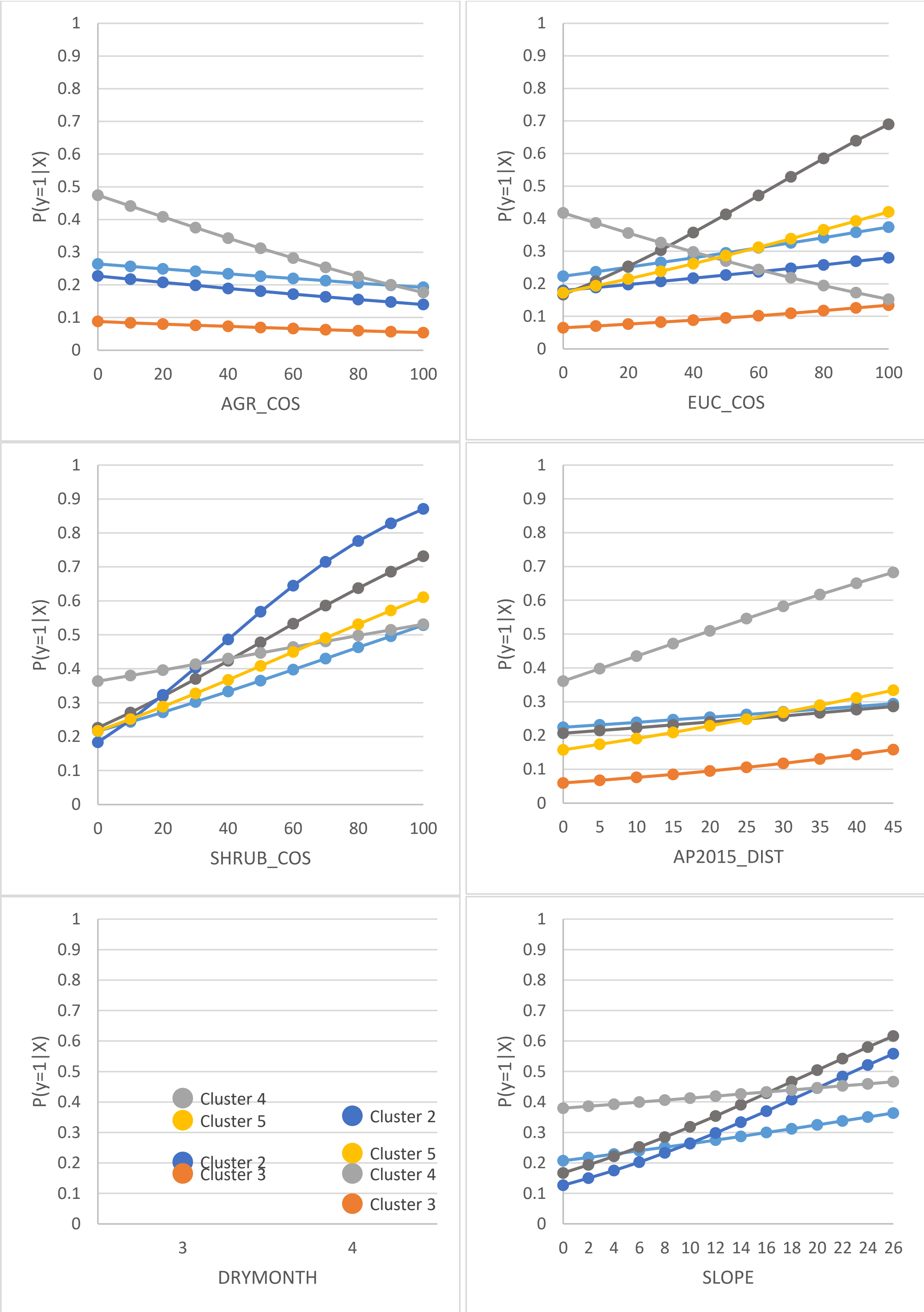
Fire spread model specification left out the number of farms per km^2 (FARMDEN_KM) and the percentage of grazing areas (GRZ_COS), suggesting that these two factors are not determinant for the propagation of large wildfires across the study area. Three other variables were deemed significant both in the global model and in all five regions: the percentage of resident population employed in primary sector activities (PRIM_PERC); the percentage of land covered by eucalyptus forests (EUC_COS); and the distance to ignition locations (IGN_DIST), which is unsurprisingly and by far the most important factor. The estimated APEs for both PRIM_PERC and IGN_DIST display the same signal direction throughout the study area, which conveys a consistency in these variables' impact on fire behaviour.

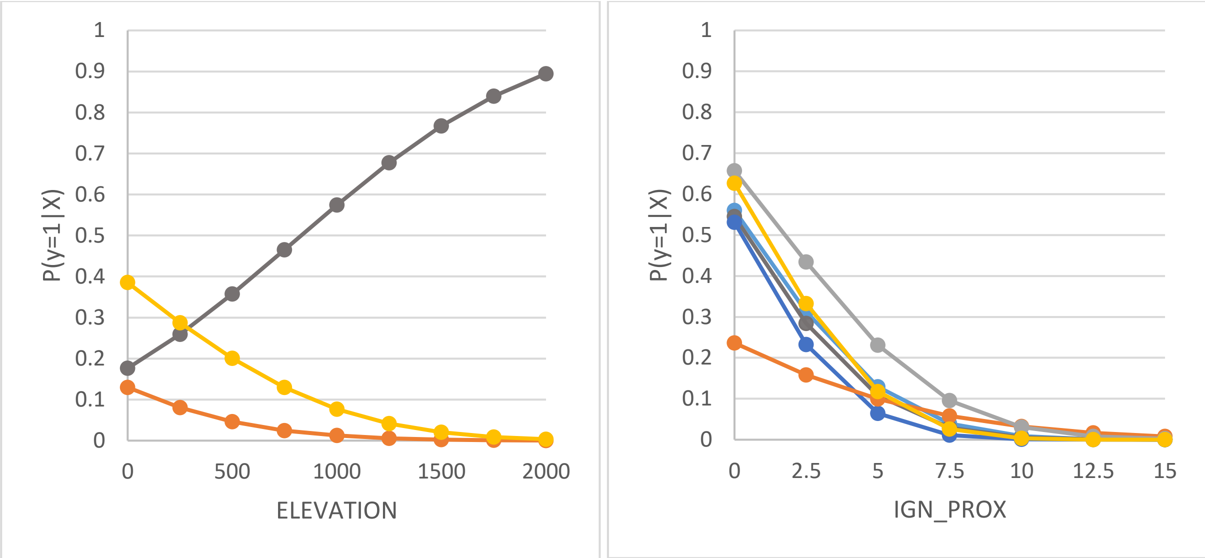
Specific driving factors have been highlighted for discussion, as before, because of their impact on large fire spread, predictive ability and inconsistency across clusters. The following figure (Figure 13) summarises the main results of the six models, displaying the associated burning probabilities for specific values within the range of the contemplated variables. For a better interpretation of these results, it is also important to consider the individual variable plots per cluster, which present the estimates' corresponding 95% confidence intervals, in Annex D.

Figure 13 – Main results of the fire spread model (part one)




P(y=1|X)
PRIM_PERC
P(y=1|X)
SAUFARM_HA
P(y=1|X)
LVSTK_NFARM
P(y=1|X)
HEADS_NSAU
P(y=1|X)
SSEHOUS_PERC
P(y=1|X)
POTENT_INDEX


P(y=1|X)
AGR_COS
0 20 40 60 80 100
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
EUC_COS
SHRUB_COS
AP2015_DIST
0 5 10 15 20 25 30 35 40 45
DRYMONTH
3 4
Cluster 4
Cluster 5
Cluster 2
Cluster 3
SLOPE
0 2 4 6 8 10 12 14 16 18 20 22 24 26



The results of the second part of the burned area modelling exercise provide further insights about the impact of each driving factor on fire propagation in the case of large wildfire events, complementing the findings from the binary response first part models. The next step of the modelling procedure fitted OLS regression lines to each of the study regions, following the methodology presented before.

Table 9 presents the six models' quality assessment measures, as well as the dimension of the clusters population ($y = 0$). The results of the modelling exercise (second part) are summarised in Table 10, with an emphasis on the coefficients' signs and statistical significance.

Table 9 – Estimated OLS models dimension and quality assessment measures (fire spread)

MODEL ASSESSMENT	GLOBAL	SOUTHERN URBAN COASTLINE	NORTHERN URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
Dimension (n)	5,343	917	1,139	192	1,994	1,101
F statistic	89.34 ***	36.31 ***	44.10 ***	6.06 ***	28.50 ***	31.47 ***
Adjusted R^2	0.188	0.316	0.294	0.096	0.181	0.200

Significance of the F statistics: *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$)

In addition to the presented quality assessment measures, residuals from all models were found to be at least approximately normally distributed. It is possible to observe that Cluster 4 is by far the region displaying the higher proportion of burned area cells (roughly 41%), while in Cluster 3 only 7% of the territory has experienced burning by large wildfire events during the 2005-2015 period.

The F statistic values are significant for all models, which is good indication of model suitability. However, it is also possible to recognise that the adjusted R^2 values are low in all cases, mainly for Cluster 3, although these results should be considered lightly. As mentioned before, the purpose of this study is merely exploring the relationships and behaviour of the different factors driving large wildfire events, and not necessarily developing a predictive model for large fire propagation, hence the flexibility in the measurements of the quality assessment statistics.

Table 10 – OLS coefficients, associated t-scores and statistical significance (fire spread)

VARIABLES	GLOBAL	SOUTHERN URBAN COASTLINE	NORTHERN URBAN COASTLINE	BEIRA BAIXA AND BORDER AREAS	NORTH-EASTERN HIGHLANDS	CENTRAL FORESTS
CONSTANT ($\hat{\beta}_0$)	103.231 *** t = 11.16	29.02 *** t = 3.48	71.0937 *** t = 4.00	40.0369 *** t = 4.13	153.46 *** t = 10.05	106.04 *** t = 3.96
PRIM_PERC		-1.29375 *** t = -4.57	0.71687 *** t = 2.76	-0.60788 ** t = -2.14		-0.47366 ** t = -2.56
FARMDEN_KM		-1.03295 ** t = -2.11				-1.59624 ** t = -2.57
SAUFARM_HA					-0.23557 ** t = -2.10	
LVSTK_NFARM	0.00107 *** t = 3.14	0.00069 ** t = 2.01				
HEADS_NSAU	0.39146 * t = 1.88	0.76439 *** t = 2.78	-1.69891 ** t = -2.19		-7.00184 *** t = -5.34	
SSEHOUS_PERC	0.13319 *** t = 3.70	0.56645 *** t = 5.43				
POTENT_INDEX	-0.67417 *** t = -6.77		-0.39672 * t = -1.86		-0.973 *** t = -5.34	-0.4858 ** t = -1.99
AGE_INDEX					-0.00324 ** t = -2.18	
POP_GRID	-0.01909 *** t = -5.21		-0.01133 *** t = -2.85		-0.03855 *** t = -2.88	
AGR_COS	-0.4425 *** t = -13.58	-0.35647 *** t = -6.20	-0.43232 *** t = -7.19	-0.23051 * t = -1.95	-0.64592 *** t = -10.47	
EUC_COS	0.22741 *** t = 10.69	0.48727 *** t = 9.76	0.26358 *** t = 6.82		-0.34679 *** t = -2.75	0.44333 *** t = 12.13
GRZ_COS	0.55225 *** t = 7.93			0.41351 *** t = 2.97	0.55249 *** t = 2.84	
OUTR_COS		0.19743 *** t = 3.94			-0.36986 *** t = -7.87	
SHRUB_COS	0.24486 *** t = 11.41	0.50703 *** t = 3.61	0.54626 *** t = 5.66		0.08008 *** t = 2.77	0.49115 *** t = 6.51
AP2015_DIST	0.10367 ** t = 2.25		-0.35040 *** t = -2.60		1.10158 *** t = 7.90	0.28249 *** t = 2.78
PROAD_DIST			1.14231 * t = 1.86		1.18853 *** t = 3.09	
DRYMONTH	-2.57044 ** t = -2.33					-8.75216 ** t = -2.30
SLOPE	0.61038 *** t = 5.73		1.52079 *** t = 5.73	2.24835 * t = 1.88	0.57521 *** t = 3.00	1.33822 *** t = 4.41
ASPECT	-0.02087 * t = -1.86	-0.07253 *** t = -2.77			-0.03574 ** t = -2.18	
ELEVATION		0.02316 ** t = 2.54			-0.01151 *** t = -3.46	-0.03843 *** t = -5.05
IGN_DIST	-1.16798 *** t = -4.3	-2.40021 *** t = -3.82	-2.83019 *** t = -4.03		-1.38465 *** t = -2.75	

Significance of average partial effects (APE): *** ($p < 0.01$), ** ($p < 0.05$), * ($p < 0.1$)

A rise in the proportion of residents employed in primary sector jobs (PRIM_PERC) was found in the previous exercise to increase the probability of large fire ignitions in most regions. Nevertheless, these results suggest that population working on the primary sector impacts fire spread in the opposite direction, regardless of geographical location, with all clusters displaying a negative relationship between this variable and the likelihood of large fire spread. The maximum burning probability (higher than 40%) is listed at 0% primary sector working force, for “North-eastern Highlands” (Figure 13).

The results from the models second part also support this evidence, with “Southern Urban Coastline”, “Beira Baixa and Border Areas” and “Central Forests” presenting a negative relationship between PRIM_PERC and the percentage of burned area per cell (Table 10). This relationship can be considered very strong in the case of “Southern Urban Coastline”. For “Northern Urban Coastline”, however, the results suggest that the presence of people employed in primary sector activities moderates fire spread probability in this region, unless a cell does burn. Even though not strong, this information diverges from the results of the models first part (probit model).

Farm dimension (SAUFARM_HA) can only be considered in terms of spread probability. The global model seems to associate larger farms to a higher burning probability, which could be linked to the fact that these are located in sparsely populated locations outside urban centres. Nevertheless, “Southern Urban Coastline” and “North-eastern Highlands” present a decrease in burning probability with an increase in farm size, quite prominent in both situations. Large agricultural holdings may be connected to land management activities, which may in turn translate into fire prevention efforts and a reduction in fuel availability.

The influence of agricultural areas (AGR_COS) on burning probability is similar to that displayed by PRIM_PERC and can be considered more pronounced in the case of “North-eastern Highlands” (Figure 13). This region exhibits the highest large fire spread probability (roughly 50%) at 0% agricultural land cover, while the behaviour of “Beira Baixa and Border Areas” is very weak (see Annex D).

When considering the percentage of burned area per cell, the effect of agricultural land cover is very similar, with all models displaying negative associations between both phenomena (Table 10). Again, the weight of this variable is particularly relevant for “Northern Urban Coastline” and “North-eastern Highlands” and for the global model, presenting very high t-scores, and barely significant for “Beira Baixa and Border Areas”.

Many arguments support the findings from this study, which generally agree that agricultural activities work as a deterrent factor of large burned areas. Wildfires occurring in agricultural land covers are not expected to develop into high intensity events, mainly due to a small amount of fuel load, predominantly dry fine fuels (Mitsopoulos et al., 2014). Many are the authors linking the process of agriculture, pastures and forestry abandonment, which occurred in European Mediterranean regions during the second half of the 20th century, to an increase in size an intensity of wildfires as a result of an accumulation of connected flammable materials (Calviño-Cancela et al., 2016; Moreira et al., 2011; Viedma, Moity, & Moreno, 2015).

Another reason explaining the low probability of fire spreading into agricultural spaces, or else a higher burning probability in wildland areas, has to do with the interaction between agriculture and topography. Farms are usually located in flat lands and slope is known to strongly affect fire spread (Calviño-Cancela et al., 2017).

Agricultural activities are also associated to the presence of humans in rural areas, which supports earlier fire detection and more efficient firefighting (Moreira et al., 2011). Farm management requires that holders be vigilant and mindful of their property (Calviño-Cancela et al., 2016).

Nunes et al. (2016) have found the same overall patterns as this work for the relationship between agricultural activities and reduced burned area, in Portugal, in a study conducted at municipal level. However, a somewhat contrasting trend has been observed by Martínez-Fernández et al. (2013), in Spain, where tightly knit forest-agriculture mosaics are deemed the most fire-vulnerable regions.

The effect of animal creation variables on burning probability is mixed and is very analogous to what was previously found for large fire ignition. A strong positive association is observable for the average number of animal per farm (LVSTK_NFARM) in “Beira Baixa and Border Areas” and “North-eastern Highlands”, with 500 animals corresponding to a burning probability of over 50% in both regions. The results for “Northern Urban Coastline” and the global model in the first part of the model are inconclusive, while the second part has shown a positive, although small contribution of this variable, possibly connected to foraging necessities and use of fire practices.

Animal density (HEADS_NSAU), on the other hand, exhibits a negative relationship with large fire spread probability for the global model, “Central Forests” and, particularly, “North-eastern Highlands”. In the latter case, burning probability drops from around 45% at no normal heads per hectare, to virtually 0% when animal density is highest. Therefore, it is interesting to observe that the outcomes of the models’ second part differ among themselves, with “Northern Urban Coastline” and “North-eastern Highlands” displaying a negative association, and the global model and “Southern Urban Coastline” presenting the opposite trend.

The contribution of grazing lands (GRZ_COS) can only be assessed in terms of burned area percentage, showing a positive relationship, namely for the model comprising the burned cells of the entire study area, which presents a very significant coefficient estimate (Table 10). “Beira Baixa and Border Areas” and “North-eastern Highlands” also display a positive association, and higher burned area percentages in grazing areas might be explained by fuel availability, as there is an evident contrast with the findings from the ignition model.

Additionally, many authors have highlighted the role of fire for clearing land for grazing purposes in Mediterranean areas across the globe and in Portugal (Álvarez-Díaz et al., 2015; Ferreira-Leite, Lourenço, et al., 2013; Ganteaume & Jappiot, 2013; Vilar et al., 2016). This phenomenon might also explain why pastures are associated with increased burning probability.

However, animal creation and livestock grazing have been found by Moreira et al. (2011) to contribute to a reduced fire hazard, because these activities naturally control fuel availability and density. This fact may help justify the effect displayed by animal density in most clusters. In this context it is important to consider animal type for the influence exerted by different species (Oliveira et al., 2014), and this element was not contemplated in this analysis, which might have had an influence in obtaining these results.

This said, evidence from the literature highlights the different impacts pastures and animal breeding display on large wildfire occurrence. Oliveira et al. (2017) have found that Portuguese parishes affected by larger burned areas still have a considerable presence of livestock and grazing, although human activities in rural areas are linked to less burned area, a trend mentioned before in connection to agriculture.

Fuel availability represents one of the most relevant contributing factors to large wildfire spread, as gathered from the models results. The extent of eucalyptus tree cover (EUC_COS) displays noticeable positive associations with burning probability in all regions, except for “North-eastern Highlands” (Figure 13). “Southern Urban Coastline” presents a very pronounced rise in burning probability in connection with eucalyptus forests (ranging from 20% to 70%), while in “North-eastern Highlands” the opposite relationship is also prominent (ranging from 40% to approximately 10%). Model results from the second part follow the same behaviour (i.e. a negative association in “North-eastern Highlands”), with “Southern Urban Coastline”, “Central Forests” and the global model presenting very high t-scores.

Shrublands (SHRUB_COS), on the other hand, have been found to encourage burned area probability to a great extent in all regions. This pattern is observable in both parts of the model, with only “North-eastern Highlands” displaying a weaker association. It is relevant to point out the burning probability corresponding to 100% shrub cover in “Northern Urban Coastline”, which is roughly 90% (Figure 13). Among the OLS coefficient estimates the global model stands out with a very high t-score (Table 10), meaning that this factor is very important for large fire spread overall.

Generally speaking, forests are considered more fire prone than farms, with eucalypt plantations displaying the same fire hazard as pine stands (Moreira et al., 2011). Effects connected to the economic value and active management of paper and pulp production eucalyptus forests might mitigate the influence of fuel availability and other favourable burning conditions (Barros & Pereira, 2014). Additionally, it is interesting to note that according to both these authors shrublands are even more susceptible to fire than forests, which might explain why the effect of eucalyptus stands in burning probability was negative in the shrub-dominated region of “North-eastern Highlands”.

Comparing these conclusions with the results from the previous ignition exercise provides that vegetation characteristics are much more relevant for fire propagation than fire incidence. For shrublands, where large ignition and burned area probability display a strong positive association with the amount of available fuel, these findings suggest that these areas should be prioritised in large fire prevention efforts.

Other forest types, apart from eucalyptus, pine trees and shrublands, present mixed results across regions, displaying a positive relationship with the percentage of burned area in “Southern Urban Coastline” and the opposite trend in “North-eastern Highlands”. Types of vegetation cover other than eucalyptus, pine trees and shrublands, such as oak forests, are known to be less fire prone (Mateus & Fernandes, 2014), and this pattern can be observed in “North-eastern Highlands”. However, in “Southern Urban Coastline”, the positive link, which is not particularly strong, might be connected to fuel availability, given the fact that this region has a distinct urban profile.

All models presented a clear positive association between the distance to protected sites (AP2015_DIST) and the probability of large fire spread, meaning that areas further from these locations are expected burn in large wildfire events. This statement is particularly true for “North-eastern Highlands”, where associated probabilities range from 35% to 70% (Figure 13). The majority of the estimated OLS coefficients also support these results, with “North-eastern Highlands” displaying a very high positive t-score (Table 10). “Northern Urban Coastline”, on the other hand, exhibits a negative, although weak connection to the percentage of burned area, which suggests that the weight of this variable fluctuates across space.

Protected areas are generally assumed to influence fire occurrence, either as a deterrent factor connected to landscape protection (Rodrigues et al., 2014, 2016), or related to an increase in ignitions resulting from conflicts about the establishment of these protected sites (Calviño-Cancela et al., 2017; Fuentes-Santos, Marey-Pérez, & González-Manteiga, 2013). The results from this study are consistent with the views of Srivastava et al. (2014) on this topic, which associate protected areas to increased success in efficient wildfire suppression, and its spatial variability is also noted by Rodrigues et al. (2014, 2016).

There is a noticeable positive link between the percentage of seasonal, secondary use or empty dwellings (SSEHOUS_PERC) and the probability of large wildfire spread. This pattern is particularly evident in "North-eastern Highlands", where 90% empty housing corresponds to 60% burning probability. However, apart from "Southern Urban Coastline" and the global model, this variable does not seem to be connected to burned area dimension in a cell affected by burning. This driving factor might be masked by the urban-rural dichotomy and its characteristics, with most empty houses being located in country regions as a result of land abandonment.

In fact, Oliveira et al. (2017) mention that secondary houses are also located in sparsely populated and ageing inland areas, to where emigrants return during the summer months, when there is an intensification of outdoor activities. These same authors associate these dwelling characteristics (as well as degraded housing conditions) with increased burned area, supporting the findings from this work and extending the scope of influence of this factor, which had already been found to stimulate fire ignition in Spain (Romero-Calcerrada et al., 2010, 2008). As an explanation, during fire suppression activities, these infrastructures are considered less of a priority when compared with primary homes.

Demographic potential (POTENT_INDEX) seems to be negatively associated with burning probability throughout the study area, with "Beira Baixa and Border Areas" and "North-eastern Highlands" displaying the most pronounced drops in large fire spread chance with an increase in the potentiality index (Figure 13). These same results are also observable in the models second part, where estimated coefficients support these negative relationships and where "North-eastern Highlands" and the global model stand out (Table 10).

The population density (POP_GRID) can also be moderately associated with burned area extent in the same direction (models second part). Although ageing (AGE_INDEX) has not exhibited a strong connection with the phenomenon under study, these outcomes suggest that large wildfires tend to spread into demographically depressed regions (less population and demographic potential).

Many literature references support these claims. Communities affected by larger burned areas in Portugal are known to have suffered from depopulation, as young and better educated people migrated from inland and mountainous areas to coastal regions, and which in turn render these locations increasingly vulnerable to fire (less intervention in prevention strategies) (Oliveira et al., 2017). The influence of population potential, which is generally connected to human presence and density of human activities, has been found to both increase the probability of fire ignition and reduce the probability of larger burned areas, in Spain (Martínez-Fernández et al., 2013; Rodrigues & De la Riva, 2014).

Biophysical factors can also be counted among the contributing factors of large fire spread. Terrain slope (SLOPE) seems to display a positive relationship with burning likelihood, particularly noticeable for “Southern Urban Coastline” and “Northern Urban Coastline” (Figure 13). The fact that these clusters stand out from the rest of the study area might be due to the fact that urban areas are located near the ocean, in flat landscapes. The probabilities associated to the different slope values range approximately 40 p.p. in these regions, with maximum gradient corresponding to more than 60% burning probability in “Southern Urban Coastline”.

The estimated OLS coefficients also exhibit positive associations for this variable, with highest t-score in “Northern Urban Coastline” and on the global model (Table 10). These results reveal that ignition and propagation patterns are similar for large wildfires in what concerns slope.

Slope is known to influence fire spread both directly and indirectly, because of fire dynamics (flames closer to ground fuel) and fuel moisture and density patterns (Holsinger et al., 2016). A similar reason has been found by Kalabokidis et al. (2007) to justify the positive relationship between slope and fire growth in a Mediterranean region in Greece.

The explanation found for slope, in connection to coastline regions, can also be used to account for the pronounced relationship found between an increase in elevation (ELEVATION) and an increase in large fire spread probability, in “Southern Urban Coastline” (Figure 13). In fact, the variable plot shows a very steep rise in burning chances for this region, while “Beira Baixa and Border Areas” and “Central Forests” exhibit the opposite trend. The models second part shows similar mixed results, if not exactly for the same clusters (Table 10). “Southern Urban Coastline” presents a moderate positive association with burned area whereas for “North-eastern Highlands” and “Central Forests” the relationship is negative and stronger.

Evidence from Australia has found that fires tend to become larger at higher altitudes, with this fact being connected to limited fire suppression activities at such elevations (Price, Penman, Bradstock, & Borah, 2016). The same pattern has also been confirmed in the European Mediterranean context (Silva, Rego, Fernandes, & Rigolot, 2010). The results obtained from the models seem to point out that elevation promotes large fire spread in densely populated regions (such as “Southern Urban Coastline”), while in inland rural regions large wildfire events occur at lower altitudes, where there is more available fuel. These findings are not exactly consistent with previous knowledge on this subject which associate larger burned areas to higher elevations in North and Central Portugal (Mateus & Fernandes, 2014), meaning further investigation into the effects of this topographic factor is needed.

The number of dry months (DRYMONTH) also shows mixed results throughout the study area, with “Northern Urban Coastline” presenting a positive association with burning likelihood and the other three clusters displaying a negative relationship (Figure 13). Contrary to previous knowledge on this subject, which states that a longer dry season increases the chances of fire spread (Ganteaume & Guerra, 2018), only in “Northern Urban Coastline” a higher number of dry months increases burning probability. The estimated OLS coefficient values exhibit weak negative associations in only two situations (Table 10), leaving these trends in need of additional corroboration and suggesting that the influence of drought on large burned area should be assessed resorting to other data and methodologies.

Last of all contemplated driving factors, the distance to ignition points presents relevant negative associations with large wildfire propagation likelihood everywhere in the study area (Figure 13). This effect on probability follows the same behaviour in all regions, with a sharp decline in spread probability with the increase in distance to ignition locations by 1km.

Although this variable represents one of the most important elements driving fire spread, it is not as associated to the extent of burned area, given that only three clusters and the global model present moderate positive t-scores on the models second part (Table 10). It appears that specific characteristics of the cells affected by fire are determinant for the percentage of burned area, with proximity to ignition playing only a minor role. Ignition locations were found to impact fire spread in Portugal, although it has also been suggested that it depends on complex interactions with biophysical elements and that it varies considerably across regions (Benali et al., 2016).

The rest of the driving factors included in this study failed to provide enough evidence to support further discussion (see Annex D). Nevertheless, it can be noted that the results of the models first part showed a positive relationship between distance to primary roads and the probability of fire spread for “Southern Urban Coastline”, “Northern Urban Coastline” and “Central Forests”, although not strong. This has also been found in the models second part for “Northern Urban Coastline” and “North-eastern Highlands”.

These results are consistent with the conclusions of Ager et al. (2014) who have found an increase in the probability of large fires with distance to roads, relating these infrastructures to enhanced fire suppression, and stressing the overall importance of this anthropogenic driving factor in the Mediterranean context. In Portugal, however, the contribution of this variable needs to be studied further, as results do not show any definite patterns.

5. CONCLUSION

Large wildfire ignition and propagation patterns in central Portugal, between 2005 and 2015, show a prevalence of fire events in inland central areas, with ignitions occurring mostly far from the eastern border and larger burned areas arising infrequently near the coastline. The overwhelming majority of these ignitions are human-caused, with many of these occurring intentionally.

Many of the most striking conclusions stemming from the analyses and following discussion, on the main factors driving large wildfire occurrence and spread in the study area, seem to corroborate information from the reference literature:

- **Agricultural activities:** Agricultural activities have been found to be connected to increased fire ignition, specifically in areas where most agricultural holders are over 65 years, which is instead an indication of the use of fire in land use activities. This study also confirms the spatial variability of this association. On the other hand, agriculture presents a negative relationship with large fire spread, a pattern that had already been found in other studies in the Portuguese context, and this connection appears to be mostly constant throughout the study area.
- **Precipitation:** Precipitation values have been found to be positively associated to large fire incidence. When confronting these results with findings from other studies, it is suggested that this relationship is valid for ignitions generating fires of all sizes in central Portugal.
- **Vegetation type:** Vegetation characteristics seem to be much more related to large fire propagation than to large fire ignition. Eucalyptus forests have been found to intensify fire spread, according to previous knowledge, although the real effect of this species varies across space. Shrubland areas, on the other hand, have a very expressive positive association with burned area, and this association seems mostly stationary.
- **Protected areas:** Protected areas are known as a deterrent factor of fire spread, as found by this work. However, these findings also confirm the spatial variability of this relationship, which had been previously identified in the reference literature.
- **Education:** Large wildfire incidence is negatively associated with the educational level, in particular higher education, and this relationship is well explained in the reference literature. In the context of this study, this aspect seems to be masking the effect of the urban areas.
- **Secondary roads:** The distance to secondary roads shows a positive relationship with large wildfire ignition, suggesting that this effect is valid for wildfires of all sizes in central Portugal.
- **Proximity to ignition:** The proximity to ignition points has been shown to be positively connected to fire spread in the entire study area. However, when considering the extent of burned area, the characteristics of the cells affected by fire appeared to be more determinant.

On the other hand, some of the findings of this research seem to contradict knowledge from previous sources and need further confirmation:

- **Livestock activities:** Livestock activities have been associated with a lower ignition probability. This is not consistent with the results from previous studies, which relate these activities to the use of fire. Additionally, this factor displays a marked variability throughout the study area.
- **Tourism activity:** Contrary to literary sources on this phenomenon, where increased tourism activity is known to impact fire ignition positively in the Mediterranean, the results from this study show the opposite pattern. This fact might question the suitability of the chosen variable.

- **Other vegetation types:** Other vegetation types, which are mostly different species of oak forests, are believed to prevent fire spread. This research suggests that in the study area fuel availability plays a more important role that fuel type when large fires are concerned. Nevertheless, this effect varies across space.
- **Slope:** This study provides evidence that the influence of slope in large fire ignition and propagation follows the same patterns (positive association). The connection between slope and fire spread confirmed through this research is well known. On the other hand, the relationship found between slope and ignition goes against findings from other studies, even though the effect displays interregional variability.
- **Elevation:** Findings are mostly inconclusive for the association between elevation and large wildfire spread. The effects of this variable are mixed throughout the study area, with some regions displaying a negative relationship between high altitudes and burned area, which goes against previous knowledge on this phenomenon.

Also requiring further investigation and confirmation are a few other results of this work, whose effects are mixed throughout the study area, whose explanation has not been deemed satisfactory in the context of this research or whose analysis suffered from significant methodological constraints:

- **Livestock activities and grazing:** Livestock activities display variable effects throughout the study area in connection with large fire spread, which challenges the identification of specific trends concerning these driving factors. Grazing areas have been found to be positively associated with large fire spread, displaying an effect contrasting with large fire ignition that needs to be explored.
- **Demographic aspects:** Population density and the potentiality index show negative relationships with fire spread, which had already been confirmed by previous studies. Nevertheless, further research on these topics is needed for a detailed account of the effect of demographic variables on large fire ignition and propagation.
- **Urban areas:** Urban areas are connected to an increase in large fire ignition in most regions of the study area, which confirms the results from previous studies. Still, the spatial variability of this effect needs to be considered in detail for a clear understanding of this relationship.
- **Unemployment:** The association between unemployment and large fire ignition was found to be positive, which confirms knowledge from the reference literature. The explanations found in previous studies, however, do not seem to apply to central Portugal and further investigation is required about the relationship between this socioeconomic aspect and arson.
- **Shrublands:** Shrublands are more prone to large fire ignition in the study area, which might be connected to the reduced value and higher flammability of this vegetation species, even though this trend has not been observed in previous research, as far as this work is concerned. The spatial variability of this effect renders the results inconclusive.
- **Drought:** The connection between the number of dry months and fire propagation failed to be confirmed by this study, mostly because of the contrasting effects displayed. This fact might be explained by methodological issues, e.g. possible unsuitability of the chosen variable.

Overall, this study confirms the presence of spatial variability in the contribution exerted by most structural factors driving large wildfire ignition and spread in central Portugal between 2005 and 2015. It is important that upcoming studies in the field of large wildfires, specifically in the Portuguese context, account for this key feature, which shapes fire behaviour considerably.

This research has also shown, through the results of the spatial clustering analysis, that a partition based on ignition driving factors is similar to a partition based on variables influencing fire propagation. Future studies featuring the same methodology as the one proposed by this work should make use of this finding and provide a sole regional clustering partition for both modelling exercises.

Some methodological issues, which might have had an influence in this study's conclusions, have been identified along this research. These are summarised in the following points:

- **Data availability:** There was no data available for some of the main driving factors acknowledged in previous studies. This problem had already been recognised by Oliveira et al. (2017). In fact, other variables may be important to consider in subsequent studies and model performance can be assessed to identify improvements.
- **Data generalisation:** The methodological framework (1x1km cell grid) required the generalisation of many of the driving factors, which might have had a detrimental effect in the overall data analysis. This aspect is closely connected to data availability.
- **Dimensionality:** This research intended to study a large set of variables and this high dimensionality might have had a harmful impact in the detailed assessment of the main structural driving factors. Kalabokidis et al. (2007) had already mentioned the unpracticality in evaluating the influence of all relevant aspects.
- **Fire spread vs. large fire ignition relationship:** Large wildfire ignitions cannot be dissociated from burned area. Therefore, the target variables' spatial patterns and overall behaviour were similar and the results of this research reflect this limitation.
- **Unbalanced target distributions:** i.e. ignition cells vs. non-ignition cells and burned cells vs. non-burned cells). This unbalance in target distributions impacted model performance on both exercises. Methods such as oversampling, undersampling or SMOTE, which generate artificial data to achieve more balanced statistical distributions (Douzas, Bacao, & Last, 2018), were not used but might be considered in future studies to overcome this constraint.

Based on the main conclusions of this work and some of the methodological issues presented above, the following research topics are proposed for future development:

- **Data mining:** Compare the results of this study with those stemming from the use of data mining techniques, especially regarding the development of predictive models. This methodology may be more suitable given the high data volume (both variables and observations).
- **Condensed set of variables:** Select a condensed set of variables for a more detailed assessment of the effect of different elements in large wildfire ignition and spread, namely those factors whose results seem to contradict previous knowledge in this field.
- **Ignition size:** In the case of ignition models, confirm if results differ substantially when the analysis is performed following the same methodology but focusing on occurrences of any size.
- **Causes of large wildfires:** In the case of ignition models, introduce ignition cause as a target variable and contribute to the process of identifying the unknown causes of large wildfires.

6. BIBLIOGRAPHY

Ager, A. A., Preisler, H. K., Arca, B., Spano, D., & Salis, M. (2014). Wildfire risk estimation in the Mediterranean area. *Environmetrics*, *25*(6), 384–396. https://doi.org/10.1002/env.2269

Almeida, A. C., Nunes, A., & Figueiredo, A. (2013). Paisagem sob risco em Portugal: um contributo. In L. Lourenço & M. Mateus (Eds.), *Riscos Naturais, Antrópicos e Mistos* (pp. 741–750).

Álvarez-Díaz, M., González-Gómez, M., & Otero-Giraldez, M. S. (2015). Detecting the socioeconomic driving forces of the fire catastrophe in NW Spain. *European Journal of Forest Research*, *134*(6), 1087–1094. https://doi.org/10.1007/s10342-015-0911-1

Asuero, A. G., Sayago, A., & González, A. G. (2006). The correlation coefficient: An overview. *Critical Reviews in Analytical Chemistry*, *36*(1), 41–59. https://doi.org/10.1080/10408340500526766

Bachmann, A., & Allgöwer, B. (2001). A consistent wildland fire risk terminology is needed. *Fire Management Today*, *61*(4), 28–33.

Bagnouls, F., & Gaussen, H. (1957). Les climats biologiques et leur classification. *Annales de Géographie*, *66*(355), 193–220. https://doi.org/https://doi.org/10.3406/geo.1957.18273

Balsa-Barreiro, J., & Hermosilla, T. (2013). Socio-geographic analysis of wildland fires: Causes of the 2006's wildfires in Galicia (Spain). *Forest Systems*, *22*(3), 497–509. https://doi.org/10.5424/fs/2013223-04165

Barros, A. M. G., & Pereira, J. M. C. (2014). Wildfire selectivity for land cover type: Does size matter? *PLoS ONE*, *9*(1). https://doi.org/10.1371/journal.pone.0084760

Benali, A., Ervilha, A. R., Sá, A. C. L., Fernandes, P. M., Pinto, R. M. S., Trigo, R. M., & Pereira, J. M. C. (2016). Deciphering the impact of uncertainty on the accuracy of large wildfire spread simulations. *Science of the Total Environment*, *569–570*, 73–85. https://doi.org/10.1016/j.scitotenv.2016.06.112

Brunsdon, C., Fotheringham, A. S., & Charlton, M. (1998). Geographically Weighted Regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, *47*(3), 431–443. https://doi.org/10.1111/1467-9884.00145

Buntin, M. B., & Zaslavsky, A. M. (2004). Too much ado about two-part models and transformation? Comparing methods of modeling Medicare expenditures. *Journal of Health Economics*, *23*(3), 525–542. https://doi.org/10.1016/j.jhealeco.2003.10.005

Calviño-Cancela, M., Chas-Amil, M. L., García-Martínez, E. D., & Touza, J. (2016). Wildfire risk associated with different vegetation types within and outside wildland-urban interfaces. *Forest Ecology and Management*, *372*, 1–9. https://doi.org/10.1016/j.foreco.2016.04.002

Calviño-Cancela, M., Chas-Amil, M. L., García-Martínez, E. D., & Touza, J. (2017). Interacting effects of topography, vegetation, human activities and wildland-urban interfaces on wildfire ignition risk. *Forest Ecology and Management*, *397*, 10–17. https://doi.org/10.1016/j.foreco.2017.04.033

Cao, X., Cui, X., Yue, M., Chen, J., Tanikawa, H., & Ye, Y. (2013). Evaluation of wildfire propagation susceptibility in grasslands using burned areas and multivariate logistic regression. *International Journal of Remote Sensing*, *34*(19), 6679–6700. https://doi.org/10.1080/01431161.2013.805280

Chuvieco, E., Aguado, I., Jurdao, S., Pettinari, M. L., Yebra, M., Salas, J., … Martínez-Vega, J. (2014).

Integrating geospatial information into fire risk assessment. *International Journal of Wildland Fire*, *23*(5), 606–619. Retrieved from http://hdl.handle.net/1885/18171

Chuvieco, E., Martínez, S., Román, M. V, Hantson, S., & Pettinari, M. L. (2014). Integration of ecological and socio-economic factors to assess global vulnerability to wildfire. *Global Ecology and Biogeography*, *23*(2), 245–258. https://doi.org/10.1111/geb.12095

Dimitrakopoulos, A., Gogi, C., Stamatelos, G., & Mitsopoulos, I. (2011). Statistical analysis of the fire environment of large forest fires (1000 ha) in Greece. *Polish Journal of Environmental Studies*, *20*(2), 327–332.

Dondo Bühler, M., de Torres Curth, M., & Garibaldi, L. A. (2013). Demography and socioeconomic vulnerability influence fire occurrence in Bariloche (Argentina). *Landscape and Urban Planning*, *110*(1), 64–73. https://doi.org/10.1016/j.landurbplan.2012.10.006

Douzas, G., Bacao, F., & Last, F. (2018). Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE. *Information Sciences*, *465*, 1–20. https://doi.org/10.1016/j.ins.2018.06.056

Ender, P. B. (2002). histbox: Stata module to generate histogram with boxplot. UCLA: Statistical Consulting Group.

Farewell, V. T., Long, D. L., Tom, B. D. M., Yiu, S., & Su, L. (2017). Two-Part and Related Regression Models for Longitudinal Data. *Annual Review of Statistics and Its Application*, *4*, 283–315. https://doi.org/10.1146/annurev-statistics-060116-054131.Two-Part

Fernandes, P. M., Davies, G. M., Ascoli, D., Fernández, C., Moreira, F., Rigolot, E., … Molina, D. (2013). Prescribed burning in southern Europe: Developing fire management in a dynamic landscape. *Frontiers in Ecology and the Environment*, *11*(SUPPL. 1), 4–14. https://doi.org/10.1890/120298

Fernandes, P. M., Gonçalves, H., Loureiro, C., Fernandes, M., Costa, T., Cruz, M. G., & Botelho, H. (2009). Modelos de Combustível Florestal para Portugal. *Actas Do 6º Congresso Florestal Nacional. SPCF, Lisboa.*, *17*(July 2015), 348–354.

Fernandes, P. M., Luz, A., & Loureiro, C. (2010). Changes in wildfire severity from maritime pine woodland to contiguous forest types in the mountains of northwestern Portugal. *Forest Ecology and Management*, *260*(5), 883–892. https://doi.org/10.1016/j.foreco.2010.06.008

Fernandes, P. M., Monteiro-Henriques, T., Guiomar, N., Loureiro, C., & Barros, A. M. G. (2016). Bottom-Up Variables Govern Large-Fire Size in Portugal. *Ecosystems*, *19*(8), 1362–1375. https://doi.org/10.1007/s10021-016-0010-2

Fernandes, P. M., Pacheco, A. P., Almeida, R., & Claro, J. (2016). The role of fire-suppression force in limiting the spread of extremely large forest fires in Portugal. *European Journal of Forest Research*, *135*(2), 253–262. https://doi.org/10.1007/s10342-015-0933-8

Ferreira-Leite, F., Bento-Gonçalves, A., Lourenço, L., Úbeda, X., & Vieira, A. (2013). Grandes Incêndios Florestais em Portugal Continental como resultado das perturbações nos regimes de fogo no mundo Mediterrâneo. *Silva Lusitana*, *21*(Especial), 129–144. Retrieved from http://www.scielo.gpeari.mctes.pt/pdf/slu/v21nEspecial/v21a09.pdf%5Cnhttp://hdl.handle.net/1822/25046

Ferreira-Leite, F., Lourenço, L., & Bento-Gonçalves, A. (2013). Large forest fires in mainland Portugal, brief characterization. *Méditerranée*, (121), 53–65. https://doi.org/10.4000/mediterranee.6863

Fuentes-Santos, I., Marey-Pérez, M. F., & González-Manteiga, W. (2013). Forest fire spatial pattern analysis in Galicia (NW Spain). *Journal of Environmental Management*, *128*, 30–42. https://doi.org/10.1016/j.jenvman.2013.04.020

Ganteaume, A., Camia, A., Jappiot, M., San-Miguel-Ayanz, J., Long-Fournel, M., & Lampin, C. (2013). A review of the main driving factors of forest fire ignition over Europe. *Environmental Management*, *51*(3), 651–662. https://doi.org/10.1007/s00267-012-9961-z

Ganteaume, A., & Guerra, F. (2018). Explaining the spatio-seasonal variation of fires by their causes: The case of southeastern France. *Applied Geography*, *90*(March 2017), 69–81. https://doi.org/10.1016/j.apgeog.2017.11.012

Ganteaume, A., & Jappiot, M. (2013). What causes large fires in Southern France. *Forest Ecology and Management*, *294*, 76–85. https://doi.org/10.1016/j.foreco.2012.06.055

González-Olabarria, J. R., Mola-Yudego, B., & Coll, L. (2015). Different Factors for Different Causes: Analysis of the Spatial Aggregations of Fire Ignitions in Catalonia (Spain). *Risk Analysis*, *35*(7), 1197–1209. https://doi.org/10.1111/risa.12339

Goudie, A. (2006). *The human impact on the natural environment - past, present and future* (6th ed.). Oxford: Blackwell Publishing.

Grala, K., Grala, R. K., Hussain, A., Cooke, W. H., & Varner, J. M. (2017). Impact of human factors on wildfire occurrence in Mississippi, United States. *Forest Policy and Economics*, *81*(April), 38–47. https://doi.org/10.1016/j.forpol.2017.04.011

Guo, F., Su, Z., Wang, G., Sun, L., Lin, F., & Liu, A. (2016). Wildfire ignition in the forests of southeast China: Identifying drivers and spatial distribution to predict wildfire likelihood. *Applied Geography*, *66*, 12–21. https://doi.org/10.1016/j.apgeog.2015.11.014

Hernandez, C., Drobinski, P., & Turquety, S. (2015). How much does weather control fire size and intensity in the Mediterranean region? *Annales Geophysicae*, *33*(7), 931–939. https://doi.org/10.5194/angeo-33-931-2015

Holsinger, L., Parks, S. A., & Miller, C. (2016). Weather, fuels, and topography impede wildland fire spread in western US landscapes. *Forest Ecology and Management*, *380*, 59–69. https://doi.org/10.1016/j.foreco.2016.08.035

Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression* (2nd ed.). John Wiley & Sons. https://doi.org/10.2307/2074954

ICNF. (2013). *IFN6 – Áreas dos usos do solo e das espécies florestais de Portugal continental*. *6º Inventário Florestal Nacional*. Lisboa. Retrieved from http://www.icnf.pt/portal/florestas/ifn/resource/ficheiros/ifn/ifn6-res-prelimv1-1

ICNF. (2014). *Análise Das Causas Dos Incêndios Florestais - 2003 - 2013*. *Instituto da Conservação da Natureza e das Florestas*. Retrieved from http://www.icnf.pt/portal/florestas/dfci/relat/relat-causa-incendios-2003-2013

INE. (2007). *Classificação Portuguesa das Actividades Económicas —Revisão 3 (CAE Rev. 3)*. (Instituto Nacional de Estatística, Ed.). Lisboa.

Jiménez-Ruano, A., Rodrigues, M., & de la Riva, J. (2017). Exploring spatialoral dynamics of fire regime features in mainland Spain. *Natural Hazards and Earth System Sciences*, *17*(10), 1697–1711. https://doi.org/10.5194/nhess-17-1697-2017

Kalabokidis, K., Koutsias, N., Konstantinidis, P., & Vasilakos, C. (2007). Multivariate analysis of landscape wildfire dynamics in a Mediterranean ecosystem of Greece. *Area*, *39*(3), 392–402.

Keeley, J. E., & Syphard, A. (2016). Climate Change and Future Fire Regimes: Examples from California. *Geosciences*, *6*(3), 37. https://doi.org/10.3390/geosciences6030037

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, *15*(3), 259–263. https://doi.org/10.1127/0941-2948/2006/0130

Künzli, N., Avol, E., Wu, J., Gauderman, W. J., Rappaport, E., Millstein, J., … Peters, J. M. (2006). Health effects of the 2003 Southern California wildfires on children. *American Journal of Respiratory and Critical Care Medicine*, *174*(11), 1221–1228. https://doi.org/10.1164/rccm.200604-519OC

Li, D. (2018). Geocoding and Reverse Geocoding. In B. Huang, T. J. Cova, & Ming-Hsiang Tsou (Eds.), *Comprehensive Geographic Information Systems - Volume 1: GIS Methods and Techniques* (pp. 95–109). Amsterdam: Elsevier.

Lu, G. Y., & Wong, D. W. (2008). An adaptive inverse-distance weighting spatial interpolation technique. *Computers & Geosciences*, *34*(9), 1044–1055. https://doi.org/10.1016/J.CAGEO.2007.07.010

Martínez-Fernández, J., Chuvieco, E., & Koutsias, N. (2013). Modelling long-term fire occurrence factors in Spain by accounting for local variations with geographically weighted regression. *Natural Hazards and Earth System Science*, *13*(2), 311–327. https://doi.org/10.5194/nhess-13-311-2013

Mateus, P., & Fernandes, P. M. (2014). Forest Fires in Portugal: Dynamics, Causes and Policies. In F. Reboredo (Ed.), *Forest Context and Policies in Portugal: Present and Future Challenges* (Vol. 19, pp. 117–154). Springer. https://doi.org/10.1007/978-3-319-08455-8

Mhawej, M., Faour, G., & Adjizian-Gerard, J. (2015). Wildfire Likelihood's Elements: A Literature Review. *Challenges*, *6*(2), 282–293. https://doi.org/10.3390/challe6020282

Miller, C., & Ager, A. A. (2013). A review of recent advances in risk analysis for wildfire management. *International Journal of Wildland Fire*, *22*(1), 1–14. https://doi.org/10.1071/WF11114

Mitsopoulos, I., Mallinis, G., & Arianoutsou, M. (2014). Wildfire Risk Assessment in a Typical Mediterranean Wildland–Urban Interface of Greece. *Environmental Management*, *55*(4), 900–915. https://doi.org/10.1007/s00267-014-0432-6

Modugno, S., Balzter, H., Cole, B., & Borrelli, P. (2016). Mapping regional patterns of large forest fires in Wildland-Urban Interface areas in Europe. *Journal of Environmental Management*, *172*. https://doi.org/10.1016/j.jenvman.2016.02.013

Moreira, F., Catry, F. X., Rego, F., & Bação, F. (2010). Size-dependent pattern of wildfire ignitions in Portugal: When do ignitions turn into big fires? *Landscape Ecology*, *25*(9), 1405–1417. https://doi.org/10.1007/s10980-010-9491-0

Moreira, F., Viedma, O., Arianoutsou, M., Curt, T., Koutsias, N., Rigolot, E., … Bilgili, E. (2011). Landscape-wildfire interactions in southern Europe: Implications for landscape management. *Journal of Environmental Management*, *92*(10), 2389–2402. https://doi.org/10.1016/j.jenvman.2011.06.028

Moreno, J. M., Vallejo, V. R., & Chuvieco, E. (2013). Current Fire Regimes, Impacts and the Likely Changes - VI: Euro Mediterranean. In J. G. Goldammer (Ed.), *Vegetation Fires and Global Change* (Global Fir, pp. 115–131). Remagen-Oberwinter: Kessel Publishing House. Retrieved from http://oro.open.ac.uk/39953/

Nunes, A., Lourenço, L., Bento-Gonçalves, A., & Vieira, A. (2013). Três décadas de incêndios florestais em Portugal: incidência regional e principais fatores responsáveis. *Cadernos de Geografia*, (32), 133–143.

Nunes, A., Lourenço, L., & Castro Meira, A. C. (2016). Exploring spatial patterns and drivers of forest fires in Portugal (1980–2014). *Science of the Total Environment*, *573*, 1190–1202. https://doi.org/10.1016/j.scitotenv.2016.03.121

Oliveira, S. L. J., Pereira, J. M. C., & Carreiras, J. M. B. (2012). Fire frequency analysis in Portugal (1975-2005), using Landsat-based burnt area maps. *International Journal of Wildland Fire*, *21*(1), 48–60. https://doi.org/10.1071/WF10131

Oliveira, S., Pereira, J. M. C., San-Miguel-Ayanz, J., & Lourenço, L. (2014). Exploring the spatial patterns of fire density in Southern Europe using Geographically Weighted Regression. *Applied Geography*, *51*, 143–157. https://doi.org/10.1016/j.apgeog.2014.04.002

Oliveira, S., Zêzere, J. L., Queirós, M., & Pereira, J. M. C. (2017). Assessing the social context of wildfire-affected areas. The case of mainland Portugal. *Applied Geography*, *88*, 104–117. https://doi.org/10.1016/j.apgeog.2017.09.004

Oliveira, T. M., Guiomar, N., Baptista, F. O., Pereira, J. M. C., & Claro, J. (2017). Is Portugal's forest transition going up in smoke? *Land Use Policy*, *66*(May), 214–226. https://doi.org/10.1016/j.landusepol.2017.04.046

Parente, J., Pereira, M. G., & Tonini, M. (2016). Space-time clustering analysis of wildfires: The influence of dataset characteristics, fire prevention policy decisions, weather and climate. *Science of the Total Environment*, *559*, 151–165. https://doi.org/10.1016/j.scitotenv.2016.03.129

Pereira, J. M. C., Carreiras, J. M. B., Silva, J. M. N., & Vasconcelos, M. J. (2006). Alguns Conceitos Básicos sobre os Fogos Rurais em Portugal. In J. S. Pereira, J. M. C. Pereira, F. C. Rego, J. M. N. Silva, & T. P. Silva (Eds.), *Incêndios Florestais em Portugal: Caracterização, Impactes e Prevenção* (pp. 133–161). ISAPress.

Pereira, M. G., Caramelo, L., Orozco, C. V., Costa, R., & Tonini, M. (2015). Space-time clustering analysis performance of an aggregated dataset: The case of wildfires in Portugal. *Environmental Modelling and Software*, *72*, 239–249. https://doi.org/10.1016/j.envsoft.2015.05.016

Phillips, S. J., Anderson, R. P., & Schapire, R. E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, *190*, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Pinho, J. (2014). Forest Planning in Portugal. In F. Reboredo (Ed.), *Forest Context and Policies in Portugal: Present and Future Challenges* (World Fore, pp. 155–183). Springer, Cham. https://doi.org/https://doi.org/10.1007/978-3-319-08455-8_6

Platt, R. V. (2010). The Wildland-Urban Interface : Evaluating the Definition Effect. *Journal of Forestry*, (February), 9–15.

Price, O. F., Penman, T. D., Bradstock, R. A., & Borah, R. (2016). The drivers of wildfire enlargement do not exhibit scale thresholds in southeastern Australian forests. *Journal of Environmental Management*, *181*, 208–217. https://doi.org/10.1016/j.jenvman.2016.06.033

Rego, F., & Silva, J. S. (2014). Wildfires and landscape dynamics in Portugal: a regional assessment and global implications. In J. C. Azevedo, A. H. Perera, & M. A. Pinto (Eds.), *Forest Landscapes and Global Change: Challenges for Research and Management* (pp. 227–248). Springer. https://doi.org/10.1007/978-1-4939-0953-7

Ribeiro, C., Valente, S., Coelho, C., & Figueiredo, E. (2015). A look at forest fires in Portugal: technical, institutional, and social perceptions. *Scandinavian Journal of Forest Research*, *30*(4), 317–325. https://doi.org/10.1080/02827581.2014.987160

Ribeiro, O. (2011). *Portugal, o Mediterrâneo e o Atlântico. Estudo Geográfico* (8ª Edição).

Ricotta, C., & Di Vito, S. (2014). Modeling the landscape drivers of fire recurrence in Sardinia (Italy). *Environmental Management*, *53*(6), 1077–1084. https://doi.org/10.1007/s00267-014-0269-z

Riley, K. L., Abatzoglou, J. T., Grenfell, I. C., Klene, A. E., & Heinsch, F. A. (2013). The relationship of large fire occurrence with drought and fire danger indices in the western USA, 1984-2008: the role of temporal scale. *International Journal of Wildland Fire*, *22*(7), 894–909. https://doi.org/10.1071/WF12149

Rodrigues, M. (2015). *Review and new methodological approaches in human-caused wildfire modeling and ecological vulnerability: risk modeling at mainland Spain*. Universidad de Zaragoza. Retrieved from https://dialnet.unirioja.es/servlet/dctes?codigo=99499

Rodrigues, M., & De la Riva, J. (2014). An insight into machine-learning algorithms to model human-caused wildfire occurrence. *Environmental Modelling and Software*, *57*, 192–201. https://doi.org/10.1016/j.envsoft.2014.03.003

Rodrigues, M., de la Riva, J., & Fotheringham, A. S. (2014). Modeling the spatial variation of the explanatory factors of human-caused wildfires in Spain using geographically weighted logistic regression. *Applied Geography*, *48*, 52–63. https://doi.org/10.1016/j.apgeog.2014.01.011

Rodrigues, M., Jiménez, A., & de la Riva, J. (2016). Analysis of recent spatial–temporal evolution of human driving factors of wildfires in Spain. *Natural Hazards*, *84*(3), 2049–2070. https://doi.org/10.1007/s11069-016-2533-4

Romero-Calcerrada, R., Barrio-Parra, F., Millington, J. D. A., & Novillo, C. J. (2010). Spatial modelling of socioeconomic data to understand patterns of human-caused wildfire ignition risk in the SW of Madrid (central Spain). *Ecological Modelling*, *221*(1), 34–45. https://doi.org/10.1016/j.ecolmodel.2009.08.008

Romero-Calcerrada, R., Novillo, C. J., Millington, J. D. A., & Gomez-Jimenez, I. (2008). GIS analysis of spatial patterns of human-caused wildfire ignition risk in the SW of Madrid (Central Spain). *Landscape Ecology*, *23*(3), 341–354. https://doi.org/10.1007/s10980-008-9190-2

Ruß, G., & Kruse, R. (2011). Machine learning methods for spatial clustering on precision agriculture data. *Frontiers in Artificial Intelligence and Applications*, *227*, 40–49. https://doi.org/10.3233/978-1-60750-754-3-40

Salis, M., Ager, A. A., Alcasena, F. J., Arca, B., Finney, M. A., Pellizzaro, G., & Spano, D. (2015). Analyzing seasonal patterns of wildfire exposure factors in Sardinia, Italy. *Environmental Monitoring and Assessment*, *187*(1). https://doi.org/10.1007/s10661-014-4175-x

San-Miguel-Ayanz, J., Moreno, J. M., & Camia, A. (2013). Analysis of large fires in European Mediterranean landscapes: Lessons learned and perspectives. *Forest Ecology and Management*, *294*, 11–22. https://doi.org/10.1016/j.foreco.2012.10.050

Sarris, D., Christopoulou, A., Angelonidi, E., Koutsias, N., Fulé, P. Z., & Arianoutsou, M. (2014). Increasing extremes of heat and drought associated with recent severe wildfires in southern Greece. *Regional Environmental Change*, *14*(3), 1257–1268. https://doi.org/10.1007/s10113-013-0568-6

SAS Institute Inc. (2009). *Getting Started with SAS Enterprise Miner 6.1*. Cary, NC.

Seol, A., Lee, B., & Chung, J. (2012). Analysis of the seasonal characteristics of forest fires in South Korea using the multivariate analysis approach. *Journal of Forest Research*, *17*(1), 45–50. https://doi.org/10.1007/s10310-011-0263-8

Sharma, S. (1996). *Applied multivariate techniques*. New York: John Wiley & Sons. https://doi.org/10.2307/1270777

Sharples, J. J., & Matthews, S. (2011). Evaluation of some simple models for predicting fuel moisture content. In *19th International Congress on Modelling and Simulation* (pp. 242–248). Perth, Australia.

Silva, J. S., Rego, F., Fernandes, P., & Rigolot, E. (Eds.). (2010). *Towards Integrated Fire Management - Outcomes of the European Project Fire Paradox*. European Forest Institute.

Srivastava, S. K., Saran, S., de By, R. A., & Dadhwal, V. K. (2014). A geo-information system approach for forest fire likelihood based on causative and anti-causative factors. *International Journal of Geographical Information Science*, *28*(3), 427–454. https://doi.org/10.1080/13658816.2013.797984

Stewart, S. I., Radeloff, V. C., Hammer, R. B., & Hawbaker, T. J. (2007). Defining the Wildland-Urban Interface. *Journal of Forestry*, (June), 201–207.

Sweet, S., & Martin, K. G. (2012). *Data Analysis with SPSS: A First Course in Applied Statistic* (4th ed.). Allyn & Bacon.

Tan, P.-N., Steinbach, M., Karpatne, A., & Kumar, V. (2018). Cluster Analysis: Basic Concepts and Algorithms. In *Introduction to Data Mining* (2nd ed., p. 864). Pearson. https://doi.org/10.1016/0022-4405(81)90007-8

Tedim, F., Leone, V., Amraoui, M., Bouillon, C., Coughlan, M., Delogu, G., … Xanthopoulos, G. (2018). Defining Extreme Wildfire Events: Difficulties, Challenges, and Impacts. *Fire*, *1*(1), 9. https://doi.org/10.3390/fire1010009

Tedim, F., Remelgado, R., Borges, C., Carvalho, S., & Martins, J. (2013). Exploring the occurrence of mega-fires in Portugal. *Forest Ecology and Management*, *294*, 86–96. https://doi.org/10.1016/j.foreco.2012.07.031

Van Wagner, C. E. (1987). *Development and structure of the Canadian forest fire weather index system*. Ottawa. https://doi.org/19927

Vasilakos, C., Kalabokidis, K., Hatzopoulos, J., & Matsinos, I. (2009). Identifying wildland fire ignition factors through sensitivity analysis of a neural network. *Natural Hazards*, *50*(1), 125–143. https://doi.org/10.1007/s11069-008-9326-3

Verde, J. C., & Zêzere, J. L. (2007). *Avaliação da Perigosidade de Incêndio Florestal*. Lisboa.

Verde, J. C., & Zêzere, J. L. (2010). Assessment and validation of wildfire susceptibility and hazard in Portugal. *Natural Hazards and Earth System Science*, *10*(3), 485–497. https://doi.org/10.5194/nhess-10-485-2010

Viana, J. D. A. (2010). *Rede Nacional De Postos De Vigia – Tendências Para O Futuro?* Academia Militar.

Viedma, O., Moity, N., & Moreno, J. M. (2015). Changes in landscape fire-hazard during the second half of the 20th century: Agriculture abandonment and the changing role of driving factors. *Agriculture, Ecosystems and Environment*, *207*, 126–140. https://doi.org/10.1016/j.agee.2015.04.011

Vilar, L., Gómez, I., Martínez-Vega, J., Echavarría, P., Riaño, D., & Martín, M. P. (2016). Multitemporal modelling of socio-economic wildfire drivers in central Spain between the 1980s and the 2000s: Comparing generalized linear models to machine learning algorithms. *PLoS ONE*, *11*(8), 1–18. https://doi.org/10.1371/journal.pone.0161344

Vilar, L., Nieto, H., & Martín, M. P. (2010). Integration of lightning- and human-caused wildfire occurrence models. *Human and Ecological Risk Assessment*, *16*(2), 340–364. https://doi.org/10.1080/10807031003670469

Westerling, A. L., & Bryant, B. P. (2007). Climate change and wildfire in California. *Climatic Change*, *87*(1 SUPPL). https://doi.org/10.1007/s10584-007-9363-z

Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach* (5th ed.). South-Western, Cengage Learning.

Yakubu, I., Mireku-Gyimah, D., & Duker, A. A. (2015). Review of methods for modelling forest fire risk and hazard. *African Journal of Environmental Science and Technology*, *9*(3), 155–165. https://doi.org/10.5897/AJEST2014.1820

ANNEX

ANNEX A

Table A 1 – Identified factors driving large wildfire ignition and spread and corresponding scientific references

FACTORS	CONSIDERED IN THIS RESEARCH	NUNES, LOURENÇO & CASTRO MEIRA (2016)	VILAR ET AL. (2016)	HOLSINGER, PARKS & MILLER (2016)	MHAWEJ, FAOUR & ADJIZIAN-GERARD (2015)	DONDO BÜHLER, TORRES CURTH & GARIBALDI (2013)	GONZÁLEZ-OLABARRIA, MOLA-YUDEGO & COLL (2015)	GRALA ET AL. (2017)	CALVIÑO-CANCELA ET AL. (2017)	GANTEAUME & JAPPIOT (2013)	VASILAKOS ET AL. (2009)	ROMERO-CALCERRADA ET AL. (2008)	HERNANDEZ, DROBINSKI & TURQUETY (2015)	ÁLVAREZ-DÍAZ, GONZÁLEZ-GÓMEZ & OTERO-GIRALDEZ (2015)	CHUVIECO ET AL. (2014)	GUO ET AL. (2016)	ROMERO-CALCERRADA ET AL. (2010)	SARRIS ET AL. (2014)	RODRIGUES, JIMÉNEZ & DE LA RIVA (2016)	NUNES ET AL. (2013)	FERNANDES ET AL. (2016)	DIMITRAKOPOULOS ET AL. (2011)	MOREIRA ET AL. (2010)	BALSA-BARREIRO & HERMOSILLA (2013)	RILEY ET AL. (2013)	RICOTTA & DI VITO (2014)	SRIVASTAVA ET AL. (2014)	RODRIGUES, DE LA RIVA & FOTHERINGHAM (2014)	GANTEAUME ET AL. (2013)	MARTÍNEZ-FERNÁNDEZ, CHUVIECO & KOUTSIAS (2013)	SALIS ET AL. (2015)	OLIVEIRA, PEREIRA, SAN-MIGUEL-AYANZ, & LOURENÇO (2014)	BENALI ET AL. (2016)
CLIMATIC FACTORS																																	
TEMPERATURE	X			X	X					X	X		X			X		X		X		X							X	X		X	X
WIND SPEED	X				X					X	X		X			X						X							X		X		X
WIND DIRECTION					X					X												X									X		X
PRECIPITATION	X				X					X	X					X		X		X				X	X					X		X	
RELATIVE HUMIDITY	X				x						X					X						X		X									X
PRESSURE / GEOPOTENTIAL HEIGHT													X																				
DROUGHT	X				X					X										X					X				X				
EVAPOTRANSPIRATION					X																												
ENERGY RELEASE COMPONENT (ERC)				X																													
INITIAL SPREAD INDEX (ISI)				X																	X												

FACTORS	CONSIDERED IN THIS RESEARCH	NUNES, LOURENÇO & CASTRO MEIRA (2016)	VILAR ET AL. (2016)	HOLSINGER, PARKS & MILLER (2016)	MHAWEJ, FAOUR & ADJIZIAN-GERARD (2015)	DONDO BÜHLER, TORRES CURTH & GARIBALDI (2013)	GONZÁLEZ-OLABARRIA, MOLA-YUDEGO & COLL (2015)	GRALA ET AL. (2017)	CALVIÑO-CANCELA ET AL. (2017)	GANTEAUME & JAPPIOT (2013)	VASILAKOS ET AL. (2009)	ROMERO-CALCERRADA ET AL. (2008)	HERNANDEZ, DROBINSKI & TURQUETY (2015)	ÁLVAREZ-DÍAZ, GONZÁLEZ-GÓMEZ & OTERO-GIRALDEZ (2015)	CHUVIECO ET AL. (2014)	GUO ET AL. (2016)	ROMERO-CALCERRADA ET AL. (2010)	SARRIS ET AL. (2014)	RODRIGUES, JIMÉNEZ & DE LA RIVA (2016)	NUNES ET AL. (2013)	FERNANDES ET AL. (2016)	DIMITRAKOPOULOS ET AL. (2011)	MOREIRA ET AL. (2010)	BALSA-BARREIRO & HERMOSILLA (2013)	RILEY ET AL. (2013)	RICOTTA & DI VITO (2014)	SRIVASTAVA ET AL. (2014)	RODRIGUES, DE LA RIVA & FOTHERINGHAM (2014)	GANTEAUME ET AL. (2013)	MARTÍNEZ-FERNÁNDEZ, CHUVIECO & KOUTSIAS (2013)	SALIS ET AL. (2015)	OLIVEIRA, PEREIRA, SAN-MIGUEL-AYANZ, & LOURENÇO (2014)	BENALI ET AL. (2016)
BUILDUP INDEX (BUI)																					X												
ILLUMINATION TIME / SUNSHINE HOURS					X											X																	
ILLUMINATION INTENSITY					X																												
TOPOGRAPHIC FACTORS																																	
SLOPE	X				X	X			X						X	X						X									X	X	
RUGGEDNESS / TOPOGRAPHIC ROUGHNESS				X						X																				X			
ASPECT	X				X	X			X		X					X						X									X	X	
ELEVATION	X	X		X	X	X	X		X	X	X					X				X		X		X		X			X	X	X	X	
ELEVATION RELIEF FATIO (ERR)				X																													
HEAT LOAD INDEX (HLI)				X																													
TOPOGRAPHIC POSITION INDEX (TPI)				X																													
VEGETATION AND SOIL FACTORS																																	
LAND COVER / TYPE OF VEGETATION	X				X				X	X					X	X				X	X		X			X	X			X	X	X	
FUEL MODELS (FLAMMABILITY)	X										X																				X		X

FACTORS	CONSIDERED IN THIS RESEARCH	NUNES, LOURENÇO & CASTRO MEIRA (2016)	VILAR ET AL. (2016)	HOLSINGER, PARKS & MILLER (2016)	MHAWEJ, FAOUR & ADJIZIAN-GERARD (2015)	DONDO BÜHLER, TORRES CURTH & GARIBALDI (2013)	GONZÁLEZ-OLABARRIA, MOLA-YUDEGO & COLL (2015)	GRALA ET AL. (2017)	CALVIÑO-CANCELA ET AL. (2017)	GANTEAUME & JAPPIOT (2013)	VASILAKOS ET AL. (2009)	ROMERO-CALCERRADA ET AL. (2008)	HERNANDEZ, DROBINSKI & TURQUETY (2015)	ÁLVAREZ-DÍAZ, GONZÁLEZ-GÓMEZ & OTERO-GIRALDEZ (2015)	CHUVIECO ET AL. (2014)	GUO ET AL. (2016)	ROMERO-CALCERRADA ET AL. (2010)	SARRIS ET AL. (2014)	RODRIGUES, JIMÉNEZ & DE LA RIVA (2016)	NUNES ET AL. (2013)	FERNANDES ET AL. (2016)	DIMITRAKOPOULOS ET AL. (2011)	MOREIRA ET AL. (2010)	BALSA-BARREIRO & HERMOSILLA (2013)	RILEY ET AL. (2013)	RICOTTA & DI VITO (2014)	SRIVASTAVA ET AL. (2014)	RODRIGUES, DE LA RIVA & FOTHERINGHAM (2014)	GANTEAUME ET AL. (2013)	MARTÍNEZ-FERNÁNDEZ, CHUVIECO & KOUTSIAS (2013)	SALIS ET AL. (2015)	OLIVEIRA, PEREIRA, SAN-MIGUEL-AYANZ, & LOURENÇO (2014)	BENALI ET AL. (2016)
FUEL DENSITY / FUEL LOADING	X			X	X	X									X							X											X
FUEL MOISTURE				X						X	X														X				X		X		X
SOIL MOISTURE					X																												
SOIL TEXTURE					X																												
SOIL ERODIBILITY															X																		
SOIL ORGANIC MATTER / DUFF COVER					X																	X											
FUEL AGE / AGE OF TREES					X																X												
PAST FIRE ACTIVITY / PREVIOUS BURNS (< 5 YEARS)				X											X																		
TREE DIAMETER AT BREAST HEIGHT / TREE GROWTH					X													X															
BASAL AREA					X																												
HARDWOOD PROPORTION					X																												
DISEASE INDEX					X																												
HUMAN FACTORS																																	
AGRICULTURAL ACTIVITY / WORKFORCE / WILDLAND-AGRICULTURAL INTERFACE	X	X	X		X						X			X					X					X		X		X	X	X			

FACTORS	CONSIDERED IN THIS RESEARCH	NUNES, LOURENÇO & CASTRO MEIRA (2016)	VILAR ET AL. (2016)	HOLSINGER, PARKS & MILLER (2016)	MHAWEJ, FAOUR & ADJIZIAN-GERARD (2015)	DONDO BÜHLER, TORRES CURTH & GARIBALDI (2013)	GONZÁLEZ-OLABARRIA, MOLA-YUDEGO & COLL (2015)	GRALA ET AL. (2017)	CALVIÑO-CANCELA ET AL. (2017)	GANTEAUME & JAPPIOT (2013)	VASILAKOS ET AL. (2009)	ROMERO-CALCERRADA ET AL. (2008)	HERNANDEZ, DROBINSKI & TURQUETY (2015)	ÁLVAREZ-DÍAZ, GONZÁLEZ-GÓMEZ & OTERO-GIRALDEZ (2015)	CHUVIECO ET AL. (2014)	GUO ET AL. (2016)	ROMERO-CALCERRADA ET AL. (2010)	SARRIS ET AL. (2014)	RODRIGUES, JIMÉNEZ & DE LA RIVA (2016)	NUNES ET AL. (2013)	FERNANDES ET AL. (2016)	DIMITRAKOPOULOS ET AL. (2011)	MOREIRA ET AL. (2010)	BALSA-BARREIRO & HERMOSILLA (2013)	RILEY ET AL. (2013)	RICOTTA & DI VITO (2014)	SRIVASTAVA ET AL. (2014)	RODRIGUES, DE LA RIVA & FOTHERINGHAM (2014)	GANTEAUME ET AL. (2013)	MARTÍNEZ-FERNÁNDEZ, CHUVIECO & KOUTSIAS (2013)	SALIS ET AL. (2015)	OLIVEIRA, PEREIRA, SAN-MIGUEL-AYANZ, & LOURENÇO (2014)	BENALI ET AL. (2016)
MECHANISATION OF AGRICULTURE	X						X																					X	X	X			
SIZE AND DENSITY OF FARMS	X																			X													
SIZE AND DENSITY OF AGRICULTURAL PLOTS																				X										X			
GRAZING / LIVESTOCK ACTIVITIES / ANIMAL DENSITY	X	X			X		X					X					X			X				X			X	X				X	
USE OF FIRE FOR AGRICULTURE AND GRAZING ACTIVITIES	X		X					X		X				X																			
FORESTRY ACTIVITY / FOREST MANAGEMENT		X																												X			
NATURAL PROTECTED AREAS / DEGREE OF PROTECTION / PRIORITY OF CONSERVATION	X		X												X				X									X					
HUNTING																								X			X						
TOURISM	X						X			X																							
ROAD NETWORK	X	X	X		X			X		X	X	X				X	X		X				X			X	X		X	X		X	
RAILWAY NETWORK	X		X					X		X						X			X									X	X				
TRACKS	X		X					X				X					X		X									X	X				
WILDLAND-URBAN INTERFACE	X		X						X						X				X									X	X	X		X	
PROXIMITY TO URBAN AREAS / INFRASTRUCTURES	X				X			X			X	X				X	X							X		X	X		X				

FACTORS	CONSIDERED IN THIS RESEARCH	NUNES, LOURENÇO & CASTRO MEIRA (2016)	VILAR ET AL. (2016)	HOLSINGER, PARKS & MILLER (2016)	MHAWEJ, FAOUR & ADJIZIAN-GERARD (2015)	DONDO BÜHLER, TORRES CURTH & GARIBALDI (2013)	GONZÁLEZ-OLABARRIA, MOLA-YUDEGO & COLL (2015)	GRALA ET AL. (2017)	CALVIÑO-CANCELA ET AL. (2017)	GANTEAUME & JAPPIOT (2013)	VASILAKOS ET AL. (2009)	ROMERO-CALCERRADA ET AL. (2008)	HERNANDEZ, DROBINSKI & TURQUETY (2015)	ÁLVAREZ-DÍAZ, GONZÁLEZ-GÓMEZ & OTERO-GIRALDEZ (2015)	CHUVIECO ET AL. (2014)	GUO ET AL. (2016)	ROMERO-CALCERRADA ET AL. (2010)	SARRIS ET AL. (2014)	RODRIGUES, JIMÉNEZ & DE LA RIVA (2016)	NUNES ET AL. (2013)	FERNANDES ET AL. (2016)	DIMITRAKOPOULOS ET AL. (2011)	MOREIRA ET AL. (2010)	BALSA-BARREIRO & HERMOSILLA (2013)	RILEY ET AL. (2013)	RICOTTA & DI VITO (2014)	SRIVASTAVA ET AL. (2014)	RODRIGUES, DE LA RIVA & FOTHERINGHAM (2014)	GANTEAUME ET AL. (2013)	MARTÍNEZ-FERNÁNDEZ, CHUVIECO & KOUTSIAS (2013)	SALIS ET AL. (2015)	OLIVEIRA, PEREIRA, SAN-MIGUEL-AYANZ, & LOURENÇO (2014)	BENALI ET AL. (2016)
PROXIMITY TO INDUSTRIAL AREAS	X											X					X																
PROXIMITY TO RECREATIONAL AREAS / TOURISTIC ZONES	X				X						X	X					X										X		X				
PROXIMITY TO CAMPSITES	X						X					X					X																
PROXIMITY TO ELECTRIC LINES	X										X								X									X	X			X	
PROXIMITY TO LANDFILLS	X										X																						
HOUSING DENSITY	X						X							X															X	X			
SECONDARY HOUSING	X											X					X																
POPULATION DENSITY	X	X	X			X	X	X		X						X	X			X			X	X					X			X	
POPULATION DYNAMICS: VARIATION AND POTENTIAL	X						X												X	X				X				X		X			
AGEING OF POPULATION	X					X														X													
AGE OF RURAL POPULATION																												X	X	X			
RURAL EXODUS		X																						X						X			
LAND USE	X	X					X		X											X													
CHANGES IN LAND COVER	X																			X								X		X			

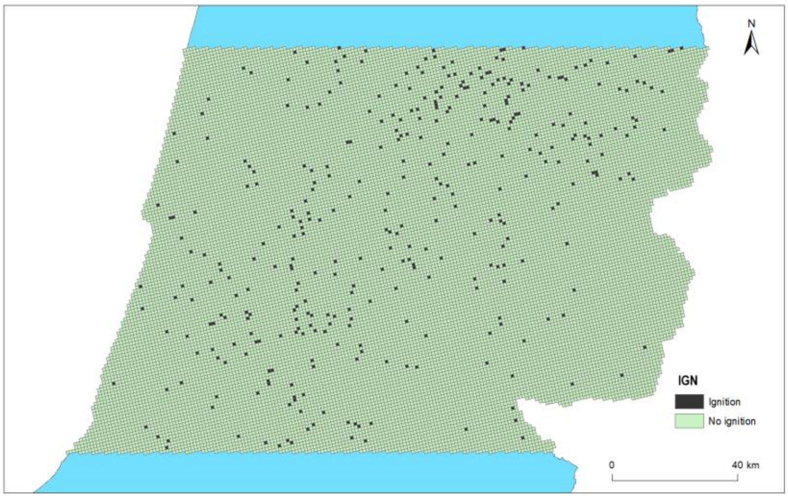
FACTORS	CONSIDERED IN THIS RESEARCH	NUNES, LOURENÇO & CASTRO MEIRA (2016)	VILAR ET AL. (2016)	HOLSINGER, PARKS & MILLER (2016)	MHAWEJ, FAOUR & ADJIZIAN-GERARD (2015)	DONDO BÜHLER, TORRES CURTH & GARIBALDI (2013)	GONZÁLEZ-OLABARRIA, MOLA-YUDEGO & COLL (2015)	GRALA ET AL. (2017)	CALVIÑO-CANCELA ET AL. (2017)	GANTEAUME & JAPPIOT (2013)	VASILAKOS ET AL. (2009)	ROMERO-CALCERRADA ET AL. (2008)	HERNANDEZ, DROBINSKI & TURQUETY (2015)	ÁLVAREZ-DÍAZ, GONZÁLEZ-GÓMEZ & OTERO-GIRALDEZ (2015)	CHUVIECO ET AL. (2014)	GUO ET AL. (2016)	ROMERO-CALCERRADA ET AL. (2010)	SARRIS ET AL. (2014)	RODRIGUES, JIMÉNEZ & DE LA RIVA (2016)	NUNES ET AL. (2013)	FERNANDES ET AL. (2016)	DIMITRAKOPOULOS ET AL. (2011)	MOREIRA ET AL. (2010)	BALSA-BARREIRO & HERMOSILLA (2013)	RILEY ET AL. (2013)	RICOTTA & DI VITO (2014)	SRIVASTAVA ET AL. (2014)	RODRIGUES, DE LA RIVA & FOTHERINGHAM (2014)	GANTEAUME ET AL. (2013)	MARTÍNEZ-FERNÁNDEZ, CHUVIECO & KOUTSIAS (2013)	SALIS ET AL. (2015)	OLIVEIRA, PEREIRA, SAN-MIGUEL-AYANZ, & LOURENÇO (2014)	BENALI ET AL. (2016)
ECONOMIC DIFFICULTIES / POVERTY		X				X		X																									
INCOME	X							X								X																	
UNEMPLOYMENT	X					X	X	X		X																		X	X				
POLITICAL ACTIVITY														X																			
EDUCATIONAL LEVEL	X					X																											
CRIME / DELINQUENCY / POLICE ARRESTS	X						X							X																			
OWNERSHIP (HOUSING, LAND)	X					X		X																X						X			
FIREFIGHTING (INITIAL ATTACK TIME, SUPPRESSION TIME, RESOURCES - FIREFIGHTERS, WATER POWER AND CAPACITY)																					X	X											
FIRE PREVENTION (FIRELINES, FIRE WATCH TOWERS)	X																										X						
IGNITIONS	X																																X

ANNEX B

Summary tables: Target and explanatory variables

Table B 1 – Target variable: IGN

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>



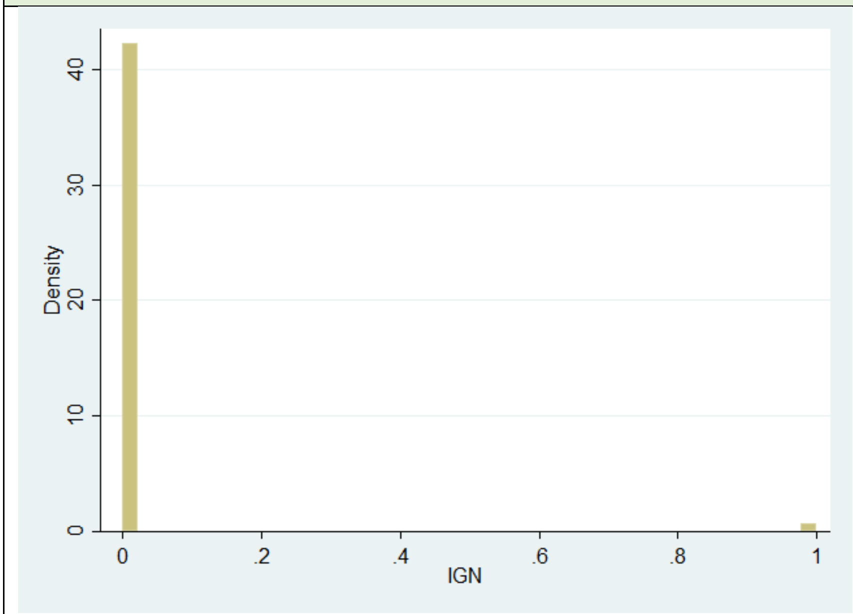
</td><td rowspan="3">

Variable:
Ignition points

Variable name:
IGN

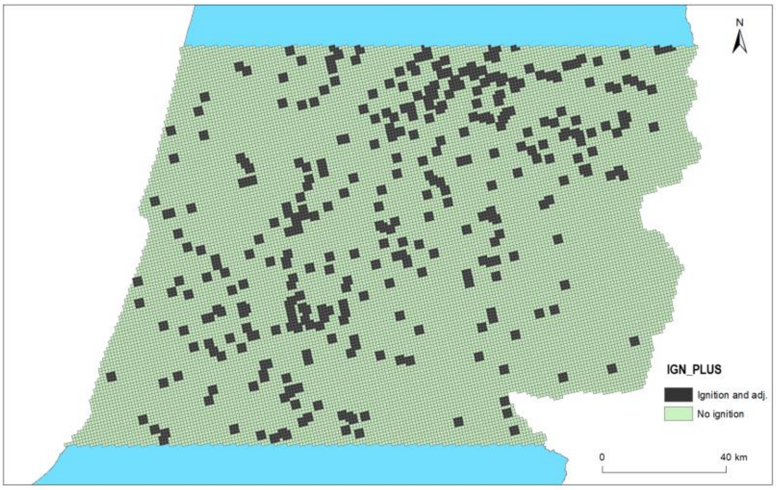
Type of variable:
Binary

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



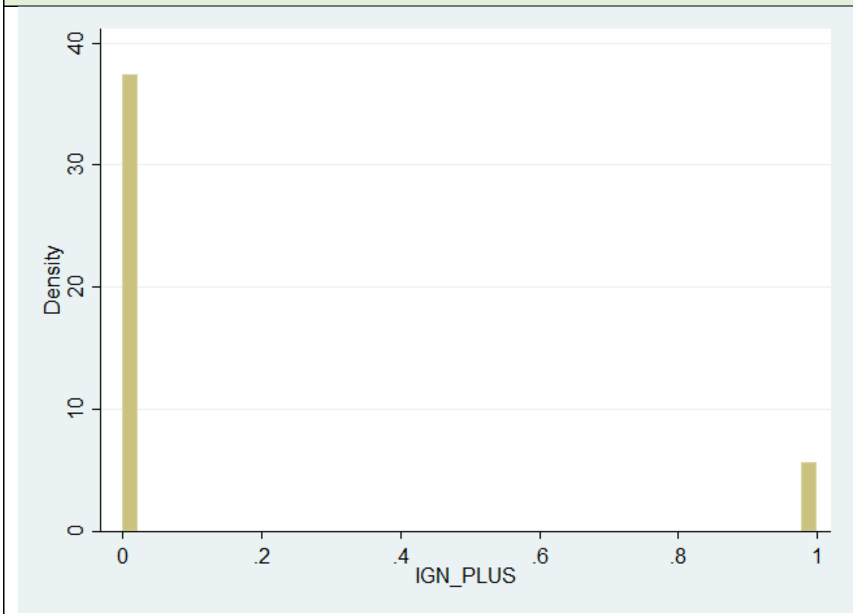
</td></tr>
</table>

Table B 2 – Target variable: IGN_PLUS

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>


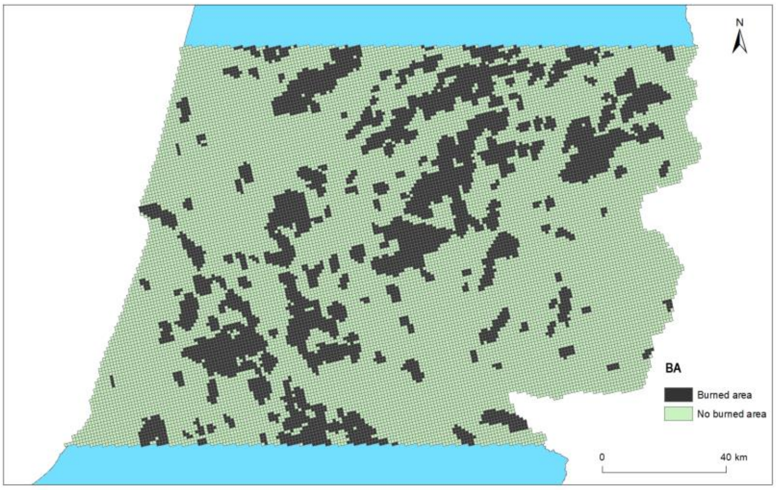
</td><td rowspan="3">
Variable:
Generalised ignition points (ignition points plus ignition adjacent cells)

Variable name:
IGN_PLUS

Type of variable:
Binary
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>


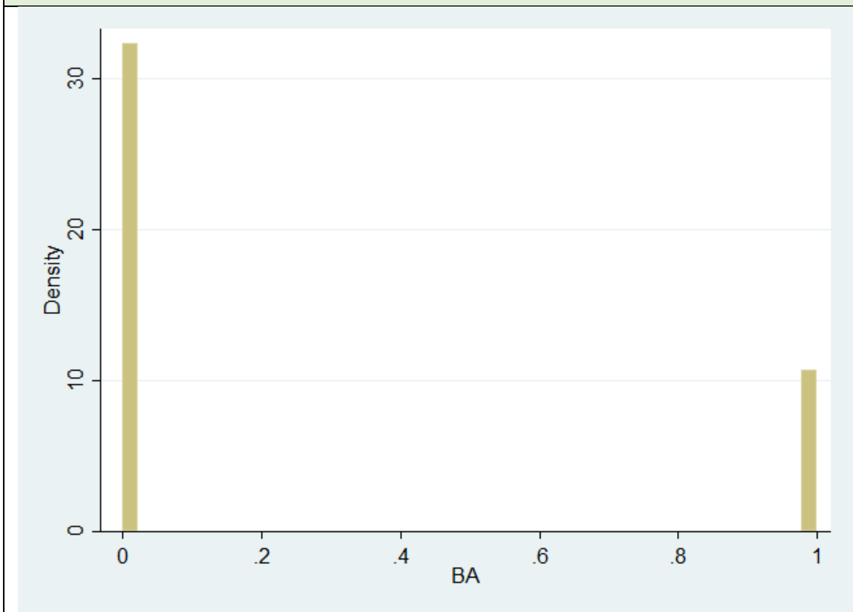
</td></tr>
</table>

Table B 3 – Target variable: BA

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>


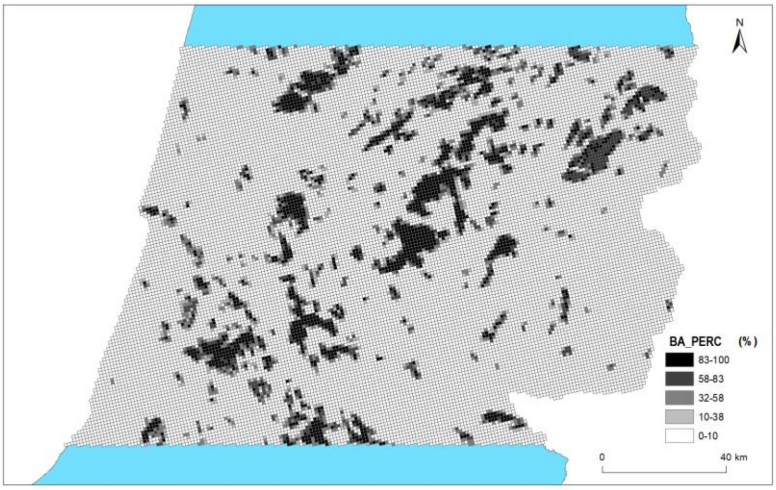
</td><td rowspan="3">Variable:
Burned area

Variable name:
BA

Type of variable:
Binary</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>


</td></tr>
</table>

Table B 4 – Target variable: BA_PERC

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>
</td><td rowspan="3">Variable:
Percentage of burned area in each cell (%)

Variable name:
BA_PERC

Type of variable:
Continuous

Range:
0% – 100%

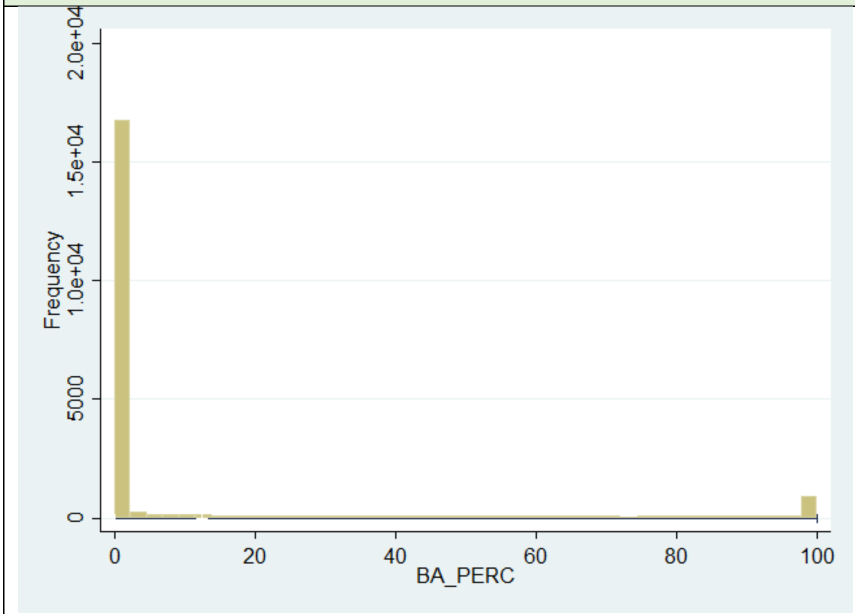
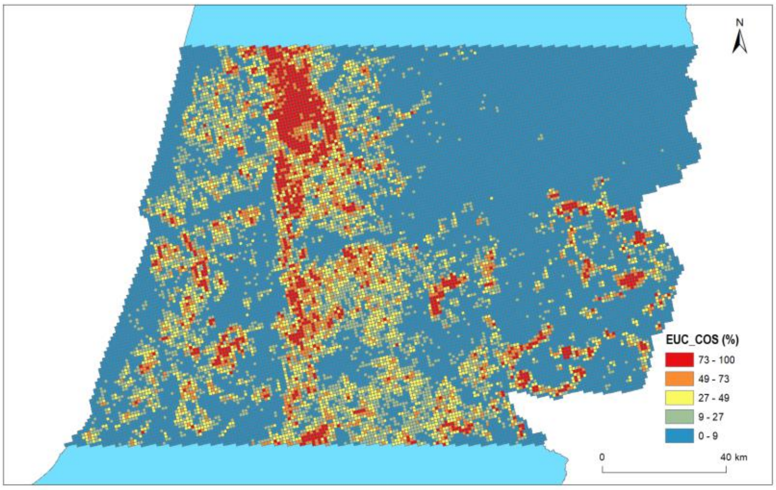
Average (burned area only):
50.1%</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 5 – Explanatory variable: EUC_COS

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Percentage of eucalyptus tree cover in each cell (%)

Variable name:
EUC_COS

Type of variable:
Continuous

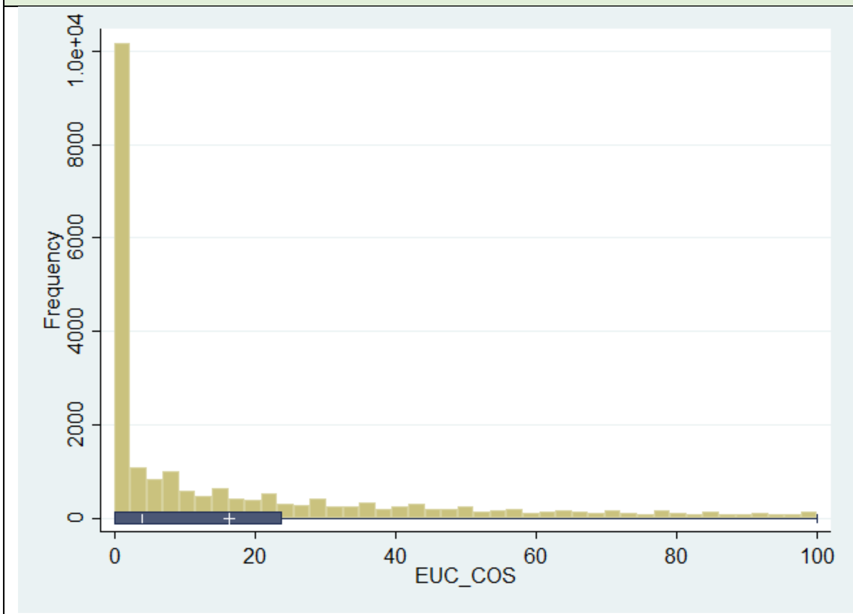
Range:
0% – 100%

Average:
16.4%

Standard deviation:
24.4%

Effect:
Ignition and spread

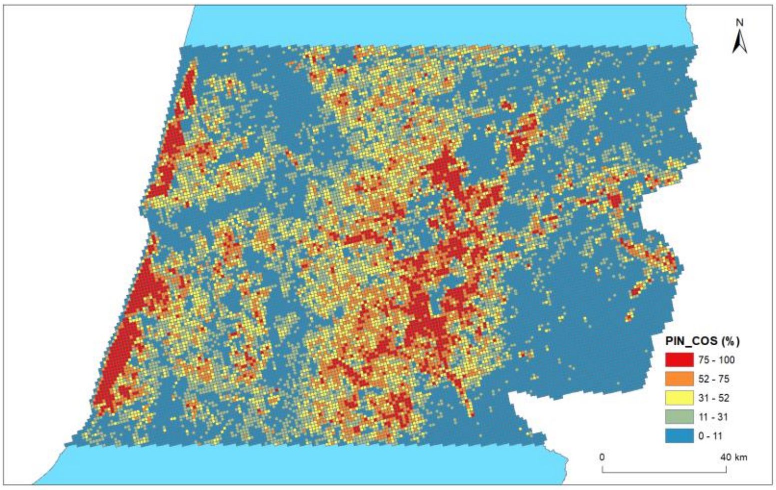
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 6 – Explanatory variable: PIN_COS

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Percentage of pine tree cover in each cell (%)

Variable name:
PIN_COS

Type of variable:
Continuous

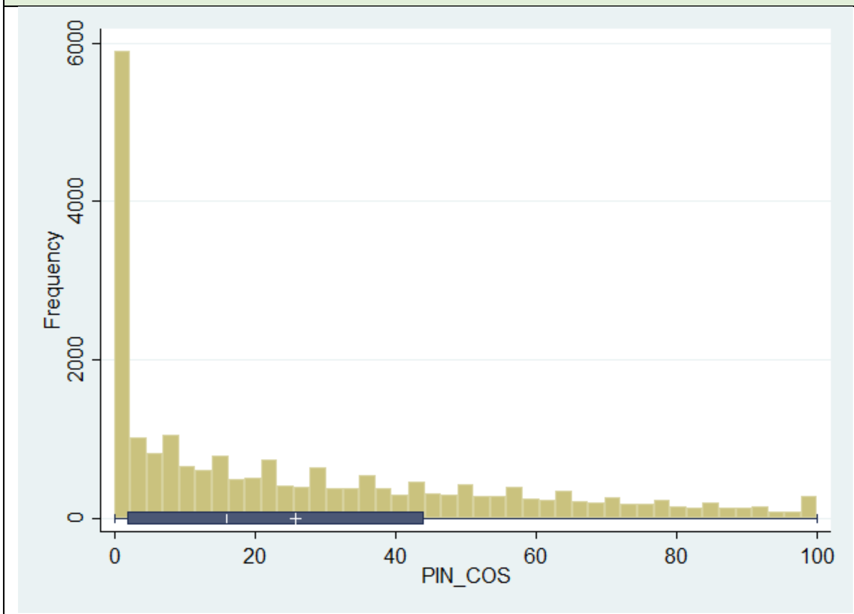
Range:
0% – 100%

Average:
25.9%

Standard deviation:
27.3%

Effect:
Ignition and spread

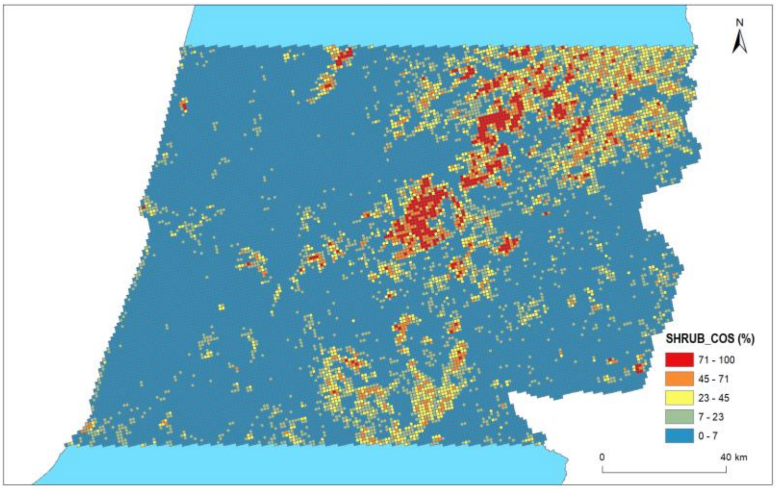
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



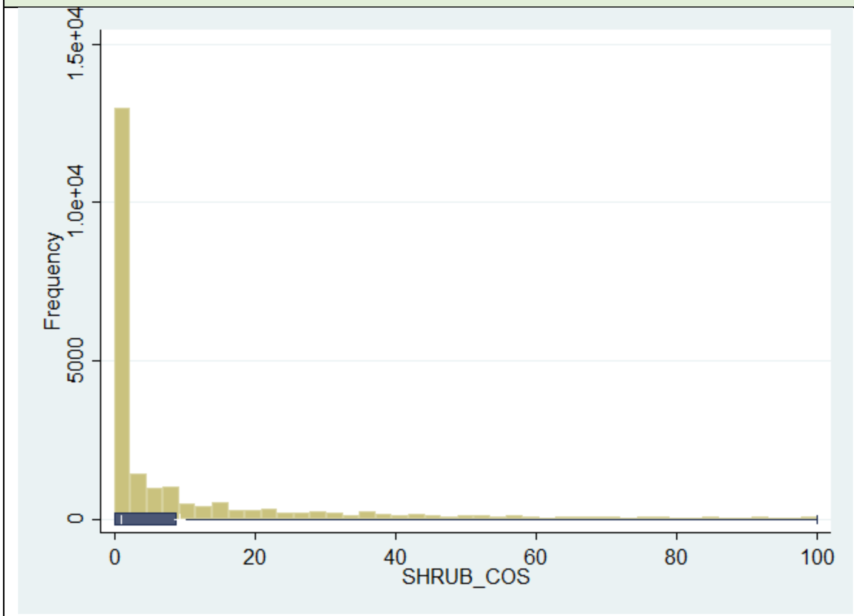
</td></tr>
</table>

Table B 7 – Explanatory variable: SHRUB_COS

Map



Histogram



DESCRIPTION

Variable:
Percentage of shrubland cover in each cell (%)

Variable name:
SHRUB_COS

Type of variable:
Continuous

Range:
0% – 100%

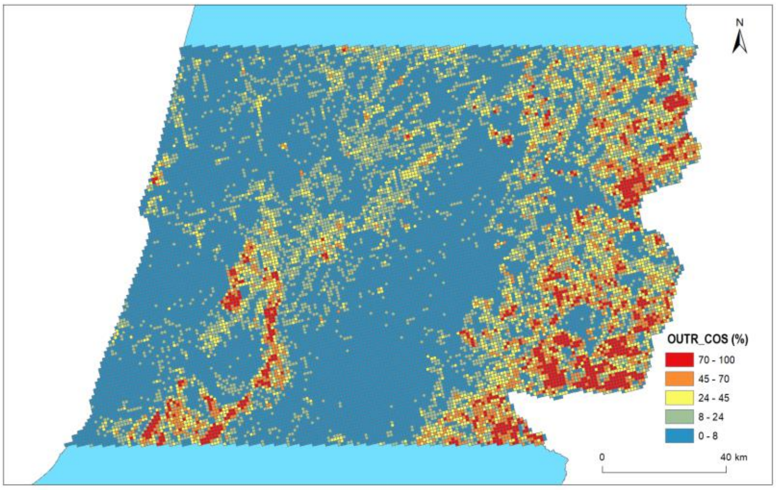
Average:
9.4%

Standard deviation:
18.3%

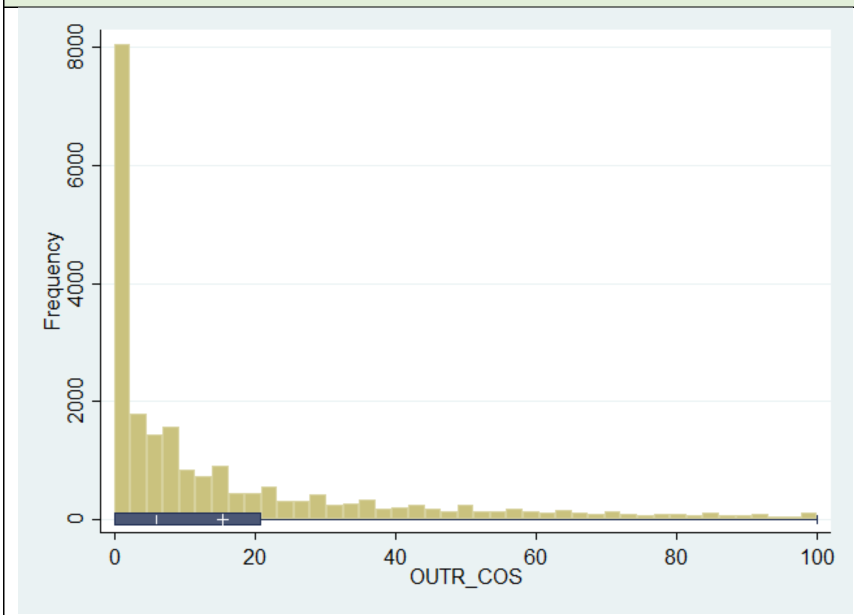
Effect:
Ignition and spread

Table B 8 – Explanatory variable: OUTR_COS

Map	DESCRIPTION



Histogram



Variable:
Percentage of other types of tree cover in each cell (%)

Variable name:
OUTR_COS

Type of variable:
Continuous

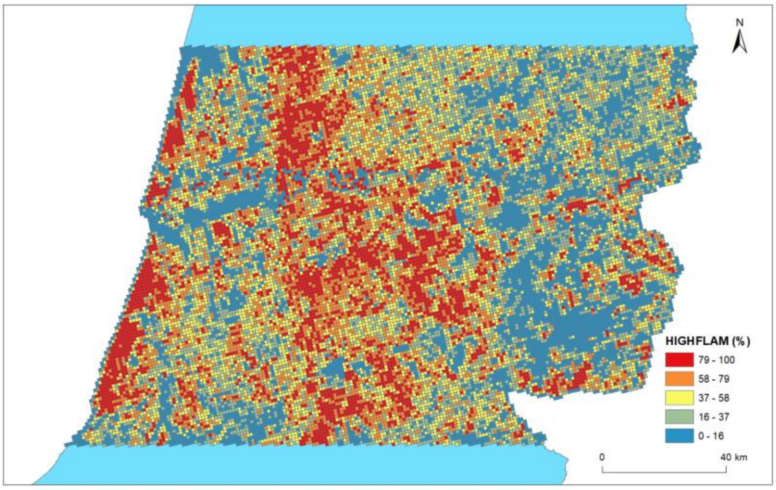
Range:
0% – 100%

Average:
15.4%

Standard deviation:
21.8%

Effect:
Ignition and spread

Table B 9 – Explanatory variable: HIGHFLAM

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>
</td><td rowspan="3">Variable:
Percentage of highly flammable vegetation in each cell (%)

Variable name:
HIGHFLAM

Type of variable:
Continuous

Range:
0% – 100%

Average:
43.9%

Standard deviation:
29.1%

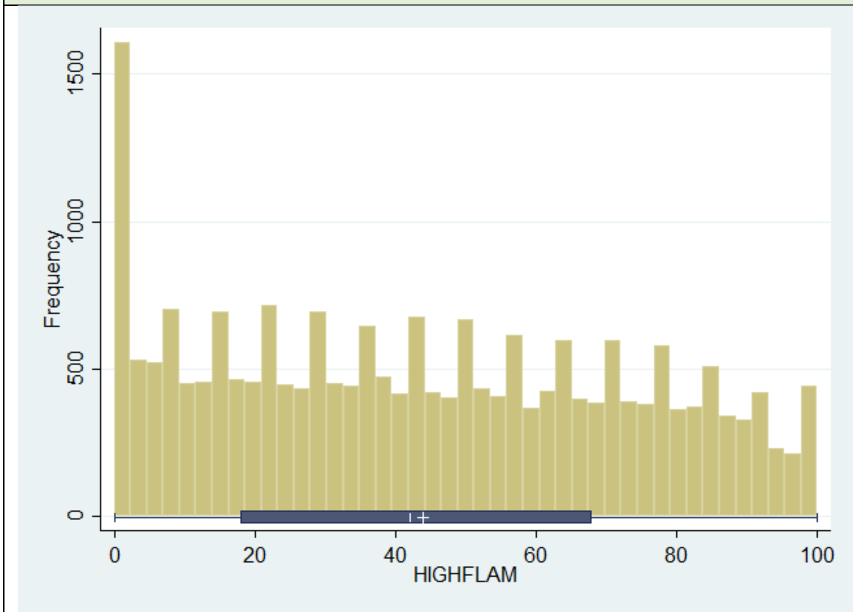
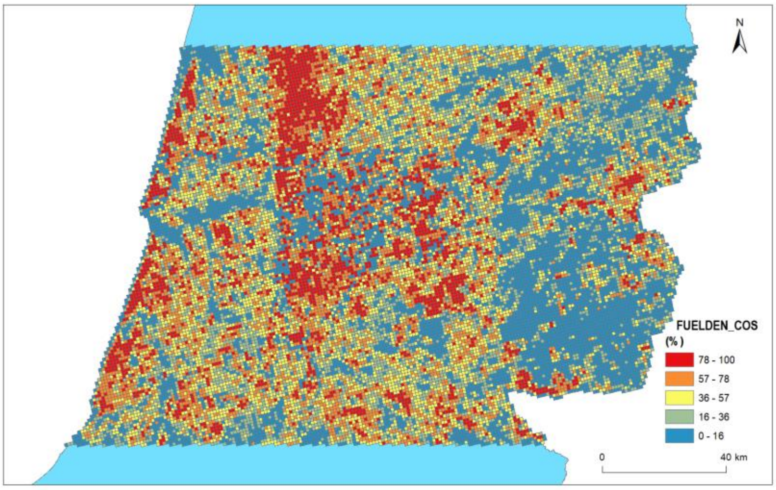
Effect:
Ignition and spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 10 – Explanatory variable: FUELDEN_COS

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>


</td><td rowspan="3">
Variable:
Percentage of dense vegetation cover in each cell (%)

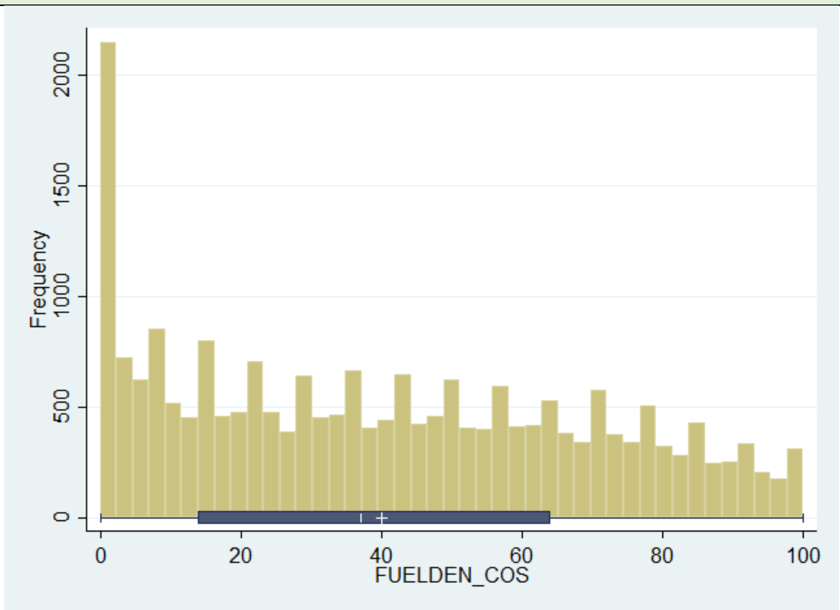
Variable name:
FUELDEN_COS

Type of variable:
Continuous

Range:
0% – 100%

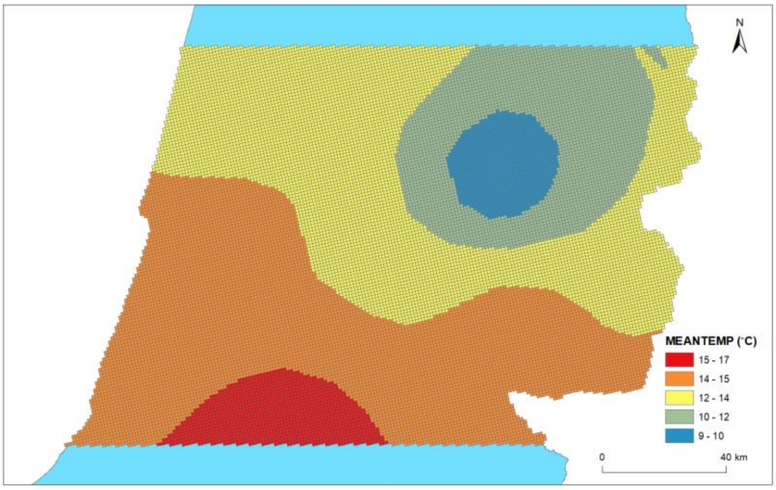
Average:
40.0%

Standard deviation:
29.0%

Effect:
Ignition and spread
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>


</td></tr>
</table>

Table B 11 – Explanatory variable: MEANTEMP

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>


</td><td rowspan="3">
Variable:
Average daily mean temperature (°C)

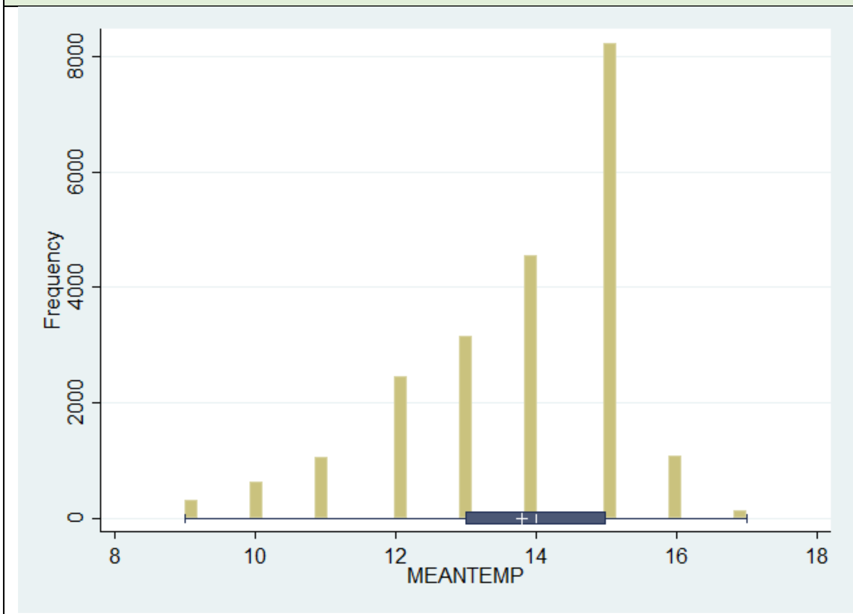
Variable name:
MEANTEMP

Type of variable:
Continuous

Range:
9°C – 17°C

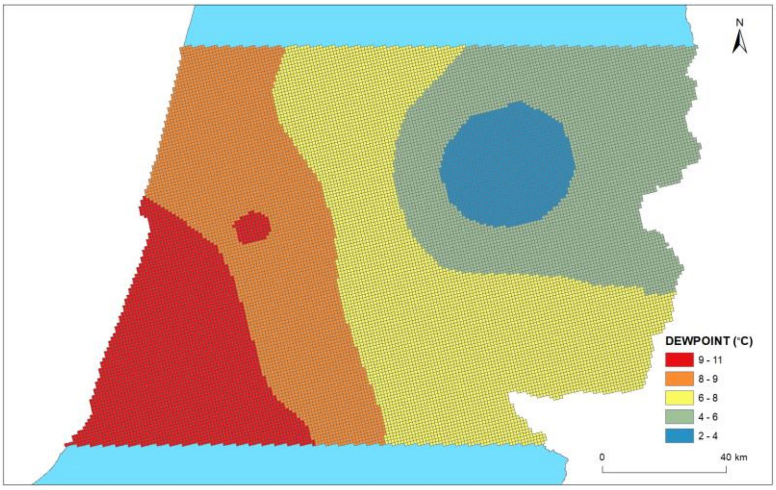
Average:
13.8°C

Standard deviation:
1.6°C

Effect:
Ignition and spread
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>


</td></tr>
</table>

Table B 12 – Explanatory variable: DEWPOINT

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>


</td><td rowspan="3">
Variable:
Average daily mean dewpoint (°C)

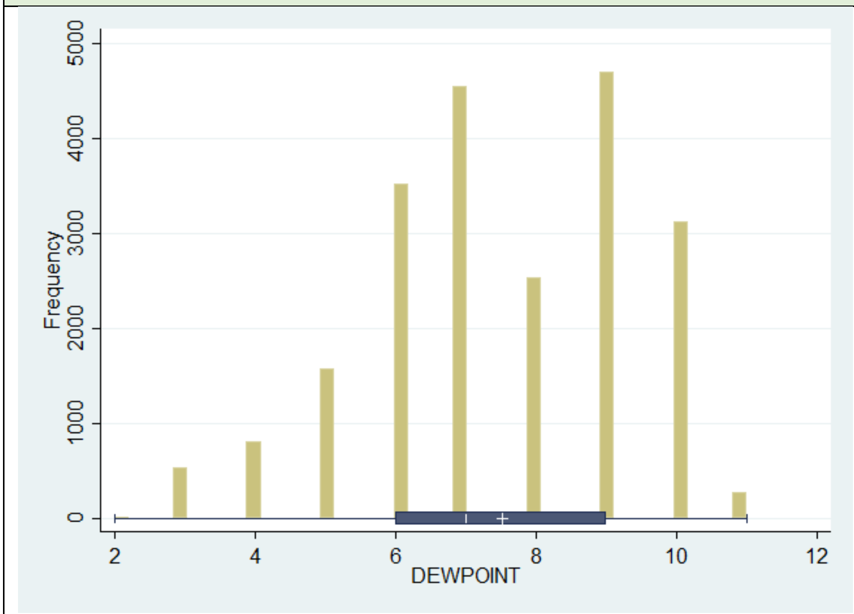
Variable name:
DEWPOINT

Type of variable:
Continuous

Range:
2°C – 11°C

Average:
7.5°C

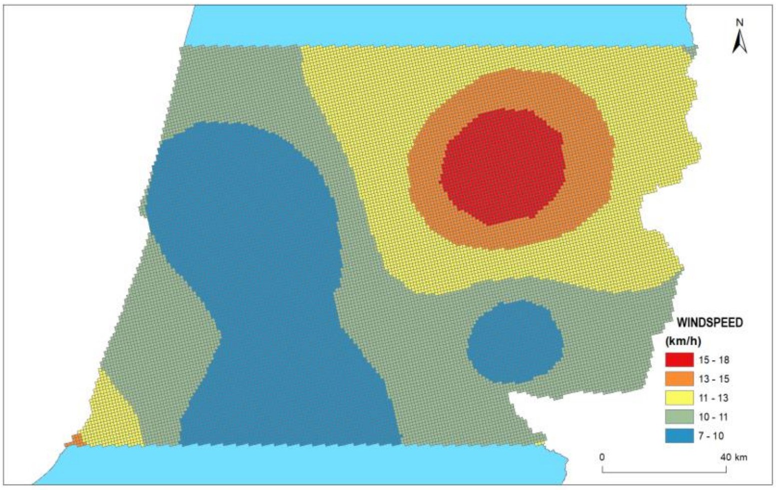
Standard deviation:
1.8°C

Effect:
Ignition and spread
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>


</td></tr>
</table>

Table B 13 – Explanatory variable: WINDSPEED

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Average daily mean wind speed (km/h)

Variable name:
WINDSPEED

Type of variable:
Continuous

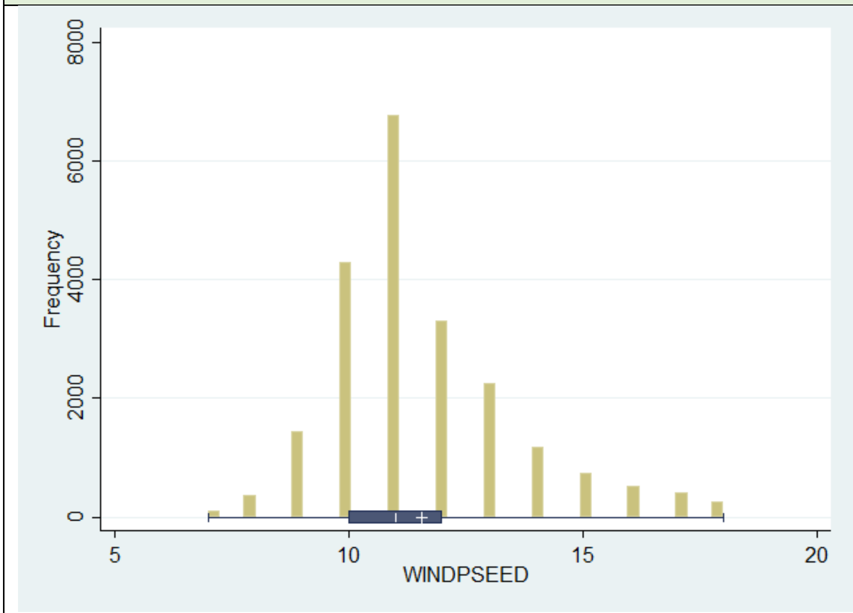
Range:
7 km/h – 18 km/h

Average:
11.6 km/h

Standard deviation:
1.9 km/h

Effect:
Spread

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 14 – Explanatory variable: PRECTOT

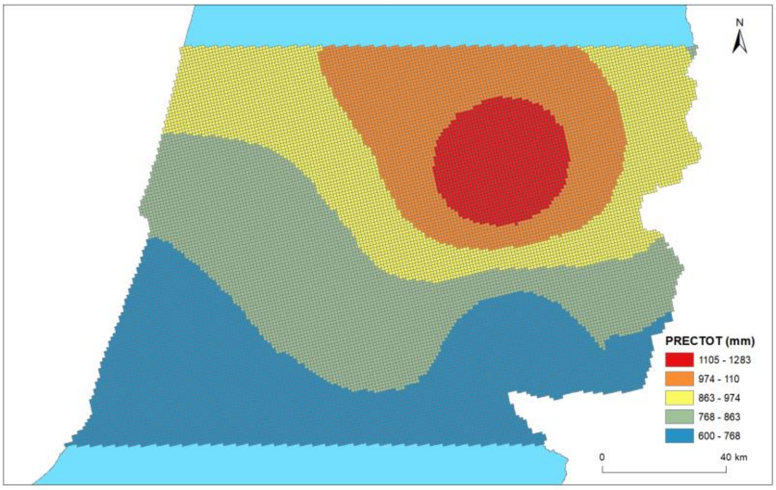
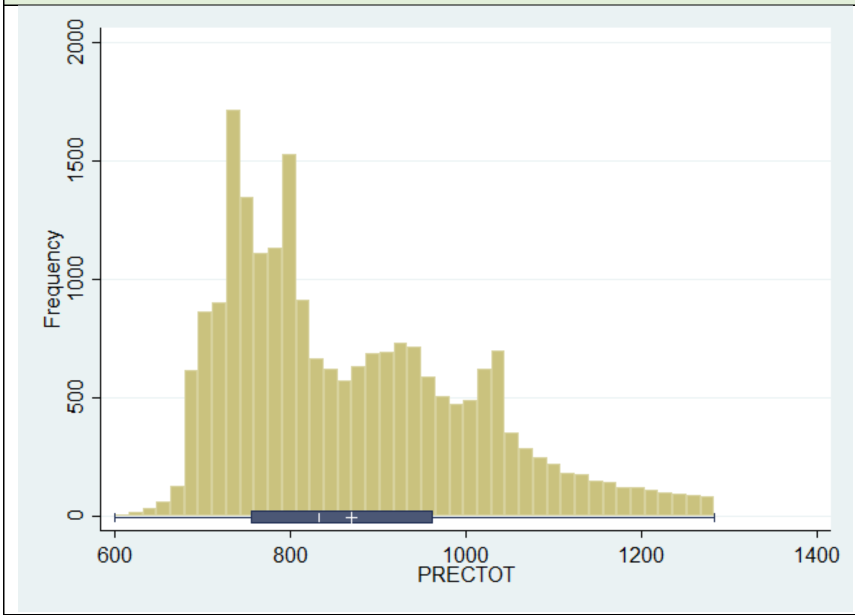
Map	DESCRIPTION
	**Variable:** Average total annual precipitation (mm) **Variable name:** PRECTOT **Type of variable:** Continuous **Range:** 600 mm – 1283 mm **Average:** 869.7 mm **Standard deviation:** 137.9 mm **Effect:** Ignition and spread
Histogram	
	

Table B 15 – Explanatory variable: DRYMONTH

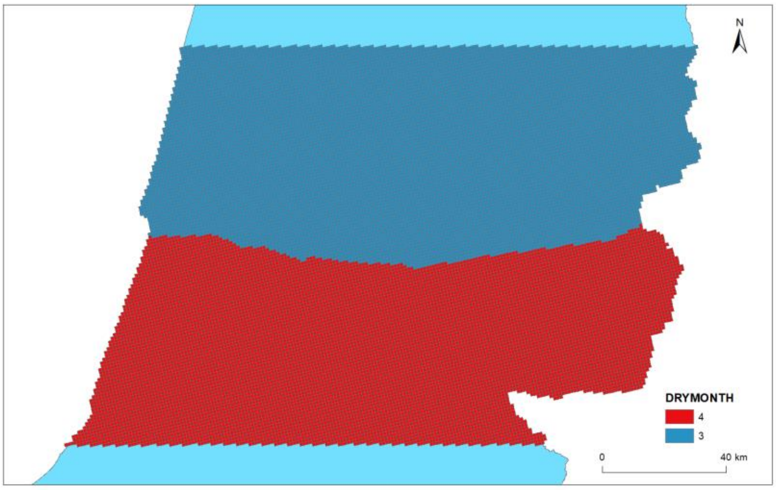
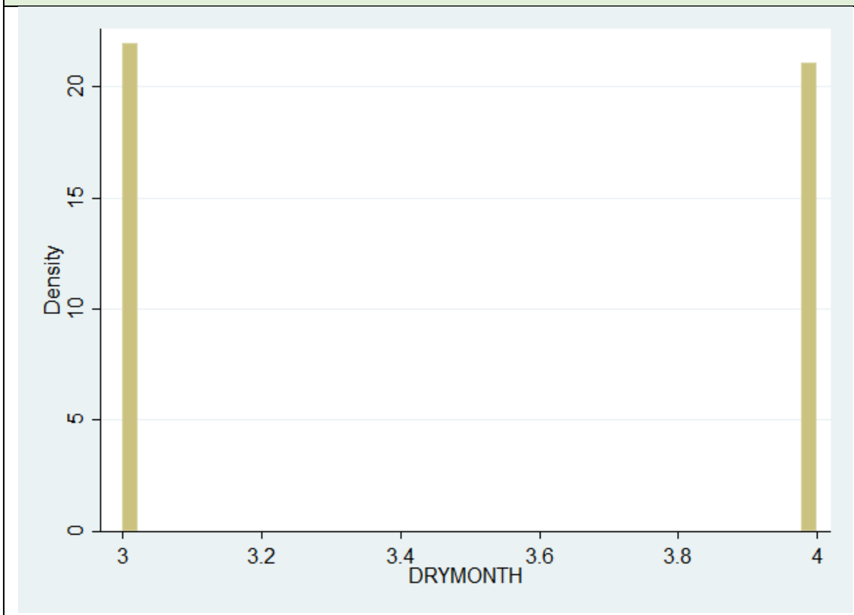
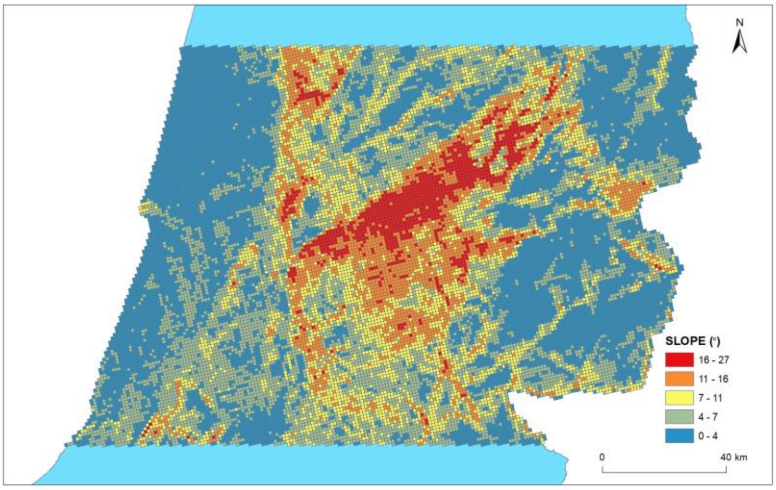
Map	**DESCRIPTION**
	**Variable:** Number of dry months (no.) **Variable name:** DRYMONTH **Type of variable:** Discrete (treated as continuous) **Range:** 3 – 4 **Effect:** Ignition and spread
Histogram	
	

Table B 16 – Explanatory variable: SLOPE

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>
</td>
<td rowspan="3">Variable:
Slope (°)

Variable name:
SLOPE

Type of variable:
Continuous

Range:
0° – 27°

Average:
6.5°

Standard deviation:
4.7°

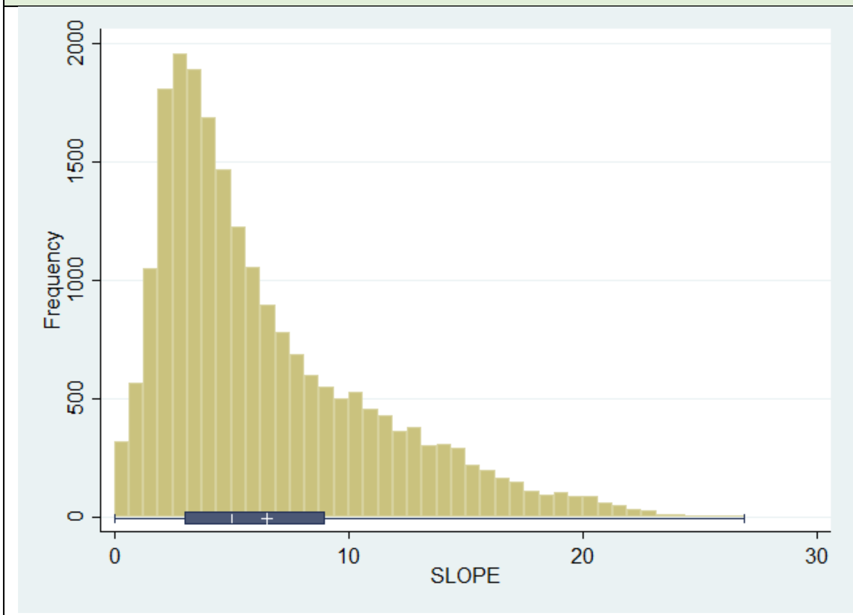
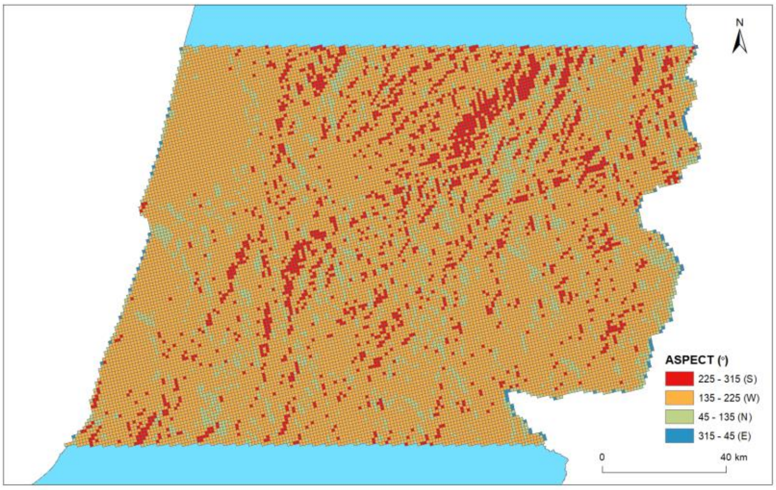
Effect:
Ignition and spread</td></tr>
<tr><th>Histogram</th></tr>
<tr><td>
</td></tr>
</table>

Table B 17 – Explanatory variable: ASPECT

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Aspect (°)

Variable name:
ASPECT

Type of variable:
Continuous

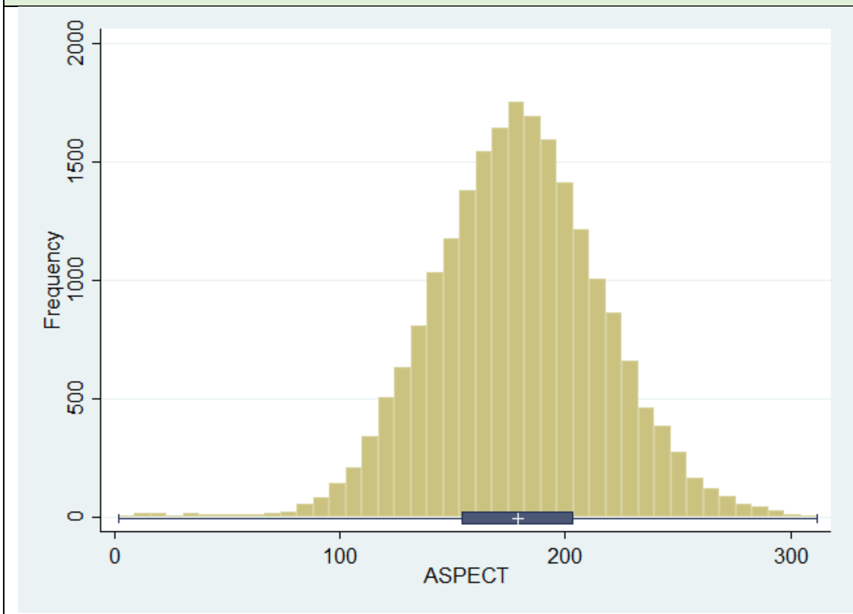
Range:
1° – 312°

Average:
179°

Standard deviation:
38.3°

Effect:
Ignition and spread

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 18 – Explanatory variable: ELEVATION

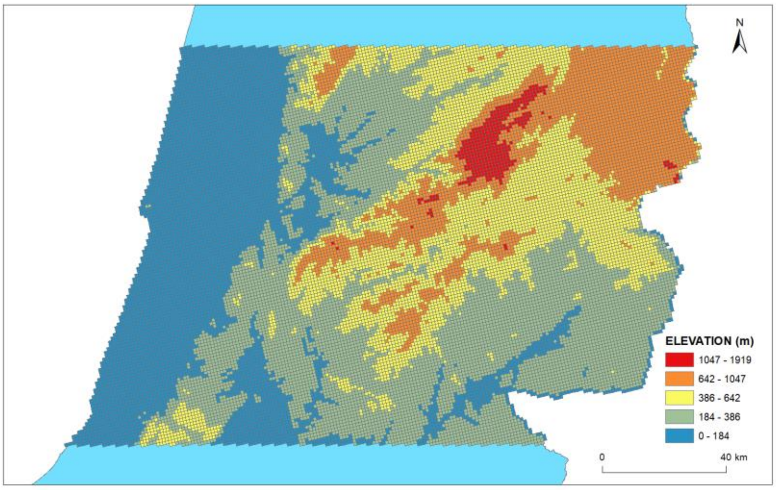
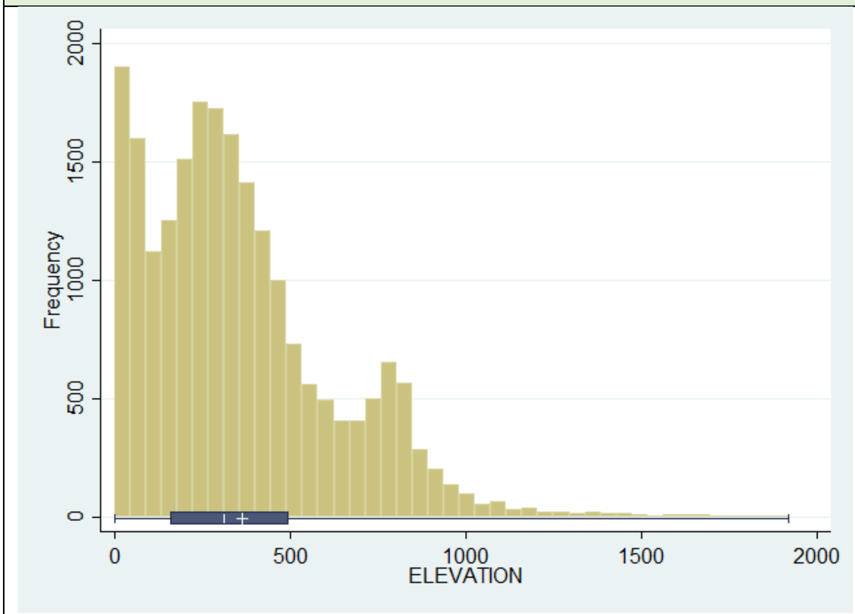
Map	DESCRIPTION
	**Variable:** Elevation (m) **Variable name:** ELEVATION **Type of variable:** Continuous **Range:** 0 m – 1919 m **Average:** 363.9 m **Standard deviation:** 274.3 m **Effect:** Ignition and spread
Histogram	
	

Table B 19 – Explanatory variable: PRIM_PERC

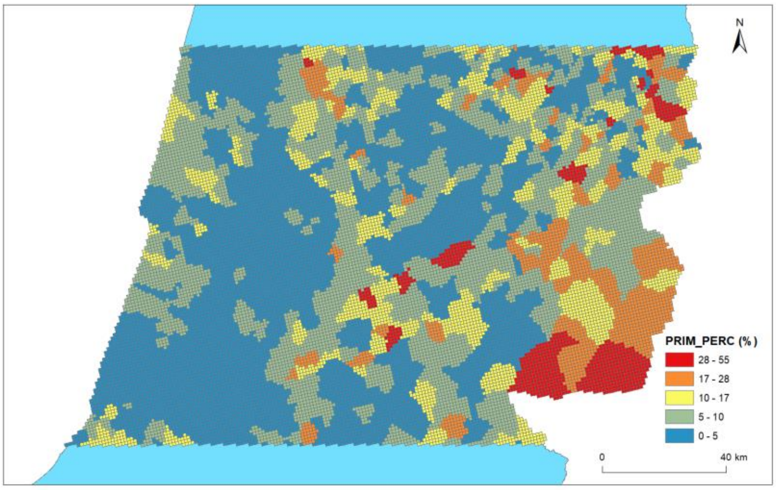
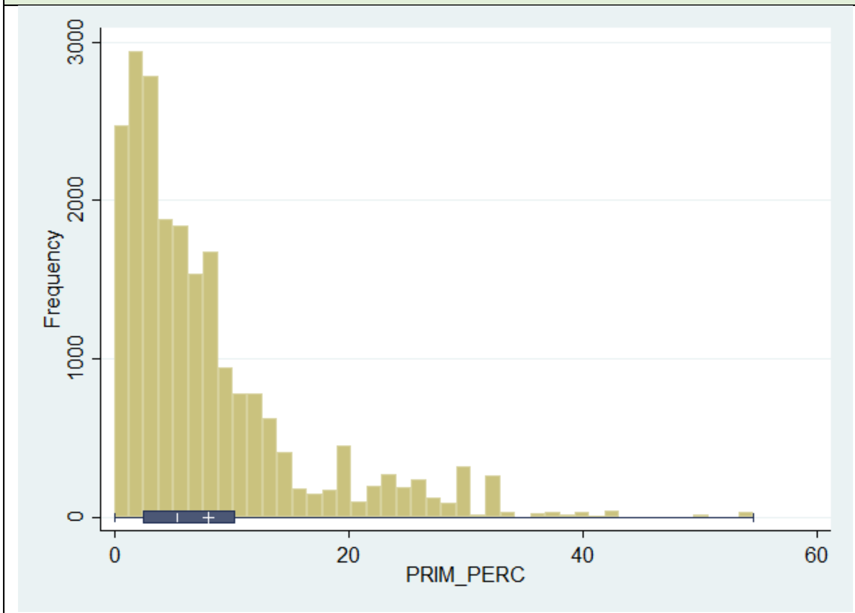
Map	**DESCRIPTION**
	**Variable:** Proportion of population employed in agriculture, animal production, fishing, forestry and hunting (CAE Rev3 A) (%) **Variable name:** PRIM_PERC **Type of variable:** Continuous **Range:** 0% – 54% **Average:** 8.0% **Standard deviation:** 8.0% **Effect:** Ignition and spread
Histogram	
	

Table B 20 – Explanatory variable: AGRMAQ_PERC

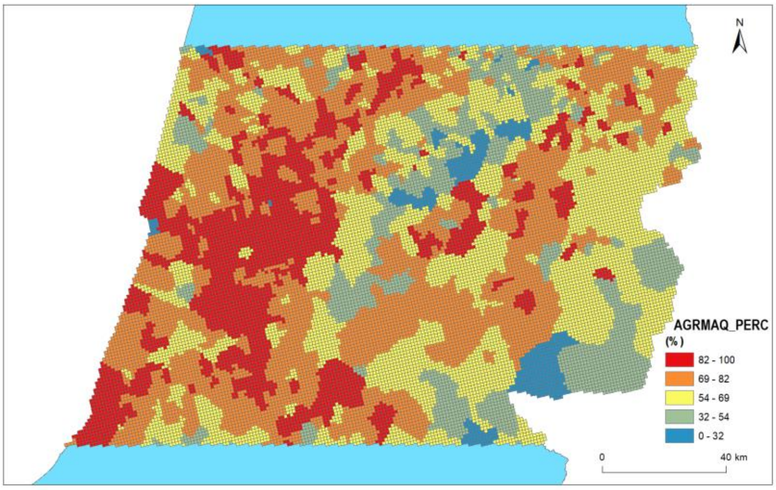
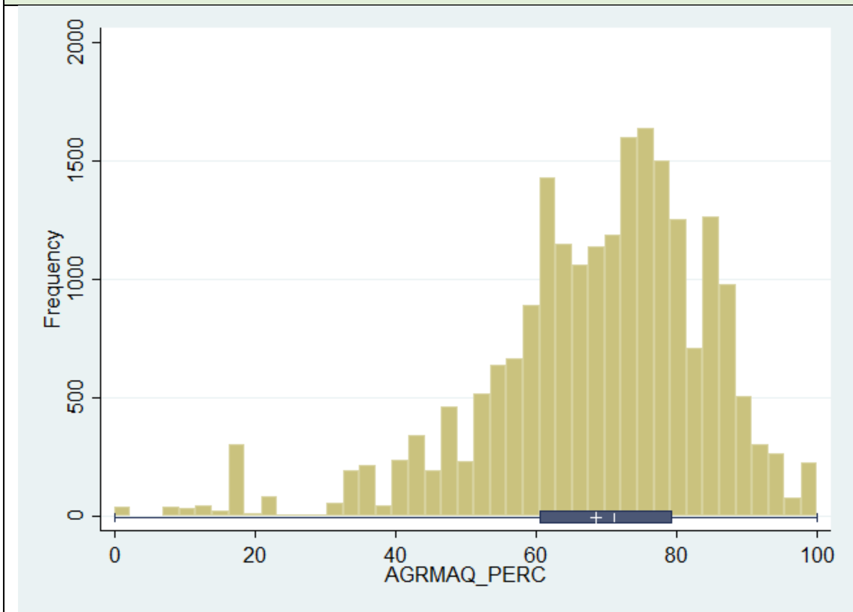
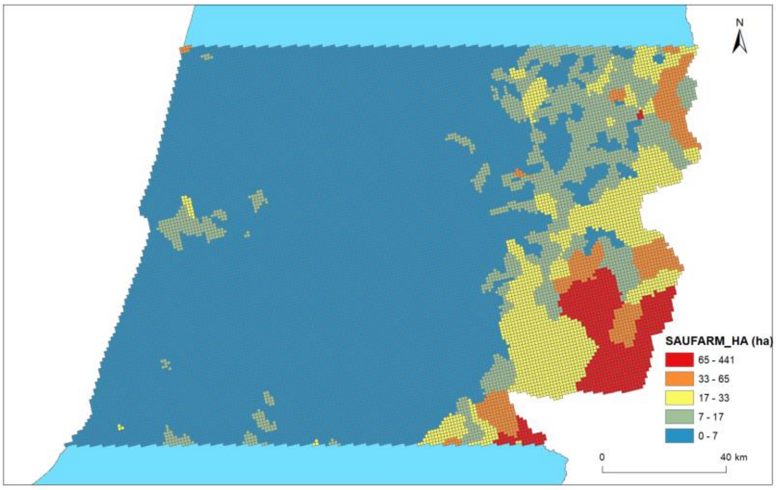
Map	DESCRIPTION
	**Variable:** Proportion of farms with agricultural machines (%) **Variable name:** AGRMAQ_PERC **Type of variable:** Continuous **Range:** 0% – 100% **Average:** 68.6% **Standard deviation:** 16.3% **Effect:** Ignition
Histogram	
	

Table B 21 – Explanatory variable: SAUFARM_HA

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>
</td><td rowspan="3">Variable:
Average used agricultural surface (SAU) per farm (ha)

Variable name:
SAUFARM_HA

Type of variable:
Continuous

Range:
0 ha – 441 ha

Average:
11.1 ha

Standard deviation:
24.5 ha

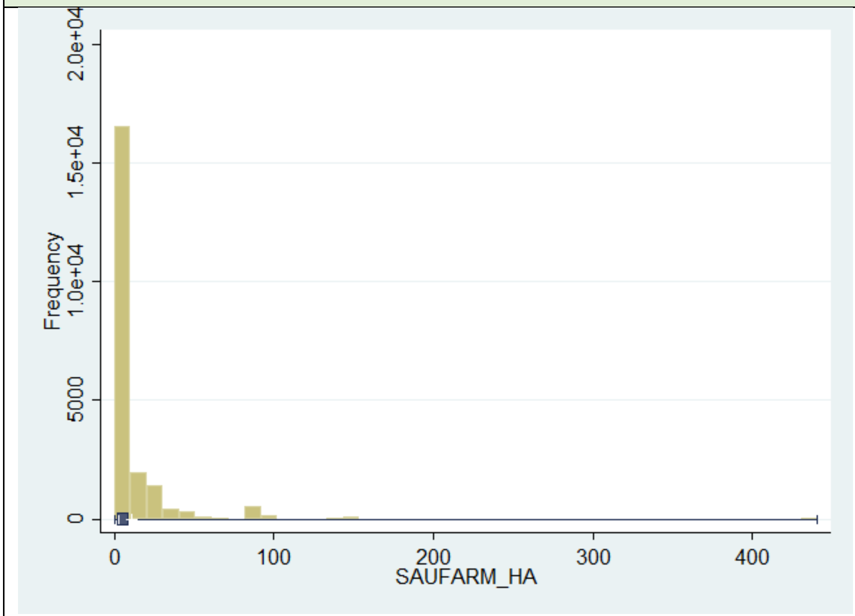
Effect:
Spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 22 – Explanatory variable: FARMDEN_KM

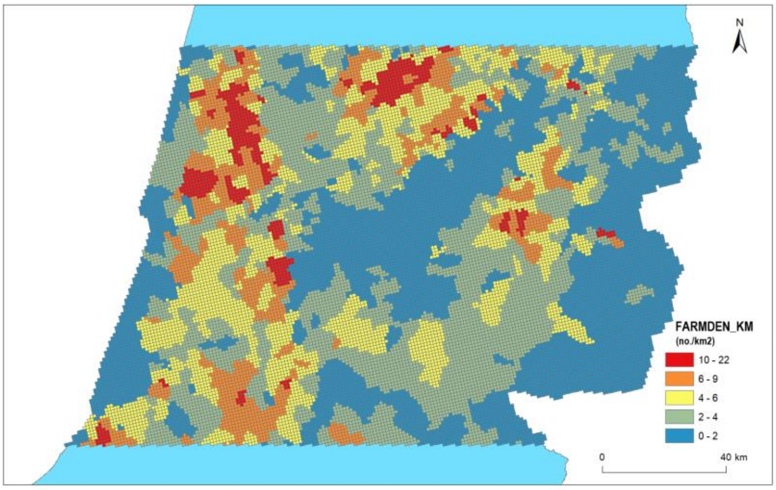
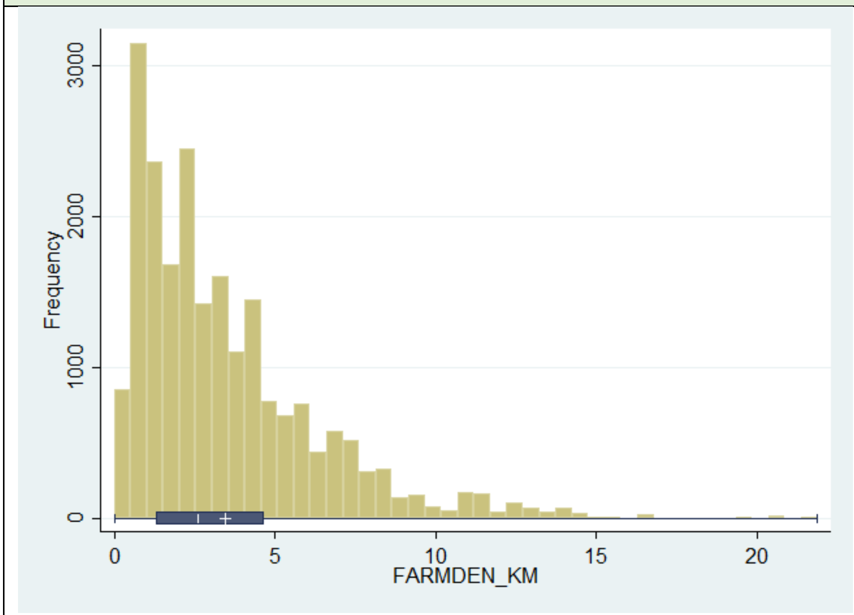
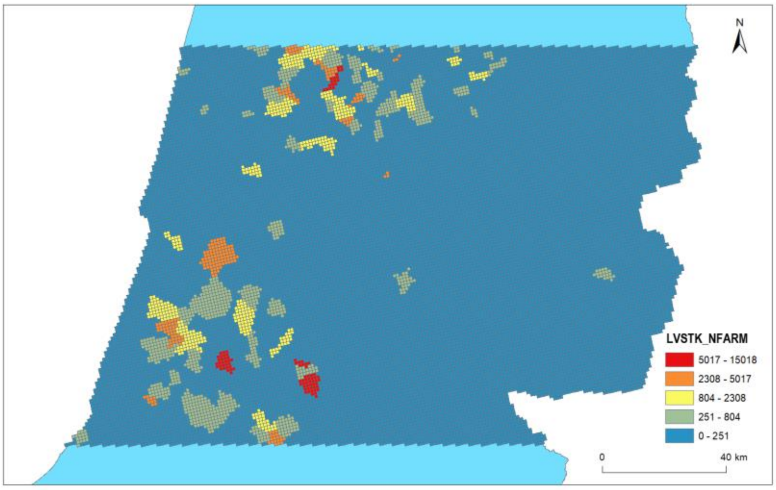
Map	DESCRIPTION
	**Variable:** Farm density (no./km^2) **Variable name:** FARMDEN_KM **Type of variable:** Continuous **Range:** 0 – 22 **Average:** 3.5 **Standard deviation:** 2.8 **Effect:** Spread
Histogram	
	

Table B 23 – Explanatory variable: LVSTK_NFARM

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>


</td>
<td rowspan="3">

Variable:
Average livestock per farm (no.)

Variable name:
LVSTK_NFARM

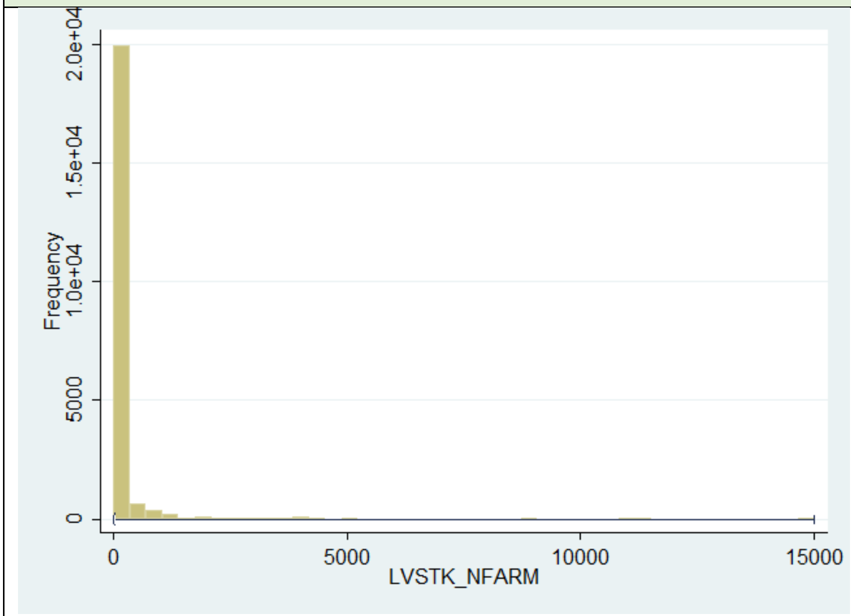
Type of variable:
Discrete (treated as continuous)

Range:
0 – 15018

Average:
186.7

Standard deviation:
925.7

Effect:
Ignition and spread

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>


</td></tr>
</table>

Table B 24 – Explanatory variable: HEADS_NSAU

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td></td><td rowspan="3">Variable:
Normal heads per used agricultural surface (SAU) (no./ha)

Variable name:
HEADS_NSAU

Type of variable:
Continuous

Range:
0 – 14

Average:
1.7

Standard deviation:
2.7

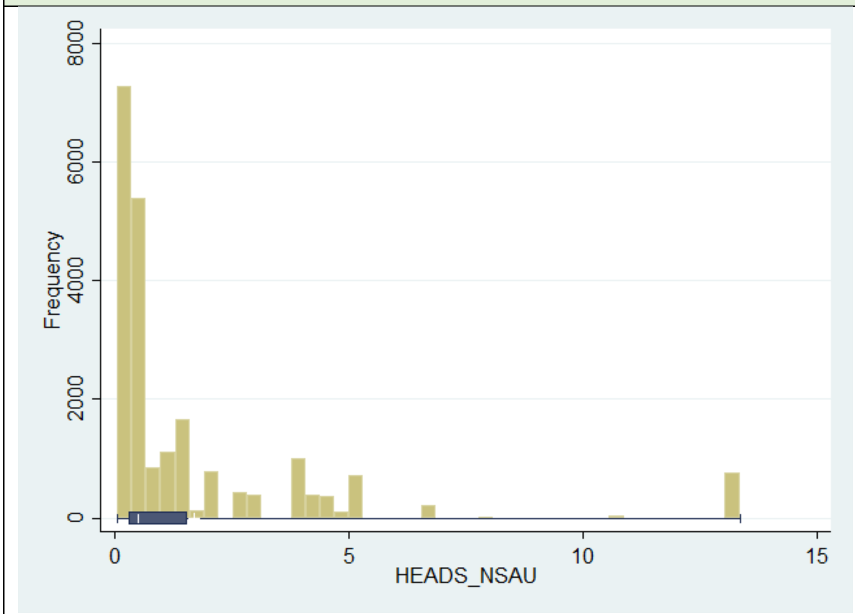
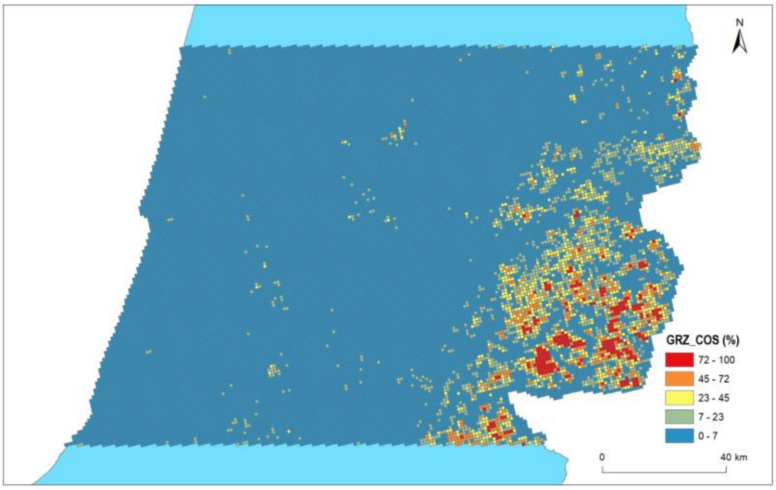
Effect:
Ignition and spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td></td></tr>
</table>

Table B 25 – Explanatory variable: GRZ_COS

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Percentage of pasture cover (grazing areas) in each cell (%)

Variable name:
GRZ_COS

Type of variable:
Continuous

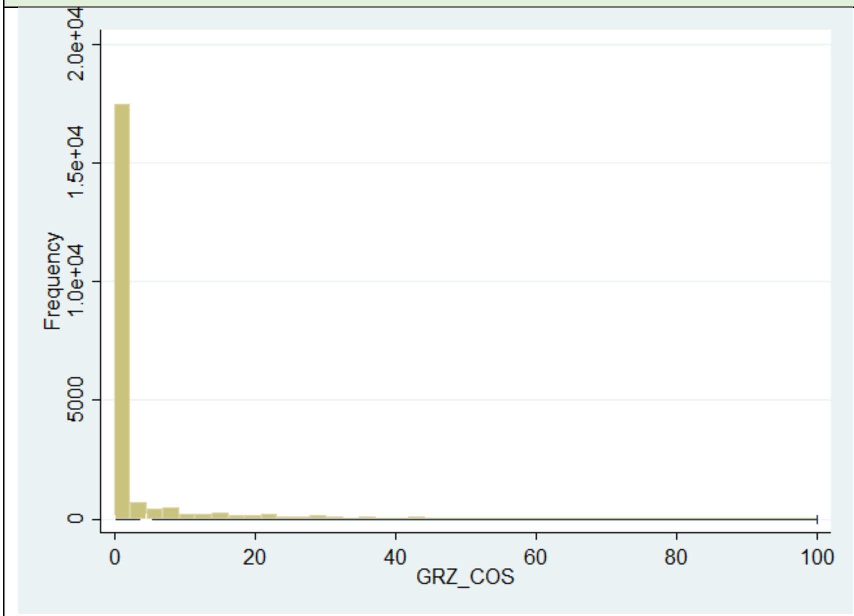
Range:
0% – 100%

Average:
4.6%

Standard deviation:
13.9%

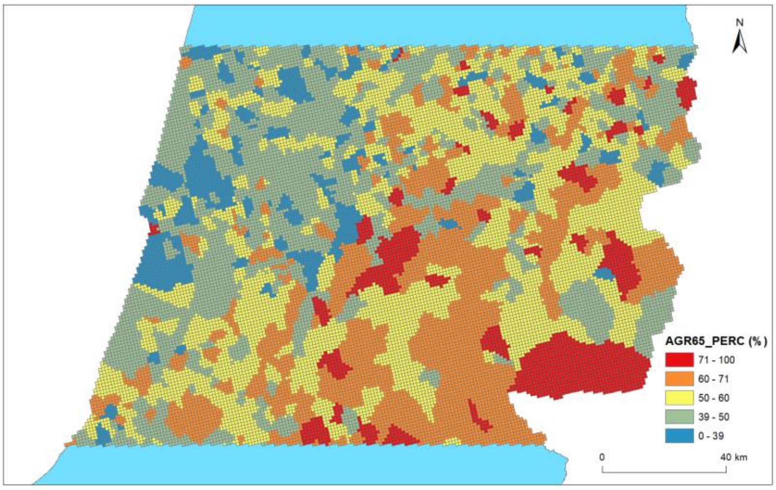
Effect:
Ignition and spread

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 26 – Explanatory variable: AGR65_PERC

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>
</td><td rowspan="3">Variable:
Proportion of single agricultural holders over 65 years of age (%)

Variable name:
AGR65_PERC

Type of variable:
Continuous

Range:
0% – 100%

Average:
55.2%

Standard deviation:
12.3%

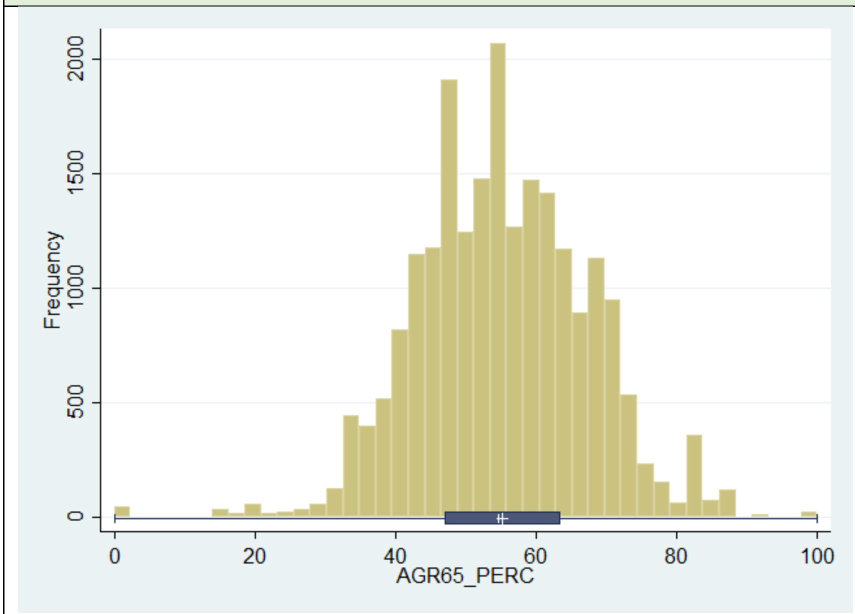
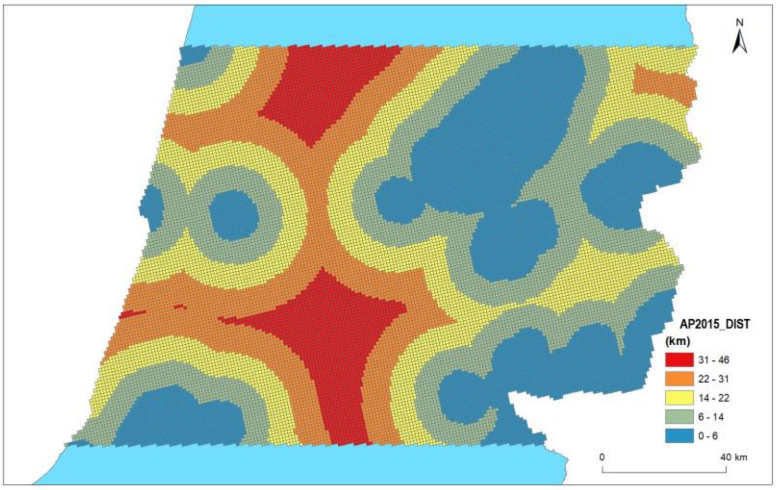
Effect:
Ignition and spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 27 – Explanatory variable: AP2015_DIST

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td></td><td rowspan="3">Variable:
Distance to protected sites (km)

Variable name:
AP2015_DIST

Type of variable:
Continuous

Range:
0 km – 46 km

Average:
14.7 km

Standard deviation:
11.0 km

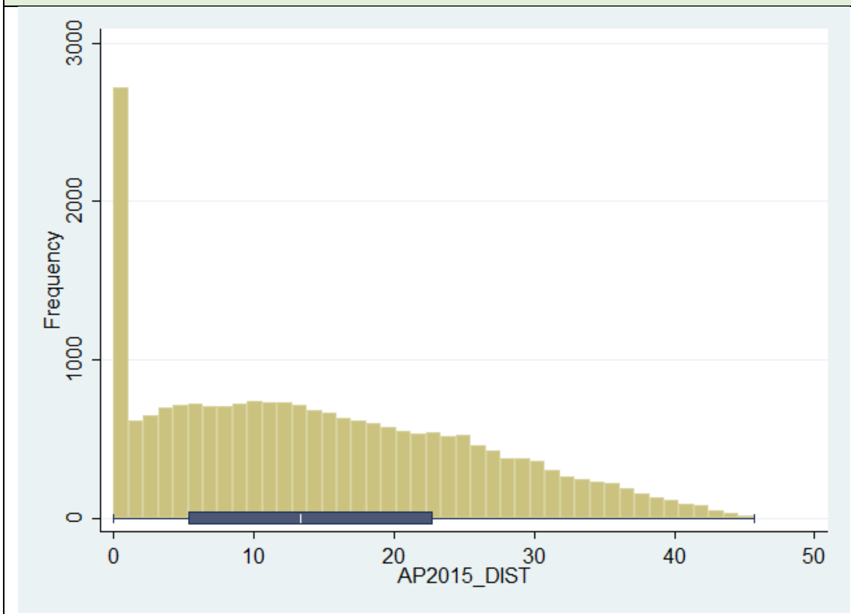
Effect:
Spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td></td></tr>
</table>

Table B 28 – Explanatory variable: NHEST_PERC

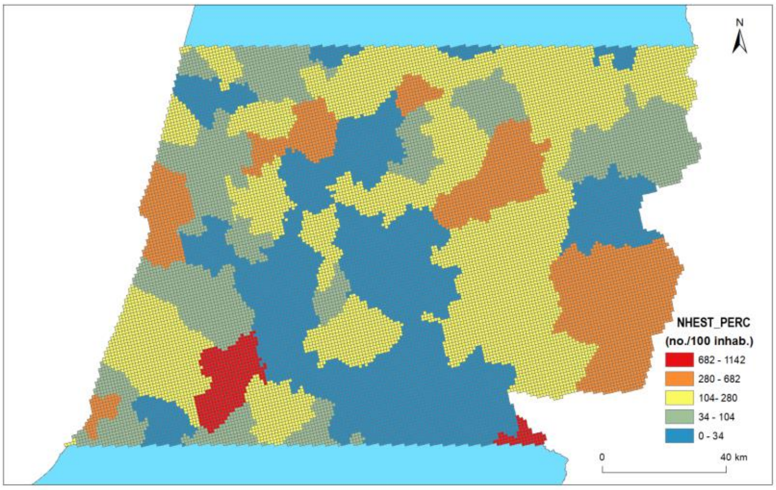
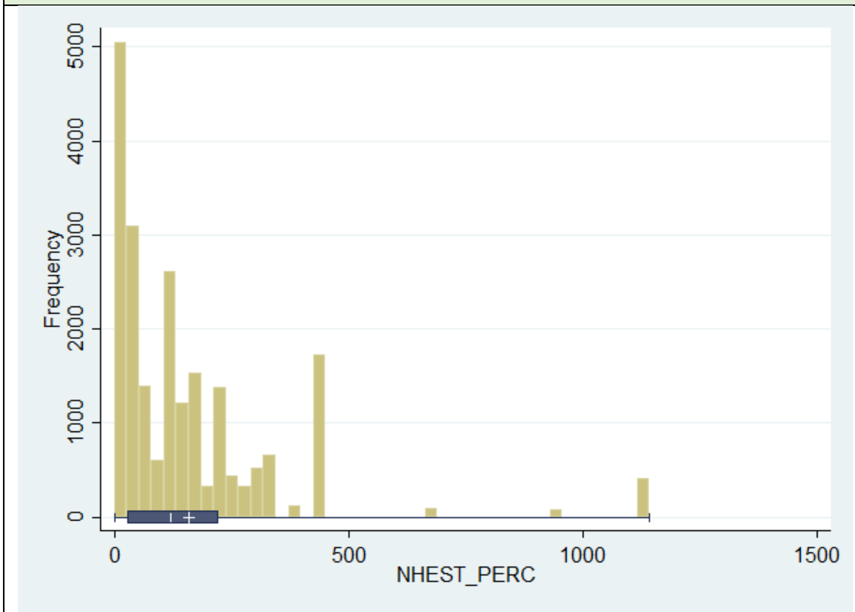
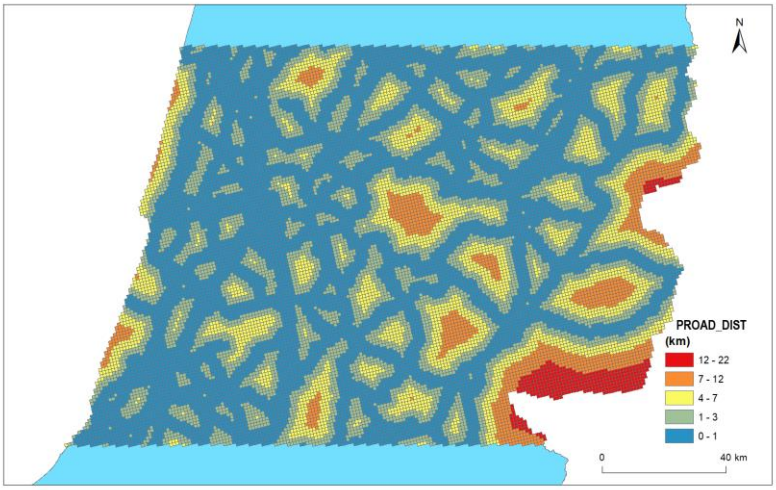
Map	**DESCRIPTION**
 	**Variable:** Nights at hotel establishments (no./100 inhab.) **Variable name:** NHEST_PERC **Type of variable:** Continuous **Range:** 0 – 1142 **Average:** 159.5 **Standard deviation:** 198.1 **Effect:** Ignition
Histogram	
 	

Table B 29 – Explanatory variable: PROAD_DIST

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Distance to primary roads (km)

Variable name:
PROAD_DIST

Type of variable:
Continuous

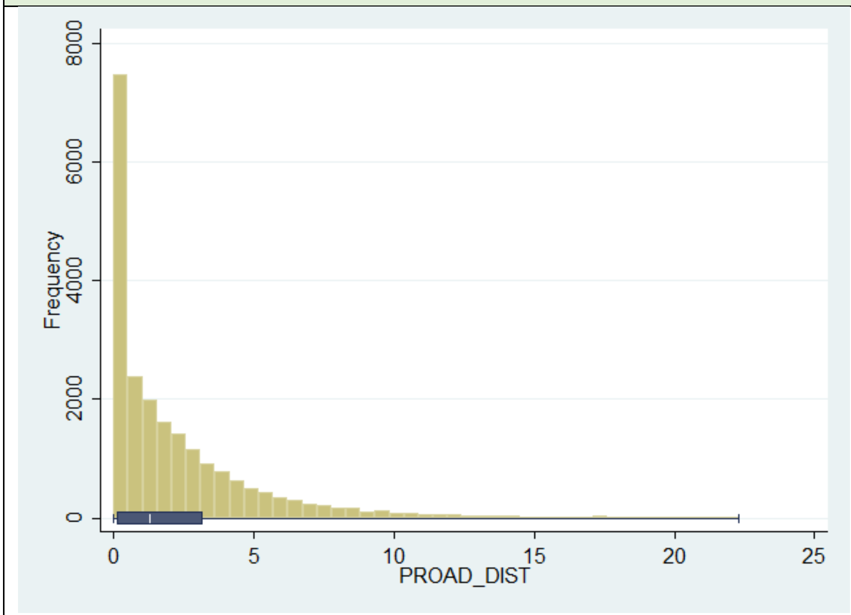
Range:
0 km – 23 km

Average:
2.3 km

Standard deviation:
3.0 km

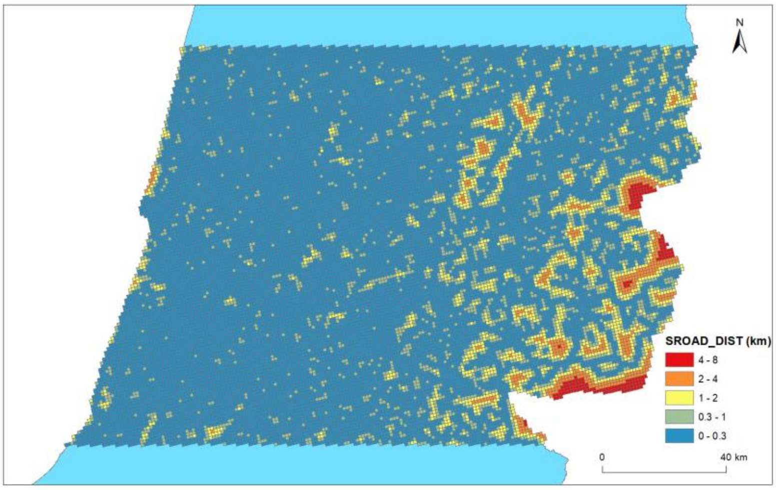
Effect:
Ignition and spread

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 30 – Explanatory variable: SROAD_DIST

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>
</td><td rowspan="3">**Variable:**
Distance to secondary roads (km)

Variable name:
SROAD_DIST

Type of variable:
Continuous

Range:
0 km – 8 km

Average:
0.3 km

Standard deviation:
0.7 km

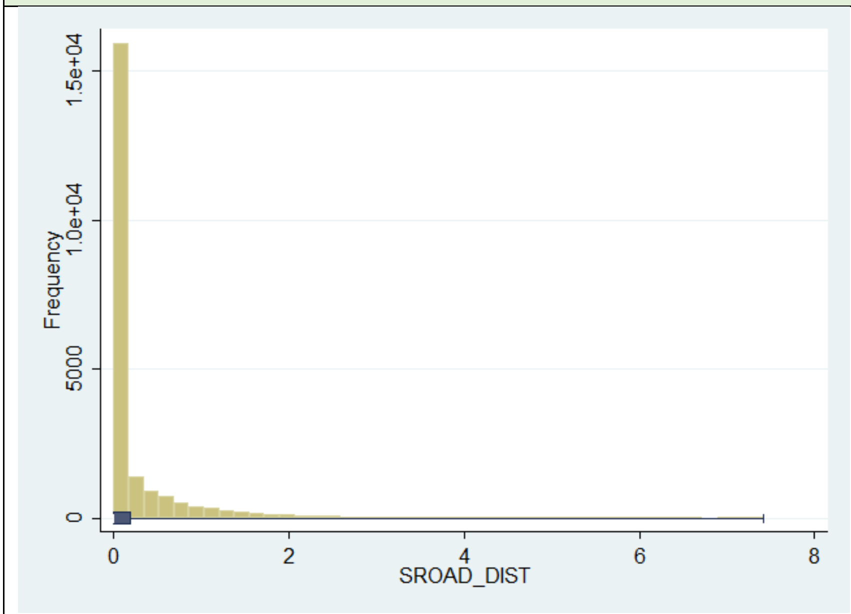
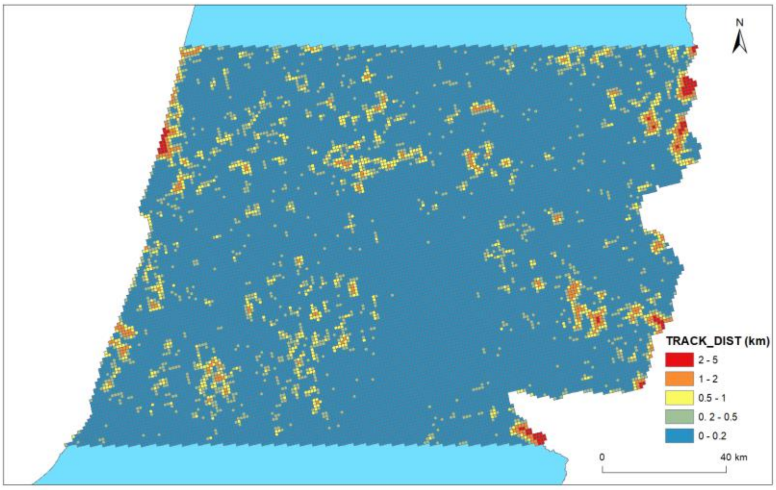
Effect:
Ignition and spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 31 – Explanatory variable: TRACK_DIST

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>
</td>
<td rowspan="3">Variable:
Distance to tracks (km)

Variable name:
TRACK_DIST

Type of variable:
Continuous

Range:
0 km – 5 km

Average:
0.1 km

Standard deviation:
0.3 km

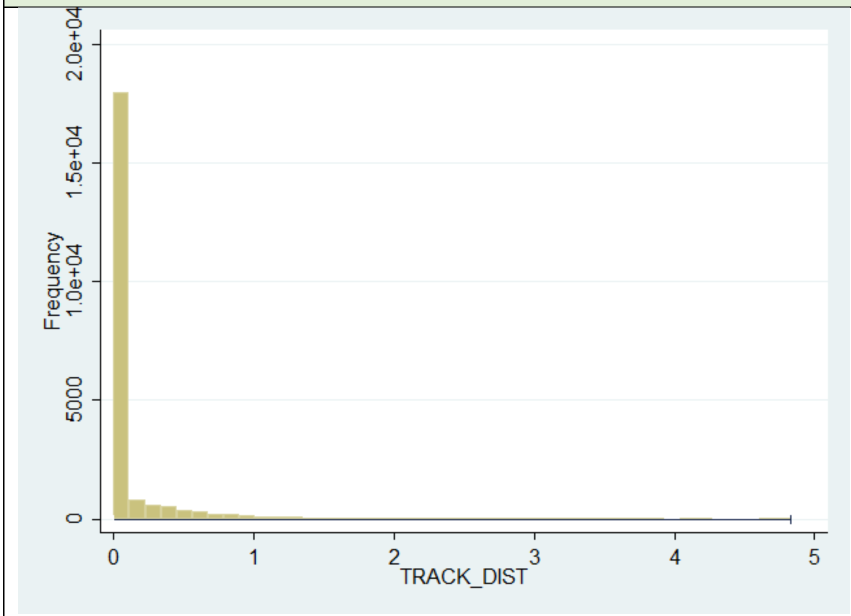
Effect:
Ignition and spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 32 – Explanatory variable: RAIL_DIST

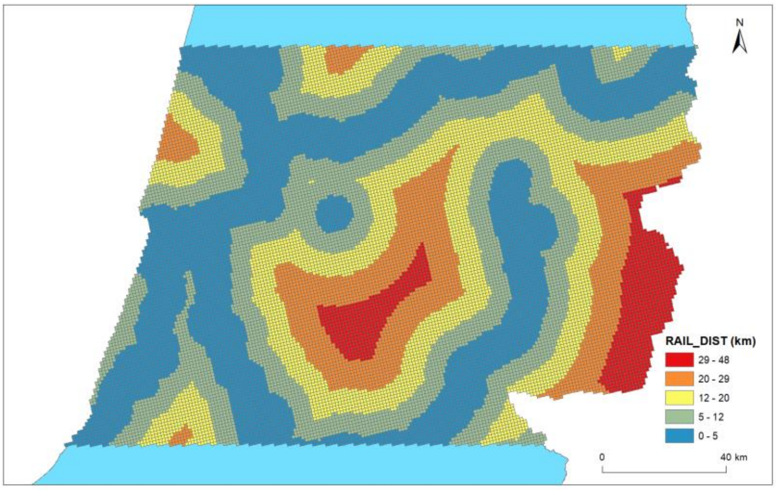
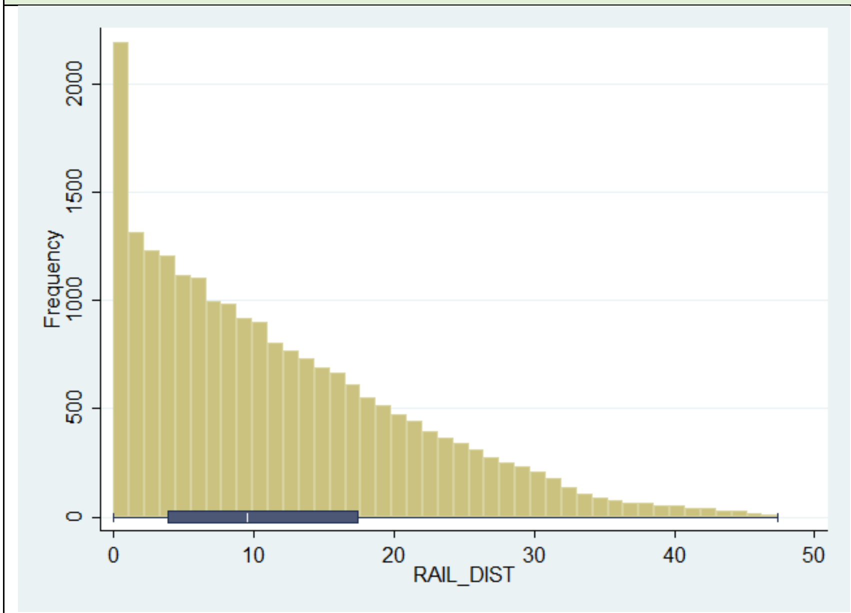
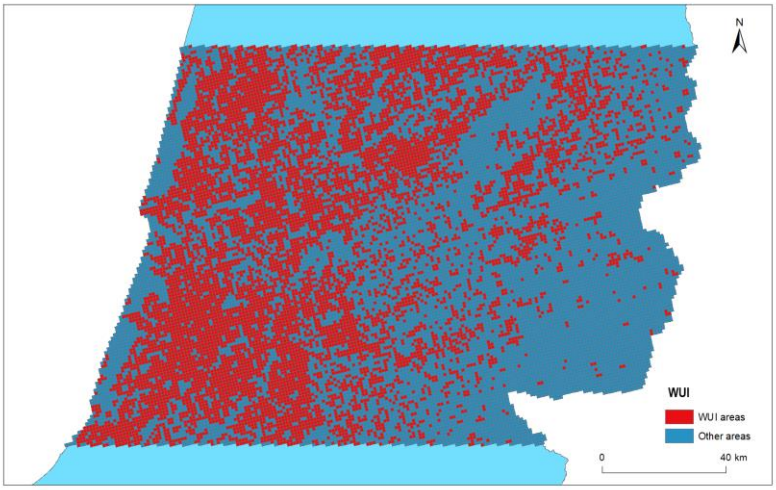
Map	DESCRIPTION
	**Variable:** Distance to railroads (km) **Variable name:** RAIL_DIST **Type of variable:** Continuous **Range:** 0 km – 48 km **Average:** 11.7 km **Standard deviation:** 9.6 km **Effect:** Ignition
Histogram	
	

Table B 33 – Explanatory variable: WUI

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>
</td><td rowspan="3">Variable:
Wildland-Urban Interface areas

Variable name:
WUI

Type of variable:
Binary

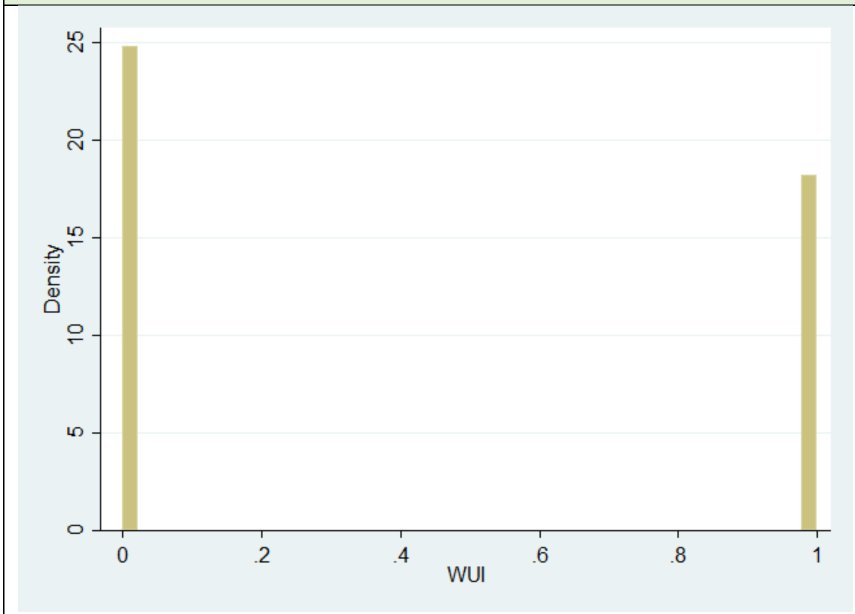
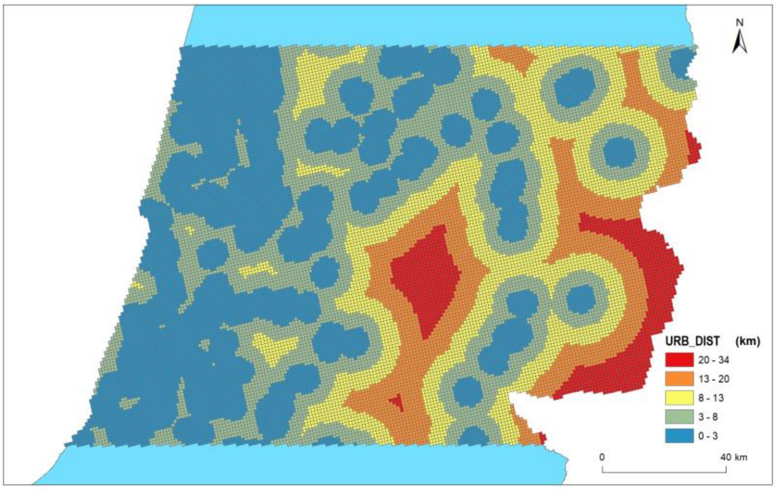
Effect:
Ignition and spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 34 – Explanatory variable: URB_DIST

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Distance to urban areas and infrastructures (km)

Variable name:
URB_DIST

Type of variable:
Continuous

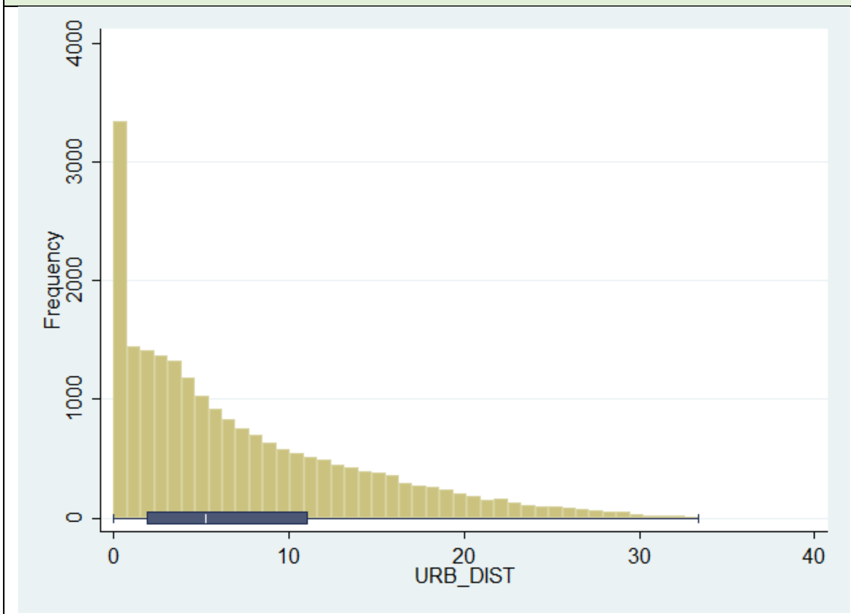
Range:
0 km – 34 km

Average:
7.3 km

Standard deviation:
6.8 km

Effect:
Ignition and spread

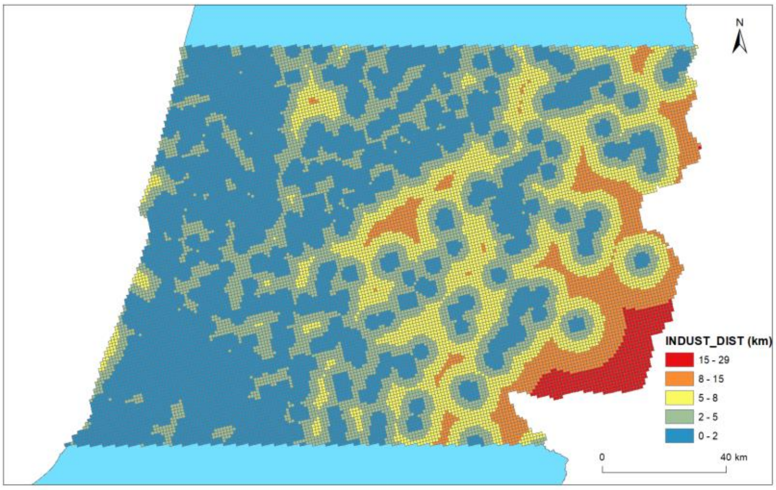
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 35 – Explanatory variable: INDUST_DIST

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Distance to industrial sites (km)

Variable name:
INDUST_DIST

Type of variable:
Continuous

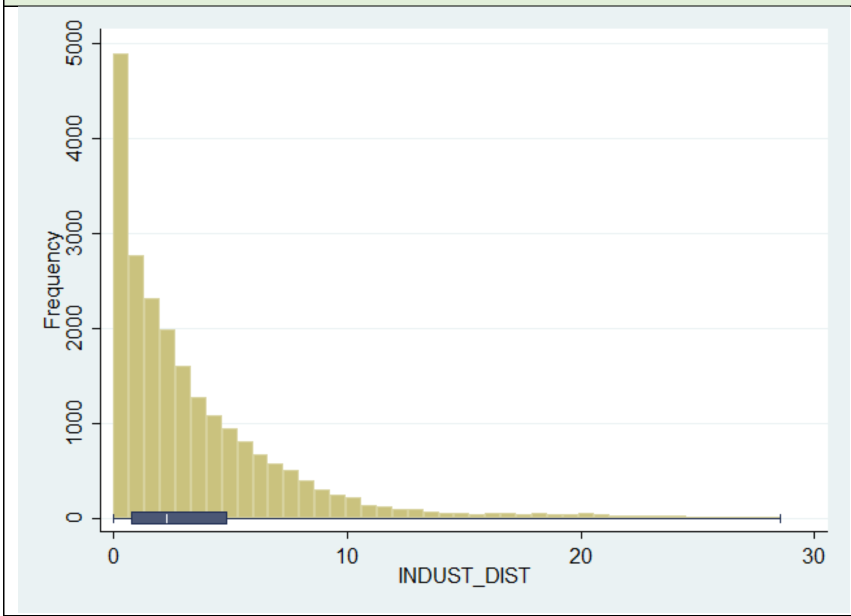
Range:
0 km – 29 km

Average:
3.5 km

Standard deviation:
3.9 km

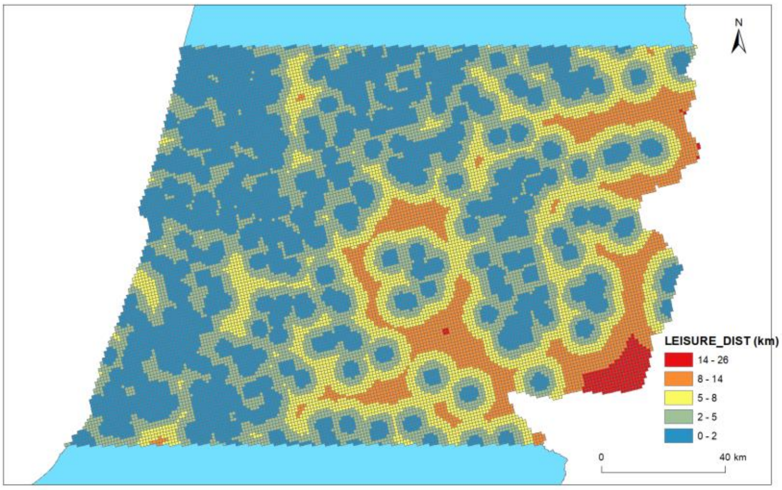
Effect:
Ignition

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 36 – Explanatory variable: LEISURE_DIST

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td></td><td rowspan="3">Variable:
Distance to recreational areas and touristic zones (km)

Variable name:
LEISURE_DIST

Type of variable:
Continuous

Range:
0 km – 26 km

Average:
3.7 km

Standard deviation:
3.2 km

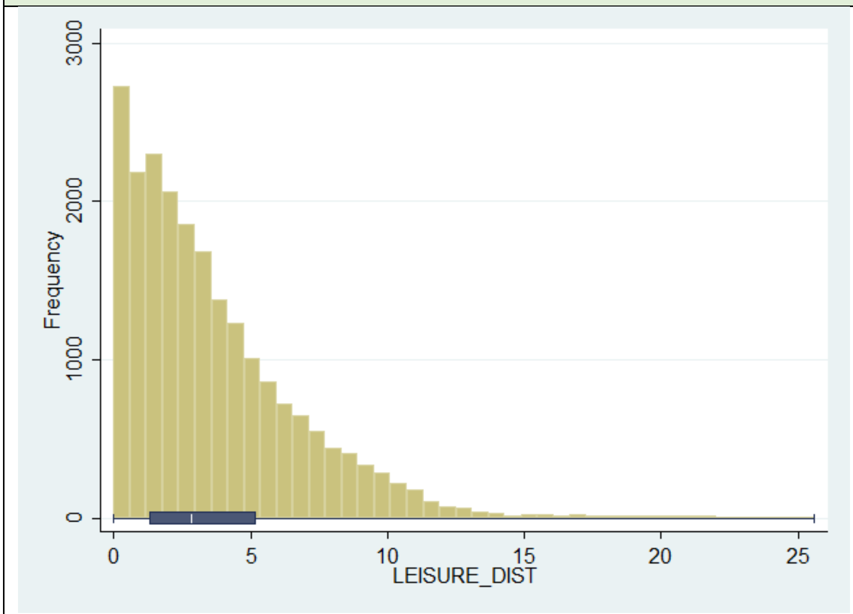
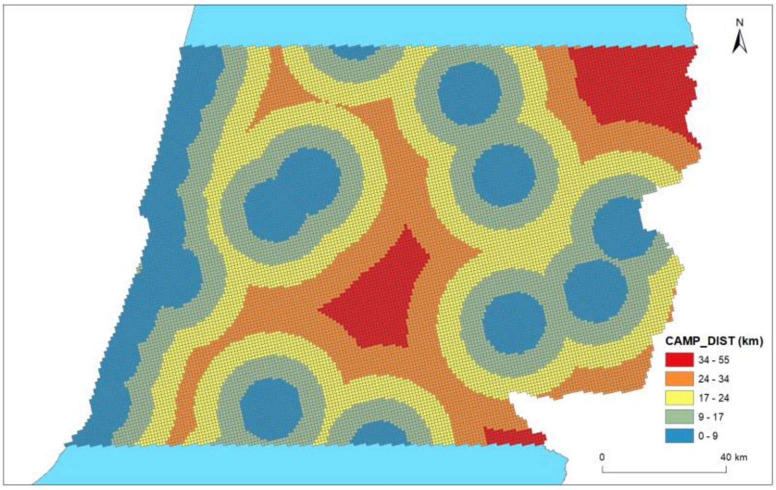
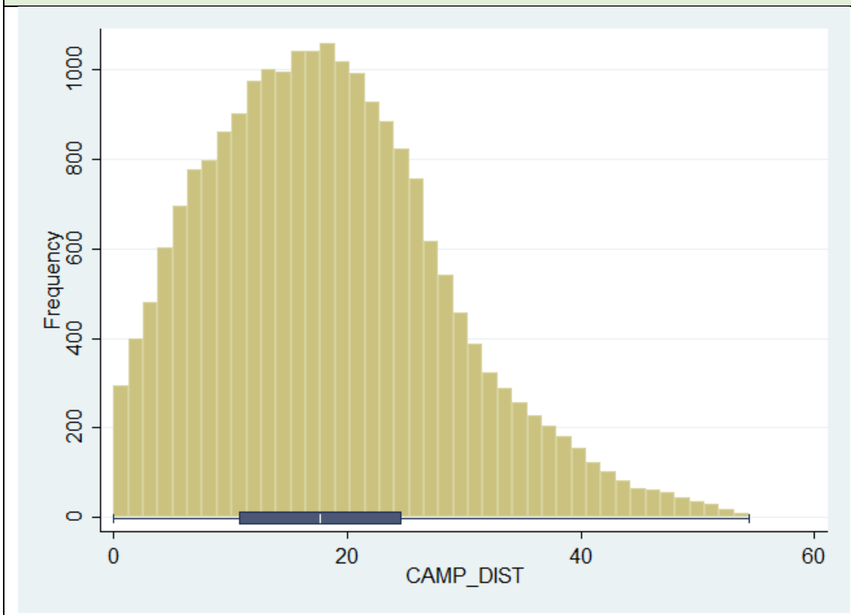
Effect:
Ignition</td></tr>
<tr><td>Histogram</td></tr>
<tr><td></td></tr>
</table>

Table B 37 – Explanatory variable: CAMP_DIST

Map



Histogram



DESCRIPTION

Variable:
Distance to campsites (km)

Variable name:
CAMP_DIST

Type of variable:
Continuous

Range:
0 km – 55 km

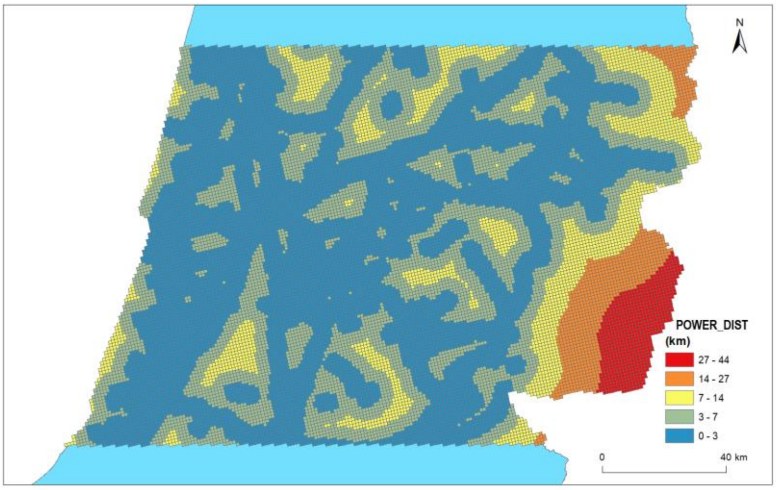
Average:
18.4 km

Standard deviation:
10.1 km

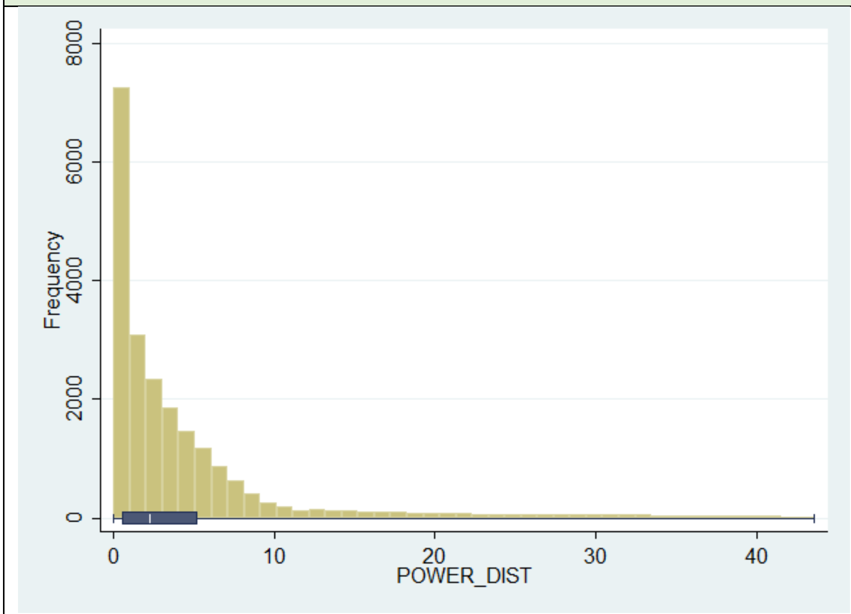
Effect:
Ignition

Table B 38 – Explanatory variable: POWER_DIST

Map



Histogram



DESCRIPTION

Variable:
Distance to electric lines (km)

Variable name:
POWER_DIST

Type of variable:
Continuous

Range:
0 km – 44 km

Average:
4.6 km

Standard deviation:
7.0 km

Effect:
Ignition

Table B 39 – Explanatory variable: LDFLL_DIST

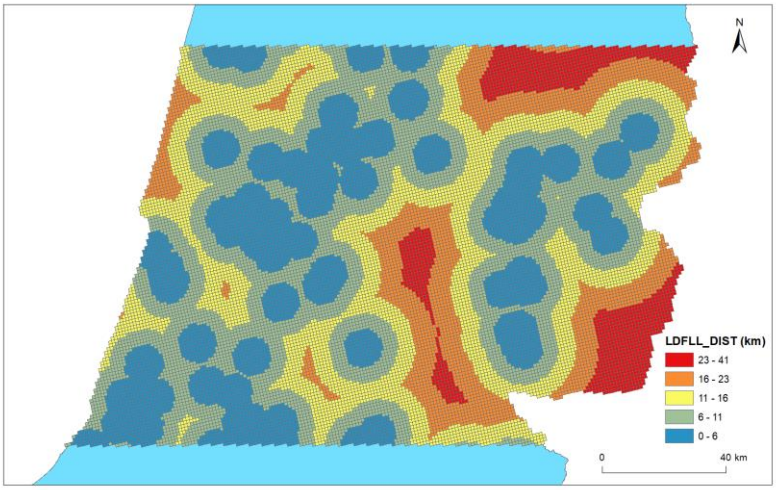
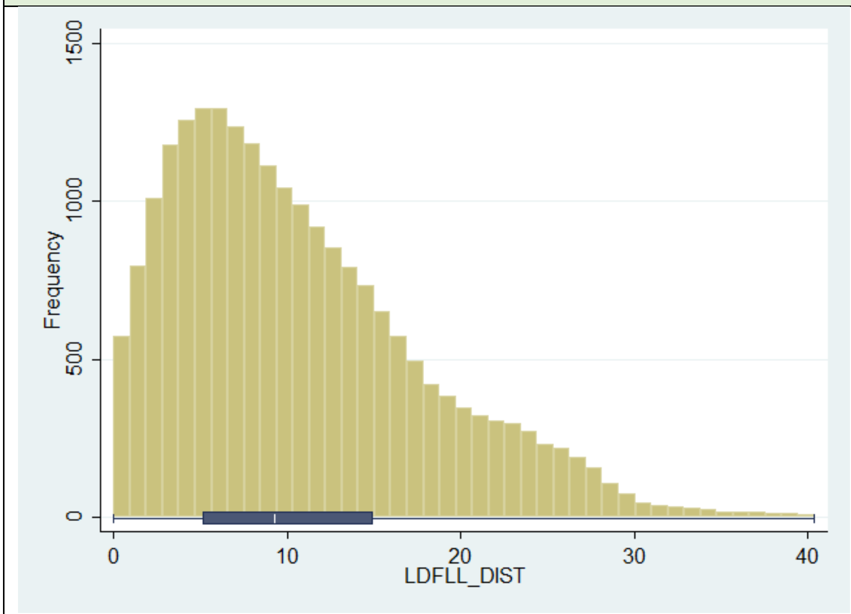
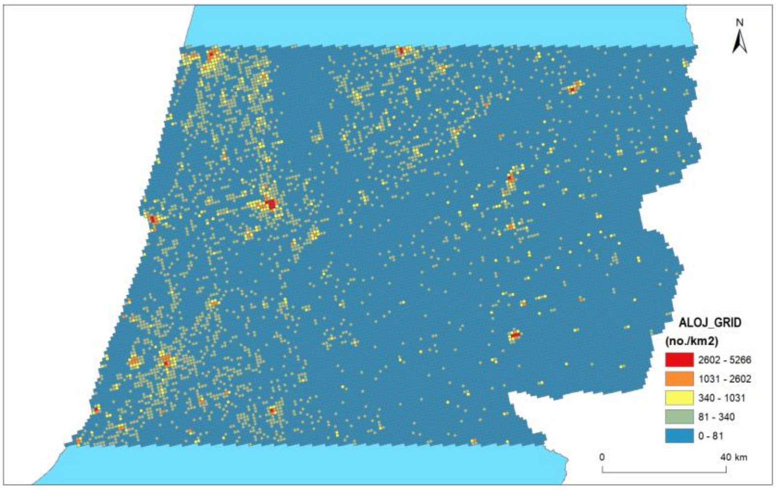
Map	DESCRIPTION
	**Variable:** Distance to landfills (km) **Variable name:** LDFLL_DIST **Type of variable:** Continuous **Range:** 0 km – 41 km **Average:** 10.8 km **Standard deviation:** 7.3 km **Effect:** Ignition
Histogram	
	

Table B 40 – Explanatory variable: ALOJ_GRID

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>
</td><td rowspan="3">Variable:
Housing density (no./km^2)

Variable name:
ALOJ_GRID

Type of variable:
Continuous

Range:
0 – 5266

Average:
47.8

Standard deviation:
175.8

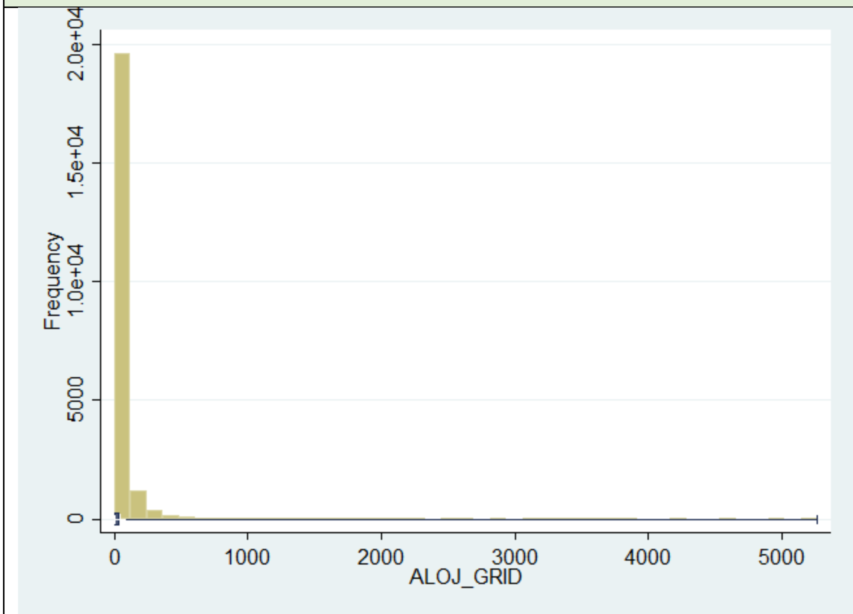
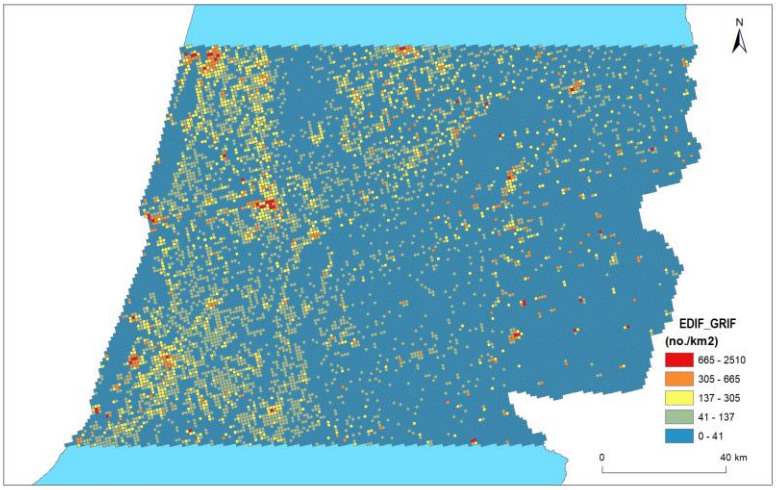
Effect:
Ignition and spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 41 – Explanatory variable: EDIF_GRID

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>
</td><td rowspan="3">

Variable:
Building density (no./km^2)

Variable name:
EDIF_GRID

Type of variable:
Continuous

Range:
0 – 2510

Average:
36.5

Standard deviation:
84.5

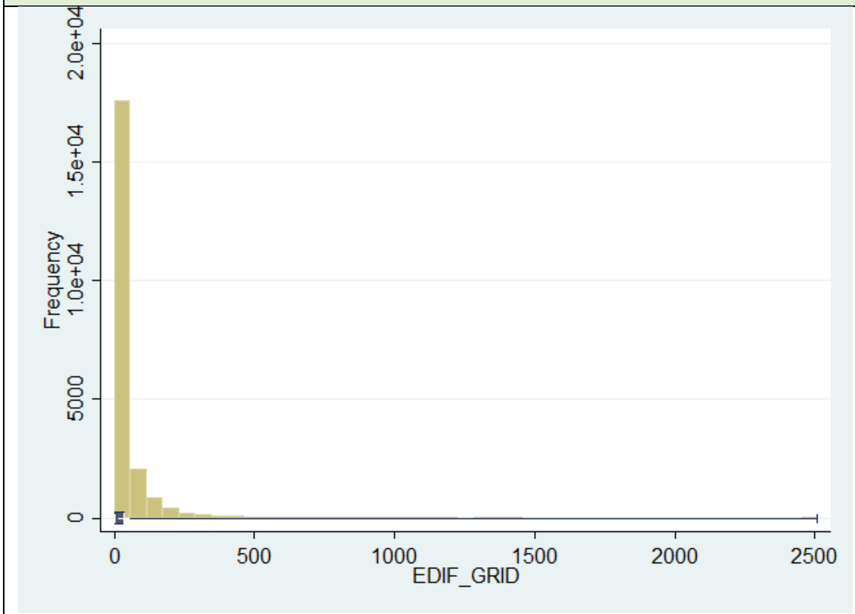
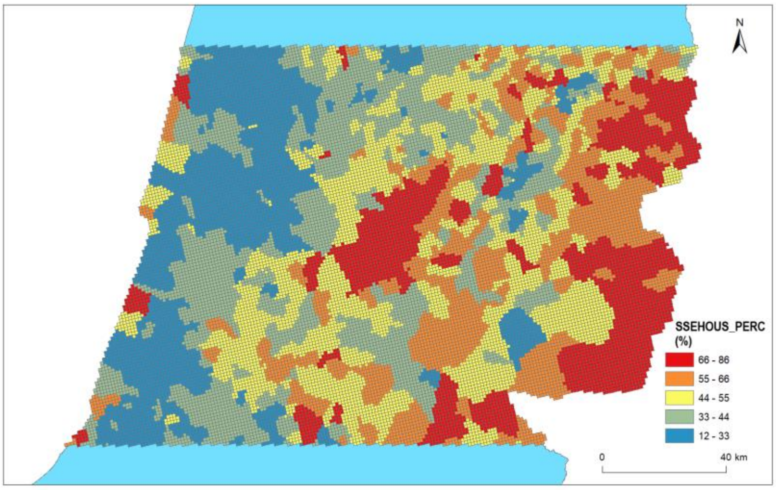
Effect:
Ignition and spread
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 42 – Explanatory variable: SSEHOUS_PERC

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Proportion of seasonal use, secondary residence or empty housing (%)

Variable name:
SSEHOUS_PERC

Type of variable:
Continuous

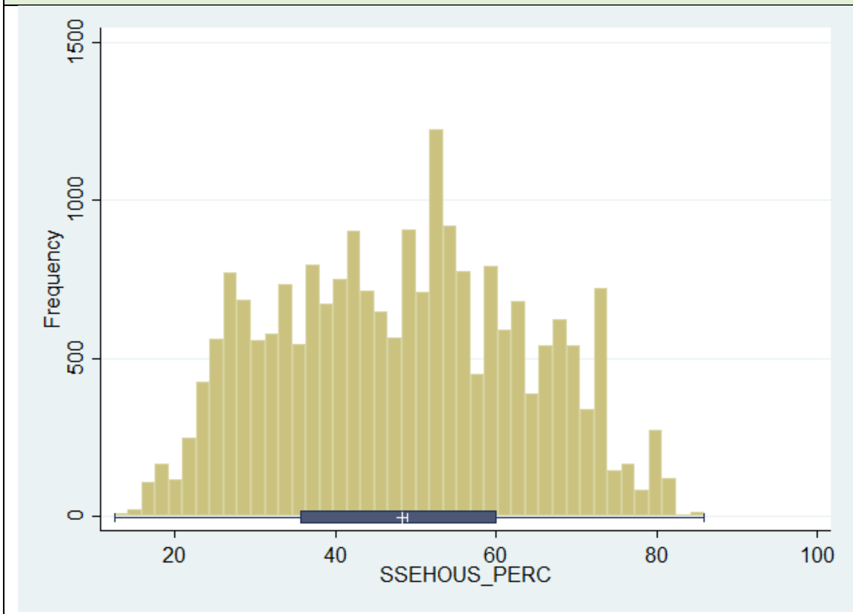
Range:
12% – 86%

Average:
48.2%

Standard deviation:
15.5%

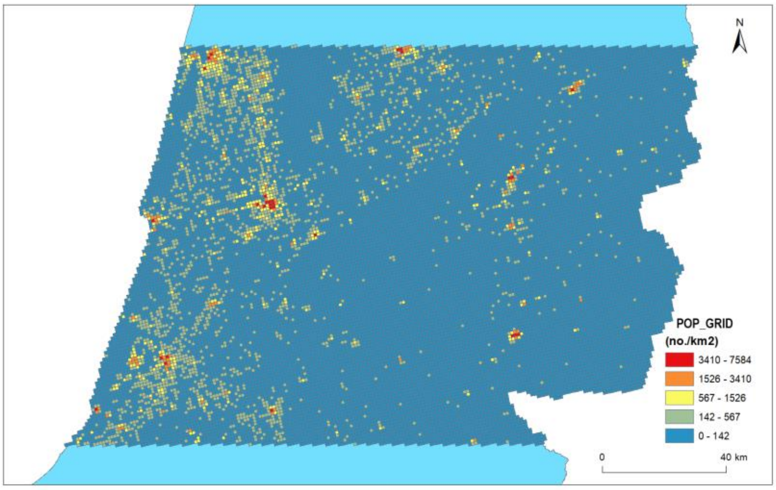
Effect:
Ignition and spread

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 43 – Explanatory variable: POP_GRID

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>
</td><td rowspan="3">Variable:
Population density (inhab./km^2)

Variable name:
POP_GRID

Type of variable:
Continuous

Range:
0 – 7584

Average:
75.6

Standard deviation:
271.7

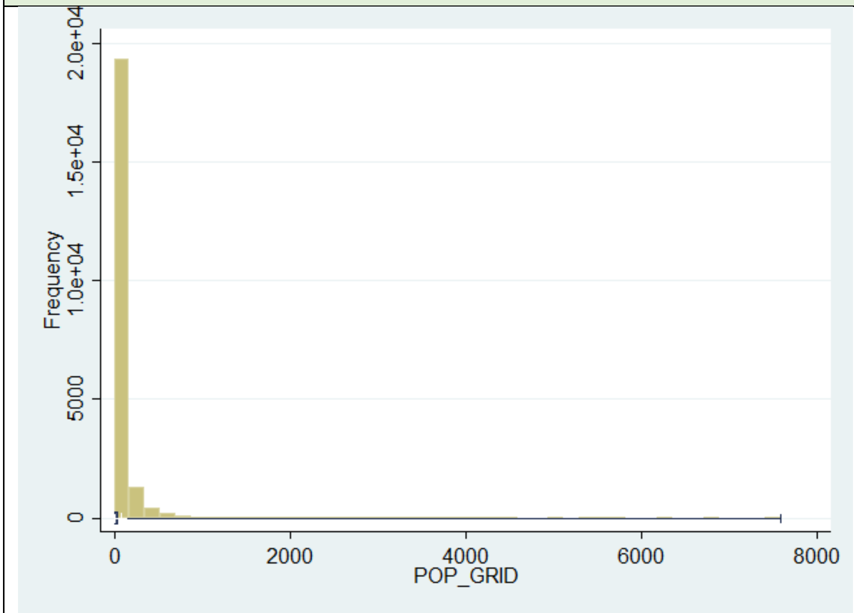
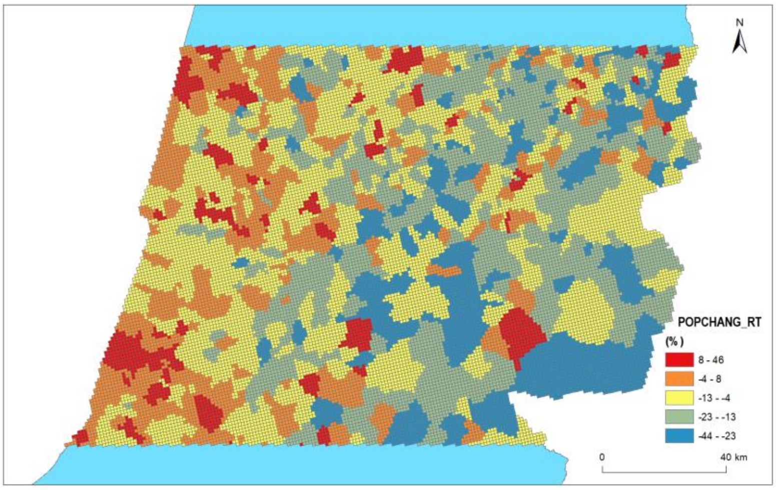
Effect:
Ignition and spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 44 – Explanatory variable: POPCHANG_RT

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Rate of population change by parish (2001-2011) (%)

Variable name:
POPCHANG_RT

Type of variable:
Continuous

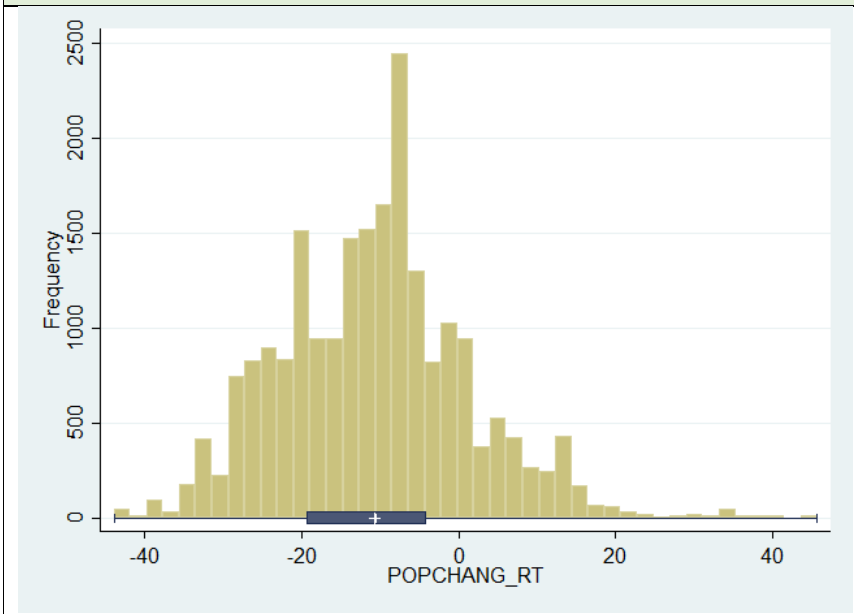
Range:
-43% – 45%

Average:
-10.8%

Standard deviation:
11.9%

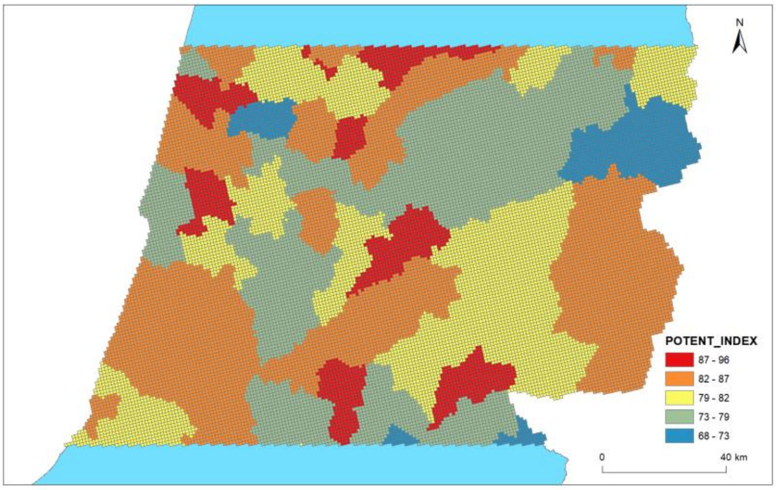
Effect:
Ignition and spread

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 45 – Explanatory variable: POTENT_INDEX

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>
</td>
<td rowspan="3">Variable:
Potentiality index by municipality

Variable name:
POTENT_INDEX

Type of variable:
Continuous

Range:
68 – 96

Average:
81.2

Standard deviation:
4.8

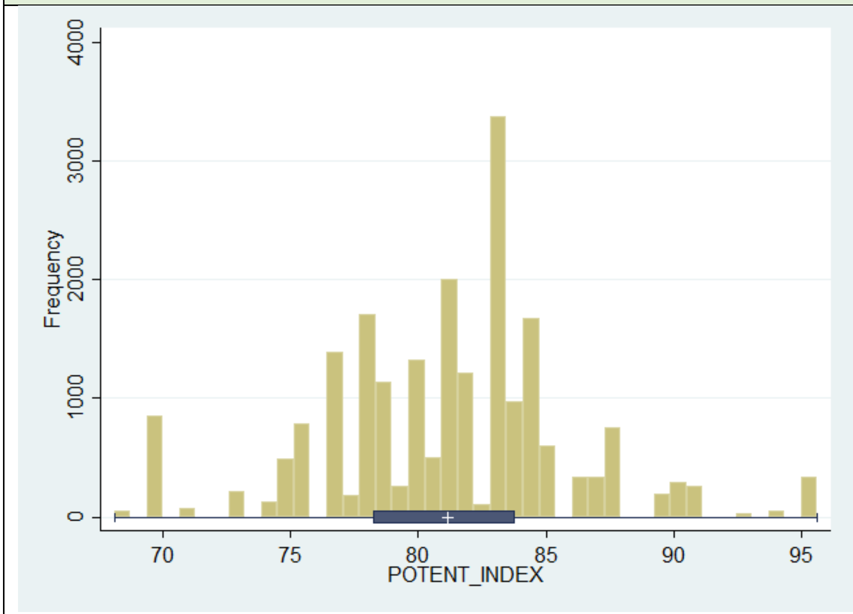
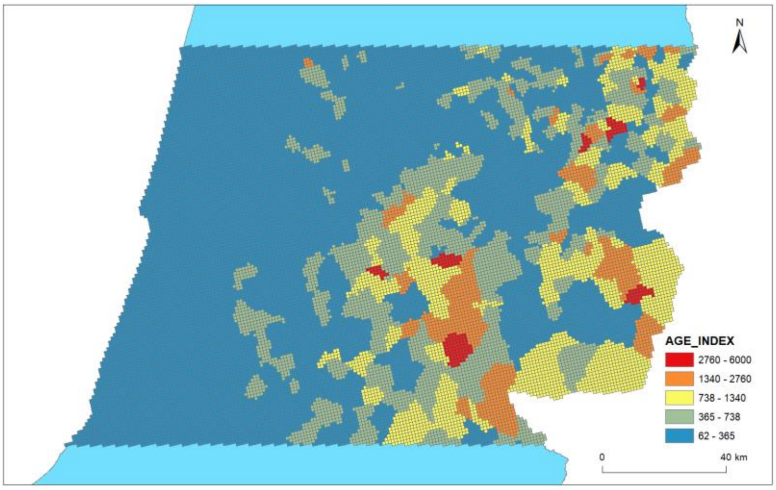
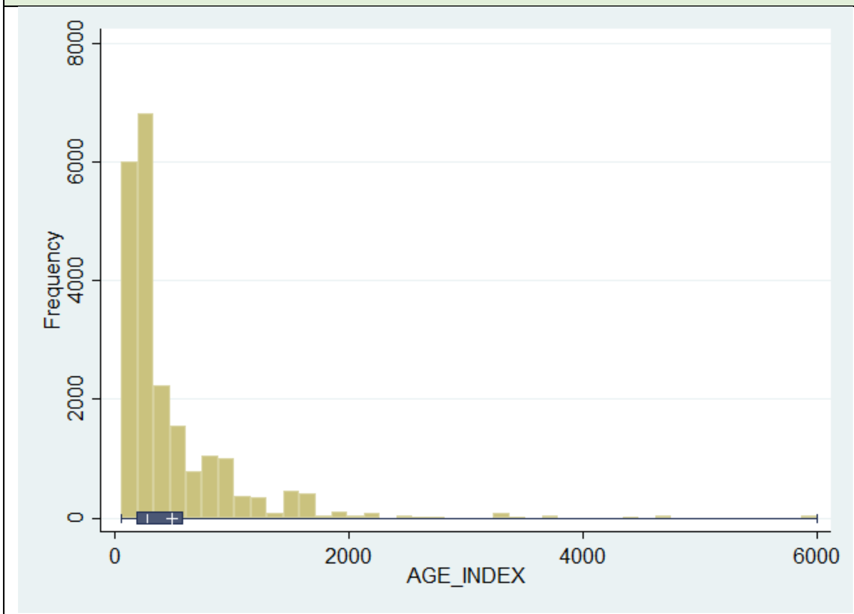
Effect:
Ignition and spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 46 – Explanatory variable: AGE_INDEX

Map



Histogram



DESCRIPTION

Variable:
Ageing index by parish

Variable name:
AGE_INDEX

Type of variable:
Continuous

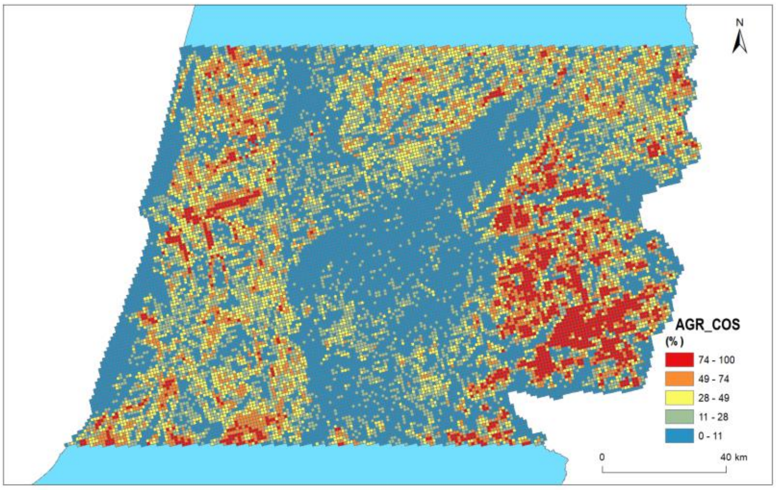
Range:
62 – 6000

Average:
491.1

Standard deviation:
565.4

Effect:
Ignition and spread

Table B 47 – Explanatory variable: AGR_COS

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>


</td><td rowspan="3">
Variable:
Percentage of agricultural land cover in each cell (%)

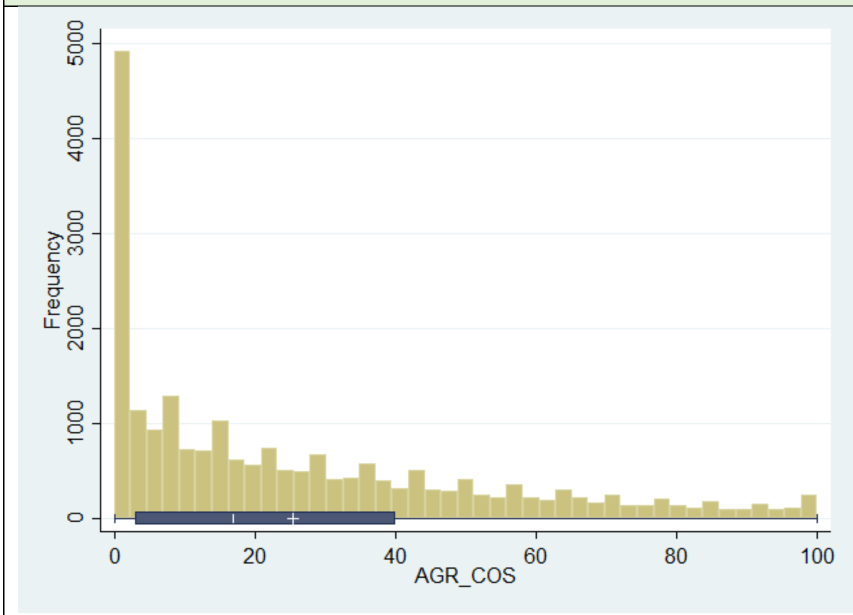
Variable name:
AGR_COS

Type of variable:
Continuous

Range:
0% – 100%

Average:
25.4%

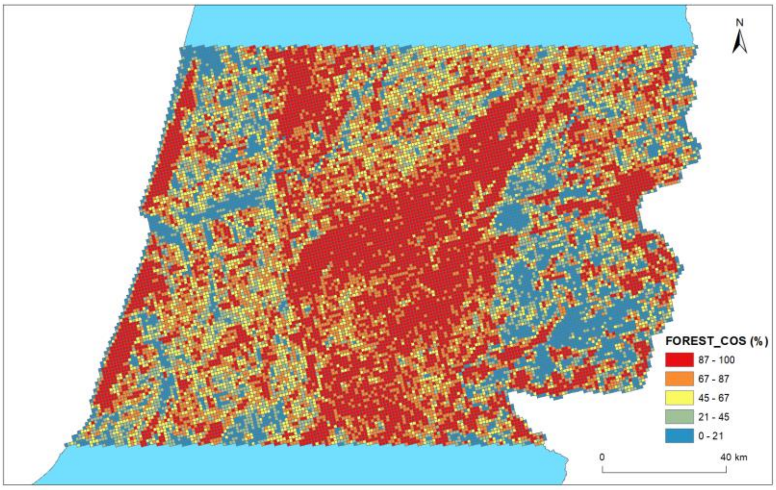
Standard deviation:
26.1%

Effect:
Ignition and spread
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>


</td></tr>
</table>

Table B 48 – Explanatory variable: FOREST_COS

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Percentage of forest cover in each cell (%)

Variable name:
FOREST_COS

Type of variable:
Continuous

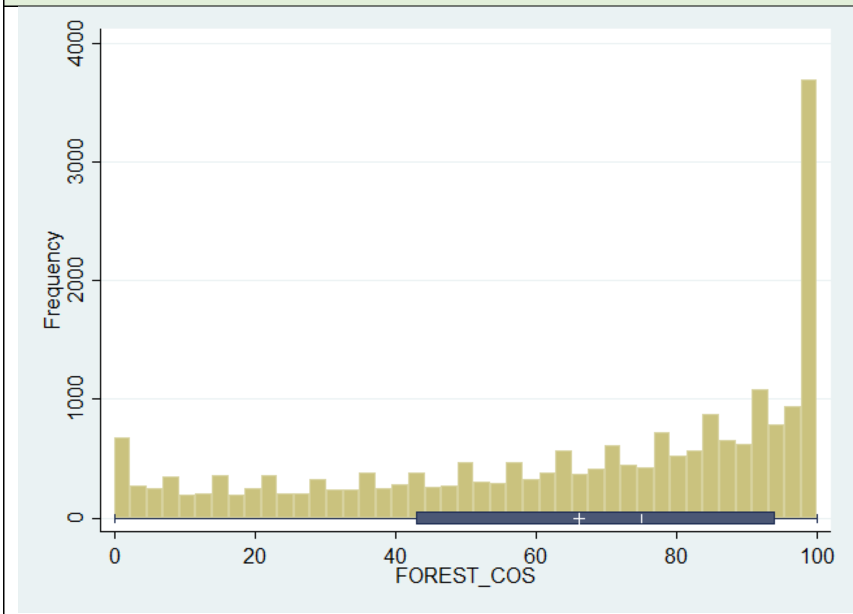
Range:
0% – 100%

Average:
66.0%

Standard deviation:
30.9%

Effect:
Ignition and spread

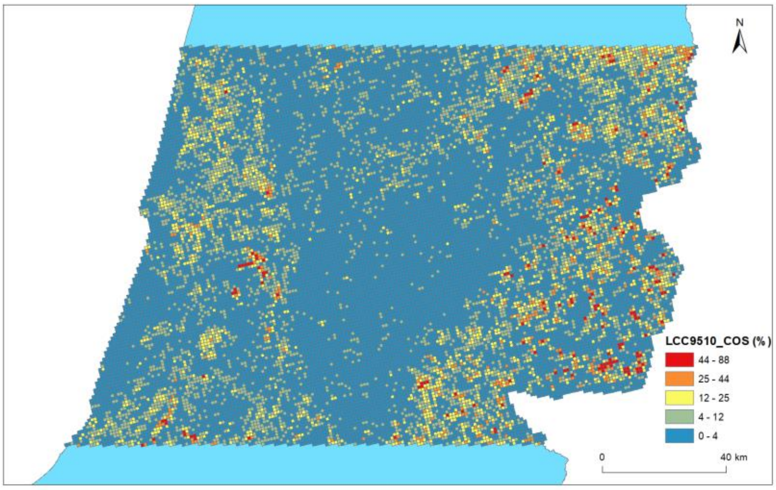
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 49 – Explanatory variable: LCC9519_COS

<table>
<tr><td>Map</td><td>DESCRIPTION</td></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Percentage of land that had changes in use/cover (agricultural areas to forested and natural areas and other areas to forested and natural areas) in each cell (1995-2010) (%)

Variable name:
LCC9510_COS

Type of variable:
Continuous

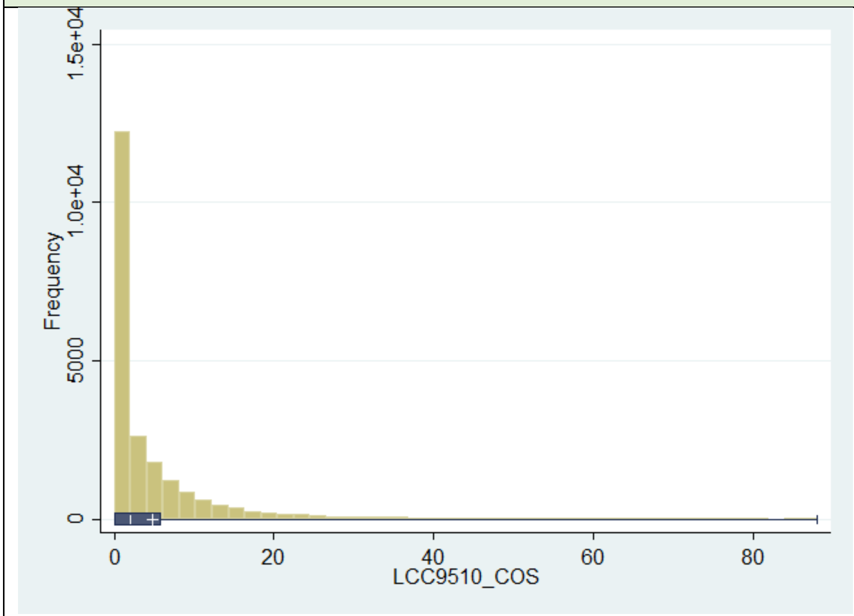
Range:
0% – 88%

Average:
4.7%

Standard deviation:
8.1%

Effect:
Ignition and spread

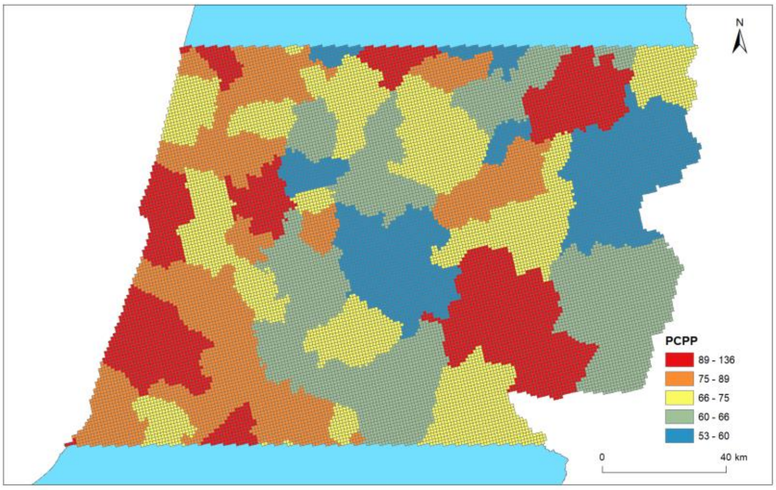
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



</td></tr>
</table>

Table B 50 – Explanatory variable: PCPP

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Per capita purchasing power by municipality

Variable name:
PCPP

Type of variable:
Continuous

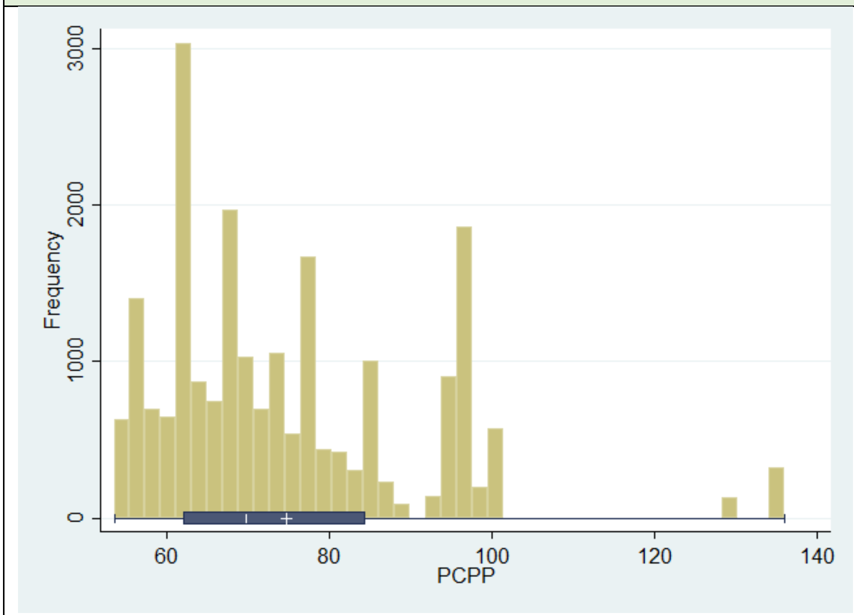
Range:
53 – 136

Average:
74.7

Standard deviation:
15.9

Effect:
Ignition

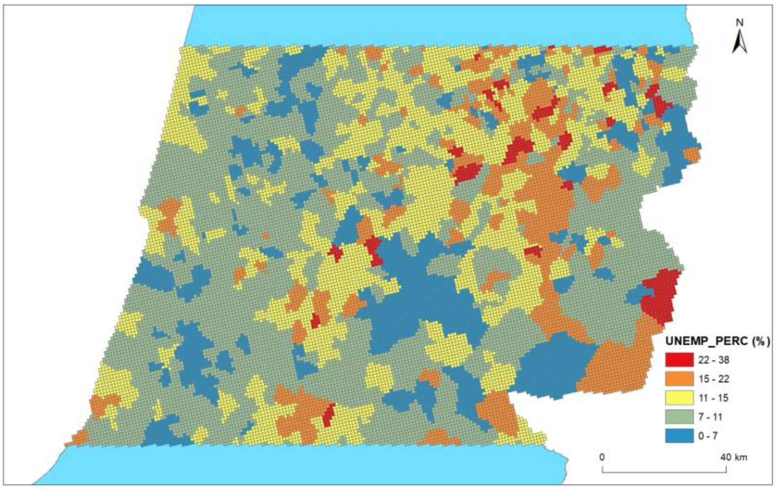
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



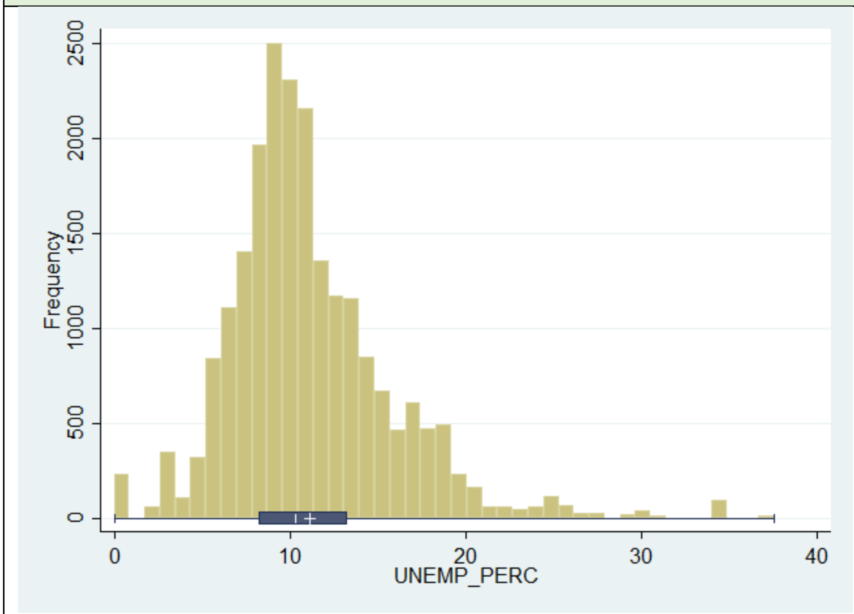
</td></tr>
</table>

Table B 51 – Explanatory variable: UNEMP_PERC

Map



Histogram



DESCRIPTION

Variable:
Unemployment rate (%)

Variable name:
UNEMP_PERC

Type of variable:
Continuous

Range:
0% – 38%

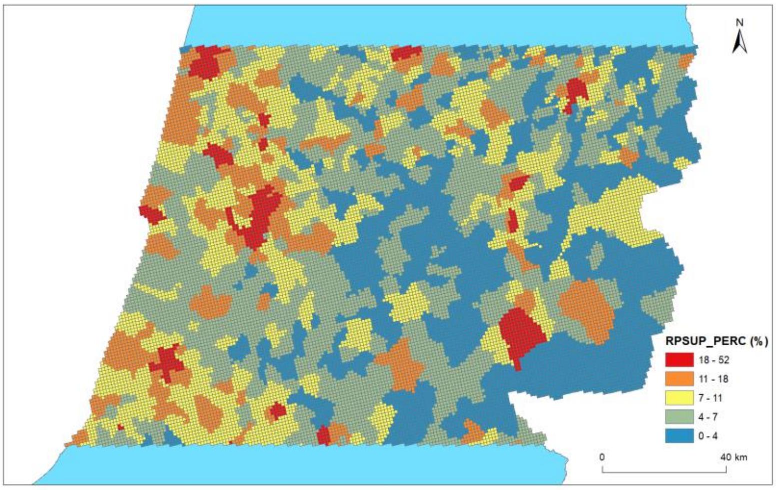
Average:
11.1%

Standard deviation:
4.7%

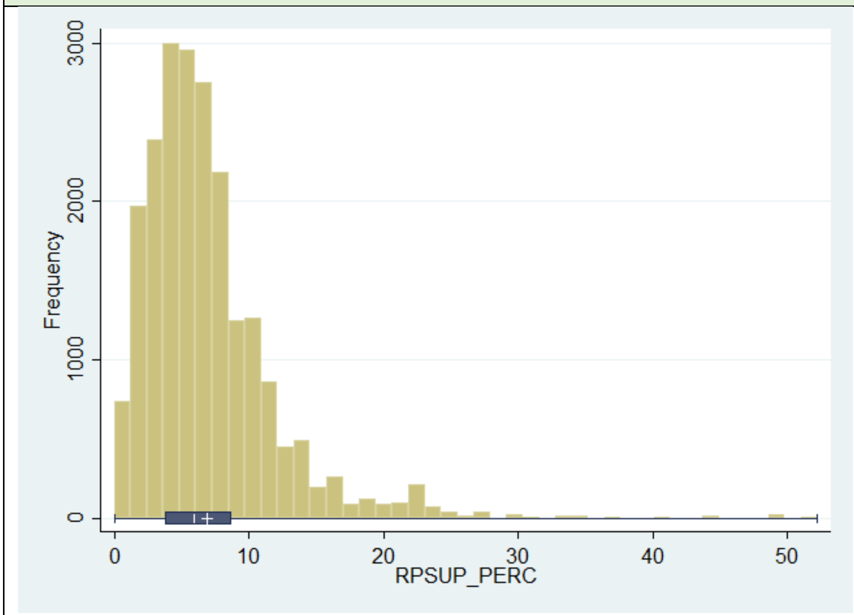
Effect:
Ignition

Table B 52 – Explanatory variable: RPSUP_PERC

Map



Histogram



DESCRIPTION

Variable:
Proportion of resident population with post-secondary education (%)

Variable name:
RPSUP_PERC

Type of variable:
Continuous

Range:
0% – 52%

Average:
6.9%

Standard deviation:
4.9%

Effect:
Ignition

Table B 53 – Explanatory variable: RPSEC_PERC

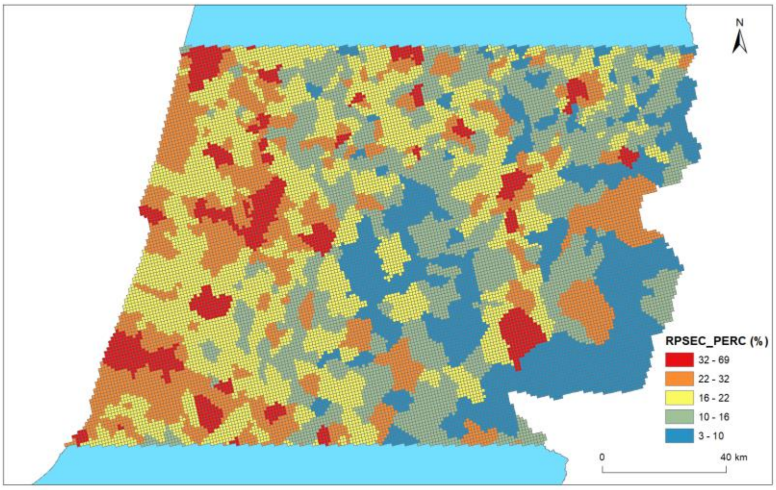
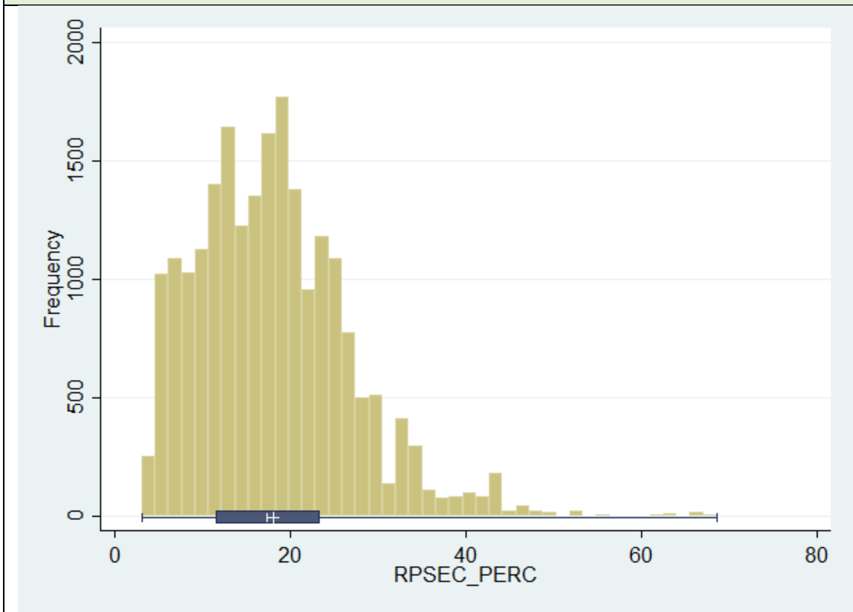
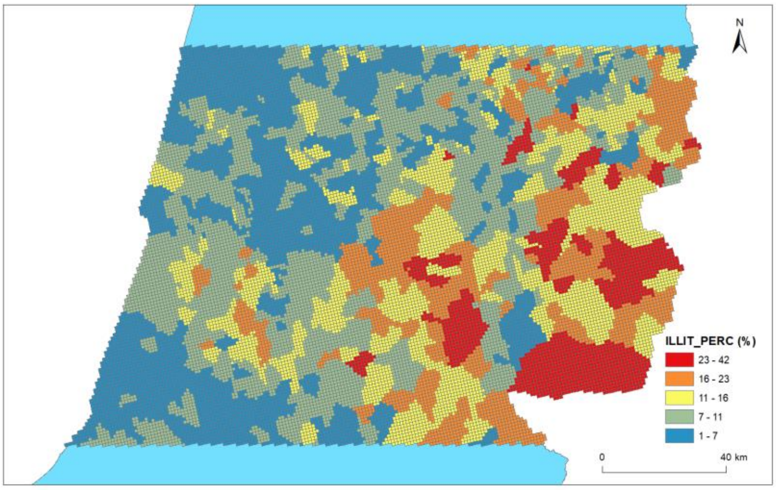
Map	**DESCRIPTION**
	**Variable:** Proportion of resident population with secondary education (%) **Variable name:** RPSEC_PERC **Type of variable:** Continuous **Range:** 3% – 69% **Average:** 18.1% **Standard deviation:** 8.8% **Effect:** Ignition
Histogram	
	

Table B 54 – Explanatory variable: ILLIT_PERC

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>



</td><td rowspan="3">

Variable:
Illiteracy rate (%)

Variable name:
ILLIT_PERC

Type of variable:
Continuous

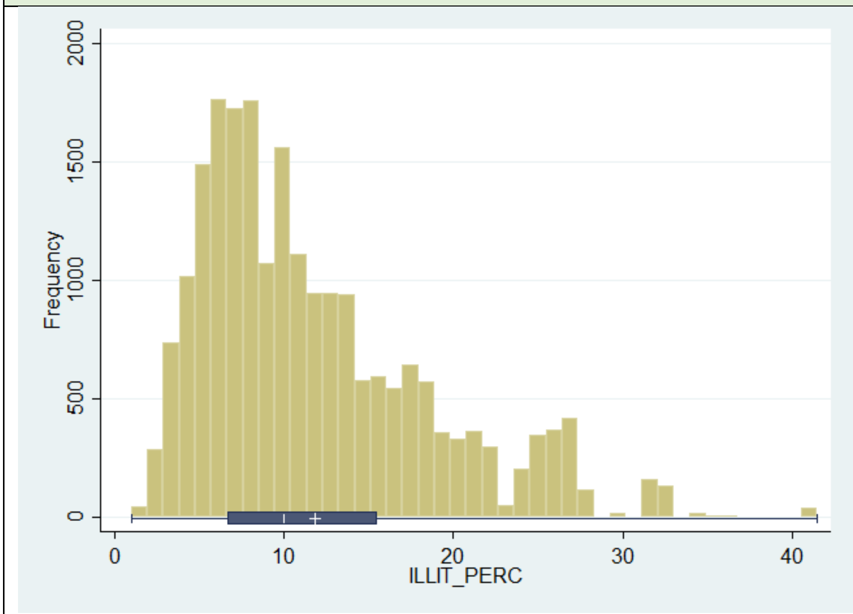
Range:
1% – 42%

Average:
11.8%

Standard deviation:
6.8%

Effect:
Ignition

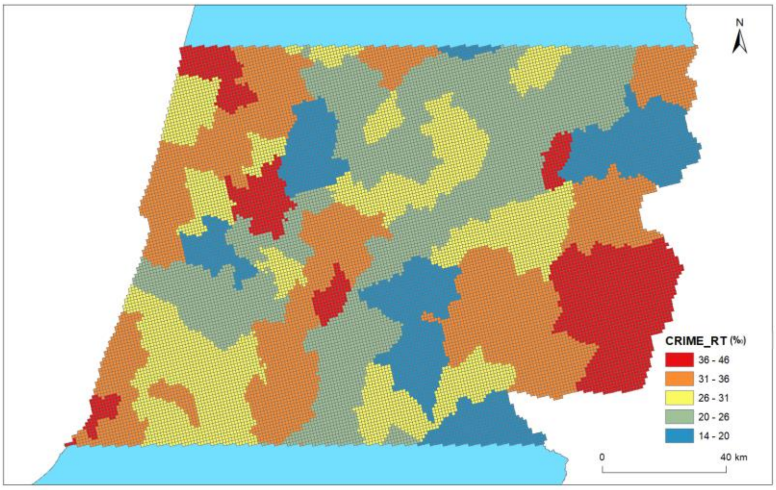
</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>



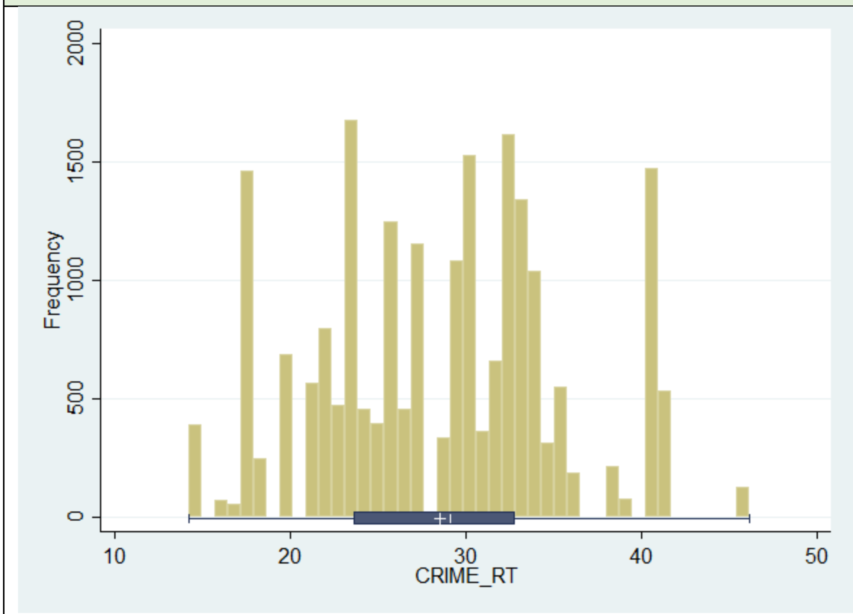
</td></tr>
</table>

Table B 55 – Explanatory variable: CRIME_RT

Map



Histogram



DESCRIPTION

Variable:
Criminality rate (‰)

Variable name:
CRIME_RT

Type of variable:
Continuous

Range:
14‰ – 46‰

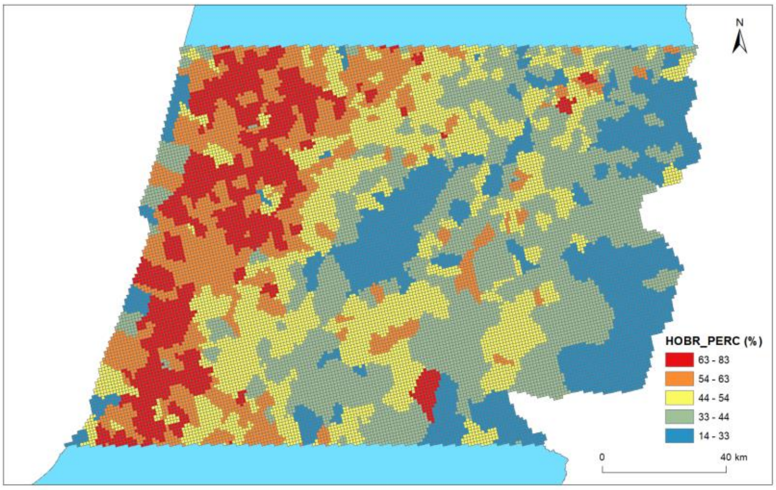
Average:
28.5‰

Standard deviation:
6.8‰

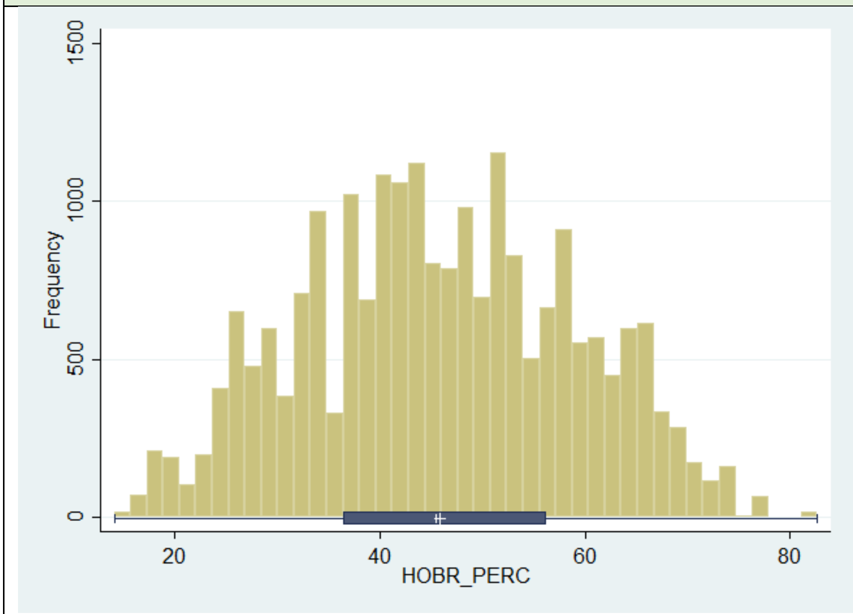
Effect:
Ignition

Table B 56 – Explanatory variable: HOBR_PERC

Map



Histogram



DESCRIPTION

Variable:
Proportion of housing owned by residents (%)

Variable name:
HOBR_PERC

Type of variable:
Continuous

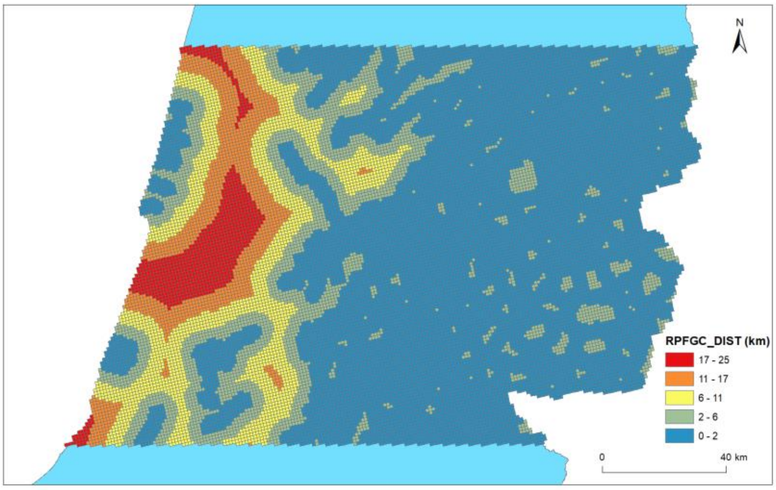
Range:
14% – 83%

Average:
45.9%

Standard deviation:
13.2%

Effect:
Spread

Table B 57 – Explanatory variable: RPFGC_DIST

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>
</td><td rowspan="3">Variable:
Distance to firelines (km)

Variable name:
RPFGC_DIST

Type of variable:
Continuous

Range:
0 km – 25 km

Average:
3.4 km

Standard deviation:
5.2 km

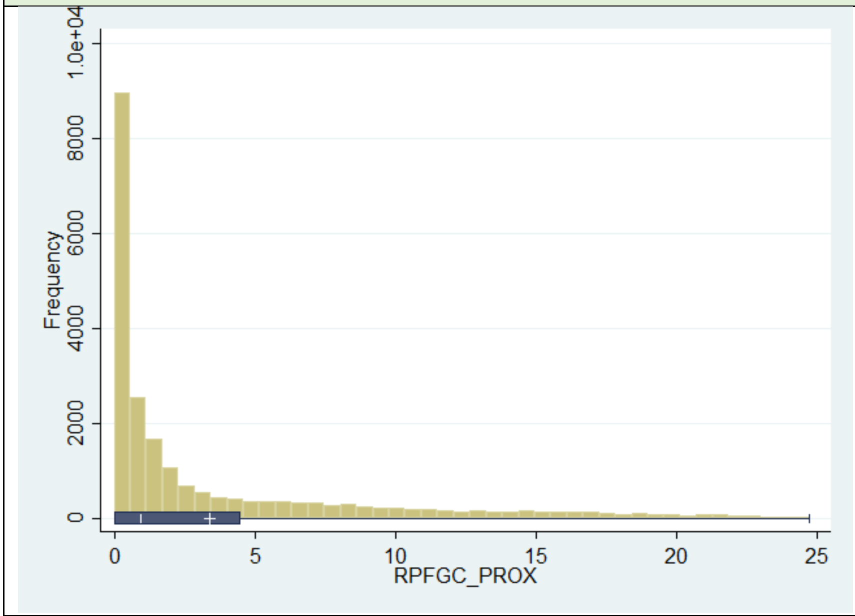
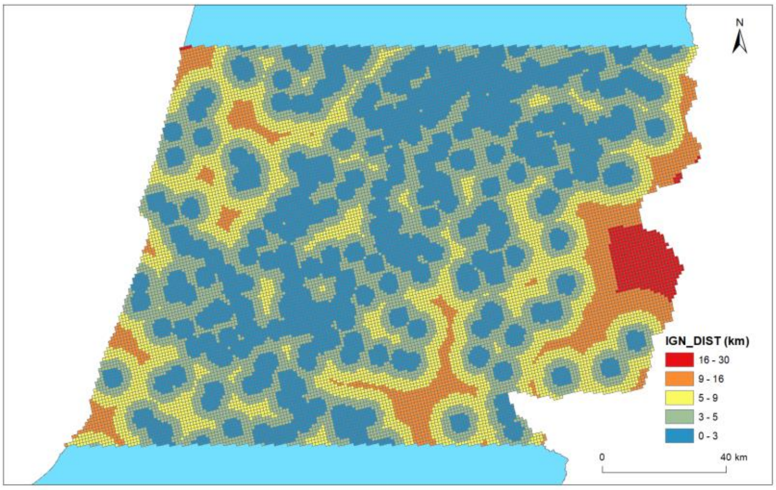
Effect:
Spread</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>
</td></tr>
</table>

Table B 58 – Explanatory variable: IGN_DIST

<table>
<tr><th>Map</th><th>DESCRIPTION</th></tr>
<tr><td>


</td>
<td rowspan="3">

Variable:
Distance to ignition points (km)

Variable name:
IGN_DIST

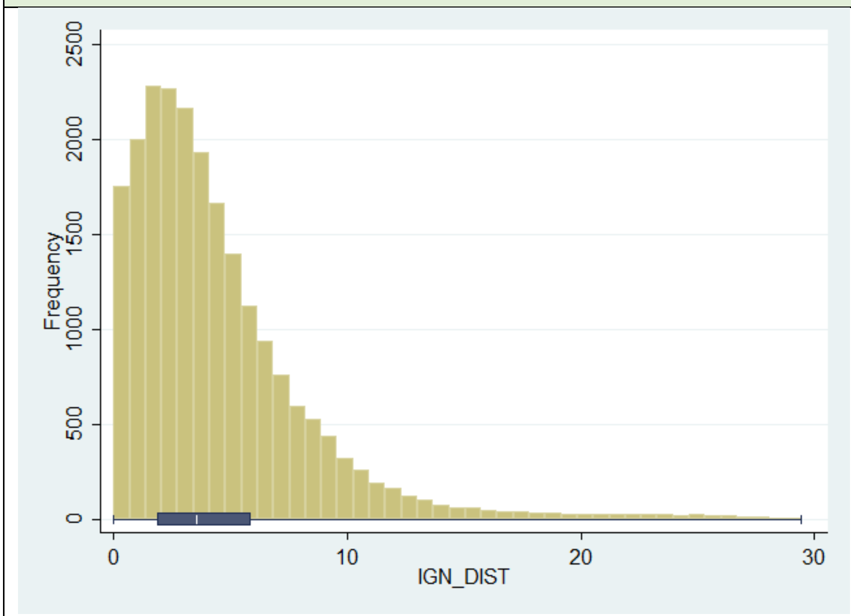
Type of variable:
Continuous

Range:
0 km – 30 km

Average:
4.5 km

Standard deviation:
3.9 km

Effect:
Spread

</td></tr>
<tr><td>Histogram</td></tr>
<tr><td>


</td></tr>
</table>

ANNEX C

ROC curve plots and classification tables: large wildfire ignition and propagation models

Figure C 1 – Large wildfire ignition model: Global

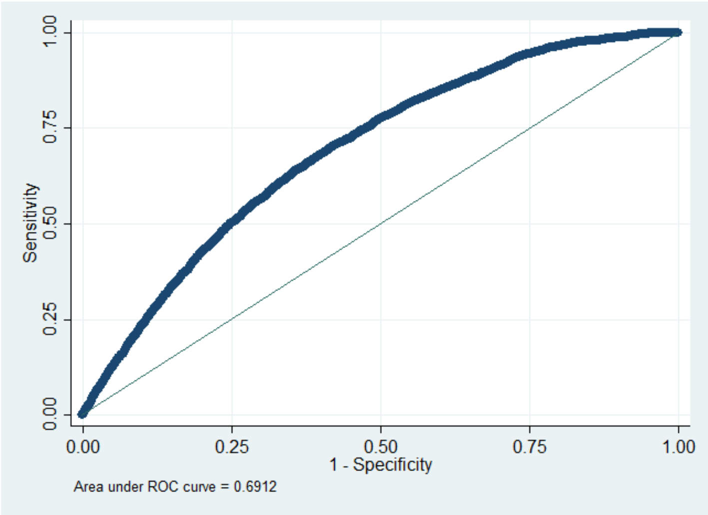


Table C 1 - Large wildfire ignition model: Global

	True		
Classified	D	~D	Total
+	2	5	7
-	2808	18755	21563
Total	2810	18760	21570

Figure C 2 – Large wildfire ignition model: Cluster 1

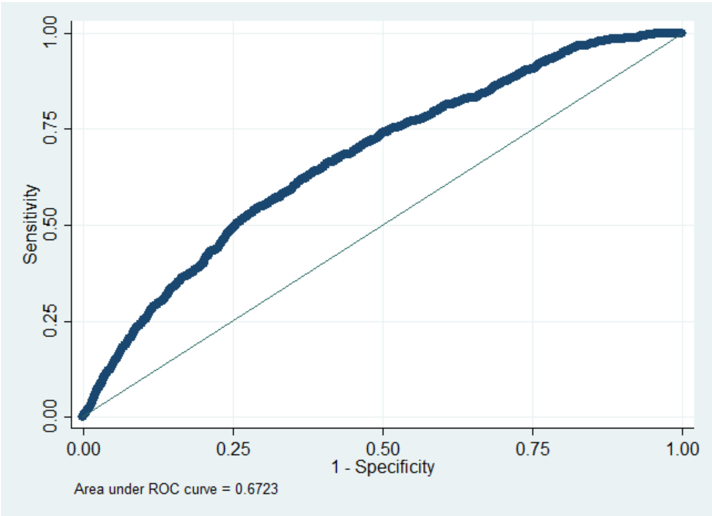


Table C 2 – Large wildfire ignition model: Cluster 1

	True		
Classified	D	~D	Total
+	0	0	0
-	931	7002	7933
Total	931	7002	7933

Figure C 3 – Large wildfire ignition model: Cluster 2

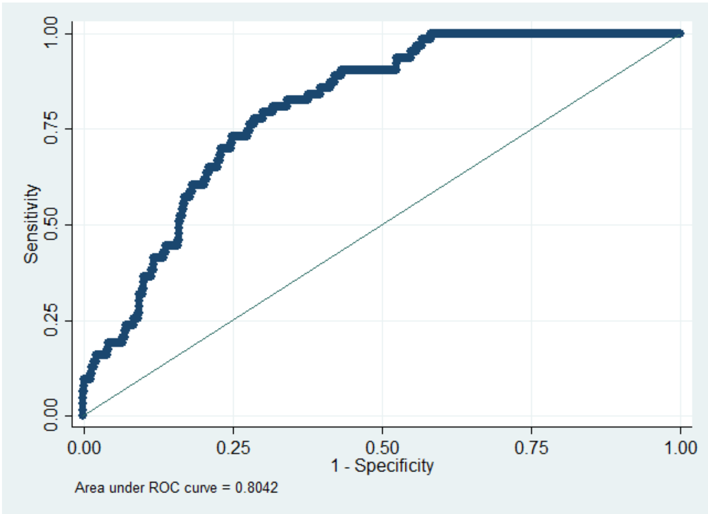


Table C 3 – Large wildfire ignition model: Cluster 2

	True		
Classified	D	~D	Total
+	0	0	0
-	63	2295	2358
Total	63	2295	2358

Figure C 4 – Large wildfire ignition model: Cluster 3

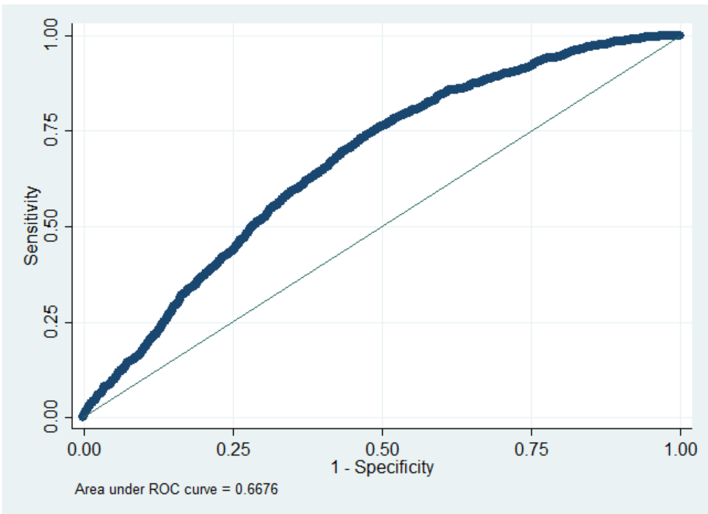


Table C 4 – Large wildfire ignition model: Cluster 3

	True		
Classified	D	~D	Total
+	18	19	37
-	1252	5322	6574
Total	1270	5341	6611

Figure C 5 – Large wildfire ignition model: Cluster 4

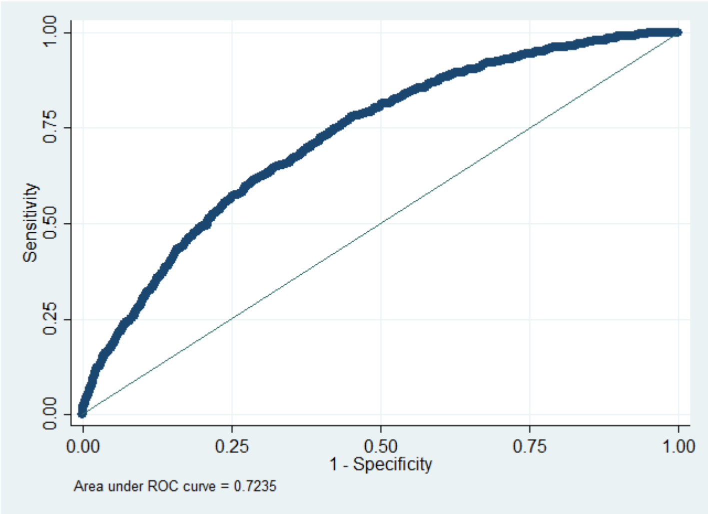


Table C 5 – Large wildfire ignition model: Cluster 4

	True		
Classified	D	~D	Total
+	12	7	19
-	534	4115	4649
Total	546	4122	4668

Figure C 6 – Large wildfire propagation model: Global

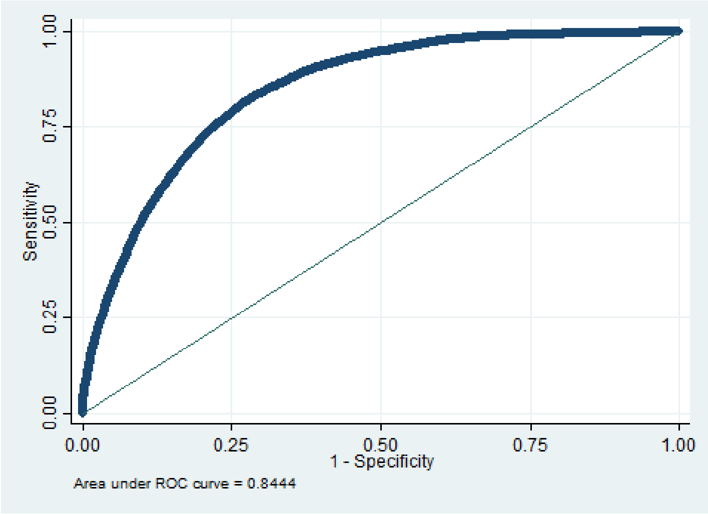


Table C 6 – Large wildfire propagation model: Global

	True		
Classified	D	~D	Total
+	2275	1234	3509
-	3068	14993	18061
Total	5343	16227	21570

Figure C 7 – Large wildfire propagation model: Cluster 1

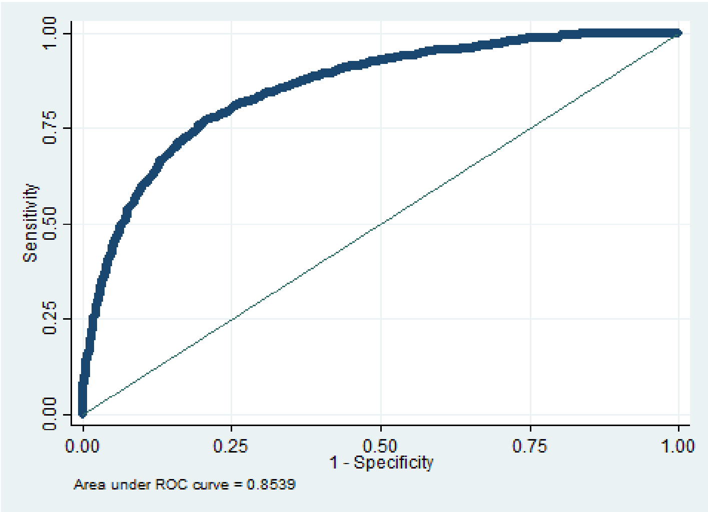


Table C 7 – Large wildfire propagation model: Cluster 1

	True		
Classified	D	~D	Total
+	438	187	625
-	479	2795	3274
Total	917	2982	3899

Figure C 8 – Large wildfire propagation model: Cluster 2

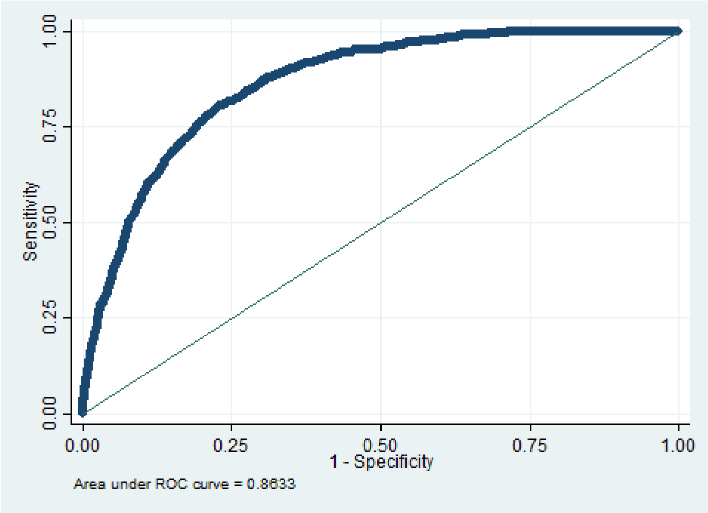


Table C 8 – Large wildfire propagation model: Cluster 2

	True		
Classified	D	~D	Total
+	448	251	699
-	691	4156	4847
Total	1139	4407	5546

Figure C 9 – Large wildfire propagation model: Cluster 3

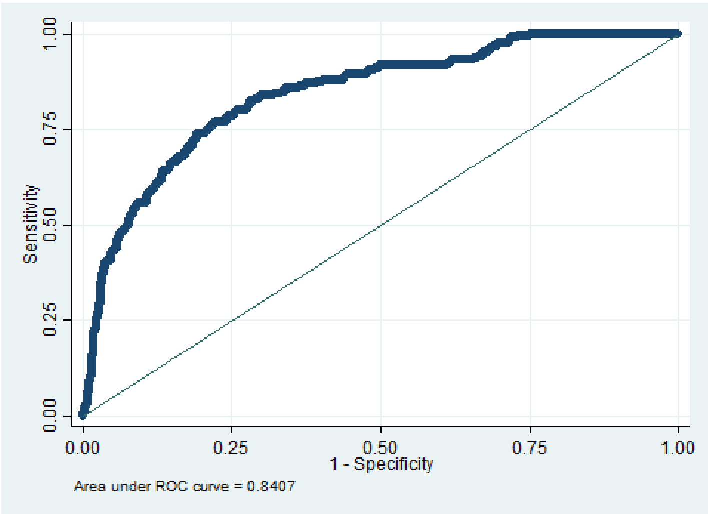


Table C 9 – Large wildfire propagation model: Cluster 3

	True		
Classified	D	~D	Total
+	4	8	12
-	188	2526	2714
Total	192	2534	2726

Figure C 10 – Large wildfire propagation model: Cluster 4

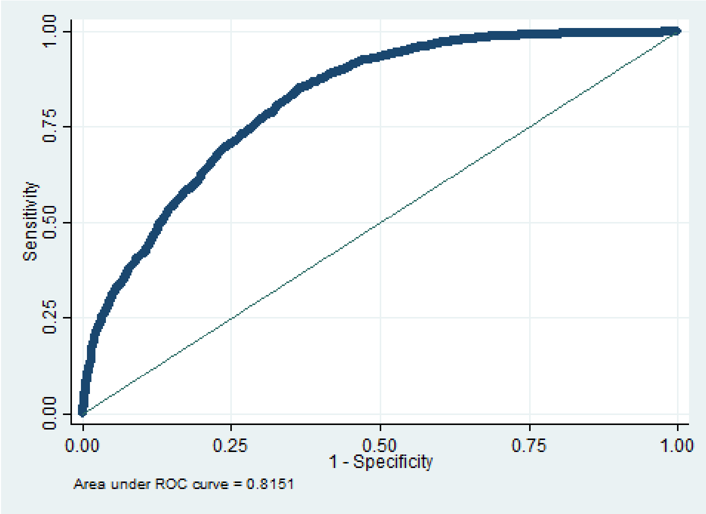


Table C 10 – Large wildfire propagation model: Cluster 4

	True		
Classified	D	~D	Total
+	1279	601	1880
-	715	2258	2973
Total	1994	2859	4853

Figure C 11 – Large wildfire propagation model: Cluster 5

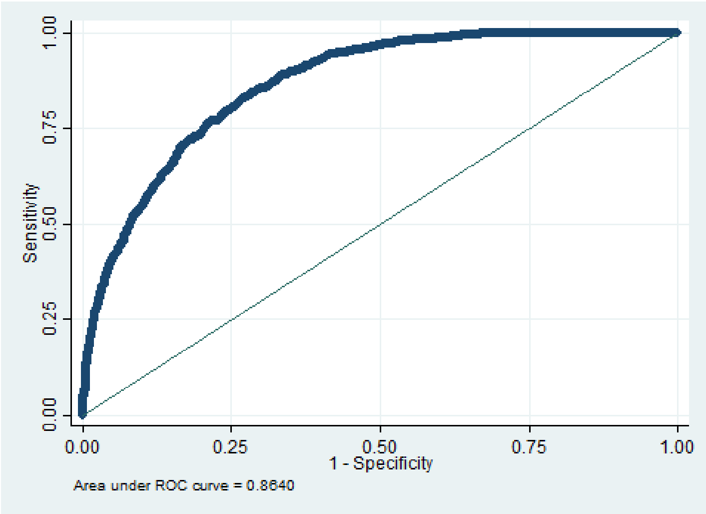


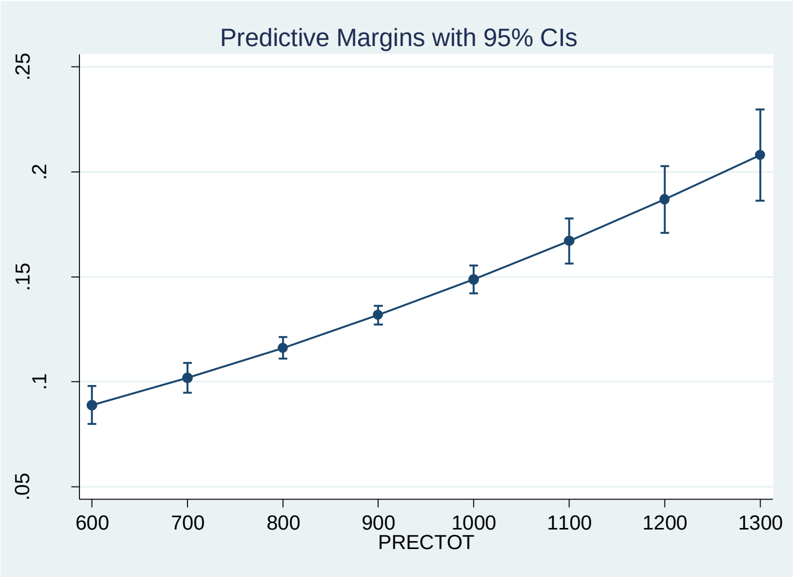
Table C 11 – Large wildfire propagation model: Cluster 5

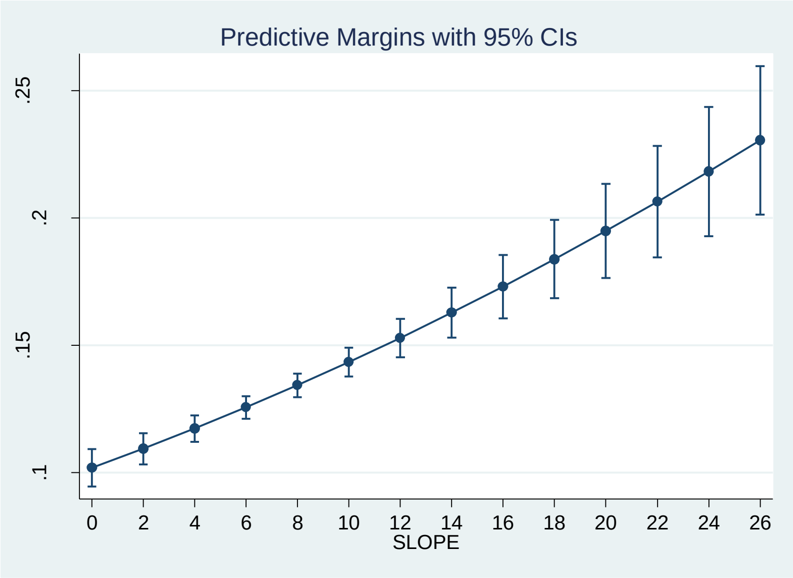
	True		
Classified	D	~D	Total
+	560	289	849
-	541	3156	3697
Total	1101	3445	4546

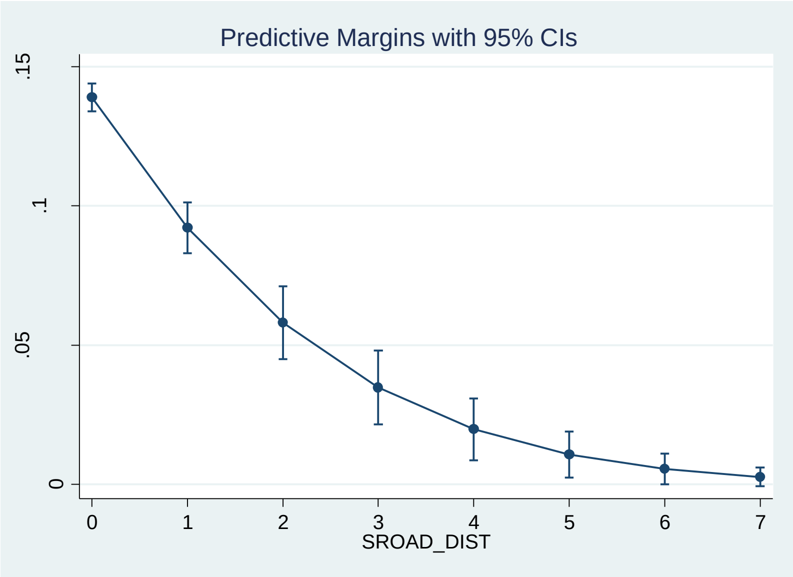
ANNEX D

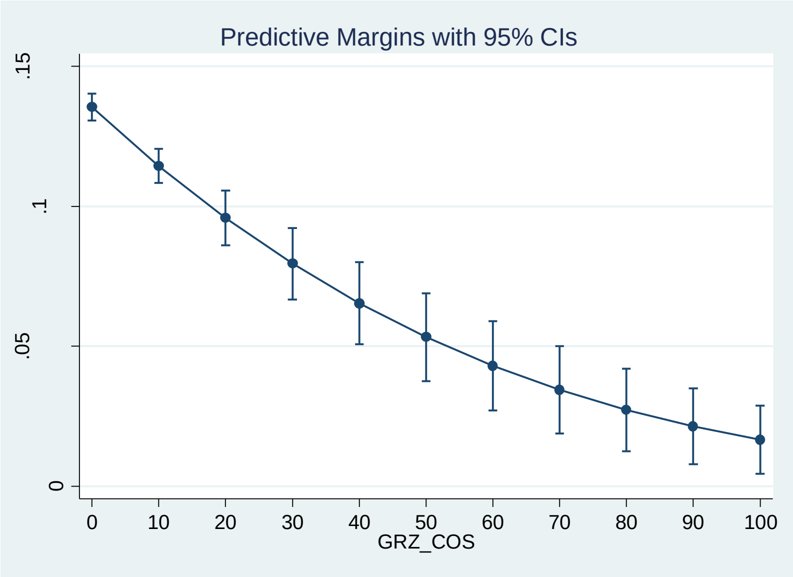
Individual variable plots: large wildfire ignition and propagation models

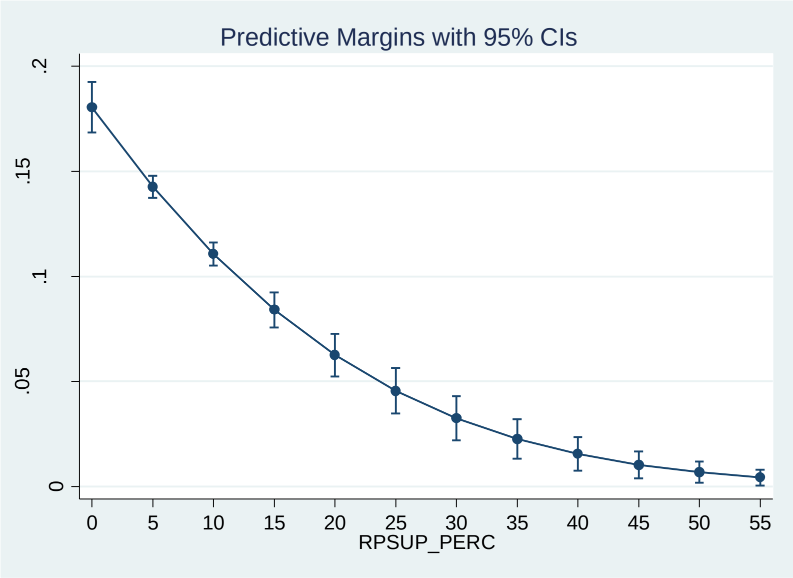
Figure D 1 – Large wildfire ignition model: Global

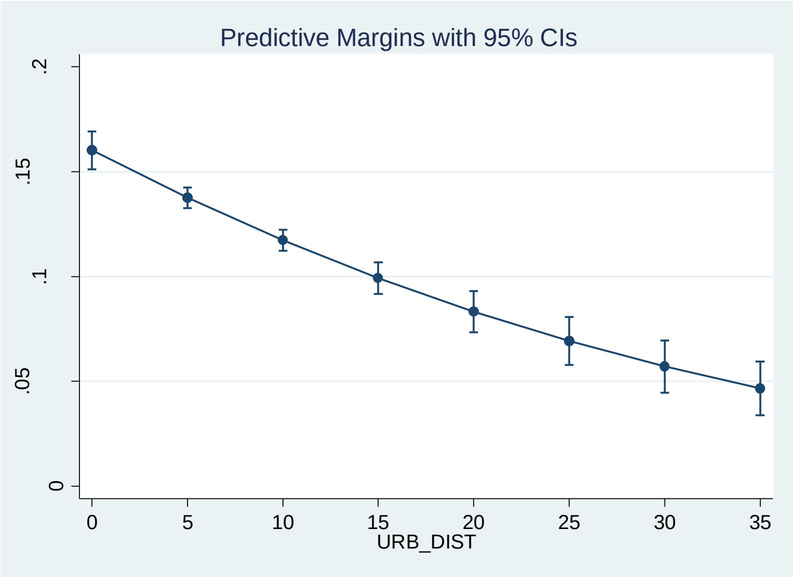


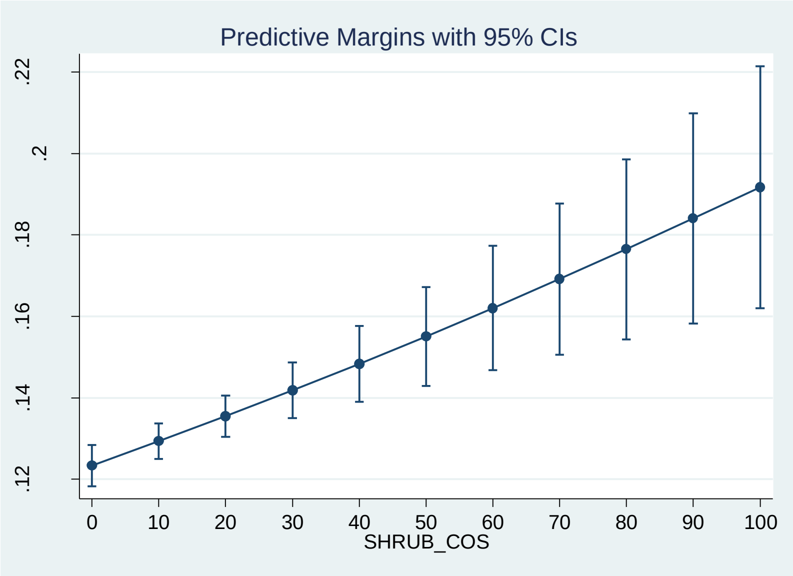

Predictive Margins with 95% CIs
.1
.15
.2
.25
0
2
4
6
8
10
12
14
16
18
20
22
24
26
SLOPE

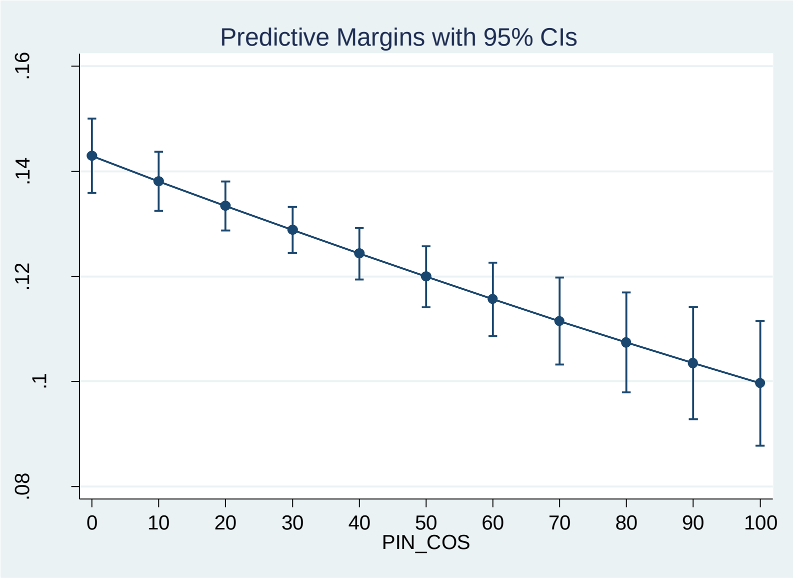

Predictive Margins with 95% CIs
0
.05
.1
.15
0
1
2
3
4
5
6
7
SROAD_DIST

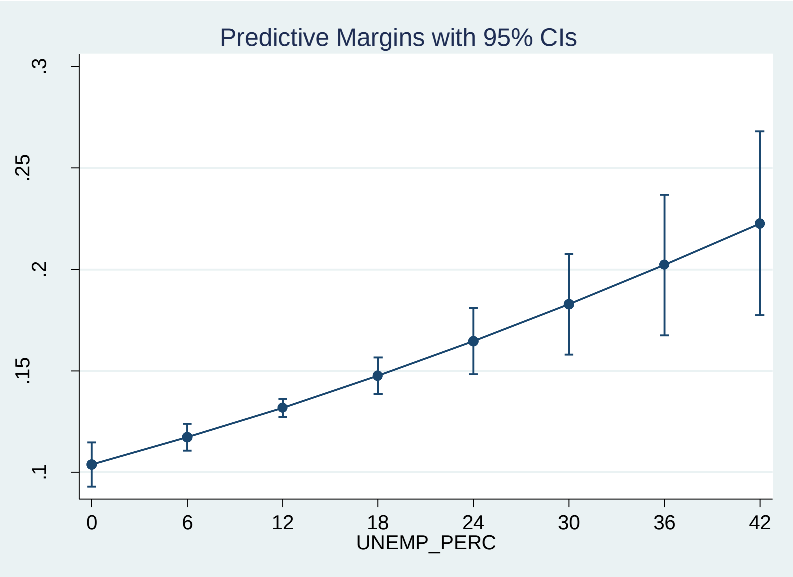

Predictive Margins with 95% CIs
0
.05
.1
.15
0
10
20
30
40
50
60
70
80
90
100
GRZ_COS

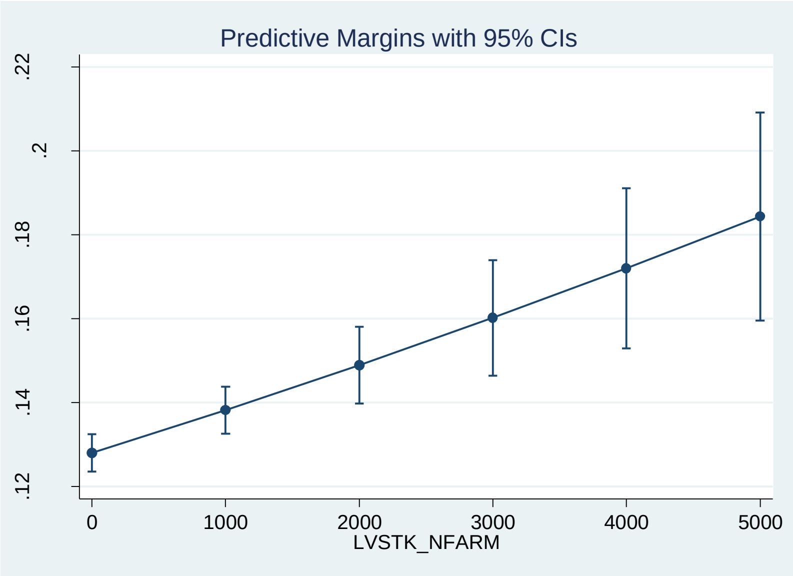

Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0 5 10 15 20 25 30 35 40 45 50 55
RPSUP_PERC

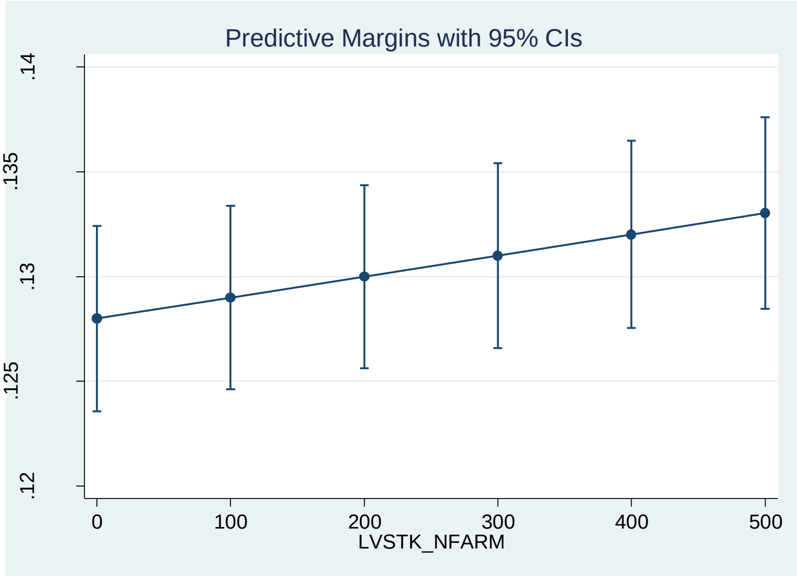

Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0 5 10 15 20 25 30 35
URB_DIST

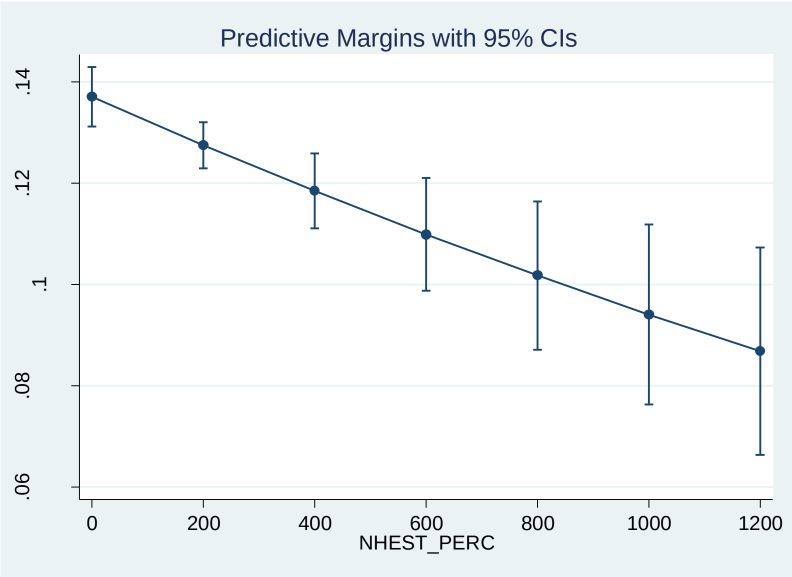

Predictive Margins with 95% CIs
.12
.14
.16
.18
.2
.22
0 10 20 30 40 50 60 70 80 90 100
SHRUB_COS

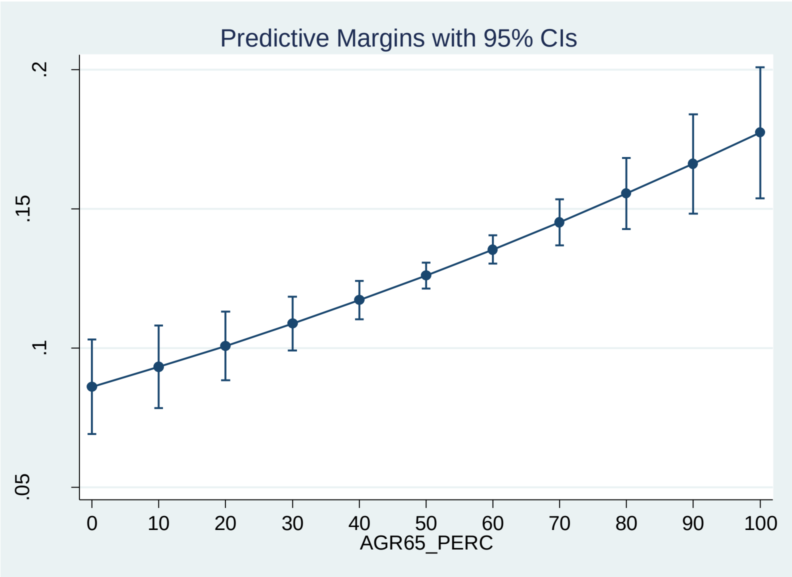

Predictive Margins with 95% CIs
.08
.1
.12
.14
.16
0 10 20 30 40 50 60 70 80 90 100
PIN_COS

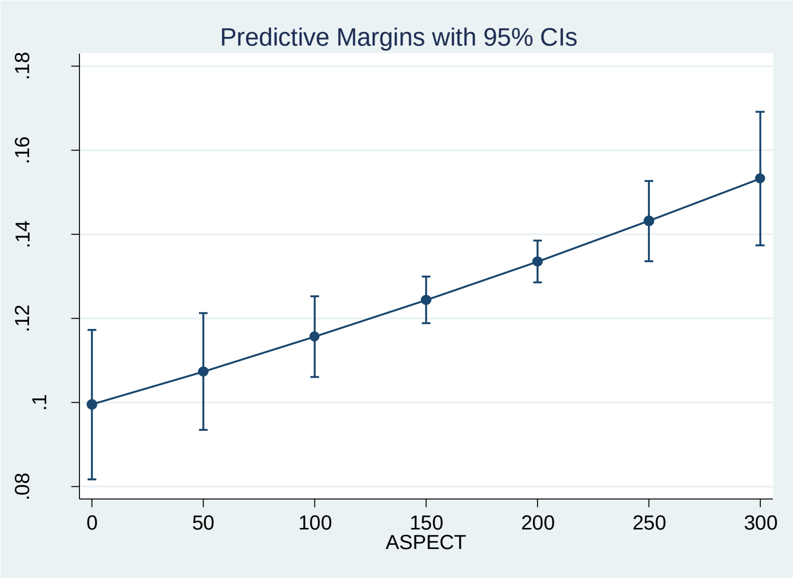

Predictive Margins with 95% CIs
.1
.15
.2
.25
.3
0 6 12 18 24 30 36 42
UNEMP_PERC

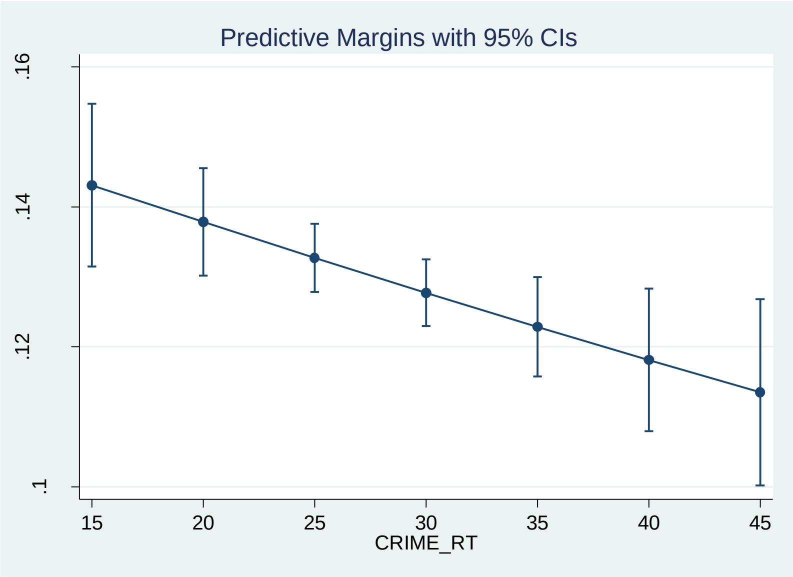

Predictive Margins with 95% CIs
.12
.14
.16
.18
.2
.22
0 1000 2000 3000 4000 5000
LVSTK_NFARM

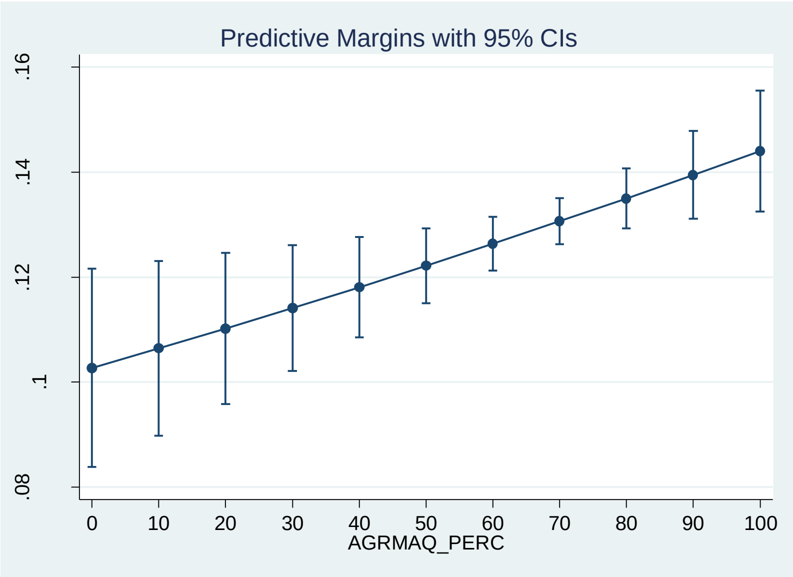

Predictive Margins with 95% CIs
.12
.125
.13
.135
.14
0
100
200
300
400
500
LVSTK_NFARM

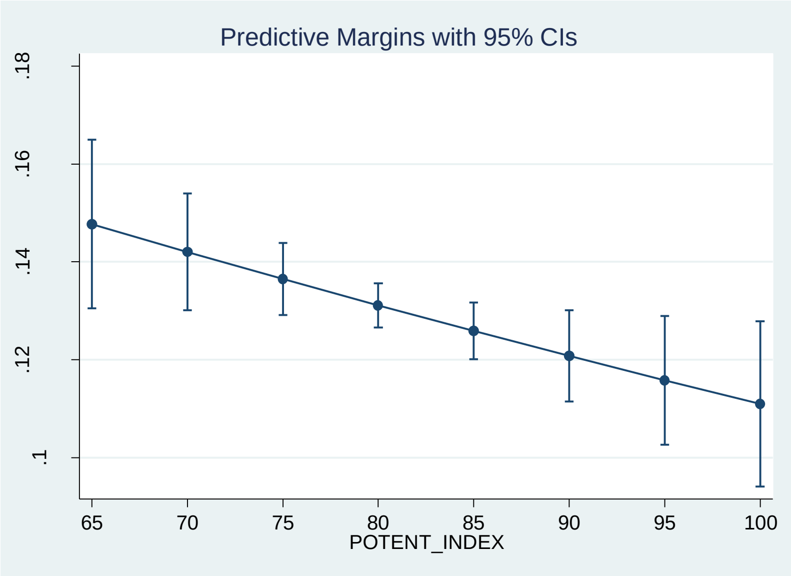

Predictive Margins with 95% CIs
.06
.08
.1
.12
.14
0
200
400
600
800
1000
1200
NHEST_PERC


Predictive Margins with 95% CIs
.05
.1
.15
.2
0
10
20
30
40
50
60
70
80
90
100
AGR65_PERC


Predictive Margins with 95% CIs
.08
.1
.12
.14
.16
.18
0
50
100
150
200
250
300
ASPECT


Predictive Margins with 95% CIs
.1
.12
.14
.16
15
20
25
30
35
40
45
CRIME_RT


Predictive Margins with 95% CIs
.08
.1
.12
.14
.16
0
10
20
30
40
50
60
70
80
90
100
AGRMAQ_PERC


Predictive Margins with 95% CIs
.1
.12
.14
.16
.18
65
70
75
80
85
90
95
100
POTENT_INDEX

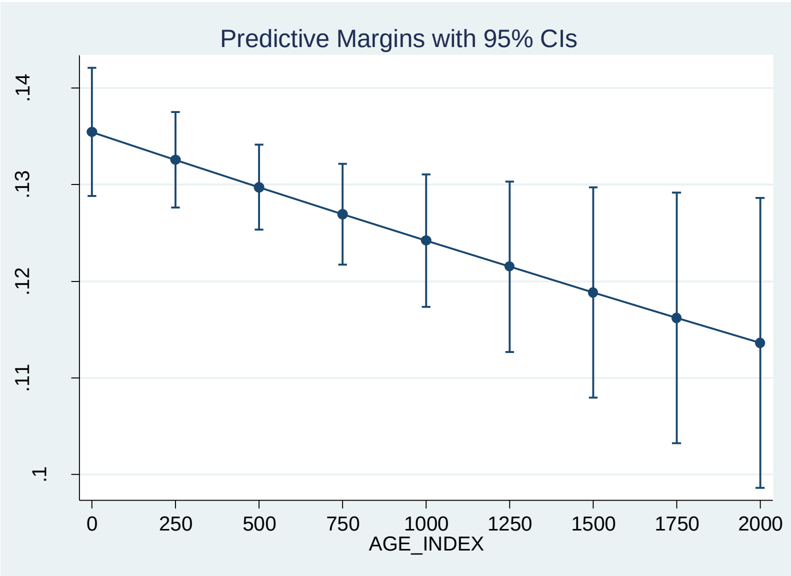
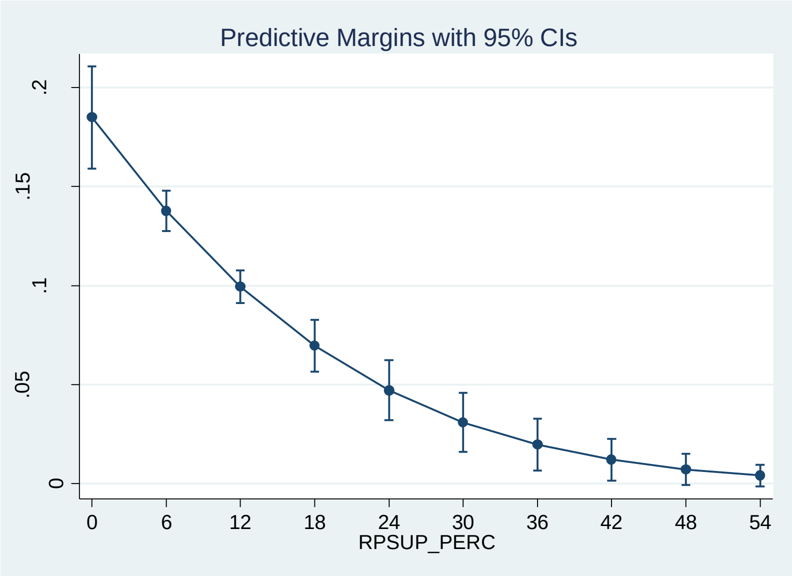
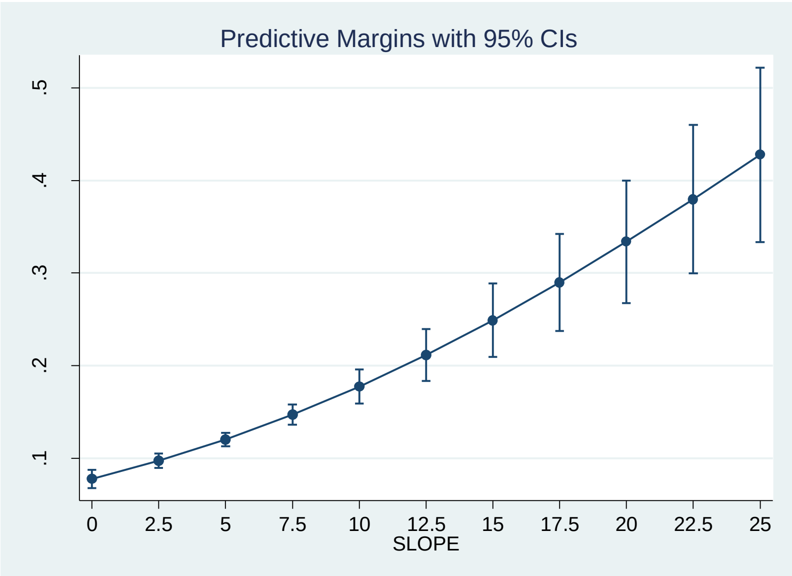
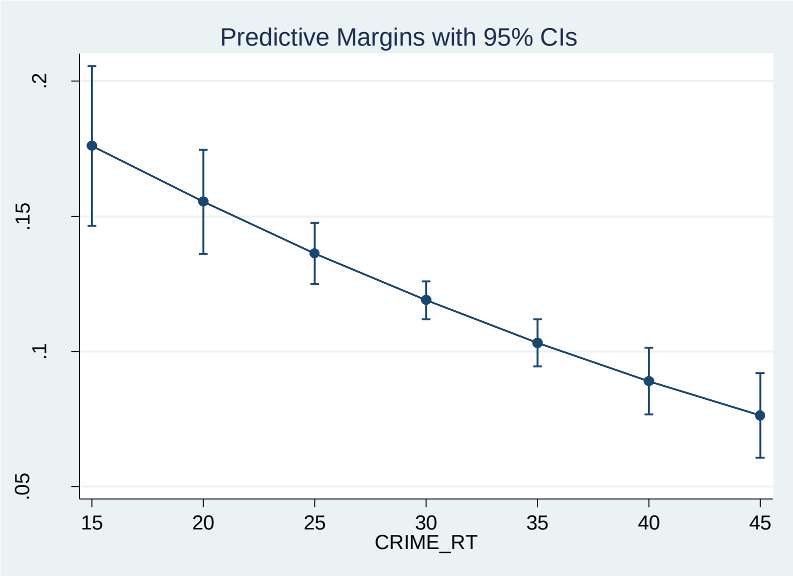
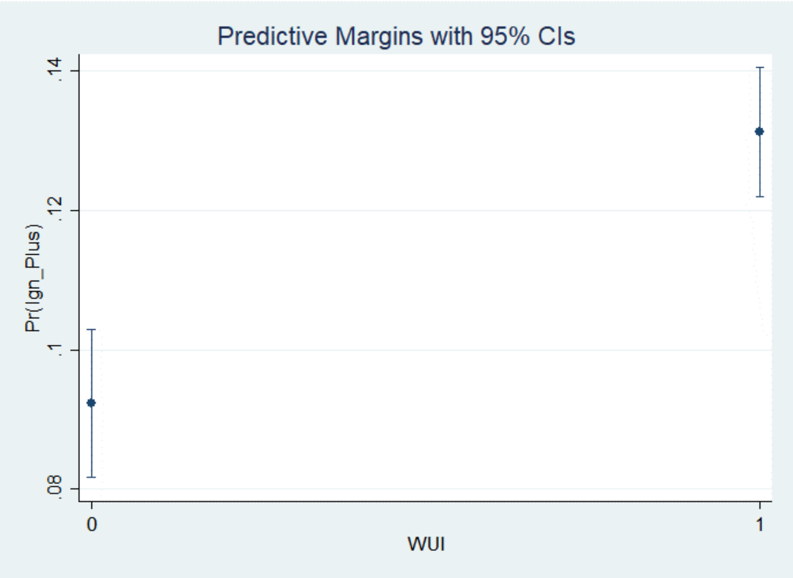

Predictive Margins with 95% CIs
.1
.11
.12
.13
.14
0
250
500
750
1000
1250
1500
1750
2000
AGE_INDEX

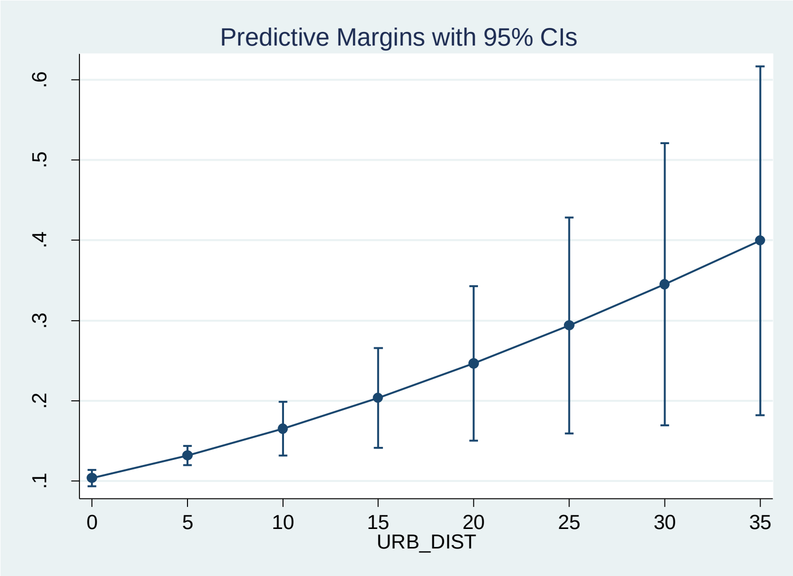
Figure D 2 – Large wildfire ignition model: Cluster 1

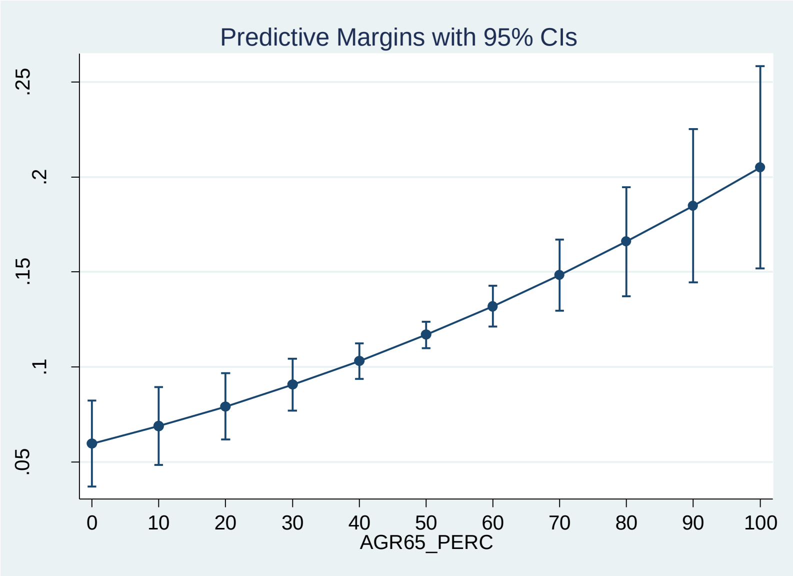


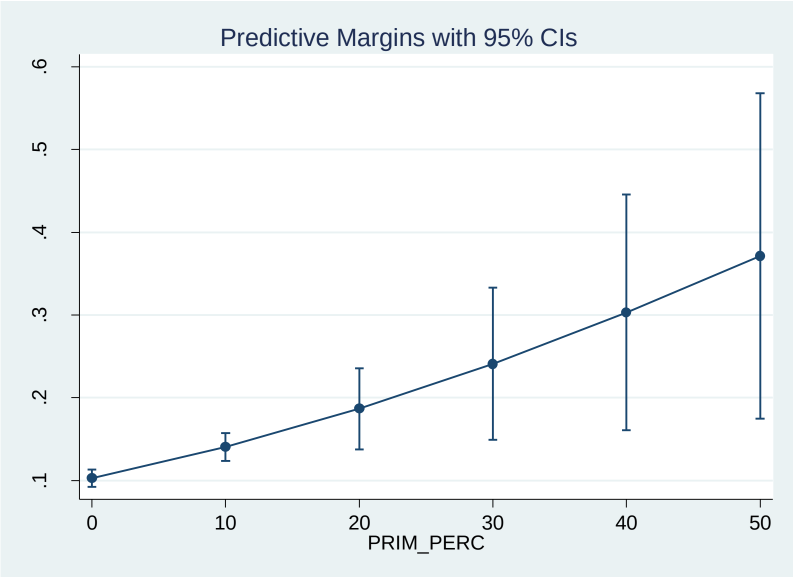


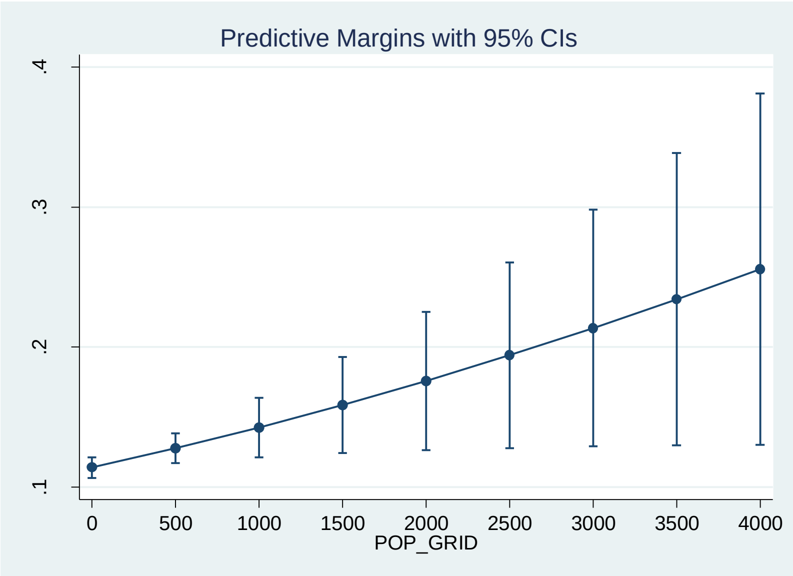

Predictive Margins with 95% CIs
.05
.1
.15
.2
15
20
25
30
35
40
45
CRIME_RT

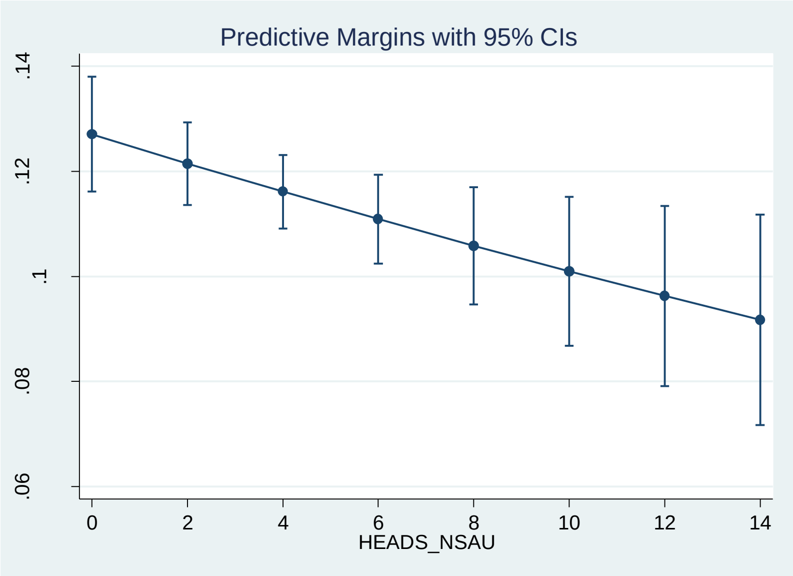

Predictive Margins with 95% CIs
Pr(Ign_Plus)
.08
.1
.12
.14
0
1
WUI


Predictive Margins with 95% CIs
.1
.2
.3
.4
.5
.6
0
5
10
15
20
25
30
35
URB_DIST

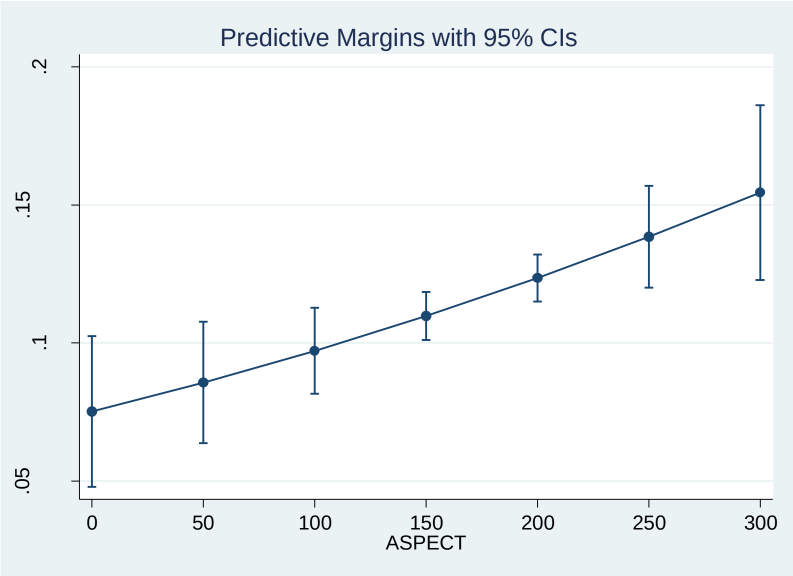

Predictive Margins with 95% CIs
.05
.1
.15
.2
.25
0 10 20 30 40 50 60 70 80 90 100
AGR65_PERC


Predictive Margins with 95% CIs
.1
.2
.3
.4
.5
.6
0 10 20 30 40 50
PRIM_PERC

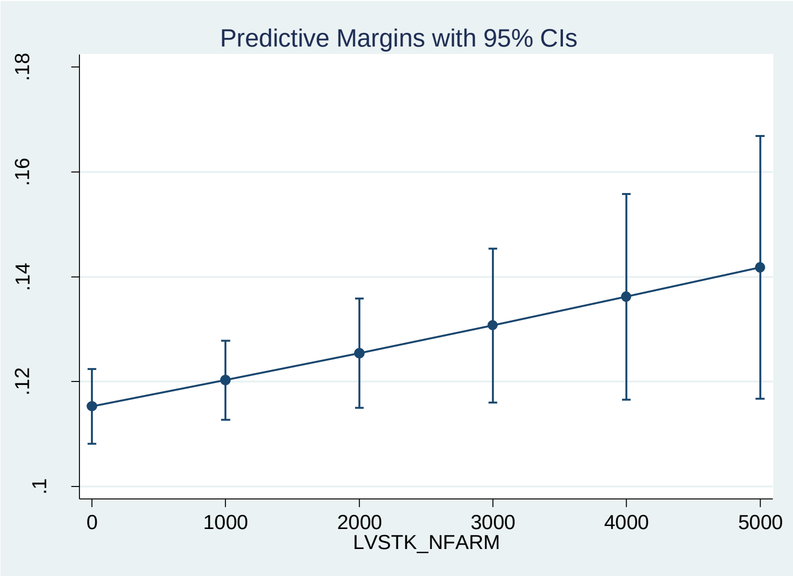

Predictive Margins with 95% CIs
.1
.2
.3
.4
0 500 1000 1500 2000 2500 3000 3500 4000
POP_GRID



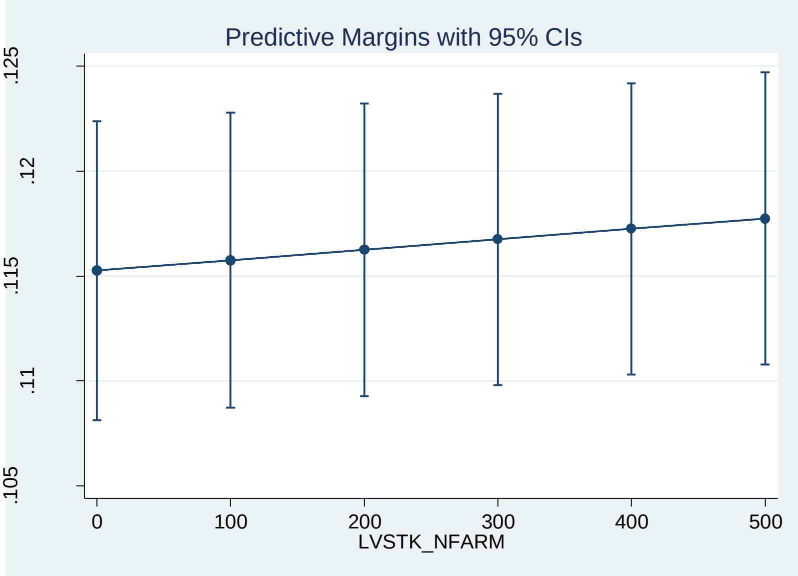
Predictive Margins with 95% CIs
.06
.08
.1
.12
.14
0
2
4
6
8
10
12
14
HEADS_NSAU

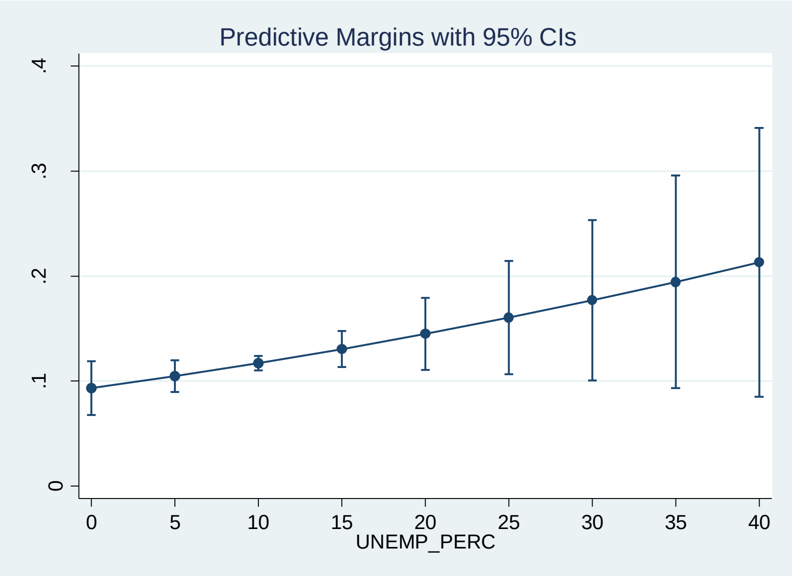


Predictive Margins with 95% CIs
.05
.1
.15
.2
0
50
100
150
200
250
300
ASPECT



Predictive Margins with 95% CIs
.1
.12
.14
.16
.18
0
1000
2000
3000
4000
5000
LVSTK_NFARM


Predictive Margins with 95% CIs
.105
.11
.115
.12
.125
0
100
200
300
400
500
LVSTK_NFARM


Predictive Margins with 95% CIs
0
.1
.2
.3
.4
0
5
10
15
20
25
30
35
40
UNEMP_PERC

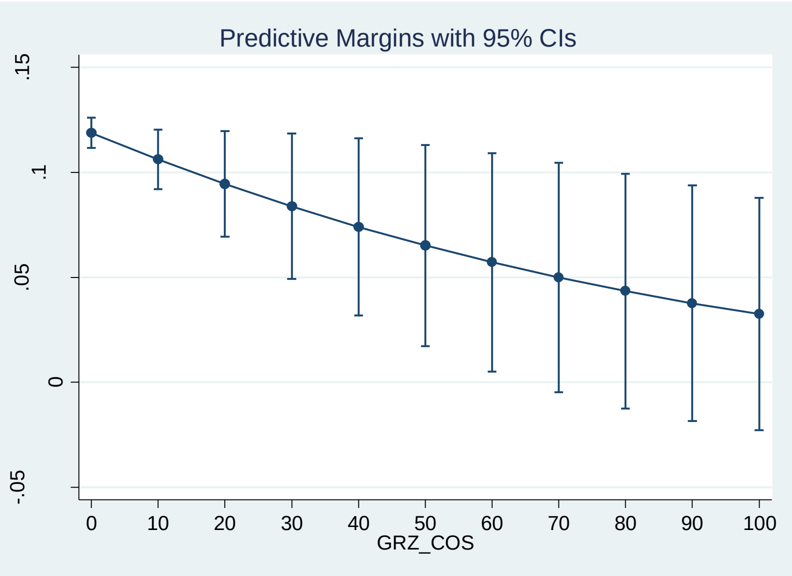
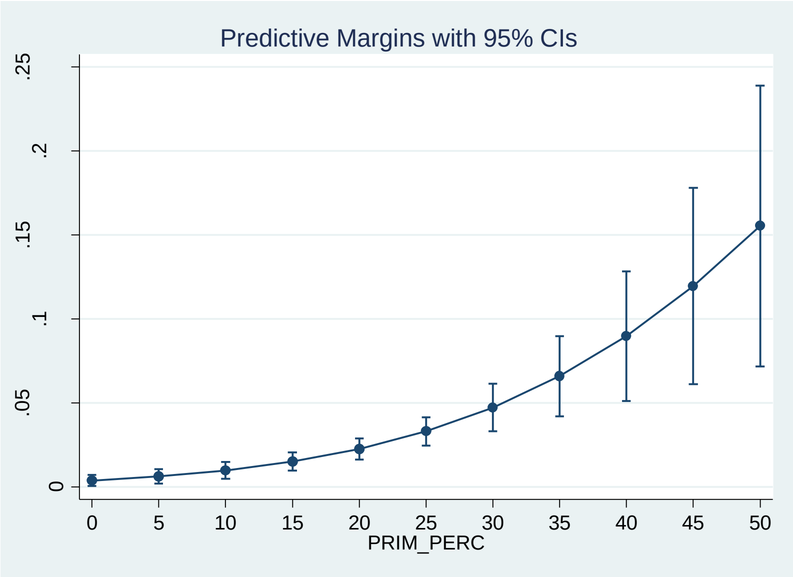
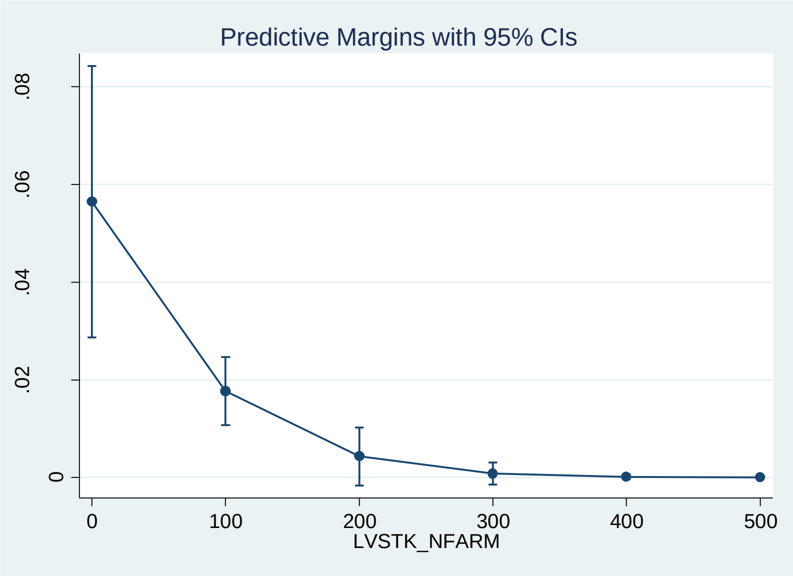
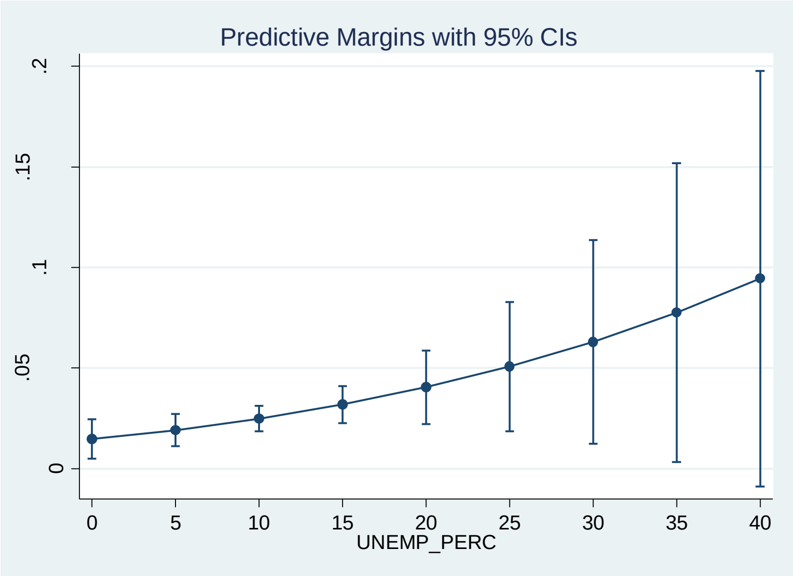

Predictive Margins with 95% CIs
-.05
0
.05
.1
.15
0
10
20
30
40
50
60
70
80
90
100
GRZ_COS

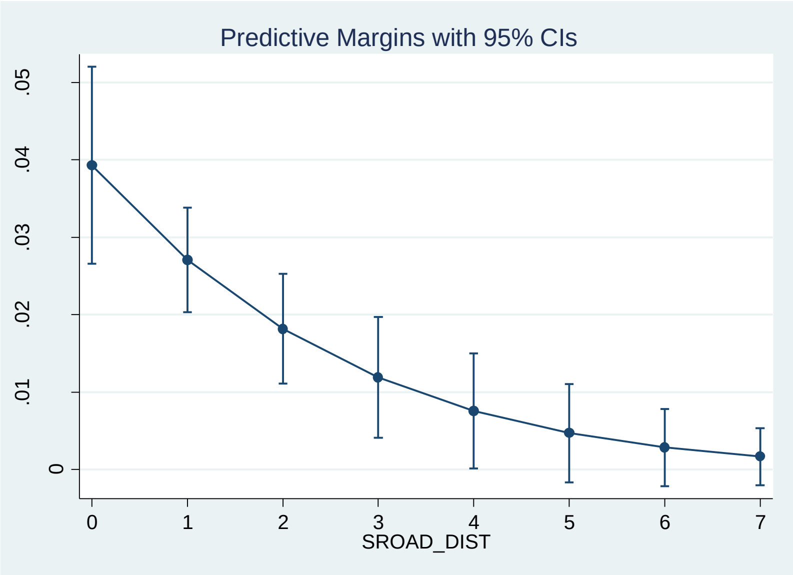
Figure D 3 – Large wildfire ignition model: Cluster 2



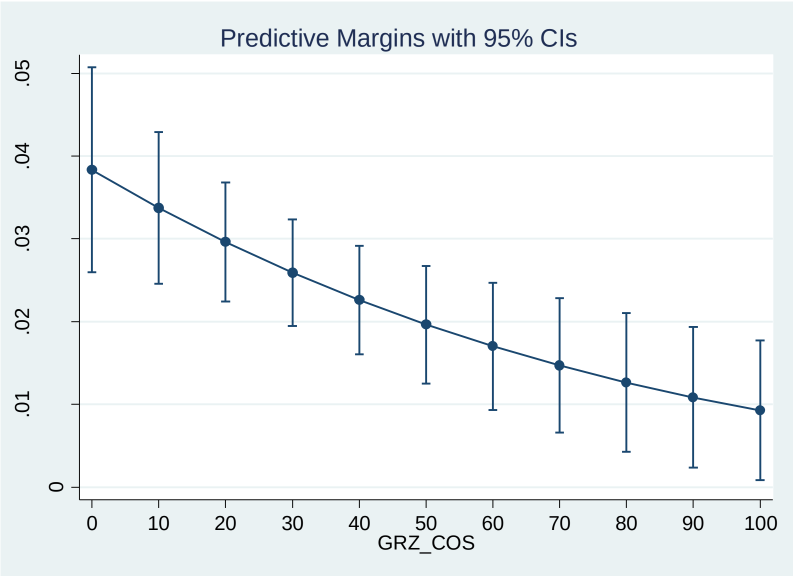




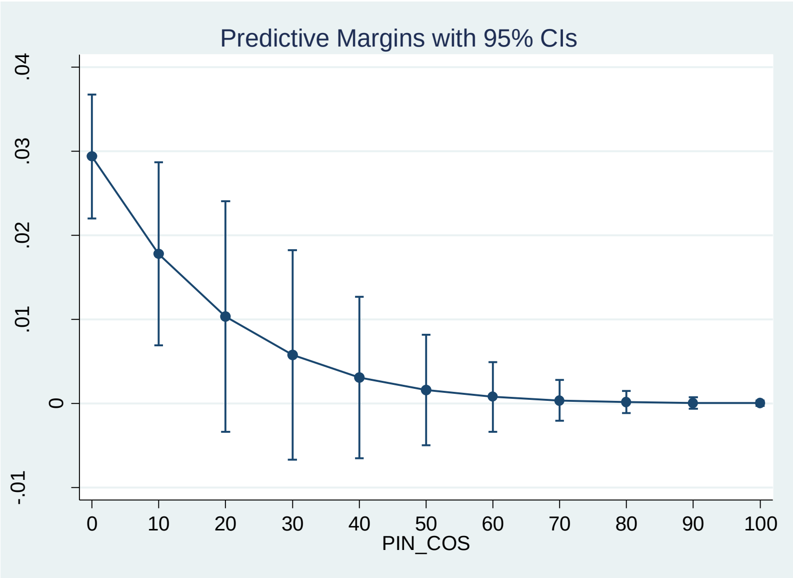
Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0
5
10
15
20
25
30
35
40
UNEMP_PERC

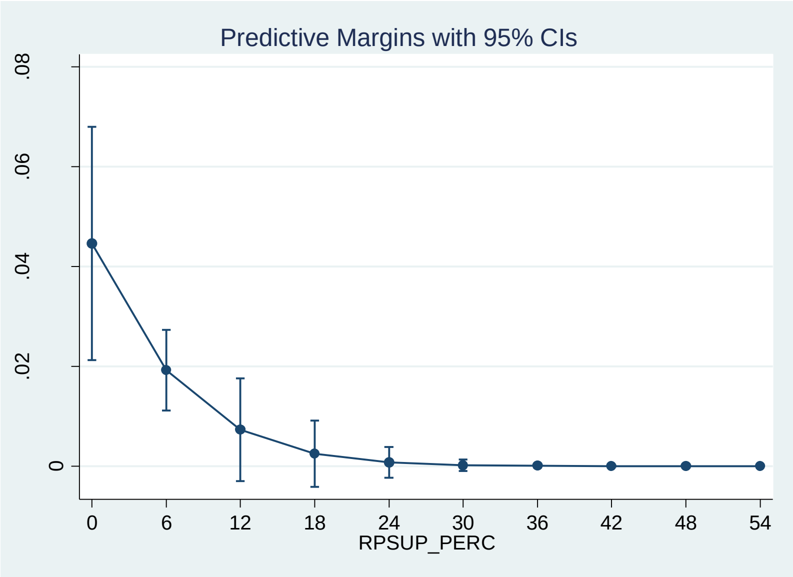


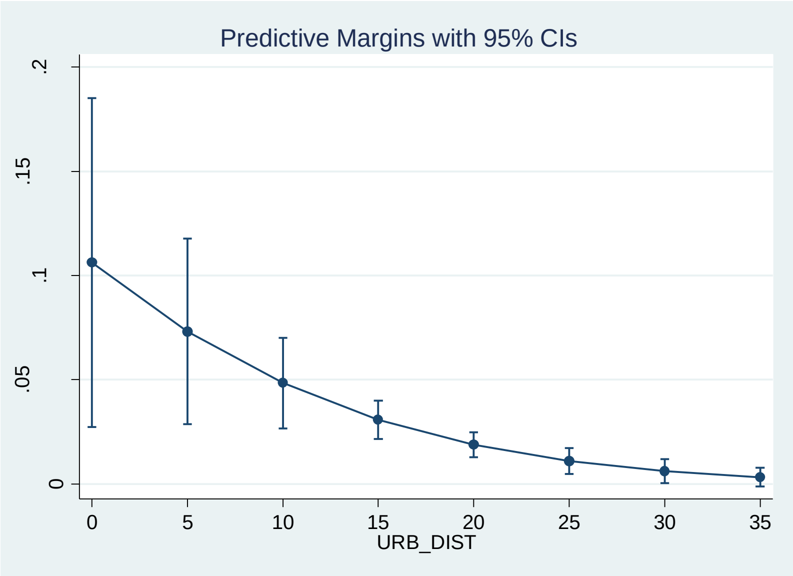
Predictive Margins with 95% CIs
0
.01
.02
.03
.04
.05
0
1
2
3
4
5
6
7
SROAD_DIST

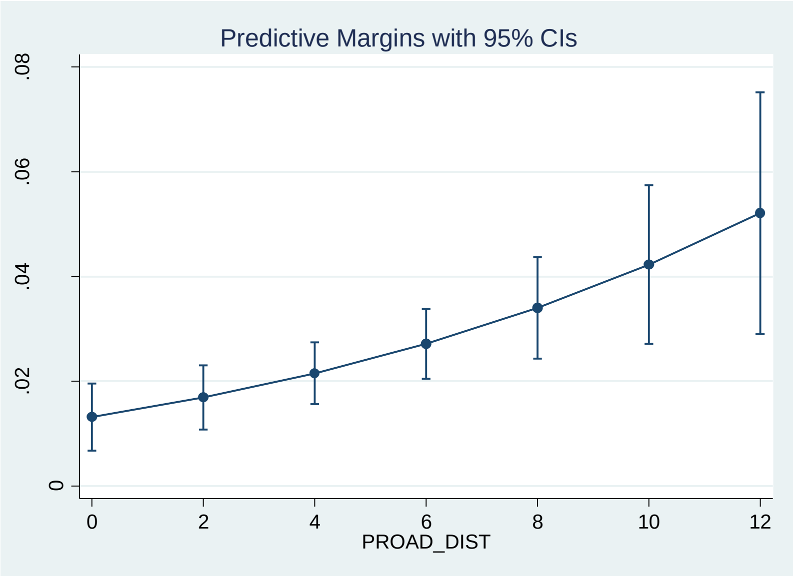


Predictive Margins with 95% CIs
0
.01
.02
.03
.04
.05
0
10
20
30
40
50
60
70
80
90
100
GRZ_COS

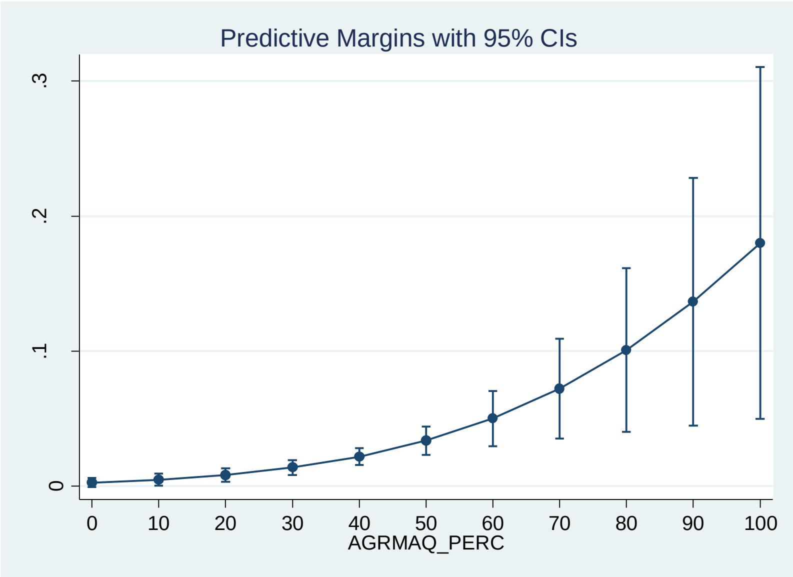

Predictive Margins with 95% CIs
-.01
0
.01
.02
.03
.04
0
10
20
30
40
50
60
70
80
90
100
PIN_COS


Predictive Margins with 95% CIs
0
.02
.04
.06
.08
0
6
12
18
24
30
36
42
48
54
RPSUP_PERC

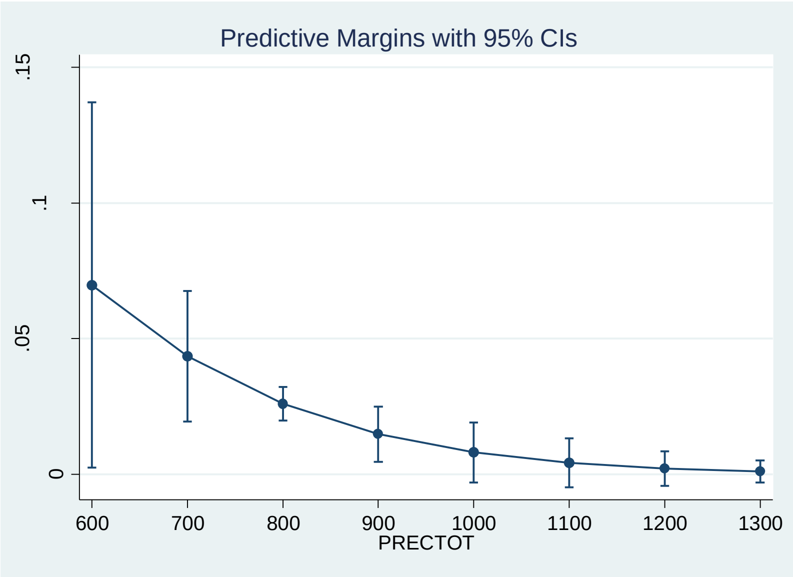

Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0
5
10
15
20
25
30
35
URB_DIST



Predictive Margins with 95% CIs
0
.02
.04
.06
.08
0
2
4
6
8
10
12
PROAD_DIST

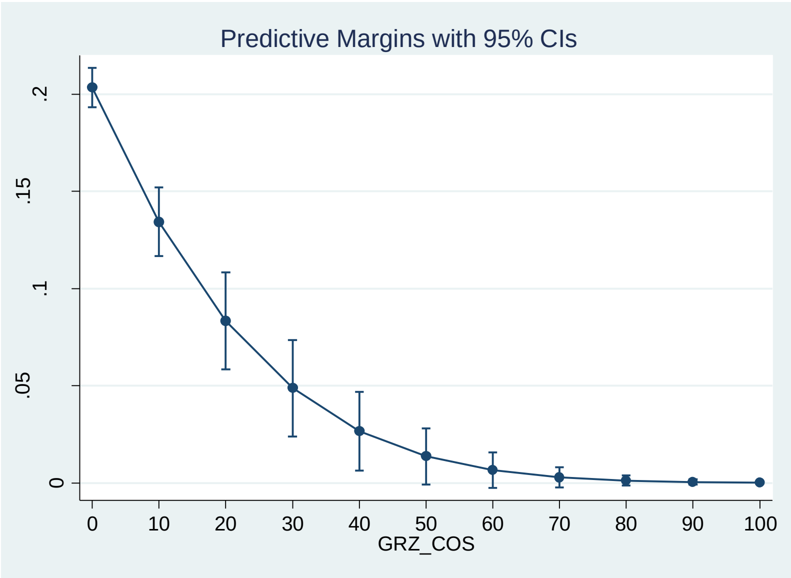


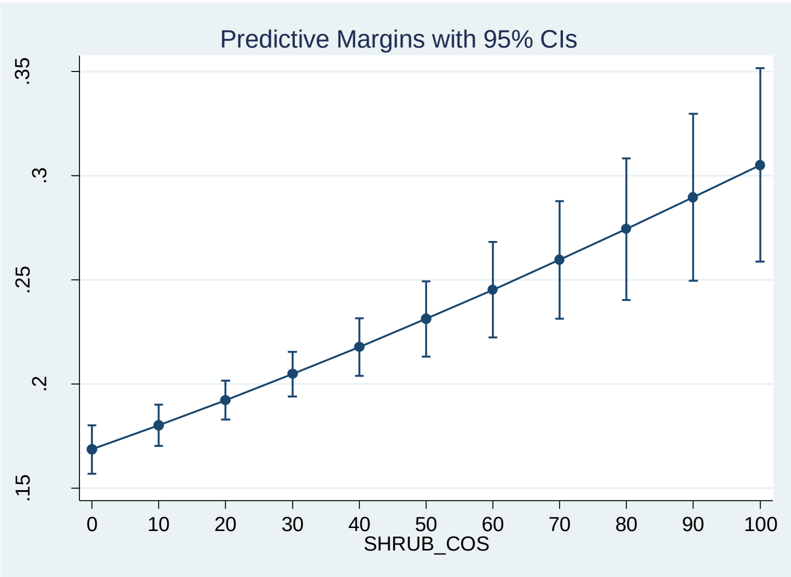
Predictive Margins with 95% CIs
0
.1
.2
.3
0
10
20
30
40
50
60
70
80
90
100
AGRMAQ_PERC

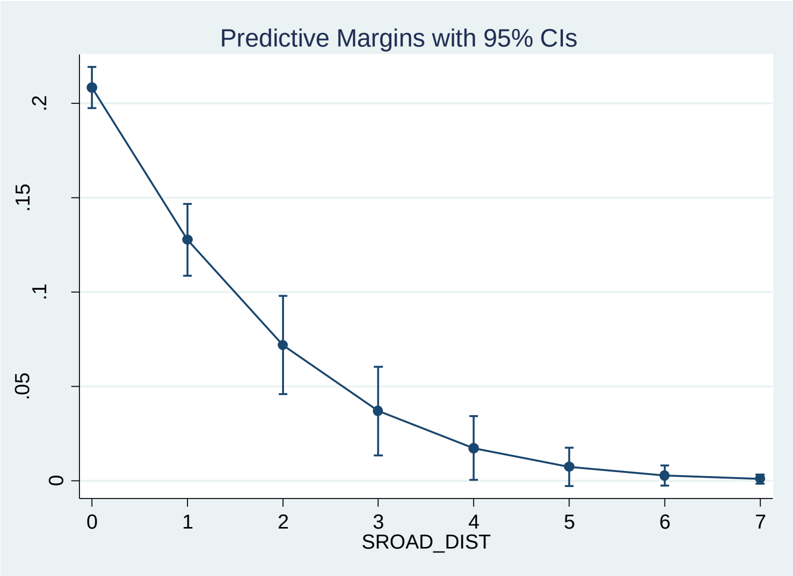


Predictive Margins with 95% CIs
0
.05
.1
.15
600
700
800
900
1000
1100
1200
1300
PRECTOT

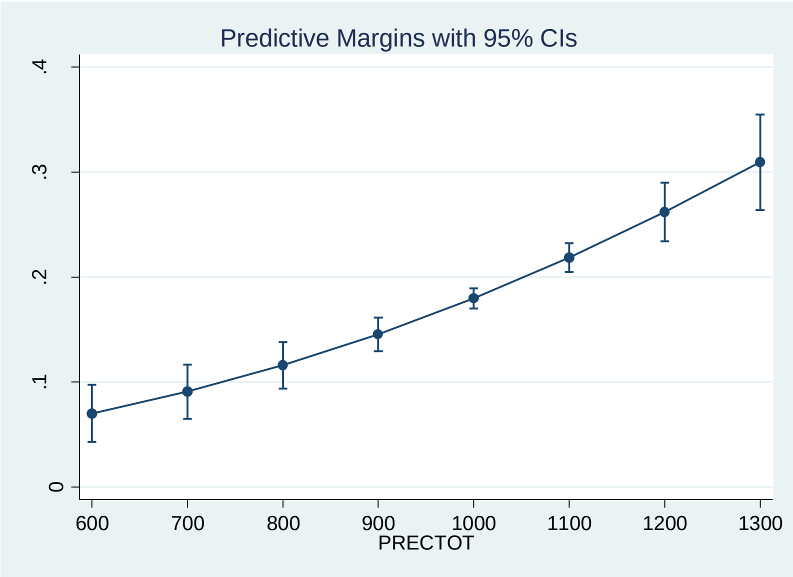
Figure D 4 – Large wildfire ignition model: Cluster 3



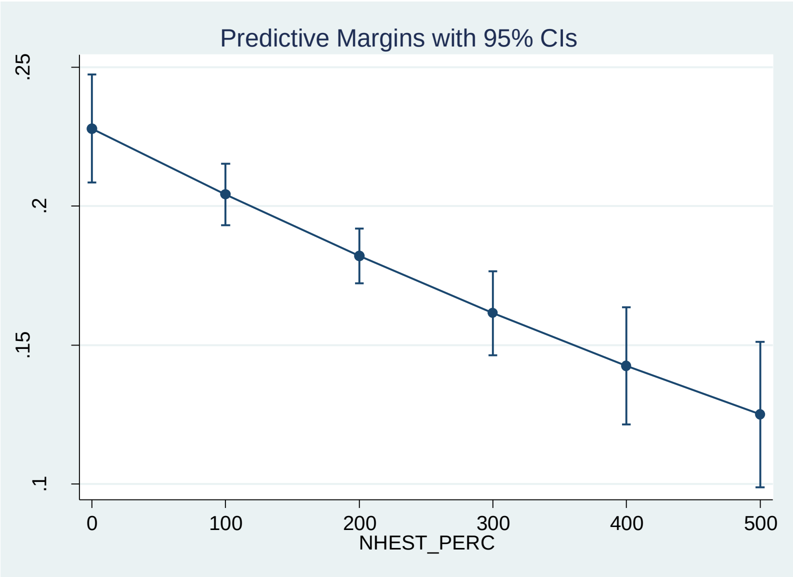




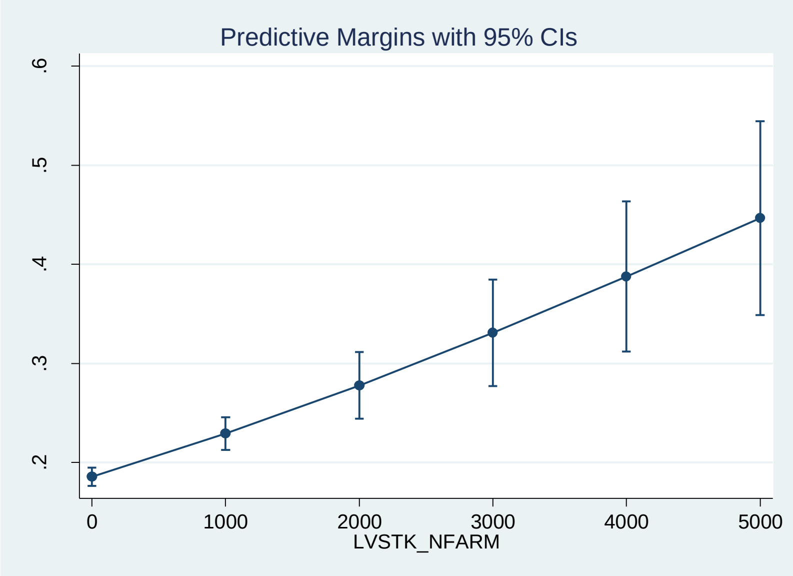
Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0 1 2 3 4 5 6 7
SROAD_DIST

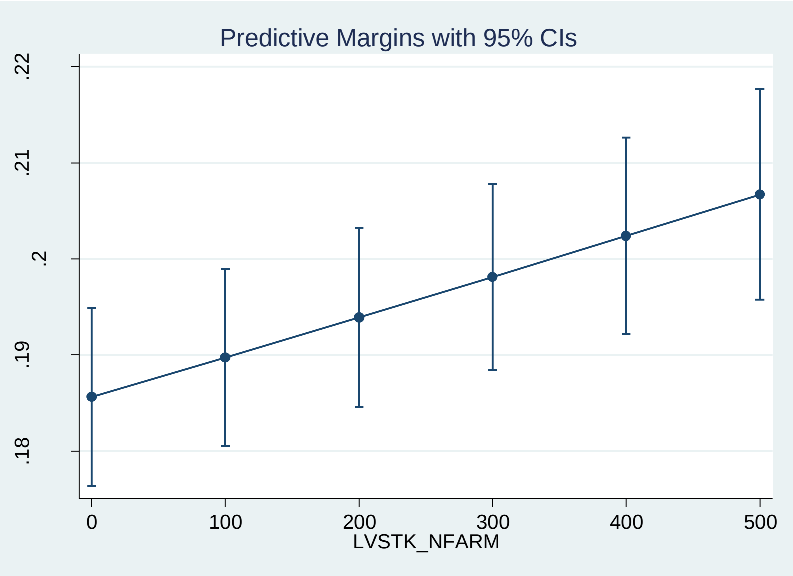


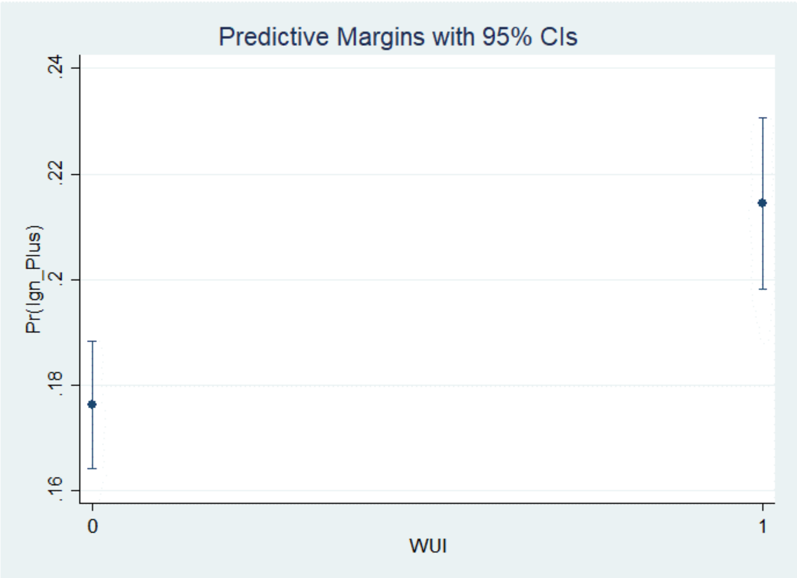
Predictive Margins with 95% CIs
0
.1
.2
.3
.4
600 700 800 900 1000 1100 1200 1300
PRECTOT

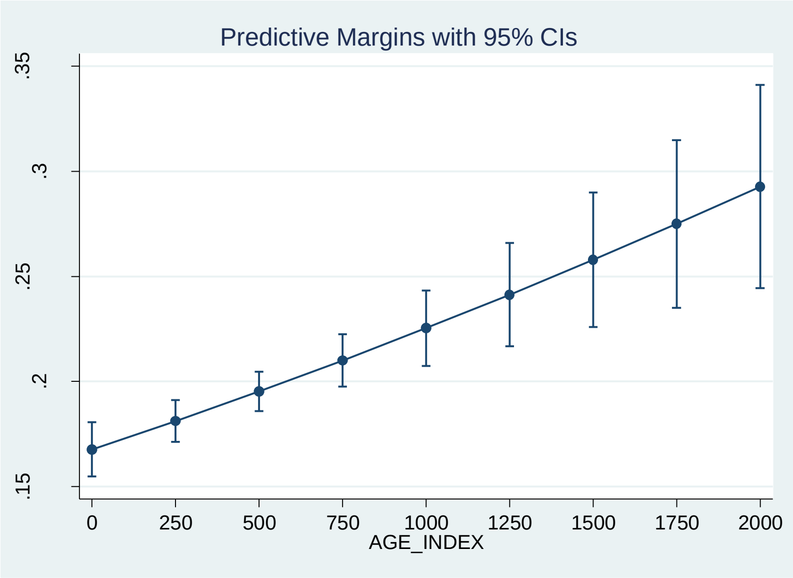


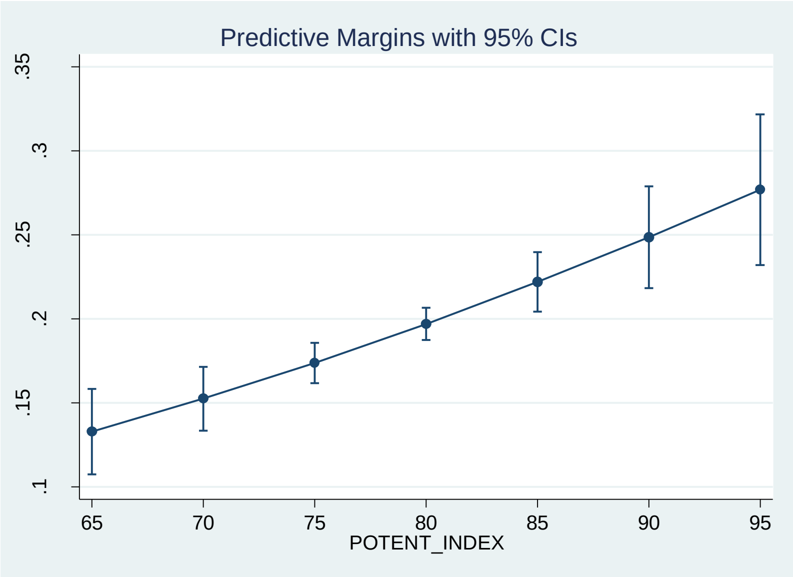
Predictive Margins with 95% CIs
.1
.15
.2
.25
0 100 200 300 400 500
NHEST_PERC

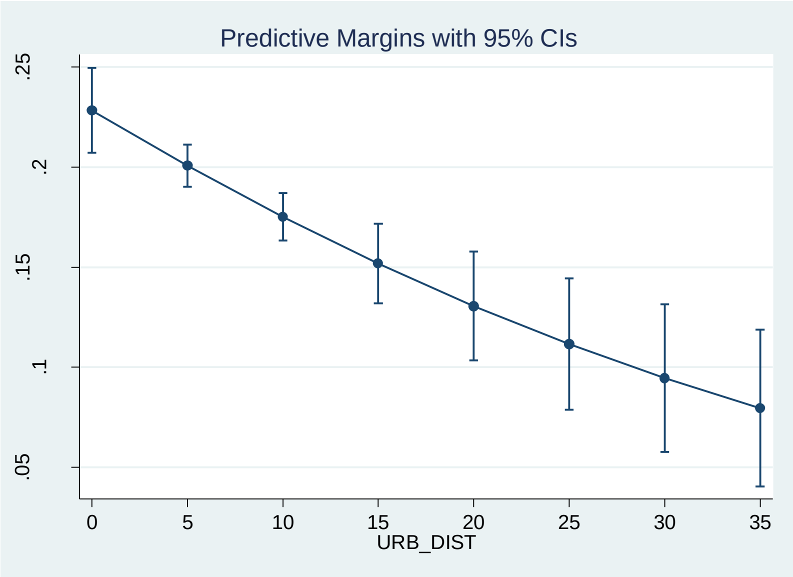

Predictive Margins with 95% CIs
.2
.3
.4
.5
.6
0
1000
2000
3000
4000
5000
LVSTK_NFARM

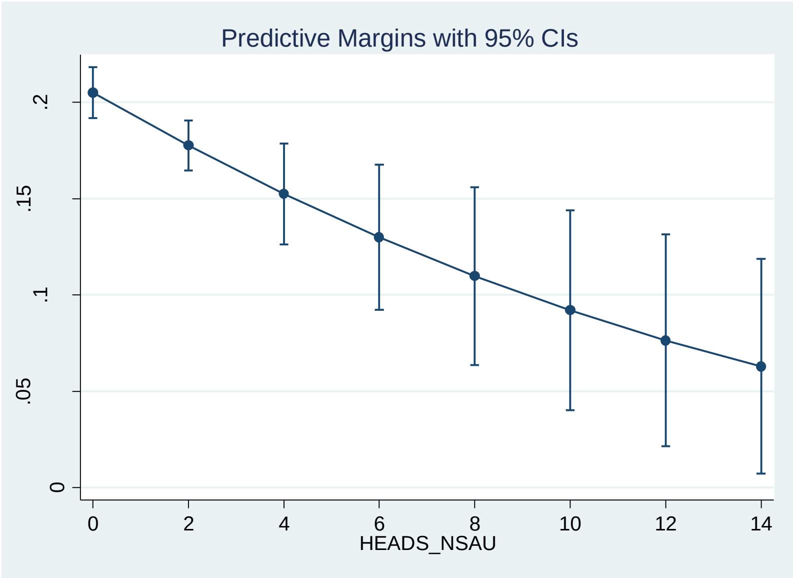

Predictive Margins with 95% CIs
.18
.19
.2
.21
.22
0
100
200
300
400
500
LVSTK_NFARM

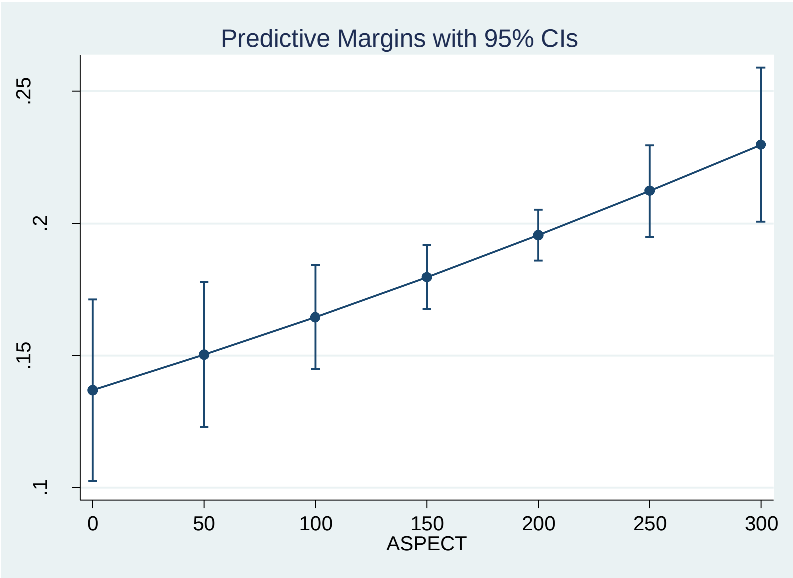

Predictive Margins with 95% CIs
Pr(Ign_Plus)
.16
.18
.2
.22
.24
0
1
WUI

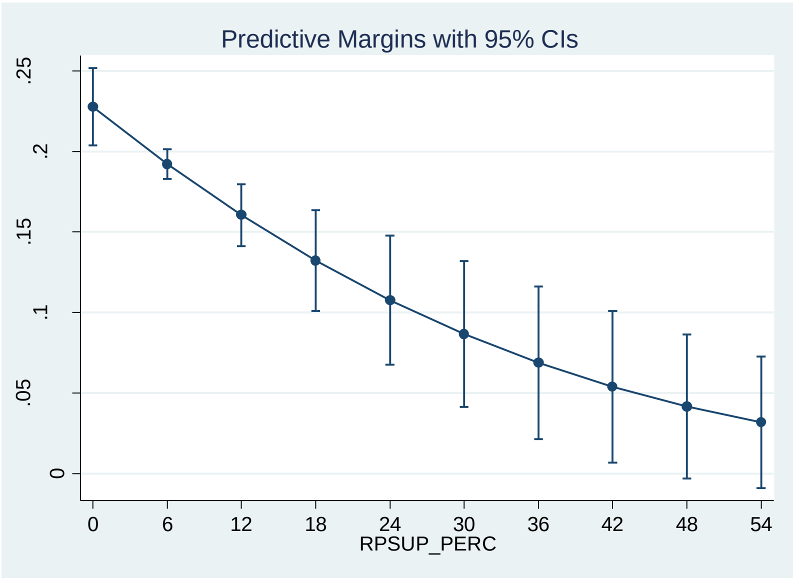

Predictive Margins with 95% CIs
.15
.2
.25
.3
.35
0
250
500
750
1000
1250
1500
1750
2000
AGE_INDEX

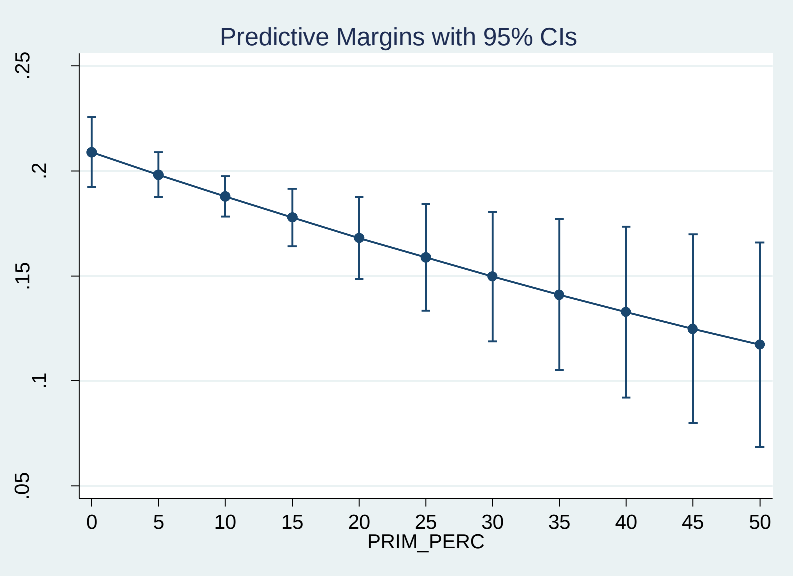

Predictive Margins with 95% CIs
.1
.15
.2
.25
.3
.35
65
70
75
80
85
90
95
POTENT_INDEX

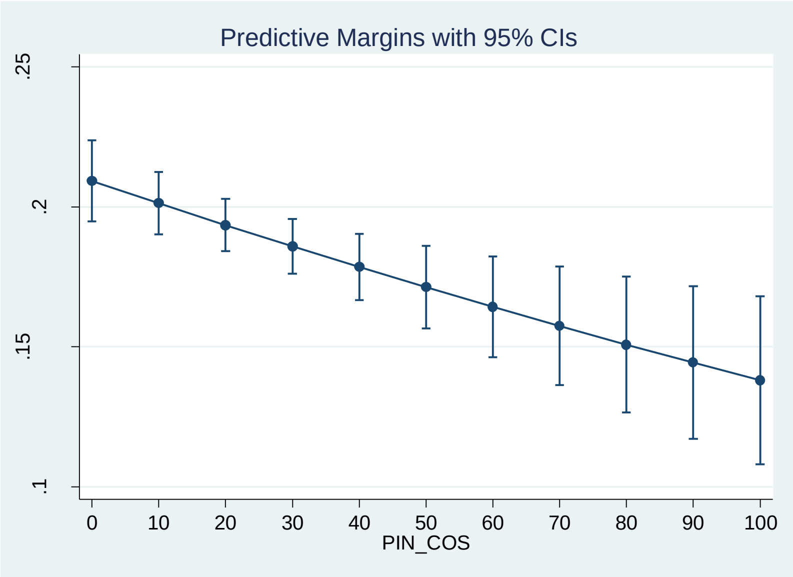

Predictive Margins with 95% CIs
.05
.1
.15
.2
.25
0
5
10
15
20
25
30
35
URB_DIST

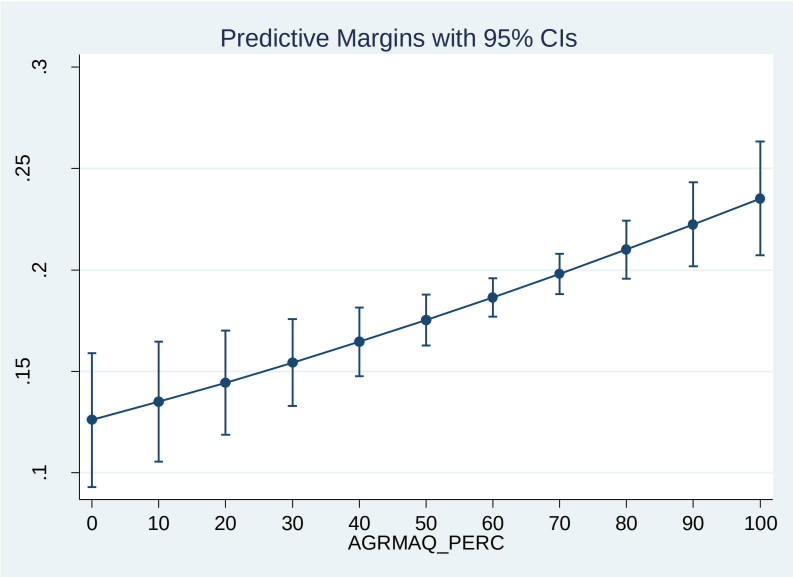

Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0
2
4
6
8
10
12
14
HEADS_NSAU

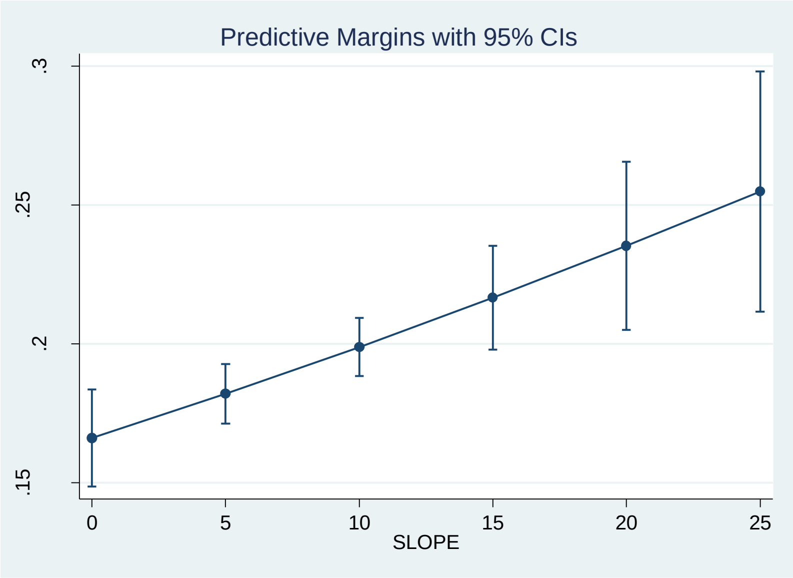

Predictive Margins with 95% CIs
.1
.15
.2
.25
0
50
100
150
200
250
300
ASPECT


Predictive Margins with 95% CIs
0
.05
.1
.15
.2
.25
0
6
12
18
24
30
36
42
48
54
RPSUP_PERC

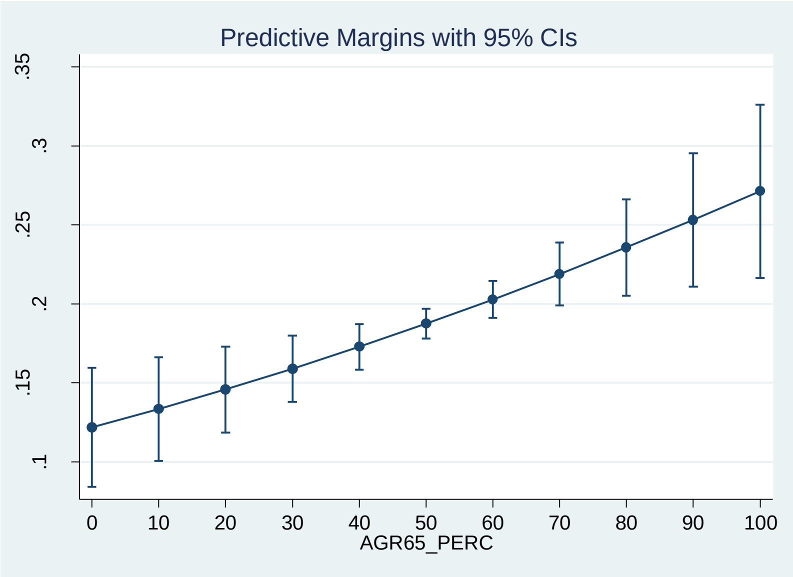

Predictive Margins with 95% CIs
.05
.1
.15
.2
.25
0 5 10 15 20 25 30 35 40 45 50
PRIM_PERC


Predictive Margins with 95% CIs
.1
.15
.2
.25
0 10 20 30 40 50 60 70 80 90 100
PIN_COS


Predictive Margins with 95% CIs
.1
.15
.2
.25
.3
0 10 20 30 40 50 60 70 80 90 100
AGRMAQ_PERC

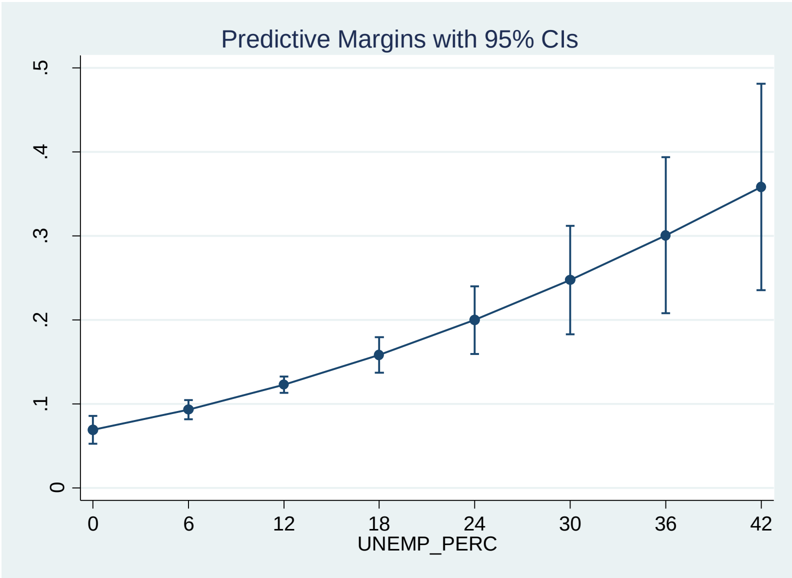


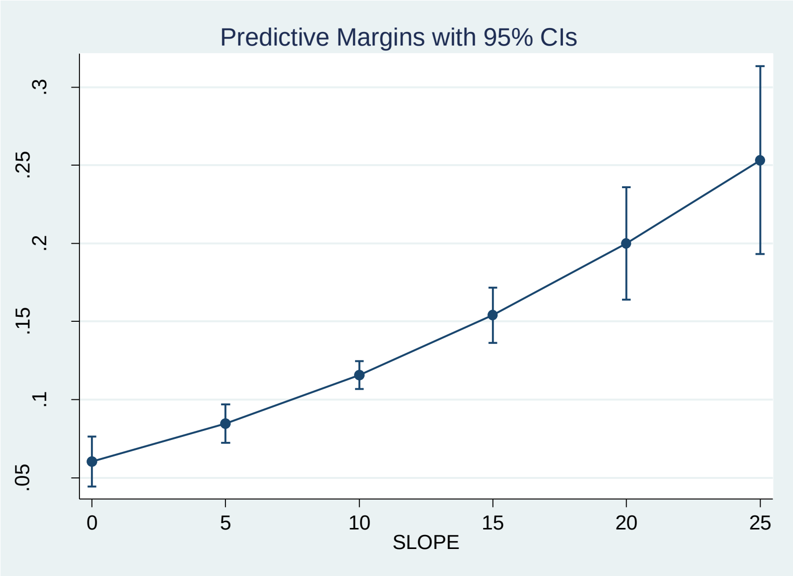
Predictive Margins with 95% CIs
.15
.2
.25
.3
0
5
10
15
20
25
SLOPE

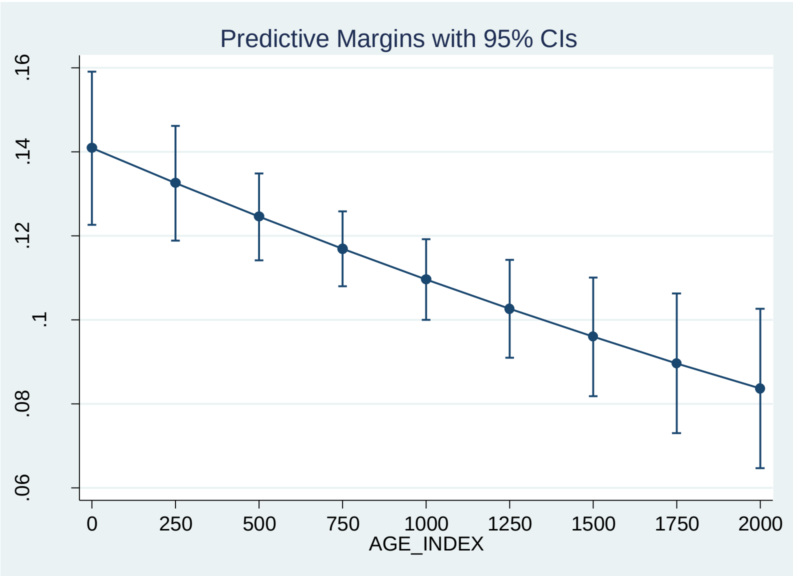


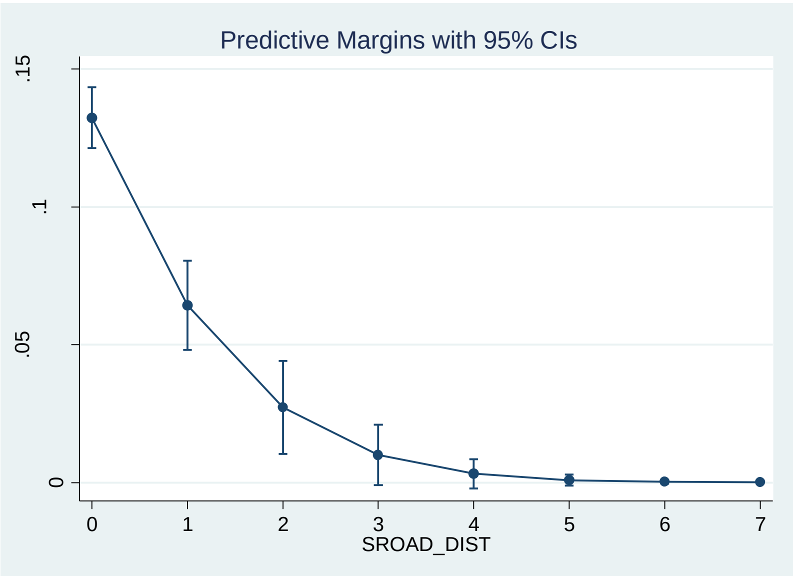
Predictive Margins with 95% CIs
.1
.15
.2
.25
.3
.35
0
10
20
30
40
50
60
70
80
90
100
AGR65_PERC

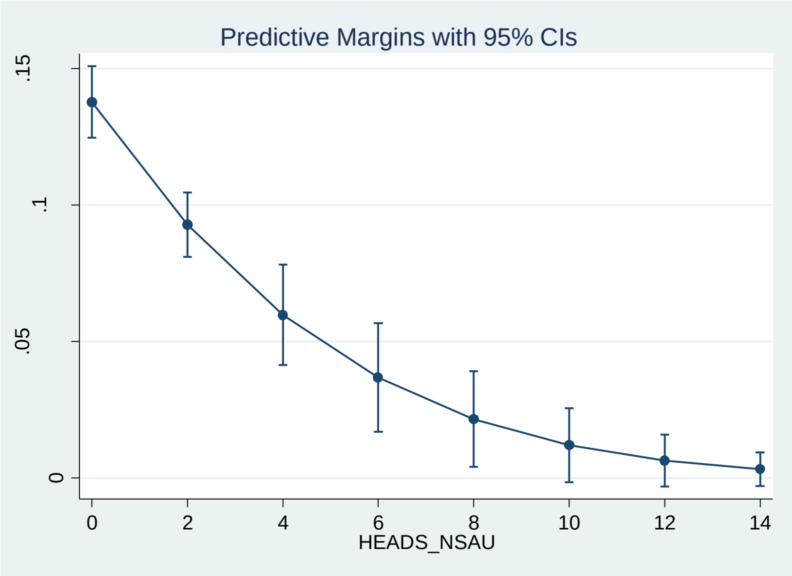
Figure D 5 – Large wildfire ignition model: Cluster 4



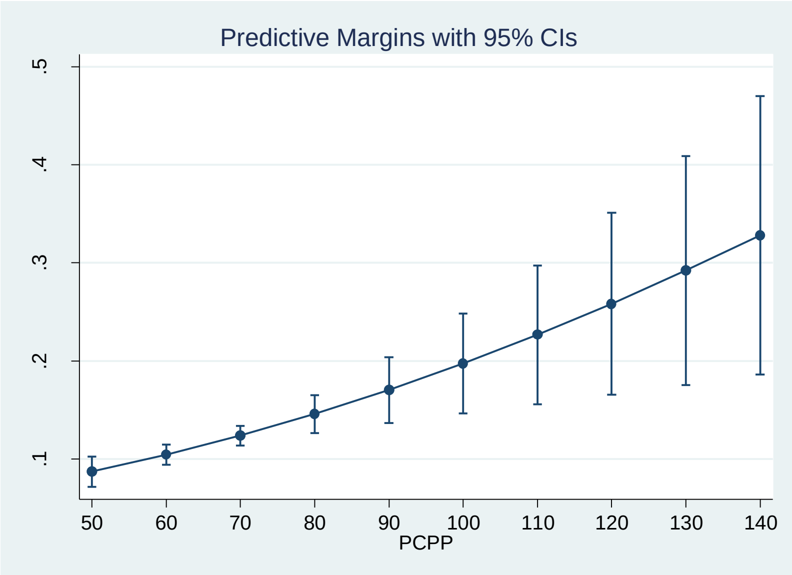

Predictive Margins with 95% CIs
.05
.1
.15
.2
.25
.3
0
5
10
15
20
25
SLOPE


Predictive Margins with 95% CIs
.06
.08
.1
.12
.14
.16
0
250
500
750
1000
1250
1500
1750
2000
AGE_INDEX

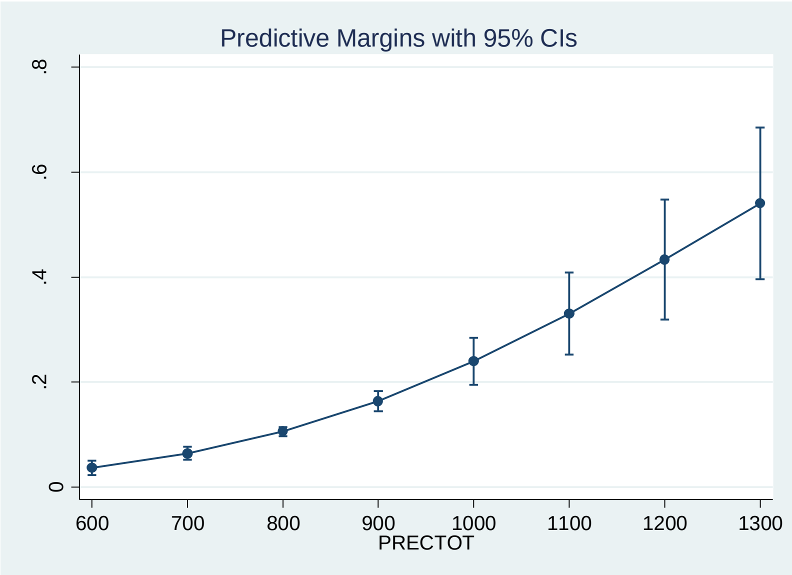

Predictive Margins with 95% CIs
0
.05
.1
.15
0
1
2
3
4
5
6
7
SROAD_DIST



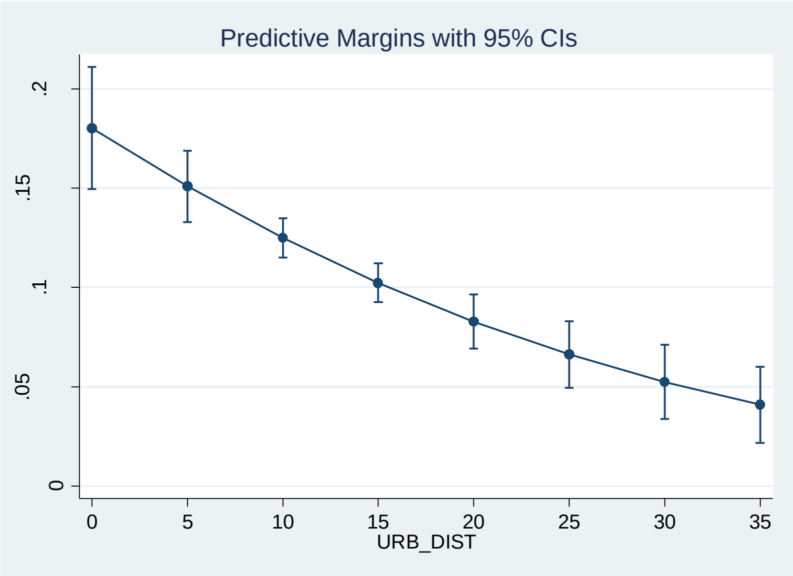
Predictive Margins with 95% CIs
0
.05
.1
.15
0
2
4
6
8
10
12
14
HEADS_NSAU

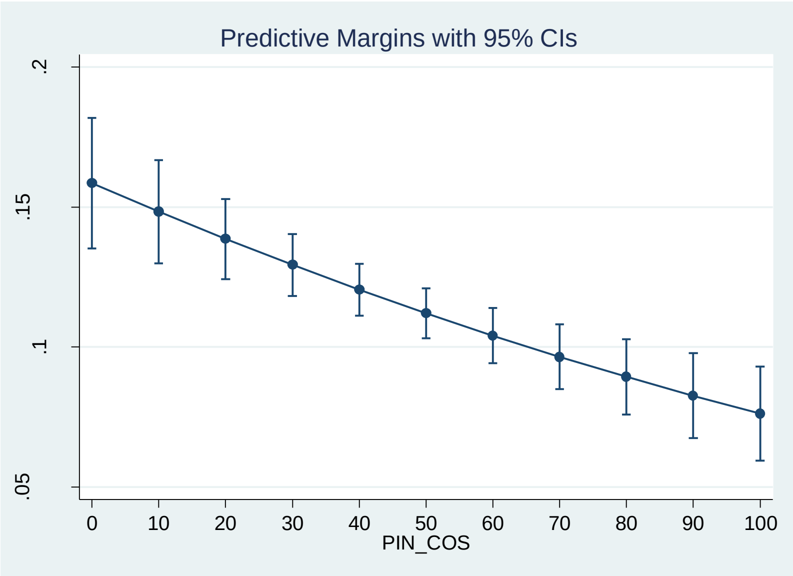


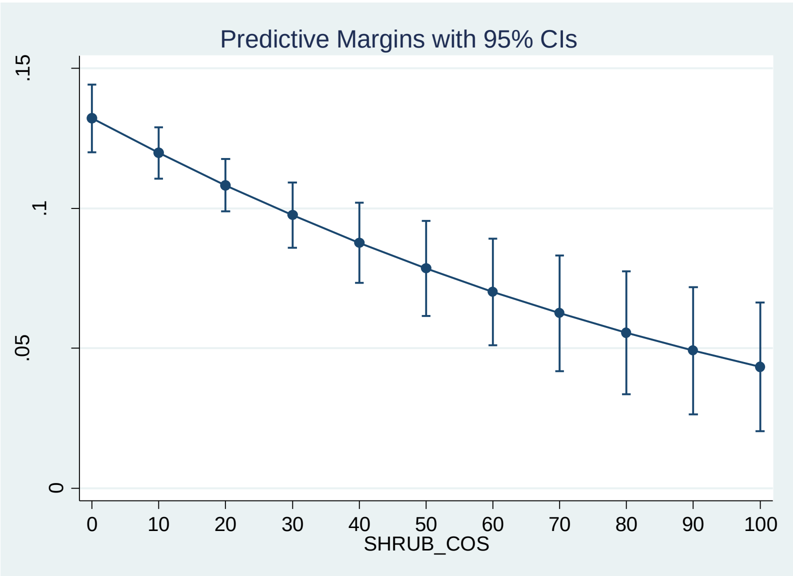
Predictive Margins with 95% CIs
.1
.2
.3
.4
.5
50
60
70
80
90
100
110
120
130
140
PCPP

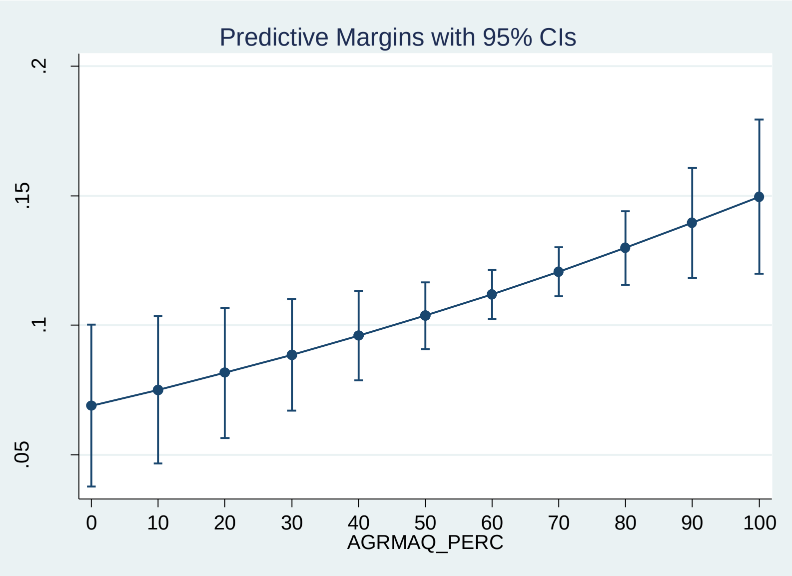


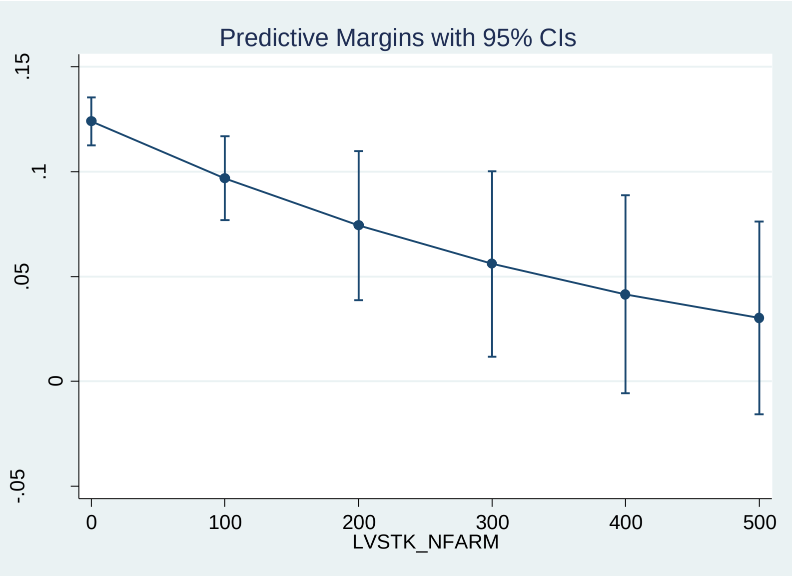
Predictive Margins with 95% CIs
0
.2
.4
.6
.8
600
700
800
900
1000
1100
1200
1300
PRECTOT

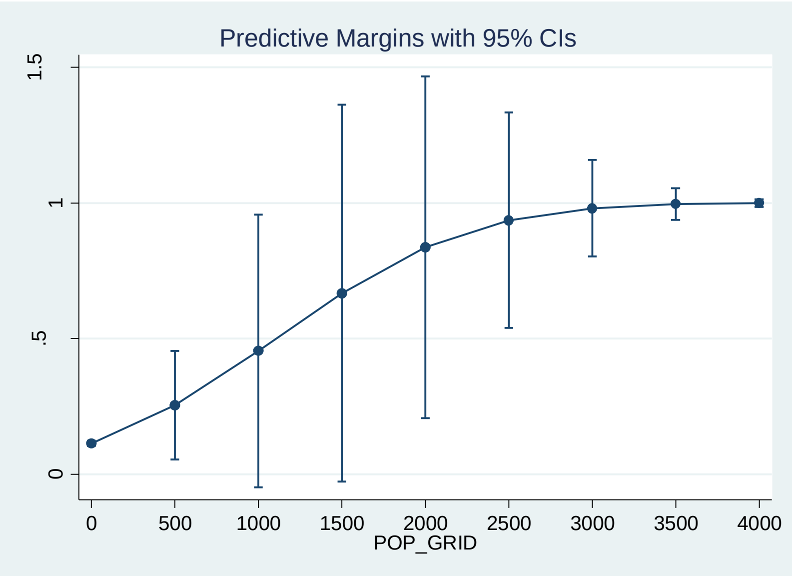

Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0
5
10
15
20
25
30
35
URB_DIST

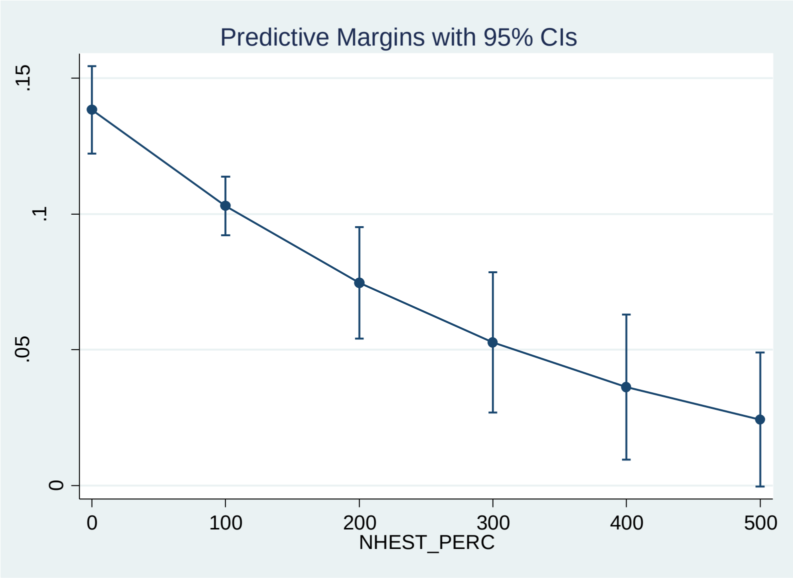

Predictive Margins with 95% CIs
.05
.1
.15
.2
0
10
20
30
40
50
60
70
80
90
100
PIN_COS

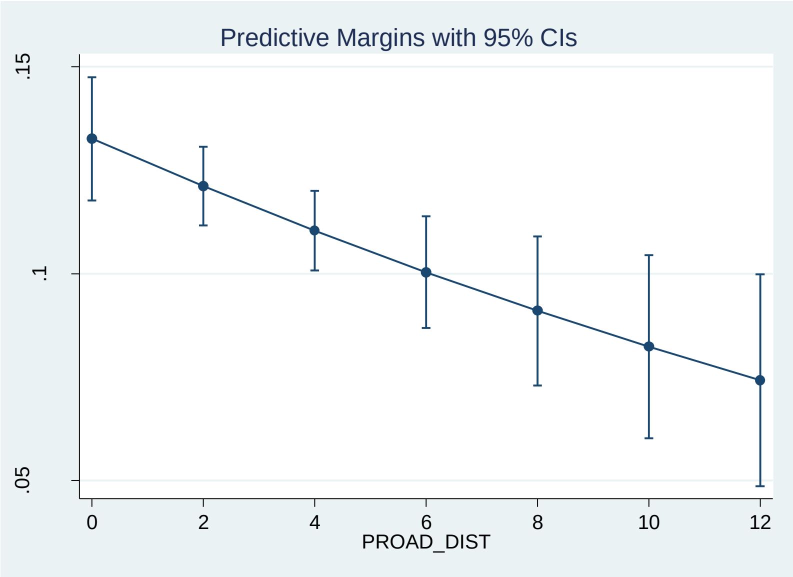

Predictive Margins with 95% CIs
0
.05
.1
.15
0
10
20
30
40
50
60
70
80
90
100
SHRUB_COS


Predictive Margins with 95% CIs
.05
.1
.15
.2
0
10
20
30
40
50
60
70
80
90
100
AGRMAQ_PERC


Predictive Margins with 95% CIs
-.05
0
.05
.1
.15
0
100
200
300
400
500
LVSTK_NFARM


Predictive Margins with 95% CIs
0
.5
.1
1.5
0
500
1000
1500
2000
2500
3000
3500
4000
POP_GRID


Predictive Margins with 95% CIs
0
.05
.1
.15
0
100
200
300
400
500
NHEST_PERC


Predictive Margins with 95% CIs
.05
.1
.15
0
2
4
6
8
10
12
PROAD_DIST

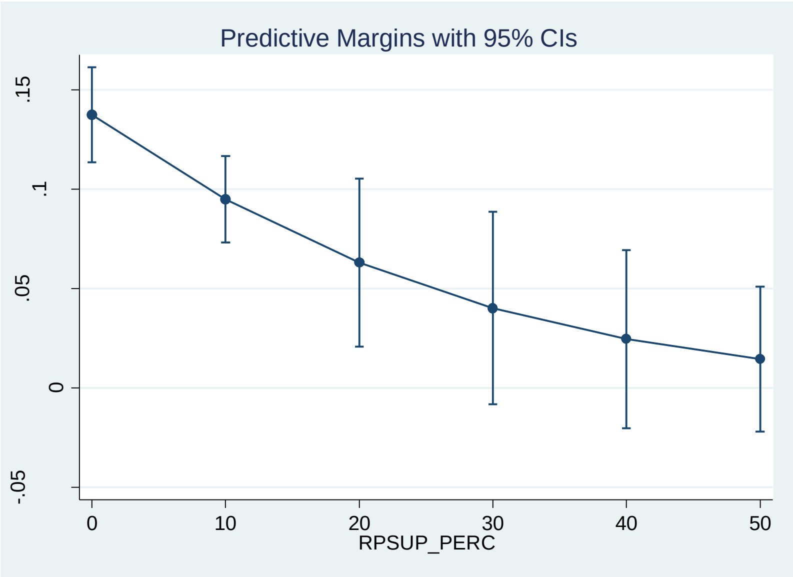
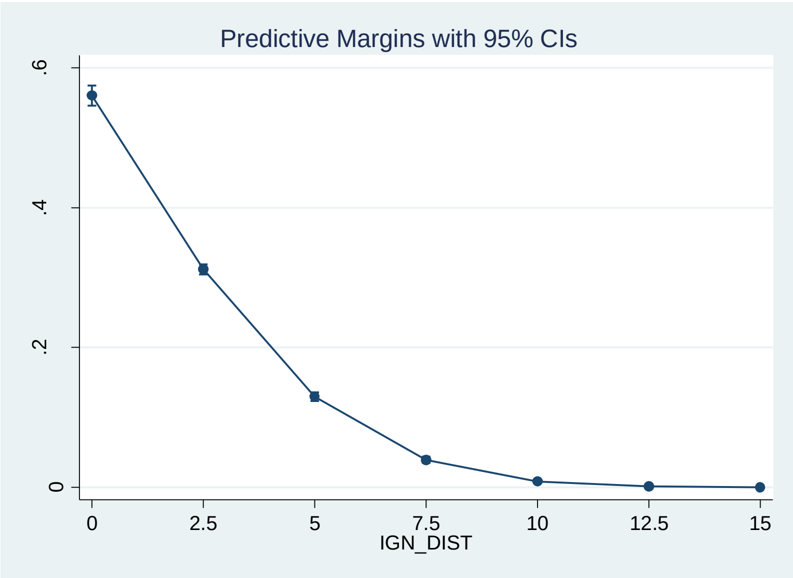
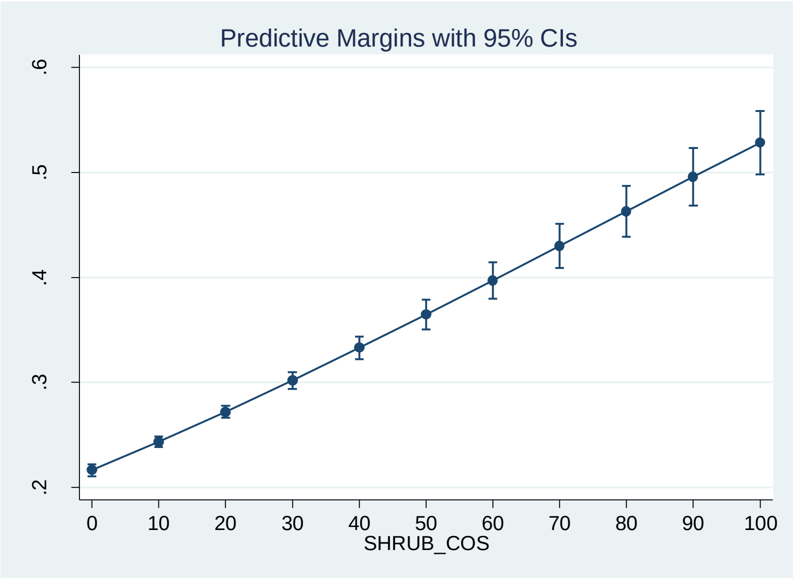
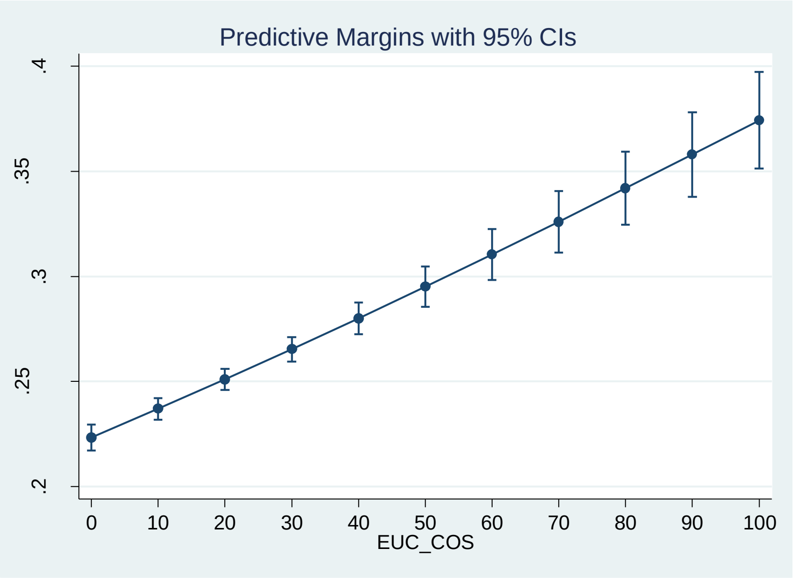
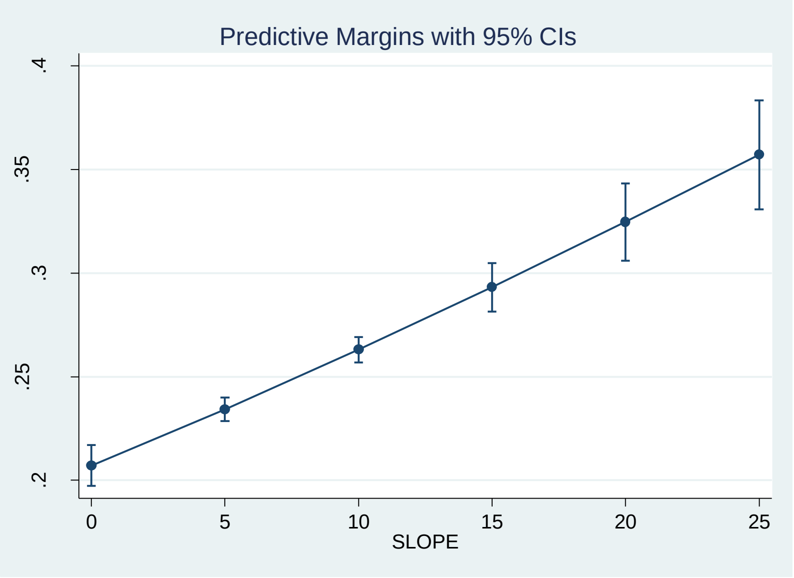

Predictive Margins with 95% CIs
-.05
0
.05
.1
.15
0
10
20
30
40
50
RPSUP_PERC

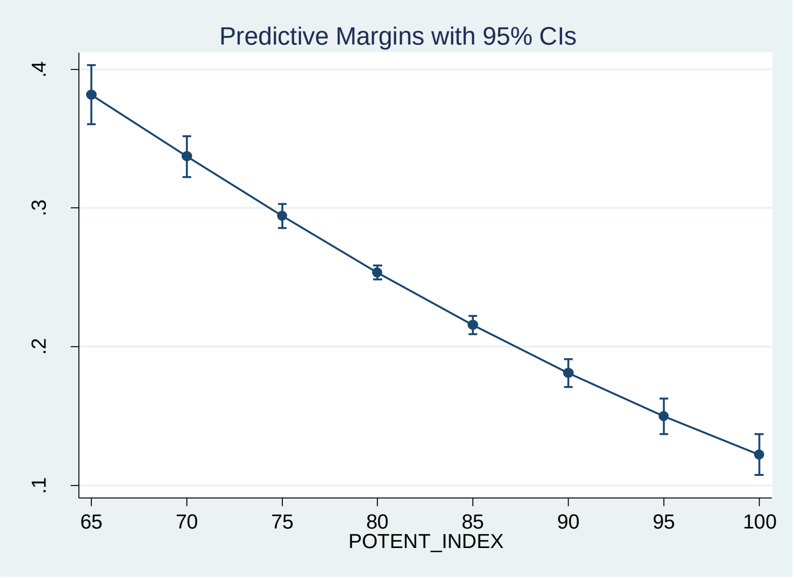
Figure D 6 – Large wildfire propagation model (first part): Global

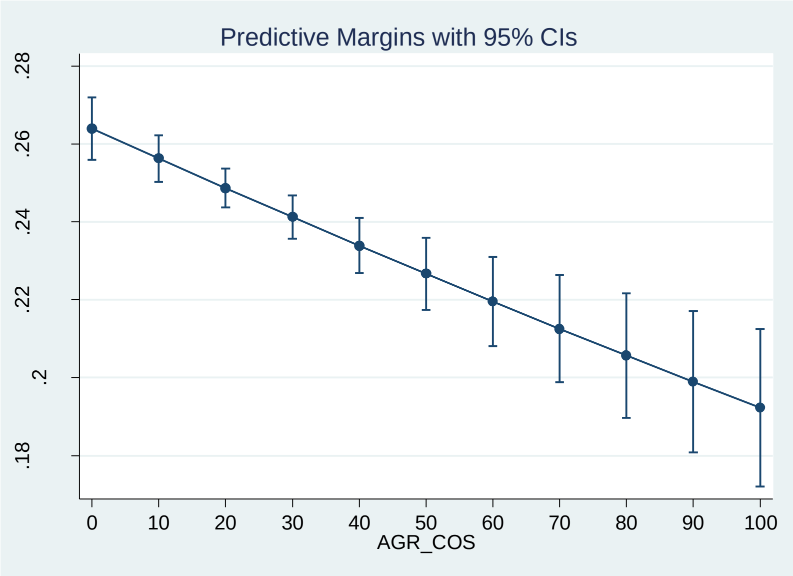


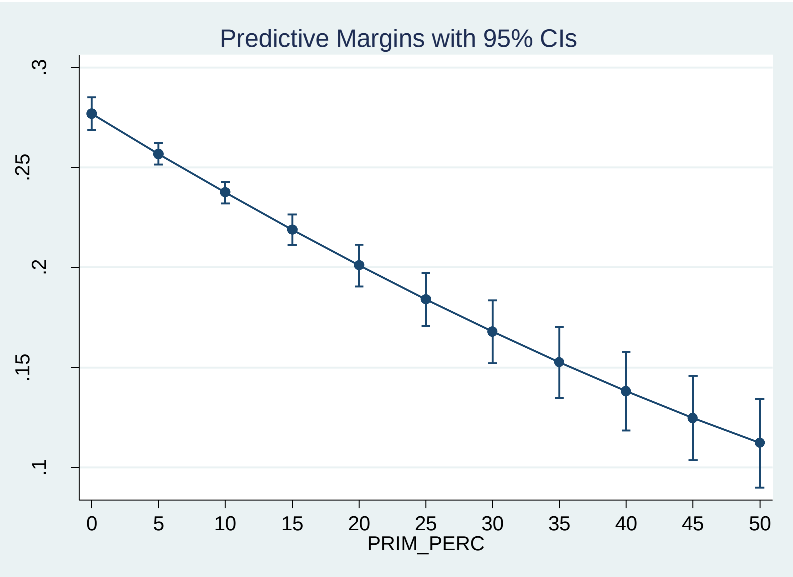


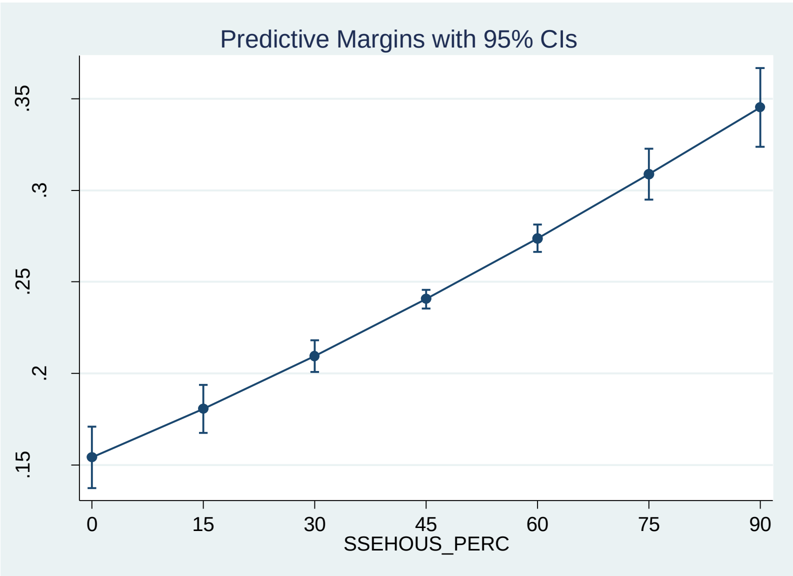

Predictive Margins with 95% CIs
.2
.25
.3
.35
.4
0
10
20
30
40
50
60
70
80
90
100
EUC_COS

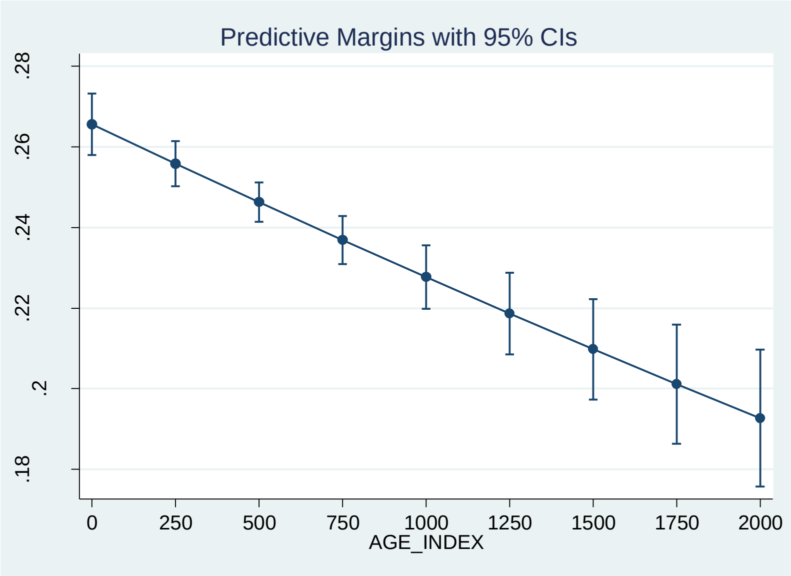

Predictive Margins with 95% CIs
.2
.25
.3
.35
.4
0
5
10
15
20
25
SLOPE


Predictive Margins with 95% CIs
.1
.2
.3
.4
65
70
75
80
85
90
95
100
POTENT_INDEX

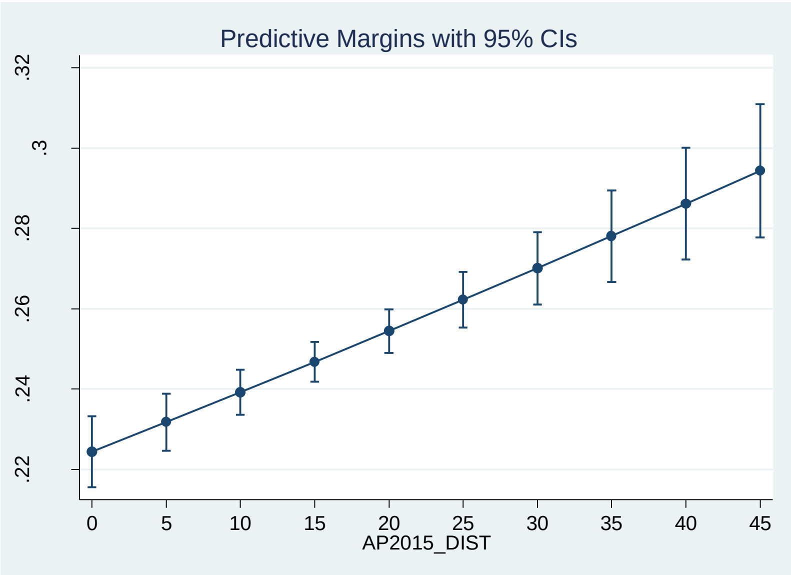

Predictive Margins with 95% CIs
.18
.2
.22
.24
.26
.28
0
10
20
30
40
50
60
70
80
90
100
AGR_COS


Predictive Margins with 95% CIs
.1
.15
.2
.25
.3
0
5
10
15
20
25
30
35
40
45
50
PRIM_PERC

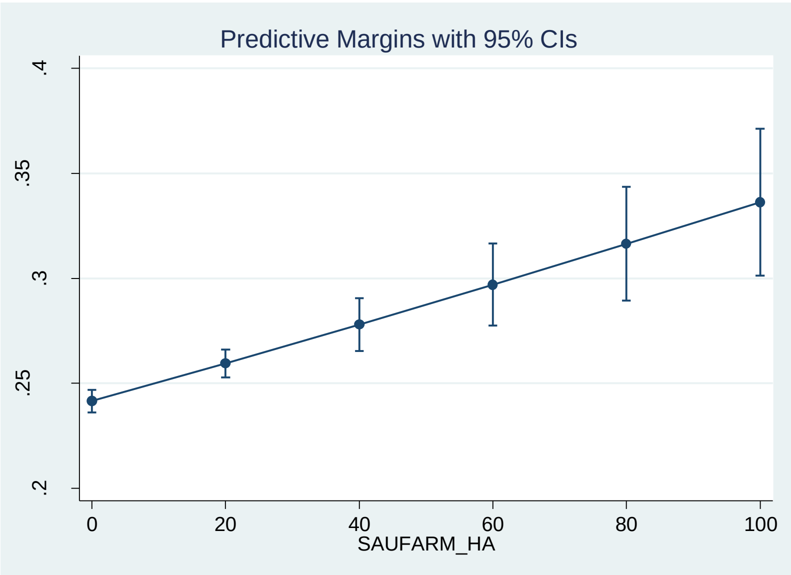

Predictive Margins with 95% CIs
.15
.2
.25
.3
.35
0
15
30
45
60
75
90
SSEHOUS_PERC



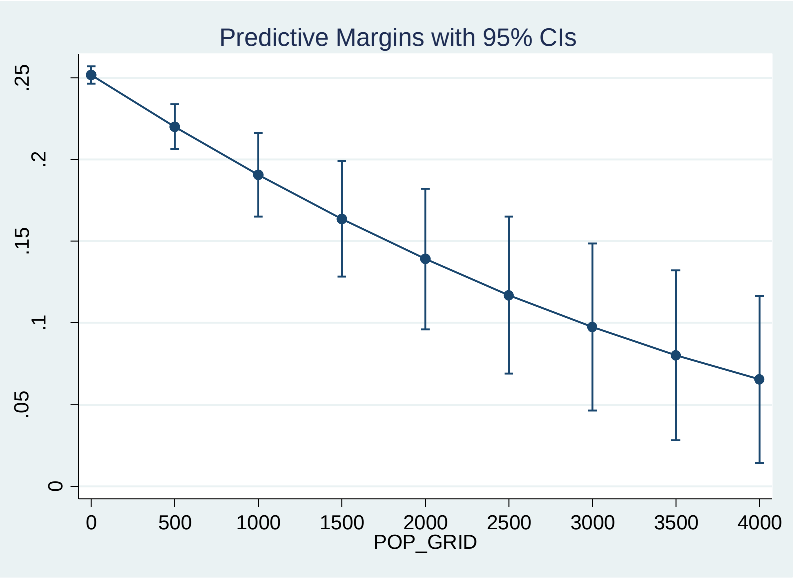
Predictive Margins with 95% CIs
.18
.2
.22
.24
.26
.28
0
250
500
750
1000
1250
1500
1750
2000
AGE_INDEX

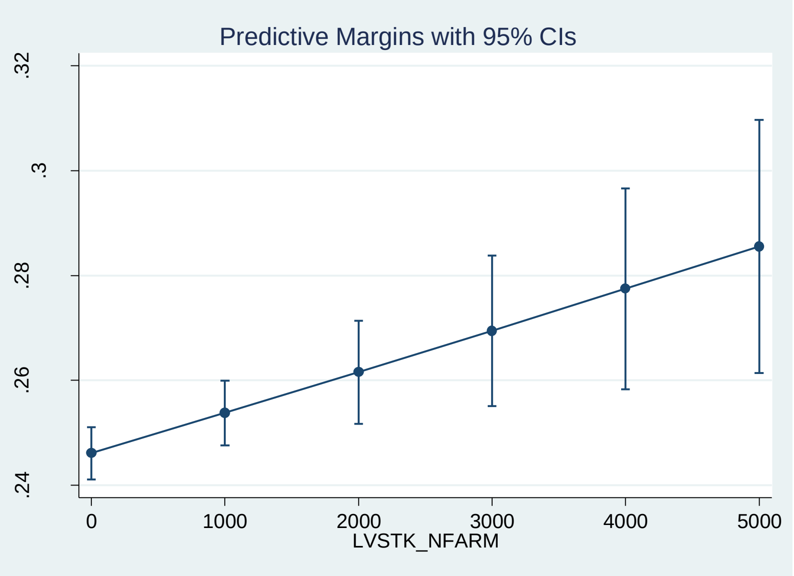


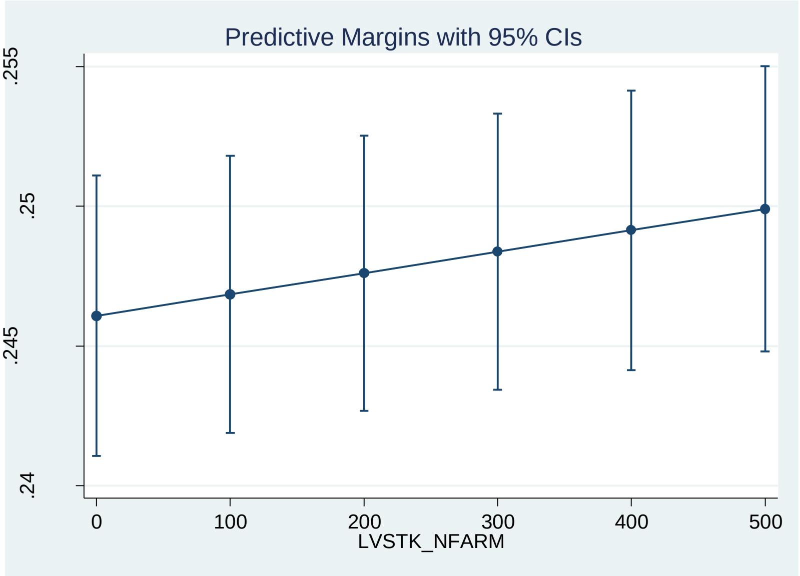
Predictive Margins with 95% CIs
.22
.24
.26
.28
.3
.32
0
5
10
15
20
25
30
35
40
45
AP2015_DIST



Predictive Margins with 95% CIs
.2
.25
.3
.35
.4
0
20
40
60
80
100
SAUFARM_HA


Predictive Margins with 95% CIs
0
.05
.1
.15
.2
.25
0
500
1000
1500
2000
2500
3000
3500
4000
POP_GRID


Predictive Margins with 95% CIs
.24
.26
.28
.3
.32
0
1000
2000
3000
4000
5000
LVSTK_NFARM


Predictive Margins with 95% CIs
.24
.245
.25
.255
0
100
200
300
400
500
LVSTK_NFARM

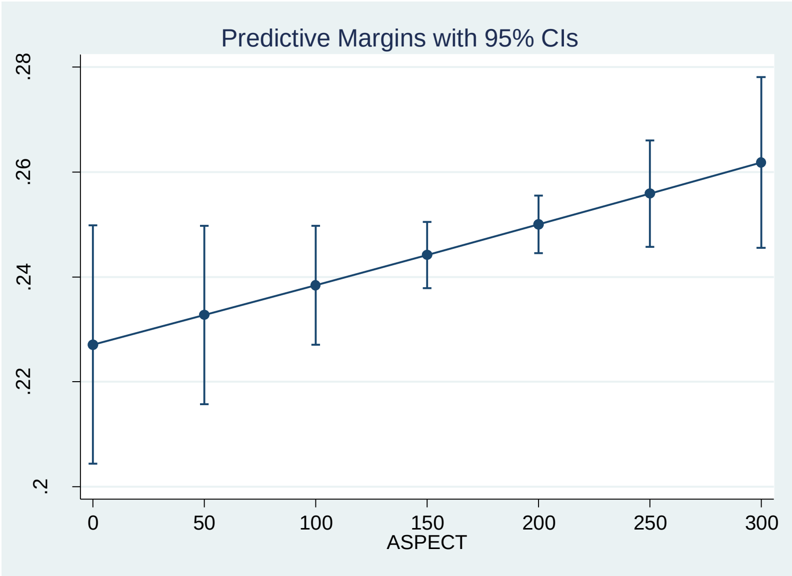
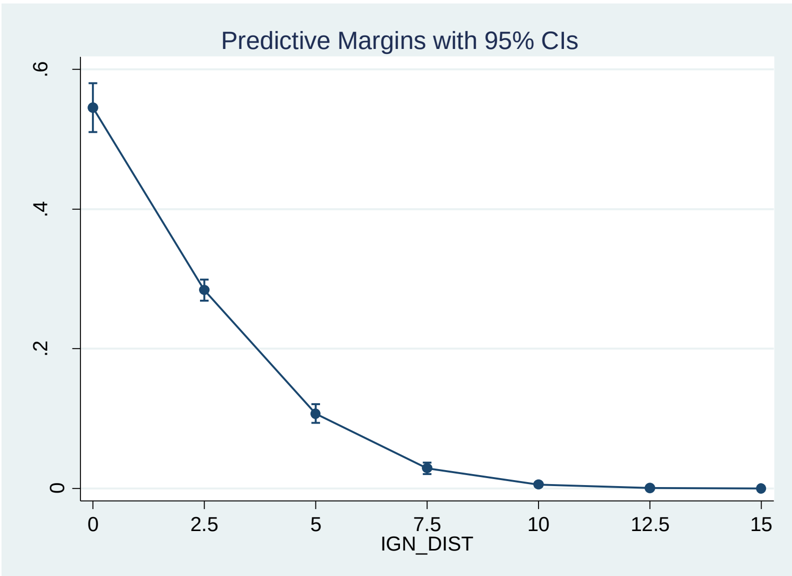
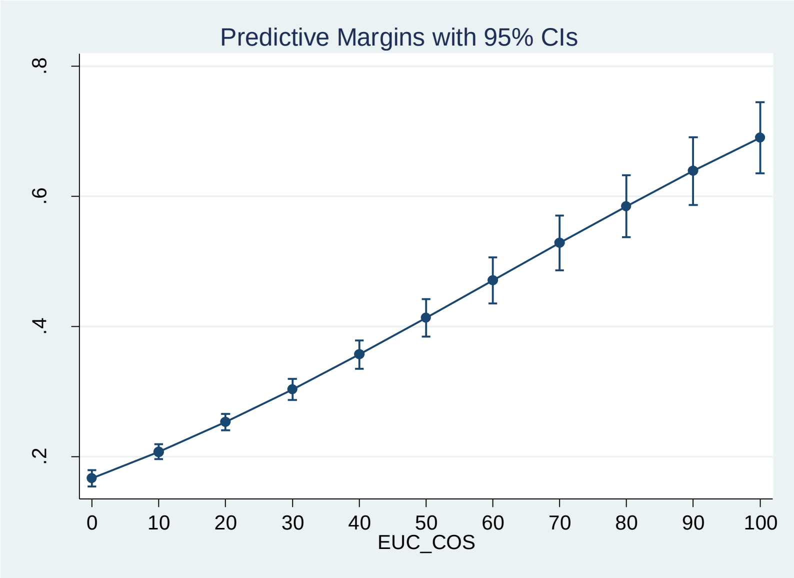
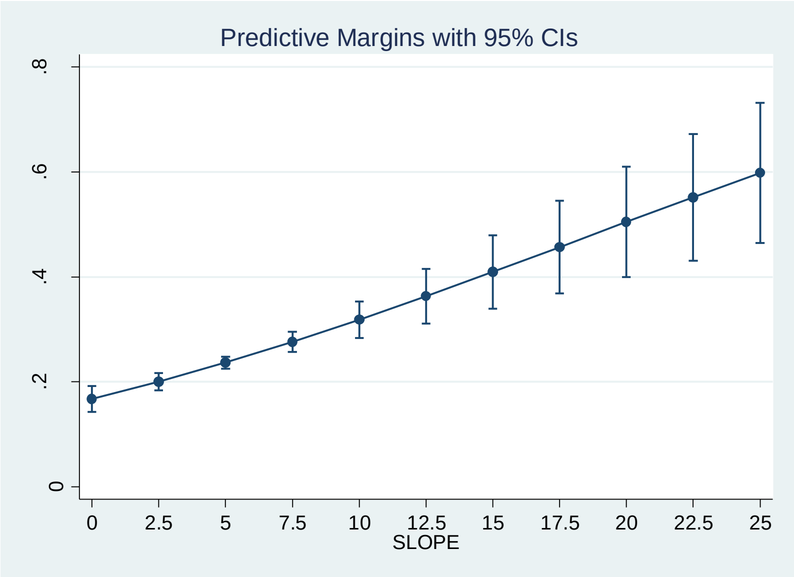
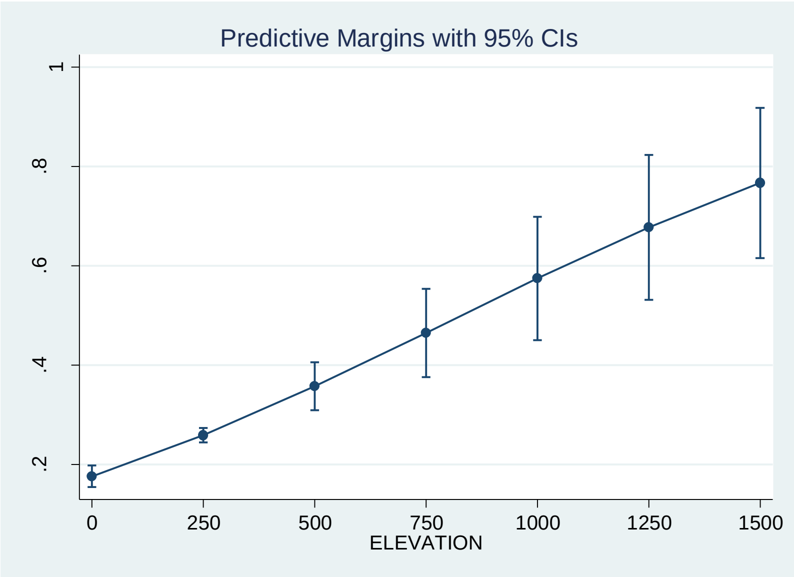


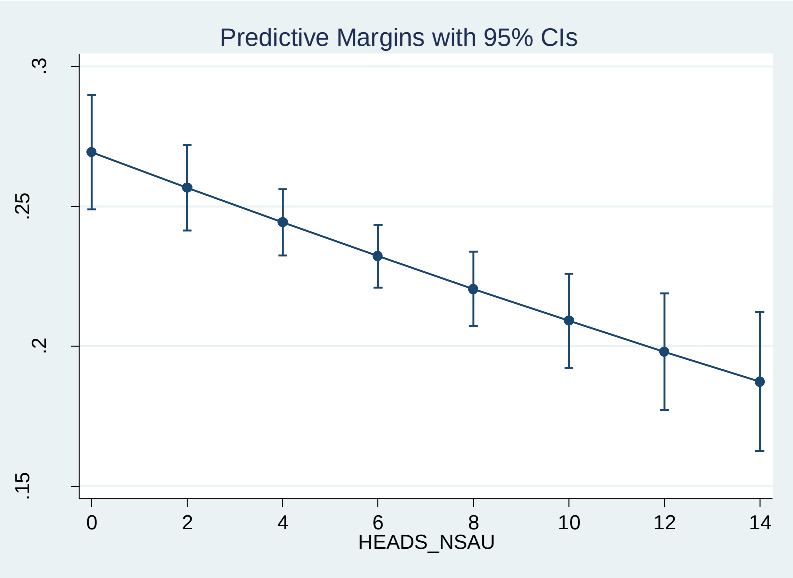
Figure D 7 – Large wildfire propagation model (first part): Cluster 1

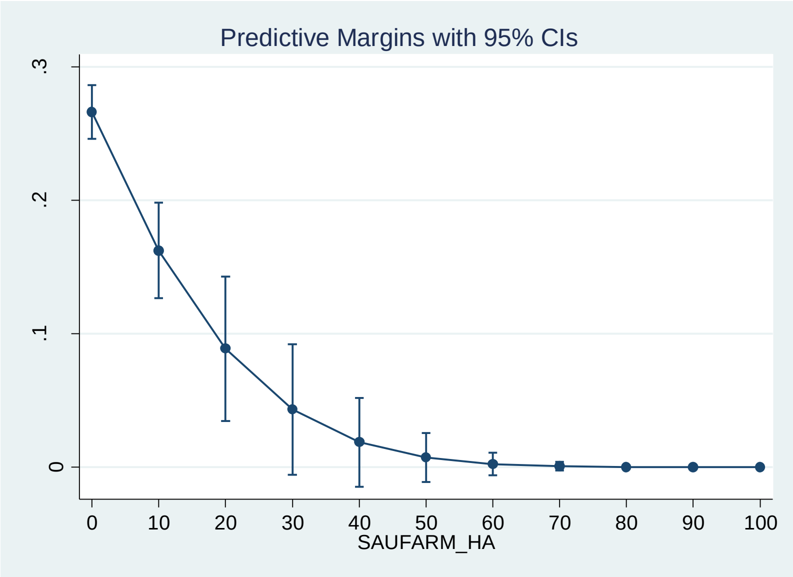


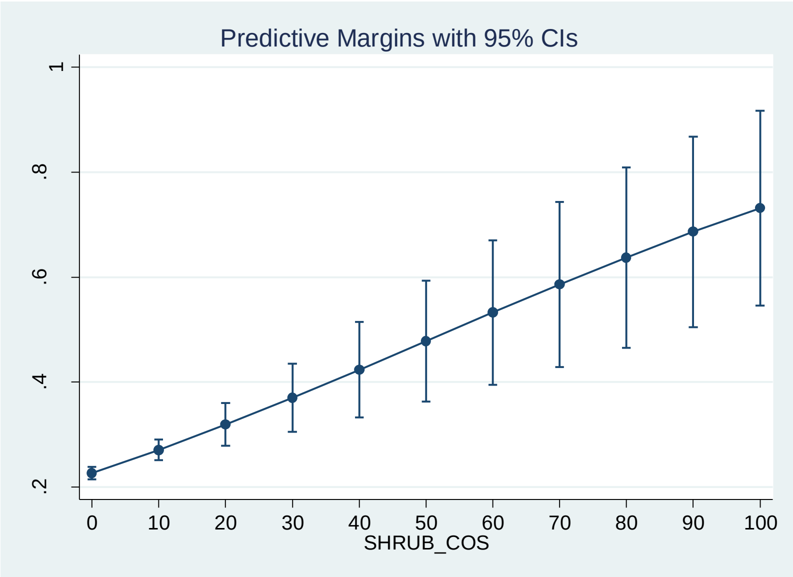

Predictive Margins with 95% CIs
.2
.4
.6
.8
0
10
20
30
40
50
60
70
80
90
100
EUC_COS

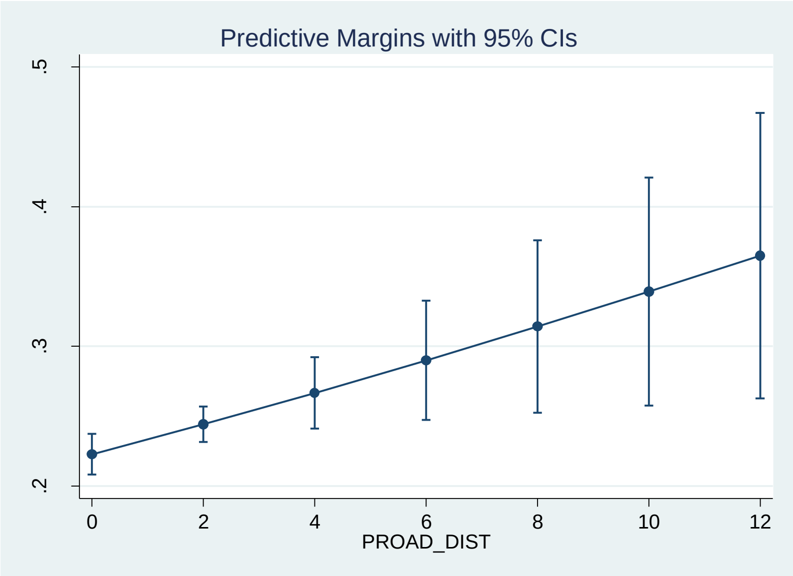

Predictive Margins with 95% CIs
0
.2
.4
.6
.8
0
2.5
5
7.5
10
12.5
15
17.5
20
22.5
25
SLOPE

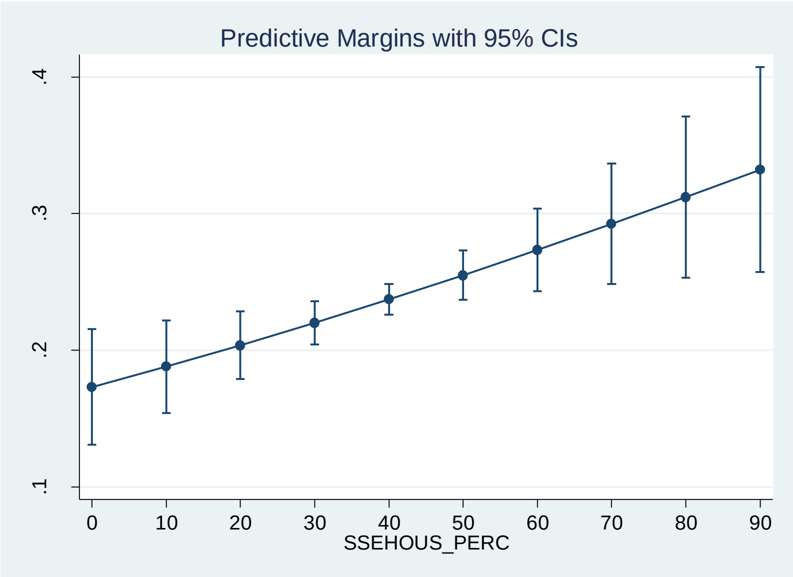

Predictive Margins with 95% CIs
.2
.4
.6
.8
.1
0
250
500
750
1000
1250
1500
ELEVATION

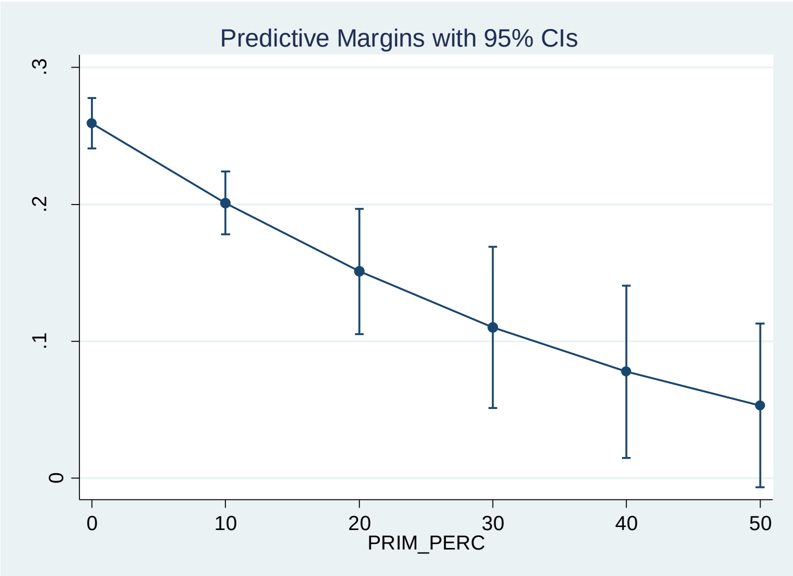

Predictive Margins with 95% CIs
.15
.2
.25
.3
0
2
4
6
8
10
12
14
HEADS_NSAU

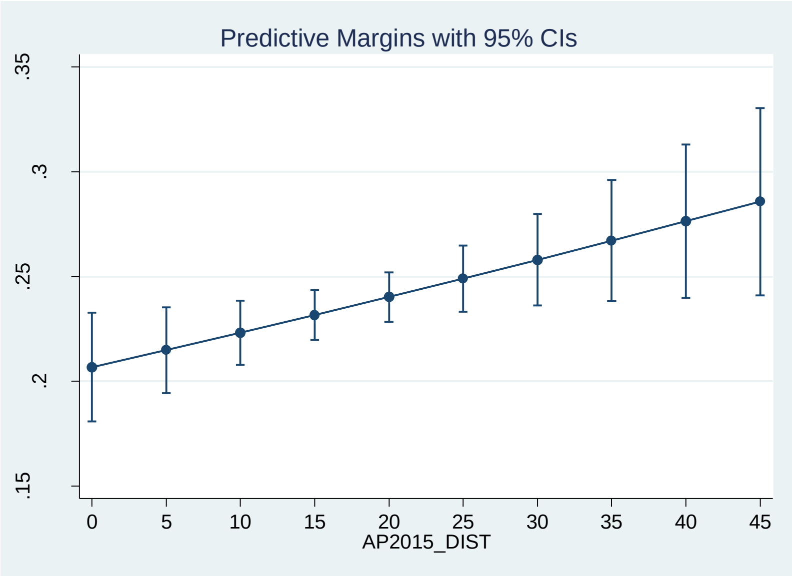

Predictive Margins with 95% CIs
0
.1
.2
.3
0
10
20
30
40
50
60
70
80
90
100
SAUFARM_HA


Predictive Margins with 95% CIs
.2
.4
.6
.8
1
0
10
20
30
40
50
60
70
80
90
100
SHRUB_COS


Predictive Margins with 95% CIs
.2
.3
.4
.5
0
2
4
6
8
10
12
PROAD_DIST


Predictive Margins with 95% CIs
.1
.2
.3
.4
0
10
20
30
40
50
60
70
80
90
SSEHOUS_PERC


Predictive Margins with 95% CIs
0
.1
.2
.3
0
10
20
30
40
50
PRIM_PERC


Predictive Margins with 95% CIs
.15
.2
.25
.3
.35
0
5
10
15
20
25
30
35
40
45
AP2015_DIST

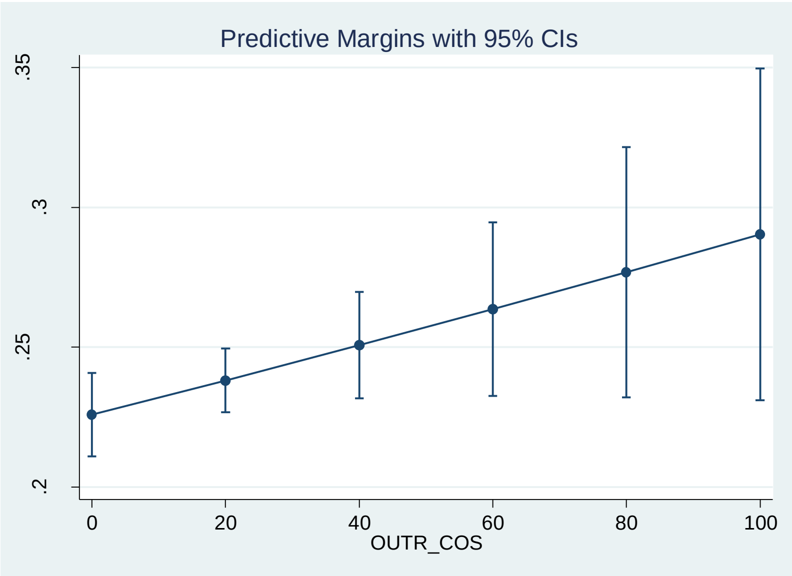
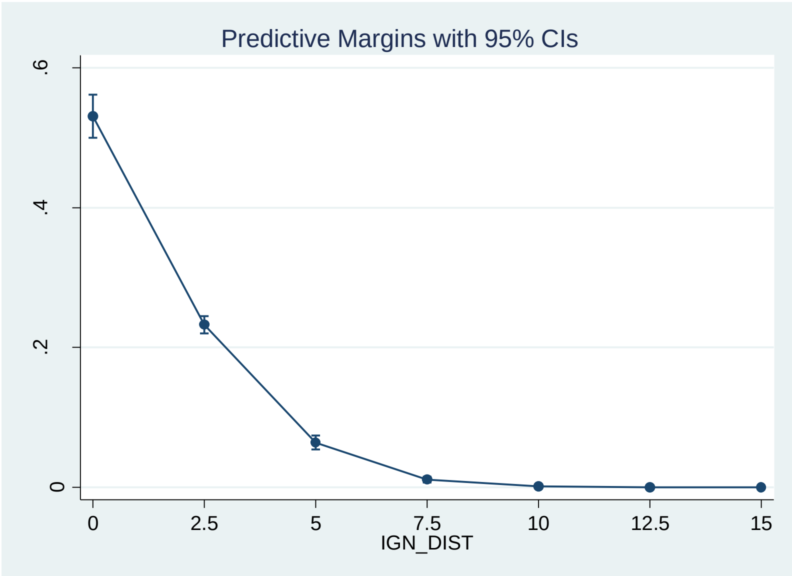
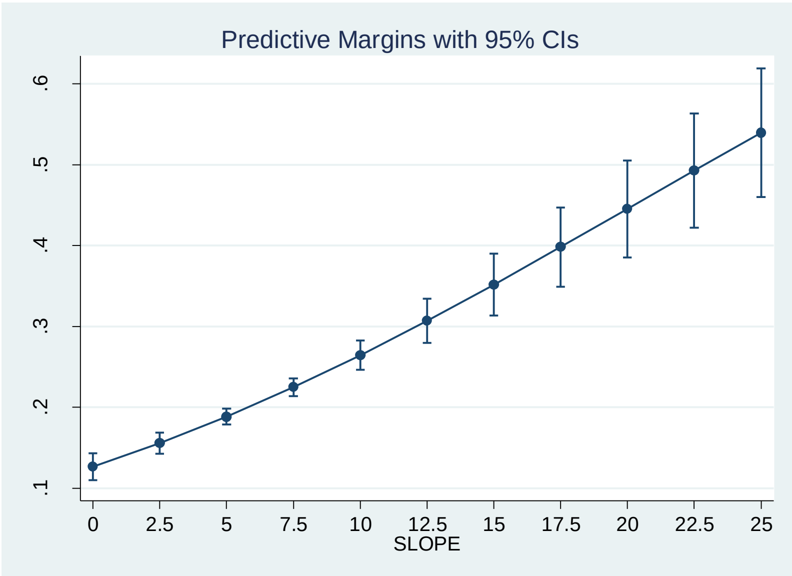
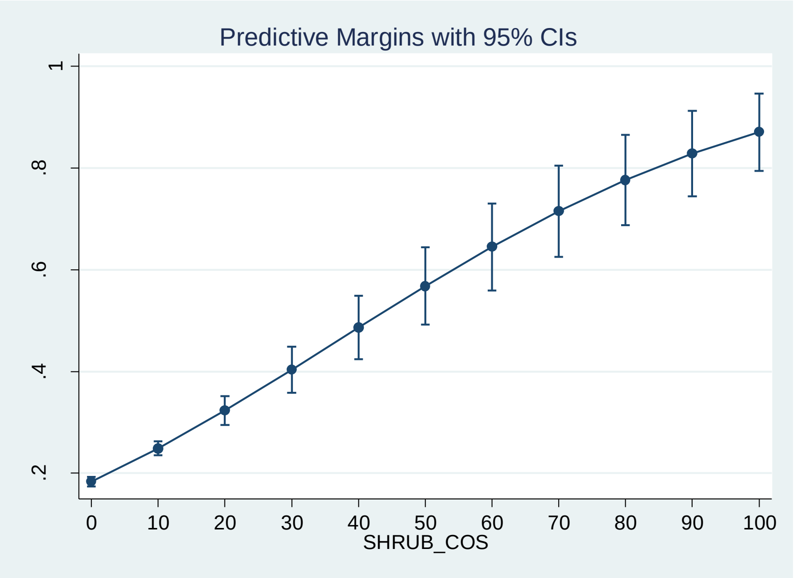
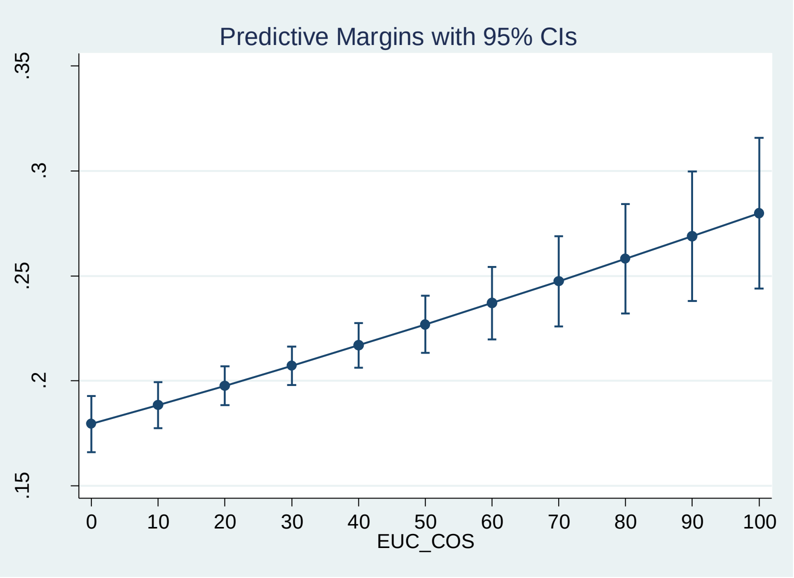

Predictive Margins with 95% CIs
.2
.25
.3
.35
0
20
40
60
80
100
OUTR_COS

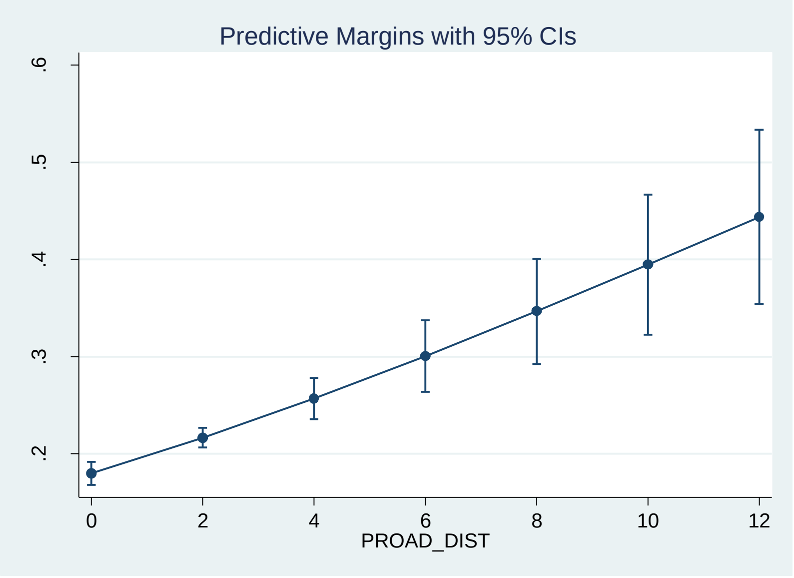
Figure D 8 – Large wildfire propagation model (first part): Cluster 2

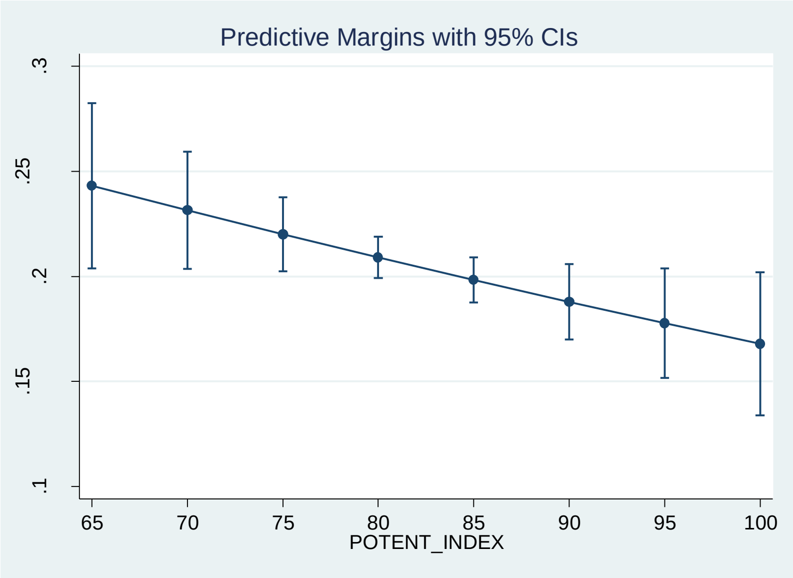


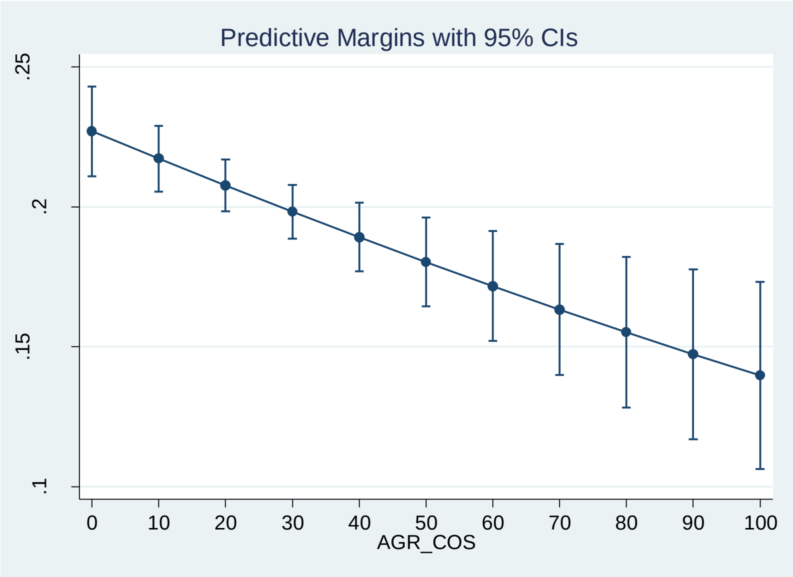


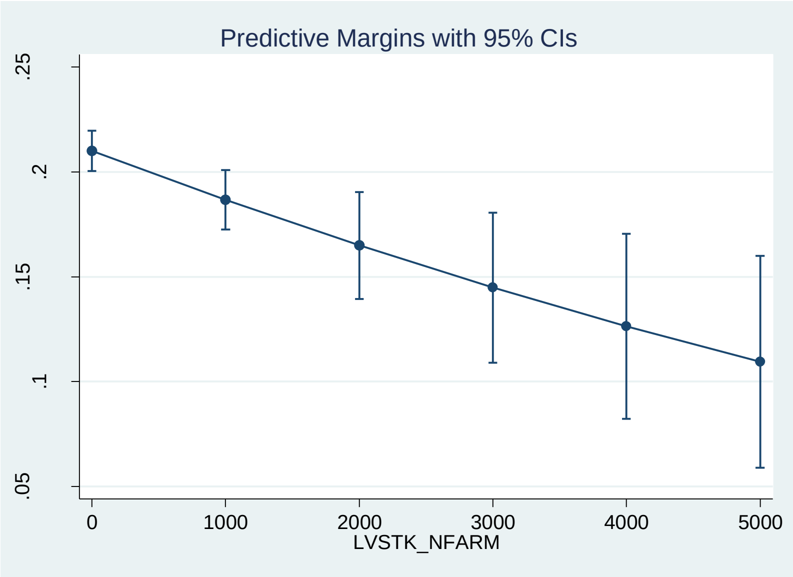

Predictive Margins with 95% CIs
.2
.4
.6
.8
1
0
10
20
30
40
50
60
70
80
90
100
SHRUB_COS

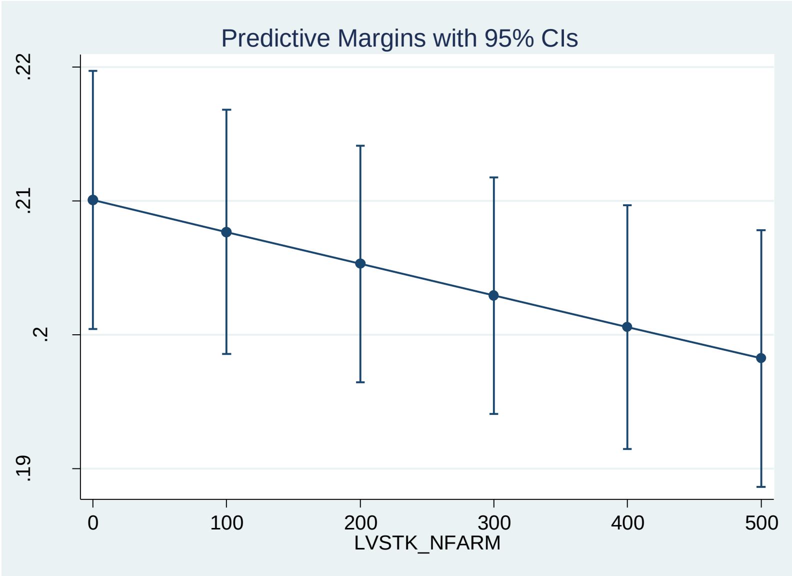

Predictive Margins with 95% CIs
.15
.2
.25
.3
.35
0
10
20
30
40
50
60
70
80
90
100
EUC_COS

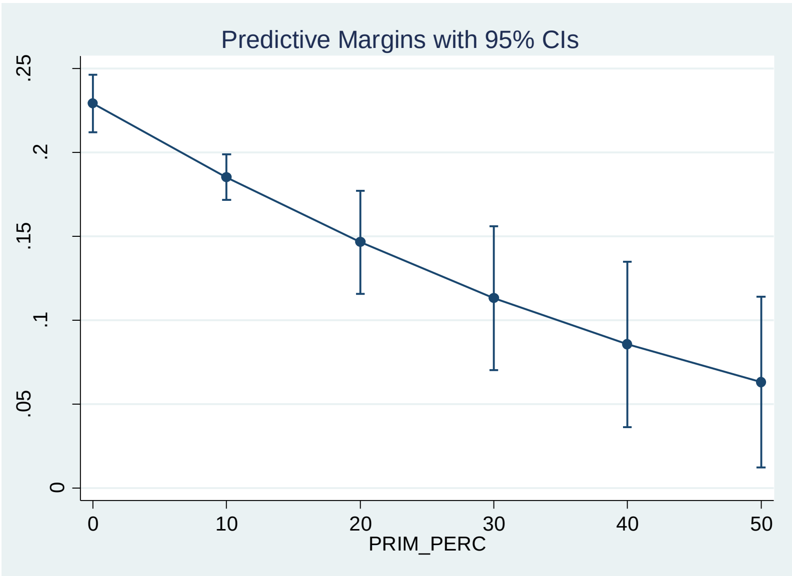

Predictive Margins with 95% CIs
.2
.3
.4
.5
.6
0
2
4
6
8
10
12
PROAD_DIST

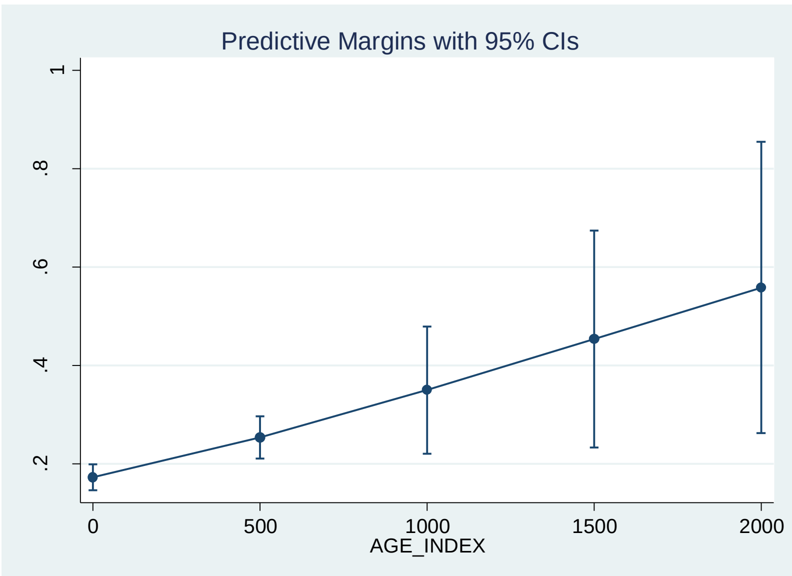

Predictive Margins with 95% CIs
.1
.15
.2
.25
.3
65
70
75
80
85
90
95
100
POTENT_INDEX

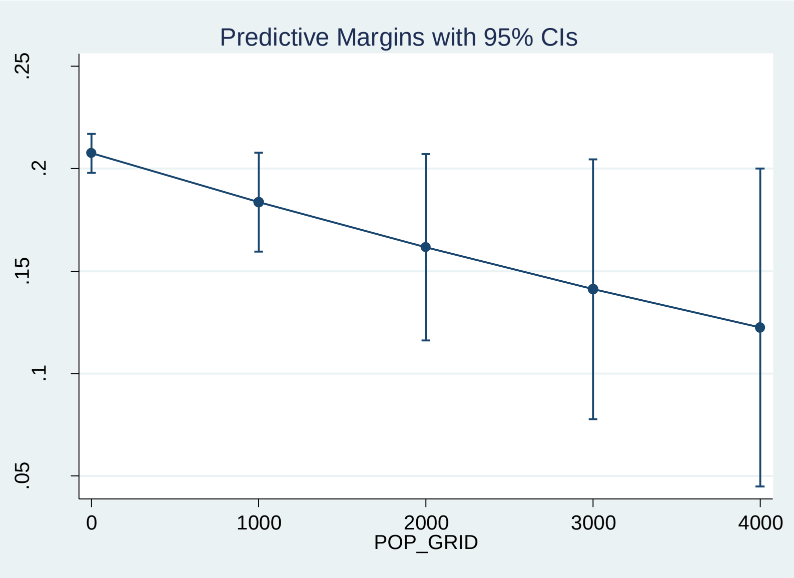

Predictive Margins with 95% CIs
.1
.15
.2
.25
0
10
20
30
40
50
60
70
80
90
100
AGR_COS


Predictive Margins with 95% CIs
.05
.1
.15
.2
.25
0
1000
2000
3000
4000
5000
LVSTK_NFARM

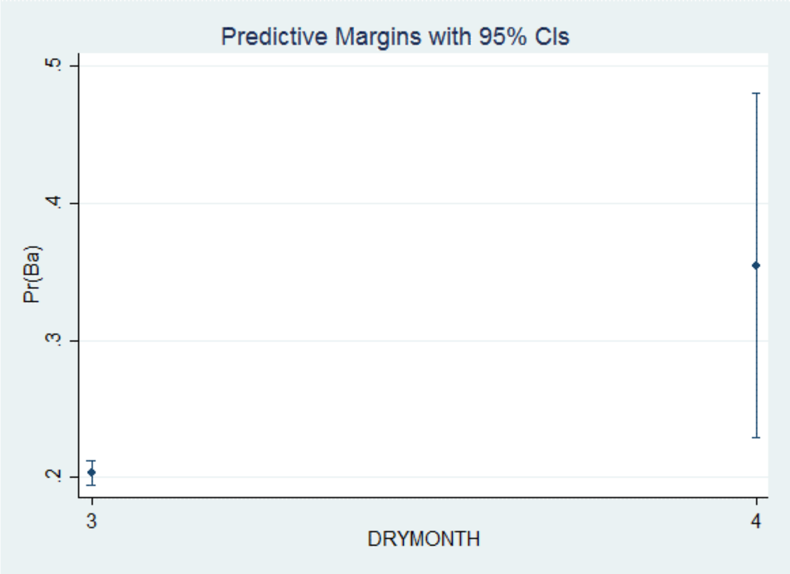

Predictive Margins with 95% CIs
.19
.2
.21
.22
0
100
200
300
400
500
LVSTK_NFARM


Predictive Margins with 95% CIs
0
.05
.1
.15
.2
.25
0
10
20
30
40
50
PRIM_PERC


Predictive Margins with 95% CIs
.2
.4
.6
.8
1
0
500
1000
1500
2000
AGE_INDEX

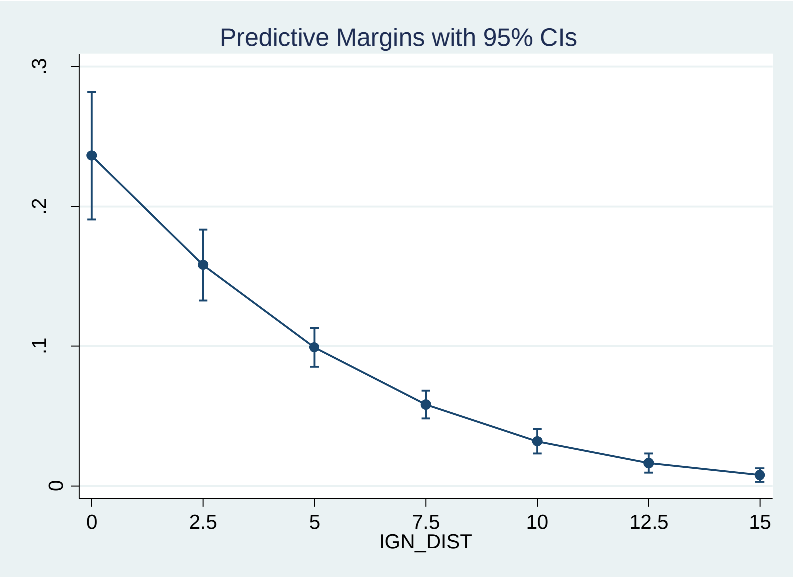


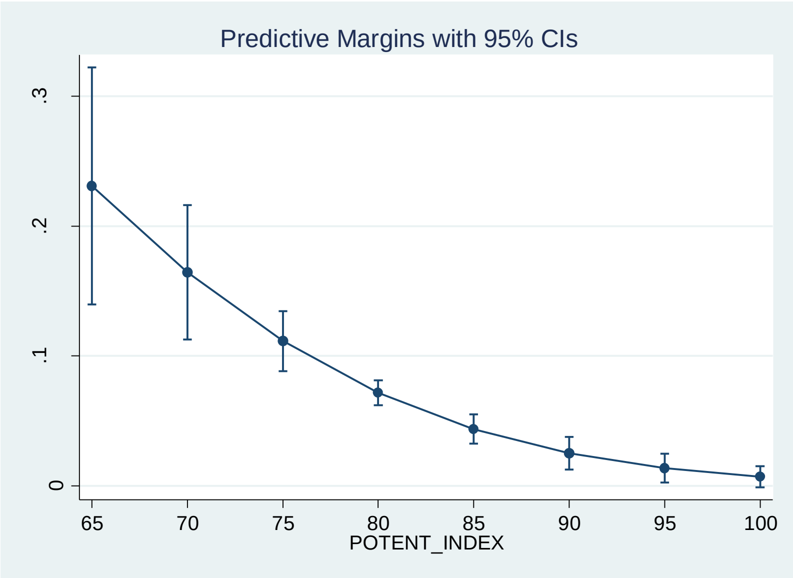
Predictive Margins with 95% CIs
.05
.1
.15
.2
.25
0
1000
2000
3000
4000
POP_GRID

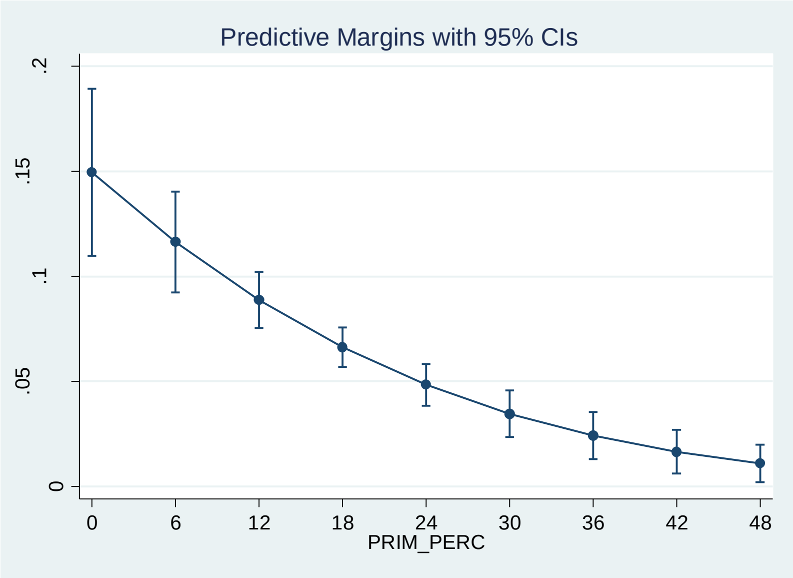


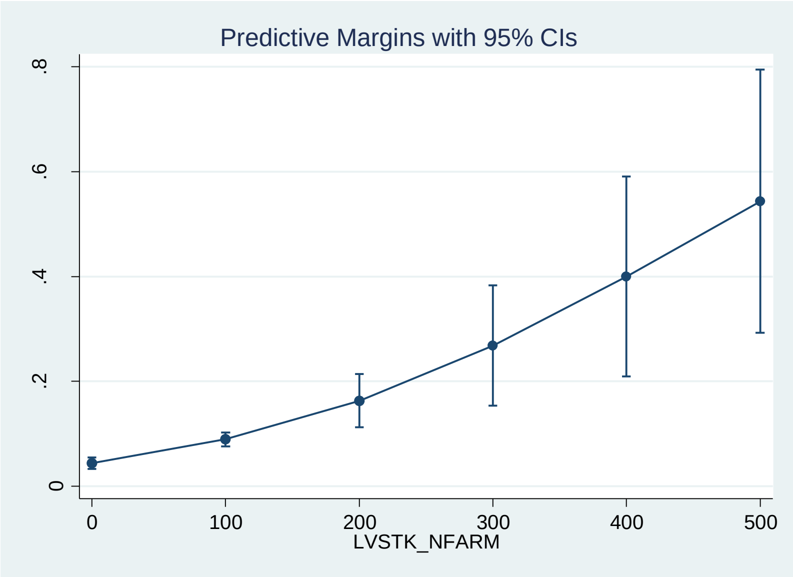
Predictive Margins with 95% CIs
Pr(Ba)
.2
.3
.4
.5
3
4
DRYMONTH

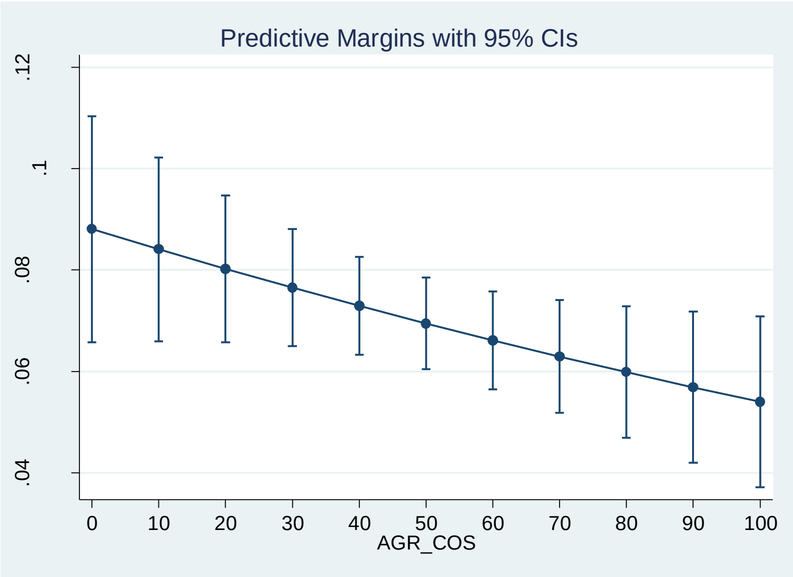
Figure D 9 – Large wildfire propagation model (first part): Cluster 3

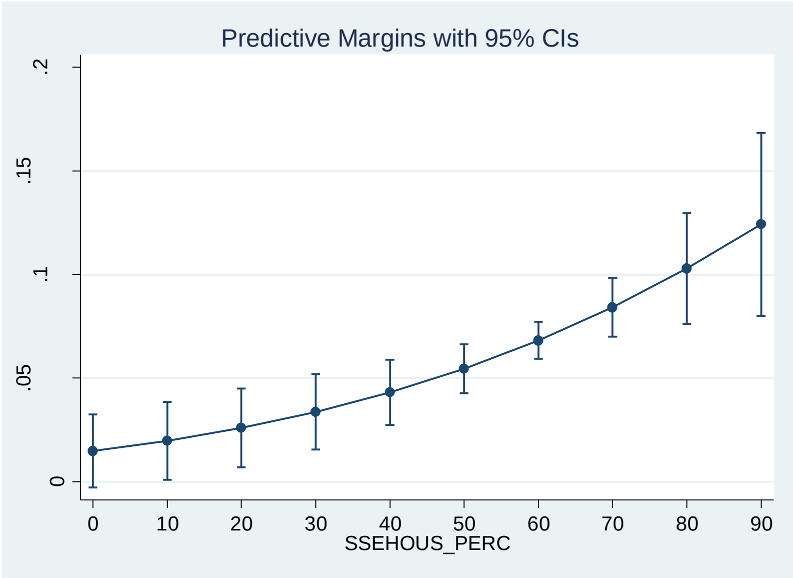




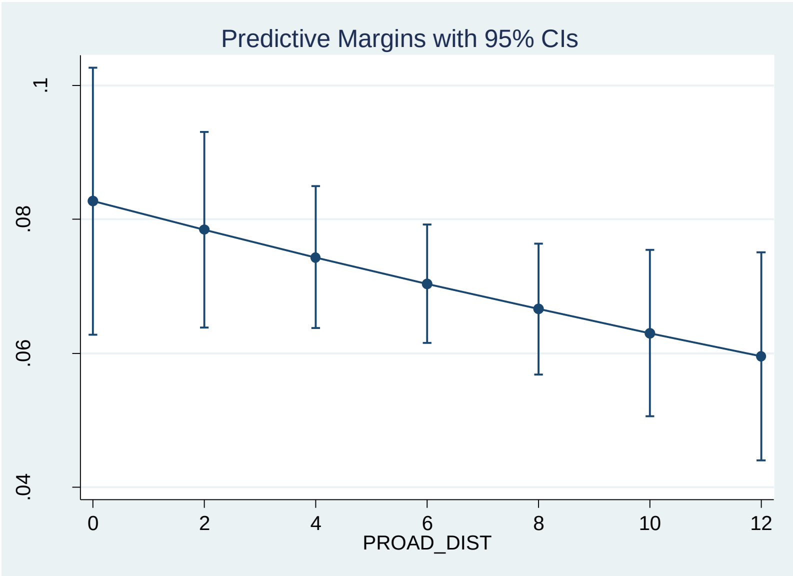

Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0
6
12
18
24
30
36
42
48
PRIM_PERC


Predictive Margins with 95% CIs
0
.2
.4
.6
.8
0
100
200
300
400
500
LVSTK_NFARM

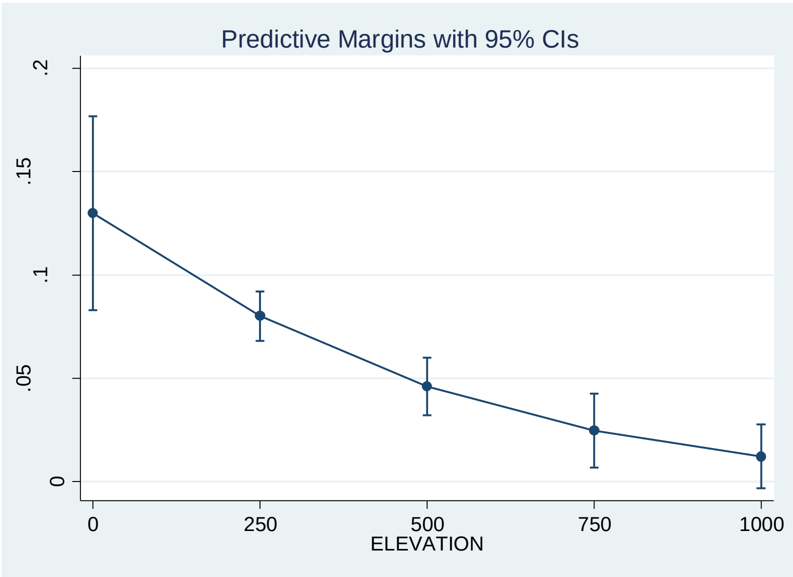

Predictive Margins with 95% CIs
.04
.06
.08
.1
.12
0
10
20
30
40
50
60
70
80
90
100
AGR_COS



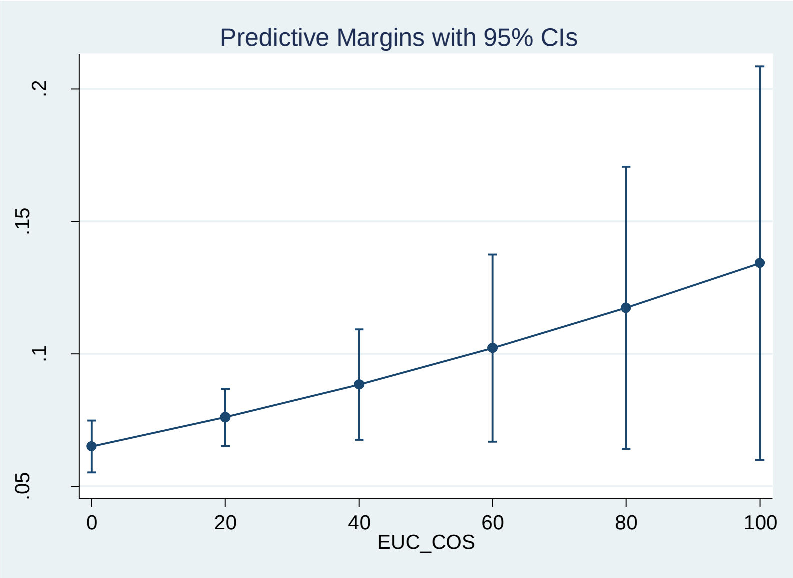
Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0
10
20
30
40
50
60
70
80
90
SSEHOUS_PERC

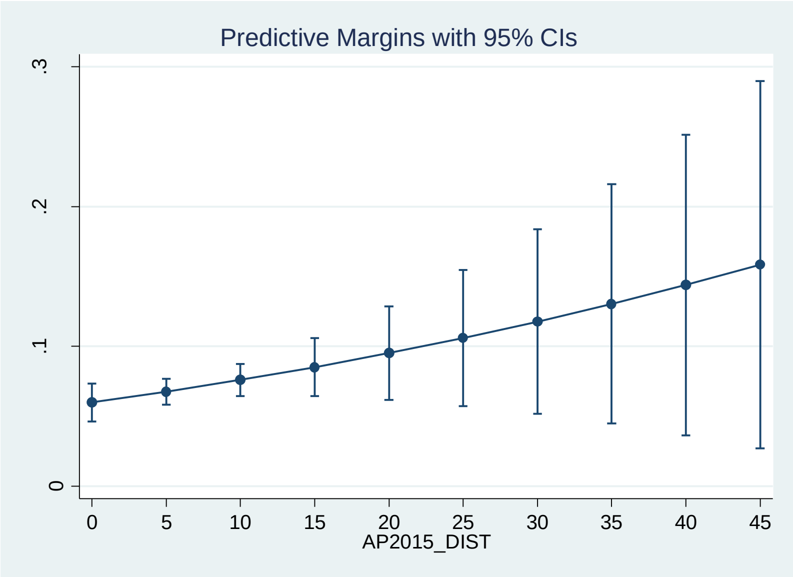


Predictive Margins with 95% CIs
.04
.06
.08
.1
0
2
4
6
8
10
12
PROAD_DIST



Predictive Margins with 95% CIs
0
.05
.1
.15
.2
0
250
500
750
1000
ELEVATION


Predictive Margins with 95% CIs
.05
.1
.15
.2
0
20
40
60
80
100
EUC_COS


Predictive Margins with 95% CIs
0
.1
.2
.3
0
5
10
15
20
25
30
35
40
45
AP2015_DIST

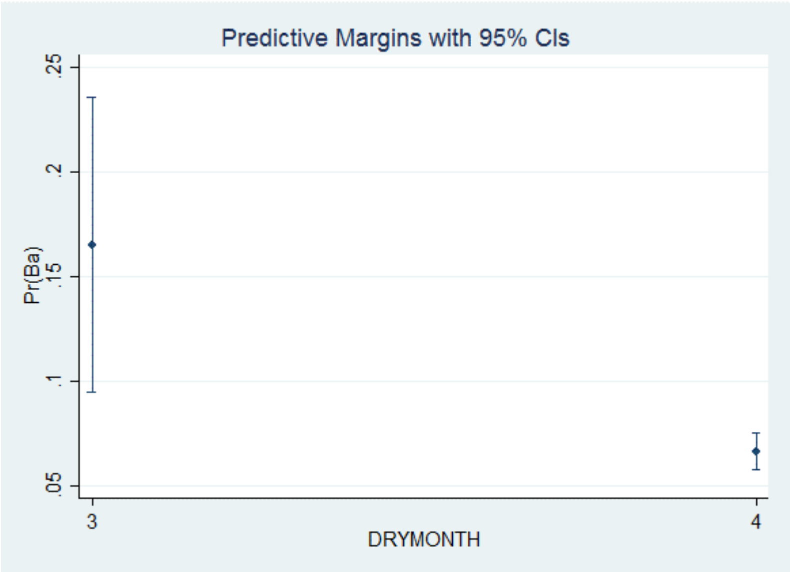
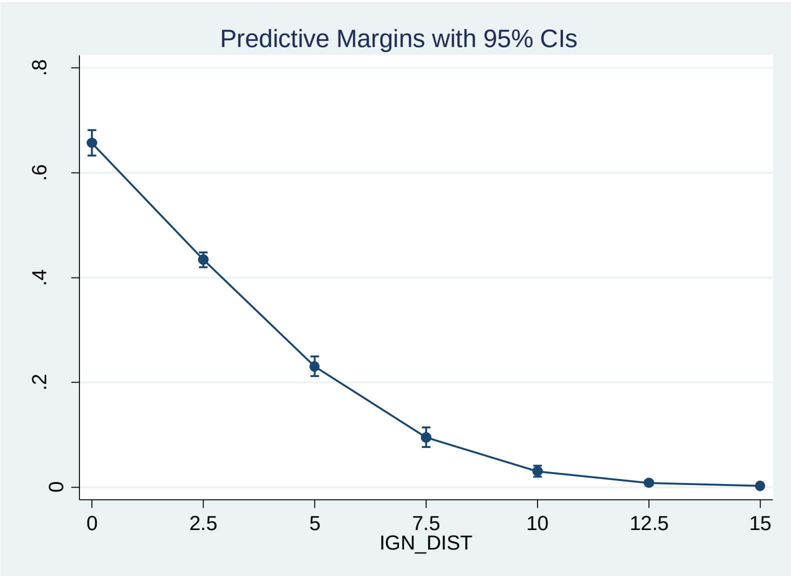
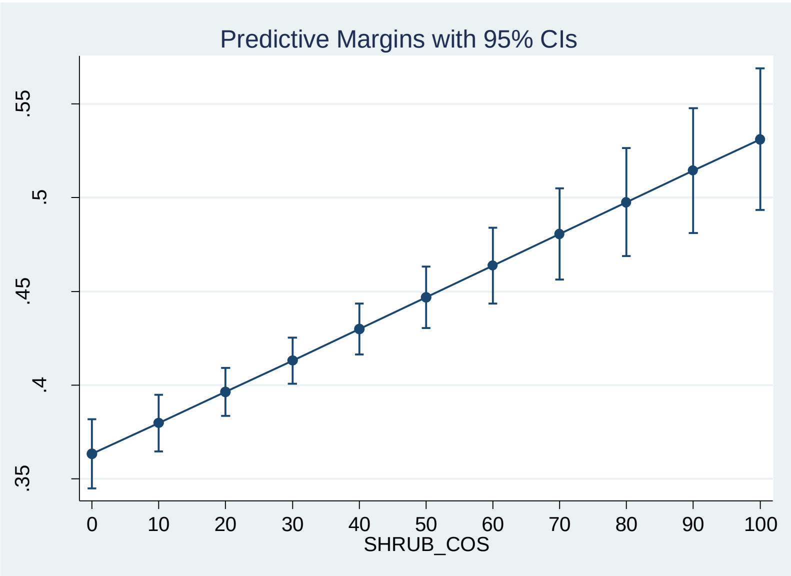
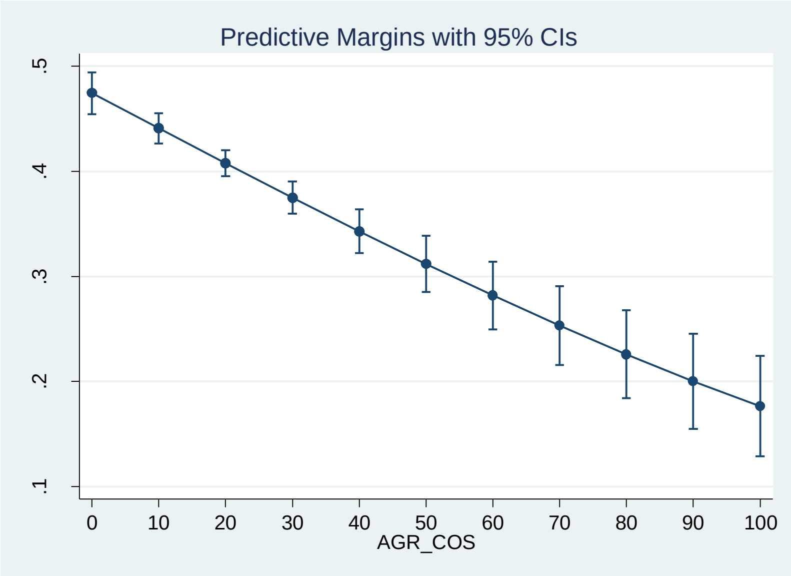
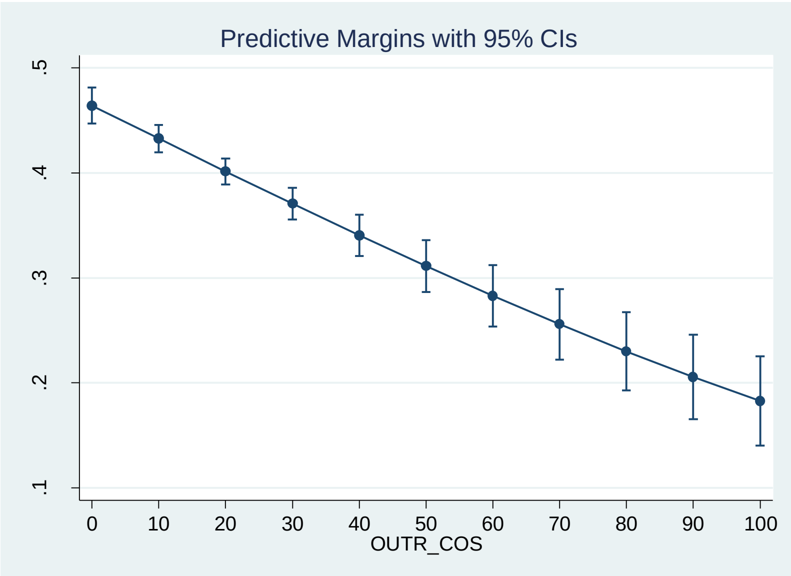

Predictive Margins with 95% CIs
Pr(Ba)
.05
.1
.15
.2
.25
3
4
DRYMONTH

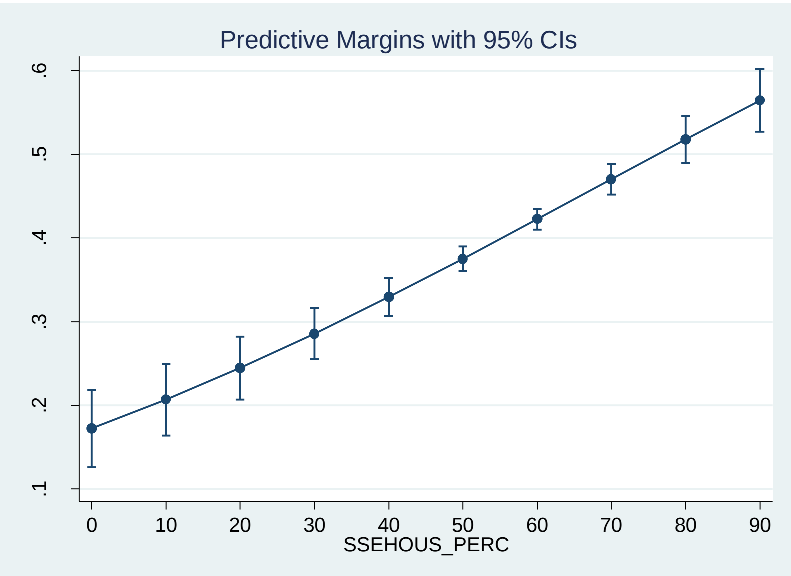
Figure D 10 – Large wildfire propagation model (first part): Cluster 4

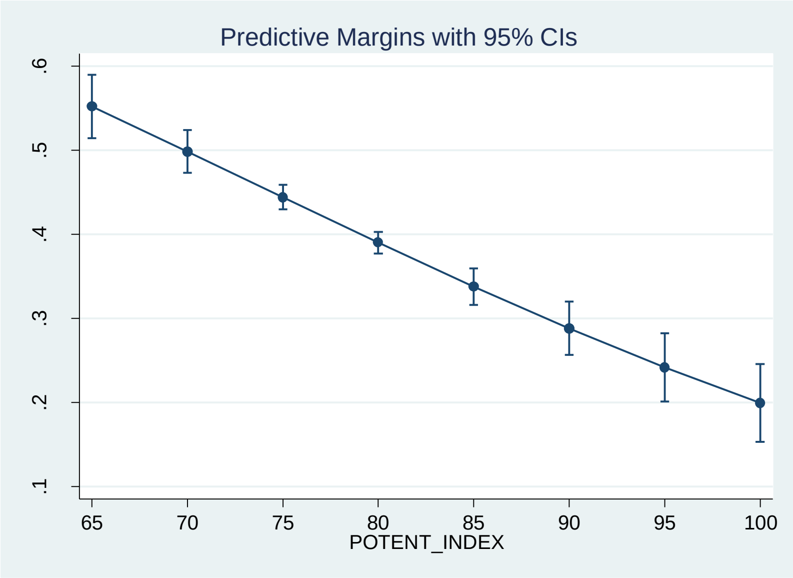


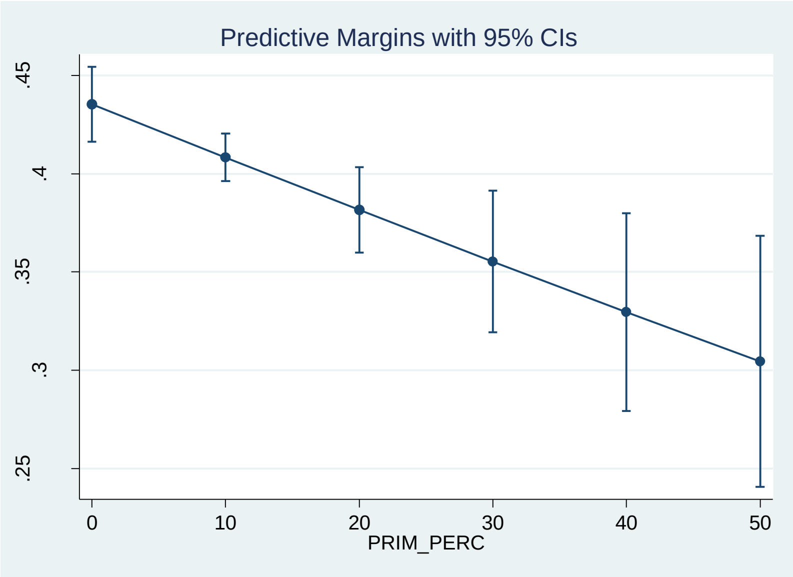


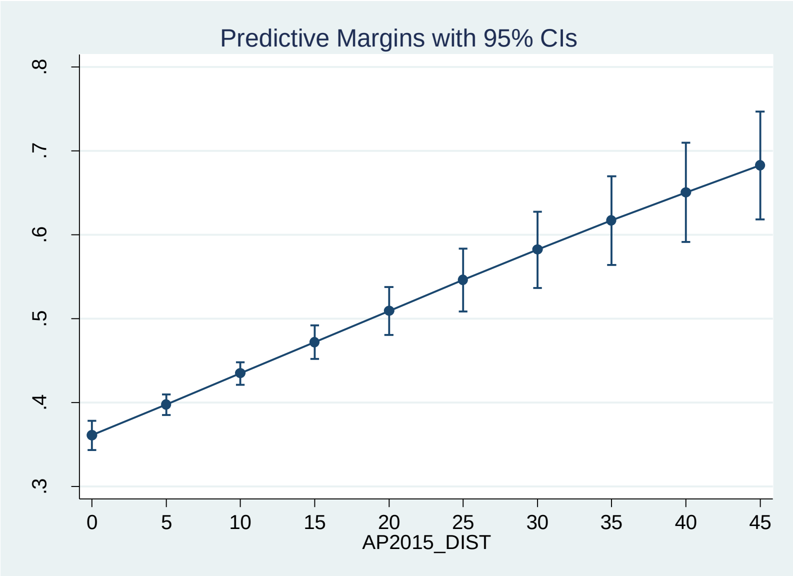

Predictive Margins with 95% CIs
.1
.2
.3
.4
.5
0 10 20 30 40 50 60 70 80 90 100
AGR_COS

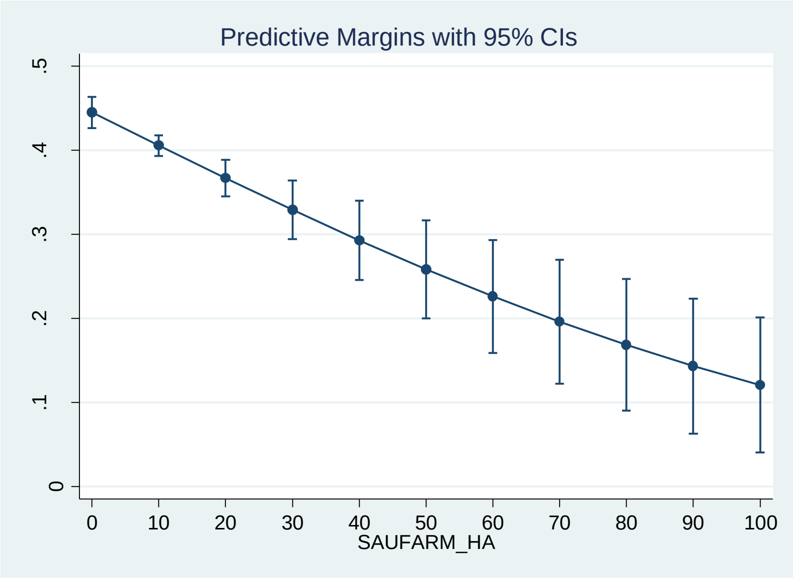

Predictive Margins with 95% CIs
.1
.2
.3
.4
.5
0 10 20 30 40 50 60 70 80 90 100
OUTR_COS

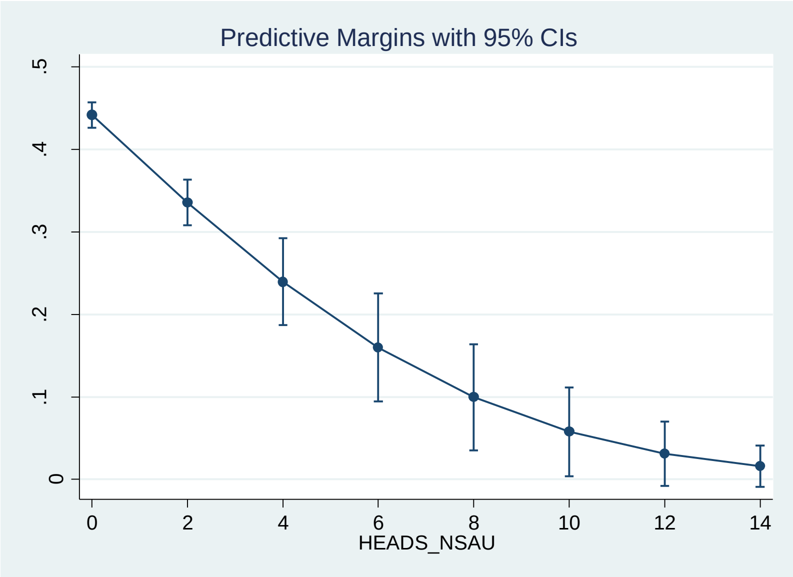

Predictive Margins with 95% CIs
.1
.2
.3
.4
.5
.6
0 10 20 30 40 50 60 70 80 90
SSEHOUS_PERC

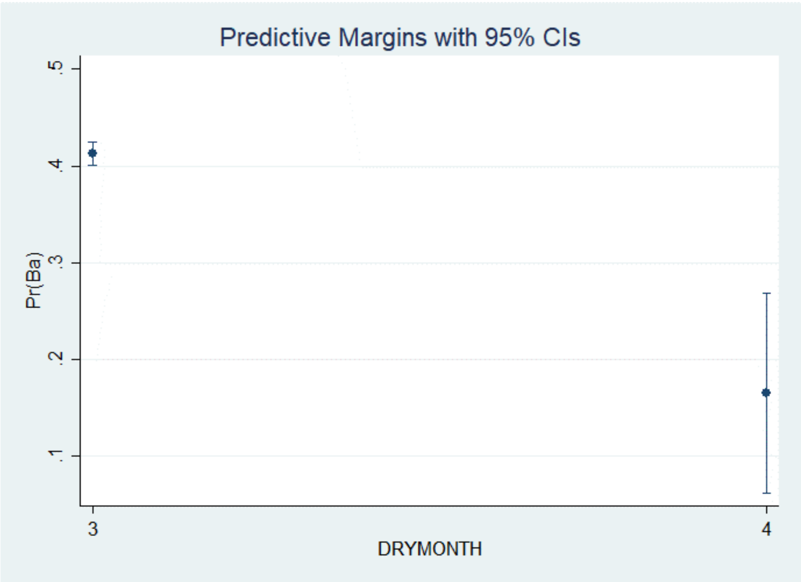

Predictive Margins with 95% CIs
.1
.2
.3
.4
.5
.6
65
70
75
80
85
90
95
100
POTENT_INDEX

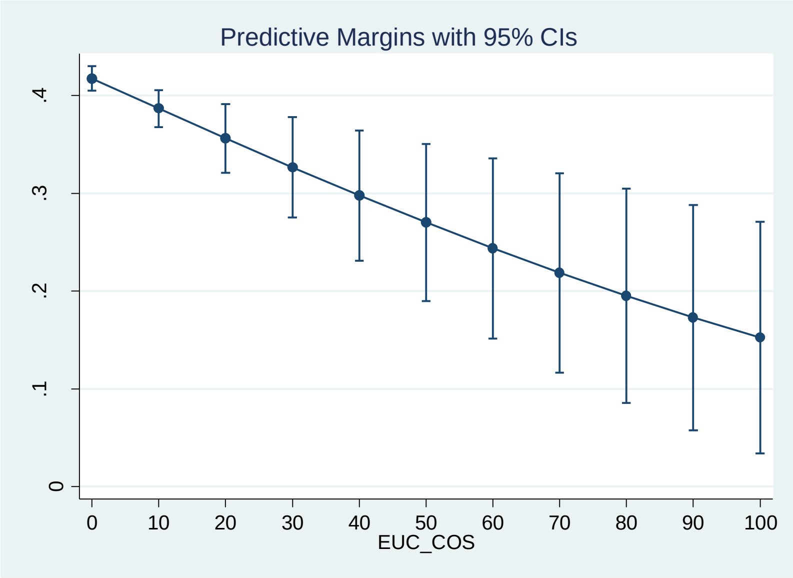

Predictive Margins with 95% CIs
.25
.3
.35
.4
.45
0
10
20
30
40
50
PRIM_PERC

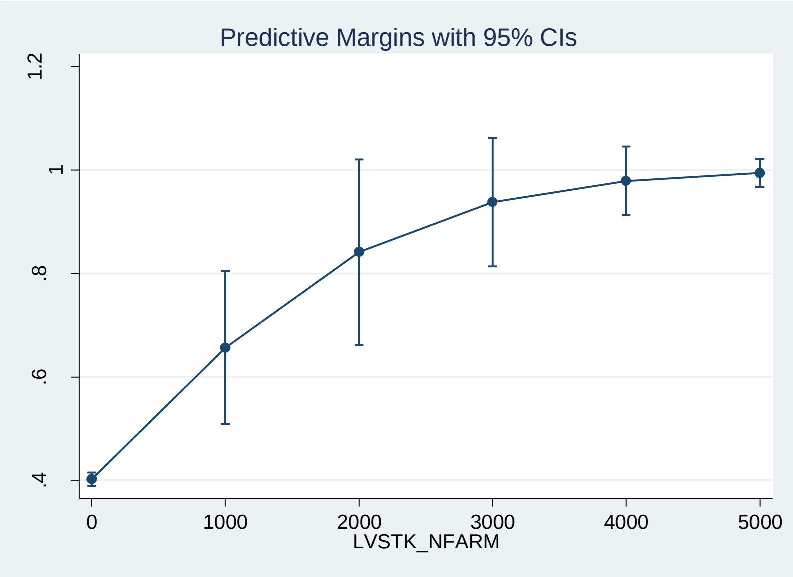

Predictive Margins with 95% CIs
.3
.4
.5
.6
.7
.8
0
5
10
15
20
25
30
35
40
45
AP2015_DIST

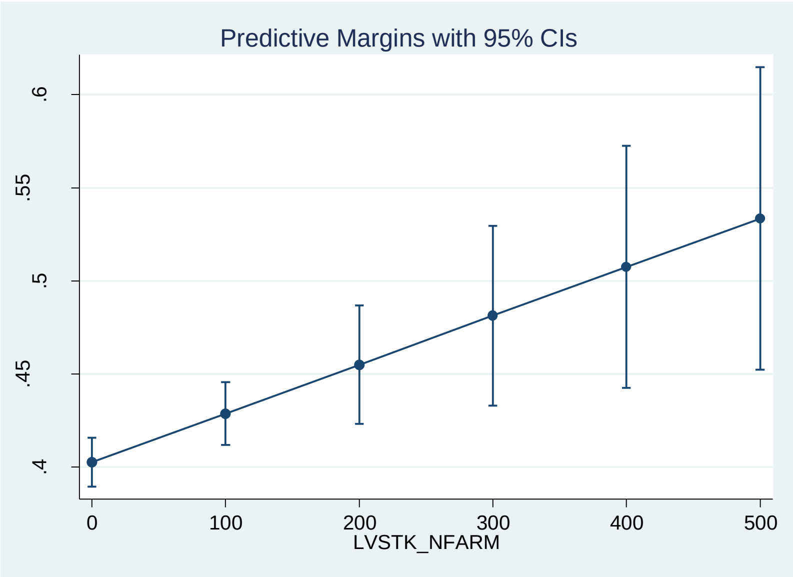

Predictive Margins with 95% CIs
0
.1
.2
.3
.4
.5
0
10
20
30
40
50
60
70
80
90
100
SAUFARM_HA

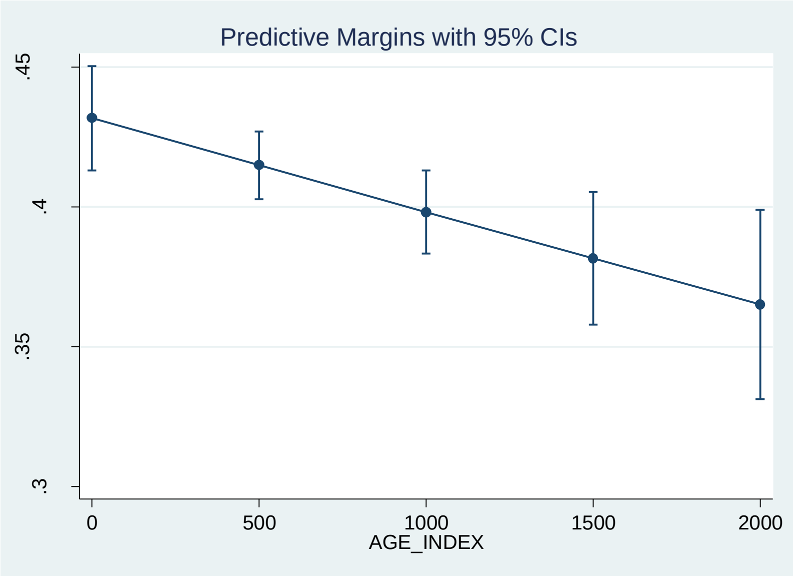

Predictive Margins with 95% CIs
0
.1
.2
.3
.4
.5
0
2
4
6
8
10
12
14
HEADS_NSAU


Predictive Margins with 95% CIs
Pr(Ba)
.1
.2
.3
.4
.5
3
4
DRYMONTH


Predictive Margins with 95% CIs
0
.1
.2
.3
.4
0
10
20
30
40
50
60
70
80
90
100
EUC_COS


Predictive Margins with 95% CIs
.4
.6
.8
1
1.2
0
1000
2000
3000
4000
5000
LVSTK_NFARM


Predictive Margins with 95% CIs
.4
.45
.5
.55
.6
0
100
200
300
400
500
LVSTK_NFARM


Predictive Margins with 95% CIs
.3
.35
.4
.45
0
500
1000
1500
2000
AGE_INDEX

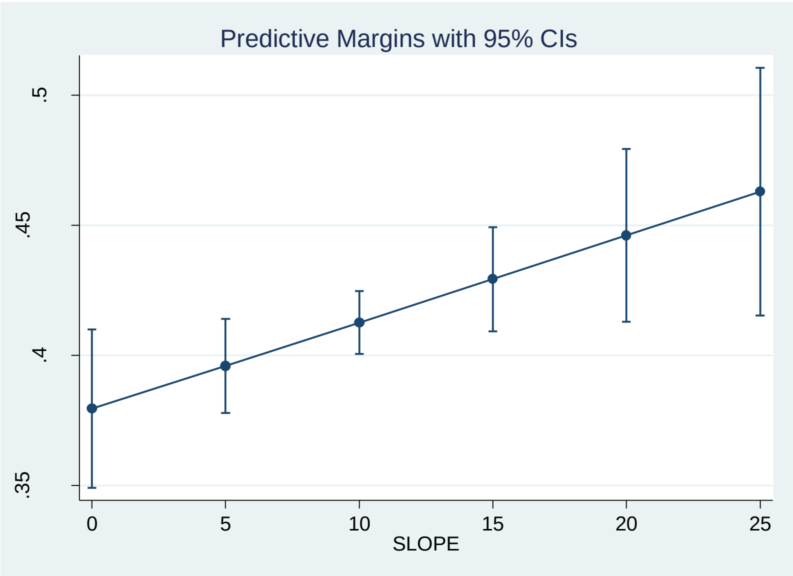
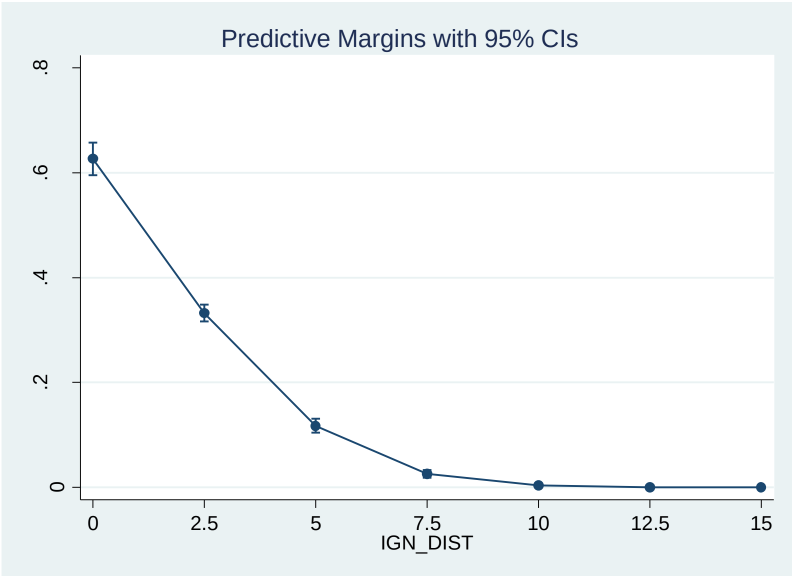
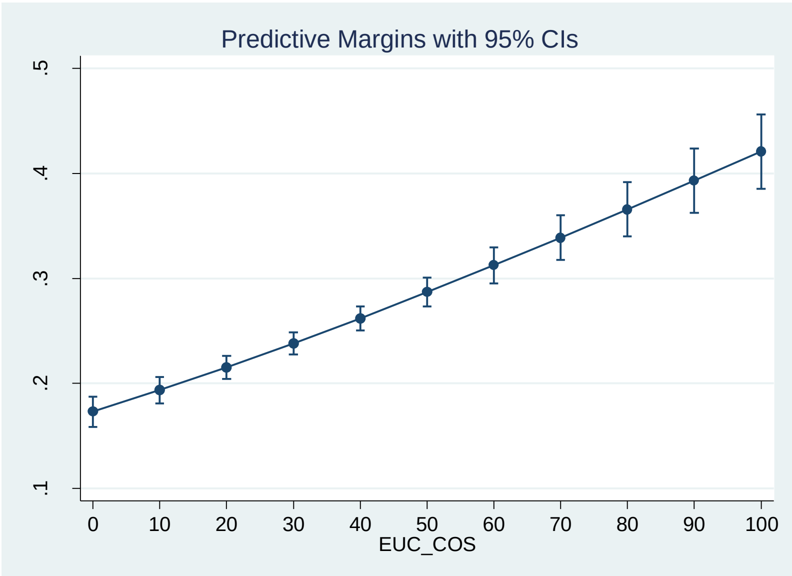
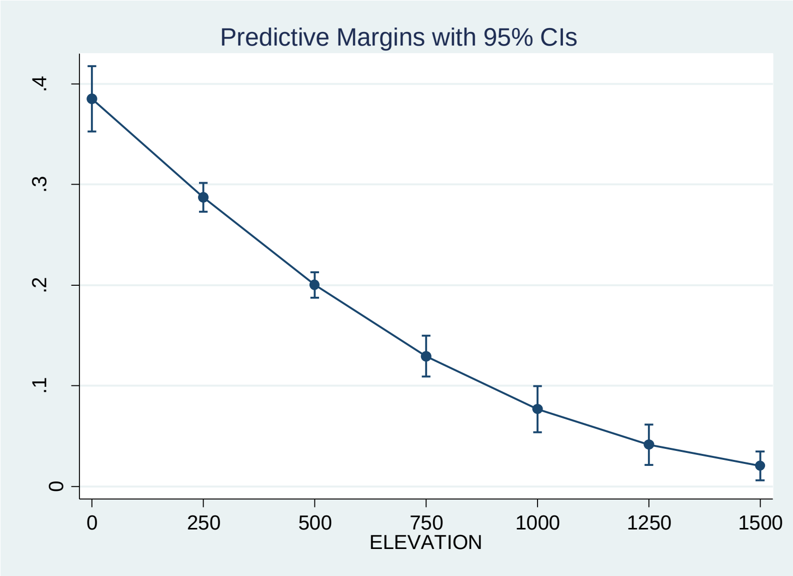

Predictive Margins with 95% CIs
.35
.4
.45
.5
0
5
10
15
20
25
SLOPE

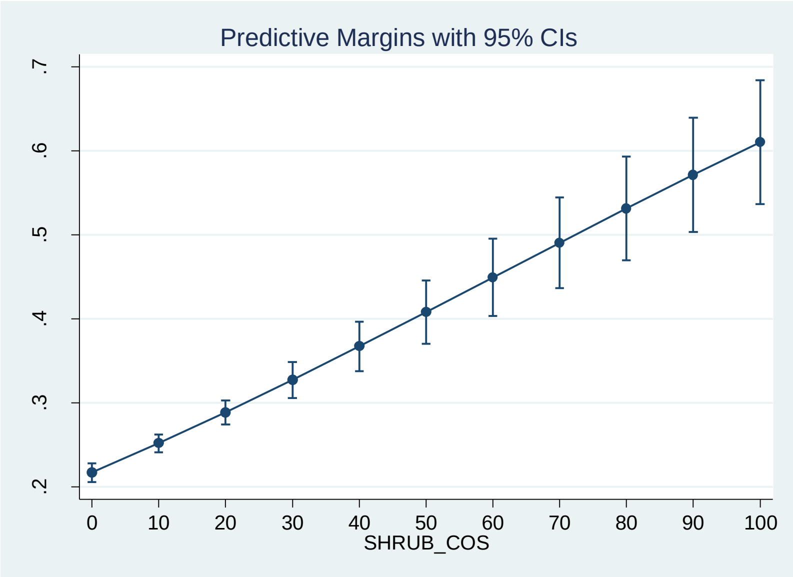
Figure D 11 – Large wildfire propagation model (first part): Cluster 5



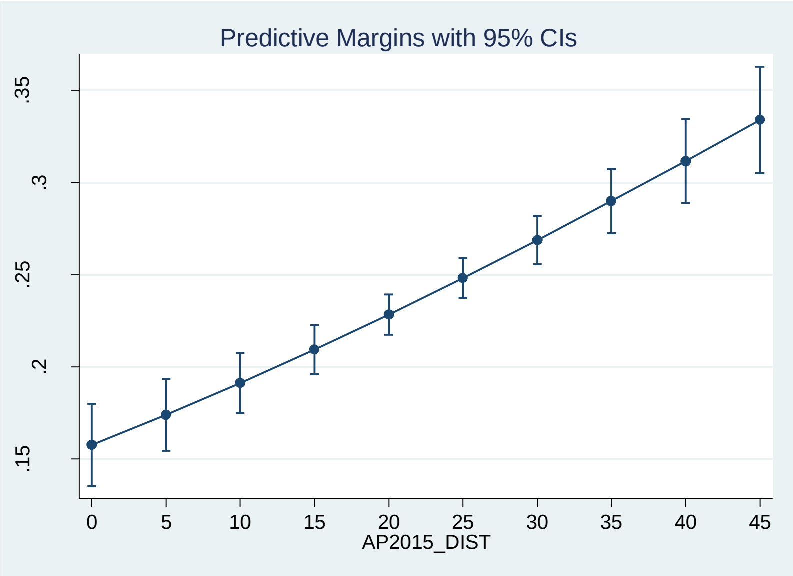




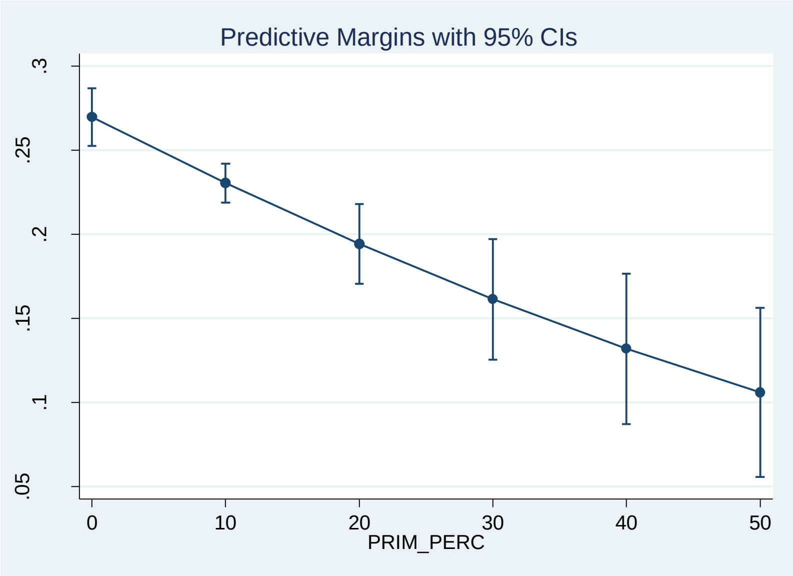
Predictive Margins with 95% CIs
0
.1
.2
.3
.4
0
250
500
750
1000
1250
1500
ELEVATION

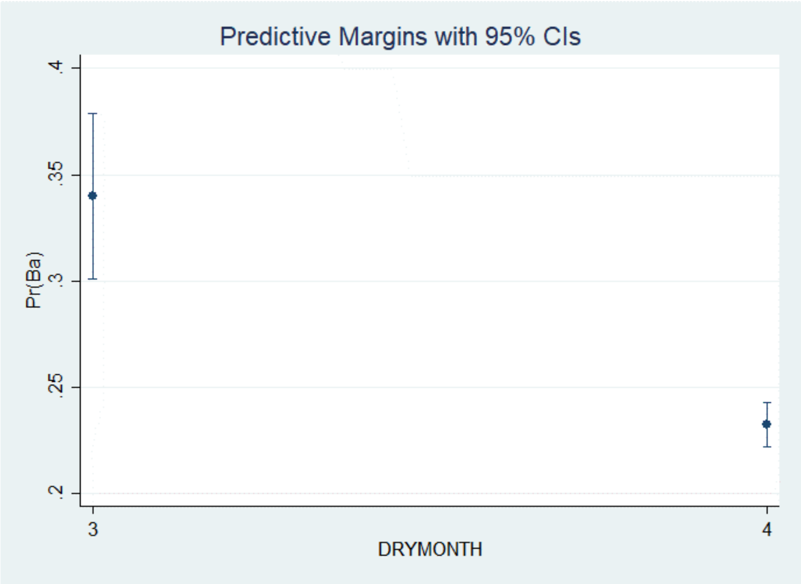


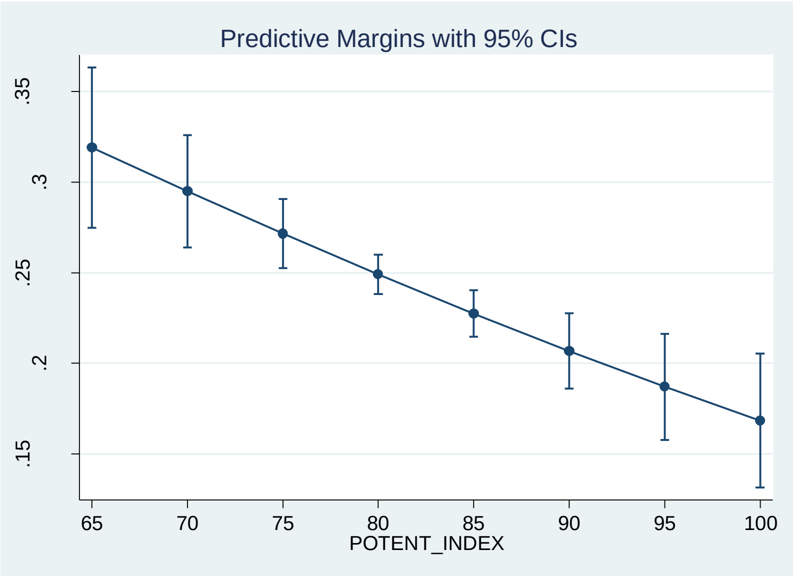
Predictive Margins with 95% CIs
.2
.3
.4
.5
.6
.7
0
10
20
30
40
50
60
70
80
90
100
SHRUB_COS

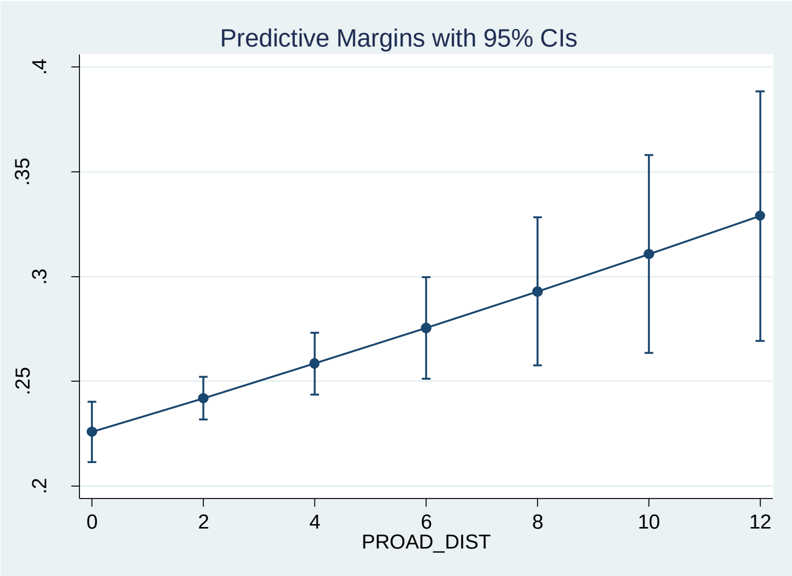


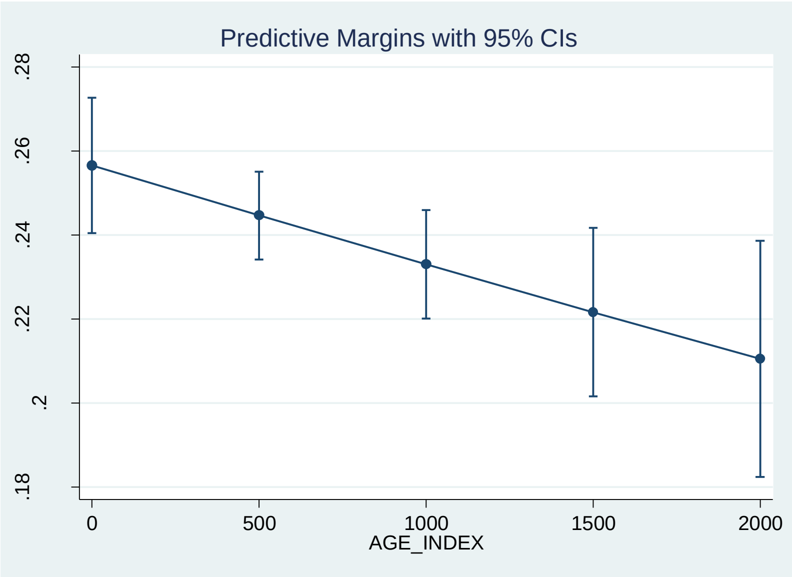
Predictive Margins with 95% CIs
.15
.2
.25
.3
.35
0
5
10
15
20
25
30
35
40
45
AP2015_DIST

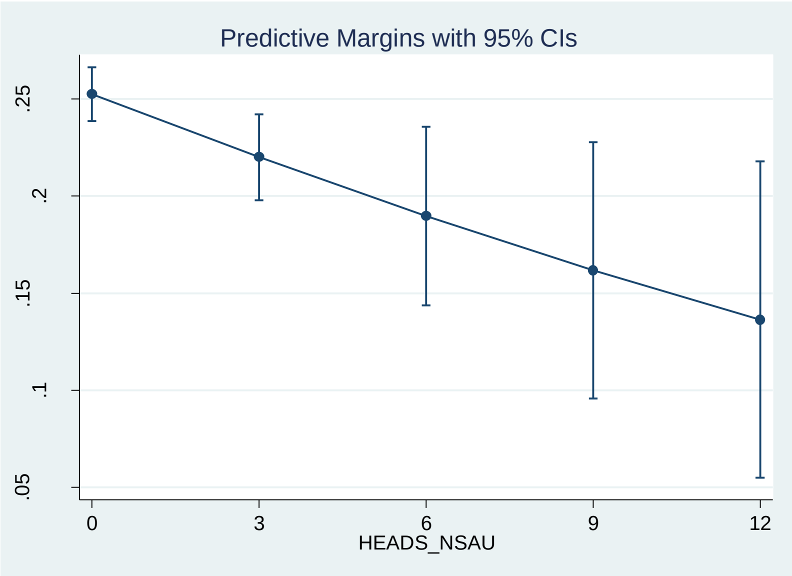

Predictive Margins with 95% CIs
.05
.1
.15
.2
.25
.3
0
10
20
30
40
50
PRIM_PERC


Predictive Margins with 95% CIs
Pr(Ba)
.2
.25
.3
.35
.4
3
4
DRYMONTH


Predictive Margins with 95% CIs
.15
.2
.25
.3
.35
65
70
75
80
85
90
95
100
POTENT_INDEX


Predictive Margins with 95% CIs
.2
.25
.3
.35
.4
0
2
4
6
8
10
12
PROAD_DIST


Predictive Margins with 95% CIs
.18
.2
.22
.24
.26
.28
0
500
1000
1500
2000
AGE_INDEX


Predictive Margins with 95% CIs
.05
.1
.15
.2
.25
0
3
6
9
12
HEADS_NSAU