



Hinc patriam sustinet

Instituto Superior de Agronomia
Universidade Técnica de Lisboa

DETERMINANTES GEOGRÁFICOS DA LOCALIZAÇÃO DO PERÍMETRO DAS ÁREAS QUEIMADAS

Alicia Horta Ameneiro

Dissertação para obtenção do Grau de Mestre em
Engenharia Florestal e dos Recursos Naturais

Orientador: Doutor José Miguel Oliveira Cardoso Pereira

Co-orientador: Doutor João Manuel das Neves Silva

Júri:

Presidente: Doutor Francisco Manuel Cardoso de Castro Rego, Professor Associado do Instituto Superior de Agronomia da Universidade Técnica de Lisboa.

Vogais: Doutor José Miguel Oliveira Cardoso Pereira, Professor Catedrático do Instituto Superior de Agronomia da Universidade Técnica de Lisboa;

Doutora Maria Helena Reis de Noronha Ribeiro de Almeida, Professora Associada do Instituto Superior de Agronomia da Universidade Técnica de Lisboa;

Doutor Bernardo Wildung Cantante Mota.

Lisboa, 2013

À minha família.

“Vale máis unha terra con arvres nos montes que un estado con ouro nos bancos.”

A. D. M. R. Castelao (1886-1950)

AGRADECIMENTOS.

“Cada pessoa que passa na nossa vida passa sozinha, porque cada pessoa é única e nenhuma substitui outra. Cada pessoa que passa pela nossa vida passa sozinha, não nos deixa só, porque deixa um pouco de si e leva um pouquinho de nós. Essa é a mais bela responsabilidade da vida e a prova de que as pessoas não se encontram por acaso.”

Charles Chaplin (1889-1977).

Com mais uma etapa importante na minha vida, gostaria de expressar o meu agradecimento a todos aqueles que me apoiaram nesta caminhada e contribuíram para a realização deste trabalho, especialmente a todos que fizeram com que a minha estadia além Minho fosse possível e inesquecível.

Agradecer ao meu Orientador, José Miguel Cardoso Pereira, pela oportunidade e sugerência de realizar esta Dissertação, pelo seu conhecimento, interesse e tempo dedicado.

Ao meu Co-orientador, João Silva, pelo apoio, dedicação, paciência, companheirismo e amizade fornecidos durante esta fase.

À equipa do Laboratório de Detecção Remota e Análise Geográfica do Instituto Superior de Agronomia, pelas interrupções feitas, pela ajuda e esclarecimentos prestados, especialmente ao Duarte Oom.

Ao grupo do Professor José Guilherme Calvão Borges, pela informação disponibilizada.

Obrigado aos colegas e amigos que me apoiaram e ajudaram durante estes anos, estando mais perto duns e mais longe de outros.

Quero agradecer aos meus pais, irmáns e avoa por dedicar e compartilhar as súas vidas conmigo. Polo seu apoio, agarimo e paciencia que sempre me demostraron en cada fase da vida. Pola forza, loita e ideais que me inculcaron desde nova. Grazas aos máis pequenos da familia por eses momentos de felicidade naqueles intres de baixa autoestima.

A todos os que merecem o meu agradecimento e não foram citados.

Aos que non viron concluída esta fase da miña carreira.

RESUMO

O fogo é um processo natural e um elemento integrante de alguns ecossistemas pelo que se considera um fator fundamental para garantir a biodiversidade destas áreas. Pode, no entanto ser um fator negativo para as nossas florestas, pois nas últimas décadas a sua incidência tem aumentado de forma drástica nos países Mediterrânicos, sobretudo devido ao abandono rural e às condições climáticas.

Em particular, Portugal verifica um grande aumento da área ardida nos primeiros anos do século XXI em relação ao fim do século XX. Sendo um país com elevada cobertura vegetal e com grande importância derivada do setor florestal, é preciso realizar os possíveis para manter o futuro das suas florestas.

A análise das fronteiras das áreas ardidas permite o estudo da correlação entre a distribuição espacial do fogo e as variáveis ambientais, procurando em particular explicar o facto de algumas zonas do território se manterem frequentemente como fronteiras de áreas ardidas. Será este o objetivo principal do estudo, baseado na exploração das variáveis mais significativas para explicar a localização das fronteiras das áreas ardidas em Portugal continental.

Mediante a realização da análise estatística de um conjunto de dados, referentes aos grandes incêndios ocorridos em Portugal no período de 1996 a 2005, tem-se provado a correlação entre a ocupação do solo e as fronteiras do fogo. O método estatístico usado são as árvores de classificação, comunmente usadas para obter modelos de predição precisos. Os resultados da análise demonstram que a maior parte das fronteiras dos fogos ocorrem em zonas de matos, coincidindo com a preferência do próprio fogo por estas áreas.

Palavras chave: incêndios florestais, gestão florestal, fronteiras do fogo, árvores de classificação.

RESUMO

O fogo é un proceso natural e un elemento que integra algúns ecosistemas, considérase un factor fundamental para garantir a biodiversidade destas áreas. Aínda así, é un factor negativo para os nosos montes, pois nas últimas décadas aumentou a súa incidencia de forma drástica nos países Mediterráneos debido, sobretudo, ao abandono rural e ás condicións climáticas.

En particular, Portugal verifica un grande aumento de área ardida nos primeiros anos do século XXI en relación ao fin do século XX. Tratándose dun país com elevada cobertura forestal e con grande importancia derivada do sector forestal, é preciso realizar o posíbel para manter o futuro dos seus montes.

A análise das fronteiras das áreas ardidadas permite estudar a correlación entre a distribución espacial do fogo e as variábeis ambientais, procurando particularmente explicar o feito de que algunhas zonas se manteñan frecuentemente como fronteiras de áreas ardidadas. Será este o obxectivo principal do estudo, baseado na exploración das variábeis máis significativas para proceder á explicación da localización das fronteiras das áreas ardidadas en Portugal continental.

Mediante a realización de análise estatística dun conxunto de datos, referentes aos grandes incendios ocorridos en Portugal no período de 1996 a 2005, tense provado a correlación entre a ocupación do solo e as fronteiras do fogo. O método estatístico usado son as árbores de clasificación, usadas para obter modelos de predición precisos. Os resultados da análise demostran que a maior parte das fronteiras dos fogos se dá en zonas de matogueira, coincidindo coa preferéncia do propio fogo por estas áreas.

Palabras chave: incendios forestais, xestión forestal, fronteiras do fogo, árbores de clasificación.

RESUMEN

El fuego es un proceso natural que integra algunos ecosistemas, se considera un factor fundamental para garantizar la biodiversidad de estas áreas. Aún así, puede ser un factor poco favorable para nuestros bosques, ya que en las últimas décadas ha aumentado la incidencia en los países Mediterráneos, sobretudo debido al abandono rural y a las condiciones climáticas.

En particular, Portugal demuestra un gran aumento de área ardida en los primeros años del siglo XXI en relación al fin del siglo XX. Tratándose de un país con elevada cubierta forestal y con gran importancia derivada del sector forestal, es necesario realizar lo posible para mantener el futuro de sus bosques.

El análisis de las fronteras de las áreas ardidadas permite estudiar la correlación existente entre la distribución espacial del fuego y las variables ambientales, y particularmente explicar el hecho de que algunas zonas se mantengan en muchos casos como frontera del

fuego. Este será el objetivo principal del estudio, basado en la exploración de las variables más significativas para interpretar la localización de dichas fronteras en Portugal continental.

Mediante la realización del análisis estadístico de un conjunto de datos, referentes a los grandes incendios ocurridos en Portugal en el período de 1996 a 2005, se provó la correlación existente entre la ocupación del suelo y las fronteras del fuego. El método estadístico usado son los árboles de clasificación, comúnmente usados para obtener modelos de predicción precisos. Los resultados del análisis demuestran que la mayoría de las fronteras de los fuegos tienen lugar en zonas de matorral coincidiendo con la preferencia del fuego por estas áreas.

Palabras clave: incendios forestales, gestión forestal, fronteras del fuego, árboles de clasificación.

ABSTRACT

Fire is a natural process and a component of many ecosystems; it is an important factor to ensure the biodiversity in these areas. However, fire is damage for the forests, in recent decades it has increased in the Mediterranean countries due to rural depopulation and climatic conditions.

Particularly, Portugal shows a great increase of surface burnt in the first years of this century in relation to the end of twentieth century. Portugal is a country with high forest cover and importance derived from forestry, so that is required every effort to maintain the future of their forests.

The analysis of boundaries of the burned areas allows study the correlation between the spatial distribution of fire and land environmental variables, and particularly explain why some areas are maintained in many cases as border fire. This will be the main objective of this study, based on the exploration of the most significant variables to interpret the location of such boundaries in mainland Portugal.

By performing statistical analysis of a set of data for large fires in Portugal from 1996 to 2005, has proved the correlation between land use and the boundaries of the fire. The statistical method used is the classification trees, used to obtain accurate prediction models. The analysis results show that most of the boundaries of fires take place in areas of undergrowth.

Key words: forest fires, forest management, wildfire boundaries, classification tree.

EXTENDED ABSTRACT

Year after year the presence of fires increases in the world, particularly in the Mediterranean being there the major cause of forest destruction. Fire is a natural process that helps to maintain biodiversity but its frequency often carries tragic consequences.

In Southern Europe, Portugal and Galiza are the countries of most affected from this phenomenon and in the early years of the twenty-first century there have been an increase burned area in mainland Portugal. The incidence of forest fires is more concentrated to the North of Tagus river and some areas in Algarve.

This is a problem that produces economic, environmental and personal damages. Essential works to prevent wildfires and decrease the number of occurrences are management of human behavior and vegetation management.

Weather in Portugal is characterized by hot and dry summers and cool and wet winters. The alternation of rain and drought favors the accumulation of fuel, being wet enough to create flammable fuel and then sufficiently dry to burn in summer season.

This paper aims to perform a study on the analysis of the boundaries of fire that may help understanding the spatial distribution of fires in relation with the variables related to land use, topography, roads and rivers. For this goal data on fire in Portugal between 1996 and 2005 was required, but it was only possible to obtain data from major fires that presented an area not less than 100 hectares.

The study is based on a prediction model that can be represented as a decision tree. For this purpose, binary trees have been calculated with one response variable and eighteen predictors, using the software CART (Classification and Regression Trees). The classification trees statistical method was used in order to obtain accurate precisions models. This analysis had two main objectives, (1) to gain a better understanding of the characteristics of the boundary burned areas in Portugal and (2) assess the potential of classifiers to future studies (adapted Yohannes and Webb, 1999).

The results obtained in this study, for all different regions in the country, indicate that the shrubland is the most relevant variable that defines the boundaries fires. This is the land cover type “preferred” by fires.

ÍNDICE GERAL

Agradecimentos	i
Resumo	ii
Resumen	iii
Abstract	iv
Extended Abstract	v
Índice Geral	vi
Lista de Abreviaturas e Acrónimos	viii
Lista de Espécies Florestais	x
Lista de Figuras	xi
Lista de Quadros	xii
Lista de Anexos	xiii
1. INTRODUÇÃO	1
1.1 Estrutura da Dissertação	1
1.2 Evolução da Floresta Portuguesa e do Fogo ao Longo da História	2
1.3 Triângulo do Fogo	3
1.4 Comportamento do Fogo	4
1.5 Causas e Efeitos dos Incêndios Florestais	5
1.6 Objetivos de Estudo	7
2. ÁREA DE ESTUDO	8
2.1 Justificação da Área de Estudo	8
2.2 Área Mediterrânea	10
2.3 O Fogo em Portugal	11
2.4 População	13
2.5 Clima	14
2.6 Descrição da Vegetação	16

3. DADOS, MATERIAL E MÉTODOS	17
3.1 Dados	17
3.1.1 Regiões PROF	17
3.1.2 (Perímetro de) Áreas Ardidas	17
3.1.3 Variáveis Topográficas	19
3.1.4 Rede Viária	19
3.1.5 Áreas Urbanas	19
3.1.6 Rede Hidrográfica	20
3.2 Material	21
3.2.1 ArcGis	21
3.2.2 GME	21
3.2.3 CORINE Land Cover	21
3.2.4 CART	23
3.2.4.1 Vantagens e Desvantagens do CART	24
3.3 Método	26
3.3.1 Árvores de Classificação	26
3.3.1.1 Modelo de Árvore	26
3.3.1.2 Regras de Divisão	27
3.3.1.3 Erro de Classificação	27
3.3.1.4 Árvore Ótima	28
3.4 Pré-processamento dos Dados	29
4. ANÁLISE DE RESULTADOS	32
5. DISCUSÃO E CONCLUSÕES	46
6. GLOSSÁRIO	50
7. REFERÊNCIAS BIBLIOGRÁFICAS	53
8. ANEXOS	61

LISTA DE ABREVIATURAS E ACRÓNIMOS

°C – graus centígrados ou graus Celsius

AEA – Agência Europeia do Ambiente

AFN – Autoridade Florestal Nacional

CART – Classification And Regression Trees

CE – Comissão Europeia

CLC – CORINE Land Cover

CLC90 – CORINE Land Cover 1990

CLC2000 – CORINE Land Cover 2000

CORINE – Co-ordination of Information on the Environment

LDRAG – Laboratório de Detecção Remota e Análise Geográfica

ESRI – Environmental Systems Research Institute

et al. - *et alii*, e outros

e.g. – *exempli gratia*, por exemplo

GIS/SIG – Geographic Information System/ Sistema de Informação Geográfica

GME – Geospatial Modelling Environment

ha – hectares

i.e. – *id est*, isto é

IFN – Inventário Florestal Nacional

IGP – Instituto Geográfico Português

IM – Instituto de Meteorologia

INAG – Instituto Nacional da Água

INE – Instituto Nacional de Estatística

MADRP – Ministério da Agricultura, do Desenvolvimento Rural e das Pescas

MAGRAMA – Ministerio de Agricultura, Alimentación y Medio Ambiente

MAMAOT – Ministério da Agricultura, do Mar, do Ambiente e do Ordenamento do Território

mm – milímetros

NUT – Nomenclatura das Unidades Territoriais para fins estatísticos

OEA – Organización de los Estados Americanos

P – precipitação

PIB – Produto Interno Bruto

PLADIGA – Plan de Prevención e Defensa Contra Incendios Forestais de Galicia

PROF – Plano Regional de Ordenamento Florestal

ROC – Receiver Operating Characteristic

T – temperatura

UE – União Europeia

LISTA DE ESPÉCIES: NOME CIENTÍFICO E NOME COMUM

Acacia spp. – acácias

Betula spp. – videiros

Ceratonia siliqua L. – alfarrobeira

Cistus spp. – estevas

Crataegus monogyna Jacq. – pilriteiro, espinheiro-branco

Cytisus spp. – giestas

Erica cinerea L. – queiroga, urze roxa

Eucalyptus globulus Labill – eucalito

Fagus sylvatica L. – faia

Fraxinus spp. – freixos

Myrtus communis L. – murta, mirto

Olea europaea var. *sylvestris* Brot. – zambujeiro

Pinus pinaster Aiton. – pinheiro bravo, pinheiro marítimo

Pinus pinea L. – pinheiro manso

Pinus sylvestris L. – pinheiro silvestre

Quercus coccifera L. – carrasco

Quercus robur L. – carvalho comum, carvalho alvarinho

Quercus rotundifolia L. – azinheira, azinheira de bolota doce (também chamado *Q. ilex* subsp. *ballota*)

Quercus suber L. – sobreiro, sobro

Ulex europaeus L. – tojo, tojo arnal

LISTA DE FIGURAS.

Figura 1. Triângulo do Fogo. (pág. 3)

Figura 2. Triângulo do Comportamento do Fogo. (pág. 4)

Figura 3. Área Total Ardida por Ano na Península Ibérica, no Período de 1996 a 2005. (pág. 8)

Figura 4. a) Análise Espacial da Temperatura Média Anual (°C) Baseada nas Normais de 1961-1990; **b)** Análise Espacial da Precipitação Acumulada Anual (mm) Baseada nas Normais de 1961-1990. (pág. 15)

Figura 5. Discriminação da Área Ardida Devido a Incêndios Florestais Ocorridos em Portugal no Período de 1991 a 2005. (pág. 18)

Figura 6. Mapa de Ocupação do Solo para Portugal segundo a Legenda do Programa CORINE Land Cover 2000. (pág. 22)

Figura 7. Representação Gráfica do Índice Gini. (pág. 24)

Figura 8. Classes de Ocupação do Solo, CORINE Land Cover 2000, localizadas no Distrito de Castelo Branco. (pág. 30)

Figura 9. Classes Simplificadas da Ocupação do Solo a partir do CORINE Land Cover 2000, localizadas no Distrito de Castelo Branco. (pág. 30)

Figura 10. Árvore Ótima Sintetizada e Curva de Custos para “Pinhal Interior Sul”. (pág. 33)

Figura 11. Curva ROC que Representa os Verdadeiros Positivos Vs Falsos Positivos. (pág.34)

Figura 12. Importância das Variáveis Preditoras para a Região do Pinhal Interior Sul. (pág. 36)

Figura 13. Árvore Ótima Obtida mediante o CART para a Região de Alto Minho. (pág.40)

Figura 14. Médias das Variáveis Preditoras mais Importantes para as 21 Regiões PROF. (pág. 44)

Figura 15. Mapa Regional de Distribuição da Importância das Variáveis. (pág. 45)

LISTA DE QUADROS

Quadro 1. Comparação do Número de Ocorrências e Área Total Ardida no Período de 1981 a 2010. (pág. 9)

Quadro 2. Distribuição da Área de Floresta por Tipo e Espécie Dominante, em Portugal Continental. (pág. 10)

Quadro 3. Repartição Regional da População Continental. (pág.14)

Quadro 4. Classes Simplificadas do CORINE Land Cover em que se Reagrupam Diferentes Classes da Legenda do CORINE Land Cover 2000 Original. (pág. 29)

Quadro 5. Matriz de Erros que Representa as Observações Totais Bem/Mal Classificadas em cada Categoria e a Precisão da Classificação. (pág. 35)

Quadro 6. Resumo do Output do CART para cada PROF. (pág. 37)

Quadro 7. Variáveis Primárias, Variáveis mais Importantes e Variáveis Competidoras para Cada PROF. (pág. 38)

Quadro 8. Classes Erradamente Classificadas para Cada Região. (pág. 39)

Quadro 9. Leitura-Resumo das Árvores de Classificação para as 21 Regiões PROF. (pág. 41)

LISTA DE ANEXOS

ANEXO 1. Número de Ocorrências e Área Ardida para os Distritos de Portugal Continental no Período de 1996 a 2005. Número de Hectares Afetados pelo Fogo no Mesmo Período.

ANEXO 2. Mapa de Identificação das Regiões PROF para Portugal Continental.

ANEXO 3. Mapa Representativo dos Grandes Fogos Ocorridos em Portugal Continental entre 1996 e 2005.

ANEXO 4. Mapa da Linhas de Cumeada para Portugal Continental.

ANEXO 5. Mapa da Rede Viária de Portugal Continental.

ANEXO 6. Mapa da Rede Hidrográfica de Portugal Continental.

ANEXO 7. Ávore de Classificação e Output Obtido para Pinhal Interior Sul.

ANEXO 8. Árvores de Classificação para as 21 Regiões PROF

1. INTRODUÇÃO

1.1 Estrutura da Dissertação

Inicialmente foi feita uma revisão bibliográfica de índole geral que abrange vários assuntos de importância de forma a introduzir o tema em estudo: história da floresta portuguesa e dos fogos, breves conceitos sobre o fogo, suas causas e efeitos, assim como uma descrição concisa do comportamento do fogo e os objetivos do estudo. A seguir, foi caracterizada a área de estudo, nomeadamente o seu clima, a sua vegetação e o problema do fogo em Portugal.

A parte do desenvolvimento do trabalho mais metodológica consiste na descrição dos dados e do seu pré-processamento, do material e do método a usar para analisar e interpretar os dados resultantes. Para sintetizar os resultados obtidos procedeu-se à discussão e a conclusões pertinentes, tendo em conta a referência a outros trabalhos publicados.

São acrescentados documentos auxiliares que aportam informação adicional, tais como anexos, lista de abreviaturas, lista de espécies florestais e um glossário de termos específicos.

1.2 Evolução da Floresta e do Fogo ao Longo da História

A paisagem foi sempre objeto de transformações no sentido da adaptação às necessidades humanas. A essência destas transformações é a sua utilidade para a sociedade, a sustentabilidade dos processos, a valorização da identidade da paisagem e o seu contributo para a satisfação do interesse público (Abreu, 2012).

A paisagem portuguesa era, outrora, constituída principalmente por árvores do género *Quercus* – carvalhos de folha caduca a Norte, carvalhos de folha perene a Sul – e em menor grau por árvores do género *Pinus* – pinheiro bravo no Litoral Norte, pinheiro manso mais a Sul (Radich e Alves, 2000). Mas este tipo de floresta foi desaparecendo, quer pelo abate das árvores, quer pelo fogo, com o objetivo de fomentar o pastoreio e a agricultura. Também o aparecimento do eucalipto (espécie alóctona) na década de sessenta, impulsionado pela implantação da indústria de pasta e papel, foi restringindo a área das espécies nativas e ocupando grandes extensões de terreno (Radich e Alves, 2000). Não menos impactante é o efeito da urbanização na transformação da paisagem, por exemplo modificando o uso do solo após o fogo, como tem acontecido nos últimos tempos até a entrada em vigor do Decreto-Lei nº 55/2007 (Artigo 1º).

O fogo foi usado, desde a antiguidade, como ferramenta essencial (Rego et al., 2010) na defesa e proteção das populações humanas, quer em guerras e conquistas quer para afastar animais e pragas dos cultivos. Também foi utilizado na expansão da área agrícola, na regeneração de rebentos para alimento do gado e mesmo como adubo misturando as cinzas com a terra devido às substâncias nutritivas libertadas pelo fogo (Pausas, 2012). Os nossos antepassados utilizaram o fogo para controlar os combustíveis vegetais e evitar o desastre dos incêndios florestais. Nos últimos anos tem-se recorrido a esta atividade ancestral como técnica de silvicultura preventiva, conhecida como uso do fogo controlado; mesmo assim, este método nem sempre é bem visto ou compreendido pela população.

O fogo é um processo natural que faz parte do ciclo de muitos ecossistemas, e é necessário para manter a biodiversidade de algumas paisagens (Pausas e Vallejo, 1999). Existem espécies vegetais pirófitas que se propagam após a passagem do fogo, por exemplo algumas espécies dos géneros *Cistus* ou *Pinus* (Ruiz del Castillo, 2003). Com tudo, os incêndios são responsáveis pela destruição do nosso património natural, não apenas pela perda de vegetação e pelas consequências que originam mas também pela entrada e dominância de espécies invasoras, como é o caso de diferentes espécies de acácias que competem com as espécies nativas alterando a disponibilidade dos recursos, o ciclo de nutrientes e/ou o regime hídrico.

A intervenção humana, nomeadamente, através da má ou deficiente gestão dos povoamentos florestais, de práticas agrícolas incorretas e de atitudes negligentes ou mesmo intencionais, tem feito aumentar a frequência de incêndios, pondo em causa a regeneração das florestas e a proteção dos próprios cidadãos e dos seus bens.

Foi após 1978 que o problema dos incêndios florestais se tornou mais sério em Portugal – com exceção dos fogos ocorridos em 1974 e 1975 –, particularmente nas regiões Norte e Centro do país (Lourenço, 1988). Estes, podem ser considerados os primeiros incêndios da era atual que virão a dizimar a floresta portuguesa, coincidindo com um grande movimento migratório populacional, pois é a partir da década de setenta que se manifesta este movimento da população rural para as cidades e para países estrangeiros (Baptista, 2003). Assim, o abandono do campo favorece o descuido das florestas, pois deixam de se utilizar as lenhas, resinas, carvão e as limpezas das matas ficam no esquecimento; o resultado é o aumento da quantidade de combustível e a continuidade do mesmo, facilitando a propagação dos incêndios e tornando a sua extinção mais difícil (Ruiz, 2001).

1.3 Triângulo do Fogo

O fogo é um processo físico-químico, resultante da combustão que consiste numa oxidação violenta com grande libertação de calor e luz. Para se produzir combustão é necessário que se encontrem no tempo e no espaço os três elementos que compõem o triângulo do fogo: combustível, comburente e energia de ativação (Figura 1).



Figura 1. Triângulo do Fogo (Rodrigues, 2009).

Combustível: refere-se a toda matéria de origem vegetal que tem capacidade para arder. Alguns vegetais contêm substâncias químicas – resinas, óleos, ceras – que favorecem a propagação do fogo.

Comburente: é o elemento que ativa o fogo, o oxigénio. Encontra-se numa proporção de 21% no ar atmosférico e sofre poucas variações quer espaciais quer temporais.

Energia de ativação: é o elemento que inicia o fogo, concretamente o calor, fazendo com que comburente e combustível entrem em reação; assim é atingido um ponto em que se produz a ignição e se mantém a combustão pela quantidade de calor libertada no processo.

Após o fogo se ter iniciado pode manter-se por si só, isto é, o fogo é um fenómeno autosustentado, o qual é explicado com a introdução de um quarto elemento chamado reação em cadeia, obtendo-se assim o tetraedro do fogo. Esta reação em cadeia é devida a uma sequência de reações que se originam quando o próprio fogo liberta gases ou calor enquanto existe combustível e oxigénio disponível, de forma que o calor libertado seca a vegetação contígua podendo provocar um outro foco de ignição.

A extinção do fogo ocorre quando se elimina um dos lados do triângulo, ou seja, quando se atua sobre um dos três elementos. Pode-se atuar na fonte de calor descendo a temperatura com água ou terra; atuar sobre o comburente mediante o abafamento, fazendo que a percentagem de oxigénio atmosférico fique abaixo de 16% (e.g., água pulverizada faz com que diminua a concentração de oxigénio no ar); atuar no combustível modificando a continuidade da vegetação (e.g., faixas livres de vegetação, contrafogos) ou dificultando a combustão com retardantes químicos que melhoram a eficácia da água (Lombardero, 2005; Pereira, 2008).

1.4 Comportamento do Fogo

O comportamento do fogo pode ser definido como o modo em que os combustíveis entram em ignição, a chama se desenvolve e o fogo se propaga. Este comportamento, e também a evolução do incêndio, depende da interação de três fatores (Figura 2): a topografia ou terreno em que se desenvolve o fogo, a vegetação que serve como combustível e a meteorologia (Pereira, 2008). Será este comportamento que condiciona as técnicas e medidas necessárias para a prevenção, o combate e a mitigação dos efeitos do fogo.



Figura 2. Triângulo do Comportamento do Fogo (Rodrigues, 2009).

Topografia: é um fator que pouco varia no tempo mas é muito variável no espaço. Refere-se a elementos como a altitude que determina a quantidade e distribuição do combustível e tem influência sobre o clima; a exposição ou orientação das encostas que afeta o comportamento do fogo devido às variações de radiação solar que recebem; o declive que favorece a continuidade vertical da vegetação e os ventos de encosta e ainda tem influência nas colunas de convecção que afetam à velocidade de propagação do fogo; os vales que têm associado o efeito chaminé facilitando o pré-aquecimento dos combustíveis.

Combustível: como foi referido anteriormente, os combustíveis são uma componente do triângulo do fogo. Como tal, as suas condições ou características – humidade, tamanho, quantidade, distribuição – são decisivas para saber como é que o fogo se propaga.

Meteorologia: as variáveis atmosféricas que mais influência têm no comportamento do fogo são a temperatura, a humidade relativa do ar e o vento. Os fatores climatológicos influenciam o estado dos combustíveis (temperatura, humidade) e a propagação do fogo (vento). Embora a temperatura elevada associada a uma baixa humidade origine materiais vegetais mais secos e portanto com maior combustibilidade (Torres, 2006), é o vento o fator meteorológico com maior relevância na evolução dos incêndios florestais, desde o dessecamento da vegetação ou mudanças na sua velocidade e direção até a propagação de chamas ou emissão de faíscas (Vélez, 2009). Estas variáveis afetam a inflamabilidade e a combustibilidade da vegetação.

Inflamabilidade e combustibilidade são duas propriedades ou características dos combustíveis, enquanto a ignição é uma ação realizada sobre um combustível. Viger et al. (2004) definiram estes três termos: a ignição é um termo referente à ação de fornecer a energia necessária para que um combustível possa entrar em combustão; a inflamabilidade define-se como a facilidade com que um combustível entra em combustão quando se põe em contacto com uma fonte de ignição; a combustibilidade definir-se-á como a facilidade que tem um combustível para continuar a arder uma vez iniciada a combustão.

1.5 Causas e Efeitos dos Incêndios Florestais

A causa dos incêndios florestais é muito variada. Podem ter a sua origem em causas estruturais que dependem de fatores intrínsecos do próprio meio natural, tais como as secas, a acumulação de carga de combustível, a inflamabilidade da vegetação; ou em causas imediatas que derivam de comportamentos antrópicos ou de agentes naturais, tais como os raios, as negligências, a intencionalidade (Porrero Rodríguez, 2001).

Algumas vezes os fogos ocorrem de forma natural mas em grande medida trata-se de fogos que têm origem humana, quer direta quer indiretamente. Estes podem resultar de negligências – queimas de restos de trabalhos florestais ou agrícolas, uso de foguetes, cigarros – e inclusive da intencionalidade. Existe, também, um elevado número de causas classificadas como desconhecidas que se deveriam reduzir mediante cursos de formação para agentes florestais e policiais e/ou mediante o aumento do número de brigadas de investigação (Vélez, 2006).

Na atualidade pode dizer-se que a maioria dos fogos são intencionais, mas as motivações são do mais diverso. Desde vinganças pessoais, áreas protegidas ou atividade cinegética até piromania ou diminuir o preço da madeira (Lombardero, 2005).

O risco de incêndio na interface urbano-rural está a aumentar significativamente nas últimas décadas devido à expansão das edificações em zonas florestais. Como tal, o combate ao fogo concentra-se na proteção das habitações e vidas humanas, de forma que se abandona a defesa da vegetação (Vélez, 2006).

Os efeitos mais óbvios que se podem observar após um incêndio são os impactes na vegetação, pois a maior parte dos seus tecidos aéreos podem ser destruídos em pouco tempo provocando a sua morte. As plantas que vivem em ambientes com elevada frequência de incêndios foram adquirindo uma capacidade adaptativa que permite a sua sobrevivência e reprodução, entre estas adaptações pode falar-se de casca espessa, serotinia, desramação natural ou germinação estimulada por calor.

Dependendo da intensidade do fogo e da sua severidade algumas espécies vegetais morrem imediatamente, outras ficam debilitadas mas com elevada probabilidade de serem atacadas por pragas e doenças ou derrubadas pelo vento (Catry et al., 2010). Com tudo, o fogo também tem efeitos consideráveis no solo, tal como a mineralização da matéria orgânica que faz com que os nutrientes fiquem em forma de sais com elevado pH ou sejam arrastados através da escorrência quer em profundidade quer em superfície. Outros efeitos são o aumento da suscetibilidade do solo à erosão derivada da perda de proteção proporcionada pelo coberto arbóreo e pela folhada ou a formação de uma camada repelente à água que dificulta a infiltração e facilita o processo de escorrência. Neste sentido, o ciclo hidrológico é afetado pelas substâncias transportadas pela água a jusante que muitas vezes chegam a aquíferos próximos (Ferreira et al., 2010).

O fogo deixa, para além destes efeitos, outros não menos importantes, como são os sofridos por parte da fauna selvagem quando o seu alimento, habitat ou refúgio são destruídos e tem como possível consequência a migração de algumas espécies e inclusive a morte

instantânea dos indivíduos mais vulneráveis (Morgado e Moreira, 2010). Também a qualidade do ar, a paisagem e a própria vida humana sofrem com este fenómeno.

1.6 Objetivos do Estudo

Neste trabalho pretende-se determinar quais as características que influenciam os limites geográficos ou fronteiras das áreas queimadas derivadas dos incêndios florestais em Portugal continental, assim como verificar se existem correlações entre as variáveis – topografia, rede hidrográfica, estradas, ocupação do solo – associadas a estas fronteiras. Para realizar este estudo será usado um método estatístico que permite analisar e classificar os dados recorrendo a árvores de classificação.

Para tal, foram seleccionados os fogos relativos ao período de 1996 a 2005, por verificar maior número de ocorrências e maior área afetada pelo fogo. Da totalidade dos incêndios registados, apenas serão usados os grandes fogos, aqueles com uma área igual ou superior aos 100 hectares. Estes fogos são os que determinam a dimensão das perdas e danos verificados num dado ano, muito mais do que o número total de fogos (Pereira et al., 2006).

O estudo das fronteiras dos fogos ajuda a perceber a influência destes na estrutura da paisagem, proporcionando também informação da relação entre os incêndios florestais e as variáveis ambientais em que os fogos se desenvolvem. Pretende-se, desta forma, realizar uma análise espácio-temporal útil para compreender melhor a frequência da paragem dos fogos nas mesma áreas durante anos diferentes.

2. ÁREA DE ESTUDO

2.1 Justificação da Área de Estudo

O presente estudo corresponde unicamente a Portugal continental, ficando excluídos dele as Regiões Autónomas dos Açores e da Madeira. Portugal situa-se na zona Ocidental da Península Ibérica, no extremo Sudoeste da Europa continental. Localiza-se entre as latitudes 37°N e 42°N e entre as longitudes 6.5°O e 9.5°O (Miranda et al., 2006).

O rio Tejo divide o país em duas grandes zonas orográficas. A Norte a geografia é montanhosa e acidentada apresentando os picos mais elevados da orografia continental – Serra da Estrela com 1993 metros de altitude ou Serra do Gerês com 1548 metros. A Sul as terras caracterizam-se pela sua uniformidade e planícies, mas também com serras esporádicas – Serra de São Mamede com 1205 metros de altitude ou Serra de Monchique com 902 metros (Ribeiro, 1998).

Portugal é um país com elevado coberto florestal, aproximadamente 39% do território português está destinado a florestas, 33% a agricultura e 22% a matos (AFN, 2011). O Noroeste da Península Ibérica, correspondente ao Norte e Centro de Portugal e à Galiza, é a região da Europa com maior número de incêndios, chegando a representar metade dos fogos registados no Velho Continente (Pereira et al., 2002; Pereira et al., 2005). Em comparação com outros países do Sul da Europa, a área ardida no Noroeste peninsular é a mais alta (Pereira et al., 2002). Na Figura 3 mostra-se a área total ardida na Península Ibérica para o período de 1996 a 2005.

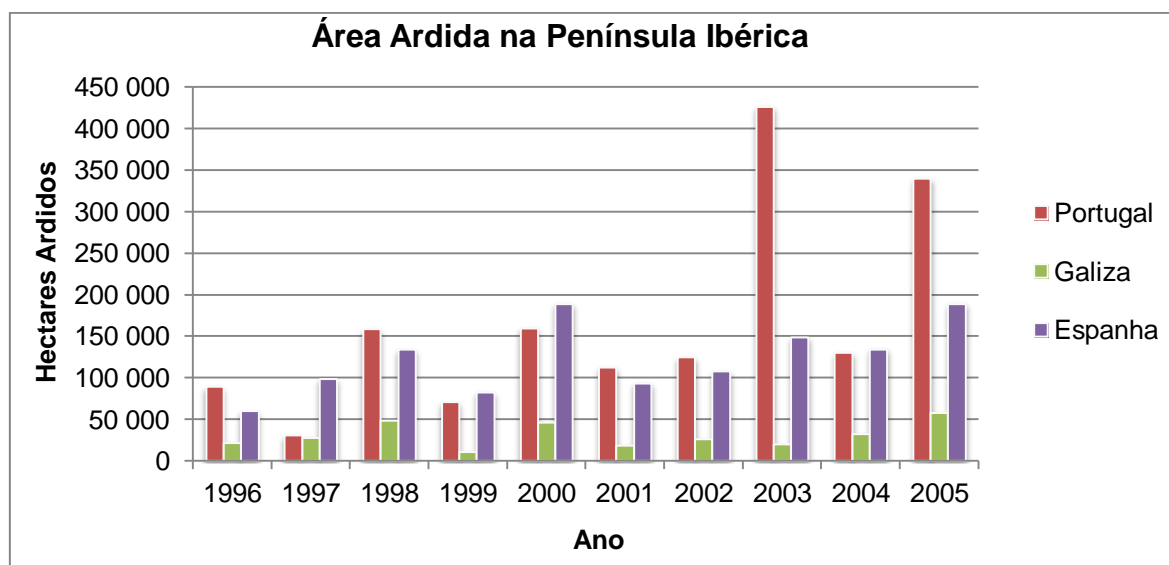


Figura 3. Área Total Ardida por Ano na Península Ibérica, no período de 1996 a 2005 (AFN, 2011; PLADIGA, 2011; MAGRAMA, 2011).

Para Portugal verificou-se um grande aumento da área ardida, resultando que nos primeiros cinco anos deste século (entre 2001 e 2005) esta área corresponde a 1 131 967 hectares (ha), mais do dobro do somatório dos cinco anos anteriores (507 989 ha). Esta eventualidade é favorecida pelos incêndios acontecidos em 2003 (425 839 ha) e 2005 (339 089 ha). Neste período destaca-se o ano 2003 com maior área ardida, sendo que apenas 0.9% dos incêndios registados correspondem a fogos com área igual ou superior a 100 ha. O ano 1997 tem registado a menor área afetada por incêndios florestais dos quais 0.2% foram responsáveis pela queima de áreas iguais ou superiores aos 100 ha. (Pereira et al., 2006).

O número de ocorrências nos primeiros dez anos do século XXI ultrapassa os 241 500 e o número total de área florestal ardida aproxima-se do milhão e meio de hectares, entre povoamentos e matos. São valores muito elevados que merecem ser estudados e analisados em profundidade para poder melhorar o futuro das nossas florestas. No Quadro 1 comparam-se, resumidamente os dados registados nos três últimos decénios, referentes ao número de ocorrências e a área assolada pelos fogos.

Quadro 1. Comparação do Número de Ocorrências e Área Total Ardida entre 1981 e 2010. Fonte: AFN, 2011.

Período	Nº Ocorrências Totais	Área Ardida Total (ha)
1981-1990	82 205	827 842
1991-2000	245 866	1 044 383
2001-2010	241 583	1 495 465

Na última revisão do Inventário Florestal Nacional (IFN 2005-2006) da Autoridade Florestal Nacional (AFN), a ocupação florestal cresceu ligeiramente (3%) no território continental, o que supõe um aumento próximo de 110 000 ha, em relação ao inventário anterior de 1995-1998. No que se refere ao povoamento total, houve um aumento de pinheiro manso, carvalhos caducifólios e eucalipto, mas também houve uma importante descida de castanheiros, azinheiras e pinheiro bravo. Por outro lado, o sobreiro tem sofrido um pequeno aumento nos últimos anos embora seja o pinheiro bravo a espécie mais representativa de Portugal continental seguido do eucalipto. Dados detalhados são fornecidos no Quadro 2, onde se comparam os dois últimos inventários em relação aos povoamentos florestais.

Quadro 2. Distribuição da Área de Floresta por Tipo e Espécie Dominante, em Portugal Continental.
Fontes: AFN, 2011 e CELPA, 2010.

Espécies Florestais	Povoamento Puro (x1000 ha)		Povoamento Misto Dominante (x1000 ha)		Povoamento Total (x1000 ha)	
	1995-1998	2005-2006	1995-1998	2005-2006	1995-1998	2005-2006
Pinheiro bravo	730	681	246	204	976	885
Eucalipto	573	566	99	173	672	740
Sobreiro	592	548	121	168	713	716
Azinhreira	387	361	74	52	462	413
Pinheiro manso	48	54	30	77	78	130
Carvalho	76	102	55	48	131	150
Castanheiro	32	24	9	6	41	30
Outras resinosas	21	12	6	13	27	25
Outras folhosas	63	57	39	29	102	86
Outras áreas					148	242
TOTAL	2522	2405	679	770	3350	3458

Portugal é um país com grande riqueza – económica, social e ambiental – derivada do setor florestal. Segundo dados referentes a 2010 do ex-Ministério da Agricultura, do Desenvolvimento Rural e das Pescas (atual Ministério da Agricultura, do Mar, do Ambiente e do Ordenamento de Território), as florestas geram 260 000 postos de trabalho, contribuem com 12% do PIB da indústria e são responsáveis por 11% das exportações totais do país. Assim sendo, cabe a investigadores, instituições e à própria cidadania a proteção das florestas para poderem continuar a ser alvo da economia nacional.

2.2 Área Mediterrânea

Na bacia mediterrânica, os incêndios florestais são um problema permanente que é sustentado/alimentado pela topografia do terreno e/ou pelas condições ambientais (Vélez, 2006). A alternância da chuva e da seca propicia e favorece a acumulação de combustível, pois o local é suficientemente húmido para criar combustível inflamável e suficientemente seco para que este possa arder. Além disso, as trovadas podem aumentar a ocorrência de grandes incêndios (Vélez, 2003; Pyne, 2006; Pausas, 2012).

A presença dos fogos aumenta ano após ano em todos os lugares do mundo, particularmente na bacia mediterrânica, onde são a causa mais importante da destruição das florestas (Pausas, 2012; Vélez, 2006). É um problema que produz perdas económicas,

ecológicas e pessoais nos países mediterrâneos. Portugal, Espanha, França, Itália e Grécia destacam-se por serem os mais afetados desta região (Vélez, 2003).

Os incêndios florestais nesta área geográfica são o produto de diferentes fatores como a meteorologia – ventos fortes, falta de humidade, temperatura –; a vegetação – espécies que dependem fisiologicamente do fogo, grau de secura e tipo de coberto –; os factores socio económicos – abandono das zonas rurais com a consequente perda de usos tradicionais como a recolha de resinas e lenhas ou pastoreio, e o aumento de áreas recreativas. A acumulação de biomassa numa atmosfera rica em oxigénio faz com que o fogo seja possível mesmo que o esforço humano tente evitá-lo (Zedler e Rego, 2006).

2.3 O Fogo em Portugal

Portugal é um dos países do Sul da Europa que mais sofre com este fenómeno e a área ardida tem vindo a aumentar, *e.g.*, na década de 1980 arderam 72 757 ha, 91 220 ha na década de 1990 e 192 186 ha entre 2000 e 2005. A maior parte das áreas ardidas estão associadas a um número reduzido de incêndios, mas estes são de grandes dimensões (Pereira et al., 2006).

A incidência dos fogos florestais concentra-se sobretudo a norte do rio Tejo e em algumas zonas do Algarve. Grande parte dos fogos ocorre em zonas densamente povoadas (*e.g.*, Porto, Braga, Lisboa), mas estes não são os mais preocupantes devido às boas infraestruturas viárias, postos de bombeiros ou condições ambientais pouco severas que fazem com que não se originem grandes incêndios (Pereira et al., 2006). No extremo oposto localizam-se zonas despovoadas e envelhecidas com extensas manchas contínuas de floresta, matagais e condições propícias – verões secos, ondas de calor – para se desenvolverem grandes fogos; é nestes lugares que se registam incêndios com grandes áreas queimadas (Pereira et al., 2006), tal como a maior mancha ardida registada em 2003 na Serra de Monchique com cerca de 70 000 ha calcinados.

A área ardida menos elevada confirma-se nos distritos de Setúbal, Lisboa, Beja e Aveiro registando valores inferiores a 40 000 ha queimados. A maior área afetada pelo fogo corresponde a Guarda, Castelo Branco, Vila Real e Viseu onde se encontram valores acima dos 150 000 ha ardidos entre povoamentos e matos. Os anos mais devastadores correspondem a 1998, 2000, 2003 e 2005, em todos eles a área atingida pelos incêndios ultrapassou os 155 000 hectares de terreno consumidos pelo fogo. Esta informação é disponibilizada nas tabelas do Anexo 1.

A relação entre o número de incêndios e a área queimada varia muito com os anos, pelo que é importante uma boa política de prevenção e combate. Nos últimos anos têm melhorado os meios de extinção e portanto os danos são menores mas o resultado não é totalmente satisfatório, pois mesmo que o sistema de prevenção, detecção e combate tenham uma elevada eficácia na extinção rápida (antes de se queimar 1 ha de terreno), existe uma pequena percentagem de fogos que escapa do ataque inicial chegando a atingir grandes dimensões e provocando uma grande proporção de área queimada (Pereira et al., 2006).

Tarefas imprescindíveis para prevenir incêndios, e assim diminuir o número de ocorrências, podem ser: (1) gestão de comportamentos humanos: através do fomento da sensibilização da população mediante campanhas de propaganda/educação com o fim de consciencializar dos efeitos e consequências do fogo não desejado; (2) gestão da vegetação: recorrendo a programas de silvicultura preventiva – desbastes parciais, queimas controladas, pastoreio, limpezas – e apostar nas infraestruturas de prevenção – corta-fogos, pontos de água, bases aéreas, pontos fixos de vigilância, mosaicos vegetais –; (3) formação de mão de obra especializada na defesa contra incêndios florestais e voluntariado para vigilância, melhora nas políticas de ação e revisão da legislação punitiva (Lombardero, 2005; DL 124/2006).

No X Congresso Florestal Mundial, celebrado em Paris em 1991, Vélez destacou a importância da silvicultura preventiva na defesa contra os incêndios florestais. Esta, tem por objetivo modificar a estrutura da massa florestal para dificultar a propagação do fogo, por meio de técnicas tais como a diversificação da massa – evitando manchas monoespecíficas e favorecendo a alternância de espécies –, a discontinuidade da vegetação quer horizontal quer vertical – mantendo faixas lineares livres de vegetação, poda de árvores, pastoreio –, ou o uso do fogo controlado como técnica para ensinar a queimar.

Borges (2001) considera que o *Ordenamento e a Gestão Florestal*, tendo como conceito básico promover a sustentabilidade das florestas, é um instrumento necessário na prevenção dos incêndios florestais e na mitigação dos seus impactes negativos. Portanto, a prevenção dos fogos deve começar na conceção de projetos de (re)florestação e continuar na silvicultura e na prática da gestão florestal. É por isso que Borges e Uva (2006) sugerem que a articulação dos processos de *Ordenamento e Gestão Florestal* e de *Prevenção e Combate a Incêndios Florestais* pode facilitar a solução do problema dos fogos.

Portugal encontra-se com um problema que dificulta as atividades de gestão, e é que a maior parte da área florestal do país pertence a proprietários privados, ocupando perto de 75% da área florestal do continente e a maioria das propriedades são de pequena dimensão (Baptista e Santos, 2005). O resultado é um país desordenado e uma paisagem demasiado

simplificada com monoculturas que ocupam vastas extensões de terreno e terras abandonadas. O Associativismo Florestal surge com o intuito de fazer frente ao abandono e degradação das áreas florestais, derivados da emigração ou da idade avançada dos habitantes do meio rural (Rebelo et al., 1999). É uma estratégia de grande importância na gestão da floresta que permite melhorar o ordenamento das matas ao mesmo tempo que se torna mais fácil evitar incêndios, melhorar a qualidade dos produtos florestais e obter maiores rendimentos (FORESTIS, 2009).

Os Planos de Ordenamento, quer nacional quer regional, são documentos inovadores que integram diversas funções do espaço florestal. Ainda assim, reconhecendo a vocação e o esforço desenvolvido, este setor tem pontos que deveriam ser mudados e melhorados para encontrar soluções adequadas e ultrapassar os obstáculos que limitam o aproveitamento florestal; e.g., evitar grandes manchas contínuas mono-específicas, melhorar o conhecimento dos modelos silvícolas, diminuir o risco de incêndio, melhorar a gestão dos espaços. Foi com o desastre dos incêndios ocorridos em 2003 que se decidiu acelerar o processo dos instrumentos de planeamento e gestão, de forma a evitar a repetição do acontecido nesse verão (Santos et al.; Rosendo e Martins).

2.4 População

A população representa um papel importante neste estudo, embora não seja utilizada como uma variável para a obtenção dos resultados finais, não poderia deixar de ser mencionada. Por um lado pode contribuir com o aumento de focos de incêndios e por outro lado pode representar grande ajuda na hora de proporcionar o alerta sobre a localização de um fogo e mesmo proceder ao ataque de extinção enquanto os meios efetivos não chegarem.

A demografia portuguesa experimentou uma grande mudança nas últimas décadas, sobretudo no referente ao mundo rural. O abandono destas terras, por diversos motivos, provoca não apenas a decrepitude nas aldeias ou o envelhecimento da população, mas sim a falta de gestão das terras agrícolas e florestais que conleva problemáticas consequências, como por exemplo a presença de grandes massas de vegetação contínuas e densas que aumentam o risco de incêndio e ao mesmo tempo tornam difícil o seu controlo quando se deflagra.

A distribuição da população é muito variada entre as regiões do litoral – onde se encontra a maior concentração populacional – e as regiões do interior do país – onde a redução da população aumentou nos últimos anos (IGP, 2011). No Quadro 3 mostra-se a densidade populacional média para as Regiões NUT II em 2010.

Quadro 3. Repartição Regional da População Continental (Média 2010). Fonte: INE, Inquérito de Emprego, 2010.

Região	População (milhares)	% do total	Densidade (hab/km ²)
Norte	3 738.8	35.2	176
Centro	2 379.5	22.4	84.3
Lisboa	2 835	26.7	943.3
Alentejo	751	7	23.8
Algarve	437.9	4.1	86.9
Total	10 142.2	95.4	1 314.3

A Região Norte apresenta maior proporção de população seguida da Região de Lisboa e da Região Centro, embora seja a Região de Lisboa a que tem maior densidade populacional seguida da Região Norte. No lado oposto encontram-se o Algarve com menor população, e o Alentejo com menor número de habitantes por quilómetro quadrado.

2.5 Clima

A Península Ibérica, e consequentemente Portugal, enquadra-se na transição de duas grandes regiões biogeográficas. A Região Eurossiberiana, caracterizada por ter uma aridez estival nula ou muito ligeira, nunca superior a dois meses com $P < 2T$. A Região Mediterrânea, distinguida por possuir um clima em que escasseiam as chuvas no verão, pelo menos dois meses onde $P < 2T$ (Costa et al., 1998). Em ambos os casos P representa a precipitação média mensal (mm) e T a temperatura média mensal (°C).

O clima de Portugal caracteriza-se por ter verões quentes e secos e invernos frescos e húmidos. É um clima temperado de influência mediterrânica mas também com influência oceânica que amortiza as alterações e amplitudes térmicas. As cadeias montanhosas atuam como barreira impedindo a passagem de ventos marítimos carregados de humidade (Fabião e Oliveira, 2007).

A distribuição espacial da temperatura média deve-se ao efeito da latitude, altitude e proximidade do litoral atlântico. Assim, as regiões a Sul do Tejo registam valores mais elevados de temperatura média anual, nomeadamente o Alentejo e o Algarve. A região Norte regista os valores mais baixos de temperatura, sobretudo nas zonas montanhosas (Miranda et al., 2006). Na Figura 4a) pode-se constatar a variação da temperatura média anual (°C) obtida no site do Instituto de Meteorologia.

Relativamente às precipitações, estas ocorrem com maior frequência na parte Norte (e.g., montanhas do Noroeste), enquanto o Sul recebe menos precipitações (e.g., Nordeste da Serra Algarvia). Também há um contraste litoral-interior que manifesta valores elevados nos relevos próximos da costa e valores muito baixos nas depressões, planícies e planaltos protegidos da influência atlântica por barreiras montanhosas (Daveau et al., 1988). Na Figura 4b) representa-se a precipitação acumulada anual (mm) obtida no Instituto de Meteorologia.

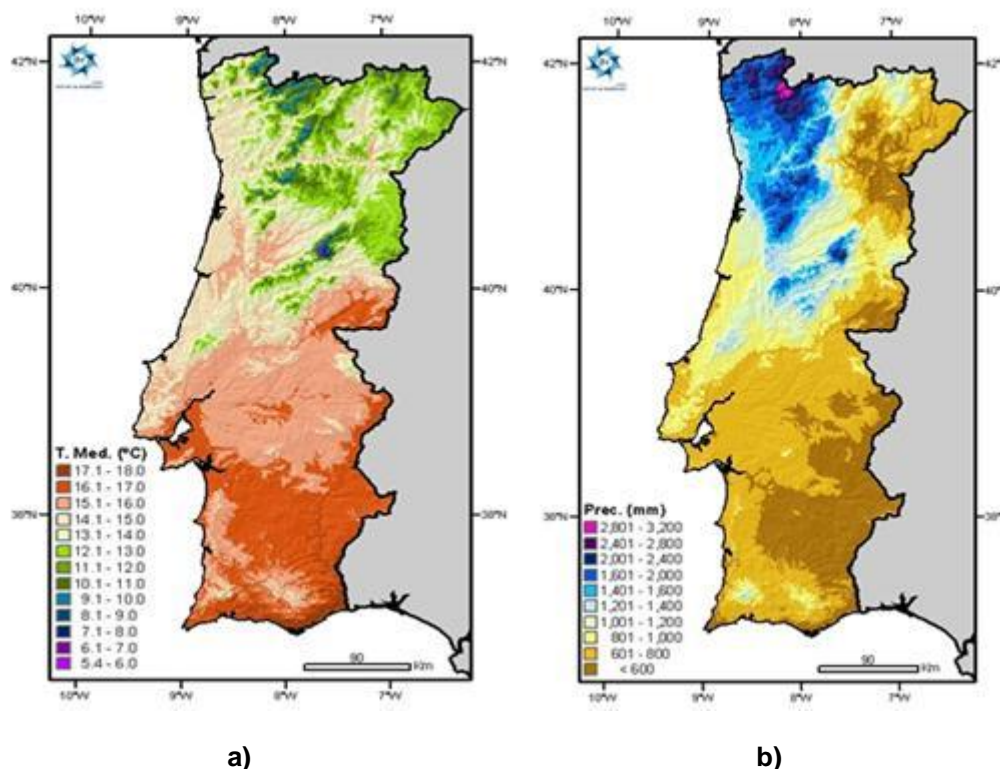


Figura 4. a) Análise Espacial da Temperatura Média Anual (°C) Baseada nas Normais de 1961-1990;
b) Análise Espacial da Precipitação Acumulada Anual (mm) Baseada nas Normais de 1961-1990.
 (IM, 2012)

Com base na figura da esquerda, pode-se advertir que a temperatura média varia entre 6°C, nas terras altas do Norte e Centro do país, e 18°C no litoral Sul. Na figura da direita, destacam os valores de precipitação superiores a 3 000 mm na região do Minho, enquanto que a Beira Interior e o interior do Alentejo recebem os mais baixos, menos de 600 mm. Em ambos os casos a variação espacial é um fator muito significativo.

As condições mais severas de clima (seca, calor e/ou vento) condicionam o comportamento do fogo e a sua propagação, ao mesmo tempo que dificulta as tarefas de combate e extinção. Neste contexto, o ecossistema Mediterrâneo possui uma época seca que é mais propícia para a ocorrência de incêndios.

2.6 Descrição da Vegetação

Segundo Rivas-Martínez (1982), a Península Ibérica, na sua maior parte, está localizada na Região Bioclimática Mediterrânica. A Região Eurossiberiana ocupa apenas o limite setentrional e marítimo da Península.

Na Região Eurossiberiana mais de 60% do território está coberto por formações lenhosas de florestas e matos. Destacando as espécies *Betula spp.*, *Cytisus spp.*, *Erica cinerea* L., *Fagus sylvatica* L., *Fraxinus spp.*, *Pinus pinaster* Aiton, *P. sylvestris* L., *Quercus robur* L. e *Ulex europaeus* L.. Na Região Mediterrânica distinguem-se espécies como *Ceratonia siliqua* L., *Cistus spp.*, *Crataegus monogyna* Jacq., *Myrtus communis* L., *Olea europaea* var. *sylvestris* Brot., *Pinus pinea* L., *Pistacia lentiscus* L., *Quercus rotundifolia* L., *Q. coccifera* L. ou *Q. suber* L. (Costa et al., 1998).

De todas as espécies florestais lenhosas que se podem encontrar no território continental português, há um destaque especial para três delas: (1) pinheiro bravo (*Pinus pinaster* Aiton), localizado principalmente na metade Norte do país; (2) eucalipto (*Eucalyptus globulus* Labill.), espécie introduzida que domina na metade ocidental e em algumas zonas do Centro e Sul e (3) sobreiro (*Quercus suber* L.), localizado na metade Sul (Nunes et al., 2005). Estas são as espécies que apresentam maior interesse na economia florestal portuguesa.

A agricultura está presente em todo o país. No Norte e Centro as propriedades são pequenas e a paisagem está mais fragmentada, dominando vinhas e olivais; no Sueste do país as áreas agrícolas são mais extensas, dominando o cultivo de cereais (Nunes et al., 2005).

Os mesmos autores têm demonstrado a preferência dos fogos pelo coberto vegetal, sendo que os fogos pequenos preferem matagais enquanto evitam áreas agrícolas e culturas anuais. Os fogos de grande tamanho não são tão seletivos, mas o pinheiro bravo destaca como a espécie florestal mais afetada por este tipo de incêndios por conter, muitas vezes, uma carga de combustível pesada que torna o comportamento do fogo mais grave.

3. DADOS, MATERIAL E MÉTODOS

Os incêndios florestais têm interesse ambiental, económico e social mas também faz sentido que a Geografia se ocupe do seu estudo considerando que existe uma íntima relação entre a análise do território e alguns fatores ambientais e humanos – tipo e estado do coberto vegetal, condições climáticas, topografia, atividades humanas. Após o fogo, a Geografia faz uma avaliação das suas consequências espaciais através da quantificação e cartografia das áreas queimadas e uma análise dos seus efeitos ambientais, nomeadamente, erosão do solo, ciclo hidrológico ou impacte visual (Chuvienco et al., 1998).

A análise espacio-temporal dos limites geográficos dos incêndios, ajuda a estudar a correlação existente entre a ocorrência dos fogos e a área queimada causada pelos mesmos, assim como fornecer ferramentas de investigação para avaliar e identificar fatores relacionados com o risco de incêndio ou implementar medidas de prevenção. O recurso a imagens satélite e outras tecnologias afins facilitam a compreensão das relações espacio-temporais da incidência dos fogos florestais. O estudo dos limites do fogo proporciona informação sobre o seu comportamento e tamanho, assim como a influência que tem na estrutura da paisagem (Narayanaraj et al., 2011).

3.1 Dados

3.1.1 Regiões PROF

Os PROF – Planos Regionais de Ordenamento Florestal – são instrumentos setoriais para a organização e gestão dos espaços florestais a nível regional. Foram criados pelo Decreto-Lei nº 204/99 de 9 de Junho. Para a realização deste trabalho optou-se por utilizar o número de fogos referentes a cada uma das 21 regiões PROF que compõem o continente português. Encontra-se no Anexo 2 o mapa que representa estas regiões.

3.1.2 (Perímetro de) Áreas Ardidas

Anualmente, em Portugal, ocorre um elevado número de incêndios de pequena dimensão, mas a área afetada por estes fogos pode considerar-se pouco significativa em comparação com a área total ardida. Por este motivo e para este estudo, utilizaram-se os dados dos grandes incêndios relativos aos anos compreendidos entre 1996 e 2005 (ambos inclusive) por verificar a maior área afetada pelo fogo e também o aumento do número de ocorrências relativamente a anos anteriores. Outro critério estabelecido para a escolha deste período é

ter uma representação em torno do ano 2000, pela sua relação com o projeto *CORINE Land Cover 2000* (CLC2000).

Neste período de dez anos encontram-se os incêndios mais devastadores da história recente de Portugal, referentes a 2003 e 2005. Na Figura 3 mostrou-se a série cronológica da área ardida devida a incêndios florestais ocorridos na Península Ibérica no período de 1996 a 2005. Analisando o gráfico desta figura, com dados compilados pela AFN (2011), pode-se confirmar que Portugal é um país muito afetado pelos fogos florestais. Na década compreendida entre 1996 e 2005, houve sete anos em que os incêndios ultrapassaram os 100 mil hectares por ano. Em 2005 registou-se um grande número de área ardida (339 089 ha), seguido dos anos 2000 e 1998, mas foi em 2003 que os fogos tiveram o seu ponto mais alto até a atualidade, tendo sido queimados mais de 425 000 ha de terreno florestal. Por sua vez, 1997 é o ano com menor área afetada pelos fogos neste período de dez anos, com 30 535 ha ardidos.

Na Figura 5 está representada, discriminada por lustros, a área afetada pelos incêndios florestais que ocorreram em Portugal entre 1991 e 2005. No Anexo 3 mostra-se o mapa com a área afetada pelos grandes fogos.

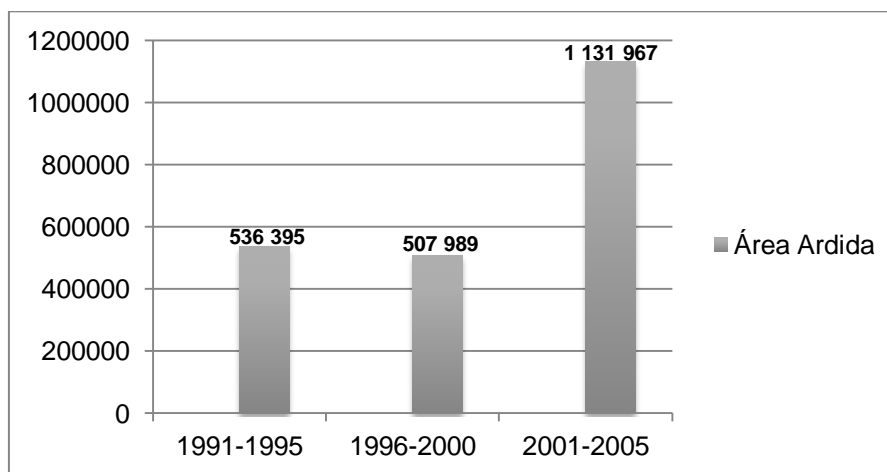


Figura 5. Discriminação, por Lustros, da Área Ardida Devido a Incêndios Florestais Ocorridos em Portugal no Período de 1991 a 2005 (AFN, 2011)

Para este estudo serão utilizados os perímetros das áreas ardidas cartografadas, a partir de imagens Landsat, pelo Laboratório de Detecção Remota e Análise Geográfica (LDRAG) do ex-Departamento de Engenharia Florestal (atual Departamento dos Recursos Naturais, Ambiente e Território – DRAT) do Instituto Superior de Agronomia, para a AFN. Será esta a

principal informação para realizar o presente trabalho. A variável a estudar será a existência ou não existência de fronteira de fogo, denominada por **Fronteira** (“CFRONTBIN”).

3.1.3 Variáveis Topográficas

A topografia é um fator muito importante não apenas na propagação do fogo mas também no seu combate. Para este estudo foi estabelecida como variável topográfica as **Linhas de Cumeada** (“COUNTRIDGE”), obtidas no Atlas da Água (<http://www.inag.pt/>). No Anexo 4 apresenta-se o mapa que representa as linhas de cumeada utilizadas no presente estudo.

O relevo acidentado da zona Centro de Portugal cria condições necessárias para a ocorrência dos fogos florestais, pois apresenta uma vegetação perenifólia e pirófito (Pereira et al., 2002). A topografia do terreno tem grande influência (1) no combustível: quantidade e continuidade da vegetação; e (2) nas condições atmosféricas: humidade e temperatura do ar e vento (Linn et al., 2007).

3.1.4 Rede Viária

Neste trabalho a rede viária é considerada uma condicionante que afeta o risco de incêndio, pois muitos fogos têm o seu início perto de estradas ou trilhos e mesmo caminhos de ferro. Mas também é um fator importante, pois facilita a intervenção dentro das florestas para sua manutenção e gestão, e também permite uma melhor ação à hora de combate ao fogo (Narayanaraj et al., 2011). Foi por isto decidido ter em conta a variável **Estradas** (“CROADSBIN”) como aquela que representa as estradas e caminhos nacionais. Os dados foram fornecidos por Estradas de Portugal (<http://www.estradasdeportugal.pt/>) e pelo Instituto Geográfico Português (<http://www.igeo.pt/>). No Anexo 5 representa-se a o mapa da rede viária de Portugal continental em que se juntaram os mapas proporcionados pelas duas entidades.

3.1.5 Área Urbanas

Foram usadas as áreas urbanas do MAOTDR – Ministério do Ambiente, do Ordenamento do Território e do Desenvolvimento Regional (atual Ministério da Agricultura, do Mar, do Ambiente e do Ordenamento do Território), por apresentar uma boa resolução espacial para o estudo em questão.

3.1.6 Rede Hidrográfica

A água que decorre pelo continente português é de grande interesse a nível de produção de energia (hidro)elétrica em barragens que se localizam nas suas bacias hidrográficas, mas também pelo seu contributo com a promoção da actividade agrícola e ainda pelo papel no auxílio prestado ao combate dos incêndios florestais (Viega da Cunha et al., 2006).

As bacias dos rios que se encontram a Norte e Oeste do país têm maior caudal e são mais regulares ao longo do ano do que os rios localizados a Leste e Sul. Mais uma vez fica marcada a diferença do Norte e do Sul do país. Rios Ave, Cávado, Leça, Lima e Minho são os que apresentam os valores mais elevados de precipitação média anual; a bacia do rio Tejo é a maior do país e marca a transição entre as regiões Norte e Sul (Viega da Cunha et al., 2006).

Trata-se de um fator de grande relevância para este estudo, pois as linhas de água podem ter um papel importante na propagação do fogo atuando como barreira física que impede o seu passo. Neste trabalho será usada como variável as **Linhas de Água** (“CRIVERBIN”) obtidas no Atlas da Água (<http://www.inag.pt/>). No Anexo 6 mostra-se o mapa da rede hidrográfica que compõe Portugal continental.

3.2 Material

3.2.1 ArcGIS

Um GIS (*Geographic Information System*) ou SIG (*Sistema de Informação Geográfica*) é uma base de dados espaciais que representam uma parte da realidade relevante para resolver problemas do mundo real, permite analisar e gerir informação geográfica correspondente a posições no terreno. Armazena informação cartográfica digital e tem anexado tabelas de atributos com dados descritivos utilizadas para realizar análises, gráficos e outros informes relativos aos dados espaciais (ESRI, 2011).

O ArcGIS é um programa SIG desenhado pela ESRI (*Environmental Systems Research Institute*), sendo o ArcMap a sua aplicação central, nele podem-se criar mapas, visualizar, consultar, analisar e apresentar os dados geográficos com as ferramentas apropriadas. Neste trabalho foi usado o programa **ArcGIS 10**.

3.2.2 GME

O GME (*Geospatial Modelling Environment*) é uma plataforma que proporciona um conjunto de ferramentas para facilitar a análise e modelação espacial, funciona em conjunto com o SIG e funciona à par do ArcMap, *i.e.*, todos os procedimentos efetuados no GME são abertos diretamente no ArcMap, na sua base de dados. Para efeitos deste estudo foi utilizada a **versão 0.7.1.0**.

3.2.3 CORINE Land Cover

O programa CORINE – *Co-ordination of Information on the Environment* – foi criado em 1985 pela Comissão Europeia com o intuito de desenvolver um sistema de informação sobre o estado do ambiente. Uma das componentes é a produção de uma cartografia da ocupação e uso do solo (Land Cover) para os países membros da União Europeia. Surge assim o projeto CLC90 (*CORINE Land Cover 1990*) como uma base de dados relativa a anos próximos de 1990 que será melhorada e atualizada em 1999 com o projeto CLC2000. O programa é coordenado pela Agência Europeia do Ambiente e desenvolvido a partir de imagens satélite (Painho e Caetano, 2006). No presente trabalho recorreu-se ao **CLC2000**.

O mapa da Figura 6 ilustra a ocupação/uso do solo em Portugal continental, segundo as classes de ocupação utilizadas no programa CORINE. As classes estão distribuídas em três níveis, correspondendo ao nível 1 as seguintes classes: 1.Territórios Artificializados, 2.Áreas

Agrícolas, 3.Florestas e meios semi-naturais, 4.Zonas Húmidas e 5.Massas de água. As classes mais representativas em Portugal são as Áreas Florestais e Agrícolas.

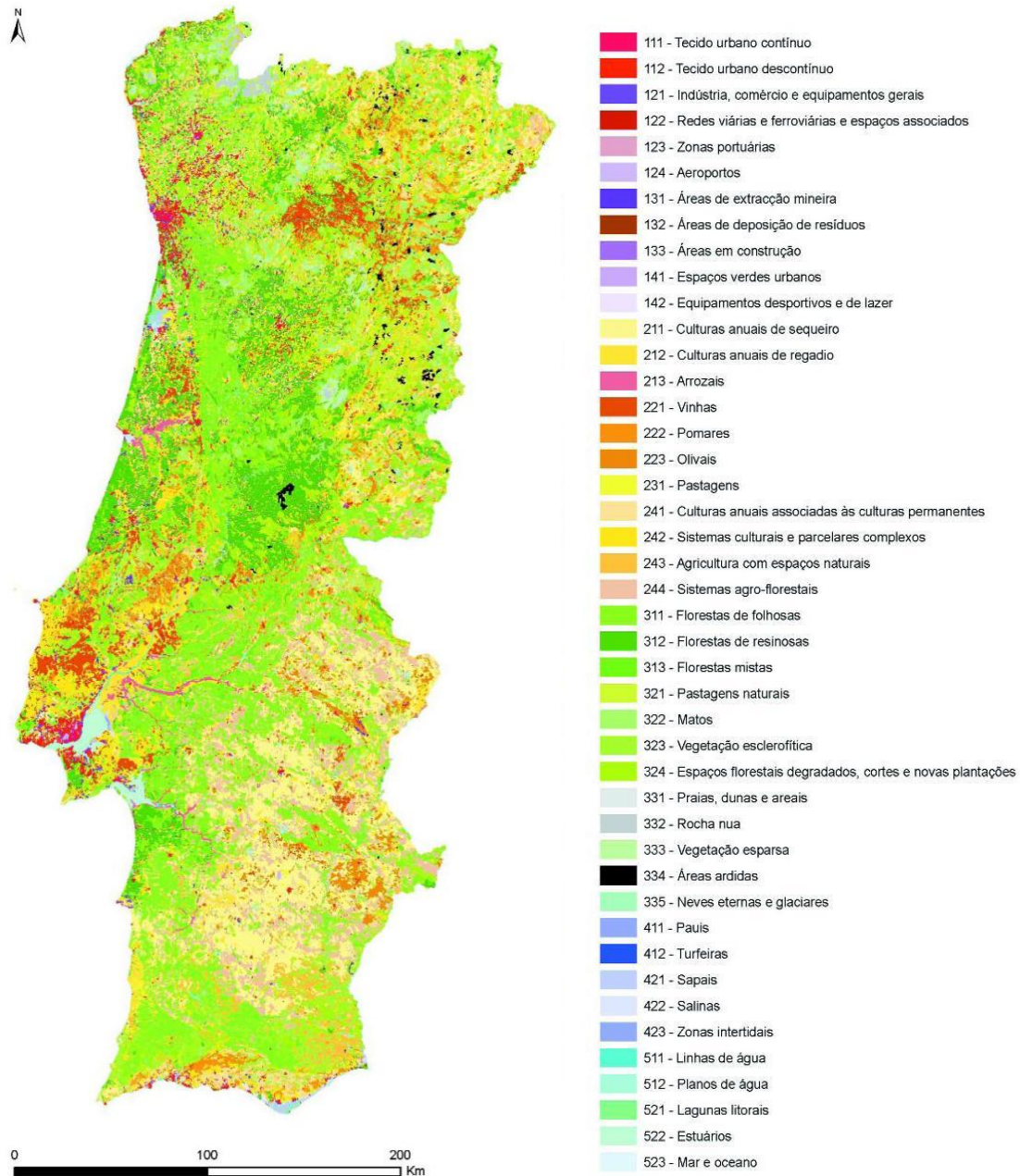


Figura 6. Mapa de Ocupação do Solo para Portugal segundo a Legenda do Programa CORINE Land Cover 2000.

3.2.4 CART

O CART – *Classification And Regression Trees* – é um programa baseado na estatística desenvolvida por quatro estatísticos na década de 1980: Leo Breiman, Jerome H. Friedman, R. A. Olshen e Charles J. Stone. A referência bibliográfica principal da análise deste estudo será a referente a Breiman et al. (1984).

O objetivo principal do CART é produzir um conjunto de classificadores mediante uma estrutura sequencial de predição para o problema em questão (Breiman et al., 1984). As árvores geradas podem ser usadas para obter modelos preditivos precisos e fiáveis (Steinberg e Golovnya).

A árvore consta de um nó inicial ou raiz, nós de divisão ou intermédios e nós terminais ou finais. Este conjunto de nós define-se através dos valores das variáveis. Os nós terminais correspondem às diferentes regiões da partição e esta é determinada por regras de divisão associadas aos nós intermédios. Para todas as variáveis e valores possíveis, o CART procura a melhor divisão correspondente àquela partição que verifica máxima homogeneidade dos dados. Este processo repete-se para cada uma das divisões dos dados resultantes (Timofeev, 2004).

Dependendo da tipologia das variáveis, o CART constrói diferentes árvores. No caso de se tratar de variáveis dependentes contínuas ou quantitativas produz árvores de regressão; se se tratar de variáveis dependentes categóricas ou qualitativas produz árvores de classificação. Neste trabalho foi usada a versão 6.6.0.091 para construir as árvores de classificação.

A metodologia da construção das árvores consiste em três partes essenciais, (1) construção da árvore máxima com maior número de nós, (2) combinação de ramos mediante poda para criar sub-árvores de diferentes tamanhos e (3) escolha da árvore ótima através das medidas precisas (Wei-Yin Loh, 2008; Yohannes e Webb, 1999).

A análise das árvores no CART começa com uma amostra heterogénea e complexa que é dividida em sub-amostras mais pequenas seguindo uns critérios ou regras de divisão, desta forma compara e determina qual das sub-amostras é mais homogénea detetando assim a estrutura da árvore menos complexa.

As regras de divisão dos nós representam questões referidas às observações. O CART inclui métodos como o *Gini* ou *Twoing* relativos às regras de divisão; estes têm demonstrado a eficiência das árvores e a informação referente aos dados. Neste trabalho foi eleito o índice de *Gini* como critério da divisão dos nós, está baseado na curva de Lorenz e é

representado por uma curva que mede o grau de distribuição/proporção das observações em que se verifica existência de fronteira de fogo (Figura 7).



Figura 7. Representação Gráfica do Índice Gini (adaptado de OEA, 1991)

A medida que melhora a equidade, isto é, diminui o valor do índice de *Gini*, a curva de Lorenz aproxima-se da linha de igualdade perfeita; quanto menor for o índice de *Gini* maior perfeição existirá na divisão dos nós.

3.2.4.1 Vantagens e Desvantagens do CART

Segundo vários autores (e.g., Breiman et al., 1984; Timofeev, 2004) existe uma série de vantagens que fazem com que este método tenha adquirido grande êxito. No entanto é também importante destacar as suas desvantagens.

Vantagens:

- método simples que fornece toda a informação necessária para descrever um modelo mediante características estatísticas;
- pode produzir resultados úteis utilizando apenas algumas das variáveis do conjunto total que foram utilizadas para a análise, pois seleciona as interações mais importantes de entre

um grande número de variáveis preditoras para explicar a variável resposta que se pretende estudar;

- técnica não paramétrica, as variáveis não seguem qualquer tipo de distribuição estatística;
- não requer a seleção prévia das variáveis, pois são identificadas as variáveis mais significativas e eliminadas as outras;
- lida facilmente com valores atípicos ou *outliers* atribuindo pesos iguais em todas as observações;
- não tem pressupostos e é rápido computacionalmente.

Desvantagens:

- pode criar árvores de decisão instáveis, pois eliminando algumas variáveis produz mudanças na complexidade da árvore, nas variáveis usadas na divisão ou nos seus valores;
- divide uma única variável de cada vez;
- a metodologia aplicada apenas assegura que cada divisão seja ótima, embora a árvore final pode não o ser;
- não está baseado num modelo probabilístico e portanto não existe um intervalo de confiança associado às predições da árvore que possa ajudar a classificar um novo conjunto de dados.

3.3 Método

3.3.1 Árvores de Classificação

O algoritmo usado apenas produz árvores de decisão binárias, ou seja, os nós inicial e intermédios são divididos unicamente em dois nós descendentes e mutuamente excludentes. Assim, as respostas às perguntas são do tipo Sim-Não ou Verdadeiro-Falso, de forma que a partir da amostra inicial é feita uma divisão em duas subamostras binárias (Yohannes e Webb, 1999).

Para construir uma árvore binária é preciso ter as observações da variável resposta ou dependente Y e das variáveis preditoras ou independentes X . As variáveis independentes ou preditoras também são chamadas covariáveis.

Esta análise tem dois objetivos fundamentais, (1) obter uma compreensão melhor das características das fronteiras das áreas ardidas em Portugal e (2) avaliar o potencial dos classificadores (fronteira do fogo) para efeitos de futuros estudos (adaptado de Yohannes e Webb, 1999).

O próprio algoritmo reduz significativamente a complexidade da estrutura dos dados enquanto a árvore cresce proporcionando a informação mais relevante das variáveis independentes. Do mesmo modo, fornece uma estimativa dos erros de classificação que se cometem nas classes da variável resposta e dos erros nos nós. Como foi referido anteriormente no ponto 3.2.4.1, as árvores são modelos não paramétricos e portanto não estão condicionados por qualquer pressuposto na distribuição de probabilidades dos dados (Chipman et al., 1998).

As fronteiras do fogo podem ser detetadas usando dados contínuos – que definem a localização no espaço – ou dados categóricos – que definem os lugares de mudança de coberto arbóreo (Fortin e Dale, 2005). Neste caso, para a construção das árvores com o CART, foi utilizado o modelo de classificação para variáveis categóricas.

3.3.1.1 Modelo de Árvore

A árvore de classificação é uma descrição da distribuição condicional de Y dado X , onde X é o vetor das variáveis preditoras. O fundamento desta metodologia consiste em repartir o espaço das variáveis independentes $X=(X_1, X_2, \dots, X_p)$ de forma tal que os valores da variável resposta Y sejam cada vez mais homogéneos dentro das classes J da partição.

Assim, para cada indivíduo X corresponde uma classe J consoante os valores observados para cada p covariáveis (Juárez e Castells, 2010).

Mediante a atribuição de critérios ou regras e partindo do nó inicial, que contém o número total de observações para a região em estudo, divide-se o conjunto em partes mais homogêneas, à sua vez cada uma destas novas partições será subdividida em outras de forma recursiva. Se houver alguma condição especial – igualdade de atributos para as variáveis, impossibilidade de fazer mais subdivisões homogêneas – produz-se a paragem da divisão. As divisões são selecionadas de modo que a impureza dos nós intermédios seja menor que a do nó inicial, *i.e.*, que haja maior homogeneidade nas observações (Timofeev, 2004).

3.3.1.2 Regras de Divisão

As regras ou critérios de divisão dos nós são referentes aos atributos das observações que definem a escolha de um valor associado a uma variável preditora (Breiman et al., 1984), de modo que cada vez os nós são mais puros do que os seus predecessores. Assim podera-se garantir a maior homogeneidade ou pureza das observações.

Existem vários critérios utilizados na divisão dos nós dependendo do algoritmo usado. Para escolher o melhor critério há que ter em conta os objetivos e o tipo de informação que se quer obter. Este estudo foi realizado com base ao método *Gini* por ser o mais adequado para variáveis resposta binárias. O índice de *Gini* representa uma estimativa do valor do erro de classificação e pode ser utilizado como medida de impureza do nó e portanto como critério para avaliar a qualidade de uma regra na obtenção da pureza de nós recorrentes em comparação com o antecedente, trata-se pois de uma regra de divisão assimétrica (Sá, 1999).

3.3.1.3 Erro de Classificação

Existem vários métodos para estimar o erro de classificação em função das características amostrais, o método preferível para este estudo é a validação cruzada (*V-fold cross validation*) que realiza um teste com 90% das observações para ter uma estimativa do modelo da árvore a usar, o teste é validado *a posteriori* com o 10% restante. Este método é repetido 10 vezes para cada região, obtendo uma maior gama na estimativa de erros, *i.e.*, o número de observações da amostra total é dividido em 10 subamostras de modo a conter um número similar de observações cada uma delas. Com base nisto é escolhida a melhor

alternativa ou árvore ótima como sendo aquela com o menor valor da estimativa do erro de classificação.

3.3.1.4 Árvore Ótima

Para escolher a árvore ótima o CART realiza uma série de operações, segue a continuação uma breve informação das ações realizadas pelo algoritmo (Breiman et al., 1984; Yohannes e Webb, 1999):

- a partir da amostra total é feita a divisão em dois grupos de nós intermédios usando o critério “Sim-Não”;
- o CART avalia a homogeneidade derivada da divisão com base no cálculo da proporção de cada categoria da variável resposta em cada nó;
- a seguir, seleciona a melhor divisão da variável que corrobore a redução da impureza (maior homogeneidade) atingida em cada uma das divisões;
- a cada nó atribui uma categoria de acordo com os critérios formulados que minimizam os custos do erro de classificação, criando uma matriz de erro que compara as categorias através da relação entre as classes observadas e as classes atribuídas pela classificação;
- estes passos são aplicados a cada nó intermédio em cada uma das sucessivas divisões;
- o CART continua com o processo de divisão e constrói uma árvore grande em que os nós terminais são mais puros;
- depois de gerar esta árvore, realiza uma poda nos resultados utilizando a validação cruzada e cria uma sequência de árvores das quais selecionará a árvore ótima em função da estimativa do erro de classificação e/ou do menor número de nós finais.

O método das árvores de classificação pode ser útil para grandes conjuntos de dados. Cada uma das divisões apresenta uma solução ótima, embora a otimização final da árvore possa ser contraditória devido à fragilidade das regras atribuídas (Maindonald e Braun, 2006). As opções escolhidas na parametrização do CART foram as seguintes:

- variáveis: especificou-se como variável resposta “**CFRONTBIN**” que indica se existe fronteira de fogo ou não dentro de uma determinada célula da grelha; como variáveis predictoras definiram-se as classes do **CORINE simplificado** (ver Quadro 4), as classes combinadas ou **interfaces** (ver ponto 7 da seção 3.4 Pré-processamento dos Dados), as **linhas de água**, as **estradas** e as **linhas de cumeada**;
- método de divisão dos nós: índice de **Gini**;
- estimação do erro de classificação: teste de validação cruzada (**V-fold cross validation**);

- seleção da melhor árvore: escolheu-se a menos complexa correspondente à que apresenta menor número de nós terminais, tendo em conta também o mínimo custo dos erros associado aos nós (***best tree within one standard error of minimum***).

3.4 Pré-processamento dos Dados

Para proceder à análise espacial das áreas ardidas em Portugal foi preciso o pré-processamento dos dados. Neste estudo, em particular, foram usados para esta análise os programas ArcGIS e GME.

1.- Inicialmente criou-se uma grelha de células de com resolução de 250 metros, utilizadas como unidade de análise, que abrange toda a área continental de Portugal.

2.- A seguir, foram excluídos os fogos com área inferior a 100 ha, ficando apenas com os fogos utilizados na análise.

3.- Simplificaram-se os perímetros das áreas ardidas aplicando o algoritmo Douglas-Peucker que alinha os pontos da imagem “pixelada”. Disponível no ArcMap.

4.- Reagruparam-se as classes do CLC2000, obtendo cinco novas classes mais adequadas ao objetivo do trabalho.

As cinco classes novas (classes simplificadas do CORINE) são identificadas a seguir, assim como o código do CORINE original. Daqui em diante poderão ser designadas resumidamente com os códigos: CLCC1 (Área Artificial), CLCC2 (Agricultura), CLCC3 (Florestas), CLCC4 (Matos) e CLCC5 (Água). As alterações feitas no CORINE original para a construção das classes simplificadas são referidas no Quadro 4. Uma expressão gráfica das Classes do CORINE Land Cover 2000 e das novas classes podem ser observadas nas Figuras 9 e 10, respetivamente (detalhe da ocupação do solo para o Distrito de Castelo Branco).

Quadro 4. Classes Simplificadas do CORINE Land Cover em que se Reagruparam Diferentes Classes da Legenda do CORINE Land Cover Original..

Código	Classes Simplificadas	Classes do CORINE Unidas
CLCC1	Área Artificial	1.1.1, 1.1.2, 1.2.1, 1.2.2, 1.2.3, 1.2.4, 1.3.1, 1.3.2, 1.3.3, 1.4.1, 1.4.2, 3.3.1, 3.3.2, 3.3.4 e 3.3.5
CLCC2	Agricultura	2.1.1, 2.1.2, 2.1.3, 2.2.1, 2.2.2, 2.2.3, 2.3.1, 2.4.1, 2.4.2, 2.4.3 e 2.4.4
CLCC3	Florestas	3.1.1, 3.1.2 e 3.1.3
CLCC4	Matos	3.2.1, 3.2.2, 3.2.3, 3.2.4 e 3.3.3
CLCC5	Água	4.1.1, 4.1.2, 4.2.1, 4.2.2, 4.2.3, 5.1.1, 5.1.2, 5.2.1, 5.2.2 e 5.2.3

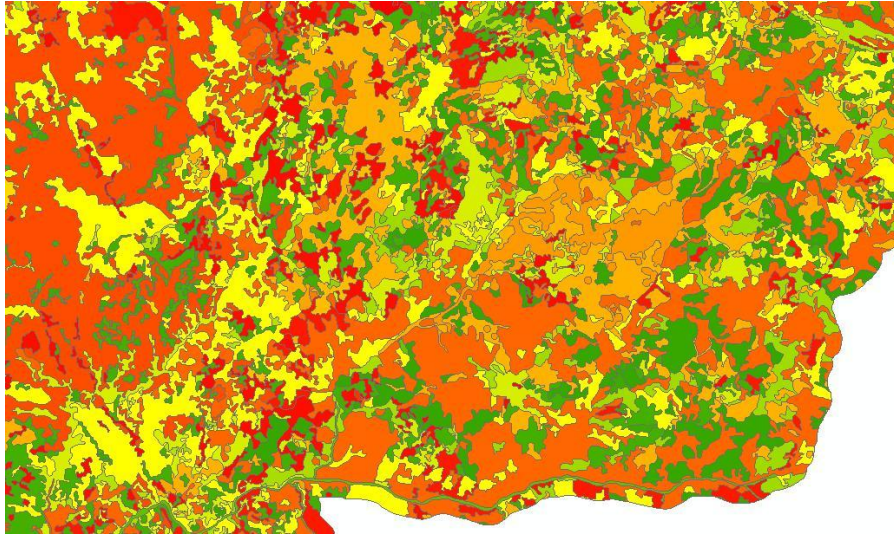
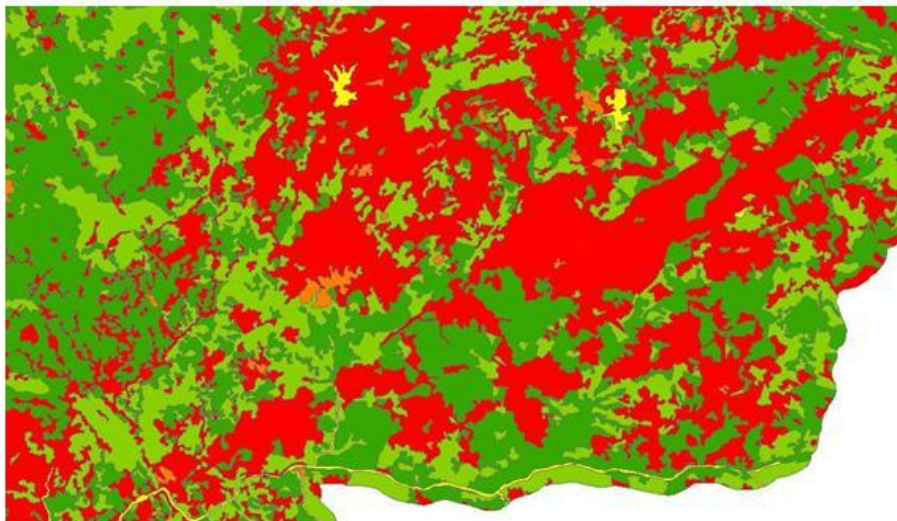


Figura 8. Classes de Ocupação do Solo, CORINE Land Cover 2000, Localizadas no Distrito de Castelo Branco.



Legenda

- Área Artificial
- Agricultura
- Florestas
- Matos
- Água

Figura 9. Classes Simplificadas de Ocupação do Solo a partir do CORINE Land Cover 2000, Localizadas no Distrito de Castelo Branco.

5.- Sendo que a informação das áreas urbanas do MAOTDR tem mais detalhe do que o CORINE Land Cover, foram compatibilizados os dados das duas fontes de informação. Assim, foi feita a união, entre as áreas urbanas do MAOTDER e as do CORINE simplificado.

6.- Realizaram-se interseções entre a quadrícula e cada um dos fatores (rede viária, ocupação do solo, rede hidrográfica e topografia).

7.- Fizeram-se combinações das classes simplificadas do CORINE. Para este efeito foi preciso recorrer a programação no MATLAB (*MATrix LABoratory*) com o fim de se combinarem apenas duas classes de ocupação do solo, aquelas que representam maior área em cada célula. As possíveis combinações de classes foram: Classe1-2 (combina a classe artificial com a classe agricultura), Classe1-3 (combina artificial e florestas), Classe1-4 (combinação de artificial e matos), Classe1-5 (combina artificial e água), Classe2-3 (combina agricultura e florestas), Classe2-4 (combina as classes agricultura e matos), Classe2-5 (combina agricultura e água), Classe3-4 (combinação de florestas e matos), Classe3-5 (combina florestas e água), Classe4-5 (combina a classe matos e a classe água).

8.- Efetou-se o cruzamento de toda a informação produzida com os limites das regiões PROF, de modo a permitir uma análise regional de relação entre as fronteiras das áreas áridas e os fatores ambientais.

4. ANÁLISE DE RESULTADOS

Recorrendo ao programa CART procedeu-se a uma análise para cada região PROF de modo a salientar diferenças regionais em Portugal continental. Assim, as 1 447 327 observações obtidas no período de 1996 a 2005 e para a totalidade do país (correspondentes aos elementos da quadrícula/célula de 250 m) foram divididas pelas 21 regiões PROF.

Foi definida como variável resposta a variável binária que regista a presença ou ausência de fronteiras de áreas aridas ao longo do período da análise (denominada por “CFRONTBIN”). Da mesma forma transformaram-se todas as variáveis predictoras em variáveis binárias: “CLCC1”, “CLCC2”, “CLCC3”, “CLCC4”, “CLCC5”, “COUNTRIDGE”, “CRIVERBIN”, “CROADSBIN” e as variáveis de interface que combinam as diferentes classes de ocupação do solo (Classe1-2, Classe1-3, Classe1-4, Classe1-5, Classe2-3, Classe2-4, Classe2-5, Classe3-4, Classe3-5 e Classe4-5).

Para poder explicar o funcionamento do programa e a modo de exemplo, foi escolhida a região “Pinhal Interior Sul”, de forma que todos os comentários relativos às escolhas feitas para a elaboração das árvores de classificação, assim como alguns dos resultados obtidos com o CART, serão baseadas neste PROF (Anexo 5). Mesmo assim procedera-se a uma revisão concisa dos resultados obtidos nas restantes regiões.

TREE SEQUENCE (Sequência das Árvores)

O *output* fornece um quadro no que se representa a sequência das árvores geradas a partir da poda da árvore de maior tamanho, identificando a que apresenta um custo mínimo e a árvore ótima. Também proporciona o custo do erro de classificação geral gerado pela poda da árvore ótima ao aplicar a validação cruzada. A medida que o número de nós diminui este custo aumenta.

Na Figura 10 pode visualizar-se a árvore ótima de classificação (a árvore original e detalhada encontra-se no Anexo 7). Para este PROF em concreto, o CART elaborou 19 árvores de classificação, desde uma árvore com um único nó terminal até uma com 135 nós finais.

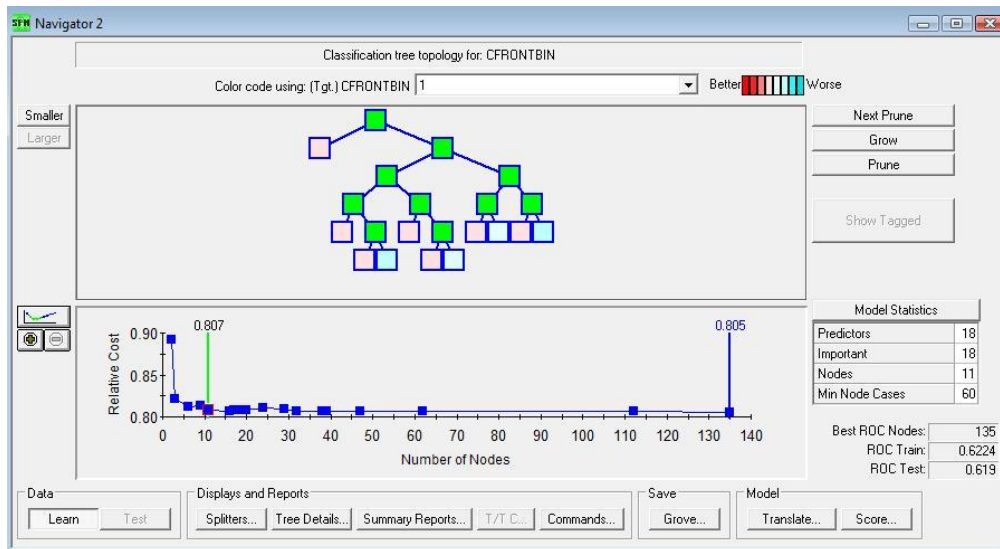


Figura 10. Árvore Ótima e Curva de Custos para o Pinhal Interior Sul.

Na parte superior da figura estão representados os nós das divisões binárias a cor verde (nó raíz e nós intermédios) e os nós terminais a cor de rosa (nós mais homogéneos) e a azul (nós menos homogéneos). Na sua elaboração foram usadas 18 variáveis predictoras e consta de 11 nós finais.

Na parte baixa da figura está representada a curva dos custos relativos, na escala de 0 a 1, de acordo com o número de nós finais que definem cada uma das árvores. A melhor alternativa tem um valor de 0.805 correspondente à Árvore 1 (135 nós finais), verificando o custo mínimo, embora o custo da árvore ótima (Árvore 14) seja 0.807. O valor deste custo é o valor do erro global de classificação que avalia a qualidade da árvore.

Outra maneira de avaliar a precisão ou sensibilidade do classificador é dada pela interpretação da curva ROC – *Receiver Operating Characteristic* –, representada na Figura 11. Os pontos da curva correspondem a diferentes limiares usados pelo classificador. Tal como o índice de *Gini*, esta curva também se baseia na curva de Lorenz.

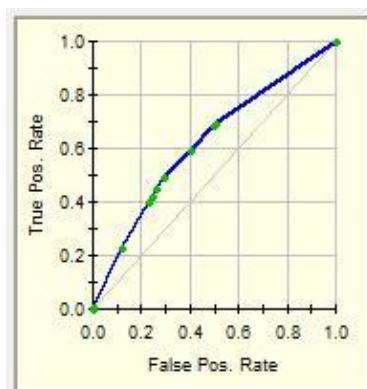


Figura 11. Curva ROC que Representa os Verdadeiros Positivos (Observações da Classe 1) Vs. Lalsos Positivos (Observações da Classe 0).

Um classificador que segue a linha de 45° não seria útil, pois indica que se classifica a mesma proporção de dados para a Classe 0 e a Classe 1, *i.e.* a não existência e a existência de fronteira, respetivamente. O melhor classificador será aquele que segue o eixo vertical, sendo que neste caso se classificaria a totalidade das observações da Classe 1 e nenhuma observação da Classe 0. Com isto pode-se afirmar que quanto mais próxima estiver a curva ROC da recta de 45° pior será o classificador quanto à capacidade de discriminar as classes, o classificador ideal será aquele cuja curva ROC se afasta mais desta recta (Hand, 1997). A exatidão é maior quanto mais se aproximar a curva ROC do canto superior esquerdo do diagrama, ou seja, quanto maior for a área entre a curva ROC e a diagonal do gráfico.

NODE INFORMATION (Informação dos Nós)

Outras das informações principais dadas pelo CART dizem respeito ao nó raiz e aos nós intermédios. O nó inicial apresenta a maior impureza ou heterogeneidade, pois nele estão presentes todas as observações possíveis para a região do Pinhal Interior Sul. Este valor é dado pelo *improvement* e será menor nos nós sucessores, de forma que para o Nó 1 (nó raiz) tem um valor de 0.0104 e para o Nó 2 o seu valor desce para 0.0077.

Uma tabela mostra a distribuição das observações nas diferentes classes e nós junto com os custos associados. Assim, o Nó 1 tem um total de 31 235 observações, das quais 26 916 passam ao Nó 2 com um custo associado de 0.4677 e 4 319 dados passam ao Nó -1 com um custo de 0.3387, o sinal negativo indica que este é um nó terminal.

Ainda na informação referente aos nós, se pode observar a relação das variáveis que competem com as variáveis das regras primárias da divisão. Qualquer uma pode ser utilizada mas tem preferência aquela com maior valor no *Improvement*, pois indica que a redução da heterogeneidade é maior relativamente às outras variáveis concorrentes.

TERMINAL NODE INFORMATION (Informação dos Nós Terminais)

Fornece a informação referente aos nós terminais. Em concreto para o Nó Terminal 1 existem 4 319 observações estando divididas nas duas categorias: Classe 0 com 2 699 observações e Classe 1 com 1 620 observações. Também apresenta o valor do erro de classificação (0.3387) que é maior quanto mais heterogéneo for o nó. De entre todos os nós observa-se um mais puro que os outros e neste PROF corresponde ao Nó Terminal 6 (0.3113) que tem 84 observações; o mais impuro é o Nó Terminal 7 com 3 401 observações e um erro de 0.4886.

MISCLASSIFICATION BY CLASS (Observações Erradas em cada Classe)

É fornecido também um quadro que resume o número de observações em cada uma das categorias assim como o número de observações mal classificadas e os custos associados à respetiva classe, resultantes da classificação errónea. Corresponde à matriz de erro dos nós finais (Quadro 5), em que para a Classe 0 o erro determinado com a validação cruzada é de 32.18% (7 688 observações mal classificadas), e para a Classe 1 é de 48.56% (3 566 observações mal classificadas). O seu objetivo é avaliar a precisão da classificação da árvore mediante a proporção das observações que foram corretamente classificadas para cada classe.

Quadro 5. Matriz de Erro que Representa as Observações Totais Bem/Mal Classificadas em cada Categoria ou Classe e a Precisão da Classificação (% correto).

Classe	Observ. Classe	Classe 1	Classe 0	% correto
1	7 344	3 778	3 566	51.44
0	23 891	7 688	16 203	67.82
TOTAL	31 235	11 466	19 769	

VARIABLE IMPORTANCE (Importância das Variáveis)

Proporciona a lista de todas as variáveis, quer as utilizadas na construção da árvore quer as não utilizadas. Compara a capacidade de cada uma delas para decidir qual é a melhor para realizar a classificação, ou seja, qual delas é a que deve ser usada para a divisão de cada nó. Dependendo da importância de cada variável avalia-se o seu contributo na divisão da variável resposta. Uma variável pode ter grande importância e mesmo assim não aparecer na divisão dos nós do classificador; isto deve-se ao efeito desta variável estar oculta por outra variável que produz uma classificação igualmente precisa (Breiman et al., 1984).

Para o Pinhal Interior Sul, observou-se que a variável correspondente com as áreas artificiais (CLCC1) é a mais importante para explicar a presença de fronteiras do fogo. Apresenta-se na Figura 12 a percentagem da importância das diferentes variáveis para esta região, sendo mais importantes aquelas relacionadas com o Land Cover.

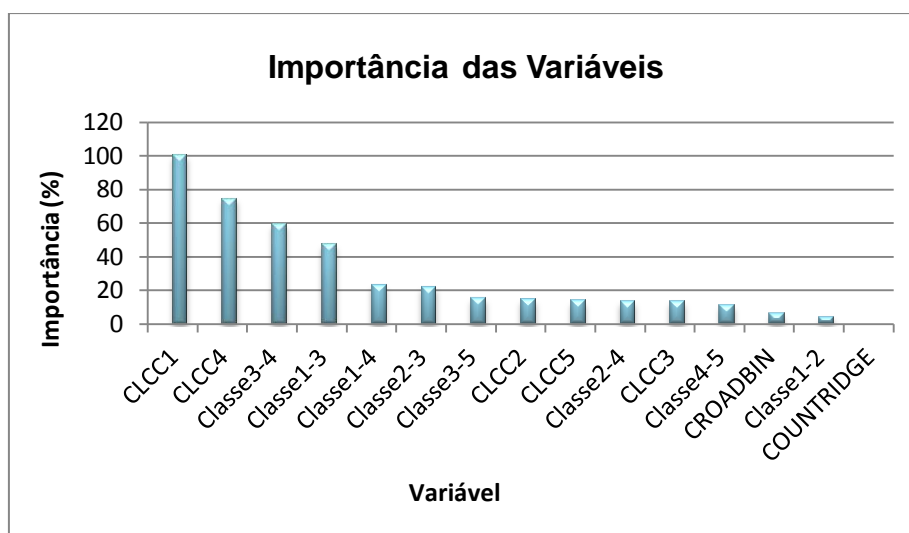


Figura 12. Importância das Variáveis Predictoras para o Pinhal Interior Sul.

No Anexo 7 apresenta-se toda a informação do *output* obtida e não mencionada neste capítulo para o Pinhal Interior Sul.

Seguidamente serão apresentados os dados obtidos com a utilização do CART para cada uma das regiões PROF. Serão usados quadros que auxiliam a interpretação dos diferentes *output* produzidos.

No Quadro 6 representam-se os principais resultados correspondentes às árvores de classificação obtidas para as diferentes regiões, em que se observa uma clara dominância da variável CLCC4 correspondente à presença de matos, que se manifesta como a variável preditora mais importante em 13 regiões.

Quadro 6. Resumo do Output obtido no CART para cada PROF.

PROF	Nós Terminais	Observações	Erro Global	Variável Import	ROC	% Correta Global
Alto Minho	4	36 177	0.6462	CLCC4	0.6877	64.19
Baixo Minho	9	40639	0.4635	CLCC4	0.8118	70.70
Barroso e Padrela	3	47 242	0.7446	CLCC4	0.6284	58.27
Nordeste	4	85 537	0.6512	CLCC4	0.6741	58.26
AMP	5	27 696	0.5228	CLCC4	0.8074	71.80
Tâmega	6	42 799	0.5692	CLCC4	0.7319	68.35
Douro	12	66 861	0.6533	CLCC4	0.6998	60.41
Centro Litoral	6	91 235	0.7399	CLCC3	0.6740	54.45
Dão-Lafões	2	56 719	0.5590	CLCC4	0.7205	68.49
Beira Interior Norte	4	102 242	0.7356	CLCC4	0.6372	51.74
Pinhal Interior Norte	6	42 783	0.8468	CLCC3	0.5890	53.09
Pinhal Interior Sul	11	31 235	0.8074	CLCC1	0.6189	63.97
Beira Interior Sul	16	61 023	0.7695	CLCC4	0.6483	58.46
Oeste	3	36 360	0.6093	CLCC4	0.7295	60.29
Ribatejo	7	106 763	0.7015	CLCC3	0.6987	62.86
Alto Alentejo	12	101 371	0.6967	CLCC4	0.6954	67.04
AML	3	48 469	0.5854	CLCC3	0.7216	61.60
Alentejo Central	8	117 201	0.7835	CLCC2	0.6242	58.27
Alentejo Litoral	10	85 473	0.8045	CLCC4	0.6440	56.95
Baixo Alentejo	5	138 274	0.6755	CLCC3	0.6803	68.54
Algarve	7	81 228	0.7025	CLCC3	0.6919	61.21
Total		1 447 327				

A árvore mais precisa que se constata neste quadro, é a referente à região do "Baixo Minho" que tem o menor erro de classificação global (0.4635) assim como o maior valor do teste ROC (0.8118). Por sua vez, a região "Pinhal Interior Norte" apresenta a árvore menos precisa tendo o maior erro de classificação (0.8468) e o menor valor para o teste ROC (0.5890).

A última coluna (% correta global ou precisão global) diz respeito da percentagem de observações totais bem classificadas (Classe 0 e Classe 1). Sendo o seu valor para o Pinhal Interior Sul de 63.97%, confirma-se que 19 981 (16203+3778) observações das 31 235 totais foram bem classificadas. Dados comparáveis com o Quadro 5.

No Quadro 7 é mostrada a variável que representa maior concorrência com a selecionada para a divisão do nó raiz, também se mostra a relação entre as variáveis utilizadas na divisão do nó raiz e as variáveis com maior importância, para cada uma das regiões.

Quadro 7. Variáveis Primárias, Variáveis Importantes e Variáveis Concorrentes para cada região PROF.

PROF	Variável Usada no Nó Raiz	Variável mais Importante	Variável mais Concorrente	Redução Impureza da Concorrente
Alto Minho	CLCC4	CLCC4	CLCC4	0.062571
Baixo Minho	CLCC4	CLCC4	CLCC4	0.135573
Barroso e Padrela	CLCC4	CLCC4	CLCC4	0.03234
Nordeste	CLCC4	CLCC4	CLCC4	0.053111
AMP	CLCC4	CLCC4	CLCC4	0.102218
Tâmega	CLCC4	CLCC4	CLCC4	0.092389
Douro	CLCC4	CLCC4	CLCC4	0.063099
Centro Litoral ⁽¹⁾	CLCC4	CLCC3	CLCC4	0.017627
Dão-Lafões	CLCC4	CLCC4	CLCC4	0.098549
Beira Interior Norte	CLCC4	CLCC4	CLCC4	0.037836
Pinhal Interior Norte ⁽¹⁾	CLCC4	CLCC3	CLCC4	0.009111
Pinhal Interior Sul	CLCC1	CLCC1	CLCC1	0.010418
Beira Interior Sul	CLCC4	CLCC4	CLCC4	0.021606
Oeste	CLCC4	CLCC4	CLCC4	0.05894
Ribatejo ⁽¹⁾	CLCC4	CLCC3	CLCC4	0.023421
Alto Alentejo	CLCC4	CLCC4	CLCC4	0.03472
AML ⁽¹⁾	CLCC4	CLCC3	CLCC4	0.041012
Alentejo Central ⁽¹⁾	CLCC3	CLCC2	CLCC3	0.015199
Alentejo Litoral	CLCC4	CLCC4	CLCC4	0.022439
Baixo Alentejo ⁽¹⁾	CLCC2	CLCC3	CLCC2	0.039778
Algarve	CLCC3	CLCC3	CLCC3	0.037513

Como se constata no Quadro 7, nem sempre é usada a variável mais importante na elaboração da árvore ótima tal como acontece nos PROF identificados com ⁽¹⁾. Nas restantes regiões a árvore é iniciada tendo em conta a variável mais importante e também a de maior concorrência, em todos os casos coincidem estas duas variáveis.

No Quadro 8 apresenta-se o número de observações totais e o número de observações mal classificadas em cada classe (usando validação cruzada) para as 21 regiões continentais. Destacam algumas regiões por apresentarem uma elevada diferença entre o número de observações de não existência de fronteira do fogo e o número de observações de presença de dita fronteira.

Quadro 8. Classes Erradamente Classificadas para Cada Região.

PROF	Observações totais		Classificações erradas		% Erro por Classe	
	Classe 0	Classe 1	Classe 0	Classe 1	Classe 0	Classe 1
Alto Minho	28 885	7 292	11 027	1 928	38.18	26.44
Baixo Minho	35 293	5 346	11 115	794	31.49	14.85
Barroso e Padrela	35 865	11 377	16 467	3 248	45.91	28.55
Nordeste	78 694	6 843	34 226	1 480	43.49	21.63
AMP	24 869	2 827	7 143	666	28.72	23.56
Tâmega	31 500	11 299	1 194	2 452	35.22	21.70
Douro	50 471	16 390	23 347	3 125	46.26	19.07
Centro Litoral	84 869	6 366	39 838	1 722	46.94	27.05
Dão-Lafões	45 835	10 884	15 461	2 413	33.73	22.17
Beira Interior Norte	76 763	25 479	45 798	3 541	59.66	13.90
Pinhal Interior Norte	36 232	6 551	17 727	2 342	48.93	35.75
Pinhal Interior Sul	23 891	7 344	7 688	3 566	32.18	48.56
Beira Interior Sul	54 330	6 693	23 038	2 312	42.40	34.54
Oeste	35 936	424	14 351	89	39.93	20.99
Ribatejo	96 871	9 892	36 428	3 219	37.60	32.54
Alto Alentejo	95 937	5 434	31 403	2 007	32.73	36.93
AML	47 508	9 61	18 420	190	38.77	19.77
Alentejo Central	114 887	2 314	48 065	845	41.84	36.52
Alentejo Litoral	83 218	2 255	35 952	840	43.20	37.25
Baixo Alentejo	136 327	1 947	42 793	704	31.39	36.16
Algarve	74 029	7 199	29 300	2 208	39.58	30.67

No Anexo 8 estão representadas as árvores de classificação obtidas para cada uma das regiões PROF. A seguir é feita uma breve interpretação de cada árvore correspondente a cada uma destas regiões. Referem-se os aspetos mais relevantes da árvore ótima, exemplificando-se com a região do “Alto Minho”, cuja árvore se mostra na Figura 13.

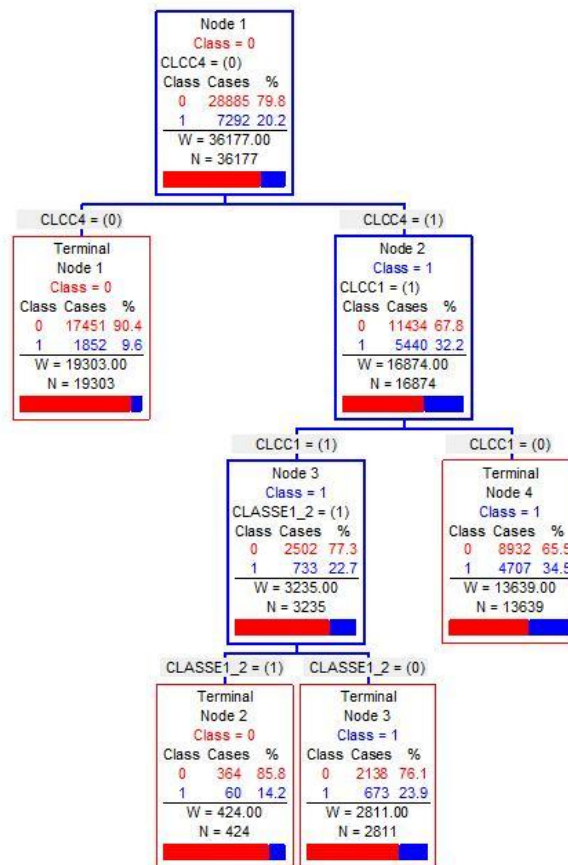


Figura 13. Árvore Ótima obtida para a Região de Alto Minho.

A árvore obtida para esta região PROF começa com a variável de ocupação do solo referente aos matos (CLCC4) produzindo uma primeira divisão com dois nós binários (0, 1). No nó da esquerda pode ler-se ausência de matos (CLCC4=0) que corresponde ao Nó Terminal 1 e não apresenta fronteira de fogo (Class=0). No nó da direita (Nó 2) a presença de matos (CLCC4=1) combina-se com a ausência de áreas artificiais (CLCC1=0) para produzir o Nó Terminal 4 correspondente à presença de fronteiras de áreas ardidas (Class=1). No Nó 3, a presença de áreas artificiais (CLCC1=1) e a presença/ausência da combinação de classes 1 e 2 (áreas artificiais e agricultura) resulta na ausência/presença de fronteiras de áreas ardidas nos Nós Terminais 2 e 3, respetivamente. Neste exemplo, o algoritmo identifica as características dos limites das áreas ardidas tendo em consideração apenas 3 das 18 variáveis disponíveis.

De maneira mais sucinta e resumida serão comentadas as árvores para as restantes regiões. A presença de fronteira do fogo pode ser explicada seguindo um ou vários conjuntos de relações que conduzem aos nós terminais pertencentes à Classe 1 (Class=1), tal como se demonstra a continuação no Quadro 9.

Quadro 9. Leitura-Resumo das Árvores de Classificação para as 21 Regiões PROF.

Alto Minho:

- presença de matos e ausência de áreas artificiais
- presença de matos e área artificial e ausência da interface artificial-agricultura

Baixo Minho:

- presença de matos e ausência de artificial e água
- presença de matos e artificial e ausência da interface artificial-agricultura
- ausência de matos, agricultura, artificial, água e rios

Barroso-Padrela:

- presença de matos e ausência de água

Nordeste:

- presença de matos
- ausência de matos e estradas e presença de área artificial

Área Metropolitana de Porto e entre Douro e Vouga (AMP):

- presença de matos
- ausência de matos, artificial e interface agricultura-floresta e presença de floresta

Tâmega:

- presença de matos e ausência de agricultura
- presença de matos, agricultura e interface agricultura-matos
- presença de matos e agricultura com ausência da interface agricultura-matos, artificial-agricultura e água

Douro:

- presença de mato e floresta
- presença de mato e agricultura e ausência de floresta e água
- presença de mato com ausência de floresta, agricultura e interface mato-água
- presença de mato e da interface mato-água, ausência de floresta, agricultura e rios
- ausência de matos, floresta, agricultura, estradas e cumeadas, presença de área artificial

Centro Litoral:

- presença de matos
- ausência de matos e presença de floresta e estradas
- matos e estradas ausentes e floresta e estradas presente
- ausência de matos, estradas e áreas artificiais com presença de floresta e agricultura

Quadro 9. Leitura-Resumo das Árvores de Classificação para as 21 Regiões PROF (Continuação).

Beira Interior Norte:

- presença de matos
- ausência de matos e estradas e presença de áreas artificiais

Pinhal Interior Norte:

- presença de matos
- ausência de matos e presença de floresta e da interface artificial-floresta
- ausência de matos, agricultura e da interface artificial-floresta com presença de florestas e estradas

Pinhal Interior Sul:

- presença de área artificial
- ausência de área artificial, matos e agricultura-floresta com presença de interface floresta-água
- ausência de área artificial e matos e presença da interface agricultura-floresta e estradas
- ausência de área artificial e cumeadas e presença de matos e floresta
- ausência de área artificial e presença de matos, floresta, cumeadas e estradas
- ausência de área artificial e floresta com presença de matos e agricultura
- ausência de área artificial, floresta e agricultura e presença de matos e água

Beira Interior Sul:

- presença de matos
- ausência de matos e presença de florestas e área artificial
- ausência de matos e área artificial com presença de floresta, agricultura e água
- ausência de matos, área artificial, água e cumeadas com presença de floresta, agricultura e rios
- ausência de matos, artificial, água, cumeadas e rios mas presença de floresta, agricultura e estradas
- ausência de matos, artificial e agricultura com presença de floresta e estradas
- ausência de matos, floresta, estradas e agricultura e presença de área artificial
- ausência de matos, floresta e estradas; presença de área artificial, agricultura e cumeadas
- ausência de matos, floresta e artificial; presença de estradas e da interface agricultura-água

Oeste:

- presença de matos
- ausência de matos e presença de floresta

Quadro 9. Leitura-Resumo das Árvores de Classificação para as 21 Regiões PROF (Continuação).

Ribatejo:

- presença de matos
- ausência de matos e presença de florestas e artificial
- ausência de matos e artificial com presença de florestas e estradas
- ausência de matos, artificial e estradas e presença de florestas e água
- ausência de matos, área artificial, estradas e água e presença de florestas e agricultura

Alto Alentejo:

- presença de matos
- ausência de matos e presença de floresta e estradas
- ausência de matos e estradas com presença de floresta e água
- ausência de matos, estradas e água mas presença de floresta e área artificial
- ausência de matos, estradas, água, área artificial e rios; presença de floresta e agricultura
- ausência de matos e florestas e presença de área artificial
- ausência de matos, floresta e artificial; presença de estradas e o binómio agricultura-água
- ausência de mato, floresta, artificial e interface agricultura-água; presença de estradas e cumeadas

Área Metropolitana de Lisboa (AML):

- presença de matos
- ausência de matos com presença de florestas

Alentejo Central:

- presença de florestas e matos
- presença de florestas e ausência de matos e agricultura
- presença de floresta, agricultura e cumeadas com ausência de matos e área artificial
- presença de floresta e agricultura e ausência de matos, artificial, cumeadas e estradas
- ausência de florestas e presença de matos

Alentejo Litoral:

- presença de matos e ausência de água e da interface artificial-mato
- ausência de matos e água com presença florestal e cumeadas
- ausência de matos, água, cumeadas, área artificial, agricultura e estradas com presença de florestas

Quadro 9. Leitura-Resumo das Árvores de Classificação para as 21 Regiões PROF (Continuação).

Baixo Alentejo:

- ausência de agricultura e área artificial
- presença de agricultura e floresta
- presença de agricultura e matos e ausência de floresta

Algarve:

- presença de floresta
- ausência de florestas e presença de matos e mato-água
- ausência de floresta, mato-água e artificial-mato, presença de matos e agricultura-mato
- ausência de floresta, mato-água, artificial-mato e agricultura-mato com presença de matos e cumeadas

Na Figura 14 é representada a média da importância das variáveis predictoras para o conjunto total das regiões.

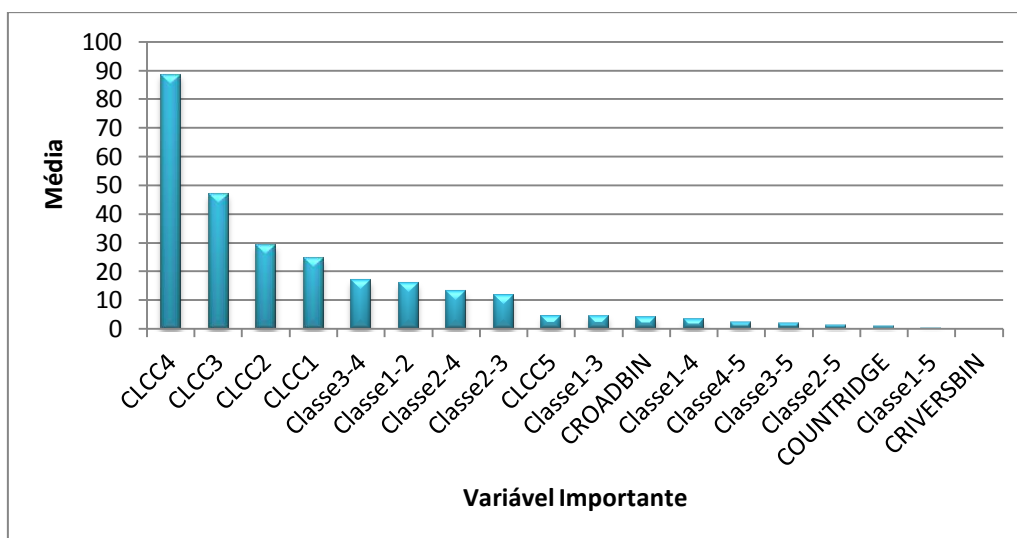


Figura 14. Médias das Variáveis Predictoras mais Importantes para as 21 Regiões PROF.

A nível geral para a totalidade das regiões continentais, as variáveis predictoras que revelam maior importância são aquelas relacionadas com o Land Cover. Anteriormente, na Figura 12 para o Pinhal Interior Sul também estas tinham tido maior relevância na fronteira do fogo.

Para poder fazer uma comparação da importância das variáveis preditoras entre os diferentes PROF em que se divide o país, criou-se um mapa (Figura 15) que representa a variável mais importante para determinar a presença de fronteiras de áreas ardidas.

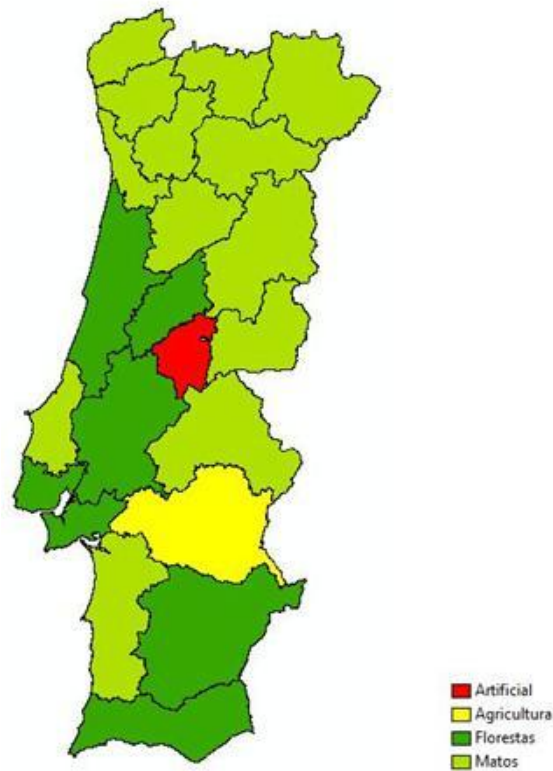


Figura 15. Mapa de Distribuição da Importância das Variáveis.

Pode observar-se na Figura 15 que a variável preditora mais importante na maior parte das regiões PROF é a presença de matos (CLCC4), seguida pelas florestas, agricultura e zonas artificiais. Esta figura e a Figura 14 indicam a preferência do algoritmo de classificação pelas variáveis relacionadas com a ocupação ou uso do solo derivadas da simplificação da legenda CORINE. Em menor medida prefere as variáveis combinadas, nomeadamente as interfaces matos-florestas, artificial-agricultura, agricultura-matos e agricultura-florestas. As variáveis menos utilizadas nas árvores de classificação correspondem a outras combinações de ocupação do solo e às linhas de água, estradas e linhas de cumeada.

5. DISCUSSÃO E CONCLUSÕES

Existem vários critérios utilizados na seleção da árvore ótima. Um dos métodos para selecionar esta árvore é estabelecido pela precisão da classificação que é obtida com a proporção das observações corretamente classificadas em cada categoria. Outro está baseado no número de nós finais, neste caso a árvore menos complexa será a escolhida, este último também está relacionado com o erro de classificação pois quanto maior é o número de nós terminais menor custo tem o erro.

Breiman et al. (1984) sugerem que a árvore ótima seja aquela com menor número de nós terminais mas que se encontre dentro dum erro estândar da árvore de custo mínimo. Deste modo, a árvore ótima eleita pelo algoritmo será aquela que minimize o erro de classificação global e que tenha um número baixo de nós terminais. Árvores muito complexas com elevado número de nós finais conduzem a um classificador mais específico em que se reduz o número de observações nos nós terminais e também à redução da possibilidade de existirem *outliers* ou observações estranhas.

Quer o erro de classificação quer a curva ROC são dois métodos igualmente válidos para avaliar a precisão do classificador. Neste contexto a árvore mais precisa é a referida à região do Baixo Minho por apresentar o erro mais baixo de todas as regiões e o valor mais elevado da curva ROC, tendo um total de 9 nós finais. A árvore correspondente ao Pinhal Interior Norte é a menos precisa de todas com o erro mais elevado e o valor da curva ROC mais baixo, apresenta 6 nós terminais (Quadro 6).

Em relação ao Quadro 8 confirma-se que as árvores com mais observações classificadas erradamente na Classe 0, correspondem às regiões do Barroso-Padrela, Nordeste, Douro, Centro Litoral, Beira Interior Norte, Pinhal Interior Norte, Beira Interior Sul, Alentejo Central e Alentejo Litoral, todas elas com erros superiores a 40%. As pior classificadas na Classe 1 correspondem a Pinhal Interior Norte, Pinhal Interior Sul, Beira Interior Sul, Ribatejo, Alto Alentejo, Alentejo Central, Alentejo Litoral, Baixo Alentejo e Algarve, com erros superiores a 30%. Assim sendo, Pinhal Interior Norte, Beira Interior Sul, Alentejo Central e Alentejo Litoral são as regiões pior classificadas. Por sua contra as melhor classificadas correspondem a Baixo Minho, AMP e Dão-Lafões tendo erros inferiores a 35% para a Classe 0, e erros inferiores a 25% para a Classe 1.

Observando o mapa gerado a partir das classes simplificadas do CORINE, a agricultura é a classe dominante nas regiões de Nordeste, Douro, Beira Interior Sul, Oeste, Ribatejo e na maior parte do Alentejo com a exceção do Alentejo Litoral em que dominam as florestas junto com o Centro Litoral, Pinhal Interior Norte e Pinhal Interior Sul. Como seria de esperar, nas

grandes áreas metropolitanas de Lisboa e Porto domina a classe artificial. Os matos destacam-seç no Alto Minho e na Beira Interior Norte. Contudo, esta relação nem sempre é a utilizada para proceder com a divisão primária das árvores de classificação, pois o algoritmo escolhe de forma recursiva e quase inexpugnável a variável referente aos matos para a realização de cada uma das árvores ótimas.

Como se observa no Quadro 7, na maioria das regiões coincidem a variável que produz maior homogeneidade (concorrente) e a variável mais importante, esta será a variável usada para iniciar a construção das árvores. Acontece nas regiões de Alto Minho, Baixo Minho, Barroso-Padrela, Nordeste, AMP, Tâmega, Douro, Dão-Lafões, Beira Interior Norte, Beira Interior Sul, Oeste, Alto Alentejo e Alentejo Litoral, em que a árvore de classificação tem início com a variável CLCC4; no Algarve a árvore é iniciada com CLCC3 e no Pinhal Interior Sul a classificação começa com CLCC1.

Existem regiões como Centro Litoral, Pinhal Interior Norte, Ribatejo e AML em que a variável com maior importância são as florestas mas a divisão da árvore começa com os matos por ser a variável que produz maior homogeneidade nos nós. No Baixo Alentejo o algoritmo inicia a classificação com a agricultura (variável mais competitiva), embora a maior importância é dada pelas florestas. No Alentejo Central a variável com maior homogeneidade é CLCC3 e com ela é iniciada a divisão da árvore (Quadro 7).

A presença de fronteira indica que uma parte da célula foi afetada pelo fogo e outra não, ou seja, dentro da célula existem critérios de paragem do fogo (objetivo da análise dos resultados obtidos). Recorrendo à informação disponibilizada nas próprias árvores para cada região PROF (Anexo 6), é possível compreender as variáveis responsáveis pelo fenómeno em estudo e a interação entre elas.

Muitas regiões apresentam nós finais pertencentes à presença de **matos** implicados na paragem dos fogos, como é o caso de Dão-Lafões, Nordeste, AMP, Centro Litoral, Beira Interior Norte, Pinhal Interior Norte, Beira Interior Sul, Oeste, Ribatejo, Alto Alentejo, AML, Alentejo Central e Baixo Alentejo. As **estradas** estabelecem paragem de fogos no Centro Litoral, Pinhal Interior Norte, Pinhal Interior Sul, Beira Interior Sul, Ribatejo e Alto Alentejo. Outros nós terminais apresentam fronteira em **áreas artificiais** no caso das regiões do Centro Litoral, Pinhal Interior Sul, Beira Interior Sul, Ribatejo e Alto Alentejo. As fronteiras dentro das **florestas** estão presentes no Douro, Oeste, AML, Baixo Alentejo e Algarve. As linhas de **cumeada** são representativas na Beira Interior Sul, Alto Alentejo, Alentejo Central, Alentejo Litoral e Algarve. A **água** é representativa nas regiões do Pinhal Interior Sul, Beira Interior Sul, Ribatejo e Alto Alentejo. A **agricultura** apenas aparece nas regiões do Centro

Litoral, Pinhal Interior Sul e Ribatejo, onde se assentam fronteiras do fogo. Os **rios** são relevantes só na região da Beira Interior Sul.

Os nós finais ligados á paragem do fogo nas **zonas de interface** estão representados em poucas regiões e não têm feito uma caraterização muito eficiente, pois não se manifesta grande dependência das fronteiras nos resultados obtidos. Um dos maiores fatores de risco dos incêndios florestais é a interface urbano-rural (artificial-florestas, artificial-agricultura ou artificial-matos) devido à influência das atividades humanas, pelo mesmo motivo são também a principal área em que se disponibilizam os meios de combate ao fogo e de proteção habitacional-pessoal (Vélez, 2006; Nunes et al., 2005). A paragem do fogo nesta área apenas foi identificada no Pinhal Interior Norte em que existe fronteira de fogo na interface artificial-floresta. Outras interfaces relacionadas com as fronteiras estão presentes no Tâmega (fronteira de fogo na interface agricultura-matos), no Pinhal Interior Sul (fronteira na interface floresta-água), no Alto Alentejo (fronteira na interface agricultura-água) e no Algarve (fronteiras em zonas de agricultura-mato e mato-água). Nos resultados deste estudo tem-se confirmado que as fronteiras do fogo nas diferentes interfaces estão correlacionadas com a paragem do fogo em áreas com ocupação de uma das classes da combinação. Por exemplo, a paragem na interface artificial-floresta da-se quando também existe paragem do fogo em zonas florestadas, pois o fogo vem da floresta e é combatido na periferia das povoações; a paragem em agricultura-mato é produzida quando o fogo proveniente das zonas de mato para em áreas agrícolas com pouco combustível; a interface mato-água ocorre quando os fogos de matagais terminam junto a uma superfície de água.

As regiões que apresentam maior número de fronteiras de fogo são Centro Litoral, Pinhal Interior Sul, Beira Interior Sul, Ribatejo, Alto Alentejo e Algarve, sendo que as suas fronteiras estão assentes em diversos tipos da ocupação do solo. Por outro lado, as regiões de Nordeste, AMP, Tâmega, Douro, Dão-Lafões, Beira Interior Norte e Alentejo Litoral apenas revelam existência de fronteira do fogo em um dos tipos de ocupação.

Em cada divisão da árvore existe um nó terminal que é mais puro do que os outros, sendo o seu erro de classificação menor, do mesmo modo existem nós totalmente puros que correspondem a observações classificadas em uma única categoria. Por exemplo, os Nós Terminais 2 e 7 da região de Baixo Minho são totalmente puros, o mesmo acontece em três nós finais do Alentejo Litoral; embora estes exemplos representem a maior pureza também têm um erro de valor zero, correspondente à não existência fronteira de fogo, o qual não é representativo do ponto de vista do estudo, pois para os nós que representam fronteira (objetivo principal desta análise) não se registou nenhum nó totalmente homogéneo.

Tem-se verificado que os matos representam a variável com maior influência na existência das fronteiras do fogo, sendo esta relação, entre a presença de matos e de fronteiras de áreas ardidas, espectável, pois os terrenos ocupados por matos podem ser o produto de fogos reincidentes que frequentemente afetam a mesma área dando lugar a formações vegetais dominadas por matagal. A abundância de áreas de matos deriva do abandono rural, mas não só, pois zonas rochosas ou com declive pronunciado não são aptas para outras formações vegetais, e ainda podem ter origem na frequência de incêndios. Assim, pode-se afirmar que os matos, mesmo não atuando diretamente como barreira na propagação do fogo, revelam o fim de grande parte dos fogos quando acaba a área de matagal; e à escala da célula usada, encontra-se mato no perímetro do fogo.

Estudos anteriores (e.g., Nunes et al., 2005) revelam que os fogos são seletivos à hora de ter preferência por um determinado tipo de coberto vegetal. Os matagais são as áreas preferidas pelo fogo seguidos pelo coberto florestal, por outra parte os fogos raramente ocorrem em zonas artificializadas. No mesmo estudo foi confirmado que em condições meteorológicas severas (seca, vento forte) os fogos que se registam são de grande tamanho e portanto menos seletivos no que diz respeito ao tipo de coberto vegetal. A pesar de tudo, no presente trabalho não se analisou a relação entre as variáveis associadas à fronteira do fogo e o seu tamanho, mas é um tema importante a explorar.

O crescimento de florestações (especialmente coníferas) para produção de madeira, pasta para papel ou conservação de bacias hidrográficas, têm aumentado a quantidade de combustível inflamável (Pausas, 2012). Estas espécies são mais susceptíveis aos fogos do que as folhosas, as quais podem reduzir o risco de fogo quer em povoamentos puros quer em povoamentos mistos pelas suas características morfológicas, anatómicas e/ou fisiológicas. Deste modo, grande percentagem dos incêndios acontecidos no continente português correspondem a povoamentos dominados por pinheiros (AFN, 2011).

Condições atmosféricas adversas, abandono rural e carga de combustível facilitam não apenas a ignição dos incêndios florestais, mas também a sua propagação e comportamento que podem dificultar os trabalhos de extinção. Neste sentido, os perímetros do fogo também dependem das condições meteorológicas, uma possibilidade a explorar em estudos posteriores.

6. GLOSSÁRIO

Algoritmo: sequência de instruções necessárias para a resolução de um problema.

Área florestal: terreno ocupado com floresta, matos e pastagens ou outras formações vegetais espontâneas, segundo os critérios definidos no Inventário Florestal Nacional.

Biomassa: fração biodegradável dos produtos, desperdícios ou resíduos de atividades agrícolas e florestais.

Biodiversidade ou Diversidade Biológica: riqueza e variedade de formas de vida, constituída por espécies e/ou populações de animais, vegetais e microorganismos dum determinado nível de observação.

Causas estruturais: causas que não provocam a ignição mas influenciam o comportamento e a propagação do fogo.

Causas imediatas: causas que provocam o início do fogo.

Coluna de convecção: transferência de uma massa de ar quente das zonas baixas às altas que deseca a vegetação existente encosta acima e transporta fagulhas incandescentes podendo gerar novos focos na parte alta da encosta.

Combustão: reação exotérmica de uma substância combustível com um comburente, susceptível de ser acompanhada de uma emissão de chama e/ou de incandescência e/ou emissão de fumo.

Contra-fogo: técnica que consiste em queimar vegetação, contra o vento, no local para onde se dirige o incêndio, destinando-se a diminuir a sua intensidade e facilitando o seu domínio e extinção.

Corta-fogo: faixa do terreno livre de vegetação até o solo mineral, usado em reflorestações; pode ser perimetral ou interior e os extremos têm de estar apoiados em lugares seguros (estradas, trilhos, linhas de água, outro corta-fogo). A largura varia entre 15 e 40 metros, dependendo das espécies instaladas na floresta.

Desbaste: corta e extração das árvores sobrantes da massa para manter a espessura próxima da normal; realiza-se nos estados de bastio e fustadio, quando as árvores têm um diâmetro entre 10 e 20 centímetros no primeiro estado e maior a 20 cm no segundo.

Efeito chaminé: fenómeno que consiste num desenvolvimento rápido da frente de chamas acompanhado por correntes de ar de modo que a velocidade de propagação do fogo

aumenta de forma súbita. Esta “explosão” acontece geralmente em vales ou em declives muito aguçados e é a causa de muitos acidentes fatais nos incêndios florestais.

Espécie alóctona: espécie vegetal que de forma acidental ou intencional se encontram fora do seu âmbito natural.

Espécie invasora: espécie susceptível de, por si própria, ocupar o território de uma forma excessiva, em área ou em número de indivíduos, assumindo o carácter de praga ou provocando uma modificação significativa nos ecossistemas.

Fogo controlado: uso do fogo na gestão de espaços florestais, sob condições, normas e procedimentos conducentes à satisfação de objetivos específicos e quantificáveis e que é executada sob responsabilidade de técnico credenciado.

Gestão florestal: conjunto de atividades técnicas e material relativo à conservação, organização, melhora e aproveitamento da floresta.

Incêndio florestal: qualquer fogo que decorre em espaços florestais, não planeado e não controlado e que, independentemente da fonte de ignição, requer ações de supressão.

A Comissão Europeia define Incêndio ou Fogo Florestal como um fogo que se deflagra e se estende por florestas e outros terrenos com coberto arbóreo ou que tem início noutros terrenos e que se propaga na floresta e outros terrenos arborizados. Esta definição exclui queimadas ou fogos controlados, geralmente destinados a diminuir ou eliminar o material combustível acumulado no terreno.

Incidência de fogos: medida que representa o número de fogos que se produzem numa área concreta ao longo dum período de tempo determinado.

Interface rural: área ou zona onde estruturas ou outros meios humanos se misturam com combustíveis vegetais e florestais.

Limpeza: eliminação de material vegetal estranho à massa principal, tendo em conta a descontinuidade vertical e horizontal da carga combustível e a gestão da biodiversidade.

Mata: floresta plantada e/ou trabalhada pelo homem, usualmente destinada à exploração.

Mato: terreno inculto dominado por espécies arbustivas e em alguns casos com árvores esparsas.

Ordenamento Forestal: conjunto de normas que regulam as intervenções nos espaços florestais com vista a garantir, de forma sustentada, o fluxo regular de bens e serviços por eles proporcionados.

Pirófito: vegetal adaptado a resistir o fogo e que é favorecido pelos incêndios, geralmente são plantas com forte geminação após fogo e com elevada inflamabilidade.

Plano Regional de Ordenamento Florestal (PROF): instrumento político setorial que estabelece normas específicas de intervenção sobre a ocupação e utilização dos espaços florestais, de modo a promover e garantir a produção sustentada do conjunto de bens e serviços a eles associados, na salvaguarda dos objetivos da política florestal nacional. Criados pelo D-Lnº 204/99 de 9 de Junho.

Silvicultura: conjunto de técnicas que tratam da conservação, melhora, aproveitamento e regeneração ou também da restauração das massas florestais.

Silvicultura Preventiva: conjunto de medidas aplicadas aos povoamentos florestais com o objetivo de dificultar a progressão do fogo e diminuir a sua intensidade, limitando os danos causados no arvoredo. Pretende-se garantir que os povoamentos possuam a máxima resistência à passagem do fogo e reduzir a dependência das forças de combate para a sua proteção.

Fontes: Relatório final de PNDFCI – Plano Nacional de Defesa da Floresta Contra Incêndios Florestais; ANPC – Autoridade Nacional de Proteção Civil; CE – Comissão Europeia.

7. REFERÊNCIAS BIBLIOGRÁFICAS

Abreu, M. C. (2012). *A Actual Orgânica do Ministério da Agricultura, do Mar, do Ambiente e do Ordenamento do Território e a Política da Paisagem*. Revista de Arquitectura Paisagista, Janeiro-Maio, 2012: 3-5. Lisboa.

Autoridade Florestal Nacional, 2011 [em linha]. *Inventário Florestal Nacional*. Disponível na Internet, em: <http://www.afn.min-gricultura.pt/portal/inf> Consulta em agosto de 2011. (Atual “Instituto da Conservação da Natureza e da Floresta”).

Autoridade Florestal Nacional, 2011 [em linha]. *Estatísticas. Dados sobre Incêndios Florestais*. Disponível na Internet, em: <http://www.afn.min-gricultura.pt/portal/dudf/estatisticas> Consulta em dezembro de 2011.

Autoridade Nacional de Proteção Civil – ANPC. Disponível em: <http://www.prociv.pt/GLOSSARIO/Pages/default.aspx> Consulta em setembro de 2012.

Baptista, F. O. (2003). *Um Rural sem Território*. Em: Portela, J. e Caldas, J. C. (2003). *Portugal Chão*. CELTA: Oeiras.

Baptista, F. O. e Santos, R. T. (2005). *Os Proprietários Florestais. Resultados de Um Inquérito*. CELTA: Oeiras.

Borges, J. G. (2001). *Ordenamento Florestal (Conclusões)*. Conferência Nacional sobre Prevenção e Investigação de Incêndios Florestais. Évora, 20 Março de 2001.

Borges, J. G. e Uva, J. S. (2006). *A Prevenção e o Combate a Incêndios Florestais no Âmbito do Ordenamento e da Gestão Florestal*. Em: Pereira, J. S., Pereira, J. M. C., Rego, F. C., Silva, J. M. N. e Silva, T. P. (2006). *Incêndios Florestais em Portugal. Caracterização, Impactes e Prevenção* (2006). ISAPress: Lisboa.

Breiman, L., Friedman, J. H., Olshen, R. A. e Stone, C. J. (1984). *Cassification and Regression Trees*. Wadsworth Int. Group: California.

Catry, F. X., Silva, J. S. e Fernandes, P. (2010). *Efeitos do Fogo na Vegetação*. Em: Moreira, F., Catry, F. X., Silva, J. S. e Rego, F. (2010). *Ecologia do Fogo e Gestão das Áreas Áridas*. ISAPress: Lisboa.

CELPA (2010) [em linha]. *Boletim Estatístico*. Associação da Indústria Papeleira. Disponível na internet, em: <http://www.celpa.pt/> . Consulta em dezembro de 2011.

Chipman, H. A., George, E. I. e McCulloch, R. E. (1998). *Bayesian CART Model Search*. Journal of American Statistical Association, Vol. 93, Nº 443: 935-960. September 1998.

Chuvieco, E., Martín, M. P., Martínez, J. e Salas. J. (1998). *Geografía e Incendios Forestales*. Serie Geográfica, Vol. 7, 1998: 11-17.

Comissão Europeia – CE (2003). *Regulamento Nº 2152/2003 do Parlamento Europeu e do Conselho de 17 de Novembro de 2003. Relativo ao Acompanhamento das Florestas e das Interações Ambientais na Comunidade (Forest Focus)*.

Costa, J.C., Aguiar, C., Capelo, J.H., Lousã, M. e Neto, C. (1998). *Biogeografia de Portugal Continental*. Revista Quercetea, Vol. 0, Bragança. Dezembro 1998.

Daveau, S., Ribeiro, O. e Lautensach, H. (1988). *Geografia de Portugal. O Ritmo Climático e a Paisagem*. Vol. 2. João Sá da Costa: Lisboa.

Decreto Lei nº 55/2007 de 12 de Março, que altera o DL nº 34/99 de 5 de Fevereiro.

Decreto-Lei nº 124/2006 de 28 de Junho.

Duarte, J. F. B. (s.d.). *Os Fogos Florestais em Portugal: o Planeamento do Espaço na Interface urbano-florestal e a Segurança das Populações*. FLUC, Coimbra. Consulta em 2011.

http://www.empreender.aip.pt/irj/go/km/docs/site-manager/www_empreender_aip_pt/conteudos/pt/centrodocumentacao/Centro%20de%20Documenta%C3%A7%C3%A3o/Sector%20Florestal/Fogos%20Florestais%20Portugal_Jorge%20Baptista%20Duarte.pdf

EP, 2009. Estradas de Portugal (<http://www.estradasdeportugal.pt/>)

ESRI, 2011 [em linha]. *Environmental Systems Research Institute*. Disponível na internet, em: <http://www.esri.com/>. Consulta em setembro de 2011.

Fabião, A. M. D. e Oliveira, A. C. (2007). *A Floresta em Portugal. Apontamentos de apoio às aulas das disciplinas de Silvicultura e Gestão da Vegetação*. ISA-UTL, Lisboa.

Ferreira, A. D., Coelho, C., Silva, J. S. e Esteves, T. (2010). *Efeitos do Fogo no Solo e no Regime Hidrológico*. Em: Moreira, F., Catry, F. X., Silva, J. S. e Rego, F. (2010). *Ecologia do Fogo e Gestão das Áreas Ardidas*. ISAPress: Lisboa.

Fortin, M. J. e Dale, M. R. T (2005). *Spatial Analysis. A Guide for Ecologists*. Cambridge University Press, New York.

FORESTIS (2009) [em linha]. *O Associativismo Florestal: seu contributo para o fomento, a gestão e a defesa dos recursos florestais*. Disponível na internet, em: <http://naturalink.sapo.pt/Natureza-e-Ambiente/Agricultura-e-Floresta/content/O-Associativismo-Florestal-seu-contributo-para-o-fomento-a-gestao-e-a-defesa-dos-recursos-florestais?bl=1>. Consulta em outubro de 2011.

Hand, D. J. (1997). *Construction and Assessment of Classification Rules*. John Wiley and Sons: Chichester, England.

IGP, 2009. Mapa de Estradas (<http://www.igeo.pt/>).

IGP, 2011 [em linha]. *Os Homens e o Meio: Um Envelhecimento Progressivo*. Atlas de Portugal. Instituto Geográfico Português. Disponível na internet, em: http://www.igeo.pt/atlas/Cap2/Cap2b_4.html. Consulta em outubro de 2011.

IM, 2012 [em linha]. *Clima de Portugal Continental*. Instituto de Meteorologia. Disponível na internet: <http://www.meteo.pt/pt/areaeducativa/otempo.eoclima/clima.pt/index.html>. Consulta em janeiro de 2012. (Atual “Instituto do Mar e da Atmosfera”).

INAG, 2012 [em linha]. *Sistema Nacional de Informação de Recursos Hídricos. Atlas da Água*. Instituto da Água. Disponível na internet: <http://www.inag.pt/>. Consulta em março de 2012.

INE, 2010. *Inquérito de Emprego*. Instituto Nacional de Estatística. Disponível em : http://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_destaques&DESTAQUESdest_boui=83328022&DESTAQUEStema=55466&DESTAQUESmodo=2. Consulta em Novembro de 2011.

Juárez, O. e Castells, E. (2010). *Modelos de Árbol de Regresión Bayesiano: Un Estudio de Caso*. Revista Investigación Operacional, Vol. 31, nº2: 109-125. México.

Linn, R., Winterkamp, J., Edminster, C., Colman, J. J. e Smith, W. S. (2007). *Coupled Influences of Topography and Wind on Wildland Fire Behaviour*. International Journal of Wildland Fire 16, 183–195.

Lombardero, B. (2005). *Apontamentos de apoio às aulas de Defesa Contra Incendios Florestais*. EPS-USC, Galiza.

Lourenço, L. (1988). *Representação Geográfica dos Incêndios Florestais Ocorridos em Portugal Continental*. Em: Lourenço, L., Golçalves, A. V. e Soares, H. (1988). *Jornadas Científicas sobre Incêndios Florestais*. Coimbra, 23-25 de Novembro de 1988.

MAGRAMA, 2011. Ministerio de Agricultura, Alimentación y Medio Ambiente. Gobierno de España. Consulta em maio de 2012.

Maindonald, J. e Braun, W. J., 2006. *Data Analysis and Graphics Using R. An Example-Based Approach*. Second Edition. Cambridge University Press.

MAMAOT, 2011. Ministério da Agricultura, do Mar, do Ambiente e do Ordenamento do Território (ex-MADRP – Ministério de Agricultura, do Desenvolvimento Rural e das Pescas). Governo de Portugal. Consulta em fevereiro de 2012.

Miranda, P. M. A., Valente, M. A., Tomé, A. R., Trigo, R., Coelho, M. F., Aguiar, A. e Azevedo, E. B. (2006). *O Clima de Portugal nos Séculos XX e XXI*. Em: Santos, F. D. e Miranda, P. M. A. (2006). *Alterações Climáticas em Portugal. Cenário, Impactos e Medidas de Adaptação*. Projecto SIAM II. Gradiva: Lisboa.

Morgado, R e Moreira, F. (2010). *Efeitos do Fogo na Fauna*. Em: Moreira, F., Catry, F. X., Silva, J. S. e Rego, F. (2010). *Ecologia do Fogo e Gestão das Áreas Ardidadas*. ISAPress: Lisboa.

Narayanaraj, G. e Wimberly, M. C. (2011). *Influences of Forest Roads on the Spatial Pattern of Wildfire Boundaries*. International Journal of Wildland Fire 20 (6): 792-803.

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

OEA, (1991). Organización de los Estados Americanos. *Proyecto de Manejo y Conservación de los Recursos Naturales Renovables de la Cuenca del Río Chixoy*. Departamento de Desarrollo Regional y Medio Ambiente, República e Guatemala.

Painho, M. e Caetano, M. (2006). *Cartografia de Ocupação do Solo. Portugal Continental 1985-2000. CORINE Land Cover 2000*. Instituto do Ambiente. Amadora, 2006.

Pausas, J. G. (2012). *Incendios Forestales*. CSIC – Ministerio de Economía y Competitividad. Catarata, Madrid.

Pausas, J. G. e Vallejo, R. M. (1999). *The Role of Fire in European Mediterranean Ecosystems*. Em: Chuvieco, E. (1999). *Remote Sensing of Large Wildfires in the European Mediterranean basin*. Springer –Verlag, 1999: 3-16.

Pereira, J. M. C. (2008). *Apontamentos de apoio às Aulas de Protecção Florestal*. ISA-UTL, Lisboa.

Pereira, J. M. C., Carreiras, J. M. C., Silva, J. M. N. e Vasconcelos, M. J. (2006). *Alguns Conceitos Básicos sobre os Fogos Rurais em Portugal*. Em: Pereira, J. S., Pereira, J. M. C., Rego, F. C., Silva, J. M. N. e Silva, T. P. (2006). *Incêndios Florestais em Portugal. Caracterização, Impactes e Prevenção* (2006). ISAPress: Lisboa.

Pereira, J. S., Correia, A. V., Correia, A. P., Branco, M., Bugalho, M., Caldeira, M. C., Souto Cruz, C., Freitas, H., Oliveira, A. C., Pereira, J. M. C., Reis, R. M. e Vasconcelos, M. J. (2002). *Forests and Biodiversity*. Em: F.D. Santos, F. D., Forbes, K. e Moita, R. *Climate Change in Portugal. Scenarios, Impacts, and Adaptation Measures* (2002). SIAM Project. Gradiva: Lisboa.

Pereira, M. G., Trigo, R. M., da Camara, C. C., Pereira, J. M. C. e Leite, S. M. (2005). *Synoptic Patterns Associated with Large Summer Forest Fires in Portugal*. *Agricultural and Forest Meteorology* 129 (2005): 11-25.

PLADIGA, 2011. *Plan de Prevención e Defensa Contra Incendios Forestais de Galicia*. Consellería do Medio Rural e do Mar, Xunta de Galicia. (<http://www.mediosruralemar.xunta.es>) Consulta em maio de 2012.

PNDFCI, 2006 [em linha]. *Plano Nacional de Defesa da Floresta Contra Incêndios Florestais*. Disponível na internet: http://www.isa.utl.pt/pndfci/relatorio_intercalar/files/ . Consulta em setembro de 2011.

Porrero Rodríguez, M. A. (2001). *Incendios Forestales: Investigación de Causas*. Mundi-Prensa: Madrid.

Pyne, S. J. (2006). *Fogo no Jardim: Compreensão do Contexto dos Incêndios em Portugal*. Em: Pereira, J. S., Pereira, J. M. C., Rego, F. C., Silva, J. M. N. e Silva, T. P. (2006). *Incêndios Florestais em Portugal. Caracterização, Impactes e Prevenção* (2006). ISAPress: Lisboa.

Radich, M. C. e Alves, A. A. M. (2000). *Dois Séculos da Floresta em Portugal*. CELPA: Lisboa.

Rebelo, V., Bento, J. e Portela, J. (1999). *Associativismo Florestal na Região Norte - Das Expectativas às Soluções*. Comunicação ao I Encontro Galiza-Portugal de Estudos Rurais, IPB. Bragança, 1999.

Rego, F. C., Rigolot, E., Fernandes, P., Montiel, C. e Silva, J. S. (2010). *Para uma Gestão Integrada do Fogo*. Policy Brief 4, EFI – European Forest Institute.

Ribeiro, O. (1998). *Portugal. O Mediterrâneo e o Atlântico*. 6ª Ed. Sá da Costa: Lisboa.

Rivas-Martínez, S. (1982). *Série de Vegetación de la Región Eurosiberiana de la Península Ibérica*. Lázaro, Vol. 4.

Rodrigues, M. M. T. (2009). *Integração das Variáveis de Natureza Social na Avaliação do Risco de Incêndio Florestal na Região de Trás-os-Montes e Alto Douro*. Tese de Mestrado em Gestão e Conservação da Natureza. Instituto Politécnico de Bragança.

Rosendo, J. e Martins, J. (s. d.). *A Comissão Regional de Reflorestação do Algarve e a Recuperação das Áreas Ardidas e Envolventes*. Portimão. <http://www.esac.pt/cernas/cfn5/docs/T1-18.pdf>

Ruiz, E. M. (2001). *Manual de Quemas Controladas. El Manejo del Fuego en la Prevención de Incendios Forestales*. Mundi-Prensa: Madrid.

Ruiz del Castillo, J. (2003). *El Fuego, Factor Ecológico*. Em: Vélez, R. (2003). *La Defensa Contra Incendios Forestales. Fundamentos y Experiencias*. McGraw-Hill: Madrid.

Sá, A. C. L. (1999). *Comparação entre a Regressão Logística e as Árvores de Classificação na Cartografia de Áreas Ardidas com Imagens do Satélite Landsat 5 TM*. Tese de Mestrado em Matemática Aplicada às Ciências Biológicas. Departamento de Matemática, Instituto Superior de Agronomia. Universidade Técnica de Lisboa.

Santos, E., Leite, A., Santos, C., Cornaglia, G. e Pinho, J. (s. d.). *Objectivos Estratégicos para a Floresta até 2045*. Lisboa. <http://www.esac.pt/cernas/cfn5/docs/T6-20.pdf>

Steinberg, D. e Golovnya, M. (s.d). *CART: Users's Guide*. Salford Systems 2002-2007. California.

Timofeev, R. (2004). *Classification and Regression Trees (CART). Theory and Applications*. Master Thesis, 2004. Berlim.

Torres, F. T. P. (2006). *Relações entre Fatores Climáticos e Ocorrências de Incêndios Florestais na Cidade de Juiz de Fora (MG)*. Caminhos de Geografia 7 (18): 162-171. Junho, 2006. UFU-Brasil.

Vélez, R. (1991). *Uso del Fuego en Selvicultura*. Em: X Congreso Forestal Mundial. Paris, 1991.

Vélez, R. (2003). *Los Incendios Forestales en la Cuenca Mediterránea*. Em: Vélez, R. (2003). *La Defensa Contra Incendios Forestales. Fundamentos y Experiencias*. McGraw-Hill: Madrid.

Vélez, R. (2006). *A Defesa Contra Incêndios Florestais. Estratégias, Recursos, Organização*. Em: Pereira, J. S., Pereira, J. M. C., Rego, F. C., Silva, J. M. N. e Silva, T. P. (2006). *Incêndios Florestais em Portugal. Caracterização, Impactes e Prevenção* (2006). ISAPress: Lisboa.

Vélez, R. (2009). *Manual de Formación para la Lucha Contra Incendios*. ICONA, Madrid.

Viega da Cunha, L., Ribeiro, L. Proença de Oliveira, R. e Nascimento, J. (2006). *Recursos Hídricos*. Em: Santos, F. D. e Miranda, P. M. A. (2006). *Alterações Climáticas em Portugal. Cenário, Impactos e Medidas de Adaptação*. Projecto SIAM II. Gradiva: Lisboa.

Viegas, D. X. e Viegas, M. T. (1994). *A Relationship between Rainfall and Burned Area for Portugal*. *International Journal of Wildland Fire* 4 (1): 11-16.

Viger, J. A., Nonell, X. N., Ferrer, E. P., Cuchi, E.P. e López, L. Z. (2004). *Manual de Ingeniería Básica para la Prevención y Extinción de Incendios Forestales*. Mundi-Prensa: Barcelona.

Wei-Yin Loh (2008). *Classification and Regression Tree Methods*. Em: Ruggeri, Kenett e Faltin (2006). *Encyclopedia of Statistics in Quality and Reliability*. Wiley. 2008: 315-323.

Yohannes, Y. e Webb, P. (1999). *Classification and Regression Trees: a Manual for Identifying Indicators of Vulnerability to Famine and Chronic Food Insecurity*. International Food Policy Research Institute, Washington D.C.. Mimeo.

Zelder, P. H e Rego, F. C. (2006). *Regimes de Fogo e Biodiversidade: Respostas dos Ecossistemas a Alternativas de Gestão*. Em: Pereira, J. S., Pereira, J. M. C., Rego, F. C., Silva, J. M. N. e Silva, T. P. (2006). *Incêndios Florestais em Portugal. Caracterização, Impactes e Prevenção* (2006). ISAPress: Lisboa.

Legislação consultada

Comissão Europeia (2003). *Regulamento Nº 2152/2003 do Parlamento Europeu e do Conselho de 17 de Novembro de 2003. Relativo ao Acompanhamento das Florestas e das Interações Ambientais na Comunidade (Forest Focus)*.

Decreto-Lei 204/99 de 9 de Junho, que regula o processo de elaboração, de aprovação, de execução e de alteração dos planos regionais de ordenamento florestal a aplicar nos espaços florestais, nos termos do artigo 5º da Lei nº 33/96 de 17 de Agosto (Lei de Base da Política Florestal).

Decreto-Lei nº 565/99 de 21 de Dezembro, que regula a introdução na Natureza de espécies não indígenas da flora e da fauna, com a excepção das consideradas invasoras.

Decreto-Lei nº 63/2004 de 22 de Março, pelo que se cria um fundo financeiro de carácter permanente destinado a apoiar a gestão florestal sustentável.

Decreto-Lei nº 156/2004 de 30 de Junho, pelo que se define um novo quadro orientador das medidas e acções a desenvolver no âmbito do Sistema Nacional de Prevenção e Protecção da Floresta contra Incêndios.

Decreto-Lei nº 124/2006 de 28 de Junho, que estabelece as medidas e acções a desenvolver no âmbito do Sistema Nacional de Defesa contra Incêndios Florestais.

Decreto-Lei nº 55/2007 de 12 de Março, que altera o DL nº 34/99 de 5 de Fevereiro. Estabelece no seu Artigo 1º, a proibição, pelo prazo de 10 anos, de várias acções nos terrenos com povoamentos florestais percorridos por incêndios, em áreas não classificadas nos planos municipais de ordenamento do território como solos urbanos.

Lei nº 33/96 de 17 de Agosto. Lei de Bases da Política Florestal.

Portaria nº 1061/2004 de 21 de Agosto, pelo que se estabelece o Regulamento do Fogo Controlado e se definem os requisitos dos técnicos especializados nesta técnica, de acordo com o disposto no DL 156/2004 de 30 de Junho.

8. ANEXOS

ANEXO 1.

Número de Ocorrências e Área Ardida para os Distritos de Portugal Continental, no Período de 1996 a 2005 (AFN, 2011).

Distrito	Nº Ocorrências	Área Ardida (ha)		
		Povoamentos	Matos	Total
Viana do Castelo	16 713	43 734	40 993	84 727
Braga	39 377	24 420	37 307	61 727
Porto	62 967	40 068	39 912	79 980
Vila Real	21 357	64 668	95 895	160 564
Bragança	10 042	25 082	74 898	99 979
Aveiro	19 804	30 714	6 704	37 418
Viseu	25 324	49 248	104 919	154 167
Guarda	14 657	62 618	172 204	234 823
Coimbra	6 150	56 030	23 204	79 234
Castelo Branco	6 452	156999	39 081	196 081
Leiria	8 860	35 493	13 303	48 796
Santarém	10 906	102 815	24 934	127 749
Lisboa	27 903	9 663	15 231	24 894
Portalegre	1 053	52 539	25 003	77 542
Évora	623	19 255	2 448	21 703
Setúbal	8 242	14 475	2 582	17 058
Beja	996	23 605	12 304	35 909
Faro	2 690	54 704	42 902	97 606
TOTAL	284 116	866 129	773 826	1 693 955

Número de Hectares Afetados pelos Incêndios entre 1996 e 2005 (AFN, 2011).

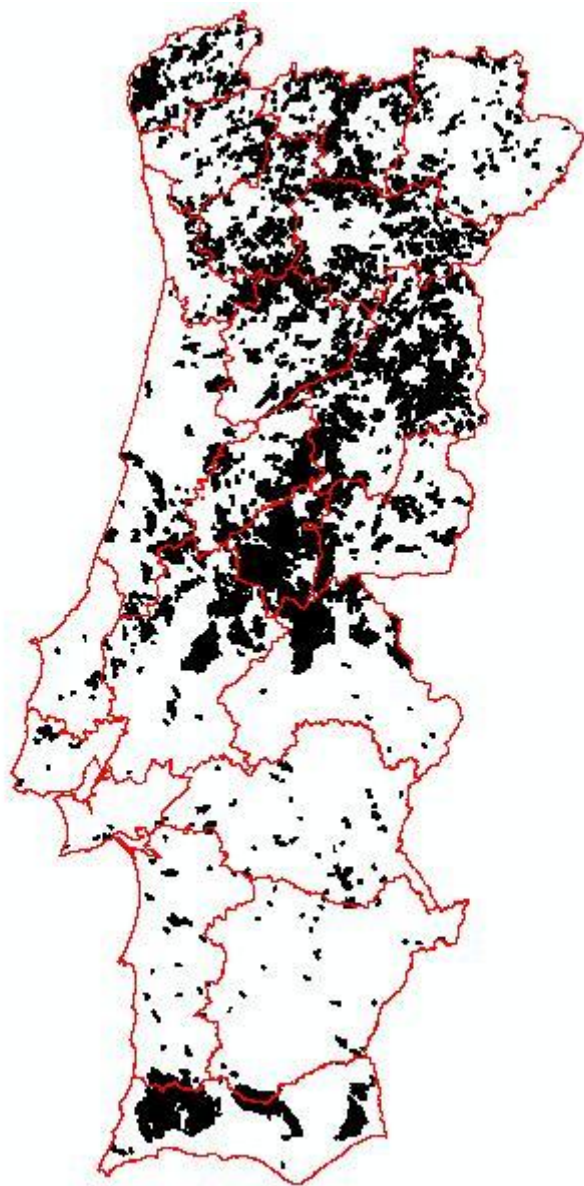
Ano	Área Ardida (ha)		
	Povoamentos	Matos	Área Total
1996	30 542	58 325	88 867
1997	11 466	19 068	30 535
1998	57 393	100 975	158 369
1999	31 052	39 561	70 613
2000	68 646	90 958	159 605
2001	45 617	66 695	112 312
2002	65 164	59 455	124 619
2003	286 055	139 784	425 839
2004	56 271	73 836	130 108
2005	213 921	125 168	339 089
TOTAL	866 129	773 826	1 639 955

ANEXO 2. Identificação das Regiões PROF para Portugal Continental.

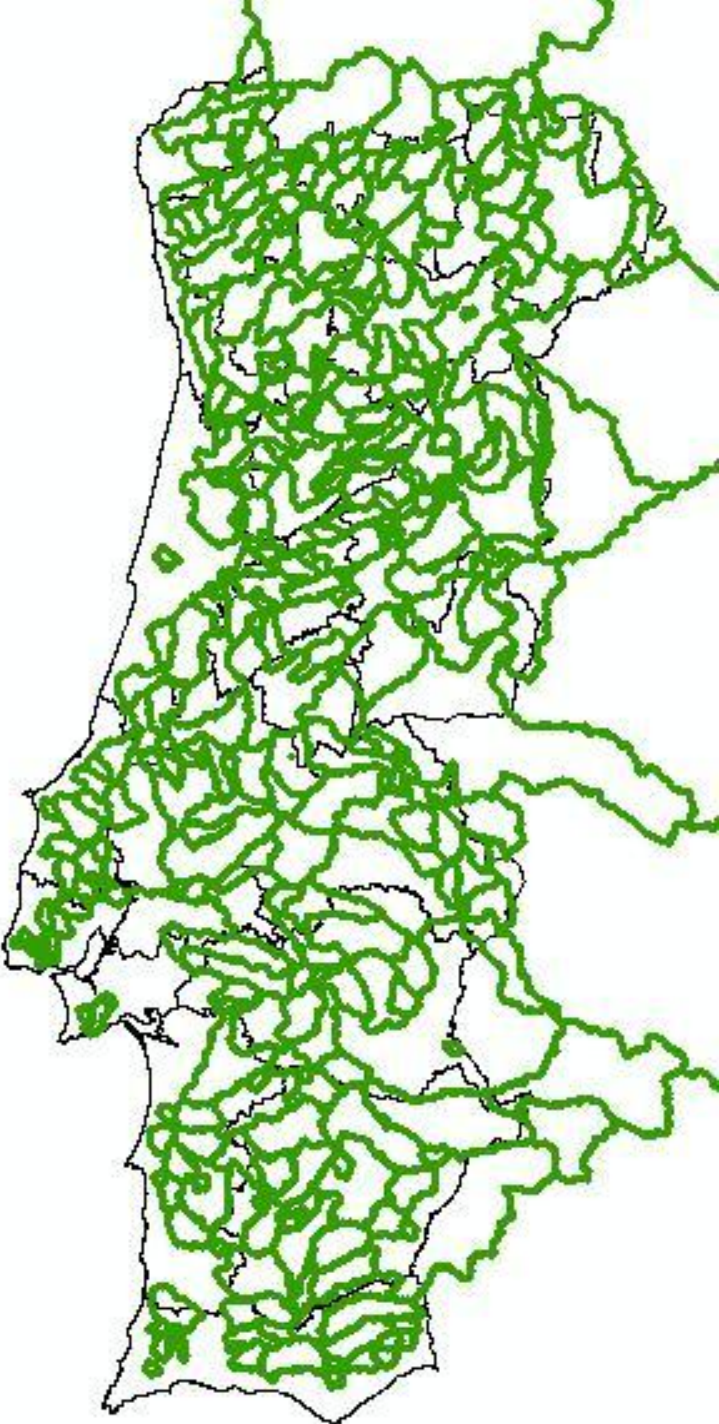


- 1- Alto Minho
- 2- Baixo Minho
- 3- Barroso e Padrela
- 4- Nordeste
- 5- Área Metropolitana de Porto e Entre Douro e Vouga
- 6- Tâmega
- 7- Douro
- 8- Centro Litoral
- 9- Dão-Lafões
- 10- Beira Interior Norte
- 11- Pinhal Interior Norte
- 12- Pinhal Interior Sul
- 13- Beira Interior Sul
- 14- Oeste
- 15- Ribatejo
- 16- Alto Alentejo
- 17- Área Metropolitana de Lisboa
- 18- Alentejo Central
- 19- Alentejo Litoral
- 20- Baixo Alentejo
- 21- Algarve

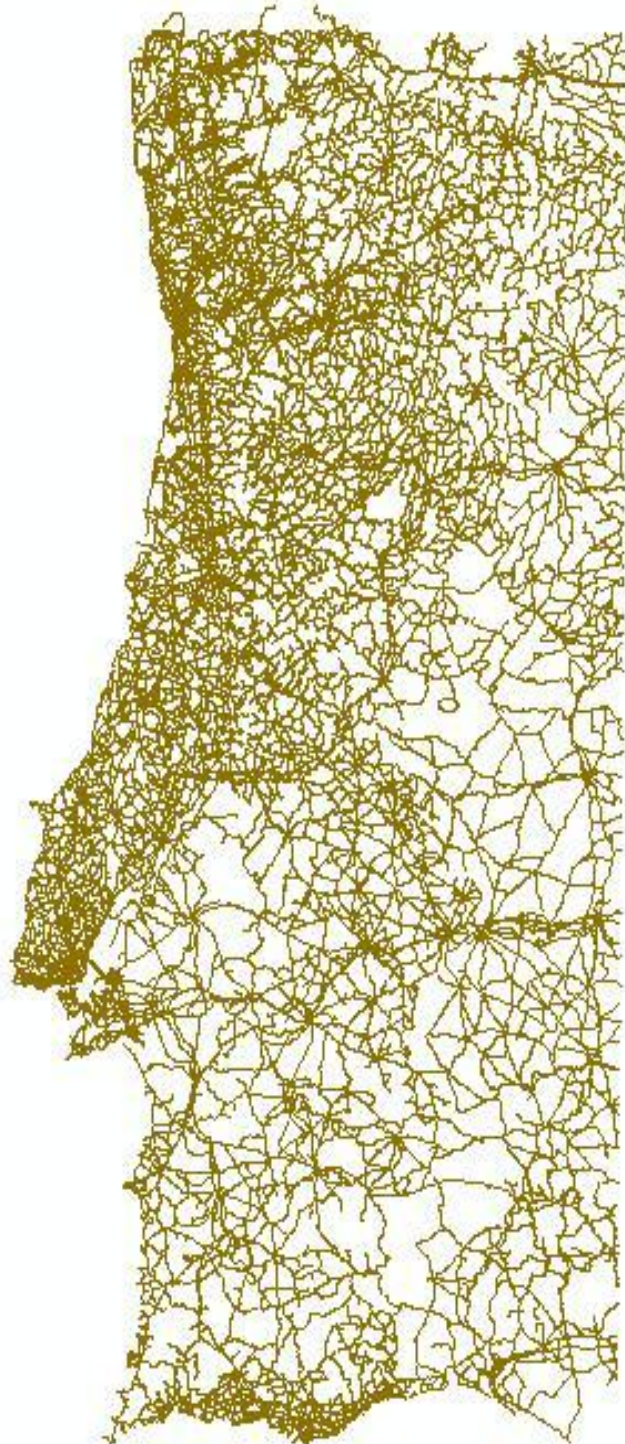
ANEXO 3. Mapa dos Fogos com Área Igual ou Superior aos 100 ha Ardidos entre 1996 e 2005 (LDRAG).



ANEXO 4. Mapa das Linhas e Cumeada para Portugal Continental (INAG, 2012).



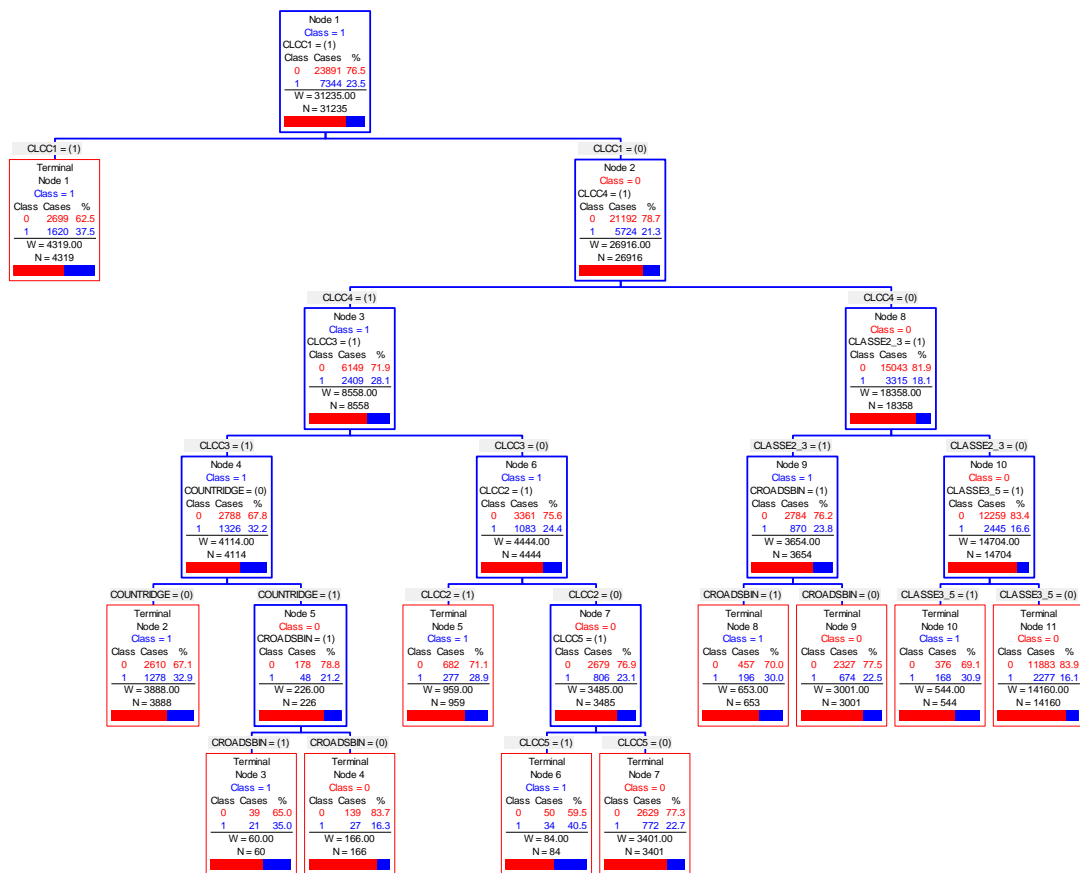
ANEXO 5. Rede Viária de Portugal Continental (EP, 2009; IGP, 2009).



ANEXO 6 . Rede Hidrográfica de Portugal Continental (INAG, 2012).



ANEXO 7. Ávore de Classificação e Output Obtido para o PROF “Pinhal Interior Sul”.



Para facilitar a leitura das árvores é mostrado o nome e código das variáveis usadas.

Nome da Variável	Código da Variável
Fronteira (resposta)	CFRONTBIN
Área Artificial (preditora)	CLCC1
Agricultura (preditora)	CLCC2
Floresta (preditora)	CLCC3
Matos (preditora)	CLCC4
Água (preditora)	CLCC5
Linha de Cumeada (preditora)	COUNTRIDGE
Estradas (preditora)	CROADSBIN
Linha de Água (preditora)	CRIVERBIN
Interface artificial-agricultura (preditora)	CLASSE1-2
Interface artificial-floresta (preditora)	CLASSE1-3
Interface artificial-matos (preditora)	CLASSE1-4
Interface artificial-água (preditora)	CLASSE1-5
Interface agricultura-floresta (preditora)	CLASSE2-3
Interface agricultura-matos (preditora)	CLASSE2-4
Interface agricultura-água (preditora)	CLASSE2-5
Interface floresta-matos (preditora)	CLASSE3-4
Interface floresta-água (preditora)	CLASSE3-5
Interface matos-água (preditora)	CLASSE4-5

>USE "E:\final\pinhal-int-sul.xlsx" ENCODING=SHIFTJIS

Model reset.

VARIABLES IN RECT FILE ARE:

IDGRID	CLCC1	CLCC2	CLCC3
CLCC4	CLCC5	COUNTRIDGE	CROADSBIN
CRIVERBIN	CFRONTBIN	CLASSE6	CLASSE10
CLASSE14	CLASSE22	CLASSE15	CLASSE21
CLASSE33	CLASSE35	CLASSE55	CLASSE77
CLASSE1_2	CLASSE1_3	CLASSE1_5	CLASSE1_4
CLASSE2_3	CLASSE2_4	CLASSE2_5	CLASSE3_4
CLASSE3_5	CLASSE4_5		

E:\final\pinhal-int-sul.xlsx: 31236 records.

Model reset.

>REM ***Setting General options

>LOPTIONS MEANS = NO, PREDICTIONS = NO, TIMING = NO, GAINS = NO, ROC = NO, PLOTS = NO

>FORMAT = 5

>REM***Setting CART options

>LOPTIONS, NOPRINT = NO, PS = NO

>BOPTIONS SURROGATES = 5 PRINT = 5, COMPETITORS = 5 CPRINT = 5, TREELIST = 10,

> BRIEF

>REM ***Setting MARS default options

>BOPTIONS PENALTY = 0.000000, SPEED = 4, INTERACTIONS = 1, MINSPAN = 0, BASIS = 15

>MARS SEED = 987654321

>BOPTIONS OLS = YES

>PRINT = TERSE

>CATEGORY

>AUXILIARY

>MODEL CFRONTBIN

Model reset: CFRONTBIN

>KEEP

>KEEP CLASSE1_2, CLASSE1_3, CLASSE1_4, CLASSE1_5, CLASSE2_3, CLASSE2_4, CLASSE2_5,

> CLASSE3_4, CLASSE3_5, CLASSE4_5, CLCC1, CLCC2, CLCC3, CLCC4, CLCC5, COUNTRIDGE,

> CRIVERBIN, CROADSBIN

>LOPTIONS UNS = NO

>CATEGORY CFRONTBIN, CLASSE1_2, CLASSE1_3, CLASSE1_4, CLASSE1_5, CLASSE2_3,

>CLASSE2_4, CLASSE2_5, CLASSE3_4, CLASSE3_5, CLASSE4_5, CLCC1, CLCC2, CLCC3,

>CLCC4, CLCC5, COUNTRIDGE, CRIVERBIN, CROADSBIN

>ERROR CROSS = 10

>BOPTIONS SERULE = 1, IMPORTANCE = 1

>BOPTIONS SURROGATES = 5 PRINT = 5, COMPETITORS = 5 CPRINT = 5

>METHOD GINI POWER = 0.0000

>BUILD

Salford Predictive Modeler: CART(R) version 6.6.0.091

Records Read: 31235

Records Kept in Learning sample: 31235

Discrete Variable	N Levels in Model	Discrete Variable	N Levels in Model
CLCC1	2	CLASSE1_3	2
CLCC2	2	CLASSE1_5	2
CLCC3	2	CLASSE1_4	2
CLCC4	2	CLASSE2_3	2
CLCC5	2	CLASSE2_4	2
COUNTRIDGE	2	CLASSE2_5	2
CROADSBIN	2	CLASSE3_4	2
CRIVERBIN	2	CLASSE3_5	2
CFRONTBIN	2	CLASSE4_5	2
CLASSE1_2	2		

Target Frequency Table

```

=====
Variable: CFRONTBIN
N Classes: 2
Data Value          N          %          Wgt Count          %
-----
0                   23891   76.49          23891   76.49
1                    7344   23.51          7344   23.51
-----
Total                31235                                31235

```

PRIORS SET EQUAL
 The data are being read ...
 31235 Observations in the learning sample.
 FILE: E:\final\pinhal-int-sul.xlsx

WARNING
 V-fold cross validation is being used on a training sample
 with 31235 records. Using a test sample will speed up the run.
 CART is running

Root Node Competitors In Order of Improvement

```

=====
Competitor          Split          Improvement  N Left N Right
-----
CLCC1                1              0.0104176   4319  26916
CLCC4                1              0.0060615   9240  21995
CLASSE1_3           1              0.0049209    993  30242
CLASSE3_4           1              0.0032951   3858  27377
CLASSE1_4           1              0.0023689    264  30971
CLCC2                1              0.0019383   9026  22209
CROADSBIN           1              0.0019312   4409  26826
CLASSE2_4           1              0.0011964   1209  30026
CLASSE2_3           1              0.0008381   4700  26535
CLASSE3_5           1              0.0005702    609  30626
CLCC5                1              0.0005291    987  30248
CLASSE4_5           1              0.0004797    149  31086
CLASSE1_2            6              0.0004347   1303  29932
COUNTRIDGE          1              0.0000974   1302  29933
CLASSE2_5            0              0.0000747  31086   149
CLCC3                0              0.0000395   7377  23858
CLASSE1_5            0              0.0000360  31224   11
CRIVERBIN           1              0.0000004   4098  27137

```

TREE SEQUENCE

```

=====
Dependent variable: CFRONTBIN

Terminal          Cross-Validated          Resubstitution          Complexity
Tree Nodes        Relative Cost            Relative Cost            Parameter
-----
 1*   135    0.80476 +/- 0.00658          0.78807          0.000000
10    20    0.80780 +/- 0.00658          0.79449          0.000141
11    19    0.80754 +/- 0.00658          0.79489          0.000207
12    17    0.80794 +/- 0.00658          0.79582          0.000243
13    16    0.80606 +/- 0.00658          0.79632          0.000261
14**  11    0.80736 +/- 0.00657          0.79998          0.000375
15     9    0.81363 +/- 0.00659          0.80212          0.000545
16     6    0.81201 +/- 0.00650          0.80704          0.000830
17     3    0.82174 +/- 0.00659          0.82174          0.002460
18     2    0.89238 +/- 0.00525          0.89238          0.035333
19     1    1.00000 +/- 0.00008          1.00000          0.053818

```

Initial misclassification cost = 0.50000
 Initial class assignment = 1

NODE INFORMATION

```

*****
*                               Node 1: CLCC1                               *
*                               N: 31235                                   *
*****
=====
=      Terminal Node 1          =      *      Node 2          *
=      N: 4319                  =      *      N: 26916         *
=====
*****
Node 1 was split on CLCC1
A case goes left if CLCC1 = (1)
Improvement = 0.010418      Complexity Threshold = 0.053808

```

Node	Cases	Wgt Counts	Cost	Class
1	31235	31235.00	0.50000	1
-1	4319	4319.00	0.33868	1
2	26916	26916.00	0.46771	0

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	23891.00	2699.00	21192.00	0.50000	0.33868	0.53229
1	7344.00	1620.00	5724.00	0.50000	0.66132	0.46771

Surrogate	Split	Assoc	Improvement
1 CLASSE1_2 s	0	0.26953	0.00043
2 CLASSE1_3 s	1	0.25410	0.00492
3 CLASSE1_4 s	1	0.07497	0.00237
4 CLASSE1_5 s	1	0.00166	0.00004

Competitor	Split	Improvement
1 CLCC4	1	0.00606
2 CLASSE1_3	1	0.00492
3 CLASSE3_4	1	0.00330
4 CLASSE1_4	1	0.00237
5 CLCC2	1	0.00194

```

*****
*                               Node 2: CLCC4                               *
*                               N: 26916                                   *
*****
*****
*      Node 3                      *      *      Node 8                      *
*      N: 8558                      *      *      N: 18358                    *
*****
Node 2 was split on CLCC4
A case goes left if CLCC4 = (1)
Improvement = 0.007743      Complexity Threshold = 0.035323

```

Node	Cases	Wgt Counts	Cost	Class
2	26916	26916.00	0.46771	0
3	8558	8558.00	0.43966	1
8	18358	18358.00	0.41755	0

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	21192.00	6149.00	15043.00	0.53229	0.43966	0.58245
1	5724.00	2409.00	3315.00	0.46771	0.56034	0.41755

Surrogate	Split	Assoc	Improvement
1 CLASSE3_4 s	0	0.45909	0.00543
2 CLCC3	1	0.37890	0.00002
3 CLASSE2_4 s	0	0.12777	0.00103
4 CLASSE4_5 s	0	0.02027	0.00065

Competitor	Split	Improvement
1 CLASSE3_4	1	0.00543
2 CROADSBN	1	0.00145
3 CLCC5	1	0.00109
4 CLASSE2_4	1	0.00103
5 CLASSE3_5	1	0.00094


```

*****
*                               Node 3: CLCC3                               *
*                               N: 8558                                   *
*****
*                               Node 4                               * *                               Node 6                               *
*                               N: 4114                               * *                               N: 4444                               *
*****

```

Node 3 was split on CLCC3
A case goes left if CLCC3 = (1)
Improvement = 0.001338 Complexity Threshold = 0.000820

Node	Cases	Wgt Counts	Cost	Class
3	8558	8558.00	0.43966	1
4	4114	4114.00	0.39258	1
6	4444	4444.00	0.48822	1

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	6149.00	2788.00	3361.00	0.43966	0.39258	0.48822
1	2409.00	1326.00	1083.00	0.43966	0.39258	0.48822

Surrogate	Split	Assoc	Improvement
1 CLASSE3_4	s 0	0.90108	0.00070
2 CLASSE2_4	s 0	0.19446	0.00004
3 CLCC2	s 1	0.12509	0.00021
4 CLASSE4_5	s 0	0.00861	0.00024
5 CROADSBIN	s 1	0.00278	0.00009

Competitor	Split	Improvement
1 CLASSE3_4	0	0.00070
2 CLCC5	0	0.00049
3 CLASSE4_5	1	0.00024
4 CLASSE3_5	1	0.00021
5 CLCC2	1	0.00021

```

*****
*                               Node 4: COUNTRIDGE                               *
*                               N: 4114                                   *
*****
=====
=          Terminal Node 2          = *          Node 5          *
=          N: 3888                  = *          N: 226          *
=====

```

Node 4 was split on COUNTRIDGE
A case goes left if COUNTRIDGE = (0)
Improvement = 0.000288 Complexity Threshold = 0.000535

Node	Cases	Wgt Counts	Cost	Class
4	4114	4114.00	0.39258	1
-2	3888	3888.00	0.38567	1
5	226	226.00	0.46731	0

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	2788.00	2610.00	178.00	0.39258	0.38567	0.53269
1	1326.00	1278.00	48.00	0.60742	0.61433	0.46731

Competitor	Split	Improvement
1 CLASSE3_4	0	0.00026
2 CLCC5	1	0.00018
3 CLCC2	1	0.00016
4 CLASSE3_5	1	0.00012
5 CRIVERBIN	1	0.00010

```

*****
*                               *
*           Node 5: CROADSBIN   *
*                               *
*           N: 226              *
*                               *
=====
=           Terminal Node 3     = =           Terminal Node 4     =
=           N: 60               = =           N: 166               =
=====

```

Node 5 was split on CROADSBIN

A case goes left if CROADSBIN = (1)

Improvement = 0.000190 Complexity Threshold = 0.000614

Node	Cases	Wgt	Counts	Cost	Class
5	226		226.00	0.46731	0
-3	60		60.00	0.36341	1
-4	166		166.00	0.38722	0

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	178.00	39.00	139.00	0.53269	0.36341	0.61278
1	48.00	21.00	27.00	0.53269	0.36341	0.61278

Surrogate	Split	Assoc	Improvement
1 CLCC2	s 1	0.12361	0.00001
2 CLASSE2_4	s 1	0.10958	0.00004
3 CLASSE3_4	s 1	0.08162	0.00004
4 CLASSE2_3	s 0	0.02100	0.00000

Competitor	Split	Improvement
1 CRIVERBIN	0	0.00007
2 CLASSE3_4	1	0.00004
3 CLASSE2_4	1	0.00004
4 CLCC2	1	0.00001
5 CLCC5	0	0.00001

```

*****
*                               *
*           Node 6: CLCC2       *
*                               *
*           N: 4444             *
*                               *
=====
=           Terminal Node 5     = *           Node 7           *
=           N: 959             = *           N: 3485           *
=====

```

Node 6 was split on CLCC2

A case goes left if CLCC2 = (1)

Improvement = 0.000284 Complexity Threshold = 0.001230

Node	Cases	Wgt	Counts	Cost	Class
6	4444		4444.00	0.48822	1
-5	959		959.00	0.43080	1
7	3485		3485.00	0.49463	0

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	3361.00	682.00	2679.00	0.48822	0.43080	0.50537
1	1083.00	277.00	806.00	0.51178	0.56920	0.49463

Surrogate	Split	Assoc	Improvement
1 CLASSE2_4	s 1	0.98720	0.00025
2 CLASSE2_5	s 1	0.00600	0.00001

Competitor	Split	Improvement
1 CLCC5	1	0.00028
2 CLASSE4_5	1	0.00027
3 CLASSE2_4	0	0.00025
4 CROADSBIN	0	0.00007
5 CRIVERBIN	1	0.00002


```

*****
*                               Node 7: CLCC5                               *
*                               N: 3485                                   *
*****
=====
= Terminal Node 6                = = Terminal Node 7                =
= N: 84                          = = N: 3401                          =
=====

```

Node 7 was split on CLCC5
A case goes left if CLCC5 = (1)
Improvement = 0.000261 Complexity Threshold = 0.001268

Node	Cases	Wgt Counts	Cost	Class
7	3485	3485.00	0.49463	0
-6	84	84.00	0.31132	1
-7	3401	3401.00	0.48856	0

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	2679.00	50.00	2629.00	0.50537	0.31132	0.51144
1	806.00	34.00	772.00	0.49463	0.68868	0.48856

Surrogate	Split	Assoc	Improvement
1 CLASSE4_5	s 0	1.00000	0.00026

Competitor	Split	Improvement
1 CLASSE4_5	0	0.00026
2 COUNTRIDGE	1	0.00001
3 CRIVERBIN	1	0.00000
4 CROADSBIN	1	0.00000

```

*****
*                               Node 8: CLASSE2_3                               *
*                               N: 18358                                   *
*****
*****
*                               Node 9                               *
*                               N: 3654                               *
*****
*                               Node 11                               *
*                               N: 14704                               *
*****

```

Node 8 was split on CLASSE2_3
A case goes left if CLASSE2_3 = (1)
Improvement = 0.002250 Complexity Threshold = 0.002450

Node	Cases	Wgt Counts	Cost	Class
8	18358	18358.00	0.41755	0
9	3654	3654.00	0.49588	1
11	14704	14704.00	0.39351	0

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	15043.00	2784.00	12259.00	0.58245	0.49588	0.60649
1	3315.00	870.00	2445.00	0.41755	0.50412	0.39351

Surrogate	Split	Assoc	Improvement
1 CLCC2	s 1	0.70881	0.00100

Competitor	Split	Improvement
1 CROADSBIN	1	0.00189
2 CLASSE3_5	0	0.00122
3 CLCC2	0	0.00100
4 CLCC3	0	0.00085
5 CLCC5	0	0.00079

```

*****
*                               *
*           Node 9: CROADSBIN   *
*           N: 3654             *
*                               *
=====
= Terminal Node 8               = * Terminal Node 9           *
=           N: 653              = *           N: 3001           *
=====
*****

```

Node 9 was split on CROADSBIN
A case goes left if CROADSBIN = (1)
Improvement = 0.000350 Complexity Threshold = 0.002813

Node	Cases	Wgt Counts	Cost	Class
9	3654	3654.00	0.49588	1
-8	653	653.00	0.41750	1
-9	3001	3001.00	0.48513	0

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	2784.00	457.00	2327.00	0.49588	0.41750	0.51487
1	870.00	196.00	674.00	0.50412	0.58250	0.48513

Competitor	Split	Improvement
1 CRIVERBIN	0	0.00029
2 COUNTRIDGE	0	0.00018
3 CLCC5	1	0.00002

```

*****
*                               *
*           Node 10: CLASSE3_5  *
*           N: 14704            *
*                               *
=====
= Terminal Node 10              = = Terminal Node 11           =
=           N: 544              = =           N: 14160           =
=====
=====

```

No Node 10 was split on CLASSE3_5
A case goes left if CLASSE3_5 = (1)
Improvement = 0.001601 Complexity Threshold = 0.003569

Node	Cases	Wgt Counts	Cost	Class
10	14704	14704.00	0.39351	0
-10	544	544.00	0.40758	1
-11	14160	14160.00	0.38399	0

Class	Weighted Counts			Within Node Probabilities		
	Top	Left	Right	Top	Left	Right
0	12259.00	376.00	11883.00	0.60649	0.40758	0.61601
1	2445.00	168.00	2277.00	0.39351	0.59242	0.38399

Surrogate	Split	Assoc	Improvement
1 CLCC5	s 1	0.74708	0.00115

Competitor	Split	Improvement
1 CROADSBIN	1	0.00124
2 CLCC5	0	0.00115
3 CLCC3	0	0.00053
4 CLCC2	0	0.00024
5 CRIVERBIN	0	0.00012

TERMINAL NODE INFORMATION

```

=====
[Breiman adjusted cost, lambda = 0.00046]
Node      N    Prob   Cost      Class
-----
[Breiman adjusted cost, lambda = 0.00043]
Node      N    Prob   Cost      Class
-----
1         4319 0.1668 0.3387    1
          Parent C.T. =    0.05381
          [ 0.3408]
          2699 0.3387    0
          1620 0.6613    1

2         3888 0.1416 0.3857    1
          Parent C.T. =    0.00054
          [ 0.3881]
          2610 0.3857    0
          1278 0.6143    1

3          60 0.0022 0.3634    1
          Parent C.T. =    0.00061
          [ 0.4937]
          39 0.3634    0
          21 0.6366    1

4         166 0.0047 0.3872    0
          Parent C.T. =    0.00061
          [ 0.4546]
          139 0.6128    0
          27 0.3872    1

5         959 0.0331 0.4308    1
          Parent C.T. =    0.00123
          [ 0.4412]
          682 0.4308    0
          277 0.5692    1

6          84 0.0034 0.3113    1
          Parent C.T. =    0.00127
          [ 0.4033]
          50 0.3113    0
          34 0.6887    1

7        3401 0.1076 0.4886    0
          Parent C.T. =    0.00127
          [ 0.4918]
          2629 0.5114    0
          772 0.4886    1

8         653 0.0229 0.4175    1
          Parent C.T. =    0.00281
          [ 0.4325]
          457 0.4175    0
          196 0.5825    1

9        3001 0.0946 0.4851    0
          Parent C.T. =    0.00281
          [ 0.4888]
          2327 0.5149    0
          674 0.4851    1

10        544 0.0193 0.4076    1
          Parent C.T. =    0.00357
          [ 0.4253]
          376 0.4076    0
          168 0.5924    1

```

11	14160	0.4037	0.3840	0	Parent C.T. =	0.00357
					[0.3849]	
	11883	0.6160		0		
	2277	0.3840		1		

Terminal Node Target Class Counts

Node	Learn	Target	Class Counts
1	4319.00	2699.00	1620.00
2	3888.00	2610.00	1278.00
3	60.00	39.00	21.00
4	166.00	139.00	27.00
5	959.00	682.00	277.00
6	84.00	50.00	34.00
7	3401.00	2629.00	772.00
8	653.00	457.00	196.00
9	3001.00	2327.00	674.00
10	544.00	376.00	168.00
11	14160.00	11883.00	2277.00

MISCLASSIFICATION BY CLASS

(Cross Validation)

Class	Prior Prob	Wgt Count	Count	Wgt Misclass	Misclass	Cost
0	0.50000	23891.00 (23891.00)	23891	6913.00 7688.00	6913 7688	0.28936 0.32179)
1	0.50000	7344.00 (7344.00)	7344	3750.00 3566.00	3750 3566	0.51062 0.48557)
Total	1.00000	31235.00 (31235.00) (31235.00)	31235	10663.00 11254.00 11933.00	10663 11254 11933)	

VARIABLE IMPORTANCE

	Relative Importance	Number Of Categories	Penalty
CLCC1	100.00000	2	
CLCC4	74.32381	2	
CLASSE3_4	59.16022	2	
CLASSE1_3	47.23684	2	
CLASSE1_4	22.73945	2	
CLASSE2_3	21.60614	2	
CLASSE3_5	15.36773	2	
CLCC2	14.41868	2	
CLCC5	13.54298	2	
CLASSE2_4	13.06004	2	
CLCC3	12.99740	2	
CLASSE4_5	10.97200	2	
CROADSBIN	6.00019	2	
CLASSE1_2	4.17288	2	
COUNTRIDGE	2.76572	2	
CLASSE1_5	0.34510	2	
CLASSE2_5	0.11400	2	
CRIVERBIN	0.00000	2	

OPTION SETTINGS

```

=====
Construction Rule          Gini (priors altered by costs)
Estimation Method         10-fold cross-validation
Misclassification Costs   Unit
Tree Selection             1.00000 se rule
Linear Combinations       No

Initial value of the complexity parameter          = 0.00000
Minimum size below which node will not be split  = 10
Minimum size for a child node                     = 1
Node size above which sub-sampling will be used  = 31235
Maximum number of surrogates used for missing values = 5
Number of surrogate splits printed                = 5
Number of competing splits printed                = 5
Maximum number of trees printed in the tree sequence = 10
Max. number of cases allowed in the learning sample = 31235
Maximum number of cases allowed in the test sample = 0
Max # of nonterminal nodes in the largest tree grown = 8191
  (Actual # of nonterminal nodes in largest tree grown = 148)
Max. no. of categorical splits including surrogates = 49152
Max. number of linear combination splits in a tree = 0
  (Actual number cat. + linear combination splits = 352)
Maximum depth of largest tree grown                = 54
  (Actual depth of largest tree grown                = 13)
Exponent for center weighting in split criterion    = 0.00000
  
```

Elapsed time: 00:00:43.49

CV-Tree Competitor List

(Type, Predictor, Split if continuous, Improvement)

Top Split	Competitors				
	Main	1	2	3	4
CV 1	Categorical CLCC1 0.01083	Categorical CLCC4 0.00607	Categorical CLASSE1_3 0.00552	Categorical CLASSE3_4 0.00335	Categorical CLASSE1_4 0.00252
CV 2	Categorical CLCC1 0.00994	Categorical CLCC4 0.00639	Categorical CLASSE1_3 0.00459	Categorical CLASSE3_4 0.00356	Categorical CLASSE1_4 0.00248
CV 3	Categorical CLCC1 0.01020	Categorical CLCC4 0.00614	Categorical CLASSE1_3 0.00501	Categorical CLASSE3_4 0.00351	Categorical CLASSE1_4 0.00236
CV 4	Categorical CLCC1 0.00955	Categorical CLCC4 0.00590	Categorical CLASSE1_3 0.00475	Categorical CLASSE3_4 0.00312	Categorical CLASSE1_4 0.00211
CV 5	Categorical CLCC1 0.01076	Categorical CLCC4 0.00602	Categorical CLASSE1_3 0.00492	Categorical CLASSE3_4 0.00321	Categorical CLASSE1_4 0.00225
CV 6	Categorical CLCC1 0.01047	Categorical CLCC4 0.00639	Categorical CLASSE1_3 0.00505	Categorical CLASSE3_4 0.00328	Categorical CLASSE1_4 0.00273
CV 7	Categorical CLCC1 0.01086	Categorical CLCC4 0.00582	Categorical CLASSE1_3 0.00464	Categorical CLASSE3_4 0.00305	Categorical CLASSE1_4 0.00247
CV 8	Categorical CLCC1 0.01071	Categorical CLCC4 0.00579	Categorical CLASSE1_3 0.00510	Categorical CLASSE3_4 0.00315	Categorical CLASSE1_4 0.00233
CV 9	Categorical CLCC1 0.01047	Categorical CLCC4 0.00588	Categorical CLASSE1_3 0.00479	Categorical CLASSE3_4 0.00346	Categorical CLASSE1_4 0.00211
CV 10	Categorical CLCC1	Categorical CLCC4	Categorical CLASSE1_3	Categorical CLASSE3_4	Categorical CLASSE1_4

	0.01043	0.00623	0.00487	0.00327	0.00235
FINAL	Categorical	Categorical	Categorical	Categorical	Categorical
	CLCC1	CLCC4	CLASSE1_3	CLASSE3_4	CLASSE1_4
	0.01042	0.00606	0.00492	0.00330	0.00237

Left Split		Competitors				
		Main	1	2	3	4
CV 1	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_3	CLASSE1_4	CLCC3	
	0.00126	0.00085	0.00082	0.00061	0.00060	
CV 2	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_4	CLASSE1_3	CLCC3	
	0.00112	0.00091	0.00063	0.00056	0.00043	
CV 3	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_3	CLASSE1_4	CLCC3	
	0.00111	0.00070	0.00069	0.00058	0.00055	
CV 4	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_3	CLCC3	CLASSE1_4	
	0.00112	0.00080	0.00067	0.00056	0.00050	
CV 5	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_3	CLCC3	CLASSE1_4	
	0.00093	0.00079	0.00060	0.00055	0.00050	
CV 6	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_4	CLASSE1_3	CLCC3	
	0.00108	0.00095	0.00073	0.00066	0.00057	
CV 7	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_4	CLCC3	CLASSE1_3	
	0.00107	0.00084	0.00057	0.00056	0.00048	
CV 8	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_3	CLCC3	CLASSE1_4	
	0.00099	0.00070	0.00066	0.00061	0.00052	
CV 9	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_3	CLCC3	CLASSE1_4	
	0.00101	0.00072	0.00058	0.00057	0.00046	
CV 10	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLASSE1_2	CLCC4	CLASSE1_3	CLASSE1_4	CLCC3	
	0.00104	0.00089	0.00061	0.00057	0.00052	
FINAL						

Right Split		Competitors				
		Main	1	2	3	4
CV 1	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLCC4	CLASSE3_4	CROADSBIN	CLCC5	CLASSE2_4	
	0.00775	0.00553	0.00140	0.00113	0.00101	
CV 2	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLCC4	CLASSE3_4	CROADSBIN	CLCC5	CLASSE2_4	
	0.00788	0.00575	0.00148	0.00092	0.00082	
CV 3	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLCC4	CLASSE3_4	CROADSBIN	CLCC5	CLASSE3_5	
	0.00804	0.00577	0.00137	0.00122	0.00110	
CV 4	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLCC4	CLASSE3_4	CROADSBIN	CLCC5	CLASSE2_4	
	0.00746	0.00509	0.00158	0.00097	0.00097	
CV 5	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLCC4	CLASSE3_4	CROADSBIN	CLASSE2_4	CLCC5	
	0.00775	0.00532	0.00135	0.00118	0.00105	
CV 6	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLCC4	CLASSE3_4	CROADSBIN	CLASSE2_4	CLCC5	
	0.00791	0.00539	0.00135	0.00102	0.00101	
CV 7	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
	CLCC4	CLASSE3_4	CROADSBIN	CLCC5	CLASSE3_5	
	0.00751	0.00520	0.00143	0.00133	0.00114	
CV 8	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical

		CLCC4	CLASSE3_4	CROADSBIN	CLASSE2_4	CLCC5
		0.00763	0.00530	0.00150	0.00115	0.00110
CV 9	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
		CLCC4	CLASSE3_4	CROADSBIN	CLCC5	CLASSE3_5
		0.00764	0.00560	0.00153	0.00120	0.00108
CV 10	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
		CLCC4	CLASSE3_4	CROADSBIN	CLASSE2_4	CLCC5
		0.00787	0.00534	0.00154	0.00125	0.00104
FINAL	Categorical	Categorical	Categorical	Categorical	Categorical	Categorical
		CLCC4	CLASSE3_4	CROADSBIN	CLCC5	CLASSE2_4
		0.00774	0.00543	0.00145	0.00109	0.00103

Resubstitution, CV and Breiman Cost Summary

Tree	Nodes	Relative Error Rates			Breiman Lambda
		Resub	CV	Breiman	
1	135	0.7880660	0.8047564	0.8047564	0.0001050
2	112	0.7881392	0.8069991	0.8069992	0.0001543
3	62	0.7896248	0.8068301	0.8068301	0.0002515
4	47	0.7903364	0.8064307	0.8064307	0.0002703
5	39	0.7910993	0.8060953	0.8060954	0.0002957
6	38	0.7912043	0.8060645	0.8060645	0.0002972
7	32	0.7919716	0.8065054	0.8065054	0.0003194
8	29	0.7924632	0.8096108	0.8096109	0.0004145
9	24	0.7934497	0.8102403	0.8102403	0.0004896
10	20	0.7944946	0.8077999	0.8077999	0.0004655
11	19	0.7948877	0.8075445	0.8075445	0.0004673
12	17	0.7958207	0.8079418	0.8079418	0.0005006
13	16	0.7963229	0.8060630	0.8060630	0.0004058
14	11	0.7999768	0.8073612	0.8073612	0.0004340
15	9	0.8021184	0.8136306	0.8136306	0.0008246
16	6	0.8070400	0.8120074	0.8120073	0.0005196
17	3	0.8217376	0.8217376	0.8217376	0.0000000
18	2	0.8923832	0.8923832	0.8923832	0.0000000
19	1	1.0000000	1.0000000	1.0000000	0.0000000

Breiman Cost Node Detail For the 11-Node Tree

```

=====
Lambda = 0.0004340
CV Relative Error = 0.8073612
RS Relative Error = 0.7999768
Min CV Relative Error = 0.8047564

Terminal Node 1
Learn Sample Class Assignment: 1
Learn Sample Cost :
0.3386843
Breiman Adjusted Cost :
0.3407731
Learn Sample Prob :
0.1667798
Terminal Node 2
Learn Sample Class Assignment: 1
Learn Sample Cost :
0.3856666
Breiman Adjusted Cost :
0.3881252
Learn Sample Prob :
0.1416329
Terminal Node 3
Learn Sample Class Assignment: 1
Learn Sample Cost :
0.3634135
Breiman Adjusted Cost :
0.4937425

Learn Sample Prob :
0.0022459
Terminal Node 4
Learn Sample Class Assignment: 0
Learn Sample Cost :
0.3872186
Breiman Adjusted Cost :
0.4546296
Learn Sample Prob :
0.0047473
Terminal Node 5
Learn Sample Class Assignment: 1
Learn Sample Cost :
0.4307956
Breiman Adjusted Cost :
0.4412012
Learn Sample Prob :
0.0331321
Terminal Node 6
Learn Sample Class Assignment: 1
Learn Sample Cost :
0.3113199

```

```

Breiman Adjusted Cost      :          Learn Sample Cost      :
0.4033499                 :          0.4851327
Learn Sample Prob         :          Breiman Adjusted Cost      :
0.0033612                 :          0.4888084
Terminal Node 7           :          Learn Sample Prob         :
Learn Sample Class Assignment: 0      :          0.0945881
Learn Sample Cost         :          Terminal Node 10           :
0.4885630                 :          Learn Sample Class Assignment: 1
Breiman Adjusted Cost     :          Learn Sample Cost         :
0.4917966                 :          0.4075766
Learn Sample Prob         :          Breiman Adjusted Cost      :
0.1075806                 :          0.4252695
Terminal Node 8           :          Learn Sample Prob         :
Learn Sample Class Assignment: 1      :          0.0193070
Learn Sample Cost         :          Terminal Node 11           :
0.4174988                 :          Learn Sample Class Assignment: 0
Breiman Adjusted Cost     :          Learn Sample Cost         :
0.4324619                 :          0.3839935
Learn Sample Prob         :          Breiman Adjusted Cost      :
0.0229085                 :          0.3848577
Terminal Node 9           :          Learn Sample Prob         :
Learn Sample Class Assignment: 0      :          0.4037165

```

CV Partition Details

CV Cycle	-----Focal Class-----		-----Other Classes-----	
	Learn	Test	Learn	Test
1	21501.00	2390.00	6610.00	734.00
2	21501.00	2390.00	6610.00	734.00
3	21502.00	2389.00	6609.00	735.00
4	21502.00	2389.00	6609.00	735.00
5	21501.00	2390.00	6610.00	734.00
6	21502.00	2389.00	6610.00	734.00
7	21502.00	2389.00	6610.00	734.00
8	21503.00	2388.00	6609.00	735.00
9	21503.00	2388.00	6609.00	735.00
10	21502.00	2389.00	6610.00	734.00
SRW	17545548.00			

CV ROC for each tree in the main sequence

Tree	Nodes	----- Learn -----		----- Test -----	
		ROC	Var(ROC)	ROC	Var(ROC)
1	135	0.6456179	0.0000135	0.6346448	0.0001373
2	112	0.6448427	0.0000135	0.6328862	0.0001371
3	62	0.6414076	0.0000133	0.6331163	0.0001364
4	47	0.6400687	0.0000133	0.6314347	0.0001360
5	39	0.6371660	0.0000132	0.6305896	0.0001348
6	38	0.6371349	0.0000132	0.6294686	0.0001347
7	32	0.6366966	0.0000133	0.6291730	0.0001344
8	29	0.6362445	0.0000133	0.6291166	0.0001343
9	24	0.6339629	0.0000133	0.6279932	0.0001339
10	20	0.6301101	0.0000132	0.6249194	0.0001328
11	19	0.6300287	0.0000132	0.6242842	0.0001329
12	17	0.6292633	0.0000132	0.6246593	0.0001326
13	16	0.6289705	0.0000132	0.6223944	0.0001329
14	11	0.6223948	0.0000132	0.6189508	0.0001324
15	9	0.6214124	0.0000132	0.6176355	0.0001309
16	6	0.6157775	0.0000128	0.6140541	0.0001279
17	3	0.5989898	0.0000121	0.5990026	0.0001210
18	2	0.5538084	0.0000069	0.5538085	0.0000690
19	1	0.5000000	0.0000000	0.5000000	0.0000000

Tree Optimality Criteria

=====

Target: CFRONTBIN

Terminal Tree	Nodes	Cross Valid Rel Error	Resub Rel Error	Test Set ROC	Test Set 10% Lift
1	135	0.8047564*	0.7880660	0.6346448*	1.8228420*
2	112	0.8069991	0.7881392	0.6328862	1.8146772
3	62	0.8068301	0.7896248	0.6331163	1.8053286
4	47	0.8064307	0.7903364	0.6314347	1.7868111
5	39	0.8060953	0.7910993	0.6305896	1.7629251
6	38	0.8060645	0.7912043	0.6294686	1.7436052
7	32	0.8065054	0.7919716	0.6291730	1.7387687
8	29	0.8096108	0.7924632	0.6291166	1.7414103
9	24	0.8102403	0.7934497	0.6279932	1.7106957
10	20	0.8077999	0.7944946	0.6249194	1.6225456
11	19	0.8075445	0.7948877	0.6242842	1.6217706
12	17	0.8079418	0.7958207	0.6246593	1.6071928
13	16	0.8060630	0.7963229	0.6223944	1.6045007
14	11	0.8073612	0.7999768	0.6189508	1.5897176
15	9	0.8136306	0.8021184	0.6176355	1.5921975
16	6	0.8120074	0.8070400	0.6140541	1.5950443
17	3	0.8217376	0.8217376	0.5990026	1.5950443
18	2	0.8923832	0.8923832	0.5538085	1.5950443
19	1	1.0000000	1.0000000	0.5000000	1.0000000

Grove file created: C:\Users\Usuario\AppData\Local\Temp\s4ik109: 266 kb

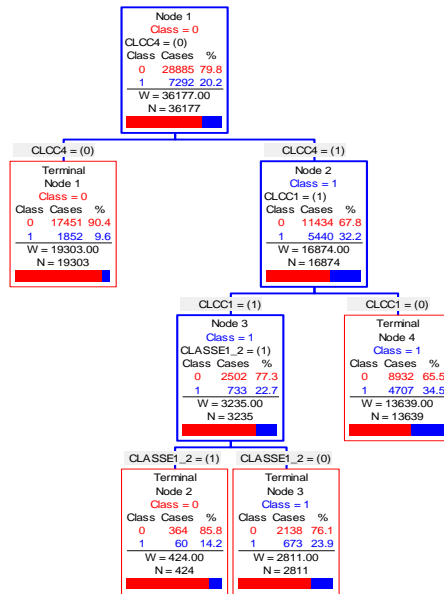
Grove file created containing:

1 CART Tree

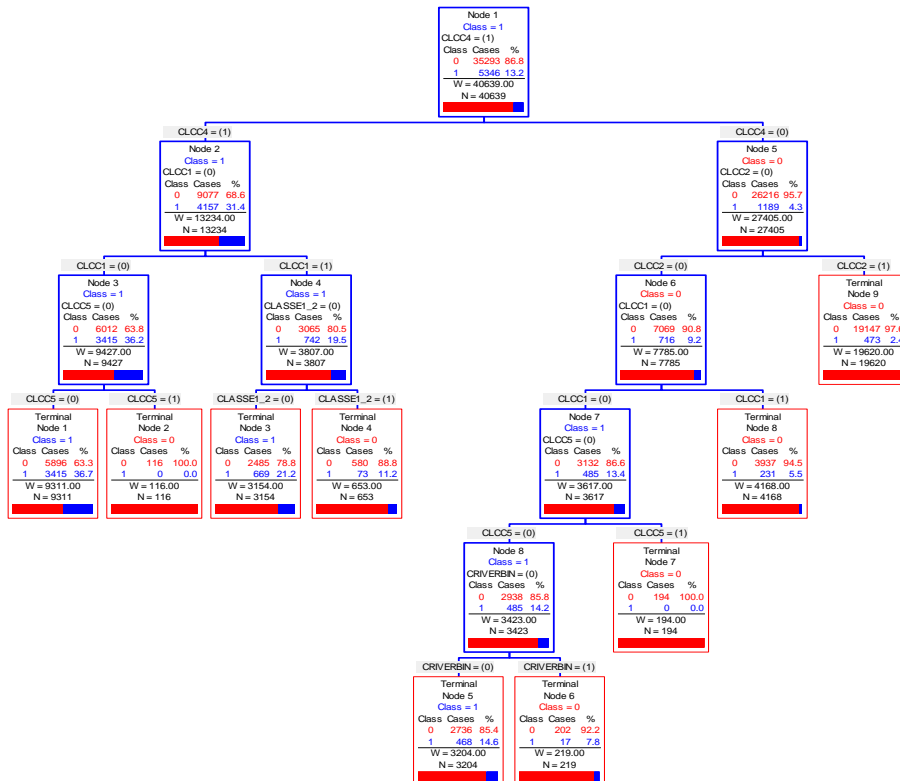
E:\final\pinhal-int-sul.xlsx: 31236 records.

ANEXO 8. Árvores de Classificação para as 21 Regiões PROF.

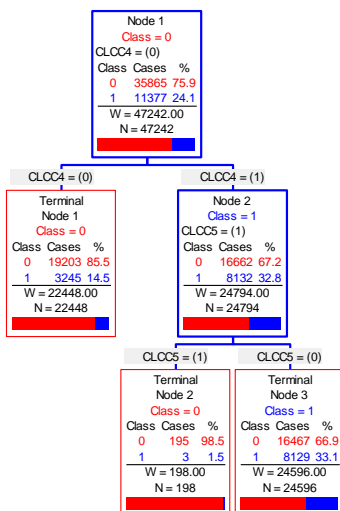
Alto Minho



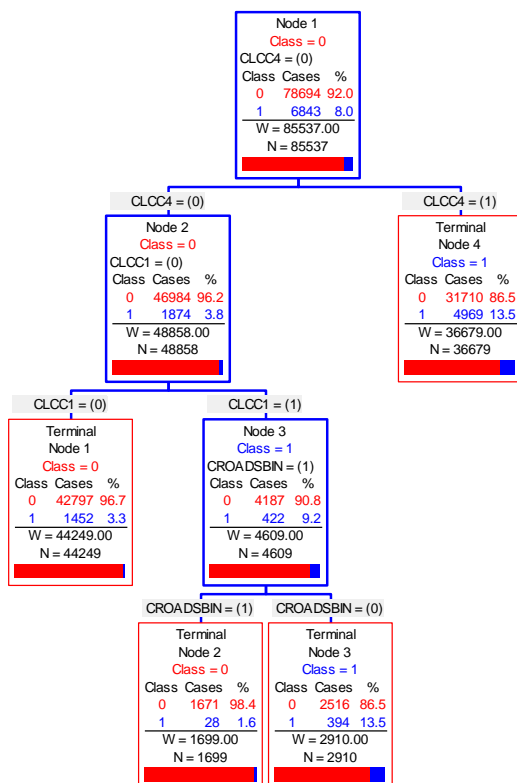
Baixo Minho



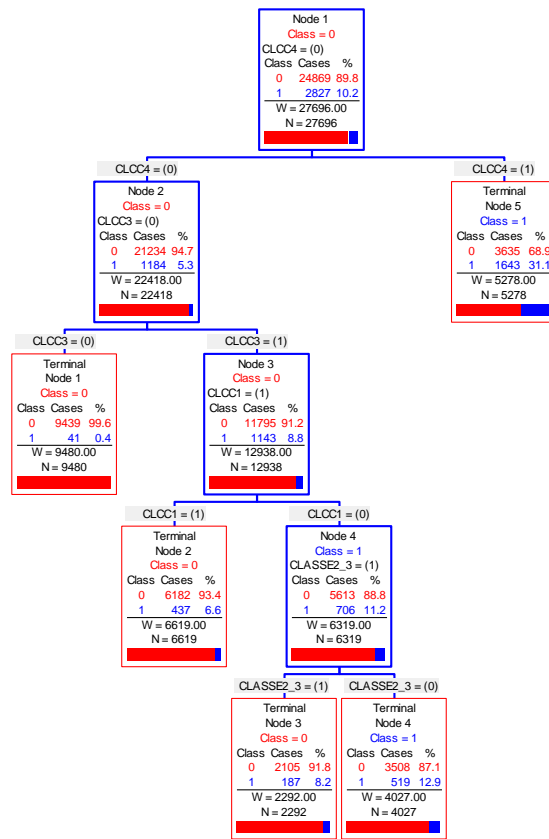
Barroso-Padrela



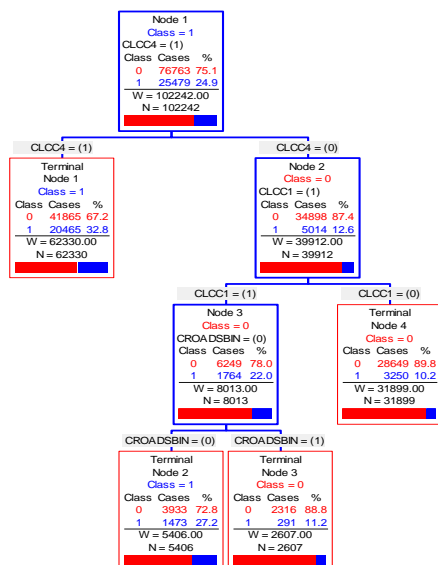
Nordeste



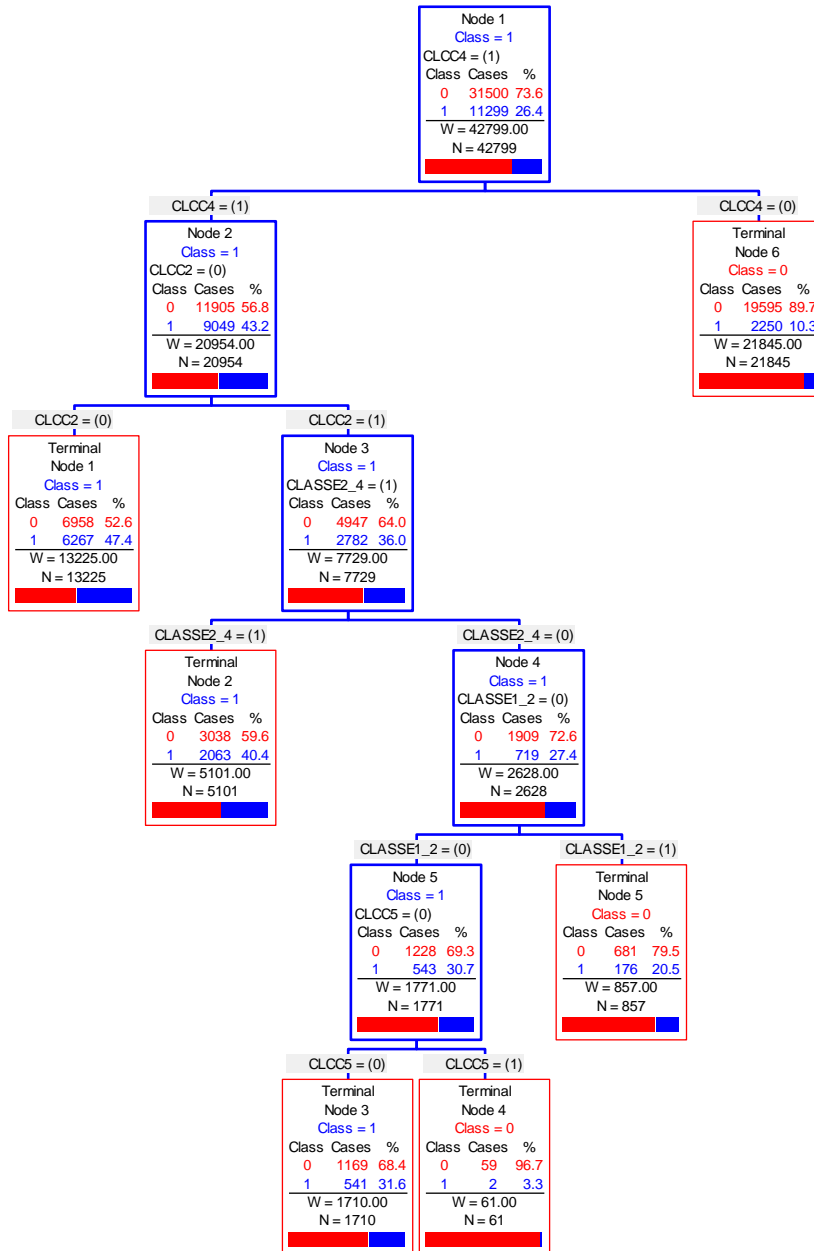
Área Metropolitana do Porto e Entre Douro e Vouga



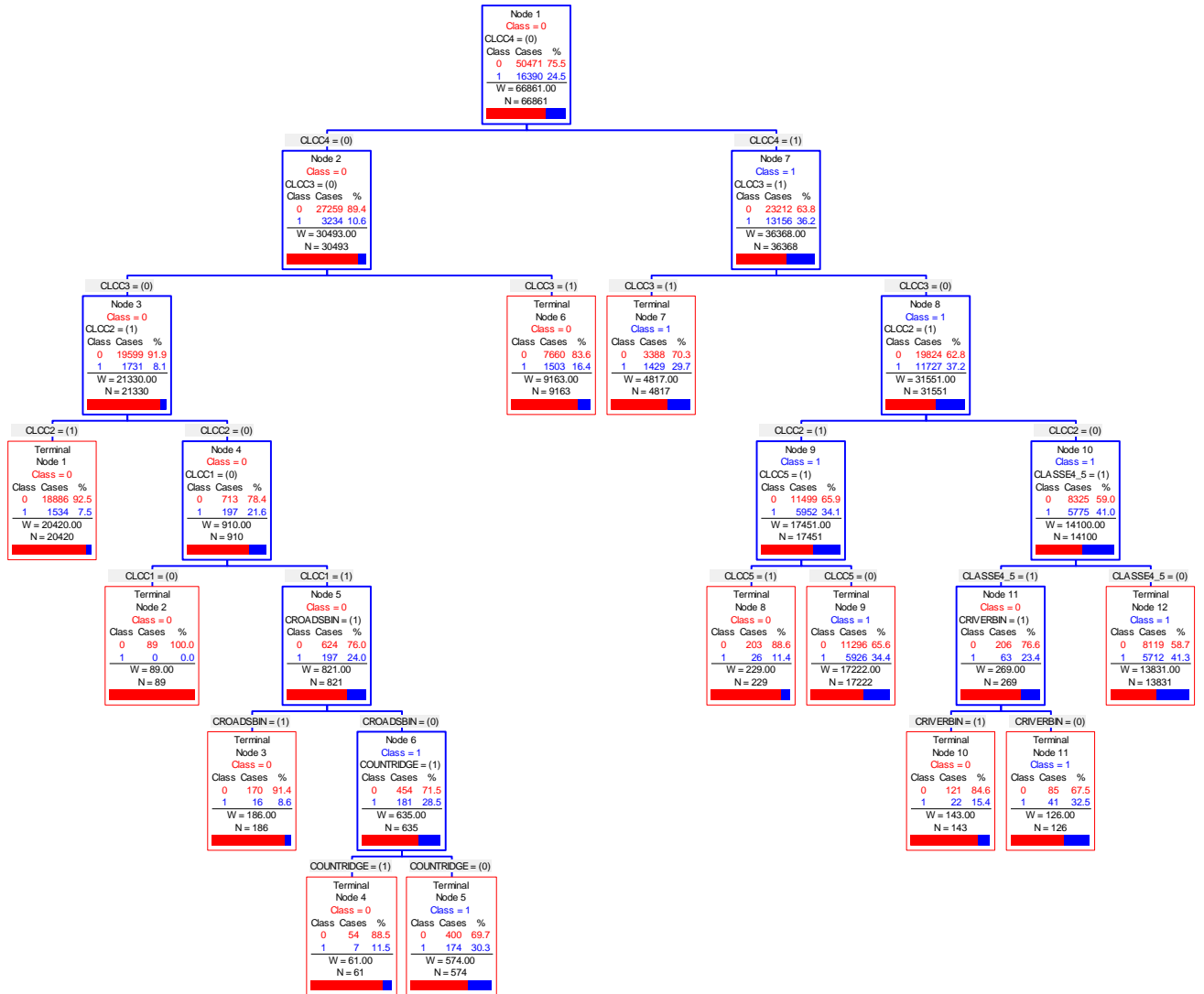
Beira Interior Norte



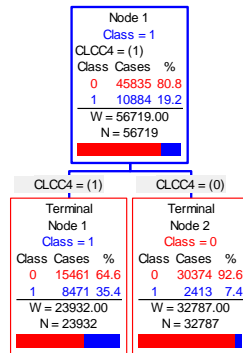
Tâmega



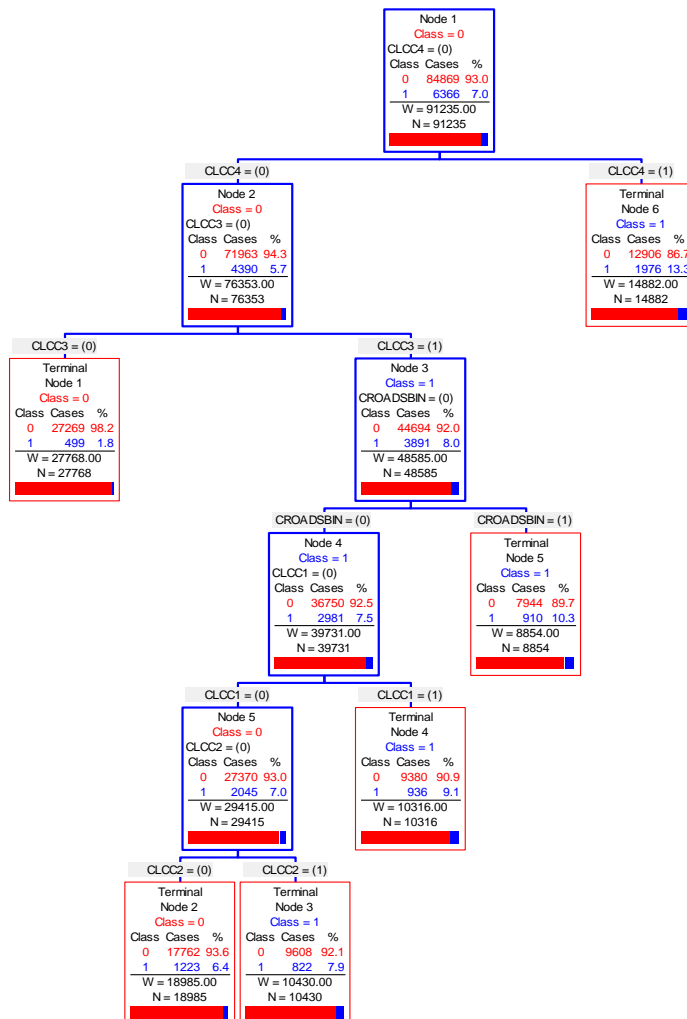
Douro



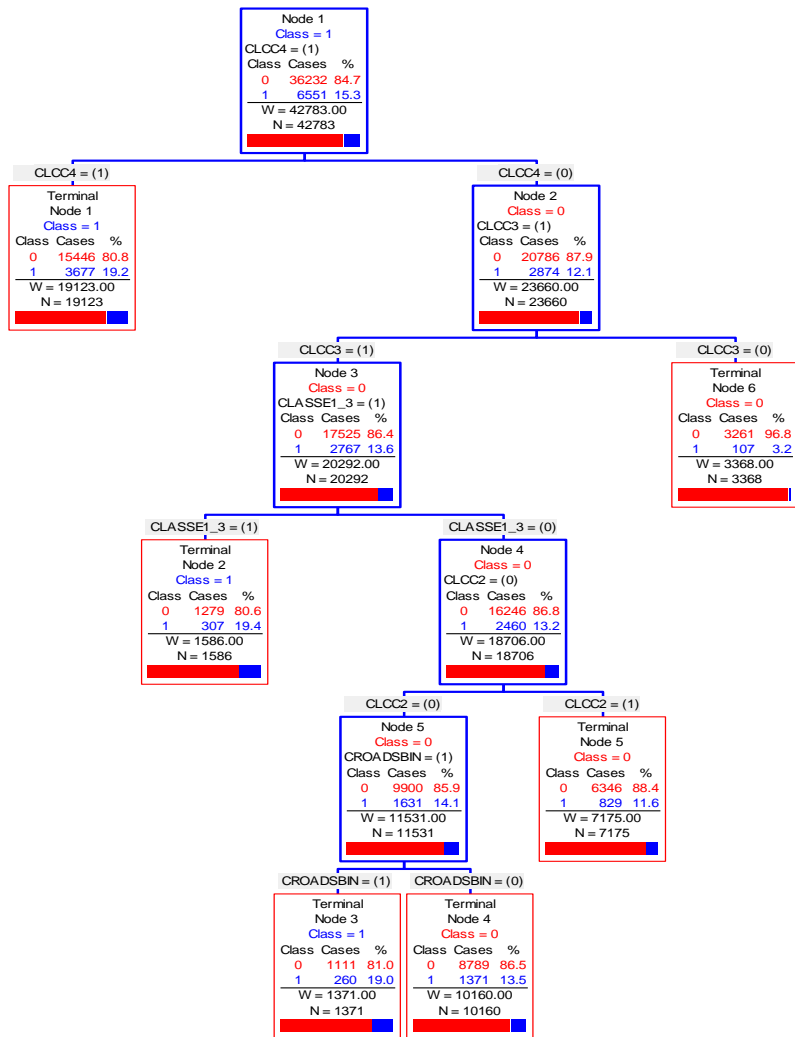
Dão-Lafões



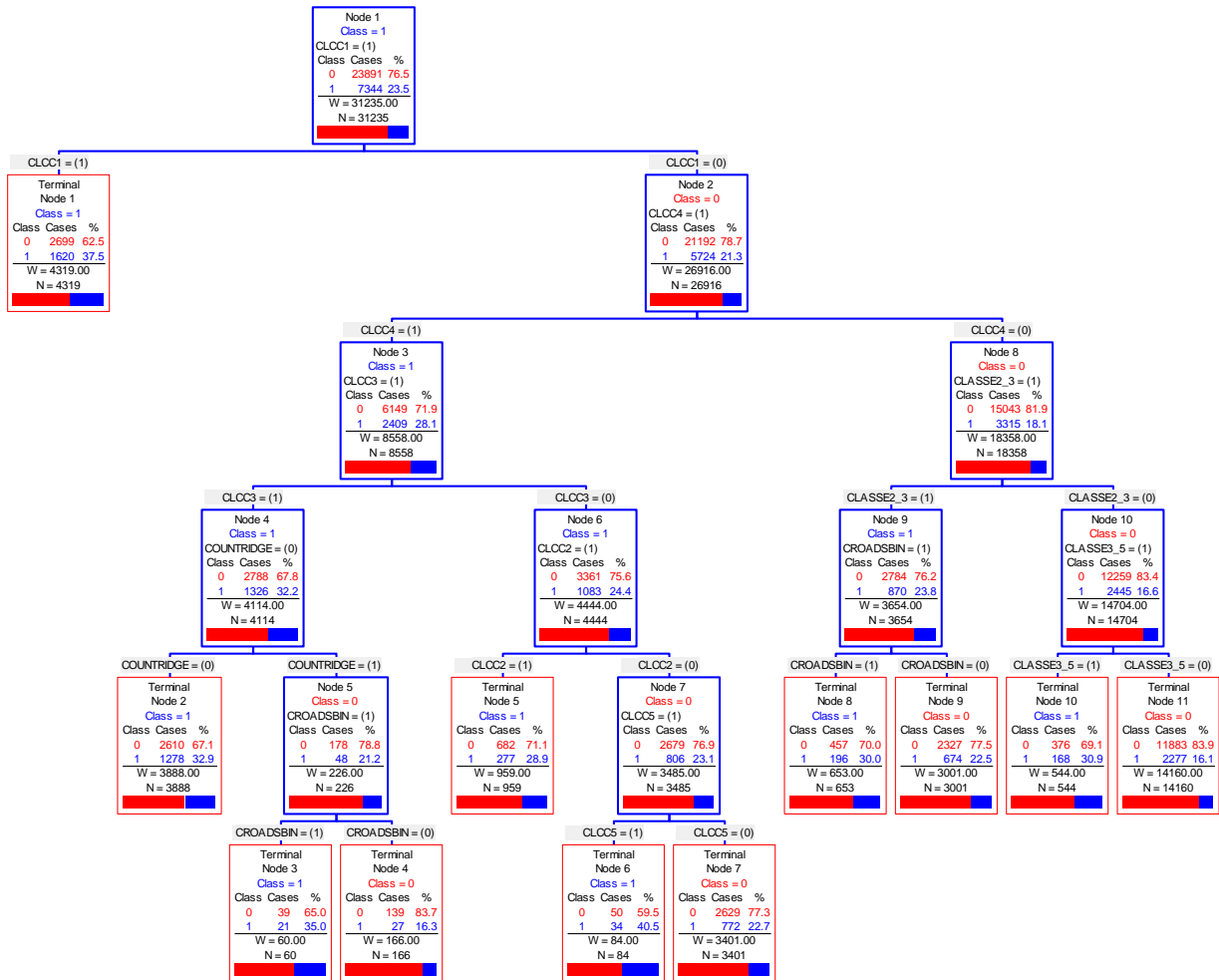
Centro Litoral



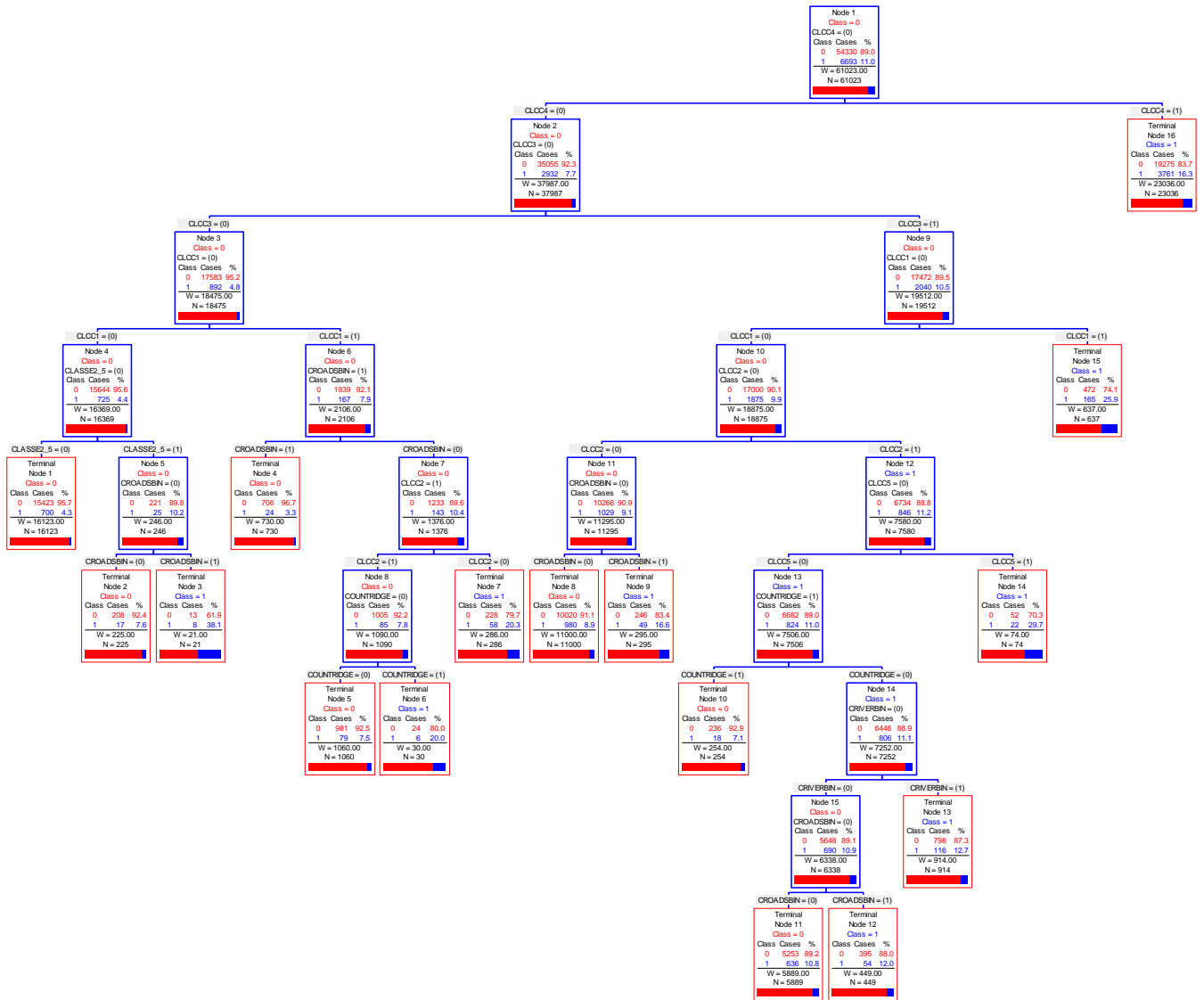
Pinhal Interior Norte



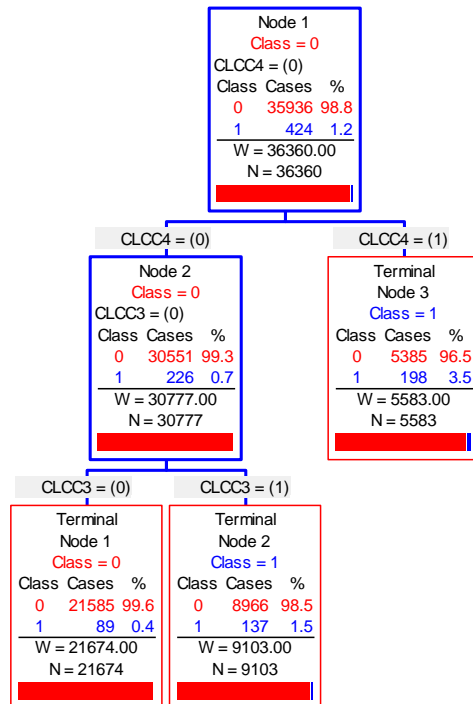
Pinhal Interior Sul



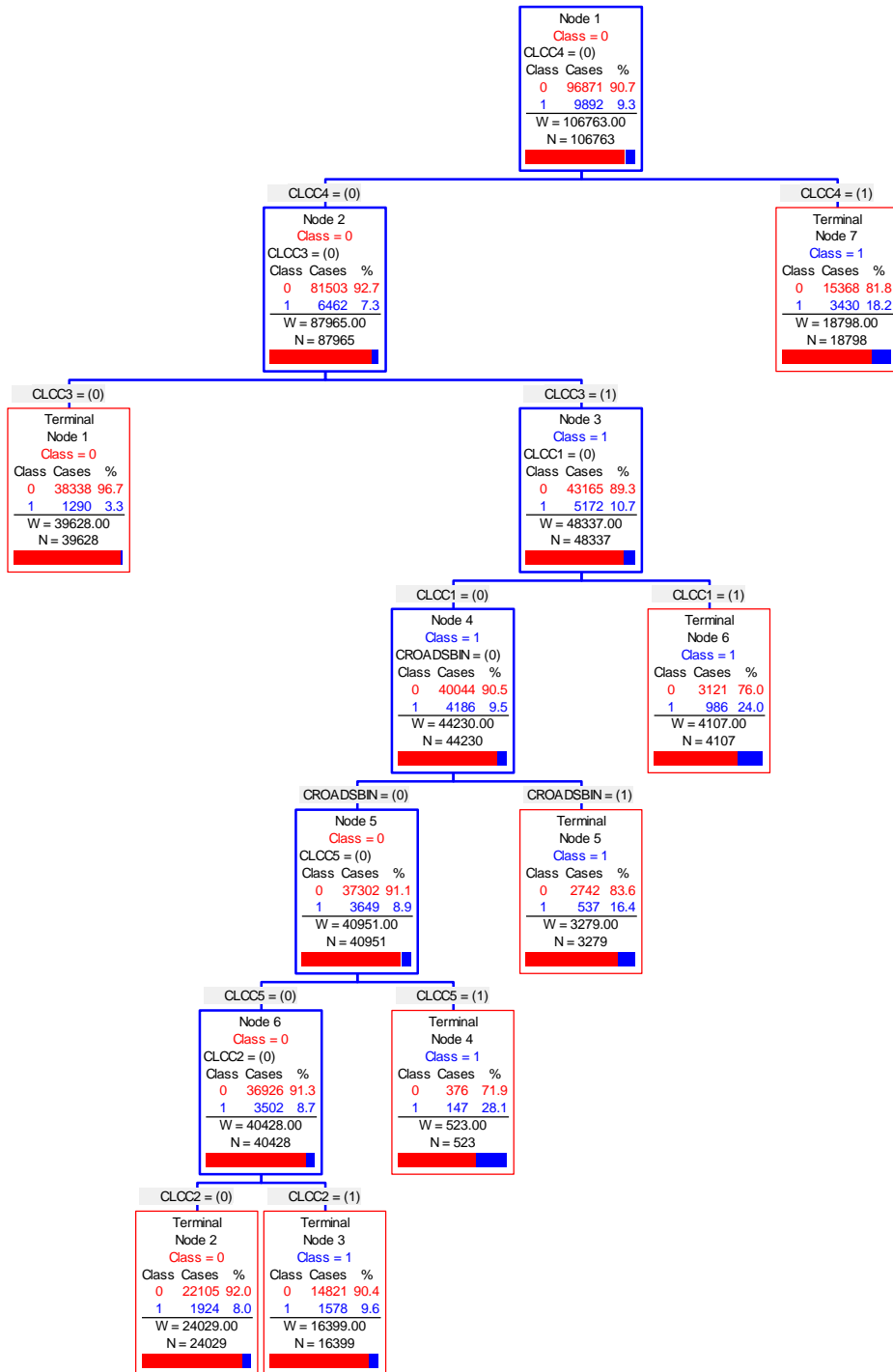
Beira Interior Sul



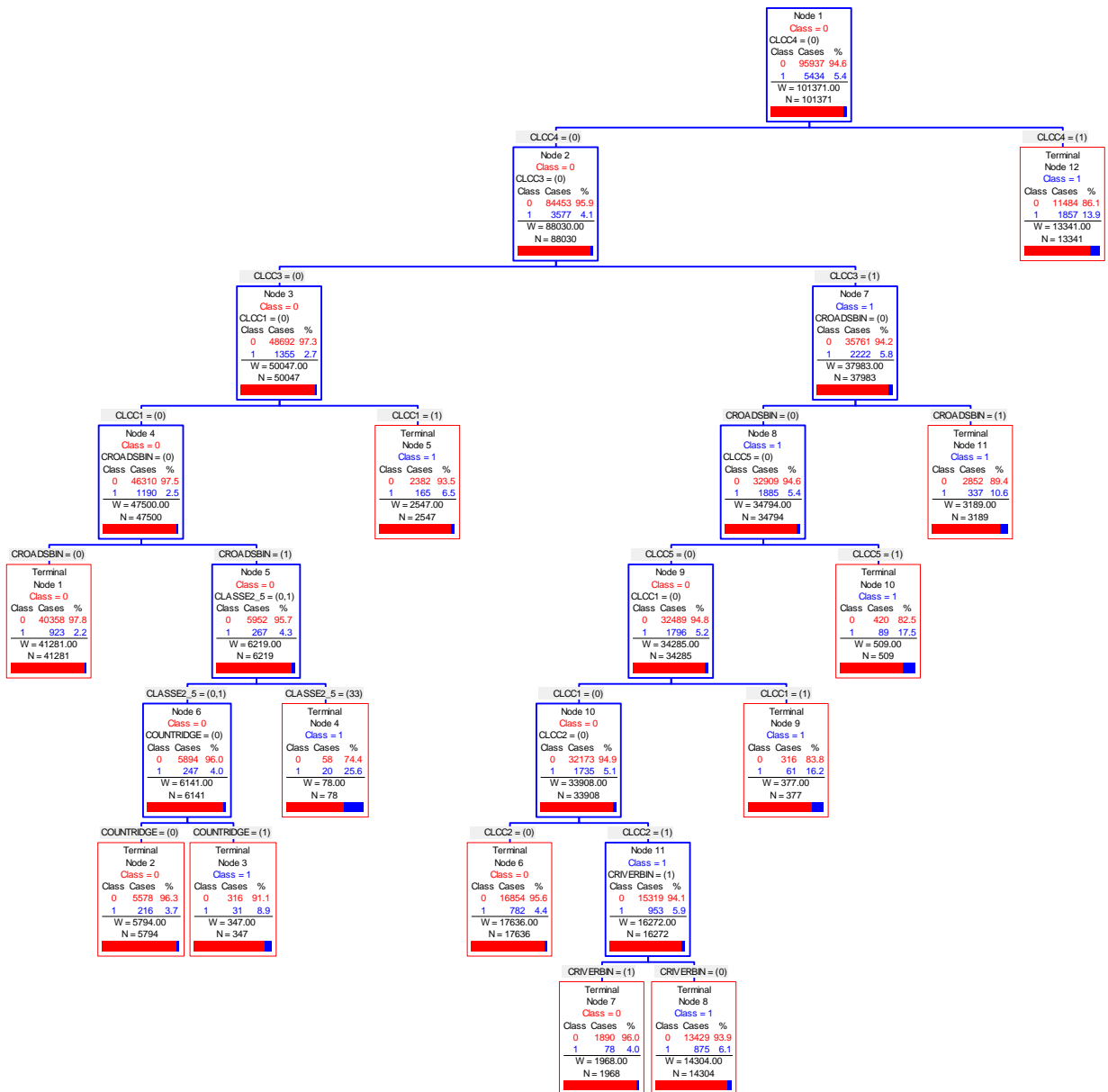
Oeste



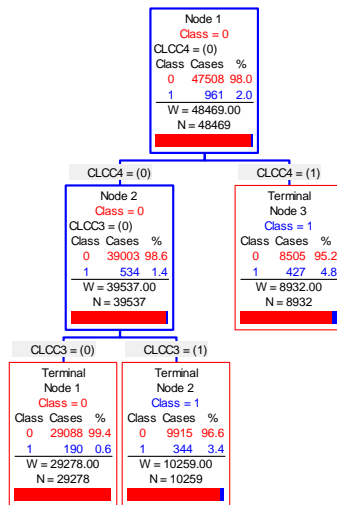
Ribatejo



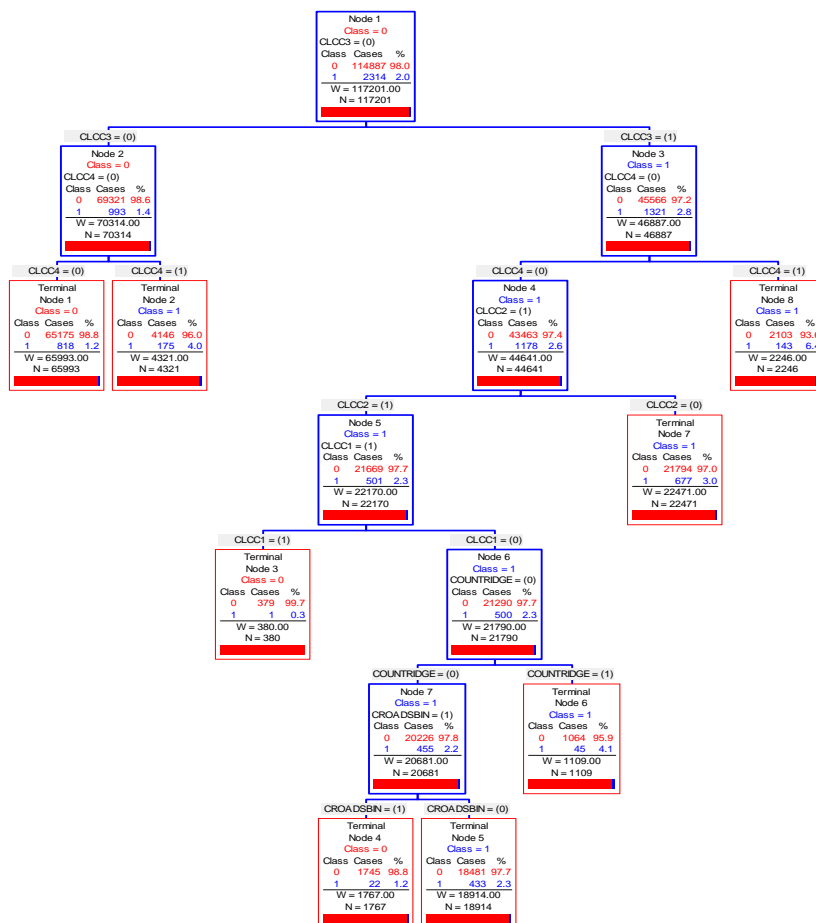
Alto Alentejo



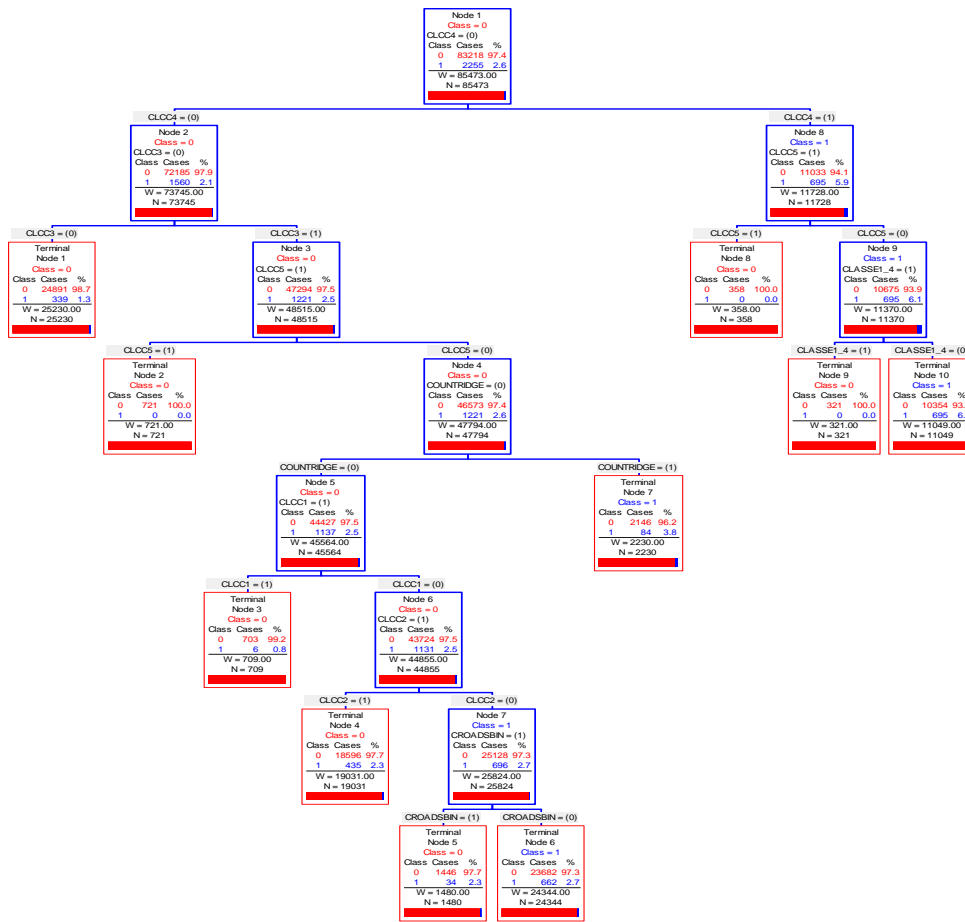
Área Metropolitana de Lisboa



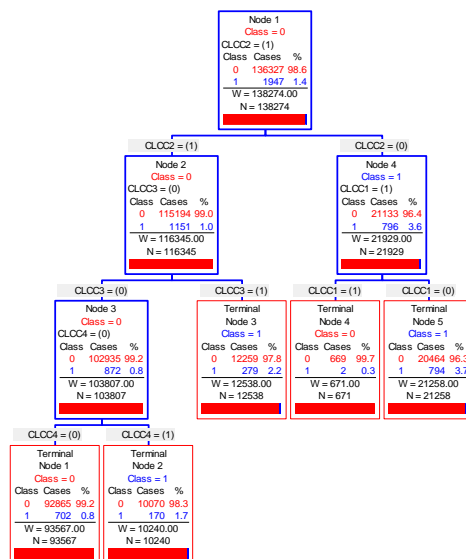
Alentejo Central



Alentejo Litoral



Baixo Alentejo



Algarve

