

CLASSIFICAÇÃO SUPERVISIONADA DE IMAGENS SENTINEL-2 PARA IDENTIFICAÇÃO DE FLORESTA NO PARQUE NATURAL DE MONTESINHO

Leticia Gonçalves Parlotti

*Dissertação apresentada à Escola Superior Agrária de Bragança para
obtenção do Grau de Mestre em Gestão de Recursos Florestais no âmbito
da dupla diplomação com a Universidade Tecnológica Federal do Paraná*

Orientada por

**José Manuel Ferreira de Castro
Raoni Wainer Duarte Bosquilia**

LETICIA GONÇALVES PARLOTTI

CLASSIFICAÇÃO SUPERVISIONADA DE IMAGENS SENTINEL-2 PARA IDENTIFICAÇÃO DE FLORESTA NO PARQUE NATURAL DE MONTESINHO

Dissertação apresentada à Escola Superior Agrária de Bragança para obtenção do Grau de Mestre em Gestão de Recursos Florestais no âmbito da dupla diplomação com a Universidade Tecnológica Federal do Paraná.

Orientador: Prof. José Manuel Ferreira de Castro.

Co-orientador: Prof. Raoni Wainer Duarte Bosquilia.

Bragança
2021

A minha Família, amigos,
meus orientadores e a UTFPR.

Abstract

Currently, with the advancement of technologies and the ease of obtaining satellite images, the monitoring and spatial analysis of terrestrial vegetation, with reduced costs and efficiency, make it possible to assemble profiles of multitemporal and seasonal images of the earth's surface. This work aims to identify the best time for classification and calculation of forest vegetation indices in the Montesinho Natural Park (PNM) and to locate its forests of greatest natural value holm oak (*Quercus rotundifolia*), black oak (*Quercus pyrenaica* Willd.), maritime pine (*Pinus pinaster* Aiton) and chestnut (*Castanea sativa* Miller) using remote sensing and geoprocessing techniques. All data processing was performed in QGIS 3.18.1 software using SCP (Semi-Automatic Classification Plugin). Vegetation indices and the supervised Random Forest (RF) classification were used. For the production of maps, training samples (ROI) were collected, based on the support material: land use and occupation charts (COS), and the national forest inventory (IFN) of 2015, made available by the ICNF, where it was possible to remotely identify the forests of the studied species. To validate the results, the Kappa Index and Global Accuracy were calculated for the three months studied: the month with the highest index was July (63%), followed by February (41%) and October (15%). We can relate these results with the season of the year of the images, being those of winter and autumn, where the plants are in vegetative rest. It is concluded that the classification made by the RF for the identification of the species of interest fits better with the real (IFN) in the month of July. However, more study applied to the area is needed..

Keywords: Random Forest; Remote Sensing; Temporal analysis.

Resumo

Atualmente, com o avanço de tecnologias e a facilidade de obtenção de imagens de satélite, o monitoramento e análises espaciais da vegetação terrestre, com custos reduzidos e de forma eficiente, possibilitam montar perfis de imagens multitemporais e sazonais da superfície terrestre. Este trabalho tem como objetivo identificar a melhor época para classificação e cálculo de índices de vegetação florestal no Parque Natural de Montesinho (PNM), e localizar os seus bosques de maior valor natural – azinheira (*Quercus rotundifolia*), carvalho negral (*Quercus pyrenaica Willd.*), pinheiro-bravo (*Pinus pinaster Aiton*) e castanheiro (*Castanea sativa Miller*) – através de técnicas de teledetecção remota e geoprocessamento. Todo o processamento dos dados foi realizado no software QGIS 3.18.1 utilizando o SCP (Semi-Automatic Classification Plugin). Foram utilizados índices de vegetação e a classificação supervisionada Random Forest (RF). Para a produção dos mapas recolheram-se as amostras de treinamento (ROI), com base no material de apoio: cartas de uso e ocupação do solo (COS), e o inventário florestal nacional (IFN) de 2015, disponibilizado pelo ICNF, onde foi possível identificar de forma remota os bosques das espécies estudadas. Para a validação dos resultados calculou-se o Índice Kappa e Exatidão Global para os três meses estudados: o mês com o maior índice foi o de julho (63%), seguido de fevereiro (41%) e outubro (15%). Podemos relacionar esses resultados com a estação do ano das imagens, sendo essas de inverno e de outono, onde as plantas estão em repouso vegetativo. Conclui-se que a classificação feita pelo RF para a identificação das espécies de interesse se ajusta melhor com o real (IFN) no mês de julho. No entanto é necessário mais estudo aplicado à área.

Palavras-chaves: Random Forest; Detecção Remota; Análise Temporal.

LISTA DE FIGURAS

Figura 1. Configuração orbital dos satélites gêmeos Sentinel-2.....	17
Figura 2. Mapa de localização do Parque Natural de Montesinho Trás-os-Montes em Portugal continental.	26
Figura 3. Fluxograma do processamento do método de classificação Random Forest.	30
Figura 4. Índice de vegetação por diferença normalizada (NDVI) da área de estudo, referente ao mês de fevereiro.....	35
Figura 5. Índice de vegetação por diferença normalizada (NDVI) da área de estudo, referente ao mês de julho.	35
Figura 6. Índice de vegetação por diferença normalizada (NDVI) da área de estudo, referente ao mês de outubro.	35
Figura 7. Mapa temático da classificação supervisionada do Parque Natural de Montesinho referente ao mês fevereiro de 2020.	38
Figura 8. Mapa temático da classificação supervisionada do PNM referente ao mês de julho de 2020.	43
Figura 9. Classificação supervisionada RF do Parque Natural de Montesinho referente ao mês de outubro de 2020.	44
Figura 10. Dispersão da azinheira em relação com o Ombrotipo do Parque Natural de Montesinho.	48
Figura 11. Dispersão da azinheira em relação com o Termotipo do Parque Natural de Montesinho.	48

LISTA DE TABELAS

Tabela 1. Número de bandas e comprimento de ondas do Sensor Sentinel-2	28
Tabela 2. Imagens utilizadas para a produção dos mapas de interesse.	29
Tabela 3 Representação da Matriz de Confusão.....	32
Tabela 4. Matriz de confusão para a análise da acurácia obtida na classificação supervisionada Random Forest, com informações de acurácia e erros de produtor e usuário, do mês de fevereiro de 2020.....	41
Tabela 5. Matriz de confusão para a análise da acurácia obtida na classificação supervisionada Random Forest, com informações de acurácia e erros de produtor e usuário, do mês de julho de 2020.	46
Tabela 6. Matriz de confusão para a análise da acurácia obtida na classificação supervisionada Random Forest, com informações de acurácia e erros de produtor e usuário, do mês de outubro de 2020.....	47

LISTA DE GRÁFICOS

Gráfico 1. Uso e Ocupação do Solo de 2018 em Portugal Continental.....	23
Gráfico 2. Valores médios do NDVI relacionado as épocas do ano com base nas espécies de interesse.....	36
Gráfico 3. Relação dos valores médios de SAVI com os meses do ano analisados, com base nas espécies de interesse.	37
Gráfico 4. Análise de acurácia das amostras, referentes ao mês de fevereiro de 2020.	39
Gráfico 5. Análise da acurácia das amostras, referentes ao mês de junho de 2020.	42
Gráfico 6. Análise da acurácia das amostras, referentes ao mês de outubro de 2020.	44

LISTA DE SIGLAS, ACRÓNIMOS E ABREVIATURAS

COS	Carta de Uso e Ocupação do Solo
DGT	Direção Geral do Território
ESA	<i>European Space Agency</i> /Agência Espacial Europeia
FAPAR	Fração da Radiação Absorvida pela Vegetação Fotossinteticamente Ativa
GNSS	<i>Global Navigation Satellite System</i>
ICNF	Instituto da Conservação da Natureza e das Florestas
IAF	Índice de Área Foliar
IPB	Instituto Politécnico de Bragança
LANDSAT	Land Remote Sensing Satellite
MSI	Multispectral Imaging
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
ND	Número digital
NIR	Reflectância do infravermelho próximo
PNM	Parque Natural de Montesinho
RF	Random Forest
RGB	<i>Red,green,blue</i>
ROI	Regions of Interest
SAVI	Soil-adjusted Vegetation Index
SCP	Semi-Automatic Classification Plugin
SIG	Sistema de Informação Geográfica
USGS	United States Geological Survey
UTFPR	Universidade Tecnológica Federal do Paraná

SUMÁRIO

1	INTRODUÇÃO	11
2	REVISÃO DE LITERATURA	13
2.1	Sensoriamento Remoto	13
2.1.1	Plataformas e sensores.....	16
2.1.2	Índice de Vegetação da Diferença Normalizada (NDVI)	19
2.1.3	Classificador Random Forest	22
2.2	A Floresta em Portugal e no PNM	23
3	MATERIAIS E MÉTODOS	26
3.1	Área de Estudo	26
3.2	Seleção e processamento das imagens	27
3.3	Tratamento e Pós Processamento das imagens	30
4	RESULTADOS E DISCUSSÃO	34
4.1	Índices de Vegetação	34
4.2	Classificação por Random Forest.....	38
5	CONCLUSÃO	49
6	REFERÊNCIAS.....	50

1 INTRODUÇÃO

Estudos utilizando técnicas de sensoriamento remoto correlacionados ao monitoramento da cobertura vegetal podem auxiliar no diagnóstico de degradação ambiental de uma determinada área (SZATMÁRI; TOBAK; NOVÁK, 2016). Assim, esta ciência tem um papel importante para estudos que visam a garantia da preservação dos recursos naturais associados com o desenvolvimento sustentável (MOREIRA, 2012).

Atualmente, com o avanço da tecnologia e a facilidade de obtenção de imagens de satélites, o monitoramento e análises espaciais da vegetação terrestre, com custos reduzidos e de forma eficiente, possibilitam montar perfis multitemporais da superfície terrestre (PONZONI; SHIMABUKURO E KUPLICH 2015). Além disso, a partir da análise de diferentes componentes da paisagem em que o ambiente estudado está inserido, é possível colaborar com a elaboração de um plano de zoneamento e propor o manejo adequado para cada situação.

Para a análise de conservação da vegetação é comum utilizar índices de vegetação que foram desenvolvidos com o objetivo de realçar o comportamento espectral da vegetação na faixa do vermelho e infravermelho próximo (PONZONI; SHIMABUKURO E KUPLICH 2015), sendo o NDVI (*Normalized Difference Vegetation Index*) o índice mais utilizado, pois facilita o mapeamento de classes e variações da vegetação destacando o vigor vegetativo (DAINESE, 2001).

Com base no exposto o intuito desde trabalho é de coletar diferentes assinaturas espectrais da vegetação ao longo do ano de 2020 para assim identificar qual a melhor época para a classificação e identificação, nomeadamente a dos bosques de azinheira (*Quercus rotundifolia*), de carvalho negral (*Quercus pyrenaica* Willd.), de pinheiro-bravo (*Pinus pinaster* Aiton) e de castanheiro (*Castanea sativa* Miller.), que são os elementos de maior valor natural na paisagem do Parque Natural de Montesinho. A melhor alternativa para tal é utilizar uma série temporal suavizada com vetores de entrada para

classificar automaticamente as imagens. Isto justifica o desenvolvimento e a implementação de um sistema de monitorização robusto e expedito.

Neste sentido o trabalho explora ferramentas de sensoriamento remoto e geoprocessamento com base na coleção de imagens de satélite disponíveis com mais de duas décadas, o 6º Inventário Nacional Florestal de 2015 e cartografia detalhada elaborada pelo IPB (Instituto Politécnico de Bragança), para o plano de gestão desta área protegida.

Para tal, foi investigada a melhor época de captação de imagens para identificação dos bosques de interesse, nomeadamente os de azinheira e de carvalho, de castanheiro e de pinheiro bravo. Adicionalmente foram exploradas assinaturas espectrais características destas ocupações de solo e testadas técnicas mistas de pesquisa, tais como interpretação visual de fotografia aérea e a cartografia de uso e ocupação do solo histórica e atual.

Foi privilegiado o uso de ferramentas, informação e programas informáticos de acesso livre, no sentido de assegurar a continuidade e sustentabilidade do sistema de monitorização proposto, assim como os classificadores de imagem automáticos e assistidos mais recentes e generalizados, nomeadamente plugin SCP do Qgis, e imagens do sensor Sentinel-2 longo do ano de 2020.

2 REVISÃO DE LITERATURA

2.1 Sensoriamento Remoto

O sensoriamento remoto é a ciência que possibilita obter informação referente a um objeto, a uma área, ou a um fenômeno, sem entrar em contato com o mesmo (LILLESAND, KIEFER e CHIPMAN, 2014). Tais informações são obtidas por satélites artificiais que se encontram em circulação na órbita terrestre; acoplados a estes satélites estão os sistemas sensores operando em diferentes faixas do espectro eletromagnético. Os sistemas sensores que são responsáveis por coletar a energia incidente do objeto e converte-la em dados passíveis de interpretação (LIMA, 2018).

Esta ciência utiliza procedimentos direcionados à obtenção de imagens com registros da energia eletromagnética (SPRING, 2018). A energia eletromagnética, ou energia radiante, é salientada, devido à forma como interage com a matéria, podendo ser refletida, absorvida ou transmitida pelo objeto onde o fluxo de energia emitida e refletida se propaga pela atmosfera (MOREIRA, 2012). Os processos de absorção, reflexão e transmissão da energia ocorrem de forma simultânea e com intensidades diferentes em relação a cada matéria investigada na região do espectro (SPRING, 2018).

Para obter informações remotamente são utilizados diversos sistemas sensores capazes de detectar a energia eletromagnética refletida ou emitida pelo objeto (MOREIRA, 2012). Dentro dos diversos elementos que podem ser imageados estão os recursos hídricos, vegetação, as diferentes formas de cobertura do terreno como solo, vegetação, rochas e áreas construídas (PONZONI; SHIMABUKURO E KUPLICH, 2015).

Os dados das “imagens” recebidos através dos sistemas sensores passam por um processamento digital, que facilita a identificação de objetos e padrões contidos na imagem, além de auxiliar nos processos de interpretação e de análise dos materiais. É realizado através de um Sistema de Informação Geográfica (SIG) que é um sistema fundamentado em computador que permite a realização de tratamento computacional de dados geográficos,

disponibilizando análises, consultas, modelagem, manipulação e recuperação de dados georreferenciados (DAINESE, 2001).

Uma das principais vantagens de um software SIG é que possibilita o uso dos dados para a elaboração de tabelas, facilitando a elaboração dos mapas (GONÇALVES, 2002).

A ideia de mapas temáticos é utilizada através do conceito de camadas, onde, para uma mesma região podem ser criadas diversas camadas de dados, sendo uma para cada tema de característica de interesse, facilitando a realização de determinadas interpretações e análises dos gerados através dos mapas (LISBOA FILHO, 1995).

Os dados recebidos em formato de imagem digital apresentam uma subdivisão em células de tamanhos iguais formando uma matriz, chamadas de pixels. A tonalidade presente em cada pixel é representada por um número digital (ND) registrado dentro de cada célula. Os sistemas sensores podem ser classificados quanto a sua resolução.

Existem quatro termos relacionados com a resolução de imagens dos sensores remotos, sendo elas: geométricas ou espacial, espectrais, radiométricas e temporal (NOVO, 2008).

A resolução espacial segundo Zani, Duarte e Cruz (2015) é definida com o tamanho de cada pixel: o tamanho que o pixel representa no terreno, assim determinando a qualidade da imagem e os níveis dos detalhes perceptíveis na imagem, ou seja, quanto maior for o detalhamento da resolução espacial do sensor melhor será a detecção de contorno dos objetos.

A resolução espectral é definida como a medida das faixas espectrais e o número de bandas que o sensor pode distinguir (MOREIRA, 2012). Como consequência, sensores com melhor resolução espectral, aumentam o número de bandas disponíveis para investigações das interações entre a radiação eletromagnética e a superfície terrestre (NOVO, 2008).

A resolução radiométrica pode ser descrita como a capacidade do sensor a distinguir variações no nível de energia emitida ou refletida por um alvo. Esta energia possui intensidades diferentes em relação aos níveis de cinza, onde cada pixel é passado por uma codificação digital e transformado em valores numéricos denominados ND (Número Digital) expressos em bits (WENG, 2018; CASTRO, 1992). Assim, quanto maior for a eficácia do sensor em diferenciar

estas intensidades de níveis de cinza maior será sua resolução radiométrica (CASTRO, 1992).

A resolução temporal é definida em relação ao intervalo de tempo que o sensor leva para cobrir um mesmo local na superfície (CASTRO, 1992; NOVO, 2008). Este intervalo de tempo está relacionado com a altura, velocidade e inclinação do satélite e o ângulo de abertura do sensor (NOVO, 2008).

2.1.1 Plataformas e sensores

O sensoriamento remoto aplicado ao estudo da vegetação, aliado a repetitividade dos satélites, torna possível o monitoramento dinâmico da vegetação durante seu ciclo de crescimento e de desenvolvimento (PONZONI; SHIMABUKURO; KUPLICH, 2015) permitindo criar um banco de dados com informações multitemporais (JACKSON e HUETE, 1991). Os dados (imagens) podem ser obtidos a partir de sensores instalados em plataformas, podendo estas serem terrestres, aéreas ou orbitais (LILLESAND, KIEFER e CHIPMAN, 2014).

Em nível orbital os dados são obtidos de sensores instalados em satélites que fazem o imageamento terrestre (NOVO, 2008). Entre os sensores orbitais existentes um dos mais antigos, pertence à série de satélites LANDSAT (*Land Remote Sensing Satellite*) gerenciados pela USGS (*United States Geological Survey*) e NASA (*National Aeronautics and Space Administration*) que oferecem um registro contínuo de imagens terrestres gratuitas desde 1973 (USGS, 2021).

O constante interesse pela obtenção de dados sobre a superfície terrestre através de sensores remotos fez com que mesmo após o lançamento do LANDSAT os avanços tecnológicos sobre novos sensores imageadores continuassem.

Neste sentido, a Agência Espacial Europeia *ESA* (*European Space Agency*) lançou um projeto com base no regulamento da União Europeia de nº 377/2014, onde o Parlamento Europeu e do Conselho criaram em 3 de abril de 2014, o Programa *Copernicus* (COPERNICUS, 2021), o qual vem desenvolvendo um conjunto de missões denominadas Sentinel, em parceria com indústrias europeias e *Eumetsat*, o que possibilitou um enorme avanço no setor (COPERNICUS, 2021).

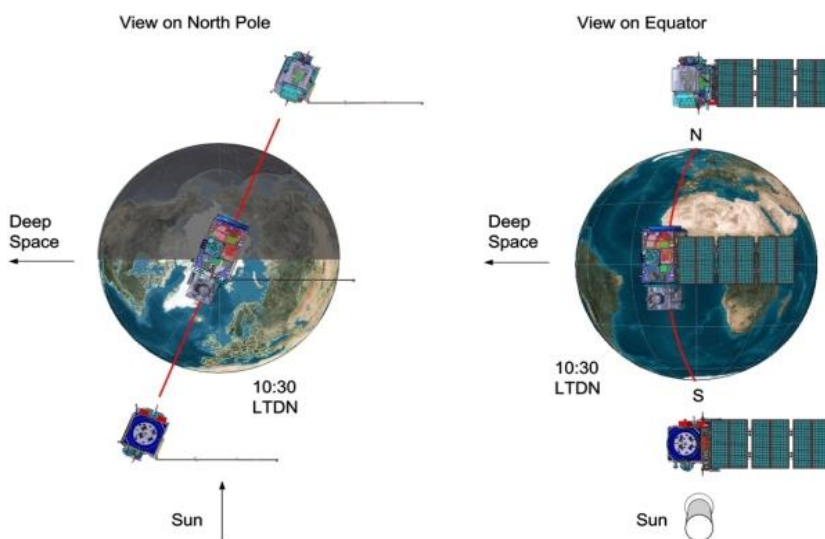
A missão foca-se em temáticas relacionadas com a atmosfera, a criosfera, a hidrosfera e a biosfera. Malenovský et al., (2012) e Torres et al., (2012) citam em seu trabalho as principais características que fazem do programa Sentinel um dos mais eficientes da área, sendo esses pontos:

a. a linha contínua de adesão de imagens com perspectiva para longo prazo que o programa possui;

- b. uma cobertura ao nível global e frequente;
- c. os sensores calibrados com alto nível de acurácia;
- d. o arquivamento e a entrega dos dados cumprindo os rigorosos requisitos de programas operacionais;
- e. uma gama de métodos de detecção ativos, passivos, e por micro-ondas (como o sensor óptico entre outros).

O sistema de satélites Sentinel-2 foi desenvolvido em parceria com a Astrium GmbH da Alemanha e Astrium SAS da França que é a responsável pelo Instrumento MultiSpectral (MSI). Em órbita síncrona solar dois satélites idênticos operam de forma simultânea, a 180° um do outro, em uma altitude média de 786km (Figura 01). A posição de cada satélite é mensurada por um receptor GNSS (*Global Navigation Satellite System*) de dupla frequência. Sua órbita é mantida por um propulsor dedicado somente a isso.

Figura 1. Configuração orbital dos satélites gêmeos Sentinel-2.



Fonte: Sentinel Online, ESA, (2021).

O satélite S-2 busca contribuir de forma significativa para estudos temáticos relacionados com as mudanças climáticas, o monitoramento da superfície da terra, a gestão de emergências como incêndios e a segurança. Com 13 bandas espectrais de 290km de faixa espectral e alta frequência de repassagem, e com o MSI (*MultiSpectral Instrument*) acoplado ao S-2 o tempo de resposta acaba sendo menor e assim pode fornecer dados de alta qualidade e com maior rapidez (ESA, 2021).

O Instrumento MSI funciona de forma passiva, coletando dados referentes a luz solar refletida da Terra, os dados são adquiridos conforme o satélite se move. O feixe de luz que adentra é dividido em um filtro e focado em dois conjuntos de planos focais. Uma para bandas do visível e infravermelho próximo e outra para ondas curtas de infravermelho. Um mecanismo obturador impede a iluminação direta do sol evitando contaminação durante a coleta de dados, o mesmo mecanismo funciona como um dispositivo de calibração (ESA, 2021).

Os instrumentos ópticos dos sensores a bordo do S-2 (Sentinel-2) fornecem informações espectrais suficientes para o mapeamento de radiação tanto da superfície da Terra quando do topo da atmosfera (TOA). Suas imagens espectrais são utilizadas para criação de uma série temporal da vegetação, ou melhor, permitem diagnosticar áreas de queimadas, de vegetação e solo (ESA, 2021), sendo possível a aplicação de diversos índices de vegetação como a Fração da Radiação Absorvida pela Vegetação Fotossinteticamente Ativa (FAPAR), Índice de Área Foliar (IAF) e Índice de Vegetação por Diferença Normalizada (NDVI) (MALENOVSKÝ et al., 2012) .

Com o lançamento da série de satélites Sentinel com seus sensores S-2,-3 e instrumentos ópticos de altíssima qualidade, onde a coleta de imagens com resolução espacial é maior e já corrigidas radiometricamente, é possível realizar uma cobertura terrestre periódica e diagnosticar as mudanças ocorridas em determinado período (ESA, 2021). Com o radar C-band SAR (*Synthetic Aperture RADAR*) totalmente dedicado ao clima, é possível mapear com eficiência águas paradas, inundações, neve, umidade do solo na superfície e medições de biomassa (ciclo do carbono) (ESA, 2021).

Malenovský et al., (2012) aponta conclusões interessantes avaliando o desempenho do satélite S-2 para o mapeamento da cobertura do solo; o mesmo afirma que a produção de mapas temáticos derivados das imagens do satélite é de excelente qualidade em relação aos dados ópticos e de radar do Sentinel que devem ser complementados por estimativas locais para produzir mapas sistemáticos e precisos. Assim, os esquemas de agregação permitirão integrar várias escalas espaciais e temporais de diferentes fontes de dados (SARAIVA, 2015; TORRES et al., 2012).

2.1.2 Índice de Vegetação da Diferença Normalizada (NDVI)

A vegetação apresenta uma característica de absorção solar pelos pigmentos da clorofila que denota baixa refletância na banda do espectro da visível e alta refletância na região do infravermelho próximo (NOVO, 2008). Nesta região de alta refletância é comum aplicarem-se índices para aumentar o contraste entre a vegetação e as diferentes superfícies. O NDVI é mais comumente utilizado para investigar tais assinaturas espectrais, além de diminuir os efeitos que influenciam na resposta espectral das bandas (PONZONI; SHIMABUKURO; KUPLICH, 2015).

Para a obtenção do NDVI utiliza-se a seguinte equação (1):

$$NDVI = \frac{(NIR-RED)}{(NIR+RED)} \quad (1)$$

Onde:

NIR = Reflectância do infravermelho próximo;

RED = Reflectância da faixa do vermelho do espectro do visível.

A utilização do NDVI é comum para elaborar perfis temporais e sazonais da vegetação, facilitando o mapeamento de classe e variações fenológicas da vegetação (DAINESE, 2001).

Os autores Crosta (2002) e Moreira (2012) destacam a importância de se levar em consideração os fatores limitantes relacionados à interpretação dos índices, sendo eles os pontos de saturação que se diferenciam de forma notória nas faixas espectrais do vermelho e infravermelho próximo; também diferenciada nessas duas regiões do espectro está a interferência atmosférica e a largura de cada banda que é variável de acordo com cada sensor. Os valores de NDVI da imagem em uma única banda são escalonados e variam entre -1 a +1. Assim quanto mais claro os tons de cinza forem maiores serão os valores de NDVI, e ao contrário ocorre em tons mais escuros de cinza, apresentando valores mais baixos (PONZONI, SHIMABUKURO E KUPLICH 2015; SHIMABUKURO, NOVO E PONZONI, 1998)

É importante salientar os valores típicos de representação da refletância em relação ao NDVI (PONZONI; SHIMABUKURO E KUPLICH 2015). A presença de nuvens, de água e de neve traz maior refletância no visível, gerando um valor de NDVI negativo. Por outro lado, solos recobertos por aflorações rochosas possuem valores similares nos dois comprimentos de onda, gerando um NDVI próximo a zero (CRÓSTA, 2002).

A vegetação é a que melhor responde ao índice NDVI devido a sua alta refletância no infravermelho, produzindo valores de NDVI entre 0,3 e 1. Assim valores altos de NDVI estão normalmente associados a densidade alta e a vegetação saudável; já valores negativos estão associados a falta de vegetação, como as construções, a água, o asfalto entre outras coisas (SHIMABUKURI; NOVO; PONZONI, 1998; FONTANA, 2001).

Brandão; Bezerra; Silva, (2007) ao utilizarem o NDVI conseguiram estimar o acúmulo de biomassa no município de Barbalha-CE, onde observaram que ao aumentar o NDVI a biomassa acumulada na superfície era incrementada proporcionalmente. Os autores Duarte; Ramos; Maciel, (2015) estimaram a cobertura vegetal avaliando a densidade das áreas de preservação permanente, utilizando o NDVI. Assim como Saraiva, (2015) pode através da utilização do índice identificar as fases fenológicas de culturas anuais, como milho e tomate, onde constatou que ambos possuem comportamentos espectrais parecidos.

Martins, (2017) afirma que a utilização do índice NDVI para cartografar áreas de incêndios na região norte de Portugal é eficiente e considera-o um dos melhores índices para a discriminação de áreas ardidas, permitindo assim diferenciar, caracterizar a vegetação e delimitar locais de queimas, utilizando as técnicas de sensoriamento remoto juntamente com os de processamento digital de imagens.

Em áreas onde a vegetação dispersa, onde apresentam mais solo exposto, Huete (1988) desenvolveu um índice de vegetação que visa reduzir a confusão gerada pelos efeitos de solo descoberto SAVI (*Soil-adjusted Vegetation Index*). O autor fez um ajuste sobre o NDVI acrescentando um fator de correção (L), que busca diminuir a influência do solo quando a cobertura da vegetação é reduzida (HUETE, 1988).

Equação:

$$SAVI = \frac{(NIR-R) \times (1+L)}{(NIR+R+L)} \quad (2)$$

Onde:

SAVI= Soil-adjusted Vegetation Index

NIR= Infravermelho Próximo;

R= Vermelho e;

L= fator de solo ajustado.

Esse fator de ajuste pode assumir valores de 0,25 a 1 dependendo da intensidade de cobertura da área estudada (HUETE, 1988), onde de 0 a 1 para áreas com alta cobertura vegetal e sem cobertura, e o valor de 0,5 é adotado quando ocorre uma densidade intermediária, 0,25 para uma alta densidade de vegetação (HUETE, 1988).

Laurentino, (2019) em seu estudo buscou identificar sobreiros com sinais de estresse e monitoramento da vitalidade dos bosques e obteve resultados satisfatórios ao aplicar o SAVI, devido à estrutura dos bosques que apresentam grandes espaços descobertos. Desta mesma forma, outros autores relatam que modelos como SAVI e NDVI têm as melhores estimativas para biomassa verde, com um coeficiente de determinação ajustado para o solo (LAURENTINO, 2019; LEITE; SANTOS; SANTOS, 2017; LEMENKOVA, 2020).

2.1.3 Classificador Random Forest

A identificação e a classificação automática de imagens para Moreira (2001) compreende a análise e a manipulação de imagens por meio de técnicas computacionais que visam a extração de dados de um objeto do mundo real.

Para a classificação das imagens referentes ao PNM utilizou-se a metodologia Random Forest (RF), por ser uma das mais completas, recentes e precisas, de acesso livre (CONGEDO, 2021). Essa metodologia é baseada na utilização de uma árvore de decisão, onde cada árvore é criada desenhando um subconjunto de amostras de treinamento por meio da substituição (BREIMAN e CUTLER, 2016). Pode-se dizer que o Random Forest (RF) recorre a dois parâmetros principais para a produção das árvores de decisão: o seu número e o de variáveis a serem testadas. Cada árvore é produzida de forma independente, sem qualquer corte, e cada nó é separado usando um número de variáveis pré-definido pelo usuário, e selecionado de forma aleatória. O número de árvores cresce até ao valor definido pelo usuário (PAL, 2005).

Lima e Andrade (2018) em seu trabalho relatam a capacidade que o algoritmo tem de provocar mudanças em uma variável ao alterar o erro de predição durante a execução do mesmo. Uma das grandes vantagens do Random Forest é a sua capacidade de trabalhar com diversas variáveis de entrada sem que se perca nenhuma variável precedente (PAL, 2005; BELGIU e DRĂGUȚ, 2016).

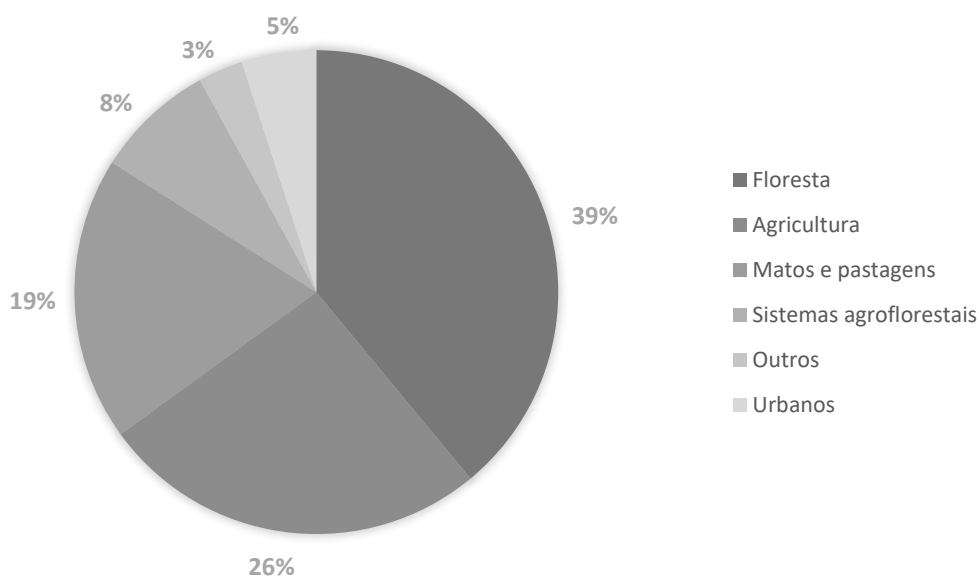
Diferentes trabalhos foram produzidos na área do PNM, buscando a identificação da vegetação e caracterização da paisagem em relação ao seu contexto histórico e cultural (CASTRO, 2010; DIAS e AZEVEDO, 2008; MAGNONI, 2017). Benito Garzón et al., (2008) utilizou do algoritmo RF para prever a distribuição atual e futura de espécies da Península Ibérica, além de avaliar como as mudanças climáticas poderão afetar na distribuição dessas espécies no futuro, o mesmo ainda conclui que o *Quercus pyrenaica* sofrerá um declínio notável até 2080.

Entretanto, são ainda necessárias mais pesquisas com técnicas de detecção remota para identificação de espécies de valor económico e cultural para a população.

2.2 A Floresta em Portugal e no PNM

A As florestas autóctones portuguesas têm um papel importante na sociedade. Além de serem um patrimônio natural de grande valor económico, socioeconómico e ecológico para o país, também representam valores paisagísticos e de recreação, assim como constituem um valor primordial relacionado com a proteção do meio ambiente e a conservação da vida selvagem, e de toda a biodiversidade encontrada dentro de um grande mosaico paisagístico (ICNF, 2019). Cerca de 40% do território nacional é composto por florestas (Gráfico 1) podendo essas ser de monoculturas exóticas como as de eucalipto, pinheiro-bravo, e castanheiro, invasoras como as de acácia, ou de nativas como os montados de sobreiro, de azinheira, e outras espécies que compõem a sua diversidade florística (ICNF, 2021).

Gráfico 1. Uso e Ocupação do Solo de 2018 em Portugal Continental.



Fonte: Adaptado pelo autor, Carta de Uso e Ocupação do Solo de Portugal Continental 2018 (COS2018), Direção-Geral do Território e Relatório Completo do 6.º Inventário Florestal Nacional ICNF.

A biocenose presente no Parque Natural de Montesinho sobressai pelas suas complexas condições climáticas (COSTA et al., 1990) e a sua rara litologia (como os afloramentos ultrabásicos), o que proporciona uma variedade de vegetação edafoixerófila e de bosques climatófilos (AGUIAR, 2000), para além

do seu particular cenário de transição ambiental entre o bioma mediterrânico e o eurosiberiano. Os solos dos afloramentos ultrabásicos na região transmontana, devido ao seu baixo teor de sílica, tornam-se tóxicos para grande parte das plantas, o que induz a ocorrência de espécies edafoxerófilas / sobre formações ultrabásicas (COSTA et al., 1990).

Os autores Sequeira; Aguiar; Meireles, (2010) listam em seu trabalho espécies endêmicas e outras plantas raras encontradas no Sector Orensano-Sanabriense de Portugal, sendo estas, *Alyssum pintodasilvae*, *Anthyllis sampaioana*, *Arenaria querioides* subsp. *fontqueri*, *Aemeria eriophylla*, *Avenula pratensis* subsp. *lusitânica* e *Festuca brigantina*. Aguiar (2000) encontrou populações de *Linaria coutinhoi* a leste do PNM um dos mais raros táxons portugueses. Estas características fazem com que o parque se destaque ao nível nacional devido a sua elevada diversidade de espécies raras e endêmicas preservadas no PNM.

Algumas das florestas da região transmontana se desenvolveram em solos ultrabásicos, com substrato rochoso que dificulta o crescimento de vegetação arbórea, o que acabou por controlar a distribuição dos bosques arbóreos da região (COSTA et al., 1990), bosques climatofilos em condições edafoxerófilas de *Genisto hystricis-Quercetum rotundifoliae* sendo das formações arbóreas mais ameaçadas pela ação do fogo.

Em Portugal, grande parte da floresta natural desapareceu ou se encontra muito alterada por ações antrópicas (QUERCUS, 2021). Tal está relacionado aos incêndios recorrentes na região mediterrânica, a forma como a vegetação responde a este ataque, sendo que cada espécie apresenta diferentes tipos de sensibilidade ao fogo. Por norma, as florestas autóctones portuguesas são constituídas por diferentes espécies com menor poder de combustão, resistentes ao fogo (ICNF,2021). É importante salientar sua resiliência à ação do fogo permite que os tecidos vivos (cambio e floema) não sejam afetados pelo efeito do fogo, conseguindo se regenerar após a ocorrência de incêndios, seja por rebentos a partir do caule ou raiz, bancos de sementes, em outras situações como é o caso do pinheiro-bravo que se adaptou ao longo do tempo aos incêndios constantes, desenvolvendo uma casca robusta com uma espessura elevada superior a outras espécies (QUERCUS, 2021).

Estudos realizados pelos autores Azevedo e Caçador (1999) discutem a presença de bosques maduros de Azinheiras mesmo com a incidência de incêndios frequentes na região, o mesmo sugere que a espécie criou seu próprio mecanismos de defesa, uma das teorias abordadas é a qual o processo de combustão das copas das árvores, que ao se darem de forma efêmera, dissipam-se de forma rápida, não adentrando para o interior dos bosques. Azevedo et al., (2013) em sua pesquisa mais recente confirma esta hipótese de defesa dos bosques de azinheira, onde o fogo acaba por ser controlado nas bordaduras, em razão da repentina mudança na estrutura da vegetação adjacente causando a diminuição da intensidade do fogo.

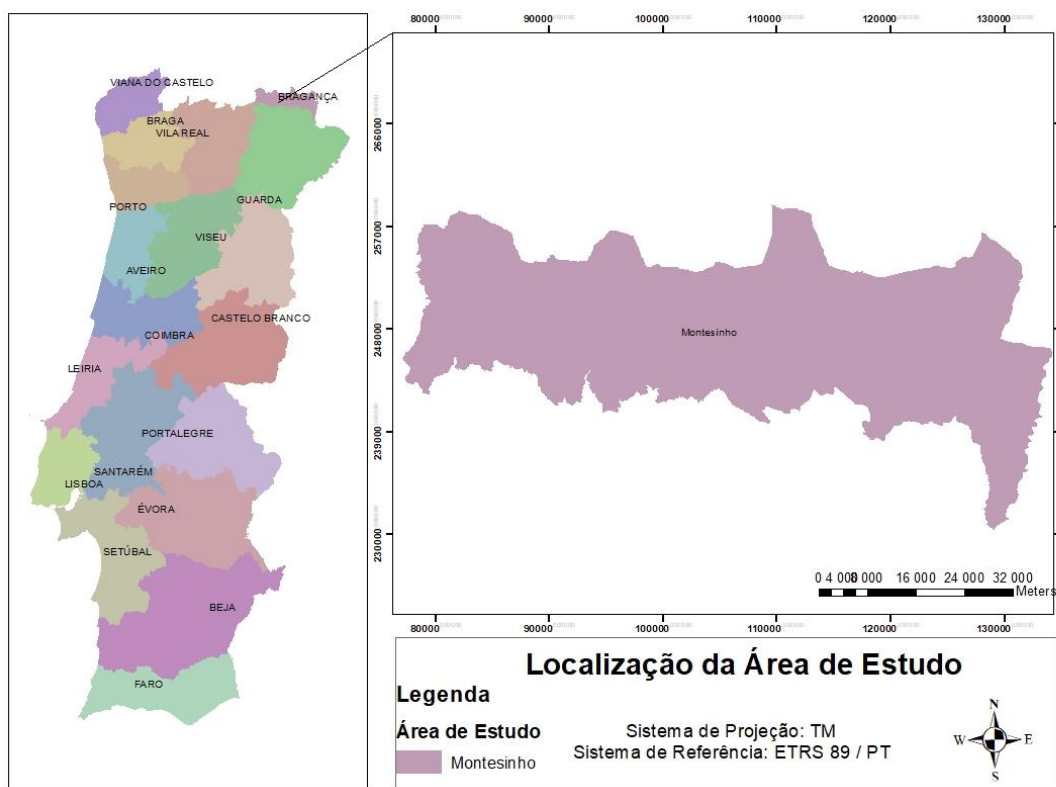
Uma das estratégias interessantes para diminuir a intensidade dos incêndios dentro dos bosques, é a utilização da atividade silvo pastoril. Alguns resultados obtidos na região de Bragança se mostram promissores: os autores Castro, Castro; Sal (2009) e Castro (2008) discutem a relevância da redução do combustível fino como um aliado ao combate aos incêndios florestais.

3 MATERIAIS E MÉTODOS

3.1 Área de Estudo

A área de estudo – Parque Natural de Montesinho – está localizada no nordeste de Portugal Continental, na remota região de Trás-os-Montes (6° 30' 53" O e 41° 43' 47" S), abrangendo parte dos municípios de Bragança e Vinhais, fazendo fronteira com a Espanha (Figura 02). Possui uma área de 748 km² constituída de inúmeras elevações e vales, o que faz com que a sua altitude varie entre 438 m e 1481 m, com uma pedologia predominante com origem em rochas básicas e ultrabásicas (AGUIAR e VILA-VIÇOSA, 2017).

Figura 2. Mapa de localização do Parque Natural de Montesinho Trás-os-Montes em Portugal continental.



Fonte: A autora, (2021).

O Parque Natural de Montesinho (PNM) foi criado em 30 de agosto de 1979 sob o Decreto-Lei nº 355/79, objetivando a proteção dos seus recursos naturais, tendo em vista que o parque possui um dos mais elevados índices de biodiversidade a nível nacional. O PNM é rico em avifauna, com a presença de espécies ameaçadas de extinção ainda em relativa abundância, além de vastas áreas de vegetação natural.

3.2 Seleção e processamento das imagens

Todo o processamento dos dados foi realizado no software QGIS 3.18.1 utilizando o SCP (Semi-Automatic Classification Plugin), um complemento que permite a classificação semi-automática ou supervisionada das imagens de satélites.

O tratamento das imagens acaba por ser uma tarefa imperceptível visualmente, porém essencial quando se trabalha com detecção remota, esta fase acrescenta maior credibilidade para se trabalhar com tais dados.

As imagens de satélite utilizadas para a produção deste trabalho são advindas do sensor Sentinel-2 que possui bandas com resolução espacial de 10,20 e 60 metros conforme expresso na Tabela 1. Utilizou-se como “ground truth”, a Carta de Uso e Ocupação do Solo de Portugal (COS 2018), juntamente com o Inventário Florestal Nacional de 2015, informação disponível no site do DGT (Diretório Geral do Território) e ICNF (Instituto da Conservação da Natureza e das Florestas).

Tabela 1. Número de bandas e resolução espacial de cada banda do Sensor Sentinel-2

Número de Bandas	Comprimento de Onda Central (nm)	Resolução Espacial (m)
B1 (Costal Aerossol)	442,7	60
B2 (Azul)	492,4	10
B3 (Verde)	559,8	10
B4 (Vermelho)	664,6	10
B5 (Red Edge 1)	704,1	20
B6 (Red Edge 2)	740,5	20
B7 (Red Edge 3)	782,8	20
B8 (Infravermelho)	832,8	10
B8A (Red Edge 4)	864,7	20
B9 (Vapor D'água)	945,1	60
B10 (Cirrus)	1373,5	60
B11 (SWIR 1)	1613,7	20
B12 (SWIR 2)	2202,4	20

Fonte: adaptado pelo autor de ESA Copernicus, 2021.

A correção radiométrica visa o ajuste das interações atmosféricas presentes na imagem, esta correção deve ser realizada quando se deseja analisar imagens obtidas em datas diferentes de passagens do satélite, e utiliza uma das passagens como referência (MOREIRA, 2012). O processo de retificação radiométrica relaciona as mudanças observadas no comportamento do espaço-tempo e espectro com a interação da radiação eletromagnética com o alvo, com o objetivo de eliminar quaisquer relações atmosféricas na reflectância do alvo (CENTENO, 2010).

Novo (2008), Ponzoni, Shimabukuro e Kuplich (2015) citam que se faz necessária a introdução de dados referentes às condições atmosféricas da data em que as imagens foram adquiridas, transformando os números digitais em radiância espectral, e aplicam-se equações lineares de calibração para realizar a conversão nos números digitais.

A fase de correção foi realizada apenas nos produtos L1C (Tabela 2) referentes ao mês de fevereiro e julho diretamente no plugin SCP, onde os dados necessários para o tratamento de cada imagem são inseridos e processados de forma automática gerando as correções atmosféricas. O mesmo tratamento não

foi realizado no produto L2A, devido a este tipo já ser corrigido pelo algoritmo Sen2Cor (ESA, 2021). Esta é considerada com uma etapa fundamental quando se deseja realizar estudo de series temporais.

Tabela 2. Imagens utilizadas para a produção dos mapas de interesse.

Época do ano	Imagem
19 de fevereiro de 2020	L1C_T29TPG_A024344_20200219T112435_2020-02-19
18 de julho de 2020	L1C_T29TPG_A026489_20200718T112448_2020-07-18
11 de outubro de 2020	L2A_T29TPG_A018796_20201011T112135_2020-10-11

Fonte: Sentinel2 Copernicus, (2021).

Dentro do band set foram inseridas as bandas já corrigidas que serviram como referência para a retirada das leituras espectrais e a criação da banda RGB que serviu como base visual das áreas de interesse.

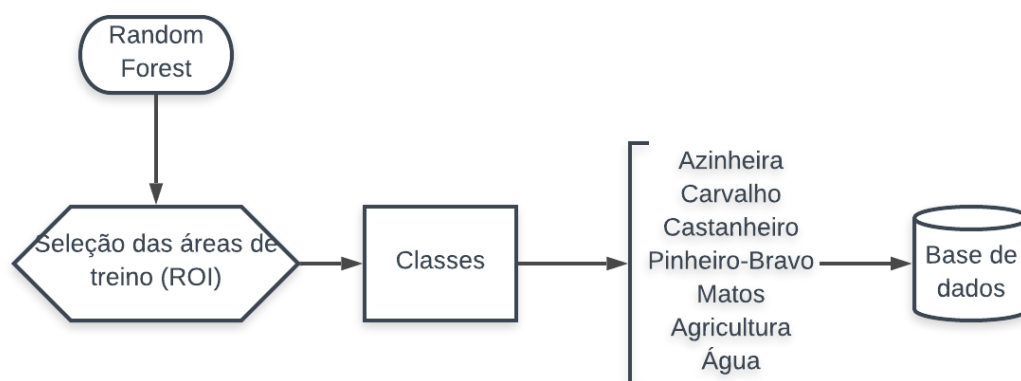
Foi produzida a análise dos índices de vegetação através da ferramenta Vegetation index [distance based] no QGIS 13.18.1. Os resultados obtidos foram utilizados para complementar o mapeamento das classes e as suas variações fenológicas.

Os índices utilizados foram o NDVI e SAVI, conforme expresso nas equações 3 e 4. Para isso ser possível, foi utilizado um fator de ajuste de 0,25 para a área de estudo, pois o PNM apresenta grande densidade de volume de copa ao se levar em consideração as espécies nativas perenifólias.

3.3 Tratamento e Pós Processamento das imagens

Para a criação das amostras de treinamento (ROI- Regions of Interest) foi utilizado como material de apoio as cartas de uso e ocupação do solo (COS) e o inventário florestal de 2015, disponibilizados pelo ICNF, onde foi possível identificar a localização remota dos bosques das espécies estudadas. A identificação das áreas foi baseada na semelhança dos pixels (Region Growing Algorithm), ferramenta disponível dentro do plugin do SCP que identifica a similaridade espectral de cada pixel a partir de um pixel central. Após a identificação das áreas foi desenhado um polígono de manualmente e criadas as classes (Figura 3) que compuseram a base de dados para a classificação supervisionada. Os polígonos foram feitos de forma manual devido à compatibilidade do SNAP (Sentinel Application Platform) com o plugin (CONGEDO, 2021); o region growin foi utilizada apenas para identificar os bosques mais homogêneos.

Figura 3. Fluxograma do processamento do método de classificação Random Forest. referente ao mês de julho.



Fonte: A autora, 2021.

As imagens foram adquiridas em meses diferentes, o que pôde reduzir a precisão de classificação do plugin. Pois, as mudanças na cobertura vegetal que se observa consoante os meses afetam a assinatura espectral: um mesmo pixel tem assinatura espectral diferente em períodos diferentes, assim como os efeitos atmosféricos afetam as imagens de maneira diferentes. Sendo assim, Congedo, (2021) sugere que se colete sempre novas ROIs e assinaturas espectrais para

cada imagem (PRATES; LACERDA, 2019). À vista disso, foi coletado novas assinaturas a cada imagem utilizada.

Na fase de processamento para gerar as classificações RF, como alguns dos polígonos abrangiam mais de 5.000 pixels optou-se por utilizar um número de amostras de treinamento de 10.000 usados aleatoriamente para delinear o modelo. Utilizou-se um número de árvores de decisão = 100, devido ao tempo de processamento e à capacidade do computador de processar os dados.

- Análise de exatidão

Ao trabalhar com detecção remota é recomendado verificar os resultados obtidos com a verdade do terreno, ou seja, com a interpretação de imagens de dados georreferenciados reais (CONGALTON, 2001). Segundo Anand (2012) a associação que existe entre a imagem classificada e uma referência pode ser expressa através de uma matriz de confusão e o coeficiente *Kappa*. Após a verificação de campo é construída uma Matriz de Confusão (ou de Erro, ou de Contingência), registando-se os resultados da verificação no campo em colunas e os resultados da interpretação das imagens em linhas. Os valores da diagonal desta matriz quadrada representam a concordância entre os resultados da interpretação e da realidade observada.

Tabela 3 Representação da Matriz de Confusão.

		Dados de Referência			Total	Precisão do Utilizador (%)	Erro de Omissão (%)
		1	2	x			
Dados Classificados	1	n_{11}	n_{12}	n_{1x}	n_{1+}	$(n_{11}/n_{1+}) * 100$	$(n_{1+}-n_{11}) / n_{1+} * 100$
	2	n_{21}	n_{22}	n_{2x}	n_{2+}	$(n_{22}/n_{2+}) * 100$	$(n_{2+}-n_{22}) / n_{2+} * 100$
	x	n_{x1}	n_{x2}	n_{xx}	n_{x+}	$(n_{xx}/n_{x+}) * 100$	$(n_{x+}-n_{xx}) / n_{x+} * 100$
	Total	n_{+1}	n_{+2}	n_{+x}	n		
Precisão do Produtor		$(n_{11}/n_{+1}) * 100$	$(n_{22}/n_{+2}) * 100$	$(n_{xx}/n_{+x}) * 100$			
Erro de Comissão (%)		$(n_{+1}-n_{11}) / n_{+1} * 100$	$(n_{+2}-n_{22}) / n_{+2} * 100$	$(n_{+x}-n_{xx}) / n_{+x} * 100$			
Exatidão Global		$(n_{+1}-n_{11}) / (n_{+1}) * 100$	$(n_{+2}-n_{22}) / (n_{+2}) * 100$	$(n_{+x}-n_{xx}) / (n_{+x}) * 100$			

Fonte: Adaptado pelo autor de Congalton e Green (1999).

Exatidão Global (overall accuracy) está relacionada com a divisão numérica total de pixel corretos (diagonal) pelo número total de pixels da matriz de erro, sendo este expresso pela Equação 2:

$$EG = \frac{A}{n} \times 100 \quad (02)$$

Onde:

EG: Exatidão Global;

A: compreende ao número de pontos amostrais com acerto; e

n: ao número de pontos amostrais.

- Coeficiente *Kappa*

Técnica usada para avaliar a precisão da classificação de uma matriz. Melhora a concordância existente entre os dados classificados e os dados de referência (ANAND, 2012). O coeficiente é calculado com base na Equação 03.

$$K = \frac{[n \times \sum_{i=1}^r X_{ij} - \sum_{i=1}^r (X_i \times X_j)]}{[n^2 - \sum_{i=1}^r (X_i \times X_j)]} \quad (03)$$

Onde:

K: é o coeficiente de concordância *Kappa*;

r: o número de linhas na matriz;

i: número de observações;

j: número de colunas; e

n: número total de observações.

Com base no valor obtido através da equação supracitada o desempenho da classificação é medido de consoante a tabela 4.

Tabela 4. Qualidade da classificação associada aos valores do Índice *Kappa*.

COEFICIENTE KAPPA	DESEMPENHO
<0%	Muito Fraco
0% - 20%	Fraco
21% - 40%	Razoável
41% - 60%	Moderado
61% - 80%	Considerável
81% - 100%	Muito bom

Fonte: Adaptado pelo autor de Landis e Koch, (1977).

Como fonte de validação da classificação optou-se por utilizar o 6.º Inventário Florestal Nacional (IFN6) do ano de 2015 adquiridos no site do ICNF e gerado uma malha de pontos no Qgis correlacionado os pontos com a classificação do RF, NDVI e SAVI. Para a criação da matriz de pontos em formato csv, juntaram-se os pontos do IFN com os feitos pelo classificador, NDVI e SAVI para a extração dos valores médios de cada espécie de interesse para este estudo.

Assim foi possível a construção da matriz de confusão, no programa Microsoft Excel conforme a metodologia de Congalton e Green (1999), para calcular os parâmetros de acurácia para a classificação. Após os dados organizados foram gerados gráficos e tabelas dinâmicas para comparação dos valores obtidos.

4 RESULTADOS E DISCUSSÃO

4.1 Índices de Vegetação

É exibido na Figura 4, o mapa das imagens processadas referentes à área de estudo com o mapa do NDVI correspondente ao dia 19 de fevereiro de 2020 denominado “NDVI 02”. Pode-se verificar na legenda que os valores mais baixos estão associados aos tons de vermelho e os valores mais altos aos tons de verde.

Os maiores valores de NDVI, compreendidos no intervalo de 0,66 a 0,91, estão relacionados às áreas de vegetação com maior vigor, representados pelas áreas de floresta mais densa, nomeadamente as dos bosques de pinheiro-bravo, de carvalho e de outros matos altos que compõe o PNM. O valor negativo compreende a barragem de montesinho e demais zonas de água, seguidas por áreas de solo exposto.

Em relação ao mês de julho, que compreende a estação mais verde do ano, o verão, a vegetação apresenta uma alta produção fotossintética, como pode ser observado na Figura 5. Já no período do outono podemos observar uma baixa atividade fotossintética da vegetação, analisando a Figura 6, nota-se um aumento nas áreas de solo exposto, devido ao processo de desfolha da vegetação para a entrada do inverno.

Figura 4. Índice de vegetação por diferença normalizada (NDVI) da área de estudo, referente ao mês de fevereiro.

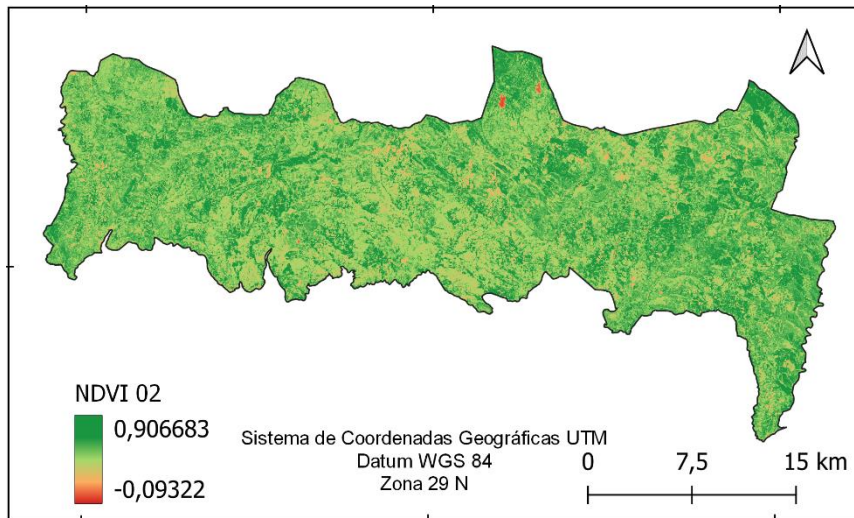


Figura 5. Índice de vegetação por diferença normalizada (NDVI) da área de estudo, referente ao mês de julho.

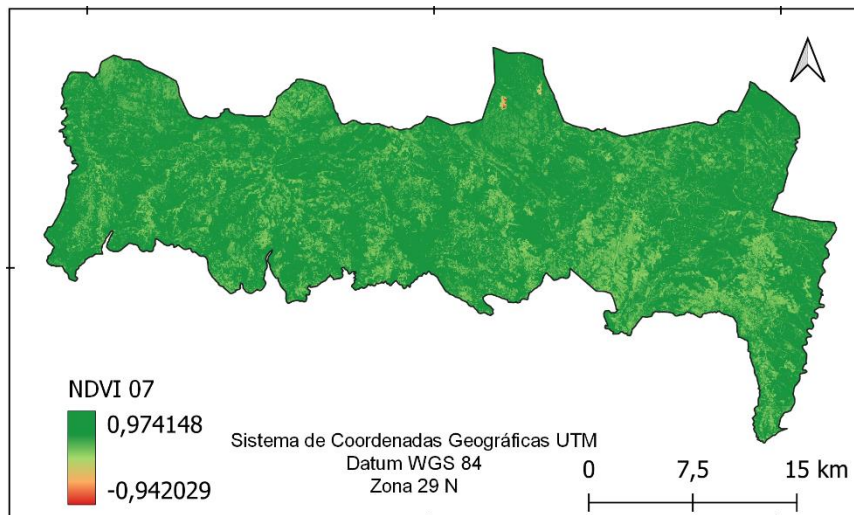
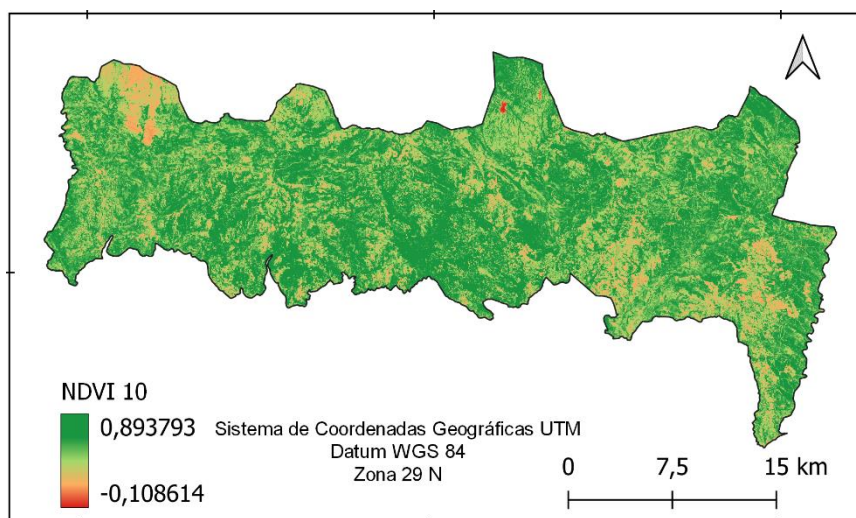


Figura 6. Índice de vegetação por diferença normalizada (NDVI) da área de estudo, referente ao mês de outubro.

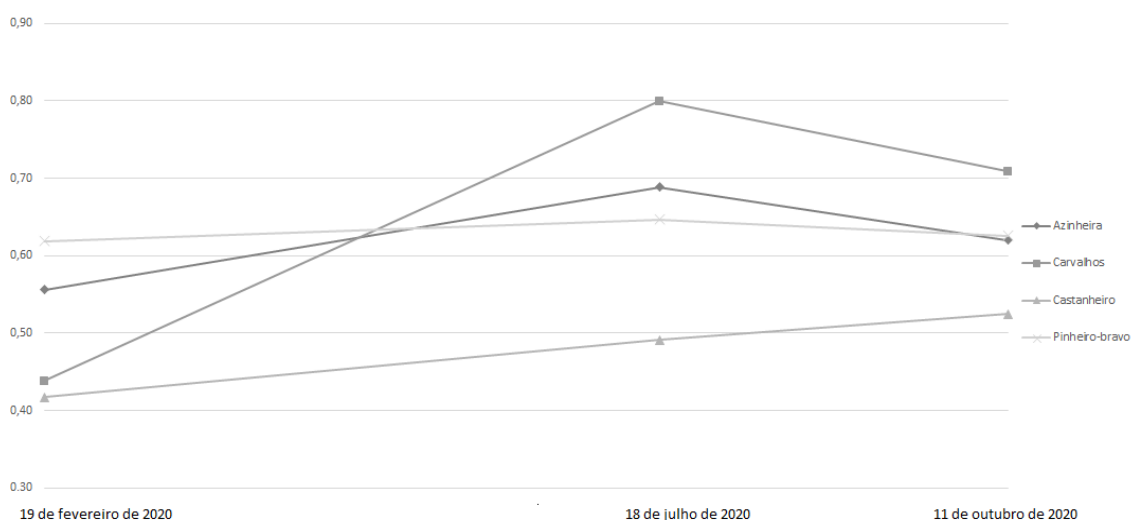


Fonte: A autora, 2021.

Ao compararmos o índice com as principais espécies de interesse da região (azinheira, carvalho, castanheiro e pinheiro-bravo), em relação às épocas do ano (Gráfico 2), podemos notar o pico acentuado dos bosques de carvalho no mês de julho, atingindo um valor médio de NDVI = 0,80. O que pode ser explicado pela maior produção de clorofila implicando uma alta atividade fotossintética da vegetação no verão.

Gráfico 2. Valores médios do NDVI relacionado a época do ano com base nas espécies de interesse.

Fonte: a autora, 2021.

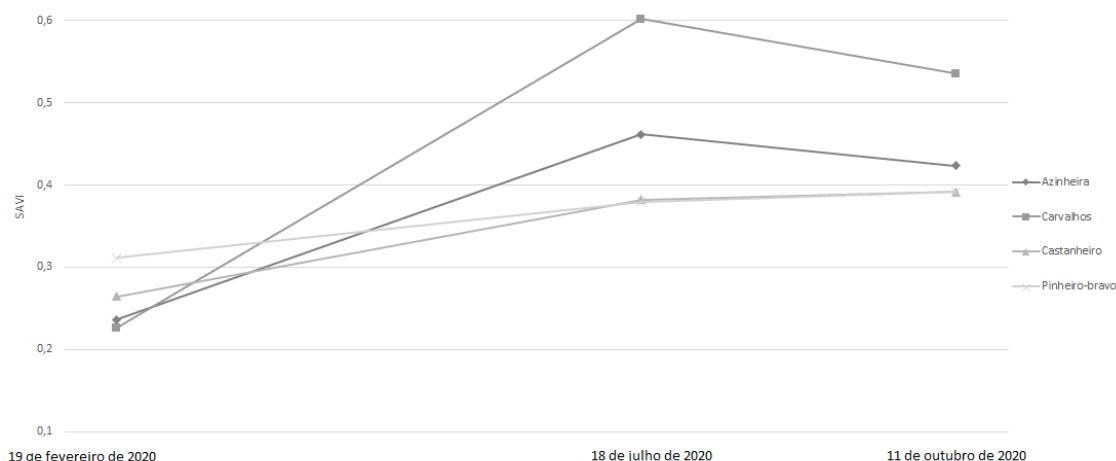


A azinheira, com um valor médio NDVI = de 0,69 no mês de julho e de 0,56 no mês de fevereiro, mantém-se estável ao longo de todo o ano, o que pode ser justificado por ter uma folhagem perene. O mesmo já não é notado no carvalho, que possui folhas caducas, ou seja, perde suas folhas no período de frio intenso, apresentando um valor de NDVI de 0,44 no mês de fevereiro. O comportamento do pinheiro-bravo é muito semelhante ao da azinheira em relação a sua estabilidade foliar ao longo do ano, apresentando um valor médio de 0,63.

Testa e Borgogno Mondino (2013) em seu trabalho trabalham com valores médios para povoamentos puros de castanheiro na região de Piemonte, onde os valores de NDVI mais baixos variaram de 0,4 para os meses de inverno e 0,9 nos meses de verão. Não foi encontrado valores altos para os bosques de castanheiro, pode-se atribuir a dimensão e homogeneidade dos mesmos.

Para o índice de vegetação ajustado sob o solo SAVI, obtiveram-se valores interessantes para a discussão, quando levamos em consideração as épocas de inverno e de outono onde a vegetação entra em repouso vegetativo, perdendo grande parte da sua área foliar (Gráfico 3).

Gráfico 3. Relação dos valores médios de SAVI com os meses do ano analisados, com base nas espécies de interesse.



Fonte: A autora, 2021.

Quando comparamos o índice em relação à data da coleta das imagens podemos identificar os valores médios de cada espécie de interesse, onde a azinheira, o carvalho e o castanheiro apresentaram um SAVI de 0,2 para o mês de fevereiro, enquanto o pinheiro-bravo apresentou um índice de 0,3.

No mês de julho a espécie que obteve o maior valor foi o carvalho (0,6) seguido pela azinheira (0,5) o pinheiro-bravo e castanheiro com o mesmo valor de 0,4.

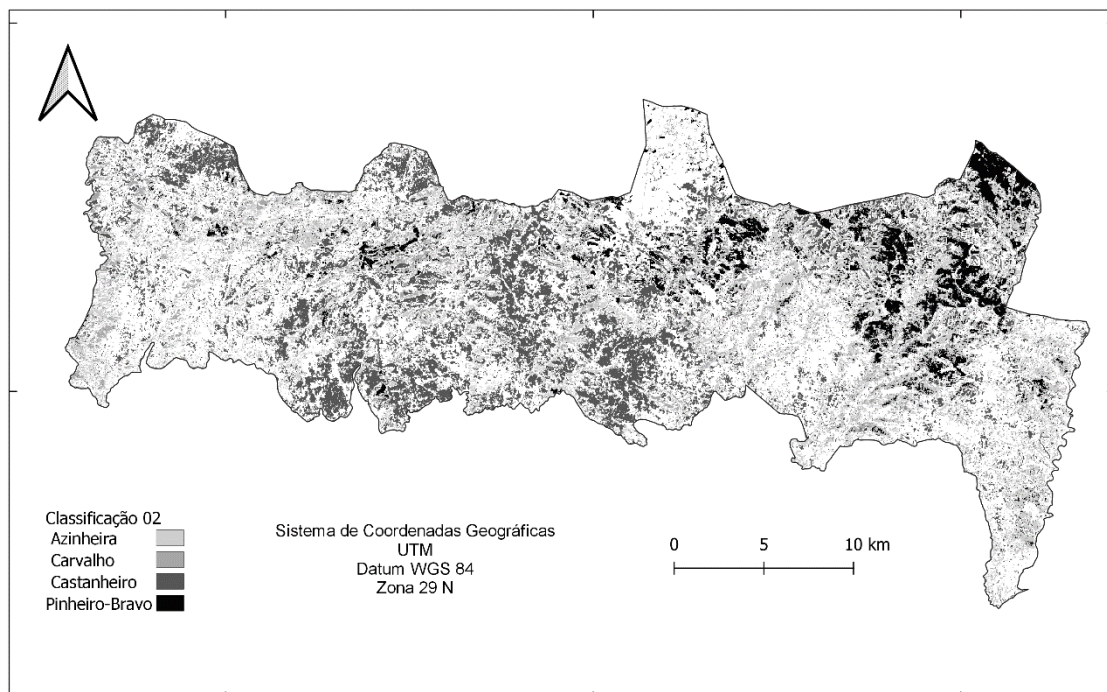
Em outubro as espécies de pinheiro-bravo, de castanheiro e de azinheira apresentaram um valor médio de 0,4 e o Carvalho de 0,5.

Laurentino (2019) em seu trabalho relata que os melhores índices para a caracterização dos sobreiros são o NDVI e EVI o mesmo ainda afirma que a utilização de imagens Sentinel-2 não são aconselháveis para a caracterização, pois a resolução de 10m acaba englobando uma maior área de solo exposto e assim, influenciando nos valores obtidos pelo SAVI. Optou-se pela utilização do mesmo, para ter um melhor entendimento da região.

4.2 Classificação por Random Forest

A partir dos resultados do método de classificação supervisionada do PNM, a avaliação de precisão e o Índice *Kappa*, por meio de uma matriz de confusão. Quando os pontos da matriz que apresentam discordância entre os dados reais e os classificados, percebe-se que a margem de erro de classificação se concentrou nas bordas dos povoamentos, em locais onde há similaridade dos pixels dentro das diferentes classes, ou em áreas muito heterogêneas (CONGEDO, 2021). Este tipo de erro já era esperado pois, quando se tem uma mistura entre as classes de vegetação, e como o objetivo inicial do trabalho é identificar bosques de azinheira, carvalho, castanheiro e Pinheiro-bravo, que são espécies que acabam por se misturar no mosaico da paisagem do PNM.

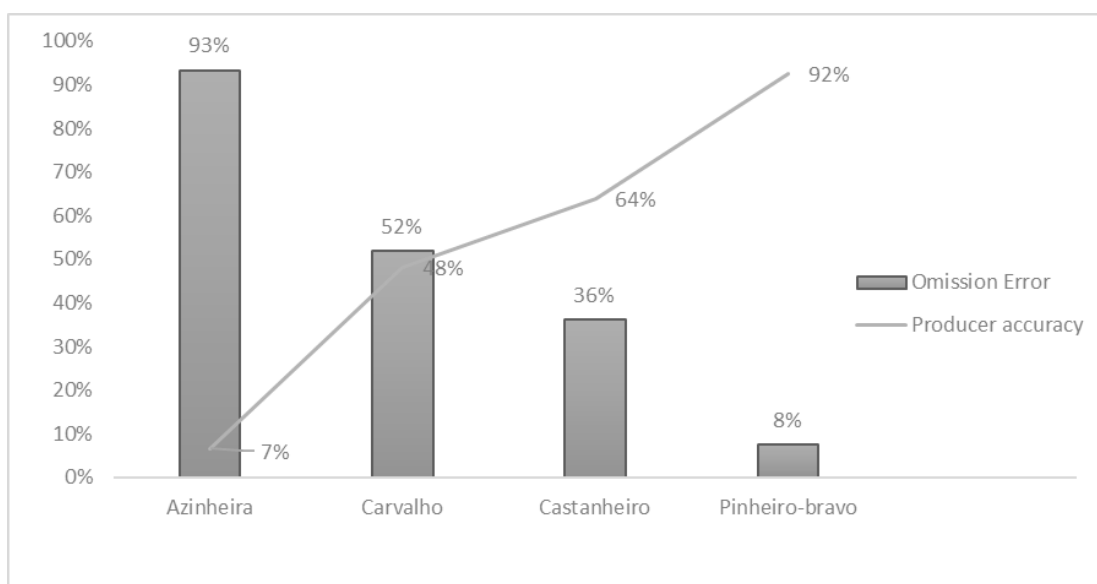
Figura 7. Mapa temático da classificação supervisionada do Parque Natural de Montesinho referente ao mês fevereiro de 2020.



Fonte: A autora, 2021.

A Figura 7 apresenta o mapa temático para fins de interpretação visual da classificação feita pelo RF (Random Forest) da área estudada no mês de fevereiro do ano de 2020 na estação fria do ano, para a produção do mapa foi introduzido junto ao band set as variáveis de NDVI para aumentar a precisão do classificador. Feng, Liu e Gong (2015) em sua pesquisa demonstram que o RF se supera quando comparado com outro classificador, o mesmo é discutido por Alegria, Cristina et al. (2021) em seu trabalho avaliando a distribuição de *Pinus pinaster Aiton*. O Gráfico 4 representa os parâmetros de Erro de Omissão e Acurácia do Produtor, utilizados para investigar o desempenho da classificação em relação a cada uma das espécies de interesse.

Gráfico 4. Análise de acurácia das amostras, referentes ao mês de fevereiro de 2020.



Fonte: A autora, 2021.

A classe de azinheira não foi bem classificada neste mês, com um erro de omissão de 93%, este índice indica a confiabilidade da amostra, ou seja, a probabilidade de uma classe de pixels no mapa representar a categoria da amostra coletada. Ao analisarmos o quão bem a área foi classificada o valor acaba por ser baixo de 7% (Gráfico 4).

A classe do Carvalho apresentou valores significativos em relação ao erro de omissão de 52% e a acurácia do produtor de 48%. O que significa que esta classe foi bem representada.

O Castanheiro e o pinheiro-bravo foram os mais bem classificados neste mês considerando os parâmetros de erro de omissão e a acurácia do produtor, apresentando valores de 36% para o castanheiro e 8% para o pinheiro-bravo referentes ao erro de omissão e 64% e 92 % para a acurácia do produtor.

Utilizando os dados de validação cruzada para a aferição da acurácia do mapa temático (Tabela 4), obteve-se um valor de 56 % para a Exatidão Global (Overall Accuracy) onde, 44 % cerca de dos pontos de amostragem foram classificados errados, equivalendo a um índice Kappa de 41%.

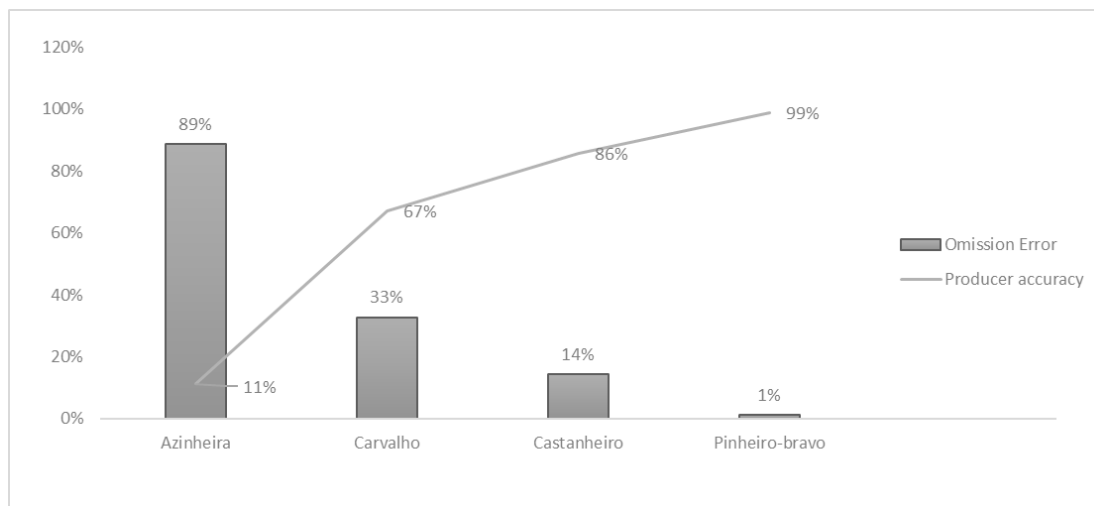
Tabela 4. Matriz de confusão para a análise da acurácia obtida na classificação supervisionada Random Forest, com informações de acurácia e erros de produtor e usuário, do mês de fevereiro de 2020.

Dados de Validação (IFN)							
Dados Classificados (RF)	Azinhreira	Carvalho	Castanheiro	Pinheiro Bravo	Total Geral	Precisão do Utilizador (%)	Erro de Omissão (%)
Azinhreira	6	5		3	14	57%	43%
Carvalho	45	25	65	4	139	82%	18%
Castanheiro	4	9	122		135	10%	90%
Pinheiro-bravo	35	13	4	86	138	38%	62%
Total	90	52	191	93	426		
Precisão do produtor	93%	52%	36%	8%			
Erro de comissão (%)	7%	48%	64%	92%			
Exatidão global	56%						
Coeficiente <i>Kappa</i>	41%						

Fonte: Autora, 2021.

Na classificação RF para o mês de julho a classe de Pinheiro-bravo foi a melhor classificada, com um valor de 99% referente a acurácia do produtor e de 1% de erro de omissão. Seguido pelo Castanheiro e Carvalho, podendo ser observado no Gráfico 5.

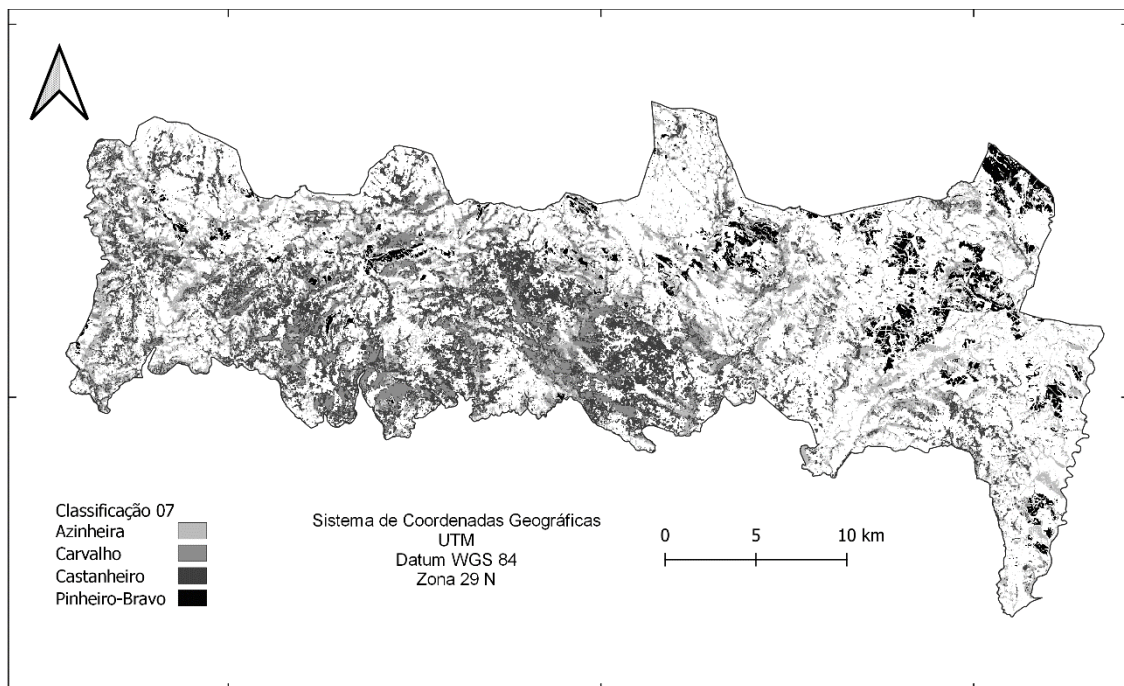
Gráfico 5. Análise da acurácia das amostras, referentes ao mês de junho de 2020.



Fonte: A autora, 2021.

Este mês foi o que melhor representou classes, em relação ao classificado e ao real (IFN), a Figura 8 apresenta o mapa temático para uma melhor interpretação visual da classificação RF do mês de julho de 2020. Onde pode-se observar que a classe de pinheiro ficou muito bem distribuída quando comparada com o real.

Figura 8. Mapa temático da classificação supervisionada do PNM referente ao mês de julho de 2020.

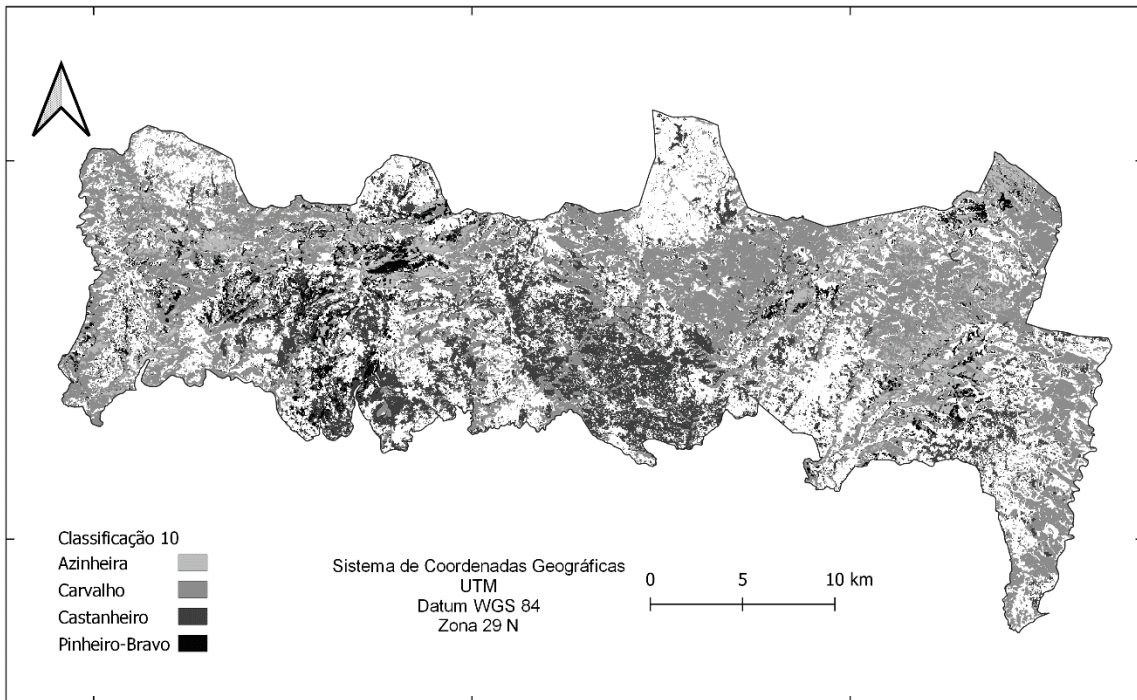


Fonte: A autora, 2021.

Nota-se que as classes da azinheira e do carvalho são as mais problemáticas, apresentando um erro de omissão muito maior que as demais. A validação dos dados através da matriz de confusão está expressa na Tabela 5, onde se obteve um valor de 77 % para o Overall Accuracy e um índice Kappa de 63 % que é considerado, segundo a literatura, um valor considerável (LANDIS; KOCH, 1977).

Ao analisarmos a classificação RF do mês de outubro (Figura 9), que coincide com a estação outonal, onde o carvalho e castanheiro perdem suas folhas para a chegada do inverno e a intensidade da atividade fotossintética da Azinheira e Pinheiro-bravo esta baixa.

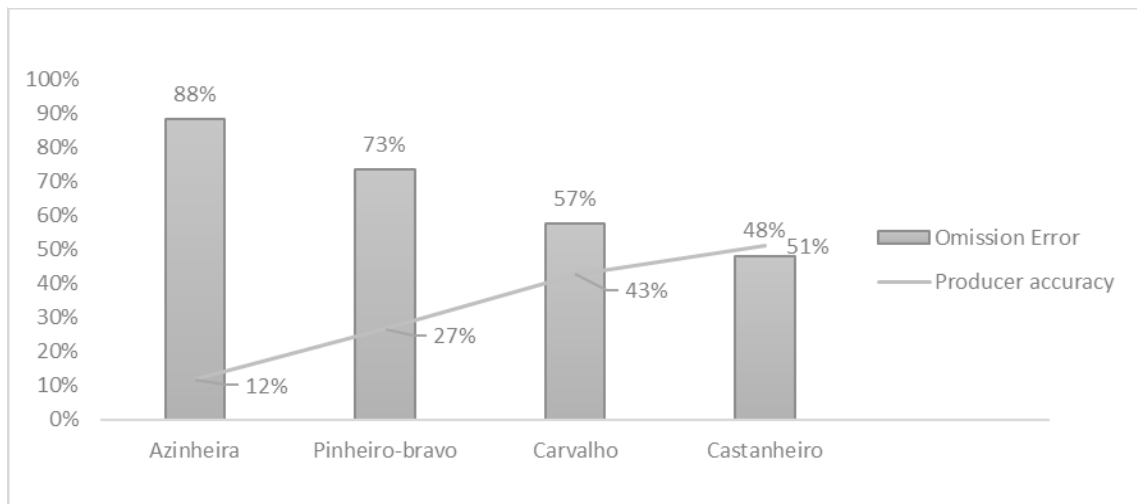
Figura 9. Classificação supervisionada RF do Parque Natural de Montesinho referente ao mês de outubro de 2020.



Fonte: A autora, 2021.

Quando comparamos os valores do erro de omissão e acurácia do produtor (Gráfico 6) todas as espécies analisadas apresentam resultados insatisfatórios correlacionando com a classificação gerada e o real (IFN).

Gráfico 6. Análise da acurácia das amostras, referentes ao mês de outubro de 2020.



Fonte: A autora, 2021.

A classe com maior erro de omissão continua a ser a da azinheira (88%) seguida pela do pinheiro-bravo (73 %) que no mês de julho foi a classe com melhor representação.

Ao analisarmos a matriz de confusão do mês de outubro podemos constatar a confusão entre as classes. A maior confusão é a da azinheira, que se mistura com todas as outras classes e em especial com a do pinheiro-bravo, assim como, as classes de carvalho e castanheiro.

Conforme expresso na Tabela 6 a validação dos dados obteve valores muito inferiores em relação aos outros meses, onde o overral accuracy (Exatidão Global) foi de 40% e o índice *Kappa* de 15 %, o que é considerado um valor fraco na classificação de Landis e Koch, (1977).

Observando os resultados, percebe-se que a classe que apresentou os maiores valores de erro de omissão e de comissão em todos os três meses analisados foi a azinheira, onde cerca de 97% dos pixels analisados nesta classe foram assinalados em outras classes.

As áreas de floresta dentro de Portugal continental não são tão homogêneas, sendo assim, qualquer mudança na densidade de copa das árvores influencia na sua leitura espectral e por fim influencia na classificação.

Tabela 5. Matriz de confusão para a análise da acurácia obtida na classificação supervisionada Random Forest, com informações de acurácia e erros de produtor e usuário, do mês de julho de 2020.

Dados Validação (IFN)							
Dados Classificados (RF)	Azinheira	Carvalho	Castanheiro	Pinheiro Bravo	Total Geral	Precisão do Utilizador (%)	Erro de Omissão (%)
Azinheira	12	-	-	-	12	-	100%
Carvalho	69	108	57	1	235	54%	46%
Castanheiro	13	17	119	-	149	20%	80%
Pinheiro-bravo	13	1	1	89	104	14%	86%
Total	107	126	177	90	428		
Precisão do produtor	89%	14%	33%	1%			
Erro de comissão (%)	11%	86%	67%	99%			
Exatidão global	77%						
Coeficiente <i>Kappa</i>	63%						

Fonte: Autora, 2021.

Tabela 6. Matriz de confusão para a análise da acurácia obtida na classificação supervisionada Random Forest, com informações de acurácia e erros de produtor e usuário, do mês de outubro de 2020.

Dados Classificados (RF)	Dados de Validação (IFN)				Total Geral	Precisão do Utilizador (%)	Erro de Omissão (%)
	Azinheira	Carvalho	Castanheiro	Pinheiro-bravo			
Azinheira	9	5	2	-	16	44%	56%
Carvalho	20	90	100	22	232	61%	39%
Castanheiro	2	17	107	14	140	14%	76%
Pinheiro-bravo	45	99	1	13	158	92%	8%
Total	76	211	210	49	546		
Precisão do produtor	88%	57%	48%	73%			
Erro de comissão (%)	12%	43%	51%	27%			
Exatidão global	40%						
Coeficiente <i>kappa</i>	15%						

Fonte: Autora, 2021

Estudos realizados por Macedo et al., (2018) utilizando imagens de alta resolução e realizados no verão permitiram um alto nível de precisão na classificação de bosque de *Quercus rotundifolia* em função de índices de vegetação. O que indica que aumentando a resolução da imagem utilizada pode-se aumentar a acurácia das amostras estudadas, assim como das áreas com maior representatividade da espécie (HERNANDEZ et al., 2020). Como a Azinheira é uma espécie que prefere zonas de ombrotipo sub humido alto e termotipo supramediterrânico baixo, sua distribuição nos limites do parque é reduzida, como pode ser observado na Figura 10 e 11. Este motivo pode ter influenciado negativamente.

Figura 10. Dispersão da azinheira em relação com o Ombrotipo do Parque Natural de Montesinho.

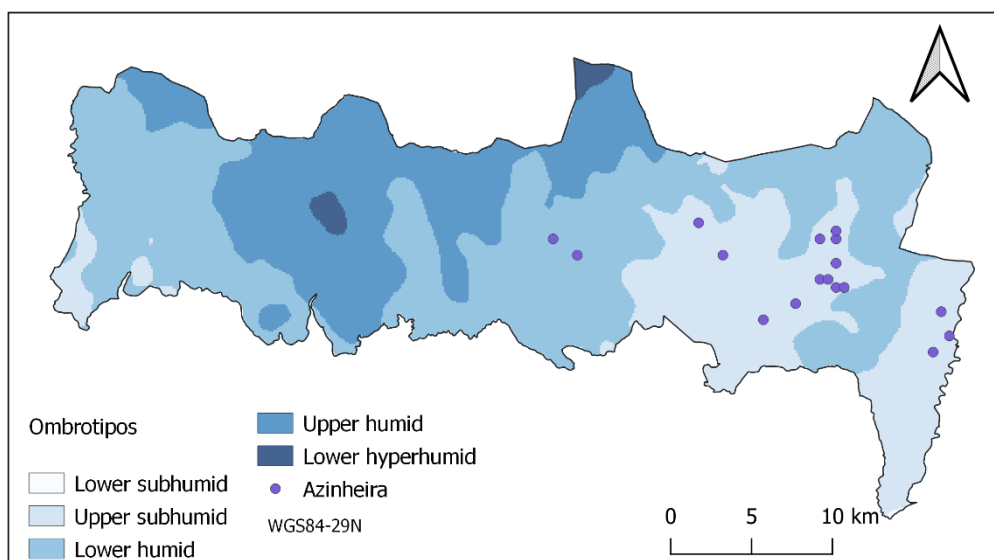
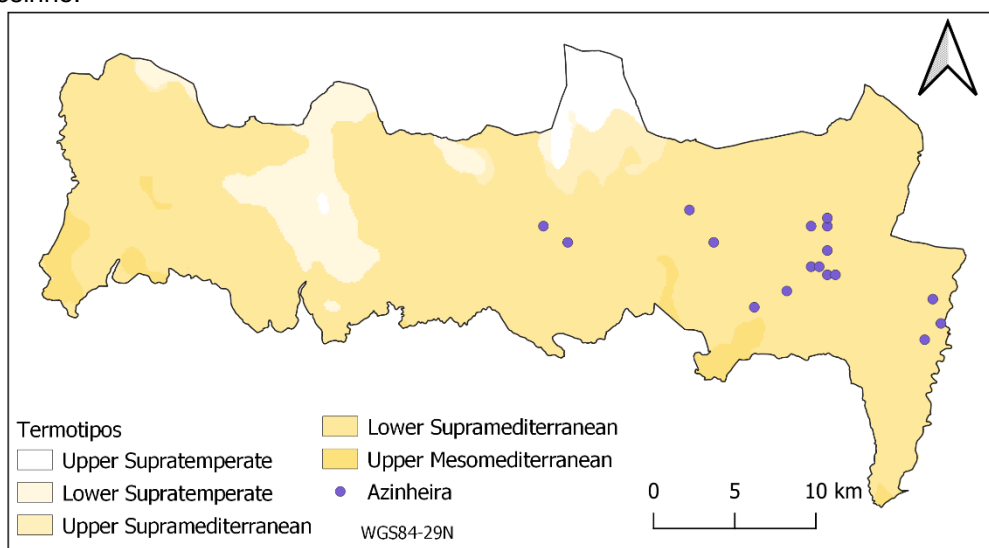


Figura 11. Dispersão da azinheira em relação com o Termotipo do Parque Natural de Montesinho.



Fonte: Adaptado pelo autor de IFN e Monteiro-henriques et al., (2016).

5 CONCLUSÃO

O mapeamento da cobertura da terra é um processo complexo e com variáveis que interferem na qualidade do produto, sendo alguns desses fatores, o padrão do mosaico da paisagem, as classes de interesse e a escala. Nesta pesquisa concentraram-se os esforços em identificar bosques de espécies de grande importância para a região. Encontraram-se algumas dificuldades na identificação destas, devido às suas leituras espectrais se assemelharem e provocando uma confusão de classificação, principalmente com a azinheira e o carvalho.

Em suma o mês com a pior classificação foi o de outubro, seguido pelo mês de fevereiro. Podemos relacionar esses resultados com a estação do ano em que as classificações foram conduzidas, sendo essas de inverno e de outono, onde as plantas estão em repouso vegetativo.

Com base nos resultados obtidos, o mês que melhor se ajusta a classificação feita pelo RF e o real (IFN) para a identificação das espécies de interesse foi o de julho no verão, onde a vegetação está com uma alta produção fotossintética. Ainda são necessários desenvolvimentos para obter resultados de Kappa e acurácia melhores. Assim, recomenda-se a utilização de imagens com melhor resolução espectral, diminuindo a importância dos pixels que juntam diferentes classes, além de uma amostragem maior das áreas de treino. Porém, para isso seria interessante buscar outros povoamentos de azinheira, com maiores dimensões de terreno e com maior representatividade.

Estas recomendações são fornecidas para que os próximos estudos apurem uma metodologia que se mostra muito promissora: o Random Forest. Este demonstra uma significativa capacidade em trabalhar com números crescentes. Seria interessante adicionar ao modelo algumas métricas espectrais e temporais que descrevessem a distribuição espectral de uma banda ou até mesmo algum índice daquele período em específico. Fornecer mais informação poderia aumentar a precisão da classificação, para aplicar algoritmos de segmentação de objetos, onde pixels segmentados poderiam ser agregados em limites mais homogêneos.

6 REFERÊNCIAS

AGUIAR, Carlos Francisco Gonçalves. **Flora e vegetação da Serra de Nogueira e do Parque Natural de Montesinho Doutoramento**. 2000. 659 f. Tese (Doutorado) - Curso de Engenharia Agronómica, Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Lisboa, 2000.

ALEGRIA, C, ROQUE N, ALBUQUERQUE T, FERNANDEZ P, RIBEIRO MM. Modelling Maritime Pine (*Pinus pinaster* Aiton) Spatial Distribution and Productivity in Portugal: Tools for Forest Management. **Forests**, v. 12, n. 3, p. 368, 2021.

ANAND, A. (2017). Unit-14 Accuracy Assessment. Processing and Classification of Remotely Sensed Images. **Remote Sensing and Image Interpretation**, p. 59-78.

AZEVEDO, J.; CAÇADOR, F. Bordaduras de bosque de *Quercus rotundifolia* Lam. no Parque Natural de Montesinho. v. 1, p. 163–176, dez. 1999.

BELGIU, M., & DRĂGUȚ, L. (2016). Random forest in remote sensing: A review of applications and future directions. **ISPRS journal of photogrammetry and remote sensing**, 114, 24-31.

BENITO GARZÓN, M.; SÁNCHEZ DE DIOS, R.; SAINZ OLLERO, H. Effects of climate change on the distribution of Iberian tree species. **Applied Vegetation Science**, v. 11, n. 2, p. 169–178, 2008.

BREIMAN, Leo *et al.* **Random Forests**. 2016. Disponível em: <https://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm>. Acesso em: 16 jul. 2021.

BRANDÃO, Z. N.; BEZERRA, M. V. C.; SILVA, B. B. Uso do NDVI para determinação da Biomassa na Chapada do Araripe. **XIII Simpósio Brasileiro de Sensoriamento Remoto**, p. 75–81, 2007.

CASTRO, J. Land Use, Landscape and Sustainability: Examples from Montesinho. In **Natural Heritage from East to West**. Springer, Berlin, Heidelberg. p. 151–154.

CASTRO, E.G. de. **Técnicas de sensoriamento remoto e sistemas geográficos de informações no estudo integrado de bacias hidrográficas**. São José dos Campos. Dissertação (mestrado em sensoriamento remoto) – Instituto Nacional de Pesquisas Espaciais, p. 142, 1992.

CONGALTON, R. G. of Wildland Fire and other spatial information. **International Journal of Wildland Fire**, v. 10, n. January 2001, p. 321–328, 2001.

CONGEDO, L. Semi-Automatic Classification Plugin Documentation. n. August, 2021.

COSTA, J. C. et al. Biogeografia de Portugal Continental. p. 46, 1990.

DAINESE, RENATA C. **Sensoriamento remoto e geoprocessamento aplicado ao estudo temporal do uso da terra e na comparação entre classificação não supervisionada e análise visual**, p. 186, 2001. 50

DIAS, R.; AZEVEDO, J. C. **Distribution and spatial configuration of holm oak woodlands in the Montesinho/Nogueira site, Portugal**. IUFRO, 2008.

Direção-Geral Território. (2017a). **Carta Administrativa Oficial de Portugal - CAOP 2019**. Lisboa, PT.

Disponível em: <https://www.dgterritorio.gov.pt/cartografia/cartografia-tematica/caop>. Acesso em: 20 de janeiro de 2020.

DUARTE, J. F. DOS SANTOS; RAMOS, A. JOSE REIS; MACIEL, M. DE NAZARÉ MARTINS. Avaliação da cobertura vegetal em áreas de preservação permanente em mojuí dos campos – PA através do NDVI. 2015.

EUROPEAN SPACE AGENCY (ESA). User Guides - Sentinel-2 MSI - Revisit and Coverage - Sentinel Online. 2020. Disponível em: <https://sentinel.esa.int/web/sentinel/userguides/sentinel-2-msi/revisit-coverage>. Acesso em: 22 de junho de 2021.

Feng, Quanlong, Jiantao Liu, and Jianhua Gong. "UAV Remote Sensing for Urban Vegetation Mapping Using Random Forest and Texture Analysis" **Remote Sensing** 7, 2015.

GONÇALVES, J. M. A.; **Utilização do Geoprocessamento para a Produção de Mapas Temáticos, Aplicação na Sub-Bacia do Rio Maranhão - Bacia Do Rio Paraopeba. Monografia** (Monografia em Geoprocessamento), Universidade Federal de Minas Gerais. Belo Horizonte, p. 31, 2002.

HERNANDEZ, I. et al. Exploring sentinel-2 for land cover and crop mapping in Portugal. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives**, v. 43, n. B3, p. 83–89, 2020.

HUETE, A. R. A Soil-Adjusted Vegetation Index (SAVI). v. 309, p. 295–309, 1988.

ICNF. Instituto da Conservação da Natureza e das Florestas (2007). **Plano de Ordenamento do Parque Natural de Montesinho**. Bragança, PT. Disponível em: <http://www.icnf.pt/portal/naturaclas/ordgest/poap/popnm/popnm>. Acesso em: 20 de novembro de 2020.

ICNF Instituto da Conservação da Natureza e das Florestas. **Gestão Florestal**. Bragança, PT. Disponível em: <http://www2.icnf.pt/portal/florestas/gf>. Acesso em: 20 de novembro de 2020.

JACKSON, R. D.; HUETE, A. R. Interpreting vegetation indices. **Preventive Veterinary Medicine**, v.11, n.3-4, p.185-200, 1991.

LANDIS, J. R.; KOCH, G. G. The Measurement of Observer Agreement for Categorical Data. **Biometrics**, v. 33, n. 1, p. 159, 1977.

LAURENTINO, J. R. S. **Condição fitossanitária do montado de sobreiro com recurso a imagens multiespectrais de alta resolução**. [s.l.] UNIVERSIDADE DE LISBOA FACULDADE DE CIÊNCIAS, 2019.

LEITE, A. N. A. P.; SANTOS, G. R.; SANTOS, J. É. O. Análise Temporal Dos 51 Índices De Vegetação Ndvi E Savi Na Estação Experimental De Itatinga Utilizando Imagens Landsat 8. **Revista Brasileira de Energias Renováveis**, p. 606–623, 2017.

LILLESAND, Thomas; KIEFER, Ralph W.; CHIPMAN, Jonathan. **Remote sensing and image interpretation**. John Wiley & Sons, p. 756. 2014.

LIMA, S. A. **Avaliação da acurácia do posicionamento e orientação de aeronaves remotamente pilotada com uso de técnicas de fotogrametria e processamento digital de imagens**. 2018. 211 p. Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2018.

LEMENKOVA, P. Hyperspectral Vegetation Indices Calculated by Qgis Using Landsat Tm Image: A Case Study of Northern Iceland. **Advanced Research in Life Sciences**, v. 4, n. 1, p. 70–78, 2020.

LISBOA FILHO, Jugurta. Introdução a SIG-Sistemas de Informações Geográficas. **Porto Alegre: CPGCC da UFRGS**, p. 69, 1995.

MACEDO, F. L. et al. Above-ground biomass estimation for *Quercus rotundifolia* using vegetation indices derived from high spatial resolution satellite images. **European Journal of Remote Sensing**, v. 51, n. 1, p. 932–944, 2018.

MAGNONI, P. H. J. **Cartografia da paisagem natural do Parque Natural de Montesinho por Sensoriamento Remoto**. [s.l.] Escola Superior Agrária de Bragança, 2017.

MALENOVSKÝ, Z. et al. Sentinels for science: Potential of Sentinel-1, -2, and -3 missions for scientific observations of ocean, cryosphere, and land. **Remote Sensing of Environment**, v. 120, p. 91–101, 2012.

MARTINS, R. Cartografia de áreas áridas do Norte de Portugal Continental com recurso a imagens multiespectrais do satélite Sentinel-2. p. 69 p, 2017.

MONTEIRO-HENRIQUES, T. et al. Bioclimatological mapping tackling uncertainty propagation: Application to mainland Portugal. **International Journal of Climatology**, v. 36, n. 1, p. 400–411, 2016.

MOREIRA, ALVES, M. **Fundamentos do Sensoriamento Remoto e Metodologia de Aplicação**. Ed. UFV. v. 1ª edição, p. 246, 2001.

NOVO, Evelyn, M. L. de Moraes. **Sensoriamento Remoto Princípios e Aplicações**. 3º ed. Revisada e Ampliada, p.363, 2008.

PAL, M. Random forest classifier for remote sensing classification. **International Journal of Remote Sensing**, v. 26, n. 1, p. 217–222, 2005. 52

PORTUGAL. ICNF. (org.). **6.º Inventário Florestal Nacional (IFN6)**. 2015. Relatório Final. Autoridade Florestal Nacional, Lisboa Disponível em: <http://www2.icnf.pt/portal/florestas/ifn/ifn6>. Acesso em: 08 fev. 2021.

PONZONI, Flávio Jorge; SHIMABUKURO, Yosio Edemir; KUPLICH, Tatiana Mora. **Sensoriamento remoto da vegetação**. 2º. ed. São José, dos Campos: Oficina de Textos, p.159, 2015.

PRATES, L. P.; LACERDA, C. S. DOS A. Comparação entre o uso de ndvi, ndbi e classificação não- supervisionada para avaliação da expansão urbana. n. 2003, 2019.

QUERCUS Organização Não Governamental de Ambiente (ONGA). **Florestas** Disponível em: <https://quercus.pt/floresta/> Acesso em: 14 de outubro de 2020.

RAMOS, M. F. Estudo da variação temporal do índice de vegetação por diferença normalizada em juiz de fora a partir de imagens de satélite. 2016.

SARAIVA, C. S. A. Avaliação do Potencial das Imagens Sentinel-1 para Identificação de Culturas Agrícolas. p. 71, 2015.

SEQUEIRA, E.; AGUIAR, C.; MEIRELES, C. **Ultramafics of Bragança Massif: Soils, Flora and Vegetation Natural Heritage from East to West: Case Studies from 6 EU Countries**, 2010.

SHIMABUKURO, Y. E.; NOVO, E. M.; PONZONI, F. J. Índice de vegetação e modelo linear de mistura espectral no monitoramento da região do Pantanal. **Pesquisa Agropecuária Brasileira**, v. 33, p. 1729-1737, 1998.

SPRING Sistema de Processamento de Informações Georreferenciadas. **INPE Instituto de Pesquisas Espaciais**. Disponível em: <http://www.dpi.inpe.br/spring/portugues/tutorial/index.html> Acesso em: 14 de outubro de 2020.

SZATMÁRI, J.; TOBAK, Z.; NOVÁK, Z. Environmental Monitoring Supported by Aerial Photography – a Case Study of the Burnt Down Bugac Juniper Forest, Hungary. **Journal of Environmental Geography**, v. 9, n. 1–2, p. 31–38, 2016.

Testa, S., & Borgogno Mondino, E. (2013, December). MODIS Vegetation Indices time series improvement considering real acquisition dates. In AGU Fall Meeting Abstracts (Vol. 2013, pp. B43C-0519).

TORRES, R. et al. GMES Sentinel-1 mission. **Remote Sensing of Environment**, v. 120, p. 9–24, 2012.

USGS GLOBAL VISUALIZATION VIEWER (Usa) (Org.). **Science for a changing world**. Disponível em: <http://glovis.usgs.gov/>. Acesso em: 20 de novembro de 2020.

WENG, Qihao (Ed.). **Remote Sensing Time Series Image Processing**. CRC Press, 2018.