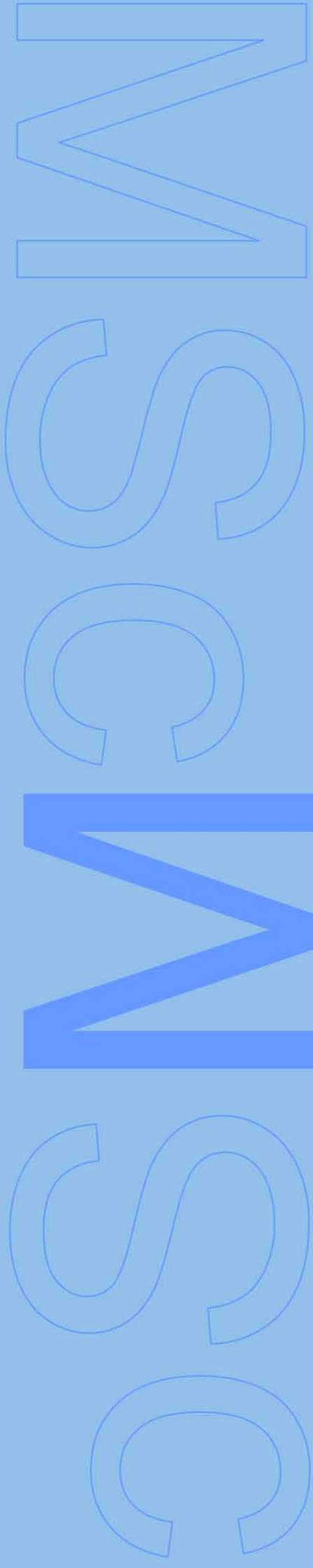


Análise Espaço- Temporal da Asma em Portugal Continental com base em Dados de Detecção Remota

Diogo Ayres Sampaio
2012



Análise Espaço- Temporal da Asma em Portugal Continental com base em Dados de Detecção Remota

Diogo Ayres Sampaio

Mestrado em Engenharia Geográfica

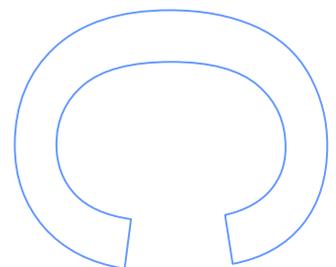
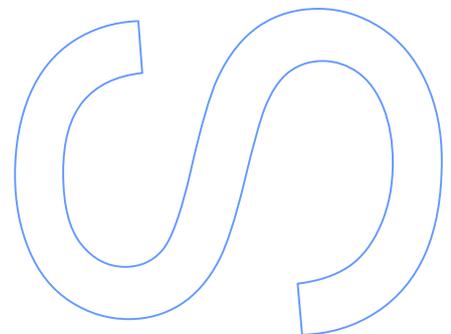
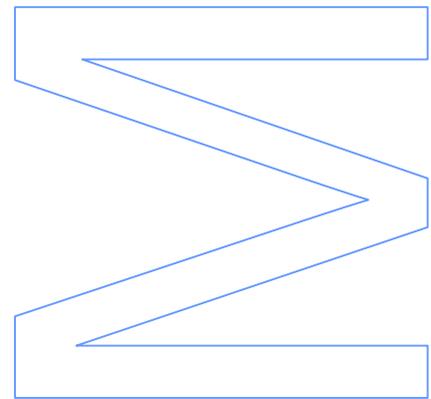
Departamento de Geociências, Ambiente e Ordenamento do Território
2012

Orientador

Ana Cláudia Moreira Teodoro, Professora Auxiliar, Faculdade
de Ciências da Universidade do Porto

Co-orientador

José Alberto da Silva Freitas, Professor Auxiliar, Faculdade
de Medicina da Universidade do Porto



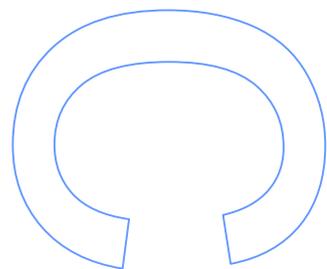
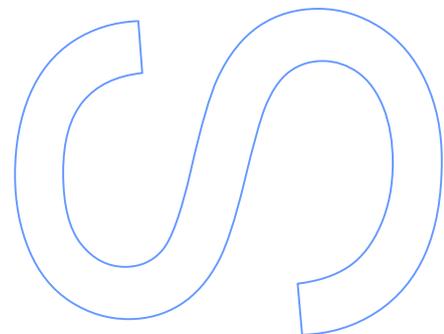
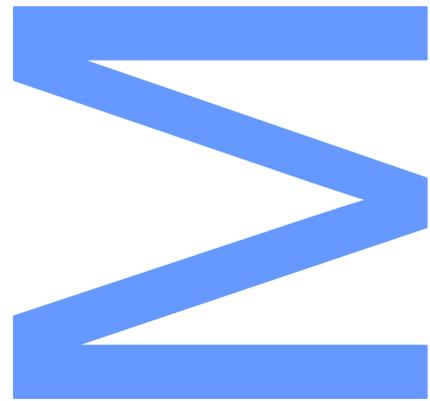
U. PORTO

FC FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ____ / ____ / ____



Agradecimentos

A realização do trabalho aqui apresentado só foi possível devido à contribuição de várias pessoas a quem quero expressar a minha sincera gratidão.

À minha orientadora, a Professora Doutora Ana Cláudia Teodoro, um agradecimento especial pela confiança que depositou no meu trabalho. Obrigado pela constante motivação, inesgotável paciência, dedicação e amizade ao longo do desenvolvimento deste projeto.

Ao meu co-orientador, o Professor Doutor Alberto Freitas, por me ter recebido no Centro de Investigação em Tecnologias e Sistemas de Informação em Saúde, pela disponibilidade e pelas sugestões fundamentais na conceção deste trabalho.

Ao Neftalí, por todo o apoio, motivação e espírito crítico constante durante a elaboração deste projeto.

Ao Professor Doutor João Fonseca, pela simpatia, pelos pertinentes comentários e pela sua importante visão clínica do trabalho.

À Professora Doutora Cristina Santos, pela ajuda na análise estatística e interesse demonstrado pelo meu trabalho.

Aos meus colegas da FCUP, nomeadamente à Cátia, Elena, Elisabete, Hélder, Miguel e Néilson, pelos momentos de boa disposição e descontração que me proporcionaram.

A todos os meus colegas de curso, em especial ao André, Bento, Nuno e Sérgio, que acompanharam o meu percurso em Engenharia Geográfica desde o início. Obrigado, sobretudo, pela amizade e companheirismo.

À minha avó, pelo orgulho e preocupação que sempre demonstrou por mim.

Agradecimentos

Aos meus pais, pelo apoio incondicional em todas as escolhas que tomei ao longo da minha vida. Obrigado pela paciência, dedicação e compreensão. Espero continuar a ser motivo do vosso orgulho.

Por fim, um agradecimento muito especial à Elsa, por tornar a minha vida repleta de felicidade. Obrigado por estares sempre presente e por me incentivares a fazer sempre mais e melhor. Sem dúvida que a tua inspiração e dedicação facilitou o meu percurso académico.

Resumo

A asma é uma doença inflamatória crónica das vias aéreas que afeta pessoas de todas as idades em todo o mundo. A manifestação desta doença pode ser influenciada por vários fatores ambientais, como alergénios, poluição do ar e condições meteorológicas. Apesar da crescente utilização da Deteção Remota (DR) em estudos epidemiológicos, as suas potencialidades ainda não foram utilizadas no estudo de doenças como a asma. Assim, o principal objetivo deste trabalho consiste em estudar a asma em Portugal Continental, relacionando um conjunto de parâmetros ambientais, determinados maioritariamente com observações de satélites, com internamentos em hospitais públicos entre 2003 e 2008.

O perfil de temperatura do produto MOD07 do *Moderate Resolution Imaging Spectroradiometer* (MODIS), juntamente com o Modelo Digital de Elevação (MDE) do *Shuttle Radar Topography Mission* (SRTM), foi utilizado para determinar a temperatura do ar ao nível da superfície (T_a). O método de Kriging foi utilizado para interpolar medições de humidade relativa (HR), efetuadas por estações meteorológicas, para todo Portugal Continental. O *Normalized Difference Vegetation Index* (NDVI) do produto MOD13A3 do MODIS foi escolhido como indicador da quantidade de vegetação. Recorrendo a modelos *Land-Use Regression* – LUR determinaram-se estimativas espaço-temporais da concentração dos poluentes dióxido de azoto (NO_2) e partículas em suspensão de diâmetro aerodinâmico inferior a $10\ \mu\text{m}$ (PM_{10}). Para além das variáveis de tráfego e ocupação de solo, foram introduzidos nos modelos parâmetros variáveis ao longo do tempo, como o *Aerosol Optical Thickness* (AOT) a 550 nm do produto MOD04 do MODIS. A associação entre a taxa de internamentos por 1000 habitantes em cada concelho com as variáveis ambientais (T_a , HR, NDVI, NO_2 e PM_{10}), durante todo o intervalo de tempo em estudo, foi efetuada a partir de análises de regressão linear

simples e múltipla. De referir, que os 18 distritos de Portugal Continental foram analisados separadamente.

Na análise univariada encontram-se relações significativas em 8 distritos (Aveiro, Braga, Castelo Branco, Faro, Lisboa, Portalegre, Porto e Setúbal). Nos distritos mais populosos as relações tendem a ser mais elevadas. Na maioria, a T_a e os poluentes correlacionaram-se positivamente com a taxa de internamentos por asma, enquanto o NDVI se correlacionou negativamente. Na análise multivariada, observou-se que nos distritos de Braga e do Porto a taxa de internamentos por asma estava associada, respetivamente, com as combinações $T_a + PM_{10}$ e $NO_2 + HR$. Compreender os fatores que afetam as várias zonas do país pode servir de auxílio ao controlo da asma.

Os dados de DR apresentam grandes potencialidades, mas possuem limitações que restringem a sua utilização em estudos epidemiológicos de doenças que exigem parâmetros de pouca complexidade. No estudo epidemiológico de doenças respiratórias, a exposição a poluentes é uma das variáveis mais importantes. Atualmente, a baixa resolução dos dados de DR, referentes à concentração de poluentes, não permite a sua correta modelação ao nível da superfície em escalas locais.

Abstract

Asthma is a chronic inflammatory disorder of the airway that affects people of all ages throughout the world. The expression of this disease can be influenced by some environmental factors such as allergens, air pollution or climate conditions. Despite the growing use of Remote Sensing (RS) data in epidemiological studies, several diseases, including asthma, have not been studied yet using RS potentialities. Therefore, the main objective of this study was to relate some environmental parameters, mainly obtained from satellite observations, with public hospital admissions in Mainland Portugal, between 2003 and 2008.

The temperature profile, retrieved from the Moderate Resolution Imaging Spectroradiometer (MODIS), and the Digital Elevation Model (DEM) from Shuttle Radar Topography Mission (SRTM) were used to compute the near-surface air temperature (T_a). Standard Kriging method was used with meteorological stations' data to compute spatial estimates of relative humidity (RH). The Normalized Difference Vegetation Index (NDVI) was used as an indicator of vegetation. Two Land-Use Regression (LUR) models were built to compute spatiotemporal estimates of the concentration of nitrogen dioxide (NO_2) and particulate matter with an aerodynamic diameter less than $10 \mu\text{m}$ (PM_{10}). In addition to the commonly used variables of traffic and land cover, variables such as Aerosol Optical Thickness (AOT) at 550 nm were introduced into the models to account for spatial and temporal variability of the pollutants. Single and multiple regression analysis were used to relate the asthma hospitalization rates per 1000 inhabitants in each municipality with the environmental variables (T_a , RH, NDVI, NO_2 , and PM_{10}), during the time interval under study. It should be noted that the 18 districts in Portugal Mainland were analyzed separately.

In single variable regression models, relationships were found for only 8 districts (Aveiro, Braga, Castelo Branco, Faro, Lisboa, Portalegre, Porto, and Setúbal). In the most

populous districts relationships tend to be higher. In most cases, T_a and pollutants were positively correlated with asthma hospitalization rates, while NDVI presented a negative correlation. In multi-variable regression models, asthma hospitalization rates were significantly correlated with $T_a + PM_{10}$ in Braga and with $NO_2 + RH$ in Porto. Understanding the factors that affect the different areas of the country might be useful in the control of asthma.

RS data have a great potential but also present limitations which restrict their use in epidemiological studies of diseases that require low complexity parameters. In epidemiological study of respiratory diseases, air pollution exposure is one of the most important variables. Currently, the low spatial resolution of RS air pollution data does not allow the correct modeling of near-surface exposure at local scales.

Índice

Agradecimentos	v
Resumo	vii
Abstract	ix
Índice de Figuras	xiii
Índice de Tabelas	xv
Abreviaturas	xvii
1 Introdução	1
1.1 Epidemiologia	1
1.1.1 Epidemiologia Espacial	2
1.2 Detecção Remota	4
1.2.1 Definição e Origens.....	4
1.2.2 Radiação Eletromagnética	5
1.2.3 A Detecção Remota Aplicada à Epidemiologia	7
1.3 Asma.....	11
1.4 Objetivo.....	13
2 Dados	15
2.1 <i>Moderate Resolution Imaging Spectroradiometer</i> (MODIS).....	15
2.1.1 Perfis Atmosféricos (MOD07)	19
2.1.2 <i>Aerosol Optical Thickness</i> (MOD04)	19
2.1.3 Índices de Vegetação (MOD13A3)	20
2.1.4 Formato dos Dados MODIS	21
2.2 Modelo Digital de Elevação	24
2.3 Dados de Estações Meteorológicas	25
2.4 Dados da Concentração de Poluentes	26

2.5 Carta Administrativa Oficial de Portugal 2011	28
2.6 Cartografia CORINE Land Cover 2006 para Portugal Continental	28
2.7 Rede Viária.....	30
2.8 Dados de Internamentos Hospitalares e Populacionais	30
3 Metodologia	33
3.1 Determinação da Temperatura e Humidade Relativa do Ar a partir do Perfil Atmosférico MODIS.....	33
3.2 Interpolação Espacial: Kriging	38
3.3 Modelos <i>Land-Use Regression</i> (LUR)	43
3.4 Análise Estatística	46
3.5 Implementação dos Métodos	48
3.5.1 Dados das Estações Meteorológicas e de Qualidade do ar	48
3.5.2 Modelo Digital de Elevação para Portugal Continental.....	50
3.5.3 Determinação da Temperatura e Humidade Relativa do Ar a partir do Perfil Atmosférico MODIS.....	50
3.5.4 <i>Aerosol Optical Thickness</i>	55
3.5.5 <i>Normalized Difference Vegetation Index</i>	56
3.5.6 Determinação de Variáveis Ambientais pelo Método de Kriging	57
3.5.7 Determinação dos Poluentes por Modelos LUR.....	61
3.5.8 Análise Estatística	64
4 Resultados	67
4.1 Variáveis Ambientais	67
4.1.1 Temperatura e Humidade Relativa do Ar (MODIS)	67
4.1.2 Método de Kriging.....	72
4.1.3 Modelos LUR.....	76
4.2 Análises de Regressão.....	80
5 Discussão.....	87
5.1 Variáveis Ambientais	87
5.2 Análise Epidemiológica da Asma.....	89
5.3 Limitações	92
6 Conclusões e Perspetivas Futuras	95
Referências	97
Anexos.....	107

Índice de Figuras

Figura 1.1. Mapa elaborado por John Snow em 1854 com a incidência da cólera em Londres.....	3
Figura 1.2. Propagação da Radiação Eletromagnética.	6
Figura 1.3. Espectro Eletromagnético.....	6
Figura 2.1. <i>Swath</i> típica de um satélite com sensor de varrimento.....	22
Figura 2.2. Projeção Sinusoidal da Terra, dividida em quadrados de 10° por 10°	24
Figura 2.3. Processo de aquisição dos dados do SRTM.....	25
Figura 2.4. Distribuição das estações meteorológicas utilizadas.....	26
Figura 2.5. Distribuição das estações de qualidade do ar utilizadas.....	27
Figura 3.1. Relação entre a pressão média mensal do ar medida pelas estações e a respetiva altitude.....	35
Figura 3.2. Relação entre a pressão média mensal do ar medida pelas estações meteorológicas e dada pelo produto MOD07.....	36
Figura 3.3. Exemplo de um semi-variograma.....	40
Figura 3.4. Exemplo gráfico de modelos de variograma.....	41
Figura 3.5. Exemplo da distribuição dos dados do poluente NO_2 : a) sem transformação; b) com transformação raiz quadrada.	46
Figura 3.6. Taxa de internamentos devido a asma por 1000 habitantes em função da população total média, referente ao intervalo de tempo de 2003 a 2008.	47
Figura 3.7. Fluxograma com a metodologia de determinação das médias mensais de T_a e HR.	54
Figura 3.8. Fluxograma com a metodologia de processamento do NDVI.	58
Figura 3.9. Fluxograma com a metodologia para a determinação das estimativas espaciais mensais das variáveis ambientais, recorrendo ao método de Kriging.	62
Figura 4.1. Relação entre a T_a medida e prevista (MOD07).....	68

Figura 4.2. Relação entre a T_d medida e prevista (MOD07).	70
Figura 4.3. Relação entre a HR medida e prevista (MOD07).	71
Figura 4.4. Exemplo de duas imagens correspondentes a médias mensais do mês de Novembro de 2005 determinadas com dados MODIS: a) T_a ; b) HR.	71
Figura 4.5. Relação entre a T_a medida e prevista (Kriging).	72
Figura 4.6. Relação entre a HR medida e prevista (Kriging).	73
Figura 4.7. Validação do PM_{10} : a) Zona Norte; b) Zona Centro.	73
Figura 4.8. Validação do NO_2 : a) Zona Norte; b) Zona Centro.	74
Figura 4.9. Exemplo de duas imagens correspondentes a médias mensais do mês de Novembro de 2005 determinadas pelo método de Kriging: a) T_a ; b) HR.	75
Figura 4.10. Exemplo de um mapa com a distribuição das estações de medição do NO_2 na zona Norte, correspondente a Novembro de 2005.	76
Figura 4.11. Histograma dos resíduos do modelo: a) PM_{10} ; b) NO_2 . Resíduos do modelo vs. valores previstos: c) PM_{10} ; d) NO_2	77
Figura 4.12. Exemplo de duas imagens correspondentes a médias mensais dos modelos LUR, correspondentes a Novembro de 2005: a) PM_{10} ; b) NO_2	80
Figura 4.13. Relação entre o R^2 , atingido com o NDVI e a taxa de internamentos devido a asma por 1000 habitantes, e respetivo NO_2 médio no distrito.	83
Figura 4.14. Relação entre o R^2 , atingido com o NO_2 e a taxa de internamentos devido a asma por 1000 habitantes, e respetivo NO_2 médio no distrito.	84
Figura 4.15. Relação entre o R^2 , atingido com o PM_{10} e a taxa de internamentos devido a asma por 1000 habitantes, e respetivo PM_{10} médio no distrito.	84

Índice de Tabelas

Tabela 1.1. Dados de satélite.	10
Tabela 1.2. Fatores que influenciam o desenvolvimento e/ou expressão da asma.....	12
Tabela 2.1. Características e aplicações das bandas do MODIS.	17
Tabela 2.2. Parâmetros MODIS utilizados no âmbito deste trabalho.	18
Tabela 2.3. Nomenclatura dos dois primeiros níveis de classes do CLC06_PT.....	29
Tabela 3.1. Variáveis independentes consideradas na criação dos modelos LUR.	45
Tabela 3.2. Classes utilizadas na construção dos modelos LUR.	46
Tabela 3.3. Bases de dados criadas com os parâmetros meteorológicos e poluentes.	48
Tabela 3.4. Programas utilizados para calcular médias mensais dos dados pontuais.	49
Tabela 3.5. Programas utilizados para calcular as médias mensais de T_a e HR a partir do MOD07.	51
Tabela 3.6. Programas utilizados no cálculo das médias mensais do AOT550 a partir do MOD04.	55
Tabela 3.7. Programas utilizados para preparar o NDVI mensal do MOD13A3.	56
Tabela 3.8. Programas utilizados para determinar, pelo método de Kriging, estimativas espaciais mensais das variáveis ambientais.	59
Tabela 3.9. Programas utilizados para preparar os dados para a análise estatística.....	64
Tabela 4.1. Validação mensal de cada combinação de dados considerada no cálculo das médias mensais de T_a	69
Tabela 4.2. Validação das variáveis determinadas com dados MODIS.	72
Tabela 4.3. Validação das variáveis ambientais determinadas com o método de Kriging.	75
Tabela 4.4. Modelo de regressão para prever a concentração de PM_{10}	78
Tabela 4.5. Modelo de regressão para prever a concentração de NO_2	79
Tabela 4.6. Validação dos modelos LUR.....	80

Tabela 4.7. Descrição das variáveis ambientais por distrito.	81
Tabela 4.8. Análise univariada entre as variáveis ambientais (variáveis independentes) e a taxa de internamentos devido a asma por 1000 habitantes (variável dependente) nos concelhos de cada distrito.	82
Tabela 4.9. Análise multivariada entre as variáveis ambientais (variáveis independentes) e a taxa de internamentos devido a asma por 1000 habitantes (variável dependente) nos concelhos de cada distrito.	85

Abreviaturas

AOT	<i>Aerosol Optical Thickness</i>
AOT550	AOT a 550 nm
AVHRR	<i>Advanced Very High Resolution Radiometer</i>
BD	Base de Dados
CAOP	Carta Administrativa Oficial de Portugal
CO	Monóxido de carbono
CORINE	<i>Co-ordination of Information on the Environment</i>
CLC2006	<i>CORINE Land Cover 2006</i>
DP	Desvio padrão
DR	Deteção Remota
EM	Erro médio
EMQ	Erro médio quadrático
EN	Estradas Nacionais
GINA	<i>Global INitiative for Asthma</i>
GNSS	<i>Global Navigation Satellite System</i>
GPS	<i>Global Positioning System</i>
HR	Humidade relativa do ar ao nível da superfície
IC	Itinerários complementares
IGP	Instituto Geográfico Português
INE	Instituto Nacional de Estatística
IP	Itinerários principais
LST	<i>Land Surface Temperature</i>
LUR	<i>Land-Use Regression</i>
MDE	Modelo Digital de Elevação
MDE5	MDE com 5 km de resolução espacial

MODIS	<i>Moderate Resolution Imaging Spectroradiometer</i>
MSS	<i>Multispectral Scanner</i>
NASA	<i>National Aeronautics and Space Administration</i>
NDVI	<i>Normalized Difference Vegetation Index</i>
NO₂	Dióxido de azoto
NOAA	<i>National Oceanic and Atmospheric Administration</i>
O₃	Ozono
OMI	<i>Ozone Monitoring Instrument</i>
OSM	<i>OpenStreetMap</i>
PM	<i>Particulate matter</i> (Partículas em suspensão)
PM_{2.5}	Partículas em suspensão de diâmetro aerodinâmico inferior a 2.5 µm
PM₁₀	Partículas em suspensão de diâmetro aerodinâmico inferior a 10 µm
PT_a	Perfil de temperatura do ar
R²	Coeficiente de determinação
REM	Radiação Eletromagnética
SIG	Sistemas de Informação Geográfica
SO₂	Dióxido de enxofre
SRTM	<i>Shuttle Radar Topography Mission</i>
T_a	Temperatura do ar ao nível da superfície
TCWV	<i>Total Column Water Vapor</i> (coluna total de vapor de água)
TM	<i>Thematic Mapper</i>

Capítulo 1

Introdução

1.1 Epidemiologia

A epidemiologia consiste no estudo da distribuição e fatores determinantes de estados ou eventos (incluindo doenças) relacionados com a saúde em populações específicas e a aplicação desse conhecimento no controlo de doenças e outros problemas de saúde [1-3]. O estudo inclui vigilância, observação, teste de hipóteses, investigação analítica e experiências [2].

O conceito de epidemiologia pode ser dividido em três componentes: frequência, distribuição e fatores determinantes [1]. A medida de frequência de uma determinada doença envolve a quantificação da existência ou ocorrência da doença. Assim, a disponibilização destes dados é um pré-requisito para qualquer estudo sistemático dos padrões de ocorrência de uma doença em populações de seres vivos [1]. A distribuição da doença refere-se à análise por tempo, espaço e classes ou subgrupos afetados numa população [2]. O conhecimento de tais distribuições é essencial para descrever os padrões de doença e formular hipóteses sobre fatores causais ou de prevenção [1]. Os determinantes da doença consistem em todos os fatores físicos, biológicos, culturais, económicos e comportamentais que influenciam a doença [2].

A epidemiologia é tão antiga quanto a própria medicina. No século V a.C., Hipócrates (considerado o pai da medicina moderna) sugeriu pela primeira vez que o desenvolvimento da doença humana poderia estar relacionado com fatores ambientais e estilo de vida [1, 3, 4]. John Snow foi, provavelmente, o primeiro investigador a reunir os três componentes do conceito de epidemiologia no estudo da cólera, em Londres, na década de 50 do século XIX [1]. Com o decorrer dos anos, o termo epidemiologia foi

sendo utilizado para designar o estudo de epidemias de doenças infecciosas [1, 2]. Contudo, após a Segunda Guerra Mundial, tornou-se evidente o aumento do impacto em países desenvolvidos de doenças não infecciosas de origem desconhecida, como o cancro e as doenças cardiovasculares [4]. Consequentemente, surgiu o conceito atual de epidemiologia [2, 4], que exige métodos mais avançados do que os inicialmente desenvolvidos por Snow.

1.1.1 Epidemiologia Espacial

A epidemiologia espacial resulta da interação entre a estatística, a geografia e a epidemiologia e consiste na análise espacial da exposição, da doença e da relação entre elas [5]. A epidemiologia espacial permite melhorar a compreensão das doenças ou dos processos dos problemas de saúde, investigar a relação entre o ambiente e a presença da doença, realizar análises de grupos, prever a propagação, avaliar alternativas de controlo da doença e evitar vários erros (p.e., erros de exposição) que de outra forma poderiam ser cometidos [6].

A primeira, e talvez mais conhecida, aplicação da epidemiologia espacial remonta ao estudo de John Snow, anteriormente referido [5-7]. Utilizando mapas, perspicácia médica, aptidões laboratoriais, conhecimentos locais, entrevistas porta-a-porta (*shoe-leather epidemiology*) e raciocínio dedutivo, Snow conseguiu combinar a informação espacial relativa às fontes de água e à incidência da cólera (Figura 1.1) para identificar uma associação, levando ao encerramento da bomba de água de Broad Street e consequentemente à decadência da cólera em Londres [6]. Contudo, os estudos de epidemiologia espacial só se tornaram comuns a partir da década de 90 do século XX, quando os *software* de Sistemas de Informação Geográfica (SIG) se tornaram amplamente disponíveis [5]. Os SIG são sistemas computacionais utilizados para guardar e manipular informação geográfica. Um SIG é desenhado para a recolha, armazenamento e análise de objetos e fenómenos, onde a localização geográfica constitui uma característica importante ou crítica para a análise [8].

O crescente uso dos SIG na epidemiologia, bem como noutras ciências, deve-se sobretudo à diminuição do preço e evoluções verificadas na velocidade e capacidade de memória dos computadores [6]. Esta ferramenta modificou drasticamente a capacidade de epidemiologistas e especialistas em saúde pública de trabalhar com dados espaciais [7, 9]. Para além dos SIG, desenvolvimentos noutras domínios, como a Deteção Remota (DR), a análise espacial (p.e., estatística espacial e geoestatística) ou os Sistemas Globais de Navegação por Satélite (*Global Navigation Satellite System – GNSS*),

contribuíram para a evolução da epidemiologia espacial. Atualmente, existem vários estudos que recorrem a ferramentas SIG e a tecnologias de análise espacial, cujos dados utilizados são adquiridos por sensores instalados em satélites artificiais de observação da Terra. Os sensores de DR têm o potencial de revolucionar os estudos epidemiológicos, pois adquirem grandes quantidades de dados sobre a superfície terrestre com cobertura sinóptica e com várias resoluções espaciais, temporais, espectrais e radiométricas [10]. É neste tipo de dados que se foca uma parte deste trabalho.



Figura 1.1. Mapa elaborado por John Snow em 1854 com a incidência da cólera em Londres (Adaptado de [7]).

1.2 Detecção Remota

1.2.1 Definição e Origens

A DR é a arte e/ou ciência de obter informação sobre um objeto, área ou fenómeno sem que haja contacto direto com esse objeto, área ou fenómeno [11].

O conceito atual de DR está diretamente relacionado com o aparecimento da fotografia e com a exploração de radiação eletromagnética para além da região do visível (radiação infravermelha por Herschel, ultravioleta por Ritter e ondas rádio por Hertz), na primeira metade do século XIX [12, 13]. A primeira fotografia aérea surgiu quase de imediato quando, em 1858, Gaspar Félix Tournachon fotografou parte da superfície terrestre a partir de um balão [11-13]. Mais tarde, no início do século XX, foi inventado o avião e, conseqüentemente, surgiu a Fotogrametria, que consistia na obtenção de informação fiável, quer métrica quer semântica, acerca de objetos por meio da medição e interpretação de imagens obtidas através de registos de radiação visível [14].

Durante as duas Guerras Mundiais a fotografia aérea teve um papel importante devido à sua utilização para fins de reconhecimento militar. Contudo, os maiores desenvolvimentos na DR verificaram-se na Segunda Grande Guerra, devido à utilização de instrumentos sensíveis à zona do infravermelho e sistemas de radar. Em meados dos anos 50 do século XX, estes sistemas começaram a ser utilizados por civis, nomeadamente os filmes de infravermelho no estudo da vegetação [13].

O termo “Detecção Remota” surgiu pela primeira vez no início dos anos 60 do século XX, num artigo não publicado escrito por Evelyn L. Pruitt, uma cientista do ramo de Geografia do *Office of Naval Research* (ONR) da marinha americana [11, 12]. Por essa altura, a era espacial estava a dar os primeiros passos com o lançamento dos satélites *Sputnik* (União Soviética), em 1957, e *Explorer I* (Estados Unidos da América – EUA), em 1958, e estavam a ser utilizados novos instrumentos (p.e., radiómetros), que permitiam recolher informação em zonas do espectro eletromagnético para além da região do visível. Assim, o termo “fotografia”, que significa “escrever com luz (visível)”, tornou-se demasiado limitador [11].

Apesar da existência de outros satélites até à data (p.e., o satélite meteorológico *Television Infrared Observation Satellite* – TIROS-1), o início da era moderna dos satélites de DR deu-se em 1972 com o *Earth Resources Technology Satellite* (ERTS) da *National Aeronautics and Space Administration* (NASA) [13]. O ERTS foi posteriormente renomeado para Landsat-1 e, atualmente (maio de 2012), a continuidade temporal do programa Landsat é garantida pelo Landsat-7 lançado em 1999 [11].

Desde o lançamento do Landsat-1 em 1972, o número e diversidade de sensores aumentou exponencialmente, permitindo medir vários tipos de variáveis e criar conjuntos de dados consistentes que abrangem longos intervalos de tempo [11, 12]. Estes dados são utilizados por uma vasta comunidade científica que os aplica em diversas áreas [11].

1.2.2 Radiação Eletromagnética

Os sensores de DR registam radiação eletromagnética (REM) proveniente de objetos da superfície terrestre. O Sol é a principal fonte de radiação para os sistemas de DR, emitindo REM em todo o espectro eletromagnético [11]. A radiação solar viaja pelo espaço, atravessa a atmosfera e, quando chega à Terra, uma parte é refletida de volta para o espaço e outra parte é absorvida, dando origem a transformações físicas do objeto (p.e., o aumento da temperatura) [12, 15]. A maioria dos sensores de DR mede a radiação solar refletida pela superfície e pela atmosfera e a radiação térmica emitida pela Terra (sensores passivos). Contudo, existem sensores que emitem a sua própria REM (sensores ativos), como é o caso do *Radio Detection And Ranging* (RADAR) ou do *Light Detection And Ranging* (LIDAR).

O conceito de REM foi formulado por James Clerk Maxwell como sendo uma onda eletromagnética que viaja pelo espaço à velocidade da luz. Esta onda é composta por dois campos de forças (Figura 1.2), um elétrico e um magnético, orientados ortogonalmente entre si, formando um plano perpendicular à direção de propagação [11]. A onda eletromagnética é caracterizada pelo comprimento de onda (λ) e pela frequência (f), cuja relação é dada por:

$$c = \lambda f \quad (1.1)$$

em que c é a velocidade da luz e é igual a $299792458 \text{ ms}^{-1}$.

A REM é constituída por fótons que, segundo a teoria quântica, recebem energia excedente resultante da transição dos eletrões atômicos para órbitas com menor nível energético [15]. A energia Q dos fótons é dada pela expressão:

$$Q = hf \quad (1.2)$$

onde $h = 6.626 \times 10^{-34} \text{ J s}$ é designada constante de Planck.

Resolvendo (1.1) em ordem a f e substituindo em (1.2) obtém-se a expressão:

$$Q = \frac{hc}{\lambda} \quad (1.3)$$

Pela expressão (1.3) verifica-se que a quantidade de energia transportada por cada fóton é inversamente proporcional ao comprimento de onda. Esta é uma propriedade importante em DR, pois sugere ser mais complicado detetar comprimentos de onda elevados, condicionando a resolução espacial das imagens. Assim, sensores que são sensíveis a bandas com comprimentos de onda mais elevados necessitam de observar elementos de terreno com maiores dimensões para que a REM que chega ao sensor seja detetável [15].

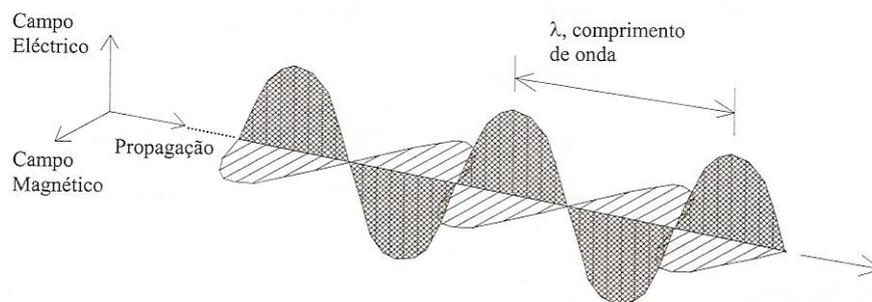


Figura 1.2. Propagação da Radiação Eletromagnética [15].

A REM encontra-se dividida num intervalo, designado por espectro eletromagnético (Figura 1.3). Este intervalo é contínuo e vai desde os pequenos comprimentos de onda dos raios gama até aos grandes comprimentos de onda das ondas rádio. O espectro divide-se ainda em vários subintervalos, sendo a radiação visível a única detetável pelo olho humano.

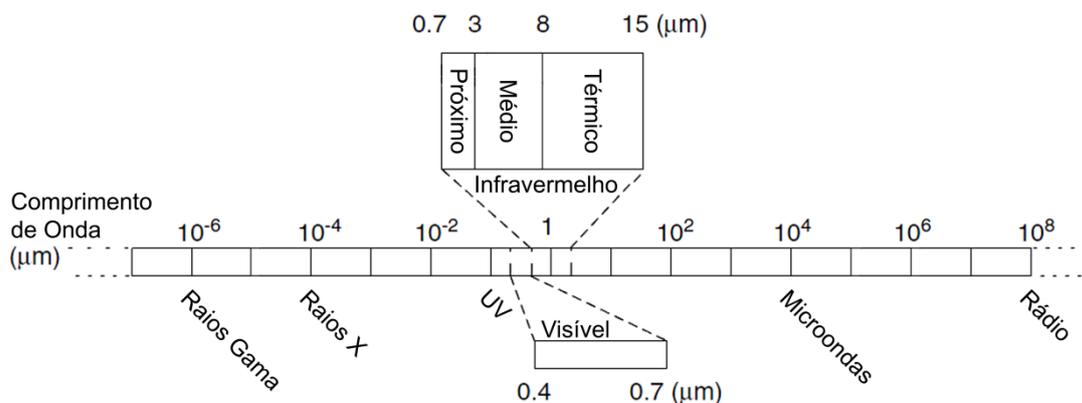


Figura 1.3. Espectro Eletromagnético (Adaptado de [16]).

A assinatura espectral (resposta espectral de um objeto, tal como registado ao longo de um intervalo de comprimentos de onda [12]) de um objeto depende das suas propriedades físicas e químicas. Por exemplo: a vegetação reflete bastante radiação no infravermelho próximo, o que permite diferenciar tipos de vegetação; a temperatura das superfícies terrestre e oceânica pode ser determinada a partir de bandas no infravermelho térmico; a radiação nas micro-ondas de comprimento de onda elevado (3-25 cm) é muito sensível à rugosidade das superfícies, permitindo detetar derrames de hidrocarbonetos nos oceanos.

Nem todas as zonas do espectro eletromagnético podem ser utilizadas, pois alguma da REM emitida pelo Sol ou pela Terra é absorvida pela atmosfera. As zonas em que a absorção por parte da atmosfera é baixa são designadas por janelas atmosféricas e são estas as utilizadas pelos sensores de DR. Este princípio não se aplica aos sensores que realizam observações em comprimentos de onda mais elevados (p.e., na zona das micro-ondas), pois não são afetados pela atmosfera.

1.2.3 A Detecção Remota Aplicada à Epidemiologia

Sendo o principal objetivo o estudo/monitorização da Terra, a DR tem sido também aplicada na área da saúde. A utilização de dados de DR tem como principal vantagem possibilitar a observação sinóptica de zonas onde a cobertura por estações de monitorização é escassa. O aumento na cobertura global e as melhorias na qualidade de dados (resoluções espacial, temporal, radiométrica e espectral) e no processamento de imagem têm permitido que os dados de DR sejam utilizados como ferramentas de estudos epidemiológicos desde 1970 [17]. Atualmente, encontram-se publicados inúmeros estudos que recorrem a estes dados para estudar vários tipos de doenças. Em Herbreteau *et al.* [17] a utilização da DR em estudos epidemiológicos de doenças infecciosas é revista para o período de 1970 a 2004. Depois de uma pesquisa exaustiva foram encontrados 86 estudos (podendo cada um possuir mais do que uma publicação), sendo que 59% incidiam em doenças provocadas por parasitas, como a malária (16%), a esquistossomose (16%) ou a tripanossomíase (10%). A grande maioria destes estudos baseia-se no facto da distribuição de um agente patogénico transmitido eficientemente ser, normalmente, limitada pela distribuição do vetor que o transmite [17-19]. Uma vez que o desenvolvimento de um determinado vetor (p.e., mosquitos) poder variar consoante a temperatura e a humidade do solo, as imagens obtidas por sensores de DR permitem estabelecer uma forte previsão dos padrões de distribuição da doença provocada pela transmissão do agente patogénico. Os restantes 41% dos estudos compreendem outras

doenças, como bacterianas (p.e., doença de Lyme ou cólera) ou virais (p.e., a dengue ou a febre do Vale de Rift).

As imagens mais utilizadas pertencem aos sensores *Advanced Very High Resolution Radiometer* (AVHRR), dos satélites da *National Oceanic and Atmospheric Administration* (NOAA), e *Multispectral Scanner* (MSS) e *Thematic Mapper* (TM), dos satélites Landsat. Estes sensores permitem gerar vários parâmetros ambientais que, juntamente com dados recolhidos no campo, possibilitam a identificação de áreas de risco. Os parâmetros ambientais mais comuns são o *Normalized Difference Vegetation Index* (NDVI) (permite medir as variações da biomassa), a temperatura do solo (*land surface temperature* – LST) e da água e mapas de ocupação de solo. Em 2002, Rogers *et al.* [19] realizaram um estudo sobre a incidência da malária em África, recorrendo a imagens da temperatura do solo e do ar, NDVI (AVHRR) e precipitação (Meteosat) para mapear a distribuição de cinco espécies de mosquitos do tipo *Anopheles* (vetor da malária). Foi encontrada uma relação entre o número de admissões mensais de crianças com malária em três hospitais no Quênia e o NDVI do mês anterior. Num estudo mais recente, Dambach *et al.* [18] recorreram a dados de sensores com maior resolução espacial para estudar a variação local da epidemiologia da malária na província de Kossi, na zona noroeste do Burkina Faso. Foram utilizados dados do *Satellite Pour l'Observation de la Terre 5* (SPOT-5) para determinar vários índices com 10 m de resolução espacial, como o já referido NDVI e o *Normalized Difference Pond Index* (NDPI), que permite classificar lagos [11]. A precipitação, o LST e a altitude foram determinadas usando o *Tropical Rainfall Measurement Mission* (TRMM), o *Moderate Resolution Imaging Spectroradiometer* (MODIS) e o *Shuttle Radar Topography Mission* (SRTM), respetivamente. Os autores verificaram que a presença do vetor se correlaciona negativamente com o LST e positivamente com a quantidade de chuva acumulada nos 15 dias anteriores e com o NDPI dentro de um raio de 500 m dos locais de recolha de dados.

A utilização da DR em epidemiologia não se resume apenas a doenças infecciosas, sendo também utilizada no estudo de doenças oncológicas derivadas da exposição contínua a fatores ambientais. Como é referido em Maxwell *et al.* [20], alguns estudos sobre cancro de pele utilizam dados da coluna total de ozono da estratosfera do sensor *Total Ozone Mapping Spectrometer* (TOMS) para estimar os níveis de radiação ultravioleta. Os mapas de ocupação derivados a partir do Landsat ou mesmo de fotografias aéreas, também têm sido utilizados na investigação de relações entre a exposição a pesticidas e os cancros da bexiga e da mama e o linfoma não-Hodgkin (LNH) [20].

Na última década, com o lançamento de novos sensores e desenvolvimento de novos algoritmos, têm sido verificados vários avanços que permitiram alargar a aplicação da DR a outras áreas. Dados recolhidos por sensores como o MODIS ou o *Multi-angle Imaging*

SpectroRadiometer (MISR) têm sido utilizados para estimar o *Aerosol Optical Thickness* (AOT), devido ao impacto que os aerossóis têm no clima e no ambiente [21]. O AOT reflete o nível de radiação solar absorvida pelos aerossóis em toda a coluna atmosférica [22]. Quanto mais elevada for a concentração de aerossóis menos radiação chega ao sensor. O AOT tem um grande potencial na determinação de estimativas espaço-temporais da qualidade do ar [23], existindo vários estudos que o relacionam com a concentração de partículas em suspensão (*particulate matter* – PM) de diâmetros aerodinâmicos inferiores a 10 μm (PM₁₀) e 2.5 μm (PM_{2.5}) medidos à superfície por estações de monitorização de poluentes [21-26]. As PM de menor dimensão são dos poluentes que mais afeta a qualidade do ar em zonas urbanas e rurais, podendo provocar doenças respiratórias e pulmonares e até causar mortes prematuras [22, 23]. Uma vez que as estações que medem concentrações de PM estão maioritariamente concentradas nas grandes cidades, existe muito interesse em poder estimá-las para as restantes zonas. No caso de Portugal, grande parte das estações encontra-se situada no Porto e em Lisboa.

A relação entre o AOT e as PM não é direta, sendo condicionada por vários fatores [24]: o AOT dado pelos satélites refere-se a uma coluna atmosférica estimada, constituída por vários tipos de aerossóis (derivados de processos naturais, atividade humana e interação entre ambos), enquanto as PM consistem em massa seca; a relação entre o AOT e as PM varia em função da localização, pois depende de fatores como fontes de emissão de aerossóis ou condições meteorológicas; a baixa resolução espacial dos produtos de AOT (normalmente 10 km) não permite captar as variações nas concentrações de PM em curtas distâncias; as estações medem as PM continuamente (p.e., intervalos de 1 hora), enquanto uma imagem de AOT diz respeito a um único instante de um determinado dia. Gupta *et al.* [23] verificaram que a correlação entre o AOT e as PM_{2.5} em várias cidades do mundo era elevada em zonas sem nuvens, da Camada Limite Atmosférica (para valores de AOT maiores que 0.1) e com baixa humidade relativa. Kumar *et al.* [24] modificaram o algoritmo desenvolvido pela NASA para a determinação do AOT MODIS para melhorar a resolução espacial de 10 km para 2 km, num estudo realizado na cidade de Cleveland do estado de Ohio dos EUA. Com o aumento da resolução espacial, o coeficiente de correlação aumentou, permitindo assim captar a variabilidade espacial provocada por fontes de emissão locais.

Apesar da maioria das estimativas da qualidade do ar, determinadas a partir de satélite, se focarem nas concentrações de PM à superfície, outros poluentes, como o monóxido de carbono (CO), o dióxido de azoto (NO₂) ou o ozono (O₃), têm sido estimados ultimamente, devido sobretudo ao aumento da resolução espacial dos sensores mais recentes [27].

Nesta secção foram abordados alguns exemplos da utilização da DR na epidemiologia, podendo ser encontrada informação mais detalhada em [17, 20, 27]. Na Tabela 1.1 são apresentados alguns dos principais sensores que podem ser utilizados nesta área e alguns dos seus produtos *standard*.

Tabela 1.1. Dados de satélite.

Satélite	Sensor	Parâmetros	Período	RE	RT
		NDVI		250 m	16 dias
		LST		1 km	12 horas
		Áreas aridas		500 m	1 mês
Terra/Aqua	MODIS	Perfis de Temperatura e Humidade	2000-	5 km	12 horas
		Temperatura do solo		1 km	12 horas
		AOT		10 km	12 horas
		Mapas de ocupação do solo		500 m	1 ano
SPOT 4 e 5	Vegetation	NDVI	1998-	1 km	10 dias
TRMM	PR	Precipitação	1997-	5 km	1 dia
ENVISAT	MERIS	AOT	2002-	1 km	3 dias
	SCIAMACHY	NO ₂ , SO ₂ , CO, O ₃		60 x 30 km	
Terra	MOPITT	CO	2000-	22 km	12 horas
	MISR	AOT		18 km	
Aqua	AIRS	SO ₂ , CO	2002-	14 km	12 horas
Aura	OMI	NO ₂ , SO ₂ , O ₃ , AOT	2004-	24 x 13 km	1 dia
	TES	CO, O ₃		8 x 5 km	
MetOp	GOME-2	NO ₂ , SO ₂ , O ₃	2006-	80 x 40 km	1 dia
	IASI	CO, O ₃		12 x 12 km	
N/A	SRTM	Altitude	N/A	90 m	N/A

SPOT – *Satellite Pour l’Observation de la Terre*;

TRMM – *Tropical Rainfall Measurement Mission*;

MODIS – *Moderate Resolution Imaging Spectroradiometer*;

PR – *Precipitation Radar*;

MERIS – *MEdium Resolution Imaging Spectrometer*;

SCIAMACHY – *SCanning Imaging Absorption SpectroMeter for Atmospheric CartographY*;

MOPITT – *Measurements of Pollution in the Troposphere*;

MISR – *Multi-angle Imaging SpectroRadiometer*;

AIRS – *Atmospheric Infrared Sounder*;

OMI – *Ozone Monitoring Instrument*;

TES – *Tropospheric Emission Spectrometer*;

GOME-2 – *Global Ozone Monitoring Experiment-2*;

IASI – *Infrared Atmospheric Sounding Interferometer*;

SRTM – *Shuttle Radar Topography Mission*;

RE – Resolução Espacial;

RT – Resolução Temporal.

1.3 Asma

A asma é uma doença inflamatória crónica das vias aéreas que leva a episódios recorrentes de pieira, dificuldade respiratória, aperto torácico e tosse, particularmente noturna e no início da manhã. Estima-se que esta doença afete cerca de 300 milhões de pessoas em todo o mundo, sendo responsável por 250 mil mortes por ano [28]. A doença apresenta ainda um forte impacto nas populações infantil e juvenil [29-33]. Apesar de tanto na perspetiva do paciente como da sociedade, os custos para controlar a asma parecerem elevados, se ela não for tratada corretamente os custos são ainda mais elevados [28]. Para além disso, a asma é também um fator importante de internamentos hospitalares e causa limitações consideráveis a nível físico, emocional, social e profissional, interferindo com a atividade normal e qualidade de vida, tanto dos pacientes como das suas famílias [30]. Segundo o programa *Global INitiative for Asthma* (GINA), que tem como principal objetivo produzir recomendações para o controlo da asma (última atualização em 2011), a prevalência da asma a nível mundial varia entre 1% e 18% [28]. Em Portugal, vários estudos estimam que a prevalência varie entre 3.3% e 15%. Contudo, as metodologias utilizadas para obter estes resultados são de precisão questionável [30].

O GINA identifica vários fatores que influenciam o desenvolvimento e/ou a expressão da asma. Estes podem ser divididos em dois grupos: fatores intrínsecos (principalmente genéticos) e fatores extrínsecos ou ambientais (Tabela 1.2).

Tabela 1.2. Fatores que influenciam o desenvolvimento e/ou expressão da asma.

Fatores intrínsecos
<ul style="list-style-type: none">• Genéticos (p.e., genes com predisposição para atopia ou hiperresponsividade das vias aéreas);• Obesidade;• Sexo.

Fatores extrínsecos
<ul style="list-style-type: none">• Alergénios<ul style="list-style-type: none">• Interior: Ácaros domésticos, animais de pêlo (cães, gatos, ratos), alergénios de baratas, fungos (p.e., bolor e leveduras);• Exterior: Pólen, fungos (p.e., bolor e leveduras);• Infecções (principalmente virais);• Sensibilizantes ocupacionais;• Fumo do tabaco (fumadores passivos e ativos);• Poluição do ar (no interior ou exterior);• Hábitos alimentares;• Condições meteorológicas.

Ao longo dos anos, vários estudos têm sido desenvolvidos em todo o mundo de forma a aferir os fatores ambientais com maior influência na expressão da asma. Muitos dos estudos focam-se na exposição a poluentes como O_3 , $PM_{2.5}$ e PM_{10} , NO_2 , CO e dióxido de enxofre (SO_2). Contudo, existem discrepâncias nos resultados [34], fazendo com que o papel dos poluentes como causadores da asma continue a ser controverso [28]. Variações no desenho do estudo e diferentes métodos de modelação podem contribuir para a falta de consistência nos resultados [29, 30, 35], enquanto a escolha do *outcome* (p.e., prevalência, consultas de urgência, internamentos, mortalidade), a escala espaço-temporal (p.e., nível de agregação dos dados, resolução temporal) e a modelação da exposição (p.e., dados de estações de monitorização ou dados pessoais de exposição) podem influenciar as relações observadas [34].

Uma grande quantidade de estudos transversais e de séries temporais reportaram variações na função pulmonar, sintomas respiratórios, e/ou recorrência a medicamentos ou cuidados de saúde após exposição a concentrações elevadas de O_3 [29, 32, 35-42], PM_{10} e/ou $PM_{2.5}$ [29, 31, 32, 34, 37-46], NO_2 [34, 37, 39-41, 43, 47], CO [34, 41, 47] e SO_2 [37, 40, 43, 47, 48]. Para além dos poluentes, outros estudos focam-se na associação com concentrações de pólen [49-51] e com as condições meteorológicas [52-

55], nomeadamente, com alterações na temperatura e na humidade relativa do ar e ocorrências de nevoeiro e trovoada (pode provocar um aumento repentino nas concentrações de pólen [52]).

1.4 Objetivo

A maioria dos estudos da asma não recorre à análise espacial de informação. Os estudos não espaciais utilizam dados recolhidos por uma única estação de monitorização ou valores médios calculados a partir de várias estações, podendo assim conduzir a erros. Assim, o principal objetivo deste trabalho consiste em estudar a asma em Portugal Continental, relacionando um conjunto de parâmetros ambientais, determinados maioritariamente com observações de satélites, com internamentos em hospitais públicos entre 2003 e 2008.

Especificamente, os objetivos deste estudo são:

- Determinar mensalmente a temperatura e humidade do ar ao nível da superfície (T_a e HR, respetivamente), quantidade de vegetação (representada pelo NDVI) e concentração dos poluentes NO_2 e PM_{10} para Portugal Continental;
- Avaliar as vantagens e limitações da utilização de dados de satélite pré-processados com calibração global neste tipo de estudos;
- Verificar se existe alguma relação entre as cinco variáveis ambientais consideradas e a expressão da asma.

Capítulo 2

Dados

As cinco variáveis ambientais consideradas neste estudo foram determinadas utilizando uma grande variedade de dados. A T_a foi determinada com dois conjuntos de dados: perfil de temperatura do ar (PT_a) do produto *standard* MOD07 do sensor MODIS e o Modelo Digital de Elevação (MDE) do *Shuttle Radar Topography Mission* (SRTM); medições de estações meteorológicas. A coluna total de vapor de água (*Total Column Water Vapor* - TCWV) foi utilizada juntamente com a T_a e com o MDE para determinar a HR. À semelhança da T_a , a HR foi também determinada com medições de estações meteorológicas. O NDVI do produto *standard* MOD13A3 do MODIS foi utilizado como medida representante da quantidade de vegetação. Na determinação dos poluentes NO_2 e PM_{10} recorreu-se a medições realizadas por estações de qualidade do ar e outros tipos de dados: T_a e HR (determinadas anteriormente), AOT a 550 nm (AOT550) do produto *standard* MOD04 do MODIS, CORINE Land Cover 2006 (CLC2006), Carta Administrativa Oficial de Portugal 2011 (CAOP2011) e duas redes viárias.

De forma a estudar espacialmente a asma em Portugal Continental, recorreu-se à CAOP2011 para relacionar as cinco variáveis ambientais com a frequência de internamentos hospitalares, obtida a partir de uma BD de episódios hospitalares. Os internamentos foram reduzidos à população total, através de informação sobre a população residente em cada concelho.

2.1 Moderate Resolution Imaging Spectroradiometer (MODIS)

O *Earth Observing System* (EOS) é um programa da NASA que disponibiliza observações globais da superfície terrestre, atmosfera e oceanos durante um período de

15 ou mais anos [56]. O EOS é constituído por uma série de satélites em órbitas polares e de baixa inclinação, uma rede de computadores para processamento, armazenamento e distribuição de dados e equipas de cientistas/investigadores que os analisam [57]. O primeiro satélite deste programa foi lançado no final dos anos 90 [56].

Face ao contínuo crescimento de uma população humana cada vez mais afetada pela variabilidade natural do ambiente e com uma participação ativa na evolução da Terra foram desenvolvidos, nos últimos anos, vários sensores EOS com o intuito de servir diferentes áreas de estudo. Um dos sensores mais importantes é o MODIS que, em comparação com os seus antecessores NOAA AVHRR ou Nimbus *Coastal Zone Color Scanner* (CZCS), oferece uma combinação única de características com melhorias nas resoluções espectral e espacial e com uma maior frequência de coberturas globais da Terra [58]. Tratando-se de um sensor versátil, permite gerar vários tipos de produtos, desde perfis de temperatura e humidade, a índices de vegetação e mapas de temperatura das superfícies oceânicas e terrestres.

O MODIS foi lançado no dia 18 de dezembro de 1999, a bordo do satélite Terra (passagem descendente pelo equador às 10h30 UTC), e no dia 4 de maio de 2002, a bordo do satélite Aqua (passagem ascendente pelo equador às 13h30 UTC) [58], tendo como objetivo fornecer observações que permitam modelar e prever mudanças nas dinâmicas globais (atmosférica, oceânica e terrestre) da Terra [56, 59]. O MODIS possui um ângulo de visão de $\pm 55^\circ$ que, combinado com uma órbita heliosíncrona de 705 km de altitude, permite adquirir informação com uma largura de varrimento de 2330 km e alcançar uma cobertura global da Terra a cada 1-2 dias. O sensor obtém imagens com uma resolução radiométrica de 12 bits da radiação refletida (solar) e emitida (térmica) por todas as regiões do globo [11]. As observações são realizadas em 36 bandas espectrais, que vão desde o visível até ao infravermelho térmico (0,4-14,4 μm), com uma resolução espacial que varia entre 250 m (bandas 1 e 2), 500 m (bandas 3 à 7) e 1 km (bandas 8 à 36) [56, 58]. As características das bandas e respetivas aplicações encontram-se sumariadas na Tabela 2.1 [56, 58]. Em Xiong e Barnes [60] é feita uma descrição aprofundada sobre as características do MODIS e respetiva calibração.

O projeto MODIS fornece à comunidade científica uma vasta quantidade de produtos *standard* calibrados globalmente. Como já foi referido, foram utilizados três produtos *standard* derivados do sensor MODIS: produto diário de nível 2 de Perfis Atmosféricos da coleção 005, derivado das observações dos satélites Aqua (MYD07) e Terra (MOD07); produto diário de nível 2 de Aerossóis da coleção 051, derivado das observações dos satélites Aqua (MYD04) e Terra (MOD04); produto mensal de nível 3 de Índices de Vegetação da coleção 005, derivado das observações do satélite Terra (MOD13A3). A

partir daqui, os termos MOD07 e MOD04 serão utilizados como referência aos produtos derivados das observações de ambos os satélites.

Tabela 2.1. Características e aplicações das bandas do MODIS.

Nº	Resolução espectral (μm)	Resolução espacial	Aplicações
1	0.620 – 0.670	250 x 250 m	Limites de Terra/Nuvens/Aerossóis
2	0.841 – 0.876	250 x 250 m	
3	0.459 – 0.479	500 x 500 m	Propriedades de Terra/Nuvens/Aerossóis
4	0.545 – 0.565	500 x 500 m	
5	1.230 – 1.250	500 x 500 m	
6	1.628 – 1.652	500 x 500 m	
7	2.105 – 2.155	500 x 500 m	
8	0.405 – 0.420	1 x 1 km	Cor do oceano, fitoplâncton, biogeoquímica
9	0.438 – 0.448	1 x 1 km	
10	0.483 – 0.493	1 x 1 km	
11	0.526 – 0.536	1 x 1 km	
12	0.546 – 0.556	1 x 1 km	
13	0.662 – 0.672	1 x 1 km	
14	0.673 – 0.683	1 x 1 km	
15	0.743 – 0.753	1 x 1 km	
16	0.862 – 0.877	1 x 1 km	
17	0.890 – 0.920	1 x 1 km	
18	0.931 – 0.941	1 x 1 km	
19	0.915 – 0.965	1 x 1 km	
20	3.600 – 3.840	1 x 1 km	Temperatura das superfícies/nuvens
21	3.929 – 3.989	1 x 1 km	
22	3.929 – 3.989	1 x 1 km	
23	4.020 – 4.080	1 x 1 km	
24	4.433 – 4.498	1 x 1 km	Temperatura atmosférica
25	4.482 – 4.549	1 x 1 km	
26	1.360 – 1.390	1 x 1 km	Nuvens Cirrus

Tabela 2.1. (Continuação).

Nº	Resolução espectral (μm)	Resolução espacial	Aplicações
27	6.535 – 6.895	1 x 1 km	
28	7.175 – 7.475	1 x 1 km	Vapor de água
29	8.400 – 8700	1 x 1 km	
30	9.580 – 9.880	1 x 1 km	Ozono
31	10.780 – 11.280	1 x 1 km	Temperatura das superfícies/nuvens
32	11.770 – 12.270	1 x 1 km	
33	13.185 – 13.485	1 x 1 km	
34	13.485 – 13.785	1 x 1 km	Altitude das nuvens
35	13.785 – 14.085	1 x 1 km	
36	14.085 – 14.385	1 x 1 km	

Os parâmetros utilizados de cada produto são apresentados na Tabela 2.2. Todos os ficheiros dos produtos de nível 2 que abrangessem a área e o intervalo de tempo em estudo foram descarregados gratuitamente a partir da página <http://ladsweb.nascom.nasa.gov/data/search.html>, enquanto os ficheiros do produto de nível 3 foram descarregados da página https://lpdaac.usgs.gov/get_data/data_pool. Devido à órbita dos satélites, os produtos de nível 2 derivados do Terra e do Aqua encontram-se disponíveis, respetivamente, entre as 10h30-12h e 13h-15h, durante o dia, e entre as 22h30-00h e 1h-3h, durante a noite. Nas secções seguintes será efetuada uma breve descrição sobre os algoritmos de produção dos parâmetros aqui considerados, assim como sobre o formato em que os produtos são distribuídos.

Tabela 2.2. Parâmetros MODIS utilizados no âmbito deste trabalho.

Parâmetro	Produto	Resolução espacial	Resolução Temporal	Unidades	Dimensões ^a
PT _a	MOD07	5 km	Diária	K	406 x 270 x 20 ^b
TCWV	MOD07	5 km	Diária	cm	406 x 270
AOT550	MOD04	10 km	Diária	N/A	203 x 135
NDVI	MOD13A3	1 km	Mensal	N/A	1200 x 1200

^a As dimensões dos produtos de nível 2 são valores típicos (406 x 270 e 203 x 135).

^b PT_a é dado com 20 níveis de pressão que variam entre 1000 hPa e 5 hPa.

2.1.1 Perfis Atmosféricos (MOD07)

O MOD07 é um produto diário com uma resolução espacial de 5 km, constituído por vários parâmetros, nomeadamente, os perfis de temperatura e humidade, coluna total de ozono, TCWV e índices de estabilidade atmosférica. Estes parâmetros podem ser utilizados para corrigir o efeito atmosférico em alguns produtos MODIS (como as temperaturas das superfícies terrestre e oceânica ou as propriedades dos aerossóis oceânicos), bem como para caracterizar o estado da atmosfera em estudos globais sobre gases do efeito estufa [57].

O algoritmo utilizado para gerar os perfis tira proveito do facto do MODIS possuir várias bandas espectrais. O algoritmo aplica uma regressão estatística sintética com uma opção para a obtenção de uma relação não linear, utilizando como principais variáveis independentes medidas de radiância em 11 bandas térmicas (entre 4,5 μm e 14,2 μm) efetuadas em condições de céu limpo (sem nuvens) sobre terra e oceano. Os coeficientes da regressão sintética são calculados com base num modelo de rápida transferência radiativa, que recebe como *input* observações do estado da atmosfera de um conjunto de radiossondas com distribuição global. A resolução espacial das bandas é reduzida para 5 km e cada pixel só é utilizado no procedimento se pelo menos 5 dos 25 píxeis de 1 km que o compõem forem classificados com ausência de nuvens. Para além disso, os perfis são gerados inicialmente com 101 níveis de pressão, sendo posteriormente guardados apenas com 20 níveis (entre 1000 hPa e 5 hPa) [61]. Antes da conversão para 20 níveis de pressão, o TCWV é determinado pela integração do perfil de humidade através da coluna atmosférica [61].

A validação do MOD07 é realizada comparando-o com produtos de outros sistemas de observação. As comparações são efetuadas com base em medições recolhidas com instrumentação terrestre, nomeadamente na *Atmospheric Radiation Measurement (ARM) Cloud and Radiation Testbed (CART)*, no sul das Grandes Planícies em Oklahoma, e em estações *Global Positioning System (GPS)*. As observações de outros satélites também são utilizadas para validação, como é o caso do *Geostationary Operational Environmental Satellites (GOES)*, do *Atmospheric Infrared Sounder (AIRS)*, do *Special Sensor Microwave/Imager (SSM/I)* e do TOMS [61].

2.1.2 Aerosol Optical Thickness (MOD04)

O MOD04 é um produto diário com uma resolução espacial de 10 km, que contém vários parâmetros de aerossóis. As principais finalidades deste produto são o balanço radiativo

e aplicações climáticas. Contudo, é de esperar que sejam relevantes em aplicações hidrológicas, oceanográficas e de qualidade do ar [62].

Os parâmetros do MOD04 encontram-se divididos em três grandes grupos: terra; oceano; combinação de terra e oceano. O AOT é produzido em três comprimentos de onda para terra (470, 550 e 660 nm), em sete para oceano (470, 550, 660, 860, 1240, 1630 e 2130 nm) e em um para a combinação de terra e oceano (550 nm). Para além disso, existem ainda outros parâmetros, como a grelha com a classificação dos aerossóis (misto, poeira, sulfato, fumo e fumo altamente absorvente), dada sobre terra, ou a grelha com a distribuição de aerossóis por tamanho, dada sobre oceano.

O algoritmo de produção do MOD04 utiliza duas abordagens independentes para determinar os aerossóis sobre a terra e sobre o oceano, devido às diferenças inerentes à interação da radiação solar com estas superfícies. Ambas as abordagens utilizam refletâncias calibradas do MODIS nos comprimentos de onda referidos anteriormente e recorrem a comprimentos de onda adicionais para identificar nuvens e sedimentos de rios [63]. O princípio básico deste algoritmo consiste em comparar as refletâncias observadas do topo da atmosfera com refletâncias pré-calculadas de uma tabela de transformação (*look up table*) para uma variedade de parâmetros de aerossóis e da superfície [62]. O produto final corresponde ao conjunto de propriedades de aerossóis associado à refletância da tabela de transformação que melhor se ajusta (método dos mínimos quadrados) à refletância medida pelo MODIS. As refletâncias na tabela de transformação são calculadas a partir de modelos de aerossóis que representam as propriedades dos aerossóis numa coluna vertical [64]. O AOT é determinado com uma resolução de 10 km, o que corresponde a 20 x 20 pixéis de 500 m (a resolução das bandas 1 e 2 tem que ser reduzida). Para cada conjunto de 400 pixéis de 500 m tem que existir no mínimo 12 pixéis válidos (não contaminados por nuvens, neve, corpos de água interiores, etc.) [62].

O AOT é validado recorrendo a dados de uma rede de estações global de monitorização de aerossóis, a *Aerosol Robotic Network* (AERONET). Os detalhes sobre o conceito e a descrição da AERONET estão disponíveis em [65].

2.1.3 Índices de Vegetação (MOD13A3)

O MOD13A3 é um produto mensal com uma resolução espacial de 1 km, contendo dois índices de vegetação, o já referido NDVI e o *Enhanced Vegetation Index* (EVI). Este último minimiza as variações das características das copas e mantém a sensibilidade sobre condições de vegetação densa. O NDVI do MODIS complementa o NDVI derivado das observações do AVHRR, disponível desde 1981 [66].

O NDVI foi publicado em 1974 por Rouse *et al.* [67] e apresenta uma boa sensibilidade às alterações na cobertura de vegetação. O índice consiste na diferença normalizada entre os sinais medidos no infravermelho próximo (*Near-Infrared* – NIR) e o vermelho (Red), tal como apresentado na equação (2.1).

$$\text{NDVI} = \frac{(\text{NIR} - \text{Red})}{(\text{NIR} + \text{Red})} \quad (2.1)$$

A normalização é realizada para minimizar várias formas de ruído (p.e., diferentes condições de iluminação, sombras derivadas de nuvens ou variações na topografia) presentes numa ou em várias imagens. O NDVI varia entre -1 e 1, sendo que os valores mais altos correspondem a valores elevados de biomassa. Normalmente, os valores negativos são ajustados a 0, passando o NDVI a variar entre 0 e 1.

Os produtos mensais são gerados a partir de um algoritmo que recolhe todos os produtos MOD13A2 (resolução espacial de 1 km e temporal de 16 dias) que abrangem um determinado mês e calcula uma média temporal pesada, caso os dados não estejam contaminados por nuvens, ou seleciona o valor máximo de NDVI, caso os dados estejam contaminados. O MOD13A2 é produzido a partir das imagens diárias de NDVI calculadas com as bandas 1 e 2 do MODIS, aplicando a expressão (2.1). O método para determinar as imagens de 16 dias é baseado no utilizado pelos produtos NDVI do AVHRR e consiste em escolher para cada pixel da imagem de 16 dias o valor máximo observado durante esse período (*Maximum Value Composite* – MVC) [66].

A validação dos índices é efetuada comparando-os com dois modelos de transferência de radiação de copa (SAIL e Myneni), medições *in-situ* e a partir de avião e dados provenientes de outros sensores, como o Landsat TM ou o AVHRR [68].

2.1.4 Formato dos Dados MODIS

Os produtos MODIS encontram-se no formato *Hierarchical Data Format – Earth Observing System* (HDF-EOS). Este formato é baseado no HDF e é um *standard* para o armazenamento de dados derivados de missões EOS [69].

No caso dos produtos de nível 2 (MOD04 e MOD07), é introduzido o conceito de HDF-EOS *Swath*. Este conceito é baseado na típica aquisição por varrimento, onde o instrumento adquire os dados linha a linha (Figura 2.1) [70].

Na estrutura *Swath* cada parâmetro é guardado numa matriz, em que cada linha corresponde a uma linha varrida pelo sensor. Associada à matriz de dados, existem outras duas matrizes (Latitude e Longitude em WGS84) com o mesmo tamanho, que

definem a posição de cada elemento da matriz de dados. Para além disso, o tempo de aquisição de cada linha é guardado num vetor (ou matriz) *Time*, normalmente em segundos.

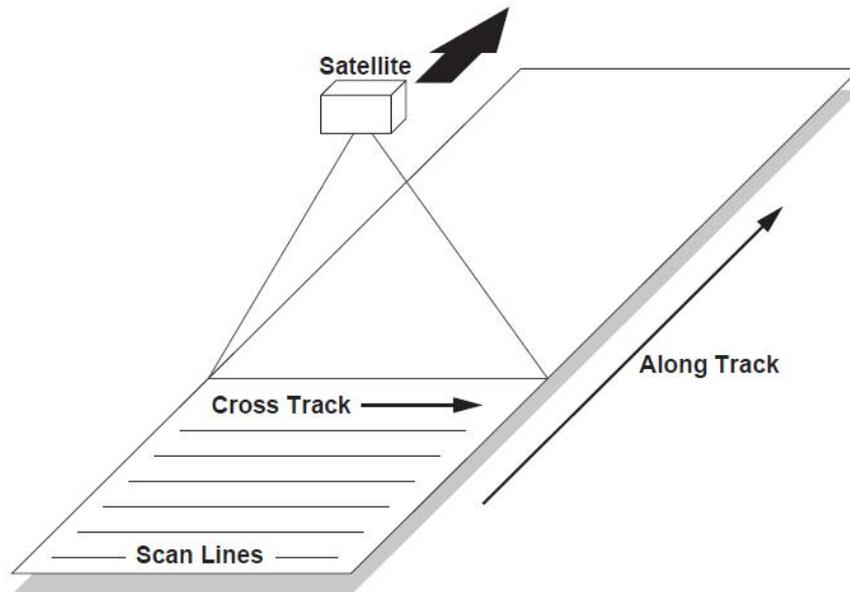


Figura 2.1. *Swath* típica de um satélite com sensor de varrimento (Adaptado de [71]).

A *Swath* é constituída essencialmente por um campo com os parâmetros (p.e., diferentes variáveis ou diferentes bandas de um sensor), o *Data Fields*, e por outro com as posições, o *Geolocation Fields*. O tempo de aquisição de cada linha encontra-se guardado num dos dois grupos. Os ficheiros de dados contêm ainda atributos associados, como é o caso dos metadados (*SwathMetadata*) [70].

Cada ficheiro dos produtos de nível 2 MOD04 e MOD07 contém 5 minutos de recolha de dados, o que resulta em ficheiros com dimensão igual a 270 x 406 (≈ 2330 km / 5 km) pixéis, para o MOD07, e 135 x 203 (≈ 2330 km / 10 km) pixéis, para o MOD04. Os ficheiros possuem ainda uma nomenclatura uniforme do tipo:

MODxx_L2.Ayyyyddd.hhmm.vvv.yyyydddhhmmss.hdf

em que:

MODxx_L2 – código do produto

Ayyyyddd – ano e dia juliano da aquisição dos dados

hhmm – horas e minutos do início da aquisição dos dados

vvv – versão da coleção (a cada nova coleção os dados são reprocessados com alterações nos algoritmos)

yyyydddhhmss – data e hora de produção do ficheiro

hdf – extensão do ficheiro

No caso dos produtos de nível 3 (MOD13A3), os dados são guardados em formato grelha HDF-EOS. O princípio é idêntico ao da *Swath*, em que existe um grupo *Data Fields* que contém os parâmetros, embora não exista o grupo *Geolocation Fields* [71]. Uma vez que os dados se encontram projetados, as grelhas contêm as informações sobre a projeção e os seus parâmetros relevantes, permitindo assim definir a posição de cada pixel.

Nos ficheiros do produto MOD13A3 é utilizada uma projeção sinusoidal, considerando uma esfera de raio igual a 6371007.181 m.

O ponto central ou origem das coordenadas retangulares coincide com o ponto de interseção entre o meridiano de Greenwich e o equador, ou seja, possui latitude e longitude igual a 0. Cada ficheiro abrange uma quadrícula de $10^{\circ} \times 10^{\circ}$, das 648 em que a Terra é dividida (Figura 2.2) e, para o caso do MOD13A3, 1200 x 1200 pixéis de dimensão. As quadrículas h17v04 e h17v05 abrangem a totalidade de Portugal Continental.

Os ficheiros possuem a seguinte nomenclatura:

MODxxAx.Ayyyyddd.hmmvnn.vvv.yyyydddhhmss.hdf

em que:

MODxxAx – código do produto

Ayyyyddd – ano e dia juliano da aquisição (no caso do MOD13A3 representa o primeiro dia do mês)

hmmvnn – código da quadrícula

vvv – versão da coleção

yyyydddhhmss – data e hora de produção do ficheiro

hdf – extensão do ficheiro

Os algoritmos utilizados na criação dos produtos MODIS, abordados anteriormente, são descritos detalhadamente em [61-64, 68] e os respetivos indicadores de qualidade encontram-se em [66, 72].

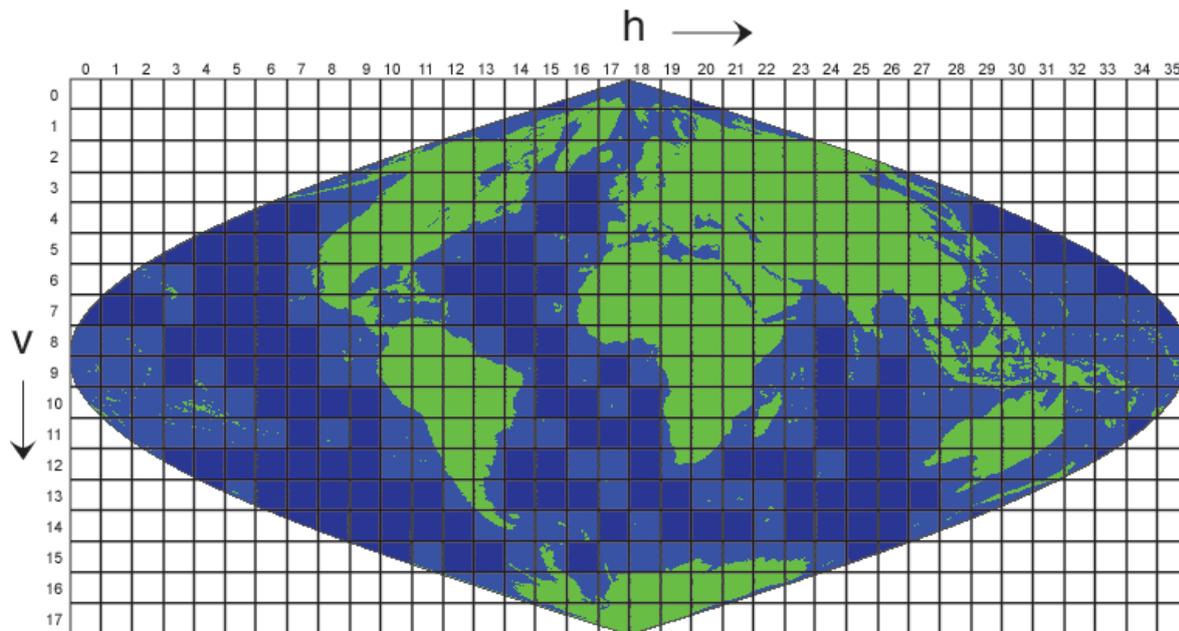


Figura 2.2. Projecção Sinusoidal da Terra, dividida em quadrados de 10° por 10° [73].

2.2 Modelo Digital de Elevação

A missão SRTM foi realizada pela plataforma espacial *Space Shuttle Endeavour* (SSE) durante 11 dias em fevereiro de 2000 e resultou de uma colaboração entre a NASA, a *National Geospatial-Intelligence Agency* (NGA) (anteriormente conhecida por *National Imagery and Mapping Agency* – NIMA) e as agências espaciais Alemã e Italiana. O objetivo desta missão era recolher informação por interferometria radar sobre a superfície terrestre para gerar um MDE quase global [74].

Mapear a topografia de uma superfície a partir de interferometria radar consiste na aquisição de duas imagens radar a partir de dois pontos (ligeiramente) diferentes. A diferença entre essas duas imagens permite calcular a altitude da superfície [11, 75]. De forma a poder adquirir duas imagens radar de perspetivas diferentes, a missão SRTM possuía duas antenas, uma colocada no corpo do SSE e outra na extremidade de um braço com 60 m (base). O SSE emitia um feixe de ondas radar que, ao chegarem à superfície terrestre, eram refletidas em várias direções, sendo posteriormente recebidas pelas duas antenas (Figura 2.3).

Os dados do SRTM estão organizados em imagens individuais que cobrem uma área de 1° por 1° de latitude e longitude. Os três tipos de resolução do SRTM são $1''$, $3''$ e $30''$, respetivamente denominados por dados SRTM1, SRTM3 e SRTM30. Como $1''$ ao longo do equador corresponde a, aproximadamente, 30 m, o SRTM1 e o SRTM3 são conhecidos como os dados de 30 m e 90 m, respetivamente. Cada ficheiro SRTM1 possui 3601×3601 píxeis, enquanto cada ficheiro SRTM3 possui 1201×1201 píxeis [74].

O SRTM com resolução de 30 m encontra-se disponível para os EUA e territórios associados e o SRTM com resolução de 90 m para todas as regiões situadas entre 60° N e 56° S de latitude, cobrindo assim cerca de 80% da área total terrestre [74].

Para validar os dados do SRTM, a NGA e a NASA realizaram uma recolha de dados à escala global, verificando-se que o SRTM cumpria com as suas especificações (p.e., erro vertical absoluto menor que 16 m e erro vertical relativo menor que 10 m) [76].

Os dados encontram-se em formato *height* (HGT) e podem ser descarregados gratuitamente a partir da página http://dds.cr.usgs.gov/srtm/version2_1/. Para abranger Portugal Continental são necessárias 23 imagens.

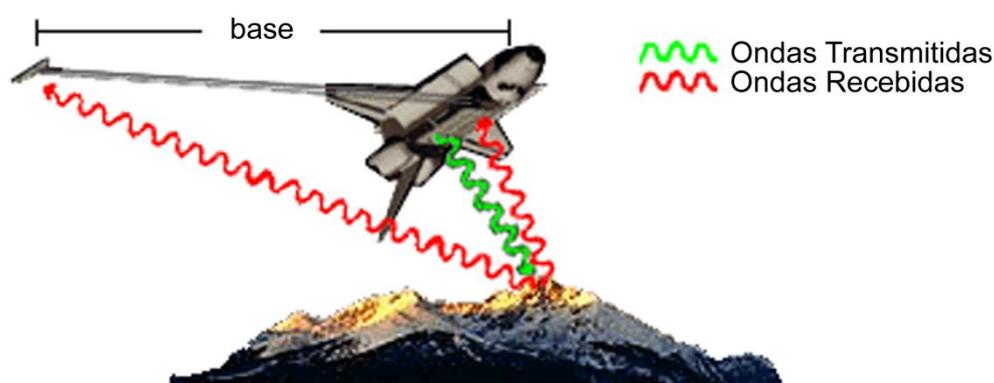


Figura 2.3. Processo de aquisição dos dados do SRTM (Adaptado de [75]).

2.3 Dados de Estações Meteorológicas

O *National Climatic Data Center* (NCDC), anteriormente conhecido como *National Weather Records Center* (NWRC), é o maior arquivo ativo de dados meteorológicos do mundo. O centro possui mais de 150 anos de dados e a cada dia recebe cerca de 224 GB de nova informação. Os dados são recebidos de várias fontes, como satélites, radares, sistemas de DR, barcos ou radiossondas [77].

O NCDC fornece dados de estações meteorológicas com várias resoluções temporais (horária, diária, mensal, etc). De acordo com a *Resolution 40* (Cg-XII) do programa *World Weather Watch* da *World Meteorological Organization* (WMO), o NCDC possibilita o acesso gratuito, com restrições, a dados diários (produto *Global Surface Summary of Day*) das estações meteorológicas de países membros da WMO [78]. O acesso a estes dados pode ser feito a partir do serviço *Climate Data Online* (CDO), que possibilita ao utilizador escolher geograficamente de que estações pretende receber dados. Este serviço encontra-se disponível na página <http://www.ncdc.noaa.gov/cdo-web/>. Os dados

são disponibilizados em formato *comma-separated values* (CSV), em que cada linha corresponde às medições diárias de uma estação.

No âmbito deste trabalho, utilizaram-se dados diários (média) da temperatura do ar e do ponto de orvalho. Para além dos dados de estações portuguesas, recorreu-se também a dados de algumas estações espanholas e marroquinas (Figura 2.4) para se realizar interpolação espacial para todo Portugal continental.

Este produto de dados meteorológicos encontra-se descrito detalhadamente em [78]. De notar, que algumas estações não possuem dados completos para o intervalo de tempo em estudo.

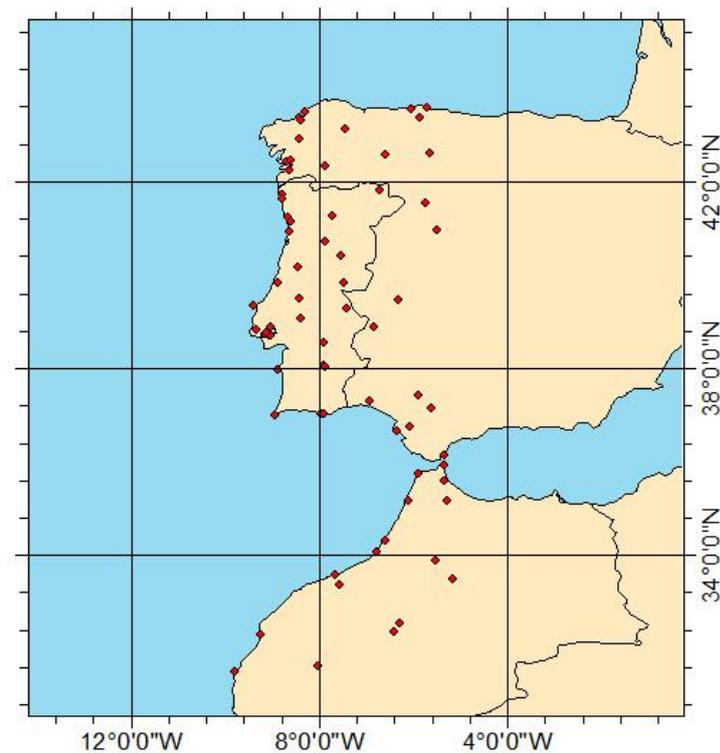


Figura 2.4. Distribuição das estações meteorológicas utilizadas (assinalado no mapa com pontos).

2.4 Dados da Concentração de Poluentes

A QualAr é uma base de dados (BD) sobre qualidade do ar em Portugal da responsabilidade da Agência Portuguesa do Ambiente (APA). O desenvolvimento da BD esteve a cargo do Departamento de Ciências e Engenharia do Ambiente da Faculdade de Ciências e Tecnologia da Universidade Nova Lisboa (DCEA – FCT/UNL), tendo como objetivos centralizar todos os dados de poluentes medidos em Portugal e disponibilizar ao público informação sobre o estado da qualidade do ar [79].

Os dados de vários poluentes são validados e fornecidos à QualAr pelas Comissões de Coordenação e Desenvolvimento Regional (Norte, Centro, Lisboa e vale do Tejo, Alentejo e Algarve), Direção Regional do Ambiente e Ordenamento do Território dos Açores e Direção Regional do Ambiente e Ordenamento do Território da Madeira [79].

Os poluentes de interesse para este trabalho são o PM_{10} e o NO_2 . Os dados da concentração destes poluentes foram descarregados gratuitamente a partir da página da QualAr [79] em formato XLS (folha de cálculo do Microsoft Excel). Cada ficheiro diz respeito a um poluente medido durante um ano, estando organizado da seguinte forma: a primeira coluna corresponde à data (ano, mês, dia, hora); as restantes colunas correspondem às medições do poluente, realizadas por cada estação. A distribuição de todas as estações utilizadas encontra-se representada na Figura 2.5. Nem todas as estações representadas possuem dados completos para o intervalo de tempo em estudo e são poucas as que recolhem dados de todos poluentes de interesse.

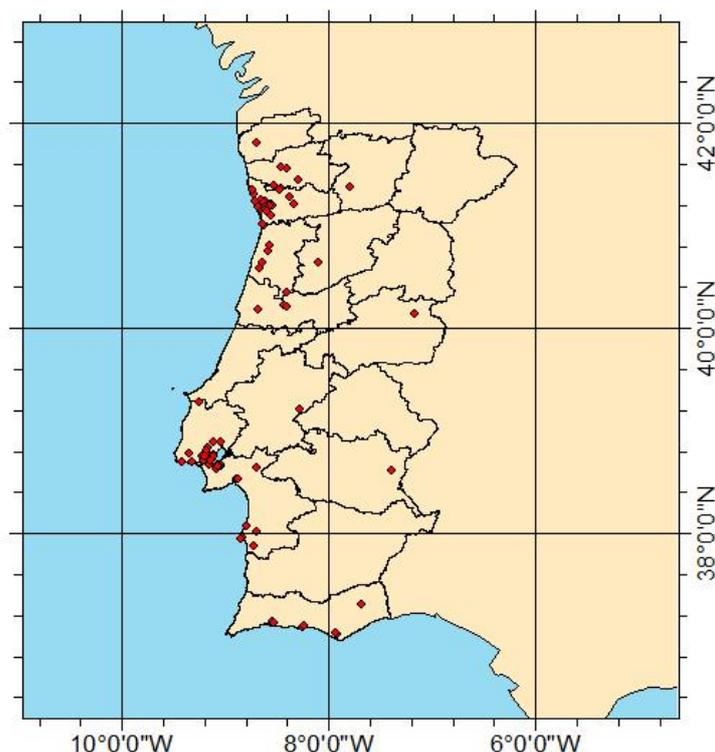


Figura 2.5. Distribuição das estações de qualidade do ar utilizadas (assinalado no mapa com pontos). O mapa de Portugal encontra-se dividido por distrito.

2.5 Carta Administrativa Oficial de Portugal 2011

A CAOP regista a delimitação e demarcação dos limites administrativos de Portugal (distrito, concelho e freguesia). A sua elaboração e atualização é da responsabilidade do Instituto Geográfico Português (IGP) desde 1999 [80].

Os limites administrativos existentes na CAOP provêm de diversas fontes de dados [80]:

- Base Geográfica de Referência de Informação (BGRI), do Instituto Nacional de Estatística (INE), elaborada no âmbito dos trabalhos realizados para os Censos 2001;
- Limites vetorizados a partir das secções de Cadastro Geométrico da Propriedade Rústica (CGPR);
- Procedimentos de Delimitação Administrativa (PDA) de freguesia/município;
- Diplomas Oficiais (p.e., referentes à extinção ou alteração de freguesias);
- Dados correspondentes a alterações propostas pelas autarquias, devidamente justificadas;
- Dados fornecidos por Institutos Oficiais na área da cartografia que contribuíram e contribuem com informação para a CAOP, como o INE, o Instituto Geográfico do Exército (IgeoE) ou o Instituto Hidrográfico (IH).

A CAOP é fornecida em formato *shapefile* (SHP) de polígonos no sistema de coordenadas nacional, o PTTM06. Cada área administrativa é identificada por um código de 6 dígitos denominado DICOFRE. Os dois primeiros dígitos correspondem ao distrito, os dois dígitos seguintes ao concelho e os dois últimos dígitos à freguesia. Neste trabalho, utilizou-se a atualização de 2011 da CAOP, descarregada gratuitamente na página <http://www.igeo.pt/produtos/cadastro/caop/versao2011.htm>.

2.6 Cartografia CORINE Land Cover 2006 para Portugal Continental

O programa *Co-ordination of Information on the Environment* (CORINE) foi criado em 1985 pela Comissão Europeia (CE) com o objetivo de desenvolver um sistema de informação sobre o estado do ambiente a nível europeu. O CORINE Land Cover (CLC) é um dos principais componentes deste programa e tem como objetivo principal a produção de cartografia de ocupação e uso do solo para os países da União Europeia [81]. Desde a criação do programa CORINE foram elaborados produtos CLC para três anos. A versão mais recente (CLC2006) é constituída por quatro produtos: cartografia CHA06_PT, que representa as alterações de uso e ocupação do solo entre 2000 e 2006; cartografia

CLC06_PT, que representa a caracterização do uso e ocupação do solo em 2006; CLC00_PT, que consiste numa versão revista do mapa CLC2000 original; CLC90_PT v2, que consiste numa versão revista do mapa CLC90 original [82].

Neste trabalho, utilizou-se o produto CLC06_PT fornecido em formato SHP de polígonos [82], descarregado gratuitamente na página do Grupo de Detecção Remota (GDR) do IGP (<http://www.igeo.pt/gdr/index.php>). O CLC06_PT tem uma escala de 1: 100 000, unidade mínima cartográfica (UMC) de 25 ha, exatidão geométrica menor que 100 m e possui uma nomenclatura hierárquica de 3 níveis (código com 3 dígitos) com 44 classes ao nível mais detalhado [82]. A nomenclatura dos dois primeiros níveis (menor detalhe) de classes do CLC06_PT encontra-se na Tabela 2.3.

Tabela 2.3. Nomenclatura dos dois primeiros níveis de classes do CLC06_PT.

Nível 1	Nível 2
1 Territórios artificializados	1.1 Tecido urbano
	1.2 Indústria, comércio e transportes
	1.3 Áreas de extração de inertes, áreas de deposição de resíduos e estaleiros de construção
	1.4 Espaços verdes urbanos, equipamentos desportivos, culturais e de lazer, e zonas históricas
2 Áreas agrícolas e agroflorestais	2.1 Culturas temporárias
	2.2 Culturas permanentes
	2.3 Pastagens permanentes
	2.4 Áreas agrícolas heterogéneas
3 Florestas e meios naturais e seminaturais	3.1 Florestas
	3.2 Florestas abertas, vegetação arbustiva e herbácea
	3.3 Zonas descobertas e com pouca vegetação
4 Zonas húmidas	4.1 Zonas húmidas interiores
	4.2 Zonas húmidas litorais
5 Corpos de água	5.1 Águas interiores
	5.2 Águas marinhas e costeiras

O CLC06_PT foi produzido pelo GDR em 2007, no âmbito da iniciativa *Global Monitoring for Environment and Security Land Fast Track Service Precursor* (GMES Land FTSP) para Portugal Continental. O GMES Land FTSP envolve 38 países e é da

responsabilidade da Agência Europeia do Ambiente (*European Environment Agency* – EEA) e da CE. O principal objetivo desta iniciativa passa por atualizar os mapas CLC produzidos anteriormente. Na produção da CLC06_PT foram utilizadas duas coberturas referentes aos períodos de primavera e verão de 2006, obtidas pelos satélites SPOT-4, SPOT-5 e *Indian Remote Sensing Satellite P6 (IRS-P6) Linear Imaging Self-Scanning Sensor-III (LISS-III)*, e uma referente ao verão de 2000, obtida pelo satélite Landsat-7. Para além das imagens de satélite, também foram utilizados outros dados, como fotografias aéreas ortorretificadas, mapas de áreas ardidas e a CAOP2008. Toda esta informação foi analisada visualmente por uma equipa de foto-intérpretes e posteriormente utilizada em ambiente SIG para gerar o produto final. À semelhança da CAOP2011, o CLC06_PT é fornecido no sistema de coordenadas nacional, o PTTM06.

2.7 Rede Viária

A rede viária de Portugal Continental foi descarregada gratuitamente em formato SHP de linhas a partir do OpenStreetMap (OSM). O OSM é um mapa global gratuito elaborado por pessoas de todo o mundo com dados provenientes de recetores GPS portáteis, fotografias aéreas e outras fontes livres. Apesar do mapa ser muito detalhado, a extração de diferentes tipos de estrada é um processo complicado, pois a codificação não é uniforme. Por esta razão, utilizou-se uma outra rede viária nacional em formato SHP de linhas, disponibilizada pelas Estradas de Portugal. S.A. (EP). Esta rede viária é pouco detalhada mas, possibilita a correta identificação de três tipos de estradas: itinerários principais (IP); itinerários complementares (IC); estradas nacionais (EN). De referir que, as redes viárias do OSM e das EP encontra-se, respetivamente, nos sistemas de coordenadas WGS84 e Hayford Gauss/*Datum* Lisboa (militar).

2.8 Dados de Internamentos Hospitalares e Populacionais

A informação com a frequência de internamentos hospitalares por freguesia devido à asma foi obtida a partir de uma BD de episódios hospitalares, disponibilizada pela Administração Central do Sistema de Saúde (ACSS). A cada episódio de internamento são associadas várias informações, como o diagnóstico principal e os secundários, o sexo e a freguesia de residência do paciente. Os diagnósticos encontram-se codificados através da Classificação Internacional de Doenças, 9ª Revisão, Modificação Clínica (CID-9-MC) (<http://icd9cm.chrisendres.com/>). As freguesias são identificadas pelo código DICOFRE.

Os episódios hospitalares referentes a admissões com diagnóstico principal de asma (código CID-9-MC 493.x) foram extraídos da BD para o intervalo de tempo em estudo. No total foram considerados 18 409 internamentos como diagnóstico principal de asma.

Os internamentos hospitalares devido a asma foram reduzidos à população residente em cada concelho. Para tal, utilizaram-se as estimativas provisórias da População Residente, Portugal, NUTS II, NUTS III e Municípios, relativas a 2010 (as estimativas definitivas não se encontravam disponíveis à data deste documento), fornecidas pelo INE.

As estimativas provisórias da população residente são calculadas a partir da população do recenseamento mais recente (neste caso, Censos 2001), através do método do seguimento demográfico [83]:

$$P_n = P_{n-1} + NV_n - O_n + I_n - E_n \quad (2.2)$$

em que:

P_n – População residente em 31/12 do ano n

P_{n-1} – População residente em 31/12 do ano n-1

NV_n – Número de nados vivos ocorridos em território nacional durante o ano n

O_n – Número de óbitos ocorridos em território nacional durante o ano n

I_n – Número de imigrantes (indivíduos que entraram em território nacional), estimado para o ano n

E_n – Número de emigrantes (indivíduos que saíram do território nacional), estimado para o ano n

Estes dados populacionais foram obtidos em formato XLS na página do INE (<http://www.ine.pt>).

Capítulo 3

Metodologia

3.1 Determinação da Temperatura e Humidade Relativa do Ar a partir do Perfil Atmosférico MODIS

A T_a é um parâmetro importante em várias aplicações, incluindo estudos de doenças infecciosas e estudos de alterações climáticas [84]. Existem várias formas de determinar a temperatura do ar à superfície a partir de dados de DR, como regressões lineares entre T_a e o LST, ou a combinação entre a temperatura e o NDVI [85].

Neste trabalho, optou-se por utilizar o produto MOD07 para determinar a temperatura por três motivos. Em primeiro lugar, algumas variáveis não podem ser determinadas de forma imediata. Assim, era importante verificar se um único produto *standard* com calibração global podia ser utilizado de forma simples e fiável para determinar esta variável numa área de estudo local. Em segundo lugar, como a análise das ocorrências de asma foi realizada com base no concelho dos pacientes, 5 km de resolução espacial é suficiente, pois são poucos os concelhos com menos de 25 km² de área. Por último, o MOD07 possui parâmetros (perfil de humidade ou TCWV) que permitem determinar outra variável de interesse, a humidade relativa.

De acordo com o U.S. *Standard Atmosphere 1976*, para uma T_a igual a 15 °C e pressão do ar ao nível do mar igual 1013.25 hPa (equivalente a mbar), o *lapse rate* de T_a na troposfera (entre 0 km e 11 km de altitude), ou seja, a diminuição da temperatura com o aumento da altitude na troposfera, é igual a -6.5 °C km⁻¹ [86]. Este *lapse rate* é bastante utilizado [86-88] mas, como se trata de um *standard* global, pode não modelar

corretamente algumas zonas (p.e., varia com a latitude [87]). Assim, a sua utilização em escalas de maior pormenor requer algum cuidado.

Normalmente, as estações meteorológicas, em particular em zonas montanhosas, encontram-se bastante separadas, por isso a utilização de um *lapse rate* preciso na construção de mapas climáticos a partir das suas medições torna-se importante. Contudo, este problema não se põe num PT_a determinado por satélite, pois o *lapse rate* pode ser determinado em cada pixel através de uma regressão linear simples. Conhecendo o *lapse rate* e a temperatura do ar ao nível do mar (dados pela equação da regressão linear) num dado ponto, para determinar a T_a basta conhecer a sua altitude. A altitude pode ser obtida a partir do MDE do SRTM mas, primeiro é necessário converter os níveis de pressão do PT_a MODIS para níveis de altitude.

A altitude pode ser estimada a partir da pressão (P_A) pela seguinte relação *standard* [89]:

$$A = -H \times \ln \frac{P_A}{P_0} \quad (3.1)$$

onde P_0 é a pressão *standard* ao nível do mar e H a escala da altitude. A escala da altitude H é dada por:

$$H = \frac{RT_a}{M_a g} \quad (3.2)$$

em que:

R – constante de gás ideal ($8.3144621 \text{ m}^3 \text{ Pa K}^{-1} \text{ mol}^{-1}$)

M_a – peso molecular médio do ar (28.97 g mol^{-1})

g – aceleração da gravidade (9.807 m s^{-2})

Na parte inferior da atmosfera, o H costuma ser considerado igual a 8 km para um $T_a = 273 \text{ K}$ ($\approx 0 \text{ }^\circ\text{C}$) [89].

À semelhança do *lapse rate*, optou-se por não utilizar um valor *standard*, uma vez que pode não refletir a realidade local. Por exemplo, existem várias estações que se encontram a uma altitude superior a 10 m, cuja pressão medida ao longo do intervalo de tempo é, em geral, superior a P_0 .

Como na parte inferior da atmosfera a variação da pressão com a altitude é, aproximadamente, linear [90, 91], recorreu-se às medidas de pressão efetuadas pelas estações e respetiva altitude para determinar uma expressão local através de uma regressão linear (Figura 3.1):

$$A = -8.8188 \times P + 8977.0354 \quad (3.3)$$

onde A é a altitude e P a pressão do ar.

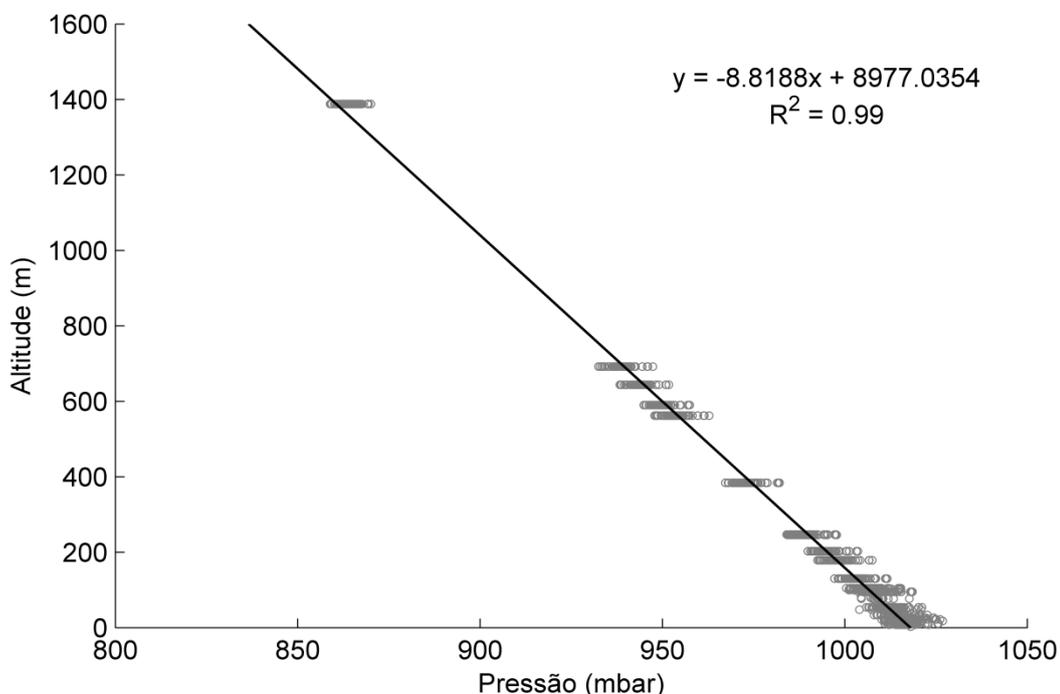


Figura 3.1. Relação entre a pressão média mensal do ar medida pelas estações e a respetiva altitude.

Após converter os níveis de pressão em níveis de altitude, a cada pixel do PT_a é ajustada uma reta para determinar os respetivos lapse rate (lr) e temperatura do ar ao nível do mar (T_0). A T_a em qualquer pixel do PT_a é determinada pela expressão:

$$T_a = lr \times A + T_0 \quad (3.4)$$

onde, lr e A são, respetivamente, o *lapse rate* e a altitude. T_0 é a temperatura do ar ao nível do mar, considerando lr .

Conforme realizado noutros estudos [90, 92], a grelha da pressão do ar ao nível da superfície contida no produto MOD07 poderia ser utilizada na regressão, em vez do MDE. Esta grelha resulta da interpolação bilinear dos pixels vizinhos do modelo de pressão dado pelo *National Environmental Prediction (NCEP) Global Data Assimilation System (GDAS)* [61]. Contudo, como se trata de um produto com calibração global, algumas zonas não são modeladas corretamente, conforme pode ser verificado pelo coeficiente de

determinação (R^2) obtido ao comparar a pressão medida pelas estações meteorológicas com a pressão dada pelo modelo (Figura 3.2). De referir, que os pontos que apresentam maior erro (conjunto de pontos assinalados com um círculo vermelho) correspondem à estação com maior altitude (Penhas Douradas, 1388 m).

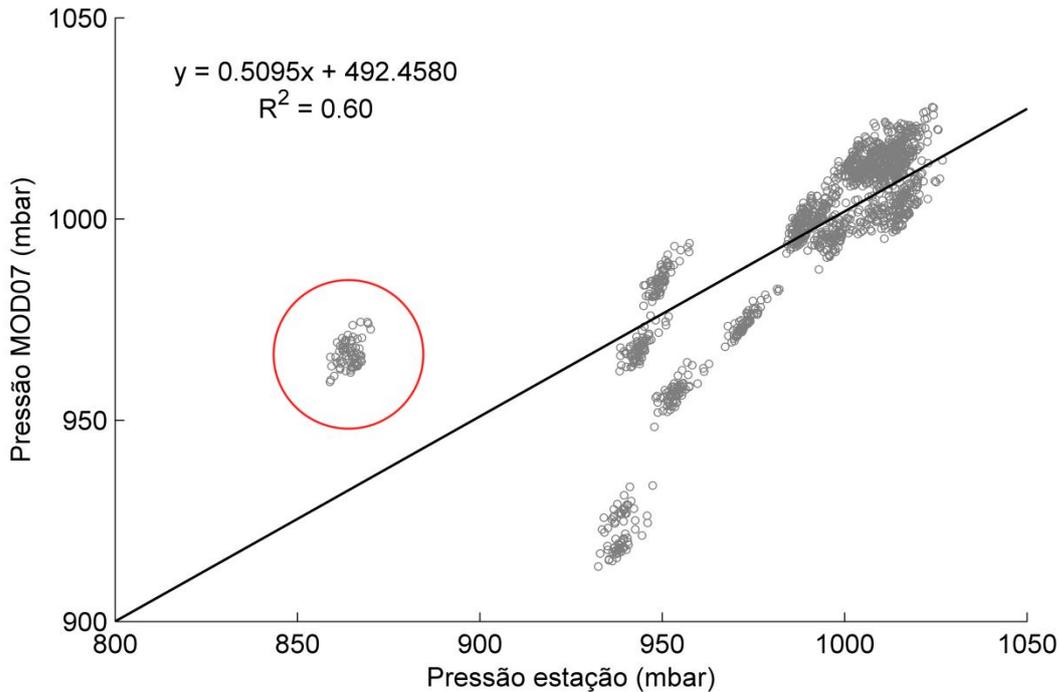


Figura 3.2. Relação entre a pressão média mensal do ar medida pelas estações meteorológicas e dada pelo produto MOD07. Os valores mensais de pressão do MOD07 foram calculados com todos os dados diários (dia e noite) dos satélites Aqua e Terra, para o intervalo de tempo 2003 a 2008.

Após a determinação de T_a diária, calcularam-se médias mensais de T_a . Inicialmente, pretendia-se recorrer a todas as observações dos satélites Aqua e Terra para o cálculo. Contudo, como se poderá verificar na validação dos resultados (secção 4.1.1), existe um erro sistemático nos dados da temperatura durante, principalmente, os meses de verão, fazendo com que se verifique um aumento no erro de T_a . Assim, decidiu-se determinar as médias mensais de T_a com quatro combinações de dados, que vão excluindo as observações

- Todas as observações dos satélites Aqua e Terra (T_aAT);
- Todas as observações do satélite Aqua e as observações do satélite Terra durante a noite (T_aATn);

- Todas as observações do satélite Terra e as observações do satélite Aqua durante a noite (T_aAnT);
- Observações durante a noite dos satélites Aqua e Terra (T_aAnTn).

Nestas combinações, as observações realizadas durante o dia foram sendo removidas, pois são sujeitas a erros relativos à radiação solar, que afeta as bandas no infravermelho térmico. O Erro Médio Quadrático (EMQ) foi calculado para cada mês e para cada combinação de dados. Em cada mês, selecionou-se a combinação de dados com menor EMQ para estimar T_a .

As médias mensais de TCWV foram calculadas recorrendo a todas as observações dos satélites Aqua e Terra, pois não se verificou nenhuma melhoria significativa dos resultados, utilizando outra combinação de dados.

Depois de determinadas as médias mensais de T_a e TCWV, calcularam-se médias mensais da humidade relativa (HR). A HR mede a quantidade de vapor de água que existe no ar em relação ao máximo que o ar poderia conter à mesma T_a . Considerando a quantidade de vapor de água constante, a HR aumenta com a diminuição da temperatura.

Os satélites não medem a HR diretamente, uma vez que esta variável depende do vapor de água e da T_a . Uma medição de humidade frequentemente utilizada é a temperatura do ponto de orvalho ou temperatura *dew point* (T_d). A T_d consiste na temperatura a que o ar tem que ser arrefecido de forma a alcançar a saturação (pressão do ar constante) [89]. Quando a saturação é alcançada, ou seja, T_a igual a T_d , HR é igual a 100%.

Segundo Reitan [93], a T_d pode ser estimada a partir de uma regressão linear com o logaritmo da coluna total de vapor de água. Deste modo, T_d foi determinada aplicando uma regressão linear múltipla, com T_d medida pelas estações meteorológicas como variável dependente. Como variáveis independentes da regressão considerou-se o logaritmo do TCWV do MOD07 e a altitude dada pelo MDE do SRTM. A altitude foi incluída na regressão, pois permitiu aumentar o R^2 do modelo de 0.793 para 0.859 e não havia multicolinearidade com o logaritmo de TCWV (correlação baixa com $R = 0.300$). Sendo assim, T_d no ponto S foi estimada pela expressão:

$$T_d = 6.971 + 8.715 \times \ln \text{TCWV} - 0.004 \times A_S \quad (3.5)$$

A T_d também poderia ser calculada por interpolação do perfil de humidade do MOD07 mas, verificou-se que o perfil possuía baixa precisão (EMQ entre 2 e 3 °C).

Após a determinação da T_d , foi determinada a HR:

$$HR = \frac{e_a}{e_s} \times 100\% \quad (3.6)$$

sendo a pressão do vapor e_a e a pressão do vapor saturado e_s dados por [94]:

$$e_a = 6.11 \times 10^{\frac{7.5 \times T_d}{237.7 + T_d}} \quad (3.7)$$

$$e_s = 6.11 \times 10^{\frac{7.5 \times T_a}{237.7 + T_a}} \quad (3.8)$$

3.2 Interpolação Espacial: Kriging

O método de Kriging foi utilizado para determinar estimativas espaço-temporais da HR e dos poluentes NO_2 e PM_{10} . A utilização deste método surgiu do facto de existirem limitações nos dados de satélite, em particular dos poluentes (p.e., a baixa resolução destes dados).

A origem do método de Kriging está associada ao trabalho desenvolvido pelo engenheiro de minas, Daniel G. Krige, que utilizava médias móveis para evitar a sobrestimação sistemática de reservas de mineração [95, 96]. Na altura, os métodos utilizados para descrever a variabilidade espacial de uma variável baseavam-se sobretudo em estatística clássica (isto é, utilização de parâmetros como média e desvio padrão para representar um determinado fenómeno), utilizando grandes conjuntos de dados. Estes métodos partem do pressuposto que as variações de um ponto para outro são aleatórias [95].

Durante o seu trabalho com dados de concentração de ouro, Krige concluiu que a informação dada pela variância não era suficiente para explicar o fenómeno em estudo e, como tal, seria necessário considerar a distância entre as observações. Mais tarde, Georges Matheron melhorou este conceito e deu uma orientação matemática a um método, surgindo assim os conceitos de variáveis regionalizadas e variograma. Matheron nomeou o método desenvolvido como Kriging [96].

A teoria de variáveis regionalizadas (VR) constitui o princípio básico do método de Kriging. Uma VR é uma função numérica com distribuição espacial, que varia de um ponto para outro com continuidade aparente. Esta teoria pressupõe que a variação de uma variável pode ser expressa pela soma de três componentes: uma componente estrutural, associada a um valor médio constante ou a uma tendência constante; uma componente aleatória espacialmente correlacionada; um ruído aleatório ou erro residual [95].

Se x representa uma posição em uma, duas ou três dimensões, então o valor da variável Z , em x , é dada por [95]:

$$Z(x) = m(x) + \varepsilon'(x) + \varepsilon'' \quad (3.9)$$

em que:

$m(x)$ – função determinística que descreve a componente estrutural de Z em x ;

$\varepsilon'(x)$ – termo estocástico, que varia localmente e depende espacialmente de $m(x)$;

ε'' – ruído aleatório não correlacionado, com distribuição normal com média igual a zero e variância σ^2 .

Tendo em conta a teoria das VR, a hipótese mais comum considerada pelo método de Kriging denomina-se “estacionariedade de 2ª ordem” [95, 96]. Esta hipótese assume que as VR possuem as seguintes propriedades [95]:

- $m(x)$ é constante, ou seja, não existem tendências na região;
- A variância (Var) das diferenças entre duas amostras depende somente da distância h entre elas:

$$\text{Var}[Z(x) - Z(x+h)] = E\{[Z(x) - Z(x+h)]^2\} = 2\gamma(h) \quad (3.10)$$

onde $\gamma(h)$ é a semi-variância e E o valor esperado.

Desta forma, a variação local das amostras pode ser caracterizada pelo variograma. Assim, a expressão (3.9) pode ser reescrita como:

$$Z(x) = m(x) + \gamma(h) + \varepsilon'' \quad (3.11)$$

A diferença entre o Kriging e outros métodos de interpolação é a forma como os pesos são atribuídos às diferentes amostras. A interpolação pelo método de Kriging recorre à determinação de um variograma empírico (semi-variograma), a partir de um conjunto de pontos, para calcular os pesos λ_i que minimizem a variância nos valores estimados [97].

O variograma representa o grau de semelhança entre duas observações (variáveis regionalizadas) separadas por uma determinada distância (*lag*). O semi-variograma pode ser calculado a partir de dados pontuais pela seguinte expressão [97]:

$$\gamma(h) = \frac{1}{2M(h)} \sum_{i=1}^{M(h)} \{z(x_i) - z(x_i+h)\}^2 \quad (3.12)$$

em que:

$\gamma(h)$ – semi-variância estimada para a distância de separação ou *lag* h ;

$M(h)$ – número de pares de valores medidos ($z(x_i)$ e $z(x_i+h)$), separados pela distância h ;

$z(x_i)$ e $z(x_i+h)$ – valores observados em x_i e x_i+h , respetivamente.

Na Figura 3.3 encontra-se representado um semi-variograma com características muito próximas do ideal, ou seja, as diferenças $\{z(x_i) - z(x_i+h)\}$ decrescem à medida que h diminui.

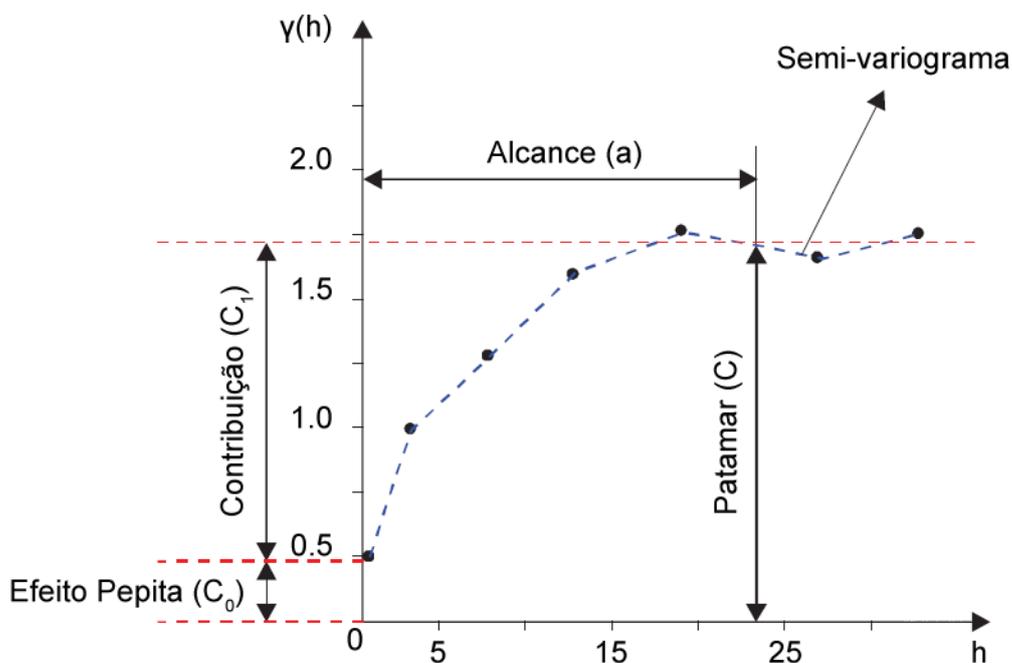


Figura 3.3. Exemplo de um semi-variograma (Adaptada de [95]).

O semi-variograma é descrito pelos seguintes parâmetros [95]:

- Alcance (a) – distância dentro da qual as amostras apresentam-se correlacionadas espacialmente. Na Figura 3.3, o alcance ocorre próximo de 25 m;
- Patamar (C) – valor no qual o semi-variograma estabiliza, ou seja, é o valor a partir do qual as amostras tornam-se independentes, devido à grande distância que as separa;

- Efeito Pepita (C_0) – valor da função semi-variograma na origem ($h=0$). Idealmente, $\gamma(0)=0$, pois duas amostras tomadas no mesmo ponto deveriam ter os mesmos valores. Contudo, na prática, à medida que h tende para zero, $\gamma(h)$ aproxima-se de um valor positivo chamado Efeito de Pepita, que revela a descontinuidade do semi-variograma para distâncias menores do que a menor distância entre amostras. Normalmente, esta descontinuidade deve-se a erros de medição ou da variabilidade natural da variável;
- Contribuição (C_1) – diferença entre C e C_0 .

Como o método de Kriging utiliza o variograma para calcular os pesos de cada observação, é necessário ajustar uma função teórica ao semi-variograma, possibilitando que as semi-variâncias possam ser calculadas para todos os *lags* [97]. Existem vários modelos (M) teóricos que podem ser utilizados no ajuste, sendo os mais comuns o exponencial, o esférico e o gaussiano (Figura 3.4).

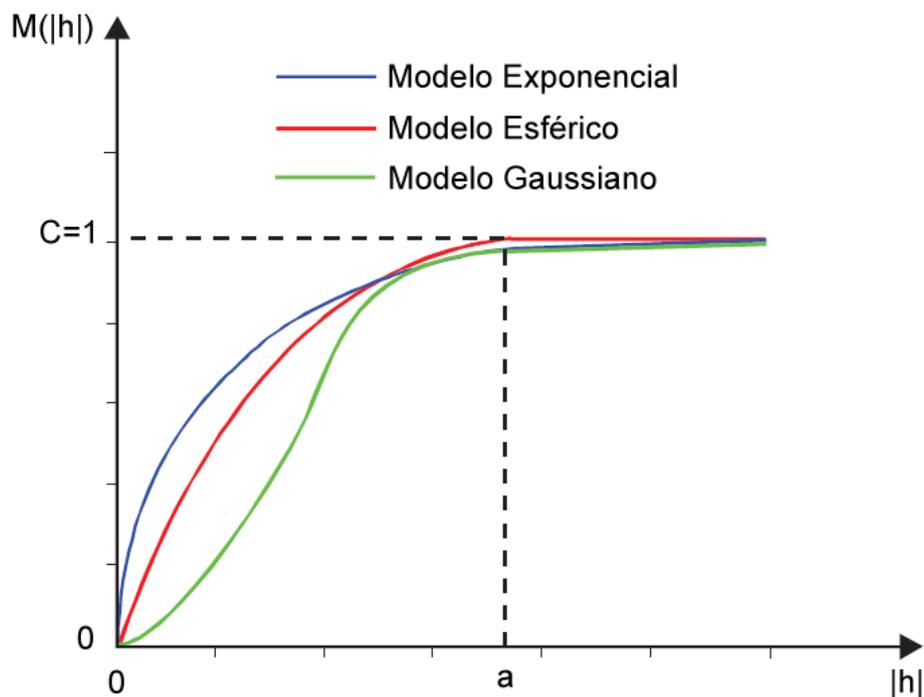


Figura 3.4. Exemplo gráfico de modelos de variograma (Adaptado de [95]).

O modelo esférico (Esf) foi escolhido para ajustar ao semi-variograma, pois trata-se de um dos mais utilizados [95, 97]. A equação normalizada deste modelo é [95]:

$$\text{Esf}(|h|) = \begin{cases} 0 & , \quad |h|=0 \\ \frac{3}{2}\left(\frac{|h|}{a}\right) - \frac{1}{2}\left(\frac{|h|}{a}\right)^3 & , \quad 0 < |h| \leq a \\ 1 & , \quad |h| > a \end{cases} \quad (3.13)$$

O modelo de variograma, ajustado aos dados pontuais, é utilizado para calcular os pesos, garantindo que a variância da estimação seja menor que a variância de qualquer combinação linear dos valores observados [97]. O semi-variograma do modelo esférico é definido por:

$$\gamma(|h|) = \begin{cases} 0 & , \quad |h|=0 \\ C_0 + C_1 \left[\frac{3}{2}\left(\frac{|h|}{a}\right) - \frac{1}{2}\left(\frac{|h|}{a}\right)^3 \right] = C_0 + C_1[\text{Esf}(|h|)] & , \quad 0 < |h| \leq a \\ C_0 + C_1 & , \quad |h| > a \end{cases} \quad (3.14)$$

Após o cálculo dos pesos, o método de Kriging é aplicado. Existem algumas variantes deste método como, Kriging Simples, Kriging Ordinário, Kriging Universal, Co-Kriging, etc. O Kriging Ordinário foi adotado para todas as interpolações por ser uma das variantes mais utilizadas e por se tratar de um método conservativo, que permite evitar erros inerentes ao desconhecimento da tendência dos dados [97].

Seja Z a temperatura do solo, medida em n pontos, com coordenadas bidimensionais representadas pelo vetor x . O valor de Z num ponto de coordenadas x_0 pode ser estimado a partir da combinação linear dos n valores observados somados a um parâmetro λ_0 [95]:

$$Z_{x_0}^* = \lambda_0 + \sum_{i=1}^n \lambda_i Z(x_i) \quad (3.15)$$

O Kriging Ordinário é um método de estimação não tendencioso [95, 97], por isso:

$$E[Z_{x_0} - Z_{x_0}^*] = 0 \quad (3.16)$$

Substituindo (3.15) em (3.16), tem-se:

$$E[Z_{x_0}] = E\left[\lambda_0 + \sum_{i=1}^n \lambda_i Z(x_i)\right] \Leftrightarrow m = \lambda_0 + \sum_{i=1}^n \lambda_i m \quad (3.17)$$

Ao contrário do Kriging Simples, onde é necessário conhecer previamente a média m [97], no Kriging Ordinário tal não é necessário. Assim, para que a igualdade (3.17) seja satisfeita é necessário que $\lambda_0=0$ e $\sum_{i=1}^n \lambda_i=1$.

Desta forma, o estimador Kriging Ordinário é dado por [95]:

$$Z_{x_0}^* = \sum_{i=1}^n \lambda_i Z(x_i), \quad \text{com } \sum_{i=1}^n \lambda_i = 1 \quad (3.18)$$

Se a variância do erro for minimizada pela condição $\sum_{i=1}^n \lambda_i=1$, os pesos são obtidos a partir do sistema [95]:

$$\gamma(h) = \begin{cases} \sum_{j=1}^n \lambda_j C(x_i, x_j) - \alpha = C(x_i, x_0) & \text{para } i=1, \dots, n \\ \sum_{j=1}^n \lambda_j = 1 \end{cases} \quad (3.19)$$

em que:

$C(x_i, x_j)$ – semi-variância entre os pontos x_i e x_j ;

$C(x_i, x_0)$ – semi-variância entre os pontos x_i e x_0 ;

α – multiplicador de Lagrange necessário para a minimização da variância do erro.

A variância minimizada do erro $\sigma_{k_0}^2$ é dada por [95]:

$$\sigma_{k_0}^2 = \text{Var}[Z_{x_0} - Z_{x_0}^*] = C(0) - \sum_{i=1}^n \lambda_i C(x_i, x_0) - \alpha \quad (3.20)$$

3.3 Modelos *Land-Use Regression* (LUR)

Os estudos epidemiológicos sobre os efeitos que a exposição a poluentes provoca na saúde necessitam que as concentrações de poluentes em locais onde não existe monitorização sejam determinadas com precisão. Existem vários métodos utilizados para esse fim, como a interpolação espacial (p.e., Kriging), modelos de proximidade (p.e., proximidade a estradas com elevado tráfego diário) ou modelos de dispersão [98] (p.e., os modelos CHIMERE, EURAD e CALIOPE, utilizados em Portugal [99]). Contudo, os

poluentes em zonas urbanas, nomeadamente os de origem automóvel, apresentam uma grande variabilidade em curtos espaços, que não é capturada pela maior parte dos métodos de interpolação [98]. A utilização dos modelos de proximidade tem também algumas limitações, como a variabilidade temporal dos poluentes não ser considerada e a ocorrência de erros de classificação de exposição (p.e., residência < 2 km = exposto, residência > 2 km = não exposto) [100]. Por um lado, os modelos de dispersão são muito versáteis, não necessitando de dados adicionais de estações de monitorização e tendo a capacidade de adaptação a diferentes áreas e poluentes. Por outro lado, possuem alguns problemas, nomeadamente, o facto de requererem uma vasta quantidade e densidade de dados (p.e., volume de tráfego, velocidade de tráfego, comprimento das estradas, altura dos edifícios, velocidade e direção do vento, estabilidade atmosférica, etc.) [100, 101].

De forma a resolver estes problemas foram desenvolvidos os modelos *Land Use Regression* (LUR) para modelar a distribuição de poluentes de origem automóvel, como o NO₂ ou as PM_{2.5} [98]. Neste trabalho, a necessidade de recorrer aos modelos LUR surgiu da incapacidade do método de Kriging em modelar espacialmente os poluentes NO₂ e PM₁₀ com precisão e para a totalidade de Portugal Continental.

Os modelos LUR utilizam os níveis dos poluentes de interesse como variável dependente e parâmetros como, volume de tráfego diário, altitude, classe de ocupação de solo, proximidade a estradas ou outros tipos de parâmetros espaciais, como variáveis independentes, num modelo de regressão linear múltipla:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (3.21)$$

onde, os $k + 1$ parâmetros β_j ($j = 0, 1, \dots, k$), são os coeficientes da regressão e ε o erro aleatório. Apesar dos modelos LUR serem utilizados sobretudo em escalas espaciais elevadas, o objetivo da utilização destes modelos consistiu em mapear a concentração mensal dos poluentes para toda a área de Portugal Continental. Isto não era possível com o método de Kriging, pois conforme se pode verificar na Figura 2.5 a distribuição das estações de qualidade do ar não permite criar estimativas para todo o país. Sendo assim, foram desenvolvidos dois modelos LUR mensais para os poluentes PM₁₀ e NO₂.

Na Tabela 3.1 são apresentadas todas as variáveis independentes consideradas para construir estes modelos. Estas variáveis foram escolhidas com base em estudos anteriores [98, 101-103] e nos dados disponíveis. Para além das variáveis normalmente utilizadas nos modelos LUR (i.e., ocupação de solo, tipo de estrada, topografia, etc.), os anos e os meses foram introduzidos como variáveis independentes, de forma a ter em conta a variação temporal da concentração dos poluentes, assim como fatores

desconhecidos. Estas variáveis possuem valores binários de 0 ou 1. Outros parâmetros variáveis ao longo do tempo, como T_a , HR ou AOT550, também foram incluídos.

Os dados dos poluentes e o AOT550 não possuíam uma distribuição simétrica, por isso foram transformados a função raiz quadrada. Na Figura 3.5 encontra-se um exemplo de dados assimétricos e respetiva transformação.

As classes de ocupação de solo foram definidas através da agregação das 44 classes do CLC06_PT. Deste processo resultaram 7 classes (Tabela 3.2). As classes urbanas foram diferenciadas pelo facto da poluição nestas zonas possuir uma grande variabilidade [98].

Depois de preparar as variáveis independentes e as associar a cada medida pontual dos poluentes construíram-se os modelos LUR. Os modelos foram construídos através das melhores combinações das variáveis da Tabela 3.1, garantindo que nos modelos finais não havia multicolinearidade entre as variáveis independentes.

Tabela 3.1. Variáveis independentes consideradas na criação dos modelos LUR.

Variável	Raio do Buffer (m)	Unidade	Origem
Latitude e Longitude	N/A	°	N/A
Altitude	N/A	m	SRTM MDE
Declive	N/A	°	SRTM MDE
Densidade de IP, IC e EN no buffer	250, 500, 750	km/km ²	rede viária EP
Densidade de Estradas no buffer	250, 500, 750	km/km ²	rede viária OSM
Distância a IP, IC e EN mais próximo	N/A	m	rede viária EP
Distância à estrada mais próxima	N/A	m	rede viária OSM
Distância à linha de costa	N/A	m	CAOP2011
Ocupação de solo Urb1, Urb2, Urb3, Urb4, Agua, Agro e Veg no buffer	250, 500, 750	%	CLC06_PT
AOT550	N/A	N/A	MOD04
T_a	N/A	°C	MOD07
HR	N/A	%	Kriging
NDVI	N/A	N/A	MOD13A3
Variáveis anuais e mensais	N/A	N/A	N/A

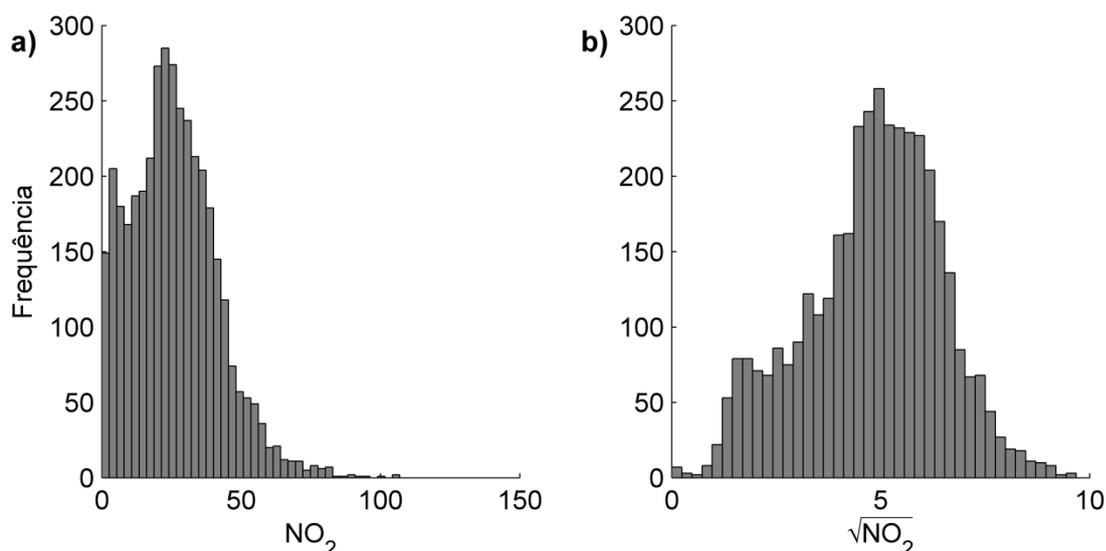


Figura 3.5. Exemplo da distribuição dos dados do poluente NO_2 : a) sem transformação; b) com transformação raiz quadrada.

Tabela 3.2. Classes utilizadas na construção dos modelos LUR.

Classe Nova	Classe Original
Urb1	Tecido urbano (nível 2)
Urb2	Indústria, comércio e transportes (nível 2)
Urb3	Áreas de extração de inertes, áreas de deposição de resíduos e estaleiros de construção (nível 2)
Urb4	Espaços verdes urbanos, equipamentos desportivos, culturais e de lazer, e zonas históricas (nível 2)
Agua	Zonas húmidas (nível 1) e Corpos de água (nível 1)
Agro	Áreas agrícolas e agroflorestais (nível 1)
Veg	Florestas e meios naturais e seminaturais (nível 1)

3.4 Análise Estatística

A associação entre os internamentos devido a asma e as variáveis ambientais foi efetuada a partir de regressões lineares simples e múltiplas (equação (3.21)). Como variável dependente da regressão considerou-se a taxa de internamentos por 1000 habitantes em cada concelho durante todo o intervalo de tempo em estudo (2003-2008). Esta informação foi determinada através da redução dos internamentos à população residente em cada concelho (dados populacionais do INE). Como variáveis

independentes considerou-se a média das cinco variáveis ambientais determinadas (T_a , HR, NDVI, PM_{10} e NO_2).

A análise foi realizada por concelho para todo o intervalo de tempo em estudo com o intuito de garantir um equilíbrio entre o nível de detalhe da exposição às variáveis ambientais e o número de internamentos. Apesar de se possuir informação ao nível da freguesia, o número reduzido de internamentos hospitalares (cada freguesia tem em média $18409/4050 = 4$ internamentos) prejudica as associações, uma vez que tende para valores residuais. A resolução temporal do estudo permite também evitar possíveis fatores de confusão resultantes da sazonalidade dos dados.

Tendo em conta que a taxa de internamentos por 1000 habitantes por asma não se correlaciona com a população total (Figura 3.6), as relações com as variáveis ambientais, em particular com os poluentes, vão variar em função da região (em geral, os distritos mais populosos são dos mais poluídos). Assim, cada distrito foi analisado separadamente.

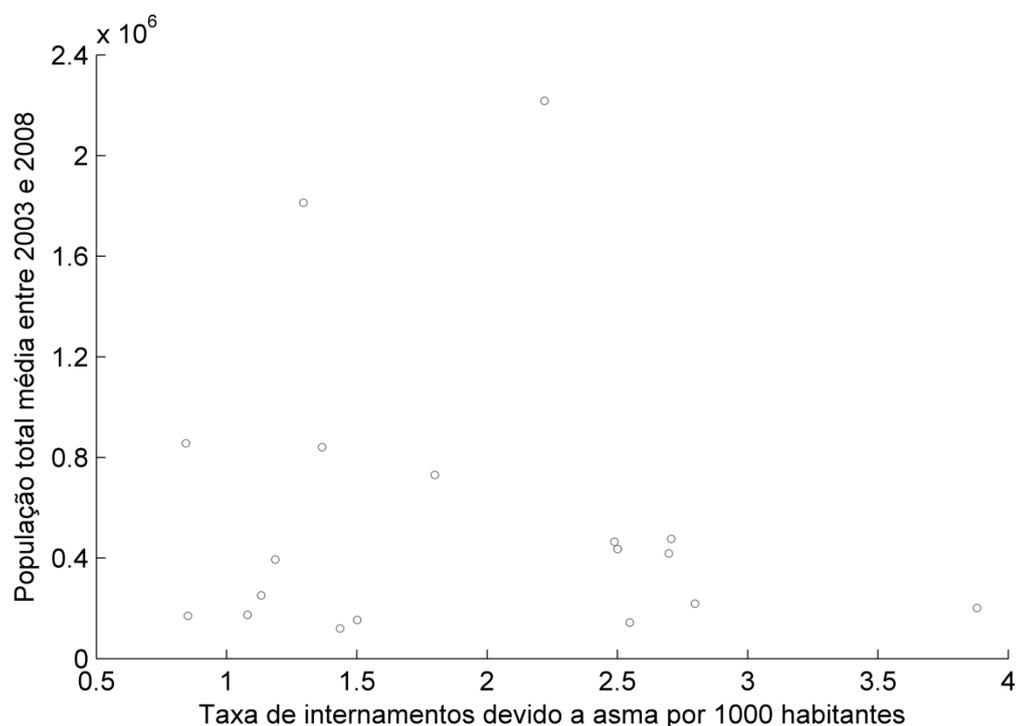


Figura 3.6. Taxa de internamentos devido a asma por 1000 habitantes em função da população total média, referente ao intervalo de tempo de 2003 a 2008.

3.5 Implementação dos Métodos

3.5.1 Dados das Estações Meteorológicas e de Qualidade do ar

Os dados pontuais de T_a , HR, PM_{10} e NO_2 foram utilizados tanto para validação como para determinação de estimativas espaciais mensais. Para tal, criaram-se ficheiros CSV com médias mensais desses parâmetros.

Antes de descrever o processo de criação dos ficheiros CSV é importante referir os critérios utilizados no cálculo das médias mensais. As médias mensais de T_a e RH são calculadas com base em medições diárias, considerando o seguinte critério:

- O cálculo da média mensal de um determinado parâmetro só é realizado para estações que possuam pelo menos 75% de medições diárias desse parâmetro nesse mês (24, 23, 22 ou 21 dias de medições, dependendo do mês).

No caso dos poluentes, as medições horárias são inicialmente agregadas para medições diárias. Uma medição diária corresponde à média das medições horárias de um dia completo (24 horas). As medições diárias são, posteriormente, utilizadas para calcular as médias mensais. Os critérios considerados nestes cálculos são os seguintes:

- O cálculo da medição diária dos poluentes só é realizado para estações que possuam pelo menos 75% de medições horárias dos poluentes nesse dia (18 horas de medições);
- O cálculo da média mensal de um determinado poluente só é realizado para estações que possuam pelo menos 75% de medições diárias desse poluente nesse mês (24, 23, 22 ou 21 dias de medições, dependendo do mês).

Para facilitar o cálculo das médias mensais foram construídas duas BDs no Microsoft Access (Tabela 3.3).

Tabela 3.3. Bases de dados criadas com os parâmetros meteorológicos e poluentes.

Nome da BD	Nº de Tabelas	Descrição
NCDC_2003_2008	2	Dados meteorológicos diários e informação das respetivas estações
QualAr_2003_2008	13 ^a	Dados dos poluentes PM_{10} e NO_2 (uma tabela por cada ano de cada poluente) e informação das respetivas estações

^a Para cada poluente manteve-se uma tabela por ano para tornar o processo mais eficiente.

Um conjunto de programas (Tabela 3.4) foi implementado em *Visual Basic for Applications* (VBA) para auxiliar a elaboração das BDs da Tabela 3.3 e criar os ficheiros CSV com as médias mensais de cada parâmetro.

Tabela 3.4. Programas utilizados para calcular médias mensais dos dados pontuais.

Nome do Programa	Objetivo
FormatarDadosQualAr.bas	Formatar os dados dos poluentes para que cada linha corresponda a uma medição horária de uma determinada estação
ImportarFolhasExcel.bas	Importar as tabelas em Excel com os dados dos poluentes para uma BD em Access com a estrutura definida na Tabela 3.3
AccessMedias.bas	Criar ficheiros CSV com as médias mensais de todos os parâmetros

Antes de utilizar os programas apresentados na Tabela 3.4, todos os ficheiros dos poluentes foram convertidos manualmente para o formato XLSM (ficheiro XLS que suporta macros).

FormatarDadosQualAr.bas (Anexo A):

No caso dos poluentes, como já foi referido, cada ficheiro diz respeito a um poluente, onde cada linha do ficheiro corresponde às medições de várias estações num dado instante. Como esta estrutura dificulta o cálculo das médias mensais, este programa é utilizado para formatar os dados dos poluentes para que cada linha corresponda a uma medição horária de uma determinada estação.

No início do programa definem-se quais os ficheiros que se pretende formatar (definindo os anos e os poluentes). De seguida, para cada ficheiro definido, o programa começa por abri-lo e cria uma nova folha de cálculo (“Dados_Formatados”), que irá receber os dados com a formatação final. Após a criação da nova folha, é preenchido o cabeçalho dos dados finais (Data, Ano, Mês, Dia, Hora, Estação, Parâmetros).

Por fim, os dados de cada estação, juntamente com o nome e a data, são copiados sequencialmente para a folha “Dados_Formatados”, de forma a que cada linha corresponda a uma medição horária de uma determinada estação.

ImportarFolhasExcel.bas (Anexo A):

Os ficheiros que se pretende importar são definidos no início (definindo os anos e os poluentes). Para cada ficheiro é importada apenas a folha de cálculo denominada

“Dados_Formatados”. De referir que, antes de executar este programa é necessário criar uma BD vazia no Access.

A importação dos dados meteorológicos é realizada manualmente, pois são constituídos apenas por uma tabela.

AccessMedias.bas (Anexo A):

Este programa deve ser devidamente editado consoante a BD onde é utilizado. Inicialmente, define-se para que parâmetros da BD e para que intervalo de tempo (anos e meses) se pretende calcular as médias mensais. De seguida, para cada mês é efetuada uma consulta SQL que retorna várias informações para cada estação, como as coordenadas WGS84 e PTTM06 da estação, o ano, o mês e a média mensal do parâmetro. Para terminar, a informação fornecida pela consulta é armazenada num ficheiro CSV.

3.5.2 Modelo Digital de Elevação para Portugal Continental

Como foi referido anteriormente, o MDE do SRTM encontra-se dividido em quadrículas de 1° de latitude por 1° de longitude, por isso foi necessário unir todas as imagens. Para isso, utilizou-se a função “gdal_merge” da biblioteca *Geospatial Data Abstraction Library* (GDAL) na linha de comandos do *software* OSGeo4W. Após a união, converteu-se a imagem final para o formato TIF, recorrendo à função “gdal_translate”. Finalmente, com recurso à função “gdalwarp” reduziu-se a resolução do MDE de 100 m para 5 km para ser utilizado na determinação de T_a com os dados MODIS. Neste caso, utilizou-se o método de reamostragem bilinear, que para calcular um pixel da nova imagem considera a média ponderada dos quatro pixels mais próximos.

3.5.3 Determinação da Temperatura e Humidade Relativa do Ar a partir do Perfil Atmosférico MODIS

Na determinação das médias mensais de T_a e HR a partir do perfil atmosférico MOD07 recorreu-se a um conjunto de seis programas, desenvolvidos no âmbito deste trabalho e implementados em MATLAB e Python (Tabela 3.5). De referir que, os dados MOD07 em formato HDF encontram-se armazenados com a seguinte estrutura de diretórios: “Satélite/Ano”.

Tabela 3.5. Programas utilizados para calcular as médias mensais de T_a e HR a partir do MOD07.

Nome do Programa	Objetivo
MODIS_AP.m	Calcular T_a e extrair o TCWV do MOD07
MODIS_Atmosphere_move_files.m	Preparar os dados diários, calculados em MODIS_AP.m, para o cálculo das médias mensais
MODIS_AP_avg.m	Calcular as médias mensais de T_a e TCWV
MODIS_L2_Geotiff.py	Converter as médias mensais de todos os parâmetros para o formato TIF na projeção PTTM06
MODIS_AP_val.py	Validar os dados de satélite com base nas medições realizadas pelas estações meteorológicas
MODIS_RH_avg.m	Calcular as médias mensais de HR

MODIS AP.m (Anexo B):

Para iniciar, é feita a leitura do MDE de 5 km (MDE5) e estimada a altitude para cada nível de pressão a utilizar, recorrendo à expressão (3.3). Como para Portugal Continental a altitude máxima (no MDE5) 1871 m corresponde a uma pressão de, aproximadamente, 805 mbar, utilizaram-se os seis níveis de pressão mais elevados dados no MOD07 (700, 780, 850, 920, 950 e 1000 mbar). De seguida, utilizando as coordenadas dos cantos do MDE5, são construídas duas grelhas regulares de latitude e longitude, que serão utilizadas na interpolação dos dados MOD07.

Posteriormente, para cada ficheiro MOD07, são extraídas as matrizes de latitude, longitude, PT_a , TCWV e indicadores de qualidade (IQT e IQWV, para PT_a e TCWV, respetivamente). Apenas as camadas dos seis níveis de pressão mais elevados do PT_a são utilizadas, sendo ignorado (colocando-os como falta de dados – NaN no MATLAB) cada pixel de cada camada não declarado como “*Best Quality*” nos indicadores de qualidade. Após a correção, os valores reais de PT_a e de TCWV são calculados, aplicando um fator de escala e um *offset*, de acordo com as especificações do produto (p.e., o fator de escala e o *offset* do perfil de T_a são, respetivamente, 0.01 e -15000):

$$\text{valor real} = \text{fator de escala} \times \text{valor armazenado} + \text{offset} \quad (3.22)$$

O PT_a e o TCWV são de seguida interpolados para grelhas regulares, utilizando as matrizes de latitude e longitude, definidas inicialmente. O método de reamostragem utilizado na interpolação é o bilinear.

Por fim, a T_a é calculada a partir do PT_a . Para tal, recorrendo à função “polyfit” do MATLAB (baseia-se no método dos mínimos quadrados), determina-se a equação da reta de melhor ajuste ao PT_a correspondente a cada pixel do MDE5. Os parâmetros da reta (declive e ponto de interseção com o eixo dos yy) em cada pixel e a altitude, são utilizados para calcular T_a pela expressão (3.4). Por fim, T_a e TCWV, dispostos em grelhas regulares, são guardados em ficheiros MAT do MATLAB.

MODIS Atmosphere move files.m (Anexo B):

Os ficheiros MAT, criados anteriormente, são agrupados em diretórios mensais (p.e., 01, corresponde a janeiro, 02 a fevereiro, etc.), através da leitura do respetivo nome. De notar que o agrupamento em diretórios mensais inclui a junção no mesmo diretório dos ficheiros MAT correspondentes a dados dos satélites Aqua e Terra. A estrutura de diretórios resultante é a seguinte: “DadosAqua&Terra/Ano/Mes”.

MODIS AP avg.m (Anexo B):

O programa percorre cada diretório anual e mensal, calculando a média mensal das variáveis T_a e TCWV, presentes nos ficheiros MAT. No cálculo das médias os pixéis sem dados (NaN) não são considerados. Como referido na secção 3.1, as médias mensais de T_a são calculadas para cada mês com quatro combinações de dados.

Para terminar, as médias de T_a são convertidas de graus Kelvin para graus Celsius e as grelhas são guardadas em ficheiros ASCII no mesmo diretório (cada mês com quatro ficheiros, correspondentes às quatro combinações de dados consideradas).

MODIS L2 Geotiff.py (Anexo B):

Este programa utiliza a função `gdal_translate` da GDAL para converter os ficheiros ASCII para o formato TIF e projetá-los de WGS84 para PTTM06.

MODIS AP val.py (Anexo B):

O programa recorre aos ficheiros CSV com as médias mensais dos dados das estações meteorológicas, criados anteriormente na secção 3.5.1, para validar os dados de satélite (em formato TIF e em PTTM06). Para cada imagem de satélite, o programa começa por abrir o ficheiro CSV correspondente e, para cada estação, guarda o código, as

coordenadas PTTM06, o mês, o ano e o valor do parâmetro, medido pela estação. Recorrendo à função “gdallocationinfo” da biblioteca GDAL, as coordenadas PTTM06 da estação são utilizadas para obter o valor do parâmetro na imagem de satélite. Juntamente com o código da estação, o ano, o mês e o valor do parâmetro medido, o valor do parâmetro dado pela imagem de satélite é armazenado num ficheiro final (formato CSV). Desta forma, o ficheiro final irá conter, para todo o intervalo de tempo em estudo, os valores do parâmetro, medidos pelas estações, e respetivos valores dados pelas imagens de satélite. Com esta informação é possível calcular medidas que permitam aferir a precisão dos dados.

MODIS RH avg.m (Anexo B):

Inicialmente é feita a leitura do MDE5, que será utilizado no cálculo de T_d . Posteriormente, para cada par de imagens mensais de T_a e TCWV é calculada T_d , a partir da expressão (3.5). Após os cálculo de T_d , a HR é calculada pela expressão (3.6). Os resultados da HR mensal são guardados em ficheiros ASCII.

Após aplicar “MODIS_AP_val.py, foi selecionado para cada mês a combinação de dados com menor EMQ para estimar T_a . Na Figura 3.7 é apresentado um fluxograma com a metodologia de determinação destas variáveis.

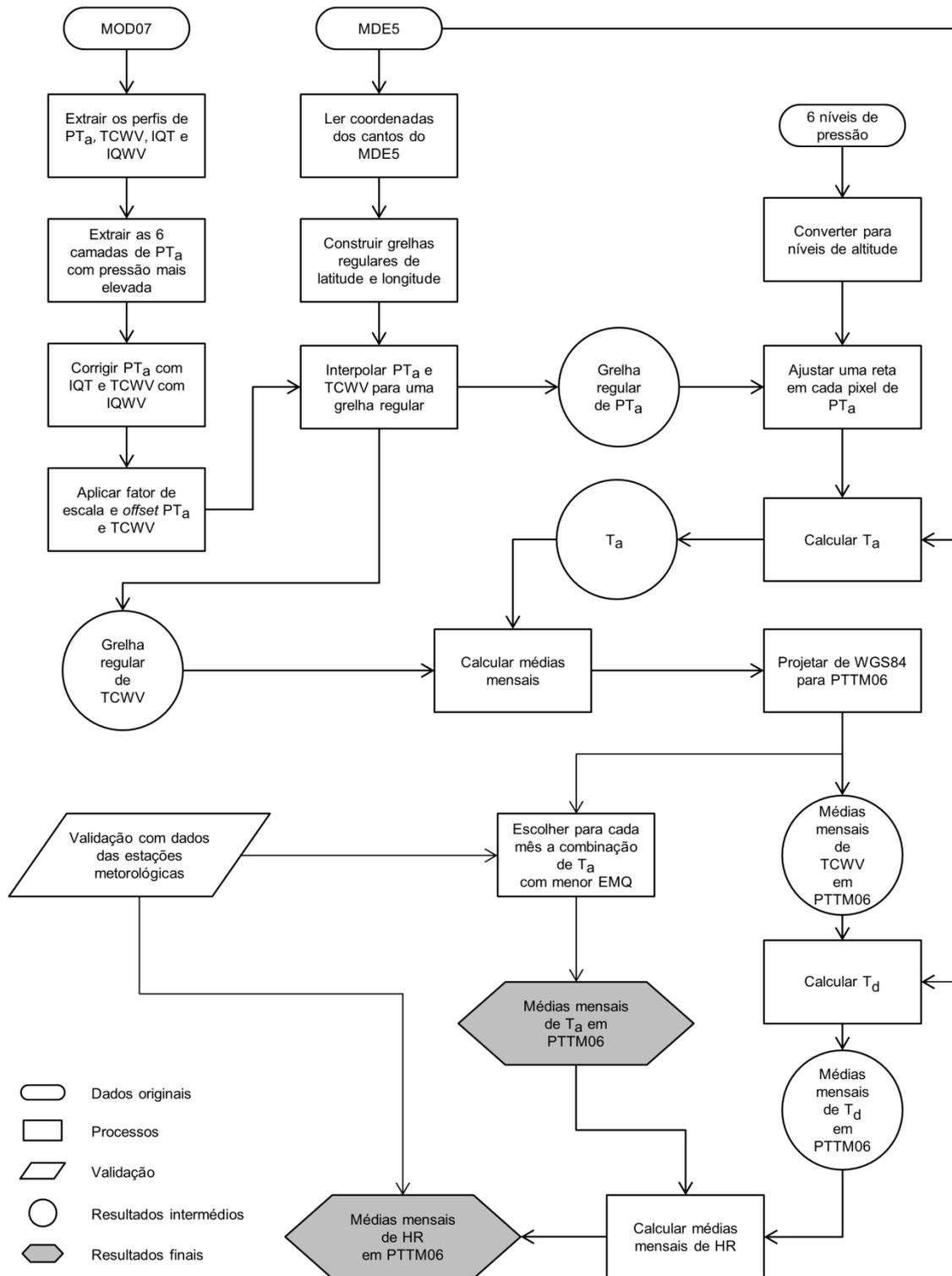


Figura 3.7. Fluxograma com a metodologia de determinação das médias mensais de T_a e HR.

3.5.4 Aerosol Optical Thickness

A determinação do AOT mensal é composta por uma sequência de quatro programas, implementados em MATLAB e Python (Tabela 3.6). Os dados MOD04 em formato HDF encontram-se armazenados com a seguinte estrutura de diretórios: “Satélite/Ano/”.

Tabela 3.6. Programas utilizados no cálculo das médias mensais do AOT550 a partir do MOD04.

Nome do Programa	Objetivo
MODIS_AOT.m	Extrair o AOT no comprimento de onda de 550 nm dos ficheiros MOD04
MODIS_Atmosphere_move_files.m	Preparar os dados diários do AOT, extraídos em MODIS_AOT.m, para o cálculo das médias mensais
MODIS_AOT_avg.m	Calcular as médias mensais de AOT
MODIS_L2_Geotiff.py	Converter médias mensais de todos os parâmetros para o formato TIF na projeção PTTM06

MODIS AOT.m (Anexo B):

O primeiro programa é utilizado para extrair o AOT550 (Optical_Depth_Land_And_Ocean). Neste caso, não é necessário realizar qualquer correção, pois o parâmetro Optical_Depth_Land_And_Ocean é determinado somente com os dados de melhor qualidade [104]. Assim, o programa simplesmente começa por gerar duas grelhas regulares de latitude e longitude. De seguida, para cada ficheiro MOD04 extrai a latitude, a longitude e o parâmetro AOT550, referido em cima. Ao AOT550 é aplicado um fator de escala e um *offset* (0.001 e 0, respetivamente), de acordo com a expressão (3.22), e os pixéis com valor menor que -0.05 são colocados a NaN. Após esta operação, o AOT550 é interpolado, utilizando as grelhas regulares de latitude e longitude, definidas inicialmente. O método de interpolação utilizado é novamente o bilinear. O AOT550 interpolado é guardado num ficheiro MAT.

MODIS Atmosphere move files.m:

Descrito na secção 3.5.3.

MODIS AOT avg.m (Anexo B):

Os diretórios anuais e mensais são percorridos pelo programa, sendo calculada a média mensal do AOT550. Neste cálculo os pixels sem dados (NaN) não são considerados. À semelhança do MODIS_AP_avg.m, as médias mensais do AOT550 são guardadas em ficheiros ASCII no mesmo diretório.

MODIS L2 Geotiff.py:

Descrito na secção 3.5.3.

Importa referir, que em algumas imagens mensais de AOT550 (total de 17 imagens) o método de Kriging foi aplicado no ArcGIS com parâmetros por defeito, para estimar o valor de pixels com falta de dados.

3.5.5 Normalized Difference Vegetation Index

Como o NDVI presente no produto MOD13A3 é um produto mensal, apenas foi necessário aplicar os indicadores de qualidade e projetar as imagens para PTTM06. Para este fim, utilizou-se uma sequência de dois programas, implementados em MATLAB e Python (Tabela 3.7).

Os dados MOD13A3 em formato HDF encontram-se armazenados no mesmo diretório.

Tabela 3.7. Programas utilizados para preparar o NDVI mensal do MOD13A3.

Nome do Programa	Objetivo
MODIS_NDVI_QA.m	Corrigir os pixels do NDVI com os respetivos indicadores de qualidade
MODIS_NDVI_reproject.py	Unir imagens NDVI do mesmo dia e reprojeta-las para PTTM06

MODIS NDVI QA.m (Anexo B):

O programa começa por ler as grelhas NDVI e indicador de qualidade (IQNDVI) e depois procede à correção do NDVI. Apenas os pixels que sejam declarados no IQNDVI como “*VI produced, good quality*” e com utilidade entre “*Highest quality*” e “*Lowest quality*” são utilizados. Os restantes são declarados como falta de dados (-3000, que representa falta de dados no NDVI, foi o valor utilizado). De seguida, a variação do NDVI é reduzida de -

2000 a 10000 para 0 a 10000 e é aplicado um fator de escala e um *offset* (0.001 e 0, respetivamente), de acordo com a expressão (3.22).

As imagens NDVI corrigidas são guardadas em ficheiros HDF (“*.h5”).

MODIS NDVI reproject.py (Anexo B):

Em primeiro lugar, os ficheiros NDVI H5 são convertidos para TIF com a função `gdal_translate`. Neste processo, a cada imagem TIF é associada a informação sobre o sistema de coordenadas dos ficheiros MOD13A3, pois esta tinha sido perdida durante a correção do NDVI no programa anterior. Para se identificar os parâmetros da projeção utilizou-se função “`gdalinfo`” na linha de comandos num ficheiro HDF MOD13A3 inicial.

Em segundo lugar, como Portugal Continental é abrangido por duas quadrículas (quadrículas h17v04 e h17v05) de dados MOD13A3, as imagens TIF antes de serem projetadas têm de ser unidas. Para tal, o programa utiliza a função “`gdal_merge`” para unir as quadrículas h17v04 e h17v05, correspondentes ao mesmo mês.

Em último lugar, as imagens TIF resultantes da união são reprojctadas de MODISSIN para PTTM06, sendo também reduzidas apenas para área de Portugal Continental (as duas quadrículas abrangem toda a Península Ibérica).

Na Figura 3.8 é apresentado um fluxograma com a metodologia de processamento do NDVI.

3.5.6 Determinação de Variáveis Ambientais pelo Método de Kriging

O método de Kriging foi utilizado juntamente com os dados das estações meteorológicas e de qualidade do ar para determinar estimativas espaciais mensais de cinco variáveis ambientais: T_a , HR, CO, NO_2 e PM_{10} . Neste trabalho, utilizou-se o método de Kriging Ordinário, implementado no *software* ArcGIS 10.

A automatização do método de Kriging no ArcGIS pode ser feita de duas formas: utilizando a função “Kriging” da ferramenta “Spatial Analyst” ou utilizando a ferramenta “Geostatistics”. A primeira forma é a mais simples mas não permite realizar validação cruzada. A validação cruzada consiste em remover um ponto e prever o seu valor com os restantes. Repetindo este processo para todos os pontos é possível aferir a precisão do modelo.

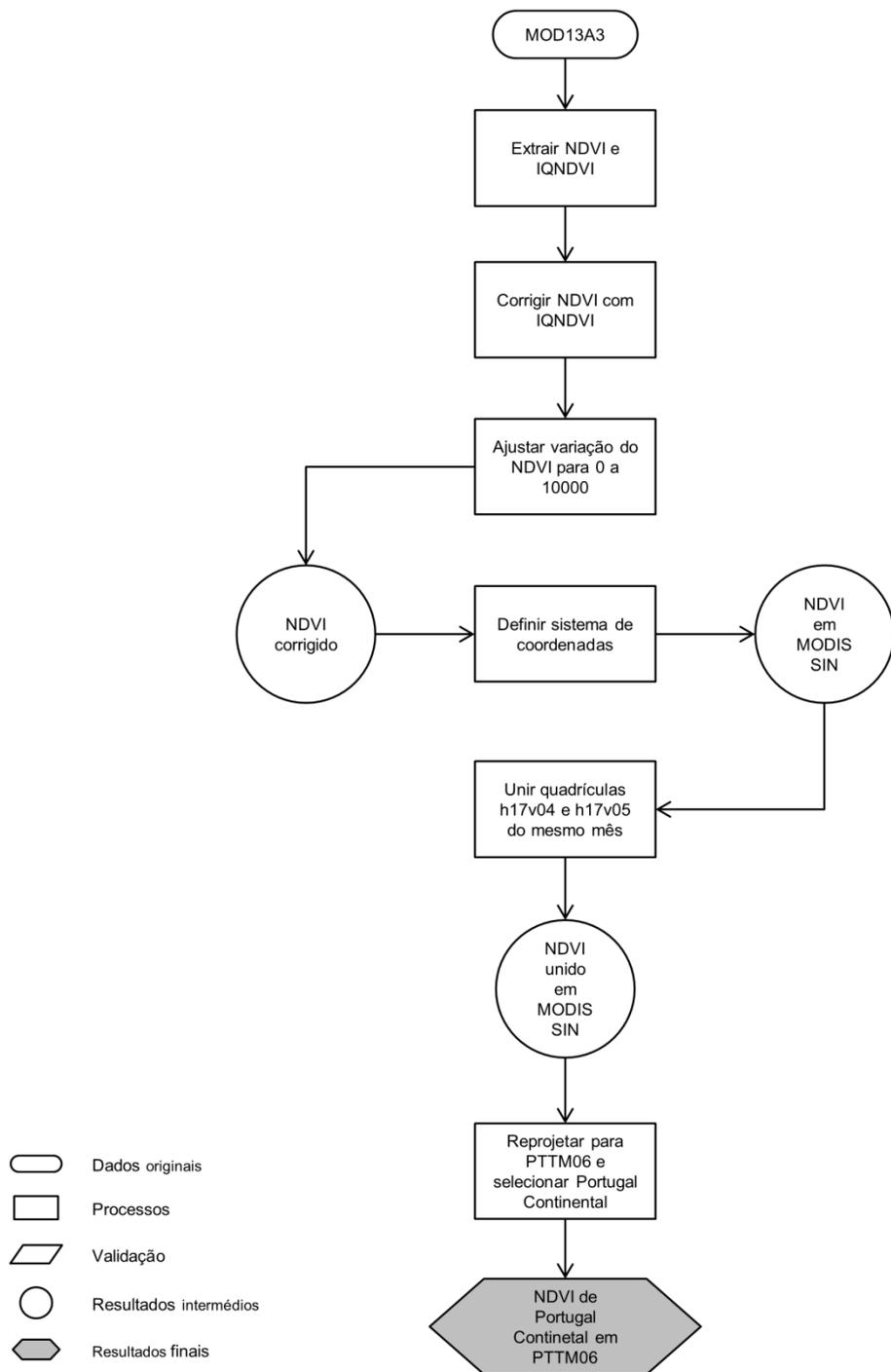


Figura 3.8. Fluxograma com a metodologia de processamento do NDVI.

Tendo em conta a limitação da primeira forma, as estimativas espaciais mensais foram determinadas a partir da ferramenta “Geostatistics”. Contudo, a automatização do processo não é direta. Face a este problema, foi implementada uma sequência de programas em MATLAB e Python (Tabela 3.8).

Tabela 3.8. Programas utilizados para determinar, pelo método de Kriging, estimativas espaciais mensais das variáveis ambientais.

Nome do Programa	Objetivo
WriteKrigingModel.m	Escrever um ficheiro XML com os parâmetros do modelo esférico ajustado ao semi-variograma dos dados de cada ficheiro CSV das variáveis ambientais
csv2shp.py	Converter ficheiros CSV para SHP de pontos
ArcGIS_Geostatistics.py	Criar estimativas espaciais mensais das variáveis ambientais, a partir do método de Kriging e realizar a respetiva validação cruzada

Para determinar automaticamente, através do “Geostatistics”, estimativas espaço-temporais de um parâmetro A pelo método de Kriging, é necessário começar por construir um modelo em linguagem XML, que dará origem a uma camada (*layer*) do tipo “Geostatistics”. Esse modelo deve conter, entre outros, os parâmetros do variograma, determinados pelo ajuste de uma função ao semi-variograma, calculado a partir das medidas pontuais do parâmetro A.

Se as estimativas forem determinadas manualmente a partir do interface gráfico do “Geostatistics”, os parâmetros do variograma são ajustados automaticamente aos dados pontuais. Através da construção do modelo XML os parâmetros podem ser determinados automaticamente durante a construção da camada “Geostatistics” mas, os valores serão sempre iguais, independentemente dos dados. Para resolver esse problema implementou-se um programa (WriteKrigingModel.m), que permite calcular os parâmetros do variograma a partir de um conjunto de dados e escrever o modelo para um ficheiro XML.

WriteKrigingModel.m (Anexo C):

O programa lê o primeiro ficheiro CSV (data mais antiga) de uma determinada variável ambiental (definida no início) e guarda as coordenadas de cada estação e os respetivos valores da variável ambiental. De seguida, recorre à função “variogram.m”, para calcular o semi-variograma, e à função “variogramfit.m”, para ajustar o modelo esférico e determinar os respetivos parâmetros. Estas funções foram implementadas por Wolfgang Schwanghart e encontram-se disponíveis gratuitamente em <http://www.mathworks.com/>

MATLABcentral/fileexchange/. À exceção do modelo do variograma, que foi modificado para esférico, as restantes opções das funções foram mantidas por defeito. De notar, que por defeito o efeito de pepita não é calculado, ou seja, é igual a 0.

Por fim, é escrito o ficheiro XML com o modelo. A sintaxe XML deste tipo de modelos pode ser obtida sempre que se realizar uma interpolação pelo método de Kriging no interface gráfico do “Geostatistics”.

Estes procedimentos são repetidos para todos os ficheiros CSV de todas as variáveis ambientais definidas no início do programa.

csv2shp.py (Anexo C):

Este programa permite converter um ficheiro de texto em formato CSV para um ficheiro SHP de pontos, utilizando da função “ogr2ogr” da biblioteca GDAL. Esta operação é essencial, pois a criação automática de uma camada do tipo “Geostatistics” não permite que os dados pontuais sejam introduzidos em formato texto.

O programa começa por ler o primeiro ficheiro CSV (data mais antiga) de uma determinada variável ambiental (definida no início) e constrói um ficheiro no formato *Virtual* (VRT), que serve de auxílio à conversão do CSV para SHP. Os ficheiros VRT são escritos em XML e devem ser estruturados mediante alguns critérios definidos em [105]. A partir destes critérios os ficheiros VRT indicam à função “ogr2ogr” várias informações sobre os ficheiros CSV (p.e., coluna que representa a coordenada x, a coordenada y, etc.).

Após a construção do ficheiro VRT, o programa utiliza a função “ogr2ogr” para converter o ficheiro CSV para SHP. Neste processo as coordenadas WGS84 das estações são também convertidas para PTTM06.

Estes procedimentos são repetidos para todos os ficheiros CSV de todas as variáveis ambientais definidas no início do programa.

ArcGIS Geostatistics.py (Anexo C):

Este programa cria as estimativas espaciais mensais das variáveis ambientais, utilizando funções do ArcGIS. Para iniciar, é utilizada a função “Create Geostatistical Layer” com o modelo XML e respetivo ficheiro SHP como argumentos, para criar uma camada “Geostatistics”. Posteriormente, o programa guarda a camada criada para um ficheiro *Layer* (LYR), a partir da função “Save To Layer File”. Após a camada ser guardada, é efetuada a sua validação cruzada, através da função “Cross Validation”.

O programa termina com a conversão do ficheiro LYR para TIF com 1000 m de resolução espacial, para os dados meteorológicos, e 100 m, para os dados dos poluentes.

Estes procedimentos são repetidos para todos os ficheiros SHP de todas as variáveis ambientais definidas no início do programa.

Na Figura 3.9 é apresentado o fluxograma com a metodologia de determinação das estimativas espaciais mensais das variáveis ambientais, recorrendo ao método de Kriging.

3.5.7 Determinação dos Poluentes por Modelos LUR

A criação dos modelos LUR compreendeu quatro fases. A primeira fase consistiu na preparação das variáveis independentes sem variabilidade temporal (p.e., distância a estradas). Na segunda fase, foi elaborado um programa em Python (“LUR.py”), que permitiu associar a cada medição pontual de um poluente, os respetivos valores das variáveis independentes. A terceira fase diz respeito à construção e validação do modelo LUR. Por fim, na quarta fase, foi elaborado um programa em Python (“LUR_Rasters.py”) para construir as imagens mensais da concentração dos poluentes, tendo em conta os modelos LUR criados.

1ª Fase – Preparação das variáveis sem variabilidade temporal:

Na 1ª fase, procedeu-se à preparação das variáveis sem variabilidade temporal, apresentadas na Tabela 3.1. As variáveis que representam “Densidades”, “Distâncias” e “Ocupação de solo” e o Declive foram preparadas no ArcGIS. As grelhas com a Latitude e com a Longitude foram elaboradas no MATLAB e guardadas em ficheiros ASCII, que, por sua vez, foram convertidos para TIF com o programa “MODIS_L2_Geotiff.py” (Tabela 3.5).

Para iniciar a elaboração dos mapas de densidades, utilizaram-se consultas por atributos para criar três ficheiros SHP com as estradas do tipo IP, IC e EN, a partir do ficheiro SHP com a rede viária da EP. Depois, recorrendo à função “Focal Statistics” da extensão “Spatial Analyst” do ArcGIS, construíram-se vários mapas de densidade para cada tipo de estrada, considerando *buffers* com 250, 500 e 750 m de raio. A mesma função foi utilizada para criar os mapas de densidade total de estradas, considerando os mesmos critérios e o ficheiro SHP com a rede viária do OpenStreetMap.

Recorrendo à função “EuclidianDistance” da extensão “Spatial Analyst”, elaboraram-se os mapas de distâncias a diferentes tipos de estradas, a partir dos ficheiros SHP com estradas do tipo IP, IC e EN e com a rede viária do OpenStreetMap (todos os tipos de estradas). Esta função foi também utilizada para criar o mapa com a distância à linha de

costa. A linha que delimita a costa foi extraída da CAOP2011, a partir de um processo de edição gráfica.

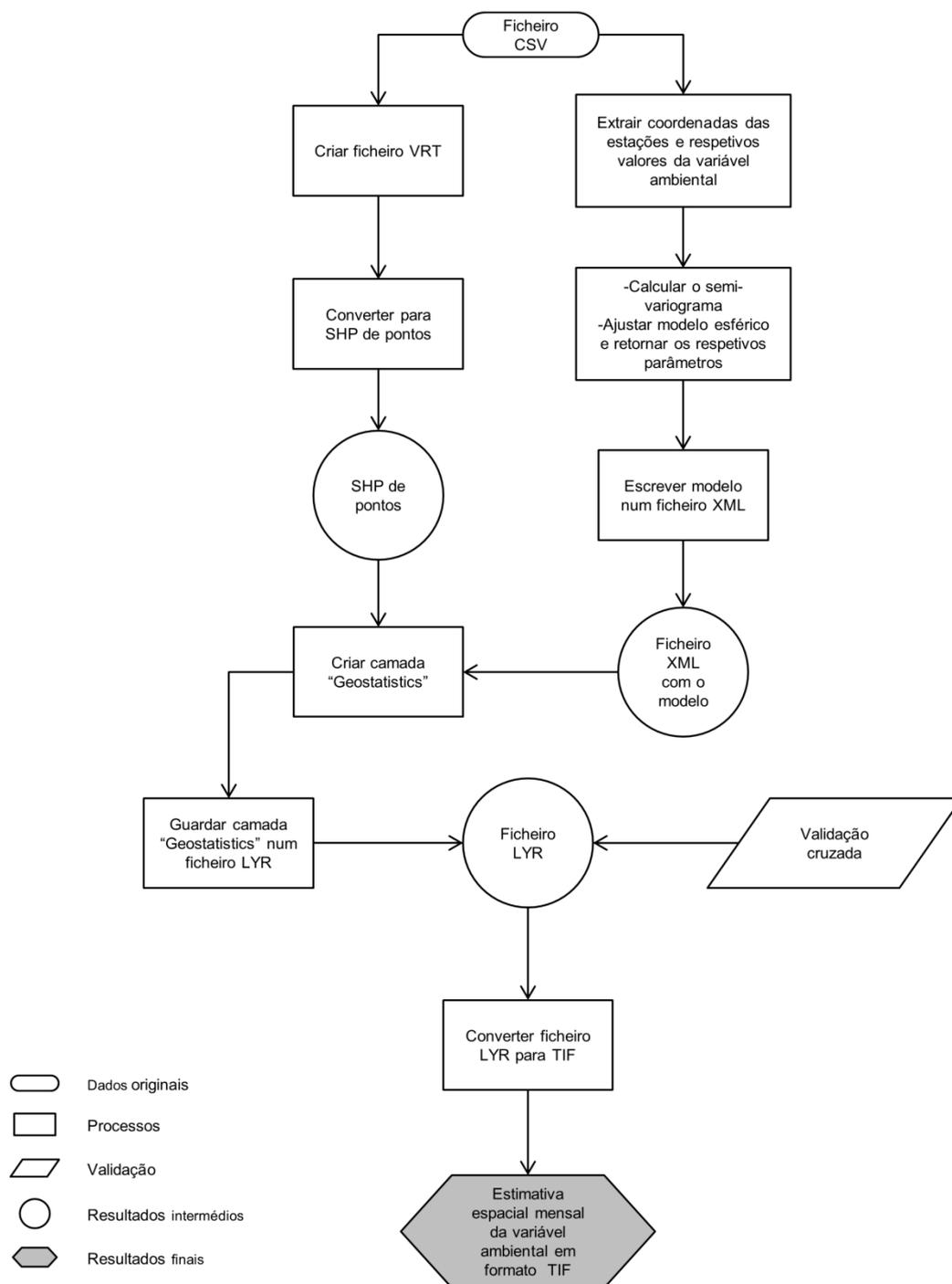


Figura 3.9. Fluxograma com a metodologia para a determinação das estimativas espaciais mensais das variáveis ambientais, recorrendo ao método de Kriging.

Para determinar os mapas com a percentagem de ocupação de solo a distâncias de 250, 500 e 750 m, foi utilizado o ficheiro SHP CLC06_PT. Através de consultas por atributos agruparam-se as classes do CLC06_PT, conforme descrito na Tabela 3.2. Este processo resultou em sete novos ficheiros SHP, cada um correspondendo à nova classe. De seguida, os novos ficheiros SHP foram convertidos para formato *raster* com dimensão binária (p.e., no *raster* da classe Urb1, pixéis com o valor 1 pertencem à classe Urb1 e pixéis com o valor 0 pertencem a outra classe) com resolução espacial de 50 m. Posteriormente, criaram-se três novos mapas para cada *raster* binário, com o número de pixéis com valor 1 que se encontram dentro de *buffers* com, respetivamente, 250, 500 e 750 m de raio. Esta operação foi realizada com a função “Focal Statistics”.

No final, sabendo o número de pixéis que abrange cada *buffer* (81, 317 e 709 pixéis, respetivamente), os resultados anteriores foram convertidos para percentagem.

O mapa de declive foi criado com uma resolução espacial de 100 m a partir do MDE, utilizando a função “Slope” da extensão “Spatial Analyst” do ArcGIS.

2ª Fase – LUR.py (Anexo D):

No início deste programa é definido para que poluentes se pretende realizar as associações com as variáveis independentes.

O restante processo é idêntico ao do programa “MODIS_AP_val.py”, exceto no facto de para cada medição mensal de uma determinada estação serem associados 44 parâmetros diferentes (Tabela 3.1) em vez de apenas um (imagem de satélite de T_a ou HR).

No final, para cada poluente, o programa guarda as respetivas associações com as variáveis independentes num ficheiro CSV.

3ª Fase – Construção e validação do modelo LUR:

O ficheiro CSV foi aberto no *software* IBM SPSS Statistics 20 e a variável dependente e respetivas variáveis independentes foram introduzidas num processo de regressão *stepwise*. Diferentes combinações de variáveis independentes foram comparadas através do R^2 e a melhor combinação foi escolhida. No decorrer do processo de regressão *stepwise*, o coeficiente de correlação entre pares de variáveis independentes foi examinado para evitar multicolinearidade.

Os modelos foram validados no IBM SPSS Statistics 20, seguindo o método de validação cruzada *10-fold*. Este método consiste em dividir aleatoriamente os dados em 10 partes (aproximadamente) iguais. De seguida, o modelo é novamente construído com 9 partes (dados de treino) e validado com a parte excluída (dados de teste). Este processo é

executado 10 vezes até que todas as partes sejam utilizadas como dados de teste [106]. O R^2 médio dos dados de teste é utilizado para avaliar o modelo.

Por fim, verificou-se se as suposições da regressão linear múltipla foram cumpridas, observando se os resíduos do modelo possuíam distribuição normal centrada em 0 (média) e se não apresentavam nenhuma tendência.

4ª Fase – LUR Rasters.py (Anexo D):

O programa aplica os modelos LUR, criados anteriormente, para determinar as estimativas mensais da concentração dos poluentes. Os vários cálculos realizados para aplicar a fórmula de cada modelo são efetuados com as funções do ArcGIS “Times” (multiplicação), “Plus” (soma), “Power” (potência) e “Exp” (exponenciação). A utilização destas funções tem como vantagem permitir cálculos entre imagens com elevadas resoluções. De referir, que todas as imagens de baixa resolução foram reamostradas para 100 m.

3.5.8 Análise Estatística

Antes de se proceder à análise estatística foi necessário realizar um pré-processamento de dados. Para isso, recorreu-se a um conjunto de programas, implementos em MATLAB e Python (Tabela 3.9).

Tabela 3.9. Programas utilizados para preparar os dados para a análise estatística.

Nome do Programa	Objetivo
var_avg.m	Calcular a média das cinco variáveis de interesse para o intervalo de tempo em estudo
ArcGIS_ZonalStatistics.py	Criar tabelas mensais (formato DBF) com o valor médio de um determinado parâmetro, por concelho (ou distrito)

var_avg.m (Anexo E):

Este programa calcula a média de cada variável ambiental para o intervalo de tempo em estudo. Cada variável é guardada como um ficheiro ASCII.

ArcGIS_ZonalStatistics.py (Anexo E):

Os parâmetros que se pretende agrupar por concelho são definidos no início do programa. Para cada imagem (ficheiro ASCII) de cada parâmetro, o programa determina

a respetiva média mensal em cada concelho, recorrendo à função “ZonalStatisticsAsTable” da extensão “Spatial Analyst” do ArcGIS. Os limites de cada concelho são definidos pela CAOP2011. Como resultado final, cada imagem de cada parâmetro possui associada uma tabela com as médias por concelho em formato DBF.

Após a utilização dos dois programas acima referidos, criou-se manualmente um ficheiro XLS no Excel com as médias por concelho das cinco variáveis ambientais. Por fim, realizou-se a análise estatística no IBM SPSS Statistics 20.

Capítulo 4

Resultados

4.1 Variáveis Ambientais

4.1.1 Temperatura e Humidade Relativa do Ar (MODIS)

A validação das médias mensais de T_a determinadas com quatro combinações de dados encontra-se apresentada na Tabela 4.1. Este processo revelou a existência de uma espécie de *offset* sazonal, que atinge valores mais elevados no verão. Na primeira combinação de dados (T_aAT), apesar do desvio padrão manter-se relativamente constante ao longo dos meses, verifica-se a existência de um erro sistemático, principalmente nos meses de verão (erro médio elevado). Na segunda e terceira combinações de dados (T_aATn e T_aAnT , respetivamente), verifica-se que a ausência das observações realizadas durante o dia produz resultados idênticos para os dois satélites. Exceto nos meses de janeiro, fevereiro, março e abril, observa-se um decréscimo no erro médio para todos os meses. De qualquer forma, o decréscimo é mais acentuado quando as observações do satélite Aqua realizadas durante o dia são removidas. Trata-se de um resultado esperado, pois a hora de passagem do satélite Aqua por Portugal Continental (entre 13h-15h) coincide, aproximadamente, com a hora de ocorrência da T_a máxima (entre 14h-16h). Na quarta e última combinação (T_aAnTn), apenas as observações dos satélites realizadas durante a noite foram consideradas. Para os meses de verão o erro continua a descer, não existindo alguma evidência da permanência do erro sistemático, conforme se pode verificar pelo erro médio próximo de 0. O aumento do EMQ nos meses

de verão poderá estar relacionado com as elevadas temperaturas características desta estação do ano, uma vez que a radiação solar afeta as bandas no infravermelho térmico. Como referido anteriormente, a combinação de dados com menor EMQ (a cinzento na Tabela 4.1) em cada mês foi selecionada como estimativa de T_a . Além disso, foi observada uma boa correlação entre a T_a medida e prevista (Figura 4.1), apresentando um declive da equação de regressão próximo de 1 e um termo constante próximo de 0. Desta forma, demonstra-se que com um perfil de temperatura derivado de observações de satélites, uma grelha com a altitude e um simples processo de interpolação e seleção é possível determinar estimativas precisas de T_a . Alternativamente, o aumento sazonal do EMQ poderia ser corrigido através da construção de um modelo que permitisse calibrar T_a (calculado com a combinação de dados T_a AT) para cada mês.

Na Tabela 4.2, para além dos resultados finais de T_a , também se encontram os resultados de T_d , calculada com o TCWV mensal e com a altitude do MDE5, e de HR, calculada com T_a e T_d . A precisão de T_d é idêntica a T_a , com um EMQ de 1.56 °C e um R^2 ligeiramente inferior (0.86). Neste caso, a relação entre a T_d medida e prevista (Figura 4.2) apresenta alguma assimetria.

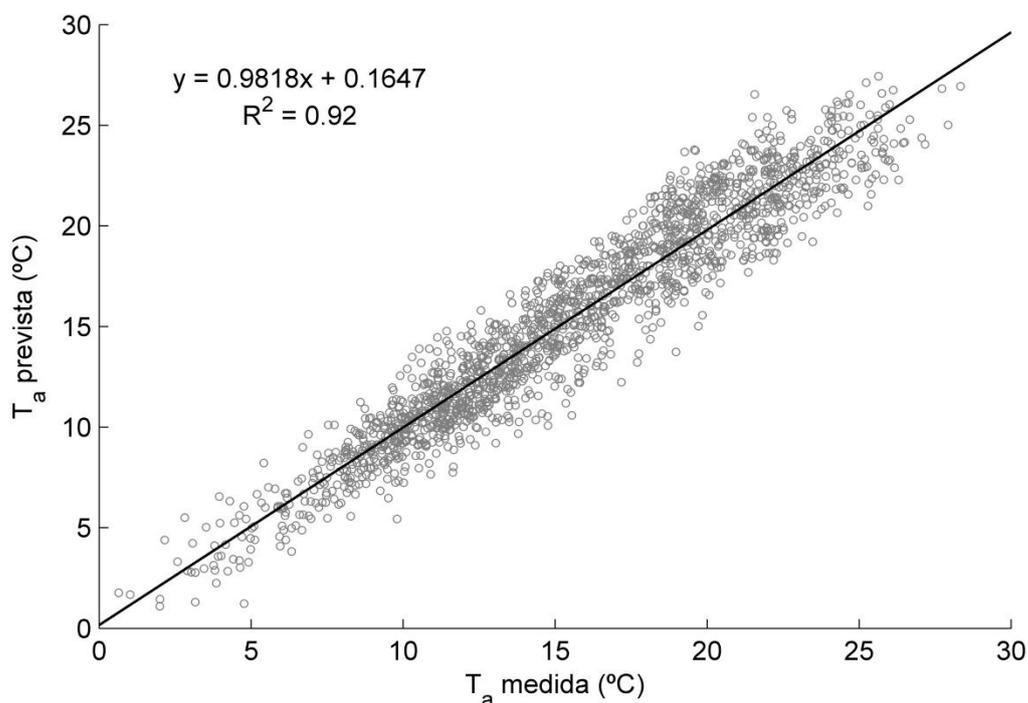


Figura 4.1. Relação entre a T_a medida e prevista (MOD07).

Tabela 4.1. Validação mensal de cada combinação de dados considerada no cálculo das médias mensais de T_a . A sombreado encontra-se a combinação de dados com menor EMQ para cada mês.

Mês	T_{aAT} (°C)			T_{aATn} (°C)			T_{aAnT} (°C)			T_{aAnTn} (°C)		
	EMQ	EM	DP	EMQ	EM	DP	EMQ	EM	DP	EMQ	EM	DP
Janeiro	1.21	0.00	1.21	1.52	-0.70	1.35	1.34	-0.52	1.24	2.36	-1.81	1.52
Fevereiro	1.05	-0.26	1.02	1.50	-0.96	1.15	1.34	-0.85	1.04	2.54	-2.17	1.33
Março	1.87	-1.24	1.40	2.45	-1.91	1.54	2.35	-1.86	1.45	3.56	-3.14	1.69
Abril	1.09	0.20	1.07	1.37	-0.61	1.23	1.39	-0.73	1.19	2.79	-2.35	1.50
Maio	1.63	0.95	1.33	1.47	-0.05	1.48	1.41	-0.12	1.41	2.72	-2.14	1.69
Junho	3.21	2.90	1.39	2.40	1.89	1.48	2.22	1.52	1.62	2.04	-0.79	1.88
Julho	4.12	3.87	1.41	3.19	2.83	1.48	2.85	2.35	1.61	1.85	-0.20	1.85
Agosto	4.38	4.12	1.48	3.60	3.24	1.59	3.14	2.68	1.65	1.97	0.42	1.93
Setembro	3.13	2.78	1.43	2.45	1.88	1.57	2.21	1.57	1.56	1.92	-0.52	1.86
Outubro	1.81	1.48	1.04	1.35	0.76	1.12	1.31	0.62	1.16	1.68	-0.86	1.44
Novembro	1.37	0.75	1.15	1.26	0.18	1.26	1.22	0.15	1.21	1.75	-0.95	1.47
Dezembro	1.30	0.76	1.06	1.21	0.02	1.21	1.15	0.33	1.11	1.69	-0.96	1.40

T_{aAT} – Todas as observações dos satélites Aqua e Terra;

T_{aATn} – Todas as observações do satélite Aqua e as observações do satélite Terra durante a noite;

T_{aAnT} – Todas as observações do satélite Terra e as observações do satélite Aqua durante a noite;

T_{aAnTn} – Observações durante a noite dos satélites Aqua e Terra;

EM – Erro Médio; DP – Desvio Padrão.

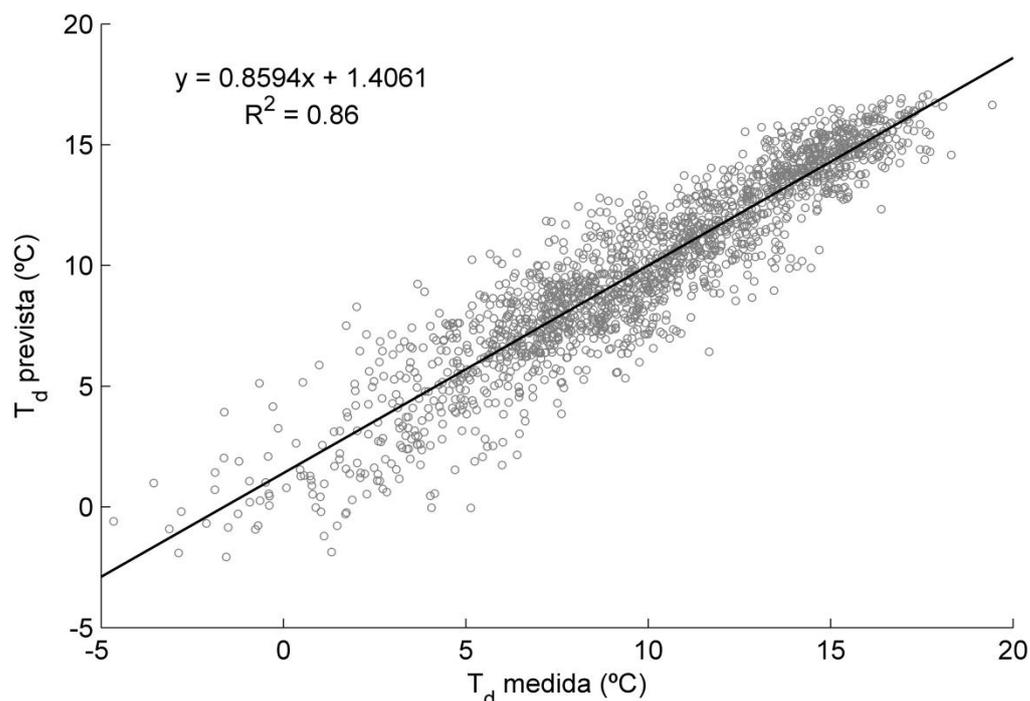


Figura 4.2. Relação entre a T_d medida e prevista (MOD07).

Como se trata de um parâmetro determinado com as medições realizadas nas estações meteorológicas, realizou-se validação cruzada *10-fold*. O R^2 da validação (0.85) e o erro médio muito próximo de 0 evidenciam o bom comportamento do modelo.

Para as médias mensais de HR obteve-se um EMQ de 8.76%, que se pode considerar aceitável. Porém, como se pode observar na Figura 4.3, a correlação entre a HR medida e prevista é baixa, resultando numa falta de simetria entre os valores medidos e previstos (declive da equação de regressão próximo de 0.5). Para além disso, o erro médio (Tabela 4.2), calculado a partir dos resíduos de HR mostra a existência de um erro sistemático negativo (-1.45%). Estes resultados eram de esperar, pois a HR é calculada a partir de uma equação não linear (equação (3.6)), que recebe como *input* duas quantidades estimadas com erros associados. Assim, o cálculo de estimativas precisas de HR através de dados de DR é uma tarefa complexa, pois os níveis de precisão aceites como precisos para T_a parecem não ser suficientes.

Na Figura 4.4 encontra-se um exemplo com uma imagem de T_a e uma de HR, ambas correspondentes a Novembro de 2005.

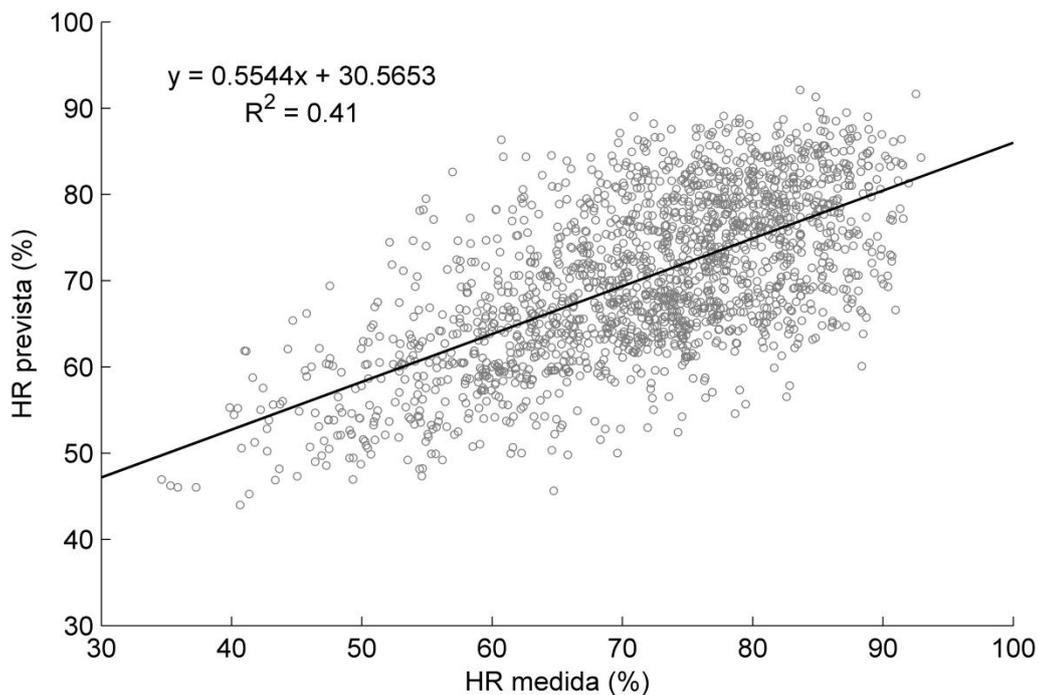


Figura 4.3. Relação entre a HR medida e prevista (MOD07).

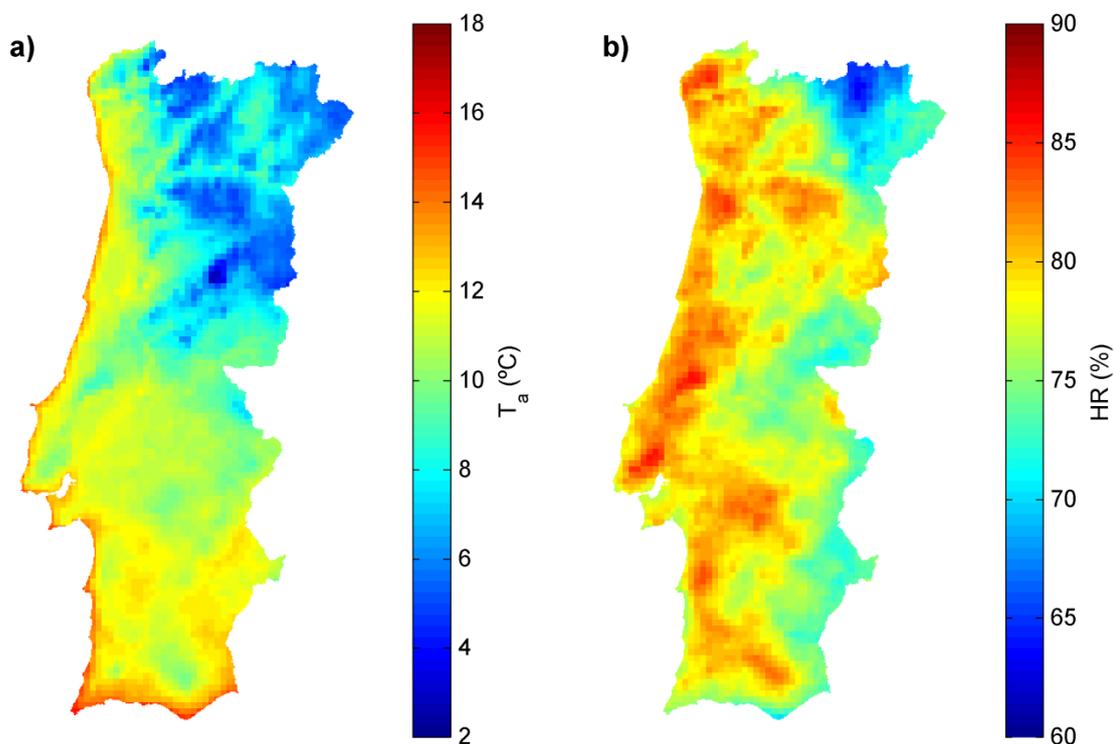


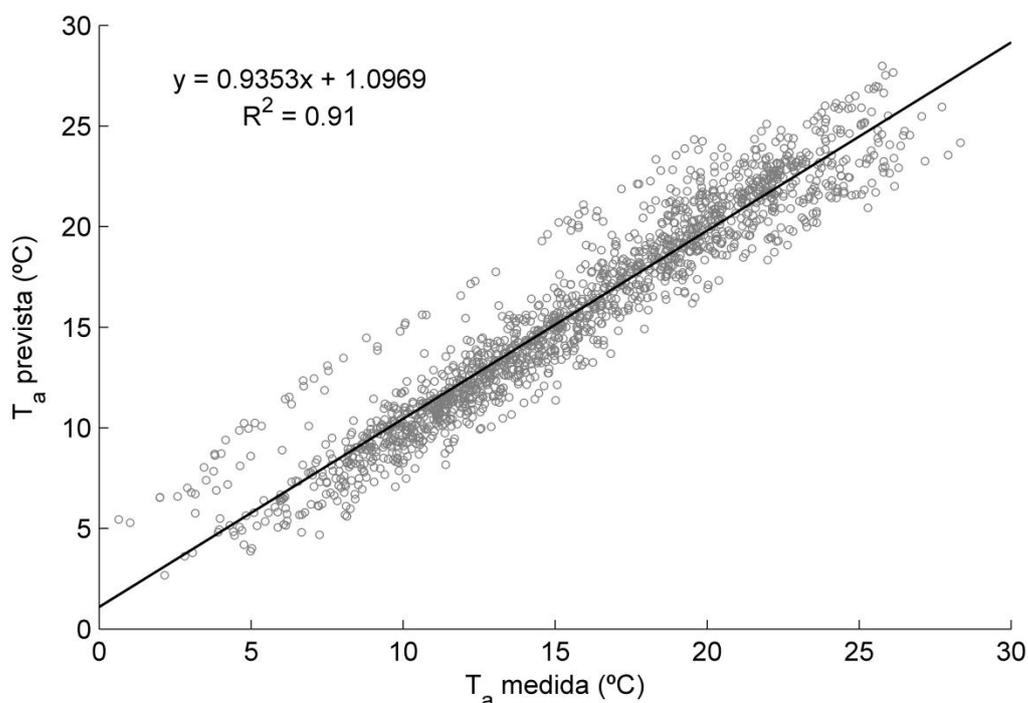
Figura 4.4. Exemplo de duas imagens correspondentes a médias mensais do mês de Novembro de 2005 determinadas com dados MODIS: a) T_a ; b) HR.

Tabela 4.2. Validação das variáveis determinadas com dados MODIS.

Variável	EMQ	EM	DP	R ²	R ² Validação cruzada 10-fold
T _a	1.55 °C	-0.11 °C	1.55 °C	0.92	N/A
T _d	1.56 °C	0.00 °C	1.56 °C	0.86	0.85 (±0.01)
HR	8.76%	-1.45%	8.64%	0.41	N/A

4.1.2 Método de Kriging

O método de Kriging foi utilizado para determinar estimativas espaciais mensais de T_a e HR, para todo Portugal Continental, e dos poluentes PM₁₀ e NO₂, para parte das zonas Norte e Centro do país. Este método de interpolação não extrapola os resultados, por isso, no caso de T_a e HR, foram utilizadas na interpolação estações de Espanha e de Marrocos para gerar mapas que abrangessem o país todo. No caso dos poluentes, a distribuição e quantidade (Figura 2.5) de estações só permitiu gerar estimavas para parte das zonas Norte e Centro. Nas Figuras 4.5 a 4.8 é possível observar as correlações entre os valores medidos e previstos de todas as variáveis (resultantes da validação cruzada) e na Tabela 4.3 a respetiva validação.

**Figura 4.5.** Relação entre a T_a medida e prevista (Kriging).

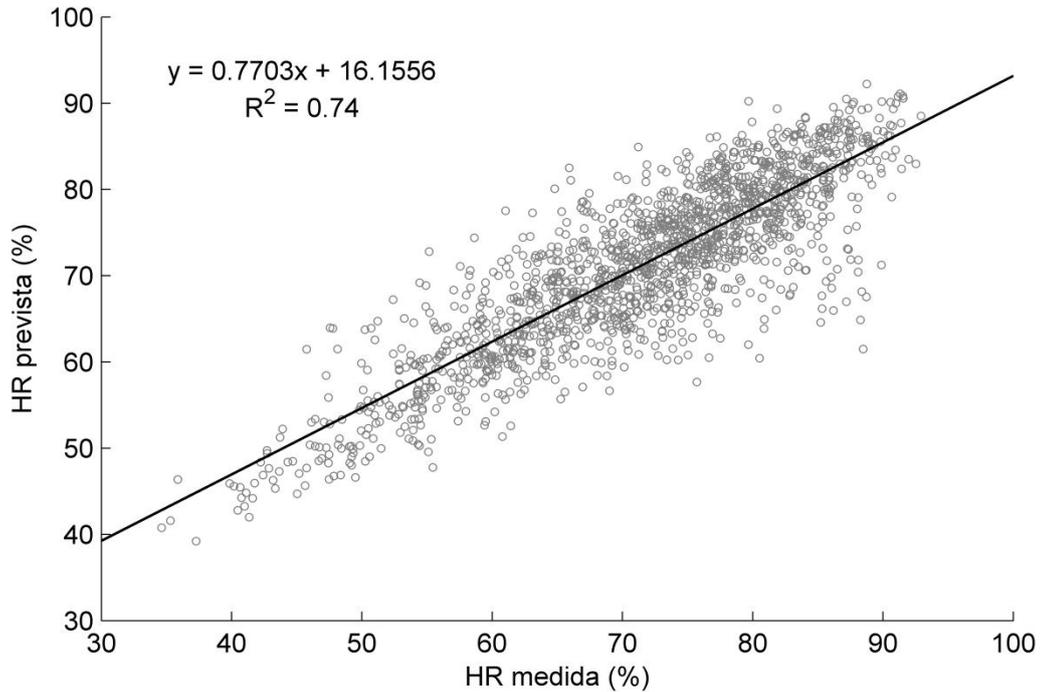


Figura 4.6. Relação entre a HR medida e prevista (Kriging).

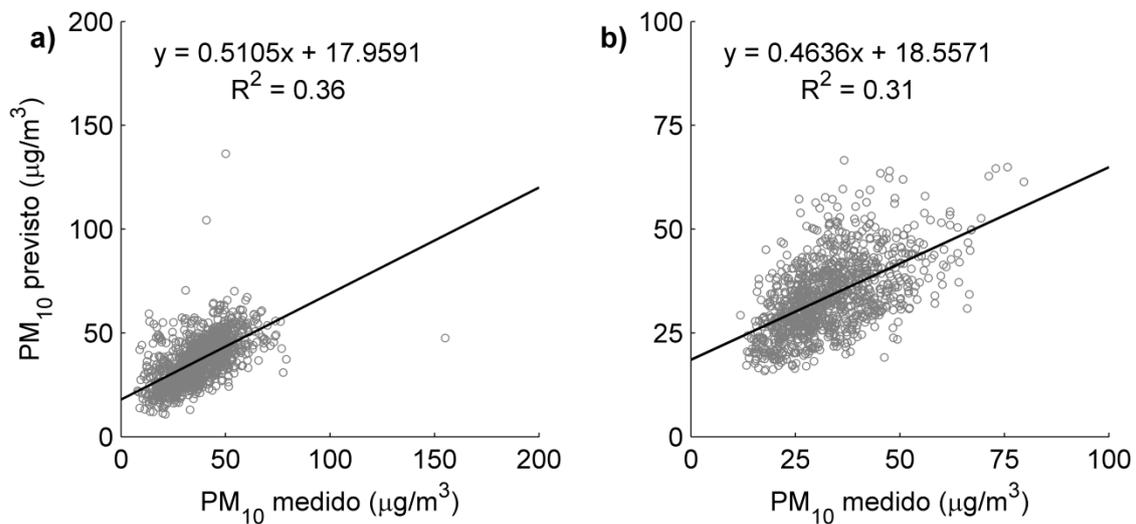


Figura 4.7. Validação do PM_{10} : a) Zona Norte; b) Zona Centro.

A T_a e a HR, determinadas por interpolação, registaram, respetivamente, um R^2 de 0.91 e 0.71 e um EMQ de 1.57 °C e 5.58%. Em relação a T_a , os valores obtidos são semelhantes aos alcançados com os dados MODIS. No caso do HR, verificou-se uma melhoria em relação aos dados de satélite, tanto no R^2 como no EMQ (0.41 e 8.76%, respetivamente). Nos mapas de T_a das Figuras 4.4a (dados MODIS) e 4.9a (método de Kriging), correspondentes a Novembro de 2005, verifica-se que ambos prevêm áreas

semelhantes. Contudo, um dos problemas em utilizar métodos de interpolação consiste no facto de, normalmente, a distribuição de estações não ser ideal, principalmente em zonas montanhosas. Apesar da distribuição das estações meteorológicas (Figura 2.4) ser razoável, o método de Kriging pode não modelar corretamente a variabilidade de T_a entre estações. Por exemplo, no gráfico da relação entre a T_a medida e prevista (Figura 4.5) é possível verificar a existência de um conjunto de pontos sobrestimados, que influenciam a correlação. Estes pontos referem-se à estação com maior altitude (Penhas Douradas, 1388 m), que ao ser estimada na validação cruzada por estações com altitude inferior (e consequentemente com T_a mais elevada), resulta num aumento significativo do erro.

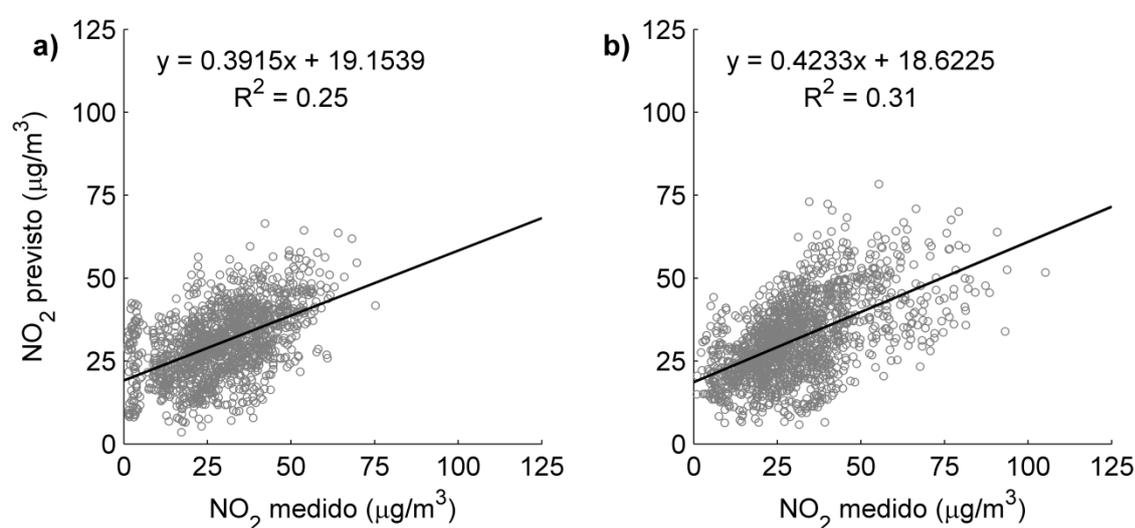


Figura 4.8. Validação do NO₂: a) Zona Norte; b) Zona Centro.

Nos mapas de HR apresentados nas Figuras 4.4b (dados MODIS) e 4.9b (método de Kriging) as diferenças são evidentes. Face à baixa precisão de HR, determinado com os dados MODIS, optou-se por utilizar o HR determinado pelo método de Kriging.

Os R^2 obtidos entre os valores medidos e previstos dos poluentes variam entre 0.25 e 0.36 (Figuras 4.7 e 4.8), revelando baixas correlações. Mais uma vez, surge o problema já referido do método de Kriging poder não ter em consideração a variabilidade dos parâmetros entre pontos. No caso destes poluentes, que apresentam uma grande variabilidade em curtos espaços, esta limitação torna-se ainda mais evidente. Por exemplo, o NO₂ pode variar significativamente dentro de uma distância de 50 m ou menos [98] e, conforme se pode observar na Figura 4.10, a distância entre estações é muito superior.

Como já foi referido anteriormente, a base do método de Kriging consiste na teoria de variáveis regionalizadas, que considera que uma determinada variável varia de um ponto

para outro com continuidade aparente. Como a distribuição e quantidade de estações, em relação à variabilidade dos poluentes, não é a ideal, deixa de haver dependência entre pontos. Se não existe dependência entre pontos, não é possível ajustar corretamente um modelo teórico ao semi-variograma e, conseqüentemente, os resultados da interpolação tendem a não se aproximarem da realidade.

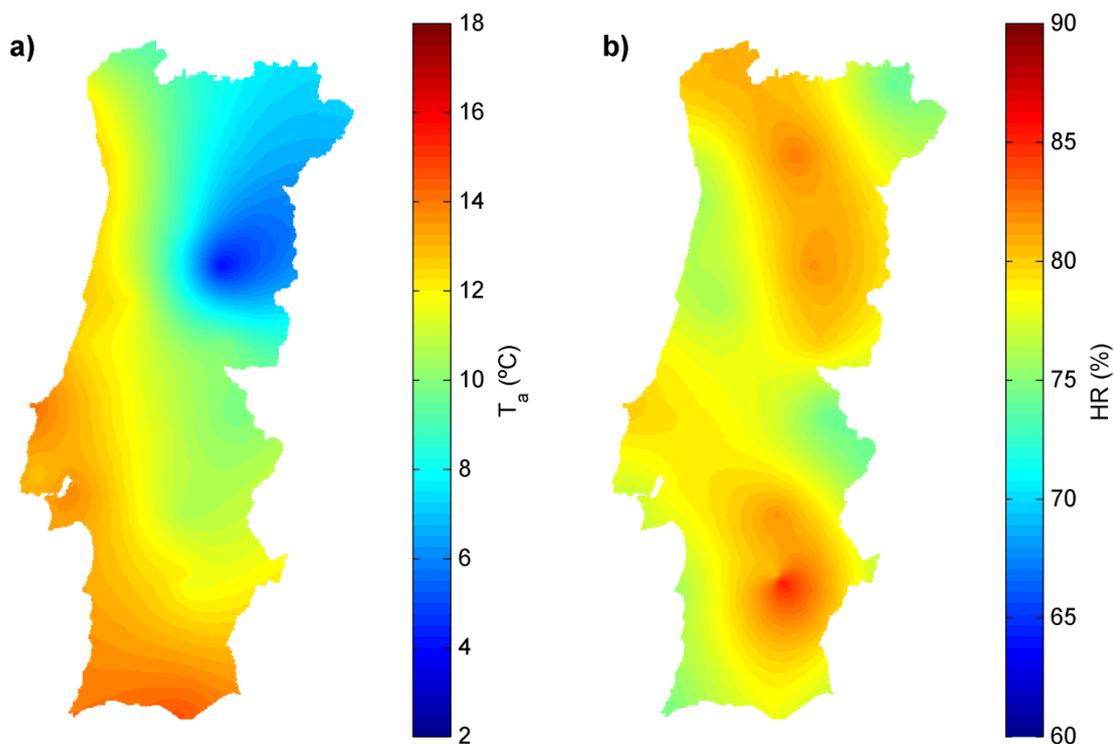


Figura 4.9. Exemplo de duas imagens correspondentes a médias mensais do mês de Novembro de 2005 determinadas pelo método de Kriging: a) T_a ; b) HR.

Tabela 4.3. Validação das variáveis ambientais determinadas com o método de Kriging.

Variável	EMQ	EM	DP	R ² Validação cruzada
T_a	1.57°C	0.09 °C	1.57 °C	0.91
HR	5.58%	-0.31%	5.57%	0.74
PM ₁₀ Norte	10.75 µg/m ³	0.01 µg/m ³	10.76 µg/m ³	0.36
PM ₁₀ Centro	9.38 µg/m ³	0.94 µg/m ³	9.34 µg/m ³	0.31
NO ₂ Norte	12.32 µg/m ³	1.47 µg/m ³	12.23 µg/m ³	0.25
NO ₂ Centro	13.46 µg/m ³	1.02 µg/m ³	13.43 µg/m ³	0.31

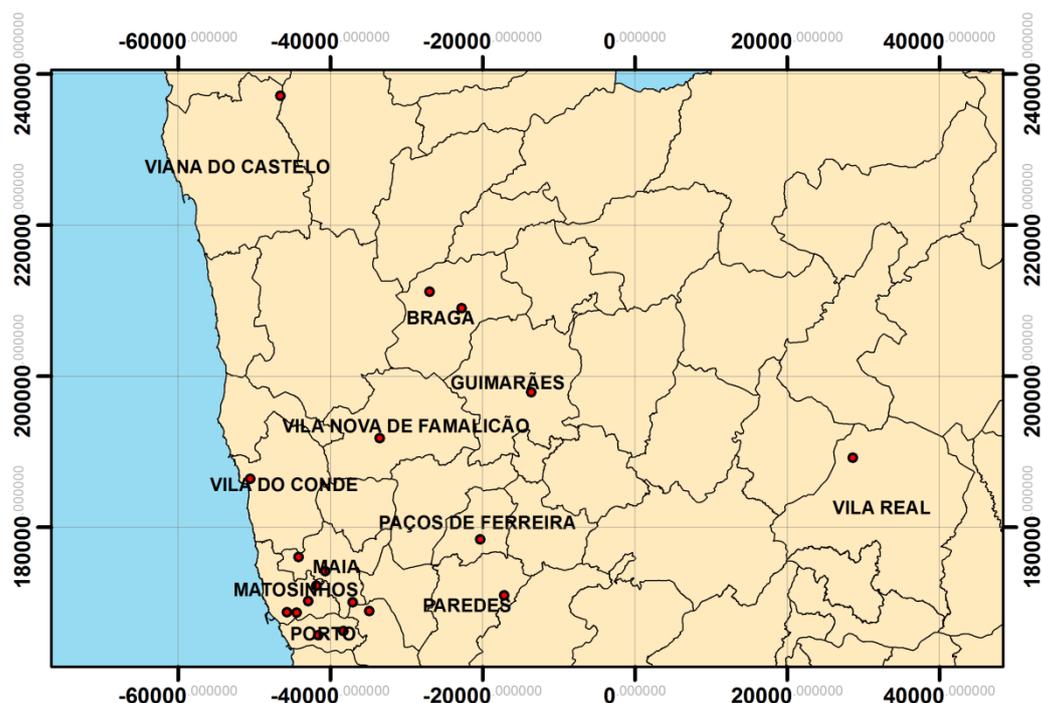


Figura 4.10. Exemplo de um mapa com a distribuição das estações de medição do NO₂ na zona Norte, correspondente a Novembro de 2005. O mapa encontra-se em PTTM06 (metros).

4.1.3 Modelos LUR

As 46 variáveis independentes e as variáveis binárias anuais e mensais foram utilizadas para construir os modelos LUR. Na análise foram utilizadas 3076 (Figura 4.11a) medições médias mensais de PM₁₀ e 3842 (Figura 4.11b) de NO₂. Após realizada a regressão *stepwise*, foram identificadas como variáveis independentes (mais as variáveis binárias) 9 variáveis para PM₁₀ e 11 variáveis para NO₂. Nas Tabelas 4.4 e 4.5 são apresentados os resultados dos dois modelos LUR e respetivos coeficientes do processo de regressão linear múltipla. O modelo LUR de PM₁₀ obteve um R² de 0.57, enquanto que o modelo de NO₂ apresentou um R² de 0.69. Os modelos são compostos por três classes de variáveis, normalmente, utilizadas neste tipo de modelos: tipo de estrada, topografia e ocupação de solo. O volume de tráfego constitui outra classe de variáveis, normalmente, utilizadas neste tipo de modelos mas, devido à escala deste estudo não foi possível incorporar tal informação. A validação dos modelos encontra-se na Tabela 4.6 e as assunções da regressão linear são testadas na Figura 4.11. Os histogramas (Figuras 4.11a e 4.11b) indicam que os resíduos são normalmente distribuídos, apresentando valores médios muito próximos de zero e desvios padrão iguais a 0.71, 0.91,

respetivamente (Tabela 4.7). Para além disso, os resíduos aparentam ser bem comportados, não mostrando nenhuma tendência (Figuras 4.11c e 4.11d).

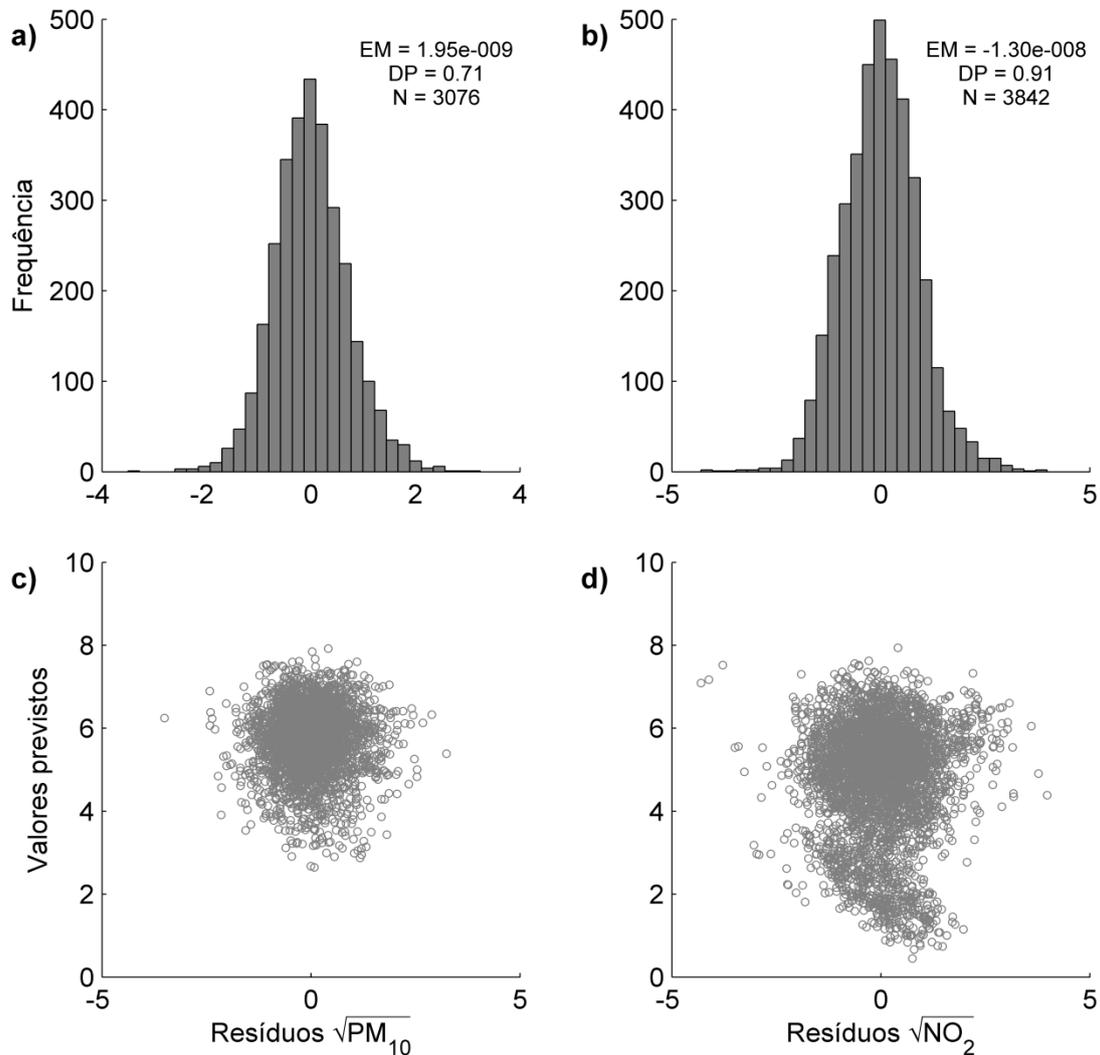


Figura 4.11. Histograma dos resíduos do modelo: a) PM₁₀; b) NO₂. Resíduos do modelo vs. valores previstos: c) PM₁₀; d) NO₂.

Os resultados dos modelos mostram a complexidade da variação espacial dos poluentes considerados, sendo cada modelo composto por várias variáveis independentes. No modelo de PM₁₀, as variáveis mais importantes são a densidade de estradas dentro de um *buffer* com 750 m de raio, ano_2008, Altitude e Latitude (Tabela 4.4). No modelo de NO₂, a ocupação de solo Urb1 dentro de um *buffer* com 750 m de raio revelou-se bastante importante (Tabela 4.5), explicando quase 50% da variabilidade das concentrações de NO₂. É de esperar que em zonas urbanas as concentrações de NO₂ sejam mais elevadas, pois o volume de tráfego e a densidade populacional são mais elevados. Para além do tráfego, o setor industrial é das principais fontes de emissão de

PM₁₀ e NO₂ [79]. Nas Tabela 4.4 e 4.5 demonstra-se que ambos os poluentes se correlacionam positivamente com a classe Urb2 num raio de 250 m, que representa, entre outros, a ocupação de solo do tipo Indústria (Tabela 3.2). Apesar do AOT ser utilizado para estabelecer uma relação com as medidas de PM₁₀ e PM_{2,5}, a análise de regressão considerou que a relação entre o AOT e as concentrações de NO₂ era estatisticamente significativa ($p < 0.001$) para o modelo. Esta relação não é surpreendente, pois o AOT trata-se de uma coluna atmosférica com todos os tipos de aerossóis.

Tabela 4.4. Modelo de regressão para prever a concentração de PM₁₀.

Variável ^a	Coefficiente	R ² Sequencial
(Constante)	-5.806	N/A
Densidade de estradas (<i>buffer</i> 750 m)	0.038	0.161
ano_2008	-0.780	0.249
Altitude	-0.002	0.310
Latitude	0.297	0.383
janeiro	1.589	0.410
HR (Kriging)	-0.030	0.427
fevereiro	0.981	0.440
Distância à EN mais próxima	0.000	0.452
Ocupação de solo Água (<i>buffer</i> 750 m)	0.027	0.467
dezembro	1.141	0.480
AOT (transformação raiz quadrada)	3.072	0.495
novembro	0.869	0.509
março	0.685	0.523
outubro	0.601	0.537
Ocupação de solo Urb2 (<i>buffer</i> 250 m)	0.008	0.546
ano_2007	-0.289	0.553
setembro	0.348	0.559
agosto	0.238	0.562
ano_2006	-0.156	0.564
Ocupação de solo Urb3 (<i>buffer</i> 750 m)	0.015	0.566

^a Valor $p < 0.001$.

Tabela 4.5. Modelo de regressão para prever a concentração de NO₂.

Variável ^a	Coefficiente	R ² Sequencial
(Constante)	-14.956	N/A
Ocupação de solo Urb1 (750 m buffer)	0.026	0.451
Longitude	-0.765	0.487
Latitude	0.356	0.531
junho	-1.731	0.545
julho	-1.778	0.559
Ocupação de solo Urb2 (<i>buffer</i> 250m)	0.020	0.573
Altitude	-0.002	0.589
maio	-1.567	0.602
agosto	-1.676	0.616
abril	-1.245	0.627
Densidade de IC (<i>buffer</i> 750m)	-0.590	0.637
HR (Kriging)	-0.031	0.645
Declive	0.048	0.652
março	-0.833	0.658
setembro	-0.924	0.665
Densidade de IP (<i>buffer</i> 250m)	1.220	0.671
AOT (transformação raiz quadrada)	1.700	0.676
outubro	-0.560	0.681
Ocupação de solo Água (<i>buffer</i> 750m)	0.016	0.684
fevereiro	-0.332	0.686
ano_2007	0.194	0.688

^a Valor $p < 0.001$.

Embora a contribuição de HR e de AOT550 nos modelos seja pequena, estas variáveis introduzem informação sobre a variabilidade espacial e temporal dos poluentes. Utilizar parâmetros variáveis ao longo do tempo e, de preferência, que resultem da observação da totalidade do espaço é de grande importância, pois permitem modelar alterações na concentração dos poluentes em locais onde não existem estações de monitorização.

Na Figura 4.12 é apresentado um exemplo com imagens de PM_{10} e NO_2 , correspondentes a Novembro de 2005. As áreas mais poluídas encontram-se localizadas perto das zonas costeiras no Norte e Centro de Portugal. Como era de esperar, no mapa de NO_2 (Figura 4.12b) verifica-se que as áreas com maior poluição correspondem às cidades do Porto e Lisboa.

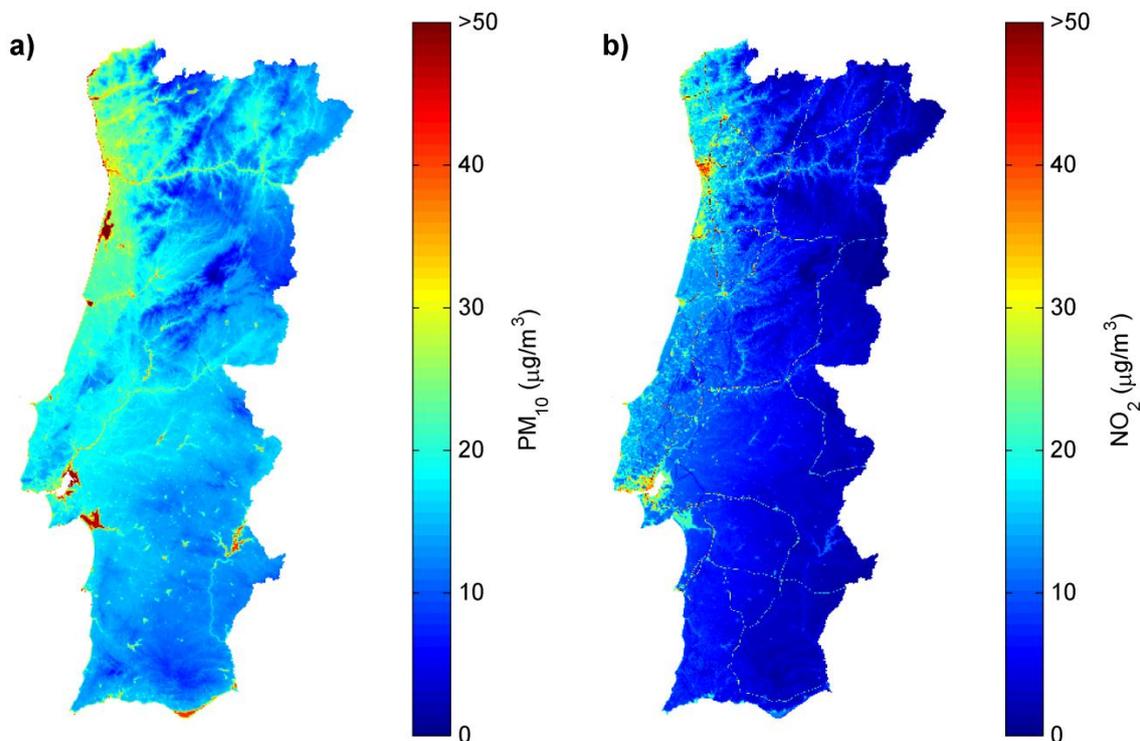


Figura 4.12. Exemplo de duas imagens correspondentes a médias mensais dos modelos LUR, correspondentes a Novembro de 2005: a) PM_{10} ; b) NO_2 .

Tabela 4.6. Validação dos modelos LUR.

Variável	EMQ	EM	DP	R^2	R^2 Validação cruzada 10-fold
PM_{10}	0.71	$1.95e^{-9}$	0.71	0.57	0.56 (± 0.02)
NO_2	0.91	$-1.30e^{-8}$	0.91	0.69	0.68 (± 0.02)

4.2 Análises de Regressão

Portugal Continental é constituído por 18 distritos, por isso, e como já foi referido, foram realizadas 18 análises independentes de regressão. As regressões foram realizadas para o intervalo de tempo de 2003 a 2008, considerando a taxa de internamentos devido a asma por 1000 habitantes por concelho (variável dependente) e as respetivas médias das

variáveis ambientais T_a , HR, NDVI, NO_2 e PM_{10} (variáveis independentes). Na Tabela 4.7 é apresentada uma descrição das variáveis ambientais por distrito (valor médio) e informação sobre a população total residente, percentagem de ocupação de solo urbano (classe 1 do nível 1 do CLC06_PT – Tabela 2.3) e taxa de internamentos devido a asma por 1000 habitantes. De referir, que os valores a sombreado dizem respeito a correlações significativas (valor $p < 0.050$) encontradas na análise univariada (Tabela 4.8).

Tabela 4.7. Descrição das variáveis ambientais por distrito. Os valores a sombreado representam correlações significativas (valor $p < 0.050$) encontradas na análise univariada.

Distrito	Pop. Total	Urb (%)	Asma1000	Valor médio				
				T_a	HR	NDVI	PM_{10}	NO_2
Aveiro	730,389	10.69	1.80	14.99	75.46	0.61	19.27	10.70
Beja	154,516	0.63	1.50	16.14	71.85	0.45	11.47	2.34
Braga	856,477	8.03	0.84	13.72	74.92	0.63	17.09	9.65
Bragança	143,679	0.50	2.55	12.23	67.15	0.49	13.84	2.45
Castelo Branco	201,049	0.84	3.88	13.99	65.96	0.52	13.68	3.10
Coimbra	435,916	3.24	2.50	14.45	73.35	0.62	15.63	7.28
Évora	170,358	1.03	0.85	15.62	66.94	0.49	13.67	3.09
Faro	418,616	3.95	2.70	16.75	72.52	0.48	11.43	2.87
Guarda	174,097	0.88	1.08	11.30	68.33	0.50	10.92	1.75
Leiria	475,891	6.52	2.71	14.96	74.34	0.59	15.31	8.17
Lisboa	2,217,440	18.57	2.22	15.77	72.34	0.49	16.70	11.80
Portalegre	119,857	0.63	1.44	15.33	62.87	0.50	14.91	3.31
Porto	1,812,797	17.94	1.29	14.51	74.68	0.58	19.28	12.34
Santarém	464,578	2.99	2.49	15.71	68.39	0.53	16.39	7.06
Setúbal	840,899	5.72	1.37	16.23	72.63	0.48	15.49	6.92
Viana do Castelo	251,645	4.61	1.13	13.45	75.89	0.65	17.40	10.98
Vila Real	219,127	1.55	2.80	11.23	70.58	0.55	12.47	3.84
Viseu	394,362	2.05	1.19	12.18	71.92	0.59	12.59	4.50

Pop. Total – População total;

Urb – Ocupação de solo urbano;

Asma1000 – Taxa de internamentos devido a asma por 1000 habitantes.

Tabela 4.8. Análise univariada entre as variáveis ambientais (variáveis independentes) e a taxa de internamentos devido a asma por 1000 habitantes (variável dependente) nos concelhos de cada distrito.

DICOFRE	Nome do distrito	Variáveis	R ²	Coefficiente	Valor p
010000	Aveiro	T _a	0.221	0.280	0.042
		T _a	0.543	0.145	0.003
030000	Braga	HR	0.431	0.113	0.011
		PM ₁₀	0.391	0.053	0.017
		NO ₂	0.410	0.052	0.014
		T _a	0.614	-1.714	0.004
050000	Castelo Branco	T _a	0.614	-1.714	0.004
		NDVI	0.286	-8.656	0.033
080000	Faro	PM ₁₀	0.290	0.175	0.031
		NDVI	0.481	-6.484	0.003
110000	Lisboa	PM ₁₀	0.272	0.125	0.038
		NO ₂	0.656	0.104	<0.001
		NO ₂	0.332	0.304	0.024
120000	Portalegre	NO ₂	0.332	0.304	0.024
		T _a	0.286	0.272	0.022
130000	Porto	NDVI	0.541	-5.935	<0.001
		PM ₁₀	0.442	0.102	0.003
		NO ₂	0.583	0.073	<0.001
		T _a	0.415	0.624	0.018
150000	Setúbal	NDVI	0.441	-4.347	0.013
		PM ₁₀	0.353	0.049	0.032
		NO ₂	0.542	0.054	0.004
		NO ₂	0.542	0.054	0.004

Na análise univariada encontram-se relações entre a taxa de internamentos devido a asma por 1000 habitantes e as variáveis ambientais, apenas para 8 distritos (Aveiro, Braga, Castelo Branco, Faro, Lisboa, Portalegre, Porto e Setúbal – Tabela 4.8). Em geral, estes distritos são dos mais poluídos e dos mais populosos. Nos distritos de Aveiro, Braga, Porto e Setúbal foram encontradas relações positivas entre os internamentos e a T_a com um R² a variar entre 0.221 e 0.543 (Tabela 4.8). A correlação mais elevada verificou-se no distrito de Castelo Branco (R² = 0.614), mas a relação obtida entre os

internamentos e a T_a foi negativa. No distrito de Braga foi também encontrada uma relação positiva com a HR ($R^2 = 0.431$).

Entre os internamentos e o NDVI foram encontradas relações negativas, em Faro, Lisboa, Porto e Setúbal, com um R^2 a variar entre 0.286 e 0.541 (Tabela 4.8). Para estes distritos, foi realizada uma comparação entre o R^2 , atingido com o NDVI e os internamentos, e o NO_2 médio (Figura 4.13). Verifica-se que a correlação tende a ser mais elevada em distritos mais poluídos por NO_2 .

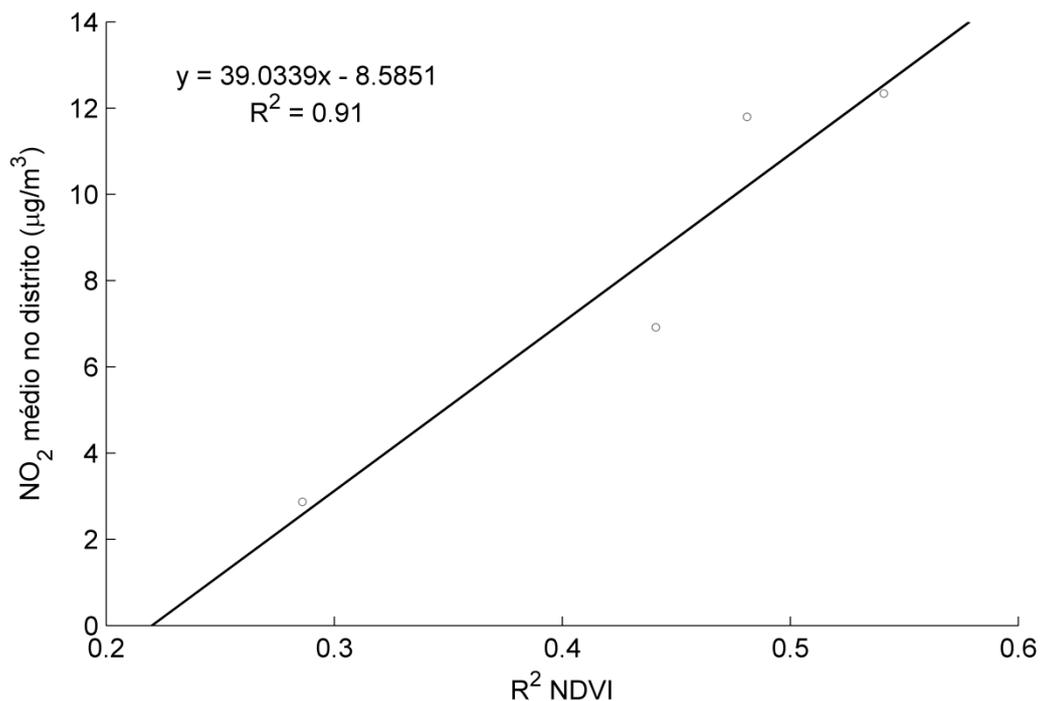


Figura 4.13. Relação entre o R^2 , atingido com o NDVI e a taxa de internamentos devido a asma por 1000 habitantes, e respetivo NO_2 médio no distrito.

Para Braga, Lisboa, Porto e Setúbal, os resultados (Tabela 4.8) mostram relações positivas entre os internamentos e os poluentes NO_2 (R^2 entre 0.410 e 0.656) e PM_{10} (R^2 entre 0.272 e 0.442). Estes são os quatro distritos mais populosos (Tabela 4.7). Em Faro e Portalegre foram encontradas relações positivas com apenas um dos poluentes (NO_2 e PM_{10} , respetivamente). Nas Figuras 4.14 e 4.15 pode-se observar que quanto maior for o nível médio de NO_2 e PM_{10} , respetivamente, maior é a correlação encontrada com os internamentos.

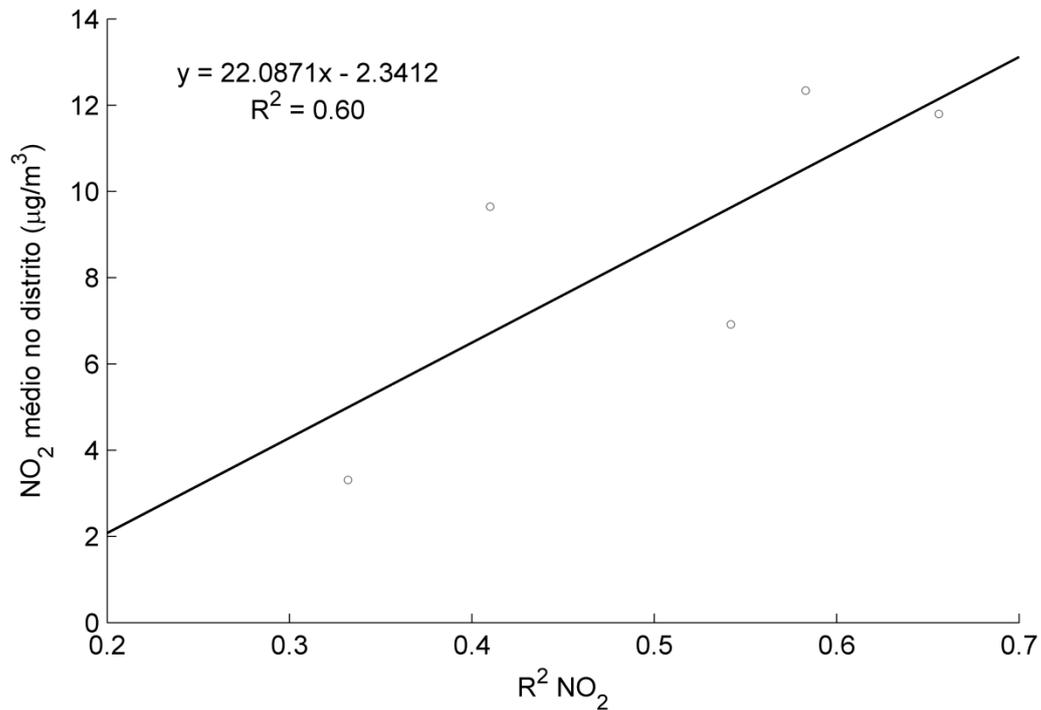


Figura 4.14. Relação entre o R^2 , atingido com o NO₂ e a taxa de internamentos devido a asma por 1000 habitantes, e respectivo NO₂ médio no distrito.

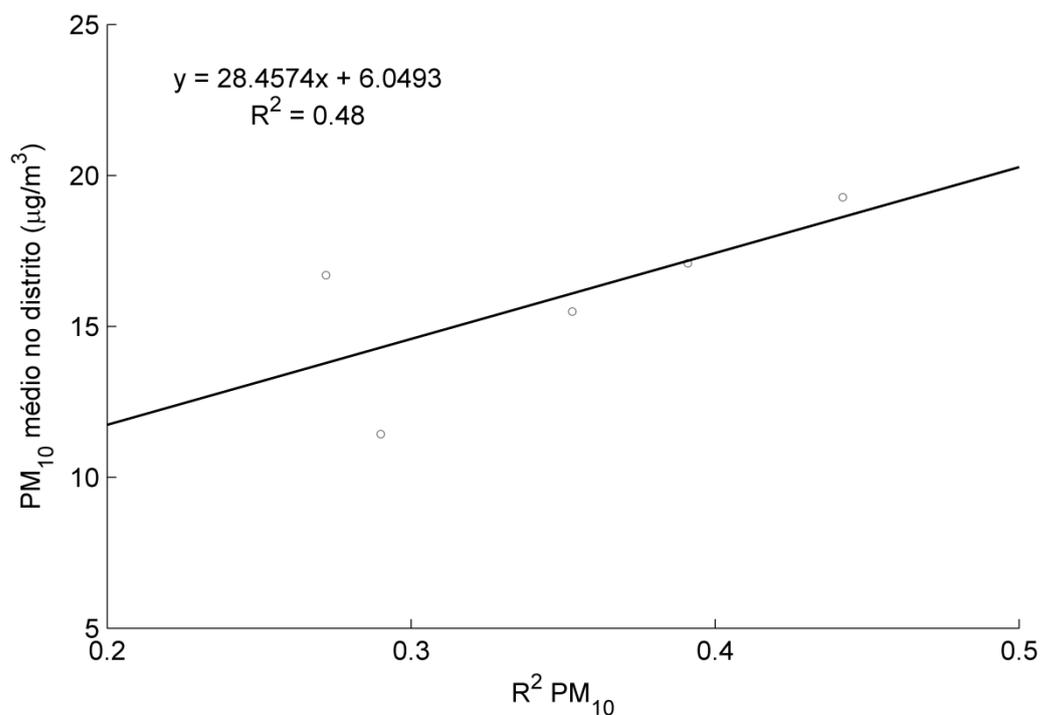


Figura 4.15. Relação entre o R^2 , atingido com o PM₁₀ e a taxa de internamentos devido a asma por 1000 habitantes, e respectivo PM₁₀ médio no distrito.

Na análise multivariada (Tabela 4.9), os distritos de Braga e Porto foram os únicos a registar relações significativas. No primeiro distrito, estabeleceu-se um modelo com elevada correlação ($R^2 = 0.815$) entre as variáveis T_a e PM_{10} e os internamentos. Esta conjugação fez com que o PM_{10} se relacionasse negativamente com os internamentos. No segundo distrito, foi estabelecido um modelo com as variáveis NO_2 e HR, sendo que a HR tem um efeito negativo.

Tabela 4.9. Análise multivariada entre as variáveis ambientais (variáveis independentes) e a taxa de internamentos devido a asma por 1000 habitantes (variável dependente) nos concelhos de cada distrito.

DICOFRE	Nome do distrito	Variáveis	R^2	Coefficiente	Valor p
030000	Braga	T_a	0.815	0.680	<0.001
		PM_{10}		-0.233	0.002
130000	Porto	NO_2	0.688	0.093	<0.001
		HR		-0.136	0.040

Capítulo 5

Discussão

5.1 Variáveis Ambientais

Inicialmente, pretendia-se utilizar maioritariamente dados de DR no estudo da asma em Portugal Continental. No entanto, com o decorrer do trabalho, verificou-se que não seria possível. Apenas duas variáveis determinadas por satélite foram utilizadas na análise da asma: T_a e NDVI. A T_a foi determinada pela interpolação do PT_a do MOD07, utilizando o MDE5 do SRTM. Através de um processo simples de seleção da combinação de dados com menor EMQ em cada mês, a T_a foi determinada com um EMQ de 1.55 °C, que se encontra dentro do limite 1-2 °C, em geral, aceite como preciso na determinação de T_a a partir de dados de DR [85]. O NDVI não sofreu qualquer processamento adicional, sendo apenas corrigido com os indicadores de qualidade.

Devido à baixa precisão da HR determinada com os dados MODIS (EMQ = 8.76% $R^2 = 0.41$), utilizou-se o método de Kriging para gerar estimativas espaciais da HR mais precisas (EMQ = 5.58% e $R^2 = 0.74$).

A utilização de dados de DR tem como principal vantagem possibilitar a observação sinóptica de zonas onde a cobertura por estações de monitorização é escassa. Contudo, existem várias limitações, que por vezes impossibilitam a sua utilização. Como já foi referido, a maior parte dos estudos epidemiológicos que recorrem a dados de DR dizem respeito a doenças infecciosas. Normalmente, estes estudos utilizam parâmetros de fácil determinação, como o LST ou o NDVI. Porém, doenças como a asma necessitam de parâmetros mais complexos, nomeadamente a concentração de poluentes ao nível da superfície.

Existe uma vasta quantidade de dados de satélite sobre a concentração de poluentes na atmosfera. Contudo, este tipo de dados apenas se revela útil em estudos globais, devido à sua baixa resolução e ao facto da grande maioria não se referir unicamente à troposfera. Numa fase inicial, testou-se a coluna troposférica de NO_2 do sensor OMI do satélite Aura para estimar os níveis de NO_2 à superfície, mas logo se verificou que não era possível estabelecer qualquer relação. Um pixel deste parâmetro (resolução espacial de 24 x 13 km) engloba várias estações com níveis distintos de NO_2 . Assim, foi necessário procurar alternativas para modelar a concentração dos poluentes de interesse para todo Portugal Continental.

Começou-se por utilizar o método de Kriging para interpolar as medições das estações de qualidade do ar. Contudo, surgiu de imediato um problema: a distribuição das estações, concentradas sobretudo no Porto e Lisboa, não permitia gerar mapas para o país todo. Como tal, considerou-se somente parte das zonas Norte e Centro. Conforme documentado na secção 4.1.2, a precisão das estimativas era baixa (R^2 entre 0.25 e 0.36), devido à distribuição e quantidade de estações, em relação à variabilidade dos poluentes, não ser a ideal, não cumprindo com o princípio básico do Kriging (dependência entre pontos). Deste modo, resolveu-se modelar a concentração dos poluentes de interesse para todo Portugal Continental com modelos LUR, que consideram independência entre pontos.

Os modelos LUR registaram precisões mais elevadas ($R^2 = 0.57$, para o modelo de PM_{10} , e $R^2 = 0.69$, para o modelo de NO_2). Em ambos os modelos, o AOT500 foi considerado como variável independente para introduzir informação sobre a variabilidade espacial e temporal dos poluentes, sobretudo no modelo de PM_{10} . Porém, a sua contribuição foi reduzida, acrescentando menos de 1% à variabilidade dos poluentes já explicada por outras variáveis. A principal referência na construção do modelo de PM_{10} foi o estudo de Mao *et al.* [102], no qual é elaborado um modelo LUR-AOT para mapear a concentração mensal de $\text{PM}_{2.5}$, durante o ano de 2005 numa escala regional (Florida). Apesar de não ser possível realizar uma comparação direta (em Mao *et al.* o poluente mapeado foi $\text{PM}_{2.5}$), o modelo de PM_{10} obtido neste trabalho atingiu um R^2 próximo do obtido no estudo referido (0.62). Os modelos LUR são utilizados maioritariamente para modelar a distribuição espacial de poluentes à escala local. Apesar disso, o modelo LUR de NO_2 atingiu um R^2 dentro dos limites, normalmente, obtidos (entre 0.54 e 0.81 [98]).

A determinação de estimativas da concentração de poluentes ao nível da superfície com dados de DR tem incidido na relação entre PM_{10} ou $\text{PM}_{2.5}$ e o AOT, mas trata-se de um processo complicado. A relação encontrada entre o AOT do MODIS e as medições de PM_{10} realizadas pelas estações de qualidade do ar é fraca ($R^2 = 0.06$). Não se trata de

um resultado surpreendente, uma vez que o PM_{10} possui uma grande variabilidade dentro de um pixel de 10 x 10 km e outras limitações já referidas (p.e.: a associação entre o AOT e o PM_{10} varia com as condições meteorológicas; o AOT refere-se a uma coluna atmosférica, enquanto PM_{10} são medidas ao nível da superfície). Assim, desenvolver um método que permita aumentar a relação entre o AOT e o PM_{10} pode melhorar a precisão do modelo LUR de PM_{10} . O aumento dessa relação deve passar sobretudo pela determinação de um AOT com maior resolução espacial.

Os dados de DR apresentam grandes potencialidades, mas possuem limitações que restringem a sua utilização em estudos epidemiológicos de doenças que exigem parâmetros de pouca complexidade.

5.2 Análise Epidemiológica da Asma

Neste trabalho, realizou-se a análise de cada distrito separadamente de forma a evitar efeitos de confusão induzidos pela variação demográfica e pela prevalência e estado do controlo da asma. Um distrito com alta prevalência de asma pode ser muito poluído, mas a doença pode estar bem controlada. As pessoas asmáticas cuja doença esteja controlada são menos suscetíveis aos fatores ambientais desfavoráveis [28]. Por exemplo, no caso do Porto e Lisboa a população residente é muito mais elevada que em distritos como Castelo Branco ou Vila Real, mas apresentam uma taxa de internamentos devido a asma por 1000 habitantes inferior. Os resultados mostram que, apesar de poder existir tanta ou mais poluição, se os doentes forem devidamente acompanhados, poderão existir benefícios em termos de saúde (qualidade de vida, consumo de recursos, etc.)

Este estudo foi realizado, em parte, porque existem poucos estudos em Portugal que avaliam a associação entre a asma e fatores ambientais. Para além disso, a maioria dos estudos realizados não utiliza informação espacial [34]. Os estudos tradicionais não espaciais utilizam dados recolhidos por uma única estação de monitorização ou valores médios calculados a partir de várias estações, podendo assim conduzir a erros de exposição. A modelação espacial das variáveis ambientais permite analisar diferentes áreas de Portugal Continental e aumentar a correspondência com a realidade.

Os resultados obtidos na análise univariada sugerem fortes relações entre a T_a , NDVI, NO_2 e PM_{10} e os internamentos devido a asma. Nos distritos de Aveiro, Braga, Porto e Setúbal verificou-se que T_a se correlacionava positivamente com a taxa de internamentos por 1000 habitantes. Contudo, em Castelo Branco observou-se uma elevada correlação negativa (0.614). Estudos demonstram que tanto uma T_a elevada como uma T_a reduzida podem ter efeitos em pessoas asmáticas [36, 54]. Por um lado, as temperaturas elevadas

podem estar associadas com a expressão da asma, pois ocorrem quando existe maior atividade solar, que por sua vez está relacionada com a formação de O_3 ao nível da superfície [36, 41, 107]. Ao contrário do O_3 da estratosfera, que se forma naturalmente na atmosfera e protege-nos dos raios ultravioleta provenientes do Sol, o O_3 ao nível da superfície é criado a partir da interação das emissões naturais e criadas pelo homem de Compostos Orgânicos Voláteis (COVs) e Óxidos de Azoto (NO_x) na presença de calor e luz solar [79, 107]. Uma vez que os automóveis são das maiores fontes de COVs e os NO_x são produzidos pela queima de combustíveis fósseis (veículos e centrais elétricas), não é surpreendente encontrar estas relações em distritos com uma percentagem elevada de tecido urbano. Por outro lado, o ar frio e seco (Castelo Branco é dos distritos com menor HR médio) pode agir como um irritante das vias aéreas [108, 109], nomeadamente na asma induzida por exercício [108].

Nos distritos do Porto, Lisboa, Setúbal e Faro encontraram-se correlações negativas entre o NDVI e os internamentos. Estas relações certamente estão relacionadas com a poluição, uma vez que em espaços verdes esta é mais reduzida. A correlação entre a taxa de internamentos por 1000 habitantes e o NDVI tende a ser maior em zonas com níveis de NO_2 mais elevados, o que revela a importância da vegetação em zonas urbanas. Seria interessante testar a variação do NDVI durante a Primavera em distritos onde os níveis de poluição são reduzidos, para assim verificar a existência de alguma correlação positiva que sugira o efeito do aumento das concentrações de pólen no ar durante essa estação. Contudo, o NDVI é um índice correspondente a todo o tipo de vegetação e nem todos os pólenes são prejudiciais para pessoas asmáticas.

O NO_2 foi a variável que registou associações (positivas) mais elevadas, sendo superiores a 0.50 em três distritos (Porto, Lisboa e Setúbal). Como já foi referido, o NO_2 é um indicador de tráfego automóvel, e como tal encontra-se em níveis mais elevados sobretudo em cidades com elevada população. Este facto é confirmado pelos distritos onde foram encontradas correlações. À exceção de Portalegre, cuja correlação é baixa (0.332), os restantes quatro distritos são os mais populosos (Porto, Lisboa, Setúbal e Braga). Para além disso, os resultados sugerem que quanto mais elevadas forem as concentrações de NO_2 no distrito, maior é a influência nos internamentos por asma.

Existem vários mecanismos biológicos pelos quais o NO_2 pode afetar a saúde respiratória [41, 110]. Este poluente torna as pessoas asmáticas mais suscetíveis a infeções virais respiratórias, que provocam a expressão da asma [110, 111]. A exposição a níveis elevados de NO_2 pode também causar sintomas crónicos de bronquite [112] e reduzir a função pulmonar de crianças asmáticas [43].

O segundo poluente em estudo, o PM_{10} , também apresentou relações positivas com os internamentos (R^2 entre 0.272 e 0.442). Apesar de não serem tão elevadas quanto as

verificadas com o NO₂, a exposição a este poluente aumenta o risco de ataques de asma. As PM₁₀ são constituídas por material sólido ou pequenas gotículas de fumo, poeiras e vapor condensado no ar e que têm origem no tráfego, indústria, obras de construção civil e em processos agrícolas [41, 79]. À semelhança do NO₂, as correlações encontradas dizem respeito a distritos com elevada população e correlacionam-se positivamente com as concentrações médias nos distritos. Conforme referido anteriormente, as PM de menor dimensão, nomeadamente as de diâmetro aerodinâmico igual ou menor que 2.5 µm (PM_{2.5}) são as mais prejudiciais, pois conseguem penetrar mais facilmente no sistema respiratório. Contudo, não foi possível utilizar este poluente, uma vez que grande parte das estações não as mede.

À semelhança deste trabalho, vários estudos reportam associações entre o aumento de internamentos hospitalares por asma e a exposição a concentrações elevadas de NO₂ [34, 39, 41, 47] e de PM₁₀ [39, 41, 44-46].

Em relação à HR, apenas no distrito de Braga se verificou uma correlação significativa ($R^2 = 0.431$). A HR elevada tem sido associada com a expressão da asma [54, 55, 113]. Através da realização de experiências controladas numa câmara de emissão, Taylor *et al.* [113] demonstraram que os pólenes são libertados em condições de elevada HR. O distrito de Braga aparenta reunir as condições necessárias para este facto, uma vez que é um distrito com bastante vegetação (NDVI médio = 0.63) e possui uma HR média das mais elevadas (74.92%).

Na análise multivariada foram encontradas correlações nos distritos de Braga e do Porto. No primeiro distrito, estabeleceu-se um modelo multivariado com T_a e PM₁₀, apresentando um R^2 elevado de 0.815. Estes resultados revelam-se algo surpreendentes, tendo em conta que quando o PM₁₀ é introduzido no modelo aparenta ter um efeito negativo nos internamentos. Mais estudos devem ser desenvolvidos com o intuito de compreender esta relação. No segundo, alcançou-se um modelo multivariado com um R^2 de 0.688, considerando as variáveis NO₂ + HR. A introdução da HR permite aumentar em cerca de 15% a variabilidade explicada pelo modelo univariado de NO₂. Os resultados sugerem que a interação entre níveis elevados de NO₂ e HR baixa aumenta a taxa de internamentos devido a asma. Alguns estudos referem que a baixa HR pode aumentar os sintomas de asma [55, 114]. De facto, uma das recomendações do GINA [28] para pacientes com asma de difícil controlo consiste em evitar atividade física extenuante sobre condições de T_a e poluição do ar elevadas e baixa HR. Delamanter *et al.* [34] encontrou uma relação semelhante num estudo efetuado na Califórnia, exceto no facto do efeito da HR ser positivo.

As interações entre diferentes variáveis na expressão da asma é um processo complexo, que deve ser estudado mais detalhadamente.

5.3 Limitações

A análise realizada apresenta algumas limitações. A resolução espacial dos internamentos devido a asma leva a que sejam feitas algumas suposições que podem influenciar os resultados. Ao considerar a média de uma determinada variável num concelho como indicador da presença da doença assume-se que os indivíduos que residem num determinado concelho estão expostos aos mesmos níveis ambientais. Por exemplo, um indivíduo que resida perto de uma estrada com elevado volume de tráfego pode estar exposto a níveis de NO₂ mais elevados do que um indivíduo que resida num local do mesmo concelho com uma densidade de vegetação elevada. Para além disso, a utilização de níveis ambientais pode resultar em erros de classificação de exposição, pois não tem em conta a presença de outros fatores. Um indivíduo pode não trabalhar no concelho de residência, estando exposto a diferentes níveis ambientais durante o período em que está fora de casa. Os internamentos devido a asma têm também a limitação de medirem apenas uma porção da presença da doença, uma vez que nem todos os indivíduos que sofram graves sintomas asmáticos necessitam de ser internados. Existem também possíveis limitações associadas ao facto de não se possuir informação acerca de casos em hospitais privados (p.e., no Porto e Lisboa existem vários). Outros fatores, como a poluição dentro de espaços fechados, total de tempo ao ar livre, tabagismo passivo ou comportamento de evitação, não são considerados.

A análise de regressão efetuada apenas é sensível a relações lineares entre os internamentos e as variáveis ambientais. Apesar de se ter encontrado fortes correlações, podem existir relações não lineares importantes que não são modeladas. Neste contexto, poder-se-ia utilizar Modelos Aditivos Generalizados (*Generalized Additive Models – GAM*) para lidar com relações não lineares. Seria também interessante repetir a análise para verificar os efeitos das variáveis ambientais em diferentes grupos etários e sexo.

Para além das limitações referidas em cima, existem também limitações associadas às variáveis ambientais, que podem introduzir erros nas associações alcançadas. As principais limitações encontram-se nos modelos LUR. Normalmente, estes modelos são utilizados para modelar a distribuição espacial de poluentes à escala local. Se a distribuição e quantidade de estações for fraca, o modelo vai estar ajustado a algumas zonas em específico, podendo cometer erros elevados na modelação de zonas não cobertas por estações. Inicialmente, estabeleceu-se um modelo LUR para o poluente CO, que tem origem no tráfego [79]. Este modelo apresentava um R² elevado (cerca de 0.68) e a principal variável era a Latitude (relação positiva com as medições das estações). Como para o intervalo de tempo em estudo não haviam estações localizadas a mais de

41.5° de Latitude, as zonas mais a Norte apresentavam concentrações de CO muito mais elevadas do que nas grandes cidades.

Capítulo 6

Conclusões e Perspetivas Futuras

Neste estudo, exploraram-se as relações entre internamentos hospitalares devido a asma e algumas variáveis ambientais em Portugal Continental. As principais relações foram encontradas sobretudo em distritos com elevada população. Na análise univariada verificaram-se associações elevadas com as cinco variáveis, nomeadamente com o poluente NO_2 . Na análise multivariada, observou-se que em dois distritos os internamentos estavam associados, respetivamente, com as combinações $T_a + \text{PM}_{10}$ e $\text{NO}_2 + \text{HR}$.

O estudo da asma aqui realizado deve ser aprofundado, utilizando métodos de análise mais complexos e dados mais recentes (posteriores a 2008) e considerando outras variáveis, como o O_3 , o nevoeiro ou concentrações de pólen. Compreender os fatores que afetam as várias zonas do país pode servir de auxílio ao controlo da asma.

Relativamente aos dados de satélite, apenas foi possível utilizar a T_a e o NDVI. Estes dados possuem grandes potencialidades, pois tratam-se de observações de todo o espaço, mas apresentam limitações que restringem a sua aplicação. No estudo epidemiológico de doenças respiratórias a exposição a concentrações de poluentes é uma das variáveis mais importantes. Atualmente, a baixa resolução dos dados de DR, referentes à concentração de poluentes, não permite a sua correta modelação ao nível da superfície em escalas locais. Enquanto isso não for possível, devem-se procurar outras alternativas. Assim, a utilização de modelos de dispersão de ar deve ser explorada, uma vez que, apesar de requererem muita informação, podem determinar a distribuição espacial de vários poluentes e são independentes de estações de monitorização.

Referências

- [1] Hennekens, C.H. and Buring, J.E. (1987) *Epidemiology in Medicine*. Little, Brown and Company, Boston / Toronto.
- [2] Porta, M. (2008) *A Dictionary of Epidemiology*. 5th ed Oxford University Press, New York.
- [3] Webb, P. and Bain, C. (2010) *Essential Epidemiology: An Introduction for Students and Health Professionals*. 2nd ed Cambridge University Press, Cambridge, UK.
- [4] Saracci, R. (2010) *Epidemiology: A Very Short Introduction*. 1st ed Oxford University Press, New York.
- [5] Meliker, J.R. and Sloan, C.D. (2011) Spatio-temporal epidemiology: Principles and opportunities. *Spatial and Spatio-temporal Epidemiology*, 2(1), 1-9.
- [6] Carpenter, T.E. (2011) The spatial epidemiologic (r)evolution: A look back in time and forward to the future. *Spatial and Spatio-temporal Epidemiology*, 2(3), 119-124.
- [7] Pfeiffer, D.U., Robinson, T.P., Stevenson, M., Stevens, K.B., Rogers, D.J., and Clements, A.C.A. (2008) *Spatial Analysis in Epidemiology*. 1st ed Oxford University Press, New York.
- [8] Aronoff, S. (1989) *Geographic Information Systems: A Management Perspective*. WDL Publications, Ottawa, Canada.
- [9] Rezaeian, M., Dunn, G., St Leger, S., and Appleby, L. (2007) Geographical epidemiology, spatial analysis and geographical information systems: a multidisciplinary glossary. *Journal of Epidemiology & Community Health*, 61(2), 98-102.
- [10] Graham, A.J., Atkinson, P.M., and Danson, F.M. (2004) Spatial analysis for epidemiology. *Acta Tropica*, 91(3), 219-225.

- [11] Jensen, J.R. (2006) Remote Sensing of the Environment: An Earth Resource Perspective. 2nd ed Prentice Hall, New Jersey.
- [12] Campbell, J.B. and Wynne, R.H. (2011) Introduction to Remote Sensing. 5th ed The Guilford Press, New York.
- [13] Rees, W.G. (2001) Physical Principles of Remote Sensing. 2nd ed Cambridge University Press, Cambridge, UK.
- [14] Berberan, A. (2003) Elementos de Fotogrametria: Aplicada à aquisição de informação geográfica. Edição do Autor.
- [15] Fonseca, A.D. and Fernandes, J.C. (2004) Detecção Remota. Lidel, Lisboa, Portugal.
- [16] Tomlinson, C.J., Chapman, L., Thornes, J.E., and Baker, C. (2011) Remote sensing land surface temperature for meteorology and climatology: a review. *Meteorological Applications*, 18(3), 296-306.
- [17] Herbreteau, V., Salem, G., Souris, M., Hugot, J.P., and Gonzalez, J.P. (2007) Thirty years of use and improvement of remote sensing, applied to epidemiology: from early promises to lasting frustration. *Health Place*, 13(2), 400-3.
- [18] Dambach, P., Machault, V., Lacaux, J.P., Vignolles, C., Sie, A., and Sauerborn, R. (2012) Utilization of combined remote sensing techniques to detect environmental variables influencing malaria vector densities in rural West Africa. *International Journal of Health Geographics*, 11, 8.
- [19] Rogers, D.J., Randolph, S.E., Snow, R.W., and Hay, S.I. (2002) Satellite imagery in the study and forecast of malaria. *Nature*, 415(6872), 710-5.
- [20] Maxwell, S.K., Meliker, J.R., and Goovaerts, P. (2010) Use of land surface remotely sensed satellite and airborne data for environmental exposure assessment in cancer research. *Journal of Exposure Science and Environmental Epidemiology*, 20(2), 176-85.
- [21] Wang, Z., Chen, L., Tao, J., Zhang, Y., and Su, L. (2010) Satellite-based estimation of regional particulate matter (PM) in Beijing using vertical-and-RH correcting method. *Remote Sensing of Environment*, 114(1), 50-63.
- [22] Péré, J.C., Pont, V., Mallet, M., and Bessagnet, B. (2009) Mapping of PM10 surface concentrations derived from satellite observations of aerosol optical thickness over South-Eastern France. *Atmospheric Research*, 91(1), 1-8.
- [23] Gupta, P., Christopher, S.A., Wang, J., Gehrig, R., Lee, Y., and Kumar, N. (2006) Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment*, 40(30), 5880-5892.

- [24] Kumar, N., Chu, A.D., Foster, A.D., Peters, T., and Willis, R. (2011) Satellite Remote Sensing for Developing Time and Space Resolved Estimates of Ambient Particulate in Cleveland, OH. *Aerosol Science and Technology*, 45(9), 1090-1108.
- [25] van Donkelaar, A., Martin, R.V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., and Villeneuve, P.J. (2010) Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environmental Health Perspectives*, 118(6), 847-55.
- [26] van Donkelaar, A., Martin, R.V., Levy, R.C., da Silva, A.M., Krzyzanowski, M., Chubarova, N.E., Semutnikova, E., and Cohen, A.J. (2011) Satellite-based estimates of ground-level fine particulate matter during extreme events: A case study of the Moscow fires in 2010. *Atmospheric Environment*, 45(34), 6225-6232.
- [27] Martin, R.V. (2008) Satellite remote sensing of surface air quality. *Atmospheric Environment*, 42(34), 7823-7843.
- [28] Global Initiative for Asthma (GINA). Global strategy for asthma management and prevention. [Internet]. 2011 [cited 2012 Jun 15]. Available from: http://www.ginasthma.org/uploads/users/files/GINA_Report2011_May4.pdf.
- [29] Akinbami, L.J., Lynch, C.D., Parker, J.D., and Woodruff, T.J. (2010) The association between childhood asthma prevalence and monitored air pollutants in metropolitan areas, United States, 2001–2004. *Environmental Research*, 110(3), 294-301.
- [30] de Sousa, J., Santo, M., Colaco, T., Almada-Lobo, F., and Yaphe, J. (2011) Asthma in an Urban Population in Portugal: A prevalence study. *BMC Public Health*, 11(1), 347.
- [31] Portnov, B.A., Reiser, B., Karkabi, K., Cohen-Kastel, O., and Dubnov, J. (2011) High prevalence of childhood asthma in Northern Israel is linked to air pollution by particulate matter: evidence from GIS analysis and Bayesian Model Averaging. *International Journal of Environmental Health Research*, 22(3), 249-269.
- [32] Gent, J.F., Triche, E.W., Holford, T.R., Belanger, K., Bracken, M.B., Beckett, W.S., and Leaderer, B.P. (2003) Association of Low-Level Ozone and Fine Particles With Respiratory Symptoms in Children With Asthma. *JAMA: The Journal of the American Medical Association*, 290(14), 1859-1867.
- [33] Direcção-Geral de Saúde (2000) Programa Nacional de Controlo da Asma. Direcção-Geral de Saúde, Lisboa, Portugal.
- [34] Delamater, P.L., Finley, A.O., and Banerjee, S. (2012) An analysis of asthma hospitalizations, air pollution, and weather conditions in Los Angeles County, California. *Science of The Total Environment*, 425, 110-118.

- [35] Sousa, S.I.V., Ferraz, C., Alvim-Ferraz, M.C.M., Martins, F.G., Vaz, L.G., and Pereira, M.C. (2011) Spirometric tests to assess the prevalence of childhood asthma at Portuguese rural areas: Influence of exposure to high ozone levels. *Environment International*, 37(2), 474-478.
- [36] Babin, S.M., Burkom, H.S., Holtry, R.S., Taberner, N.R., Stokes, L.D., Davies-Cole, J.O., DeHaan, K., and Lee, D.H. (2007) Pediatric patient asthma-related emergency department visits and admissions in Washington, DC, from 2001-2004, and associations with air quality, socio-economic status and age group. *Environmental Health*, 6(1), 9.
- [37] Delfino, R.J., Gong, H., Jr., Linn, W.S., Pellizzari, E.D., and Hu, Y. (2003) Asthma symptoms in Hispanic children and daily ambient exposures to toxic and criteria air pollutants. *Environmental Health Perspectives*, 111(4), 647-56.
- [38] Delfino, R.J., Zeiger, R.S., Seltzer, J.M., and Street, D.H. (1998) Symptoms in pediatric asthmatics and air pollution: differences in effects by symptom severity, anti-inflammatory medication use and particulate averaging time. *Environmental Health Perspectives*, 106(11), 751–761.
- [39] Ko, F.W.S., Tam, W., Wong, T.W., Lai, C.K.W., Wong, G.W.K., Leung, T.F., Ng, S.S.S., and Hui, D.S.C. (2007) Effects of air pollution on asthma hospitalization rates in different age groups in Hong Kong. *Clinical & Experimental Allergy*, 37(9), 1312-1319.
- [40] Mortimer, K.M., Neas, L.M., Dockery, D.W., Redline, S., and Tager, I.B. (2002) The effect of air pollution on inner-city children with asthma. *European Respiratory Journal*, 19(4), 699-705.
- [41] Villeneuve, P.J., Chen, L., Rowe, B.H., and Coates, F. (2007) Outdoor air pollution and emergency department visits for asthma among children and adults: a case-crossover study in northern Alberta, Canada. *Environmental Health*, 6, 40.
- [42] Wilhelm, M., Meng, Y.Y., Rull, R.P., English, P., Balmes, J., and Ritz, B. (2008) Environmental public health tracking of childhood asthma using California health interview survey, traffic, and outdoor air pollution data. *Environmental Health Perspectives*, 116(9), 1254-60.
- [43] O'Connor, G.T., Neas, L., Vaughn, B., Kattan, M., Mitchell, H., Crain, E.F., Evans, R., 3rd, Gruchalla, R., Morgan, W., Stout, J., Adams, G.K., and Lippmann, M. (2008) Acute respiratory health effects of air pollution on children with asthma in US inner cities. *Journal of Allergy and Clinical Immunology*, 121(5), 1133-1139 e1.
- [44] Tecer, L.H., Alagha, O., Karaca, F., Tuncel, G., and Eldes, N. (2008) Particulate matter (PM(2.5), PM(10-2.5), and PM(10)) and children's hospital admissions for

- asthma and respiratory diseases: a bidirectional case-crossover study. *Journal of Toxicology and Environmental Health, Part A*, 71(8), 512-20.
- [45] Erbas, B., Kelly, A.M., Physick, B., Code, C., and Edwards, M. (2005) Air pollution and childhood asthma emergency hospital admissions: estimating intra-city regional variations. *International Journal of Environmental Health Research*, 15(1), 11-20.
- [46] Schwartz, J., Slater, D., Larson, T.V., Pierson, W.E., and Koenig, J.Q. (1993) Particulate air pollution and hospital emergency room visits for asthma in Seattle. *American Review of Respiratory Disease*, 147(4), 826-31.
- [47] Lavigne, E., Villeneuve, P.J., and Cakmak, S. (2012) Air pollution and emergency department visits for asthma in Windsor, Canada. *Canadian Journal of Public Health*, 103(1), 4-8.
- [48] Alvim-Ferraz, M.C.M., Pereira, M.C., Ferraz, J.M., Almeida e Mello, A.M.C., and Martins, F.G. (2005) European Directives for Air Quality: Analysis of the New Limits in Comparison with Asthmatic Symptoms in Children Living in the Oporto Metropolitan Area, Portugal. *Human and Ecological Risk Assessment: An International Journal*, 11(3), 607-616.
- [49] Burr, M.L., Emberlin, J.C., Treu, R., Cheng, S., and Pearce, N.E. (2003) Pollen counts in relation to the prevalence of allergic rhinoconjunctivitis, asthma and atopic eczema in the International Study of Asthma and Allergies in Childhood (ISAAC). *Clinical & Experimental Allergy*, 33(12), 1675-80.
- [50] Erbas, B., Akram, M., Dharmage, S.C., Tham, R., Dennekamp, M., Newbigin, E., Taylor, P., Tang, M.L., and Abramson, M.J. (2012) The role of seasonal grass pollen on childhood asthma emergency department presentations. *Clinical & Experimental Allergy*, 42(5), 799-805.
- [51] Heguy, L., Garneau, M., Goldberg, M.S., Raphoz, M., Guay, F., and Valois, M.F. (2008) Associations between grass and weed pollen and emergency department visits for asthma among children in Montreal. *Environmental Research*, 106(2), 203-11.
- [52] Taylor, P.E. and Jonsson, H. (2004) Thunderstorm asthma. *Current Allergy and Asthma Reports*, 4(5), 409-13.
- [53] Villeneuve, P.J., Leech, J., and Bourque, D. (2005) Frequency of emergency room visits for childhood asthma in Ottawa, Canada: the role of weather. *International Journal of Biometeorology*, 50(1), 48-56.
- [54] Celenza, A., Fothergill, J., Kupek, E., and Shaw, R.J. (1996) Thunderstorm associated asthma: a detailed analysis of environmental factors. *British Medical Journal*, 312(7031), 604-7.

- [55] Ehara, A., Takasaki, H., Takeda, Y., Kida, T., Mizukami, S., Hagiwara, M., and Yamada, Y. (2000) Are high barometric pressure, low humidity and diurnal change of temperature related to the onset of asthmatic symptoms? *Pediatrics International*, 42(3), 272-4.
- [56] King, M.D. (2003) EOS Data Product Handbook - Volume 1. NASA, Washington.
- [57] Savtchenko, A., Ouzounov, D., Ahmad, S., Acker, J., Leptoukh, G., Koziara, J., and Nickless, D. (2004) Terra and Aqua MODIS products available from NASA GES DAAC. *Advances in Space Research*, 34(4), 710-714.
- [58] Xiong, X., Chiang, K., Sun, J., Barnes, W.L., Guenther, B., and Salomonson, V.V. (2009) NASA EOS Terra and Aqua MODIS on-orbit performance. *Advances in Space Research*, 43(3), 413-422.
- [59] Guenther, B., Xiong, X., Salomonson, V.V., Barnes, W.L., and Young, J. (2002) On-orbit performance of the Earth Observing System Moderate Resolution Imaging Spectroradiometer; first year of data. *Remote Sensing of Environment*, 83(1–2), 16-30.
- [60] Xiong, X. and Barnes, W.L. (2006) An Overview of MODIS Radiometric Calibration and Characterization. *Advances in Atmospheric Sciences*, 23(1), 69-79.
- [61] Seemann, S.W., Borbas, E.E., Li, J., Menzel, W.P., and Gumley, L.E. MODIS Atmospheric Profile Retrieval Algorithm Theoretical Basis Document. [Internet]. 2006 [cited 2012 May 31]. Available from: http://modis-atmos.gsfc.nasa.gov/_docs/MOD07MYD07ATBDC005.pdf.
- [62] Levy, R.C., Remer, L.A., Tanré, D., Mattoo, S., and Kaufman, Y.J. Algorithm for Remote Sensing of Tropospheric Aerosol Over Dark Targets from MODIS: Collections 005 and 051: Revision 2. [Internet]. 2009 [cited 2012 Jun 1]. Available from: http://modis-atmos.gsfc.nasa.gov/_docs/ATBD_MOD04_C005_rev2.pdf.
- [63] Remer, L.A., Kaufman, Y.J., Tanre, D., Mattoo, S., Chu, D.A., Martins, J.V., Li, R.R., Ichoku, C., Levy, R.C., Kleidman, R.G., Eck, T.F., Vermote, E., and Holben, B.N. (2005) The MODIS aerosol algorithm, products, and validation. *Journal of the Atmospheric Sciences*, 62(4), 947-973.
- [64] Levy, R.C., Remer, L.A., Tanre, D., Kaufman, Y.J., Ichoku, C., Holben, B.N., Livingston, J.M., Russell, P.B., and Maring, H. (2003) Evaluation of the Moderate-Resolution Imaging Spectroradiometer (MODIS) retrievals of dust aerosol over the ocean during PRIDE. *Journal of Geophysical Research-Atmospheres*, 108(D19).
- [65] Holben, B.N., Eck, T.F., Slutsker, I., Tanré, D., Buis, J.P., Setzer, A., Vermote, E., Reagan, J.A., Kaufman, Y.J., Nakajima, T., Lavenue, F., Jankowiak, I., and Smirnov, A. (1998) AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization. *Remote Sensing of Environment*, 66(1), 1-16.

- [66] Solano, R., Didan, K., Jacobson, A., and Huete, A. MODIS Vegetation Index User's Guide (MOD13 Series). [Internet]. 2010 [cited 2012 Jun 2]. Available from: http://vip.arizona.edu/documents/MODIS/MODIS_VI_UsersGuide_01_2012.pdf.
- [67] Rouse, J., Haas, R.H., Schell, J.A., and Deering, D.W., "Monitoring Vegetation Systems in the Great Plains with ERTS," presented at the Third Earth Resources Technology Satellite-1 Symposium, Greenbelt, 1974.
- [68] Huete, A., Justice, C., and van Leeuwen, W. MODIS Vegetation Index (MOD 13): Algorithm Theoretical Basis Document. [Internet]. 1999 [cited 2012 Jun 2]. Available from: http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf.
- [69] The HDF Group. HDF-EOS Project. [Internet]. 2011 [cited 2012 Jun 1]. Available from: <http://www.hdfgroup.org/hdfEOS.html>.
- [70] Hughes Applied Information Systems. The HDF-EOS swath Concept. [Internet]. 1995 [cited 2012 Jun 3]. Available from: <http://edhs1.gsfc.nasa.gov/waisdata/docsw/pdf/wp1700301.pdf>.
- [71] Raytheon Systems Company. HDF-EOS Library User's Guide for the ECS Project, Volume 1: Overview and Examples. [Internet]. 1999 [cited 2012 Jun 3]. Available from: <http://edhs1.gsfc.nasa.gov/waisdata/sdp/pdf/tp17050001.pdf>.
- [72] Hubanks, P.A. MODIS Atmosphere QA Plan for Collection 005 and 051. [Internet]. 2012 [cited 2012 Apr 19]. Available from: http://modis-atmos.gsfc.nasa.gov/_docs/QA_Plan_2012_01_12.pdf.
- [73] National Snow & Ice Data Center (NSIDC). MODIS Sinusoidal Grid. [Internet]. [cited 2012 Jul 7]. Available from: http://nsidc.org/data/docs/daac/MODIS_V5/landgrid_v5.html.
- [74] U. S. Geological Survey. SRTM Topography. [Internet]. 2009 [cited 2012 Jun 3]. Available from: http://dds.cr.usgs.gov/srtm/version2_1/Documentation/SRTM_Topo.pdf.
- [75] U. S. Geological Survey. Interferometry and SRTM - An Overview. [Internet]. 2008 [cited 2012 Jun 3]. Available from: <http://srtm.usgs.gov/data/interferometry.php>.
- [76] Rodriguez, E., Morris, C.S., Belz, J.E., Chapin, E.C., Martin, J.M., Daffer, W., and Hensley, S. An assessment of the SRTM topographic products. [Internet]. 2005 [cited 2012 Jun 4]. Available from: http://www2.jpl.nasa.gov/srtm/SRTM_D31639.pdf.
- [77] NOAA National Climate Data Center (NCDC). What is NCDC. [Internet]. 2010 [cited 2012 Jun 4]. Available from: <http://www.ncdc.noaa.gov/oa/about/whatisncdc.html>.

- [78] NOAA National Climate Data Center (NCDC). Global Surface Summary of Day Data. [Internet]. 2006 [cited 2012 Feb 18]. Available from: http://www7.ncdc.noaa.gov/CDO/GSOD_DESC.txt.
- [79] Agência Portuguesa do Ambiente (APA). QualAr - Base de Dados On-line sobre Qualidade do Ar. [Internet]. [cited 2012 Jun 5]. Available from: <http://www.qualar.org/>.
- [80] Instituto Geográfico Português (IGP). Carta Administrativa Oficial de Portugal (CAOP). [Internet]. 2012 [cited 2012 Jun 5]. Available from: <http://www.igeo.pt/produtos/cadastro/caop/inicial.htm>.
- [81] Painho, M. and Caetano, M. (2006) Cartografia de ocupação do solo: Portugal continental, 1985-2000: CORINE Land Cover 2000. Instituto do Ambiente, Amadora, Portugal.
- [82] Caetano, M., Nunes, V., and Nunes, A., "CORINE Land Cover 2006 for Continental Portugal," Technical Report, Instituto Geográfico Português, 2009.
- [83] Departamento de Estatísticas Demográficas e Sociais - Serviço de Estatísticas Demográficas, "Estimativas Anuais da População Residente 113/1.1," Documento Metodológico, Instituto Nacional de Estatística (INE), 2011.
- [84] Vancutsem, C., Ceccato, P., Dinku, T., and Connor, S.J. (2010) Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. *Remote Sensing of Environment*, 114(2), 449-465.
- [85] Benali, A., Carvalho, A.C., Nunes, J.P., Carvalhais, N., and Santos, A. (2012) Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sensing of Environment*, 124, 108-121.
- [86] Yang, S.-K. and Smith, G.L. (1985) Further Study on Atmospheric Lapse Rate Regimes. *Journal of the Atmospheric Sciences*, 42(9), 961-966.
- [87] Blandford, T.R., Humes, K.S., Harshburger, B.J., Moore, B.C., Walden, V.P., and Ye, H. (2008) Seasonal and Synoptic Variations in Near-Surface Air Temperature Lapse Rates in a Mountainous Basin. *Journal of Applied Meteorology and Climatology*, 47(1), 249-261.
- [88] Rolland, C. (2003) Spatial and Seasonal Variations of Air Temperature Lapse Rates in Alpine Regions. *Journal of Climate*, 16(7), 1032-1046.
- [89] Seinfeld, J.H. and Pandis, S.N. (1998) *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. 1st ed Wiley-Interscience, New York.
- [90] Feng, C., Xiaofeng, Z., Hong, Y., and Huiying, H., "Retrieving land surface temperature from Landsat TM using different atmospheric products as ancillary data," presented at the Spatial Data Mining and Geographical Knowledge Services (ICSDM), 2011 IEEE International Conference on, 2011.

- [91] Gravenstein, J.S., Jaffe, M.B., Gravenstein, N., and Paulus, D.A. (2011). Capnometry monitoring in high- and low-pressure environments. In (Eds.), *Capnography* (2nd ed.) (pp. 115-126). New York: Cambridge University Press.
- [92] Wenhua, G., Fengsheng, Z., Youfu, X., and Xuan, F. (2008) Validation of the Surface Air Temperature Products Retrieved From the Atmospheric Infrared Sounder Over China. *Geoscience and Remote Sensing, IEEE Transactions on*, 46(6), 1783-1789.
- [93] Reitan, C.H. (1963) Surface Dew Point and Water Vapor Aloft. *Journal of Applied Meteorology*, 2(6), 776-779.
- [94] National Oceanic and Atmospheric Administration (NOAA). Vapor Pressure. [Internet]. [cited 2012 Feb 23]. Available from: <http://www.wrh.noaa.gov/slc/projects/wxcalc/formulas/vaporPressure.pdf>.
- [95] Câmara, G. and Medeiros, J.S. (1998) *Geoprocessamento para Projetos Ambientais*. 2ª ed Instituto Nacional de Pesquisas Espaciais, São Paulo, Brasil.
- [96] Sarma, D.D. (2009) *Geostatistics with Applications in Earth Sciences*. 2nd ed Springer, New Delhi, India.
- [97] Wong, D.W., Yuan, L., and Perlin, S.A. (2004) Comparison of spatial interpolation methods for the estimation of air quality data. *Journal of Exposure Analysis and Environmental Epidemiology*, 14(5), 404-415.
- [98] Ryan, P.H. and LeMasters, G.K. (2007) A review of land-use regression models for characterizing intraurban air pollution exposure. *Inhalation Toxicology*, 19 Suppl 1, 127-33.
- [99] Borrego, C., Monteiro, A., Pay, M.T., Ribeiro, I., Miranda, A.I., Basart, S., and Baldasano, J.M. (2011) How bias-correction can improve air quality forecasts over Portugal. *Atmospheric Environment*, 45(37), 6629-6641.
- [100] Zou, B., Wilson, J.G., Zhan, F.B., and Zeng, Y. (2009) Air pollution exposure assessment methods utilized in epidemiological studies. *Journal of Environmental Monitoring*, 11(3), 475-90.
- [101] Briggs, D.J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebret, E., Pryn, K., Van Reeuwijk, H., Smallbone, K., and Van Der Veen, A. (1997) Mapping urban air pollution using GIS: a regression-based approach. *International Journal of Geographical Information Science*, 11(7), 699-718.
- [102] Mao, L., Qiu, Y., Kusano, C., and Xu, X. (2012) Predicting regional space-time variation of PM_{2.5} with land-use regression model and MODIS data. *Environmental Science and Pollution Research*, 19(1), 128-138.
- [103] Mercer, L.D., Szpiro, A.A., Sheppard, L., Lindstrom, J., Adar, S.D., Allen, R.W., Avol, E.L., Oron, A.P., Larson, T., Liu, L.J., and Kaufman, J.D. (2011) Comparing

- universal kriging and land-use regression for predicting concentrations of gaseous oxides of nitrogen (NO_x) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Atmospheric Environment*, 45(26), 4412-4420.
- [104] National Aeronautics and Space Administration (NASA). MOD04_L2 - Format & Content. [Internet]. 2012 [cited 2012 Jul 12]. Available from: http://modis-atmos.gsfc.nasa.gov/MOD04_L2/format.html.
- [105] Geospatial Data Abstraction Library (GDAL). OGR Virtual Format. [Internet]. [cited 2012 Jul 1]. Available from: http://www.gdal.org/ogr/drv_vrt.html.
- [106] Witten, I.H. and Frank, E. (2005) *Data Mining: Practical Machine Learning Tools and Techniques*. 2nd ed Morgan Kaufmann, California.
- [107] United States Environmental Protection Agency (EPA). The Ozone Problem. [Internet]. 2012 [cited 2012 Sep 23]. Available from: http://www.epa.gov/region1/airquality/oz_prob.html.
- [108] Davis, M.S., Malayer, J.R., Vandeventer, L., Royer, C.M., McKenzie, E.C., and Williamson, K.K. (2005) Cold weather exercise and airway cytokine expression. *Journal of Applied Physiology*, 98(6), 2132-6.
- [109] Jamason, P.F., Kalkstein, L.S., and Gergen, P.J. (1997) A synoptic evaluation of asthma hospital admissions in New York City. *American Journal of Respiratory and Critical Care Medicine*, 156(6), 1781-8.
- [110] Chauhan, A.J., Inskip, H.M., Linaker, C.H., Smith, S., Schreiber, J., Johnston, S.L., and Holgate, S.T. (2003) Personal exposure to nitrogen dioxide (NO₂) and the severity of virus-induced asthma in children. *Lancet*, 361(9373), 1939-44.
- [111] Chauhan, A.J. and Johnston, S.L. (2003) Air pollution and infection in respiratory illness. *British Medical Bulletin*, 68, 95-112.
- [112] McConnell, R., Berhane, K., Gilliland, F., Molitor, J., Thomas, D., Lurmann, F., Avol, E., Gauderman, W.J., and Peters, J.M. (2003) Prospective study of air pollution and bronchitic symptoms in children with asthma. *American Journal of Respiratory and Critical Care Medicine*, 168(7), 790-7.
- [113] Taylor, P.E., Flagan, R.C., Valenta, R., and Glovsky, M.M. (2002) Release of allergens as respirable aerosols: A link between grass pollen and asthma. *Journal of Allergy and Clinical Immunology*, 109(1), 51-56.
- [114] Bury, J.D. (1972) Climate and chest disorders. *British Medical Journal*, 4(5840), 613.

Anexos

Anexo A

FormatarDadosQualAr.bas:

Sub Macro1()

```
Dim ndias As Integer
Dim nestacoes As Integer
Dim nome_estacao, inicio_dados_estacao, fim_dados_estacao As String
Dim col, nome_workbook As String
Dim Poluentes(1) As String
Dim k As Integer

' Definir poluentes
Poluentes(0) = "NO2"
Poluentes(1) = "PM10"

' Diretório dos poluentes
file_dir = ActiveWorkbook.Path
poluentes_dir = Left(file_dir, 3) & "Dados\QualAr\"

For Each p In Poluentes

' Definir intervalo de tempo dos dados
ano_inicial = 2003
ano_final = 2008

For k = ano_inicial To ano_final

    Debug.Print (poluentes_dir & p & "\" & p & "_" & k & ".xlsm")

' Abrir o ficheiro excel
Workbooks.Open Filename:=poluentes_dir & p & "\" & p & "_" & k & _
".xlsm", ReadOnly:=False
' Ativar o ficheiro anteriormente aberto
Workbooks(2).Activate

' Apagar folha Dados_Formatados existente
If Sheets.Count > 1 Then
' Desactivar mensagem que pede para confirmar o apagamento da folha
Application.DisplayAlerts = False
Worksheets("Dados_Formatados").Delete
Application.DisplayAlerts = True
End If

' Criar folha Dados_Formatados
Worksheets.Add(After:=Sheets(Sheets.Count)).Name = "Dados_Formatados"

' Escrever cabeçalho
Worksheets("Dados_Formatados").Range("A:H").ClearContents
Worksheets("Dados_Formatados").Range("A1").FormulaLocal = "Data"
Worksheets("Dados_Formatados").Range("B1").FormulaLocal = "Ano"
```

```
Worksheets("Dados_Formatados").Range("C1").FormulaLocal = "Mes"
Worksheets("Dados_Formatados").Range("D1").FormulaLocal = "Dia"
Worksheets("Dados_Formatados").Range("E1").FormulaLocal = "Hora"
Worksheets("Dados_Formatados").Range("F1").FormulaLocal = "Estacao"
Worksheets("Dados_Formatados").Range("G1").FormulaLocal = p
```

With Worksheets("Dados")

```
ndias = .Range("A3", .Range("A3").End(xlDown).Address).Count
nestacoes = .Range("B1", .Range("B1").End(xlToRight).Address).Count
```

' Percorrer todas as estações

For i = 0 To nestacoes - 1

' Determinar o endereço com o nome da estação

```
nome_estacao = .Range("B1").Offset(0, i).Address
```

' Determinar o endereço com o início dos dados da estação

```
inicio_dados_estacao = .Range("B3").Offset(0, i).Address
```

' Determinar o endereço com o fim dos dados da estação

```
str_split = Split(nome_estacao, "$")
```

```
col = str_split(1)
```

```
fim_dados_estacao = col & ndias + 2
```

' Copiar dias

```
.Range("A3", .Range("A3").End(xlDown).Address).Copy _
Destination:=Worksheets("Dados_Formatados").Range("A" & ndias * i + 2)
```

' Separar ano, mês, dia e hora

```
Worksheets("Dados_Formatados").Range("B" & ndias * i + 2). _
FormulaLocal = "=VALUE(LEFT(A2;4))"
```

```
Worksheets("Dados_Formatados").Range("B" & ndias * i + 2, _
Worksheets("Dados_Formatados").Range("A2").End(xlDown). _
Offset(0, 1).Address).FillDown
```

```
Worksheets("Dados_Formatados").Range("C" & ndias * i + 2). _
FormulaLocal = "=VALUE(LEFT(RIGHT(A2;14);2))"
```

```
Worksheets("Dados_Formatados").Range("C" & ndias * i + 2, _
Worksheets("Dados_Formatados").Range("A2").End(xlDown). _
Offset(0, 2).Address).FillDown
```

```
Worksheets("Dados_Formatados").Range("D" & ndias * i + 2). _
FormulaLocal = "=VALUE(LEFT(RIGHT(A2;11);2))"
```

```
Worksheets("Dados_Formatados").Range("D" & ndias * i + 2, _
Worksheets("Dados_Formatados").Range("A2").End(xlDown). _
Offset(0, 2).Address).FillDown
```

```
Worksheets("Dados_Formatados").Range("E" & ndias * i + 2). _
FormulaLocal = "=VALUE(LEFT(RIGHT(A2;8);2))"
```

```
Worksheets("Dados_Formatados").Range("E" & ndias * i + 2, _
Worksheets("Dados_Formatados").Range("A2").End(xlDown). _
Offset(0, 2).Address).FillDown
```

' Copiar nome das estações

```
Worksheets("Dados_Formatados").Range("F" & ndias * i + 2). _
FormulaLocal = .Range(nome_estacao).Value
```

```
Worksheets("Dados_Formatados").Range("F" & ndias * i + 2, _
Worksheets("Dados_Formatados").Range("A2").End(xlDown). _
```

```
Offset(0, 3).Address).FillDown
```

```
' Copiar dados da estação
```

```
.Range(inicio_dados_estacao, fim_dados_estacao).Copy _
```

```
Destination:=Worksheets("Dados_Formatados").Range("G" & ndias * i + 2)
```

```
Next i
```

```
End With
```

```
Workbooks(2).Close savechanges:=True
```

```
Next k
```

```
Next p
```

```
End Sub
```

ImportarFolhasExcel.bas:

```
Sub Macro1()
```

```
Dim Poluentes(1) As String  
Dim bd_dir As String
```

```
' Definir poluentes
```

```
Poluentes(1) = "NO2"
```

```
Poluentes(2) = "PM10"
```

```
' Diretório das BDs
```

```
bd_dir = CurrentDb().Name
```

```
poluentes_dir = Left(bd_dir, 3) & "Dados\QualAr\"
```

```
For Each p In Poluentes
```

```
  ' Definir intervalo de tempo dos dados a copiar
```

```
  ano_inicial = 2003
```

```
  ano_final = 2008
```

```
  For i = ano_inicial To ano_final
```

```
    ' Copiar folha Dados_Formatados para Tabela do Access
```

```
    nome_tabela = p & "_" & i
```

```
    DoCmd.TransferSpreadsheet acImport, acSpreadsheetTypeExcel9, _
```

```
    nome_tabela, poluentes_dir & p & "\" & nome_tabela & ".xism", _
```

```
    True, "Dados_Formatados!"
```

```
  Next i
```

```
Next p
```

```
End Sub
```

AccessMedias.bas:

Sub Macro1()

```
Dim con As ADODB.Connection
Dim cmd As ADODB.Command
Dim rst As ADODB.Recordset
Set con = CurrentProject.Connection
con.CursorLocation = adUseClient
Set cmd = New ADODB.Command
cmd.ActiveConnection = con
Dim dias As Integer
```

```
Dim Poluentes(1) As String
Dim bd_dir As String
```

```
' Definir poluentes
Poluentes(0) = "NO2"
Poluentes(1) = "PM10"
```

```
' Diretório do output
bd_dir = CurrentDb().Name
output_dir = Left(bd_dir, 3) & "LUR\"
```

For Each p In Poluentes

```
' Definir intervalo de tempo das médias (ano)
ano_inicial = 2003
ano_final = 2008
For i = ano_inicial To ano_final
```

```
' Definir intervalo de tempo das médias (mês)
mes_inicial = 1
mes_final = 1
For j = mes_inicial To mes_final
```

```
' Definir critérios das médias mensais
If i = 2004 Or i = 2008 Then
' Ano bissexto
If j = 2 Then
dias = 22 ' fevereiro
End If
Else
If j = 1 Or j = 3 Or j = 5 Or j = 7 Or j = 8 Or j = 10 Or j = 12 Then
dias = 24 ' meses de 31 dias
Elseif j = 2 Then
dias = 21 ' fevereiro
Else
dias = 23 ' meses com 30 dias
End If
End If
```

```
' Query
cmd.CommandText = _
```

```
"SELECT * " & _
"FROM (SELECT aux1.cod,aux1.lon,aux1.lat,aux1.alt,aux1.x1,aux1.y1," & _
"aux1.nom, aux1.A,aux1.M,COUNT(*) AS N2,AVG(aux1.pol) " & _
"FROM (SELECT E.Codigo AS cod,E.Lon_dec AS lon,E.Lat_dec AS lat," & _
"E.Altitude AS alt,E.X AS x1,E.Y AS y1,E.Nome AS nom,D.Ano AS A," & _
"D.Mes AS M,D.Dia,COUNT(*) AS N,AVG(D." & p & ") AS pol " & _
"FROM Estacoes AS E INNER JOIN " & p & " " & i & " AS D ON " & _
"E.Nome=D.Estacao " & _
"WHERE " & p & " Is Not Null And D.Mes = " & j & " " & _
"GROUP BY E.Codigo,E.Lat_dec,E.Lon_dec,E.Altitude,E.X,E.Y,E.Nome," & _
"D.Ano, D.Mes, D.Dia) AS aux1 " & _
"WHERE aux1.N >= 18 " & _
"GROUP BY aux1.cod,aux1.lon,aux1.lat,aux1.alt,aux1.x1,aux1.y1,aux1.nom," & _
"aux1.A, aux1.M) AS aux2 " & _
"WHERE aux2.N2 >= " & dias & ";"
```

```
' Executar query
```

```
Set rst = cmd.Execute
```

```
' String do mês
```

```
If j > 9 Then
```

```
    Mes = j
```

```
Else
```

```
    Mes = "0" & j
```

```
End If
```

```
' Abrir ficheiro CSV
```

```
File = output_dir & "QualAr_" & p & "_" & i & Mes & ".csv"
```

```
f1 = FreeFile()
```

```
Open File For Append As f1
```

```
' Escrever o Cabeçalho
```

```
Print #f1, "COD" & ";" & "LONGITUDE" & ";" & "LATITUDE" & ";" & _
```

```
"ALTITUDE" & ";" & "X" & ";" & "Y" & ";" & "NOME" & ";" & "ANO" & _
```

```
;" & "MES" & ";" & p
```

```
' Escrever o resultado da query para o ficheiro CSV
```

```
While rst.EOF = False
```

```
    Print #f1, rst.Fields(0).Value & ";" & rst.Fields(1).Value & ";" & _
```

```
    rst.Fields(2).Value & ";" & rst.Fields(3).Value & ";" & _
```

```
    rst.Fields(4).Value & ";" & rst.Fields(5).Value & ";" & _
```

```
    rst.Fields(6).Value & ";" & rst.Fields(7).Value & ";" & _
```

```
    rst.Fields(8).Value & ";" & rst.Fields(10).Value & ";"
```

```
    rst.MoveNext
```

```
Wend
```

```
Close #f1
```

```
Next j
```

```
Next i
```

```
Next p
```

```
End Sub
```

Anexo B

MODIS AP.m:

```
clear all; close all; clc;

cd ..
cd Dados\SRTM_DEM

% Ler o MDE de Portugal
[SRTM_DEM, R, bbox] = geotiffread('SRTM_DEM_PT_5km.tif');
SRTM_DEM = double(SRTM_DEM);
SRTM_DEM(SRTM_DEM == -32768) = NaN;

[linDEM colDEM] = size(SRTM_DEM);
DEMpixelsize = (bbox(2,2)-bbox(1,2))/linDEM;

% Declarar os vector com os niveis de pressão que se irá utilizar
pressure_level = [700 780 850 920 950 1000];

% Calcular a altitude pela equação determinada através das observações das
% estações meteorologicas
% -----
elevation = -8.8188*pressure_level + 8977.0354; % metros
% -----

% Construir matrizes de longitude e latitude para interpolar dados para
% grelhas regulares
% NOTA: As coordenadas dos cantos do SRTM referem-se ao canto do pixel
% e não ao centro
MODISpixelsize = 0.05;
new_lat_vector=[42.4 - MODISpixelsize/2:-MODISpixelsize:36.8 + MODISpixelsize/2];
new_lon_vector=[-9.7 + MODISpixelsize/2:MODISpixelsize:-6.1 - MODISpixelsize/2];

[new_lon_matrice new_lat_matrice]=meshgrid(new_lon_vector,new_lat_vector);

cd ..
cd Atmospheric_Profile_MOD07_L2
%cd Atmospheric_Profile_MYD07_L2\

first_year = 2003;
last_year = 2008;
for j = first_year:last_year

    year = num2str(j);

    cd(year);
    files = dir('M*.hdf')

    for i=1:numel(files)
```

```

% Ler variáveis
% -----
latitude = hdfread(files(i).name, 'Latitude');
longitude = hdfread(files(i).name, 'Longitude');
% Perfil de temperatura
temperature = hdfread(files(i).name, 'Retrieved_Temperature_Profile');

% Vapor de água (TCWV)
water_vapor = hdfread(files(i).name, 'Water_Vapor');

% Indicadores de qualidade
QA = hdfread(files(i).name, 'Quality_Assurance');
QAI = hdfread(files(i).name, 'Quality_Assurance_Infrared');
% -----

latitude = double(latitude);
longitude = double(longitude);
temperature = double(temperature);
temperature = temperature(15:20,:);
water_vapor = double(water_vapor);
water_vapor(water_vapor<0) = NaN;

% valor real = fator de escala * valor armazenado + offset
water_vapor = 0.001 * water_vapor;

QA_tm = QA(:,:,1);
QAI_wv = QAI(:,:,1);

% Dimensões do cubo de dados temperature
[layerMODIS linMODIS colMODIS] = size(temperature);

% Corrigir perfil de temperatura e vapor de água
% -----
for u=1:linMODIS
    for v=1:colMODIS
        flag = dec2bin(QA_tm(u,v),8);
        % Temperatura
        if strcmp(flag(8),'0') == 1
            temperature(:,u,v) = NaN;
        elseif strcmp(flag(7),'0') == 1
            temperature(:,u,v) = NaN;
        end

        % Vapor de água
        flag2 = dec2bin(QAI_wv(u,v),8);
        if strcmp(flag2(8),'0') == 1
            water_vapor(u,v) = NaN;
        elseif strcmp(flag2(7),'0') == 1
            water_vapor(u,v) = NaN;
        end
    end
end
% -----

% Formatar os dados de temperatura e vapor de água

```

```
% -----  
[temperature,linMODIS,colMODIS,layerMODIS] = ...  
    format_temp(temperature,latitude,longitude,new_lon_matrice,new_lat_matrice);  
  
F = TriScatteredInterp(longitude(:),latitude(:),water_vapor(:));  
wv_interp = F(new_lon_matrice,new_lat_matrice);  
% -----  
  
% Calcular temperatura à superfície  
% -----  
near_surface_temperature = nan(linDEM,colDEM);  
for u=1:linDEM  
    for v=1:colDEM  
  
        if isnan(SRTM_DEM(u,v)) == 0  
  
            notnan = find(isnan(temperature(u,v,:))==0);  
            if isempty(notnan) == 0  
  
                tempaux = temperature(u,v,notnan);  
                temp2 = zeros(1,numel(tempaux));  
                for i2 = 1:numel(temp2)  
                    temp2(i2) = tempaux(i2);  
                end  
  
                % determinar os parâmetros da equação da recta que  
                % melhor se ajusta aos pontos do perfil  
                p1 = polyfit(elevation(notnan), temp2, 1);  
                m1 = p1(1);  
                b1 = p1(2);  
  
            else  
                m1 = NaN;  
                b1 = NaN;  
            end  
  
        end  
  
        % Calcular a temperatura ao nível da superfície  
        near_surface_temperature(u,v) = m1 * (SRTM_DEM(u,v)) + b1;  
  
    end  
end  
% -----  
  
% guardar os resultados  
save([files(i).name(1:end-4),'.mat'], ...  
    'near_surface_temperature','wv_interp','-mat')  
  
clear near_surface_temperature;  
clear temperature; clear QA; clear QA_tm; clear QAI; clear QA_wv;
```

```
clear latitude; clear longitude; clear tempaux; clear moistaux; clear F;

end
clear files;
cd ..

end
```

format temp.m:

```
function [cubo_interp,l,c,layer] =
format_temp(cubo_inicial,lat,lon,new_lon_matrice,new_lat_matrice)

% O Perfil de temperatura vem no formato: (nível de pressão, linha, coluna)
% Por uma questão de simplicidade pretende-se converte-lo para o formato
% (linha, coluna, nível de pressão)
[layer l c] = size(cubo_inicial);
for i=1:l
    for j=1:c
        for k=1:layer
            cubo_inv(i,j,k) = cubo_inicial(k,i,j);
        end
    end
end
end

cubo_inv(cubo_inv<=0) = NaN;
% valor real = fator de escala * valor armazenado + offset
cubo_inv = 0.01 * (cubo_inv - (-15000));
for k=1:layer
    % Interpolar o cubo de dados
    cubo_aux = cubo_inv(:, :, k);
    F = TriScatteredInterp(lon(:),lat(:),cubo_aux(:));
    cubo_interp(:, :, k) = F(new_lon_matrice,new_lat_matrice);
end

end
```

MODIS Atmosphere move files.m:

```
clear all; close all; clc;

cd ..
cd Dados\Atmospheric_Profile_MOD07_L2\
%cd Dados\Atmospheric_Profile_MYD07_L2\

first_year = 2003;
last_year = 2008;
for j = first_year:last_year

    year = num2str(j);

    cd(year);
    files = dir('M*.mat')

    % Criar pastas mensais
    for i=1:12
        if i <=9
            command = cat(2,'md 0',num2str(i));
        else
            command = cat(2,'md ',num2str(i));
        end
        system(command);
    end

    for i=1:numel(files)
        % Ler o dia do ano
        doy = files(i).name(15:17);

        % Mover cada ficheiro consoante o seu dia do ano
        if (strcmp(year,'2000') == 1) || (strcmp(year,'2004') == 1) || (strcmp(year,'2008') == 1)
            if str2double(doy) <= 31
                system(cat(2,'move ./',files(i).name,' ./01'));
            elseif str2double(doy) <= 60
                system(cat(2,'move ./',files(i).name,' ./02'));
            elseif str2double(doy) <= 91
                system(cat(2,'move ./',files(i).name,' ./03'));
            elseif str2double(doy) <= 121
                system(cat(2,'move ./',files(i).name,' ./04'));
            elseif str2double(doy) <= 152
                system(cat(2,'move ./',files(i).name,' ./05'));
            elseif str2double(doy) <= 182
                system(cat(2,'move ./',files(i).name,' ./06'));
            elseif str2double(doy) <= 213
                system(cat(2,'move ./',files(i).name,' ./07'));
            elseif str2double(doy) <= 244
                system(cat(2,'move ./',files(i).name,' ./08'));
            elseif str2double(doy) <= 274
                system(cat(2,'move ./',files(i).name,' ./09'));
            elseif str2double(doy) <= 305
                system(cat(2,'move ./',files(i).name,' ./10'));
            elseif str2double(doy) <= 335
```

```
        system(cat(2,'move ./',files(i).name, ' ./11'));
elseif str2double(doy) <= 366
    system(cat(2,'move ./',files(i).name, ' ./12'));
end
else
if str2double(doy) <= 31
    system(cat(2,'move ./',files(i).name, ' ./01'));
elseif str2double(doy) <= 59
    system(cat(2,'move ./',files(i).name, ' ./02'));
elseif str2double(doy) <= 90
    system(cat(2,'move ./',files(i).name, ' ./03'));
elseif str2double(doy) <= 120
    system(cat(2,'move ./',files(i).name, ' ./04'));
elseif str2double(doy) <= 151
    system(cat(2,'move ./',files(i).name, ' ./05'));
elseif str2double(doy) <= 181
    system(cat(2,'move ./',files(i).name, ' ./06'));
elseif str2double(doy) <= 212
    system(cat(2,'move ./',files(i).name, ' ./07'));
elseif str2double(doy) <= 243
    system(cat(2,'move ./',files(i).name, ' ./08'));
elseif str2double(doy) <= 273
    system(cat(2,'move ./',files(i).name, ' ./09'));
elseif str2double(doy) <= 304
    system(cat(2,'move ./',files(i).name, ' ./10'));
elseif str2double(doy) <= 334
    system(cat(2,'move ./',files(i).name, ' ./11'));
elseif str2double(doy) <= 365
    system(cat(2,'move ./',files(i).name, ' ./12'));
end
end
end

clear files;
cd ..

end
```

MODIS AP avg.m:

```
clear all; clc; close all;
format long g;

cd ..
cd Dados\Atmospheric_Profile_Terra&Aqua

comb = 'TaAT'; % combinação de dados a calcular
calc_wv = 1; % 1 - calcula a média do vapor de água
           % 0 - não calcula a média do vapor de água

% Definir coordenadas do canto inferior esquerdo e tamanho do pixel
nyllcorner = 36.8;
nxllcorner = -9.7;
psize = 0.05;

first_year = 2003;
last_year = 2008;
for j = first_year:last_year

    year = num2str(j);
    cd(year);

    first_month = 1;
    last_month = 12;
    for k=first_month:last_month

        if k <= 9
            month = cat(2,'0',num2str(k));
        else
            month = num2str(k);
        end
        cd(month)

        % Listar todos os ficheiros MAT
        files = dir('*.mat')

        % Declarar variáveis
        load(files(1).name)
        [lin col] = size(near_surface_temperature);
        temp_sum = zeros(lin,col);
        temp_npixéis = zeros(lin,col);
        wv_sum = zeros(lin,col);
        wv_npixéis = zeros(lin,col);
        for i=1:numel(files)

            load(files(i).name)

            hour = str2double(files(i).name(19:20));

            for u=1:lin
                for v=1:col
```

```

% Temperatura

% Combinações de dados
if strcmp(comb,'TaAT')
    % Combinação com todas as observações
    if isnan(near_surface_temperature(u,v)) == 0
        temp_sum(u,v) = temp_sum(u,v) + near_surface_temperature(u,v);
        temp_npixéis(u,v) = temp_npixéis(u,v) + 1;
    end
elseif strcmp(comb,'TaATn')
    % Combinação sem as observações do Terra realizadas
    % durante o dia
    if (strcmp(files(i).name(1:3),'MYD') == 0) || (hour < 9)
        if isnan(near_surface_temperature(u,v)) == 0
            temp_sum(u,v) = temp_sum(u,v) + near_surface_temperature(u,v);
            temp_npixéis(u,v) = temp_npixéis(u,v) + 1;
        end
    end
elseif strcmp(comb,'TaAnT')
    % Combinação sem as observações do Aqua realizadas
    % durante o dia
    if (strcmp(files(i).name(1:3),'MOD') == 0) || (hour < 9)
        if isnan(near_surface_temperature(u,v)) == 0
            temp_sum(u,v) = temp_sum(u,v) + near_surface_temperature(u,v);
            temp_npixéis(u,v) = temp_npixéis(u,v) + 1;
        end
    end
elseif strcmp(comb,'TaAnTn')
    % Combinação sem as observações do Terra e do Aqua
    % realizadas durante o dia
    if hour < 9
        if isnan(near_surface_temperature(u,v)) == 0
            temp_sum(u,v) = temp_sum(u,v) + near_surface_temperature(u,v);
            temp_npixéis(u,v) = temp_npixéis(u,v) + 1;
        end
    end
end

% Vapor de água
if calc_wv == 1
    if isnan(wv_interp(u,v)) == 0
        wv_sum(u,v) = wv_sum(u,v) + wv_interp(u,v);
        wv_npixéis(u,v) = wv_npixéis(u,v) + 1;
    end
end

end
end
end

% Calcular as médias
temp_avg = temp_sum./temp_npixéis;
wv_avg = wv_sum./wv_npixéis;

% Converter para graus celsius

```

```
temp_avg = temp_avg - 273.15;

% Colocar valores sem dados como -9999
temp_avg(isnan(temp_avg) == 1) = -9999;
wv_avg(isnan(wv_avg) == 1) = -9999;

% Abrir e escrever o cabeçalho dos ficheiros ASCII
fileID1 = fopen(['H:\Dados\Atmospheric_Profile_Terra&Aqua\' , ...
    comb,'_MODIS_TEMP_',year,month,'.asc'],'w');
fprintf(fileID1,['ncols %d\nnrows %d\nxllcorner %f\nyllcorner %f\ncellsz ' , ...
    '%f\nnodata_value -9999'],col,lin,nxllcorner,nyllcorner,psize);

if calc_wv == 1
    fileID2 = fopen(['H:\Dados\Atmospheric_Profile_Terra&Aqua\MODIS_WV_', ...
        year,month,'.asc'],'w');
    fprintf(fileID2,['ncols %d\nnrows %d\nxllcorner %f\nyllcorner %f\ncellsz ' , ...
        '%f\nnodata_value -9999'],col,lin,nxllcorner,nyllcorner,psize);
end

% Escrever ficheiros ASCII
for p=1:lin
    fprintf(fileID1,'\n');
    fprintf(fileID1,'%0.8ft',temp_avg(p,:));
    if calc_wv == 1
        fprintf(fileID2,'\n');
        fprintf(fileID2,'%0.8ft',wv_avg(p,:));
    end
end

fclose(fileID1);
if calc_wv == 1
    fclose(fileID2);
end

cd ..

end

cd ..

end
```

MODIS L2 Geotiff.py:

```
# coding: utf-8
# Importar o módulos
import os
import glob

idir = os.getcwd()
os.chdir(os.pardir) # ir para o diretório anterior
aux = os.getcwd()

# Definir diretório com os dados
data_path = aux[0:2] + "\\Dados\\Atmospheric_Profile_Terra&Aqua"
os.chdir(data_path)

# código EPSG do sistema de coordenadas do ficheiro de entrada
cs_in = "4326" # WGS84
# código EPSG do sistema de coordenadas do ficheiro de saída
cs_out = "3763" # PTTM06
# resolução do ficheiro de saída
res = "5000"

for i in glob.glob("*.asc"):

    os.system("gdalwarp -s_srs EPSG:" + cs_in + " -t_srs EPSG:" + cs_out + " -te \
-140000 -315000 180000 290000 -tr " + res + " " + res + " -r bilinear \
-dstnodata -9999 " + i + " PTTM06_" + i[0:len(i)-4] + ".tif")

# voltar ao diretório inicial
os.chdir(idir)
```

MODIS AP_val.py:

```
# coding: utf-8
# Importar o módulos
import os
import glob
import csv

idir = os.getcwd()
os.chdir(os.pardir) # ir para o diretório anterior
aux = os.getcwd()

# Definir diretório com os dados
data_path = aux[0:2] + "\\Dados\\Atmospheric_Profile_Terra&Aqua"
os.chdir(data_path)

# Definir dados a validar
file_list = ["TaAT", "TaATn", "TaAnT", "TaAnTn"]

for f in file_list:

    # Abrir ficheiro final com a validação
    fout = open(data_path + "\\ " + f + "_val.csv", "w")
    fout.write("COD;ANO;MES;ESTACAO;MODIS")

    for i in glob.glob(f + "*.tif"):

        # Abrir ficheiros CSV mensais com as medições das estações meteorológicas
        dados = csv.reader(open("TEMP_" + i[len(i)-10:len(i)-4] + ".csv", 'rb'), delimiter=';')

        # Percorrer cada estação
        for j in dados:

            lin_dados = j

            if lin_dados[0] != "COD":

                # Código e coordenadas PTTM06 da estação e respetivo ano, mês e valor
                medido
                cod = lin_dados[0]
                x = lin_dados[4]
                y = lin_dados[5]
                ano = lin_dados[7]
                mes = lin_dados[8]
                med = lin_dados[10]

                # Determinar nos rasters o valor do ponto de coordenadas (x,y)
                os.system("gdallocationinfo -geoloc " + i + " " + x + " " + y + " \
> auxiliar_GDALLOCATION.txt")
                faux = open(data_path + "\\auxiliar_GDALLOCATION.txt", "r")
                text_aux = faux.readlines()
                faux.close()
```

```
# NOTA: text_aux[3][11:len(text_aux[3])-1] contém o valor do ponto lido no
raster
aux1 = text_aux[3][11:len(text_aux[3])-1]

fout.write("\n")
fout.write(cod + ";" + ano + ";" + mes + ";" + med + ";" + aux1)

# Fechar ficheiro com a validação
fout.close()

# Voltar ao diretório inicial
os.chdir(idir)
```

MODIS RH_avg.m:

```
clear all; clc; close all;
format long g;
cd ..
cd Dados\Atmospheric_Profile_Terra&Aqua

% Ler o MDE de Portugal
[SRTM_DEM, R, bbox] = geotiffread('PTTM06_SRTM_DEM_5km.tif');
SRTM_DEM = double(SRTM_DEM);
SRTM_DEM(SRTM_DEM == -32768) = NaN;
% Definir coordenadas do canto inferior esquerdo e tamanho do pixel
nyllcorner = bbox(1,2);
nxllcorner = bbox(1,1);
psize = 5000;

% Listar todas as imagens de vapor de água
files = dir('PTTM06_MODIS_WV*.tif')
[lin col] = size(SRTM_DEM);
for i=1:numel(files)

    % Ler imagem de vapor de água
    [WV, R1, bbox1] = geotiffread(files(i).name);
    WV = double(WV);
    WV(WV < -32768) = NaN;

    % Ler imagem de temperatura
    [TEMP, R2, bbox2] = geotiffread(['PTTM06_MODIS_TEMP_', ...
        files(i).name(end-9:end)]);
    TEMP = double(TEMP);
    TEMP(TEMP < -32768) = NaN;

    % Calcular a temperatura do ponto de orvalho (dew point)
    logWV = log(WV);
    DEWP = 6.971 + 8.715*logWV - 0.004*SRTM_DEM;

    % Calcular a humidade relativa
    ea = 6.11 * (10.^(7.5*DEWP./(237.7+DEWP)));
    es = 6.11 * (10.^(7.5*TEMP./(237.7+TEMP)));
    RH = ea*100./es;
    RH(isnan(RH) == 1) = -9999;

    % Abrir e escrever o cabeçalho dos ficheiros ASCII
    fileID = fopen(cat(2,'RH_',files(i).name(end-9:end-4),'.txt'),'w');
    fprintf(fileID,['ncols %d\nnrows %d\nnxllcorner %f\nnyllcorner %f\ncellsz ' , ...
        '%f\nnodata_value -9999'],col,lin,nxllcorner,nyllcorner,psize);
    % Escrever ficheiros ASCII
    for p=1:lin
        fprintf(fileID,'\n');
        fprintf(fileID,'%0.8f\t',RH(p,:));
    end

    fclose(fileID);
end
```

MODIS_AOT.m:

```

clear all; close all; clc;

% Construir matrizes de longitude e latitude para interpolar dados para
% grelhas regulares
MODISpixelsize = 0.1;
new_lat_vector=[42.4 - MODISpixelsize/2:-MODISpixelsize:36.8 + MODISpixelsize/2];
new_lon_vector=[-9.7 + MODISpixelsize/2:MODISpixelsize:-6.1 - MODISpixelsize/2];
[new_lon_matrice new_lat_matrice]=meshgrid(new_lon_vector,new_lat_vector);

cd ..
cd Dados\AOT_MOD04_L2\
% cd Dados\AOT_MYD04_L2\

first_year = 2003;
last_year = 2008;
for j = first_year:last_year

    year = num2str(j);

    cd(year);
    files = dir('M*.hdf')

    for i=1:numel(files)

        latitude = hdfread(files(i).name, 'Latitude');
        longitude = hdfread(files(i).name, 'Longitude');
        % AOT 550 nm
        AOT_550_best = hdfread(files(i).name, 'mod04', 'Fields', ...
            'Optical_Depth_Land_And_Ocean','Index',{[],[],[]});

        latitude = double(latitude);
        longitude = double(longitude);
        AOT_550_best = double(AOT_550_best);
        % valor real = fator de escala * valor armazenado + offset
        AOT_550_best = 0.001 * AOT_550_best;
        AOT_550_best(AOT_550_best < -0.05) = NaN;

        % Interpolar para uma grelha regular (bilinear)
        F = TriScatteredInterp(longitude(:),latitude(:),AOT_550_best(:));
        AOT_550_best_interp = F(new_lon_matrice,new_lat_matrice);

        % Guardar grelha regular do AOT
        save([files(i).name(1:end-4),'.mat'],'AOT_550_best_interp','-mat')

        clear F; clear latitude; clear longitude;

    end

clear files;
cd ..

end

```

MODIS AOT avg.m:

```
clear all; clc; close all;
format long g;

cd ..
cd Dados\AOT_Terra&Aqua\

% Definir coordenadas do canto inferior esquerdo e tamanho do pixel
nyllcorner = 36.8;
nxllcorner = -9.7;
psize = 0.1;

first_year = 2003;
last_year = 2008;
for j = first_year:last_year

    first_month = 1;
    last_month = 12;
    for k=first_month:last_month

        if k <= 9
            month = cat(2,'0',num2str(k));
        else
            month = num2str(k);
        end
        cd(month)

        % Listar todos os ficheiros MAT
        files = dir('*.*mat')

        % Declarar variáveis
        load(files(1).name)
        [lin col] = size(AOT_550_best_interp);
        AOT550_sum = zeros(lin,col);
        AOT550_npixels = zeros(lin,col);

        for i=1:numel(files)

            load(files(i).name)

            for u=1:lin
                for v=1:col

                    if isnan(AOT_550_best_interp(u,v)) == 0
                        AOT550_sum(u,v) = AOT550_sum(u,v) + AOT_550_best_interp(u,v);
                        AOT550_npixels(u,v) = AOT550_npixels(u,v) + 1;
                    end

                end
            end

        end

    end

end
```

```
% Calcular a média
AOT550_avg = AOT550_sum./AOT550_npixels;

% Colocar valores sem dados como -9999
AOT550_avg(isnan(AOT550_avg) == 1) = -9999;

% Abrir e escrever o cabeçalho dos ficheiros ASCII
fileID = fopen(['H:\Dados\AOT_Terra&Aqua\MODIS_AOT550_', ...
    year,month,'.asc'], 'w');
fprintf(fileID,['ncols %d\nnrows %d\nxllcorner %f\nylcorner %f\ncellsz', ...
    '%f\nndata_value -9999'],col,lin,nxllcorner,nyllcorner,psize);

% Escrever ficheiros ASCII
for p=1:lin
    fprintf(fileID, '\n');
    fprintf(fileID, '%.8ft', AOT550_avg(p,:));
end

fclose(fileID);
cd ..

end

cd ..
end
```

MODIS NDVI QA.m:

```

clear all; close all; clc;

cd ..
cd Dados\NDVI_MYD13A3\

% Listar todos os ficheiros HDF do NDVI
files = dir('*.*hdf');

for i=1:numel(files)

    % Ler variáveis
    NDVI = hdfread(files(i).name, ['MOD_Grid_monthly_1km_VI', 'Fields', ...
    '1 km monthly NDVI', 'Index', {}, {}, {}]);
    QA = hdfread(files(i).name, ['/MOD_Grid_monthly_1km_VI/Data Fields', ...
    '1 km monthly VI Quality', 'Index'], {}, {}, {});

    % Realizar a correção dos NDVI com os indicadores de qualidade QA
    [lin col] = size(NDVI);
    for j=1:lin
        for k=1:col
            flag = dec2bin(QA(j,k));
            % bit 0-1
            % 00 VI produced, good quality
            % 01 VI produced, but check other QA
            % 10 Pixel produced, but most probably cloudy
            % 11 Pixel not produced due to other reasons than clouds
            % NDVI(j,k) = 1;
            if strcmp(flag(end-1:end), '01') == 1
                % bit 2-5
                % VI usefulness
                % 0000 Highest quality
                % 0001 Lower quality
                % 0010 Decreasing quality
                % 0100 Decreasing quality
                % 1000 Decreasing quality
                % 1001 Decreasing quality
                % 1010 Decreasing quality
                % 1100 Lowest quality
                % 1101 Quality so low that it is not useful --> 1101 = 13
                % 1110 L1B data faulty
                % 1111 Not useful for any other reason/not processed
                aux_flag = bin2dec(flag(end-5:end-2));
                if aux_flag >= 13
                    NDVI(j,k) = -3000;
                end
            elseif (strcmp(flag(end-1:end), '10') == 1) || (strcmp(flag(end-1:end), '11') == 1)
                NDVI(j,k) = -3000;
            end
        end
    end
end

% Reduzir variação do NDVI para 0 a 10000
NDVI((NDVI>=-2000) & (NDVI<=0)) = 0;

```

```
% Aplicar fator de escala
NDVI(NDVI>=0) = NDVI(NDVI>=0)*0.0001;

% Guardar NDVI corrigido num ficheiro H5
hdf5write(cat(2,files(i).name(1:end-4),'_flag','.h5'), '/dataset1', NDVI);
end
```

MODIS NDVI reproject.py:

```

# coding: utf-8
# Importar o módulos
import os
import glob

var_dir = "./Dados/NDVI_MOD13A3"
output_name = "MOD13A3"

os.chdir(var_dir)
print os.getcwd()

# Sistema de coordenadas sinusoidal do MOD13A3
modis_sin_cs = "PROJCS[\"unnamed\",\\
GEOGCS[\"Unknown_datum_based_upon_the_custom_spheroid\",\\
DATUM[\"Not_specified_based_on_custom_spheroid\",\\
SPHEROID[\"Custom_spheroid\",6371007.181,0]],\\
PRIMEM[\"Greenwich\",0],\\
UNIT[\"degree\",0.0174532925199433]],\\
PROJECTION[\"Sinusoidal\"],\\
PARAMETER[\"longitude_of_center\",0],\\
PARAMETER[\"false_easting\",0],\\
PARAMETER[\"false_northing\",0],\\
UNIT[\"metre\",1,AUTHORITY[\"EPSG\",\"9001\"]]]]"

# Associar a cada imagem o sistema e coordenadas sinusoidal e coordenadas dos cantos
for i in glob.glob("*v04*flag.h5"):
    os.system("gdal_translate -a_srs " + modis_sin_cs + " \
-a_ullr -1111950.519667 5559752.598333 0.000 4447802.078667 \
-a_nodata -3000 " + i + " " + i[0:len(i)-3] + ".tif")

for i in glob.glob("*v05*flag.h5"):
    os.system("gdal_translate -a_srs " + modis_sin_cs + " \
-a_ullr -1111950.519667 4447802.078667 0.000 3335851.559000 \
-a_nodata -3000 " + i + " " + i[0:len(i)-3] + ".tif")

# Unir imagens
for i in glob.glob("*v04*flag.tif"):
    doy_v04 = i[9:16]
    for j in glob.glob("*v05*flag.tif"):
        doy_v05 = j[9:16]
        if doy_v04 == doy_v05:
            os.system("gdal_merge.bat -o \
./SIN_" + output_name + "_" + doy_v04 + ".tif " + i + " " + j + " -a_nodata -3000")

# Re projetar imagens para PTTM06
for i in glob.glob("SIN*.tif"):

    # Os ficheiros NDVI originais possuem o ano e o primeiro dia do ano do mês a que dizem respeito
    # Os ficheiros NDVI em PTTM06 serão identificados pelo ano e mês
    if (int(i[len(i)-11:len(i)-7]) == 2004) or (int(i[len(i)-11:len(i)-7]) == 2008):

```

```
a = 1
else:
    a = 0

if int(i[len(i)-7:len(i)-4]) == 1:
    mes = "01"
elif int(i[len(i)-7:len(i)-4]) == 31 + 1:
    mes = "02"
elif int(i[len(i)-7:len(i)-4]) == 59 + 1 + a:
    mes = "03"
elif int(i[len(i)-7:len(i)-4]) == 90 + 1 + a:
    mes = "04"
elif int(i[len(i)-7:len(i)-4]) == 120 + 1 + a:
    mes = "05"
elif int(i[len(i)-7:len(i)-4]) == 151 + 1 + a:
    mes = "06"
elif int(i[len(i)-7:len(i)-4]) == 181 + 1 + a:
    mes = "07"
elif int(i[len(i)-7:len(i)-4]) == 212 + 1 + a:
    mes = "08"
elif int(i[len(i)-7:len(i)-4]) == 243 + 1 + a:
    mes = "09"
elif int(i[len(i)-7:len(i)-4]) == 273 + 1 + a:
    mes = "10"
elif int(i[len(i)-7:len(i)-4]) == 304 + 1 + a:
    mes = "11"
elif int(i[len(i)-7:len(i)-4]) == 334 + 1 + a:
    mes = "12"

# Reprojecção das imagens
os.system("gdalwarp -t_srs EPSG:3763 -srcnodata -3000 \
-dstnodata -3000 -r bilinear -tr 1000 1000 \
-te -140000 -315000 180000 290000 \
NDVI_MODIS_" + i[len(i)-11:len(i)-7] + mes + ".tif")

# voltar ao diretório com os programas
os.chdir(prog_dir)
```

Anexo C

WriteKrigingModel.m:

```
clear all; close all; clc;
format long g

cd ..
cd Dados\NOAA_Climate_Data_Online\Medias_Mensais
%cd Dados\QualAr\Medias_Mensais

var = {'TEMP','RH'};
%var = {'NO2','PM10'};
for j=1:numel(var)

    cd(var{j})

    % Listar ficheiros CSV com dados pontuais
    files = dir('*.*csv')

    for i=1:numel(files)

        % Abrir o ficheiro
        fid = fopen(files(i).name);
        % Ler o cabeçalho do ficheiro
        C_text = textscan(fid, '%s', 10, 'delimiter', ',');
        % Ler os dados
        C = textscan(fid, '%d %f %f %f %f %f %s %d %d %f', 'delimiter', ',');
        if isempty(C{1}) == 0

            % Guardar as coordenadas os dados das estações
            cod = C{1};
            x = C{5};
            y = C{6};
            z = C{10};

            % Calcular o variograma experimental
            v = variogram([x y],z,'plotit',false,'maxdist',10000000);

            % Ajustar um modelo ao variogram experimental
            [dum,dum,dum,vstruct] =
            variogramfit(v.distance,v.val,[],[],[],'model','spherical','plotit',false);

            % Criar ficheiro xml com o modelo
            fileID = fopen(cat(2,files(i).name(1:end-4),'_mod','.xml'),'w');
            % Escrever código com o modelo
            fprintf(fileID,['<?xml version="1.0"?>\n', ...
            '<model xml:lang="pt" sDecimal="." name="Kriging">\n', ...
            '<dataset ID="1" Label="Dataset" dataset-type="DVA"/>\n', ...
            '<dataset Label="Dataset 2" dataset-type="DVA" optional="true"/>\n', ...
            '<dataset Label="Dataset 3" dataset-type="DVA" optional="true"/>\n', ...
            '<dataset Label="Dataset 4" dataset-type="DVA" optional="true"/>\n', ...
```

```

'<dataset Label="Decluster"s Clipping Dataset" dataset-type="Generic" ', ...
'sub-type="polygon" optional="true" visible="false"/>\n', ...
'<dataset Label="Decluster"s Clipping Dataset 2" dataset-type="Generic" ', ...
'sub-type="polygon" optional="true" visible="false"/>\n', ...
'<dataset Label="Decluster"s Clipping Dataset 3" dataset-type="Generic" ', ...
'sub-type="polygon" optional="true" visible="false"/>\n', ...
'<dataset Label="Decluster"s Clipping Dataset 4" dataset-type="Generic" ', ...
'sub-type="polygon" optional="true" visible="false"/>\n', ...
'<enum name="KrigingMethodType">Ordinary</enum>\n', ...
'<enum name="KrigingResultType">Prediction</enum>\n', ...
'<items name="Datasets">\n', ...
'<item name="Dataset">\n', ...
'<enum name="TrendType">None</enum>\n', ...
'<model xml:lang="pt" sDecimal="." name="NeighbourSearch" options="">\n', ...
'<enum name="Type">Standard</enum>\n', ...
'<bool name="Continuous">>false</bool>\n', ...
'<value name="NeighboursMax" auto="false">5</value>\n', ...
'<value name="NeighboursMin" auto="false">2</value>\n', ...
'<enum name="SectorType">Four45</enum>\n', ...
'<value name="Angle">0</value>\n', ...
'<value name="MajorSemiaxis" auto="true"></value>\n', ...
'<value name="MinorSemiaxis" auto="true"></value>\n', ...
'</model>\n', ...
'</item>\n', ...
'</items>\n', ...
'<model xml:lang="pt" sDecimal="." name="Variogram">\n', ...
'<value name="DatalayerCount">1</value>\n', ...
'<value name="NumberOfLags" auto="false">20</value>\n', ...
'<value name="LagSize" auto="false">', num2str(vstruct.h(2)-vstruct.h(1)), '</value>\n', ...
'<enum name="PairsType" auto="false">Semivariogram</enum>\n', ...
'<bool name="NuggetOn">>true</bool>\n', ...
'<value name="Nugget" auto="false">0</value>\n', ...
'<value name="MeasurementError100">100</value>\n', ...
'<bool name="ShiftON">>false</bool>\n', ...
'<bool name="VariogramModelAuto">>false</bool>\n', ...
'<model xml:lang="pt" sDecimal="." name="VariogramModel">\n', ...
'<enum name="ModelType">Spherical</enum>\n', ...
'<value name="Range" auto="false">', num2str(vstruct.range), '</value>\n', ...
'<bool name="Anisotropy">>false</bool>\n', ...
'<value name="Sill" auto="false">', num2str(vstruct.sill), '</value>\n', ...
'</model>\n', ...
'</model>\n', ...
'</model>\n'];

        fclose(fileID);
        fclose(fid);

    end

end

cd ..

end

```


ArcGIS_Geostatistics.py:

```

# coding: utf-8
# Importar o módulos
import arcpy
import os
import glob

idir = os.getcwd()
os.chdir(os.pardir) # ir para o diretório anterior
aux = os.getcwd()

# Definir diretório com os dados
data_path = aux[0:2] + "\\Dados\\NOAA_Climate_Data_Online\\Medias_Mensais"
os.chdir(data_path)
arcpy.env.workspace = data_path

for i in glob.glob("*.csv"):
    # Determinar o nome do parâmetro do ficheiro CSV
    var = ""
    k = 0
    while i[k] != "_":
        var = var + i[k]
        k = k + 1

    # Definir as variáveis de input e output
    inLayer = i[0:len(i)-4] + "_mod.xml"
    inData = i[0:len(i)-4] + ".shp " + var
    outLayer = i[0:len(i)-4]

    # Verificar se a extensão ArcGIS Geostatistical Analyst está disponível
    arcpy.CheckOutExtension("GeoStats")

    # Criar um layer do tipo Geostatistics
    arcpy.GACreateGeostatisticalLayer_ga(inLayer, inData, outLayer)

    # Guardar o layer criado em memória para um ficheiro lyr
    outLayerFile = i[0:len(i)-4] + ".lyr"
    arcpy.SaveToLayerFile_management(outLayer, outLayerFile, "ABSOLUTE")

    # Efectuar a cross-validation
    cvResult = arcpy.CrossValidation_ga(outLayerFile, i[0:len(i)-4] + "_val")

    # Converter a layer para grelha
    inLayer = i[0:len(i)-4] + ".lyr"
    outGrid = i[0:len(i)-4] + ".tif"
    cellSize = 1000
    cellptsHor = 1
    cellptsVer = 1
    arcpy.GALayerToGrid_ga(inLayer, outGrid, cellSize, cellptsHor, cellptsVer)

# Voltar ao diretório inicial
os.chdir(idir)

```

Anexo D

LUR.py:

```
# coding: utf-8
# Importar o módulos
import os
import glob
import math
import csv

idir = os.getcwd()
os.chdir(os.pardir) # ir para o diretório anterior
aux = os.getcwd()

# Definir diretório com os dados
data_path = aux[0:2] + "\\LUR"
os.chdir(data_path)

# Definir poluentes
pol_list = ['NO2','PM10']

for ipol in pol_list:

    # Criar ficheiro final
    fout = open(data_path + "\\LUR_" + ipol + "2003-2008.csv", "w")

    fout.write("COD;ANO;LON;LAT;" + ipol + ";ALTITUDE;DECLIVE;\
Dens_EN_250;Dens_EN_500;Dens_EN_750;\
Dens_ET_250;Dens_ET_500;Dens_ET_750;\
Dens_IC_250;Dens_IC_500;Dens_IC_750;\
Dens_IP_250;Dens_IP_500;Dens_IP_750;\
Dist_EN;Dist_ET;Dist_IC;Dist_IP;Dist_LC;\
Urb1_250;Urb1_500;Urb1_750;\
Urb2_250;Urb2_500;Urb2_750;\
Urb3_250;Urb3_500;Urb3_750;\
Urb4_250;Urb4_500;Urb4_750;\
Agua_250;Agua_500;Agua_750;\
Agro_250;Agro_500;Agro_750;\
Veg_250;Veg_500;Veg_750;\
AOT550;TEMP;RH;NDVI")

    for i in glob.glob("QualAr_" + ipol + "/*.csv"):
        dados = csv.reader(open(i, 'rb'), delimiter=',')

        for j in dados:

            lin_dados = j

            if lin_dados[0] != "COD":
```

```

# Código e coordenadas PTTM06 da estação e respetivo ano, mês e valor
medido
cod = lin_dados[0]
lon = lin_dados[1]
lat = lin_dados[2]
x = lin_dados[4]
y = lin_dados[5]
ano = lin_dados[7]
pol = lin_dados[9]

# Parâmetros
par = ["Altitude_PTTM06.tif", "Declive_Percentagem.tif",
      "Densidade_EN_250.tif", "Densidade_EN_500.tif", "Densidade_EN_750.tif",
      "Densidade_ET_250.tif", "Densidade_ET_500.tif", "Densidade_ET_750.tif",
      "Densidade_IC_250.tif", "Densidade_IC_500.tif", "Densidade_IC_750.tif",
      "Densidade_IP_250.tif", "Densidade_IP_500.tif", "Densidade_IP_750.tif",
      "Distancia_EN.tif", "Distancia_ET.tif", "Distancia_IC.tif",
      "Distancia_IP.tif", "Distancia_LC.tif",
      "PUrb1_250.tif", "PUrb1_500.tif", "PUrb1_750.tif",
      "PUrb2_250.tif", "PUrb2_500.tif", "PUrb2_750.tif",
      "PUrb3_250.tif", "PUrb3_500.tif", "PUrb3_750.tif",
      "PUrb4_250.tif", "PUrb4_500.tif", "PUrb4_750.tif",
      "PAgua_250.tif", "PAgua_500.tif", "PAgua_750.tif",
      "PAgro_250.tif", "PAgro_500.tif", "PAgro_750.tif",
      "PVeg_250.tif", "PVeg_500.tif", "PVeg_750.tif",
      "AOT550_" + i[len(i)-10:len(i)-4] + ".tif",
      "TEMP_MODIS_" + i[len(i)-10:len(i)-4] + ".tif",
      "RH_KRIGING_" + i[len(i)-10:len(i)-4] + ".tif",
      "NDVI_MODIS_" + i[len(i)-10:len(i)-4] + ".tif"]

# i[len(i)-10:len(i)-4] corresponde ao ano e ao mês do ficheiro CSV

par_val = []
for k in par:

    # Determinar nos rasters o valor do ponto de coordenadas (x,y)
    os.system("gdallocationinfo \
-geoloc " + k + " " + x + " " + y + " >" + data_path + "\\auxiliar_GDALLOCATION.txt")
    faux = open(data_path + "\\auxiliar_GDALLOCATION.txt", "r")
    text_aux = faux.readlines()
    faux.close()

    # NOTA: text_aux[3][11:len(text_aux[3])-1] contém o valor do ponto lido no
raster
    aux1 = text_aux[3][11:len(text_aux[3])-1]

    par_val.append(aux1)

# Escrever valores do ponto nos rasters no ficheiro final
fout.write("\n")
fout.write(cod + "," + ano + "," + lon + "," + lat + "," + pol + "," + "\
+ par_val[0] + "," + par_val[1] + "," + par_val[2] + "," + "\
+ par_val[3] + "," + par_val[4] + "," + par_val[5] + "," + "\
+ par_val[6] + "," + par_val[7] + "," + par_val[8] + "," + "\
+ par_val[9] + "," + par_val[10] + "," + par_val[11] + "," + "\

```

```
+ par_val[12] + " " + par_val[13] + " " + par_val[14] + " "\
+ par_val[15] + " " + par_val[16] + " " + par_val[17] + " "\
+ par_val[18] + " " + par_val[19] + " " + par_val[20] + " "\
+ par_val[21] + " " + par_val[22] + " " + par_val[23] + " "\
+ par_val[24] + " " + par_val[25] + " " + par_val[26] + " "\
+ par_val[27] + " " + par_val[28] + " " + par_val[29] + " "\
+ par_val[30] + " " + par_val[31] + " " + par_val[32] + " "\
+ par_val[33] + " " + par_val[34] + " " + par_val[35] + " "\
+ par_val[36] + " " + par_val[37] + " " + par_val[38] + " "\
+ par_val[39] + " " + par_val[40] + " " + par_val[41] + " "\
+ par_val[42] + " " + par_val[43])
```

```
fout.close()
```

```
# Voltar ao diretório inicial
os.chdir(idir)
```

LUR Rasters.py:

```

# coding: utf-8
# Importar o módulos
import os
import glob
import arcpy
from arcpy import env
from arcpy.sa import *

idir = os.getcwd()
os.chdir(os.pardir) # ir para o diretório anterior
aux = os.getcwd()

# Definir diretório com os dados
data_path = aux[0:2] + "/LUR/"
os.chdir(data_path)

# Set environment settings
env.workspace = data_path

p_list = ["NO2", "PM10"]

for poluente in p_list:

    ano_inicial = 2003
    ano_final = 2008

    for i in range(ano_inicial, ano_final + 1):

        mes_inicial = 1
        mes_final = 12

        for k in range(mes_inicial, mes_final + 1):

            ano_str = str(i)

            if k < 10:
                mes_str = "0" + str(k)
            else:
                mes_str = str(k)

            # Check out the ArcGIS Spatial Analyst extension license
            arcpy.CheckOutExtension("Spatial")

            if poluente == "PM10":
                print poluente

# MODELO PM10 -----
#####

# Densidade total de estradas num raio de 750 m -----
RD_T_750 = "roads_density_750m.tif"
RD_T_750_coef = 0.038

```

```
RD_T_750_final = Times(RD_T_750, RD_T_750_coef)
# -----
# Altitude -----
Altitude = "PT_DEM.tif"
Altitude_coef = -0.002
Altitude_final = Times(Altitude, Altitude_coef)
# -----
# Latitude -----
Latitude = "Latitude_PTTM06.tif"
Latitude_coef = 0.297
Latitude_final = Times(Latitude, Latitude_coef)
# -----
# Humidade Relativa -----
RH = "RH_" + ano_str + mes_str + ".tif"
RH_coef = -0.030
RH_final = Times(RH, RH_coef)
# -----
# Percentagem de ocupação do solo do tipo Agua num raio de 750 m -----
LC_Agua_750 = "FAgua_750m.tif"
LC_Agua_750_coef = 0.027
LC_Agua_750_final = Times(LC_Agua_750, LC_Agua_750_coef)
# -----
# AOT -----
AOT = "SQRT_AOT_" + ano_str + mes_str + ".tif"
AOT_coef = 3.072
AOT_final = Times(AOT, AOT_coef)
# -----
# Percentagem de ocupação do solo do tipo Urb2 num raio de 250 m -----
LC_Urb2_250 = "FUrb2_250m.tif"
LC_Urb2_250_coef = 0.008
LC_Urb2_250_final = Times(LC_Urb2_250, LC_Urb2_250_coef)
# -----
# Percentagem de ocupação do solo do tipo Urb3 num raio de 750 m -----
LC_Urb3_750 = "FUrb3_750m.tif"
LC_Urb3_750_coef = 0.015
LC_Urb3_750_final = Times(LC_Urb3_750, LC_Urb3_750_coef)
# -----
# Ano -----
if i == 2006:
    ano_coef = -0.156
    print str(ano_coef) + " " + str(i)
elif i == 2007:
    ano_coef = -0.289
    print str(ano_coef) + " " + str(i)
elif i == 2008:
    ano_coef = -0.780
    print str(ano_coef) + " " + str(i)
```

```

else:
    ano_coef = 0
    print str(ano_coef) + " " + str(i)
# -----

# Ano -----
if k == 1:
    mes_coef = 1.589
    print str(mes_coef) + " " + str(k)
elif k == 2:
    mes_coef = 0.981
    print str(mes_coef) + " " + str(k)
elif k == 3:
    mes_coef = 0.685
    print str(mes_coef) + " " + str(k)
elif k == 8:
    mes_coef = 0.238
    print str(mes_coef) + " " + str(k)
elif k == 9:
    mes_coef = 0.348
    print str(mes_coef) + " " + str(k)
elif k == 10:
    mes_coef = 0.601
    print str(mes_coef) + " " + str(k)
elif k == 11:
    mes_coef = 0.869
    print str(mes_coef) + " " + str(k)
elif k == 12:
    mes_coef = 1.141
    print str(mes_coef) + " " + str(k)
else:
    mes_coef = 0
    print str(mes_coef) + " " + str(k)
# -----

# Somas -----
soma1 = Plus(-5.806, RD_T_750_final)
soma2 = Plus(soma1, Altitude_final)
soma3 = Plus(soma2, Latitude_final)
soma4 = Plus(soma3, RH_final)
soma5 = Plus(soma4, LC_Agua_750_final)
soma6 = Plus(soma5, AOT_final)
soma7 = Plus(soma6, LC_Urb2_250_final)
soma8 = Plus(soma7, LC_Urb3_750_final)
soma9 = Plus(soma8, ano_coef)
soma10 = Plus(soma9, mes_coef)
# -----

# Potência -----
PM10 = Power(soma10, 2)
# -----

# Guardar o raster final -----
PM10.save(data_path + "rasters/PM10_LUR_" + ano_str + mes_str + ".tif")
# -----

```

```
# FIM MODELO PM10 -----  
#####  
  
    elif poluente == "NO2":  
        print poluente  
# MODELO NO2 -----  
#####  
  
# Percentagem de ocupação do solo do tipo Urb1 num raio de 750 m -----  
    LC_Urb1_750 = "FUrb1_750m.tif"  
    LC_Urb1_750_coef = 0.026  
    LC_Urb1_750_final = Times(LC_Urb1_750, LC_Urb1_750_coef)  
# -----  
  
# Longitude -----  
    Longitude = "Longitude_PTTM06.tif"  
    Longitude_coef = -0.765  
    Longitude_final = Times(Longitude, Longitude_coef)  
# -----  
  
# Latitude -----  
    Latitude = "Latitude_PTTM06.tif"  
    Latitude_coef = 0.356  
    Latitude_final = Times(Latitude, Latitude_coef)  
# -----  
  
# Percentagem de ocupação do solo do tipo Urb2 num raio de 250 m -----  
    LC_Urb2_250 = "FUrb2_250m.tif"  
    LC_Urb2_250_coef = 0.020  
    LC_Urb2_250_final = Times(LC_Urb2_250, LC_Urb2_250_coef)  
# -----  
  
# Altitude -----  
    Altitude = "PT_DEM.tif"  
    Altitude_coef = -0.002  
    Altitude_final = Times(Altitude, Altitude_coef)  
# -----  
  
# Densidade total de ICs num raio de 750 m -----  
    RD_IC_750 = "roads_density_IC_750m.tif"  
    RD_IC_750_coef = -0.590  
    RD_IC_750_final = Times(RD_IC_750, RD_IC_750_coef)  
# -----  
  
# Humidade Relativa -----  
    RH = "RH_" + ano_str + mes_str + ".tif"  
    RH_coef = -0.031  
    RH_final = Times(RH, RH_coef)  
# -----  
  
# Declive do terreno -----  
    Declive = "slope.tif"  
    Declive_coef = 0.048  
    Declive_final = Times(Declive, Declive_coef)
```

```

# -----
# Densidade total de IPs num raio de 250 m -----
RD_IP_250 = "roads_density_IP_250m.tif"
RD_IP_250_coef = 1.220
RD_IP_250_final = Times(RD_IP_250, RD_IP_250_coef)
# -----

# AOT -----
AOT = "SQRT_AOT_" + ano_str + mes_str + ".tif"
AOT_coef = 1.700
AOT_final = Times(AOT, AOT_coef)
# -----

# Percentagem de ocupação do solo do tipo Agua num raio de 750 m -----
LC_Agua_750 = "FAgua_750m.tif"
LC_Agua_750_coef = 0.016
LC_Agua_750_final = Times(LC_Agua_750, LC_Agua_750_coef)
# -----

# Ano -----
if i == 2007:
    ano_coef = 0.194
    print str(ano_coef) + " " + str(i)
else:
    ano_coef = 0
    print str(ano_coef) + " " + str(i)
# -----

# Mes -----
if k == 2:
    mes_coef = -0.332
    print str(mes_coef) + " " + str(k)
elif k == 3:
    mes_coef = -0.833
    print str(mes_coef) + " " + str(k)
elif k == 4:
    mes_coef = -1.245
    print str(mes_coef) + " " + str(k)
elif k == 5:
    mes_coef = -1.567
    print str(mes_coef) + " " + str(k)
elif k == 6:
    mes_coef = -1.731
    print str(mes_coef) + " " + str(k)
elif k == 7:
    mes_coef = -1.778
    print str(mes_coef) + " " + str(k)
elif k == 8:
    mes_coef = -1.676
    print str(mes_coef) + " " + str(k)
elif k == 9:
    mes_coef = -0.924
    print str(mes_coef) + " " + str(k)
elif k == 10:

```

```
    mes_coef = -0.560
    print str(mes_coef) + " " + str(k)
else:
    mes_coef = 0
    print str(mes_coef) + " " + str(k)
# -----

# Somas -----
soma1 = Plus(-14.956, LC_Urb1_750_final)
soma2 = Plus(soma1, Longitude_final)
soma3 = Plus(soma2, Latitude_final)
soma4 = Plus(soma3, LC_Urb2_250_final)
soma5 = Plus(soma4, Altitude_final)
soma6 = Plus(soma5, RD_IC_750_final)
soma7 = Plus(soma6, RH_final)
soma8 = Plus(soma7, Declive_final)
soma9 = Plus(soma8, RD_IP_250_final)
soma10 = Plus(soma9, AOT_final)
soma11 = Plus(soma10, LC_Agua_750_final)
soma12 = Plus(soma11, ano_coef)
soma13 = Plus(soma12, mes_coef)
# -----

# Potência -----
NO2 = Power(soma13, 2)
# -----

# Guardar o raster final -----
NO2.save(data_path + "rasters/NO2_LUR" + ano_str + mes_str + ".tif")
# -----

# FIM MODELO NO2 -----
#####

# Voltar ao diretório inicial
os.chdir(idir)
```

Anexo E

var_avg.m:

```

clear all; clc; close all;
format long g;

cd ..
cd Dados_Finais\Variaveis_Ambientais

% Ler uma imagem qualquer para obter as coordenadas dos cantos
[img_aux, R, bbox] = geotiffread('TEMP_MODIS_200301.tif');
nyllcorner = bbox(1,2);
nxllcorner = bbox(1,1);

par = {'TEMP','RH','NDVI','NO2','PM10'};

for j = 1:numel(par)

    % Listar ficheiros da variável par{j}
    files = dir([par{j}, '*.tif'])
    psize = 100;
    outfile = [par{j}, '_2003-2008.asc'];

    % Declarar variáveis
    img = geotiffread(files(1).name);
    [lin col] = size(img);
    img_sum = zeros(lin,col);
    img_npixels = zeros(lin,col);
    for i = 1:numel(files)

        % Ler ficheiros das variáveis
        img = geotiffread(files(i).name);
        img = double(img);
        img(img < -999) = NaN;

        for u=1:lin
            for v=1:col
                % Somar todas as imagens. Se os pixéis corresponderem a
                % falta de dados não devem ser considerados na média
                if isnan(img(u,v)) == 0
                    img_sum(u,v) = img_sum(u,v) + img(u,v);
                    img_npixels(u,v) = img_npixels(u,v) + 1;
                end
            end
        end
    end

    % Cálculo da média
    img_avg = img_sum./img_npixels;

```

```
% Colocar valores sem dados como -9999
img_avg(isnan(img_avg) == 1) = -9999;

% Abrir e escrever o cabeçalho dos ficheiros ASCII
fileID1 = fopen(outfile,'w');
fprintf(fileID1,['ncols %d\nnrows %d\nxllcorner %f\nyllcorner %f\ncellsz ' , ...
    '%f\nndata_value -9999'],col,lin,nxllcorner,nyllcorner,psize);

% Escrever ficheiros ASCII
for p=1:lin
    fprintf(fileID1,'\n');
    fprintf(fileID1,'%0.8f\t',img_avg(p,:));
end

fclose(fileID1);

end
```

ArcGIS_ZonalStatistics.py:

```
# coding: utf-8
# Importar os modulos
import os
import glob
import arcpy
from arcpy import env
from arcpy.sa import *

idir = os.getcwd()
os.chdir(os.pardir) # ir para o diretório anterior
aux = os.getcwd()

# Definir diretório com os dados
data_path = aux[0:2] + "/Dados_Finais/Variaveis_Ambientais"
os.chdir(data_path)

# Diretório com a CAOP2011
caop_dir = idir[0:2] + "/Dados/CAOP2011/CAOP2011.shp"

# Definir o workspace
env.workspace = data_path

# Verificar se a extensão Spatial Analyst está disponível
arcpy.CheckOutExtension("Spatial")

for i in glob.glob("*2003-2008.asc"):

    # Definir parâmetros da função ZonalStatisticsAsTable
    inZoneData = caop_dir
    zoneField = "DICOFRE_D" # este código permite identificar cada concelho
    inValueRaster = i
    outTable = data_path + "/Medias_por_Concelho/" + i[0:len(i)-4] + ".dbf"

    # Executar a função ZonalStatisticsAsTable
    outZSaT = ZonalStatisticsAsTable(inZoneData, zoneField, inValueRaster,
                                     outTable, "DATA", "ALL")

# Voltar ao diretório inicial
os.chdir(idir)
```