



FACULDADE DE LETRAS
UNIVERSIDADE D
COIMBRA

Ana Cristina Silva Batista

**ELABORAÇÃO DE UM MODELO PREDITIVO DE
IDENTIFICAÇÃO DO POTENCIAL ARQUEOLÓGICO
PALEOLÍTICO NA REGIÃO DE LEIRIA
(CENTRO DE PORTUGAL)**

Dissertação de mestrado em Tecnologias de Informação Geográfica, no ramo Ambiente e Ordenamento do Território, orientada pelo Professor Doutor Luca Antonio Dimuccio e pelo Professor Doutor Rui Ferreira de Figueiredo, apresentada ao Departamento de Geografia e Turismo da Faculdade de Letras da Universidade de Coimbra

Dezembro de 2021

FACULDADE DE LETRAS

FACULDADE DE LETRAS

ELABORAÇÃO DE UM MODELO PREDITIVO DE IDENTIFICAÇÃO DO POTENCIAL ARQUEOLÓGICO PALEOLÍTICO NA REGIÃO DE LEIRIA (CENTRO DE PORTUGAL)

Ficha Técnica

Tipo de trabalho	Dissertação
Título	Elaboração de um modelo preditivo de identificação do potencial arqueológico Paleolítico na região de Leiria (Centro de Portugal)
Autor/a Orientador/a(s)	Ana Cristina Silva Batista Prof. Doutor Luca Antonio Dimuccio Prof. Doutor Rui Ferreira de Figueiredo
Júri	Presidente: Prof.^a Doutora Adélia Jesus Nobre Nunes Vogais: 1. Prof. Doutor António Alberto Teixeira Gomes 2. Prof. Doutor Luca Antonio Dimuccio
Identificação do Curso Especialidade/Ramo	2º Ciclo em Tecnologias de Informação Geográfica Ambiente e Ordenamento do Território
Data da defesa Classificação	14-12-2021 18 valores

Agradecimentos

A presente dissertação foi o culminar de dois anos de trabalho, esforço e dedicação. Esta só foi possível concluir com o apoio de algumas pessoas e é a elas que quero deixar o meu enorme reconhecimento.

Saudação aos Senhores Professores Doutores Luca Antonio Dimuccio e Rui Ferreira de Figueiredo, por terem conduzido este trabalho a bom porto, com os seus sábios conselhos, apreciações e correções na elaboração dos modelos cartográficos e na revisão dos textos.

À Doutora Cristina Gameiro, por todo o apoio, especialmente nos momentos mais difíceis, com as suas palavras de incentivo e compreensão.

Ao Senhor Professor Doutor Lúcio José Sobral da Cunha, pelo tempo dispensado, pelos seus valiosos conselhos e confiança transmitida.

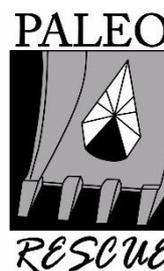
Ao colega Tiago Tamagusko, pelas explicações e dedicação aos problemas que pareciam não ter solução, nunca desistindo e acreditando até ao fim.

Os colegas Telmo Gomes e Telmo Brás, pelos debates de ideias que tanto contribuíram para a elaboração desta dissertação.

Um agradecimento especial ao Senhor Professor Nataniel Lopes, por todas as revisões feitas, dedicação, compreensão, companheirismo, motivação, conselhos e apoio, que tão importantes foram nos momentos mais difíceis.

À minha Família, em particular ao meu pai, à minha mãe, à minha irmã e ao meu cunhado, que desde o início me apoiaram para perseguir o meu sonho, incentivando-me a fazer mais e melhor.

Apoios institucionais e financeiros



Bolsa de Investigação (BI) no âmbito do Projeto PALEORESCUE (PTDC/HAR-ARQ/30779/2017), financiado pela Fundação para a Ciência e a Tecnologia, I.P. através do Orçamento de Estado - OE.



UNIVERSIDADE
DE LISBOA



LETRAS
LISBOA



CENTRO DE ARQUEOLOGIA
DA UNIVERSIDADE
DE LISBOA

uniarq

Com o apoio do Projeto CLIMATE@COA (COA/CAC/0031/2019), financiado pela Fundação para a Ciência e a Tecnologia, I.P. através do Orçamento de Estado - OE.



climate@coa

COA/CAC/0031/2019

Cofinanciado por:



UNIÃO EUROPEIA
Fundo Europeu
de Desenvolvimento Regional

1 2



9 0

FACULDADE DE LETRAS
UNIVERSIDADE DE
COIMBRA

**DEPARTAMENTO
DE GEOGRAFIA
E TURISMO**

Resumo

A presente dissertação tem como objetivo principal a realização de um conjunto de ensaios cartográficos, com o apoio dos Sistemas de Informação Geográfica (SIG), cuja finalidade se prende, fundamentalmente, com a escolha dos processos analíticos e metodológicos mais adequados para a identificação de áreas com maior ou menor potencial à preservação do registo Paleolítico, num contexto geológico/geomorfológico de tipo fluvial - o Baixo Vale do Rio Lis - na região de Leiria (Centro de Portugal). Os seguintes fatores ambientais condicionantes foram recolhidos, processados e organizados num projeto SIG: declive, índice de posição topográfica, distância ponderada de custo para acesso aos cursos de água, índice de carga de calor e litologia. A estes juntaram-se dados arqueológicos correspondentes à localização de 40 sítios paleolíticos conhecidos, até ao presente, na área de estudo. Três modelos preditivos de identificação do potencial arqueológico Paleolítico foram desenvolvidos com recurso a técnicas de Análise de Regressão (AR) e de Análise Multicritério (AMC), utilizadas individualmente, como foi o caso do Modelo Aditivo Generalizado (GAM) para o Modelo 1, ou em combinação, como foi o caso da Razão de Frequência Probabilística (RFP) e do processo de Análise Hierárquica por Pares (AHP) para os Modelos 2 e 3. Todos os resultados cartográficos produzidos para a área de estudo pelos vários métodos usados são aceitáveis, embora a combinação da RFP + AHP pareça evidenciar uma maior eficácia em comparação com o GAM. O desempenho dos três modelos foi avaliado e comparado entre si através das curvas ROC (*receiver operating characteristic*), juntamente com a métrica associada e correspondente a AUC (*area under the curve*), utilizando os dados dos sítios paleolíticos conhecidos - o Modelo 3 parece oferecer o melhor desempenho (90%) na deteção de locais com maior potencial arqueológico, seguido pelo Modelo 2 (87%) e pelo Modelo 1 (68%), que parece ser o modelo menos eficaz. O mapa mais robusto e fiável (Modelo 3), juntamente com os procedimentos analíticos e metodológicos subjacentes, pode também servir de complemento e auxílio à tomada de decisão nas políticas públicas de planeamento e ordenamento do território (à escala local/regional) possibilitando uma antevisão espacial relativamente pormenorizada do potencial arqueológico Paleolítico do território, com célere identificação dos sítios no âmbito da Arqueologia Preventiva e, permitir assim, uma redução substancial de tempo e dos custos para os promotores de obra ou mesmo durante os estudos de impacto ambiental que, por norma, precedem a realização da obra propriamente dita (seja ela pública ou privada).

Palavras-chave: Sistemas de Informação Geográfica; Modelo Aditivo Generalizado; Razão de Frequência Probabilística; Análise Hierárquica por Pares; Baixo Vale do Rio Lis; Paleolítico.

Abstract

The main objective of this dissertation is to carry out a set of cartographic results, with the support of Geographic Information Systems (GIS), related to the choice of the most appropriate analytical and methodological processes for the identification of areas with greater or lesser potential to the preservation of the Palaeolithic record, in a fluvial geological/geomorphological context - Lower Valley of the River Lis - in the Leiria region (Central Portugal). The following data about environmental factors were collected, processed, and organized into a GIS project: slope, topographical position index, cost-weighted distance to access the watercourses, heat load index, and lithology. Additionally, archaeological data corresponding to the location of 40 known Palaeolithic sites in the study area was also used. For identifying the Palaeolithic archaeological potential three predictive models were developed using Regression Analysis (RA) and Multi-Criteria Analysis (MCA) techniques, used individually, as was the case of the Generalized Additive Model (GAM) for Model 1, or in combination, as was the case of the Probabilistic Frequency Ratio (PFR) and the Analytic Hierarchy Process (AHP) for Models 2 and 3. All cartographic results produced for the study area by the various methods are acceptable, although the combination of PFR + AHP seems to evidence greater efficacy compared to the GAM. The performance of the three models was evaluated and compared to each other through ROC (receiver operating characteristic) curves, together with the metric associated with the corresponding AUC (area under the curve), using data from known Palaeolithic sites - the Model 3 seems to offer the best performance (90%) in detecting sites with the greatest archaeological potential, followed by Model 2 (87%) and by Model 1 (68%), which seems to be the least effective model. The more robust and reliable map (Model 3), together with the underlying analytical and methodological procedures, can also serve as a complement and aid to decision-making in public policies for spatial planning (at a local and regional scale) enabling a preview relatively detailed of the archaeological potential of the territory, with quick identification of the sites in the context of Preventive Archaeology and, thus, allowing a substantial reduction in time and costs for the work promoters or even during the environmental impact assessment which, by standard, precedes the implementation of the project (whether public or private).

Keywords: Geographic Information Systems; Generalized Additive Model; Probabilistic Frequency Ratio; Analytic Hierarchy Process; Lower Valley of the River Lis; Palaeolithic.

Índice geral

Agradecimentos.....	5
Apoios institucionais e financeiros	7
Resumo.....	9
Abstract	11
1 INTRODUÇÃO.....	15
1.1 Modelação espacial “preditiva”	15
1.1.1 Análise de Regressão (AR).....	18
1.1.2 Análise Multicritério (AMC).....	19
1.2 Objetivos gerais do trabalho	20
2 ÁREA DE ESTUDO.....	23
2.1 Critérios de seleção e enquadramento geográfico	23
2.2 Caraterização fisiográfica	26
2.3 Contexto arqueológico Paleolítico	30
3 MATÉRIAS E MÉTODOS.....	33
3.1 Recolha e processamento dos dados.....	34
3.1.1 Sítios paleolíticos conhecidos.....	34
3.1.2 Fatores ambientais condicionantes	36
3.2 Implementação dos modelos cartográficos preditivos.....	45
3.2.1 Modelo Aditivo Generalizado (GAM)	46
3.2.2 Razão de Frequência Probabilística (RFP)	48
3.2.3 Análise Hierárquica por Pares (AHP)	52
3.3 Avaliação da capacidade preditiva dos modelos	56
4 RESULTADOS.....	59
4.1 Modelo 1	59
4.2 Modelo 2	61
4.3 Modelo 3	63

5	DISCUSSÃO	67
5.1	Comparação entre modelos	67
5.2	Uso do mapa de identificação do potencial arqueológico Paleolítico	69
6	CONSIDERAÇÕES FINAIS	73
	BIBLIOGRAFIA	77

1 INTRODUÇÃO

Enquadrada nas tarefas definidas para o projeto PALEORESCUE - O Paleolítico Superior e a Arqueologia Preventiva em Portugal: desafios e oportunidade (PTDC/HAR-ARQ/30779/2017), e com o apoio do projeto CLIMATE@COA - Clima e adaptação humana durante o último Período Glaciar na região do Vale do Côa (Portugal) (COA/CAC/0031/2019), esta dissertação de mestrado em Tecnologias de Informação Geográfica, no ramo Ambiente e Ordenamento do Território, visa, antes de mais, consolidar esta área de formação científica e académica através do aprofundamento dos conhecimentos sobre o desenvolvimento e a aplicação de modelos cartográficos preditivos no contexto da Arqueologia.

O reconhecimento em Portugal de um conjunto de sítios pré-históricos importantes, não apenas à escala nacional, mas mesmo à escala mundial, como é (por exemplo) o caso das gravuras rupestres do Vale do Rio Côa no nordeste do país (Zilhão, 1996; Aubry, 2009), durante as diversas fases dos trabalhos da Arqueologia de Salvamento (ou, também, designada de “Arqueologia Preventiva” - *sensu* Silva, 2003; Gonçalves, 2012), justifica a necessidade de se aprofundarem os conhecimentos e os métodos para a implementação de um processo de avaliação empírico e objetivo que permita produzir representações cartográficas, em diversos contextos geológicos/geomorfológicos (fluvial, cársico, fluvio-cársico, eólico, planalto, etc.), que identifiquem e definam setores espaciais com maior ou menor potencial à preservação do registo arqueológico, com óbvias vantagens práticas para as equipas técnicas que, no campo, seguem as prospeções e, eventualmente, realizam as escavações (Aubry *et al.*, 2017a; Gameiro e Dimuccio, 2019). No âmbito específico desta dissertação, e pelas razões explicitadas mais adiante, privilegia-se como período temporal de referência o Plistocénico e, portanto, a ocupação humana do Paleolítico, em contexto fluvial.

1.1 Modelação espacial “preditiva”

No que diz respeito a modelação espacial preditiva, historicamente, esta surge como [...] *“a technique that seeks to predict the probability of encountering a phenomenon in unsampled areas based on knowledge gained from sampled areas, may provide insight into a suspected pattern”* (Judge e Sebastian, 1988; Kvamme, 1990; Conolly e Lake, 2006). Largamente utilizada em geografia e geologia desde a década de 1970, as aplicações deste tipos de técnicas na arqueologia e na gestão dos recursos naturais têm vindo a ganhar cada vez mais adeptos (Kohler

e Parker, 1986; Warren, 1990; Wescott e Brandon, 2000; Wheatley e Gillings, 2002; Warren e Asch, 2000, 2005; Mehrer e Wescott, 2006; Kvamme, 2006; Mark *et al.*, 2006; Verhagen, 2007; Verhagen e Whitley, 2012; Balla *et al.*, 2014; Wachtel *et al.*, 2018; entre outros), sendo a sua integração com os Sistemas de Informação Geográfica (SIG) e a Detecção Remota uma das abordagens mais frequentes e valorizadas na arqueologia quantitativa digital (Espa *et al.*, 2006; Lasaponara e Masini, 2006, 2013; Siart *et al.*, 2008; Noviello e al., 2018; Kempf, 2019; Parow-Souchon *et al.*, 2021; e referências neles incluídas). A implementação dos modelos preditivos em arqueologia pode ser realizada considerando duas estratégias diferentes: (a) baseando-se no prévio conhecimento da teoria do comportamento humano que explica a distribuição dos dados arqueológicos, sendo assim de caráter dedutivo (Verhagen *et al.*, 2007); (b) baseando-se em observações empíricas que procuram determinar a correlação entre os sítios arqueológicos de referência conhecidos e os seus atributos ambientais e paisagísticos, ou seja, num processo indutivo (Wheatley e Gillings, 2002). Na realidade, estas abordagens não são mutuamente exclusivas (Verhagen, 2007; Wheatley e Gillings, 2002), sendo que, nesta investigação, segue-se exclusivamente uma aproximação indutiva.

Importa referir que a aplicação da modelação espacial preditiva em arqueologia é ainda controversa (Wheatley e Gillings, 2002; Kamermans, 2010; Grøn, 2018), sobretudo tendo em conta às incertezas que estão quase sempre subjacentes à escolha dos dados arqueológicos a serem utilizados nos procedimentos analíticos. Algumas destas incertezas relacionam-se com: (i) a subestimação dos números de sítios arqueológicos conhecidos que, muitas vezes, não reflete o quadro real dos assentamentos humanos passados, estando a sua representatividade atual dependente da evolução geomorfológica e paleoclimática da área de estudo (Hancock *et al.*, 2011); (ii) a tipologia e intensidade das prospeções e eventuais escavações arqueológicas conduzidas na área de estudo (Wheatley, 2004); e (iii) a visibilidade arqueológica dos diferentes tipos de assentamentos (Cavulli e Grimaldi, 2007).

Seja como for, a modelação espacial preditiva foi e continua a ser utilizada com sucesso em muitas regiões do mundo, e em diferentes contextos crono-culturais (Vaugh e Crawford, 2009; Sánchez *et al.*, 2017; Carleton *et al.*, 2017; Caracausi *et al.*, 2018; Zhu *et al.*, 2018; Kempf, 2019; Klehm *et al.*, 2019; Nicu *et al.*, 2019; Parow-Souchon *et al.*, 2021; entre outros), onde as técnicas estatísticas mais usadas para definir a importância relativa das variáveis de entrada envolvem ambas as aproximações qualitativas e quantitativas, tais como, por exemplo: a Regressão Logística (Wescott e Brandon, 2000; Charlton *et al.*, 2002; Wachtel *et al.*, 2018), a Comparação Hierárquica por Pares ou Processo Hierárquico Analítico (Nsanziyera *et al.*, 2018), o método da Máxima Entropia (Noviello *et al.*, 2018; Wachtel *et al.*, 2018) e a Lógica Fuzzy (Mokarram e Aminzadeh, 2010; Banerjee *et al.*, 2018).

Em Portugal, os poucos trabalhos desenvolvidos no âmbito da elaboração de modelos preditivos no contexto da arqueologia dizem respeito à determinação da potencial localização de Vilas Romanas, aplicando o método da Adição Binária (Rua, 2004, 2009), às estações da Idade do Ferro localizadas no Noroeste, onde se utiliza a Regressão Logística (Costa, 2009), aos povoados proto-histórico da Beira Interior, mais uma vez com recurso à Regressão Logística (Estanqueiro, 2016, 2017), ao Mesolítico do Tejo (Gonçalves, 2009, 2011) e entre os vales do Tejo e do Sado (Gonçalves, 2014), com a Adição Binária, ao Paleolítico Médio do território nacional, com a ajuda dos modelos de Nichos Eco-culturais (Manuel *et al.*, 2014; Silva, 2015), ao Paleolítico da região costeira da Estremadura (Sims, 2014) e ao Paleolítico Médio do Barlavento Algarvio, onde, mais uma vez, volta a ser utilizado o método da Adição Binária (Maio, 2018). No conjunto, trata-se de investigações académicas (dissertações e teses) de cariz essencialmente metodológico.

Nota-se, ainda, que a elaboração de todos estes trabalhos, nacionais e internacionais, tiveram sempre subjacente uma abordagem científico-metodológica em que a arqueologia, a geografia e a geologia se integram e se complementam na tentativa de obter uma visão sistémica (*sensu* Bertrand, 1982; Rocha e Morgado, 2007; Mendes e Dimuccio, 2013; e referências neles incluídas), e onde os SIG representam o elemento aglutinador, permitindo a integração e a interação de dados espacializados provenientes de cada uma destas disciplinas na leitura dos territórios (Fig. 1) (Soares *et al.*, 2010). De facto, todas estas investigações utilizam, de uma forma mais ou menos aprofundada, um sistema espacial de apoio à decisão (*sensu* Ferreira, 2001), incorporando alguns métodos das análises de regressão e multicritério (amplamente descritos mais adiante), no contexto operacional dos SIG.

Para contextualizar e suportar a investigação que se pretende realizar, em particular da possibilidade de utilizar um leque diversificado de variáveis dependentes das condições físicas do território (altimetria, hidrografia, declive, exposição das vertentes, litologia, etc.) na modelação espacial preditiva aplicada à arqueologia paleolítica, importa referir três pressupostos básicos e fundamentais que devem ser considerados: (1) a distribuição dos sítios arqueológicos na paisagem não é aleatória, mas antes, reflete uma escolha cultural e é influenciada pelas condições naturais (topográficas, hidrológicas, geológicas, geomorfológicas, edáficas e climáticas) e de disponibilidade/acesso aos recursos essenciais à vida no período temporal de referência; (2) a preservação dos depósitos arqueológicos de cronologia pré-histórica é essencialmente controlada pelos processos naturais; (3) os fatores ambientais que condicionaram a fixação das comunidades e a eventual preservação dos vestígios estão ainda presentes na paisagem atual, podendo ser identificados, classificados e quantificados utilizando diversos métodos de recolha de dados, tais como os levantamentos de pormenor no campo, a

cartografia já publicada, a teledeteção ou outras fontes de informação georreferenciada, para depois serem organizados e processados utilizando os SIG.

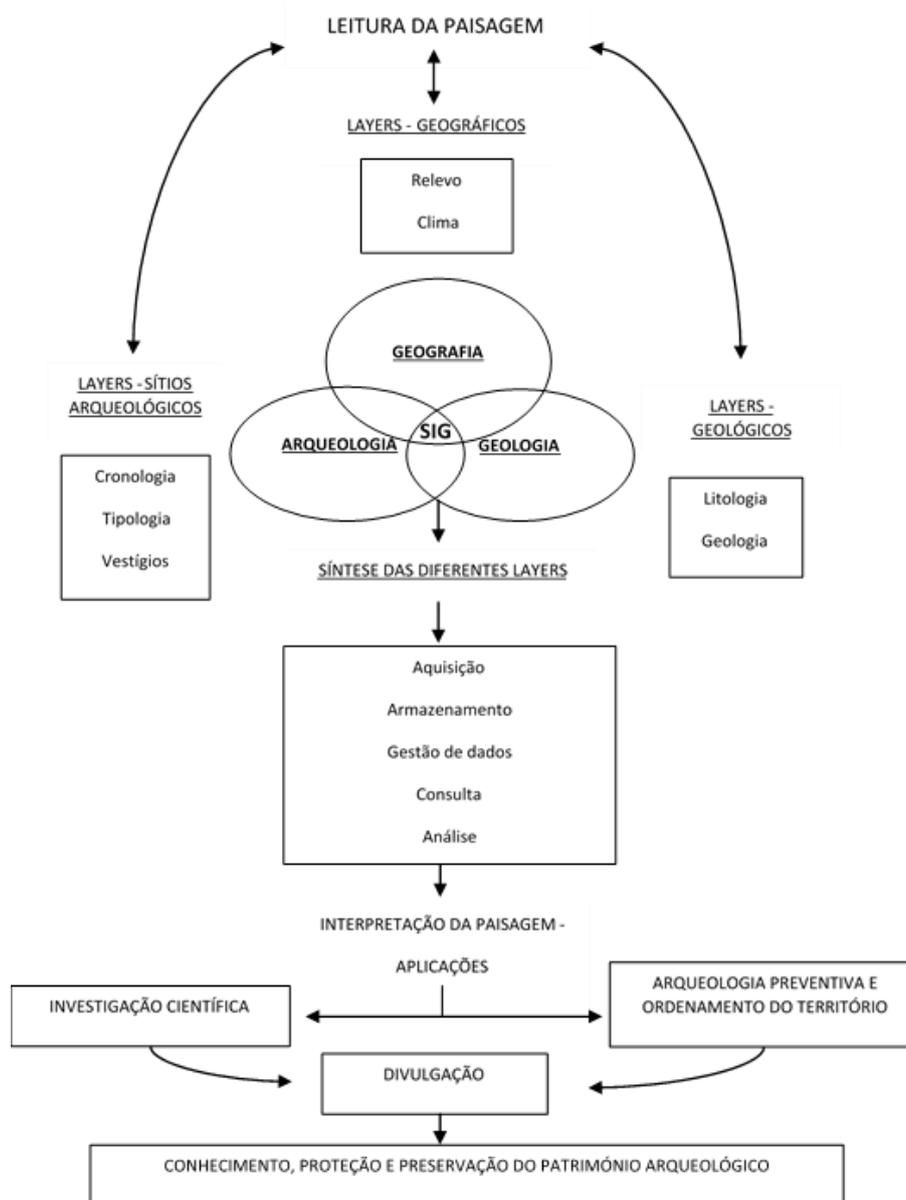


Figura 1 - Esquematização conceitual sobre a articulação entre os conhecimentos geográficos, geológicos e arqueológicos adquiridos na leitura e interpretação dos territórios, integrados entre si através dos SIG (adaptado de Soares *et al.*, 2010).

1.1.1 Análise de Regressão (AR)

A “Análise de Regressão” (AR) corresponde a um amplo conjunto de técnicas estatísticas usadas para estimar as relações entre variáveis e prever o valor da variável dependente (ou de

resposta) em função de um conjunto de variáveis independentes (ou preditores) (Maroco, 2003). Neste contexto, o modelo de regressão linear normal e o modelo logístico linear assumem uma variação linear (ou outra forma paramétrica) das covariáveis e, pela sua simplicidade, são os mais frequentemente utilizados (Hosmer e Lemeshow, 2000; Tolles, 2016).

Neste trabalho, contudo, optou-se por usar, como técnica da análise de regressão, o Modelo Aditivo Generalizado (GAM) (Hastie e Tibshirani, 1986; Wood, 2017), que é um caso particular (não paramétrico) dos Modelos Lineares Generalizados (MLG), desenvolvidos com o intuito de se ultrapassarem as limitações dos modelos lineares de regressão, nomeadamente, através da expansão à família exponencial das opções de distribuição de probabilidades de erros associados ao modelo. A distribuição Normal, Beta normal inversa, Binomial, Binomial negativa, Poisson, etc. são frequentemente utilizadas no ajuste dos MLG (Nelder e Wedderburn, 1972; Madsen e Thyregod, 2010).

Os métodos GAM apresentam como principal vantagem a flexibilidade do processo de ajustamento aos dados de base. De facto, nos GAM, a soma das funções de suavização das covariáveis é não-paramétrica, permitindo que sejam os próprios dados de base a guiarem a relação com o preditor que, frequentemente, se apresenta como não-linear. Assim, a curva alisada do modelo permite descrever de uma forma mais ajustada os dados reais, incluindo as não-linearidades eventualmente existentes nas relações estudadas, já que não está sujeita à rigidez da função paramétrica. De notar, no entanto, que este processo de suavização apenas se aplica às covariáveis contínuas (Hastie e Tibshirani, 1986; Wood, 2017).

1.1.2 Análise Multicritério (AMC)

Em sentido amplo, e de acordo com Ferreira (2001), a “Análise multicritério” (AMC) pode ser considerada [...] *“uma técnica orientada para a resolução de problemas que envolvem múltiplas alternativas ou acções, avaliadas através de múltiplos critérios, por vezes conflitivos ou dificilmente mensuráveis, visando apoiar o agente decisor na sua tarefa árdua de compreender o contexto dos dados que compõem o seu problema, para assim poder chegar a uma solução de compromisso ou hierarquizar alternativas de acordo com a sua valoração.”* Nesta tipologia de análise, existem vários métodos estatísticos, mais ou menos sistematizados, para se definir as prioridades assumidas por cada um dos critérios, destacando os de ordenação, de proporção e de comparação por pares (Voogd, 1983; Malczewski, 1999; Saaty, 1988).

Ainda que as técnicas multicritérios tenham sido inicialmente desenvolvidas num contexto não espacial, a sua integração no SIG permite expandir significativamente a sua utilidade, e âmbito de aplicação, ao incorporarem as variações espaciais decorrentes da heterogeneidade geográfica inerente à realidade do mundo real (Roy, 1996; Triantaphyllou *et al.*, 1998; Malczewski, 1999, 2004, 2006; Chaktar e Mosseau, 2007; Mokarram e Aminzadeh, 2010; Munier, 2011; Malczewski e Rinner, 2015; entre outros). Esta dimensão espacial é, de facto, a característica que melhor define o conceito de SIG, ainda que o seu espetro de abrangência seja bastante amplo e diversificado e as múltiplas definições existentes reflitam, essencialmente, as perspetivas e as experiências de quem as apresenta (mais detalhes em Ferreira, 2001). Assim, um SIG pode ser visto como uma ferramenta composta por uma determinada conjugação de *hardware* e *software* que permite, para além das operações básicas de recolha, armazenamento, verificação e visualização, a implementação de diversas metodologias de análise de dados que estão espacialmente referenciados à superfície terrestre (*e.g.*, Burrough e McDonnell, 1998; Parker e Asencio, 2008; Lloyd, 2010; Wheeler *et al.*, 2010; Martinelli, 2011).

A convergência (aqui explorada) entre as potencialidades dos SIG e alguns dos métodos geoestatísticos (*i.e.*, técnicas estatísticas facilmente implementáveis num contexto espacial) de AMC, tais como a Razão de Frequência Probabilística (RFP) (Cunha e Dimuccio, 2002; Dimuccio *et al.*, 2011) e o processo de Análise Hierárquica por Pares (AHP) (Saaty, 1988; Ferreira, 2001; Ramos *et al.*, 2014), torna-se uma abordagem valiosa para resumir os processos de decisão que envolvem múltiplos fatores espacializados a serem ponderados quantitativamente, de acordo com a sua percebida importância (Carver, 1991; Malczewski, 2006; Lloyd, 2010), permitindo, no final, a elaboração de modelos cartográficos explicativos do fenómeno em análise - como, por exemplo, o potencial à preservação do registo arqueológico.

1.2 Objetivos gerais do trabalho

O principal objetivo deste trabalho prende-se com a exploração dos processos analíticos e metodológicos mais adequados para a identificação de áreas com maior ou menor potencial à preservação do registo Paleolítico da região de Leiria (Centro de Portugal). De forma a minimizar as incertezas nos processos de modelação espacial, e tendo em conta as características locais de implantação dos sítios paleolíticos conhecidos, o estudo centrar-se-á numa unidade

geográfica e geomorfológica bem distinta e individualizada na paisagem atual - o Baixo Vale do Rio Lis.

Neste contexto, serão também evidenciados os diversos aspetos críticos dos modelos e dos resultados cartográficos por eles produzidos, sobretudo no que diz respeito aos métodos e técnicas geoestatísticas utilizadas, singularmente (como é o caso do GAM) ou em combinação (RFP + AHP), assim como a aplicabilidade (ou não) dos mesmos procedimentos para a concretização de modelos preditivos em contextos geográficos e geomorfológicos diversos da área de estudo (*e.g.*, grutas/abrigos em terrenos cársicos e/ou fluvio-cársicos, dunas na zona litoral, planaltos em substratos rochosos).

Entre os resultados cartográficos produzidos dar-se-á destaque ao protótipo (mapa) que parece oferecer os resultados mais robustos e fiáveis para a área de estudo, com o intuito de identificar aquelas áreas potencialmente mais adequadas para a continuação das atividades de prospeção arqueológica (eventualmente escavações) e assim melhor direcionar as atividades no campo. Com base nestes resultados, e num esforço indutivo, proceder-se-á, depois, à discussão das suas possíveis aplicações práticas no contexto da Arqueologia Preventiva, assim como no panorama mais geral das políticas públicas de Planeamento e Ordenamento do Território, com particular referência para a delimitação de áreas de salvaguarda nos Planos Diretores Municipais.

2 ÁREA DE ESTUDO

2.1 Critérios de seleção e enquadramento geográfico

Num estudo científico-metodológico desta natureza, a presença de características físicas específicas (topográficas, hidrográficas, geológicas e geomorfológicas) e humanas (arqueológicas) dos territórios, em detrimento de outras, juntamente com a disponibilidade (ou não) da informação espacial de base que as permita descrever, classificar, quantificar e integrar em SIG, devem ser previamente consideradas na escolha da área de estudo e na definição dos seus limites, tornando-se assim numa tarefa desafiante, mas, ao mesmo tempo, facilmente criticável. No entanto, tal escolha deve ser feita de forma ponderada e conscienciosa, sempre balanceando estes dois aspetos fulcrais que, direta ou indiretamente, têm a capacidade intrínseca de condicionar (positivamente ou negativamente) os procedimentos analíticos e, desta forma, influenciar as escolhas metodológicas e os resultados finais da modelação espacial que se pretende realizar.

Para além disso, considerando mais uma vez os objetivos previamente definidos neste trabalho, a escolha da área de estudo deve necessariamente compreender um grau de heterogeneidade espacial tão reduzido quanto possível, sobretudo em termos geológicos e geomorfológicos. Tal pressuposto é essencial, na medida em que condiciona a escolha das variáveis a incluir nos procedimentos analíticos. Em ambientes geológicos/geomorfológicos diversos, tais como o cársico vs. o fluvial vs. o litoral (todos estes presentes na região de Leiria), a preservação do registo Paleolítico, no espaço e no tempo, pode ser condicionada por fatores distintos ou com um peso de importância relativa entre eles bastante variável. A eventual preferência para uma área de estudo caracterizada por um território muito heterogéneo, onde dentro dos seus limites se passa de um ambiente geológico/geomorfológico a outro, implicaria necessariamente um aumento exponencial da complexidade (entropia) do geossistema em análise e, em consequência, uma maior incerteza relativa à tomada de decisão nas diferentes fases processuais dos procedimentos de análise espacial a implementar.

Assim sendo, e considerando os critérios de seleção aqui definidos, foi escolhida como área de estudo um espaço geográfico de aproximadamente 357 km², correspondendo, *grosso modo*, ao Baixo Vale do Rio Lis na região de Leiria (Centro de Portugal). Trata-se de um território com características predominantemente fluviais (*i.e.*, pouco heterogéneo, de um ponto de vista geológico/geomorfológico) e localizado a Sul e a Oeste das colinas, serras e planícies aluviais que definem o sector meridional do Baixo Mondego (Almeida *et al.*, 1990; Cunha, 1990; Cunha *et*

al., 2020), a Norte das serras e planaltos essencialmente carbonatados e carsificados do Maciço Calcário Estremenho (Martins, 1949; Crispim, 1995; Rodrigues, 1998), a Leste da faixa litoral atlântica (Dinis, 1996; Gonçalves, 2007) (Fig. 2).

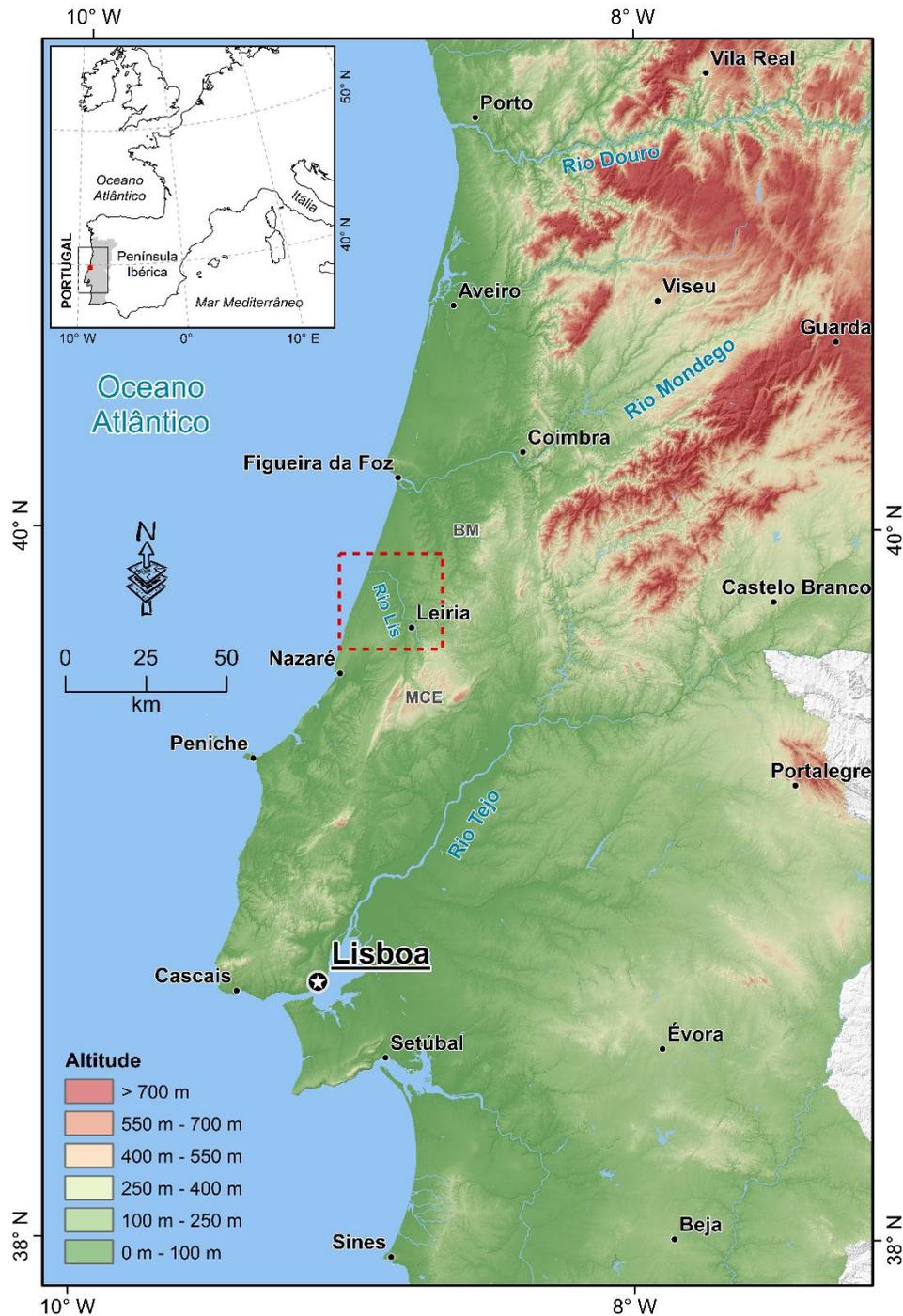


Figura 2 - Enquadramento geográfico regional da região de Leiria (Centro de Portugal), onde se insere o Baixo Vale do Rio Lis (no interior do polígono a traço descontinuo vermelho). BM = Baixo Mondego; MCE = Maciço Calcário Estremenho. O modelo digital de elevação de Portugal, com uma resolução espacial de 10 m x 10 m de *pixel*, foi construído utilizando curvas de nível, pontos cotados e linhas de águas vetorizados a partir das Cartas Militares de Portugal em escala 1:25 000 (Série M888; IGeoE, 2003).

A área selecionada inscreve-se nas folhas n.º 272, 273, 284, 285, 296 e 297 da Carta Militar de Portugal na escala 1:25 000 do Instituto Geográfico do Exército (Série M888; IGeoE, 2003), abrangendo parte dos concelhos de Pombal, Leiria, Marinha Grande e Batalha (Fig. 3).



Figura 3 – Localização geográfica da área de estudo: Baixo Vale do Rio Lis na região de Leiria (Centro de Portugal). A base cartográfica administrativa (à escala dos concelhos - em castanho), e viária, foi retirada de <http://qoto.arcgisonline.com/maps/ESRI Street Map World2D> (último acesso em junho 2021) (Copyright:© 2009 ESRI, AND, TANA, ESRI Japan, UNEP-WCMC).

2.2 Caracterização fisiográfica

Toda a bacia hidrográfica do Rio Lis, onde se inclui também a porção montante mais meridional e oriental, fora da área de estudo, é limitada a NE pela bacia hidrográfica do Rio Mondego, a SE pela bacia hidrográfica do Rio Tejo e a NW e SW por pequenas bacias litorais. Esta bacia hidrográfica estende-se por 831 Km², se tivermos apenas em consideração o escoamento superficial, ou 915 Km² se também incluirmos a drenagem sub-superficial associada ao Maciço Calcário Estremenho (Cunha *et al.*, 1980; Almeida *et al.*, 1989; Dinis, 1996; Carvalho, 2002) que se ergue na porção de montante (Fig. 2).

O clima caracteriza-se por uma amplitude térmica anual reduzida e regular (Daveau, 1985), com a temperatura média mensal inferior ou igual aos 20°C (em agosto), enquanto a mínima se fica pelos ~10°C (em janeiro) (Fig. 4). A pluviosidade, com uma média anual das precipitações na ordem dos 935 mm, segue um regime diferenciado que permite a ocorrência de uma estação seca pouco acentuada, entre junho e setembro. Segundo a classificação de Köppen, este é do subtipo Csb (clima mediterrâneo de verão quente) caracterizado por um verão quente e suave e um inverno chuvoso, mas com uma influência atlântica (Ribeiro, 1945), já que a região se situa próximo da costa.

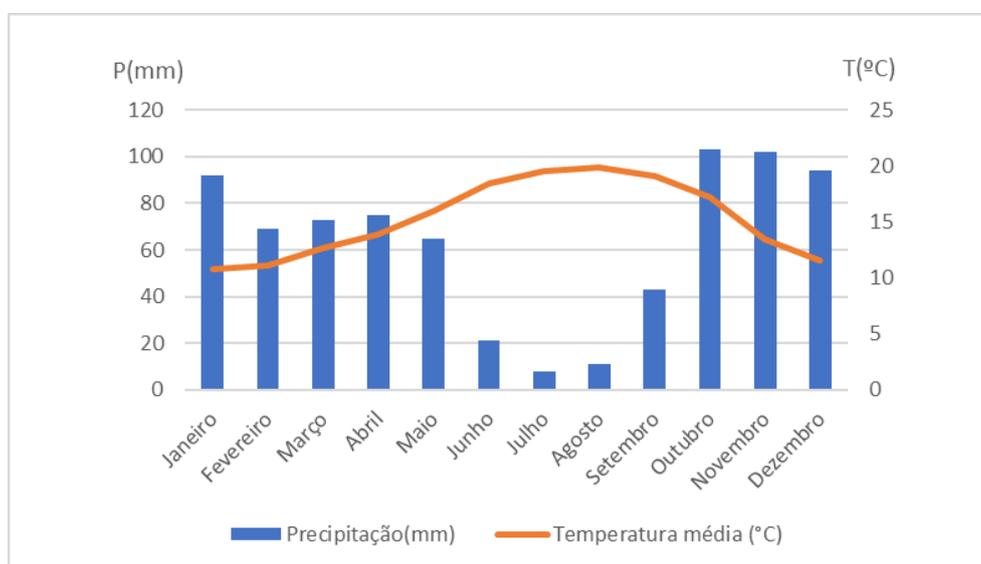


Figura 4 - Diagrama termopluviométrico da estação de Monte Real (intervalo temporal 1960 - 1990). Os dados foram retirados do <https://www.ipma.pt/pt/index.html> (último acesso em abril 2021).

Em termos geológicos, a região de Leiria caracteriza-se pelo afloramento de unidades litostratigráficas bem distintas que vão dos pelitos gipsíferos do Triásico-Hetangiano até às areias, conglomerados e argilas do Quaternário, passando pelas rochas essencialmente

carbonatadas do Jurássico e dos arenitos e argilas do Cretácico, do Paleogénico e do Neogénico (Fig. 5), situando-se integralmente na Orla Meso-Cenozoica Ocidental Portuguesa (Azerêdo *et al.*, 2003; Callapez *et al.*, 2010; Pais *et al.* 2012; Kullberg *et al.*, 2013; e referências neles incluídas).

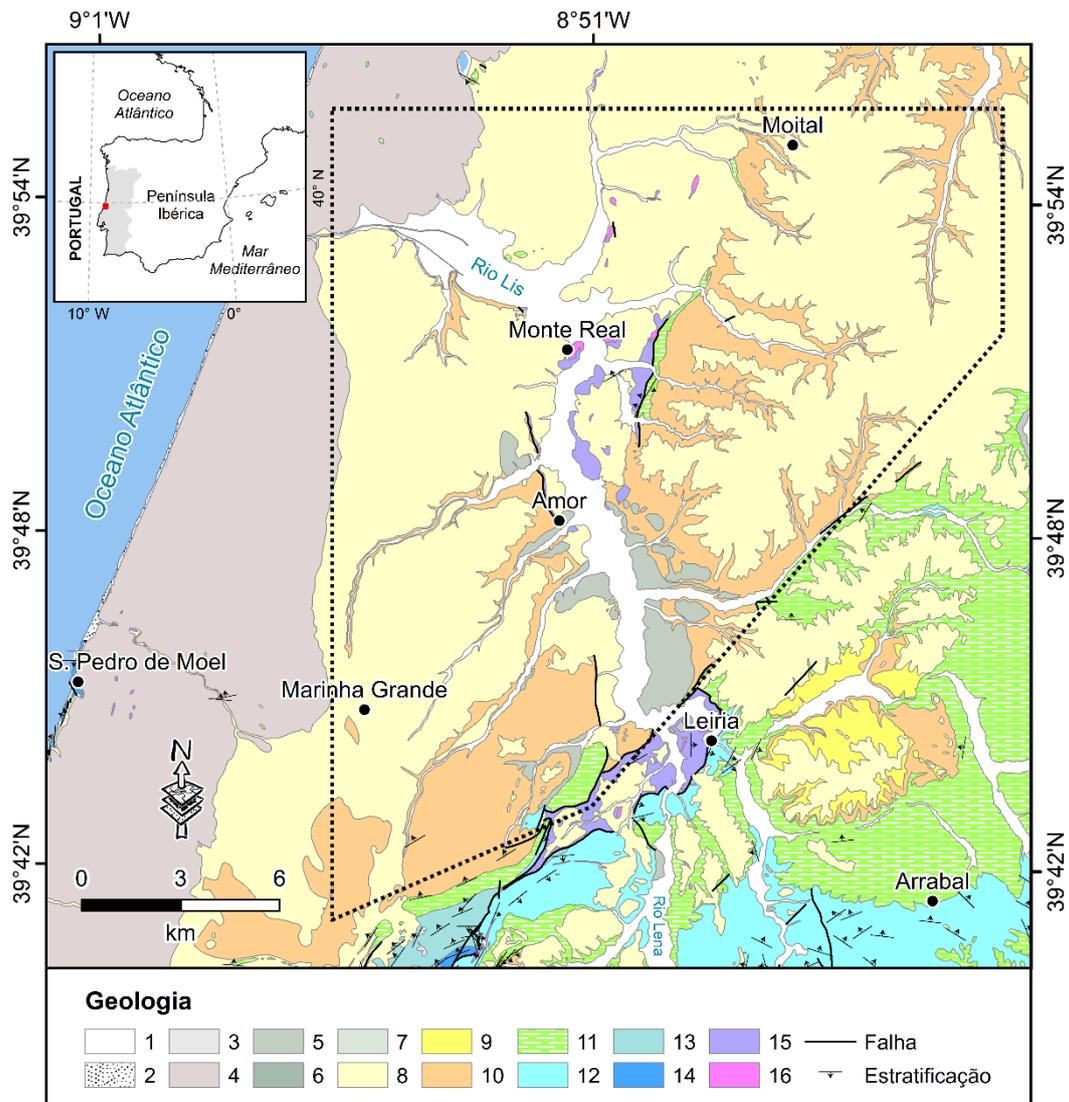


Figura 5 - Mapa geológico de enquadramento da área de estudo (polígono a traço descontinuo preto). Base cartográfica de acordo com o Mapa Geológico de Portugal, Folha 23-C, na escala 1:50 000, e respetiva Notícia Explicativa (Teixeira *et al.*, 1968). (1) Aluviões - Holocénico; (2) Areias de praia - Holocénico; (3) Depósitos de vertente - Holocénico; (4) Dunas e areias eólicas - Plistocénico a Holocénico; (5) Depósitos de terraços fluviais - Plistocénico; (6) Tufos calcários - Plistocénico; (7) Depósitos de terraços marinhos - Plistocénico; (8) Areias, cascalheiras e argilas - Pliocénico; (9) Arenitos, arenitos argilosos, conglomerados e argilas - Miocénico; (10) Argilas, margas, arenitos argilosos e conglomerados - Paleogénico indiferenciado; (11) Arenitos, conglomerados e argilitos com calcários margosos, calcários e tufo vulcânico - Cretácico; (12) Argilas e arenitos, calcários margosos e conglomerados, calcários e margas - Jurássico Superior; (13) Calcários, calcários margosos e margas - Jurássico Médio; (14) Calcários margosos e margas, calcários, calcários dolomíticos - Jurássico Inferior; (15) Margas evaporíticas e calcários, calcários dolomíticos - Triássico-Hetangiano; (16) Filões e massas de andesitos, doleritos e basaltos.

A porção central e oriental da área de estudo caracteriza-se, essencialmente, pelo afloramento das unidades siliciclásticas do Cretácico, Paleogénico e Neogénico (Dinis, 1999, 2001; Callapez *et al.*, 2010; Dinis *et al.*, 2008; Pais *et al.*, 2012), onde a atividade diapírica, ligada aos afloramentos dispersos da Formação de Dagorda do Triásico-Hetangiano (Zbyszewski, 1959; Palain, 1976; Soares *et al.*, 1986, 2012), determina a presença de domos e elevações morfológicas, como o morro do Castelo de Leiria, que se salientam numa paisagem essencialmente fluvial (Fig. 5). Corpos ígneos aparecem próximos dos diapiros de Monte Real e Leiria (Ferreira & Macedo, 1983). No limite meridional da área de estudo, as unidades carbonatadas do Jurássico Inferior e Médio, juntamente com as estruturas tectónicas que as organizam (Azerêdo *et al.*, 2003; Azerêdo, 2007; Carvalho, 2013, 2018), definem parte do rebordo ocidental e mais setentrional do Maciço Calcário Estremenho, caracterizado pela presença de uma típica paisagem cársica onde se podem encontrar lúpias, dolinas, uvalas, fôrnea (= reculé = anfiteatro natural), abrigos sob rocha e espetaculares cavidades subterrâneas (Martins, 1949; Ferreira *et al.*, 1988; Crispim, 1995; Rodrigues, 1998, 2020; e referências neles incluídas). Vários níveis de terraços fluviais do Plistocénico, muitas vezes mascarados por uma cobertura de coluviões e areias eólicas, são particularmente evidentes na margem direita do Rio Lis, sendo que as aluviões holocénicas estão associadas ao fundo do vale principal e aos seus tributários mais desenvolvidos. As areias de praia e os depósitos dunares encontram-se na faixa litoral e próximo da foz do rio (Teixeira *et al.*, 1968; Carvalho, 1972; Almeida *et al.*, 1989; Dinis, 1996; Gonçalves, 2007).

As estruturas tectónicas são essencialmente fraturas (falhas e diaclases) que afetam todas as unidades líticas meso-cenozoicas, e alguns dobramentos como o sinclinal de Pousos no extremo meridional e oriental do mapa geológico na Figura 5 (já fora da área de estudo). Observam-se famílias de falhas com orientação preferencial NE-SW a NNE-SSW, subordinadamente NW-SE e N-S. Todas elas resultam da reativação recorrente, durante a orogenia Alpina (incluindo as fases pirenaica e bética), de fraturas no soco varisco, sendo coerentes com algumas das mais importantes direções de fraturação tardi-hercínicas (Ribeiro *et al.*, 1979; Ribeiro, 2013). A conjugação daquelas famílias de fraturação confere à área uma compartimentação interna em mosaico, com abatimento dos blocos mais setentrionais. No limite meridional da área de estudo (Fig. 5), a direção NE-SW está bem marcada pela presença de um dos acidentes tectónicos mais importantes da geologia portuguesa - a “Falha de Nazaré” - considerado de idade tardi-hercínica e correspondente a uma flexura em compressão, com levantamento do bordo SE até Leiria (Ribeiro *et al.*, 1996; Kullberg *et al.*, 2013). Esta estrutura tectónica de significado regional apresenta a ocorrência de deslocamentos plio-pleistocénicos com uma importante componente de movimentação vertical - com descida do sector de NW (Cabral,

1995). Com muita probabilidade, na área de estudo, o mesmo traçado meridiano a sub-meridiano do Rio Lis será o reflexo superficial de uma estrutura tectónica com esta mesma orientação N-S, dando lugar ao extenso corpo aluvionar da sua zona central (Fig. 5).

Em termos morfológicos, e uma vez que a bacia hidrográfica do Rio Lis ocupa apenas terrenos litorais, raramente se atingem cotas superiores aos 400 m sobre o nível médio do mar. Verifica-se que 66.7% da sua área está abaixo da cota de 150 m, 78.3% abaixo de 200 m e apenas 3.6% ultrapassa a cota de 400 m. Esta situação deve estar relacionada com a grande extensão dos setores aplanados, englobados na plataforma litoral, que se posicionam entre as cotas de 50 e 150 metros (Dinis, 1996). Os terrenos mais elevados situam-se no Maciço Calcário Estremenho, onde se atinge a cota de 589 m. Ao avançar para Norte observa-se uma série de colinas progressivamente mais baixas, onde predominam cotas entre os 100 m e os 300 m. Para Oeste e Noroeste temos terrenos geralmente a cotas inferiores a 100 m, aplanados, que se relacionam com uma antiga superfície de abrasão marinha, englobada na plataforma litoral e apenas recortada por uma rede de drenagem suavemente encaixada e pelos fenómenos diapíricos de Leiria e Monte Real. Outros terrenos aplanados, a uma cota geralmente inferior aos 30 m, são os campos aluvionares de extensão variável. Verifica-se uma assimetria do relevo em torno dum eixo definido pelo Rio Lis até à região de Monte Real. Os afluentes da margem direita apresentam um perfil longitudinal mais inclinado, posicionando-se numa área marcada por um relevo mais acentuado do que os da margem esquerda.

O Rio Lis, com desenvolvimento principal N-S (assim como já se referiu), rodando para NW-SE a W-E na sua porção final onde desagua no Oceano Atlântico (junto à Praia da Vieira), tem como principal afluente o Rio Lena e como tributários secundários, na margem esquerda, as Ribeiras de Escoura, de Amor, do Fagundo e do Picheleiro; na margem direita as Ribeiras da Caranguejeira e dos Milagres (Fig. 5). No Maciço Calcário Estremenho, esta rede de drenagem superficial faz-se em vales estreitos e encaixados por intermédio de muitas linhas de água de reduzida extensão, em configuração dendrítica, embora com alguma tendência quadrangular e “*trellis*”. Mais para jusante (*i.e.*, na área de estudo deste trabalho), os cursos de água passam a ocupar vales de largura variável, da ordem das centenas de metros, mantendo o carácter dendrítico, mas por vezes semelhante ao tipo “*trellis*”. De facto, a jusante de Leiria, e após a confluência com o Lena, o vale do Lis alarga-se significativamente. Este alargamento é acompanhado de uma ligeira quebra do declive no perfil longitudinal marcando-se o início da zona baixa do rio. Posteriormente, após atravessar os corpos ígneos da zona de Monte Real, atingem-se as maiores larguras, com uma extensão de cerca de dois quilómetros (Fig. 5).

2.3 Contexto arqueológico Paleolítico

Até os finais do século passado, o conhecimento científico adquirido sobre a arqueologia paleolítica da região de Leiria não tinha uma grande expressão na literatura da especialidade (*vide* Zilhão, 1997; Cunha-Ribeiro, 1992, 1999, 2002; Almeida, 2000; Raposo, 2002), pelo menos até à fortuita descoberta do “Menino do Lapedo” no Algar Velho (Zilhão e Trinkaus, 2002; Almeida *et al.*, 2002; Angelucci, 2003; Almeida, 2005; Carvalho, 2011), despertando assim a curiosidade dos especialistas que intensificaram e alargaram os seus estudos a toda a bacia hidrográfica do Rio Lis, conjuntamente com um aumento das intervenções no contexto da Arqueologia Preventiva (Braz e Gaspar, 2003; Cunha-Ribeiro, 2005; Aubry *et al.*, 2005; Carvalho e Tavares, 2005; Carvalho e Pajuelo, 2005; Braz *et al.*, 2006; Carvalho e Carvalho, 2007; Pereira, 2010; Gameiro, 2019), tendo sido assim revelado um número significativo de outros sítios arqueológicos.

A existência de um povoamento continuado desde o Paleolítico Inferior, em toda a região de Leiria, pode justificar-se, em grande parte, pelas características geológicas, geomorfológicas e paleoclimáticas locais e regionais. Destaca-se a presença de afloramentos carbonatados no setor mais meridional e oriental que, num contexto tectónico e climático variável no espaço e no tempo, permitiram o desenvolvimento de formas cársicas superficiais e subterrâneas, entre as quais destacam-se as grutas e os abrigos sucessivamente utilizadas pelos caçadores-recolectores (*e.g.*, Zilhão *et al.*, 2021). Outra característica bastante importante é a presença de um retículo hidrográfico bem desenvolvido na dependência de uma cobertura detrítica siliciclástica, sendo a água um recurso vital para a sobrevivência (Carvalho, 2011).

Em termos de fontes de matérias-primas das indústrias de pedra lascada, durante o Paleolítico, destaca-se a utilização do sílex local e do quartzito para criar utensílios utilizados nas mais diversas tarefas do quotidiano - *e.g.*, a caça (Zilhão, 1997; Cunha-Ribeiro, 1999, 2002; Raposo, 2002; Zilhão e Almeida, 2002).

As primeiras descobertas arqueológicas feitas na região ocorreram em 1879, pelas mãos do Prof. Carlos Ribeiro, na região de Marrazes e Barosa, onde dois artefactos líticos foram recolhidos e atribuídos ao Paleolítico Inferior (Carvalho, 2011). Também Émile de Cartailhac descobriu mais um artefacto (biface), desconhecendo-se a sua origem. Achados idênticos foram descobertos por Tavares Proença Júnior, junto às margens do Rio Lena e Mesquita Figueiredo, nas imediações da povoação de Milagres (Cunha-Ribeiro, 2002). Manuel Heleno centrou aí as suas campanhas arqueológicas, sobretudo nos anos quarenta do século XX, identificando 60 sítios arqueológicos (Heleno, 1922; Texier e Cunha-Ribeiro, 1991/1992; Cunha-Ribeiro, 2005).

Apesar destas investigações prévias, os achados arqueológicos mais interessantes foram encontrados aquando dos trabalhos de campo levados a cabo para a elaboração da Folha 23-C da Carta Geológica de Portugal, na escala 1:50 000 (Teixeira *et al.*, 1968), com a descoberta de três sítios atribuídos ao Paleolítico Inferior - dois em terraços fluviais e o outro em coluviões (Zbyszewski *et al.*, 1980). Entre 1984 e 2005, o Prof. Cunha-Ribeiro desenvolveu trabalhos relativos ao Paleolítico Inferior em toda a bacia hidrográfica do Rio Lis, recolhendo mais artefactos nos sítios já conhecidos, como Pousias/Quinta de Cónego (Oosterbeek *et al.*, 2010) e Outeiro Pelado, assim como em novos sítios associados aos depósitos do Plistocénico que materializam terraços fluviais e coluviões ao longo das vertentes (Cunha-Ribeiro, 1992, 1999).

Quanto ao Paleolítico Médio, não são conhecidos muitos locais de ocupação no Baixo Vale do Rio Lis, destacando-se apenas o sítio arqueológico da Praia do Pedrogão na faixa litoral, sendo um local de particular relevância no panorama da investigação arqueológica nacional, na medida em que fornece as evidências de uma estratégia de exploração do território e dos recursos relacionados com o litoral (Aubry *et al.*, 2005; Carvalho e Carvalho, 2007).

Também o conhecimento sobre o Paleolítico Superior na área de estudo tem uma reduzida expressão, embora as várias campanhas de escavação realizadas nos ambientes cársicos do Maciço Calcário Estremenho tenham permitido um avanço substancial, sendo o exemplo mais emblemático o já citado “Menino do Lapedo” do Lagar Velho (Zilhão & Trinkaus, 2002). Ainda em ambiente cársico, onde se evidenciam os traços de uma evolução superficial de tipo fluviocársica (depressões fechadas ou semifechadas), foram recuperados materiais líticos atribuídos ao Paleolítico Superior como, por exemplo, no sítio ao ar-livre da Portela II (em 2009), aquando da intervenção realizada pela Câmara Municipal de Leiria na abertura de uma vala para a instalação de uma conduta (Gameiro, 2019).

3 MATÉRIAS E MÉTODOS

A escolha da AR e da AMC como abordagens gerais neste trabalho prende-se, fundamentalmente, com a capacidade intrínseca destas para levarem em conta a variabilidade espacial de diferentes tipologias de dados (ambientais e arqueológicos) e de os combinar entre si de uma forma muito mais simples e eficaz por comparação com as aptidões da estatística espacial tradicional. Por sua vez, os SIG desempenham um papel fundamental, possibilitando a gestão de dados geográficos georreferenciados, em várias camadas informativas e escalas, permitindo a execução das operações geoestatísticas e de sobreposição espacial necessárias. A Figura 6 ilustra, esquematicamente, as três etapas metodológicas gerais implementadas neste trabalho, onde se distingue o pré processamento, o processamento e a análise.

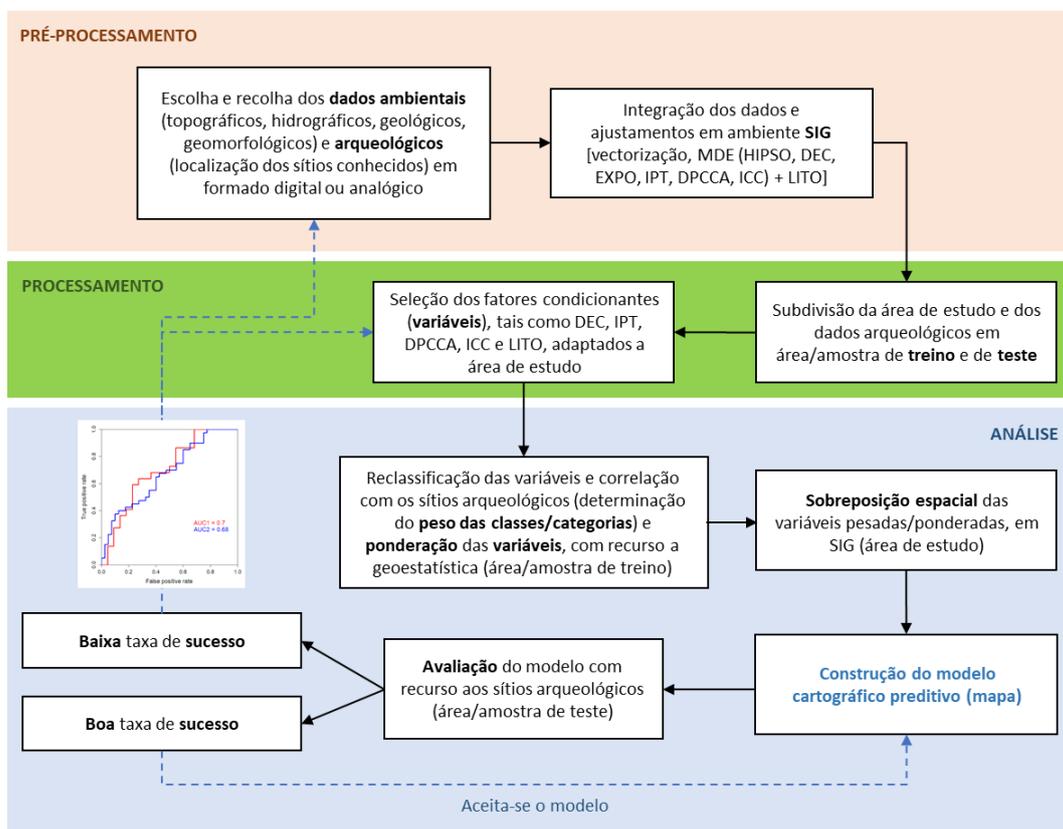


Figura 6 - Fluxograma metodológico utilizado neste trabalho. O pré-processamento diz respeito, essencialmente, a recolha dos dados (analógicos ou digitais) e a sua integração e ajustamento em SIG, onde se procede a vectorização e gera-se o modelo digital de elevação (MDE) que é sucessivamente utilizado para derivar a hipsometria (HIPSO), o declive (DEC), a exposição (EXPO), o índice de posição topográfica (IPT), a distância ponderada de custo para acesso aos cursos de água (DPCCA), quando considerado juntamente com a rede hidrográfica vectorizada, e o índice de carga de calor (ICC); se acrescenta a esta lista a litologia (LITO). Durante o processamento definem-se a área/amostra de treino e de teste e são selecionados os fatores ambientais condicionantes a utilizar como variáveis de entrada do modelo, adaptados a área de estudo. Na última etapa (análise), todos os conjuntos de dados recolhidos e processados nas duas etapas anteriores são pesados/ponderados e integrados, em SIG, para a implementação do modelo cartográfico preditivo (mapa) sujeito a revisão e avaliação no que diz respeito ao seu desempenho/eficácia.

3.1 Recolha e processamento dos dados

Na Tabela 1 sistematiza-se a base de dados espacial construída para este estudo. Esta configuração compreende a seleção de um conjunto de fatores ambientais condicionantes, derivados em SIG a partir dos seguintes dados de base previamente recolhidos em formato digital ou analógico: topografia (curvas de nível, pontos cotados); hidrografia (rede hidrográfica e lagoas); geologia (litologia e estruturas geológicas); e geomorfologia (terraços fluviais). A estes, junta-se a recolha dos dados arqueológicos correspondentes à localização dos sítios paleolíticos conhecidos até a data atual na área de estudo.

3.1.1 Sítios paleolíticos conhecidos

Os dados arqueológicos foram extraídos a partir do Sistema de Informação e Gestão Arqueológica - Endovélico - da Direção Geral do Património Cultural (DGPC), utilizando o GeoPortal do Arqueólogo (<https://arqueologia.patrimoniocultural.pt/> - último acesso 31 de maio 2021). Importa referir que, embora se trate de uma base de dados arqueológica oficial do governo português, é possível que alguns dos sítios aí incluídos estejam erroneamente localizados e/ou com uma atribuição cronológica cultural incerta, sobretudo os que foram identificados exclusivamente através das prospeções de superfície, sem ter havido verdadeiras escavações. Por outro lado, as diferentes estratégias de escavação adotadas (decapagem mecânica, escavação manual, sondagens ou sanjas, etc.), muitas vezes condicionadas pela execução dos trabalhos no âmbito da arqueologia de salvamento vs. científico-académica, também podem ter criado diferenças nas abordagens e, conseqüentemente, na própria caracterização dos sítios. Desta forma, e para evitar, tanto quanto possível, a inclusão de erros na lista final, os sítios iniciais foram verificados, ordenados e cruzados com as informações complementares disponíveis na literatura através da leitura de artigos científicos e/ou relatórios (trabalho este sistematizado pelo arqueólogo Armando Melo, através de uma prestação de serviço, no âmbito do projeto PALEORESCUE). Todas as entradas desta lista de dados incluíram a respetiva informação geográfica (designação, freguesia, lugar, acesso e altitude), assim como um conjunto de outros elementos descritivos (tipologia de sítio, cronologia, implantação, hidrografia e espólio). Este processo permitiu selecionar um total de 40 sítios ao ar-livre georreferenciados e oficialmente atribuídos ao Paleolítico Inferior, Médio, Superior ou indiferenciado (Fig. 7).

Tabela 1 - Base de dados geoespacial, construída em SIG, para a elaboração do modelo preditivo de identificação do potencial arqueológico Paleolítico na região de Leiria (Centro de Portugal). CIMRL = Comunidade Intermunicipal da Região de Leiria; IGeoE = Instituto Geográfico do Exército; LNEG = Laboratório Nacional de Engenharia e Geologia; DGPC = Direção Geral do Património Cultural; Endovélico = Sistema de Informação e Gestão Arqueológica da DGPC. A resolução espacial das camadas *rasters* derivadas é de 10 m x 10 m de *pixel*. Vetorial: modelo de dados georreferenciado, em SIG, que permite representar os elementos geográficos como pontos, linhas ou polígonos e onde a cada elemento está associada uma tabela de atributos. *Raster* (ou matricial): modelo de dados, em SIG, que define o espaço como uma quadricula, de igual tamanho de células, constituída por linhas e colunas, e onde cada célula contém um atributo (valor) e as coordenadas de localização; grupos de células que partilham o mesmo valor representam a mesma tipologia de elemento geográfico. Em negrito as camadas *rasters* utilizadas como variáveis de entrada na modelação espacial preditiva.

Configuração (base de dados espacial)	Tipologia de dados em SIG	Recolha dos dados para a área de estudo		Camadas <i>rasters</i> derivadas (Fatores condicionantes) em SIG
		Escala ou resolução	Fonte ou citação	
Topografia	Curvas de nível (vetorial - linhas)	1:10 000	CIMRL (2019)	- Hipsometria ^B
	Pontos cotados (vetorial - pontos)	1:10 000	CIMRL (2019)	- Declive^B
	Cartas militares (<i>raster</i>)	2.5 m x 2.5 m	IGeoE (2003)	- Exposição ^B
Hidrografia	Rede hidrográfica (vetorial - linhas) ^A	1:25 000	IGeoE (2003)	- Índice de posição topográfica^C
	Lagoas (vetorial - polígonos) ^A	1:25 000	IGeoE (2003)	- Distância ponderada de custo para acesso aos cursos de água^D
Geologia	Litologia (vetorial - polígonos)	1:50 000	LNEG (2017)	- Índice de carga de calor^E
	Estruturas tectónicas (vetorial - linhas e pontos)	1:50 000	LNEG (2017)	- Litologia^G
Geomorfologia	Terraços fluviais (vetorial - polígonos) ^F	1:25 000	IGeoE (2003)	
Arqueologia	Localização dos sítios arqueológicos (vetorial - pontos) ^H	1:100 000	DGPC (Endovélico)	- Sítios paleolíticos conhecidos

^ADados vetoriais digitalizados a partir da Carta Militar de Portugal, na escala 1:25 000 (Série M888; IGeoE, 2003)

^BDerivado do modelo digital de elevação (MDE - tamanho do pixel de 10 m x 10 m), por sua vez gerado com a ferramenta " *TopoToRaster*" em ArcGIS 10.8.1

^CDerivado do MDE através da utilização da ferramenta " *Topographic Tools Toolbox*" em ArcGIS 10.8.1

^DDerivado da combinação entre a rede hidrográfica e o MDE através a utilização da ferramenta " *Cost-weighted distance*" em ArcGIS 10.8.1

^EDerivado do MDE através da utilização da ferramenta " *Geomorphometry and Gradient Metric Tools*" em ArcGIS 10.8.1

^FDados vetoriais interpretados e digitalizados a partir da Carta Militar de Portugal, na escala 1:25 000 (Série M888; IGeoE, 2003), suportada pelo MDE

^GOnde se inclui a identificação dos terraços fluviais do Plistocénico

^HDisponível em <https://arqueologia.patrimoniocultural.pt/> (último acesso 31 de maio de 2021), utilizando o GeoPortal do Arqueólogo

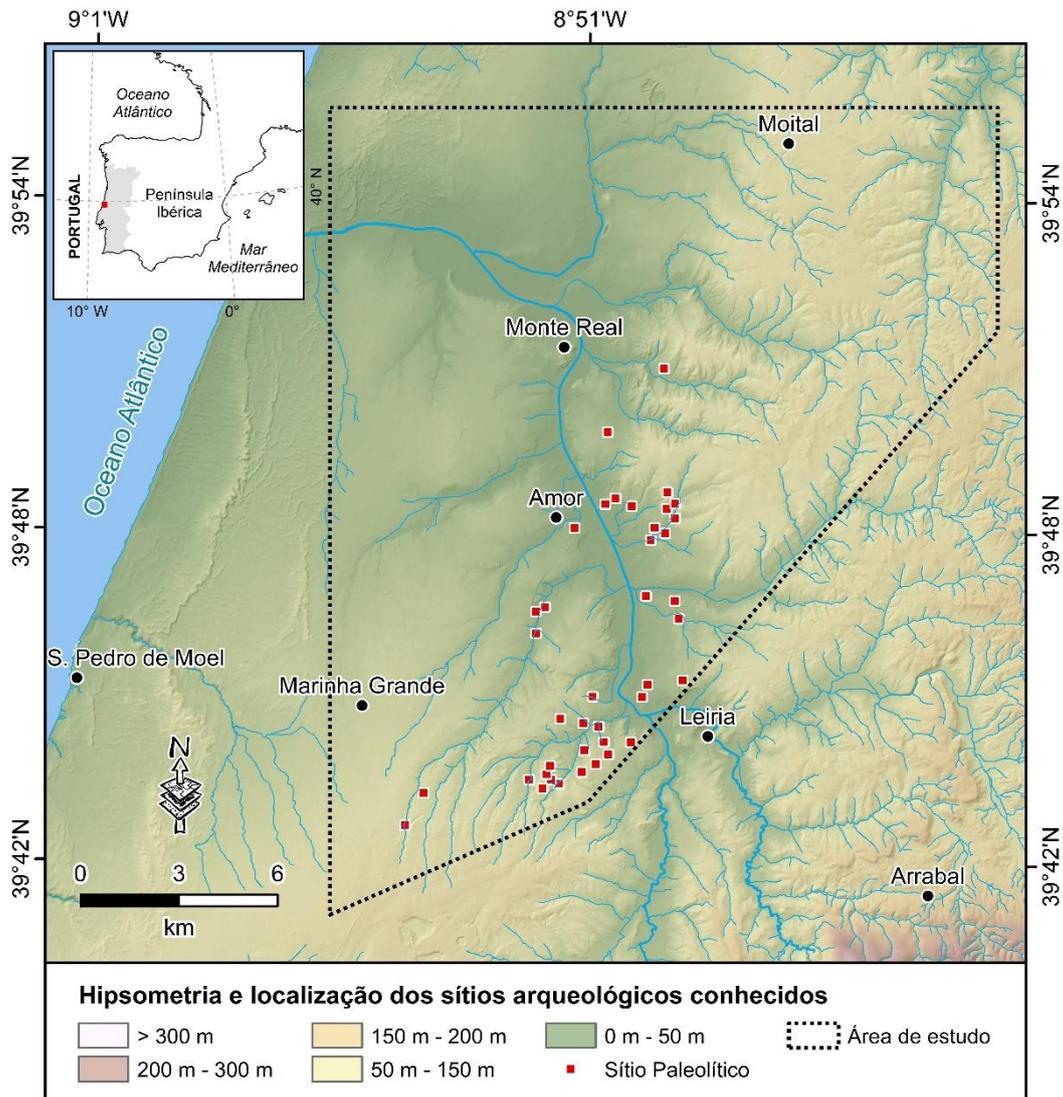


Figura 7 - Mapa hipsométrico da área de estudo (excluído da modelação espacial preditiva - mais detalhes no texto) e localização dos 40 sítios arqueológicos paleolíticos ao ar-livre selecionados para este trabalho. O modelo digital de elevação, com uma resolução espacial de 10 m x 10 m de *pixel*, foi construído utilizando curvas de nível, pontos cotados, lagoas e linhas de águas vetorizadas a partir das Cartas Militares de Portugal em escala 1:25 000 (Série M888; IGeoE, 2003).

3.1.2 Fatores ambientais condicionantes

A seleção de todos os fatores ambientais condicionantes da Tabela 1 (hipsometria, declive, exposição, índice de posição topográfica, distância ponderada de custo para acesso aos cursos de água, índice de carga de calor e litologia) baseou-se, antes de mais, na observação direta, no campo, das características físicas atuais da área de estudo, tendo sempre em consideração os pressupostos pré-definidos no âmbito da modelação espacial preditiva aplicada à arqueologia pré-histórica (*vide* o subcapítulo 1.1). Após um procedimento iterativo e

demorado de inclusão-exclusão de fatores, chegou-se à conclusão de que um número elevado destes prejudicaria a exequibilidade dos procedimentos analíticos; ao mesmo tempo, a utilização de um número limitado evitaria também os eventuais problemas de multicolinearidade (Rose e Altschul, 1988). Como resultado deste exercício interativo, antecipase que, após realizar uma série de ensaios (combinando todos os fatores da Tabela 1 entre si), a hipsometria e a exposição foram excluídas da lista dos fatores ambientais condicionantes a serem utilizados como variáveis de entrada na modelação espacial preditiva, pois verificou-se que a sua inclusão reduzia significativamente a capacidade preditiva dos modelos aplicados. Para a determinação da multicolinearidade espacial entre fatores recorreu-se ao coeficiente de correlação de Pearson, estatisticamente analisado com um nível de significância de $P < 0.05$ (Pearson, 1895; Landau e Everitt, 2004). De facto, na comparação entre a hipsometria com a distância ponderada de custo para acesso aos cursos de água, assim como entre a exposição com o índice de carga de calor, o coeficiente de Pearson apresenta valores perto de 1 (que expressa uma correlação perfeita positiva entre as duas variáveis). Para além disso, é preciso considerar que a altimetria (hipsometria) está sempre incluída nos procedimentos metodológicos que permitiram derivar, em SIG, todos os outros fatores (com exclusão da litologia); por sua vez, a orientação das vertentes (exposição) está também indiretamente expressa, juntamente com a altimetria, no índice de carga de calor.

Importa também referir que, embora seja um fato bem estabelecido que em ecologia da paisagem a vegetação tende a formar mosaicos dinâmicos que influenciam, na escala de pormenor, o comportamento animal e humano, levando assim a variações significativas nas escolhas culturais no espaço e no tempo (Warden *et al.*, 2017; Grøn, 2018; e referências neles incluídas), este fator não foi aqui considerado devido à nossa incapacidade em julgar as condições edáficas e de uso do solo a uma escala analítica e espacial suficientemente pormenorizadas.

Para proceder à caracterização do relevo foi gerado, em SIG, um modelo digital de elevação (MDE), com o tamanho do *pixel* de 10 m x 10 m, a partir da integração das curvas de nível, dos pontos cotados, da rede hidrográfica e das lagoas (dados vetoriais obtidos através da digitalização da Carta Militar de Portugal, na escala 1:25 000, ou fornecidos com base num protocolo de colaboração com a Comunidade Intermunicipal da Região de Leiria - CIMRL), utilizando a ferramenta “*TopoToRaster*” do software ArcGIS 10.8.1. Este MDE foi utilizado para derivar a **hipsometria**, o **declive** e a **exposição**, utilizando a ferramenta “*Surface*” do “*Spatial Analyst Tools*” em ArcGIS, onde os valores dos *pixéis* identificam a altimetria em metros, a inclinação das vertentes em graus e a direção das vertentes em relação aos pontos cardinais, respetivamente. Trata-se, no conjunto, das variáveis topográficas mais utilizadas na literatura

para a elaboração de modelos espaciais preditivos aplicados à arqueologia pré-histórica por estarem, direta ou indiretamente, todas relacionadas com a capacidade de deslocação e de fixação do ser humano no território (Conolly e Lake, 2006), pressupondo que para esta fixação a escolha recaía nas áreas com topografia mais suave, pouco íngreme e ensolaradas (que recebem mais luz solar e calor durante o dia), em detrimento dos terrenos mais acidentados, declivosos e com uma exposição que restringia a quantidade de radiação solar recebida à superfície (Kvamme, 1992; Siart *et al.*, 2008). É de salientar também que nas áreas onde o declive é mais acentuado os processos de remobilização das coberturas sedimentares atuam com maior facilidade e, eventualmente, de uma forma mais intensa, podendo condicionar a preservação do registo arqueológico. Na área de estudo, a hipsometria e o declive foram ambas ordenados em 5 classes (Figs. 7 e 8); para a exposição se consideraram 9 classes (Fig. 9).

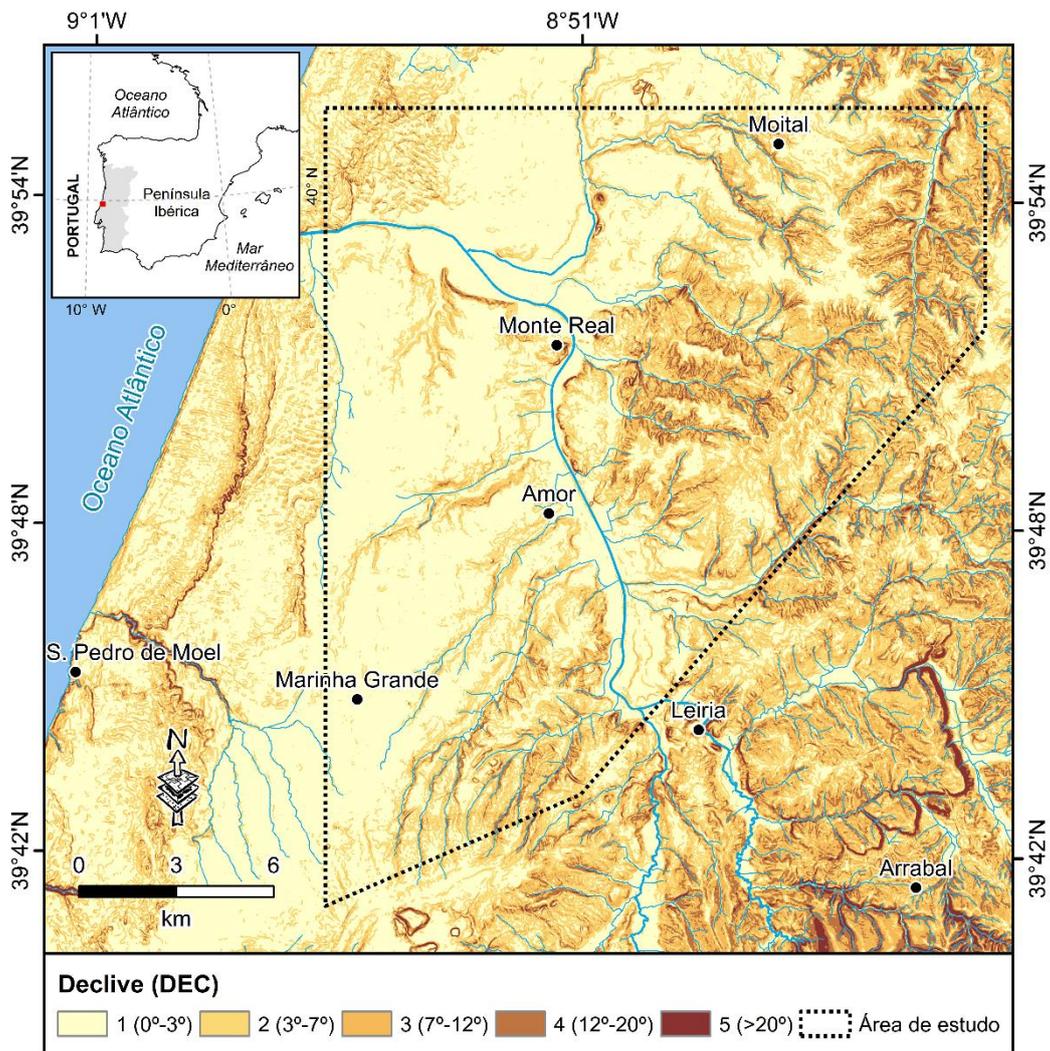


Figura 8 - Mapa de declive da área de estudo, utilizado como variável de entrada na modelação espacial preditiva.

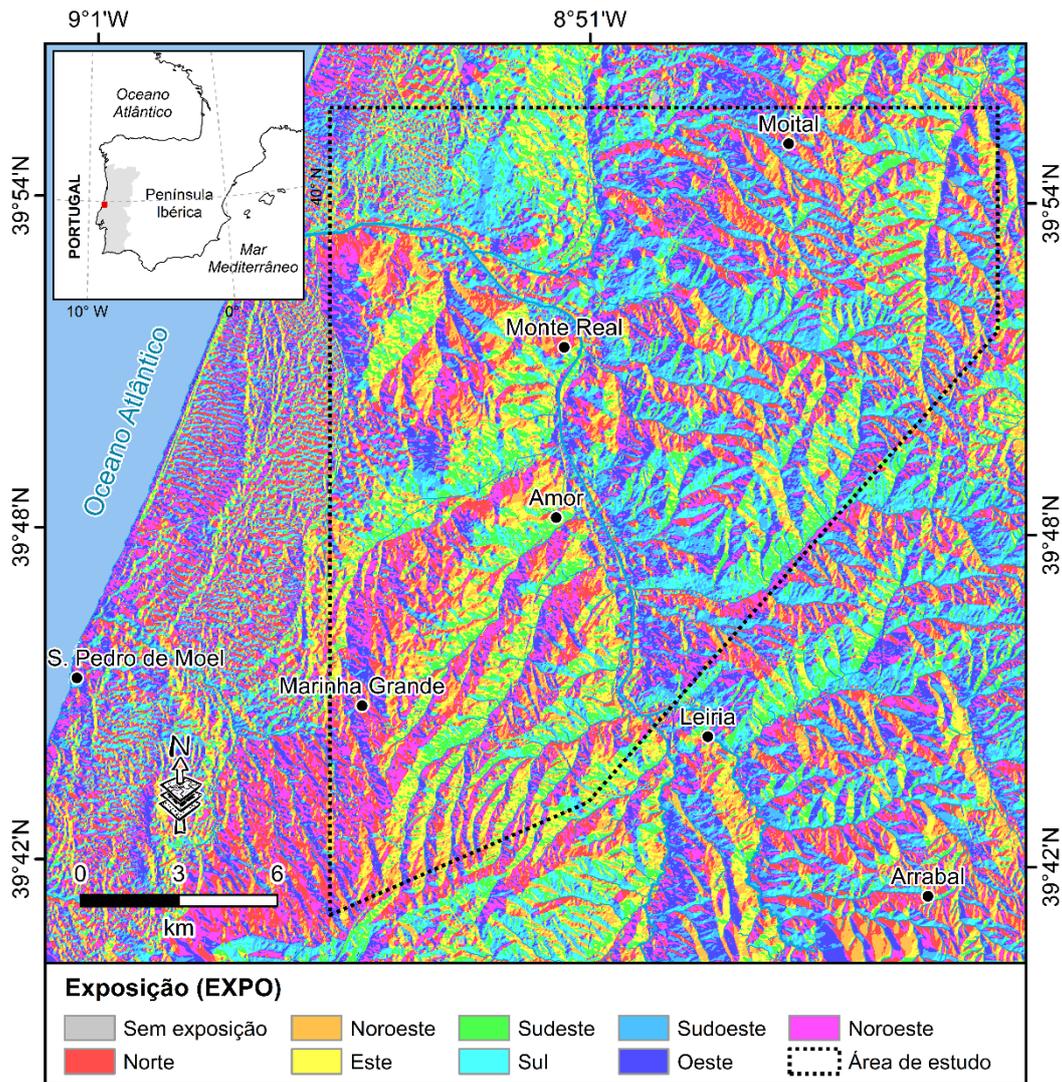


Figura 9 - Mapa de exposição da área de estudo (excluído da modelação espacial preditiva - mais detalhes no texto).

O **índice de posição topográfica** foi obtido a partir do MDE com recurso à ferramenta "Topographic Tools Toolbox" (<http://www.gisandbeers.com/topography-tools-toolbox-herramientas-analisis-topografico/> - último acesso em julho 2021), integrada na *toolbox* do ArcGIS, que permite uma classificação semiautomática das formas do relevo com base nas diferentes posições identificáveis ao longo da encosta e na categorização das formas (Weiss, 2001; Dilts, 2015). Os valores obtidos através da aplicação deste algoritmo identificam a diferença entre a altimetria numa célula (*pixel*) e a altimetria média das células vizinhas ao seu redor. Valores positivos indicam que a célula está numa posição topográfica superior relativamente às células vizinhas, enquanto valores negativos indicam que a célula se encontra numa posição abatida (Tagil e Jenness, 2008). A posição da célula na vertente é definida pela extensão desta diferença. Em outras palavras, os valores do índice de posição topográfica

forneem um meio expedito e eficaz para classificar a paisagem em classes morfológicas (Jenness, 2006). A escolha do tamanho e forma das vizinhanças, que influencia a classificação final do índice, é feita de forma interativa através de várias tentativas com diferentes opções que devem ter sempre em conta a escala analítica que está a ser utilizada para caracterizar a paisagem. Neste trabalho, uma vizinhança retangular foi usada, produzindo uma variação continua de valores desde -43.1 até 45.4. Os valores mais altos correspondem às áreas perto do topo das colinas, os mais baixos aos fundos dos vales, enquanto os valores em torno de 0 correspondem a superfícies planas (inclinação muito baixa) ou aos setores no meio da encosta, se a inclinação for mais significativa (Jenness, 2006). Para a área de estudo, estes valores contínuos foram ordenados em 6 classes (Fig. 10).

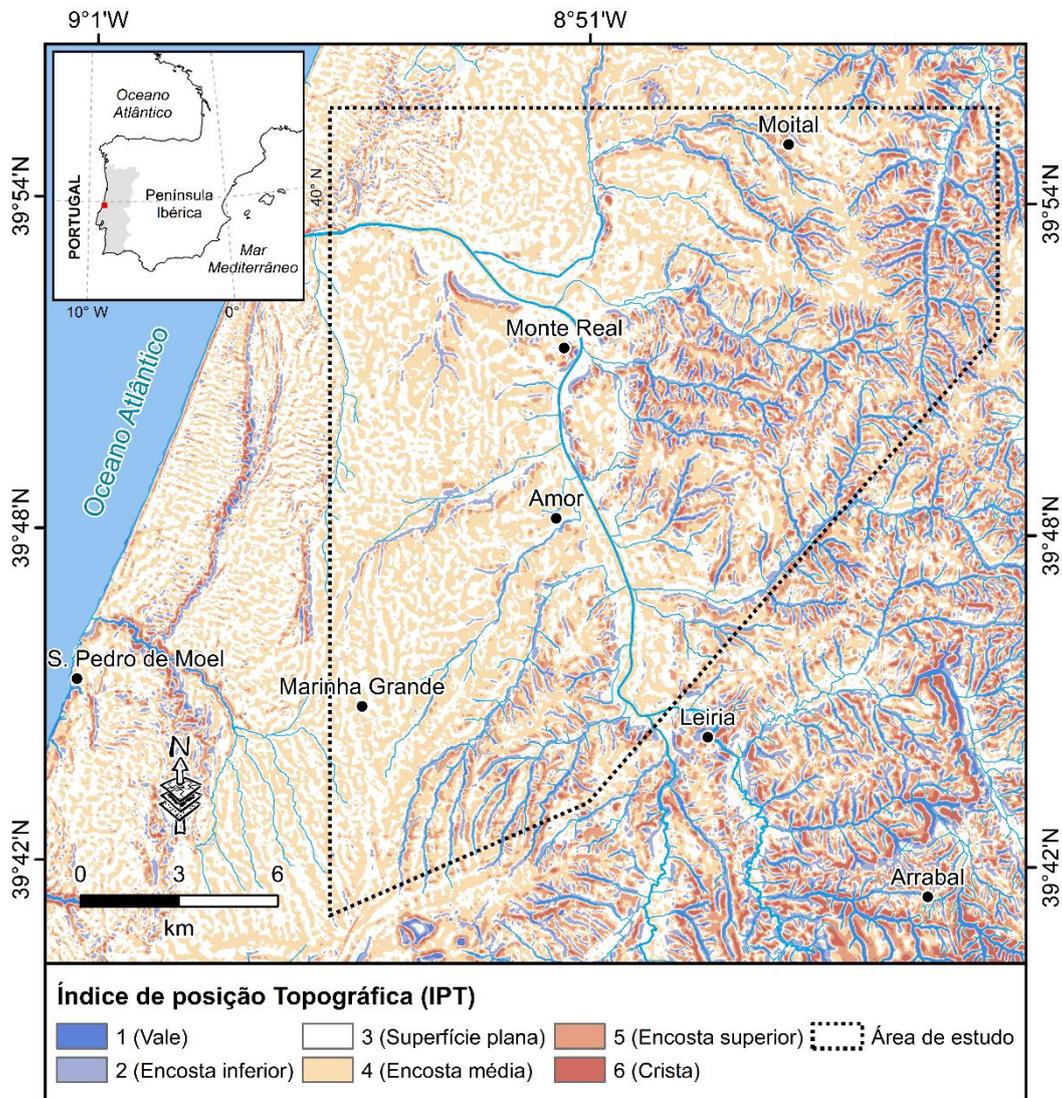


Figura 10 - Mapa do índice de posição topográfica da área de estudo, utilizado como variável de entrada na modelação espacial preditiva.

A combinação do MDE com a rede hidrográfica vetorizada permitiu a elaboração da **distância ponderada de custo para acesso aos cursos de água**, utilizando a ferramenta "*Cost-weighted distance*" do "*Spatial Analyst Tools*" em ArcGIS. A importância da hidrografia, assim como dos fatores topográficos atrás referidos (hipsometria, declive e índice de posição topográfica), na modelação espacial preditiva aplicada à arqueologia pré-histórica, relaciona-se com a constatação de que a implantação clássica dos sítios paleolíticos conhecidos no território nacional (mas não só; *vide* Brown, 1997) parece depender, sobretudo em contexto fluvial, sempre, ou quase sempre, da existência de uma topografia aplanada (pequenas plataformas na margem ou no meio de meandros de rios de pequena ou média dimensão) ou pouco inclinada na encosta superior para proteção das inundações e dos movimentos de matérias ao longo da vertente, da proximidade de recursos naturais essenciais (*e.g.*, água e combustível lenhoso) e do acesso a matérias-primas líticas disponíveis nos depósitos superficiais (*e.g.*, aluviões, depósitos de terraços fluviais e de vertente) (Aubry, 2009; Almeida, 2013; Gaspar *et al.*, 2016a, 2016b; entre outros). Para além da necessidade de água (para consumo humano ou como recurso indispensável em certas tarefas domésticas) e de combustível lenhoso (Karmon, 1971), as margens dos rios podem funcionar também como zonas de bebedouros para animais, sendo que a sua concentração num determinado local, ao longo do rio, facilitaria a caça (Binford, 1983; Zilhão, 1997). Por outro lado, em Portugal, alguns dos vales dos principais rios têm sido interpretados como eixos viários e zonas preferenciais de deslocação durante a pré-história. O vale do Vouga, por exemplo (Gameiro e Aubry, 2020), está situado a meio caminho entre a concentração de sítios da Estremadura e de algumas das fontes de sílex com boa aptidão ao talhe utilizadas também no Vale do Rio Côa (Aubry *et al.*, 2016). Neste contexto, diversos autores defendem que é mesmo a conjugação da altimetria com a distância às linhas de água (*i.e.*, a distância ponderada de custo para acesso aos cursos de água) a variável que melhor traduz o verdadeiro esforço necessário para ter acesso ao recurso (*e.g.*, Wheatley e Gillings, 2002). Para a área de estudo, os valores contínuos da distância ponderada de custo para acesso aos cursos de água foram ordenados em 10 classes, com a classe 1 a corresponder à menor distância e a 10 à distância maior (Fig. 11).

O **índice de carga de calor** também foi derivado do MDE, utilizando a ferramenta "*Geomorphometry and Gradient Metric Tools*" (<https://www.arcgis.com> - último acesso em julho 2021), integrada na *toolbox* do ArcGIS. Este índice ganha relevância, em relação a outros fatores como a exposição ou a radiação solar, na medida em que existem locais à superfície da terra que, embora recebam a mesma quantidade de radiação solar, tendem a evidenciar valores de

temperatura média diferentes. É este o caso das vertentes voltadas a sudeste e sudoeste, que recebem ambas uma quantidade de radiação solar equivalente, mas as segundas evidenciam-se como relativamente mais quentes (McCune e Keon, 2002). O método de McCune e Keon (2002) [...] *“accounts for this by “folding” the aspect so that the highest values are southwest and the lowest values are northeast”*; ou seja, o algoritmo utilizado para derivar este índice praticamente “dobra” a exposição de forma que os valores mais altos fiquem atribuídos as encostas viradas para sudoeste e os valores mais baixos às viradas para nordeste. Outra vantagem na utilização deste índice, por comparação com a exposição ou a radiação solar, diz respeito ao facto de que a inclinação das vertentes ser também considerada no seu cálculo (McCune e Keon, 2002). Este índice apresenta valores contínuos entre 0 (mais frio) e 1 (mais quente). Para a área de estudo, os valores contínuos do índice de carga de calor foram ordenados em 5 classes (Fig. 12).

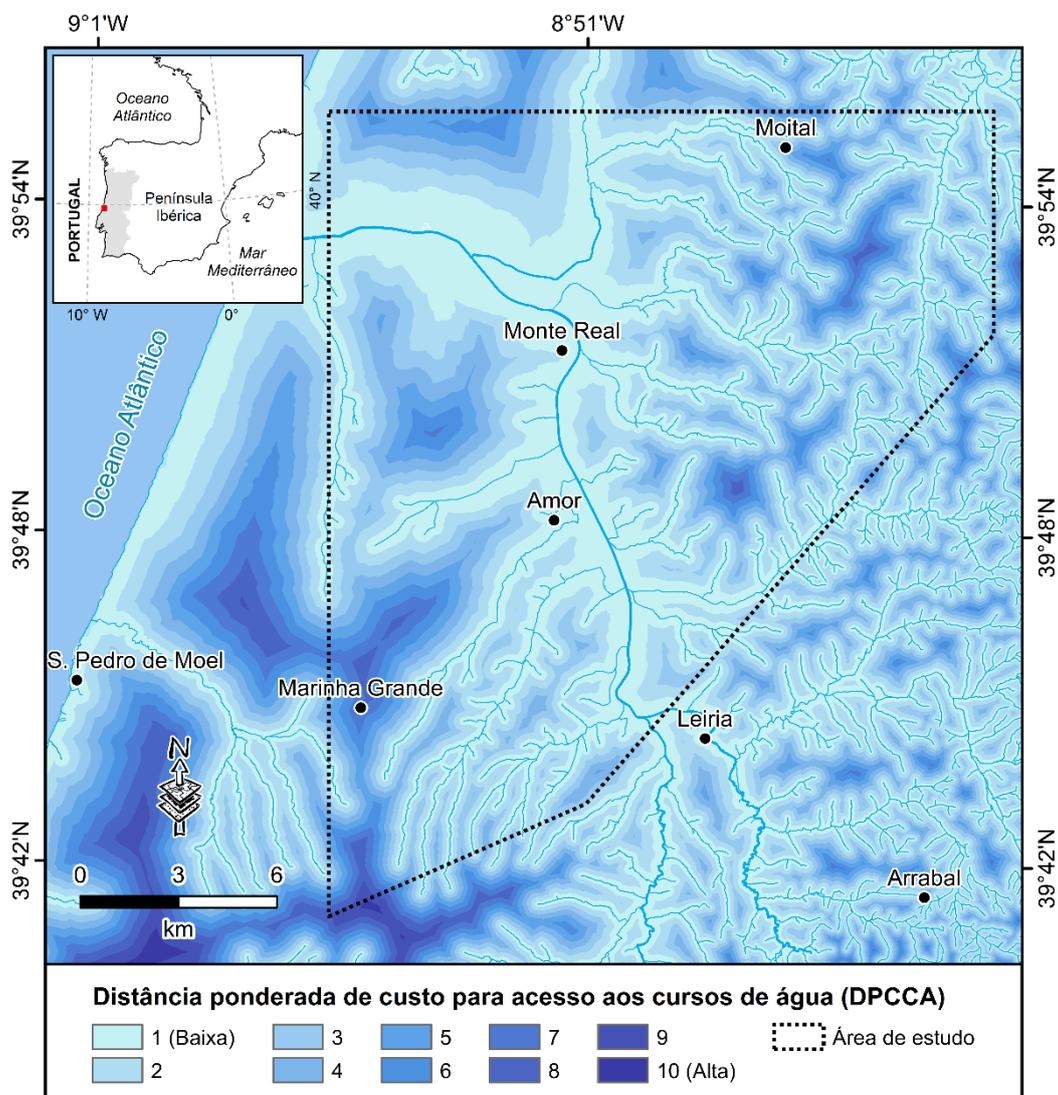


Figura 11 - Mapa da distância ponderada de custo para acesso aos cursos de água da área de estudo, utilizado como variável de entrada na modelação espacial preditiva.

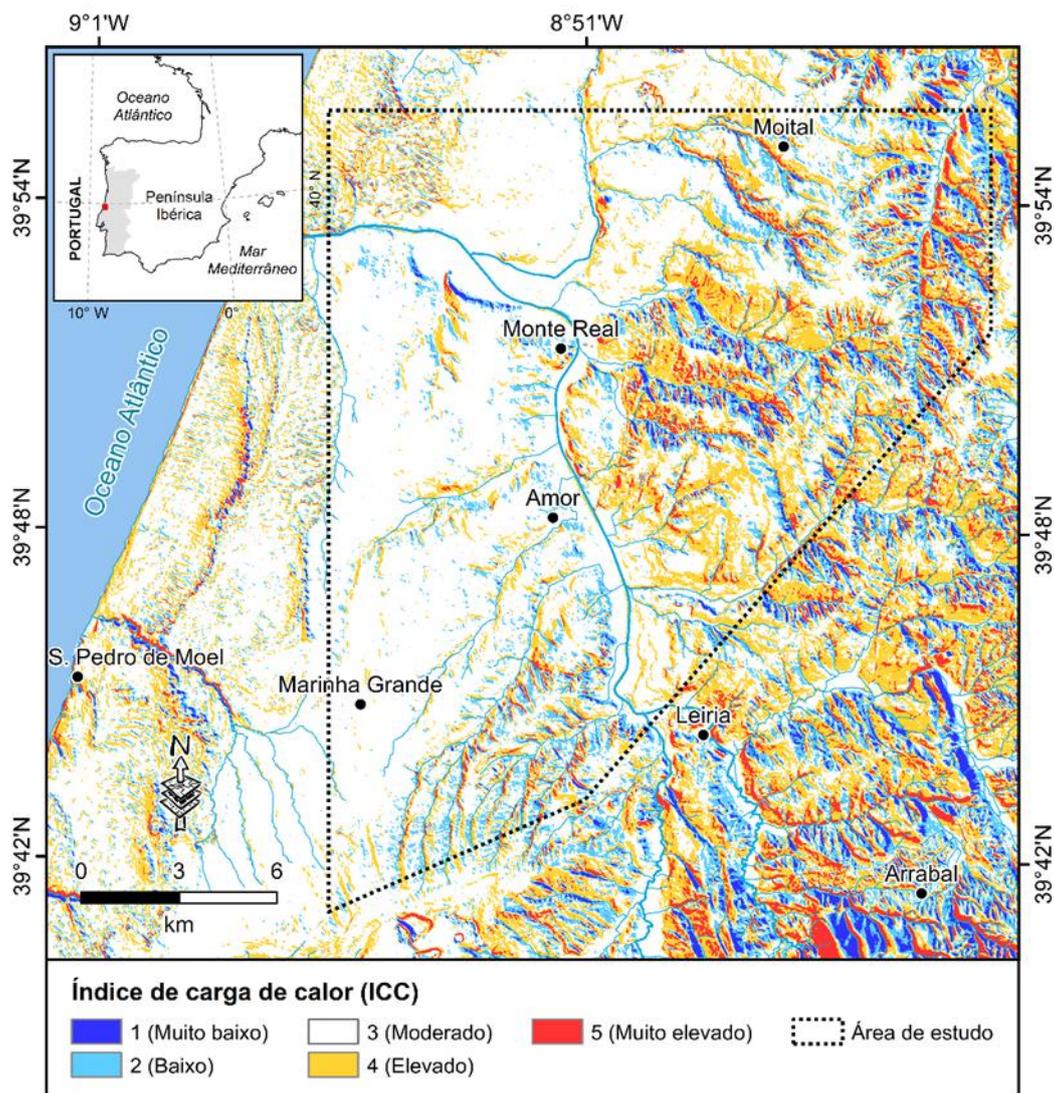


Figura 12 - Mapa do índice de carga de calor da área de estudo, utilizado como variável de entrada na modelação espacial preditiva.

A **litologia** foi obtida através da simples transformação em *raster* dos polígonos que identificam as unidades líticas aflorantes na área de estudo (dados vetoriais disponibilizados pelo Laboratório Nacional de Engenharia e Geologia - LNEG), assim como dos terraços fluviais do Plistocénico digitalizados a partir da Carta Militar de Portugal, na escala 1:25 000, com apoio no MDE (Fig. 13). Em contexto fluvial, os depósitos que atualmente identificam os restos, geralmente degradados, de antigas planícies aluviais (terraços) são os mais propensos à preservação do registo arqueológico devido aos próprios processos/mecanismos de sedimentação que rapidamente preservam os sítios, tornando-os ao mesmo tempo invisíveis na paisagem (Brown, 1997).

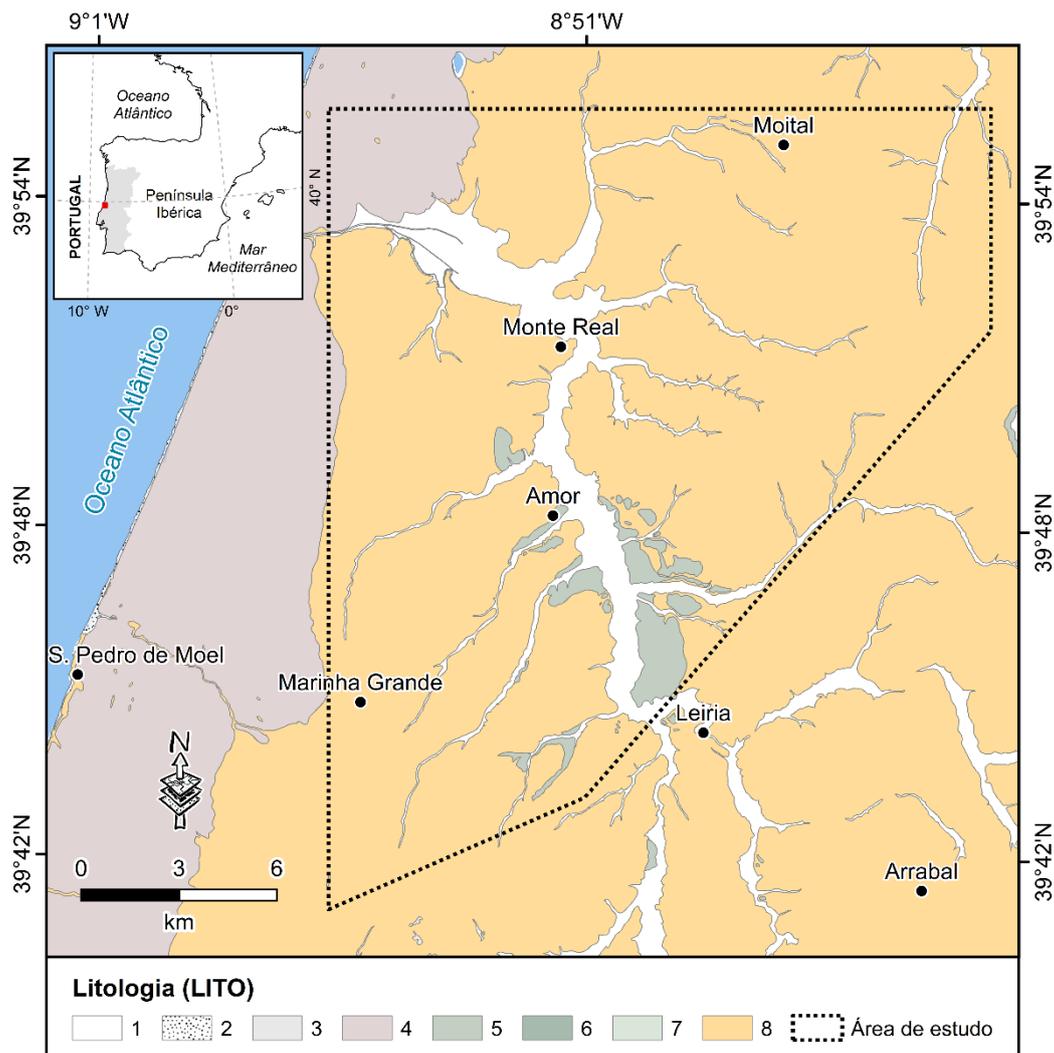


Figura 13 - Mapa da litologia da área de estudo, utilizado como variável de entrada na modelação espacial preditiva. (1) Aluviões - Holocénico; (2) Areias de praia - Holocénico; (3) Depósitos de vertente - Holocénico; (4) Dunas e areias eólicas - Plistocénico a Holocénico; (5) Depósitos de terraços fluviais - Plistocénico; (6) Tufos calcários - Plistocénico; (7) Depósitos de terraços marinhos - Plistocénico; (8) Substrato pré-Plistocénico. Para a identificação das unidades que compõem o substrato pré-Plistocénico ver a legenda da Figura 5.

As áreas de deposição dos coluviões também podem servir como locais de retenção para os materiais arqueológicos remobilizados ao longo da vertente, por força da gravidade e/ou da escorrência superficial, a partir de sítios localizados na parte superior da encosta ou mesmo na crista. Importa referir que a delimitação das manchas que na cartografia geológica oficial identificam os afloramentos de depósitos de terraços fluviais do Plistocénico, por vezes, foram ligeiramente retificadas com base nas observações de campo. De facto, durante o exercício de correlação entre o mapa geológico da área de estudo e os 40 sítios paleolíticos conhecidos verificou-se que alguns deles se localizavam mesmo na interface entre os depósitos de terraço fluviais e o substrato pré-Plistocénico, indicando assim um provável erro de georreferenciação dos sítios ou mesmo uma imprecisão no levantamento geológico (com esta última resultando,

sobretudo, das limitações impostas pela escala cartográfica utilizada - 1:50 000 - na representação analógica original deste mapa). Para a área de estudo, a litologia foi ordenada em 8 categorias, onde se consideram, separadamente, todos os depósitos do Quaternário reconhecíveis (de 1 a 7) e o substrato pré-Pleistocénico (8) (*vide* as legendas das Figuras 5 e 13).

As cinco variáveis escolhidas (declive, índice de posição topográfica, distância ponderada de custo para acesso aos cursos de água, índice de carga de calor e litologia), juntamente com os sítios paleolíticos conhecidos, foram reclassificadas, com resolução espacial comum de 10 m x 10 m de *pixel*, e utilizadas como camadas *rasters* de entrada nos procedimentos metodológicos sucessivos. A camada de informação arqueológica foi sempre utilizada como simples ocorrência ou não-ocorrência binária (1/0) dos sítios. Para a definição das classes, no caso das variáveis contínuas (*i.e.*, não discretas), foi usado o método das quebras naturais onde as classes se baseiam em agrupamentos naturais inerentes aos dados (são identificadas as quebras de classe que melhor agrupam valores semelhantes e que maximizam as diferenças entre as classes - os valores contínuos dos *pixéis* são divididos em classes cujos limites se definem onde existem diferenças relativamente grandes entre os valores). Antes dos cálculos inerentes a análise, ou seja, antes da implementação e verificação dos modelos, tanto a área de estudo, como os sítios arqueológicos conhecidos, foram subdivididos em dois conjuntos de dados: (i) área de treino (correspondente a $\sim 1/3$ da área de estudo) e amostra de treino (18 sítios), respetivamente; e (ii) área de teste (restantes $\sim 2/3$ da área de estudo) e amostra de teste (restantes 22 sítios), respetivamente.

3.2 Implementação dos modelos cartográficos preditivos

Na modelação espacial preditiva, a atribuição de um valor numérico que represente a importância relativa na explicação do fenómeno em causa, para cada variável ambiental condicionante (ponderação) e para as classes/categorias (peso) de cada uma destas variáveis, torna-se num passo crucial durante a análise (Fig. 6). Esta atribuição de pesos/ponderações pode ser feita utilizando um conjunto diferenciado de metodologias, mais ou menos convencionais, onde os valores numéricos mais elevados (dos pesos/ponderações) indicam uma maior importância relativa. Estas atribuições geralmente são feitas com base na experiência e conhecimento que o investigador detém sobre a área de estudo e do fenómeno em causa, sendo assim um procedimento bastante subjetivo e não isento de preconceitos potencialmente

prejudiciais. De forma a minimizar esta subjetividade, assim como os próprios preconceitos na atribuição destes pesos/ponderações, vários métodos quantitativos podem ser aplicados, incluindo as análises (geo-)estatísticas e determinísticas, os modelos probabilísticos e as abordagens livres de distribuição (entre outros).

Neste trabalho, de forma a realizar uma comparação capaz também de evidenciar eventuais vantagens e limitações na aplicação de métodos (ou combinação de métodos), utilizaram-se as seguintes e já referidas técnicas geoestatísticas: o Modelo Aditivo Generalizado (GAM); a Razão de Frequência Probabilística (RFP); e o processo de Análise Hierárquica por Pares (AHP). Três modelos especializados foram assim desenvolvidos para a área de estudo, no contexto operacional dos SIG, que, para além de conferir clareza e transparência aos processos analíticos, minimiza a incerteza e os riscos inerentes à tomada de decisão, permitindo soluções eficientes que espelham melhor a realidade através do estudo do problema com múltiplos critérios. O método GAM permitiu desenvolver um primeiro elaborado cartográfico preditivo (Modelo 1), enquanto a RFP e o processo de AHP combinaram-se para desenvolver mais dois cenários distintos (Modelo 2 e Modelo 3).

3.2.1 Modelo Aditivo Generalizado (GAM)

O método GAM foi usado para definir o grau de probabilidade de presença/ausência do fenómeno em estudo como base num rol de variáveis explicativas (contínuas ou ordinais). Trata-se de um método extremamente útil para modelar a probabilidade de ocorrência de um evento (*e.g.*, a preservação de sítios arqueológicos paleolíticos) em função da combinação entre as características de outros elementos contextuais (neste caso, os fatores ambientais condicionantes).

No caso deste trabalho, os fatores ambientais condicionantes previamente escolhidos (Tabela 1) correspondem às variáveis independentes, enquanto os sítios paleolíticos conhecidos representam a variável binária dependente. No processo de operacionalização do método GAM, é necessário identificar a presença e a ausência da variável dependente, pelo que, adicionalmente à identificação dos pontos correspondentes aos sítios arqueológicos conhecidos, foi também necessário definir a localização espacial de um conjunto de “não-sítios” arqueológicos. A identificação dos pontos que identificam os não-sítios foi obtida através de uma distribuição aleatória na área de estudo, considerando parâmetros de afastamento mínimo relativamente a cada um dos pontos e também aos sítios arqueológicos conhecidos. Importa,

contudo, referir que a identificação destes pontos de ausência do fenómeno não significa, necessariamente, uma ausência da efetiva presença de sítios arqueológicos preservados, apenas que esses não são conhecidos, pelo que se incluiu um número significativamente maior de pontos (não-sítios) de modo a diluir, tanto quanto possível, a possibilidade de se estar a contaminar o processo de análise com erros criados nesta fase (Kvamme, 1992). Assim, foi utilizada uma relação de 1/5 entre “sítios” e “não-sítios”, para ambas as amostras de treino e de teste. Selecionaram-se, no total, 200 localizações aleatórias de não-sítios, mas com uma distância mínima relativamente aos sítios paleolíticos conhecidos de 50 m. Igualmente, foram excluídas desta seleção aleatória as áreas urbanas existentes no território de estudo.

De uma forma muito geral, o modelo preditivo GAM examina a relação entre as variáveis independentes e a variável dependente (sítios/não-sítios), calculando o valor de probabilidade desta última em todos os *pixéis* do mapa (em SIG). A regressão linear standard utilizada pode ser genericamente descrita através desta simples equação:

$$Z = B_0 + \sum B_p X_p$$

onde Z representa a variável dependente, B_0 é uma constante, B_p é um coeficiente que mede o contributo de cada variável independente de entrada e X_p representa uma variável independente para cada variável na equação. No caso do método GAM esta função genérica da regressão linear é substituída por uma função do tipo:

$$Z = \sum f(x)$$

em que $f(x)$ é uma função não paramétrica (*i.e.*, cuja forma não é especificada), estimada através de curvas de alisamento (*e.g.*, *splines*).

Para a conversão dos resultados criados pelo GAM em termos de probabilidade, considera-se uma distribuição binomial para a variável dependente utilizando a seguinte equação:

$$P(k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

em que P é a probabilidade de sucesso numa tentativa individual, k corresponde ao número total de sucessos nas n tentativas independentes (com probabilidade de sucesso p^k) e

$(n - k)$ corresponde aos fracassos nessas tentativas [com probabilidade de fracasso correspondente a $(1 - p)^{n-k}$].

Para a implementação desta análise foi utilizada a ferramenta “*Marine Geospatial Ecology Tools*” (<https://mgel.env.duke.edu/mget/> - último acesso julho 2021), integrada na *toolbox* do ArcGIS, tendo sido realizadas múltiplas execuções com ligeiras variações no ajustamento do GAM, a fim de verificar a consistência dos resultados. O produto final corresponde a um *raster* onde a cada *pixel* da área de estudo está associado o valor de probabilidade de preservação do registo Paleolítico (Modelo 1) - valores elevados indicam uma maior probabilidade. O método das quebras naturais foi utilizado para a categorização final deste modelo em 5 classes (muito alta, alta, moderada, baixa e muito baixa) e assim produzir o primeiro mapa de probabilidade de preservação do registo Paleolítico da área de estudo, em formado *raster* e com resolução espacial de 10 m x 10 m de *pixel* (*vide* Fig. 15).

3.2.2 Razão de Frequência Probabilística (RFP)

Amplamente utilizada como técnica de AMC nas “Ciências da Terra”, com particular ênfase para a avaliação da suscetibilidade aos movimentos de materiais em vertente, aluviões e incêndios florestais (Cunha e Dimuccio, 2002; Dimuccio *et al.*, 2005, 2008, 2011, 2020; Cunha *et al.*, 2017; Nicu, 2018; Khoshtinat *et al.*, 2019; Tehrany *et al.*, 2019; entre outros), a RFP tem sido pouco aplicada na modelação espacial preditiva em arqueologia (Aubry *et al.*, 2012, 2017b; Nicu *et al.*, 2019). Trata-se de uma técnica estatística que se presta bem à utilização de dados espacializados, permitindo que todas as variáveis previamente escolhidas e classificadas sejam correlacionadas com as evidências no terreno (neste caso, os sítios arqueológicos paleolíticos conhecidos) de forma a priorizar, dentro de cada uma, a importância das respetivas classes/categorias.

Neste trabalho, a relação entre o número de sítios paleolíticos conhecidos e as classes/categorias dos fatores ambientais condicionantes foi matematicamente estabelecida. Resumidamente, o método da razão de frequência probabilística foi utilizado para associar um peso de importância relativa às classes/categorias de cada variável independente (fator) com base na sua relação espacial com a ocorrência dos sítios paleolíticos conhecidos (*i.e.*, a razão entre a frequência do número de sítios e a frequência da área em *pixéis*, para cada classe/categoria). Esta razão expressa a relevância proporcional entre a localização dos sítios conhecidos e as áreas “sem sítios” conhecidos - o valor 1 indica proporções iguais e é utilizado

como marco de referência. Um valor desta razão > 1 indica que aquela classe/categoria é mais importante para a preservação dos sítios paleolíticos, enquanto um valor < 1 aponta para uma importância relativa dessa classe/categoria menos significativa. A média e o desvio padrão dos valores desta razão foram calculados, para cada fator, de forma a poder aplicar-se um procedimento de normalização capaz de facilitar a interpretação. Os valores normalizados foram transformados numa escala ordinal entre 0 e 100, representando assim os “pesos finais” (r_i) para cada classe/categoria de fator. O método da RFP foi aplicado duas vezes: numa primeira abordagem utilizaram-se todos os 40 sítios paleolíticos conhecidos na área de estudo (Tabela 2), enquanto na segunda usaram-se só os 18 sítios da área de treino (Tabela 3).

Tabela 2 - Procedimentos para a aplicação da RFP utilizando todos os 40 sítios paleolíticos conhecidos da área de estudo, na assinação dos pesos de importância relativa as classes/categorias de cada fator ambiental condicionante (r_i = pesos finais normalizados e transformados numa escala ordinal com valores entre 0 e 100).

Camadas <i>rasters</i> derivadas (Fatores condicionantes)	Classes (ou categorias)	Número de sítios na área de estudo	Frequência número de sítios na área de estudo (%)	Área de estudo em pixels	Frequência área de estudo em pixels (%)	Cálculo dos pesos				Pesos finais (área de estudo)
						Razão (ab^{-1})	Média (m)	Desvio padrão (σ)	Normalização ($[(ab^{-1} - m)\sigma^{-1}]$)	
Declive (DEC)	1 (0º - 3º)	13	32,5	1969198	55,2	0,6	0,9	0,7	-0,5	35
	2 (3º - 7º)	17	42,5	903083	25,3	1,7			1,1	100
	3 (7º - 12º)	8	20,0	478833	13,4	1,5			0,8	89
	4 (12º - 20º)	2	5,0	200326	5,6	0,9			-0,1	53
	5 (> 20º)	0	0,0	14663	0,4	0,0			-1,4	0
Índice de posição topográfica (IPT)	1 (Vale)	1	2,5	68662	1,9	1,3	1,4	0,6	-0,2	45
	2 (Encosta inferior)	7	17,5	345454	9,7	1,8			0,7	79
	3 (Superfície plana)	10	25,0	1405696	39,4	0,6			-1,3	0
	4 (Encosta média)	13	32,5	1359828	38,1	0,9			-1,0	15
	5 (Encosta superior)	8	20,0	337440	9,5	2,1			1,2	100
	6 (Crista)	1	2,5	49023	1,4	1,8			0,7	80
Distância ponderada de custo para acesso aos cursos de água (DPCCA)	1 (Baixa)	24	60,0	1343307	37,7	1,6	0,4	0,6	2,2	100
	2	10	25,0	944155	26,5	0,9			1,0	59
	3	5	12,5	556215	15,6	0,8			0,8	50
	4	1	2,5	323809	9,1	0,3			-0,2	17
	5	0	0,0	201065	5,6	0,0			-0,6	0
	6	0	0,0	116296	3,3	0,0			-0,6	0
	7	0	0,0	59661	1,7	0,0			-0,6	0
	8	0	0,0	19013	0,5	0,0			-0,6	0
	9	0	0,0	2565	0,1	0,0			-0,6	0
	10 (Alta)	0	0,0	17	0,0	0,0			-0,6	0
Índice de carga de calor (ICC)	1 (Muito baixo)	0	0,0	74575	2,1	0,0	1,0	0,7	-1,5	0
	2 (Baixo)	5	12,5	340900	9,6	1,3			0,4	73
	3 (Moderado)	24	60,0	2334345	65,5	0,9			-0,2	51
	4 (Elevado)	8	20,0	667140	18,7	1,1			0,1	60
	5 (Muito elevado)	3	7,5	149143	4,2	1,8			1,2	100
Litologia (LITO)	1 (Aluviões - Holocénico)	4	10,0	381364	10,7	0,9	3,9	10,6	-0,3	3
	2 (Areias de praia - Holocénico)	0	0,0	0	0,0	0,0			-0,4	0
	3 (Depósitos de vertente - Holocénico)	0	0,0	0	0,0	0,0			-0,4	0
	4 (Dunas e areias eólicas indiferenciadas - Plistocénico a Holocénico)	0	0,0	152859	4,3	0,0			-0,4	0
	5 (Depósitos de terraços fluviais - Plistocénico)	36	90,0	106323	3,0	30,2			2,5	100
	6 (Tufo calcários - Plistocénico)	0	0,0	0	0,0	0,0			-0,4	0
	7 (Depósitos de terraços marinhos - Plistocénico)	0	0,0	0	0,0	0,0			-0,4	0
	8 (Substrato pré-Plistocénico)	0	0,0	2925557	82,0	0,0			-0,4	0

Tabela 3 - Procedimentos para a aplicação da RFP utilizando só os 18 sítios paleolíticos conhecidos da área de treino, na assinação dos pesos de importância relativa as classes/categorias de cada fator ambiental condicionante (r_i = pesos finais normalizados e transformados numa escala ordinal com valores entre 0 e 100).

Camadas <i>rasters</i> derivadas (Fatores condicionantes)	Classes (ou categorias)	Número de sítios na área de treino	Frequência número de sítios na área de treino (%)	Área de treino em pixels	Frequência área de treino em pixels (%)	Cálculo dos pesos				Pesos finais (área de treino)
						Razão (ab^{-1})	Média (m)	Desvio padrão (σ)	Normalização ($[(ab^{-1} - m)\sigma^{-1}]$)	
Declive (DEC)	1 (0º - 3º)	7	38,9	272637	58,4	0,7	1,0	0,7	-0,5	43
	2 (3º - 7º)	7	38,9	126481	27,1	1,4			0,6	92
	3 (7º - 12º)	3	16,7	49677	10,6	1,6			0,8	100
	4 (12º - 20º)	1	5,6	17279	3,7	1,5			0,7	96
	5 (> 20º)	0	0,0	998	0,2	0,0			-1,5	0
Índice de posição topográfica (IPT)	1 (Vale)	0	0,0	3740	0,8	0,0	1,1	1,3	-0,9	0
	2 (Encosta inferior)	3	16,7	38935	8,3	2,0			0,7	60
	3 (Superfície plana)	5	27,8	198990	42,6	0,7			-0,4	20
	4 (Encosta média)	6	33,3	190640	40,8	0,8			-0,2	25
	5 (Encosta superior)	4	22,2	31339	6,7	3,3			1,7	100
	6 (Crista)	0	0,0	3428	0,7	0,0			-0,9	0
Distância ponderada de custo para acesso aos cursos de água (DPCCA)	1 (Baixa)	14	77,8	277062	59,3	1,3	0,2	0,4	2,4	100
	2	3	16,7	130745	28,0	0,6			0,8	45
	3	1	5,6	44203	9,5	0,6			0,8	45
	4	0	0,0	13617	2,9	0,0			-0,6	0
	5	0	0,0	1440	0,3	0,0			-0,6	0
	6	0	0,0	1	0,0	0,0			-0,6	0
	7	0	0,0	1	0,0	0,0			-0,6	0
	8	0	0,0	1	0,0	0,0			-0,6	0
	9	0	0,0	1	0,0	0,0			-0,6	0
	10 (Alta)	0	0,0	1	0,0	0,0			-0,6	0
Índice de carga de calor (ICC)	1 (Muito baixo)	0	0,0	4472	1,0	0,0	1,1	0,7	-1,5	0
	2 (Baixo)	2	11,1	37609	8,1	1,4			0,4	73
	3 (Moderado)	10	55,6	326257	69,9	0,8			-0,4	42
	4 (Elevado)	5	27,8	85022	18,2	1,5			0,6	81
	5 (Muito elevado)	1	5,6	13712	2,9	1,9			1,0	100
Litologia (LITO)	1 (Aluviões - Holocénico)	2	11,1	104337	22,3	0,5	0,8	2,1	-0,1	8
	2 (Areias de praia - Holocénico)	0	0,0	1	0,0	0,0			-0,4	0
	3 (Depósitos de vertente - Holocénico)	0	0,0	1	0,0	0,0			-0,4	0
	4 (Dunas e areias eólicas indiferenciadas - Plistocénico a Holocénico)	0	0,0	1	0,0	0,0			-0,4	0
	5 (Depósitos de terraços fluviais - Plistocénico)	16	88,9	69180	14,8	6,0			2,5	100
	6 (Tufo calcários - Plistocénico)	0	0,0	1	0,0	0,0			-0,4	0
	7 (Depósitos de terraços marinhos - Plistocénico)	0	0,0	1	0,0	0,0			-0,4	0
	8 (Substrato pré-Plistocénico)	0	0,0	293550	62,8	0,0			-0,4	0

3.2.3 Análise Hierárquica por Pares (AHP)

A ponderação dos cinco fatores ambientais condicionantes previamente escolhidos (Tabela 1), para utilizar na AMC, foi feita através de uma matriz de comparação por pares (método qualitativo), como descrito em Saaty (1988, 1990, 2008) e Saaty e Vargas (2001). Trata-se de uma técnica de AMC que utiliza uma escala de pontuação que varia entre 1 a 9 (ou 1 a 1/9), para classificar a importância relativa de cada critério com base numa comparação par-a-par (Malczewski, 1999). Com o intuito de facilitar esta comparação, uma expressão linguística é atribuída a cada valor da escala de pontuação (Tabela 4), onde, por exemplo, ao número 1 corresponde à associação de uma igual importância entre os critérios na explicação do fenómeno em causa, enquanto o número 9 (ou 1/9) descreve a situação em que um dos critérios é absolutamente mais (ou menos) importante do que o outro. A pontuação global para cada critério é determinada através da soma ponderada da pontuação que se obteve em relação a todos os critérios.

Quando se utiliza esta abordagem, é comumente aceite que, ao considerar os valores numéricos associados às expressões linguísticas, uma terminologia imprecisa pode criar grandes ambiguidades em relação aos resultados alcançados. Os valores numéricos são traduções quantificadas úteis para ponderar as variáveis e a validade dessa valoração pode ser avaliada pelas ponderações e a consistência do processo de cálculo (Ayalew *et al.*, 2004; Kayastha, 2013). Ainda assim, a comparação par-a-par carrega sempre uma grande componente de subjetividade, sendo a qualidade dos resultados extremamente dependente do julgamento pessoal que o investigador (ou grupo de especialistas) faz durante este exercício (*e.g.*, De Steiguer *et al.*, 2003; Davidsson *et al.*, 2005).

Tabela 4 - Escala de pontuação do processo de AHP proposta em Saaty (1988).

Valor de importância de cada fator (critério)	Expressão linguística	Explicação
1	Mesma importância	Os dois critérios contribuem de igual forma para o objetivo.
3	Importância moderada	A experiência e o julgamento favorecem moderadamente um critério em relação ao outro.
5	Importância essencial ou forte	A experiência e o julgamento favorecem fortemente um critério em relação ao outro.
7	Importância muito forte	Um critério é fortemente favorecido e a sua dominância é demonstrada na prática.
9	Importância extrema	A evidência favorece um critério sobre o outro com um elevado grau de certeza.
2, 4, 6, 8	Valores associados a julgamentos intermédios	Quando se deseja um maior compromisso.

O processo de AHP foi integrado na modelação espacial preditiva através de uma estrutura com múltiplos níveis (Fig. 14). Primeiro definiu-se uma meta ou objetivo geral a alcançar e, depois, consideraram-se os objetivos específicos de cada um dos critérios (*i.e.*, variáveis explicativas) relativamente à sua influência na presença/ausência de sítios arqueológicos que, no caso concreto deste trabalho, são representadas pelos fatores ambientais condicionantes. O último nível corresponde às escolhas de alternativas de localização. As linhas entre os níveis indicam as relações entre os fatores, as escolhas e o objetivo a alcançar. No caso específico deste estudo só foi utilizado um nível de subcritério caracterizado por uma matriz de comparação com os cinco fatores ambientais condicionantes previamente escolhidos (DEC, IPT, DPCCA, ICC e LITO). Desta forma, esta matriz de comparação apresenta-se com uma dimensão de 5 x 5, considerando que cada escolha está conectada a cada fator.

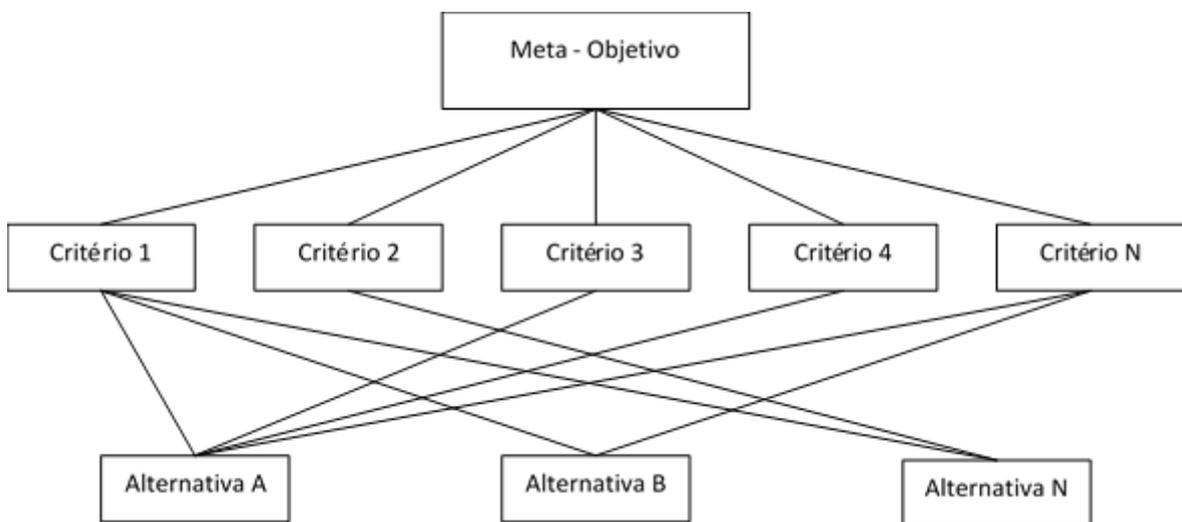


Figura 14 - Estrutura de um processo de AHP onde se reconhecem três diferentes níveis (Saaty, 1990) (mais detalhes no texto).

A matriz de comparação por pares $A = [a_{ij}]_{n \times n}$ pode ser expressa matematicamente desta forma:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ a_{21} & a_{22} & a_{2n} \\ a_{n1} & a_{n2} & a_{nn} \end{bmatrix}$$

Esta matriz tem uma propriedade recíproca que é expressa matematicamente como:

$$a_{ji} = \frac{1}{a_{ij}}$$

Uma vez construída a matriz de comparação par-a-par entre os fatores considerados, e de forma a formalizar esta comparação em termos aritméticos precisos, torna-se necessário operacionalizar estes juízos através da determinação do vetor-próprio principal normalizado da matriz de relação $w = [w_1, w_2, w_3 \dots, w_n]$ de acordo com o método linear de Saaty (Goepel, 2013). Este processo de ponderação inclui duas etapas: primeiro a matriz de correlação (A) é normalizada e, de seguida, as ponderações para cada critério (w_i) são calculadas utilizando as seguintes equações, respetivamente:

$$a_{ij} = \frac{a_{ij}}{\sum_{i=1}^n a_{ij}}$$

para todos os $j = 1, 2, 3 \dots, n$

$$w_i = \frac{\sum_{i=1}^n a_{ij}}{n}$$

e para todos os $i = 1, 2, 3, \dots, n$

A Tabela 5 mostra os resultados da matriz de comparação par-a-par e as ponderações finais dos fatores ambientais condicionantes previamente escolhidos para este trabalho. O índice de consistência (IC) desta matriz foi determinado utilizando a seguinte fórmula:

$$IC = \frac{(\lambda_{max} - n)}{(n - 1)}$$

onde λ_{max} representa o valor mais elevado ou principal da matriz (neste trabalho assume o valor de 5.012) e n corresponde a ordem da matriz (*i.e.*, o número total de critérios utilizados), sendo aqui igual a 5. A qualidade (ou grau de consistência) da comparação é descrita através da razão de consistência (RC), calculada utilizando a seguinte fórmula:

$$RC = \frac{IC}{IA}$$

onde IA é um índice de consistência calculado para uma matriz gerada aleatoriamente, com ordens que variam entre 1 e 15, e que, em função da ordem da matriz utilizada e de acordo com Saaty (1988), assume neste trabalho um valor de 1.12. Para a comparação ser satisfatória/consistente, Saaty (1988) recomenda um valor de RC inferior a 0.10. A RC neste estudo foi de 0.003, apontando para uma matriz de julgamentos razoavelmente consistente e, portanto, utilizável nos processos de AMC para ponderar os fatores ambientais condicionantes. O índice de consistência geométrica (ICG = 0.01) e a inconsistência ordinal (ou dissonância - Psi = 0.0%) também foram avaliadas, na elaboração desta matriz, de forma a traduzir o erro de julgamento dos valores numérico da classificação par-a-par realizada (sendo considerados aceitáveis valores até 0.1). O processo de AHP foi aqui implementado utilizando a Folha Excel de Goepel (2013) (versão 15.09.2018 - disponível para download em <http://bpmmsg.com>).

Tabela 5 - Resultados da matriz de comparação par-a-par e ponderações finais (w_i) dos fatores ambientais condicionantes utilizados neste trabalho. DEC = Declive; IPT = Índice de posição topográfica; DPCCA = distância ponderada de custo para acesso aos cursos de água; ICC = índice de carga de calor; LITO = Litologia.

	DEC	IPT	DPCCA	ICC	LITO	Ponderação final (w_i)
DEC	1					21
IPT	1/4	1				5
DPCCA	1	4	1			22
ICC	1/3	1	1/4	1		6
LITO	2	9	2	8	1	45

Razão de consistência (RC) = 0.003

Índice de consistência geométrica (ICG) = 0.01

Inconsistência ordinal (ou dissonância - Psi) = 0.0%

Os resultados finais de ambos os procedimentos que combinam a RFP com a AHP correspondem a dois mapas *rasters* que resultam da sobreposição espacial dos fatores ambientais condicionantes, selecionados da Tabela 1, aplicando a seguinte equação no “*raster calculator*” em ArcGIS:

$$\text{Índice preditivo} = \sum r_i w_i$$

onde r_i indica o peso das classes/categorias dos cinco fatores ambiental condicionantes, determinados através da RFP, e w_i é a ponderação calculada para cada um destes fatores, utilizando o processo de AHP. Este índice preditivo representa, para cada *pixel* da área de estudo, o valor de probabilidade de preservação do registo Paleolítico (Modelos 2 e 3) - valores

elevados indicam uma maior probabilidade. O método das quebras naturais foi utilizado para a categorização final deste índice em 5 classes (muito alta, alta, moderada, baixa e muito baixa) e assim produzir o segundo e terceiro mapa de probabilidade de preservação do registo Paleolítico da área de estudo, em formato *raster* e com uma resolução espacial de 10 m x 10 m de *pixel* (*vide* Figs. 17 e 19).

3.3 Avaliação da capacidade preditiva dos modelos

Depois da elaboração dos três modelos cartográficos preditivos, para a área de estudo, os respetivos resultados foram avaliados através das curvas ROC (*receiver operating characteristic*), juntamente com a métrica associada e correspondente a AUC (*area under the curve*), criadas para cada um deles. Este é um método gráfico simples e robusto que evidencia a variação de sensibilidade e especificidade dos modelos (Braga, 2001). Na prática, todos os valores de probabilidade calculados para cada *pixel* da área de estudo são classificados em ordem decrescente e divididos em 100 classes com intervalo cumulativo constante de 1%. A curva ROC resulta da tabulação cruzada entre o resultado cartográfico preditivo e os dados arqueológicos de teste (*i.e.*, os 22 sítios não utilizados nas fases de treino de atribuição dos pesos). Para comparar os resultados quantitativamente, a AUC é calculada com base numa área total igual a 1 ou 100% (indicando uma concordância perfeita). Para julgar os valores da AUC, o seguinte raciocínio foi sugerido por Sweets (1988) e sucessivamente utilizado por vários autores (Araújo *et al.*, 2005; Noviello *et al.*, 2018): excelente quando $AUC > 90\%$; bom quando $90\% < AUC < 80\%$; razoável quando $80\% < AUC < 70\%$; medíocre quando $70\% < AUC < 60\%$; falhou quando $AUC < 60\%$. Outros autores simplesmente consideram que um modelo com valor de $AUC > 75\%$ apresenta uma capacidade preditiva aceitável (Howey *et al.*, 2016) e que, embora evidenciam trata-se de uma medida bidimensional de classificação do desempenho bastante simples e eficaz, a sua aplicação é útil para discriminar entre “bons” e “maus” modelos, mas não necessariamente entre “bons” modelos (Marzban, 2004).

As curvas ROC/AUC foram processadas utilizando o ROCR package in R (Sing *et al.*, 2005) para o modelo desenvolvido utilizando o método GAM e através uma simples folha de cálculos em Excel para os restantes dois modelos.

Para além disso, e como meio de verificação da robustez dos três modelos cartográficos preditivos desenvolvidos, calculou-se também o ganho (*G*) de cada modelo de acordo com a equação de Kvamme (1988):

$$G = 1 - \frac{\%PS}{\%OS}$$

onde *PS* corresponde ao total da superfície ocupada pelas zonas com muito alta probabilidade de preservação do registo Paleolítico e *OS* representa o número de sítios paleolítico conhecidos que caem nas zonas com muito alta probabilidade. O *G* exprime, de modo expedito, a capacidade do modelo em prever a localização de sítios arqueológicos preservados com uma precisão aceitável. O valor de *G* varia entre 1 e -1, sendo que valores positivos sugerem que o modelo é robusto e, portanto, pode ser aceite.

4 RESULTADOS

4.1 Modelo 1

A Figura 15 mostra os resultados cartográficos do primeiro modelo preditivo desenvolvido para a área de estudo (Modelo 1), onde se identificam espacialmente as áreas com maior ou menor potencial à preservação do registo Paleolítico com base na aplicação do método GAM.

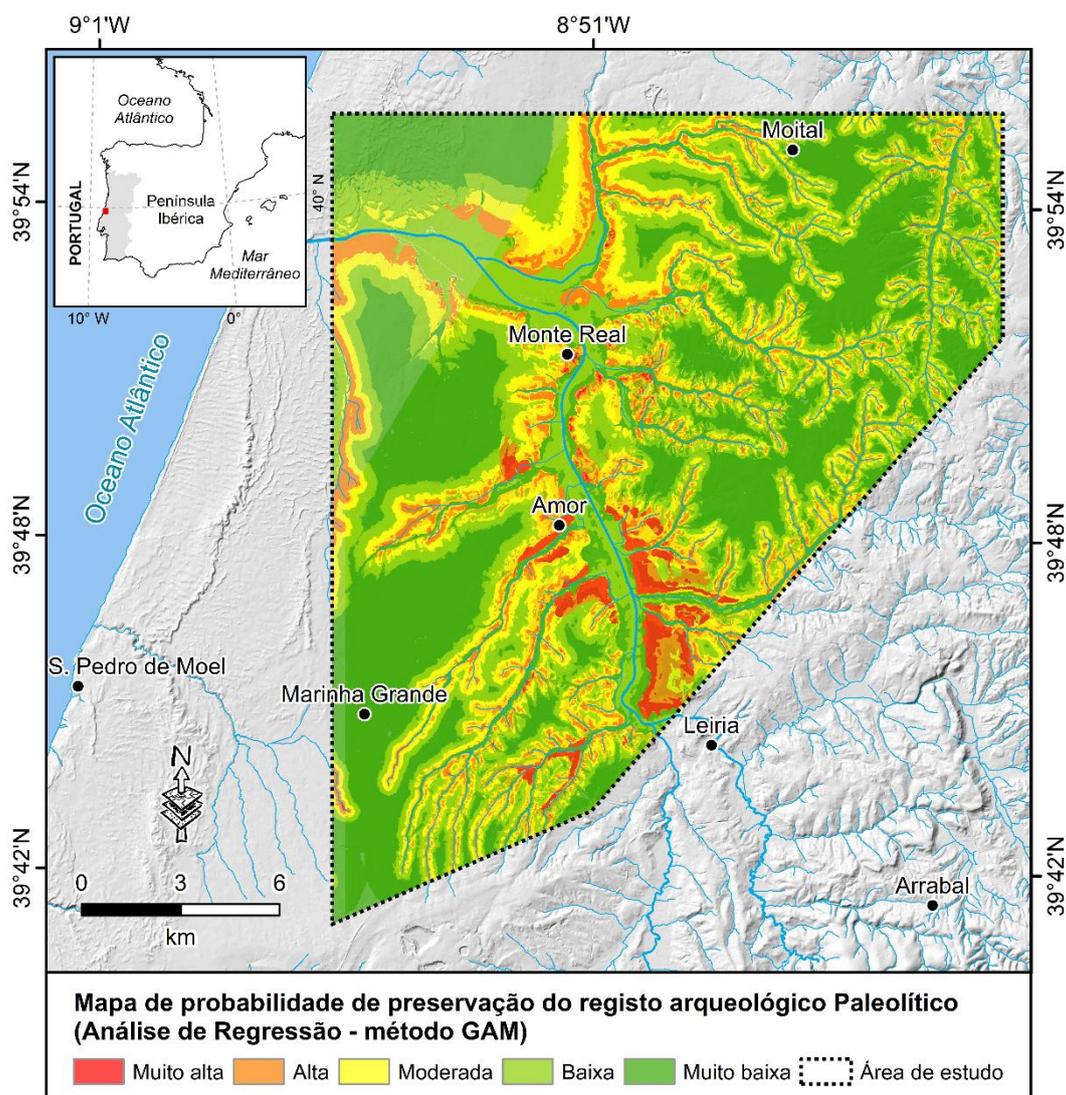


Figura 15 - Resultados da modelação espacial preditiva de identificação do potencial arqueológico Paleolítico na região de Leiria, utilizando o método GAM no âmbito da AR (Modelo 1).

Neste primeiro cenário, a maioria da área de estudo apresenta valores muito baixos (33.3%) a baixos (32.4%) de probabilidade de preservação do registo Paleolítico. Pouco mais de ¼ da área de estudo apresenta valores moderados (18.3%), altos (13.2%) a muito altos (2.8%). O resultado cartográfico da Figura 15 mostra uma maior probabilidade de se encontrarem registos arqueológicos paleolíticos preservados perto de ambas as margens do Rio Lis e seus tributários, com uma clara correspondência com os afloramentos que identificam, na área de estudo, os depósitos de terraços fluviais do Plistocénico (Fig. 13). Outras importantes porções de território com alta a muito alta probabilidade de preservação do registo Paleolítico correspondem a numerosas e pequenas manchas alongadas, paralelamente à rede hidrográfica principal (rio Lis e ribeiras afluentes) e, numa posição lateral, reduzidas manchas, onde a cartografia geológica oficial (pelo menos a escala de 1:50 000) não mostra evidentes afloramentos de depósitos de terraços fluviais. Para além disso, os resultados desta primeira modelação permitem evidenciar como as zonas com alta a muito alta probabilidade de preservação do registo Paleolítico correspondem as porções do território com as seguintes características: 84% com declive nulo ou muito pouco acentuado (0-7º), 15% com declive entre 7-20º e 1% com declive > 20º; 46% são encostas médias, 26% encostas superiores, 19% superfícies planas e 7% cristas, sendo os restantes 2% encostas inferiores ou vales; 64% das vertentes apresentam um índice de carga de calor moderado, 26% elevado, 6% baixo e os restantes 5% muito elevado a muito baixo.

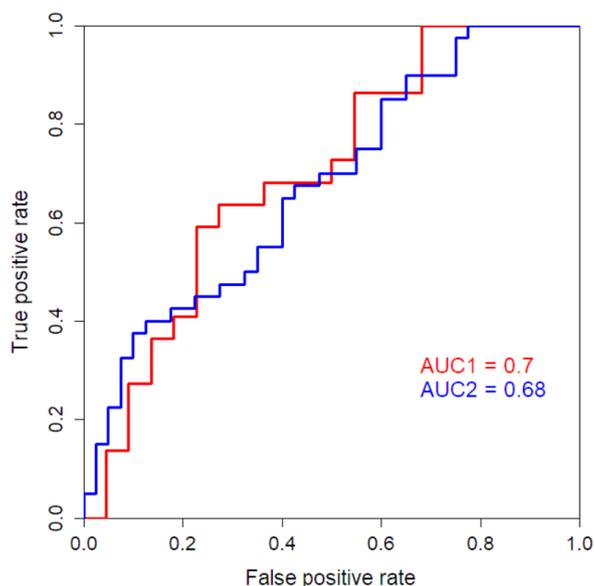


Figura 16 - Avaliação do Modelo 1 através das curvas ROC/AUC. A curva ROC1/AUC1 (em vermelho) mostra a relação entre a localização dos 22 sítios arqueológicos conhecidos na área de teste, não usados nas fases de treino de atribuição dos pesos, com a classificação do índice de probabilidade preditiva do Modelo 1 (este último calculado utilizando o método GAM). Por sua vez, a curva ROC2/AUC2 (em azul) mostra a relação entre a localização de todos os 40 sítios arqueológicos conhecidos da área de estudo com a classificação do índice de probabilidade preditiva do Modelo 1 (mais detalhes no texto).

A comparação espacial, em SIG, entre o *raster* obtido utilizando o método GAM (Fig. 15) e a localização dos sítios arqueológicos conhecidos da área de teste (*i.e.*, os 22 sítios não usados nas fases de atribuição dos pesos) permitiu avaliar a capacidade preditiva do Modelo 1 através da elaboração da curva ROC1 (Fig. 16). A AUC1 foi então utilizada para quantificar o desempenho do Modelo 1, apontando para um valor de 0.7 (= 70%), portanto, ligeiramente abaixo do limite de aceitabilidade. No entanto, o ganho (*G*) deste Modelo 1 é de 0.22, sugerindo que este cenário tem condições para ser aceite. Além disso, aponta-se uma taxa de sucesso do Modelo 1 de 68% (curva ROC2/AUC2 na Fig. 16) calculada através da comparação espacial entre o *raster* da Figura 15 e a localização de todos os 40 sítios arqueológicos conhecidos da área de estudo. A curva ROC2 demonstra que com apenas 10% da área de estudo classificada como muito alta probabilidade de preservação do registo arqueológico paleolítico, no Modelo 1, se conseguem enquadrar só o 35% dos sítios arqueológicos conhecidos. Quando se consideram os 20% da área de estudo classificada como alta a muito alta probabilidade verifica-se que ela integra apenas 40% dos sítios.

4.2 Modelo 2

A Figura 17 mostra os resultados cartográficos do segundo modelo preditivo desenvolvido para a área de estudo (Modelo 2), onde se identificam espacialmente as áreas com maior ou menor potencial à preservação do registo Paleolítico com base na aplicação do método da RFP, usando todos os sítios arqueológicos paleolíticos conhecidos na área de estudo (Tabela 2), em combinação com o processo de AHP (Tabela 5).

Neste segundo cenário, a maioria da área de estudo apresenta valores moderados (35.7%) a baixos (30.1%) de probabilidade de preservação do registo Paleolítico. O restante território apresenta valores altos (16.5%), muito baixos (14.7%) a muito altos (3.0%). Mais uma vez, o resultado desta segunda modelação cartográfica mostra uma maior probabilidade de se encontrarem registos arqueológicos paleolíticos preservados perto de ambas as margens do Rio Lis e seus tributários (Fig. 17), com uma clara correspondência com os afloramentos que identificam, na área de estudo, os depósitos de terraços fluviais do Plistocénico (Fig. 13). Outras importantes porções de território com alta probabilidade de preservação do registo Paleolítico correspondem a faixas alongadas (agora mais contínuas, quando comparadas com o resultado do modelo anterior). Igualmente, em posições paralelas à rede hidrográfica principal (rio Lis e ribeiras afluentes), surgem machas muito recortadas, por vezes, descontínuas e classificadas

como de alta probabilidade. Mais uma vez, isso poderá estar relacionado com a existência de restos de depósitos de terraços fluviais não representados na escala original da cartografia geológica oficial utilizada. Para além disso, os resultados desta segunda modelação permitem evidenciar como zonas com alta a muito alta probabilidade de preservação do registo Paleolítico correspondem as porções do território com as seguintes características: 69% com declive nulo ou muito pouco acentuado (0-7º), 31% com declive entre 7-20º e 0% com declive > 20º; 32% são superfícies planas, 27% encostas médias, 19% encostas inferiores e 16% encostas superiores, sendo os restantes 5% e 1% vales e cristas, respetivamente; 40% das vertentes apresentam um índice de carga de calor elevado, 36% moderado, 19% baixo e os restantes 4% e 1% muito elevado a muito baixo, respetivamente.

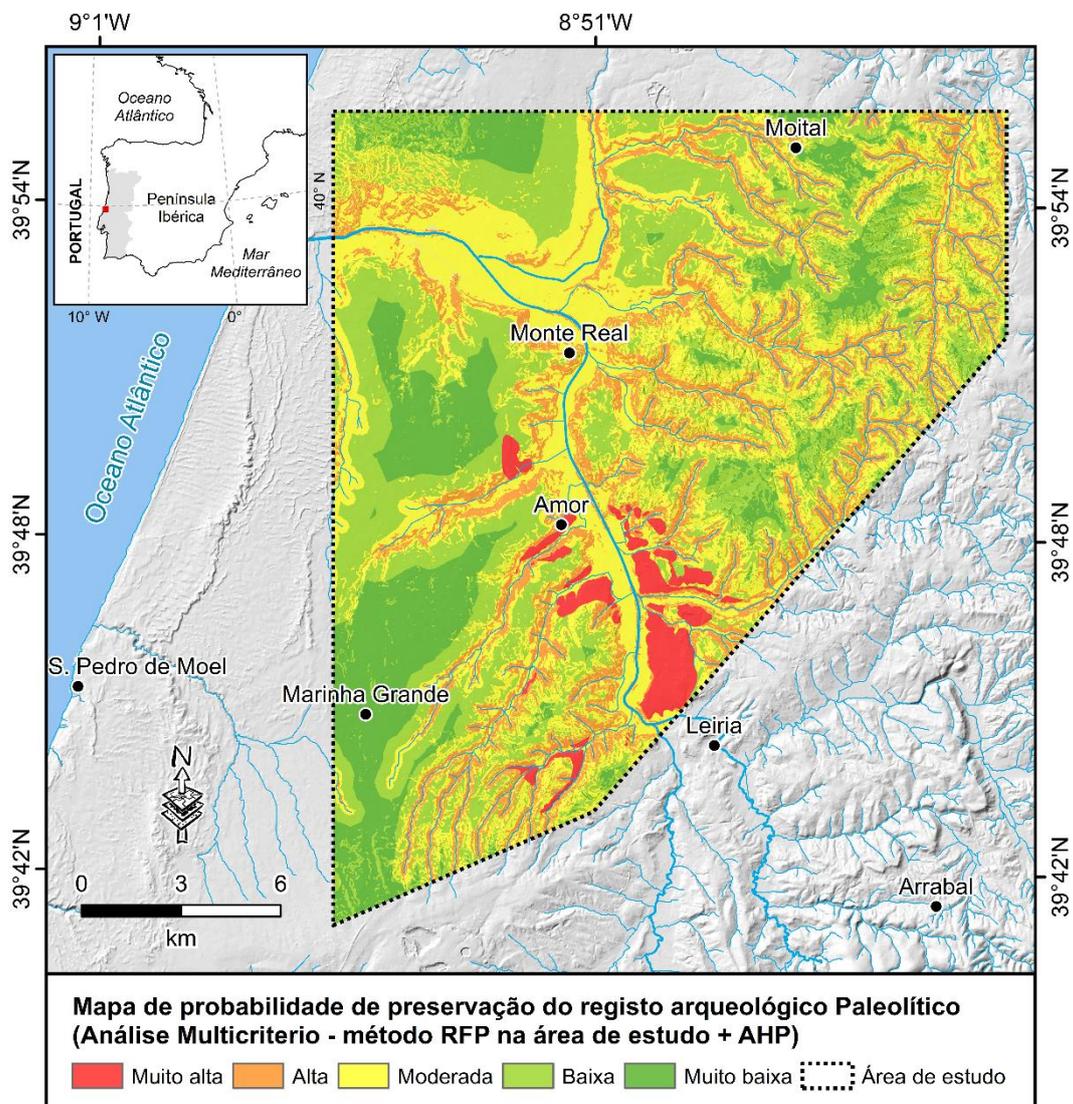


Figura 17 - Resultados da modelação espacial preditiva de identificação do potencial arqueológico Paleolítico na região de Leiria, utilizando o método da RFP (com todos os 40 sítios paleolíticos conhecidos na área de estudo) em combinação com o processo de AHP no âmbito da AMC (Modelo 2).

A comparação espacial, em SIG, entre o *raster* obtido utilizando o método da RFP em combinação com o processo de AHP (Fig. 17) e a localização de todos os 40 sítios arqueológicos conhecidos da área de estudo, permitiu avaliar a taxa de sucesso do Modelo 1 através da elaboração da respetiva curva ROC (Fig. 18). Aponta-se para o Modelo 2 uma taxa de sucesso de 87% (curva ROC3/AUC3 na Fig. 18). O ganho (*G*) deste Modelo 2 resulta ser de 0.17, sugerindo que pode ser aceite. A curva ROC3 demonstra que com apenas 10% da área de estudo classificada como muito alta probabilidade de preservação do registo arqueológico paleolítico, no Modelo 2, se conseguem enquadrar só 42% dos sítios arqueológicos conhecidos. Quando se consideram os 20% da área de estudo classificada como alta a muito alta probabilidade verifica-se que ela integra pouco mais de 80% dos sítios.

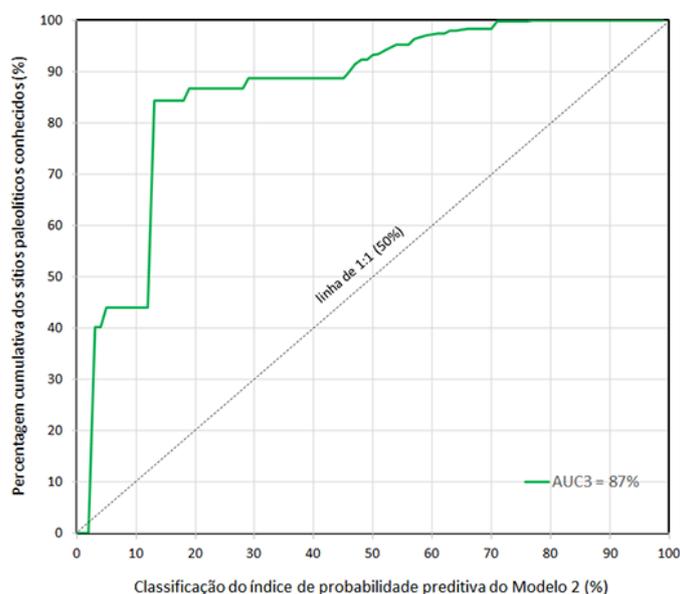


Figura 18 - Avaliação do Modelo 2 através das curvas ROC/AUC. No eixo horizontal do gráfico, o índice de probabilidade preditiva está cumulativamente indicado por ordem decrescente. A curva ROC3/AUC3 (em verde) mostra a relação entre a localização de todos os 40 sítios arqueológicos conhecidos da área de estudo com a classificação do índice de probabilidade preditiva do Modelo 2 (este último calculado utilizando o método da RFP - na área de estudo - em combinação com o processo de AHP) (mais detalhes no texto).

4.3 Modelo 3

A Figura 19 mostra os resultados cartográficos do terceiro e último modelo preditivo desenvolvido para a área de estudo (Modelo 3), onde se identificam espacialmente as áreas com maior ou menor potencial à preservação do registo Paleolítico com base na aplicação do método

da RFP, usando só os sítios arqueológicos paleolíticos conhecidos na área de treino (Tabela 3), em combinação com o processo de AHP (Tabela 5).

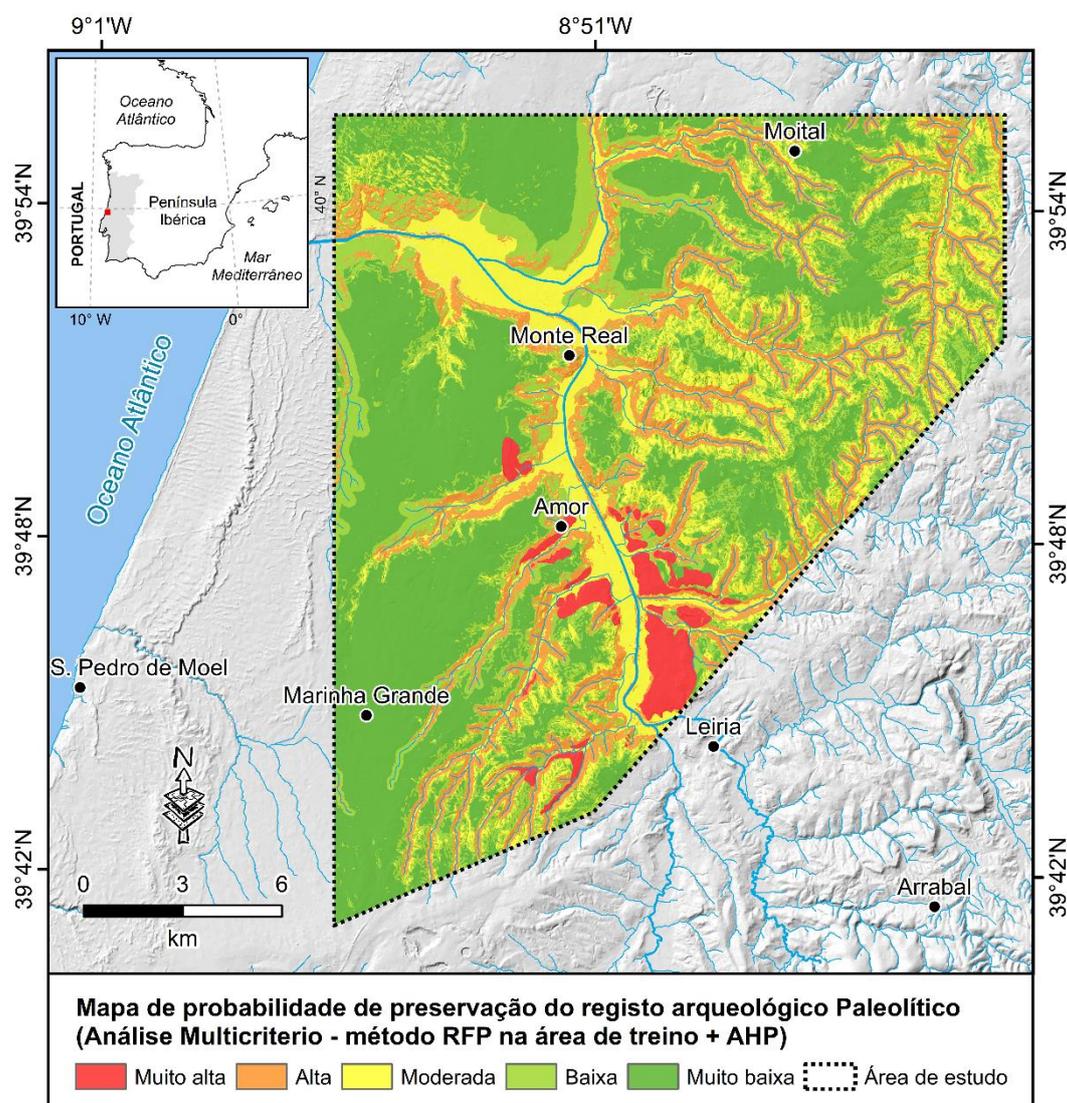


Figura 19 - Resultados da modelação espacial preditiva de identificação do potencial arqueológico Paleolítico na região de Leiria, utilizando o método da RFP (só com os 18 sítios paleolíticos conhecidos da área de treino) em combinação com o processo de AHP no âmbito da AMC (Modelo 3).

Neste terceiro cenário, a maioria da área de estudo apresenta valores muito baixos (38.3%) a baixos (22.2%) de probabilidade de preservação do registo Paleolítico. O restante território apresenta valores moderados (18.5%), altos (18.0%) a muito altos (3.0%). Também neste caso, o resultado cartográfico mostra uma maior probabilidade de se encontrarem registos arqueológicos paleolíticos preservados perto de ambas as margens do Rio Lis e seus tributários (Fig. 19), com uma clara correspondência com os afloramentos que identificam, na área de estudo, os depósitos de terraços fluviais do Plistocénico (Fig. 13). Outras importantes

porções de território com alta probabilidade de preservação do registo Paleolítico correspondem a faixas alongadas e contínuas, paralelamente à rede hidrográfica principal. Para além disso, os resultados desta última modelação permitem evidenciar como as zonas com alta a muito alta probabilidade de preservação do registo Paleolítico correspondem as porções do território com as seguintes características: 57% com declive nulo ou muito pouco acentuado (0-7º), 43% com declive entre 7-20º e 0% com declive > 20º; 33% são superfícies planas, 28% encostas médias, 23% encostas inferiores e 10% encostas superiores, sendo os restantes 5% e 1% vales e cristas, respetivamente; 36% das vertentes apresentam um índice de carga de calor elevado, 33% moderado, 18% baixo e os restantes 9% e 4% muito elevado a muito baixo, respetivamente.

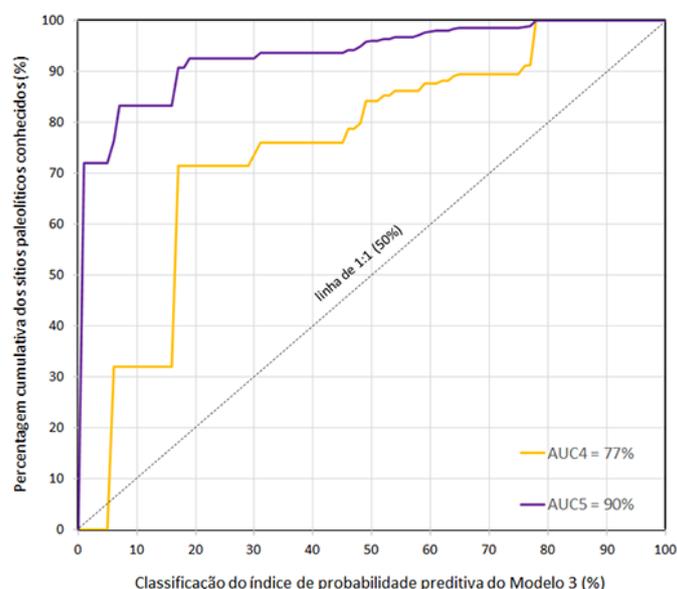


Figura 20 - Avaliação do Modelo 3 através das curvas ROC/AUC. No eixo horizontal do gráfico, o índice de probabilidade preditiva está cumulativamente indicado por ordem decrescente. A curva ROC4/AUC4 (em amarelo) mostra a relação entre a localização dos 22 sítios arqueológicos conhecidos na área de teste, não usados nas fases de treino de atribuição dos pesos, com a classificação do índice de probabilidade preditiva do Modelo 3 (este último calculado utilizando o método da RFP - na área de treino - em combinação com o processo de AHP). Por sua vez, a curva ROC5/AUC5 (em roxo) mostra a relação entre a localização de todos os 40 sítios arqueológicos conhecidos da área de estudo com a classificação do índice de probabilidade preditiva do Modelo 3 (mais detalhes no texto).

O confronto espacial, em SIG, entre o modelo *raster* obtido utilizando o método da RFP em combinação com o processo de AHP (Fig. 19) e a localização dos sítios arqueológicos conhecidos da área de teste (*i.e.*, os 22 sítios não usados nas fases de treino de atribuição dos pesos) permitiu avaliar a capacidade preditiva do Modelo 1 através da elaboração da curva ROC4 (Fig. 20). A AUC4 foi então utilizada para quantificar o desempenho do Modelo 3, apontando para um valor aceitável de 77%. o ganho (*G*) deste Modelo 3 resulta ser de 0.17, sugerindo, mais uma vez, que pode ser aceite. Além disso, aponta-se uma taxa de sucesso do Modelo 3 de 90%

(curva ROC5/AUC5 na Fig. 20) calculada através da comparação espacial entre o *raster* da Figura 19 e a localização de todos os 40 sítios arqueológicos conhecidos da área de estudo. A curva ROC5 demonstra que com apenas 10% da área de estudo classificada como muito alta probabilidade de preservação do registo arqueológico paleolítico, no Modelo 3, se conseguem enquadrar mais de 80% dos sítios arqueológicos conhecidos. Quando se consideram os 20% da área de estudo classificada como alta a muito alta probabilidade verifica-se que ela integra mais de 90% dos sítios.

5 DISCUSSÃO

Os resultados alcançados com este trabalho de investigação podem ser discutidos de acordo com duas perspetivas gerais: (A) uma mais virada para os aspetos analíticos e metodológicos (vantagens e limitações) considerados na elaboração dos três modelos cartográficos preditivos capazes de identificar áreas com maior ou menor potencial à preservação do registo Paleolítico na área de estudo, juntamente com a escolha do modelo que se afigura mais robusto e fiável - desse modo, pode perspetivar-se a aplicabilidade destes procedimentos em contextos geográficos e geomorfológicos diversos do fluvial; (B) outra mais ligada às possíveis utilidades ou aplicações práticas do protótipo de mapa de probabilidade de preservação do registo Paleolítico desenvolvido para a área de estudo, sobretudo nos contextos da Arqueologia Preventiva e das políticas públicas de Planeamento e Ordenamento do Território.

5.1 Comparação entre modelos

A comparação entre os resultados dos três modelos cartográficos preditivos produzidos neste trabalho, utilizando vários métodos de AR e de AMC, singularmente (GAM) ou em combinação (RFP + AHP), permitiu avaliar e comparar as respetivas taxas de sucesso tendo em conta todos os 40 sítios paleolíticos conhecidos na área de estudo (Figs. 16, 18 e 20): o Modelo 3 parece oferecer o melhor desempenho (90%) na deteção de locais com maior potencial de preservação do registo arqueológico, seguido pelo Modelo 2 (87%) e pelo Modelo 1, que parece ser o modelo menos eficaz (68%). O uso do método da RFP, utilizando só os dados arqueológicos de treino na atribuição dos pesos de importância relativa das classes/categorias de cada fator ambiental condicionante previamente escolhido, em combinação com o processo de AHP para a ponderação dos fatores, parece ser, no caso específico deste trabalho, a abordagem metodológica com melhor desempenho em comparação com as outras duas. A escolha de um número limitado, mas informativo, de fatores (5 neste caso), a utilizar como variáveis de entrada na modelação, também fica demonstrada ser uma tática adequada, embora mais experiências metodológicas são sempre desejáveis neste sentido.

No entanto, embora de uma forma geral aceitamos esta hierarquização dos resultados cartográficos preditivos com base na análise das curvas ROC/AUC, importa também referir que, geralmente, quando o padrão das curvas de sucesso apresenta grandes saltos quantitativos em

degraus sucessivos (como é o caso da ROC3 do Modelo 2 e da ROC4 do Modelo 3 - Figs. 18 e 20) isso pode estar relacionado com situações de *overfitting* (artificial) do processo de modelação, sendo que a situação normal corresponde a um incremento grande inicial da curva seguido por incrementos mais pequenos (situação ideal) ou incrementos mais ou menos idênticos ao longo de toda a curva (como é o caso da ROC1 e da ROC2 do Modelo 1 - Fig. 16). No caso específico deste trabalho, embora não podemos excluir a priori esta hipótese de *overfitting*, parece-nos que a configuração por vezes não-ideal das curvas de sucesso deve-se, provavelmente, ao reduzido número de dados disponíveis da variável dependente (os 40 sítios arqueológicos conhecidos) e pelo seu padrão de distribuição espacial, na área de estudo, em clusters muito localizados (*vide* Fig. 7).

Uma abordagem de atribuição dos pesos com claro carácter objetivo, como é o caso da RFP, permite evitar as ambiguidades associadas à utilização de métodos mais subjetivos e dependentes de um conhecimento detalhado da área de estudo (*e.g.*, Rua, 2009; Gonçalves, 2009, 2014; Sims, 2014; Maio, 2018). Este prévio conhecimento detalhado torna-se virtualmente impossível no caso em que se considera uma área de estudo extensa e, sobretudo, lateralmente heterogénea de um ponto de vista geológico e geomorfológico. De facto, a modelação espacial preditiva em arqueologia pré-histórica aplicada a uma área de estudo pouco heterogénea, de um ponto de vista geológico/geomorfológico (fluvial, neste caso), pode ser invocada como sendo uma das principais razões que justificam o sucesso computacional alcançado neste trabalho (sobretudo com o Modelo 3), em comparação com outras abordagens similares (analiticamente e metodologicamente) mas aplicadas em contextos geográficos diversos e mais diversificados (Caracausi *et al.*, 2018; Noviello *et al.*, 2018; Wachtel *et al.*, 2018; Nicu *et al.*, 2019; Parow-Souchon *et al.*, 2021; só para citar os mais recentes).

Todo isto reforça a necessidade de utilizar múltiplos métodos na elaboração destes modelos, mas também uma preocupação acrescida em avaliar caso a caso a possível aplicabilidade dos procedimentos adotados em áreas com características geológicas e geomorfológicas mais diversificadas, como, por exemplo, os sistemas fluviais e cárnicos que, num contexto espaço-temporal variável, se combinaram (sob a influência das variações paleoclimáticas) na definição de formas da paisagem “fluviocársica” e onde a escolha dos fatores ambientais condicionantes, a utilizar no âmbito da modelação espacial preditiva aplicada à arqueologia paleolítica, deve necessariamente ter em conta estas complexas interações genéticas.

Considerando que a informação utilizada nesta modelação espacial preditiva se baseia num conjunto bastante diferenciado de dados (*vide* Tabela 1), a aplicação de um procedimento standard que permita avaliar/verificar a eficácia dos resultados obtidos (curvas ROC/AUC), ou

seja testar o desempenho do modelo, torna-se numa tarefa adequada e sempre desejável. De facto, embora tenha sido realizada uma prévia análise de multicolinearidade entre os fatores iniciais, permitindo excluir, a priori, alguns deles dos procedimentos analíticos efetuados, a verdade é que ainda assim existem, em nosso entender, casos de equidirecionalidade paramétrica entre algumas das variáveis selecionadas. Por exemplo, o mesmo valor de declive está, embora indiretamente, contabilizado em ambos os índices de posição topográfica e de carga de calor. Para além disso, a distribuição espacial do declive está também, em parte, relacionada com os afloramentos geológicos (rochas consolidadas vs. rochas friáveis), e o mesmo índice de carga de calor pode também ser considerado, pelo menos em parte, como uma consequência da topografia. A tudo isto, acrescentam-se ainda as incertezas cartográficas relacionadas com os procedimentos de interpolação/generalização espacial que muitas vezes desempenham um papel importante na exatidão dos dados de base digitalizados, a que se junta a problemática da recolha dos dados arqueológicos com intervalos de anos e por diferentes investigadores. Neste contexto, a escala analítica escolhida para a área de estudo (Baixo Vale do Rio Lis - na região de Leiria - Centro de Portugal), juntamente com as diferenciadas escalas cartográficas de representação dos dados iniciais (*vide* Tabela 1), muitas vezes impostas pela sua disponibilidade (ou indisponibilidade no formato “*open source*”), assumem aqui um papel preponderante no condicionamento das sucessivas opções metodológicas e dos resultados finais da aplicação dos diversos modelos. Diferentes generalizações cartográficas de dados podem também introduzir ruído numa análise espacial desta natureza.

O mapa de probabilidade de preservação do registo Paleolítico mais robusto e fiável aqui construído necessariamente reflete algumas das vantagens e limitações metodológicas e práticas evidenciadas (Fig. 19). Trata-se, no conjunto, de aspetos muito difíceis de se julgar estatisticamente, embora estejamos conscientes de que a melhor abordagem é, e será sempre, um trabalho extensivo de prospeção no campo que permita a recolha de uma amostragem realística da informação de base e das evidências que permitem suportar uma avaliação/verificação dos resultados alcançados.

5.2 Uso do mapa de identificação do potencial arqueológico Paleolítico

Em termos práticos, o protótipo cartográfico preditivo mais robusto e fiável produzido neste trabalho - o Modelo 3 da Figura 19 - onde se regista que o $3\% + 18\% = 21\%$ da área de estudo corresponde a zonas com alto a muito alto potencial à preservação do registo Paleolítico,

pode tornar-se num instrumento de suporte para a investigação arqueológica *p.d.*, mas também como complemento (ou auxílio) à tomada de decisão nas políticas públicas de planeamento e ordenamento do território. Por exemplo, a sua utilidade nos Planos Diretores Municipais prende-se com a capacidade intrínseca em evidenciar, espacialmente, num mapa e na escala mais apropriada (Fig. 21), aqueles elementos territoriais que devem ser investigados antes de autorizar qualquer tipo de intervenção pública ou privada na paisagem através da delimitação das zonas mais suscetíveis em preservar um registo arqueológico (Paleolítico, neste caso específico).

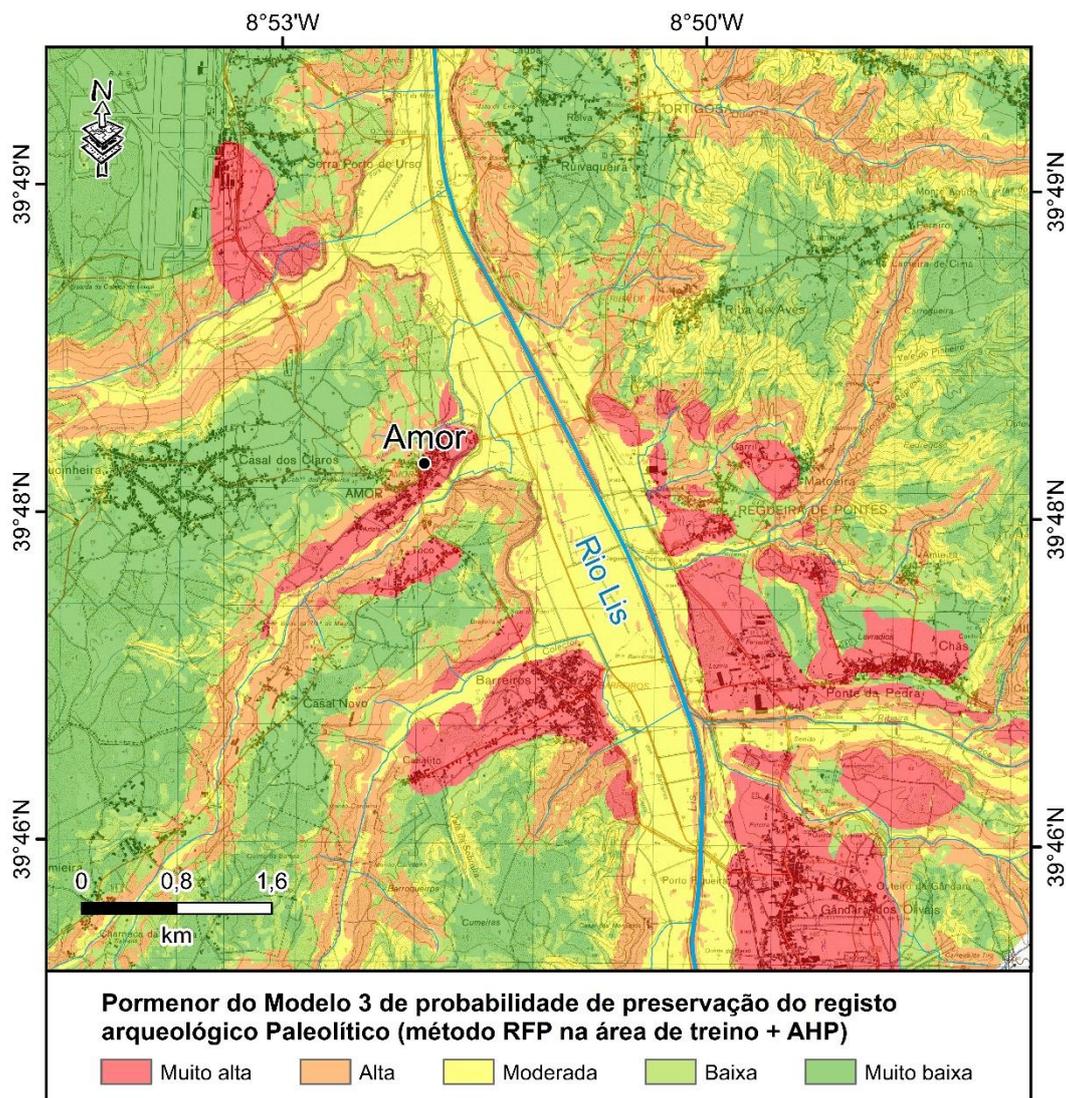


Figura 21 - Pormenor do Modelo 3 utilizado como exemplo para evidenciar a importância que esta tipologia de cartografia pode ter no âmbito dos Planos Diretores Municipais.

Este desafio pode certamente contribuir para um melhor conhecimento, proteção, preservação e gestão do património cultural (a escala local e regional), bem como para reduzir

de forma substancial os custos dos promotores de obras integrando os procedimentos analíticos e metodológicos aqui desenvolvidos na fase dos estudos de Impacte Ambiental, sobretudo no contexto da Arqueologia Preventiva que tem como principal papel a consciencialização de quem vai intervir (*e.g.*, através do acompanhamento as obras públicas e privadas) no território.

Por outras palavras, a Arqueologia Preventiva trata de estabelecer várias ações que evidenciam e minimizem a potencial degradação do património arqueológico existente nos locais a intervir. Ao longo dos últimos anos, este conceito operacional tem vindo a modificar-se, obrigando a uma participação mais ativa e exigente das equipas de arqueólogos, algo que colocou a Arqueologia Preventiva num lugar de destaque nas tarefas de planeamento e gestão dos territórios, com uma importância relevante durante a elaboração dos Planos de Ordenamento do Território, sendo estes os principais instrumentos de planeamento cuja missão principal se prende com a utilização sustentável dos recursos e a valorização do território. De facto, é preciso ter em consideração que no caso dos depósitos arqueológicos trata-se de um património quase sempre vulnerável devido ao facto de nem sempre ser facilmente identificável para a maioria da população, podendo, desta forma, estar em risco se não for devidamente salvaguardado.

Esta questão do património arqueológico está carregada de uma especificidade muito particular, pois o desenvolvimento das tarefas necessária para a expansão do conhecimento estão intimamente relacionadas com uma dupla característica dos processos de escavação arqueológica que, ao mesmo tempo, constituem fonte de dados e de destruição desses mesmos dados (Martins, 2012). É durante os trabalhos de escavação que se torna necessário tomar a decisão do que se deve destruir e do que se deve preservar, existindo aqui uma clara escolha subjetiva que depende muito da leitura que o profissional faz do contexto arqueológico e do tipo de intervenção (Moreira, 2006).

Outro ponto que se deveria considerar, sobretudo junto das populações locais, é a desmistificação destes instrumentos de proteção, que muitas vezes são vistos pelos intervenientes como uma ameaça ou penalização, na medida em que condicionam a utilização do território, contribuindo assim para que muitos proprietários omitam a existência de achados arqueológicos nos próprios terrenos e levando, muitas vezes, à sua destruição. É aqui que os resultados cartográficos de uma modelação espacial preditiva aplicada à Arqueologia Preventiva se tornam imprescindíveis.

6 CONSIDERAÇÕES FINAIS

Em “Ciências da Terra” onde, de uma forma quase provocadora, mas com evidente cognição de causa, incluímos também a arqueologia pré-histórica, com particular ênfase para o período mais antigo da história da humanidade (o Paleolítico), a preservação dos depósitos arqueológicos, integrados nas sucessões sedimentares quaternárias de sítios ao ar-livre, grutas ou abrigos, é essencialmente controlada e condicionada pelos processos naturais. De facto, e assim como já se referiu na introdução deste trabalho, a distribuição dos sítios arqueológicos na paisagem não é aleatória, mas antes, reflete uma escolha cultural e é influenciada pelas condições naturais e de disponibilidade/acesso aos recursos essenciais à vida no período temporal de referência. Para além disso, os fatores ambientais que condicionaram a fixação das comunidades e a eventual preservação dos vestígios estão ainda presentes na paisagem atual, podendo ser identificados, classificados e quantificados. Desta forma, a capacidade em descobrir novos sítios paleolíticos depende muito do entendimento prévio que os arqueólogos, ou quem conjuntamente com eles trabalham (equipas técnicas multidisciplinares), adquiriram sobre os processos e os elementos naturais que moldaram ou influenciaram esta distribuição, em diferentes contextos geológicos/geomorfológicos (fluvial, cársico, fluviocársico, eólico, planalto, etc.), claramente variáveis no espaço e no tempo.

Tendo em conta estas considerações, este trabalho teve como objetivo principal fornecer uma visão crítica na elaboração de modelos cartográficos preditivos em arqueologia pré-histórica, assim como na respetiva classificação do território em zonas com diferentes graus de probabilidade de preservação do registo Paleolítico. Para tal, utilizou-se como área de estudo um contexto natural pouco diferenciado de um ponto de vista geológico e geomorfológico (fundamentalmente de tipo fluvial), correspondente ao Baixo Vale do Rio Lis, na região de Leiria (Centro de Portugal).

Iniciou-se pela criação de uma base de dados geoespaciais onde se incluíram todos os elementos topográficos, hidrográficos, geológicos e geomorfológicos que, nesta específica área de estudo, foram considerados preponderantes no condicionamento, no tempo e no espaço, da preservação deste registo. A partir desta base de dados, e depois de várias tentativas de inclusão/exclusão de fatores, os seguintes fatores ambientais condicionantes foram escolhidos: declive, índice de posição topográfica, distância ponderada de custo para acesso aos cursos de água, índice de carga de calor e litologia. A estes fatores juntaram-se também os dados arqueológicos, recuperados dos registos oficiais, e correspondentes a localização de 40 sítios paleolíticos conhecidos, até ao presente, na área de estudo.

Todos estes dados espacializados foram recolhidos, processados e organizados em camadas informativas em modo matricial (*raster*), com resolução espacial de 10 m x 10 m de *pixel*, utilizando o pacote de softwares ArcGIS Desktop 10.8.1 da ESRI® - conjuntamente com a utilização de ferramentas específicas (*i.e.*, “*Topographic Tools Toolbox*”, “*Geomorphometry and Gradient Metric Tools*” & “*Marine Geospatial Ecology Tools*”). As camadas *rasters* assim produzidas foram usadas como variáveis de entrada nos processos analíticos conducentes à elaboração do modelo preditivo de identificação do potencial arqueológico Paleolítico na região de Leiria.

Para desenvolver e, depois, avaliar o desempenho dos modelos cartográficos, tanto a área de estudo, como os 40 sítios arqueológicos conhecidos foram subdivididos em dois conjuntos de dados: (i) área de treino (correspondente a $\sim 1/3$ da área de estudo) e amostra de treino (18 sítios), respetivamente; e (ii) área de teste (restantes $\sim 2/3$ da área de estudo) e amostra de teste (restantes 22 sítios), respetivamente.

Três modelos cartográficos espacializados foram desenvolvidos em SIG, com recurso a três métodos geoestatísticos distintos:

- O Modelo Aditivo Generalizado (GAM), para desenvolver um primeiro resultado cartográfico preditivo (= Modelo 1) examinando a relação espacial entre os fatores ambientais condicionantes e a presença/ausência dos 18 sítios paleolíticos conhecidos (treino) e escolhidos aleatoriamente na área de estudo;
- A Razão de Frequência Probabilística (RFP), para pesar as classes/categorias dos fatores ambientais condicionantes através da determinação da relação espacial existente entre estas e a localização de todos os 40 sítios paleolíticos conhecidos na área de estudo, a que se junta o processo de Análise Hierárquica por Pares (AHP) para a ponderação dos fatores e que, combinados, permitiram produzir o segundo elaborado cartográfico preditivo (= Modelo 2);
- A RFP, utilizada para pesar as classes/categorias dos fatores ambientais condicionantes através da determinação da relação espacial existente entre estas e a localização dos 18 sítios arqueológicos paleolíticos conhecidos na área de treino, a que se junta o processo de AHP para a ponderação dos fatores e que, combinados, permitiram produzir o terceiro e último elaborado cartográfico preditivo (= Modelo 3).

No caso dos Modelos 1 e 2, a comparação espacial dos resultados cartográficos com a localização dos 22 sítios (amostra/área de teste) não utilizados nas fases de atribuição dos pesos, através da utilização das curvas ROC/AUC, permitiu avaliar a capacidade preditiva dos dois cenários, apontando para os seguintes valores: 70% para o Modelo 1 e 77% para o Modelo 3. Para além disso, em termos de taxa de sucesso calculada para a área de estudo, utilizando para tal todos os 40 sítios arqueológicos conhecidos, os três elaborados cartográficos preditivos foram comparados entre si verificando-se que o Modelo 3 mostra o melhor desempenho (90%), seguido pelo Modelo 2 (87%) e pelo Modelo 1, que parece ser o modelo menos eficaz (68%).

No caso do Modelo 3, ou seja, o mais robusto e fiável, as características topográficas e geomorfológicas mais significativas em condicionar a preservação do registo Paleolítico parecem ser, juntamente com uma clara e obvia vizinhança aos cursos de água principais, a presença de superfícies planas (ou restos delas) com declive nulo ou pouco acentuado, a meio das encostas e em vertentes geralmente mais quentes o com índice de carga de calor moderado. Em termos geológicos, os depósitos que identificam os terraços fluviais do Plistocénico são as unidades mais propensas a esta preservação. Neste contexto, e considerando o exercício de comparação realizado entre os resultados dos três modelos espaciais preditivos desenvolvidos, importa destacar como as áreas que todos eles identificam como de elevada a muito elevada probabilidade de preservação do registo Paleolítico se apresentam sempre, ou quase sempre, com características fisiográficas similares nos três cenários, embora ocupando porções de território com dimensões variáveis. Desta forma, embora fica demonstrada a maior fiabilidade de um modelo respeito a outro, a verdade é que todos eles têm que ser visto como ferramentas complementares e não destinados a excluírem-se mutuamente; nada impede que com outros dados da variável dependente (os sítios arqueológicos conhecidos), ou outras variáveis independentes (fatores ambientais condicionantes), os resultados não fossem diferentes.

Finalmente, e num contexto mais aplicativo, os modelos preditivos em arqueologia podem ser vistos como uma ferramenta indispensável em auxiliar a delimitação das áreas mais sensíveis e com maior potencial arqueológico ainda não identificado, podendo este estar em risco se forem descobertas tardiamente ou sem um acompanhamento profissionalizado (papel este da dita Arqueologia Preventiva).

BIBLIOGRAFIA

- Almeida, F., 2000. The Terminal Gravettian of Portuguese Estremadura: Technological Variability of the Lithic Industries. Tese de Doutoramento, Dallas, Southern Methodist University.
- Almeida, F., 2005. Abrigo do Lagar Velho. O Paleolítico Superior da Bacia do Lis. In Carvalho, S. (coord.), *Habitantes e Habitats - Pré e Proto-História na Bacia do Lis*. Câmara Municipal de Leiria, Leiria, 68-83.
- Almeida, F., (Dir) 2013. Testemunhos do Paleolítico no Regolfo de Alqueva. Resultados do Bloco 1 do Plano de Minimização de impactes sobre o Património Arqueológico, Beja: EDIA.
- Almeida, F., Gameiro, C., Zilhão, J., 2002. The artifact assemblages. In Zilhão, J., Trinkaus, E. (Eds), *Portrait of the Artist as a child. The Gravettian Human skeleton from the Abrigo do Lagar Velho and its Archaeological Context*, Instituto Português de Arqueologia, Lisboa, *Trabalhos de Arqueologia* 22, 202-219.
- Almeida, A.C., Gama, A., Cunha, L., Jacinto, R., Boura, I., Medeiros, J., Brandão, J., 1989. A Bacia Hidrográfica do Rio Lis - contributo para o estudo da organização do espaço e dos problemas de ambiente. Comissão de Coordenação Regional do Centro, Câmara Municipal de Leiria.
- Almeida, A.C., Soares, A.F., Cunha, L., Marques, J.F., 1990. Proémio ao estudo do Baixo Mondego. *Biblios LXVI*, 17-47.
- Angelucci, D.E., 2003. A partir da Terra: a contribuição da Geoarqueologia. In Mateus, J., Moreno-García (Eds.), *Paleoecologia Humana e Arqueociências. Um programa multidisciplinar para a Arqueologia sob a tutela da cultura*, Lisboa, Instituto Português de Arqueologia, *Trabalhos de Arqueologia* 29, 35-84.
- Araújo, M. B., Pearson, R. G., Thuiller, W., Erhard, M., 2005. Validation of species–climate impact models under climate change. *Global Change Biology* 11, 1504-1513.
- Aubry, T., (Eds.) 2009. 200 séculos da história do Vale do Côa: incursões na vida quotidiana dos caçadores-artistas do Paleolítico. *Trabalhos de Arqueologia* 52.
- Aubry, T., Luís, L., Dimuccio, L.A., 2012. Nature vs. Culture: present-day spatial distribution and preservation of open-air rock art in the Côa and Douro River Valleys (Portugal). *Journal of Archaeological Science* 39, 848-866.
- Aubry, T., Cunha-Ribeiro, J.P., Angelucci, D., 2005. Testemunhos da ocupação pelo Homem de Neandertal: o sítio da Praia do Pedrogão. In Carvalho, S, (Coord.), *Habitantes e Habitats- Pré e Proto História na Bacia do Lis*, Câmara Municipal de Leiria, 26-33.
- Aubry, T., Gameiro, C., Mangado Llach, J., Luís, L., Matias, H., Pereiro, T., 2016. Upper Palaeolithic lithic raw material sourcing in Central and Northern Portugal as an aid to reconstructing

- hunter-gatherer societies. Proceedings of the 10th International Symposium on Knappable Materials, Barcelona, Journal of Lithic Studies 3 (2).
- Aubry, T., Luís, L., Dimuccio, L.A., 2017b. Porque é que a arte do Coa se concentra na margem esquerda? Condicionantes geológicas e ambientais para a formação e conservação dos suportes artísticos do Vale do Coa. O Arqueólogo Português 4/5 (Série V), 133-174.
- Aubry, T., Dimuccio, L.A., Moura, H., 2017a. Paleoambientes e Culturas do Paleolítico Superior no Centro e Norte de Portugal: balanço e perspectivas de investigação. Estudos do Quaternário 17, 29-43.
- Ayalew, L., Yamagishi, H., Ugawa, N., 2004. Landslide susceptibility mapping using GIS-based weighted linear combination, the case in Tsugawa area of Agano River, Niigata prefecture, Japan Landslides, 1 (2004), pp. 73-81
- Azerêdo, A.C., 2007. Formalização da litostratigrafia do Jurássico inferior e médio do Maciço Calcários Estremenho (Bacia Lusitaniana). Comunicações Geológicas 94, 29-51.
- Azerêdo, A.C., Duarte, L.V., Henriques, H.M., Manupela, G., (Eds.) 2003. Da dinâmica continental no Triásico aos mares do Jurássico Inferior e Médio. Cadernos de Geologia de Portugal, Lisboa.
- Balla, A., Pavlogeorgatos, G., Tsiafakis, D., Pavlidis, G., 2014. Recent advances in archaeological predictive modeling for archaeological research and cultural heritage management. Mediterranean Archaeology and Archaeometry 14 (4), 143-153.
- Banerjee, R., Srivastava, P.K., Pike, A.W.G., 2018. Petropoulos, G.P. Identification of painted rock-shelter sites using GIS integrated with a decision support system and fuzzy logic. ISPRS International Journal of Geo-Information 7, 326.
- Bertrand, G., 1982. Construire la Geographie Physique. Herodote 26, 90-116.
- Binford, L., 1983. In Pursuit of the Past, London, Thames and Hudson.
- Braga, A.C., 2001. Curvas ROC: aspectos funcionais e aplicações. Tese de doutoramento, Universidade do Minho.
- Braz, A., Gaspar, R., 2003. Intervenção de emergência no vale da Ribeira das Chitas: o caso de dois abrigos com Pré-história antiga. Al-Madan 12 (II Série), 186-187.
- Braz, A.F., Gaspar, F., Pereira, T., 2006. Vale da Ribeira das Chitas - sondagens de diagnóstico. Relatório Final - Fase 1, Torres Novas, junho de 2006.
- Brown, A.G., 1997. Alluvial Geoarchaeology. Floodplain archaeology and environmental change. Cambridge University Press.
- Burrough, P.A., McDonnell, R.A., 1998. Principles of Geographical Information Systems. Oxford University Press, Oxford.
- Cabral, J., 1995. Neotectónica em Portugal continental. Mem. Inst. Geológico e Mineiro 31.

- Callapez, P.M., Dinis, J.L., Soares, A.F., Marques, J.F., 2010. O cretácico superior da Orla Meso-Cenozoica Ocidental de Portugal (Cenomaniano a Campaniano Inferior). Ensino e Investigação e sua história, Volume I, Capítulo III, Paleontologia e Estratigrafia, 331-340.
- Caracausi, S., Berruti, G.L.F., Daffara, S., Berté, D., Borel R.F., 2018. Use of a GIS predictive model for the identification of high altitude prehistoric human frequentations. Results of the Sessera valley project (Piedmont, Italy), *Quaternary International* 490, 10-20.
- Carleton, W.C., Cheong, K.F., Savage, D., Barry, J., Conolly, J., Iannone, G., 2017. A comprehensive test of the Locally-Adaptive Model of Archaeological Potential (LAMAP). *Journal of Archaeological Science: report* 11, 59-68.
- Carvalho, A.M.G., 1972. Caracterização sedimentológica dos depósitos Plio-pleistocénicos da região de Leiria. *Rev. Fac. Ciências, 2º Sec., C, Ciências Naturais, Lisboa, vol. XVII, faz. 1,* 197-206.
- Carvalho, J.M., 2013. Tectónica e caracterização da fraturação do Maciço Calcário Estremenho, Bacia Lusitânica. Contributo para a prospeção de rochas ornamentais e ordenamento da atividade extrativa. Tese de Doutoramento, Departamento de Geologia da Faculdade de Ciências da Universidade de Lisboa.
- Carvalho, J.M., 2018. Jointing patterns and tectonic evolution of the Maciço Calcário Estremenho, Lusitanian Basin, Portugal. *Journal of Structural Geology* 110, 155-171.
- Carvalho, S., Tavares, J., 2005. A jazida de sílex e Oficina de talhe do Povo da Martinela. In Carvalho, S., (Coord.), *Habitantes e Habitats - Pré e Proto-História na Bacia do Lis,* Câmara Municipal de Leiria, 26-33.
- Carvalho, S., Carvalho, V., 2007. Relatório de progresso da Carta Arqueológica de Leiria (2004 - 2007), Câmara Municipal de Leiria (relatório técnico não publicado).
- Carvalho, V., 2011. O Abrigo do Laga Velho e o Paleolítico em Leiria, Portugal: análise dos dados arqueológicos no actual contexto da evolução humana. Dissertação de mestrado apresentada à Faculdade de Ciências e Tecnologia da Universidade de Coimbra.
- Carvalho, V., Pajuelo, A., 2005. Novas realidades no campo da investigação arqueológica – minimização de impactos e arqueologia preventiva: projecto Simlis 2002 a 2005. In Carvalho, S., (Coord.), *Habitantes e Habitats - Pré e Proto-História na Bacia do Lis,* Câmara Municipal de Leiria, 135-156.
- Carver, S., 1991. Integrating multi-criteria evaluation with geographical information systems. *International Journal of Geographical Information Systems* 5 (3), 321-339.
- Cavulli, F., Grimaldi, S., 2007. To see or not to see. Archaeological data and surface visibility as seen by an AIS (Archaeological Information System) approach. In Figueiredo, A., Velho,

- G., (Eds.), Proceedings of the XXXIII Computer Application and Quantitative Methods in Archaeology. The World Is in Your Eyes, Tomar, 21-24 January 2005, 413-422.
- Chaktar, S., Mosseau, V., 2007. Spatial Multicriteria Decision Making. LAMSADE University of Paris Dauphine, France.
- Charlton, M.E., Brunson, M., Fotheringham, A.S., 2002. Geographically weighted regression: the analysis of spatially varying relationships. John Wiley & Sons, Chichester.
- Conolly, J., Lãke, M., 2006. Geographical Information Systems in Archaeology. Cambridge: Cambridge University Press.
- Costa, A., 2009. A aplicabilidade dos SIG e das imagens na identificação de áreas com potencial arqueológico: estações arqueológicas da Idade do Ferro. Dissertação de Mestrado em Sistemas de Informação Geográfica e Ordenamento do Território à Faculdade de Letras da Universidade do Porto.
- Costanzo, D., Irigaray, C., 2020. Comparing Forward Conditional Analysis and Forward Logistic Regression Methods in a Landslide Susceptibility Assessment: A Case Study in Sicily. *Hydrology* 7 (3), 37.
- Crispim, J.A., 1995. Dinâmica cársica e implicações ambientais nas depressões de Alvados e Minde. Tese de Doutoramento, em Geologia, Faculdade de Ciências da Universidade de Lisboa.
- Cunha, L.V., Gonçalves, A.S., Figueredo, V.A., Lino, M., 1980. A gestão da água. Princípios fundamentais e sua aplicação em Portugal. Fundação Calouste Gulbenkian.
- Cunha, L., 1990. As Serras Calcárias de Condeixa-Sicó-Alvaiázere. Estudo de Geomorfologia. Instituto Nacional de Investigação Científica.
- Cunha, L., Dimuccio, L.A., 2002. Considerações sobre riscos naturais num espaço de transição. Exercícios cartográficos numa área a Sul de Coimbra. *Territorium* 9, 37-51.
- Cunha, L., Dimuccio, L.A., Figueiredo, R., 2017. Analyse multi-aléas sur le territoire de la municipalité de Coimbra (centre-ouest du Portugal). L'omniprésence du climat et l'importance anthropique. *Geo-Eco-Trop* 41 (3) n.s., 399-419.
- Cunha, L., Dimuccio, L.A., Paiva, I., 2020. The Sicó Massif: Morphostructural Aspects, Hydrology and Karstification. In: G. Vieira, G., Zêzere, J.L., Moura, C., (eds.), *Landscapes and Landforms of Portugal, World Geomorphological Landscapes*, 211-227.
- Cunha-Ribeiro, J.P., 1992. Contribuição para o estudo do Paleolítico do Vale do Lis no seu contexto crono-estratigráfico. *PORTVGALIA*, Nova Série, Vols. XIII-XIV.
- Cunha-Ribeiro, J.P., 1999. O Acheulense no Centro de Portugal: o vale do Lis. Contribuição para uma abordagem tecno-tipológica das suas indústrias líticas e problemática do seu

- contexto cronoestratigráfico. Tese de Dissertação de Doutoramento, Faculdade de Letras da Universidade de Lisboa.
- Cunha-Ribeiro, J.P., 2002. O Paleolítico Inferior em Portugal no final do século XX: balanço das investigações e novos desafios. In *Arqueologia 2000 – Balanço de um século de investigação arqueológica em Portugal*. *Arqueologia e História* 94, 13-24.
- Cunha-Ribeiro, J.P., 2005. O Paleolítico Inferior os primeiros habitantes da bacia hidrográfica do Rio Lis. In Carvalho, S., (Coord.), *Habitantes e Habitats – Pré e Proto-História na Bacia do Lis*, Câmara Municipal de Leiria, 36-53.
- Daveau, S. *et al.*, 1985. *Mapas Climáticos de Portugal. Nevoeiro e Nebulosidade. Contrastes Térmicos*. Centro de Estudos Geográficos, Memórias 7, Lisboa.
- Davidsson, P., Johansson, S., Svahnberg, M., 2005. Using the analytic hierarchy process for evaluating multi-agent system architecture candidates. *AOSE'05 Proceedings of the 6th international conference on Agent-Oriented Software Engineering*, 205-217.
- De Steiguer, J.E., Duberstein, J., Lopes, V., 2003. The Analytic Hierarchy Process as a Means for Integrated Watershed Management - Proceedings of the 1st Interagency Conference on Research on the Watersheds, US Department of Agriculture, Agricultural Research Service, Benson, Arizona.
- Dilts, T.E., 2015. *Topography Tools for ArcGIS 10.1*. University of Nevada Reno, <http://www.arcgis.com/home/item.html?id=b13b3b40fa3c43d4a23a1a09c5fe96b9>.
- Dimuccio, L.A., Cunha, L., Figueiredo, R., 2020. Defining a multi-hazard index in territorial planning framework: application on the Coimbra municipality (western-central Portugal). Abstract EGU 2020-9598, EGU General Assembly 2020. <https://doi.org/10.5194/egusphere-egu2020-9598>
- Dimuccio, L.A., Ferreira, R., Cunha, L., Campar, A.A., 2011. Regional forest-fire susceptibility analysis in central Portugal using a probabilistic ratings procedure and artificial neural network weights assignment. *International Journal of Wildland Fire* 20, 776-791.
- Dimuccio, L.A., Ferreira, R., Cunha, L., Campar de Almeida, A., 2008. Susceptibilidade aos Incêndios Florestais na Região Centro de Portugal. Utilização de ferramentas SIG e de um Modelo de Redes Neurais para ponderar os factores condicionantes. *Cadernos de Geografia* 26/27, 179-188.
- Dimuccio, L.A., Ferreira, R., Cunha, L., 2005. Aplicação de um modelo de redes neuronais (Neural Network) na elaboração de mapas de susceptibilidade a movimentos de vertente: um exemplo numa área a Sul de Coimbra. In: *Acta do 2º Congresso nacional de Geomorfologia*, Vol. III, 281-289.

- Dinis, P.A., 1996. Dinâmica sedimentar e evolução do Estuário do Lis. Tese de Mestrado, Departamento de Ciências da Terra, Universidade de Coimbra, Coimbra.
- Dinis, J.L., 2001. Definição da Formação da Figueira da Foz; Aptiano a Cenomaniano do sector central da margem oeste ibérica. *Comunicações do Instituto Geológico e Mineiro*, t. 88, 127-160.
- Dinis, J.L., 1999. Estratigrafia e sedimentologia da Formação de Figueira da Foz. Aptiano a cenomaniano do sector norte da Bacia Lusitânica. Tese de Doutoramento apresentada no Departamento de Ciências da Terra da Universidade de Coimbra.
- Dinis, J.L., Rey, J., Cunha, P.P., Callapez, P., Pena dos Reis, R., 2008. Stratigraphy and allogenic controls of the western Portugal Cretaceous: and update synthesis. *Cretaceous Research* 29, 772-780.
- Espa, G., Benedetti, R., De Meo, A., Ricci, U., Espa, S., 2006. GIS based models and estimation methods for the probability of archaeological site location. *Journal of Cultural Heritage* 7, 147-155.
- Estanqueiro, M., 2016. Modelo Preditivo Logístico Aplicado aos Povoados Proto-Históricos da Beira Interior. Dissertação de Mestrado em Arqueologia e Território na especialidade de Proto-História apresentada à Faculdade de Letras da Universidade de Coimbra.
- Estanqueiro, M., 2017. Modelos Preditivos em Arqueologia: uma aplicação aos povoados proto-históricos da Beira Interior. *Al-Madan (II Série)* 21 (3), 122-127.
- Ferreira, M.P., Macedo, C.R., 1983. Igneous rocks in the diapiric areas of the western Portuguese border: the K-Ar ages and settings of the Upper Jurassic suite. *Mem. Not. Pub. Mus. Lab. Min. Geol. Univ. Coimbra* 96, 159-181.
- Ferreira, A.B., Rodrigues, M.L., Zêzere, J.L., 1988. Problemas de evolução Geomorfológica do Maciço Calcário Estremenho. *Finisterra* XXIII (45), 5-28.
- Ferreira, R., 2001. Modelação cartográfica em ambiente SIG para apoio à decisão. Dissertação de Mestrado, Faculdade de Letras, Universidade de Coimbra.
- Gameiro, C. 2019. Portela II (Maceira, Leiria) – Porque é que um sítio arqueológico, alvo de uma escavação preventiva, é incluído num projeto de investigação? *Anais Leirienses – estudos & documentos* 4, 85-98.
- Gameiro, C., Dimuccio, L.A., 2019. O Paleolítico Superior e a Arqueologia Preventiva em Portugal: desafios e oportunidades. Apresentação do projeto PALEORESCUE. *Al-Madam* 22, 55-60.
- Gameiro, C., Aubry, T., 2020. Rodo, Vau e Bispeira 8: integração crono-cultural e contextualização no quadro do paleolítico Superior Peninsular. In: EDP (Edit.) O

- Património Histórico e Arqueológico do Vale do Vouga. O aproveitamento Hidroelétrico de Ribeiradio-Erminda.
- Gaspar, R., Ferreira, J., Carrondo, J., Silva, M.J., 2016b. Open-air Gravettian lithic assemblages from Northeast Portugal: the Foz do Medal site (Sabor Valley). *Quaternary International* 406, 44-64.
- Gaspar, R., Ferreira, J., Carrondo, J., Silva, M.J., 2016a. The use of quartz during Upper Paleolithic and Early Mesolithic in Sabor valley (NW Iberia): the Foz do Medal case. *Quaternary International* 424, 98-112.
- Goepel, K.D., 2013. Implementing the Analytic Hierarchy Process as a Standard Method for Multi-Criteria Decision Making. In *Corporate Enterprises – A New AHP Excel Template with Multiple Inputs. Proceedings of the International Symposium on the Analytic Hierarchy Process*, 1-10.
- Gonçalves, C., 2009. Modelos preditivos em SIG na localização de sítios arqueológicos de cronologia mesolítica no Vale do Tejo. Tese de mestrado em Arqueologia apresentada à Faculdade de Ciências Humanas e Sociais da Universidade do Algarve.
- Gonçalves, C., 2011. Os sistemas de informação geográfica como ferramenta de análise de ocupação do território na época mesolítica - O caso de Vale do Tejo. *Estrat Crític* 5 (3), 145-155.
- Gonçalves, C., 2014. Modelos preditivos de ocupação do território no Mesolítico entre os Vales do Tejo e do Sado. Tese de doutoramento em Arqueologia apresentada à Faculdade de Ciências Humanas e Sociais da Universidade do Algarve.
- Gonçalves, P., 2007. A Evolução Holocénica do Rio Lis e da Laguna da Pederneira. Dissertação apresentada ao Departamento de Ciências da Terra, para obtenção do grau de Mestre em Geociências Ramo de Ambiente e Ordenamento do Território. Faculdade de Ciências e Tecnologia da Universidade de Coimbra.
- Gonçalves, R.M, 2012. Arqueologia preventiva: modelos e perspetivas ao serviço da memória cultural. Dissertação de mestrado apresentado à Universidade Católica Portuguesa, Lisboa.
- Grøn, O., 2018. Some problems with modelling the positions of prehistoric hunter-gatherer settlements on the basis of landscape topography. *Journal of Archaeological Sciences: reports* 20, 192-199.
- Hancock, G.R., Coulthard, T.J., Martinez, C., Kalma, J., 2011. An evaluation of landscape evolution models to simulate decadal and centennial scale soil erosion in grassland catchments. *Journal of Hydrology* 308, 171-183.
- Hastie, T., Tibshirani, R., 1986. Generalized Additive Models. *Statist. Sci.* 1 (3), 297-310.

- Heleno, M., 1922. *Antiguidades de Monte Real*. Lisboa Imprensa Nacional de Lisboa.
- Hosmer, D.W., Lemeshow, S., 2000. *Applied logistic regression*. 2ª Edição, John Wiley & Sons, Inc., New York.
- Howey, M.C., Palace, M.W., McMichael, C.H., 2016. Geospatial modeling approach to monument construction using Michigan from AD 1000–1600 as a case study. *Proceedings of the National Academy of Sciences* 113 (27), 7443–7448.
- IGeoE, 2003. *Cartas Militar de Portugal, Série M888, Escala 1:25000*. Instituto Geográfico do Exército, Lisboa.
- Jenness, J., 2006. Topographic Position Index (tpi_jen.avx) extension for ArcView 3.x, v. 1.2. Jenness Enterprises, <http://www.jennessent.com/arcview/tpi.htm>.
- Jude, W.J., Sebastian, L., 1988. *Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling*. U.S. Bureau of Land Management, Department of Interior, Denver.
- Kamermans, H., 2010. The application of predictive modelling in archaeology: problems and possibilities. In Niccolucci, F., Hermon, S., (Eds.), *Beyond the Artifact. Digital Interpretation of the Past*. *Archaeolingua*, 273-277.
- Karmon, Y., 1971. *Israel: a Regional Geography*. Wiley Interscience, London.
- Kayastha, P., Dhital, M.R., De Smedt, F., 2013. Application of the analytical hierarchy process (AHP) for landslide susceptibility mapping: A case study from the Tinau watershed, west Nepal. *Computers & Geosciences Volume 52*, 398-408.
- Kempf, M., 2019. The application of GIS and satellite imagery in archaeological land-use reconstruction: A predictive model? *Journal of Archaeological Science: reports* 25, 116-128.
- Khoshtinat, S., Aminnejad, B., Hassanzadeh, Y., Ahmadi, H., 2019. Groundwater potential assessment of the Sero plain using bivariate models of the frequency ratio, Shannon entropy and evidential belief function. *Journal Earth System Science* 128, 152.
- Klehm, C., Barnes, A., Follett, F., Simon, K., Kiahtipes, C., Mothulatshipi, S., 2019. Toward archaeological predictive modeling in the Bosutswe region of Botswana: Utilizing multispectral satellite imagery to conceptualize ancient landscapes. *Journal of Anthropological Archaeology* 54, 68-83.
- Kohler, T.A., Parker, S.C., 1986. Predictive models for archaeological resource location. *Advances in Archaeological Method and Theory* 9, 397-452.
- Kullberg, J.C., Rocha, R.B., Soares, A.F., Rey, J., Terrinha, P., Azerêdo, A.C., Callapez, P., Duarte, L.V., Kullberg, M.C., Martins, L., Miranda, R., Alves, C., Mata, J., Madeira, J., Mateus, O., Moreira, M., Nogueira, C.R., 2013. *A Bacia Lusitaniana: Estratigrafia, Paleogeografia e*

- Tectónica. In Dias, R., Araújo, A., Terrinha, P., Kullberg, J.C., (Eds.), *Geologia de Portugal*, Escola Editora, V. II, 195-347.
- Kvamme, K., 1988, Development and testing of quantitative models in quantifying the present and predicting the past. In Jugde, J., Sebastian, L., Denver, *Quantifying the present and predicting the past: Theory, method, and application of archaeological predictive modeling*, US Department of the Interior, Bureau of Land Management, 325-418.
- Kvamme, K., 1990. The fundamental principles and practice of predictive archaeological modeling. In *Mathematics and Information Science in Archaeology: A Flexible Framework* edited by Voorrips A., Bonn: Holos Verlag, 257-295.
- Kvamme, K., 1992. A predictive site location model on the High Plains: an example with an independent test. *Plains Anthropologist* 37 (138), 19-40.
- Kvamme, K., 2006. There and Back Again: Revisiting Archaeological Locational Modeling". In Mehrer, M., Wescott, K., (Eds.), *GIS and Archaeological Site Location Modeling*, Taylor & Francis, London, 4-38.
- Landau, S., Everitt, B.S., 2004. *A handbook of statistical analyses using SPSS*. Boca Raton: Chapman & Hall.
- Lasaponara, R., Masini, N., 2006. Identification of archaeological buried remains based on the normalized difference vegetation index (NDVI) from Quickbird satellite data. *IEEE Geoscience Remote Sensing Letters* 3 (3), 325-328.
- Lasaponara, R., Masini, N., 2013. Satellite synthetic aperture radar in archaeology and cultural landscape: an overview. *Archaeological Prospection* 20 (2), 71-78.
- Lloyd, C.D., 2010. *Spatial Data Analysis: an Introduction for GIS Users*. Oxford University Press, Oxford, New York.
- LNEG, 2017. *Laboratório Nacional de Engenharia e Geologia*, Portugal.
- Madsen, H., Thyregod, P., 2010. *Introduction to General and Generalized Linear Models*. CRC Press.
- Maio, D., 2018. *Ocupação Paleolítica no Barlavento Algarvio: modelos preditivos com recurso aos SIG*. Dissertação de Mestrado em Arqueologia apresentada à Faculdade de Ciências Humanas e Sociais da Universidade do Algarve.
- Malczewski, J., 1999. *GIS and multicriteria decision analysis*. John Wiles & Sons, New York.
- Malczewski, J., 2004. GIS - based land-use suitability analysis: a critical overview. *Progress in planning* 2 (1), 3-65.
- Malczewski, J., 2006. GIS-based multicriteria decision analysis: a survey of the literature. *Int. J. Geogr. Inf. Sci.* 20 (7), 703–726.

- Malczewski, J., Rinner, C., 2015. Multicriteria Decision Analysis in Geographic Information Science. *Advances in Geographic Information Science*. Springer, Berlin.
- Manuel, C., Gomes, A., Monteiro-Rodrigues, S., 2014. Modelação de Nichos Eco-culturais para o Paleolítico Médio em Portugal Continental. 'A Jangada de Pedra'. *Geografias Ibero-Afro-Americanas. Atas do XIV Colóquio Ibérico de Geografia, 1805-1811*.
- Mark, W., Mehrer, K., Wescott, L., 2006. *GIS and Archaeological Site Location Modeling*", CRC Press, Boca Raton.
- Maroco, J., 2003. *Análise Estatística - Com utilização do SPSS, 2ª edição*. Edições Sílabo.
- Martinelli, M., 2011. Técnicas de Cartografia Temática. In Venturi, L.A.B., (Org.), *Geografia: práticas de campo, laboratório e sala de aula*. São Paulo: Editora Sarandi, Capítulo 8, 203-230.
- Martins, A.M., 2012. A salvaguarda do património arqueológico no âmbito dos processos de avaliação de impacte ambiental e de ordenamento territorial: reflexões a partir do direito do património cultural, do ambiente e da gestão do território. *Revista Portuguesa de Arqueologia*.
- Martins, A.F., 1949. Maciço Calcário Estremenho. Contribuição para um estudo de Geografia Física. Tese de Doutoramento, Ciências Geográficas, Faculdade de Letras, Universidade de Coimbra.
- Marzban, C., 2004. The ROC Curve and the Area under It as Performance Measures. *American Meteorological Society, Weather and Forecasting* 19, 1106-1114.
- McCune, B., Keon, D., 2002. Equations for potential annual direct incident radiation and heat load index. *Journal of Vegetation Science* 13, 603-606.
- Mehrer, M., Wescott, K., 2006. *GIS and Archaeological Site Location Modeling*. Taylor & Francis Group.
- Mendes, G.A., Dimuccio, L.A., 2013. Geomorfologia: a construção de uma identidade. In *Riscos naturais, antrópico e mistos, Livro de Homenagem ao Prof. Doutor Fernando Rebelo, Departamento de Geografia, Faculdade de Letras, Universidade de Coimbra, Simões & Linhares Lda., (Eds.), 779-795*.
- Mokarram, M., Aminzadeh, F., 2010. GIS-based Multicriteria Land Suitability Evaluation using Ordered Weight Averaging with Fuzzy Quantifier: A Case Study in Shavur Plain, Iran. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 38 (Part II), 508-512.
- Moreira, C.D., 2006. Introdução. In *Patrimónios e identidades: ficções contemporâneas*. Lisboa: Celta.

- Munier, N., 2011. A Strategy for Using Multicriteria Analysis in Decision-Making: A Guide for Simple and Complex Environmental Projects. Springer-Verlag, Berlin.
- Nelder, J.A., Wedderburn, R.W.M., 1972. Generalized Linear Models. *Journal of the Royal Statistical Society. Serie A (General)* 135 (3), 370-384.
- Nicu, I.C., Asandulesei, A., 2018. GIS-based evaluation of diagnostic areas in landslide susceptibility analysis of Bahluet, River Basin (Moldavian Plateau, NE Romania). *Are Neolithic sites in danger? Geomorphology* 314, 27-41.
- Nicu, I.C., Pintilie, A.M., Williamson, J., 2019. GIS-Based and Statistical Approaches in Archaeological Predictive Modelling (NE Romania). *Sustainability* 11 (21) 5969.
- Noviello, M., Cafarelli, B., Calulli, C., Sarris, A., Mairota, P., 2018. Investigating the distribution of archaeological sites: Multiparametric vs probability models and potentials for remote sensing data. *Applied Geography* 95, 34-44.
- Nsanziyera, A.F., Rhiname, H., Oujaa, A., Mubea, K., 2018. GIS and Remote-Sensing Application in Archaeological Site Mapping in the Awsard Area (Morocco). *Geosciences* 8 (6), 207.
- Oosterbeek, L., Grimaldi, S., Rosina, P., Cura, S., Cunha, P., Martins, A., 2010. The earliest Pleistocene archaeological sites in western Iberia: Present evidence and research prospects. *Quaternary International* 223, 399-407.
- Pais, J., Cunha, P.P., Pereira, D., Legoinha, P., Dias, R., Moura, D., Brum Ferreira da Silva, A., Kullberg, J.C., Gonzáles-Delgado, J.A., 2012. The Paleogene and Neogene of Western Iberia (Portugal). A Cenozoic record in the European Atlantic Domain. Springer-Verlag Berlin Heidelberg.
- Palain, C., 1976. Une serie detritique terrigene. Les Gres de Silves, Trias et Lias inferieur du Portugal. *Mem. 25 (N.S.), Serv. Geol. Portugal, Lisboa.*
- Parker, R. N., Asencio, E.K., 2008. *GIS and Spatial Analysis for the Social Sciences: Coding, Mapping, and Modeling*, Routledge, New York.
- Parow-Souchon, H., Zickel, M., Manner, H., 2021. Upper Palaeolithic sites and where to find them: A predictive modelling approach to assess site expectancy in the Southern Levant. *Quaternary International* (in press).
- Pearson, K., 1895. Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London* 58, 240-242.
- Pereira, T., 2010. A exploração do quartzito na faixa Atlântica peninsular no final do Plistocénico. Tese de Dissertação de Doutoramento. Faculdade de Ciências Humanas e Sociais da Universidade do Algarve. Faro.

- Ramos, A., Cunha, L., Proença Cunha, P., 2014. Application de la Méthode de l'Analyse Multicritère Hiérarchique à l'étude des glissements de terrain dans la région littorale du centre du Portugal: Figueira da Foz – Nazaré. *Geo-Eco-Trop* 38 (1), 33-44.
- Raposo, L., 2002. Um século de estudos no Paleolítico Médio em Portugal: balanço e perspectivas. In *Arqueologia 2000 – Balanço de um século de investigação arqueológica em Portugal*. *Arqueologia e História* 94, 25-39.
- Ribeiro, A., 2013. Evolução geodinâmica de Portugal; os ciclos ante-mesozóicos. In Dias, R., Araújo, A., Terrinha, P., Kullberg, J.C., (Eds.), *Geologia de Portugal*, Escola Editora, V. II, 15-57.
- Ribeiro, A., Cabral, J., Baptista, R., Matias, L., 1996. Stress pattern in Portugal mainland and the adjacent Atlantic region, West Iberia. *Tectonics* 15 (2), 641-659.
- Ribeiro, A., Antunes, M.T., Ferreira, M.P., Rocha, R.B., Soares, A.F., Zbyszewski, G., Almeida, F.M., Carvalho, D., Monteiro, J.H., 1979. *Introduction a la geologie generale du Portugal*. Serv. Geol. Portugal, Lisboa.
- Ribeiro, O., 1945. *Portugal, o Mediterrâneo e o Atlântico*, 4ª Edição, Livraria Sá da Costa Editora, Lisboa.
- Rocha, J., Morgado, P., 2007. A complexidade em geografia. In *Geophilia - o sentir e os sentidos da geografia*. Lisboa, Centro de Estudo Geográfico, Faculdade de Letras, Universidade de Lisboa, 137-153.
- Rodrigues, M.L., 2020. The Limestone Massif of Estremadura. In Vieira, G., Zêzere, J.L., Moura, C., (Eds.), *Landscapes and Landforms of Portugal*, *World Geomorphological Landscapes*, 229-250.
- Rodrigues, M.L., 1998. *Evolução geomorfológica quaternária e dinâmica actual, aplicações ao ordenamento do território, exemplos no Maciço Calcário Estremenho*. Tese de Doutoramento em Geografia Física, Faculdade de Letras da Universidade de Lisboa.
- Rose, M., Altschul, J., 1988. An overview of statistical method building. In James, W.J., Sebastian, I., (Eds.). *Quantifying the Present and Predicting the Past: Theory, Method and Application of Archaeology Predictive Modelling*. Bureau of Land Management Service Center, Denver, 173-256.
- Roy B., 1996. *Multicriteria methodology for decision aiding*. Dordrecht. Kluwer Academic.
- Rua, H., 2004. *Os sistemas de informação geográfica na deteção de villae em meio rural no Portugal Romano: um modelo preditivo*. Tese de doutoramento apresentada ao Instituto Superior Técnico.
- Rua, H., 2009. Geographic information systems in archaeological analysis: a predictive model in the detection of rural Roman villae. *Journal of Archaeological Science* 36, 224-235.

- Saaty, T., 1988. The Analytic hierarchy process. Planning, priority setting, resource allocation. University of Pittsburg, Pittsburg.
- Saaty, T., 1990. How to make a decision: the analytic hierarchy process. *European Journal of Operational Research* 48, 9-26.
- Saaty, T., 2008. Decision making with the analytic hierarchy process. *International Journal of Services Sciences* 1 (1), 16.
- Saaty, T., Vargas, L.G., 2001. *Models, Methods, Concepts and Applications of the Analytic Hierarchy Process*. Kluwer Academic, Boston.
- Sánchez, J.C., Pelgrom, J., Stek, T.D., 2017. Comparing re-surveys in Isernia and Venosa (Molise and Basilicata, Italy). *Mediterranean Archaeology Archaeometry* 17, 39-52.
- Siart, C., Eitel, B., Panagiotopoulos, D., 2008. Investigation of past archaeological landscapes using remote sensing and GIS; a multi-method case study from Mount Ida, Crete. *Journal of Archaeological Science* 35 (11), 2918-2926.
- Silva, A.C., 2003. Arqueologia de Salvamento - amostragem mínima obrigatória ou direito á livre escolha? *Al-Madan* 12 (IIª série), 65- 69.
- Silva, A.C., 2015. *Modelação de nichos eco-culturais para o Paleolítico Médio*. Dissertação de Mestrado em Sistemas de Informação Geográfica e Ordenamento do Território à Faculdade de Letras da Universidade do Porto.
- Sims, C.M., 2014. *Applications of GIS analyses and ecological approaches to paleolithic sites in Portugal*. Dissertação de Mestrado, Univesrity of Louisville.
- Sing, T., Sander, O., Beerenwinkel, N., Lengauer, T., 2005. ROCr: visualizing classifier performance in R. *Bioinformatics* 21 (20), 7881, <http://rocr.bioinf.mpi-sb.mpg.de>.
- Soares, A.F., Kullberg, J.C., Marques, J.F., Rocha, R.B., Callapez, P.M., 2012. Tectono-sedimentary model for the evolution of the Silves Group (Triassic, Lusitanian Basin, Portugal). *Bull. Soc. Géol. France* 183 (3), 203-216.
- Soares, A.F., Lapa, M.L.R., Marques, J.F., 1986. Contribuição para o conhecimento da litologia das unidades Meso-Cenozoicas da Bacia Lusitaniana a Norte do «Acidente» da Nazaré (sub-zone setentrional). *Memórias e Noticias Publ. Mus. Lab. Min. Geol.* 102, Universidade de Coimbra.
- Soares, L., Costa, A., Gomes, A., 2010. Geografia, Arqueologia e sistemas de informação geográfica: exemplos prospetivos de articulação. https://sigarra.up.pt/flup/pt/pub_geral.pub_view?pi_pub_base_id=76916&pi_pub_r1_id=
- Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. *Science* 240 (4857), 1285-1293.

- Tagil, S., Jenness, J., 2008. GIS-Based Automated Landform Classification and Topographic, Landcover and Geologic Attributes of Landforms Around the Yazoren Polje, Turkey. *Journal of Applied Sciences* 8 (6), 910-921.
- Tehrany, M.S., Kumar, L., Jebur, M.N., Shabani, F., 2019. Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods. *Geomat. Nat. Haz. Risk* 10, 79-101.
- Teixeira, C., Zbyszewski, G., Torre de Assunção, C., Manupella, G., 1968. Mapa geológico e Notícia explicativa da Folha 23-C, Leiria - Carta Geológica de Portugal, na escala de 1/50 000, Lisboa.
- Texier, J.P., Cunha-Ribeiro, J.P., 1991/1992. Les formations quaternaires du bassin du Lis: leur importance pour la chronostratigraphie de l'acheuléen portugais. *Cadernos de Arqueologia* 8 e 9.
- Tolles, J., Meurer, W.J., 2016. Logistic Regression Relating Patient Characteristics to Outcomes. *JAMA Guide to Statistics and Methods* 316 (5), 533-534.
- Triantaphyllou, E., Shu, B., Nieto Sanchez, S., Ray, T., 1998. Multi-Criteria Decision Making: An Operations Research Approach. *Encyclopedia of Electrical and Electronics Engineering* ~15, 175-186, J.G. Webster, John Wiley & Sons, New York.
- Vaughn, S., Crawford, T., 2009. A predictive model of archaeological potential: An example from northwestern Belize. *Applied Geography* 29, 542-555.
- Verhagen, P., 2007. Case studies in archaeological predictive modelling. Leiden University Press, the Netherlands.
- Verhagen, P., Whitley, T.G., 2012. Integrating Archaeological Theory and Predictive Modeling: A Live Report from the Scene. *Journal of Archaeological Method Theory* 19, 49-100.
- Verhagen, P., Kamermans, H., Leusen, M., Deeben, J., Hallewas, D., Zoetbrood, P., 2007. First thoughts on the incorporation of cultural variables into predictive modelling. *Verhagen Ph, Case Studies in Predictive Modeling* 14, 203-210.
- Voogs, H., 1983. Multicriteria evaluation for urban and regional planning. Pion Limited, Londres.
- Wachtel, R.Z., Shimon, G., Gideon, S., 2018. Predictive modeling for archaeological site locations: Comparing logistic regression and maximal entropy in north Israel and north-east China. *Journal of Archaeological Science* 92, 28-36.
- Warden, L., Moros, M., Neumann, T., *et al.*, 2017. Climate induced human demographic and cultural change in northern Europe during the mid-Holocene. *Sci. Rep.* 7, 15251.
- Warren, R.E., 1990. Predictive modelling of archaeological site location: a primer. In Allen, K.M.S., Green, S.W. (Eds.), *Interpreting Space: GIS and Archaeology*. Taylor & Francis, London, 90-111.

- Warren, R.E., Asch, D., 2000. A Predictive Model of Archaeological Site Location in the Eastern Prairie Peninsula. Taylor & Fisher, London, 5-25.
- Warren, R.E., Asch, D., 2005. A predictive model of archaeological site location in the eastern prairie Peninsula. In Westcott, K.L., Brandon, R.J., (Eds.), Practical Applications of GIS for Archaeologists: A Predictive Modelling Kit. Taylor & Francis e-Library, Philadelphia, 6-36.
- Weiss, A., 2001. Topographic Position and Landform Analysis. Poster presentation. ESRI user conference, San Diego, CA.
- Wescott, K., Brandon, R., 2000. Practical Applications of GIS for Archaeologists: A Predictive Modeling Kit. Taylor & Francis, London.
- Wheatley, D., 2004. Making space for an archaeology of place. Internet Archaeology 15 <https://doi.org/10.11141/ia.15.10>.
- Wheatley, D., Gillings, M., 2002. Spatial Technology and Archaeology: The Archaeological Applications of GIS. Taylor & Francis, London.
- Wheeler, P., Gordon-Brown, L., Ward, M., 2010. Geographical information systems in Victorian secondary schools: current constraints and opportunities. International Research in Geographical and Environmental Education 19 (2), 155-170.
- Wood., S.N., 2017. Generalized Additive Models. An introduction with R, Second Edition. Chapman and Hall/CRC.
- Zbyszewski, G., 1959. Étude structurale de l'aire typhonique de Caldas da Rainha (Portugal). Memórias do Serviço Geológico de Portugal 3, Lisboa.
- Zbyszewski, G., Ferreira, O.V., Penalva, C., Teixeira, A.J.M., 1980. Nova contribuição para o conhecimento do Paleolítico da Quinta do Cónego (Cortes, Leiria). Guimarães, Sociedade Martins Sarmento, Revista de Guimarães 90, 181-189.
- Zhu, X., Chen, F., Guo, H., 2018. A Spatial Pattern Analysis of Frontier Passes in China's Northern Silk Road Region Using a Scale Optimization BLR Archaeological Predictive Model. Heritage 1, 15-32.
- Zilhão, J., (Coord.) 1996. Arte Rupestre e Pré-História do Vale do Côa. Trabalhos de 1995-1996, Relatório Científico ao Governo da República Portuguesa elaborado nos termos da resolução do concelho de ministros n.º 4/96, de 1 de janeiro. Ministério da Cultura, 453 p.
- Zilhão, J., 1997. O Paleolítico Superior da Estremadura Portuguesa. 2 vols. Lisboa: Colibri
- Zilhão, J., Almeida, F., 2002. The Archaeological Framework. In Zilhão, J., Trinkaus, E., (Eds.) Portrait of the artist as a child. The Gravettian Human Skeleton from the Abrigo do Lagar Velho and its Archaeological Context". Lisboa, Instituto Português de Arqueologia. Trabalhos de Arqueologia 22, 29-57.

- Zilhão, J; Trinkaus, E., (Eds.) 2002. Portrait of the Artist as a Child. The Gravettian Human Skeleton from the Abrigo do Lagar Velho and its Archaeological Context. Lisboa, Instituto Português de Arqueologia. Trabalhos de Arqueologia 22.
- Zilhão, J., Almeida, F., 2002. The Archaeological Framework. In Zilhão, J.; Trinkaus, E., (Eds.) Portrait of the artist as a child. The Gravettian Human Skeleton from the Abrigo do Lagar Velho and its Archaeological Context. Lisboa, Instituto Português de Arqueologia, Trabalhos de Arqueologia 22, 29-57.
- Zilhão, J., Angelucci, D., Arnold, L., Demuro, M., Hoffmann, D.L., Pike, A.W.G., 2021. A revised, Last Interglacial chronology for the Middle Palaeolithic sequence of Gruta da Oliveira (Almonda karst system, Torres Novas, Portugal). Quaternary Science Reviews 258, 106885.

