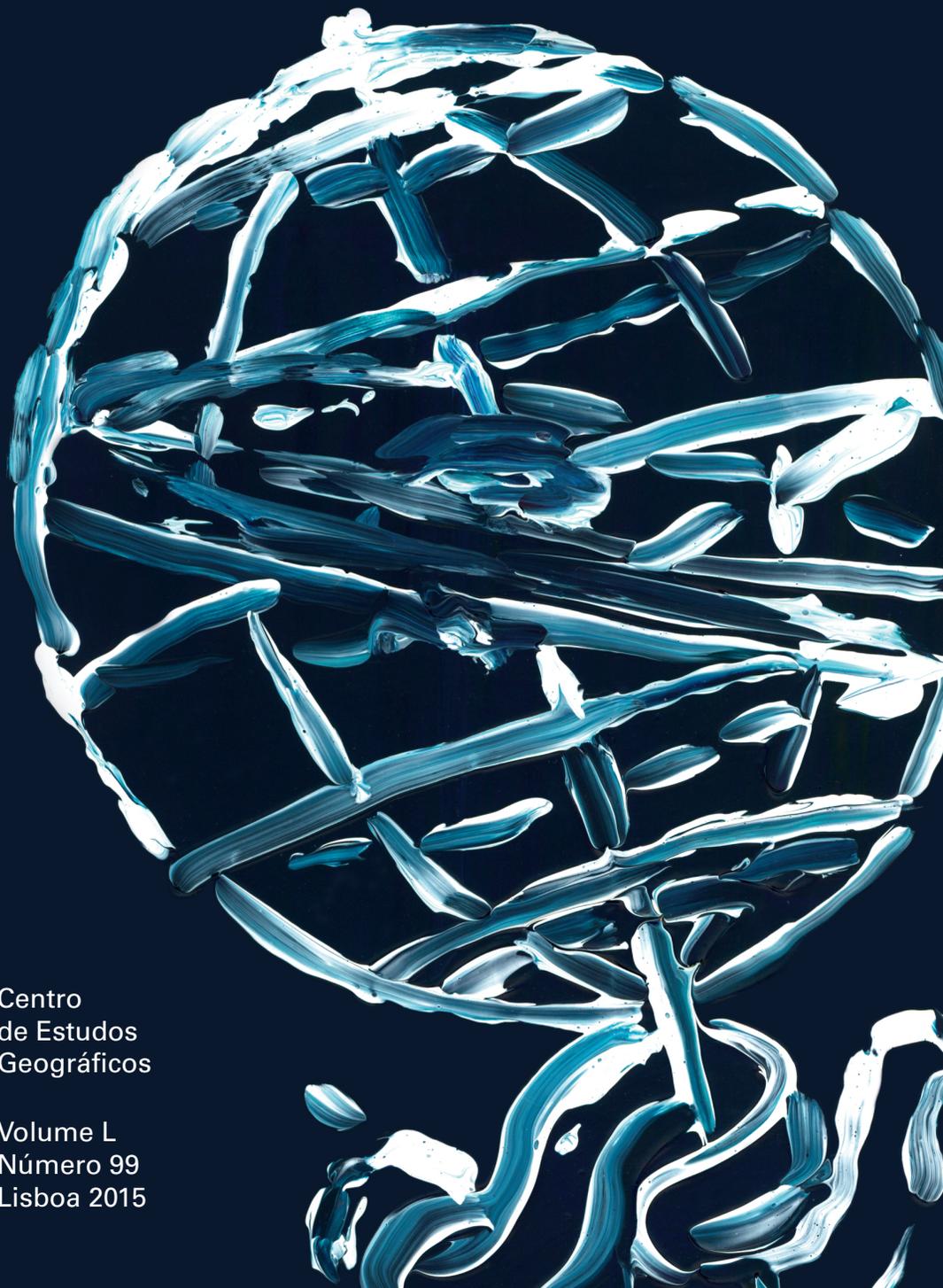


# FINISTERRA

Revista Portuguesa de Geografia



Centro  
de Estudos  
Geográficos

Volume L  
Número 99  
Lisboa 2015

## MODELAÇÃO PREDICTIVA DA VEGETAÇÃO NATURAL POTENCIAL DO CONCELHO DE LOURES

FRANCISCO GUTIERRES<sup>1</sup>

LEANDRO GABRIEL<sup>1</sup>

ANTÓNIO EMÍDIO<sup>1</sup>

PAULA MENDES<sup>2</sup>

CARLOS NETO<sup>1</sup>

EUSÉBIO REIS<sup>1</sup>

**Resumo** – O conceito de Vegetação Natural Potencial (VNP) e a sua representação cartográfica assume uma importância primordial para a maioria dos países europeus nas questões relacionadas com o restauro de habitats. Dada a relação existente entre as séries de vegetação de um território e os factores ambientais, este artigo visa o desenvolvimento de modelação predictiva da VNP para o concelho de Loures (AML). Atendendo à possibilidade de integração de um vasto conhecimento empírico, utilizou-se uma abordagem de modelação precedida de classificação de séries de vegetação (*classification-then-modelling*). Para averiguar a relação entre 6 séries de vegetação e um conjunto de 8 variáveis ambientais recorreu-se a Modelos de Distribuição de Espécies (SDM) aplicados ao nível da comunidade, suportados em Sistemas de Informação Geográfica (SIG) e em modelos de regressão, *machine learning* e *rule-based*. Os resultados obtidos permitiram aferir o modo como os gradientes ecológicos determinam a ocorrência das séries de vegetação. A cartografia predictiva da VNP resultante do modelo da Máxima Entropia, foi ainda validada com a cartografia oficial da VNP do concelho de Loures (precisão global de 88%). Dado que a gestão e conservação da biodiversidade é frequentemente desenvolvida a escalas de grande detalhe, o SDM possibilita a integração de observações directas de comunidades vegetais, e uma interpretação da distribuição local da VNP ao longo de gradientes ambientais.

**Palavras-chave:** Vegetação Natural Potencial, séries de vegetação, modelação espacial e ordenamento do território, Loures (Portugal).

---

Recebido: Outubro 2013. Aceite: Abril 2014.

<sup>1</sup> Instituto de Geografia e Ordenamento do Território da Universidade de Lisboa (IGOT-UL). Edifício IGOT. Avenida Prof. Gama Pinto. 1649-003 Lisboa. E-mail: franciscogutierres@campus.ul.pt; leandrogabriel@campus.ul.pt; a.emidio@campus.ul.pt; cneto@campus.ul.pt; eusebioreis@campus.ul.pt

<sup>2</sup> Bolseiro de Investigação do Departamento de Paisagem Ambiente e Ordenamento, Escola de Ciências e Tecnologia da Universidade de Évora, Colégio Luis Verney, Rua Romão Ramalho, 57 7000. E-mail: paulabm@uevora.pt

**Abstract** – MODELLING THE POTENTIAL NATURAL VEGETATION IN THE LOURES MUNICIPALITY (LISBON METROPOLITAN AREA). The concept of Potential Natural Vegetation (PNV) and its mapping have become extremely important within the scope of habitat restoration in almost every European country. The aim of this study is to predict the PNV in Loures based on the vegetation series and the main environmental variables. The modelling approach is based on the distribution of communities referred to as *classification-then-modelling*. Subsequently, several statistical model-fitting techniques, such as regression models, machine learning and rule-based, were successfully applied to the survey data (vegetation series; and 8 environmental/predictor variables). The spatial database was organized as a Geographic Information System (GIS) and was also used to perform the Species Distribution Models (SDM) at community level. The results show a high correspondence between the vegetation series and the environmental gradients. The predicted PNV maps based on the Maximum Entropy Model were validated with the official map of the PNV of Loures, and presented an overall accuracy of 88%. Often, conservation planning and biodiversity resource management is carried out at more detailed scales, where SDM allows integration of community direct observations and improve our interpretation of PNV local distributions along environmental gradients.

**Key words:** Potential Natural Vegetation, vegetation series, spatial modeling and land use planning, Loures (Portugal).

**Résumé** – UM MODÈLE PRÉDICTIF DE LA VÉGÉTATION NATURELLE DU *CONCELHO* DE LOURES. Le concept de Végétation Naturelle Potentielle (VNP) et sa représentation cartographique ont une importance primordiale dans la plupart des États européens, quand il s'agit de restaurer des habitats. En prenant en compte le rapport existant entre les séries de végétation et les facteurs environnementaux du territoire, on a cherché une modélisation prédictive de la VNP du *concelho* de Loures (AML). Tout en intégrant une importante connaissance empirique, on a introduit dans ce modèle une classification des séries de végétation (*classification-then-modelling*). Pour établir la relation existant entre 6 séries de végétation et 8 variables environnementales, on a eu recours à des Modèles de Distribution des Espèces (SDM), applicables au niveau des communautés en Systèmes d'Information Géographique (SIG) et en modèle de régression, *machine learning* et *rule-based*. Les résultats montrent comment les gradients écologiques déterminent les séries de végétation. La cartographie prédictive de la VNP, selon le modèle d'Entropie Maximale, a ainsi été validée à 88 % par rapport à la cartographie officielle de la VNP du *concelho* de Loures. Étant donné que la gestion et conservation de la biodiversité sont souvent pratiquées à des échelles de grand détail, le SDM permet d'intégrer les observations directes de communautés végétales et l'interprétation de la répartition locale de la VNP au long de gradients environnementaux.

**Mots clés:** Végétation Naturelle Potentielle, séries de végétation, modèle spatial, gestion du territoire, Loures (Portugal).

**摘要** – 斐列斯市潜在天然植被预测模型的建造。潜在天然植被这个概念和它的制图对大部分欧洲国家在恢复生境问题上至为重要。以存在于一个地域的植被系列和环境因素之间的关系为基础, 本文将阐述斐列斯市潜在天然植被预测模型的建造。由于能够综合到广泛的经验性资料, 我们在建造过程中采取了先把植被系列分类然后再制图的方法(*classification-then-modelling*)。为了分析6个植被系列和8个环境变数之间的关系, 我们选择了物种分布模式(SDM), 把它用于群体层次, 以基于机器学习和模式识别的地理信息系统和回归模式去支持。所得结果允许我们衡量环境梯度如何影响植被系列的形成, 接着再透过最大熵模式制成了潜在天然植被的预测模型, 在与斐列斯市的官方潜在天然植被地图作比较之后, 其准确性得到肯定(大致上有88%吻合)。由于生物多样性的管理和保护通常是以较大的细节为尺度, 物种分布模式容许我们直接观察植物群体, 以及了解潜在天然植被循着环境梯度而变的地方性分布。

**关键字:** 潜在天然植被、植被系列、空间模拟、地域规划。

## I. INTRODUÇÃO

### 1. Vegetação Natural Potencial (VNP) e os Modelos de Distribuição de Espécies (SDM)

A Vegetação Natural Potencial (VNP) materializa o estágio de máximo desenvolvimento da vegetação num dado habitat, que ocorreria num determinado território e suas envolventes se não existisse qualquer influência humana (Tüxen, 1956; Westhoff e van der Maarel, 1978; Breidenkamp *et al.*, 1998). A VNP define-se através de comunidades vegetais que atingiram o seu máximo ecológico estável, estando em equilíbrio com as condições do biótopo (Neto *et al.*, 2008). Assim, a série de vegetação compreende o tipo de vegetação representativo da etapa madura ou cabeça de série, as comunidades iniciais ou subseriais que a substituem, bem como os espaços ocupados pelas comunidades existentes e os factores mesológicos que configuram os seus habitats naturais (Rivas-Martinez *et al.*, 1999). No entanto, deve ser salientado que a sua permanência como VNP persiste até que essas condições se alterem, ou seja, apenas se estas condições permanecerem inalteradas a comunidade manterá a sua estrutura, composição e funcionamento interno (Ricotta *et al.*, 2000; Neto *et al.*, 2008; Mucina, 2010; Loidi e Fernández-González, 2012). Recentemente, a modelação da VNP tem registado avanços significativos com o desenvolvimento de diversos modelos e *software* com grande capacidade de processamento. A sua determinação revela-se especialmente importante na reconstituição da vegetação natural após acção antrópica, tendo em conta o carácter predictivo associado ao conceito (Zampieri e Lionello, 2008; Neto *et al.*, 2008). Com base no cariz relativamente determinístico do processo de sucessão ecológica e na sua relação com os principais factores ambientais – climáticos, geomorfológicos e edáficos – é possível modelar e prever a vegetação potencial para uma determinada área, através da análise da comunidade vegetal actualmente presente e do conhecimento das séries de vegetação desse território.

Durante a última década, a Europa tem vindo a assistir a uma crescente consciencialização sobre a importância da cartografia da VNP e sua utilidade como uma ferramenta para apoiar as actuais estratégias de Ordenamento e Planeamento do Território (Bohn *et al.*, 2003; Cross, 2006; Capelo *et al.*, 2007; Zampieri e Lionello, 2008; Neto *et al.*, 2008; Mucina, 2010). Os mapas de distribuição da VNP, e também de distribuição potencial de habitats e de espécies, representam um importante contributo para as Florestas, Agricultura, Recursos hídricos, Conservação da Natureza e Ordenamento do Território. Ricotta *et al.* (2002) considera que os mapas de distribuição da VNP representam uma alternativa ecológica significativa para os modelos neutros de paisagem, especialmente quando a avaliação dos efeitos da estrutura da paisagem sobre os processos ecológicos está em causa.

Guisan e Zimmermann (2000), por outro lado, assumem que os modelos predictivos foram desenvolvidos para abranger aspectos tão diversos como a Biogeografia, gestão e restauro de habitats, gestão de espécies invasoras, modelação de comunidades e ecossistemas, predição dos efeitos das alterações climáticas nas espécies e

ecossistemas e, deste modo, a distribuição potencial de uma espécie ou comunidade constitui claramente o primeiro passo na investigação aplicada.

Nos últimos anos tem-se verificado um aumento crescente da utilização de métodos estatísticos (análise de regressão e outros), de Inteligência Artificial (*machine learning*) ou baseado em regras (*rule-based*) com interligação às Tecnologias de Informação Geográfica, com o intuito de modelar espacialmente as distribuições de espécies ou comunidades em função de dados biológicos e ambientais (Gutierrez, 2014).

Esta prática é usualmente conhecida como Modelação de Distribuição de Espécies (SDM), podendo também ser referida como modelação do nicho ecológico de Hutchinson<sup>1</sup>, bioclimática, ambiental ou de adequabilidade de habitat (Gutierrez, 2014).

O desenvolvimento de um SDM é suportado em dados de ocorrências de espécies ou comunidades (podendo ser presença ou presença/ausência) e variáveis ambientais (predictivas) que se julga influenciar a adequabilidade do habitat, e consequentemente a espécie ou comunidade (Ferrier *et al.*, 2002a,b; Hirzel *et al.*, 2002; Franklin *et al.*, 2005; Phillips *et al.*, 2009; Múcher *et al.*, 2009).

Diversos autores (*e.g.* Guisan e Zimmermann, 2000; Austin *et al.*, 2006) alertam que a escolha das variáveis preditivas tem uma forte influência no desempenho dos SDM. Actualmente tem sido destacada a importância de uma melhor e mais precisa definição das relações dos gradientes directos (variáveis proximais com efeito fisiológico directo, tal como o teor de água, luz, nutrientes) com a distribuição das espécies ou comunidades, diferenciadas das variáveis predictivas indirectas (variáveis distais, sem efeito fisiológico directo, tal como a altitude, declive e exposição) a integrar no processo de modelação. Barry e Elith (2006) acrescentam que os estudos actuais de modelação não são susceptíveis de incluírem um conjunto significativo de variáveis predictivas proximais (especializadas). Neste sentido, é recomendado o estabelecimento de modelos conceptuais que expressem as relações das distribuições bióticas com as variáveis predictivas distais, devendo ser seleccionadas as variáveis mais apropriadas em função dos dados de base disponíveis, para que os modelos apresentem níveis de erro aceitáveis. Contudo, é necessário ter em conta que os modelos com múltiplas variáveis predictivas poderão apresentar problemas de autocorrelação espacial, dependência espacial (escalas globais *versus* escalas locais) e ainda de multicolinearidade, pelo facto destas variáveis representarem variáveis indirectas (distais), e não variáveis directas (proximais) relacionadas com a adequabilidade ou abundância de espécies ou comunidades (Lennon, 2002; Miller e Franklin, 2002; Diniz-Filho *et al.*, 2003; Segurado *et al.*, 2006; Miller *et al.*, 2007; Dormann *et al.*, 2007; Osborne *et al.*, 2007; Zhang *et al.*, 2008).

A área de estudo foi seleccionada em virtude da diversidade litológica, pedológica e topográfica presente, o que determina a ocorrência de uma grande diversidade de séries de vegetação com comunidades vegetais importantes para protecção e conservação, incluídas no Anexo I da Directiva Habitats. Deste modo, dada a disponibilidade de informação cartográfica e bases de dados, justifica-se o interesse do desenvolvimento da modelação predictiva da VNP neste território exíguo. O outro factor de selecção está relacionado com a crescente necessidade de disponibilização de

cartografia predictiva da VNP a escalas de grande detalhe, baseadas em abordagens de modelação que possibilitem a integração de observações directas de comunidades vegetais, e uma interpretação da distribuição local da VNP ao longo de gradientes ambientais (Franklin, 2009). McPherson *et al.* (2006) e Bustamante e Seoane (2004) descrevem a limitação da aplicabilidade de mapas de distribuição de pequena resolução publicados em atlas ou guias de campo (caso da Carta da VNP da Europa na escala 1:2 500 000 produzida pelo *Institute für Bundesamt für Naturschutz (BfN)* (Bohn *et al.*, 2003) ou da Carta da VNP para Portugal Continental à escala 1:1 000 000 (Capelo *et al.*, 2007)) no Ordenamento e Planeamento do Território à escala local. Acresce o facto dos factores ambientais que controlam a distribuição das espécies e comunidades se modificarem em função da escala, o que torna difícil a sua redução (*downscaling*) e a sua aplicabilidade em pequenas áreas. Ao invés, o carácter predictivo do SDM permite a obtenção da distribuição das espécies, comunidades e adequabilidade de habitat, podendo ainda este modelo integrar e responder a impactos de mudanças ambientais, quer sejam motivadas por alterações climáticas ou devido às dinâmicas de transformação do uso e ocupação do solo (Wu e Smeins, 2000; Gibson *et al.*, 2004; Rushton *et al.*, 2004; Araújo *et al.*, 2005a; Burgman *et al.*, 2005; Wisz *et al.*, 2008). Portanto, é fundamental definir o padrão e a escala de análise (Elton, 1927; Scott *et al.*, 2002), em que o objectivo poderá basear-se apenas na caracterização geral dos factores que determinam a distribuição e ocorrência das espécies e comunidades proveniente de cartografia de referência a uma escala regional ou global ou, por outro lado, no desenvolvimento de cartografia predictiva desde a escala local à global assente em diferentes técnicas de modelação (Cumming, 2000; Segurado e Araújo, 2004; Thuiller *et al.*, 2004; Araújo *et al.*, 2005b). No presente estudo, assume-se que o conceito da VNP é fortemente determinado por múltiplos factores bióticos e abióticos e a sua utilização visa a inferência ecológica das relações séries de vegetação-variáveis ambientais com recurso a SDM.

Neste sentido, o presente estudo tem o objectivo principal de produzir a carta da VNP do concelho de Loures, tal como depreender os gradientes ecológicos associados às séries de vegetação, com recurso à aplicação de várias técnicas de modelação predictiva em ambiente SIG. Constitui ainda objectivo testar o efeito de selecção das variáveis predictivas no processo de modelação da VNP, de forma a definir modelos parcimoniosos e que utilizem o conjunto mais apropriado de variáveis. Deste modo, pretende-se contribuir para a definição de modelos calibrados de adequabilidade das séries de vegetação baseados em regras de decisão, comparar e testar diferentes alternativas metodológicas e fornecer orientações para a investigação da VNP a escalas de pormenor.

## 2. A área de estudo (Concelho de Loures)

A área de estudo desta análise corresponde ao concelho de Loures. Pertencente à Área Metropolitana de Lisboa, situa-se na margem direita do estuário do Tejo a norte da cidade de Lisboa (fig. 1).

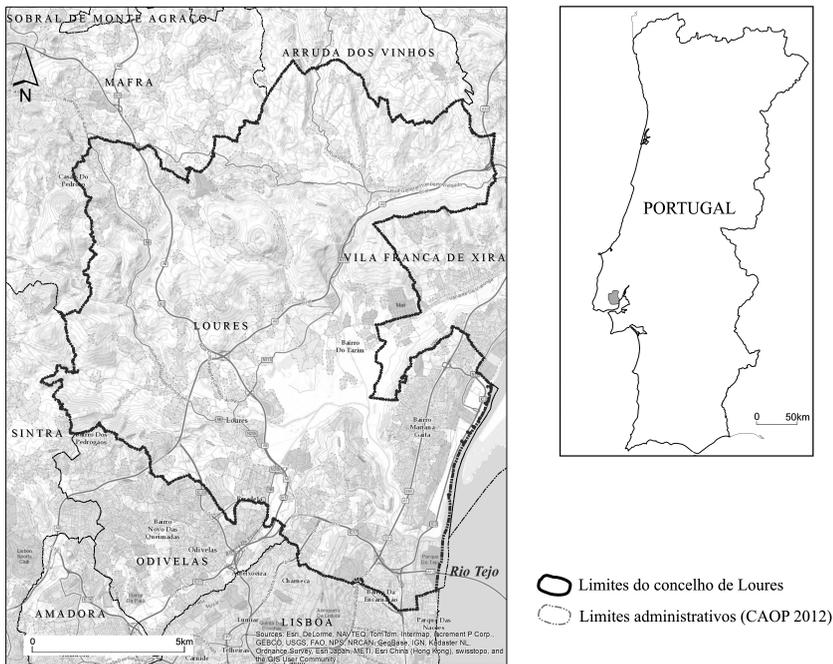


Fig. 1 – Enquadramento da área de estudo.

*Fig. 1 – Location of the study area.*

Com uma área de aproximadamente 169 km<sup>2</sup>, o concelho de Loures apresenta grande diversidade morfológica. No sector Norte e Noroeste predominam as áreas de maior altitude (de 200 a 400 m) e forte declive, devido à existência de vales fortemente encaixados por onde escoam o rio Trancão e seus afluentes. No sector Sul e Sudeste situam-se as áreas mais planas, destacando-se a várzea de Loures e a frente ribeirinha do Tejo. Esta diversidade morfológica constitui um importante factor explicativo da distribuição da vegetação uma vez que determina a distribuição das áreas de erosão, acumulação e escoamento de água, assim como a diversidade de microclimas. De acordo com Magalhães *et al.* (2002), podem-se definir três situações morfológicas distintas no concelho de Loures: a) Zonas adjacentes às linhas de água – áreas planas correspondentes a solos aluvionares; b) Cabeços – áreas mais expostas aos ventos e à erosão hídrica; e c) Encostas – zonas situadas entre as áreas adjacentes às linhas de água e os cabeços, caracterizadas por uma forte erosão hídrica e, conseqüentemente, uma menor profundidade do solo.

Ao situar-se na bacia ceno-antropozóica do Tejo, esta área “apresenta uma enorme variedade de materiais de enchimento que vão do Oligocénico até à actualidade. As diferentes formações litológicas exibem afloramentos superficiais, com extensão e características variáveis. O mosaico de comunidades vegetais está condicionado pela natureza dos sedimentos que constituem as várias formações litológicas, sobretudo no que respeita à granulometria dos materiais e à coesão das partículas.” (Neto *et al.*, 2007: 202).

Em termos biogeográficos, de acordo com Costa *et al.* (1998) e Rivas-Martínez (2007), o concelho de Loures insere-se nas seguintes unidades biogeográficas: Reino Holártico, Região Mediterrânea, Sub-Região Mediterrânea Ocidental, Província Lusitano-Andalusa Litoral, Subprovíncia Divisório Portuguesa-Sadense, Sector Divisório Português, Subsector Divisório Português Setentrional e Distrito Estremenho Português, apenas na parte norte (freguesia de Bucelas) e Subsector Divisório Português Meridional e Distrito Olissiponense na restante área de estudo.

Do ponto de vista bioclimático, de acordo com Rivas-Martínez *et al.* (2011) e nos mapas de zonamento de Monteiro-Henriques (2010), o concelho de Loures encontra-se no Macrobioclima Mediterrâneo pluviestacional-oceânico, nos andares termo/mesomediterrâneo de ombrótipo sub-húmido a húmido.

Deste modo, em termos de vegetação climatófila domina o cercal de carvalho-cerquinho pertencente à associação *Arisaro simorrhini-Quercetum broteroi* (Vila-Viçosa, 2012). Em termos de diagnose este cercal corresponde à série mediterrânea pluviestacional, termo a mesomediterrânea, hiperoceânica a euoceânica, sub-húmida superior a húmida basófila e calcária, dominada por *Quercus broteroi* (Vila-Viçosa, *op.cit.*). Este bosque climácico contacta catenalmente com zambujais de *Viburno tinii-Oleetum sylvestris* sobre solos calcários margosos de carácter vértico, sujeitos a fendilhamento no período seco. Em posição edafoxerófila, contacta com azinhais de *Lonicero implexae-Quercetum rotundifoliae* e com o sobreiral de *Asparagoaphylli-Quercetum suberis*, que ocorre sobre arenitos e solos basálticos. Este contacta ainda com associações edafo-higrófilas dos salgueirais de *Clematido campaniflorae-Salicetum neotrichae* (Costa *et al.*, 2010). A vegetação ripícola é constituída essencialmente por freixiais de *Irido foetidissimae-Fraxinetum angustifoliae* (Costa *et al.*, 2012).

Neste território dominam ainda as Permaséries Rupícolas da comunidade termo a mesomediterrânea, sub-húmida a húmida, de gretas largas do Divisório Português e do Arrabidense associadas a rochas calcárias (*Sileno longiciliae-Antirrhinetum linkiani*); a comunidade mesomediterrânea e luso-extremadurense, que atinge o Oeste-Estremenho de forma empobrecida (*Asplenio ceterach-Cheilanthesetum acrosticae*); a comunidade mesomediterrânea, de fendas estreitas de calcários do Divisório Português e do Arrabidense (*Narcisso calcicolae-Asplenietum rutae-murariae*) e a comunidade de rochas calcárias planas do Sudoeste da Península Ibérica (*Sedetum micrantho-sediformis*).

Relativamente aos biótopos de salgados destaca-se a presença das comunidades halofíticas *Spartinetum maritimae*, *Halimiono portulacoidis-Salicornietum patulae* e *Holoschoeno-Juncetum acuti*. A primeira caracteriza-se por ser uma comunidade hidrófila, pioneira, mono ou pauciespecífica, dominada por *Spartina maritima*, e que se instala em sedimentos marinhos ou fluvio-marinhos, mais ou menos finos, saturados na maré baixa e sujeitos à influência diária das marés. A comunidade de *Halimiono portulacoidis-Salicornietum patulae* caracteriza-se por ser uma associação mediterrânea, cuja sinecologia está associada a locais planos arenosos ou limosos saturados dos níveis baixos dos sapais inundados diariamente, caracterizada pela presença de *Salicornia ramosissima*, *Suaeda albescense* *Puccinellia iberica* (Costa *et al.*, 2012).

## II. METODOLOGIA

A modelação espacial da distribuição da VNP do concelho de Loures, baseou-se num conjunto de variáveis predictivas, consideradas representativas das características biofísicas presentes no território estudado e em dados de distribuição de comunidades vegetais. Sobre estes conjuntos de dados geográficos foi aplicado um conjunto de ferramentas de análise exploratória e de modelação (estatística *machine learning* e *rule-based*), que possibilitaram a criação de modelos predictivos SDM ao nível da comunidade, e por sua vez, a determinação da VNP.

### 1. Amostragem

Os levantamentos das presenças das comunidades vegetais foram efectuados directamente no terreno com recurso a um GPS com correcção diferencial, com base na carta da vegetação actual disponibilizada pela Câmara Municipal (CM) de Loures e ainda com recurso a foto-interpretação de ortofotomapas de 2005 (DGT).

Os pontos de presença foram determinados segundo a amostragem baseada na alocação aleatória proporcional de amostras entre estratos (também conhecida por *proportional random-stratified sampling*) (Ruxton e Colegrave, 2006; Theobald, 2007).

Os pontos de ausência, igualmente determinados por uma amostragem aleatória estratificada, correspondem às áreas sociais onde existe uma menor probabilidade (ou mesmo ausência) de se encontrarem comunidades vegetais inseridas na dinâmica serial das respectivas séries de vegetação. Na alocação das pseudo-ausências para cada série de vegetação foram ainda consideradas as áreas onde se observam condições abióticas extremas entre os biogeossistemas em que se integram as respectivas séries de vegetação, como por exemplo, biótopos halofíticos *versus* biótopos calcícolas. Foi adoptada uma prevalência com valor 1 isto é, mesma quantidade de dados de presença e ausência (quadro I), por produzir um equilíbrio ideal entre erros de omissão e comissão nos modelos SDM (Jiménez-Valverde *et al.*, 2009).

Quadro I – Dados amostrais.  
*Table I – Sampled data.*

Séries de Vegetação	Presença	Ausência
Geossérie Ripícola	93	93
<i>Asparago aphylli-Quercetum suberis</i>	138	138
<i>Arisaro simorrhini-Quercetum broteroi</i>	276	276
<i>Geopermasigmetum</i> Rupícola	21	21
<i>Geopermasigmetum</i> Halófitico	36	36
<i>Viburno tini-Oleetum sylvestris</i>	210	210
	774	774

## 2. Selecção das Variáveis Predictivas

O hipervolume de Hutchinson sugere a existência de múltiplos factores determinantes na distribuição das espécies e comunidades, constituindo um dos principais aspectos a considerar na operacionalização do SDM (Biondi *et al.*, 2004; Franklin, 2009; Monteiro-Henriques, 2010). As variáveis ambientais a integrar neste tipo de modelos são um dos principais factores passíveis de condicionar o seu desempenho (Hirzel e Le Lay, 2008; Franklin, 2009).

As variáveis predictivas integradas nos modelos SDM geralmente incluem uma mistura de variáveis contínuas e categóricas, não sendo expectável que as relações entre estas e a variável resposta sejam lineares, sendo deste modo, passível a ocorrência de interacções entre variáveis (Franklin, 2009). A mesma autora adverte que estes factores deverão ser considerados na formulação estatística do SDM.

Neste sentido, as variáveis predictivas seleccionadas resultam do modelo conceptual da VNP e da disponibilidade de cartografia de base (curvas de nível (1:10 000) (DGT), carta geológica (1:25 000), carta de solos (1:25 000) (DGADR) e carta de vegetação actual (1:25 000) (CM Loures). Estes dados de base foram integrados em formato matricial (com uma resolução de 5 m) e modelados nos *softwares* ArcGIS 10 (ESRI) e SAGAGIS, tendo sido determinadas as variáveis predictivas consideradas explicativas e necessárias para a modelação da VNP: Altitude, Declive, Exposição, Índice Topográfico de Humidade, *Sky View Factor*, Solos, Geologia e Distância ao Tejo<sup>ii</sup>.

Convém salientar que nas ciências geobotânicas tem sido amplamente reconhecida a importância da edafologia e geologia como factores que controlam os padrões de distribuição das comunidades vegetais (Neto, 2002; Blasi *et al.*, 2004, 2005; Capelo *et al.*, 2007; Monteiro-Henriques, 2010; Biondi *et al.*, 2011; Costa *et al.*, 2012; Portela-Pereira, 2013). Contudo, em função da escala de análise adoptada, que requer uma maior diferenciação territorial, é vista como fundamental a incorporação adicional de variáveis morfométricas<sup>iii</sup> para melhor aferição dos gradientes ecológicos (curvas de resposta) e da adequação óptima das séries de vegetação (Perring, 1959; Moore *et al.*, 1991; Mackey, 1993; Franklin, 1995, 2009; Lane *et al.*, 1998; Franklin *et al.*, 2000; Hutchinson e Gallant, 2000; Wilson e Gallant, 2000; Pausas e Austin, 2001; Jetz e Rahbek, 2002; Shriner *et al.*, 2002; Pausas *et al.*, 2003). Por outro lado, van Manen *et al.* (2002) e Fleishman *et al.* (2002) evidenciam a proximidade a cursos de água como um factor importante na análise da adequabilidade de habitat. Austin *et al.* (2006) acrescentam que a selecção das variáveis predictivas tem bastante importância na performance do SDM. Neste contexto, apesar de Franklin (2009) e Biondi *et al.* (2011) sugerirem a utilização da radiação solar potencial na modelação da vegetação, por se tratar de uma variável que exprime um gradiente directo e reflecte o efeito combinado do declive e da exposição no regime térmico/hídrico, optou-se por testar o efeito separado destas duas variáveis (de gradiente indirecto) (McCune, 2007; Franklin, 2009).

Por outro lado, a exclusão de variáveis predictivas deverá incidir sobre as variáveis menos correlacionadas com a variável de resposta, podendo estas relações serem determinadas mediante uma análise exploratória. Alternativamente poderá ser considerada a

aplicação de métodos de normalização numa das variáveis auto-correlacionadas (Leathwick *et al.*, 2005; Franklin, 2009). No entanto, apesar de ser reconhecida a utilidade da normalização de variáveis predictivas, não se trata de um procedimento comum no SDM (Elith e Leathwick, 2009), apesar de também se admitir que algumas ferramentas de modelação são beneficiadas com a normalização das variáveis (Rocha, 2012). Porém, dados os pressupostos dos modelos SDM (Franklin, 2009), não foram aplicados procedimentos de normalização das variáveis utilizadas no desenvolvimento do modelo da VNP do concelho de Loures. A análise exploratória das variáveis predictivas desempenha um papel fundamental na modelação adoptada (*classification-then-modelling*), em particular a Análise Canónica de Correspondências (CCA) e a Correlação de Pearson (Ferrier *et al.*, 2002b). Modelos predictivos baseados em análises de ordenação restritas revelam maior performance quando aplicados ao nível da comunidade, em que é contemplada a diversidade de espécies de múltiplas comunidades vegetais, (Ohmann e Spies, 1998; Legendre e Legendre, 1998; Guisan *et al.*, 1999; Dirnbock *et al.*, 2003). Neste sentido, a análise CCA define a dispersão máxima dos dados de espécies ou comunidades, em que os eixos da ordenação estão restritos a ser combinações lineares de variáveis ambientais, sendo possível observar a variação contínua na composição das comunidades e relacionar esta variação com gradientes ambientais (Franklin, 2009). A CCA é uma técnica de análise multivariada ou de ordenação, que incorpora regressão múltipla, na qual os eixos são extraídos não só tendo em conta a maximização da inércia (variância) explicada, mas também de modo a que a sua correlação com outro conjunto de variáveis ambientais predictivas seja também maximizada (ter Braak, 1988). A vantagem da CCA sobre os demais métodos é que pode ser utilizada para detectar a inter-relação das comunidades vegetais com os factores ambientais e ainda estudar os casos específicos da resposta de espécies às variáveis predictivas (Kent e Ballard, 1988).

As variáveis predictivas excluídas do processo de modelação deverão ser aquelas que estão fortemente correlacionadas entre si e menos correlacionadas com a variável de resposta. Deste modo, o grau e força da relação entre as variáveis predictivas pode ser determinado mediante análises exploratórias, que expressam as correspondências entre cada variável (independente) e resposta das comunidades (dependente). Antes ou durante a selecção do modelo, as variáveis predictivas podem ser testadas para avaliar a existência e grau de correlação entre as variáveis predictivas (problemas de multicolinearidade) e entre as variáveis predictivas (factores ambientais) e a variável dependente (séries de vegetação).

Os métodos baseados na ordenação CCA são sensíveis à colinearidade entre variáveis e, portanto, a inserção de variáveis redundantes leva à introdução de ruído no modelo e a interpretações erradas. Deste modo, poderão ser eliminadas as variáveis que apresentem multicolinearidade (factor de variância-expansão (VIF)  $\geq 10$  e  $R \geq 0,9$ ). A determinação da significância para o conjunto das variáveis predictivas desenvolve-se com o teste de simulação de Monte Carlo (9999 simulações).

Como a colinearidade significa que as variáveis independentes são correlacionadas, esta pode ser igualmente detectada através da matriz de correlação entre as variáveis. Assim, de forma a prevenir a ocorrência de multicolinearidade entre

variáveis, foi calculada uma matriz de coeficientes de correlação de Pearson entre as variáveis independentes (Rocha, 2012).

A ordenação CCA foi efetuada pelo método de análise por gradiente directo no *software* CANOCO for Windows 4.5, tendo sido consideradas 774 ocorrências, correspondentes a 6 séries de vegetação e 8 variáveis predictivas. A natureza e a força das relações entre cada variável predictiva e as séries de vegetação foram ainda exploradas com recurso à Correlação de Pearson no *software* SPSS 19 (IBM).

### 3. Modelação da VNP

Tendo por base o modelo conceptual da VNP, com a base de dados espaciais e com a possibilidade de integração de um vasto conhecimento empírico, aplicou-se a abordagem *classification-then-modelling* (fig. 2).

As etapas metodológicas desta estratégia de modelação consistem, numa primeira fase, na análise exploratória das variáveis predictivas e na ligação ou estabelecimento de relações entre os dados georreferenciados das comunidades com as variáveis predictivas, sendo necessário extrair para uma matriz de dados os valores das localizações das comunidades e das variáveis independentes (Franklin, 1995). Posteriormente, foram aplicados métodos baseados em algoritmos estatísticos (Modelos Lineares Generalizados - GLM, *machine learning* (Máxima Entropia - Maxent), Redes Neurais Artificiais (RNA), Máquinas de Suporte Vetorial (SVM) e *rule-based* (Algoritmo Genético para Produção de um Conjunto de Regras - GARP), que descrevem/integram as relações entre as ocorrências e as variáveis predictivas (Franklin, 1998, 2009; Guisan *et al.*, 1999; Rocha, 2012).

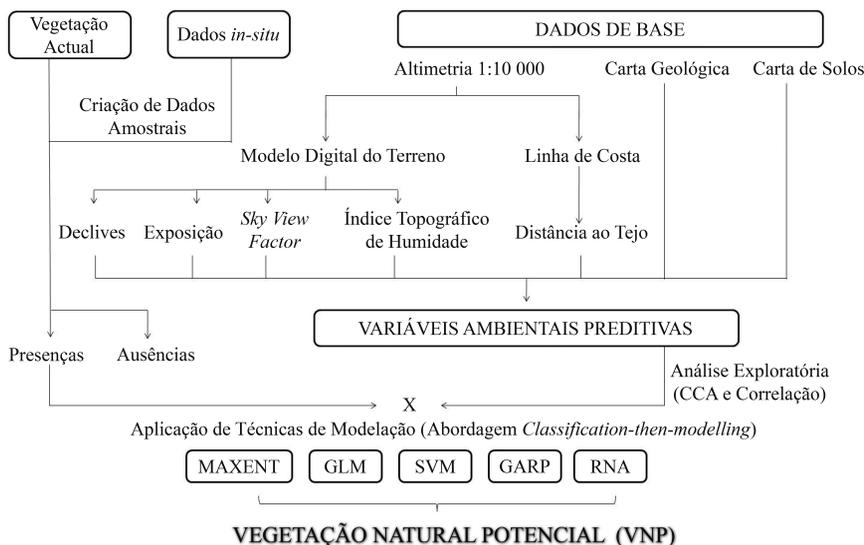


Fig. 2 – Modelo conceptual.  
Fig.2 – Conceptual model.

Foram utilizadas cinco técnicas de modelação com o objectivo de determinar quais os modelos que produzem um melhor ajustamento entre a variável dependente e variáveis independentes, e que podem ser eficazmente utilizados para fazer a representação espacial contínua das séries de vegetação em apreciação (quadro II). Vários estudos demonstraram que os modelos testados responderam satisfatoriamente à análise de padrões e distribuição de espécies e comunidades vegetais, nomeadamente: o modelo Maxent, aplicado com sucesso por Elith *et al.* (2006) e Phillips e Dudík (2008); o modelo GARP por Stockwell e Peters (1999); o modelo SVM por Guo *et al.* (2005) e Drake *et al.* (2006); os modelos GLM por Guisan e Zimmermann (2000) e Guisan *et al.* (2002) e as RNA por Hilbert e van den Muyzenberg (1999), Hilbert e Ostendorf (2001) e Linderman *et al.* (2004).

Considerando que, na modelação da ocorrência potencial das séries de vegetação, não foram utilizadas verdadeiras ausências mas sim pseudo-ausências, foi adoptada a designação de mapas de localização adequada (*habitat suitability map*) (Hirzel *et al.*, 2002; Phillips *et al.*, 2009).

De forma a testar o efeito de selecção das variáveis predictivas no processo de modelação das séries de vegetação e a garantir modelos parcimoniosos, foram testados dois cenários de modelação: 1) com o total de variáveis predictivas e de acordo com a estratégia de modelação aplicada; 2) com as variáveis significativas resultantes da análise de correlação de Pearson.

Quadro II – Técnicas de modelação.  
*Table II – Modeling techniques.*

Técnicas (algoritmos)	Tipo de modelo	Software	Parametrização	Referências
Maxent	Presença	<i>Maxent software for species habitat modelling</i>	<i>Random test percentage: 0; Regularization multiplier: 1; Max number of background points: 10000; Replicates: 1; Replicated run type: Crossvalidate; Maximum iterations: 500; Convergence threshold: 0,00001; Adjust sample radius: 0</i>	Phillips <i>et al.</i> (2006).
GARP	Presença	<i>OpenModeller</i>	<i>Max generations: 400; Convergence limit: 0,01; Population size: 50; Resamples: 2500</i>	Stockwell (1999).
SVM	Presença	<i>OpenModeller</i>	<i>SVM type: C-SVC; Kernel type: radial basis function: <math>\exp(-\gamma u-v ^2)</math>; Degree: 3; Gamma: 0; Coeff0: 0; Cost: 1; Nu: 0,5; Probabilistic output: 1; Number of pseudo-absences: 0</i>	Cristianini e Shawe-Taylor (2000).
GLM	Presença/ Ausência	<i>R [20], library: glm</i>	<i>Model: gaussian, link: identity; Number of Fisher Scoring iterations: 2</i>	Venables e Ripley (2002).
RNA	Presença/ Ausência	<i>Tiberius Data Mining</i>	<i>Inputs:8; Hidden neurons: 8; Linear neurons: 1; Train: 10 hours/net</i>	Lek e Guégan (2000).

#### 4. Validação dos Modelos

A avaliação dos SDM baseia-se na sua performance predictiva, e é considerada a medida mais importante no processo de validação, durante o qual é frequente dividir os dados (Guisan e Zimmerman, 2000), com uma parte utilizada para calibrar o modelo, designados dados de treino, e outra para validar as predições, designados dados de validação ou teste (Smith, 1994; Miller e Franklin, 2002). Na modelação predictiva da VNP aplicou-se um rácio de dados de treino e teste de 70:30 (Franklin, 2002; Phillips *et al.*, 2006; Gutierrez *et al.*, 2011). Em termos de medidas estatísticas de avaliação da capacidade predictiva, a análise da curva ROC (*Receiver Operating Characteristic*) constituiu um importante método na medição e especificação de problemas no desempenho dos modelos. Permitiu avaliar a sensibilidade da superfície criada pelo modelo, ou seja, uma representação gráfica dos pares sensibilidade (taxa dos verdadeiros positivos) e especificidade (taxa dos falsos positivos). A área abaixo da curva (AUC), delimitada pela união de todos os pares sensibilidade/especificidade, é encarada como uma medição da performance do modelo, independentemente do *threshold* (linha de corte das respostas de adequabilidade devolvidas pelo modelo) seleccionado para representar o modelo final da VNP.

Para além da avaliação individual da AUC dos SDM individuais de cada série de vegetação, procedeu-se à aferição da fracção correctamente classificada da cartografia predictiva da VNP (decorrente dos dois cenários de modelação e com posterior aplicação de valores de *threshold* do tipo *Maximum training sensitivity plus specificity*) com a cartografia oficial da VNP da CM Loures (escala 1:25 000), sendo apresentada sob a forma de tabela de contingência. Foi igualmente determinada a qualidade do mapa como um todo sem ter em consideração a qualidade das classes individuais (Precisão Global: PG).

### III. RESULTADOS

Na ordenação CCA das séries de vegetação e das variáveis predictivas, observou-se que o eixo 1, com valor próprio 0,327, apresenta um peso relativo superior ao eixo 2, cujo valor próprio é 0,311 (fig. 3). Nesta análise, destaca-se um gradiente de Altitude e Humidade, em que os extremos correspondem a séries de vegetação com as seguintes condições: baixa Altitude, Índice Topográfico de Humidade elevado e menores valores de *Sky View Factor* (Geosérie Ripícola), no quadrante I; e Altitude elevada, declives elevados e Índice Topográfico de Humidade baixo (*Arisaro simorrhini-Quercetum broteroi*), e ainda a elevados valores de *Sky View Factor* (*Viburno tini-Oleatum sylvestris*), no quadrante III.

Relativamente ao Solo, Geologia e Distância ao Tejo, do quadrante II para o IV, verifica-se a sua enorme influência na ocorrência e composição particular das séries de vegetação, em que num dos extremos destes gradientes ecológicos (menor

Distância ao Tejo, ocorrência de aluviões e solos com elevada salinidade) pode observar-se a presença do *Geopermasigmetum* Halófitico. No extremo oposto deste gradiente observa-se um aumento da distância ao Tejo (áreas mais interiores e de maior altitude do concelho de Loures), presença de arenitos e solos basálticos (*Asparago aphylli-Quercetum suberis*) e ainda de afloramentos rochosos com calcários (*Geopermasigmetum Rupícola*).

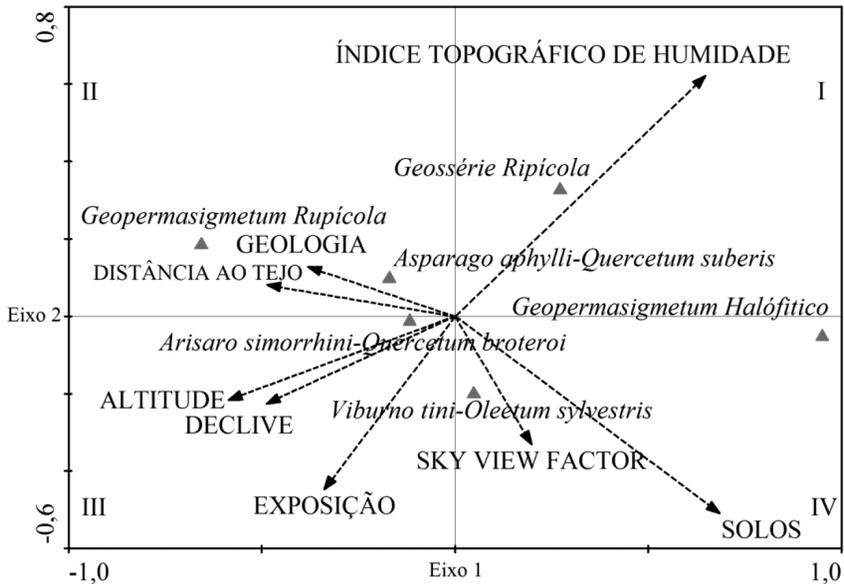


Fig. 3 – Análise Canónica de Correspondências (CCA). Eixo 1 (Valor Próprio): 0,327. Eixo 2 (Valor Próprio): 0,311.

Fig. 3 – Canonical Correspondence Analysis (CCA). Axis 1 (Eigenvalues): 0,327. Axis 2 (Eigenvalues): 0,311.

Deste modo, numa primeira fase, tendo em consideração a inexistência de Multicolinearidade ( $VIF < 10$ ), auto-correlação ( $R < 0,9$ ) e inter-relação com as séries de vegetação, foram utilizadas as 8 variáveis predictivas no primeiro cenário de modelação da VNP.

Numa segunda fase, resultante da análise de correlação foram seleccionadas as variáveis com maior significância e correlação com as séries de vegetação, que foram incorporadas no segundo cenário de modelação da VNP (quadro III).

Todos os modelos foram comparados segundo o valor de AUC, quer entre as várias séries de vegetação, quer entre os modelos desenvolvidos pelas diferentes técnicas. Observa-se, no quadro IV, que para todas as séries de vegetação os modelos Maxent, GLM e SVM apresentam melhor desempenho predictivo face aos restantes (valores médios de  $AUC > 0.85$ ). Deste modo, as omissões sobre as amostras de

validação ajustam-se bem à taxa de omissão prevista, o que significa um bom desempenho dos modelos e que a AUC gerada foi maior do que o valor aleatório (0,5), o que também é revelador de um bom desempenho dos modelos.

Quadro III – Correlação das variáveis predictivas significativas com as séries de vegetação.  
Table III – Correlation between significant predictive variables and vegetation series.

	Altitude	Declive	Exposição	I.T.H. <sup>†</sup>	SVF <sup>++</sup>	Solos	Geologia	Distância ao Tejo
Geossérie Ripícola	-,144**	-,153**	-,154**	,299**	-,080*	,020	-,133**	,001
<i>Asparago aphylli-Quercetum suberis</i>	,045	-,019	-,031	-0,38	,001	-,129**	-,141**	,097**
<i>Arisaro simorrhini-Quercetum broteroi</i>	,055	,145**	,064	-,157**	-0,54	-,188**	,092*	-,046
<i>Geopermasigmetum Rupícola</i>	,178**	,144**	,058	-,090*	-,080*	-,193**	,142**	,122**
<i>Geopermasigmetum Halófitico</i>	-,298**	-,221**	-,184**	,305**	,130**	,328**	-,090*	-,320**
<i>Viburno tini-Oleetum sylvestris</i>	,084*	,025	,137**	-,129**	,084*	,214**	-,132**	,072*

\* Nível de Significância de 0,05.

\*\* Nível de Significância de 0,01.

<sup>†</sup>Índice Topográfico de Humidade; <sup>++</sup> *Sky View Factor*

No presente estudo os modelos do tipo *machine learning* baseados em dados de presença (Maxent e SVM) apresentam uma performance predictiva mais elevada face aos modelos de presença-ausência. São técnicas com elevado ajustamento, devido à implementação de funções de respostas não lineares e aplicação de funções mais complexas ao nível de separação das classes das variáveis predictivas (Phillips e Dudík, 2008).

A elevada performance do modelo Maxent na modelação da VNP baseia-se nas seguintes características: a) a probabilidade de distribuição é definida matematicamente e, portanto, a formulação do modelo é relativamente transparente; b) permite considerar interações entre as variáveis predictivas; c) permite considerar transformações polinomiais dos preditores; d) permite avaliar a influência de cada variável predictiva na distribuição final da VNP; e e) apresenta um desempenho relativamente bom com amostras pequenas. No entanto apresenta as seguintes desvantagens: a) não possui procedimentos de selecção de variáveis predictivas; b) podem ocorrer fragilidades ao lidar com amostras enviesadas (Rocha, 2012).

O modelo SVM apresenta igualmente uma alta performance predictiva como resultado de conferir suporte às tarefas de regressão e classificação e conseguir lidar com múltiplas variáveis contínuas e categóricas. Segundo Rocha (2012), a SVM é “um classificador binário que procura um hiperplano óptimo

como uma função de decisão num espaço de dimensões maiores”. Acresce ainda a vantagem da possibilidade de definição das designadas variáveis soltas (ou variáveis de folga) no processo de treino, em que os padrões de treino não separáveis linearmente no espaço de características podem ser tratados. Por último, a SVM recorre a um algoritmo de treino iterativo, que é utilizado para minimizar uma função de erro.

Apesar de ser reconhecida melhor eficiência das técnicas de modelação baseadas em dados de presença-ausência (Brotons *et al.*, 2004 ; Franklin, 2009; Palialexis *et al.*, 2011), no presente estudo apenas o modelo GLM corrobora esta tendência. O GLM apresenta as seguintes vantagens: a) considera interações entre variáveis e relações não lineares entre as variáveis predictivas e de resposta; b) permite analisar a influência que cada preditor tem no padrão de distribuição da VNP; c) o algoritmo apresenta numerosas derivações para a resolução de problemas tais como a auto-correlação espacial, preditores correlacionados e amostragem de enviesamento amostral e d) as medidas de incerteza na estimativa dos coeficientes de regressão e análise residual permitem investigar as fontes e a distribuição espacial dos erros do modelo. No entanto, apresenta algumas desvantagens: a) os procedimentos de selecção de variáveis (Critério de Informação de Akaike - AIC) não seleccionam com a frequência desejável o modelo com a exactidão potencial mais alta. b) permite determinar a importância relativa de cada variável na capacidade predictiva do modelo da VPN; e c) a distribuição predita é extremamente sensível à relação entre a presença e a ausência de ocorrências nos dados (Rocha, 2012).

Deste modo, a boa performance predictiva do GLM na modelação da VNP advém do facto de se tratar de um modelo de regressão linear com elevada capacidade predictiva, em virtude da sua flexibilidade para todos os tipos de dados e das funções de ligação que utiliza para relacionar as variáveis independentes (tanto contínuas como categóricas). E ainda pelo facto de se terem utilizado equilibradamente as quantidades de ausências e de presença para simular a prevalência estimada das séries de vegetação na área de estudo.

Relativamente às RNA, apesar de ser descrita como uma técnica de modelação adequada para modelos ecológicos complexos não lineares, ficou aquém dos resultados esperados. As RNA apresentam a vantagem de não requererem uma compreensão a priori das relações entre as variáveis predictivas. Contudo, elas são sensíveis à composição do conjunto de dados e à formação inicial dos parâmetros da rede (Özesmi *et al.*, 2006). Existem numerosos tipos de algoritmos de aprendizagem específicos para determinados modelos de redes neuronais, que diferem entre si principalmente pelo modo como os pesos atribuídos à rede são modificados. Deste modo, a pior performance predictiva na modelação da VNP estará relacionada com os padrões de treino colocados na entrada da rede neuronal, que conduziu a um problema de aprendizagem da rede. Trata-se portanto de um modelo que requer a reiniciação do processo, compreensão das dificuldades (caixa-negra) e correcção das falhas.

Por último, o modelo GARP obteve igualmente uma performance inferior ao potencial expectável para a modelação da VNP. Esta técnica é descrita como um algoritmo de aprendizagem máquina para desenvolver um conjunto de regras. O GARP procura iterativamente correlações não-aleatórias entre as presenças e as ausências (pseudo-ausências geradas pelo algoritmo) e os factores predictivos. No entanto apresenta diversas desvantagens: a) gera pseudo-ausências e não permite que sejam substituídas; b) os mapas predictivos não são determinísticos. As saídas vão ser diferentes em diferentes execuções do GARP, mesmo que sejam utilizados os mesmos dados de ocorrência; c) é difícil de interpretar (caixa-negra); d) há tendência para erros de comissão; e e) não possui procedimentos de selecção de variáveis.

Quadro IV – Comparação dos modelos predictivos da VNP (valores de AUC).  
Table IV – Comparison of several predictive models of PNV (AUC values).

	Maxent		GARP		SVM		GLM		RNA	
	TV*	VS**	TV	VS	TV	VS	TV	VS	TV	VS
Geossérie Ripícola	0,98	0,96	0,66	0,71	0,86	0,82	0,87	0,87	0,56	0,56
<i>Asparago aphylli-Quercetum suberis</i>	0,97	0,96	0,60	0,64	0,86	0,87	0,83	0,81	0,68	0,47
Arisaro simorrhini-Quercetum broteroi	0,86	0,82	0,60	0,63	0,71	0,66	0,83	0,78	0,47	0,36
<i>Geopermasigmatum</i> Rupícola	1,00	1,00	0,86	0,93	0,96	0,96	0,94	0,94	1,00	1,00
<i>Geopermasigmatum</i> Halófitico	0,99	0,99	0,95	0,96	0,97	0,97	0,99	0,99	0,98	0,97
Viburno tini-Oleetum sylvestris	0,94	0,94	0,70	0,73	0,81	0,81	0,88	0,87	0,54	0,59
Média	0,96	0,95	0,73	0,77	0,86	0,85	0,89	0,88	0,70	0,66

\* Total de variáveis.

\*\* Variáveis significativas.

Foi ainda analisada a importância das variáveis predictivas no modelo com melhor capacidade predictiva (AUC=0,96) (quadro V). Foi possível analisar a contribuição individual de cada variável predictiva para o modelo Maxent, tal como o desempenho do modelo quando cada uma das variáveis não é incluída na análise (*stepwise with backward elimination*).

Verifica-se, no quadro V, que em todas as séries de vegetação as variáveis Solos e Geologia apresentam um elevado contributo para a sua distribuição no território estudado. Convém ainda destacar a importância das variáveis Altitude, *Sky View Factor*, Índice Topográfico de Humidade e Distância ao Tejo na diferenciação da Geossérie Ripícola, cercal de *Arisaro simorrhini-Quercetum broteroi* e comunidades halófiticas permanentes.

Os modelos desenvolvidos por cada uma das cinco técnicas de modelação foram regionalizados, utilizando a cartografia temática de base de cada variável predictiva, de forma a espacializar a distribuição potencial das séries de vegetação. Através de processos de análise espacial em ambiente SIG, os mapas de localização adequada de cada série de vegetação foram reclassificados com base nos *threshold*. Os mapas binários produzidos a partir do modelo Maxent com inclusão de todas as variáveis predictivas (adequabilidade de habitat *versus* não adequabilidade de habitat) foram integrados por combinação lógica num único output cartográfico, resultando deste modo, o mapa da VNP do concelho de Loures (fig. 4). Finalmente foi comparada a performance da cartografia predictiva produzida com cartografia oficial da VNP da CM Loures. O mapa da VNP resultante do primeiro cenário de modelação apresenta uma PG de 88%, enquanto o modelo da VNP resultante do segundo cenário de modelação apresenta uma PG de 78%.

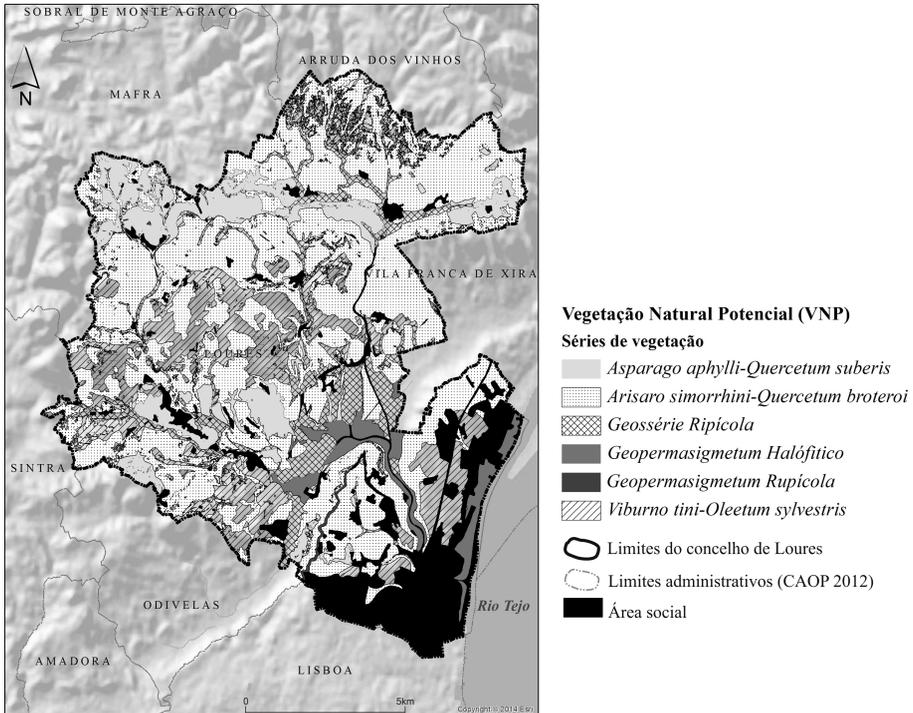


Fig. 4 – Carta da Vegetação Natural Potencial do concelho de Loures (versão a cores online).  
 Fig. 4 – Potential Natural Vegetation map of the Loures Council (coloured figure online).

Quadro V – Contribuição das variáveis predictivas no modelo Maxent.  
 Table V – Contribution of predictive variables used in the Maxent model.

	Geossérie Ripícola		<i>Asparago aphylli- Quercetum suberis</i>		<i>Arisaro simorhini- Quercetum broteroi</i>		<i>Geopermasigmatum Rupicola</i>		<i>Geopermasigmatum Halófitico</i>		<i>Viburno tini- Oleetum sylvestris</i>	
	% Contri.*	IP**	% Contri.	IP	% Contri.	IP	% Contri.	IP	% Contri.	IP	% Contri.	IP
Altitude	3,6	8,8	0,2	0,4	10,3	8,7	0,6	0	10,5	83,1	3,6	4,3
Declive	0,2	0,2	0,5	0,5	1,2	2	0	0	2,8	8	1,4	1,8
Exposição	0,8	4,3	1,5	2,1	1,4	4	0,7	0,1	0,2	0,3	0,9	1
Índice Topográfico de Humidade	0,2	0,2	1,4	2,8	1,9	3,4	0,7	0	0,2	0	0,5	0,7
<i>Sky View Factor</i>	6,9	1,7	2,3	0,7	0,5	0,7	0	0	0,2	0,1	2,6	1
Solos	<b>19,9</b>	<b>12,2</b>	<b>81,2</b>	<b>76</b>	<b>67</b>	<b>57,7</b>	<b>95,2</b>	<b>99,7</b>	<b>46,9</b>	<b>8,4</b>	<b>56,3</b>	<b>63,9</b>
Geologia	<b>62,4</b>	<b>69,9</b>	<b>6,6</b>	<b>4</b>	<b>11,7</b>	<b>7,5</b>	<b>2,8</b>	<b>0,2</b>	<b>39,1</b>	<b>0,1</b>	<b>29,9</b>	<b>11,4</b>
Distância ao Tejo	6	2,7	6,4	13,5	6	15,9	0	0	0,1	0	4,8	15,9

\*% Contribuição

\*\* Importância de permutação

#### IV. DISCUSSÃO

Os modelos da VNP produzidos possibilitaram uma melhor compreensão dos principais factores biofísicos associados à distribuição das séries de vegetação no território estudado. O modelo Maxent, além de apresentar uma elevada performance predictiva, possibilita uma análise integrada da adequabilidade ao longo de gradientes ambientais. Observam-se alguns aspectos interessantes relacionados com as curvas de resposta, nomeadamente a forma como cada variável *per se* afecta o modelo. Contudo, a interpretação do contributo das variáveis neste modelo pode apresentar algumas limitações, em virtude do modo automático como são relacionadas as diferentes classes (linear, quadrática, polinomial). A influência de colinearidade e interacção entre as variáveis predictivas também não é explicitada de forma clara, podendo originar dificuldades na interpretação. Deste modo, a aplicação de análise exploratória para a remoção de multicolinearidade permite uma melhor interpretação do contributo de cada variável predictiva. No entanto, os efeitos de interacção podem permanecer e serem impossíveis de diferenciar.

Os resultados dos modelos predictivos da VNP, a partir das curvas de resposta do modelo Maxent com a variante TV, estão de acordo com o conhecimento actual da ecologia das séries de vegetação estudadas (Costa *et al.*, 2012).

Para a Geossérie Ripícola a Geologia contribuiu em cerca de 62%, destacando-se a elevada importância da classe *Aluviões, Aterros*. A análise relativa às variáveis explicativas confirma que a adequação óptima desta série reflecte-se nas seguintes condições ecológicas: altitudes compreendidas entre 0 e 25 m (fig. 5), declive abaixo dos 10%, exposições Norte e Oeste, solos mal drenados, vales encaixados, presença de aluviossilos e com distribuição independente da continentalidade.

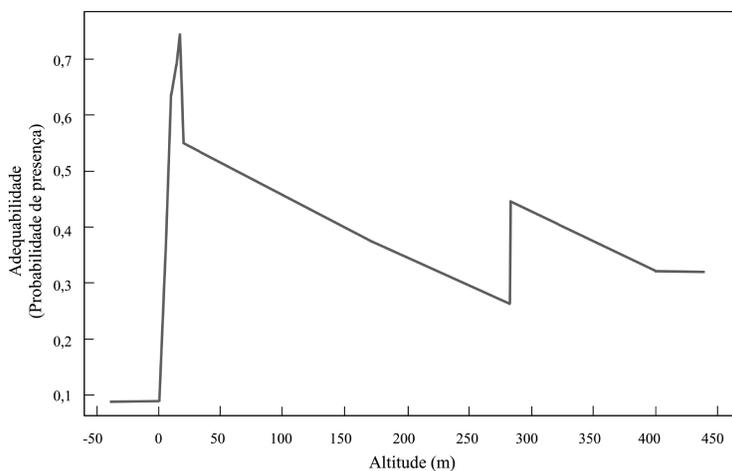


Fig. 5 – Curva de Resposta relativa à presença da Geossérie Ripícola para a variável Altitude.

Fig. 5 – Response curve to Altitude of the Riparian Geoserie.

O *Geopermasigmetum* Rupícola obteve uma melhor resposta no tipo de solo *Afloramentos rochosos com calcários ou dolomias*, ocorre maioritariamente sob a *Formação de Bica: calcários com rudistas (inclui o nível com *Neolobites vibrayanus*)* (fig. 6, classe geológica 17). Situa-se entre 400 e 450 m, em declives compreendidos entre 50 e 60%, na exposição Norte, em vales encaixados, na presença de solos bem drenados, em posições mais interiores do território e com uma maior xericidade (de maior continentalidade, a 12 500 m do Tejo), o que estabelece um elevada concordância com as características ecológicas desta permasérie.

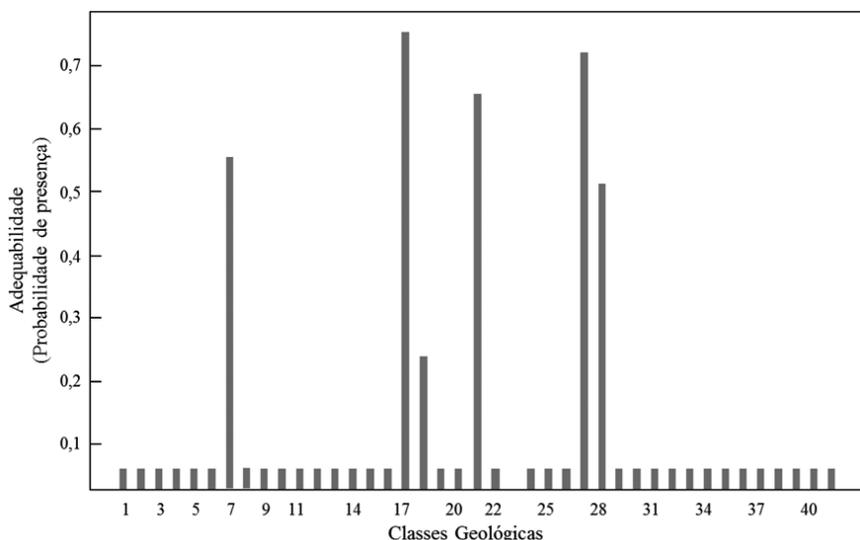


Fig. 6 – Curva de Resposta relativa à presença da *Geopermasigmetum* Rupícola para a variável Geologia.

Fig. 6 – Response curve to Geology of the rupicolous *Geopermasigmetum*.

Todavia, verificam-se efeitos marginais nas variáveis Altitude, Declive, *Sky View Factor* e Distância ao Tejo, pois quando uma destas variáveis é permutada e as outras todas assumem o valor médio, apresentam-se constantes ao longo da Curva de Resposta. Este facto poderá ser explicado pela elevada correlação entre as variáveis, que apresentam uma menor contribuição para o modelo. Conforme será discutido, uma selecção iterativa das variáveis ou normalização das variáveis contínuas poderia constituir uma estratégia alternativa para a selecção do conjunto mais apropriado de variáveis.

Em relação à série de vegetação do sobreiral *Asparago aphylli-Quercetum suberis* verifica-se mais uma vez que o tipo de solo explica 81% do modelo, nomeadamente o solo que corresponde à classe *Solos Litólicos Não Húmicos Pouco Insaturados, Normais de arenitos finos micáceos (de textura arenosa a franco-arenosa), fase normal* (com uma adequabilidade de 91%). Segundo

Cardoso (1965) trata-se de solos em que a rocha-mãe, sujeita a intensa meteorização física e a uma reduzida alteração química e que apresentam um baixo teor orgânico e espessura efectiva. Assim, a sua textura tende a ser predominantemente arenosa, solta e friável no horizonte A, chegando a franco-arenosa no horizonte B. Estas características evidenciam a diagnose desta série, cuja sinecologia se encontra em termos pedológicos associada a arenitos e solos basálticos (Costa *et al.*, 2012). O modelo prediz também que a presença da *Formação de Porto da Calada: arenitos, pelitos, calcários e dolomitos*, declives abaixo dos 10%, distribuição entre 200 e 225 m de altitude, na exposição Norte (fig. 7), em vales abertos e igualmente em situações de maior continentalidade (a uma distância de 15 000 m do Tejo) são condições ecológicas favoráveis para a ocorrência desta série de vegetação.

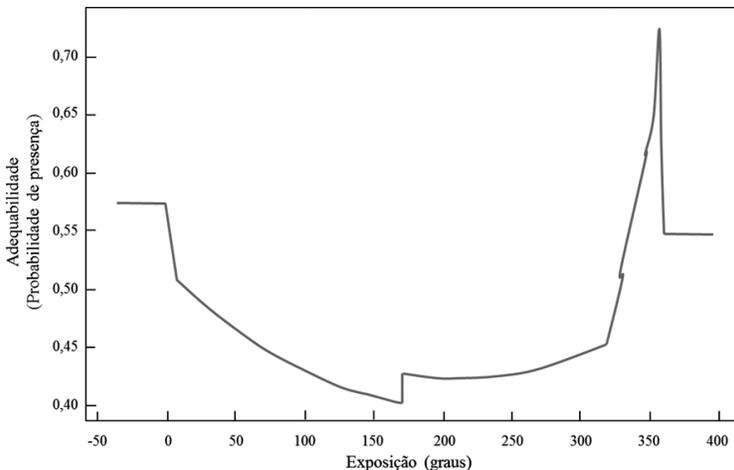


Fig. 7 – Curva de Resposta relativa à presença da série de vegetação *Asparagus aphylli-Quercetum suberis* para a variável Exposição.

Fig. 7 – Response curve to Aspect of the vegetation series *Asparagus aphylli-Quercetum suberis*.

No que concerne à série de vegetação do cercal de *Arisaro simorrhini-Quercetum broteroi* verifica-se mais uma vez que os solos explicam 67% do modelo, sendo várias as classes que melhor responderam com uma adequabilidade superior a 70% de presença desta série. A classe de solos que apresenta uma maior adequação (cerca de 87%) é a referente a *Solos Incipientes Coluviosolos ou Solos de Baixas Calcários (Para-Solos Calcários) de textura ligeira, fase normal*. Os restantes solos correspondem de uma forma geral a solos mediterrâneos de origem calcária, sendo a resposta consonante com a diagnose da série (Vila-Viçosa, 2012). O modelo também identificou áreas de óptima adequabilidade situadas na *Formação de Benfica: intercalações calcárias (Calcários de Alfornelos)*, nas altitude de 45 m e também entre

350 e 450 m, em zonas com declives entre 50 e 60%, em vales encaixados (fig. 8), nas exposições Sul e Sudoeste e na presença de solos bem drenados.

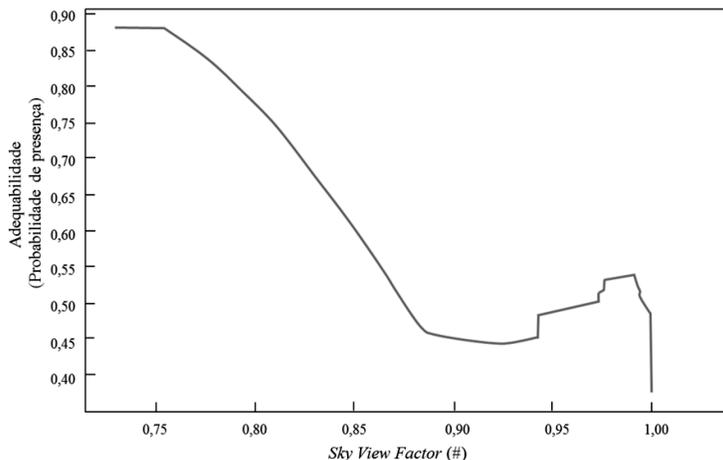


Fig. 8 – Curva de Resposta relativa à presença da série de vegetação *Arisaro simorrhini-Quercetum broteroi* para a variável *Sky View Factor*.  
 Fig. 8 – Response curve to *Sky View Factor* of the vegetation series *Arisaro simorrhini-Quercetum broteroi*.

A série de vegetação do zambujal *Viburno tini-Oleetum sylvestris* também apresentou melhor resposta com a variável tipo de Solo (56%) seguida da Geologia (30%). Os tipos de solos com maiores probabilidades de correspondência foram os *Barros Pardos Calcários Pouco Descarbonatados, de arenitos argilosos, argilas ou argilitos, calcários* e *Solos Argiluvitados Pouco Insaturados Mediterrâneos Pardos de Materiais Não Calcários Para-Barros de arenitos finos, argilas ou argilitos, fase normal*. Estes solos caracterizam-se por apresentarem nos horizontes A e B uma textura argilosa a franco-argilo-arenosa extremamente aderente, plástica, sendo solos firmes e rijos a extremamente rijos, sujeitos a fendilhamento quando secam, com valores de pH elevados (6.5 a 8.5), o que está de acordo com a diagnose desta série (Costa *et al.*, 2012). A série apresenta igualmente uma elevada adequabilidade sob as formações geológicas do *Complexo Vulcânico de Lisboa e Formação de Benfica: intercalações calcárias (Calcários de Alfovelos)*, a 300 m de altitude, em declives entre os 5 e 10%, em vales abertos, nas exposições Sudoeste e Oeste, na presença de solos bem drenados e em zonas de maior continentalidade (pontualmente aos 1 000 m, mas atingindo os valores mais elevados a uma distância compreendida entre os 10 000 e os 15 000 m do Tejo) (fig. 9).

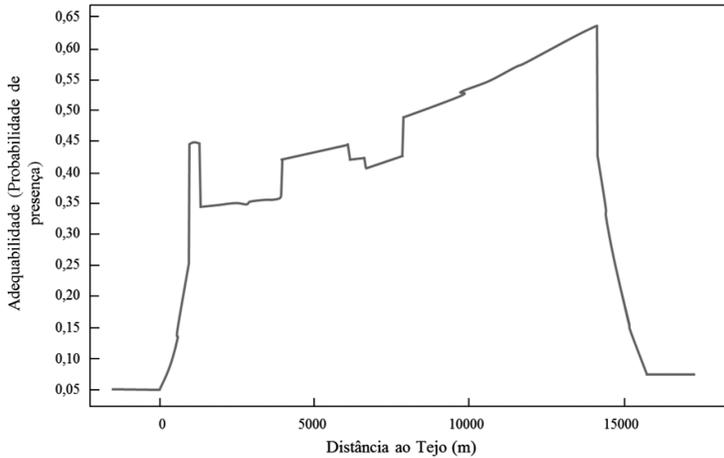


Fig. 9 – Curva de Resposta relativa à presença da série de vegetação *Viburno tini-Oleetum sylvestris* para a variável Distância ao Tejo.

Fig. 9 – Response curve to Distance to 'Tejo' of the vegetation series *Viburno tini Oleetum sylvestris*.

Finalmente, no que concerne ao *Geopermasigmetum halófitico* mais uma vez a variável Solos apresenta as melhores respostas (47%), sendo de destacar as elevadas probabilidades de correspondência com as classes *Solos Halomórficos Salinos de Salinidade Elevada de Aluviões de textura pesada, calcários, fase inundável*, *Solos Halomórficos Salinos de Salinidade Elevada de Aluviões de textura pesada, calcários, fase normal* e *Solos Halomórficos Salinos de Salinidade Moderada de Aluviões de textura ligeira, calcários*. As restantes curvas de resposta desta permasérie revelam padrões extremamente interessantes, expressos pela forma como cada variável afecta *per se* o modelo. Deste modo, é possível observar que, com o aumento de Altitude (fig. 10), a adequabilidade da permasérie tende a diminuir, a par da Distância ao Tejo (fig. 11), o que se reflecte numa maior proximidade ao mar, menor continentalidade, maior teor de humidade e influência da salsugem. Deste modo, a elevada adequabilidade desta permasérie é determinada pela presença de *Aluviões*, *Aterros*, solos com má drenagem, por baixas altitudes (entre 0 e 3 m), declives abaixo dos 4%, proximidade ao Tejo (aprox. 200 a 300 m), áreas planas (sem exposição definida) e de relevo suave (*Sky View Factor* próximo de 1).

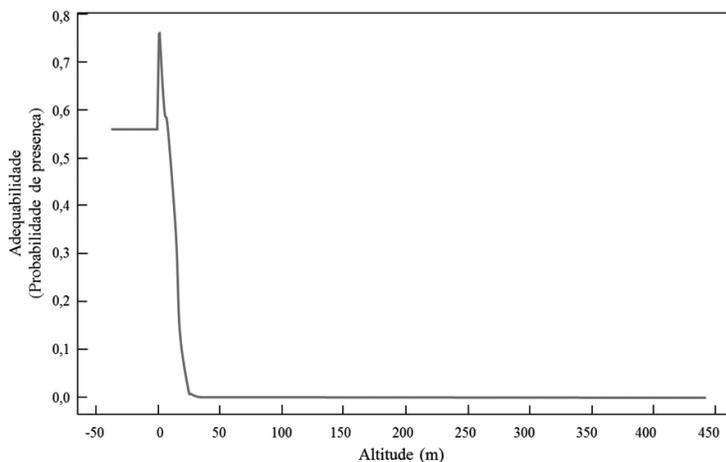


Fig. 10 – Curva de Resposta relativa à presença de *Geopermasigetum* halofítico para a variável Altitude.

Fig. 10 – Response curve to Altitude of the halophytic *Geopermasigetum*.

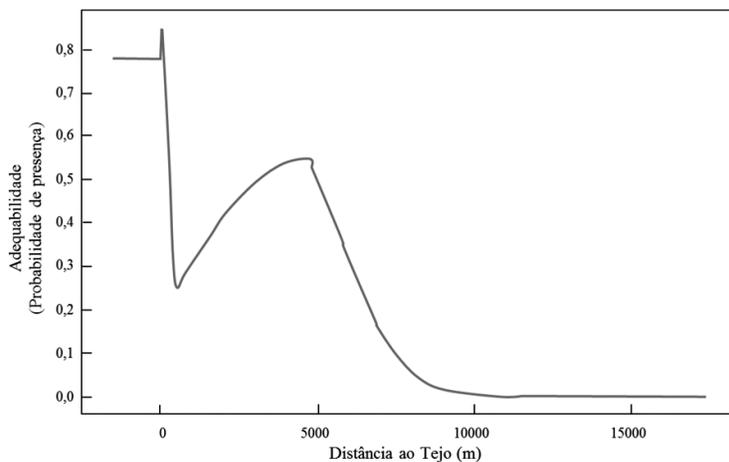


Fig. 11 – Curva de Resposta relativa à presença de *Geopermasigetum* halofítico para a variável Distância ao Tejo.

Fig. 11 – Response curve to Distance to 'Tejo' of the halophytic *Geopermasigetum*.

## V. CONCLUSÃO

A estratégia de modelação predictiva ao nível da comunidade (*classification-then-modelling*) permitiu a quantificação e avaliação da distribuição espacial da VNP no concelho de Loures. A classificação canónica na modelação da VNP possi-

bilitou a integração da classificação das séries de vegetação e modelação espacial num único nível de análise. De referir que, nos SDM ao nível da espécie a única preocupação centra-se na selecção da técnica de modelação a implementar (e.g. Maxent, GLM). Ao invés, na modelação ao nível da comunidade esta selecção apresenta-se como uma componente da estratégia analítica global, permitindo um melhor conhecimento da posição exacta das séries de vegetação ao longo de gradientes ambientais, sendo expressa na sua menor ou maior adequabilidade territorial.

No entanto, ficou evidente a necessidade de uma selecção de preditores adequados, sobretudo variáveis que tenham em consideração os factores fisiológicos directos ou efeitos históricos (uso e ocupação do solo) que actuam a escalas espaciais mais detalhadas, cuja ausência poderá conduzir a uma sobrestimação da amplitude espacial da distribuição potencial das séries de vegetação. No entanto, os resultados obtidos com os modelos predictivos e um bom conhecimento ecológico das séries de vegetação permitem a utilização eficaz de modelos de adequabilidade para descrever e estudar a distribuição das séries e as suas respostas a determinados factores geográficos e ambientais.

Actualmente têm-se vindo a desenvolver esforços no sentido de encontrar a melhor estratégia em diferentes contextos. Tal facto depende dos objectivos analíticos e da natureza do conjunto de dados disponíveis. Na estratégia implementada foi considerado que o conceito de Nicho Ecológico de Hutchinson deve ser o mais explícito possível, ao nível da selecção e interacção entre as variáveis ambientais predictivas, funções de resposta, tipo de modelo, tipo de dados bióticos e interpretação das predições resultantes. Deste modo, o conceito de nicho realizado aponta similaridades com os pressupostos subjacentes ao conceito de VNP. Assim sendo, para além da elevada performance predictiva do modelo Maxent, os resultados obtidos por este modelo permitiram aferir o modo como os gradientes ecológicos, expressos ao longo das curvas de resposta, determinam o nicho realizado para as diversas séries de vegetação.

Conforme ficou demonstrado neste estudo, mesmo considerando a existência de um vasto conhecimento empírico, centrado nos factores biofísicos determinantes para a modelação predictiva da VNP, e a selecção prévia das variáveis com maior significância e correlação com as séries de vegetação, dever-se-á optar por uma estratégia de modelação que permita a inclusão do total das variáveis predictivas. No entanto, previamente devem ser analisadas as assumpções dos modelos, ao nível da eliminação de situações de multicolinearidade, avaliação da normalidade e linearidade das variáveis predictivas e as possíveis melhorias no desempenho predictivo dos modelos se for tida em consideração a standardização das variáveis independentes.

Por outro lado, os resultados obtidos demonstram que, mesmo que as variáveis predictivas apresentem um baixo nível de significância, contribuem para uma melhoria da performance predictiva. Deste modo, na espacialização da VNP ficou demonstrado que o cenário com todas as variáveis permite obter melhores resultados.

Dever-se-á ainda delinear a amostragem que assegure maior representatividade da variabilidade espacial das espécies/comunidades a modelar, incorpore número de variáveis proximais e integre dados de base das variáveis categóricas (e.g. edafologia) com grande desagregação territorial.

Apesar do carácter hipotético da cartografia da VNP, a distribuição potencial das séries de vegetação serve de base para uma comparação quantitativa com a distribuição da vegetação actual, e revela-se fundamental para a avaliação dos efeitos de perturbação, sucessão vegetal e diversidade.

Conclui-se que a expressão cartográfica da VNP no território assume um papel importante no Ordenamento e Planeamento do Território, nomeadamente, ao nível das políticas de protecção dos Solos e da Biodiversidade e nos Planos de Ordenamento do Território em vigor. A sua aplicação na reconstituição da vegetação natural após acção antrópica apresenta igualmente enorme importância.

## AGRADECIMENTOS

Este trabalho foi desenvolvido no âmbito da Bolsa de Doutoramento financiada pela Fundação para a Ciência e Tecnologia (FCT) com a referência SFRH / BD / 45147 / 2008. Os autores agradecem à Câmara Municipal de Loures a cedência da geo-informação, que se revelou fundamental para a elaboração deste trabalho e estão gratos pelos comentários dos *referees* e dos editores da Finisterra.

## BIBLIOGRAFIA

- Alcoforado M J (1987) Brisas estivais do Tejo e do Oceano na região de Lisboa. *Finisterra-Revista Portuguesa de Geografia*, XXII (43): 71-112.
- Alcoforado M J, Andrade H, Lopes A, Vasconcelos J (2009) Application of climatic guidelines to urban planning. The example of Lisbon (Portugal). *Landscape and Urban Planning*, 90 (1-2): 56-65.
- Araújo M B, Pearson R G, Thuiller W, Erhard M (2005a) Validation of species-climate impact models under climate change. *Global Change Biology*, 11: 1504-1513.
- Araújo M B, Thuiller W, Williams P H, Reginster I (2005b) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography*, 14: 17-30.
- Austin M P, Belbin L, Meyers J A, Doherty M D, Luoto M (2006) Evaluation of statistical models used for predicting plant species distributions: role of artificial data and theory. *Ecological Modelling*, 199: 197-216.
- Barry S, Elith J (2006) Error and uncertainty in habitat models. *Journal of Applied Ecology*, 43: 413-423.
- Biondi E, Casavecchia S, Pesaresi S (2011) Phytosociological synrelevés and plant landscape mapping: from theory to practice. *Plant Biosystems*, 145 (2): 261-273.
- Biondi E, Feoli E, Zuccarello V (2004) Modelling Environmental Responses of Plant Associations: A Review of Some Critical Concepts in Vegetation Study. *Critical Reviews in Plant Sciences*, 23 (2): 149-156.
- Blasi C, Capotorti G, Fronzoni R (2005) Defining and mapping typological models at the landscape scale. *Plant Biosystems*, 139: 155-163.
- Blasi C, Filibeck G, Fronzoni R, Rosati L, Smiraglia D (2004) The map of the vegetation series of Italy. *Fitosociologia*, 41 (1 Suppl. 1): 21-25.
- Bohn U, Gollub G, Hettwer C, Neuhäuslová Z, Schlüter H, Weber H (2003) *Map of the natural vegetation of Europe 1-3*. Federal Agency for Nature Conservation, Bonn-Bad Godesberg.

- Bredenkamp G, Chytrý M, Fischer H S, Neuhäuslová Z, van der Maarel E (1998) Vegetation mapping: Theory, methods and case studies: Introduction. *Applied Vegetation Science*, 1: 162-164.
- Brotans L, Thuiller W, Araújo M B, Hirzel A H (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. *Ecography*, 27: 437-448.
- Burgman M A, Lindenmayer D B, Elith J (2005) Managing landscapes for conservation under uncertainty. *Ecology*, 86: 2007-2017.
- Bustamante J, Seoane J (2004) Predicting the distribution of four species of raptors (Aves: Accipitridae) in southern Spain: statistical models work better than existing maps. *Journal of Biogeography*, 31: 295-306.
- Capelo J, Mesquita S, Costa J C, Ribeiro S, Arsénio P, Neto C, Monteiro-Henriques T, Aguiar C, Honrado J, Espírito-Santo D, Lousã M (2007) A methodological approach to potential vegetation modeling using GIS techniques and phytosociological expert-knowledge: application to mainland Portugal. *Phytocoenologia*, 17: 399-415.
- Cardoso J C (1965) *Os Solos de Portugal – Sua Classificação. Caracterização e Gênese I – a Sul do Rio Tejo*, Direcção Geral dos Serviços Agrícolas, Secretaria de Estado da Agricultura, Lisboa.
- Costa J C, Aguiar C, Capelo J H, Lousã M, Neto C (1998) Biogeografia de Portugal Continental. *Quercetea*, 0: 5-56.
- Costa J C, Neto C, Aguiar C, Capelo J, Espírito Santo M D, Honrado J, Pinto Gomes C, Monteiro-Henriques T, Sequeira M, Lousã M (2012) Vascular Plant Communities in Portugal (Continental, the Azores and Madeira). *Global Geobotany*, 2: 1-180.
- Costa J C, Espírito-Santo, D, Arsénio P (2010) Guia geobotânico da excursão ao Parque Natural da Serras de Aire e Candeeiros, *Quercetea* 10: 5-106.
- Cristianini N, Shawe-Taylor J (2000) *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press, London.
- Cross J R (2006) The potential natural vegetation of Ireland. *Biology and Environment: Proceedings of the Royal Irish Academy*, 106B (2): 65-116.
- Cumming G S (2000) Using habitat models to map diversity: pan-African species richness of ticks (Acari: Ixodida). *Journal of Biogeography*, 27: 425-440.
- Diniz-Filho J A F, Bini L M, Hawkins B A (2003) Spatial autocorrelation and red herrings in geographical ecology. *Global Ecology & Biogeography*, 12: 53-64.
- Dirnböck T, Dullinger S, Gottfried M, Ginzler C, Grabherr G (2003) Mapping alpine vegetation based on image analysis, topographic variables and Canonical Correspondence Analysis. *Applied Vegetation Science*, 6: 85-96.
- Dormann C F, McPherson J M, Araújo M B, Bivand R, Bolliger J, Carl G, Davies R G, Hirzel A, Jetz W, Kissling D W, Kühn I, Ohlemüller R, Peres-Neto P R, Reineking B, Schröder B, Schurr F, Wilson R (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30(5): 609-628 (2007).
- Drake J M, Randin C, Guisan A (2006) Modelling ecological niches with support vector machines. *Journal of Applied Ecology*, 43: 424-432.
- Elith J, Graham C H, Anderson R P. *et al.* (2006) Novel methods improve prediction of species distributions from occurrence data. *Ecography*, 29: 129-151.
- Elith J, Leathwick J (2009) Conservation prioritisation using species distribution modelling. In Moilanen A, Wilson K A, Possingham H (eds.) *Spatial conservation prioritization: quantitative methods and computational tools*. Oxford, UK, Oxford University Press: 70-93.
- Elton C (1927) *Animal Ecology*. Chicago: University of Chicago Press.
- Ferrier S, Drielsma M, Manion G, Watson G (2002a) Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. Community-level modeling. *Biodiversity and Conservation*, 11: 2309-2338.
- Ferrier S, Watson G, Pearce J, Drielsma M (2002b) Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. *Biodiversity and Conservation*, 11: 2275-2307.
- Fleishman E, Murphy D D, Sjögren-Gulve P (2002) Modeling species richness and habitat suitability for taxa of conservation interest. In Scott J M, Heglund P J, Morrison, M L *et al.*

- (eds.) *Predicting Species Occurrences: Issues of Accuracy and Scale*. Covelo, CA: Island Press: 507-517.
- Franklin J (2009) *Mapping species distributions. Spatial inference and prediction*. Cambridge University Press, London.
- Franklin J (2002) Enhancing a regional vegetation map with predictive models of dominant plant species in chaparral. *Applied Vegetation Science*, 5: 135-146.
- Franklin J (1998) Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science*, 9: 733-748.
- Franklin J (1995) Predictive vegetation mapping: geographic modeling of biospatial patterns in relation to environmental gradients. *Progress in Physical Geography*, 19: 474-499.
- Franklin J, McCullough P, Gray C (2000) Terrain variables used for predictive mapping of vegetation communities in Southern California. In Wilson J, Gallant J (eds.) *Terrain analysis: principles and applications*. Wiley & Sons, New York: 331-353.
- Franklin J, Syphard A D, He H S, Mladenoff D J (2005) The effects of altered fire regimes on patterns of plant succession in the foothills and mountains of southern California. *Ecosystems*, 8: 885-898.
- Gibson L A, Wilson B A, Cahill D M, Hill J (2004) Modelling habitat suitability of the swamp antechinus (*Antechinus minimus maritimus*) in the coastal heathlands of southern Victoria, Australia. *Biological Conservation*, 117: 143-150.
- Guisan A, Edwards T C Jr, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, 157: 89-100.
- Guisan A, Weiss S, Weiss A (1999) GLM versus CCA spatial modeling of plant species distributions. *Plant Ecology*, 143: 107-122.
- Guisan A, Zimmermann N E (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, 135: 147-186.
- Guo Q H, Kelly M, Graham C H (2005) Support vector machines for predicting distribution of sudden oak death in California. *Ecological Modelling*, 182: 75-90.
- Gutierrez F (2014) *Estrutura e Dinâmica dos Habitats e da Paisagem dos Sítios Estuário do Sado e Comporta/Galé - Um contributo para a Gestão e Restauro Ecológico*. Dissertação de doutoramento, Instituto de Geografia e Ordenamento do Território, Universidade de Lisboa.
- Gutierrez F, Gil A, Reis E, Lobo A, Neto C, Calado H, Costa J C (2011) *Acacia saligna* (Labill.) H. Wendl in the Sesimbra Council: Invaded habitats and potential distribution modelling. *Journal of Coastal Research*, SI 64: 403-407.
- Hilbert D W, Ostendorf B (2001) The utility of artificial neural networks for modelling the distribution of vegetation in past, present and future climates. *Ecological Modelling*, 146: 311-327.
- Hilbert D W, van den Muyzenberg J (1999) Using an artificial neural network to characterise the relative suitability of environments for forest types in a complex tropical vegetation mosaic. *Diversity and Distributions*, 5: 263-274.
- Hirzel A H, Hausser J, Chessel D, Perrin N (2002) Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data?. *Ecology*, 83 (7): 2027-2036.
- Hirzel A H, Le Lay G (2008) Habitat suitability modelling and niche theory. *Journal of Applied Ecology*, 45: 1372-1381.
- Hutchinson M F, Gallant J (2000) Digital elevation models and representation of terrain shape. In Wilson J, Gallant J (eds.) *Terrain Analysis: Principles and Applications*. John Wiley & Sons, New York: 29-50.
- Jetz W, Rahbek C (2002) Geographic range size and determinants of avian species richness. *Science*, 297: 1548-1551.
- Jimenez-Valverde A, Lobo J M, Hortal J (2009) The effect of prevalence and its interaction with sample size on the reliability of species distribution models. *Community Ecol.*, 10: 196-205.
- Kent M, Ballard J (1988) Trends and problems in the application of classification and ordination methods in plant ecology. *Vegetatio*, 78: 109-124.
- Lane S N, Richards K S, Chandler J H (1998) *Landform monitoring, modelling and analysis*. John Wiley and Sons, Chichester, UK.
- Leathwick J R, Rowe D, Richardson J, Elith J, Hastie T (2005) Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology*, 50: 2034-2052.

- Legendre P, Legendre L (1998) *Numerical Ecology*. Elsevier, Amsterdam.
- Lek S, Guégan J F (2000) *Artificial Neural Networks: Application to Ecology and Evolution*. Springer-Verlag, Berlin.
- Lennon J J (2002) Red shifts and red herrings in geographical ecology. *Ecography*, 23: 101-113.
- Linderman M A, Liu J, Qi J, An L, Ouyang Z, Yang J, Tan Y (2004) Using artificial neural networks to map the spatial distribution of understory bamboo from remotely sensed data. *International Journal of Remote Sensing*, 25: 1685-1700.
- Loidi J, Fernández-González F (2012) Potential Natural Vegetation: reburying or reboring? *Journal of Vegetation Science*, 23: 596-604.
- Mackey B G (1993) Predicting the potential distribution of rain-forest structural characteristics. *Journal of Vegetation Science*, 4: 43-54.
- McCune B (2007) Improved estimates of incident radiation and heat load using non-parametric regression against topographic variables. *Journal of Vegetation Science*, 18: 751-754.
- Magalhães M R, Silva A P G, Cunha N S, Campo S L (2002) Delimitação da estrutura ecológica municipal de Loures, métodos de análise espacial para interpretação da paisagem. *Encontro de Utilizadores de Sistemas de Informação Geográfica*, Tagus-Park, Oeiras (Portugal). [Acedido em 27 de Junho de 2013]. [http://www.isa.utl.pt/ceap/pvcloures/publicacoes/esig\\_outubro\\_02.pdf](http://www.isa.utl.pt/ceap/pvcloures/publicacoes/esig_outubro_02.pdf).
- McPherson J M, Jetz W, Rogers D J (2006) Using coarse-grained occurrence data to predict species distributions at finer spatial resolution – possibilities and limitations. *Ecological Modelling*, 192: 499-522.
- Miller J, Franklin J (2002) Predictive vegetation modeling with spatial dependence – vegetation alliances in the Mojave Desert. *Ecological Modelling*, 57: 227-247.
- Miller J, Franklin J, Aspinall R (2007) Incorporating spatial dependence in predictive vegetation models. *Ecological Modelling*, 202: 225-242.
- Monteiro-Henriques T (2010) *Fitossociologia e paisagem da bacia hidrográfica do Rio Paiva*. Tese de Doutoramento, Universidade Técnica de Lisboa - Instituto Superior de Agronomia.
- Mora C (2006) *Climas da Serra da Estrela. Características regionais e particularidades locais dos planaltos e do alto vale do Zêzere*. Tese de Doutoramento, Faculdade de Letras, Universidade de Lisboa.
- Moore I D, Burch G J, Mackenzie D H. (1988) Topographic effects on the distribution of surface soil water and the location of ephemeral gullies. *Transactions of the American Society of Agricultural Engineers*, 31: 1098-1107.
- Moore I D, Grayson R, Ladson A (1991) Digital terrain modelling: a review of hydrological, geomorphological, and biological applications. *Hydrological Processes*, 5: 3-30.
- Mücher C A, Hennekens S M, Bunce R G H, Schaminee J H J, Schaepman M E (2009) Modelling the spatial distribution of Natura 2000 habitats across Europe. *Landscape and Urban Planning*, 92: 148-159.
- Mucina L (2010) Floristic-phytosociological approach, potential natural vegetation, and survival of prejudice. *Lazaroo*, 31: 173-182.
- Neto C (2002) A Flora e a Vegetação do superdistrito Sadense (Portugal). *Guineana*, 8: 1-269.
- Neto C, Costa J C, Capelo J, Gaspar N, Monteiro-Henriques T (2007) Os Sobreirais da Bacia Ceno-Antropozóica do Tejo (Província Lusitano-Andaluza Litoral), Portugal. *Acta Botanica Malacitana*, 32: 201-210.
- Neto C, Pereira E, Reis E, Costa J C, Capelo J, Henriques C (2008) Carta da Vegetação Natural Potencial de Caldas da Rainha. *Finisterra – Revista Portuguesa de Geografia*, XLIII(86): 31-56.
- Ohmann J L, Spies T A (1998) Regional gradient analysis and spatial pattern of woody plant communities of Oregon forests. *Ecological Monographs*, 68: 151-182.
- Osborne P E, Foody G M, Suarez-Seoane S (2007) Non-stationarity and local approaches to modelling the distribution of wildlife. *Diversity and Distributions*, 13: 313-323.
- Özesmi S, Tan C, Özesmi U (2006) Methodological issues in building, training, and testing artificial neural networks in ecological applications. *Ecological Modelling*, 195: 83-93.
- Palialexis A, Georgakarakos S, Karakassis I, Lika K, Valavanis V D (2011) Prediction of marine species distribution from presence– absence acoustic data: comparing the fitting efficiency and the predictive capacity of conventional and novel distribution models. *Hydrobiologia*, 670: 241-266.

- Pausas J G, Austin M P (2001) Patterns of plant species richness in relation to different environments: an appraisal. *Journal of Vegetation Science*, 12: 153-166.
- Pausas J G, Carreras J, Ferre A, Font X (2003) Coarse-scale plant species richness in relation to environmental heterogeneity. *Journal of Vegetation Science*, 14: 661-668.
- Perring F (1959) Topographical gradients of chalk grassland. *Journal of Ecology*, 47: 447-481.
- Phillips S J, Anderson R P, Schapire R E (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190: 231-259.
- Phillips S J, Dudík M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, 31: 161-175.
- Phillips S J, Dudík M, Elith J, Graham C H, Lehmann A, Leathwick J, Ferrier S (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, 19 (1): 181-197.
- Portela-Pereira E (2013) *Análise geobotânica dos bosques e galerias ripícolas da bacia hidrográfica do Tejo em Portugal*. Dissertação de doutoramento, Instituto de Geografia e Ordenamento do Território, Universidade de Lisboa.
- Ricotta C, Carranza M L, Avena G, Blasi C (2002) Are potential natural vegetation maps a meaningful alternative to neutral landscape models? *Applied Vegetation Science*, 5: 271-275.
- Ricotta C, Carranza M L, Avena G, Blasi C (2000) Quantitative comparison of the diversity of landscapes with actual vs. potential natural vegetation. *Applied Vegetation Science*, 3: 157-162.
- Rivas-Martínez S (2007) Mapa de series, geoseries y geopermaseries de vegetación de España I. *Itinera Geobotanica*, 17: 1-435.
- Rivas-Martínez S, Sanchez-Mata D, Costa M (1999) North American boreal and western temperate forest vegetation (syntaxonomical synopsis of the potential natural plant communities of North America, 11). *Itinera Geobotanica*, 12: 5-316.
- Rivas-Martínez S, Rivas Sáenz S, Penas A (2011) World bioclimatic classification systems. *Global Geobotany*, 1: 1-634.
- Rocha J (2012) *Sistemas Complexos, Modelação e Geosimulação da Evolução de Padrões de Uso e Ocupação do Solo*. Dissertação de doutoramento, Instituto de Geografia e Ordenamento do Território, Universidade de Lisboa.
- Rushton S P, Ormerod S J, Kerby G (2004) New paradigms for modelling species distributions? *Journal of Applied Ecology*, 41: 193-200.
- Ruxton G D, Colegrave N (2006) *Experimental design for the life sciences. Second Edition*. Oxford University Press, Oxford.
- Scott J M, Heglund P J, Samson F, Haufler J, Morrison M, Raphael M, Wall B (2002) *Predicting Species Occurrences: Issues of Accuracy and Scale*. Covelo, CA: Island Press.
- Segurado P, Araújo M B (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, 31: 1555-1568.
- Segurado P, Araújo M B, Kunin W E (2006) Consequences of spatial autocorrelation for niche-based models. *Journal of Applied Ecology*, 43: 433-444.
- Shriner S A, Simons T R, Farnsworth G L (2002) A GIS-based habitat model for Wood Thrush, *Hyocichla mustelina*, in Great Smokey Mountains National Park. In Scott J M, Heglund P J, Morrison M L (eds.) *Predicting Species Occurrences: Issues of Accuracy and Scale*. Covelo, CA, Island Press: 529-535.
- Smith P A (1994) Autocorrelation in logistic regression modelling of species distributions. *Global Ecology and Biogeography Letters*, 4: 47-61.
- Stockwell D R B (1999) Genetic algorithms II. In Fielding A H (ed.) *Machine learning methods for ecological applications*. Kluwer Academic Publishers, Boston: 123-144.
- Stockwell D R B, Peters D P (1999) The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13: 143-158.
- ter Braak C (1988) *Canoco – A Fortran Program for Canonical Community Ordination by [Partial] [Detrended] [Canonical] Correspondence Analysis (version 2.1.)* Agricultural Mathematics Group, Wageningen (Netherlands).
- Theobald D M (2007) Using GIS to generate spatially balanced random survey designs for

- natural resource applications. *Environmental Management*, 134-146.
- Thuiller W, Brotons L, Araújo M B, Lavorel S (2004) Effects of restricting environmental range of data to project current and future species distributions. *Ecography*, 27: 165-172.
- Tüxen R (1956) Die heutige potentielle natürliche Vegetation als Gegenstand der Vegetationskartierung. *Angew. Pflanzensoz. (Stolzenau)*, 13: 5-42.
- van Manen F T, Clark J D, Schlarbaum S E, Johnson K, Taylor G (2002) A model to predict the occurrence of surviving butternut trees in the southern Blue Ridge Mountains. In Scott J M, Heglund P J, Morrison M L, et al. (eds.) *Predicting species occurrences: issues of accuracy and scale*. Covelo, CA: Island Press: 491-497.
- Vila-Viçosa C (2012) *Os Carvalhais Marcescentes do Centro e Sul de Portugal*. Dissertação de mestrado, Universidade de Évora: 266.
- Venables W M, Ripley B D (1994) *Modern Applied Statistics with S-Plus*. Springer-Verlag, New York.
- Westhoff V, van der Maarel E (1978) *The Braun-Blanquet approach*. In Whittaker R H (ed.) Classification of plant communities. Junk, The Hague: 287-374.
- Wilson J, Gallant J (2000) *Terrain analysis: principles and applications*. John Wiley & Sons, New York.
- Wisz M S, Dendoncker N, Madsen J, Rounsevell M, Jespersen M, Kuijken E, Courtens W, Verschuur C, Cottaar F (2008) Modelling pink-footed goose (*Anser brachyrhynchus*) wintering distributions for the year 2050: potential effects of land-use change in Europe. *Diversity and Distributions*, 14: 721-731.
- Wu X B, Smeins F E (2000) Multiple-scale habitat modeling approach for rare plant conservation. *Landscape and Urban Planning*, 51: 11-28.
- Zampieri M, Lionello P (2008) A statistical model for potential natural vegetation, application to mediterranean climate. *EGU Geophysical Research Abstracts*, Vol. 10, EGU General Assembly 2008.
- Zerbe S (1998) Potential natural vegetation: validity and nature conservation. *Applied Vegetation Science*, 1: 165-172.
- Zhang L, Ma Z, Guo L (2008) Spatially assessing model errors of four regression techniques for three types of forest stands. *Forestry*, 81: 209-22.

---

<sup>i</sup> Hutchinson (1987) define nicho ecológico como “... o hipervolume definido pelas dimensões ambientais dentro do qual uma espécie pode sobreviver e reproduzir”. Hutchinson distingue ainda nicho fundamental (fisiológico ou potencial), definido como a resposta das espécies ao ambiente (recursos) na ausência de interações bióticas (competição, predação, facilitação); de nicho realizado (ecológico, actual), como sendo as dimensões ambientais onde as espécies actualmente ocorrem e podem sobreviver e reproduzir, incluindo os efeitos das interações bióticas.

<sup>ii</sup> Foi considerada a distância euclidiana à margem direita do estuário do Tejo (em metros) como factor explicativo da influência da continentalidade nos regimes de temperatura e humidade atmosférica associada ao efeito das brisas do estuário do Tejo (Alcoforado, 1987; Alcoforado et al., 2009), e ainda para discriminação do efeito da salinidade (*spray* salino) e da penetração da toalha freática salgada.

<sup>iii</sup> O Modelo Numérico de Elevação (MNE) e índices topográficos derivados foram determinados tendo como base a extensão territorial definida pelas folhas 403, 416 e 417 da altimetria à escala 1:10 000, sendo os mesmos representados ao nível do concelho de Loures. De referir que foram consideradas as linhas de fecho na delimitação da área de cálculo do Índice Topográfico de Humidade. Este índice caracteriza as zonas de saturação de água superficial e o conteúdo de água nos solos, tendo sido consideradas as seguintes situações: solos bem drenados (<0,3), solos moderadamente drenados (0,3 a 0,6) e solos mal drenados (0,6 a 1) (Moore et al., 1988). Relativamente à variável *Sky View Factor* foram consideradas as seguintes formas de relevo: vales muito encaixados (SVF<0,7), vales encaixados (0,7<SVF<0,8) e vales abertos (SVF>0,8) (Mora, 2006).