THE EFFECTS OF PRESENCE/ABSENCE VS. CONTINUOUS SUITABILITY DATA ON RESERVE SELECTION

SPYROS TSIFTSIS[1],[*], IOANNIS TSIRIPIDIS[1], PANAYIOTIS TRIGAS[2], and VASSILIKI KARAGIANNAKIDOU[1]

[1] Department of Botany, School of Biology, Aristotle University of Thessaloniki, GR-54124 Thessaloniki, Greece
[2] Laboratory of Systematic Botany, Department of Agricultural Biotechnology, Agricultural University of Athens, GR-11855 Athens, Greece
[*] Corresponding author: stsiftsi@bio.auth.gr

ABSTRACT

Species distribution models are widely applied for generating the appropriate data for designing networks of reserve systems. The aim of the present study is to compare the use of presence/absence and continuous suitability data, derived from species distribution models, in reserve selection, and to detect any systematic trends in the reserve networks produced on the basis of these two types of data. The MAXENT model was applied to predict orchids' potential distribution in east Macedonia (northeast Greece). One presence/absence and one with suitability values data set was made and used in the ZONATION decision support tool in order to prioritize the cells of the study area for inclusion in a reserve network. In the selection procedure, species weighting has been applied by using the species specialization index. Results show that the presence/absence data set favors the selection of cells with more extreme climatic conditions and more distant from the mean habitat of the study area. Furthermore, the continuous suitability data set provides higher suitability values for the specialist taxa in comparison with the presence/absence data set, while the opposite is happening for the generalist taxa. The present study reveals that the suitability data outperform the presence/absence data in reserve selection because: (a) they better represent the average environmental conditions of the study area in the selected networks; (b) they ensure higher suitability values for the specialist species, which are more prone to extinction in the future; and (c) they take full advantage of species weighting according to their habitat specificity.

Keywords: Conservation planning, east Macedonia, efficiency, orchids, species distribution models

Introduction

Indirect and direct human activities, including climate change, are the main threats for plant species in the Mediterranean Basin (Montmollin and Strahm 2005; Kladis et al. 2011). Setting up conservation area networks is the main tool for in situ conservation of biodiversity across the world (Ravenel and Redford 2005). The first national park worldwide was created over a century ago and the increasing need for biodiversity conservation has resulted in many countries establishing networks of protected areas. However, the criteria for the establishment of these networks were in many cases their outstanding physical features, natural beauty, or the distribution of particular species (Margules and Sarkar 2007), rather than their contribution to the overall goal, protecting examples of as much variation in biodiversity as possible.

In recent times, several methodologies for conservation network selection have been developed to redress this problem (Csuti et al. 1997; Rey Benayas and de la Montana 2003; Pawar et al. 2007). Due to the fact that the establishment and management of conservation reserves are expensive (Rodrigues et al. 2000), the aim of these methodologies is to maximize biodiversity subject to a cost constraint, or minimize cost subject to a set of biodiversity constraints (targets). Most traditional systematic conservation planning approaches do the latter. Zonation attempts to trade-off cost and biodiversity achievement therefore does not really maximize or minimize either. Systematic conservation planning is usually based on

biodiversity patterns, although the involvement of biodiversity processes and dynamic threats has also been recognized as necessary (Pressey et al. 2007). The former concern usually biodiversity features that can be recorded and mapped in an area. As conservation planning is inherently spatial (Pressey et al. 2007), the effectiveness (this concerns the ability of a network to conserve in the long-term, the highest amount of biodiversity; Rodrigues et al. 2000) of reserve networks depends much on the quality and the resolution of the data set that will be used (Brotons et al. 2004; Thuiller et al. 2005b; Araújo et al. 2006; Margules and Sarkar 2007), as well as on its uncertainty (Rocchini et al. 2011).

Biodiversity distribution data, even in the case of few and well-known biodiversity surrogates, are in many cases lacking, and their collection takes time, funds, and other human resources, which are often unaffordable (Margules and Sarkar 2007). In order to deal with the scarcity of distribution data, novel techniques such as species distribution models (SDMs) can be applied. These methods use the relationship between species known distribution and environmental conditions to predict species potential distribution (Guisan et al. 2002; Pawar et al. 2007), by identifying areas that are ecologically suitable for a species. The outcome of these models is a continuous suitability map, which in conservation and environmental management practice, is usually converted into presence/absence data (Liu et al. 2005).

Despite the fact that the results of species distribution models often suffer from high levels of uncertainty due

to several factors, concerning biased species distribution data, errors in environmental variables used as predictors, spatial resolution and the modeling process – see Elith and Graham (2009) and Rocchini et al. (2011) for further information – SDMs have become widely accepted tools to predict species distributions – see Rocchini et al. (2011) and references therein.

The outputs of these models can be used in two ways. The suitability map can be converted into a presence/absence map by applying a specific threshold. While there are many studies that deal with optimal threshold selection, no selection rules have yet been developed (Liu et al. 2005; Phillips et al. 2006). Pearson et al. (2007) state that most of the threshold selection techniques are applicable only in cases where absence data are available, while they suggest the use of the lowest presence threshold (LPT) for presence-only data as it can be better interpreted ecologically. Alternatively, conservation planning can proceed using the continuous suitability data directly. This approach has the advantage of being able to account for differences amongst all sites in persistence probabilities. The use of presence/absence data instead of the suitability ones causes an unnecessary loss of information and introduces an increased degree of uncertainty according to Rocchini et al. (2011), which concerns the crucial role of threshold selection. Furthermore, the use of presence/absence data cause a reduction of flexibility in choosing the reserve networks (Wilson et al. 2005; Pawar et al. 2007) because they give equal weight to all the sites were a species has suitability values above the threshold, while suitability data give weight according to the species' suitability values (for the same reason the threshold selection is more crucial in the case of presence/absence data than in the suitability data). It is well known that the future persistence of the species under concern can be accomplished by selecting the sites that are most suitable for them to occur and persist (Rodrigues et al. 2000). Areas where species have high suitability values usually provide more appropriate habitats and lower local probabilities of extinction compared to areas where species have low suitability values (Araújo et al. 2002). The long-term persistence of biodiversity is the ultimate aim of conservation area networks planning (Matthies et al. 2004; Early and Thomas 2007).

The orchid family, with approximately 25 000 species (Chase et al. 2003), is an important group with respect to conservation biology (Pillon and Chase 2006), being at the front-line of extinction (Swarts and Dixon 2009). Their importance pertains to their complex biology (Rasmussen 1995; van der Cingel 1995), to their sensitivity to environmental changes (Vakhrameeva et al. 2008), as well as to their high extinction risk, compared to other plant families, as a result of natural and/or anthropogenic causes (Hutchings 1989; Kull et al. 2006). However, decrease of many orchid species happened in whole Europe as a result mainly of the loss or even alteration of their habitats (Wotavová et al. 2004; Janečková et al. 2006; Efimov 2011;

Feldmann and Prat 2011). Additionally, orchids are reputed as flagship species for many nature conservation organizations, while can be used as environmental quality indicators (Kindlmann et al. 2006).

The aim of the present study is to compare the conservation networks that result applying both presence/absence and suitability species distribution models derived data and using the orchid family as a case study. Both types of data sets have been generated on the basis of extensive distribution data on the orchids in east Macedonia and they were obtained by using the Maxent modeling technique. The aim of the comparison is to detect any systematic trends in the results produced by the use of these two data sets and thus investigate their possible effects in the efficiency and effectiveness of the reserve networks under different sources of uncertainty.

Material and Methods

Study site

The study area is situated in east Macedonia (longitude 23°20´ to 24°48´E, latitude 40°42´ to 41°34´N), corresponding to an area of about 15,000 km² and includes the mountains Menikion, Orvilos, Falakron, Pangeon, Vrondous, Rhodopi, Lekani, Symvolon, as well as a part of Mt Kerdilion; for more details on the study area see Tsiftsis et al. (2009). The orchid flora of the study area consists of 65 taxa according to Tsiftsis (2009) and supplementary unpublished data. The climate of east Macedonia can be characterized as Mediterranean in the areas of lower altitude, while it becomes more continental at higher altitudes as well as in the northern areas (Flocas et al. 1983; Kotini-Zampaka 1983).

Geologically, the study area belongs to the Rhodopi crystalline massif, and is dominated by crystalline bedrocks (marble, gneiss, and schist) and igneous rocks (granite, rhyolites, and granodiorite) (Mountrakis 1985). The vegetation of the investigated area is highly diverse due to meso- and microclimate and the substrate type differentiation, and the different types and history of human impact.

Species distribution data

The geographical distribution of the orchids was predicted by applying a specific ecological niche modeling technique. The prediction was based on real species' occurrences and on few – among many – environmental variables that were not highly correlated each other, while the derived orchids' distributions were further used in two different ways in order to determine the important areas for orchid conservation.

We used presence-only data of species distribution in the study area at a spatial resolution of 1 × 1 km. Distribution data were obtained from Baumann and Baumann

(1988), Hölzinger and Künkele (1988), Willing (1990), Syska (1995) and mainly from Tsiripidis et al. (2007), and Tsiftsis (2009). Also, unpublished data (detailed orchid records) of the authors of the present study, as well as of other colleagues (see acknowledgments) were added in the dataset. All the records were geo-referenced prior to the application of species distribution models and for the ones referring to grid cells of 1 × 1 km size, the coordinates of grids' centroid were used. However, the majority of the data (ca. more than 80%) used in the study represent field data with precise geographical coordinates, recorded mainly by the authors. In total, the data set comprised of 3,576 records of 65 orchid species occurrences, while the number of presence cells of each orchid species is presented in Table 1.

The geographical distribution of 54 out the 65 orchids was predicted by applying the species distribution models generated by Maxent software version 3.3.1 (Phillips et al. 2006; Pearson et al. 2007; Phillips and Dudík 2008). Maxent is considered as an appropriate technique in modeling species distributions, even with very small sample sizes (Elith et al. 2006; Pearson et al. 2007). Furthermore, it is preferable for conservation purposes, because although it generates high omission errors or false negative rates, it can simultaneously avoid commission errors (Anderson et al. 2003; Loiselle et al. 2003; Phillips et al. 2006). Maxent estimates species distribution based on presence occurrence data and background absences, by using the maximum entropy principle and computes a probability distribution based on environmental variables spread over the entire study area, corresponding to species' habitat suitability. Eleven taxa were excluded from distribution modeling, which concerned species growing in wet meadows (seven taxa), as well as those that were distributed in less than four presence cells (four taxa). The former species were excluded because they depended on rare and small-sized habitat types, which are not included in vegetation maps, and as a result, could not be modeled.

Initially, 54 environmental variables were selected as predictors. These were mapped at a 30 sec resolution (approximately 1 km^2) and 19 of them were bioclimatic variables, while the other 35 concerned altitude, vegetation type (18 categories of habitat types) and soil type (16 categories). The bioclimatic variables, as well as the altitude were obtained from the WorldClim database (Hijmans et al. 2005), while the vegetation layer was obtained by digitizing the vegetation maps of the study area. For the soil data we used the soil types from the soil map of Eurasia at a scale of 1 : 1,000,000 (Anonymous 2004).

In order to deal with multi-collinearity the Variation Inflation Factor (VIF) was calculated using SPSS v. 16.0 (Anonymous 2007), and the continuous variables presenting VIF values higher than 10 were excluded from the analysis.

Maxent was applied using the auto-features mode, as suggested by Phillips and Dudík (2008) and the default settings. As the occurrence data are always biased, to-

tal species presence cells were added to the background sample (3,145 presence cells) (Phillips and Dudík 2008) incorporating the same bias into the pseudo-absences (Phillips et al. 2009). Seventy-five percent of the sample point data was used to generate species distribution models (training dataset) and the remaining 25% to test the accuracy of each model (test dataset). Finally, only those models with AUC values higher than 0.75 in the training data sets (Elith et al. 2006), and those in which the test data curve overcame the random prediction curve were used. Although there have been criticisms about the performance of AUC (Lobo et al. 2008), it is consider as a useful tool for the evaluation of models performance (Elith et al. 2006; Elith and Leathwick 2009). Nevertheless, any bias produced by the application of AUC is common in the analysis of the two data sets and we do not consider that affects the validity of the comparison of the results produced by the two data sets.

The output of the Maxent model is a continuous suitability map and in order to convert it to presence/absence data (Liu et al. 2005), the lowest presence threshold (LPT) was used. This threshold was preferred due to its advantage in identifying sites that are at least as suitable as those where a species' presence has been recorded (Pearson et al. 2007).

Place prioritization

Place prioritization was determined using the Zonation software (v.2.0) (Moilanen et al. 2005; Moilanen 2007). Zonation is a site prioritization method (Moilanen et al. 2005) that produces a hierarchical prioritization of a landscape, based on the biological value of sites (cells). The algorithm uses the complementarity principle together with connectivity, and as a result, landscapes can be zoned according to their potential for conservation. This is done by iteratively removing cells from the full landscape whose loss will cause the smallest marginal decrease in the overall conservation value of the remaining landscape. By removing cells only from the edge of the landscape the algorithm promotes structural connectivity in the remaining area, while previously removed areas are shown as buffer zones (Moilanen and Kujala 2008). The prioritization was based on modeled (54 taxa), as well as, observed (11 taxa) species distributions. The removal rule used was the core-area Zonation, because the primary aim of the analysis was to ensure the conservation of the total orchid flora in the study area (Moilanen et al. 2005). In core-area Zonation, cell removal is done by picking cell i that has the smallest value (minimizing biological loss) for the most valuable occurrence over all species in the cell. However, when a part of the distribution of a species is removed, the proportion located in each remaining cell goes up. This means that Zonation tries to retain core areas of all species until the end of cell removal, even if the species is initially widespread and common (Moilanen et al. 2009).

In order to determine the important areas for orchid conservation two data sets were used. The first was the presence/absence data set, which was made considering that a taxon was present in a cell, when its suitability value in this was higher or equal to the lowest presence threshold. The second data set included the suitability values of taxa when these were higher or equal to the lowest presence threshold, while in the opposite case the taxa were considered as absent. The use of suitability values instead of presence/absence data, in the algorithm of Zonation, intended to give extra weight to those cells with high suitability values of taxa, while the consideration of taxa as absent in cells with very low suitability values (lower than the lowest presence threshold) was aimed at securing the non selection of cells unsuitable for taxa conservation.

Species Specialization Index (SSI) (Tsiftsis et al. 2009) was used to ascribe different weights to species. Species weighting is considered appropriate in conservation action planning, as different species do not have the same importance with regard to nature and biodiversity conservation (Arponen et al. 2005; Girardello et al. 2009). Species weighting is also a critical component of the Zonation algorithm. Although the full distribution of each species is assigned with the same value by that algorithm, weighting affects the order in which cells are removed during the selection procedure. Consequently, cells that include a part of the distribution of a valuable species (high weight) remain later in the iterative cell removal process than cells containing only low-weight species, assuming everything else is equal between the occurrences (Moilanen and Kujala 2008). Species specialization index (SSI) was based on the outputs of the Outlying Mean Index (OMI) analysis (Dolédec et al. 2000). Specifically, the two data sets described above were used in OMI analysis, to measure species tolerance and then by using the formula described by Tsiftsis et al. (2009), to calculate the species specialization index (SSI). Therefore, two indices resulted, which corresponded to the two data sets used. In the OMI analyses, the taxa occurring in only one locality were excluded and the maximum SSI value was given to them.

The solutions concerning the top fractions of 1, 2, 5, and 10% of the landscape prioritization, as well as the minimum number of cells required to safeguard at least one representation of each orchid in the study area were compared for the two data sets, using the Cohen's Kappa coefficient.

The Mann-Whitney U test was applied to compare the solutions obtained using the two data sets regarding: (a) the values of the environmental variables of the selected cells, (b) the Euclidean distances of the selected cells from the mean habitat of the study area calculated in the OMI analysis, and (c) the suitability values of each species in the selected cells. The comparison was made for all the top fractions mentioned above for (a)

and (b), while the minimum number of cells required to safeguard at least one representation of each the orchids was excluded in (c), because of the very small number of cases.

Finally, we compared the habitat suitability values of species between the cells selected by the two data sets, in order to explore if there is any systematic difference in them. Specifically, we found all the possible combinations of cells selected by the two methods and then we calculated the differences of habitat suitability values between them for each species. From these differences we found an average value for each species and then we calculated the non-parametric Spearman's correlation coefficient between these averages and species SSI. This was done in order to investigate if there is any systematic bias of habitat suitability of species according the data set used in network selection and in relation to their SSI. The average difference between all the combinations of cells of the two networks for each species, which reaches in some cases to several thousands, was considered that represents well the mean difference of the suitability values of species between the selected cells using the two data sets. Spearman's correlation coefficient was calculated for all species and for those ones found having significant differences of probabilities of occurrence between the cells selected by the two data sets. The SSI calculated on the basis of the suitability data was used in the above-mentioned correlation analysis. Mann-Whitney U test and correlation analysis were done using SPSS v. 16.0 (Anonymous 2007).

Results

From the total number of the continuous variables (19 bioclimatic variables and the altitude) eight of them (BIO2: mean diurnal range, BIO3: isothermality, BIO5: maxi-mum temperature of the warmest month, BIO6: minimum temperature of the coldest month, BIO8: mean temperature of the wettest quarter, BIO13: precipitation of the wettest month, BIO14: precipitation of the driest month, and ALT: altitude) had values of VIF < 10, and together with the vegetation and the soil type layers were included in the model.

Models generated with the Maxent software for the 54 taxa were sufficiently accurate (AUC ≥ 0.854; mean ± SD: 0.95 ± 0.04) and were used in the Zonation software together with the taxa that were not modeled. The most important variables contributing to 64.5% of the species distribution models were those concerning soil type (mean ± SD: 24.63% ± 15.74), vegetation (mean ± SD: 20.03% ± 15.19) and altitude (mean ± SD: 19.89% ± 17.79), while those with less importance were the mean diurnal range (mean ± SD: 0.95% ± 0.04) and the mean temperature of the wettest quarter (mean ± SD: 1.89% ± 4.45) (Table 1).

Table 1 Number of records (N), ratio of presence cells against background absences (R in % or training prevalence), Area Under the ROC Curve (AUC), and relative contributions of the used explanatory variables (in %). Alt : Altitude; Soil: soil type; Veg: vegetation type.

	N	R	AUC	Bio2	Bio3	Bio5	Bio6	Bio8	Bio13	Bio14	Alt	Soil	Veg
Anacamptis coriophora ssp. *coriophora*	11	0.3	0.961	0	0	0	0	0	14	58.4	5.2	5.3	17.1
Anacamptis coriophora ssp. *fragrans*	41	1.3	0.934	10.6	17.6	17.3	2.3	0.5	0	2.8	9.1	10.7	29
Anacamptis laxiflora	19	0.6	0.958	2.4	2.8	4.5	0.4	0	0.5	27.8	0.5	30.8	30.2
Anacamptis morio	45	1.4	0.95	0	0.6	0.1	0.9	0.6	9.4	9.7	41.1	23.4	14.1
Anacamptis papilionacea	34	1.1	0.917	22.1	5.6	4.7	0	0	13.9	7.5	16.2	25.9	4
Anacamptis picta	215	6.8	0.866	9.7	17.7	32.8	0.1	0.3	1.6	2.7	15.6	12.1	7.4
Anacamptis pyramidalis	115	3.7	0.878	12.6	27.5	8.2	0.2	19.7	10.7	4	5.7	6.7	4.7
Cephalanthera damasonium	120	3.8	0.927	0	3	0.3	7.5	15.8	6.1	5.5	41.7	7.8	12.3
Cephalanthera longifolia	105	3.3	0.854	24.9	0.1	2.2	8.6	0.1	13.1	11.3	1.3	18.1	20.2
Cephalanthera rubra	126	4.0	0.92	0.4	0.1	51.3	0.1	2.3	6.6	0.2	16.4	7.8	14.8
Coeloglossum viride	30	1.0	0.958	0.4	0.2	0.4	0.2	0	1.8	3.4	47.5	34	12.2
Corallorhiza trifida	89	2.8	0.978	0.2	0	7.2	4.1	0.2	0.6	37.5	46.2	1.8	2.1
Dactylorhiza romana	5	0.2	0.989	3.4	1.8	0	0.6	0	0.9	0	0.1	53.9	39.4
Dactylorhiza saccifera	17	0.5	0.975	0.7	0.5	0	1.2	17.3	0.1	1.5	2	20.8	56
Dactylorhiza sambucina	123	3.9	0.951	5.8	11	18	2	10.2	0.7	4.3	36.6	8.7	2.8
Epipactis atrorubens ssp. *atrorubens*	24	0.8	0.988	3.7	0.5	15.7	1.2	0	0.4	1.9	31.6	39.2	5.8
Epipactis gracilis	93	3.0	0.979	0.9	0.1	0.7	1.4	0.1	14.3	39.9	23.8	8.6	10.1
Epipactis helleborine	206	6.6	0.881	1.7	6.2	28.3	0.5	5.7	3	3.3	37.7	4.8	8.8
Epipactis microphylla	33	1.0	0.975	1.7	2.5	2.4	2	1.1	23.8	5.1	22.4	15.1	23.9
Epipactis nauosaensis	4	0.1	0.998	0.5	2.9	0	0.1	0	3.7	0.4	9.6	29.9	52.9
Epipactis pontica	4	0.1	0.999	0.6	3.9	3.4	9.2	0	2.1	19.7	8	16.4	36.7
Epipogium aphyllum	4	0.1	0.993	0.5	0.1	0.2	0.4	5.2	0	6.4	22.7	12.6	52.1
Goodyera repens	43	1.4	0.991	0	0.3	4.1	0.2	1.4	0.2	5.3	30	29.6	29
Gymnadenia conopsea	70	2.2	0.957	0	0.6	3.5	0	0.6	0.8	0.3	70.9	12.9	10.4
Gymnadenia rhellicani	5	0.2	0.998	0	0.2	0	0	0	0	0	53.1	22.4	24.3
Himantoglossum caprinum	119	3.8	0.901	11.1	15.1	25.3	6.8	1	6.4	1.5	11.6	9.9	11.4
Limodorum abortivum	75	2.4	0.883	11	1.1	5.9	7.1	0	11	0.6	29.1	31.4	2.8
Neotinea tridentata	160	5.1	0.937	5.2	21.2	4.4	1	10.6	2.9	14.1	13.1	20.8	6.7
Neotinea ustulata	48	1.5	0.955	2.8	0.4	14.9	2.8	0	2.3	8.3	43.9	11.3	13.2
Neottia cordata	5	0.2	0.996	4.9	0	0	2.5	0	0	0	11.8	24.4	56.4
Neottia nidus-avis	267	8.5	0.931	1.5	3.1	0.2	1.8	2.4	3	6.1	39.1	18.7	24.2
Neottia ovata	47	1.5	0.956	0	0.8	16	0	0.6	6.5	13.3	34	17.5	11.3
Ophrys apifera	18	0.6	0.924	0	7.2	0	23.3	0	1.4	0	0	20.8	47.3
Ophrys cornuta	87	2.8	0.865	22.2	10.1	2.4	0.2	0.5	6.4	1.8	6.7	40.7	9
Ophrys epirotica	5	0.2	0.997	15	0.1	0	0	0	0	26	6.4	42.3	10.2
Ophrys grammica	21	0.7	0.974	1.2	7.4	24.8	0.3	0	14	4.4	12.6	24.2	11.2
Ophrys mammosa	135	4.3	0.932	18.5	26.5	6.6	0.6	3.4	4.2	5.1	17.6	11	6.5
Ophrys reinholdii	7	0.2	0.993	0.1	0.8	0	1.1	0	5.7	7.9	0	57.2	27.3
Ophrys zeusii	5	0.2	0.984	0.6	0.4	0	0	0	4.7	0	16.9	32.9	44.5
Orchis italic	38	1.2	0.982	4.2	4.9	1.6	2.1	0	25	21	8.2	24.5	8.6
Orchis militaris	4	0.1	0.996	0	0.3	0	0	0.7	8.6	0.6	44.1	36.6	9.1
Orchis pallens	10	0.3	0.991	0.2	0	0.7	0.5	0	3.9	2.4	15.1	31.2	46
Orchis pauciflora	8	0.3	0.987	0.1	0	22.2	0.6	0.2	3.4	0	0.7	45.1	27.7
Orchis pinetorum	90	2.9	0.924	0.3	0.4	0	0	0.1	2.2	2.3	63.3	19	12.4
Orchis provincialis	5	0.2	0.998	0.6	0	0	2.8	0	6.5	0	0	69.9	20.2
Orchis purpurea	64	2.0	0.904	0.1	0	0	0.9	0	22.3	0.1	38.3	21.4	17

	N	R	AUC	Bio2	Bio3	Bio5	Bio6	Bio8	Bio13	Bio14	Alt	Soil	Veg
Orchis quadripunctata	39	1.2	0.986	1.7	0.7	1.5	0.6	0.1	21.7	6	15.5	43.7	8.5
Orchis simia	36	1.1	0.946	0.3	2.5	2.1	4.2	0	2	4.3	8.1	48	28.6
Platanthera bifolia	8	0.3	0.991	0	3.3	0	0	0	6.4	8.2	0.7	50.1	31.1
Platanthera chlorantha	160	5.1	0.901	0.4	1.3	57.3	0.1	1.3	17.3	0	11.2	3.9	7.2
Serapias bergonii	8	0.3	0.992	0	0.8	0	4.4	0.2	0.9	62.7	5.5	15.9	9.6
Serapias cordigera	4	0.1	0.989	0	0.2	0	7	0	0	0	7	57	28.8
Serapias vomeracea	26	0.8	0.984	1.9	0	0	1.1	0	12	54	12.3	15.2	3.5
Spiranthes spiralis	65	2.1	0.932	7.6	3.7	0.8	3.3	0	17	13.8	6.3	26.3	21.1
Mean	58.24	1.85	0.95	3.97	4.04	7.26	2.19	1.89	6.41	9.69	19.89	24.63	20.03
SD	61.73	1.96	0.04	6.36	6.73	12.57	3.79	4.45	6.85	14.96	17.79	15.74	15.19

Prioritization of east Macedonia using presence/absence data, as well as probabilities of occurrence is presented in Figure 1. The areas that had the highest conservation priority when the presence/absence data set was used were located mainly in Rhodopi and Falakron, and to a lesser extent on Mt. Pangeon and in the southwestern parts of Mt. Lekani. Furthermore, scattered cells were located on Mt. Orvilos and on Mt. Vrondous as well (Fig. 1a). When the continuous suitability values were used, the highest conservation priority areas were located on the same mountains, but had a more scattered distribution (Fig. 1b). The minimum number of cells required to safeguard at least one representation of all orchids of the study area was 13, for both data sets.

The comparison of the solutions for each data set, as measured by the Kappa coefficient, indicated a low agreement when the top 1% fraction was used, as well as when the minimum number of cells required for the conservation of all orchids was used ($k = 0.32$ and $k = 0.38$, respectively). A moderate agreement was found when the top 2% fraction was used ($k = 0.41$), and a substantial agreement when 5 and 10% of the top fractions were used ($k = 0.61$ and $k = 0.72$, respectively) (Figs 2–6).

The results of the Mann-Whitney U test concerning the environmental variables, as well as the Euclidean distances of the selected cells from the mean habitat conditions are presented in Table 2. These results show that in the case of the presence/absence data set, the selected cells have higher altitudes, usually lower values of mean diurnal range, maximum temperature of the warmest month, minimum temperature of the coldest month and mean temperature of the wettest quarter, and higher values of isothermality and precipitation of the wettest month. However, statistical differences ($P < 0.05$) for altitude, maximum temperature of the warmest month, minimum temperature of the coldest month and mean temperature of the wettest quarter were found only for the top 5 and 10% fractions, while for mean diurnal range and precipitation of the driest month only for the latter top fraction. Isothermality was not found to differ significantly in any of the top fractions.
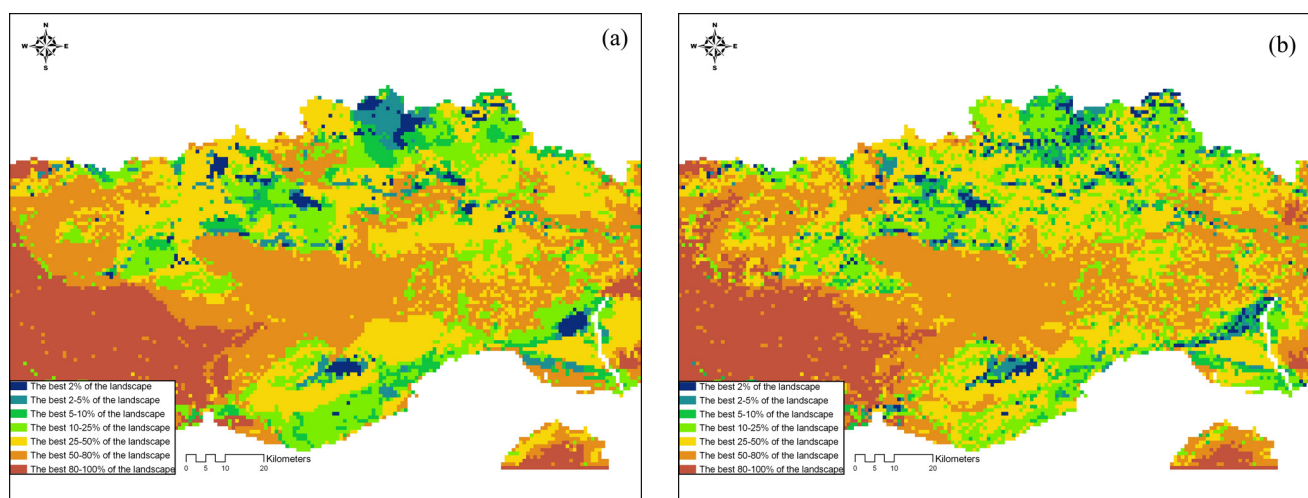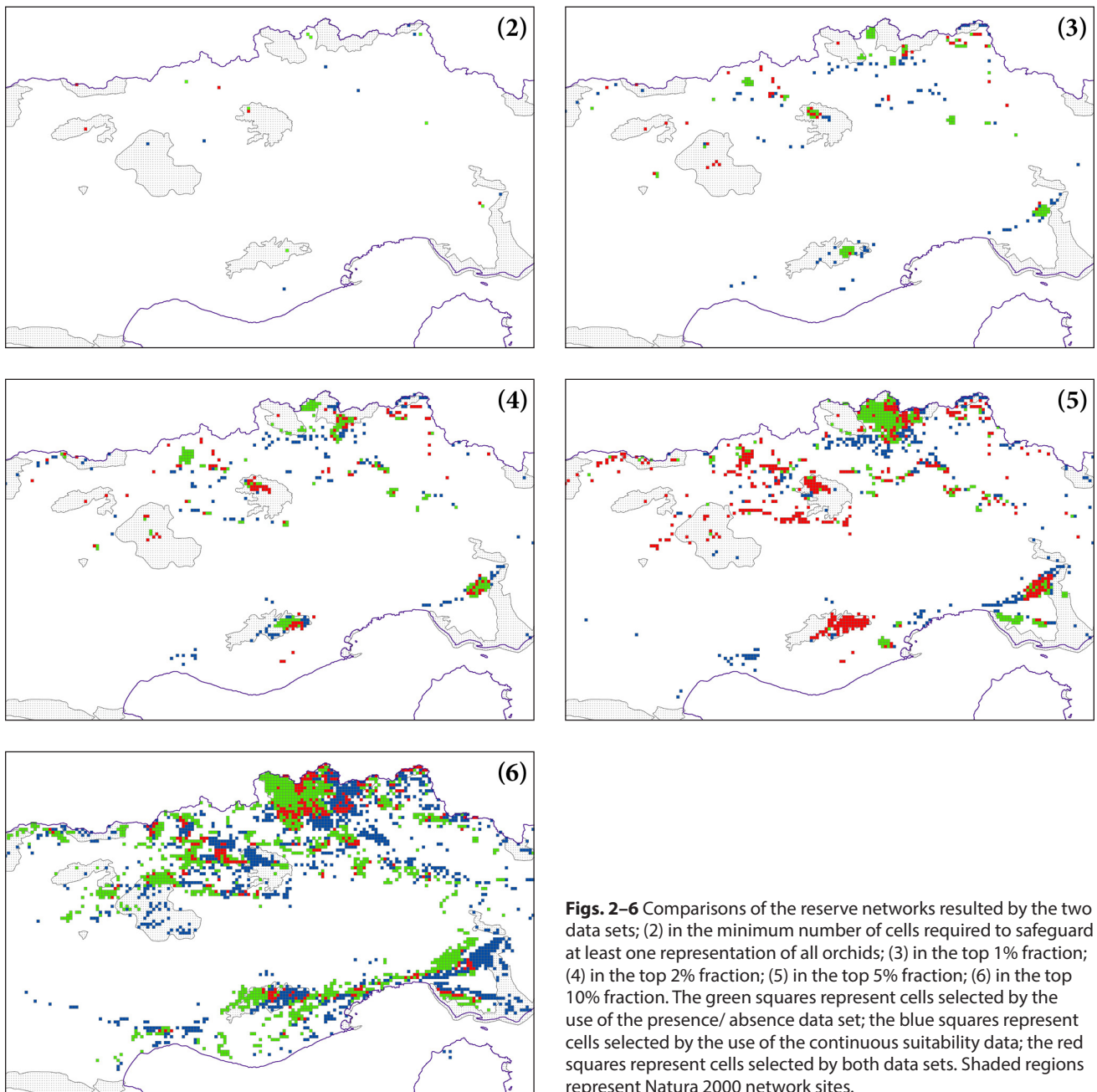

Fig. 1 Results of Zonation analysis using (a) presence/absence distribution data, and (b) continuous suitability data.

Figs. 2–6 Comparisons of the reserve networks resulted by the two data sets; (2) in the minimum number of cells required to safeguard at least one representation of all orchids; (3) in the top 1% fraction; (4) in the top 2% fraction; (5) in the top 5% fraction; (6) in the top 10% fraction. The green squares represent cells selected by the use of the presence/ absence data set; the blue squares represent cells selected by the use of the continuous suitability data; the red squares represent cells selected by both data sets. Shaded regions represent Natura 2000 network sites.

Table 2 Univariate comparisons (Mann-Whitney *U* test) of environmental variables and the Euclidean distances of the selected cells from the mean habitat conditions in OMI analyses. Each row represents the different top fractions of the landscape considered; RC: minimum number of cells required to safeguard at least one representation of each species; 1: indicate higher values for the presence/absence data set; 2 indicate higher values for the data set of the suitability values.

	Alt	Bio2	Bio3	Bio5	Bio6	Bio8	Bio13	Bio14	Euc. Dist.
RC	1	1	1	2	1	2	2	2	1
1%	1	2	2	2	2	2	1	1	1
2%	1	1	1	1	1	1	2	2	2
5%	1*	2	1	2*	2*	2*	1	1	1**
10%	1*	2*	1	2*	2*	2*	1	1*	1**

*p < 0.05, **p < 0.01

The Euclidean distances of the selected cells using the presence/absence data set were higher for all the top fractions used, except from the top 2% fraction. Also, in this case the test was found to be significant ($P < 0.01$) for the top 5 and 10% fractions of the landscape.

The results of the pair-wise Mann-Whitney comparisons with regard to the species' suitability values in the selected cells by each data set, are presented in Appendix 1. The results are shown together with the species SSI, in order to investigate any trends concerning their niche breadth. The number of species having significantly different suitability values between the two data sets fluctuated from 15 to 19 species, depending on the top fraction of the landscape. Species presenting significant difference between the two data sets were usually those with a narrow niche breadth.

Spearman's correlation coefficient revealed that there are significant correlations between the average difference of suitability values of species in the cells selected using the two data sets and species SSI (see Table 3). Significant correlations were found for all the top fractions of the landscape tested when all the species were used and for all except the 2% top fraction when the species showing significant differences of suitability values between the cells selected by the two data sets were used. Taking into account that the differences were calculated abstracting from the suitability values of species in cells selected using the presence/absence data set those in the cells selected using the suitability data, and that the correlation coefficients are always negative, it may be concluded that the higher is species SSI the lower is the average difference. In other words, the presence/absence data set favors the suitability values of the generalists in the resulted networks, while the data set with the continuous suitability values favors the ones of the specialists.

Table 3 Spearman's Rho correlation coefficient between the average difference of suitability values of species in the selected cells using the two data set, and species SSI (calculated on the basis of the continuous suitability data). Second column presents the correlations calculated on the basis of all taxa, third column the ones for the taxa having significant differences of suitability values between the cells selected by the two data sets, and fourth column the number of the above-mentioned taxa.

	SSI all species	SSI sig. species	Cases
RC	−0.428**		
1%	−0.544**	−0.624*	15
2%	−0.467**	−0.335	17
5%	−0.432**	−0.588*	19
10%	−0.557**	−0.792**	17

*p < 0.05, **p < 0.01

Discussion

In previous studies, the outputs of species distribution models have been used as input data in area selection algorithms either as presence/absence data (Araújo et al. 2004; Wilson et al. 2005) or as suitability data (Pawar et al. 2007; Leathwick et al. 2008; Girardello et al. 2009; Tsiftsis et al. 2011). In the present study both presence/absence and continuous suitability data were used in the Zonation algorithm with the core-area removal rule and the resulted networks were compared and found that differed in terms of the climatic and habitat extremes sampled, habitat suitability values of generalist vs. specialists taxa, and the level of fragmentation of the reserve system. These differences were greater in the cases where the reserve systems are relatively small. This is logical as when more sites are included in the selected networks, there are more possibilities of having similar sites between the networks. However, taking into account that resources are limited and reserve systems are often very small fractions of a landscape (Margules and Sarkar 2007), our results are of crucial importance for the practice of conservation planning.

The use of the presence/absence data set resulted in the selection of cells at higher altitudes, with more extreme climatic conditions and more distant from the mean habitat of the study area. This trend was found for most top fractions of landscape selected, and the fact that it was found to be statistically significant for the highest ones, may be attributed to the higher number of selected cells in these top fractions. The cells with extreme conditions may be favored because they host one or few rare species and as the continuous habitat suitability values in a cell are not taking into account using the presence/absence data set, the selection procedure is driven mainly by the occurrence of rare species. On the other hand, the conservation value of such cells may be greatly downweighted in the case of the continuous suitability data because of the possible low habitat suitability for most of the other species occurring in them. Therefore, the presence/absence data set prioritizes higher cells with rare taxa, which tend to occur in more extreme ecological conditions, while the suitability data gives higher priority to cells with high habitat suitability for species and moderate ecological conditions.

The continuous suitability data ensures higher suitability values for the specialist taxa in comparison with the presence/absence data set, while the opposite is happening for the generalist taxa. This trend was found for all the top fractions of landscape tested. This result seems, at first glance, to be contradictory, with the trend of the first data set to select cells with more extreme conditions. However, many of the rare taxa in our data set were not found as specialists. Examples of such taxa are the *Orchis pallens, Anacamptis coriophora* ssp. *coriophora,* and *Epipactis atrorubens* ssp. *atrorubens*, which occur in few scattered localities, representing, however, different environmental conditions. These rare species are usually of northern origin, having their southernmost limits in the study area and their rarity is not related with their habitat specificity. Furthermore, the high altitude areas, which are selected

more often by using the presence/absence data set, mostly harbor generalist species (e.g. *Orchis pallens*, *Dactylorhiza sambucina*, *Gymnadenia conopsea*, *Corallorhiza trifida*), with high suitability values, while the cells having high suitability values for specialist species (e.g. *Serapias cordigera*, *Serapias bergonii*, *Ophrys zeusii*, *Ophrys epirotica*, *Orchis provincialis*) are found at medium altitude and in moderate climatic and edaphic conditions. The latter cells are more frequently selected by using the suitability data. This finding supports the opinion that species which are rare at local or regional scale, but not at broader scales or they do not owe their rarity to their habitat specificity but rather to their vagrant occurrence may bias the selection procedure in favor to the areas and habitats they occur (He and Gaston 2000; Hartley and Kunin 2003; Tsiftsis et al. 2009).

In addition, the continuous suitability data, taking into account both species suitability values in each cell and species weighting according to the SSI (niche breadth), drives the selection algorithm to prioritize cells with specialist taxa having high suitability values in them. On the other hand, the use of the presence/absence data in combination with the application of SSI cannot ensure that the cells selected for the conservation of the weighted specialist taxa also have high suitability values in them. Therefore, the presence/absence data does not fully take into advantage taxa weighting, while the suitability data does.

Finally, another difference between the networks as a result of the two data sets concerns their fragmentation. The presence/absence data set resulted in more compact networks than that with the suitability values (Figs 1–6). Connectivity is one desirable feature of conservation networks (Williams et al. 2005; Pressey et al. 2007), which theoretically enhances the probabilities of species persistence through metapopulation dynamics and reduction of edge effects, and it also facilitates conservation management (Moilanen and Wintle 2006). However, differences in connectivity between the results of the two data sets should not be used for comparison, as many methods have been developed to improve the connectivity of conservation networks (e.g. aggregation methods in Zonation or boundary length modifier in Marxan; Game and Grantham 2008; Moilanen and Kujala 2008), which, however, have not been applied in the present study.

According to Rodrigues et al. (2000), the species extinction rate is significantly and negatively correlated with the local species abundance and density, and the endangered species should receive a higher conservation investment. Thus, the continuous suitability data, ensuring higher suitability values for the specialist species, which for several reasons are also more prone to extinction (Travis 2003; Thuiller et al. 2005a), and weighting them in an appropriate and logical way, should lead to higher possibilities of long-term species persistence. The latter is

of the most urgent challenges that systematic conservation planning must face (Pressey et al. 2007).

Furthermore, the fact that the suitability data favor the selection of cells representing mostly average and not extreme environmental conditions may also be considered to improve long-term species persistence, especially in the face of climate change. Sites representing the centers of the geographical and environmental distribution of species may be anticipated to provide the optimal conditions for their long-term persistence, rather than the geographical and environmental marginal sites (Whittaker et al. 2005). Although there are arguments that marginal sites may harbor genotypes with higher capacity to adapt to suboptimal conditions (e.g. produced by adverse climatic change conditions; Smith et al. 2001), core areas should be prioritized first, as we already know that populations occurring in them are generally better adapted to cope with the current environmental conditions (Whittaker et al. 2005). This is true, especially in cases where limited area may be allocated to reserve networks and a small representation target has been set for the species.

In conclusion, suitability data offer advantages over presence/absence data in terms of (a) offering higher suitability values for the specialist species, and (b) taking full advantage of species weighting. These advantages are expected to have positive effects on species persistence in a selected conservation network. Suitability data may tend to offer a better representation of the average environmental conditions of the study area in the selected networks, but a more fragmented network, when compared with presence/absence data. Both of these factors should therefore be addressed explicitly in a conservation planning objective. Finally, our results reveal that species distribution models may not differ only in the degree and causes of uncertainty (see Rocchini et al. 2011), but they may also have systematic trends in the differentiation of their results.

Acknowledgements

REFERENCES

Anderson RP, Lew D, Peterson AT (2003) Evaluating predictive models of species' distributions: criteria for selecting optimal models. Ecol Model 162: 211–232.

Anonymous (2004) European Soil Database (v 2.0). European Soil Bureau Network and the European Commission, EUR 19945 EN.

Anonymous (2007) SPSS 16.0 for Windows. Rel. 16.0.1, SPSS Inc.

Araújo MB, Cabeza M, Thuiller W, Hannah L, Williams PH (2004) Would climate change drive species out of reserves? An assess-

ment of existing reserve-selection method. Global Change Biol 10: 1618–1626.

Araújo MB, Thuiller W, Pearson RG (2006) Climate warming and the decline of amphibians and reptiles in Europe. J Biogeogr 33: 1712–1728.

Araújo MB, Williams PH, Fuller RJ (2002) Dynamics of extinction and the selection of nature reserves. Proc R Soc Lond Ser B – Biol Sci 269: 1971–1980.

Arponen A, Heikkinen R, Thomas C, Moilanen A (2005) The value of biodiversity in reserve selection: representation, species weighting, and benefit functions. Conserv Biol 19: 2009–2014.

Baumann B, Baumann H (1988) Ein Beitrag zur Kenntnis der Gattung Epipactis Zinn im Mittelmeergebiet. Mitt Bl Arbeitskr Heim Orch Baden-Württ 20: 1–68.

Brotons L, Thuiller W, Araújo MB, Hirzel AH (2004) Presence-absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27: 437–448.

Chase MW, Cameron KM, Barrett RL, Freudebstein JV (2003) DNA data and Orchidaceae systematics: a new phylogenetic classification. In: Dixon KW, Kell SP, Barrett RL, Cribb PJ (eds) Orchid conservation. Natural History Publications, Kota Kinabalu, Sabah, pp 69–89.

Csuti B, Polasky S, Williams PH, Pressey RL, Camm JD, Kershaw M, Kiesterg AR, Downsh B, Hamiltone R, Husoi M, Sahri K (1997) A comparison of reserve selection algorithms using data on terrestrial vertebrates in Oregon. Biol Conserv 80: 83–97.

Dolédec S, Chessel D, Gimaret-Carpentier C (2000) Niche separation in community analysis: a new method. Ecology 81: 2914–2927.

Early R, Thomas CD (2007) Multispecies conservation planning: identifying landscapes for the conservation of viable populations using local and continental species priorities. J Appl Ecol 44: 253–262.

Efimov PG (2011) Revealing the decline and expansion of orchids on NW European Russia. Eur J Environ Sci 1: 7–17.

Elith J, Graham CH (2009) Do they? How do they? WHY do they differ? On finding reasons for differing performances of species distribution models. Ecography 32: 66–77.

Elith J, Graham CH, Anderson RP, Dudik M, Ferrier S, Guisan A, Hijmans RJ, Huettmann F, Leathwick JR, Lehmann A, Li J, Lohmann LG, Loiselle BA, Manion G, Moritz C, Nakamura M, Nakazawa Y, Overton JM, Peterson AT, Phillips SJ, Richardson K, Scachetti-Pereira R, Schapire RE, Soberon J, Williams S, Wisz MS, Zimmermann NE (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129–151.

Elith J, Leathwick J (2009) The contribution of species distribution modelling to conservation prioritization. In: Moilanen A, Wilson AK, Possingham HP (eds) Spatial conservation prioritization. Quantitative methods & computational tools. Oxford University Press Inc., New York, pp 70–93.

Feldmann P, Prat D (2011) Conservation recommendations from a large survey of French orchids. Eur J Environ Sci 1: 18–27.

Flocas AA, Giles BD, Angouridakis VE (1983) On the estimation of annual and monthly mean values of air temperature over Greece using stepwise multiple regression analysis. Arch Meteorol Geophys Bioclimatol Ser B Climatol Environ Meteorol Radiat Res 32: 287–295.

Game ET, Grantham HS (2008) Marxan User Manual: For Marxan version 1.8.10. University of Queensland, St. Lucia, Queensland, Australia, and Pacific Marine Analysis and Research Association, Vancouver, British Columbia, Canada.

Girardello M, Griggio M, Whittingham MJ, Rushton SP (2009) Identifying important areas for butterfly conservation in Italy. Anim Conserv 12: 20–28.

Grand J, Cummings MP, Rebelo TG, Ricketts TH, Neel MC (2007) Biased data reduce efficiency and effectiveness of conservation reserve networks. Ecol Lett 10: 364–374.

Guisan A, Edwards TC, Hastie T (2002) Generalized linear and generalized additive models in studies of species distributions: setting and scene. Ecol Model 157: 89–100.

Guisan A, Zimmermann NE (2000) Predictive habitat distribution models in ecology. Ecol Model 135: 147–186.

Hartley S, Kunin W (2003) Scale dependency of rarity, extinction risk, and conservation priority. Conserv Biol 17: 1559–1570.

He F, Gaston KJ (2000) Occupancy-abundance relationships and sampling scales. Ecography 23: 503–511.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. Int J Climatol 25: 1965–1978.

Hölzinger J, Künkele S (1988) *Dactylorhiza macedonica* sp. nova. Mitt Bl Arbeitskr Heim Orch Baden-Württ 20: 185–195.

Hutchings MJ (1989) Population biology and conservation of Ophrys sphegodes. In: Pritchard HW (ed) Modern methods in orchid conservation: the role of physiology, ecology and management. Cambridge University Press, Cambridge, pp 101–115.

Janečková P, Wotavová K, Schödelbauerová I, Jersáková J, Kindlmann P (2006) Relative effects of management and environmental conditions on performance and survival of populations of a terrestrial orchid, *Dactylorhiza majalis*. Biol Conserv 129: 40–49.

Kindlmann P, Kull T, Whigham D, Willems J (2006) Ecology and population dynamics of terrestrial orchids: An introduction. Folia Geobot 41: 1–2.

Kladis G, Panitsa M, Tsiripidis I, Sarris D, Dimopoulos P (2011) Vegetation ecology and diversity relationships in a riparian forest remnant of Western Greece. J Biol Res 16: 237–254.

Kotini-Zampaka S (1983) Contribution to the monthly study of the climate of Greece. Ph.D. thesis, Aristotle University of Thessaloniki (In Greek).

Kull T, Kindlmann P, Hutchings M, Primack B (2006) Conservation biology of orchids: Introduction to the special issue. Biol Conserv 129: 1–3.

Leathwick J, Moilanen A, Francis M, Elith J, Taylor P, Julian K, Hustie T, Duffy C (2008) Novel methods for the design and evaluation of marine protected areas in offshore waters. Conserv Lett 1: 91–102.

Liu C, Berry PM, Dawson TP, Pearson RG (2005) Selecting threshold of occurrence in the prediction of species distributions. Ecography 28: 385–393.

Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. Global Ecol Biogeogr 17: 145–151.

Loiselle BA, Howell CA, Graham CH, Goerck JM, Brooks T, Smith KG, Williams PH (2003) Avoiding pitfalls of using species distribution models in conservation planning. Conserv Biol 17: 1591–1600.

Margules CR, Sarkar S (2007) Systematic Conservation Planning. Cambridge University Press.

Matthies D, Brauer I, Maibom W, Tscharntke T (2004) Population size and the risk of local extinction: empirical evidence from rare plants. Oikos 105: 481–488.

Moilanen A (2007) Landscape zonation, benefit functions and target-based planning: unifying reserve selection strategies. Biol Conserv 134: 571–579.

Moilanen A, Franco AMA, Early RI, Fox R, Wintle B, Thomas CD (2005) Prioritizing multiple-use landscapes for conservation: methods for large multi-species planning problems. Proc R Soc Lond Ser B – Biol Sci 272: 1885–1891.

Moilanen A, Kujala H (2008) Zonation spatial conservation planning framework and software v. 2.0. User manual. http://www.helsinki.fi/bioscience/consplan.

Moilanen A, Kujala H, Leathwick JR (2009) The Zonation framework and software for conservation prioritization. In: Moilanen A, Wilson AK, Possingham HP (eds) Spatial conservation prioritization. Quantitative methods & computational tools. Oxford University Press Inc., New York, pp 196–210.

Moilanen A, Wintle BA (2007) The boundary-quality penalty: a quantitative method for approximating species responses to fragmentation in reserve selection. Conserv Biol 21: 355–364.

Montmollin B de, Strahm W (eds) (2005) The top 50 Mediterranean island plants: wild plants at the brink of extinction, and what is needed to save them. IUCN/SSC Mediterranean Islands Plant Specialist Group. IUCN.

Mountrakis D (1985) Geology of Greece. Thessaloniki (In Greek).

Pawar S, Koo MS, Kelley C, Ahmed MF, Chaudhuri S, Sarkar S (2007) Conservation assessment and prioritization of areas in Northeast India: Priorities for amphibians and reptiles. Biol Conserv 136: 346–361.

Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J Biogeogr 34: 102–117.

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190: 231–259. Software available on: http://www.cs.princeton.edu/~schapire/maxent.

Phillips SJ, Dudík M (2008) Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. Ecography 31: 161–175.

Phillips SJ, Dudík M, Elith J, Graham C, Lehmann A, Leathwick JR, Ferrier S (2009) Sample Selection Bias and Presence-Only Models of Species Distributions: Implications for Selection of Background and Pseudo-absences. Ecol Appl 19: 181–197.

Pillon Y, Chase M (2006) Taxonomic exaggeration and its effects on orchid conservation. Conserv Biol 21: 263–265.

Pressey RL, Cabeza M, Watts ME, Cowling RM, Wilson KA (2007) Conservation planning in a changing world. Trends Ecol Evol 22: 583–592.

Rasmussen HN (1995) Terrestrial orchids from seed to mycotrophic plant. Cambridge University Press, Cambridge.

Ravenel RM, Redford KH (2005) Understanding IUCN protected area categories. Nat Area J 25: 381–389.

Rey Benayas JM, de la Montana E (2003) Identifying areas of high-value vertebrate diversity for strengthening conservation. Biol Conserv 114: 357–370.

Rocchini D, Hortal J, Lengyel S, Lobo JM, Jiménez-Valverde A, Ricotta C, Bacaro G, Chiarucci A (2011) Accounting for uncertainty when mapping species distributions: The need for maps of ignorance. Prog Phys Geog 35: 211–226.

Rodrigues AS, Gregory RD, Gaston KJ (2000) Robustness of reserve selection procedures under temporal species turnover. Proc R Soc Lond Ser B – Biol Sci 267: 49–55.

Smith TB, Kark S, Schneider CJ, Wayne RK, Moritz C (2001) Biodiversity hotspots and beyond: the need for preserving environmental transitions. Trend Ecol Evol 16: 431.

Swarts ND, Dixon KW (2009) Terrestrial orchid conservation in the age of extinction. Ann Bot 104: 543–556.

Syska M (1995) Die Orchideenflora des westlichen Nestos-Deltas und des angrenzenden Berglandes (Nordost-Griechenland): Verbreitung – Ökologie – Gefahrdung. Jour Eur Orch 27: 339–552.

Thuiller W, Lavorel S, Araújo MB (2005a) Niche properties and geographical extent as predictors of species sensitivity to climate change. Global Ecol Biogeogr 14: 347–357.

Thuiller W, Lavorel S, Araújo MB, Sykes MT, Prentice IC (2005b) Climate change threats to plant diversity in Europe. P Natl Acad Sci USA 102: 8245–8250.

Travis JMJ (2003) Climate change and habitat destruction: a deadly anthropogenic cocktail. Proc R Soc Lond Ser B – Biol Sci 270: 467–473.

Tsiftsis S (2009) The orchids (Orchidaceae) of E. Macedonia: distribution, ecology and high conservation value areas. Ph.D. thesis, Aristotle University of Thessaloniki (In greek with English summary).

Tsiftsis S, Tsiripidis I, Karagiannakidou V (2009) Identifying areas of high importance for orchid conservation in east Macedonia (NE Greece). Biodivers Conserv 18: 1765–1780.

Tsiftsis S, Tsiripidis I, Trigas P (2011) Identifying important areas for orchid conservation in Crete. Eur J Environ Sci 1: 28−37.

Tsiripidis I, Karagiannakidou V, Alifragis D, Athanasiadis N (2007) Classification and gradient analysis of the beech forest vegetation of the southern Rodopi (Northeast Greece). Folia Geobot 42: 249–270.

Vakhrameeva MG, Tatarenko IV, Varlygina TI, Torosyan GK, Zagulski MN (2008) Orchids of Russia and adjacent countries (within the borders of the former USSR). ARG Gantner Verlag KG, Ruggell/Liechtenstein.

van der Cingel NA (1995) An atlas of orchid pollination. European orchids. AA Balkema, Rotterdam.

Whittaker RJ, Araújo MB, Jepson P, Ladle RJ, Watson JEM, Willis KJ (2005) Conservation Biogeography: assessment and prospect. Divers Distrib 11: 3–23.

Williams JC, ReVelle CS, Levin SA (2005) Spatial attributes and reserve design models: a review. Environ Model Assess 10: 163–181.

Willing B (1990) *Dactylorhiza* in Nordwestgriechenland – neue Erkenntnisse. Jber Naturw Ver Wuppertal 43: 58–70.

Wilson KA, Westphal MI, Possingham HP, Elith J (2005) Sensitivity of conservation planning to different approaches to using predicted species distribution data. Biol Conserv 122: 99–112.

Wotavová K, Balounová Z, Kindlmann P (2004) Factors affecting persistence of terrestrial orchids in wet meadows and implications for their conservation in a changing agricultural landscape. Biol Conserv 118: 271–279.

Appendix 1 Species specialization index (on the basis of the suitability data), and number of cells where each taxon is distributed in the resulted networks for the top fractions considered. Numbers in bold indicate where the data set resulted higher suitability values (Suit: suitability data; Bin: presence/absence data set); *p < 0.05%.

Taxa	SSI	RC		1%		2%		5%		10%	
	Suit.	Bin.	Suit.	Bin.	Suit.	Bin.	Suit.	Bin.	Suit.	Bin.	Suit.
Anacamptis coriophora ssp. *coriophora*	0.571	1	2	**20**	29	31	**54**	**85**	119	**202**	207
Anacamptis coriophora ssp. *fragrans*	0.613	7	8	81	**86**	161	**167**	448	**420***	873	**856***
Anacamptis laxiflora	0.659	2	2	27	**27***	62	**50***	193	**149**	432	**391**
Anacamptis morio	0.622	3	3	30	**34**	62	**62**	183	**156**	361	**318**
Anacamptis papilionacea	0.634	2	5	48	**54**	116	**110**	294	273	620	**622**
Anacamptis picta	0.610	8	8	**79**	84	**167**	159	**393***	435	**810**	852
Anacamptis pyramidalis	0.609	7	8	73	**83***	148	**153***	370	**400**	764	**783***
Cephalanthera damasonium	0.497	10	8	108	**80**	197	162	**475***	433	**863***	827
Cephalanthera longifolia	0.571	11	11	**114**	109	223	**229**	560	**540**	1095	**1063***
Cephalanthera rubra	0.528	10	9	**111**	90	**203**	187	**498**	463	**909**	888
Coeloglossum viride	0.519	6	5	47	**42**	84	85	211	219	**343**	377
Corallorhiza trifida	0.461	7	6	**56**	42	**104**	90	271	**196**	438	**361**
Dactylorhiza romana	0.765	1	2	14	**12***	31	**23***	78	**62***	182	**144***
Dactylorhiza saccifera	0.554	5	3	**50**	30	**91**	63	**221**	151	**304**	258
Dactylorhiza sambucina	0.399	9	9	83	**78**	156	156	**420**	375	**729**	676
Epipactis atrorubens ssp. *atrorubens*	0.685	3	1	31	**21**	58	**51**	135	**134**	191	**197**
Epipactis gracilis	0.610	5	2	61	**34**	108	**73**	249	**187**	390	353
Epipactis helleborine	0.536	12	13	**141**	137	**280**	273	**704**	665	**1299**	1294
Epipactis microphylla	0.648	3	2	40	**29**	67	**61**	154	**143**	235	**227**
Epipactis nauosaensis	0.700	1	1	11	**11**	25	**20***	47	**43**	47	47
Epipactis pontica	0.650	1	1	13	**8**	13	13	13	13	13	13
Epipogium aphyllum	0.637	1	1	12	**10***	26	**22***	64	**57***	150	**123***
Goodyera repens	0.419	2	1	**36***	21	**70***	48	**198***	111	**267***	207
Gymnadenia conopsea	0.476	8	7	80	**68**	144	**138**	389	**322**	640	569
Gymnadenia rhellicani	0.708	1	1	12	**11***	27	**23***	65	**63**	65	65
Himantoglossum caprinum	0.648	6	7	67	**82**	142	**149**	332	**381**	770	**802***
Limodorum abortivum	0.637	6	9	80	**86***	168	**168***	420	**432***	859	**863**
Neotinea tridentata	0.577	10	9	77	**80**	143	**150**	356	**360**	614	**645**
Neotinea ustulata	0.548	8	5	69	**53**	127	**114***	289	**278**	492	464
Neottia cordata	0.711	1	1	15	**12***	32	**24***	80	**60***	106	106
Neottia nidus-avis	0.490	10	8	**108**	89	**191**	176	**446***	420	**841**	782
Neottia ovata	0.582	5	7	**58***	62	**112**	119	**301***	275	**603***	559
Ophrys apifera	0.629	2	1	34	**37**	84	77	**177***	192	394	**392**
Ophrys cornuta	0.640	7	9	79	**94**	166	**181**	431	**486***	980	**1046***
Ophrys epirotica	0.802	1	1	14	**12***	31	**26***	78	**70**	82	82
Ophrys grammica	0.614	3	5	24	**28***	56	**61**	158	**168**	303	**328**
Ophrys mammosa	0.686	4	5	**42**	40	**90**	85	214	**235**	431	**472***
Ophrys reinholdii	0.742	1	1	16	**17**	36	42	88	**100**	168	**165***
Ophrys zeusii	0.872	1	1	15	**17**	34	**34**	83	**89***	194	**193***
Orchis italica	0.777	2	2	24	**16**	53	**35**	76	**93***	169	**186***
Orchis militaris	0.734	2	1	14	**13***	30	**25***	69	**61**	121	**113**
Orchis pallens	0	1	2	13	**15**	22	**23**	50	51	86	**68**
Orchis pauciflora	0.602	1	2	**14**	13	**26**	29	62	**74***	145	**141***
Orchis pinetorum	0.485	9	9	**100**	90	183	185	**475**	432	**806**	770
Orchis provincialis	0.787	1	1	16	**13**	36	**27***	65	65	65	65
Orchis purpurea	0.552	6	7	74	**74**	138	**136***	347	**334**	712	675

Orchis quadripunctata	0.641	1	1	16	**14**	29	**44**	107	**125**	211	**221***
Orchis simia	0.687	2	4	32	**38**	60	**88**	143	**183**	284	**326**
Platanthera bifolia	0.814	2	2	14	**13***	31	**26***	77	**71***	181	**157***
Platanthera chlorantha	0.527	10	10	110	**100**	210	**209**	541	504	**997**	960
Serapias bergonii	0.890	1	1	16	**14***	35	**29***	86	**74***	186	**163**
Serapias cordigera	0.941	1	1	17	**15***	37	**30***	92	**84***	179	179
Serapias vomeracea	0.622	1	1	17	**15**	37	31	**79**	102	249	**229**
Spiranthes spiralis	0.655	1	3	26	**38**	73	79	**219***	221	488	**449**