

Allocating active conservation measures using species distribution models: a case study of red-footed falcon breeding site management in the Carpathian Basin

P. Fehérvári^{1,2}, S. Solt², P. Palatitz², K. Barna³, A. Ágoston³, J. Gergely³, A. Nagy⁴, K. Nagy² & A. Harnos¹

¹ Department of Biomathematics and Informatics, Faculty of Veterinary Sciences, Szent István University, Budapest, Hungary

² Red-footed Falcon Conservation Working Group, MME/BirdLife Hungary, Budapest, Hungary

³ Bird Protection and Study Society of Serbia, Novi Sad, Serbia

⁴ Milvus Group, Tirgu Mures, Romania

Keywords

Falco vespertinus; random forest; generalized boosted model; BIOMOD; grasslands; raptor.

Correspondence

Péter Fehérvári, István St. 2, Budapest, Hungary H-1078.

Tel: +36 20 9731311

Email: fpeter17@gmail.com

Editor: Res Altwegg

Associate Editor: Vincenzo Penteneriani

Received 30 November 2011; accepted 9 May 2012

doi:10.1111/j.1469-1795.2012.00559.x

Abstract

The red-footed falcon *Falco vespertinus* is an enigmatic colonial raptor of high international conservation concern. One of the identified threatening factors responsible for the recent worldwide population decline is the shortage of suitable colonial nesting sites. In theory, this problem can easily be resolved by establishing artificial colonies. However, the key to a successful large scale nest-box scheme is to provide these artificial colonies in habitats suitable for the species. A Hungarian–Serbian project aimed to establish such nesting facilities in northern Serbia, although the lack of recent full-scale habitat surveys hindered the designation of the locations of these artificial nesting sites. We used five different species distribution models to model the distribution of nest sites on a 10 × 10 km grid in Hungary and in Romania. We then used the ensemble predictions of the best performing models to project the probability of red-footed falcon nest site presence in northern Serbia (predicted area). The models showed that three variables (grasslands, pastures and broad-leaved forests) had the highest importance in describing the spatial pattern of nest sites in the modelling area. The extent of grasslands and pastures had positive effects, while broad-leaved forests had negative impact on the probability of nest site presence. The predictions classified all the currently known colonies in the predicted area correctly. Our results suggest that the potential breeding distribution in Serbia is similar to that of two decades ago, thus large-scale land use changes are presumably not responsible for the reported population decline. We have also reduced the extent of conservation target areas to 11.5%, allowing to pinpoint locations for these future nest box colonies, and also provided a basis for future conservation measures like allocating monitoring efforts and designating future Natura 2000 sites in Serbia.

Introduction

The recent advancement in statistical sciences, geographical information systems (GIS) and the computing power available led to a boom in the number of modelling approaches available for species distribution modelling (Hirzel & Le Lay, 2008; Elith & Leathwick, 2009). These species distribution models (SDMs) are applied in an array of fields ranging from evolutionary perspectives (Titeux *et al.*, 2007) to responses to climate change (Marini *et al.*, 2009) and others (Guisan & Thuiller, 2005). SDMs can play a crucial role in identifying key sites for endangered species; however, they are less often applied in conservation sciences compared with their relative potential (Engler, Guisan & Rechsteiner, 2004; Parviainen *et al.*, 2009; Wilson, Roberts & Reid, 2010).

One of the most important challenges of modern nature conservation is to prioritize activities to allocate scarce funding and resources effectively (Brooks *et al.*, 2006; Wilson, Carwardine & Possingham, 2009; Segan *et al.*, 2010). Increasing cost-effectiveness may be as simple as defining target areas where resources can be allocated. In case of certain rare, endangered and/or flagship species, major regional threats have already been assessed and methods to eliminate them developed (Heredia, Rose & Painter, 1996; Meyburg *et al.*, 2001; Palatitz *et al.*, 2009). Prioritization in these cases can be narrowed down to target areas based on various aspects like expert knowledge, expected conservation results or potential presence of the species in question (Bessa-Gomes & Petrucci-Fonseca, 2003; Olsson & Rogers, 2009; Yosef & Wineman, 2010; Zduniak & Yosef, 2012).

The red-footed falcon is a species of high international conservation concern ('near-threatened' in the International Union for Conservation of Nature Red List, ANNEX I of European Commission's Birds Directive 79/409/EEC) due to the drastic breeding population decline of the past decades (Palatitz *et al.*, 2009). This small raptor is a facultative colonial breeder (i.e. breeding in colonies and in solitary pairs) that does not build a nest; falcons naturally breed in rook *Corvus frugilegus* colonies (rookeries, Horváth, 1964; Purger & Tepavcevic, 1999) or in loose aggregations of magpie *Pica pica* nests (Végyvári, Magnier & Nogues, 2001).

Red-footed falcons have suffered from loss of foraging and breeding habitats during the past decades as their foraging habitats have been, while the drastic decrease in rookeries within the Carpathian Basin has had a serious impact on the available nesting sites. In Hungary, the landscape scale distribution of rookeries remained stable, while the density and size of rookeries decreased and their location shifted to human settlements (Fehérvári *et al.*, 2009), while similar patterns have been reported from other European countries (Orlowski & Czapulak, 2007). The reasons of rookery declines can be attributed to a large-scale persecution in the mid-80s, resulting in a massive 90% population crash. The obvious consequences for red-footed falcons was that most of the potential breeding colonies disappeared, causing a shrinkage in distribution range, and a decline in the number of breeding pairs (Fehérvári *et al.*, 2009 and see references therein). Moreover, the ratio of solitary breeders increased, with only 40–50% of the whole population used colonies for breeding at the turn of the century (Bagyura & Palatitz, 2004).

A previous conservation project (Conservation of *Falco vespertinus* in the Pannonian Region LIFE05 NAT/H/000122) has already developed a method to compensate for the lack of nesting sites by establishing artificial nest box colonies in habitats thought to be optimal for the birds. However, individual experience, anecdotic historical data and the protection status of the area had more roles in the designation of the location of these sites than verified scientific knowledge. The project's monitoring scheme later provided precise data on the spatial pattern and occupancy rate of these artificial nest sites.

While red-footed falcon monitoring and conservation efforts have been implemented on a wide spatial scale in Hungary and in Romania, there is scant information on recent population trends, and distribution from northern Serbia, where approximately 5–10% of the total European Union population is thought to breed (Palatitz *et al.*, 2009). However, reports of population decline and spatial distribution are available from the early 2000s (Purger, 2008). A recently initiated international project aimed to fill in the knowledge gaps by adapting the monitoring scheme and fund the implementation of conservation activities by placing 800 nest boxes for red-footed falcons in northern Serbia. The short duration of the project (15 months) hindered the implementation of a thorough red-footed falcon breeding site survey of the target areas, and thus designating

the optimal sites for new colonies based on recent field experience was not feasible.

In the present study, we focused on utilizing the vast and accurate data available on red-footed falcon distribution in Hungary and Western Romania, through assessing the landscape scale habitat preference of red-footed falcons with SDMs. Initially, we focused on understanding the relationship between landscape scale habitat variables and red-footed falcon presence. This information was later used to predict the presence of suitable potential nest sites in northern Serbia, to identify target areas for surveying and to aid the designation of the location of artificial colonies.

Materials and methods

Red-footed falcon distribution

Red-footed falcon distribution data were derived from a monitoring scheme, developed and first applied in Hungary and in Western Romania within the framework of the 'Conservation of *F. vespertinus* in the Pannonian Region' LIFE Nature program (<http://www.falcoproject.hu>). The monitoring concerns all nesting types, that is natural colonies (rookeries, dense assemblages of magpie nests), artificial colonies and solitary nests. In case of the latter two types, the applied scheme allows to precisely estimate the number of breeding pairs. However, in case of rookeries, typically presence/absence data are more reliable than the vague expert estimates given. Intuitively, there is also a considerable difference in the probability of discovering the location of red-footed falcon breeding sites considering the different nesting types. Obviously, the probability is close to 1 in case of artificial colonies; however, in case of the other two types, it is highly variable. Including artificial breeding sites in a distribution model may cause bias as the fact that breeding birds present at these nests is the product of human preconception of suitable habitats and the choice of the birds. The percentage of falcons breeding in natural nest sites during the study period was approximately 40 and 90% in Hungary and in Romania, respectively. This large deviation is due to the remarkably different number of rookeries on the two sides of the border. Additionally, there is a considerable difference in monitoring efforts made in the two countries, as the number of participants is approximately 10-fold larger in Hungary.

Our primary objective was to build models that can predict potential presence of the species (i.e. presence if nest sites are available or made available) and not to predict true presence. Using the data simultaneously from the two countries for model building allowed estimation of the broadest spectrum of potential nest sites, regardless of the nest building species.

We used the presence only data of the location of all monitored nesting sites, regardless of the number of breeding birds. As the position of all breeding sites was measured with handheld global positioning systems, we considered the

Table 1 Variables of the CORINE 2006 Land Cover project that were identified in the modelling area, and used as predictors in case of all models

| Var. code | Name (level 3) | ACRONYM | Modelling area surface coverage | | Predicted area surface coverage | |
|-----------|--------------------------------------|---------|---------------------------------|-------------|---------------------------------|-------------|
| | | | Mean % | \pm range | Mean % | \pm range |
| 1.1.1. | Continuous urban fabric* | CONURB | 0.01 | 0.00–0.70 | 0.00 | 0.00–0.35 |
| 1.1.2. | Discontinuous urban fabric | DISURB | 4.57 | 0.00–32.84 | 5.04 | 0.00–46.16 |
| 1.2.1 | Industrial and comm. units | INDCOM | 0.70 | 0.00–9.21 | 0.52 | 0.00–10.29 |
| 1.2.4. | Airports | AIRPOR | 0.08 | 0.00–5.32 | 0.05 | 0.00–7.49 |
| 1.3.1. | Mineral extraction sites | MINEXT | 0.06 | 0.00–9.39 | 0.07 | 0.00–2.92 |
| 1.3.2. | Dump sites | DUMPST | 0.00 | 0.05–7.09 | 0.01 | 0.00–2.69 |
| 1.3.3. | Construction sites | CONSTIT | 0.03 | 0.00–2.09 | 0.00 | 0.00–0.50 |
| 1.4.1. | Green urban areas | GREURB | 0.04 | 0.00–2.85 | 0.11 | 0.00–9.68 |
| 1.4.2. | Sport and leisure facilities | SPOLEI | 0.17 | 0.00–5.32 | 0.06 | 0.00–3.96 |
| 2.1.1. | Non-irrigated arable land | NONIRR | 65.16 | 2.51–98.98 | 70.00 | 0.30–99.44 |
| 2.1.3. | Rice fields | RICEFI | 0.26 | 0.00–18.51 | 0.00 | 0.00–0.00 |
| 2.2.1. | Vineyards | VINEYA | 1.25 | 0.00–47.85 | 0.20 | 0.00–18.30 |
| 2.2.2. | Fruit and berry plantations | FRUITS | 0.59 | 0.00–15.66 | 0.19 | 0.00–10.03 |
| 2.3.1. | Pastures | PASTUR | 8.79 | 0.00–43.81 | 2.50 | 0.00–24.17 |
| 2.4.2. | Complex cultivation | COMPLE | 2.29 | 0.00–43.69 | 5.51 | 0.00–35.38 |
| 2.4.3. | Agriculture – nat. veg. ^a | AGR NAT | 1.39 | 0.00–11.49 | 2.39 | 0.00–24.88 |
| 3.1.1. | Broad-leaved forests | BROLEA | 6.01 | 0.00–60.69 | 5.48 | 0.00–85.26 |
| 3.1.2. | Coniferous forests | CONFOR | 0.40 | 0.00–17.90 | 0.09 | 0.00–13.44 |
| 3.1.3. | Mixed forests | MIXFOR | 0.59 | 0.00–13.68 | 0.18 | 0.00–15.23 |
| 3.2.1. | Natural grasslands | NATGRA | 3.48 | 0.00–83.13 | 1.61 | 0.00–32.72 |
| 3.2.4. | Transitional woodland shrub | WOOSHR | 1.40 | 0.00–17.75 | 2.56 | 0.00–34.19 |
| 3.3.1. | Beaches dunes and sand plains* | BEASAN | 0.00 | 0.00–0.24 | 0.00 | 0.00–0.97 |
| 4.1.1. | Inland marshes | INMARS | 0.78 | 0.00–24.69 | 1.00 | 0.00–2.61 |
| 4.1.2. | Peat bogs | PEATBO | 0.07 | 0.00–13.98 | 0.00 | 0.00–0.00 |
| 5.1.1. | Water courses | WATCOU | 0.78 | 0.00–9.66 | 1.37 | 0.00–21.74 |
| 5.1.2. | Water bodies | WATBOD | 0.91 | 0.00–21.60 | 0.79 | 0.00–15.09 |

Var. code and Name (level 3) refers to the original nomenclature of the CORINE 2006. Variables indicated with an * variables left out of model building due to low variability.

^aLand principally occupied by agriculture, with significant areas of natural vegetation.

data accurate without any further adjustments. We used the data of all discovered sites between 2006 and 2009 in Hungary and in Western Romania. As red-footed falcons are non-territorial, in theory, they may have large inter-annual spatial variability in breeding distribution. Using the cumulative distribution pattern of several consecutive years inherently addresses this potential bias.

The last large-scale red-footed falcon survey in Serbia was conducted in 2000–2001 (Purger, 2008), although sporadic information on breeding distribution is available from 2009 to 2010. The latter could be used to evaluate model predictions; however, the recent survey effort made is spatially biased, hence does not reflect the true distribution. Therefore, we only used the most recent data for illustration and confirmation of result outputs instead of incorporating them into the modelling procedure (see Fig. 5b).

Land cover variables, modelling and predicted areas

We used the CORINE 2006 Land Cover GIS database (Corine Land Cover, 2006 – version 13 available from <http://www.eea.europa.eu/data-and-maps/data/clc-2006-vector->

data-version), as this is the sole reliable source of information on the country-wide habitat composition in Serbia. The CORINE 2006 has a 1:100 000 scale, the minimum mapping unit is 25 ha and the minimum width of linear elements is 100 m. The CORINE nomenclature consists of 44 different layer types describing the surface coverage. Twenty-nine of these variables were found in the modelling and predicted areas, and were used as predictors in the models (Table 1).

The fact that red-footed falcons have a 3–4 km diameter potential foraging area around a given colony (Palatitz *et al.*, 2011), and that the resolution of the environmental variables is relatively coarse, led us to choose a considerably large 10 × 10 km grid scale for modelling. As a first step, we created an arbitrary 10 × 10 km grid covering the whole Carpathian Basin. We then defined the modelling area as the smallest rectangle around the known breeding sites within Hungary and Western Romania. Where possible, a minimum of one cell distance was kept from all peripheral breeding sites to avoid edge effects (Ries *et al.*, 2004) (Fig. 1). The prediction area for Serbia was defined as the area north of the latitude of Belgrade (Fig. 1). All habitat variables were clipped with the pre-defined grid and were transformed to their relative coverage within a given cell.

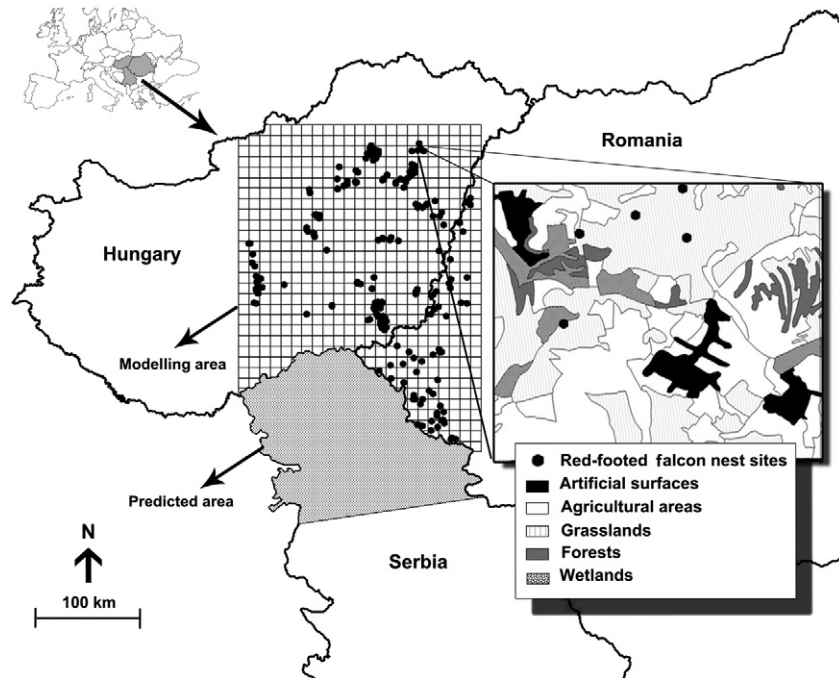


Figure 1 Hungarian and western Romanian red-footed falcon *Falco vespertinus* distribution between 2006 and 2010 on a custom made 10 × 10 km grid system, the modelling area used in the analyses and the extent of the area used for model predictions. The modelling area was assessed so that the distance of the edges is one cell from the peripheral occupied cells. The predicted area is confined by the national borders of Serbia in the north, east and west, while the latitude of Belgrade was used as the southern limit. A sample of the habitat structure as defined by the CORINE 2006 of a given UTM cell is also presented. Level 1 grouping of CORINE variables (see Table 1.) was used, albeit natural grasslands and forests are depicted separately for clear visualization. The sample habitat map of a given cell shows that the resolution of the variables used is relatively coarse. Note that the latter is only an illustration of the data, the actual distribution of red-footed falcon colonies in the depicted cell is not shown.

The modelling area consisted of a 555 grid cells, out of which 137 were identified as occupied, while the area to be predicted consisted of 277 cells.

Model building

The impressive diversity of tools lately available for creating SDMs also yields considerable variance in model predictions, thus making the right choice of modelling procedure difficult for non-statisticians (Elith *et al.*, 2006). Thuiller (2003) proposed to use a framework of statistical approaches and to evaluate their relative performance on predicting a given species distribution before using them for novel predictions.

Machine learning (ML) techniques are increasingly used as SDMs because of their flexibility, robustness against outliers, non-linear relationships and finally because they often outcompete conventional frequentist statistical models (Olden, Lawler & Poff, 2008). In the current study, we applied three ML techniques, namely feed-forward unsupervised automated neural networks (Haykin, 1994), generalized boosting models (GBM, Friedman, 2001; Friedman & Meulman, 2003; Elith, Leathwick & Hastie, 2008), and random forests (RF, Svetnik *et al.*, 2003; Elith *et al.*, 2006; Strobl *et al.*, 2007) together with classification and regression trees (Breiman *et al.*, 1984; De'ath, 2002; Hastie *et al.*, 2005) and multivariate adaptive regression splines (Friedman, 1991; Munoz, Felicísimo & Kland, 2004; Leathwick, Elith & Hastie, 2006) to describe the relationship between predictors and the distribution of red-footed falcons in the model area.

All models were applied simultaneously within the framework of biodiversity modelling (BIOMOD) (Thuiller *et al.*,

2009) in the R software (R Development Core Team, 2011). Model accuracy was tested with splitting the data 10 times, using 70% of the observations in the modelling area in each random split (Araújo *et al.*, 2005). In each run, the models were evaluated on the remaining 30% of observations. The final run used 100% of the data in case of all models.

As our red-footed falcon distribution data were considered as presence only, we used randomly selected pseudo-absence data for modelling (Phillips *et al.*, 2009). The scale of the study was set to that of the current distribution of the red-footed falcons in Hungary and in Western Romania (Fig. 1), therefore we considered that the absence of the species from a given cell is of a random distribution, and that no other niche-limiting factors may have an influence. Thus, we chose a total of random 275 cells (approximately half of the modelling area) where no red-footed falcons were recorded as pseudo-absence cells. This randomization was carried out on three independent occasions, and all models were run for all three sets with a total of 10, summing to a total of 165 models built altogether.

Selecting the best performing models was carried out using three methods: (1) the area under the relative operating characteristic curve (AUC) (James & Barbara, 1982; Hanley & McNeil, 1983), (2) Cohen's kappa (Cohen, 1960), and (3) the true skill statistic (Allouche, Tsoar & Kadmon, 2006). The BIOMOD framework also allows for model comparison through a multiple cross-validation procedure, which we have adopted to assess individual predictive power.

We used the randomization technique described by Thuiller *et al.* (2009) to conclude variable importance in case of all models. This technique uses a model-independent approach, which allows for the making of direct comparison

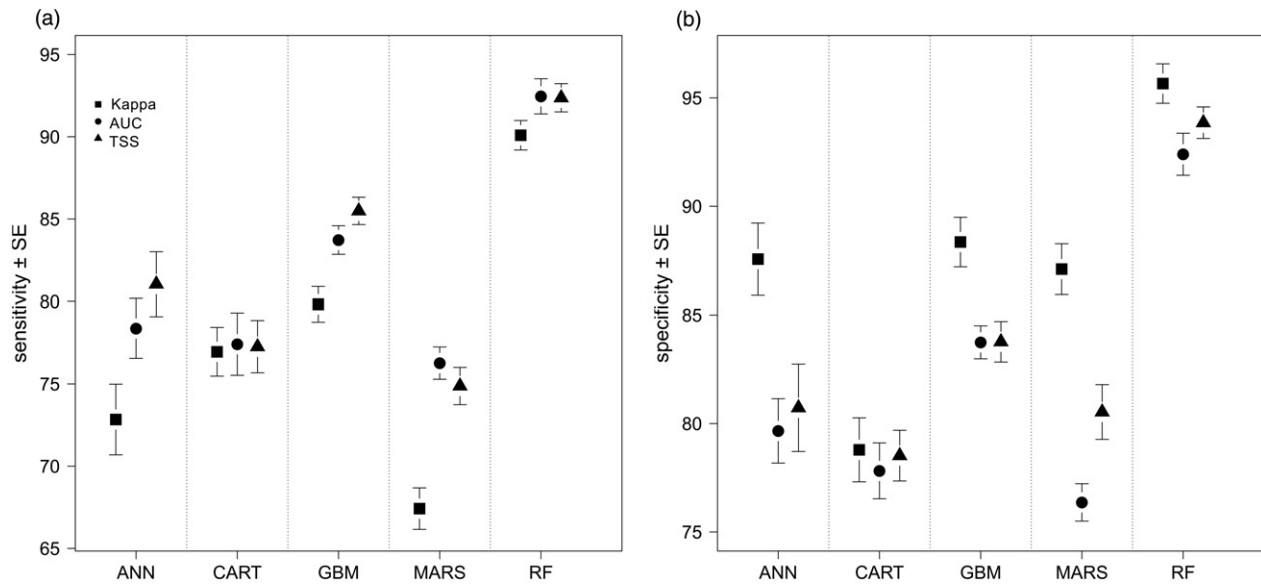


Figure 2 Sensitivity and specificity of the five model types used in the analyses averaged across all three randomization procedures and all random splittings. (a) Mean sensitivity values \pm SE while (b) mean specificity values \pm SE as assessed by the kappa, TSS and AUC, and defined thresholds ($n = 33$, in case of all models). The probability thresholds were calculated to maximize the kappa and TSS values, and to maximize the percentage of correctly classified presences and absences in case of AUC, respectively. Random forests have the highest mean values in case of both sensitivity and specificity, while GBMs show the second highest performance. ANN, automated neural networks; CART, classification trees; GBM: generalized boosted regressions; MARS, multivariate adaptive regression splines; RF, random forests; SE, standard error; TSS, true skill statistic; GBMs, generalized boosting models; AUC, area under the relative operating characteristic curve.

of variable importance across the models. We also used evaluation strips (Elith *et al.*, 2005) to determine the response curves of the three most influential variables.

Evaluating the probability of presence of nest sites in the predicted area was carried out by ensemble forecast of the two best performing models. The projected distributions were calculated with the weighted mean approach, using the cross-validation results to weigh predicted probabilities for a given grid cell (Marmion *et al.*, 2009) The sensitivity-specificity sum maximization threshold (Liu *et al.*, 2005; Jiguet, Barbet-Massin & Chevallier, 2011) was used to transform model prediction probabilities to presence/absence predictions.

Results

RFs and boosted regression trees proved to have the highest predictive power when considering the three runs with 100% of the data. These two models had the highest mean AUC values (0.98 and 0.94, respectively), and had the highest overall sensitivity and specificity values ranged across the threshold independent methods used for model evaluation (Fig. 2). Moreover, RFs were chosen as best models in 25% while GBMs in 21% of the model runs.

Both models agreed on the first three most influential variables, namely natural grasslands (NATGRA.), broad-leaved forests (BROLEA) and pastures (PASTUR) (Fig. 3). These variables were shown to have different effects on the

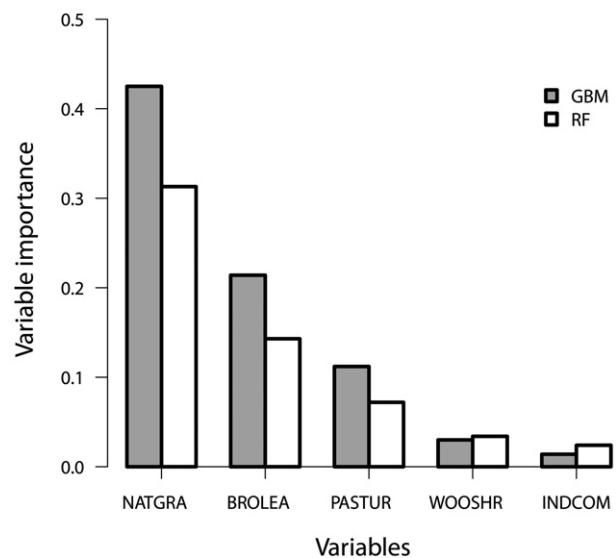


Figure 3 Variable importance measures as defined by BIOMOD for RFs and GBMs. Natural grasslands (NATGRA), broad-leaved forests (BROLEA) and pastures (PASTUR) have the highest importance measures compared with all predictors used in the models. Only the first five most relevant variables are shown. For variable codes and names, see Table 1. BIOMOD, biodiversity modelling; RF, random forests; GBM, generalized boosting models.

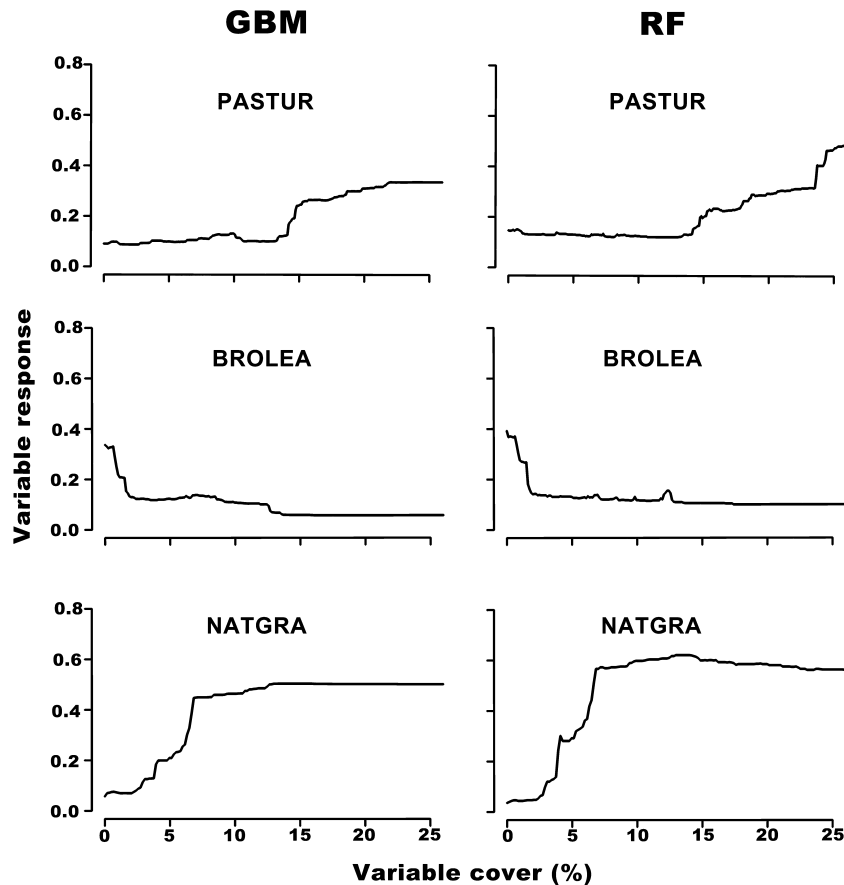


Figure 4 Evaluation strips of the three most influential variables for RFs and GBMs. Prediction units are relative likelihoods scaled from 0 to 1. Both models agreed that pastures (PASTUR) and natural grasslands (NATGRA) have positive effects on the probability of red-footed falcon *Falco vespertinus* breeding site presence while broad-leaved forests (BROLEA) have a negative impact. Pastures have a moderate positive impact in lower coverage ranges, and only increase the likelihood of red-footed falcon nest site presence in the upper (10–25%) cover ranges. Broad-leaved forest coverage has an overall negative impact; however, it seems that small 0–5% coverage has the most prominent negative effect. RF, random forests; GBM, generalized boosting models.

probability of red-footed falcon nest site presence (Fig. 4). The increase of natural grasslands and pastures had a positive effect on nest site presence, while the increase of broad-leaved forests negatively influenced the probability of nest site presence. Natural grasslands and pastures can be considered scarce and localized in both the modelling and predicted areas (Table 1); however, they are a key to distinguish occupied and unoccupied cells in the modelling area.

The ensemble projection shows that, in general, there is an east-west gradient in predicted probabilities with values on average higher in the northeast part of the predicted area (Fig. 5a). When transformed to presence data, a total of 32 cells out of the 277 cells were predicted to have red-footed falcon nest sites, which is 11.5% of the whole predicted area (Fig. 5b). Moreover, the ensemble prediction of presence was 100% accurate in finding the currently known breeding sites (nine cells) in northern Serbia.

Discussion

Landscape scale breeding habitat preference

Our modelling approach proved to be successful in describing the landscape scale habitat composition of red-footed falcon breeding sites in the modelling area. Both RFs and

GBMs agreed that natural grasslands and pastures have a considerable impact on the probability of nest site presence, corroborating our previous findings from different spatial scales (Fehérvári *et al.*, 2009).

The effect of natural grasslands and/or pastures is obvious in case of a steppe species as these are most often utilized for foraging (Palatitz *et al.*, 2011). On the other hand, the negative impact of broad-leaved forest may vary with the coverage percentage. Cells with vast forest coverage (> 15%) may be unsuitable for nesting due to low percentage of foraging habitats. The smallest mapping units of the predictors is 25 ha, thus any presence of forests in a given cell yields at least one larger homogeneous forest patch. These patches may be suitable breeding sites for goshawks *Accipiter gentilis*, a widespread raptor in both Hungary and Romania (Snow, Perrins & Gillmor, 1998). Goshawks are also one of the handful of avian predators of red-footed falcons, thus the avoidance of cells with seemingly low forest coverage may be the result of a predator-avoidance strategy in nest site selection (Fehérvári *et al.*, 2009).

Ensemble prediction of nest sites

There is increasing evidence that SDMs and their prediction in space and time may lead to erroneous conclusions if they are solely based on environmental predictors. For example,

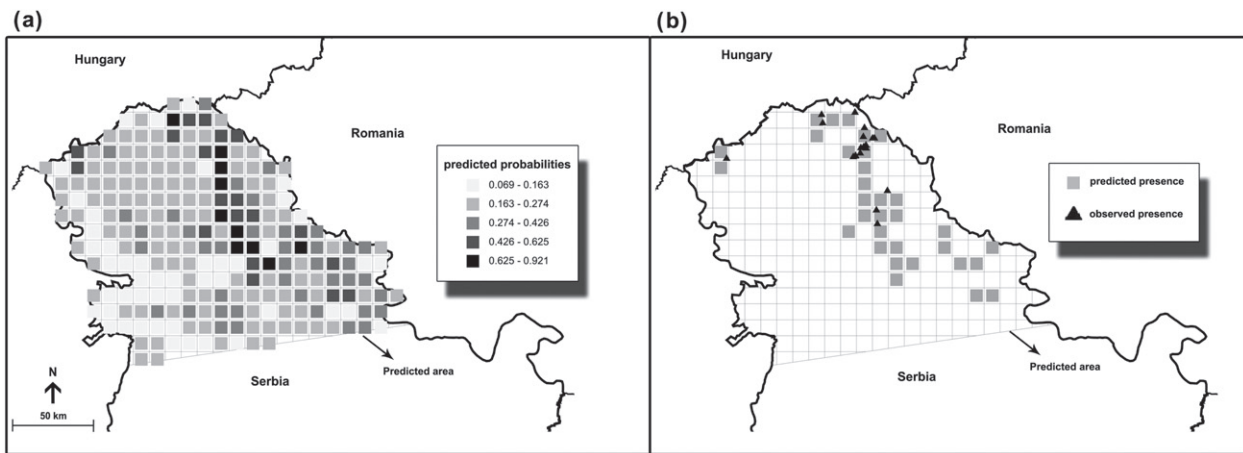


Figure 5 Ensemble model prediction of probabilities (a) and presence (b) of red-footed falcon *Falco vespertinus* nest sites in the predicted area. There is a clear east-west gradient in the predicted probabilities. Model predictions classified 100% of the known (2009–2010) breeding sites correctly.

landscape scale decisions of breeding site choice may be influenced by the presence of intraguild predators (Sergio *et al.*, 2007), or by non-environmental related behavioural traits of individuals (Penteriani, Ferrer & Delgado, 2011). Moreover, habitat quality is not solely dependent of landscape structure, local/regional conditions or land use may have a substantial part in shaping distribution patterns. For instance, red-footed falcon foraging habitat use is sensitive to vegetation cover that may be affected by environmental factors like precipitation or certain land use practices (Palatitz *et al.*, 2011). The type and timing of different land use forms, especially in case of grasslands, are governed by subsidy systems and regulations within the EU; however, Serbia does not have a coherent legislative approach. Therefore, our results are to be treated with caution, and it has to be noted that they only resemble a possible scenario of potential distribution. Nonetheless, they present a solid backbone for designing and elaborating future conservation measures and allow to fundament important knowledge concerning distribution shaping effects.

The spatial pattern of predicted nest site presence probabilities (Fig. 5a) shows a southwest-northeast gradient. This corroborates the lack of reported data from Srem County (Serbia), in the past decade (Purger, 2008). Once the predicted probabilities are translated to presence/absence predictions (Fig. 5b), our findings resemble the spatial distribution of breeding sites in 1990–1991 (Purger, 1996, 2008). Grasslands, the main variables defining the current distribution in the modelling area, are scarce (1.6% mean coverage in the predicted area). Therefore, even small alterations in land use practices, especially converting grasslands to arable land, may have a drastic effect on the potential distribution. The predictions indicate that in theory, the landscape scale habitat structure has not limited the distribution of these falcons in the past two decades; hence, the observed decline cannot be attributed directly to large-scale changes in land use. However, in Serbia – where only a handful of

artificial colonies exist to date – red-footed falcons are dependent on rookeries for colonial breeding. Red-footed falcons are known to breed in rookeries found in the vicinity of, or in small villages in Romania, and Serbia. However, this is less likely in Hungary, probably because the traditionally widespread village pastures have been converted to arable land and/or industrial area. Presumably, urbanized rookeries in Hungary are often at a greater distance from the suitable foraging habitats. The distance of foraging grounds, thus the probability of finding prey within a unit of time, is thought to influence the viability of falcon colonies (Rodriguez, Johst & Bustamante, 2006). Rookeries may show a similar trend to urbanization in Serbia, thus the lack of rookeries may be one of the key factors to the observed red-footed falcon population decline. We, therefore, propose to survey predicted cells for red-footed falcon presence and for rookeries. We hypothesize that these areas are – at least on a landscape scale – suitable for red-footed falcon breeding, and that nest site shortage rather than altered land use will be the primary cause of the lack of red-footed falcon breeding.

If proven true, the absence of nest sites can be managed with artificial colonies. However, the provisioning of nest boxes is only the first albeit probably a crucial step in conservation, and there are a number of limitations to consider. In general, Central European landscapes with high grassland proportions are scarce in trees suitable for rookeries and for installing nest boxes. Our experience shows that the chosen trees for the artificial colonies have to fulfil at least three important criteria: (1) they have to be in the proximity of the foraging areas, (2) they have to be in suitable condition to be able to hold nest boxes for at least a decade, and (3) they have to be less prone to illegal felling, a generally growing problem within the region (e.g. Purger, 2008).

None of the above-mentioned factors were included in our modelling approach as there are no reliable, spatially defined data available for the predicted area. These aspects

can only be assessed with regular visits, which have high labour and travel costs. Our results allow authorities and local non-governmental organizations to spatially weight conservation efforts focusing on areas with high predicted probabilities.

Serbia is a likely candidate for EU membership, and as such, has to fulfil nature conservation obligations like designating Natura 2000 sites. Red-footed falcons can be considered as a classic umbrella species for eastern grassland habitats as they have an eminent role in the Natura 2000 designation process (http://ec.europa.eu/environment/nature/natura2000/index_en.htm). Therefore, locating their breeding sites or creating breeding sites by providing artificial colonies with the aid of our results has additional conservation and socio-economic benefits for both wildlife and the local human population.

Acknowledgements

Our study was funded by the LIFE Nature project 'Conservation of *F. vespertinus* in the Pannonian Region' (LIFE05 NAT/H/000122) and the 'Conservation Management and Animal Health Monitoring of Natura 2000 Bird Species' HU-SRB IPA CBC PROJECT (HU-SRB 0901/122/120). We are grateful for Prof Reuven Yosef and for two anonymous referees for their valuable suggestions on earlier versions of this paper.

References

- Allouche, O., Tsoar, A. & Kadmon, R. (2006). Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232.
- Araújo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005). Validation of species-climate impact models under climate change. *Glob. Change Biol.* **11**, 1504–1513.
- Bagyura, J. & Palatitz, P. (2004). *Species conservation plan, red-footed falcon (Falco vespertinus)*. Budapest: Ministry of Environment and Water.
- Bessa-Gomes, C. & Petrucci-Fonseca, F. (2003). Using artificial neural networks to assess wolf distribution patterns in Portugal. *Anim. Conserv.* **6**, 221–229.
- Breiman, L., Friedman, J., Olshen, R.A. & Stone, C.J. (1984). *Classification and decision trees*. Belmont: Wadsworth.
- Brooks, T.M., Mittermeier, R.A., da Fonseca, G.A., Gerlach, J., Hoffmann, M., Lamoreux, J.F., Mittermeier, C.G., Pilgrim, J.D. & Rodrigues, A.S. (2006). Global biodiversity conservation priorities. *Science* **313**, 58–61.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **20**, 37–46.
- Corine Land Cover. (2006). Seamless vector data – version 13 (02/2010) – Datasets – EEA. Available at <http://www.eea.europa.eu/data-and-maps/data/clc-2006-vector-data-version> (accessed 1 May 2011).
- De'ath, G. (2002). Multivariate regression trees: a new technique for modeling species-environment relationships. *Ecology* **83**, 1105–1117.
- Elith, J., Ferrier, S., Huettmann, F. & Leathwick, J. (2005). The evaluation strip: a new and robust method for plotting predicted responses from species distribution models. *Ecol. Modell.* **186**, 280–289.
- Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R. & Lehmann, A. (2006). Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151.
- Elith, J. & Leathwick, J.R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Syst.* **40**, 677–697.
- Elith, J., Leathwick, J.R. & Hastie, T. (2008). A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813.
- Engler, R., Guisan, A. & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo absence data. *J. Appl. Ecol.* **41**, 263–274.
- Fehérvári, P., Harnos, A., Neidert, D., Solt, S. & Palatitz, P. (2009). Modeling habitat selection of the red-footed falcon (*Falco vespertinus*): a possible explanation of recent changes in breeding range within Hungary. *Appl. Ecol. Environ.* **7**, 59–69.
- Friedman, J.H. (1991). Multivariate adaptive regression splines. *Ann. Stat.* **19**, 1–67.
- Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *Ann. Stat.* **19**, 1189–1232.
- Friedman, J.H. & Meulman, J.J. (2003). Multiple additive regression trees with application in epidemiology. *Stat. Med.* **22**, 1365–1381.
- Guisan, A. & Thuiller, W. (2005). Predicting species distribution: offering more than simple habitat models. *Ecol. Lett.* **8**, 993–1009.
- Hanley, J.A. & McNeil, B.J. (1983). A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* **148**, 839–843.
- Hastie, T., Tibshirani, R., Friedman, J. & Franklin, J. (2005). The elements of statistical learning: data mining, inference and prediction. *Math. Intell.* **27**, 83–85.
- Haykin, S. (1994). *Neural networks: a comprehensive foundation*. New York: Macmillan College Publishing Company.
- Heredia, B., Rose, L. & Painter, M. (1996). *Globally threatened birds in Europe: action plans*. Germany: Council of Europe Publishing. Not Avail.
- Hirzel, A.H. & Le Lay, G. (2008). Habitat suitability modelling and niche theory. *J. Appl. Ecol.* **45**, 1372–1381.
- Horváth, L. (1964). A kék vércse (*Falco vespertinus* L.) és a kis örgébics (*Lanius minor* Gm.) élettörténetének összehasonlító vizsgálata II. A fiókák kikelésétől az őszi vonulásig. *Vertebrata Hungarica*. **VI**, 13–19.

- James, A.H. & Barbara, J.M. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* **143**, 29–36.
- Jiguet, F., Barbet-Massin, M. & Chevallier, D. (2011). Predictive distribution models applied to satellite tracks: modelling the western African winter range of European migrant Black Storks *Ciconia nigra*. *J. Ornithol.* **152**, 1–8.
- Leathwick, J.R., Elith, J. & Hastie, T. (2006). Comparative performance of generalized additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecol. Modell.* **199**, 188–196.
- Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005). Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* **28**, 385–393.
- Marini, M.N., Barbet-Massin, M., Lopes, L.E. & Jiguet, F. (2009). Predicted climate-driven bird distribution changes and forecasted conservation conflicts in a neotropical savanna. *Conserv. Biol.* **23**, 1558–1567.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009). Evaluation of consensus methods in predictive species distribution modelling. *Divers. Distrib.* **15**, 59–69.
- Meyburg, B.U., Haraszthy, L., Strazds, M. & Schäffer, N. (2001). *European species action plan for lesser spotted eagle (Aquila pomarina)*. European Union action plans for eight priority species. Luxembourg: European Commission.
- Munoz, J., Felicísimo, Á.M. & Kland, R.H. (2004). Comparison of statistical methods commonly used in predictive modelling. *J. Veg. Sci.* **15**, 285–292.
- Olden, J.D., Lawler, J.J. & Poff, N.L. (2008). Machine learning methods without tears: a primer for ecologists. *Q. Rev. Biol.* **83**, 171–194.
- Olsson, O. & Rogers, D.J. (2009). Predicting the distribution of a suitable habitat for the white stork in Southern Sweden: identifying priority areas for reintroduction and habitat restoration. *Anim. Conserv.* **12**, 62–70.
- Orlowski, G. & Czapulak, A. (2007). Different extinction risks of the breeding colonies of Rooks *Corvus frugilegus* in rural and urban areas of SW Poland. *Acta Ornithol.* **42**, 145–155.
- Palatitz, P., Fehérvári, P., Solt, S. & Barov, B. (2009). European Species Action Plan for the red-footed falcon *Falco vespertinus* Linnaeus, 1766. Available at: http://ec.europa.eu/environment/nature/conservation/wildbirds/action_plans/per_species_en.htm (accessed 3 August 2012).
- Palatitz, P., Fehérvári, P., Solt, S., Kotymán, L., Neidert, D. & Harnos, A. (2011). Exploratory analyses of foraging habitat selection of the red-footed falcon (*Falco vespertinus*). *Acta Zoolog. Acad. Sci. Hung.* **57**, 255–268.
- Parviainen, M., Marmion, M., Luoto, M., Thuiller, W. & Heikkinen, R.K. (2009). Using summed individual species models and state-of-the-art modelling techniques to identify threatened plant species hotspots. *Biol. Conserv.* **142**, 2501–2509.
- Penteriani, V., Ferrer, M. & Delgado, M.M. (2011). Floater strategies and dynamics in birds, and their importance in conservation biology: towards an understanding of non-breeders in avian populations. *Anim. Conserv.* **14**, 233–241.
- 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. *Ecol. Appl.* **19**, 181–197.
- Purger, J.J. (1996). Numbers and distribution of red-footed falcon (*Falco vespertinus*) nests in Voivodina (northern Serbia). *J. Raptor Res.* **30**, 165–168.
- Purger, J.J. (2008). Numbers and distribution of red-footed falcons (*Falco vespertinus*) breeding in Voivodina (northern Serbia): a comparison between 1990–1991 and 2000–2001. *Belg. J. Zool.* **138**: 3–7.
- Purger, J.J. & Tepavcevic, A. (1999). Pattern analysis of red-footed falcon (*Falco vespertinus*) nests in the rook (*Corvus frugilegus*) colony near Torda (Voivodina, Yugoslavia), using fuzzy correspondences and entropy. *Ecol. Modell.* **117**, 91–97.
- R Development Core Team. (2011). *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Ries, L., Fletcher, R.J. Jr, Battin, J. & Sisk, T.D. (2004). Ecological responses to habitat edges: mechanisms, models, and variability explained. *Ann. Rev. Ecol. Evol. Syst.* **35**, 491–522.
- Rodríguez, C., Johst, K. & Bustamante, J. (2006). How do crop types influence breeding success in lesser kestrels through prey quality and availability? A modelling approach. *Ecology* **43**, 587–597.
- Segan, D.B., Bottrill, M.C., Baxter, P.W. & Possingham, H.P. (2010). Using conservation evidence to guide management. *Conserv. Biol.* **25**, 200–202.
- Sergio, F., Marchesi, L., Pedrini, P. & Penteriani, V. (2007). Coexistence of a generalist owl with its intraguild predator: distance-sensitive or habitat-mediated avoidance? *Anim. Behav.* **74**, 1607–1616.
- Snow, D.W., Perrins, C.M. & Gillmor, R. (1998). *The birds of the Western Palearctic*. In *Non-passerines*. Oxford: Oxford University Press.
- Strobl, C., Boulesteix, A.L., Zeileis, A. & Hothorn, T. (2007). Bias in random forest variable importance measures: illustrations, sources and a solution. *BMC Bioinformatics* **8**, 25.
- Svetnik, V., Liaw, A., Tong, C., Culbertson, J.C., Sheridan, R.P. & Feuston, B.P. (2003). Random forest: a classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **43**, 1947–1958.
- Thuiller, W. (2003). BIOMOD-optimizing predictions of species distributions and projecting potential future shifts under global change. *Glob. Change Biol.* **9**, 1353–1362.

- Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009). BIOMOD—a platform for ensemble forecasting of species distributions. *Ecography* **32**, 369–373.
- Titeux, N., Dufrene, M., Radoux, J., Hirzel, A.H. & Defourny, P. (2007). Fitness-related parameters improve presence-only distribution modelling for conservation practice: the case of the red-backed shrike. *Biol. Conserv.* **138**, 207–223.
- Végyvári, Z., Magnier, M. & Nogues, J.-B. (2001). Kék vércsék (*Falco vespertinus*) fészekválasztása és állományváltozása a vetési varjak (*Corvus frugilegus*) állományváltozásának tükrében 1995–1999 között a Hortobágyon. *Aquila* **107/108**, 9–14.
- Wilson, C.D., Roberts, D. & Reid, N. (2010). Applying species distribution modelling to identify areas of high conservation value for endangered species: a case study using *Margaritifera margaritifera* (L). *Biol. Conserv.* **144**, 821–829.
- Wilson, K.A., Carwardine, J. & Possingham, H.P. (2009). Setting conservation priorities. *Ann. N. Y. Acad. Sci.* **1162**, 237–264.
- Yosef, R. & Wineman, A. (2010). Differential stopover of blackcap (*Sylvia atricapilla*) by sex and age at Eilat, Israel. *J. Arid Environ.* **74**, 360–367.
- Zduniak, P. & Yosef, R. (2012). Crossing the desert barrier: migration ecology of the Lesser Whitethroat (*Sylvia curruca*) at Eilat, Israel. *J. Arid Environ.* **77**, 32–38.