Biological Conservation 181 (2015) 150–161

Contents lists available at ScienceDirect

Biological Conservation

journal homepage: www.elsevier.com/locate/biocon

Rapid multi-nation distribution assessment of a charismatic conservation species using open access ensemble model GIS predictions: Red panda (*Ailurus fulgens*) in the Hindu-Kush Himalaya region

Kamal Kandel [a], Falk Huettmann [b,*], Madan Krishna Suwal [a], Ganga Ram Regmi [a], Vincent Nijman [c], K.A.I. Nekaris [c], Sonam Tashi Lama [a], Arjun Thapa [d], Hari Prasad Sharma [e], Tulsi Ram Subedi [f]

[a] Global Primate Network-Nepal, GPO Box 26288, Kathmandu, Nepal
[b] -EWHALE LAB-, Inst. of Arctic Biology, Biology & Wildlife Dept., University of Alaska Fairbanks (UAF), USA
[c] Anthropology Centre for Conservation, Environment and Development, Oxford Brookes University, Oxford, UK
[d] Small Mammals Conservation and Research Foundation, Kathmandu, Nepal
[e] Central Department of Zoology, Tribhuvan University, Kathmandu, Nepal
[f] Himalayan Nature, Kathmandu, Nepal

ARTICLE INFO

Article history:
Received 24 April 2014
Received in revised form 3 October 2014
Accepted 5 October 2014
Available online 1 December 2014

Keywords:
Red panda (*Ailurus fulgens*)
Hindu-Kush Himalaya (HKH)
Open access GBIF data
Machine learning
Ensemble model GIS prediction

ABSTRACT

The red panda (*Ailurus fulgens*) is a globally threatened species living in the multi-national Hindu-Kush Himalaya (HKH) region. It has a declining population trend due to anthropogenic pressures. Human-driven climate change is expected to have substantial impacts. However, quantitative and transparent information on the ecological niche (potential as well as realized) of this species across the vast and complex eight nations of the HKH region is lacking. Such baseline information is not only crucial for identifying new populations but also for restoring locally-extinct populations, for understanding its bio-geographical evolution, as well as for prioritizing regions and an efficient management.

First we compiled, and made publicly available through an institutional repository (dSPACE), the best known 'presence only' red panda dataset with ISO compliant metadata. This was done through the International Centre for Integrated Mountain Development (ICIMOD.org) data-platform to the Global Biodiversity Information Facility (GBIF.org). We used data mining and machine learning algorithms such as high-performance commercial Classification and Regression Trees, Random Forest, TreeNet, and Multivariate Adaptive Regression Splines implementations. We averaged all these Geographic Information System (GIS) models for the first produced ensemble model for this species in the HKH region.

Our predictive model is the first of its kind and allows to assess the red panda distribution based on empirical open access data, latest methods and the major signals and drivers of the ecological niche. It allows to assess and fine-tune earlier habitat area estimates. Our models promote 'best professional practices'. It can readily be used by the red panda Recovery Team, the red panda Action Plan, etc. because they are robust, transparent, publicly available, fit for use, and have a good accuracy, as judged by several independent assessment metrics (Receiver Operating Characteristics (ROC-AUC) curves, expert opinion, assessed by known absence regions, 95% confidence intervals and new field data).

1. Introduction

The distribution of a species is an inherent part of its ecology (Krebs, 2009); such information is essential to know for a successful conservation-dependent species' management (Braun, 2005). The prediction of species' distribution is central to diverse applications in ecology, evolution and conservation science (Guisan and Zimmermann, 2000; Elith et al., 2006) and when the precautionary principle is of increasing importance (Cushman and Huettmann, 2010; Drew et al., 2011; Huettmann, 2012). The potential

* Corresponding author at: 419 Irving 1, University of Alaska-Fairbanks, 99775 Fairbanks, AK, USA.
E-mail addresses: kandel.kamal82@gmail.com (K. Kandel), fhuettmann@alaska.edu (F. Huettmann), madanksuwal@gmail.com (M.K. Suwal), regmigr1978@gmail.com (G. Ram Regmi), vnijman@brookes.ac.uk (V. Nijman), anekaris@brookes.ac.uk (K.A.I. Nekaris), sonam.tac@gmail.com (S.T. Lama), thapa.nature@gmail.com (A. Thapa), himalayanhari@gmail.com (H.P. Sharma), tulsi.biologist@gmail.com (T.R. Subedi).

http://dx.doi.org/10.1016/j.biocon.2014.10.007
0006-3207/

distribution and quantified ecological niche describes where conditions are suitable for an occurrence and for supporting the survival of the species (Phillips et al., 2004; Cushman and Huettmann, 2010; Drew et al., 2011). Distribution models can describe such ecological niches and establish a quantitative relationship between the relative occurrence of the species and their bio-physical and environmental conditions in the landscape (Guisan and Zimmermann, 2000; Elith et al., 2006; Phillips et al., 2004, 2006; Phillips and Dudik, 2008). This can even be the case when the survey effort in a landscape is somewhat incomplete and uneven (Kadmon et al., 2004) and when data are 'messy' (Craig and Huettmann, 2008). Such models provide essential information for monitoring and restoration of declining populations in the natural habitat. They help to establish core conservation areas and biological corridors for instance (Wilson et al., 2009). In addition to just quantifying the ecological niche of a species, species distribution models (SDMs) are also widely used in many other ecological applications (Guisan and Zimmermann, 2000; Drew et al., 2011). For example, they can be used for testing biogeographical, ecological and evolutionary hypotheses (Leathwick, 1998; Anderson et al., 2002; Graham et al., 2004), for assessing species invasion and proliferation (Beerling et al., 1995; Peterson, 2003), for modeling species assemblages from individual species predictions (Leathwick et al., 1996; Guisan and Theurillat, 2000; Ferrier et al., 2002) and for future potential distribution (forecasting) (Nielsen et al., 2008; Murphy et al., 2012a,b). They can also be used for locating biogeographic regions (Murphy et al., 2012a,b) and for improving the calculation of ecological distances between patches in landscape meta-population dynamics and gene flow models (Guisan and Thuiller, 2005; Cushman and Huettmann, 2010). SDMs are nowadays a standard for virtually all conservation management projects; it is also used by IUCN (www.iucnredlist.org) to map the species global distribution range. By now, ignoring species model predictions and for species management must basically present a professional oversight (Braun, 2005). As a 'best professional practice' (Drew et al., 2011; Huettmann, 2012) they should be done for any species of conservation concern. This is even more so when most data and tools required for predictive modeling are easily and cheaply available now (e.g. Ohse et al., 2009; Huettmann et al., 2011) and rank among the very best software tools on the (commercial) market.

Therefore, our work goes beyond classic SDMs because (a) we use a commercial algorithm implementation that is among the best known data mining methods, (b) we have compiled extensive datasets not used before, (c) we use open access data with metadata, (d) we also make our models publicly available, and (e) we employ many metrics to convince on the validity and predictive performance (statistical and expert review based on latest information). All of this comes as a multivariate package and as a single powerful workflow.

The red panda (*Ailurus fulgens*; Taxonomic Serial Number TSN: 621846) is still a little known Himalayan member of Carnivora (Glatston, 2010) that has adapted to a herbivorous diet in a humid environment (Yonzon, 1989; Pradhan et al., 2001; Choudhary, 2001, Wei et al., 1999a,b; Glatston, 2010). Some authors now consider it to be two subspecies: *A.f. fulgens* (in the western region such as Nepal and adjacent Buthan and Sikkhim), and *A.f. styani* mostly in China (Groves, 2011). Red panda is generally described to inhabit multiple vegetation types, including evergreen, mixed broad-leaf, deciduous and coniferous forests. The species is documented to live between an elevation range of 2000–4000 m above sea level (asl), which is associated with dense bamboo thicket understories (Yonzon, 1989; Wei et al., 1999a,b; Pradhan et al., 2001; Choudhary, 2001; Glatston, 2010). It prefers steep north- and eastward facing slopes due to the associated rainfall regime, and because of its vegetation/food (Yonzon, 1989; Wei et al., 1999a,b; Wei and Zhang, 2011). In the wild, red panda are difficult to detect, track, observe and study though (Wei et al., 1999a,b). This is due to their elusive nature, arboreal habit and the sole occurrence in remote and inaccessible areas (Roberts and Gittleman, 1984). Conflicts with humans are rampant (e.g. Fox et al., 2002; Glatston and Gebaur, 2011), but no reliable population estimates exist (Ziegler et al., 2010). The red panda is listed as 'Vulnerable' because its population is estimated <10,000 mature individuals with a continuing decline of greater than 10% over the next 3 generations (estimated at 30 years). Geo-referenced survey data on red panda distribution with a research design for such cryptic and rare species are often sparse and clustered. It is here where predictive modeling can help overcome sampling problems and generate reliable, consistent and transparently derived estimates over wider areas (*sensu* Drew et al., 2011).

Since proven so successfully elsewhere (Guisan and Zimmermann, 2000; Cushman and Huettmann, 2010; Drew et al., 2011), a consistent model-based prediction of red panda distribution across the Hindu-Kush Himalaya (HKH) region would contribute in documenting its spatial distribution, habitat preferences and as a baseline aiding management practices. These are all details that are highly needed and beneficial for the red panda Action Plan and the Recovery Team and PVAs for instance (Jnawali et al., 2012).

As a best practice role model, here we attempt to develop an immediate and rapid assessment of quantitative and accurate information on red panda distributions in eight nations of the HKH region and for a science-based conservation management for this globally threatened species. To bring the conservation community closer together, we aim that this work can be used and assessed by others within the global community for efficient red panda conservation. While this can only be one step for good progress, it should still serve as a powerful template for similar species, habitats and applications (see Ohse et al., 2009; Huettmann et al., 2011). If successful, we think this can set a suitable standard for other species management schemes to build on for progress and for a more international scale even.

2. Materials and methods

2.1. Study area, field data collection and metadata

The Hindu-Kush Himalaya (HKH) region is a significant global landscape (Ives, 2012). It consists of eight vast and diverse Asian nations: Afghanistan, Pakistan, China, Nepal, India, Bhutan, Bangladesh and Myanmar; it comprises of 3,441,719 km^2 (Table 1, Fig. 1). The HKH region is tightly linked with most global economic and ecological processes, e.g. for climate, as well as a resource provider to India, China, and partly to Russia and the western world (Winters and Yusuf, 2007; Huettmann, 2012). Glacier-fed river systems cover the HKH nations and include major rivers such as Ganges, Yangtsekiang, Yellow River and Brahmaputra. Several oceans also rely for their freshwater inflow on the HKH region. In addition to its complex ecological, human and religious set up (Huettmann, 2012), this vast region is infamous for its difficulty to survey. Rugged mountainous terrain covered by dense forest with often dense bamboo undergrowth creates difficulties to observe red panda directly in the field (Wei et al., 1999a,b; Pradhan et al., 2001; Zhang et al., 2004). Therefore, feces (latrine sites) of red panda can be considered as the most reliable indirect evidence of red panda presence because animals usually defecate a midden of feces at the feeding site (Wei et al., 1999a,b; Jnawali et al., 2012; Pradhan et al., 2001). Tracks however are usually too soft to be found on the ground. Consequently, most of our 'presence only' data come from geo-referenced fecal midden ($n = 1093$), with a smaller proportion consisting of direct sightings ($n = 27$) in Nepal (see data in the

Table 1
The eight Hindu-Kush Himalaya (HKH) nations and their areas, the predicted red panda distribution within, and the outcome of the predictive Random Forest model. It is noteworthy that our expert opinion and the available literature are all in general agreement with the predicted presence or pseudo-absence of red pandas in the respective nations.

Nations	Total national area in km^2	Part of the nation falling in HKH as km^2 (proportion of total country area)	Predicted red panda area in km^2 (proportion of area in HKH)		Predicted presence/absence from Random Forest model outputs
			Based on expert cutoff (% of total predicted)	Based on 95 percentile threshold (% of total predicted)	
China	9,596,960	1,647,725 (17%)	8500 (26.1%)	13,100 (27.8%)	Presence
Nepal	147,181	147,181 (100%)	17,400 (53.4%)	22,400 (47.6%)	Presence
India	3,287,263	482,920 (14%)	3200 (9.8%)	5700 (12.1%)	Presence
Myanmar	677,000	317,640 (47%)	2900 (8.9%)	5000 (10.6%)	Presence
Bhutan	38,394	38,394 (100%)	600 (1.8%)	900 (1.9%)	Presence
Pakistan	796,096	404,195 (51%)	0 (0%)	0 (0%)	Absence
Bangladesh	147,570	13,189 (9%)	0 (0%)	0 (0%)	Absence
Afghanistan	652,000	390,475 (60%)	0 (0%)	0 (0%)	Absence
Total	15,342,464	3,441,719 (22%)	32,600 (100%)	47,100 (100%)	

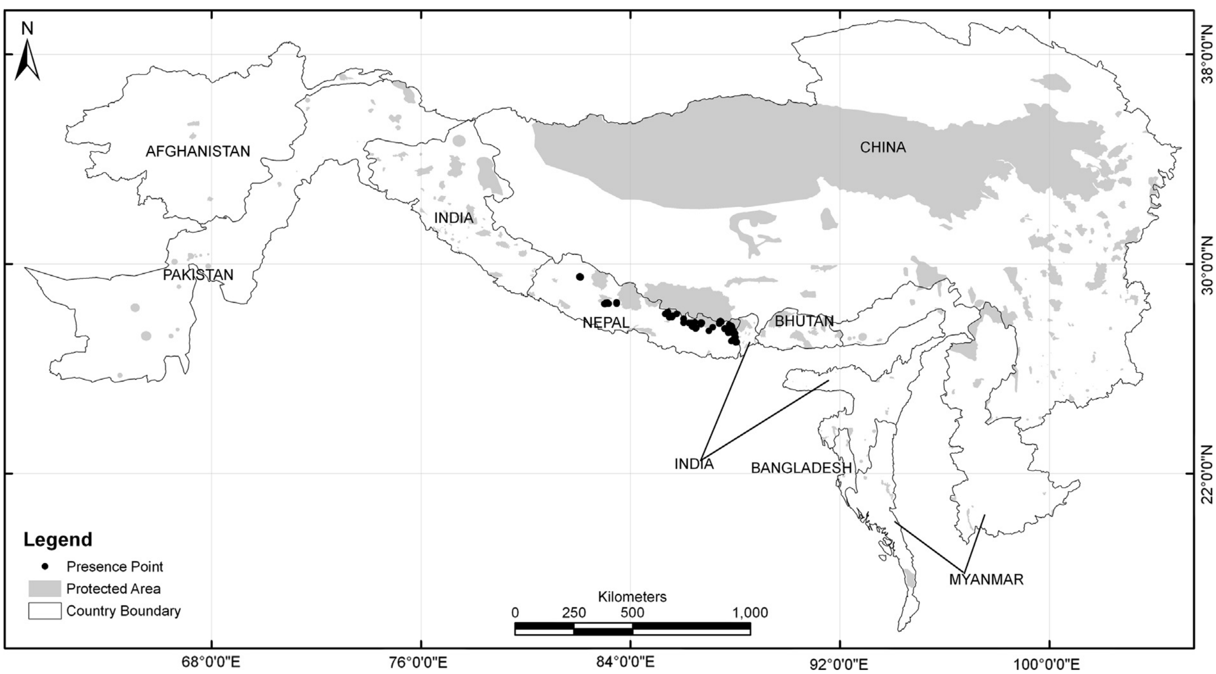

Fig. 1. Map of study area, protected areas and best publicly available 'presence only' points for red panda in the Hindu Kush-Himalaya (HKH; from ICIMOD.org).

institutional repository https://scholarworks.alaska.edu/handle/11122/1012 as well as metadata in https://scholarworks.alaska.edu/handle/11122/1012 and published with NBII online http://goo.gl/vUWJZ4).

The overall work flow for our predictive model and its field work and digital data compilation is presented in Fig. 2. The Department of National Parks and Wildlife Conservation of Government of Nepal granted research permission to carry out surveys. Searching along elevational contour lines, as well as transect surveys and surveys along small forest trails and opportunistic sightings of species presence are the most common data collection techniques in the mountain topography for the elusive and shy red panda (Jnawali et al., 2012; Pradhan et al., 2001). While maintaining all relevant data details, here we pool sightings from all survey techniques to obtain the best possible 'presence only' record dataset for the public record as a legacy. The geo-referencing of these data was done with an eTREX 12 channel GPS (Global Positioning System) and we recorded decimal degrees (5 decimals); the WGS 84 geographic datum was used. Most records include altitude information in meters measured by an altimeter in the field. The field data were collected between 2007 and 2011 from various parts of Nepal by the authors (for more details see Metadata or contact authors; see Fig. 1 for data point distributions).

2.2. Compilation of public and open access spatial (GIS and environmental) data

We compiled bioclimatic data (derived from the monthly temperature and precipitation values) (www.worldclim.org; 30 arcsecond pixels for bio1 to bio19 (Hijmans et al., 2005) and masked them with the boundary of HKH defined by the International Centre for Integrated Mountain Development (ICIMOD) in a Geographic Information System (ArcGIS 10, ESRI, 2011) freely available to us as a campus license. Similarly, we downloaded a Digital Elevation Model (DEM) with a 30 m resolution for HKH from the USGS site (http://gdex.cr.usgs.gov/gdex/) and masked it with the study area boundary. We fixed a small set of DEM errors (e.g. a few coastline pixels) with a DEM created from 20 m to 40 m contours (USGS, 2004). We used the corrected final DEM to prepare

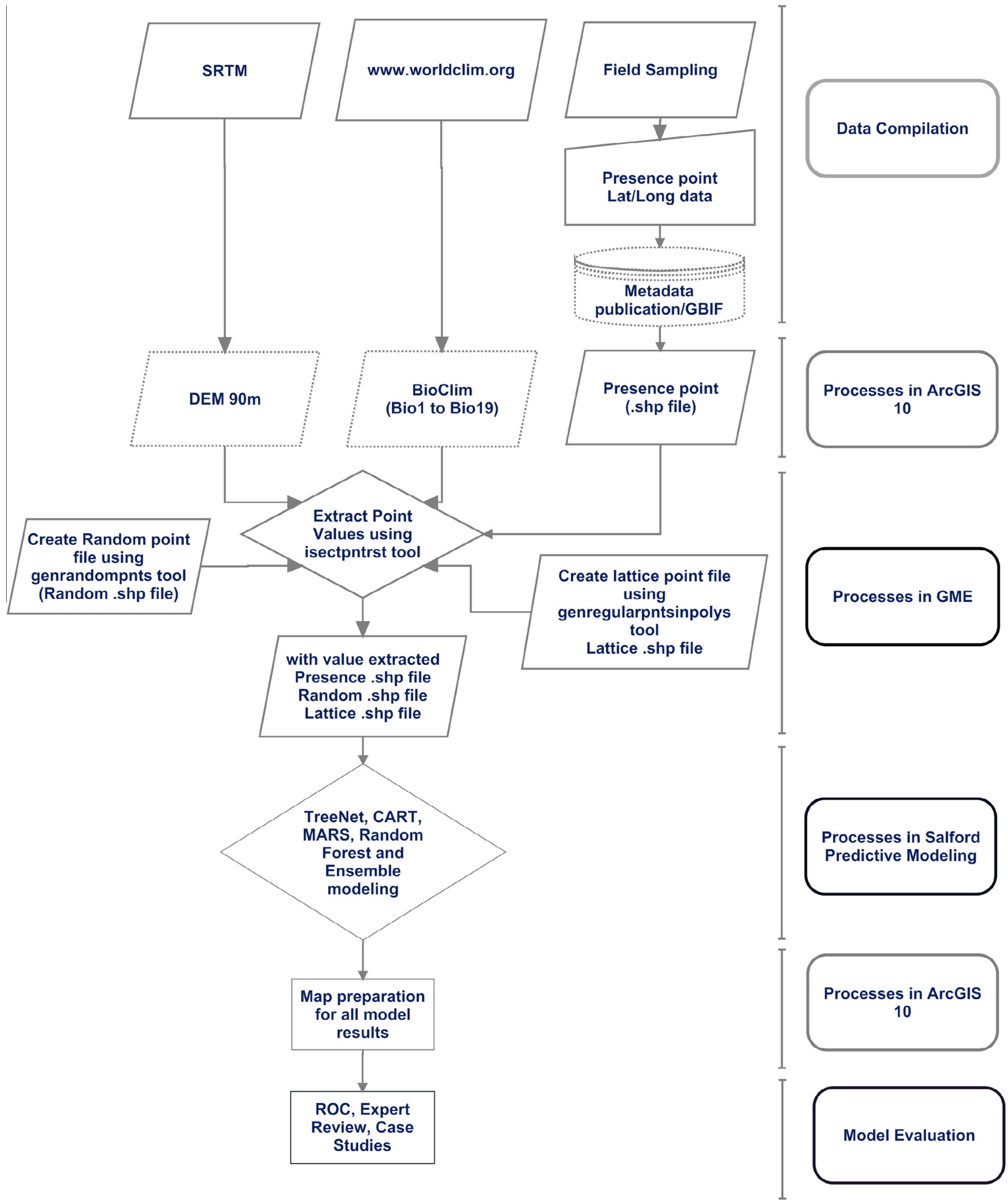

Fig. 2. The work flow for the predictive red panda model and its field work and digital data compilation and analysis.

an aspect map (circular degrees) and a slope map (degrees steepness) in ArcGIS 10.

2.3. Preparation and processing of spatial data

As a next step, we used the publicly available Geospatial Modeling Environment (GME) software (http://www.spatialecology.com/gme/) to generate a total of 62,880 pseudo-absence (random) points in the HKH region for habitat modeling using the genrandompnts (Generate Random Points). For a representative landscape sampling we used a 10 km distance rule to avoid any clumping in the random points for representative background sampling. Initially, we also used the pseudo-absences only for Nepal as a test but this approach was dismissed because it did not generalize well and for the entire HKH region. Next, a total of 79,597 lattice points (=a set of regular 10 km-spaced grid points) were generated for the HKH study area using Generate Regular Points in Polygons in order to create the 'prediction to' data. We then used the GME tool isectpntrst to extract (='drill down') the underlying point values for the different GIS layers. We extracted underlying values for 1119 presence points, 62,880 random points, and for 42,435 (predict to) lattice points from all predictor layers.

2.4. Model building process, predictions and ecological distribution niche assessment

We built our ensemble model using the individual algorithms of Random Forest (RF), TreeNet (TN), Multiple Adaptive Regression Splines (MARS) and Classification and Regression Trees (CART). These algorithms are proven to be very powerful for many applications and they are widely used (Elith et al., 2006; Magness et al., 2008; Huettmann et al., 2011), including in major commercial and global security sectors. While several implementations exist for such algorithms, e.g. in R, here we used the Salford Systems Ltd version and which is known to perform extremely well (see Herrick, 2013 for a comparison) and as a global market-leader for those algorithms. We used this high-end software for data mining in a highly relevant but complex international conservation context because (a) it offers a realistic and rapid approach to complex and 'messy data, (b) it is widely available and as a free trial version (time-limited) to interested students and scholars, (c) it can easily be linked with GIS (via tables in txt, ASCII format), (d) it is created and highly endorsed by statistical and mathematical experts in the field, (e) it is able to extract the major signals from data even when data are skewed and imbalanced, and (f) it usually performs superior to the R implementations (Herrick, 2013) and when compared with other versions and implementations. The software we used creates robust models even when using a small and/or poor quality number of 'presence only' data (Herrick, 2013, see also Hernandez et al., 2006 for a general comparison elsewhere) using 'balanced' option. RF and TN can be considered an ensemble in itself (that's because they employ bagging and stochastic gradient boosting employing hundreds of trees). It is well-established that in niche modeling, and as far as the potential niche is concerned, incomplete and somewhat biased sampling (presence and random habitat) is usually not a problem for the inference and results (Kadmon et al., 2004; Drew et al., 2011; Tessarolo et al., 2014). Also, using Worldclim data in concert with elevation, slope and aspect is not a problem then when such advanced machine learning software is used (rather than MaxEnt and R implementations such as in BioMod http://r-forge.r-project.org/projects/biomod/). It can handle such data, interactions, stopping rules, weighting and complexities (Craig and Huettmann, 2008) (see http://www.salford-systems.com/ and (Drew et al., 2011) for details on bagging (out of bag OOB sampling) and stochastic gradient boosting algorithms we used). Our data are distributed over wide areas and thus generally lack high resolution autocorrelation. Further, the tree-based algorithms we use and for the ensemble overall are rather robust against autocorrelation also (Cushman and Huettmann, 2010; Drew et al., 2011). Many approaches exist to obtain and create best and most suitable ensemble models (Araujo and New, 2007; Jones-Farrand et al., 2011). Here we created and tested two versions of ensemble models from the individual models. The first one used the average from indices obtained from each model (as done in Hardy et al. (2011)). The average was derived within the ArcGIS attribute table and its columns using 'Field Calculator' based on all predicted individual relative occurrence indices (ROI) for the algorithms used. For a test and comparison, we then created a second ensemble model, which forced all indices into 0 (pseudo-absence) and 1 (confirmed presence), and which was then also averaged (for methods and formulae, see https://scholarworks.alaska.edu/handle/11122/2496). For area estimation of red panda habitat in the HKH study area, we counted for each nation the number of lattice points with an Ensemble ROI value above the threshold (cut-off), and the number counted was multiplied by $10 * 10\,km^2$ because they were created with a 10 km-spaced distance between lattice points.

We believe that this approach can help to obtain more meaningful results because (i) it is based on empirical data, (ii) it uses a consistent algorithm, and (iii) it corrects through the standardization of the indices for contradicting averages (in machine learning, the resulting response variables are not automatically and symmetrically distributed between 0 and 1 (Breiman, 2001a,b; Cutler et al., 2007). Ensembles are a relatively new approach, and with many unexplored options and settings still (Araujo and New, 2007; Hardy et al., 2011). Here we started a first model approach, but we make all data available for each model algorithm to other users who may want to apply them for a wider assessment and in other combinations and optimizations.

Finally, our models got assessed for validity and prediction quality in five ways (similar done than Elith et al., 2006; Huettmann et al., 2011): as a first metric we used the usual ROC-Area under the Curve (AUC) metric, as commonly done (Pearce and Ferrier, 2000). Using ROC and with unconfirmed absence (contaminated data) can though potentially result into a somewhat flawed assessment (usually still underestimating the performance because the pseudo-absences could include some presences and blurs the presence–absence signal). Our spatial approach, and where we include the entire HKH region with known national absence locations (e.g. for HKH nations like Afghanistan and Pakistan) 'to predict to' should buffer against that. It performs like a natural experiment in space for entire nations within the HKH study area confronting our model with reality. It allows to assess the predictions for areas where red panda is not known to occur (our second performance metric). Third, we used some expert opinions and known red panda distributions on the ground (Drew and Perera, 2011). Fourth, we used a 95% confidence interval of the predictions and compared it with the expert knowledge and whether our maps remain valid for a trend. Last and fifth, we were able to obtain brand new field data from co-workers and confronted them with our model also for a test.

3. Results

3.1. Model performance

Based on empirical data, here we present for the first time a rapid assessment of an open access spatial predictive ensemble model for a globally threatened species: the red panda. All obtained data (https://scholarworks.alaska.edu/handle/11122/2496), models (https://scholarworks.alaska.edu/handle/11122/2496) and diagnostics (see ROC curves below and https://scholarworks.alaska.edu/handle/11122/2496) can be used for immediate use in the HKH nations, or can be tested, assessed and fine-tuned further. We were able to compile the best available red panda data and describe the ecological niche for this species in the HKH region in quantitative terms, and subsequently will be able to provide assessment metrics. Except for MARS, all models within the ensemble performed rather well on these data, e.g. a ROC over 95%. The tree-based models, Random Forests and CART, are among the leading algorithms for red panda distribution predictions, followed by TreeNet; their single models did not differ much from the overall ensemble model prediction. However, from our assessments of the highly performing ensemble, the Random Forest is the best model prediction of the red panda, thus far (it should be repeated here that some modelers refer to Random Forest as an ensemble model in itself due to the bagging procedure it employs).

3.2. Variable importance for different models

From the four machine learning algorithms, the results for the top predictors to describe model predictions (Tables 2 and 3) generally showed that temperature variables are more important than precipitation. As a summarizing result of the Ensemble Model runs,

Table 2
Summary and meaning of variables for the ensemble models.

Predictor	Biological meaning	How often used in top 5	Mean importance value in ensemble (percent)
BIO1	Annual mean temperature	3	24.9
BIO2	Mean diurnal range (mean of monthly (max temp–min temp))	3	25.1
BIO4	Temperature seasonality (standard deviation ∗ 100)	4	48.4
BIO5	Max temperature of warmest month	2	35.6
BIO6	Min temperature of coldest month	3	20.9
BIO7	Temperature annual range (BIO5-BIO6)	4	100.0
BIO10	Mean temperature of warmest quarter	3	28.0
BIO11	Mean temperature of coldest quarter	2	16.7
BIO12	Annual precipitation	2	14.8
BIO13	Precipitation of wettest month	2	26.9
BIO15	Precipitation seasonality (coefficient of variation)	2	5.5
BIO16	Precipitation of wettest quarter	1	24.4
BIO17	Precipitation of driest quarter	2	2.8

Table 3
Ensemble model results for ROC and for variable importance and by percent (for explanations of variable see Table 2).

Variable importance rank	CART (% of importance)	MARS (% of importance)	TreeNet (% of importance)	Random Forest (% of importance)
1	BIO7 (100)	BIO7 (100)	BIO7 (100)	BIO7 (100)
2	BIO4 (99)	BIO6 (67)	BIO5 (36)	BIO5 (36)
3	BIO2 (77)	BIO1 (49)	BIO4 (35)	BIO4 (35)
4	BIO13 (66)	BIO12 (46)	BIO10 (22)	BIO10 (22)
5	BIO18 (62)	BIO13 (33)	BIO19 (9)	BIO19 (9)
ROC (variance explained)	0.99 (>98%)	0.99 (>98%)	0.99 (>98%)	0.99 (>98%)

it stands out that only a few variables from the larger WorldClim BIO set were favored. Generally, for best predictive performance no strong single predictor was chosen but instead a multivariate set of predictors was located. Predictors BIO4 and BIO7 were chosen four times and in tree-based models. Variables BIO1, BIO6 and BIO10 were also more strongly selected. Other variables occurred usually just once in the selected set. This is an interesting aspect and outcome when applying ensemble models for complex conservation applications and inference.

While not all possible predictors such as soil and vegetation were tested here, we used powerful proxy variables. All of these details point toward the notion that a connected multivariate set of climate predictors on a landscape-scale of HKH are among the drivers of red panda distribution. This is not well-modeled with singular, parsimonious linear algorithms, nor it is well described for red pandas, yet (e.g. Glatston, 2010). Mostly, the interactions are related to landscape-climate, and thus, somewhat linked with altitude (Hof et al., 2012). Altitude was not selected specifically, pointing to climate being a better and more detailed predictor than altitude *per se* for red panda conservation. This suggests that the micro-climate of the HKH region makes for a powerful driver of the red panda ecological niche and for its distribution. The coarse link between climate, vegetation and altitude is well-known (Hof et al., 2012) and should still be considered in any interpretation.

3.3. Model validation and area estimates

Prior to generalizing a model for a landscape it should be assessed for its predictive performance (Drew et al., 2011). Thus, model validation is essential for any predictive models (Pearce and Ferrier, 2000; Huettmann and Gottschalk, 2011). If not provided, model results are subsequentially of lower inferential value, just representing an untested hypothesis and a poorly substantiated claim (Huettmann and Gottschalk, 2011). While data mining and predictive modeling can always be used for hypothesis creation, ideally, they should be used and assessed first for more applications and to provide real answers and a more robust inference (Breiman, 2001a,b; Cushman and Huettmann, 2010; Drew et al., 2011). As this is the first model of its kind, we lack a good and valid high quality assessment data set for the study area yet. Such data are either not available, not made public, or in most parts, were never collected in a useful design and format. We believe that fixing this situation would make for great progress in future conservation research for this species and with red panda Action Plans for instance (*sensu* Huettmann and Gottschalk, 2011; compare with Jnawali et al., 2012).

In our models and based on ROC values, mathematically, Random Forest and CART are among the leading model predictions (though they are relatively close to one another). Then we decided to find the threshold value to define potential and not potential areas in two different ways, by (1) expert analysis and by (2) statistical methods:

(1) Expert analysis: Maps were prepared with different cut-off values (such as above 0.1, above 0.2 and so on till above 0.9) and then got shared with some red panda experts available to us and from the literature for their assessment (see disclosure of these 8 experts in https://scholarworks.alaska.edu/handle/11122/2496). After assessing the spatial output visually, it was found that the Random Forest model performed 'best' (Fig. 3).
(2) Statistical analysis (95% confidence interval): For covering the entire study area, we first developed an Inverse Distance Weighting (IDW) raster surface for the predicted indices from all the algorithms using the IDW tool in ArcGIS. Next, the IDW surface value for the presence data set was extracted in GME (see methods for details). Then a 95 percentile value (where 95% of presence points are within the range) of the extracted values was determined and this value was used as the threshold value for partitioning the predicted index into suitable and unsuitable climatic niche (Fig. 3).

Table 1 presents how well the predictions match reality and opinion of some experts. Table 1 also includes estimates and proportions of the red panda habitat in HKH, based on the 'best' model (following the assessment of our experts). These findings were derived from the 10 km scale pixel predictions. While initially this might appear relatively coarse, they cover the entire HKH landscape and indicate the overall trend and for valleys, regions and microclimates. Such findings are the first empirical quantitative estimates for the vast HKH region, and provide us with a robust quantitative baseline for distribution, population and future climate and impact modeling (our pixel size is also finer than most IPCC models, for instance). We are also happy to report that the area estimates resulting from a cut off value from experts' opinion vs the statistical 95% method are very close to each other and that trends are almost equal (Table 4). We found that instead the cut-off method *per se*, the actual algorithm is of bigger relevance for the model prediction accuracy.

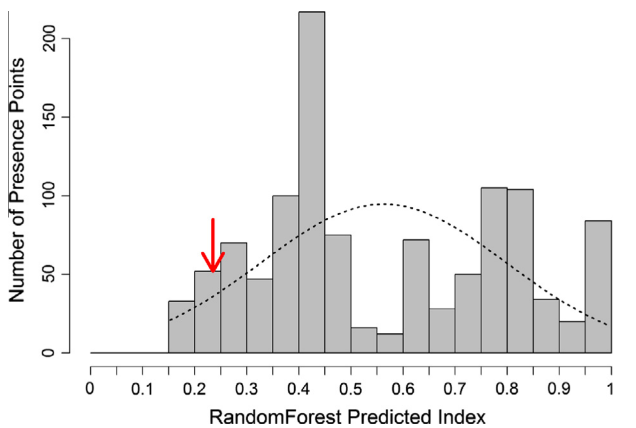

Fig. 3. Frequency histogram of predicted vs reality red panda occurrence, showing the 95 percentile threshold and the expert cut-off value (at the 87.13 percentile) for the Random Forest model (a boosted density curve is also shown).

Matching up our predictions with the known status for this species in each of the HKH nations indicates a link with known red panda hot- and cold spots. Further, this makes for a real-world test (e.g. for Afghanistan and Pakistan where red panda is absent and predicted by us to be absent). Thus, this large-scale model enables for an expert-driven quality assessment.

3.4. Response curves (partial dependence plots) for identified main predictors

Response curves can be very useful when based on stable models that are shown to generalize (=predict) well, and as it is the case here. The Random Forest GUI implementation we used provided for the best model, but it does not really provide partial dependence plots yet. Thus, we show the similarly-high performing and tree-based TreeNet plots and to assess the top predictors that were selected consistently. These figures read similar to linear regression plots and resource selection functions (Ferrier et al., 2002; Drew et al., 2011): it shows how the y-axis (response as a relative index of occurrence) relates to changes along the x-axis (a single predictor when all else is held constant and corrected for). The response shape of the two top predictors, BIO4 and BIO7 are presented in Fig. 5. Both response curves show the existence of thresholds, and they are located at 5000 units of the x-axis for BIO4, and at 245 units for BIO7 (see methods and Table 2 for the detailed description of predictors and their units). BIO4 indicates that red panda can be found where temperature seasonality (standard deviation $*$ 100) peaks between 4000 and 5000 units, and with an intermediate occurrence of the species between 2000 and 4000 units. BIO7 shows us that the red panda niche is located where Temperature Annual Range (BIO5-BIO6) lies between 18 and 24 units (degrees Celsius). In summary, this indicates that red panda are found in lower mid-temperature ranges of HKH and that it fluctuates in the medium range of the spectrum. Interpreting the response curves for these top predictors is meaningful because they have been tested across ensemble model algorithms and for the large HKH region based on thousands of pixels (which basically each make for a localized test). This specific niche description is the first of its kind for HKH. It shows in what peculiar regions red panda occur in the landscape. This finding has wider implications helping with topics like population estimates for a species otherwise difficult to detect and to monitor in the field, for its management, forecasting, prioritization and climate change questions.

4. Discussion

4.1. New data and results

We present the first publicly available 'presence only' data for the globally endangered red panda and the HKH region. These data are documented with ISO metadata and available free of charge in dSPACE (UAF) and GBIF.org via ICIMOD.org. Further, we used the free and open access WorldClim and DEM data as powerful proxies for red panda habitat preference modeling, which are also available online, robust and widely used. While we lack soil and vegetation predictors, we make the obtained ensemble model predictions also freely available as GIS files and for a further public test and scrutiny. Further, we assessed our model with five performance metrics and showed that the obtained 'best' model is accurate, robust, and generalizing well for the HKH region. It allows for rapid assessments and area estimations for instance (Ohse et al., 2009). Our results and model trends remain very consistent, even when we confront our model predictions with new data (a data set from SL with 35 sightings for the Rolpa region, and a data set with 11 sightings by KK; see details in the https://scholarworks.alaska.edu/handle/11122/2496). Noteworthy is the model prediction of known outlier populations in the Meghalaya plateau, India (Glatston, 2010). The consequent gain of our compiled and published data, the workflow overall and the resulting model analysis and products provide for a major progress regarding conservation management, transparency and habitat assessments when it comes to red panda and for the HKH region and its 8 nations overall.

4.2. Habitat area estimates for red panda

Choudhary (2001) estimated the global potential habitat of red panda at about 142,000 km^2, with China alone accounting for more than half of the area. Similarly, Wei et al. (1999a,b) estimated about 37,436 km^2 as the potential red panda habitat within China. Whereas based on our best Random Forest model prediction, and for areas predicted above a determined 0.4 threshold level of ROI (see Table 1), we estimated the current global potential habitat of red panda as 47,000 km^2 (just c. 33.17% of the area estimated by Choudhary (2001)). Within that, China only holds 13,100 km^2,

Table 4
Comparing the 95 percentile threshold and expert opinion cutoff value a for suitable and unsuitable area for red panda and their respective number of included lattice points (10 km scale).

Algorithms	95 Percentile threshold	No. of lattice points within 95% threshold	%	Cut-off value from expert model	No. of lattice points from expert model	%	% Difference between 95 percentile threshold and expert cut-off
CART	0.00000001	154	0.36	0.1	154	0.36	0
MARS	0.52523510	1726	4.06	0.6	798	1.88	2.19
Random-forest	0.23562860	471	1.10	0.4	326	0.76	0.34
TreeNet	0.11955670	299	0.70	0.1	325	0.76	−0.06
Ensemble	0.26993830	358	0.84	0.3	314	0.73	0.10

which is slightly above a quarter (27.8%) of the total suitable red panda area (see Table 1 for details; Ziegler et al., 2010). Based on best public empirical data and consistent prediction methods, these estimates help to expose, assess and revise for 2014 with scrutiny the previous assessment of the distribution of red panda and of their published numbers; many of them were not really based on transparent procedures and not using quantitative and repeatable methods and best available public data. If red panda has subspecies, then our findings would even be more worrisome because we would just be dealing with population fragments which are known to be even more problematic (Krebs, 2009; Worboys et al., 2010).

We fully acknowledge that our species model was built on Nepal data and of course with the wide HKH region somewhat undersampled. However, that is precisely what machine learning in a niche modeling context can deal with and overcome (see Elith et al., 2006; Drew et al., 2011, see also Kadmon et al., 2004 for examples); it's the strength of our highly cost-effective approach. Despite the sampling limit, our five independent model assessment metrics showed that our model performs reliably (see Tables 3 and 4 and https://scholarworks.alaska.edu/handle/11122/2496), confirming our approach with Breiman (2001a). Our model building approach is realistic and provides such a good outcome because Nepal features the complex set of the niche space for model training. Next, we used high-end machine learning algorithms for such complex data. Also, the potential niche can easily be predicted as long as similar data and habitat GIS layers exist (and as it is the case here). It remains a peculiar fact that we have a well-achieving model, but predict less red panda habitat overall, as well as in China, than previously reported. We think that this pattern is very real and provides now a study emphasize and for species needs in its range, including for China's habitats and populations.

Similarly, Choudhary (2001) estimated 25,500 km^2 (17.95%) potential red panda habitat in India but we predicted only 3200 km^2 (based on expert cut-off) and 5700 km^2 based on the 95 percentile threshold. In the same way, Choudhary (2001) estimated 13,000 km^2 (9.15%) as potential red panda habitat in Myanmar. However, our unique and advanced machine learning prediction with GIS shows that Myanmar holds only 2900 km^2 and 5000 km^2 as a predicted red panda habitat by expert cut-off and percentile threshold method respectively. So we think we have a consistent pattern and model performance here, but which is not in full agreement with published expert knowledge for area and population estimates (Ziegler et al., 2010; Jnawali et al., 2012). For red panda, no accepted population estimates really exist (Ziegler et al., 2010). Such disagreements have been described before for international species that are difficult to detect in the field. This also occurs when assessing expert knowledge and which was coined as somewhat subjective, data deficient and not repeatable by Drew et al. (2011, see also Huettmann and Gottschalk, 2011). Here we present a quantitative alternative based on best available science and make it available for more assessments.

A first small-scale GIS-based overlay analysis (Yonzon et al., 1991) estimated only 912 km^2 of area suitable for the red panda in Nepal which is only 0.61% of the total country area (147,181 km^2) whereas our Random Forest model has predicted 22,400 km^2 (56.25% of total potential red panda area in HKH). This is 15.20% of the total area of Nepal and makes Nepal a central focus for this species and specifically for *A.f. fulgens*. We can provide good support for our results and explanations. We think the reason for this finding is because for Nepal (Yonzon et al., 1991) used only three parameter area estimates: fir forest (*Abies spectabilis*), an elevation range of 3000–4000 m, and an annual precipitation >2000 mm. However, data show that the red panda has clearly been recorded beyond this elevation range, e.g., at elevations as low as 2200 m in Ilam and Panchthar districts of eastern Nepal (Kamal Kandel unpublished data) and at 2400 m in Singhalila National Park in India (Pradhan et al., 2001; see also new field data in https://scholarworks.alaska.edu/handle/11122/2496). Likewise, the red panda has been recorded in other forest types beyond *A. spectabilis*, for instance, Rhododendron forest and mixed broadleaf forest in eastern Nepal and in Singhalila National Park (Pradhan et al., 2001). Therefore, it is safe to assume that Yonzon et al. (1991) somewhat underestimated the red panda distribution in Nepal. Our predicted estimate there is c. 47% larger than Yonzon et al. (1991). And this is explained because we have used a more realistic and complete data set and large-scale approach for all of Nepal to describe red panda habitat: 19 bioclimatic variables and contemporary non-linear techniques and methods based on real empirical newly compiled field data which were not used previously to predict the potential ecological niche of the red panda in all of the HKH region. From our work here using consistent GIS pixels and methods, we believe that our estimates are more inclusive, transparent and repeatable, and thus should be more science-based, consistent and realistic for all of the red panda habitat in HKH, and as the model assessments convincingly show. We open them up for a wider assessment by a global audience in this publication.

We acknowledge that our models could predict the potential, not the realized niche (Cushman and Huettmann, 2010), and thus our model actually provides potentially some overestimates. But if this would be true, the conservation status of red panda would be even bleaker, making our model findings even more relevant and urgent. On the other side, our models are done conservative, consistent, and overall they predict well on the entire HKH landscape scale and match currently known distributions. This holds we when evaluated it by our expert opinion, with ROCs from the training data, and from what is known in the literature (we used 3 metrics) and also when confronted with new data (see https://scholarworks.alaska.edu/handle/11122/2496 for 2 datasets by the co-authors SL & KK). For instance, the new habitat patch for red panda has been identified in Jajarkot, western Nepal (pers. comm. Harihar Singh Rathour) which lies within our predicted range and confirming our model further.

At minimum, our models provide publicly available numeric and scientific estimates and trends for the distribution of red panda obtained in a scientifically sound manner (e.g. use of best-available data and high-end algorithms that are available to use freely). It helps to obtain a better foundation for a science-based management of this species and for the entire HKH region and its nations. This is achieved by providing information on where the species is predicted to occur and survive. It also provides a foundation for red panda reintroduction programs in the new areas and for setting management priorities and for more pointed study efforts. Our findings make for a common situation and when models are applied for the first time and for the conservation management of species (see Yen et al., 2004 for an example using endangered species). Either way, our models help to form new hypothesis, present new public science-based data, and invoke more ground-truthing and fieldwork toward better red panda models and for management in a readily available form for a global public peer-review.

4.3. Climate warming and next conservation management steps

The map shown in Fig. 4 provides first support for wildlife managers in the region for preparing a site-specific Red Panda Conservation Action Plan (Jnawali et al., 2012 for a Population and Habitat Viability Assessment PHVA), especially when accompanied later with tools like GIS and Marxan (Ball et al., 2009; Huettmann, 2012) in the poorly known and highly threatened HKH region

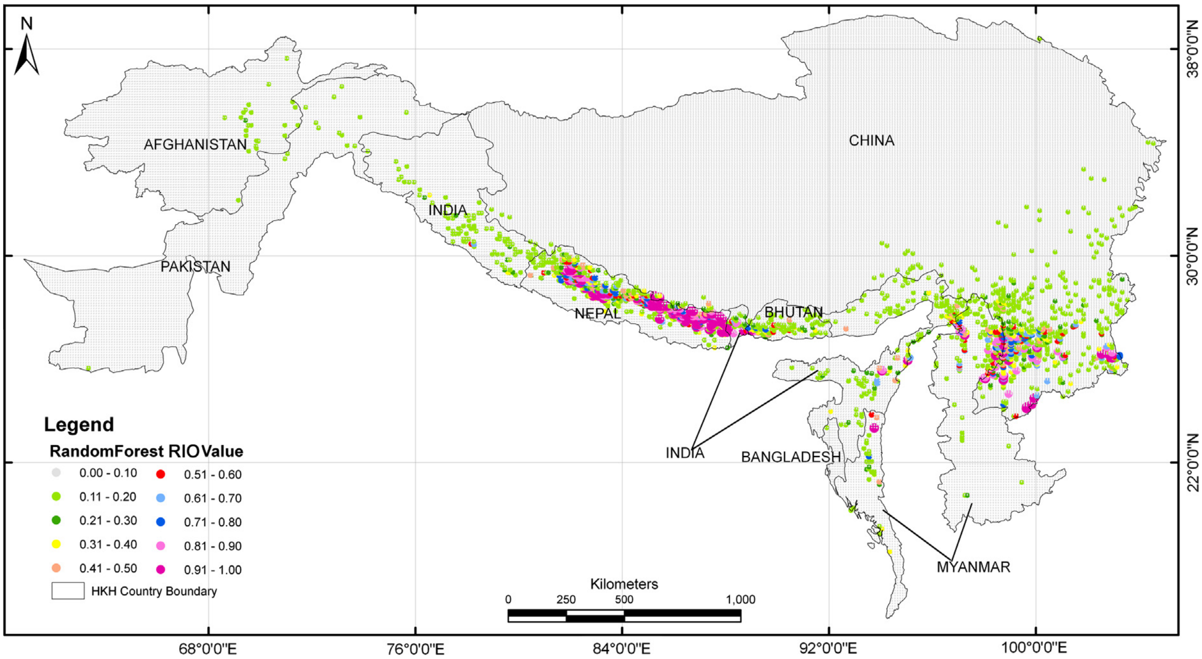

Fig. 4. Random Forest prediction map (prediction assessed as 'best' from the ensemble).

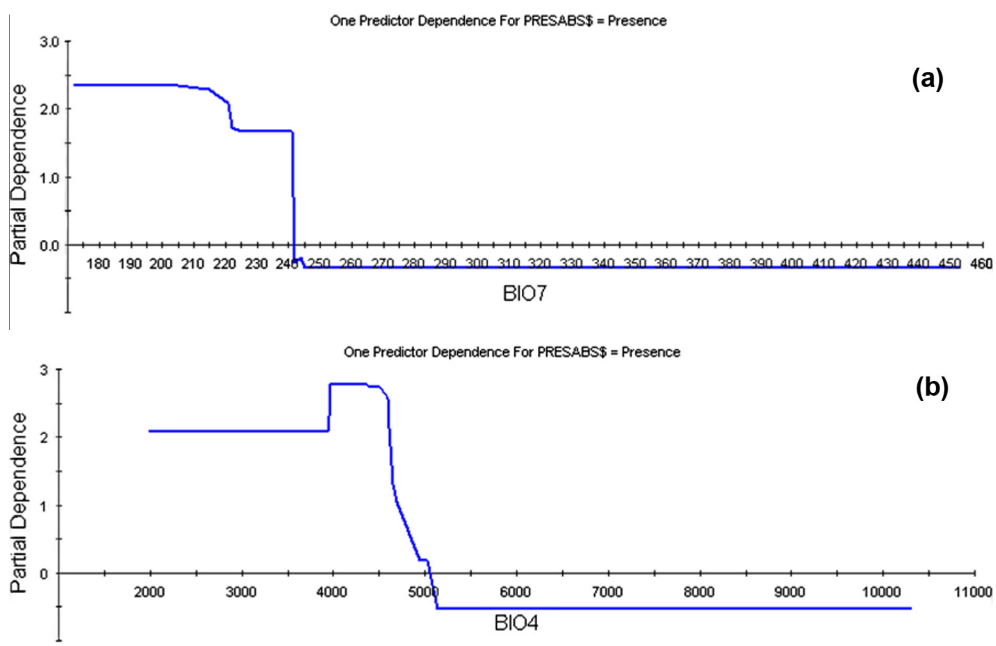

Fig. 5. Shape of the two top predictors ((a) BIO7 and (b) BIO4; WorldClim dataset) to explain red panda occurrence (for details of the predictors please see Tables 1 and 2, and text).

(Chettri et al., 2012). Based on such predictions (Magness et al., 2010; Huettmann and Gottschalk, 2011; Lawler et al., 2011), they can initiate to develop and assess further a biological corridor and its gaps (Doko et al., 2008; Worboys et al., 2010; Huettmann, 2012) between remaining fragmented habitats and protected areas; this can for instance help to keep habitats intact and to reduce potential inbreeding depression of isolated red panda subspecies and their subpopulations (Wei et al., 1999a,b; Zhang et al., 2006). Fig. 6 indicates already the urgent need to align existing protected areas with the endangered red panda hotspots; currently, there is hardly a good match yet.

We find that GIS-based high-end open access ensemble models can optimize model predictions by drawing from the benefits of each model algorithm. When using more than over 20 landscape predictors it becomes virtually untestable and for applying a traditional hypothesis in the classic concept of frequency statistics and Bayesian predictions (Oppel et al., 2009; Breiman, 2001a; Elith et al., 2006; Drew et al., 2011). This is a powerful and new approach never applied to red panda before. However, it is a known and applied technique elsewhere (see Araujo and New, 2007, for overview and Drew et al., 2011; Phillips and Dudik, 2008; Jones-Farrand et al., 2011; Hardy et al., 2011 for examples).

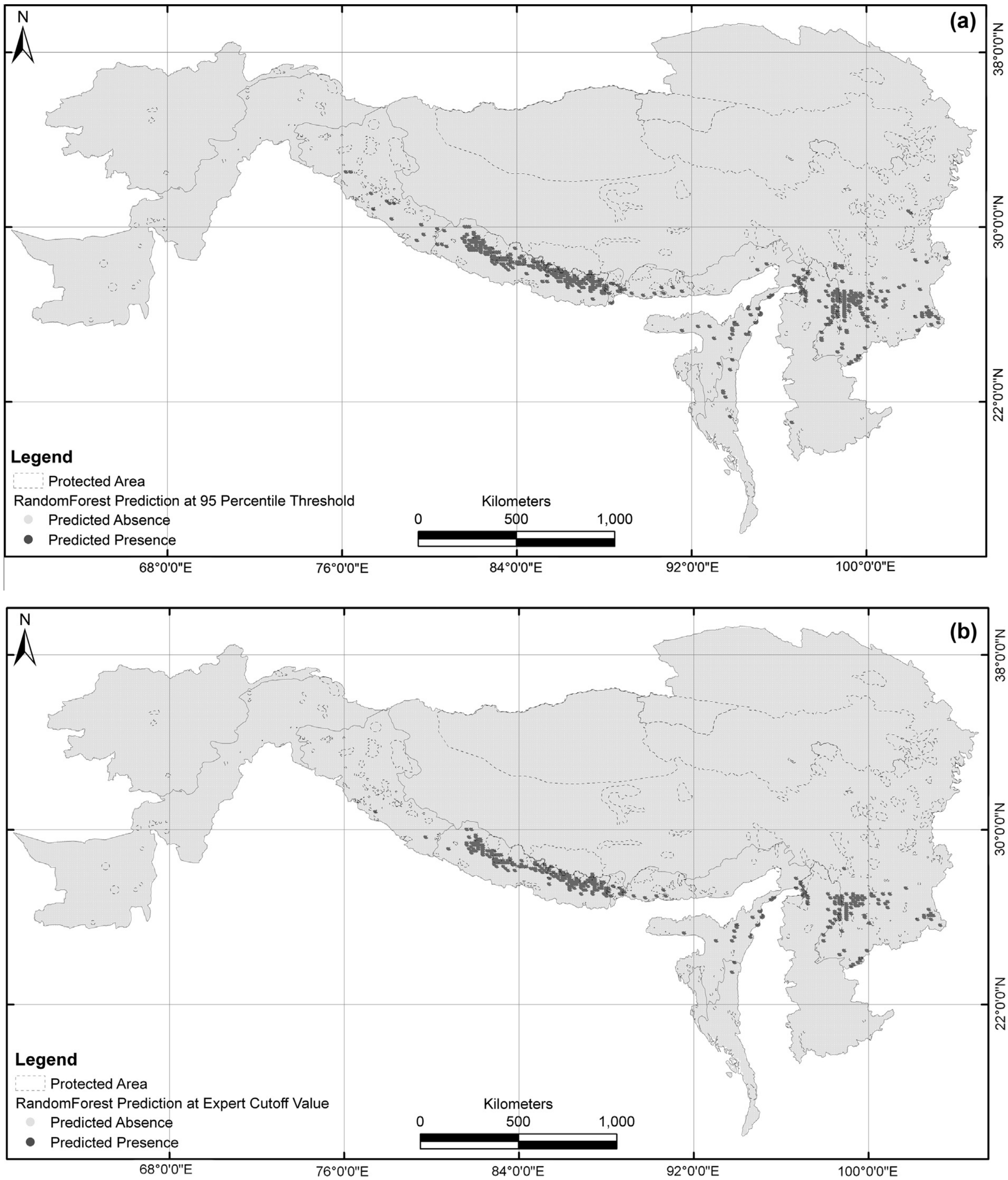

Fig. 6. Locations of currently protected areas in the study area and red panda predictions: see mismatch of protected areas vs predicted occurrences (compare with Figs. 3 and 4 for best predicted red panda occurrence models. Here we present two resulting figures but which are very similar and for the general trend we report: (a) shows the predictions based on a 95% percentile threshold; (b) shows the predictions based on an expert cut-off).

We propose more of those efforts made in regards to fine-tuning these models and techniques, also triggering more data releases, more thought, and a better policy. One of the current shortcomings still is the lack of independent ground-truthing plots and data with a research design, and to compare all available models for the entire HKH region of over 8 nations.

Moving into truly open access data, with metadata and source code published in appendices and institutional repositories as new paradigms presents a ‘best practice’, major progress and how wildlife and habitat gets managed (see Braun, 2005; Huettmann, 2005 for management, as well as Huettmann et al., 2011). This is well promoted in the peer-reviewed literature and with professional societies, e.g. Huettmann, 2005; Bluhm et al., 2010; Zuckerberg et al., 2011) but still hardly realized and funded for conservation (Costello et al., 2014). Still, such data concepts and workflows provide progress and modern approaches in dealing with natural resources in remote regions, and how they are sustainably managed with adaptive methods and in a transparent fashion.

Climate warming is likely to drive accelerating shifts in species distributions (Tse-ring et al., 2010; Carvalho et al., 2011; Araujo et al., 2004), e.g. moving along altitudes. Thus, the climate-driven habitat shifts would force the species to move out from protected areas to unprotected areas and vice versa (Carvalho et al., 2011; Forrest et al., 2012; Murphy et al., 2012a,b). Related impacts in association with anthropogenic land use and resource use is also known to affect the survival of red panda. Therefore, climate change and its possible threats have to be integrated more into efficient species conservation plans (Araujo et al., 2004; Forrest et al., 2012) in space and time. This need is rather serious for species like red panda and which are already on the 'edge', on the verge of extinction and which are sensitive to even a slight alteration in land use pattern in the Himalaya (Yonzon, 1991; Yonzon and Hunter, 1991; Yonzon et al., 1991; Zuckerberg et al., 2011; Huettmann, 2012). Noteworthy here is the rapid increase of the human population in the vicinity of protected areas of the HKH region (Ives, 2012), making for additional threats of many wildlife and their habitat, including red panda.

Taking into account all of these reality facts for the HKH region (Glatston, 2010; Ziegler et al., 2010; Ives, 2012; see Huettmann, 2012 for context and synthesis) and on the basis of these model outputs, the next research goals should be to put these models to a consistent test, ground-truth and institutionalize them with GIS for consistent updates. The red panda distribution is to be investigated under future telecoupled climate-change scenarios and for a large-scale spatial population viability assessment for this species (Regmi et al., in prep, Jnawali et al., 2012; see Oppel and Huettmann, 2010 for an example based on Random Forest). The mountainous regions of HKH provide us here with a unique experimental test ground due to altitudinal gradients and the overall multicomplexity for saving this charismatic but endangered species from irreversible gene loss and ultimate extinction.

Acknowledgements

ICIMOD, GMBA and their staff provided kindly facilitation and workshops, and the EWHALE lab and students of University of Alaska Fairbanks (UAF) provided technical support. The Department of National Parks and Wildlife Conservation of Government of Nepal granted research permission to carry out surveys. We thank Salford Systems Ltd for a free trial version of their data mining and machine learning software to the conservation research community. We also thank the Geospatial Modeling Environment (GME, H. Beier) for the free provision of software. We thank our local field assistants for their help. H.S. Baral, H. Beyer, R. Hijmans, G. Humphries, M. Lindgren, S. Linke, P. Regmi, M. Schmid, E. Spehn and L. Strecker are kindly thanked for general support and answering queries, questions and administrative post and banking tasks. Funding and in-kind support was gratefully received from Nagao Natural Environment Foundation, Japan (to G.R.R.), Oregon Zoo Foundation, U.S.A. (K.K.), The Rufford Small Grants Foundation, U.K. (T.R.S, A.T. and H.P.S.), People's Trust for Endangered Species, U.K. (H.P.S.), Chester Zoo, U.K. (A.T.) and Idea Wild, U.S.A. (G.R.R. and K.K). This is EWHALE lab publication # 125.

References

Anderson, R.P., Peterson, A.T., Gomez-Laverde, M., 2002. Using niche-based GIS modeling to test geographic predictions of competitive exclusion and competitive release in South American pocket mice. Oikos 98, 3–16.

Araujo, M.B., New, M., 2007. Ensemble forecasting of species distributions. Trends Ecol. Evol. 22k, 42–47.

Araujo, M.B., Cabeza, M., Thuiller, W., Hannah, L., Williams, P.H., 2004. Would climate change drive species out of reserves? An assessment of existing reserve-selection methods. Glob. Change Biol. 10, 1618–1626.

Ball, I.R., Possingham, H.P., Watts, M.E., 2009. Marxan and relatives: software for spatial conservation prioritization. In: Moilanen, A., Wilson, K.A., Possingham, H.P. (Eds.), Spatial Conservation Prioritization – Quantitative Methods and Computational Tools. Oxford University Press, Oxford, pp. 185–195.

Beerling, D., Huntley, J.B., Bailey, J.P., 1995. Climate and the distribution of *Fallopia japonica*: use of an introduced species to test the predictive capacity of response surface. J. Veg. Sci. 6, 269–282.

Bluhm, B.A., Watts, D., Huettmann, F., 2010. Free database availability, metadata and the internet: an example of two high latitude components of the census of marine life. In: Huettmann, F., Cushman, S.A. (Eds.), Spatial Complexity, Informatics and Wildlife Conservation. Springer, Tokyo, pp. 233–244.

Braun, C.E. (Ed.), 2005. Techniques for Wildlife Investigations and Management. The Wildlife Society (TWS), Bethesda, Maryland, USA.

Breiman, L., 2001a. Statistical modeling: the two cultures (with comments and a rejoinder by the author). Stat. Sci. 16, 199–231.

Breiman, L., 2001b. Random forests. Mach. Learn. 45 (supplement 1), 5–32.

Carvalho, S.B., Brito, J.C., Crespo, E.G., Watts, M.E., Possingham, H.P., 2011. Conservation planning under climate change: toward accounting for uncertainty in predicted species distributions to increase confidence in conservation investments in space and time. Biol. Conserv. 144, 2020–2030.

Chettri, N., Shrestha, A.B., Zhaoll, Y., Bajracharya, B., Sharma, E., Ouyang, H., 2012. Real world protection for the "third pole and its people". In: Huettmann, F. (Ed.), Protection of the Three Poles. Springer, Tokyo, pp. 113–133.

Choudhary, A.U., 2001. An overview of the status and conservation of the red panda in India, with reference to its global status. Oryx 35, 250–259.

Costello, M.J., Appeltans, W., Bailly, N., Berendsohn, W.G., de Yong, Y., Edwards, M., Froese, R., Huettmann, F., Los, W., Mees, J., 2014. Strategies for the sustainability of online open-access biodiversity databases. Biol. Conserv. 173, 155–165.

Craig, E., Huettmann, F., 2008. Using "blackbox" algorithms such as TreeNet and random forests for data-mining and for finding meaningful patterns, relationships and outliers in complex ecological data: an overview, an example using golden eagle satellite data and an outlook for a promising future. In: Wang, Hsiao-fan (Ed.), Intelligent Data Analysis: Developing New Methodologies through Pattern Discovery and Recovery. IGI Global, Hershey, PA, USA, Chapter IV.

Cushman, S.A., Huettmann, F. (Eds.), 2010. Spatial Complexity, Informatics, and Wildlife Conservation. Springer, Tokyo, Japan.

Cutler, D.R., Edwards Jr., T.C., Beard, K.H., Cutler, A., Hess, K.T., Gibson, J., Lawler, J.J., 2007. Random forests for classification in ecology. Ecology 88, 2783–2792.

Doko, T., Kooiman, A., Toxopeus, A.G., 2008. Modeling of species geographic distribution for assessing present needs for the ecological networks. In: ISPRS 2008. Proceedings of the XXIth ISPRS Congress, ISPRS, Beijing, China, pp. 267–276.

Drew, C.A., Perera, A.H., 2011. Expert knowledge as a basis for landscape ecological predictive models. In: Drew, C.A., Wiersma, Y., Huettmann, F.H. (Eds.), Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications. Springer, New York, NY, Chapter 12.

Drew, C.A., Wiersma, Y.F., Huettmann, F. (Eds.), 2011. Predictive Species and Habitat Modelling in Landscape Ecology, first ed. Springer, Tokyo, Japan.

Elith, J. et al., 2006. Novel methods improve prediction of species' distributions from occurrence data. Ecography 29, 129–151.

ESRI, 2011. ArcGIS Desktop: Release 10. Environmental Systems Research Institute, Redlands, CA.

Ferrier, S., Drielsma, M., Manion, G., Watson, G., 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in north-east New South Wales. II. Community-level modelling. Biodivers. Conserv. 11, 2309–2338.

Forrest, J.L., Wikramanayake, E., Shrestha, R., Areendran, G., Gyeltshen, K., Maheshwari, A., Mazumdar, S., Naidoo, R., Thapa, G.J., Thapa, K., 2012. Conservation and climate change: assessing the vulnerability of snow leopard habitat to treeline shift in the Himalaya. Biol. Conserv. 150, 129–135.

Fox, J., Yonzon, P., Podger, N., 2002. Mapping conflicts between biodiversity and human needs in Langtang Park, Nepal. Conserv. Biol. 10, 562–569.

Glatston, A. (Ed.), 2010. Red Panda: Biology and Conservation of the First Panda (Noyes Series in Animal Behavior, Ecology, Conservation, & Management).

Glatston, A., Gebaur, A., 2011. The red panda's role in economy and culture. In: Glatston, A.R. (Ed.), Red Panda, the Biology and Conservation of the First Panda. Academic Press, London, UK, pp. 11–25.

Graham, C.H., Ron, S.R., Santos, J.C., Schneider, C.J., Moritz, C., 2004. Integrating phylogenetics and environmental niche models to explore speciation mechanisms in dendrobatid frogs. Evolution 58, 1781–1793.

Groves, C., 2011. Taxonomy and phylogeny of ailurus. In: Glatston, A.R. (Ed.), Red Panda, the Biology and Conservation of the First Panda. Academic Press, London, UK, pp. 101–124.

Guisan, A., Theurillat, J.P., 2000. Equilibrium modeling of alpine plant distribution: how far can we go? Phytocoenologia 30, 353–384.

Guisan, A., Thuiller, W., 2005. Predicting species distribution: offering more than simple habitat models. Ecol. Lett. 8, 993–1009.

Guisan, A., Zimmermann, N.E., 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135, 147–186.

Hardy, S.M., Lindgren, M., Konakanchi, H., Huettmann, F., 2011. Predicting the distribution and ecological niche of unexploited snow crab (*Chionoecetes opilio*) populations in Alaskan waters: a first open-access ensemble model. Integr. Comp. Biol. 51, 608–622. http://dx.doi.org/10.1093/icb/icr102.

Hernandez, P., Graham, C., Mater, L., Albert, D., 2006. The effect of sample size and species characteristic on performance of different species distribution modelling methods. Ecography 29, 773–785.

Herrick, K., 2013. Predictive Modeling of Avian Influenza in Wild Birds. Unpublished PhD thesis, University of Alaska-Fairbanks (UAF).

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.
Hof, A.R., Jansson, R., Nilsson, C., 2012. Usefulness of elevation as a predictor variable in species distribution modeling. Ecol. Model. 246, 86–90.
Huettmann, F., 2005. Databases and science-based management in the context of wildlife and habitat: towards a certified ISO standard for objective decision-making for the global community by using the internet. J. Wildl. Manage. 69, 466–472.
Huettmann, F., 2012. Protection of the Three Poles. Springer, New York, p. 257.
Huettmann, F.T., Gottschalk, 2011. Simplicity, complexity and uncertainty in spatial models applied across time. In: Drew, C.A., Wiersma, Y., Huettmann, F. (Eds.), Predictive Modeling in Landscape Ecology. Springer, New York, pp. 189–208.
Huettmann, F., Artukhin, Y., Glig, O., Humphries, G., 2011. Predictions of 27 Arctic pelagic seabird distributions using public environmental variables, assessed with colony data: a first digital IPY and GBIF open access synthesis platform. Mar. Biodivers. 41, 141–179.
Ives, J.D., 2012. Environmental change and challenge in the Himalaya. A historical perspective. Pirineos 167, 29–68.
Jnawali, S., Leus, K., Molur, S., Glatston, A., Walker, S. (Eds.), 2012. Red Panda (*Ailurus fulgens*) in Nepal. A Population and Habitat Viability Assessment (PHVA) and Species Conservation Strategy (SCS) Workshop, 2–6 September 2010. National Trust for Nature Conservation NTNC, Kathmandu, Nepal. <http://redpandanetwork.org/blog/wp-content/documents/RedPanda_PHVA-2010_Nepal_WorkshopReport_(ed2012-Final)_s.pdf>.
Jones-Farrand, D.T.T.M., Fearer, W.E., Thogmartin, F., Thompson, R., Nelson, M.D., Tirpak, J.M., 2011. Comparison of statistical and theoretical habitat models for conservation planning: the benefit of ensemble prediction. Ecol. Appl. 21, 2269–2282.
Kadmon, R., Farber, O., Danin, A., 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. Ecol. Appl. 14, 401–413.
Krebs, C.J., 2009. Ecology: The Experimental Analysis of Distribution and Abundance, sixth ed. Benjamin Cummings, San Francisco, p. 655.
Lawler, J., Wiersma, J.Y., Huettmann, F., 2011. Designing predictive models for increased utility: using species distribution models for conservation planning, forecasting and risk assessment. In: Drew, C.A., Wiersma, Y., Huettmann, F. (Eds.), Predictive Modeling in Landscape Ecology. Springer, New York, pp. 271–290.
Leathwick, J.R., 1998. Are New Zealand's Nothofagus species in equilibrium with their environment? J. Veg. Sci. 9, 719–732.
Leathwick, J.R., Whitehead, D., McLeod, M., 1996. Predicting changes in the composition of New Zealand's indigenous forests in response to global warming: a modelling approach. Environ. Softw. 11, 81–90.
Magness, D.R., Huettmann, F., Morton, J.M., 2008. Using random forests to provide predicted species distribution maps as a metric for ecological inventory & monitoring programs. In: Smolinski, T.G., Milanova, M.G., Hassanien, A.-E. (Eds.), Applications of Computational Intelligence in Biology: Current Trends and Open Problems. Studies in Computational Intelligence, vol. 122. Springer-Verlag, Berlin Heidelberg, pp. 209–229, 428.
Magness, D., Morton, J.M., Huettmann, F., 2010. How spatial information contributes to the management and conservation of animals and habitats. In: Cushman, S., Huettmann, F. (Eds.), Spatial Complexity, Informatics and Wildlife Conservation. Springer, Tokyo, Japan, pp. 429–444.
Murphy, K., Huettmann, F., Fresco, N., Morton, J., 2012. Connecting Alaska landscapes into the future. U.S. Fish and Wildlife Service, and the University of Alaska. Prepared by the Scenarios Network for Arctic Planning (SNAP). <http://www.snap.uaf.edu/attachments/SNAP-connectivity-2010-complete.pdf>.
Murphy, K., Reynolds, J., Jenkins, J., Whitten, E., Fresco, N., Lindgren, M., Huettmann, F., 2012. Predicting future potential climate-biomes for the Yukon, Northwest Territories, and Alaska: a climate-linked cluster analysis approach to analyzing possible ecological refugia and areas of greatest change. Prepared by the Scenarios Network for Arctic Planning (SNAP) and the EWHALE lab, University of Alaska-Fairbanks on behalf of The Nature Conservancy Canada, Government Northwest Territories. <http://www.snap.uaf.edu/attachments/Cliomes-FINAL.pdf>.
Nielsen, S.E., Stenhouse, G.B., Beyer, H.L., Huettmann, F., Boyce, M.S., 2008. Can natural disturbance-based forestry rescue a declining population of grizzly bears? Biol. Conserv. 141, 2193–2207.
Ohse, B., Huettmann, F., Ickert-Bond, S., Juday, G., 2009. Modeling the distribution of white spruce (*Picea glauca*) for Alaska with high accuracy: an open access role-model for predicting tree species in last remaining wilderness areas. Polar Biol. 32, 1717–1724.
Oppel, S., Huettmann, F., 2010. Using a random forest model and public data to predict the distribution of prey for marine wildlife management. In: Cushman, S., Huettmann, F. (Eds.), Spatial Complexity, Informatics and Wildlife Conservation. Springer, Tokyo, Japan, pp. 151–164.
Oppel, S., Strobel, C., Huettmann, F., 2009. Alternative Methods to Quantify Variable Importance in Ecology. Technical Report Number 65. Department of Statistics, University of Munich.
Pearce, J., Ferrier, S., 2000. Evaluating the predictive performances of habitat models developed using logistic regression. Ecol. Model. 133, 225–245.
Peterson, A.T., 2003. Predicting the geography of species' invasions via ecological niche modeling. Quart. Rev. Biol. 78, 419–433.
Phillips, S.J., Dudik, M., 2008. Modelling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31 (2), 161–175.
Phillips, S.J., Dudik, M., Schapire, R.E., 2004. A maximum entropy approach to species distribution modelling. In: Proceedings of the 21st International Conference on Machine Learning. ACM Press, New York, pp. 655–662.
Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modelling of species geographic distributions. Ecol. Model. 190 (supplement 3–4), 231–259.
Pradhan, S., Saha, G.K., Khan, J.A., 2001. Ecology of the red panda (*Ailurus fulgens*) in the Singhalia National Park, Darjeeling, India. Biol. Conserv. 98, 11–18.
Roberts, M.S., Gittleman, J.L., 1984. Mammalian species: *Ailurus fulgens*. No. 222: 1–8. The American Society of Mammalogists.
Tessarolo, G., Rangel, T.F., Araujo, M.B., Hortal, J., 2014. Uncertainty associated with survey design in species distribution models. Divers. Distrib. 1, 1–12. http://dx.doi.org/10.1111/ddi.12236.
Tse-ring, K., Sharma, E., Chettri, N., Shrestha, A.B. (Eds.), 2010. Climate change vulnerability of mountain ecosystems in the eastern Himalayas. International Center for Integrated Mountain Development, Kathmandu, Nepal.
USGS, 2004. Shuttle Radar Topography Mission, 1 Arc Second scene SRTM_u03_n008e004, Unfilled Unfinished 2.0, Global Land Cover Facility, University of Maryland, College Park, Maryland, February 2000. <http://glcf.umiacs.umd.edu/data/srtm>.
Wei, F., Zhang, Z., 2011. Red panda ecology. In: Glaston, A.R. (Ed.), Red Panda: Biology and Conservation of the First Panda. Academic Press, pp. 193–212.
Wei, F.W., Feng, Z.J., Wang, Z.W., Zhou, A., Hu, J.W., 1999a. Use of nutrients in bamboo by the red panda (*Ailurus fulgens*). J. Zool. 248, 535–541.
Wei, F.W., Feng, Z.J., Wang, Z.W., Li, M., 1999b. Feeding strategy and resource partitioning between giant and red pandas. Mammalia 63, 417–430.
Wilson, K.A., Carwardine, J., Possingham, H.P., 2009. Setting conservation priorities. Ann. N. Y. Acad. Sci. 1162, 237–264.
Winters, A., Yusuf, S. (Eds.), 2007. Dancing With Giants: China, India, and the Global Economy. World Bank Publications, p. 288.
Worboys, G.L., Francis, W.L., Lockwood, M. (Eds.), 2010. Connectivity Conservation Management: A Global Guide – with Particular Reference to Mountain Connectivity Conservation. Earthscan, London, p. 382.
Yen, P., Huettmann, F., Cooke, F., 2004. Modelling abundance and distribution of Marbled Murrelets (*Brachyramphus marmoratus*) using GIS, marine data and advanced multivariate statistics. Ecol. Model. 171, 395–413.
Yonzon, P.B., 1989. Ecology and Conservation of the Red Panda in the Nepal-Himalayas. PhD Thesis. University of Maine, USA.
Yonzon, P.B., 1991. Conservation of the red panda (*Ailurus fulgens*). Biol. Conserv. 57 (supplement 1), 1–11.
Yonzon, P.B., Hunter, M.L., 1991. Cheese, tourists, and red pandas. Conserv. Biol. 5, 196–202.
Yonzon, P.B., Jones, R., Fox, J., 1991. Geographic information systems for assessing habitat and estimating population of red pandas in Langtang National Park, Nepal. Ambio 20, 285–288.
Zhang, Z., Wei, F., Li, M., Zhang, B., Liu, X., Hu, J., 2004. Microhabitat separation during winter among sympatric giant pandas, red pandas, and tufted deer: the effects of diet, body size, and energy metabolism. Can. J. Zool. 82, 1451–1458.
Zhang, Z., Wei, F., Li, M., Hu, J., 2006. Winter microhabitat separation between Giant and Red Pandas in Bashania faberi Bamboo Forest in Fengtongzhai Nature Reserve. J. Wildl. Manage. 70, 231–235.
Ziegler, S.A. Gebauer, Melisch, R., Sharma, B.K., Ghose, P.S., Chakraborty, R., Shrestha, P., Ghose, D., Legshey, K., Pradhan, H., Bhutia, N.T., Tambe, S., Sinha, S., 2010. Sikkim – Im Zeichen des Roten Panda. Zeitschrift des Kölner Zoos 53, 79–92 (English Version available at: <http://www.traffic.org/non-traffic/non-traffic_pub12.pdf>).
Zuckerberg, B., Huettmann, F., Friar, J., 2011. Proper data management as a scientific foundation for reliable species distribution modeling. In: Drew, C.A., Wiersma, Y., Huettmann, F. (Eds.), Predictive Modeling in Landscape Ecology. Springer, New York, pp. 45–70, Chapter 3.