Climate change enlarges China's Great Bustards' (*Otis tarda dybowskii*) suitable wintering distribution in the 21st century

Chunrong Mi, Huettmann Falk, Yumin Guo

Rapidly changing climate makes humans realize that there is a critical need to rethink the current conservation and incorporate climate change adaptation into conservation planning. Whether Great Bustards' (*Otis tarda dybowskii*), a globally endangered species whose population is approximately 1,500~2,200 individuals in China, would survive in a changing climate environment is an important protection issue. In this study, we selected the most suitable species distribution model for bustards from four machine learning models, combining two modelling approaches (TreeNet and Random Forest) with two sets of variables (correlated variables removed or not), using common evaluation methods (AUC, Kappa and TSS) and independent testing data. We found Random Forest with all environmental variables outperformed in all assessment methods. Projected the best model to the latest IPCC-CMIP5 climate scenarios (RCP 2.6, 4.5 and 8.5 in BCC-CSM1-1), we found suitable wintering habitats in the current bustards distribution would increase during the 21st century, and dramatically extend eastwards, lightly northwards and westwards, with ongoing climate change. Northeast Plain and the south of North China and the North of East China would become two major suitable wintering habitats of bustards. However, some current suitable habitats will experience a reduction, such as in Middle and Lower Yangtze. Although our results suggest the habitats quantity and quality would widen with climate changing, greater efforts should be undertaken on human disturbance, such as pollution, hunting, unsuitable agriculture development, infrastructure construction, habitat fragmentation, oil and mine exploitation. All of which are negatively and intensely linked with global change.

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1263v1 | CC-BY 4.0 Open Access | rec: 26 Jul 2015, publ: 26 Jul 2015

Climate change enlarges China's Great Bustards' (*Otis tarda dybowskii*) suitable wintering distribution in the 21st century

Chunrong Mi[1,2], Falk Huettmann[3], Yumin Guo[1]

[1] College of Nature Conservation, Beijing Forestry University, P.O. Box 159, Beijing 100083, China;

[2] Institute of Geographic Sciences and Natural Resources Research, University of Chinese Academy of Science, Beijing 100049, China;

[3] EWHALE Lab, Department of Biology and Wildlife, Institute of Arctic Biology, University of Alaska Fairbanks, 419 Irving I, P.O. Box 757000, AK 99775, USA

Corresponding author: Yumin Guo

guoyumin@bjfu.edu.cn

Publish time: 2015/12

ABSTARCT

Rapidly changing climate makes humans realize that there is a critical need to rethink the current conservation and incorporate climate change adaptation into conservation planning. Whether Great Bustards' (*Otis tarda dybowskii*), a globally endangered species whose population is approximately 1,500~2,200 individuals in China, would survive in a changing climate environment is an important protection issue. In this study, we selected the most suitable species distribution model for bustards from four machine learning models, combining two modelling approaches (TreeNet and Random Forest) with two sets of variables (correlated variables removed or not), using common evaluation methods (AUC, Kappa and TSS) and independent testing data. We found Random Forest with all environmental variables outperformed in all assessment methods. Projected the best model to the latest IPCC-CMIP5 climate scenarios (RCP 2.6, 4.5 and 8.5 in BCC-CSM1-1), we found suitable wintering habitats in the current bustards distribution would increase during the 21st century, and dramatically extend eastwards, lightly northwards and westwards, with ongoing climate change. Northeast Plain and the south of North China and the North of East China would become two major suitable wintering habitats of bustards. However, some current suitable habitats will experience a reduction, such as in Middle and Lower Yangtze. Although our results suggest the habitats quantity and quality would widen with climate changing, greater efforts should be undertaken on human disturbance, such as pollution, hunting, unsuitable agriculture development, infrastructure construction, habitat fragmentation, oil and mine exploitation. All of which are negatively and intensely linked with global change.

Keywords: Climate change, Species distribution models (SDMs), Great Bustards (*Otis tarda*

dybowskii), Random Forest， China

INTRODUCTION

Climate is among the most dominant factors that affect species distributions across broad spatial scales (Woodward 1987, Pearson and Dawson 2003) and ecological niche models are able to successfully quantify these relationships (Drew et al. 2011). Long-term studies indicate that the anomalous climate of the last half-century is already affecting the physiology, distribution, and phenology of many species, especially for many of the already endangered species in ways that are consistent with theoretical predictions (Sykes and Prentice 1996, Hughes 2000). Increasing attention has been given to the ability to predict potential species distributions under various climate change scenarios (Dyer 1995, Iverson and Prasad 1998, Prasad et al., 2006, Wu et al., 2012). A major challenge in conservation planning for these species, in particular, is to incorporate climate change impacts into species conservation strategies (Araújo and Rahbek 2006, Strange et al., 2011).

Species distribution represents an essential foundation in conservation biology (Araujo and Guisan 2006, Tanneberger et al., 2010, Drew et al., 2011). Understanding where species emerge temporally and spatially across large geographic areas is important to conserving, monitoring, and managing species effectively (Wu and Smeins 2000). For this purpose, species distribution models (SDMs), including process-based and bioclimatic envelope approaches, have been suggested as an effective tool to meet these needs (Guisan and Thuiller 2005, Elith et al., 2006, Hu and Jiang 2010). There has been rapid progress in this field of SDMs, and tools and workflows are now available to assess distributions and the impacts of climate change on species and habitats (Peterson et al., 2002, Hijmans and Graham 2006, Drew et al., 2011). Most of the traditionally used models are too complex for parameterization, model fit, assessment of met assumptions, and validation in model application (Pearson and Dawson 2003, Li et al., 2015). Although the bioclimatic envelope models have various limitations (such as the assumption of equilibrium, the assumption of complete sampling of species niche, and insufficient inclusion of adaptation, evolution, and dispersal), they are still applied by many researchers (Huntley et al., 2010, Araújo and Peterson 2012). With a good understanding of the modelling techniques, careful choice of explanatory predictors, and appropriate model validation and testing, these SDMs can still provide important information on the potential impact of climate change on species range shifts, and help inform conservation decisions in a changing climate (Hijmans and Graham 2006, Araújo and Peterson 2012). The use of machine learning specifically offers a great extension of traditional SDMs, see for instance Drew et al. (2011) and Kamel et al. (2015).

The Great Bustard is one of the world's heaviest flying birds, occupying grassland habitats; it is categorized as a globally threatened (VU) species according to the IUCN. Its world population for 2010 was estimated to be 44,100 to 57,000 individuals; approximately 4~10% of the global population is located in China and is believed to be declining (Alonso and Palacín 2010). The Great Bustard is divided into two subspecies: *Otis tarda tarda* and *Otis tarda dybowskii*. The *O. t. dybowskii* population is small and consists of approximately 1,500~2,200 individuals (Goroshko 2010). This subspecies is distributed throughout eastern Asia in areas such as Russia, Mongolia, China, and South and North Korea (Kong and Li 2005). In China, *O. t. dybowskii* makes for a typical agricultural steppe bird, distributed during winter in Heilongjiang, Jilin, Inner Mongolia, and Hebei (Jiang 2003, Wang and Yan 2002).

In the recent 10 years and more, we found it's hard for humans to observe this subspecies in the middle and lower reaches of the Yangtze River (such as Poyang Lake and Dongting Lake) where they were known stay during winter in early times. We think there must be many reasons for this change, such as habitat loss, hunting and climate change etc. Here we investigate the wintering distribution of *O. t. dybowskii* and how they may be affected by climate change. We employed species distribution models based on machine learning (TreeNet and Random Forest). According to our knowledge, this is the first predictive, spatial model of wintering Great Bustards. It presents a step toward developing a national or Northeast Asia conservation effort to assess management uncertainties. More specifically, the goal of this study is to estimate the spatial impacts of climate change on the future wintering distribution of Great Bustards. The results of this study will provide information on what habitat changes may occur, and guide future sampling, surveying, and conservation efforts across China. We try to infer for the wider status of this bird during times of Global Change.

MATERIALS AND METHODS

Study area and data

The species data we used in this study was from our own fieldwork investigations of 2012 and 2013. Also we used previously published literature data and mapped it in ArcGIS10.1 (see Supplement S1. Overall, we used 102 geo-referenced bird sighting locations for a time period 1990-2013 across China and used them as confirmed presence points to explore the relationships between the Great Bustard and bioclimatic variables for a climate envelope. To accommodate potential species dispersal under future climate change, we extended the study area to all of China (Figure 1).

Put Figure 1 here

Nineteen bioclimatic variables at a 30s resolution were obtained from the Worldclim database (Hijmans et al., 2005, http://www.worldclim.org/) for current climate (1950–2000) and future climate scenarios for 2050 and 2070 (average for 2041-2060 and 2061–2080). The data applied here are the most recent IPCC-CMIP5 climate projections from Global Circulation Models (BCC-CSM1-1, develops by China) under three representative concentration pathways (RCP 2.6, 4.5 and 8.5). Other environmental variables that are considered to be important drivers of Great Bustards distributions were also used to build the bustards' habitat distribution model, including topographical factors (altitude, slope, and aspect), water factors (distance to river, distance to lake, distance to coastline), inference factor (distance to road, distance to rail road, and distance to settlement), and land-cover. Aspect and slope layer were derived from altitude layer in ArcGIS10.1, which was obtained from Worldclim database; Road, rail road, river, lake and coastline and settlement map were obtained from Natural Earth database; Land cover map was from ESA database (detailed information is provided in Supplement S2)All spatial layers of these environmental variables were resampled to a resolution of 30s to correspond to that of bioclimatic variables. Because reliable future projections of these variables (land cover, distance to road, rail road, settlement, river, and lake) are not available, and because including static variables in SDMs alongside dynamic variables can improve model performance (Stanton et al., 2012), we kept these variables static in our projections. Therefore, our future scenarios are underestimates for these stressors and the situation overall. To address the argument of potentially overfitting of in the models, we first calculated correlations among 19 bioclimatic variables and other 10 environmental variables in ArcGIS. We removed a variable when correlation coefficient >|0.90| was obtained (see correlation matrix in Supplement S3 (Costa et al., 2010). A total of 11 bioclimate variables were removed, 4 bioclimatic variables and 10 environmental variables were left. Consequently, we constructed two sets of bustard distribution models: one was based on the result of correlation test that kept the 4 bioclimatic variables and 10 environmental predictors were used to construct SDMs; the other approach was to use all of the 19 bioclimatic and 10 environmental variables for model construction.

Species distribution modeling and testing

Species distribution models combine information from point occurrence data and environmental variables to predict the geographical distribution of species and communities. Various statistical approaches have been employed in this type of modelling whereas simplistic linear methods have been shown to perform less optimal (Elith et al., 2006 and Fernandez-Delgado et al., 2014). Therefore, we chose TreeNet (hereafter TN, also called boosted regression trees (BRT), stochastic gradient boosting, Friedman 2002) and Random Forest (hereafter RF, Breiman 2001) as our species distribution models. TN and RF are stand-alone software products from Salford Systems Ltd (Salford Predictive Modeler (SPM)), and each performs one specific techniques. Boosting is based on the concept in machine learning that many weak learners allow for a superior learner. It employs an optimized set of trees in a sequence where each new tree explains the unexplained variance of the previous tree (Friedman 2002). Bagging is based on a similar concept, but it uses each time a specific subset of rows and predictors, and then averages out these trees for the best outcome (Breiman 2001 and Hegel et al., 2010 for overview). These two approaches have shown to be two of the best performing models in predicting ecological species distributions with presence-absence data (Zhai and Li 2003, Elith et al., 2006, Lei et al., 2011). It also has been extensively applied to project species range and vegetation shifts under climate change (Drew et al., 2011). For more details on TreeNet and Random Forest, we refer readers to read the userguide (https://www.salford-systems.com/products/spm/userguide). About 10,000 pseudo-absence points were taken by random sampling across China by using the freely available Geospatial Modeling Environment (GME; Hawth's Tools). We used 10 fold cross-validation procedure for TreeNet, where it divided our dataset using 80% of the data for model calibration and retaining 20% of the data for evaluation; and out of bag (OOB) data used to test Random Forest. In addition, balanced was set as class weights, and 1000 trees were built for all models to find the optimum. The test and background points were used to calculate receiver operating characteristic (ROC) curves, the Kappa Statistic (Kappa) and the True Skill Statistic (TSS) (Graham and Hijmans 2006). The area under the ROC curve (AUC) is commonly used to evaluate models in species distributional modelling (Manel et al., 2001, McPherson et al., 2004). A ROC curve is created by plotting the true-positive fraction against the false-positive fraction for all test points across all possible probability thresholds (Fielding & Bell, 1997); Kappa (Cohen 1960) is a measure of model performance, the advantages of kappa are its simplicity, the fact that both commission and omission errors are accounted for in one parameter, and its relative tolerance to zero values in the confusion matrix (Manel et al., 2001, Allouche et al., 2006); whereas the True Skill Statistic (TSS) was also used to evaluate the model performance. TSS is increasingly applied as a simple and intuitive measure for the performance of species distribution models that are built from presence-only data and binary predictions are produced by applying thresholds (Allouche et al., 2006). The best suitable SDM for bustards was finally determined by comparing the AUC, Kappa and TSS of all models, and also boxplots with 95% confidence intervals of relative index of occurrence (RIO) for independent testing data (see Supplement S4 from literature (the Threaten Birds of Asia, Collar et al., 2001). Subsequently, we applied the sensitivity-specificity equality approach (Liu et al., 2005) as the most suitable threshold to define the presence–absence distribution of great bustards wintering habitats, after comparing three approaches (Kappa maximization approach (Liu et al., 2005), the sensitivity-specificity equality approach (Liu et al., 2005), and lowest presence threshold approach (Pearson et al., 2007, see Supplement S5). A GCM (BCC-CSM1-1, developed by China) was used to produce locally valid probability outputs for each scenario (RCP 2.6, 4.5 and 8.5). The overall work flow for great bustards' best suitable model select and project to future climate scenarios is presented in Figure 2.

Habitat suitability and classification

We adopted the thresholds of 0.5, 0.7 and 0.85 (0.85 created by sensitivity-specificity equality approach, see Supplement S5) to identify the distributions of four different classes of habitat suitability for great bustards' final model. Our selection of thresholds was based on our knowledge for great bustards' true current distribution across China. Finally, we reclassified bustards' habitats into unsuitable, marginally suitable, moderately suitable and highly suitable.

Spatial analysis of potential effects of climate change

We applied ArcGIS 10.1 to calculate the area of different classes of great bustards habitats for

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1263v1 | CC-BY 4.0 Open Access | rec: 26 Jul 2015, publ: 26 Jul 2015

three periods of time (current, 2050 and 2070) in three scenarios (RCP 2.6, 4.5 and 8.5). We also used the tabulate area analysis (in ArcGIS 10.1) to assess the distribution patterns and potential changes of different classes of the Great Bustard habitats. This allowed us to identify areas of the habitat range that are projected to be lost, gained or remain under future climate scenarios. While these projections do not take into account stochastic and other effects yet, they allow to access business as usual scenarios using best-available science. In addition, we can locate the directionality of these habitat changes and where they move, e.g. east, south, west and north to obtain the birds' distribution shift trends.

Put Figure 2 here

RESULTS

The high AUC values (>0.91) for all four models of great bustards (Table 1) indicated that our models can accurately capture bustards relationships, and values above 0.75 generally indicate adequate model performance for most applications (Pearce and Ferrier 2000). AUCs of Random Forest model were higher than TreeNet, and SDMs with 29 variables were higher than more parsimonious models with 14 variables. Values of what consist of acceptable max-Kappa vary in the literature, but generally 0.6 and above are considered'good' (Czaplewski and Forest 1994, Fielding and Bell, 1997). Therefore, TreeNet models did not perform suitable enough in this study for bustards. The high TSS (>0.82) for this subspecies also indicated that our model performs well in projecting bustards habitat distributions. TSS had the same trends as AUC and Kappa, Random Forest performed better than TreeNet, more predictors (29) models were better than relative few predictors (14) model (14 predictors are still a high number of predictors in model studies). Boxplots created by the independent testing data from literature (the Threaten Birds of Asia) (Figure 3) indicated same results than above, Random Forest showed a higher relative index of occurrence (hereafter RIO) than TreeNet, and a stronger focus on a narrow range of values (>0.9), 29 variables models performed a little better than the 14 variables model. According to the above four consistent results, we selected Random Forest model with 29 predictors as our final SDM and projected to future climate.

Put Table 1 here

Put Figure 3 here

Our results indicated that climate change would enlarge the suitable area of great bustards wintering habitats (Figure 4 and Table 2). Under the IPCC CMIP5 representative concentration pathway (RCP) 8.5 climate change scenario, the total habitat area for great bustards would improve from the current 0.92 million km^2 to 2.07 million km^2 by 2050, an improvement of 124.06%; to 2.47 million km^2 by 2070, an improvement of 19.10% from 2050. Under RCP4.5, a median radiative forcing, climate change would result in a habitat increase of 81.33% by 2050, 11.96% again by 2070 from 2050. And under RCP 2.6, the lowest radiative forcing, the habitat area would still increase by 71.44% by 2050 and by 1.26% to 2070 (Table 2).

Put Figure 4 here

In addition to an increase in the predicted area of suitable habitat, climate change would also enhance habitat quality. The moderately suitable habitats are likely to be affected the most by climate change. Under RCP 2.6, moderately suitable habitats would increase from the current 0.27 million km^2 to 0.56 million km^2 by 2050, and to 0.58 million km^2 by 2070

(Table 2). Marginally suitable and highly suitable habitats would also increase in proportion obviously, although it wouldn't increase so much as moderately suitable habitats. In contrast, the unsuitable area would lightly decrease in proportion but dramatically in area (Table 2). RCP 8.5 and RCP 4.5 would produce even more positive impacts on great bustards' habitats than RCP 2.6 (Table 2). Detailed information of different classes of suitable habitats transformed between the current and three RCPs by 2050 and 2070 could see in Table 3. Under all three RCPs, over 45% marginally suitable habitats would become moderately suitable, 16% become highly suitable area, and 45% original moderately suitable habitats transform to highly suitable range from current to 2050. From 2050 to 2070, 16% marginally suitable habitats would change to moderately suitable habitats, highly suitable area gained above 7% from moderately suitable land. In addition, according to the information of Figure 4, under RCP 2.6, the highly suitable habitats in Northeast Plain would increase and the quality would enhance. This makes it an important and large habitat for wintering bustards such as the habitats in the south of North China and the northern part of East China (including areas south of Hebei, Henan, west of Shandong, north of Anhui and west of Jiangsu Province) and the Hetao Plain (the central part of Inner Mongolia). However, the habitat in the Middle and Lower Yangtze (bustards major distribution areas in Poyang and Donting Lakes) would suffer the most severe fragmentation under future climate changes (Figure 4). RCP 4.5 and RCP 8.5 shows similar impacts on great bustards' habitats with RCP 2.6.

Put Table 2 here

Put Table 3 here

Moreover, climate change would lead to both horizontal and vertical changes in Bustards habitat distribution, with the three RCPs producing similar trends of impact (Figure 4 and Table 4). More specifically, in the Northeast Plain, projected habitat area gains to the east, particularly in the east of Heilongjiang Province, the boundary between China and Russia (Figure 4). In the western part, it would produce a westward shift of marginally suitable habitats raised in Xinjiang Province where possibly the distribution range is located for *Otis tarda tarda* living in China (Figure 4). Table 4 illustrated that the highly suitable habitats of Great bustards would dramatically shift eastward (nearly by 7°) in all three RCPs, lightly moving northward (3′-2 °) and westward (19′- 1° 28′). It should be noted that the habitats of North Plain, and the habitat in south of North China and Northwest of East China would become two major wintering distribution range of Great Bustards.

Put Table 4 here

DISCUSSION

Effective conservation of great bustards requires protection and restoration of their suitable habitats. Our model is the first to predict and map, with high accuracy (AUC: 0.982, Kappa: 0.704, TSS: 0.94), the wintering distribution of *O. t. dybowskii* in China. Our results indicated that these suitable wintering habitats in the current bustard distribution would increase during the 21st century (Table 2 and 3), and dramatically extend eastwards, lightly northwards and westwards, with ongoing climate change (Figure 4 and Table 4). However, some current suitable habitats will experience a reduction, such as in Middle and Lower Yangtze, where birds' observers haven't seen great bustards in recent ten years. Such a finding is very relevant for the improved understanding and for prioritizing conservation efforts.

Furthermore, these results can also be used in the future for impact studies similar to Nielsen et al., (2008) following spatial Population Viability Analysis (sPVA).

Though the suitable habitats of bustards were projected to increase, this does by no means mean that it will have the same trend in population size. That's because the population size of the birds is correlated with other factors in addition to climate, particularly one should consider human disturbance (including pollution, hunting, unsuitable agriculture development, infrastructure construction, habitat loss, habitat fragmentation, oil and mine exploitation for instance).

According to our modelling, the Northeast Plain would become one of the major wintering distribution areas for this species. Originally, there actually were several male individuals overwintering in the Northeast Plain (Liu 1997). Here we speculated that more bustards, both male and female individuals, may remain there. Based on this result, we inferred this habitat will become a residential area or that bustards will have a shorter distance for migration than in earlier times. This situation has already been observed in the Red-crowned Crane (Masatomi et al., 2007) and the Oriental White Stork (Yang et al., 2007).

The suitable wintering habitats of bustards in Northeast Plain are located southeast of the Greater Khingan, southwest of the Lesser Khingan Mountains as well as northwest of the Changbai Mountains. It is possible that these mountains might become a natural barrier to the habitat expansion of this subspecies. These areas are used for agriculture and susceptible to urban expansion, thus the question of how to leave enough space and how to protect and maintain this species under such a situation should be taken into serious consideration before new policies and conservation plans are made.

To our knowledge, ecological applications of TreeNet (also named BRT,) or Random Forest prediction models for birds are still not widespread (but see Booms et al., 2010, Drew et al., 2011, Hardy et al., 2011, Steen and Powell 2012). Increasingly though researchers have used MaxEnt (Hu et al., 2010, Shao et al., 2011, Wu et al., 2012), and remain using Generalized Additive Models (GAMs) (Li et al., 2010), Generalized Linear Models (GLMs) (de Boer et al., 2011, Zhang et al., 2011) and Discriminant Function Analysis (DFA) (Wang et al., 2011, Xia et al., 2012). However, there are several advantages of TreeNet and Random Forest of Salford's SPM software (Herrick 2013), when compared with other techniques that are often used for building habitat models such as GLM or DFA: (i) it is non-parametric, (ii) the GUI is very user friendly but powerful, (iii) it automatically selects the important predictor variables thus no prior variable selection or data reduction is required, (iv) the results are quite invariant with regard to modifications of the data such as transformation or rescaling,(v) it approaches missing values automatically and in a good possible way, (vi) it is quite immune to outliers in predictors or the target variable, i.e., if samples are coded incorrectly and the model prediction starts to diverge substantially from observed data, the data will not be used in further updates. TreeNet and Random Forest construct models conveniently and without the time consuming pre-processing of the data. Furthermore, they are remarkably resistant to overfitting (Wickert 2007, Breiman 2001).

Despite the statistical superiority of Random Forest (Fernandez-Delgado et al., 2014), care is needed in deciding whether to use climatic variables for prediction models because they may emphasize the fundamental rather than the realized niche. However, habitat and climate models can provide new insights into factors limiting species distributions and how they may respond to climate change (Suárez‐Seoane et al., 2004). Our results show that model construction does not have to remove correlation variables first; it means to loose information to obtain better predictions! In the process of model building, we found models with correlation variables (29 predictors' models) could be more accurate than model without correlate variables (Table 1 and Figure 3), which stands clearly against the notion of parsimony (Burnham and Anderson 2002).

In summary, there is a critical need to rethink the current approach to parsimony and conservation, and to incorporate climate change adaptation into our conservation planning with a rapidly changing climate. Based on concrete data and a robust modelling approach, our model would be useful to managers currently addressing conservation issues in China. For example, the model could be more combined with existing regionalized IPCC climate models to forecast future *O. t. dybowskii* population size and changes under varying climate scenarios. In addition, distribution maps, created in-time, could overlay maps of the current and predicted locations of oil, gas, mineral, and wind resources to identify areas of potential future conflict, estimate the potential size or severity of impacts caused by a specific activity, and prioritize conservation strategies geographically (such as Marxan applications etc.; Beiring 2014 for parts of Asia).

ACKNOWLEDGEMENTS

This research was possible because of the large investment of field effort, money, time, personal interest, and dedication by researchers for the past 24+ years. We heartily thank all those who contributed to the International Great Bustard Census, as well as students from the EWHALE lab, UAF and Salford Systems Ltd and China Great Bustard Protection and

Monitoring Network (http://www.otistarda.org/en).

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This research was funded by The National Forestry Bureau of China.

Author Contributions

Chunrong Mi analyzed the data, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.

Falk Huettmann analyzed the data and reviewed drafts of the paper.

Yumin Guo reviewed drafts of the paper.

REFERENCE

Alonso, J C. and Palacín, C. 2010. The world status and population trends of the Great Bustard (Otis tarda): 2010 update. *Chinese Birds* 1(2): 141-147

Allouche, O., Tsoar, A. and Kadmon, R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology.* 43: 1223-1232.

Araújo, MB., Peterson, A.T., 2012. Uses and misuses of bioclimatic envelope modeling. *Ecology* 93: 1527-1539.

Araújo MB, and Rahbek C. 2006. How Does Climate Change Affect Biodiversity? *Science* 313:21-22.

Araujo, MB. and Guisan, A. 2006. Five (or so) challenges for species distribution modelling. *Journal of Biogeography* 33: 1677-1688.

Araújo, MB. and Peterson, AT. 2012. Uses and misuses of bioclimatic envelope modeling. *Ecology* 93: 1527-1539.

Beiring, M. 2014. Modeling Migratory Passerines in the Pacific Rim. Unpublished M.Sc. Thesis, University of Vienna, Austria

Booms, TL., Huettmann, F. and Schempf, PF. 2010. Gyrfalcon nest distribution in Alaska based on a predictive GIS model. *Polar biology* 33: 347-358.

Breiman, L. 2001. Random Forests. *Machine Learning* 45(supplement 1): 5-32.

Burnham, KP. and Anderson, DR. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer.

Cohen, J. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20: 37-46.

Costa, G., Nogueira, C., Machado, R., and Colli, G. 2010 Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. *Biodiversity and Conservation* 19: 883-899.

Collar, N.J., Crosby, R., and Crosby, M. 2001. Threatened birds of Asia: the BirdLife International red data book. BirdLife International Cambridge, UK.

Czaplewski, RL. and Forest, RM. 1994. Variance approximations for assessments of classification accuracy. US Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station Fort Collins, CO.

de Boer, WF., Cao, L., Barter, M., Wang, X., Sun, M., van Oeveren, H., de Leeuw, J., Barzen, J. and Prins, H. 201. Comparing the community composition of European and Eastern Chinese Waterbirds and the influence of human factors on the China Waterbird Community. *Ambio* 40: 68-77.

Drew, CA., Wiersma, Y. and Huettmann, F. 2011. Predictive species and habitat modeling in *landscape ecology*: Springer.

Dyer, JM. 1995. Assessment of climatic warming using a model of forest species migration. *Ecological Modelling* 79: 199-219.

Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J., Lehmann, A., Li, J., G. Lohmann, L., A. Loiselle, B., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., McC. M. Overton, J., Townsend Peterson, A., J. Phillips, S., Richardson, K., Scachetti-Pereira, R., E. Schapire, R., Soberón, J., Williams, S., S. Wisz, M. and E. Zimmermann, N. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography* 29: 129-151.

Fernández-Delgado, M., Cernadas, E., Barro, S. and Amorim, D. 2014. Do we need hundreds of classifiers to solve real world classification problems? *The Journal of Machine Learning Research* 15: 3133-3181.

Fielding, AH. and Bell, JF. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental conservation* 24: 38-49.

Friedman, J. 2002. Stochastic gradient boosting. *Computational Statistics & Data Analysis* 38: 367-378.

Goroshko, OA. 2010. Present status of population of Great Bustard (Otis tarda dybowskii) in Dauria and other breeding grounds in Russia and Mongolia: distribution, number and dynamics of population, threats, conservation. pp. First International Symposium on Conservation of Great Bustard. Beijing.

Graham, CH. and Hijmans, RJ. 2006. A comparison of methods for mapping species ranges and species richness. *Global ecology and biogeography* 15: 578-587.

Guisan, A. and Thuiller, W. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology letters* 8: 993-1009.

Hardy, SM., Lindgren, M., Konakanchi, H. and 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. *Integrative and comparative biology* 51: 608-622.

Hegel, TSA., Cushman, J E, and Huettmann, F. 2010. Current State of the Art for Statistical Modelling of Species Distributions. Chapter 16, pp. 273 – 312. Cushman and F. Huettmann. Spatial Complexity, Informatics and Wildlife

Conservation. Springer Tokyo, Japan. pp. 273-312

Herrick, KA., Huettmann, F. and Lindgren, MA. 2013. A global model of avian influenza prediction in wild birds: the importance of northern regions. *Vet Res* 44(1): 42.

Hijmans, RJ. and Graham, CH. 2006. The ability of climate envelope models to predict the effect of climate change on species distributions. *Global change biology* 12: 2272-2281.

Hu, J., Hu, H. and Jiang, Z. 2010. The impacts of climate change on the wintering distribution of an endangered migratory bird. *Oecologia* 164: 555-565.

Hughes, L. 2000. Biological consequences of global warming: is the signal already apparent? *Trends in Ecology & Evolution* 15: 56-61.

Huntley, B., Barnard, P., Altwegg, R., Chambers, L., Coetzee, B.W., Gibson, L., Hockey, P.A., Hole, DG., Midgley, GF. and Underhill, LG. 2010. Beyond bioclimatic envelopes: dynamic species' range and abundance modelling in the context of climatic change. *Ecography* 33: 621-626.

Iverson, LR. and Prasad, AM. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. *Ecological Monographs* 68: 465-485.

Jiang, J. 2003. The status of resource and conservation of Great Bustard in China. M.Sc. Thesis, Northeast Forestry University (in Chinese).

Kandel, K., Huettmann, F., Suwal, MK., Regmi, GR., Nijman, V., Nekaris, K., Lama, ST., Thapa, A., Sharma, HP. and Subedi, TR. 2015. 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. *Biological Conservation* 181: 150-161.

Kong, Y. and Li, F. 2005. The Status and Research Trends of the Great Bustard. Chinese *Journal of Zoology* 40: 111-115 (in Chinese).

Lei, Z., Liu, S., Sun, P., and Wang, T. 2011. Comparative evaluation of multiple models of the effects of climate change on the potential distribution of Pinus massoniana. *Chinese Journal of Plant Ecology* 35: 1091-1105.

Li, R., Tian, H. and Li, X. 2010. Climate change induced range shifts of Galliformes in China. *Integrative zoology* 5: 154-163.

Li, R., Xu, M., Wong, M.H.G., Qiu, S., Li, X., Ehrenfeld, D. and Li, D. 2015. Climate change threatens giant panda protection in the 21st century. *Biological Conservation* 182: 93-101.

Liu, B. 1997. The Status and Protecion of Great Bustards in Northeast. *Natural Resources Study* 61-63 (in Chinese).

Liu, C., Berry, P., Dawson, T. and Pearson, R. 2005. Selecting thresholds of occurrence in the prediction of species distributions. *Ecography* 28: 385-393.

Manel, S., Williams, HC. and Ormerod, S. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of applied Ecology* 38: 921-931.

Masatomi, Y., Higashi, S. and Masatomi, H. 2007. A simple population viability analysis of Tancho (Grus japonensis) in southeastern Hokkaido, Japan. *Population ecology* 49: 297-304.

McPherson, J., Jetz, W. and Rogers, D.J. 2004. The effects of species' range sizes on the accuracy of distribution models: ecological phenomenon or statistical artefact? *Journal of Applied Ecology* 41: 811-823.

Nielsen, S.E., Stenhouse, GB., Beyer, HL., Huettmann, F. and Boyce, M S. 2008. Can natural disturbance-based forestry rescue a declining population of grizzly bears? *Biological Conservation* 141: 2193-2207.

Pearce, J. and Ferrier, S. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133: 225-245.

Pearson, RG. and Dawson, TP. 2003 Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global ecology and biogeography* 12: 361-371.

Pearson, RG., Dawson, TP. and Liu, C. 2004. Modelling species distributions in Britain: a hierarchical integration of climate and land‐cover data. *Ecography* 27: 285-298.

Pearson, RG., Raxworthy, C.J., Nakamura, M. and Townsend Peterson, A. 2007. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography* 34: 102-117.

Peterson, AT., Ortega-Huerta, MA., Bartley, J., Sánchez-Cordero, V., Soberón, J., Buddemeier, RH. and Stockwell, DR. 2002. Future projections for Mexican faunas under global climate change scenarios. *Nature* 416: 626-629.

Shao, J., Zhou, Y., Li, J., Wang, X., Luo, Z. and Yan, B. 2011. Spatial Distribution Analysis of Wild Bird Migration in Qinghai Lake based on Maximum Entropy Modeling. In Networking and Distributed Computing (ICNDC). 2011 Second International Conference on. pp. 140-144. IEEE.

Stanton, JC., Pearson, RG., Horning, N., Ersts, P., and Reşit Akçakaya, H. 2012. Combining static and dynamic variables in species distribution models under climate change. *Methods in Ecology and Evolution* 3: 349-357.

Steen, VA. and Powell, AN. 2012. Wetland selection by breeding and foraging Black Terns in the Prairie Pothole region of the United States. *The Condor* 114: 155-165.

Strange, N., Thorsen, B.J., Bladt, J., Wilson, K.A. and Rahbek, C. 2011. Conservation policies and planning under climate change. *Biological Conservation* 144, 2968-2977.

Suárez - Seoane, S., Osborne, PE. and Rosema, A. 2004. Can climate data from METEOSAT improve wildlife distribution models? *Ecography* 27: 629-636.

Sykes, M. and Prentice, IC. 1996. Climate change, tree species distributions and forest dynamics: A case study in the mixed conifer/northern hardwoods zone of northern Europe. *Climatic Change* 34: 161-177.

Tanneberger, F., Flade, M., Preiksa, Z. and Schroeder, B. (2010) Habitat selection of the globally threatened Aquatic Warbler Acrocephalus paludicola at the western margin of its breeding range and implications for management. *Ibis* 152: 347-358.

Wang, Q. and Yan, C. 2002. Chinese Cranes, Rails and Bustards, Taiwan: National Fenghuangu Bird Park (in Chinese).

Wickert, C. 2007. Breeding White Storks (Ciconia ciconia) in former East Prussia: comparing predicted relative occurrences across scales and time using a stochastic gradient boosting method (TreeNet), GIS and public data. M. Sc. Thesis, University of Potsdam, Germany.

Woodward, FI. 1987. Climate and plant distribution: Cambridge University Press.

Wu, W., Gu, S., Wu, J., Cao, M., Juncheng, L. and Xu, H. 2012 Impact of Climate Change on Distribution of Breeding Sites of Red-Crowned Crane in China. J*ournal of Ecology and Rural Environment* 3: 004.

Wu, X. B. and Smeins, F. E. 2000. Multiple-scale habitat modeling approach for rare plant conservation. *Landscape and Urban Planning* 51: 11-28.

Xia, C., Lin, X., Liu, W., Lloyd, H. and Zhang, Y. 2012. Acoustic identification of individuals within large avian populations: a case study of the Brownish-flanked Bush Warbler, South-Central China. *PloS one* 7: e42528.

Yang, C., Zhou, L., Zhu, W. and Hou, Y. 2007. A preliminary study on the breeding biology of the oriental whitestork Ciconia boyciana in its wintering area. *Acta Zoologica Sinica* 53: 215-226 (in Chinese).

Zhai, T. and Li. X. 2003. Climate change induced potential range shift of the crested ibis based on ensemble models. *Acta Ecologica Sinica* 9: 1353-1362 (in Chinese).

Zhang, S., Lei, F., Liu, S., Li, D., Chen, C. and Wang, P. 2011.Variation in baseline corticosterone levels of Tree Sparrow (Passer montanus) populations along an urban gradient in Beijing, China. *Journal of Ornitholog*. 152: 801-806.

Figure Legend:

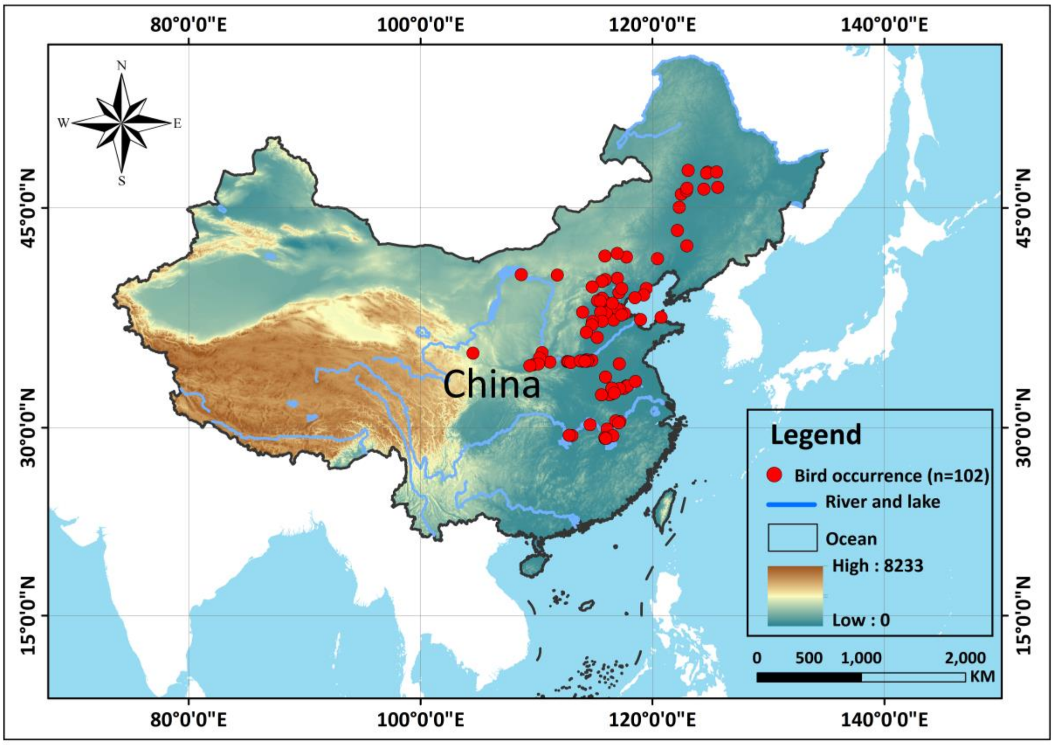

Figure 1 The study area of great bustards. 102 presence records of bustards are shown; the elevation of this study area ranges from 0 to 8,233 m.

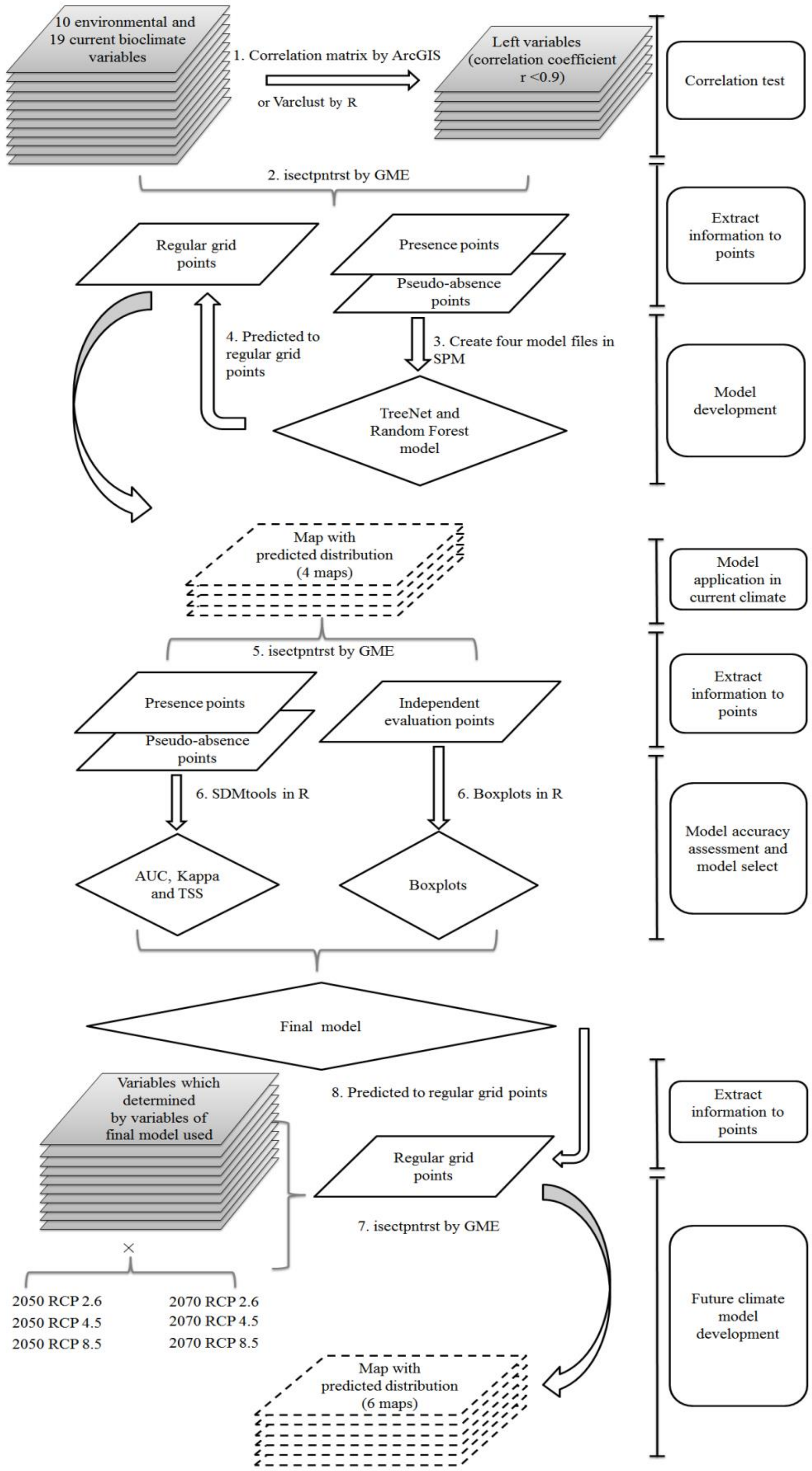

Figure 2 The work flow for great bustards best suitable model select and project to future climate scenarios.

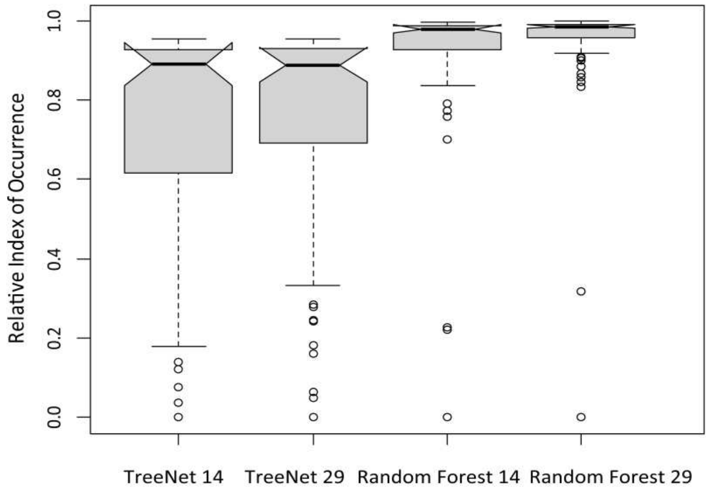

Figure 3 Boxplots of independent testing data from literature (the Threaten Birds of Asia) extracted from four great bustards distribution models

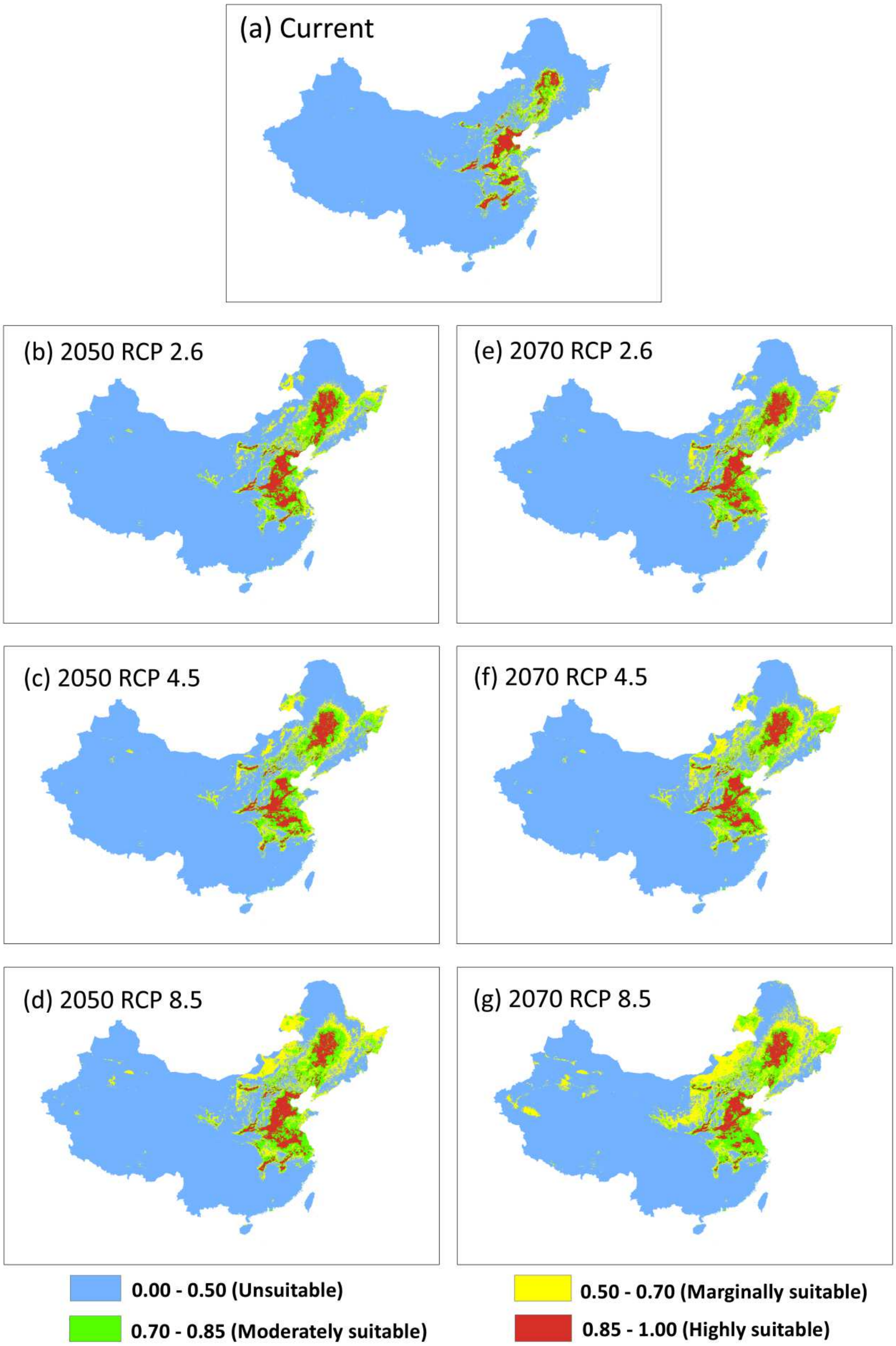

Figure 4 Projected wintering distributions of Great Bustards extracted from the Random Forest with 29 predictors model for the current, 2050 and 2070 in China: (a) current suitability; (b)–(d) suitability in 2050, (e)-(g) suitability in 2070 projected by GCM (BCC-CSM1-1, BC) for the three RCPs(2.6, 4.5, 8.5).

Table Legend

Table 1 The AUC, Kappa and TSS value of four great bustards' distribution models

	TreeNet 14	TreeNet 29	Random Forest 14	Random Forest 29
AUC	0.914	0.923	0.961	**0.982**
Kappa	0.386	0.386	0.656	**0.704**
TSS	0.828	0.846	0.922	**0.965**

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1263v1 | CC-BY 4.0 Open Access | rec: 26 Jul 2015, publ: 26 Jul 2015

Table 2 the patch area (km^2) of great bustards' habitats in different suitability classes and the entire habitat landscape under the current and three RCPs by 2050 and 2070 (total suitable was the sum of marginally suitable, moderately suitable and highly suitable. The percentage (%) of habitat changed was obtained from comparing three scenarios of 2050 with current situation, and scenarios of 2070 comparing with related scenarios of 2050).

Time Period	Unsuitable (km^2)(%)	Marginally suitable (km^2) (%)	Moderately suitable (km^2) (%)	Highly suitable (km^2) (%)	Total suitable (km^2) (%)
Current	8,671,912	362,778	274,669	290,641	928,088
2050 RCP2.6	8,008,852 (-7.65)	584,976 (+61.25)	565,263 (+105.80)	440,910 (+51.70)	1,591,149 (+71.44)
2050 RCP4.5	7,917,098 (-8.70)	646,672 (+78.26)	620,788 (+126.01)	415,441 (+42.94)	1,682,901 (+81.33)
2050 RCP8.5	7,520,554 (-13.28)	917,185 (152.82)	689,439 (+151.01)	472,822 (+62.68)	2,079,446 (+124.06)
2070 RCP2.6	7,988,755 (-0.25)	586,750 (+0.30)	585,887 (+3.65)	438,608 (-0.52)	1,611,245 (+1.26)
2070 RCP4.5	7,715,845 (-2.54)	808,308 (+25.00)	670,526 (+8.01)	405,321 (-2.44)	1,884,155 (+11.96)
2070 RCP8.5	7,123,483 (-5.28)	1,255,645 (+36.90)	821,434 (+19.15)	399,438 (-15.52)	2,476,517 (+19.10)

Table 3 The patch area (km^2) of great bustards' habitats in different suitability classes transformed between the current and three RCPs by 2050 and 2070

xxxx means the area of lower suitable habitats transformed into higher suitable habitats, xxxx means the area of original class of suitable habitats remained (the percentage (%) of habitats transformed was obtained from comparing three scenarios of 2050 with current situation, and scenarios of 2070 comparing with related scenarios of 2050).

2050 RCP2.6 / Current	Unsuitable (km^2)(%)	Marginally suitable (km^2) (%)	Moderately suitable (km^2) (%)	Highly suitable (km^2) (%)
Unsuitable	7,941,160 (91.57)	481,759 (13.18)	229,776 (2.65)	19,153 (0.22)
Marginally suitable	47,803 (13.18)	76,501 (21.09)	175,929 (48.49)	62,544 (17.24)
Moderately suitable	13,749 (5.01)	17,746 (6.46)	112,474 (40.92)	130,764 (47.61)
Highly suitable	6,139 (2.11)	8,969 (3.09)	47,084 (16.2)	228,449 (78.6)

2070 RCP2.6 / 2050 RCP2.6	Unsuitable (km2) (%)	Marginally suitable (km2) (%)	Moderately suitable (km2) (%)	Highly suitable (km2) (%)
Unsuitable	7,782,226 (97.16)	196,058 (2.45)	28,458 (0.36)	2,110 (0.03)
Marginally suitable	195,946 (33.50)	284,902 (48.70)	99,124 (16.94)	5,004 (0.86)
Moderately suitable	10,408 (1.84)	103,217 (18.26)	372,530 (65.91)	79,107 (13.99)
Highly suitable	176 (0.04)	2,574 (0.58)	85,774 (19.45)	352,386 (79.93)

2050 RCP4.5 / Current	Unsuitable (km^2) (%)	Marginally suitable (km^2) (%)	Moderately suitable (km^2) (%)	Highly suitable (km^2) (%)
Unsuitable	7,857,688 (90.62)	525,069 (6.05)	265,413 (3.06)	23,678 (0.27)
Marginally suitable	48,427 (13.35)	85,438 (23.55)	163,123 (44.96)	65,790 (18.14)
Moderately suitable	9,625 (3.5)	27,723 (10.09)	113,225 (41.21)	124,161 (45.20)
Highly suitable	1,359 (0.47)	8,442 (2.90)	79,027 (27.19)	201,813 (69.44)

2070 RCP4.5 / 2050 RCP4.5	Unsuitable (km^2) (%)	Marginally suitable (km^2) (%)	Moderately suitable (km2) (%)	Highly suitable (km^2) (%)
Unsuitable	7,631,893 (96.40)	279,050 (3.52)	6,155 (0.08)	0 (0.00)
Marginally suitable	78,452 (12.13)	462,174 (71.48)	105,951 (16.38)	96 (0.01)
Moderately suitable	5,452 (0.88)	63,295 (10.2)	501,935 (80.85)	50,106 (8.07)
Highly suitable	48 (0.01)	3,789 (0.91)	56,485 (13.60)	355,120 (85.48)

2050 RCP8.5 / Current	Unsuitable (km^2) (%)	Marginally suitable (km^2) (%)	Moderately suitable (km^2) (%)	Highly suitable (km^2) (%)
Unsuitable	7,488,099 (86.35)	824,696 (9.51)	328,948 (3.79)	30,105 (0.35)
Marginally suitable	126,172 (7.21)	75,318(20.76)	195,642 (53.93)	65,646 (18.10)
Moderately suitable	5,564 (2.03)	14,645 (5.33)	126,735 (46.13)	127,790 (46.51)
Highly suitable	719 (0.25)	2,526 (0.87)	38,115 (13.11)	249,281 (85.77)

2070 RCP8.5 / 2050 RCP8.5	Unsuitable (km^2) (%)	Marginally suitable (km^2) (%)	Moderately suitable (km^2) (%)	Highly suitable (km^2) (%)
Unsuitable	7,015,358 (93.28)	493,861 (6.57)	11,335 (0.15)	0 (0.00)
Marginally suitable	99,716 (10.87)	668,751 (72.92)	148,606 (16.20)	112 (0.01)
Moderately suitable	7,290 (1.06)	89,419 (12.82)	551,177 (78.85)	41,552 (7.27)
Highly suitable	1,119 (0.24)	3,613(0.76)	110,315 (23.5)	357,774 (75.50)

Table 4 The distribution range of highly suitable habitats shift under the current and three RCPs by 2050 and 2070

	East	South	West	North
Current	126°16′19.2″	28°20′38.4″	104°15′54.0″	47°38′42.0″
2050 RCP2.6	133°05′27.6″	28°26′45.6″	103°56′52.8″	47°48′06.0″
2050 RCP4.5	133°09′50.4″	28°25′58.8″	103°05′31.2″	47°52′40.8″
2050 RCP8.5	133°07′40.8″	28°23′49.2″	102°47′13.2″	47°56′42.0″
2070 RCP2.6	133°06′14.4″	28°26′13.2″	103°14′24.0″	47°52′01.2″
2070 RCP4.5	133°11′34.8″	28°36′03.6″	103°03′54.0″	47°59′06.0″
2070 RCP8.5	133°04′37.2″	28°26′13.2″	102°48′28.8″	49°34′26.4″