Nest-Site Selection Analysis of Hooded Crane (*Grus monacha*) in Northeastern China Based on a Multivariate Ensemble Model

Author(s): Shengwu Jiao, Yumin Guo, Falk Huettmann and Guangchun Lei
Source: Zoological Science, 31():430-437. 2014.
Published By: Zoological Society of Japan
DOI: http://dx.doi.org/10.2108/zs130248
URL: http://www.bioone.org/doi/full/10.2108/zs130248

ZOOLOGICAL SCIENCE **31**: 430–437 (2014)

Nest-Site Selection Analysis of Hooded Crane (· ·· · ·· · · · · · ·) in Northeastern China Based on a Multivariate Ensemble Model

Shengwu Jiao[1], **Yumin Guo**[1*], **Falk Huettmann**[2], **and Guangchun Lei**[1]

[1]*College of Nature Conservation, Beijing Forestry University, Beijing 100-083, China*
[2]*Institute of Arctic Biology, Biology & Wildlife Department, -EWHALE lab-,*
University of Alaska Fairbanks (UAF) Fairbanks, Alaska 99775, USA

Avian nest-site selection is an important research and management subject. The hooded crane (· ·· ·· · · · · · ·) is a vulnerable (VU) species according to the IUCN Red List. Here, we present the first long-term Chinese legacy nest data for this species (1993–2010) with publicly available metadata. Further, we provide the first study that reports findings on multivariate nest habitat preference using such long-term field data for this species. Our work was carried out in Northeastern China, where we found and measured 24 nests and 81 randomly selected control plots and their environmental parameters in a vast landscape. We used machine learning (stochastic boosted regression trees) to quantify nest selection. Our analysis further included varclust (R Hmisc) and (TreeNet) to address statistical correlations and two-way interactions. We found that from an initial list of 14 measured field variables, water area (+), water depth (+) and shrub coverage (–) were the main explanatory variables that contributed to hooded crane nest-site selection. Agricultural sites played a smaller role in the selection of these nests. Our results are important for the conservation management of cranes all over East Asia and constitute a defensible and quantitative basis for predictive models.

Key words: hooded crane (*Grus monacha*), nest-site selection, boosted regression trees (TreeNet), Northeastern China

INTRODUCTION

Nest-site selection of birds is closely related to their reproductive success and demography (Cody, 1981; Robertson, 1995). A detailed understanding of habitat is essential for endangered species (Li et al, 2012), and it is a classic item of management-related information in ornithology and for general wildlife and landscape management (Forman, 1995; Manly et al., 2002; Braun, 2005; Cushman and Huettmann, 2010). Nests must provide protection for adult birds, eggs and nestlings. Dense vegetation and camouflaged eggs can offer good protection for ground-nesting birds and their eggs (Colwell et al., 2011). Classic studies on this issue demonstrate that food limitation and interspecific competition are among the major underlying factors in nest-site selection (Lack, 1948; Macarthur, 1958). Further, Martin (1993) noted that predation is also an important factor, and breeding birds choose nests next to other potential nest-sites to reduce the risk of predation.

The hooded crane (*Grus monacha*, taxonomic serial number TSN 176186, Avibase ID 38F36091DBC85095) is a vulnerable (VU) species according to the IUCN Red List. The estimated world population of this species is 11,160 individuals, c. 10,500 of which winter in Izumi, Japan (Birdlife International 2014).. The hooded crane breeds in landscapes of Eastern Russia and Northeastern China (Neufeldt, 1977; Guo, 2005; Simonov and Dahmer, 2008). It generally nests in forest swamps, mostly within the permafrost zone (Pukinskiy and Ilyinskii, 1977; Pukinskiy et al., 1982). Understanding the nest-site habitat and selection of cranes is critical for their successful habitat and landscape management. Numerous studies of the related sandhill crane (*Grus canadensis*), red-crowned crane (*G. japonensis*) and Eurasian crane (*G. grus*) species have confirmed (Depkin et al., 1994; Li et al., 1999; Leito et al., 2005). However, relatively little hooded crane research has been conducted, and detailed studies of their nesting landscape are lacking. Only a small amount of quantitative landscape research has been conducted. To date, this mostly refers only to the nest habitat of this species in parts of Russia and China (Pukinskiy and Ilyinskii, 1977; Pukinskiy et al., 1982; Guo et al., 2005a, b). There may be hooded crane breeding pairs in the Ussuri River in Heilongjiang, and at the Hulun and Beier Lakes in Inner Mongolia (Cheng, 1987). However, confirmed evidence of nesting is lacking in these areas. A wider and more holistic landscape scale perspective is missing with regard to this species' breeding in the international zone of the Heilong–Amur River (Simonov and Dahmer, 2008).

To ensure the conservation success of this migratory species, conservation should be based on science-based management (Braun, 2005) on a landscape scale (Forman, 1995; Drew et al., 2011). It is extremely important to know the latest (best) nest descriptions, and then show how these cranes select their nest habitat in those landscapes and for

---

* Corresponding author. Tel. : +86-010-62336065;
Fax : +86-010-62336065;
E-mail: guoyumin@bjfu.edu.cn

doi:10.2108/zs130248

reproductive success. If this can be generalized with a model-prediction, then it can be tested by others for mutual agreement and institutionalized for successful implementation. Here, our objective was to present a robust and first investigation and analysis based on 24 nests that were found during eight years 2003 to 2010. This study presents new information and provides a legacy nest data set in China. A second objective was to report for the first time the quantitative nest habitat preferences of the hooded crane using the latest computational analysis methods to establish a coherent and robust picture. In the present work we chose the statistical method and quantitative concept of boosting (Breiman, 2001a; for a pluralistic view see Stephens et al., 2005) to overcome inherent mathematical limitations that traditional analysis carries for such landscapes (Hastie et al., 2001; Drew et al., 2011; Oppel et al., 2012).

MATERIALS AND METHODS

Study areas

This study was conducted at two nature reserves and their surrounding landscape areas in Heilongjiang Province, Northeastern China. The study region belongs to the Lesser Khingan Mountain Range landscape (Fig. 1). Wu (1980) shows that forest wetlands relevant to hooded cranes are distributed north of the Greater Khingan Mountain and in the transition zone of the Greater and Lesser Khingan Mountains. The landscape under study consists of typical forest swamp areas with permafrost islands, and is located where a seasonally frozen soil layer develops. The first study site is the Greater Zhanhe Wetland National Nature Reserve (127°57′E, 48°23′N), with an area of 211,618 ha; the second site is the Xinqing Hooded Crane National Nature Reserve (130°13′E, 48°30′N) with an area of 62,576 ha. The distance between the two study sites (Greater Zhanhe and Xinqing) is approximately 150 km, and they are treated as landscape replicates for the purposes of this study. The topography of the study areas consists of flat rolling hills with altitudes of 197–480 m above sea level. This region is subject to a temperate monsoon climate. Forest and marsh are the dominant habitat types, and red deer (*Cervus elaphus*), Siberian roe deer (*Capreolus pygargus*) and the Asian grass frog (*Rana chensinensis*) are commonly found. Some small, scattered villages and towns, which rely on logging and farming, are distributed in the area.

Nest searching

Nest searching began at Greater Zhanhe during the breeding season (middle April to early May) and it covered the years 2003 to 2010, whereas at Xinqing, the surveys were conducted from 2006 to 2010. For searching routes and potential habitats we followed the experience of local forest workers and farmers who had encountered hooded cranes. In addition, playback calls of hooded cranes were used to elicit reactions and responses from present wild cranes. This innovative method allows for surveys in large areas in a reliable fashion (see Braun, 2005 for such robust methods). Based on the call of "alerted cranes," or incubating cranes at their nests, it allowed us to locate their nest locations and also the actual birds; a GPS was used for geo-referencing (decimal latitude and longitude; geographic datum WGS84).

Nest data collection

Once a nest was found, we collected plot description data that were robust and quick to measure. The nests were left again as soon as possible to reduce any interference with incubation. The other data we report were actually gathered immediately after breeding. All data were measured on a quadrate of 10 m by 10 m centered on the nest (Table 1).

For our analysis, we intended to pursue a robust and basic comparison of used vs. available sites (Manly et al., 2002; Braun, 2005). Therefore, we sampled 'confirmed absence' in a random fashion within the available nesting landscape. This should allow us to obtain a general estimate of habitat preference (Johnson, 1980;

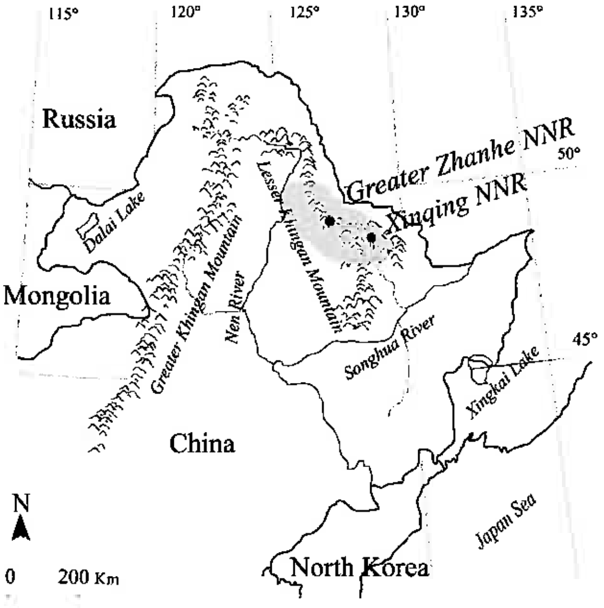

Fig. 1. Location of Greater Zhanhe and Xinqing, in the Heilongjiang province, China.

Table 1. Explanatory nest habitat descriptors for analyzing crane nest preferences.

Variables	Measurement units	Mean ± SE	
		Nests	Control points
elevation	Meters	378.38 ± 12.04	376.60 ± 6.29
aspect	Degrees ▲	2.38 ± 0.26	2.06 ± 0.09
slope position	Degrees slope ◆	2.21 ± 0.13	2.53 ± 0.12
distance to the nearest tree	Meters	3.375 ± 0.43	21.59 ± 8.00
distance to the nearest road	Meters	2503.75 ± 443.47	1789.75 ± 191.30
distance to the nearest human settlement	Meters	5738.75 ± 569.35	3672.59 ± 271.22
distance to the nearest feeding site	Meters	3295.83 ± 437.35	2628.64 ± 207.02
distance to the nearest skidding road	Meters	300.83 ± 77.36	282.04 ± 26.41
water surface area around the nest	Square Meters	53.33 ± 2.99	5.56 ± 1.17
average water depth around the nest	Centimeter	15.08 ± 1.09	3.53 ± 1.16
canopy coverage	Percent	0.35 ± 0.05	0.35 ± 0.03
shrub coverage	Percent	0.07 ± 0.01	0.31 ± 0.03
grass coverage	Percent	0.55 ± 0.04	0.60 ± 0.02
number of trees	Count	4.42 ± 1.05	7.59 ± 0.97

▲ 1-east, 2-south, 3-west, 4-north; ◆ranked: 1-up slope, 2-middle slope, 3-down slope.

Resource Selection Functions RSF Type II, Manly et al., 2002; Braun, 2005; see also, Jones 2001 for terminology we followed). This study is basically a landscape-scale investigation because nests were searched and located in the wider regions representative of the Amur region (Russia and China; Simonov and Dahmer, 2008). Secondly, the control plots were selected randomly within a 5 km range around the nests, and then a high number, 81, of landscape-scale control points were randomly selected in the vast landscape (except for points that were located on ponds, intense farmlands or roads). The control plots of 10 m by 10 m were centered at the 81 plots, where we measured the same parameters as at the nest sites.

Data analysis

Our samples were independent landscape samples and entirely driven by nest locations, as based on nest finding surveys. Our data were not affected much by autocorrelation, as the nest data were collected across different years and at different regions; only a few nests were from the same year. The nearest distance between nests during the same year was 2.4 km (this distance was somewhat closer for locations found during different years). While we could not entirely exclude whether the same individual was sampled twice over the two-year period, we had new samples each year and used those for an underlying analysis of nest selection with non-parametric boosted regression trees (not linear regression lines suffering from the need to meet parametric assumptions; Breiman, 2001a). This is a valid approach, as each year the landscape changes and birds must make new assessments and decisions regarding their site selection for each nest. All data were assembled in a digital database, described with metadata and put in a public institutional repository (dSPACE) following best practices (Zuckerberg et al., 2011). First digital data operations were performed in MS Excel. The value of 1 was assigned for nests, and the value of the randomly selected control plots was 0; the collected habitat attributes were appended to the nest and random locations accordingly. Therefore, the following analysis steps were performed:

We used varclust in R to visualize correlations among predictors. TreeNet is used for data analysis, which is a powerful algorithm in the large and deep machine-learning suite (Breiman, 2001a; Hastie et al., 2001; Elith et al., 2006). It presents us with a good explanation for the data pooled across study sites, and when using Cross Entropy approximately 98% of the 'noise' in these data was explained by TreeNet and its configuration. An AUC curve of over 90% was achieved showing that the findings are generalizable from the data. TreeNet is a boosting algorithm from Salford Ltd. (http://www-stat.stanford.edu/~jhf/ftp/stobst.pdf) and a 30-day trial version can be used for free. This algorithm is based on classification and regression trees (CARTs) and related concepts; it is increasingly used in ecological studies due to its great speed, convenience, and high performance (Breiman, 2001a; Elith et al., 2008; Oppel et al., 2009; Drew et al., 2011). Boosting is a non-parametric method basically free of most common assumptions (Breiman, 2001a; Hastie et al., 2001). The method is not significantly influenced by correlations, for instance. Boosting is part of machine learning and has been shown to perform very well in complex data analysis (Elith et al., 2006; Oppel et al., 2012). It is used in many landscape and ecological applications and its concept and mathematical performance have been well described (Hastie et al., 2001; Elith et al., 2008; Cushman and Huettmann, 2010; Drew et al., 2011). TreeNet is 'stochastic gradient boosting' and in essence represents optimized CARTs in a sequence, with each of the new statistical trees explaining the remaining variance of the previous tree. It also borrows concepts from bagging (Friedman, 1999; Breiman, 2001b), which makes it an ensemble model (Araujo and New, 2007; Hardy et al., 2011), and it maintains a running performance metric to obtain the best prediction. In a confirmatory fashion, we applied TreeNet to the pooled nest data for both study sites. As is commonly the case, our data were not derived from a completely surveyed landscape, but instead were obtained by sampling. However, these constitute the best available data on nests and random plots in China for the hooded crane and form a national data legacy. Here, we followed the Resource Selection Functions RSF Type II (Manly et al., 2002; Braun, 2005) research design. Specifically, the machine learning approach and methods were chosen here to find the 'best signal' in the data, when data are 'messy,' complex, pooled and correlated and when they carry many interactions (Cushman and Huettmann, 2010; Drew et al., 2011). In addition, this method has been widely ignored in nest site and habitat selection studies (but see Popp et al., 2007), and large opportunities still exist for its application and contribution to studies such as this when conducted at a landscape scale and for determining thresholds.

RESULTS

Number of nests

The conservation status of the hooded crane has become much worse since 1993 (when the first hooded crane nest was found in China; Li, 1993). In 2003, we were able to find one nest in the Greater Zhanhe Forest Region. Due to the seclusion of nests and the small number of breeding pairs in the Lesser Khingan Mountain Range, 24 nests were eventually found and made available for our analysis of nest-site selection at two breeding sites through 2010.

Nest habitat characteristics

Open forest swamps are an important breeding area for the hooded crane. These swamps made up a typical landscape mosaic and matrix for this study. They were all surrounded by open water, consisting of either seasonal streams or small pools, and were part of a large watershed and water table. Hummocks played an important role for crane nests. The cranes mainly constructed the nest substrate with herbaceous vegetation and with some branches of trees that were located in the vicinity of the nests. Typical vegetation species around the nests identified by us included *Larix gmelini* and ifor trees, *Spiraea salicifolia* and *Betula ovalifolia* for shrubs, and *Carex schmidtii* and *Calamagrostis angustifelia* for herbs.

Nest-site selection

We feel that for landscape studies it is important to obtain control and background samples characterizing the landscape. Thus, here we compared the 24 nests with 81 random samples of control plots, resulting in a ratio of 24/81 (0.29). Although at first sight this might appear like a mathematical skew and an oversampling of landscape control plots, this is not a real problem in RSF studies, when the landscape is sampled (Manly et al., 2002). It has also been shown to work well for data mining and ensemble model methods, which tend to be quite robust (Craig and Huettmann, 2008) when weighting is used.

When investigating the data with 'pairs' and 'varclust' plots in R (Hmisc library by F. Harrell), Fig. 2 helps to visualize which variables are correlated with each other and to what degree. It shows that the water surface and depth predictors are tightly correlated. Shrub cover belongs to the same group, whereas for instance canopy cover and number of trees are correlated but form an own group. Similar correlations are observed for distance to human settlement

Table 2. Importance ranking (in percent) of explanatory variables using the TreeNet algorithm (stochastic gradient boosting).

Variables	Score (% relative importance)
Water surface around nest	100
Average water depth around nest	49.63
Shrub coverage	26.57
Distance to the nearest human settlement	20.07
Distance to the nearest skidding road	12.08
Distance to the nearest road	8
Number of trees	6.77
Aspect	6.13
Distance to the nearest tree	5.37
Canopy coverage	5.33
Altitude	4.95
Distance to the nearest farmland or feeding site	4.55
Grass coverage	2.71
Slope position	1.15

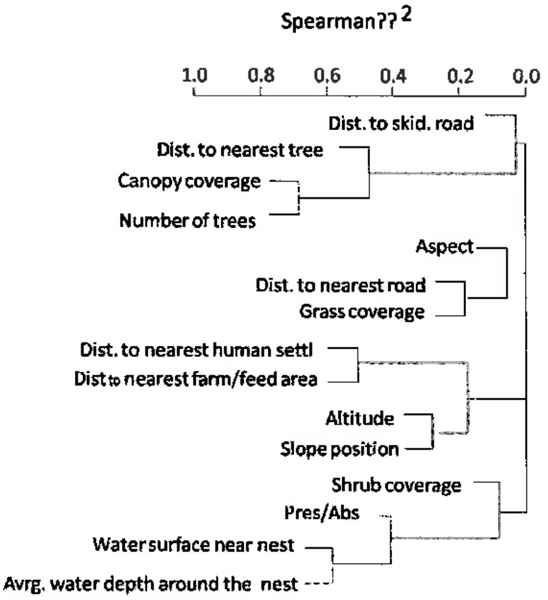

Fig. 2. Varclust of the nest control data for Hooded Cranes in China (pooled analysis across study sites).

and farming and feeding areas.

Figure 3 presents the 'loess' function, which shows how much variance was explained with an increasing number of statistical trees in TreeNet (0 to 400). Overall, TreeNet started on the y-axis with a variance of approximately 0.68 and eventually it drops to 0.2. The training and testing data followed similar patterns as well. This supported our conclusion that the data, the obtained model, and the findings were homogenous in themselves, that findings were relatively free of 'noise' and bias and that they were representative of the dataset.

In boosting, the ranking score was expressed as a percentage (Table 2), and the predictor variable 'water surface area around the nest' ranked the highest (100%, in relative units across all predictors). The next predictor variable 'average water depth around the nest' was ranked in the middle range (49%). Shrub coverage ranked third (26%). All other predictors were ranked lower. We suggest rather than interpreting variables individually, the top three variables should rather be assessed on a multivariate basis.

Boosting is able to produce partial dependence plots. These are essentially resource selection functions (Manly et al., 2002) using 'non-linear functions' in a multivariate predictor context based on TreeNet (stochastic gradient boosting; instead of using classic logistic regressions or similar linear functions as mentioned in the section above). Such figures are powerful as they can show how the environment links with the presence/absence of the crane nest plots in an ecological multivariate setting (Oppel et al., 2009; see Popp et al., 2007), based on the best-available field data for this species. In a machine-learning context, it allows the user to find 'the major signal,' which distinguishes the presence of nests from assumed absences in a complex landscape (as is commonly done in RSF studies, Manly et al., 2001). The y-axis was to be interpreted as an index of the presence/random comparison of nests (thus, in relative units it was expressed as high to low). Figure 4 shows that cranes preferred nest areas with a large open water surface. On the x-axis, a break-through point of 23 m^2 of 'water surface area around the nest' was located and the 'presence' signal of nests disappeared when less water area was available.

The subsequent two variables, average water depth around the nest and shrub coverage, were not as important when compared to 'water surface area around the nest'; all other variables after 'shrub coverage' played a contributing but smaller role and were not pursued here further. The TreeNet model suggested that a multivariate combination of a few variables drives crane nest selection: cranes prefer water areas with c. 23 m^2, a depth of c. 8 cm (Fig. 5), and a shrub cover of < 20% (Fig. 6). Figure 7 shows a two-dimensional habitat preference plot for hooded crane nest preferences using the top predic-

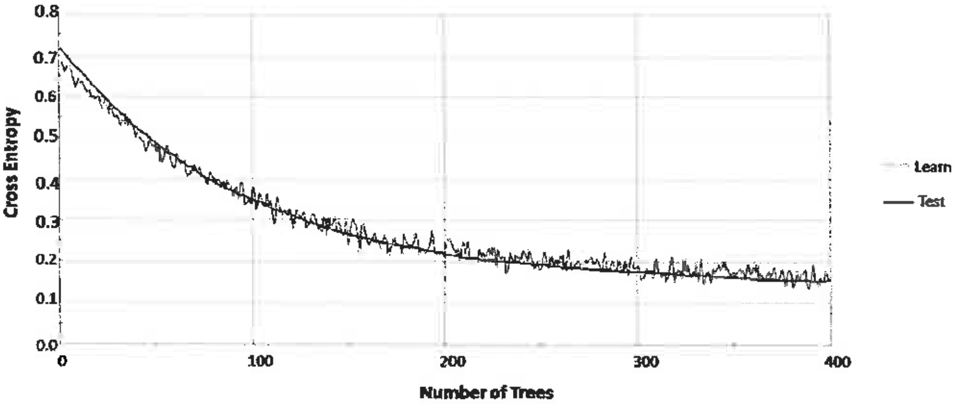

Fig. 3. Cross Entropy of TreeNet algorithm (pooled analysis across study sites).

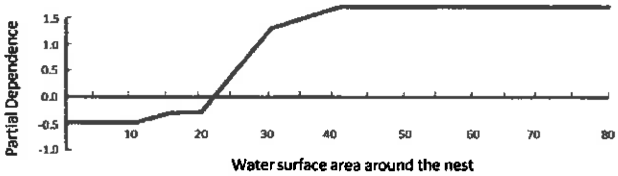

Fig. 4. Partial dependence plot (TreeNet algorithm) for 'water surface area around the nest'.

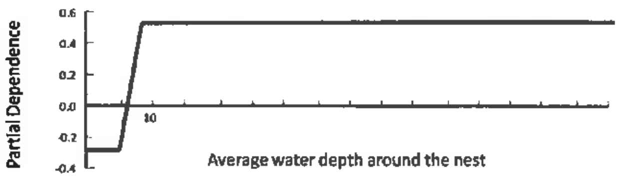

Fig. 5. Partial dependence plot (TreeNet algorithm) for 'average water depth around the nest'.

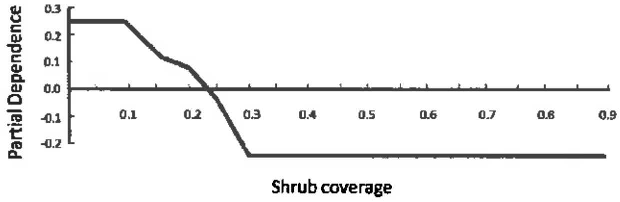

Fig. 6. Partial dependence plot (TreeNet algorithm) for 'shrub coverage'.

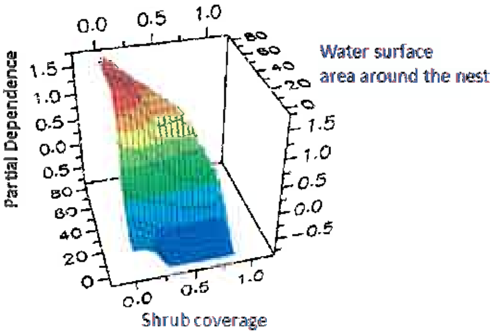

Fig. 7. Two-dimensional Partial Dependence Plot.

tor set. The less important the predictor variables, the lower they were ranked; they started to 'bounce' (this means they have no strong clear signal, do not generalize well anymore, and should not be interpreted; thus, most of them are not shown here).

Further, Treenet allows us to document and to track statistical interactions (Salford Systems Ltd). Table 3 shows the two-way interactions. Shrub coverage, water surface area around the nest and average water depth around the nest ranked highest for interactions and compose a multivariate package to describe crane nests. There were other interactions, but they appeared to be of lower relevance for the predictors we identified in this analysis and for our interpretation.

Table 3. Two-way interaction statistics for Hooded Crane nest selection using the TreeNet algorithm (stochastic gradient boosting).

Variables	Absolute	Relative
Shrub coverage	1.65	100
Water surface around nest	1.16	70.55
Average water depth around nest	0.98	59.37
Distance to the nearest road	0.74	44.98
Altitude	0.39	23.89
Distance to the nearest human settlement	0.3	18.6
Number of trees	0.22	13.65
Distance to the nearest skidding road	0.18	11.29
Distance to the nearest tree	0.14	8.99
Canopy coverage	0.13	7.88
Distance to the nearest farmland or feeding site	0.07	4.64
Slope position	0.68	4.15
Aspect	0.56	3.42
Grass coverage	0.01	0.87

DISCUSSION

We describe for the first time newly found nests of hooded crane in Chinese landscapes. This information constitutes a Chinese legacy data set covering eight years of fieldwork. Further, we have used these data to determine the nest preference of the hooded crane in a quantitative fashion on a landscape scale. Our study followed an analysis method that accounts for habitat preference (Manly et al., 2002; Braun, 2005). However, we presented a general analysis template that is robust enough to handle any 'messy' data beyond the assumptions and constraints of frequency statistics (see for instance O'Connor et al., 1996). For this reason, we went beyond traditional methods to utilize machine learning in a landscape ecology analysis of complex crane data. We compared nest sites with confirmed absences from random samples conducted nearby in the landscapes. This helped to find and describe a 'pure' signal and trend from empirical field data in landscapes where cranes select nest sites, instead of simply describing nest locations that were affected by other factors and were analyzed with little or no context.

Our study may suffer from biases in nest finding and detection ability, as well as from the fact that the random control points for comparison were from a pooled analysis and did not cover ponds, farmland or roads. However, we believe this results in only a small bias (if any), and we have used an analytical approach to extract trends, even from messy data. Instead, excluding meaningless comparison sites made the findings stronger because better comparisons were made. Here we present the best and most long-term field data available for a species of international conservation concern and used a combined suite of the best available research methods to determine the species' nesting habitat. Relative to the major breeding area in Siberia (though there is no research about nest selection of the hooded crane at present), only a small population breeds in our research sites. Thirty individuals were counted during an earlier investigation at Zhanhe valley (Guo et al., 2005b). For this reason, only 24 nests were found in the 8-year period. As a result, this study also contributed to the conservation of the species and habitat, by focusing on these fragmented populations.

Based on the results of the analytical concepts of this study, water was rather important for the nest habitat of hooded crane. With the exception of demoiselle crane (*Anthropoides virgo*) nests located in crop fields in Kazakhstan and Central Asia, and blue crane (*Anthropoides paradiseus*) nests found on dry ground, all other crane species nest in shallow water areas (Johnsgard, 1983). Some standing water is necessary for nesting sandhill cranes. Otherwise, the cranes are known to delay their breeding behavior (Layne, 1981). Guo (2005) infers that water area around the nests is one of the main influential factors for the nest habitat of hooded cranes (based on a cluster analysis at Greater Zhanhe region). The presence of a water-body is a necessary requirement for breeding hooded cranes and for providing food (e.g., *Rana limnocharis*, *Salamandrella keyserlingii*) for nestlings. The available water area around the nests can determine the growth of chicks. Certain water depths (15.08 ± 1.042) around nests also can protect eggs or nestlings against some small mammals (e.g., *Capreolus pygargus*, *Sus scrofa*). Hooded cranes have difficulty wading in deep water, e.g., deeper than 24 cm. This maximum wading depth is also similar to those for other cranes found in the same region, such as 0–20 cm for the white-naped crane and 0–15 cm for the Eurasian crane (Bradter et al., 2005; Leito et al., 2005). During stochastic events, such as flooding (Guo, 2005), hooded cranes have been described as preferring flat areas as nest-sites where water accumulated temporarily, forming a small pool during the rainy season. But nests suffer from a risk of flooding. Such a peculiar situation also occurs with sandhill cranes (Depkin et al., 1994).

In addition to the finding that water affected the nest-site selection of Hooded Cranes as with most other cranes, some special characteristics were also found in this study due to the surrounding environment. Open shrub coverage < 20% was one of the important predictors. Suitable shrub coverage contributes to the breeding success of cranes. On one hand, when the canopy coverage is too dense, it makes it difficult to escape by flight in an emergency and when stressed; on the other hand, a canopy coverage that is extremely sparse easily exposes the nests of hooded cranes to dangers (Meine and Archibald, 1996; Guo, 2005).

Nest fidelity is very common with whooping cranes (*Grus americana*; Parasharya, 1989). But of four known hooded crane nests, only one was reused in the Primorye District of the Bikin River Basin (Pukinskiy and Ilyinskii, 1977). Of our 24 nests, though no nest was re-used, a crane pair did choose another grass hummock in the same pool as its base (Guo, 2005). Based on opportunistic field video monitoring, which had been set up to monitor around the nests, it was found that two cranes actually went back to their nests, but they did not ultimately nest at those sites. We speculate that such low nest fidelity in our study site may be related to logging in winter, which changed the landscape habitat and represents a disturbance to these birds.

It is worthwhile to note the longer list of variables that cannot be determined well in TreeNet. While several of these were actually found in the literature about crane nest selection, even the powerful machine-learning algorithm TreeNet cannot identify strong trends or a relevant explanatory power. We thought that this set of variables was of less relevance when taken together in a more ecological landscape prediction context. Instead, TreeNet and similar machine learning algorithms are known for their high power to detect signals and trends, even from multivariate and 'messy' data (Elith et al., 2006; Craig and Huettmann, 2008). If such trends cannot be detected from such predictors (as we found here), it shows that the best data available do not lend themselves to confirming such earlier findings. Consequently, earlier studies and findings that promoted these predictors for crane nests should be scrutinized when not in agreement with Figs. 2–6 and when conducted at a similar scale.

In our results, water area can provide food for chick cranes; water depth can keep some mammals out of the nests; the shrub can protect some predators against nests. These support some classic theories about nest selection. (Lack, 1948; Macarthur, 1958; Martin, 1993). Our use of a new analysis method added to a more complete, robust, state-of-the-art, and better ecological description of nest-sites. Such multivariate, analytical and pluralistic approaches are demanded (Stephens et al., 2005; see also Worm and Myer, 2003; Araújo and New, 2007) and are becoming increasingly common (Drew et al., 2011; Oppel et al., 2012) with an increase in computing power (Breiman, 2001a; Cushman and Huettmann, 2010). However, these methods are still not sufficiently applied and considerable potential remains for them to be used with GIS and on a landscape-scale (O'Connor, 2000). Our hooded crane nest preference study, for instance, is the first of its type using machine learning. Here, we have shown with a long-term Chinese legacy data set that such work can be performed relatively easily and that it adds to our understanding of species of conservation concern. Until better methods are shown, we recommend this type of analysis as a rapid method of choice ('best professional practice') for a more ecological and well-rounded conclusion based on empirical and transparent, high-quality data (Huettmann, 2005; Ohse et al., 2009; Zuckerberg et al., 2011) that is virtually free of bias for species of such wide conservation relevance as the hooded crane. In our study, we analyzed the species' nest-site selection pattern at many microhabitat sites that are distributed throughout the entire landscape in Northeast China. We think that this allows for a first landscape assessment in China and beyond. However, for a large and wide-ranging species such as the hooded crane, and for their flyways, there may be additional nest-site selection patterns on a macro-scale in the international landscapes in Russia, China (the Amur River basin), and Japan (wintering sites), taking into account connecting sites. This can be assessed using remote sensing (Gottschalk et al., 2005) which requires further modelling (Cushman and Huettmann, 2010; Drew et al., 2011). This subject will be a worthwhile research topic for the future, and also will feed into wider management decisions and policies for this vulnerable species on the global flyway-scale. Finally, we also need to concern that: Whether the Far East in Russia or Northeastern China, the two breeding sites for Cranes belong to Permafrost area. Because of global warming, the south limit has moved north more than 50 km, and the total area reduced ~35% in Greater and Lesser Khingan Mountain Ranges (Jin et al., 2006). The possible influence is breeding habitat degradation of hooded cranes.

ACKNOWLEDGMENTS

Many people and institutions kindly contributed to this study. We would like to thank the Administration Bureau of Greater Zhanhe National Nature Reserve and Xinqing National Nature Reserve for their assistance in our field survey and data collection. We also would like to acknowledge the EWHALE lab students for input, suggestions, discussions, and work in the Far East. We are grateful for the software support from Salford Systems Ltd (Dan Steinberg and his great team). This project was supported by the National Natural Science Foundation of China (Project NO. 30770287), the Whitley Fund for Nature (WFN), the State Forestry Administration of China and Grus Monacha International Aid (GMIA). The experiments comply with the current laws of China in which they were performed. This is EWHALE lab publication # 130.

REFERENCES

Araújo MB, New M (2007) Ensemble forecasting and species distributions. Trends Ecol Evol 22: 42–47

BirdLife International (2014) IUCN Red List for birds. Downloaded from http://www.birdlife.org on 09/02/2014

Bradter U, Gombobaatar S, Uuganbayar C, Grazia TE, Klaus-Michael Exo (2005) Reproductive performance and nest-site selection of White-naped Cranes *Grus vipio* in the Ulz river valley, north-eastern Mongolia. Bird Conserv Int 15: 313–326

Braun C (2005) Techniques for Wildlife Investigations and Management, 6th ed. The Wildlife Society, Bethesda, Maryland, USA

Breiman L (2001a) Statistical Modeling: The Two Cultures. Stat Sci 16: 199–231

Breiman L (2001b) Random forests. Mach Learn 5–32

Cheng T (1987) A Synopsis of the Avifauna of China. Beijing, Science Press

Cody ML (1981) Habitat selection in birds: The roles of vegetation structure, competitors and productivity. Bioscience 31: 107–113

Colwell MA, Meyer JJ, Hardy MA, Mcallister SE, Transou AN, Levalley RR, Dinsmore SJ (2011) Western Snowy Plovers *Charadrius alexandrinus nivosus* select nesting substrates that enhance egg crypsis and improve nest survival. Ibis 153: 303–311

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 "Intelligent Data Analysis: Developing New Methodologies through Pattern Discovery and Recovery" Ed by Wang Hsiao-fan, IGI Global, Hershey, PA, USA

Cushman S, Huettmann F (2010) Spatial Complexity, Informatics and Wildlife Conservation. Springer Tokyo, Japan, 448

Depkin FC, Brandt LA, Mazzotti JF (1994) Nest sites of Florida Sandhill Cranes in southwestern Florida. Fla Field Naturalist 22: 39–47

Drew AS, Wiersma YF, Huettmann F (2011) Predictive Species and Habitat Modeling in Landscape Ecology: Concepts and Applications. Springer, New York

Elith J, Graham C, NCEAS working group (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29: 129–151

Elith J, Leathwick JR, Hastie T (2008) A working guide to boosted regression trees. J Anim Ecol 77: 802–813

Forman RT (1995) Land Mosaics: The Ecology of Landscapes and Regions. Cambridge University Press

Friedman JH (1999) Greedy Function Approximation: a Gradient Boosting Machine. Technical report, Dept. of Statistics, Stanford University

Gottschalk T, Huettmann F, Ehlers M (2005) Thirty years of analyzing and modelling avian habitat relationships using satellite imagery data: a review. Int J Remote Sens 26: 2631–2656

Guo YM (2005) Breeding Behavioral Ecology of Hooded Crane (*Grus monacha*) in the Lesser Xingan Mountains. Dissertation, Northeast Forestry University

Guo YM, Qi FW, Liu XL, Xu CZ, Ma JZ (2005a) Preliminary report on breeding habit of Hooded Crane *Grus monacha* in Lesser Xingan Mountains. Acta Zool Sinica 51: 903–908

Guo YM, Liu XL, Xu CZ, Qian FW, Li L (2005b) Preliminary investigation of breeding population of Hooded Crane in Lesser Xingan Mountains. Chin J Zool 40: 51–54

Hardy SM, 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. Integrative and Comparative Biology 51: 608–622; doi: 10.1093/icb/icr102

Hastie T, Tibshirani R, Friedman J (2001) The Elements of Statistical Learning in Data Mining, Inference, and Prediction. Springer Series in Statistics, New York

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 Wildlife Manage 69: 466–472

Jin HJ, Yu SP, Lv LZ, Guo DX, Li YW (2006) Degradation of permafrost in the Greater and Lesser Khingan Mountain Ranges Northeast China and preliminary assessment of its trend. J of Glaciology and Geocryology 28: 467–475 (in Chinese)

Johnsgard PA (1983) Cranes of the World. Indiana University Press, Bloomington; Croom Helm, Beckenham, UK

Johnson DH (1980) The Comparison of usage and availability measurements for evaluating resource preference. Ecology 6: 65–71

Jones J (2001) Habitat selection studies in avian ecology: A critical review. Auk 118: 557–562

Lack D (1948) The significance of clutch size. Ibis 90: 25–45

Layne JN (1981) Dry ground nest of Florida Sandhill Cranes. Fla Field Naturalist 10: 55–56

Leito A, Ojaste I, Truu J, Palo A (2005) Nest site selection of the Eurasian Crane *Grus grus* in Estonia: an analysis of nest record cards. Ornis Fennica 82: 44–54

Li F, Yang HJ, Zhang HH, Gao ZX (1999) Nest-site selection of Red-crowned Crane in the Zhalong Wetland. J Northeast Forest Univ 27: 57–60

Li HQ, Yue BS, Lian ZM, Zhao HF, Zhao DL, Xiao XM (2012) Modeling spring habitat requirements of the endangered Brown Eared Pheasant *Crossoptilon mantchuricum* in the Huanglong Mountains, Shaanxi Province, China. Zool Sci 29: 593–598

Li L (1993) The first breeding area of Hooded Crane in China. Chin J Wildlife 14: 16

MacArthur RH (1958) Population ecology of some warblers in northeastern coniferous forests. Ecology 39: 599–619

Manly B, McDonald LL, Thomas DL, McDonald TL, Erickson WP (2002) Resource selection by animals: statistical design and analysis for field studies, 2nd ed. Dordrecht, the Netherlands

Martin TE (1993) Nest predation and nest sites. Bioscience 43: 523–532

Meine CD, Archibald GW (1996) The cranes: status survey and conservation action plan. IUCN, Gland, Switzerland, and Cambridge, UK

Neufeldt I (1977) The distribution of the Hooded Crane Grus monacha in light of the available information. Ornitologiya 13: 56–61

O'Connor R (2000) Why ecology lags behind biology. Scientist 14: 35–37

O'Connor RJ, Jones MT, White D, Hunsaker CT, Loveland T, Jones B, Peston E (1996) Spatial partitioning of environmental correlates of avian biodiversity in the conterminous United States.

Biodivers Lett 3: 97–110

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, Strobel C, Huettmann F (2009) Alternative Methods to Quantify Variable Importance in Ecology. Technical Report Number 65, Department of Statistics, University of Munich

Oppel S, Meirinho A, Ramírez I, Gardner B, O'Connell AF, Miller PI, Louzao M (2012) Comparison of five modelling techniques to predict the spatial distribution and abundance of seabirds. Biol Conserv 156: 94–104

Popp J, Neubauer D, Paciulli L, Huettmann F (2007) Using TreeNet for identifying management thresholds of mantled howling monkeys' habitat preferences on Ometepe Island, Nicaragua, on a tree and home range scale. J Med Biol Sci 1: 1–14

Pukinskiy YB, Ilyinskii IV (1977) On the biology and behaviour of Hooded Crane in the breeding season, Primorye District, Bikin River Basin. Byulleten Moskovskogo Obshchestva Zspytatelei Prirody Otdelenie Biologicheskii 82: 5–17

Pukinskiy YB, Ilyinskii IV, Shibnev YB (1982) The number and distribution of Hooded Crane in Bikin River Basin. In "Cranes of East Asia. Far East Science Centre" Ed by NM Litvinenko, IA Neufeldt, Academy of Sciences of the USSR, Vladivostok, pp 44–48

Robertson GJ (1995) Factors affecting nest site selection and nesting success in the Common Eider *Somateria mollissima*. Ibis 137: 109–115

Simonov EA, Dahmer T (2008) Amur-Heilong River Basin Reader. Ecosystems Limited, Hong Kong

Stephens PA, Buskirk SW, Hayward G, Martinez DRC (2005) Information theory and hypothesis testing: a call for pluralism. J Appl Ecol 42: 4–12

Worm B, Myers RA (2003) Meta-analysis of cod-shrimp interactions reveals top-down control in oceanic food webs. Ecology 84: 162–173

Wu ZM (1980) Vegetation of China. Beijing, Science Press

Zuckerberg B, Huettmann F, Friar J (2011) Proper data management as a scientific foundation for reliable species distribution modeling. In "Predictive Modeling in Landscape Ecology" Ed by CA Ashton, Y Wiersma, F Huettmann, Springer, New York

(Received November 28, 2013 / Accepted February 24, 2014)