Modeling the predicted suitable habitat distribution of Javan hawk-eagle *Nisaetus bartelsi* in the Java Island, Indonesia

**ILYAS NURSAMSI¹, RUHYAT PARTASASMITA²*, NURVITA CUNDANINGSIH¹, HASNA SILMI RAMADHANI¹**

²Department of Biology, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran. Jl. Raya Bandung-Sumedang Km 21, Jatinangor, Sumedang 45363, West Java, Indonesia. Tel.: +62-22-7796412, *email: rp2010rikkyo@gmail.com; ruhyat.partasasmita@unpad.ac.id*


**Abstract.** Nursamsi I, Partasasmita R, Cundaningsih N, Ramadhani HS. 2018. Modeling the predicted suitable habitat distribution of Javan hawk-eagle *Nisaetus bartelsi* in the Java Island, Indonesia. *Biodiversitas* 19: 1539-1551. Javan hawk-eagle (*Nisaetus bartelsi*) is an endemic raptor of Java Island. The conservation status of Javan hawk-eagle (JHE) according to IUCN is endangered (EN) and included in CITES Appendix II list, and this species is also protected by the Indonesian government law based on act no. 5, year 1990. The position of Javan hawk-eagle as a top predator is now very threatened by habitat fragmentation, wildlife trade, and the declining quality of its habitat. The primary purpose of this study was to give preliminary information about the distribution of predicted suitable habitat for JHE as a means of finding potential releasing sites, as an evaluation for habitat protection, and even as an option for the quality of its habitat. The primary purpose of this study was to give preliminary information about the distribution of predicted suitable habitat for JHE as a means of finding potential releasing sites, as an evaluation for habitat protection, and even as an option for the development of new JHE protected areas. However, mapping the spatial distribution of potential habitat for JHE using terrestrial survey is problematic because it requires enormous time, fund, and human resources. The most possible approach is by using Ecological Niche Modeling (ENM)/species distribution modeling (SDM). In this study, modeling exercise was conducted by using a maximum entropy method as an adaptation from Maxent software ver. 3.4.1, with the utilization of JHE-nest coordinate data and 16 environmental variables datasets as the main input. The predicted suitable habitat distribution map has shown a good match with historical and present records of JHE and has fairly succeeded in capturing a wide range of habitat patches from tiny spots to quite large suitable habitat. Modeling results also showed that altitude, annual mean temperature, and two types of land cover (closed shrub, and forest area) are considered to be most important variables affecting the distribution of potential habitat for JHE. Moreover, about 17.77% (23,209 km²) area of Java Island has been projected to be suitable for Havan Hawk-Eagle's habitat, which mostly spread in mountainous areas while also appear in several lowland areas. This study suggests the importance of topographic, climatic, and land cover as pivotal predictors in determining the suitability of habitat for JHE. This study also shows that the modeling results have a good match with the historical records of JHE across the island, which suggests the overall accuracy of the model.

**Keywords:** Distribution habitat modelling, ecological niche modelling, Javan hawk-eagle, maximum entropy, *Nisaetus bartelsi*

**INTRODUCTION**

Biogeographic regions with a significant reservoir of biodiversity, known by the term “biodiversity hotspots,” cover only 2.3% of the earth’s land surface but host 42% of the world’s vertebrates (Jha and Bawa 2006). Currently, none of these hotspots have more than one-third of its pristine habitat left, and all these whole regions face deforestation threat caused by human population growth and development; this pressure being exceptionally high in the tropical regions (Brooks et al. 2002). The tropical rainforests in Sundaland are part of the biodiversity hotspot. Sundaland is one part of the country of Indonesia with a high diversity of bird species, with the highest rates of deforestation in East Asia (Giam et al. 2011). The high biodiversity in Java is concentrated only in certain areas, and this is because it is strongly influenced by the extreme population growth and deforestation (Miettinen et al. 2011; Partasasmita 2009; Safana et al. 2018).

The response of species in areas threatened by deforestation is important to know (John and Skorupa 1987). While some species may be able to positively adapt to landscape changes made by human (Sodhi et al. 2010; Partasasmita et al. 2009), for most, the landscape changes that lead to decreased quality of habitat adversely affect the survival of the species (Fahrig 2007). In the effort to develop a conservation plan for the species threatened by deforestation and changes in habitat quality, the collection of data on their habitat preferences, as well as the availability of suitable habitats in an area, is vital (John and Skorpura 1987). Conservation methods to assess the suitability of habitat for the species has grown to be able to build a model which relates species distribution to environmental characteristics (Guisan et al. 2006). The model could be a useful tool in the selection of protected areas for maximizing biodiversity conservation (Rodriguez et al. 2007).

Javan hawk-eagle (*Nisaetus bartelsi*) is an endemic raptor of Java Island which is currently severely threatened by both wildlife trade and habitat loss. Javan hawk-eagle (JHE) became one of the most endangered raptors in Indonesia (Rakhman 2012; Partasasmita et al. 2016);
consequently, Javan hawk-eagles are classified as ‘Endangered’ (EN) on the IUCN Red List and listed in Appendix II of CITES. Despite Javan hawk-eagles facing a severe threat to its existence, our knowledge of their habitat preferences and the distribution of its habitats in Java Island is limited. Various attempts to mapping the habitat distribution of Javan hawk-eagles habitats were limited only to recording occupied habitat; therefore data about habitat suitability for Javan hawk-eagle are often neglected. However, collecting these data through direct field surveys will require enormous human resources, funding, and time, then a different approach is needed.

In the last few decades, attention toward understanding the characteristics of preferred habitat of certain species and the distribution of potentially suitable habitat leads to a marked increase of interest in the use of Ecological Niche Modeling (ENM) (Merow et al. 2013; Fourcade et al. 2014). ENM, also known as Species Distribution Modeling (SDM), which was developed in the mid-1980s (Booth et al. 2014), comprehensively involving the utilization of statistic, ecology, Geographic Information System (GIS), and even Remote Sensing (RS) to develop estimation of suitable niche for species across predefined landscapes (Franklin and Miller 2009), while also can be extrapolated through different space and time (Guisan and Thuiller 2005; Elith and Leathwick 2009; Franklin and Miller 2009). This modeling will be useful for JHE conservation efforts as a means of finding potential releasing sites, as an evaluation for habitat protection, and even as an option for the development of new JHE protected areas.

Such modeling can be conducted using a variety of alternatives, including heuristic models (e.g., BIOCLIM—Beaumont and Hughes 2007), combinatorial optimization (e.g. GARP-Fitzpatrick et al. 2007), statistical models (e.g. GAMs-Jensen et al. 2008), and machine learning (e.g. ANN—Ostendorf et al. 2001; Berry et al. 2002; Harrison et al. 2006; MAXENT—Phillips et al. 2006) (Sinclair et al. 2010). Each of these approaches, indeed, has their advantages and disadvantages. Nevertheless, one of the most growing approaches of ENM is through the use of Maximum Entropy (Maxent) algorithms (Belgacem and Louhaichi 2013). Maxent modeling has a high potential for identifying distributions and habitat selection of wildlife given its reliance on only presence locations and has shown higher predictive accuracy than many other methods (Phillips et al. 2006; Baldwin 2009; Franklin and Miller 2009; Elith and Frankling 2013; Peterson et al. 2011; Remya et al. 2015). Being a general-purpose machine learning method, Maxent offers a precise and straightforward mathematical formulation to characterize probability distribution across a user-defined landscape (Phillips et al. 2006; Merow et al. 2013). Maxent is a software package which was developed for ENM/SDM given presence-only species records (without the need of absence data) and a “background” sample of environments in the region of interest (Phillips et al. 2006; Phillips and Dudik 2008). Therefore, we refer to this type of data requirements as “Presence-Background” (PB) data.

In general, Maxent works by applying the maximum entropy principle (Jaynes 1957) to fit the model so that the estimated distribution diverges from a uniform distribution as minimally as required to explain the observations (Guillera-Arroita 2014). Elith et al. (2011) have explained this software intuitively from a statistical point of view by looking at the environmental domain, even though the algorithm within this software works in geographic space (Phillips et al. 2006). Herewith, we try to follow this interpretation to explain this software briefly. PB data are used to obtain a set of environmental characteristics at the presence site of species and at background location of study (which are a regular or random sample of the landscape or could be intended to match the recognized biases in the sampling process). Otherwise stated, Maxent examines the ratio of $f_1/f$, where $f_1$ is the probability distribution defining the characteristic of sites wherein the species occurs, and $f$ is the probability distribution depicting the environmental attributes of the sites wherein the species is absent ($y = 0$) and present ($y = 1$). This ratio corresponds to the probability of presence given the environmental covariates $z$, thus $Pr(y = 1|x) = \psi(z)$, but the scaling factor (the prevalence of the species over the landscape, $Pr(y = 1)$) cannot be identified from PB data only (Elith et al. 2011; Hatic and Fithian 2013; Phillips and Elith 2013; Guillera-Arroita 2014). The simple output or the ‘raw output’ of Maxent is the estimate of $f_1/f$ scaled to sum to 1 over the fitted background and representing the probability distribution over sites $x$, given the species exist at site $x; Pr(x|y = 1)$ (Phillips et al. 2006). Therefore, as currently implemented, that raw output represents $\psi(z)/(n.\psi)$, where $n$ is the total number of points in the background sample. Nonetheless, the raw output only represents relative suitability, as it is proportional to inhabitancy probabilities by a factor that is not identifiable without external data (Guillera-Arroita 2014).

With the purpose of estimating the ratio $f_1/f$, Maxent fits an exponential model so that $\log((f_1/f) = \eta(z)$, where $\eta(z)$ is a linear term of a set of features. Features, which in Maxent can belong to six classes (i.e., linear, product, quadratic, hinge, threshold, and categorical), represent a broader set of transformations of the covariates (Phillips and Dudik 2008). Moreover, features offer a great flexibility to fit complex environmental relationships, and Maxent utilizes regulation to control the trade-off between model complexity and model fit as a means to avoid overfitting (Guillera-Arroita 2014). Nevertheless, instead of working with Maxent’s “raw output,” it is more common for users to work with "logistic output." Using a logistic transformation which based on a user-specified parameter ‘$\tau’,” this output scales the raw values into a scale of relative suitability ranging between 0 and 1 (Elith et al. 2011). This parameter is the prevalence for sites with ‘average’ environmental conditions under $f_1$ and by default is set to an arbitrary value ($\tau = 0.5$) (Elith et al. 2011). Therefore, the output represents the probability of the species occurs on each location/sites or can be interpreted as a probability of suitable habitat for the species based on the environmental variables included in the model.

This study was then conducted as an attempt to model the distribution of potential habitat for JHE across the Java Island, while also tried to understand the environmental characteristics of preferred habitat based on the available

$\Pr (y = 1|\mathbf{z}) = \psi(\mathbf{z})$, where $\eta(\mathbf{z})$, but the scaling factor (the prevalence of the species over the landscape, $Pr (y = 1) = \psi$) cannot be identified from PB data only (Elith et al. 2011; Hatic and Fithian 2013; Phillips and Elith 2013; Guillera-Arroita 2014). The simple output or the ‘raw output’ of Maxent is the estimate of $f_1/f$ scaled to sum to 1 over the fitted background and representing the probability distribution over sites $x$, given the species exist at site $x; Pr (x|y = 1)$ (Phillips et al. 2006). Therefore, as currently implemented, that raw output represents $\psi(z)/(n.\psi)$, where $n$ is the total number of points in the background sample. Nonetheless, the raw output only represents relative suitability, as it is proportional to inhabitancy probabilities by a factor that is not identifiable without external data (Guillera-Arroita 2014).

With the purpose of estimating the ratio $f_1/f$, Maxent fits an exponential model so that $\log((f_1/f) = \eta(z)$, where $\eta(z)$ is a linear term of a set of features. Features, which in Maxent can belong to six classes (i.e., linear, product, quadratic, hinge, threshold, and categorical), represent a broader set of transformations of the covariates (Phillips and Dudik 2008). Moreover, features offer a great flexibility to fit complex environmental relationships, and Maxent utilizes regulation to control the trade-off between model complexity and model fit as a means to avoid overfitting (Guillera-Arroita 2014). Nevertheless, instead of working with Maxent’s "raw output," it is more common for users to work with "logistic output." Using a logistic transformation which based on a user-specified parameter ‘$\tau’,” this output scales the raw values into a scale of relative suitability ranging between 0 and 1 (Elith et al. 2011). This parameter is the prevalence for sites with ‘average’ environmental conditions under $f_1$ and by default is set to an arbitrary value ($\tau = 0.5$) (Elith et al. 2011). Therefore, the output represents the probability of the species occurs on each location/sites or can be interpreted as a probability of suitable habitat for the species based on the environmental variables included in the model.

This study was then conducted as an attempt to model the distribution of potential habitat for JHE across the Java Island, while also tried to understand the environmental characteristics of preferred habitat based on the available
environmental variables. Climatic, topographic, Land use/Land cover, and NDVI are variables used as predictors in determining the distribution of potential habitat suitability.

MATERIALS AND METHODS

Study sites
This study was carried out in Java Island which has been known as one of the 25 identified biodiversity hotspots by Myers et al. (2000), overlapped with the four closest biologically richest hotspots such as Indo-Burma, Peninsular Malaysia, Wallacea, and The Philippines. The field survey was conducted from February until December 2016 at nineteen locations across the Island, that we considered being representative of the niche breadth of the JHE based on reports from Sözer and Nijman (1924), Setiadi et al. (2000), and Prawiradilaga (2006). Field data collection was conducted to collect coordinate points of every JHE's nesting sites that were observed.

Java Island has approximately 133,930 km² of land area, with its altitude ranges from 0 m asl to about 3,676 m asl (meter above sea level). All of the locality points were collected both in the lowland and highland areas of Java. This island, like other areas in the equatorial zone, has only two seasons: wet season (during October-April) and Dry Season (during May-September). Java Island has a wide range of precipitation divided into different categories of the area. The western region of Java (Banten and West Java Provinces) and Central region of Java (Central Java and Special Region of Yogyakarta) have the same average rainfall at about 2,000 mm per year, but in some mountainous areas in western Java the number could reach up to 3,000-5,000 mm per year. Eastern area of Java, compared to the western region, has a less rate of precipitation at about 1900 mm per year (Qian et al. 2010). Likewise, the average temperature in Java can be ranged differently according to its altitude feature. Coastal areas have average temperature between 22°C and 32°C, while in higher areas with an altitude of between 400 and 1350 m asl, the average temperature ranges between 18°C and 29°C. Higher altitude generally means a lower range of temperature, in this case, the lowest temperature in Java Island can reach minus 4°C which was recorded in Ranu Pani area (slopes of Mount Semeru) (Hariyati et al. 2013).

Procedures
Field data collection
Occurrence points of JHE were compiled from two primary sources, i.e., field survey and historical database. Field survey, aiming to collect the locality data of JHE's nest coordinates, was conducted by visiting sites mentioned above accompanied by local ornithologist or student who acknowledged the exact location of JHE's nest. The observation was conducted at a safe distance of about 50-70 meters from the nest-tree of the bird, and coordinates were taken using GPS handheld receiver (Garmin© eTrex 30). To determine the exact coordinate of the nest, we recorded the direction from observation point to the nest tree using a compass (Suunto© A-10 NH) and the distance between the nest tree and the observation point we took. These data were then used to adjust the position recorded at the observation point to the nest point using QuantumGIS ver 2.6 software. A total of 31 occurrence points were collected during this field survey.

The second source of occurrence points were collected from the Global Biodiversity Information Facility (GBIF 2017) database, which provides freely accessible occurrence points in its website (http://www.gbif.org) and birdlife database (http://www.birdlife.org/datazone/speciesfactsheet). From these databases, we collected occurrence points that were recorded no older than 2010 and recorded by human observation (not specimen records). All of the occurrence points were then carefully verified, and errors that may occur were corrected using Google Earth software (Google Earth Pro 2017). Indeed, strong geographic sampling biases may often present in such database which were derived.

Figure 1. Study site and survey location of N. bartelsi across Java Island: Green points are surveyed locations. (Base map: Google Physical Maps, 2014)
from opportunistic observation and collection of records (Stolar and Nielsen 2014). Therefore, sampling bias correction is highly important and strongly advised to be conducted to minimize its strong influence on modeling prediction ability and later interpretation (Fourcade et al. 2013; Kramer-schadt et al. 2013; Fourcade et al. 2014). In their study, Fourcade et al. (2014) explained four types of biases that might be contained in such datasets, and then proposed five option methods of sampling bias correction which were carefully designed to overcome or minimize the effect of those biases. Subsequently, after identifying the type of sampling data biases contained in our occurrence data, we conducted two out of five sampling bias correction methods, i.e., (i) spatial filtering, performed by creating a grid of 3 km x 3 km cell size and randomly select only one point of occurrence per grid cell. Nonetheless, it should be noted that the size of this grid is not the representation of approximate species' dispersal capabilities, but rather as a result of modifying the 10-km radius rule of spatial filtering proposed by Kramer-Schadt et al. (2013) and Boria et al. (2014). The grid creation and point selections were conducted using QuantumGIS software ver. 2.18.14 (QGIS Development Team 2017). ii) Bias file creation, bias file is a probability surface represented by cell value which reflects the intensity of sampling effort across the area of study and gives a gradual weight to random background data used for modeling (Fourcade et al. 2014). Bias file can be artificially estimated using the aggregation of occurrences from closely related species (Phillips et al. 2009). However, in the real situation, this information is limited. Therefore, by following Elith et al. (2010), we produced a Gaussian kernel density map of the occurrence locations then rescaled it from 1 to 20 to be derived as bias file, instead of using our knowledge to create artificial bias file (Fourcade et al. 2014). This bias file was then included into Maxent modeling process through setting options (Dudik et al. 2005; Elith et al. 2010; Phillips et al. 2017). The remaining 17 points were then compiled with the occurrence points from the field survey to be used as input data.

Table 1. Climate and environmental variables used to build the models

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alt</td>
<td>Altitude</td>
<td>m asl</td>
</tr>
<tr>
<td>bio_1</td>
<td>Annual Mean Temperature</td>
<td>°C×10</td>
</tr>
<tr>
<td>bio_2</td>
<td>Mean Diurnal Range</td>
<td>°C×10</td>
</tr>
<tr>
<td>bio_3</td>
<td>Isothermality</td>
<td>×100</td>
</tr>
<tr>
<td>bio_4</td>
<td>Temperature Seasonality</td>
<td>×100</td>
</tr>
<tr>
<td>bio_12</td>
<td>Annual Precipitation</td>
<td>mm</td>
</tr>
<tr>
<td>bio_13</td>
<td>Precipitation of Wettest Month</td>
<td>mm</td>
</tr>
<tr>
<td>bio_15</td>
<td>Precipitation Seasonality</td>
<td>mm</td>
</tr>
<tr>
<td>bio_18</td>
<td>Precipitation of Warmest Quarter</td>
<td>mm</td>
</tr>
<tr>
<td>bio_19</td>
<td>Precipitation of Coldest Quarter</td>
<td>mm</td>
</tr>
<tr>
<td>LandCover_Java</td>
<td>Land Cover in Java Island 2016</td>
<td></td>
</tr>
<tr>
<td>broadleafjava</td>
<td>Broadleaf forest coverage</td>
<td></td>
</tr>
<tr>
<td>Evergreenjava</td>
<td>Evergreen forest coverage</td>
<td></td>
</tr>
<tr>
<td>Deciduousjava</td>
<td>Deciduous forest coverage</td>
<td></td>
</tr>
<tr>
<td>treecoverJava</td>
<td>Tree coverage</td>
<td></td>
</tr>
<tr>
<td>ndvi_all</td>
<td>Averaged annual NDVI value</td>
<td></td>
</tr>
</tbody>
</table>

Environmental variables data collection and Maxent modelling

We selected the environmental variables in this study on the basis of earlier screenings of related variables expected to influence the existence of species (e.g., Prawiradiagla 2006; Syartinilia and Tsuyuki 2008; Fernandez and Gurrutxaga 2010; Sohl 2014; Ferrer-Sanchez and Rodriguez-estrella 2016). Environment variables datasets collected to be used in this study amounted to 26 variables which included nineteen Bioclimate layers, altitude, land cover data, NDVI, and forest coverage data in Java Island. Bioclim data were extracted from WorldClim Bioclim datasets ver. 2.0 that provides 19 climatic variables that were interpolated and modeled from observations and averaged over the period 1970 until 2000 at one km² resolution (http://www.worldclim.org). Elevation data was also acquired from the WorldClim database. The Normalized Difference Vegetation Index (NDVI) is an index of plant “greenness” or photosynthetic activity from satellite imagery instruments (http://earthobservatory.nasa.gov) and also used to analyze the density of vegetation and to separate the healthy vegetation and unhealthy or sparse vegetation (Devadas 2008; Gene et al. 2008; Szilárd et al. 2016). An average of annual NDVI data value was extracted from the SPOT-vegetation platform at 1 km² resolution (http://free.vgt.vito.be). Java Island land cover in 2016 was obtained from http://gdfs.umd.edu/data/lc/. Forest coverage data were acquired from http://gdfs.umd.edu/data/vcf/which consist of 4 types of forest coverage data, i.e., Evergreen forest coverage, Deciduous forest coverage, broadleaf forest coverage, and tree coverage data generally in Java. All of these layers were processed through several steps including resampling data, image cutting, and type file converting by using Qgis Software ver. 2.18.14 (QGIS 2017).

It has been proved that the high inter-dependency among the bioclimatic variables gives a raise to the issue of redundancy and multicollinearity (Bedia et al. 2012). Even though neglecting this multi-collinearity issue will not affect the predictive quality of the model significantly (Elith et al. 2011), it does, however, negatively affecting model interpretability, limiting any inference of the contribution of any correlated variables, and also hampering the ability of the model for extrapolation (Brauner and Shacham 1998; Van Gils et al. 2012, 2014). Consequently, we omitted the bioclimatic variables yielding correlation values above 0.95 (Spearman's rho coefficient) in the pairwise cross-correlation matrix of each dataset (intra-dataset correlations) (Bedia et al. 2013). SDM toolbox ver. 2.0 (Brown 2014) in ArcGIS ver.10.3 was used to perform the calculation and automatically removed each one of the two correlated variables. Finally, the remaining nine bioclimatic variables along with altitude, NDVI, Land cover, and Forest coverage layers were then compiled to be used as predictor variables in this study (table 1). Supplementary table on land cover class and its code representation could be found here (http://gdfs.umd.edu/data/lc/).
Maxent modelling

Modeling exercise in this study was conducted using Maxent software ver. 3.4.1 (Phillips et al. 2017). Maxent has been proved to provide better results than other modeling algorithms with the basis of presence-only data (PO) and environmental variables (Phillips and Dudik 2008; Summers et al. 2012). The popularization of Maxent is also due to its higher predictive accuracy than any other methods (Elith et al. 2006; Summers et al. 2012), with more than 1000 published studies using this software since 2005 (Merow et al. 2013; Fourcade et al. 2014). Furthermore, Maxent also offers a wide variety of setting options which will be different in each case and occasionally requires species-specific settings (Merow et al. 2013). Therefore, in this study, we tried to ensure that the setting options were adjusted to our specific study aims, hypothesis, and our intended a priori assumptions (Peterson et al. 2011; Araujo and Peterson 2012; Merow et al. 2013). The adjusted parameters were: (i) maximum iterations was set to 5,000 for each run to allow the model to have adequate time for converging, (ii) Convergence threshold was set to $1 \times 10^{-6}$, (iii) The model calculation was set to ten times (the averaged value was the one used as the result) using "cross-validate" as the replicated run type. Using "cross-validate" means to split the data ten times (10% per partition) then train the model ten times on 90% of the data, while testing it each time on the 10% partition alternately. To avoid over-fitting and to assume that the species responds directly to the predictors (vs. to correlated factors), we decided to "smooth" the model by choosing only hinge features (Elith et al. 2010). Furthermore, we doubled the "regularization multiplier" to reduce over-fitting to a lower level (Radosavljevic and Anderson 2013).

Data analysis

One of the main outputs of Maxent is a predictive map representing the distribution of potentially suitable habitat for the species. The degrees of predicted suitable habitat are linearly scaled from 0 (lowest) to 1 (highest) probability (Phillips and Dudik 2008). Additionally, Maxent also calculated the variables’ relative contribution to the model and how these variables affect the prediction. Alternate estimation of variable importance was also collected by running the jackknife test. The results of the jackknife test show which variable have the most useful information by itself and which variable appears to have the most information that is not present in other variables (Phillips et al. 2009). The predictive maps, which by default are in ASCII format, were further analyzed using QuantumGIS software ver. 2.18.14 (QGIS Development Team 2017). To allow us to quantify the geographical distribution of predicted suitable habitat, we applied binary calculation and categorizing the values into two categories (i.e., suitable and unsuitable) using the selected threshold rule in the setting option of Maxent. Selecting the threshold rule, indeed, is one of the many sources of bias that should be minimized by Maxent user (Phillips and Dudik 2008; Nenzen and Araujo 2011; Bean et al. 2012; Syfert et al. 2013). Selecting threshold rule should incorporate consideration of relative importance difference between commission error and omission error (Phillips and Dudik 2008; Nenzen and Araujo 2011; Bean et al. 2012; Syfert et al. 2013). Considering that reducing omission error is more important determinant than reducing commission error, Norris (2014) proposed “minimum training presence” or “fixed cumulative value 1” as the most appropriate rule. Liu et al. (2016) in their study also supporting those proposed threshold rule as appropriate for modeling a rare species. Therefore, we selected “minimum training presence” threshold rule to be used in this study. The predicted suitable habitat was then reclassified into three classes: low suitability (25 - 50% probability of occurrence), medium suitability, (51 - 75% probability of occurrence), and high suitability (>75% probability of occurrence), by using the "natural breaks (Jenks) classification method in Reclassify Analysis of ArcMap ver. 10.3.

Maxent will calculate an Area Under the receiver operating characteristic (ROC) Curve (AUC) to evaluate the model performance. AUC value ranges between 0 (lowest) and 1 (highest), whereby value from 0 to 0.5 represents that the model is no better than just random prediction, value below 0.7 is low, value between 0.7 and 0.9 is good, and value above 0.9 indicates high discrimination or indicates that the model is far better than random prediction (Araujo et al. 2005). Despite having been proved that AUC does not necessarily provide useful information to assess and/or to evaluate the model performance (Lobo et al. 2008; Bahn and McGill 2013; Aguirre-Gutiérrez et al. 2013), we reported it to illustrate that the predictions in this study perform better than any model with a set of random predictors. Additionally, we conducted True Skill Statistic (TSS) (also known as the Youden index) calculation as an additional measurement to evaluate the performance of the model by calculating the summary of sensitivity and specificity minus one (Youden 1950; Allouche et al. 2006). Several studies have also demonstrated the use of Kappa statistic as a post-hoc evaluation for the Maxent model (e.g., Duan et al. 2014; Ali and Hossein 2016; Bagheri et al. 2017). However, Kappa value is highly correlated to the prevalence of the locality points and the size of the study area which would generate some bias or misunderstanding (Lobo et al. 2008; Fourcade et al. 2017). Moreover, due to the fact that both AUC and Kappa are weighting commission and omission errors equally (Allouche et al. 2006; Lobo et al. 2008; Jimenez-Valverde 2012, 2014; Fourcade et al. 2017), Kappa, just like AUC, is more reliable if it is applied in PA (Presence-Absence) model. Consequently, in case of this study where presence only data were used, we assume that the use of TSS is more suitable than the Kappa statistic.

RESULTS AND DISCUSSION

Identifying important environmental variables and model evaluation

We generated the predicted distribution of potentially suitable habitat for JHE based on observed nest locations compounded with sixteen climatic and environmental variables. According to the calculation of the relative
contributions of environmental variables to the Maxent model, there were three variables considered to have the most contribution to the model. These variables are representative of all major aspects of environmental variables included in the model, i.e., altitude (alt), Annual mean temperature (bio_1), and land cover, accounting for 58.6%, 21.6%, and 6.4%, contribution to the final result of the prediction, respectively (Table 2). The cumulative contributions of these variables contributed in a total of 86.6% to the model, whereas the remaining variables, each contributed less than 5% to the model, suggesting that the suitability of habitat for JHE are strongly influenced by the altitudinal factor, average temperature, and types of land cover in the area.

Additionally, we retrieved the alternate estimation of variable importance through the utilization of the jackknife test. The results showed that for the model, the environmental factor with the highest gain when used in isolation is annual mean temperature and altitude, which means that these variables appear to have the most useful information by itself (Phillips et al. 2008). Furthermore, temperature seasonality (bio_4) variable seemed to have the most unique information that is not present in the other variables as, according to the test, omitting this variable will reduce the regularized training gain the most (Phillips et al. 2008).

A post-hoc evaluation of ecological niche modeling is commonly performed to assess the statistical significance and the predictive performance of the model, before being used or interpreted in any procedure (Peterson et al. 2011). The Area Under the receiver operating characteristic (ROC) Curve (AUC) value may have been highlighted; it can be misleading and may poorly reflect the model accuracy (Lobo et al. 2008; Peterson et al. 2008; Jimenez-Valverde et al. 2013; Fourcade et al. 2014). Nevertheless, we retrieved the AUC value of 0.893 to indicate that the model built in this study performs better than any model with a set of random predictors, and to show the discrimination ability of the model (Lobo et al. 2008; Peterson et al. 2008; Jimenez-Valverde 2012, 2014; Fourcade et al. 2017). Furthermore, an additional evaluation of the model was conducted by calculating the True Skill Statistic (TSS) value. By calculating the summary of sensitivity and specificity minus one, we retrieved a TSS value of 0.87 for the model. This value, which is above 0.70, gives the impression that the model built for this study have a good degree of agreement, good predictive capacity, and also can be interpreted as preliminary evidence for the real ecological phenomenon, based on the environmental variables being used, rather than just statistical artifacts (Allouche et al. 2006; Li and Guo 2013).

The relationship between the models and the dominant environmental variables
In this study, we presented the response curves to illustrate how each of the prevailing environmental variables affects the model prediction. We used the response curve which represents model created using only the corresponding variable to reflect the dependence of predicted suitability both on the selected variable and on dependencies induced by the correlation between the selected variable and other variables (Phillips et al. 2008). The relationship between probability of presence and altitude (alt) (Figure 2.A) depicted that the probability of presence increased gradually along with the increase of altitude, reach above the 50% of probability of presence at the altitude of 1,200 m asl and reached its peak

![Figure 2. Response curves from Maxent for the most important variables for the species distribution model of JHE. A. Response curve for altitude variable in m asl (meters above sea level; B. Response curve for Annual Mean Temperature in °C*10; C. Response curve for land cover variable with every category represented by numbers](image)
(probability of presence = >90%) at the altitude of about 1800 m asl. Thus, it can be expected that the optimum altitude for JHE's habitat is at above 1800 m asl. The annual mean temperature (bio_1) was another critical variable that affected the suitability of habitat for JHE (Figure 2.B). The curve shows that there was a negative correlation between the probability of presence and the average temperature, presenting that the probability of presence decreased with the increase of average temperature, wherein the lowest probability of presence occurred at the temperature above 24°C. Regarding the land cover in Java Island, the response curve depict that forest area (no 8 in Figure 2.C) and closed shrubland (no. 6 in Figure 2.C) as having the most probability of species presence, with the probability value of 0.90 and 0.87, respectively.

**Predicted distribution of potential suitable habitat for JHE**

Projected distribution of predicted suitable habitat for JHE under current climate and environmental condition is shown in Figure 2. Subsequent to categorizing the output into two categories (suitable vs unsuitable) using the aforementioned threshold rule (see method), the predicted suitable habitat were then reclassified into three classes: low suitability (25-50% probability of occurrence), medium suitability (51-75% probability of occurrence), and high suitability (>75% probability of occurrence) to allow us to compare the changes in every class of probability under future climate projection. According to the result, about 17.77% (23,209 km²) area of Java Island has been projected to be suitable for the JHE’s habitat. The number consisted of 9.31% (12,163 km²), 5.81% (7,585 km²), and 2.65% (3,461 km²) of low, medium, and high probability areas, respectively (Figure 2).

The predicted suitable habitat distributed mainly on the mountainous areas of Java Island. Nonetheless, the surrounding lowland areas also predicted to be suitable as JHE's habitat, despite mostly predicted to have low probabilities of presence. Altitudinally, the low probability areas were mainly distributed in lowland areas at the altitude of between 200 and 1100 m asl, whereas medium probability areas were mainly distributed at altitude of between 1100 and 1500 m asl. Furthermore, the models predicted that the high probability areas were mainly distributed in the highland region at the altitude of above 1500 m asl (Figure 3). □

**Discussion**

This study represents an attempt to model the distribution of potential habitat for JHE across the island while also trying to gather information on its preferences of climatic and environmental conditions. In our modeling exercise, altitude and annual mean temperature were predicted as two of the most important factors determining the suitability of JHE's habitat (Table 2). The model in this study depicts a positive correlation between the probability of presence and increase in altitude, wherein the high probability of presence (above 50%) was mainly predicted at the altitude of above 1200 m asl. Regarding this altitudinal factor, it is in accordance with Van Balen et al. (2001) study which found that JHE is generally encountered in undulating, hilly, or mountainous terrain. Altitudinally, JHE species will be found at the altitude of between sea level and about 2500 m asl. However, this species are more frequently encountered in secondary forest and evergreen forest at above 1200 m asl (Partasasmita et al. 2017). This elevational preference is similar to some of the other raptor species such as Eleonora's falcon (Falco eleonorae), bearded vulture (Gypaetus barbatus), and lesser kestrel (Falco naumanni) (Donazar et al. 1993; Bustamante 1997; Urios and Martinez-abrain 2005).
The variation of local temperature from that of altitude. It has been commonly known that altitude and temperature are still vague and unclear. However, we can define as mean average temperature (Lennon et al. 2000). suitability is energy availability, which can be crudely defined as sum size of medium and high probability of presence areas. Status: NP, National Park; GFP, Grand Forest Park, NR, Nature Reserve; WS, Wildlife Sanctuary; PF, Protection Forest. Past Record References: Sozer and Nijman 1995; Hapsoro et al. 1998; Afianto 1999; Van Balen et al. 1999; 2001; Setiadi et al. 2000; BridLife International 2001; Utami 2001; Suparman 2002; Hendarsah 2003; Yuda et al. 2003; Mikoyan 2004.)

Furthermore, one of the main predictors of habitat suitability is energy availability, which can be crudely defined as mean average temperature (Lennon et al. 2000). It has been commonly known that altitude and temperature are highly correlated and it can be difficult to disentangle the variation of local temperature from that of altitude. The causal relationship between habitat suitability and temperature are still vague and unclear. However, we can identify two broad categories of how temperature affects the degree of habitat suitability, i.e., indirect and direct mechanisms. In the indirect mechanism, temperature affects the suitability of habitat through complex pathways involving its effects on resources availability, the density of population, competition, and other biotic interactions. Furthermore, the temperature may control the quantity and seasonal availability of prey species and, therefore, influence the population dynamics of JHE. On the other hand, a direct mechanism of how temperature may affect the species has been proposed by Turner et al. (1988; 1996), wherein the temperature is affecting the species directly on the energy budget of its homeotherms. Areas which have colder temperature impose greater thermoregulatory loads of species and force the species to devote relatively more energy to regulate their body temperature. As a consequence, less energy is available for other activities, such as for growing and reproducing. Whereas areas which have higher temperature will affect the physiological traits of the bird such as affecting breeding time, laying eggs pattern, and hatching time (Blondel 1985; Perrins and McCleery 1989; Woodburn 1997; Buse et al.1999; Parmesan and Yohe 2003; Parmesan 2007; Visser et al. 2009; Both et al. 2014). The characteristic of actually suitable habitat for JHE which is mainly on the higher altitude, wherein tend to have a lower temperature than the lowland areas, is in line with the pattern on the response curve model that suggests the most optimum range of temperature for JHE’s suitable habitat is at between 8° and 20°C.

The model exercise suggested that forest area and closed shrubland area as having the highest probability of presence accounting the probability value of 0.90 and 0.87, respectively. Accordingly, JHE’s nests, during field observations, were generally encountered in mountain and hills areas which is still covered by the remaining natural forest of Java Island. Generally, natural forests in this island have been cleared due to various anthropogenic activities, and the remnants are now constrained to mountain areas (above 1200 m asl) with only limited areas of natural lowland forest (below 1200 m asl) remaining (Whitten et al. 1996; Prasetyo et al. 2009; Partasasmita et al. 2017). In the study conducted by Prawiradilaga (2006), it is stated that the home-range of JHE covers not only forest area but also production forest, cultivated area, and plantation areas. Nevertheless, an intensive study conducted by Kuswandono et al. (2003) and Nijman and Prawiradilaga (2003) showed that the forest areas (secondary and evergreen forests) are more frequently used by JHE as nest location than other habitat types. Furthermore, the model in this study also suggested that shrubland plays a vital role in the degree of habitat suitability as one of the preferred habitat types. Shrubland areas in Java are often found as an ecotone of between the forest areas and cultivated areas or plantation areas. Moreover, most of the prey species of JHE are small mammals such as squirrels (Sciuridae), small rodents (Muridae), tree shrews (Tupidae), and Lesser Mouse Deer (Tragulus javanicus) (Prawiradilaga 2006) which often found in shrubland areas. Accordingly, most of the JHE nest trees are located on the steep slope between the forest and shrubland areas.
Various attempts on habitat modeling of JHE have been conducted using different approaches. By using logistic and autologistic regression models, Syartinilia and Tsuyuki (2008) modeled the distribution of JHE’s habitat in Mt. Gede-Pangrango National Park (TNGP) and its surrounding area. Utilizing presence nests data, pseudo-absence data, which were generated through a random selection process, and several environmental variables, this study successfully identified the preferred habitat of JHE based on its environmental variables and modeling the habitat patches distribution surrounding the TNGP area. By validating the results with the historical data, this study also showed a significant overall accuracy of the results and the results also could be useful for conservation management activities of this species. Subsequently, Syartinillia et al. (2014) conducted a GIS-based habitat model of JHE by using an inductive approach in the entire Java Island. They demonstrated that by using autologistic regression, it is possible to extrapolate the preceding local study into the whole area of Java Island. Moreover, this study also developed an estimation of JHE population by dividing the area of predicted suitable habitat by assumed minimum and maximum home-range size. However, the predicted suitable area in this study was mostly distributed in the mountainous areas and failed to identify the suitability habitat of JHE in lowland areas, even though the species had been found in these locations. Afterward, in 2017, Nurfatimah et al. conducted a study to model the potential habitat of JHE in Central Java Province by utilizing multi-scale approach at different image resolutions, i.e., 30 m², 90 m², and 250 m². This study demonstrated the utilization of logistic regression to model JHE's habitat patches in multi-scale images. The results of this study were able to highlight the feasibility of using different image resolutions to model the distribution of predicted habitat for JHE, while also provided more options for the conservationist to choose the most suitable image scale for managing, planning, and monitoring species based on the scale of application.

Overlaying the World Database Protected Area of Java into the predicted map gives preliminary information that most of the predicted suitable areas fall into protected areas, wherein almost all of them are predicted to have medium to high suitability. Moreover, predictive maps also depict some of the predicted suitable habitats that fall outside the protected areas. The non-protected areas are mostly cultivated area and forest plantation area which are mainly owned and operated by the state-owned company. The predictive map also has shown a good match with historical records of JHE, with only three locations which were predicted to be suitable but neither historical nor current records can confirm it. Despite various measurements have been taken to minimize errors in the model caused by bias on the sampling data (see method), it is inevitable that such omission and commission errors may still present in the result of the model due to several reasons, i.e., (i) neither the dispersal rate nor the demography of meta-population of species was included in the model, since these variables are currently unavailable. A raptor species, such as *N. bartelsi*, has a wide dispersal
range, and for this reason, modeling the niche for this type of species is considerably better if absence data were included because it is known that the ENM/SDM model utilizes the actual absence data, rather than pseudo-absences data. This is intended to produce a lower level of overprediction (Vaclavik and Meentemeyer 2009). (ii) The predictors used in this study have not yet comprehensively represented all of the environmental factors affecting the existence of the species, e.g., biotic interaction factor). iii) micro-climate variations affect the existence of species in the predicted areas but were not included in the model due to limited availability of data. Therefore, it is important to note that, like most of the ENM (SDM), the "predicted" distribution of suitable habitat does not represent the "true" prediction of the distribution of species, but rather the prediction of the distribution of "suitable" habitat based on the predictors used in this study. Nonetheless, we may treat the results of this model as an appropriate representation of how the current climate condition and other variables shape the distribution of suitable habitat for JHE. Moreover, such modeling exercise provides critical information that can be utilized for research, planning, and management needs at landscape scales.

Building more ideal model requires the availability of multiple compounding factors which are expected to have either direct or indirect effect on the target species and its associated biota. Currently, however, such ideal packages of data are limited. This limitation, in the availability of more detailed ecological and physiological data, prevents the construction of more ideal models (Morin and Thriller 2009; Sinclair et al. 2010; Ellis 2011). Nonetheless, the recent development of new climate models and the refining of current climate models provide the opportunity to build more precise and ideal model. Further modeling attempt should also incorporate potential human-induced land use/land cover changes, biotic interactions between species in the regional ecosystems, more detailed ecological data, dispersal rates of species, meta-population demography, and better presence data which accurately represent the variability of ecological niche of species. In spite of all of the aforementioned limitations, this study provides the baseline of understanding the influence of macroclimate on shaping the distribution of potentially suitable habitat for JHE while finding out other possible areas which are currently unoccupied but likely can be occupied in the future. By using a different technique of species distribution modeling, such as profile technique (e.g., DOMAIN, ENFA) and Regression-based technique (e.g., GLM, GAM, and MARS), they may presents slightly different quantitative results and discrepancies in the potential distribution habitat may occur. Nonetheless, we believe that by using currently available resources of data, the overall trend and projection results would be similar.

ACKNOWLEDGEMENTS

We thank the DIKTI for partially funding this research. We also thank all parties who provided some nest site of Javan hawk-eagle coordinates data used by us. This research is part of our big research project to modeling the predicted suitable habitat for Javan hawk-eagle across the Java Island which later funded by Muhammad bin Zayed Conservation Fund. This paper publication is supported by ALG (Academic Leadership Grant) of Prof. Johan Iskandar. Therefore, authors would like to thank Rector of Padjadjaran University, Prof. Tri Hanggono has supported the publication.

REFERENCES


GBIF.org (29th February 2016) GBIF Occurrence Download https://doi.org/10.15468/dlywhpmz


Woodburn RJW. 1997. The breeding ecology of the blue tit and great tit and the possible effect of climate change. [Thesis], University of Oxford, UK.
