Maximum Entropy Modelling for Sumatran Orangutan Habitat Suitability in Ketambe Research Station

Nurul Islamidini [1], *Zumaidar* Zumaidar[2*], *Nurdin* Sulistiyono[3] and *Muhammad* Alfarabi[4]

[1] Master Student, Biology Department Universitas Syiah Kuala, 23111 Banda Aceh, Indonesa
[2] Biology Department Universitas Syiah Kuala, 23111 Banda Aceh, Indonesa
[3] Forestry Department, Universitas Sumatera Utara 20155, Medan, Indonesia
[4] Department of Biochemistry, Faculty of Medicine, Universitas Kristen Indonesia. 13630 Jakarta, Indonesia

Abstract. This research aims to examine the spatial influence on the presence of orangutans. The study models four significant environmental variables: distance from food trees, distance from rivers, elevation, and slope, using the Maximum Entropy (Maxent) software with 10 replications. The results indicate that the environmental factor with the most substantial impact on the presence of orangutans is the proximity to food trees, contributing 90.9%. On the other hand, rivers, elevation, and slope contribute minimally to the modelling. Rivers are recognized as a determinant for orangutan nest presence, given their role as a vital life source for the forest ecosystem. However, in this modelling, the distance from rivers has limited impact due to the prevalence of numerous small rivers throughout the research area. Elevation and slope play a supportive role with modest influence, attributed to orangutans' arboreal nature. These models identify areas sharing similar environmental variables with locations where orangutans are found, suggesting that conditions matching these variables are suitable for orangutans. The habitat highly suitable for Sumatran orangutans encompasses an area of 17.43 hectares (4.14%).

1 Introduction

Wildlife selects specific habitats deemed beneficial for survival and reproductive success [1]. Habitat selection is influenced by factors such as food availability, water sources, shelter, and breeding sites. Orangutans are believed to exhibit specific habitat preferences in certain regions. According to Sulistiyono [2], orangutans exhibit preferences for areas in close proximity to riverbanks, characterized by high canopy density, elevations exceeding 200 meters above sea level, gentle slopes, and considerable distance from human presence. However, there has been no subsequent research examining the impact of orangutan preferences for nest trees, feeding trees, and the spatial attributes of orangutan presence at the Ketambe Research Station (KRS).

* Corresponding author: zumaidar@usk.ac.id

Habitat suitability modelling reveals the potential presence of Sumatran orangutans. The construction of habitat suitability models can be performed using various software, including Maximum Entropy (MaxEnt). MaxEnt is a method employed in distribution modelling to identify priority conservation areas for Sumatran orangutans [3]. Considering spatial use patterns, nest tree preferences, and feeding tree tendencies observed in the KRS, Gunung Leuser National Park, it is hypothesized that orangutan preferences in habitat utilization are constrained to specific regions. This research aims to model the habitat suitability for orangutans.

2 Methods

2.1 Collection Data

The orangutan findings data consist of direct encounters (individual orangutans) and indirect encounters. The collection of orangutan presence data was conducted through free-roaming surveys, supplemented by secondary data [4-6]. The data were consolidated into a single shapefile as the Y variable.

2.2 Processing Data

The X variable data consist of environmental variables logically influencing the presence of orangutans: distance from food trees (DFFT), distance from the river (DFR), elevation, and slope. Variables such as distance from settlements and roads are not considered due to the geographical isolation of the KRS area from human activities. Environmental variable data undergo atmospheric and radiometric correction, then are converted into ASCII raster format by standardizing path, row, and pixel resolution (10m). Subsequently, the data undergo multicollinearity testing, and Pearson correlation is employed to examine the relationships between environmental variables.

2.3 Processing and Interpretation

The modelling process was validated using Receiver Operating Characteristic (ROC) graphs. Orangutan presence data were obtained from both primary and secondary sources (Figure 4.13). In executing the model in Maxent software, the 209 presence points were divided into two sets: 75% for training (157 points) and 25% for testing (52 points) (Figure 4.14). This division was repeated 10 times for replication. Environmental variables used in the model creation include continuous variables such as elevation, slope, distance from water sources, and distance from feeding trees. After configuration, Maxent generated model performance data, response curves, variable contribution analysis, and prediction maps for orangutan presence.

Model validation is conducted to determine the accuracy of orangutan habitat suitability classification. In modelling endangered species like orangutans, dividing the dataset for validation poses a challenge due to the extremely limited sample size. Therefore, this study employs the Jackknife procedure to examine the importance of individual variables for model prediction [7]. The Jackknife procedure aims to predict the importance of environmental factors based on the increase in the area under the curve (AUC) for three different scenarios (without variables, with only one variable, and with all variables). Additionally, this research uses the Receiver Operating Characteristic (ROC) method and AUC to evaluate model fitness. The model with the highest AUC value is considered the best model. AUC measures the predictive performance of the model by comparing its ability to random predictions, with values ranging from 0 to 1 [8]. The final suitability map has values ranging from 0 to 1, interpreted [9] (Table 1).

Table 1. AUC interpretation of AUC value [9]

Value	Interpretation
0-0,2	Habitat is not suitable
0,2-0,4	Habitat is less suitable
0,4-0,6	Habitat is moderately suitable
0,6-1	Habitat is highly suitable

3 Results

3.1 Preprocessing Test

Environmental variables, such as DFFT and DFR, are represented from red to green, where the shade becomes redder as the proximity increases. Meanwhile, for slope and elevation, red hues signify higher and steeper terrain (Figure 2). Multicollinearity tests (Table 2.) and Pearson correlation tests (Table 3) were conducted for these four variables. The results of the tests indicate that all four variables can be used for modelling, and the outcome states that there is no variable exhibiting symptoms of multicollinearity. This suggests that all four environmental variables can be utilized in the development of Maxent models. Furthermore, Maxent also provides output in the form of percentage contribution values, where food trees contribute the highest percentage to the presence of orangutans.

Table. 2. Multicolinearity Test Result

Variable	Colinearity Statistics	
	Tolerance	VIF
Slope	0,941	1,062
Elevation	0,963	1,039
DFFT	0,944	1,059
DFR	0,949	1,054

Table. 3. Pearson Correlation Test Result

Variable	Slope	Elevation	DFFT	DFR
Slope	1	0,025	-0,062	-0,026
Elevation	0,025	1	0,013	0,131
DFFT	-0,062	0,013	1	0,154*
DFR	-0,026	0,131	0,154*	1

3.2 Orangutan Presence and Environmental Variables

Elith et al., as presented by Putri [10], assert that test gain is a statistical parameter gauging how well distribution predictions align with presence data compared to a uniform distribution. Higher AUC and test gain values signify a more significant influence of environmental variables on model formation. Additional criteria for assessing the impact of environmental variables involve observing the decrease in test gain and AUC values when a variable is excluded from the model. Additionally, environmental variables are evaluated based on their highest contribution to test gain and AUC when used individually, comparing their contribution to the values obtained when all variables are considered. If the difference in test gain and AUC is not significant, the variable is deemed to have the most substantial contribution to model formation. In the Jackknife test graph (Figure 3). The food tree variable has the most significant impact on the modelling, demonstrated by the dark blue graph being longer than the light blue one. Slope is a less important environmental variable, indicated by the light blue graph being longer than the dark blue one.

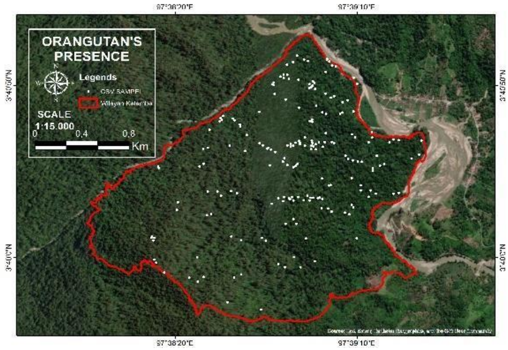


Fig. 1. Locations of orangutan presence.

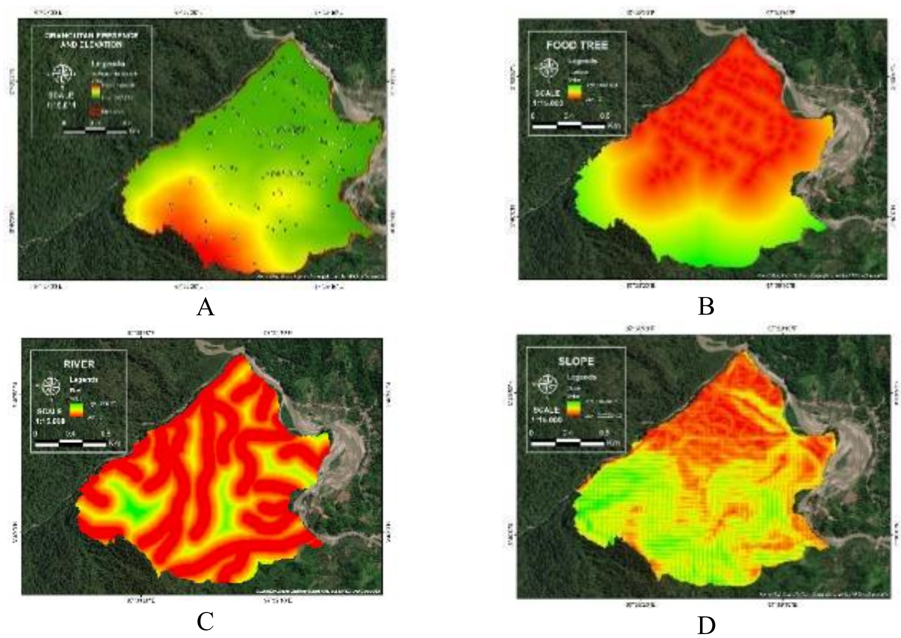


Fig. 2. Environmental factors used in the model development (A=Elevation, B=Presence of feeding trees, C=Slope, D=Distance from the river)

https://doi.org/10.1051/bioconf/202412301010

Table. 4. Environmental variable contribution percentage

Environmental Factor	Percentage Contribution	Cumulative Percentage
Food tree	90.9%	90.9%
River	4.5%	95.4%
Elevation	2.7%	98.1%
Slope	1.9%	100%

The Response Curve Prediction graph (Figure 4.) represents the presence of Sumatran orangutans concerning supporting variables (feeding trees, rivers, elevation, and slope). The river distance prediction curve shows a positive relationship within the range of 0-160 meters, indicating an increase in the probability of orangutan presence. Generally, the probability curve for Sumatran orangutan presence has a negative trend, suggesting that the farther away from the river, the lower the probability of orangutan presence.

The feeding tree prediction curve shows a positive relationship, indicating that the closer the distance to feeding trees, the higher the probability of orangutan presence. The slope prediction curve shows a negative impact, with the probability of orangutan presence increasing at slope values of 0-15% and decreasing at slope values >15%. This indicates that steeper slopes result in lower habitat suitability probability for orangutans.

The environmental variables' impact on AUC (Area Under the Curve) is depicted in Figure 4. where higher AUC values indicate a more significant influence of environmental variables on the model. The results suggest that feeding trees contribute the most, followed by rivers, elevation, and slope. The overall impact of the variables is reflected in the test gain and AUC values.

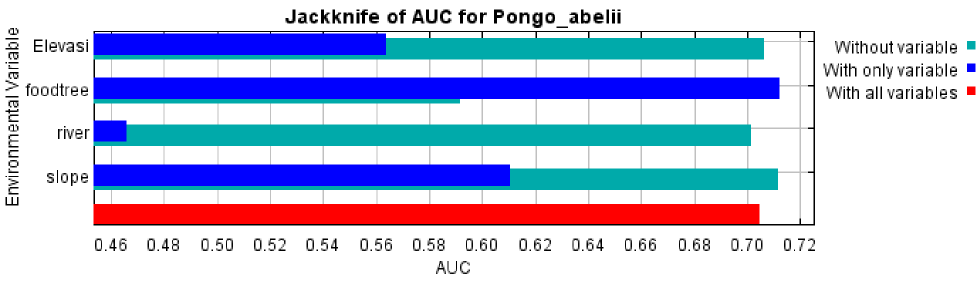


Fig. 3. Environmental variable and the impact for AUC value (The light blue color represents the contribution of the environmental variable when only that specific variable is analyzed. Dark blue color illustrates the contribution of the environmental variable when the data is analyzed without that variable. The red color indicates the contribution of all environmental variables to the modelling results)

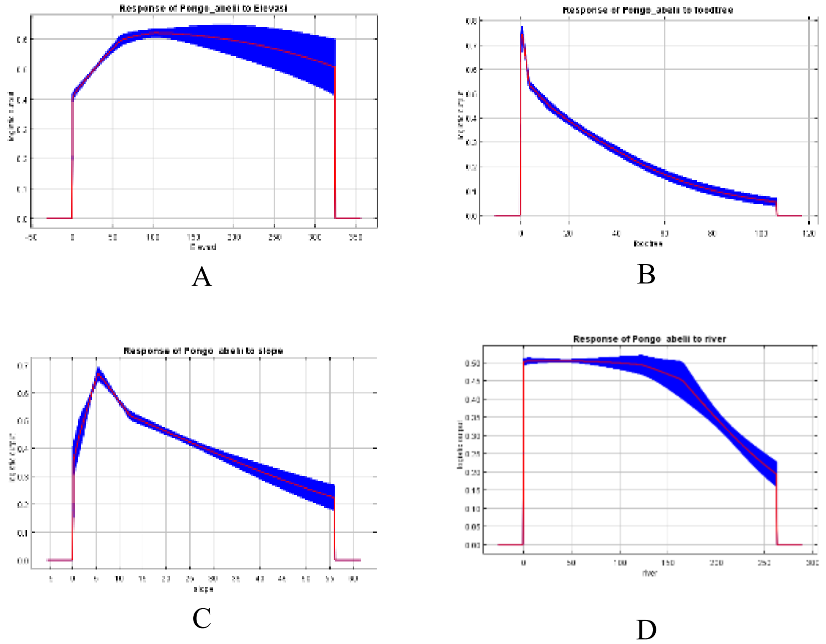


Fig. 4. Response curve at the variable A) elevation, B) DFFT, C) slope, D) DFFR

3.3 Validity Test

The AUC test value is obtained from testing 25% of randomly selected samples. The AUC value obtained is 0.705 (Figure 5) with a standard deviation of 0.023. This value indicates that the created modelling can be used and has high accuracy. The selection of the AUC method in the validation process is due to its status as a standard method for testing the validity of a model [11]. Additionally, AUC provides an advantage by avoiding subjectivity in the threshold selection process.

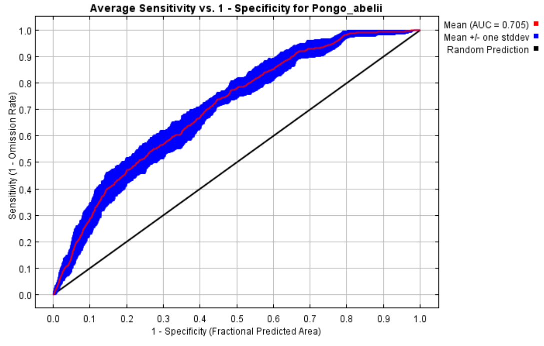
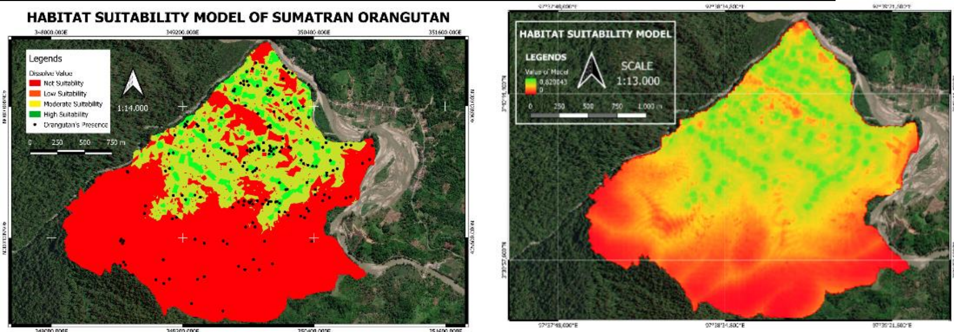


Fig. 5. AUC value for modelling

3.4 Potential Distribution

Suitability Classes	Area (Ha)	Persentage (%)	Color symbol
Not suitability	162,43	38,39%	Red
Less suitability	137,99	32,64%	Orange
Moderately suitability	104,94	24,82%	Yellow
Highly suitability	17,43	4,14%	Green
Total area	422,99	100%	



A B

Fig. 6. Orangutan suitability model, A) Vector version, B) Raster

Table. 5. Suitability Area Classes

The modelling results indicate that the areas with red color (low suitability) cover an extent of 280 hectares (66.38%), yellow areas (moderate suitability) cover 118.58 hectares (28.02%), and green areas (high suitability) cover 23.62 hectares (5.59%) (Figure 6.). The findings reveal that the least extensive category of habitat suitability is the high suitability category when compared to the areas classified in other suitability classes. Areas considered highly suitable are those with environmental conditions similar to where orangutans are found.

3.5 Discussion

The quality and quantity of the habitat will determine the composition, distribution, and productivity of a wildlife species. High-quality habitats are expected to support wildlife with good quality of life. Conversely, lower-quality habitats are predicted to result in lower wildlife quality, making them susceptible to infections and having low reproductive capabilities [12]. The interaction of biotic and abiotic factors shapes a habitat [13]. In this study, environmental factors include the presence of feeding trees, rivers, elevation, and slope (Fig. 2.). Each environmental factor will undergo classification and overlay with the presence of orangutans.

In this study, there is novelty in the introduction of the presence of feeding tree variables. Previous research has not incorporated this variable in orangutan habitat modelling. Modelling is a simplification or abstraction of real-world conditions. Species distribution models (SDMs) do not predict the exact presence of wildlife, such as orangutans in this study. Instead, these models identify areas with similar environmental variables to the points where the animals are found and infer that conditions matching these are suitable for orangutans. Despite their capabilities, SDMs do not precisely identify the specific locations of orangutans. Nevertheless, it's noteworthy that the KRS area is perceived as relatively small when compared to the expansive home ranges of orangutans. Adult female Sumatran orangutans typically establish home ranges covering approximately 850 hectares, while subadult and adult males utilize ranges of at least 2500 hectares, and potentially even larger areas [14]. Therefore, further research is expected, involving modelling orangutans across the entire island of Sumatra, including *Pongo tapanuliensis*.

The habitat suitability modelling for Sumatran orangutans was conducted using Jackknife operations [15]. Jackknife involves sequentially excluding one environmental variable from the model and running the model with the remaining variables. The modelling results indicate that feeding trees (food trees) contribute significantly, with a high contribution of 90.9% (Table 4). Field observations suggest that orangutans tend to visit trees with food and quickly move to other trees in search of food. This aligns with Rijksen [16], stating that orangutans tend to approach feeding trees to fulfill their needs.

Research by Rahman (2010b) also highlights the importance of food as a limiting factor influencing the presence of wildlife. Kartono [17] study on orangutan feeding activities in KRS showed that orangutans spend 38.34% of their time eating, while Azhar [18] in KRS reported a dominant activity of eating, accounting for 40.9% of the time. This behavior is considered dominant as adult orangutans have high energy requirements.

Other environmental variables contribute less, with rivers contributing 4.5%, elevation 2.7%, and slope 1.9%. The river factor plays a role in determining the presence of orangutan nests. Rivers are considered crucial for the forest ecosystem, influencing orangutan habitat suitability. The proximity to a river increases vegetation diversity, including species of food trees for orangutans. Distance from the river (DFFR) is identified as one of the factors influencing the presence of orangutans. Rivers serve as a crucial life source for the forest ecosystem. The proximity to the river is hypothesized to be associated with increased species diversity, including the tree species that serve as food for Sumatran orangutans. Consequently, there is a potential for orangutan sightings in the vicinity of rivers. Therefore, it is essential to consider the distance from the river as a factor influencing the presence of Sumatran orangutan nests. Field observations and communication with research assistants have revealed instances where orangutans come down to the forest floor to drink water using their hands. This aligns with Sulistiyono, Maulana [2] statement that the most suitable habitat for Sumatran orangutans is areas near rivers. Meijaard, Rijksen [19] also suggest that the optimal habitat for orangutans is along riverbanks, as rivers are essential life sources. This is because rivers provide a continuous flow of water, serving as a vital water source for the surrounding flora and fauna, as stated by Rahman (2010a).

Orangutans tend to prefer lower elevations. Based on field observations and direct communication with research assistants, orangutans tend to avoid areas with high elevations. This is because, at higher elevations, the sound of blowing wind becomes more thunderous, causing the orangutans to feel threatened. Additionally, at higher elevations, vegetation is sparser, restricting the movement of orangutans and making them more vulnerable to predators. This aligns with the statement by Prayogo, Thohari [20], revealing that the majority of orangutan nests are found at elevations between 0-500 meters above sea level. This statement is also supported by research from Van Schaik, Priatna [21], indicating that orangutans tend to prefer living in lowlands, with the highest density observed between elevations of 200-400 meters above sea level.

Orangutans are more commonly found in the slope class 1 (0-8%), which represents flat terrain. Despite covering only 16% of the total area, this flat slope class accounts for 74 sightings of orangutans (35%). Orangutans are arboreal animals, meaning that the slope does not significantly impact their presence. However, the slope does play a role in the type and quantity of vegetation. Observations indicate that as the slope increases, the growth of vegetation becomes more limited. Consequently, it becomes more challenging to find orangutans in such areas. This observation is supported by Yang, El-Kassaby [22], who found significant differences in slope aspects concerning species composition, vegetation structure, and biodiversity patterns. This is attributed to the fact that in areas with steep slopes, nutrients in the soil tend to be carried away by water flow to lower-slope areas. As a result, the nutrient content in high-slope areas is lower than in low-slope areas, leading to fewer vegetation and making orangutan sightings more difficult. According to Rijksen [16], the topographic profile, including slopes ranging from gentle to steep, influences the selection of nesting locations. Nests can be found on hillside slopes and in trees at the edge of cliffs. However, this study aligns with Prayogo, Thohari [20], revealing that orangutan nests are predominantly found in areas with gentle slopes. In contrast, nests are scarce in hilly regions.

The Response Curve Prediction graph represents the presence of Sumatran orangutans concerning supporting variables (feeding trees, rivers, elevation, and slope). The river distance prediction curve shows a positive relationship within the range of 0-160 meters, indicating an increase in the probability of orangutan presence. Generally, the probability curve for Sumatran orangutan presence has a negative trend, suggesting that the farther away from the river, the lower the probability of orangutan presence. The feeding tree prediction curve shows a positive relationship, indicating that the closer the distance to feeding trees, the higher the probability of orangutan presence. The slope prediction curve shows a negative impact, with the probability of orangutan presence increasing at slope values of 0-15% and decreasing at slope values >15%. This indicates that steeper slopes result in lower habitat suitability probability for orangutans.

4 Conclusion

The environmental factor with the highest percentage contribution to the presence of orangutans is the food trees (90.9%). River, elevation, and slope contribute minimally to the modelling. Rivers is identified as a determinant of orangutan nest presence, given that rivers serve as a life source for the forest ecosystem. In this modelling, the distance from rivers has limited impact due to the prevalence of numerous small rivers across the entire research area. Elevation and slope serve as supporting factors with modest influence, due to orangutans being arboreal creatures. These models identify areas with similar environmental variables to the points where the animals are found and infer that conditions matching these are suitable for orangutans. The habitat highly suitable for Sumatran orangutans covers an area of 17.43 hectares (4.14%). The limitation of this study is that the KRS area is perceived as relatively small in comparison. The researcher suggests conducting the next study in a more extensive area.

5. Acknowledgement

Special thanks to the Orangutan Information Centre for funding this research. We would like to express our gratitude to the field staff at Ketambe RS, including Mr. Muhammad Firhan Al Azhar, Mr. Arwin, Mr. Ibnu Hajar Al Haitami, Mr. Abdul Khalik Hasibuan, Mr. Isma Kusara, Mr. Sukurudin and Mr. Makmur Jaya.

References

1. A. R. E. Sinclair, J.M. Fryxell. Wildlife Ecology, Conservation, and Management Second Edition (Blackwell Publishing Ltd; 2006)
2. N. Sulistiyono, M. Maulana, I. Pramuja, P. Patana, A. Purwoko, editors. Spatial model of human-sumatran orangutan (Pongo abelii) conflict vulnerability in Besitang, Indonesia. IOP Conference Series: Earth and Environmental Science (2021).
3. C. Merow, M. J. Smith, J. A. Silander Jr. A practical guide to MaxEnt for modelling species' distributions: what it does, and why inputs and settings matter. Ecography. 36(10), 1058-69 (2013).
4. N. Islamidini. Inventarisasi Jenis Pohon Sarang Orangutan (Pongo abelii Lesson, 1827) dengan Metode Jelajah Bebas di Stasiun Penelitian Ketambe, Taman Nasional Gunung Leuser (KKP). Thesis, Universitas Syiah Kuala (2022)
5. N. Islamidini. Kajian Preferensi Pohon Sarang Orangutan (Pongo abelii Lesson, 1827) di Stasiun Penelitian Ketambe, Taman Nasional Gunung Leuser sebagai Upaya Konservasi. Tugas Akhir, Universitas Syiah Kuala Fakultas Matematika dan Ilmu Pengetahuan Alam (2023)
6. N. Nurlaili. Estimasi Populasi Orangutan Sumatera (Pongo abelii Lesson, 1827) berdasarkan Jumlah Sarang di Stasiun Penelitian Ketambe, Taman Nasional Gunung Leuser. Thesis, Universitas Syiah Kuala Fakultas Matematika dan Ilmu Pengetahuan Alam (2019)
7. R. G. Pearson, C. J. Raxworthy, M. Nakamura, A. Townsend Peterson. Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of biogeography. 34(1), 102-17 (2007).
8. S. J. Phillips, R. P. Anderson, R. E. Schapire. Maximum entropy modelling of species geographic distributions. Ecological modelling. 190(3-4), 231-59 (2006).
9. M. Ansari, A. Ghoddousi. Water availability limits brown bear distribution at the southern edge of its global range. Ursus. 29(1), 13-24 (2018).
10. A. Putri, M. D. Kusrini, L. B. Prasetyo. Pemodelan Kesesuaian Habitat Katak Serasah (Leptobrachium Hasseltii Tschudi 1838) dengan Sistem Informasi Geografis di Pulau Jawa. Jurnal Pengelolaan Sumberdaya Alam dan Lingkungan (Journal of Natural Resources and Environmental Management). 10(1), 12-24 (2020)
11. J. M. Lobo, A. Jiménez-Valverde, R. Real. AUC: a misleading measure of the performance of predictive distribution models. Global ecology and Biogeography. 17(2), 145-51 (2008).
12. H. Alikodra. Pengelolaan Satwaliar, Jilid I dan II. Departemen Pendidikan dan Kebudayaan Direktorat Jenderal Pendidikan Tinggi Pusat Antar Universitas Ilmu Hayat IPB, Bogor (1990)
13. F. Kartono. Potensi dan Preferensi Pakan Orangutan Sumatra di Taman Nasional Gunung Leuser. Jurnal Pengelolaan Sumberdaya Alam dan Lingkungan. 10(4), 627-38 (2021)

14. H. Prayogo, T. Thohari, A. Machmud, S. Solihin, D. Duryadi, P. Prasetyo, et al. Permodelan Kesesuaian Habitat Orangutan Kalimantan (*Pongo pygmaeus pygmaeus*) di Koridor Satwa Kapuas Hulu Kalimantan Barat. Jurnal Penelitian Hutan dan Konservasi Alam. 13(2), 137-50 (2016)
15. J. Yang, Y.A. El-Kassaby, W. Guan. The effect of slope aspect on vegetation attributes in a mountainous dry valley, Southwest China. Scientific reports. 10(1), 16465 (2020)
16. K. Urban. "Pongo abelii" (On-line), Animal Diversity Web 2008 [Available from: https://animaldiversity.org/accounts/Pongo_abelii/]
17. H.E. Negga. Predictive modelling of amphibian distribution using ecological survey data: a case study of Central Portugal. Thesis, ITC (2007)
18. H.D. Rijksen. A field study on Sumatran orang utans (*Pongo pygmaeus abelii* Lesson, 1827): ecology, behaviour and conservation. Ph.D. thesis, Wageningen University and Research (1978)
19. A. Azhar. Teknik Pendataan Perilaku Harian dan Pemetaan Jalur Jelajah Orangutan Sumatra (*Pongo abelii* Lesson, 1827). Thesis, Universitas Syiah Kuala (2018)
20. E. Meijaard, H.D. Rijksen, S. Kartikasari. Diambang kepunahan!: kondisi orangutan liar di awal abad ke-21. Penyunting S.N. Kartikasari, The Gibbon Foundation Indonesia, Jakarta (2001)