ROADKILL PROBABILITY MODELLING FOR MALAYAN TAPIR BASED ON LAND COVER CHANGES USING TIME-SERIES OF LANDSAT IMAGERY

[1]Muhammad Daniel Iman bin Hussain, [2]Masahiko Nagai and [3]Tsuyoshi Eguchi
[123]Yamaguchi University, 2-16-1, Tokiwadai, Ube-shi, Yamaguchi 755-0097, Japan,
Email: c053vew@yamaguchi-u.jp

KEY WORDS: Roadkill, Malayan Tapir, land cover, risk prediction, highway

ABSTRACT: Malayan Tapir is an endangered species, located in Peninsular Malaysia, with decreasing populations each year due to roadkill and habitat loss due to deforestation. This research aims to perform time-series land cover change analysis between the year 2000 and 2020 using Landsat imagery, determine the relationship between the land cover change and the number of roadkill occurrences and analyze which factors contribute to the distribution of roadkill hotspots and develop a roadkill risk map for the year 2025 based on future land cover change. Terengganu, Malaysia was selected as the study area. Landsat 7 and 8 imagery, roadkill data, road and water maps and digital elevation models were used. Landsat imagery was preprocessed and classified with supervised classification into (1) urban, (2) forest, (3) water, (4) bare land and (5) plantation. The accuracy of 2000, 2010 and 2020 land cover maps were assessed, and change detection analysis was performed. 2020 land cover was predicted and validated with the actual 2020 land cover map. After model validation, predicted 2025 land cover map was generated. Using 2000-2020 land cover maps, changes that affect roadkill were analyzed. From 2000 to 2020, forest areas decreased by 28.53% while plantation areas increased by 17.29%. Between 2010 and 2020, there was a drastic decrease in forest and increase in plantation. Forest to plantation change was found to be the major land cover change in the study area. After model validation, predicted 2025 land cover map was generated. Out of 40 roadkill locations, 27 were found to be influenced by forest to plantation change, indicating it as the major roadkill factor. Thus, the areas that are predicted to change from forest to plantation in 2025 surrounding roads are classified as roadkill risk areas.

1. INTRODUCTION

Tapirus Indicus, or more commonly known as the Malayan Tapir is the largest of four species of tapir, located primarily in Southeast Asia. Weighing in about 350 kg and growing 1.8 meters long, they are easily recognizable by their trademark white patch around their torso, and black heads and hind quarters. They play a major role in maintaining the biodiversity of tropical ecosystems, dispersing seeds through their scat. Tapir population is widely distributed in Peninsular Malaysia's forest especially in swamps and lowland forests, similar to the distribution of other big mammals such as elephants and tigers. Malayan Tapir is considered endangered throughout its range (Traeholt, 2016). The population of tapir in Peninsular Malaysia has been estimated between 1100 to 1500 by the Department of Wildlife and National Park (DWNP) (DWNP, 2012). Habitat loss and fragmentation are the primary threats to the survival of the Malayan Tapir. Due to rapid infrastructure development, mainly road expansion, wildlife movements have been severely impacted. This can be seen from countless wildlife-vehicle collisions (WVC), also known as roadkill, sightings, and reports throughout the country. A total 115 Malayan tapir roadkill were recorded from 2006 to 2019 in Peninsular Malaysia. The number of roadkill will most likely increase with rapid road development and land cover changes along the road (Alamgir, 2018). This shows that analyzing future land cover changes is vital to prevent the number of roadkill from increasing. This study aims to model roadkill probability using free time-series Landsat satellite imagery based on future land cover changes.

2. RESEARCH OBJECTIVES

This study has three technical objectives: To perform time-series land cover change analysis between the year 2000 and 2020 using Landsat satellite imagery, to determine the relationship between the land cover change and the number of roadkill occurrences and analyze which factors contribute to the distribution of roadkill hotspots and to develop a roadkill risk map for the year 2025 based on future land cover changes.

3. LITERATURE REVIEW

3.1. Malayan Tapir Habitat And Ecosystem

The Malayan Tapir is distributed throughout tropical lowlands in Southeast Asia and typically inhabit primary forests, secondary forests, mature rubber plantations, forest edges and grasslands. Although they have been found at altitudes up to 2000m, they generally live in lower slopes and valley bottoms below 400m. They are also found commonly near villages and within 5 km of major cities. They have been found to eat a wide variety of fruits and leaves of more than 115 species of aquatic leaves, buds, soft twigs from the forest floor and low-hanging shrub and play a major role in maintaining the biodiversity of tropical ecosystems, dispersing seeds through their scat. The Malayan tapir is mainly solitary, has poor eyesight but excellent hearing and sense of smell (Denver, 2018). They are primarily active at night when they are vulnerable to predators and poachers. They usually prefer to live near water, and capable of climbing steep slopes. The main threat to the survival of the Malayan tapir is habitat loss due to deforestation and fragmentation (Tapirs, 2022).

3.2. Roadkill In Peninsular Malaysia

One study assessed the spatiotemporal patterns of Malayan tapir roadkill using records from the Department of Wildlife and National Park, Peninsular Malaysia (Magintan, 2021). A total 115 Malayan tapir roadkill were recorded from 2006 to 2019 in Peninsular Malaysia (Figure 1). Figure 1 represents the number of roadkill according to month from 2006 to 2019 2019 from the Department of Wildlife and National Parks, Peninsular Malaysia. April, June and September had the highest number of roadkill.

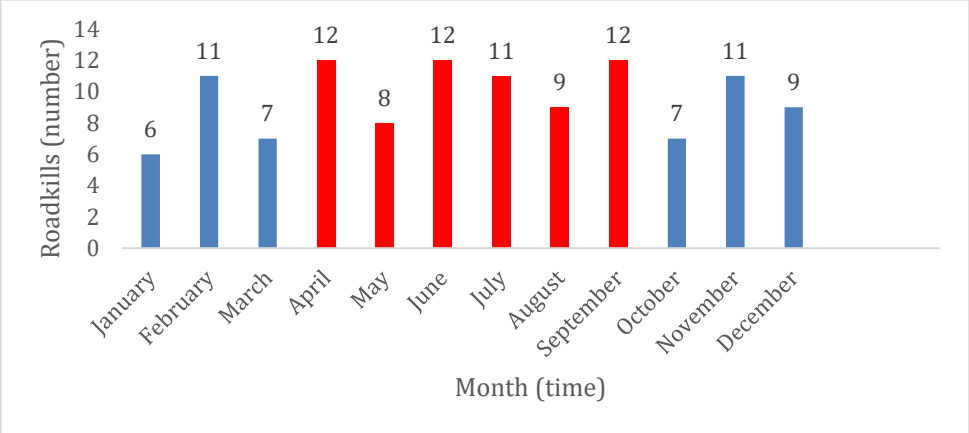


Figure 1 Roadkill in Peninsular Malaysia according to month from 2006 to 2019 (red:dry ,blue:wet)

There was no significant difference in roadkill occurrences between dry and wet seasons. Terengganu showed the highest number of roadkill occurrences. This is due to the construction of the East Coast Expressway Phase 2, with the increase of high-speed vehicles on the road. Most roadkill for Malayan tapir occurred at night when drivers tend to drive faster, and driving is more dangerous due to darkness (Magintan, 2010). (Clements, 2014) revealed that roads construction is one of the drivers of declines in mammal populations and extinctions in tropical forests (due to roadkill).

3.3. Remote Sensing Analysis For Roadkill

Countless studies have found that road width, proximity to water and proximity to forest are the common factors that affect roadkill. Generally, the wider the road and the higher the proximity to water and vegetation, the higher the risk of roadkill (Pagany, 2020). Geographic Information System (GIS) has been useful in analyzing the spatial and temporal patterns of roadkill. Spatial analysis is used primarily to assess roadkill patterns. The Kernel Density Estimation (KDE)

method has been utilized in several studies to detect roadkill hotspots. (Bartonicka, 2018) used the KDE+ method to determine the presence or absence of roadkill in a cluster. The measurement included explanatory variables such as biological, traffic and environmental and found that environmental variables such as distance to forest, distance to stream and road width. Remote sensing data is commonly used to create land use of a particular year, typically using Landsat imagery, that is used often to assess the surrounding landscape and vegetation that may affect the number of roadkill. However, land use land cover change is rarely discussed as an explanatory variable for roadkill occurrences. This creates a need for areas with high cases of land use and cover changes due to deforestation and rapid urban development that increase the risk of roadkill.

4. DATA COLLECTION

Optical satellite surface reflectance imagery from Landsat 7 Enhanced Thematic Mapper (ETM+) sensor from the year 2000 until 2012 and Landsat 8 Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) sensors from the year 2013 until 2020 were obtained from the United States Geological Survey (USGS) website. All images were geometrically corrected and acquired in level 1T (L1T). Only scenes with less than 30% of cloud cover were selected. The roadkill data of Malayan Tapir in Terengganu from 2008 to 2020 was provided by the Department of Wildlife and National Parks, Peninsular Malaysia. The provided data is in Excel format, and includes the time, date, location, and coordinates of the roadkill occurrences. Road and water map were acquired from BBBike.org using the extract service. BBBike extracts enables users to extracts areas from Planet.osm in OSM, PBF, o5m, Garmin, Navit, maps.me, mbtiles, OsmAnd, Esri shapefile, mapsforge, OPL, GeoJSON, SQLite, text or CSV format. Global Multi-resolution Terrain Elevation Data (GMTED2010) of the study area was acquired from the United States Geological Survey (USGS) website. The GMTED2010 elevation products have been produced using the following aggregation methods: minimum elevation, maximum elevation, mean elevation, median elevation, standard deviation of elevation, systematic subsample, and breakline emphasis. The resolution is 30-arc-second (1 kilometer).

Google Earth Engine (GEE) was used initially for satellite imagery acquisition, preprocessing and image classification. GEE is a free cloud-based platform for planetary-scale geospatial analysis which have supercomputing ability for complex calculation or massive processing problems. Quantum Geographic Information System (QGIS) 3.18 was used for roadkill analysis, particularly for kernel density estimation, Euclidean distance calculation for road and water layer, creating slope maps derived from DEM data, converting vector layers to raster and ASCII formats, and further land cover analysis. QGIS is a free and open-source cross-platform desktop geographic information system (GIS) application that supports viewing, editing, and analysis of geospatial data. The Land Change Modeler tool inside Terrset was used for land cover change prediction and validation. TerrSet (formerly IDRISI) is an integrated geographic information system (GIS) and remote sensing software developed by Clark Labs at Clark University for the analysis and display of digital geospatial information. The LCM (Land Change Modeler) embedded in the TerrSet Geospatial Monitoring and Modeling System (TGMMS) software was used for prediction of future LULC for 2025 based on the classified satellite images from 2000 to 2020. The LCM determines how the factors influence future LULC change, how much land cover change took place between earlier and later LULC, and then calculates a relative amount of transitions. The future land use scenarios were based on recent trends, historical land use information, and anticipated future changes.

5. METHODOLOGY

5.1. Remote Sensing-Based Analysis

Terengganu, Malaysia was chosen as the study area. A polygon of the study area (ROI=region of interest) was created using the geometry tool in GEE. After loading a raw Landsat ImageCollection for a single year for each year, the image collection was reduced by calculating the median. The median composite was then clipped to the ROI. Clipping is done so that only the data inside the ROI is processed. Single Landsat scenes do not cover the whole study area, thus multiple scenes of the same year are combined to create an annual composite. Clouds were masked using the Quality Assessment band. To fill the gaps in Landsat 7 images from 2003 to 2012, focal mean or focal median function were applied to the images in the collection. These apply a morphological filter to each band of an image by inputting the pixels in a custom kernel and then applying a blend. A radius of 1.5, a square kernel type, pixel kernel, and 8 iterations were defined.

Supervised classification was performed using the Classification and Regression Trees (CART) classifier within the Google Earth Engine platform. 2000 Landsat 7 and 2020 Landsat 8 imagery were used, due to low cloud cover, to create the training dataset for Landsat 7 and Landsat 8 imagery respectively. A total of 200 sample points (70%) were selected for the training dataset and 85 points (30%) for the validation dataset. The training datasets were exported as a GEE asset to be used for training Landsat imagery for the remaining years. The composite images were classified into five land cover types: (1) Urban, (2) Forest, (3) Water, (4) Bare land, (5) Plantation. Values less than 1 were classified as no data (0).

Accuracy assessment was performed for the classified maps using Google Earth images of the respective years for reference. To assess the accuracy of the classified images, confusion matrices, such as overall accuracy (OA), producer accuracy (PA) and user accuracy (UA) and Kappa coefficient were calculated in GEE.

In this study, the Land Change Modeler (LCM) was used to predict land cover changes by following four steps: (1) change analysis, (2) transition potential and determination of land cover change drivers, (3) change prediction and (4) model validation. Change detection analysis was computed using LCM and performed by using the classified maps (2000, 2010, and 2020) and the predicted LULC (2025) to demonstrate the change patterns. The LULC dynamics in each study period were assessed using the numerical values extracted from the classified images. To determine the change pattern, the images classified from consecutive periods were cross-tabulated and compared to each other. The probability matrix was created between 2000 and 2010, 2010 and 2020, 2020, and 2025 using LCM. In QGIS 3.18, using the Semi-Automatic Classification Plugin (SCP), classification reports each classified image were created to identify the number of pixels, percentage, and area for each land cover class. Land cover change drivers such as slope, distance from roads and distance from water maps were calculated using Euclidean distance algorithm in QGIS 3.18.

Change prediction was performed using the CA-Markov Chain model, which is a stochastic modelling process used to simulate future land cover changes based on the transition probability maps. Validation was performed by calculating Kappa statistics to measure the goodness of fit or the best value between the predicted and the reference maps. Kappa values were categorized as poor below 0.40, fair to good from 0.40 to 0.75, and excellent over 0.75 [22]. Also, an error matrix was produced using ERRMAT in Terrset software. Errors of omission and errors of commission were calculated. After the model was validated for the predicted 2020 land cover map, the simulation process was repeated to predict the 2025 land cover using 2010 and 2020 classified maps.

5.2. Roadkill-Based Analysis

Figure 4.11 shows the roadkill location for Malayan Tapir in the study area and the surrounding changes that influence the number of roadkill occurrences. Roadkill data, that comprises of 40 locations of roadkill occurrences in the study area from 2008 to 2020, and classified Landsat images from 2000 to 2020 were imported into QGIS to identify the factors and transitions that influence roadkill. Land cover changes at each location were analyzed by looking at the land cover of several years before and after the roadkill occurred.

Comparing the actual 2020 land cover map and the predicted 2025 land cover map, areas predicted to change from forest to plantation surrounding road areas in 2025 are recognized as roadkill risk areas. For roadkill hotspot prediction, spatial analysis was performed using 3 raster layers in QGIS 3.18: (1) land cover change map (2) road map (3) roadkill location map. A buffer of 500 meters was created on each side of the road to detect the surrounding changes from forest to plantation. Circular buffers with radius of 1.5km were created for each roadkill location. The roadkill risk was divided into 3 categories: low risk, medium risk, and high risk. Medium risk areas have surrounding areas that change from forest to plantation while high risk areas have that and more than 3 past roadkill occurring in the roadkill radius.

6. RESULTS AND DISCUSSION

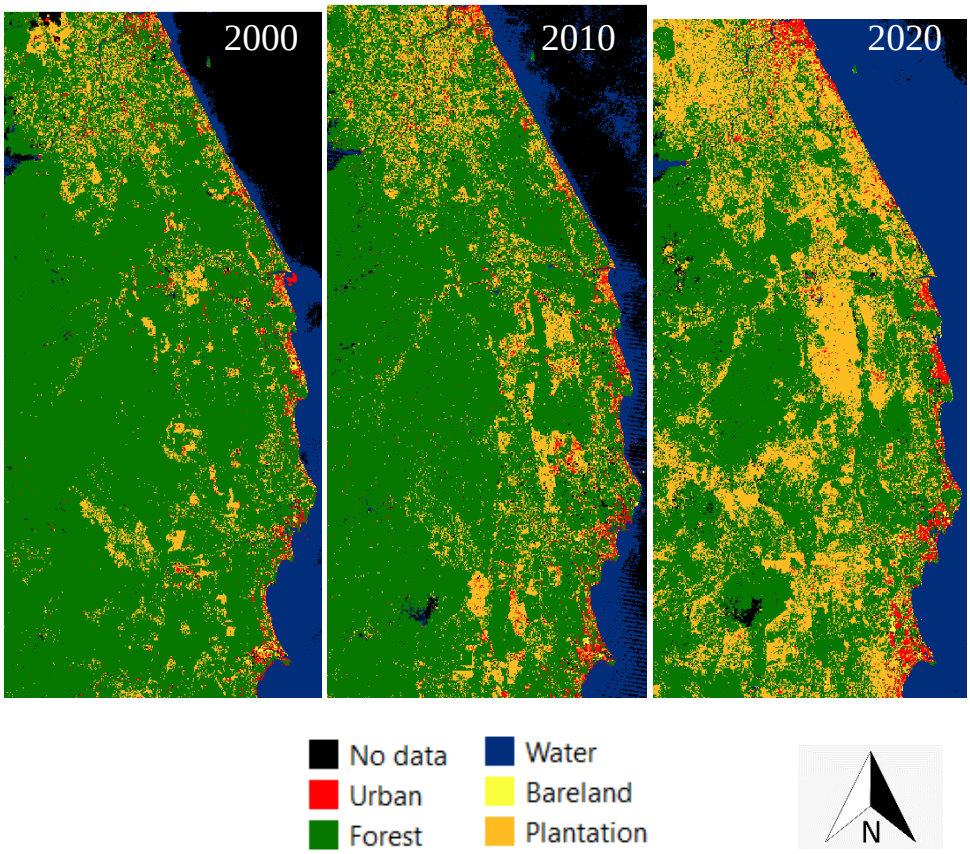


Figure 2 The classified land cover maps of the study area for 2000, 2010 and 2020.

Figure 2 shows the classified land cover maps of the study area for 2000, 2010 and 2020. Black represents no data, red represents urban, green represents forest, blue represents water, yellow represents bareland and orange represents plantation areas. All three classified maps showed high accuracy and were accepted for land cover change prediction.

Table 1 Area statistics of the land cover classes for 2000, 2010 and 2020.

	2000		2010		2020	
Class	%	km^2	%	km^2	%	km^2
Urban	2.09	220.61	2.70	284.56	2.67	313.53
Forest	78.93	8347.50	73.16	7709.50	50.40	5909.40
Water	8.66	915.98	8.95	942.69	18.60	2178.60
Bareland	0.21	21.89	0.27	28.74	0.96	112.78
Plantation	10.11	1069.40	14.92	1572.20	27.40	3217.60

Table 1 shows the area statistics of the land cover classes for 2000, 2010 and 2020. The forest is the dominant land cover which covered 8347.50 km2 (78.93%) of the study area in 2000, 7709.50 km2 (73.16%) in 2010, and 5909.40 km2 (50.40%) in 2020. During 2000-2010, the forest land decreased from 8347.50 km2 (78.93%) to 7709.50 km2 (73.16%). However, forest land area drastically declined during from 2010 to 2020 by 1800.10 km2 (23.35%). This is due to forests being cut down to create plantation areas, particularly palm. This can be seen from 2010 to 2020 where plantation areas increased by almost twice from 1572.20 km2 (14.92%) to 3217.60 km2 (27.40%).

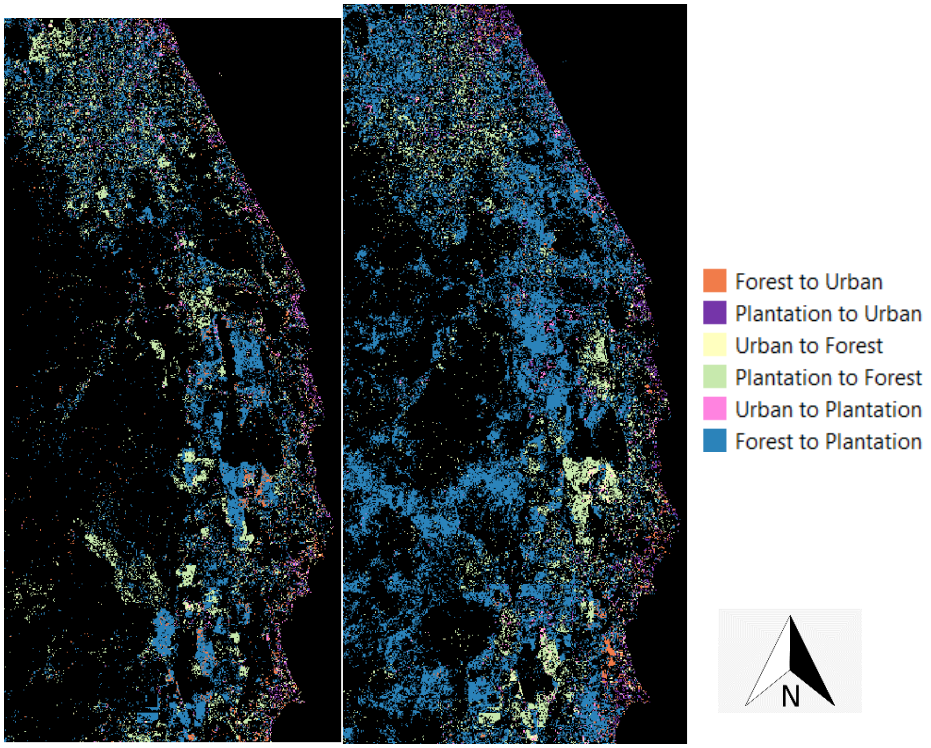


Figure 3 Overall change map between 2000-2010 (left) and 2010-2020 (right)

Figure 3 shows the overall change map between 2000-2010 and 2010-2020. Orange represents forest to urban, purple represents plantation to urban, yellow represents urban to forest, green represents plantation to forest, pink represents urban to plantation and blue represents forest to plantation. From the figure, deforestation by transition of forest to plantation was found to be the major land cover change for both periods between 2000-2010 and 2020 in the study area.

The land cover change results indicated that the decrease of forest cover was significantly influenced by the increase in plantation areas. Forest-plantation transition was considered in LCM for transition potential modelling. Change drivers such as slope, distance from roads and distance from water were added to improve the accuracy of the model. Transition potential modeling is assessing the likelihood of land cover change from one class to another depending on the suitability transition of area and the presence of change drivers. The transition probability matrices were created for the years 2020 (using 2000 and 2010 land cover map) and 2025 (using 2010 and 2020 land cover map) as shown in Table 2 and Table 3 respectively. The transition probability matrix shows the probability of a land cover changing to another class, within the specified time.

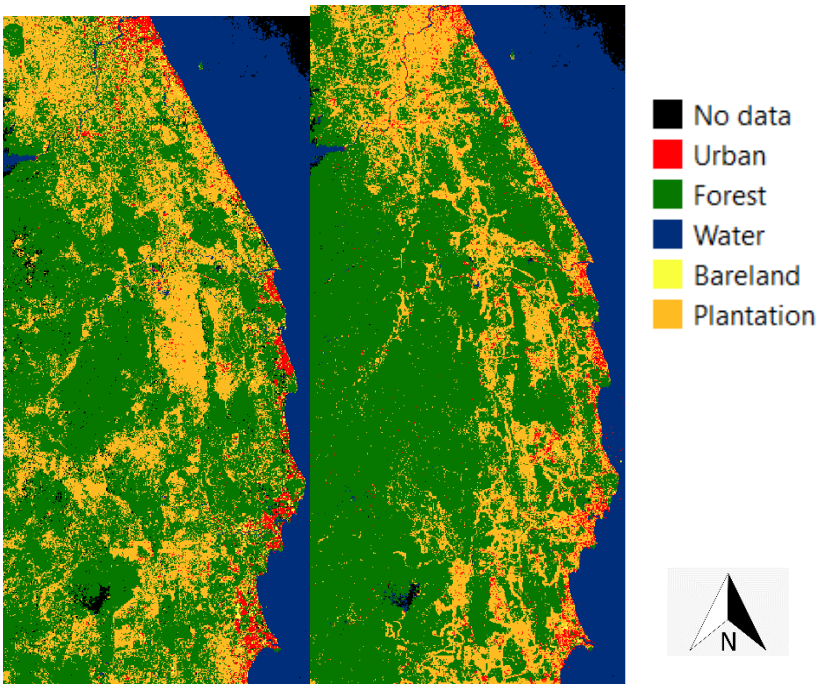
Table 2 Transition probability matrix of land use land cover classes for the year 2020

	Urban	Forest	Water	Bareland	Plantation
Urban	**0.3288**	0.265	0.0681	0.01	0.3281
Forest	0.0147	**0.8639**	0.0016	0.0019	0.1179
Water	0.0232	0.0157	**0.9577**	0.0005	0.0029
Bareland	0.3371	0.1021	0.0241	**0.224**	0.3128
Plantation	0.0567	0.4594	0.0017	0.0048	**0.4773**

Table 3 Transition probability matrix of land use land cover classes for the year 2025

	Urban	Forest	Water	Bareland	Plantation
Urban	**0.6119**	0.1198	0.0502	0.0369	0.1812
Forest	0.0034	**0.8227**	0.0013	0.0025	0.1701
Water	0.0015	0.0028	**0.9936**	0.001	0.0011
Bareland	0.2044	0.1297	0.0028	**0.5338**	0.1292
Plantation	0.043	0.1954	0	0.0164	**0.7452**

Table 2 shows that water is the most stable class for 2020 with a probability of 0.96. Urban, plantation and forest are the most dynamic classes with transition probabilities of 0.33, 0.48, and 0.87. In these land cover classes; urban was mainly converted into plantation and forest was converted into plantation with respective probabilities of 0.33 and 0.12.
Table 3 shows that water remains the most stable class for 2025 with a probability of 0.99. Probabilities of forest converting to plantation and urban to plantation was shown as 0.17 and 0.18.



Actual 2020 LC map Predicted 2020 LC map

Figure 4 The actual 2020 land cover map (left) and the predicted 2020 land cover map (right)

Table 5 Area statistics for the actual 2020 LC map and predicted 2020 LC map

	Actual 2020		Predicted 2020	
Class	%	km^2	%	km^2
Urban	2.67	313.53	2.67	273.21
Forest	50.37	5909.40	66.29	6770.08
Water	18.57	2178.60	6.71	685.07
Bareland	0.96	112.780	0.27	27.63
Plantation	27.43	3217.60	24.05	2455.40

Figure 4 shows the actual 2020 land cover map with the predicted 2020 land cover map. Table 5 shows the area statistics of the actual 2020 land cover map and the predicted 2020 map. Urban areas were accurately predicted at 2.67%, Forest cover of the predicted map and actual map was 50.37 and 66.29% respectively with a difference 15.92%. This is because the model could not predict the drastic decrease of forest cover from 2010 and 2020. Plantation shows strong agreement with a difference of 3.38%. The model was only trained using forest-plantation transition probabilities as they were shown to be the major land change drivers, so water and bare land areas were not accurately predicted. In model validation, Overall Kappa was 0.4, indicating a fair agreement between the predicted and reference image. After model validation, a predicted land cover map for 2025 was created using 2010 and 2020 land cover map.

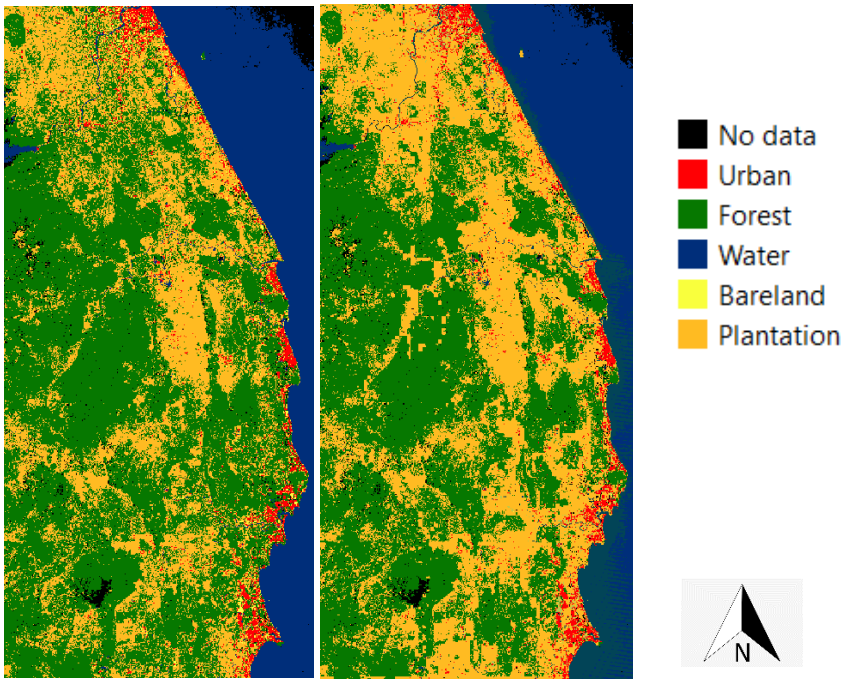


Figure 5 2020 LC map and predicted 2025 LC map

Table 6 Area statistics for 2020 LC map and 2025 LC map

	2020		predicted 2025	
Class	%	km2	%	km2
Urban	2.67	313.53	2.82	284.4396
Forest	50.37	5909.4	48.02	4839.5925
Water	18.57	2178.6	7.06	711.675
Bareland	0.96	112.78	1.04	104.6799
Plantation	27.43	3217.6	41.06	4138.2783
	Total	11731.91	Total	10078.6653

Figure 5 shows the 2020 land cover map and predicted 2025 land cover map. Table 6 shows the area statistics for 2020 and 2025 land cover maps. The 2025 land cover area shows decreasing values due to some pixels converting into no data pixels, particularly water areas. However, we can see that urban and plantation areas increases by 0.15% and 13.63% respectively while forest cover decreases by 2.35%. Bareland areas shows minimal change with 0.08%.

Land cover changes at each 40 roadkill locations were spatially analyzed by looking at the land cover around the road area of several years before and after the roadkill occurred using classified Landsat images between 2000-2020. Out of

40 locations, 27 showed changes from (1) forest to plantation, 6 showed (2) no significant changes, 4 showed (3) changes from forest to urban and 3 showed (4) changes from plantation to forest. Forest to plantation change surrounding the road was found to be the major factor of roadkill. Thus, areas that will change from forest to plantation in 2025 are recognized as roadkill risk areas.

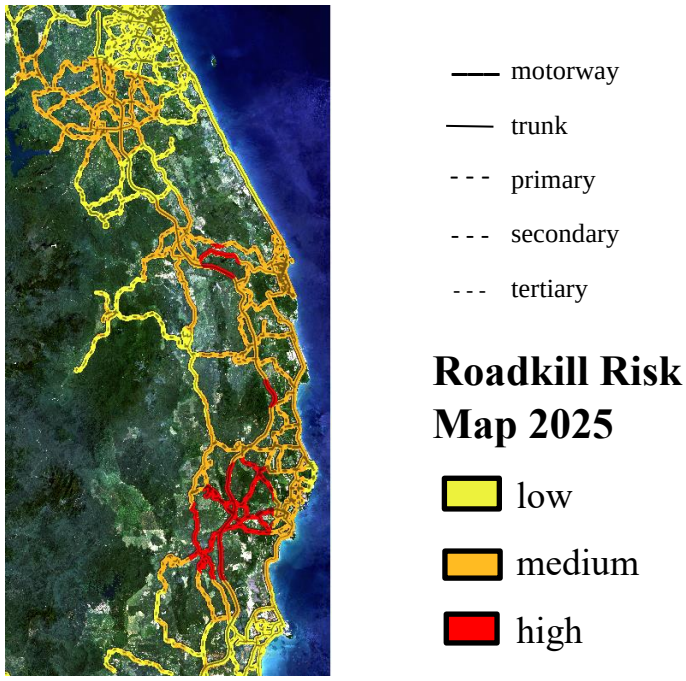


Figure 6 Roadkill risk map for the year 2025

Figure 6 represents the roadkill risk map for the year 2025. Motorways (highways) were found to have the highest roadkill risk overall. The roads near Kemaman River showed the highest roadkill risk. Malayan Tapir tend to roam in plantation areas, where the food is abundant, and near source of water. Therefore, roads intersecting these areas have high roadkill risk.

Using the classified images between 2000-2020, roadkill occurrences were found to be significantly influenced by changes from forest to plantation areas. Based on the predicted 2025 land cover, a roadkill risk map was produced based on past roadkill occurrences and the predicted land cover changes that influence roadkill.

This study showed that GIS and remote sensing data are valuable tools to identify factors that affect roadkill and predict areas with roadkill risk as a prevention method. High spatial resolution images, sufficient roadkill data and other explanatory variables that drive land cover change may be helpful in improving prediction accuracy. This research proved that roadkill risk prediction based on land cover changes to supplement limited roadkill data can be helpful in detecting risky areas. This can enable officials to take the appropriate action and warn drivers, thus preventing more Malayan Tapir deaths by roadkill.

7. REFERENCES

1. C. Traeholt, W. Novarino, S. Saaban, N. M. Shwe, A. Lynam, Z. Zainuddin, B. Simpson and S. Mohd, "The IUCN Red List of Threatened Species," 2016. [Online]. Available: https://www.iucnredlist.org/species/. [Accessed 24 January 2022].

2. "Department of Wildlife and National Parks (DWNP)," 2012. [Online]. Available: http://wildlife.gov.my/images/stories/penerbitan/kertas_maklumat/terbaru/Tapir%20BI.pdf. [Accessed 24 January 2022].

3. M. Alamgir, M. Campbell, S. Sloan, E. P. Wong and W. F. Laurence, "Road risks and environmental impact assessments in Malaysian road infrastructure projects," 2018.

4. "Denver Zoo," [Online]. Available: https://denverzoo.org/wp-content/uploads/2018/09/Malayan-Tapir.pdf. [Accessed 24 January 2022].

5. "Malayan Tapir," [Online]. Available: https://tapirs.org/tapirs/malayan-tapir/#:~:text=Threats%20to%20Their%20Survival,major%20factor%20in%20habitat%20loss.. [Accessed 24 January 2022].

6. D. Magintan, T. A. Rahman, E. Jiliun, Y. Adib, A. A. H. Abd Aziz, M. S. Mohd Suri, M. N. Ismail and A. K. A. Hashim, "Malayan tapir roadkill in Peninsular Malaysia from 2006 to 2019," *Journal of Wildlife and Parks,* vol. 36, 2021.

7. D. Magintan, M. B. M. Rufino and N. Cosmas, "Activity pattern on Malayan tapir (Tapirus indicus) in Temenggor Forest Reserve, Perak, through the use of the camera trapping technique," *Journal of Wildlife and Parks,* vol. 26, pp. 1-4, 2010.

8. G. R. Clements, A. J. Lynam, D. Gaveau, W. L. Yap, S. Lhota, M. Ghoosem, S. Laurance and W. F. Laurance, "Where and how are roads endangering mammals in Southeast Asia's forests?," *PLoS ONE,* vol. 9, no. 12, 2018.

9. R. Pagany, "Wildlife-vehicle collisions - Influencing factors, data collection and research methods," *Biological Conservation,* vol. 251, 2020.

10. T. Bartonicka, R. Andrášik, M. Duľa, J. Sedoník and M. Bíl, "Identification of local factors causing clustering of animal-vehicle collisions," *Journal of Wildlife Management,* vol. 82, no. 2, 2018.