A Cellular Automata Artificial Neural Network Approach for Land Use and Land Cover Prediction in Camiguin Province Using QGIS and MOLUSCE

Usher Raymond Abainza, Ruszed Ayad, Karylle dela Cruz, Ian Jure Macalisang, and Nicole Menorias

Abstract. This study employed a Cellular Automata Artificial Neural Network (CA-ANN) approach within the MOLUSCE plugin in QGIS to predict future Land Use and Land Cover (LULC) changes in Camiguin Province, Philippines. Landsat 8 imagery and ancillary datasets were utilized, and the Spectral Angle Mapper (SAM) method demonstrated high accuracy in LULC classification. Using the CA-ANN model, LULC changes were projected for 2028, indicating ongoing forest conversion to built-up areas and a slight expansion of water bodies. These findings emphasize the importance of implementing sustainable land management practices in Camiguin to mitigate potential environmental challenges and disaster risks.

Keywords: LULC prediction; Cellular Automata Simulation; QGIS; SCP plugin; MOLUSCE plugin.

1 Introduction

Understanding future land use and land cover (LULC) changes is crucial for effective environmental management and sustainable development [1]. Accurate LULC predictions are vital for assessing and mitigating risks from natural hazards such as floods and landslides [2]. They also guide sustainable land management practices, ensuring responsible land use planning and resource allocation to minimize environmental impacts and promote biodiversity conservation [3]. Furthermore, LULC predictions inform urban planning initiatives, facilitating efficient urban growth and optimized infrastructure development [4]. Finally, these predictions are essential for assessing the impacts of LULC changes on water availability and quality, supporting sustainable water resource management and contributing to effective climate change mitigation strategies [5]. This study employs advanced machine learning techniques, specifically the Cellular Automata Artificial Neural Network (CA-ANN) method within the MOLUSCE plugin in QGIS, to predict future LULC changes. This approach aligns with recent research emphasizing the effectiveness of machine learning for LULC forecasting and its importance in informing sustainable practices.

2 Materials and Methods

2.1 Study Area Selection

Camiguin Island was selected as the study area for this LULC prediction research due to its unique characteristics and pressing environmental management needs. As a small volcanic island in the Philippines, Camiguin possesses diverse ecosystems, ranging from coastal areas and mangroves to upland forests and volcanic peaks [6]. This diversity, coupled with increasing population pressure and tourism development, makes the island highly susceptible to LULC changes with significant implications for biodiversity, water resources, and disaster risk [7]. Recent studies have highlighted the vulnerability of Camiguin's ecosystems to deforestation, land degradation, and the impacts of climate change [8]. Furthermore, the island's volcanic nature poses significant hazards, underscoring the need for accurate LULC predictions to inform hazard mapping and disaster preparedness strategies [9]. By focusing on Camiguin, this research aims to provide valuable insights for local policymakers and stakeholders to promote sustainable land management practices and enhance the island's resilience to environmental challenges.

2.2 Data Collection and Preprocessing

In this study, Landsat 8 satellite imagery of Camiguin was acquired from the United States Geological Survey (USGS) EarthExplorer platform for the years 2013, 2018, and 2023. Landsat 8 imagery was selected due to its consistent temporal coverage, high spatial resolution, and suitability for analyzing land use/land cover changes over time.

To ensure the accuracy of the analysis, preprocessing steps were conducted on the Landsat 8 imagery. Atmospheric correction was applied using the Semi-Automatic Classification Plugin (SCP) in QGIS.

This process involved converting the Top of Atmosphere (TOA) reflectance to Surface Reflectance (SR) to

minimize atmospheric effects, such as scattering and absorption caused by gases and aerosols. The corrected surface reflectance values provided more reliable inputs for subsequent analyses by enhancing the accuracy of spectral indices and classification results.
To enhance spatial analysis, ancillary datasets were incorporated, including a Digital Elevation Model (DEM) from the Shuttle Radar Topography Mission (SRTM) for terrain analysis, vector datasets on roads, waterways, and buildings to assess infrastructure development impacts, tropical cyclone tracks to evaluate extreme weather effects, and administrative boundaries for spatial organization. These datasets provide a comprehensive understanding of Camiguin's land use and environmental changes.

Table 1. Data Collections and Data Sources

Data Type	Data Source	Data Description
Landsat 8 Imagery for 2013, 2018, and 2023	https://earthexplorer.usgs.gov/	Satellite imagery for the specified years
Digital Elevation Model (DEM)	https://earthexplorer.usgs.gov/	SRTM Arc-Second Global (Shuttle Radar Topography Mission)
Subnational Administrative Boundaries	https://gadm.org/download_country.html	Administrative boundaries of Camiguin
Roads (Lines), Waterways (Lines & Polygons), Buildings (Polygons), Seaports (Polygons), Water Courses (Lines), Tropical Cyclones Tracks (Lines)	https://data.humdata.org/	Infrastructure and disaster-related data

2.3 Land Use/ Land Cover Classification

Three primary classes were identified: **water**, **built-up area**, and **forest** to analyze land use and land cover (LULC) in the study area. These classes were selected based on their ecological and socio-economic significance to the region. A total of 10 training polygons and 8 validation polygons were delineated for each class to ensure robust classification and validation processes. The training polygons were carefully selected to capture the spectral variability within each LULC class, while validation polygons were chosen independently to avoid bias and ensure accurate performance evaluation.
To assess classification performance, validation points were systematically generated and used to evaluate the accuracy of three widely applied classification algorithms.

Minimum Distance Classification (MDC). This method calculates the Euclidean distance between a pixel's spectral signature and the mean spectral signature of each class, assigning the pixel to the class with the shortest distance.

Maximum Likelihood Classification (MLC). This probabilistic approach assumes a normal distribution of spectral data and assigns pixels to the class with the highest probability of membership based on statistical variance and covariance.

Spectral Angle Mapper (SAM). A physically-based algorithm that measures the spectral angle between the pixel and reference spectra in multi-dimensional space, making it particularly effective for materials with distinct spectral features.

2.4 Accuracy Assessment

Among the three methods, the **Spectral Angle Mapper (SAM)** consistently demonstrated the highest accuracy, making it the preferred method for our analysis.

Table 2. Overall Accuracy of the Classification Methods

Study Year	MDC	MLC	SAM
2013	95.94%	16.25%	96.17%
2018	95.26%	92.48%	95.31%
2023	95.24%	96.40%	97.70%

The superior performance of SAM can be attributed to its robustness in distinguishing spectral differences, particularly in complex and heterogeneous landscapes. Based on these results, SAM was employed for the final classification of land cover in the Landsat 8 images. This ensured a reliable and consistent mapping of LULC changes over the analyzed years, providing valuable insights into the spatiotemporal dynamics of the study area.

2.5 Preparation of Spatial Variables

Spatial variables were prepared by importing the LULC map and study area boundary into QGIS, ensuring CRS consistency. Ancillary data were converted to raster format, aligned with the study area, and transformed into proximity rasters. Each raster was then clipped to the study area boundary to maintain focus on the region of interest.

2.6 Evaluating Correlation

Correlation analysis within MOLUSCE revealed key relationships between predictor variables.

Table 3. Pearson's Correlation Between Predictors

	Buildings	Elevation	Roads	Seaports	Tropical Cyclone Tracks	Water Courses	Waterways_1	Waterways_2
Buildings	–	0.9142	0.6756	0.1665	-0.0988	0.4960	-0.0687	0.2636
Elevation		–	0.6146	0.1244	-0.0924	0.4638	-0.1051	0.2514
Roads			–	0.0624	-0.0682	0.2856	0.0145	0.1817
Seaports				–	-0.0609	-0.0843	0.0653	-0.2071
Tropical Cyclone Tracks					–	-0.0394	0.0770	-0.0119
Water Courses						–	0.006	0.2333
Waterways_1							–	0.0617
Waterways_2								–

A strong positive correlation (0.914) between buildings and elevation suggests urbanization in higher areas, possibly driven by factors like cooler temperatures or scenic views. Moderate positive correlations were also found between buildings and roads (0.676) and elevation and roads (0.615), indicating infrastructure development patterns. These insights will inform our LULC prediction model within MOLUSCE, enabling us to capture the complex dynamics of land use change.

2.7 Area Change Analysis

Area change analysis, facilitated by MOLUSCE, quantifies and characterizes LULC transformations through the comparison of LULC maps from different time points. This analysis yields valuable metrics, such as gross change, net change, and transition matrices, which provide insights into the magnitude, direction, and spatial distribution of LULC changes, along with trends and potential tipping points. This study conducted two such analyses, comparing LULC maps from 2013 to 2018 and 2018 to 2023, to comprehensively examine LULC dynamics across two distinct time intervals.

2.8 Cellular Automata Simulation

Employing the 2023 LULC map, this study utilized Cellular Automata (CA) simulation within MOLUSCE to predict future land use changes. By representing the landscape as a grid of cells and simulating transitions based on predefined rules and neighboring interactions, CA effectively captures spatial dynamics inherent in LULC changes. Integration with Artificial Neural Networks (ANN) allows MOLUSCE to leverage historical LULC data and predictor variables for calibrating transition rules, enabling accurate and spatially explicit predictions. This CA-ANN approach facilitates the simulation of complex interactions between LULC classes and their driving forces, generating robust predictions to inform sustainable land management decisions.

2.9 Validation

Validation is crucial for assessing accuracy and reliability of LULC predictions. Comparing the MOLUSCE-generated 2023 LULC map with a reference map derived from high-resolution satellite imagery quantifies agreement between predicted and actual LULC patterns. This process allows for identification of discrepancies, refinement of model parameters for improved performance, and ensures LULC predictions are robust and trustworthy for informing sustainable land management and planning decisions.

3 Results and Discussion

3.1 Neural Network Training

A multi-layer perceptron artificial neural network (ANN) with four hidden layers was employed to discern complex patterns and forecast LULC transitions. Utilizing a learning rate of 0.001 and a momentum value of 0.010, the network was trained for 500 iterations on a stratified dataset of 5000 samples, with a neighborhood size of 1 pixel. This configuration enabled effective learning of LULC change dynamics and generation of robust predictions.

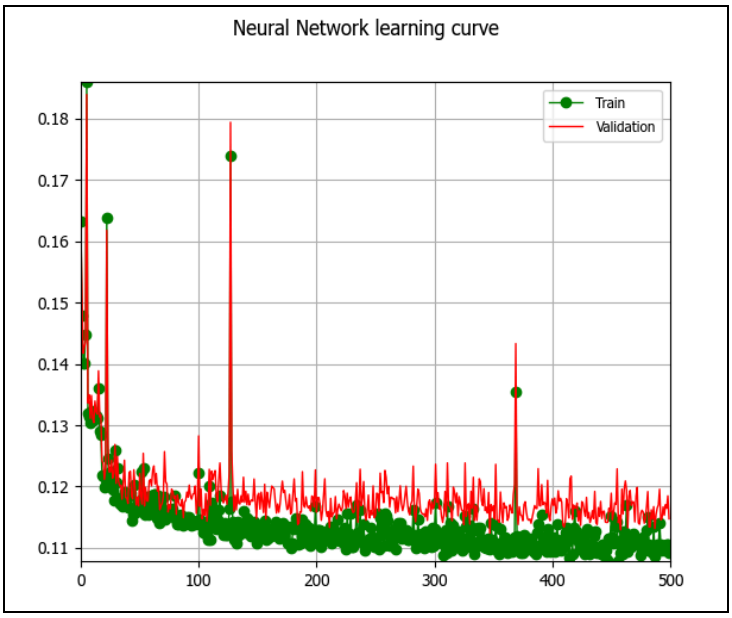

Fig. 1. This depicts the neural network's learning process, with the *green line* representing training error and the *red line* representing validation error. Initially, both errors are high, but the training error steadily decreases as the model learns. The validation error stabilizes after approximately 200 iterations, with occasional spikes potentially due to noisy data or overfitting. By iteration 400, both errors stabilize near 0.11, suggesting the model has reached its optimal state.

Table 4. Performance Metrics of Neural Network Training

Parameter	Value
Overall Accuracy	-0.00235
Min Validation Overall Error	0.11239
Current Validation Kappa	0.77886

The performance metrics presented in Table 4 offer insights into the predictive capacity of the LULC model. Minimal overfitting is suggested by a slight decrease in overall accuracy, while strong generalization ability is evidenced by the low validation error. Furthermore, the high Kappa value signifies accurate LULC class prediction, underscoring the model's efficacy in capturing land cover dynamics.

3.2 LULC Change Analysis

This section presents the area changes observed in Camiguin's land use and land cover (LULC) categories across the years 2013, 2018, and 2023, including a prediction for 2028.

Table 5. Class Statistics

Year	Forest	Built-up Areas	Water
2013	231.07 sq. km.	11.91 sq. km.	1.13 sq. km
2018	225.00 sq. km.	17.78 sq. km.	1.33 sq. km
2023	208.88 sq. km.	34.44 sq. km.	0.79 sq. km
2028	208.59 sq. km.	33.96 sq. km	1.56 sq. km.

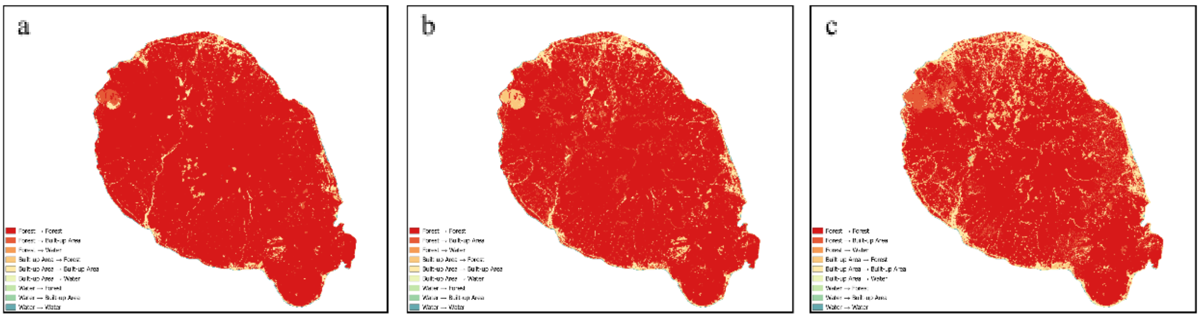

Fig. 2. This figure depicts the LULC transitions between the years *(a) 2013 to 2018, (b) 2018 to 2023*, and potential transitions from *(c) 2023 to 2028*. *Red* indicates areas where LULC remained unchanged (forest remaining forest). Transitions between forest, built-up areas, and water are represented by other colors as per the legend. Transitions are most notable along the coastlines and near existing settlements. (*a) (2013-2018)*: Dominated by forest conversion to built-up areas, with minimal transition to water. *(b) (2018-2023*): While forest-to-built-up transitions persist, some built-up areas revert to forest. Water body expansion is more pronounced, mainly replacing forest areas. *(c) 2023-2028 (predicted)*: Projects substantial forest loss to both built-up areas and water.

Table 6. Transition Matrix

Year		2028		
	LULC Category	Forest	Built-up Areas	Water
2023	Forest	0.9038	0.0955	0.0007
	Built-up Areas	0.5742	0.4051	0.0207
	Water	0.0274	0.0810	0.8917

Table 6 projects LULC change probabilities in Camiguin for 2023-2028. Forest exhibits a high probability of persistence, but with notable transitions to built-up areas. Built-up areas also show persistence, yet with considerable reversion to forest. Water bodies are predicted to remain relatively stable. These projections highlight continued pressure on forest resources due to urbanization, potential for reforestation, and relative stability in water body extent.

3.3 Validation

The MOLUSCE plugin integrated with Cellular Automata Artificial Neural Network (CA-ANN) approach employed LULC data from 2013 and 2018 along with spatial variables to project LULC for 2023. A validation ANN Kappa value showed 0.77, and after obtaining the projected LULC, the classified LULC map of 2023 and the projected 2023 were compared. The correctness percent showed 85.89% and an overall kappa value of 0.37. Figure 4 and Table 6 show the classified and predicted maps and statistics for 2028.

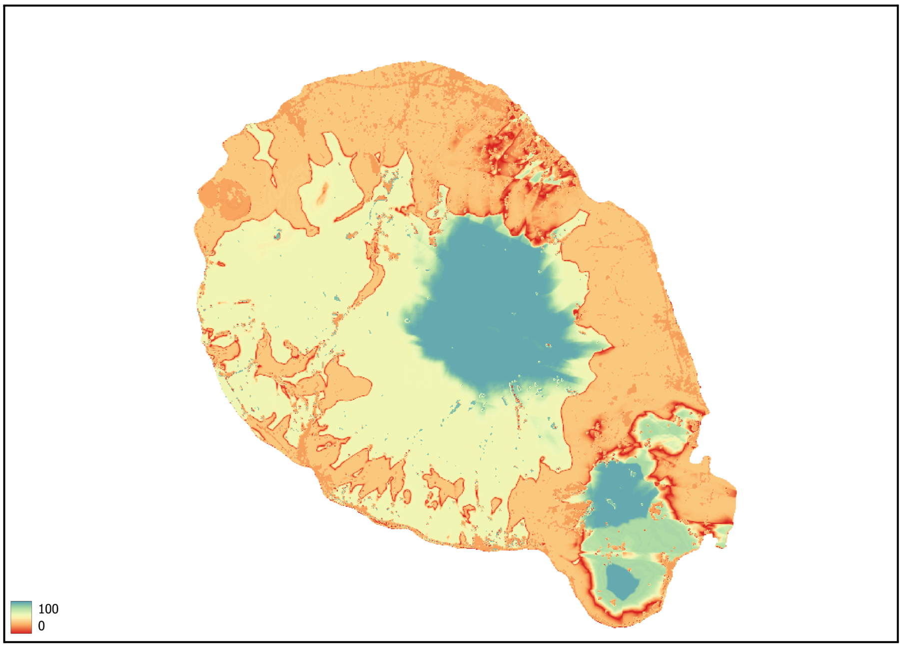

Fig. 3. This figure depicts the certainty function for the 2023 LULC prediction in Camiguin. *High certainty (cool colors)* is observed in areas with distinct topographic features and homogenous land use, such as the central volcanic peaks and coastal plains, indicating greater confidence in the model's predictions. Conversely, *lower certainty (warm colors)* is associated with transition zones and heterogeneous landscapes, highlighting the challenges of predicting LULC changes in these areas.

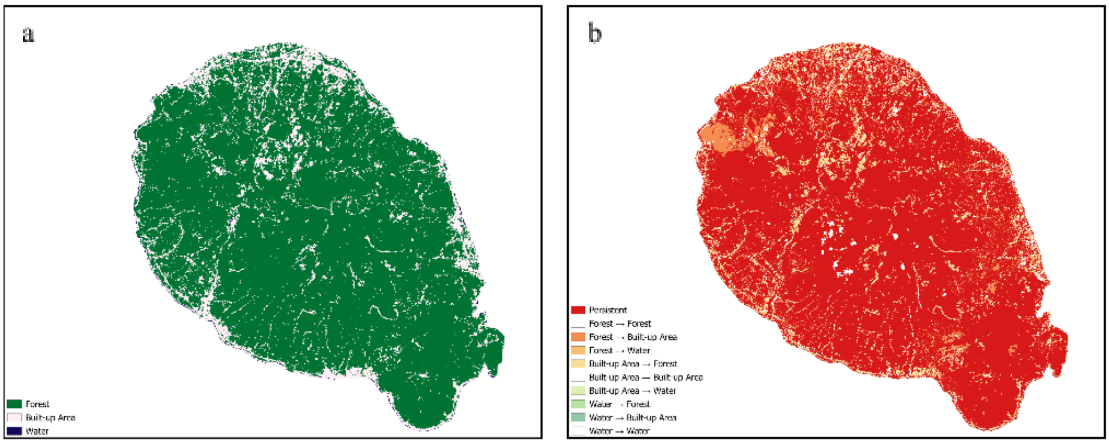

Fig. 4. This figure presents a visual comparison between the *(a) actual* and *(b) predicted* LULC for Camiguin in 2023. While the model generally captures the spatial distribution of persistent forest cover, particularly in the central volcanic region and coastal areas, discrepancies arise in predicting transitions. Notably, the model overestimates forest persistence and underestimates built-up area expansion, particularly along coastlines and major roads. Furthermore, limitations in capturing fine-scale patterns and transitions involving water bodies are evident. This comparative analysis highlights the need for further model refinement to enhance accuracy in predicting LULC dynamics.

Table 6. Validation Metrics

Correctness	Kappa (overall)	Kappa (histogram)	Kappa (location)
85.89%	0.37	0.83	0.44

Table 6 presents the validation metrics derived from comparing the predicted 2023 LULC map of Camiguin with the corresponding actual LULC map. The overall correctness of 85.89% indicates a high degree of agreement between the predicted and actual maps, suggesting that the model accurately classified a majority of the pixels. However, the Kappa statistics provide a more nuanced perspective, accounting for the possibility of agreement by chance. The overall Kappa value of 0.37, while indicating fair agreement, suggests that there is room for improvement in the model's predictive performance. The Kappa histogram value of 0.83 highlights a strong agreement in the frequency distribution of LULC classes, indicating that the model effectively captures the overall proportions of different land cover types. Conversely, the Kappa location value of 0.44 suggests moderate agreement in the spatial allocation of LULC classes, implying that while the model predicts the correct proportions of land cover types, it may not always place them in the precise locations. These metrics collectively provide a comprehensive assessment of the model's performance, highlighting its strengths and areas for potential refinement.

4 Conclusion

This study used the CA-ANN method to predict LULC changes in Camiguin Island, Philippines. Landsat 8 imagery and ancillary datasets were utilized. The SAM method achieved high accuracy for LULC classification. The CA-ANN model was used to predict LULC changes for 2028, projecting a continued trend of forest conversion to built-up areas and a slight expansion of water bodies. This indicates a trend of increasing urbanization in Camiguin, which can lead to environmental challenges such as deforestation, habitat loss, and increased pressure on natural resources.

The study projects continued deforestation in Camiguin, which can have significant environmental and socio-economic impacts. Deforestation can lead to soil erosion, loss of biodiversity, and decreased water quality. It can also impact the livelihoods of communities that depend on forest resources. The predicted slight expansion of water bodies requires careful monitoring and management of water resources in Camiguin. Changes in water availability can affect agriculture, domestic water supply, and the island's overall ecosystem.

Camiguin's vulnerability to natural hazards, such as typhoons and volcanic eruptions, necessitates effective disaster risk management strategies. Changes in LULC can influence the island's susceptibility to natural disasters. For instance, deforestation can increase the risk of landslides, while urbanization can impact drainage patterns and increase flood risk.
The predicted LULC changes highlight the need for sustainable land management practices in Camiguin. This includes promoting responsible land use planning, reforestation efforts, and sustainable agricultural practices. The accuracy of the CA-ANN model was validated by comparing the predicted 2023 LULC map with a classified LULC map derived from the same Landsat data. The validation process showed a high degree of agreement between the predicted and actual LULC maps, with an overall correctness of 85.89% and a Kappa value of 0.37. This study highlights the potential of the CA-ANN model in predicting LULC changes and its value for informing sustainable land management decisions in Camiguin Province. The findings contribute to a better understanding of LULC change dynamics and support informed decision-making for sustainable land management practices.

References

1. Song, X.P., Hansen, M.C., Stehman, S.V., Potapov, P.V., Tyukavina, A., Vermote, E.F., Townshend, J.R.: Global land change from 1982 to 2016. Nature[1] 560(7720), 639--643 (2018)
2. Depietri, Y., McPhearson, T.: Integrating the social and ecological dimensions of urban green infrastructure: a policy-oriented framework. Urban Forestry & Urban Greening 21, 100--111 (2017)
3. Lausch, A., et al.: Understanding and managing connectedness in the Earth’s land system. Current Opinion in Environmental Sustainability 14, 16--23 (2015)
4. Liu, Y., et al.: Effects of land use and land cover change on regional climate: a review. Journal of Geographical Sciences 31(1), 3--21 (2021)
5. Guerrero, M.C., et al.: Assessing the impact of land use/land cover change on the water balance of a tropical watershed in the Philippines. Water 14(14), 2188 (2022)
6. Ong, P.S., Afuang, L.E., Rosell-Ambal, R.G.B. (eds.): Philippine Biodiversity Conservation Priorities: A Second Iteration of the National Biodiversity Strategy and Action Plan. Department of Environment and Natural Resources - Protected Areas and Wildlife Bureau, Conservation International Philippines, Biodiversity Conservation Program - University of the Philippines Center for Integrative and Development Studies, Quezon City, Philippines (2016)
7. Guerrero, M.C., et al.: Assessing the impact of land use/land cover change on the water balance of a tropical watershed in the Philippines. Water 14(14), 2188 (2022)
8. Lasco, R.D., et al.: Climate change and forest ecosystems in the Philippines: vulnerability, adaptation and mitigation. Philippine Journal of Science 148(1), 151--167 (2019)
9. Philippine Institute of Volcanology and Seismology (PHIVOLCS). (2018). Camiguin Volcano Hazard Maps.

Video Presentation

A CA-ANN Approach for Land Use and Land Cover Prediction in Camiguin Province Using QGIS and …