Application of AI-based Modelling and Remote Sensing to Assess Inselberg Habitats in Gamapaha District in Sri Lanka

Jayasekara D. [1*] and Herath H.M.B.[2]

[1]Department of Zoology, Faculty of Applied Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka.

[2] Department of Geography, Faculty of Humanities and Social Sciences, University of Sri Jayewardenepura, Gangodawila, Nugegoda, Sri Lanka.

*dulan@sjp.ac.lk

Abstract: *Inselbergs are isolated rock formations that develop as a result of differential weathering and erosion advances during the formation of topographic etch plains. Inselbergs possess unique ecosystem characteristics, cultural significance, and geomorphological importance worldwide, including in Sri Lanka. However, studies on the application of geospatial technologies to identify the spatial distribution of Inselberg habitats to evaluate their geomorphological and ecological features in Sri Lanka are still lacking. Identification of the spatial distribution of these habitats and mapping of Inselbergs for later retrieval and reference will enhance the management and conservation interventions. In this context, the main objective of this study was to develop models to identify the spatial distribution of Inselberg habitats in the Gampaha district using AI-based machine learning algorithms and remote sensing data. Random forest (RF) and Support Vector Machines (SVM) algorithms were used to model the Inselberg areas based on 41 topographic and spectral variables in R platform. Main R packages used for the analysis were "caret", "randomForest", "raster", "sf", and "e1071". Model training was conducted using a training sample of 1035 points. The best-performing model was identified as SVM based model with a Kappa accuracy of 70.1%, a global accuracy of 94.5% and a sensitivity of 98.9%. Digital Elevation ModelFerrous Mineral Ratio, Sentinel Band 8 (NIR), Soil Adjusted Vegetation Index, Terrain Ruggedness Index, , slope, Terrain Surface Convexity, Downslope Difference and Clay Minerals were identified as raster covariates that contributed most to the generated models. The findings of this study have shown significant potential of machine learning integrated remote sensing data analysis in identifying Inselberg habitats. The generated results can be effectively used for the conservation of Inselberg habitats in Gamapaha district, and to preserve the associated unique cultural identity and geodiversity. The models developed in this study can also be applied to the rest of the areas of Sri Lanka.*

Keywords: Random Forest, Support Vector Machines, Classification, Machine Learning, Sentinel, Geospatial Science

Introduction

Inselbergs are rocky outcrops generally composed of granitic or gneiss matrix, with monolithic proportions (da Silveira et al., 2022). Inselbergs have been defined as remnants of erosion processes forming isolated mountains which can range in elevation from few to several hundred meters (Burke, 2003; Coor et al., 1993). These landscapes are identified as useful for the study of biogeographical patterns and processes (Barthlott & Porembski, 2000), and rich in species and endemism as a result of low dispersibility and high clonality of native species (Hopper et al., 2021).

Environmental mapping and land cover monitoring for conservation purposes are among the main applications of satellites dedicated to Earth observation (Rodriguez-Galiano et al., 2012). For this purpose, a variety of classification methods have been developed (Rodriguez-Galiano et al., 2012). Satellite image classification can be performed by unsupervised algorithms including K-means or ISODATA, supervised parametric algorithms such as Maximum Likelihood (Jensen, 2005), and various alternative machine learning algorithms (da Silveira et al., 2022). Artificial Intelligence (AI) encompasses various methods and techniques that enable machines to mimic human intelligence, while machine learning focuses on using data-driven approaches to enable systems to learn patterns and make decisions. Machine learning which specifically refers to algorithms that learn from data can be considered as a subset of AI. AI techniques can be integrated into Geographic Information Systems (GIS) to automate land use analysis and machine learning approach has been widely accepted, as evidenced by their use for mapping land cover over large geographical areas (Huang et al., 2020; Maxwell et al., 2018). The use of such algorithms allows to consider a wide variety of predictive covariates (Souza et al., 2018). Studies have demonstrated that those algorithms are better than conventional parametric classifiers, due to their greater precision and better performance in supervised classification, especially owing to their efficient database processing (Di Shi and Yang, 2016; Gašparović and Jogun, 2018; Ghimire et al., 2012; Souza et al., 2018; Yu et al., 2014; Zeraatpisheh et al., 2017). Due to their excellent ability to deal with nonlinear relationships between dependent and independent variables (Voyant et al., 2017a), machine learning algorithms employed in the supervised classification of satellite images are efficient and capable of modelling spectral signatures of different classes of land use and coverage (Emadi et al., 2020; Maxwell et al., 2018).

Studies on geomorphological and ecological features in Sri Lanka's Inselbergs are still incipient and limited to the work conducted on geotourism, geoheritage (Katupotha & Ravibhanu, 2020), individual sites (Katupotha & Kodituwakku, 2015) and some preliminary work on investigating the ecology and biodiversity value. Due to the Gondwanaland origin of Sri Lanka, being one of the oldest islands on earth, the island is rich with a vast array of Inselbergs distributed throughout the island. However, in addition to the general topographic mapping approaches, no focused efforts have been put on to model these significant habitat types identifying the factors that contribute to their geomorphological and ecological structure. Hence, such an endeavour would facilitate the accurate mapping of Sri Lanka's Inselberg habitats which would help decision making and management interventions on conservation.

Wet zone of Sri Lanka harbours the most disturbed landscapes (Premakantha et al., 2021) because of significant anthropogenic influence over the last several centuries. Moreover, having the highest human population density of the island the Western province faces critical disruptions in its natural landform structures (Nawarathna Banda, 2008; Premakantha et al., 2021). Inselberg habitats of Gamapaha district host the last remnants of important forest fragments, mainly due to difficulty of access and due to the custodianships of Buddhist monasteries. Despite the distribution of a considerable number of Inselberg habitats in Gampaha district, no geomorphological or geospatial investigations have been conducted so far. In addition, not many conservation units have been created to protect such environments to facilitate the conservation of the last remnants of the districts forest fragments that surround the rock outcrops.

Given the environmental, cultural and social relevance of the typical landscapes of Inselbergs in Gampaha district, it is possible to carry out more accurate and systematic mapping based on remote sensing and machine learning approaches for purposes of protection and conservation of strategic environments for islands biodiversity.

Given the above, developing models to identify and map the Inselberg habitats in the Gampaha district, Sri Lanka will facilitate the effective management and conservation of Inselberg habitats as well as environmental values related to those habitats.

Literature Review

Inselberg ecosystems harbour extraordinarily high numbers of geographically restricted and threatened species, functioning as terrestrial habitat islands (Bussell & James, 1997). Whether or not Inselbergs can be considered 'islands' in the classic sense of island biogeography (MacArthur & Wilson, 1967) is debatable. However, most of the Inselbergs bestow the impression of the concept of island biogeography theory due to the isolation from the surrounding areas in both a geomorphological and ecological sense (S Porembski, 2000). Inselberg habitats in Gampaha district carry multiple social, cultural, and environmental values which are not properly understood. One of the major drawbacks in this regard is not identifying Inselberg habitats as unique ecosystems with specialized characteristics. Therefore, identification of the spatial distribution of these habitats and mapping them for later retrieval and reference will enhance the management and conservation interventions. Furthermore, environmental values of these habitats are not properly understood. The remaining natural forest cover of the district is a mere 1.2% (Premakantha et al., 2021). The general distribution of the Inselberg habitats seems to be associated with the remaining forest areas of the district based on the observations of satellite images. Therefore, further investigating whether there's a significant association in the two factors is important. One of the most significant issues related to the Inselberg habitats of Gampaha district is the destroying these habitats to establish rock quarries and extraction of granite for the construction industry. Furthermore, the large-scale development projects such as building of highways will have significant impacts on the integrity of these unique ecosystems. Some of the Inselbergs in the area are preserved due to the custodianship of Buddhist monasteries (Katupotha & Gamage, 2023). However, there remains a question up to which level these habitats can be modified to keep the natural phenomena and associated ecological values intact.

Land cover mapping and classification

Data obtained from dedicated earth observation satellites are highly applied for environmental mapping and land cover monitoring (Galiano et al., 2012). A variety of classification methods are available for this purpose (Galiano et al., 2012) and the technology keeps updating with advanced and improved imagery as well as novel analysis techniques (da Silveira et al., 2022).

The Sentinel satellite constellation, developed by the European Space Agency (ESA) as part of the Copernicus program, has become a cornerstone in modern remote sensing (Jutz & Milagro-Pérez, 2018). Equipped with advanced sensors, including radar and multispectral instruments, Sentinel satellites offer a comprehensive and consistent Earth observation platform. These satellites provide high-resolution and frequent imaging of the Earth's surface, allowing for a wide range of applications such as environmental monitoring, disaster management, and land cover analysis.

Machine learning in satellite image classification

Due to their excellent ability to deal with nonlinear relationships between dependent and independent variables (Voyant et al., 2017b), machine learning algorithms employed in the supervised classification of satellite images are efficient and capable of modeling spectral signatures of different classes of land use and coverage (Emadi et al., 2020; Maxwell et al., 2018). Support Vector Machine (SVM) (Cortes & Vapnik, 1995; Mountrakis et al., 2011), Random Forest – RF (Breiman, 2001), KKNN (Hechenbichler & Schliep, 2004) and Artificial Neural Networks (Mas & Flores, 2008) are some of the prominent machine learning algorithms currently used (da Silveira et al., 2022).

Remote Sensing in inselberg mapping

The application of Remote Sensing in inselberg mapping spans various disciplines, including geology, ecology, hydrology, archaeology, and urban planning. It enhances our understanding of these unique landforms and supports sustainable management and conservation efforts. Most of the focused research on Inselberg ecosystems have been initiated after the late 90's following the work of (Barthlott and Porembski (2000) where inselbergs have been identified as model ecosystems by their work. (Burke (2003) has identified global trends of inselbergs with regard to: bioclimatic positions, potential for providing habitat niches, and human impacts that may influence ecological processes. Several studies are available in the literature where the inselberg associated plant communities and diversity have been assessed (da Silveira et al., 2022; S Porembski, 2000; Stefan Porembski & Barthlott, 2000; Yates et al., 2019). The only available work on modelling the inselberg ecosystems has been conducted by (da Silveira et al. (2022) which has been conducted focusing the Atlantic Forest and Caatinga domains in Brazil.

Inselberg mapping approaches in Sri Lanka

The number of studies focused on inselberg ecosystems in Sri Lanka is scarce. However, (Stefan Porembski & Barthlott, 2000) identify India and Sri Lanka as well-known areas for the widespread occurrence of inselbergs. (Chandrajith (2020) briefly mentions the distribution of inslebergs in the dry zone of coastal plain of Sri Lanka. Work of Katupotha highlights the related work in Sri Lanka with multifaceted approaches linking vegetation diversity and culture in localities such as Pidurangala (Katupotha, 2015; Katupotha & Kodituwakku, 2015), Manewakanda and Danduwellawa in Kala Oya basin (De Jayawardena, 2015), geo-tourism and geo-heritage (Katupotha & Ravibhanu, 2020), and one conference proceedings regarding the overall biodiversity and ecology of Inselbergs of Sri Lanka in general. Several species-specific zoological studies briefly acknowledge the importance of inselberg ecosystems of Sri Lanka (Bauer & Das, 2000; Kittle et al., 2018; Weerasekara et al., 2021).

Geospatial technology and machine leaning integrated remote sensing assessments to identify Inselberg habitats in Sri Lanka are not available in the literature. Therefore, this research marks a first step towards such an endeavour and can be used as a baseline for future Inselberg mapping of the island.

Methodology

Study Area

The study was conducted covering the area of Gampaha district situated in the Western province of Sri Lanka (Figure 2). Gampaha district is bound by the districts of Kurunegala and Puttalam to the North, Kegalle to the East and Colombo to the South (Nawarathna Banda, 2008). The northern natural boundary is "*Maha Oya*" river and Southern boundary is "*Kelani*" river. The Indian Ocean borders the western margin of the district. Gampaha district spans for an extent of 1,387 square kilometres and it is also the second most human populated district of the country with a population of 2,443,000 with an estimated 1,800/km^2 population density (Fernando, 2018; Withanage et al., 2018). The district resides in the coastal plain of the island with an altitude ranging from sea level to a maximum of 450 m a.s.l (Fernando, 2018). However, most of the district display a plain topography and the North and Eastern areas consists of Inselbergs in a scattered distribution (Fernando, 2018).

Due to the high population density and anthropogenic activities, the remaining natural forest cover of the area is limited to 1.2% of the total area (Premakantha et al., 2021).

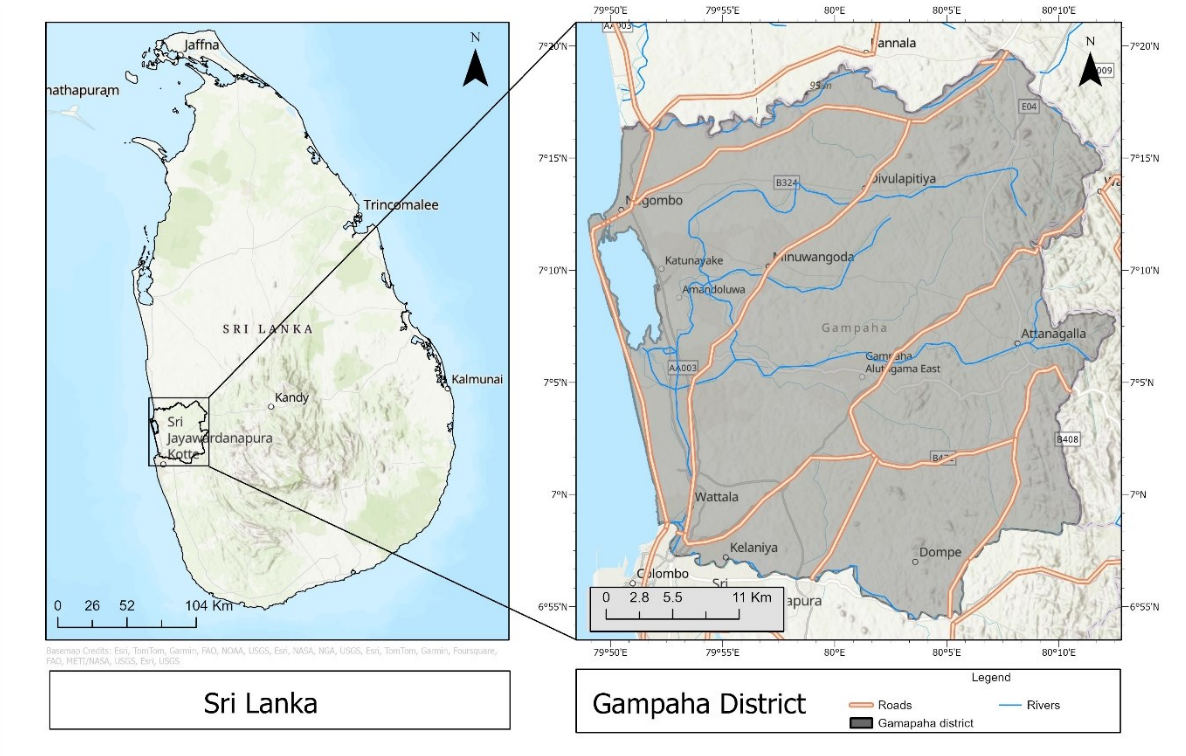


Figure 1: Study area map

Data Acquisition

Two types of covariates were used for the model development: spectral covariates and topographic covariates. The workflow followed in this process is depicted in Figure 2.

Google Earth Engine (GEE) geospatial processing service in the Google Cloud Platform was used to filter atmospherically corrected Surface Reflectance (SR) Sentinel 2A satellite images from Copernicus Open Access Hub database. Cloud and cirrus correction were obtained using the “maskS2clouds” function in GEE. The mosaic function in Google Earth Engine (GEE) was used to combine multiple Sentinel 2A images from the time period 1st January 2020 to 31st December 2024 into a single image. A multiband composite image was generated including the spectral bands B2, B3, B4, B7, B8, B11 and B12. The generated multispectral composite image was clipped to the area of Gamapaha district and downloaded from the GEE for further analysis in Arc GIS Pro 3.2 (ESRI, Redlands).

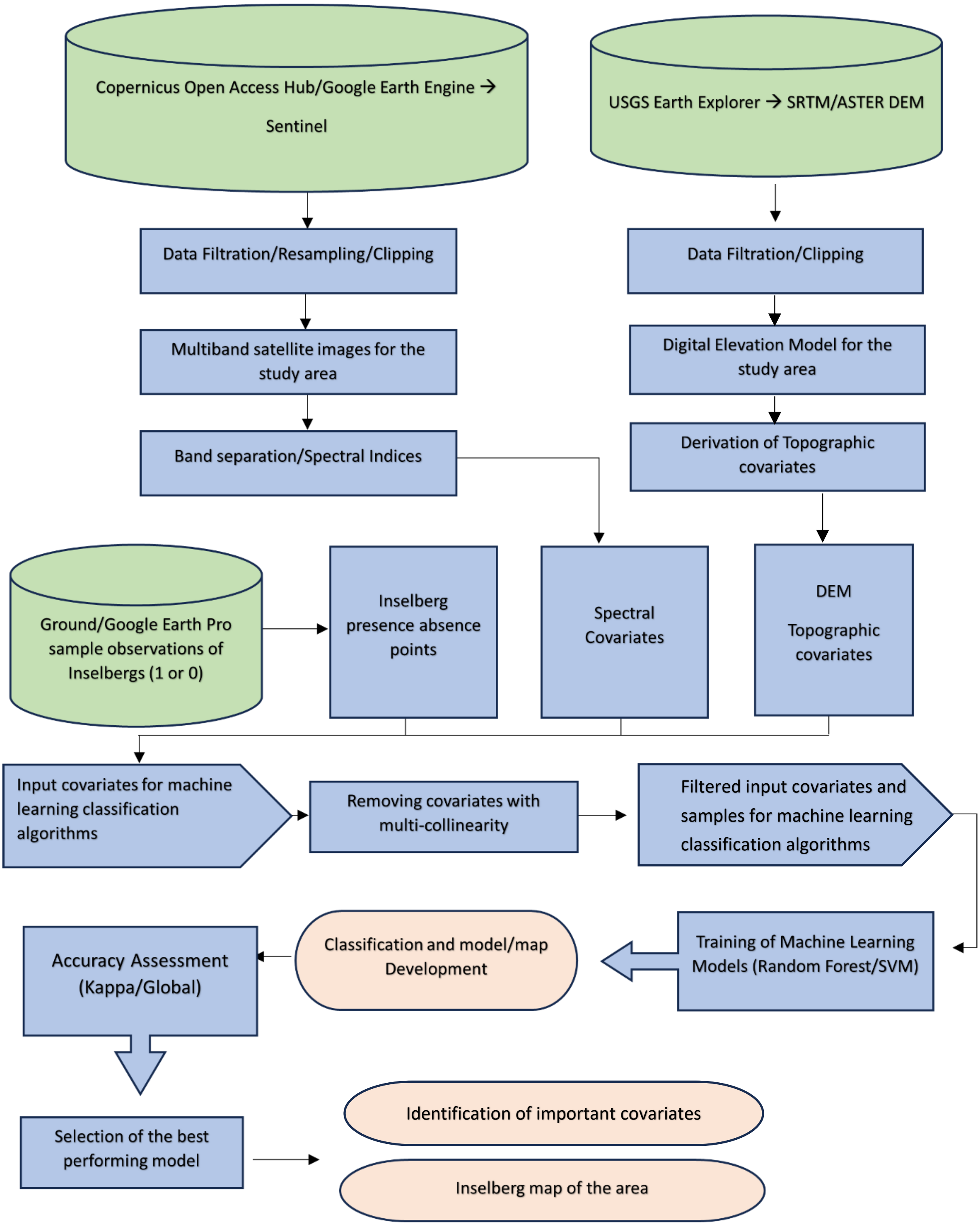


Figure 2: Methodological flowchart of data acquisition and processing for the generation of covariates used in modeling

Band separation and generation of spectral indices

The obtained multiband satellite image was utilized to generate separate spectral bands from B2-B12. Additionally, 8 spectral indices were generated using raster calculator tool in Arc GIS Pro. Table 1 summarizes specifications of 14 spectral covariates that were generated for modelling.

Table 1: Spectral covariates considered for modelling.

Spectral covariate	Description
NDVI	Normalized Difference Vegetation Index
SAVI	Soil Adjusted Vegetation Index
GNDVI	Green Normalized Difference Vegetation Index
MNDWI	Modified Normalized Difference Water Index
REDENDVI	Red Edge Normalized Difference Vegetation Index
CMR	Clay minerals ratio
IOR	Iron Oxide Ratio
FMR	Ferrous minerals ratio
Sentinel Band 2	Blue
Sentinel Band B3	Green
Sentinel Band B4	Red
Sentinel Band B8	NIR
Sentinel Band B11	SWIR-1
Sentinel Band B12	SWIR-2

Generation of topographic covariates

ALOS PALSAR Digital Elevation Model (DEM) at 12.5m spatial resolution was obtained from USGS Earth Explorer database. The DEM was clipped to the area of Gampaha district and resampled to achieve a spatial resolution similar to Sentinel images (10m). Using the resampled DEM, 25 topographic covariates in raster format (Table 2) were derived using the RSAGA (Brenning et al., 2023) package in R programming interface (R Core Team, 2024). All raster layers were having the same spatial extent and resolution of DEM.

Table 2: Topographic covariates considered for modelling.

Topographic covariate	Description
Aspect	Slope orientation
Convergence Index	Convergence Index
Cross Sectional Curvature	Transverse Curvature
Curvature Classification	Curvature Classification
DEM	Digital Elevation Model
Downslope Difference	Difference in vertical slope
Downslope Gradient	Difference in hydrological gradient
Generalized Surface	General curvature
Hillshade	Hillshade
Longitudinal Curvature	Longitudinal Curvature
Mass Balance Index	Balance index between erosion and deposition
Maximal Curvature	Maximum curvature
Minimal Curvature	Minimum curvature
Morphometric Features	Morphometric Features
MRRTF	Multiresolution Ridge Top Flatness
MRVBF	Multiresolution Valley bottom Flatness
Plan Curvature	Plan Curvature
Profile Curvature	Profile Curvature
Slope	Declivity
Surface Texture	Surface Texture
Terrain Ruggedness Index	Quantitative topography roughness index
Terrain Surface Classification	Terrain Surface Classification
Terrain Surface Convexity	Surface terrain convexity
TPI	Topographic Position Index - Vertical difference between base and summit of standardized slope
Vector Ruggedness Measure	Variation in slope terrain roughness

Ground sample collection

Based on different classes of land use and land cover, point observations were obtained in the Google Earth Pro computational application. At the end of this step, points identified as Inselberg habitats were assigned the value "1" (da Silveira et al., 2022). The remaining points referred to the other classes of land use and cover, namely: paddy, plantations, built-up areas, home gardens and water bodies, were assigned the value "0". Gound observations were conducted where necessary to confirm the ground truth conditions.

Covariate selection for modelling

Covariate selection process was carried out in three steps: (1) removal of covariates with variance close to zero; (2) removal by correlation; and (3) removal by importance (da Silveira et al., 2022). This step ensured that statistically insignificant covariates, multicollinearity and covariates without contextual relevance were removed prior to modelling to facilitate the development of a more accurate and meaningful model.

Machine learning algorithms, training and model development

Two machine learning algorithms; Random Forest – RF (Breiman, 2001) and Support Vector Machines (Cortes & Vapnik, 1995) were utilized for modelling in R programming platform (R core team, 2024). Several R packages including *"caret", "randomForest", "raster", "sf", "e1071",* and *"iml"* were used during this process. Training was performed using the ideal subset in the importance selection phase (RFE) (da Silveira et al., 2022). The performance of the algorithms was evaluated by Kappa and Global Accuracy indexes (Di Shi & Yang, 2016) derived from the confusion matrix. According with kappa (k), detection ranges were identified as poor ($k < 0$) to high ($0.7 < k \leq 1.0$) agreement. The global accuracy index indicates the probability that studied and classified classes correspond to true data. It also presents values ranging from zero (0) to one (1), according to aforementioned k values.

Model selection and map generation

Best performing model was selected according to the Kappa and Global Accuracy indices. Based on the best performing model, Inselberg map for the Gampaha district was generated and covariates that were significantly contributing to the model were further investigated to identify the relationships. For this purpose, the variable importance was obtained based on Mean Decrease Gini Index using the R package "*iml*". Permutation importance percentage was obtained for the best performing model to further investigate the contribution of different covariates on the final model.

Results and Discussion

The filtering process of covariates by correlation resulted in a total of 36 covariates available for modeling while four covariates were excluded from the analysis due to their high collinearity. Therefore, the number of topographic and spectral covariates used for initial model development was 35.

Both RF and SVM models were relatively accurate at predicting the inselberg habitats based on the training provided. However, SVM model scored higher in both Kappa and Global accuracy scores at 70.1% and 94.5% respectively. The accuracy scores of RF model were 59.1% and 92.7% respectively. The sensitivity of the two models marginally differed having 98.8% for SVM and 98,2% for RF. These values indicate the overall greater predictive power of SVM model when compared to RF. This accuracy is visually depicted in Figure 3 comparing the spatial distribution of Inselberg habitat with ground references.

The contribution of selected topographic and spectral covariates the prediction of Inselbergs varied according to each machine learning algorithm. Recursive Feature Elimination (RFE) eliminated further six covariates in SVM model to result in 28 covariates in the final model.

The most contributing covariates based on the Mean Decrease Gini Index of Random Forest model were DEM (24.8%), Ferrous Mineral Ratio (23.5%), Sentinel Band 8 (NIR) (23.5%), Soil Adjusted Vegetation Index (18.1%) and Terrain Ruggedness Index (17.7%) (Figure 4). DEM (17.3%), Ferrous Mineral Ratio (15.9%), Sentinel Band 8 (15.0%) were the top covariates of SVM model as well (Figure 5). Meanwhile, slope, Terrain Surface Convexity, Downslope Difference and Clay Minerals can be identified as important covariates that contributed to both models. After the selection of SVM as the best performing model, the per mutation importance analysis revealed that in addition to the above covariates, some important contributions from MINDWI and Minimal Curvature covariates for the model predictions (Figure 6).

The final Inselberg habitat map for Gampaha district was created based on SVM model and it was divided into a four-map series for better visualization (A1-Northwestern, A2-Northeastern, B1-Southwestern, B2-Southeastern regions) (Figure 7-10). When compared to the total extent of 1387 km^2 land area in Gampaha district, Inselberg habitats comprise of 3.4 km^2 area. In A1 which is a coastal region that holds cities such as Negombo, Katunayake and Anamaduwa, Inselberg habitats were completely absent. The only available Inselberg in this region was in Aluthepola temple area close to Minuwangoda city. There, was a similar situation in the southern coastal region of Gamapaha district (B1) where there was a very limited number of Inselberg habitats. The eastern regions of the district A2 and B2 were observed to harbor a higher number of Inselberg habitats with a greater distribution. In A2 region, areas like Warakapola, Ambepussa and Diwulapitiya host clusters of Inselberg habitats. In B2 region, Yakkala, Alawala and Dompe areas included clusters of Inselberg habitats. Interestingly, most of the Inselberg habitats were associated with Buddhist monasteries where a considerable level of protection was observed. However, evidences of quarrying were observed in other Inselbergs indicating the increasing anthropogenic destruction of these vital habitats.

The findings of the present study conform the applicability of AI-based machine learning algorithms integrating with remote sensing data to identify Inselberg habitats as described by da Silveira et al. (2022). The pioneering work of (Katupotha & Gamage, 2023) are evidenced by this work with a thorough geospatial approach in Gamapaha district and the models developed can be applied to investigate the remaining Inselbergs in Sri Lanka as well.

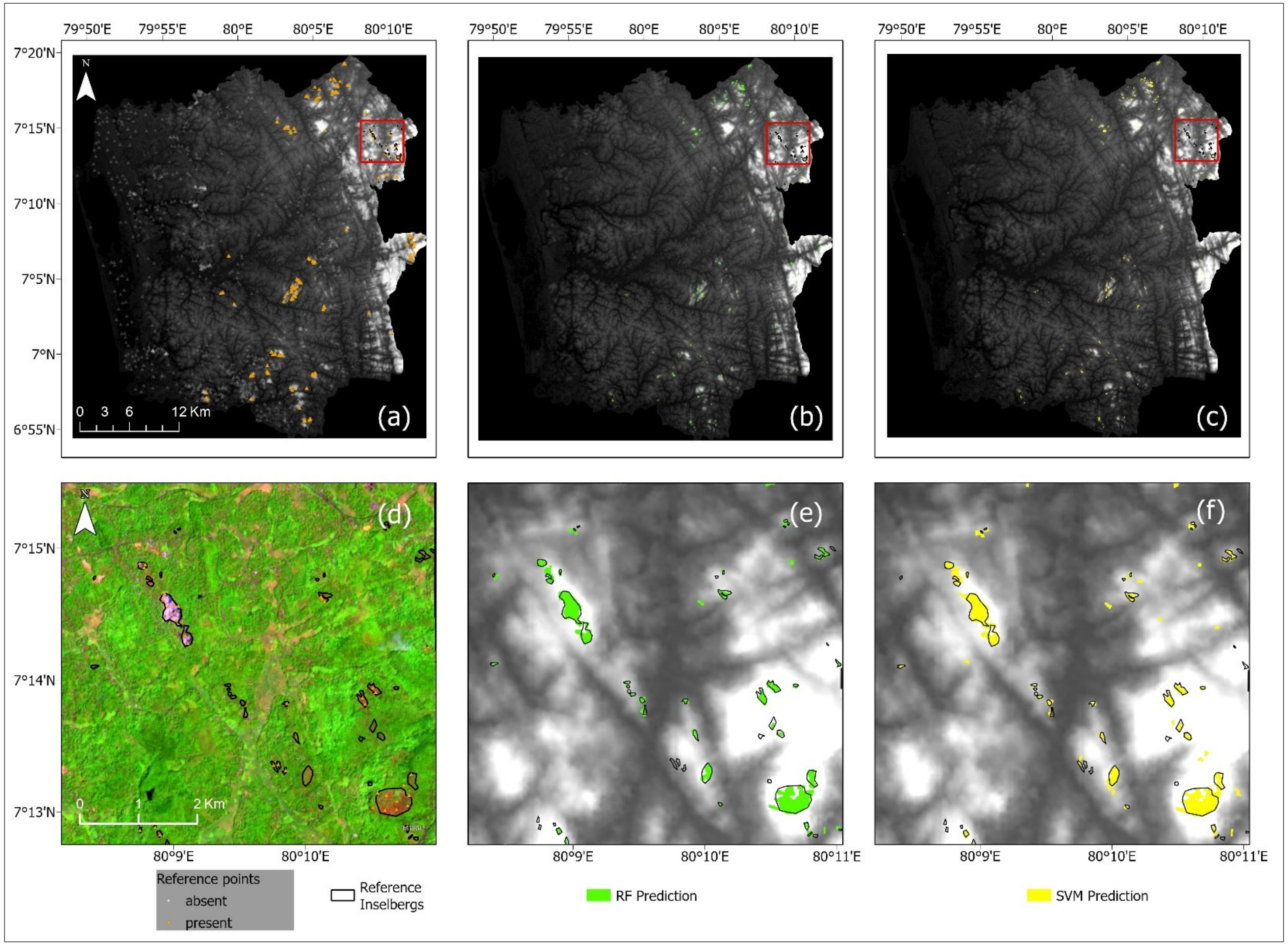


Figure 3: Inselberg mapping for the Gampaha district (a) generated by the two classifiers RF (b), and SVM(c). (d), (e) and (f) are the zoomed in regions of the three maps. (orange triangles-Reference Inselberg points, green-RF model predictions, yellow-SVM model predictions)

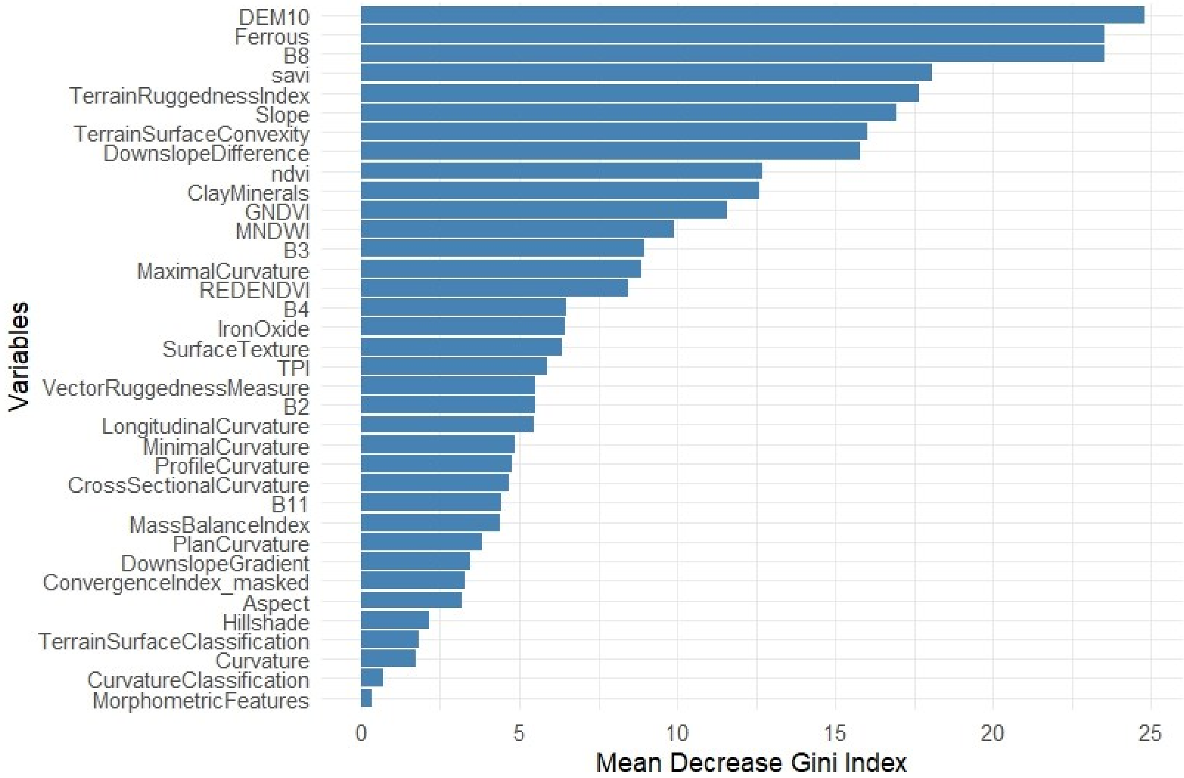


Figure 4: Covariate importance scores in the RF model prediction based on Gini Index

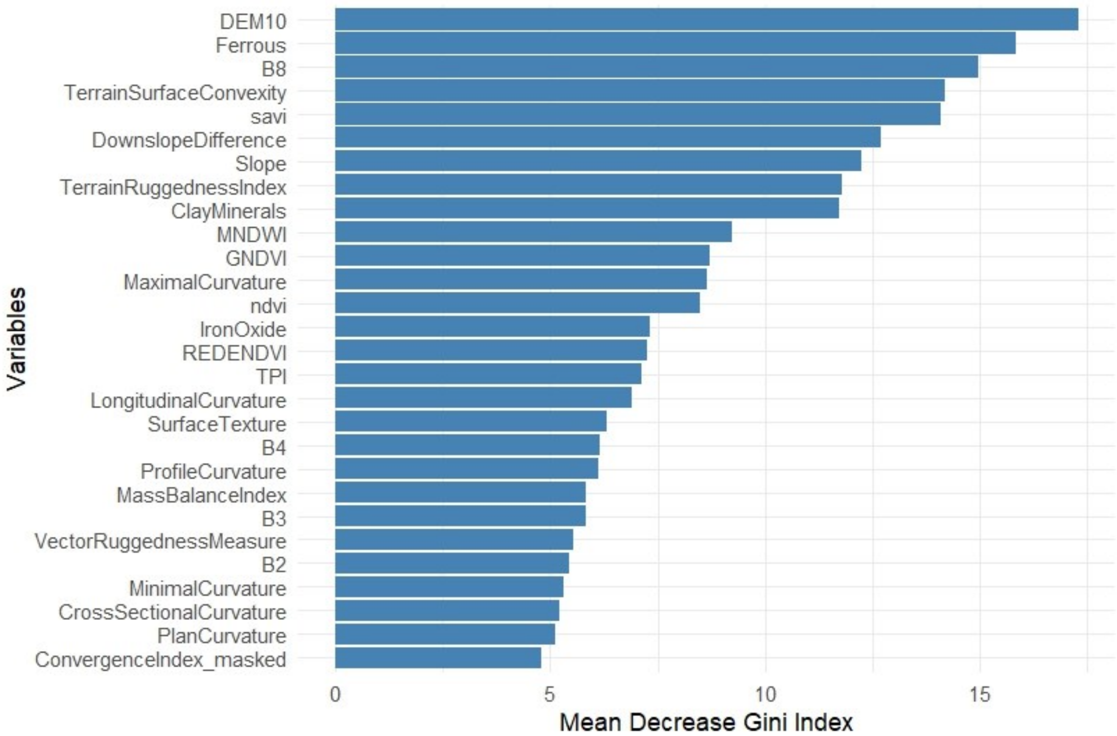


Figure 5: Covariate importance scores in the SVM model prediction based on Gini Index

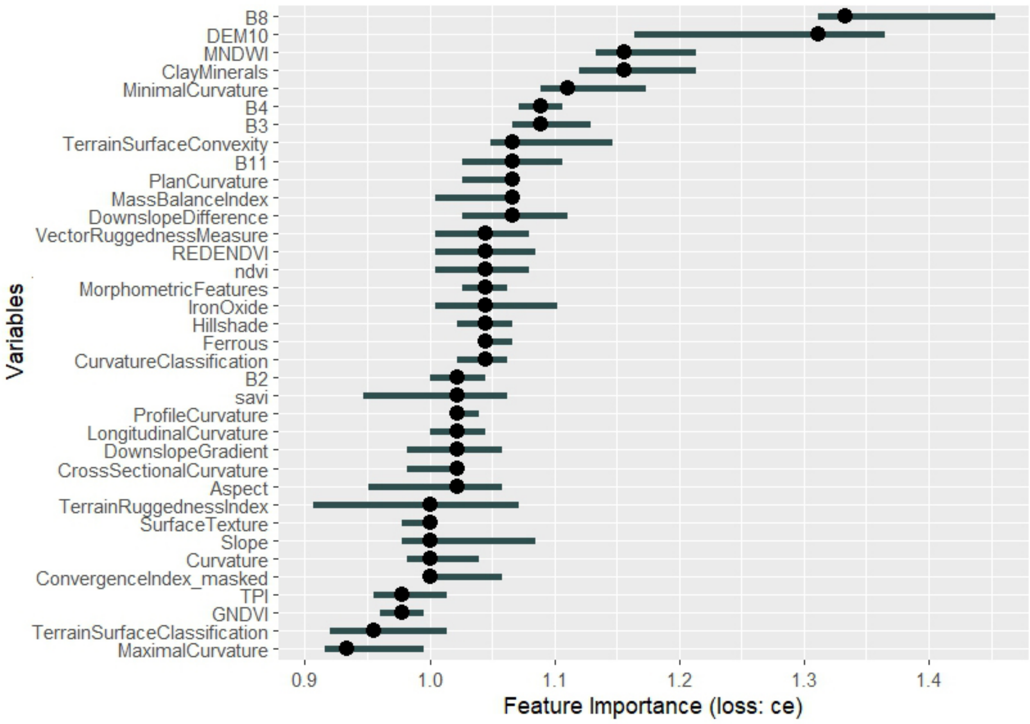


Figure 6: Permutation importance of covariates in SVM model

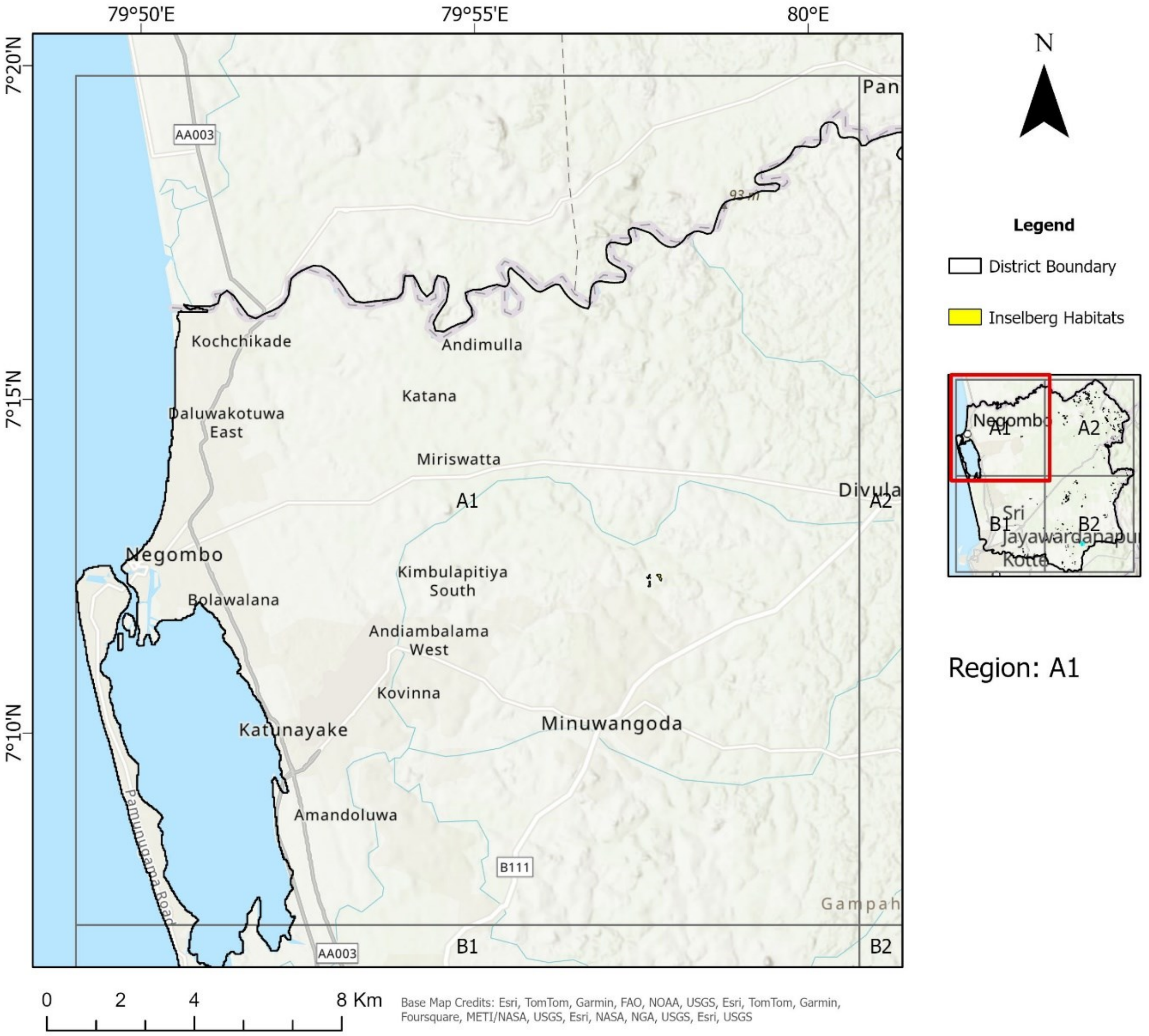


Figure 7: Inselberg map of Gampaha district (Region A1)

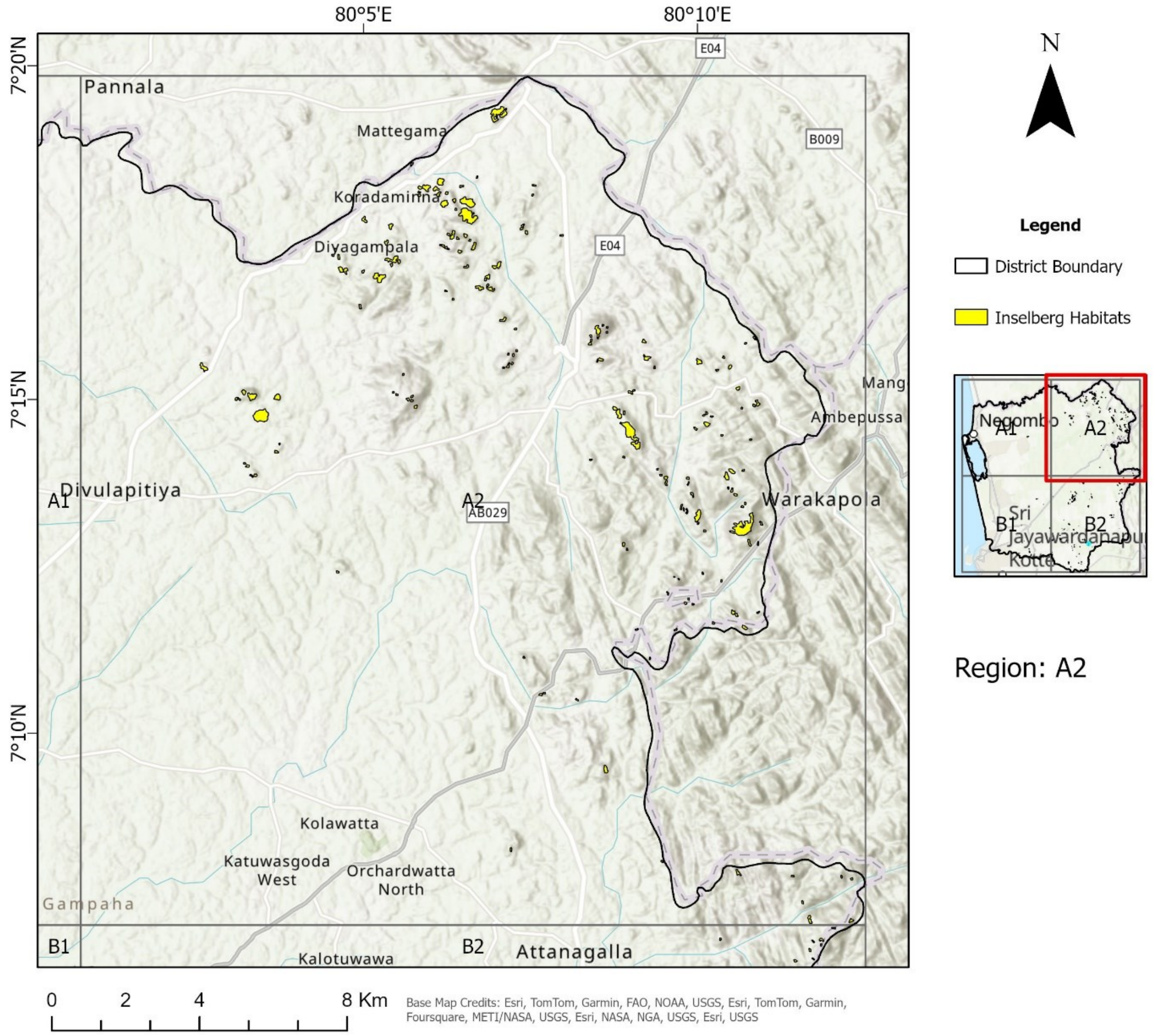


Figure 8: Inselberg map of Gampaha district (Region A2)

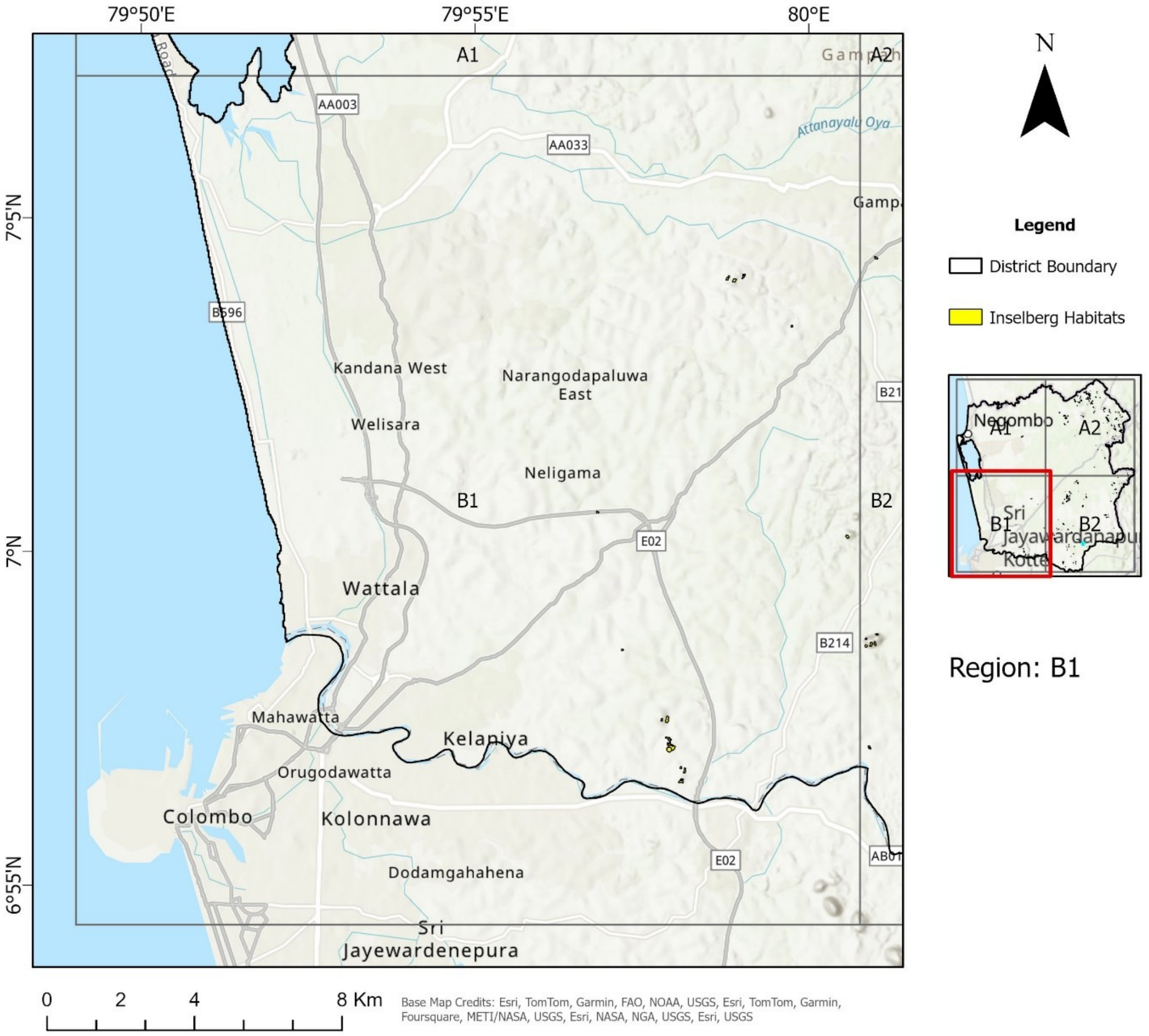


Figure 9: Inselberg map of Gampaha district (Region B1)

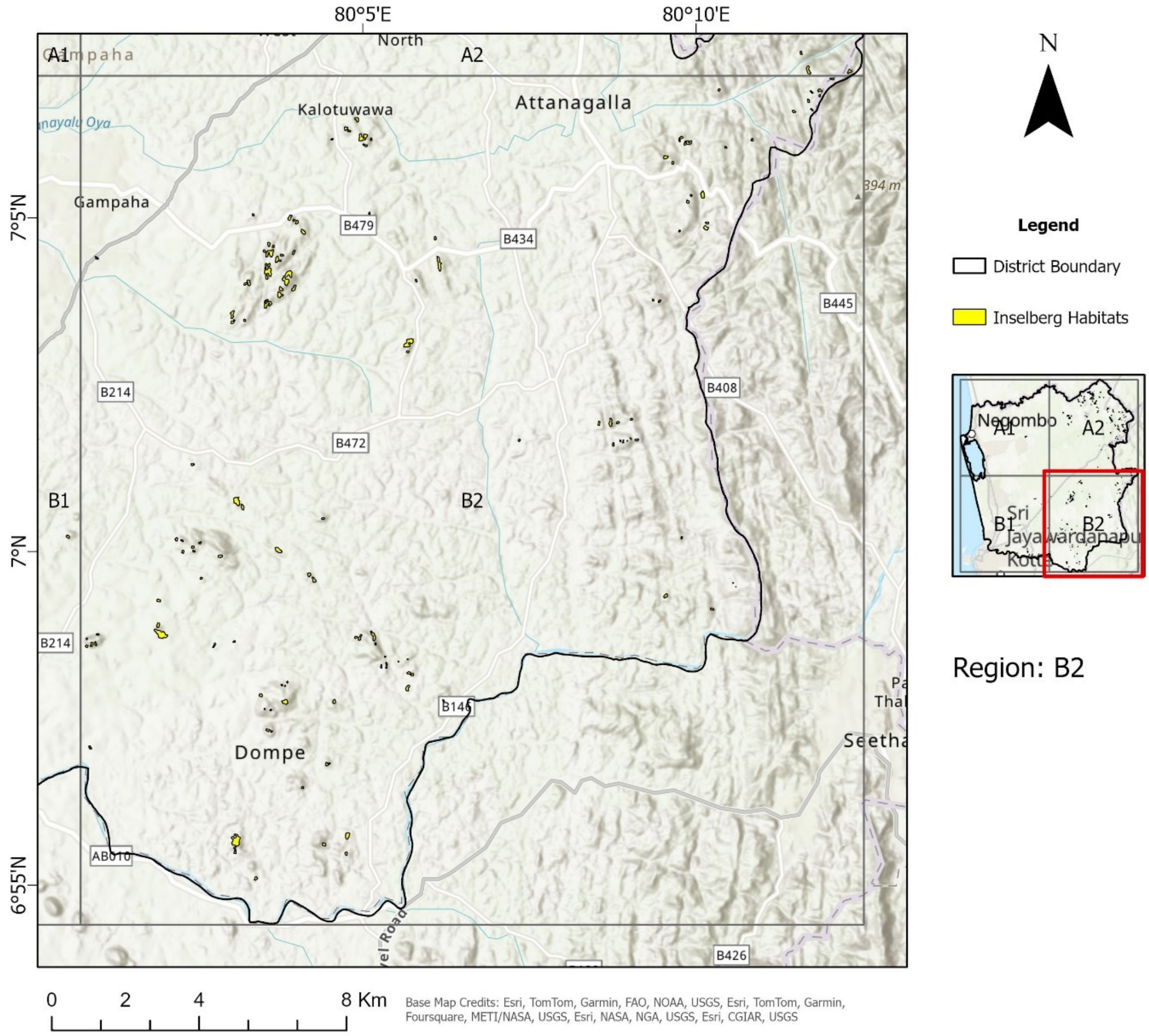


Figure 10: Inselberg map of Gampaha district (Region B2)

Conclusion and Recommendation

AI-based machine learning algorithms can be effectively utilized to analyze remotely sensed data for the identification of Inselberg habitats. SVM algorithm based model was the most accurate for the identification and predictive modelling of Inselbergs in the study area. Covariates such as DEM, ferrous minerals ratio, Band 8 (NIR) in Sentinel 2, terrain surface convexity, terrain surface ruggedness and slope were identified as most important factors that influence the occurrence of Inselbergs. Inselberg habitats in Gamapaha district were accurately identified by the SVM model and the total area of Inselbergs in the district was relatively small at 3.4 km^2.

The findings of this study can be effectively used to manage and conserve the remaining Inselberg habitats of the district which are of significant ecological, geological, and cultural values. It is recommended to increase the protection status of the last remaining Inselberg habitats in order to minimize the harmful anthropogenic activities such as quarrying, deforestation, and unregulated developments. Furthermore, this study can be used as an appropriate AI based model coupled with GIS and Remote Sensing to map inselbergs in any region of the country.

References

Barthlott, W. & Porembski, S. (2000). Why Study Inselbergs? In S. Porembski & W. Barthlott (Eds.), *Inselbergs*. Springer.

Bauer, A. M. & Das, I. (2000). A review of the gekkonid genus Calodactylodes (Reptilia: Squamata) from India and Sri Lanka. In *J. South Asian Nat. Hist* (Vol. 5, Issue 1).

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brenning, A., Bangs, D. & Becker, M. (2023). *RSAGA: SAGA Geoprocessing and Terrain Analysis. R package version 1.4.1, https://github.com/r-spatial/rsaga.* (R package version 1.4.1 https://github.com/r-spatial/rsaga).

Burke, A. (2003). Inselbergs in a changing world - Global trends. *Diversity and Distributions*, *9*(5), 375–383. https://doi.org/10.1046/j.1472-4642.2003.00035.x

Bussell, J. D. & James, S. H. (1997). Rocks as museums of evolutionary processes. In *Journal of the Royal Society of Western Australia* (Vol. 80, Issue 3).

Chandrajith, R. (2020). *Geology and Geomorphology* (pp. 23–34). https://doi.org/10.1007/978-3-030-44144-9_3

Coor, R., Warren, R. & Goudie, A. (1993). *Desert morphology*. UCL Press.

Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

da Silveira, V. A., Veloso, G. V., de Paula, H. B., dos Santos, A. R., Schaefer, C. E. G. R., Fernandes-Filho, E. I. & Francelino, M. R. (2022). Modeling and mapping of Inselberg habitats for environmental conservation in the Atlantic Forest and Caatinga domains, Brazil. *Environmental Advances*, *8*, 100209. https://doi.org/10.1016/j.envadv.2022.100209

De Jayawardena, U. S. (2015). A study on the present states of different rocks of ancient monuments in Sri Lanka. In *Engineering Geology for Society and Territory - Volume 8: Preservation of Cultural Heritage* (pp. 43–46). Springer International Publishing. https://doi.org/10.1007/978-3-319-09408-3_4

Di Shi & Yang, X. (2016). An Assessment of Algorithmic Parameters Affecting Image Classification Accuracy by Random Forests. *Photogrammetric Engineering & Remote Sensing*, *82*(6), 407–417. https://doi.org/10.14358/PERS.82.6.407

Emadi, M., Taghizadeh-Mehrjardi, R., Cherati, A., Danesh, M., Mosavi, A. & Scholten, T. (2020). Predicting and Mapping of Soil Organic Carbon Using Machine Learning Algorithms in Northern Iran. *Remote Sensing*, *12*(14), 2234. https://doi.org/10.3390/rs12142234

Fernando, G. M. T. S. (2018). *Identification of Urban Heat Islands &Its Relationship withVegetation Cover: A Case Study of Colombo & Gampaha Districts in Sri Lanka*. http://www.nasa.gov/.

Gašparović, M. & Jogun, T. (2018). The effect of fusing Sentinel-2 bands on land-cover classification. *International Journal of Remote Sensing*, *39*(3), 822–841. https://doi.org/10.1080/01431161.2017.1392640

Ghimire, B., Rogan, J., Galiano, V. R., Panday, P. & Neeti, N. (2012). An Evaluation of Bagging, Boosting, and Random Forests for Land-Cover Classification in Cape Cod, Massachusetts, USA. *GIScience & Remote Sensing*, *49*(5), 623–643. https://doi.org/10.2747/1548-1603.49.5.623

Hechenbichler, K. & Schliep, K. P. (2004). Weighted k-nearest- neighbor techniques and ordinal classification. Discussion paper 399. In *Sonderforschungsbereich 386* (pp. 1–16). Ludwig-Maximilians University.

Hopper, S. D., Fiedler, P. L. & Yates, C. J. (2021). Inselberg floristics exemplify the coast to inland OCBIL transition in a global biodiversity hotspot. In *Biological Journal of the Linnean Society* (Vol. 133). https://academic.oup.com/biolinnean/article/133/2/624/6118897

Huang, F., Cao, Z., Guo, J., Jiang, S.-H., Li, S. & Guo, Z. (2020). Comparisons of heuristic, general statistical and machine learning models for landslide susceptibility prediction and mapping. *CATENA*, *191*, 104580. https://doi.org/10.1016/j.catena.2020.104580

Jutz, S. & Milagro-Pérez, M. P. (2018). Copernicus Program. In *Comprehensive Remote Sensing* (pp. 150–191). Elsevier. https://doi.org/10.1016/B978-0-12-409548-9.10317-3

Katupotha, J & Ravibhanu, A. (2020). *Geotourism and geoheritage potentials: vistas and unique opportunities of Sri Lanka* (Vol. 8, Issue 2).

Katupotha, J. (2015). *Diversity of Vegetation Types of the Pidurangala Granitic Inselberg with Ancient Forest Monastery, Near Sigiriya, Sri Lanka: A Preliminary Study*. https://www.researchgate.net/publication/280234472

Katupotha, J & Gamage, S. (2023). Inselbergs in Sri Lanka: Their ecology and biodiversity value. *WILDLANKA International Symposium – 2023*. https://doi.org/10.13140/RG.2.2.30080.61440

Katupotha, J & Kodituwakku, K. (2015). Utility of Inselbergs and Boulders as Rock-Shelter Monasteries: A Geoarchaeological Study of Sigiriya and Its Environs. In *Journal of Social Sciences and Humanities* (Vol. 1, Issue 4). http://www.aiscience.org/journal/jsshhttp://creativecommons.org/licenses/by-nc/4.0/

Kittle, A. M., Watson, A. C., Cushman, S. A. & Macdonald, D. W. (2018). Forest cover and level of protection influence the island-wide distribution of an apex carnivore and umbrella species, the Sri Lankan leopard (Panthera pardus kotiya). *Biodiversity and*

Conservation, *27*(1), 235–263. https://doi.org/https://doi.org/10.1007/s10531-017-1431-8

MacArthur, R. H. & Wilson, E. O. (1967). The Theory of Island Biogeography. In *Monographs in population biology*. Princeton University Press.

Mas, J. F. & Flores, J. J. (2008). The application of artificial neural networks to the analysis of remotely sensed data. *International Journal of Remote Sensing*, *29*(3), 617–663. https://doi.org/10.1080/01431160701352154

Maxwell, A. E., Warner, T. A. & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing*, *39*(9), 2784–2817. https://doi.org/10.1080/01431161.2018.1433343

Mountrakis, G., Im, J. & Ogole, C. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(3), 247–259. https://doi.org/10.1016/j.isprsjprs.2010.11.001

Nawarathna Banda, H. M. (2008). *Land fragmentation and environmental degradation in Sri Lanka (with reference to Gampaha district)*.

Porembski, S. (2000). *Biodiversity of Terrestrial Habitat Islands-the Inselberg Evidence*.

Porembski, Stefan & Barthlott, W. (2000). Granitic and gneissic outcrops (inselbergs) as centers of diversity for desiccation-tolerant vascular plants. In *Plant Ecology* (Vol. 151).

Premakantha, K. T., Chandani, R. P. D. S., Kingsly, S. A. D., Dias, H. U. & Kekulandara, N. S. B. (2021). Forest cover assessment in Sri Lanka using high resolution satellite images. In *The Sri Lanka Forester* (Vol. 40).

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M. & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, 93–104. https://doi.org/10.1016/j.isprsjprs.2011.11.002

Souza, C. M. P. de, Thomazini, A., Schaefer, C. E. G. R., Veloso, G. V., Moreira, G. M. & Fernandes Filho, E. I. (2018). Multivariate Analysis and Machine Learning in Properties of Ultisols (Argissolos) of Brazilian Amazon. *Revista Brasileira de Ciência Do Solo*, *42*(0). https://doi.org/10.1590/18069657rbcs20170419

Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F. & Fouilloy, A. (2017a). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, *105*, 569–582. https://doi.org/10.1016/j.renene.2016.12.095

Voyant, C., Notton, G., Kalogirou, S., Nivet, M.-L., Paoli, C., Motte, F. & Fouilloy, A. (2017b). Machine learning methods for solar radiation forecasting: A review. *Renewable Energy*, *105*, 569–582. https://doi.org/10.1016/j.renene.2016.12.095

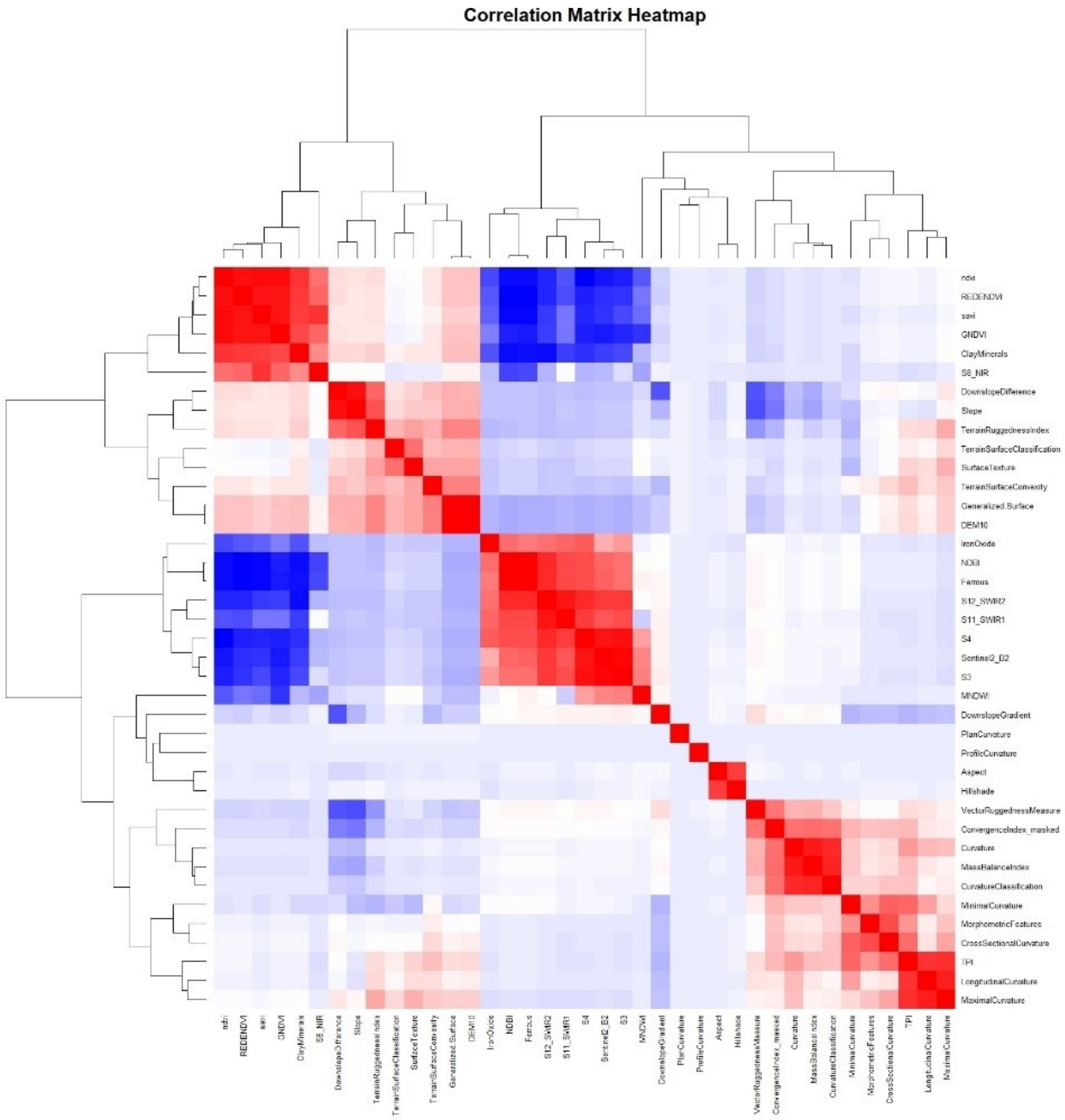
Weerasekara, L. S., Wijesooriya, K., Ranawana, K. & G Vandercone, R. P. (2021). Survey of relative abundance of diurnal primates in the Giritale Nature Reserve, Sri Lanka. *Asian Primates Journal*, *9*(1). https://www.researchgate.net/publication/350462740

Withanage, G. P., Viswakula, S. D., Nilmini Silva Gunawardena, Y. I. & Hapugoda, M. D. (2018). A forecasting model for dengue incidence in the District of Gampaha, Sri Lanka. *Parasites and Vectors*, *11*(1). https://doi.org/10.1186/s13071-018-2828-2

Yates, C. J., Robinson, T., Wardell-Johnson, G. W., Keppel, G., Hopper, S. D., Schut, A. G. T. & Byrne, M. (2019). High species diversity and turnover in granite inselberg floras highlight the need for a conservation strategy protecting many outcrops. *Ecology and Evolution*, *9*(13), 7660–7675. https://doi.org/10.1002/ece3.5318

Yu, L., Liang, L., Wang, J., Zhao, Y., Cheng, Q., Hu, L., Liu, S., Yu, L., Wang, X., Zhu, P., Li, X., Xu, Y., Li, C., Fu, W., Li, X., Li, W., Liu, C., Cong, N., Zhang, H., … Gong, P. (2014). Meta-discoveries from a synthesis of satellite-based land-cover mapping research. *International Journal of Remote Sensing*, *35*(13), 4573–4588. https://doi.org/10.1080/01431161.2014.930206

Zeraatpisheh, M., Ayoubi, S., Jafari, A. & Finke, P. (2017). Comparing the efficiency of digital and conventional soil mapping to predict soil types in a semi-arid region in Iran. *Geomorphology*, *285*, 186–204. https://doi.org/10.1016/j.geomorph.2017.02.015



Supplementary Figure 1: Correlation matrix of raster covariates