Model of Peatland Vegetation Species using HyMap Image and Machine Learning

Muhammad Dayuf Jusuf*, Projo Danoedoro², Bangun Muljo Sukoko³, Hartono⁴

¹ PhD student remote sensing, Fakultas Geografi – Universitas Gadjah Mada, Yogyakarta
², ³ Fakultas Geografi – Universitas Gadjah Mada, Yogyakarta
³ Fakultas Teknik Geomatika, Institut Teknologi Sepuluh Nopember – ITS, Surabaya

Abstract. Species Tumih / Parepat (Combretocarpus-rotundatus Mig. Dancer) family Anisophylleaceae and Meranti (Shorea Belangerang, Shorea Teysmanniana Dyer ex Brandis) family Dipterocarpaceae is a group of vegetation species distribution model. Species pioneer is predicted as an indicator of the succession of ecosystem restoration of tropical peatland characteristics and extremely fragile (unique) in the endemic hot spot of Sundaland. Climate change projections and conservation planning are hot topics of current discussion, analysis of alternative approaches and the development of combinations of species projection modelling algorithms through geospatial information systems technology. Approach model to find out the research problem of vegetation level based on the machine learning hybrid method, wavelet and artificial neural networks. Field data are used as a reference collection of natural resource field sample objects and biodiversity assessment. The testing and training ANN data set iterations times 28, achieve a performance value of 0.0867 MSE value is smaller than the ANN training data, above 50%, and spectral accuracy 82.1 %. Identify the location of the sample point position of the Tumih / Parepat vegetation species using HyMap Image is good enough, at least the modelling, design of the species distribution can reach the target in this study. The computation validation rate above 90% proves the calculation can be considered.

Keywords: Peatland species, HyMap image, WANN hybrid algorithm model.

1. Introduction

Some methods of uncertainty around the effectiveness of biodiversity make the researchers and resource managers ambiguous, as well as conservation planners. Summarized the model distribution method of species distribution models (SDMs), niche concepts of environmental habitat suitability and climate prediction distribution, spatial species attribute data of scale, space, and time. In the last five years there have been many published papers on various aspects of modelling methods of species distribution that enable the achievement of various advances. Theoretical modelling of species, ecological data provides comprehensive guidance for the production of statistical models, machine learning and environments for the classification of spatial distribution species [3].

The intention of modelling is usually done based on the interpolation and spatial distribution of geospatial species of non-linear geographically broad. Such modelling practices are widespread and supported by particular modelling contexts in various comparative results. The species distribution
model is an empirical model related to field observations for predictors of world-level environmental response variables, statistically or theoretically based. Vegetation species data may be present or not and appear to be abundant or not, depending on the results of field observation, done at random or stratified sampling for ease of observation of the object in nature, the layout can be shown in Figure 1.

The actual species distribution maps or habitat potentials are required in many facets of environmental inquiry, such as: resource management, and preservation planners. It includes biodiversity assessment, biological availability design, habitat management and restoration, species preservation, habitat design, population survival analysis, environmental hazard assessment, invasive species management, community ecosystem modelling, and predicted impacts of global environmental ecosystem changes in species [1].

The dispersion model of species results and benefits are relatively inadequate in the appropriate resource analysis. Integration of Geographic Information System (GIS) model mapping is not even complete enough to pump for the standardization of the organization model of the environment variable and produce prediction map of species, predictive evaluation method of errors and doubt, can

---

**Figure 1. Research Flowchart**

The spectral recording of the HyMap Image type has a narrow wavelength (<10 nm) with a fairly wide wavelength range. HyMap bands channel range, sensitive recording of the unique species of vegetation. The unique nature of Indonesia's unique tropical peatland vegetation bands is not encountered in other ecosystems, the use of Hyperspectral imagery is very appropriate, recording the features of this unique spectrum.
inhibit effective modelling of collecting key components of problem approach apply all over the world [7] [3].

Figure 2. Research study area

The subject area of this research was covered area surrounding HEF (hampengan ecology, forest) disturb forest in Katingan district - Central Kalimantan, Indonesia is shown in figure 2. This subject area selected due to various objects types and vegetation density. The use of Hyperspectral (HyMap) imagery with high spatial resolution expected to be capable to distinguish the various objects types and vegetation in the study area accurately. The Hyperspectral imagery used was (HyMap) imagery obtained from an aerial scan survey. The imagery was recorded in 15-16 July\textsuperscript{th} 2011 and it was already corrected automatically since proceed it on air. Thus, the Hyperspectral (HyMap) imagery had already 4.2 m spatial resolution in a multi spectral band in 4 groups [VINIR, NIR (VISNIR & NIR), SWIR (NIR) and SWIR 1 _2]. Detailed information about Hyperspectral (HyMap) imagery shown in Table 1.
Table 1. Detail information about the Hyperspectral (HyMap) imagery used in this research.

| Channel Type | Bands | Channel Length |
|--------------|-------|---------------|
| VNIR         | VISNIR| 425.50000 - 714.50000 |
| NIR          | VISNIR| 729.70000 - 895.60000 |
|              | NIR   | 868.30000 - 1119.40000 |
| SWIR         | SWIR-1| 1134.20000 - 1333.19995 |
|              | SWIR-2| 1959.00000 - 2490.00000 |

2. Methods
Machine learning, is a discipline of science that develops algorithms or models that can memorize from the data from the knowledge, information that is on [5]. Briefly, learning machine can be interpreted as the procedure of changing data into information. Learning machine is divided into Supervised Learning and Unsupervised Learning.

Simulation of Wavelet transformation and Artificial Neural Network (WANN) is performed by using MatLab software. These hybrid transformation programming processes, takes four main steps to receive the test solutions and simulation practice. The Sequence programming includes: 1) design features files dataset species, vegetation, 2) extraction sample feature dataset in DWT decomposition + power average calculation, 3) constructs programming species pioneer identification Tumih and several pioneer species peatlands, 4) Training and testing dataset in ANN. The results indicate an increase in accuracy and speed in recognition of high level error and opposition.

Figure 3. Result execution of ANN training dataset in these experiments, achievement spectral accuracy (82.1%) for 3 class species pioneers (Tumih, Meranti and others).
The basic simulation design as input parameters 11 dataset and input 15 works on the Artificial Neural Network (ANN) layer of Levenberg–Marquardt algorithm (LMA or just LM), used to solve nonlinear least squares problems in mathematics and computing. These minimization problems arise, especially in least squares curve fitting [4]. Experimental results obtained vary widely in pigment concentration (per unit leaf area and unit area) and leaf area index, based on a plant canopy scale, and independent variation between chlorophyll and carotenoid leaf area index. Survey data were used to evaluate the nature and strength of the relationship between the amount of spectral transformation and pigment concentration (per unit area land) [8].

The hyperspectral narrowband reflectance in the region's electromagnetic spectrum appears to be related to chlorophyll concentrations, while the wavelength ratio index works near-infrared (NIR), especially the more successful green wavelength. While the red edge position (REP) of the red edge is connected to the chlorophyll concentration, the foremost and second amplitude characteristics of the reflectance and pseudo absorption are stronger correlated with chlorophyll. The research approaches the information, analysis method of Hymap image integration and study data using algorithm design:

![Figure 4. Development methods WANN for peatland species pioneer.](image)

No spectral reflectance transformation was tested for Carotenoid concentrations, but the two index ratios were strongly associated with carotenoid ratios for chlorophyll a. Leaf area index (LAI) is not related to all spectral ratio indices, including broadband normalization difference and simple vegetation ratio index. This indicates that the spectral index may take in a limited application to estimate the LAI of vegetation canopy when the leaves comprising chlorophyll concentrations (per unit of leaf area) of the canopy have been different [5].

The multi sensor data application of the remote sensing field develops combine satellite data sources, opening up a new field research. Basic investigation of satellite imagery, related to the land cover classification factor. Optical data in the collection of high-resolution radiometer sensors for thematic mapping, in the specific classification capabilities of specific spectral signatures, some cover varies by scale and type of information. Traditional image processing resampling leads to significant differences in the quality of information extracts. The role of the spectral signature scale helps to secure the correct interpretation of the classification results [8].

Hyperspectral data analysis provides detailed spectral information of electromagnetic spectrum band subjects. The content of hyperspectral image information is extremely correlated to the neighbouring bands beyond and riding the size of its large dimension. Hyperspectral endmember validation analysis data are simulated in ENVI programming. The information of spectral bands of
overlapping hyperspectral data should be extracted by the characteristics of the estimate of the exact properties of vegetation and soil [2].

Whereas a strong relationship exists between the image’s spectral values and the object, the image can be used for further modelling. The undisturbed environment is not homogeneous, in contrast to highly heterogeneous-related topography. The slope of the slope and the direction of the slope affect the magnitude of the energy of the reflected object, and recording the result. The model relationship between spectral values and vegetation species should be corrected for topographic effects [5].

3. Results and Discussion
To obtain the testing and training ANN data set, shown are: several components: epoch iterations times 28, performance 0.067 means that a good enough because above 50 %, and spectral accuracy 82.1 %. However, the calculations show that the MSE (Minimum Square Error) value is smaller than the ANN training data that can achieve a performance value of 0.0837 and MSE = 0.0547. Identify the location of the sample point position of the Tumih / Parepat vegetation species using HyMap Image is good enough, at least the modelling, design of the species distribution can reach the target in this study. The computation validation rate above 90% proves the calculation can be considered. The aim of the development programming algorithm of peatland vegetation species can be utilized for other optical images such as: Multispectral, Hyperspectral, and Drone image.

4. Conclusions
Hyperspectral remote sensing image, the recorded spectrum has a narrow wavelength (<10 NM), but the overall wavelength range is broad enough. In this example, a sensor that receives only 20 bands, only includes a wavelength of 500 ~ 700 NM, can be categorized included in hyperspectral. The narrow reach of each channel of the hyperspectral channel is really sensitive to recording for unique objects. This is called for a spectral library field of peatland vegetation species.

The channels in the hyperspectral image data in this survey should consist of 10 image strips, but used only 6 (six) image strips where the availability of successful field samples in the collection: HU1_01, HU1_02, HU1_03, HU1_04, HU1_05 and HU1_06.

Acknowledgement
I would like to thank are conveyed on behalf of deceased our brother Mr. Muhammad Evri (Agency for the Assessment and Application of Technology, BPPT) who delivered previous talks and advised me to this research and Mr. Hendrik Segah from the Faculty of Forestry, University of Palangkaraya (UNPAR) - Central Kalimantan, which has been willing to provide a means of handheld spectrometer measurements in the field. Which supply me of hyperspectral data (HyMap) images from cooperation projects between BPPT- UNPAR, Indonesia and JST-JICA Japan.

References
[1] Burhanuddin. 2011. “The Association of Arbuscular Mycorrhizal Mushrooms with Perepat (Combretocarpus Rotundatus Miq) and Jelutung (Dyera Lowii Hook) In Peatlands.” Post-Graduate Program - Gadjah Mada University Yogyakarta, Indonesia.
[2] Harris, A., R. Charnock, and R.M. Lucas. 2015. “Hyperspectral Remote Sensing of Peatland Floristic Gradients.” Remote Sensing of Environment 162. Elsevier Inc.: 99–111. Doi:10.1016/j.rse.2015.01.029.
[3] Janet Franklin, Jennifer A. Miller. 2010. Mapping Species Distributions: Spatial Inference and Prediction. Edited by, and Jeremy Wilson, Paul Adam, H. J. B. Birks, Lena Gustafsson, Jeff McNeely, R. T. Paine, David Richardson, Birds. Cambridge CB2 8RU, UK: Cambridge University Press. www.cambridge.org/9780521876353.
[4] Kalacska, M., G. a. Sanchez-Azofeifa, B. Rivard, T. Caelli, H. Peter White, and J. C. Calvo-Alvarado. 2007. “Ecological Fingerprinting of Ecosystem Succession: Estimating Secondary Tropical Dry Forest Structure and Diversity Using Imaging Spectroscopy.” Remote Sensing
of Environment 108 (1): 82–96. Doi:10.1016/j.rse.2006.11.007.
[5] Kalacska, M., M. Lalonde, and T.R. Moore. 2015. “Estimation of Foliar Chlorophyll and Nitrogen Content in an Ombrotrophic Bog from Hyperspectral Data: Scaling from Leaf to Image.” Remote Sensing of Environment 169. Elsevier Inc.: 270–79. doi:10.1016/j.rse.2015.08.012.
[6] Murfi, H. 2017. Model Combination. Available at: http://ocw.ui.ac.id/course/view.php?id=21 [Access date: 14 Agustus. 2017].
[7] Raes, L.G. Saw, P.C. Van Welzen, T. Yahara. 2013. “Legume Diversity as Indicator for Botanical Diversity on Sundaland, South East Asia.” South African Journal of Botany 89: 265–72.
[8] Heikkinen, J, and R Makipaa. 2010. “Testing Hypotheses on Shape and Distribution of Ecological Response Curves.” Ecological Modelling 221 (3): 388–99. Doi: 10.1016/j.ecolmodel.2009.10.030.