Suitability model using support vector machine for land use planning scenarios in Java Island, Indonesia

Safitri, S., Sumarto, I., Riqqi, A., Deliar, A., Norvyani, D A, Taradini, J.

Department of Geodesy and Geomatics Engineering, Institute Technology of Bandung
10 Ganesha Street, 40132 Bandung, Indonesia

e-mail: sitaranisafitri@gmail.com, irawansumarto@gmail.com, akhmad.riqqi@gmail.com, albertus.deliar@gmail.com, dinorv19@gmail.com, jesikaturadini@gmail.com

Abstract. Java Island has experienced rapid population growth in the past four decades. Along with population growth, human needs are also increasing. This is also driven by economic and social growth. However, land resources to meet population needs are becoming increasingly limited. Human needs change ecosystems by creating ecological pressures, such as land cover change, resource extraction and depletion, and emissions pollution. Thus, sustainable land use planning is needed to meet human needs in the future. Furthermore, it is necessary to consider the suitability of the land in land use planning. In this paper, a suitability evaluation method was developed, which synthetically considered the topographic, meteorological, ecoregion, and water supply conditions. The suitability evaluation model is developed using Support Vector Machine (SVM) to classify various designations. SVM training algorithm aims to find a hyperplane that separates the dataset into a discrete predefined number of classes in a fashion consistent with the training examples. The term optimal hyperplane is used to refer to the decision boundary that minimizes misclassifications, obtained in the training step. Model calibration and validation were performed based on the land-use status in 2016. Subsequently, the validated simulations were conducted based on the planning scenarios. This model is expected to be implemented in the short-term and long-term land planning scenarios, respectively. This study provides a synthetic suitability evaluation method for creating a land-use planning scenario, which overcomes the shortcoming of the traditional way of assigning land-use scenarios that being lack of objectivity.

1. Introduction

The land can produce products in the form of goods and services (supply) to meet human needs (demand). The intended needs include food, water, air, homeostasis, rest, and excretion as the most basic human physiological needs that must be met based on Maslow's Hierarchy of Needs [1]–[4]. Land products include food, fiber, and environmental services such as water supply, and airflow management. The ability of the land to produce such products depends on the ecological quality [5], [6]. Environmental quality is affected by environmental stress (environmental distress syndrome) due to changes in ecosystems to meet human needs, such as changes in land cover, resource extraction and depletion (such as deforestation and overfishing), disposal and pollution emissions and modification and movement of organisms [7], [8]. The resulting environmental impacts include, but are not limited to, climate change, land degradation, loss of biodiversity, and environmental pollution [9]–[11]. These environmental problems have led to reduced land productivity, and this has become a global issue in recent decades [4], [12]. Therefore, sustainable planning is needed that takes into account not only social and economic needs but also environmental sustainability.
Land suitability is an essential component in development activities to meet human needs. Meeting human needs through development activities requires land allocation. Land-use allocation is the determination of the amount of land for a particular use (or not used) through legal and administrative steps, which will ultimately lead to the implementation of planning [13]. Thus, land use allocation determines economic development performance and environmental quality [14]. Therefore, the allocation of land becomes a critical stage in development planning.

The determination of land use allocation can be done using the concept of spatial pattern recognition and land suitability. The concept of land suitability aims at optimal and sustainable land use [15]. This can be achieved by providing information on the relationship between land characteristics. Land characteristics are land attributes that can be measured or estimated, such as slope, rainfall, soil texture, and vegetation [15], [16]. There are several land suitability assessments or land evaluation assessments built based on the classification framework of the United States Department of Agriculture (USDA) (1961) and the Food and Agriculture Organization (FAO) (1976) [15], [17], [18]. The models developed are generally based on location (spatial) and based on deterministic or probabilistic approaches. The land characteristics of the assessment system are assumed to be used in this study to determine the direction of land-use allocation locations.

2. Materials and Methods

2.1. Case Study (Java Island, Indonesia)

The aim of this study is to analyze the land-use suitability using Support Vector Machine (SVM), with Java Island as a case study. The land and sea area of Java Island provides a variety of natural resources, both biotic and abiotic, which are of high economic and ecological value. The population growth rate in Java in 2018 reaches 1.23% with a population density of 1,317 people / km², and 56.7% of the total population lives in urban areas [19]. The population of Java Island is almost the same as the total population of Kalimantan, Sulawesi, Maluku, Papua, and NTT. Energy consumption in Indonesia is focused on Java, or more than 60% of the total national consumption because 57% of consumers are in Java [20]. The volume of energy consumption per capita reached 883.91 kg of oil equivalent in 2014 and increased by 5.9% in 2016 [21]. Indonesia's CO2 emissions increased from 0.41 metric tons per capita in 1975 to 1.8 metric tons per capita in 2014 [22], [23]. Energy consumption and CO2 emissions have been on an upward trend in recent years. Empirical evidence shows that energy consumption increases carbon emissions, and economic growth is a significant contributor [24].

![Map of Java Island](image-url)
2.2. Suitability Modelling using SVM

Ecology and environmental changes can be ascertained by land cover and land use as the ground assumption [25]. The land cover portrays heterogenous Earth surface’s physical properties, for instance, waterbody, wetland, vegetation, and so forth. The physical conditions of Earth’s surface will potentially prompt the utilization of land for social and economic activities, described as land use, for example, industrial and agricultural areas [17], [26], [27]. Proper planning and management on land use could enable development with the least ecosystem damage [28].

In order to optimize and ensure the sustainability of land use, the concept of land suitability is employed in this research, by providing information on the relationship between the characteristics and the quality of the land. Land suitability indicates how to fit a particular class of land for the specific utilizations [28]. Land suitability could be assessed for both presents (actual) or after improvement (potential). It is evaluated by matching the land characteristics and its uses based on standard defined criteria. The results serve planners to identify optimal land use [29].

Land characteristics are commonly used as parameters in land suitability analysis. In this study, ecoregion, slope, elevation, and water provisioning are the parameters besides land-use for the suitability analysis. Socio-economic parameters can also be used for the analysis, such as population density, distance to the main road, and so forth. Socio-economic parameters are needed when searching for suitable locations for socio-economic activities, such as industry, trade, settlements, and others. Land use/land cover data, as the main parameter in this research, is originally classified into 22 classes. Reclassification to 8 classes is made based on the standard of land cover classes in the interpretation of Medium-Resolution Optical Imagery, including forest, wetland agriculture, dryland agriculture, plantation, built-area, grassland, embankments, and conservation area. Table 1 shows the data used in this study.

| Data                                      | Source                        | Year  |
|-------------------------------------------|-------------------------------|-------|
| Landuse/Landcover (1:250,000 scale resolution) | Ministry of Environment and Forestry | 2016  |
| The population of Java Island             | Statistics Indonesia          | 2016  |
| Land Surface Temperature                   | USGS and NASA                 | 2016  |
| Precipitation                             | World Climate                 | 1970 – 2000 |

This research utilizes the Support Vector Machine (SVM) to generate a suitability evaluation model by classifying various designations. SVM is an approach of classification algorithm which produces a hyperplane that represents the optimal division of linearly separable classes [25]. It is commonly used for hyperspectral image classification and object detection, and also developed in many applications, including multispectral analysis [25], [30]. In this study, SVM was employed with the kernel trick, which is a probabilistic approach as a computational algorithm. Kernel trick allows SVM to work on a non-linear problem, makes SVM a link from linear to nonlinear that simplify the dimensional
complexity of the calculation [31], [32]. The classification deals with two foremost steps, determining the system of classification and deciding the training samples.

![Image of parameters to generate land use suitability probabilities]

**Figure 2.** Parameters to generate land use suitability probabilities

### 3. Results and Discussions

#### 3.1. Significance Test of Parameters and Sampling

The parameters of elevation, slope, water availability, precipitation, temperature, and ecoregion are assumed to be related to land-use. The statement becomes a hypothesis (H0) that needs to be tested with a significance test. A significant understanding in research is the level of confidence in a hypothesis, whether the hypothesis was accepted or rejected. The level of significance in statistics is expressed as a number in percent, which indicates the likelihood or risk of errors in the results of tests conducted. The
significance levels commonly used by researchers are 5% and 1%. Statistical test results show a significance value between 0 (0%) and 0.024 (2.4%) for all parameters (Table 2). This means that based on test results, there is still a 2.4% chance of error and a 97.6% chance of correctness. This study uses the commonly used confidence level, which is 0.95 or 95%. It means a 95% chance that the results can be trusted. Thus, for a 95% confidence level, the hypothesis that all parameters have a relationship to land-use is accepted.

Table 2. Significance value of independent variables

| Parameter     | Significance |
|---------------|--------------|
| Elevation     | 0.000        |
| Slope         | 0.000        |
| Water availability | 0.000   |
| Precipitation | 0.000        |
| Temperature   | 0.024        |
| Ecoregion     | 0.000        |

Sampling was conducted on a population of 159,757 grids in Java Island. The sample size in this model is 1,050 grids by taking 150 sample grids for each type of land-use so that there are 1,050 sample grids for seven types of land-use (other than protected/conservation areas). The number of sample points refers to the standard number of samples in SNI ISO 19157: 2015, ie, with a population of data of 159,757 grids (between 150,001 - 500,000 grids) a minimum number of 800 grid samples is used [33].

Sampling was carried out using the stratified random sampling method. This method allows the population to be divided into strata or classes that do not overlap or be divided into several subpopulations that each sub has homogeneous samples [33]. Each class, in this case, the land-use class, has its characteristics. For each type of land-use, the determination of 150 sample grid locations is carried out randomly in all regions of Java.

3.2. Calculation of Land Suitability Probability Values

In this model, all grids in Java already have the six attributes of the physical parameter values and land-use class in 2016. In the previous stage, samples were also determined. The sample grid is used as a training point for classifying other grids. Based on a sample grid that contains physical parameter values and land use classes, SVM forms hyperplane lines with certain margins that separate one land-use class from another. In theory, one land-use class can be changed to all other land use classes, but with varying degrees of suitability. Probability values are used to represent the level of suitability of land use. SVM does not provide probability values at the time of classification. Therefore, SVM is combined with the error-correcting output coding (ECOC) algorithm. ECOC is a kind of method which converts a multi-class classification problem into two-class classification problems. In ECOC, the code bit values in each column of the error-correcting code matrix are used to reclassify the training samples. Thus, the probability value of land suitability of the seven land use classes can be calculated.

Table 3 is the probability value for an n-grid in Java Island. With the ECOC SVM method, each grid in Java has seven probability values for each land-use class, which is between 0% to 100%. The land-use class, with the highest probability value, is considered the most suitable land-use class for a grid. For example, grid A has the highest probability value in the wetland agriculture class, which is 88.20% and the lowest in the embankments class, which is 0.01%. As such, grid A is best suited for the wetland
agriculture class, and least suitable for embankments. The probability value can later be used for the spatial allocation of land uses.

Table 3. Probability Values of Each Grid for Seven Land Use Classes

| Grid | Forest  | Wetland agriculture | Dryland agriculture | Plantation | Built-area | Grassland | Embankments |
|------|---------|---------------------|---------------------|------------|------------|-----------|-------------|
| A    | 0.48%   | 88.20%              | 2.55%               | 0.02%      | 8.75%      | 0.00%     | 0.01%       |
| B    | 5.32%   | 66.03%              | 14.94%              | 0.86%      | 12.66%     | 0.14%     | 0.05%       |
| C    | 0.91%   | 88.63%              | 6.73%               | 0.01%      | 3.69%      | 0.01%     | 0.02%       |
| D    | 18.85%  | 0.28%               | 76.50%              | 0.03%      | 0.86%      | 3.45%     | 0.02%       |
| E    | 1.51%   | 85.91%              | 2.11%               | 2.88%      | 7.55%      | 0.04%     | 0.00%       |
| F    | 0.81%   | 4.58%               | 75.00%              | 0.10%      | 18.31%     | 0.90%     | 0.31%       |
| G    | 0.28%   | 82.60%              | 1.63%               | 0.01%      | 0.29%      | 0.00%     | 15.18%      |
| H    | 0.10%   | 98.75%              | 1.10%               | 0.01%      | 0.01%      | 0.00%     | 0.02%       |
| I    | 19.61%  | 1.66%               | 74.26%              | 3.99%      | 0.38%      | 0.08%     | 0.03%       |

The highest probability value also indicates the level of confidence of the land-use suitability model. The highest probability value in Java reached 0.995004 or 99.5004%, with an average of 0.693878 or 69.3878%. Based on the distribution graph, the highest probability values on the entire grid converge above the average (> 69.3878%). The statement was also supported by the results of the visualization of the distribution of the highest probability values in Java, which showed that from 159,757 grids in Java Island, the probability values were dominated between 75,001% - 100% (Figure 3).

Figure 3. Distribution of maximum probability values of each grid

The validity of the land-use suitability model can be seen by cross-tabulation between the total quantification of the model and the existing land use (Table 4). The area difference between the model and PL 2016 shows the accuracy of the land suitability model. Also, the broad percentage pattern of each land-use class in the model shows the same pattern as the existing land use. This probability value is influenced by parameters calculated in the model. From the results of the modeling, the highest probability on each grid shows the suitability of the type of land use. Some land-use classes have a larger area of modeling results than existing land uses. This means that when the class has a wide deficit, it can be partially fulfilled with a grid that has land suitability. For example, plantation classes have a broad percentage based on the model of 40.31%, while existing conditions are 35.73%. If the allocation
of the plantation class experiences a deficit, then it can be fulfilled from the difference in the area of land that has the suitability of the plantation land.

**Table 4.** Cross Tabulation of Land-Use Model and Existing Land Use

| LU 2016 | LU Model (6 Parameters) | Total |
|--------|-------------------------|-------|
|        | 1                       | 2     | 3    | 4    | 5    | 6    | 7    |       |
| 1      | 11.74%                  | 0.57% | 5.40%| 0.12%| 0.10%| 0.01%| 0.01%|17.95% |
| 2      | 0.52%                  | 26.41%| 1.94%| 0.01%| 1.20%| 0.00%| 0.09%|32.16% |
| 3      | 3.12%                  | 1.55% | 29.82%| 0.14%| 1.08%| 0.01%| 0.00%|35.73% |
| 4      | 0.40%                  | 0.13% | 1.85%| 0.73%| 0.03%| 0.00%| 0.00%|3.14%  |
| 5      | 0.24%                  | 0.96% | 1.18%| 0.03%| 5.31%| 0.01%| 0.03%|9.25%  |
| 6      | 0.14%                  | 0.02% | 0.11%| 0.01%| 0.02%| 0.12%| 0.00%|0.42%  |
| 7      | 0.01%                  | 0.49% | 0.02%| 0.00%| 0.04%| 0.00%| 0.79%|1.35%  |
| Total  | 16.17%                 | 32.13%| 40.31%| 1.04%| 9.27%| 0.15%| 0.92%|100.00%|

Notes:
1 = forest  
2 = wetland agriculture  
3 = dryland agriculture  
4 = plantation  
5 = built-area  
6 = grassland  
7 = embankments

4. Conclusions

The determination of alternative locations for potential land use can be done probabilistically using the SVM method by taking into account the physical parameters of the land, namely height, slope, water availability, ecoregion, precipitation, and temperature. However, most grid probability values are quite high (>75%). Probability values of land use suitability can be increased by adding parameters that certainly have a significant relationship with land use.

5. References

[1] U. Abulof, “Introduction: Why We Need Maslow in the Twenty-First Century,” *Society*, vol. 54, no. 6, pp. 508–509, 2017.
[2] R. H. M. Fallatah and J. Syed, “A Critical Review of Maslow’s Hierarchy of Needs,” in *Employee Motivation in Saudi Arabia*, Cham, Switzerland: Springer, 2013, pp. 19–60.
[3] M. Lane, L. Dawes, and P. Grace, “The essential parameters of a resource-based carrying capacity assessment model: An Australian case study,” *Ecol. Modell.*, vol. 272, pp. 220–231, 2014.
[4] D. H. Meadows, J. Randers, and D. L. Meadows, *Limits to Growth: The-30-Year Update*, 3rd ed. London, UK: Chelsea Green Publishing, 2005.
[5] M. Borucke et al., “Accounting for demand and supply of the biosphere’s regenerative capacity: The National Footprint Accounts’ underlying methodology and framework,” *Ecol. Indic.*, vol. 24, pp. 518–533, 2013.
[6] M. Wackernagel, D. Lin, L. Hanscom, A. Galli, and K. Iha, “Ecological Footprint,” *Earth Syst. Environ. Sci.*, vol. 4, no. March 2018, pp. 270–282, 2019.
[7] W. Steffen et al., *Global Change and the Earth System: A Planet Under Pressure*, 1st ed. Heidelberg, Germany: Springer-Verlag Berlin Heidelberg, 2005.
[8] UNEP, *Global Environment Outlook 5*. Nairobi: United Nations Environment Programme (UNEP), 2012.
[9] H. E. Daly, “Toward some operational principles of sustainable development,” *Ecol. Econ.*
[10] IPCC, “Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change,” Cambridge University Press, Cambridge UK and New York, USA, 2014.

[11] M. S. Mancini et al., “Ecological Footprint: Refining the carbon Footprint calculation,” Ecol. Indic., vol. 61, pp. 390–403, 2016.

[12] M. Lane, “The development of a carrying capacity assessment model for the Australian socio-environmental context,” Queensland University of Technology, 2014.

[13] Land and Water Development Division FAO, Planning for sustainable use of land resources: Towards a new approach. Rome, Italy: Food and Agriculture Organization of The United Nations, 1995.

[14] D. Djajundin, R. Oktaviani, S. Hartoyo, and H. Dwiprabowo, “Modelling of land allocation behavior in Indonesia,” Procedia Environ. Sci., vol. 33, pp. 78–86, 2016.

[15] R. W. Mcdowell et al., “The land use suitability concept: Introduction and an application of the concept to inform sustainable productivity within environmental constraints,” Ecol. Indic., vol. 91, no. April, pp. 212–219, 2018.

[16] M. Rusdi, R. Roosli, and M. S. S. Ahamad, “Land evaluation suitability for settlement based on soil permeability, topography and geology ten years after tsunami in Banda Aceh, Indonesia,” Egypt. J. Remote Sens. Sp. Sci., vol. 18, no. 2, pp. 207–215, 2015.

[17] FAO, A framework for land evaluation. FAO Soils Bulletin No. 32, no. M–51. 1976.

[18] C. A. van Diepen, H. van Keulen, J. Wolf, and J. A. A. Berkhouit, “Land Evaluation: From Intuition to Quantification,” in Advances in Soil Science, vol. 15, B. A. Stewart, Ed. New York, USA: Springer, 1991, pp. 139–140.

[19] Badan Pusat Statistik, Statistik Indonesia 2018. Jakarta, Indonesia: Badan Pusat Statistik (BPS), 2018.

[20] Kementerian Energi dan Sumber Daya Mineral RI, Statistik Ketenagalistrikan 2016. Jakarta, Indonesia: ESDM, 2017.

[21] International Energy Agency, “Market Report Series: Energy Efficiency 2017,” Jakarta, Indonesia, 2017.

[22] R. Kurniawan, Y. Sugiawan, and S. Managi, “Cleaner energy conversion and household emission decomposition analysis in Indonesia,” J. Clean. Prod., vol. 201, pp. 334–342, 2018.

[23] A. Wijaya et al., “How Can Indonesia Achieve Its Climate Change Mitigation Goal? An Analysis of Potential Emissions Reductions From Energy and Land-Use Policies,” Jakarta, Indonesia, 2017.

[24] M. Shahbaz, Q. M. A. Hye, A. K. Tiwari, and N. C. Leitao, “Economic growth, energy consumption, financial development, international trade and CO2 emissions in Indonesia,” Renew. Sustain. Energy Rev., vol. 25, pp. 109–121, 2013.

[25] B. R. Deilmai, B. Bin Ahmad, and H. Zabihi, “Comparison of two classification methods (MLC and SVM) to extract land use and land cover in Johor Malaysia,” IOP Conf. Ser. Earth Environ. Sci., vol. 20, no. 1, p. 6 pp, 2014.

[26] H. George, “An overview of land evaluation and land use planning at FAO,” Rome, Italy, 2005.

[27] D. Rudrapal and M. S. Subhedar, “Land Cover Classification using Support Vector Machine,” Int. J. Eng. Res. Technol., vol. 4, no. 9, pp. 584–588, 2015.

[28] J. Gong, Y. Liu, and W. Chen, “Land suitability evaluation for development using a matter-element model: A case study in Zengcheng, Guangzhou, China,” Land use policy, vol. 29, no. 2, pp. 464–472, 2012.

[29] S. Ritung, Wahyunto, F. Agus, and H. Hidayat, Land Suitability Evaluation with A Case Map of Aceh Barat District. Bogor, Indonesia: Indonesian Soil Research Institute and World Agroforestry Centre, 2007.

[30] G. M. Foody and A. Mathur, “A relative evaluation of multi-class image classification by support vector machines,” IEEE Trans. Geosci. Remote Sens., vol. 42, pp. 1335–1343, 2004.
[31] Y. Liu, Z. Zhang, and Y. Zhou, “Efficiency of construction land allocation in China: An econometric analysis of panel data,” *Land use policy*, vol. 74, pp. 261–272, 2018.

[32] C. J. C. Burges, “A Tutorial on Support Vector Machines for Pattern Recognition,” *Data Min. Knowl. Discov.*, vol. 2, pp. 121–167, 1998.

[33] BSN and Badan Informasi Geospasial, *SNI ISO/TS 19157:2015 Informasi geografis - Kualitas data*. Indonesia: Sekretariat BSN, 2015, p. 170.