Integrating Data Mining Techniques and GIS Technology in Land Suitability Assessment - Case Study for Durian Trees in Cho Lach District, Ben Tre Province, Vietnam

Huy Anh Nguyen[1], Thi Ngoc Diem Vo[1, 2], Van Anh Nguyen[1]

[1]Faculty of Land Administration, Ho Chi Minh City University of Natural Resources and Environment, Vietnam

[2] Tien Giang Department of Natural Resources and Environment, Vietnam

E-mail: anhnh@hcmunre.edu.vn

1. Introduce

Decision trees in data mining are considered an effective method for processing training data sequences, specifically by efficiently handling discrete variables and categorical variables of soil characteristics. This is a new method for quantifying the relationships between soil properties and crop yields.

In this study, data mining involves applying analytical and machine learning techniques to extract information from data on soil characteristics and crop yields to assess land suitability for durian trees in Cho Lach district, Ben tre Province.

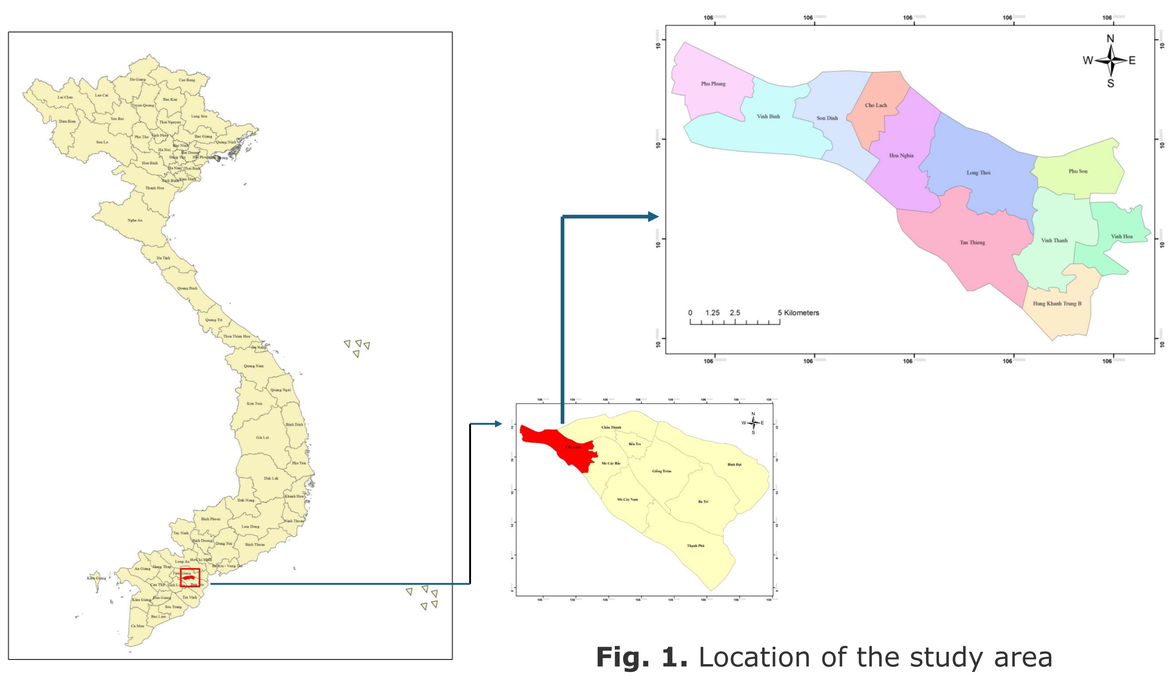

Fig. 1. Location of the study area

2. Method

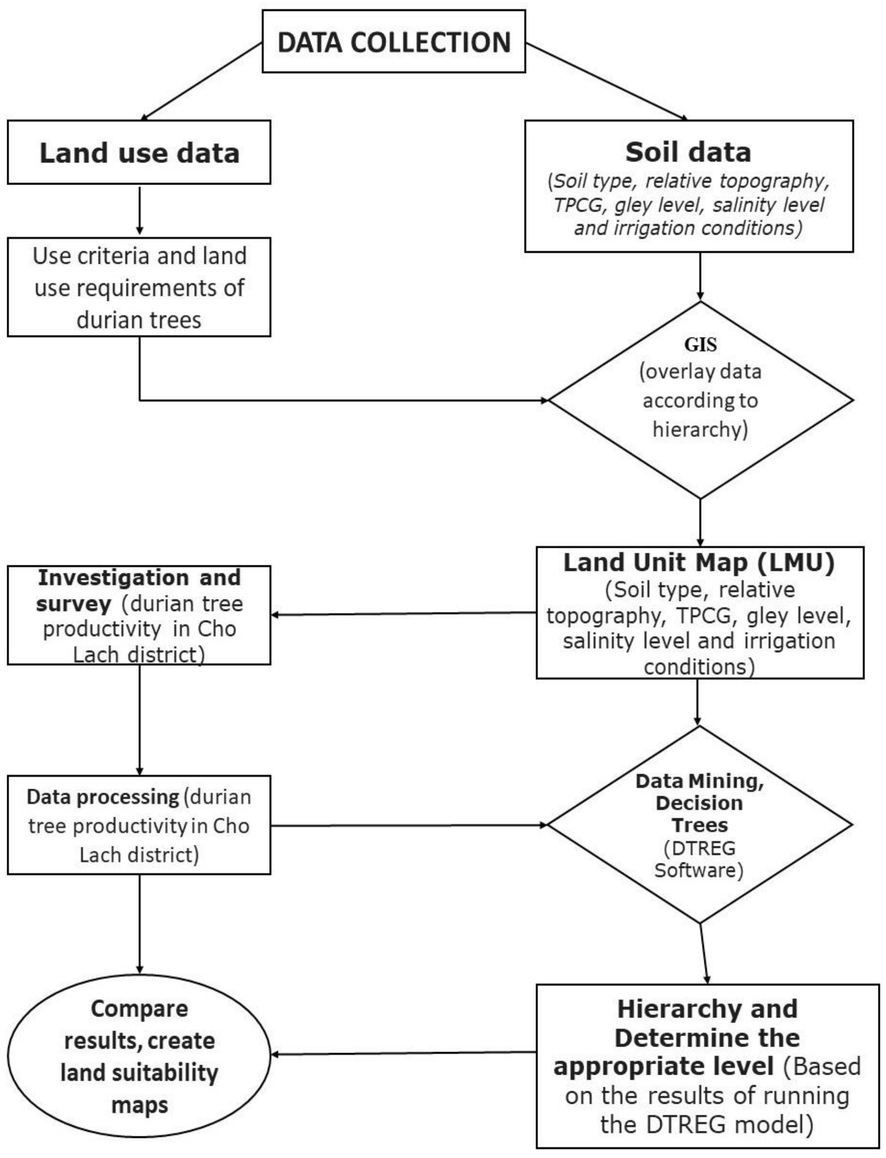

Fig. 2. Research process and methods

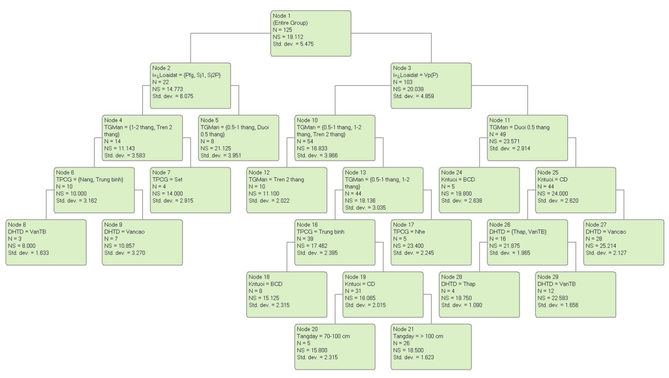

Fig. 10. The Decision Tree model evaluates the suitability for durian trees

3. Results and Discussion

The study has built 6 component maps: soil type, soil thickness, mechanical composition, terrain, salinity level, irrigation capacity.

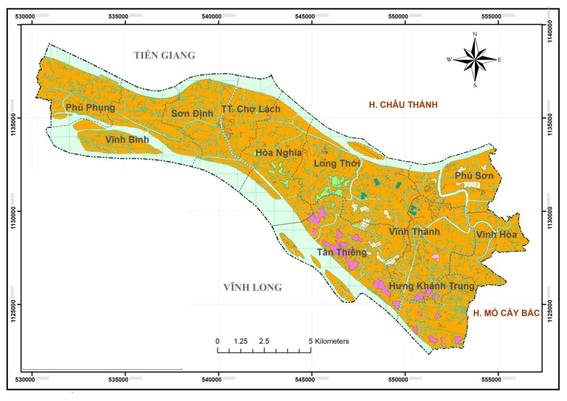

Fig. 2. Soil type map

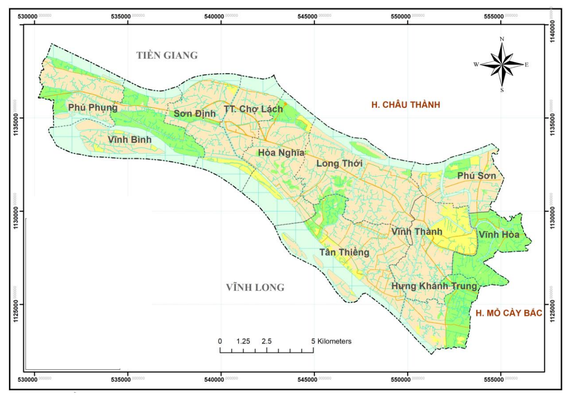

Fig. 3. Terrain map

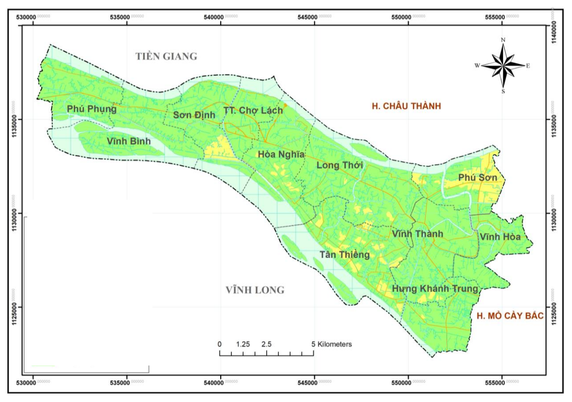

Fig. 4. Soil thickness map

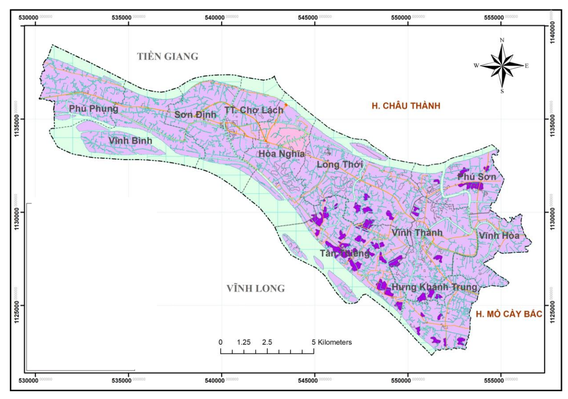

Fig. 5. Mechanical composition map

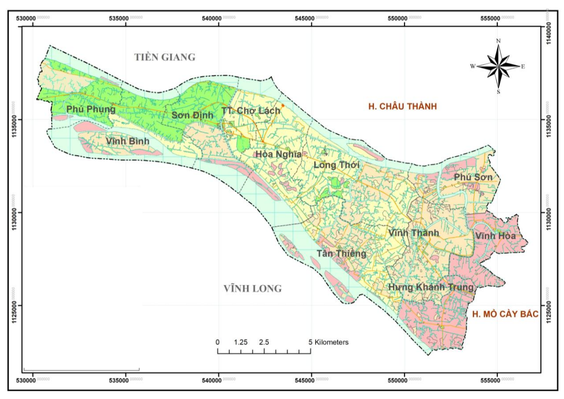

Fig. 6. Salinity level

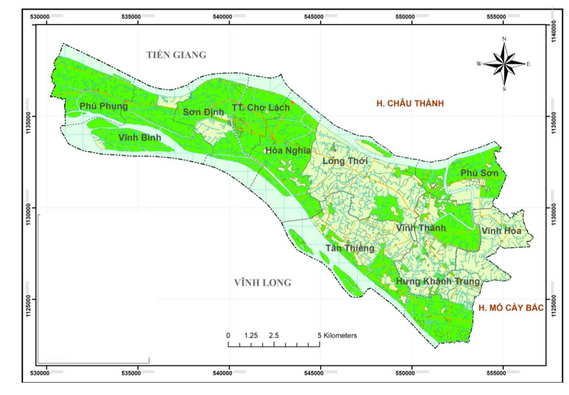

Fig. 7. irrigation map

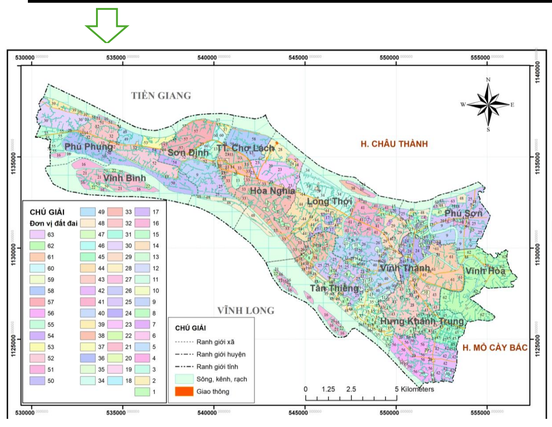

Fig. 8. LMUs map

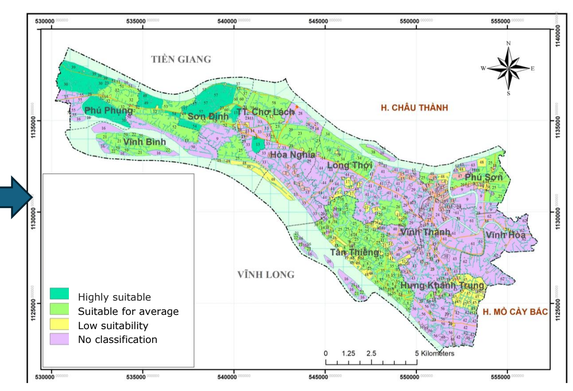

Fig. 9. Classification map

4. Conclusions

In Cho Lach district, Ben Tre province, 36 units (with durian trees planted) were evaluated for land suitability using the decision tree model using DTREG software. The findings indicate that the land is appropriate. 36.78% of the land is unevaluated, meaning that durian is either not planted at all or has not been steadily harvested with inappropriate or unsuitable outcomes. High (S1) accounts for 6.82%, medium suitable (S2) serves for 28.66%, and low suitable (S3) 810.93 hectares accounts for 4.8%.

The data mining method based on the decision tree model is very meaningful and reliable to supplement current land assessment methods, but it also has limitations so it is necessary to research and use it in combination. with other assessment methods to classify land suitability for crops