# THE INTEGRATION OF THE CLUE-S MODEL WITH GEOGRAPHIC INFORMATION SYSTEMS (GIS) ENHANCES LAND USE PLANNING

**Nguyen Duc Tri[1,2,3] , Dinh Thi Kim Phuong[1,2, *]**

[1]*Faculty of Environment and Natural Resources, Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam*

[2]*Vietnam National University Ho Chi Minh City, Ho Chi Minh City, Vietnam*

[3]*Sub-National Institute of Agricultural Planning and Projection*

*Email: *dtkphuong@hcmut.edu.vn*

Received: 2 May 2025; Revised: 20 May 2025; Accepted: 31 May 2025

## ABSTRACT

Land use planning consists of two main components: land evaluation and spatial allocation of land use plans. Land evaluation is conducted using the maximum limitation method and multi-criteria evaluation (MCE). The CLUE-S model enhances this process by applying logistic regression analysis to assess land suitability. For spatial allocation, CLUE-S employs Cellular Automata (CA) as a primary method, integrating it with decision rules and Markov analysis. In this study, the integration of CLUE-S and Geographic Information Systems (GIS) was applied in Cat Tien district, Lam Dong province, focusing on three land use types: rice, cashew, and vegetables. Three logistic regression models were developed based on 17 property classes of land characteristics, with the area under the ROC curve (AUC) used to measure the Goodness of Fit (GoF). Parameters β (significance code < 5%) were utilized for land suitability mapping. Using demand tables and land suitability maps, CLUE-S allocated land use for each year up to 2030, producing simulated land use maps for the planning period. This integration of CLUE-S and GIS serves as a valuable tool for land use planners and policymakers, facilitating the analysis of land use changes and complementing existing methods for effective land use planning.

*Key words:* CLUE-S model, logistic regression analysis, land use planning (LUP), Markov analysis, GIS.

## 1. INTRODUCTION

Land use planning is a fundamental component of state land management. A key deliverable of the planning process is the land use planning map, which visualizes the spatial distribution of land use types (LUTs) at the intended time horizon. Developing such maps generally involves two sequential steps: (1) assessing land suitability for different LUTs based on ecological criteria, and (2) allocating land uses spatially according to this suitability and demand dynamics. According to Vietnam's Ministry of Natural Resources and Environment's guidelines, the classification of agricultural land often stops at general categories such as perennial (CLN) or annual crops (HNK), without identifying specific crop types. However, effective agricultural land use planning demands greater granularity, specifying crop types like rice, cashew, coffee, tea, and vegetables. This level of detail requires a robust understanding of each crop's ecological requirements to perform land suitability assessments that guide precise spatial allocations.

***(1) Land evaluation methods***

Traditional land evaluation methods, such as the FAO framework (1976) and multi-criteria evaluation (MCE), have been widely employed in assessing land suitability. However, this study adopts a statistical modeling approach using the CLUE-S (Conversion of Land Use and its Effects at Small regional extent) framework to generate crop-specific suitability maps across the study region. Unlike conventional methods, the statistical approach excels in predictive modeling and allows for a robust analysis of the relationships between land characteristics and specific land use types. Its comprehensive data analysis capabilities enable the investigation of complex interactions that are often beyond the scope of deterministic or rule-based methods.

***(2) Land use allocation***

Land use allocation involves applying technical and algorithmic methods to assign land use types (LUTs) spatially, guided by land suitability assessments, land use demand, conversion potential among LUTs, and the exclusion of non-agricultural areas. A core challenge lies in determining optimal locations and extents for each LUT. In this study, the CLUE-S model is employed as a complementary tool to existing methodologies, utilizing statistical evaluation outcomes, decision rules, and Markov analysis for spatial allocation. CLUE-S effectively integrates empirical statistical approaches—particularly logistic regression—with decision rules to simulate location-specific land use changes. However, its limitation lies in the lack of direct map outputs, as it relies on ASCII data formats, which can hinder practical usability. In contrast, Geographic Information Systems (GIS), widely adopted since the 1960s, offer robust capabilities in spatial data handling, analysis, and map generation. GIS's compatibility with diverse data formats, including ASCII, effectively mitigates CLUE-S's limitations. Therefore, integrating CLUE-S with GIS emerges as a practical and urgent solution to enhance the spatial planning of land use.

## 2. MATERIAL

### 2.1. Overview of the Integration of the CLUE-S Model and GIS in Land Use Planning

#### *2.1.1. Structure and Operation of the CLUE-S Model*

In 1996, the CLUE model was first introduced globally through its initial implementation as CLUE-CR, applied to Costa Rica [2]. Subsequently, the model was further developed by P.H. Verburg into a more detailed version known as CLUE-S. The fundamental difference between the original CLUE model and CLUE-S lies in the level of spatial detail, with CLUE-S offering higher spatial resolution for more precise land use analysis.

Due to spatial scale differences, the original CLUE model cannot be directly applied to small regional extents. When the study area is relatively small, the way data is represented differs significantly from applications at broader spatial scales. As a result, the modeling approach was modified, leading to the development of CLUE-S (Conversion of Land-Use and its Effects at Small regional extent).

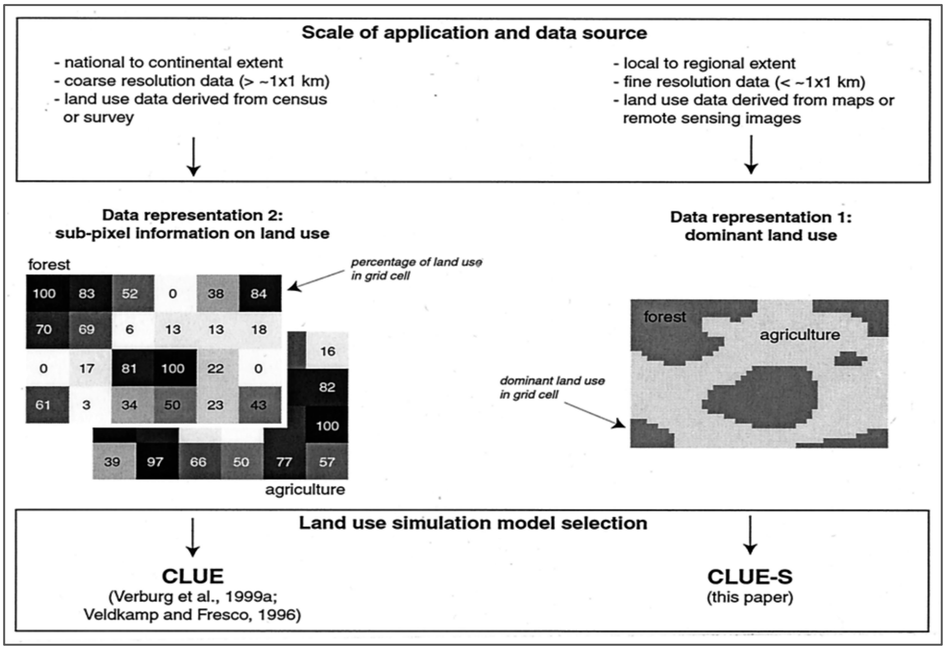


*Figure 1.* Scale of application and data source of CLUE & CLUE-S [1]

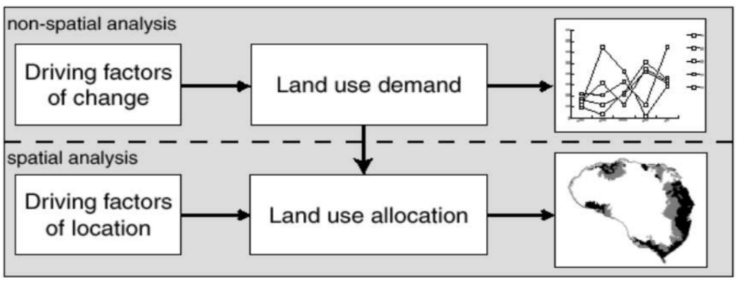


*Figure 2.* Overview of the modelling procedure [1]

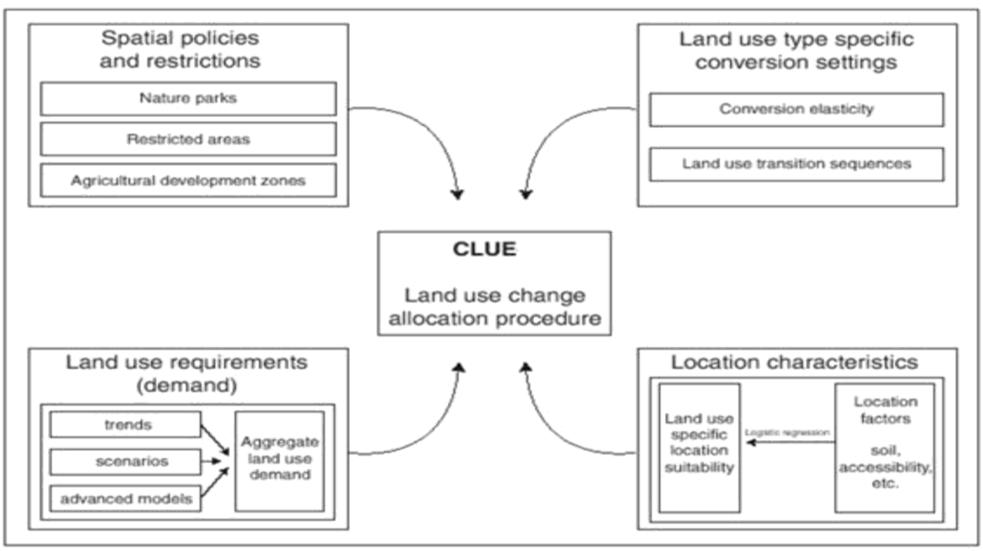


*Figure 3.* Overview of the information flow in the CLUE-s model [1]

The CLUE-S model consists of two distinct components: a non-spatial and a spatial module. The non-spatial component calculates the aggregate future land use demands based on various scenario inputs. These outputs are then transferred to the spatial component, which allocates land use changes across different locations. The simulation results include annual tabular data detailing land use dynamics and corresponding maps illustrating yearly land use change.

Spatial allocation in the CLUE-S model is based on a combination of empirical analysis and spatial modeling. Figure 4 provides an overview of the key components required to run the CLUE-S model. These components are categorized into four groups, which collectively establish the conditions and probability surfaces that the model uses to compute optimal land use allocation through an iterative procedure.

The components are categorized into four groups as follows: Delineation of agricultural and non-agricultural land boundaries; Configuration of CLUE-S model parameters; Land use demand determined through socio-economic assessment; Outcomes from land use specific location suitability.

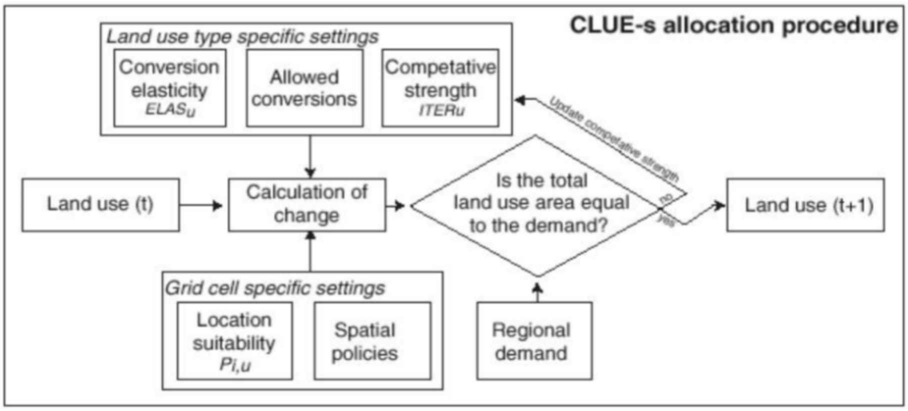


*Figure 4.* Flow chart of the allocation module of the CLUE-S model [1]

*i. Land use specific location suitability*

Land use changes are projected to occur at locations with the highest "preference or suitability" scores for specific land use types. These preferences reflect the outcomes of interactions among various agents and decision-making processes, ultimately shaping a spatial land use configuration.

A statistical model can be developed using a binary logit framework, with two alternatives: whether or not location $i$ is converted to land use type *k*. The associated probability function links this decision to both natural and socio-economic location factors, as expressed in the following logistic regression model:

$$log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_n x_{in} \quad (1)$$

Here, $p_i$ denotes the probability that pixel $i$ is assigned to a given land use type, and X represents the set of location-specific variables. The coefficients β are estimated using logistic regression, where current land use classes serve as the dependent variable. Most location factors are site-specific characteristics such as soil properties and elevation.

*ii. Land use type specific conversion settings*

The **land use conversion matrix** plays a critical role in the CLUE-S model, enabling calculations of transitions between different land use types (LUTs). In CLUE-S, a binary convention is used: a value of **1** indicates that conversion between LUTs is allowed, whereas **0** signifies that conversion is not permitted. For instance, in the case study of Cát Tiên District, a land use conversion matrix was constructed to define feasible transitions between land use categories.

*iii. Land use change allocation procedure in CLUE-S model*

Once all input factors are provided, the CLUE-S model performs calculations over discrete time steps, typically using an annual time interval. The entire allocation process is summarized in Figure 4. The following steps are executed to allocate land use changes:

*2.1.2. Raster-Based Spatial Analysis in GIS*

A core function of Geographic Information Systems (GIS) is the ability to analyze spatial and attribute data to support decision-making. Spatial analysis is conducted to address real-world questions, which can vary in complexity—from simple logical or arithmetic operations to sophisticated spatial modeling. It is this spatial analytical capability that distinguishes GIS from conventional database management systems.

In this study, most analyses were performed using raster data models within the GIS. These analytical functions are essential for managing both the spatial input and output data of the CLUE-S model.

It is important to emphasize that the CLUE-S model cannot function without the integration of GIS for pre-processing its input data. GIS spatial analysis capabilities are critical for tasks such as standardizing coordinate systems, ensuring spatial consistency of the study area, and customizing attributes within the raster data model.

## 2.2. Global Applications and Customization of the CLUE-S Model

Over nearly 30 years of development across various versions, the CLUE model—and in particular CLUE-S—has been widely applied in diverse landscapes and ecological zones around the world. It has been used extensively to simulate land use change and support policy planning across a range of environments, from tropical regions (Asia, South America, Central Africa) to temperate zones (Europe, the United States), and from monsoon climates in Southeast Asia to desert regions such as Inner Mongolia, China. Versions of the CLUE model, including CLUE-S and Dyna-CLUE, have been deployed in numerous land use contexts, including agriculture, forestry, and urbanization.

Notable case studies include applications in Sibuyan Island [3], Selangor River [4], Taips County [6], Pennsylvania County [7], Northern Thailand [8], Sangong Watershed [9], etc.

In Vietnam, the CLUE and CLUE-S models have been applied to study land use change at two primary levels: national and provincial (e.g., Bắc Kạn and Quảng Nam).

- National level: The study titled *"Land Use Change, Food Security, and Climate Change in Vietnam – A Global-to-Local Modeling Approach"* was supported by CDKN and the Netherlands Ministry of Economic Affairs. Its objective was to optimize Vietnam's land use policies. The study integrated the MAGNET economic model (for socio-economic assessment) with the CLUE land allocation model to simulate three development scenarios from 2008 to 2030 [11].
- Provincial/Regional level: The CLUE-S model has been applied in several subnational contexts, including Bắc Kạn Province, Quảng Nam Province, and the Bé River basin [12].

General Assessment of CLUE-S Applications Globally and in Vietnam: The CLUE-S model has been widely adopted worldwide and has proven to be a practical tool for land use planners. Most applications tend to follow a common pattern in selecting land evaluation input variables, often structured around predefined frameworks.

However, in the case of the Cát Tiên study area, a large volume of detailed soil property data was available from prior surveys. This study departs from the conventional framework by leveraging this unique dataset. The planning approach adopted here follows the GTZ land use planning methodology, which emphasizes participatory planning and integrates the CLUE-S model with GIS tools to support land suitability evaluation and spatial allocation in the land use planning process.

## 2.3. Data Collection

### *2.3.1. Development of Current Land Use Data and Land Characteristics Data*

The current land use status map was derived from the land inventory data and the 2020 current land use status map of Cat Tien district (in .dgn format) provided by the Lam Dong Province Department of Natural Resources and Environment. Based on this, field verification, adjustments, and additions were conducted to establish the 2020 current land use status map of the district in a raster GIS environment with a spatial resolution of 100m. Additionally, input data regarding land properties are crucial for operating the CLUE-S model. Following an analysis of the practical issues within the study area, and based on the actual conditions of Cat Tien district and the ecological requirements of various land use types, the following land properties were classified into two types of data variables for input into the model:Continuous Variables: Slope, Elevation; Binary Categorical Variables: Soil type (7 categories), Soil depth (4 classes), Irrigation (4 levels of irrigation potential). From these data, thematic maps of integrated land properties were processed, including: Elevation map, Slope map, Soil map, Soil depth map, and Irrigation potential map. In this study are intended to support agricultural zoning and enhance land use planning by providing technical input for the land use map of Lâm Đồng province at a scale of 1:100.000. Accordingly, all input data used for simulation in the CLUE-S model are at a scale of 1:100,000 or finer to ensure spatial accuracy.

Furthermore, climatic factors such as humidity, evaporation, and rainfall also influence land use processes. However, given that these indicators exhibit nearly uniform values across Cat Tien district, it was deemed unnecessary to incorporate this information into the model.

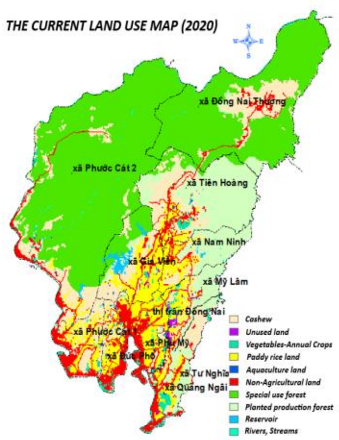


*Figure 5.* The current LU map (2020)

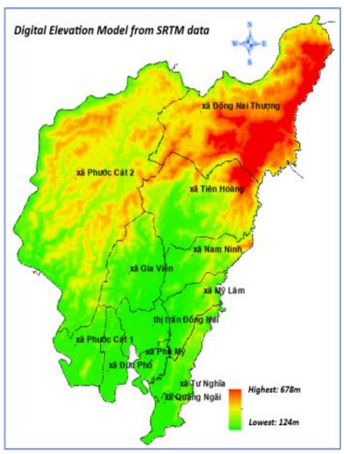


*Figure 6.* Elevation (DEM - SRTM2000)

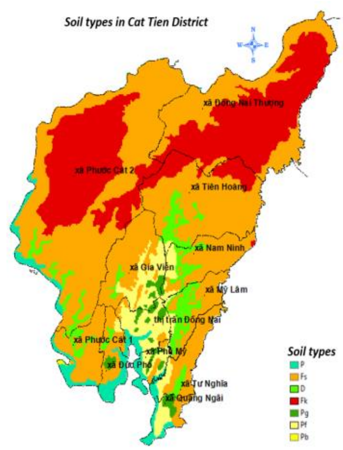


*Figure 7.* Soil type

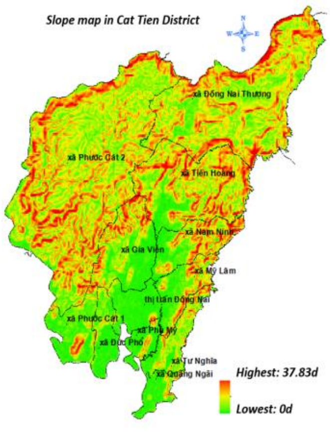


Figure 8. Slope

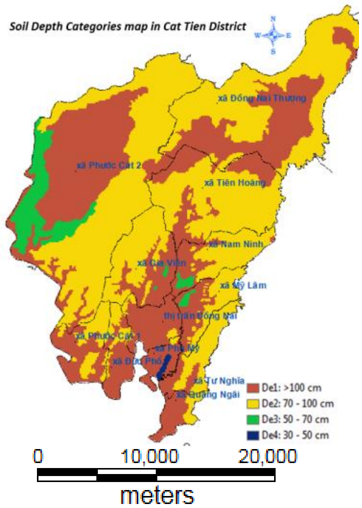


Figure 9. Soil Depth

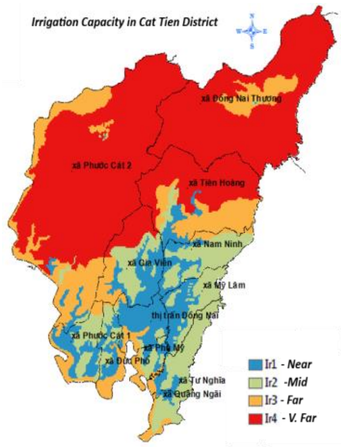


Figure 10. Irrigation capacity

*Source: Sub-NIAPP (2020)*

*2.3.2. Defining Permissible Zones for Agricultural Land Allocation*

In spatial land use modeling, especially within the CLUE-S framework, defining the areas where land use types can be allocated is a crucial step that directly influences the accuracy and feasibility of simulation results. For agricultural land allocation, this involves delineating zones where cultivation is legally, physically, and environmentally permissible.

The zoning process begins with the exclusion of areas unsuitable for agricultural purposes, such as existing urban settlements, water bodies, protected natural reserves, and steep or erosion-prone terrains. These areas are designated as "exclusion zones" and are masked out in the spatial model to prevent allocation of any agricultural land use types.

Subsequently, the remaining land is assessed for its biophysical and socio-economic suitability for specific crops, based on the classified land suitability maps. Only those areas falling within the "suitable" categories (e.g., S1, S2, or S3, as defined in FAO, 1976) are considered permissible for future allocation of agricultural LUTs.

This zoning process ensures that the spatial allocation procedures are constrained to realistic and policy-compliant zones, thus enhancing the robustness of model outputs. The integration of exclusion zones and suitability thresholds plays a pivotal role in aligning land use simulation with real-world planning constraints and environmental sustainability goals.

The agricultural land use planning area was delineated in coordination with the Department of Natural Resources and Environment of Cát Tiên District and other relevant local agencies. This process involved surveying sectoral land use plans and gathering input from the community, including local preferences and land use expectations. Based on a harmonized demand assessment across agricultural and non-agricultural land categories, spatial analysis was conducted within a GIS environment to identify the boundaries and total area designated for agricultural use. This geospatial framework ensures that all land use transitions among different LUTs are confined strictly within the predefined agricultural planning zone. No land use changes are permitted outside this designated boundary, maintaining consistency with both policy priorities and spatial planning constraints.

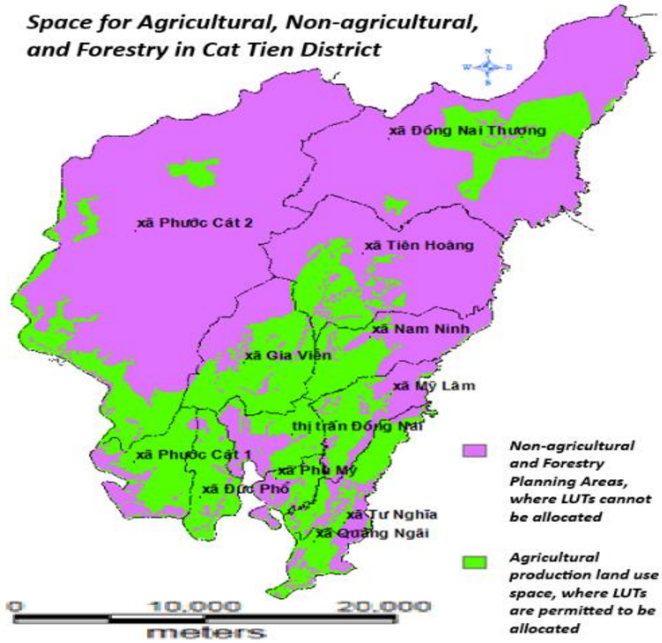


*Figure 11.* Space for agricultural, Non-Agricultural and Forestry in Cát Tiên District

## 3. METHODOLOGY

### 3.1. Relevant theories and methods

*3.1.1. Logistic regression for land use–specific location suitability*

Logistic regression is a widely adopted statistical technique used to model the probability of a categorical outcome based on a set of predictor variables. In the context of land use modeling, particularly within the CLUE-S framework, logistic regression is instrumental in evaluating the spatial suitability of specific land use types. The approach involves constructing a binomial logistic model in which the dependent variable represents the presence or absence of a particular land use type at a given location (e.g., pixel or spatial unit). Independent variables typically include a combination of biophysical and socio-economic factors - such as elevation, slope, soil characteristics, distance to roads, distance to water sources, and proximity to urban centers. In land use allocation procedures, these probabilities are interpreted as "preference values" or suitability scores that guide the spatial allocation of land use types. High suitability values indicate that a location has characteristics similar to those where a particular land use type is currently dominant, thus making it a strong candidate for future allocation. Logistic regression offers several advantages in this context: it is relatively easy to implement, interpretable, and capable of handling both continuous and categorical predictors. Furthermore, it accommodates non-linear relationships through transformations or interaction terms. The model's output, in the form of probability surfaces, serves as a critical input for spatial allocation routines in the CLUE-S model.

*(i). Importance of logistic regression*

In this study, the analysis involves both input (driving) variables and an output (dependent) variable. The driving factors encompass a diverse range of data types, including binary (e.g., presence/absence of infrastructure) and continuous variables (e.g., elevation, slope, distance to features). In contrast, the outcome variable—representing land use types (LUTs)—is binary in nature, taking values of 0 or 1 to indicate absence or presence at a specific location. Because the dependent variable is not continuous, conventional linear regression models are not suitable. Instead, the study employs the Logistic Regression Model (LRM), a method specifically developed to handle binary dependent variables. LRM is well-established in statistical modeling and is particularly appropriate for spatial land use applications.

Logistic regression thus provides a rigorous and interpretable framework for identifying the influence of various environmental and socio-economic factors on land use patterns. It is a foundational tool in generating suitability surfaces for land use allocation in models such as CLUE-S.

*(ii). Parameter estimation of the logistic regression model using maximum likelihood estimation (MLE)*

In logistic regression, parameter estimation is performed using the Maximum Likelihood Estimation (MLE) method, which identifies the coefficients ($\beta_0, \beta_1, \beta_2, \ldots$) by maximizing the likelihood function:

$$L(\beta_0, \beta_1, \beta_2 \ldots) = \prod_{i=1}^{N} \pi_i^{y_i} (1-\pi_i)^{n_i - y_i} = \prod_{i=1}^{N} \frac{exp\left(y_i(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots)\right)}{1 + exp(\beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots)} \quad (2)$$

In practice, statistical software such as **R**, **SAS**, or **SPSS** can be used to compute MLEs for logistic regression models. In the R programming environment, the estimation is commonly implemented using the glm() function for generalized linear models or the lrm() function provided in the rms (Regression Modeling Strategies) package. In this study, a custom R script will be developed utilizing built-in functions to estimate logistic regression parameters for each land use type (LUT). Variables with a confidence level of **≥ 95%** will be retained for further analysis. The model's performance for each outcome variable (i.e., LUT) will then be evaluated based on statistical significance and predictive accuracy.

*(iii). Evaluation of the logistic regression model using the receiver operating characteristic (ROC) curve*

To evaluate the predictive performance of the logistic regression model (LRM), this study employs the Receiver Operating Characteristic (ROC) curve, a standard method for assessing the discriminative ability of binary classifiers. The ROC curve plots the true positive rate (sensitivity) against the false positive rate (1 - specificity) at various threshold settings.

The area under the ROC curve (AUC) is a commonly used summary measure that reflects the model's ability to correctly distinguish between locations where a specific land use type is present versus absent. An AUC value of 0.5 indicates no discriminative power (equivalent to random guessing), while an AUC closer to 1.0 indicates excellent model performance.

In this study, ROC curves and AUC values will be computed for each land use type modeled using LRM. The results provide insight into how well the selected driving factors predict the spatial occurrence of land use, thereby validating the suitability of the logistic regression models before their integration into the CLUE-S land allocation module. AUC quantifies the model's ability to distinguish between classes and is computed by integrating the area under the ROC curve. The trapezoidal rule is applied for approximation:

$$AUC = \frac{1}{2}\sum_{t=1}^{n}[x_{i+1} - x_i][(y_{i+1} + y_i)] \qquad (3)$$

This research employs the R programming language to compute AUC values ($AUC_{LUTj}$) for each LUT (j=1÷k). Key findings include: AUC values ≥ 0.7 indicate robust model performance for generating land suitability maps, aligning with established thresholds in land-use research [13], [14], [15], [16].

*3.1.2. Markov Analysis*

Markov Analysis[17]: Helps determine the land use conversion frequency of each land use type (LUT) from the past to the current state, for example, the period 2000-2010. This conversion frequency is the ELAS coefficient for each LUT in the study area, which is necessary for declaration during simulation. Most LUTs can be converted to other LUTs with a reversible conversion probability lying between the other two cases. There is a characteristic coefficient for the reversibility of conversion back to the original LUT, which is Elasticity (ELAS). In the case of ELAS = 1, it is impossible to return to the initial state. Conversely, an ELAS coefficient of 0 indicates an easy return to the initial state. The coefficient $ELAS \in [0,1]$

### 3.2. The integrated CLUE-S and GIS modeling framework supports land use planning

Given that CLUE-S is a software package solely capable of suitability assessment and spatial allocation, all its input and output data are in ASCII file format. Geographic Information Systems (GIS) are utilized as the "eyes" for spatial processing of input maps, visualization, spatial analysis, and the generation of reports from the CLUE-S model's output data. The intermediate data format for communication between CLUE-S and GIS is ASCII files. The GIS environment provides tools for reading and converting formats between ASCII and Raster, as well as from Raster to Vector. Various GIS and Remote Sensing software packages can perform these conversions, such as ArcGIS by Esri, IDRISI by Clark University, and ENVI software by the defense industry group Elexis. This study employs the format conversion tools available within the ArcGIS suite by Esri, specifically the Conversion toolbox, including: Polygon to Raster, Raster to Polygon, Raster to ASCII, and ASCII to Raster.

## 4. RESULTS AND DISCUSSION

### 4.1. Analysis of agricultural land use dynamics

The primary objective of analyzing land use dynamics is to quantify historical land use conversion frequencies, employing the Markov analysis framework detailed in Section (3.1.2). These frequencies serve as the empirical basis for determining the Elasticity of Land Allocation for specific agricultural land use types (ELASu) within the CLUE-S model.

*Table 1.* Changes in agricultural land use area (2011-2020)

| | LUTs | Mã | Current LU (ha) | | LU change | (ELASu) |
|---|---|---|---|---|---|---|
| | | | Year 2011 | Year 2020 | 2011-2020 | |
| **1** | **Agricultural land** | **NNP** | **39.125** | **40.456** | **1.331** | |
| | *Including:* | | | | | |
| 1.1 | Paddy Rice Land | LUA | 4.085 | 4.575 | 490 | 0.107 |
| 1.2 | Cashew | CLN | 3.110 | 7.085 | 3.975 | 0.564 |
| 1.3 | Vegetables and Other Annual Crops | HNK | 421 | 621 | 200 | 0.322 |
| 1.4 | Other Agriculture | NKH | 116 | 138 | 22 | 0.159 |
| 1.5 | Forest Land | | 33.393 | 27.187 | -4.206 | |

*Source: Sub-NIAPP*

### 4.2. Logistic regression for land use–specific location suitability

Building upon the developed maps of land properties and the selection of agricultural land use types within the study area (crop types), a land suitability assessment was conducted using the logistic regression method. As detailed in Section (3.1.1) of the theoretical framework, the practical application involved sampling between LUTs and land properties using the Convert.exe program (a sub-program of CLUE-S). This sample dataset was subsequently analyzed using a program written in the R language. Here, the criteria for selecting the β parameters were based on a 95% confidence interval.

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

*Table 2.* LRM No. 1 – LUT 1: Paddy Rice Land

| Var | Properties | Var code [X] (file name.fil) | Estimate β | Std. Error | z value | Pr(>\|z\|) | S.C |
|---|---|---|---|---|---|---|---|
| | | Intercept | 5.018 | 1.052 | 4.771 | 1.84E-06 | *** |
| Slope | [0 - 90] | Slp (sc1gr0) | -0.058 | 0.015 | -3.770 | 1.63E-04 | *** |
| Soil type | (P) | So1 (sc1gr1) | -2.871 | 0.245 | -11.698 | < 2e-16 | *** |
| | (Pg) | So3 (sc1gr3) | 0.920 | 0.300 | 3.065 | 2.18E-03 | ** |
| | (Fk) | So5 (sc1gr5) | 11.190 | 1.202 | 9.313 | < 2e-16 | *** |
| Soil depth (cm) | >100 | De1 (sc1gr8) | 1.012 | 0.354 | 2.857 | 4.28E-03 | ** |
| | 70 – 100 | De2 (sc1gr9) | 0.647 | 0.294 | 2.198 | 2.79E-02 | * |
| Irrigation capacity | Near | Ir1 (sc1gr12) | 7.799 | 1.057 | 7.375 | 1.64E-13 | *** |
| | Mid | Ir2 (sc1gr13) | 7.266 | 1.057 | 6.876 | 6.15E-12 | *** |
| | Far | Ir3 (sc1gr14) | 6.929 | 1.066 | 6.502 | 7.93E-11 | *** |
| Elevation | [124 - 168] | El (sc1gr16) | -0.082 | 0.005 | -15.880 | < 2e-16 | *** |

$AUC_{Paddy\ Rice\ Land}$ = 0.9725, Signif. Codes (S.C): 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

*Table 3.* LRM No. 2 – LUT 2: Cashew

| Var | Properties | Var code [X] (file name.fil) | Estimate β | Std. Error | z value | Pr(>\|z\|) | S.C |
|---|---|---|---|---|---|---|---|
| | | Intercept | 0.688 | 0.164 | 4.191 | 2.77E-05 | *** |
| Slope | [0 - 90] | Slp (sc1gr0) | -0.018 | 0.004 | -4.896 | 9.76E-07 | *** |
| Soil type | (P) | So1 (sc1gr1) | -1.981 | 0.147 | -13.433 | < 2e-16 | *** |
| | (Pb) | So2 (sc1gr2) | -2.673 | 0.825 | -3.240 | 1.20E-03 | ** |
| | (Pg) | So3 (sc1gr3) | -4.876 | 0.717 | -6.796 | 1.08E-11 | *** |
| | (Pf) | So4 (sc1gr4) | -2.128 | 0.127 | -16.799 | < 2e-16 | *** |
| | (Fk) | So5 (sc1gr5) | 0.516 | 0.127 | 4.052 | 5.08E-05 | *** |
| | (Fs) | So6 (sc1gr6) | -0.301 | 0.117 | -2.563 | 1.04E-02 | * |
| Slope soil type | >100 | De1 (sc1gr8) | -0.765 | 0.121 | -6.312 | 2.76E-10 | *** |
| | 70 – 100 | De2 (sc1gr9) | -0.668 | 0.093 | -7.185 | 6.70E-13 | *** |
| Irrigation capacity | Near | Ir1 (sc1gr12) | 0.776 | 0.114 | 6.806 | 1.00E-11 | *** |
| | Mid | Ir2 (sc1gr13) | 1.720 | 0.063 | 27.114 | < 2e-16 | *** |
| | Far | Ir3 (sc1gr14) | 1.986 | 0.055 | 36.300 | < 2e-16 | *** |
| Elevation | [124 - 168] | El (sc1gr16) | -0.002 | 0.000 | -7.744 | 9.66E-15 | *** |

$AUC_{Cashew}$ = 0.7588, Signif. Codes (S.C): 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

*Table 4.* LRM No. 3 – LUT 3: Vegetables and Other Annual Crops

| Var | Properties | Var code [X] (file name.fil) | Estimate β | Std. Error | z value | Pr(>\|z\|) | S.C |
|---|---|---|---|---|---|---|---|
| | | Intercept | 4.812 | 0.883 | 5.449 | 5.06E-08 | *** |
| Irrigation Capacity | Mid | Ir2 (sc1gr13) | 1.285 | 0.407 | 3.155 | 1.60E-03 | ** |
| | Far | Ir3 (sc1gr14) | 2.383 | 0.390 | 6.118 | 9.48E-10 | *** |
| Elevation | [124 - 168] | El (sc1gr16) | -0.032 | 0.004 | -7.957 | 1.76E-15 | *** |

$AUC_{\text{Vegetables-Other Annual Crops}}$ = 0.8919, Signif. Codes (S.C): 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

| | | |
|---|---|---|
| $AUC_{\text{Paddy Rice Land}}$ = 0,9725 | $AUC_{\text{Cashew}}$ = 0,7588 | $AUC_{\text{Vegetables-Other Annual Crops}}$ = 0,8919 |

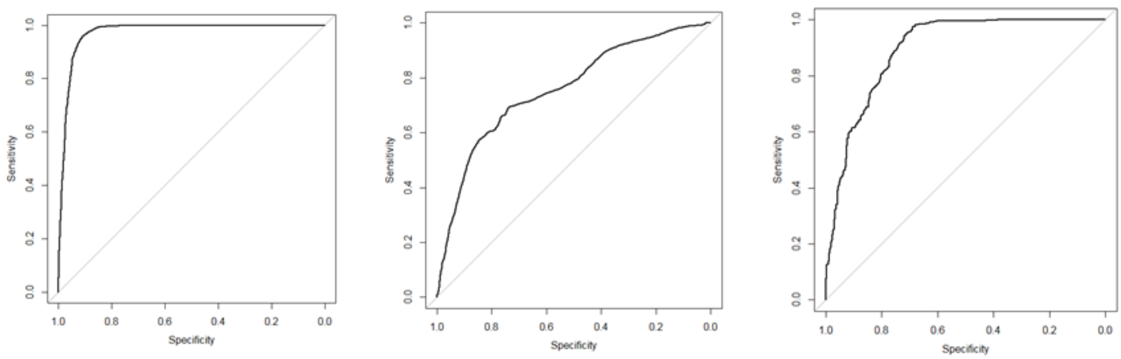


*Figure 12.* AUC of 3 Land Use Types

### 4.3. Classification of Land Suitability Maps for Crop-Specific Allocation

To generate land suitability maps for individual crop types, the suitability probability maps were classified into categories based on the FAO (1976) guidelines [17], [18]. Each pixel was assigned to a suitability class using the following thresholds:Highly Suitable (S1): $0.8 < S1 \leq 1.0$; Moderately Suitable (S2): $0.4 < S2 \leq 0.8$; Marginally Suitable (S3): $0.2 < S3 \leq 0.4$ and Not Suitable (N): $0 < N \leq 0.2$. This classification allows for the standardized comparison and spatial visualization of land use potential for different crop types, serving as a key input for land allocation decisions in the CLUE-S model.

### 4.4. Land Use Demand During the Planning Period

The land use demand table is the result of socio-economic analysis conducted by policy planners in Cát Tiên District. In this study, we adopt this demand projection as a scenario input to simulate land use changes over the planning period.

*Table 5.* Land use demand during the planning period (2021–2030)

| *Freq of LUC (ELAS)* | *0,107* | *0,564* | *0,322* | | |
|---|---|---|---|---|---|
| Years | Paddy rice (ha) | Cashew (ha) | Vegetables - other annual crops (ha) | Non-Agricultural and forestry land (ha) | Total land (ha) |
| 2021 | 4.758 | 6.892 | 1.629 | 29.392 | 42.671 |
| 2022 | 4.737 | 6.697 | 1.845 | 29.392 | 42.671 |
| 2023 | 4.748 | 6.802 | 1.729 | 29.392 | 42.671 |
| 2024 | 4.743 | 7.001 | 1.535 | 29.392 | 42.671 |
| 2025 | 4.743 | 7.025 | 1.511 | 29.392 | 42.671 |
| 2026 | 4.743 | 7.040 | 1.496 | 29.392 | 42.671 |
| 2027 | 4.743 | 6.354 | 1.350 | 29.392 | 42.671 |
| 2028 | 4.743 | 6.242 | 1.302 | 29.392 | 42.671 |
| 2029 | 4.743 | 6.130 | 1.271 | 29.392 | 42.671 |
| 2030 | 4.743 | 7.321 | 1.215 | 29.392 | 42.671 |

## 4.5. Configuration of the land use conversion matrix and model parameters

The land use conversion matrix is a critical component in the CLUE-S model, as it governs the allowable transitions between different land use types (LUTs). In the CLUE-S framework, a binary coding system is applied: a value of 1 indicates that a land use conversion is permitted, while 0 means it is not allowed.For the case study in Cát Tiên District, a conversion matrix was constructed to define feasible land use transitions, reflecting both policy constraints and land suitability considerations.

*Table 6.* Land use conversion matrix

| Land use type specific conversion settings | | Future Land-Use | | | |
|---|---|---|---|---|---|
| | | Paddy Rice Land | Cashew | Vegetables-Annual Crops | Non-agricultural And Forestry Land |
| Current Land-Use Status | Paddy Rice Land | 1 | 0 | 1 | 0 |
| | Cashew | 0 | 1 | 1 | 0 |
| | Vegetables-Annual Crops | 1 | 1 | 1 | 0 |
| | Non-agricultural and Forestry Land | 0 | 0 | 0 | 1 |

## 4.6. Results of land use allocation using the CLUE-S model

The land use allocation process in the CLUE-S model produced a projected land use map for the year 2030, based on spatial suitability, conversion rules, and land use demand inputs. This output allows for a comparative analysis between the **2020 baseline land use map** and the **simulated land use planning map for 2030**.


Figure 13: Current Land Use Map (2020)

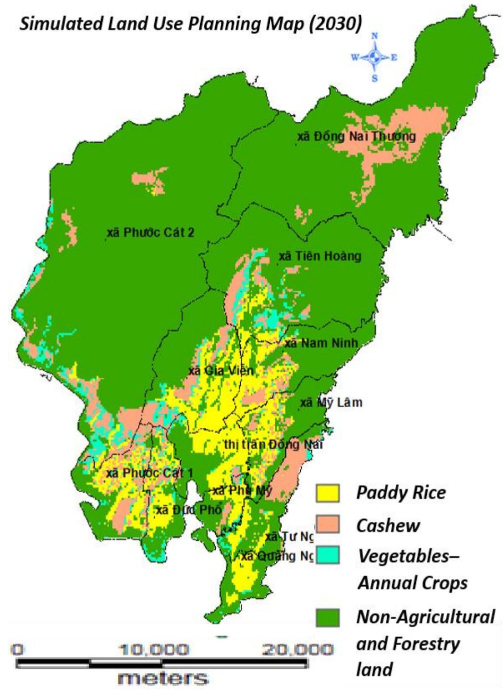


*Figure 14:* Simulated Land Use Planning Map (2030)

### 4.7. Model evaluation

The land use allocation results generated by the CLUE-S model demonstrate a high level of accuracy when compared to the projected land use demand. Specifically, the allocated land areas closely align with the planned land use areas developed by the Southern Institute of Agricultural Planning and Design, achieving an accuracy of over 85%.

The minor discrepancies observed between the CLUE-S simulated map and the official planning targets are primarily attributed to the iterative parameter configuration within the CLUE-S model. These settings influence the convergence behavior of land allocation and may result in slight variations in spatial distribution, even when overall area targets are met with high precision.

## 5. CONCLUSION

This research applied the land use planning theory developed by the German Agency for International Cooperation (GTZ), 1999 [19]. The approach emphasizes integrated land use planning with participatory involvement from local communities. The CLUE-S land use change model was constructed based on this planning framework.

Powerful spatial raster analysis within the GIS was utilized to process input data, visualize planning maps, and conduct dynamic and intuitive map analyses. The CLUE-S model has been successfully applied both globally and in Vietnam. The integration of CLUE-S with GIS enables the generation of land suitability maps in probabilistic units and simulations of future land use changes.

As land use planning increasingly demands high reliability and must rely on scientific evidence rather than solely the planner's experience, additional supportive tools are required alongside existing methodologies. The integrated CLUE-S and GIS model was developed to support policy-makers in spatial land allocation based on both land demand and environmental suitability. A pilot application was conducted in Cat Tien District, Lam Dong Province, yielding highly positive results. Regarding land suitability evaluation via statistical methods, the approach produced objective results. The accuracy of suitability maps generated through this method depends on the quality of input data, which requires expert knowledge of the region's biological, chemical, and climatic characteristics.

For land allocation: empirical validation demonstrated high spatial congruence. A key factor in ensuring accuracy in CLUE-S-based allocation is the precise identification of exclusion zones—specifically, the delineation between agricultural and non-agricultural areas. When this is done effectively, the model produces highly reliable spatial allocations.

**Advantages**: Rapid scalability to other localities.

**Limitations**: The model cannot allocate crops that have not previously been cultivated in the target region.

**Recommendations and Future Research Directions:**

Future development of this study should aim to deepen both land suitability evaluation and land use allocation methodologies.

1. **Land Suitability Evaluation**: The current research focused on evaluating natural land suitability. Future research should incorporate socio-economic factors into the assessment framework to better reflect the multifaceted nature of land use planning.
2. **Statistical Modeling in CLUE-S Land Suitability Assessment**: When applying statistical methods within the CLUE-S model, and specifically when constructing Logistic Regression Models (LRM), two critical issues should be considered:

**Issue 1: Model Accuracy and Variable Interactions -** A model's value lies in its ability to approximate reality with minimal deviation. Developing a well-fitting model requires expert knowledge to identify meaningful predictors. Importantly, interaction effects between predictors - often overlooked - can provide novel insights and should be systematically tested, particularly when sample sizes allow. A related concern is multicollinearity, which arises when predictor variables are correlated. Linear regression assumes predictor independence, but in reality, correlations often exist. When multicollinearity is present, it can: Alter parameter estimates depending on included variables; Reduce the precision of estimates; Increase residual sum of squares; Lead to inconsistent or misleading conclusions; Reduce the accuracy of predicted values. The most practical method to detect multicollinearity remains the use of statistical diagnostics within the R programming environment.

**Issue 2: Model Parsimony**- A good statistical model should explain maximum variability using a minimal number of predictors-an approach grounded in the principle of parsimony. A parsimonious model should: Be simple, using fewer predictors; Make minimal assumptions; Produce predictions closely matching observations. Overfitting (too many predictors) can cause noise and reduce predictive performance, while underfitting (too few predictors) omits important information and leads to poor generalization. Selecting the optimal number of predictors requires expert judgment and an understanding of regional ecological and landscape characteristics. Blindly including all collected data in regression analysis is not advisable, as overfitting can compromise both simplicity and accuracy.

**Acknowledgement**

The authors would like to sincerely thank the leaders of Sub-NIAPP, Professor Peter Verburg from the Institute for Environmental Studies (IVM), and expert Wilbert Van Rooij from Aidenvironment for their enthusiastic support in providing consultation on model application and supplying the necessary data for this research.

## REFERENCES


1. Verburg, P. H. - The CLUE model: Hands-on exercises. IVM (2010).
2. Veldkamp, A., & Fresco, L. O. - CLUE: A conceptual model to study the conversion of land use and its effects. Ecological Modelling **85** (2-3) (1996) 253–270.
3. Welmoed Soepboer. - CLUE-S An application for Sibuyan Island, The Philippines, Laboratory of Soil Science and Geology, Environmental Sciences, Wageningen University (2001).
4. Engelsman, W. - Simulating land use changes in an urbanizing area in Malaysia: An application of the CLUE-S model in the Selangor river basin. Department of Environmental Sciences, Laboratory of Soil Science and Geology, Wageningen University (2002).
5. Verburg, P. H., & Overmars, K. P. - Dynamic simulation of land-use change trajectories with the CLUE-S model. In Koomen, E., Stillwell, J., Bakema, A., & Scholten, H. J. (Eds.), Modelling land-use change: Progress and applications (2007) 321–335. Springer.
6. Zhan, J. Y., Deng, X., Jiang, Q. O., & Shi, N. - The application of system dynamics and CLUE-S model in land use change dynamic simulation: A case study in Taips County, Inner Mongolia of China. Proceedings of the 2007 Conference on Systems Science, Management Science and System Dynamics: Sustainable Development and Complex Systems **1** (2007) 2781-2789.
7. Batisani, N., & Yarnal, B. - Urban expansion in Centre County, Pennsylvania: Spatial dynamics and landscape transformations. Applied Geography **29** (2) (2009) 235–249.
8. Trisurat, Y., Alkemade, R., & Verburg, P. H. - Projecting land-use change and its consequences for biodiversity in Northern Thailand. Environmental Management **45** (3) (2010) 626–639.
9. Luo, G., Yin, C., Chen, X., Xu, W., & Lu, L. - Combining system dynamic model and CLUE-S model to improve land use scenario analyses at regional scale: A case study of Sangong watershed in Xinjiang, China, Ecological Complexity **7** (2010) 198–207.
10. Zhu, Z., Liu, L., Chen, Z., Zhang, J., & Verburg, P. H. - Land-use change simulation and assessment of driving factors in the loess hilly region: A case study as Pengyang County. Environmental Monitoring and Assessment **164** (1-4) (2010) 133–142.
11. van Dijk, M., Hilderink, H., van Rooij, W., Rutten, M., Ashton, R., & Kartikasari, K., & Lan V.C. - Land-use change, food security and climate change in Vietnam. LEI Wageningen UR (2013).
12. Pham Thi Loi, Dao Nguyen Khoi. - Simulation of land–use change in the Be river basin by using CLUE–s model, Journal of Hydro-Meteorology **2021**, 724, 82-93. https://doi.org/10.36335/VNJHM.2021(724).82-93
13. Pontius, R. G., Jr., & Schneider, L. C. - Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems & Environment **85** (1-3) (2001) 239–248.
14. Moulds, S., Buytaert, W., & Mijic, A. - Modelling land cover change in the Upper Ganga basin. Grantham Institute for Climate Change, Imperial College London (2013).

15. Schneider, L. C., & Pontius, R. G., Jr. - Modeling land-use change in the Ipswich watershed, Massachusetts, USA. Agriculture, Ecosystems & Environment **85** (1-3) (2001) 83–94.
16. Overmars, K. P., Verburg, P. H., & Veldkamp, A. - Comparison of a deductive and an inductive approach to specify land suitability in a spatially explicit land use model. Land Use Policy **24** (3) (2007) 584–599.
17. Peña, J., Bonet, A., Bellot, J., Sánchez, J. R., Eisenhuth, D., Hallett, S., & Aledo, A. - Driving forces of land-use change in a cultural landscape of Spain: A preliminary assessment of the human-mediated influences. In Koomen, E., Stillwell, J., Bakema, A., & Scholten, H. J. (Eds.), Modelling land-use change: Progress and applications (2007) 97–115. Springer.
18. FAO - A framework for land evaluation (1976).
19. Dinh, L. C. - Integration of ALES and GIS in land evaluation. Master's thesis, Ho Chi Minh City University of Technology (2005).
20. GTZ. Land Use Planning Methods, Strategies and Tools: Working group on Integrated land use planning. Germany: GTZ published, 1999.

## TÓM TẮT

### NÂNG CAO HIỆU QUẢ QUY HOẠCH SỬ DỤNG ĐẤT BẰNG TÍCH HỢP MÔ HÌNH CLUE-S VÀ HỆ THỐNG THÔNG TIN ĐỊA LÝ (GIS)

Nguyễn Đức Trí[1,2,3], Đinh Thị Kim Phượng[1,2,*]

[1]*Khoa Môi trường và Tài nguyên, Trường Đại học Bách khoa – Thành phố Hồ Chí Minh*

[2]*Đại học Quốc gia - Thành phố Hồ Chí Minh, Việt Nam*

[3]*Phân viện Quy hoạch và Thiết kế nông nghiệp miền Nam, Việt Nam*

*Email: *dtkphuong@hcmut.edu.vn*

Quy hoạch sử dụng đất bao gồm hai thành phần chính: đánh giá đất và phân bố không gian của các kế hoạch sử dụng đất. Việc đánh giá đất được thực hiện thông qua phương pháp giới hạn tối đa và đánh giá đa tiêu chí (MCE). Mô hình CLUE-S tăng cường quá trình này bằng cách áp dụng phân tích hồi quy logistic để đánh giá mức độ thích hợp của đất. Trong phân bố không gian, CLUE-S sử dụng phương pháp Cellular Automata (CA) như một công cụ chính, được tích hợp với các quy tắc ra quyết định và phân tích Markov. Trong nghiên cứu này, việc tích hợp mô hình CLUE-S với Hệ thống Thông tin Địa lý (GIS) đã được áp dụng tại huyện Cát Tiên, tỉnh Lâm Đồng, tập trung vào ba loại hình sử dụng đất: lúa, điều và rau màu. Ba mô hình hồi quy logistic đã được xây dựng dựa trên 17 lớp thuộc tính đặc trưng của đất, với chỉ số diện tích dưới đường cong ROC (AUC) được sử dụng để đo lường mức độ phù hợp của mô hình (Goodness of Fit - GoF). Các tham số β (với mức ý nghĩa < 5%) được sử dụng để lập bản đồ thích hợp đất. Dựa trên bảng nhu cầu và bản đồ thích hợp đất, mô hình CLUE-S đã phân bố sử dụng đất cho từng năm đến năm 2030, tạo ra các bản đồ mô phỏng sử dụng đất cho giai đoạn quy hoạch. Việc tích hợp mô hình CLUE-S và GIS là một công cụ hữu ích cho các nhà quy hoạch sử dụng đất và các nhà hoạch định chính sách, hỗ trợ phân tích biến động sử dụng đất và bổ sung cho các phương pháp hiện có nhằm nâng cao hiệu quả công tác quy hoạch sử dụng đất.

*Từ khóa:* Mô hình CLUE-S, phân tích hồi quy logistic, quy hoạch sử dụng đất, Phân tích Markov, hệ thống thông tin địa lý (GIS).